Extractors and condensers from univariate polynomials

Venkatesan Guruswami*
Department of Computer Science and Engineering
University of Washington
Seattle, WA 98195
venkat@cs.washington.edu

Christopher Umans†
Computer Science Department
California Institute of Technology
Pasadena, CA 91125
umans@cs.caltech.edu

Salil Vadhan‡
Division of Engineering and Applied Sciences
Harvard University
Cambridge, MA 02138
salil@eecs.harvard.edu

December 3, 2006

Abstract

We give new constructions of randomness extractors and lossless condensers that are optimal to within constant factors in both the seed length and the output length. For extractors, this matches the parameters of the current best known construction [LRVW03], with an improvement in case the error parameter is small (e.g. $1/\text{poly}(n)$). For lossless condensers, the previous best constructions achieved optimality to within a constant factor in one parameter only at the expense of a polynomial loss in the other.

Our constructions are based on the Parvaresh-Vardy codes [PV05], and our proof technique is inspired by the list-decoding algorithm for those codes. The main object we construct is a condenser that loses *only* the entropy of its seed plus one bit, while condensing to entropy rate $1-\alpha$ for any desired constant $\alpha > 0$. This construction is simple to describe, and has a short and completely self-contained analysis. Our other results only require, in addition, standard uses of randomness-efficient hash functions (to obtain a lossless condenser) or expander walks (to obtain an extractor).

Our techniques also show for the first time that a natural analogue of the Shaltiel–Umans extractor [SU05] based on univariate polynomials (i.e., Reed-Solomon codes) yields a condenser that retains a $1-\alpha$ fraction of the source min-entropy, for any desired constant $\alpha > 0$, while condensing to constant entropy rate and using a seed length that is optimal to within constant factors.

*Supported by NSF CCF-0343672, a Sloan Research Fellowship, and a David and Lucile Packard Foundation Fellowship.
†Supported by NSF CCF-0346991, BSF 2004329, a Sloan Research Fellowship, and an Okawa Foundation research grant.
‡Supported by NSF CCF-0133096, ONR N00014-04-1-0478, and US-Israel BSF 2002246.

1 Introduction

In this paper, we construct randomness extractors and condensers with the best parameters to date. Perhaps more importantly, we do this by introducing a new algebraic construction based on the ingenious variant of Reed-Solomon codes discovered by Parvaresh and Vardy [PV05]. Our proof technique is inspired by the list-decoding algorithm for the Parvaresh-Vardy codes, which builds on the list-decoding results of [Sud97, GS99]. The resulting extractors and condensers are simple to describe and have short, self-contained analyses. In the remainder of the introduction, we describe our results more precisely, and place them in context within the large body of literature on extractors and related objects.

A long line of research beginning in the late 1980s has been devoted to the goal of constructing explicit *randomness extractors*. (See the survey of Shaltiel [Sha02].) Extractors are efficient functions that take an n-bit string sampled from a "weak" random source together with a short truly random seed, and output a nearly uniform distribution. Extractors have turned out to be a powerful tool in a number of application areas. These include algorithms [WZ99], hardness of approximation [Zuc96a, Uma99, MU02, Zuc06], distributed protocols [Zuc97, RZ01], coding theory [TSZ04, Gur04], and the complexity of randomized algorithms [Sip88, NZ96, GZ97].

The randomness in the source is measured by *minentropy*: a random variable $\mathbf{X}$ has minentropy at least k iff $\Pr[\mathbf{X} = x] \leq 2^{-k}$ for all x. A random variable $\mathbf{Z}$ is *ε-close* to a distribution D if for all events A, $\Pr[\mathbf{Z} \in A]$ differs from the probability of A under the distribution D by at most ε. An extractor is defined as follows:

Definition 1.1 ([NZ96]). *A function $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k, ε)* extractor *if for every $\mathbf{X}$ with minentropy at least k, $E(\mathbf{X}, \mathbf{Y})$ is ε-close to uniform, when $\mathbf{Y}$ is uniformly distributed on $\{0,1\}^d$. An extractor is* explicit *if it is computable in polynomial time.*

The competing goals when constructing extractors are to obtain a short seed, and a long output length. Nonconstructively, it is possible to simultaneously have a seed length $d = \log n + 2\log(1/\varepsilon) + O(1)$ and an output length of $m = k + d - 2\log(1/\varepsilon) - O(1)$. It remains open to match these parameters with an explicit construction.

A major theme in extractor constructions since the breakthrough result of Trevisan [Tre01], has been the use of error-correcting codes. Trevisan's extractor construction, which is based on the Nisan-Wigderson pseudorandom generator [NW94], encodes the source with an error-correcting code with good distance, and uses the seed to select (via certain combinatorial designs) a subset of m bits of the codeword to output.

A more algebraic approach, exploiting the specific structure of polynomial error-correcting codes was pioneered by Ta-Shma, Zuckerman and Safra [TZS06]. There the source is encoded with a multivariate polynomial code (Reed-Muller code), the seed is used to select a starting point, and the extractor outputs m successive symbols along a line[1]. Better parameters were achieve with a variant introduced by Shaltiel and Umans [SU05], which exploits the fact that Reed-Muller codes are *cyclic*. There the m output symbols are simply m successive coordinates of the codeword, when written in the cyclic ordering. A common feature of these algebraic constructions is that their analysis relies crucially on the *local-decodability* properties of the underlying error-correcting code. This paper diverges from the previous works on exactly this point, as our constructions use only univariate polynomial codes, which are not locally decodable.

A second major theme dating to [RSW06] and [RR99][2] is the use of a relaxation of extractors, called

[1]In this discussion we are ignoring the distinction between outputting m symbols from a large alphabet and outputting m bits.

[2]Actually, since the formal definition we give does not explicitly require that the min-entropy rate increase, such objects were already considered as far back as the original papers of [Zuc96b, NZ96]. However, we will be interested in condensers that do actually increase the min-entropy rate.

condensers, as an intermediate goal:

Definition 1.2. *A function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is an $k \to_\varepsilon k'$* condenser *if for every $\mathbf{X}$ with minentropy at least k, $C(\mathbf{X}, \mathbf{Y})$ is ε*-close *to a distribution with minentropy k', when $\mathbf{Y}$ is uniformly distributed on $\{0,1\}^d$. A condenser is* explicit *if it is computable in polynomial time. A condenser is called* lossless *if $k' = k + d$.*

Observe that a $k \to_\varepsilon k'$ condenser with output length $m = k'$ is an extractor, because the unique distribution on $\{0,1\}^m$ with minentropy m is the uniform distribution. Condensers are a natural stepping-stone to constructing extractors, as they can be used to increase the *entropy rate* (the ratio of the minentropy in a random variable to the length of the strings over which it is distributed), and it is often easier to construct extractors when the entropy rate is high. Condensers have also been used extensively in less obvious ways to build extractors, often as part of complex recursive constructions (e.g., [ISW00, RSW06, LRVW03]). Nonconstructively, one can hope for *lossless* condensers with seed length $d = \log n + \log(1/\varepsilon) + O(1)$, and output length $m = k + d + \log(1/\varepsilon) + O(1)$.

Our central result is a completely elementary construction of a condenser that retains all but the seed min-entropy (plus one bit), and condenses to *any* constant entropy rate using a seed length that is optimal up to constant factors. This is the most basic object from which we derive most of the other results:

Theorem 1.1 (main). *For all $\alpha > 0$, all positive integers $n \geqslant \ell$ and all $\varepsilon \geqslant 2^{-\ell}$, there is an explicit construction of a*

$$(k = \ell t + \log(1/\varepsilon)) \to_{3\varepsilon} (k-1)$$

condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\ell d}$ with $t = \log\lceil (2n^2/\varepsilon)^{1/\alpha} \rceil$ and $d = \lfloor (1+\alpha)t \rfloor$.

For intuition about the parameters, consider taking α to be a small constant. Then the seed length is $d = O(\log(n/\varepsilon))$, which is optimal up to a constant factor. The condenser takes takes any distribution of min-entropy $k \approx \ell t$ and outputs a string of length $\ell d \approx (1+\alpha)k$ that still has min-entropy at least $k - 1$. Thus the min-entropy *rate* of the output is at least $(k-1)/(\ell d) \approx 1/(1+\alpha)$, which is arbitrarily close to 1.

In recent years, condensers have been studied in their own right. Lossless condensers are of particular interest, as they are equivalent to unbalanced bipartite *expander graphs* with extremely good expansion (of greater than half the left degree of the graph).[3] This turns out to be useful in a number of applications (see [CRVW02, §1.3] for a survey). Constructions of lossless condensers appear in [RR99, TUZ01, CRVW02, TU06].

For lossless condensers, the competing goals are short seed length, and *short* output length (thus achieving the greatest "condensing" of the source minentropy). Constructions are known that achieve essentially optimal parameters for very large k [CRVW02], and very small k [RR99], but for general k, the best known constructions can achieve optimality to within a constant factor in one parameter only at the expense of a polynomial loss in the other. Specifically, the best known constructions (stated here for constant ε) achieve seed length $t = O(\log^2 n)$ and output length $m = O(k)$ [TUZ01], or seed length $t = O(\log n)$ and output length $m = k^{1+\alpha}$ for any constant $\alpha > 0$ [TUZ01]. Recently Ta-Shma and Umans [TU06] showed that if optimal *derandomized curve samplers* can be constructed, then a construction of lossless condensers based on [SU05] would achieve seed length $t = O(\log n)$ and output length $m = k \cdot \operatorname{poly}\log(n)$; they obtain near-optimal derandomized curve samplers that produce lossless condensers with somewhat worse parameters.

[3]Technically, the usual notion of expander corresponds to condensers that are simultaneously lossless for all min-entropies k up to some threshold (in contrast to Definition 1.2, which refers to a single value of k). Our constructions actually achieve this stronger property, as shown in the more detailed statements of the theorems in the body of the paper.

Using Theorem 1.1, we obtain a new construction of lossless condensers that are optimal to within constant factors in both the seed length and the output length. This uses an idea from [RR99]: because the condenser of Theorem 1.1 is only missing a small amount of minentropy, it can be made lossless by appending a hash from an "almost-2-universal" hash family; we pay only with a constant factor increase in the seed length. We obtain:

Theorem 1.2 (lossless condenser)**.** *For every constant $\alpha > 0$, For all positive integers $n \geqslant k$ and all $\varepsilon > 0$, there is an explicit construction of a*

$$k \to_\varepsilon k + d$$

lossless condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = O(\log n + \log(1/\varepsilon))$ and $m = (1+\alpha)(k+d)$.

We now return to extractors. There is a great diversity of extractor constructions; see Shaltiel's survey [Sha02] for a nearly-up-to-date summary. The current champion is the construction of Lu, Reingold, Vadhan, and Wigderson [LRVW03] which achieves optimality to within a constant factor in the seed length and output length simultaneously, for any minentropy k. (As with lossless condensers, for small k, better constructions are known; e.g., [GW97, SZ99, TUZ01]). Again using the condenser of Theorem 1.1, we can match this best known construction with a simple, direct, and self-contained construction and analysis. We simply need to "finish" the condenser of Theorem 1.1 with an extractor that extracts any desired constant fraction of the minentropy, with a seed length that is optimal up to constant factors. Since this extractor can start from a constant entropy rate arbitrarily close to 1, we can even use a standard extractor based on expander walks. When ε is sub-constant, we use Zuckerman's extractor [Zuc97] to obtain the proper dependence on ε. Altogether we obtain:

Theorem 1.3 (extractor)**.** *For all constants $\alpha > 0$: for all positive integers n, k and all $\varepsilon > \exp(-n/2^{O(\log^* n)})$, there is an explicit construction of a (k, ε) extractor $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = O(\log n + \log \frac{1}{\varepsilon})$ and $m \geqslant (1-\alpha)k$.*

In fact this result slightly improves upon [LRVW03], for general error $\varepsilon = \varepsilon(n)$. They can handle error as small as $n^{-1/\log^{(c)} n}$ for any constant c, but for general ε, they must pay with either a larger seed length of $t = O((\log^* n)^2 \log n + \log(\frac{1}{\varepsilon}))$, or a smaller output length of $m = \Omega(k/\log^{(c)} n)$ for any constant c.

1.1 Our technique

In this section we give a high-level description of our construction and proof technique. Our condensers are based on Parvaresh-Vardy codes [PV05], which in turn are based on Reed-Solomon codes. A Reed-Solomon codeword is a univariate degree n polynomial $f \in \mathbb{F}_q[Y]$, evaluated at all points in the field. A Parvaresh-Vardy codeword is a bundle of several related degree $n-1$ polynomials $f_0, f_1, f_2, \ldots, f_{m-1}$, each evaluated at all points in the field. The evaluations of the various f_i at a given field element are packaged into a symbol from the larger alphabet $\mathbb{F}_{q^m}$. The purpose of this extra redundancy is to enable a better list-decoding algorithm than is possible for Reed-Solomon codes.

The main idea in [PV05] is to view degree $n-1$ polynomials as elements of the extension field $\mathbb{F} = \mathbb{F}_q[Y]/E(Y)$, where E is some irreducible polynomial of degree n. The f_i (now viewed as elements of $\mathbb{F}$) are chosen so that $f_i = f_0^{h_i}$ for $i \geq 1$, and positive integers h_i. In order to list-decode, one produces a nonzero univariate polynomial Q^* over $\mathbb{F}$ from the received word, with the property that f_0 is a root of Q^* whenever the codeword has sufficient agreement with the received word. We use the same technique in the analysis of our condenser, and below we describe how the interpolating polynomial is set up and how the relationship between the f_i's helps in the context of our analysis.

Our condenser construction works as follows. We view the source string x as describing a degree $n-1$ polynomial $f(Y) \in \mathbb{F}_q[Y]$. We then define $f_i \stackrel{\text{def}}{=} f^{h^i} \bmod E$ for some parameter h, and irreducible E of degree n. Given a seed $y \in \mathbb{F}_q$, our output is $f_0(y), f_1(y), \ldots, f_{m-1}(y)$.

Since [Tre01], a common technique in analyzing extractors has been to show that for every subset $D \subseteq \{0,1\}^m$, there are very few, say $\ll 2^k$, source strings x that are "bad" with respect to D; i.e., much fewer than 2^k strings x satisfy

$$\left| \Pr_y[E(x,y) \in D] - \Pr_z[z \in D] \right| > \varepsilon.$$

From this, it follows that a source with min-entropy k is unlikely to output a string that is bad with respect to any given D. Thus the output of E on such a source must hit all D's with probability close to the density of D, and so E is an extractor for minentropy k. We use the same general outline to show that our construction is a condenser. We only wish to show that the output is close to having minentropy k', rather than close to being uniform, and this is equivalent to showing that the output hits sets S of size about $2^{k'}$ with less than ε probability (see Section 2.1 for a precise statement of this fact). We do this by arguing that there are very few source strings x that are "bad" with respect to S; i.e., very few x satisfy $\Pr_y[C(x,y) \in S] > \varepsilon$.

Let's consider what $\Pr_y[C(x,y) \in S] > \varepsilon$ means for our construction. First of all, x is interpreted as a degree $n-1$ polynomial f_0. Then, f_0 being "bad" means that for more than εq of the seeds y, we have

$$(f_0(y), f_1(y), \ldots, f_{m-1}(y)) \in S.$$

The first step in our analysis is to produce a non-zero polynomial $Q : \mathbb{F}_q^m \to \mathbb{F}_q$ that vanishes on S. We arrange to have $n \deg Q < \varepsilon q$, so that the univariate polynomial $Q(f_0(Y), f_1(Y), \ldots, f_{m-1}(Y))$ is *identically zero* for bad f_0. Viewing the f_i as elements of the extension field $\mathbb{F} = \mathbb{F}_q[Y]/E(Y)$, and Q as a multivariate polynomial over $\mathbb{F}$, we have that $(f_0, f_1, \ldots, f_{m-1})$ is a *root* of Q. Just as in the list-decoding algorithm of [PV05], we define the polynomial $Q^*(Z) \stackrel{\text{def}}{=} Q(Z, Z^h, Z^{h^2}, \ldots, Z^{h^{m-1}})$, and observe that every bad f_0 is a root of this *univariate* polynomial. Thus the degree of Q^* is a bound on the number of such f_0, and it turns out that this bound is nearly optimal: the number of bad f_0 is shown to be at most the size of S.

To summarize, the analysis has two main steps: first, we encode S into a low-degree multivariate polynomial Q, and argue that for every bad polynomial $f_0(Y)$, $Q(f_0(Y), \ldots, f_{m-1}(Y))$ is in fact identically zero. Then, we produce a univariate polynomial Q^* from Q that has all of the bad f_0 as roots (when everything is viewed over the extension field $\mathbb{F}$). The degree of Q^* is an upper bound on the number of bad strings.

1.2 Additional results

In Section 6 we discuss some variations on the basic construction.

Using the "multiple roots" idea from Guruswami-Sudan [GS99], we optimize the seed length of our condenser, making it $(1+\gamma)$ times the optimal seed length, while still retaining almost all the entropy and outputting a source with a constant entropy rate of $\Omega(\gamma)$ (Theorem 6.2). For constant error ε, one can then extract almost all the entropy using the extractor from [Zuc06] which uses an additional seed of at most $\log k + O(1)$ bits. The total seed length is thus $(1+\gamma)\log n + \log k + O(1)$, which approaches the optimal $\log n + O(1)$ bound for $k = n^{o(1)}$. This result appears as Theorem 6.5. A different setting of the condenser parameters (Corollary 6.3) allows us to obtain an *exactly* optimal seed length, while retaining a constant fraction (arbitrarily close to 1) of the entropy, at the expense of an output entropy rate of $\Omega(1/\log(n/\varepsilon))$, which is nonconstant, but still quite good.

With a small change to the original proof, we can say something about the variant of the main condenser in which the seed is included in the output. One can hope to capture the entire seed entropy (which we do in Theorem 1.2, but that involves the extra step of appending a hash); here we are able to capture all but $O(\log(1/\varepsilon))$ bits of the seed entropy directly.

Finally, using one of the main ideas from the Guruswami-Rudra codes [GR06], we argue that a variant of our main construction is the natural precursor of [SU05], in which that basic construction is applied Reed-Solomon codes. It has been an intriguing question for some time to determine what (if any) pseudorandom object(s) can be obtained from this very natural construction. This question is studied in [KU06], where they show that the Reed-Solomon construction "fools" certain kinds of low-degree tests. Our results in this paper, which show that this construction is a very good condenser, seem to provide the correct (or nearly-correct) answer, as we also describe an example that shows that the entropy rate and the constant factor entropy loss for this construction cannot be improved substantively.

1.3 Organization

In Section 2, we present some basic lemmas about min-entropy and condensers that will provide the framework for analyzing our constructions. In Section 3, we describe and analyze our basic condenser, proving Theorem 1.1. In Section 4, we convert our basic condenser into a lossless condenser, establishing Theorem 1.2. In Section 5, we use our basic condenser to construct extractors, obtaining Theorem 1.3. In Section 6, we present some variants of our basic condenser, in particular giving a version that optimizes the constant in the seed length, a version that retains some of the seed min-entropy, and a version that is a univariate analogue of [SU05]. Section 7 contains some conclusions and open problems.

2 Preliminaries

Throughout this paper, we use boldface capital letters for random variables (e.g., "$\mathbf{X}$"), capital letters for indeterminates, and lower case letters for elements of a set. Also throughout the paper, $\mathbf{U_t}$ is the random variable uniformly distributed on $\{0,1\}^t$. The *support* of a random variable $\mathbf{X}$ is $\mathrm{supp}(\mathbf{X}) \stackrel{\text{def}}{=} \{x : \Pr[\mathbf{X} = x] > 0\}$. The *statistical distance* between random variables (or distributions) $\mathbf{X}$ and $\mathbf{Y}$ is $\max_T |\Pr[\mathbf{X} \in T] - \Pr[\mathbf{Y} \in T]|$. We say $\mathbf{X}$ and $\mathbf{Y}$ are ε*-close* if their statistical distance is at most ε. All logs are base 2.

We record some standard facts about minentropy:

Proposition 2.1. *For $K \in \mathbb{N}$, a distribution D has minentropy at least $\log K$ iff D is a convex combination of flat distributions on sets of size exactly K.*

Proposition 2.2. *For any $k > 0$, the distance from a distribution D to a closest distribution with minentropy k is exactly $\sum_{a:D(a)\geq 2^{-k}}(D(a) - 2^{-k})$.*

Proposition 2.3. *A distribution D with minentropy $\log(K - c)$ is c/K-close to some distribution with minentropy $\log K$.*

Proof. By Proposition 2.2, the distance from D to the closest distribution with minentropy $\log K$ is

$$\sum_{a:D(a)\geq 1/K} (D(a) - 1/K) \leqslant 1 - (K - c) \cdot 1/K = c/K.$$

□

2.1 Analysis of condensers

The next lemma gives a useful sufficient condition for a distribution to be close to having large minentropy:

Lemma 2.4. *Let* $\mathbf{Z}$ *be a random variable. If for all sets* S *of size* K, $\Pr[\mathbf{Z} \in S] \leq \varepsilon$ *then* $\mathbf{Z}$ *is* ε*-close to having minentropy at least* $\log(K/\varepsilon)$*.*

Proof. Let S be a set of the K heaviest elements x (under the distribution of $\mathbf{Z}$). Let $2^{-\ell}$ be the average probability mass of the elements in S. Then $\varepsilon \geq \Pr[\mathbf{Z} \in S] = 2^{-\ell}K$, so $\ell \geq \log(K/\varepsilon)$. But every element outside S has weight at most $2^{-\ell}$, and with all but probability ε, $\mathbf{Z}$ hits elements outside S. □

This lemma establishes the framework within which we will prove our constructions are condensers:

Lemma 2.5. *Let* $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be a function. For each subset* S*, define*

$$\mathrm{BAD}(S,\varepsilon) = \left\{ x : \Pr_y[C(x,y) \in S] > \varepsilon \right\}.$$

For $K \in \mathbb{N}$*, define* $B(K,\varepsilon) = \max_{S:|S|=K} |\mathrm{BAD}(S,\varepsilon)|$*. Then the function* C *is a*

$$\log(B(K,\varepsilon)/\varepsilon) \to_{2\varepsilon} \log(K/\varepsilon)$$

condenser.

Proof. We have a random variable $\mathbf{X}$ with minentropy $\log(B(K,\varepsilon)/\varepsilon)$. For a fixed S of size K, the probability that $\mathbf{X}$ is in $\mathrm{BAD}(S,\varepsilon)$ is at most ε; if that does not happen, then the probability $C(\mathbf{X},\mathbf{U_t})$ lands in S is at most ε. Altogether the probability $C(\mathbf{X},\mathbf{U_t})$ falls in S is at most 2ε. Now apply Lemma 2.4. □

3 The main construction

Fix the field $\mathbb{F}_q$ and let $E(Y)$ be an irreducible polynomial of degree n over $\mathbb{F}_q$. View elements of $\mathbb{F}_q^n$ as describing univariate polynomials over $\mathbb{F}_q$ with degree at most $n-1$. Fix an integer parameter h.

We describe a function $C : \mathbb{F}_q^n \times \mathbb{F}_q \to \mathbb{F}_q^m$ that is the basis of all of our constructions:

$$C(f,y) \stackrel{\text{def}}{=} [f(y), (f^h \bmod E)(y), (f^{h^2} \bmod E)(y), \cdots, (f^{h^{m-1}} \bmod E)(y)].$$

For ease of notation, we will refer to $(f^{h^i} \bmod E)$ as "f_i."

Lemma 3.1. *Defining* $\mathrm{BAD}(S,\varepsilon)$ *and* $B(K,\varepsilon)$ *with respect to* C *as in Lemma 2.5, we have*

$$B(K = h^m - 1, \varepsilon) \leq K,$$

provided $q \geq (n-1)(h-1)m/\varepsilon$*.*

Proof. Fix a set $S \subseteq \mathbb{F}_q^m$ of size at most K. We want to show that $|\mathrm{BAD}(S,\varepsilon)| \leqslant K$.

First, we observe that there exists a *nonzero* m-variate polynomial $Q \in \mathbb{F}_q[Z_1, Z_2, \ldots, Z_m]$ that vanishes on S, and whose degree in each variable is at most $h-1$. (For each $z \in S$, the condition $Q(z) = 0$ is a homogenous linear constraint on the h^m coefficients of Q. Since $|S| \leqslant K < h^m$, we have fewer constraints than unknowns, so this linear system has a nonzero solution.)

Consider any polynomial $f(Y) \in \mathrm{BAD}(S, \varepsilon)$. By the definition of $\mathrm{BAD}(S, \varepsilon)$, it holds that

$$\Pr_y[Q(f_0(y), f_1(y), \ldots, f_{m-1}(y)) = 0] > \varepsilon.$$

Therefore, the univariate polynomial $R_f(Y) \stackrel{\mathrm{def}}{=} Q(f_0(Y), \ldots, f_{m-1}(Y))$ has more than εq zeroes, and degree at most $(n-1)(h-1)m$. Since $(n-1)(h-1)m \leq \varepsilon q$, $R_f(Y)$ must be identically zero, and so

$$Q(f_0(Y), \ldots, f_{m-1}(Y)) = 0$$

as a formal polynomial. Now recall that $f_i(Y) \equiv f(Y)^{h^i} \pmod{E(Y)}$. Thus,

$$Q(f(Y), f(Y)^h, \ldots, f(Y)^{h^{m-1}}) \equiv Q(f_0(Y), \ldots, f_{m-1}(Y)) \equiv 0 \pmod{E(Y)}.$$

So if we interpret $f(Y)$ as an element of the extension field $\mathbb{F} = \mathbb{F}_q[Y]/E(Y)$, then $f(Y)$ is a root of the univariate polynomial

$$Q^*(Z) \stackrel{\mathrm{def}}{=} Q(Z, Z^h, Z^{h^2}, \ldots, Z^{h^{m-1}})$$

over the field $\mathbb{F}$. Since this holds for every $f(Y) \in \mathrm{BAD}(S, \varepsilon)$, we deduce that Q^* has at least $|\mathrm{BAD}(S, \varepsilon)|$ roots in $\mathbb{F}$.

On the other hand, Q^* is a non-zero polynomial, because the individual degrees of Q are all less than h (so distinct monomials in Q map to distinct monomials in Q^*). Thus, the number of roots of Q^* is bounded by its degree, which is at most

$$(h-1)(1 + h + h^2 + \cdots + h^{m-1}) = h^m - 1 = K.$$

We conclude that $|\mathrm{BAD}(S, \varepsilon)| \leqslant K$, as desired. $\square$

Remark 1. *The above proof works even if the distribution on the seed y in the definition of* $\mathrm{BAD}(S, \varepsilon)$ *is not uniform on $\mathbb{F}_q$, but comes from any distribution on $\mathbb{F}_q$ of min-entropy at least* $\log((n-1)(h-1)m/\varepsilon)$. *This means that the construction yields a condenser that works even if the seed comes from a weak random source.*

We can now prove our main theorem, Theorem 1.1. Here we state it in a stronger form, with the most significant change being that it asserts that a single construction works for many different values of the source min-entropy k (as opposed to the construction being tailored to a particular value of k). This will allow us, in the next section, to construct condensers that are lossless for all min-entropies up to a given threshold. The significance of this property is that the lossless condensers then correspond to the standard notion of expander graphs, where expansion holds for all sets up to a given size. Intuitively, the reason that our condenser works for many different source min-entropies is that every prefix of the condenser is a condenser of the same form, but corresponding to a smaller value of $B(K, \varepsilon) \leqslant K$ in Lemma 3.1 (and $\log(B(K, \varepsilon)/\varepsilon)$ corresponds to the source min-entropy, when we apply Lemma 2.5).

Theorem 3.2 (Thm. 1.1, strengthened)**.** *For all $\alpha > 0$, all positive integers $n \geqslant n'$, all $\varepsilon > 0$, and all integers $2^t \geqslant (2nn'/\varepsilon)^{1/\alpha}$, there is an explicit function $C : \{0,1\}^{nd} \times \{0,1\}^d \to \{0,1\}^{n'd}$ with $d = \lfloor (1+\alpha)t \rfloor$ such that for all positive integers $\ell \in [\log(1/\varepsilon)/t, n']$, C is a*

$$(k = \ell t + \log(1/\varepsilon)) \to_{3\varepsilon} (k-1)$$

condenser.

To see how the original form of Theorem 1.1 follows, take $t = \lceil \log(2nn'/\varepsilon)^{1/\alpha} \rceil$, change the input length from nd to n (a condenser for a given input length yields a condenser for shorter input lengths by just padding the input with zeroes), and fix $\ell = n'$

Proof. We describe how to set parameters in the condenser of Lemma 3.1 and then apply Lemma 2.5. Let $h = 2^t \geqslant (2nn'/\varepsilon)^{1/\alpha}$, $d = \lfloor (1+\alpha)t \rfloor$ and $q = 2^d$. Note that $q \geqslant h^{1+\alpha}/2$.

Let $C : \mathbb{F}_q^n \times \mathbb{F}_q \to \mathbb{F}_q^{n'}$ be the condenser of Lemma 3.1 with parameter h and $m = n'$ output symbols. Note the input length, output length, seed length, and the value of t match the parameters claimed in the theorem. Moreover, a representation of $\mathbb{F}_q$ for $q = 2^d$ (i.e. an irreducible polynomial of degree d over $\mathbb{F}_2$) as well as an irreducible polynomial $E(Y)$ of degree n over $\mathbb{F}_q$ can be found in time $\mathrm{poly}(d, n)$ [Sho90], and thus the construction is explicit.

Now, given any $\ell \leqslant n'$, let C' denote the first ℓ symbols of the output of C; this is also a condenser of the type analyzed in Lemma 3.1. We will show that C' is a

$$(k = \ell t + \log(1/\varepsilon)) \to_{3\varepsilon} k - 1,$$

which implies that C is also a condenser with these parameters.

For consistency with Lemma 3.1, we write $m = \ell$ for the rest of the proof. Note that

$$q \geqslant h \cdot (h^\alpha/2) \geqslant h \cdot (nn'/\varepsilon) \geqslant hnm/\varepsilon.$$

Thus, by Lemma 3.1 and Lemma 2.5, C is a

$$\log((h^m - 1)/\varepsilon)) \to_{2\varepsilon} \log((h^m - 1)/\varepsilon) - 1$$

condenser. All that remains is numerical manipulation to express this in the same way as it is stated in the theorem. First, note that

$$\log((h^m - 1)/\varepsilon) < \log(h^m/\varepsilon) = mt + \log(1/\varepsilon) \; .$$

Also, by Proposition 2.3, a distribution with $\log((h^m - 1)/\varepsilon) - 1$ minentropy is $1/h^m$-close to having minentropy

$$\log(h^m/\varepsilon) - 1 = mt + \log(1/\varepsilon) - 1.$$

Since $1/h^m = 1/2^{mt} \leqslant \varepsilon$, C' is a $mt + \log(1/\varepsilon) \to_{3\varepsilon} mt + \log(1/\varepsilon) - 1$ condenser as claimed. ☐

Remark 2. *In this proof we work in a field $\mathbb{F}_q$ of characteristic 2, which has the advantage of yielding a polynomial-time construction even when we need to take q to be superpolynomially large (which occurs when $\varepsilon(n) = n^{-\omega(1)}$). When $\varepsilon \geqslant 1/poly(n)$, then we could take a prime $q \geqslant 2^d$ instead, with some minor adjustments to the construction (e.g. only using 2^d elements of $\mathbb{F}_q$ for the seed, as per Remark 1) and the parameters claimed in the theorem.*

4 Lossless condensers that are optimal up to constant factors

We begin with the general method to recover "missing" minentropy, due to [RR99]. Given a $k \to_\varepsilon k'$ condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$, we say it has entropy loss $\ell = k + d - k'$. We can make the condenser lossless by appending a random hash into $\ell + \log(1/\varepsilon)$ bits. When d is small, the extra randomness can also be small, provided we use a randomness-efficient family of hash functions. Specifically, we can use a family of "almost 2-universal" hash functions:

Theorem 4.1 ([AGHP92, SZ99]). *For every n', m', there exists an explicit family H of hash functions from n' to m' bits, of cardinality $O((n'm'2^{m'})^2)$, that satisfies the following property:*

$$\forall w_1 \neq w_2 \quad \Pr_{h \in H}[h(w_1) = h(w_2)] \leq 2 \cdot 2^{-m'}. \tag{1}$$

A random $h \in H$ can be sampled using $\log |H|$ bits, and given these bits, h can be computed in poly(n', m') time.

Note that a truly 2-universal hash function would satisfy (1) with the right-hand-side replaced by $2^{-m'}$ – but the price would be that $|H| \geq 2^{n'}$, which is far too large to be useful for us. Now we show that appending a random hash makes a condenser lossless.

Lemma 4.2. *Let $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a $k \to_\varepsilon k'$ condenser. Let H be a family of hash functions from $n' = n + d$ bits to $m' \geqslant k + d - k' + \log(1/\varepsilon) + 2$ bits satisfying (1). Then the function $C' : \{0,1\}^n \times \{0,1\}^{d'=d+\log|H|} \to \{0,1\}^{m+\log|H|+m'}$ defined by:*

$$C'(x; y, h \in H) \stackrel{\text{def}}{=} (C(x,y), h, h(x,y))$$

is a $k \to_{3\varepsilon} k + d'$ lossless condenser.

Proof. Let $\mathbf{X}$ be a random variable distributed uniformly on an arbitrary set of size 2^k. We prove that C' is the stated condenser when its source is $\mathbf{X}$, which by Proposition 2.1 suffices. We denote by $\mathbf{H}$, the random variable that is uniformly distributed over the hash functions in H. We also take $\mathbf{Y}$ to be a random variable uniformly distributed on $\{0,1\}^d$.

Call $z \in \{0,1\}^m$ *good* if $\Pr[C(\mathbf{X},\mathbf{Y}) = z] \leq 2^{-k'+1}$, and *bad* otherwise. By Proposition 2.2, $C(\mathbf{X},\mathbf{Y})$ is good with all but 2ε probability. (If z is bad, then $(\Pr[C(\mathbf{X},\mathbf{Y}) = z] - 2^{-k}) \geqslant \Pr[C(\mathbf{X},\mathbf{Y})]/2$, so each bad z contributes at least half of its probability mass to the distance from min-entropy k.) Note that if z is good, then $S_z = \{(x,y) \in \mathrm{supp}(\mathbf{X},\mathbf{Y}) : C(x,y) = z\}$ is of size $2^{k+d} \cdot \Pr[C(\mathbf{X},\mathbf{Y}) = z] \leqslant 2^{k+d-k'+1}$.

Call h *good* with respect to (x,y) if $h(x',y') \neq h(x,y)$ for all $(x',y') \in S_z \setminus \{(x,y)\}$, where $z = C(x,y)$; that is, (x,y) does not collide with any other element of S_z under h. Notice that if $z = C(x,y)$ is good, then

$$\begin{aligned}
\Pr[\mathbf{H} \text{ is bad w.r.t. } (x,y)] &\leqslant \sum_{(x',y') \in S_z \setminus \{(x,y)\}} \Pr[\mathbf{H}(x',y') = \mathbf{H}(x,y)] \\
&\leqslant \frac{|S_z|}{2^{m'-1}} \\
&\leqslant \frac{2^{k+d-k'+1}}{2^{m'-1}} \\
&\leqslant \varepsilon.
\end{aligned}$$

Since $C(\mathbf{X},\mathbf{Y})$ is good with all but 2ε probability, we conclude that $\mathbf{H}$ is good with respect to $(\mathbf{X},\mathbf{Y})$ with all but 3ε probability.

Now, for every (x,y,h) such that h is good with respect to (x,y), we have that $C'(x;y,h) = (C(x,y), h, h(x,y))$ uniquely determines $(x;y,h)$ among the elements in the support of $(\mathbf{X},\mathbf{Y},\mathbf{H})$. In particular, $C'(x;y,h)$ has probability mass exactly $2^{-(k+d')}$ under $C'(\mathbf{X};\mathbf{Y},\mathbf{H})$.

We have shown that except with 3ε probability, we hit an output string with probability mass $2^{-(k+d')}$. This implies that $C'(\mathbf{X};\mathbf{Y},\mathbf{H})$ is 3ε-close to having min-entropy $k + d'$, as required. ∎

Applying this transformation to the condenser from Theorem 3.2, we obtain our second main theorem, restated here:

Theorem 4.3 (Thm. 1.2, strengthened)**.** *For every constant $\alpha > 0$, there is a constant c such that the following holds. For all positive integers n, m and all $\varepsilon > 0$ satisfying $n \geqslant m \geqslant c\log(n/\varepsilon)$, there is an explicit construction of a function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$, with $d = O(\log n + \log(1/\varepsilon))$, such that for all $k \leqslant (1-\alpha)m$, C is a*

$$k \to_{\varepsilon} k + d$$

lossless condenser.

Proof. Let $\varepsilon_0 = \varepsilon/6$, $\alpha_0 = \alpha/2$, $t = \log\lceil(2n^2/\varepsilon_0)^{1/\alpha_0}\rceil$, $d_0 = \lfloor(1+\alpha_0)t\rfloor$, $n_0 = \lceil n/d_0\rceil$, and $m_0 = \lfloor m/d_0\rfloor - 20$. Then, Theorem 3.2 gives us $C_0 : \{0,1\}^{n_0 d_0} \times \{0,1\}^{d_0} \to \{0,1\}^{m_0 d_0}$ such that for all positive integers $\ell \in [\log(1/\varepsilon_0)/t, m_0]$, C_0 is a

$$(k_0 = \ell t + \log(1/\varepsilon_0)) \to_{3\varepsilon_0} k_0 - 1$$

condenser.

Since $n_0 d_0 \geqslant n,$ we can view C_0 as having source length n (padding any input with zeroes). To obtain our condenser C, we combine C_0 with an almost 2-universal hash function as in Lemma 4.2. We use hash functions with output length $m' = d_0 + t + 3\log(1/\varepsilon_0) + O(1)$, so the number of bits needed to sample from H is $2m' + 2\log n + 2\log m' + O(1)$. The resulting condenser C has seed length $O(\log n + \log(1/\varepsilon_0))$ and has output length at most $m_0 d_0 + 2m' + 2\log n + 2\log m' + O(1) \leqslant m_0 d_0 + 20 d_0 \leqslant m.$

We now argue that it is lossless. Consider any min-entropy threshold $k \leqslant (1-\alpha)m$. First, note that

$$k \leqslant (1-\alpha)m \leqslant (1-\alpha)(m_0+20)d_0 \leqslant (1-\alpha)(m_0+20)(1+\alpha_0)t \leqslant m_0 t,$$

where the last inequality follows from the fact that $m \geqslant c\log(n/\varepsilon_0)$. Thus we can view C_0 as a condenser for sources of min-entropy k by setting $\ell = \lfloor(k - \log(1/\varepsilon_0))/t\rfloor \in [\log(1/\varepsilon_0)/t, m_0]$. The entropy loss will be at most $d_0 + t + 2\log(1/\varepsilon_0)$ bits. This is because we lose the d_0 bits of the seed, at most t bits due to rounding ℓ down, and in case $k < t + 2\log(1/\varepsilon_0)$ we can lose all of the min-entropy (because then ℓ is too small for C_0 to work).

Since we have chosen hash functions with output length $m' = d + t + 3\log(1/\varepsilon_0) + O(1)$, we will recover all of the min-entropy, by Lemma 4.2.

□

5 Extractors that are optimal up to constant factors

Once we have condensed almost all of the entropy into a source with entropy rate close to 1 (as in Theorem 1.1), extracting (most of) that entropy is not that difficult. All we need to do is to compose the condenser with an extractor that works for entropy rates close to 1. The following standard fact makes this formal:

Proposition 5.1. *Suppose $C : \{0,1\}^n \times \{0,1\}^{t_1} \to \{0,1\}^{n'}$ is an $(n,k) \to_{\varepsilon_1} (n',k')$ condenser, and $E : \{0,1\}^{n'} \times \{0,1\}^{t_2} \to \{0,1\}^m$ is a (k', ε_2)-extractor, then $E \circ C : \{0,1\}^n \times \{0,1\}^{t_1+t_2} \to \{0,1\}^m$ defined by $(E \circ C)(x, y_1, y_2) \stackrel{\text{def}}{=} E(C(x,y_1), y_2)$ is a $(k, \varepsilon_1 + \varepsilon_2)$-extractor.*

For the best dependence on the error parameter ε, the extractor we will use is due to Zuckerman:

Theorem 5.2 ([Zuc97])**.** *For all constants $\alpha, \delta > 0$: for all positive integers n, k and all $\varepsilon > \exp(-n/2^{O(\log^* n)})$, there is an explicit construction of a $(k = \delta n, \varepsilon)$ extractor $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ with $t = O(\log n + \log(1/\varepsilon))$ and $m \geqslant (1-\alpha)k$.*

We now prove our main extractor theorem, restated here:

Theorem 5.3 (Thm. 1.3, restated)**.** *For all constants $\alpha > 0$: for all positive integers n, k and all $\varepsilon > \exp(-n/2^{O(\log^* n)})$, there is an explicit construction of a (k, ε) extractor $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = O(\log n + \log \frac{1}{\varepsilon})$ and $m \geqslant (1-\alpha)k$.*

Proof. Consider the condenser of Theorem 1.1, with its parameter ε set to the one sixth of the present ε, and its parameter α set to (say) $1/2$. This condenser has seed length $d \leqslant 3t/2$ where $t = O(\log n + \log(1/\varepsilon))$. We set its parameter $\ell = \lfloor (k - \log(6/\varepsilon))/t \rfloor$. The result is a $k \to_{2\varepsilon} k - t - 1$ condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$, with $m \leqslant (3/2)(k - \log(6/\varepsilon)) \leqslant 3k/2$. (The loss of up to t bits comes from the rounding.) We may assume that $k - t - 1 \geqslant (1 - \alpha/2)k$, or else a trivial extractor that outputs its seed of length $\lceil 2(t+1)/\alpha \rceil$ would satisfy the theorem. Applying Proposition 5.1 to this condenser and the extractor of Theorem 5.2 (with its error parameter ε set to half the present ε) gives the claimed extractor. $\square$

In the fairly common case that ε is a constant, we can use the much simpler "expander-walk" extractor (in place of the extractor of Theorem 5.2) which extracts almost all of the entropy for entropy rates close to 1. Note that our condenser from Theorem 1.1 achieves a constant entropy rate arbitrarily close to 1, and so can be combined with any extractor for such high min-entropy rates. A standard construction achieving this is based on expander walks [Gil98, Zuc97, Zuc06]. Specifically, such an extractor can be obtained by combining the equivalence between extractors and 'averaging samplers' [Zuc97], and the fact that expander walks are an averaging sampler, as established by the Chernoff bound for expander walks [Gil98]. [4]

Theorem 5.4. *For all constants $\alpha, \varepsilon > 0$, there is a constant $\delta < 1$ for which the following holds: for all positive integers n, there is an explicit construction of a $(k = \delta n, \varepsilon)$ extractor $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ with $t \leqslant \log(\alpha n)$ and $m \geq (1-\alpha)n$.*

For completeness, we present the short proof:

Proof. Let $m = \lceil (1-\alpha)n \rceil$, and for some absolute constants $c > 1$ and $\lambda < 1$, let G be an explicit 2^c-regular expander on 2^m vertices (identified with $\{0,1\}^m$) and second eigenvalue $\lambda = \lambda(G) < 1$. Let L be the largest power of 2 at most $(n-m)/c$ (so $L > (n-m)/(2c)$), and let $t = \log L \leqslant \log(\alpha n)$. The extractor E is constructed as follows. Its first argument x is used to describe a walk $v_1, v_2, \ldots, v_L$ of length L in G by picking v_1 based on the first m bits of x, and each further step of the walk from the next c bits of x — so in all, L must satisfy $n = m + (L-1)c$. The seed y is used to pick one of the vertices of the walk at random. The output $E(x, y)$ of the extractor is the m-bit label of the chosen vertex.

Let $\mathbf{X}$ be a random variable with minentropy $k = \delta n$. We wish to prove that for any $S \subseteq \{0,1\}^m$, the probability that $E(\mathbf{X}, \mathbf{U_t})$ is a vertex in S is in the range $\mu \pm \varepsilon$ where $\mu = |S|/2^m$. Fix any such subset S. Call an $x \in \{0,1\}^n$ "bad" if

$$\left| \Pr_y[E(x,y) \in S] - \mu \right| > \varepsilon/2.$$

[4] The papers [IZ89, CW89] prove hitting properties of expander walks, and observe that these imply objects related to (but weaker than) extractors, known as dispersers.

The known Chernoff bounds for random walks on expanders [Gil98] imply that the number of bad x's is at most

$$2^n \cdot e^{-\Omega(\varepsilon^2(1-\lambda)L)} = 2^n \cdot e^{-\Omega(\varepsilon^2(1-\lambda)\alpha n/c)} = 2^n \cdot 2^{-\Omega(\varepsilon^2 \alpha n)}$$

(since c, λ are absolute constants). Therefore the probability that $\mathbf{X}$ is bad is at most $2^{-\delta n} \cdot 2^n \cdot 2^{-\Omega(\varepsilon^2 \alpha n)}$, which is exponentially small for large enough $\delta < 1$. Therefore

$$|\Pr[E(\mathbf{X}, \mathbf{U_t}) \in S] - \mu| \leq \varepsilon/2 + 2^{-\Omega(n)} \leq \varepsilon,$$

implying that E is a (k, ε)-extractor. $\square$

Combining Theorem 1.1 with Theorem 5.4 via Proposition 5.1, as in the proof of Theorem 1.3, we obtain the following extractor, which has the advantage that its proof is short and self-contained:

Theorem 5.5. *For every constant $\alpha > 0$: for all positive integers n, k, and all constant $\varepsilon > 0$, there is an explicit construction of a (k, ε) extractor $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = O(\log n + \log(1/\varepsilon))$ and $m \geqslant (1-\alpha)k$.*

6 Variations on the main condenser

In this section we show how minor modifications to the proof allow us to optimize the seed length or the output entropy. We also show that a small modification to the construction yields condensers from Reed-Solomon codes.

6.1 Optimizing the seed length

The condenser of Theorem 1.1 retains all the source minentropy (except for 1 bit) and achieves an entropy rate of $1 - \delta$ for any desired $\delta > 0$. Its main shortcoming is the large seed length, which is greater than $(\log n)/\delta$, whereas the optimal condenser achieves a seed length of $\log n + \log(1/\varepsilon) + O(1)$.

We now show that the seed length can be improved to $(1+\gamma)(\log n + \log(1/\varepsilon))$ — the new condenser still retains a $(1 - O(\frac{1}{\log n}))$ fraction of the input entropy and the output entropy rate is $\Omega(\gamma)$. While the entropy rate is not close to 1 as it was before, it is still a constant, and extractors with seed length of $1 \cdot \log n + O(1)$ were recently constructed for sources of any constant minentropy rate, and constant error ε [Zuc06] (Theorem 6.4 below.) Composing the condenser with such an extractor gives an extractor that extracts $(1-\alpha)k$ bits from a source with minentropy k, using seed length $(1+\gamma)\log n + \log k + O(1)$, for arbitrary constants $\alpha, \gamma > 0$. Note that when $k = n^{o(1)}$, the seed length is near-optimal.

The improved analysis that permits us to optimize the seed length is in the following lemma (compare to Lemma 3.1):

Lemma 6.1. *Defining $\mathrm{BAD}(S, \varepsilon)$ and $B(K, \varepsilon)$ with respect to C as in Lemma 2.5, for any integer parameter $s \geq 1$, we have*

$$B\left(K = \left\lfloor \frac{h^m - 1}{\binom{m+s-1}{s-1}} \right\rfloor, \varepsilon\right) \leq h^m - 1,$$

provided $q \geq m(n-1)(h-1)/(s\varepsilon)$.

Proof. Let $S \subseteq \mathbb{F}_q^m$ be an arbitrary set of size at most K. The proof follows along the lines of the proof of Lemma 3.1, with the main change being that we make sure that the interpolated polynomial $Q(Z_1, Z_2, \ldots, Z_m)$ has a root of multiplicity at least s at each element $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in S$. (Note that Lemma 3.1 is the special case of the current theorem with $s = 1$.) By a 'root of multiplicity at least s', we mean that that the polynomial

$$Q_\alpha(Z_1, \ldots, Z_m) \stackrel{\text{def}}{=} Q(\alpha_1 + Z_1, \ldots, \alpha_m + Z_m)$$

has no monomials of degree $s-1$ or smaller with nonzero coefficients, which amounts to $\binom{m+s-1}{s-1}$ homogeneous linear constraints on the coefficients of Q. Since $h^m > |S|\binom{m+s-1}{s-1}$, such a nonzero polynomial Q of degree at most $(h-1)$ in each variable exists. Fix Q to be any such nonzero polynomial.

Suppose $f(Y) \in \mathrm{BAD}(S, \varepsilon)$. Let $y \in \mathbb{F}_q$ be such that $C(f, y) \in S$. Then, by the choice of Q,

$$Q(f_0(y), f_1(y), \ldots, f_{m-1}(y)) = Q(C(f, y)) = 0.$$

In fact, since $C(f, y)$ is a root of multiplicity s, we can show that the the polynomial

$$R_f(Y) \stackrel{\text{def}}{=} Q(f_0(Y), f_1(Y), \ldots, f_{m-1}(Y))$$

has a root of multiplicity s at y. To see this, note that

$$\begin{aligned} R_f(y+Y) &= Q(f_0(y+Y), f_1(y+Y), \ldots, f_{m-1}(y+Y)) \\ &= Q(f_0(y) + Y \cdot g_0(Y), f_1(y) + Y \cdot g_1(Y), \ldots, f_{m-1}(y) + Y \cdot g_{m-1}(Y)) \\ &= Q_{C(f,y)}(Y \cdot g_0(Y), Y \cdot g_1(Y), \ldots, Y \cdot g_{m-1}(Y)) \end{aligned}$$

for some polynomials $g_0, \ldots, g_{m-1}$. Since every monomial in $Q_{C(f,y)}$ has degree at least s, when we substitute $Y \cdot g_i(Y)$ for the variables we get a univariate polynomial divisible by Y^s. Thus $Y^s | R_f(y+Y)$, i.e. R_f has a root of multiplicity s at y. Equivalently, $(Y-y)^s | R_f(Y)$. We conclude that if $f(Y) \in \mathrm{BAD}(S, \varepsilon)$, i.e., if

$$\Pr_y[Q(f_0(y), f_1(y), \ldots, f_{m-1}(y)) = 0] > \varepsilon \,,$$

then $R(Y)$ has more than εsq roots counting multiplicities. On the other hand the degree of $R(Y)$ is at most $(n-1)(h-1)m$. Therefore, since $\varepsilon sq \geq (n-1)(h-1)m$, we must have $R(Y) = 0$.

From this point on, the proof proceeds identically to that of Theorem 1.1, leading to the desired conclusion $|\mathrm{BAD}(S, \varepsilon)| \leq h^m - 1$. □

Picking parameters suitably, and following the outline of the proof of Theorem 1.1, we obtain the following condenser:

Theorem 6.2. *For every $\gamma > 0$: for all positive integers $n \geqslant \ell$ and all $\varepsilon > 0$, there is an explicit construction of a*

$$(k = \ell t + \log(1/\varepsilon)) \rightarrow_{2\varepsilon} (k - 3\ell - 1)$$

condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\ell d}$ with $t = \log\lceil (2n/\varepsilon)^\gamma \rceil$ and $d = \lfloor (1 + 1/\gamma) t \rfloor$, provided $t \geq 4$ and $\ell t \geqslant \log(1/\varepsilon)$.

Proof. We describe how to set parameters, and then apply Lemmas 6.1 and 2.5. We set $h = \lceil (2n/\varepsilon)^\gamma \rceil$, $t = \log h$, $d = \lfloor (1 + 1/\gamma)t \rfloor$, and $q = 2^d$. Set $m = s = \ell$. We have $q \geq nmh/(\varepsilon s) = nh/\varepsilon$ as required.

By Lemma 6.1, and Lemma 2.5, C is a

$$\log((h^m - 1)/\varepsilon) \rightarrow_{2\varepsilon} \log(K/\varepsilon) - 1$$

condenser. Now, $K = \lfloor (h^m - 1)/\binom{2m-1}{m-1} \rfloor \geq (h^m - 1)/2^{2m-1} - 1 \geq (h/8)^m$, as long as $h \geq 10$. The theorem follows, using the fact that $\log(h^m) = \ell t$ and $\log(h/8)^m = \ell \cdot (t - 3)$. $\square$

In the previous theorem, γ may be subconstant, and in the following corollary we show that it can be set to produce an a seed length that is optimal up to the *additive* constant, while still retaining a constant fraction of the minentropy, at the expense of an output entropy rate of $\Omega(1/\log(n/\varepsilon))$, which is subconstant, but still quite good.

Corollary 6.3. *For every integer constant $c \geq 4$: for all positive integers $n \geqslant \ell$ and all $\varepsilon > 2^{-c\ell}$, there is an explicit construction of a*

$$\left(k = c\ell + \log \frac{1}{\varepsilon}\right) \rightarrow_{2\varepsilon} \left(\left(1 - \frac{3}{c}\right) k - 1\right)$$

condenser $C : \{0,1\}^n \times \{0,1\}^d \rightarrow \{0,1\}^{n'}$ with $d = \log n + \log(1/\varepsilon) + O(1)$ and $n' = \left(1 + \frac{\log(2n/\varepsilon)}{c}\right) c\ell$.

Proof. Set $\gamma = c/\log(2n/\varepsilon)$ in Theorem 6.2. $\square$

We now combine the condenser of Theorem 6.2 with Zuckerman's recent extractor. (This extractor in turn starts by applying a condenser due to Raz [Raz05] that has constant seed length and can increase the entropy rate from δ to $1 - \delta$ for any constant $\delta > 0$, while retaining a constant fraction of the minentropy.)

Theorem 6.4 ([Zuc06])**.** *For all constants $\alpha, \delta, \varepsilon > 0$: for all positive integers n, there is an explicit construction of a $(k = \delta n, \varepsilon)$ extractor $E : \{0,1\}^n \times \{0,1\}^d \rightarrow \{0,1\}^m$ with seed length $d = \log n + O(1)$ and output length $m \geqslant (1 - \alpha)k$.*

Combining Theorem 6.2 with Theorem 6.4 via Proposition 5.1, as in the proof of Theorem 1.3, we obtain the following extractor, which has a near-optimal seed length:

Theorem 6.5. *For all constants $\alpha, \gamma, \varepsilon > 0$: for all positive integers n, k, there is an explicit construction of a (k, ε) extractor $E : \{0,1\}^n \times \{0,1\}^d \rightarrow \{0,1\}^m$ with seed length $d = (1 + \gamma)\log n + \log k + O(1)$ and output length $m \geqslant (1 - \alpha)k$, provided $k \geq cd/\alpha$ for a universal constant c.*

6.2 Increasing the output entropy

The condenser of Theorem 1.1 is missing only the entropy of the seed, which is small enough that it can be "recovered" using the hashing technique of Lemma 4.2. However, one can ask how far our new proof technique can go in isolation. More precisely, we modify the function C as follows

$$C'(f, y) \stackrel{\text{def}}{=} (y, C(f, y)),$$

and ask how much entropy is retained for this "strong" variant of the basic construction. In the language of Lemma 2.5, ideally we could hope for $B(K, \varepsilon) \leq K/q$, when the seed length is $\log q$. This would correspond to recovering all of the entropy of the source and seed together.

In this section we show that a minor modification to the proof allows us to argue that $B(K,\varepsilon) \le K/r$ for r approaching εq. This corresponds to recovering all but $\log(1/\varepsilon) + O(1)$ of the total entropy, although we don't know of a direct application for this improvement. We show the improved result by recording a variant of Lemma 3.1 for C' as defined above:

Lemma 6.6. *Defining* $\mathrm{BAD}(S,\varepsilon)$ *and* $B(K,\varepsilon)$ *with respect to* C' *as in Lemma 2.5, we have*

$$B(K = rh^m - 1, \varepsilon) < K/r,$$

for any positive integer r *such that* $q \ge [(n-1)(h-1)m + r]/\varepsilon$.

Proof. Fix a set $S \subseteq \mathbb{F}_q \times \mathbb{F}_q^m$ of size at most K. Let $Q \in \mathbb{F}_q[Y, Z_1, Z_2, \ldots, Z_m]$ be a nonzero $m+1$-variate polynomial that vanishes on S, with degree at most $r-1$ in Y, and individual degrees at most $h-1$ for the remaining m variables. By definition, for every $f(Y) \in \mathrm{BAD}(S,\varepsilon)$, it holds that

$$\Pr_y[Q(y, f_0(y), f_1(y), \ldots, f_{m-1}(y)) = 0] > \varepsilon.$$

Therefore, the univariate polynomial $R_f(Y) \stackrel{\text{def}}{=} Q(Y, f_0(Y), \ldots, f_{m-1}(Y))$ has more than εq zeroes, and degree at most $r + (n-1)(h-1)m$. Since $r + (n-1)(h-1)m \le \varepsilon q$, $R_f(Y)$ must be identically zero, and so $Q(Y, f_0(Y), \ldots, f_{m-1}(Y)) = 0$ for every bad $f(Y)$.

Now, view Q as a polynomial in $\mathbb{F}_q[Y][Z_1, Z_2, \ldots, Z_m]$, and factor out the largest power of $E(Y)$. Since $E(Y)$ has no roots in $\mathbb{F}_q$, the resulting polynomial still vanishes on S. Also, the resulting polynomial is non-zero modulo $E(Y)$; let Q^* be the resulting polynomial after reducing modulo $E(Y)$.

Now, view Q^* as a multivariate polynomial (in variables $Z_1, Z_2, \ldots, Z_m$) over the extension field $\mathbb{F} = \mathbb{F}_q[Y]/E(Y)$, and define

$$Q^{**}(Z) = Q^*(Z, Z^h, Z^{h^2}, \ldots, Z^{h^{m-1}}).$$

Because the individual degrees of Q^* are all less than h, Q^{**} is a non-zero polynomial (because distinct monomials in Q^* map to distinct monomials in Q^{**}).

For every $f(Y) \in \mathrm{BAD}(S,\varepsilon)$, now viewed as an element of $\mathbb{F}$, we have $Q^{**}(f) = 0$; i.e., f is a root of Q^{**}. Thus $|\mathrm{BAD}(S,\varepsilon)| \le \deg(Q^{**})$. The degree of Q^{**} is at most

$$(h-1)(1 + h + h^2 + \cdots + h^{m-1}) = h^m - 1 < K/r.$$

□

6.3 Reed-Solomon version

We use one of the main ideas from [GR06] to argue that a small modification to our construction gives a good condenser from Reed-Solomon codes, answering a question raised in [KU06].

Let q be an arbitrary prime power, and let $\zeta \in \mathbb{F}_q$ be a generator of the multiplicative group $\mathbb{F}_q^*$. Then the polynomial $E(Y) = Y^{q-1} - \zeta$ is irreducible over $\mathbb{F}_q$ [LN86, Chap. 3, Sec. 5]. The following identity holds for all $f(Y) \in \mathbb{F}_q[Y]$:

$$f(Y)^q \equiv f(Y^q) \equiv f(Y^{q-1}Y) \equiv f(\zeta Y) \pmod{E(Y)}.$$

In this case, if we modify our basic function $C : \mathbb{F}_q^n \times \mathbb{F}_q \to \mathbb{F}_q^m$ slightly so that we raise f to successive powers of q rather than h, we get:

$$\begin{aligned} C(f,y) &\stackrel{\text{def}}{=} (f(y), (f^q \bmod E)(y), (f^{q^2} \bmod E)(y), \cdots, (f^{q^{m-1}} \bmod E)(y) \\ &= (f(y), f(\zeta y), \cdots, f(\zeta^{m-1} y)). \end{aligned} \tag{2}$$

In other words, our function interprets its first argument as describing a univariate polynomial over $\mathbb{F}_q$ of degree at most $n-1$ (i.e., a Reed-Solomon codeword), it uses the seed to select a random location in the codeword, and it outputs m successive symbols of the codeword. This is precisely the analogue of the Shaltiel-Umans q-ary extractor construction [SU05] for univariate polynomials, rather than multivariate polynomials.

With a minor modification to the proof of Lemma 3.1, we show that this is good condenser:

Lemma 6.7. *Defining* $\mathrm{BAD}(S,\varepsilon)$ *and* $B(K,\varepsilon)$ *with respect to the function* C *of Equation (2) as in Lemma 2.5, we have*

$$B(K = h^m - 1, \varepsilon) \le (q^m - 1)(h-1)/(q-1),$$

provided $q \ge (n-1)(h-1)m/\varepsilon$.

Proof. The proof is the same as the proof of Lemma 3.1 except that we define Q^* differently:

$$Q^*(Z) \stackrel{\text{def}}{=} Q(Z, Z^q, Z^{q^2}, \ldots, Z^{q^{m-1}}).$$

As before, every $f(Y) \in \mathrm{BAD}(S,\varepsilon)$, is a root of Q^*. Thus $|\mathrm{BAD}(S,\varepsilon)| \le \deg(Q^*)$. The degree of Q^* is at most

$$(h-1)(1 + q + q^2 + \cdots + q^{m-1}) = (h-1)((q^m-1)/(q-1)).$$

□

We obtain the following condenser:

Theorem 6.8 (Reed-Solomon condenser)**.** *For every constant* $\alpha > 0$*: for all positive integers* $n \geqslant \ell$ *and all* $\varepsilon > 0$*, there is an explicit construction of a*

$$(\ell d + \log(1/\varepsilon)) \to_{3\varepsilon} (\ell t + \log(1/\varepsilon) - 1)$$

condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\ell d}$ *with* $t = \lceil \log(2n\ell/\varepsilon)^{1/\alpha} \rceil$ *and* $d = \lfloor (1+\alpha)t \rfloor$*, provided* $\ell t \geqslant \log(1/\varepsilon)$.

The main difference between this theorem and our basic condenser (Theorem 1.1) is that the input and output min-entropies no longer differ by one bit. Instead, the ratio is roughly $d/t \approx (1+\alpha)$, which means that we retain only a $1/(1+\alpha)$ fraction of the min-entropy.

Proof. The proof is identical to that of Theorem 1.1, with the only change being that we fix $n' = \ell = m$, and due to the difference between Lemma 6.7 and Lemma 3.1, the input min-entropy required is

$$\log(q^m/\varepsilon) = \ell d + \log(1/\varepsilon).$$

□

For the Reed-Solomon-based construction, a relatively simple argument shows that the entropy rate and the ratio of output minentropy to input minentropy must both be constants less than 1. The example below comes from [GHSZ02, TZ04]:

Theorem 6.9. *For every positive integer* p *such that* $p|(q-1)$*, there is a source* $\mathbf{X}$ *with minentropy at least* $\lfloor n/p \rfloor \log q$ *for which the support of* $C(\mathbf{X}, \mathbf{U_t})$*, as defined in Equation (2), is entirely contained within a set of size* w^m*, where* $w = (q-1)/p + 1$*. Thus* $C(\mathbf{X}, \mathbf{U_t})$ *is* not ε*-close to having minentropy* $\log(\frac{1}{1-\varepsilon} w^m)$*, .*

Proof. Take the source to be p-th powers of all degree $\lfloor n/p \rfloor$ polynomials. Every output symbol of C is an evaluation of such a polynomial, and therefore must be a p-th power, or 0. There are thus only $w = (q-1)/p + 1$ possible output symbols, so the output is contained within a set of size w^m, which by Proposition 2.2 is not ε-close to any distribution with minentropy $\log(\frac{1}{1-\varepsilon} w^m)$. $\square$

This example can be interpreted as follows. For any $m \leq \lfloor n/p \rfloor$, we have enough entropy to hope for C's output (which has length $m \log q$) to be close to uniform. However, if we choose $p = n^\delta$ for some constant $\delta > 0$, then the output minentropy can be no larger than $\log(O(w^m)) = m \log(q^{1-\delta'})$, for some constant $\delta' > 0$, as long as $q = \text{poly}(n)$ (which is required for seed length $O(\log n)$). This example shows that the output minentropy rate being a constant strictly less than 1, as well as the output minentropy being a constant factor smaller than the input minentropy are inherent in the present construction; they are not artifacts of the analysis. That is, it is not possible to resolve those issues by simply giving a different, improved analysis for our generic construction.

7 Conclusions

This paper introduces a new proof technique for analyzing algebraic extractor constructions, which does not rely on local decodability of the underlying error-correcting codes. It is thus natural to ask whether these new techniques can help in other settings. For example, can we use them to argue about *computational* analogues of the objects in this paper – pseudorandom generators and pseudoentropy generators? Or, can variants of our constructions yield so-called “2-source” objects, in which both the source and the seed are only weakly random?

Of course a significant remaining open problem is to construct truly optimal extractors, ones that are optimal up to *additive* constants in the seed length and/or output length. Towards this end, we wonder if there is some variant of our constructions with a better entropy rate – the next natural threshold is to have entropy *deficiency* only $k^{o(1)}$. Another interesting question is whether some variant of these constructions can give a block-wise source directly. Depending on the actual parameters, either of these two improvements have the potential to lead to extractors with optimal output length (i.e. ones extract all the minentropy). Alternatively, if we can find an extractor with optimal output length for high min-entropy (say $.99n$), then, by composing it with our condenser, we would get one for arbitrary min-entropy.

Acknowledgements. This paper began with a conversation at the BIRS workshop “Recent Advances in Computation Complexity.” We would like to thank the organizers for inviting them, and BIRS for hosting the workshop. We also thank Oded Goldreich, Prahladh Harsha, Omer Reingold, and Ronen Shaltiel for helpful comments on the write-up.

References

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

[CRVW02] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson. Randomness conductors and constant-degree expansion beyond the degree/2 barrier. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 659–668, 2002.

[CW89] A. Cohen and A. Wigderson. Dispersers, deterministic amplification, and weak random sources (extended abstract). In *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, pages 14–19, 1989.

[GHSZ02] V. Guruswami, J. Hastad, M. Sudan, and D. Zuckerman. Combinatorial bounds for list decoding. *IEEE Transactions on Information Theory*, 48(5):1021–1035, 2002.

[Gil98] D. Gillman. A Chernoff bound for random walks on expander graphs. *SIAM J. Comput.*, 27(4):1203–1220 (electronic), 1998.

[GR06] V. Guruswami and A. Rudra. Explicit capacity-achieving list-decodable codes. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 1–10, 2006.

[GS99] V. Guruswami and M. Sudan. Improved decoding of Reed-Solomon and Algebraic-Geometry codes. *IEEE Transactions on Information Theory*, 45(6):1757–1767, 1999.

[Gur04] V. Guruswami. Better extractors for better codes? In *STOC*, pages 436–444, 2004.

[GW97] O. Goldreich and A. Wigderson. Tiny families of functions with random properties: A quality-size trade-off for hashing. *Random Structures & Algorithms*, 11(4):315–343, 1997.

[GZ97] O. Goldreich and D. Zuckerman. Another proof that BPP subseteq PH (and more). Technical Report TR97-045, Electronic Colloquium on Computational Complexity, 1997.

[ISW00] R. Impagliazzo, R. Shaltiel, and A. Wigderson. Extractors and pseudo-random generators with optimal seed length. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pages 1–10, 2000.

[IZ89] R. Impagliazzo and D. Zuckerman. How to recycle random bits. In *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, pages 248–253, 1989.

[KU06] S. Kalyanaraman and C. Umans. On obtaining pseudorandomness from error-correcting codes. *Electronic Colloquium on Computational Complexity (ECCC)*, (128), 2006.

[LN86] R. Lidl and H. Niederreiter. *Introduction to Finite Fields and their applications*. Cambridge University Press, 1986.

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

[MU02] E. Mossel and C. Umans. On the complexity of approximating the vc dimension. *J. Comput. Syst. Sci.*, 65(4):660–671, 2002.

[NW94] N. Nisan and A. Wigderson. Hardness vs. randomness. *Journal of Computer and System Sciences*, 49:149–167, 1994.

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

[PV05] F. Parvaresh and A. Vardy. Correcting errors beyond the Guruswami-Sudan radius in polynomial time. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 285–294, 2005.

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

[RR99] R. Raz and O. Reingold. On recycling the randomness of states in space bounded computation. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 159–168, 1999.

[RSW06] O. Reingold, R. Shaltiel, and A. Wigderson. Extracting randomness via repeated condensing. *SIAM J. Comput.*, 35(5):1185–1209, 2006.

[RZ01] A. Russell and D. Zuckerman. Perfect information leader election in log* n+o (1) rounds. *J. Comput. Syst. Sci.*, 63(4):612–626, 2001.

[Sha02] R. Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the European Association for Theoretical Computer Science*, 77:67–, June 2002. Columns: Computational Complexity.

[Sho90] V. Shoup. New algorithms for finding irreducible polynomials over finite fields. *Mathematics of Computation*, 54(189):435–447, 1990.

[Sip88] M. Sipser. Expanders, randomness, or time versus space. *Journal of Computer and System Sciences*, 36(3):379–383, 1988.

[SU05] R. Shaltiel and C. Umans. Simple extractors for all min-entropies and a new pseudorandom generator. *Journal of the ACM*, 52(2):172–216, 2005. Conference version appeared in FOCS 2001.

[Sud97] M. Sudan. Decoding of Reed Solomon codes beyond the error-correction bound. *J. Complexity*, 13(1):180–193, 1997.

[SZ99] A. Srinivasan and D. Zuckerman. Computing with very weak random sources. *SIAM Journal on Computing*, 28:1433–1459, 1999.

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

[TSZ04] A. Ta-Shma and D. Zuckerman. Extractor codes. *IEEE Transactions on Information Theory*, 50(12):3015–3025, 2004.

[TU06] A. Ta-Shma and C. Umans. Better lossless condensers through derandomized curve samplers. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, 2006. To appear.

[TUZ01] A. Ta-Shma, C. Umans, and D. Zuckerman. Loss-less condensers, unbalanced expanders, and extractors. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 143–152, 2001.

[TZ04] A. Ta-Shma and D. Zuckerman. Extractor codes. *IEEE Transactions on Information Theory*, 50(12):3015–3025, 2004.

[TZS06] A. Ta-Shma, D. Zuckerman, and S. Safra. Extractors from Reed-Muller codes. *J. Comput. Syst. Sci.*, 72(5):786–812, 2006.

[Uma99] C. Umans. Hardness of approximating Σ_2^p minimization problems. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 465–474, 1999.

[WZ99] A. Wigderson and D. Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. *Combinatorica*, 19(1):125–138, 1999.

[Zuc96a] D. Zuckerman. On unapproximable versions of NP-complete problems. *SIAM Journal on Computing*, 25:1293–1304, 1996.

[Zuc96b] D. Zuckerman. Simulating BPP using a general weak random source. *Algorithmica*, 16(4-5):367–391, 1996.

[Zuc97] D. Zuckerman. Randomness-optimal oblivious sampling. *Random Struct. Algorithms*, 11(4):345–367, 1997.

[Zuc06] D. 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*, pages 681–690, 2006.