Extractors for Three Uneven-Length Sources

Anup Rao[⋆,1] and David Zuckerman[⋆⋆,2]

[1] School of Mathematics, Institute for Advanced Study
arao@ias.edu
[2] Department of Computer Science, University of Texas at Austin
diz@cs.utexas.edu

Abstract. We construct an efficient 3-source extractor that requires one of the sources to be significantly shorter than the min-entropy of the other two sources. Our extractors work even when the longer, n-bit sources have min-entropy $n^{\Omega(1)}$ and the shorter source has min-entropy $\log^{10} n$. Previous constructions for independent sources with min-entropy n^γ required $\Theta(1/\gamma)$ sources [Rao06]. Our construction relies on lossless condensers [GUV07] based on Parvaresh-Vardy codes [PV05], as well as on a 2-source extractor for a block source and general source [BRSW06].

1 Introduction

Motivated by the widespread use of randomness in computer science, researchers have sought algorithms to extract randomness from a distribution that is only weakly random. A general weak source is one with some min-entropy: a distribution has min-entropy k if all strings have probability at most 2^{-k}. We would like to extract randomness from a weak source knowing only k and not the exact distribution. However, this is impossible, even for more restricted sources [SV86].

Therefore, Santha and Vazirani showed how to extract randomness from two independent restricted weak sources [SV86]. Can we design such efficient *randomness extractors* for general *independent sources*? These are efficiently computable functions $\mathsf{Ext} : (\{0,1\}^n)^{\mathsf{C}} \to \{0,1\}^m$ with the property that for any product distribution $X_1, \ldots, X_{\mathsf{C}}$, the output $\mathsf{Ext}(X_1, \ldots, X_{\mathsf{C}})$ is close to uniformly distributed as long as each X_i has high enough min-entropy. Our primary goals are to minimize the number of sources required and the amount of entropy needed. Secondarily, we'd like to maximize the length of the output and minimize the error, which is the distance of the output from uniform.

Extractors for independent sources have been useful in constructing deterministic extractors for space-bounded sources [KRVZ06] and in new constructions of network extractor protocols [KLRZ08].

1.1 Previous Results

The question of finding such an extractor came up as early as in the works of Santha and Vazirani [SV86] and Chor and Goldreich [CG88]. After that, following the work of Nisan and Zuckerman [NZ96], most work focused on constructing *seeded* extractors. This is simply a two source extractor where one source is assumed to be very short and uniformly distributed (in this case the problem only makes sense if the extractor outputs more than what's available in the short seed). Extractors for this model have found application in constructions of communication networks and good expander graphs [WZ99,CRVW02], error correcting codes [TZ04,Gur04], cryptographic protocols [Lu04,Vad04], data structures [MNSW98] and samplers [Zuc97]. Seeded extractor constructions are now available that can extract uniform bits from a source with small entropy using a seed of length only $O(\log n)$ [LRVW03,GUV07].

In recent years there have been several works improving the state of the art for independent sources [BIW04,BKS$^+$05,Raz05,Bou05,Rao06,BRSW06]. We now know how to extract from two sources when the entropy in each is at least something like $.4999n$ [Bou05], from three sources if the entropy in each is at

⋆ Supported in part by NSF Grants CCF-0634811 and CCR-0324906.
⋆⋆ Supported in part by NSF Grant CCF-0634811.

least $n^{0.99}$ [Rao06] and from $O(1/\gamma)$ sources if the entropy in each is at least n^γ [Rao06,BRSW06]. All of these constructions have exponentially small error, and by a result of Shaltiel [Sha06], the output length can be made almost best possible. Thus, the tradeoff that is most interesting is the one between the number of sources required and the entropy requirements.

1.2 Our Work

In this paper, we construct extractors which require only three sources with polynomial min-entropy. However, there is a key caveat: one of the sources must be significantly shorter than the min-entropy of the other two sources. On the plus side, while the longer, n-bit sources must have min-entropy $n^{\Omega(1)}$, the shorter source need only have min-entropy $\log^{10} n$.

Extractors for uneven-length sources may be more interesting than they appear, for two reasons. First, the extractors with the most applications are seeded extractors, which are extractors for uneven-length sources. Second, in other settings the uneven-length case was more difficult. For example, in the two-source setting with entropy rate bigger than 1/2, the even length case was known for decades (before extractors were defined), but the uneven-length case was only proved by Raz in 2005 [Raz05].

We now state our three source extractor precisely.

Theorem 1 (3-Source Extractor). *There exists a constant d and a polynomial time computable function $\mathsf{3Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \times \{0,1\}^{n_3} \to \{0,1\}^m$ which is an extractor for three sources with min-entropy requirements $k_1, k_2, k_3 = \sqrt{k_1}$, error $2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}}$ and output length $m = k_1 - o(k_1)$ as long as:*

- $\frac{\log k_1}{\log n_2} > d \frac{\log(n_1+n_3)}{\log k_1}$
- $k_2 > d \log n_1$

One example of how to set parameters is in the following corollary:

Corollary 1. *There exist constants d, h such that for every constant $0 < \gamma < 1$, there is a polynomial time computable function $\mathsf{3Ext} : \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^{n^{\gamma/h}} \to \{0,1\}^{n^\gamma - o(n^\gamma)}$ which is an extractor for three sources with min-entropy requirements $k = n^\gamma, n^\gamma, \log^d n$, and error $2^{-\Omega(\log^{10} n)}$.*

The corollary follows by setting $n_1 = n_3 = n$, $n_2 = n^{\gamma/h}$, $k_1 = n^\gamma, k_2 = \log^d n$ and choosing h to be a large enough constant.

For smaller min-entropy, the first constraint in the theorem forces the length of the shorter source to be much shorter than the other two sources.

It turns out that we don't really need three mutually independent sources to get our results. We obtain an extractor even when we have just two sources, but one of them is a block source with a short block followed by a long block.

Theorem 2 ((Short, Long)-Block-Source Extractor). *There exists a constant d and a polynomial time computable function $\mathsf{3Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \times \{0,1\}^{n_3} \to \{0,1\}^m$ which is an extractor for a block source with min-entropy requirements k_1, k_2 and an independent source with min-entropy $k_3 = \sqrt{k_2}$, error $2^{-\Omega(k_1)} + 2^{-k_2^{\Omega(1)}}$ and output length $m = k_2 - o(k_2)$ as long as:*

- $\frac{\log k_2}{\log n_1} > d \frac{\log(n_2+n_3)}{\log k_2}$
- $k_1 > d \log n_2$

Salil Vadhan observed that a slightly different analysis allows us to reverse the order of the short and long sources.

Theorem 3 ((Long, Short)-Block-Source Extractor). *There exists a constant d and a polynomial time computable function $\mathsf{3Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \times \{0,1\}^{n_3} \to \{0,1\}^m$ which is an extractor for a block source with min-entropy requirements k_1, k_2 and an independent source with min-entropy $k_3 = \sqrt{k_1}$, error $2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}}$ and output length $m = k_1 - o(k_1)$ as long as:*

- $\frac{\log k_1}{\log n_2} > d \frac{\log(n_1+n_3)}{\log k_1}$
- $k_2 > d \log n_1$

1.3 Techniques

We build on the work of Guruswami et al. [GUV07] and Barak et al. [BRSW06]. Guruswami et al. showed how to build a very good *seeded condenser*. This is a function $\mathsf{Cond} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ that guarantees that if X has sufficient min-entropy k and U_d is an independent random variable that's uniform on d bits, $\mathsf{Cond}(X, U_d)$ is close to having very high min-entropy. Inspired by ideas from the list decodable code constructions of [PV05,GR06], they showed that the following function is a condenser: $\mathsf{Cond}(f, y) = f(y), f^h(y), \ldots, f^{h^t}(y)$, where f is a low degree polynomial, f^{h^i} is the powered polynomial modulo a suitably chosen irreducible polynomial and h, t are suitably chosen parameters. Specifically they showed that the output of this function is close to having min-entropy $0.9m$.

We analyze Cond when the second source is *not* uniform, but merely has high entropy (say αd). In this case, we show that the output must have entropy close to $0.9\alpha m$. At first this may not seem useful, since we end up with a distribution where the entropy is even less concentrated than the source that we started with. However, we show that if the output is $m = m_1 + m_2 + \cdots + m_{\mathsf{C}}$ bits, each consecutive block of m_i bits must have entropy close to $0.9\alpha m_i$. Using an idea from [Zuc96], we choose the m_i's to increase geometrically. This implies that the output must be a *block source*: not only does each consecutive block of m_i bits have a reasonable amount of entropy, it has significant entropy conditioned on any fixing of the previous blocks. Intuitively, since each subsequent block is significantly larger than the previous ones, the previous ones cannot contain enough information to significantly reduce the entropy in the current block upon conditioning.

Block sources are a well studied object in extractor constructions. Indeed, earlier works have shown [BKS$^+$05,Rao06,BRSW06] that even two source extractors are easy to obtain under the assumption that both or even one of the sources have some block source structure. In particular, a theorem from [BRSW06] shows that we can extract from one block source and one independent source, if each block and the independent source have entropy n^γ, and the number of blocks in the block source is at least $O(1/\gamma)$.

This completes the construction. We first apply Cond to convert the first two sources into a single block source, and then use the extractor from [BRSW06] and an additional source to get random bits.

2 Preliminaries

For a distribution X, we let $H_\infty(X)$ denote the min-entropy of the distribution. We call a distribution *flat* if it is uniformly distributed on some subset of the universe.

Fact 1 *Every distribution X with min-entropy at least k is a convex combination of flat distributions with min-entropy k.*

Definition 1. *Let D and F be two distributions on a set S. Their* statistical distance *is*

$$|D - F| \stackrel{def}{=} \max_{T \subseteq S}(|D(T) - F(T)|) = \frac{1}{2}\sum_{s \in S}|D(s) - F(s)|$$

*If $|D - F| \le \epsilon$ we shall say that D is ϵ-*close *to F.*

This measure of distance is nice because it is robust in the sense that if two distributions are close in this distance, then applying any functions to them cannot make them go further apart.

Proposition 1. *Let D and F be any two distributions over a set S s.t. $|D - F| \le \epsilon$. Let g be any function on S. Then $|g(D) - g(F)| \le \epsilon$.*

When we manipulate sources which are close to having some min-entropy, it will be convenient to have the following definition. For a distribution D, $D(a)$ denotes the probability that D places on a.

Definition 2. *We call a distribution D ϵ-*close to k-light *if, when X is chosen according to D,* $\Pr[D(X) > 2^{-k}] \le \epsilon$.

The following is then immediate:

Lemma 1. *If D is ϵ-close to k-light, then D is ϵ-close to min-entropy k.*

The following lemma gives a sufficient condition to lowerbound the min-entropy of a source.

Lemma 2 ([GUV07]). *Let X be a random variable taking values in a set of size larger than 2^k such that for every set S of size less than $\epsilon 2^k$, $\Pr[X \in S] < \epsilon$. Then X is ϵ-close to k-light.*

Proof. First note that $|\mathsf{supp}(X)| \geq \epsilon 2^k$, or else the hypothesis of the lemma is contradicted by setting $S = \mathsf{supp}(X)$.

Let S be the $\epsilon 2^k$ heaviest elements under X, breaking ties arbitrarily. Then for every $x \notin S$ we must have that $\Pr[X = x] \leq 2^{-k}$, or else every element in S would have weight greater than 2^{-k}, which would contradict the hypothesis. Thus, the set of elements that have weight more than 2^{-k} are hit with probability at most ϵ.

A block source is a source broken up into a sequence of blocks, with the property that each block has min-entropy even conditioned on previous blocks.

Definition 3 (Block sources). *A distribution $X = X_1, X_2, \cdots, X_{\mathsf{C}}$ is called a $(k_1, k_2, \ldots, k_{\mathsf{C}})$-block source if for all $i = 1, \ldots, \mathsf{C}$, we have that for all $x_1 \in X_1, \ldots, x_{i-1} \in X_{i-1}$, $H_\infty(X_i | X_1 = x_1, \ldots, X^{i-1} = x_{i-1}) \geq k_i$, i.e., each block has high min-entropy even conditioned on the previous blocks. If $k_1 = k_2 = \cdots = k_{\mathsf{C}} = k$, we say that X is a k-block source.*

The following lemma is useful to prove that a distribution is close to a block source.

Lemma 3. *Let $X = X_1, \ldots, X_t$ be t dependent random variables. For every $i = 1, 2, \ldots, t$, let X^i denote the concatenation of the first i variables. Suppose each X^i takes values in $\{0,1\}^{n_i}$ and for every $i = 1, 2, \ldots, t$, X^i is ϵ_i-close to k_i-light, with $\sum_i \epsilon_i < 1/10$. Then for every $\ell > 10 \log t$ we must have that X is $\sum_{i=1}^t \epsilon_i + t2^{-\ell}$-close to a block source, where each block X_i has min-entropy $k_i - n_{i-1} - 1 - \ell$.*

Proof. We will need to define the notion of a *submeasure*. Let $n = n_t$. Say that $M : \{0,1\}^n \to [0,1]$ is a submeasure on $\{0,1\}^n$ if $\sum_{m \in \{0,1\}^n} M(m) \leq 1$. Note that every probability measure is a submeasure. We abuse notation and let $M(x^i)$ denote the marginal measure induced on the first i coordinates.

We say a submeasure on $\{0,1\}^n$ is ϵ-close to k-light if

$$\sum_{m \in \{s : M(s) > 2^{-k}\}} M(m) \leq \epsilon.$$

As usual, for any event $A \subset \{0,1\}^n$, we denote $\Pr[M \in A] = \sum_{m \in A} M(m)$.

We now define the submeasures $M_{t+1} = X$, and for $i = t, t-1, t-2, \ldots, 1$,

$$M_i(m) = \begin{cases} 0 & M_{i+1}^i(m^i) > 2^{-k_i} \vee M_{i+1}^i(m^i) < 2^{-n_i - \ell} \\ M_{i+1}(m) & \text{otherwise} \end{cases}$$

Let $M = M_1$. Now note that for every $j < i$, M_i^j is ϵ_j-close to k_j-light, since we only made points lighter in the above process. Further, for all m and $i \leq j$, $M_i^j(m^j) \leq 2^{-k_j}$, since we reduced the weight of all m's that violated this to 0. We also have that for every m, i, $M^i(m^i) = 0$ or $M^i(m^i) \geq 2^{-n_i - \ell}$ by our construction.

Now define the sets $B_i = \{m \in \{0,1\}^n : M_i(m) \neq M_{i+1}(m)\}$. Set $B = \cup_i B_i$. Then note that $\Pr[X \in B] \leq \sum_{i=2}^t \Pr[M_{i+1} \in B_i]$. Each B_i, contains two types of points: points that were removed when moving from M_{i+1} to M_i because they were too heavy, and points that were removed because they were too light. We set $C_i = \{m : M_{i+1}(m^i) > 2^{-k_i}\}$, namely the "too heavy" points. We see that $\Pr[M_{i+1} \in C_i] \leq \epsilon_i$, since M_{i+1}^i is ϵ_i-close to k_i-light. Set $D_i = \{m : M_{i+1}(m^i) < 2^{-n_i - \ell}\}$, namely the "too light" points. We get $\Pr[M_{i+1} \in D_i] < 2^{-\ell}$ by the union bound. Using both these estimates, we get that $\Pr[X \in B] \le \sum_{i=1}^{t} \Pr[M_{i+1} \in B_i] \le \sum_{i=1}^{t} \Pr[M_{i+1} \in C_i] + \Pr[M_{i+1} \in D_i] \le \sum_i \epsilon_i + t2^{-\ell}$.

Now define the distribution $Z = X | X \notin B$. Then Z is $\sum_i \epsilon_i + t2^{-\ell}$-close to X. For every i and $z \in \mathsf{supp}(Z)$, we have that $\Pr[Z^i = z^i | Z^{i-1} = z^{i-1}] = \Pr[Z^i = z^i] / \Pr[Z^{i-1} = z^{i-1}] \le 2^{-k_i+1}/2^{-n_{i-1}-\ell}$ (since every point at most doubles in weight over M), which proves the lemma.

Theorem 4 (Block vs General Source Extractor [BRSW06]). *There exists constants c_1, c_2 such that for every n, k, with $k > \log^{10} n$ there exists a polynomial time computable function $\mathsf{BExt} : \{0,1\}^{\mathsf{C}n} \times \{0,1\}^n \to \{0,1\}^m$ with $\mathsf{C} = O(\frac{\log n}{\log k})$ s.t. , if $X = X^1, \cdots, X^{\mathsf{C}}$ is a k-block source and Y is an independent k-source*

$$\Pr_{x \leftarrow_R X}[|\mathsf{BExt}(x, Y) - U_m| < 2^{-k^{c_1}}] > 1 - 2^{-k^{c_1}},$$

where $m = c_2 k$ and U_m denotes the uniform distribution on m bit strings.

3 The Extractor

In this section we describe our construction. Our extractor uses as a key component a randomness condenser, constructed by Guruswami, Umans and Vadhan [GUV07], which is in turn based on recent constructions of good list decodable codes ([GR06,PV05]), though we give a self contained proof of everything we need in this section.

First let us give a high level description of our algorithm and analysis. Although it seems hard to build extractors for two independent sources, the problem seems considerably easier when one of the sources is a block source. Indeed, our new algorithm will be obtained by reducing to this case. We will give an algorithm that given two independent sources, can turn them into a single block source, with many blocks. Once we have this algorithm, we will simply use one additional source and our extractor from Theorem 5.

3.1 Converting Two Independent Sources Into A Block Source

Fix a finite field $\mathbb{F}$. The following algorithm is from [GUV07].

Algorithm 1 ($\mathsf{Cond}(f, y)$)

Input: $f \in \mathbb{F}^{t+1}$, $y \in \mathbb{F}$ and an integer r.
Output: $z \in \mathbb{F}^r$.

Sub-Routines and Parameters:
Let $g \in \mathbb{F}[X]$ be an irreducible polynomial of degree $t+1$. Set $h = |\mathbb{F}|^{0.8\alpha}$ for some parameter α.

1. We interpret f as a degree t univariate polynomial with coefficients in $\mathbb{F}$.
2. For every $i = 0, 1, \ldots, m-1$, let $f_i \in \mathbb{F}[x]$ be the polynomial $f^{h^i} \mod g$.
3. Output $f_0(y), f_1(y), \ldots, f_{r-1}(y)$.

Guruswami et al. were interested in building seeded condensers, so they used the above algorithm with y sampled uniformly at random. Below, we show that the algorithm above is useful even when y is a high min-entropy source. We can prove the following lemma, which is a slight generalization of a lemma in [GUV07]:

Lemma 4. *Suppose F is a distribution on $\mathbb{F}^{t+1}$ with min-entropy k and Y is an independent distribution on $\mathbb{F}$ with min-entropy rate α and*

– $rt < \epsilon |\mathbb{F}|^{0.1\alpha}$
– $k > \log(2/\epsilon) + (0.8\alpha r)\log|\mathbb{F}|$.

Then $\mathsf{Cond}(F, Y)$ *is* ϵ*-close to* $.7\alpha r \log |\mathbb{F}|$*-light, and hence it is* ϵ*-close to having min-entropy rate* 0.7α*.*

Remark 1. In order to avoid using too many variables, we have opted to use constants like 0.1 and 0.7 in the proof. We note that we can easily replace the constants $0.7, 0.8$ with constants that are arbitrarily close to 1, at the price of making 0.1 closer to 0.

Proof (Lemma 4). We will repeatedly use the basic fact that any non-zero polynomial of degree d can have at most d roots.

By Fact 4, it suffices to prove the lemma when F and Y are flat sources.

We will prove that the output is close to having high min-entropy via Lemma 2. To do this, we need to show that for every set $S \subset \mathbb{F}^r$ of size $\epsilon|\mathbb{F}|^{0.7\alpha r}$, $\Pr[\mathsf{Cond}(F, Y) \in S] < \epsilon$. Fix a set S.

Let $Q(Z_1, \ldots, Z_r) \in \mathbb{F}[Z_1, \ldots, Z_r]$ be a non-zero r variate polynomial whose degree is at most $h-1$ in each variable, such that $Q(s) = 0$ for every $s \in S$. Such a polynomial must exist since the parameters have been set up to guarantee $h^r = |\mathbb{F}|^{0.8\alpha r} > |S| = \epsilon|\mathbb{F}|^{0.7r\alpha}$.

Now call $f \in \mathsf{supp}(F)$ bad for S if

$$\Pr_{y \leftarrow_{\mathrm{R}} Y}[\mathsf{Cond}(f, y) \in S] \geq \epsilon/2$$

We will bound the number of bad f's. Fix any such bad f. Then consider the univariate polynomial

$$R(X) = Q(f_0(X), f_1(X), \ldots, f_{r-1}(X)) \in \mathbb{F}[X]$$

This polynomial has degree at most $tr(h-1)$. But $tr(h-1) < \epsilon|\mathbb{F}|^{0.1\alpha}|\mathbb{F}|^{0.8\alpha} < \epsilon|\mathbb{F}|^{\alpha}/2 = (\epsilon/2)|\mathsf{supp}(Y)|$, thus this polynomial must be the zero polynomial. In particular, this means that $R(X) = 0 \mod g(X)$. This in turn implies that f must be a root of the polynomial

$$Q'(Z) = Q(Z, Z^h, Z^{h^2}, \ldots, Z^{h^{r-1}}) \in (\mathbb{F}[X]/g(X))[Z]$$

which is a univariate polynomial over the extension field $\mathbb{F}[X]/g(X)$, since $Q'(f(X)) = R(X) \mod g(X)$ by our choice of $f_0, \ldots, f_{r-1}$.

Recall that Q had degree at most $h-1$ in each variable. This means that Q' has degree at most $h^r - 1$ and is non-zero, since no two monomials can clash when making the substitution Z^i for Z_i in Q. The number of bad f's can be at most $h^r - 1 < |\mathbb{F}|^{0.8\alpha r}$, since every bad f is a root of this low degree non-zero polynomial. This implies that $\Pr[F \text{ is bad}] < |\mathbb{F}|^{0.8\alpha r}/2^k < \epsilon/2$, since the constraint on k implies that $2^k > |\mathbb{F}|^{0.8\alpha r}2/\epsilon$.

Hence $\Pr[\mathsf{Cond}(F, Y) \in S] \leq \Pr[F \text{ is bad}] + \Pr[\mathsf{Cond}(F, Y) \in S | F \text{ is not bad}] < \epsilon/2 + \epsilon/2 = \epsilon$.

Note that a seeded condenser corresponds to the special case of $\alpha = 1$ in the above lemma. When α is small, it seems like the lemma doesn't say anything useful, since the min-entropy rate of the output is bounded above by α. But note that the lemma works for a very wide range of r's. The above function is more than a condenser, it *spreads* the entropy out across the output. Specifically, if we look at the first r' symbols in the output, they must also have min-entropy rate close to 0.7α. We can use this to construct a block source with geometrically increasing block lengths, as in the following lemma:

Lemma 5. *Let* $\mathsf{Cond}, F, Y, \alpha, r, t, \epsilon$ *be as in Algorithm 6 and Lemma 4. Let* $r_1, r_2, \ldots, r_{\mathsf{C}} = r$ *be positive integers. For* $i = 1, 2, \ldots, \mathsf{C}$*, set* Z^i *to be the first* r_i *field elements in the output of* $\mathsf{Cond}(F, Y)$*. Then let* $Z_1, \ldots, Z_{\mathsf{C}}$ *be such that* $Z^i = Z_1, \ldots, Z_i$ *for every* i*. Then for every* $\ell > 10 \log \mathsf{C}$ *we have that* $Z_1, Z_2, \ldots, Z_{\mathsf{C}}$ *is* $\mathsf{C}(\epsilon + 2^{-\ell})$*-close to a block source with entropy* $(0.7\alpha r_i - r_{i-1})\log(|\mathbb{F}|) - 1 - \ell$ *in each block.*

Proof. We will apply Lemma 3.

Note that for each i, Z^i is simply the output of the condenser upto the first r_i elements. Since $r_i \leq r$, r_i satisfies the constraints of Lemma 4, so Z^i is ϵ-close to $0.7\alpha|Z^i|$-light.

We set parameters to get the following theorem:

Theorem 5. *There exists a polynomial time computable function* $\mathsf{BlockConvert} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^{m_1} \times \{0,1\}^{m_2} \times \cdots \times \{0,1\}^{m_\mathsf{C}}$*, such that for every min-entropy* k_1 *source* X *over* $\{0,1\}^{n_1}$ *and every min-entropy* k_2 *source* Y *over* $\{0,1\}^{n_2}$ *satisfying*

- $\mathsf{C}(\log \frac{10n_2}{k_2}) + 2\log(n_1) < 0.095k_2$
- $\sqrt{k_1} > k_2(10n_2/k_2)^\mathsf{C}$,

$\mathsf{BlockConvert}(X,Y)$ *is* $\mathsf{C}(2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}})$*-close to a block source with* $\sum_i m_i \le (10n_2/k_2)^\mathsf{C}\sqrt{k_1}$ *and min-entropy* $2\sqrt{k_1}$ *in each block.*

Proof. We show how to set parameters and apply Lemma 5.

Set $\mathbb{F}$ to be the finite field of size 2^{n_2}. Set $t = n_1/n_2, \epsilon = 2^{-0.05k_2}$ and $k = k_1$. Set $\alpha = k_2/n_2$.

Set $r_i = (10n_2/k_2)^i\sqrt{k_1}$, so $\sum_i m_i = r = k_1^{1/2}(10n_2/k_2)^\mathsf{C}$.

Using the first assumption,

$$rt = \sqrt{k_1}\left(\frac{10n_2}{k_2}\right)^\mathsf{C}\frac{n_1}{n_2} \le n_1^2\left(\frac{10n_2}{k_2}\right)^\mathsf{C} < 2^{0.095k_2} = 2^{-0.05k_2}2^{0.1k_2} = \epsilon|\mathbb{F}|^{0.1\alpha}$$

to satisfy the first constraint of Lemma 4.

We have that

$$k_1 = k > 1 + 0.05k_2 + 0.8\cdot 10^\mathsf{C} k_2\sqrt{k_1}(n_2/k_2)^\mathsf{C} = \log(2/\epsilon) + (0.8r\alpha)\log(|\mathbb{F}|)$$

to satisfy the second constraint of Lemma 4.

Set $\ell = k_1^{0.1}$. Note that the second constraint implies that $\mathsf{C} < \log k_1$.

Then let us use the algorithm Cond as promised by Lemma 5 with the above settings. We get that the final output is $\mathsf{C}(\epsilon + 2^{-\ell+1}) \le \mathsf{C}(2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}})$- close to a block source with min-entropy $(0.7\alpha r_i - r_{i-1})\log(|\mathbb{F}|) - 1 - 2\ell$ in each block. We can lower bound this as follows:

$$\begin{aligned}
&(0.7\alpha r_i - r_{i-1})\log(|\mathbb{F}|) - 1 - 2\ell \\
&= \left(0.7\frac{k_2}{n_2}\left(\frac{10n_2}{k_2}\right)^i\sqrt{k_1} - \left(\frac{10n_2}{k_2}\right)^{i-1}\sqrt{k_1}\right)n_2 - 1 - 2k_1^{0.1} \\
&= (0.7\cdot 10 - 1)\left(\frac{10n_2}{k_2}\right)^{i-1} n_2\sqrt{k_1} - 1 - 2k_1^{0.1} \\
&= 6\left(\frac{10n_2}{k_2}\right)^{i-1} n_2\sqrt{k_1} - (1 + 2k_1^{0.1}) \\
&\ge 2\sqrt{k_1}
\end{aligned}$$

3.2 Putting it all together

All that remains is to put together the various components to get our extractor.

Algorithm 2 ($\mathsf{IExt}(a, b, c)$)

Input: $a \in \{0,1\}^{n_1}, b \in \{0,1\}^{n_2}, c \in \{0,1\}^{n_3}$.
Output: $z \in \{0,1\}^m$ for a parameter m that we will set.

Sub-Routines and Parameters:
Let $\mathsf{BlockConvert}$ be the algorithm promised by Theorem 7, set up to operate on two sources with entropy k_1, k_2 and lengths n_1, n_2 respectively.
Let BExt be the algorithm promised by Theorem 5, set up to extract from a block source with C blocks of length $(10n_2/k_2)^{\mathsf{C}}\sqrt{k_1}$, each with entropy $\sqrt{k_1}$ conditioned on previous blocks, and an independent source with length n_3 and min-entropy k_3.

1. Run $\mathsf{BlockConvert}(a, b)$ to get the blocks $x = x_1, x_2, \ldots, x_{\mathsf{C}}$.
2. Output $\mathsf{BExt}(x, c)$.

We can now prove the main theorem.

Proof (Theorem 1). Let t be a constant so that BExt requires $\mathsf{C} = t\log(n_1 + n_3)/\log(k_1)$ blocks to extract bits from an $(n_3, k_3 = \sqrt{k_1})$ source and an independent block source with blocks of length n_1, each with entropy $\sqrt{k_1}$ conditioned on previous blocks. The error of this extractor is promised to be $2^{-k_1^{\Omega(1)}}$.

We check each of the constraints needed for $\mathsf{BlockConvert}$ to succeed.

First we have that

$$
\begin{aligned}
&\mathsf{C}\left(\log\frac{10n_2}{k_2}\right) + \log n_1 \\
&< \mathsf{C}10\log n_2 + \log n_1 \\
&\le 10t\frac{\log(n_1+n_3)}{\log k_1}\log n_2 + \log n_1 \\
&\le (10t/d)\log k_1 + \log n_1 && \text{by the first assumption} \\
&< 0.095 d\log n_1 && \text{for } d \text{ large enough} \\
&< 0.095 k_2 && \text{by the second assumption}
\end{aligned}
$$

For the next constraint,

$$
\begin{aligned}
&\log(k_2(10n_2/k_2)^{\mathsf{C}}) \\
&= \mathsf{C}\log(10n_2/k_2) + \log k_2 \\
&\le t\frac{\log(n_1+n_3)}{\log k_1}(\log(n_2) + \log 10) + \log n_2 \\
&< 3(t/d)\log k_1 && \text{by the first assumption} \\
&< (1/2)\log k_1 && \text{for } d \text{ large enough}
\end{aligned}
$$

We are not yet done, since the algorithm above will only output $m_1 = \sqrt{k_1} - o(\sqrt{k_1})$ bits. However, we do have that:

$$\Pr_{x_1 \leftarrow_{\mathrm{R}} X_1}[|\mathsf{IExt}(x_1, Y, Z) - U_{m_1}| > 2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}}] < 2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}}$$

since BExt is strong.

Thus we have that $|X, \mathsf{IExt}(X, Y, Z) - X, U_{m_1}| < 2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}}$, which implies that if Ext is any strong seeded extractor set up to extract from a min-entropy k_1 source with seed length m_2, $\mathsf{Ext}(X, U_{m_1})$ is $2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}}$ close to $\mathsf{Ext}(X, \mathsf{IExt}(X, Y, Z))$. This is our final extractor.

3.3 Extension to Block Sources

We now sketch the proofs of Theorems Theorem 2 and Theorem 3.

For Theorem 2, our block source will be (b, c), and our extractor will be $\mathsf{IExt}(a, b, c)$. We will show that $\mathsf{BlockConvert}$ is strong in the sense that with high probability, even conditioned on b, $\mathsf{BlockConvert}(X, b)$ will be a block source. The proof then proceeds as before.

The following lemma shows that any condenser with good parameters is also strong, with slightly weaker parameters.

Lemma 6. *Let $\mathcal{X}$ denote a collection of sources. Suppose the function C is a condenser in that for independent $X \in \mathcal{X}$ and Y with $H_\infty(Y) \geq \ell$, $C(X, Y)$ is ϵ-close to k-light. Then for any such X, Y, when y is chosen from Y,*

$$\Pr_y[C(X, y) \text{ is } \delta\text{-close to } (k - \ell)\text{-light}] \geq 1 - \epsilon/\delta.$$

Proof. Fix any such X and Y. Let $S = \{z | \Pr[C(X, Y) = z] > 2^{-k}\}$ denote the set of heavy elements. Note that for any y in the support of Y and $z \notin S$, $\Pr[C(X, y) = z] \leq 2^{\ell - k}$. Now let $p_y = \Pr[C(X, y) \in S]$. Then $E[p_y] \leq \epsilon$, so by Markov $\Pr[p_y \geq \delta] \leq \epsilon/\delta$, which gives the lemma.

We will use this lemma with the condenser Cond and with $\delta = \sqrt{\epsilon}$. We then modify Lemma 5 so that with high probability over the choice of y, $\mathsf{Cond}(F, y)$ is a block source. This immediately yields a strong version of Theorem 7.

Theorem 6. *There exists a polynomial time computable function* $\mathsf{BlockConvert} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^{m_1} \times \{0,1\}^{m_2} \times \cdots \times \{0,1\}^{m_\mathsf{C}}$*, such that for every min-entropy k_1 source X over $\{0,1\}^{n_1}$ and every min-entropy k_2 source Y over $\{0,1\}^{n_2}$ satisfying*

- $\mathsf{C}(\log \frac{10 n_2}{k_2}) + 2\log(n_1) < 0.095 k_2$
- $\sqrt{k_1} > k_2 (10 n_2/k_2)^\mathsf{C}$,

the following holds. When y is chosen according to Y, with probability $1 - \mathsf{C}(2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}})$, $\mathsf{BlockConvert}(X, y)$ *is $\mathsf{C}(2^{-\Omega(k_2)} + 2^{-k_1^{\Omega(1)}})$-close to a block source with $\sum_i m_i \leq (10 n_2/k_2)^\mathsf{C} \sqrt{k_1}$ and min-entropy $2\sqrt{k_1}$ in each block.*

Now, when we analyze $\mathsf{IExt}(a, b, c)$ where (b, c) is a block source, we argue that with high probability over the choice of b, we are in the same situation as before, and our proof continues in the same manner.

For Theorem 3, our block source will be (a, b), and our extractor will be $\mathsf{IExt}(a, b, c)$. In the analysis of $\mathsf{Cond}(f, y)$, we analyzed a bad f by counting the number of y that cause $\mathsf{Cond}(f, y) \in S$. The key observation is that this analysis remains unchanged if we choose y from a set Y that depends on f. This is easily verified by looking at the proof of Lemma 4. Hence (f, y) can be from a block source.

Acknowledgements

We would like to thank Salil Vadhan and Chris Umans for useful discussions. In particular, Salil showed us how to get Theorem 3. Thanks to the referees for several useful comments.

References

[BIW04] Boaz Barak, R. Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 384–393, 2004.

[BKS+05] Boaz Barak, Guy Kindler, Ronen Shaltiel, Benny Sudakov, and Avi Wigderson. Simulating independence: New constructions of condensers, Ramsey graphs, dispersers, and extractors. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 1–10, 2005.

[BRSW06] Boaz Barak, Anup Rao, Ronen Shaltiel, and Avi Wigderson. 2 source dispersers for $n^{o(1)}$ entropy and Ramsey graphs beating the Frankl-Wilson construction. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.

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

[CRVW02] M. Capalbo, Omer Reingold, Salil Vadhan, and Avi Wigderson. Randomness conductors and constant-degree lossless expanders. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 659–668, 2002.

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

[Gur04] V. Guruswami. Better extractors for better codes? In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 436–444, 2004.

[GR06] Venkatesan Guruswami and Atri Rudra. Explicit capacity-achieving list-decodable codes. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.

[GUV07] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. Unbalanced expanders and randomness extractors from Parvaresh-Vardy codes. In *Proceedings of the 22nd Annual IEEE Conference on Computational Complexity*, 2007.

[KLRZ08] Yael Kalai, Xin Li, Anup Rao, and David Zuckerman. Network extractor protocols. Unpublished manuscript, 2008.

[KRVZ06] Jesse Kamp, Anup Rao, Salil Vadhan, and David Zuckerman. Deterministic extractors for small space sources. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.

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

[Lu04] Chi-Jen Lu. Encryption against storage-bounded adversaries from on-line strong extractors. *J. Cryptology*, 17(1):27–42, 2004.

[MNSW98] Peter Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. On data structures and asymmetric communication complexity. *Journal of Computer and System Sciences*, 57:37–49, 1 1998.

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

[PV05] Farzad Parvaresh and Alexander 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.

[Rao06] Anup Rao. Extractors for a constant number of polynomially small min-entropy independent sources. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.

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

[SV86] Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. *Journal of Computer and System Sciences*, 33:75–87, 1986.

[Sha06] Ronen Shaltiel. How to get more mileage from randomness extractors. In *Proceedings of the 21th Annual IEEE Conference on Computational Complexity*, pages 49–60, 2006.

[TZ04] Amnon Ta-Shma and David Zuckerman. Extractor codes. *IEEE Transactions on Information Theory*, 50, 2004.

[Vad04] Salil P. Vadhan. Constructing locally computable extractors and cryptosystems in the bounded-storage model. *J. Cryptology*, 17(1):43–77, 2004.

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

[Zuc96] David Zuckerman. Simulating BPP using a general weak random source. *Algorithmica*, 16:367–391, 1996.

[Zuc97] David Zuckerman. Randomness-optimal oblivious sampling. *Random Structures and Algorithms*, 11:345–367, 1997.