Electronic Colloquium on Computational Complexity, Report No. 106 (2005)

Extractors for a Constant Number of Polynomial Min-Entropy Independent Sources

Anup Rao
University of Texas at Austin
arao@cs.utexas.edu

September 26, 2005

Abstract

We consider the problem of bit extraction from independent sources. We construct an extractor that can extract from a constant number of independent sources of length n, each of which have min-entropy n^γ for an arbitrarily small constant $\gamma > 0$. Our constructions are different from recent extractor constructions [BIW04, BKS$^+$05, Raz05, Bou05] for this problem in the sense that they do not rely on any results from additive number theory. They are obtained by composing previous constructions of strong seeded extractors in simple ways.

ISSN 1433-8092

1 Introduction

The use of randomness is widespread in computer science. Many of the best performing algorithms and protocols in many different areas of computer science are randomized. To guarantee their performance these algorithms usually rely on a perfect source of uncorrelated uniformly random bits, yet such a source may not be easy to obtain. We might instead have access to an imperfect random source where the bits are correlated and not uniformly random.

This motivates the study of objects called *extractors*. Loosely speaking, an extractor is an explicit efficiently computable function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ that takes as input bits from an imperfect random source and produces bits that are close to uniformly random (the distance of the output distribution from the uniform distribution is called the error of the extractor). If we had access to such a function, we could use it to extract truly random bits from an imperfect random source. We would then use the extracted bits in our application. Thus we could achieve performance guarantees even with imperfect sources of randomness. Building on work by Zuckerman [Zuc90, Zuc96], extractors were first considered and studied by Nisan and Zuckerman in [NZ93] with exactly this goal in mind. A long sequence of works further developed extractor constructions and applications. Extractors are now known to have applications in a wide range of problems and are interesting objects in their own right. For surveys of the applications and constructions we refer the interested reader to [Nis96, NT99, Sha02].

1.1 Modeling the source

To formalize the problem of randomness extraction, we must decide on a model for the types of imperfect sources that the extractors can handle. If we intend to extract m random bits, information theoretic considerations show that the imperfect source must contain at least m bits of entropy. The goal is to construct extractors which output the most number of random bits for a source with given entropy, have small error and work for very general models. The most general model that has been considered to date is what we will call a *weak source* [Zuc96]. The only constraint on a weak source that supplies n total bits is that the probability of getting any particular string from the source is at most 2^{-k}, where k is called the *min-entropy* of the source. Such a source is called an (n, k)-source. Unfortunately it can be shown that there is no deterministic extractor that can extract from general weak sources.

One way to get around this problem is to restrict the source so that it consists of a sample from a weak source and an additional much shorter independent *seed* of truly uniformly random bits. An extractor for such sources is called a *seeded extractor*. For any $n, k \in \mathbb{N}$ we now know how to construct extractors that can extract a constant fraction of k bits which are almost uniformly random using a very short (only a constant multiple of $\log n$) length seed from any (n, k)-source [LRVW03]. This is sufficient to simulate any algorithm that relies on truly uniformly random bits for its efficiency with the aid of a weak source in polynomial time. The assumption that we have access to a truly uniformly random seed is restrictive. These extractors are not appropriate for many applications where randomness is needed for more than improving efficiency. For instance, many cryptographic applications need to pick a key uniformly at random. Such an operation cannot be simulated using a seeded extractor and a weak source without access to an independent uniformly random seed.

Several other models for sources have been considered [vN51, Blu84, Vaz85, CFG$^+$85, CG88, CW89, TV00, MU02, KZ03, GRS04, GR05]. In this paper, we assume we have access to a few independent sources, each of which have high enough min-entropy. The probabilistic method shows that extractors for this model exist even when given access to $(n, \mathsf{polylog}(n))$-sources. The challenge, then, is to construct an efficiently computable function that is a good extractor for such sources.

1.2 History and Previous Results

The problem of building extractors for independent weak sources is closely related to the Bipartite Ramsey Graph Problem. In the Bipartite Ramsey Graph Problem, the goal is to find an explicit 2-coloring of the complete bipartite graph with N vertices in each bipartition so that there is no large monochromatic clique. Every extractor for 2 independent sources immediately gives a solution to the Bipartite Ramsey Graph Problem.

The problem of extracting from several independent sources was first considered by Santha and Vazirani [SV86]. Their extractors can extract from $O(\log n)$ independent $(n, \delta n)$-sources, but they placed additional restrictions on the sources that we will not discuss here. Chor and Goldreich [CG88] demonstrated extractors for 2 independent $(n, (1/2 + \alpha)n)$-sources, for all constant $\alpha \in (0, 1/2]$.

In recent work, Barak, Impagliazzo and Wigderson [BIW04] showed how to extract from a constant number of independent $(n, \delta n)$-sources, where δ (the *min-entropy rate* of the source) is allowed to be any arbitrarily small constant. The number of sources used depends on δ. Subsequently Barak et. al. [BKS$^+$05] showed how to extract a constant number of bits with constant error from 3 $(n, \delta n)$-sources, where δ is an arbitrarily small constant. In this work they also present 2-source *dispersers* (a disperser is an object similar to but somewhat weaker than an extractor) that output a constant number of bits with constant error and work for min-entropy rate δ where δ is an arbitrarily small constant.

Raz [Raz05] gave an extractor for 2 independent sources where one source needs to have min-entropy rate greater than and bounded away from 1/2 and the other source may have polylogarithmically small min-entropy. In this case his extractor can extract a linear fraction of the min-entropy with exponentially small error. He also gave an extractor for 3 independent sources where one source must have constant min-entropy rate and the other two need polylogarithmic min-entropy. In this case his extractor can extract a constant number of bits with constant error.

Bourgain [Bou05] gave another extractor construction for 2 independent sources. His extractor can extract from 2 $(n, (1/2-\alpha_0)n)$-sources, where α_0 is some small universal constant. This is the first extractor to break the 1/2 min-entropy rate barrier for 2 sources. His extractor outputs a linear fraction of the min-entropy, with exponentially small error.

All four of these recent constructions were made possible by new breakthroughs on the sum-product estimate in additive number theory [BKT04, Kon03]. A common feature of [Raz05, BKS$^+$05] is that they reduce the general problem of extracting from independent sources to the problem of extracting from independent sources that come from a much more restricted class, called *somewhere-random* sources. They then build extractors for these sources. A key step in our constructions is building much better extractors for *somewhere-random* sources.

1.3 Overview of New Results and Techniques

1.3.1 New Results

- We construct a polynomial time computable extractor which extracts $k - o(k)$ random bits from 2 independent blockwise sources (blockwise sources were first defined by Chor and Goldreich [CG88]). Each source is required to have at most $O(\frac{\log n}{\log k})$ blocks of length n, where each block has min-entropy $k(n) > \log^4 n$ conditioned on the previous blocks. The error for the extractor is $1/n^{\Omega(1)}$.

- As a special case of the previous construction, we obtain a polynomial time computable extractor which extracts k random bits from $O(\frac{\log n}{\log k})$ independent (n, k)-sources with error $1/n^c$ for any $k(n) > \log^4 n$ and any constant $c > 1$. An interesting setting of parameters is when $k = n^\gamma$ for some $0 < \gamma < 1$. In this case we get an extractor for a constant number of sources which extracts all the min-entropy with polynomially small error.

- We construct a polynomial time computable 2-source disperser which outputs a linear number of bits with exponentially small error when the min-entropy rate of the source is an arbitrarily small constant.

- We construct a polynomial time computable extractor for 3 sources which extracts a linear number of bits with exponentially small error when one source has min-entropy rate that is an arbitrarily small constant and the other two may have min-entropy that is polylogarithmically small.

Our results are the first to extract from a constant number of sources which have polynomially small min-entropy. In fact we are not even aware of previous constructions of explicit dispersers (or equivalently explicit Ramsey Graphs) of this type for such low min-entropy. The first three of the results above have the additional feature that they do not rely on the sum-product estimate. We compose existing constructions of strong seeded extractors in simple ways to get our extractors/dispersers. For the case of constant min-entropy rate sources, our results are obtained by modifying the constructions of [BKS$^+$05, Raz05].

1.3.2 Techniques

Many extractor constructions in the past have been based on the paradigm of iterative condensing [RSW00, TSUZ01, CRVW02, LRVW03, BIW04, Zuc05]. The idea is to start with some distribution that has low min-entropy and apply a function (called a *condenser*) whose output has a better min-entropy rate. Repeating this process, we eventually obtain a distribution which has very high min-entropy rate. Then we can apply some other extractor construction which works for such a high min-entropy rate to obtain random bits. The construction in this paper can also be viewed as an example of this paradigm, with a slight twist. Instead of improving the min-entropy rate of the distributions we are working with, we impose a certain structure on the distributions and find some measure of the quality of this structure. Then we iteratively improve the quality of the distribution until it is so good that extracting randomness becomes easy.

We make progress by considering a more restricted model for a source, which we call an *elementary somewhere random source*, ES-source for short. ES-sources are slightly more restricted than somewhere-random sources, first considered by Ta-Shma [TS96]. A source is a $(t \times r)$ ES-source if it can be broken up into a collection of t blocks, each of length r, such that at least one block contains uniformly random bits. A somewhere-random source is merely a convex combination of ES-sources. In fact all of our constructions for ES-sources work even for somewhere-random sources, but it is less cumbersome and no loss of generality to make the arguments for ES-sources.

It is easy to use a strong seeded extractor to convert any general weak source into an ES-source. Given a sample from the weak source, we simply evaluate the extractor on the sample with all possible seeds, getting one block of the output for each fixed seed. For any fixed weak source, the strong extractor property guarantees that most seeds will give a distribution that is statistically close to uniform. As long as the seed length required by the extractor is $O(\log n)$, we get a polynomial time algorithm to convert any weak source to a distribution that is statistically close to an ES-source with $\mathsf{poly}(n)$ blocks.

It turns out that it is easy to extract from independent ES-sources when each source has very few blocks relative to the length of each of the blocks. In the extreme case, when an ES-source has just one block, it is a uniformly random string. The measure of quality that we will be concerned with is the ratio of the length of each block in the source to the number of blocks.

We proceed by exploiting the structure of ES-sources to build *condensers* for such sources in the following sense: starting from a collection of independent ES-sources, we turn the distribution into one that is essentially the distribution of independent ES-sources with fewer blocks. The condensers are also built using previous constructions of strong seeded extractors. To condense a particular ES-source, we will use the other ES-sources to generate a short list of candidate seeds. Then we apply a strong extractor to the ES-source being condensed with all the candidate seeds. This condensing process introduces dependencies between the output ES-sources. Still, we will argue that the output distribution is statistically close to a convex combination of independent ES-sources which are suitable for further condensing. Iterating this condensing process, we eventually reduce the number of blocks to just 1, at which point we can just output the blocks to get bits which are statistically close to uniformly distributed.

2 Preliminaries

2.1 Notation

Throughout this paper, we will use capital letters to denote distributions and sets. When it isn't ambiguous we will sometimes use the same capital letter to denote a distribution and its support. We will usually use the same small letter to denote an instantiation of the capital letter, for e.g. for a set X, we would use x to denote an element in X. If R is a random variable, we will write $r \in R$ when we really mean $r \in \mathsf{supp}(R)$.

Given a distribution X over $\{0,1\}^n$ and a set $S \subseteq [n]$, we will use X_S to denote the restriction of X onto the indices in S.

We use the convention that $N = 2^n$, $M = 2^m$ and $K = 2^k$.

All logarithms are meant to be base 2.

We will use U_m to denote the uniformly random distribution on $\{0,1\}^m$.

We will use the symbol $\circ$ to denote concatenation.

We will construct and compose several explicit extractors and condensers in this paper. We will usually use long names to make clear what kinds of sources the functions are meant to manipulate. When a function f can be applied to a very restricted family of sources (the restrictions may not be apparent from the naming), we will usually name it $\overline{f}$ to indicate that its use is restricted.

2.2 Min-Entropy and Special Sources

We will be concerned with the treatment of various kinds of distributions that are *nice* in that they contain a lot of usable randomness. Here we discuss some ways to measure this niceness:

Definition 2.1. The *min-entropy* of a random variable R is defined to be:

$$H_\infty(R) = -\log(\max_{x \in R}(R(x)))$$

Definition 2.2. The *min-entropy rate* of a distribution R on $\{0,1\}^n$ is $H_\infty(R)/n$.

Definition 2.3. An (n,k)-source denotes some random variable X over $\{0,1\}^n$ with $H_\infty(X) \geq k$.

Definition 2.4. A distribution $X^1 \circ X^2 \circ \cdots \circ X^u$ is called a $(k_1, k_2, \ldots, k_u)$-blockwise source if for all $i = 1, \ldots, u$, 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.

Proposition 2.5. *Let $X^1, \ldots, X^u$ be independent sources with $H_\infty(X^i) = k_i$ for $i = 1, \ldots, u$.Then $X^1 \circ \cdots \circ X^u$ is a $(k_1, k_2, \ldots, k_u)$-blockwise source.*

In some situations we will be interested in the min-entropy of a random variable when it is conditioned on *typical* instantiations. We will need the following proposition:

Proposition 2.6. *Let X be a random variable with $H_\infty(X) = k$. Let A be any event in the same probability space. Then*

$$H_\infty(X|A) < k' \Rightarrow \Pr[A] < 2^{k'-k}$$

Definition 2.7. A $(t \times r)$ *elementary somewhere-random* source (*ES-source* for short) denotes some random variable X over t blocks of $\{0,1\}^r$ s.t. X is distributed uniformly randomly over one of the blocks. Every other block may depend on the random block in arbitrary ways.

ES-sources are a special case of *somewhere-random* sources, introduced by Ta-Shma [TS96].

Definition 2.8. A $(t \times r)$ *somewhere-random* source denotes some random variable X over t blocks of $\{0,1\}^r$ s.t. X is a convex combination of *elementary somewhere-random* sources.

For ease of discussion, throughout this paper we will consider ES-sources, even though all our constructions for extracting from ES-sources succeed even when the sources are somewhere-random sources.

Definition 2.9. We will say that a $(t \times r)$ source (i.e., a distribution over t blocks of $\{0,1\}^r$) X has *elementary somewhere-min-entropy* k, if X has min-entropy k in one of its blocks. We will say that a source has *somewhere min-entropy* k if it is a convex combination of sources which have *elementary somewhere-min-entropy* k.

2.3 Statistical Distance

Sometimes the distributions we get are not exactly the distributions we want, but they may be *close* enough. The measure of *closeness* we will use is this one:

Definition 2.10. Let D and F be two distributions on a set S. Their *statistical distance* is

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

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

Proposition 2.11. *Let D and F be any two distributions over a set S s.t. $\|D - F\| \leq \epsilon$. Let g be any function on S. Then $\|g(D) - g(F)\| \leq \epsilon$.*

In a few of our proofs we will need to change a distribution that we are working with to a statistically close distribution while maintaining its independence from various other distributions.

Proposition 2.12. *If $X^1, \ldots, X^l$ are independent random variables with $\| X^i - Y^i \| < \epsilon$, then $X^1 \circ X^2 \circ \cdots \circ X^l$ is $l\epsilon$-close to $Y^1 \circ Y^2 \circ \cdots \circ Y^l$ where the random variables Y^i are independent of each other.*

2.4 Convex Combinations

Definition 2.13. Let $\mathcal{P}$ be a property of sources. Let X be some random variable over some universe. We will say that X is a convex combination of sources with property $\mathcal{P}$ if there exists some random variable I over an arbitrary universe s.t. for all $i \in \mathsf{supp}(I)$, $X|I = i$ has property $\mathcal{P}$.

A key observation that is essential to our constructions is that random variables that are convex combinations of sources with some good property are usually good themselves. This is captured in the following easy propositions:

Proposition 2.14. *Let X, Y be random variables s.t. X is a convex combination of sources which are ϵ close to Y. Then X is ϵ close to Y.*

Proposition 2.15. *Let X, I be random variables s.t. X is a convex combination of random variables $\{X_i\}_{i \in I}$. Let f be some function s.t. for all $i \in I$, $f(X_i)$ is a convex combination of sources that have some property $\mathcal{P}$. Then $f(X)$ is a convex combination of sources that have property P.*

2.5 Extractors and Dispersers

Definition 2.16. A function $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ is a strong (n, k, ϵ) seeded extractor if for any (n, k) source X and for Y chosen uniformly at random from $\{0,1\}^t$, we have

$$\| Y \circ \mathsf{Ext}(X, Y) - Y \circ U_m \| < \epsilon$$

where U_m is independent of Y.

Definition 2.17. A function $\mathsf{IExt} : (\{0,1\}^n)^u \to \{0,1\}^m$ is an (n, k, ϵ) extractor for u independent sources if for any independent (n, k) sources $X^1, \dots, X^u$ we have

$$\| \mathsf{IExt}(X^1, \dots, X^u) - U_m \| < \epsilon$$

Definition 2.18. A function $\mathsf{IExt} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ is an (n, k, ϵ) 2-source extractor if for any independent (n, k) sources X, Y we have

$$\| \mathsf{IExt}(X, Y) - U_m \| < \epsilon$$

We will say that IExt is a *strong* 2-source extractor if

$$\| Y \circ \mathsf{IExt}(X, Y) - Y \circ U_m \| < \epsilon$$

where U_m is independent of Y.

Definition 2.19. A function $\mathsf{IDisp} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ is an (n, k, ϵ) 2-source disperser if for all sets $A, B \subseteq \{0,1\}^n$, with $|A|, |B| \geq 2^k$, $|\mathsf{IDisp}(A, B)| \geq (1 - \epsilon)2^m$.

In our constructions we will need the notion of a *strong* disperser.

Definition 2.20. A function $\mathsf{IDisp} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ is a *strong* (n, k, ϵ, μ) 2-source disperser if for all independent (n, k) sources $A, B \subseteq \{0,1\}^n$, $\Pr_{a \leftarrow_R A}[|\mathsf{IDisp}(a, B)| \geq (1 - \epsilon)2^m] \geq \mu$.

If we omit the parameter μ in the above definition, then it is assumed to be some positive quantity. Specifically:

Definition 2.21. We will say that a function $\mathsf{IDisp} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ is a *strong* (n, k, ϵ) 2-source disperser if for all sets $A, B \subseteq \{0,1\}^n$, with $|A|, |B| \geq 2^k$, there exists $a \in A$ s.t. $|\mathsf{IDisp}(a, B)| \geq (1 - \epsilon)2^m$.

Many of our constructions use previous extractor constructions as black boxes. The parameters of our final constructions depend on the parameters of the strong seeded extractors used. Here we list the previous constructions that we will need to achieve the parameters claimed.

2.6 Seeded Extractors

Theorem 2.22. *[LRVW03] For any constant $\alpha \in (0,1)$, every $n \in \mathbb{N}$ and $k \leq n$ and every $\epsilon \in (0,1)$ where $\epsilon > exp(-\sqrt{k})$, there is an explicit (n,k,ϵ) seeded extractor* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^{O(\log n + \log(n/k)\log(1/\epsilon))} \to \{0,1\}^{(1-\alpha)k}$.

Theorem 2.23. *[RRV99] For every $n,k,m \in \mathbb{N}$ and $\epsilon > 0$, such that $m \leq k \leq n$, there is an explicit (n,k,ϵ)-strong seeded extractor* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with* $d = O\left(\frac{\log^2(n/\epsilon)}{\log(k/m)}\right)$.

We shall be interested in the following two instantiations of this theorem, obtained by setting the parameters appropriately:

Corollary 2.24. *[RRV99] For every $n \in \mathbb{N}$, constants $r > 0, \gamma < 1$, there is an explicit (n, n^γ, n^{-r})-strong seeded extractor* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{n^{\gamma'}}$ *with* $d = O(\log(n))$.

Corollary 2.25. *[RRV99] For every $n,k \in \mathbb{N}$, there is an explicit (n,k,ϵ)-strong seeded extractor* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\Omega(k)}$ *with* $d = O(\log^2(n/\epsilon))$.

The first instantiation will be used when we need an extractor that has a good seed length. The second will be used when we need an extractor that has good output length.

Composing Corollary 2.25 with a construction from [WZ99], we get the following.

Corollary 2.26. *For every $n,k \in \mathbb{N}$, $\epsilon > 0$, there is an explicit (n,k,ϵ)-strong extractor* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{k - O(\log^3(n/\epsilon))}$ *with* $d = O(\log^3(n/\epsilon))$.

2.7 Few Source Extractors

Recently Jean Bourgain constructed a strong 2-source extractor for min-entropy rate slightly less than half.

Theorem 2.27. *[Bou05] There exists a polynomial time computable function* $\mathsf{Bou} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ *and a universal constant α_0 s.t. if X,Y are independent $(n, (1/2 - \alpha_0)n)$ sources,*

$$\| Y \circ \mathsf{Bou}(X,Y) - Y \circ U_m \| \leq 2^{-\Omega(n)}$$

where $m = \Omega(n)$ and U_m is independent of Y.

3 Extracting from Independent Sources by Condensing ES-Sources

Our main result is a new deterministic extractor that can extract from a constant number of independent sources which have min-entropy that is polynomially small in their length. Actually the construction can handle a slightly more general class of sources. We can extract from just 2 independent sources, where each source is a blockwise source with a constant number of blocks. To simplify the presentation, we will first present the construction assuming that we have access to truly independent sources. In the next section we will show how to prove that the construction succeeds even when given just two independent blockwise sources.

Our algorithms will repeatedly condense ES-sources. Starting with a number of independent ES-sources, we will iteratively reduce the number of blocks in each of the sources, until the number of blocks is so small that the problem becomes easy.

In this section we will prove the following theorem:

Theorem 3.1 (Main Theorem). *For every constant $c > 0$, there exists a polynomial time computable function* $\mathsf{IExt} : (\{0,1\}^n)^u \to \{0,1\}^k$ *s.t. for constant $\gamma < 1$ if $X^1, X^2, \ldots, X^u$ are independent (n,k) sources with $k > \log^4 n$, $u = O(\frac{\log n}{\log k})$ then*

$$\| \mathsf{IExt}(X^1, \ldots, X^u) - U_k \| < \epsilon$$

with $\epsilon = 1/n^c$.

Setting the parameters appropriately gives the following corollary:

Corollary 3.2. *For every constant $c > 0$, there exists a polynomial time computable function* $\mathsf{IExt} : (\{0,1\}^n)^u \to \{0,1\}^k$ *s.t. for constant $\gamma < 1$ if $X^1, X^2, \ldots, X^u$ are independent (n, n^γ) sources with u some large enough constant, k polynomial in n, then*

$$\| \mathsf{IExt}(X^1, \ldots, X^u) - U_k \| < \epsilon$$

with $\epsilon = 1/n^c$.

Our first step will be to convert each of the sources to an ES-source. The following proposition, which we state without proof, shows how to do such a conversion.

Proposition 3.3. *Let* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be a seeded (n,k,ϵ) strong extractor. Let X be any (n,k) source. Let $\{0,1\}^d = \{s_1, s_2, \ldots, s_{2^d}\}$. Then* $\mathsf{Ext}(X, s_1) \circ \mathsf{Ext}(X, s_2) \circ \cdots \circ \mathsf{Ext}(X, s_{2^d})$ *is ϵ-close to a $(2^d, m)$ ES-source.*

Using any good seeded strong extractor with seed length $O(\log n)$, we can do the conversion in polynomial time.

Definition 3.4. We will say that a collection of ES-sources $X^1, \ldots, X^u$ is *aligned* if there is some i for which the i'th block of every ES-source in the collection is uniformly distributed.

If the strong extractor that we used to convert the input general sources to ES-sources has error ϵ, at most $\sqrt{\epsilon}$ fraction of the blocks in each source are not $\sqrt{\epsilon}$-close to uniform. Thus, if we are given u sources, as long as $u\sqrt{\epsilon} < 1$, we will have one aligned block in *every* source which is $\sqrt{\epsilon}$-close to uniform. Using Proposition 2.12 these sources are $u\sqrt{\epsilon}$-close to being the distribution of independent aligned ES-sources.

Proposition 3.5. *Let* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be a seeded (n,k,ϵ) strong extractor. Let $X^1, \ldots, X^u$ be independent (n,k) sources, with $u\sqrt{\epsilon} < 1$. Let $\{0,1\}^d = \{s_1, s_2, \ldots, s_{2^d}\}$. Let Z^i denote* $\mathsf{Ext}(X^i, s_1) \circ \mathsf{Ext}(X^i, s_2) \circ \cdots \circ \mathsf{Ext}(X^i, s_{2^d})$. *Then $Z^1 \circ \cdots \circ Z^u$ is $u\sqrt{\epsilon}$-close to the distribution of u independent aligned $(2^d, m)$ ES-sources.*

If Ext is a strong seeded extractor with seed length $O(\log n)$ and output length m, we can use Proposition 3.5 to reduce the problem of extracting from independent sources to the problem of extracting from aligned independent $(\mathsf{poly}(n) \times m)$ ES-sources. It turns out that it is easy to extract from independent aligned ES-sources when each ES-source contains very few blocks.

The rest of this section is organized in the following way:

1. We will describe a couple of ways to extract from a few independent aligned $(c \times n)$ ES-sources when c is a constant.

2. We will show how to use the extractors from the previous step to build condensers for independent aligned ES-sources. We will use the condensers to give a basic extractor that can extract from $O(\log n)$ independent aligned ES-sources which have $\mathsf{poly}(n)$ blocks. Using Proposition 3.5, this will give an extractor for $O(\log n)$ general independent sources.

3. We will add a few more tricks to bring down the number of sources required to $O(\log n / \log k)$.

3.1 Strong Extractors for independent aligned $(2 \times n)$ ES-sources

If we were given just 2 independent aligned $(2 \times n)$ ES-sources X^1 and X^2, it is easy to see that

$$\| X^1_{\{1\} \times [n]} \oplus X^2_{\{2\} \times [n]} - U_n \| = 0$$

i.e. to get random bits we just have to XOR the first block from the first source with the second block from the second source.

We will actually need something stronger: a strong extractor for such sources. We can get such a strong extractor for 3 $(2 \times n)$ aligned independent ES-sources by composing the XOR function with a strong seeded extractor. The following theorem is easy to see. We state it without proof.

Theorem 3.6. *Let* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be an* (n, k, ϵ) *strong seeded extractor. Let* $\overline{\mathsf{BasicESExt}} : \{0,1\}^n \times \{0,1\}^{2d} \times \{0,1\}^{2d} \to \{0,1\}^m$ *be defined as* $\overline{\mathsf{BasicESExt}}(x, y^1, y^2) = \mathsf{Ext}(x, y^1_{\{1\} \times [d]} \oplus y^2_{\{2\} \times [d]})$. *Then if* X, Y^1, Y^2 *are independent sources, with* X *an* (n, k) *source and* Y^1, Y^2 *aligned* $(2 \times d)$ *ES-sources,*

$$\| \overrightarrow{Y} \circ \overline{\mathsf{BasicESExt}}(X, \overrightarrow{Y}) - \overrightarrow{Y} \circ U_m \| < \epsilon$$

where $\overrightarrow{Y}$ *denotes* $Y^1 \circ Y^2$ *and* U_m *is independent of* $\overrightarrow{Y}$.

Claim 3.7. *Let* X *and* Y *be as in Theorem 3.6. Then*

$$\Pr_{\vec{y} \leftarrow_R \overrightarrow{Y}} [\| \overline{\mathsf{BasicESExt}}(X, \vec{y}) - U_m \| \geq \sqrt{\epsilon}] < \sqrt{\epsilon}$$

To give an example of the kinds of parameters that can be achieved, if we start with 3 independent aligned $(2 \times n)$ ES-sources and use the extractor promised by Corollary 2.26, we can get $n - o(n)$ random bits out, with error that is exponentially small in n. Setting parameters appropriately, we can get the following corollaries.

Corollary 3.8. *Let* Ext *be the extractor from Corollary 2.26. Setting* $k = d = r, n = 2r$, $\overline{\mathsf{BasicESExt}}$ *can be set up to output* $m = r - O(\log^3(r/\epsilon))$ *random bits with error* ϵ.

Corollary 3.9. *Let* Ext *be the extractor from Corollary 2.26. Setting* $k = d = r, n = 2r$, $\overline{\mathsf{BasicESExt}}$ *can be set up to output* $m = r - O(\log^6 r)$ *random bits with error* $\epsilon < 2^{-\log^2 r}$.

Corollary 3.10. *Let* Ext *be the extractor from Corollary 2.26. Setting* $k = d = r, n = 2r$, $\overline{\mathsf{BasicESExt}}$ *can be set up to output* $m = r - \sqrt{r}$ *random bits with error* $\epsilon < 2^{-r^{\Omega(1)}}$.

Another way to get a strong extractor from just two ES-sources of this type is to use Bourgain's recent construction Theorem 2.27. This already gives a strong extractor for two ES-sources with min-entropy rate half. However Bourgain's extractor extracts only a constant fraction of the randomness. We can remedy this by composing it with a strong seeded extractor to get almost all the randomness out. We state the following theorem without proof.

Theorem 3.11. *Let* $\mathsf{Ext} : \{0,1\}^{2n} \times \{0,1\}^d \to \{0,1\}^m$ *be a* (n, k, ϵ) *strong seeded extractor. Let* Bou *and* α_0 *be as in Theorem 2.27. Let* n, k' *be such that* $n - 100k' > k$ *and* d *be the output length of* Bou *when applied to two independent* $(2k', k')$ *sources. Let* $\overline{\mathsf{Basic2ESExt}} : \{0,1\}^{2n} \times \{0,1\}^{2n} \to \{0,1\}^m$ *be defined as* $\overline{\mathsf{Basic2ESExt}}(X, Y) = \mathsf{Ext}(X_{\{1,2\} \times [n]}, \mathsf{Bou}(Y_{\{1,2\} \times [k']}, X_{\{1,2\} \times [k']}))$. *Then*

$$\| Y \circ \overline{\mathsf{Basic2ESExt}}(X, Y) - Y \circ U_m \| < \epsilon + 2^{-\Omega(k')}$$

where U_m *is independent of* Y.

As with the previous construction, depending on the parameters of the strong seeded extractor used, we can get various tradeoffs between the error and the output length.

Claim 3.12. *Let X and Y be as in Theorem 3.11. Then*

$$\Pr_{y \leftarrow_R Y}[\| \overline{\mathsf{Basic2ESExt}}(X,y) - U_m \| \geq \sqrt{\epsilon + 2^{-\Omega(k')}}] < \sqrt{\epsilon + 2^{-\Omega(k')}}$$

3.2 Warm-up: Extracting randomness from $O(\log n)$ independent sources

To illustrate some of the ideas behind the main construction of a few source extractor, we will first describe how to extract from $O(\log n)$ independent sources. The final extractor construction will be slightly more involved but will use the major ideas that we develop here.

In this section, we will prove the following theorem:

Theorem 3.13. *There exists a polynomial time computable function $\mathsf{IExt}' : (\{0,1\}^n)^u \to \{0,1\}^m$ s.t. for constants $\gamma < 1$ if $X^1, X^2, \ldots, X^u$ are independent (n, n^γ) sources with $u = O(\log(n)$, $k = \Omega(\log^4 n)$, $m = \Omega(k)$ then*

$$\| \mathsf{IExt}'(X^1, \ldots, X^u) - U_m \| < \epsilon$$

with $\epsilon = 1/n^{\Omega(1)}$.

As discussed, we will start by converting the independent (n, k) sources to distributions which are statistically close to being independent aligned $(n^d \times m)$ ES-sources using Proposition 3.5. Here m is the output length of the strong seeded extractor used in the conversion. There is a tradeoff between m and the error incurred in the conversion. This tradeoff in general depends on the relationship between n and k. To give a feel for the parameters, if $k = n^\gamma$ for some $\gamma \in (0,1)$, we can do the conversion with $m = k - o(k)$ and error that is polynomially small.

Our construction will be obtained by iteratively condensing the original distribution. We will start with $O(\log n)$ independent aligned $(n \times k)$ ES-sources. In each step we will reduce the number of independent ES-sources by 2, but will reduce the number of blocks in each of the sources by a *factor* of 2. Actually, in each step we will convert ES-sources with t blocks into ES-sources with $\lceil t/2 \rceil$ blocks. To simplify the presentation we will assume that t is always even, this does not really affect any of the claims made. After $O(\log n)$ steps, we will have reduced the number of blocks to 2. We can then use the XOR function to extract random bits. Now we describe one condensing step in detail.

Let $S_1 = \{1,2\} \times [r], S_2 = \{3,4\} \times [r], \ldots, S_{t/2} = \{t-1, t\} \times [r]$. In the rest of this section we will write X_j to denote X_{S_j}.

Construction: $\mathsf{ICond}(X^1, \ldots, X^u)$
Input: $X^1, \ldots, X^u$, $(t \times r)$ ES-sources.
Output: $Z^1, \ldots, Z^{u-2}$.
Let $\overline{\mathsf{BasicESExt}}$ be as in Theorem 3.6.

1. For $1 \leq i \leq u-2$, $1 \leq j \leq t/2$ let $Z^i_j = \overline{\mathsf{BasicESExt}}(X^i_j, X^{u-1}_j, X^u_j)$.

2. For $1 \leq i \leq u-2$, let Z^i be the ES-source whose blocks are $Z^i_1, Z^i_2, \ldots, Z^i_{t/2}$.

Lemma 3.14. *If $X^1, \ldots, X^u$ are independent aligned $(t \times r)$ ES-sources, $Z^1, \ldots, Z^{u-2}$ are $2u\sqrt{\epsilon}$-close to being a convex combination of independent aligned $(t/2 \times m)$ ES-sources, where m and ϵ are the output length and error of $\overline{\mathsf{BasicESExt}}$.*

Assuming this lemma, we can prove the theorem for this section.

Proof of Theorem 3.13. If we use the strong ES-source extractor promised by Corollary 3.10, after applying the condenser $O(\log n)$ times, we will have reduced the number of blocks in each of the sources to 2 and the length of each of the sources will still be $k - O(\sqrt{k} \log n)$. We can then use the XOR function to get a distribution which is statistically close to uniform. The error adds in each step, but since the error of $\overline{\mathsf{BasicESExt}}$ can be made as small as $2^{-k^{\Omega(1)}}$, for large k this does not affect the final error. The dominant error comes in the first step, when we convert the general sources to ES-sources. This concludes the proof of Theorem 3.13. ■

Proof of Lemma 3.14. Let $\overrightarrow{Y}$ denote the concatenation of X^{u-1}, X^u. Let h be s.t. S_h contains the truly random block. Let $\overrightarrow{Y_h}$ denote the concatenation of X_h^{u-1}, X_h^u. Let m be the output length of $\overline{\mathsf{BasicESExt}}$. We will prove the lemma by partitioning the support of $\overrightarrow{Y}$ into a good set and a bad set s.t.

Claim 3.15. *For good $\vec{y}$, the distribution $Z^1|\overrightarrow{Y} = \vec{y} \circ \cdots \circ Z^{u-2}|\overrightarrow{Y} = \vec{y}$ is $u\sqrt{\epsilon}$-close to being a collection of independent aligned $(\lceil t/2 \rceil \times m)$ ES-sources.*

Claim 3.16. $\Pr[\overrightarrow{Y}$ *is not good*$] < u\sqrt{\epsilon}$

We will call $\vec{y}$ good for X^i if

$$\| \overline{\mathsf{BasicESExt}}(X_h^i, \vec{y}_h) - U_m \| < \sqrt{\epsilon}$$

We will call $\vec{y}$ good if it is good for all $1 \leq i \leq u-2$.

Since $\overline{\mathsf{BasicESExt}}$ is a strong extractor, for any i, a $\sqrt{\epsilon}$ fraction of the seeds are good for X^i by Claim 3.7. The second claim then follows by the union bound.

For any fixed $\vec{y}$, $Z^1|\overrightarrow{Y} = \vec{y}, \ldots, Z^u|\overrightarrow{Y} = \vec{y}$ are independent. When $\vec{y}$ is good, each of the Z^i's is $\sqrt{\epsilon}$-close to being a $(t/2 \times m)$ ES-source, with the hth block being the random one. By the union bound and Proposition 2.12 we get the first claim and the lemma follows. □

3.3 Extracting from fewer sources

In this section we will prove Theorem 3.1. We will obtain the final construction in the following steps.

1. We will show how to extract from 3 independent aligned $(n^\gamma \times n)$ ES-sources for any constant $\gamma < 1$.
2. We will show how to use the construction from the previous step to extract from $O(\frac{\log n}{\log k})$ independent aligned $(n \times k)$ ES-sources when $k > \log^4 n$.

3.3.1 Extracting from 3 independent aligned $(n^\gamma \times n)$ ES-sources, for $\gamma < 1$

Theorem 3.17. *There exists a polynomial time computable function $\overline{\mathsf{3ESExt}} : (\{0,1\}^n)^3 \rightarrow \{0,1\}^m$ s.t. for constant $\gamma < 1$ if X^1, X^2, X^3 are independent aligned $(n^\gamma \times n)$ ES-sources,*

$$\| \overline{\mathsf{3ESExt}}(X^1, X^2, X^3) - U_m \| < 2^{-n^{\Omega(1)}}$$

where $m = n - O(n^\beta)$ for some $\beta \in (0,1)$.

Our starting point is the construction that can extract from $O(\log n)$ sources. We'd like to do more or less the same thing in this situation, but somehow *reuse* the sources. As in that situation, our construction will work by repeated condensing, but this time we will not discard any sources. Starting from 3 independent aligned ES-sources, in each step we will output a distribution that is a convex combination of 3 independent aligned ES-sources. The number of block in each of the sources will be reduced by a factor of 2 and the length of each block will be reduced by a little bit.

Intuitively what we will try and do is: to condense the i'th source, we will get the other two sources to conspire against it. We will use the strong independent aligned $(2 \times n)$ ES-sources extractor of Theorem 3.6 on the i'th source, with small slices of the other sources as 'seed'. Conditioned on all the small sections of the sources that we've used as seed, we show that the condensing succeeds, we obtain 3 new ES-sources which have half the number of blocks as the original sources. Since we're conditioning on the only part that's involved in the interactions between the sources, after conditioning, the output of the condensing step is a collection of independent sources. Iterating this condensing process, we will eventually obtain a single string that is statistically close to uniformly distributed.

Now we describe one condensing step in detail. As in the previous section, we will assume that t is even.

We are given: X^1, X^2, X^3, independent aligned $(t \times r)$ ES-sources.

Let w and l be parameters that we will choose later (we will have to set w to roughly polylog(r) and l to roughly r^μ for some constant $\mu < 1$). Let $S_1 = \{1,2\} \times [w], S_2 = \{3,4\} \times [w], \dots, S_{t/2} = \{t-1,t\} \times [w]$.

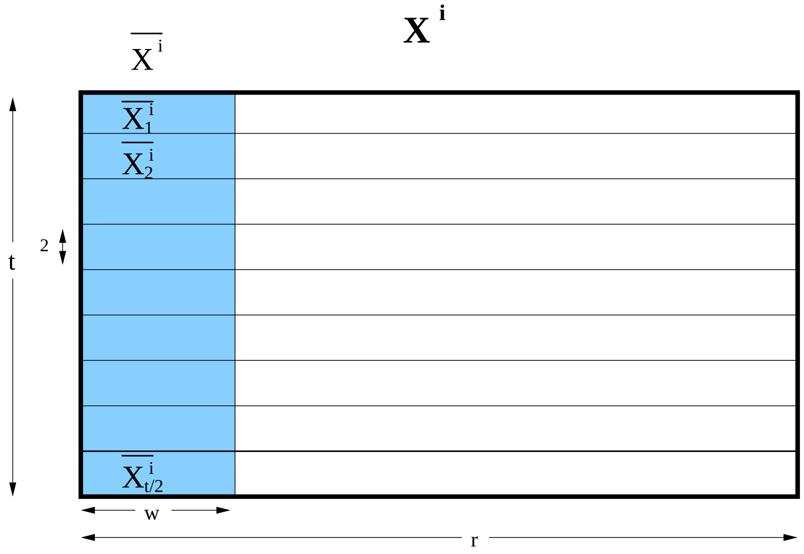

Figure 1: Notation in one source

We will adhere to the following notational guidelines:

- The superscript of an expression (if any) indicates which one of the independent sources we are referring to.
- The subscript of an expression (if any) indicates which of the blocks we are referring to.
- An over-line (for example: $\overline{X}$) indicates whether we are referring to the entire block or just a small section of the blocks.

For all i, j, we introduce the following notation:

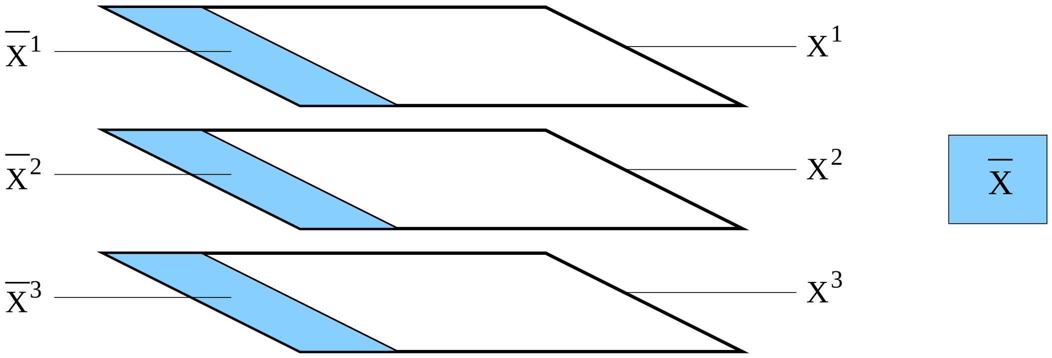

Figure 2: The region $\overline{X}$

- $\overline{X^i_j} = X^i_{S_j}$, a small slice of the jth pair of blocks in X^i.
- $\overline{X^i} = X^i_{[t]\times[w]} = \overline{X^i_1} \circ \cdots \circ \overline{X^i_{t/2}}$, a small slice of X^i.
- $\overline{X} = \overline{X^1} \circ \overline{X^2} \circ \overline{X^3}$.
- $\overline{X_j} = \overline{X^1_j} \circ \overline{X^2_j} \circ \overline{X^3_j}$.
- $\overline{X^{\neq i}_j}$ denotes the concatenation of $\overline{X^v_j}$ for all $v \neq i$.

Construction: $\overline{\mathsf{ICond}}(X^1, X^2, X^3)$
Input: X^1, X^2, X^3, independent aligned $(t \times r)$ ES-sources.
Output: Z^1, Z^2, Z^3.

Let $\overline{\mathsf{BasicESExt}}$ be the extractor that can extract from 3 independent sources A, B, C when A is a $(tr, r - l)$ source and B, C are independent aligned $(2 \times w)$ sources promised by Theorem 3.6.

1. For all i, j let $Z^i_j = \overline{\mathsf{BasicESExt}}(X^i, \overline{X^{\neq i}_j})$.
2. For all i, let Z^i be the ES-source whose blocks are $Z^i_1, Z^i_2, \ldots, Z^i_{t/2}$.
3. For all i, output Z^i.

Lemma 3.18. *Let $\overline{\mathsf{ICond}}$ be as above. If X^1, X^2, X^3 are independent aligned $(t \times r)$ ES-sources, Z^1, Z^2, Z^3 are $3(2\sqrt{\epsilon} + 2^{-(l-tw)})$-close to being a convex combination of independent aligned $(t/2 \times m)$ ES-sources, where m and ϵ are the output length and error of $\overline{\mathsf{BasicESExt}}$.*

Proof. We will roughly follow the proof for the situation in which we had $O(\log n)$ sources. There are a few additional complications that we have to deal with now.

Let h be s.t. S_h contains the truly random block. We will prove the lemma by partitioning the support of $\overline{X}$ into a good set and a bad set s.t.

Claim 3.19. *For good $\overline{x}$, the distribution $(Z^1 | \overline{X} = \overline{x}) \circ (Z^2 | \overline{X} = \overline{x}) \circ (Z^3 | \overline{X} = \overline{x})$ is $3\sqrt{\epsilon}$-close to being a collection of independent aligned $(t/2 \times m)$ ES-sources.*

Claim 3.20. $\Pr[\overline{X}$ *is not good*$] < 3(\sqrt{\epsilon} + 2^{-(l-tw)})$.

Here we will need a more involved notion of good. For any fixed i, we will say $\overline{x}$ is good for i if

$$\| \overline{\mathsf{BasicESExt}}(X^i|\overline{X^i}=\overline{x^i}, \overline{x_h^{\neq i}}) - U_m \| < \sqrt{\epsilon}$$

We will call $\overline{x}$ good if it is good for all i.

For any fixed $\overline{x}$, $Z^1|\overline{X}=\overline{x}$, $Z^2|\overline{X}=\overline{x}$, $Z^3|\overline{X}=\overline{x}$ are independent. When $\overline{x}$ is good, we have that for any i,

$$\begin{aligned}(Z_h^i|\overline{X}=\overline{x}) &= \overline{\mathsf{BasicESExt}}(X^i|\overline{X}=\overline{x}, \overline{x_h^{\neq i}}) \\ &= \overline{\mathsf{BasicESExt}}(X^i|\overline{X^i}=\overline{x^i}, \overline{x_h^{\neq i}}) && \text{since } X^i \text{ is independent of } X^j \text{ for } j \neq i\end{aligned}$$

which is $\sqrt{\epsilon}$-close to the uniform distribution by our notion of *good*. Thus we get that Z^i is $\sqrt{\epsilon}$-close to being a $(t/2 \times m)$ ES-source, with the hth block being the random one. By the union bound, we then get the first claim.

Now we prove the second claim.

Proof of Claim 3.20. We will first bound the probability that $\overline{X}$ is bad for a fixed i and then use the union bound to bound the probability that it is bad for all i. In the following statements, we will explicitly state what the probabilities are over to avoid confusion. The probability we are trying to bound is this one:

$$\begin{aligned}&\Pr_{\overline{x}\leftarrow_{\mathrm{R}}\overline{X}}[\overline{x} \text{ is bad for } i] \\ &= \Pr_{\overline{x}\leftarrow_{\mathrm{R}}\overline{X}}[\| \overline{\mathsf{BasicESExt}}((X^i|\overline{X^i}=\overline{x^i}), \overline{x_h^{\neq i}}) - U_m \| \geq \sqrt{\epsilon}]\end{aligned}$$

We will rewrite this probability in this way:

$$\begin{aligned}&\Pr_{\overline{x}\leftarrow_{\mathrm{R}}\overline{X}}[\overline{x} \text{ is bad for } i] \\ &= \sum_{p\in\mathsf{supp}(\overline{X^i})} \Pr_{\overline{x}\leftarrow_{\mathrm{R}}(\overline{X}|\overline{X}^i=p)}[\overline{x} \text{ is bad for } i] \Pr_{v\leftarrow_{\mathrm{R}}\overline{X}^i}[v=p]\end{aligned}$$

We'd like to argue that for every term in this sum, either the first probability is small, or the second probability is small. To this end, we partition the support of $\overline{X}^i$ into two sets. Recall that $H_\infty(X^i) \geq r$. We will say p is *atypical* if $H_\infty(X^i|\overline{X^i}=p) < r - l$. Otherwise we will say that p is *typical*. By Proposition 2.6,

Claim 3.21. *For* atypical p, $\Pr_{v\leftarrow_R \overline{X}^i}[v=p] < 2^{-l}$.

On the other hand, when p is *typical*,

$$\begin{aligned}&\Pr_{\overline{x}\leftarrow_{\mathrm{R}}(\overline{X}|\overline{X^i}=p)}[\overline{x} \text{ is bad for } i] \\ &= \Pr_{\overline{x}\leftarrow_{\mathrm{R}}(\overline{X}|\overline{X^i}=p)}[\| \overline{\mathsf{BasicESExt}}(X^i|\overline{X^i}=p, \overline{x_h^{\neq i}}) - U_m \| \geq \sqrt{\epsilon}] \\ &= \Pr_{\overline{x'}\leftarrow_{\mathrm{R}}(X_h^{\neq i}|\overline{X^i}=p)}[\| \overline{\mathsf{BasicESExt}}(X^i|\overline{X^i}=p, \overline{x'}) - U_m \| \geq \sqrt{\epsilon}]\end{aligned}$$

Now observe that

$$\begin{aligned}&\overline{X_h^{\neq 1}}|\overline{X^1}=p\\ &=\overline{X_h^2}|\overline{X^1}=p\circ\overline{X_h^3}|\overline{X^1}=p\\ &=\overline{X_h^2}\circ\overline{X_h^3}\end{aligned}$$

since $\overline{X_h^2}$ and $\overline{X_h^3}$ are independent of $\overline{X^1}$. We get similar statements for $\overline{X_h^{\neq 2}}|\overline{X^2}=p$ and $\overline{X_h^{\neq 3}}|\overline{X^3}=p$. Therefore, for p that is *typical*, using Claim 3.7 we have,

$$\begin{aligned}&\Pr_{\overline{x}\leftarrow_{\mathrm{R}}(\overline{X}|\overline{X^i}=p)}[\overline{x}\text{ is bad for }i]\\ &=\Pr_{\overline{x'}\leftarrow_{\mathrm{R}}\overline{X_h^{\neq i}}}[\|\ \overline{\mathsf{BasicESExt}}(X^i|\overline{X^i}=p,\overline{x'})-U_m\ \|\geq\sqrt{\epsilon}]\\ &<\sqrt{\epsilon}\end{aligned}$$

Thus,

Claim 3.22. *For* typical *p,* $\Pr_{\overline{x}\leftarrow_R(\overline{X}|\overline{X^i}=p)}[\overline{x}$ *is bad for* $i]<\sqrt{\epsilon}$*.*

Going back to the quantity we were trying to bound,

$$\begin{aligned}&\Pr_{\overline{x}\leftarrow_{\mathrm{R}}\overline{X}}[\overline{x}\text{ is bad for }i]\\ &=\sum_{p\in\mathsf{supp}(\overline{X^i})}\Pr_{\overline{x}\leftarrow_{\mathrm{R}}(\overline{X}|\overline{X^i}=p)}[\overline{x}\text{ is bad for }i]\Pr_{v\leftarrow_{\mathrm{R}}\overline{X^i}}[v=p]\\ &=\sum_{p\text{ is typical}}\Pr_{\overline{x}\leftarrow_{\mathrm{R}}(\overline{X}|\overline{X^i}=p)}[\overline{x}\text{ is bad for }i]\Pr_{v\leftarrow_{\mathrm{R}}\overline{X^i}}[v=p]+\sum_{p\text{ is atypical}}\Pr_{\overline{x}\leftarrow_{\mathrm{R}}(\overline{X}|\overline{X^i}=p)}[\overline{x}\text{ is bad for }i]\Pr_{v\leftarrow_{\mathrm{R}}\overline{X^i}}[v=p]\\ &\leq\sqrt{\epsilon}+2^{-l}2^{tw}\end{aligned}$$

The first sum was bounded using the fact that $\sum_p\Pr_{v\leftarrow_{\mathrm{R}}\overline{X^i}}[v=p]\leq 1$ and Claim 3.22. The second sum was bounded using the fact that the total number of possible *atypical* p's is at most 2^{tw} and Claim 3.21. Using the union bound over all i, Claim 3.20 follows.

□

This concludes the proof of Lemma 3.18. □

Proof of Theorem 3.17. Let $\overline{\mathsf{3ESExt}}$ be the following function.

Construction: $\overline{\mathsf{3ESExt}}(X^1,X^2,X^3)$
Input: X^1,X^2,X^3, independent aligned $(n^\gamma\times n)$ ES-sources, $\gamma\in(0,1)$.
Output: Z.

1. Repeatedly condense the sources using $\overline{\mathsf{ICond}}$ from Lemma 3.18 until each source has just one block.

2. Output Z, the block from the first source.

Since we need to repeat the condensation step at most $\lceil \log n \rceil$ times, by Lemma 3.18 the final error is $O((\sqrt{\epsilon} + 2^{-(l-tw)}) \log n)$. If we use $\overline{\mathsf{BasicESExt}}$ as in Corollary 3.8, the final output length is at least $n - O(l \log^4(n/\epsilon))$.

Setting $l = 2n^{(1+\gamma)/2}$, $w = l/(2t)$ and $\epsilon = 2^{-n^{\Omega(1)}}$, we get a total error of $2^{-n^{\Omega(1)}}$ with final output length at least $n - n^{\beta}$ for some $\beta \in (0, 1)$.

■

Replacing the extractor from Theorem 3.6 with Bourgain's extractor from Theorem 3.11, we can extract from just 2 independent $(n^{\gamma} \times n)$ ES-sources for any $\gamma \in (0, 1)$. In addition, we can actually show that the extractor is strong. To summarize, we obtain the following theorem, which we state without proof:

Theorem 3.23. *There exists a polynomial time computable function* $\overline{\mathsf{2ESExt}} : (\{0,1\}^n)^2 \to \{0,1\}^m$ *s.t. for constant* $\gamma < 1$ *if* X, Y *are independent aligned* $(n^{\gamma} \times n)$ *ES-sources,*

$$\| Y \circ \overline{\mathsf{2ESExt}}(X, Y) - Y \circ U_m \| < 2^{-n^{\Omega(1)}}$$

where U_m *is independent of* Y *and* $m = n - O(n^{\beta})$ *for some* $\beta \in (0, 1)$.

Remark 3.24. Using Bourgain's extractor here seems like overkill. Bourgain's extractor can extract from *any* 2 sources with min-entropy rate slightly less than half, where as our sources have a lot of structure. It would be interesting to find a simple construction like that in Theorem 3.6 which is a strong extractor for 2 independent aligned $(2 \times n)$ ES-sources.

3.3.2 Extracting from $O(\frac{\log n}{\log k})$ aligned independent $(n \times k)$ ES-sources

In this section we describe the final trick needed to complete the construction of the independent sources extractor. To do so we will use the extractor $\overline{\mathsf{3ESExt}}$ from Theorem 3.17 as a black box. We will set up $\overline{\mathsf{3ESExt}}$ to extract from three independent $(\sqrt{r} \times r)$ ES-sources. The extractor is obtained by repeatedly using the construction from the previous section to obtain ES sources with few and fewer blocks.

Let $S_1 = \{1, \ldots, \sqrt{r}\} \times [r], S_2 = \{\sqrt{r}+1, \ldots, 2\sqrt{r}\} \times [r], \ldots, S_{t/\sqrt{r}} = \{t - \sqrt{r} + 1, \ldots, t\} \times [r]$.

Construction: $\mathsf{ICond}(X, Y^1, Y^2, Y^3)$
Input: X, an (n, k) source and Y^1, Y^2, Y^3, $(t \times r)$ ES-sources.
Output: Z.
Let Ext be a (n, k, ϵ) extractor with output length m and seed length r.

1. For all $1 \le j \le t/\sqrt{r}$, let $Z_j = \mathsf{Ext}(X, \overline{\mathsf{3ESExt}}(Y^1_{S_j}, Y^2_{S_j}, Y^3_{S_j}))$.
2. Let $Z = Z_1 \circ \cdots \circ Z_{t/\sqrt{r}}$.

The following lemma is easy to see given our previous work:

Lemma 3.25. *If* X, Y^1, Y^2, Y^3 *are independent sources, with* X *an* (n, k) *source and* Y^1, Y^2, Y^3 *aligned* $(t \times r)$ *ES-sources, then* Z *is* ϵ *close to a* $(t/\sqrt{r} \times m)$ *ES-sources.*

Here the error can be made exponentially small in r. In addition, notice that if X^1, X^2, X^3 are independent (n, k) sources and Y^1, Y^2, Y^3 are as before, $\mathsf{ICond}(X^1, Y^1, Y^2, Y^3) \circ \mathsf{ICond}(X^2, Y^1, Y^2, Y^3) \circ \mathsf{ICond}(X^3, Y^1, Y^2, Y^3)$ is $\sqrt{\epsilon}$ close to being a convex combination of independent aligned $(t/\sqrt{r} \times m)$ ES-sources as long as $3\sqrt{\epsilon} < 1$. Using this basic tool, we can now prove the main theorem.

Proof of Theorem 3.1. First consider the following function to extract from independent (n, k) sources. Here d is a constant that depends on the seed length of the strong seeded extractor used and the output length of the strong extractor.

Construction: $\mathsf{IExt}(X^1, \ldots, X^{3d\frac{\log n}{\log k}+3})$

1. First use a strong seeded extractor to convert three of the sources to aligned independent ES-sources.
2. Iteratively run ICond on three general sources using the three ES-sources as seed. In each step we obtain three ES-sources with fewer blocks. After $d\log(n)/\log(k)$ iterations, we will have brought the number of blocks down to small enough.
3. Finally apply $\overline{\mathsf{3ESExt}}$ to get random bits.

The error adds in each step, but as long as $k > \log^4 n$, the dominant error comes from the conversion of the general source to an ES-source (where we incur error of $1/\mathsf{poly}(n)$).

In this way we get an extractor for $O(\frac{\log n}{\log k})$ independent (n, k) sources with output length $k - o(k)$ and error of $1/\mathsf{poly}(n)$. To get k bits, we can simply run the extractor twice on two disjoint sets of independent sources to double the output length. To reduce the error, we will use the following lemma from [BIW04]:

Lemma 3.26. *[BIW04] Let $Z^1, \ldots, Z^v$ be independent distributions over $\{0,1\}^k$ with $\| Z^i - U_k \| < \epsilon$ for every $i = 1, \ldots, v$. Then*

$$\| Z^1 \oplus Z^2 \oplus \cdots \oplus Z^v - U_k \| < \epsilon^v$$

Using this lemma, we can increase the number of sources used by a constant factor to reduce the error to less than $1/n^c$ for any constant c.

■

4 Extracting from 2 Independent Blockwise Sources

In this section we show how to use essentially the same construction from the previous section to obtain an extractor for 2 independent blockwise sources with very few blocks. This is analogous to an idea from [BKS$^+$05], where they show how to use a 4-source somewhere extractor to get a somewhere extractor for 2 blockwise sources with just 2 blocks each. The main theorem of this section is the following:

Theorem 4.1 (2 Blockwise Source Extractor). *There exists a polynomial time computable function $\overline{\mathsf{IExt}} : \{0,1\}^{un} \times \{0,1\}^{un} \to \{0,1\}^k$ s.t. for constant $\gamma < 1$ if $X = X^1 \circ \cdots \circ X^u$ and $Y = Y^1 \circ \cdots \circ Y^u$ are independent $(k, \ldots, k)$ block-wise sources with $k > \log^4 n$, $u = O(\frac{\log n}{\log k})$ then*

$$\| \overline{\mathsf{IExt}}(X, Y) - U_k \| < \epsilon$$

with $\epsilon = 1/n^{\Omega(1)}$.

The extractor here is essentially the same one as the one we constructed in the previous section for a few truly independent sources, if we use $\overline{\mathsf{2ESExt}}$ from Theorem 3.23 at the lowest level instead of $\overline{\mathsf{3ESExt}}$.

Let $S_1 = \{1, \ldots, \sqrt{r}\} \times [r], S_2 = \{\sqrt{r}+1, \ldots, 2\sqrt{r}\} \times [r], \ldots, S_{t/\sqrt{r}} = \{t - \sqrt{r}+1, \ldots, t\} \times [r]$.

Construction: $\mathsf{ICond}(X^1 \circ \cdots \circ X^u, Y^1 \circ \cdots \circ Y^u)$

Input: $X = X^1 \circ \cdots \circ X^u$ and $Y = Y^1 \circ \cdots \circ Y^u$, independent blockwise sources with X^1 and Y^1 independent aligned $(t \times r)$ ES-sources.

Output: $A = A^1 \circ \cdots \circ A^{u-1}$ and $B = B^1 \circ \cdots \circ B^{u-1}$.

Let Ext be a (n, k, ϵ) extractor with output length m and seed length r.

1. For all $1 \leq j \leq t/\sqrt{r}$, let $A^1_j = \mathsf{Ext}(X^2, \overline{\mathsf{2ESExt}}(Y^1_{S_j}, X^1_{S_j}))$.
2. For all $1 \leq j \leq t/\sqrt{r}$, let $B^1_j = \mathsf{Ext}(Y^2, \overline{\mathsf{2ESExt}}(X^1_{S_j}, Y^1_{S_j}))$.
3. Let $A^1 = A^1_1 \circ \cdots \circ A^1_{t/\sqrt{r}}$.
4. Let $B^1 = B^1_1 \circ \cdots \circ B^1_{t/\sqrt{r}}$.
5. For all $2 \leq i \leq u-1$, let $A^i = X^{i+1}$, $B^i = Y^{i+1}$.

The following lemma can be obtained by applying the techniques from the previous section. We state it without proof.

Lemma 4.2. *If $X = X^1 \circ \cdots \circ X^u$ and $Y = Y^1 \circ \cdots \circ Y^u$ are independent $(k, \ldots, k)$-blockwise sources with each block except the first one of length n, and X^1 and Y^1 independent aligned $(t \times r)$ ES-sources, then $A = A^1 \circ \cdots \circ A^{u-1}$ and $B = B^1 \circ \cdots \circ B^{u-1}$ are statistically close to being a convex combination of independent $(k, \ldots, k)$-blockwise sources with each block except the first one of length n with A^1 and B^1 independent aligned $(t/\sqrt{r} \times m)$ ES-sources.*

Here the error can be made exponentially small in r. Iteratively applying this condenser, we obtain the extractor for Theorem 4.1.

Remark 4.3. It can be shown that the extractor from Theorem 4.1 is strong. The extractor can also be made to work when the two sources and all blocks are of different lengths with different min-entropies, as long as the parameters are all polynomially related.

4.1 A new 2-source somewhere-extractor

One of the main results in the work of Barak et. al. [BKS+05] was the construction of a new bipartite Ramsey Graph (a 2-source disperser). Their construction was based on a technique called the challenge response mechanism. A basic tool used as a black box in their construction is a *somewhere-extractor*:

Definition 4.4. A function $\mathsf{SExt} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to (\{0,1\}^m)^u$ is a (k_1, k_2, ϵ) 2-source somewhere-extractor if for all X, Y independent (n_1, k_1) and (n_2, k_2) sources respectively, SExt is ϵ close to a $(u \times m)$ somewhere-random source.

For the construction of the 2-source disperser, the techniques of [BKS+05] require that the somewhere-extractor used outputs a constant number of blocks with each block of length some arbitrarily large constant and error that is some arbitrarily small constant when given two independent constant min-entropy rate sources. Using their constant seed condenser, they construct a 2-source somewhere-extractor that outputs a constant number of blocks of linear length, with exponentially small error. Here we outline how to use our previous construction to give an alternate somewhere-extractor for two constant min-entropy rate sources that outputs a constant number of blocks of linear length with polynomially small error. This is good enough for the application of building a 2-source disperser. Thus we obtain a 2-source disperser which is not based on the results from additive number theory.

Construction: $\mathsf{SExt}(X, Y)$
Input: X, Y, two independent $(n_1, \delta_1 n_1)$, $(n_2, \delta_2 n_2)$ sources.
Output: $Z = Z^1 \circ \cdots \circ Z^v$
Let IExt be as in Theorem 4.1.

1. Let u be the number of blocks required by IExt in each source when the two sources have min-entropy rate $\min(\delta_1/2, \delta_2/2)$.
2. Let γ be s.t. $u\gamma \ll \min(\delta_1/100, \delta_2/100)$.
3. Partition the bits of X and Y into $1/\gamma$ blocks, each of equal length. $X = X^1, \ldots, X^{1/\gamma}$, $Y = Y^1, \ldots, Y^{1/\gamma}$.
4. For $i = 1, 2, \ldots, \binom{1/\gamma}{u}u!$, let $\pi_i(X)$ denote the source obtained by choosing u blocks from $X^1, \ldots, X^u$ and permuting them in the i'th way. Similarly define $\pi_i(Y)$.
5. For all pairs $(i, j) \in [\binom{1/\gamma}{u}u!] \times [\binom{1/\gamma}{u}u!]$, let the (i, j)'th block of Z be $\mathsf{IExt}(\pi_i(X), \pi_j(Y))$.
6. Output Z.

Lemma 4.5. *If X is an $(n_1, \delta_1 n_1)$ source and Y is an independent $(n_2, \delta_2 n_2)$ source,* $\mathsf{SExt}(X, Y)$ *is $1/n^{\Omega(1)}$ close to being somewhere-random.*

Proof Sketch. It can be shown that every $(n_1, \delta_1 n_1)$ source X is statistically close to a convex combination of sources $\{X^l\}_{l \in I}$ s.t. for each X^l, there exists an i for which $\pi_i(X^l)$ is a blockwise source. A similar statement is true for every source Y that is a $(n_2, \delta_2 n_2)$ source. Thus we get that $X \circ Y$ is statistically close to a convex combination of sources s.t. for every source in the combination there exists some (i, j) s.t. $\pi_i(X), \pi_j(Y)$ are independent blockwise sources. The extractor succeeds in extracting random bits for that choice of (i, j). Further, the number of possible choices for (i, j) is just a constant. □

Remark 4.6. Applying Raz's merger [Raz05] to the sources before we partition them, we can actually ensure that almost all blocks in the output are statistically close to uniformly random.

5 Improving the constructions of [BKS+05, Raz05]

The constructions of [BKS+05, Raz05] work by first converting the input sources to a convex combination of two or more independent somewhere-random sources, where each somewhere-random source has a constant number of blocks. At this point they take a constant sized section of each of the sources and do a brute force search for an optimal independent sources extractor. In this way they obtain a constant number (you can actually get say $\log\log\log n$) random bits with large error (say $1/\log\log n$).

Instead of doing brute force search in these construction, we can apply our techniques to get almost all the random bits out of the somewhere random sources, with exponentially small error.

Applying this to the constructions from [BKS+05], we get the following results:

- We get an extractor for 3 sources that can extract a linear fraction of the min-entropy with exponentially small error when the sources have min-entropy rate any small constant.
- We get a 2-source disperser that outputs a linear number of bits with exponentially small error when the min-entropy rate of the sources is an arbitrarily small constant.

Applying our techniques to [Raz05] we can get this result:

- We get an extractor for 3 sources where one source needs to have constant min-entropy rate and the other two may have $k = \text{polylog}(n)$ min-entropy, s.t. the extractor outputs $\Omega(k)$ bits with error exponentially small in k.

Another way to compose our techniques with previous work to get something new is the following, due to Avi Wigderson.

Theorem 5.1. *There exists a polynomial time computable function* $\mathsf{Ext} : (\{0,1\}^n)^3 \to \{0,1\}^{\Omega(n^\delta)}$ *s.t. for any small constant* $0 < \delta \le 1$ *there exists a constant* $0 < \alpha = \alpha(\delta) < 1$ *so that if* X^1 *is an* $(n, n^{1-\alpha})$ *source and* X^2, X^3 *are* (n, n^δ) *sources, with all sources independent of each other,* $\mathsf{Ext}(X^1, X^2, X^3)$ *is* ϵ*-close to the uniform distribution with with* $\epsilon < 2^{-n^{\Omega(1)}}$.

Proof Sketch. If α is small enough, we can use the condenser construction of [BKS$^+$05] to convert the first source to a source with n^γ blocks, so that the source has somewhere-min-entropy rate 0.9. We now interpret this source as n^γ candidate 0.9-min-entropy rate seeds. We use these seeds with Raz's strong extractor construction [Raz05] and the other two sources to obtain two sources which, conditioned on the seeds, are statistically close to independent aligned $(n^\gamma \times n^\delta)$ somewhere random sources. As long as $\delta > \gamma$, we can then use our techniques to get $\Omega(n^\delta)$ bits which are exponentially close to uniformly distributed. ■

In this way we obtain an extractor that can extract from just 3 sources which need have only polynomial min-entropy (the polynomial cannot be arbitrarily small).

6 Acknowledgments

I would like to thank Amnon Ta-Shma for pointing out that the independent sources construction from $O(\log n)$ sources works even when the min-entropy of the sources is as small as $\text{polylog}(n)$ and Avi Wigderson for his observations on obtaining a 3 source extractor: Theorem 5.1. Thanks to Vladimir Trifonov, Salil Vadhan, Adam Klivans and David Zuckerman for useful comments and discussions.

References

[BIW04] B. Barak, R. Impagliazzo, and A. 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] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. 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.

[Blu84] M. Blum. Independent unbiased coin flips from a correlated biased source: a finite state Markov chain. In *Proceedings of the 25th Annual IEEE Symposium on Foundations of Computer Science*, pages 425–433. IEEE, 1984.

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

[BKT04] J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *Geometric and Functional Analysis*, 14:27–57, 2004.

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

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

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

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

[GR05] A. Gabizon and R. Raz. Deterministic extractors for affine sources over large fields. In *To appear*, 2005.

[GRS04] A. Gabizon, R. Raz, and R. Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, 2004.

[KZ03] J. Kamp and D. Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 92–101, 2003.

[Kon03] S. Konyagin. A sum-product estimate in fields of prime order. Technical report, Arxiv, 2003. `http://arxiv.org/abs/math.NT/0304217`.

[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. *Journal of Computer and System Sciences*, 65:660–671, 2002.

[Nis96] N. Nisan. Extracting randomness: How and why – a survey. In *Proceedings of the 11th Annual IEEE Conference on Computational Complexity*, pages 44–58, 1996.

[NT99] N. Nisan and A. Ta-Shma. Extracting randomness: A survey and new constructions. *Journal of Computer and System Sciences*, 58:148–173, 1999.

[NZ93] N. Nisan and D. Zuckerman. More deterministic simulation in logspace. In *Proceedings of the 25th Annual ACM Symposium on Theory of Computing*, pages 235–244, 1993.

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

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

[RSW00] O. Reingold, R. Shaltiel, and A. Wigderson. Extracting randomness via repeated condensing. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, pages 22–31, 2000.

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

[Sha02] R. Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the European Association for Theoretical Computer Science*, 77:67–95, 2002.

[TS96] A. Ta-Shma. Refining randomness. In *ECCCTH: Electronic Colloquium on Computational Complexity, theses*, 1996.

[TSUZ01] 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.

[TV00] L. Trevisan and S. Vadhan. Extracting randomness from samplable distributions. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, pages 32–42, 2000.

[Vaz85] U. Vazirani. Towards a strong communication complexity theory or generating quasi-random sequences from two communicating slightly-random sources (extended abstract). In *Proceedings of the 17th Annual ACM Symposium on Theory of Computing*, pages 366–378, 1985.

[vN51] J. von Neumann. Various techniques used in connection with random digits. *Applied Math Series*, 12:36–38, 1951. Notes by G.E. Forsythe, National Bureau of Standards. Reprinted in *Von Neumann's Collected Works*, 5:768-770, 1963.

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

[Zuc90] D. Zuckerman. General weak random sources. In *Proceedings of the 31st Annual IEEE Symposium on Foundations of Computer Science*, pages 534–543, 1990.

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

[Zuc05] D. Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. Technical Report TR05-100, ECCC: Electronic Colloquium on Computational Complexity, 2005.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'