Deterministic Extractors For Small-Space Sources

Jesse Kamp [*]
Department of Computer Science
University of Texas
Austin, TX 78712
kamp@cs.utexas.edu

Anup Rao [†]
Department of Computer Science
University of Texas
Austin, TX 78712
arao@cs.utexas.edu

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

David Zuckerman [§]
Department of Computer Science
University of Texas
Austin, TX 78712
diz@cs.utexas.edu

ABSTRACT

We give polynomial-time, deterministic randomness extractors for sources generated in small space, where we model space s sources on $\{0,1\}^n$ as sources generated by width 2^s branching programs: For every constant $\delta > 0$, we can extract $.99\delta n$ bits that are exponentially close to uniform (in variation distance) from space s sources of min-entropy δn, where $s = \Omega(n)$. In addition, assuming an efficient deterministic algorithm for finding large primes, there is a constant $\eta > 0$ such that for any $\zeta > n^{-\eta}$, we can extract $m = (\delta - \zeta)n$ bits that are exponentially close to uniform from space s sources with min-entropy δn, where $s = \Omega(\beta^3 n)$. Previously, nothing was known for $\delta \leq 1/2$, even for space 0.

Our results are obtained by a reduction to a new class of sources that we call *independent-symbol* sources, which generalize both the well-studied models of independent sources and symbol-fixing sources. These sources consist of a string of n independent symbols over a d symbol alphabet with min-entropy k. We give deterministic extractors for such sources when k is as small as $\mathsf{polylog}(n)$, for small enough d.

Categories and Subject Descriptors

G.3 [**Probability and Statistics**]: Random Number Generation; G.2 [**Discrete Mathematics**]: Combinatorics, Graph Theory; F.2 [**Analysis of Algorithms and Problem Complexity**]: General

General Terms

Theory, Algorithms

Keywords

Randomness Extractors, Pseudorandomness

1. INTRODUCTION

True randomness is needed for many applications, yet most physical sources of randomness are not truly random, and in fact seem quite weak in that they can have substantial biases and correlations. Weak random sources can also arise in cryptography when an adversary learns some partial information about a random string. A natural approach to dealing with weak random sources is to apply an *extractor* — a function that transforms a weak random source into an almost-perfect random source. For example, Intel's random number generator (cf., [16]) uses the extractor of von Neumann [34] as one of its components.

There was a significant body of work in the 80's focused on this problem of randomness extraction, with researchers considering richer and richer models of weak sources, e.g. [5, 26, 10, 32, 9, 4, 3, 20]. But attempts to handle sources that do not have a significant amount of independence ran into strong impossibility results showing that it is impossible to devise a single function that extracts even one bit of randomness from sufficiently general classes of sources [26].

These impossibility results led researchers to focus on the weaker task of simulating probabilistic algorithms with weak random sources [33, 10, 31, 11, 36]. This line of work culminated in the introduction, by Nisan and Zuckerman [23], of the notion of a *seeded* extractor, which uses a small number of additional *truly random* bits, known as the *seed*, as a catalyst for the randomness extraction. When simulating probabilistic algorithms with weak random sources, the need

[*] Supported in part by NSF Grant CCR-0310960.

[†] Supported in part by NSF Grant CCR-0310960 and an MCD Fellowship.

[‡] Supported by NSF grant CCF-0133096, ONR grant N00014-04-1-0478, and US-Israel BSF grant 2002246.

[§] Supported in part by a David and Lucile Packard Fellowship for Science and Engineering, NSF Grant CCR-0310960, a Radcliffe Institute Fellowship, and a Guggenheim Fellowship.

STOC'06, May 21–23, 2006, Seattle, Washington, USA.
Copyright 2006 ACM 1-59593-134-1/06/0005 ...$5.00.

for truly random bits can be eliminated by enumerating over all choices of the seed. Seeded extractors have turned out to have a wide variety of other applications and were found to be closely related to many other important pseudorandom objects. Thus, they were the main focus of attention in the area of randomness extraction in the 90's, with a variety of very efficient constructions. (See [22, 27] for surveys.)

In the last few years, however, there has been a resurgence of interest in the original concept of a "seedless" (or deterministic) extractor, cf. [30, 13]. This is motivated in part by the realization that seeded extractors do not seem suitable for many settings where we need randomness, such as cryptography. In addition, seedless extractors for specific classes of sources were found to be useful in mitigating partial key exposure in cryptography [8, 13]. Recent attention on seedless extractors has focused on several classes of sources, the main ones being *independent sources,* which consist of several independent parts, each of which has some randomness [10, 1, 2, 25, 24]; *bit-fixing sources*, where some of the bits are perfectly random and the rest are fixed [9, 11, 17, 15]; and *samplable sources*, where the source is generated by an efficient algorithm [30]. Our work relates to all of these models; indeed, we establish connections between them. But our main motivation is a form of samplable sources — namely ones generated by algorithms that have small space.

Before proceeding, we recall a few standard definitions: the *min-entropy* k of a source X is defined as $H_\infty(X) = \min_s(\log(1/\Pr[X = s]))$. (Here and throughout, all logarithms are base 2 unless otherwise specified.) The *min-entropy rate* δ for a source on $[d]^n$ is defined as $\delta = H_\infty(X)/n \log d$. The *variation distance* between random variables X_1 and X_2 on Ω is defined as $|X_1 - X_2| = \max_{S \subseteq \Omega} |\Pr[X_1 \in S] - \Pr[X_2 \in S]| = \frac{1}{2}\sum_{s \in \Omega} |\Pr[X_1 = s] - \Pr[X_2 = s]|$. A function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ is an *ϵ-extractor* for a class of random sources $\mathcal{X}$, if for every $X \in \mathcal{X}$, $\mathsf{Ext}(X)$ is ϵ-close to uniform in variation distance.

1.1 Small-Space Sources

Trevisan and Vadhan [30] proposed the study of extraction from weak random sources that are generated by a process that has a bounded amount of computational resources. This seems to be a plausible model for physical random sources and generalizes a number of the previously studied models. They focused on the case that the source is sampled by either a small circuit or an algorithm with a limited running time. Their main result is a construction of polynomial-time extractors for such sources based on some strong but plausible complexity assumptions. It would be nice to have unconditional constructions (as well as ones that are more efficient and have better error). But they showed that complexity assumptions are needed for the original model of sources generated by a time-bounded algorithms. Thus, they suggested, as a research direction, that we might be able to construct unconditional extractors for sources generated by *space-bounded* algorithms. This model is our focus.

Small space sources are very general in that most other classes of sources that have been considered previously can be computed with a small amount of space. This includes von Neumann's model of a coin with unknown bias [34], Blum's finite Markov chain model [5], symbol-fixing sources [17], and sources that consist of many independent and shorter sources. Strong results in this last model will not follow directly from strong results in the small-space model, but our results do generalize, for example, the results of [1]. In fact, the only model for which deterministic extractors have been given that appears unrelated to our model is "affine sources". Yet despite the small-space model being so natural, very little in the way of explicit constructions for such sources was known.

Table 1: Small space extractors for sources on $\{0,1\}^n$ that extract 99% of the min-entropy. In this table c and C represent sufficiently small and large constants, respectively.

Reference	Min-entropy Rate	Space	Error
Thm 1.1	$\delta = n^{-c}$	$c\delta^3 n$	$\exp(-n^c)$
Thm 1.3	Any constant δ	cn	$\exp(-\tilde{\Omega}(n))$
Thm 1.4	$\delta = C/\log n$	$c\delta \log n$	$\exp(-n^{.99})$

The first example of an explicit construction was due to Blum [5], who showed how to extract from sources generated by a finite Markov chain with a constant number of states. His results generalized the earlier results of von Neumann [34] for extracting from an independent coin with unknown bias. However, the finite Markov chain model is very restricted; it has a constant-size description and the transitions must be the same at each time step.

The exact model for small-space sources we consider is similar to the one previously considered by Koenig and Maurer [18, 19]. It is a generalization of the Markov chain model to time-dependent Markov chains, which yields a much richer class of sources. Our model of a space s source is basically a source generated by a width 2^s branching program. The exact model we consider is that at each step the process generating the source is in one of 2^s states. This can be modelled by a layered graph with each layer corresponding to a single time-step and consisting of vertices corresponding to each of the states. From each node v in layer i, the edges leaving v (going to layer $i+1$) are assigned a probability distribution as well as an output bit for each edge. Unlike in Blum's model [5], the transitions can be different at each time-step.

It can be shown using the probabilistic method that there exist extractors even when the space s is almost as large as the min-entropy k, even when the min-entropy is logarithmically small. Our goal is to provide *efficient* and *deterministic* constructions with parameters that come as close to these bounds as possible.

Koenig and Maurer [18, 19] gave the first explicit constructions of extractors for space-bounded sources. Their extractors require the min-entropy rate to be least $1/2$. We do not know of any other constructions for space-bounded sources, even space 0 sources, which are simply sources of independent bits each of which has a different, unknown, bias.

1.1.1 Our Results

For space s sources with min-entropy $k = \delta n$, we have several constructions, all of which are able to extract almost all of the entropy in the source. These extractors are summarized in Table 1. The first extracts whenever $\delta > n^{-\eta}$ for some fixed constant η and extracts almost all of the entropy.

Theorem 1.1. *Assume we can find primes with length in $[r, 2r]$ deterministically in time* $\text{poly}(r)$. *Then there is a constant $\eta > 0$ such that for every $n \in \mathbb{N}$, and $\delta > \zeta > n^{-\eta}$, there is an polynomial-time computable ϵ-extractor* $\mathsf{Ext} : \{0,1\}^n \rightarrow \{0,1\}^m$ *for space s sources with min-entropy rate δ, where $s = \Omega(\zeta^3 n)$, $m = (\delta - \zeta)n$, and $\epsilon = 2^{-n^{\Omega(1)}}$.*

Remark 1.2. *The assumption about finding primes follows from Cramer's conjecture on the density of primes [12].*

We also have constructions that do not depend on the ability to find large primes. Though the parameters of these constructions are mostly subsumed by the previous construction, they are considerably simpler and achieve somewhat better error. For constant min-entropy rate sources, we have a construction that extracts any constant fraction of the entropy.

Theorem 1.3. *For any constants $\delta > \zeta > 0$ and every $n \in \mathbb{N}$, there is a polynomial-time computable ϵ-extractor* $\mathsf{Ext} : \{0,1\}^n \rightarrow \{0,1\}^m$ *for space s sources with min-entropy rate δ, where $s = \Omega(n)$, $m = (\delta - \zeta)n$, and $\epsilon = 2^{-\Omega(n/\log^3 n)}$.*

The last extractor works with min-entropy rate as low as $\delta = \Omega(1/\log n)$ and space as large as $O(\delta \log n)$.

Theorem 1.4. *For every $n \in \mathbb{N}$ and $\delta > \zeta > 28/\log n$ and $s \leq (\zeta \log n)/28$, there is a polynomial-time computable ϵ-extractor* $\mathsf{Ext} : \{0,1\}^n \rightarrow \{0,1\}^m$ *for space s sources with min-entropy rate δ, where $m = (\delta - \zeta)n$ and*
$\epsilon = \exp(-n/(2^{O(s/\zeta)} \cdot \log^5 n))$.

In comparison to the previous results (e.g. [18, 19]) we have reduced the min-entropy required from $n/2$ to $n^{1-\Omega(1)}$ (in Theorem 1.1). However, we are still far from what can be achieved nonconstructively, where we can extract when the min-entropy is logarithmically small. We also have a gap in terms of the space tolerated. Nonconstructively we can get s to be almost the min-entropy δn while our results require s to be smaller than $\delta^3 n$.

In a partial attempt to close the entropy gap for the case of space 1 sources, we also have an extractor that extracts about $\Omega(k^2/n)$ bits from a more restricted model when $k > n^{4.01/5}$. The extra restriction is that the output bit is required to be the same as the state. Details of this extractor will appear in the full version.

1.2 Independent-Symbol Sources

Our extractors for small-space sources are all obtained via a reduction from a new model of sources we introduce called *independent-symbol sources*. The reduction we use is based on that of Koenig and Maurer [18, 19], who used it to generalize extractors for two independent sources. Independent-symbol sources consist of a string of n independent-symbols over an alphabet of size d such that the total min-entropy of the source is at least k. In addition to being a natural model, these sources are a common generalization of two of the main models studied for seedless extraction, namely symbol-fixing sources [9, 17] and independent sources [10, 1], which we proceed to discuss below.

1.2.1 Independent Sources

One of the most well-studied models of sources is that of extracting from a small number of *independent sources*, each of which has a certain amount of min-entropy, a model essentially proposed by Chor and Goldreich [10]. They constructed extractors for two independent sources with entropy rate greater than 1/2. Recently, similar extractors have been obtained for multiple independent sources with any constant and even subconstant entropy rate, but each of these require at least 3 independent sources [1, 2, 25, 24]. This model is appealing because the individual sources can have arbitrary correlations and biases, and it seems plausible that we can ensure independence between a few such sources. However, such extractors require knowing that all of the sources have large entropy. This motivates our generalization of independent sources to independent-symbol sources, where we only require that the *total* min-entropy over all of the symbols (sources) is high. Another difference between what we consider is that the usual independent source model consists of few sources that are long, whereas independent-symbol sources are interesting even if we have many short sources.

1.2.2 Oblivious Bit-Fixing and Symbol-Fixing Sources

Another particular class that has been studied a great deal is that of *bit-fixing sources*, where some subset of the bit-positions in the source are fixed and the rest are chosen uniformly at random. The first extractors for bit-fixing sources extracted perfectly random bits [9, 11] and required the source to have a large number of random positions. Kamp and Zuckerman [17] constructed extractors that worked for sources with a much smaller number of random bits. They also generalized the notion of bit-fixing sources to symbol-fixing sources, where instead of bits the values are taken from a d symbol alphabet. Gabizon, Raz, and Shaltiel [15] gave a construction that converts any extractor for bit-fixing sources into one that extracts almost all of the randomness, which they apply to the extractor from [17].

Independent-symbol sources can be seen as a generalization of symbol-fixing sources. The difference is that instead of each symbol being only fixed or uniformly random, the symbols in independent-symbol sources are allowed to have any distribution as long as the symbols are chosen independently according to those distributions. Naturally, we place a lower bound on the total min-entropy rather than just the number of random positions. Usually, symbol-fixing sources are thought of as having many symbols that come from a small alphabet (e.g. $\{0,1\}$). This restriction is not necessary to the definition, however, and here we consider the full range of parameters, including even the case that we have a constant number of symbols from an exponentially large "alphabet" (e.g. $\{0,1\}^n$).

1.2.3 Our Results

Our extractors for independent-symbol sources are all based on generalizing various techniques from extractors for independent and symbol-fixing sources.

Koenig and Maurer [18, 19] showed how any extractor for two independent sources with certain algebraic properties can be translated into an extractor for many sources where only two of the sources have sufficient entropy. Their result generalizes to extractors for more than two sources. We show that this also yields extractors for independent-symbol sources. In particular, we apply this to extractors for independent sources that follow from the exponential sum estimates of Bourgain, Glibichuk, and Konyagin [7] (see Bour-

Table 2: Independent-symbol extractors for sources on $[d]^n$ that extract 99% of the min-entropy. In this table c and C represent sufficiently small and large constants, respectively.

Reference	Min-entropy Rate	Error
Thm 1.5	$\delta = (n \log d)^{-c}$	$\exp(-(n \log d)^c)$
Thm 1.6	Any constant δ	$\exp(-\tilde{\Omega}(n \log d))$
Thm 1.7	$\delta = C \frac{d \log^{3/2} n}{(n \log d)^{\frac{1}{2}-\gamma}}$	$\exp(-(n \log d)^{2\gamma})$
Thm 1.8	$\delta = (d \log n)^C / n$	$(\delta n \log d)^{-c}$

gain [6]), and thereby obtain extractors for independent-symbol sources of any constant min-entropy rate.

We also show how to use ideas from the work of Rao [24] for extracting from several independent sources to get extractors for independent-symbol sources that extract from sources of min-entropy $n^{1-\tilde{\Omega}(1)}$.

For small alphabet size d, we use ideas from the work of Kamp and Zuckerman [17] about bit-fixing sources to construct extractors for independent-symbol sources with min-entropy k. We can extract many bits when $k > d\sqrt{n \log d}$, and for $k = \Omega(d^2 \log d)$ we can still extract $\Omega(\log k)$ bits. The base extractor simply takes the sum of the symbols modulo p for some $p > d$, similar to the cycle walk extractor in [17]. Using this extractor we can extract $\Omega(\log k)$ bits. To extract more bits when k is sufficiently large, we divide the source into blocks, apply the base extractor to each block, and then use the result to take a random walk on an expander as in [17].

Unlike the first two extractors, the extractors obtained using this technique use the full generality of the independent-symbol sources. In the first two constructions, using a Markov argument we can essentially first reduce the independent-symbol sources into sources where some of the input symbols have sufficiently high min-entropy while the rest may or may not have any min-entropy. These reductions also cause some entropy to be lost. In this last construction, however, we benefit even from those symbols that have very little min-entropy. Thus we are able to take advantage of all of the entropy, which helps us extract from smaller values of k.

We also show how to generalize the construction of Gabizon et al. [15] to independent-symbol sources to enable us to extract more of the entropy. Note that we use it to improve not only the extractors based on [17] (analogous to what was done in [15] for bit-fixing sources), but also our extractors based on techniques developed for independent sources. Independently of our work, Shaltiel [28] has recently generalized the ideas in [15] to give a framework for constructing deterministic extractors which extract almost all of the entropy from extractors which extract fewer bits. Our extractor can be seen to fit inside this framework.

Applying the technique based on [15] to our extractors based on the independent sources techniques of Rao [24], the results of [7], and the bit-fixing source extractor from [17], respectively, we get the following three theorems. These theorems are directly used to obtain the small-space extractors from Theorem 1.1, Theorem 1.3, and Theorem 1.4. Table 2 presents a summary of these extractors.

THEOREM 1.5. *Assuming we can find primes with length in $[r, 2r]$ deterministically in time $\mathrm{poly}(r)$, there is a constant η such that for every $n, \ell \in \mathbb{N}$ and $\delta > \zeta > (n\ell)^{-\eta}$, there is a polynomial-time computable ϵ-extractor for min-entropy rate $\delta > \zeta$ independent-symbol sources* $\mathsf{Ext} : (\{0,1\}^\ell)^n \to \{0,1\}^m$ *where $m = (\delta - \zeta)n\ell$ and $\epsilon = \exp(-(n\ell)^{\Omega(1)})$.*

We note that in the independent sources model this extractor gives comparable results to the extractor from [1] as a corollary.

The following extractor extracts a constant fraction of the entropy from any constant rate source.

THEOREM 1.6. *For any constants $\delta > \zeta > 0$ and every $n \in \mathbb{N}$, there is a polynomial-time computable ϵ-extractor for min-entropy rate δ independent-symbol sources* $\mathsf{Ext} : [d]^n \to \{0,1\}^m$ *where $m = (\delta - \zeta)n \log d$ and $\epsilon = 2^{-\Omega((n \log d)/\log^3(n \log d))}$.*

For the following extractor we can take $\zeta = \tilde{O}(1/\sqrt{n})$ and can then extract randomness from sources with min-entropy rate as small as $\delta = \tilde{O}(1/\sqrt{n})$.

THEOREM 1.7. *For every $n \in \mathbb{N}$, $2 \leq d \leq \sqrt{n}$ and $\zeta > \sqrt{d^2 \log^3 n / n \log d}$ there is a polynomial-time computable ϵ-extractor for min-entropy rate $\delta > \zeta$ independent-symbol sources* $\mathsf{Ext} : [d]^n \to \{0,1\}^m$ *where $m = (\delta - \zeta)n \log d$ and $\epsilon = \exp(-\Omega((\zeta^2 n \log d)/(d^2 \log^3 n)))$ and .*

Gabizon et al. also give a technique which improves extractors that only extract $\Omega(\log k)$ bits. We show that this technique also generalizes to independent-symbol sources, so we use it together with our extractor based on ideas from [17] that extracts $\Omega(\log k)$ bits to get the following theorem. This theorem shows that even for polylogarithmic k, for small enough d we can extract almost all of the min-entropy.

THEOREM 1.8. *There exists a constant $C > 0$ such that for every $n \in \mathbb{N}$, $d \geq 2$, $k \geq (d \log n)^C$, there exists a polynomial-time computable ϵ-extractor* $\mathsf{Ext} : [d]^n \to \{0,1\}^m$ *for independent symbol sources with min-entropy k, where $m = k - k^{1-\Omega(1)}$ and $\epsilon = k^{-\Omega(1)}$.*

Using the probabilistic method, it can be shown that there exist (nonconstructive) extractors that extract even when the min-entropy k is as small as $O(\log d + \log n)$. Note that we always need $k > \log d$, since otherwise all of the entropy could be in a single symbol, and thus extraction would be impossible. This last extractor comes closest to meeting this bound on k, but only works for small d and has suboptimal error, so there is still much room for improvement.

1.3 Organization

In Section 3 we describe our reduction from small-space sources to independent-symbol sources. The rest of the paper is then focused on extracting from independent-symbol sources. Only the basic extractors for these sources are described here. To get the extractors described in the introduction that extract almost all of the entropy, we also require the generalization of the techniques of Gabizon et al. [15], which are deferred to the full version. In Section 4 we describe the extractor that provides the basis for the extractor from Theorem 1.6. In Section 5 we describe the extractor that provides the basis for the extractor from Theorem 1.5. In Section 6 we describe the extractors that provide the basis for the extractors from Theorem 1.7 and Theorem 1.8.

2. PRELIMINARIES

Notation: We use $[d]$ to denote the set $\{1, \ldots, d\}$. Given a string $x \in [d]^n$ and a set $S \subseteq [n]$ we use x_S to denote the string obtained by restricting x to the indices in S. We use $\circ$ to denote concatenation.

2.1 Classes of Sources

We formally define the various classes of sources we will study.

Definition 2.1. *A* space s source X *on* $\{0,1\}^n$ *is a source generated by a width* 2^s *branching program. That is, the branching program is viewed as a layered graph with* $n+1$ *layers with a single start vertex in the first layer and* 2^s *vertices in each subsequent layer. Each edge is labeled with a probability and a bit value. From a single vertex we can have multiple edges corresponding to the same output bit. The source is generated by taking a random walk starting from the start vertex and outputting the bit values on every edge.*

This definition is very similar to the general Markov sources studied by Koenig and Maurer [18, 19].

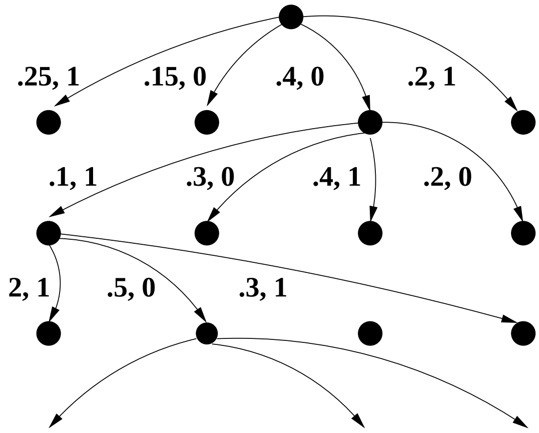

Figure 1: Part of a space 2 source

The other important class of sources we study are independent-symbol sources.

Definition 2.2. *A source* X *on* $[d]^n$ *is an* independent-symbol source *if the* n *symbols are independent. We call such a source* flat *if every symbol is also uniformly distributed over a non-empty subset of* $[d]$.

Note that when $d = 2$, flat independent-symbol sources are the same as oblivious bit-fixing sources.

Definition 2.3. *Let* X *be a random variable taking values in* $\{0,1\}^{t \times r}$*, viewed as* $t \times r$ *matrices with entries in* $\{0,1\}$*. We say that* X *on* $(\{0,1\}^r)^t$ *is* $(t \times r)$ somewhere-random [1] *(*SR-source *for short) if it is a random variable on* t *rows of* r *bits each such that* X *is distributed uniformly randomly over one of the rows. Every other row may depend on the random row in arbitrary ways. We will say that a collection* $X_1, \ldots, X_m$ *of* $(t \times r)$ *SR-sources is* aligned *if there is some* i *for which the* i*'th row of each* X_k *is uniformly distributed.*

We will also need a relaxed notion of the previous definition to where the "random" row is not completely random, but only has some min-entropy.

Definition 2.4. *We say that a* $(t \times r)$ *source* X *on* $(\{0,1\}^r)^t$ *has* somewhere-min-entropy k*, if* X *has min-entropy* k *in one of its* t *rows.*

3. SMALL-SPACE SOURCES AS CONVEX COMBINATIONS OF INDEPENDENT-SYMBOL SOURCES

Here we show how small-space sources can be converted into convex combinations of independent-symbol sources. Thus we will be able to use our extractor constructions from subsequent sections to extract from small-space sources. The idea is simple: to extract from a space s source X, we divide the n bits in X into n/t blocks of size t and view each block as a binary number. Viewed this way, X is a symbol source with alphabet size 2^t. If we condition on the states of the source at the start of each block, all of the blocks become independent, so we end up with an independent-symbol source. We show, using techniques similar to Koenig and Maurer [18, 19], that with high probability these sources will have sufficient min-entropy. To show this we use the following standard lemma (for a direct proof see Lemma 5 in Maurer and Wolf [21], although it has been used implicitly earlier in, e.g., [35]).

Lemma 3.1. *Let* X *and* Y *be random variables and let* $\mathcal{Y}$ *denote the range of* Y*. Then for all* $\epsilon > 0$

$$\Pr_Y \left[H_\infty(X|Y=y) \geq H_\infty(X) - \log|\mathcal{Y}| - \log\left(\frac{1}{\epsilon}\right) \right] \geq 1 - \epsilon$$

This basically gives us the result we want. All of our extractors for small-space sources are obtained by combining the following lemma with the corresponding extractor for independent-symbol sources. We note that the reduction in this lemma is only interesting when the min-entropy rate $\delta > 1/\sqrt{n}$, since otherwise the entropy of the independent-symbol source would be less than the length of an individual symbol. In this case all of the entropy could be in a single symbol and thus extraction would be impossible.

Lemma 3.2. *Let* X *be a space* s *source on* $\{0,1\}^n$ *with min-entropy rate* δ*. Then for any* $0 < \alpha < 1$*,* X *is* $2^{-\alpha\delta n/2}$*-close to a convex combination of independent-symbol sources on* $[d']^{n'}$ *with min-entropy rate* δ'*, where* $d' = 2^{2s/(\alpha\delta)}$*,* $n' = \alpha\delta n/2s$ *and* $\delta' = (1-\alpha)\delta$*.*

Proof. Divide X into $\alpha\delta n/2s$ blocks of size $2s/\alpha\delta$. Let Y represent the values of the initial states for each block. Then each $(X|Y=y)$ is an independent-symbol source with each block viewed as a number less than $2^{2s/(\alpha\delta)}$ in binary. By Lemma 3.1, since $|\mathcal{Y}| = (2^s)^{(\alpha\delta n)/(2s)} = 2^{\alpha\delta n/2}$, with probability $1 - 2^{-\alpha\delta n/2}$ the sources $(X|Y=y)$ have min-entropy $(1-\alpha)\delta n$ and thus min-entropy rate $(1-\alpha)\delta$. □

[1] This definition is slightly different from the original one used by Ta-Shma [29]. The original definition considered the closure under convex combinations of the class defined here (i.e. convex combinations of sources that have one random row). We use this definition because we can do so without loss of generality and it considerably simplifies the presentation.

4. EXTRACTING FROM INDEPENDENT-SYMBOL SOURCES BY REDUCING TO INDEPENDENT SOURCES

In this section, we will show how to construct extractors for independent-symbol sources using techniques from independent sources.

The following Markov-like lemma will be used to show that if we divide a source into blocks, many of the blocks will have a large entropy rate.

LEMMA 4.1. *For any partition of an independent-symbol source X on $[d]^n$ with min-entropy δ into t blocks of size n/t, the number r of blocks with min-entropy rate greater than $\delta/2$ satisfies $r > \delta t/2$.*

Therefore we can view this source as an independent-symbol source on $[d^{n/t}]^t$ where at least $\delta t/2$ of the symbols have min-entropy rate greater than $\delta/2$.

PROOF. We know that r blocks have min-entropy rate greater than $\delta/2$ and at most 1. In each of the remaining blocks the min-entropy rate is at most $\delta/2$. Since the total entropy rate is δ and min-entropies add for independent-symbol sources, $\delta \leq (r + (t-r)(\delta/2))/t$, which after a simple calculation gives the desired result. □

Once we are in this model, we can generalize the result from Koenig and Maurer [18, 19] that states that any two source extractor of the form $f(x_1 \cdot x_2)$, where the x_i are elements of some group, can be extended to any number of sources where only two of the sources have sufficient min-entropy.

LEMMA 4.2. *Let $(\mathcal{G}, *)$ be a group and let $Ext(x_1, x_2, \ldots, x_r) := f(x_1 * x_2 \cdots * x_r)$ be an extractor for r independent sources over $\mathcal{G}$, each of which has min-entropy rate at least δ. Then $F(x_1, \ldots, x_n) := f(x_1 * \cdots * x_n)$ is an extractor for n independent sources over $\mathcal{G}$, r of which have min-entropy rate at least δ.*

The proof is essentially the same as in [18, 19]. The key idea is that the n sources can be divided into r blocks, each of which contains exactly one of the high entropy sources.

Bourgain, Glibichuk, and Konyagin [7] gave bounds on the exponential sums of the function $f(x_1, \ldots, x_r) = \prod_{i=1}^r x_i$ over large subsets of fields without large subfields, in particular $GF(p)$ and $GF(2^p)$. As observed by Bourgain in [6], this estimate gives an extractor for independent sources. Bourgain only explicitly gives an extractor that outputs a single bit, but his result can be easily generalized using his techniques together with Vazirani's XOR lemma [31] to get the following.

THEOREM 4.3. *[7] Let the finite field K be either $GF(p)$ or $GF(2^p)$ for some prime p. Let $f(x_1, \ldots, x_r) = \prod_{i=1}^r x_i$ and view the output of the function as an integer from 0 to $|K|-1$. Then there exist functions $C_1(\delta)$ and $C_2(\delta)$ such that the function $\mathsf{BGK}(x_1, \ldots, x_r) = \lfloor (2^m f(x_1, \ldots, x_r))/|K| \rfloor$ (i.e. taking the m most significant bits of $f(x_1, \ldots, x_r)/|K|$) is an ϵ-extractor for r independent min-entropy rate δ sources over K for $r \geq C_1(\delta)$, $m = \Theta(C_2(\delta) \log |K|)$, and $\epsilon = 2^{-\Omega(m)}$.*

Note that for constant δ, we can extract $\Theta(\log |K|)$ bits with only a constant number of sources. For $GF(p)$, [7] make explicit the relationship between δ and the number of sources and entropy. It turns out in this case that we can handle slightly subconstant δ, down to $\delta = \Omega(1/(\log \log |K|)^{(1/C)})$ for some constant C. For $GF(2^p)$, it's not clear whether or not a similar result can be achieved.

Combining this theorem with Lemma 4.2, restricting the sources to be over the multiplicative group K^*, we get the following corollary.

COROLLARY 4.4. *Let the finite field K be either $GF(p)$ or $GF(2^p)$ for some prime p. Let $f(x_1, \ldots, x_n) = \prod_{i=1}^n x_i$ and view the output of the function as a number from 0 to $|K|-1$. Then there exist functions $C_1(\delta)$ and $C_2(\delta)$ such that the function $\mathsf{BGK}(x_1, \ldots, x_n) = \lfloor (2^m f(x_1, \ldots, x_n))/|K| \rfloor$ is an ϵ-extractor for n independent sources over K^*, at least $C_1(\delta)$ of which have min-entropy rate at least δ, and with $m = \Theta(C_2(\delta) \log |K|)$ and $\epsilon = 2^{-\Omega(m)}$.*

It will also be useful to formulate the following corollary.

COROLLARY 4.5. *For every constant $\delta > 0$, there exists a constant $v(\delta)$ and a polynomial time computable function $\mathsf{BGK} : (\{0,1\}^\ell)^n \to \{0,1\}^m$ that is an ϵ-extractor for n independent sources on $\{0,1\}^\ell$, such that at least $v(\delta)$ of them have min-entropy rate δ where $m = \Omega(\ell)$ and $\epsilon = 2^{-\Omega(\ell)}$.*

Now we can combine this extractor with Lemma 4.1 to get an extractor for independent-symbol sources with constant min-entropy rate.

THEOREM 4.6. *For any constant δ, we can construct a polynomial-time computable ϵ-extractor $\mathsf{Ext} : [d]^n \to \{0,1\}^m$ for rate δ independent-symbol sources with $m = \Theta(n \log d)$ and $\epsilon = 2^{-\Omega(m)}$. This extractor can be computed in time* $\mathrm{poly}(n, \log d)$.

PROOF. Given a source X, divide it into $t = 2C_1(\delta/2)/\delta$ blocks of size n/t, where $C_1(\delta)$ is the constant from Corollary 4.4. Then by Lemma 4.1, we can view X as an independent-symbol source over $[d^{n/t}]^t$, where at least $\delta t/2 = C_1(\delta/2)$ of the symbols have min-entropy rate at least $\delta/2$. Find the smallest prime $p > (n \log d)/t$. By Bertrand's postulate, $p \leq 2(n \log d)/t$, so we can find such a prime in time $\mathrm{poly}(n, \log d)$ by brute force search. Then we can embed each of our symbols over $d^{n/t} < 2^p$ into $GF(2^p)^*$ and apply the extractor from Corollary 4.4 to get the stated result. □

5. EXTRACTING FROM POLYNOMIAL ENTROPY RATE

In this section we will show how to extract from independent-symbol sources when the min-entropy of the sources is polynomially small, using techniques based on those of Rao [24]. As in the previous section, we will reduce the problem to another model: we will try to extract from a few independent sources when just some of them have a polynomial amount of entropy, but we don't know exactly which ones. The probabilistic method shows that extractors exist for this model even when just two sources contain logarithmic min-entropy and the total number of sources is polynomially large. In this section we consider sources over symbols with large alphabet sizes, so we let $d = 2^\ell$ and view the symbols as being over $\{0,1\}^\ell$. Our main theorem is as follows.

Theorem 5.1. *Assuming we can find primes with length in $[r, 2r]$ in time* $\mathrm{poly}(r)$*, there is a constant β such that there exists a polynomial-time computable ϵ-extractor* $\mathsf{Ext} : (\{0,1\}^\ell)^n \rightarrow \{0,1\}^m$ *for independent-symbol sources with min-entropy rate $\delta \geq \ell^{-\beta}$, where $n = \Theta(1/\delta^2)$, $m = \ell^{\Omega(1)}$ and $\epsilon = 2^{-\ell^{\Omega(1)}}$.*

We can also get the following corollary for when we have a larger number of smaller sources.

Corollary 5.2. *Assuming we can find primes with length in $[r, 2r]$ in time* $\mathrm{poly}(r)$*, there exists a constant η such that for any $\delta \geq (n\ell)^{-\eta}$, there exists a polynomial-time computable ϵ-extractor* $\mathsf{Ext} : (\{0,1\}^\ell)^n \rightarrow \{0,1\}^m$ *for independent-symbol sources with min-entropy rate δ, where $m = (\delta^2 n\ell)^{\Omega(1)}$ and $\epsilon = 2^{-(\delta^2 n\ell)^{\Omega(1)}}$.*

Proof. Divide the source into $\Theta(1/\delta^2)$ blocks of $\Theta(\delta^2 n)$ symbols each and apply Theorem 5.1. □

In this section we will describe a generic technique to turn any extractor for the model where a few sources have min-entropy rate less than half into an extractor that can extract when the min-entropy is as small as $\ell^{1-\alpha_0}$ for some universal constant α_0. There are two major ingredients that will go into our construction:

- The first ingredient is recent constructions of randomness efficient condensers [2, 25]. As observed by Avi Wigderson (the theorem appears in [24]), these condensers imply that there are small constants α, γ and a polynomial time computable function that can convert any source on ℓ bits with min-entropy $\ell^{1-\alpha}$ into a source that outputs ℓ^γ different rows of bits, such that most of them come from sources with extremely high min-entropy rate. An important property that we will need is that the length of each of the rows is $\ell^{2\gamma}$, which is much higher than the number of rows. Formally:

  Theorem 5.3. *(By Wigderson [24]) Assuming we can find primes with length in $[r, 2r]$ in time* $\mathrm{poly}(r)$*, for every sufficiently small constant $\gamma > 0$ there exists constants $\alpha = \alpha(\gamma) > 0$ and $\mu(\gamma) > 2\gamma$ and a polynomial time computable function* $\mathsf{Cond} : \{0,1\}^\ell \rightarrow (\{0,1\}^{\ell^\mu})^{\ell^\gamma}$ *such that for any min-entropy $\ell^{1-\alpha}$ source X,* $\mathsf{Cond}(X)$ *is ϵ-close to a source with somewhere min-entropy rate 0.9.*

  If we use the condenser from Raz's work [25] with the improved analysis of Dvir and Raz (Lemma 3.2 in [14]), to get the theorem above, we can even ensure that most of the output rows are statistically close to having high min-entropy. It will be more convenient to think of this as a seeded condenser. The analysis of Dvir and Raz ensures that we obtain a seeded condenser which for all but an arbitrarily small constant fraction of the seeds, succeeds in condensing the input source to give a distribution that is exponentially close to having high min-entropy.

  Theorem 5.4. *Assuming we can find primes with length in $[r, 2r]$ in time* $\mathrm{poly}(r)$*, for all small enough constants $\gamma, \epsilon > 0$ there exist constants $\alpha = \alpha(\gamma), \mu(\gamma) > 2\gamma, \beta(\gamma, \epsilon) > 0$, and a polynomial time computable function* $\mathsf{Cond} : \{0,1\}^\ell \times \{0,1\}^a \rightarrow (\{0,1\}^{\ell^\mu})$ *with $a = \gamma \log \ell$ such that for any min-entropy $\ell^{1-\alpha}$ source X,*

  $$\Pr_{s \leftarrow_R U_w}[\mathsf{Cond}(X, s) \text{ is } 2^{-\ell^\beta}\text{-close to a source with min-entropy rate } 0.9] > 1 - \epsilon.$$

- The second ingredient is the technique of condensing independent somewhere random sources from the work of Rao [24]. We will prove a generalization of a theorem from that work. We will show how to extract from independent sources with only a few of them being aligned somewhere random sources that have rows that are much longer than the number of rows. Formally, we get the following, a proof of which is deferred to the full version:

  Theorem 5.5. *For every constant $\gamma < 1$ there exists a polynomial time $2^{-\ell^{\Omega(1)}}$-extractor* $\mathsf{SRExt} : (\{0,1\}^{\ell^{\gamma+1}})^u \rightarrow \{0,1\}^m$ *for u independent sources, of which v are independent aligned $(\ell^\gamma \times \ell)$ SR-sources, where $m = \ell - \ell^{\Omega(1)}$.*

We will first describe how to use these two ingredients to extract from an intermediate model. Then we will see that independent-symbol sources can be easily reduced to this intermediate model to prove Theorem 5.1.

5.1 Extracting From The Intermediate Model

The intermediate model we work with is defined as follows.

Definition 5.6. *A (u, v, α) intermediate source X consists of u^2 sources $X^1, \ldots, X^{u^2}$, each on $\{0,1\}^\ell$. These sources are partitioned into u sets $S_1, \ldots, S_u$ such that v of the sets have the property that v of their sources have min-entropy at least $\ell^{1-\alpha}$.*

Now we show that for certain constant v and $\alpha > 0$ we can extract from this model.

Theorem 5.7. *Assuming we can find primes with length in $[r, 2r]$ in time* $\mathrm{poly}(r)$*, for some constants v and $\alpha > 0$ there exists a polynomial time computable $2^{-\ell^{\Omega(1)}}$-extractor* IExt *for (u, v, α) intermediate sources, where $m = \ell^{\Omega(1)}$.*

Using this theorem together with Lemma 4.1, we can prove Theorem 5.1.

Proof. (Of Theorem 5.1.) Let X be an independent-symbol source on $(\{0,1\}^\ell)^n$ with min-entropy rate $\delta \geq 4\ell^{-\alpha}$, where α is the constant from Theorem 5.7 and $n = u^2$ where u will be chosen later. Divide the source into u blocks with u symbols each. By Lemma 4.1, $\delta u/2$ of the blocks have min-entropy rate at least $\delta/2$. Now further divide each of the blocks into u subblocks of one symbol each. By Lemma 4.1, for the blocks with min-entropy rate at least $\delta/2$, at least $\delta u/4$ of the subblocks have min-entropy rate $\delta/4 \geq \ell^{-\alpha}$. Let $u = 4v/\delta$, where v is the constant from Theorem 5.7. Then X is a (u, v, α) intermediate source satisfying the conditions of Theorem 5.7, which immediately gives us the theorem. □

Here is the algorithm promised by Theorem 5.7:

Construction: $\mathsf{IExt}(x^1, \ldots, x^{u^2})$

```
Input: $x^1, \ldots, x^{u^2}$ partitioned into sets $S_1, \ldots, S_u$
Output: $z$.
Let $v$ be a constant that we will pick later.
Let $\mathsf{BGK}$ be as in Corollary 4.5 - an extractor for independent sources when $v-1$ of them have min-entropy.
Let $\mathsf{Cond}$ be as in Theorem 5.4 - a condenser that converts sources with sublinear min-entropy into sources with somewhere min-entropy rate 0.9.
Let $\mathsf{SRExt}$ be as in Theorem 5.5 - an extractor for independent sources that works when just $v$ of the inputs come from aligned somewhere random sources.
Set $\epsilon = 1/v^3$. Let $\gamma$ be a small enough constant to apply Theorem 5.4 with $\gamma, \epsilon$ in the hypothesis. Let $\alpha, a$ be as in the conclusion of the theorem.
Let $\{0,1\}^a = \{s_1, s_2, \ldots, s_{2^a}\}$.
  1. For every seed $s_i$:
     (a) Run $\mathsf{Cond}(., s_i)$ on every source in the input to get supposedly condensed strings $a^1, \ldots, a^{u^2}$. Each $a^j$ is of length $\ell^\mu \geq \ell^{2\gamma}$.
     (b) For every $l \in [u]$, let $b_i^l$ be the string obtained by applying $\mathsf{BGK}$ to the $a^j$'s from $S_l$.
     We think of $b^l$ as a sample from a somewhere random source with $\ell^\gamma$ rows, one for each seed $s_i$.
  2. Output $\mathsf{SRExt}(b^1, \ldots, b^u)$.
```

Proof of Theorem 5.7. If we used a completely random seed s_i in the first step, we know that the condensing succeeds for a single source with probability $(1-\epsilon)$. By the union bound, the condensing succeeds for all the v^2 high min-entropy sources with probability $(1 - v^2\epsilon)$. This quantity is greater than 0, so there must be some seed s_i for which the condensing succeeds for all the v^2 high min-entropy sources. When this s_i is used with any of the sources that contain sufficient entropy X^j, $\mathsf{Cond}(X^j, s_i)$ is $2^{-\ell^\beta}$-close to a source with min-entropy rate 0.9, where β is as in Theorem 5.4. Then BGK succeeds in extracting from the row corresponding to this s_i.

Thus the result of the first step in the algorithm is a distribution that is $2^{-\ell^{\Omega(1)}}$-close to a collection of u independent sources, v of which are independent aligned somewhere-random sources.

This is exactly the kind of distribution that our extractor SRExt can handle, so our algorithm succeeds in extracting from the input. □

6. EXTRACTING FROM SMALL ALPHABET INDEPENDENT-SYMBOL SOURCES

Now we show how for smaller alphabet sizes d we can do better than the constructions in the previous sections by generalizing previous constructions for symbol-fixing sources. The base extractor simply takes the sum of the symbols modulo p for some $p > d$. Then we divide the source into blocks, apply the base extractor to each block, and then use the result to take a random walk on an expander as in [17].

6.1 Reducing to Flat Independent-Symbol Sources

It will be simpler to analyze our extractor for flat independent-symbol sources. We show that any extractor that works for flat independent-symbol sources also works for general independent-symbol sources because any independent-symbol source is close to a convex combination of flat sources, a proof of which is deferred to the full version. In particular, since for $d = 2$ flat independent-symbol sources are the same as bit-fixing sources, this lemma shows that any extractor for bit-fixing sources is also an extractor for independent-symbol sources with $d = 2$.

Lemma 6.1. *Any ϵ-extractor for the set of flat independent-symbol sources on $[d]^n$ with min-entropy $k/(2\log 3)$ is also an $(\epsilon + e^{-k/9})$-extractor for the set of independent-symbol sources on $[d]^n$ with min-entropy k.*

6.2 Extracting From Flat Independent-Symbol Sources

Now we show how to extract from flat independent-symbol sources for small d. Our initial extractor simply takes the sum modulo p of the symbols.

Theorem 6.2. *Let $d \geq 2$ and $p \geq d$ a prime. Then $Sum_p : [d]^n \to [p]$, where $Sum_p(x) = \sum_i x_i \mod p$, is an ϵ-extractor for the set of flat independent-symbol sources on $[d]^n$ with min-entropy k, where $\epsilon = \frac{1}{2} 2^{-2k/p^2}\sqrt{p}$.*

Combining Theorem 6.2 with Lemma 6.1 we get an extractor for independent-symbol sources.

Theorem 6.3. *Let $d \geq 2$ and $p \geq d$ a prime. Then Sum_p is an ϵ-extractor for the set of independent-symbol sources on $[d]^n$ with min-entropy $k \geq \Omega(p^2 \log p)$, where $\epsilon = 2^{-\Omega(k/p^2)}$.*

We will prove Theorem 6.2 via the following lemma.

Lemma 6.4. *Let $d \geq 2$ and $p \geq d$ a prime. Then for all flat independent symbol sources X on $[d]^n$ with min-entropy k, $Sum_p(X)$ has ℓ_2 distance from uniform at most $2^{-2k/p^2}$.*

It is well known that if X_1 and X_2 are both distributed over a universe of size p, then $|X_1 - X_2| \leq \frac{1}{2}\sqrt{p}||X_1 - X_2||_2$. Theorem 6.2 then follows by combining Lemma 6.4 with this relation between ℓ_2 and variation distance.

To analyze the distance from uniform of the sum modulo p, we use the following lemma that relates this distance to the characters of $\mathbb{Z}_p$. For $\mathbb{Z}_p$, where p is a prime, the ith character is defined as $\chi_j(a) = e^{\frac{2\pi i j a}{p}}$.

Lemma 6.5. *For any function $f : \{0,1\}^n \to \mathbb{Z}_p$ and random variable X over $\{0,1\}^n$,*

$$||f(X) - U_p||_2^2 = \frac{1}{p}\sum_{j=1}^{p-1} |\mathbb{E}[\chi_j(f(X))]|^2 < \max_{j \neq 0} |\mathbb{E}[\chi_j(f(X))]|^2,$$

where U_p denotes the uniform distribution over $\mathbb{Z}_p$.

Proof. Let $Y = f(X) - U_p$. The jth Fourier coefficient of Y is given by $\hat{Y}_j = \sum_{y=0}^{p-1} Y(y)\chi_j(y)$. By Parseval's Identity and using the fact that $\sum_{y=0}^{p-1} \chi_j(y) = 0$ when $j \neq 0$ we get

$$\begin{aligned}
||Y||_2^2 &= \frac{1}{p}\sum_{j=0}^{p-1} |\hat{Y}_j|^2 \\
&= \frac{1}{p}\sum_{j=0}^{p-1} \left|\sum_{y=0}^{p-1} Y(y)\chi_j(y)\right|^2 \\
&= \frac{1}{p}\sum_{j=0}^{p-1} \left|\sum_{y=0}^{p-1} \Pr[f(X) = y]\chi_j(y) - \frac{1}{p}\sum_{y=0}^{p-1}\chi_j(y)\right|^2 \\
&= \frac{1}{p}\sum_{j=1}^{p-1} |\mathbb{E}[\chi_j(f(X))]|^2 \\
&< \max_{j \neq 0} |\mathbb{E}[\chi_j(f(X))]|^2.
\end{aligned}$$

□

Using the previous lemma we can now prove Theorem 6.2.

Proof. Let $f(X) = \sum_{i=1}^n X_i$ and fix $j \neq 0$. Then $|\mathbb{E}[\chi_j(f(X))]|^2 = \prod_{i=1}^n |\mathbb{E}[\chi_j(X_i)]|^2$. Suppose X_i has min-entropy k_i, so $k = \sum_i k_i$. Then since each X_i is a flat source, X_i is uniformly distributed over $K_i = 2^{k_i}$ values. Our goal is to upper bound $|\mathbb{E}[\chi_j(X_i)]|^2$ over all possible choices of X_i. Doing so, we get

$$\begin{aligned}
|\mathbb{E}[\chi_j(X_i)]|^2 &\leq \max_{\substack{X_i : \mathbb{Z}_p \to \{0, 1/K_i\} \\ \sum_x X_i(x) = 1}} |\mathbb{E}[\chi_j(X_i)]|^2 \\
&= \max_{\substack{X_i : \mathbb{Z}_p \to \{0, 1/K_i\} \\ \sum_x X_i(x) = 1}} \left|\sum_{x \in \mathbb{Z}_p} X_i(x)\chi_j(x)\right|^2 \\
&= \max_{y, |y|=1} \left(\max_{\substack{X_i : \mathbb{Z}_p \to \{0, 1/K_i\} \\ \sum_x X_i(x) = 1}} \left(\left(\sum_{x \in \mathbb{Z}_p} X_i(x)\chi_j(x) \right) \odot y \right)^2 \right) \\
&= \max_{\substack{X_i : \mathbb{Z}_p \to \{0, 1/K_i\} \\ \sum_x X_i(x) = 1}} \left(\max_{y, |y|=1} \left(\sum_{x \in \mathbb{Z}_p} X_i(x)(\chi_j(x) \odot y) \right)^2 \right),
\end{aligned}$$

where $\odot$ denotes the complex dot product, where the complex numbers are viewed as two dimensional vectors, and the third line follows from the observation that the dot product is maximized when y is in the same direction as $(\sum_{x \in \mathbb{Z}_p} X_i(x)\chi_j(x))$, in which case we get exactly the square of the length. Now we further note that $\chi_j(x) \odot y$ is greatest for values of x for which $\chi_j(x)$ is closest to y. Thus we achieve the maximum when X_i is distributed over the K_i values closest to y. Without loss of generality we can assume these values correspond to $x = 0$ to $K_i - 1$ (since we only care about the magnitude). Thus

$$\begin{aligned}
|\mathbb{E}[\chi_j(X_i)]|^2 &\leq \left| \frac{1}{K_i} \sum_{\ell=0}^{K_i - 1} e^{\frac{2\pi i \ell}{p}} \right|^2 = \left| \frac{1}{K_i} \frac{1 - e^{\frac{2\pi i j K_i}{p}}}{1 - e^{\frac{2\pi i}{p}}} \right|^2 \\
&= \left| \frac{1}{K_i} \frac{e^{\frac{\pi i K_i}{p}} \left(e^{\frac{-\pi i K_i}{p}} + e^{\frac{\pi i K_i}{p}} \right)}{e^{\frac{\pi i}{p}} \left(e^{\frac{-\pi i}{p}} + e^{\frac{\pi i}{p}} \right)} \right|^2 \\
&= \left(\frac{1}{K_i} \frac{\sin\left(\frac{\pi K_i}{p}\right)}{\sin\left(\frac{\pi}{p}\right)} \right)^2 \\
&= \left(\frac{1}{K_i} \frac{\frac{\pi K_i}{p} \prod_{m=1}^{\infty} \left(1 - \frac{K_i^2}{p^2 m^2}\right)}{\frac{\pi}{p} \prod_{m=1}^{\infty} \left(1 - \frac{1}{p^2 m^2}\right)} \right)^2 \\
&= \left(\prod_{m=1}^{\infty} \left(1 - \frac{K_i^2 - 1}{p^2 m^2 - 1}\right) \right)^2 \\
&< \left(1 - \frac{K_i^2 - 1}{p^2 - 1}\right)^2 < e^{-2(K_i^2 - 1)/(p^2 - 1)},
\end{aligned}$$

where in the fourth line we use the infinite product representation of sine. So

$$\begin{aligned}
|\mathbb{E}[\chi_j(f(X))]|^2 &= \prod_{i=1}^n |\mathbb{E}[\chi_j(X_i)]|^2 \\
&< \prod_{i=1}^n e^{-2(K_i^2 - 1)/(p^2 - 1)} < e^{2n/p^2} e^{-2(\sum_i K_i^2)/p^2}.
\end{aligned}$$

By the power mean inequality, $\sum_{i=1}^n K_i^2 \geq n \cdot (\prod_{i=1}^n K_i)^{2/n} = n 2^{2k/n}$. Thus $|\mathbb{E}[\chi_j(f(X))]|^2 < e^{-2n(2^{2k/n} - 1)/p^2}$. Let $k = \delta n$. Then this quantity is $e^{-(2k/p^2)((2^{2\delta} - 1)/\delta)}$. Since $(2^{2\delta} - 1)/\delta$ is an increasing function of δ and goes to $2 \ln 2$ as δ goes to 0, we have

$$|\mathbb{E}[\chi_j(f(X))]|^2 < e^{-(2k/p^2)((2^{2\delta} - 1)/\delta)} < e^{-4(\ln 2)k/p^2} = 2^{-4\frac{k}{p^2}}$$

Then by Lemma 6.5 $||f(X) - U_p||_2^2 < \max_{j \neq 0} |\mathbb{E}[\chi_j(f(X))]|^2$, so $||f(X) - U_p||_2 < 2^{-2k/p^2}$. □

Now we show that if we divide the source into blocks and take the sum modulo p for each block, we get a convex combination of "almost" symbol-fixing sources, which we can then use an expander walk to extract from, as in [17]. The proof of the following theorem is essentially the same as that in [17] and is deferred to the full version.

Theorem 6.6. *There exists an ϵ-extractor for the set of flat independent symbol sources on $[d]^n$ with min-entropy k that outputs $m = \Omega(k^2/(nd^2 \log d))$ bits and has error $\epsilon = 2^{-m}$. This extractor is computable in time* poly(n, d).

Combining this theorem with our reduction from general to flat sources, we get that this same extractor works for general independent-symbol sources.

Theorem 6.7. *There exists an ϵ-extractor for the set of independent-symbol sources on $[d]^n$ with min-entropy k that outputs $m = \Omega(k^2/nd^2 \log d)$ bits and has error $\epsilon = 2^{-m}$. This extractor is computable in time* poly(n, d).

Proof. Combine Theorem 6.6 and Lemma 6.1. □

7. REFERENCES

[1] B. Barak, R. Impagliazzo, and A. Wigderson. Extracting randomness using few independent sources. In *45th FOCS*, pages 384–393, 2004.

[2] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson. Simulating independence: new constructions of condensers, ramsey graphs, dispersers, and extractors. In *37th STOC*, pages 1–10, 2005.

[3] M. Ben-Or and N. Linial. Collective coin flipping. In S. Micali, editor, *Randomness and Computation*, pages 91–115. Academic Press, New York, 1990.

[4] C. H. Bennett, G. Brassard, and J. Robert. Privacy amplification by public discussion. *SICOMP*, 17(2):210–229, Apr. 1988.

[5] M. Blum. Independent unbiased coin flips from a correlated biased source: a finite Markov chain. *Combinatorica*, 6(2):97–108, 1986.

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

[7] J. Bourgain, A. Glibichuk, and S. Konyagin. Estimates for the number of sums and products and for exponential sums in fields of prime order. *J. London Math. Soc.*, 2005. To appear.

[8] R. Canetti, Y. Dodis, S. Halevi, E. Kushilevitz, and A. Sahai. Exposure-resilient functions and all-or-nothing transforms. In B. Preneel, editor, *Advances in Cryptology — EUROCRYPT 2000*, volume 1807 of *LNCS*, pages 453–469. Springer-Verlag, May 2000.

[9] B. Chor, J. Friedman, O. Goldreich, J. Håstad, S. Rudich, and R. Smolensky. The bit extraction problem or t–resilient functions. In *26th FOCS*, pages 396–407, 1985.

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

[11] A. Cohen and A. Wigderson. Dispersers, deterministic amplification, and weak random sources. In *30th FOCS*, pages 14–19, 1989.

[12] H. Cramer. On the order of magnitude of the difference between consecutive prime numbers. *Acta Arithmetica*, pages 23–46, 1937.

[13] Y. Dodis. Impossibility of black-box reduction from non-adaptively to adaptively secure coin-flipping. Unpublished manuscript, April 2000.

[14] Z. Dvir and R. Raz. Analyzing linear mergers. Technical Report TR05-25, ECCC, 2005.

[15] A. Gabizon, R. Raz, and R. Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. In *45th FOCS*, pages 394–403, 2004.

[16] B. Jun and P. Kocher. The Intel random number generator, 1999. `http://www.cryptography.com/resources/whitepapers/IntelRNG.pdf`.

[17] J. Kamp and D. Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. In *44th FOCS*, pages 92–101, 2003.

[18] R. Koenig and U. Maurer. Extracting randomness from generalized symbol-fixing and markov sources. In *Proceedings of 2004 IEEE International Symposium on Information Theory*, page 232, June 2004.

[19] R. Koenig and U. Maurer. Generalized strong extractors and deterministic privacy amplification. In N. Smart, editor, *Cryptography and Coding 2005*, volume 3796 of *LNCS*, pages 322–339. Springer-Verlag, Dec. 2005.

[20] D. Lichtenstein, N. Linial, and M. Saks. Some extremal problems arising from discrete control processes. *Combinatorica*, 9(3):269–287, 1989.

[21] U. Maurer and S. Wolf. Privacy amplification secure against active adversaries. In B. S. K. Jr., editor, *Advances in Cryptology — CRYPTO '97*, volume 1294 of *LNCS*, pages 307–321. Springer-Verlag, Aug. 1997.

[22] N. Nisan and A. Ta-Shma. Extracting randomness: A survey and new constructions. *JCSS*, 58:148–173, 1999.

[23] N. Nisan and D. Zuckerman. Randomness is linear in space. *JCSS*, 52(1):43–52, 1996.

[24] A. Rao. Extractors for a constant number of polynomially small min-entropy independent sources. In *38th STOC*, 2006.

[25] R. Raz. Extractors with weak random seeds. In *37th STOC*, pages 11–20, 2005.

[26] M. Santha and U. V. Vazirani. Generating quasi-random sequences from semi-random sources. *JCSS*, 33:75–87, 1986.

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

[28] R. Shaltiel. How to get more mileage from randomness extractors. Technical Report TR05-145, ECCC, 2005.

[29] A. Ta-Shma. On extracting randomness from weak random sources. In *28th STOC*, pages 276–285, 1996.

[30] L. Trevisan and S. P. Vadhan. Extracting randomness from samplable distributions. In *41st FOCS*, pages 32–42, 2000.

[31] U. V. Vazirani. *Randomness, Adversaries and Computation*. PhD thesis, EECS, University of California at Berkeley, 1986.

[32] U. V. Vazirani. Strong communication complexity or generating quasirandom sequences from two communicating semirandom sources. *Combinatorica*, 7(4):375–392, 1987.

[33] U. V. Vazirani and V. V. Vazirani. Random polynomial time is equal to slightly-random polynomial time. In *26th FOCS*, pages 417–428, 1985.

[34] J. von Neumann. Various techniques used in connection with random digits. *National Bureau of Standards, Applied Mathematics Series*, 12:36–38, 1951.

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

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