How to get more mileage from randomness extractors

Ronen Shaltiel*
University of Haifa
Haifa 31905, Israel
ronen@cs.haifa.ac.il

Abstract

Let $\mathcal{C}$ be a class of distributions over $\{0,1\}^n$. A deterministic randomness extractor for $\mathcal{C}$ is a function $E : \{0,1\}^n \rightarrow \{0,1\}^m$ such that for any X in $\mathcal{C}$ the distribution $E(X)$ is statistically close to the uniform distribution. A long line of research deals with explicit constructions of such extractors for various classes $\mathcal{C}$ while trying to maximize m.

In this paper we give a general transformation that transforms a deterministic extractor E that extracts "few" bits into an extractor E' that extracts "almost all the bits present in the source distribution". More precisely, we prove a general theorem saying that if E and $\mathcal{C}$ satisfy certain properties, then we can transform E into an extractor E'.

Our methods build on (and generalize) a technique of Gabizon, Raz and Shaltiel (FOCS 2004) that present such a transformation for the very restricted class $\mathcal{C}$ of "oblivious bit-fixing sources". Loosely speaking the high level idea is to find properties of E and $\mathcal{C}$ which allow "recycling" the output of E so that it can be "reused" to operate on the source distribution. An obvious obstacle is that the output of E is correlated with the source distribution.

Using our transformation we give an explicit construction of a two-source extractor *$E : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}^m$ such that for every two independent distributions X_1 and X_2 over $\{0,1\}^n$ with min-entropy at least $k = (1/2+\delta)n$, $E(X_1, X_2)$ is ϵ-close to the uniform distribution on $m = 2k - C_\delta \log(1/\epsilon)$ bits. This result is optimal except for the precise constant C_δ and improves previous results by Chor and Goldreich (SICOMP 1988), Vazirani (Combinatorica 1987) and Dodis et al. (RANDOM 2004).*

We also give explicit constructions of extractors for samplable distributions *that extract many bits even out of "low-entropy" samplable distributions. This improves some previous results by Trevisan and Vadhan (FOCS 2000).*

*This research was supported by BSF grant 2004329.

1. Introduction

1.1. Background

A very successful paradigm in computer science is that of randomized algorithms and protocols. It is known that having access to random bits allow parties to perform tasks that cannot be performed by deterministic parties. (The reader is referred to [18, 33] for textbooks on randomized algorithms). The most obvious example is the area of Cryptography which inherently relies on the assumption that parties have access to random bits (that is to a string of independent coin tosses).

A large body of research is concerned with obtaining such a sequence of random bits. A common strategy is to try and sample from some source distribution X (say on n bit strings) that is available to a computer (some examples are: electro-magnetic noise, key strokes of user and timing of past events). However, it is unlikely that this gives bits that are independent coin tosses. A successful paradigm is the "randomness-extraction paradigm" in which one attempts to extract bits that are (close to) uniform from the source distribution X.

1.1.1 Deterministic randomness extractors

A "deterministic randomness extractor" is a function that "extracts" bits that are (statistically close to) uniform from "weak sources of randomness" which may be very far from uniform.

Definition 1 (deterministic extractor). *Let $\mathcal{C}$ be a class of distributions on $\{0,1\}^n$. A function $E : \{0,1\}^n \rightarrow \{0,1\}^m$ is a* deterministic ϵ-extractor *for $\mathcal{C}$ if for every distribution X in $\mathcal{C}$ the distribution $E(X)$ (obtained by sampling x from X and computing $E(x)$) is ϵ-close to the uniform distribution on m bit strings.*[1]

[1] Two distributions P and Q over $\{0,1\}^m$ are ϵ-close (denoted by $P \sim_\epsilon Q$) if for every event $A \subseteq \{0,1\}^m$, $|P(A) - Q(A)| \leq \epsilon$.

Given a class $\mathcal{C}$ the goal of this field is to design *explicit* (that is polynomial time computable) deterministic extractors that extract as many random bits as possible.

1.1.2 Some related work on deterministic randomness extraction.

Various classes $\mathcal{C}$ of distributions were studied in the literature: The first construction of deterministic extractors can be traced back to von Neumann [38] who showed how to use many independent tosses of a biassed coin (with unknown bias) to obtain an unbiased coin. Blum [5] considered sources that are generated by a finite Markov-chain. Santha and Vazirani [27], Vazirani [35, 36], Chor and Goldreich [6], Dodis et al. [9], Barak, Impagliazzo and Wigderson [1], Barak et al. [2] and Raz [24] studied sources that are composed of several independent samples from various classes of "high entropy" distributions. Chor et al. [7], Ben-Or and Linial [4], Cohen and Wigderson [8], Kamp and Zuckerman [16] and Gabizon, Raz and Shaltiel [12] studied bit-fixing sources which are sources in which a subset of the bits are uniformly distributed. Trevisan and Vadhan [32] studied sources which are "samplable" by small circuits. Barak et al. [2] and Gabizon and Raz [11] studied sources which are uniform over an affine subspace.

1.1.3 Seeded extractors

A negative result was given by Santha and Vazirani [27] that exhibit a very natural class of high-entropy sources that does not have deterministic extractors. This led to the development of a different notion of extractors called "seeded extractors". Such extractors are allowed to use a short seed of few truly random bits when extracting randomness from a source. (The notion of "seeded extractors" emerged from attempts to simulate probabilistic algorithms using weak random sources [37, 6, 8, 39, 40] and was explicitly defined by Nisan and Zuckerman [22].) Unlike deterministic extractors, seeded extractors can extract randomness from the most general class of sources: Sources with high (min)-entropy.

Definition 2 (min-entropy). *Given a random variable X taking values in $\{0,1\}^n$ the* min-entropy *of X denoted $H_\infty(X)$ is given by $min_{x\in\{0,1\}^n} \log(1/\Pr[X=x])$.*

A seeded randomness extractor is a function which receives two inputs: In addition to a sample from a source X, a seeded extractor also receives a short "seed" Y of few uniformly distributed bits. Loosely speaking, the extractor is required to output many more random bits than the number of bits "invested" as a seed.

Definition 3 (seeded extractors for high min-entropy sources). *A function $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a* strong (k,ϵ)-extractor *if for every random variables X, Y such that $H_\infty(X) \geq k$ and Y is independent of X and uniformly distributed over $\{0,1\}^d$: $(E(X,Y),Y)$ is ϵ-close to the uniform distribution on $m+d$ bits.*

A long line of research focuses on constructing such seeded extractors with as short as possible seed length that extract as many as possible bits. There are explicit constructions of (k,ϵ)-extractors that use seed of length $\text{polylog}(n) + O(\log(1/\epsilon))$ to extract k random bits. The reader is referred to [19, 20, 34] for surveys on applications of seeded randomness extractors and to [28] for a survey that focuses on recent explicit constructions of seeded randomness extractors.

1.2. How to get more mileage from deterministic extractors

Let $\mathcal{C}$ be some class of distributions over $\{0,1\}^n$ and assume that for any X in $\mathcal{C}$, $H_\infty(X) \geq k$. The randomness extraction problem is to design an explicit extractor E for $\mathcal{C}$ which extracts as many random bits as possible. It is natural to hope to extract $m \approx k$ random bits as all distributions in $\mathcal{C}$ "contain" k random bits. Suppose we already have an explicit ϵ-extractor E for $\mathcal{C}$ that extracts $t < k$ random bits it is natural to try to "get more mileage from E". That is to try and extract more bits by "recycling the output of E" as follows:

$$E'(x) = E_1(x, E(x))$$

where $E_1 : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ is a seeded extractor for min-entropy threshold k. If t is large enough (say $t > \text{polylog}(n)$) then there are explicit constructions of a seeded extractor E_1 that extracts *all* the k random bits from the source distribution.

There is however an obvious obstacle. Let X be a distribution in $\mathcal{C}$. While it is true that $E(X)$ is close to uniform it is inherently dependent on the distribution X. Thus, when we run E_1 we run it with two dependent distributions $X, E(X)$ and cannot conclude that the output is close to uniform.

1.2.1 The technique of Gabizon et al.

Gabizon et al. [12] focus on the class of (oblivious) bit-fixing sources with min-entropy k. These are distributions X over n bit strings such that there is a set $S \subseteq [n]$ of k indices such that X restricted to S is uniformly distributed and the remaining bits are fixed. They show a specific construction of a function $E_1 : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ such that if E is an ϵ-extractor which extracts t random bits from "low-entropy" bit-fixing sources then for every distribution X in the class, $E'(X) = E_1(X, E(X))$ is $O(\epsilon \cdot 2^t)$-close to the uniform distribution on $m \approx k$ bits.

That is at least for the restricted case of bit-fixing sources, if one starts with an extractor E that extracts few bits out of X then it is possible to come up with a function E_1 which gives rise to a deterministic extractor E' that extracts almost all the bits out of the source distribution. It is important to notice that to achieve this we have to require that E works for a class $\mathcal{C}$ which is larger than our "target class": It is not sufficient that E extracts t random bits from bit-fixing sources with min-entropy k. It is crucial that E extracts randomness even from bit fixing sources with min-entropy $k' < k$. We also remark that E_1 needs to have a certain special structure. We elaborate more on the construction of [12] later on.

Following [12], Gabizon and Raz [11] used related ideas to get more mileage out of deterministic extractors for a "affine sources over large fields". (These are sources in which $X = (X_1, \ldots, X_n)$ where each X_i belongs to a field of size at least poly(n) and there exists a subset $S = \{i_1, \ldots, i_k\}$ of indices such that X restricted to S is uniformly distributed and for every $i \notin S$, X_i is given by a linear function of $X_{i_1}, \ldots, X_{i_k}$.) We elaborate more on the construction of [11] later on.

1.2.2 Our result: a general transformation

We now explain our main result. We are given a distribution X over $\{0,1\}^n$ and functions $E : \{0,1\}^n \to \{0,1\}^t$, $F : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^r$. We want to get that $F(X, E(X))$ is close to $F(X \otimes U_t)$.[2] We show that a sufficient condition is that there is a class $\mathcal{C}$ of distributions such that:

- X belongs to $\mathcal{C}$.
- E is an ϵ-extractor for $\mathcal{C}$ with $\epsilon < 2^{-t}$.
- For every $y \in \{0,1\}^t$ and $a \in \{0,1\}^r$, $(X|F(X,y) = a)$ belongs to $\mathcal{C}$. (We refer to this condition as a "closeness condition").

Thus, for example if F is a seeded extractor we get that $F(X, E(X))$ is close to $F(X \otimes U_t)$ which is in turn close to uniform. We now give a formal statement of our main result:

Theorem 4 (main theorem). *Let $\mathcal{C}$ be a class of distributions over $\{0,1\}^n$. Let $E : \{0,1\}^n \to \{0,1\}^t$ be an ϵ-extractor for $\mathcal{C}$. Let $F : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^r$. Let X be a distribution in $\mathcal{C}$ and assume that:*

Closeness condition: *For every $y \in \{0,1\}^t$ and $a \in \{0,1\}^r$, $(X|F(X,y) = a)$ belongs to $\mathcal{C}$.*

Then $F(X, E(X)) \sim_{\epsilon \cdot 2^{t+3}} F(X \otimes U_t)$.

[2]For two distributions P, Q we use $P \otimes Q$ to denote the distribution of pairs (p, q) sampled independently from P and Q.

We remark that the distribution $(X|F(X,y) = a)$ typically has lower min-entropy than the distribution X. Thus, to apply the Theorem on a distribution X with min-entropy k we typically need an extractor E that extracts randomness from distributions with min-entropy smaller than k.

Overview of the argument: We prove (a more general version of) Theorem 4 in Section 3. We now explain the intuition behind the proof. It is helpful to oversimplify the situation and assume that E is errorless, that is that E is a 0-extractor for $\mathcal{C}$.

Our goal is to show that $F(X, E(X))$ is equivalent to $F(X \otimes U_t)$. For this purpose it is sufficient to show that for every $y \in \{0,1\}^t$, $F(X,y)$ is equivalent to $(F(X,E(X))|E(X) = y)$. This is because by the assumption that X belongs to C we have that $E(X)$ is equivalent to U_t.

For this purpose, it is sufficient to show that for every $y \in \{0,1\}^t$, $F(X,y)$ is independent of $E(X)$ because then

$$(F(X,E(X))|E(X) = y) \sim (F(X,y)|E(X) = y) \sim F(X,y)$$

However, we have that for every $y \in \{0,1\}^t$ and $a \in \{0,1\}^r$, $(X|F(X,y) = a)$ is a distribution in $\mathcal{C}$. It follows that $(E(X)|F(X,y) = a)$ is the uniform distribution. As this is true for every $a \in \{0,1\}^r$ we have that $E(X)$ is independent of $F(X,y)$. The actual proof imitates the outline above while taking into account that $\epsilon > 0$.

1.3. Applications

We show that our technique is applicable to a large variety of classes of sources.

1.3.1 Extractors for two independent sources

We now consider the class of random variables X that are composed of two independent random variables X_1 and X_2 such that $H_\infty(X_1) \geq k_1$ and $H_\infty(X_2) \geq k_2$.

Definition 5 (two source extractors). *A function $E : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ is a $(k_1, k_2; \epsilon)$-*two-source extractor *if for every two independent random variables X_1, X_2 such that $H_\infty(X_1) \geq k_1$ and $H_\infty(X_2) \geq k_2$: $E(X_1, X_2)$ is ϵ-close to the uniform distribution on m bits.*

The function E is strong in the first source *if $(X_1, E(X_1, X_2))$ is ϵ-close to (X_1, Z) where Z is an independent random variable that is uniformly distributed over $\{0,1\}^m$.*

As noted earlier in Section 1.1.2, a long line of research is concerned with explicit constructions of such extractors. We now mention only results that are directly relevant to this paper. We focus on the scenario in which $n_1 = n_2$ and we denote this length by n. Furthermore we are concerned with entropy threshold $k = (1/2 + \delta)n$ for some constant $\delta > 0$ and set $k_1 = k_2 = k$.

Chor and Goldreich [6] proved that the function $E(x, y) = IP(x, y)\text{mod}2$ (here IP is the inner-product function) is a $(k, k; \epsilon)$-two source extractor (with error $\epsilon = 2^{-\eta n}$ where $\eta > 0$ is a constant that depends only on δ). Note that this extractor extracts only a single bit out of the source distribution. Vazirani [35, 36] (see also [10, 9]) constructed a function E that extracts $\Omega(\delta n)$ bits with the same properties. Dodis et al. [9] observed that Vazirani's extractor is strong in the first source (and also in the second one). A consequence is that it is possible to use the output of E to operate on the first source with a seeded extractor. More specifically, that the construction $E'(x_1, x_2) = E_1(x_1, E(x_1, x_2))$ where E_1 is a seeded extractor yields a two-source extractor. Using this method, [9] construct an extractor that extracts $k + \Omega(\delta n)$ bits (that is about half the entropy present in the source distribution).

In this paper we use our technique (and a recent extractor construction of [24]) to improve this result. In particular we show how to extract $2k - C\log(1/\epsilon)$ bits (for a constant C depending only on δ). That is, we extract almost all the $2k$ random bits present in the source distribution. This is stated formally in the next Theorem:

Theorem 6. *For every constant $\delta > 0$ there is a constant $C > 0$ such that for large enough n, let $k_1 = k_2 = (1/2 + \delta)n$ and let $\epsilon \leq 2^{-\log^4 n}$ then there is an explicit $(k_1, k_2; \epsilon)$-two-source extractor $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ for $m = k_1 + k_2 - C\log(1/\epsilon)$.*

Radhakrishnan and Ta-Shma [23] showed that any two-source extractor must suffer an entropy loss of $2\log(1/\epsilon)$. It follows that the output length of our extractor is optimal except for the value of the constant C.

Overview of the argument: We show that the function

$$E'(x_1, x_2) = E_1(x_1, E(x_1, x_2)), E_2(x_2, E(x_1, x_2))$$

is a two-source extractor at least when E is a two source extractor with somewhat stronger properties. Note that we use $E(x_1, x_2)$ as a seed for seeded extractors that operate both on x_1 and x_2. We find this surprising as unlike the case of [9] we have that (X_1, X_2) is inherently dependent on $E(X_1, X_2)$. The basic intuition is that we can use our transformation whenever the extractor E is able to extract randomness from sources of type $((X_1, X_2)|E_2(X_2, y) = a)$. (This is because by the observation of Dodis et al. [9] mentioned above, we can later apply E_1 on x_1 using $E(x_1, x_2)$ as a seed. As X_1, X_2 are independent we have that:

$$((X_1, X_2)|E_2(X_2, y) = a) \sim (X_1, (X_2|E_2(X_2, y) = a))$$

We can set the parameters of E_2 so that $(X_2|E_2(X_2, y) = a)$ has (a small amount of) entropy. Finally, the extractor of [24] can handle scenarios in which one source has entropy $k = (1/2 + \delta)n$ and the other has low entropy. The exact details are given in Section 4. We remark that any future improvement in constructing two-source extractors for low entropy threshold can be plugged into our technique.

1.3.2 Extractors for samplable distributions

Trevisan and Vadhan [32] suggested studying the class of distributions that are samplable by polynomial size circuits.

Definition 7 (Samplable distributions). *A function $f : \{0,1\}^r \to \{0,1\}^n \cup \{\perp\}$ is a* sampler *for a distribution P over $\{0,1\}^n$ if for a random variable R that is uniformly distributed in $\{0,1\}^r$:*

- $(f(R)|f(R) \neq \perp) \sim P$.
- $\Pr[f(R) = \perp] \leq 1/3$.

Given a class $\mathcal{F}$ of functions we define a class of distributions samplable by $\mathcal{F}$ *which is the class of all distributions P such that there is an f in $\mathcal{F}$ which is a sampler for P.*

We remark that this definition is slightly different than that in [32].[3]

Definition 8 (Extractors for samplable distributions). *A function $E : \{0,1\}^n \to \{0,1\}^m$ is a (k, ϵ)-extractor for a class $\mathcal{F}$ of functions if E is an ϵ-extractor for the class of distributions X samplable by $\mathcal{F}$ which satisfy $H_\infty(X) \geq k$.*

Following [32] we are interested in extracting randomness from distributions X on n bit strings that are samplable by size $s = n^{O(1)}$ circuits. Trevisan and Vadhan showed that such extractors exist (and in fact are computable by size $s^{O(1)}$ circuits). They also showed that any extractor E for such distributions cannot have a size s circuit. This means that with our current state of knowledge we cannot expect to have unconditional explicit extractors as such extractors are polynomial time computable functions that cannot be computable by size s circuits. Trevisan and Vadhan show that explicit extractors can be constructed assuming the existence of polynomial time computable functions that are hard on average for size s Σ_1-circuits.[4]

[3] We allow the sampling function to output $\perp$. This is a minor difference and we take this approach to allow sampling from distributions which have probabilities that aren't powers of two. We stress that all our results also follow in the definition of [32].

[4] A Σ_1-circuit is a circuit which in addition to the standard boolean gates is also allowed to use gates which compute a Σ_1-complete problem (e.g. satisfiability). The precise definition appears in Definition 31.

Theorem 9. *[32] Let $f : \{0,1\}^n \to \{0,1\}$ be a function such that for any Σ_1-circuit C of size s, $\Pr[C(X) = f(X)] \le \epsilon$ (here X is a uniformly distributed random variable over $\{0,1\}^n$). Then for any $\Delta > 0$, f is a $(n - \Delta, O(2^{\Delta}\epsilon))$-extractor for distributions samplable by size $(2^{\Delta}\epsilon s)^{\Omega(1)}$ (deterministic) circuits.*

It is natural to extend this approach to extract more than one bit. Note that a function f as above satisfies that the distribution $(X, f(X))$ is indistinguishable from uniform by size s Σ_1-circuits. By using the same argument as in [32] it is possible to extract $t < \log s$ bits using an analogous assumption that there is a pseudorandom generator G such that the distribution $(X, G(X))$ is indistinguishable from uniform by small Σ_1-circuits.

Lemma 10. *Let $G : \{0,1\}^n \to \{0,1\}^t$ be a function such that for any Σ_1-circuit D of size s:*

$$|\Pr[D(X, G(X)) = 1] - \Pr[D(X, Y) = 1]| \le \epsilon$$

where X and Y are independent random variables distributed uniformly over $\{0,1\}^n$ and $\{0,1\}^t$ respectively. If $t \le \log s - 1$ then for any $\Delta > 0$, G is a $(n - \Delta, O(2^{\Delta}\epsilon))$-extractor for distributions samplable by size $(2^{\Delta}\epsilon s)^{\Omega(1)}$ (deterministic) circuits.

In Section 6 we discuss whether the assumptions of Theorem 9 and Lemma 10 are equivalent. We remark that this argument only works for $t < \log s$.[5] Given a size bound $s = n^{O(1)}$ if there exist a generator G as in the lemma then we get an extractor. We remark that as we want G to run in polynomial time we must assume that this polynomial is larger than s and can extract at most $t = O(\log n)$ bits.

Using our general approach we show how to convert an extractor E that extracts $t = O(\log n)$ bits into an extractor E' that extracts almost all the bits present in the source distribution. Our construction improves a different construction implicit in [32] that achieves the same goal for the special case when the min-entropy threshold is $k = (1 - \nu)n$ for a small constant $\nu > 0$. The advantage of our construction (that is presented in Section 5) is that it works for any $k > \log^4 n$. One subtlety in our construction (that also occurs in the construction of [32]) is that it is not sufficient that the initial extractor E extracts randomness from distributions samplable by deterministic circuits. Instead we make the stronger requirement that E extracts randomness from distributions samplable by Σ_1-circuits.

Thus, we need to move everything "up one level in the hierarchy". Following [32] we notice that Lemma 10 relativizes and therefore if one assumes that G fools Σ_2-circuits then one gets that G is an extractor for distributions samplable by Σ_1-circuits. We now show how to use our transformation to improve some constructions of extractors for samplable distributions. Nevertheless, we remark that in this setting the improvements are less significant as we do not have good extractors to "start from". Putting the ideas sketched above together we get that if the assumption of Lemma 10 holds for Σ_2-circuits with $\epsilon = 2^{-(1-\beta)n}$ for some constant $\beta > 0$ then for any k and constant $\alpha > 0$ we get an explicit extractor E that on distribution samplable by size $s^{\Omega(1)}$ circuits extract $(1 - \alpha)k - \beta n - O(\log n)$ bits. Exact details are given in Section 5.

Overview of the argument: Our basic construction is given by $E'(x) = E_1(x, E(x))$ where E is a deterministic extractor for distributions samplable by small Σ_1-circuits and E_1 is a seeded extractor. The main problem is that to apply Theorem 4 we need $(X|E_1(X, y) = a)$ to be a smaplable distribution. We have that E_1 is polynomial time computable, however this does not suffice to get that the distribution above is samplable by deterministic circuits. It does follow (by using results on sampling NP witnesses [30, 15, 3]) that the distribution above is samplable by a Σ_1-circuit and this is why we require that E is an extractor for such distributions.

We remark that Tevisan and Vadhan also give another construction of extractors for samplable distributions that starts from worst case hardness. Even with the strongest possible assumptions this construction only gives an extractor for entropy threshold $k = (1 - \nu)n$ for some small constant $\nu > 0$. Thus, at the moment our transformation does not give an improvement in this setup.

1.3.3 Extractors for bit-fixing sources

For completeness we now sketch how the main result of Gabizon, Raz and Shaltiel [12] follows in our framework. We are interested in extracting randomness from bit-fixing sources with min-entropy k. We are given an extractor E that extracts few bits (say t bits) and want to convert it into an extractor E' that extracts almost all the bits present in the source distribution. Interestingly, in this scenario we cannot use the construction $E'(x) = E_1(x, E(x))$ as used in the previous applications. Instead we use Theorem 4 to construct a "seed-obtainer". This is an object $A(x)$ introduced in [12] which on a bit-fixing source X with min-entropy k outputs two distributions X' and Z such that X' is a bit-fixing source with min-entropy $k' \approx k$, Z is short and is (close to) uniform and X' and Y are (close to) being independent. Given a seed obtainer we can use Z as a seed to a seeded extractor and extract randomness from X' and the point is that we are guaranteed that X' and Z are (close to) independent.

[5] Loosely speaking, this is because the proof converts a distinguisher D' for $G(X)$ where X is a samplable distribution into a distinguisher D as above. The assumption that $t < \log s$ guarantees that D' can be expressed as a size $2^t < s$ circuit.

To construct a seed obtainer, [12] use an "averaging sampler" this is a function $Samp$ that given a random seed of $t/2$ random bits produces a subset $T \subseteq [n]$ which "behaves like a random set" in the sense that it intersects every fixed set T in essentially the same way as if it was a random set. (The reader is referred to [13] for a survey on averaging samplers.) To construct the seed-obtainer we consider the following function $F : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^n$: Given x and y, F splits the t bits long y into two $t/2$ bit long blocks y_1, y_2. It uses y_1 as a seed to a sampler to produce a set T and sets x' to be $x|_T$ (x restricted to the indices in T) and $z = y_2$. The parameters are set so that with high probability (over the choice of the seed to the sampler y_1) T hits approximately k' indices of bits that are random in X.

It follows that if we apply F on X and an independent uniformly distributed random variable Y then we obtain X' and Z such that indeed X' is (a convex combination of) bit-fixing sources with min-entropy k' and $Z = Y_2$ is independent and uniformly distributed. Thus, if we can meet the requirements of Theorem 4 then we can argue that $F(X, E(X))$ is close to $F(X, Y)$ and this means that $A(x) = F(x, E(x))$ gives a seed obtainer. To meet the requirements we note that $(X|F(X, y) = (x', y_2))$ is a bit fixing source with $k - k'$ bits. (This is because the conditioning misses $k - k'$ indices in which X is random). Thus, if E is an extractor for bit-fixing sources with min-entropy $k - k'$ then this transformation works and yields a seed obtainer.

1.3.4 Extractors for affine sources

Recall that Gabizon and Raz [11] used an analysis analogous to that in [12] to get more mileage out of extractors for affine sources. For completeness we sketch how this transformation follows in our framework.

We are given an extractor $E : \{0,1\}^n \to \{0,1\}^t$ that extracts randomness from distributions that are uniform over an affine space of dimension k'. We want to get more mileage by using $E'(x) = E_1(x, E(x))$ where E_1 is a seeded extractor. Given an affine source X of dimension $k > k'$ we want to apply Theorem 4. We need that $(X|E_1(X, y) = a)$ is an affine source. This follows when $E_1(X, y)$ is a linear function. Thus, we can get more mileage out of affine sources by using seeded extractors which for every seed y are linear functions. Gabizon and Raz give a construction of a deterministic extractor which extracts few bits and a seeded extractor which is linear and as a consequence get a deterministic extractor that extracts many bits.

We remark that Gabizon and Raz focus on the case where field is of polynomial size (rather than say size two) because their initial deterministic extractor only works in this case. Nevertheless, we want to point out that the transformation works for any field. In fact, the transformation can also be applied in the affine-source disperser of Barak et al. [2] but this requires getting into the details of that paper (essentially because the final object is a disperser and not an extractor).

1.4. Can we get more mileage from seeded extractors?

In Section 6 we give a counterexample showing that our method cannot be directly applied to get more mileage out of seeded extractors. One reason that our analysis does not work for seeded extractor is that we require the error ϵ of the initial extractor E to satisfy $\epsilon \leq 2^{-t}$ where t is the output length of E. However, by the lower bounds of [22, 23] an extractor with error ϵ must have a seed of length at least $2\log(1/\epsilon) > t$ and thus it is not useful as it spends more bits than it extracts.

1.5. Outline of this paper

Due to space limitations some proofs are omitted from this version. In Section 2 we give some necessary preliminaries. In Section 3 we explain our main transformation, restate Theorem 4 in a more general way and prove it. In Section 4 we show how to apply our technique to two-source extractors. In Section 5 we show how to apply our technique to exractors for samplable distributions. Finally, we discuss limitations of our technique and present open problems in Section 6.

2. Preliminaries

Notations: We use $[n]$ to denote the set $\{1, \ldots, n\}$. We denote the length of a string x by $|x|$. Logarithms will always be taken with base 2. We use U_n to denote the uniform distribution over n bits. Given a distribution A we use $w \leftarrow A$ to denote the experiment in which w is chosen randomly according to A.

2.1. Probability distributions

Distributions: Some of the proofs in this paper require careful manipulations of probability distributions. We use the following notation. We denote the probability of an event B under a probability distribution P by $\Pr_P[B]$. We say that two distributions P and Q are ϵ-close (denoted $P \sim_\epsilon Q$) if for any event B, $|\Pr_P[A] - \Pr_Q[A]| \leq \epsilon$.

Random variables: A random variable R that takes values in U is a function $R : \Omega \to U$ (where Ω is a probability space). We sometimes refer to R as a probability distribution over U (the distribution of the output of R). For example, given a random variable R and a distribution P we sometimes write "$R \sim P$" and this means that the distribution of the output of R is equal to P.

Pairs of distributions and variables: Given two random variables R_1, R_2 over the same probability space Ω we use (R_1, R_2) to denote the random variable induced by the function $(R_1, R_2)(\omega) = (R_1(\omega), R_2(\omega))$. Given two probability distributions P_1, P_2 over domains Ω_1, Ω_2 we define $P_1 \otimes P_2$ to be the product distribution of P_1 and P_1 which is defined over the domain $\Omega_1 \times \Omega_2$.

Definition 11 (conditioning distributions and random variables). *Given a probability distribution P over some domain U and an event $A \subseteq U$ such that $\Pr_P[A] > 0$ we define a distribution $(P|A)$ over U as follows: Given an event $B \subseteq U$, $\Pr_{(P|A)}(B) = \Pr_P[B|A] = \frac{\Pr_P[A \cap B]}{\Pr_P[A]}$.*

We extend this definition to random variables $R : \Omega \to U$. Given an event $A \subseteq \Omega$ we define $(R|A)$ to be the probability distribution over U given by $\Pr_{(R|A)}[B] = \Pr[R \in B|A]$.

We also need the following standard technical lemmas. We omit the proofs in this version.

Lemma 12. *Let R_1, R_2 be random variables taking values in A. Let $f : A \to B$ be some function. If $R_1 \sim_\epsilon R_2$ then $f(R_1) \sim_\epsilon f(R_2)$.*

Lemma 13. *Let R_1 be a random variable taking values in A_1 and R_2 be a random variable taking values in $\{0,1\}^v$. Assume that $H_\infty(R_1) \geq k$. Then for every $\rho > 0$ there exists a set $G \subseteq \{0,1\}^t$ such that:*

1. $\Pr[R_2 \in G] \geq 1 - \rho$.
2. *For every $a \in G$, $H_\infty((R_1|R_2 = a_2)) \geq k - (v + \log(1/\rho))$.*

Lemma 14. *Let R_1, V_1 be random variables taking values in A_1 and R_2, V_2 be random variables taking values in A_2. Suppose that:*

1. $R_2 \sim_{\epsilon_2} V_2$.
2. *For every $a \in A_2$, $(R_1|R_2 = a) \sim_{\epsilon_1} (V_1|V_2 = a)$.*

Then $(R_1, R_2) \sim_{(\epsilon_1+\epsilon_2)} (V_1, V_2)$.

Lemma 15. *Let R_1, V_1 be random variables taking values in A_1 and R_2 be random variables taking values in A_2. Suppose that there exists a set $G \subseteq A_2$ such that:*

1. $\Pr[R_2 \in G] \geq 1 - \epsilon_2$.
2. *For every $a \in G$, $(R_1|R_2 = a_2) \sim_{\epsilon_1} V_1$.*

Then $(R_1, R_2) \sim_{(\epsilon_1+\epsilon_2)} (V_1, R_2)$.

We also use the following easy Lemma proven in [12].

Lemma 16. *[12] (Lemma 2.6) Let R_1 be a random variable taking values in A_1 and R_2 be a random variable taking values in $\{0,1\}^t$. Assume that $(R_1, R_2) \sim_\epsilon (R_1 \otimes U_t)$. Then for every $b \in \{0,1\}^t$, $(R_1|R_2 = b) \sim_{\epsilon \cdot 2^{t+1}} R_1$.*

3. A transformation for a general family of sources

3.1. The main theorem

For some of our intended applications we need to restate Theorem 4 in a more general form. We first need the definition of a *strong deterministic extractor*.

Definition 17 (strong deterministic extractor). *Let $\mathcal{C}$ be a class of distributions on $\{0,1\}^n$. A function $E : \{0,1\}^n \to \{0,1\}^m$ is a* strong deterministic ϵ-extractor *for $\mathcal{C}$ with respect to some function* Part $: \{0,1\}^n \to \{0,1\}^p$ *if for every distribution X in $\mathcal{C}$: $(E(X), \mathsf{Part}(X))$ is ϵ-close to the distribution $(Z, \mathsf{Part}(X))$ where Z is an independent random variable that is uniformly distributed over $\{0,1\}^m$.*

As the statement of Theorem 18 below is rather technical, we first explain the main differences from the more simple Theorem 4.

A weaker closeness condition: We replace the "closeness condition" of Theorem 4 by a wekaer condition which essentially says that the closeness condition holds with high probability rather than with probability one.

A stronger guarantee: We don't only claim that $F(X, E(X))$ is close to $F(X \otimes U_t)$ but rather that for any $y \in \{0,1\}^t$, $(F(X, E(X))|E(X) = y)$ is close to $F(X, y)$.

Allowing E to be strong: We allow E to be a strong extractor (with respect to some function Part). In this case we get the stronger conclusion that $F(X, E(X))$ is close to being independent from $\mathsf{Part}(X)$.

Recycling the bits of $E(X)$: We replace the former guarantee that $F(X, E(X))$ is close to $F(X \otimes U_t)$ by the stronger guarantee that $(F(X, E(X)), E(X))$ is close to $(F(X, Y), Y)$ where Y is independent of X and is uniformly distributed in $\{0,1\}^t$.

We now state the more general version.

Theorem 18 (main theorem: more general version). *Let $\mathcal{C}$ be a class of distributions over $\{0,1\}^n$. Let $E : \{0,1\}^n \to \{0,1\}^t$ be an ϵ-extractor for $\mathcal{C}$ that is strong with respect to a function* Part $: \{0,1\}^n \to \{0,1\}^p$. *Let $F : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^r$. Let X be a distribution in $\mathcal{C}$ and assume that:*

Weak closeness condition: *For every $y \in \{0,1\}^t$ there exists a set $G_y \subseteq \{0,1\}^r$ such that:*

1. $\Pr[F(X,y) \notin G_y] \le \epsilon$.
2. *For every $a \in G_y$, $(X|F(X,y)=a)$ belongs to $\mathcal{C}$.*
3. $\mathsf{Part}(X)$ *is independent of* $F(X,y)$.

Then,

1. *For every $y \in \{0,1\}^t$,* $(\mathsf{Part}(X), F(X,E(X))|E(X)=y) \sim_{\epsilon \cdot 2^{t+2}} (\mathsf{Part}(X), F(X,y))$
2. *Let Y be a random variable that is independent of X and is uniformly distributed over $\{0,1\}^t$. We have that:* $(\mathsf{Part}(X), F(X,E(X)), E(X)) \sim_{\epsilon \cdot 2^{t+3}} (\mathsf{Part}(X), F(X,Y), Y)$.

Note that Theorem 4 is indeed a special case of Theorem 18 by choosing $\mathsf{Part}(X)$ to be some constant function. It is therefore sufficient to prove Theorem 18. This is done in the next subsection.

3.2. Proof of correctness

In this Section we prove Theorem 18. Let X be a distribution in $\mathcal{C}$ which satisfies the "weak closeness condition". Fix some $y \in \{0,1\}^t$ and let G_y be the set guaranteed by the "weak closeness condition".

Lemma 19. *For every $a \in G_y$,* $(E(X), \mathsf{Part}(X)|F(X,y)=a) \sim_\epsilon (U_t \otimes \mathsf{Part}(X))$.

Proof. This follows because for this choice of a the "weak closeness condition" gives that $X' = (X|F(X,y)=a)$ is a distribution in $\mathcal{C}$. The function E is an ϵ-extractor for $\mathcal{C}$ that is strong with respect to Part and the lemma follows by applying the extractor E on X' and using the fact that $\mathsf{Part}(X)$ is independent of $F(X,y)$. □

Lemma 20. $(E(X), \mathsf{Part}(X), F(X,y)) \sim_{2\epsilon} (U_t \otimes (\mathsf{Part}(X), F(X,y))$.

Proof. The lemma will follow by applying Lemma 15. We choose:

- $R_1 = (E(X), \mathsf{Part}(X))$.
- $R_2 = F(X,y)$.
- $V_1 = (U_t \otimes \mathsf{Part}(X))$.

By the "weak closeness condition" and Lemma 19 we have that:

1. $\Pr[R_2 \notin G_y] \le \epsilon$.
2. For every $a \in G_y$, $(R_1|R_2=a) \sim_\epsilon V_1$.

Applying Lemma 15 we conclude that $(R_1,R_2) \sim_{2\epsilon} (V_1,R_2)$ or in other words:

$$(E(X), \mathsf{Part}(X), F(X,y)) \sim_{2\epsilon} ((U_t \otimes \mathsf{Part}(X)), F(X,y))$$

$$\sim (U_t \otimes (\mathsf{Part}(X), F(X,y)))$$

□

Lemma 21. $(\mathsf{Part}(X), F(X,y)|E(X)=y) \sim_{2\epsilon \cdot 2^{t+1}} (\mathsf{Part}(X), F(X,y))$.

Proof. The lemma will follow by applying Lemma 16. We choose $R_1 = (\mathsf{Part}(X), F(X,y))$ and $R_2 = E(X)$. By Lemma 21 we have that $(R_1,R_2) \sim_{2\epsilon} (R_1 \otimes U_t)$. It follows from Lemma 16 that for every $b \in \{0,1\}^t$, $(R_1|R_2=b) \sim_{2\epsilon \cdot 2^{t+1}} R_1$. In particular choosing $b=y$ we get that:

$$(\mathsf{Part}(X), F(X,y)|E(X)=y) \sim_{2\epsilon \cdot 2^{t+1}} (\mathsf{Part}(X), F(X,y))$$

as required. □

We are now ready to prove the 1st item of Theorem 18.

$$(\mathsf{Part}(X), F(X,E(X))|E(X)=y)$$

$$\sim (\mathsf{Part}(X), F(X,y)|E(X)=y)$$

$$\sim_{\epsilon \cdot 2^{t+2}} (\mathsf{Part}(X), F(X,y)) \qquad (1)$$

Where the last move is by Lemma 21. The 1st item follows. (We remark that for the 1st item we did not use the assumption that X is in $\mathcal{C}$). We now prove the second item. This will follow by applying Lemma 14. For this purpose we add to our probability space an independent random variable Y that is uniformly distributed over $\{0,1\}^t$. We choose:

- $R_1 = (\mathsf{Part}(X), F(X,E(X)))$.
- $R_2 = E(X)$.
- $V_1 = (\mathsf{Part}(X), F(X,Y))$.
- $V_2 = Y$.

By the assumption that X is in $\mathcal{C}$ and the fact that E is an ϵ-extractor for $\mathcal{C}$ we have that $R_2 \sim_\epsilon V_2$. By equation (3.2) we have that for every $y \in \{0,1\}^t$ $(R_1|R_2=y) \sim_{\epsilon \cdot 2^{t+2}} (V_1|V_2=y)$. (Here we also used the fact that Y is independent of X). Applying Lemma 14 gives that $(R_1,R_2) \sim_{(\epsilon \cdot 2^{t+2}+\epsilon)} (V_1,V_2)$. Note that $\epsilon \cdot 2^{t+2} + \epsilon \le \epsilon \cdot 2^{t+3}$. We conclude that:

$$(\mathsf{Part}(X), F(X,E(X)), E(X)) \sim_{\epsilon \cdot 2^{t+3}} (\mathsf{Part}(X), F(X,Y), Y)$$

This concludes the proof of Theorem 18.

4. Extractors for two independent sources

In this section we focus on extractors for two independent sources. In Section 4.1 we suggest a way to get more mileage out of a given two source extractor. In Section 4.2 we prove the correctness of this construction. Finally, we show how to use this method to get improved two-source extractors in Section 4.3.

4.1. The transformation

We now show how to transform a given two-source extractor E that extracts few bits into a two-source extractor E' that extracts almost all the bits present in the source distribution.

The transformation below involves many parameters. The rough intuition is as follows: We start with a two-source extractor E that extracts t bits with error $\epsilon/2^t$. We require that this extractor is strong in the first source (that is that it extracts randomness from the second source). We use k_1 to denote the entropy threshold of the first source. We allow the extractor E to lose most of the bits that are present in the second source. More precisely, to extract t bits out of the second source we only require that the second source contains $t+\ell$ bits where ℓ (that may be much larger than t) is a parameter measuring the number of bits that E "loses". We now show how to transform E into an extractor E' which extracts many more bits with error ϵ. To achieve this goal we need the entropy thresholds of E' in the second source to be larger than that of E. More precisely, the entropy threshold of E' are k_1 (which is the same as E) and k_2 which needs to be somewhat larger than the entropy threshold of E. The gain is that the extractor E' extracts $k_1+k_2-(\ell+O(t+\log(1/\epsilon)))$ bits. If we set k_2 large enough so that ℓ, t and $\log(1/\epsilon)$ are small compared to k_1+k_2 we get that E' extracts almost all the k_1+k_2 bits of randomness that are present in the source. The exact details are given below.

Construction 22.

Parameters:

- n_1, n_2 *: The length of the two input sources.*
- k_1, k_2 *: The entropy threshold of the two input sources.*
- ϵ *: The required error.*

Goal: *Construct a* $(k_1,k_2;\epsilon)$*-two-source extractor* $E' : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *(for as large as possible* m*).*

Ingredients: *(note that the ingredients below involve additional parameters* t, ℓ, m_1, m_2*).*

- *A* $(k_1, t+\ell; \epsilon/2^{t+10})$*-two-source extractor* $E : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^t$ *that is strong in the first source.*[6]
- *A strong* $(k_1,\epsilon/4)$*-extractor* $E_1 : \{0,1\}^{n_1} \times \{0,1\}^t \to \{0,1\}^{m_1}$.
- *A strong* $(k_2,\epsilon/4)$*-extractor* $E_2 : \{0,1\}^{n_2} \times \{0,1\}^t \to \{0,1\}^{m_2}$.

Requirements: $0 \le m_2 \le k_2 - (2t+\ell+\log(1/\epsilon)+5)$.

Description of $E'(x_1,x_2)$
$= E_1(x_1, E(x_1,x_2)), E_2(x_2, E(x_1,x_2)), E(x_1,x_2)$.

Output length of E'**:** *Note that this gives that the output length of* E' *is* $m = m_1+m_2+t$*. If we use seeded extractors* E_1, E_2 *such that:* E_1 *extracts all the entropy from* X_1 *(apart from the unavoidable* $O(\log(1/\epsilon))$ *entropy loss) and* E_2 *satisfies the requirement above on* m_2 *with an equality, we get that* $m = k_1+k_2-(\ell+O(t+\log(1/\epsilon))$*. Thus, if* ℓ, t *and* $\log(1/\epsilon)$ *are small compared to* k_1+k_2 *we extract almost all the* k_1+k_2 *bits of randomness that are present in the source distribution.*

We now state a Theorem that says that E' is a two-source extractor. We find this quite surprising as the output $E(X_1,X_2)$ is used as a seed to operate on *both* sources X_1, X_2. This should be compared with the previous result of Dodis et al. [9] who observed that when E is strong in the first source one can use $E(X_1,X_2)$ as a seed to operate on X_1 (this is because $E(X_1,X_2)$ is essentially independent of X_1). Note that this is *not* the case in our setting and $E(X_1,X_2)$ is inherently dependent on (X_1,X_2).

Theorem 23. *Given parameters and ingredients as in Construction 22 then* E' *is a* $(k_1,k_2;\epsilon)$*-two-source extractor.*

We prove Theorem 23 in Section 4.2. Using off the shelf constructions of seeded extractors and fixing some of the parameters we obtain the following Corollary:

Corollary 24. *There is a universal constant* $A>0$ *such that: Let* $t \ge \log^4 n$*. Let* $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^t$ *be a* $(k_1, t+\ell; 2^{-1.1\cdot t})$*-two-source extractor that is strong in the first source. Then for any* $k_2 \ge \ell + At$ *we can construct a* $(k_1,k_2;2^{-\Omega(t)})$*-two-source extractor* $E' : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ *for* $m = k_1+k_2-(\ell+O(t))$*. Furthermore, if* E *is explicit then so is* E'.

The proof that Corollary 24 follows from Theorem 23 is omitted from this version. Note that for the particular setting of say $\ell = 100t$ (that is an extractor E that extracts a small yet constant fraction of the entropy present in the second source we get that by choosing k_2 such that $t =$

[6]That is, E is strong with respect to the function $\mathsf{Part}(x_1,x_2) = x_1$.

$o(k_2)$ the extractor E extracts a $1-o(1)$ fraction of the entropy.[7]

4.2. Proof of Correctness

We now prove Theorem 23. We are given the parameters and ingredients in Construction 22. Let $n = n_1 + n_2$. When given an n bit string x, we use x_1 to denote the first n_1 bits of x and x_2 to denote the last n_2 bits of x. With this notation E is a function from n bits to t bits. Let X_1 and X_2 be independent random variables over $\{0,1\}^{n_1}$ and $\{0,1\}^{n_2}$ such that $H_\infty(X_1) \geq k_1$ and $H_\infty(X_2) \geq k_2$. Let $X = (X_1, X_2)$. Our goal is to prove that $E'(X_1, X_2)$ is ϵ-close to the uniform distribution on $\{0,1\}^m$. Let Y be an independent random variable that is uniformly distributed over $\{0,1\}^t$. We first prove the following Lemma which asserts that $E(X_1, X_2)$ can replace Y and be used as a seed for E_2 to extract randomness from X_2.

Lemma 25. $(X_1, E_2(X_2, E(X_1, X_2)), E(X_1, X_2)) \sim_{\epsilon/4} (X_1, E_2(X_2, Y), Y)$.

Proof. We define a function $F : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^{m_2}$ by $F(x, y) = E_2(x_2, y)$. We define a function $\mathsf{Part} : \{0,1\}^n \to \{0,1\}^{n_1}$ by $\mathsf{Part}(x_1, x_2) = x_1$. Let $\mathcal{C}$ be the class of distributions X over n bit strings such that X_1, X_2 are independent, $H_\infty(X_1) \geq k_1$ and $H_\infty(X_2) \geq t + \ell$. The lemma will follow by applying Theorem 18 on $\mathcal{C}$, E, F and Part. We first note that X belongs to $\mathcal{C}$ because $k_2 \geq t + \ell$. We also observe that for every $y \in \{0,1\}^t$, $\mathsf{Part}(X) = X_1$ is independent of $F(X, y) = E_2(X_2, y)$. To apply the Theorem we need to check that the "weak closeness condition" holds. We set $\rho = \epsilon/2^{t+5}$ and our goal will be to show that the condition holds when choosing ρ as ϵ. More precisely, we need to show that: for every $y \in \{0,1\}^t$ there exists a set $G_y \subseteq \{0,1\}^{m_2}$ such that:

1. $\Pr[F(X, y) \notin G_y] \leq \rho$.
2. For every $a \in G_y$, $(X|F(X, y) = a)$ belongs to $\mathcal{C}$.

Note that as X_1, X_2 are independent we have that for every $y \in \{0,1\}^t$ and $a \in \{0,1\}^{m_2}$:

$$(X|F(X, y) = a) \sim ((X_1, X_2)|E_2(X_2, y) = a)$$

$$\sim (X_1 \otimes (X_2|E_2(X_2, y) = a))$$

This means that to prove that $(X|F(X, y) = a)$ belongs to $\mathcal{C}$ we only need to show that $H_\infty((X_2|E_2(X_2, y) = a)) \geq t + \ell$. To obtain G_y we use Lemma 13 choosing $R_1 = X_2$ and $R_2 = E_2(X_2, y)$ and using ρ. We indeed conclude that for every $y \in \{0,1\}^t$ there exists a set G_y such that:

1. $\Pr[R_2 \notin G_y] \leq \rho$.
2. For every $a \in G_y$, $H_\infty((X_2|E_2(X_2, y) = a)) \geq k_2 - (m_2 - \log(1/\rho)) \geq t + \ell$.

where the inequality follows because we have required that $m_2 \leq k_2 - (2t + \ell + \log(1/\epsilon) + 5) = k_2 - (t + \ell + \log(1/\rho))$ which gives the inequality above.

We have verified that the conditions of Theorem 18 are met and therefore using the second item of the Theorem we can conclude that

$$(\mathsf{Part}(X), F(X, E(X)), E(X)) \sim_{\rho \cdot 2^{t+3}} (\mathsf{Part}(X), F(X, Y), Y)$$

Which is exactly what we wanted:

$$(X_1, E_2(X_2, E(X)), E(X)) \sim_{\epsilon/4} (X_1, E_2(X_2, Y), Y)$$

□

We also need the following easy Lemma which asserts that applying Y as a seed on both X_1 and X_2 produces a distribution that is close to uniform.

Lemma 26. $(E_1(X_1, Y), E_2(X_2, Y), Y) \sim_{\epsilon/2} (U_{m_1} \otimes U_{m_2} \otimes U_t)$

Proof. We have that $(X_1, X_2, Y) \sim (X_1 \otimes X_2 \otimes U_t)$. Using Lemma 12 with the function $f(a_1, a_2, a_3) = a_1, E_2(a_2, a_3), a_3$ and using the fact that E_2 is a strong $(t + \ell, \epsilon/4)$-extractor we have that:

$$(X_1, E_2(X_2, Y), Y) \sim_{\epsilon/4} (X_1 \otimes U_{m_2} \otimes U_t)$$

We now use Lemma 12 with the function $f(a_1, a_2, a_3) = E_1(a_1, a_3), a_2, a_3)$ and using the fact that E_1 is a strong $(k_1, \epsilon/4)$-extractor we have that:

$$(E_1(X_1, Y), E_2(X_2, Y), Y) \sim_{(\epsilon/4 + \epsilon/4)} (U_{m_1} \otimes U_{m_2} \otimes U_t)$$

□

We are now ready to prove Theorem 23.

Proof. (of Theorem 23) By Lemma 25 we have that

$$(X_1, E_2(X_2, E(X)), E(X)) \sim_{\epsilon/4} (X_1, E_2(X_2, Y), Y)$$

We consider the function $f(a_1, a_2, a_3) = (E_1(a_1, a_3), a_2, a_3)$. By applying Lemma 12 on the triples above we get that:

$$(E_1(X_1, E(X)), E_2(X_2, E(X)), E(X))$$
$$\sim_{\epsilon/4} (E_1(X_1, Y), E_2(X_2, Y), Y)$$

Using Lemma 26 we conclude that:

$$(E_1(X_1, E(X)), E_2(X_2, E(X)), E(X)) \sim_{\epsilon} (U_{m_1} \otimes U_{m_2} \otimes U_t)$$

and the Theorem follows. □

[7] We are not making any attempt to present the most optimized or general corollary. The assumption that $t \geq \log^4 n$ is made so that there will be explicit constructions of strong seeded extractors that on entropy threshold k in a source of length n use a seed of length t to extract $k - O(t)$ bits with error $2^{-\Omega(t)}$ (e.g., [25]). We also remark that 1.1 above can be replaced with any constant larger than 1 and yield the same result.

4.3. Two source extractors which extract almost all the randomness

In this section we use the method above to get "more mileage" out of a recent extractor construction by Raz [24]. We use the following Theorem due to Raz. We first state the Theorem and then state a less general corollary which we use for our application.

Theorem 27. *[24] For any n_1, n_2, b_1, b_2, m and any $0 < \delta < 1/2$, such that: $n_1 \geq 6 \log n_1 + 2 \log n_2$, $b_1 \geq (1/2 + \delta)n_1 + 3 \log n_1 + \log n_2$, $b_2 \geq 5 \log(n_1 - b_1)$, $m \leq \delta \cdot min(n_1/8, b_2/40) - 1$ there is an explicit $(b_1, b_2; 2^{-1.5m})$-two-source extractor $E : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ that is strong in the first source.*

We use the following Corollary of the Theorem above:

Corollary 28. *For every $0 < \delta < 1/2$ and n such that $\delta n \geq 10 \log n$, let $b_1 = (1/2 + \delta)n$ and let b_2 be an integer such that $b_2 \geq 5 \log n$. There is an explicit $(b_1, b_2; 2^{-1.5t})$-two-source extractor $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^t$ with $t = \delta b_2/100$. Furthermore, E is strong in the first source.*

Applying the method of Section 4.1 gives the following Theorem (the proof is omitted from this version).

Theorem 29. *There is a universal constant $B > 0$ such that for every constant $0 < \delta < 1/2$ there exists a constant $C > 0$ such that: Let n be large enough, let $\epsilon \leq 2^{-\log^4 n}$, let $k_1 = (1/2 + \delta)n$, and let $k_2 \geq \frac{B \log(1/\epsilon)}{\delta}$. Then there is an explicit $(k_1, k_2; \epsilon)$-two-source extractor $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ with $m = k_1 + k_2 - C \log(1/\epsilon)$.*

Remark 30. *We made no attempt to present the most general result and Theorem 29 can be generalized in several respects. For example, we do not have to require that $\delta > 0$ is a constant and in that case $C = O(1/\delta)$ and the Theorem holds as long as $\delta n \geq 10 \log n$.*

In particular we obtain Theorem 6 as a corollary.

5. Extractors for Samplable distributions

In this section we focus on extractors for distributions samplable by poly-size circuits. In Section 5.2 we suggest a way to get more mileage out of a given extractor. In Section 5.3 we prove the correctness of this construction. Finally, we show how to use this method to get improved extractors in Section 5.4.

5.1. Preliminaries on Σ_i-circuits

We start with defining Σ_i-circuits.

Definition 31 (Σ_i-circuit). *A Σ_i-circuit is a circuit which in addition to the standard boolean gates can also use gates that compute a function that is complete for Σ_i (e.g. QBF with i alternations).*

We need the following result on Σ_1-circuits which follows from results on "sampling of NP witnesses" [30, 15, 3].

Theorem 32 (Uniform Sampling of NP witnesses [3]). *For every circuit T over $\{0,1\}^n$ of size t and $\epsilon > 0$ there exists a Σ_1-circuit T' over $\{0,1\}^{poly(n)}$ of size $poly(n,t)$ such that T' is a sampler for the distribution $(Z|T(Z) = 1)$ where Z is random variable that is uniformly distributed over $\{0,1\}^n$.*

5.2. The transformation

We now show how to transform a given extractor E that extracts few bits out of samplable distributions into an extractor E' that extracts almost all the bits present in the source distribution.

Trevisan and Vadhan [32] showed how to transform an extractor which extracts $t >$ polylog(n) random bits into one which extracts many more bits. More precisely, implicit in the paper is a general transformation (inspired by Goldreich and Wigderson [14], see also Reingold, Vadhan and Wigderson [26]). A drawback of this transformation is that it only works when the entropy threshold is very high (say $k = (1 - \nu)n$ where $\nu > 0$ is some small constant). Loosely speaking, when given an extractor E for samplable distributions with entropy threshold $k = (1 - \nu)n$ the paper shows how to construct an extractor E' which extracts $m = (1 - O(\nu)n)$ random bits. It is important to notice the following subtlety: To make the argument go through and deduce that E' works on distributions samplable by *deterministic circuits*, one needs to assume that E extracts randomness even from distributions that are samplable by Σ_1-*circuits*.[8]

[8] Loosely speaking, the idea is that when the entropy threshold $k = (1 - \nu)n$ for a small $\nu > 0$ then one can partition the n bit source into two "blocks", X_1, X_2 where X_1 is of length say $(1 - 100\nu)n$ and X_2 is of length $100\nu n$, and it follows that there must be entropy in both blocks. More precisely it follows that X_1, X_2 form a block-wise source (block-wise sources were defined by [6]). Following [22] (see also [19, 20, 28]) in this setting one can hope to extract bits from X_2 and use the bits extracted as a seed for a seeded extractor that extracts all the randomness in X_1. A subtlety is that for this argument to go through it is not sufficient that the initial extractor E can extract randomness from X_2 but rather that for every value x_1 of X_1, E can extract randomness from $(X_2|X_1 = x_1)$. The fact that X_2 is samplable by deterministic circuits does not necessarily means that the distribution $(X_2|X_1 = x_1)$ is samplable by deterministic circuits. Nevertheless, it does follow from Theorem 32 that the distribution $(X_2|X_1 = x_1)$ is samplable by Σ_1-circuits. This explains why the transformation requires that E works for distributions samplable by Σ_1-circuits. We want to also stress that the argument above completely fails when $k < n/2$ as then there is no way to partition the source into two blocks and ensure that both of them "contain" entropy.

We now present a different transformation which has essentially the same flavor (that is it transforms an extractor that extracts few bits from distributions samplable by Σ_1-circuits into an extractor that extracts many bits from distributions that are samplable by deterministic circuits). The advantage is that this transformation works even when the min-entropy threshold k is very small.

In the transformation below we start with an extractor E that on entropy threshold $t+\ell$ loses ℓ bits and extracts only t bits. We transform this extractor into an extractor E' which for entropy threshold sufficiently larger than $t+\ell$ extracts essentially all the randomness from the source.

Construction 33.

Parameters:

- n *: The length of the input sources.*
- k *: The entropy threshold of the input sources.*
- ϵ *: The required error.*
- s *: The size bound on the sampling circuit.*

Goal: *Construct a* $(k;\epsilon)$*-extractor* $E' : \{0,1\}^n \to \{0,1\}^m$ *for distributions samplable by size* s *circuits (for as large as possible* m*).*

Ingredients: *(note that the ingredients below involve additional parameters* t, ℓ, m_1, s'*).*

- *A* $(t+\ell, \epsilon/2^{t+10})$*-extractor* $E' : \{0,1\}^n \to \{0,1\}^t$ *for distributions samplable by size* s' Σ_1*-circuits.*
- *An explicit strong* $(k, \epsilon/2)$*-extractor* $E_1 : \{0,1\}^{n_1} \times \{0,1\}^t \to \{0,1\}^{m_1}$.

Requirements:

- $0 \le m_1 \le k - (2t + \ell + \log(1/\epsilon) + 5)$.
- $s' \ge p(s+n)$ *where* p *is some fixed polynomial to be determined later.*

Description of E': $E'(x) = E_1(x, E(x)), E(x)$.

Output length of E': *Note that this gives that the output length of* E' *is* $m = m_1 + t$*. If we use a seeded extractor* E_1 *such that* E_1 *satisfies the requirement above on* m_1 *with an equality, we get that* $m = k - (\ell + O(t + \log(1/\epsilon))$*. Thus, if* ℓ, t *and* $\log(1/\epsilon)$ *are small compared to* k *we extract almost all the* k *bits of randomness that are present in the source distribution.*

Theorem 34. *Given parameters and ingredients as in Construction 33 then* E' *is a* (k,ϵ)*-extractor for distributions samplable by size* s *circuits.*

5.3. Proof of correctness

We now prove Theorem 34. We are given the parameters and ingredients in Construction 33. Let X be a random variable taking values in $\{0,1\}^n$ that is samplable by a size s circuit C and $H_\infty(X) \ge k$. Our goal is to show that $E'(X)$ is ϵ-close to the uniform distribution on m bit strings. We start with the following lemma:

Lemma 35. *There exists some polynomial* p *such that For every* $y \in \{0,1\}^t$ *and* $a \in \{0,1\}^{m_1}$ *the distribution* $(X|E_1(X,y) = a)$ *is samplable by a size* $p(s+n)$ Σ_1*-circuit.*

Proof. Consider the circuit

$$T(z) = \begin{cases} 1 & E_1(C(z),y) = a \text{ and } C(z) \neq \perp \\ 0 & \text{otherwise} \end{cases}$$

Note that as E_1 is explicit, we have that $T(x)$ is a (deterministic) circuit of size poly(n,s). By Theorem 32 we get that there exists a Σ_1 circuit T' of size poly(n,s) and an independent random variable Z that is uniformly distributed such that T' is a sampler for the distribution $(Z|T(Z) = 1)$. Consider the Σ_1-circuit

$$T''(z) = \begin{cases} C(T'(z)) & T'(z) \neq \perp \\ \perp & T'(z) = \perp \end{cases}$$

It follows that T'' is a sampler for

$$(C(Z)|T(Z) = 1) \sim (C(Z)|E_1(C(Z),y) = a \text{ and } C(z) \neq \perp)$$

$$\sim (X|E_1(X,y) = a)$$

and thus this distribution is samplable by a Σ_1-circuit of size $poly(n+s)$. ☐

We now prove Theorem 33

Proof. (of Theorem 33) We define a function $F : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^{m_1}$ by $F(x,y) = E_1(x,y)$. Let $\mathcal{C}$ be the class of distributions over n bit strings that are samplable by size s' Σ_1-circuits and have min-entropy at least $t+\ell$. Let $\mathsf{Part}(x)$ be a constant function. The theorem will follow by using Theorem 18 on $\mathcal{C}$, E and F. We first note that X belongs to $\mathcal{C}$ because $k \ge t+\ell$ and $s \le s'$. To apply the Theorem we need to check that the "weak closeness condition" holds. We set $\rho = \epsilon/2^{t+5}$ and our goal will be to show that the condition holds when choosing ρ as ϵ. More precisely, we need to show that: for every $y \in \{0,1\}^t$ there exists a set $G_y \subseteq \{0,1\}^{m_2}$ such that:

1. $\Pr[F(X,y) \notin G_y] \le \rho$.
2. For every $a \in G_y$, $(X|F(X,y) = a)$ belongs to $\mathcal{C}$.

To obtain G_y we use Lemma 13 choosing $R_1 = X$ and $R_2 = E_1(X,y)$ and using ρ. We indeed conclude that for every $y \in \{0,1\}^t$ there exists a set G_y such that:

1. $\Pr[R_2 \notin G_y] \leq \rho$.
2. For every $a \in G_y$, $H_\infty((X|E_1(X,y) = a)) \geq k - (m_1 - \log(1/\rho)) \geq t + \ell$.

where the inequality follows because we have required that $m_1 \leq k-(2t+\ell+\log(1/\epsilon)+5) = k-(t+\ell+\log(1/\rho))$ which gives the inequality above. By Lemma 35 we have that for every $a \in G_y$, $(X|E_1(X,y) = a)$ is samplable by a size $p(s+n) \leq s'$ Σ_1-circuit. Thus, we conclude that $(X|F(X,y) = a)$ belongs to $\mathcal{C}$.

We have verified that the conditions of Theorem 18 are met. We add an independent random variable Y that is uniformly distributed over $\{0,1\}^t$ to our probability space. By the second item of the Theorem we can conclude that:

$$(F(X,E(X)),E(X)) \sim_{\rho \cdot 2^{t+3}} (F(X,Y),Y)$$

Or in other words that:

$$(E_1(X,E(X)),E(X)) \sim_{\epsilon/4} (E_1(X,Y),Y) \sim_{\epsilon/2} (U_{m_1} \otimes U_t)$$

□

5.4. Extractors for samplable distributions that extract almost all the randomness

In order to use the transformation of Theorem 34 we need an extractor for distributions samplable by Σ_1-circuits. By observing that Lemma 10 relativizes we can "go up a level in the hierarchy" and get the following corollary:

Corollary 36. *Let $G : \{0,1\}^n \to \{0,1\}^t$ be a function such that for any Σ_2-circuit D of size s:*

$$|\Pr[D(X,G(X)) = 1] - \Pr[D(X,Y) = 1]| \leq \epsilon$$

where X and Y are independent random variables distributed uniformly over $\{0,1\}^n$ and $\{0,1\}^t$ respectively. If $t \leq \log s - 1$ then for any $\Delta > 0$, G is a $(n-\Delta, O(2^\Delta \epsilon))$-extractor for distributions samplable by size $(2^\Delta \epsilon s)^{\Omega(1)}$ Σ_1-circuits.

Combining Corollary 36 with Theorem 34 and using an explicit seeded extractor by [17] we obtain the following corollary:

Corollary 37. *For every constant $\alpha > 0$ there is a constant $C > 1$ such that suppose that there exist a function $G = \{G_n\}$ such that $G_n : \{0,1\}^n \to \{0,1\}^{C \log n}$ is computable in time poly(n) and for any Σ_2-circuit D of size $s(n) > n^{3C}$:*

$$|\Pr[D(X,G(X)) = 1] - \Pr[D(X,Y) = 1]| \leq \epsilon(n)$$

where X and Y are independent random variables distributed uniformly over $\{0,1\}^n$ and $\{0,1\}^t$ respectively. then for every $k(n) \geq n - \log(1/\epsilon(n)) + O(C \log n)$ and constant $0 < \alpha < 1$ there is a function polynomial time computable function $E = \{E_n\}$ such that for every n, $E_n : \{0,1\}^n \to \{0,1\}^{m(n)}$ is a $(k(n), 1/100)$-extractor for distributions samplable by circuits of size $s(n)^{\Omega(1)}$ with $m(n) = (1-\alpha)k(n) - (n - \log(1/\epsilon(n)) + O(C \log n))$.

If the assumption of the Theorem holds with $\epsilon = 2^{-(1-\beta)n}$ for some constant $0 < \beta < 1$ then we get a $(k(n), 1/100)$-extractor that extracts $(1-\alpha)k(n) - \beta n - O(\log n)$ random bits.

We remark that we can improve the error of the final extractor from $1/100$ to $s(n)^{-\Omega(1)}$ (which by increasing $s(n)$ can be made n^{-c} for an arbitrary constant c). However, in this case we need to use a different seeded extractor which has both seed $O(\log n)$ and error $n^{-\Omega(1)}$. At the moment the best constructions for this setup [29, 17] do not extract $\Omega(k)$ bits but rather $k/\text{polylog}k$ bits and as a consequence our final extractor will extract fewer bits.

6. Discussion and open problems

We hope that presenting the transformation in general form will allow finding more applications. We now discuss several future directions for this research.

6.1. Can we get more mileage from seeded extractors?

We were able to get more mileage out of various kinds of deterministic extractors. It is natural to ask whether it is possible to get more mileage from seeded extractors. More specifically, the current situation in explicit construction of seeded extractors (see [28] for precise details) is that explicit constructions can optimize any of the two main parameters (seed length and output length) at the expense of losing in the other parameter. A natural approach to achieve optimality in both parameters simultaneously is to take an explicit extractor $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^t$ with t being sufficiently larger than d such that there is an explicit seeded extractor $E_1 : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ which extracts that maximum possible amount of random bits from the source. It is natural to ask whether $E'(x,y) = E_1(X, E_2(x,y))$ can yield an extractor. Following our paradigm one may ask whether there exist properties of E that will allow such a transformation to go through.

We now show that for any extractor E there is an extractor E_1 so that the suggeted transformation completely fails. In fact, we state the following more general lemma.

Lemma 38 (Impossibility result). *Let $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^t$ be any function and let $F : \{0,1\}^n \to \{0,1\}^t \to \{0,1\}^m$ be any function. There exists a function $F' : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m \cup \{\perp\}$ such that:*

- *For every random variable X over $\{0,1\}^n$, $F(X \otimes U_t) \sim_{2^{d-t}} F'(X \otimes U_t)$.*
- *For every $x \in \{0,1\}^n$ and $y \in \{0,1\}^d$, $F'(x, E(x,y)) = \perp$.*

The Lemma says for example that no matter how you pick E, if F is a seeded extractor then there exist a function F' that behaves like F when given a uniform seed (and therefore is also an extractor E_1 with essentially the same parameters). Furthermore, when trying to use $E(x,y)$ instead of a uniformly chosen seed in $E_1 = F'$ one always gets a worthless constant.

Proof. We define F' as follows:

$$F'(x,z) = \begin{cases} \perp & \exists y \in \{0,1\}^d : \ z = E(x,y) \\ F(x,y) & \text{otherwise} \end{cases}$$

It is easy to verify that the two conditions in the Lemma hold. ☐

6.2. Research directions for two-source extractors

Our technique may be useful to get more mileage out of two-source extractors for low entropy threshold once such extractors are explicitly constructed. Note that to apply our transformation directly we require extractors that are strong in the first source and can tolerate a low-entropy threshold in the second source.

We remark that there are recent constructions of two-source dispersers for entropy threshold $k = \Omega(n)$ [2]. We believe that our technique is also applicable in that setting. However, it seems that showing this may require specific properties of these dispersers. We do not know to apply our techniques directly on an arbitrary two-source disperser.

6.3. Research directions for samplable distributions

We showed that given an explicit extractor for distributions samplable by poly-size Σ_1-circuits we can get get more mileage from it (at least for distributions samplable by standard circuits). We have very few techniques for constructing such extractors. In particular, we do not have any technique to construct extractors for samplable distributions with error $\epsilon = n^{-\omega(1)}$. We find it interesting to achieve this goal using any plausible assumption.

Another important problem is to find minimal assumptions sufficient to construct extractors for samplable distributions and in particular obtain constructions that only rely on worst-case hardness. Trevisan and Vadhan [32] present such a construction (assuming worst-case hardness for Σ_5-circuits) but their construction only works for very large values of k. Interestingly, a barrier that prevents achieving constructions for lower entropy threshold is that any such construction (with certain black-box properties) translates into a two-source extractor. (This follows essentially in the same way that certain pseudorandom generators yield extractor [31] and was observed in [32]). It is natural to expect that recent advances in two-source extractors and dispersers can be translated to this framework and yield extractors for samplable distributions.

Finally, it is natural to ask whether the assumption of Lemma 10 is stronger than that of Theorem 9. We remark that following [21] it is often the case that a function that is hard on average entails a pseudorandom generator. However, this is not known for the range of parameters we are considering.

Acknowledgements I am grateful to Boaz Barak, Ariel Gabizon, Oded Goldreich, Russell Impagliazzo, Guy Kindler, Ran Raz, Omer Reingold, Benny Sudakov, Amnon Ta-Shma, Chris Umans, Salil Vadhan, Avi Wigderson and David Zuckerman for many discussions on various aspects of randomness extractors. I also want to thank Oded Regev and Amnon Ta-Shma for their encouragement. I am grateful to Zeev Dvir and Anup Rao for helpful comments.

References

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

[2] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson. Simulating independence: New constructions of condenesers, ramsey graphs, dispersers, and extractors. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 1–10, 2005.

[3] M. Bellare, O. Goldreich, and E. Petrank. Uniform generation of NP-witnesses using an NP-oracle. *INFCTRL: Information and Computation (formerly Information and Control)*, 163, 2000.

[4] M. Ben-Or and N. Linial. Collective coin flipping. *ADVCR: Advances in Computing Research*, 5:91–115, 1989.

[5] 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, 1984.

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

[7] B. Chor, O. Goldreich, J. Hastad, J. Friedman, 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.

[8] 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 14–25, 1989.

[9] Y. Dodis, A. Elbaz, R. Oliviera, and R. Raz. Improved randomnness extractors from two independent sources. In *Proccedings of thw 8th International Workshop on Randomization and Computation*, 2004.

[10] A. Elbaz. Improved constructions for extracting quasi-random bits from sources of weak randomness. Master's thesis, Weizmann institute, August 2003.

[11] A. Gabizon and R. Raz. Deterministic extractors for affine sources over large fields. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, 2005.

[12] 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.

[13] O. Goldreich. A sample of samplers - a computational perspective on sampling. *Electronic Colloquium on Computational Complexity (ECCC)*, 4(20), 1997.

[14] O. Goldreich and A. Wigderson. Tiny families of functions with random properties: A quality-size trade-off for hashing. *RSA: Random Structures and Algorithms*, 11, 1997.

[15] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43(2-3):169–188, 1986.

[16] 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.

[17] C. 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*, 2003.

[18] R. Motwani and P. Raghavan. *Randomized algorithms*. Cambridge University press, 1995.

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

[20] N. Nisan and A. Ta-Shma. Extracting randomness: A survey and new constructions. *Journal of Computer and System Sciences*, 58(1):148–173, 1999.

[21] N. Nisan and A. Wigderson. Hardness vs randomness. *Journal of Computer and System Sciences*, 49(2):149–167, Oct. 1994.

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

[23] J. Radhakrishnan and A. Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13(1):2–24, Feb. 2000.

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

[25] R. Raz, O. Reingold, and S. Vadhan. Error reduction for extractors. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 425–433, 1999.

[26] O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, 2000.

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

[28] R. Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, 77:67–95, 2002.

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

[30] L. Stockmeyer. The complexity of approximate counting. In ACM, editor, *Proceedings of the 15th Annual ACM Symposium on Theory of Computing*, 1983.

[31] L. Trevisan. Construction of extractors using pseudorandom generators. In *Proceedings of the 31st ACM Symposium on Theory of Computing*, 1999.

[32] 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.

[33] E. Upfal and M. Mitzenmacher. *Probability and comnputing*. Cambridge university press, 2005.

[34] S. Vadhan. Randomness extractors and their many guises. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 9–12, 2002.

[35] U. Vazirani. Efficient considerations in using semi-random sources. In *Proceedings of the 19th Annual ACM Symposium on the Theory of Computing*, pages 160–168, 1987.

[36] U. Vazirani. Strong communication complexity or generating quasi-random sequences from two communicating semi-random sources. *Combinatorica*, 7:375–392, 1987.

[37] U. V. Vazirani and V. V. Vazirani. Random polynomial time is equal to semi-random polynomial time. In *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science*, pages 417–428, 1985.

[38] J. von Neumann. Various techniques used in connection with random digits. *Applied Math Series*, 12:36–38, 1951.

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

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