Extracting Randomness from Samplable Distributions, Revisited

Marshall Ball
NYU
New York, USA
marshall.ball@cs.nyu.edu

Eli Goldin
NYU
New York, USA
eg3293@nyu.edu

Dana Dachman-Soled
University of Maryland
College Park, USA
danadach@umd.edu

Saachi Mutreja
Columbia University
New York, USA
saachi@berkeley.edu

***Abstract*—Randomness extractors provide a generic way of converting sources of randomness that are merely unpredictable into almost uniformly random bits. While in general, deterministic randomness extraction is impossible, it is possible if the source has some structural constraints.**

While much of the literature on deterministic extraction has focused on sources with strong independence properties, a natural class where deterministic extraction is possible is sources that can sampled by a polynomial size circuit, Levin [SIAM J Comp'86]. Trevisan and Vadhan [FOCS'00] explicitly constructed deterministic randomness extractors for this class of sources, assuming very strong circuit lower bounds.

We suggest that there is perhaps an even more reasonable model of natural sources of randomness than Levin's: sources sampled by polynomial size quantum circuits. Under a suitable circuit lower bound, we show that Trevisan and Vadhan's extractor indeed works for this class.

Along the way, we substantially improve their analysis in the classical case, showing that a circuit lower bound against NP-circuits suffice in the classical case (as opposed to a lower bounds on Σ_5-circuits, as shown by Trevisan and Vadhan). Moreover, we show that under this assumption, it is possible to handle sources sampled by postselecting circuits (a variant of nondeterministic circuits). We show that this model is sufficient to capture randomness extraction in the presence of efficiently computable leakage.

***Index Terms*—average-case complexity, quantum computing, randomness in computing**

I. Introduction

Randomness is an essential resource in computing. It is necessary for nearly all cryptographic tasks, such as achieving semantically secure symmetric key encryption. Similarly, many fundamental tasks in the domain of distributed computing, such as byzantine agreement or testing equality of two strings with low communication, are impossible to achieve deterministically. In a similar vein, certain tasks in differential privacy are also impossible without randomness.

Yet, where do these random bits come from? When constructing randomized protocols or procedures, the protocol/procedure designer almost always assumes access to independent, unbiased random bits. However, natural sources of randomness available to our machines are almost invariably far from such idealized sources of randomness, and moreover the particulars of their distributions are unknown to us. The question then becomes what algorithmic tasks can be accomplished with access to weakly random sources. Randomness extractors provide a generic means of deterministically converting a weakly random source into (almost) uniformly random independent bits, so that we may use constructions in our idealized models. This motivates the following general question:

> Can we deterministically extract uniformly random bits from naturally occurring weakly random sources?

It is well known that deterministic extraction from arbitrary weakly random sources is impossible, but is possible if the sources have some structure. While one rich line of work on deterministic (or seedless) randomness extractors has studied sources with strong independence properties (known as two-source extractors), it is unclear if naturally occurring sources can be assumed to be independent.[1] The extended Church-Turing thesis motivates another model for naturally occurring entropic sources: *sources sampled by polynomial size circuits*. [Lev86] Indeed, Trevisan and Vadhan constructed efficient extractors for such classes. [TV00]

The starting point of the present work is the simple observation that the universe, and hence natural sources, are generated by quantum phenomena. And even if quantum computing never materializes in practice, it is quite plausible that local natural physical phenomena cannot be efficiently simulated by classical circuits, but can be efficiently simulated by *theoretical* quantum circuits.

A. Our Results

We demonstrate that it is possible to deterministically (classically) extract random bits from weakly random sources sampled by polynomial size quantum circuits, assuming lower bounds on quantum circuits with postselection, a nondeterministic analog of quantum circuits. In fact, we show it is possible to deterministically extract random bits from a significantly larger, nondeterministic class of sources. Importantly, this implies the ability to extract randomness after observing arbitrary efficiently computed leakage on the source (provided some entropy remains after seeing the leakage).

Dana Dachman-Soled is supported in part by NSF grants CNS-2154705 and CNS-1933033.

[1] Moreover, even very limited quantitative relaxations of independence quickly render extraction impossible. [CG85], [BGM22]

Additionally, we improve what is known in the classical case. We show that it is possible to extract randomness from sources that are samplable by polynomial size circuits assuming lower bounds on nondeterministic circuits (as opposed to circuits with gates computing Σ_5-complete problems, as is the case in [TV00]).[2] Again, we show that indeed it is possible to extract from a larger class that includes sources uniform over a set recognized by a polynomial size circuit (also known as recognizable sources [Sha09]), from the same assumption. Prior to our work, similar results were only known assuming lower bounds on Σ_3-circuits. [AASY16][3]

In all cases, we can extract almost all the randomness from sources with linear min-entropy (there exists $\gamma > 0$ such that for all x, $\Pr[X = x] \geq 2^{(1-\gamma)n}$, where n is the length of the source). Unfortunately, like all prior work, the output of our extractors is only inverse-polynomially close to uniform.[4]

a) Classical and quantum postselecting samplers.: We consider a notion of nondeterministic samplers that generalizes samplable sources. We say that a source X (supported on $\{0,1\}^n$) is *sampled by a postselecting circuit* C (in a class $\mathcal{C}$), if C outputs $n+1$ bits, $C \to (x, b)$ such that X is identically distributed to the distribution sampled by C, conditioned on $b = 1$, namely $\Pr[X = x'] = \Pr_{C \to (x,b)}[x = x' | b = 1]$ for all x'. In particular, we are concerned with the case that the class $\mathcal{C}$ is either (randomized) polynomial size classical circuits, in which case we say the source is samplable by postselecting circuits, or $\mathcal{C}$ is polynomial size quantum circuits (with sufficient min-entropy), in which case we say the source is samplable by postselecting quantum circuits.

Note that sources sampled by postselecting classical polynomial size circuits correspond to sources sampled by polynomial size circuits whose random bits may themselves be drawn uniformly from a set recognized by a polynomial size circuit. Clearly, this class generalizes both samplable and recognizable sources.[5]

A motivation for considering such classes of sources is that they capture samplable sources induced by external observation or side-channel leakage. For example, it is unlikely that a physical source exists in a vacuum and is only observed by the extractor itself. If the extractor works for nondetermnistic samplable sources, then so long as the source has enough conditional min-entropy, then the output of the extractor will be independent of the leakage (and safe to use in a sensitive task).

b) Nondeterministic circuit models and hardness.: Before stating our results, we must briefly describe the circuit classes we assume hardness against.

A classical nondeterministic circuit, C can be thought of as a deterministic circuit C' that takes input, x and a witness, w: for any input x, $C(x) = 1$ if and only there exists w such that $C'(x, w) = 1$.

In the quantum regime, we are concerned with a fairly strong analog of nondeterminism: quantum circuits with postselection [Aar04]. These are quantum circuits (with classical description) that can *condition on a measurement being 1* before the output is measured. We say that such a circuit decides a language if such a circuit (conditioned on the first measurement outcome being 1) disagrees with the language on any input x with probability at most $1/3$.

Both of these circuit classes are quite strong. In particular, uniform polytime quantum computation with postselection, PostBQP is known to be equivalent to PP [Aar04]. However, it is nonetheless reasonable to conjecture that there are classical deterministic computations which do not admit superpolynomial speedups even if the computation is both non-uniform and nondeterministic or non-uniform, quantum, and postselecting.

The former classical assumption has been considered before in the context of derandomizing AM. [MV99a] We are not aware of a situation where the latter assumption has been made, but the connection with PP gives a classical interpretation: a set admits a postselecting quantum circuit family if and only if there is a family of randomized classical circuits that accept every string in the language with probability strictly greater than $1/2$, and reject every string not in the set with probability at least $1/2$.

c) Main Theorems.: Now we can state our results. Our main classical result is the following:

Informal Theorem 1 (Extractors for Classical Sources (Theorem II.4))**.** If there is a problem in $E = DTIME(2^{O(n)})$ with nondeterministic circuit complexity $2^{\Omega(n)}$, then for any constant c, there is an explicit deterministic extractor for sources samplableleable by size n^c postselecting circuits with linear min-entropy (whose output is $1/\mathrm{poly}(n)$-close to uniform).

Our main quantum theorem is the following:

Informal Theorem 2 (Extractors for Quantum Sources (Theorem II.5))**.** If there is a problem in $E = DTIME(2^{O(n)})$ with postselecting quantum circuit complexity $2^{\Omega(n)}$, then for any constant c, there is an explicit deterministic extractor for sources samplable by size n^c postselecting quantum circuits with linear min-entropy (whose output is $1/\mathrm{poly}(n)$-close to uniform).

We remark that regardless of whether this strong hardness assumption is true, explicit hard functions for postselecting quantum circuits are *required* to extract from this source class.

In both cases, our extractor is essentially the same extractor as that of Trevisan and Vadhan. [TV00] If f is an E-complete problem, $\widetilde{f}$ is its low degree extension, and $2Ext$ is a sufficiently good two-source extractor, our extractor will

[2] If the samplable source has very high min entropy, $n - O(\log n)$, then it was known how to extract from hardness against non-deterministic circuits. [TV00]

[3] Again, if the recognizable source was known to have very high min-entropy, $n - O(\log n)$ it is known how to extract from lower bounds on *deterministic* circuits. [LZ19].

[4] Applebaum et al. showed that this inherent in all black-box nondeterministic reductions. [AASY16]

[5] Guo et al. [GVJZ23] consider a similar analog in the algebraic setting: sources sampled by polynomials evaluated on varieties (generalizing polynomial sources [DGW07] and variety sources [Dvi09].

simply be

$$EXT(x,i) = 2EXT(\widetilde{f}(x), i).$$

Where our result differs from Trevisan and Vadhan's is that we give a novel analysis of the extractor. At the core of our analysis are new nondeterministic algorithms for an optimal parameter agnostic learning problem we call *gap probability maximization*. We refer the reader to the detailed technical overview below for details. For a complete proof, see the full version of our paper [BDSGM23].

II. Detailed Technical Overview

In this section, we explain in detail our approach to lifting Trevisan and Vadhan's proof that a hard function for Σ_5-circuits gives a good extractor for samplable sources [TV00] to the quantum realm. Through this explanation, it will become clear how we extend this result to nondeterministically samplable sources, as well as how we reduce the Σ_5 hardness requirement all the way to Σ_1.

As a warmup, we will describe how to lift Trevisan and Vadhan's proof that a boolean function f hard *on average* for NP-circuits is itself a good extractor [TV00]. The classical argument for this goes as follows: Let $\mathcal{S}$ be a flat (i.e. all outputs in the support have equal probability) source biasing f to 1. Then the following NP-circuit can compute $f(x)$:

On input x, nondeterministically check if x is in the range of $\mathcal{S}$, and if so output 1. Otherwise, output a random bit.

This approach can be augmented to non-flat $\mathcal{S}$ as long as we can solve the probability estimation problem, which asks that given a randomized circuit C and an output x, compute $\Pr_r[C(r) \to x]$ up to $(1 \pm \epsilon)$ multiplicative error. However, it is known that NP-circuits can solve the probability estimation problem in size polynomial in $size(C)$. This means that if f is hard for NP-circuits of size s, then it is an extractor for sources samplable by $s - O(n)$-size circuits for some concrete polynomial.

To extend this argument to the quantum world, all that is necessary is that we be able to do quantum probability estimation. That is, we need some model that can solve the following problem: given a quantum circuit C and an output x, compute $\Pr_r[C(r) \to x]$ up to $(1 \pm \epsilon)$ multiplicative error.

It turns out that to solve this problem for quantum circuits, we require quantum circuits with *postselection*. Postselection refers to the ability for algorithms to conditionally sample. In the quantum setting, this refers to the ability for quantum algorithms to produce the residual state resulting from measuring in the standard basis and receiving result 1. Thus, a postselecting circuit is a quantum circuit with the additional ability to postselect.

Quantum circuits with postselection are considered in depth by Aaronson in [Aar04]. It is not known how to implement postselection with a quantum computer, but it does not directly contradict the laws of quantum mechanics. In particular, Aaronson shows that PostBQP, the class of uniform postselecting circuits, is equivalent to PP.

Solving probability estimation using postselecting circuits implies that if a function is hard on average for quantum circuits with postselection, then it is an extractor for quantum samplable sources. However, it would be better to be able to show extraction from a worst-case assumption. In fact, Trevisan and Vadhan were able to extend their average case classical result to a worst-case hardness assumption, resulting in the following theorem

Theorem II.1. *If there is a problem in $E = DTIME(2^{O(n)})$ with Σ_5-circuit complexity $2^{\Omega(n)}$, then for every sufficiently small δ and for every s, there is a $(1-\delta, 1/n)$-extractor $EXT : \{0,1\}^n \to \{0,1\}^{1-O(\delta)n}$ against sources samplable by size s circuits. Furthermore, EXT is computable in time $poly(s)$ (with exponent depending on δ).*

As postselecting hardness was enough to lift the average case variation of this theorem to the quantum setting, we postulate (and will later prove) the following quantumization:

Proposition II.2. *If there is a problem in $E = DTIME(2^{O(n)})$ with postselecting quantum circuit complexity $2^{\Omega(n)}$, then for every sufficiently small δ and for every s, there is a $(1-\delta, 1/n)$-extractor $EXT : \{0,1\}^n \to \{0,1\}^{1-O(\delta)n}$ against sources samplable by size s quantum circuits. Furthermore, EXT is computable in time $poly(s)$ (with exponent depending on δ).*

One would hope that the same technique as in the average case hardness argument would apply when quantumizing this result. However, the proof for Theorem II.1 relies strongly on the fact that Σ_i circuits can do probability estimation for Σ_{i-1} circuits. In fact, the proof involves several instances of such "ladder-climbing", accomplishing some task on Σ_{i-1} circuits using Σ_i circuits.

It is not clear how one would do "ladder-climbing" for postselecting quantum circuits. One may hope that postselecting quantum circuits themselves can accomplish tasks like probability estimation for postselecting quantum circuits. Unfortunately, this seems unlikely to be true. We note that since $PostBQP = PP$, $PH \subseteq P^{PostBQP} = P^{\#P}$. This doesn't say anything definitive, but it is not clear how to reduce adaptive counting queries to a single threshold query. To provide more concrete evidence, we show in the full paper a concrete problem on quantum circuits, solvable by postselecting quantum circuits, for which the natural approach will not extend to a solution for postselecting quantum circuits.

One may wonder whether "ladder-climbing" is necessary to construct extractors from worst-case hardness assumptions. We show that it is not necessary, giving us the following improvement to Trevisan and Vadhan's classical result.

Proposition II.3. *If there is a problem in $E = DTIME(2^{O(n)})$ with nondeterministic circuit complexity $2^{\Omega(n)}$, then for every sufficiently small δ and for every s, there is a $(1-\delta, 1/n)$-extractor $EXT : \{0,1\}^n \to \{0,1\}^{1-O(\delta)n}$*

against sources samplable by size s circuits. Furthermore, EXT is computable in time $poly(s)$ (with exponent depending on δ).

Our proof for this improvement will indeed lift easily to the quantum setting, allowing us to prove Proposition II.2. In fact, in our main body we will prove both theorems simultaneously.

Note that Theorem II.1 has a nondeterminism gap. That is, we require hardness against Σ_5-circuits to obtain extractors for deterministic sources.

In fact, a small modification to our proof technique improves the result so that it provides extractors for *postselecting* (classically) samplable sources, and this improvement trivially lifts to the quantum setting. Informally, we call a source $\mathcal{S}$ a postselecting samplable source if there exists a circuit C outputting x, b such that $\mathcal{S}$ is the distribution on x conditioned on $b = 1$. That is, if $b = f(r)$ for some efficient f, we give the circuit the ability to uniformly sample from $f^{-1}(1)$. In general, this task can be implemented by an NP-circuit [JVV86], [BGP00], and so this class of sources is slightly weaker than those samplable by NP-circuits.[6]

We observe that both samplable and recongnizable sources are samplable by postselecting circuits. Any sampling circuit C gives a postselecting sampling circuit C' by setting $C'(r) = (C(r), 1)$. Any recognizing circuit C gives a postselecting sampling circuit C' by setting $C'(r) = (r, C(r))$.

Formally, our main results are captured by the following theorems:

Theorem II.4. *If there is a problem in $E = DTIME(2^{O(n)})$ with nondeterministic circuit complexity $2^{\Omega(n)}$, then for every sufficiently small δ and for every s, there is a $(1-\delta, 1/n)$-extractor $EXT : \{0,1\}^n \to \{0,1\}^{1-O(\delta)n}$ against sources samplable by size s postselecting circuits. Furthermore, EXT is computable in time $poly(s)$ (with exponent depending on δ).*

Theorem II.5. *If there is a problem in $E = DTIME(2^{O(n)})$ with postselecting circuit complexity $2^{\Omega(n)}$, then for every sufficiently small δ and for every s, there is a $(1-\delta, 1/n)$-extractor $EXT : \{0,1\}^n \to \{0,1\}^{1-O(\delta)n}$ against sources samplable by size s postselecting quantum circuits. Furthermore, EXT is computable in time $poly(s)$ (with exponent depending on δ).*

[Aar04] shows that a PP oracle can simulate postselecting quantum computation. Thus, we get the following corollary

Corollary II.6. *If there is a problem in $E = DTIME(2^{O(n)})$ with PP-circuit complexity $2^{\Omega(n)}$, then for every sufficiently small δ and for every s, there is a $(1-\delta, 1/n)$-extractor $EXT : \{0,1\}^n \to \{0,1\}^{1-O(\delta)n}$ against sources samplable by size s postselecting quantum circuits. Furthermore, EXT is computable in time $poly(s)$ (with exponent depending on δ).*

[6] In fact, it turns out that our notion of postselecting samplable sources is more powerful than sources samplable by "single-valued nondeterministic circuits," [MV99b], [SU01] a generalization of NP $\cap$ coNP to (a) computing functions, and (b) the non-uniform setting (which we won't formally define here). A simple rejection sampling argument implies that any extractor for the class of sources samplable by single-valued nondeterministic sources must be a hard to compute function for this class. It follows that hardness for this computational is indeed necessary in order to extract from postselecting samplable sources. Moreover, E being hard for exponential size nondeterministic circuits is in fact equivalent to E being hard for exponential size single-valued nondeterministic circuits. [SU01], [AKRR03]

A. Classical extractors from ladder-climbing

To generate an extractor from worst-case hardness, Trevisan and Vadhan rely on two ladder climbing techniques:

1) Probability estimation: Given a Σ_{i-1}-circuit C, there is a Σ_i-circuit estimating $\Pr_r[C(r) \to x]$ to a $(1 \pm \epsilon)$ multiplicative factor.
2) Uniform sampling: Given a boolean Σ_{i-1}-circuit C, there is a Σ_i-circuit sampling uniformly from $C^{-1}(1)$.

These ladder climbing techniques can prove the following two claims:

1) There is a very good worst-to-average case reduction for polynomial evaluation stepping up the ladder.
2) For functions $T(\cdot,\cdot)$ satisfying a "combinatorial list-decoding" property, there is a very efficient way to find points biasing $T(\cdot, \mathcal{S})$ for samplable $\mathcal{S}$

Stated more formally,

1) For a degree d polynomial $p : \mathbb{F}^t \to \mathbb{F}$, given a size-$s$ Σ_i which computes p correctly on a $c\sqrt{d/|\mathbb{F}|}$ fraction of points, there is a size-$poly(s)$ Σ_{i+2} circuit which computes p everywhere.
2) There exists a probabilistic Σ_{i+2}-circuit $DECODE$ of polynomial size such that the following holds: Let $\mathcal{S}$ be a source of density δ samplable by size-s Σ_i-circuits. Let $T(\cdot,\cdot)$ be a boolean function computable by size-$poly(n)$ circuits satisfying combinatorial list-decoding. If $T(w, C(r))$ is ϵ-biased to 1, then $C(\mathcal{S}, \epsilon) = w$ with probability $\Omega(\delta\epsilon^2)$

Trevisan and Vadhan use these claims to show that if you have a samplable distribution (X, I) of density δ which biases $T(p(x), i)$, then there is a Σ_5-circuit computing $p(x)$ everywhere. The approach here is simple: if I_x is the distribution I conditioned on $X = x$, then $DECODE(I_x)$ is a Σ_3-circuit computing $p(x)$ with some small probability. Then, the very efficient worst-to-average case reduction gives a Σ_5-circuit computing $p(x)$ everywhere. It also turns out that a sufficiently good two-source extractor $2EXT$ satisfies combinatorial list-decoding.

Putting all this together, we get that

$$EXT(x, i) := 2EXT(p(x), i)$$

is a good 1-bit extractor. Some additional care (but no further levels of nondeterminism) are required to extend this proof to multi-bit outputs.

For the proof of both of these claims, the two ladder climbing techniques described at the beginning of this section are used in sequence. For expository purposes, we will sketch the proof of the worst to average reduction for polynomial decoding.

B. Strong worst to average case reductions from ladder climbing

The proof they use for this claim relies on the following lemma from [STV99].

Lemma 1. Let C be any function and let $L_{z,x} := (1-t)z+x$ denote the line between z and x. Then there exists a $z \in \mathbb{F}$, $\gamma > 0$, such that for $15/16$ of the values of x,
- $\Pr_{\mathbb{F}\to u}[p(L_{z,x}(u)) = C(L_{z,x}(u))] \geq \gamma$
-For all univariate degree d $h : \mathbb{F} \to \mathbb{F}$ such that $h \neq p \circ L_{z,x}$, either $h(0) \neq z$ or $\Pr_{\mathbb{F}\to u}[h(u) = C(L_{z,x}(u))] \leq \frac{\gamma}{2}$.

Let C be a Σ_i circuit evaluating $p(x)$ on a $c\sqrt{d/|\mathbb{F}|}$ fraction of points.

We define C' to be the circuit which takes in a univariate h, chooses a random u from $\mathbb{F}$, and outputs 1 if $h(0) = z$ and $h(u) = C(L_{z,x}(u))$.

We further define C'' to be the Σ_{i+1}-circuit which takes an input x, gets an estimate $\widetilde{\gamma}$ for $\Pr[C'(x) = 1]$, and outputs 1 if $\widetilde{\gamma} \geq \frac{3}{4}\gamma$. The key lemma immediately shows that the only input accepted by C'' is $p \circ L_{z,x}$

It is then clear that running uniform sampling on C'' will find $p \circ L_{z,x}$ with high probability, and so outputting $(p \circ L_{z,x})(1)$ will find $p(x)$ with high probability.

C. Our techniques

As stated earlier, we will show that the extractor

$$EXT(x, i) = 2EXT(p(x), i)$$

defined by Trevisan-Vadhan still works when assuming hardness against nondeterministic circuits, and if we assume hardness against postselecting quantum circuits, this function will extract from quantum samplable sources.

Our key observation comes from the fact that the purpose of running uniform sampling and probability estimation in sequence is to solve a task we call the gap probability maximization problem. We define this problem as follows:

Say we are given a boolean randomized algorithm $\widetilde{C}$ and a constant γ with the following promise:

1) There exists some x^* such that $\Pr[\widetilde{C}(x^*) \to 1] \geq \gamma$
2) For all $x \neq x^*$, $\Pr[\widetilde{C}(x) \to 1] \leq \frac{\gamma}{2}$

The GPM problem asks us to find x^*.

We show in Section III-F that the gap maximization problem can be solved for an input circuit $\widetilde{C}$ by an NP-circuits C. Moreover, this circuit C only needs non-adaptive calls to the NP gates. This means that one only needs to step up the hierarchy once in order to achieve highly efficient worst to average case reductions for polynomial decoding, as well as bias finding for codes satisfying combinatorial list decoding. Thus, this observation immediately reduces our hardness assumption to hardness for Σ_3-circuits.

To get us all the way down to nondeterministic circuits, we note that gap probability maximization can be used to directly compute $p(x)$ using a source which biases our extractor. Once we have trimmed the layers nondeterminism produced by stacking uniform sampling and probability estimation, the extra layers of nondeterminism are purely an artifact of modularity. This gets us to NP-circuits that use their NP gates non-adaptively. From there, we can apply a result of Shaltiel and Umans [SU06] that implies that if E is hard for exponential size nondeterministic circuits, then E is hard for exponential size non-adaptive NP-circuits.

The essential ingredient for combining the two claims used by Trevisan and Vadhan is the following improved key lemma

Lemma 2. Let $p : \mathbb{F}^t \to \mathbb{F}$ be any function, let $T : \mathbb{F} \times \{0,1\}^s \to \{0,1\}^m$ satisfy sufficiently strong combinatorial list decoding, and let $\mathcal{S}$ be any distribution of density δ such that

$$\left|\Pr_{\mathcal{S}\to(u,i)}[T(p(u), i) = 1] - \frac{1}{2}\right| \geq \epsilon.$$

There exists constants c_0, c_1 such that if $\epsilon^{c_0}\delta^{c_1} \leq \sqrt{\frac{d}{q}}$ then the following holds: There exists a z such that for $\frac{15}{16}$ values of x,
- $\left|\Pr_{S\to(u,i)}[T(p(u), i) = 1 | u \in L_{z,x}] - \frac{1}{2}\right| \geq \frac{\epsilon}{3}$
- For all $h : \mathbb{F} \to \mathbb{F}$ such that $h \neq p \circ L_{z,x}$, either $\left|\Pr_{S\to(u,i)}[T(h(L_{z,x}^{-1}(u)), i) = 1 | u \in L_{z,x}] - \frac{1}{2^m}\right| \leq \frac{\epsilon}{6}$ or $h(0) \neq p(z)$.

Once this lemma has been proved, a gap probability maximization solver immediately gets us the result. Let $\mathcal{S}$ be some distribution of density δ biasing $EXT(x, i) = T(p(x), i)$. Our algorithm for evaluating $p(x)$ on operates as follows:

Algorithm 1: GPM evaluator for $p(x)$

We will define $C'(h)$ as follows: **if** $h(0) \neq p(z)$ **then**
  output 0.
end
Sample $\mathcal{S} \to (u, i)$.
Say the test passes if u lies on the line and $T(h(L_{z,x}^{-1}(u)), i) = 1$.
Output 1 if and only if the test passes $O(1/\epsilon)$ times.
The key lemma tells us that we can run gap probability maximization on C' to compute $p \circ L_{z,x}$.
Output $(p \circ L_{z,x})(1) = p(x)$.

We remark that the proof of this key lemma is highly non-trivial. Full details are included in the full version [BDSGM23]. Roughly, the first half of this lemma comes from a double application of Chebyshev's inequality. The second half of the lemma comes from proving a list decoding property,

which implies that the number of univariate polynomials which bias $T(h(L(u)), i)$ on a random line L is small.

To lift this result, all we need to do is show that gap probability maximization can be solved for quantum circuits using postselecting quantum circuits. We show this in the full paper.

D. Improvement to postselecting samplers

We remark that it is easy to extend our result to postselecting samplers. We simply replace the line "sample $\mathcal{S} \to (u, i)$" with "sample $\mathcal{S} \to (u, i, b)$ and fail if $b = 0$". The idea here is that nondeterminism allows us to condition our distribution for free, and so it costs nothing to add in the additional condition stemming from using a nondeterministic sampler.

E. Leakage Resilience

Finally, we consider a notion of leakage-resilient extractors against samplable sources. Informally, leakage resilience requires that the output of the extractor remain close to uniform even when some side information about the underlying source is revealed. Note that arbitrary leakage resilience is impossible, as the leakage could be the output of the extractor itself. However, as our goal is to model physical extractors, it is natural to consider leakage which is itself samplable.

It is natural to consider leakage any samplable function of the randomness source. However, we choose to consider a stronger notion, where the leakage is provided by the sampling circuit itself. This way, the leakage can depend on the randomness used to generate the source (including in the quantum setting). Note that providing resilience against arbitrary length leakage, even in this restricted model, is impossible as the leakage may simply be the underlying randomness used to generate the source. Thus, we additionally require that the source have high min-entropy conditioned on its leakage.

We show that given a deterministic extractor against nondeterministic samplable sources, either classical or quantum, then we get a leakage-resilient deterministic extractor for free. To show this, we rely on the fact that our extractor works against the source defined by conditioning the original source on the leakage being any particular value.

Theorem II.7. *Let EXT be a (k, ϵ)-deterministic extractor against nondeterministic sources samplable by size-s (quantum) circuits. Then, for all $c > 0$, EXT is a leakage-resilient $(k+c, \epsilon+2^{-c})$-deterministic extractor against sources samplable by size-$O(s)$ postselecting (quantum) circuits.*

We remark that in our notion of leakage-resilience, we only consider *classical* leakage. In the quantum setting, there is also a notion of min-entropy, which was defined originally in [Ren06]. This definition has been previously been used to capture randomness extraction [BFW12], [DPVR12], [BFSS14] in the presence of *quantum* side-information.

We do not make any claims about quantum leakage-resilience. If quantum computers do not exist, the power of quantum computing must come only from the real world. Therefore, any adversary wishing to use quantum side information to distinguish the output of an extractor from random must first make some efficient measurement, which equivalently could be made by the source directly. Nevertheless, constructing deterministic extractors against quantum samplable sources secure even in the presence of quantum side information is an interesting open question.

III. Preliminaries

A. Types of Nondeterministic Circuits

Let $\mathcal{P}$ be any complexity class and fix some $\mathcal{P}$-complete problem $\pi_{\mathcal{P}}$. A $\mathcal{P}$-circuit is a circuit with access to oracle gates for $\pi_{\mathcal{P}}$. We will primarily be concerned with Σ_i circuits. We will refer to Σ_1-circuits primarily as NP-circuits. The class of circuits that has all its Σ_1 gates in the same layer, i.e. circuits making SAT queries non-adaptively, is referred to as $\mathrm{NP}_{||}$-circuits.

We rely on a collapse theorem for E due to Shaltiel and Umans [SU06] of which the following is a special case:

Theorem III.1 (Corollary of [SU06, Theorem 3.2])**.** *If every language in E has $\mathrm{NP}_{||}$-circuits of size $s(n)$, then every language in E has non-deterministic circuits of size $s(n)^{O(1)}$.*

A quantum circuit of size s is a sequence of unitaries $U_1, \ldots, U_s$ where each U_i is taken from some universal gate set. A quantum circuit has an input register of length m, $\ell \leq s$ ancilla qubits, and an output register of length n. We use $C(x)$ to refer to the distribution on the output register of $(U_s \ldots U_1)(|x\rangle \otimes |0\rangle^{\otimes \ell} \otimes |0\rangle^{\otimes n})$ after measuring in the standard basis.

We will use the following formulation of postselecting quantum circuits. A postselecting quantum circuit C has the same format as a quantum circuit, except it has an additional postselection register. We use $C(x)$ to refer to the distribution on the output register of $(U_s \ldots U_1)(|x\rangle \otimes |0\rangle^{\otimes \ell} \otimes |0\rangle^{\otimes n})$ after measuring in the standard basis, conditioned on the measurement of the postselection register being 1. Note that here the size of a postselecting quantum circuit is the size of the corresponding quantum circuit.

Aaronson proved in [Aar04] that this model is equivalent to the model of quantum circuits with the ability to perform arbitrary postselections. Note that we must be somewhat careful here, as it is necessary that postselecting quantum circuits not be allowed to intersperse measurement and postselection.

Definition III.2. The Σ_i-circuit complexity of a boolean function f is the size of the smallest Σ_i-circuit C such that $C(x) = f(x)$ for all x.

Definition III.3. The postselecting quantum circuit complexity of a boolean function f is the size of the smallest postselecting quantum circuit C such that for all x,

$$f(x) = 1 \Rightarrow \Pr[C(x) = 1] \geq \frac{2}{3}$$

$$f(x) = 0 \Rightarrow \Pr[C(x) = 1] \leq \frac{1}{3}$$

Definition III.4. Let $\mathfrak{C}$ be a circuit model of computation and let L be a language. We define $f_n^L : \{0,1\}^n \to \{0,1\}$ by $f_n^L(x) = 1 \iff x \in L$. The $\mathfrak{C}$-complexity of L is the function $s(n) :=$ the $\mathfrak{C}$-complexity of f_n^L.

B. Min-entropy and density

Definition III.5. Let X, Y be random variables. We define the min-entropy of X conditioned on Y:

$$H_\infty(X|Y) := -\log \mathbb{E}_{Y \to y}[\max_x \Pr[X = x|Y = y]]$$

The unconditional min-entropy of X is the min-entropy of X conditioned on a constant. That is,

$$H_\infty(X) := -\log(\max_x \Pr[X = x])$$

Oftentimes, it is more convenient for us to reframe min-entropy from the framework of density. Formally,

Definition III.6. Let X be a random variable over some space $\mathcal{X}$. We say X has density δ if

$$\max_x \Pr[X = x] \leq \frac{1}{\delta |\mathcal{X}|}$$

Note that a random variable X over $\{0,1\}^n$ has density δ if and only if the min-entropy of X is $\geq n - \log \frac{1}{\delta}$. Density also satisfies the following useful property.

Proposition III.7. *Let (X, Y) be a random variable over some product space $\mathcal{X} \times \mathcal{Y}$. If (X, Y) has density δ, then X and Y both have density δ.*

To see this, observe that

$$\begin{aligned}\Pr[X \to x] &= \sum_y \Pr[(X, Y) \to (x, y)] \\ &\leq |Y| \frac{1}{\delta |X||Y|} = \frac{1}{\delta |X|}\end{aligned}$$

C. Classes of sources

A randomness source $\mathcal{S}$ is a distribution over some space $\mathcal{X}$. A class of sources $\mathfrak{S}$ is a set of randomness sources. We define several relevant classes of sources.

Definition III.8. We say that a distribution $\mathcal{S}$ is samplable by size-s circuits if there exists a circuit C of size s such that

$$\Pr_r[C(r) = x] = \Pr[\mathcal{S} \to x]$$

for all x.

Definition III.9. We say that a distribution $\mathcal{S}$ is samplable by size-s quantum circuits if there exists a quantum circuit C of size s such that

$$\Pr[C \to x] = \Pr[\mathcal{S} \to x]$$

for all x.

Note that as quantum circuits can sample their own randomness, we no longer need to quantify the probability that C outputs x by some randomness space.

We also define a notion of postselecting samplable sources. A source is samplable by postselecting circuits if we allow the circuit to condition on one of its outputs being 1.

Definition III.10. Let $\mathfrak{S}$ be a class of sources over $\mathcal{X} \times \{0,1\}$. We define a new class of sources $\mathfrak{S}_{nd}$ over $\mathcal{X}$, which we call postselecting $\mathfrak{S}$. We say that $\mathcal{S}' \in \mathfrak{S}_{nd}$ if there exists a source $(\mathcal{S}, b) \in \mathfrak{S}$ such that for all x',

$$\Pr_{\mathcal{S}' \to x}[x = x'] = \Pr_{\mathcal{S} \to (x,b)}[x = x'|b = 1].$$

Observe that when $\mathfrak{S}$ is the class of sources samplable by size-s quantum circuits, we see that $\mathfrak{S}_{nd}$ is the class of sources samplable by size-s postselecting quantum circuits.

D. Statistical Distance and Extractors

Definition III.11. For two distributions X, Y, the statistical distance between X and Y is

$$SD(X, Y) := \frac{1}{2} \sum_x |\Pr[X = x] - \Pr[Y = x]|$$

We say that X and Y are ϵ-close if $SD(X, Y) \leq \epsilon$.

Definition III.12. Let $\mathfrak{S}$ be some class of sources. We say that a function $EXT : \{0,1\}^n \to \{0,1\}^m$ is a (k, ϵ) (deterministic) extractor against $\mathfrak{S}$ if for every distribution $\mathcal{S} \in \mathfrak{S}$ such that $H_\infty(\mathcal{S}) \geq k$, $EXT(\mathcal{S})$ is ϵ-close to U_m.

Definition III.13. Let $\mathfrak{S}$ be some class of sources. We say that a function $EXT : \{0,1\}^n \to \{0,1\}^m$ is a (k, ϵ) leakage-resilient extractor against $\mathfrak{S}$ if for every distribution $(\mathcal{S}, L) \in \mathfrak{S}$ such that $H_\infty(\mathcal{S}|L) \geq k$, $EXT(\mathcal{S})$ is ϵ-close to U_m.

E. Combinatorial list decoding

Definition III.14. We say that a function $E : \{0,1\}^{n'} \times \{0,1\}^{n'} \to \{0,1\}^m$ satisfies (ϵ, δ, t)-combinatorial list decoding if, for all $a \in \{0,1\}^m$, the following holds:

Let $\mathcal{S}$ be any distribution over $\{0,1\}^{n'}$ of density δ. Then,

$$\left| \left\{ \omega : \left| \Pr_{\mathcal{S} \to i}[T(w, i) = a] - \frac{1}{2^m} \right| \geq epsilon \right\} \right| \leq t$$

Note that functions satisfying combinatorial list-decoding have been constructed in [TV00], [AASY16]. In particular, we have the following proposition

Claim 3 (Lemma 7.7 from [AASY16]. Also [DEOR04])**.** For all constants $c > 0$, there exists a constant $\alpha > 0$ such that for every sufficiently large n', for every $m \leq \alpha n'$ and $\epsilon \geq 2^{-cm}$, there exists a function $E : \{0,1\}^{n'} \to \{0,1\}^{n'} \to \{0,1\}^m$ satisfying $(\epsilon, 2^{-0.1n'}, 2^{0.2n'})$-combinatorial list decoding.

F. The Gap Probability Maximization Problem

We repeat our definition of the gap maximization problem (GPM) from the technical overview:

Say we are given a boolean randomized algorithm $\widetilde{C}$ and a constant γ with the following promise:
- There exists some x^* such that $\Pr[\widetilde{C}(x^*) \to 1] \geq \gamma$

- For all $x \neq x^*$, $\Pr[\widetilde{C}(x) \to 1] \leq \frac{\gamma}{2}$

The GPM problem asks us to find x^*.

We remark show that the GPM problem can be efficiently solved for classical circuits using NP-circuits. Formally,

Theorem 4. Let $\gamma, s > 0$, and let $\mathcal{C}$ be the class of boolean valued circuits of size s with input space X such that there exists an $x^* \in X$ satisfying
- $\Pr[\widetilde{C}(x^*) \to 1] \geq \gamma$,
- For all $x \neq x^*$, $\Pr[\widetilde{C}(x) \to 1] \leq \frac{\gamma}{2}$.

Then $C(\widetilde{C}) = x^*$. There exists a $poly(s)$-size $\mathrm{NP}_{||}$-circuit C such that for all $\widetilde{C} \in \mathcal{C}$, $C(\widetilde{C}) = x^*$.

In fact, the size of C will be polynomial in γ and s. But since $\gamma \leq 1$, we can thus upper bound the size of C by a polynomial in s. Note that this means the problem is easier for small values of γ. Intuitively, this is because the difficulty lies in reducing the number of witnesses for "bad" $x \neq x^*$.

Note that nondeterminism is indeed necessary to solve this problem, as such a C can be used to solve SAT. In particular, if $\widetilde{C}_\phi(x; r)$ is a circuit evaluating ϕ on input x, then for all ϕ with exactly one witness x^*, $C_1(\widetilde{C}_\phi) = x^*$.

A quantum analogue of this theorem also holds.

Theorem III.15. *Let $\gamma, s > 0$, and let $\mathcal{C}$ be the class of boolean valued quantum circuits of size s with input space X and ℓ ancilla qubits such that there exists an $x^* \in X$ satisfying*
- $\Pr[\widetilde{C}(x^*) \to 1] \geq \gamma$,
- *For all $x \neq x^*$, $\Pr[\widetilde{C}(x) \to 1] \leq \frac{\gamma}{2}$.*

Then there exists a $poly(s)$-size PostBQP-circuit C such that for all $\widetilde{C} \in \mathcal{C}$, $\Pr[C(\widetilde{C}) \to x^*] \geq \frac{2}{3}$.

We also remark that it is not necessary that these algorithms be given γ, since they can try every $\gamma = \frac{1}{2^i}$ for each $1 \leq i \leq s$ in time $poly(s)$.

Proofs of both these statements are included in the full version of the paper [BDSGM23].

IV. Extractors from hardness assumptions

In this section, we will prove the following theorem by relying on a key lemma.

Theorem IV.1. *Let $p : \mathbb{F}^t \to \mathbb{F}$ be any polynomial of degree d with $|\mathbb{F}| = q$. Let $T : \mathbb{F} \times \{0,1\}^s \to \{0,1\}$ satisfy $\left(\epsilon, \delta, \frac{1}{\delta \cdot \epsilon^2}\right)$-combinatorial list decoding.*

There exists constants c_0, c_1 such that if $\epsilon^{c_0}\delta^{c_1} \leq \sqrt{\frac{d}{q}}$ then the following holds:

If there exists a size-s postselecting (quantum/classical) samplable source S of density δ with output space $\mathbb{F}^t \times \{0,1\}^s$ such that

$$\left|\Pr_{S \to (x,i)}[T(p(x), i) = b] - \frac{1}{2}\right| \geq \epsilon$$

then there exists a (postselecting/$\mathrm{NP}_{||}$)-circuit C of size $poly\left(s, \frac{1}{\epsilon}\right)$ computing p everywhere.

This theorem is enough for us to build extractors from the hardness of E, netting us Theorems II.4 and II.5. First we instantiate Theorem IV.1 with the T from Proposition 3. Theorems II.4 and II.5 then follow by applying the same argument used by [TV00] in proving Theorem 5.8 from Theorem 5.3. For the classical case (Theorem II.4), Theorem IV.1 is about $\mathrm{NP}_{||}$-circuits, but Theorem II.4 assumes hardness against nondeterministic circuits. To close this gap, we apply Theorem III.1 that says: if E is hard for exponential size nondeterministic circuits, then E is hard for exponential size $\mathrm{NP}_{||}$-circuits.

To prove Theorem IV.1, we will rely on the following (restated) key lemma as well as the existence of an efficient implementation of a circuit solving the gap probability maximization problem. The proof of the key lemma is deferred to the full version [BDSGM23].

Lemma 5. KEY LEMMA:

Let $p : \mathbb{F}^t \to \mathbb{F}$ be any function, let $T : \mathbb{F} \times \{0,1\}^s \to \{0,1\}^m$ satisfy $\left(\epsilon, \delta, \frac{1}{\delta \cdot \epsilon^2}\right)$-combinatorial list decoding, and let $\mathcal{S}$ be any distribution of density δ such that

$$\left|\Pr_{\mathcal{S} \to (u,i)}[T(p(u), i) = 1] - \frac{1}{2}\right| \geq \epsilon.$$

There exists constants c_0, c_1 such that if $\epsilon^{c_0}\delta^{c_1} \leq \sqrt{\frac{d}{q}}$ then the following holds: There exists a z such that for $\frac{15}{16}$ values of x,
- $\left|\Pr_{S \to (u,i)}[T(p(u), i) = 1 | u \in L_{z,x}] - \frac{1}{2}\right| \geq \frac{\epsilon}{3}$
- For all $h : \mathbb{F} \to \mathbb{F}$ such that $h \neq p \circ L_{z,x}$, either $\left|\Pr_{S \to (u,i)}[T(h(L_{z,x}^{-1}(u)), i) = 1 | u \in L_{z,x}] - \frac{1}{2^m}\right| \leq \frac{\epsilon}{6}$ or $h(0) \neq p(z)$.

We now prove Theorem IV.1 from Lemma 5.

Proof. Without loss of generality, assume that

$$\Pr_{S \to (x,i)}[T(p(x), i) = 1] \geq \frac{1}{2} + \epsilon.$$

Let z be as in the key lemma applied to $p, \mathcal{S}$. We will develop a randomized algorithm $\widetilde{C}_x$ of size $poly(s, 1/\epsilon)$ taking in a univariate polynomial $h : \mathbb{F} \to \mathbb{F}$ of degree d such that $\Pr_{\mathcal{S} \to (x,i)}[\widetilde{C}_x(p \circ L_{z,x}) \to 1] \geq 2\Pr[\widetilde{C}_x(h) \to 1]$ for all $h \neq p \circ L_{z,x}$. Then, our circuit C will simply run a GPM solver on $\widetilde{C}$ to find $p \circ L_{z,x}$ with probability $\frac{15}{16}$, and then will output $(p \circ$

$L_{z,x})(1) = p(x)$. This gives us an efficient circuit evaluating $p(x)$ with probability $\frac{15}{16}$ with a circuit of size $poly(s, 1/\epsilon)$.

We define $\tilde{C}_x(h)$ as follows:

Algorithm 2: $\widetilde{C}_x(h)$

0: If $h(0) \neq p(z)$, output 0.

0: Sample $(u_1, i_1, b_1), \ldots, (u_k, i_k, b_k) \xleftarrow{\$} SAMP$.

0: Output 1 if and only if for all j:

- $b_j = 1$
- $u_j \in L_{z,x}(\mathbb{F})$ or
- $T(h(L_{z,x}^{-1}(u_j)), i_j) = 1$.

The key lemma implies that

$$\Pr[\widetilde{C}_x(p \circ L_{z,x}) \to 1] \geq \Pr[PASS] \left(\frac{1}{2} + \frac{\epsilon}{3}\right)^k$$

and

$$\Pr[\widetilde{C}_x(h') \to 1] \geq \Pr[PASS] \left(\frac{1}{2^m} - \frac{\epsilon}{6}\right)^k$$

for all $h' \neq p \circ L_{z,x}$. Then, for some $k = O(n^2)$, we can say that

$$\left(\frac{1}{2} + \frac{\epsilon}{3}\right)^k \geq 2 \left(\frac{1}{2^m} - \frac{\epsilon}{6}\right)^k$$

So nonuniformly fixing $\gamma = \Pr[PASS] \left(\frac{1}{2^m} + \frac{\epsilon}{3}\right)^k$ in Theorem 4 (or Theorem III.15 for the quantum case) gives us the result.

Observe that, $\widetilde{C}_x$ is of size $poly\left(s, \frac{1}{\epsilon}\right)$, and so we are done. Note that in order to run the GPM solver, we need C to be a $\mathrm{NP}_{||}$-circuit or a PostBQP-circuit, depending on whether we are operating in the classical or quantum world.

□

References

[Aar04] Scott Aaronson. Quantum computing, postselection, and probabilistic polynomial-time, 2004.

[AASY16] Benny Applebaum, Sergei Artemenko, Ronen Shaltiel, and Guang Yang. Incompressible functions, relative-error extractors, and the power of nondeterministic reductions. *computational complexity*, 25(2):349–418, Jun 2016.

[AKRR03] Eric Allender, Michal Koucký, Detlef Ronneburger, and Sambuddha Roy. Derandomization and distinguishing complexity. In *18th Annual IEEE Conference on Computational Complexity (Complexity 2003), 7-10 July 2003, Aarhus, Denmark*, pages 209–220. IEEE Computer Society, 2003.

[BDSGM23] Marshall Ball, Dana Dachman-Soled, Eli Goldin, and Saachi Mutreja. Extracting randomness from samplable distributions, revisited. arXiv, 2023. https://arxiv.org/.

[BFSS14] Mario Berta, Omar Fawzi, Volkher Scholz, and Oleg Szehr. Variations on classical and quantum extractors. In *2014 IEEE International Symposium on Information Theory*. IEEE, jun 2014.

[BFW12] Mario Berta, Omar Fawzi, and Stephanie Wehner. Quantum to classical randomness extractors. In Reihaneh Safavi-Naini and Ran Canetti, editors, *CRYPTO 2012*, volume 7417 of *LNCS*, pages 776–793. Springer, Heidelberg, August 2012.

[BGM22] Marshall Ball, Oded Goldreich, and Tal Malkin. Randomness extraction from somewhat dependent sources. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA*, volume 215 of *LIPIcs*, pages 12:1–12:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[BGP00] Mihir Bellare, Oded Goldreich, and Erez Petrank. Uniform generation of np-witnesses using an np-oracle. *Information and Computation*, 163(2):510–526, 2000.

[CG85] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity (extended abstract). In *26th Annual Symposium on Foundations of Computer Science, Portland, Oregon, USA, 21-23 October 1985*, pages 429–442. IEEE Computer Society, 1985.

[DEOR04] Yevgeniy Dodis, Ariel Elbaz, Roberto Oliveira, and Ran Raz. Improved randomness extraction from two independent sources. In Klaus Jansen, Sanjeev Khanna, José D. P. Rolim, and Dana Ron, editors, *Approximation, Randomization, and Combinatorial Optimization, Algorithms and Techniques, 7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2004, and 8th International Workshop on Randomization and Computation, RANDOM 2004, Cambridge, MA, USA, August 22-24, 2004, Proceedings*, volume 3122 of *Lecture Notes in Computer Science*, pages 334–344. Springer, 2004.

[DGW07] Zeev Dvir, Ariel Gabizon, and Avi Wigderson. Extractors and rank extractors for polynomial sources. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007), October 20-23, 2007, Providence, RI, USA, Proceedings*, pages 52–62. IEEE Computer Society, 2007.

[DPVR12] Anindya De, Christopher Portmann, Thomas Vidick, and Renato Renner. Trevisan's extractor in the presence of quantum side information. *SIAM Journal on Computing*, 41(4):915–940, jan 2012.

[Dvi09] Zeev Dvir. Extractors for varieties. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity, CCC 2009, Paris, France, 15-18 July 2009*, pages 102–113. IEEE Computer Society, 2009.

[GVJZ23] Zeyu Guo, Ben Lee Volk, Akhil Jalan, and David Zuckerman. Extractors for images of varieties. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 46–59. ACM, 2023.

[JVV86] Mark R. Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986.

[Lev86] Leonid A. Levin. Average case complete problems. *SIAM J. Comput.*, 15(1):285–286, 1986.

[LZ19] Fu Li and David Zuckerman. Improved extractors for recognizable and algebraic sources. In Dimitris Achlioptas and László A. Végh, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2019, September 20-22, 2019, Massachusetts Institute of Technology, Cambridge, MA, USA*, volume 145 of *LIPIcs*, pages 72:1–72:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[MV99a] Peter Bro Miltersen and N. V. Vinodchandran. Derandomizing arthur-merlin games using hitting sets. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99, 17-18 October, 1999, New York, NY, USA*, pages 71–80. IEEE Computer Society, 1999.

[MV99b] Peter Bro Miltersen and N. V. Vinodchandran. Derandomizing arthur-merlin games using hitting sets. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99, 17-18 October, 1999, New York, NY, USA*, pages 71–80. IEEE Computer Society, 1999.

[Ren06] Renato Renner. Security of quantum key distribution, 2006.

[Sha09] Ronen Shaltiel. Weak derandomization of weak algorithms: Explicit versions of yao's lemma. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity, CCC 2009, Paris, France, 15-18 July 2009*, pages 114–125. IEEE Computer Society, 2009.

[STV99] Madhu Sudan, Luca Trevisan, and Salil P. Vadhan. Pseudoran-

dom generators without the XOR lemma (extended abstract). In *31st ACM STOC*, pages 537–546. ACM Press, May 1999.

[SU01] Ronen Shaltiel and Christopher Umans. Simple extractors for all min-entropies and a new pseudo-random generator. In *42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, 14-17 October 2001, Las Vegas, Nevada, USA*, pages 648–657. IEEE Computer Society, 2001.

[SU06] Ronen Shaltiel and Christopher Umans. Pseudorandomness for approximate counting and sampling. *Comput. Complex.*, 15(4):298–341, 2006.

[TV00] Luca Trevisan and Salil P. Vadhan. Extracting randomness from samplable distributions. In *41st FOCS*, pages 32–42. IEEE Computer Society Press, November 2000.