Multiplicative Extractors for Samplable Distributions*

Ronen Shaltiel†

November 3, 2024

Abstract

Trevisan and Vadhan (FOCS 2000) introduced the notion of (seedless) extractors for samplable distributions as a possible solution to the problem of extracting random keys for cryptographic protocols from weak sources of randomness. They showed that under a very strong complexity theoretic assumption, there exists a constant $\alpha > 0$ such that for every constant $c \geq 1$, there is an extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{\Omega(n)}$, such that for every distribution X over $\{0,1\}^n$ that has $H_\infty(X) \geq (1-\alpha) \cdot n$, $\mathsf{Ext}(X)$ is ϵ-close to uniform for $\epsilon = \frac{1}{n^c}$, and furthermore, Ext is computable in time $\mathsf{poly}(n^c)$.

Recently, Ball, Goldin, Dachman-Soled and Mutreja (FOCS 2023) gave a substantial improvement, and achieved the same conclusion under the weaker (and by now standard) assumption that there exists a constant $\beta > 0$, and a problem in $\mathsf{E} = \mathsf{DTIME}(2^{O(n)})$ that requires size $2^{\beta n}$ nondeterministic circuits.

In this paper we give an alternative proof of this result with the following advantages:

- Our extractors have "multiplicative error". More specifically, it is guaranteed that for every event $A \subseteq \{0,1\}^m$, $\Pr[\mathsf{Ext}(X) \in A] \leq (1+\epsilon) \cdot \Pr[U_m \in A]$. (This should be contrasted with the standard notion that only implies $\Pr[\mathsf{Ext}(X) \in A] \leq \epsilon + \Pr[U_m \in A]$).

  Consequently, unlike the extractors of Trevisan and Vadhan, and Ball et al., our multiplicative extractors guarantee that in the application of selecting keys for cryptographic protocols, if when choosing a random key, the probability that an adversary can steal the honest party's money is $n^{-\omega(1)}$, then this also holds when using the output of the extractor as a key.

  A related notion of multiplicative extractors was defined by Applebaum, Artemenko, Shaltiel and Yang (CCC 2015) who showed that black-box techniques cannot yield extractors with additive error $\epsilon = n^{-\omega(1)}$, under the assumption assumed by Ball et al. or Trevisan and Vadhan. This motivated Applebaum et al. to consider multiplicative extractors, and they gave constructions based on the original hardness assumption of Trevisan and Vadhan.
- Our proof is significantly simpler, and more modular than that of Ball et al. (and arguably also than that of Trevisan and Vadhan). A key observation is that the extractors that we want to construct, easily follow from a seed-extending pseudorandom generator against nondeterministic circuits (with the twist that the error is measured multiplicatively, as in computational differential privacy). We then proceed to construct such pseudorandom generators under the hardness assumption. This turns out to be easier (utilizing amongst other things, ideas by Trevisan and Vadhan, and by Ball et al.)

Trevisan and Vadhan also asked whether lower bounds against nondeterministic circuits are *necessary* to achieve extractors for samplable distributions. While we cannot answer this question, we show that the proof techniques used in our paper (as well as those used in previous work) produce extractors which imply seed-extending PRGs against nondeterministic circuits, which in turn imply lower bounds against nondeterministic circuits.

*In memory of Luca Trevisan.

†University of Haifa, Email: ronen@cs.haifa.ac.il. Ronen Shaltiel was supported by ISF grant 1006/23.

Contents

1 Introduction

1.1 Multiplicative Pseudorandomness

Pseudorandomness is a viewpoint that says that a distribution Z over $\{0,1\}^m$ is "similar" to the uniform distribution U_m from the point of view of a function $C : \{0,1\}^m \to \{0,1\}$, if the quantities $p_1 = \Pr[C(U_m) = 1]$ and $p_2 = \Pr[C(Z) = 1]$ are "similar". Typically, this similarity is measured by choosing a parameter $0 < \epsilon \leq 1$ and using the relation $\overset{ad}{\sim}_\epsilon$ on $[0,1]$ defined as follows:

$$p_1 \overset{ad}{\sim}_\epsilon p_2 \iff |p_2 - p_1| \leq \epsilon,$$

We can generalize this approach to define pseudorandomness with respect to different relations.

Definition 1.1 (Pseudorandomness with respect to a relation)**.** *Let $\sim$ be a relation on $[0,1]$. Given a function $C : \{0,1\}^m \to \{0,1\}$, a distribution Z over $\{0,1\}^m$ is* **pseudorandom for C with respect to $\sim$,** *if*

$$\Pr[C(U_m) = 1] \sim \Pr[C(Z) = 1].$$

We will abbreviate "with respect to" as "w.r.t." for brevity. Given a class $\mathcal{C}$ of functions $C : \{0,1\}^m \to \{0,1\}$ we say that Z is **pseudorandom for $\mathcal{C}$ w.r.t. $\sim$,** *if it is pseudorandom for every C in $\mathcal{C}$ w.r.t. $\sim$. Z is* **close to uniform w.r.t. $\sim$,** *if it is pseudorandom w.r.t. $\sim$ for the class of all boolean functions on m bits.*

The standard notion of ϵ-pseudorandomness is obtained when taking the relation $\overset{ad}{\sim}_\epsilon$. If the class $\mathcal{C}$ is closed under complement then the standard notion is also obtained when using the (one sided) relation

$$p_1 \overset{a}{\sim}_\epsilon p_2 \iff p_2 \leq p_1 + \epsilon,$$

in which the absolute value is removed. The generalized formulation of Definition 1.1 allows other relations. This generality is used in differential privacy [DMNS06] that uses the following *multiplicative* relation:

$$p_1 \overset{m}{\sim}_\epsilon p_2 \iff p_2 \leq e^\epsilon \cdot p_1.$$

Note that for $0 \leq \epsilon \leq 1$, $e^\epsilon = 1 + \Theta(\epsilon)$, and therefore, pseudorandomness with respect to $\overset{m}{\sim}_\epsilon$, implies pseudorandomness with respect to $\overset{a}{\sim}_\epsilon$.[1] The field of differential privacy also considers a generalization of $\overset{m}{\sim}$ with two parameters: a "large" multiplicative ϵ, and a "small" additive δ, defined as follows:

$$p_1 \overset{m}{\sim}_{(\epsilon,\delta)} p_2 \iff p_2 \leq e^\epsilon \cdot p_1 + \delta.$$

Note that $\overset{m}{\sim}_\epsilon$ is obtained as $\overset{m}{\sim}_{(\epsilon,0)}$, and pseudorandomness w.r.t. $\overset{m}{\sim}_{(\epsilon,\delta)}$ implies pseudorandomness w.r.t. $\overset{a}{\sim}_{\epsilon+\delta}$. We call the pseudorandomness obtained by these relations "*multiplicative pseudorandomness*".[2] We can continue and define two fundamental objects of pseudorandomness (pseudorandom generators and seedless extractors) in this generalized way (which we will later use with the multiplicative relations).

Definition 1.2 (Pseudorandom generators and extractors w.r.t. a relation)**.** *Let $\sim$ be a relation on $[0,1]$.*

[1] Pseudorandomness w.r.t. $\overset{m}{\sim}_\epsilon$ makes sense also for $\epsilon \geq 1$, and such choices are sometimes used in differential privacy. However, in this paper we will only consider the case where $0 \leq \epsilon \leq 1$, so that $1 + \epsilon \leq e^\epsilon \leq 1 + 3\epsilon$.

[2] It is known that in the standard definition (w.r.t. $\overset{ad}{\sim}_\epsilon$ or $\overset{a}{\sim}_\epsilon$) Z is ϵ-close to uniform iff Z has statistical distance at most ϵ from U_m. The multiplicative notion also has a natural information theoretic meaning, specifically note that Z is close to uniform w.r.t. $\overset{m}{\sim}_\epsilon$ iff for every $z \in \{0,1\}^m$, $\Pr[Z = z] \leq e^\epsilon \cdot 2^{-m}$.

- *$G : \{0,1\}^d \to \{0,1\}^m$ is a* **PRG** *for a class $\mathcal{C}$ w.r.t. to $\sim$ (which we will also shorten to "$\sim$-PRG for $\mathcal{C}$") if $G(U_d)$ is pseudorandom for $\mathcal{C}$ w.r.t. $\sim$. G is* **seed-extending** *if the function $G'(x) = (x, G(x))$ is a PRG for the considered class $\mathcal{C}$, w.r.t. the considered relation $\sim$.*
- *A function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ is a $(k, \sim)$-***extractor** *for a class $\mathcal{D}$ of distributions over $\{0,1\}^n$, if for every distribution X in $\mathcal{D}$ with $H_\infty(X) \geq k$, the distribution $\mathsf{Ext}(X)$ is close to uniform w.r.t. $\sim$.*

Once again, the standard notions of extractors and pseudorandom generators are obtained for the relation $\overset{ad}{\sim}_\epsilon$ (the same holds also for $\overset{a}{\sim}_\epsilon$ for extractors, and also for PRGs in case $\mathcal{C}$ is closed under complement).

We will consider multiplicative variants of pseudorandom generators and seedless extractors. Some multiplicative versions of these objects (with a slightly different definition) have been considered before [AASY15, AIKS16, LZ19, SS24] (and we elaborate on these works below). The main contribution of this paper is improved constructions, under weaker hardness assumptions.

A motivating example: using seedless extractors to select keys for cryptographic protocols. Consider a cryptographic protocol which is known to be secure when the key of an honest party is chosen according to U_m. That is, the probability that an adversary can steal the honest party's money is smaller than some "negligible" $\alpha > 0$. A signature application of seedless extractors is choosing keys for cryptographic protocols by extracting randomness from weak random sources. (Note that seeded extractors do not apply for this application). When using a seedless extractor, the key will be "close to uniform" rather than "truly uniform".

If the key is chosen according to a distribution that is ϵ-close to uniform (using to the standard notion) then we are only guaranteed that the adversary's probability to cheat is smaller than $\alpha + \epsilon$, which may be unacceptable if ϵ is "large" compared to α.

In contrast, if we replace the standard notion by the multiplicative notion (w.r.t. $\overset{m}{\sim}_\epsilon$), then the probability that the adversary can cheat is bounded by $e^\epsilon \cdot \alpha \leq (1 + 3\epsilon) \cdot \alpha$, which is still very small even for constant ϵ. The same holds when using the multiplicative version with two parameters ϵ and δ (that is w.r.t. $\overset{m}{\sim}_{(\epsilon,\delta)}$), even if ϵ is large, as long as δ is sufficiently small and is comparable to α.

Indeed, this advantage of the multiplicative notion over the additive notion is the rational for using these multiplicative relations in differential privacy (where it is often impossible or expensive to obtain small ϵ).

1.2 Extractors for Samplable Distributions

An influential paper by Trevisan and Vadhan [TV00] introduced the notion of (seedless) extractors for samplable distributions. Their goal was to identify a class of distributions that contains "sources of randomness that are available to computers" and allows seedless extractors that run in poly-time.

We say that a distribution X over $\{0,1\}^n$ is *sampled* by a circuit $A : \{0,1\}^r \to \{0,1\}^n$ if $X = A(U_r)$ (See more formal definition in Section 3.1). Trevisan and Vadhan considered extractors for distributions that are samplable by poly-size circuits, namely distributions samplable by circuits of size n^c for some constant parameter c. They showed that such extractors cannot run in time smaller than n^c, and considered extractors that run in time $\mathsf{poly}(n^c)$. They showed that such extractors imply circuit lower bounds, and so, motivated by the hardness vs. randomness paradigm, they gave a conditional construction based on hardness assumptions.

Hardness assumptions against various types of nondeterministic circuits. We say that "E is hard for exponential size circuits of some type", if there exist a problem $L \in \mathsf{E} = \mathsf{DTIME}(2^{O(n)})$ and a constant $\beta > 0$, such that for every sufficiently large n, circuits of size $2^{\beta \cdot n}$ (of the specified type) fail to compute the characteristic function of L on inputs of length n. (See Section 3.3 for a more formal definition).

The assumptions that E is hard for exponential size (deterministic) circuits was used by the celebrated paper of Impagliazzo and Wigderson [IW97] to imply that $\mathsf{BPP} = \mathsf{P}$. The stronger assumption that E is hard

for exponential size nondeterministic circuits[3], originated in works on hardness versus randomness for AM, and is now standard, and used in many results [AK02, KvM02, MV05, SU05, BOV07, GW02, GST03, SU06, SU09, Dru13, AASY15, BV17, AIKS16, HNY17, DMOZ22, BDL22, CT22, BGDM23, BSS24, SS24]. It can be viewed as a scaled, nonuniform version of the widely believed assumption that $\mathsf{EXP} \neq \mathsf{NP}$.

In their seminal paper on extractors for samplable distributions, Trevisan and Vadhan [TV00] introduced a version of the assumption for a stronger circuit class. A Σ_i-circuit, is a circuit that in addition to the standard gates, is also allowed to use a special gate (with large fan-in) that solves the canonical complete language for the class Σ_i^{P} (the i'th level of the polynomial time hierarchy).[4] The extractor of Trevisan and Vadhan [TV00] relies on the extremely strong assumption that E is hard for exponential size Σ_5-circuits.[5]

Previous work on extractors for samplable distributions. The main result of Trevisan and Vadhan [TV00] is that under a hardness assumption for Σ_5-circuits, there is an extractor for distributions samplable by poly-size circuits with $k = n - \Delta$, where $\Delta = \alpha n$ for some constant $\alpha > 0$. Below is a precise statement.[6]

Theorem 1.3 ([TV00])**.** *If E is hard for exponential size Σ_5-circuits then there exists a constant $\alpha > 0$, such that for every constant $c > 1$, and for every sufficiently large n, there is a function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{\alpha n}$ that is a $((1-\alpha)\cdot n, \overset{a}{\sim}_\epsilon)$-extractor for distributions samplable by circuits of size n^c, where $\epsilon = n^{-c}$. Furthermore, Ext is computable in time $\mathsf{poly}(n^c)$.*[7]

Note that this extractor only achieves an additive error of $\epsilon = n^{-c}$. It is not known how to achieve a smaller error of $\epsilon = n^{-\omega(1)}$. Moreover, Applebaum et al. [AASY15] showed that "black-box techniques" cannot be used to achieve $\epsilon = n^{-\omega(1)}$ in Theorem 1.3, even if one replaces Σ_5-circuits with Σ_i-circuits for any number i. We note that all existing results (including the one in this paper) use "black-box techniques".

This led Applebaum et al. [AASY15] to consider multiplicative extractors.[8] Applebaum et al. [AASY15] showed that the construction and proof of Trevisan and Vadhan [TV00] can be extended to yield multiplicative extractors in Theorem 1.3. Recently, Ball et al. [BGDM23] improved upon Theorem 1.3 in two respects:

- The assumption was significantly improved to assuming that E is hard for exponential size nondeterministic circuits. This is a significant improvement as this assumption is weaker and more standard.
- The class of distributions was extended to include "samplable distributions with postselection".[9] This is a richer class of distributions. More specifically, a distribution X over $\{0,1\}^n$ is samplable with postselection by size s circuits, if there is a size s "sampling circuit" $A : \{0,1\}^r \to \{0,1\}$ and a size

[3] A precise definition of nondeterministic circuits appears in Section 3.2.

[4] A Σ_i-circuit is a nonuniform analogue of the class $P^{\Sigma_i^{\mathsf{P}}}$ that contains Σ_i^{P}, and recall that $\mathsf{P} = \Sigma_0^{\mathsf{P}}$ and $\mathsf{NP} = \Sigma_1^{\mathsf{P}}$. See Section 3.2 for a formal definition.

[5] We remark that following [TV00] there is some later work that relies on hardness for Σ_i-circuits for $i > 1$ [GW02, AS14, AASY15, AIKS16, BDL22].

[6] The parameter Δ is known as the "entropy deficiency", and a result is stronger when Δ is larger. All the known results, as well as our new results, achieve $\Delta = \alpha n$, for some constant $\alpha > 0$. See discussion and open problem in Section 6.1.

[7] The result stated in [TV00] gives an extractor with shorter output length of $m = \Theta(\log n)$. Nevertheless, Applebaum et al. [AASY15] observe that the result of [TV00] extends to larger output length, as stated in Theorem 1.3.

[8] Applebaum et al. [AASY15] use a more stringent definition of multiplicative extractors than the one we use here. In our terminology, they consider extractors w.r.t to the (double-sided) relation:

$$p_1 \overset{md}{\sim}_\epsilon p_2 \iff p_1 \overset{m}{\sim}_\epsilon p_2 \text{ and } p_2 \overset{m}{\sim}_\epsilon p_1,$$

which they call "relative-error extractors". However, for the suggested applications of such extractors (for example, the motivating application of selecting keys for cryptographic protocols), extractors w.r.t. $\overset{m}{\sim}_\epsilon$ suffice, and one does not benefit from considering extractors w.r.t. $\overset{md}{\sim}_\epsilon$. For this reason, we focus on extractors w.r.t. $\overset{m}{\sim}_\epsilon$ in this paper. Jumping ahead, we remark that our technique is also applicable to construct extractors w.r.t. $\overset{md}{\sim}_\epsilon$, and we discuss the two notions in Section 6.2.

[9] Ball et al. [BGDM23] also consider distributions samplable by quantum circuits, which we will not discuss in this paper.

s "postselection circuit" $P : \{0,1\}^r \to \{0,1\}$, such that $X = (A(Y)|P(Y) = 1)$ for $Y \leftarrow U_r$. (See precise definition in Section 3.1). Loosely speaking, this allows A to first sample $A(Y)$, and then, "postselect" the obtained distribution, and condition it on the event $\{P(Y) = 1\}$. This class of distributions contains samplable distributions, as well as recognizable distributions (Defined in [Sha09] and studied in [KvMS09, AASY15, LZ19, SS24], see precise definition in Section 3.1).

Ball et al. [BGDM23] use the same construction as [TV00, AASY15], however their analysis is significantly more complicated, and introduces new conceptual ideas, as well as some considerable technical sophistication. The price of achieving a weaker hardness assumption is that the proof is less modular, and significantly more complicated than that of [TV00]. Additionally, Ball et al. [BGDM23] only achieve standard (additive) extractors.

1.3 Our Results

1.3.1 Multiplicative Extractors for Samplable Distributions

In this paper we prove a version of Theorem 1.3 that achieves multiplicative extractors w.r.t. $\overset{m}{\sim}_\epsilon$, together with the two improvements of Ball et al. [BGDM23]. This achieves the best of both worlds. The precise result (stated below) is identical to Theorem 1.3, except for the weaker hardness assumption, the addition of "postselection", and that $\overset{a}{\sim}_\epsilon$ is replaced by $\overset{m}{\sim}_\epsilon$.

Theorem 1.4 (Multiplicative extractors for samplable distributions)**.** *If* E *is hard for exponential size nondeterministic circuits then there exists a constant* $\alpha > 0$*, such that for every constant* $c > 1$*, and for every sufficiently large* n*, there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{\alpha n}$ *that is a* $((1-\alpha)\cdot n, \overset{m}{\sim}_\epsilon)$*-extractor for distributions samplable with postselection by circuits of size* n^c*, where* $\epsilon = n^{-c}$*. Furthermore,* Ext *is computable in time* $\mathsf{poly}(n^c)$*.*

We use (essentially) the same construction as the previous papers [TV00, AASY15, BGDM23] but suggest a different analysis that is more modular, and significantly simpler than the approach of Ball et al. [BGDM23]. In fact, in our opinion, this approach is simpler and more natural than that of the original work of Trevisan and Vadhan [TV00]. Loosely speaking, our proof borrows some of the new ideas of Ball et al. [BGDM23], but avoids the technical complications by considering an intermediate object that is a "multiplicative PRG". (We elaborate on our approach and compare it to that of [TV00, BGDM23] in Section 2).

Theorem 1.4 achieves $m = \Omega(n)$ bits. As in previous work [TV00, BGDM23], we can also obtain extractors that extract almost all the randomness (rather than a constant fraction) under the same assumption. In this result we obtain multiplicative extractors w.r.t $\overset{m}{\sim}_{(\epsilon,\delta)}$, for an exponentially small additive error term δ.

Theorem 1.5 (Multiplicative extractors with larger output length)**.** *If* E *is hard for exponential size nondeterministic circuits then for every sufficiently small constant* $\gamma > 0$*, every constant* $c > 1$*, and for every sufficiently large* n*, there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{(1-O(\gamma))\cdot n}$ *that is a* $((1-\gamma)\cdot n, \overset{m}{\sim}_{(\epsilon,\delta)})$*-extractor for distributions samplable with postselection by circuits of size* n^c*, where* $\epsilon = n^{-c}$ *and* $\delta = 2^{-\Omega(\gamma\cdot n)}$*. Furthermore,* Ext *is computable in time* $\mathsf{poly}(n^c)$*.*

Both [TV00] and [BGDM23] got extractors with the same output length. However, their extractor is additive (w.r.t. $\overset{a}{\sim}_\epsilon$ for $\epsilon = n^{-c}$) and are not suitable for the application of selecting keys for cryptographic protocols. In contrast, our extractors are mutliplicative w.r.t $\overset{m}{\sim}_{(\epsilon,\delta)}$ for the same ϵ, and an exponentially small $\delta = 2^{-\Omega(n)}$ which (as explained before) is suitable for the intended application.

As in the previous works [TV00, BGDM23], enlarging the output length is achieved by composing the basic extractor with a seeded-extractor (which can be set up to have exponentially small additive error δ). We observe that when the basic extractor is multiplicative (as in the case of Theorem 1.4) one obtains a multiplicative extractor (with two parameters ϵ and δ) as in Theorem 1.5 (see Section 4.3).

1.3.2 Consequences and Necessary Assumptions for Extractors for Samplable Distributions

Hardness assumptions for extractors for samplable distributions. As mentioned previously, Trevisan and Vadhan [TV00] observed that extractors for samplable distributions imply circuit lower bounds. Specifically that if $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ is an $(n-1, \overset{a}{\sim}_{\frac{1}{5}})$-extractor for distributions samplable by size s circuits, then Ext cannot be computed by circuits of size slightly smaller than s. This lower bound seems significantly weaker than the hardness assumption used in Theorem 1.4. A natural question is whether hardness against nondeterministic circuits (as in Theorem 1.4) is necessary for obtaining extractors for samplable distributions?

Following [BGDM23], we observe that our proof of Theorem 1.4 (as well as the proofs of the previous results [TV00, BGDM23]) yield extractors for a richer class of distributions: The class of distributions that are samplable by size $s = n^c$ deterministic circuits, with postselection by size s *nondeterministic* circuits.[10] This is a richer class than both distributions samplable by size n^c circuits (as in Theorem 1.3) and distributions samplable with postselection by size n^c circuits (as in [BGDM23] and Theorem 1.4). See precise definition of this class in Section 3.1, and a formal statement and discussion in Section 4.

We show that an extractor for this richer class of distributions *does imply* circuit lower bounds against nondeterministic circuits. More specifically, that Ext cannot be computed by nondeterministic circuits of size slightly smaller than s.

In fact, we get the stronger conclusion that computing the extractor is *hard on average* for nondeterministic circuits. Specifically, we show that an $(n - \log(1/\epsilon), \overset{a}{\sim}_\epsilon)$-extractor for this richer class is a function that is hard on average for nondeterministic circuits, meaning that every nondeterministic circuit of size slightly smaller than s, computes the extractor correctly on at most a $(\frac{1}{2} + O(\epsilon))$-fraction of the inputs. (In the case of extractors for the original class, this result gives average-case lower bounds against deterministic circuits).

Summing up, our results imply that hardness assumptions against nondeterministic circuits *cannot be avoided* as long as one uses proof techniques that immediately give extractors against this richer class of distributions (as is the case for all previous work). We remark that the lower bounds that we get are quantitatively weaker than the hardness assumption used in Theorem 1.4. See discussion in Section 5.

Extractors for samplable distributions and seed-extending PRGs. We show that extractors for the richer class imply a stronger object than a hard on average function. Specifically, an extractor w.r.t $\overset{m}{\sim}_\epsilon$ (rather than $\overset{a}{\sim}_\epsilon$) for the richer class (as is the case in Theorem 1.4) *is* a seed-extending $\overset{a}{\sim}_{O(\epsilon)}$-PRG for nondetrministic circuits of size slightly smaller than s. This result holds for every output length m, and is achieved by adapting an argument of Kinne, van Melkebeek and Shaltiel [KvMS09].

Jumping ahead, we remark that the key idea in our construction of multiplicative extractors of Theorem 1.4, is to construct (multiplicative) seed-extending PRGs for nondeterministic circuits, and show that such PRGs *are* multiplicative extractors. See Section 2 for a detailed explanation.

Together, these results give a formal connection between seed-extending PRGs for nondeterministic circuits and extractors (at least for some ranges of parameters) and we believe that it may be beneficial to explore further connections between these objects. See section 5 for a detailed discussion.

[10] Loosely speaking, the property of samplable distributions that is used in [TV00, BGDM23] (and this paper) is that for a smaplable distribution X sampled by a poly-size circuit A, a nondeterministic poly-size circuit can check whether a given x is in the support of X (or more generally that a poly-size Σ_1-circuit can compute a multiplicative approximation to $\Pr[X = x]$). These properties also hold for postselecting samplable distributions even if one allows nondeterministic postselection.

2 Technique

2.1 A Brief Overview of the Approach Used in the Previous Work

Extractors from functions that are very hard functions on average. Trevisan and Vadhan [TV00] started from a simple observation that if a function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ is sufficiently hard on average (against Σ_1-circuits) then Ext is an extractor (that outputs a single bit) for distributions samplable by poly-size (deterministic) circuits. More specifically, using our terminology they showed that:

Lemma 2.1 (Extractors from very hard on average functions [TV00])**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a function such that for every Σ_1-circuit C of size $s \geq n$, it holds that $\Pr_{X \leftarrow U_n}[C(X) = f(X)] \leq \frac{1}{2} + \frac{\epsilon}{2^\Delta}$. Then, f is an $(n - \Delta, \overset{a}{\sim}_{4\epsilon})$-extractor for distributions samplable by circuits of size $s' = (s\epsilon)^{\Omega(1)}$.*

This means that constructing extractors for samplable distributions from worst-case assumptions can be potentially achieved by "hardness amplification" which is the task of converting worst-case hard functions (as in hardness assumptions) into functions that are sufficiently hard on average. This seems promising as hardness amplification is a successful paradigm with many classical results [IW97, STV01].

Unfortunately, even if we settle for constant ϵ, unless the entropy deficiency is very small, and $\Delta = O(\log n)$, there are no known hardness amplification results with suitable parameters. (Note that in Theorems 1.3 and 1.4, a much larger entropy deficiency of $\Delta = \Omega(n)$ is obtained).

In fact, later work [AASY15] shows that it is impossible to use "black-box techniques" to start from the assumption that E is hard for Σ_i-circuits, and obtain a function that is this hard on average (and this holds for every i). This means that results like Theorem 1.3 cannot be obtained by hardness amplification.

Bypassing the barrier of obtaining functions that are very hard on average. Because of this barrier, Trevisan and Vadhan (and following work) could not use Lemma 2.1 directly. Instead, Trevisan and Vadhan used a construction by Sudan, Trevisan and Vadhan [STV01] (that is based on error-correcting codes, and was used to obtain hardness amplification) to design their function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$.

As they could not show that Ext is sufficiently hard on average, they instead directly showed that Ext is an extractor for samplable distributions. This leads to technical complications (that do not arise when analyzing Ext in the realm of hardness amplification). Specifically, in the case of hardness amplification one is interested in the behavior of Ext on a *uniform* $X \leftarrow U_n$. In contrast, when analyzing Ext as an extractor, one needs to analyze Ext on an arbitrary samplable distribution X with $H_\infty(X) \geq n - \Delta$.

On a technical level, the function Ext designed by Trevisan and Vadhan (which we will soon review in detail) relies on an error-correcting code with block length 2^n. Analyzing it on a distribution X that is substantially different than U_n runs into difficulties, as error-correcting codes give the "same importance" to every one of the symbols of the 2^n bit long codeword, whereas the distribution X does not.

The recent and exciting work of Ball et al. [BGDM23] uses the same function Ext, and uses considerable technical sophistication to analyze the behavior of Ext on distributions X that have high min-entropy but are not uniform. This indeed allows Ball et al. to make the reduction use "less levels of nondeterminism" and start from a weaker hardness assumption, but leads to a complicated and technical proof (as modularity is sacrificed in order make the reduction use less levels of nondeterminism). Moreover, [BGDM23] do not get a multiplicative extractor.

2.2 Multiplicative Extractors from Seed-Extending Multiplicative PRGs

In Lemma 2.2 of this paper, we introduce a new approach to construct extractors from samplable distributions. More specifically, we prove an analogous result to Lemma 2.1 with the difference that rather than starting from

a function that is hard on average for nondeterministic circuits, we start from a seed-extending multiplicative PRG for nondeterministic circuits (as in Definition 1.2). This has several advantages:

- This approach is applicable to any output length m, and not just to $m = 1$, as is the case of Lemma 2.1.
- The approach gives multiplicative extractors w.r.t. $\overset{m}{\sim}_\epsilon$ rather than additive extractors w.r.t. $\overset{a}{\sim}_\epsilon$.[11]
- Most importantly, as we shoe in Theorem 2.3 below, the starting point of the new approach in Lemma 2.2, *can* be achieved under the weak assumption that E is hard for exponential size nondeterministic circuits. In contrast (as we explained previously) there are barriers to obtaining the starting point of Lemma 2.1 even under significantly stronger assumptions against Σ_i-circuits [AASY15], and these barriers apply even for (additive) extractors that output a single bit.

The new approach is stated in the lemma below.

Lemma 2.2 (Multiplicative extractors from Multiplicative PRGs)**.** *If $G : \{0,1\}^n \to \{0,1\}^m$ is a seed-extending PRG for nondeterministic circuits of size $s \geq n \geq m$ w.r.t. $\overset{m}{\sim}_{(\epsilon, \frac{\epsilon}{2^{\Delta+m}})}$. Then, G is an $(n-\Delta, \overset{m}{\sim}_{12\epsilon})$-extractor for distributions samplable by circuits of size $s' = (s\epsilon)^{\Omega(1)}$.*

We will soon show in Section 2.3 that essentially the same function Ext used by Trevisan and Vadhan, can be shown to be a seed-extending multiplicative PRG, with parameters that yield Theorem 1.4 using Lemma 2.2. This approach will lead to a simple and modular proof that produces a multiplicative extractor.

A more general version of Lemma 2.2 (which applies to the richer class of samplable distributions with postselection) is stated and proven in Section 4.1. Below is a proof sketch.

Proof sketch for Lemma 2.2. Let us now explain the idea behind the proof of Lemma 2.2. For this purpose, we will consider a simpler case in which we are only interested in extracting from samplable distributions W over $\{0,1\}^n$ which are flat. That is, that W is uniform over a set $T \subseteq \{0,1\}^n$ of size $2^{n-\Delta}$. The advantage of assuming that W is flat, is that this immediately implies that there is a small nondeterministic circuit B which given $x \in \{0,1\}^n$, answers one iff $x \in T$. This follows because if $A : \{0,1\}^r \to \{0,1\}^n$ is the size s' sampling circuit such that $W = A(U_r)$, then when given x, B can verify that $x \in T$ by "guessing" $v \in \{0,1\}^r$ such that $A(v) = x$, and v serves as a witness that $x \in T$.

Assume that $G : \{0,1\}^n \to \{0,1\}^m$ is not an $(n-\Delta, \overset{m}{\sim}_{12\epsilon})$-extractor for W. This means that there exists $z \in \{0,1\}^m$ such that $\Pr[G(W) = z] > e^{12\epsilon} \cdot 2^{-m}$. We now design a nondeterministic circuit $D : \{0,1\}^n \times \{0,1\}^m \to \{0,1\}$ of size s that shows that $G'(x) = (x, G(x))$ is not a $\overset{m}{\sim}_{(\epsilon, \frac{\epsilon}{2^{\Delta+m}})}$-PRG.

On input $(x,y) \in \{0,1\}^n \times \{0,1\}^m$, $D(x,y)$ will answer one if $x \in T$ and $y = z$. (Note that D can do this by using the circuit B). By construction D is a nondeterministic circuit of size s for s slightly larger than s'. Let us consider the random variables $X \leftarrow U_n$, and $Y \leftarrow U_m$, we compute:

$$\begin{aligned} p_1 &= \Pr[D(X,Y) = 1] = \Pr[X \in T \wedge Y = z] = \Pr[X \in T] \cdot \Pr[Y = z] = 2^{-\Delta} \cdot 2^{-m}. \\ p_2 &= \Pr[D(G(X)) = 1] = \Pr[D(X, G(X)) = 1] = \Pr[X \in T \wedge G(X) = z] \\ &= \Pr[X \in T] \cdot \Pr[G(X) = z | X \in T] = 2^{-\Delta} \cdot \Pr[G(W) = z] > 2^{-\Delta} \cdot e^{12\epsilon} \cdot 2^{-m}, \end{aligned}$$

In particular, we have that

$$p_2 > e^{12\epsilon} \cdot p_1 \geq (1 + 12\epsilon) \cdot p_1 = (1 + 11\epsilon) \cdot p_1 + \epsilon \cdot p_1 \geq e^{\epsilon} \cdot p_1 + \epsilon \cdot 2^{-(\Delta+m)},$$

[11] Note that for small output length m, say $m = 1$, the multiplicative and additive notions of "close to uniform" essentially coincide. The difference between the additive and multiplicative notions increases with m. The fact that the new approach works directly for large m is one of the reasons that allow it to get multiplicative extractors.

and we indeed conclude that $p_1 \overset{m}{\not\sim}_{(\epsilon, \frac{\epsilon}{2\Delta+m})} p_2$, and get a contradiction.

The case where W is not flat is handled in the formal proof in Section 4.1 by replacing the check whether x is in the support of W, by a quantitative check that approximates $\Pr[W = x]$, and using the fact that nondeterministic circuits can approximate this quantity (in the sense that using Goldwasser-Sipser protocol [GS86], nondeterministic circuits can verify that this quantity is approximately larger than a given threshold, see Section 3.8 for details).

2.3 A Construction of Seed-Extending Multiplicative PRGs

In light of Lemma 2.2, in order to obtain the extractor stated in Theorem 1.4, it is sufficient to start from the assumption that E is hard for exponential size nondeterministic circuits, and construct multiplicative seed-extending PRGs for nondeterministic circuits.

Our PRG construction (which is specified formally in Figure 1) is essentially the same as that of the extractor of Trevisan and Vadhan [TV00], which builds on an adaptation of [STV01]. More specifically, we start from a hard function $f : \{0,1\}^\ell \to \{0,1\}$ (which is the characteristic function of the language in the hardness assumption). When shooting to get the extractor of Theorem 1.4, we set $\ell = O(\log n)$, so that the assumption gives that f cannot be computed by nondeterministic circuits of size slightly larger than n^c. With this choice, f is computable in time $2^{O(\ell)} = \mathsf{poly}(n^c)$. As is common in this area, the first step is to extend $f : \{0,1\}^\ell \to \{0,1\}$ into a low degree polynomial $\hat{f} : \mathbb{F}_q^d \to \mathbb{F}_q$ using the "low degree extension" a.k.a. the Reed-Muller code. Specifically, for an appropriately chosen constant d, we set $h = 2^{\ell/d}$, and set $\hat{f} : \mathbb{F}_q^d \to \mathbb{F}_q$ to be a polynomial of individual degree h (and total degree hd) that coincides with f on a "subcube" of size $h^d = 2^\ell$ of the q^d inputs (see Figure 1 for a more formal description). In coding theoretic terms, this means that the truth table of $\hat{f}$ is a Reed-Muller encoding of the truth table of f. As in [TV00], we choose a non-standard and huge alphabet size q, which will be exponential in n when proving Theorem 1.4.

The next step is to use "code concatenation" to obtain an m bit output. This concatenation is done using the seeded-extractor $\mathsf{SExt} : \{0,1\}^{\log q} \times \{0,1\}^{\log q} \to \{0,1\}^m$ of the leftover hash lemma [ILL89] (see precise statement in Theorem 3.10).[12] The final PRG $G(x)$ is obtained by thinking of $x \in \{0,1\}^n$ as a pair $(w, y) \in \mathbb{F}_q^d \times \{0,1\}^{\log q}$ and defining $\mathsf{Ext}(x) = \mathsf{SExt}(\hat{f}(w), y)$. A precise formal description appears in Figure 1. We will prove the following theorem.

Theorem 2.3. *[Multiplicative PRG] If E is hard for exponential size nondeterministic circuits, then there exists a constant $a \geq 1$ such that for every sufficiently large s, $m \leq s$, and $\frac{1}{2^s} \leq \rho \leq \frac{1}{s}$. The function $G : \{0,1\}^{a \cdot (m + \log(1/\rho))} \to \{0,1\}^m$ defined in Figure 1 is a seed-extending $\overset{m}{\sim}_{(\frac{1}{s}, \rho)}$-PRG for nondeterministic circuits of size s. Furthermore, G can be computed in time $\mathsf{poly}(s)$.*

Theorem 1.4 follows directly from Lemma 2.2 and Theorem 2.3, the details appear in Section 4. Note that in Theorem 2.3 the output length m is smaller than the input length. While this is suitable for our application, this raises the question of whether PRGs with larger stretch are possible. See Section 6.3 for a discussion and related results by Artemenko et al. [AIKS16].

In the remainder of this section we prove Theorem 2.3. That is, we will show that given a size s nondeterministic circuit D that breaks the PRG, we can construct a small nondeterministic circuit A that computes f and contradicts the hardness assumption.

The main technical difficulty in this argument is that we want A to have size polynomial in s even though $\frac{1}{\rho}$ and q are not polynomial in s. This means that we cannot run list-decoding algorithms for the underlying code, and this holds also if we restrict $\hat{f}$ to a line in $\mathbb{F}_q^d$ (which corresponds to a Reed-Solomon code) as the

[12] Here there is a slight difference from previous work [TV00, AASY15, BGDM23] and we can use a seeded-extractor SExt, whereas previous work required SExt to be a 2-source extractor (which is a stronger requirement).

Figure 1: Construction of multiplicative PRG

Hardness assumption: We are assuming that E is hard for exponential size nondeterministic circuits. Namely, that there exist constants $0 < \beta < 1 < B$ and a function $f : \{0,1\}^* \to \{0,1\}$ such that:

Easiness: f is computable in time $2^{B\ell}$ on inputs of length ℓ.

Hardness: For every sufficiently large ℓ, nondeterministic circuits of size $2^{\beta\ell}$ fail to compute f on inputs of length ℓ.

Input parameters: We are given integers $m \leq s$ and $\rho > 0$, such that $\frac{1}{2^s} \leq \rho \leq \frac{1}{s}$, and are assuming that s is sufficiently large.

Goal: A seed-extending $\overset{m}{\sim}_{(\frac{1}{s},\rho)}$-PRG for size s nondeterministic circuits, with output length m and seed length $O(m + \log(1/\rho))$.

Construction:

Low degree extension: Let c_0, c_q be sufficiently large universal constants that will be chosen in the proof. We set $h = s$, $d = \frac{c_0}{\beta}$, $\ell = d \log h$ and $q = \frac{2^m}{\rho^{c_q}}$. Let $\mathbb{F}_q$ be the field with q elements (we will be assuming that q is a power of 2) and fix some set $H \subseteq \mathbb{F}_q$ of size h. Note that $|H^d| = 2^\ell$, and we can identify between $\{0,1\}^\ell$ and H^d. We define $\hat{f} : \mathbb{F}_q^d \to \mathbb{F}_q$ be the "low degree extension" of f (a precise statement is given in Lemma 3.12). This is a polynomial of degree $\hat{h} = hd$ such that for every $x \in \{0,1\}^\ell$, $\hat{f}(x) = f(x)$ (where in the l.h.s. we view x as element in $H^d \subseteq \mathbb{F}_q^d$). We have that $\hat{f}$ is computable in time $\mathsf{poly}(2^{B\ell}, \log q) = \mathsf{poly}(s)$.

Leftover hash lemma seeded extractor: Let $\epsilon = \frac{\rho}{100s}$, and $\mathsf{SExt} : \{0,1\}^{\log q} \times \{0,1\}^{\log q} \to \{0,1\}^m$ be the $(m + 2\log(1/\epsilon), \epsilon)$-strong seeded extractor of the "Leftover hash lemma" (formally specified in Theorem 3.10). Note that by choosing c_q to be sufficiently large, we have that the $\log q > m + 2\log(1/\epsilon)$.

Construction of seed-extending PRG: We define $G : \{0,1\}^{(d+1)\log q} \to \{0,1\}^m$ as follows: Given a seed $x \in \{0,1\}^{(d+1)\log q}$ we interpret it as a pair $(w, v) \in \mathbb{F}_q^d \times \{0,1\}^{\log q}$ and define:

$$G(x) = \mathsf{SExt}(\hat{f}(w), v).$$

Note that by our choices, the seed length is $(d+1)\log q = a \cdot (m + \log(1/\rho))$, for some constant a that depends only on β. Furthermore, G is computable in time $2^{B \cdot \ell + 1} + \mathsf{poly}(2^\ell, \log q) = \mathsf{poly}(s)$ (where the exponent of the polynomial in s is a universal constant times $\frac{B}{\beta}$).

line is of length $q \geq 1/\rho$. Instead, following Trevisan and Vadhan [TV00] (who in turn attribute the idea to Feige and Lund [FL97]) we will use nondeterminism to "speed up" such computation, so that it does run in time $\mathsf{poly}(s)$. Our approach also builds on some of the improvements of Ball et al. [BGDM23] to reduce the number of "levels of nondeterminism" used in the argument. The full proof appears in Section 2.3.1 below.

2.3.1 Proof of Theorem 2.3

Assume that G is not a seed-extending $\overset{m}{\sim}_{(\frac{1}{s},\rho)}$-PRG for nondeterministic circuits of size s. Let $D : \mathbb{F}_q^d \times \{0,1\}^{\log q} \times \{0,1\}^m \to \{0,1\}$ be a size s nondeterministic circuit that breaks G. Throughout this proof we will consider a probability space with the following independently chosen random variables:

$$W \leftarrow \mathbb{F}_q^d, V \leftarrow \{0,1\}^{\log q}, R \leftarrow U_m, T \leftarrow \mathbb{F}_q \setminus \{0\}.$$

We define $p_1 = \Pr[D(W, V, R) = 1]$ and $p_2 = \Pr[D(W, V, \mathsf{SExt}(\hat{f}(W), V)) = 1]$. By the assumption that D breaks G we have that $p_1 \overset{m}{\not\sim}_{(\frac{1}{s},\rho)} p_2$, meaning that $p_2 > e^{\frac{1}{s}} \cdot p_1 + \rho$.

We will design a nondeterministic circuit A that computes f and contradicts the hardness assumption. This will be done by using (a variant of) the celebrated list-decoding algorithm of Sudan, Trevisan and Vadhan [STV01]. In this variant we will use "curves" instead of "lines" in a way that resembles the PRG of [SU05]. This will be simpler to analyze, and will produce a self contained proof.[13]

Definition 2.4 (Degree r curve passing through given $r+1$ points)**.** *For distinct $r+1$ elements $t_0,\ldots,t_r \in \mathbb{F}_q$ and (not necessarily distinct) $y_0,\ldots,y_r \in \mathbb{F}_q^d$ we define $C_{\substack{t_0,\ldots,t_r\\ y_0,\ldots,y_r}} : \mathbb{F}_q \to \mathbb{F}_q^d$ to be the unique degree r polynomial such that for every $0 \leq j \leq r$, $C_{\substack{t_0,\ldots,t_r\\ y_0,\ldots,y_r}}(t_j) = y_j$.*

The main technical lemma in the proof of Theorem 2.3 is the following lemma. It shows that for every $x \in \mathbb{F}_q^d$, there is a low-degree univariate polynomial $\hat{p}_x$ such that $\hat{p}_x(0) = \hat{f}(x)$, and furthermore, there is a specific test (specified in the lemma) that $\hat{p}_x$ passes, but no other low degree polynomial does.

More specifically, the lemma shows that for some sufficiently large constant r, there exist $t_1,\ldots,t_r \in \mathbb{F}_q \setminus \{0\}$ and $y_1,\ldots,y_r \in \mathbb{F}_q^d$, such that for every $x \in \mathbb{F}_q^d$, if we define $C_x = C_{\substack{0,t_1,\ldots,t_r\\ x,y_1,\ldots,y_r}}$ to be the degree r curve passing through the points $(0,x),(t_1,y_1),\ldots,(t_r,y_r)$, then the polynomial $\hat{p}_x = \hat{f} \circ C_x$ (which indeed satisfies $\hat{p}_x(0) = \hat{f}(x)$) is the only low-degree polynomial p such that $\Pr[D(C_x(T),V,\mathsf{SExt}(p(T),V)) = 1]$ is large. This is useful (as we will explain in detail below) as we aim to construct a nondeterministic circuit and this circuit will guess a polynomial p, verify that it passes the test, and then we have that $\hat{f}(x) = \hat{p}_x(0)$.

Lemma 2.5. *Let $r = c_r \cdot d = \frac{c_r \cdot c_0}{\beta}$ for a sufficiently large universal constant c_r, let $\gamma_1 = p_1 + 4\epsilon$, and $\gamma_2 = p_2 - \epsilon$. There exist distinct $t_1,\ldots,t_r \in \mathbb{F}_q \setminus \{0\}$ and $y_1,\ldots,y_r \in \mathbb{F}_q^d$ such that for every $x \in \mathbb{F}_q^d$, setting $C_x = C_{\substack{0,t_1,\ldots,t_r\\ x,y_1,\ldots,y_r}}$, we have that $\gamma_2 > e^{\frac{1}{4s}} \cdot \gamma_1$, and furthermore:*

***The correct polynomial passes*:** *For the degree $\hat{h} \cdot r$ polynomial $\hat{p}_x : \mathbb{F}_q \to \mathbb{F}_q$ defined by $\hat{p}_x = \hat{f} \circ C_x$, we have that for every $j \in [r]$, $\hat{p}_x(t_j) = \hat{f}(y_j)$, $\hat{p}_x(0) = \hat{f}(x)$ and*

$$\Pr[D(C_x(T),V,\mathsf{SExt}(\hat{p}_x(T),V)) = 1] \geq \gamma_2.$$

***No incorrect polynomial passes*:** *For every degree $\hat{h} \cdot r$ polynomial $p : \mathbb{F}_q \to \mathbb{F}_q$ such that $p \neq \hat{p}_x$, that satisfies that for every $j \in [r]$, $p(t_j) = \hat{f}(y_j)$, we have that*

$$\Pr[D(C_x(T),V,\mathsf{SExt}(p(T),V)) = 1] \leq \gamma_1.$$

Showing that Theorem 2.3 follows from Lemma 2.5. We will use the conclusion of Lemma 2.5 to contradict the hardness assumption, and show that there is a nondeterministic circuit A of size $2^{\beta\ell}$ that computes f. The circuit A will be hardwired with γ_1,γ_2, $t_1,\ldots,t_r$, $y_1,\ldots,y_r$ and $\hat{f}(y_1),\ldots,\hat{f}(y_r)$. Given input $x \in \{0,1\}^\ell$ (which we can think of as $x \in \mathbb{F}_q^d \subseteq H^d$ so that it is an input to $\hat{f}$) the nondeterministic circuit A will guess a polynomial $p : \mathbb{F}_q \to \mathbb{F}_q$ of degree $\hat{h} \cdot r$, (by guessing its coefficients) and do the following:

- Verify that for every $j \in [r]$, $p(t_j) = \hat{f}(y_j)$.
- Construct the nondeterministic circuit $D_x(t,i) = D(C_x(t),i,\mathsf{SExt}(p(t),i))$, which is of size $\mathsf{poly}(s,\log q)$.
- Goldwasser and Sipser [GS86] showed that there is an AM-protocol that given a circuit C of size s and $\gamma_2 > e^{1/s} \cdot \gamma_1$ solves the promise problem of distinguishing whether $\Pr[C(U_n) = 1] \geq \gamma_2$ or

[13] The argument of [STV01] requires an additional step of "self correction", and in our setting, showing that this step can be performed by a nondeterministic circuit (rather than a Σ_1-circuit) requires repeating the self-correction argument. Moreover, the argument we present here is arguably simpler and more direct than that used in [STV01, TV00]. It should be noted however, that our argument gives slightly inferior parameters as a list-decoding algorithm, but this difference is immaterial when proving Theorem 2.3.

$\Pr[C(U_n) = 1] \leq \gamma_1$. It is standard that this protocol extends to the case where C is nondeterministic.[14] Using Adleman's argument that $\mathsf{AM} \subseteq \mathsf{NP/poly}$, our size $\mathsf{poly}(s)$ nondeterministic circuit A, can indeed verify that $\Pr[D_x(T, V) = 1] \geq \gamma_2$.

- If all verification steps pass, then A outputs $p(0)$.

Overall, using Lemma 2.5, this gives a nondeterministic circuit A that computes f.[15] This circuit is of size $\mathsf{poly}(s, \log q) = s^{c_0}$ for some universal constant c_0, and we have that $s^{c_0} = h^{d\beta} = 2^{\beta\ell}$ as required.

2.3.2 Proof of Lemma 2.5

A calculation gives $\gamma_2 > e^{\frac{1}{4s}} \cdot \gamma_1$. Specifically, let $\eta = \frac{1}{s}$, and recall that $p_2 > e^{\eta} \cdot p_1 + \rho > \max(\rho, e^{\eta} \cdot p_1)$, implying $\epsilon = \frac{\rho\eta}{100} \leq \frac{p_2\eta}{100}$. Using that $\forall x \in [0, 1]$, $1 + x \leq e^x \leq 1 + 3x$ and $1 - x \leq e^{-x} \leq 1 - x/3$, we get:

$$\frac{\gamma_2}{\gamma_1} = \frac{p_2 - \epsilon}{p_1 + 4\epsilon} > \frac{p_2 - \frac{p_2\eta}{100}}{p_2 \cdot e^{-\eta} + \frac{4p_2\eta}{100}} = \frac{p_2 \cdot (1 - \frac{\eta}{100})}{p_2 \cdot (e^{-\eta} + \frac{4\eta}{100})} \geq \frac{e^{-\frac{3\eta}{100}}}{1 - \frac{\eta}{3} + \frac{4\eta}{100}} \geq \frac{e^{-\frac{3\eta}{100}}}{e^{-(\frac{\eta}{3} - \frac{4\eta}{100})}} = e^{\frac{\eta}{3} - \frac{4\eta}{100} - \frac{3\eta}{100}} > e^{\frac{\eta}{4}}.$$

We will use the probabilistic method to show the existence of $t_1, \ldots, t_r$ and $y_1, \ldots y_r$. For this purpose we consider a probability space in which we choose $y_1, \ldots, y_r \leftarrow \mathbb{F}_q^d$ and distinct $t_1, \ldots, t_r \leftarrow \mathbb{F}_q \setminus \{0\}$. For every $x \in \mathbb{F}_q^d$ we define (the random variable) $C_x = C_{\substack{0, t_1, \ldots, t_r \\ x, y_1, \ldots, y_r}}$. It is standard that the random variables $(C_x(t))_{t \neq 0}$ are r-wise independent.[16] Lemma 2.5 follows from the following claim by a union bound over the q^d choices of $x \in \mathbb{F}_q^d$.

Claim 2.6. *For every $x \in \mathbb{F}_q^d$, except for probability $\frac{1}{5q^d}$ over the choice of $t_1, \ldots, t_r$, $y_1, \ldots, y_r$ we have that $\hat{p}_x = \hat{f} \circ C_x$ satisfies:*

- *For every $j \in [r]$, $\hat{p}_x(t_j) = \hat{f}(y_j)$, and $\Pr[D(C_x(T), V, \mathsf{SExt}(\hat{p}_x(T), V)) = 1] \geq \gamma_2$.*
- *For every degree $\hat{h} \cdot r$ polynomial $p \neq \hat{p}_x$ such that $\Pr[D(C_x(T), V, \mathsf{SExt}(p(T), V)) = 1] > \gamma_1$, there exists a $j \in [r]$ such that $p(t_j) \neq \hat{f}(y_j)$,*

Proof of Claim 2.6. By a standard application of an r-wise independent tail inequality [BR94] we get that for every $x \in \mathbb{F}_q^d$, the values $p_1 = \Pr[D(W, V, R) = 1]$ and $p_2 = \Pr[D(W, V, \mathsf{SExt}(\hat{f}(W), V)) = 1]$ (which are probabilities over the choice $W \leftarrow \mathbb{F}_q^d$) are approximated by values $p_{x,1}, p_{x,2}$ (which are defined below by replacing W with $C_x(T)$ for $T \leftarrow \mathbb{F}_q \setminus \{0\}$). This is stated formally in the next claim.

Claim 2.7 (Sampling preserves p_1 and p_2)**.** *For every $x \in \mathbb{F}_q^d$, except for probability $\frac{1}{10q^d}$ over the choice of $t_1, \ldots, t_r$, $y_1, \ldots, y_r$ we have that:*

- *$p_{x,1} = \Pr[D(C_x(T), V, R) = 1] \leq p_1 + \epsilon$, and*
- *$p_{x,2} = \Pr[D(C_x(T), V, \mathsf{SExt}(\hat{f}(C_x(T)), V)) = 1] \geq p_2 - \epsilon = \gamma_2 > \gamma_1$.*

[14] This follows as the AM protocol of Goldwasser and Sipser [GS86] works in the framework of "constant round" AM-protocols, which allows Merlin to speak many times, and such protocols are later collapsed to a 2-message public coin AM protocol. See precise statement in Section 3.8.

[15] In fact, the circuit that we construct is a "single-valued nondeterministic circuit". This means that on every input x there is an accepting nondeterministic guess that outputs $f(x)$, and there does not exist an accepting nondeterministic guess that outputs a value different than $f(x)$.

[16] Note that this holds even though one of the points on the curve (specifically $C_x(0)$) is fixed to x, and is not random. Indeed this is where we can see the advantage of using curves over lines, as they give us r-wise independence, even when some points are fixed. Another advantage is that by increasing r, we get more independence, which allows us to use stronger tail inequalities.

The proof of Claim 2.7 follows by a straightforward application of the r-wise independent tail inequality of [BR94] (stated in Theorem 3.15). The calculation appears in Section 4.5.

We continue with the proof of Claim 2.6. Fix some $x \in \mathbb{F}_q^d$. By Claim 2.7 with probability $1 - \frac{1}{10q^d}$ of the choices of $t_1, \ldots, t_r$ and $y_1, \ldots, y_r$ we have that $p_{x,1} \leq p_1 + \epsilon$ and $p_{x,2} \geq p_2 - \epsilon$. Fix some specific choice of $t_1, \ldots, t_r$ which satisfies this condition. We define List_x to be the set of all degree $\hat{h} \cdot r$ polynomials $p : \mathbb{F}_q \to \mathbb{F}_q$ such that $\Pr[D(C_x(T), V, \mathsf{SExt}(p(T), V)) = 1] > \gamma_1$. We have seen that $\hat{p}_x = \hat{f} \circ C_x \in \mathsf{List}_x$. For every polynomial $p \in \mathsf{List}_x$ we have that:

$$\Pr[D(C_x(T), V, \mathsf{SExt}(p(T), V)) = 1] - \Pr[D(C_x(T), V, R) = 1] > \gamma_1 - p_{x,1} > (p_1 + 4\epsilon) - (p_1 + \epsilon) = 3\epsilon.$$

As T is uniform over $\mathbb{F}_q \setminus \{0\}$ and independent of (V, R), by an averaging argument, it follows that there exist a subset $V_{x,p} \subseteq \mathbb{F}_q \setminus \{0\}$ of size $\epsilon(q-1)$ such that for every $t \in V_{x,p}$, we have that:

$$\Pr[D(C_x(t), V, \mathsf{SExt}(p(t), V)) = 1] - \Pr[D(C_x(t), V, R) = 1] > 2\epsilon.$$

For every $t \in \mathbb{F}_q \setminus \{0\}$ we define:

$$\mathsf{List}_{x,t} = \{a : \Pr[D(C_x(t), V, \mathsf{SExt}(a, V)) = 1] - \Pr[D(C_x(t), V, R) = 1] > \epsilon\},$$

so that for $t \in V_{x,p}$, we have that $p(t) \in \mathsf{List}_{x,t}$. As SExt is a (k, ϵ)-strong extractor for $k = m + 2\log(1/\epsilon)$, we have that for every $t \in \mathbb{F}_q \setminus \{0\}$, $|\mathsf{List}_{x,t}| \leq 2^k$ (as otherwise the uniform distribution on $\mathsf{List}_{x,t}$ violates the guarantee of strong extractors (see Definition 3.9) with respect to the distinguisher $D_t(i, z) = D(C_x(t), i, z)$).

We have the setup of the celebrated Reed-Solomon list-decoding algorithm of Sudan [Sud97] (stated formally in Theorem 3.8). More precisely, there are $\mathsf{prs} = (q-1) \cdot 2^k$ points (namely, all pairs (t, y) for $t \in \mathbb{F}_q \setminus \{0\}$ and $y \in \mathsf{List}_{x,t}$) such that every degree $\mathsf{deg} = \hat{h} \cdot r$ polynomial $p \in \mathsf{List}_x$, p passes through $\mathsf{agr} = \epsilon \cdot (q-1)$ of the points. By Sudan's theorem, if $\mathsf{agr} > \sqrt{2 \cdot \mathsf{prs} \cdot \mathsf{deg}}$ then $|\mathsf{List}_x| \leq \frac{2\mathsf{prs}}{\mathsf{agr}} = \frac{2 \cdot 2^k}{\epsilon} = \frac{2^{m+1}}{\epsilon^3}$.[17] The requirement on prs translates to $q - 1 > \frac{2 \cdot 2^m \cdot \hat{h} \cdot r}{\epsilon^3}$. Recall that $\hat{h} = h \cdot d \leq s^2$, $r \leq s$, $\epsilon = \frac{\rho}{100 \cdot s}$, and we have that $s \leq \frac{1}{\rho}$. We choose the constant c_q to be sufficiently large so that $q = \frac{2^m}{\rho^{c_q}}$ satisfies the requirement.

Using that $(t_1, \ldots, t_r)$ and C_x are independent to trim the list. We observe that for every $x \in \mathbb{F}_q^d$, in the probability space of choosing $t_1, \ldots, t_r$ and $y_1, \ldots, y_r$, the random variables $p_{x,1}, p_{x,2}$ and List_x depend only on the "shape" of the curve C_x. More formally, $p_{x,1}, p_{x,2}$ and List_x are determined by the set $\{(t, C_x(t)) : t \in \mathbb{F}_q\}$ which is determined by the polynomial C_x. However, for every specific fixing of the polynomial C_x, every choice of distinct values for $t_1, \ldots, t_r \in \mathbb{F}_q \setminus \{0\}$ is still possible, and equally likely. More formally, this says that the random variable C_x is independent of the random variable $(t_1, \ldots, t_r)$).

Consider conditioning the probability space of choosing $t_1, \ldots, t_r$ and $y_1, \ldots, y_r$, on a specific fixing of C_x, such that $p_{x,1} \leq p_1 + \epsilon$ and $p_{x,2} \geq p_2 - \epsilon$. By Claim 2.7 such a fixing occurs with probability $1 - \frac{1}{10q^d}$. We've seen that List_x is fixed, and yet $(t_1, \ldots, t_r)$ are distributed like t random distinct values in $\mathbb{F}_q \setminus \{0\}$. We also have that $\hat{p}_x \in \mathsf{List}_x$, and that for every $j \in [r]$, $\hat{p}_x(t_j) = \hat{f}(C_x(t_j)) = \hat{f}(y_j)$. Every $p \in \mathsf{List}_x$ that is different from $\hat{p}_x$ agrees with $\hat{p}_x$ in at most $\hat{h} \cdot r$ elements. Therefore, the probability (in this conditioned probability space) that p and $\hat{p}_x$ agree on the (still random) $t_1, \ldots, t_r$ is at most $\left(\frac{\hat{h} \cdot r}{q-1}\right)^r \leq \frac{1}{10q^d}$. The last inequality follows as $\hat{h} \cdot r = hdr = s \cdot c_r \cdot d^2 \leq s^3 \leq \frac{1}{\rho^3}$ and by taking the constant c_q to be sufficiently large, we have that for $q = \frac{2^m}{\rho^{c_q}}$, $\left(\frac{\hat{h} \cdot r}{q-1}\right)^r \leq \left(\frac{1}{\sqrt{q}}\right)^{c_r \cdot d} \leq \frac{1}{10q^d}$ by choosing the constant c_r to be sufficiently large.

Overall, we have that except for probability $\frac{1}{10q^d} + \frac{1}{10q^d} = \frac{1}{5q^d}$ over the choice of $t_1, \ldots, t_r$ and $y_1, \ldots, y_r$, $\Pr[D(C_x(T), V, \mathsf{SExt}(\hat{p}_x(T), V)) = 1] \geq \gamma_2$ and for every degree $\hat{h} \cdot r$ polynomial $p \neq \hat{p}_x$ such that $\Pr[D(C_x(T), V, \mathsf{SExt}(p(T), V)) = 1] > \gamma_1$, there exist $j \in [r]$ such that $p(t_j) \neq \hat{p}_x(t_j) = \hat{f}(y_j)$.

[17] Note that here (similar to [TV00, BGDM23] and in contrast to [STV01]) we only use combinatorial list-decoding, and do not rely on the efficiency of Sudan's algorithm, and we could have used a combinatorial list-decoding result like the Johnson bound.

Paper Organization

In Section 3 we review some of the definitions and components used in this paper. In Section 4 we restate Lemma 2.2 in a more general way, and use it (together with Theorem 2.3) to prove the two main theorems (Theorems 1.4 and 1.5). In Section 5 we give the formal statements of the results described in Section 1.3.2 on consequences of extractors for samplable distributions, and elaborate on our interpretation of these results. Finally, in Section 6 we discuss open problems.

3 Preliminaries

Probabilistic notation: For a distribution D, we use the notation $X \leftarrow D$ to denote the experiment in which X is chosen according to D. For a set A, we use $X \leftarrow A$ to denote the experiment in which X is chosen uniformly from the set A. We often also identify a distribution X, with the random variable X chosen from this distributions. For a random variable X and an event A we use $(X|A)$ to denote the distribution which chooses an element according to X, conditioned on A. We use U_n to be the uniform distribution on n elements.

Relations from the introduction. For completeness, we repeat the definition of the various relations defined in the Section 1.

Definition 3.1 (Definitions of relations from Section 1)**.** *Given numbers $p_1, p_2, \epsilon, \delta \in [0,1]$, we define the following relations:*

$$\begin{aligned}
p_1 \overset{ad}{\sim}_\epsilon p_2 &\iff |p_2 - p_1| \leq \epsilon. \\
p_1 \overset{a}{\sim}_\epsilon p_2 &\iff p_2 \leq p_1 + \epsilon. \\
p_1 \overset{m}{\sim}_\epsilon p_2 &\iff p_2 \leq e^\epsilon \cdot p_1. \\
p_1 \overset{m}{\sim}_{(\epsilon,\delta)} p_2 &\iff p_2 \leq e^\epsilon \cdot p_1 + \delta. \\
p_1 \overset{md}{\sim}_\epsilon p_2 &\iff p_1 \overset{m}{\sim}_\epsilon p_1 \text{ and } p_2 \overset{m}{\sim}_\epsilon p_1.
\end{aligned}$$

Note that while some of these relations (e.g. $\overset{m}{\sim}_\epsilon$) are interesting for $\epsilon > 1$, in this paper we will always have that $0 \leq \epsilon \leq 1$ so that $1 + \epsilon \leq e^\epsilon \leq 1 + 3\epsilon$, and $1 - \epsilon \leq e^{-\epsilon} \leq 1 - \frac{\epsilon}{3}$. We will use these inequalities throughout this paper.

3.1 Samplable Distributions and Postselection

Samplable distributions. We use the following standard definition of samplable distributions. In the definition below, we will typically be interested in the case where $\mathcal{C}$ is the class of functions computable by size n^c circuits for a constant parameter c.

Definition 3.2 (Samplable distributions)**.** *We say that a distribution X over $\{0,1\}^n$ is* **sampled** *by a function $A : \{0,1\}^r \to \{0,1\}^n$, if $X = A(U_r)$. Let $\mathcal{C}$ be a class of functions. A distribution X over $\{0,1\}^n$ is* **samplable** *by $\mathcal{A}$, if there exists a function $A : \{0,1\}^r \to \{0,1\}^n$ in the class $\mathcal{A}$ such that X is sampled by A.*

Samplable distributions with postselection. Ball et al. [BGDM23] consider a more general class which allows the sampling circuit to perform "postselection". More specifically, given a "sampling procedure" $A : \{0,1\}^r \to \{0,1\}^n$ and a "postselection" procedure $P : \{0,1\}^r \to \{0,1\}$, we will say that the distribution sampled by the pair (A, P) is the distribution X over $\{0,1\}^n$ obtained by taking $Y \leftarrow U_r$ and setting $X = (A(Y)|P(Y) = 1)$. A formal definition appears below.

Definition 3.3 (Samplable distributions with postselection)**.** *We say that a distribution X over $\{0,1\}^n$ is* **sampled** *by a function $A : \{0,1\}^r \to \{0,1\}^n$ with postselection by $P : \{0,1\}^n \to \{0,1\}$, if $X = (A(Y)|P(Y) = 1)$ for $Y \leftarrow U_r$. Let $\mathcal{A}$ and $\mathcal{P}$ be classes of functions. A distribution X over $\{0,1\}^n$ is* **samplable** *by $\mathcal{C}$ with* **postselection** *by $\mathcal{P}$ if there exists a function $A : \{0,1\}^r \to \{0,1\}^n$ in the class $\mathcal{A}$, and $P : \{0,1\}^r \to \{0,1\}$ in the class $\mathcal{P}$ such that X is sampled by A with postselection by P. In the case that $\mathcal{A}$ and $\mathcal{P}$ coincide, we will say that X is* **samplable with postselection** *by $\mathcal{A}$.*

Following Ball et al. [BGDM23] we are interested in distributions that are samplable with postselection by size $s = n^c$ circuits. Obviously, every distribution that is samplable by circuits of size s is also samplable with postselection by circuits of size s. However, sampling with postelection allows conditioning on events $\{Y : P(Y) = 1\} \subseteq \{0,1\}^r$ that occur with low probability, and seems to give a richer class of distribution.

Distributions samplable by deterministic circuits with postselection by nondeterministic circuits. The reason that we allow the class $\mathcal{A}$ (of sampling circuits) to be different than the class $\mathcal{P}$ (of postselecting circuits) is that we want to consider the yet richer class of distributions that are samplable by size s (deterministic) circuits with postselection by size s nondeterministic circuits. See discussion in Section 4 and Section 5.

Recognizable distributions. Distributions that are Samplable with postselection can also be seen as a generalization of the notion of "recognizable distribution" defined by Shaltiel [Sha09], see also [KvMS09, AASY15, LZ19, SS24], which in this terminology is the special case of distribution samplable with postselection, but restricted to the case that the sampling circuit A is the identity function (so that $A(U_n)$ samples the uniform distribution on n bits).

3.2 Definition of Circuits of Various Types

We formally define the circuit types that will be used in this paper.

Definition 3.4 (randomized circuits, nondeterministic circuits, oracle circuits and Σ_i-circuits)**.** *A* randomized *circuit C has additional wires that are instantiated with uniform and independent bits.*

A nondeterministic *circuit C has additional "nondeterministic input wires". We say that the circuit C evaluates to 1 on x iff there exist an assignment to the nondeterministic input wires that makes C output 1 on x.*

An oracle circuit $C^{(\cdot)}$ is a circuit which in addition to the standard gates uses an additional gate (which may have large fan in). When instantiated with a specific boolean function A, C^A is the circuit in which the additional gate is A. Given a boolean function $A(x)$, an A-circuit is a circuit that is allowed to use A gates (in addition to the standard gates). An $A_{||}$-circuit is a circuit that makes nonadaptive queries to its oracle A. (Namely, on every path from input to output, there is at most a single A gate).

An NP-circuit is a SAT-circuit (where SAT is the satisfiability function) a Σ_i-circuit is an A-circuit where A is the canonical Σ_i^P-complete language. The size of all circuits is the total number of wires and gates.[18]

[18] An alternative approach to define these circuit classes is using the Karp-Lipton notation for Turing machines with advice. For $s \geq n$, a size $s^{\Theta(1)}$ deterministic circuit is equivalent to $\mathsf{DTIME}(s^{\Theta(1)})/s^{\Theta(1)}$, a size $s^{\Theta(1)}$ nondeterministic circuit is equivalent to $\mathsf{NTIME}(s^{\Theta(1)})/s^{\Theta(1)}$, a size $s^{\Theta(1)}$ NP-circuit is equivalent to $\mathsf{DTIME}^{\mathsf{NP}}(s^{\Theta(1)})/s^{\Theta(1)}$, and a size $s^{\Theta(1)}$ Σ_i-circuit is equivalent to $\mathsf{DTIME}^{\Sigma_i^P}(s^{\Theta(1)})/s^{\Theta(1)}$.

3.3 Hardness Assumptions

We will rely on assumptions of the following form, introduced by Impagliazzo and Wigderson [IW97]

Definition 3.5 (E is hard for exponential size circuits)**.** *We say that "E is hard for exponential size circuits of type X" if there exist constants $0 < \beta < B$, and a language L in $\mathsf{E} = \mathsf{DTIME}(2^{B \cdot n})$, such that for every sufficiently large n, the characteristic function of L on inputs of length n is hard for circuits of size $2^{\beta n}$ of type X.*

We will also consider functions that are hard on average.

Definition 3.6 (average-case hard functions)**.** *We say that a function $f : \{0,1\}^m \to \{0,1\}^{m'}$ is ϵ-**hard** for a class $\mathcal{C}$, if for every $C \in \mathcal{C}$, such that $C : \{0,1\}^m \to \{0,1\}^{m'}$,*

$$\Pr_{X \leftarrow U_m}[C(X) = f(X)] < \epsilon.$$

We say that f is **hard** *for $\mathcal{C}$ if f is 1-hard for $\mathcal{C}$.*

3.4 Averaging Arguments for Multiplicative Relations

In the standard setup of pseudorandomness (that is w.r.t. $\overset{a}{\sim}_\epsilon$) we have that if a randomized circuit D distinguishes the output of some $G : \{0,1\}^r \to \{0,1\}^m$ from uniform, in the sense that $\Pr[D(U_m) = 1] \not\overset{a}{\sim}_\epsilon \Pr[D(G(U_r)) = 1]$, then there exists a fixing to the random coins of D, such that the non-randomized circuit obtained by employing this fixing also distinguishes.

The next lemma states that this property also holds for pseudorandomness w.r.t. $\overset{m}{\sim}_{(\epsilon,\delta)}$.

Lemma 3.7. *Let $G : \{0,1\}^r \to \{0,1\}^m$ be a function, and let D be a randomized circuit. If $\Pr[D(U_m) = 1] \not\overset{m}{\sim}_{(\epsilon,\delta)} \Pr[D(G(U_r)) = 1]$, then there exists a fixing to the random coins of D such that the obtained (non-randomized) circuit D' satisfies $\Pr[D(U_m) = 1] \not\overset{m}{\sim}_{(\epsilon,\delta)} \Pr[D(G(U_r)) = 1]$.*

Proof. Let $D(x, y)$ denote the output of D on input x and random coins y, and let:

$$p_1 = \Pr[D(U_m) = 1] = \mathbb{E}_y[\Pr[D(U_m, y) = 1]].$$
$$p_2 = \Pr[D(G(U_r)) = 1] = \mathbb{E}_y[\Pr[D(G(U_r), y) = 1]].$$

We have that $p_2 > e^\epsilon \cdot p_1 + \delta$. If there does not exists a fixing y' such that

$$\Pr[D(G(U_r)), y') = 1] > e^\epsilon \cdot \Pr[D(U_m, y') = 1] + \delta,$$

then

$$\begin{aligned} p_2 - p_1 &= \mathbb{E}_y[\Pr[D(G(U_r), y) = 1] - \Pr[D(U_m, y) = 1]] \\ &\leq \mathbb{E}_y[\Pr[D(U_m, y) = 1] \cdot (e^\epsilon - 1) + \delta] \\ &= (e^\epsilon - 1) \cdot \mathbb{E}_y[\Pr[D(U_m, y) = 1]] + \delta \\ &= (e^\epsilon - 1) \cdot p_1 + \delta, \end{aligned}$$

which implies the contradiction $p_2 \leq e^\epsilon \cdot p_1 + \delta$. □

3.5 Sudan's List-Decoding Algorithm

We will rely on Sudan's celebrated list-decoding algorithm for the Reed-Solomon code [Sud97].

Theorem 3.8 (Sudan's list-decoding algorithm [Sud97])**.** *Let* $\mathsf{prs}, \mathsf{agr}, \mathsf{deg}$ *be integers. Given* prs *distinct pairs* (x_i, y_i) *in field* F *with* $\mathsf{agr} > \sqrt{2 \cdot \mathsf{deg} \cdot \mathsf{prs}}$*, there are at most* $2\mathsf{prs}/\mathsf{agr}$ *polynomials* g *of degree* deg *such that* $g(x_i) = y_i$ *for at least* agr *pairs. Furthermore, a list of all such polynomials can be computed in time* $\mathsf{poly}(\mathsf{prs}, \log |F|)$.

We remark that in this paper (as in the previous work [TV00, BGDM23]) we will rely only existence of small lists, and do not use the efficiency of list-decoding algorithm.

3.6 Seeded Extractors

We use the following standard definition of seeded extractors. We remark that in many cases these are called "extractors" and this paper we use the term "seeded extractor" to differentiate them from seedless extractors.

Definition 3.9 (Seeded extractors)**.** *A function* $\mathsf{SExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a* (k, ϵ)*-seeded extractor if for every distribution* X *over* $\{0,1\}^n$*, with* $H_\infty(X) \geq k$*,* $\mathsf{SExt}(X)$ *is* ϵ*-close to* U_m*.*

SExt *is a strong* (k, ϵ)*-seeded extractor if the function* $\mathsf{SExt}' : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{d+m}$ *defined by* $\mathsf{SExt}'(x, y) = (y, \mathsf{SExt}(x, y))$ *is a* (k, ϵ)*-seeded extractor.*

We use the following result known as the "leftover hash lemma" by Impagliazzo, Levin and Luby [ILL89]

Theorem 3.10 (Leftover hash lemma [ILL89])**.** *For every integers* $m \leq n$*, and* $\epsilon > 0$*, there is a* $(m + 2\log(1/\epsilon), \epsilon)$*-strong extractor* $\mathsf{SExt} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$*. Furthermore,* SExt *can be computed in time* $\mathsf{poly}(n)$.

We remark that in some sources this lemma is stated with $d = 2n$ rather than $d = n$, but the statement also holds for $d = n$ (as stated above).

We also use the following result by Guruswami, Umans and Vadhan [GUV07].

Theorem 3.11 ([GUV07])**.** *For every constant* $\alpha > 0$*, and for every* $k \leq n$ *and* $\epsilon > 0$*, there is a* (k, ϵ)*-seeded extractor* $\mathsf{SExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *for* $d = O(\log n + \log(1/\epsilon))$ *and* $m = (1-\alpha)k$*. Furthermore,* SExt *can be computed in time* $\mathsf{poly}(n)$.

3.7 The Low Degree Extension

Many results in complexity theory and derandomization rely on the low-degree extension. Loosely speaking, this is a technique to extend a given function $f : \{0,1\}^\ell \to \{0,1\}$ to a low-degree d-variate polynomial $\hat{f} : \mathbb{F}_q^d \to \mathbb{F}_q$. The standard precise statement is given below.

Lemma 3.12. *Let* $f : \{0,1\}^\ell \to \{0,1\}$ *be a function and* $d \leq h \leq q$ *be integers such that* $h^d \geq 2^\ell$ *and* q *is a power of* 2*. Given* $H \subseteq \mathbb{F}_q$ *of size* h*, and a one-to-one map* $\phi : \{0,1\}^\ell \to H^d$*, there is a degree* $\hat{h} = h \cdot d$ *polynomial* $\hat{f} : \mathbb{F}_q^d \to \mathbb{F}$ *such that for every* $x \in \{0,1\}^\ell$*,* $f(x) = \hat{f}(\phi(x))$*. Furthermore,* $\hat{f}$ *can be computed in time* $\mathsf{poly}(2^\ell, \log q)$ *given oracle access to* f.

3.8 The Goldwasser-Sipser AM Protocol and Consequences

A classical result by Goldwasser and Sipser [GS86] shows that there is an AM protocol for showing that the fraction of accepting inputs of a given circuit is above some threshold. The same approach translates immediately to the case where the given circuit is nondeterministic (rather than deterministic). Below is a formal definition.

Definition 3.13 (The nondeterministic large set promise problem)**.** *Given $\lambda > 0$, we define a promise problem* $\mathsf{NondetLarge}_\lambda$ *over pairs (C, γ) where C is a nondeterministic circuit, and $0 \leq \gamma \leq 1$.*

- *The Yes instances are pairs (C, γ) such that C accepts at least a γ-fraction of its inputs.*
- *The No instances are pairs (C, γ) such that C accepts less than a $\gamma \cdot e^{-\lambda}$-fraction of its inputs.*

Note that a circuit C of size s can have at most s input bits. Throughout the paper we will always assume w.l.o.g. that C has s input bits (and may ignore some of them). We also note that because the number of possible inputs to C is at most 2^s, we can always assume that the number of bits needed to represent γ is at most s (which implies that the input to the promise problem is of length that is dominated by the length of the description of C, which is $O(s \log s)$).

Theorem 3.14 (Goldwasser and Sipser [GS86])**.** *For every integer s and $\lambda > 0$, there is a nondeterministic circuit A of size* $\mathsf{poly}(s, \frac{1}{\lambda})$ *which solves the promise problem* $\mathsf{NondetLarge}_\lambda$.

Theorem 3.14 is stated in a somewhat nonstandard way. The more standard formulation discusses deterministic circuits C, and gives an AM protocol that solves the promise problem. However, the same result immediately applies to nondeterministic circuits. This is because in the Goldwasser-Sipser AM protocol, Merlin sends inputs x to C on which $C(x) = 1$, and if C is nondeterministic, whenever Merlin sends an x, he can also supply a witness showing that $C(x) = 1$. This gives an AM-protocol with time $\mathsf{poly}(s, \frac{1}{\lambda})$ for $\mathsf{NondetLarge}_\lambda$, and the result in the theorem follows because one can transform an AM-protocol into a nondeterministic circuits, as in the proof that $\mathsf{AM} \subseteq \mathsf{NP/poly}$.

3.9 An r-wise Independent Tail Inequality

We need the following tail inequality by Bellare and Rompel [BR94].

Theorem 3.15 (r-wise independent tail inequality [BR94])**.** *Let $r > 4$ be an even integer. Suppose $X_1, X_2, \ldots, X_n$ are r-wise independent random variables taking values in $[0, 1]$. Let $X = \sum X_i$, $\mu = \mathbb{E}[X]$ and $A > 0$. Then:*

$$\Pr[|X - \mu| \geq A] \leq 8 \cdot \left(\frac{r\mu + r^2}{A^2} \right)^{r/2}.$$

In particular, if $r \leq n$, setting $A = \epsilon n$, for some $\epsilon > 0$, it follows that:

$$\Pr[|X - \mu| \geq \epsilon n] \leq 8 \cdot \left(\frac{2r}{\epsilon^2 n} \right)^{r/2}.$$

4 Constructions of Multiplicative Extractors for Samplable Distributions

In this section we prove our main theorems (Theorem 1.4 and Theorem 1.5). We start by restating the two theorems for the richer class of distributions samplable by deterministic circuits with postselection by nondeterministic circuits.

Remark 4.1 (Distributions samplable by deterministic circuits with postselection)**.** *As we explained in the introduction, Ball et al. [BGDM23] obtained extractors not only for distributions samplable by (deterministic) circuits, but also for the richer class of distributions samplable with postselection by deterministic circuits.*

In fact, their approach also immediately extends to distributions samplable by deterministic circuits with postselection by nondeterministic circuits. This is also the case for our results stated in Theorem 4.2 and 4.3. In all cases, handling this richer class of distributions requires no extra effort, and is immediate from the proof. See Remark 4.6.

The reason that we find this interesting is that (as explained in Section 1.3.2) extractors for this richer class of distributions imply lower bounds against nondeterministic circuits. This observation shows that hardness assumptions against nondeterministic circuits cannot be avoided as long as we use proof techniques that immediately give extractors against the richer class of distributions samplable by deterministic circuits with postselection by nondeterministic circuits. See Section 5 for a discussion.

Theorem 4.2 (Multiplicative extractors for samplable distributions)**.** *If* E *is hard for exponential size nondeterministic circuits then there exists a constant* $\alpha > 0$*, such that for every constant* $c > 1$*, and for every sufficiently large* n*, there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{\alpha n}$ *that is a* $((1-\alpha)\cdot n, \overset{m}{\sim}_\epsilon)$*-extractor for distributions samplable by deterministic circuits of size* n^c *with postselection by nondeterministic circuits of size* n^c*, where* $\epsilon = n^{-c}$*. Furthermore,* Ext *is computable in time* $\mathsf{poly}(n^c)$*.*

Theorem 4.3 (Multiplicative extractors with larger output length)**.** *If* E *is hard for exponential size nondeterministic circuits then for every sufficiently small constant* $\gamma > 0$*, every constant* $c > 1$*, and for every sufficiently large* n*, there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{(1-O(\gamma))\cdot n}$ *that is a* $((1-\gamma)\cdot n, \overset{m}{\sim}_{(\epsilon,\delta)})$*-extractor for distributions samplable by deterministic circuits of size* n^c *with postselection by nondeterministic circuits of size* n^c*, where* $\epsilon = n^{-c}$ *and* $\delta = 2^{-\Omega(\gamma\cdot n)}$*. Furthermore,* Ext *is computable in time* $\mathsf{poly}(n^c)$*.*

Remark 4.4 (About the constants in Theorems 4.2 and 4.3)**.** *The assumption that* E *is hard for exponential size nondeterministic circuits (defined in Assumption 3.5) asserts the existence of two constants* β, B*. Some of the constants in Theorems 4.2 and 4.3 depend on* β *and* B*. We now list the precise dependence.*

- *The polynomial specified by the* $\mathsf{poly}(n^c)$ *term in both theorems depends on both* β *and* B*, and in both cases the extractor runs in time* $n^{O(c\cdot\frac{B}{\beta})}$*.*
- *The constant* α *in Theorem 4.2 and the constant hidden in the* $O(\gamma)$ *term in Theorem 4.3 depend on the constant* β*, but not on* B*. In both cases the dependence is of the form* $O(\frac{1}{\beta})$*.*
- *The constant hidden in the* $\Omega(\gamma\cdot n)$ *in Theorem 4.3 is a universal constant.*

Outline for this section. In Section 4.1 we show that multiplicative PRGs yield extractors. More specifically, we restate Lemma 2.2 in a more general way, and prove it. In Section 4.2 we use Lemma 2.2 to derive Theorem 4.2 from the multiplicative PRG construction of Theorem 2.3. In Section 4.3 we observe that one can compose a seeded extractor with a multiplicative extractor for samplable distributions and obtain a multiplicative extractor, and in Section 4.4 we use this composition to derive Theorem 4.3 from Theorem 4.2 and known seeded extractors. Finally in Section 4.5 we provide a proof of Claim 2.7.

4.1 Multiplicative Seed-Extending PRGs imply Multiplicative Extractors

The following lemma generalizes Lemma 2.2 and shows that a multiplicative extractor (with the same parameters as stated in Lemma 2.2) yields an extractor not only for distributions samplable with postselection ny deterministic circuits, but also to the richer class of distribution samplable by deterministic circuits with postselection by nondeterministic circuits.

Lemma 4.5. *For every $\epsilon > 0$ and $\Delta > 0$, if $G : \{0,1\}^n \to \{0,1\}^m$ is a seed-extending $\overset{m}{\sim}_{(\epsilon,\rho)}$-PRG for nondeterministic circuits of size $s \geq n \geq m$, where $\rho = \frac{\epsilon}{2^{m+\Delta}}$, then G is an $(n-\Delta, \overset{m}{\sim}_{12\epsilon})$-extractor for distributions samplable by deterministic circuits of size $s' = (s\epsilon)^{\Omega(1)}$ with postselection by nondeterministic circuits of size s'.*

The proof below generalizes the argument sketched in Section 2. The technique is inspired by results by Kinne et al. [KvMS09] and Shaltiel and Silbak [SS24] (that show that PRGs with sufficiently low (additive) error for deterministic circuits imply extractors for distributions recognizable by deterministic circuits). Here, we adapt the argument to handle the case where the PRG is multiplicative and fools nondeterministic circuits, and want to obtain an extractor for distributions samplable by deterministic circuits. This presents more complications (partly because classes of languages accepted by nondeterministic circuits are not closed under complement). We also remark that Li and Zuckerman [LZ19] gave a quantitatively better version of the result of [KvMS09], and while we believe that it may be possible to adapt the (more complicated) proof of Li and Zuckerman to this setting, this will not result in improvements in the parameters of the final extractor.

Proof. (of Lemma 4.5) Let W be a distribution over $\{0,1\}^n$ with $H_\infty(W) \geq n-\Delta$ that is sampled by a size s' circuit $A : \{0,1\}^{s'} \to \{0,1\}^n$, with postselection by a size s' nondeterministic circuit $P : \{0,1\}^{s'} \to \{0,1\}$.

If G is not a extractor for W w.r.t $\overset{m}{\sim}_\eta$ for $\eta = 12\epsilon$, then there exists $z \in \{0,1\}^m$, such that for $p_1 = \Pr[U_m = z] = 2^{-m}$, and $p_2 = \Pr[G(W) = z]$, we have that $p_1 \overset{m}{\not\sim}_\eta p_2$, meaning that $p_2 > e^\eta \cdot p_1 = e^\eta \cdot 2^{-m}$.

We will show that there exists a size s nondeterministic circuit $D : \{0,1\}^{n+m} \to \{0,1\}$ that breaks the seed-extending PRG. We will start by designing a randomized nondeterministic circuit $D_\lambda : \{0,1\}^{n+m} \to \{0,1\}$, which will also rely on a parameter $0 \leq \lambda \leq 1$. Later on, we will choose the parameter λ appropriately, and use an averaging argument to convert the randomized nondeterministic circuit, into a non-randomized nondeterministic circuit. The description of D_λ appears in Figure 2.

Figure 2: The distinguisher $D_\lambda : \{0,1\}^{n+m} \to \{0,1\}$

We first describe a randomized nondeterministic circuit $B_\lambda : \{0,1\}^n \to \{0,1\}$ (that will be used as a component D_λ). The circuit $B_\lambda(x)$ works as follows:

- Given $x \in \{0,1\}^n$, we prepare a nondeterministic circuit C_x that is defined as follows: On input $v \in \{0,1\}^{s'}$, C_x will run $A(v)$, and output one iff $A(v) = x$ and $P(v) = 1$.
  Consequently, for every $x \in \{0,1\}^n$, $\Pr[C_x(U_{s'}) = 1] = \Pr[W = x]$.
- Pick a uniform $\alpha \leftarrow [0,1]$.
- Let $\tau = 2^{-(n-\Delta)}$, so that by the min-entropy requirement on W, $\Pr[C_x(U_{s'}) = 1] = \Pr[W = x] \leq \tau$. By Theorem 3.14 there is a circuit A_λ of size $\mathsf{poly}(s', \frac{1}{\lambda})$ which solves the promise problem $\mathsf{NondetLarge}_\lambda$ on input $(C_x, \alpha \cdot \tau)$.
- Output $A_\lambda(C_x, \alpha \cdot \tau)$.

The randomized nondeterministic circuit $D_\lambda(x, y)$ is hardwired with the string $z \in \{0,1\}^m$, and works as follows:

- Given $x \in \{0,1\}^n$, and $y \in \{0,1\}^m$, output one iff $B_\lambda(x) = 1$ and $y = z$.

By Theorem 3.14, for every $\lambda > 0$, D_λ is a nondeterministic circuit of size $\mathsf{poly}(s', \frac{1}{\lambda})$. We will first analyze the case where $\lambda = 0$ (namely, the case in which A solves $\mathsf{NondetLarge}_0$, meaning that A is "errorless", and such an A may have exponential size). Later, we will show that the analysis also applies with small losses for $\lambda = O(\epsilon)$.

Note that for $\lambda = 0$, the circuit $B_0(x)$ answers one iff $\Pr[W = x] \geq \alpha \cdot \tau$. In other words, $B_0(x)$ is a randomized nondeterministic circuit that outputs a random bit that is one with probability $\frac{\Pr[W=x]}{\tau}$. For

$X \leftarrow U_n$ we conclude that:

$$\Pr[B_0(X)=1] = \sum_{x\in\{0,1\}^n} \Pr[X=x \wedge B_0(x)=1] = \sum_{x\in\{0,1\}^n} 2^{-n}\cdot\frac{\Pr[W=x]}{\tau} = \frac{2^{-n}}{\tau} = 2^{-\Delta}.$$

which implies that:

$$\Pr[X=x|B_0(X)=1] = \frac{\Pr[X=x\wedge B_0(X)=1]}{\Pr[B_0(X)=1]} = \frac{\Pr[X=x\wedge B_0(x)=1]}{2^{-\Delta}} = \frac{2^{-n}\cdot\frac{\Pr[W=x]}{\tau}}{2^{-\Delta}} = \Pr[W=x],$$

which gives that the distribution $(X|B_0(X)=1)$ is distributed precisely like W.

We will use this to show that D_0 breaks the PRG G. More specifically, we consider $X \leftarrow U_n$ and $Y \leftarrow U_m$ and define: $p_1^{(\lambda)} = \Pr[D_\lambda(X,Y)=1]$ and $p_2^{(\lambda)} = \Pr[D_\lambda(X,G(X))=1]$. We will start by showing that $p_1^{(0)} \overset{m}{\not\sim}_\eta p_2^{(0)}$.

$$p_1^{(0)} = \Pr[D_0(X,Y)=1] = \Pr[B_0(X)=1\wedge Y=z] = \Pr[B_0(X)=1]\cdot\Pr[Y=z] = 2^{-\Delta}\cdot 2^{-m}.$$
$$p_2^{(0)} = \Pr[D_0(X,G(X))=1] = \Pr[B_0(X)=1\wedge G(X)=z] = \Pr[B_0(X)=1]\cdot\Pr[G(X)=z|B_0(X)=1].$$

Using the fact that $(X|B_0(X)=1)$ is distributed like W, we conclude that:

$$p_2^{(0)} = \Pr[B_0(X)=1]\cdot\Pr[G(X)=z|B_0(X)=1] = 2^{-\Delta}\cdot\Pr[G(W)=z] = 2^{-\Delta}\cdot p_2 > e^\eta\cdot 2^{-\Delta}\cdot 2^{-m}.$$

This means that $p_1^{(0)} \overset{m}{\not\sim}_\eta p_2^{(0)}$. However, for $\lambda=0$ we do not control the size of the circuit D_λ. Therefore, we will try to argue that approximately the same inequality holds for small $\lambda > 0$. Indeed, by the definition of D^λ, it immediately follows that for every $0\le\lambda\le 1$ and every $x\in\{0,1\}^n$ and $y\in\{0,1\}^m$,

$$\Pr D_\lambda(x,y)=1] \le \Pr[D_0(x,y)=1] \le e^\lambda\cdot\Pr D_\lambda(x,y)=1].$$

This implies that if we set $\mu = 2^{-(m+\Delta)}$, for every $0\le\lambda\le 1$

$$p_1^{(\lambda)} \le p_1^{(0)} = 2^{-(m+\Delta)} = \mu.$$
$$p_2^{(\lambda)} \ge p_2^{(0)}\cdot e^{-\lambda} > e^{\eta-\lambda}\cdot 2^{-(m+\Delta)} = e^{\eta-\lambda}\cdot\mu,$$

meaning that setting $\lambda=\eta/2$, we have that $p_1^{(\lambda)} \overset{m}{\not\sim}_{\frac{\eta}{2}} p_2^{(\lambda)}$, and for $\rho = \frac{\epsilon}{2^{m+\Delta}} \le \frac{\eta\mu}{4}$ we have that

$$p_2^{(\lambda)} > e^{\frac{\eta}{2}}\cdot\mu \ge (1+\frac{\eta}{2})\cdot\mu = (1+\frac{\eta}{4})\cdot\mu + \frac{\eta}{4}\cdot\mu \ge e^{\frac{\eta}{12}}\cdot\mu + \rho = e^\epsilon\cdot p_1^{(\lambda)} + \rho,$$

giving that $p_1^{(\lambda)} \overset{m}{\not\sim}_{(\epsilon,\rho)} p_2^{(\lambda)}$, as required. We have obtained that D_λ breaks the PRG. We now turn to analyzing the complexity of D_λ. For $\lambda=\eta/2=6\epsilon$, the size of the randomized nondeterministic circuit D_λ is $s = \mathsf{poly}(s', \frac{1}{\epsilon})$. By Lemma 3.7 the coins of this randomized nondeterministic circuit can be fixed to yield a (non-randomized) nondeterministic of the same size. □

Remark 4.6. *We remark that in the proof above, no additional effort was needed to argue that the extractor works not only for samplable distributions, but also for distribution samplable by deterministic circuits with postselection by nondeterministic circuits. The only place where this came up, is in the first item in the definition of the circuit D_λ. More specifically, in the first item, the nondeterministic circuit C_x applies the postselecting circuit P.*

In fact, the same approach can be applied in the previous proofs of [TV00, BGDM23] and lead to extractors for the richer class of distribution samplable by deterministic circuits with postselection by nondeterministic circuits.

4.2 Proof of Theorem 4.2

Theorem 4.2 now follows immediately from composing the PRG of Theorem 2.3 with Lemma 4.5. The precise calculation follows.

We are assuming that E is hard for exponential size nondeterministic circuits. Let $a \geq 1$ be the constant guaranteed by Theorem 2.3 under this assumption. We set $\alpha = \frac{1}{4a}$. We are then given a constant c, and are aiming to construct the extractor guaranteed in Theorem 4.2. By Lemma 4.5, for sufficiently large n, in order to obtain an $((1-\alpha)\cdot n, \overset{m}{\sim}_{n^{-c}})$-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{\alpha n}$ for distributions samplable by deterministic circuits of size n^c with postselection by nondeterministic circuits of size n^c, it is sufficient to obtain a seed extending $\overset{m}{\sim}_{(\frac{1}{s},\rho)}$-PRG $G : \{0,1\}^n \to \{0,1\}^{\alpha n}$ for nondeterministic circuits of size $s = n^{c'}$, for some constant $c' > c$, with $\rho = 2^{-3\alpha n} \leq \frac{1}{n^c \cdot 2^{m+\alpha n}}$.

When ρ is expressed as a function of s, $\rho = 2^{-3\alpha n} = 2^{-3\alpha \cdot s^{\frac{1}{c'}}}$ satisfies the requirements of Theorem 2.3. We apply Theorem 2.3 and conclude that for every sufficiently large s, taking $n = s^{\frac{1}{c'}}$ and $m = \alpha n = \alpha \cdot s^{\frac{1}{c'}} \leq s$ we obtain a seed-extending $\overset{m}{\sim}_{(\frac{1}{s},\rho)}$-PRG $G : \{0,1\}^{a\cdot(m+\log(1/\rho))} \to \{0,1\}^{\alpha n}$ for nondeterministic circuits of size $s = n^{c'}$. By Theorem 2.3, G is computable in time $\mathsf{poly}(s) = \mathsf{poly}(n)$.[19] All that remains is to check that the seed length

$$a \cdot (m + \log(1/\rho)) = a \cdot (\alpha n + 3\alpha n) = n,$$

as required.

4.3 Composing a Multiplicative Extractor with an (Additive) Seeded Extractor

In this section we prove the following composition lemma, which shows that under certain conditions, one can use the output of an extractor Ext for samplable distribution as a seed to a seeded extractor SExt, resulting in an extractor for samplable distributions E with larger output length than Ext.

This approach was used in both [TV00, BGDM23] to increase the output length of their (additive) extractors. The lemma below asserts that if Ext is multiplicative, then the resulting extractor E is multiplicative (with two parameters ϵ and δ) where the first is inherited from the Ext and the second is inherited from SExt.

Lemma 4.7 (Composition Lemma)**.** *Let $\Delta \leq n_2 \leq n_1 \leq s$ and let $\epsilon > 0$. Assume that we have:*

- *An $(n_2 - \Delta - \log(1/\epsilon), \overset{m}{\sim}_\eta)$-extractor $\mathsf{Ext} : \{0,1\}^{n_2} \to \{0,1\}^{m_2}$ for distributions samplable by size $3s$ deterministic circuits with postselection by size $3s$ nondeterministic circuits.*
- *A seeded $(n_1 - \Delta, \epsilon)$-extractor $\mathsf{SExt} : \{0,1\}^{n_1} \times \{0,1\}^{m_2} \to \{0,1\}^{m_1}$.*

The function $E : \{0,1\}^{n_1+n_2} \to \{0,1\}^{m_1}$ defined by $E(x_1,x_2) = \mathsf{SExt}(x_1, \mathsf{Ext}(x_2))$ is an $(n_1 + n_2 - \Delta, \overset{m}{\sim}_{(\eta, e^\eta\cdot\epsilon+\epsilon)})$-extractor for distributions samplable by size s deterministic circuits with postselection by size s nondeterministic circuits.

In Section 4.4 we use this lemma to derive Theorem 4.3. In the next section we prove Lemma 4.7.

4.3.1 Proof of the Composition Lemma

So far, we have only considered "multiplicative distance" from the uniform distribution. However, the same concept can be defined in a more general way (which will be useful in the proof below).

[19] A more careful inspection of the parameters reveals that G is computable in time $n^{ecB/\beta}$ for some universal constant $e > 1$. and where B and β are the constants from the hardness assumption, as in Figure 1.

Definition 4.8. *Let $\sim$ be a relation on $[0,1]$. Let X, Y be two distributions over $\{0,1\}^m$. We say that X is ϵ-close to Y w.r.t. $\sim$ (and write $X \sim Y$) if for every function $C : \{0,1\}^m \to \{0,1\}$, $\Pr[D(X) = 1] \sim \Pr[D(Y) = 1]$.*

In the two propositions below we observe that distance w.r.t $\overset{m}{\sim}$ satisfies some of the standard properties satisfied by $\overset{a}{\sim}$ (such as a data processing inequality, and a suitable version of the triangle inequality). We will use these properties in the proof.

Proposition 4.9 (A form of the triangle inequality)**.** *Let X_1, X_2, X_3 be distributions over the same domain.*

- *If $X_1 \overset{a}{\sim}_\epsilon X_2$ and $X_2 \overset{m}{\sim}_{(\eta,\delta)} X_3$ then $X_1 \overset{m}{\sim}_{(\eta, e^\eta \cdot \epsilon + \delta)} X_3$.*
- *If $X_1 \overset{m}{\sim}_{(\eta,\delta)} X_2$ and $X_2 \overset{a}{\sim}_\epsilon X_3$ then $X_1 \overset{m}{\sim}_{(\eta, \epsilon+\delta)} X_3$.*

Proposition 4.10 (Data processing inequality for $\overset{m}{\sim}$)**.** *If $X_1 \overset{m}{\sim}_{(\eta,\delta)} X_2$ then for every function f, $f(X_1) \overset{m}{\sim}_{(\eta,\delta)} f(X_2)$.*

Proof of Lemma 4.7. Let $X = (X_1, X_2)$ be a distribution over $\{0,1\}^{n_1+n_2}$ that is samplable by a size s deterministic circuit $A : \{0,1\}^s \to \{0,1\}^{n_1+n_2}$ with postselection by a size s nondeterministic circuit $P : \{0,1\}^s \to \{0,1\}$. We start with the following claim.

Claim 4.11.

- *For every $x_1 \in \mathrm{Supp}(X_1)$, the distribution $(X_2|X_1 = x_1)$ is samplable by a size $3s$ deterministic circuit with postselection by a size $3s$ nondeterministic circuit.*
- *There is a set $B \subseteq \{0,1\}^{n_1}$ such that $\Pr[X_1 \in B] \leq \epsilon$ and for every $x_1 \notin B$, $H_\infty(X_2|X_1 = x_1) \geq n_2 - \Delta - \log \frac{1}{\epsilon}$.*

Proof. We start with the first item. For every $x_1 \in \mathrm{Supp}(X_1)$, the distribution $(X_2|X_1 = x_1)$ is samplable with postselection as follows: Let A_1, A_2 be deterministic circuits such that on input v, $A_1(v)$ outputs the first n_1 bits of $A(v)$ and $A_2(v)$ outputs the last n_2 bits of $A(v)$. Let $P_{x_1}(v)$ be the nondeterministic circuit that on input v answers one if $P(v) = 1$ and $A_1(v) = x_1$. Note that P_{x_1} is a nondeterministic circuit of size $3s$. It is immediate that for every $x_1 \in \mathrm{Supp}(X_1)$, the distribution $(X_2|X_1 = x_1) = (A_2(Y)|P_{x_1}(Y) = 1)$ for $Y \leftarrow U_s$, and is therefore samplable by the deterministic circuit A_2, with postselection by the nondeterministic circuit P_{x_1}.

For the second item we define $B = \left\{x_1 \in \{0,1\}^{n_1} : \Pr[X_1 = x_1] < 2^{-(n_1 + \log(1/\epsilon))}\right\}$, and observe that:

$$\Pr[X_1 \in B] \leq \sum_{x_1 \in B} \Pr[X_1 = x_1] \leq 2^{n_1} \cdot 2^{-(n_1+\log(1/\epsilon))} = \epsilon,$$

and for every $x_1 \notin B$, and $x_2 \in \{0,1\}^{n_2}$.

$$\Pr[X_2 = x_2 | X_1 = x_1] \leq \frac{\Pr[X_2 = x_2 \wedge X_1 = x_1]}{\Pr[X_1 = x_1]} \leq \frac{2^{-(n_1+n_2-\Delta)}}{2^{-(n_1+\log(1/\epsilon))}} = 2^{-(n_2 - \Delta - \log \frac{1}{\epsilon})}.$$

□

By Claim 4.11 we have that for every $x_1 \notin B$, $(\mathsf{Ext}(X_2)|X_1 = x_1)$ is close to uniform w.r.t. $\overset{m}{\sim}_\eta$. We will use this to prove the following claim.

Claim 4.12. $(X_1, U_{m_2}) \overset{m}{\sim}_{(\eta,\epsilon)} (X_1, \mathsf{Ext}(X_2))$.

Proof. Let $D : \{0,1\}^{n_1+m_2} \to \{0,1\}$ be some function.

$$\begin{aligned}
\Pr[D(X_1, \mathsf{Ext}(X_2)) = 1] &\leq \Pr[X_1 \in B] + \Pr[D(X_1, \mathsf{Ext}(X_2)) = 1 \wedge X_1 \notin B] \\
&\leq \epsilon + \sum_{x_1 \notin B} \Pr[D(X_1, \mathsf{Ext}(X_2)) = 1 \wedge X_1 = x_1] \\
&= \epsilon + \sum_{x_1 \notin B} \Pr[X_1 = x_1] \cdot \Pr[D(x_1, \mathsf{Ext}(X_2)) = 1 | X_1 = x_1] \\
&\leq \epsilon + \sum_{x_1 \notin B} \Pr[X_1 = x_1] \cdot e^{\eta} \cdot \Pr[D(x_1, U_{m_2}) = 1] \\
&\leq \epsilon + e^{\eta} \cdot \sum_{x_1 \in \mathrm{Supp}(X_1)} \Pr[X_1 = x_1] \cdot \Pr[D(x_1, U_{m_2}) = 1] \\
&= \epsilon + e^{\eta} \cdot \Pr[D(X_1, U_{m_2}) = 1]
\end{aligned}$$

The fourth line follows because for every $x_1 \notin B$, we have obtained that $U_{m_2} \overset{m}{\sim}_{\eta} (\mathsf{Ext}(X_2)|X_1 = x_1)$ and can consider the function $D_{x_1} : \{0,1\}^{m_2} \to \{0,1\}$ defined by $D_{x_1}(z) = D(x_1, z)$. □

We have that $H_\infty(X_1, X_2) \geq n_1 + n_2 - \Delta$ which implies that $H_\infty(X_1) \geq n_1 + n_2 - \Delta - n_2 = n_1 - \Delta$. By the guarantee of SExt, we have that

$$U_{m_1} \overset{a}{\sim}_{\epsilon} \mathsf{SExt}(X_1, U_{m_2}).$$

By Claim 4.12 and the data processing inequality of Proposition 4.10 we can conclude that

$$\mathsf{SExt}(X_1, U_{m_2}) \overset{m}{\sim}_{(\eta,\epsilon)} \mathsf{SExt}(X_1, \mathsf{Ext}(X_2))$$

and note that the latter distribution is $E(X_1, X_2)$. Using proposition 4.9 we conclude that

$$U_{m_1} \overset{m}{\sim}_{(\eta, e^{\eta}\cdot\epsilon+\epsilon)} E(X_1, X_2)$$

as required.

4.4 Proof of Theorem 4.3

Theorem 4.3 now follows immediately from composing the extractor of Theorem 4.2 with a seeded extractor (say the one by Guruswami, Umans and Vadhan [GUV07]) using Lemma 4.7. The precise calculation follows.

We are assuming that E is hard for exponential size nondeterministic circuits. Let $\alpha > 0$ be the constant guaranteed by Theorem 4.2. (This constant is not universal, and depends on the constants in the hardness assumption). Let $\gamma > 0$ be a constant, such that $\gamma < \frac{\alpha}{100}$ and let $c > 1$ be a constant. We are aiming to construct $E : \{0,1\}^n \to \{0,1\}^{(1-O(\gamma))\cdot n}$ that is a $((1-\gamma)\cdot n, \overset{m}{\sim}_{(n^{-c}, 2^{-\Omega(\gamma\cdot n)})})$-extractor for distributions samplable by size n^c deterministic circuits with postselection by size n^c nondeterministic circuits.

We set $t = \frac{\alpha}{2\gamma} \geq 50$. We plan to use Lemma 4.7 and define $\Delta = \gamma \cdot n$ and write n as $n = n_1 + n_2$ for $n_2 = \frac{n}{t}$ and $n_1 = \frac{t-1}{t} \cdot n = (1 - \frac{1}{t}) \cdot n$. We now choose the two components Ext and SExt for lemma 4.7.

- We take $\mathsf{Ext} : \{0,1\}^{n_2} \to \{0,1\}^{\alpha n_2}$ to be a $((1-\alpha)\cdot n_2, \overset{m}{\sim}_{n^{-(c+1)}})$-extractor for distributions samplable by deterministic circuits of size n^{c+1} with postselection by nondeterministic circuits of size n^{c+1} that is guaranteed by Theorem 4.2.

- By Theorem 3.11, there exists a constant a such that for $\epsilon \leq \frac{1}{n}$, the extractor guaranteed in Theorem 3.11 has seed length $a \cdot \log(1/\epsilon)$. We will assume w.l.o.g. that $a \geq 2$. We set $\epsilon = 2^{-\frac{\gamma \cdot n_1}{a}}$ and using Theorem 3.11, we obtain a seeded $((1 - \frac{t}{t-1} \cdot \gamma) \cdot n_1, \epsilon)$-seeded extractor $\mathsf{SExt} : \{0,1\}^{n_1} \times \{0,1\}^{d_1} \to \{0,1\}^{m_1}$ for $d_1 = a \log(1/\epsilon) = \gamma n_1$ and

$$m_1 = (1-\gamma) \cdot (1 - \frac{t}{t-1} \cdot \gamma) \cdot n_1 \geq (1 - 3\gamma) \cdot n_1 \geq (1 - 3\gamma)(1 - \frac{1}{t}) \cdot n = (1 - O(\gamma)) \cdot n,$$

  where the hidden constant depends on α which depends on the constant β from the hardness assumption.

To apply Lemma 4.7 we need to check that:

- The output length of Ext is at least the seed length of SExt. We need that $\alpha \cdot n_2 \geq d_1 = \gamma \cdot n_1 = \gamma \cdot (t-1) \cdot n_2$. We indeed have that $\alpha \geq \gamma t$, by our choice of $t = \frac{\alpha}{2\gamma}$.
- The entropy deficiency of Ext is larger than $\Delta + \log(1/\epsilon)$. We need that $\alpha \cdot n_2 \geq \gamma \cdot n + \frac{\gamma \cdot n_1}{a}$. We indeed have that $\alpha \cdot n_2 = \alpha \cdot \frac{n}{t} = 2\gamma \cdot n \geq \gamma \cdot n + \frac{\gamma \cdot n_1}{a}$.

By Lemma 4.7 we conclude that for $\eta = \frac{1}{n^c}$, and $\rho = e^{\eta} \cdot \epsilon + \epsilon$, the function $E : \{0,1\}^n \to \{0,1\}^{(1-O(\gamma)) \cdot n}$ defined in the lemma is a $((1-\gamma) \cdot n, \overset{m}{\sim}_{(e^{\eta}, \rho)})$-extractor for distributions samplable by size n^c deterministic circuits with postselection by size n^c nondeterministic circuits. We have that $\rho \leq 3\epsilon = 3 \cdot 2^{-\frac{\gamma n_1}{a}} = 2^{-\Omega(\gamma \cdot n)}$.

4.5 Proof of Claim 2.7

We have that: $p_1 = \Pr[D(W, V, R) = 1]$ and $p_2 = \Pr[D(W, V, \mathsf{SExt}(\hat{f}(W), V)) = 1]$. Define $v_1(w) = \Pr[D(w, V, R) = 1]$ and $v_2(w) = \Pr[D(w, V, \mathsf{SExt}(\hat{f}(w), V)) = 1]$, so that $p_1 = \mathbb{E}_{w \leftarrow \mathbb{F}_q^d}[v_1(w)]$ and $p_2 = \mathbb{E}_{w \leftarrow \mathbb{F}_q^d}[v_2(w)]$. Recall that we are considering a probability space where r distinct elements $t_1, \ldots, t_r$ are chosen from $\mathbb{F}_q \setminus \{0\}$ and an independent experiment where r (not necessarily distinct) elements $y_1, \ldots, y_r \leftarrow \mathbb{F}_q^d$ are chosen. For every $x \in \mathbb{F}_q^d$ we have defined the random variable $C_x = C^{0, t_1, \ldots, t_r}_{x, y_1, \ldots, y_r}(t)$ which is the unique degree r curve that passes through the points $(0, x), (t_1, y_1), \ldots, (t_r, y_r)$. For every $x \in \mathbb{F}_q^d$ and every $t \in \mathbb{F}_q \setminus \{0\}$ we define the random variable $R_t^x = C_x(t)$. Note that for every $x \in \mathbb{F}_q^d$, the random variables $(R_t^x)_{t \in \mathbb{F}_q \setminus \{0\}}$ are r-wise independent. This follows in a standard way because although the curve C_x passes through the fixed point $(0, x)$ (and is not completely random) it passes through r random points, and there are still "sufficiently many degrees of freedom" to get r-wise independence. More formally, note that for every fixing of $t_1, \ldots, t_r$, and every choice of t distinct $t'_1, \ldots, t'_r \in \mathbb{F}_q \setminus \{0\}$, there is an invertible map from the r-tuple $y_1 = C_x(t_1), \ldots, y_d = C_x(t_r)$ to the r tuple $C_x(t'_1), \ldots, C_x(t'_r))$.

This means that we can apply the r-wise tail inequality from Theorem 3.15 to argue that for every $x \in \mathbb{F}_q^d$, p_1 and p_2 are with high probability approximated by $p_{x,1}$ and $p_{x,2}$. More specifically, for every $x \in \mathbb{F}_q^d$, Theorem 3.15 implies that the probability that $|p_1 - p_{x,1}| > \epsilon$ is at most

$$8 \cdot \left(\frac{2r}{\epsilon^2 (q-1)} \right)^{r/2} \leq \frac{1}{20 q^d},$$

where the last inequality follows because we can choose the constants c_r, c_q in the definition of $r = c_r \cdot d$ and $q = \frac{2^m}{\rho^{c_q}}$ to be sufficiently large so that $\frac{2r}{\epsilon^2 \cdot (q-1)} \leq \frac{1}{\sqrt{q}}$ and $r \geq 10d$. The same reasoning gives that the probability that $|p_2 - p_{x,2}| > \epsilon$ is at most $\frac{1}{20q^d}$, and the claim follows by a union bound over these two events.

5 Consequences of Extractors for Samplable Distributions

The extractors for samplable distributions of Theorem 4.2 relies on the assumption that E is hard for exponential size nondeterministic circuits, which is then used to obtain a multiplicative seed-extending PRG for nondeterministic circuits in Theorem 2.3, which is then shown to be an extractor in Lemma 4.5. This raises two natural questions:

- Is the hardness assumption (which assumes hardness against nondeterministic circuits) necessary for such extractors?
- Do extractors for samplable distributions imply seed-extending PRGs for nondeterministic circuits?

While we do not give a definitive answer to these questions, we prove results that shed more light on these two questions. We state and discuss our results in Section 5.1 and the proofs are given in Section 5.2.

5.1 Results and Discussion

5.1.1 Extractors that Imply Lower Bounds Against Nondeterministic Circuits

While we do not know whether extractors for samplable distributions imply lower bounds against nondeterministic circuits, we can show that extractors for distributions samplable by deterministic circuits with postselection by nondeterministic circuits *do* imply lower bounds against nondeterministic circuits.

Recall that our Theorems 4.2 and 4.3 yield extractors for this richer class of distributions, and (as explained in Remark 4.1) the proof techniques used in previous work (as well as the proofs this paper) immediately extend to give extractors for this richer class.

This means that hardness for nondeterministic circuits cannot be avoided as long as we use proof techniques that give extractors for this richer class of distributions.

This result is stated formally in the lemma below.

Lemma 5.1 (Extractors for the richer class imply lower bounds for nondeterministic circuits)**.** *If* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ *is an* $(n-10, \overset{a}{\sim}_{\frac{1}{100}})$*-extractor for distributions samplable by size* $s \geq n$ *deterministic circuits with postselection by size* s *nondeterministic circuits, then* Ext *cannot be computed by circuits of size* $s - O(1)$ *nondeterministic circuits.*

Trevisan and Vadhan [TV00] showed that if Ext is an $(n-1, \overset{a}{\sim}_{\frac{1}{5}})$-extractor for distributions samplable by size s circuits then Ext cannot be computed by *deterministic* circuits of slightly smaller size. In Lemma 5.1 we get a lower bound against *nondeterministic circuits* if the extractor is for the richer class.

We remark that the constants stated in Lemma 5.1 can be improved by a more careful argument, and the same conclusion holds for an $(n-1, \overset{a}{\sim}_{\frac{1}{5}})$-extractor. We also remark that a stronger conclusion in which the concluded lower bound is an average-case lower bound is stated below in Lemma 5.3.

A detailed comparison between the conclusion of Lemma 5.1 and the hardness assumption of Theorem 4.2 appears in Section 5.1.3

5.1.2 Extractors that Imply Seed-Extending PRGs for Nondeterministic Circuits

Lemma 4.5 shows that (for certain parameters) multiplicative PRGs for nondeterministic circuits are extractors for the richer class of distribution samplable by deterministic circuits with postselection by nondeterministic circuits. We are able to prove a partial converse, stated below.

Lemma 5.2. *Let* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *be a function.*

- *If* Ext *is an* $(n-\Delta, \overset{m}{\sim}_\epsilon)$*-extractor for distributions samplable by size* $s \geq n$ *deterministic circuits with postselection by size* s *nondeterministic circuits, and* $\Delta \geq m + \log(1/\epsilon)$ *then* Ext *is a seed-extending* $\overset{a}{\sim}_{4\epsilon}$*-PRG for nondeterministic circuits of size* $s' = s - O(m)$.
- *If* $m = 1$ *and* Ext *is an* $(n-\Delta, \overset{a}{\sim}_\epsilon)$*-extractor for distributions samplable by size* $s \geq n$ *deterministic circuits with postselection by size* s *nondeterministic circuits, and* $\Delta \geq \log(1/\epsilon)$ *then* Ext *is a seed-extending* $\overset{a}{\sim}_{4\epsilon}$*-PRG for nondeterministic circuits of size* $s' = s - O(1)$.
- *If* $m = 1$ *and Ext is an* $(n-\Delta, \overset{a}{\sim}_\epsilon)$*-extractor for distributions samplable by size* $s \geq n$ *deterministic circuits, and* $\Delta \geq \log(1/\epsilon)$ *then* Ext *is a seed-extending* $\overset{a}{\sim}_{5\epsilon}$*-PRG for deterministic circuits of size* $s' = \Omega(\frac{s}{\log(1/\epsilon)})$.

The proof of Lemma 5.2 appears in Section 5.2.1.

Comparing the conclusion of Lemma 5.2 to the assumption of Lemma 4.5.

- The first item of Lemma 5.2 asserts that for every output length m, a multiplicative extractor for the richer class implies an (additive) seed-extending PRG for nondeterministic circuits. As we only get an additive PRG in the conclusion of the first item of Lemma 5.2, the conclusion is not sufficiently strong to apply Lemma 4.5. We do not know whether the conclusion of the first item of Lemma 5.2 can be strengthened to give a multiplicative PRG.
- The second item of Lemma 5.2 asserts that for output length $m = 1$, even the weaker notion of an additive extractor implies an (additive) seed-extending PRG for nondeterministic circuits. Note that it is trivial that a $\overset{a}{\sim}_{4\epsilon}$-seed-extending PRG is automatically also a $\overset{m}{\sim}_{(4\epsilon,4\epsilon)}$-PRG. This means that the conclusion of the second item of Lemma 5.2 is sufficiently strong to apply Lemma 4.5.

  This means that for $s = n^{O(1)}$, $m = 1$ and $\epsilon = \frac{1}{\mathsf{poly}(n)}$ we have the following: An $(n - \log(1/\epsilon), \overset{a}{\sim}_\epsilon)$-extractor for the richer class of distributions samplable by deterministic circuits with postselection by nondeterministic circuits, is *essentially equivalent* to a seed-extending $\overset{a}{\sim}_\epsilon$-PRGs for nondeterministic circuits (in the sense that starting from one, one can get the other while only increasing ϵ by a constant factor).

  We stress however that the extractors of [TV00, AASY15, BGDM23] and this paper, consider distributions with min-entropy $n - \Delta$ for $\Delta = \Omega(n)$, and not $\Delta = \log(1/\epsilon) = O(\log n)$, which is the parameter regime for which the equivalence above holds.
- The third item of Lemma 5.2 discusses the original class of distributions samplable by deterministic circuits (rather than the richer class discussed in the two items above) and in this case the PRG in the conclusion fools deterministic circuits (and not nondeterministic circuits).

5.1.3 Extractors that imply hard on average functions

The results of Lemma 5.1 can be strengthened to imply functions that are not only hard on the worst-case for nondeterministic circuits, but rather hard on average for nondeterministic circuits. This is stated next using the terminology of Definition 3.6.

Lemma 5.3 (Extractors for the richer class imply average-case hardness for nondeterministic circuits)**.** *If* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ *is an* $(n-\Delta, \overset{a}{\sim}_\epsilon)$*-extractor for distributions samplable by size* $s \geq n$ *deterministic circuits with postselection by size* s *nondeterministic circuits, and* $\Delta \geq \log(1/\epsilon)$ *then* Ext *is* $(\frac{1}{2} + O(\epsilon))$*-hard for nondeterministic circuits of size* $s - O(1)$.

Comparing the conclusion of Lemma 5.3 to the assumption of Theorem 4.2. The extractors of [BGDM23] and of our Theorems 4.2 and Theorem 4.3 are constructed under the assumption that E is hard for exponential size nondeterministic circuits. This assumption is stronger than the conclusion of Lemma 5.3 for $s = n^c$.

While it is true that both assumptions involve lower bounds against nondeterministic circuits. If we apply Lemma 5.3 with $s = n^c$ and an extractor that is computed in time $\mathsf{poly}(n^c)$, we obtain a function $f : \{0,1\}^n \to \{0,1\}$ such that:

- f is $(\frac{1}{2} + \frac{1}{n^c})$-hard for nondeterministic circuits of size n^c (for every sufficiently large n).
- f is computable in time $\mathsf{poly}(n^c)$.

However, if we scale the assumption that E is hard for exponential size nondeterministic circuits in the same way (by taking the input length of the hard function from Definition 3.5 to be $\ell = \frac{c}{\beta} \cdot \log n$) then we obtain a function $g : \{0,1\}^{\frac{c}{\beta} \cdot \log n} \to \{0,1\}$ such that:

- g cannot be computed by nondeterministic circuits of size n^c (for every sufficiently large n).
- g is computable in time $\mathsf{poly}(n^c)$.

In this range (where the function is computed in time exponential in its input length) there are hardness amplification results [IW97, STV01] that transform a worst-case hard function into an average-case hard function. Furthermore, there are such results that work against nondeterministic circuits [SU06], and using these results would immediately give that g is $(\frac{1}{2} + \frac{1}{n^c})$-hard for nondeterministic circuits of size n^c (achieving the same hardness as the function f, but on a shorter input length).

The reason that we highlight the difference between worst-case and average-case hardness in the consequence of Lemma 5.3, is that in the parameter regime of the function f (where the running time is polynomial in the input length) we do not know how to transform worst-case hard functions into average-case hard functions. Furthermore, when working against nondeterministic circuits, even tools like Yao's XOR lemma (that transform weak average-case hardness into strong average-case hardness) are not applicable. Indeed, this is why we feel that there is a big difference between showing that extractors imply worst-case hardness (as in Lemma 5.1) and showing that extractors imply average-case hardness (as in Lemma 5.3).

Summing up, the function g that is obtained from the hardness assumption that E is hard for exponential size nondeterministic circuits, implies the function f above, as we can always artificially lengthen the input length of g to match that of f. The converse is not true, and it seems that assuming the existence of a function like g is a stronger assumption than the existence of a function like f.

An average-case hard function in the case of samplable distributions. We can also prove a version of Lemma 5.3 for the original class of distributions samplable by size s circuits (rather than the richer class of distributions samplable by size s circuits with postselection by size s nondeterministic circuits) but here, in the conclusion we only get an average-case lower bound against *deterministic circuits*.

Lemma 5.4 (Extractors for samplable distributions are hard on average for deterministic circuits)**.** *If* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ *is an* $(n - \log(1/\epsilon), \overset{a}{\sim}_\epsilon)$*-extractor for distributions samplable by deterministic circuits of size* $s \geq n$*, then* Ext *is* $(\frac{1}{2} + O(\epsilon))$*-hard for deterministic circuits of size* $s' = \Omega(\frac{s}{\log(1/\epsilon)})$.

5.2 Proofs

In this section we prove the results stated in the previous section.

5.2.1 Proof of Lemma 5.2

The proof of Lemma 5.2 adapts and extends an argument by Kinne, Shaltiel and van Melkebeek [KvMS09], that shows that extractors with exponentially small error for distributions recognizable by deterministic circuits yield seed-extending PRGs for deterministic circuits.

Proof. We start with proving the first item. Let $T : \{0,1\}^n \times \{0,1\}^m \to \{0,1\}$ be a size s' nondeterministic circuit. We consider the probability space where $X \leftarrow U_n$ and $R \leftarrow U_m$ are chosen independently. We will show that the function $G(x) = (x, \mathsf{Ext}(X))$ is a $\overset{a}{\sim}_{4\epsilon}$-PRG for T. For this purpose we compute:

$$\begin{aligned}
p_1 &= \Pr[T(X,R)=1] \\
&= \sum_{r\in\{0,1\}^m} \Pr[T(X,r)=1 \wedge R=r] \\
&= \sum_{r\in\{0,1\}^m} \Pr[T(X,r)=1]\cdot\Pr[R=r] \\
&= \sum_{r\in\{0,1\}^m} \Pr[T(X,r)=1]\cdot 2^{-m}. \\
p_2 &= \Pr[T(X,\mathsf{Ext}(X))=1] \\
&= \sum_{r\in\{0,1\}^m} \Pr[T(X,r)=1 \wedge \mathsf{Ext}(X)=r] \\
&= \sum_{r\in\{0,1\}^m} \Pr[T(X,r)=1]\cdot\Pr[\mathsf{Ext}(X)=r|T(X,r)=1]
\end{aligned}$$

For every $r \in \{0,1\}^m$, we define a nondeterministic circuit $P_r : \{0,1\}^n \to \{0,1\}$ as follows: $P_r(x) = 1$ iff $T(x,r) = 1$. Note that for every $r \in \{0,1\}^m$, P_r is a nondeterministic circuit of size $s' + O(m) = s$. Let $G = \left\{r : \Pr[T(X,r)=1] \geq 2^{-\Delta}\right\}$. Note that for every $r \in G$, the distribution

$$X_r = (X|T(X,r)=1) = (X|P_r(X)=1)$$

has min-entropy at least $n - \Delta$. Furthermore X_r is the distribution sampled by the circuit $A(x)$ which computes the identity function, with postselection by the nondeterministic circuit P_r.[20] It follows that for every $r \in G$, we have that

$$\Pr[\mathsf{Ext}(X)=r|T(X,r)=1] - 2^{-m} = \Pr[\mathsf{Ext}(X_r)=r] - 2^{-m} \leq e^{\epsilon}\cdot 2^{-m} - 2^{-m} \leq 3\epsilon\cdot 2^{-m}.$$

[20] In fact, X_r is a distribution recognizable by the nondeterministic circuit P_r which is of size s, and this proof works assuming that Ext is an extractor for such distributions.

Therefore, we have that:

$$\begin{aligned}
p_2 - p_1 &= \sum_{r\in\{0,1\}^m} \Pr[T(X,r)=1]\cdot(\Pr[\mathsf{Ext}(X)=r|T(X,r)=1]-2^{-m}) \\
&= \sum_{r\in G} \Pr[T(X,r)=1]\cdot(\Pr[\mathsf{Ext}(X)=r|T(X,r)=1]-2^{-m}) \\
&+ \sum_{r\in\{0,1\}^m\setminus G} \Pr[T(X,r)=1]\cdot(\Pr[\mathsf{Ext}(X)=r|T(X,r)=1]-2^{-m}) \\
&\leq \sum_{r\in G} \Pr[T(X,r)=1]\cdot 3\epsilon\cdot 2^{-m} + \sum_{r\in\{0,1\}^m\setminus G} \Pr[T(X,r)=1] \\
&\leq 2^m\cdot 3\epsilon\cdot 2^{-m} + (2^m-|G|)\cdot 2^{-\Delta} \\
&\leq 3\epsilon+\epsilon \\
&\leq 4\epsilon.
\end{aligned}$$

We now prove the second item. We are now assuming that $m=1$, and observe that the only place where we used the assumption that the extractor is w.r.t $\overset{m}{\sim}_\epsilon$ and not $\overset{a}{\sim}_\epsilon$ is when we argued that

$$\Pr[\mathsf{Ext}(X)=r|T(X,r)=1]-2^{-m} \leq e^\epsilon\cdot 2^{-m}-2^{-m} \leq 3\epsilon\cdot 2^{-m}.$$

Note that for $m=1$ and an extractor w.r.t. $\overset{a}{\sim}_\epsilon$ we have that:

$$\Pr[\mathsf{Ext}(X)=r|T(X,r)=1]-2^{-m} = \Pr[\mathsf{Ext}(X_r)=r]-2^{-1} \leq 2^{-1}+\epsilon-2^{-1} \leq \epsilon = 2\epsilon\cdot 2^m,$$

and exactly the same proof applies, noting that there is at most one $r\in\{0,1\}^m\setminus G$.

We now consider the third item. We are now assuming that $m=1$ and that the relation is $\overset{a}{\sim}_\epsilon$ (as in the second item). However, here we are assuming that the class of distributions that Ext extracts from is without postselection, and are aiming to show that the PRG is against deterministic circuits, rather than nondeterministic circuits. We can repeat computation of p_1, p_2 under the assumption that T is a *deterministic* circuit of size s'.

In the proof of the first two items, we argued that for every $r\in G$, $X_r = (X|T(X,r)=1)$ is a source from the class for which Ext was designed. This no longer holds, as it is not clear that for $r\in G$, X_r is samplable by small deterministic circuits. Nevertheless, we can argue instead that for every $r\in G$, there exists a distribution X'_r over $\{0,1\}^n$ such that:

- X'_r is samplable by size $s=O(s'\cdot\log(1/\epsilon))$ circuits.
- X'_r is ϵ-close to X_r.
- $H_\infty(X'_r)\geq H_\infty(X_r)$.

We will sample X'_r as follows. Choose $t=\log(1/\epsilon)+1$ independent $X_1,\ldots,X_t\leftarrow U_n$. If there does not exist $i\in[t-1]$ such that $T(X_i,r)=1$, we output $X'_r=X_t$. Otherwise let i' be the smallest $i\in[t-1]$ such that $T(X_{i'},r)=1$, and we output $X_{i'}$.

The three properties above immediately follow for every $r\in G$, and we can conclude that for every $r\in G$,

$$\Pr[\mathsf{Ext}(X)=r|T(X,r)=1] = \Pr[\mathsf{Ext}(X_r)=r] \leq \Pr[\mathsf{Ext}(X'_r)=r]+\epsilon \leq 2^{-1}+2\epsilon.$$

This gives that for $m=1$:

$$\Pr[\mathsf{Ext}(X_r)=r]-2^{-m} \leq 2\epsilon = 4\epsilon\cdot 2^{-m}.$$

and we can conclude the proof as in the second item. □

5.2.2 A Seed-extending PRG for Nondeterministic Circuits is an Average-Case Hard Function

In this section we prove Lemma 5.1, Lemma 5.3 and Lemma 5.4. This is done by the following argument: By Lemma 5.2, the assumptions of the three considered Lemmata imply certain seed-extending PRGs, and the next lemma shows that such PRGs are functions that are hard on average. This immediately implies the three Lemmata.

Lemma 5.5 (Seed-extending PRGs for nondeterministic circuits are hard on average for nondeterministic circuits)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a function that is a seed-extending $\overset{a}{\sim}_\epsilon$-PRG for nondeterministic circuits of size $s \geq n$, then f is $(\frac{1}{2} + 9\epsilon)$-hard for nondeterministic circuits for circuits of size $s - O(1)$.*

Remark 5.6. *It is easy and standard to prove a version of Lemma 5.5 in the case of* deterministic *circuits. More specifically, given a function $f : \{0,1\}^n \to \{0,1\}$, if the distribution $(X, f(X))$ for $X \leftarrow U_n$ is pseudorandom for size s deterministic circuits, w.r.t $\overset{a}{\sim}_\epsilon$, then for every circuit $C : \{0,1\}^n \to \{0,1\}$ that attempts to compute f, we can define the circuit $D : \{0,1\}^n \times \{0,1\} \to \{0,1\}$ that on input (x, y) answers one iff $C(x) = y$.*

It immediately follows that $\Pr[D(X, f(X)) = 1] = \Pr[C(X) = f(X)]$ and $\Pr[D(X, U_1) = 1] = \frac{1}{2}$, which provides the required result as D (that is of size slightly larger than C) cannot distinguish between $(X, f(X))$ and (X, U_1).

Note however, that if we assume that C is a nondeterministic circuit of size s, then it does not follow that D can be implemented by a small nondeterministic circuit. This issue (that a nondeterministic procedure that uses a nondeterministic procedure does not yield a nondeterministic procedure) causes many arguments that hold for deterministic circuits to fail for nondeterministic circuits.

Somewhat surprisingly (at least to us) it turns out that in this case the problem can be bypassed. This is done by considering a different distinguisher $D(x, y)$ that checks whether $C(x) = 1$ and $y = 1$. Note that this computation can be performed by a small nondeterministic circuit (assuming C is a small nondeterministic circuit) and in the proof below, we will show that it proves the Lemma.

Proof. Note that by assumption we have that

$$\frac{1}{2} - \epsilon \leq \Pr_{X \leftarrow U_n}[f(X) = 1] \leq \frac{1}{2} + \epsilon.$$

This is because we can consider the deterministic circuits $D(x, y) = y$ and $D(x, y) = 1 - y$, and we are guaranteed that $G(x) = (x, f(x))$ is a $\overset{a}{\sim}_\epsilon$-PRG against both of them.

Assume for contradiction that there exists a nondeterministic circuit C of size s' such that

$$\Pr_{X \leftarrow U_n}[C(X) = f(X)] > \frac{1}{2} + 9\epsilon.$$

We will construct a size s nondeterministic circuit $D : \{0,1\}^n \times \{0,1\} \to \{0,1\}$ such that the function $G(x) = (x, f(x))$ is not an $\overset{a}{\sim}_\epsilon$-PRG for D. Given inputs $(x, y) \in \{0,1\}^n \times \{0,1\}$, $D(x, y)$ will answer one iff $C(x) = 1$ and $y = 1$. Note that D is indeed a nondeterministic circuit of size $s' + O(1) = s$. We will consider an experiment in which $X \leftarrow U_n$ and $R \leftarrow \{0,1\}$ are chosen independently. We have that:

$$p_1 = \Pr[D(X, R) = 1] = \Pr[C(X) = 1 \wedge R = 1] = \Pr[C(X) = 1] \cdot \Pr[R = 1] = \frac{1}{2} \cdot \Pr[C(X) = 1].$$

$$p_2 = \Pr[D(X, f(X)) = 1] = \Pr[C(X) = 1 \wedge f(X) = 1] = \Pr[f(X) = 1] \cdot \Pr[C(X) = 1 | f(X) = 1].$$

We will get a contradiction by showing that $p_2 > p_1 + \epsilon$. Let us denote $a_i := \Pr[C(X) = 1 | f(x) = i]$, and compute:

$$\begin{aligned}
\frac{1}{2} + 9\epsilon &< \Pr[C(X) = f(X)] \\
&= \Pr[f(X) = 1] \cdot \Pr[C(X) = 1 | f(X) = 1] + \Pr[f(X) = 0] \cdot \Pr[C(X) = 0 | f(X) = 0] \\
&= \Pr[f(X) = 1] \cdot a_1 + \Pr[f(X) = 0] \cdot (1 - a_0) \\
&= \Pr[f(X) = 0] + \Pr[f(X) = 1] \cdot a_1 - \Pr[f(X) = 0] \cdot a_0 \\
&\leq \frac{1}{2} + \epsilon + (\frac{1}{2} + \epsilon) \cdot a_1 - (\frac{1}{2} - \epsilon) \cdot a_0 \\
&\leq \frac{1}{2} + 3\epsilon + \frac{1}{2} \cdot (a_1 - a_0)
\end{aligned}$$

We have obtained that:

$$\frac{a_1 - a_0}{2} > 9\epsilon - 3\epsilon = 6\epsilon$$

We observe that:

$$a := \Pr[C(X) = 1] = \Pr[f(X) = 0] \cdot a_0 + \Pr[f(X) = 1] \cdot a_1 \leq (\frac{1}{2} + \epsilon)(a_0 + a_1) \leq \frac{a_0 + a_1}{2} + 2\epsilon.$$

As the average $\frac{a_0 + a_1}{2} = a_1 - \frac{a_1 - a_0}{2}$, we conclude that: $a - 2\epsilon \leq \frac{a_0 + a_1}{2} = a_1 - \frac{a_1 - a_0}{2}$ which gives:

$$a_1 - a \geq \frac{a_1 - a_0}{2} - 2\epsilon > 4\epsilon.$$

Finally, we can compute

$$\begin{aligned}
p_2 - p_1 &= \Pr[f(X) = 1] \cdot \Pr[C(X) = 1 | f(X) = 1] - \frac{1}{2} \cdot \Pr[C(X) = 1] \\
&\geq (\frac{1}{2} - \epsilon) \cdot a_1 - \frac{1}{2} a \\
&\geq \frac{a_1 - a}{2} - \epsilon \\
&> \epsilon.
\end{aligned}$$

□

6 Discussion and Open Problems

We now describe some open problems and research directions.

6.1 Extractors for Samplable Distributions with Lower Min-Entropy

The constructions of extractors for samplable distribution in the literature [TV00, AASY15, BGDM23] and the results of this paper, achieve extractors only for very large min-entropy. Specifically, min-entropy $k = n - \alpha n$ where $\alpha > 0$ is a small constant.

It is a longstanding open problem to construct extractors for lower min-entropy. The current approach of [TV00] and subsequent work (including this paper) fails for $k < n/2$, and it seems that new ideas are required.

6.2 Other Notions of Multiplicative Extractors

In this paper we focus on extractors w.r.t. $\overset{m}{\sim}_\epsilon$. Previous work by Applebaum et al. [AASY15] (that we have already described) considered a version which they call "extractors with relative error" which in the notation of this paper is an extractor w.r.t. the relation $\overset{md}{\sim}_\epsilon$ define by:

$$p_1 \overset{md}{\sim}_\epsilon p_2 \iff p_1 \overset{m}{\sim}_\epsilon p_1 \text{ and } p_2 \overset{m}{\sim}_\epsilon p_1.$$

This double sided multiplicative notion is stronger than the one we consider here. However, we are not aware of an application which calls for extractors w.r.t. $\overset{md}{\sim}_\epsilon$ rather than extractors w.r.t. $\overset{m}{\sim}_\epsilon$. (Recall that extractors w.r.t. $\overset{m}{\sim}_\epsilon$ do give the standard double sided additive notion, as explained in Section 1).

Nevertheless, we mention that the results of this paper easily imply extractors w.r.t. $\overset{md}{\sim}_\epsilon$ under the assumption that E is hard for exponential size Σ_2-circuits, improving upon the result of Applebaum et al. [AASY15] which achieves such extractors under the assumption that E is hard for exponential size Σ_4-circuits.

It is open whether extractors w.r.t. $\overset{md}{\sim}_\epsilon$ follow under the weaker assumption used to obtain extractors w.r.t. $\overset{m}{\sim}_\epsilon$ in Theorem 4.2.

Loosely speaking, the reason that we need a stronger assumption to achieve the double sided relation is that while nondeterministic circuits can verify that the number of accepting inputs of a given nondeterministic circuit is larger than some threshold (as stated precisely in Section 3.8) it seems that they cannot verify that the fraction is smaller than some threshold, and this (as far as we know) seems to require Σ_2-circuits. This difference makes the reduction more expensive, resulting in the need to assume a stronger hardness assumption.

6.3 Multiplicative PRGs with Larger Stretch

The multiplicative seed-extending PRG of Theorem 2.3 achieves modest stretch. Its output length is shorter than its input length. (This means that even when including the seed in the output the PRG has mild stretch). While this is suitable for the application of constructing extractors, it is natural to ask whether it is possible to construct a multiplicative PRG with larger stretch.

A multiplicative PRG with larger stretch was constructed by Artemenko et al. [AIKS16]. This PRG builds on a stronger hardness assumption (the assumption that E is hard for exponential size Σ_3-circuits) and fools deterministic circuits. It stretches a seed length of length $n^{\Omega(1)}$ into n bits and is a $\overset{m}{\sim}_{(\frac{1}{n^{O(1)}},\rho)}$-PRG for $\rho = 2^{-\Omega(\sqrt{n})}$. As pointed out in [AIKS16] one could expect a better dependence of the seed length on the error parameter ρ.

We believe that using the techniques on this paper (as well as several additional ideas) it is possible to construct a multiplicative PRG with large stretch and correct dependence of the seed length on the error parameter. Moreover, we believe that this can be achieved under the weak assumption that E is hard for exponential size nondeterministic circuits.

6.4 Minimal Assumptions for Extractors for Samplable Distribution

We showed that a lower bound against nondeterministic circuits follows from extractors for the richer class of distributions samplable by deterministic circuits with postselection by nondeterministic circuits. Does this follow for the original class of distributions samplable by deterministic circuits? Maybe one can hope to construct extractors for this class without assuming a lower bound for nondeterministic circuits?

Even for the richer class, the lower bound obtained as a consequence of extractors in Section 5 is not known to imply these extractors. Can we bridge this gap?

6.5 Connection between Extractors for Samplable Distributions and Seed-Extending PRGs

The results of this paper show that extractors for distributions samplable by deterministic circuits with post-selection by nondeterministic circuits are closely related to seed-extending PRGs for nondeterministic circuits (at least for some choice of parameters). Can this connection be extended to other parameter regimes? A more precise statement of this problem appears in Section 5.

References

[AASY15] B. Applebaum, S. Artemenko, R. Shaltiel, and G. Yang. Incompressible functions, relative-error extractors, and the power of nondeterministic reductions. In *30th Conference on Computational Complexity*, pages 582–600, 2015.

[AIKS16] S. Artemenko, R. Impagliazzo, V. Kabanets, and R. Shaltiel. Pseudorandomness when the odds are against you. In *31st Conference on Computational Complexity, CCC*, volume 50, pages 9:1–9:35, 2016.

[AK02] V. Arvind and J. Köbler. New lowness results for ZPP^{NP} and other complexity classes. *J. Comput. Syst. Sci.*, 65(2):257–277, 2002.

[AS14] S. Artemenko and R. Shaltiel. Pseudorandom generators with optimal seed length for non-boolean poly-size circuits. In *Symposium on Theory of Computing, STOC*, pages 99–108, 2014.

[BDL22] M. Ball, D. Dachman-Soled, and J. Loss. (nondeterministic) hardness vs. non-malleability. In *Advances in Cryptology - CRYPTO 2022 - 42nd Annual International Cryptology Conference*, volume 13507, pages 148–177, 2022.

[BGDM23] M. Ball, E. Goldin, D. Dachman-Soled, and S. Mutreja. Extracting randomness from samplable distributions, revisited. In *64th IEEE Annual Symposium on Foundations of Computer Science, FOCS*, pages 1505–1514, 2023.

[BOV07] B. Barak, S. J. Ong, and S. P. Vadhan. Derandomization in cryptography. *SIAM J. Comput.*, 37(2):380–400, 2007.

[BR94] M. Bellare and J. Rompel. Randomness-efficient oblivious sampling. In *35th Annual Symposium on Foundations of Computer Science*, pages 276–287, 1994.

[BSS24] M. Ball, R. Shaltiel, and J. Silbak. Non-malleable codes with optimal rate for poly-size circuits. In *Advances in Cryptology - EUROCRYPT*, volume 14654 of *Lecture Notes in Computer Science*, pages 33–54, 2024.

[BV17] N. Bitansky and V. Vaikuntanathan. A note on perfect correctness by derandomization. In *Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, volume 10211, pages 592–606, 2017.

[CT22] L. Chen and R. Tell. When arthur has neither random coins nor time to spare: Superfast derandomization of proof systems. *Electron. Colloquium Comput. Complex.*, TR22-057, 2022.

[DMNS06] C. Dwork, F. McSherry, K. Nissim, and A. D. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography, Third Theory of Cryptography Conference, TCC*, volume 3876 of *Lecture Notes in Computer Science*, pages 265–284. Springer, 2006.

[DMOZ22] D. Doron, D. Moshkovitz, J. Oh, and D. Zuckerman. Nearly optimal pseudorandomness from hardness. *J. ACM*, 69(6):43:1–43:55, 2022.

[Dru13] Andrew Drucker. Nondeterministic direct product reductions and the success probability of SAT solvers. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS*, pages 736–745, 2013.

[FL97] U. Feige and C. Lund. On the hardness of computing the permanent of random matrices. *Computational Complexity*, 6(2):101–132, 1997.

[GS86] S. Goldwasser and M. Sipser. Private coins versus public coins in interactive proof systems. In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing*, pages 59–68, 1986.

[GST03] Dan Gutfreund, Ronen Shaltiel, and Amnon Ta-Shma. Uniform hardness versus randomness tradeoffs for arthur-merlin games. *Computational Complexity*, 12(3-4):85–130, 2003.

[GUV07] V. Guruswami, C. Umans, and S. P. Vadhan. Unbalanced expanders and randomness extractors from parvaresh-vardy codes. In *CCC*, pages 96–108, 2007.

[GW02] O. Goldreich and A. Wigderson. Derandomization that is rarely wrong from short advice that is typically good. In *APPROX-RANDOM*, pages 209–223, 2002.

[HNY17] P. Hubácek, M. Naor, and E. Yogev. The journey from NP to TFNP hardness. In *8th Innovations in Theoretical Computer Science Conference, ITCS*, volume 67, pages 60:1–60:21, 2017.

[ILL89] R. Impagliazzo, L. A. Levin, and M. Luby. Pseudo-random generation from one-way functions (extended abstracts). In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, pages 12–24, 1989.

[IW97] R. Impagliazzo and A. Wigderson. $P = BPP$ if E requires exponential circuits: Derandomizing the XOR lemma. In *STOC*, pages 220–229, 1997.

[KvM02] A. Klivans and D. van Melkebeek. Graph nonisomorphism has subexponential size proofs unless the polynomial-time hierarchy collapses. *SIAM J. Comput.*, 31(5):1501–1526, 2002.

[KvMS09] J. Kinne, D. van Melkebeek, and R. Shaltiel. Pseudorandom generators and typically-correct derandomization. In *APPROX-RANDOM*, pages 574–587, 2009.

[LZ19] F. Li and D. 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*, volume 145 of *LIPIcs*, pages 72:1–72:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[MV05] P. Bro Miltersen and N. V. Vinodchandran. Derandomizing arthur-merlin games using hitting sets. *Computational Complexity*, 14(3):256–279, 2005.

[Sha09] R. Shaltiel. Weak derandomization of weak algorithms: explicit versions of yao's lemma. In *CCC*, 2009.

[SS24] R. Shaltiel and J. Silbak. Explicit codes for poly-size circuits and functions that are hard to sample on low entropy distributions. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC*, pages 2028–2038, 2024.

[STV01] M. Sudan, L. Trevisan, and S. P. Vadhan. Pseudorandom generators without the xor lemma. *J. Comput. Syst. Sci.*, 62(2):236–266, 2001.

[SU05] R. Shaltiel and C. Umans. Simple extractors for all min-entropies and a new pseudorandom generator. *J. ACM*, 52(2):172–216, 2005.

[SU06] R. Shaltiel and C. Umans. Pseudorandomness for approximate counting and sampling. *Computational Complexity*, 15(4):298–341, 2006.

[SU09] R. Shaltiel and C. Umans. Low-end uniform hardness versus randomness tradeoffs for am. *SIAM J. Comput.*, 39(3):1006–1037, 2009.

[Sud97] M. Sudan. Decoding of Reed Solomon codes beyond the error-correction bound. *Journal of Complexity*, 13, 1997.

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