Extractors for Samplable Distribution with Polynomially Small Min-Entropy*

Ronen Shaltiel†

April 24, 2025

Abstract

Trevisan and Vadhan (FOCS 2000) introduced the notion of (seedless) extractors for samplable distributions. They showed that under a very strong complexity theoretic hardness assumption (specifically, that there exists a problem in $\mathsf{E} = \mathsf{DTIME}(2^{O(n)})$ that cannot be computed by size $2^{\Omega(n)}$ circuits that have an oracle to Σ_6^{P}) there are extractors for samplable distributions with large min-entropy of $k = (1 - \gamma) \cdot n$, for some small constant $\gamma > 0$. Recently, Ball, Shaltiel and Silbak (STOC 2025) were able to reduce the min-entropy threshold to $k = n^{1-\gamma}$. Ball et al., point out that their approach does not work for $k < \sqrt{n}$ (and this holds even for stronger hardness assumptions, in which 6 is replaced with any other constant).

In this paper, we show how to further reduce the min-entropy threshold to $k = n^{0.34} < \sqrt{n}$ under the same hardness assumption used by Trevisan and Vadhan. More generally, for every positive integer $i \geq 2$, and every $\alpha > \frac{1}{i}$, we construct an extractor for samplable distributions with min-entropy $k = n^\alpha$, under a hardness assumption in which 6 is replaced with $i + 3$ (the aforementioned result is a obtained for $i = 3$). We also provide a multiplicative version of our extractors (under a stronger hardness assumption) addressing an open problem of Ball et al.

Our work builds on the approach of Ball et al., who reduced the task of constructing extractors for samplable distributions with min-entropy k, to the task of constructing errorless condensers for samplable distributions with min-entropy k. Our main technical contribution is a new construction of errorless condensers for samplable distributions with $k = n^\alpha$ under the hardness assumption stated above, improving upon the min-entropy threshold achieved in Ball et al. (which cannot achieve $k < \sqrt{n}$).

Our insight is that the technique used by Ball et al. to reduce the task of constructing extractors to that of constructing errorless condensers, can *itself* be used to construct errorless condensers for polynomially small min-entropy when combined with "win-win analysis" approaches that are inspired by some early work on seeded extractors and dispersers. In order to do this, we adapt these approaches from the information theoretic scenario of seeded extractors and dispersers to the computational scenario of errorless condensers for samplable distributions.

Contents

1 Introduction

An influential paper by Trevisan and Vadhan [TV00] introduced the notion of (seedless) extractors for samplable distributions.

Definition 1.1 (Seedless extractor). *A function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *is a* (k, ϵ)-**extractor** *for a class* $\mathcal{D}$ *of distributions, if for every distribution* X *in* $\mathcal{D}$, *that is over* $\{0,1\}^n$, *such that* $H_\infty(X) \geq k$, $\mathsf{Ext}(X)$ *is* ϵ-*close to* U_m.[1]

The goal of Trevisan and Vadhan 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.

Definition 1.2 (Sampling procedures and samplable distributions). *For a function* $A : \{0,1\}^r \to \{0,1\}^n$, *we use* $Z \leftarrow A$ *to denote the experiment in which* $W \leftarrow U_r$, *and* $Z = A(W)$, *and say that* Z *is* **sampled** *by* A. *We say that the distribution* Z *is* **samplable** *by a class* $\mathcal{C}$ *of functions, if there exists* $A \in \mathcal{C}$ *that samples* Z.

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 n^d for a constant $d > c$. 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 exists 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 2.5 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 BPP = P. The stronger assumption that E is hard for exponential size nondeterministic circuits[2], originated in works on hardness versus randomness for AM, and is used in many results [AK02, KvM02, MV05, SU05, BOV07, GW02, GST03, SU06, SU09, Dru13, AASY15, BV17, AIKS16, HNY17, DMOZ22, BDL22, CT22, BGDM23, BSS24, SS24, Sha24]. 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).[3] The extractor of Trevisan and Vadhan [TV00] relies on the strong assumption that E is hard for exponential size Σ_6-circuits (which can be viewed as a scaled, nonuniform version of the widely believed assumption that $\mathsf{EXP} \neq \Sigma_6^\mathsf{P}$). 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, BSS25].

Previous work on extractors for samplable distributions. The main result of Trevisan and Vadhan [TV00] is that under a hardness assumption for Σ_6-circuits, there is an extractor for distributions samplable by poly-size circuits with $k = (1 - \gamma) \cdot n$, for some small constant $\gamma > 0$. Below is a precise statement.[4]

---

[1] See Section 2 for the standard definitions of min-entropy and statistical distance.

[2] A precise definition of nondeterministic circuits appears in Section 2.1.

[3] 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 2.1 for a formal definition.

[4] The statement of Theorem 1.3 given here is taken from the conference version [TV00]. In a later unpublished version, Trevisan and Vadhan notice that the assumption can be weakened to assume hardness for Σ_5-circuits.

Theorem 1.3 ([TV00]). *If E is hard for exponential size Σ_6-circuits then for every sufficiently small constant $\gamma > 0$, and every constant $c > 1$, there is a constant d such that 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 $((1-\gamma)\cdot n, \epsilon)$-extractor for distributions samplable by circuits of size n^c, where $\epsilon = n^{-c}$. Furthermore, Ext is computable in time n^d.*

Recent work on extractors for samplable distributions improved Theorem 1.3 in two respects:

- Ball, Goldin, Dachman-Soled and Mutreja [BGDM23] achieved the conclusion of Theorem 1.3 under the weaker, and more standard assumption that E is hard for exponential size nondeterministic circuits.

- Ball, Shaltiel and Silbak [BSS25] showed how to reduce the min-entropy threshold in Theorem 1.3 from $k = (1 - \gamma) \cdot n$ to $k = n^{1-\gamma}$, where in both results $\gamma > 0$ is some unspecified constant. This is achieved under a hardness assumption against Σ_5-circuits.

1.1 Our Results

In this paper we construct extractors for samplable distributions with polynomially small min-entropy threshold $k = n^\alpha$, where $\alpha > 0$ is an arbitrary constant. More specifically, for every positive integer constant $i \geq 2$, and for every constant $\alpha > \frac{1}{i}$, we obtain extractors for samplable distributions for min-entropy threshold $k \geq n^\alpha$ which extract almost all the randomness present in the source distribution, under the hardness assumption that E is hard for exponential size Σ_{i+3}-circuits. The precise result is stated below.

Theorem 1.4 (Extractor for samplable distributions polynomially small min-entropy). *For every positive integer constant $i \geq 2$, and every constant $\alpha > \frac{1}{i}$, if E is hard for exponential size Σ_{i+3}-circuits, then for every constants $c > 1$, and $\lambda > 0$, there exists a constant d, such that for every sufficiently large n, and every $k \geq n^\alpha$, there is a (k, ϵ)-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{(1-\lambda)\cdot k}$ for distributions samplable by circuits of size n^c, where $\epsilon = n^{-c}$. Furthermore, Ext can be computed in time n^d.*

Comparison to previous work on extractors for samplable distributions.

- For $i = 2$, the hardness assumption in Theorem 1.4 is against Σ_5-circuits. This is precisely the hardness assumption used by Ball, Shaltiel and Silbak [BSS25], and the statement of Theorem 1.4 is identical to their result, with the sole difference that the min-entropy threshold is improved from $k \geq n^{1-\gamma}$ for some unspecified constant $\gamma > 0$, to $k \geq n^{\frac{1}{2}+\gamma}$.

- For $i = 3$, the hardness assumption in Theorem 1.4 is against Σ_6-circuits. This is precisely the hardness assumption used by Trevisan and Vadhan [TV00] in Theorem 1.3, and we achieve min-entropy threshold $k \geq n^{\frac{1}{3}+\gamma}$, for any constant $\gamma > 0$. Note that already for $i = 3$, this min-entropy threshold is smaller than $\sqrt{n}$. As explained in detail in [BSS25], the technique of [BSS25] cannot give extractors for $k < \sqrt{n}$, and this holds even when assuming hardness against Σ_i-circuits, for an arbitrary constant i.

- Theorem 1.4 is incomparable to the extractor of Ball, Goldin, Dachman-Soled and Mutreja [BGDM23]. On the one hand the extractor of [BGDM23] only works for very large min-entropy (specifically, for $k = (1 - \gamma) \cdot n$ for some unspecified constant $\gamma > 0$) but on the other hand, the assumption used is weaker than what we assume, and only assumes hardness for nondeterministic circuits.

Samplable distributions with postselection. Ball, Goldin, Dachman-Soled and Mutreja [BGDM23] introduced a generalization of samplable distributions, and showed that their extractor applies for this more general class. More specifically, they considered distributions that are samplable with postselection.

Definition 1.5 (Samplable distributions with postselection). *For functions $A : \{0,1\}^r \to \{0,1\}^n$, and $P : \{0,1\}^r \to \{0,1\}$, we use $X \leftarrow A \mid P$ to denote the experiment in which $W \leftarrow U_r$, and $X = (A(W) \mid P(W) = 1)$, and say that X is **sampled** by A with postselection by P.*

2

Theorem 1.4 (as well as our other results below in Theorems 1.6 and Theorem 1.8) hold also when replacing "distributions samplable by circuits of size n^c" by "distributions samplable by circuits of size n^c with postselection by size n^c circuits". A precise definition is given in Definition 2.5.

Extractors with larger output length and higher error. As is the case in the construction of Ball, Shaltiel and Silbak [BSS25], our approach is to first construct an extractor that outputs $m = O(\log n)$ bits, and then use a transformation of Shaltiel [Sha08] to increase the output length. When using this transformation, one can also obtain extractors with $m = (1 - o(1)) \cdot k$, at the cost of having larger error (as done in [BSS25]). The precise result is stated below, and is identical to a corresponding result in [BSS25], except for the modifications in the min-entropy threshold, and the hardness assumption.

Theorem 1.6 (Extractor for samplable distributions with larger output length, and higher error). *For every positive integer constant $i \geq 2$, and every constant $\alpha > \frac{1}{i}$, if E is hard for exponential size Σ_{i+3}-circuits, then for every constants $c > 1$, every constant $0 < \eta < 1$, and every constant $b > 1$, there exists a constant d, such that for every sufficiently large n, and every $k \geq n^\alpha$, there is a (k, ϵ)-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{(1 - \frac{1}{\log^b n})k}$ for distributions samplable by circuits of size n^c, where $\epsilon = \frac{1}{2^{\log^\eta n}}$. Furthermore, Ext can be computed in time n^d.*

In Theorem 1.6, the improved output length comes with a cost of a larger error ϵ. We remark that our techniques can potentially achieve output length $m = (1 - o(1)) \cdot k$ with error $\epsilon = n^{-c}$ (which is the error achieved in Theorem 1.3 and Theorem 1.4) and the missing component is a seeded extractor that achieves $m = (1 - o(1)) \cdot k$ for $\epsilon = n^{-c}$ with seed length $O(\log n)$.

1.2 Multiplicative Extractors for Samplable Distributions

A signature application of seedless extractors is choosing keys for cryptographic protocols by extracting randomness from weak random sources. (Indeed, this was the original motivation of Trevisan and Vadhan [TV00] for introducing extractors for samplable distributions). More specifically, 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" $\eta > 0$.

If the key is chosen using the output of an extractor (which is only ϵ-close to uniform) then we are only guaranteed that the adversary's probability to cheat is smaller than $\eta + \epsilon$, which may be unacceptable if ϵ is "large" compared to η.

Applebaum, Artemenko, Shaltiel and Yang [AASY15] showed that "current techniques" cannot yield extractors for samplable distributions with error $\epsilon = n^{-\omega(1)}$, under a hardness assumption against Σ_i-circuits, and this holds for every constant i. (See [AASY15] for a precise formulation). This means that with current extractors for samplable distributions (which only achieve $\epsilon = n^{-c}$) we cannot expect cryptographic protocols to maintain negligible probability of cheating, when selecting the key with an extractor.

In order to address this problem, Applebaum et al. [AASY15], and later Shaltiel [Sha24], introduced a notion of "multiplicative extractors" which are designed to maintain negligible probability of cheating in cryptograhic protocols. Below, we use the definition of [Sha24] which is inspired by choices made in definitions in differential privacy [DMNS06].

Definition 1.7 (Multiplicative seedless extractor [Sha24]). *A function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ is a (k, ϵ)-**multiplicative extractor** for a class $\mathcal{D}$ of distributions, if for every distribution X in $\mathcal{D}$, that is over $\{0,1\}^n$, such that $H_\infty(X) \geq k$, and every event $A \subseteq \{0,1\}^m$*

$$\Pr[\mathsf{Ext}(X) \in A] \leq e^\epsilon \cdot \Pr[U_m \in A].$$

3

It is easy to see that every multiplicative extractor is also a (regular) extractor (see Proposition 2.8). However, unlike regular extractors, multiplicative extractors apply for the application of selecting keys for cryptographic protocols, even if $\epsilon > 0$ is a small constant. This is because such extractors give that the adversary's probability to cheat is at most $e^\epsilon \cdot \eta \le (1 + O(\epsilon)) \cdot \eta$.[5]

Shaltiel [Sha24] constructed a (k, ϵ)-multiplicative extractor for samplable distributions, with $k \ge (1 - \gamma) \cdot n$ for some constant $\gamma > 0$, $\epsilon = n^{-c}$, and $m = \Omega(k)$, under a hardness assumption for nondeterministic circuits.

The aforementioned extractor achieved by Ball, Shaltiel and Silbak [BSS25] is not multiplicative. As explained in [BSS25] (and as we also explain in Section 1.3.6) this is because current constructions of 2-source extractors for low min-entropy [CZ16, Li16] do not achieve low error (this is a well known open problem).

In this paper, we are able to circumvent this difficulty, and construct multiplicative extractors for samplable distributions for polynomially small min-entropy threshold of $k = n^\alpha$, for every constant $\alpha > 0$. As is the case in our Theorem 1.4, our result gives a tradeoff between α and the hardness assumption used. At this point, we have not tried to optimize this tradeoff, and only show that for every constant $\alpha > 0$, there is a constant $j \ge 1$ such that a hardness assumption is against Σ_j-circuits suffices. A precise statement is given below. (A more general formulation appears in Theorem 4.8).

Theorem 1.8 (multiplicative extractor for samplable distributions)**.** *For every constants $\alpha > 0$, there exist constants $j \ge 1$, and $\beta > 0$ such that if E is hard for exponential size Σ_j-circuits, then for every constant $c > 1$, there exists a constant d, such that for every sufficiently large n, there is a function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{n^\beta}$ that is an $(n^\alpha, \frac{1}{n^c})$-multiplicative extractor for distributions samplable by circuits of size n^c. Furthermore, Ext can be computed in time $\mathsf{poly}(n^d)$.*

A closer inspection of the argument reveals that $j = c_0 + \lceil \frac{1}{\alpha} \rceil$ for some universal constant c_0. However, bounding the constant c_0 will require computing unspecified constants in some components that we use.

Perspective. Trevisan and Vadhan made the philosophical argument that every weak source of randomness from nature is *necessarily* efficiently samplable. If one agrees with this argument, then extractors for samplable distributions capture all natural weak sources of randomness that are available to computers. Our extractors extend the usefulness of extractors for samplable distributions to sources with polynomially small min-entropy.

1.3   Technique

In this section, we give a detailed informal overview of the main ideas that we use. The later technical sections contain full definitions, statements and proofs and do not build on the informal explanation of this section. The readers can skip to the technical section if they wish.

1.3.1   Errorless Condensers are Sufficient for Constructing Extractors

Ball, Shaltiel and Silbak [BSS25] extended the techniques of Trevisan and Vadhan [TV00] and showed that the task of constructing extractors for samplable distributions can be reduced to that of constructing errorless condensers for samplable distributions. We start with a definition of errorless condensers.

---

[5]This argument applies to "unpredictability security games" where one bounds the probability that the adversary can cheat, but not necessarily to "indistinguishability security games" where one bounds the probability that the adversary can distinguish between two distributions. See e.g. [DY13] for a discussion.

Definition 1.9 (Errorless condenser). *A function* $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n_1}$ *is a* (k, k_{out})**-errorless condenser** *for a class $\mathcal{D}$ of distributions, if for every distribution X in $\mathcal{D}$, that is over $\{0,1\}^n$, such that $H_\infty(X) \geq k$, $H_\infty(\mathsf{Cnd}(X)) \geq k_{\mathsf{out}}$.*

More specifically, following the work of [BSS25] (which we will explain in detail below) for every $\alpha > 0$, in order to construct extractors for samplable distributions with min-entropy threshold $k = n^\alpha$, it is sufficient to construct a $(k = n^\alpha, k_{\mathsf{out}} = n^\delta)$-errorless condenser $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{k/10}$ for samplable distributions, for some constant $\delta > 0$ that may depend on α. We remark that in this reduction, it is not necessary for the output distribution $\mathsf{Cnd}(X)$ to be "more condensed" than X, and it is allowed that δ is tiny compared to α. What is important is that the output length is smaller than k.

Having established this reduction, Ball, Shaltiel and Silbak [BSS25] proceed to show how to construct an errorless condenser for samplable distributions, under a hardness assumption. Their condenser construction achieves min-entropy threshold $k = n^{1-\gamma}$ for some $\gamma > 0$, and this threshold is inherited by their extractor construction. As explained in [BSS25], the errorless condenser construction of Ball, Shaltiel and Silbak [BSS25] (which relies on "hard to sample functions" [SS24] and "seeded dispersers with short seed and large error" [Zuc07]) cannot work for $k < \sqrt{n}$ (regardless of the hardness assumption that is used).

1.3.2 A New Construction of Errorless Condensers

In this paper, we give a new (and different) construction of errorless condensers that achieves a lower min-entropy threshold, and is used to derive our extractors using the aforementioned reduction of Ball, Shaltiel and Silbak [BSS25]. Our condenser can achieve $k \approx n^{1/i}$, under hardness for Σ_{i+3}-circuits, and this min-entropy threshold and hardness assumption is inherited by our extractors, as stated in Theorem 1.4 and Theorem 1.6. Our main result on errorless condenser is stated below (a more general statement appears in Theorem 3.1).

Theorem 1.10. *For every positive integer constant $i \geq 2$, and every constant $\alpha > \frac{1}{i}$, there exist constants $0 < \delta_1 < \delta_2 \leq \frac{\alpha}{10}$, such that if E is hard for exponential size Σ_{i+3}-circuits, then for every constants $c > 1$, there exists a constant d, such that for every sufficiently large n, and every $k \geq n^\alpha$, there is a (k, n^{δ_2})-errorless condenser $\mathsf{FCnd} : \{0,1\}^n \to \{0,1\}^{n^{\delta_1}}$. Furthermore, Ext can be computed in time n^d.*

The condenser above is termed FCnd for "final condenser", and it will be constructed in two steps: We will first construct a "basic condenser" BCnd with large output length of roughly $n^{1-\alpha}$ bits, and then use BCnd to construct FCnd. We will explain both these constructions in detail in Sections 1.3.4 and 1.3.5.

The important thing to notice is that the min-entropy threshold in Theorem 1.10 is $k = n^\alpha$ for an arbitrarily small $\alpha > 0$, and that the output length is indeed smaller than $\frac{k}{10}$, meaning that FCnd is suitable for the reduction, and implies extractors.

We will now focus on explaining our new construction of errorless condensers. At a high level, our insight to try and use the argument that constructs extractors for samplable distributions from errorless condensers for samplable distributions, in order to *directly* construct errorless condensers. We will start by explaining the reduction of [BSS25], which relies on, and extends the argument of Trevisan and Vadhan [TV00].

1.3.3 A Review of the Technique of [TV00, BSS25]

We will now explain how Ball, Shaltiel and Silbak [BSS25] used errorless condensers to construct extractors. We will later extend these ideas in order to directly construct errorless condensers. The construction of [BSS25] uses 2-source extractors.

Definition 1.11 (Two-source extractor). *A function* $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *is a (k_1, k_2, ϵ)-2-source extractor if for every two independent distributions X_1, X_2 with $H_\infty(X_1) \geq k_1$ and $H_\infty(X_2) \geq k_2$, $\mathsf{TExt}(X_1, X_2)$ is ϵ-close to U_m.*

The construction of [BSS25] (that we now describe) produces an extractor which only outputs $m = O(\log n)$ output bits. The output length can later be increased using a result of Shaltiel [Sha08]. We will only describe the first part.

For every $\alpha > 0$, in order to construct an extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{m=O(\log n)}$ for samplable distributions X with min-entropy threshold $k = n^\alpha$, and error $\epsilon > 0$ one requires:

- An explicit $(k = n^\alpha, k_{\mathsf{out}} = n^\delta)$-errorless condenser $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{\frac{k}{10}}$. (Here it is sufficient that $k_{\mathsf{out}} = n^\delta$ for some (possibly very small) constant $\delta > 0$ that depends on α.

- A function $g : \{0,1\}^n \to \{0,1\}^n$, such that for $Y = \mathsf{Cnd}(X)$, and every Σ_2-circuit C of size slightly larger than the circuit size of Cnd, we have that: $\Pr[C(Y) = g(Y)] \leq 2^{-\Omega(k_{\mathsf{out}})}$.

  Ball, Shaltiel and Silbak [BSS25] showed how to construct such a function under a hardness assumption for Σ_4-circuits. See precise definition and statement in Section 2.6.[6]

- A (k', k', ϵ')-2-source extractor $\mathsf{TExt} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ set up for min-entropy threshold k' that is slightly smaller than k_{out} (say $k' = \sqrt{k_{\mathsf{out}}}$) and $\epsilon' = O(\epsilon/2^m)$. Explicit constructions of such 2-source extractors were given by Chattopadhyay and Zuckerman [CZ16], and Li [Li16] for $m = O(\log n)$ (see Section 2.4.4 for precise statements).

Given these components, the extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ is defined by:[7]

$$\mathsf{Ext}(X) = \mathsf{TExt}(g(\mathsf{Cnd}(X)), X).$$

As we already have explicit constructions of the function g, and the 2-source extractor TExt, this is indeed a reduction, showing that errorless condensers suffice for extractors. We now survey the argument used to prove the correctness of this construction.

Recall that we are assuming that $H_\infty(X) \geq k$, whereas the output length of Cnd is $k/10$. This intuitively means that X has min-entropy even conditioned on $\mathsf{Cnd}(X)$. More precisely, it can be argued that $(\mathsf{Cnd}(X), X)$ is (close to) a *block-wise source*.

Definition 1.12 (block-wise sources). *A distribution (W_1, W_2) is a (k_1, k_2)-block-wise source if $H_\infty(W_1) \geq k_1$, and for every $w_1 \in \mathrm{Supp}(W_1)$, $H_\infty(W_2 \mid W_1 = w_1) \geq k_2$.*

Indeed, we can argue that $(\mathsf{Cnd}(X), X)$ is (close to) an (n^δ, n^δ)-block-wise source. The correctness of the extractor follows by showing that if the output distribution $\mathsf{Ext}(X)$ is not close to uniform, then one can use the sampling circuit A of the samplable distribution X, to construct a Σ_2-circuit C (of slightly larger size) that breaks the security guarantee of the function g.

More specifically, if $\mathsf{Ext}(X)$ is not ϵ-close to uniform, then there exists an element $z \in \{0,1\}^m$, such that $\Pr[\mathsf{Ext}(X) = z] > (1 + \epsilon) \cdot 2^{-m}$. By an averaging argument, it can be shown that with probability at least say $\epsilon \cdot 2^{-m}/4$ over choosing $y \leftarrow Y = \mathsf{Cnd}(X)$, we have that y is "useful", meaning that:

$$\Pr[\mathsf{TExt}(g(y), X) = z \mid Y = y] = \Pr[\mathsf{Ext}(X) = z \mid Y = y] > (1 + \frac{\epsilon}{2}) \cdot 2^{-m} = 2^{-m} + \frac{\epsilon \cdot 2^{-m}}{2}.$$

---

[6]In fact, the construction of [BSS25] gives that the function g (which they termed HOS for "Hard on Samplable distributions") is hard on average, not just on $Y = \mathsf{Cnd}(X)$, but on any distribution Y with $H_\infty(Y) \geq k_{\mathsf{out}}$ that is samplable by circuits of the same size as Cnd (and $Y = \mathsf{Cnd}(X)$ is such a distribution as one can efficiently sample X, and then apply Cnd).

[7]We remark that in the HOS construction of [BSS25], the output length of g is n^{a_0} for some universal constant $a_0 > 1$ rather than n. See precise formulation in Theorem 2.22. This means that TExt should work for source length n^{a_0} rather than n, and that some inputs need to be padded with zeros. We ignore these technicalities in this high level overview, and the precise details appear in Section 2.6.

We will set the error parameter ϵ' of the 2-source extractor to be $\epsilon' = \frac{\epsilon \cdot 2^{-m}}{2}$.[8] This gives that for a useful $y \in \{0,1\}^n$, $g(y)$ is an element in the set

$$T_y = \left\{ v : \Pr[\mathsf{TExt}(v, (X \mid Y = y)) = z] > 2^{-m} + \epsilon' \right\}.$$

We have that $(X \mid \mathsf{FCnd}(X) = y)$ meets the entropy threshold k' of the 2-source extractor. This can be used to conclude that for a useful y, $|T_y| \leq 2^{k'}$. (This follows as otherwise the distribution V_y that is uniform over T_y, and the distribution $W_y = (X \mid \mathsf{FCnd}(X) = y)$ are independent distributions on which $\Pr[\mathsf{Ext}(V_y, W_y) = z] > 2^{-m} + \epsilon'$, violating the guarantee of 2-source extractors, as $\Pr[U_m = z] = 2^{-m}$).

Trevisan and Vadhan [TV00] showed that this observation can be used to construct a small Σ_2-circuit C which given $Y = \mathsf{Cnd}(X)$ computes $g(Y)$ with not too small probability, violating the guarantee of g.[9]

Summing up, we see that in the specified construction, the requirement that the output length of Cnd is smaller than k, can be removed if we can guarantee that the distribution $(\mathsf{Cnd}(X), X)$ is an (n^δ, n^δ)-block-wise source. A formal statement of this observation is given in Lemma 3.2 in Section 3.

1.3.4 The Basic Condenser

We will now explain how to use the approach presented above in order to construct a "basic condenser" BCnd which will be designed for min-entropy threshold $k = n^\alpha$ (which is what we want) but will output strings of length $\approx n^{1-\alpha}$ (and recall that for the reduction we need output length $k/10 < n^\alpha$). This condenser is stated below, a stronger, and more general formulation appears in Theorem 3.9.

Theorem 1.13. *If E is hard for exponential size Σ_5-circuits, then for every constants $c \geq 1$, $0 < \alpha < 1$, and $0 < \gamma < \alpha/4$, there exist constants $\delta > 0$ and $d \geq 1$ such that for every sufficiently large n, there is a function $\mathsf{BCnd} : \{0,1\}^n \to \{0,1\}^{O(n^{1-\alpha+\gamma})}$ that is an (n^α, n^δ)-errorless condenser for distributions samplable by size n^c circuits. Furthermore, BCnd can be computed in time n^d.*

The output length of BCnd is roughly $n^{1-\alpha}$, and this is shorter than $k = n^\alpha$ if $\alpha > \frac{1}{2}$. This means that BCnd is a suitable errorless condenser for the reduction that converts errorless condenser into extractors, if $k \geq n^\alpha$ for $\alpha > \frac{1}{2}$ (and this already gives the result mentioned for this min-entropy threshold in Section 1.1).

While this is suitable for $k > \sqrt{n}$, this is not suitable for $k < \sqrt{n}$, an indeed, later on, in Section 1.3.5, we will explain how to prove Theorem 1.10, and construct FCnd using BCnd.

We now focus on explaining how to construct the basic condenser BCnd (using the ideas outlined in Section 1.3.3). Our approach is also inspired by some of the first constructions of seeded extractors and dispersers due to Saks, Srinivasan and Zuckerman [SSZ98] and Ta-Shma [Ta-96].

In order to outline the idea, it will be useful to cheat and assume an unjustified simplifying assumption:

Simplifying assumption: For every function B, if $H_\infty(X) \geq k$, and $t_1 \leq H_\infty(B(X)) \leq t_2$, then $(B(X), X)$ is a $(t_1, k - t_2)$-block-wise source.

This unjustified assumption would have held if min-entropy had a chain rule. In that case, it would indeed follow that if $H_\infty(X) \leq t_2$, then the remaining $k - t_2$ bits of entropy must remain in X, conditioned on $B(X)$.[10] (We will later explain how to remove this unjustified assumption).

---

[8]Current constructions of 2-source extractors for low min-entropy [CZ16, Li16] can only achieve $\epsilon' = n^{-O(1)}$, and this is why we need to set $m = O(\log n)$, so that $\epsilon' = \frac{\epsilon \cdot 2^{-m}}{2}$ is $n^{-O(1)}$, and recall that we are shooting for $\epsilon = n^{-c}$.

[9]For completeness, we survey this argument. Trevisan and Vadhan [TV00] first show that there is a Σ_1-circuit that given y, v, is able to check whether $v \in T_y$. This is done by approximating $\Pr[\mathsf{TExt}(v, X) = z \mid \mathsf{FCnd}(X) = y] = \frac{\Pr[\mathsf{TExt}(v,X)=z \wedge \mathsf{Cnd}(X)=y]}{\Pr[\mathsf{Cnd}(X)=y]}$ (which can be done by using classical results on "approximate counting of NP witnesses" [Sto83, Sip83, JVV86]). They then use classical results on "uniform sampling of NP witnesses" [JVV86, BGP00] to construct a Σ_2-circuit C which for every good y, is able to sample a uniform output from T_y. Overall, this gives a small Σ_2-circuit C that computes $g(y)$ with probability $2^{-k'}$ on any useful y. This means that for $Y = \mathsf{Cnd}(X)$, $\Pr[C(Y) = g(Y)] > 2^{-\Omega(k_{\text{out}})}$, violating the security guarantee of g.

[10]This does not hold as it could be the case for example that $B(X)$ outputs the first $n/2$ bits of X, and X is a convex combination

The construction of BCnd **(under an unjustified simplifying assumption).** Recall that we aim to construct an errorless condenser for $k = n^\alpha$. Following a seeded disperser construction of Saks, Srinivasan and Zhou [SSZ98], we will divide the n bit long input source X into $\ell = 10 \cdot n^{1-\alpha}$ blocks of length $b = n/\ell = n^\alpha/10$. Let $B_i(X) = X[1, \ldots, i]$ denote the first i blocks of X and let $k_i = H_\infty(B_i(X))$. Let i^* be the first index such that $k_{i^*} \geq k/2$, and note that such an i^* exists. Furthermore, because the length of blocks is less than $k/10$, we can hope to get that $H(B_{i^*}(X)) \leq k/2 + 2k/10 \leq 3k/4$, and therefore (by the unjustified assumption) that $(B_{i^*}(X), X)$ is a $(k/4, k/4)$-block-wise source.

This means that if we apply the construction of Section 1.3.3 using the function B_{i^*} instead of Cnd (namely set $\mathsf{Ext}(X) = \mathsf{TExt}(g(B_{i^*}(X)), X)$) then as we explained in Section 1.3.3, the analysis given there shows that $\mathsf{Ext}(X)$ is close to uniform.

We do not know how to find i^*, and cannot get an extractor for samplable distributions. Instead, we will try all possible $i \in [\ell]$, which is good enough when constructing a condenser. More precisely, we define:

$$\mathsf{BCnd}(X) = \mathsf{TExt}(g(B_1(X)), X), \ldots, \mathsf{TExt}(g(B_\ell(X)), X),$$

and can indeed conclude that as one of the substrings in $\mathsf{BCnd}(X)$ is (close to) uniform, the distribution $\mathsf{BCnd}(X)$ is (close to) having high min-entropy.

Note that even with the unjustified assumption, this is not what we wanted, as the condenser BCnd has error (inherited from the error of TExt) and is not errorless. Moreover, as explained in Section 1.3.3, the analysis using 2-source extractors, can at best produce individual outputs of length $m = O(\log n)$ whereas we want a much larger m, as we want the condenser to output significantly more bits of min-entropy.

We address both problems by noting that when constructing condensers for samplable distributions, we can replace the 2-source extractor TExt, by a 2-source condenser TCnd. Fortunately, there are explicit constructions of 2-source condensers due to Ben-Aroya et al. [BCDT19] which for min-entropy threshold $k = n^\alpha$ are able to output $m = n^\gamma$ bits, with min-entropy $k_{\mathsf{out}} = n^{\Omega(\gamma)}$, for some constant $\gamma > 0$. Moreover, we observe that the error parameter of these 2-source condensers is sufficiently low, so that by decreasing the constant hidden in the $\Omega(\cdot)$ notation above, we can "swallow the error" and obtain that these 2-source condensers are errorless. (A precise statement of the result of Ben-Aroya et al. [BCDT19] and its interpretation as an errorless 2-source errorless condenser is stated in Section 2.4.5). This gives that by replacing the 2-source extractor TExt with the 2-source errorless condenser TCnd, and taking:

$$\mathsf{BCnd}(X) = \mathsf{TCnd}(g(B_1(X)), X), \ldots, \mathsf{TCnd}(g(B_\ell(X)), X),$$

we obtain an errorless condenser for samplable distributions with output length $\ell \cdot m = O(n^{1-\alpha+\gamma})$ (as required). We also have that $\mathsf{BCnd}(X)$ has min-entropy at least $n^{\Omega(\gamma)}$, which is good enough for our purposes.

Removing the unjustified assumption. We would like to remove the unjustified assumption. This is often done by proving a structural result on distributions X with $H_\infty(X) \geq k$, showing that X can be written as a convex combination of distributions $\{X^i\}$ where each for X^i, the index i satisfies that $(B_i(X^i), X^i)$ is a block-wise source. This approach was used by Ta-Shma [Ta-96] and is sufficient assuming the analysis can be done separately for each component X^i.

In our setting, this is more difficult, as we will need to decompose X into a convex combination of distributions $\{X^i\}$ where in addition to the aforementioned condition, we will need that each X^i is *efficiently samplable*.

Following Ta-Shma [Ta-96] we would like to design a "selector function" $S : \{0,1\}^n \to [\ell]$, that assigns each $x \in \mathrm{Supp}(X)$ to an $i \in [\ell]$, such that for every $i \in [\ell]$:

$X = \frac{1}{2}X_1 + \frac{1}{2}X_2$, where X^1 is a distribution in which $H_\infty(B_1(X^1)) = k$, and the last $n/2$ bits of X^1 are fixed, and X^2 is a distribution where $H_\infty(B(X^2)) = k - t_1$, and the last $n/2$ bits of X^2 are independent of $B(X^2)$ and have min-entropy $k - t_1$.

- the distribution $X^i = (X \mid S(X) = i)$ is efficiently samplable.
- $(B(X^i), X^i)$ is a block-wise source.

This would be sufficient for our purposes, as the aforementioned analysis can now be applied on each X^i separately.

Following Ta-Shma [Ta-96], we show that a suitable selector function S that satisfies the second condition exists. Furthermore, we show that S can be computed by a Σ_1-circuit of size slightly larger than that of the circuit A that samples X. This is not sufficient to achieve the first item as stated, as this does not imply that $X^i = (X \mid S(X) = i)$ is efficiently samplable by deterministic circuits. However, it does follow (directly from the definition) that X^i is samplable by deterministic circuits with postselection by the Σ_1-circuit S_i.

This turns out to be sufficient for implementing the analysis, with the cost of "pushing the hardness assumption one level up the polynomial time hierarchy" and assuming hardness for Σ_5-circuits, rather than Σ_4-circuits (so that the function g will be hard not only on samplable distributions, but also on samplable distributions with postselection by Σ_1-circuits). Indeed, this is why the assumption in Theorem 1.13 is stated with Σ_5-circuits, rather than Σ_4-circuits.

We implement the selector function using ideas from [SSZ98, Ta-96] together with classical results on "approximate counting of NP witnesses" [Sto83, Sip83, JVV86]. These classical results (stated precisely in Section 2.7) imply that for a samplable distribution X, and for every $i \in [\ell]$, a small Σ_1-circuit can approximate $k_i(x) = -\log \Pr[B_i(X) = B_i(x)]$. Loosely speaking, this can be viewed as the amount of "entropy" in the (fixed) string $B_i(x)$, and the high level idea is that the selector function $S(x)$ will output the smallest i, such that $k_i(x) \geq k/2$. (We are hiding many details here, and the reader is referred to Section 3.2 for the precise argument).

1.3.5 The Final Condenser

We will now continue to implement our plan and show how the "final condenser" FCnd of Theorem 1.10 is constructed from the "basic condenser" BCnd of Theorem 1.13.

Our goal is to prove Theorem 1.10. We will first focus on the case where $i = 3$, meaning that we are shooting to construct an errorless FCnd that for $\alpha > 1/3$, and min-entropy threshold $k \geq n^\alpha$, outputs a distribution over significantly less than $n^{1/3}$ bits with min-entropy $n^{\Omega(1)}$. For this choice of α slightly larger than $1/3$, BCnd outputs roughly $n^{1-\alpha} \approx n^{2/3}$ bits, where the output min-entropy is n^δ for some very small constant $\delta > 0$. Our goal is to make the output length shorter, and in particular significantly shorter than $n^{1/3}$. (Recall that such a short output length is required to transform the condenser into an extractor, as explained in Section 1.3.1).

One natural approach to improve a given condenser is composition (or "repeated condensing"). However, in our scenario, as δ may be very small, the "entropy rate" of the output distribution of BCnd may be inferior to that of the input distribution, and we cannot hope to make progress by straightforward composition.

Instead, we use a "win-win approach" that is inspired by a seeded extractor constriction of Reingold, Shaltiel and Wigderson [RSW06]. Once again, we will start by assuming the unjustified assumption of the previous section regarding block-wise sources. Given a samplable distribution X with $H_\infty(X) \geq k$, we are guaranteed that $H_\infty(\mathsf{BCnd}(X)) \geq n^\delta$. We will consider two scenarios:

- If $H_\infty(\mathsf{BCnd}(X)) \leq k - n^\delta$, then (by the unjustified assumption, taking $B = \mathsf{BCnd}$) we have that $(\mathsf{BCnd}(X), X)$ is an (n^δ, n^δ)-block-wise source. In this case, if we use BCnd as the function B from Section 1.3.4 (and replace TExt with TCnd as we did in the previous section) we can obtain that $\mathsf{TCnd}(g(\mathsf{BCnd}(X)), X)$ is significantly shorter than $n^{1/3}$ and has $n^{\Omega(1)}$ bits of min-entropy.
- If $H_\infty(\mathsf{BCnd}(X)) > k - n^\delta$ then we have made progress as $H_\infty(\mathsf{BCnd}(X)) > n^{1/2}$, and now we can hope to apply another instantiation BCnd_2 of the basic condenser (set up for min-entropy threshold n^{α_2}

9

where $\alpha_2 > \frac{1}{2}$) and the distribution $\mathsf{BCnd}_2(\mathsf{BCnd}(X))$ has length significantly shorter than n^α, while having $n^{\Omega(1)}$ bits of min-entropy.

Overall, this leads to the following construction:

$$\mathsf{FCnd}(X) = \mathsf{TCnd}(g(\mathsf{BCnd}(X)), X), \mathsf{BCnd}_2(\mathsf{BCnd}(X)),$$

and by the case analysis above, there exists a substring of $\mathsf{FCnd}(X)$ that has min-entropy $n^{\Omega(1)}$, implying that FCnd is an errorless condenser for min-entropy threshold k that is slightly larger than $n^{1/3}$, and output length significantly less than $n^{1/3}$ bits, with the guarantee that $\mathsf{FCnd}(X)$ outputs $n^{\Omega(1)}$ bits of min-entropy (as guaranteed for $i = 3$ in Theorem 1.10).

There are technical difficulties in implementing this approach. More specifically, the two different instantiations of BCnd need to be set up for distributions that are samplable by two different circuit sizes. This is because, BCnd_2 should be set up to work on the distribution $\mathsf{BCnd}(X)$, which is efficiently samplable in roughly the time it takes to compute BCnd (which in turn is larger than the time it takes to sample X). Nevertheless, this complexity leveraging can be done (and a complexity leveraging argument with a similar flavor already appears in [BSS25]).

For $i > 3$, we can apply the approach above recursively. At each step either we obtain a block-wise source and can output a string with high min-entropy, or we obtain a more condensed distribution (and our task becomes easier). Repeating this roughly i times, we eventually output a string with high min-entropy.

There is however significant difficulty in making this argument work without the simplifying assumption. This is because (as we previously did when removing the simplifying assumption in the construction of BCnd) at each step, we will need to design a selector circuit, which "divides" the inputs x in the support of X between the two scenarios.

The precise argument (which we will not describe here) is quite technical, and appears in Section 3.3. Additional complication is that (at least the way we are able to do it) each time we define a selector function for a recursive step, we use "approximate-counting of NP witnesses" to show that the new selector circuit is computed by a Σ_1-circuit that uses the selector circuit of the previous iteration as oracle. This means that we "pick up more levels of the polynomial time hierarchy" in each recursive step. To account for that, we need to strengthen the hardness assumption for each recursive step, and this is why Theorem 1.10 (and consequently Theorems 1.4 and Theorem 1.6) are stated under a hardness assumption for Σ_j circuits where $j = i + 3$ is a constant that increases with i. We do not know whether this loss is necessary, and it may be possible to analyze the suggested construction without this loss. (For example, no such loss is incurred under the unjustified assumption).

1.3.6 Constructing a Multiplicative Extractor

In this section we explain how to prove Theorem 1.8 and obtain a multiplicative extractor for samplable distributions. We would like to extend the argument of Ball et al. [BSS25] that we described in Section 1.3.3 and show how to use our errorless condenser FCnd from Theorem 1.10 to get an extractor that is *multiplicative*.

The difficulty in obtaining a multiplicative extractor.   It should be noted that (as can be seen from our description in Section 1.3.3) the approach of [BSS25] does indeed give a multiplicative extractor $\mathsf{Ext}(x) = \mathsf{TExt}(g(\mathsf{FCnd}(x)), x)$ that outputs m bits, if the 2-source extractor TExt has error ϵ' that is sufficiently smaller than 2^{-m}. This follows because we start from the assumption that there exists $z \in \{0, 1\}^m$ such that $\Pr[\mathsf{Ext}(X) = z] > (1 + \epsilon) \cdot 2^{-m}$ (namely that Ext is not a multiplicative extractor) and obtain a contradiction by showing that TExt is not an extractor with error $\epsilon' \approx \epsilon \cdot 2^{-m}$.

Unfortunately, the best known current constructions of 2-source extractors [CZ16, Li16] (as well as subsequent work) have running time that is $\mathrm{poly}(1/\epsilon')$ for error ϵ'. This means that as we want these 2-source

extractors to run in time poly(n), we must have that $\epsilon' = n^{-O(1)}$, and as the argument requires that $\epsilon' < 2^{-m}$, one can at best set $m = O(\log n)$ and obtain an extractor for samplable distributions with output length $m = O(\log n)$. (For this output length, and $\epsilon = n^{-c}$, the distinction between multiplicative extractors and regular extractors is immaterial).

In order to obtain a large output length (as stated in final results) Ball, Shaltiel and Silbak [BSS25] use a result by Shaltiel [Sha08] that shows that one can use the output $\mathsf{Ext}(X)$, as a seed to a seeded extractor, and extract almost all the randomness from X. (Note that as X and $\mathsf{Ext}(X)$ are correlated, this is far from obvious). The crucial point is that using this transformation, one loses multiplicativity, and the final extractor is not multiplicative.

Modifying the extractor construction to get a multiplicative extractor. We would like to modify the technique of Ball, Shaltiel and Silbak [BSS25], so that given our final condenser FCnd, we can *directly* transform it into an extractor for samplable distributions with *large output length* that *is multiplicative* (without relying on the composition of [Sha08]).

We would therefore like to replace the 2-source extractor TExt used in Section 1.3.3 with a different extractor that *does* has large output length m and error $\epsilon' < 2^{-m}$, in a way that will still enable us to perform the argument of [BSS25, TV00].

Fortunately, Li [Li15] gave a construction of an extractor IExt which for $k' = n^{\Omega(1)}$ has $m = (k')^{\Omega(1)}$ and exponentially small error $\epsilon' < 2^{-m}$. This extractor is designed for two independent sources, with the caveat that the second source needs to be a (k', k')-block-wise source. (See Theorem 2.15 for a precise formulation).

We will adapt the approach of [BSS25] to work with this extractor. More specifically, for min-entropy threshold $k = n^{\alpha}$, we define:

$$\mathsf{Ext}(X) = \mathsf{IExt}(g(\mathsf{FCnd}_2(\mathsf{FCnd}_1(X))), \mathsf{FCnd}_1(X), X).$$

Here, FCnd_1 will be an instantiation of Theorem 1.10 set up to output n^{α_1} bits of min-entropy from X (for some $\alpha_1 > 0$), and FCnd_2 will be yet another instantiation of Theorem 1.10 set up to output n^{α_2} bits of min-entropy from $\mathsf{FCnd}_1(X)$ (for some $\alpha_2 > 0$). Note that we can indeed obtain the errorless condenser FCnd_2, as $\mathsf{FCnd}_1(X)$ is guaranteed to have min-entropy n^{α_1} and is efficiently samplable by the composition of the circuit A that samples X and FCnd_1.

The idea is that when analyzing the construction of $\mathsf{Ext}(X) = \mathsf{TExt}(g(\mathsf{Cnd}(X)), X)$ in Section 1.3.3, we first fixed the first input of TExt by fixing $Y = \mathsf{Cnd}(X)$ to some fixed useful y, and then considered the distribution of the second input conditioned on the first input being y, and it turned out to be sufficient for the argument that the second source $(X \mid Y = y)$ meets the requirement that TExt makes from its second source.

Similarly, in the analysis of the modified construction (that appears in Section 4.3) we will fix the first input by fixing $Y = \mathsf{FCnd}_2(\mathsf{FCnd}_1(X))$ to some fixed useful y, and then consider the distribution of the second input conditioned on the first input being y, namely $((\mathsf{FCnd}_1(X), X) \mid Y = y)$. We will be able to show that this distribution is a block-wise source, and so meets the requirement that IExt makes from its second source. Loosely speaking, this intuition will allow us to prove the correctness of the construction. We will not go into additional details here, and the construction and proof appear in Section 4.3.

This enables us to use IExt instead of TExt, and as we explained earlier, the fact that the error of IExt is exponentially smaller, will translate to give that Ext which outputs $m = n^{\Omega(1)}$ bits, and is a multiplicative extractor, proving Theorem 1.8.

We stress that for this argument, it is crucial that our errorless condenser FCnd of Theorem 1.10 works for any polynomial small min-entropy. This is because even if the initial threshold α is relatively large, we end up needing to apply FCnd_2 for threshold α_1 that may be very small. (Consequently, this approach would not have worked with the errorless condenser of [BSS25]). Note also that having to set up FCnd_2 for low (and in fact unspecified) threshold $\alpha_1 > 0$, means that we need to set up the hardness assumption for this task, and

this is why we need a stronger hardness assumption in Theorem 1.8 than that in Theorem 1.4. One obvious way to reduce this hardness assumption is to argue that one can perform the argument with α_1 that is not much smaller than α. This may be doable by figuring out the precise constants in previous work that we use.

Organization of the paper

In Section 2 we give some preliminaries, as well as definitions, and past results that we use. In Section 3 we present our new constructions of errorless condensers. In Section 4 we prove our main theorems and construct extractors for samplable distributions using the errorless condenser of Section 3. Finally, in Section 5 we give some open problems.

2 Preliminaries

In this section, we present notation, definitions, and past work that we use. For completeness, we will also repeat definitions from the introduction.

2.1 Definition of Circuits of Various Types

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

Definition 2.1 (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 exists an assignment to the nondeterministic input wires that makes C output 1 on x.*

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 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.[11]

2.2 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 $(X \mid A)$ 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.

Two distributions X, Y over the same finite domain S are ϵ-close if for every $A \subseteq S$, $|\Pr[X \in A] - \Pr[Y \in A] \leq \epsilon$.

The *min-entropy* of a distribution X over a finite set S, is defined by $H_\infty(X) := \min_x \log \frac{1}{\Pr[X=x]}$, where the minimum is taken over all strings x in the support of X.

We repeat the standard definition of block-wise sources, that appeared in Section 1 as Definition 1.12.

---

[11]An alternative approach to defining 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 $\mathsf{DTIME}^{\Sigma_i^P}(s^{\Theta(1)})/s^{\Theta(1)}$.

Definition 2.2 (block-wise sources). *A distribution (W_1, W_2) is a (k_1, k_2)-block-wise source if $H_\infty(W_1) \geq k_1$, and for every $w_1 \in \text{Supp}(W_1)$, $H_\infty(W_2 \mid W_1 = w_1) \geq k_2$.*

We use the following standard lemma.

Lemma 2.3. *Let X, Y be random variables, such that $H_\infty(X) \geq k$ and Y is over $\{0,1\}^m$. For every $\eta > 0$, with probability at least $1 - \eta$ over choosing $y \leftarrow Y$, we have that $H_\infty(X \mid Y = y) \geq k - m - \log \frac{1}{\eta}$.*

2.3 Samplable Distributions

We repeat the standard definition of samplable distributions, that appeared in Section 1 as Definition 1.2.

Definition 2.4 (Sampling procedures and samplable distributions). *For a function $A : \{0,1\}^r \to \{0,1\}^n$, we use $Z \leftarrow A$ to denote the experiment in which $W \leftarrow U_r$, and $Z = A(W)$, and say that Z is **sampled** by A. We say that the distribution Z is **samplable** by a class $\mathcal{C}$ of functions, if there exists $A \in \mathcal{C}$ that samples Z.*

We now give a more general definition of samplable distributions with postselection which generalizes Definition 1.5.

Definition 2.5 (Samplable distributions with postselection by Σ_i-circuits). *For functions $A : \{0,1\}^r \to \{0,1\}^n$, and $P : \{0,1\}^r \to \{0,1\}$, we use $Z \leftarrow A \mid P$ to denote the experiment in which $W \leftarrow U_r$, and $Z = (A(W) \mid P(W) = 1)$, and say that Z is **sampled** by A with postselection by P (or that Z is sampled by $A \mid P$ for brevity).*

*We say that the distribution Z is (s, i)-**samplable**, if there exists circuit A of size s and a Σ_i-circuit of size s such that Z is sampled by $A \mid P$.*

We remark that for $i = 0$, the notion of $(s, 0)$-samplable distributions captures distributions samplable by circuits of size s, with postselection by circuits of size s. This is the notion that was considered in the introduction. We will later also consider cases where $i > 0$.

2.4 Extractors and Related Objects

We will be interested in several flavors of extractors and related objects.

2.4.1 Seedless Extractors

We repeat the standard definition of seedless extractors, that appeared in Section 1 as Definition 1.1.

Definition 2.6 (Seedless extractor). *A function $\text{Ext} : \{0,1\}^n \to \{0,1\}^m$ is a (k, ϵ)-**extractor** for a class $\mathcal{D}$ of distributions, if for every distribution X in $\mathcal{D}$, that is over $\{0,1\}^n$, such that $H_\infty(X) \geq k$, $\text{Ext}(X)$ is ϵ-close to U_m.*

We repeat the definition of "multiplicative extractors" given in Section 1 as Definition 1.7

Definition 2.7 (Multiplicative seedless extractor [Sha24]). *A function $\text{Ext} : \{0,1\}^n \to \{0,1\}^m$ is a (k, ϵ)-**multiplicative extractor** for a class $\mathcal{D}$ of distributions, if for every distribution X in $\mathcal{D}$, that is over $\{0,1\}^n$, such that $H_\infty(X) \geq k$, $\text{Ext}(X)$, and every event $A \subseteq \{0,1\}^m$*

$$\Pr[\text{Ext}(X) \in A] \leq e^\epsilon \cdot \Pr[U_m \in A].$$

Using the fact that for $0 < \epsilon \leq 1$, $e^\epsilon \leq 1 + 2\epsilon$, the following proposition immediately follows:

Proposition 2.8 (Multiplicative extractors imply standard extractors). *For every class $\mathcal{D}$ and $0 < \epsilon \le 1$, a (k, ϵ)-multiplicative-extractor for $\mathcal{D}$ is a $(k, 2\epsilon)$-extractor for $\mathcal{D}$.*

The motivation behind the definition of multiplicative extractors, is that even with large error of say $\epsilon = 1/10$, multiplicative extractors guarantee that an event $A \subseteq \{0,1\}^m$ that occurs with probability at most $n^{-\omega(1)}$ under the uniform distribution, occurs with probability $n^{-\omega(1)}$ under the distribution $\mathsf{Ext}(X)$. This is beneficial because (as discussed in detail in [AASY15, Sha24] there are barriers for obtaining extractors for samplable distributions with $\epsilon = n^{-\omega(1)}$ [AASY15].

2.4.2 Errorless Condensers

We repeat the definition of errorless condensers, that appeared in Section 1 as Definition 1.9.

Definition 2.9 (Errorless condenser). *A function $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n_1}$ is a (k, k_{out})-**errorless condenser** for a class $\mathcal{D}$ of distributions, if for every distribution X in $\mathcal{D}$, that is over $\{0,1\}^n$, such that $H_\infty(X) \ge k$, $H_\infty(\mathsf{Cnd}(X)) \ge k_{\mathsf{out}}$.*

2.4.3 Seeded Extractors

We need the following standard definition of strong seeded extractors.

Definition 2.10 (Strong extractors). *A function $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a strong (k, ϵ)-extractor if for every distribution X over $\{0,1\}^n$ with $H_\infty(X) \ge k$, the distribution $Z = (Y, E(X, Y))$ where $Y \leftarrow U_d$ is ϵ close to U_{m+d}.*

We now state several explicit constructions of extractors.

Theorem 2.11 (Strong extractors with logarithmic seed and low error [GUV07]). *There exists a constant $c_1 > 1$ such that for every constant $\alpha > 0$, every sufficiently large n, and every $k > c_1 \log n$, there is a strong (k, ϵ)-extractor $E : \{0,1\}^n \times \{0,1\}^{O(\log \frac{n}{\epsilon})} \to \{0,1\}^{(1-\alpha) \cdot k}$. Furthermore, E can be computed in time $\mathsf{poly}(n)$.*

Theorem 2.12 (Strong extractors with logarithmic seed and larger output length [TU12]). *For every constants $0 < \eta < 1$ and $b \ge 1$, there exists a constant c_1 such that for every sufficiently large n, and every $k \ge 2^{c_1 \cdot \log^\eta n}$, there is a strong $(k, 2^{-\log^\eta n})$-extractor $E : \{0,1\}^n \times \{0,1\}^{O(\log n)} \to \{0,1\}^{k - O(\frac{k}{\log^b n} + \log n)}$. Furthermore, E can be computed in time $\mathsf{poly}(n)$.*

2.4.4 Two-Source Extractors

We repeat the standard definition of 2-source extractors, that appeared in Section 1 as Definition 1.11.

Definition 2.13 (Two-source extractors). *A function $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ is a (k_1, k_2, ϵ)-2-source extractor if for every two independent distributions X_1, X_2 with $H_\infty(X_1) \ge k_1$ and $H_\infty(X_2) \ge k_2$, $\mathsf{TExt}(X_1, X_2)$ is ϵ-close to U_m.*

We use the following explicit construction of 2-source extractors, due to Chattopadhyay and Zuckerman [CZ16], with a later improvement by Li [Li16].

Theorem 2.14 ([CZ16, Li16]). *There exists constant c_0 such that for every constant c_1, every sufficiently large n, and every $k > \log^{(c_0 + c_1)} n$ there is a $(k, k, \frac{1}{n^{c_1}})$-2-source extractor $\mathsf{TExt} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{\Omega(k)}$. Furthermore, TExt can be computed in time $\mathsf{poly}(n^{c_1})$.*

Theorem 2.14 was proven by Chattopadhyay and Zuckerman [CZ16] for the case $m = 1$. This proof was extended by Li [Li16] to handle larger m. The statement in Li's paper is weaker than the one we state here, and only applies for some fixed constant c_1 (rather than any constant c_1). Nevertheless, Li's proof can be extended to yield the statement here by choosing the parameters in the way done by Chattopadhyay and Zuckerman [CZ16].

We will also use an extractor construction of Li [Li15] which works for "one source, and one block-wise source".

Theorem 2.15 ([Li15]). *For every sufficiently large n, and every $k \geq \log^{12} n$ there is a function IExt : $\{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{0.9k}$ such that for every three distributions X_1, X_2, X_3 over $\{0,1\}^n$ such that X_1 is independent of (X_2, X_3), $H_\infty(X_1) \geq k$, and (X_2, X_3) is a (k,k)-block-wise source, $\mathsf{IExt}(X_1, X_2, X_3)$ is ϵ-close to U_m, for $\epsilon = 2^{-k^{\Omega(1)}}$. Furthermore, IExt can be computed in time $\mathsf{poly}(n)$.*

2.4.5 Two-Source Condensers

We will also be interested in 2-source condensers, and especially in ones that are errorless.

Definition 2.16 (Two-source condensers). *A function $\mathsf{TCnd} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ is a (k_1, k_2, k', ϵ)-2-source condenser if for every two independent distributions X_1, X_2 with $H_\infty(X_1) \geq k_1$ and $H_\infty(X_2) \geq k_2$, $\mathsf{TExt}(X_1, X_2)$ is ϵ-close to some distribution Z over $\{0,1\}^m$ with $H_\infty(Z) \geq k'$. We omit ϵ, and say that TCnd is an errorless (k_1, k_2, k')-2-source condenser if it is a $(k_1, k_2, k', 0)$-2-source condenser.*

We use the following explicit construction of 2-source condensers, due to Ben-Aroya, Cohen, Doron and Ta-Shma [BCDT19].

Theorem 2.17 ([BCDT19]). *There exists a constant $c \geq 1$, such that for every $n \geq k$, and every $\epsilon > 0$ such that $k \geq (\log \frac{n}{\epsilon})^c$ there is a (k, k, k', ϵ)-2-source condenser $\mathsf{TCnd} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ for $m = k - 5\log(1/\epsilon) - O(1)$ and $k' = m - o(\log(1/\epsilon))$. Furthermore, TCnd can be computed in time $\mathsf{poly}(n, \log(1/\epsilon))$.*

For our purposes, it will be more convenient to state the following corollary, which aims for weaker parameters, but achieves an errorless 2-source condensers.

Corollary 2.18 ([BCDT19]). *There exist constants $c_0 \geq 1$ and $\delta_0 > 0$, such that for every sufficiently large n, k, and every m, such that $k \geq 2m \geq (\log n)^{c_0}$, there is a (k, k, m^{δ_0})-errorless 2-source condenser $\mathsf{TCnd} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ that can be computed in time $\mathsf{poly}(n)$.*

Proof. Let c be the constant from Theorem 2.17. We set $\alpha = \frac{1}{2c}$, $c_0 = 2c$ and $\epsilon = 2^{-m^\alpha}$. Given a sufficiently large n, k, m such that $k \geq 2m \geq (\log n)^{c_0}$, we want to apply Theorem 2.17, and need to verify that $k \geq (\log \frac{n}{\epsilon})^c$, and this holds because by our choices, $k^{1/2} \geq (\log n)^c$ and

$$(\log \frac{n}{\epsilon})^c \leq (\log n)^c \cdot (\log \frac{1}{\epsilon})^c \leq (\log n)^c \cdot m^{\alpha c} \leq (\log n)^c \cdot m^{1/2} \leq (\log n)^c \cdot k^{1/2} \leq k,$$

giving that the assumption of Theorem 2.17 is satisfied, and we obtain (k, k, k', ϵ)-2-source condenser TCnd : $\{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{m'}$ for $m' = k - 5\log(1/\epsilon) - O(1) = k - 5 \cdot m^\alpha - O(1) \geq k/2$, and $k' = m' - g$ for $g = o(\log(1/\epsilon)) = o(m^\alpha)$. We will use the function $\mathsf{TCnd}' : \{0,1\}^n \to \{0,1\}^n \to \{0,1\}^m$ which is defined by $\mathsf{TCnd}'(x_1, x_2) = \mathsf{TCnd}'(x_1, x_2)_{1,\dots,m}$ (namely, the first m output bits). It is standard that if a distribution Z on m' bits is ϵ-close to having min-entropy $m' - g$, then truncating Z to the first m bits, gives a distribution that is ϵ-close to having min-entropy $m - g$. It follows that TCnd' is a $(k, k, m - g, \epsilon)$-2-source condenser, and this implies that TCnd' is a (k, k, m^{δ_0})-errorless condenser for $\delta_0 = \frac{\alpha}{2} = \frac{1}{4c}$, because for every two

15

independent distributions X_1, X_2 such that $H_\infty(X_1) \geq k$, and $H_\infty(X_2) \geq k$, and every $z \in \{0,1\}^m$, we have that:

$$\Pr[\mathsf{TCnd}'(X_1, X_2) = z] \leq 2^{-(m-g)} + \epsilon \leq 2^{-(m-m^\alpha)} + 2^{-m^\alpha} \leq 2^{-m^{\alpha/2}} = 2^{-m^{\delta_0}}.$$

which gives that $H_\infty(\mathsf{TCnd}'(X_1, X_2) = z) \geq m^{\delta_0}$. $\qquad\qquad\square$

2.5 Impagliazzo-Wigderson Style Hardness Assumptions

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

Definition 2.19 (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.*

Remark 2.20 (Ladder Climbing). *The assumption that E is hard for exponential size Σ_i circuits is typically used to construct functions that are secure (in some sense) against circuits of size n^c, and are computable in larger time $\mathsf{poly}(n^c)$.*

Typically, these proofs allow "ladder climbing", meaning that they immediately extend to show that for every $j \geq 0$, if E is hard for exponential size Σ_{i+j} circuits then the construction gives functions that are secure against Σ_j-circuits of size n^c, and are computable in time $\mathsf{poly}(n^c)$.

This immediately follows because the proofs typically use the hardness of the problem in the hardness assumption to argue that the function is secure (in a relativizing argument) and so prove the statement relative to a Σ_j^P-oracle. On the other hand, the fact that the function is easy to compute, and is computable in time $\mathsf{poly}(n^c)$ follows by a separate and independent argument that only relies on the easiness of the problem in the hardness assumption.

This observation is used in many of the past works, starting with [TV00], and we will use it extensively in this paper.

2.6 Functions that are Hard on Samplable Distributions with Sufficient Min-Entropy (HOS)

Ball, Shaltiel and Silbak [BSS25] introduced the notion of a function that is hard on average every samplable distribution with sufficient min-entropy (HOS), which is closely related to a notion of functions that are hard to sample (HTS) introduced in [SS24].

Definition 2.21 (A function that is hard on samplable distributions (HOS) [BSS25]). *A function $g : \{0,1\}^n \to \{0,1\}^m$ is a (k, ρ)-HOS for a class $\mathcal{D}$ of distributions, against a class a class $\mathcal{C}$, if for every distribution $Y \in \mathcal{D}$ over $\{0,1\}^n$ that has $H_\infty(Y) \geq k$, and every function $C : \{0,1\}^n \to \{0,1\}^m$ in $\mathcal{C}$,*

$$\Pr[C(Y) = g(Y)] \leq \rho.$$

The definition below considers a slightly more general scenario than the one considered in [BSS25]. In [BSS25], the class $\mathcal{D}$ was fixed to be the class of distributions samplable by (deterministic) circuits. In this paper, we consider a more general scenario where the class $\mathcal{D}$ is the class of (s, j)-samplable distributions.

We will use the following theorem from [BSS25].

Theorem 2.22 ([BSS25]). *There exists a constant $a_0 > 1$ such that for every integer $j \geq 0$, if E is hard for exponential size Σ_{j+4}-circuits, then for every constant $c > 1$, and every constant $\nu > 0$, there is a constant $d > 1$ such that every sufficiently large n, and every $k \geq 4n^\nu$ there is a function $g : \{0,1\}^n \to \{0,1\}^{n^{a_0}}$ that is a $(k, 2^{-\Omega(k)})$-HOS for the class of $(n^c, j+1)$-samplable distributions, against the class of Σ_{j+2}-circuits of size n^c. Furthermore, f is computable in time n^d.*

The statement that we make here of Theorem 2.22 is more general than the one made in [BSS25]. Nevertheless, as we now explain, this more general statement immediately follows from the proof of [BSS25]. More specifically, the statement made in [BSS25] is weaker in two respects:

- In [BSS25] it is stated only for $j = 0$, whereas here we state it for general j. The more general statement is achieved by "ladder climbing", see Remark 2.20.

- In [BSS25] it is stated where the class $\mathcal{D}$ is the class of distributions samplable by size n^c circuits, whereas here we state it for the stronger class of $(n^c, j + 1)$-samplable distributions. Nevertheless, the proof in [BSS25] immediately gives this stronger statement, as we now explain: The proof in [BSS25] actually constructs a stronger object called a "min-entropy HTS" against Σ_{j+2}-circuits of size n^c. This immediately yields an HOS where the class $\mathcal{D}$ is the class of distributions samplable by Σ_{j+2}-circuits of size n^c, and the class $\mathcal{C}$ is the class of Σ_2-circuits of size n^c.

  It is easy to see that every distribution that is $(n^c, j + 1)$-samplable, can be sampled (to within a tiny statistical error of less than say 2^{-n}) by a Σ_{j+2}-circuit of size $\mathsf{poly}(n^c)$. This immediately gives that (by choosing the constant c to be a bit larger) an HOS for the class $\mathcal{D}$ above is also an HOS for the class of $(n^c, j + 1)$-samplable distributions.

We remark that while the theorem applies to the class of $(n^c, j + 1)$-samplable distributions, in this paper it would have been sufficient to state it for (n^c, j)-samplable distributions.

2.7 Approximate Counting and Uniform Sampling of NP Witnesses

We use the classical result on approximate counting and uniform sampling of NP-witnesses [Sto83, Sip83, JVV86, BGP00], which we state below in a way that is convenient for our application.

Definition 2.23 (Relative approximation). *We say that a number p is an ϵ-relative approximation to q if $(1 - \epsilon) \cdot p \leq q \leq (1 + \epsilon) \cdot p$, and an ϵ-additive approximation to q if $|p - q| \leq \epsilon$.*

It is useful to note that if $0 \leq p \leq 1$ is an ϵ-relative approximation to q, then it is also an additive approximation to q. For $\epsilon \leq \frac{1}{2}$, we also have the following: If p is an ϵ-relative approximation to q, then q is an $O(\epsilon)$-relative approximation to p. If p is an ϵ-relative approximation to q and q is an ϵ-relative approximation to w, then p is an $O(\epsilon)$-relative approximation to w. If p' is an ϵ-relative approximation to p and q' is an ϵ-relative approximation to q, then a p'/q' is an $O(\epsilon)$-relative approximation to p/q. (The last property does not hold if we replace relative approximations with additive approximations).

Theorem 2.24 (Approximate counting [Sto83, Sip83, JVV86]). *For every i, every sufficiently large s, and every $\epsilon > 0$, there is a size $\mathsf{poly}(s/\epsilon)$ Σ_{i+1}-circuit that given a size s Σ_i-circuit C, outputs an ϵ-relative approximation of $|\{x : C(x) = 1\}|$.*

Theorem 2.25 (Uniform sampling [JVV86, BGP00]). *For every i, every sufficiently large s, and every $\delta > 0$, there is a size $\mathsf{poly}(s, \log(1/\delta))$ randomized Σ_{i+1}-circuit A that given a size s Σ_i-circuit $C : \{0, 1\}^n \to \{0, 1\}$, outputs a value in $\{0, 1\}^n \cup \bot$ such that $\Pr[A(C) = \bot] \leq \delta$ and the distribution $(A(C) \mid A(C) \neq \bot)$ is uniform over $\{x : C(x) = 1\}$.*

Regarding the formulation of Theorems 2.24 and 2.25. We state Theorems 2.24 and Theorem 2.25 for general i, whereas typically they are only stated for $i = 0$.

The formulation in the two theorems only requires that the tasks be achieved by (nonuniform) circuits. The classical results in this area, are in fact stronger. For $i = 0$, Theorem 2.25 holds for A that is a randomized uniform algorithm with an NP oracle (which is stronger than the statement we give here). Similarly, for $i = 0$,

Theorem 2.24 holds for a counting procedure that is a randomized uniform algorithm with an NP oracle. Here, we state it for a circuit (which is nonuniform, and non-randomized). This immediately follows by Adleman's proof that BPP $\subseteq$ P/poly which extends to BPP^{NP} $\subseteq$ P^{NP}/poly.

3 A Construction of an Errorless Condenser

In this section we present our constructions of errorless condensers, and prove Theorems 1.10 which we now restate in a more general way. In Section 4 we will use this errorless condenser to derive our extractors for samplable distributions.

Theorem 3.1 (Final condenser). *For every constants $\alpha > 0$ and $0 < \xi \leq \frac{\alpha}{10}$, such that $\alpha + \xi < 1$, there exist constants $j = \lceil \frac{1-\alpha-\xi}{\alpha-\xi} \rceil + 4$, and $0 < \delta_1 < \delta_2 \leq \frac{\alpha}{10}$ such that if E is hard for exponential size Σ_j-circuits, then for every constant $c \geq 1$, there is a constant $d \geq 1$, such that for every sufficiently large n, there is a function $\mathsf{FCnd} : \{0,1\}^n \to \{0,1\}^{n^{\delta_2}}$ that is an (n^α, n^{δ_1})-errorless condenser for $(n^c, 0)$-samplable distributions. Furthermore, FCnd can be computed in time n^d.*

Showing that the formulation in Theorem 1.10 follows from Theorem 3.1. Theorem 1.10 follows from this more general formulation as the function $f(x) = \frac{1-x}{x}$ is decreasing in $(0, 1]$. This means that for $\alpha > \frac{1}{i}$, $\frac{1-\alpha}{\alpha} < \frac{1-1/i}{1/i} = i - 1$, and we can take $\xi > 0$ to be sufficiently small so that $\frac{1-\alpha-\xi}{\alpha-\xi} < i - 1$, which gives that $j = \lceil \frac{1-\alpha-\xi}{\alpha-\xi} \rceil + 4 \leq i + 3$, and indeed Theorem 1.10 is stated with $j = i + 3$.

Outline for this section. The remainder of this section is devoted to proving Theorem 3.1. In Section 3.1 we state and prove a lemma (Lemma 3.2) that will allow us to construct errorless condensers, whenever we can "split" a given samplable source into a block-wise source. (We explained the high level intuition of this argument in Section 1.3.3, and the formulation and proof of Lemma 3.2 below, closely follows this intuition).

In Section 3.2 we will use the "oracle condenser" of Lemma 3.2 to construct a "basic condenser" which has output length that is longer than what we want in Theorem 3.1. (We explained the high level intuition of this argument in Section 1.3.4). Finally, in Section 3.3 we will use the "basic condenser" in order to construct the final condenser and prove Theorem 3.1.

3.1 An Oracle Condenser

In this section we state and prove a lemma that is the formal instantiation of the ideas explained in Section 1.3.3, and will be used in the construction of our basic condenser and final condenser. More specifically, we devise an oracle procedure $\mathsf{OCnd}^{(\cdot)}$ which we call an "oracle condenser". When given a samplable distribution X as input, OCnd expects to receive oracle access to a function $B : \{0,1\}^n \to \{0,1\}^t$ for $t \leq n$, such that B "splits X into a block-wise source" (more formally, that $(B(X), X)$ is a block-wise source). If OCnd is indeed supplied with such a function B, it acts as an errorless condenser and produces a short output that has some min-entropy. The precise formulation is given below.

Lemma 3.2. *There exist constants $c_0 \geq 1$ and $\delta > 0$, such that for every constants $c \geq 1$, $j \geq 0$ and $\nu > 0$, if E is hard for exponential size Σ_{j+4}-circuits, then there exists a constant d such that for every sufficiently large n, and every $(\log n)^{c_0} \leq m \leq n^{\nu/2}$, there is an oracle procedure $\mathsf{OCnd}^{(\cdot)} : \{0,1\}^n \to \{0,1\}^m$ such that for every $t \leq n$ and every function $B : \{0,1\}^n \to \{0,1\}^t$, that has a circuit of size n^c, OCnd^B can be computed in time n^d, and is an (n^ν, m^δ)-errorless condenser for all (n^c, j)-samplable distributions X over $\{0,1\}^n$, that have the additional property that $(B(X), X)$ is a (n^ν, n^ν)-block-wise source.*

We plan to use OCnd as a procedure in our constructions of basic condenser and final condenser (that appear later in this section). We start with proving Lemma 3.2.

3.1.1 Proof of Theorem 3.2

Figure 1: Construction of function OCnd

Parameters:

- Let c_0 and δ_0 be the constants from Corollary 2.18. We choose $\delta = \delta_0/2$.
- We are given constants $c \geq 1$, $j \geq 0$, $\nu > 0$.
- Let $a \geq 1$ be a constant that will be chosen to be sufficiently large in the proof.
- We are assuming that n is sufficiently large, and receive an additional parameter m, such that $(\log n)^{c_0} \leq m \leq n^{\nu/2}$.

Assumption: We are assuming that E is hard for exponential size Σ_{j+4}-circuits.

Ingredients: We will require the following ingredients:

- *HOS:* By Theorem 2.22, under the hardness assumption, there is a function $g : \{0,1\}^n \to \{0,1\}^{n^{a_0}}$ (where $a_0 > 1$ is a universal constant) that is an $(n^\nu, 2^{-\Omega(n^\nu)})$-HOS for the class of $(n^{c \cdot a}, j)$-samplable distributions, against the class of size $n^{c \cdot a}$ Σ_{j+2}-circuits. We have that g can be computed in time n^{d_g} for some constant d_g that depends on c, a and ν.
- *2-source errorless condenser:* Let $k' = \frac{n^\nu}{a}$ and let TCnd $: \{0,1\}^{n^{a_0}} \times \{0,1\}^{n^{a_0}} \to \{0,1\}^m$ be the (k', k', m^{δ_0})-errorless 2-source condenser that is guaranteed by Corollary 2.18. We have TCnd is can be computed in time $\mathrm{poly}(n)$.

Construction: We define $\mathsf{OCnd}^{(\cdot)} : \{0,1\}^n \to \{0,1\}^m$, as follows: Given oracle access to a function $B : \{0,1\}^n \to \{0,1\}^t$, with $t \leq n$, we define:
$$\mathsf{OCnd}^B(x) = \mathsf{TCnd}(g(B(x)), x).$$

Note that g expects inputs of length n, and TCnd expects inputs of length n^{a_0}, and so:

- If $t \leq n$, we pad $B(x)$ with $n - t$ zeros, before applying g on $B(x)$.
- We pad x with $n^{a_0} - n$ zeros, before giving it as a second input to TCnd.

Note that by construction, there indeed exists a constant d, such that if B can be computed in time n^c, then OCnd^B can be computed in time n^d (where d depends on d_g, and c).

The construction of the oracle procedure OCnd appears in Figure 1 and implements the idea explained in Section 1.3.3.

Let n be sufficiently large, and assume for the purpose of contradiction that there is a function $B : \{0,1\}^n \to \{0,1\}^t$, and an (n^c, j) samplable distribution X such that:

1. B has a circuit of size n^c.

2. $(B(X), X)$ is an (n^ν, n^ν)-block-wise source.

3. There exists a $z \in \{0,1\}^m$, such that $\Pr[\mathsf{TCnd}(g(B(X)), X) = z] > 2^{-m^\delta}$.

Our goal will be to obtain a contradiction by showing that there is a Σ_{j+2}-circuit C of size $n^{c \cdot a}$, and a $(n^{c \cdot a}, j)$-samplable distribution Y (we will choose $Y = B(X)$) with $H_\infty(Y) \geq n^\nu$, on which $\Pr[C(Y) = g(Y)]$ is too large, and contradicts the HOS guarantee of g.

We have chosen $\delta = \delta_0/2$, which gives that the third item above implies that for $\mu = 2^{-m^{\delta_0}}$ and some

constant $\eta > 0$,

$$\Pr[\mathsf{TCnd}(g(B(X)), X) = z] > (1 + \eta) \cdot \mu.$$

We will say that $y \in \{0, 1\}^t$ is *useful* if

$$\Pr[\mathsf{TCnd}(g(y), X) = z \mid B(X) = y] > (1 + \frac{\eta}{2}) \cdot \mu.$$

Claim 3.3. $\Pr[B(X) \text{ is useful}] > \frac{\eta \cdot \mu}{4}$.

Proof. Let G denote the set of all useful $y \in \{0, 1\}^t$. If the claim does not hold then

$$\begin{aligned}
\Pr[\mathsf{TCnd}(g(B(X)), X) = z] &\leq \Pr[B(X) \in G] + \Pr[\mathsf{TCnd}(g(B(X)), X) = z \cap B(X) \notin G] \\
&\leq \frac{\eta \cdot \mu}{4} + \sum_{y \notin G} \Pr[\mathsf{TCnd}(g(B(X)), X) = z \cap B(X) = y] \\
&= \frac{\eta \cdot \mu}{4} + \sum_{y \notin G} \Pr[\mathsf{TCnd}(g(B(X)), X) = z \mid B(X) = y] \cdot \Pr[B(X) = y] \\
&\leq \frac{\eta \cdot \mu}{4} + \sum_{y \notin G} (1 + \frac{\eta}{2}) \cdot \mu \cdot \Pr[B(X) = y] \\
&\leq \frac{\eta \cdot \mu}{4} + (1 + \frac{\eta}{2}) \cdot \mu \\
&\leq (1 + \eta) \cdot \mu
\end{aligned}$$

which is a contradiction. $\qquad\square$

Let $n' = n^{a_0}$ be the output length of g. For every $y \in \{0, 1\}^t$ and every $0 \leq \alpha \leq 1$, we define:

$$T_{y,\alpha} = \left\{ v \in \{0, 1\}^{n'} : \Pr[\mathsf{TCnd}(v, X) = z \mid B(X) = y] > (1 + \alpha) \cdot \mu \right\}.$$

With this definition we immediately have that for every useful y, $g(y) \in T_{y,\eta/2}$. We now observe that for every y and $\alpha \geq 0$, $T_{y,\alpha}$ is a small set.

Claim 3.4. *For every $y \in \{0, 1\}^t$, and every $\alpha \geq 0$, $|T_{y,\alpha}| < 2^{k'}$.*

Proof. If this does not hold, then there exists a $y \in \{0, 1\}^t$ and $\alpha \geq 0$, such that $|T_{y,\alpha}| \geq 2^{k'}$. We consider the following two distributions: The first is V_y that is uniform over $T_{y,\alpha}$, and the second $W_y = (X \mid B(X) = y)$. These two distributions are independent, and have min-entropy at least k', and therefore, by the guarantee of TCnd, we have that $\Pr[\mathsf{TCnd}(V_y, W_y) = z] \leq \mu$. This is a contradiction as we also have that,

$$\begin{aligned}
\Pr[\mathsf{TCnd}(V_y, W_y) = z] &= \Pr[\mathsf{TCnd}(V_y, X) = z \mid B(X) = y] \\
&= \mathbb{E}_{v \leftarrow T_{y,\alpha}}[\Pr[\mathsf{TCnd}(v, X) = z \mid B(X) = y]] \\
&> (1 + \alpha) \cdot \mu.
\end{aligned}$$

$\qquad\square$

We have that X is (n^c, j)-samplable which means that $X \leftarrow A \mid P$, for some circuit $A : \{0, 1\}^r \to \{0, 1\}^n$ of size n^c, and some Σ_j-circuit $P : \{0, 1\}^r \to \{0, 1\}$ of size n^c. We now define the following circuits.

Definition 3.5. *For every $y \in \{0, 1\}^t$, and $v \in \{0, 1\}^{n'}$ we define two Σ_j circuits $C^1_{y,v} : \{0, 1\}^r \to \{0, 1\}$ and $C^2_y : \{0, 1\}^r \to \{0, 1\}$ of size $\mathsf{poly}(n^c)$ as follows:*

- $C^1_{y,v}(w)$ *answers one iff* $\mathsf{TCnd}(v, A(w)) = z \land B(A(w)) = y \land P(w) = 1$.
- $C^2_y(w)$ *answers one iff* $B(A(w)) = y \land P(w) = 1$.

We also define:

- $p^1_{y,v} = \Pr[C^1_{y,v}(U_r) = 1]$.
- $p^2_y = \Pr[C^2_y(U_r) = 1]$.

Claim 3.6. *For every* $y \in \{0,1\}^t$ *and* $v \in \{0,1\}^{n'}$, *if* $p^2_y \neq 0$, *then*

$$\frac{p^1_{y,v}}{p^2_y} = \Pr[\mathsf{TCnd}(v, X) = z \mid B(X) = y].$$

Proof. For every $y \in \{0,1\}^t$ and $v \in \{0,1\}^{n'}$, if $p^2_y \neq 0$ then for $W \leftarrow U_r$, we have that:

$$
\begin{aligned}
\frac{p^1_{y,v}}{p^2_y} &= \frac{\Pr[\mathsf{TCnd}(v, A(W)) = z \land B(A(W)) = y \land P(W) = 1]}{\Pr[B(A(W)) = y \land P(W) = 1]} \\
&= \frac{\Pr[\mathsf{TCnd}(v, A(W)) = z \land B(A(W)) = y \mid P(W) = 1] \cdot \Pr[P(W) = 1]}{\Pr[B(A(W)) = y \mid P(W) = 1] \cdot \Pr[P(W) = 1]} \\
&= \frac{\Pr[\mathsf{TCnd}(v, A(W)) = z \land B(A(W)) = y \mid P(W) = 1]}{\Pr[B(A(W)) = y \mid P(W) = 1]} \\
&= \frac{\Pr[\mathsf{TCnd}(v, X) = z \land B(X) = y]}{\Pr[B(X) = y]} \\
&= \Pr[\mathsf{TCnd}(v, X) = z \mid B(X) = y].
\end{aligned}
$$

$\square$

This means that for every $y \in \{0,1\}^t$ and $0 \leq \alpha \leq 1$, we can decide whether a given $v \in \{0,1\}^t$ is in $T_{y,\alpha}$ if we can check whether $p^2_y = 0$ and compute $p^1_{y,v}$ and p^2_y. By Theorem 2.7 a small Σ_{j+1}-circuit, can compute relative approximations to $p^1_{y,v}$ and p^2_y. We will now use this idea to prove the following:

Claim 3.7. *For every* $y \in \{0,1\}^t$, *there is a* Σ_{j+1}-*circuit* $C_y : \{0,1\}^{n'} \to \{0,1\}$ *of size* $\mathsf{poly}(n^c)$ *such that:*

- *For every* $v \in \{0,1\}^{n'}$ *such that* $C_y(v) = 1$, *we have that* $v \in T_{y,\frac{\eta}{8}}$.
- *If* y *is useful, then* $C_y(g(y)) = 1$.

Proof. When given $v \in \{0,1\}^{n'}$, the circuit C_y works as follows:

- C_y checks whether there exists $w \in \{0,1\}^r$, such that $B(A(w)) = y$ and $P(w) = 1$. If there does not exist such a w, it answers zero (as this means that $p^2_y = 0$)
- Let $\lambda = \eta/a'$ for a universal constant $a' > 1$ to be chosen later. C_y applies Theorem 2.24 to compute a λ-relative approximations $\hat{p}^1_{y,v}$ and $\hat{p}^2_y$, of $p^1_{y,v}$ and p^2_y respectively. It can do this by computing approximations of the number of accepting inputs of the circuits $C^1_{y,v}$ and C^2_y, respectively.
- C_y computes $\hat{p}_{y,v} = \frac{\hat{p}^1_{y,v}}{\hat{p}^2_y}$ and note that as this is an $O(\lambda)$-relative approximation to $p_{y,v} = \frac{p^1_{y,v}}{p^2_y} = \Pr[\mathsf{TCnd}(v, X) = z \mid B(X) = y]$.
- C_y outputs one if $\hat{p}_{y,v} > (1 + \frac{\eta}{4}) \cdot \mu$ and zero otherwise.

21

By choosing the constant a' to be sufficiently large, we can make $\lambda = \eta/a'$ sufficiently small, to guarantee that checking whether the $O(\lambda)$ approximation $\hat{p}_{y,v}$ is larger than $(1+\frac{\eta}{4})\cdot\mu$ distinguishes between the case that $p_{y,v} > (1+\frac{\eta}{2})\cdot\mu$ and the case that $p_{y,v} \le (1+\frac{\eta}{8})\cdot\mu$. This gives that if $C_y(v) = 1$ then $p_{y,v} > (1+\frac{\eta}{8})\cdot\mu$ which gives that $v \in T_{y,\frac{\eta}{8}}$. We have that for every useful y, $g(y) \in T_{y,\frac{\eta}{2}}$, which means that $p_{y,g(y)} > (1+\frac{\eta}{2})\cdot\mu$, and indeed, $C_y(g(y)) = 1$.

Finally, by definition C_y is a circuit of size $\mathsf{poly}(n^c, \frac{1}{\lambda}) = \mathsf{poly}(n^c)$. $\qquad\square$

We are finally ready to complete the proof, with the next claim.

Claim 3.8. *There is a Σ_{j+2}-circuit C of size $\mathsf{poly}(n^c)$ such that for $Y = B(X)$,*

$$\Pr[C(Y) = g(Y)] \ge 2^{-k'} \cdot \frac{\eta \cdot \mu}{8}$$

Proof. We will first construct a randomized Σ_{j+2}-circuit C', and then use a standard averaging argument to convert it to a non-randomized Σ_{j+2}-circuit. The randomized circuit C' is defined as follows: On input $y \in \{0,1\}^t$:

- C' constructs the Σ_{j+1}-circuit C_y. Note that the circuit C_y is specified precisely in the proof of Claim 3.7, and so, the circuit C' (that can be hardwired with A, P, B, z, and the circuit from Theorem 2.24) can construct the circuit C_y.

- C' uses the Σ_{j+2}-circuit guaranteed in Theorem 2.25 (choosing $i = j+1$ and $\delta = \frac{1}{2}$) to output a uniform element in $\{v : C_y(v) = 1\}$.

By definition, the circuit C' is a randomized Σ_{j+2}-circuit of size $\mathsf{poly}(n^c)$. We conclude that:

$$\begin{aligned}
\Pr[C'(Y) = g(Y)] &\ge \Pr[C'(Y) = g(Y) \mid Y \text{ is useful}] \cdot \Pr[Y \text{ is useful}] \\
&\ge \Pr[C'(Y) = g(Y) \mid Y \text{ is useful}] \cdot \frac{\eta \cdot \mu}{4} \\
&\ge \frac{1}{2} \cdot 2^{-k'} \cdot \frac{\eta \cdot \mu}{4} \\
&= 2^{-k'} \cdot \frac{\eta \cdot \mu}{8}.
\end{aligned}$$

where the first inequality follows from Claim 3.3, and the last inequality follows because by Claim 3.7, for every useful y, $g(y) \in \{v : C_y(v) = 1\}$ which by Claim 3.4, is of size at most $2^{k'}$, and each element in the set is obtained with probability $\frac{1}{2} \cdot 2^{-k'}$.

Finally, by a standard averaging argument, there exists a (non-randomized) Σ_{j+2}-circuit of size $\mathsf{poly}(n^c)$ with the same success probability. $\qquad\square$

We have obtained a Σ_{j+2}-circuit C of size $\mathsf{poly}(n^c)$. We can choose the constant a to be sufficiently large so that the size of C is bounded by $n^{c\cdot a}$. We can view the distribution $Y = B(X)$ (which is over $\{0,1\}^t$) as a distribution over $\{0,1\}^n$ by padding $B(X)$ with zeros (recall that $t \le n$). We have that:

- $H_\infty(Y) \ge n^\nu$, and
- Y is $(n^{c\cdot a}, j)$-samplable (it is sampled by $(B \circ A) \mid P$).

We also have that,

$$\begin{aligned}
\Pr[C(Y) = G(Y)] &\ge 2^{-k'} \cdot \frac{\eta}{8} \cdot \mu \\
&\ge \Omega(2^{-2k'}) \\
&\ge 2^{-\Omega(\frac{n^\nu}{a})},
\end{aligned}$$

where the first inequality follows because η is constant, and for large enough n, $\mu = 2^{-m^{\delta_0}}$ and $m \leq n^{\nu/2} \leq n^{\nu}/a = k'$. Finally, by choosing the constant a to be sufficiently large, we can make sure that the success probability of C violates the HOS guarantee of g.

3.2 The Basic Condenser

In this section we present our construction of the basic condenser, and prove Theorems 1.13 which we now restate in a more general way.

Theorem 3.9. *There exist constants $c_0 \geq 1$ and $\delta_0 > 0$, such that for every constants $c \geq 1$, $0 < \alpha < 1$ and $j \geq 0$, if E is hard for exponential size Σ_{j+5}-circuits then there exists a constant $d \geq 1$, such that for every sufficiently large n, and every $(\log n)^{c_0} \leq m \leq n^{\frac{\alpha}{4}}$, there is a function $\mathsf{BCnd} : \{0,1\}^n \to \{0,1\}^{10n^{1-\alpha} \cdot m}$ that is an $(n^{\alpha}, m^{\delta_0})$-errorless condenser for (n^c, j)-samplable distributions. Furthermore, BCnd can be computed in time n^d.*

The basic condenser will be used in our final condenser (that appears later in this section). We first prove Theorem 3.9.

3.2.1 Proof of Theorem 3.9

The construction and proof of Theorem 3.9 closely follow the high level explanation given in Section 1.3.4. As explained there, the construction and its analysis use (amongst other things) ideas by Saks, Srinivasan and Zhou [SSZ98] and Ta-Shma [Ta-96].

We will choose the constants $c_0 \geq 1$ and $\delta_0 > 0$ specified in Theorem 3.9 relying on the choices of c_0, δ in Lemma 3.2. More specifically, c_0 will be the same as in Lemma 3.2, and $\delta_0 = \delta/2$, where δ is the constant in Lemma 3.2. We are now given constants $c \geq 1$, $0 < \alpha < 1$ and $j \geq 0$. Our goal is to construct an errorless condenser BCnd for distributions X over $\{0,1\}^n$ that are (n^c, j)-samplable. Before presenting the construction of the errorless condenser BCnd, we will make some preparations.

Preparation and notation. For a sufficiently large n, let $k = n^{\alpha}$, $b = k/10$, $\ell = n/k = 10 \cdot n^{1-\alpha}$, and $\epsilon = 2^{-k/100}$. Given a string $x \in \{0,1\}^n$, we will partition it into ℓ blocks of length b. For $i, i' \in [\ell]$, we will use $x[i]$ to denote the b bits long i'th block of x, and $x[i, \ldots, i']$ to denote the $b(i' - i + 1)$ bits long concatenation of the blocks $x[i] \circ \ldots \circ x[i']$.

Let X be an (n^c, j)-samplable distribution over $\{0,1\}^n$. That is, $X \leftarrow A \mid P$ where $A : \{0,1\}^r \to \{0,1\}^n$ is a circuit of size n^c, and $P : \{0,1\}^r \to \{0,1\}$ is a Σ_j-circuit of size n^c.

For $i \in [\ell]$ and $v \in \{0,1\}^{bi}$, we define:

$$k_i(v) = -\log \Pr[X[1, \ldots, i] = v].$$

For $0 \leq i \leq \ell$ and $x \in \{0,1\}^n$, we define $k_i(x) = k_i(x[1, \ldots, i])$. With this notation we have that for every $i \in [\ell]$ and every $x \in \{0,1\}^n$, $\Pr[X[1, \ldots, i] = x[1, \ldots, i]] = 2^{-k_i(x)}$.

A circuit R that approximates $k_i(x)$. We now observe that there is a Σ_{j+1}-circuit of size $\mathsf{poly}(n^c)$ that computes an approximation of $k_i(v)$.

Claim 3.10. *There is a Σ_{j+1}-circuit R of size $\mathsf{poly}(n^c)$ which given input $0 \leq i \leq \ell$ and $v \in \{0,1\}^{ib}$, outputs a number $\hat{k}_i(v)$ such that: $|\hat{k}_i(v) - k_i(v)| \leq 1$.*

Proof. If $i = 0$, then R outputs zero. When the circuit R receives an input $i \in [\ell]$ and $v \in \{0, 1\}^{ib}$, We first consider the Σ_j-circuit $A' : \{0, 1\}^r \to \{0, 1\}$ defined by $A'(w) = 1$ iff $A(w)[1, \ldots, i] = v$ and $P(w) = 1$. Note that:

$$
\begin{aligned}
2^{-k_i(v)} &= \Pr[X[1, \ldots, i] = v] \\
&= \Pr_{W \leftarrow U_r}[A(W)[1, \ldots, i] = v \mid P(W) = 1] \\
&= \frac{\Pr_{W \leftarrow U_r}[A'(W) = 1]}{\Pr_{W \leftarrow U_r}[P(W) = 1]}.
\end{aligned}
$$

By Theorem 2.24, the circuit R (which will be a Σ_{j+1}-circuit of size $\mathrm{poly}(n^c)$) can compute λ-relative approximations to both the enumerator and denominator, for say constant $\lambda = \frac{1}{100}$. R can then compute the division, which is guaranteed to be a $1/10$-relative approximation of $2^{-k_i(v)}$, using this approximation, R can output the required approximation $\hat{k}_i(v)$. $\qquad\square$

Once again, it will be convenient to allow R to receive an input $x \in \{0, 1\}^n$, rather than $v \in \{0, 1\}^{ib}$. More formally, for $i \in [\ell]$ and $x \in \{0, 1\}^n$, we define $\hat{k}_i(x) = k_i(x[1, \ldots, i])$, and will assume w.l.o.g. that $R(i, x) = \hat{k}_i(x)$.

Partitioning the support of X. In the next definition, we partition $\mathrm{Supp}(X)$ into $\ell + 1$ sets, $T'_0, \ldots, T'_{\ell+1}$. We will later claim that $\Pr[X \in T'_0]$ is small, and for every $i \in [\ell]$, the distribution $(X[1, \ldots, i], X)$ is a $(k/4, k/4)$-block-wise source when X is conditioned on the event $\{X \in T'_i\}$.

Definition 3.11.

- *For every $x \in \mathrm{Supp}(x)$, let $i(x)$ denote the smallest $i \in [\ell]$, such that $\hat{k}_i \geq k/2$. (Such an i exists because $k_\ell(x) \geq k$, and $\hat{k}_\ell(x) \geq k_\ell(x) - 1$). Furthermore, note that given $i \in [\ell]$ and $x \in Supp(X)$, whether $i(x) = i$ is determined by i and $x[1, \ldots, i]$.*

- *We will say that $x \in \mathrm{Supp}(X)$ is useful if $\hat{k}_{i(x)} \leq k/2 + b + 2\log n + \log(1/\epsilon)$.*

- *For every $i \in [\ell]$, we define $T_i = \{x \in \mathrm{Supp}(X) : i(x) = i, \text{ and } x \text{ is useful}\}$.*

- *We define $T_0 = \{x : x \text{ is not useful}\}$*

- *We define $B = \{i : \Pr[X \in T_i] \leq \frac{\epsilon}{n^2}\}$.*

- *For every $i \in [\ell]$ we define $T'_i = \begin{cases} T_i, & \text{if } i \notin B \\ \emptyset, & \text{if } i \in B \end{cases}$*

- *We define $T'_0 = \cup_{i \in B \cup \{0\}} T_i$.*

This definition is made so that we obviously have that:

- T'_0 and $(T'_i)_{i \notin B}$ are a partition of $\mathrm{Supp}(X)$.

- For every $i \in [\ell]$ such that $i \notin B$, $\Pr[X \in T'_i] \geq \frac{\epsilon}{n^2}$.

- For every $x \in \mathrm{Supp}(X)$ that is useful, if $i(x) \notin B$, then $x \in T'_{i(x)}$, and otherwise $x \in T'_0$.

- For every $x \in \mathrm{Supp}(X)$ that is not useful, $x \in T'_0$.

- For every $x \in \mathrm{Supp}(X)$ and $i \in [\ell]$, whether or not $x \in T'_i$, is determined by i and $x[1, \ldots, i]$. (It is sufficient to check whether $i \in B$, and examine $\hat{k}_1(x), \ldots, \hat{k}_i(x)$ which are determined by $x[1, \ldots, i]$).

We start by showing that the probability that X is not useful is small.

24

Claim 3.12. $\Pr[X \text{ is not useful}] \le \frac{\epsilon}{2}$

Proof. For every $i \in [\ell]$, we define

$$P_i = \{x \in \mathrm{Supp}(X) : x \text{ is not useful, and } i(x) = i\},$$

so that $\{x \in \mathrm{Supp}(X) : x \text{ is not useful}\} = \cup_{i \in [\ell]} P_i$.

We observe that for every $i \in [\ell]$ and $x \in \mathrm{Supp}(X)$, whether or not $x \in P_i$ is determined by i and $x[1, \ldots, i]$. This follows because given i and $x[1, \ldots, i]$, we can determine whether $i(x) = i$. If $i(x) \ne i$, then $x \notin P_i$. If $i(x) = i$, whether or not $x \in P_i$ depends on whether x is useful, which is determined by $\hat{k}_i(x)$ which is determined by $x[1, \ldots, i]$.

This means that for every $i \in [\ell]$, we can define a set $P_i' \subseteq \{0,1\}^{ib}$, such that $x \in P_i$ if and only if $x[1, \ldots, i] \in P_i'$. This gives that:

$$\Pr[X \text{ is not useful}] = \Pr[X \in \cup_{i \in [\ell]} P_i] \le \sum_{i \in [\ell]} \Pr[X \in P_i] = \sum_{i \in [\ell]} \Pr[X[1, \ldots, i] \in P_i'].$$

It therefore remains to upper-bound $\Pr[X[1, \ldots, i] \in P_i']$. For this purpose, we fix some $i \in [\ell]$ and note that for every $z \in \{0,1\}^{(i-1)b}$ and $y \in \{0,1\}^b$, such that $v = z \circ y \in P_i'$, we have that:

$$
\begin{aligned}
2^{-k_i(v)} &= \Pr[X[1, \ldots, i] = v] \\
&= \Pr[X[1, \ldots, i-1] = z] \cdot \Pr[X[i] = y \mid X[1, \ldots, i-1] = z] \\
&= 2^{-k_{i-1}(v)} \cdot \Pr[X[i] = y \mid X[1, \ldots, i-1] = z].
\end{aligned}
$$

Recall that we have that $|\hat{k}_i(v) - k_i(v)| \le 1$, and so, rearranging we get that:

$$\Pr[X[i] = y \mid X[1, \ldots, i-1] = z] \le 2^{-(k_i(v) - k_{i-1}(v))} \le 2^{-(\hat{k}_i(v) - \hat{k}_{i-1}(v) - 2)} \le 2^{-(b + 2\log n + \log(1/\epsilon) - 2)},$$

where the last inequality follows because for $v \in P_i'$, we have that $\hat{k}_{i-1}(v) < k/2$ (as i is the smallest index such that $k_i(v) \ge k/2$) and $\hat{k}_i(v) > k/2 + b + 2\log n + \log(1/\epsilon)$ (because $v \in P_i'$ and therefore v not useful, and $i(v) = i$).

We now proceed with our plan, to upper-bound $\Pr[X[1, \ldots, i] \in P_i']$.

$$\Pr[X[1, \ldots, i] \in P_i'] = \sum_{z \in \{0,1\}^{(i-1)b}} \Pr[X[1, \ldots, i] \in P_i' \mid X[1, \ldots, i-1] = z] \cdot \Pr[X[1, \ldots, i-1] = z].$$

Thus, it is suffcient to upper-bound $\Pr[X[1, \ldots, i] \in P_i' \mid X[1, \ldots, i-1] = z]$ for every $z \in \{0,1\}^{(i-1)b}$. For every $z \in \{0,1\}^{(i-1)b}$, we have that:

$$
\begin{aligned}
\Pr[X[1, \ldots, i] \in P_i' \mid X[1, \ldots, i-1] = z] &= \sum_{y \in \{0,1\}^b : \, z \circ y \in P_i'} \Pr[X[i] = y \mid X[1, \ldots, i-1] = z] \\
&\le \sum_{y \in \{0,1\}^b : \, z \circ y \in P_i'} 2^{-(b + 2\log n + \log(1/\epsilon) - 2)} \\
&\le 2^b \cdot 2^{-(b + 2\log n + \log(1/\epsilon) - 2)} \\
&\le \frac{\epsilon}{2n}
\end{aligned}
$$

25

It now remains to put everything together:

$$
\begin{aligned}
\Pr[X \text{ is not useful}] &\leq \sum_{i \in [\ell]} \Pr[X[1, \ldots, i] \in P_i'] \\
&\leq n \cdot \sum_{z \in \{0,1\}^{(i-1)b}} \Pr[X[1, \ldots, i] \in P_i' \mid X[1, \ldots, i-1] = z] \cdot \Pr[X[1, \ldots, i-1] = z] \\
&\leq n \cdot max_{z \in \{0,1\}^{(i-1)b}} \Pr[X[1, \ldots, i] \in P_i' \mid X[1, \ldots, i-1] = z] \\
&\leq n \cdot \frac{\epsilon}{2n} \\
&\leq \frac{\epsilon}{2}.
\end{aligned}
$$

$\square$

A "selector circuit" S. We will now show that there exists a "selector circuit" Σ_{j+1}-circuit S of size $\mathsf{poly}(n^c)$, which given $x \in \mathrm{Supp}(X)$ computes the unique $0 \leq i \leq \ell$ such that $x \in T_i'$. Furthermore, we will show that $\Pr[S(X) = 0]$ is very small, and for every $i \in [\ell]$, the distribution $(X[1, \ldots, i], X)$ is a $(k/4, k/4)$-block-wise source when X is conditioned on the event $\{S(X) = i\}$.

Lemma 3.13. *For every distribution X that is (n^c, j)-samplable, and has $H_\infty(X) \geq k$, there is a Σ_{j+1}-circuit S of size $\mathsf{poly}(n^c)$, such that:*

- *There exists a set $B \subset [\ell]$, such that for every $x \in \mathrm{Supp}(X)$, $S(X) \in ([\ell] \setminus B) \cup \{0\}$, and for every $i \notin B$, $\Pr[S(X) = i] > 0$.*

- $\Pr[S(X) = 0] \leq \epsilon$.

- *For every $i \in [\ell] \setminus B$, the distribution $(X \mid S(X) = i)$ is $(\mathsf{poly}(n^c), j + 1)$-samplable.*

- *For every $i \in [\ell] \setminus B$, the distribution $((X[1, \ldots, i], X) \mid S(X) = i)$ is a $(\frac{k}{4}, \frac{k}{4})$-block-wise source.*

Proof. The circuit S will be hardwired with the set B from definition 3.11. Given x, the circuit S will do the following.

- Use the circuit R to compute $\hat{k}_0(x), \ldots, \hat{k}_\ell(x)$.

- Find $i(x)$ (recall that this is the smallest i such that $\hat{k}_i \geq k/2$).

- If $i(x) \in B$, output zero (as in this case $x \in T_0'$).

- Check whether x is useful (by checking if $\hat{k}_{i(x)} \leq k/2 + b + 2 \log n + \log(1/\epsilon)$).

- If x is useful, output $i(x)$, and otherwise output zero.

It is obvious by construction that for $x \in \mathrm{Supp}(X)$, $S(x) = i$ if and only if $x \in T_i'$. The first item immediately follows as we have seen that the sets T_0' and $(T_i')_{i \notin B}$ are a partition of $\mathrm{Supp}(X)$.

We now turn our attention to the second item. We have defined $T_0' = \cup_{i \in B \cup \{0\}} T_i$. For every $i \in B$, by definition $\Pr[X \in T_i] \leq \frac{\epsilon}{n^2}$, and by Claim 3.12, $\Pr[X \in T_0] = \Pr[X \text{ is not useful}] \leq \frac{\epsilon}{2}$. Overall, we have that by a union bound:

$$
\Pr[S(X) = 0] = \Pr[X \in T_0'] \leq \ell \cdot \frac{\epsilon}{n^2} + \frac{\epsilon}{2} \leq \epsilon.
$$

The third item follows as for every $i \in [\ell]$, such that $\Pr[S(X) = i] > 0$, the distribution $(X \mid S(X) = i)$ is sampled by $A|P'$, where $P'(w)$ is a Σ_{j+1} circuit of size $\mathsf{poly}(n^c)$, defined by $P'(w) = P(w) \wedge S(A(w)) = i$. This is because for $W \leftarrow U_r$, $(A(W) \mid P'(W) = 1)$ is identical to $(X \mid S(X) = i)$.

We will now prove the fourth item. Let $i \in [\ell]$ be an index such that $\Pr[S(X) = i] > 0$, and in particular, $i \notin B$. Let $x \in \mathrm{Supp}(X)$, be such that $S(x) = i$, which means that $x \in T_i'$. We have that:

$$\Pr[X[1, \ldots, i] = x[1, \ldots, i] \mid S(X) = i] = \frac{\Pr[X[1, \ldots, i] = x[1, \ldots, i]] \wedge S(X) = i]}{\Pr[S(X) = i]}$$

$$\leq \frac{\Pr[X[1, \ldots, i] = x[1, \ldots, i]]]}{\Pr[X \in T_i']}$$

$$\leq \frac{2^{-k_i(x)}}{\epsilon/n^2}$$

$$\leq 2^{-(\hat{k}_i(x) - 1 - 2\log n - \log(1/\epsilon))}$$

$$\leq 2^{-k/2 - 1 - 2\log n - \log(1/\epsilon))}$$

$$\leq 2^{-k/4},$$

where the third line follows because $i \notin B$, and so $\Pr[X \in T_i'] \geq \epsilon/n^2$. The fourth line follows because $|\hat{k}_i(x) - k_i(x)| \leq 1$. The fifth line follows because for $x \in T_i'$, we have that $k_i(x) \geq k/2$. The last line follows because $k = n^\alpha$ and $\epsilon = 2^{-k/100}$. This gives that $H_\infty(X[1, \ldots, i] \mid S(X) = i) \geq \frac{k}{4}$.

It remains to analyze $H_i(X \mid X[1, \ldots, X_i] = x[1, \ldots, x_i] \wedge S(X) = i)$. For this purpose, we compute:

$$\Pr[X = x \mid X[1, \ldots, i] = x[1, \ldots, i] \wedge S(X) = i] = \frac{\Pr[X = x \wedge X[1, \ldots, i] = x[1, \ldots, i] \wedge S(X) = i]}{\Pr[X[1, \ldots, i] = x[1, \ldots, i] \wedge S(X) = i]}$$

$$\leq \frac{\Pr[X = x]}{\Pr[X[1, \ldots, i] = x[1, \ldots, i] \wedge X \in T_i']}$$

$$= \frac{\Pr[X = x]}{\Pr[X[1, \ldots, i] = x[1, \ldots, i]]}$$

$$= \frac{2^{-k}}{2^{-k_i(x)}}$$

$$\leq 2^{-(k - \hat{k}_i(x) - 1)}$$

$$\leq 2^{-(k - (k/2 + b + 2\log n + \log(1/\epsilon)) - 1)}$$

$$\leq 2^{-k/4},$$

where the third line follows because for $x \in T_i'$, $\{X[1, \ldots, i] = x[1, \ldots, i]\} \subseteq \{X \in T_i'\}$. More specifically, this is because whether or not $X \in T_i'$ is determined by $X[1, \ldots, i]$ and as $X[1, \ldots, i] = x[1, \ldots, i]$ for $x \in T_i'$, it follows that $X[1, \ldots, i] = x[1, \ldots, i]$ implies $X \in T_i'$. The fourth line follows because $|\hat{k}_i(x) - k_i(x)| \leq 1$. The fifth line follows because for x in T_i', $i(x) = i$, and x is useful, which gives that $\hat{k}_i \leq k/2 + b + 2\log n + \log(1/\epsilon)$. The last line follows because $k = n^\alpha$, $b = k/10$ and $\epsilon = 2^{-k/100}$.

This gives that $H_\infty(X \mid X[1, \ldots, i] = x[1, \ldots, x_i] \wedge S(X) = i) \geq k/4$, and overall, we have that $((X[1, \ldots, i], X) \mid S(X) = i)$ is a $(\frac{k}{4}, \frac{k}{4})$-block-wise source. $\square$

The construction of BCnd: Let $a > 1$ be a large constant to be chosen later. We are assuming that E is hard for exponential size Σ_{j+5}-circuits, and will apply Lemma 3.2, with the following choices $c^{(3.2)} = a \cdot c$, $j^{(3.2)} = j + 1$, $\nu = \alpha/2$, and the m specified in Theorem 3.9 (which indeed satisfies that $(\log n)^{c_0} \leq m \leq n^{\nu/2}$, as required by Lemma 3.2). We obtain an oracle procedure $\mathsf{OCnd}^{(\cdot)} : \{0, 1\}^n \to \{0, 1\}^m$ such that for every $t \leq n$ and every function $B : \{0, 1\}^n \to \{0, 1\}^t$ that has a circuit of size $n^{a \cdot c}$, OCnd^B can be computed in time $n^{d^{(3.2)}}$ (for some constant $d^{(3.2)}$). Furthermore, setting $k' = m^{\delta_0}$, we obtain that OCnd^B is a $(n^{\alpha/2}, k')$-errorless condenser for all $(n^{a \cdot c}, j + 1)$-samplable distributions V over $\{0, 1\}^n$, that have the additional property that $(B(V), V)$ is a $(n^{\alpha/2}, n^{\alpha/2})$-block-wise source.

For every $i \in [\ell]$, we define the function $B_i(x) = x[1, \ldots, i]$. We now define $\mathsf{BCnd} : \{0,1\}^n \to \{0,1\}^{\ell m}$, by:

$$\mathsf{BCnd}(x) = \mathsf{OCnd}^{B_1}(x), \ldots, \mathsf{OCnd}^{B_\ell}(x).$$

(Note that the output length of BCnd is $\ell m = 10n^{1-\alpha} \cdot m$ as promised by Theorem 3.9). In order to prove Theorem 3.9, we need to show that $H_\infty(\mathsf{BCnd}(X)) \geq m^\delta$. We first observe that:

Claim 3.14. *For every $i \in [\ell]$ such that $i \notin B$, for $V_i = (X \mid S(X) = i)$ we have that $H_\infty(\mathsf{OCnd}^{B_i}(V_i)) \geq k'$.*

Proof. By Claim 3.13 we have that V_i is $(\mathsf{poly}(n^c), j+1)$-samplable, and that $(B_i(V_i), V_i)$ is a $(k/4, k/4)$-block-wise source. We can choose the constant a to be sufficiently large so that V_i is $(n^{a \cdot c}, j+1)$-samplable. Recall that $k/4 = n^\alpha/4 \geq n^{\alpha/2}$ and therefore, the function B_i and the distribution V_i meets the guarantee of OCnd. Therefore by the guarantee of OCnd, we indeed have that $H_\infty(\mathsf{OCnd}^{B_i}(V_i)) \geq k'$. $\qquad\square$

By Claim 3.13, X can be expressed as a convex combination of $(X \mid S(X) = 0)$ (which has a coefficient of at most $\epsilon = 2^{-n^\eta/100}$) and the distributions $(V_i)_{i \notin B}$.

We obtain that the distribution $\mathsf{BCnd}(X)$ is ϵ-close to a distribution with min-entropy at least k'. (This follows because a convex combination of distributions with min-entropy at least k', has min-entropy at least k').

This is not exactly what is guaranteed in Theorem 3.9, as we wanted BCnd to be errorless. Nevertheless, the error ϵ is smaller than $2^{-k'}$ and we can "swallow it" at the cost of slightly reducing k'. More precisely, we have that for every $z \in \{0,1\}^{\ell \cdot m}$,

$$\Pr[\mathsf{BCnd}(X) = z] \leq 2^{-k'} + \epsilon \leq 2^{-(k'-1)} \leq m^\delta,$$

where the first inequality holds because $\epsilon = 2^{-n^\alpha/100}$ and $k' = m^{\delta_0} \leq m \leq n^{\alpha/4}$, and the second inequality follows because $\delta = \delta_0/2$, so that $k' - 1 \geq m^\delta$. Overall, we conclude that indeed $H_\infty(\mathsf{BCnd}(X)) \geq m^\delta$. Furthermore, by construction, there exists a constant d (that depends on c, a and $d^{(3.2)}$) such that BCnd is computable in time n^d.

3.3 Proof of Theorem 3.1

In this section we Prove Theorem 3.1. The construction and proof closely follow the informal explanation given in Section 1.3.5.

3.3.1 The Construction of the Final Condenser

Setting up instantiations of the basic condenser. We are given constants $\alpha > 0$ and $0 < \xi \leq \alpha/10$, such that $\alpha + \xi < 1$. We will set the constant j as chosen in Theorem 3.1. We are assuming that E is hard for exponential size Σ_j-circuits, and given a constant c, we assume that n is sufficiently large, and aim to construct an errorless condenser for $(n^c, 0)$-samplable distributions X that have $H_\infty(X) \geq n^\alpha$.

We will now set up some additional parameters that will be used in the construction: Let $k = n^\alpha$. Let δ_0 be the constant from Theorem 3.9, and δ be the constant from Lemma 3.2. We set $\gamma = \xi/2$, $\lambda = \gamma \cdot \delta_0/2$, $\nu = \lambda/2 = \gamma \cdot \delta_0/4$ and $\epsilon = 2^{-n^\nu}$.

We plan to use the hardness assumption to set up a constant number instantiations (where the constant will depends on the constant α) of the basic condenser BCnd. For every $i \geq 0$, We plan to make each such instantiation BCnd_i to be a function $\mathsf{BCnd}_i : \{0,1\}^{n_i} \to \{0,1\}^{n_{i+1}}$, where $n_0 = n$, and $n_{i+1} \approx n_i^{1-\alpha}$. For this purpose, we make the following definitions. For every $i \geq 0$, we define:

- $n_0 = n$, and for $i \geq 0$, $n_{i+1} = n_i^{1-\alpha+2\gamma}$, which gives that: $n_i = n^{1-i\cdot(\alpha-2\gamma)}$.

- $k_0 = k$, and for $i \geq 0$, $k_{i+1} = k_i - n^\lambda$, which gives that $k_i = k - i \cdot n^\lambda$.

Let ℓ be the smallest integer, such that $n_{\ell+1} \leq n^{\alpha-\xi}$. By the definition of the sequence $\{n_i\}$, this holds for

$$\ell = \lceil \frac{1-\alpha-\xi}{\alpha-\xi} \rceil - 1.$$

We set $j = \ell + 5 = \lceil \frac{1-\alpha-\xi}{\alpha-\xi} \rceil + 4$ (as stated in Theorem 3.1). Using these choices, the construction of FCnd is specified in Figure 2.

3.3.2   Analysis of the Construction

We start by listing some properties that are maintained as an invariant throughout the iterative construction. We will then show that Theorem 3.1 directly follows from these properties.

In the lemma below, we will show that for every i there exists a "selector circuit" S_i which when given input $x \in \mathrm{Supp}(X)$ outputs an element in $\{E, W, R\}$. The intuition is that:

- If $S_i(x) = E$, then we think of x as an erroneous element which does not contribute to the success of FCnd, and our goal will be to maintain that the probability of these elements is small.

- If $S_i(x) = W$, then we think of x as an element on which FCnd already wins in one of the previous iterations. Here winning, means that conditioned on the event $\{S_i(X) = W\}$ the output of $\mathrm{FCnd}(X)$ (in fact even the first $i-1$ blocks) contain the required amount of min-entropy. Our hope is that eventually, we will win on all $x \in \mathrm{Supp}(X)$, except for the few erroneous ones.

- If $S_i(X) = R$, then we think of x as "remaining", and we have the property that conditioned on the event $\{S_i(X) = R\}$, the distribution $B_i(X)$ (which is the distribution that we are holding in the i'th step) has min-entropy at least $k/2$. In every iteration, the length of $B_i(x)$ (which is n_i) decreases. This means that the distribution that we are holding at this step is more condensed than the ones we held in previous iterations, and so we make progress, and eventually we will have to win.

The precise statement is given below.

Lemma 3.15. *Let X over $\{0,1\}^n$ be an $(n^c, 0)$-samplable distribution, such that $H_\infty(X) \geq k$. For every $0 \leq i \leq \ell + 1$ we have that:*

- *There exists a circuit $S_i : \{0,1\}^n \to \{E, W, R\}$ such that:*

  - *S_i is a Σ_i-circuit of size n^{c_i-1}.*
  - *The sets $T_i^E = \{x \in \mathrm{Supp}(X) : S_i(x) = E\}$, $T_i^W = \{x \in \mathrm{Supp}(X) : S_i(x) = W\}$, $T_i^R = \{x \in \mathrm{Supp}(X) : S_i(x) = R\}$ are a partition of $\mathrm{Supp}(X)$.*

- *$\Pr[S_i(X) = E] \leq i \cdot \epsilon$.*
- *If $\Pr[S_i(X) = R] > 0$, then $H_\infty(B_i(X) \mid S_i(X) = R) \geq k_i \geq k/2$.*
- *If $\Pr[S_i(X) = W] > 0$ then the distribution $Z = (\mathrm{FCnd}(X)_{0,\ldots,i-1} \mid S_i(X) = W)$ has $H_\infty(Z) \geq (m')^\delta$.*

The proof of Lemma 3.15 appears in Section 3.3.4. Before proving Lemma 3.15 we will show that Theorem 3.1 follows from Lemma 3.15.

29

Figure 2: Construction of function FCnd

Recall that $\gamma = \xi/2$, and set $m = n^\gamma$. Let $a \geq 1$ be a sufficiently large universal constant that will be chosen in the proof. We are assuming that n is sufficiently large, and will construct FCnd by the following iterative process:

Initialization: We define $c_0 = c + 1$, and a function $B_0(x) = x$.

Iterative step: Assume that for some $0 \leq i \leq \ell$ we have already defined: $c_0, \ldots, c_i$ and $B_0, \ldots, B_i$ (and note that this holds for $i = 0$). We will continue step i as follows:

Basic condenser: We apply Theorem 3.9, choosing parameters c_i and $j = i$ and as we are assuming that E is hard for $\Sigma_{\ell+5}$-circuits, and as $i \leq \ell$, we can conclude that there is a function

$$\mathsf{BCnd}_i : \{0,1\}^{n_i} \to \{0,1\}^{n_{i+1}},$$

such that:

- BCnd_i is a $(\frac{k}{2}, m^{\delta_0})$-errorless condenser for (n^{c_i}, i)-samplable distributions (and recall that $k_i \geq \frac{k}{2}$).
- Furthermore, there exists a constant $d_i \geq c_i$ that is determined by c_i such that BCnd_i can be computed in time n^{d_i}.

We now verify that the choices of $n_0, \ldots, n_{\ell+1}$ indeed meet the requirements of Theorem 3.9. This follows because for every $0 \leq i \leq \ell$, the output length of BCnd_i specified in Theorem 3.9 is

$$\frac{10 \cdot n_i \cdot m}{n^\alpha/2} = 20n^{1-i\cdot(\alpha-2\gamma)-\alpha+\gamma} \leq n^{1-(i+1)\cdot(\alpha-2\gamma)} = n_{i+1},$$

and we can artificially increase the output length of BCnd_i to n_{i+1}.

Block function: We define the function: $B_{i+1}(x) = \mathsf{BCnd}_i(B_i(x))$, and note that $B_{i+1} : \{0,1\}^n \to \{0,1\}^{n_{i+1}}$ can be computed in time $n^{d_i'}$, for some constant $d_i' \geq d_i \geq c_i$,

Oracle condenser: We will now apply lemma 3.2. Recall that we are assuming that E is hard for exponential size $\Sigma_{\ell+5}$-circuits. We use the constant $\nu = \gamma \cdot \delta_0/4$ chosen earlier, and will choose the parameter m from Lemma 3.2 to be $m' = n^{\nu/2}$. We plan to supply the function $B_{i+1} : \{0,1\}^n \to \{0,1\}^{n_{i+1}}$ as oracle to the oracle condenser. More specifically, we set $o_i = a \cdot d_i'$ and apply Lemma 3.2 choosing c to be o_i, and j to be $i + 1$. We obtain an oracle procedure

$$\mathsf{OCnd}_i^{(\cdot)} : \{0,1\}^n \to \{0,1\}^{m'=n^{\nu/2}},$$

such that:

- $\mathsf{OCnd}_i^{B_{i+1}}$ is an $(n^\nu, (m')^\delta = n^{\frac{\nu \cdot \delta}{2}})$-errorless condenser for all $(n^{a \cdot d_i'}, i+1)$-samplable distributions V_i such that $(B_{i+1}(V_i), V_i)$ is an (n^ν, n^ν)-block-wise source.
- $\mathsf{OCnd}_i^{B_{i+1}}$ can be computed in time $n^{d_i''}$ for some constant d_i'' that is determined as a function of d_i' and the universal constant a.

Setting up the next iteration: At this point, we choose $c_{i+1} = a \cdot d_i'$, and continue to the next iterative step. We stop this process when we complete step $i = \ell$, and note that at the end of step ℓ, $B_{\ell+1}$ is defined.

Construction of FCnd: We define $\mathsf{FCnd} : \{0,1\}^n \to \{0,1\}^{(\ell+1)\cdot m'}$ as follows:

$$\mathsf{FCnd}(x) = \mathsf{OCnd}_0^{B_1}(x), \ldots, \mathsf{OCnd}_\ell^{B_{\ell+1}}(x).$$

By construction, the output length of FCnd is $(\ell+1) \cdot m' = (\ell+1) \cdot n^{\nu/2} \leq n^{\delta_2}$ for some constant $0 < \delta_2 \leq \frac{\alpha}{10}$ (as required) and there exists a constant d, such that FCnd can be computed in time n^d.

3.3.3 Showing that Theorem 3.1 Follows from Lemma 3.15

Let X be an $(n^c, 0)$-samplable distribution, such that $H_\infty(X) \geq k$. Our goal is to show that $H_\infty(\mathsf{FCnd}(X)) \geq n^{\delta_1}$, for some constant $\delta_1 > 0$. We will show that after the final step no elements remain.

Claim 3.16. $\Pr[S_{\ell+1}(X) = R] = 0$.

Proof. This holds because ℓ was chosen so that $n_{\ell+1}$ (which is the length of the string $B_{\ell+1}(X)$) is at most $n^{\alpha-\xi}$, which for sufficiently large n is less than $n^\alpha/4 = k/4$.

By Lemma 3.15 (choosing $i = \ell + 1$), if $\Pr[S_{\ell+1}(X) = R] > 0$, then $H_\infty(B_{\ell+1}(X) \mid S_{\ell+1}(X) = R) \geq k/2$. However, the latter cannot hold because the bit length of $B_{\ell+1}(X)$ is smaller than $k/4$ and the min-entropy of a random variable cannot exceed its length. $\square$

By Lemma 3.15 we also have that the sets $T_{\ell+1}^E, T_{\ell+1}^W, T_{\ell+1}^R$ are a partition of $\mathrm{Supp}(X)$, and that

$$\Pr[S_{\ell+1}(X) = E] \leq (\ell + 1) \cdot \epsilon.$$

We can therefore conclude that $\Pr[S_{\ell+1}(X) = W] \geq 1 - (\ell + 1) \cdot \epsilon$.

Lemma 3.15 (for $i = \ell + 1$) also gives that the distribution

$$Z = (\mathsf{FCnd}(X)_{0,\ldots,\ell} \mid S_{\ell+1}(X) = W) = (\mathsf{FCnd}(X) \mid S_{\ell+1}(X) = W)$$

has $H_\infty(Z) \geq (m')^\delta$.

This gives that for every possible output string $z \in \{0,1\}^{(\ell+1)\cdot m'}$ of FCnd,

$$\begin{aligned}
\Pr[\mathsf{FCnd}(X) = z] &= \Pr[\mathsf{FCnd}(X) = z \mid S_{\ell+1}(X) = W] \cdot \Pr[S_{\ell+1}(X) = W] \\
&\quad + \Pr[\mathsf{FCnd}(X) = z \mid S_{\ell+1}(X) = E] \cdot \Pr[S_{\ell+1}(X) = E] \\
&\leq \Pr[\mathsf{FCnd}(X) = z \mid S_{\ell+1}(X) = W] + \Pr[S_{\ell+1}(X) = E] \\
&\leq 2^{-(m')^\delta} + (\ell + 1) \cdot \epsilon. \\
&\leq 2^{-(m')^{\delta/2}}
\end{aligned}$$

where the last inequality follows because $m' = n^{\nu/2}$, ℓ is constant, and $\epsilon = 2^{-n^\nu}$.

Recall that $m' = n^{\nu/2}$ We set $\delta_1 = \delta \cdot \nu/4$, and note that we indeed have that

$$H_\infty(\mathsf{FCnd}(X)) \geq (m')^{\delta/2} = n^{\delta_1},$$

as required.

3.3.4 Proof of Lemma 3.15

We have that X is $(n^c, 0)$-samplable. Let $A : \{0,1\}^r \to \{0,1\}^n$ and $P : \{0,1\}^r \to \{0,1\}$ be size n^c circuits, such that X is sampled by A with postselection by P.

We will prove Lemma 3.15 by induction on i. For the base case where $i = 0$, we define the circuit $S_0(x)$ to be the constant function that answers R. This is indeed a Σ_0-circuit of size n. With this choice $(B_0(X) \mid S_0(X) = R)$ is the distribution X, which has $H_\infty(X) \geq k_0 = k$ and the invariant holds.

We now assume that the invariant holds for some $0 \leq i \leq \ell$, and will show that it holds for $i + 1$.

If $\Pr[S_i(X) = R] = 0$, then we define $S_{i+1} = S_i$, and we are done. Therefore, we will assume that $\Pr[S_i(X) = R] > 0$.

Showing that $(B_i(X) \mid S_i(X) = R)$ **meets the requirements of** BCnd_i. The distribution $(B_i(X) \mid S_i(X) = R)$ is (n^{c_i}, i)-samplable. This is because it is sampled by the circuit $A_i(w) = B_i(A(w))$ with postselection by the circuit $P_i(w)$ which answers one iff $P(w) = 1$ and $S_i(W) = R$. Note that for $i = 0$, A_0 has size $n^c \leq n^{c_0}$, and P_0 is a Σ_0-circuit of size $O(n^c) \leq n^{c+1} = n^{c_0}$. For $i \geq 0$, A_i has size $n^{d'_{i-1}} + n^c \leq n^{c_i}$, where this holds because $c_i = a \cdot d'_{i-1}$ for a sufficiently large constant $a \geq 1$. We also have that P_i is a Σ_i-circuit of size $n^c + n^{c_i - 1} \leq n^{c_i}$, where this holds because for every $i \geq 0$, $c_i \geq c + 1$.

By the invariant we have that $H_\infty(B_i(X) \mid S_i(X) = R) \geq k_i \geq k/2$. Therefore, the distribution $V_i = (B_i(X) \mid S_i(X) = R)$ meets the requirements of BCnd_i, and we can conclude that $H_\infty(\mathsf{BCnd}_i(V_i)) \geq m^{\delta_0}$, which gives that:

$$H_\infty(B_{i+1}(X) \mid S_i(X) = R) = H_\infty(\mathsf{BCnd}_i(B_i(X)) \mid S_i(X) = R) \geq H_\infty(\mathsf{BCnd}_i(V_i)) \geq m^{\delta_0}.$$

Constructing the circuit S_{i+1}. For every $x \in T_i^R$, we define:

$$k_i(x) = -\log \Pr[B_{i+1}(X) = B_{i+1}(x) \mid S_i(X) = R],$$

so that $\Pr[B_{i+1}(X) = B_{i+1}(x) \mid S_i(X) = R] = 2^{-k_i(x)}$. Note that we have just seen that for every $x \in T_i^R$, $H_\infty(B_{i+1}(X) \mid S_i(X) = R) \geq m^{\delta_0}$, which gives that $k_i(x) \geq m^{\delta_0}$.

Our next step is to show that there is a Σ_{i+1}-circuit that approximates $k_i(x)$.

Claim 3.17. *There is a Σ_{i+1}-circuit R of size $\mathsf{poly}(n^{d'_i})$ which given input $x \in T_i^R$, outputs a number $\hat{k}_i(x)$ such that: $|\hat{k}_i(x) - k_i(x)| \leq 1$.*

Proof. We have that $(B_{i+1}(X) \mid S_i(X) = R)$ is sampled by $A' = B_{i+1} \circ A$ with postselection by the Σ_i-circuit $P'(x)$ that answers one iff $P(x) = 1 \wedge S_i(X) = R$. When the circuit R receives an input $x \in \{0,1\}^n$, it will construct the circuit $A_x(w)$ which answers one iff $A'(w) = B_{i+1}(x) \wedge P'(w) = 1$. Note that:

$$\begin{aligned}
2^{-k_i(x)} &= \Pr[B_{i+1}(X) = B_{i+1}(x) \mid S_i(X) = R] \\
&= \Pr_{W \leftarrow U_r}[A'(W) = B_{i+1}(x) \mid P'(W) = 1] \\
&= \frac{\Pr_{W \leftarrow U_r}[A'(W) = B_{i+1}(x) \wedge P'(W) = 1]}{\Pr_{W \leftarrow U_r}[P'(W) = 1]} \\
&= \frac{\Pr_{W \leftarrow U_r}[A_x(W) = 1]}{\Pr_{W \leftarrow U_r}[P'(W) = 1]}.
\end{aligned}$$

By Theorem 2.24, the circuit R (which will be a Σ_{i+1}-circuit of size $\mathsf{poly}(n^{d'_i})$ can compute β-relative approximations to both the enumerator and denominator, for say constant $\beta = \frac{1}{100}$. R can then compute the division, which is guaranteed to be a $1/10$-relative approximation of $2^{-k_i(x)}$, and using this approximation, R can output the required approximation $\hat{k}_i(x)$. $\square$

We will partition the set T_i^R into two sets as follows:

$$M_i = \left\{ x \in T_i^R : \hat{k}_i(x) \leq k_i - \frac{n^\lambda}{2} \right\}$$

$$L_i = \left\{ x \in T_i^R : \hat{k}_i(x) > k_i - \frac{n^\lambda}{2} \right\}$$

We now define the Σ_{i+1} circuit S_{i+1} that we need to maintain the invariant for $i + 1$.

Claim 3.18. *There is a Σ_{i+1}-circuit S_{i+1} of size $\mathsf{poly}(n^{d'_i})$ which given input $x \in \mathrm{Supp}(X)$, outputs an element in $\{E, W, R\}$ such that:*

- S_i is a Σ_{i+1}-circuit of size $\mathsf{poly}(n^{d_i'})$.
- The sets $T_{i+1}^E = \{x \in \mathrm{Supp}(X) : S_{i+1}(x) = E\}$, $T_{i+1}^W = \{x \in \mathrm{Supp}(X) : S_{i+1}(x) = W\}$, $T_{i+1}^R = \{x \in \mathrm{Supp}(X) : S_{i+1}(x) = R\}$ are a partition of $\mathrm{Supp}(X)$.
- If $S_i(x) \in \{E, W\}$ then $S_{i+1}(x) = S_i(x)$.
- If $S_i(x) = R$ then

   - If $x \in M_i$ and $\Pr[X \in M_i \mid S_i(X) = R] \geq \epsilon$ then $S_{i+1}(x) = W$.
   - If $x \in L_i$ and $\Pr[X \in L_i \mid S_i(X) = R] \geq \epsilon$ then $S_{i+1}(x) = R$.
   - Otherwise, $S_{i+1}(x) = E$, and in particular $\Pr[S_{i+1}(X) = E \mid S_i(X) = R] \leq \epsilon$.

Proof. The circuit S_{i+1} will be hardwired with two advice bits M, L, where $M = 1$ iff $\Pr[X \in M_i \mid S_i(X) = 1] < \epsilon$, and $L = 1$ iff $\Pr[X \in L_i \mid S_i(X) = 1] < \epsilon$. When the circuit S_{i+1} receives input x, it first runs $S_i(x)$ and answers the same answer if $S_i(x) \in \{E, W\}$. If $S_i(x) = R$, then S_{i+1} uses the circuit R to compute $\hat{k}_i(x)$ and determine whether $x \in M_i$ or $x \in L_i$. If $x \in M_i$ and $M = 0$, S_{i+1} answers W. If $x \in L_i$ and $L = 0$, S_{i+1} answers R. Otherwise, S_{i+1} answers E.

The correctness of S_{i+1} follows by construction, and by the size guarantees on S_i and R, S_{i+1} is a Σ_{i+1}-circuit of size $\mathsf{poly}(n^{d_i'}, n^{c_i-1}) = \mathsf{poly}(n^{d_i'})$, where the later follows vecause $d_i \geq c_{i-1}$. $\qquad\square$

Handling the first item in the invariant. Recall that in Figure 2, we have chosen $c_{i+1} = a \cdot d_i'$. By choosing a to to be sufficiently large we can make sure that the size of S_{i+1}, which is $n^{\mathsf{poly}(d_i')}$ is smaller than $n^{c_{i+1}-1} = n^{a \cdot d_i'-1}$. This gives that S_{i+1} is a Σ_{i+1}-circuit of size $n^{c_{i+1}-1}$ and the circuit S_{i+1} indeed meets the first item in the invariant of Lemma 3.15.

Handling the second item in the invariant. We note that by Claim 3.18,

$$
\begin{aligned}
\Pr[S_{i+1}(X) = E] &= \sum_{y \in \{W,R,E\}} \Pr[S_{i+1}(X) = E \mid S_i(X) = y] \cdot \Pr[S_i(X) = y] \\
&\leq \Pr[S_i(X) = E] + \Pr[S_{i+1}(X) = E \mid S_i(X) = R] \\
&\leq i \cdot \epsilon + \epsilon \\
&= (i+1) \cdot \epsilon.
\end{aligned}
$$

Handling the third item in the invariant. We assume that $\Pr[S_{i+1}(X) = R] > 0$. We need to show that $H_\infty(B_{i+1}(X) \mid S_{i+1}(X) = R) \geq k_{i+1}$.

For this purpose, let $x \in \mathrm{Supp}(X)$ be such that $S_{i+1}(x) = R$, and observe that:

$$
\begin{aligned}
\Pr[B_{i+1}(X) = B_{i+1}(x) \mid S_{i+1}(X) = R] &= \Pr[B_{i+1}(X) = B_{i+1}(x) \mid S_i(X) = R \wedge S_{i+1}(X) = R] \\
&\leq \frac{\Pr[B_{i+1}(X) = B_{i+1}(x) \mid S_i(X) = R]}{\Pr[S_{i+1}(X) = R \mid S_i(X) = R]} \\
&\leq \frac{2^{-k_i(x)}}{\epsilon} \\
&\leq 2^{-(\hat{k}_i(x)-1-\log(1/\epsilon))} \\
&\leq 2^{-(k_i-\frac{n^\lambda}{2}-1-\log(1/\epsilon))}.
\end{aligned}
$$

The first line follows because $\{S_{i+1}(X) = R\} \subseteq \{S_i(X) = R\}$. The second line follows because for every events A, B, C, $\Pr[A \mid B \cap C] \leq \frac{\Pr[A|B]}{\Pr[C|B]}$. The fifth lines follows because we have that $S_{i+1}(x) = R$, and this implies that $x \in L_i$ and $\Pr[S_{i+1}(X) = R \mid S_i(X) = R] = \Pr[X \in L_i \mid S_i(X) = R] \geq \epsilon$.

33

Overall, we conclude that the third item holds, as:

$$H_\infty(B_{i+1}(X) \mid S_{i+1}(X) = R) \geq k_i - \frac{n^\lambda}{2} - 1 - \log(1/\epsilon) \geq k_i - \frac{n^\lambda}{2} - 1 - n^\nu \geq k_i - n^\lambda = k_{i+1}.$$

This follows because $\epsilon = 2^{-n^\nu}$, $\nu = \gamma \cdot \delta_0/4$ and $\lambda = \gamma \cdot \delta_0/2$.

Handling the fourth item in the invariant. We assume that $\Pr[S_{i+1}(X) = W] > 0$. We need to show that the distribution $Z_{i+1} = (\mathsf{FCnd}(X)_{0,\dots,i} \mid S_{i+1}(X) = W)$ has $H_\infty(Z_{i+1}) \geq (m')^\delta$. By the invariant, we have that the distribution $Z_i = (\mathsf{FCnd}(X)_{0,\dots,i-1} \mid S_i(X) = W)$ has $H_\infty(Z_i) \geq (m')^\delta$, which implies that $H_\infty((\mathsf{FCnd}(X)_{0,\dots,i} \mid S_i(X) = W)) \geq (m')^\delta$. Therefore, it is sufficient to show that the distribution $Z'_i = (\mathsf{FCnd}(X)_{0,\dots,i} \mid S_i(X) = R \wedge S_{i+1}(X) = W)$ has $H_\infty(Z'_i) \geq (m')^\delta$. This is because the distribution Z_{i+1} is a convex combination of Z_i and Z'_i.

In order to prove that $H_\infty(Z'_i) \geq (m')^\delta$, using the fact that $\mathsf{FCnd}_i(x) = \mathsf{OCnd}^{B_{i+1}}(x)$, it is sufficient to prove that:

$$H_\infty(\mathsf{OCnd}_i^{B_{i+1}}(X) \mid S_i(X) = R \wedge S_{i+1}(X) = W) \geq (m')^\delta,$$

as this will imply that:

$$H_\infty(\mathsf{FCnd}_i(X) \mid S_i(X) = R \wedge S_{i+1}(X) = W) \geq (m')^\delta.$$

We now turn our attention to proving that:

$$H_\infty(\mathsf{OCnd}_i^{B_{i+1}}(X) \mid S_i(X) = R \wedge S_{i+1}(X) = W) \geq (m')^\delta,$$

This will immediately follow (by the choices of parameters for OCnd made in Figure 2) if we show that for $V_i = (X \mid S_i(X) = R \wedge S_{i+1}(X) = W)$ we have that

- V_i is $(n^{a \cdot d'_i}, i+1)$-samplable, and
- $(B_{i+1}(V_i), V_i)$ is an (n^ν, n^ν)-block-wise source.

The distribution $V_i = (X \mid S_i(X) = R \wedge S_{i+1}(X) = W)$ is obviously samplable by the circuit A, with postelection by a circuit $P'(w)$ that answers one iff $P'(w) = 1 \wedge S_i(w) = R \wedge S_{i+1}(w) = W$, and this gives that it is $\mathsf{poly}(n^{d'_i}, i+1)$-samplable, and by choosing the constant a to be sufficiently large, the first requirement holds.

Consequently it remains to prove the second requirement. We start by computing $H_\infty(B_{i+1}(X) \mid S_i(X) = R \wedge S_{i+1}(X) = W)$. For this purpose, let $x \in \mathrm{Supp}(X)$ be such that $S_i(x) = R$ and $S_{i+1}(x) = w$.

$$
\begin{aligned}
\Pr[B_{i+1}(X) = B_{i+1}(x) \mid S_i(X) = R \wedge S_{i+1}(X) = W] &\leq \frac{\Pr[B_{i+1}(X) = B_{i+1}(x) \mid S_i(X) = R]}{\Pr[S_{i+1}(X) = W \mid S_i(X) = R]} \\
&\leq \frac{2^{-m^{\delta_0}}}{\epsilon} \\
&\leq 2^{-(m^{\delta_0} - \log(1/\epsilon))}.
\end{aligned}
$$

The first line follows because for every events A, B, C, $\Pr[A \mid B \cap C] \leq \frac{\Pr[A|B]}{\Pr[C|B]}$. The second line follows as we have already seen that $H_\infty(B_{i+1}(X) \mid S_i(X) = R) \geq m^{\delta_0}$. We conclude that:

$$H_\infty(B_i(V_i)) = H_\infty(B_i(X) \mid S_i(X) = R \wedge S_{i+1}(X) = W) \geq m^{\delta_0} - \log(1/\epsilon).$$

We now compute $H_\infty(X \mid B_{i+1}(X) = B_{i+1}(x) \wedge S_i(X) = R \wedge S_{i+1}(X) = W)$. For this purpose, let $x \in \mathrm{Supp}(X)$ be such that $S_i(x) = R$ and $S_{i+1}(x) = w$. Let

$$p = \Pr[X = x \mid B_{i+1}(X) = B_{i+1}(x) \wedge S_i(X) = R \wedge S_{i+1}(X) = W].$$

We compute:

$$
\begin{aligned}
p &\leq \frac{\Pr[X = x \mid S_i(X) = R]}{\Pr[B_{i+1}(X) = B_{i+1}(x) \wedge S_{i+1}(X) = W \mid S_i(X) = R]} \\
&= \frac{\Pr[X = x \mid S_i(X) = R]}{\Pr[B_{i+1}(X) = B_{i+1}(x) \mid S_i(X) = R]} \\
&\leq \frac{2^{-k_i}}{2^{-\hat{k}_i(x)}} \\
&\leq 2^{-(k_i - \hat{k}_i(x) - 1)} \\
&\leq 2^{-(k_i - (k_i - \frac{n^\lambda}{2}) - 1)} \\
&\leq 2^{-(\frac{n^\lambda}{2} - 1)}.
\end{aligned}
$$

The first line follows because for every events A, B, C, $\Pr[A \mid B \cap C] \leq \frac{\Pr[A|B]}{\Pr[C|B]}$. The second line follows because conditioned on $\{S_i(X) = R\}$, $\{B_{i+1}(X) = B_{i+1}(x)\} \subseteq \{S_{i+1}(X) = w\}$. More formally, knowing that $B_{i+1}(X) = B_{i+1}(x)$ and $S_i(X) = R$, and that there exists x such that $S_i(x) = R$, and $S_{i+1}(x) = W$, we can conclude that $S_{i+1}(X) = W$. This is because knowing that $S_i(X) = R$, whether or not $X \in M_i$ is determined by $B_{i+1}(X)$, and as we have that $B_{i+1}(X) = B_{i+1}(x)$, we can conclude that $X \in M_i$. We also know that there exists x, such that $S_i(x) = R$ and $S_{i+1}(x) = W$, and this implies that every $x \in M_i$ satisfies $S_{i+1}(x) = W$. We can therefore conclude that $S_{i+1}(X) = W$.

The third line follows because we have that $H_\infty(X \mid S_i(X) = R) \geq H_\infty(B_i(X) \mid S_i(X) = R) \geq k_i$. The fifth line follows because $S_i(x) = R$ and $S_{i+1}(x) = W$ implies $x \in M_i$.

Overall, we conclude that

$$H_\infty(X \mid B_{i+1}(X) = B_{i+1}(x) \wedge S_i(X) = R \wedge S_{i+1}(X) = W) \geq \frac{n^\lambda}{2} - 1,$$

and we have that $(B_{i+1}(V_i), V_i)$ is an $(m^{\delta_0} - \log(1/\epsilon), \frac{n^\lambda}{2} - 1)$-block-wise source. We have that:

$$m^{\delta_0} - \log(1/\epsilon) = n^{\gamma \cdot \delta_0} - n^\nu = n^{4\nu} - n^\nu \geq n^\nu,$$

where this follows because $m = n^\gamma$ and $\nu = n^{\frac{\gamma \cdot \delta_0}{4}}$. We also have that:

$$\frac{n^\lambda}{2} - 1 = n^{\frac{\gamma \cdot \delta_0}{2}} - 1 \geq n^\nu.$$

Overall, we indeed conclude that $(B_{i+1}(V_i), V_i)$ is an (n^ν, n^ν)-block-wise source.

4 A Construction of an Extractor for Samplable Distributions

In this section we use the errorless condenser of Section 3 to construct an extractor for samplable distributions. This is done by adapting a reduction of Ball, Shaltiel and Silbak [BSS25] (as we explained in Section 1.3.3). (We cannot use the formal statement of the reduction in [BSS25] as the parameters and components are somewhat different. However, the argument is essentially the same as the one given in [BSS25]).

The first step is to construct an extractor with small output length $m = O(\log n)$. This extractor is stated in Theorem 4.1 below.

Theorem 4.1 (extractor for with small output length). *For every constants $\alpha > 0$ and $0 < \xi \leq \frac{\alpha}{10}$, such that $\alpha + \xi < 1$, there exists a constant $j = \lceil \frac{1-\alpha-\xi}{\alpha-\xi} \rceil + 4$, such that if E is hard for exponential size Σ_j-circuits, then for every constants $c, c' > 1$, there exists a constant d such that for every sufficiently large n, there is a function $\mathsf{Ext} : \{0,1\}^n \rightarrow \{0,1\}^{c' \cdot \log n}$ that is an $(n^\alpha, \frac{1}{n^{c'}})$-extractor for $(n^c, 0)$-samplable distributions. Furthermore, Ext can be computed in time $\mathsf{poly}(n^d)$.*

By the same calculation made in the beginning of Section 3, for every $i \geq 2$, and $\alpha > \frac{1}{i}$, the j stated in Theorem 4.1 satisfies $j \leq i + 3$, matching the choice made in Theorems 1.4 and 1.6.

The proof of Theorem 4.1 is given in Section 4.1. In Section 4.2 we explain how to increase the output length of the extractor in Theorem 4.1 and prove Theorems 1.4 and 1.6. Finally, in Section 4.3 we show how to extend the proof of Theorem 4.1, so that it gives the multiplicative extractor of Theorem 1.8

4.1 Proof of Theorem 4.1

The proof of Theorem 4.1 essentially repeats the approach of Ball, Shaltiel, and Silbak [BSS25], using the specific choices of this paper. The proof is also quite similar to the proof of Lemma 3.2 (which is also based on the approach of [BSS25], as explained in Section 1.3.3).

The construction of Ext appears in Figure 3. Let n be sufficiently large, let $\epsilon = n^{-c'}$ and assume for the purpose of contradiction that there is an $(n^c, 0)$-samplable distribution X such that $H_\infty(X) \geq n^\alpha$, and $\mathsf{Ext}(X)$ is not ϵ-close to uniform, and in particular that there exists a set $T \subseteq \{0,1\}^m$ such that:

$$\Pr[\mathsf{TExt}(g(\mathsf{FCnd}(X)), X) \in T] > \frac{|T|}{2^m} + \epsilon.$$

In a similar manner to the proof of Lemma 3.2, our goal will be to obtain a contradiction by showing that there is a Σ_2-circuit C of size $n^{\bar{c}}$, and an $(n^{\bar{c}}, 0)$-samplable distribution Y (we will choose $Y = \mathsf{FCnd}(X)$) with $H_\infty(Y) \geq n^\nu$, on which $\Pr[C(Y) = g(Y)]$ is too large, and contradicts the HOS guarantee of g.

We have that:

$$\Pr[\mathsf{TExt}(g(\mathsf{FCnd}(X)), X) \in T] > \frac{|T|}{2^m} + \epsilon \geq \frac{|T|}{2^m} \cdot (1 + \epsilon).$$

By an averaging argument, there exists a $z \in T$ such that:

$$\Pr[\mathsf{TExt}(g(\mathsf{FCnd}(X)), X) = z] > (1 + \epsilon) \cdot 2^{-m}.$$

Let $t = n^{\delta_2}$. We will say that $y \in \{0,1\}^t$ is *useful* if the following two conditions hold:

- $\Pr[\mathsf{TExt}(g(y), X) = z \mid \mathsf{FCnd}(X) = y] > (1 + \frac{\epsilon}{2}) \cdot 2^{-m}.$
- $H_\infty(X \mid \mathsf{FCnd}(X) = y) \geq k'.$

The claim below is similar in spirit to Claim 3.3 in the proof of Lemma 3.2.

Claim 4.2. $\Pr[\mathsf{FCnd}(X) \text{ is useful}] > \frac{\epsilon \cdot 2^{-m}}{4}.$

Proof. We define:

$$B_1 = \left\{ y \in \{0,1\}^t : \Pr[\mathsf{TExt}(g(y), X) = z \mid \mathsf{FCnd}(X) = y] \leq (1 + \frac{\epsilon}{2}) \cdot 2^{-m} \right\}$$
$$B_2 = \left\{ y : H_\infty(X \mid \mathsf{FCnd}(X) = y) < k' \right\}$$

This is done so that y is useful iff $y \notin B_1 \cup B_2$. We will proceed to bound $\Pr[\mathsf{FCnd}(X) \in B_1]$ and $\Pr[\mathsf{FCnd}(X) \in B_2]$, separately.

Figure 3: Construction of function Ext

Parameters:

- Given constants $\alpha > 0$ and $0 < \xi < \frac{\alpha}{10}$, such that $\alpha + \xi < 1$, we set $j = \lceil \frac{1-\alpha-\xi}{\alpha-\xi} \rceil + 4$.
- We are given constants $c, c' > 1$.
- Let $a \geq 1$ be a sufficiently large constant to be chosen in the proof.
- We are assuming that n is sufficiently large.

Assumption: We are assuming that E is hard for exponential size Σ_j-circuits.

Ingredients: We will require the following ingredients:

- *Errorless condenser:* By Theorem 3.1, under the hardness assumption, there exist constants $0 < \delta_1 < \delta_2 \leq \frac{\alpha}{10}$, and a function $\mathsf{FCnd} : \{0,1\}^n \to \{0,1\}^{n^{\delta_2}}$ that is an (n^α, n^{δ_1})-errorless condenser for $(n^c, 0)$-samplable distributions. Furthermore, there is a constant $d_{\mathsf{FCnd}} \geq 1$ such that FCnd can be computed in time $n^{d_{\mathsf{FCnd}}}$.
- *HOS:* Let $\bar{c} > d_{\mathsf{FCnd}}$ be a sufficiently large constant that will be chosen in the proof. By Theorem 2.22, setting $\nu = \delta_1$, under the hardness assumption, there is a function $g : \{0,1\}^n \to \{0,1\}^{n^{a_0}}$ (where $a_0 > 1$ is a universal constant) that is an $(n^\nu, 2^{-\Omega(n^\nu)})$-HOS for the class of $(n^{\bar{c}}, 0)$-samplable distributions, against the class of size $n^{\bar{c}}$ Σ_2-circuits. We have that g can be computed in time n^{d_g} for some constant d_g that depends on $\bar{c}$.
- *2-source extractor:* Let $k' = \frac{n^\nu}{a} = \frac{n^{\delta_1}}{a}$, $\epsilon' = \frac{n^{-2c'}}{16}$, and $m = c' \cdot \log n$ and let $\mathsf{TExt} : \{0,1\}^{n^{a_0}} \times \{0,1\}^{n^{a_0}} \to \{0,1\}^m$ be the (k', k', ϵ')-2-source extractor that is guaranteed by Theorem 2.14. We have that TExt can be computed in time $n^{d_{\mathsf{TExt}}}$ for a constant d_{TExt} that depends on c'.

Construction: We define $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$, as follows:

$$\mathsf{Ext}(x) = \mathsf{TExt}(g(\mathsf{FCnd}(x)), x).$$

Note that g expects inputs of length n, and TExt expects inputs of length n^{a_0}, and so:

- We pad $\mathsf{FCnd}(x)$ with $n - n^{\delta_2}$ zeros, before applying g on $B(x)$.
- We pad x with $n^{a_0} - n$ zeros, before giving it as a second input to TExt.

Note that by construction, there indeed exists a constant d such that Ext can be computed in time n^d.

By Lemma 2.3 with probability at least $1 - \frac{\epsilon \cdot 2^{-m}}{4}$ over choosing $y \leftarrow \mathsf{FCnd}(X)$, we have that

$$H_\infty(X \mid \mathsf{FCnd}(X) = y) \geq n^\alpha - n^{\delta_2} - \log(1/\epsilon) - m - 2 \geq n^\alpha - n^{\frac{\alpha}{10}} - 2c' \cdot \log n - 2 \geq n^{\frac{\alpha}{10}} \geq n^\nu \geq k',$$

where this follows because $k' = n^\nu/a$ and $\nu = \delta_1 \leq \frac{\alpha}{10}$. This gives that $\Pr[\mathsf{FCnd}(X) \in B_2] \leq \frac{\epsilon \cdot 2^{-m}}{4}$.

We claim that $\Pr[\mathsf{FCnd}(X) \in B_1] \leq 1 - \frac{\epsilon \cdot 2^{-m}}{2}$. This follows as otherwise, $\Pr[\mathsf{FCnd}(X) \notin B_1] \leq \frac{\epsilon \cdot 2^{-m}}{2}$,

which implies:

$$\Pr[\mathsf{TExt}(g(\mathsf{FCnd}(X)), X) = z] \leq \Pr[\mathsf{FCnd}(X) \notin B_1] + \Pr[\mathsf{TExt}(g(\mathsf{FCnd}(X)), X) = z \cap \mathsf{FCnd}(X) \in B_1]$$

$$\leq \frac{\epsilon \cdot 2^{-m}}{2} + \sum_{y \in B_1} \Pr[\mathsf{TExt}(g(\mathsf{FCnd}(X)), X) = z \cap \mathsf{FCnd}(X) = y]$$

$$= \frac{\epsilon \cdot 2^{-m}}{2} + \sum_{y \in B_1} \Pr[\mathsf{TExt}(g(\mathsf{FCnd}(X)), X) = z \mid \mathsf{FCnd}(X) = y] \cdot \Pr[\mathsf{FCnd}(X) = y]$$

$$\leq \frac{\epsilon \cdot 2^{-m}}{2} + \sum_{y \in B_1} (1 + \frac{\epsilon}{2}) \cdot 2^{-m} \cdot \Pr[\mathsf{FCnd}(X) = y]$$

$$\leq \frac{\epsilon \cdot 2^{-m}}{2} + (1 + \frac{\epsilon}{2}) \cdot 2^{-m}$$

$$\leq (1 + \epsilon) \cdot 2^{-m}$$

which is a contradiction.

Putting things together, we conclude that:

$$\Pr[\mathsf{FCnd}(X) \text{ is useful}] = \Pr[\mathsf{FCnd}(X) \notin B_1 \wedge \mathsf{FCnd}(X) \notin B_2]$$

$$= 1 - \Pr[\mathsf{FCnd}(X) \in B_1 \cup B_2]$$

$$\geq 1 - (\Pr[\mathsf{FCnd}(X) \in B_1] + \Pr[\mathsf{FCnd}(X) \in B_2])$$

$$\geq 1 - (1 - \frac{\epsilon \cdot 2^{-m}}{2} + \frac{\epsilon \cdot 2^{-m}}{4})$$

$$= \frac{\epsilon \cdot 2^{-m}}{4}$$

$\square$

Let $n' = n^{a_0}$ be the output length of g. For every $y \in \{0, 1\}^t$ and every $0 \leq \alpha \leq 1$, we define:

$$T_{y,\alpha} = \left\{ v \in \{0, 1\}^{n'} : \Pr[\mathsf{TExt}(v, X) = z \mid \mathsf{FCnd}(X) = y] > (1 + \alpha) \cdot 2^{-m} \right\}.$$

With this definition we immediately have that for every useful y, $g(y) \in T_{y,\epsilon/2}$. We now observe that for every useful y, $T_{y,\frac{\epsilon}{8}}$ is a small set. The next Claim is similar to Claim 3.4 in the proof of Lemma 3.2.

Claim 4.3. *For every useful $y \in \{0, 1\}^t$, $|T_{y,\frac{\epsilon}{8}}| < 2^{k'}$.*

Proof. If this does not hold, then there exists a useful $y \in \{0, 1\}^t$, such that $|T_{y,\frac{\epsilon}{8}}| \geq 2^{k'}$. We consider the following two distributions: The first is V_y that is uniform over $T_{y,\frac{\epsilon}{8}}$, and the second $W_y = (X \mid \mathsf{FCnd}(X) = y)$. These two distributions are independent, and have min-entropy at least k', and therefore, by the guarantee of TExt, we have that $\Pr[\mathsf{TExt}(V_y, W_y) = z] \leq 2^{-m} + \epsilon'$. This is a contradiction as we also have that,

$$\Pr[\mathsf{TExt}(V_y, W_y) = z] = \Pr[\mathsf{TExt}(V_y, X) = z \mid \mathsf{FCnd}(X) = y]$$

$$= \mathbb{E}_{v \leftarrow T_{y,\alpha}}[\Pr[\mathsf{TExt}(v, X) = z \mid \mathsf{FCnd}(X) = y]]$$

$$> (1 + \frac{\epsilon}{8}) \cdot 2^{-m}$$

$$> 2^{-m} + \frac{n^{-2c'}}{16}$$

$$= 2^{-m} + \epsilon',$$

where this follows because $\epsilon = n^{-c'}$, $m = c' \cdot \log n$, and $\epsilon' = \frac{n^{-2c'}}{16}$. $\square$

The proof will proceed using a similar argument to the proof of Lemma 3.2. We have that X is $(n^c, 0)$-samplable which means that $X \leftarrow A \mid P$, for some circuits $A : \{0,1\}^r \to \{0,1\}^n$ and $P : \{0,1\}^r \to \{0,1\}$ of size n^c. We now define the following circuits.

Definition 4.4. *For every $y \in \{0,1\}^t$, and $v \in \{0,1\}^{n'}$ we define two circuits $C^1_{y,v} : \{0,1\}^r \to \{0,1\}$ and $C^2_y : \{0,1\}^r \to \{0,1\}$ of size $\mathsf{poly}(n^{d_{\mathsf{FCnd}}})$ as follows:*

- *$C^1_{y,v}(w)$ answers one iff $\mathsf{TExt}(v, A(w)) = z \land \mathsf{FCnd}(A(w)) = y \land P(w) = 1$.*
- *$C^2_y(w)$ answers one iff $\mathsf{FCnd}(A(w)) = y \land P(w) = 1$.*

We also define:

- *$p^1_{y,v} = \Pr[C^1_{y,v}(U_r) = 1]$.*
- *$p^2_y = \Pr[C^2_y(U_r) = 1]$.*

Claim 4.5. *For every $y \in \{0,1\}^t$ and $v \in \{0,1\}^{n'}$, if $p^2_y \neq 0$, then*

$$\frac{p^1_{y,v}}{p^2_y} = \Pr[\mathsf{TExt}(v, X) = z \mid \mathsf{FCnd}(X) = y].$$

Proof. For every $y \in \{0,1\}^t$ and $v \in \{0,1\}^{n'}$, if $p^2_y \neq 0$ then for $W \leftarrow U_r$, we have that:

$$
\begin{aligned}
\frac{p^1_{y,v}}{p^2_y} &= \frac{\Pr[\mathsf{TExt}(v, A(W)) = z \land \mathsf{FCnd}(A(W)) = y \land P(W) = 1]}{\Pr[\mathsf{FCnd}(A(W)) = y \land P(W) = 1]} \\
&= \frac{\Pr[\mathsf{TExt}(v, A(W)) = z \land \mathsf{FCnd}(A(W)) = y \mid P(W) = 1] \cdot \Pr[P(W) = 1]}{\Pr[\mathsf{FCnd}(A(W)) = y \mid P(W) = 1] \cdot \Pr[P(W) = 1]} \\
&= \frac{\Pr[\mathsf{TExt}(v, A(W)) = z \land \mathsf{FCnd}(A(W)) = y \mid P(W) = 1]}{\Pr[\mathsf{FCnd}(A(W)) = y \mid P(W) = 1]} \\
&= \frac{\Pr[\mathsf{TExt}(v, X) = z \land \mathsf{FCnd}(X) = y]}{\Pr[\mathsf{FCnd}(X) = y]} \\
&= \Pr[\mathsf{TExt}(v, X) = z \mid \mathsf{FCnd}(X) = y].
\end{aligned}
$$

$\square$

This means that for every $y \in \{0,1\}^t$ and $0 \leq \alpha \leq 1$, we can decide whether a given $v \in \{0,1\}^t$ is in $T_{y,\alpha}$ if we can check whether $p^2_y = 0$ and compute $p^1_{y,v}$ and p^2_y. By Theorem 2.7 a small Σ_1-circuit, can compute relative approximations to $p^1_{y,v}$ and p^2_y. We will now use this idea to prove the following:

Claim 4.6. *For every $y \in \{0,1\}^t$, there is a Σ_1-circuit $C_y : \{0,1\}^{n'} \to \{0,1\}$ of size $\mathsf{poly}(n^{d_{\mathsf{FCnd}}}, \frac{1}{\epsilon})$ such that:*

- *For every $v \in \{0,1\}^{n'}$ such that $C_y(v) = 1$, we have that $v \in T_{y, \frac{\epsilon}{8}}$.*
- *If y is useful, then $C_y(g(y)) = 1$.*

Proof. When given $v \in \{0,1\}^{n'}$, the circuit C_y works as follows:

- C_y checks whether there exists $w \in \{0,1\}^r$, such that $\mathsf{FCnd}(A(w)) = y$ and $P(w) = 1$. If there does not exist such a w, it answers zero (as this means that $p^2_y = 0$)

39

- Let $\lambda = \epsilon/a'$ for a universal constant $a' > 1$ to be chosen later. C_y applies Theorem 2.24 to compute a λ-relative approximations $\hat{p}^1_{y,v}$ and $\hat{p}^2_y$, of $p^1_{y,v}$ and p^2_y respectively. It can do this by computing approximations of the number of accepting inputs of the circuits $C^1_{y,v}$ and C^2_y, respectively.

- C_y computes $\hat{p}_{y,v} = \frac{\hat{p}^1_{y,v}}{\hat{p}^2_y}$ and note that as this is an $O(\lambda)$-relative approximation to $p_{y,v} = \frac{p^1_{y,v}}{p^2_y} = \Pr[\mathsf{TExt}(v, X) = z \mid \mathsf{FCnd}(X) = y]$.

- C_y outputs one if $\hat{p}_{y,v} > (1 + \frac{\epsilon}{4}) \cdot 2^{-m}$ and zero otherwise.

By choosing the constant a' to be sufficiently large, we can make $\lambda = \epsilon/a'$ sufficiently small, to guarantee that checking whether the $O(\lambda)$ approximation $\hat{p}_{y,v}$ is larger than $(1 + \frac{\epsilon}{4}) \cdot 2^{-m}$ distinguishes between the case that $p_{y,v} > (1 + \frac{\epsilon}{2}) \cdot 2^{-m}$ and the case that $p_{y,v} \leq (1 + \frac{\epsilon}{8}) \cdot 2^{-m}$. This gives that if $C_y(v) = 1$ then $p_{y,v} > (1 + \frac{\epsilon}{8}) \cdot 2^{-m}$ which gives that $v \in T_{y,\frac{\epsilon}{8}}$. We have that for every useful y, $g(y) \in T_{y,\frac{\epsilon}{2}}$, which means that $p_{y,g(y)} > (1 + \frac{\epsilon}{2}) \cdot 2^{-m}$, and indeed, $C_y(g(y)) = 1$.

Finally, by definition C_y is a circuit of size $\mathsf{poly}(n^{d_{\mathsf{FCnd}}}, \frac{1}{\lambda}) = \mathsf{poly}(n^{d_{\mathsf{FCnd}}}, \frac{1}{\epsilon})$. $\qquad\square$

We are finally ready to complete the proof, with the next claim.

Claim 4.7. *There is a Σ_2-circuit C of size $\mathsf{poly}(n^{d_{\mathsf{FCnd}}}, \frac{1}{\epsilon})$ such that for $Y = \mathsf{FCnd}(X)$,*

$$\Pr[C(Y) = g(Y)] \geq 2^{-k'} \cdot \frac{\epsilon \cdot 2^{-m}}{8}$$

Proof. We will first construct a randomized Σ_2-circuit C', and then use a standard averaging argument to convert it to a non-randomized Σ_2-circuit. The randomized circuit C' is defined as follows: On input $y \in \{0,1\}^t$:

- C' constructs the Σ_1-circuit C_y. Note that the circuit C_y is specified precisely in the proof of Claim 4.6, and so, the circuit C' (that can be hardwired with A, P, FCnd, z, and the circuit from Theorem 2.24) can construct the circuit C_y.

- C' uses the Σ_2-circuit guaranteed in Theorem 2.25 (choosing $i = 1$ and $\delta = \frac{1}{2}$) to output a uniform element in $\{v : C_y(v) = 1\}$.

By definition, the circuit C' is a randomized Σ_2-circuit of size $\mathsf{poly}(n^{d_{\mathsf{FCnd}}}, \frac{1}{\epsilon})$. We conclude that:

$$\Pr[C'(Y) = g(Y)] \geq \Pr[C'(Y) = g(Y) \mid Y \text{ is useful}] \cdot \Pr[Y \text{ is useful}]$$
$$\geq \Pr[C'(Y) = g(Y) \mid Y \text{ is useful}] \cdot \frac{\epsilon \cdot 2^{-m}}{4}$$
$$\geq \frac{1}{2} \cdot 2^{-k'} \cdot \frac{\epsilon \cdot 2^{-m}}{4}$$
$$= 2^{-k'} \cdot \frac{\epsilon \cdot 2^{-m}}{8}.$$

where the first inequality follows from Claim 4.2, and the last inequality follows because by Claim 4.6, for every useful y, $g(y) \in \{v : C_y(v) = 1\}$ which by Claim 3.4, is of size at most $2^{k'}$, and each element in the set is obtained with probability $\frac{1}{2} \cdot 2^{-k'}$.

Finally, by a standard averaging argument, there exists a (non-randomized) Σ_2-circuit of size $\mathsf{poly}(n^{d_{\mathsf{FCnd}}}, \frac{1}{\epsilon})$ with the same success probability. $\qquad\square$

We have obtained a Σ_2-circuit C of size $\mathsf{poly}(n^{d_{\mathsf{FCnd}}}, \frac{1}{\epsilon})$. Recall that $\epsilon = n^{-c}$ and so, we can choose the constant $\bar{c}$ to be sufficiently large so that the size of C is bounded by $n^{\bar{c}-1}$. We can view the distribution $Y = \mathsf{FCnd}(X)$ (which is over $\{0,1\}^t$) as a distribution over $\{0,1\}^n$ by padding $\mathsf{FCnd}(X)$ with zeros (recall that $t = n^{\delta_2} \leq n$). We have that:

- $H_\infty(Y) \geq n^{\delta_1}$, and

- Y is $(n^{\bar{c}}, 0)$-samplable (it is sampled by $(\mathsf{FCnd} \circ A) \mid P$).

We also have that,

$$
\begin{aligned}
\Pr[C(Y) = G(Y)] &\geq 2^{-k'} \cdot \frac{\epsilon}{8} \cdot 2^{-m} \\
&= 2^{-(k' + 2c' \log n - 3)} \\
&\geq 2^{-2k'} \\
&\geq 2^{-\Omega(\frac{n^\nu}{a})},
\end{aligned}
$$

where the second line follows because $\epsilon = n^{-c'}$, and $m = c' \log n$. The third line follows because $k' = n^\nu/a$. Finally, by choosing the constant a to be sufficiently large, we can make sure that the success probability of C violates the HOS guarantee of g.

4.2 Obtaining Extractors with Large Output Length

Shaltiel [Sha08] showed how to take an extractor for samplable distributions with small output length, and transform it into one that has large output length. This transformation works by first extracting t bits (using the initial extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^t$ for samplable distributions) and then using the output as a seed to a seeded strong extractor $\mathsf{SExt} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$. Note that the original source X, and the seed $\mathsf{Ext}(X)$ (that is used for the seeded extractor) are correlated, and so, it is not clear that such a transformation should work. Nevertheless, Shaltiel [Sha08] showed that if the error of the initial extractor Ext is smaller than 2^{-t}, then this transformation does work, assuming Ext can extract from distributions samplable by Σ_1-circuits.

An inspection of Shaltiel's argument shows that it is in fact sufficient that Ext is an extractor for distributions which are $(n^c, 0)$-samplable, that is, it is sufficient that Ext is an extractor for distributions that are samplable with postselection (rather than samplable by Σ_1-circuits).

As the extractor of Theorem 4.1 works for $(n^c, 0)$-samplable, we can increase the output length by composing with the seeded extractors of Theorems 2.22 and Theorem 2.12. This composition gives Theorem 1.4 and Theorem 1.6 respectively.

We omit the precise argument, as an identical argument (with the same choices of seeded extractors) appears in [BSS25], and the proof is identical to that proof.

4.3 Obtaining a Multiplicative Extractor

In this section we prove Theorem 1.8, which we now restate in a stronger form, in which the extractor applies to samplable distributions with postselection.

Theorem 4.8 (multiplicative extractor). *For every constants $\alpha > 0$, there exist constants $j \geq 1$, and $\beta > 0$ such that if E is hard for exponential size Σ_j-circuits, then for every constant $c > 1$, there exists a constant d such that for every sufficiently large n, there is a function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{n^\beta}$ that is an $(n^\alpha, \frac{1}{n^c})$-multiplicative extractor for $(n^c, 0)$-samplable distributions. Furthermore, Ext can be computed in time $\mathrm{poly}(n^d)$.*

As explained in Section 1.3.6, the proof of Theorem 1.8 will try to imitate the proof of Theorem 4.1 with a more complicated construction that relies on the extractor of [Li15] (which is designed for one source and one block-wise source) instead of standard 2-source extractors. While the overall argument is similar in spirit, there are additional technical complications. The full proof appears below.

Parameters:

- Given a constant $\alpha > 0$, let j be a sufficiently large constant to be determined later, and $\beta > 0$ be a sufficiently small constant to be determined later.
- We are given a constants $c > 1$.
- We are assuming that n is sufficiently large.

Assumption: We are assuming that E is hard for exponential size Σ_j-circuits.

Ingredients: We will require the following ingredients:

- *First Errorless condenser:* Under the hardness assumption, for a sufficiently large j, by Theorem 3.1, there exist constants $0 < \delta_1^1 < \delta_2^1 \leq \frac{\alpha}{10}$, and a function $\mathsf{FCnd}_1 : \{0,1\}^n \to \{0,1\}^{n^{\delta_2^1}}$ that is an $(n^\alpha, n^{\delta_1^1})$-errorless condenser for $(n^c, 0)$-samplable distributions. Furthermore, there is a constant $d_{\mathsf{FCnd}_1} \geq 1$ such that FCnd_1 can be computed in time $n^{d_{\mathsf{FCnd}_1}}$.

- *Second Errorless condenser:* Under the hardness assumption, for a sufficiently large j, by Theorem 3.1, there exist constants $0 < \delta_1^2 < \delta_2^2 \leq \frac{\delta_1^1}{10}$, and a function $\mathsf{FCnd}_2 : \{0,1\}^n \to \{0,1\}^{n^{\delta_2^2}}$ that is an $(n^{\delta_1^1}, n^{\delta_1^1})$-errorless condenser for $(n^{d_{\mathsf{FCnd}_1}+1}, 0)$-samplable distributions. Furthermore, there is a constant $d_{\mathsf{FCnd}_2} \geq 1$ such that FCnd_2 can be computed in time $n^{d_{\mathsf{FCnd}_2}}$

- *HOS:* Let $\bar{c} > d_{\mathsf{FCnd}_2}$ be a sufficiently large constant that will be chosen in the proof. By Theorem 2.22, setting $\nu = \delta_1^2$, under the hardness assumption, there is a function $g : \{0,1\}^n \to \{0,1\}^{n^{a_0}}$ (where $a_0 > 1$ is a universal constant) that is an $(n^\nu, 2^{-\Omega(n^\nu)})$-HOS for the class of $(n^{\bar{c}}, 0)$-samplable distributions, against the class of size $n^{\bar{c}}$ Σ_2-circuits. We have that g can be computed in time n^{d_g} for some constant d_g that depends on $\bar{c}$.

- *Extractor for one general source and one block-wise source:* Let $k' = n^{\frac{\delta_1^2}{2}}$, let $m = n^\beta$, let $\epsilon' = \frac{n^{-c} \cdot 2^{-m}}{16}$, and let $\mathsf{IExt} : \{0,1\}^{n^{a_0}} \times \{0,1\}^{n^{a_0}} \times \{0,1\}^{n^{a_0}} \to \{0,1\}^m$ be the extractor from Theorem 2.15, set for min-entropy threshold k' and error ϵ'. Note that the error ϵ' guaranteed in Theorem 2.15 is $\epsilon' = 2^{-(k')^{\Omega(1)}}$. Recall that we have chosen $m = n^\beta$, and we can choose the constant $\beta > 0$ to be sufficiently small so that $\epsilon' = 2^{-(k')^{\Omega(1)}} \leq \frac{n^{-c} \cdot 2^{-m}}{16}$. We have that IExt can be computed in time $n^{d_{\mathsf{IExt}}}$

Construction: We define $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$, as follows:

$$\mathsf{Ext}(x) = \mathsf{IExt}(g(\mathsf{FCnd}_2(\mathsf{FCnd}_1(x))), \mathsf{FCnd}_1(x), x).$$

Note that g expects inputs of length n, and IExt expects inputs of length n^{a_0}, and so whenever an input is too short, we pad it with zeros to the appropriate length. Note that by construction, there indeed exists a constant d such that Ext can be computed in time n^d.

4.3.1 Proof of Theorem 4.8

The construction of Ext appears in Figure 4 and closely follows the intuition explained in Section 1.3.6. Let n be sufficiently large, let $\epsilon = n^{-c}$ and assume for the purpose of contradiction Ext is not an (n^α, ϵ)-multiplicative extractor. Specifically, that there is an $(n^c, 0)$ samplable distribution X such that $H_\infty(X) \geq n^\alpha$, and a $z \in \{0,1\}^m$ such that:

$$\Pr[\mathsf{IExt}(g(\mathsf{FCnd}_2(\mathsf{FCnd}_1(X))), \mathsf{FCnd}_1(X), X) = z] > (1 + \epsilon) \cdot 2^{-m}.$$

Once again, our goal will be to obtain a contradiction by showing that there is a Σ_2-circuit C of size $n^{\bar{c}}$,

and a $(n^{\bar{c}}, 0)$-samplable distribution Y (we will choose $Y = \mathsf{FCnd}_2(\mathsf{FCnd}_1(X))$) with $H_\infty(Y) \geq n^\nu$, on which $\Pr[C(Y) = g(Y)]$ is too large, and contradicts the HOS guarantee of g. In order to avoid clutter, we define $F_1(x) = \mathsf{FCnd}_1(x)$ and $F_2(x) = \mathsf{FCnd}_2(\mathsf{FCnd}_1(x))$. Note that we have chosen the parameters so that $H_\infty(F_1(X)) \geq n^{\delta_1^1}$ and $H_\infty(F_2(X)) \geq n^{\delta_1^2}$.

The proof below will use the same structure as the proof of Theorem 4.1. Let $t = n^{\delta_2^2}$. We will say that $y \in \{0,1\}^t$ is *useful* if the following three conditions hold:

- $\Pr[\mathsf{IExt}(g(y), F_1(X), X) = z \mid F_2(X) = y] > (1 + \frac{\epsilon}{2}) \cdot 2^{-m}$.
- $H_\infty(F_1(X) \mid F_2(X) = y) \geq k'$.
- $H_\infty(X \mid F_2(X) = y) \geq n^\alpha/2$.

Claim 4.9. $\Pr[F_2(X) \text{ is useful}] > \frac{\epsilon \cdot 2^{-m}}{4}$.

Proof. We define:

$$B_1 = \left\{ y \in \{0,1\}^t : \Pr[\mathsf{IExt}(g(y), F_1(X), X) = z \mid F_2(X) = y] \leq (1 + \frac{\epsilon}{2}) \cdot 2^{-m} \right\}$$
$$B_2 = \left\{ y : H_\infty(F_1(X) \mid F_2(X) = y) < k' \right\}$$
$$B_3 = \left\{ y : H_\infty(X \mid F_2(X) = y) < n^\alpha/2 \right\}$$

This is done so that y is useful iff $y \notin B_1 \cup B_2 \cup B_3$. We will proceed to bound $\Pr[F_2(X) \in B_1]$, $\Pr[F_2(X) \in B_2]$ and $\Pr[F_2(X) \in B_3]$ separately.

By Lemma 2.3 with probability at least $1 - \frac{\epsilon \cdot 2^{-m}}{8}$ over choosing $y \leftarrow F_2(X)$, we have that

$$H_\infty(F_1(X) \mid F_2(X) = y) \geq n^{\delta_1^1} - n^{\delta_2^2} - \log(1/\epsilon) - m - 3 \geq n^{\delta_1^1} - n^{\delta_2^2} - c \cdot \log n - n^\beta - 3 \geq k',$$

where this follows because $\delta_1^1 > \delta_2^2$, $\beta > 0$ can be chosen to be sufficiently small, and $k' = n^{\frac{\delta_1^2}{2}}$ This gives that $\Pr[F_2(X) \in B_2] \leq \frac{\epsilon \cdot 2^{-m}}{8}$.

By Lemma 2.3, with probability at least $1 - \frac{\epsilon \cdot 2^{-m}}{8}$ over choosing $y \leftarrow F_2(X)$, we have that

$$H_\infty(X \mid F_2(X) = y) \geq n^\alpha - n^{\delta_2^1} - \log(1/\epsilon) - m - 3 \geq n^\alpha - n^{\delta_2^1} - c \cdot \log n - n^\beta - 3 \geq n^\alpha/2,$$

where this follows because $\delta_2^1 < \alpha/10$, and $\beta > 0$ can be chosen to be sufficiently small. This gives that $\Pr[F_2(X) \in B_3] \leq \frac{\epsilon \cdot 2^{-m}}{8}$.

We claim that $\Pr[F_2(X) \in B_1] \leq 1 - \frac{\epsilon \cdot 2^{-m}}{2}$. This follows as otherwise, $\Pr[F_2(X) \notin B_1] \leq \frac{\epsilon \cdot 2^{-m}}{2}$, which implies that for $p = \Pr[\mathsf{IExt}(g(F_2(X)), F_1(X), X) = z]$ we have that:

$$
\begin{aligned}
p &= \Pr[\mathsf{IExt}(g(F_2(X)), F_1(X), X) = z] \\
&\leq \Pr[F_2(X) \notin B_1] + \Pr[\mathsf{IExt}(g(F_2(X)), F_1(X), X) = z \cap F_2(X) \in B_1] \\
&\leq \frac{\epsilon \cdot 2^{-m}}{2} + \sum_{y \in B_1} \Pr[\mathsf{IExt}(g(F_2(X)), F_1(X), X) = z \cap F_2(X) = y] \\
&= \frac{\epsilon \cdot 2^{-m}}{2} + \sum_{y \in B_1} \Pr[\mathsf{IExt}(g(F_2(X)), F_1(X), X) = z \mid F_2(X) = y] \cdot \Pr[F_2(X) = y] \\
&\leq \frac{\epsilon \cdot 2^{-m}}{2} + \sum_{y \in B_1} (1 + \frac{\epsilon}{2}) \cdot 2^{-m} \cdot \Pr[F_2(X) = y] \\
&\leq \frac{\epsilon \cdot 2^{-m}}{2} + (1 + \frac{\epsilon}{2}) \cdot 2^{-m} \\
&\leq (1 + \epsilon) \cdot 2^{-m}
\end{aligned}
$$

which is a contradiction.

Putting things together, we conclude that:

$$\begin{aligned}
\Pr[F_2(X) \text{ is useful}] &= \Pr[F_2(X) \notin B_1 \wedge F_2(X) \notin B_2 \wedge F_2(X) \notin B_3] \\
&= 1 - \Pr[F_2(X) \in B_1 \cup B_2 \cup B_3] \\
&\geq 1 - (\Pr[F_2(X) \in B_1] + \Pr[F_2(X) \in B_2] + \Pr[F_2(X) \in B_3]) \\
&\geq 1 - (1 - \frac{\epsilon \cdot 2^{-m}}{2} + \frac{\epsilon \cdot 2^{-m}}{8} + \frac{\epsilon \cdot 2^{-m}}{8}) \\
&= \frac{\epsilon \cdot 2^{-m}}{4}
\end{aligned}$$

$\square$

Let $n' = n^{a_0}$ be the output length of g. For every $y \in \{0,1\}^t$ and every $0 \leq \alpha \leq 1$, we define:

$$T_{y,\alpha} = \left\{ v \in \{0,1\}^{n'} : \Pr[\mathsf{IExt}(v, F_1(X), X) = z \mid F_2(X) = y] > (1+\alpha) \cdot 2^{-m} \right\}.$$

With this definition we immediately have that for every useful y, $g(y) \in T_{y,\epsilon/2}$. We now observe that for every useful y, $T_{y,\frac{\epsilon}{8}}$ is a small set.

Claim 4.10. *For every useful $y \in \{0,1\}^t$, $|T_{y,\frac{\epsilon}{8}}| < 2^{k'}$.*

Proof. If this does not hold, then there exists a useful $y \in \{0,1\}^t$, such that $|T_{y,\frac{\epsilon}{8}}| \geq 2^{k'}$. We consider the following three distributions:

- V_y that is uniform over $T_{y,\frac{\epsilon}{8}}$.
- $W_y^2 = (X \mid F_2(X) = y)$.
- $W_y^1 = F_1(W_y^2)$. This choice is made so that (W_y^1, W_y^2) is distributed like $((F_1(X), X) \mid F_2(X) = Y)$.

Let $W_y = (W_y^1, W_y^2)$. By definition, we have that V_y and W_y are independent distributions. We also claim that $W_y = (W_y^1, W_y^2)$ is $2^{-n^\alpha/10}$-close to a (k', k')-block-wise source. This holds because we have that

$$H_\infty(W_y^2) = H_\infty(F_1(X) \mid F_2(X) = y) \geq k',$$

and by Lemma 2.3 we have that with probability $1 - 2^{-n^\alpha/10}$ over choosing $w^1 \leftarrow W_y^1$, we have that

$$\begin{aligned}
H_\infty(W_y^2 \mid W_y^1 = w^1) &\geq H_\infty(W_y^2) - n^{\delta_2^1} - n^\alpha/10 \\
&= H_\infty(X \mid F_2(X) = y) - n^{\delta_2^1} - n^\alpha/10 \\
&\geq n^\alpha/2 - n^{\delta_2^1} - n^\alpha/10 \\
&\geq k',
\end{aligned}$$

where the third line follows because y is useful, and the fourth line follows because $\delta_2^1 \leq \alpha/10$, and $k' = n^{\frac{\delta_1^2}{2}}$ and $\delta_1^2 \leq \alpha/10$.

We conclude that the distribution (V_y, W_y) is $2^{-n^\alpha/10}$-close to a distribution that satisfies the requirements of IExt. We therefore have that that $\Pr[\mathsf{IExt}(V_y, W_y) = z] \leq 2^{-m} + \epsilon' + 2^{-n^\alpha/10}$. This is a contradiction as

we also have that,

$$\Pr[\mathsf{IExt}(V_y, W_y) = z] = \Pr[\mathsf{IExt}(V_y, F_1(X), X) = z \mid F_2(X) = y]$$
$$= \mathbb{E}_{v \leftarrow T_{y,\alpha}}[\Pr[\mathsf{IExt}(v, F_1(X), X) = z \mid F_2(X) = y]]$$
$$> (1 + \frac{\epsilon}{8}) \cdot 2^{-m}$$
$$> 2^{-m} + \epsilon' + 2^{-n^\alpha/10},$$

where this follows because $\epsilon' = \frac{\epsilon \cdot 2^{-m}}{16}$, and $\epsilon = n^{-c}$. $\qquad\square$

We have that X is $(n^c, 0)$-samplable which means that $X \leftarrow A \mid P$, for some circuits $A : \{0,1\}^r \to \{0,1\}^n$ and $P : \{0,1\}^r \to \{0,1\}$ of size n^c. We now define the following circuits.

Definition 4.11. *For every $y \in \{0,1\}^t$, and $v \in \{0,1\}^{n'}$ we define two circuits $C^1_{y,v} : \{0,1\}^r \to \{0,1\}$ and $C^2_y : \{0,1\}^r \to \{0,1\}$ of size $\mathsf{poly}(n^{d_{\mathsf{FCnd}}})$ as follows:*

- *$C^1_{y,v}(w)$ answers one iff $\mathsf{IExt}(v, F_1(A(w)), A(w)) = z \wedge F_2(A(w)) = y \wedge P(w) = 1$.*
- *$C^2_y(w)$ answers one iff $F_2(A(w)) = y \wedge P(w) = 1$.*

We also define:

- *$p^1_{y,v} = \Pr[C^1_{y,v}(U_r) = 1]$.*
- *$p^2_y = \Pr[C^2_y(U_r) = 1]$.*

Claim 4.12. *For every $y \in \{0,1\}^t$ and $v \in \{0,1\}^{n'}$, if $p^2_y \neq 0$, then*

$$\frac{p^1_{y,v}}{p^2_y} = \Pr[\mathsf{IExt}(v, F_1(X), X) = z \mid F_2(X) = y].$$

Proof. For every $y \in \{0,1\}^t$ and $v \in \{0,1\}^{n'}$, if $p^2_y \neq 0$ then for $W \leftarrow U_r$, we have that:

$$\frac{p^1_{y,v}}{p^2_y} = \frac{\Pr[\mathsf{IExt}(v, F_1(A(W)), A(W)) = z \wedge F_2(A(W)) = y \wedge P(W) = 1]}{\Pr[F_2(A(W)) = y \wedge P(W) = 1]}$$
$$= \frac{\Pr[\mathsf{IExt}(v, F_1(A(W)), A(W)) = z \wedge F_2(A(W)) = y \mid P(W) = 1] \cdot \Pr[P(W) = 1]}{\Pr[F_2(A(W)) = y \mid P(W) = 1] \cdot \Pr[P(W) = 1]}$$
$$= \frac{\Pr[\mathsf{IExt}(v, F_1(A(W)), A(W)) = z \wedge F_2(A(W)) = y \mid P(W) = 1]}{\Pr[F_2(A(W)) = y \mid P(W) = 1]}$$
$$= \frac{\Pr[\mathsf{IExt}(v, F_1(X), X) = z \wedge F_2(X) = y]}{\Pr[F_2(X) = y]}$$
$$= \Pr[\mathsf{IExt}(v, F_1(X), X) = z \mid F_2(X) = y].$$

$\qquad\square$

This means that for every $y \in \{0,1\}^t$ and $0 \leq \alpha \leq 1$, we can decide whether a given $v \in \{0,1\}^t$ is in $T_{y,\alpha}$ if we can check whether $p^2_y = 0$ and compute $p^1_{y,v}$ and p^2_y. By Theorem 2.7 a small Σ_1-circuit, can compute relative approximations to $p^1_{y,v}$ and p^2_y. We will now use this idea to prove the following:

Claim 4.13. *For every $y \in \{0,1\}^t$, there is a Σ_1-circuit $C_y : \{0,1\}^{n'} \to \{0,1\}$ of size $\mathsf{poly}(n^{d_{\mathsf{FCnd2}}}, \frac{1}{\epsilon})$ such that:*

45

- *For every $v \in \{0,1\}^{n'}$ such that $C_y(v) = 1$, we have that $v \in T_{y,\frac{\epsilon}{8}}$.*

- *If y is useful, then $C_y(g(y)) = 1$.*

The proof of this claim is essentially identical to the proof of Claim 4.6 in the proof of Theorem 4.1 and we omit it. We also need the next claim.

Claim 4.14. *There is a Σ_2-circuit C of size $\mathsf{poly}(n^{d_{\mathsf{FCnd}}}, \frac{1}{\epsilon})$ such that for $Y = F_2(X)$,*

$$\Pr[C(Y) = g(Y)] \geq 2^{-k'} \cdot \frac{\epsilon \cdot 2^{-m}}{8}$$

Once again the proof is identical to the proof of Claim 4.7 and we omit it.

We have obtained a Σ_2-circuit C of size $\mathsf{poly}(n^{d_{\mathsf{FCnd_2}}}, \frac{1}{\epsilon})$. Recall that $\epsilon = n^{-c}$ and so, we can choose the constant $\bar{c}$ to be sufficiently large so that the size of C is bounded by $n^{\bar{c}-1}$. We can view the distribution $Y = F_2(X)$ (which is over $\{0,1\}^t$) as a distribution over $\{0,1\}^n$ by padding $F_2(X)$ with zeros. We have that:

- $H_\infty(Y) \geq n^{\delta_1^2}$, and

- Y is $(n^{\bar{c}}, 0)$-samplable (it is sampled by $(F_2 \circ A) \mid P$).

We also have that,

$$\begin{aligned}
\Pr[C(Y) = G(Y)] &\geq 2^{-k'} \cdot \frac{\epsilon}{8} \cdot 2^{-m} \\
&= 2^{-(k' + c \log n - 3 - n^\beta)} \\
&\geq 2^{-2k'} \\
&\geq 2^{-\Omega(n^\nu)},
\end{aligned}$$

where the second line follows because $\epsilon = n^{-c}$, and $m = n^\beta$. The third line follows because we can choose $\beta > 0$ to be sufficiently small, and $k' = n^{\frac{\delta_1^2}{2}}$, and $\nu = \delta_1^2$. This gives that the success probability of C violates the HOS guarantee of g.

5 Conclusion and Open Problems

In this paper we construct extractors for samplable distributions with polynomially small min-entropy. There are several natural open problems.

Using a weaker hardness assumption. The extractors of Theorem 1.4 and Theorem 1.6 use a hardness assumption for Σ_j-circuits, where j increases as the min-entropy threshold is reduced. Can this be avoided? This is even more severe in the case of the multiplicative extractors of Theorem 1.8.

More ambitiously, is it possible to achieve similar extractors under hardness for nondeterministic circuits?

Further reducing the min-entropy threshold. It is interesting to try and construct extractors for min-entropy threshold $k = n^{o(1)}$.

For both aforementioned open problems, a natural direction is to give an improved construction of error-less condensers (for lower min-entropy threshold and/or weaker hardness assumption).

Improved multiplicative extractors for low min-entropy. The previous work of Ball, Shaltiel and Silbak [BSS25] gave extractors that are not multiplicative. In this work, we are able to obtain multiplicative extractors, however, both the assumption used, and the output length are inferior compared to our (standard extractors) and it is a natural open problem to try and improve them.

Acknowledgenemt

We are grateful to Gil Cohen for a helpful discussion, and for pointing our attention to [Li15].

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.

[BCDT19] A. Ben-Aroya, G. Cohen, D. Doron, and A. Ta-Shma. Two-source condensers with low error and small entropy gap via entropy-resilient functions. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*, volume 145 of *LIPIcs*, pages 43:1–43:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[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.

[BGP00] M. Bellare, O. Goldreich, and E. Petrank. Uniform generation of np-witnesses using an np-oracle. *Inf. Comput.*, 163(2):510–526, 2000.

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

[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.

[BSS25] M. Ball, J. Silbak, and R. Shaltiel. Extractors for samplable distributions with low min-entropy. *To appear in STOC 2025*, 2025.

[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.

[CZ16]   E. Chattopadhyay and D. Zuckerman. Explicit two-source extractors and resilient functions. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC*, pages 670–683, 2016.

[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.

[DY13]   Y. Dodis and Y. Yu. Overcoming weak expectations. In *Theory of Cryptography - 10th Theory of Cryptography Conference, TCC*, volume 7785 of *Lecture Notes in Computer Science*, pages 1–22, 2013.

[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.

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

[JVV86]   M. Jerrum, L. G. Valiant, and V. V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theor. Comput. Sci.*, 43:169–188, 1986.

[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.

[Li15]   X. Li. Three-source extractors for polylogarithmic min-entropy. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS*, pages 863–882, 2015.

[Li16]   X. Li. Improved two-source extractors, and affine extractors for polylogarithmic entropy. In *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS*, pages 168–177. IEEE Computer Society, 2016.

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

[RSW06]   O. Reingold, R. Shaltiel, and A. Wigderson. Extracting randomness via repeated condensing. *SIAM J. Comput.*, 35(5):1185–1209, 2006.

[Sha08]   R. Shaltiel. How to get more mileage from randomness extractors. *Random Struct. Algorithms*, 33(2):157–186, 2008.

[Sha24]   R. Shaltiel. Multiplicative extractors for samplable distributions. *Electronic Colloquium on Computational Complexity (ECCC)*, TR24-168, 2024.

[Sip83]   M. Sipser. A complexity theoretic approach to randomness. In *STOC*, pages 330–335, 1983.

[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.

[SSZ98]   M. E. Saks, A. Srinivasan, and S. Zhou. Explicit or-dispersers with polylogarithmic degree. *J. ACM*, 45(1):123–154, 1998.

[Sto83]   L. J. Stockmeyer. The complexity of approximate counting. In *STOC*, pages 118–126, 1983.

[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.

[Ta-96]   Amnon Ta-Shma. On extracting randomness from weak random sources (extended abstract). In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA, May 22-24, 1996*, pages 276–285, 1996.

[TU12]   A. Ta-Shma and C. Umans. Better condensers and new extractors from parvaresh-vardy codes. In *Proceedings of the 27th Conference on Computational Complexity, CCC*, pages 309–315. IEEE Computer Society, 2012.

[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.

[Zuc07]   David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. *Theory Comput.*, 3(1):103–128, 2007.