# Extractors for Samplable Distributions with Low Min-Entropy[*]

Marshall Ball[†] Ronen Shaltiel[‡] Jad Silbak[§]

December 16, 2024

## 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, there are extractors for samplable distributions with large min-entropy of $k = (1 - \gamma) \cdot n$, for some small constant $\gamma > 0$. Recent work by Ball, Goldin, Dachman-Soled and Mutreja (FOCS 2023) weakened the hardness assumption. However, since the original paper by Trevisan and Vadhan, there has been no improvement in the min-entropy threshold $k$.

In this paper we give a construction of extractors for samplable distributions with low min-entropy of $k = n^{1-\gamma}$ for some constant $\gamma > 0$, and in particular we achieve $k < \frac{n}{2}$ (which is a barrier for the construction of Trevisan and Vadhan).

Our extractors are constructed under a hardness assumption that is weaker than the one used by Trevisan and Vadhan, and stronger than that used by Ball, Goldin, Dachman-Soled and Mutreja. Specifically, that there exists a constant $\beta > 0$, and a problem in $\mathsf{E} = \mathsf{DTIME}(2^{O(n)})$ that cannot be computed by size $2^{\beta n}$ circuits that have an oracle to $\Sigma_5^{\mathsf{P}}$.

Our approach builds on the technique of Trevisan and Vadhan, while introducing new objects and ideas. We introduce and construct two objects: an errorless (seedless) condenser for samplable distributions, and functions that are hard to compute on every samplable distributions with sufficient min-entropy. We use techniques by Shaltiel and Silbak (STOC 2024), as well as additional tools and ideas, to construct the two new objects, under the hardness assumption. We then show how to modify the construction of Trevisan and Vadhan, using these new objects, so that the barrier of $k = n/2$ can be bypassed, and we can achieve an extractor for samplable distributions with low min-entropy.

[*] In memory of Luca Trevisan.

[†] New York University, Email:marshall.ball@cs.nyu.edu. Research supported in part by generous support from Simons Foundation. This research was conducted in part while this author was visiting Tel Aviv University.The visit was supported by the European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme (grant agreement No. 101042417, acronym SPP).

[‡] University of Haifa, Email: ronen@cs.haifa.ac.il. Research supported by ISF grants 1006/23.

[§] Northeastern University, Email: jadsilbak@gmail.com. Research supported by the Khoury College Distinguished Postdoctoral Fellowship.

# Contents

# 1 Introduction

## 1.1 Extractors for Samplable Distributions

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* $(k,\epsilon)$**-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* $\epsilon$*-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 $\mathsf{poly}(n^c)$. They showed that such extractors imply circuit lower bounds, and so, motivated by the hardness vs. randomness paradigm, they gave a conditional construction based on hardness assumptions.

**Hardness assumptions against various types of nondeterministic circuits.** We say that "$\mathsf{E}$ 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 2.4 for a more formal definition).

The assumptions that $\mathsf{E}$ 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 $\mathsf{E}$ is hard for exponential size nondeterministic circuits[2], originated in works on hardness versus randomness for $\mathsf{AM}$, 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 $\Sigma_i$-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 $\Sigma_i^{\mathsf{P}}$ (the $i$'th level of the polynomial time hierarchy).[3] The extractor of Trevisan and Vadhan [TV00] relies on the strong assumption that $\mathsf{E}$ is hard for exponential size $\Sigma_6$-circuits (which can be viewed as a scaled, nonuniform version of the widely believed assumption that $\mathsf{EXP} \neq \Sigma_6^{\mathsf{P}}$).[4]

**Previous work on extractors for samplable distributions.** The main result of Trevisan and Vadhan [TV00] is that under a hardness assumption for $\Sigma_6$-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.[5]

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

[2] A precise definition of nondeterministic circuits appears in Section 2.3.

[3] A $\Sigma_i$-circuit is a nonuniform analogue of the class $P^{\Sigma_i^{\mathsf{P}}}$ that contains $\Sigma_i^{\mathsf{P}}$, and recall that $\mathsf{P} = \Sigma_0^{\mathsf{P}}$ and $\mathsf{NP} = \Sigma_1^{\mathsf{P}}$. See Section 2.3 for a formal definition.

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

[5] 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 $\Sigma_5$-circuits.

**Theorem 1.3** ([TV00]). *If* $\mathsf{E}$ *is hard for exponential size* $\Sigma_6$*-circuits then for every sufficiently small constant* $\gamma > 0$*, every constant* $c > 1$*, and for every sufficiently large* $n$*, there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{(1-O(\gamma))\cdot n}$ *that is a* $((1-\gamma)\cdot n, \epsilon)$*-extractor for distributions samplable by circuits of size* $n^c$*, where* $\epsilon = n^{-c}$*. Furthermore,* $\mathsf{Ext}$ *is computable in time* $\mathsf{poly}(n^c)$.

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

- Ball et al. [BGDM23] achieved the conclusion of Theorem 1.3 under the weaker, and more standard assumption that $\mathsf{E}$ is hard for exponential size nondeterministic circuits.
- Applebaum et al. [AASY15] and Shaltiel [Sha24] considered extractors with "multiplicative error", meaning that the extractor satisfies the stronger guarantee that for every output string $z \in \{0,1\}^m$, $\Pr[\mathsf{Ext}(X) = z] \leq e^{\epsilon} \cdot 2^{-m}$. See Definition and discussion in Section 2.2.1. Very recently, Shaltiel [Sha24] constructed an extractor with multiplicative error (for slightly shorter output length) under the weaker assumption used by Ball et al. [BGDM23]. See [AASY15, Sha24] for a discussion on multiplicative extractors.

Since the original paper of Trevisan and Vadhan [TV00] there was no improvement in the min-entropy threshold of explicit extractors for samplable distributions. All constructions of explicit extractors work only when the min-entropy threshold $k$ is very large, and apply only to sources with very high min-entropy of $k = (1-\gamma)\cdot n$ for some constant $\gamma > 0$. As we will explain in detail in Section 1.3, the technique of [TV00] and subsequent work, cannot give extractors with $k < \frac{n}{2}$.

## 1.2 Our Results

In this paper, we give the first explicit construction of extractors for samplable distributions with low min-entropy. We are able to construct extractors with $k \ll n/2$, and achieve extractors with $k = n^{1-\gamma}$, where $\gamma > 0$ is some small constant.

**Theorem 1.4** (Extractor for samplable distributions with low min-entropy). *There exists a constant* $\gamma > 0$ *such that if* $\mathsf{E}$ *is hard for exponential size* $\Sigma_5$*-circuits, then for every constants* $c > 1$*, every constant* $\alpha > 0$*, every sufficiently large* $n$*, and every* $k \geq n^{1-\gamma}$*, there is a* $(k, \epsilon)$*-extractor* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{(1-\alpha)\cdot k}$ *for distributions samplable by circuits of size* $n^c$*, where* $\epsilon = n^{-c}$*. Furthermore,* $\mathsf{Ext}$ *can be computed in time* $\mathsf{poly}(n^c)$.

We remark that there are barriers preventing current techniques from achieving extractors with $\epsilon = n^{-\omega(1)}$, see [AASY15] for precise details.

The assumption used in Theorem 1.4 is weaker than that used in Theorem 1.3, and stronger than that used in [BGDM23, Sha24]. The past success of [BGDM23] in weakening the original hardness assumption of [TV00] gives hope that this may be doable also for the low min-entropy case. The reader is referred to [Sha24] for a discussion regarding the necessity of hardness for nondeterministic circuits for extractors for samplable distributions. We remark that a less modular argument could have replaced $\Sigma_5$-circuits with $\Sigma_4$-circuits in Theorem 1.4, see Remark 4.11.

While Theorem 1.4 does not achieve multiplicative extractors, our technique can potentially yield multiplicative extractors, and the missing component is an explicit construction of a low-error 2-source extractor for low min-entropy, see Remark 4.9 and Section 5.

We also obtain extractors with $m = (1-o(1))\cdot k$. for every $k \geq n^{1-\gamma}$. (This yields an improved output length over Theorem 1.3 even for large values of $k$).

**Theorem 1.5** (Extractor for samplable distributions with larger output length, and higher error). *There exists a constant* $\gamma > 0$ *such that if* $\mathsf{E}$ *is hard for exponential size* $\Sigma_5$*-circuits, then for every constant* $c > 1$*,* *every constant $0 < \eta < 1$, every constant $b > 1$, every sufficiently large $n$, and every $k \geq n^{1-\gamma}$, there is a $(k, \epsilon)$-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* poly$(n^c)$.

In Theorem 1.5, the improved output length comes with a cost of a larger error $\epsilon$. We remark that our techniques can potentially achieve output length $m = (1 - o(1)) \cdot k$ with error $\epsilon = n^{-c}$ (as 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 $O(\log n)$. See Section 5 for more details.

**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. Previous extractors for samplable distributions could only handle sources with very large min-entropy. Our extractors extend the usefulness of extractors for samplable distributions to sources with low 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 proofs and do not build on the content of this section. The reader can skip to the technical section if they wish. In Section 1.3.1 we give a high-level overview of the construction and proof of [TV00], and explain why this construction does not work for low min-entropy. In Section 1.3.2 we explain our construction (which relies on two new components that we introduce: an errorless condenser for samplable distributions, and a function that is hard on average on samplable distributions (HOS)). In Section 1.3.3 we explain how we construct the new components.

### 1.3.1 An Overview of the Construction of Trevisan and Vadhan

The construction of Trevisan and Vadhan uses average-case hard functions, and 2-source extractors.

**Definition 1.6** (average-case hard functions)**.** *We say that a function $f : \{0,1\}^n \to \{0,1\}^m$ is $\rho$-**hard** for a class $\mathcal{C}$, if for every $C : \{0,1\}^n \to \{0,1\}^m$ in $\mathcal{C}$, $\Pr_{X \leftarrow U_n}[C(X) = f(X)] \leq \rho$.*

**Definition 1.7** (Two-source extractors)**.** *A function* TExt $: \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *is a $(k_1, k_2, \epsilon)$-**2-source extractor** if for every two independent distributions $X_1, X_2$ with $H_\infty(X_1) \geq k_1$ and $H_\infty(X_2) \geq k_2$,* TExt$(X_1, X_2)$ *is $\epsilon$-close to $U_m$.*

Given parameters $n$ and $c$, we aim to construct a $(k, \epsilon)$-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ for distributions samplable by size $n^c$ circuits, with $\epsilon = n^{-c}$ and as large as possible $k$. For simplicity, we will consider the case that $m = 1$. Let $n_1 = n/2$. Trevisan and Vadhan start by showing that the hardness assumption implies a function $\hat{f} : \{0,1\}^{n_1} \to \{0,1\}^{n_1}$ such that:

- $\hat{f}$ is $2^{-\Omega(n)}$-hard for $\Sigma_2$-circuits of size $n^{c'}$ for a constant $c'$ somewhat larger than $c$.
- $\hat{f}$ is computable in time poly$(n^{c'})$.

This construction is a major contribution of [TV00], which we will not survey in detail here.[6] The second ingredient used by [TV00] is an explicit $(k_1', k_2', \epsilon')$-2-source extractor TExt $: \{0,1\}^{n_1} \times \{0,1\}^n \to \{0,1\}$, where $\epsilon' = \frac{\epsilon}{16} = \frac{n^{-c}}{16}$, and we will explain how to choose $k_1'$ and $k_2'$ later.

[6] This result uses the "low-degree extension" a.k.a "Reed-Muller code" that was previously used by Sudan, Trevisan and Vadhan [STV01] to yield "local list-decoding". Loosely speaking, the contribution of [TV00] is showing that the local list-decoding algorithm of [STV01] can be sped up using nondeterminism, and this in turn means that by assuming a strong hardness assumption against $\Sigma_i$-circuits, one can obtain a $\rho$-hard function which is computable in time poly$(n)$, with very small $\rho = 2^{-\Omega(n)}$. The reader is referred to [AASY15] for a discussion of the power of nondeterministic reductions.

Given $x \in \{0,1\}^n$, let $\mathsf{Left}(x)$ denote the first $n/2$ bits of $x$, and $\mathsf{Right}(x)$ denote the last $n/2$ bits of $x$, so that $x = (\mathsf{Left}(x), \mathsf{Right}(x))$. The final extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ is defined by:

$$\mathsf{Ext}(x) = \mathsf{TExt}(\hat{f}(\mathsf{Left}(x)), x). \tag{1}$$

We remark that Trevisan and Vadhan use a slightly different construction than the one we state here.[7]

**The role of 2-source extractors.** For a samplable distribution $X$ with $H_\infty(X) \geq k$, it is not the case that $\mathsf{Left}(X)$ and $X$ are independent. Therefore, it is not immediately clear why a 2-source extractor should be helpful. Indeed, the property of 2-source extractors that is used in the analysis is the following:

**Proposition 1.8** (List-decoding view of 2-source extractors)**.** *If* $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^n \to \{0,1\}^m$ *is a* $(k'_1, k'_2, \epsilon')$*-2-source extractor, then for every* $z \in \{0,1\}^m$*, and every distribution* $W$ *over* $\{0,1\}^n$ *such that* $H_\infty(W) \geq k'_2$*, the set* $T_{W,z} := \{v \in \{0,1\}^{n_1} : \Pr[\mathsf{TExt}(v, W) = z] > 2^{-m} + \epsilon'\}$ *satisfies* $|T_{W,z}| \leq 2^{k'_1}$.

Proposition 1.8 says that for every distribution $W$ that has min-entropy at least $k'_2$ (the min-entropy requirement of the second source) there is only a small number of strings $v \in \{0,1\}^{n_1}$, such that the distribution $\mathsf{TExt}(v, W)$ is biased. Proposition 1.8 follows by noting that if it does not hold, then taking $R$ to be the uniform distribution over $T_{W,z}$, one obtains two independent distributions $R, W$, that contradict the guarantee of 2-source extractors. We will now explain how Trevisan and Vadhan use this property in the analysis.

**The analysis of Trevisan and Vadhan.** We assume that $\mathsf{Ext}$ is not a $(k, \epsilon)$-extractor for distributions of size $n^c$, and will get a contradiction by showing that there is a $\Sigma_2$-circuit $C$ of size $n^{c'}$ that breaks the hardness of $\hat{f}$. Our starting point is that there exists a distribution $X$ with $H_\infty(X) \geq k$ that is samplable by a size $n^c$ circuit, and $\mathsf{Ext}(X)$ is not $\epsilon$-close to uniform. This gives that there exists a $z \in \{0,1\}^m$ such that $\Pr[\mathsf{Ext}(X) = z] > 2^{-m} + \epsilon$, and we have that:

$$\Pr[\mathsf{TExt}(\hat{f}(\mathsf{Left}(X)), X) = z] > 2^{-m} + \epsilon.$$

Let $Y = \mathsf{Left}(X)$. We will say that a string $y \in \{0,1\}^{n_1}$ is *useful* if:

- $H_\infty(X|Y = y) \geq k'_2$. (I.e. conditioning on y leaves a lot entropy in $X$).
- $\Pr[\mathsf{TExt}(\hat{f}(y), X) = z | Y = y] > 2^{-m} + \epsilon'$. (I.e. conditioning on y biases the extractor).

This definition is made so that for every useful $y \in \{0,1\}^{n_1}$, if we denote the distribution $X$ conditioned on the event $\{Y = y\}$, by $W_y = (X|Y = y)$, it follows that:

- $\hat{f}(y) \in T_{W_y,z}$. (This follows directly from the second item, and the definition of $T_{W_y,z}$).
- $|T_{W_y,z}| \leq 2^{k'_1}$. (This follows from the first item and Proposition 1.8).

Recall that our goal is to construct a size $n^{c'}$ $\Sigma_2$-circuit $C$ that given $y \in \{0,1\}^{n_1}$ is able to compute $\hat{f}(y)$ "too well" and contradict the hardness of $\hat{f}$. The two items above give that for every useful $y \in \{0,1\}^{n_1}$, the desired output $\hat{f}(y)$ is in the set $T_{W_y,z}$ which is small. This means that on a useful $y$, a circuit $C$ can output $\hat{f}(y)$ with probability $2^{-k'_1}$, if it can sample a uniform element from $T_{W_y,z}$. Trevisan and Vadhan used classical results on approximate counting and uniform sampling of NP witnesses [Sto83, Sip83, JVV86, BGP00] in order to show that a (randomized) $\Sigma_2$-circuit $C$ of size $\mathsf{poly}(n^c) = n^{c'}$ can indeed sample a uniform element from $T_{W_y,z}$ when given a useful $y$. We will not present this argument in detail in this overview.[8]

[7] Trevisan and Vadhan use the construction $\mathsf{Ext}(x) = \mathsf{TExt}(\hat{f}(\mathsf{Left}(x)), \mathsf{Right}(x))$. However, their analysis also applies to the construction we state here if one has an approprate 2-source extractor (which they didn't have at the time). We present the modified construction, as it will be easier to explain and extend. In the construction of Trevisan and Vadhan [TV00], $\hat{f}$ is the low-degree extension, and $\mathsf{TExt}$ is the "Hadamard extractor" of Chor and Goldreich [CG88]. With these choices, the construction of Trevisan and Vadhan is very natural and corresponds to concatenating the Reed-Muller code with the Hadamard code.

[8] On a high level, this is achieved in two steps. In the first step, Trevisan and Vadhan show that for every useful $y$, there is a small

**A barrier at $k = n/2$.** We would like to argue that there are many useful $y$. Note that the first requirement in the definition of a useful $y$ requires that $H_\infty(X|Y = y) \geq k'_2$. In order to guarantee that there exists even a single $y \in \{0,1\}^{n_1}$, such that $H_\infty(X|Y = y) \geq 0$, Trevisan and Vadhan have to assume that $k = H_\infty(X) \geq n/2$. Otherwise, it could be that $X$ is uniformly distributed on the first $n_1 = n/2$ bits, so that $Y = \mathsf{Left}(X)$ is uniformly distributed and for every $y \in \{0,1\}^{n_1}$, $H_\infty(X|Y = y) = 0$.

**Choosing the threshold $k'_2$.** While this rules out choosing $k < \frac{n}{2}$, it turns out that if one chooses $k$ to be sufficiently larger than $n_1 = n/2$, then a simple averaging argument shows that:

$$k'_2 \leq k - n_1 - O(\log n) \Rightarrow \Pr[Y \text{ is useful}] \geq \frac{\epsilon}{8}. \tag{2}$$

This means that by choosing $k$ to be sufficiently larger than $n_1 = n/2$, and choosing $k'_2 = k - n_1 - O(\log n)$, we obtain that the size $n^{c'}$ $\Sigma_2$-circuit $C$ satisfies:

$$\Pr[C(Y) = \hat{f}(Y)] \geq \frac{\epsilon}{8} \cdot 2^{-k'_1}. \tag{3}$$

Recall that $\hat{f}$ is hard on average on the uniform distribution $U_{n_1}$, whereas the statement in (3) discusses a distribution $Y = \mathsf{Left}(X)$, which is not necessarily uniform. In order to contradict the hardness of $\hat{f}$, Trevisan and Vadhan need to obtain a version of (3) on the uniform distribution. They observe that for every $\Delta > 0$, if we assume that $H_\infty(X) \geq k \geq n - \Delta$, then we can conclude that $H_\infty(\mathsf{Left}(X)) \geq n_1 - \Delta$ (as only $\Delta$ bits of min-entropy can be "missing" from $Y = \mathsf{Left}(X)$) and this allows to convert (3) to the uniform distribution by paying a penalty of $2^{-\Delta}$ in the success probability. This follows because for every event $A \subseteq \{0,1\}^{n_1}$, and for every distribution $Y$ with $H_\infty(Y) \geq n_1 - \Delta$, it holds that $\Pr[U_{n_1} \in A] \geq 2^{-\Delta} \cdot \Pr[Y \in A]$. More specifically, we get that if $k \geq n - \Delta$ then:

$$\Pr_{V \leftarrow U_{n_1}}[C(V) = \hat{f}(V)] \geq \frac{\epsilon}{8} \cdot 2^{-k'_1} \cdot 2^{-\Delta}. \tag{4}$$

**Choosing the threshold $k'_1$.** We can set the parameters so that this gives a contradiction to the hardness of $\hat{f}$. For this purpose we choose $\Delta = \gamma \cdot n$ for a sufficiently small constant $\gamma > 0$ (and this will dictate that we can only get an extractor for very high min-entropy of $k = n - \Delta = (1-\gamma) \cdot n$, as in Theorem 1.3). We also need to set $k'_1 = \eta \cdot n$ for a sufficiently small constant $\eta > 0$, so that the probability in (4) is sufficiently large to contradict the hardness of $\hat{f}$. This concludes the argument of Trevisan and Vadhan [TV00].

### 1.3.2 Our Construction of Extractors for Samplable Distributions with Low Min-Entropy

We aim to use the same proof structure as the one used by Trevisan and Vadhan [TV00]. However, we would like to bypass the barriers, and achieve an extractor for low min-entropy threshold $k = n^{1-\gamma}$ for some constant $\gamma > 0$ (as guaranteed in Theorem 1.4). We will need two new components.

$\Sigma_1$-circuit $C_y$ that on input $v \in \{0,1\}^{n_1}$ can decide whether $v \in T_{W_y,z}$. This circuit works by using "approximate counting of NP witnesses" (see Section 2.7 for a precise statement) to approximate $\Pr[\mathsf{TExt}(v, W_y) = z | Y = y]$, and compare it to the threshold in the definition of $T_{W_y,z}$. In the second step, Trevisan and Vadhan use "uniform sampling of NP witnesses" (see Section 2.7) to obtain a $\Sigma_2$-circuit $C$ that given $y \in \{0,1\}^{n_1}$ samples a uniform element from $\{v : C_y(v) = 1\}$. In our formal proof, we repeat this argument, and the proof appears in Section 4. We remark that in this argument, the circuit $C$ that is obtained has size $\mathsf{poly}(n^c, \frac{1}{\epsilon})$, and as we aim for a $\mathsf{poly}(n)$-size circuit, this dictates that $\epsilon \geq n^{-\Omega(1)}$.

**Errorless condenser for samplable distributions.** We will replace the function $\mathsf{Left} : \{0,1\}^n \to \{0,1\}^{n_1}$ in the construction of Trevisan and Vadhan that is specified in (1) with an "errorless condenser for samplable distributions", which we now define.

**Definition 1.9** (Errorless condenser)**.** *A function* $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n_1}$ *is a* $(k,k')$-**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* $H_\infty(X) \geq k$*,* $H_\infty(\mathsf{Cnd}(X)) \geq k'$*.*

One of the contributions of this paper is showing that the hardness assumption implies the following errorless condenser.

**Theorem 1.10** (Errorless condenser)**.** *There exists a constant* $\gamma > 0$ *such that if* $\mathsf{E}$ *is hard for exponential size* $\Sigma_3$*-circuits, then for every constant* $c > 1$*, and every sufficiently large* $n$*, there is a function* $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n^{0.9}}$ *that is an* $(n^{1-\gamma}, n^{0.7})$*-errorless condenser for distributions samplable by circuits of size* $n^c$*. Furthermore,* $\mathsf{Cnd}$ *can be computed in time* $\mathsf{poly}(n^c)$*.*

Note that this is not a "condenser" as typically construed because the entropy rate may decrease (that is, the entropy rate $k'/n_1$ of the output distribution is smaller than the entropy rate $k/n$ of the input distribution). Theorem 1.10 is stated in a more general way in Theorem 3.1. We give a high level overview of the proof of Theorem 1.10 in Section 1.3.3.

Recall that in the construction of Trevisan and Vadhan, we had to choose $k \geq n/2$, because we had to guarantee that both $Y = \mathsf{Left}(X)$, and $X$ conditioned on $Y$, have high min-entropy. We will modify the construction of Trevisan and Vadhan, described in (1) as follows: We choose $n_1 = n^{0.9}$ rather than $n_1 = n/2$, and use $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n_1}$ instead of Left. That is, we define:

$$\mathsf{Ext}(x) = \mathsf{TExt}(\hat{f}(\mathsf{Cnd}(x)), x). \tag{5}$$

For $k = n^{1-\gamma}$, and a samplable distribution $X$ with $H_\infty(X) \geq k$, we now have that $Y = \mathsf{Cnd}(X)$ has $H_\infty(Y) \geq n^{0.7}$, and furthermore because $Y$ is now shorter than $k$, it cannot "steal all the entropy of $X$", and therefore we expect $X$ to have roughly $k - n_1 = \Omega(k)$ bits of min-entropy, even conditioned on $Y$.

This observation gives that for this modified construction, we can repeat the argument described in Section 1.3.1, exactly as before, and obtain that by choosing $k_2' \leq k - n_1 - O(\log n)$ (as instructed in (2), which will be easy to satisfy) we can obtain a $\Sigma_2$-circuit $C$ of size $n^{c'}$ that satisfies (3), and specifically,

$$\Pr[C(Y) = \hat{f}(Y)] \geq \frac{\epsilon}{8} \cdot 2^{-k_1'}. \tag{6}$$

As we explained in Section 1.3.1, at this point, the argument of Trevisan and Vadhan relies on the fact that they choose $k$ to be extremely large, to convert (3) which discusses success probability on $Y$, to (4) which considers success probability on the uniform distribution on $\{0,1\}^{n_1}$, and obtain a contradiction to the hardness of $\hat{f}$. We cannot afford this step, which requires $k$ to be very large.

**Functions that are hard on samplable distributions with sufficiently high min-entropy (HOS).** We will replace the function $\hat{f}$ with a different function Hrd, which we will set up to be "sufficiently hard" so that (6) is *already* a contradiction to its hardness, without having to move to the uniform distribution. This leads to the following notion of functions that are hard on samplable distributions with sufficiently high min-entropy, which we abbreviate by HOS.

**Definition 1.11** (A function that is hard on samplable distributions (HOS))**.** *A function* $\mathsf{Hrd} : \{0,1\}^n \to \{0,1\}^m$ *is an* $(s,k,\rho)$*-HOS for a class* $\mathcal{C}$*, if for every distribution* $Y$ *over* $\{0,1\}^n$ *that is samplable by size* $s$ *circuits, with* $H_\infty(Y) \geq k$*, and every function* $C : \{0,1\}^n \to \{0,1\}^m$ *in* $\mathcal{C}$*,* $\Pr[C(Y) = \mathsf{Hrd}(Y)] \leq \rho$*.*

Note that an HOS is a generalization of a hard on average function, in the sense that rather than being hard on average only on the uniform distribution, an HOS is hard on average on every samplable distribution with sufficiently high min-entropy. One of the contributions of this paper is showing that the hardness assumption implies the following HOS.

**Theorem 1.12** (HOS for $\Sigma_2$-circuits)**.** *There exists a constant $a_0 > 1$ such that if $\mathsf{E}$ is hard for exponential size $\Sigma_4$-circuits, then for every constant $c > 1$, and every sufficiently large $n$, there is a function $\mathsf{Hrd} : \{0,1\}^{n^{0.9}} \to \{0,1\}^{n^{a_0}}$ that is an $(n^c, n^{0.7}, 2^{-\Omega(n^{0.7})})$-HOS for size $n^c$ $\Sigma_2$-circuits. Furthermore, $\mathsf{Hrd}$ is computable in time $\mathsf{poly}(n^c)$.*

Theorem 1.10 is stated in a more general way in Theorem 3.2. We give a high level overview of the proof of Theorem 1.12 in Section 1.3.3.

**Complexity leveraging against the errorless condenser.** Recall that we want that (6) is a contradiction to the hardness of $\mathsf{Hrd}$. To achieve this, we will need to argue that $Y = \mathsf{Cnd}(X)$ is a samplable distribution on which the guarantee of $\mathsf{Hrd}$ applies. We indeed do have that $H_\infty(Y) \geq n^{0.7}$ which is the required min-entropy threshold. We will also need that $Y$ is efficiently samplable. This seems problematic as $\mathsf{Cnd}$ itself is a type of "hard function" against circuits of size $n^c$, and we do not have that $Y = \mathsf{Cnd}(X)$ is samplable by size $n^c$ circuits.

Fortunately, we can use "complexity leveraging" against $\mathsf{Cnd}$. More specifically, we have that $\mathsf{Cnd}$ can be computed in time $\mathsf{poly}(n^c) = n^{d_{\mathsf{Cnd}}}$ for some constant $d_{\mathsf{Cnd}} > c$. Recall that $X$ is a distribution that is samplable by size $n^c$ circuits, and has $H_\infty(X) \geq k$. It follows that the distribution $Y = \mathsf{Cnd}(X)$ is samplable by circuits of size $n^{d_{\mathsf{Cnd}}} + n^c$, and has $H_\infty(Y) \geq n^{0.7}$.

We will apply Theorem 1.12 with a sufficiently large constant $c' > d_{\mathsf{Cnd}}$, to obtain a function $\mathsf{Hrd} : \{0,1\}^{n_1 = n^{0.9}} \to \{0,1\}^{n^{a_0}}$ that is an $(n^{c'}, n^{0.7}, 2^{-\Omega(n^{0.7})})$-HOS for size $n^{c'}$ $\Sigma_2$-circuits, and is computable in time $\mathsf{poly}(n^{c'}) = \mathsf{poly}(n^c)$. (Loosely speaking, this creates a "hierarchy of hardness" where the size $n^c$ circuit that samples $X$ is at the bottom, $\mathsf{Cnd}$ is harder, and $\mathsf{Hrd}$ is hard even against $\mathsf{Cnd}$).

We modify the construction of $\mathsf{Ext}$ in (5) and replace the function $\hat{f}$ with $\mathsf{Hrd}$. We can repeat the argument of Section 1.3.1 to obtain the $\Sigma_2$-circuit $C$. (The circuit $C$ will now also run $\mathsf{Cnd}$ as a subroutine, and is therefore of size $\mathsf{poly}(n^{d_{\mathsf{Cnd}}})$). We get that (6) holds, and specifically that

$$\Pr[C(Y) = \mathsf{Hrd}(Y)] \geq \frac{\epsilon}{8} \cdot 2^{-k_1'} \tag{7}$$

By choosing $c'$ to be sufficiently large, and choosing $k_1' = n^{0.6}$, so that $\frac{\epsilon}{8} \cdot 2^{-k_1'} \geq 2^{-\Omega(n^{0.7})}$, we get that $C$ contradicts the hardness of $\mathsf{Hrd}$. More specifically, that on the distribution $Y = \mathsf{Cnd}(X)$ that is samplable by circuits of size $n^c + n^{d_{\mathsf{Cnd}}} \leq n^{c'}$, the size $n^{c'}$ $\Sigma_2$-circuit $C$, satisfies $\Pr[C(Y) = \mathsf{Hrd}(Y)] \geq \frac{\epsilon}{8} \cdot 2^{-k_1'} \geq 2^{-\Omega(n^{0.7})}$. This gives a contradiction to the HOS guarantee of $\mathsf{Hrd}$ and shows the correctness of the constructed extractor.

**The final extractor construction for one bit output.** We now combine the two ideas together, choose parameters, and review the argument. Let $\gamma > 0$ be the constant from Theorem 1.10 and let $k = n^{1-\gamma}$. We will construct a $(k, \epsilon)$-extractor, $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ for distributions samplable by size $n^c$ circuits, with $\epsilon = n^{-c}$ as follows:

- Let $n_1 = n^{0.9}$, and Let $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n_1}$ be the $(k, n^{0.7})$-errorless condenser of Theorem 1.10.
- We have that $\mathsf{Cnd}$ can be computed in time $n^{d_{\mathsf{Cnd}}}$ for some constant $d_{\mathsf{Cnd}} > c$. Let $c' > d_{\mathsf{Cnd}}$ be a sufficiently large constant. Let $\mathsf{Hrd} : \{0,1\}^{n^{0.9}} \to \{0,1\}^{n^{a_0}}$ be an $(n^{c'}, n^{0.7}, 2^{-\Omega(n^{0.7})})$-HOS for size $n^{c'}$ $\Sigma_2$-circuits, which we get from Theorem 1.12.

- Let $k_1' = k_2' = n^{0.6}$ and let $\mathsf{TExt} : \{0,1\}^{n^{a_0}} \times \{0,1\}^n \to \{0,1\}$ be a $(k_1', k_2', \epsilon')$-2-source extractor for $\epsilon' = \epsilon/16$. By the breakthrough result of Chattopadhyay and Zuckerman [CZ16], we can obtain such an explicit 2-source extractor.[9]

The extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ is given by:

$$\mathsf{Ext}(x) = \mathsf{TExt}(\mathsf{Hrd}(\mathsf{Cnd}(x)), x). \tag{8}$$

With these choices, we can repeat the argument outlined in Section 1.3.1. A full formal construction and proof for this argument appear in Section 4.

**Constructing extractors with large output length.** In Section 4 we extend the argument described above for larger output length $m$. However, in order to obtain an extractor with output length $m$, we need to take the error of the 2-source extractor $\mathsf{TExt}$ to be $\epsilon' = \frac{\epsilon}{8 \cdot 2^m}$. Unfortunately, the recent explicit constructions of low min-entropy 2-source extractors (specifically, [CZ16] and subsequent work) do not achieve $\epsilon' = n^{-\omega(1)}$. This means that using current 2-source extractors we can only obtain extractors for samplable distributions with output length $m = O(\log n)$. This result is stated in Theorem 4.8. We remark that our approach can give larger $m$ directly, if low error explicit constructions of 2-source extractors are achieved.

We get extractors for samplable distributions with large output length (as stated in Theorem 1.4 and Theorem 1.5) using a general transformation by Shaltiel [Sha08]. Shaltiel [Sha08] showed that one can boost an extractor for samplable distributions that extracts $O(\log n)$ bits, into one that extracts almost all the min-entropy from sources with slightly larger min-entropy. This transformation is described in Section 4.4 and yields Theorem 1.4 and Theorem 1.5.[10].

### 1.3.3 Constructing an Errorless Condenser and an HOS

In the previous section we explained how to construct an extractor for samplable distributions with low min-entropy using two components: an errorless condenser, and an HOS. In this section we give an overview to the proofs of Theorem 1.10 and Theorem 1.12. Both components are similar to extractors for samplable distributions in the sense that they expect their input distribution $X$ to be a samplable distribution with sufficient min-entropy, and can be thought of as "hard functions". (This is clear for an HOS, and note that by definition, a $(k, k')$-errorless condenser for distributions samplable by size $s$ circuits is equivalent to an $(s, k, 2^{-k'})$-HOS for the class $\mathcal{C}$ of *constant functions*).

Our construction will rely on recent work by Shaltiel and Silbak [SS24] that show how to construct a variant of hard function which they termed HTS, for "hard to sample". We will discuss these functions below. Our construction works as follows:

- We observe that an HTS with certain parameters implies both errorless condensers and HOS. The parameters that we aim for, are different, and more challenging than those achieved by the constructions of [SS24].
- Nevertheless, by using tools and ideas from [SS24] (as well as additional ideas that we will explain below) we are able to construct a good enough HTS, under the hardness assumptions of Theorem 1.10 and Theorem 1.12.

[9] A technicality is that because the length of the first source is $n^{a_0}$ that is larger than the length of the second source which is $n$, we need to pad the second source to length $n^{a_0}$. This is not a problem, as the min-entropy thresholds $k_1' = k_2' = n^{0.6}$, are large when thought of as a function of the source length $n^{a_0}$, and recall that the 2-source extractor of [CZ16] works even when the min-entropy is logarithmic in the source length.

[10] As stated in [Sha08], this transformation comes with the cost of strengthening the hardness assumption from hardness for $\Sigma_i$-circuits, to hardness for $\Sigma_{i+1}$-circuits. However, a more careful proof of both our result, and Shaltiel's transformation can avoid this cost. See Remark 4.11.

**Functions that are hard to sample (HTS).** Functions that are hard to sample (rather than just hard to compute) have been studied in several contexts in recent years (see e.g., [AST+98, Vio12, Vio14, Vio20]). Inspired by the work of Viola [Vio12], Shaltiel and Silbak [SS24] considered the following variant of functions that are hard to sample: Let $f : \{0,1\}^n \to \{0,1\}^m$, and fix some class $\mathcal{A}$ of "sampling procedures" (as in Definition 1.2) that sample a distribution $(X,Y) \in \{0,1\}^n \times \{0,1\}^m$ (for concreteness, let us focus on the case that $\mathcal{A}$ is the class of circuits of size $n^c$).

- We say that $f$ is a $(k,\rho)$-min-entropy HTS for $\mathcal{A}$, if for every $A \in \mathcal{A}$, if $H_\infty(X) \geq k$, then $\Pr[Y = f(X)] \leq \rho$.
- We say that $f$ is a $(k,\rho)$-min-entropy cHTS, if the requirement above holds for $A \in \mathcal{A}$, for which the distribution $Y$ is fixed to some constant value $y$ (rather than being a random variable).

The definition of min-entropy HTS that we give here is different than the definition used by Shaltiel and Silbak [SS24]. Shaltiel and Silbak [SS24] used a different definition, that we refer to as a "small-set HTS" in this paper. A precise definition of the two notions can be found in Section 3.1.[11]

**Errorless condensers and HOS follow from HTS/cHTS.** It immediately follows that a min-entropy HTS/cHTS implies the components that we want to construct.

**Proposition 1.13** (HTS implies errorless condenser and HOS)**.**

- *A $(k, 2^{-k'})$-min-entropy cHTS for circuits of size $s$ is in particular an $(s, k, 2^{-k'})$-HOS for the class of constant functions, and therefore (as explained above) is also a $(k, k')$-errorless condenser for distributions samplable by size $s$ circuits.*
- *A $(k,\rho)$-min-entropy HTS for $\Sigma_2$-circuits of size $2s$ is an $(s,k,\rho)$-HOS for $\Sigma_2$-circuits of size $s$.*

Both implications above trivially follow because if some procedure $C$ computes $f$ too well on some samplable distribution $X$ with $H_\infty(X) \geq k$, then $f$ is not an HTS/cHTS, by considering the sampling circuit $A$ that samples $X$, and computes $Y = C(X)$. See Section 3.1.2 for a more formal proof.

**Constructing the errorless condenser of Theorem 1.10.** By Proposition 1.13, in order to prove Theorem 1.10, it is sufficient to explicitly construct a $(k = n^{1-\gamma}, 2^{-n^{0.7}})$-min-entropy cHTS $f : \{0,1\}^n \to \{0,1\}^{n^{0.9}}$ for circuits of size $n^c$, under the hardness assumption.

We will construct $f$ in two steps. In the first step, we will obtain a min-entropy cHTS $f_1$ (that will based on multiplicative extractors for samplable distributions) and will only work for very high min-entropy. In the second, we will show how to reduce the min-entropy threshold of $f_1$ using an approach of Shaltiel and Silbak [SS24] that is based on strong seeded dispersers.

**A min-entropy cHTS for large min-entropy.** For some constant $\alpha > 0$, we set $n_1 = n^{0.99}$, $m_1 = n^{0.8}$, and will construct a $(k_1 = (1-\alpha)\cdot n_1, \rho_1 = 2^{-(m_1-1)})$-min-entropy cHTS $f_1 : \{0,1\}^{n_1} \to \{0,1\}^{m_1}$. Note that here, the min-entropy threshold is very high, namely $k_1 = (1-\alpha)\cdot n_1$. This is the range of parameters for which we already have extractors for samplable distributions. We will take the function $f_1$ to be the multiplicative extractor for samplable distribution of Shaltiel [Sha24]. As explained in Section 1, the guarantee of this multiplicative extractor is that for every samplable distribution $X$ with $H_\infty(X) \geq k_1$, and every $z \in \{0,1\}^{m_1}$, $\Pr[f_1(X) = z] \leq e^{\epsilon} \cdot 2^{-m_1} \leq 2^{-(m_1-1)}$, if we take $\epsilon = \frac{1}{2}$. By definition, we obtain that $f_1$ is indeed a $(k_1,\rho_1)$-min-entropy cHTS. (Note that here, it is crucial to have multiplicative error, as otherwise, we only get that $\rho_1 = n_1^{-\Theta(1)}$ which is not good enough for our purposes).

[11] Shaltiel and Silbak [SS24] used an HTS as a tool to construct error-correcting codes for computationally bounded channels. In this setting, the computationally bounded adversary $A$ is not necessarily required to act in a way that leads to a distribution $X$ that has large min-entropy. For that reason, Shaltiel and Silbak considered a stronger notion that applies to every adversary $A$, and asks that for every $A$ there exists a small set $H$ of inputs, such that it is unlikely that $A$ samples $(X,Y)$ such that $Y = f(X)$ and $X \notin H$.

**Reducing the min-entropy threshold of a cHTS.** Shaltiel and Silbak [SS24] showed how to reduce the min-entropy threshold of an HTS. The technique of Shaltiel and Silbak is designed to start with a different variant of HTS (which is the one defined in [SS24]) which we call "small-set HTS". We will not define this variant in this high level overview, and a formal definition is given in Section 3.1.

The notion of small-set HTS is stronger than a min-entropy HTS. More specifically, while it is obvious that a small-set HTS is a min-entropy HTS (with a slight loss in parameters), we do not know whether this applies in the other direction. Nevertheless, we show that if a function is a min-entropy HTS (resp. cHTS) for $\Sigma_2$-circuits (and not just for deterministic circuits) then it *is* a small-set HTS (resp. cHTS) for deterministic circuits (of slightly smaller size, and with a slight loss in parameters). See Lemma 3.7 for a precise formulation.

This means that in order to use the min-entropy reduction of Shaltiel and Silbak, we need to set $f_1$ to be a min-entropy HTS, not just against deterministic circuits, but rather for the stronger class of $\Sigma_2$-circuits. We can achieve this by replacing the hardness assumption assumed in [Sha24] (that is against $\Sigma_1$-circuits) with an assumption against $\Sigma_{1+2}$-circuits. Indeed, this is why Theorem 1.10 uses a hardness assumption against $\Sigma_3$-circuits. Summing up, under the hardness assumption of Theorem 1.10 we can make sure that $f_1$ is an HTS of the type for which the min-entropy reduction of Shaltiel and Silbak applies.

We now describe the construction of Shaltiel and Silbak. In this high level overview, we will cheat and ignore the difference between a min-entropy HTS and a small-set HTS, and simply call a function an HTS or cHTS, without mentioning the precise variant. This will make things easier to explain, and a precise formal argument appears in Section 3.

The construction of Shaltiel and Silbak [SS24] relies on strong seeded dispersers.[12] A strong seeded $(k, \epsilon)$-disperser is a function $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ such that for every distribution $X$ over $\{0,1\}^n$ with $H_\infty(X) \geq k$, if we take $Y \leftarrow U_d$, then the support of the distribution $(Y, E(X,Y))$ is of size at least $(1-\epsilon) \cdot 2^{m+d}$.

Shaltiel and Silbak [SS24], suggested the following construction: Let $f_1 : \{0,1\}^{n_1} \to \{0,1\}^{m_1}$ be a $(k_1, \rho_1)$-cHTS, and let $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a strong seeded $(k, \epsilon)$-disperser, where $m = n_1$, so that the output length of $E$ and the input length of $f_1$ coincide. It is also required that $\epsilon \leq 1 - 2^{k_1 - m}$ (and note that this is an unusual, and weak requirement, as $\epsilon$ is allowed to be close to $1$ rather than close to $0$). Let $D = 2^d$ be the "degree" of the disperser, and think of $E$ as a function $E : \{0,1\}^n \times [D] \to \{0,1\}^m$. Shaltiel and Silbak [SS24] show that under these conditions, the function $f : \{0,1\}^n \to \{0,1\}^{D \cdot m_1}$ defined by

$$f(x) = f_1(E(x,1)), \ldots, f_1(E(x,D)),$$

is a $(k, \rho_1 \cdot D)$-HTS.

Loosely speaking, this means that one can start from a function $f_1$, where the min-entropy threshold $k_1$ is very large compared to the input length $n_1$ (in our case $k_1 = (1-\alpha)n_1$) and obtain a function $f$ on $n > n_1$ bits, where the min-entropy threshold $k$, is small compared to $n$.

Loosely speaking, this is where it is more beneficial to work with an HTS than with extractors for samplable distributions. More specifically, while we do not have a good technique to reduce the min-entropy threshold of an extractor for samplable distributions, the construction above *is* able to reduce the min-entropy threshold of an HTS/cHTS, and this enables us to construct HTS/cHTS for distributions with low min-entropy, which in turn translate into an errorless condenser, and an HOS, and can be used to construct extractors for samplable distributions with low min-entropy by the approach explained in Section 1.3.2.

[12] We remark that in [SS24], and also in our formal proof in Section 3, the construction is described in terms of list-recoverable codes, rather than strong seeded dispersers, and then the required list-recoverable codes are constructed using strong seeded dispersers. In this high level overview we will describe the construction in the terminology of strong dispersers, which will be more natural in terms of parameters, and easier to explain.

We now use this approach to prove Theorem 1.10. More specifically, we have already explained how to get a $((1-\alpha)\cdot n_1, 2^{-(m_1-1)})$-cHTS $f_1 : \{0,1\}^{n_1=n^{0.99}} \to \{0,1\}^{m_1=n^{0.8}}$. In order to obtain our desired $(k=n^{1-\gamma}, 2^{-n^{0.7}})$-cHTS $f : \{0,1\}^n \to \{0,1\}^{n^{0.9}}$, by the explanation above, we need a strong $(k,\epsilon)$-seeded disperser $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $k = n^{1-\gamma}$, $\epsilon \leq 1 - 2^{k_1 - m} = 1 - 2^{-\alpha \cdot m}$ and $D = 2^d \leq n^{0.1}$.

The requirement that $D \leq n^{0.1}$ seems like a problem. This is because Radhakrishnan and Ta-Shma [RTS00] showed that if $\epsilon < \frac{1}{2}$, then seeded dispersers must have degree $D \geq n - k = (1-o(1))\cdot n$. (Note that with $D \geq n$ we cannot hope to use this transformation to obtain a function $f$ where the output length $D \cdot m_1$ is shorter than the input length $n$, making this approach unsuitable to construct an errorless condenser).

Fortunately, for the case of high-error dispersers with $\epsilon = 1 - o(1)$, Zuckerman [Zuc07] gave an explicit construction of strong seeded $(k,\epsilon)$-dispersers which for a sufficiently small constant $\gamma > 0$ and $k = n^{1-\gamma}$, obtains the required $\epsilon \leq 1 - 2^{k_1 - m} = 1 - 2^{-\alpha \cdot m}$ with degree $D \leq n^{0.1}$. (See Theorem 2.10 for a precise statement). Using this strong seeded disperser (and accounting correctly for the difference between a min-entropy cHTS and a small-set cHTS) we obtain that $f$ is a $(k = n^{1-\gamma}, 2^{-n^{0.7}})$-cHTS as required. The precise argument appears in Section 3.

**Constructing the HOS of Theorem 1.12.** We have already observed in Proposition 3.5 that the desired HOS of Theorem 1.12 follows if we can construct an HTS with certain parameters.

We will construct such an HTS using the same approach used to prove Theorem 1.10. Namely, we will start from an intial HTS $f_1$ for high min-entropy, and reduce the min-entropy threshold using strong seeded dispersers. Note however that the setting here is somewhat different than that of Theorem 1.10.

- Here we aim to construct an HTS rather than a cHTS. It is harder to construct an HTS than a cHTS, and in particular, we cannot use the same approach that was used to get the starting function $f_1$ in the proof of Theorem 1.10. Instead, we will use an HTS $f_1$ that is based on the function $\hat{f}$ of Trevisan and Vadhan that was described in Section 1.3.1. We will not go into details in this high level overview.
- On the other hand, in Theorem 1.12 we aim to construct a function $f$ with output length that is *longer* than the input length. This makes our life significantly easier, as we do not have to use strong seeded dispersers with small degree $D$, and are allowed to use strong seeded dispersers with seed length $d = O(\log n)$ which translates to degree $D = n^{O(1)}$. This means that we can use strong seeded dispersers (or in fact seeded extractors) with error $\epsilon = \frac{1}{2}$, and such explicit constructions are known for every choice of $k$, [LRVW03, GUV07]. The fact that we are allowed to construct a function $f$ with a large output length, is also fortunate as the function $f_1$ that we use in this argument has large output length, and this would have prevented us from obtaining a final function $f$ where the output length if shorter than the input length, even if we used strong seeded dispersers with small $D$.

The precise details appear in Section 3.

# 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.

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

Two distributions $X, Y$ over the same finite domain $S$ are $\epsilon$-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 use the following standard lemma.

**Lemma 2.1.** *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 $H_\infty(X|Y = y) \geq k - m - \log \frac{1}{\eta}$.*

## 2.1 Samplable Distributions

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

**Definition 2.2** (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 $Z$.*

## 2.2 Extractors

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

### 2.2.1 Seedless Extractors

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

**Definition 2.3** (Seedless extractor)**.** *A function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ is a $(k, \epsilon)$-***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 $\epsilon$-close to $U_m$.*

The following "multiplicative variant" was defined in [AASY15, Sha24].

**Definition 2.4** (Multiplicative seedless extractor)**.** *A function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ is a $(k, \epsilon)$-***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 for every $z \in \{0,1\}^m$,*

$$\Pr[\mathsf{Ext}(X) = z] \leq e^\epsilon \cdot 2^{-m}.$$

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

**Proposition 2.5** (Multiplicative extractors imply standard extractors)**.** *For every class $\mathcal{D}$ and $0 < \epsilon \leq 1$, a $(k, \epsilon)$-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 $\epsilon = \frac{1}{2}$, 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].

### 2.2.2 Errorless Condensers

In Definition 1.9, which appears in Section 1.3.2 we introduced a notion of an errorless condenser, and we repeat it here.

**Definition 2.6** (Errorless condenser)**.** *A function* $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n_1}$ *is a* $(k, k')$**-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* $H_\infty(X) \geq k$*,* $H_\infty(\mathsf{Cnd}(X)) \geq k'$*.*

### 2.2.3 Seeded Extractors and Dispersers

We need the following standard definition of strong seeded extractors and dispersers.

**Definition 2.7** (Strong extractors and dispersers)**.** *A function* $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a strong* $(k, \epsilon)$*-extractor if for every distribution* $X$ *over* $\{0,1\}^n$ *with* $H_\infty(X) \geq k$*, the distribution* $Z = (Y, E(X, Y))$ *where* $Y \leftarrow U_d$ *is* $\epsilon$ *close to* $U_{m+d}$*.* $E$ *is a strong* $(k, \epsilon)$*-disperser if the support size of* $Z$ *is at least* $(1 - \epsilon) \cdot 2^{m+d}$*.*

We now state several explicit constructions of extractors and dispersers.

**Theorem 2.8** (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* $(k, \epsilon)$*-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 2.9** (Strong extractors with logarithmic seed and larger output length [TU12])**.** *For every constants* $0 < \eta < 1$ *and* $b \geq 1$*, there exists a constant* $c_1$ *such that for every sufficiently large* $n$*, and every* $k \geq 2^{c_1 \cdot \log^\eta n}$*, 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* $\mathsf{poly}(n)$*.*

**Theorem 2.10** (Strong high-error seeded dispersers with very short seed length [Zuc07])**.** *There exist constant* $c_1, c_2$ *such that for every functions* $\delta(n), s(n)$ *and every sufficiently large* $n$*, there is a strong* $(\delta(n) \cdot n, 1 - s(n))$*-disperser* $E : \{0,1\}^n \times \{0,1\}^{\log D} \times \{0,1\}^m$ *for* $D = O(\frac{n}{\delta(n)^{c_1} \cdot \log \frac{1}{s(n)}})$ *and* $m = O(\delta(n)^{c_2} n)$*. Furthermore,* $E$ *can be computed in time* $\mathsf{poly}(n)$*.*

### 2.2.4 Two-Source Extractors

We repeat the standard definition of 2-source extractors, that appeared in Section 1 as Definition 1.7.

**Definition 2.11** (Two-source extractors)**.** *A function* $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *is a* $(k_1, k_2, \epsilon)$*-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* $\epsilon$*-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 2.12** ([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,* $\mathsf{TExt}$ *can be computed in time* $\mathsf{poly}(n^{c_1})$*.*

Theorem 2.12 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 $c_2$ (rather than any constant $c_2$). 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].

### 2.3 Definition of Circuits of Various Types

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

**Definition 2.13** (randomized circuits, nondeterministic circuits, oracle circuits and $\Sigma_i$-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 $\Sigma_i$-circuit is an $A$-circuit where $A$ is the canonical $\Sigma_i^P$-complete language. The size of all circuits is the total number of wires and gates.*[13]

### 2.4 Impagliazzo-Wigderson Style Hardness Assumptions

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

**Definition 2.14** ($\mathsf{E}$ is hard for exponential size circuits)**.** *We say that "$\mathsf{E}$ 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 2.15** (Ladder Climbing)**.** *The assumption that $\mathsf{E}$ is hard for exponential size $\Sigma_i$ 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 $\mathsf{E}$ is hard for exponential size $\Sigma_{i+j}$ circuits then the construction gives functions that are secure against $\Sigma_j$-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 $\Sigma_j^{\mathsf{P}}$-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.*

[13] 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 $s^{\Theta(1)}$ $\Sigma_i$-circuit is equivalent to $\mathsf{DTIME}^{\Sigma_i^P}(s^{\Theta(1)})/s^{\Theta(1)}$.

## 2.5 Functions that are Hard on Average

We repeat the definition of functions that are hard on average, that appeared in Section 1 as Definition 1.6.

**Definition 2.16** (Average-case hard functions)**.** *We say that a function $f : \{0,1\}^n \to \{0,1\}^m$ is $\rho$-**hard** for a class $\mathcal{C}$, if for every $C : \{0,1\}^n \to \{0,1\}^m$ in $\mathcal{C}$,*

$$\Pr_{X \leftarrow U_n}[C(X) = f(X)] \le \rho.$$

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

In Definition 1.11 in Section 1.3.2, we introduced the notion of functions that are hard on samplable distributions. We repeat this definition below.

**Definition 2.17** (A function that is hard on samplable distributions (HOS))**.** *A function $\mathsf{Hrd} : \{0,1\}^n \to \{0,1\}^m$ is an $(s,k,\rho)$-HOS for a class $\mathcal{C}$, if for every distribution $Y$ over $\{0,1\}^n$ that is samplable by size $s$ circuits, with $H_\infty(Y) \ge k$, and every function $C : \{0,1\}^n \to \{0,1\}^m$ in $\mathcal{C}$,*

$$\Pr[C(Y) = \mathsf{Hrd}(Y)] \le \rho.$$

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

It is useful to note that if $0 \le p \le 1$ is an $\epsilon$-relative approximation to $q$, then it is also an additive approximation to $q$. For $\epsilon \le \frac{1}{2}$, we also have the following: If $p$ is an $\epsilon$-relative approximation to $q$, then $q$ is an $O(\epsilon)$-relative approximation to $p$. If $p$ is an $\epsilon$-relative approximation to $q$ and $q$ is an $\epsilon$-relative approximation to $w$, then $p$ is an $O(\epsilon)$-relative approximation to $w$. If $p'$ is an $\epsilon$-relative approximation to $p$ and $q'$ is an $\epsilon$-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 2.19** (Approximate counting [Sto83, Sip83, JVV86])**.** *For every $i$, every sufficiently large $s$, and every $\epsilon > 0$, there is a size $\mathsf{poly}(s/\epsilon)$ $\Sigma_{i+1}$-circuit that given a size $s$ $\Sigma_i$-circuit $C$, outputs an $\epsilon$-relative approximation of $|\{x : C(x) = 1\}|$.*

**Theorem 2.20** (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 $\Sigma_{i+1}$-circuit $A$ that given a size $s$ $\Sigma_i$-circuit $C : \{0,1\}^n \to \{0,1\}$, outputs a value in $\{0,1\}^n \cup \bot$ such that $\Pr[A(C) = \bot] \le \delta$ and the distribution $(A(C)|A(C) \ne \bot)$ is uniform over $\{x : C(x) = 1\}$.*

**Regarding the formulation of Theorems 2.19 and 2.20.** We state Theorems 2.19 and Theorem 2.20 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 2.20 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 2.19 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 $\mathsf{BPP} \subseteq \mathsf{P/poly}$ which extends to $\mathsf{BPP}^{\mathsf{NP}} \subseteq \mathsf{P}^{\mathsf{NP}}/\mathsf{poly}$.

## 2.8 List-Recoverable Codes

We will use the following less standard definition of list-recoverable codes.

**Definition 2.21** (List-recoverable codes)**.** *A function* $\mathsf{LR} : \{0,1\}^n \to (\{0,1\}^m)^D$ *is* $(\ell, L)$**-list-recoverable** *if for every* $S_1, \ldots, S_D \subseteq \{0,1\}^m$*, such that for every* $i \in [D]$*,* $|S_j| \le \ell$*, the set*

$$\mathsf{List}_{\mathsf{LR}}(S_1, \ldots, S_D) = \{x \in \{0,1\}^n : \forall i \in [D] : \mathsf{LR}(x)_i \in S_i\},$$

*is of size at most* $L$*.*

We remark that the definition above is for an "errorless version" of list-recoverable codes. In the more general setting, the definition of the set $\mathsf{List}_{\mathsf{LR}}(S_1, \ldots, S_D)$ has an additional "agreement parameter", measuring the fraction of $i$, for which $\mathsf{LR}(x)_i \in S_i$.

Ta-Shma and Zuckerman [TZ04] observed that there is a tight connection between extractors and (the more general version) of list-recoverable codes. In the proposition below, we state a version of this connection for errorless list-recoverable codes and strong dispersers.

**Proposition 2.22** (Strong dispersers give list-recoverable codes [TZ04])**.** *For a function* $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$*, we define* $D = 2^d$*, and* $\mathsf{LR} : \{0,1\}^n \to (\{0,1\}^m)^D$ *by* $\mathsf{LR}(x)_i = E(x,i)$*. If* $E$ *is a strong* $(k, \epsilon)$*-disperser, then* $\mathsf{LR}$ *is* $(\ell, 2^k)$*-list-recoverable for every* $\ell < (1-\epsilon) \cdot 2^m$*.*

*Proof.* If $\mathsf{LR}$ is not $(\ell, 2^k)$-list-recoverable, then there exist $S_1, \ldots, S_D \subseteq \{0,1\}^m$, such that each of the sets is of size $\le \ell$, and the set $A = \mathsf{List}_{\mathsf{LR}}(S_1, \ldots, S_D)$ is of size at larger than $2^k$. Let $X$ be the uniform distribution over $A$, and note that $H_\infty(X) \ge k$. Let $Z = (Y, E(X,Y))$ where $Y \leftarrow U_d$. By definition the support of $Z$ is contained in the set $T = \left\{(i,w) : i \in \{0,1\}^d,\ w \in S_i\right\}$ which is a set of size $\ell \cdot 2^d$, and we get a contradiction if $\ell < (1-\epsilon) \cdot 2^m$. □

This means that the strong dispersers of Theorem 2.10 and Theorem 2.8 can be stated as list-recoverable codes. Specifically, using Theorem 2.10 we get the following Corollary.

**Corollary 2.23** (List-Recoverable code with small block-length)**.** *There exists a constant* $\gamma > 0$ *such that for every sufficiently large* $n$*, there is an* $(\ell, 2^{n^{1-\gamma}})$*-list recoverable code* $\mathsf{LR} : \{0,1\}^n \to (\{0,1\}^m)^D$ *with* $n^{0.99} \le m \le n$*,* $\ell = 2^{(1-\frac{2}{n^\gamma})\cdot m}$*, and* $D \le n^{0.1}$*. Furthermore,* $\mathsf{LR}$ *can be computed in time* $\mathsf{poly}(n)$*.*

*Proof.* Let $\gamma > 0$ be a constant that we will choose later. We apply Theorem 2.10 choosing $\delta(n) = n^{-\gamma}$ (so that $\delta(n) \cdot n = n^{1-\gamma}$). Note that $m = O(\delta(n)^{c_2} \cdot n) = O(n^{1-c_2\cdot\gamma})$ is determined, and we can choose $s(n) = 2^{-\frac{m}{n^\gamma}}$. We obtain a strong $(n^{1-\gamma}, 1-2^{-\frac{m}{n^\gamma}})$-disperser $E : \{0,1\}^n \times [D] \to \{0,1\}^m$ for

$$D = O\Big(\frac{n}{\delta(n)^{c_1} \cdot \log\frac{1}{s(n)}}\Big) = O\Big(\frac{n \cdot n^{c_1\cdot\gamma} \cdot n^\gamma}{n^{1-c_2\cdot\gamma}}\Big) = O(n^{(c_1+c_2+1)\cdot\gamma}).$$

We can choose the constant $\gamma > 0$ to be sufficiently small so that for sufficiently large $n$, we have that $D \le n^{0.1}$ and $m \ge n^{0.99}$. Applying Proposition 2.22 we get that $E$ can be viewed as a function $\mathsf{LR} : \{0,1\}^n \to (\{0,1\}^m)^D$ that is $(\ell, 2^{n^{1-\gamma}})$-list recoverable for every $\ell < s(n)\cdot 2^m = 2^{(1-\frac{1}{n^\gamma})\cdot m}$ and in particular for $\ell = 2^{(1-\frac{2}{n^\gamma})\cdot m}$. □

Using Theorem 2.8 we get the following corollary.

**Corollary 2.24** (List-Recoverable code with small $L$)**.** *There exist constant* $c_0, c_1$ *such that for every sufficiently large* $n$*, every* $k \ge \log^{c_1} n$*, and* $m \le \frac{k}{2}$ *there is a* $(2^{m-1}, 2^k)$*-list recoverable code* $\mathsf{LR} : \{0,1\}^n \to (\{0,1\}^m)^D$ *with* $D = n^{c_0}$*. Furthermore,* $\mathsf{LR}$ *can be computed in time* $\mathsf{poly}(n)$*.*

*Proof.* We use Theorem 2.8 to obtain a strong $(k, \frac{1}{4})$-disperser $E : \{0,1\}^n \times \{0,1\}^{c_0\cdot\log n} \to \{0,1\}^m$, where $c_0$ is a universal constant. This translates into the required list-recoverable code, using Proposition 2.22. □

# 3 A Construction of Errorless Condensers and HOS

As explained in Section 1.3, our construction of extractor for samplable distributions will rely on two new components: an errorless condenser, and an HOS. In this section we construct these two objects.

The next theorem states the constructions of errorless condensers, and generalizes Theorem 1.10 from Section 1.3.2.

**Theorem 3.1** (Errorless condenser)**.** *There exists a constant $\gamma > 0$ such that if $\mathsf{E}$ is hard for exponential size $\Sigma_3$-circuits, then for every constant $c > 1$, every sufficiently large $n$, and every $m \leq n$, there is a function $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^m$ that is an $(n^{1-\gamma}, \frac{m}{n^{0.1}} - 0.1 \cdot \log n - 3)$-errorless condenser for distributions samplable by circuits of size $n^c$. Furthermore, $\mathsf{Cnd}$ can be computed in time $\mathsf{poly}(n^c)$.*

Theorem 1.10 follows by setting $m = n^{0.8}$ in Theorem 3.1.

The next theorem states the constructions of HOS, and generalizes Theorem 1.12 from Section 1.3.2.

**Theorem 3.2** (HOS for $\Sigma_2$-circuits)**.** *There exists a constant $a_0 > 1$ such that if $\mathsf{E}$ is hard for exponential size $\Sigma_4$-circuits, then for every constant $c > 1$, every constant $\nu > 0$, every sufficiently large $n$, and every $k \geq 4n^{\nu}$ there is a function $\mathsf{Hrd} : \{0,1\}^n \to \{0,1\}^{n^{a_0}}$ that is an $(n^c, k, 2^{-\Omega(k)})$-HOS for size $n^c$ $\Sigma_2$-circuits. Furthermore, $f$ is computable in time $\mathsf{poly}(n^c)$.*

Theorem 1.12 guarantees a function $\mathsf{Hrd} : \{0,1\}^{n^{0.9}} \to \{0,1\}^{n^{a_0}}$ with $k = n^{0.7}$ and $\rho = 2^{-\Omega(k)}$ and easily follows from Theorem 3.2.

**Outline for this section.** As explained in Section 1.3.3, the proofs of Theorem 3.1 and Theorem 3.2 will use techniques developed by Shaltiel and Silbak [SS24] that discuss functions that are hard to sample on distribution with sufficient min-entropy (HTS). In Section 3.1 we give a formal definition of (several variants) of HTS. We also show several connections between these variants, errorless condensers and HOS. These connections will be used to prove Theorem 3.1 (which is proven in Section 3.2) and Theorem 3.2 (which is proven in Section 3.3).

## 3.1 Functions That Are Hard to Sample (HTS)

We describe the work of Shaltiel and Silbak on [SS24], that will be used in our constructions.

### 3.1.1 Definition of Functions That Are Hard to Sample (HTS)

Complexity theory is mostly concerned with understanding functions that are hard to compute. In recent years, there has also been work on functions that are also hard to sample. Loosely speaking, a function $f : \{0,1\}^m \to \{0,1\}^{m'}$ is hard to sample for a class of functions $\mathcal{A}$, if for every $A \in \mathcal{A}$ that samples a distribution $Z = (X,Y)$ over $\{0,1\}^m \times \{0,1\}^{m'}$, the distribution $(X,Y) \leftarrow A$ is far in statistical distance from $(W, f(W))$ where $W \leftarrow U_m$.

Early examples are Ambainis et al. [AST$^+$98] (in the context of quantum communication complexity) and Viola [Vio12] (in the context of constant depth circuits).

Shaltiel and Silbak [SS24] were interested in defining a variant where one does not measure success on the uniform distribution $U_m$, and instead asks that for every $A \in \mathcal{A}$, $\Pr_{(X,Y)\leftarrow A}[Y = f(X)]$ is small. This does not make sense as stated, as $\mathcal{A}$ is typically a nonuniform complexity class, and therefore $A$ can be hardwired with a pair $(x,y)$ such that $y = f(x)$. This problem can be bypassed in two ways:

- One can require that $X$ has min-entropy larger than some threshold $k$. We will call this notion a min-entropy HTS.

- One can require that for every $A \in \mathcal{A}$, there exists a small set $H \subseteq \{0,1\}^m$ such that the success of $A$ in coming up with pairs $(X,Y)$ such that $Y = f(X)$ is only measured for $X \notin H$. We call this notion a small-set HTS, and this is the notion that was defined under the term HTS in [SS24].

Shaltiel and Silbak [SS24] used the second approach, as it was more suitable for their coding theoretic application. In Definition 3.3 below, we give a formal definition of both notions. In both cases we also consider a weakened variant called cHTS, in which the guarantee holds only for $A \in \mathcal{A}$ that samples a distribution $Y$ where $Y$ is fixed to some constant $y$.

**Definition 3.3** (Hard To Sample functions (HTS))**.** *Let $f : \{0,1\}^m \to \{0,1\}^{m'}$ be a function, and $\mathcal{A}$ be a class of functions.*

- *$f$ is a $(k,\rho)$-min-entropy HTS for $\mathcal{A}$, if for every $A \in \mathcal{A}$ that samples a distribution $Z = (X,Y)$ over $\{0,1\}^m \times \{0,1\}^{m'}$, with $H_\infty(X) \geq k$,*

$$\Pr_{(X,Y)\leftarrow A}[Y = f(X)] \leq \rho.$$

- *$f$ is a $(k,\rho)$-min-entropy cHTS for $\mathcal{A}$, if the requirement above holds for every $A \in \mathcal{A}$ that samples a distribution $Z = (X,Y)$ where $Y$ is a fixed value $y \in \{0,1\}^{m'}$.*
- *$f$ is an $(h,\rho)$-small-set HTS for $\mathcal{A}$, if for every $A \in \mathcal{A}$ that samples a distribution $Z = (X,Y)$ over $\{0,1\}^m \times \{0,1\}^{m'}$, there exists a set $H \subseteq \{0,1\}^m$ of size at most $h$, such that:*

$$\Pr_{(X,Y)\leftarrow A}[X \notin H \text{ and } Y = f(X)] \leq \rho.$$

- *$f$ is an $(h,\rho)$-small-set cHTS for $\mathcal{C}$, if the requirement above holds for every $A \in \mathcal{A}$ that samples a distribution $Z = (X,Y)$ where $Y$ is a fixed value $y \in \{0,1\}^{m'}$.*

It should be noted that when comparing a small-set HTS to a min-entropy HTS, the parameters are scaled differently. While in a min-entropy HTS the parameter $k$ measures min-entropy, in a small-set HTS the parameter $h$ measures set size, and a good intuition is that the two notions are comparable the same under the translation $h = 2^k$. This intuition is made more precise in Section 3.1.3.

### 3.1.2 HTS implies Errorless Condensers and HOS

We now observe that the components that we want to construct (errorless condensers and HOS) follow from min-entropy HTS/cHTS with certain parameters.

The two propositions below are a formal statement of Proposition 1.13 from Section 1.3.3.

**Proposition 3.4** (A min-entropy cHTS is an errorless condenser)**.** *If $f : \{0,1\}^n \to \{0,1\}^m$ is a $(k, 2^{-k'})$-min-entropy cHTS for size $s+m$ circuits, then $f$ is a $(k,k')$-errorless condenser for distribution samplable by size $s$ circuits.*

*Proof.* Let $f : \{0,1\}^n \to \{0,1\}^m$ be a $(k, 2^{-k'})$-min-entropy cHTS for size $s$ circuits, and let $X$ be a distribution over $\{0,1\}^n$ that is samplable by a size $s$ circuit $A$, and has $H_\infty(X) \geq k$. For every $y \in \{0,1\}^m$ we can consider the size $s+m$ sampling circuit $A_y$ that samples $X \leftarrow A$, and outputs $(X,y)$. By the guarantee of a min-entropy cHTS, we have that for every $y \in \{0,1\}^m$,

$$\Pr_{X\leftarrow A}[f(X) = y] = \Pr_{(X,y)\leftarrow A_y}[f(X) = y] \leq 2^{-k'}.$$

and we can conclude that $H_\infty(f(X')) \geq k'$. □

**Proposition 3.5** (A min-entropy HTS is an HOS)**.** *For every $i$, if $f$ is a $(k,\rho)$-min-entropy HTS for size $2s$ $\Sigma_i$-circuits, then $f$ is an $(s,k,\rho)$-HOS for $\Sigma_i$-circuits of size $s$.*

*Proof.* Let $f : \{0,1\}^n \to \{0,1\}^m$ be a $(k,\rho)$-min-entropy HTS for size $2s$ $\Sigma_i$-circuits. Let $X$ be a distribution over $\{0,1\}^n$ that is samplable by a size $s$ circuit $A$, and has $H_\infty(X) \geq k$. Let $C : \{0,1\}^n \to \{0,1\}^m$ be a size $s$, $\Sigma_i$ circuit. We consider the $\Sigma_i$ sampling circuit $B$, that samples $X \leftarrow A$ and outputs $(X, C(X))$. Note that $B$ is of size $2s$. By the guarantee of a min-entropy HTS we have that:

$$\Pr_{X \leftarrow A}[C(X) = f(X)] = \Pr_{(X,Y)\leftarrow B}[Y = f(X)] \leq \rho.$$

□

### 3.1.3 Converting Between a Small-Set HTS and a Min-Entropy HTS

In this section we discuss relations between the two variants of HTS. It is immediately clear that a small-set HTS for a class $\mathcal{A}$ is a min-entropy HTS for the same class, with a slight loss in parameters.

**Proposition 3.6** (A small-set HTS is a min-entropy HTS)**.** *If $f$ is an $(h,\rho)$-small-set HTS (resp. cHTS) for $\mathcal{A}$ then $f$ is a $(\log h + \log(1/\rho), 2\rho)$-min-entropy HTS (resp. cHTS) for $\mathcal{A}$.*

*Proof.* Let $f : \{0,1\}^m \to \{0,1\}^{m'}$ be an $(h,\rho)$-small-set HTS. For every $A \leftarrow \mathcal{A}$ that samples a distribution $(X,Y)$ such that $H_\infty(X) \geq \log h + \log(1/\rho)$, it follows that for every set $H \subseteq \{0,1\}^m$ of size $2^h$,

$$\Pr[X \in H] \leq \frac{|H|}{2^{h+\log\frac{1}{\rho}}} \leq \rho.$$

By the definition of small-set HTS, $A$ has a set $H$ of such $2^h$, such that:

$$\Pr[Y = f(X) \text{ and } X \notin H] \leq \rho.$$

It follows that:

$$\Pr[Y = f(X)] = \Pr[Y = f(X) \text{ and } X \notin H] + \Pr[Y = f(X) \text{ and } X \in H] \leq 2\rho.$$

Note that the same proof applies for the case of cHTS. □

It is not clear whether the converse is true, and a min-entropy HTS is a small-set HTS. What we can show that is that a min-entropy HTS against $\Sigma_{i+2}$ circuits, is a small-set HTS against $\Sigma_i$-circuits, with a slight loss in parameters.

**Lemma 3.7** (From min-entropy HTS to small-set HTS)**.** *For every $i$, if $f$ is a $(k,\rho)$-min-entropy HTS (resp. cHTS) for $\Sigma_{i+2}$ circuits of size $s$, then $f$ is an $(h = \frac{2^{k+3}}{\rho}, 2\rho)$-small-set HTS (resp. cHTS) for $\Sigma_i$-circuits of size $s' = s^{\Omega(1)}$.*

*Proof.* Let $A : \{0,1\}^r \to \{0,1\}^m \times \{0,1\}^{m'}$ be a size $s' \geq r$ $\Sigma_i$-circuit that samples a distribution $(X,Y)$ over $\{0,1\}^m \times \{0,1\}^{m'}$. By Theorem 2.19 there is a $\Sigma_{i+1}$-circuit $B$ of size $\mathsf{poly}(s')$ that given $x \in \{0,1\}^m$, computes a $\frac{1}{8}$-relative approximation $B(x)$ of $\Pr[A(U_r) = x]$.

Let $h = \frac{2^{k+3}}{\rho}$ and let $H = \left\{x : B(x) \geq \frac{2}{h}\right\}$. For every $x \in H$, $\Pr[A(U_r) = x] > \frac{1}{2} \cdot B(x) \geq \frac{1}{h}$, and it follows that $|H| \leq h$. We will show that $H$ is a suitable small set for the sampling circuit $A$.

If $\Pr[X \notin H] \leq \rho$ then we are done, as $\Pr[Y = f(X)$ and $X \notin H] \leq \rho \leq 2\rho$. Otherwise, we consider the distribution

$$(X', Y') = ((X, Y)|X \notin H).$$

We have that for every $x \notin H$, $\Pr[X = x] \leq \frac{4}{h}$ and therefore,

$$\Pr[X' = x] = \Pr[X = x|X \notin H] \leq \frac{\Pr[X = x]}{\Pr[X \notin H]} \leq \frac{4}{h \cdot \rho}.$$

This gives that $H_\infty(X') \geq \log h - \log \frac{1}{\rho} - 2$. We will now argue that there is a $\Sigma_{i+2}$ circuit $A'$ of size $\mathsf{poly}(s')$ that samples a distribution that is very close to $(X', Y')$.

For this purpose we observe that there is a $\Sigma_{i+1}$-circuit $C : \{0,1\}^r \rightarrow \{0,1\}$ of size $\mathsf{poly}(s')$ that answers one on $w \in \{0,1\}^r$ iff $A(w) \notin H$. The circuit $C$ simply answers one iff $B(A(w)) < \frac{2}{h}$.

We now consider the following $\Sigma_{i+1}$-circuit $C : \{0,1\}^r \rightarrow \{0,1\}$. When given input $w \in \{0,1\}^r$, $C$ outputs one iff $B(A(w)) < \frac{2}{h}$. This is a $\Sigma_{i+1}$-circuit of size $\mathsf{poly}(s')$ and this definition is made so that $(X', Y')$ is exactly the distribution obtained by sampling $W \leftarrow U_r$ and considering $(A(W)|C(W) = 1)$.

By Theorem 2.20 choosing $\delta = \frac{1}{2 \cdot 2^{s'}}$, there is a randomized $\Sigma_{i+2}$-circuit $\hat{A}$ of size $\mathsf{poly}(s', \log(1/\delta)) = \mathsf{poly}(s')$ that samples a distribution that is $\delta$-close to $(W|C(W) = 1)$ for $W \leftarrow U_r$. By taking $A'(w) = A(\hat{A}(w))$ we obtain that $A'$ is a randomized $\Sigma_{i+2}$-circuit of size $\mathsf{poly}(s')$ that samples a distribution $(X'', Y'')$ that is $\delta$-close to $(X', Y')$. We have that for every $x \in \{0,1\}^m$, $\Pr[X'' = x] \leq \Pr[X' = x] + \delta$, and we have chosen $\delta$ to be so small, that we can conclude that

$$H_\infty(X'') \geq H_\infty(X') - 1 \geq \log h - \log \frac{1}{\rho} - 3 \geq k$$

Therefore, $A'$ is a $\Sigma_{i+2}$ circuit of size $\mathsf{poly}(s') \leq s$ that is a potential adversary for $f$. We note that in the case that $Y$ is fixed (that corresponds to the case that $f$ is a cHTS) we have that $Y''$ is fixed, so that $A'$ is an adversary for a cHTS.
Overall, we conclude that:

$$\Pr[Y'' = f(X'')] \leq \rho,$$

which implies (using the statistical distance between $(X'', Y'')$ and $(X', Y')$ that:

$$\Pr[Y' = f(X')] \leq 2\rho.$$

Finally, we conclude that:

$$\begin{aligned}\Pr[X \notin H \text{ and } Y = f(X)] &= \Pr[Y = f(X)|X \notin H] \cdot \Pr[X \notin H] \\ &\leq \Pr[Y = f(X)|X \notin H] \\ &= \Pr[Y' = f(X')] \\ &\leq 2\rho.\end{aligned}$$

□

### 3.1.4 Reducing the Min-Entropy Threshold of a Small-Set HTS

Shaltiel and Silbak [SS24] showed how to take a given small-set HTS on $m$ bits, with set size $h$ that is large compared to $2^m$, and convert it into a small-set HTS on $n > m$ bits with set size $h$ that is small compared to $2^n$. Lemma 3.8 below, is a generalization of a similar Lemma of [SS24] that is proven by the same argument, and also considers $\Sigma_i$-circuit, and notes that the transformation applies also to cHTS.

**Lemma 3.8** (Improving a small-set HTS using list-recoverable codes)**.**

- *Let $f : \{0,1\}^m \to \{0,1\}^{m'}$ be an $(h,\rho)$-small-set HTS (resp. cHTS) for size $s$ $\Sigma_j$-circuits.*
- *Let $\mathsf{LR} : \{0,1\}^n \to (\{0,1\}^m)^D$ be an $(h,L)$-list-recoverable code, such that $\mathsf{LR}$ has a size $s_{\mathsf{LR}}$ circuit.*

*Then the function $f' : \{0,1\}^n \to \{0,1\}^{Dm'}$ defined by*

$$f'(x) = f(\mathsf{LR}(x)_1), \dots, f(\mathsf{LR}(x)_D)$$

*is a $(L, \rho \cdot D)$-small-set HTS (resp. cHTS) for $\Sigma_j$-circuits of size $s' = s - s_{\mathsf{LR}} - \log D$.*

*Proof.* Let $A$ be a sampling $\Sigma_j$-circuit of size $s'$ that samples a pair $(X,Y) \in \{0,1\}^n \times \{0,1\}^{Dm'}$. We think of $y$ as a sequence $Y = (Y_1, \dots Y_d)$ where for every $i \in [D]$, $Y_i \in \{0,1\}^{m'}$. For every $i \in [D]$, we can define a sampling $\Sigma_j$ circuit $A_i$, as follows:

- $A_i$ applies $A$ to sample $X, Y$.
- $A_i$ computes $X_i = \mathsf{LR}(X)_i$.
- $A_i$ outputs the pair $(X_i, Y_i) \in \{0,1\}^m \times \{0,1\}^{m'}$.

Note that by definition, for every $i \in [D]$, $A_i$ is a $\Sigma_j$-circuit of size $s' + s_{\mathsf{LR}} + \log D \le s$. Furthermore, in the case that $A$ samples $(X,Y)$ such that $Y$ is fixed (namely, $A$ is an adversary for a cHTS rather than an HTS) then for every $i \in [D]$, $A_i$ samples $(X_i, Y_i)$ such that $Y_i$ is fixed.

By the definition of small-set HTS/cHTS, for every $i \in [D]$, the sampling circuit $A_i$ has a set $H_i \subseteq \{0,1\}^m$ of size at most $h$ such that:

$$\Pr_{(X_i,Y_i) \leftarrow A_i}[X_i \notin H_i \text{ and } Y_i = f(X_i)] \le \rho.$$

By the list-recoverability of $\mathsf{LR}$, we have that the set $H = \mathsf{List}_{\mathsf{LR}}(H_1, \dots, H_D)$ is of size at most $L$. It follows that if $x \notin H$, then there exists $i \in [D]$ such that $\mathsf{LR}(x)_i \notin H_i$.

We will show that $H$ satisfies the definition of an HTS/cHTS for $A$. In the computation below, probabilities are in the experiment $(X,Y) \leftarrow A$. We have that:

$$\begin{aligned}
\Pr[X \notin H \text{ and } Y = f(X)] &= \Pr[X \notin H \text{ and } \forall i \in [D] : (X_i = \mathsf{LR}(X)_i \text{ and } Y_i = f(X_i))] \\
&\le \Pr[\exists i \in [D] : \mathsf{LR}(X)_i \notin H_i \text{ and } \forall i \in [D] : (X_i = \mathsf{LR}(X)_i \text{ and } Y_i = f(X_i))] \\
&\le \Pr[\exists i \in [D] : (\mathsf{LR}(X)_i \notin H_i \text{ and } X_i = \mathsf{LR}(X)_i \text{ and } Y_i = f(X_i))] \\
&\le \Pr[\exists i \in [D] : (X_i \notin H_i \text{ and } Y_i = f(X_i))] \\
&\le \sum_{i \in [D]} \Pr[X_i \notin H_i \text{ and } Y_i = f(X_i)] \\
&\le D \cdot \rho.
\end{aligned}$$

Here, the last inequality follows because the distribution of $(X_i, Y_i)$ in our probability space is identical to $(X_i, Y_i) \leftarrow A_i$. □

## 3.2 Proof of Theorem 3.1

We will start from a min-entropy cHTS that follows from the recent construction of multiplicative extractors for samplable distributions by Shaltiel [Sha24].

**Theorem 3.9** (A min-entropy cHTS with short output length [Sha24])**.** *For every $i$, if* $\mathsf{E}$ *is hard for exponential size $\Sigma_{i+1}$-circuits, then there exists a constant $\alpha > 0$, such that for every constant $c > 1$, every sufficiently large $n$, and for every $m \leq \alpha \cdot n$, there is a function $E : \{0,1\}^n \rightarrow \{0,1\}^m$ that is a $((1-\alpha) \cdot n, 2^{-(m-1)})$-min-entropy cHTS for $\Sigma_i$-circuits of size $n^c$. Furthermore, $E$ can be computed in time* $\mathsf{poly}(n^c)$.

We state Theorem 3.9 in our terminology using the notion of a min-entropy cHTS. In [Sha24] it is stated for $i = 0$, as a $((1-\alpha) \cdot n, \frac{1}{2})$-multiplicative extractor for distributions samplable by size $n^c$ circuits. The definition of a multiplicative extractor (see Definition 2.4) says for every distribution $X$ over $\{0,1\}^n$ with $H_\infty(X) \geq (1-\alpha) \cdot n$ that is samplable by size $n^c$ circuits, and for every $z \in \{0,1\}^m$, $\Pr[E(X) = z] \leq 2 \cdot 2^{-m} = 2^{-(m-1)}$, which indeed translates into a min-entropy cHTS for size $n^c$ circuits. The extension to $i > 0$ is by standard "ladder climbing", see Remark 2.15.

By applying Lemma 3.7 we can transform the min-entropy cHTS into a small-set cHTS, of roughly the same parameters, at the cost of assuming an assumption for $\Sigma_{i+3}$-circuits rather than $\Sigma_{i+1}$-circuits. More specifically, we obtain the following:

**Claim 3.10.** *For every $i$, if* $\mathsf{E}$ *is hard for exponential size $\Sigma_{i+3}$-circuits, then there exists a constant $\alpha > 0$, such that for every constant $c > 1$, every sufficiently large $n$, and for every $m \leq \alpha \cdot n$, there is a function $E : \{0,1\}^n \rightarrow \{0,1\}^m$ that is a $(2^{(1-\alpha)\cdot n}, 2^{-(m-2)})$-small-set cHTS for $\Sigma_i$-circuits of size $n^c$. Furthermore, $E$ can be computed in time* $\mathsf{poly}(n^c)$.

*Proof.* Given $i \geq 0$, we apply Theorem 3.9 for $i+2$. We obtain a constant $\alpha' > 0$. When shooting to get a small-set HTS for $\Sigma_i$-circuits of size $n^c$, we apply Theorem 3.9 using a constant $c'$ larger than $c$ (to account for the loss in size in Lemma 3.7). We obtain a function $E : \{0,1\}^n \rightarrow \{0,1\}^m$ that is a $((1-\alpha')n, 2^{-(m-1)})$-min-entropy cHTS for $\Sigma_{i+2}$-circuits of size $n^{c'}$. By Lemma 3.7, $E$ is also a $(\frac{2^{(1-\alpha')n+3}}{2^{-(m-1)}}, 2^{-(m-2)})$-small-set cHTS for $\Sigma_i$-circuits of size $n^{\Omega(c')} = n^c$. We will choose the constant $\alpha$ that we aim for in Claim 3.10 to be $\alpha = \frac{\alpha'}{3}$, so that for sufficiently large $n$, using the requirement that $m \leq \alpha \cdot n$ we have that

$$\frac{2^{(1-\alpha')n+3}}{2^{-(m-1)}} \leq 2^{(1-\alpha)\cdot n}.$$

□

In the small-set cHTS above we have that $h = 2^{(1-\alpha)\cdot n}$. We would like to obtain a min-entropy cHTS for $k = o(n)$ which corresponds to $h = 2^k = 2^{o(n)}$. For this purpose, we will apply Lemma 3.8 using the list-recoverable code of Corollary 2.23. This gives the following:

**Claim 3.11.** *There exists a constant $\gamma > 0$ such that if* $\mathsf{E}$ *is hard for exponential size $\Sigma_3$-circuits, then for every constant $c > 1$, every sufficiently large $n$, and every $m' \leq n^{0.98}$, there is a function $f : \{0,1\}^n \rightarrow \{0,1\}^{n^{0.1}\cdot m'}$ that is a $(2^{n^{1-\gamma}}, 4n^{0.1} \cdot 2^{-m'})$-small-set cHTS for circuits of size $n^c$. Furthermore, $f$ can be computed in time* $\mathsf{poly}(n^c)$.

*Proof.* Let $\gamma > 0$ be the constant from Corollary 2.23. Given an integer $n$, let $\mathsf{LR} : \{0,1\}^n \rightarrow (\{0,1\}^m)^D$ be the $(\ell, 2^{n^{1-\gamma}})$-list recoverable code from Corollary 2.23. We have that $n^{0.99} \leq m \leq n$, $\ell = 2^{(1-\frac{2}{n^\gamma})\cdot m}$. and $D \leq n^{0.1}$. Furthermore, $\mathsf{LR}$ can be computed in time $\mathsf{poly}(m) = \mathsf{poly}(n)$.

We apply Claim 3.10 to obtain a function $E : \{0,1\}^m \rightarrow \{0,1\}^{m'}$. More specifically, when shooting for a constant $c$, we will apply Claim 3.10 using $i = 0$, constant $c' > c$ that we will choose later, input length $m$ and any output length $m' \leq n^{0.98} \leq \alpha n$. We obtain a function $E : \{0,1\}^m \rightarrow \{0,1\}^{m'}$ that is a $(2^{(1-\alpha)\cdot m}, 2^{-(m'-2)})$-small-set cHTS for circuits of size $m^{c'}$.

We now apply Lemma 3.8 using $\mathsf{LR}$ and the function $E$. Note that:

- $\ell = 2^{(1-\frac{2}{n^{\gamma}})\cdot m} \leq 2^{(1-\alpha)m}$.
- $\mathsf{LR}$ can be computed by a circuit of size $n^{c_{\mathsf{LR}}}$ for some universal constant $c_{\mathsf{LR}}$.

This means that we can indeed apply Lemma 3.8 to obtain a function $f : \{0,1\}^n \to \{0,1\}^{Dm'}$ that is a $(2^{n^{1-\gamma}}, D \cdot 2^{-(m'-2)})$-small-set cHTS for circuits of size $m^{c'} - n^{c_{\mathsf{LR}}} - \log D \geq n^c$ (where for the inequality to hold we choose $c'$ to be sufficiently large and recall that $m \geq n^{0.99}$). Note that we indeed have that $D \cdot 2^{-(m'-2)} \leq 4n^{0.1} \cdot 2^{-m'}$. □

Theorem 3.1 immediately follows from Claim 3.11 by applying Proposition 3.6 to convert the small-set cHTS into a min-entropy cHTS, and Proposition 3.4 to argue that the latter is an errorless condenser. There are slight losses in the parameters in these conversions, and these are overcome by halving the constant $\gamma$.

### 3.3 Proof of Theorem 3.2

Our starting point is the following Theorem of Trevisan and Vadhan [TV00].

**Theorem 3.12** ([TV00][14])**.** *For every $i$, if $\mathsf{E}$ is hard for exponential size $\Sigma_{i+1}$-circuits, then there exists a constant $\alpha > 0$, such that for every constant $c > 1$, and every sufficiently large $n$, there is a function $f : \{0,1\}^n \to \{0,1\}^m$ that is $\epsilon$-hard by size $n^c$ $\Sigma_i$-circuits for $m = \alpha n$ and $\epsilon = 2^{-(m/3)} = 2^{-\Omega(n)}$. Furthermore, $f$ is computable in time* $\mathsf{poly}(n^c)$.

We use the following Lemma by Shaltiel and Silbak [SS24] that shows that an $\epsilon$-hard function for $\Sigma_{i+1}$-circuits is a small-set HTS with $h \approx \epsilon \cdot 2^m$ for $\Sigma_i$-circuits.

**Lemma 3.13** (HTS from functions that are average-case hard for $\Sigma_1$-circuits)**.** *For every $i$, if $f : \{0,1\}^m \to \{0,1\}^{m'}$ is $\epsilon$-hard for size $s$ $\Sigma_{i+1}$-circuits, then $f$ is an $(h,\rho)$-small-set HTS for $\Sigma_i$-circuits of size $s' = \frac{s^{\Omega(1)}}{\log(1/\rho)}$, where $h = \epsilon^{\frac{1}{2}} \cdot 2^m$, and $\rho = 32 \cdot \epsilon^{\frac{1}{4}}$.*

An immediate corollary of Theorem 3.12 and Lemma 3.13 is the following.

**Claim 3.14.** *If $\mathsf{E}$ is hard for exponential size $\Sigma_4$-circuits, then there exist constants $\alpha > \alpha' > 0$, such that for every constant $c > 1$, and for every sufficiently large $n$, there is a function $f : \{0,1\}^n \to \{0,1\}^{\alpha n}$ that is a $(2^{(1-\alpha')\cdot n}, 2^{-\Omega(n)})$-small-set HTS for size $n^c$ $\Sigma_2$-circuits. Furthermore, $f$ is computable in time* $\mathsf{poly}(n^c)$.

In the small-set HTS above we have that $h = 2^{(1-\alpha')\cdot n}$. We would like to obtain a min-entropy cHTS for smaller $k$ which corresponds to $h = 2^k = 2^{o(n)}$. For this purpose, we will apply Lemma 3.8 using the list-recoverable code of Corollary 2.23. This gives the following:

**Claim 3.15.** *There exists a constant $a_0 > 1$ such that if $\mathsf{E}$ is hard for exponential size $\Sigma_4$-circuits, then for every constants $c > 1$, $\nu > 0$, every sufficiently large $n$, and every $k \geq 2n^{\nu}$ there is a function $f : \{0,1\}^n \to \{0,1\}^{n^{a_0}}$ that is a $(2^k, 2^{-\Omega(k)})$-small-set HTS for size $n^c$ $\Sigma_2$-circuits. Furthermore, $f$ is computable in time* $\mathsf{poly}(n^c)$.

*Proof.* Let $a_0 = c_0 + 1$ where $c_0$ is the constant from Corollary 2.24. Given $c$, $\nu$, a sufficiently large $n$ and $k \geq 2n^{\nu}$, we choose $m = k/2$, and let $\mathsf{LR} : \{0,1\}^n \to (\{0,1\}^m)^D$ be the $(2^{m-1}, 2^k)$-list recoverable code of Corollary 2.24, and we have that $D = n^{c_0}$, and that $\mathsf{LR}$ can be computed by a circuit of size $n^{c_{\mathsf{LR}}}$.

We apply Claim 3.14 to obtain a function $f : \{0,1\}^m \to \{0,1\}^{m'=\alpha m}$ that is a $(2^{(1-\alpha')\cdot m}, 2^{-\Omega(m)})$ for circuits of size $n^c + n^{c_{\mathsf{LR}}} + n^{c_0}$ and note that as $m = k/2 \geq n^{\nu}$, we can handle circuits of this size, as this size is smaller $m^{c'}$ for a constant $c'$ that depends on $c, c_{\mathsf{LR}}, c_0$ and $\nu$.

[14] Theorem 3.12 is not stated in this form in [TV00]. Nevertheless, it directly follows from [TV00]. See [AIKS16] (Section 7) for an explanation. The statement that we give here also uses "ladder climbing". See remark 2.15.

We apply Lemma 3.8 using the list-recoverable code LR. Note that we indeed have that $h = 2^{(1-\alpha')\cdot m} \leq 2^{m-1}$.[15] We obtain a function $f' : \{0,1\}^n \to \{0,1\}^{Dm'}$ that is a $(2^k, D \cdot 2^{-\Omega(m)})$-small-set HTS for circuits of size $n^c$. We note that $Dm' \leq n^{c_0+1} = n^{a_0}$ and $D \cdot 2^{-\Omega(m)} = n^{c_0} \cdot 2^{-\Omega(k)} = 2^{-\Omega(k)}$, by the requirement that $k \geq 2 \cdot m^{\nu}$, for sufficiently large $n$. □

Using Proposition 3.6 we can get a min-entropy HTS with essentially the same parameters. Specifically, we obtain:

**Claim 3.16.** *There exists a constant $a_0 > 1$ such that if* E *is hard for exponential size $\Sigma_4$-circuits, then for every constants $c > 1$, $\nu > 0$, every sufficiently large $n$, and every $k \geq 4n^{\nu}$ there is a function $f : \{0,1\}^n \to \{0,1\}^{n^{a_0}}$ that is a $(k, 2^{-\Omega(k)})$-min-entropy HTS for size $n^c$ $\Sigma_2$-circuits. Furthermore, $f$ is computable in time* $\mathsf{poly}(n^c)$.

By Proposition 3.5 such an HTS is an HOS with the required parameters.

# 4 Extractors for Samplable Distributions with Low Min-Entropy

In this section we present our construction of extractors for samplable distributions with low entropy and prove Theorem 1.4 and Theorem 1.5. In Section 4.1 we present our construction (that was described informally in Section 1.3) assuming certain ingredients are provided. In Section 4.2 we prove a general theorem showing that our construction gives an extractor for samplable distributions. In Section 4.3 we plug in the errorless condenser and HOS that we constructed in Section 3 into the extractor construction, and derive an extractor with small output length. Finally, in Section 4.4 we describe a general transformation by Shaltiel [Sha08] that transforms extractors with small output length into extractors with large output length, and use this transformation to prove Theorem 1.4 and Theorem 1.5.

## 4.1 The Construction

We now present our construction. The construction is specified in Figure 1. Unlike the presentation of Section 1.3.2 we state the constructions for ingredients with general parameters (and only choose the precise parameters later on in Section 4.3). This more general presentation will be used later to argue that future improvements in the parameters of the ingredients, will result in an improved construction. See Remark 4.9 for a discussion.

The correctness of the construction is established in the next theorem.

**Theorem 4.1.** *For every $n, k, c$ and $c_1$, if the ingredients* Cnd, Hrd, TExt *specified in Figure 1 satisfy the specified requirements, then the function* Ext $: \{0,1\}^n \to \{0,1\}^m$ *defined in Figure 1, is a $(k, \frac{1}{n^{c_1}})$-extractor for distributions samplable by circuits of size $n^c$.*

The proof of Theorem 4.1 is given in Section 4.2.

## 4.2 Proof of Theorem 4.1

Let us denote $\epsilon = \frac{1}{n^{c_1}}$ and assume for contradiction that Ext is not a $(k, \epsilon)$-extractor for distributions samplable by size $n^c$ circuits. That is, that there exists $r \leq n^c$ and a circuit $A : \{0,1\}^r \to \{0,1\}^n$ of size $n^c$ that samples a distribution $X = A(U_r)$ such that $H_{\infty}(X) \geq k$, and yet, the distribution Ext$(X)$ is not $\epsilon$-close to uniform.

[15] In fact, $h$ is much smaller, and there is plenty of slack in this argument. We could have used a list-recoverable code with much smaller $\ell = 2^{(1-\alpha')\cdot m}$ for this argument.

Figure 1: Construction of extractor for samplable distributions

**Goal:** Given parameters $n, k, c, c_1$, construct a $(k, \frac{1}{n^{c_1}})$-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ for distributions samplable by size $n^c$ circuits.

**Ingredients:** We will require the following ingredients, which introduce additional internal parameters: $n_{\mathsf{Cnd}}, k_{\mathsf{Cnd}}, c_{\mathsf{Cnd}}, d_{\mathsf{Cnd}}, c_{\mathsf{Hrd}}, \rho_{\mathsf{Hrd}}, n_1, k'_1, k'_2, d'$. The precise requirements on these parameters are specified below.

- *Errorless condenser:* A function $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n_{\mathsf{Cnd}}}$ that is a $(k, k_{\mathsf{Cnd}})$-errorless condenser for distributions samplable by size $n^c$ circuits. Furthermore, we require $\mathsf{Cnd}$ can be computed by a size $n^{d_{\mathsf{Cnd}}}$ circuit.
- *HOS:* A function $\mathsf{Hrd} : \{0,1\}^{n_{\mathsf{Cnd}}} \to \{0,1\}^{n_1}$ that is a $(n^{c_{\mathsf{Hrd}}}, k_{\mathsf{Cnd}}, \rho_{\mathsf{Hrd}})$-HOS for $\Sigma_2$-circuits of size $n^{c_{\mathsf{Hrd}}}$. Namely, for every distribution $Y$ over $\{0,1\}^{n_{\mathsf{Cnd}}}$ that is samplable by a circuit of size $n^{c_{\mathsf{Hrd}}}$ circuit, and has $H_\infty(Y) \geq k_{\mathsf{Cnd}}$, and for every $\Sigma_2$-circuit $C$ of size $n^{c_{\mathsf{Hrd}}}$,
$$\Pr[C(Y) = \mathsf{Hrd}(Y)] \leq \rho_{\mathsf{Hrd}}.$$
- *2-source extractor:* A $(k'_1, k'_2, \epsilon')$-2-source extractor $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^n \to \{0,1\}^m$, where $\epsilon' = \frac{1}{8 \cdot n^{c_1} \cdot 2^m}$. We require that $\mathsf{TExt}$ can be computed cy a size $n^{d'}$ circuit.

**Requirements:** We make the following requirements:

- $k'_2 \leq k - n_{\mathsf{Cnd}} - m - c_1 \log n - 2$.
- $\rho_{\mathsf{Hrd}} < \frac{1}{8n^{c_1} \cdot 2^m \cdot 2^{k'_1}}$.
- $c_{\mathsf{Hrd}} \geq \max(c, c_1, d_{\mathsf{Cnd}}, d') + c_0$, where $c_0$ is a universal constant chosen in the proof.

**Construction:** We define $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ by:
$$\mathsf{Ext}(x) = \mathsf{TExt}(\mathsf{Hrd}(\mathsf{Cnd}(x)), x).$$

Note that $\mathsf{Ext}$ can be computed in polynomial time if $\mathsf{Cnd}, \mathsf{Hrd}, \mathsf{TExt}$ can be computed in time polynomial in $n$.

In particular, there exists an $S \subseteq \{0,1\}^m$ such that $\Pr[\mathsf{Ext}(X) \in S] > \frac{|S|}{2^m} + \epsilon > (1+\epsilon) \cdot \frac{|S|}{2^m}$, which implies that there exists a $z \in \{0,1\}^m$ such that:
$$\Pr[\mathsf{Ext}(X) = z] > (1+\epsilon) \cdot 2^{-m}. \tag{9}$$

Let $Y = \mathsf{Cnd}(X)$, and note that by the properties of $\mathsf{Cnd}$ we have that $H_\infty(Y) \geq k_{\mathsf{Cnd}}$. Our approach will be to obtain a contradiction by constructing a $\Sigma_2$-circuit $C$ of size $n^{c_{\mathsf{Hrd}}}$ which contradicts the hardness requirement of $\mathsf{Hrd}$ on the source $Y$ (which is samplable by the size $n^c + n^{d_{\mathsf{Cnd}}}$ circuit $\mathsf{Cnd} \circ A$).

At this point we have that:
$$\Pr[\mathsf{TExt}(\mathsf{Hrd}(Y), X) = z] > (1+\epsilon) \cdot 2^{-m}.$$

We say that $y \in \{0,1\}^{n_{\mathsf{Cnd}}}$ is *useful* if:

- $\Pr[\mathsf{TExt}(\mathsf{Hrd}(y), X) = z | Y = y] > (1 + \frac{\epsilon}{2}) \cdot 2^{-m}$.
- $H_\infty(X | Y = y) \geq k - n_{\mathsf{Cnd}} - m - c_1 \log n - 2$.

We are planning to construct a randomized circuit $C(y)$ that will succeed in computing $\mathsf{Hrd}(y)$ (with not too small probability) when $y$ is useful. For this purpose, we claim that the probability that $Y$ is useful is not too small.

**Claim 4.2.** $\Pr[Y$ *is useful*$] \geq \frac{\epsilon}{4 \cdot 2^m}$.

*Proof.* Let $\eta = \frac{\epsilon}{4 \cdot 2^m}$. By applying Lemma 2.1 on $X$ and $Y$, we have that with probability at least $1 - \eta$ over choosing $y \leftarrow Y$,

$$H_\infty(X|Y = y) \geq k - n_{\mathsf{Cnd}} - \log \frac{1}{\eta} = k - n_{\mathsf{Cnd}} - c_1 \cdot \log n - m - 2.$$

We have that

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

By applying an averaging argument, we get that with probability at least $\frac{\epsilon}{2 \cdot 2^m}$ over choosing $y \leftarrow Y$, we have that:

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

By a union bound, we get that with probability at least $\frac{\epsilon}{4 \cdot 2^m}$ over choosing $y \leftarrow Y$, we obtain a $y$ that satisfies the two properties and is useful. □

For every $y \in \{0,1\}^{n_{\mathsf{Cnd}}}$ and every $0 \leq \alpha \leq 1$, we define:

$$T_{y,\alpha} = \left\{ v \in \{0,1\}^{n_1} : \Pr[\mathsf{TExt}(v, X) = z | Y = y] > (1 + \alpha) \cdot 2^{-m} \right\}.$$

We now claim that whenever $y$ is useful, $\mathsf{Hrd}(y)$ is in a set that is not very large.

**Claim 4.3.** *For every useful* $y \in \{0,1\}^{n_{\mathsf{Cnd}}}$,

- $\mathsf{Hrd}(y) \in T_{y,\frac{\epsilon}{2}}$.
- $|T_{y,\frac{\epsilon}{8}}| \leq 2^{k_1'}$.

*Proof.* The first item follows immediately from the definition of a useful $y$. The second item follows from the list-decoding view of extractors in Proposition 1.8. More specifically, we assume for contradiction that $|T_{y,\frac{\epsilon}{8}}| > 2^{k_1'}$, and consider the following two independent distributions:

- $V \leftarrow T_{y,\frac{\epsilon}{8}}$, and note that $H_\infty(V) \geq k_1'$.
- $W = (X|Y = y)$ and note that as $y$ is useful, $H_\infty(W) \geq k - n_{\mathsf{Cnd}} - m - c_1 \log n - 2 \geq k_2'$, where the last inequality is by the requirement in Figure 1.

By definition, we have that for every $v \in T_{y,\frac{\epsilon}{8}}$,

$$\Pr[\mathsf{TExt}(v, W) = z] = \Pr[\mathsf{TExt}(v, X) = z | Y = y] > (1 + \frac{\epsilon}{8}) \cdot 2^{-m}.$$

Therefore, we have that:

$$\Pr[\mathsf{TExt}(V, W) = z] > (1 + \frac{\epsilon}{8}) \cdot 2^{-m} = 2^{-m} + \frac{\epsilon}{8 \cdot 2^m}.$$

However, this is a contradiction as $V$ and $W$ are independent distributions that meet the entropy thresholds $k_1', k_2'$ of the 2-source extractor $\mathsf{TExt}$, and recall that $\mathsf{TExt}$ has error $\epsilon' = \frac{1}{8 \cdot n^{c_1} \cdot 2^m} = \frac{\epsilon}{8 \cdot 2^m}$. □

We now proceed with our plan of constructing the circuit $C$. Our first step will be to show that for every $y$, and $\alpha$, the set $T_{y,\alpha}$ can be recognized by a small $\Sigma_1$-circuit. For this purpose, we make the following definition:

**Definition 4.4.** *For every $y \in \{0,1\}^{n_{\mathsf{Cnd}}}$, and $v \in \{0,1\}^{n_1}$ we define two deterministic circuits $C^1_{y,v} : \{0,1\}^r \to \{0,1\}$ and $C^2_y : \{0,1\}^r \to \{0,1\}$ as follows:*

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

*We also define:*

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

This definition is made so that for every $y \in \{0,1\}^{n_{\mathsf{Cnd}}}$ and $v \in \{0,1\}^{n_1}$, if $p^2_y \neq 0$ then

$$\begin{aligned}\frac{p^1_{y,v}}{p^2_y} &= \frac{\Pr[\mathsf{TExt}(v, A(U_r)) = z \wedge \mathsf{Cnd}(A(U_r)) = y]}{\Pr[\mathsf{Cnd}(A(U_r) = y]} \\ &= \frac{\Pr[\mathsf{TExt}(v, X) = z \wedge \mathsf{Cnd}(X) = y]}{\Pr[\mathsf{Cnd}(X) = y]} \\ &= \Pr[\mathsf{TExt}(v, X) = z | Y = y].\end{aligned}$$

This means that for every $y \in \{0,1\}^{n_{\mathsf{Cnd}}}$ and $0 \leq \alpha \leq 1$, we can decide whether a given $v \in T_{y,\alpha}$ 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 $\Sigma_1$-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 4.5.** *For every $y \in \{0,1\}^{n_{\mathsf{Cnd}}}$, there is a $\Sigma_1$-circuit $C_y : \{0,1\}^{n_1} \to \{0,1\}$ of size $\mathsf{poly}(n^c, n^{c_1}, n^{d_{\mathsf{Cnd}}}, n^{d'})$ such that:*

- *For every $v \in \{0,1\}^{n_1}$ such that $C_y(v) = 1$, we have that $v \in T_{y,\frac{\epsilon}{8}}$.*
- *If $y$ is useful, then $C_y(\mathsf{Hrd}(y)) = 1$.*

*Proof.* When given $v \in \{0,1\}^{n_{\mathsf{Cnd}}}$, the circuit $C_y$ works as follows:

- $C_y$ checks whether there exists $w \in \{0,1\}^r$, such that $\mathsf{Cnd}(A(w)) = y$. If there does not exist such a $w$, it answers zero (as this means that $p^2_y = 0$)
- Let $\lambda = \epsilon/a$ for a universal constant $a > 1$ to be chosen later. $C_y$ applies Theorem 2.19 to compute a $\lambda$-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 $p_{y,v} = \frac{p^1_{y,v}}{p^2_y} = \Pr[\mathsf{TExt}(v, X) = z | Y = y]$.
- $C_y$ 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 $\lambda = \epsilon/a = \frac{1}{a \cdot n^{c_1}}$ 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 $p_{y,v} \leq (1 + \frac{\epsilon}{8}) \cdot 2^{-m}$.[16]

[16] Note that in order to distinguish between $(1 + \frac{\epsilon}{2}) \cdot 2^{-m}$ and $(1 + \frac{\epsilon}{8}) \cdot 2^{-m}$, it is sufficient to have an additive $\lambda$-approximation rather than a relative $\lambda$-approximation. However, as we approximate $p_{y,v}$ by approximating the enumerator $p^1_{y,v}$ and denominator $p^2_y$, we need to use a relative approximation to these two quantities in order to derive an additive approximation to $p_{y,v}$.

This gives that if $C_y(v) = 1$ then $p_{y,v} > (1 + \frac{\epsilon}{8}) \cdot 2^{-m}$ which gives that $v \in T_{y,\frac{\epsilon}{8}}$. By Claim 4.3 we have that for every useful $y$, $\mathsf{Hrd}(y) \in T_{y,\frac{\epsilon}{2}}$, which means that $p_{y,\mathsf{Hrd}(y)} > (1 + \frac{\epsilon}{2}) \cdot 2^{-m}$, and indeed, $C_y(\mathsf{Hrd}(y)) = 1$.

Finally, by definition $C_y$ is a circuit of size $\mathsf{poly}(n^c, n^{c_1}, n^{d_{\mathsf{Cnd}}}, n^{d'})$. □

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

**Claim 4.6.** *There is a $\Sigma_2$-circuit $C$ of size $\mathsf{poly}(n^c, n^{c_1}, n^{d_{\mathsf{Cnd}}}, n^{d'})$ such that:*

$$\Pr[C(Y) = \mathsf{Hrd}(Y)] \geq 2^{-k'} \cdot \frac{\epsilon}{4 \cdot 2^m}$$

*Proof.* We will first construct a randomized $\Sigma_2$-circuit $C'$, and then use a standard averaging argument to convert it to a non-randomized $\Sigma_2$-circuit. The randomized circuit $C'$ is defined as follows: On input $y \in \{0,1\}^{n_{\mathsf{Cnd}}}$:

- $C'$ constructs the $\Sigma_1$-circuit $C_y$. Note that the circuit $C_y$ is specified precisely in the proof of Claim 4.5, and so, the circuit $C'$ (that can be hardwired with $A$, $z$, and the circuit from Theorem 2.19) can construct the circuit $C_y$.
- $C'$ uses the $\Sigma_2$ circuit from Theorem 2.20 (choosing $\delta = \frac{1}{2}$) to output a uniform element in $\{v : C_y(v) = 1\}$.

By definition, the circuit $C'$ is a randomized $\Sigma_2$-circuit of size $\mathsf{poly}(n^c, n^{c_1}, n^{d_{\mathsf{Cnd}}}, n^{d'})$. We conclude that:

$$\begin{aligned}\Pr[C'(Y) = \mathsf{Hrd}(Y)] &\geq \Pr[C'(Y) = \mathsf{Hrd}(Y) | Y \text{ is useful}] \cdot \Pr[Y \textit{ is useful}] \\ &\geq \Pr[C'(Y) = \mathsf{Hrd}(Y) | Y \text{ is useful}] \cdot \frac{\epsilon}{4 \cdot 2^m} \\ &\geq \frac{1}{2} \cdot 2^{-k'_1} \cdot \frac{\epsilon}{4 \cdot 2^m} \\ &= 2^{-k'_1} \cdot \frac{\epsilon}{8 \cdot 2^m}\end{aligned}$$

where the first inequality follows by Claim 4.2, and the last inequality follows because by Claim 4.5, for every useful $y$, $\mathsf{Hrd}(y) \in \{v : C_y(v) = 1\}$ which by Claim 4.3, is of size at most $2^{k'_1}$, and each element in the set is obtained with probability $(1 - \delta)2^{-k'_1} = \frac{1}{2} \cdot 2^{-k'_1}$.

Finally, by a standard averaging argument, there exists a (non-randomized) $\Sigma_2$-circuit of size $\mathsf{poly}(n^c, n^{c_1}, n^{d_{\mathsf{Cnd}}}, n^{d'})$ with the same success probability. □

We have obtained a $\Sigma_2$-circuit $C$ of size $\mathsf{poly}(n^c, n^{c_1}, n^{d_{\mathsf{Cnd}}}, n^{d'}) = n^{\max(c, c_1, d_{\mathsf{Cnd}}, d') + c_0}$ for some universal constant $c_0$. By the requirements in Figure 1, we can get that $C$ is of size $n^{c_{\mathsf{Hrd}}}$. We also have that there is a distribution $Y$ over $\{0,1\}^{n_{\mathsf{Cnd}}}$ with $H_\infty(Y) \geq k_{\mathsf{Cnd}}$, that is samplable by the size $n^c + n^{d_{\mathsf{Cnd}}} \leq n^{c_{\mathsf{Hrd}}}$ circuit $\mathsf{Cnd} \circ A$, such that

$$\Pr[C(Y) = \mathsf{Hrd}(Y)] \geq 2^{-k'_1} \cdot \frac{\epsilon}{8 \cdot 2^m} = \frac{1}{n^{c_1} \cdot 8 \cdot 2^m} > \rho_{\mathsf{Hrd}},$$

where the inequality follows from the requirements in Figure 1, and this is a contradiction to the guarantee on $\mathsf{Hrd}$.

**Remark 4.7** (Theorem 4.1 yields multiplicative extractors)**.** *Theorem 4.1 states that the construction of Figure 1 yields an extractor for samplable distributions. We remark that the extractor that is achieved is multiplicative in the sense of Definition 2.4. More specifically, in the proof of Theorem 4.1, when we assume that* $\mathsf{Ext}$ *is not an extractor, we use it to derive (9) that states that there exists a $z \in \{0,1\}^m$ such that*

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

*We proceed from there to get a contradiction to this assumption, and this means that the same proof shows that* $\mathsf{Ext}$ *is a multiplicative extractor.*

*Unfortunately, as explained in Remark 4.9, current constructions of 2-source extractors do not achieve sufficiently low error to provide multiplicative extractors for samplable distributions with large output length. More specifically, as explained in Remark 4.9 using current constructions of 2-source extractors, we only obtain extractors for samplable distributions with output length* $m = O(\log n)$*, and* $\epsilon = n^{-O(1)}$*. The difference between standard extractors and multiplicative extractors becomes interesting when* $m \gg \log \frac{1}{\epsilon}$*, which is not the case here.*

## 4.3 An Extractor With Small Output Length

We now show how to choose specific ingredients to the construction of Section 4.1 and obtain an extractor for samplable distributions with low min-entropy. Using the best currently known 2-source extractors [CZ16, Li16] the resulting extractor for samplable distributions will only have short output length of $m = O(\log n)$. This result is stated below.

**Theorem 4.8** (extractor for with small output lenghth)**.** *There exists a constant* $\gamma > 0$ *such that if* $\mathsf{E}$ *is hard for exponential size* $\Sigma_4$*-circuits, then for every constants* $c, c_1 > 1$*, and every sufficiently large* $n$*, there is an* $(n^{1-\gamma}, \frac{1}{n^{c_1}})$*-extractor* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{c_1 \cdot \log n}$ *for distributions samplable by circuits of size* $n^c$*. Furthermore,* $\mathsf{Ext}$ *can be computed in time* $\mathsf{poly}(n^c, n^{c_1})$*.*

The reason that Theorem 4.8 only obtains $m = O(\log n)$ is because the best currently known 2-source explicit extractors for $k = o(n)$ [CZ16, Li16] cannot achieve $\epsilon = n^{-\omega(1)}$. In the construction described in Figure 1, when shooting for output length $m$, we require that the error $\epsilon'$ of the 2-source extractor $\mathsf{TExt}$ is $\epsilon' < \frac{1}{2^m}$, and this means that with current 2-source extractors, we can at best get $m = O(\log n)$ (and this is indeed what we obtain in Theorem 4.8).

*Proof.* (of Theorem 4.8) Theorem 4.8 will follow directly from Theorem 4.1 by choosing specific ingredients for the construction in Figure 1. More specifically:

- We choose $\gamma > 0$ to be the constant guaranteed in Theorem 3.1.
- We are assuming that $\mathsf{E}$ is hard for exponential size $\Sigma_4$-circuits.

Given constants $c, c_1$, and a sufficiently large $n$, we will instantiate the parameters and ingredients for the construction of Figure 1, as follows:

- We choose $k = n^{1-\gamma}$, and use the given parameters $c, c_1$ and $n$ as given for the parameters of Figure 1.
- We choose $m = c_1 \cdot \log n$ to be required output length.
- We choose $\epsilon' = \frac{1}{8 \cdot n^{c_1} \cdot 2^m}$, as is done in Figure 1, and note that for $c' = 3c_1$, for sufficiently large $n$, we have that that $\epsilon' \geq n^{-c'}$.
- We choose $n_{\mathsf{Cnd}} = n^{0.9}$ and $k_{\mathsf{Cnd}} = n^{0.7}$. We use the hardness assumption to apply Theorem 3.1, choosing $m = n_{\mathsf{Cnd}} = n^{0.9}$ to obtain a $(k, n^{0.8} - O(\log n))$-errorless condenser $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n_{\mathsf{Cnd}}}$ for distributions samplable by circuits of size $n^c$. Note that for sufficiently large $n$, we indeed have that $\mathsf{Cnd} : \{0,1\}^n \to \{0,1\}^{n_{\mathsf{Cnd}}}$ is a $(k, k_{\mathsf{Cnd}})$-errorless condenser for distributions samplable by circuits of size $n^c$, as required in Figure 1. Furthermore, we have that $\mathsf{Cnd}$ is computable in time $\mathsf{poly}(n^c)$, and more specificaly that there exists a constant $d_{\mathsf{Cnd}}$ such that $\mathsf{Cnd}$ can be computed by a circuit of size $n^{d_{\mathsf{Cnd}}}$.

- Let $a_0 > 1$ be the universal constant from Theorem 3.2. Theorem 2.12 specifies a running time for a 2-source extractor with error $\epsilon' \geq n^{-c'}$, and source length $n_1 = n^{a_0}$. This running time is polynomial is $n^{a_0}$ and $n^{c'}$, which is polynomial in $n^{c_1}$. Let $d'$ be a constant such that such a 2-source extractor can be computed by a circuit of size $n^{d'}$.
- At this point, the constants $c, c_1, d_{\mathsf{Cnd}}$ and $d'$ are determined. We choose a constant $c_{\mathsf{Hrd}}$ so that $c_{\mathsf{Hrd}}$ meets the requirements in Figure 1. More specifically, we choose $c_{\mathsf{Hrd}} = \max(c, c_1, d_{\mathsf{Cnd}}, d') + c_0$, where $c_0$ is the universal constant chosen by the proof of Theorem 4.1.
- We use the hardness assumption to apply Theorem 3.2. We are shooting for a function $\mathsf{Hrd}$ on input length $n_{\mathsf{Cnd}} = n^{0.9}$, entropy threshold $k_{\mathsf{Cnd}} = n^{0.7}$, and constant $c_{\mathsf{Hrd}}/0.9$. We indeed meet the requirement of Theorem 3.2 that $k_{\mathsf{Cnd}} \geq 4n_{\mathsf{Cnd}}^{\nu}$, and therefore, we can obtain a function $\mathsf{Hrd} : \{0,1\}^{n_{\mathsf{Cnd}}} \to \{0,1\}^{n^{a_0}}$ that is a $(n^{c_{\mathsf{Hrd}}}, k_{\mathsf{Cnd}}, \rho_{\mathsf{Hrd}})$-HOS for size $n^{c_{\mathsf{Hrd}}}$ $\Sigma_2$-circuits, for $\rho_{\mathsf{Hrd}} = 2^{-\Omega(k_{\mathsf{Cnd}})}$. Furthermore, $\mathsf{Hrd}$ is computable in time $\mathsf{poly}(n^{c_{\mathsf{Hrd}}}) = \mathsf{poly}(n^c, n^{c_1})$.
- Finally, we apply Theorem 2.12 to obtain a 2-source extractor. We have already chosen the error parameter $\epsilon' = \frac{1}{8 \cdot n^{c_1} \cdot 2^m} > n^{-c'}$, and the source length to be $n^{a_0}$. Indeed, Figure 1, calls for a $(k'_1, k'_2, \epsilon')$, $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^n \to \{0,1\}^m$. We note that we have already chosen $n_1 = n^{a_0} > n$, and so we can use a 2-source extractor $\mathsf{TExt} : \{0,1\}^{n_1} \times \{0,1\}^{n_1} \to \{0,1\}^m$ (by padding the second source). Using Theorem 2.12 we can get such a $(k', k', \epsilon')$-2-source extractor for $k' = n^{0.6} = n_1^{0.6/a_0} = n_1^{\Omega(1)}$.
- It remains to check that we meet the three requirements of Figure 1.
  - We have already made sure to meet the requirement on the constant $c_{\mathsf{Hrd}}$.
  - Our choices meet the requirement that $\rho_{\mathsf{Hrd}} < \frac{1}{8n^{c_1} \cdot 2^m \cdot 2^{k'_1}}$. This is because $m = c_1 \cdot \log n$, $k'_1 = k' = n^{0.6}$ and $\rho_{\mathsf{Hrd}} = 2^{-\Omega(k_{\mathsf{Cnd}})} = 2^{-\Omega(n^{0.7})}$, and so the requirements hold for sufficiently large $n$.
  - We meet the requirement that $k'_2 \leq k - n_{\mathsf{Cnd}} - m - c_1 \log n - 2$ for sufficiently large $n$, because $k = n^{1-\gamma}$, $n_{\mathsf{Cnd}} = n^{0.7}$ and $k'_2 = k' = n^{0.6}$.

We have chosen all the parameters and ingredients in Figure 1 in a way that satisfies the requirements. Using Theorem 4.1 we conclude that the constructed function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{c_1 \cdot \log n}$ is an $(n^{1-\gamma}, \frac{1}{n^{c_1}})$-extractor for distributions samplable by circuits of size $n^c$. Furthermore, by construction the running time of $\mathsf{Ext}$ is $\mathsf{poly}(n^c, n^{c_1})$. □

**Remark 4.9** (Potential improvements if the ingredients are improved)**.** *We now discuss potential improvements to Theorem 4.8 that can be achieved if the ingredients are improved.*

**Toward obtaining large output length directly.** *When using Theorem 4.1 to obtain an extractor for samplable distributions with output length m, we need to choose the error $\epsilon'$ of the 2-source extractor* $\mathsf{TExt}$ *to be $\epsilon' \leq 2^{-m}$. We need 2-source extractors with min-entropy threshold $\mathrm{n}^{\Omega(1)}$ and while the recent breakthrough 2-source extractors of Chattopadhyay and Zuckerman [CZ16, Li16] achieve such a min-entropy threshold, they run in time* $\mathsf{poly}(\frac{1}{\epsilon'})$ *and so, cannot achieve error $\epsilon' = n^{-\omega(1)}$. Consequently, when using Theorem 4.1 directly, we can only hope for extractors for samplable distributions with $m = O(\log n)$.*

*Fortunately, we can achieve extractors with large output length by using a transformation of Shaltiel [Sha08] that is discussed in Section 4.4. This transformation comes with a cost of a stronger hardness assumption (although as explained in Remark 4.11, this cost can be avoided). Future improvements in the error of 2-source extractors will translate to a direct construction using Theorem 4.1. Such improvements will also have additional consequences as we explain next.*

**Toward multiplicative extractors.** *As observed in Remark 4.7, Theorem 4.1 yields multiplicative extractors. However, with the current 2-source extractors this does not yield interesting results. This is because multiplicative extractors are superior to (standard) additive extractors, only when $m \gg \log\frac{1}{\epsilon}$ and as observed above, current 2-source extractors are not sufficient to achieve extractors for samplable distributions with such parameters. Future improvements in 2-source extractors will immediately yield multiplicative extractors. The reader is referred to [Sha24] for a discussion on the benefits of multiplicative extractors.*

*We also remark that the construction of Theorem 4.1 cannot achieve error of $\epsilon = n^{-\Omega(1)}$. Moreover, Applebaum et al. [AASY15] showed that black-box techniques cannot be used to obtain extractors for samplable distributions with $\epsilon = n^{-\omega(1)}$ that run in time $\mathsf{poly}(n)$ from the type of hardness assumptions that we use.*

## 4.4 Increasing the Output Length of Extractors for Samplable Distributions

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 (specified in Figure 2) 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 $\Sigma_1$-circuits. The exact details are given in Figure 2 and Theorem 4.10.

Figure 2: Increasing the output length of extractors for samplable distributions

**Parameters:**

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

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

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

- A $(t + \ell, \epsilon/2^{t+10})$-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^t$ for distributions samplable by size $s'$ $\Sigma_1$-circuits.
- An explicit strong $(k, \epsilon/2)$-extractor $\mathsf{SExt} : \{0,1\}^{n_1} \times \{0,1\}^t \to \{0,1\}^{m_1}$.

**Requirements:**

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

**Description of $\mathsf{Ext}'$:** $\mathsf{Ext}'(x) = \mathsf{SExt}(x, \mathsf{Ext}(x)), \mathsf{Ext}(x)$.

**Output length of $\mathsf{Ext}'$:** Note that this gives that the output length of $\mathsf{Ext}'$ is $m = m_1 + t$.

**Theorem 4.10.** *Given parameters and ingredients as in Figure 2, the function $\mathsf{Ext}'$ is a $(k, \epsilon)$-extractor for distributions samplable by size $s$ circuits.*

Theorem 1.4 and Theorem 1.5 will follow by using this transformation, and the difference will be in the choice of seeded extractors.

### 4.4.1 Proof of Theorem 1.4

We are assuming that $\mathsf{E}$ is hard for exponential size $\Sigma_5$-circuits. We have taken $i = 5$, rather than $i = 4$, so that by "ladder climbing" (see Remark 2.15) when applying Theorem 4.8, under this stronger assumption for $\Sigma_5$-circuits, we obtain that $\mathsf{Ext}$ is an extractor for distributions samplable by $\Sigma_1$-circuits.

More specifically, let $\gamma > 0$ be the constant guaranteed by Theorem 4.8. We will use the constant $\gamma' = \gamma/2$ as the constant that we should guarantee in Theorem 1.4. Given a constants $c > 1$, a constant $\alpha > 0$, a sufficiently large $n$, and $k \geq n^{1-\gamma'}$, we set $\epsilon = n^{-c}$ and apply Theorem 2.8 to obtain a strong $(k, \frac{\epsilon}{2})$-extractor $\mathsf{SExt} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^{m_1}$, for $m_1 = (1-\alpha)k$. By Theorem 2.8 we have that $t = c_0 \cdot c \cdot \log n$ for some constant $c_0$ (hidden in the O-notation in Theorem 2.8). We set $c_1$ to be sufficiently large so that $c \cdot c_0 \leq c_1$, and $\frac{\epsilon}{2^{t+10}} = 2^{-10} \cdot n^{-c_0 \cdot c - c} \geq n^{-c_1}$. This will hold for $c_1 = O(c)$. These choices are made so that $t \leq c_1 \cdot \log n$ and $n^{-c_1} \leq \frac{\epsilon}{2^{t+10}}$.

We will use the hardness assumption to apply Theorem 4.8 to obtain an $(n^{1-\gamma}, n^{-c_1})$-extractor for distributions samplable by size $q(n^c + n)$, where $q$ is the polynomial from Figure 2. The output will be $c_1 \cdot \log n$, and we can cut it to length $t \leq c_1 \cdot \log n$. We obtain that $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^t$ can be computed in time $\mathsf{poly}(n^c, n^{c_1}) = \mathsf{poly}(n^c)$. Let $\ell = n^{1-\gamma} - t$ so that $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^t$ is a $(t + \ell, \frac{\epsilon}{2^{t+10}})$-extractor for distributions samplable by size $q(n^c + n)$ $\Sigma_1$-circuit (as required in Figure 2).

In order to meet the requirement made in Figure 2 that

$$m_1 = (1-\alpha) \cdot k \leq k - (2t + \ell + \log(1/\epsilon) + 5),$$

we note that $k \geq n^{1-\gamma'} = n^{1-\gamma/2}$ is significantly larger than $\ell \leq n^{1-\gamma}$. More specifically, for sufficiently large $n$,

$$k - (2t + \ell + \log(1/\epsilon) + 5) \geq k - O(\ell) \geq n^{1-\gamma/2} - O(n^{\gamma}) \geq (1-\alpha) \cdot n^{1-\gamma/2} \geq (1-\alpha)k,$$

and so the requirement is met. Using Theorem 4.10 we obtain a function $\mathsf{Ext}' : \{0,1\}^n \to \{0,1\}^{(1-\alpha)\cdot k}$ that is a $(k, \epsilon)$-extractor for distributions samplable by circuits of size $n^c$ as required. By construction, $\mathsf{Ext}'$ runs in time $\mathsf{poly}(n^c)$.

**Remark 4.11.** *As stated in Theorem 4.10 and Figure 2, the approach of Shaltiel [Sha08] requires the initial extractor* $\mathsf{Ext}$ *to be an extractor for distributions samplable by $\Sigma_1$-circuits. However, a closer inspection of the Proof of Theorem 4.10 reveals that the same result follows under a weaker condition. It is sufficient that* $\mathsf{Ext}$ *is an extractor for distributions that are samplable* with postselection *by (deterministic) circuits of size $s'$.*

*The notion of distributions samplable with postselection was introduced by Ball et al. [BGDM23]. A distribution $X$ is samplable with postselection by circuits of size $s$, if there exists two size $s$ circuits, $A : \{0,1\}^r \to \{0,1\}^n$ and $P : \{0,1\}^r \to \{0,1\}$, such that $X$ is obtained as the distribution $(A(W)|P(W) = 1)$ for $W \leftarrow U_r$.*

*Ball et al. [BGDM23] and later work by Shaltiel [Sha24] achieved extractors for samplable distributions with postselection (rather than just samplable distributions). We remark that Theorem 4.8 can be extended to yield extractors for distributions samplable with postselection by size $n^c$ circuits (under the same hardness assumption).*

*This can be done by a more careful argument (that replaces samplable distributions, with distributions samplable with postselection in both the extractor construction of Section 4, and the components in Section 3.*

*Together, these two improvements can avoid the loss in the transformation of Theorem 4.10, resulting in an improved assumption that replaces $\Sigma_5$-circuits with $\Sigma_4$-circuits in Theorems 1.4 and Theorem 1.5. We defer the details to a later version.*

### 4.4.2 Proof of Theorem 1.5

The proof of Theorem 1.5 is identical to the proof of Theorem 1.4 except that we use the seeded extractor of Theorem 2.9 instead of Theorem 2.8.

## 5 Conclusion and Open Problems

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

**Improving the min-entropy threshold.** Our extractors achieve $k = n^{1-\gamma}$ for some constant $\gamma > 0$. It is natural to try an obtain extractors for lower values of $k$. We remark that the specific approach of this paper cannot give $k < \sqrt{n}$.

Let us now technically explain this statement. As explained in Section 1.3.2, for min-entropy $k$, our construction requires an errorless condenser for min-entropy threshold $k$ with output length $m_{\mathsf{Cnd}} < k$. Our technique for constructing an errorless condenser for min-entropy threshold $k$, relies in turn on a strong seeded $(k, \epsilon)$-disperser $E : \{0,1\}^n \times [D = 2^d] \to \{0,1\}^m$ (as well as other components). Even ignoring the other components, when using our approach, the output length $m_{\mathsf{Cnd}}$ of the obtained errorless condenser satisfies $m_{\mathsf{Cnd}} > D$ (it is in fact, somewhat larger). It is easy to show that in every nontrivial strong seeded disperser (namely, one in which $(1-\epsilon) \cdot 2^{m+d} \geq 2^{d+1}$ which is equivalent to $\epsilon \leq 1 - 2^{-(m-1)}$) it holds that $D > \frac{n-k}{m}$. We use strong seeded dispersers with $m \leq k$, and therefore must have that

$$m_{\mathsf{Cnd}} > D > \frac{n-k}{m} > \frac{n-k}{k},$$

which implies that if we want $m_{\mathsf{Cnd}} < k$, then $k$ cannot be significantly smaller than $\sqrt{n}$. This means that even when using optimal strong seeded dispersers, as long as $m \leq k$, this technique does not yield an errorless condenser with $m_{\mathsf{Cnd}} < k$ for $k = o(\sqrt{n})$.[17]

Nevertheless, it may be possible to construct the required errorless condenser by other means, or alternatively maybe one can hope to use errorless condensers with $m_{\mathsf{Cnd}} > k$, and somehow avoid the requirement that $m_{\mathsf{Cnd}} < k$.

**Weakening the assumption.** The assumption used in Theorem 1.4 and Theorem 1.5 is that $\mathsf{E}$ is hard for exponential size $\Sigma_5$-circuits. This assumption can be weakened to replace $\Sigma_5$-circuits by $\Sigma_4$-circuits, as explained in Remark 4.11. While this assumption is weaker than that used by Trevisan and Vadhan [TV00]. It was recently shown by Ball et. al. [BGDM23] how to achieve the extractor of [TV00] under the weaker assumption that $\mathsf{E}$ is hard for exponential size nondeterministic circuits. (See [Sha24] for a discussion on the

[17] This begs the question of whether this limitation applies to strong seeded dipsersers where $m$ is much larger than $k$. We first remark that it is obvious that strong seeded dispersers with $\epsilon < \frac{1}{2}$ (which is the more standard setting of parameters) must have $m \leq k$. However, in this paper we are allowing $\epsilon$ to approach one. This allows $m$ to be much larger than $k$, and in this setting dispersers are often referred to as "unbalanced bipartite expanders". Indeed, in this setting it is more natural to set $k'$ so that $k' = (1-\epsilon) \cdot 2^m$, so that for every set $S \subseteq \{0,1\}^n$ on the "left hand side", the set $\Gamma(S) = \cup_{i \in [D]} \{(E(S,i), i)\} \subseteq \{0,1\}^{m+d}$ of neighbors of $S$, expands to size $D \cdot 2^{k'}$. With this parametrization, using the same argument as above, we can observe that if we aim for low $k = n^{\alpha}$ for some constant $\alpha > 0$, the disperser must have $m = \Omega(n^{1-\alpha})$ and $k' \leq k$.

The way in which we use such a disperser to construct an errorless condenser, can be viewed as reducing the task of constructing an errorless condenser on $\{0,1\}^n$ for min-entropy threshold $k$, to the task of constructing an errorless condenser on $\{0,1\}^m$ for min-entropy threshold $k'$. This means that for the parameters that we now consider (where $m = \Omega(n^{1-\alpha})$ and $k' \leq k$) the task that we are reducing to is at best, only slightly easier than what we want to achieve. It is not clear to us whether there is hope of iteratively applying this reduction in order to make progress, and in any case, current explicit constructions of unbalanced expanders for this regime are very far from optimal (ruining any potential gain).

necessity of assuming hardness for nondeterministic circuits). The improvement of Ball et al. [BGDM23] gives hope that maybe similar ideas can achieve our low min-entropy extractor under a weaker assumption.

**Multiplicative extractors with large output length.** As explained in Remark 4.7, and Remark 4.9, our technique can give multiplicative extractors in Theorem 4.8. This is not impressive as Theorem 4.8 only achieves an output length of $m = \Theta(\log n)$, for $\epsilon = \frac{1}{\mathsf{poly}(n)}$, and multiplicative extractors are not interesting unless $m = \omega(\log(1/\epsilon))$. The short output length is a result of the parameters of the current best explicit constructions of 2-source extractor for min-entropy $k = o(n)$, and our approach will give multiplicative extractors for samplable distributions with output length $m$, and multiplicative error $\epsilon = n^{-c}$, if there are explicit constructions of 2-source extractors for min-entropy threshold $k' = n^{\Omega(1)}$, output length $m$ and error $\epsilon' = o(\epsilon/2^m)$.

**Achieving extractors with $m = (1 - o(1)) \cdot k$ and $\epsilon = n^{-c}$.** It is natural to ask whether one can obtain an extractor that combines the advantages of Theorems 1.4 and Theorem 1.5, and achieves both $m = (1-o(1)) \cdot k$ and $\epsilon = n^{-c}$. Our approach will immediately give such an extractor if there are explicit constructions of strong seeded extractors which for $k = n^{1-\gamma}$ achieve seed length $O(\log n)$ for $m = (1 - o(1)) \cdot k$, and $\epsilon = n^{-c} \cdot 2^{-d}$. This immediately follows from the composition of Section 4.4. Current construcctions of seeded extractors with $m = (1 - o(1)) \cdot k$ [DKSS09, TU12] do not achieve small error (and indeed Theorem 1.5 inherits its error $\epsilon$ from the strong seeded extractors of Ta-Shma and Umans [TU12]).

# 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.

[AST$^+$98] A. Ambainis, L. J. Schulman, A. Ta-Shma, U. V. Vazirani, and A. Wigderson. The quantum communication complexity of sampling. In *39th Annual Symposium on Foundations of Computer Science, FOCS*, pages 342–351. IEEE Computer Society, 1998.

[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.

[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.

[CG88] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, April 1988.

[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.

[DKSS09] Z. Dvir, S. Kopparty, S. Saraf, and M. Sudan. Extensions to the method of multiplicities, with applications to kakeya sets and mergers. In *FOCS*, pages 181–190, 2009.

[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.

[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.

[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.

[LRVW03] C.J. Lu, O. Reingold, S. P. Vadhan, and A. Wigderson. Extractors: optimal up to constant factors. In *STOC*, pages 602–611. ACM, 2003.

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

[RTS00] J. Radhakrishnan and A. Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13(1):2–24, 2000.

[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.

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

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

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

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

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

[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.

[TZ04] A. Ta-Shma and D. Zuckerman. Extractor codes. *IEEE Trans. Inf. Theory*, 50(12):3015–3025, 2004.

[Vio12] E. Viola. The complexity of distributions. *SIAM J. Comput.*, 41(1):191–218, 2012.

[Vio14] E. Viola. Extractors for circuit sources. *SIAM J. Comput.*, 43(2):655–672, 2014.

[Vio20] E. Viola. Sampling lower bounds: Boolean average-case and permutations. *SIAM J. Comput.*, 49(1):119–137, 2020.

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