Electronic Colloquium on Computational Complexity, Report No. 89 (2026)

Near Optimal Extractors for Samplable Sources under Nondeterministic Hardness

Marshall Ball*
New York University
marshall.ball@cs.nyu.edu

Eshan Chattopadhyay†
Cornell University
eshan@cs.cornell.edu

Mohit Gurumukhani†
Cornell University
mgurumuk@cs.cornell.edu

Yunya Zhao†
Cornell University
yunya@cs.cornell.edu

Abstract

We study the problem of constructing randomness extractors for samplable sources, introduced by Trevisan and Vadhan (FOCS 2000), a natural computational model of imperfect randomness, where the source $\mathbf{X}$ (on n bits) is generated by a polynomial-size circuit. They showed how to extract from sources with min-entropy $(1-\alpha)n$ (for small $\alpha > 0$) assuming exponential hardness against Σ_6-circuits.

A line of recent works has made significant progress, either achieving extraction from such high min-entropy under weaker assumptions (e.g., nondeterministic circuits), or handling polynomially small min-entropy under stronger assumptions (e.g., hardness against Σ_i-circuits). These works output nearly all the randomness with polynomially small error. In a recent work, Oh and Shaltiel (STOC 2026) constructed extractors that work for logarithmic min-entropy under incomparable assumptions, outputting 1 random bit with constant error.

Our main result gives an extractor for samplable sources with min-entropy $\mathrm{polylog}(n)$ with polynomially small error and outputting almost all of the randomness, assuming hardness against nondeterministic circuits. The extractor also works for sources samplable with postselection by nondeterministic circuits. We can further reduce the entropy requirement to $O(\log n)$ at the expense of making the error constant and outputting only 1 bit, matching the extractor of Oh and Shaltiel (STOC 2026). By prior work of Shaltiel (CCC 2025), our extractors imply hardness against nondeterministic circuits, and thus our assumption is essentially minimal.

*Supported by NSF Award CCF-2443735.
†Supported by NSF Award CCF-2514586.

ISSN 1433-8092

Contents

1 Introduction

Randomness is a vital resource in computation, with important applications in the design of secure cryptographic protocols, machine learning, algorithm design, sampling, and more. The performance of these tasks crucially relies on the random bits being of high quality, i.e., forming a stream of uniform, independent bits. However, real-world sources of randomness are often biased or have limited entropy, necessitating the use of *randomness extractors*, which transform such sources into (almost) uniform random bits.

Towards modeling such imperfect sources arising in nature, Trevisan and Vadhan [TV00] introduced the class of *samplable distributions*—distributions generated by efficient computation. This is motivated by the perspective that while true randomness exists in nature (e.g., from quantum phenomena), the distributions accessible in practice (e.g., via TRNGs) are low-complexity transformations of underlying uniform random bits. [TV00] specifically modeled efficient computation by polynomial size circuits. Several other works considered other models such as algorithms with limited memory [KRVZ06], algebraic structure on the sampling algorithm [GR08; DGW09], or other concrete models of computation [Vio14].

Another prominent line of work assumes independence between multiple weak sources [CG88]. In this setting, one has access to two (or more) independent sources of randomness. This approach has seen a lot of progress in recent years [CZ19; Li23]. In contrast, the samplable source model does not assume independence, but instead imposes computational structure on a single source.[1]

We now formalize these notions. It is standard to measure quality of an imperfect source $\mathbf{X}$ using min-entropy, defined as $H_\infty(\mathbf{X}) = \min_{x \in \mathrm{supp}(\mathbf{X})} \log \frac{1}{\Pr[\mathbf{X}=x]}$. An (n,k)-source $\mathbf{X}$ is a distribution on $\{0,1\}^n$ with min-entropy at least k.[2] Equivalently, if $\mathbf{X}$ is an (n,k)-source, then for any $x \in \{0,1\}^n$, $\Pr[\mathbf{X}=x] \le 2^{-k}$. Hence $0 \le k \le n$. We call the quantity k/n the *entropy rate* of the source.

Definition 1.1. *A distribution $\mathbf{X}$ over $\{0,1\}^n$ is said to be* samplable by circuits of size s *if there exists a Boolean circuit $C : \{0,1\}^r \to \{0,1\}^n$ of size s such that $\mathbf{X} = C(\mathbf{U}_r)$, where $\mathbf{U}_r$ denotes the uniform distribution over $\{0,1\}^r$.*

We in fact study a richer class of sources called samplable sources with postselection. Here, along with C as above, we associate another circuit $C_{\mathsf{Post}} : \{0,1\}^r \to \{0,1\}$, where C and C_{Post} share the same input bits and $\mathbf{X}$ is sampled by C conditioned on C_{Post} outputting 1 (see Definition 4.7). This captures samplable sources subject to leakage by an efficient adversary.

We next define randomness extractors. To evaluate the quality of the output distribution of an extractor, we use statistical distance $|\mathbf{X} - \mathbf{Y}| = \frac{1}{2} \cdot \sum_{w \in \Omega} |\Pr[\mathbf{X}=w] - \Pr[\mathbf{Y}=w]|$.

Definition 1.2. *A function $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ is a (k,ε)-extractor for a family of distributions $\mathcal{X}$ if for all distributions $\mathbf{X} \in \mathcal{X}$ with $H_\infty(\mathbf{X}) \ge k$, we have $|\mathsf{Ext}(\mathbf{X}) - \mathbf{U}_m| \le \varepsilon$. The parameters m, ε are the output length and error of the extractor, respectively.*

The main focus of this paper is constructing efficient extractors for sources sampled by polynomial size circuits.

[1] It is folklore that, in general, one cannot deterministically extract even a single random bit from a single weak source of randomness without additional structural assumptions.

[2] When the length is implicit, we suppress it and simply refer to $\mathbf{X}$ as a k-source.

1.1 Extracting from samplable sources

Trevisan and Vadhan [TV00] demonstrated that unconditionally extracting from such sources will imply unknown separations in complexity classes. Specifically, the existence of an extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ computable in time $t(n)$ with constant additive error for sources of min-entropy $n-1$ sampled by circuits of size $s(n)$ would imply a complexity separation: there must exist a language in $\mathsf{DTIME}(t(n))$ with circuit complexity $\Omega(s(n))$. Therefore, in order to extract from samplable sources, it is necessary to make hardness assumptions.

Before we discuss the results of previous work, we give a brief overview of hardness assumptions made against different circuit classes.

Hardness assumptions against various circuit classes We say "E is hard for exponential size circuits in a certain class" if there is a language $L \in \mathsf{E} = \mathsf{DTIME}(2^{O(n)})$ and a constant $\delta > 0$ such that for sufficiently large n, L cannot be computed by circuits (of that class) of size $2^{\delta n}$ on inputs of length n. Below are some circuit classes of interest:

Σ_i-circuits are allowed oracle gates to the canonical Σ_i^P-complete language. For example, NP-circuits, or Σ_1-circuits, are allowed oracle SAT-gates. This is formally presented in Definition 4.4.

A nondeterministic circuit [3] C can be thought of as a deterministic circuit C' that takes in an input x along with a witness w. On input x, $C(x) = 1$ if and only if there exists a witness w such that $C'(x, w) = 1$. This is formally defined in Definition 4.1.

In this work, we assume hardness against nondeterministic circuits, a standard assumption in derandomization. Among other applications, such assumptions have been used to derandomize AM (see, e.g., [KM02]) and to obtain fast derandomization of BPP (e.g., [DMOZ22]); see also Impagliazzo-Wigderson [IW97], who showed $\mathsf{P} = \mathsf{BPP}$ under exponential size hardness assumption against deterministic circuits.

Prior works The main result in [TV00] states that if E is hard for exponential size Σ_6-circuits,[4] then there exists an efficiently computable extractor for sources samplable by poly-size circuits with min-entropy $(1-\alpha)n$ for a small constant $\alpha > 0$. Since the work of Trevisan and Vadhan [TV00], improvements have been made from two different directions: (a) relaxing the hardness assumption, and (b) extending the class of sources that can be extracted from, including samplable sources *with postselection* and sources with lower min-entropy.

Ball, Goldin, Dachman-Soled, and Mutreja [BGDM23] made substantial progress, many years after [TV00], in both directions: they relaxed the hardness assumption from Σ_6-circuits to nondeterministic circuits, and extracted from a richer class of distributions, namely samplable sources with postselection (see Definition 4.7). Their min-entropy requirement was similar to [TV00], namely at least $(1-\alpha)n$ for a small constant $\alpha > 0$. Shaltiel [Sha25b] gave a more modular proof of the same result in terms of hardness assumption and min-entropy requirement, while handling an even richer class of sources that are samplable by deterministic circuits with postselection from nondeterministic circuits.

[3] not to be confused with NP-circuits

[4] This assumption is made by the conference version [TV00]. There is a later unpublished version which only needs to assume Σ_5 hardness. In later parts of this paper, we refer to [TV00] construction as the one assuming Σ_5 hardness.

Ball, Shaltiel, and Silbak [BSS25] made the first progress towards handling lower min-entropy, assuming hardness against Σ_5-circuits. They constructed extractors for samplable sources of min-entropy $n^{1-\alpha}$, for a small constant $\alpha > 0$. In another recent work, Shaltiel in [Sha25a] improved the min-entropy requirement to n^{α}, for every $\alpha > 0$, but needed to make a hardness assumption against Σ_i circuits, where $i = \lceil \frac{1}{\alpha} \rceil + 3$.

The error of the extractors in all these works is polynomially small. [AASY16] provides evidence that achieving negligible error against samplable sources (under hardness assumptions against Σ_i-circuits) is beyond the reach of current techniques. Also, the output length of the extractor in all these works is $0.99k$.

Using a different and incomparable hardness assumption that E is hard for large exponential time with exponential advice, Oh and Shaltiel [OS26] achieved (constant error) extractors, outputting 1 bit for samplable sources (without postselection) with min-entropy $O(\log n)$.

1.2 Our results

In this work, we only make hardness assumptions against nondeterministic circuits and construct the following two extractors.

In our first result, we construct extractors with exponentially better entropy requirement than previous works: polylogarithmic in the input length, while also being able to output almost all randomness and obtain polynomially small error.

Theorem 1. *Assume E is hard for exponential size nondeterministic circuits. Then, there exists a universal constant c_0 such that for every constant $c > 1$ and all $n, k \geq (\log(n))^{c_0}$, there exists an explicit (k, ε)-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{0.99k}$ for distributions sampled by size n^c circuits with postselection by size n^c nondeterministic circuits, where $\varepsilon = n^{-c}$.*

For our second result, we construct extractors with parameters similar to those of [OS26]. As mentioned above, the result of [OS26] is under a different, incomparable assumption. We note that the extractors of [OS26] work for samplable sources, while our extractors work for the richer class of samplable sources with postselection.

Theorem 2. *Assume E is hard for exponential size nondeterministic circuits. Then, for every constant $c > 1$, there exists a constant c_0 such that for all $n, k \geq c_0 \log(n)$, there exists an explicit (k, ε)-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ for distributions sampled by size n^c circuits with postselection by size n^c nondeterministic circuits, where $\varepsilon = \frac{1}{c}$.*

Shaltiel [Sha25b] showed that an extractor for such sources (that are generated by postselection by nondeterministic circuits) implies circuit lower bounds against nondeterministic circuits; in particular, computing the extractor is hard on average for nondeterministic circuits. Therefore, hardness assumptions against nondeterministic circuits are essentially "minimal" for the class of sources we consider.

Table 1 summarizes the hardness assumptions and min-entropy requirements of prior works along with our results, omitting [OS26] since their assumptions are incomparable.

Organization The rest of the paper is organized as follows. We begin with an overview of the proof in Section 2, followed by concluding remarks in Section 3. We present basic definitions and relevant prior work in Section 4. The main technical ingredient appears in Section 5, where we show

Work	Hardness	Entropy requirement	Error	Output length
[TV00]	Σ_6-circuits	$(1-\varepsilon)n$	$1/\operatorname{poly}(n)$	$0.99k$
[BGDM23; Sha25b]	nondeterministic circuits	$(1-\varepsilon)n$	$1/\operatorname{poly}(n)$	$0.99k$
[BSS25]	Σ_5-circuits	$n^{(1-\varepsilon)}$	$1/\operatorname{poly}(n)$	$0.99k$
[Sha25a]	Σ_{i+3}-circuits	$n^{1/i}$	$1/\operatorname{poly}(n)$	$0.99k$
Theorem 1	nondeterministic circuits	$\log(n)^C$	$1/\operatorname{poly}(n)$	$0.99k$
Theorem 2	nondeterministic circuits	$C\log(n)$	ε	1

Table 1: C and ε represent some large and small unspecified constants respectively.

how to extract uniform bits from an elementary somewhere shrunk and random source together with the original source. In Section 6, we combine these ingredients to construct our extractors and prove Theorem 1 and Theorem 2.

2 Proof overview

In this section, we sketch proofs of all our main theorems. In Section 2.1, we sketch previous techniques used to construct extractors for samplable sources. Building on these, we provide details of our techniques to obtain Theorem 1 and Theorem 2 in Section 2.2.

2.1 Previous techniques

We begin by reviewing previously used techniques to construct extractors for samplable sources.

2.1.1 Extractor for $0.99n$ entropy

We first sketch the high level technique of the work of [TV00] who obtained extractors for $0.99n$[5] entropy samplable sources under a hardness assumption for Σ_5-circuits. They used the following ingredients:

1. Two source extractor: A two source extractor $\mathsf{2Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ with min-entropy requirement (k_1, k_2) and error ε has the property that for any two independent sources $\mathbf{X}_1 \sim \{0,1\}^{n_1}$ and $\mathbf{X}_2 \sim \{0,1\}^{n_2}$ such that $\mathbf{H}_\infty(\mathbf{X}_1) \geq k_1, \mathbf{H}_\infty(\mathbf{X}_2) \geq k_2$, the distribution $\mathsf{2Ext}(\mathbf{X}_1, \mathbf{X}_2)$ is ε-close to $\mathbf{U}_m$.

   From this, one can see that $\mathsf{2Ext}$ satisfies the following combinatorial list decoding property:

   Proposition 2.1 (Combinatorial list decoding property of two source extractors)**.** *Let $\mathbf{X}_2 \sim \{0,1\}^{n_2}$ be any source such that $\mathbf{H}_\infty(\mathbf{X}_2) \geq k_2$. Let $z \in \{0,1\}^m$ be arbitrary. Then, $|\{x_1 \in \{0,1\}^{n_1} : \Pr[\mathsf{2Ext}(x_1, \mathbf{X}_2) = z] > 2^{-m} + \varepsilon\}| < 2^{k_1}$.*

   Proof. We prove this by contradiction. Suppose there exists such z and $\mathbf{X}_2$. Let $B = \{x_1 \in \{0,1\}^{n_1} : \Pr[\mathsf{2Ext}(x_1, \mathbf{X}_2) = z] > 2^{-m} + \varepsilon\}$. By assumption, $|B| \geq 2^{k_1}$. Then, let $\mathbf{X}_1$ be the source that is independent of $\mathbf{X}_2$ and is uniform over B. We can easily compute that the statistical distance of $\mathsf{2Ext}(\mathbf{X}_1, \mathbf{X}_2)$ from $\mathbf{U}_m$ will be more than ε, a contradiction. □

[5] Here and everywhere in the overview, we use 0.99 to mean $1 - \alpha$ for a small constant $\alpha > 0$.

For our arguments, we will use excellent explicit two source extractors from [CZ19; Li16] such that $\mathsf{2Ext} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{0.99k}$ where $k = \text{polylog}(n)$ is the min-entropy requirement of both sources and the output error is $\frac{1}{\text{poly}(n)}$. This will be used in obtaining our poly$(\log(n))$ entropy extractors from Theorem 1. For $O(\log(n))$ entropy extractors from Theorem 2, we will use $O(\log(n))$ entropy extractors from [Li23] that require $O(\log(n))$ entropy, output 1 bit and have constant error.

2. Hard to approximate function. In particular, starting with the assumption that E is worst case hard for Σ_i-circuits, [TV00] construct a function $g : \{0,1\}^n \to \{0,1\}^n$ such that g is average case hard for Σ_{i-3}-circuits. In particular, for any polynomial size Σ_{i-3}-circuit C, g has the guarantee that $\Pr_{x\sim\{0,1\}^n}[C(x) = g(x)] \leq 2^{-\Omega(n)}$. This is surprising since this g itself is computable in larger polynomial time. To construct such a function g, they followed proof strategy of [STV01] and use low-degree extension to yield fast local list decoding algorithms of Reed-Muller code using nondeterminism.

3. Approximate counting and sampling using nondeterminism [Sto83; Sip83; JVV86; BGP00]. This result states that there exist Σ_{i+1}-circuits C_{approx} as well as C_{sample} that each take as input a Σ_i-circuit $C : \{0,1\}^n \to \{0,1\}^m$ and $y \in \{0,1\}^m$. Then, C_{approx} outputs a good relative approximation of the number of x such that $C(x) = y$ and C_{sample}, using randomness, samples a near uniform element from the preimage of y under C.

Construction Using all these ingredients, we can define extractor Ext as follows.[6] On input $x \in \{0,1\}^n$, let $x_{\text{pre}} \in \{0,1\}^{n/2}$ be the length $n/2$ prefix of x. Let $g : \{0,1\}^{n/2} \to \{0,1\}^{n/2}$ be a function that is $2^{-\Omega(n)}$ hard to approximate for Σ_2-circuits. Let $\mathsf{2Ext} : \{0,1\}^{n/2} \times \{0,1\}^n \to \{0,1\}^m$ be the [CZ19; Li16] two-source extractor with output length $O(\log(n))$.[7]

Using this, output $\mathsf{2Ext}(g(x_{\text{pre}}), x)$. Ext will have error $\varepsilon = O\left(\varepsilon_{\mathsf{2Ext}} \cdot 2^m\right)$ where $\varepsilon_{\mathsf{2Ext}}$ is the error of $\mathsf{2Ext}$.

Correctness To argue correctness, we proceed by contradiction and assume Ext fails on a samplable source $\mathbf{X}$. The strategy will be to obtain a contradiction to the hardness of g. Since Ext fails, by an averaging argument, there exists $z_{\text{heavy}} \in \{0,1\}^m$ such that $\Pr[\mathsf{2Ext}(g(\mathbf{X}_{\text{pre}}), \mathbf{X}) = z_{\text{heavy}}] \geq (1+\varepsilon) \cdot 2^{-m}$. As $\varepsilon_{\mathsf{2Ext}} \leq 10^{-6} \cdot \varepsilon \cdot 2^{-m}$, this implies $\Pr[\mathsf{2Ext}(g(\mathbf{X}_{\text{pre}}), \mathbf{X}) = z_{\text{heavy}}] \geq 2^{-m} + 10^6 \cdot \varepsilon_{\mathsf{2Ext}}$. Since $\mathbf{X}$ has entropy $0.99n$ and $\mathbf{X}_{\text{pre}}$ has length at most $0.5n$, using chain rule for min-entropy (Lemma 4.10), for most $x_{\text{pre}} \sim \mathbf{X}_{\text{pre}}$, $\mathbf{X}|(\mathbf{X}_{\text{pre}} = x_{\text{pre}})$ has entropy at least $0.49n$. We make the simplifying assumption that this holds for all x_{pre}.

Next, we say x_{pre} is "biasing" if $\Pr[\mathsf{2Ext}(g(x_{\text{pre}}), \mathbf{X}|(\mathbf{X}_{\text{pre}} = x_{\text{pre}})) = z_{\text{heavy}}] \geq 2^{-m} + \varepsilon_{\mathsf{2Ext}}$. By an averaging argument, we can show that most $x_{\text{pre}} \sim \mathbf{X}_{\text{pre}}$ are biasing. By the combinatorial list-decoding property of $\mathsf{2Ext}$, for every biasing x_{pre}, number of $y \in \{0,1\}^{n/2}$ such that $\Pr[\mathsf{2Ext}(y, (\mathbf{X}|\mathbf{X}_{\text{pre}} = x_{\text{pre}})) = z_{\text{heavy}}] \geq 2^{-m} + \varepsilon_{\mathsf{2Ext}}$ is at most 2^{k_1}. Let $A_{x_{\text{pre}}}$ be the set of such y. If we could sample a uniform element from $A_{x_{\text{pre}}}$, then with probability at least 2^{-k_1}, we would have correctly guessed $g(x_{\text{pre}})$. This is exactly what approximate counting and sampling lets one accomplish. Using approximate counting, one can construct an NP circuit C_1 that given inputs

[6] The proof in [TV00] is specialized to the Hadamard two-source extractor. We follow the more general exposition of this by [BSS25].

[7] By appropriately padding the two-source extractor with zeroes, we can let input lengths of $\mathsf{2Ext}$ be unequal. Doing so does not meaningfully affect the parameters.

y and x_{pre}, can check if $y \in A_{x_{\text{pre}}}$. Then, we can use approximate sampling for C_1, with second element fixed to x_{pre} to sample a uniform element from $A_{x_{\text{pre}}}$. Doing so yields a Σ_2-circuit, that agrees with g with high probability, under distribution $\mathbf{X}$. As $\mathbf{X}$ has very high entropy, we can incur a $2^{-\delta n}$ reduction in approximation factor, for very small $\delta > 0$, to obtain a polynomial size circuit that approximates g with probability at least $2^{-\Omega(n)}$, a contradiction.

2.1.2 Handling smaller entropy

The work of [BSS25] provided a general framework for handling smaller entropy such as $n^{0.99}$ entropy. Later works such as [Sha25a] utilized this framework to obtain extractors with even $n^{0.01}$ entropy, under various hardness assumptions. We briefly review their framework.

The first observation they made is that the argument from the previous section also works if $\mathbf{X}_{\text{pre}}$, instead of being a prefix source, was any other function of $\mathbf{X}$, provided the length of such a source is smaller than the min-entropy in $\mathbf{X}$. Therefore, if there exists a function $\mathsf{Cond} : \{0,1\}^n \to \{0,1\}^m$ such that for a samplable source $\mathbf{X}$ with min-entropy k, we have that $m \leq k/2$ and $\mathbf{H}_\infty(\mathsf{Cond}(\mathbf{X})) \geq 0.99m$, then we would be able to construct an extractor for $\mathbf{X}$ by using the construction $\mathsf{2Ext}(g(\mathsf{Cond}(\mathbf{X})), \mathbf{X})$ (provided $m \geq n^{\Omega(1)}$).

However, they were unable to construct such a Cond, with such large entropy guarantees. They did however construct Cond that was shrinking; Cond had output length smaller than the entropy in $\mathbf{X}$, and could guarantee that $\mathsf{Cond}(\mathbf{X})$ had some amount of entropy, i.e., at least $n^{\Omega(1)}$. This required many non-trivial ideas and connections. One important thing that they still required was that Cond was errorless, i.e., that $\mathsf{Cond}(\mathbf{X})$ has high min-entropy. This differs from the usual output guarantee of condensers which is that the output distribution is statistically close to a distribution which has high min-entropy.

To utilize only the shrinking properties of Cond, they strengthened the guarantees of the hard to approximate function g. Formally, under hardness assumptions, they completely changed the construction of g so that it was hard to approximate not only under uniform distribution, but also under any samplable source with reasonable amount of entropy. Using these objects they constructed the desired extractor.

To construct such a shrinking errorless condenser as well as to construct such g (that is hard on samplable distributions), they used connections between extraction, various notions of distributions being hard to sample, and other pseudorandom objects. From their techniques, especially that of transferring between various notions of distributions that are hard to sample, there does not seem to be an easy way to reduce their assumptions, from hardness against at least Σ_3-circuits, all the way down to hardness against nondeterministic circuits.

2.1.3 Relaxing assumptions

The only known construction of an extractor under hardness against nondeterministic circuits is that of [BGDM23], where they construct extractors for samplable sources with min-entropy $0.99n$.

Their construction was exactly the same as that of [TV00]. However, instead of separately constructing a hard to sample function g and then going through the approximate counting and sampling argument, they analyzed the entire construction in one shot. Doing so raised a lot of technical challenges. A large part of our proof strategy follows their analysis, and so when outlining our techniques, we will mention the corresponding part there.

For their analysis, [BGDM23] required to solve what they coined as the "Gap Probability Maximization problem" (GPM). Here, the input is a randomized circuit $C : \{0,1\}^n \to \{0,1\}$ with the following guarantee: there exists $\gamma > 0$ and some $x^* \in \{0,1\}^n$ such that $\Pr[C(x^*)] = 1 \geq \gamma$, and for $x \neq x^* \in \{0,1\}^n$, $\Pr[C(x) = 1] \leq \gamma/2$. The goal is to output such x^*.

[BGDM23] constructed a polynomial size $\mathsf{NP}_{||}$ circuit [8] solving GPM and doing so, sufficed to prove their result. For our analysis, we will also need to solve GPM, but for the stronger class of randomized nondeterministic circuits.

2.2 Our techniques

We provide a general framework that reduces the task of constructing extractors for samplable sources under nondeterministic hardness assumptions to constructing extractors for two independent sources. Our construction consists of two steps:

1. *Obtaining an elementary somewhere shrunk and random source (elementary SSR source).*[9] Given a samplable source $\mathbf{X}$ with min-entropy k, we obtain a source $\mathbf{W} = (\mathbf{W}_1, \ldots, \mathbf{W}_\ell)$ where each $\mathbf{W}_i$ has length n_w with $n_w = k/3$ and $\ell = \text{poly}(n)$; such that, for some index $j \in [\ell]$, $\mathbf{W}_j \approx_{1/\text{poly}(n)} \mathbf{U}_{n_w}$.

2. *Transforming elementary SSR source to uniform bits.* Here, we use $\mathbf{W}$ as given above, along with the original source $\mathbf{X}$ to obtain the desired extractor.

Our construction deviates from all previous works in two crucial aspects: First, unlike previous works, we do not require that the entire source $\mathbf{W}$ be shrunk. We only need that the length of each block $\mathbf{W}_i$ is smaller than the entropy of $\mathbf{X}$. This is because in the reconstruction argument, the number of $y \in (\{0,1\}^{n_w})^\ell$ that will be "biasing" will be at most $2^{\text{polylog}(n_w \cdot \ell)}$, which is equal to $2^{\text{polylog}(n)}$ since $\ell = \text{poly}(n)$. So, the number of $y_j \in \{0,1\}^{n_w}$ for which there will exist $y_1, \ldots, y_\ell \in \{0,1\}^{n_w}$ such that the entire string will be biasing will be very small, and we will obtain a good approximation to our function g, which we will construct to be $2^{-\Omega(n_w)} = 2^{-\text{polylog}(n)}$ hard to approximate. Second aspect where we differ from previous works is that we allow $\mathbf{W}_j$ to be statistically close to the uniform distribution instead of requiring $\mathbf{W}_j$ itself to have very high min-entropy. We handle this by making the initial error much smaller than the "biasing" probability and by using the fact that $\mathbf{W}_j$, with high probability, outputs from a high min-entropy distribution. We further discuss our proof strategy in Section 2.2.2.

Remark 2.2. *Our transformation from the second step is robust enough that even if $\mathbf{W}_j$ were statistically close to a source with 0.99 entropy rate, then our transformation would produce uniform random bits.*

2.2.1 Obtaining an elementary SSR source

We show how to transform $\mathbf{X}$ as above to obtain such $\mathbf{W}$. This is easily accomplished using well known constructions of explicit strong seeded extractors [GUV09].

[8] An $\mathsf{NP}_{||}$ circuit is allowed to make non-adaptive queries to an NP oracle. See Definition 4.3 for a formal definition.

[9] An elementary somewhere random source consists of many blocks, one of which is "random." This differs from the notion of somewhere random sources which are defined to be convex combinations of elementary somewhere random sources. Our constructions require the former notion.

We say $\mathsf{sExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{n_w}$ is a (k,ε) strong seeded extractor if for all $\mathbf{A}$ with $\mathbf{H}_\infty(\mathbf{A}) \geq k$ and $\mathbf{B} = \mathbf{U}_d$, it holds that $(\mathbf{B}, \mathsf{sExt}(\mathbf{A},\mathbf{B})) \approx_\varepsilon (\mathbf{B}, \mathbf{U}_{n_w})$. This implies that $\mathbb{E}_{b\sim\mathbf{U}_d}[|\mathsf{sExt}(\mathbf{A},b) - \mathbf{U}_m|] \leq \varepsilon$. Therefore, by an averaging argument, there exists $b^* \in \{0,1\}^d$ such that $\mathsf{sExt}(\mathbf{A},b^*) \approx_\varepsilon \mathbf{U}_m$.

For $b \in \{0,1\}^d$, we let $\mathbf{W}_b = \mathsf{sExt}(\mathbf{X},b)$ to obtain the desired transformation to elementary SSR source $\mathbf{W}$. We set the output length of the seeded extractor to $n_w = k/3$ where k is such that we have $\mathbf{H}_\infty(\mathbf{X}) \geq k$. Hence, $\mathbf{W}$ indeed has the desired properties.

We note that the seeded extractor of [GUV09] requires seed length $d = O(\log(n/\varepsilon))$ and so the number of blocks in $\mathbf{W}$ is $\ell = 2^d \leq \mathrm{poly}(n/\varepsilon)$. We set $\varepsilon = 1/\mathrm{poly}(n)$ as well so that $\ell \leq \mathrm{poly}(n)$. Moreover, since sExt is explicit, we can prepare such a source $\mathbf{W}$ in $\mathrm{poly}(n)$ time.

2.2.2 Extracting using elementary SSR source

We now show how given $\mathbf{X}$ and such $\mathbf{W}$ constructed from above, one can obtain uniform random bits.

Remark 2.3. *By a well known reduction of [Sha08] to increase output length of extractors, it suffices to only construct extractors for samplable sources (technically those that are sampled with postselection) that output $O(\log(n))$ bits. Here throughout, we will always assume $m = O(\log(n))$.*

Extracting under Σ_5 hardness We first show how the proof technique of [TV00] outlined in Section 2.1.1 easily yields extractors using such $\mathbf{W}, \mathbf{X}$ under larger complexity assumptions. We make two simplifying assumptions below. First, we assume that there exists $j \in [\ell]$ such that $\mathbf{W}_j \equiv \mathbf{U}_{n_w}$ (instead of $\mathbf{W}_j$ merely being statistically close to $\mathbf{U}_{n_w}$). Second, we assume $k \geq n^\delta$ for some small constant $\delta > 0$.

Under these simplifying assumptions, we proceed as follows. We let $g : \{0,1\}^{n_w} \to \{0,1\}^{n_w}$ be a function that is $2^{-\Omega(n_w)}$ hard to approximate by $\mathrm{poly}(n_w)$ size Σ_2-circuits over the uniform distribution; we can obtain such g using hardness against Σ_5-circuits. Since $n_w = n^\delta$, we can re-parameterize and ensure that g is $2^{-\Omega(n_w)}$ hard to approximate by $\mathrm{poly}(n)$ sized circuits. Using this, our extraction construction will be $\mathsf{2Ext}(g(\mathbf{W}_1), \ldots, g(\mathbf{W}_\ell); \mathbf{X})$ where $\mathsf{2Ext}$ is the two source extractor from [CZ19; Li16] with $\mathrm{polylog}(n)$ entropy requirement.

For the proof of correctness, we argue as follows. Assume the extractor fails and there exists $z_{\text{heavy}} \in \{0,1\}^m$ such that $\Pr[\mathsf{2Ext}(\mathbf{W},\mathbf{X}) = z_{\text{heavy}}] \geq (1+\varepsilon)\cdot 2^{-m}$. Since $\mathbf{W}_j$ has length much smaller than the entropy of $\mathbf{X}$, by chain rule for min-entropy (Lemma 4.10), for most $w_j \sim \mathbf{W}_j$, $\mathbf{X}|(\mathbf{W}_j = w_j)$ has high entropy. For simplicity, assume that this holds for all w_j. Using combinatorial list decoding property of $\mathsf{2Ext}$, for each w_j, the number of "biasing" $y \in \{0,1\}^{n_w\cdot\ell}$—$y$ such that $\Pr[\mathsf{2Ext}(y,\mathbf{X}|(\mathbf{W}_j = w_j)) = z_{\text{heavy}}] \geq 2^{-m} + \varepsilon_{\mathsf{2Ext}}$ is at most $2^{k_1} = 2^{\mathrm{polylog}(\ell\cdot n_w)} = 2^{\mathrm{polylog}(n)}$. Using approximate counting, for a fixed $w_j \sim \mathbf{W}_j$, we can construct a Σ_1 circuit that identifies whether a $y \in \{0,1\}^{n_w\cdot\ell}$ is such that $\Pr[\mathsf{2Ext}(y,\mathbf{X}|(\mathbf{W}_j = w_j)) = z_{\text{heavy}}] \geq 2^{-m} + \varepsilon_{\mathsf{2Ext}}$. Then, using uniform sampling ideas, we can construct a Σ_2-circuit that on input $w_j \sim \mathbf{W}_j$, uniformly samples such a biasing $y \sim \{0,1\}^{n_w\cdot\ell}$. Since $g(w_j)$ is biasing, this $\mathrm{poly}(n)$ size Σ_2 circuit computes $g(w_j)$ with probability at least $2^{-\mathrm{polylog}(n)}$. However, this contradicts the fact that g is $2^{-\Omega(n_w)} = 2^{-n^\delta}$ hard to approximate by $\mathrm{poly}(n)$ sized circuits. This establishes that our construction indeed outputs m uniform random bits.

To extend the argument for the case of when $\mathbf{W}_j$ is statistically close to $\mathbf{U}_{n_w}$, we will open up the construction of the hard to approximate function g and argue that morally, g is still exponentially

hard to approximate for samplable sources that are statistically close to the uniform distribution. Also to extend this to the case of when $k = \text{polylog}(n)$ or $k = O(\log(n))$ and reduce our hardness assumptions, we do not use a modular argument as above and instead execute the above strategy in "one shot", as done in previous works such as [BGDM23; Sha25b].

Extracting under nondeterministic hardness To reduce our assumptions all the way to hardness against nondeterministic circuits, we closely follow the proof strategy of [BGDM23]. So, instead of first constructing a hard to approximate function and then using uniform counting and sampling, we do a holistic analysis. We provide the details below.

Construction We first apply a result of [SU06] that shows that if E is hard for exponential size nondeterministic circuits, then E is also hard for exponential size $\mathsf{NP}_{||}$ circuits. Hence, we work with the latter assumption and let $f : \{0,1\}^r \to \{0,1\}$ be such that f is computable in time $2^{O(r)}$ but requires $2^{\beta r}$ size $\mathsf{NP}_{||}$ circuits (see Definition 4.3). We (carefully) set $r = O(\log(n))$ (it will depend on the size of circuits that are sampling our sources) and consider the finite field vector space $\mathbb{F}_q^t$ where $t = O(1)$ and $q^t = 2^{n_w}$. We then consider the low degree extension $\widehat{f} : \mathbb{F}_q^t \to \mathbb{F}_q$ of f so that $\widehat{f}(\tau(x)) = f(x)$ for some efficiently computable bijective function τ.[10] Moreover, we can ensure that $\widehat{f}$ has total degree $d \leq t \cdot 2^{r/t} = \text{poly}(n)$ and that $\widehat{f}$ is explicitly computable. As $q^t = 2^{n_w}$, we can interpret each $\mathbf{W}_i$ as being a source over $\mathbb{F}_q^t$. With this, our extractor Ext will output $\mathsf{2Ext}(\widehat{f}(\mathbf{W}_1), \ldots, \widehat{f}(\mathbf{W}_\ell); \mathbf{X})$. To obtain polylog$(n)$ entropy extractors, we use the two source extractors from [CZ19; Li16] and for $O(\log(n))$ entropy, we use the two source extractors from [Li23]. We take advantage of the fact that for poly(n) input length and input length n, the parameters of these extractors are affected by a small constant and so if $k \geq \text{polylog}(n)$ or $k \geq C\log(n)$ for a large enough constant C, then we still satisfy the min-entropy requirements of these extractors.

Correctness - setup We proceed by contradiction and assume that our extractor fails. This means there exists some outcome $z_{\text{heavy}} \in \{0,1\}^m$ such that $\Pr[\mathsf{2Ext}(\widehat{f}(\mathbf{W}_1), \ldots, \widehat{f}(\mathbf{W}_\ell); \mathbf{X}) = z_{\text{heavy}}] \geq (1+\varepsilon) \cdot 2^{-m}$. We will use $(\mathbf{W}, \mathbf{X})$ and above property to reconstruct $\widehat{f}$, i.e., come up with a small $\mathsf{NP}_{||}$ circuit computing $\widehat{f}$. Doing so will also let us compute our said hard function f using a small $\mathsf{NP}_{||}$ circuit, a contradiction.

Correctness - high level argument We here sketch our complete high level idea, and provide additional details about a couple of components below.

We use the notation $L_{v_1,v_2} : \mathbb{F} \to \mathbb{F}^t$ to denote the unique line such that $L_{v_1,v_2}(0) = v_1$ and $L_{v_1,v_2}(1) = v_2$. We show that there exists an element $v^* \in \mathbb{F}^t$ such that for 0.99 fraction of $v \in \mathbb{F}^t$, we can test (using randomness and nondeterminism) whether a degree d univariate polynomial $h : \mathbb{F} \to \mathbb{F}$ equals $\widehat{f} \circ L_{v^*,v}$. We then show how to explicitly construct a small $\mathsf{NP}_{||}$ circuit that outputs the description of such h by solving the gap probability maximization problem - GPM for randomized nondeterministic circuits. Finally, we compute $\widehat{f}(v) = \widehat{f} \circ L_{v^*,v}(1) = h(1)$. Since this property holds for 0.99 fraction of v, we have obtained a small $\mathsf{NP}_{||}$ circuit that agrees with $\widehat{f}$ for 0.99 fraction of inputs. We then use well known results regarding self correction of low degree

[10] We are making many simplifying assumptions here regarding divisibility but for the sake of exposition, we will continue doing so.

polynomials [GLRSW91] to obtain a small $\mathsf{NP}_{||}$ circuit that computes $\widehat{f}$ everywhere. This gives us the final contradiction.

Correctness - distinguishing property For $y \in L_{v_1,v_2}$, we write $L^{-1}_{v_1,v_2}(y)$ for the unique $c \in \mathbb{F}$ such that $L_{v_1,v_2}(c) = y$. Using combinatorial list decoding property of $\mathsf{2Ext}$ as well as properties of low degree polynomials and lines, we obtain the following distinguishing test:

Lemma 2.4 (Informal version of Lemma 5.5)**.** *There exists $v^* \in \mathbb{F}^t$ such that for 0.99 fraction of $v \in \mathbb{F}^t$, each of the following holds.*

1. *Let $h : \mathbb{F} \to \mathbb{F}$ be a degree d univariate polynomial such that $h = \widehat{f} \circ L_{v^*,v}$. Then,*

$$\Pr_{(w,x)\sim((\mathbf{W},\mathbf{X})|\mathbf{W}_j \in L_{v^*,v})}\left[\exists y \in \mathbb{F}^\ell : \mathsf{2Ext}\left(y_1, \ldots, h(L^{-1}_{v^*,v}(w_j)), \ldots, y_\ell; x\right) = z_{heavy}\right] \geq 2^{-m}(1+\varepsilon/3).$$

2. *Let $h : \mathbb{F} \to \mathbb{F}$ be a degree d univariate polynomial such that $h \neq \widehat{f} \circ L_{v^*,v}$. Then, either $h(0) \neq \widehat{f}(v^*)$ or the following holds:*

$$\Pr_{(w,x)\sim((\mathbf{W},\mathbf{X})|\mathbf{W}_j \in L_{v^*,v})}\left[\exists y \in \mathbb{F}^\ell : \mathsf{2Ext}\left(y_1, \ldots, h(L^{-1}_{v^*,v}(w_j)), \ldots, y_\ell; x\right) = z_{heavy}\right] \leq 2^{-m}(1+\varepsilon/6).$$

For the first item above, since $h = \widehat{f} \circ L_{v^*,v}$, $h(L^{-1}_{v^*,v}(w_j)) = \widehat{f}(w_j)$. In fact, we argue that the property above will hold for the choice of $y_i = \widehat{f}(w_i)$ in that item. Note that such a condition already holds for (w, x) sampled from $(\mathbf{W}, \mathbf{X})$ by our contradiction assumption. Hence, we carefully pick lines $L_{v^*,v}$ so that the high bias towards z_{heavy} is preserved. For the second item, we carefully use properties of low degree polynomials, lines, the combinatorial list decoding property of $\mathsf{2Ext}$, and the fact that $\mathbf{W}_j$ has length smaller than $\mathbf{H}_\infty(\mathbf{X})$ to obtain the result. To get around the fact that $\mathbf{W}_j$ is δ-close to $\mathbf{U}_{n_w}$ instead of being truly uniform, we express $\mathbf{W}_j$ as a convex combination $\mathbf{W}_j = (1-2\delta)\mathbf{A} + 2\delta B$ where $\mathbf{A}$ is such that $\mathbf{H}_\infty(\mathbf{A}) \geq n_w - 1$ and $\mathbf{B}$ is an arbitrary distribution.[11] We ensure that $\delta < 10^{-6}\varepsilon^2 \cdot 2^{-2m}$ and so barring tiny error terms that get absorbed in the distinguishing error, and in the probability of "good lines" (lines where the distinguishing property holds), we morally pretend that $\mathbf{W}_j$ behaves just like a very high entropy distribution such as $\mathbf{A}$. See Section 5.3 for more details regarding the proof.

Correctness - distinguishing tester For fixed $v \in \mathbb{F}^t$ we construct a small randomized non-deterministic circuit C_{test} that receives as input a degree d univariate polynomial and satisfies the following:

- If $h = \widehat{f} \circ L_{v^*,v}$, then C_{test} outputs 1 with probability at least γ (for some real $\gamma > 0$ that is independent of v). Formally, for such h, with probability at least γ over its randomness, there exists a witness accepting h.
- If $h \neq \widehat{f} \circ L_{v^*,v}$, then C_{test} outputs 1 with probability at most $\gamma/2$ (for the same γ as above). Formally, for such h, with probability at most $\gamma/2$ over its randomness, there exists a witness accepting h.

[11] We do not require that $\mathbf{A}$ or $\mathbf{B}$ are samplable.

We construct C_{test} on input h as follows:

1. If $h(0) \neq \widehat{f}(v^*)$ then output 0.
2. Using randomness, sample $(w, x) \sim (\mathbf{W}, \mathbf{X})$, outputting 0 if $w_j \notin L_{v^*,v}(\mathbb{F})$.
3. Using nondeterminism, output whether there exist $y \in \mathbb{F}^\ell$ such that $\mathsf{2Ext}(y_1, \ldots, h(L^{-1}_{v^*,v}(w_j)), \ldots, y_\ell; x) = z_{\text{heavy}}$.

Using the distinguishing property from above, we see that if $h = \widehat{f} \circ L_{v^*,v}$, then C_{test} accepts with probability at least $\Pr[\mathbf{W}_j \in L_{v^*,v}(\mathbb{F})] \cdot 2^{-m}(1 + \varepsilon/3)$. Furthermore, if $h \neq \widehat{f} \circ L_{v^*,v}$, then C_{test} accepts with probability at most $\Pr[\mathbf{W}_j \in L_{v^*,v}(\mathbb{F})] \cdot 2^{-m}(1 + \varepsilon/6)$. By sampling more elements from $(\mathbf{W}, \mathbf{X})$, we can amplify this difference between acceptance probabilities and obtain C_{test} with the stronger properties as laid out above. See Section 5.1 for further details.

Correctness - solving GPM To finish off the proof, we construct a small $\mathsf{NP}_{||}$ circuit C_{GPM} that solves the gap probability maximization problem for randomized nondeterministic circuits. Here, the input is a randomized nondeterministic circuit $C : \{0,1\}^n \to \{0,1\}$ with the guarantee that there exists $\gamma > 0$ and some $x^* \in \{0,1\}^n$ such that with probability at least γ, there exists a witness accepting x^*. Moreover, for $x \neq x^* \in \{0,1\}^n$, the probability that there exists a witness for x is at most $\gamma/2$. The goal of C_{GPM} is to output such x^*. In particular, we will pass C_{test} from above to C_{GPM} and C_{GPM} will output the description of $\widehat{f} \circ L_{v^*,v}$.

As mentioned earlier, this problem was introduced by [BGDM23] and they solved for the case of when the input circuit C was a randomized circuit (without nondeterminism). We use the exact same algorithm as [BGDM23], and show that it in fact solves GPM against randomized nondeterministic circuits. The main observation we need here is that an NP oracle can test whether a nondeterministic circuit is satisfiable. To see this, by definition, a nondeterministic circuit is satisfiable if there exists an input and a witness where the circuit outputs 1. This is indeed an NP predicate. We provide a completely self-contained construction of C_{GPM} in Section 5.2.

3 Conclusion

We presented two constructions of extractors for samplable sources under nondeterministic hardness: (1) Extractors with $k = \text{poly}(\log(n))$ entropy requirement that output $m = 0.99k$ random bits that are $1/\text{poly}(n)$ close to uniform and (2) extractors with $k = O(\log(n))$ entropy requirement that output 1 random bit that is $O(1)$-close to uniform. We note that these extractors are at the "frontier" of parameter tradeoffs for samplable sources. The situation is quite similar for known explicit constructions of two-source extractors. Among explicit constructions that can handle $1/\text{poly}(n)$ error or larger, there are two extractors that are at the frontier: (1) Extractors for two independent sources by [CZ19; Li16] that require min-entropy $\text{poly}(\log(n))$, output $m = 0.99k$ uniform random bits with error $1/\text{poly}(n)$ and (2) extractors for two independent sources by [Li23] that require min-entropy $O(\log(n))$ that output 1 uniform random bit with $O(1)$ error.

This is not a coincidence. Our constructions use these two extractors as black box objects. Moreover, our techniques are robust and flexible enough that any further improvement in the frontier for extractors for two independent sources will lead to a similar improvement in extractors for samplable sources. It also raises the following natural question: Is it possible to break this

apparent two-source extractor barrier for samplable sources? More concretely, it remains open to construct an extractor for samplable sources that works for min-entropy $O(\log n)$ with polynomially small error.

Our extractors, like prior constructions for samplable sources, rely on two-source extractors in a black box or in a white box way. Hence, it is unclear how using previous techniques one may bypass this. We leave open this question for future investigation.

4 Preliminaries

Notation For an integer $n \in \mathbb{N}$, we let $[n] := \{1, 2, \ldots, n\}$. We use $\circ$ to denote function composition. Given $x_i \in \{0,1\}^n$ and index set $[D]$, we write $\{x_i\}_{i\in[D]}$ to denote the vector which is the concatenation of x_i's. All logarithms are of base 2. We use bold font, such as $\mathbf{X}$, to denote random variables. The support of a random variable $\mathbf{X}$, denoted $\mathrm{Supp}(\mathbf{X})$, contains all elements x such that $\Pr[\mathbf{X} = x] > 0$, and we write $\mathbf{X} \sim V$ if $\mathrm{Supp}(\mathbf{X}) \subseteq V$. For a distribution $\mathbf{D}$, we write $x \sim \mathbf{D}$ to indicate x is sampled from $\mathbf{D}$. We use $\mathbf{U}_m$ to denote the uniform random variable over $\{0,1\}^m$. In the context of pseudorandomness, we often refer to a random variable as a "source."

4.1 Circuits

4.1.1 Basic definitions

Definition 4.1 (Nondeterministic circuit)**.** *For $n, w \in \mathbb{N}$, we say $C : \{0,1\}^n \to \{0,1\}$ is a nondeterministic circuit if C is fed two kinds of input wires—actual input from $\{0,1\}^n$ and witnesses from $\{0,1\}^w$. We interpret C as accepting $x \in \{0,1\}^n$ if there exists $y \in \{0,1\}^w$ such that $C(x; y) = 1$.*

Definition 4.2 (Randomized Nondeterministic circuit)**.** *For $n, r, w \in \mathbb{N}$, we say $C : \{0,1\}^n \to \{0,1\}$ is a randomized nondeterministic circuit if C is fed three kinds of input wires—actual input from $\{0,1\}^n$, random bits from $\{0,1\}^r$ and witnesses from $\{0,1\}^w$. We interpret C as accepting $x \in \{0,1\}^n$ with probability $p = \Pr_{y\sim\mathbf{U}_r}[\exists z \in \{0,1\}^w : C(x; y; z) = 1]$.*

Definition 4.3 ($\mathsf{NP}_{||}$—Non-adaptive NP circuit)**.** *For $n, m \in \mathbb{N}$, we say $C : \{0,1\}^n \to \{0,1\}^m$ is a non-adaptive NP-circuit, denoted as $\mathsf{NP}_{||}$-circuit, if C is additionally allowed access to gates solving an NP-complete problem with the restriction that queries to such gates must be made non-adaptively.* [12]

Definition 4.4 (Σ_i-circuit)**.** *For $i, n, m \in \mathbb{N}$, we say $C : \{0,1\}^n \to \{0,1\}^m$ is a Σ_i-circuit if C is allowed gates solving the canonical Σ_i-complete language.*

We emphasize the distinction between NP-circuits which are allowed NP oracle gates, and nondeterministic circuits which are circuits that take witnesses along with actual inputs.

4.1.2 Equivalence of hardness assumptions

We will need the following result relating hardness assumptions for E:

[12] To rigorously formalize non-adaptivity, we say that we can associate two circuits C_1 and C_2 with C where C_1 receives the input and outputs queries to be made to the NP oracle gates. Then, C_2 receives the input, as well as the answers to the queries that C_1 made, and provides the output.

Theorem 4.5 ([SU06])**.** *The following are equivalent:*

1. *There exists a language $L \in \mathsf{E} = \mathsf{DTIME}(2^{O(n)})$ such that L requires $2^{\Omega(n)}$ size $\mathsf{NP}_{||}$-circuits.*
2. *There exists a language $L \in \mathsf{E} = \mathsf{DTIME}(2^{O(n)})$ such that L requires $2^{\Omega(n)}$ size nondeterministic circuits.*

4.2 Defining samplable sources

Definition 4.6 (Samplable sources)**.** *We say $\mathbf{X} \sim \{0,1\}^n$ is samplable by a circuit $C : \{0,1\}^r \to \{0,1\}^n$, if $\mathbf{X} \equiv C(\mathbf{U}_r)$.*

We say $\mathbf{X}$ is samplable by size s circuits to mean that there exists a deterministic circuit C of size s such that $\mathbf{X}$ is samplable by C.

Definition 4.7 (Samplable sources with postselection)**.** *We say $\mathbf{X} \sim \{0,1\}^n$ is samplable by circuits $C : \{0,1\}^r \to \{0,1\}^n$ with postselection by circuit $P : \{0,1\}^r \to \{0,1\}$ if $\mathbf{X} \equiv (C(\mathbf{Y})|P(\mathbf{Y}) = 1)$ where $\mathbf{Y} = \mathbf{U}_r$.*

We say $\mathbf{X}$ is samplable with postselection by size s circuits to mean that there exist deterministic circuits C, P of size s each such that $\mathbf{X}$ is samplable by C, with postselection by P.

We will also naturally extend this notion to consider sources where $\mathbf{X}$ is samplable by size s deterministic circuits with postselection by size s nondeterministic circuits.

4.3 Probability

We recall some useful notions and tools from probability.

Definition 4.8 (Statistical distance)**.** *The statistical distance between two random variables $\mathbf{X}, \mathbf{Y} \sim V$ is*

$$|\mathbf{X} - \mathbf{Y}| := \max_{S \subseteq V} |\Pr[\mathbf{X} \in S] - \Pr[\mathbf{Y} \in S]| = \frac{1}{2} \sum_{v \in V} |\Pr[\mathbf{X} = v] - \Pr[\mathbf{Y} = v]|$$

We say $\mathbf{X}, \mathbf{Y}$ are ε-close, written $\mathbf{X} \approx_\varepsilon \mathbf{Y}$, if $|\mathbf{X} - \mathbf{Y}| \leq \varepsilon$. Note that this ε error is *additive*.

Definition 4.9 (Min-entropy)**.** *The min-entropy of $\mathbf{X}$ is defined as*

$$\mathbf{H}_\infty(\mathbf{X}) = \min_{x \in \mathrm{supp}(\mathbf{X})} \log \frac{1}{\Pr[\mathbf{X} = x]}$$

We will need the following useful chain rule for min-entropy.

Lemma 4.10 (Min-entropy chain rule, [MW97])**.** *For any random variables $\mathbf{X} \sim X$ and $\mathbf{Y} \sim Y$ and $\varepsilon > 0$,*

$$\Pr_{y \sim \mathbf{Y}}[H_\infty(\mathbf{X} \mid \mathbf{Y} = y) \geq H_\infty(\mathbf{X}) - \log |\mathrm{Supp}(\mathbf{Y})| - \log(1/\varepsilon)] \geq 1 - \varepsilon.$$

We will also need the following observation regarding the effect of conditioning on min-entropy:

Lemma 4.11. *Let $\mathbf{X} \sim X$ be an arbitrary random variable and let E be an event that occurs with probability at least $p > 0$. Then, $\mathbf{H}_\infty(\mathbf{X}|E) \geq \mathbf{H}_\infty(\mathbf{X}) - \log(1/p)$.*

Proof. Let $x \in \mathbf{X}$ be arbitrary. It suffices to show that $\Pr[(\mathbf{X}|E) = x] \leq 2^{-\mathbf{H}_\infty(\mathbf{X})+\log(1/p)}$. We easily compute that,

$$\Pr[(\mathbf{X}|E) = x] = \frac{\Pr[\mathbf{X} = x \wedge E]}{\Pr[E]} \leq \frac{2^{-\mathbf{H}_\infty(\mathbf{X})}}{p} = 2^{-\mathbf{H}_\infty(\mathbf{X})+\log(1/p)},$$

proving the desired claim. □

4.3.1 Characterizing smooth min entropy

Definition 4.12 (Smooth min-entropy)**.** *We say* $\mathbf{X}$ *has smooth min-entropy* $\mathbf{H}^\gamma_\infty(\mathbf{X}) \geq k$ *if there exists* $\mathbf{Y}$ *such that* $|\mathbf{X} - \mathbf{Y}| \leq \gamma$ *and* $\mathbf{H}_\infty(\mathbf{Y}) \geq k$*.*

We will need the following well known claim:

Claim 4.13. *Let* $\mathbf{X} \sim \{0,1\}^n$ *be such that* $\mathbf{H}^\gamma_\infty(\mathbf{X}) \geq k$*. Then, for all* $S \subset \{0,1\}^n$*, it holds that*

$$\Pr_{x \sim \mathbf{X}}[x \in S] \leq |S| \cdot 2^{-k} + \gamma.$$

Proof. Let $\mathbf{X}' \sim \{0,1\}^n$ be such that $|\mathbf{X} - \mathbf{X}'| \leq \gamma$ and $\mathbf{H}_\infty(\mathbf{X}') \geq k$. Then, $\Pr_{x' \sim \mathbf{X}'}[x' \in S] \leq |S| \cdot 2^{-k}$. By definition of statistical distance, we have that

$$\Pr_{x \sim \mathbf{X}}[x \in S] \leq \Pr_{x' \sim \mathbf{X}'}[x' \in S] + \gamma \leq |S| \cdot 2^{-k} + \gamma$$

as desired. □

From this, we will obtain the following result regarding smooth min-entropy and its weight on heavy elements.

Claim 4.14. *Let* $\mathbf{X} \sim \{0,1\}^n$ *be such that* $\mathbf{H}^\gamma_\infty(\mathbf{X}) \geq k$*. Let* $H = \{x \in \{0,1\}^n : \Pr[\mathbf{X} = x] \geq 2^{-k+1}\}$*. Then,* $\Pr_{x \sim \mathbf{X}}[x \in H] \leq 2\gamma$*.*

Proof. We will show that $|H| \leq \gamma \cdot 2^k$. Using Claim 4.13, we will obtain our desired claim. We now show the upper bound on size of H. Let $\mathbf{X}' \sim \{0,1\}^n$ be such that $|\mathbf{X} - \mathbf{X}'| \leq \gamma$ and $\mathbf{H}_\infty(\mathbf{X}') \geq k$. By definition of statistical distance,

$$\sum_{x \in H} \Pr[\mathbf{X} = x] - \Pr[\mathbf{X}' = x] \leq \gamma. \tag{*}$$

We observe that $\Pr[\mathbf{X}' = x] \leq 2^{-k}$ for all x and $\Pr[\mathbf{X} = x] > 2^{-k+1}$ for all $x \in H$. From this we obtain that

$$\sum_{x \in H} \Pr[\mathbf{X} = x] - \Pr[\mathbf{X}' = x] \geq \sum_{x \in H} \left(2^{-k+1} - 2^{-k}\right) \geq |H| \cdot 2^{-k}.$$

Combining this with Equation (*), we infer that $|H| \leq \gamma \cdot 2^k$ as desired. □

4.3.2 Inequalities and tail bounds

We find it useful to record the following well known inequality:

Claim 4.15. *For all $0 \leq \varepsilon \leq 1$, the following holds:*

$$1 + \varepsilon \leq e^{\varepsilon} \leq 1 + 3\varepsilon$$

We state Markov's inequality.

Lemma 4.16 (Markov's inequality)**.** *Let $\mathbf{X} \sim \mathbb{R}$ be such that all elements in $\mathrm{Supp}(\mathbf{X})$ are non-negative. Then, for $a > 0$:*

$$\Pr[\mathbf{X} \geq a] \leq \frac{\mathbb{E}[\mathbf{X}]}{a}.$$

We record the standard (multiplicative) Chernoff bound.

Lemma 4.17 (Chernoff bound)**.** *Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be independent random variables taking values in $\{0, 1\}$. Let $\mathbf{X} = \sum_i \mathbf{X}_i$. Let $\mu = \mathbb{E}[\mathbf{X}]$. Then, for all $\delta \geq 0$, the following holds:*

$$\Pr[\mathbf{X} \geq (1 + \delta)\mu] \leq \exp(-\delta^2 \mu/(2 + \delta))$$

Also for $0 \leq \delta \leq 1$, the following holds:

$$\Pr[\mathbf{X} \leq (1 - \delta)\mu] \leq \exp(-\delta^2 \mu/2)$$

We will rely on the following standard bound on variance of a bounded random variable.

Lemma 4.18 (Popoviciu's inequality on variances)**.** *Let $a \leq b \in \mathbb{R}$ and let $\mathbf{X} \sim \mathbb{R}$ be a random variable such that for all $x \in \mathrm{Supp}(\mathbf{X})$, $a \leq x \leq b$. Then,*

$$\mathrm{Var}(\mathbf{X}) \leq \frac{1}{4} \cdot (b - a)^2$$

We will also make use of the well known tail bound from variance.

Lemma 4.19 (Chebyshev's inequality)**.** *For any random variable $\mathbf{X} \sim \mathbb{R}$ and any $\varepsilon > 0$,*

$$\Pr[|\mathbf{X} - \mu| \geq \varepsilon] \leq \frac{\mathrm{Var}(\mathbf{X})}{\varepsilon^2}$$

Using the above two, we easily obtain the following tail bound regarding sums of bounded pairwise independent random variables.

Lemma 4.20. *Let $a, b \in \mathbb{R}$ and let $\mathbf{X}_1, \ldots, \mathbf{X}_n \sim \mathbb{R}$ be pairwise independent random variables such that for all $i \in [n]$, and $x_i \in \mathrm{Supp}(\mathbf{X}_i)$, it holds that $a \leq x_i \leq b$. Let $\mathbf{X} = \sum_{i=1}^n \mathbf{X}_i$. Then, for any $\varepsilon > 0$ it holds that*

$$\Pr\left[|\mathbf{X} - \mathbb{E}(\mathbf{X})| \geq \varepsilon\right] \leq \frac{n(b - a)^2}{4\varepsilon^2}.$$

Proof. We apply Lemma 4.18 to each $\mathbf{X}_i$ to obtain that $\mathrm{Var}(\mathbf{X}_i) \leq \frac{(b-a)^2}{4}$. Since $\{\mathbf{X}_i\}_{i=1}^n$ are pairwise independent, $\mathrm{Var}(\mathbf{X}) = \sum_{i=1}^n \mathrm{Var}(\mathbf{X}_i) \leq \frac{n(b-a)^2}{4}$. Applying Lemma 4.19 to $\mathbf{X}$, we obtain the desired result. □

4.4 Pseudorandom objects

4.4.1 Basic definitions

Definition 4.21 (Code)**.** *We say $\mathcal{C}$ is an $[N, D]_q$ code if $\mathcal{C} \subset [q]^N$ is such that for all $C_1 \neq C_2 \in \mathcal{C}$, the Hamming distance between C_1 and C_2 is at least D.*

Definition 4.22 (Pairwise independent hash functions)**.** *For $n, M \in \mathbb{N}$, we say a family of functions $\mathcal{H} = \{h : \{0,1\}^n \to [M]\}$ is pairwise independent if:*

- *For all $x \in \{0,1\}^n$, and $h \sim \mathbf{U}_{\mathcal{H}}$, the random variable $\mathbf{h}(\mathbf{x})$ is uniform over $[M]$.*
- *For all $x_1 \neq x_2 \in \{0,1\}^n$, and $h \sim \mathbf{U}_{\mathcal{H}}$, the random variables $\mathbf{h}(\mathbf{x_1})$ and $\mathbf{h}(\mathbf{x_2})$ are independent.*

We next define condensers that are relaxations of extractors.

Definition 4.23 (Condenser)**.** *A function $\mathsf{Cond} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(k \to \ell, \gamma)$-condenser if for every (n,k)-source $\mathbf{X}$, $\mathsf{Cond}(\mathbf{X}, \mathbf{U}_d)$ is γ-close to some (m, ℓ)-source.*

We recall the definition of two-source extractors, an ingredient that is crucially used in our work.

Definition 4.24 (Two-source extractor)**.** *For $n_1, n_2, m, k_1, k_2 \in \mathbb{N}$ and $\varepsilon > 0$, we say $\mathsf{2Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ is a (k_1, k_2, ε) two-source extractor if for all random variables $\mathbf{X}_1 \sim \{0,1\}^{n_1}, \mathbf{X}_2 \sim \{0,1\}^{n_2}$ such that $\mathbf{H}_\infty(\mathbf{X}_1) \geq k_1$ and $\mathbf{H}_\infty(\mathbf{X}_2) \geq k_2$, it holds that*

$$|\mathsf{2Ext}(\mathbf{X}_1, \mathbf{X}_2) - \mathbf{U}_m| < \varepsilon.$$

We will need the following combinatorial list decoding property of two-source extractors:

Lemma 4.25. *Let $\mathsf{2Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ be a (k_1, k_2, ε) two-source extractor. Then, for all $\mathbf{X}_2 \sim \{0,1\}^{n_2}$ such that $\mathbf{H}_\infty(\mathbf{X}_2) \geq k_2$ and $z \in \{0,1\}^m$, it holds that*

$$|\{x \in \{0,1\}^{n_1} : \Pr[\mathsf{2Ext}(x, \mathbf{X}_2) = z] \geq 2^{-m} + \varepsilon\}| \leq 2^{k_1}.$$

Proof. We proceed by contradiction and assume this does not hold. Let $B = \{x \in \{0,1\}^{n_1} : \Pr[\mathsf{2Ext}(x, \mathbf{X}_2) = z] \geq 2^{-m} + \varepsilon\}$. Then, by assumption, we have that $|B| \geq 2^{k_1}$. Let $\mathbf{X}_1 \sim \{0,1\}^{n_1}$ be a random variable that is uniformly distributed over B and is independent of $\mathbf{X}_2$. Then, we see that $\mathbf{H}_\infty(\mathbf{X}_1) \geq k_1$. Also by assumption,

$$\Pr[\mathsf{2Ext}(\mathbf{X}_1, \mathbf{X}_2) = z] = \sum_{x \in B} \frac{1}{|B|} \cdot \Pr[\mathsf{2Ext}(x, \mathbf{X}_2) = z] \geq 2^{-m} + \varepsilon$$

However, this contradicts the fact that $\mathsf{2Ext}$ is a (k_1, k_2, ε) two-source extractor. □

Definition 4.26 (Seeded extractor)**.** *For $k, n \in \mathbb{N}$ and $\varepsilon \geq 0$, we say $\mathsf{sExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a strong seeded extractor if for all $\mathbf{X} \sim \{0,1\}^n$ such that $\mathbf{H}_\infty(\mathbf{X}) \geq k$ and $\mathbf{Y} = \mathbf{U}_d$, we have that $(\mathbf{Y}, \mathsf{sExt}(\mathbf{X}, \mathbf{Y})) \approx_\varepsilon (\mathbf{Y}, \mathbf{U}_m)$.*

4.4.2 Previous work

We will use the following constructions of seeded extractors.

Theorem 4.27 ([GUV09]). *There exists a universal constant C such that for all n, k, ε where $k \geq O(\log(1/\varepsilon))$, there exists an explicit (k, ε)-strong seeded extractor $\mathsf{sExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = C\log(n/\varepsilon)$ and $m \geq k/2$. Furthermore, the runtime of the extractor only depends on n.*

We will need the following explicit construction of two-source extractors.

Theorem 4.28 ([CZ19; Li16]). *There exists a universal constant $C > 0$ such that for all $n, k \in \mathbb{N}$ and $\varepsilon > 0$ such that $(\log(n/\varepsilon))^C \leq k \leq n$, there exists an explicit (k, k, ε) two-source extractor $\mathsf{2Ext} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ running in time $\mathrm{poly}(n, 1/\varepsilon)$ where $m = k/C$.*

Theorem 4.29 ([Li23]). *For all constant $\varepsilon > 0$, there exists $C = C(\varepsilon) \geq 1$ such that for all $n \in \mathbb{N}$, there exists an explicit $(C\log(n), C\log(n), \varepsilon)$-two-source extractor $\mathsf{2Ext} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. Furthermore, $\mathsf{2Ext}$ is computable in time n^C.*

We will need the following explicit construction of pairwise independent hash functions:

Lemma 4.30. *For all $n, M \in \mathbb{N}$, there exists a family of pairwise independent hash functions $\mathcal{H} = \{h : \{0,1\}^n \to [M]\}$ such that we can sample $h \sim \mathbf{U}_{\mathcal{H}}$ using $O(n + \log(M))$ random bits and the function h can be computed in time $\mathrm{poly}(n, \log(M))$.*

4.5 Lines in finite fields

We introduce some convenient notation for lines over finite fields.

Definition 4.31. *For a finite field $\mathbb{F}$, and $x, y \in \mathbb{F}^t$, we define line $L_{x,y} : \mathbb{F} \to \mathbb{F}^t$ as $L_{x,y}(c) = (1-c) \cdot x + c \cdot y$. In particular, $L_{x,y}(0) = x$ and $L_{x,y}(1) = y$.*

We say a line $L_{x,y}$ is non-trivial if $x \neq y$. For a non-trivial line $L_{x,y}$ and a point $v \in \mathbb{F}^t$ such that $v \in L_{x,y}(\mathbb{F})$, we let $L_{x,y}^{-1}(v)$ be the unique element $c \in \mathbb{F}$ such that $L_{x,y}(c) = v$.

5 Extracting from elementary SSR sources

In this section, we will show how to extract uniform bits from an elementary somewhere shrunk and random source (elementary SSR source) along with a samplable source.

Theorem 5.1 (The main theorem). *Assume E is hard for exponential size nondeterministic circuits. Then, there exists a constant C such that for all constants $C_{out}, C_{\mathsf{Samp}}$ and all large enough integer parameters n_w, n, m, ℓ, k_y, and all $\Delta, \varepsilon, \varepsilon_{in}$ satisfying*

1. $\varepsilon_{in} \leq \varepsilon^2 \cdot 2^{-2m}/C$.
2. $(\log(n))^C \leq n_w \leq k_y/3$.
3. $\Delta \leq n_w/C$.
4. $2^m = 1/\varepsilon = n^{C_{out}}$.

the following holds. There exists a $\mathrm{poly}(n)$ *time computable function* $\mathsf{Ext} : (\{0,1\}^{n_w})^\ell \times \{0,1\}^n \to \{0,1\}^m$ *with the following property: for all* $\mathbf{W} = (\mathbf{W}_1, \ldots, \mathbf{W}_\ell) \sim (\{0,1\}^{n_w})^\ell, \mathbf{Y} \sim \{0,1\}^n$ *such that*

- $(\mathbf{W}, \mathbf{Y})$ *are samplable by size* $n^{C_{\mathsf{Samp}}}$ *circuits with postselection by size* $n^{C_{\mathsf{Samp}}}$ *nondeterministic circuits,*
- *there exists* $j \in [\ell]$ *such that* $\mathbf{H}_\infty^{\varepsilon_{in}}(\mathbf{W}_j) \geq n_w - \Delta$*, and*
- $\mathbf{H}_\infty(\mathbf{Y}) \geq k_y$*,*

we have

$$\mathsf{Ext}(\mathbf{W}; \mathbf{Y}) \approx_\varepsilon \mathbf{U}_m.$$

To prove this, we will show that there exists a small circuit that reconstructs any low degree polynomial over a finite field if the extractor fails. Before formally defining it, we find it useful to explicitly state the assumptions that we will use for that theorem, and also in many helper lemmas for it.

Assumption 5.2. *Let* $\ell, d, t, k_1, k_2, n, m, q \in \mathbb{N}$*, let* $0 < \varepsilon, \varepsilon_{in}, \varepsilon_{\mathsf{2Ext}} < 1$ *and let* $\Delta > 0$*. Assume that* q *is a power of two and let* $\mathbb{F} = \mathbb{F}_q$ *be the finite field of size* q*. Let* $\widehat{f} : \mathbb{F}_q^t \to \mathbb{F}_q$ *be a polynomial with total degree* d*. Let* $\mathsf{2Ext} : \{0,1\}^{\ell \cdot \log(q)} \times \{0,1\}^n \to \{0,1\}^m$ *be an explicit* $(k_1, k_2, \varepsilon_{\mathsf{2Ext}})$ *two-source extractor. Let* $\mathbf{W} = (\mathbf{W}_1, \ldots, \mathbf{W}_\ell) \sim (\mathbb{F}_q^t)^\ell$ *and let* $\mathbf{Y} \sim \{0,1\}^n$*. Let* $j \in [\ell]$ *be such that* $\mathbf{H}_\infty^{\varepsilon_{in}}(\mathbf{W}_j) \geq t \log(q) - \Delta$*. Assume there exists* $z_{heavy} \in \{0,1\}^m$ *such that*

$$\Pr[\mathsf{2Ext}(\widehat{f}(\mathbf{W}_1), \ldots, \widehat{f}(\mathbf{W}_\ell); \mathbf{Y}) = z_{heavy}] \geq (1 + \varepsilon) \cdot 2^{-m}.$$

Furthermore, assume that the following constraints are met:

1. $\varepsilon_{\mathsf{2Ext}} \leq \frac{\varepsilon}{48} \cdot 2^{-m}$
2. $\varepsilon_{in} \leq \varepsilon_{\mathsf{2Ext}}^2 \cdot 10^{-6}$
3. $\mathbf{H}_\infty(\mathbf{Y}) - 2t \log(q) - m - \log(24/\varepsilon) \geq k_2$
4. $q \geq \frac{18432 \cdot d \cdot 2^{2\Delta + 2m + 2k_1}}{\varepsilon^2}$.
5. $q^{-t} + \frac{2^{2(\Delta+1)}}{q} + \frac{9 \cdot 2^{2(m+\Delta)}}{\varepsilon^2 q} + 4\sqrt{\varepsilon_{in}} + \frac{48 \cdot 2^{2k_1 + m}}{\varepsilon} \cdot \frac{d}{q} \leq 0.01$.

With this, our reconstruction theorem is formally defined as follows.

Theorem 5.3 (Reconstruction from biased extractor)**.** *Suppose we are in the setting of Assumption 5.2. We additionally make the following computational assumptions:*

1. *Assume that* $(\mathbf{W}; \mathbf{Y})$ *is samplable by size* s_{Samp} *circuits with postselection by size* s_{Samp} *nondeterministic circuits.*
2. *Assume that the two source extractor* $\mathsf{2Ext}$ *runs in time* s_{2Ext} *on any input.*

Then, there exists an $\mathsf{NP}_{||}$ *circuit of size* $\mathrm{poly}(1/\varepsilon, s_{\mathsf{Samp}}, s_{\mathsf{2Ext}}, \ell, t, \log(q), n, d)$ *computing* $\widehat{f}$*.*

We prove this result in Section 5.1. Using this, we prove our main theorem by setting parameters appropriately as follows:

Proof of Theorem 5.1. Before we describe our construction, we apply Theorem 4.5 to infer that our assumption implies that E is exponentially hard for $\mathsf{NP}_{||}$ circuits. Therefore, there exist $0 < \beta \leq 1 \leq B$ such that for all r, there exists a function $f : \{0,1\}^r \to \{0,1\}$ such that $f \in \mathsf{DTIME}(2^{B \cdot r})$ and that f requires $\mathsf{NP}_{||}$ circuits of size at least $2^{\beta \cdot r}$.

Our constructions will have two ingredients—low degree extension of a hard function and an appropriate two source extractor. Our final extractor will encode each $\mathbf{W}_i$ using the low degree extension and pass them to the two-source extractor. Then, using reconstructive properties from Theorem 5.3, we will show that if our extractor fails, then our circuit lower bound assumption will be violated.

Setting parameters for the construction We observe that $\ell \leq n^{C_{\mathsf{Samp}}}$ and we will keep using this fact below. Also without loss of generality, we assume $C_{\text{out}} \leq C_{\mathsf{Samp}}$.

We let $0 < \alpha < 10^{-6}$ be a small enough constant, that we set later. Let $t = \left\lceil \frac{1}{\alpha \cdot \beta} \right\rceil$. Let $n_{w'} = \lceil n_w/(t+1) \rceil \cdot t$. Let $q = 2^{\lceil n_w/(t+1) \rceil}$. Let $r = \left\lceil \frac{C_{\mathsf{Samp}} \log(n)}{\alpha \cdot \beta} \right\rceil$. We will set our universal constant C to be much larger than t.

Setting up low degree extension We first setup the low degree extension. Let $f : \{0,1\}^r \to \{0,1\}$ be the function from our circuit lower bound assumption which will be hard for size $n^{\frac{C_{\mathsf{Samp}}}{\alpha}}$ $\mathsf{NP}_{||}$-circuits. Let $H \subset \mathbb{F}_q$ be any subset of size $\lceil 2^{r/t} \rceil$. Then, we can define $\tau : \{0,1\}^r \to H^t$ to be any natural injective map. We let $\widehat{f} : \mathbb{F}_q^t \to \mathbb{F}_q$ be the low degree extension of f—such that for all $x \in \{0,1\}^r$, $\widehat{f}(\tau(x)) = f(x)$. Moreover, we can ensure that $\widehat{f}$ is computable in time $\text{poly}(n, 2^r) = \text{poly}(n^{C_{\mathsf{Samp}}})$ and that $\widehat{f}$ will have total degree at most $t \cdot 2^{r/t} \leq t \cdot n^{C_{\mathsf{Samp}}+1} \leq n^{C_{\mathsf{Samp}}+2}$.

Setting up the two-source extractor By appropriately padding with zeroes, we can instantiate the two source extractor from Theorem 4.28 to obtain a $(\text{polylog}(n), \text{polylog}(n), \varepsilon_{\mathsf{2Ext}})$ two-source extractor $\mathsf{2Ext} : \{0,1\}^{\ell \log(q)} \times \{0,1\}^n \to \{0,1\}^m$ where $\varepsilon_{\mathsf{2Ext}} = \varepsilon \cdot 2^{-m}/C^{1/3}$ where we set the constant C later. We can indeed instantiate such an extractor since $\ell \cdot \log(q) \leq \ell \cdot n \leq n^{C_{\mathsf{Samp}}+1}$ and the two source extractor of Theorem 4.28 requires $\text{poly}(\log(n))$ entropy, which is only affected by a constant factor even when padding with $\text{poly}(n)$ many zeroes.

We mention here that we will take the universal constant C to be large enough so that the entropy requirement of $\mathsf{2Ext}$ in each source will be smaller than $(\log(n))^{C/100}$. We will repeat this fact again later on. Note that padding input to $\mathsf{2Ext}$ with $\text{poly}(n)$ many zeroes only affects this inequality by a constant factor which is completely fine.

Also, since $\ell \cdot \log(q) \leq \ell \cdot n \leq n^{C_{\mathsf{Samp}}+1}$, we have that $\mathsf{2Ext}$ runs in time $\text{poly}(n, \varepsilon_{\mathsf{2Ext}}) = \text{poly}(n)$.

Interpreting input in finite field Since $t \log(q) = \lceil n_w/(t+1) \rceil \cdot t \leq n_w$, for any $v \in \{0,1\}^{n_w}$, we can appropriately truncate the source and interpret v as an element of $\mathbb{F}_q^t$.

The construction On input $w = (w_1, \ldots, w_\ell)$ and y, for each $i \in [\ell]$, we truncate each w_i as described above to obtain $w_i' \in \mathbb{F}_q^t$. Using these, we output $\mathsf{2Ext}(\widehat{f}(w_1'), \ldots, \widehat{f}(w_\ell'); y)$.

Correctness We let $\mathbf{W}' = (\mathbf{W}'_1, \ldots, \mathbf{W}'_\ell) \sim (\mathbb{F}_q^t)^\ell$ where $\mathbf{W}'_i$ is the projection of the source $\mathbf{W} \sim \{0,1\}^{n_w}$ onto $\left\lceil \frac{n_w}{t+1} \right\rceil \cdot t$ bits. We observe that the entropy gap of $\mathbf{W}'_j$ is also at most the entropy gap of $\mathbf{W}_j$. In particular, $\mathbf{H}_\infty^{\varepsilon_{\text{in}}}(\mathbf{W}'_j) \geq n_{w'} - \Delta$, where $n_{w'}$ is the length of each w'_i.

We now proceed by contradiction and assume the extractor Ext fails, i.e., $|\mathsf{Ext}(\mathbf{W}, \mathbf{Y}) - \mathbf{U}_m| > \varepsilon$. By an averaging argument, this implies there exists $z \in \{0,1\}^m$ such that $\Pr[\mathsf{Ext}(\mathbf{W}, \mathbf{Y}) = z] \geq (1+\varepsilon) \cdot 2^{-m}$. By construction of Ext, this means that $\Pr[\mathsf{2Ext}(\widehat{f}(\mathbf{W}'_1), \ldots, \widehat{f}(\mathbf{W}'_\ell); \mathbf{Y}) = z] \geq (1+\varepsilon) \cdot 2^{-m}$. We now apply Theorem 5.3 to obtain a small $\mathsf{NP}_{||}$ circuit computing $\widehat{f}$. Before we do that, we need to verify that we indeed satisfy the assumptions laid out in Assumption 5.2. We briefly sketch the argument for each of the 5 assumptions laid out there (we will implicitly keep using the fact that our constant C is large enough):

1. This follows from our setting of $\varepsilon_{\mathsf{2Ext}}$.
2. This follows from our setting of ε_{in}.
3. By plugging in parameters, the inequality simplifies to showing:

   $$k_y - 2n_w - m - \log(24/\varepsilon) \geq k_2.$$

   By assumption, $k_y - 2n_w \geq (\log(n))^C$ and the remaining subtracted terms on the left side are $O(\log(n))$. We ensure that our constant C is large enough so that $(\log(n))^{C/100} \geq k_2$ and the above inequality is satisfied.
4. After taking logarithm and plugging in parameters, the inequality simplifies to showing:

   $$\alpha \cdot \beta \cdot n_{w'} - 2\Delta \geq O(1) + (C_{\mathsf{Samp}} + 2)\log(n) + 2m + 2k_1$$

   As $\Delta \leq n_w/C$ (we take it to be bigger than $1/(\alpha\beta) = t$), and $n_{w'} \geq n_w/2$, the left side is at least $\text{poly}(\log(n))$. Since the right side is only $O(\log(n))$, the requirement is satisfied.
5. The first term is easily $o(1)$. For the second term, we take logarithms and simplify, and argue as above that $\alpha \cdot \beta \cdot n_{w'} - 2\Delta \geq \text{poly}(\log(n))$ to obtain that this term is $o(1)$. Using our assumption regarding ε_{in}, the third term is also $o(1)$. For the fourth term, we again take logarithms and simplify and first argue as above that $\alpha \cdot \beta \cdot n_{w'} - k_1 \geq \text{poly}(\log(n))$ (this follows since our initial constant C is large enough) and that the remaining terms are $O(\log(n))$, so that the contribution of the fourth term is also $o(1)$ as desired.

Therefore, we indeed satisfy all the constraints and can apply Theorem 5.3 to obtain an $\mathsf{NP}_{||}$ circuit $D_{\widehat{f}} : \mathbb{F}_q^t \to \mathbb{F}$ of size $\text{poly}(1/\varepsilon, n^{C_{\mathsf{Samp}}}, \ell, t, \log(q), n)$ computing $\widehat{f}$ (we here incorporated the runtime of $\mathsf{2Ext}$ and utilized the fact that $d \leq n^{C_{\mathsf{Samp}}+2}$). Using $D_{\widehat{f}}$, and injective property of τ, we can obtain an $\mathsf{NP}_{||}$ circuit $D_f : \{0,1\}^r \to \{0,1\}$ such that $D_f(x) = f(x)$ for all $x \in \{0,1\}^r$. Moreover, the size of D_f will be same as that of $D_{\widehat{f}}$. So, to obtain a contradiction to the hardness assumption of f, it suffices to show that size of $D_{\widehat{f}}$ is at most $2^{\beta r} = n^{C_{\mathsf{Samp}}/\alpha}$.

We can without loss of generality upper bound the size of $D_{\widehat{f}}$ by $n^{C_{\mathsf{Samp}} \cdot C_1}$ where $C_1 \geq 1$ is a universal constant, independent of C_{Samp}. Therefore, we can set $\alpha = \frac{1}{10^6 \cdot C_1}$ to obtain the desired contradiction. □

5.1 Reconstructing low degree polynomial from failed extractor

In this subsection, we will prove Theorem 5.3. Towards proving this, we will need the following three ingredients.

The first ingredient we need is a nondeterministic circuit that solves "Gap Probability Maximization" problem for the class of randomized nondeterministic circuits. Formally, we need:

Lemma 5.4 (GPM solver)**.** *Let $n, s \in \mathbb{N}$ and $0 < \gamma = \gamma(s) < 1$ be arbitrary. Let $\mathcal{C}$ be a class of randomized nondeterministic circuits of size s with n inputs such that each $C \in \mathcal{C}$ satisfies the following.*

- *There exists a unique $x^* \in \{0,1\}^n$ such that $\Pr[C(x^*) = 1] \geq \gamma$.*
- *For all $x \in \{0,1\}^n$ such that $x \neq x^*$, it holds that $\Pr[C(x) = 1] \leq \frac{\gamma}{2}$.*

Then, there exists a $\mathrm{poly}(s)$ size $\mathsf{NP}_{||}$-circuit $C_{\mathsf{GPM}} : \{0,1\}^{s\log(s)} \to \{0,1\}^n$ such that $C_{\mathsf{GPM}}(C) = x^*$.

We construct such a solver in Section 5.2. We will use the parameter-free version of this solver. This follows from the construction in Section 5.2 by trying all guesses $G \in \{1, (1+\eta), (1+\eta)^2, \ldots, 2^s\}$ for $\gamma 2^s$, where $\eta > 0$ is a sufficiently small universal constant. One guess lies in the constant-factor range needed for the hashing argument, and the final deterministic part selects a scale whose accepting input is uniquely determined. This only increases the circuit size by a polynomial factor.

The second ingredient that we need is a distinguishing test that can decide whether a low degree univariate polynomial equals the target multivariate polynomial composed with a line or not. Here is the precise key lemma:

Lemma 5.5 (Distinguishing correct polynomial)**.** *Assume we are in the setting of Assumption 5.2. Then, there exists $v^* \in \mathbb{F}^t$ such that for 0.99 fraction of $v \in \mathbb{F}^t$, each of the following holds.*[13]

1. *Let $h : \mathbb{F} \to \mathbb{F}$ be a degree d univariate polynomial such that $h = \widehat{f} \circ L_{v^*,v}$. Then,*

$$\Pr_{(w,y)\sim((\mathbf{W},\mathbf{Y})|\mathbf{W}_j \in L_{v^*,v}(\mathbb{F}))}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1, \ldots, h(L^{-1}_{v^*,v}(w_j)), \ldots, x_\ell; y\right) = z_{heavy}\right] \geq 2^{-m}(1+\varepsilon/3).$$

2. *Let $h : \mathbb{F} \to \mathbb{F}$ be a degree d univariate polynomial such that $h \neq \widehat{f} \circ L_{v^*,v}$. Then, either $h(0) \neq \widehat{f}(v^*)$ or the following holds:*

$$\Pr_{(w,y)\sim((\mathbf{W},\mathbf{Y})|\mathbf{W}_j \in L_{v^*,v})}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1, \ldots, h(L^{-1}_{v^*,v}(w_j)), \ldots, x_\ell; y\right) = z_{heavy}\right] \leq 2^{-m}(1+\varepsilon/6).$$

We will prove this property in Section 5.3. The third ingredient we need is a well known result regarding self correcting properties of low degree polynomials:

Theorem 5.6 ([GLRSW91])**.** *For all $d, t \in \mathbb{N}$ and a finite field $\mathbb{F}$ such that $|\mathbb{F}| \geq 2(d+1)$, there exists a randomized oracle machine M such that the following holds. For all degree d polynomials $p : \mathbb{F}^t \to \mathbb{F}$ and $C : \mathbb{F}^t \to \mathbb{F}$ such that $\Pr_{x\sim\mathbf{U}_{\mathbb{F}^t_q}}[p(x) = C(x)] \geq \frac{15}{16}$, M, with oracle access to C, computes p correctly on all inputs with probability at least $2/3$. Furthermore, the runtime of M is at most $\mathrm{poly}(d, t, \log(|\mathbb{F}|))$ and all the oracle queries made by M to C are non-adaptive.*

[13] Recall from Definition 4.31 that $L_{v^*,v} : \mathbb{F} \to \mathbb{F}^t$ is the unique line where $L_{v^*,v}(0) = v^*$ and $L_{v^*,v}(1) = v$.

Using these, we obtain our desired reconstructive procedure as follows:

Proof of Theorem 5.3. Since we are in the setting of Assumption 5.2, Lemma 5.5 holds. Let $C_{\mathsf{Samp}}, P_{\mathsf{Samp}}$ be deterministic and nondeterministic circuits respectively such that $(\mathbf{W}, \mathbf{Y})$ is obtained by sampling from C_{Samp} with postselection from P_{Samp}.

- We first define the following $\mathsf{NP}_{||}$-circuit $C_{0.99} : \mathbb{F}_q^t \to \mathbb{F}_q$ that computes the answer correctly for 0.99 fraction of the inputs $v \in \mathbb{F}_q^t$:

  Circuit $C_{0.99}$.

  - For each $v \in \mathbb{F}_q^t$, define randomized nondeterministic circuit D_v that takes as input a degree d univariate polynomial, and has range $\{0,1\}$ as follows:

    Circuit D_v.

    * Non uniformly hardwire v^*, $\widehat{f}(v^*)$, and $j \in [\ell]$ (the index of the good block) in the description of D_v.
    * Let the input degree d polynomial be $h : \mathbb{F} \to \mathbb{F}$.
    * If $h(0) \neq \widehat{f}(v^*)$, then output 0.
    * Use randomness and nondeterminism to sample $(w^{(1)}, y^{(1)}), \ldots, (w^{(r)}, y^{(r)}) \sim (\mathbf{W}, \mathbf{Y})$ from $C_{\mathsf{Samp}}, P_{\mathsf{Samp}}$ where $r = \frac{36}{\varepsilon}$. When sampling, if the postselecting circuit P_{Samp} rejects any sample, then stop and output 0.
    * Output 1 if for all $i \in [r]$, it holds that $w_j^{(i)} \in L_{v^*,v}(\mathbb{F})$ and, using nondeterminism, if there exist $(x_1^{(i)}, \ldots, x_\ell^{(i)}) \in \mathbb{F}^\ell$ such that the following condition is met.

    $$\mathsf{2Ext}\left(x_1^{(i)}, \ldots, h(L_{v^*,v}^{-1}(w_j^{(i)})), \ldots, x_\ell^{(i)}; y^{(i)}\right) = z_{\text{heavy}}$$

  - On input $v \in \mathbb{F}_q^t$, proceed as follows.
  - Using the parameter-free version of Lemma 5.4, let C_{GPM} be the $\mathsf{NP}_{||}$-circuit that takes as input circuits of size the same as that of D_v and solves the gap probability maximization problem. For analysis, the relevant gap parameter is $\gamma = \beta \cdot \left(2^{-m} + \frac{\varepsilon}{3} \cdot 2^{-m}\right)^r$ where β is the probability that the postselecting circuit P_{Samp} accepts all r samples and $w_j^{(i)} \in L_{v^*,v}$ for all $i \in [r]$.
  - Let $h' : \mathbb{F} \to \mathbb{F}$ be the degree d polynomial that is the output of running C_{GPM} on input D_v.
  - Output $h'(1)$.

- We apply Theorem 5.6 with oracle circuit as $C_{0.99}$ to obtain a randomized $\mathsf{NP}_{||}$-circuit C_1 [a] that computes the correct answer with probability at least $2/3$.

- We amplify the success probability of C_1 by repeating the process $O(t \log(q))$ times and

taking the majority vote to increase the success probability of C_1 to be at least $1 - 2 \cdot q^t$; then fix a choice of randomness to obtain our final $\mathsf{NP}_{||}$-circuit C_{final}, claiming to compute $\widehat{f}$. [b]

[a] Since the self corrector makes non-adaptive queries to $C_{0.99}$, which itself makes non-adaptive queries to NP-gates, the composed circuit makes non-adaptive queries to NP-gates.

[b] We crucially use the fact that in the amplification process, the non-adaptive nature of queries is preserved

Size of C_{final} We here upper bound size of our constructed circuit C_{final}. We begin by upper bounding the size of D_v (for any v). Since the circuit samples r samples from C_{Samp} and also runs the two source extractor that many times, the final circuit size is at most $\text{poly}(1/\varepsilon, s_{\mathsf{Samp}}, s_{\mathsf{2Ext}}, \ell, t, \log(q), n, d)$. Since size of C_{GPM} is polynomial in the input size the same bound also applies to size of $C_{0.99}$. Lastly, since the procedure from Theorem 5.6 runs in polynomial time, C_{final} also has the same size bound.

Correctness We claim that $h' = \widehat{f} \circ (L_{v_*,v})$ above. If this were true then $h'(1) = \widehat{f}(v)$, and C_{final} indeed computes $\widehat{f}$ as claimed. By the self correcting property guaranteed by Theorem 5.6, if $C_{0.99}$ indeed computes $\widehat{f}$ for 0.99 fraction of v, then C_{final} computes $\widehat{f}$ everywhere.

By construction of circuit D_v and the guarantees of Lemma 5.5, for 0.99 fraction of v, the following holds: If $h = \widehat{f} \circ L_{v^*,v}$, then D_v accepts h with probability at least $\beta\,(2^{-m}(1+\varepsilon/3))^r$. And, if $h \neq \widehat{f} \circ L_{v^*,v}$, then either D_v rejects with probability 1 (if $h(0) \neq \widehat{f}(v^*)$) or accepts with probability at most $\beta\,(2^{-m}(1+\varepsilon/6))^r$.

The only thing that needs to be checked then is if the promise of the gap probability maximization problem is satisfied with the claimed parameter γ. In particular, to show that we meet the guarantees for Lemma 5.4, it suffices to show that

$$\beta\,(1+\varepsilon/3)^r \geq 2 \cdot \beta\,(1+\varepsilon/6)^r$$

We compute that

$$\begin{aligned}
\left(\frac{1+\varepsilon/3}{1+\varepsilon/6}\right)^r &= \left(1 + \frac{\varepsilon/6}{1+\varepsilon/6}\right)^r \\
&\geq \left(1 + \frac{\varepsilon}{12}\right)^r \\
&\geq \exp((\varepsilon/36) \cdot r) \\
&\geq e \geq 2
\end{aligned}$$

where the third last inequality follows using Claim 4.15 and the second last inequality follows by our setting of r. Hence, C_{final} indeed computes $\widehat{f}$ as desired. □

5.2 Small circuits for the Gap Probability Maximization problem

In this subsection we will construct small $\mathsf{NP}_{||}$-circuits to solve the gap probability maximization problem.

Proof of Lemma 5.4. Let $C \in \mathcal{C}$ be an arbitrary circuit. We will use notation $C(x; y, z)$ to mean C is fed with input $x \in \{0,1\}^n$, randomness $y \in \{0,1\}^s$ and witness $z \in \{0,1\}^s$. Since s is size

of the circuit C, without loss of generality we assume that it is fed with length s randomness and witnesses.

We construct a randomized $\mathsf{NP}_{||}$-circuit that for each circuit in $\mathcal{C}$, outputs the corresponding x^* with probability at least $1 - 2^{-2s\log(s)}$. Since the number of circuits of size s is at most $2^{s\log(s)}$, we can non-uniformly fix the randomness and obtain the desired $\mathsf{NP}_{||}$ circuit.

Let $k = 10^6 \cdot s\log(s), \tau = k/2, \ell = 48, M = \frac{\gamma}{64} \cdot 2^s$. Let $\mathcal{H} = \{\{0,1\}^s \to [M]\}$ be family of explicit pairwise independent hash functions (for formal statement and properties see Lemma 4.30). [14]

Using this, we define our circuit as follows.

- Sample $h_1, \ldots, h_k \sim \mathbf{U}_{\mathcal{H}}$ (this is the only place where our circuit uses randomness).[a]
- Construct nondeterministic circuit $D : \{0,1\}^{n+(s\cdot k\cdot \ell)} \to \{0,1\}$, that takes witnesses from $\{0,1\}^{s\cdot k\cdot \ell}$ as follows:

  Let input be $x \in \{0,1\}^n$ and $y = \{y_{i,j}\}_{(i,j)\in[k]\times[\ell]}$ where each $y_{i,j} \in \{0,1\}^s$. Also let $z = \{z_{i,j}\}_{(i,j)\in[k]\times[\ell]}$ be the witness input where each $z_{i,j} \in \{0,1\}^s$. With this, the circuit operates as follows:

  - If there exists $i \in [k]$ and $j \neq j' \in [\ell]$ such that $y_{i,j} = y_{i,j'}$, then output 0.
  - If there exists $i \in [k]$ and $j \in [\ell]$ such that $h_i(y_{i,j}) \neq 0$, then output 0.
  - For $i \in [k]$, let $\mathrm{test}_i^x = 1$ if for all $j \in [\ell]$, $C(x, y_{i,j}, z_{i,j}) = 1$.
  - Output 1 if $\sum_{i\in[k]} \mathrm{test}_i^x \geq \tau$.

  [a] We can specify the value of γ in s bits and hardwire that into the circuit so that it can correctly instantiate such a hash function family $\mathcal{H}$.

- For $i \in [n]$ let $D_i : \{0,1\}^{n+s\cdot k\cdot \ell} \to \{0,1\}$ be the nondeterministic circuit defined as follows: On input x, y, output 1 if $x_i = 1$ and $D(x, y) = 1$.
- For all $i \in [n]$, make parallel NP oracle calls to whether D_i is satisfiable and let $\widehat{x}_i$ be the outcome of each of these queries.
- Output $(\widehat{x}_1, \ldots, \widehat{x}_n)$.

Validity of the procedure We observe that the problem of checking whether a nondeterministic circuit is satisfiable is indeed in NP. In our procedure, we make NP oracle calls non-adaptively and hence, it's a valid $\mathsf{NP}_{||}$-circuit.

Bounding size of the circuit Using the fact that $k, \ell, \log(M)$ are $\mathrm{poly}(s)$, and that $\mathcal{H}$ is an explicit hash function family, we easily verify our circuit has size at most $\mathrm{poly}(s)$ as desired.

Correctness As mentioned earlier, we here show that with probability at least $1 - 2^{-2s\log(s)}$, our circuit constructed above indeed succeeds and outputs the desired x^*. Note that all randomness

[14] Since the acceptance probability is over s random bits, the relevant gap parameter satisfies $\gamma \geq 2^{-s}$ and so $M \geq \frac{1}{64}$. By first adding $O(1)$ dummy random bits that the circuit ignores, we increase $\gamma 2^s$ by a constant factor without changing any acceptance probabilities. Thus we may assume $M = \gamma 2^s/64 \geq 1$.

here is over the choice of $h_1, \ldots, h_k$ from $\mathcal{H}$.

For $y \in \{0,1\}^{s \cdot k \cdot \ell}$, we say y is "valid" if it passes the first two tests in the circuit D. For the rest of the correctness proof, we always assume that input y is valid, as this does not change our claim.

Let E_1 be the event that there does not exist $y \in \{0,1\}^{s \cdot k \cdot \ell}$ such that D accepts (x^*, y). Let E_2 be the event that there exists some $x \neq x^* \in \{0,1\}^n$, for which there exists some $y \in \{0,1\}^{s \cdot k \cdot \ell}$ such that D accepts (x, y). We will show that the probability that any of E_1 or E_2 occurring is at most $2^{-2s\log(s)}$. We observe that when none of E_1 or E_2 occur then, our circuit correctly outputs x^* and hence our desired claim follows.

For $i \in [k]$ and $x \in \{0,1\}^n$, let $\mathbf{T}_{x,i}$ be indicator random variables for whether there exists $y_i \in \{0,1\}^{s \cdot \ell}$ such that $\mathrm{test}_i^x = 1$, i.e., there exists a witness $z_i \in \{0,1\}^{s \cdot \ell}$ such that $C(x, y_{i,j}, z_{i,j}) = 1$ for all $j \in [\ell]$.

We claim the following:

Claim 5.7. *For all $i \in [k]$, the following holds:*

1. $\mathbb{E}\left[\mathbf{T}_{x^*,i}\right] \geq \frac{3}{4}$.
2. *For all $x \neq x^* \in \{0,1\}^n$, it holds that* $\mathbb{E}\left[\mathbf{T}_{x,i}\right] \leq \frac{1}{8}$.

We see that E_1 occurs when $\sum_{i \in [k]} \mathbf{T}_{x^*,i} < \tau = \frac{k}{2}$. Since $\mathbf{T}_{x^*,1}, \ldots, \mathbf{T}_{x^*,k}$ are independent random variables and their sum has expectation $\frac{3}{4} \cdot k$, we apply Chernoff bound (Lemma 4.17) to obtain that

$$\Pr\left[\sum_{i \in [k]} \mathbf{T}_{x^*,i} < \tau\right] \leq \exp(-k/24).$$

Similarly, E_2 occurs when there exists $x \neq x^* \in \{0,1\}^n$ such that $\sum_{i \in [k]} \mathbf{T}_{x,i} \geq \tau = \frac{k}{2}$. We will union bound over each of the $\leq 2^n$ many such x. For each such $x \neq x^*$, $\mathbf{T}_{x,1}, \ldots, \mathbf{T}_{x,k}$ are independent random variables and their sum has expectation $\frac{1}{8} \cdot k$. Therefore, we apply Chernoff bound (Lemma 4.17) to obtain that

$$\Pr\left[\sum_{i \in [k]} \mathbf{T}_{x,i} \geq \tau\right] \leq \exp(-9k/40).$$

Together, we obtain that E_1 or E_2 occurs with probability at most $\exp(-k/24) + 2^n \cdot \exp(-k/24)$. As $n \leq s$ and $k = 10^6 \cdot s\log(s)$, we infer that probability that either of them occurs is indeed at most $2^{-2s\log(s)}$ as desired.

We finally prove our remaining claim regarding expectations of $\mathbf{T}_{x,i}$.

Proof of Claim 5.7. Before we prove either of the parts, we will use the following simple claim:

Claim 5.8. *Let $S \subset \{0,1\}^s$ be arbitrary and let $h \sim \mathbf{U}_{\mathcal{H}}$. Let $\alpha = |S \cap \{y : h(y) = 0\}|$. Then, $\mathbb{E}[\alpha] = \frac{|S|}{M}$ and $\mathrm{Var}(\alpha) \leq \frac{|S|}{M}$.*

Proof. For $x \in S$, let $\mathbf{X}$ be the indicator of whether for $h \sim \mathcal{H}$, it holds that $h(x) = 0$. By definition of pairwise independent hashing, we have that $\mathbb{E}[\mathbf{X}] = \frac{1}{M}$ and that $\mathrm{Var}(\mathbf{X}) = \frac{1}{M}\left(1 - \frac{1}{M}\right) \leq \frac{1}{M}$. Using linearity of expectation, we obtain that $\mathbb{E}[\alpha] = \frac{|S|}{M}$. Similarly, using the fact that each of these indicators is pairwise independent, we obtain that $\mathrm{Var}(\alpha) \leq \frac{|S|}{M}$ as desired. □

Using this, we prove each part of our claims one after another. Before that, we define a couple of useful quantities. For $x \in \{0,1\}^n$, define $R(x) = \{r \in \{0,1\}^s : \exists v \in \{0,1\}^s \text{ s.t. } C(x;r;v) = 1\}$. For $h \in \mathcal{H}$ and $x \in \{0,1\}^n$, let $R_h(x) = R(x) \cap \{r : h(r) = 0\}$. By assumption regarding validity, each of $y_{i,1}, \ldots, y_{i,\ell}$ is distinct inside the circuit D. Therefore, for all $x \in \{0,1\}^n$, it holds that $\mathbb{E}[\mathbf{T}_{x,i}] = \Pr_{h\sim\mathcal{H}}[|R_h^x| \geq \ell]$. With this, we prove each of the parts of the proof.

1. We compute that

$$\begin{aligned}\mathbb{E}[\mathbf{T}_{x^*,i}] &= \Pr_{h\sim\mathcal{H}}[|R_h(x^*)| \geq \ell] \\ &= 1 - \Pr_{h\sim\mathcal{H}}[|R_h(x^*)| < \ell] \\ &\geq 1 - \Pr_{h\sim\mathcal{H}}\left[\left|\frac{|R(x^*)|}{M} - |R_h(x^*)|\right| > \frac{|R(x^*)|}{M} - \ell\right]\end{aligned}$$

   Using bounds on expectation and variance from Claim 5.8, as well as the fact that $|R(x^*)| \geq \gamma \cdot 2^s$, we apply Chebyshev's inequality (Lemma 4.19) to infer that

$$\mathbb{E}[\mathbf{T}_{x^*,i}] \geq 1 - \frac{\gamma \cdot 2^s}{M} \cdot \frac{1}{\left(\frac{\gamma \cdot 2^s}{M} - \ell\right)^2} = 1 - \frac{64}{(64-48)^2} = \frac{3}{4}$$

   as desired.

2. For $x \neq x^*$, we compute that

$$\begin{aligned}\mathbb{E}[\mathbf{T}_{x,i}] &= \Pr_{h\sim\mathcal{H}}[|R_h^x| \geq \ell] \\ &\leq \Pr_{h\sim\mathcal{H}}\left[\left||R_h(x)| - \frac{|R(x)|}{M}\right| \geq \ell - \frac{|R(x)|}{M}\right]\end{aligned}$$

   Using bounds on expectation and variance from Claim 5.8, as well as the fact that $|R(x)| \leq \gamma \cdot 2^{s-1}$, we apply Chebyshev's inequality (Lemma 4.19) to infer that

$$\mathbb{E}[\mathbf{T}_{x,i}] \leq \frac{\gamma \cdot 2^{s-1}}{M} \cdot \frac{1}{\left(\ell - \frac{\gamma \cdot 2^s}{M}\right)^2} = \frac{32}{(48-32)^2} = \frac{1}{8}$$

   as desired.

□

□

5.3 Distinguishing the correct polynomial from failed extraction

In this subsection, we will prove Lemma 5.5—our key lemma regarding distinguishing between appropriate univariate polynomials.

Towards proving this, we will rely on the following helper lemma that shows various key properties in this setting:

Lemma 5.9. *Assume that we are in the setting of Assumption 5.2. Let* $\mathsf{Heavy}_j = \{w_j \in \mathbb{F}_q^t : \Pr[\mathbf{W}_j = w_j] \geq 2^{-t\log(q)+\Delta+1}\}$. *Then,*

1. *Let* $L : \mathbb{F}_q \to \mathbb{F}_q^t$ *be a random non-trivial line.*

   Then, with probability $1 - \frac{2^{2(\Delta+1)}}{q} - \frac{9\cdot 2^{2(m+\Delta)}}{\varepsilon^2 q} - 4\sqrt{\varepsilon_{in}}$ *over* L*, the following holds:*

   (a)
   $$\frac{1}{2}\cdot\frac{q}{q^t} \leq \Pr[\mathbf{W}_j \in L(\mathbb{F}) | (\mathbf{W}_j \notin \mathsf{Heavy}_j)] \leq \frac{3}{2}\cdot\frac{q}{q^t}.$$

   (b)
   $$\Pr[\mathbf{W}_j \in \mathsf{Heavy}_j | (\mathbf{W}_j \in L(\mathbb{F}))] \leq \sqrt{\varepsilon_{in}}.$$

   (c)
   $$\Pr_{(w_j,y)\sim((\mathbf{W}_j,\mathbf{Y})|\mathbf{W}_j\in L(\mathbb{F}))}\Big[\exists x_1,\ldots,x_\ell \in \mathbb{F} : \mathsf{2Ext}\big(x_1,\ldots,h(L^{-1}(w_j)),\ldots,x_\ell;y\big) = z_{heavy}\Big] \;\geq\; 2^{-m}\left(1+\varepsilon/3\right).$$

2. *For a line* $L : \mathbb{F} \to \mathbb{F}^t$ *and a degree* d *univariate polynomial* $h : \mathbb{F} \to \mathbb{F}$*, we say* h *is 'biasing' if*

   $$\Pr_{(w_j,y)\sim((\mathbf{W}_j,\mathbf{Y})|\mathbf{W}_j\in L(\mathbb{F}))}\Big[\exists x_1,\ldots,x_\ell \in \mathbb{F} : \mathsf{2Ext}\big(x_1,\ldots,h(L^{-1}(w_j)),\ldots,x_\ell;y\big) = z_{heavy}\Big] \;\geq\; 2^{-m}\left(1+\varepsilon/6\right).$$

   For any line L *such that* $\Pr[\mathbf{W}_j \in L(\mathbb{F}) | (\mathbf{W}_j \notin \mathsf{Heavy}_j)] \geq \frac{1}{2}\cdot\frac{q}{q^t}$ *and* $\Pr[\mathbf{W}_j \in \mathsf{Heavy}_j | (\mathbf{W}_j \in L(\mathbb{F}))] \leq \sqrt{\varepsilon_{in}}$*, the number of biasing polynomials* h *is at most* $\frac{24\cdot 2^{2k_1+m}}{\varepsilon}$.

We will prove each of the parts of the helper lemma in Section 5.4, and Section 5.5 respectively. With this, let's see how we obtain the desired property regarding distinguishing the correct polynomial:

Proof of Lemma 5.5. The result for when $h = \widehat{f}\circ L_{v^*,v}$ can be restated as:

$$\Pr_{(w,y)\sim((\mathbf{W},\mathbf{Y})|\mathbf{W}_j\in L_{v^*,v})}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1,\ldots,\widehat{f}(w_j),\ldots,x_\ell;y\right) = z_{\text{heavy}}\right] \geq 2^{-m}(1+\varepsilon/3).$$

Let $\mathbf{v},\mathbf{v}^*$ be uniformly randomly chosen points from $\mathbb{F}^t$ and let $L_{\mathbf{v}^*,\mathbf{v}}$ be the line through them. Then, with probability $1-q^{-t}$, $L_{\mathbf{v},\mathbf{v}^*}$ will be a non-trivial line and by Lemma 5.9, for $1 - q^{-t} - \frac{2^{2(\Delta+1)}}{q} - \frac{9\cdot 2^{2(m+\Delta)}}{\varepsilon^2 q} - 4\sqrt{\varepsilon_{\text{in}}}$ fraction of lines L, we have that

$$\Pr_{(w,y)\sim((\mathbf{W},\mathbf{Y})|\mathbf{W}_j\in L(\mathbb{F}))}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1,\ldots,\widehat{f}(w_j),\ldots,x_\ell;y\right) = z_{\text{heavy}}\right] \geq 2^{-m}(1+\varepsilon/3),$$

proving the first part of our desired result. Also for all such lines L, the condition to apply second item of Lemma 5.9 is met. Therefore, the number of degree d polynomials h such that

$$\Pr_{(w_j,y)\sim((\mathbf{W}_j,\mathbf{Y})|\mathbf{W}_j\in L(\mathbb{F}))}[\exists x_1,\ldots,x_\ell \in \mathbb{F} : \mathsf{2Ext}(x_1,\ldots,h(L^{-1}(w_j)),\ldots,x_\ell;y) = z_{\text{heavy}}] \geq 2^{-m}(1+\varepsilon/6)$$

is at most $\frac{24\cdot 2^{2k_1+m}}{\varepsilon}$. We use terminology from Lemma 5.9 and call such polynomial h as a "biasing" polynomial and let $\mathrm{bad}(L)$ be the set of biasing polynomials for a line L.

Let E be the event that for $L \sim L_{\mathbf{v}^*,\mathbf{v}}$, $|\mathrm{bad}(L)| \leq \frac{24\cdot 2^{2k_1+m}}{\varepsilon}$. We claim the following.

Claim 5.10.

$$\Pr_{L\sim(L_{\mathbf{v}^*,\mathbf{v}}|E)}\left[\exists h \in bad(L) : h \neq \widehat{f}\circ L \wedge h(0) = \widehat{f}(\mathbf{v}^*)\right] \leq \frac{24\cdot 2^{2k_1+m}}{\varepsilon}\cdot\frac{d+1}{q}.$$

We will prove this claim towards the end of this subsection. Let's see how we finish our proof from it. From above discussion and Claim 5.10, we see that for $1 - q^{-t} - \frac{2^{2(\Delta+1)}}{q} - \frac{9\cdot 2^{2(m+\Delta)}}{\varepsilon^2 q} - 4\sqrt{\varepsilon_{\text{in}}} - \frac{48\cdot 2^{2k_1+m}}{\varepsilon}\cdot\frac{d}{q}$ fraction of lines $L_{\mathbf{v}^*,\mathbf{v}}$, both conditions of the results are satisfied. Therefore by an averaging argument, there exists a choice of $v^* \in \mathbb{F}^t$ such that for $1 - q^{-t} - \frac{2^{2(\Delta+1)}}{q} - \frac{9\cdot 2^{2(m+\Delta)}}{\varepsilon^2 q} - 4\sqrt{\varepsilon_{\text{in}}} - \frac{48\cdot 2^{2k_1+m}}{\varepsilon}\cdot\frac{d}{q}$ fraction of points $v \in \mathbb{F}^t$, the line $L_{v^*,v}$ satisfies both the desired properties. By our assumption, $1 - q^{-t} - \frac{2^{2(\Delta+1)}}{q} - \frac{9\cdot 2^{2(m+\Delta)}}{\varepsilon^2 q} - 4\sqrt{\varepsilon_{\text{in}}} - \frac{48\cdot 2^{2k_1+m}}{\varepsilon}\cdot\frac{d}{q} \geq 0.99$ and hence, our desired result holds.

We now prove our remaining claim regarding lines with a small number of biasing polynomials.

Proof of Claim 5.10. As h and $\widehat{f}\circ L$ are univariate degree d polynomials, they can agree on at most $d+1$ points on the line L. So, instead of comparing whether $h(0) = \widehat{f}\circ L(0)$, if we had picked a random element $\alpha \in \mathbb{F}_q$ and tested if $h(\alpha) = \widehat{f}\circ L(\alpha)$, then by a union bound, the claim would indeed hold. We argue that even when fixing the choice to 0, the claim holds. We do this by rethinking our sampling process. Instead of our process, where we pick $\mathbf{v}^*, \mathbf{v} \in \mathbb{F}^t$ and then consider the line L parameterized by them (subject to the event E occurring), we can instead consider the process where we first pick a random line L (subject to it satisfying the condition on $|\mathrm{bad}(L)|$) and then let $\mathbf{v}^*, \mathbf{v}$ be two random points on L; in this latter test, the test $h(0) = \hat{f}(v^*)$ is equivalent to testing $h(\alpha) = \hat{f}(L(\alpha))$ at a uniformly random point $\alpha \in \mathbb{F}_q$. Hence, for the latter process, the bound indeed holds. It is easy to see that both processes are equivalent and for a fixed line, the latter process just randomly reparameterizes it. This implies the bound also holds for the first process and our claim follows. □

□

5.4 Proving the first part of the helper lemma

We here prove the first part of Lemma 5.9 that establishes various helpful properties regarding lines, and low degree polynomials in our setting.

Proof of first part of Lemma 5.9. We prove each of the properties one by one. For $v \in \mathbb{F}_t$, let $p_v = \Pr[\mathbf{W}_j = v | (\mathbf{W}_j \notin \mathsf{Heavy}_j)]$. We will use this notation when proving both properties.

(a) We will show that this property holds for $1 - \frac{2^{2(\Delta+1)}}{q}$ fraction of lines L.

We see that $\Pr[\mathbf{W}_j \in L(\mathbb{F}) | (\mathbf{W}_j \notin \mathsf{Heavy}_j)] = \sum_{x\in\mathbb{F}} p_{L(x)}$. Observe that for a random line $\mathbf{L}$, $\{\mathbf{L}(x)\}_{x\in\mathbb{F}}$ are pairwise independent random variables, each of which is uniformly distributed over $\mathbb{F}_q^t$. Therefore, $\{p_{\mathbf{L}(x)}\}_{x\in\mathbb{F}}$ are pairwise independent random variables. For all $x \in \mathbb{F}$, we see that $\mathbb{E}[p_{\mathbf{L}(x)}] = \mathbb{E}_{u\sim\mathbf{U}_{\mathbb{F}^t}}[p_u] = q^{-t}$. Also, for all $v \in \mathbb{F}_t$, using the fact that $\mathbf{W}_j$ has

min-entropy gap of Δ and that the conditional entropy has entropy gap of $\Delta+1$, we have that $0 \le p_v \le q^{-t} \cdot 2^{\Delta+1}$.

Using the above properties, we apply a tail bound on sums of pairwise independent random variables, with bounded output range - Lemma 4.20 to $\{p_{\mathbf{L}(x)}\}_{x\in\mathbb{F}}$ to infer that

$$\Pr\left[\left|\left(\sum_{x\in\mathbb{F}} p_{\mathbf{L}(x)}\right) - q\cdot q^{-t}\right| \ge \frac{1}{2}\cdot q\cdot q^{-t}\right] \le \frac{q\cdot q^{-2t}\cdot 2^{2(\Delta+1)}}{4\cdot\frac{1}{4}\cdot q^2\cdot q^{-2t}} = \frac{2^{2(\Delta+1)}}{q}.$$

Since for every line L, $\Pr[\mathbf{W}_j \in L(\mathbb{F})|(\mathbf{W}_j \notin \mathsf{Heavy}_j)] = \sum_{x\in\mathbb{F}} p_{L(x)}$, the claim follows.

(b) We will show that this property, along with the above property, holds for $1 - \frac{2^{2\Delta}}{q} - 4\sqrt{\varepsilon_{\text{in}}}$ fraction of lines L. We first bound the quantity $\Pr[\mathbf{W}_j \in (\mathsf{Heavy}_j \cap L(\mathbb{F}))]$ for most lines. Formally

Claim 5.11. *For $1-4\sqrt{\varepsilon_{in}}$ fraction of lines L, it holds that*

$$\Pr[\mathbf{W}_j \in (\mathsf{Heavy}_j \cap L(\mathbb{F}))] \le \frac{1}{2}\cdot\sqrt{\varepsilon_{in}}\cdot q^{-t+1}.$$

We will prove this later. First, using this let us prove the desired claim. Using the lower bound from the previous part (Item (a)) and the above claim, we have that for at least $1-4\sqrt{\varepsilon_{\text{in}}} - \frac{2^{2(\Delta+1)}}{q}$ fraction of lines L,

$$\begin{aligned}\Pr[\mathbf{W}_j \in \mathsf{Heavy}_j | (\mathbf{W}_j \in L(\mathbb{F}))] &= \frac{\Pr[\mathbf{W}_j \in (\mathsf{Heavy}_j \cap L(\mathbb{F}))]}{\Pr[\mathbf{W}_j \in L(\mathbb{F})]} \\ &\le \frac{\frac{1}{2}\cdot\sqrt{\varepsilon_{\text{in}}}\cdot q^{-t+1}}{\frac{1}{2}\cdot q^{-t+1}} \\ &\le \sqrt{\varepsilon_{\text{in}}}\end{aligned}$$

as desired. We now prove our remaining claim regarding upper bounding the combined event.

Proof of Claim 5.11. For $v \in \mathbb{F}_t$, let $q_v = \begin{cases} 0 & v \notin \mathsf{Heavy}_j \\ \Pr[\mathbf{W}_j = v] & v \in \mathsf{Heavy}_j \end{cases}$. For any line L, we see that $\Pr[\mathbf{W}_j \in (\mathsf{Heavy}_j \cap L(\mathbb{F}))] = \sum_{x\in\mathbb{F}} q_{L(x)}$. We see that for a random line $\mathbf{L}$ and all $x \in \mathbb{F}$, we have that $\mathbb{E}[q_{\mathbf{L}(x)}] = \mathbb{E}_{u\sim\mathbf{U}_{\mathbb{F}^t}}[q_u] = \Pr[\mathbf{W}_j \in \mathsf{Heavy}_j]\cdot q^{-t} \le 2\varepsilon_{\text{in}}\cdot q^{-t}$ where the last inequality follows from Claim 4.14. Therefore, $\mathbb{E}\left[\sum_{x\in\mathbb{F}} q_{\mathbf{L}(x)}\right] = 2\varepsilon_{\text{in}}\cdot q^{-t+1}$. Applying Markov's inequality (Lemma 4.16) to the sum, the claim follows. □

(c) We will show that among lines satisfying Item (a) and Item (b), at least $1 - \frac{9\cdot 2^{2(m+\Delta)}}{\varepsilon^2\cdot q}$ fraction of lines L satisfy the following:

$$\Pr_{(w;y)\sim((\mathbf{W},\mathbf{Y})|\mathbf{W}_j\in L(\mathbb{F}))}[2\mathsf{Ext}(\widehat{f}(w_1),\ldots,\widehat{f}(w_\ell);y) = z_{\text{heavy}}] \ge 2^{-m}\left(1+\varepsilon/3\right).$$

This suffices to prove the desired claim by setting $x_i = \widehat{f}(w_i)$.

We first compute that for any line L satisfying the previous two properties:

$$\begin{aligned}
&\Pr_{(w;y)\sim((\mathbf{W},\mathbf{Y})|\mathbf{W}_j\in L(\mathbb{F}))}[\mathsf{2Ext}(\widehat{f}(w_1),\ldots,\widehat{f}(w_\ell);y)=z_{\mathrm{heavy}}]\\
&\geq \Pr[\mathbf{W}_j\notin \mathsf{Heavy}_j|\mathbf{W}_j\in L(\mathbb{F})]\cdot \Pr_{(w;y)\sim((\mathbf{W},\mathbf{Y})|(\mathbf{W}_j\in L(\mathbb{F}),\mathbf{W}_j\notin\mathsf{Heavy}_j))}[\mathsf{2Ext}(\widehat{f}(w_1),\ldots,\widehat{f}(w_\ell);y)=z_{\mathrm{heavy}}]\\
&\geq \Pr_{(w;y)\sim((\mathbf{W},\mathbf{Y})|(\mathbf{W}_j\in L(\mathbb{F}),\mathbf{W}_j\notin\mathsf{Heavy}_j))}[\mathsf{2Ext}(\widehat{f}(w_1),\ldots,\widehat{f}(w_\ell);y)=z_{\mathrm{heavy}}]-\sqrt{\varepsilon_{\mathrm{in}}}.
\end{aligned}$$

Hence, it suffices to show:

$$\Pr_{(w;y)\sim((\mathbf{W},\mathbf{Y})|(\mathbf{W}_j\in L(\mathbb{F})\setminus\mathsf{Heavy}_j))}[\mathsf{2Ext}(\widehat{f}(w_1),\ldots,\widehat{f}(w_\ell);y)=z_{\mathrm{heavy}}]\geq 2^{-m}\left(1+\varepsilon/3\right)+\sqrt{\varepsilon_{\mathrm{in}}}.$$

For $v\in\mathbb{F}_t$, let

$$b_v=\Pr_{(w;y)\sim((\mathbf{W},\mathbf{Y})|\mathbf{W}_j=v)}\left[\mathsf{2Ext}\left(\widehat{f}(w_1),\ldots,\widehat{f}(w_\ell);y\right)=z_{\mathrm{heavy}}\right].$$

Let

$$b'_v=b_v-2^{-m}.$$

We compute that:

$$\begin{aligned}
&\Pr_{(w_1,\ldots,w_\ell;y)\sim((\mathbf{W},\mathbf{Y})|\mathbf{W}_j\in L(\mathbb{F})\setminus\mathsf{Heavy}_j)}[\mathsf{2Ext}(\widehat{f}(w_1),\ldots,\widehat{f}(w_\ell);y)=z_{\mathrm{heavy}}]\\
&=\sum_{v\in L(\mathbb{F})\setminus\mathsf{Heavy}_j} b_v\cdot\frac{\Pr[\mathbf{W}_j|(\mathbf{W}_j\notin\mathsf{Heavy}_j)=v]}{\Pr[\mathbf{W}_j|(\mathbf{W}_j\notin\mathsf{Heavy}_j)\in L(\mathbb{F})]}\\
&=\frac{1}{\Pr[\mathbf{W}_j\in L(\mathbb{F})\setminus\mathsf{Heavy}_j]}\sum_{v\in L(\mathbb{F})}(b'_v+2^{-m})\cdot p_v\\
&=2^{-m}+\frac{1}{\Pr[\mathbf{W}_j\in L(\mathbb{F})\setminus\mathsf{Heavy}_j]}\sum_{v\in L(\mathbb{F})}b'_v\cdot p_v\\
&\geq 2^{-m}+\frac{2}{3}\cdot q^{t-1}\cdot\sum_{v\in L(\mathbb{F})}b'_v\cdot p_v
\end{aligned}$$

where in the last line we used the fact that $\Pr[\mathbf{W}_j\in L(\mathbb{F})|(\mathbf{W}_j\notin\mathsf{Heavy}_j)]\leq\frac{3}{2}\cdot\frac{q}{q^t}$. By rearranging, we see that to prove the desired claim, it suffices to show that for $1-\frac{9\cdot 2^{2m+2\Delta}}{\varepsilon^2\cdot q}$ fraction of random non-trivial line $\mathbf{L}$, it holds that

$$\sum_{v\in\mathbf{L}(\mathbb{F})}b'_v\cdot p_v\geq\left(\frac{\varepsilon}{2}\cdot 2^{-m}+\frac{3}{2}\cdot\sqrt{\varepsilon_{\mathrm{in}}}\right)\cdot\frac{q}{q^t}.$$

Using our assumptions, regarding $\varepsilon_{\mathrm{in}},\varepsilon_{\mathsf{2Ext}}$, and ε, we have that $\varepsilon_{\mathrm{in}}\leq\frac{\varepsilon^2\cdot 2^{-2m}}{324}$. Using this, we infer that it suffices for us to show:

$$\sum_{v\in\mathbf{L}(\mathbb{F})}b'_v\cdot p_v\geq\frac{7\varepsilon}{12}\cdot 2^{-m}\cdot\frac{q}{q^t}.$$

As points on a line are pairwise independent, $\{b'_v p_v\}_{v \in \mathbf{L}(\mathbb{F})}$ are pairwise independent random variables. We compute that

$$\begin{aligned}
\mathbb{E}\left[\sum_{v \in \mathbf{L}(\mathbb{F})} b'_v p_v\right] &= q \cdot \mathop{\mathbb{E}}_{u \sim \mathbf{U}_{\mathbb{F}_t}}[b'_v p_v] \\
&= \frac{q}{q^t} \cdot \left(\sum_{v \in \mathbb{F}_t} b_v p_v\right) - 2^{-m} \cdot \frac{q}{q^t} \\
&= \frac{q}{q^t} \cdot \left(\Pr_{(w;y) \sim (\mathbf{W},\mathbf{Y}) | (\mathbf{W}_j \notin \mathsf{Heavy}_j)} \left[\mathsf{2Ext}\left(\widehat{f}(w_1), \ldots, \widehat{f}(w_\ell); y\right) = z_{\text{heavy}}\right]\right) - 2^{-m} \cdot \frac{q}{q^t} \\
&\geq \left(\varepsilon \cdot 2^{-m} - 2\varepsilon_{\text{in}}\right) \cdot \frac{q}{q^t} \\
&\geq \frac{11\varepsilon}{12} \cdot 2^{-m} \cdot \frac{q}{q^t}
\end{aligned}$$

where the second-to-last inequality follows from the assumption that the extractor fails and from $\Pr[\mathbf{W}_j \in \mathsf{Heavy}_j] \leq 2\varepsilon_{\text{in}}$ (Claim 4.14); and the last inequality follows from our assumption that $\varepsilon_{\text{in}} \leq \frac{\varepsilon}{24} \cdot 2^{-m}$.

We next bound the range of $b'_v p_v$ for arbitrary $v \in \mathbb{F}^t$. First, since b_v is a probability, we have that $b_v \in [0,1]$ and so, $b'_v \in [-2^{-m}, 1-2^{-m}]$. Also as $\mathbf{W}_j | (\mathbf{W}_j \notin \mathsf{Heavy}_j)$ has min-entropy gap $\Delta + 1$, $p_v \in [0, q^{-t} \cdot 2^{\Delta+1}]$. Together, this implies

$$-2^{-m} \cdot 2^{\Delta+1} \cdot q^{-t} \leq b'_v p_v \leq (1-2^{-m}) \cdot 2^{\Delta+1} \cdot q^{-t}.$$

We finally apply tail bound inequality—Lemma 4.20 to random variables $\{b'_v p_v\}_{v \in \mathbf{L}(\mathbb{F})}$ to infer that

$$\Pr\left[\left|\frac{11\varepsilon}{12} \cdot 2^{-m} \cdot \frac{q}{q^t} - \left(\sum_{v \in \mathbf{L}(\mathbb{F})} b'_v p_v\right)\right| \geq \frac{\varepsilon}{3} \cdot 2^{-m} \cdot \frac{q}{q^t}\right] \leq \frac{q \cdot (2^{\Delta+1} q^{-t})^2}{4 \cdot \left(\frac{\varepsilon}{3} \cdot 2^{-m} \cdot \frac{q}{q^t}\right)^2} = \frac{9 \cdot 2^{2m+2\Delta}}{\varepsilon^2 \cdot q}.$$

Rearranging, the claim follows.

□

5.5 Proving the second part of the helper lemma

We here prove the second part of Lemma 5.9. To help prove this, we will need the following result regarding arbitrary codes from [BGDM23], which in turn is an extension of a similar result from [GRS00]:

Lemma 5.12 (Proposition 6 from [BGDM23])**.** *Let $\mathcal{C}$ be an arbitrary $[N, D]_q$ code. Let $S \sim [N]$ be an arbitrary distribution such that $\mathbf{H}_\infty(S) \geq \log(N) - \Delta$. Then, for all $\varepsilon \geq \sqrt{2^{\Delta+1} \cdot (1 - D/N)}$ and all $R \in [q]^N$, the number of codewords $C \in \mathcal{C}$ such that $\Pr_{i \sim S}[C_i = R_i] \geq \varepsilon$ is at most $2/\varepsilon$.*

Proof of second part of Lemma 5.9. Fix a line L such that $\Pr[\mathbf{W}_j \in L(\mathbb{F})|(\mathbf{W}_j \notin \mathsf{Heavy}_j)] \geq \frac{1}{2} \cdot \frac{q}{q^t}$ and $\Pr[\mathbf{W}_j \in \mathsf{Heavy}_j|(\mathbf{W}_j \in L(\mathbb{F}))] \leq \sqrt{\varepsilon_{\text{in}}}$.

Let $(\overline{\mathbf{W}}, \overline{\mathbf{Y}}) = (\mathbf{W}, \mathbf{Y})|\mathbf{W}_j \in L(\mathbb{F}) \setminus \mathsf{Heavy}_j$. For $v \in \mathbb{F}^t$, define

$$\text{bad}^X(v) = \{x \in \mathbb{F}^\ell : \Pr_{y \sim (\overline{\mathbf{Y}}|\overline{\mathbf{W}}_j = v)}[\mathsf{2Ext}(x, y) = z_{\text{heavy}}] \geq 2^{-m} + \varepsilon_{\mathsf{2Ext}}\}.$$

Let $\text{good}^{W_j} = \{v \in \mathbb{F}^t : \left|\text{bad}^X(v)\right| \leq 2^{k_1}\}$. We set $\gamma = \frac{\varepsilon}{24} \cdot 2^{-m}$ and claim the following:

Claim 5.13.

$$\Pr[\overline{\mathbf{W}}_j \in \mathit{good}^{W_j}] \geq 1 - \gamma.$$

We will prove this claim in Section 5.5.1. Let's continue the proof assuming it.

With this, we let

$$\text{bad}^{\mathcal{H}} = \{h : \mathbb{F} \to \mathbb{F}, \deg(h) = d : \Pr_{(w_j,y) \sim (\mathbf{W}_j, \mathbf{Y})|(\mathbf{W}_j \in L(\mathbb{F}))}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}(x_1, \ldots, h(L^{-1}(w_j)), \ldots, x_\ell; y) = z_{\text{heavy}}\right] \geq 2^{-m}(1 + \varepsilon/6)\}.$$

Fix $h \in \text{bad}^{\mathcal{H}}$. We compute

$$\begin{aligned}
&\Pr_{(w_j,y) \sim (\mathbf{W}_j, \mathbf{Y})|(\mathbf{W}_j \in L(\mathbb{F}))}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1, \ldots, h(L^{-1}(w_j)), \ldots, x_\ell; y\right) = z_{\text{heavy}}\right] \\
\leq &\Pr[\mathbf{W}_j \in \mathsf{Heavy}_j|(\mathbf{W}_j \in L(\mathbb{F}))] + \Pr[\overline{\mathbf{W}}_j \notin \text{good}^{W_j}] + \\
&\Pr_{(w_j,y) \sim ((\overline{\mathbf{W}}_j, \overline{\mathbf{Y}})|\overline{\mathbf{W}}_j \in \text{good}^{W_j})}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1, \ldots, h(L^{-1}(w_j)), \ldots, x_\ell; y\right) = z_{\text{heavy}}\right] \\
\leq &\sqrt{\varepsilon_{\text{in}}} + \gamma + \Pr_{(w_j,y) \sim ((\overline{\mathbf{W}}_j, \overline{\mathbf{Y}})|\overline{\mathbf{W}}_j \in \text{good}^{W_j})}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1, \ldots, h(L^{-1}(w_j)), \ldots, x_\ell; y\right) = z_{\text{heavy}}\right]
\end{aligned}$$

where the last inequality followed using Item (b) and Claim 5.13. Using the fact that $h \in \text{bad}^{\mathcal{H}}$ and rearranging, we infer that

$$\begin{aligned}
&\Pr_{(w_j,y) \sim ((\overline{\mathbf{W}}_j, \overline{\mathbf{Y}})|\overline{\mathbf{W}}_j \in \text{good}^{W_j})}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1, \ldots, h(L^{-1}(w_j)), \ldots, x_\ell; y\right) = z_{\text{heavy}}\right] \\
\geq &2^{-m}(1 + \varepsilon/6) - \gamma - \sqrt{\varepsilon_{\text{in}}} \qquad (1)
\end{aligned}$$

We now prove an upper bound on this quantity. We compute that

$$\begin{aligned}
&\Pr_{(w_j,y) \sim ((\overline{\mathbf{W}}_j, \overline{\mathbf{Y}})|\overline{\mathbf{W}}_j \in \text{good}^{W_j})}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1, \ldots, h(L^{-1}(w_j)), \ldots, x_\ell; y\right) = z_{\text{heavy}}\right] \\
\leq &\Pr_{(w_j,y) \sim ((\overline{\mathbf{W}}_j, \overline{\mathbf{Y}})|\overline{\mathbf{W}}_j \in \text{good}^{W_j})}\left[h(L^{-1}(w_j)) \in \text{bad}^X(w_j)\right] + \\
&\Pr_{(w_j,y) \sim ((\overline{\mathbf{W}}_j, \overline{\mathbf{Y}})|(\overline{\mathbf{W}}_j \in \text{good}^{W_j}, h(L^{-1}(w_j)) \notin \text{bad}^X(w_j))}\left[\exists x \in \mathbb{F}^\ell : \mathsf{2Ext}\left(x_1, \ldots, h(L^{-1}(w_j)), \ldots, x_\ell; y\right) = z_{\text{heavy}}\right] \\
\leq &\Pr_{(w_j,y) \sim ((\overline{\mathbf{W}}_j, \overline{\mathbf{Y}})|\overline{\mathbf{W}}_j \in \text{good}^{W_j})}\left[h(L^{-1}(w_j)) \in \text{bad}^X(w_j)\right] + 2^{-m} + \varepsilon_{\mathsf{2Ext}}
\end{aligned}$$

Rearranging and using Equation (1), we infer that

$$\Pr_{w_j \sim (\overline{\mathbf{W}}_j | \overline{\mathbf{W}}_j \in \text{good}^{W_j})} \left[h(L^{-1}(w_j)) \in \text{bad}^X(w_j) \right] \geq \frac{\varepsilon}{6} \cdot 2^{-m} - \varepsilon_{\mathsf{2Ext}} - \gamma - \sqrt{\varepsilon_{\text{in}}}$$
$$\geq \frac{\varepsilon}{12} \cdot 2^{-m} \tag{2}$$

where we used the fact that $\gamma = \frac{\varepsilon}{24} \cdot 2^{-m}, \varepsilon_{\mathsf{2Ext}} \leq \frac{\varepsilon}{48} \cdot 2^{-m}$, and that $\varepsilon_{\text{in}} \leq \frac{\varepsilon^2}{(48)^2} \cdot 2^{-2m}$.

By definition of good^{W_j}, for all $v \in \text{good}^{W_j}$ it holds that $\left|\text{bad}^X(v)\right| \leq 2^{k_1}$. Therefore, for each $v \in \text{good}^{W_j}$ with $\text{bad}^X(v) \neq \emptyset$, we can define an onto function $\phi_v : [2^{k_1}] \to \text{bad}^X(v)$. For $v \in \text{good}^{W_j}$ with $\text{bad}^X(v) = \emptyset$, fix $\phi_v : [2^{k_1}] \to \mathbb{F}^\ell$ arbitrarily. Using this, Equation (2) implies that for all $h \in \text{bad}^{\mathcal{H}}$:

$$\Pr_{w_j \sim (\overline{\mathbf{W}}_j | \overline{\mathbf{W}}_j \in \text{good}^{W_j})} \left[\exists i \in [2^{k_1}] : h(L^{-1}(w_j)) = \phi_{w_j}(i) \right] \geq \frac{\varepsilon}{12} \cdot 2^{-m}$$

This implies that for all $h \in \text{bad}^{\mathcal{H}}$, there exists $i_h \in [2^{k_1}]$ such that

$$\Pr_{w_j \sim (\overline{\mathbf{W}}_j | \overline{\mathbf{W}}_j \in \text{good}^{W_j})} \left[h(L^{-1}(w_j)) = \phi_{w_j}(i_h) \right] \geq \frac{\varepsilon}{12} \cdot 2^{-m-k_1}$$

By an averaging argument, we infer that there exists an index $i^* \in [2^{k_1}]$ and a set $\text{bad}^{\mathcal{H}}_{i^*} \subset \text{bad}^{\mathcal{H}}$ with $\left|\text{bad}^{\mathcal{H}}_{i^*}\right| \geq \left|\text{bad}^{\mathcal{H}}\right| \cdot 2^{-k_1}$ such that for all $h \in \text{bad}^{\mathcal{H}}_{i^*}$, we have that

$$\Pr_{w_j \sim (\overline{\mathbf{W}}_j | \overline{\mathbf{W}}_j \in \text{good}^{W_j})} \left[h(L^{-1}(w_j)) = \phi_{w_j}(i^*) \right] \geq \frac{\varepsilon}{12} \cdot 2^{-m-k_1} \tag{3}$$

Let $\mathcal{C}$ be code where for every univariate degree d polynomial $h : \mathbb{F} \to \mathbb{F}$, we have a codeword $C_h = (h(L^{-1}(a_1)), \ldots, h(L^{-1}(a_r)))$ where $\{a_1, \ldots, a_r\} = \text{good}^{W_j}$ in $\mathcal{C}$. Since each $h \circ L^{-1}$ is a univariate degree d polynomial, we infer that $\mathcal{C}$ is an $[r, r-d-1]_q$ code. Let $C^* \in [q]^r$ be defined as $C^*(v) = \phi_v(i^*)$ for all $v \in \text{good}^{W_j}$. We rewrite Equation (3) in this coding terminology as saying that there exist $\geq \left|\text{bad}^{\mathcal{H}}\right| \cdot 2^{-k_1}$ many codewords $C \in \mathcal{C}$ such that

$$\Pr_{w_j \sim (\overline{\mathbf{W}}_j | \overline{\mathbf{W}}_j \in \text{good}^{W_j})} \left[C(w_j) = C^*(w_j) \right] \geq \frac{\varepsilon}{12} \cdot 2^{-m-k_1}$$

We now apply Lemma 5.12 to $\mathcal{C}$. In order to apply it, we need to claim that we satisfy the required preconditions. We make the following two claims that we will prove in Section 5.5.1.

Claim 5.14. $\mathbf{H}_\infty \left(\overline{\mathbf{W}}_j | \left(\overline{\mathbf{W}}_j \in \mathit{good}^{W_j} \right) \right) \geq \log(q) - \Delta - 3$.

Claim 5.15.

$$\frac{\varepsilon}{12} \cdot 2^{-m-k_1} \geq \sqrt{\frac{d-1}{\left|\mathit{good}^{W_j}\right|} \cdot 2^{\Delta+4}}$$

Using these claims, we indeed satisfy the preconditions and can apply Lemma 5.12 to obtain that

$$\left|\text{bad}^{\mathcal{H}}\right| \cdot 2^{-k_1} \leq \frac{2}{\frac{\varepsilon}{12} \cdot 2^{-m-k_1}}$$

Hence, we conclude that $\left|\text{bad}^{\mathcal{H}}\right| \leq \frac{24 \cdot 2^{2k_1+m}}{\varepsilon}$ as desired. □

5.5.1 Proof of deferred claims

We here prove the deferred claims from proof of second property of Lemma 5.9 that we proved in Section 5.5. These deferred claims are Claim 5.13, Claim 5.14, and Claim 5.15.

We first make a couple of helpful observations:

Observation 5.16.

$$\mathbf{H}_\infty(\overline{\mathbf{W}}_j) \geq \log(q) - \Delta - 2$$

Proof. Recall that $\overline{\mathbf{W}}_j = \mathbf{W}_j|(\mathbf{W}_j \in L(\mathbb{F}) \setminus \mathsf{Heavy}_j)$. By our assumption on L, we know that $\Pr[\mathbf{W}_j|(\mathbf{W}_j \notin \mathsf{Heavy}_j) \in L(\mathbb{F})] \geq \frac{1}{2} \cdot \frac{q}{q^t}$. Also, by assumption on $\mathbf{W}_j|(\mathbf{W}_j \notin \mathsf{Heavy}_j)$, we have that $\mathbf{H}_\infty(\mathbf{W}_j) \geq t\log(q) - \Delta - 1$. With this, we apply Lemma 4.11 to obtain the desired claim. □

Observation 5.17.

$$\mathbf{H}_\infty(\overline{\mathbf{Y}}) \geq \mathbf{H}_\infty(\mathbf{Y}) - t\log(q)$$

Proof. We once again recall that $\overline{\mathbf{Y}}_j = \mathbf{W}_j|(\mathbf{W}_j \in L(\mathbb{F}) \setminus \mathsf{Heavy}_j)$. From our assumption on L, $\Pr[\mathbf{W}_j \in L(\mathbb{F}) \setminus \mathsf{Heavy}_j] \geq \frac{1}{2} \cdot \frac{q}{q^t} \geq q^{-t}$. Using this, we apply Lemma 4.11 to obtain the claim. □

We now prove our first deferred claim.

Proof of Claim 5.13. We apply chain rule for min-entropy (Lemma 4.10) to infer that with probability at least $1-\gamma$ over $v \sim \overline{\mathbf{W}}_j$, it holds that $\mathbf{H}_\infty(\overline{\mathbf{Y}}|(\overline{\mathbf{W}}_j = v)) \geq \mathbf{H}_\infty(\overline{\mathbf{Y}}) - t\log(q) - \log(1/\gamma)$. Using Observation 5.17 and the fact that $\gamma = \frac{\varepsilon}{24} \cdot 2^{-m}$, we have that with at least $1-\gamma$ probability over $v \sim \overline{\mathbf{W}}_j$, it holds that

$$\mathbf{H}_\infty(\overline{\mathbf{Y}}|(\mathbf{W}_j = v)) \geq \mathbf{H}_\infty(\mathbf{Y}) - 2t\log(q) - m - \log(24/\varepsilon) \geq k_2.$$

where the last inequality follows from our assumptions. Using the fact that $\mathsf{2Ext}$ is a $(k_1, k_2, \varepsilon_{\mathsf{2Ext}})$ two-source extractor, we apply Lemma 4.25 to infer that for each such $v \sim \overline{\mathbf{W}}_j$, it holds that

$$\left|\left\{x \in \mathbb{F}^\ell : \Pr_{y \sim (\overline{\mathbf{Y}})}[\mathsf{2Ext}(x,y) = z_{\text{heavy}}] \geq 2^{-m} + \varepsilon_{\mathsf{2Ext}}\right\}\right| \leq 2^{k_1}.$$

By definition of $\text{bad}^X(v)$, the claim follows. □

Using this, we easily obtain our second deferred claim.

Proof of Claim 5.14. From Observation 5.16, we have that $\mathbf{H}_\infty(\overline{\mathbf{W}}_j) \geq \log(q) - \Delta - 2$. Also, using Claim 5.13, we know that $\Pr[\overline{\mathbf{W}}_j \in \text{good}^{W_j}] \geq 1-\gamma$. With this we apply Lemma 4.11 to infer that $\mathbf{H}_\infty\left(\overline{\mathbf{W}}_j | \left(\overline{\mathbf{W}}_j \in \text{good}^{W_j}\right)\right) \geq \log(q) - \Delta - 2 - \log(1/(1-\gamma)) \geq \log(q) - \Delta - 3$ as desired. □

We finally prove our remaining deferred claim:

Proof of Claim 5.15. We first observe that $\left|\text{good}^{W_j}\right| \geq \frac{(1-\gamma)\cdot q}{2^{\Delta+2}}$. Indeed, using Observation 5.16, we infer that for all $v \in \text{good}^{W_j}$, it holds that $\Pr[\overline{\mathbf{W}}_j = v] \leq \frac{2^{\Delta+2}}{q}$. Using Claim 5.13, we have that $\Pr[\overline{\mathbf{W}}_j \in \text{good}^{W_j}] \geq 1-\gamma$. Using an averaging argument, the desired bound follows.

Using this, it suffices to show the following inequality:

$$\frac{\varepsilon}{12} \cdot 2^{-m-k_1} \geq \sqrt{\frac{(d-1) \cdot 2^{\Delta+2}}{(1-\gamma) \cdot q} \cdot 2^{\Delta+4}}$$

Squaring and rearranging, we infer that the above inequality is equivalent to showing that

$$q \geq \frac{9216 \cdot (d-1) \cdot 2^{2\Delta+2m+2k_1}}{(1-\gamma) \cdot \varepsilon^2}.$$

By our assumption on q, we have that $q \geq \frac{18432 \cdot d \cdot 2^{2\Delta+2m+2k_1}}{\varepsilon^2}$ and hence, the above inequality indeed holds. □

6 Putting things together

In this section, we combine all the ingredients we have developed so far, to construct various extractors for samplable sources. We construct extractors for polylogarithmic entropy and logarithmic entropy samplable sources in Section 6.1, and Section 6.2 respectively.

6.1 Extractors for $\mathrm{polylog}(n)$ entropy sources

We obtain the following extractors for samplable sources.

Theorem 6.1 (Technical version of Theorem 1)**.** *Assume* E *is hard for exponential size nondeterministic circuits. There exists a universal constant* C *such that for all constants* $C_{out}, C_{\mathsf{Samp}} > 1$ *and large enough* n*, there exists an explicit function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *that is a* $(k, n^{-C_{out}})$*-extractor, where* $k \geq (\log(n))^C, m = (1 - (1/C_{out}))k$*, for the class of sources samplable by size* $n^{C_{\mathsf{Samp}}}$ *circuits with postselection by size* $n^{C_{\mathsf{Samp}}}$ *nondeterministic circuits.*

Reduction to fewer output bits In order to obtain the claimed output length, we use the following well-known transformation due to [Sha08], together with an observation of [BSS25] that allows one to output more bits from extractors for samplable sources with postselection.

Theorem 6.2 ([Sha08], [BSS25])**.** *For all* $\alpha > 0$*, there exists a universal constant* $C > 0$ *such that the following holds. Let* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^t$ *be a* (k, ε) *extractor for sources samplable by size* s *deterministic circuits with postselection by size* s *(non)deterministic circuits. Furthermore, assume that* $\varepsilon < 2^{-100}$ *and that* $t = C \log(n)$*.*

Then, there exists a (k', ε') *extractor* $\mathsf{Ext}' : \{0,1\}^n \to \{0,1\}^m$ *where* $m = (1-\alpha)k'$ *for sources samplable by size* s' *deterministic circuits with postselection by size* s' *(non) deterministic circuits. Here,* $k' = C \cdot k, \varepsilon' = \varepsilon \cdot 2^{t+C}, s' = s \cdot n^{-C}$*.*

Remark 6.3. *The transformation of [Sha08] requires the extractor to work for sources samplable by nondeterministic circuits. However, [BSS25] observe (see Remark 4.11) that it suffices to consider extractors for sources samplable with postselection.*

Using this, it suffices for us to obtain extractors outputting $O(\log(n))$ bits. Formally, it suffices for us to get:

Theorem 6.4. *Assume* E *is hard for exponential size nondeterministic circuits. There exists a universal constant* C *such that for all constants* $C_{out}, C_{\mathsf{Samp}}$ *and large enough* n*, there exists an explicit function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *that is a* $(k, n^{-C_{out}})$*-extractor, where* $k \geq (\log(n))^C, m = C_{out}\log(n)$*, for the class of sources samplable by size* $n^{C_{\mathsf{Samp}}}$ *circuits with postselection by size* $n^{C_{\mathsf{Samp}}}$ *nondeterministic circuits.*[15]

Proof. We let C be a large enough constant that we set later. Let $\mathbf{Y}$ be a samplable source with $\mathbf{H}_\infty(\mathbf{Y}) \geq k = (\log(n))^C$. Let sExt be the strong seeded extractor from Theorem 4.27 with error parameter set to $n^{-10\cdot C_{\text{out}}}$ and output length $k/4$. Let seed length of sExt be $d = C_{\mathsf{sExt}}(\log(n))$. For $i \in [2^d]$, let $\mathbf{W}_i = \mathsf{sExt}(\mathbf{Y}, i)$. By the property of the strong seeded extractor, $\mathbb{E}_{i\in[2^d]}[|\mathbf{W}_i - \mathbf{U}_{k/4}|] \leq n^{-10\cdot C_{\text{out}}}$. Therefore, by an averaging argument, there must exist some $j \in [2^d]$ such that $\mathbf{W}_j \approx_{n^{-10\cdot C_{\text{out}}}} \mathbf{U}_{k/4}$.

We apply the transformation from Theorem 5.1 to $(\mathbf{W}, \mathbf{Y})$ to obtain the desired m bits, close to uniform. Here, we use the fact that the source $(\mathbf{W}, \mathbf{Y})$ is samplable by polynomial sized circuits with postselection by polynomial sized nondeterministic circuits. We finally sketch that we satisfy each of the constraints required to apply the above transformation. First, the input error $\varepsilon_{\text{in}} = n^{-10\cdot C_{\text{out}}}$ is indeed small enough by our setting of ε and m. Next, using the fact that C is a large enough constant, and that the length of $\mathbf{W}_i$ is smaller than $\mathbf{H}_\infty(\mathbf{Y})/4$, we satisfy the second constraint. For the third constraint, we have $\Delta = 0$, trivially satisfying the requirement. The fourth constraint is also indeed met by our setting of $m, \varepsilon, C_{\text{out}}$. Hence, we indeed satisfy all the constraints and the claim follows. □

6.2 Extractors for $O(\log(n))$ entropy sources

Theorem 6.5 (Technical version of Theorem 2)**.** *Assume* E *is hard for exponential size nondeterministic circuits. Then, for all constants* $\varepsilon > 0$ *and* $C_{\mathsf{Samp}} \geq 1$*, there exists a constant* $C = C(\varepsilon, C_{\mathsf{Samp}})$ *such that for all* $n \in \mathbb{N}$*, there exists an explicit* (k, ε)*-extractor* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ *for the class of sources samplable by size* $n^{C_{\mathsf{Samp}}}$ *circuits with postselection by size* $n^{C_{\mathsf{Samp}}}$ *nondeterministic circuits where* $k \geq C\log(n)$*.*

Proof. We let C be a large enough constant that we set later. Let $\mathbf{Y}$ be a samplable source with $\mathbf{H}_\infty(\mathbf{Y}) \geq k = C\log(n)$. We will first apply a seeded extractor to obtain an elementary SSR source $\mathbf{W}$ and then use our assumption to obtain low degree extension of a hard function, and a two-source extractor to construct our extractor. We will encode each $\mathbf{W}_i$ using the low degree extension and pass them to the two-source extractor. Then, using reconstructive properties from Theorem 5.3, we will show that if our extractor fails, then our circuit lower bound assumption will be violated.

Obtaining elementary SSR source Let sExt be the strong seeded extractor from Theorem 4.27 with error parameter set to $1/n$ and output length $n_w = k/4$. Let the seed length of sExt be $d = C_{\mathsf{sExt}}(\log(n))$ and let $\ell = 2^d$. For $i \in [\ell]$, let $\mathbf{W}_i = \mathsf{sExt}(\mathbf{Y}, i)$.

Let C'_{Samp} be such that the distribution $(\mathbf{W}; \mathbf{Y})$ can be generated by $n^{C'_{\mathsf{Samp}}}$ sized circuits with postselection by $n^{C'_{\mathsf{Samp}}}$ sized nondeterministic circuits. The length of each $\mathbf{W}_i$ is $n_w = (C/4)\log(n)$ and $k_y \geq C\log(n)$. By the property of the strong seeded extractor, we have that

[15] Technically, the theorem stated here has $\varepsilon = 2^{-m}$ but by taking appropriate length prefixes, we can meet the error requirement from Theorem 6.2.

$\mathbb{E}_{i\in[\ell]}[|\mathbf{W}_i - \mathbf{U}_{n_w}|] \le 1/n$. So, by an averaging argument, there must exist some $j \in [\ell]$ such that $\mathbf{W}_j \approx_{1/n} \mathbf{U}_{n_w}$.

Setting up the hard function Using our assumption and Theorem 4.5 we obtain that E is exponentially hard for $\mathsf{NP}_{||}$-circuits. Therefore, there exist $0 < \beta \le 1 \le B$ such that for all r, there exists a function $f : \{0,1\}^r \to \{0,1\}$ such that $f \in \mathsf{DTIME}(2^{B\cdot r})$ and that f requires $\mathsf{NP}_{||}$-circuits of size at least $2^{\beta\cdot r}$.

Setting parameters for the two-source extractor and low degree extension We let $0 < \alpha < 10^{-6}$ be a small enough constant, that we set later. Let $t = \left\lceil \frac{1}{\alpha\cdot\beta} \right\rceil$. Let $n_{w'} = \lceil n_w/(t+1) \rceil \cdot t$. Let $q = 2^{\lceil n_w/(t+1)\rceil}$. Let $r = \left\lceil \frac{C'_{\mathsf{Samp}} \log(n)}{\alpha\cdot\beta} \right\rceil$. We will set our universal constant C to be much larger than t.

Setting up low degree extension Let $f : \{0,1\}^r \to \{0,1\}$ be the function from our circuit lower bound assumption which will be hard for size $n^{\frac{C'_{\mathsf{Samp}}}{\alpha}}$ $\mathsf{NP}_{||}$-circuits. Let $H \subset \mathbb{F}_q$ be any subset of size $\lceil 2^{r/t} \rceil$. Then, we can define $\tau : \{0,1\}^r \to H^t$ to be any natural injective map. We let $\widehat{f} : \mathbb{F}_q^t \to \mathbb{F}_q$ be the low degree extension of f—such that for all $x \in \{0,1\}^r, \widehat{f}(\tau(x)) = f(x)$. Moreover, we can ensure that $\widehat{f}$ is computable in time $\mathrm{poly}(n, 2^r) = \mathrm{poly}(n^{C'_{\mathsf{Samp}}})$ and that $\widehat{f}$ will have total degree at most $t \cdot 2^{r/t} \le t \cdot 2^{C'_{\mathsf{Samp}} \log n} \le n^{C'_{\mathsf{Samp}}+1} \le n^{C'_{\mathsf{Samp}}+2}$.

Setting up the two-source extractor Let $\mathsf{2Ext} : \{0,1\}^{n\cdot\ell} \times \{0,1\}^{n\cdot\ell} \to \{0,1\}$ be the two-source extractor from Theorem 4.29 with error parameter $\varepsilon_{\mathsf{2Ext}} = \varepsilon/C_0$ for a large constant C_0 that we fix later. Let the entropy requirement of $\mathsf{2Ext}$ be $C_1 \cdot \log(n \cdot \ell) \le 2C_1 C_{\mathsf{sExt}} \cdot \log(n)$ where C_1 is a constant, that solely depends on ε/C_0. Since $\ell \cdot \log(q) \le \ell \cdot n$, we can pad with zeros and obtain $\mathsf{2Ext}' : \mathbb{F}_q^\ell \times \{0,1\}^n \to \{0,1\}$. Also, since $\ell \cdot n \le n^{C'_{\mathsf{Samp}}+1}$, $\mathsf{2Ext}$, and hence $\mathsf{2Ext}'$, can be computed in $\mathrm{poly}(n)$ time.

Interpreting each $\mathbf{W}_i$ in finite field Since $t \log(q) = \lceil n_w/(t+1) \rceil \cdot t \le n_w$, for any $v \in \{0,1\}^{n_w}$, we can appropriately truncate the source and interpret v as an element of $\mathbb{F}_q^t$.

Obtaining final output from $(\mathbf{W}, \mathbf{Y})$ We truncate each $\mathbf{W}_i$ as described above to obtain $\mathbf{W}' = (\mathbf{W}'_1, \ldots, \mathbf{W}'_\ell)$ and output $\mathsf{2Ext}'(\widehat{f}(\mathbf{W}'_1), \ldots, \widehat{f}(\mathbf{W}'_\ell); \mathbf{Y})$.

Correctness We proceed by contradiction and assume that the output distribution has statistical distance at least ε from $\mathbf{U}_1$. By an averaging argument, there exists $z \in \{0,1\}$ such that $\Pr[\mathsf{2Ext}'(\widehat{f}(\mathbf{W}'_1), \ldots, \widehat{f}(\mathbf{W}'_\ell); \mathbf{Y}) = z] \ge (1+\varepsilon) \cdot 2^{-1}$. We apply Theorem 5.3 to obtain a small $\mathsf{NP}_{||}$-circuit computing $\widehat{f}$. First, we argue that we indeed satisfy each of the 5 constraints laid out in Assumption 5.2:

1. As $m = 1, \varepsilon_{\mathsf{2Ext}} \le \varepsilon/C_0$, and we set C_0 to be large enough, the constraint is satisfied. Note that doing so still leaves C_0 to be a universal constant, independent of all other constants here.

2. For our setting, $\varepsilon_{\text{in}} = 1/n, \varepsilon_{\mathsf{2Ext}}^2 = O(1)$, the constraint is easily met.

3. By plugging in parameters, the inequality simplifies to showing:

$$(C/2)\log(n) - 1 - \log(24/\varepsilon) \geq k_2 = 2C_1 C_{\mathsf{sExt}} \log(n).$$

Since $\varepsilon = O(1)$, it suffices to show that $C\log(n) \geq 5C_1 C_{\mathsf{sExt}} \log(n)$. We will set our constant C to be much larger than $C_1 C_{\mathsf{sExt}}$ so that the inequality is satisfied. Here we crucially use the fact that C_{sExt} as well as the runtime of sExt does not depend on the initial entropy and hence, does not depend on C. This implies that $\ell \cdot n$ is independent of C, and so C_1 is also independent of C (here we also use the fact that C_0 is an independent universal constant). Therefore, we can set C to be much larger than $C_1 C_{\mathsf{sExt}}$ as desired.

4. After taking logarithm and plugging in parameters, the inequality simplifies to showing:

$$\alpha \cdot \beta \cdot n_{w'} \geq O(1) + (C'_{\mathsf{Samp}} + 2)\log(n) + 2k_1 = O(1) + (C'_{\mathsf{Samp}} + 2)\log(n) + 2C_1 C_{\mathsf{sExt}} \log(n).$$

We set our constant C to be much larger than $C'_{\mathsf{Samp}} + 2$ and $2C_1 C_{\mathsf{sExt}}$, taking advantage of the fact that α, β are universal constants so that the constraint is satisfied. As discussed above, we carefully observe that we are allowed to make C depend on C_1, C_{sExt}. We additionally observe that the runtime of sExt does not depend on C. As $C_{\mathsf{Samp}'}$ solely depends on C_{Samp} and the runtime of sExt, it is independent of C. So, we can set C to be much larger than C'_{Samp} to satisfy the above inequality.

5. The first four terms of this constraint are easily seen to be $o(1)$. For the last term, by taking logarithm and simplifying, it suffices to show:

$$\alpha \cdot \beta \cdot n_{w'} \geq O(1) + (C'_{\mathsf{Samp}} + 2)\log(n) + 2k_1.$$

As $k_1 \leq 2C_1 C_{\mathsf{sExt}} \log(n)$, we can use above arguments to set C large enough so that this inequality is also satisfied.

Therefore, we satisfy all the constraints and can apply Theorem 5.3 to obtain an $\mathsf{NP}_{||}$-circuit $D_{\widehat{f}} : \mathbb{F}_q^t \to \mathbb{F}$ of size $\text{poly}(n^{C_{\mathsf{Samp}}}, \ell, \log(q), d, n)$ computing $\widehat{f}$ (we here incorporated the runtime of $\mathsf{2Ext}$, using the fact that C_0 is an independent universal constant). Using $D_{\widehat{f}}$, and the injective function τ, we can obtain an $\mathsf{NP}_{||}$-circuit $D_f : \{0,1\}^r \to \{0,1\}$ such that $D_f(x) = f(x)$ for all $x \in \{0,1\}^r$. Moreover, the size of D_f will be same as that of $D_{\widehat{f}}$. So, to obtain a contradiction to the hardness assumption of f, it suffices to show that size of $D_{\widehat{f}}$ is at most $2^{\beta r} = n^{C'_{\mathsf{Samp}}/\alpha}$.

As $d \leq n^{C_{\mathsf{Samp}'}+2}$ and $\ell \leq n^{C'_{\mathsf{Samp}}}$, we upper bound the size of $D_{\widehat{f}}$ by $n^{C'_{\mathsf{Samp}} \cdot C_2}$ where $C_2 \geq 1$ is a universal constant, independent of C'_{Samp}. We set $\alpha = \frac{1}{10^6 \cdot C_2}$ to obtain the desired contradiction. □

Acknowledgements

We thank an anonymous reviewer for very helpful feedback.

References

[AASY16] Benny Applebaum, Sergei Artemenko, Ronen Shaltiel, and Guang Yang. "Incompressible functions, relative-error extractors, and the power of nondeterministic reductions". In: *Computational complexity* 25.2 (2016), pp. 349–418 (cit. on p. 3).

[BGDM23] Marshall Ball, Eli Goldin, Dana Dachman-Soled, and Saachi Mutreja. "Extracting randomness from samplable distributions, revisited". In: *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE. 2023, pp. 1505–1514 (cit. on pp. 2, 4, 6, 7, 9, 11, 31).

[BSS25] Marshall Ball, Ronen Shaltiel, and Jad Silbak. "Extractors for samplable distributions with low min-entropy". In: *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*. 2025, pp. 596–603 (cit. on pp. 3–6, 35).

[BGP00] Mihir Bellare, Oded Goldreich, and Erez Petrank. "Uniform generation of NP-witnesses using an NP-oracle". In: *Information and Computation* 163.2 (2000), pp. 510–526 (cit. on p. 5).

[CZ19] Eshan Chattopadhyay and David Zuckerman. "Explicit two-source extractors and resilient functions". In: *Annals of Mathematics* 189.3 (2019), pp. 653–705 (cit. on pp. 1, 5, 8, 9, 11, 17).

[CG88] Benny Chor and Oded Goldreich. "Unbiased bits from sources of weak randomness and probabilistic communication complexity". In: *SIAM Journal on Computing* 17.2 (1988), pp. 230–261 (cit. on p. 1).

[DMOZ22] Dean Doron, Dana Moshkovitz, Justin Oh, and David Zuckerman. "Nearly optimal pseudorandomness from hardness". In: *Journal of the ACM* 69.6 (2022), pp. 1–55 (cit. on p. 2).

[DGW09] Zeev Dvir, Ariel Gabizon, and Avi Wigderson. "Extractors and rank extractors for polynomial sources". In: *computational complexity* 18.1 (2009), pp. 1–58 (cit. on p. 1).

[GR08] Ariel Gabizon and Ran Raz. "Deterministic extractors for affine sources over large fields". In: *Combinatorica* 28.4 (2008), pp. 415–440 (cit. on p. 1).

[GLRSW91] Peter Gemmell, Richard Lipton, Ronitt Rubinfeld, Madhu Sudan, and Avi Wigderson. "Self-testing/correcting for polynomials and for approximate functions". In: *STOC*. Vol. 91. 1991, pp. 32–42 (cit. on pp. 10, 21).

[GRS00] Oded Goldreich, Ronitt Rubinfeld, and Madhu Sudan. "Learning polynomials with queries: The highly noisy case". In: *SIAM Journal on Discrete Mathematics* 13.4 (2000), pp. 535–570 (cit. on p. 31).

[GUV09] Venkatesan Guruswami, Christopher Umans, and Salil P. Vadhan. "Unbalanced expanders and randomness extractors from Parvaresh-Vardy codes". In: *J. ACM* 56.4 (2009), 20:1–20:34 (cit. on pp. 7, 8, 17).

[IW97] Russell Impagliazzo and Avi Wigderson. "P= BPP if E requires exponential circuits: Derandomizing the XOR lemma". In: *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*. 1997, pp. 220–229 (cit. on p. 2).

[JVV86] Mark R Jerrum, Leslie G Valiant, and Vijay V Vazirani. "Random generation of combinatorial structures from a uniform distribution". In: *Theoretical computer science* 43 (1986), pp. 169–188 (cit. on p. 5).

[KRVZ06] Jesse Kamp, Anup Rao, Salil Vadhan, and David Zuckerman. "Deterministic extractors for small-space sources". In: *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*. ACM. 2006, pp. 691–700 (cit. on p. 1).

[KM02] Adam R Klivans and Dieter van Melkebeek. "Graph Nonisomorphism Has Subexponential Size Proofs Unless the Polynomial-Time Hierarchy Collapses". In: *SIAM Journal on Computing* 31.5 (2002), pp. 1501–1526 (cit. on p. 2).

[Li16] Xin Li. "Improved two-source extractors, and affine extractors for polylogarithmic entropy". In: *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE. 2016, pp. 168–177 (cit. on pp. 5, 8, 9, 11, 17).

[Li23] Xin Li. "Two source extractors for asymptotically optimal entropy, and (many) more". In: *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE. 2023, pp. 1271–1281 (cit. on pp. 1, 5, 9, 11, 17).

[MW97] Ueli Maurer and Stefan Wolf. "Privacy amplification secure against active adversaries". In: *Advances in Cryptology—CRYPTO'97: 17th Annual International Cryptology Conference Santa Barbara, California, USA August 17–21, 1997 Proceedings 17*. Springer. 1997, pp. 307–321 (cit. on p. 13).

[OS26] Justin Oh and Ronen Shaltiel. "Extractors for Samplable Distributions from the Two-Source Extractor Recipe". In: *STOC*. 2026 (cit. on p. 3).

[Sha08] Ronen Shaltiel. "How to get more mileage from randomness extractors". In: *Random Structures & Algorithms* 33.2 (2008), pp. 157–186 (cit. on pp. 8, 35).

[Sha25a] Ronen Shaltiel. "Extractors for Samplable Distributions with Polynomially Small Min-Entropy". In: *66th IEEE Annual Symposium on Foundations of Computer Science, FOCS*. 2025 (cit. on pp. 3, 4, 6).

[Sha25b] Ronen Shaltiel. "Multiplicative Extractors for Samplable Distributions". In: *40th Computational Complexity Conference, CCC 2025*. Ed. by Srikanth Srinivasan. Vol. 339. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2025, 22:1–22:22 (cit. on pp. 2–4, 9).

[SU06] Ronen Shaltiel and Christopher Umans. "Pseudorandomness for approximate counting and sampling". In: *computational complexity* 15.4 (2006), pp. 298–341 (cit. on pp. 9, 13).

[Sip83] Michael Sipser. "A complexity theoretic approach to randomness". In: *Proceedings of the fifteenth annual ACM symposium on Theory of computing*. 1983, pp. 330–335 (cit. on p. 5).

[Sto83] Larry Stockmeyer. "The complexity of approximate counting". In: *Proceedings of the fifteenth annual ACM symposium on Theory of computing*. 1983, pp. 118–126 (cit. on p. 5).

[STV01] Madhu Sudan, Luca Trevisan, and Salil P. Vadhan. "Pseudorandom Generators without the XOR Lemma". In: *J. Comput. Syst. Sci.* 62.2 (2001), pp. 236–266 (cit. on p. 5).

[TV00] Luca Trevisan and Salil Vadhan. "Extracting randomness from samplable distributions". In: *Proceedings 41st Annual Symposium on Foundations of Computer Science*. IEEE. 2000, pp. 32–42 (cit. on pp. 1, 2, 4–6, 8).

[Vio14] Emanuele Viola. "Extractors for circuit sources". In: *SIAM Journal on Computing* 43.2 (2014), pp. 655–672 (cit. on p. 1).

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il