Multiplicative and Additive Extractors for Low Entropy Samplable Sources Under Reduced Assumptions

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

Ronen Shaltiel‡
University of Haifa
ronen@cs.haifa.ac.il

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

Abstract

Trevisan and Vadhan [TV, FOCS'00] constructed (seedless) extractors for distributions samplable by poly-size circuits that have very high min-entropy, under the very strong complexity theoretic hardness assumption that $\mathsf{E} = \mathsf{DTIME}(2^{O(n)})$ requires size $2^{\Omega(n)}$ circuits that have oracle access to Σ_6^{P}. In this paper we give three new explicit constructions of extractors for samplable distributions.

First, we construct an extractor that extracts one bit, with constant error, for min-entropy $k = O(\log n)$, which matches the non-constructive probabilistic argument. This result uses hardness for nondeterministic circuits, and achieves the same parameters as the recent extractor of Oh and Shaltiel [OS, STOC'26] under an incomparable hardness assumption. In contrast to the extractor of [OS26], our extractor (and, similarly, the constructions below) also works for the more general class of distributions samplable with postselection (a model introduced by [BGDM23] that generalizes samplable and recognizable sources).

Second, we construct an extractor that extracts almost all the entropy, with polynomially small error, for min-entropy $k = n^{\alpha}$, where $\alpha > 0$ is an arbitrary constant. This result also uses hardness for nondeterministic circuits, and significantly improves upon the recent extractor of [Shaltiel, FOCS'25] which required a much stronger assumption. This extractor is also *multiplicative*: a stronger notion suggested in [AASY, Comp. Complexity'16; Shaltiel, CCC'25] to make extractors for samplable distributions suitable for selecting keys for cryptographic protocols.

Third, we construct an extractor that extracts almost all the entropy, with polynomially small error, for min-entropy $k = \mathrm{polylog}(n)$. This result uses hardness for circuits that have oracle access to Σ_2^{P}.

Prior to this work, the lowest min-entropy threshold k that was achieved using hardness for nondeterministic circuits (or even circuits with oracle access to Σ_2^{P}) was larger than $0.99n$ [BGDM, FOCS'23]. Our results significantly improve both the min-entropy threshold and the hardness assumptions compared to all past work that extracts more than one bit.

Our constructions build upon earlier work [TV, FOCS'00; BGDM, FOCS'23; BSS, STOC'25; Shaltiel, FOCS'25] while introducing several new ideas: (1) We simplify and build upon the ideas of [BGDM, FOCS'23] to construct our extractors under nondeterministic hardness. (2) We introduce a notion of robust errorless condensers and construct and use them to obtain our multiplicative extractors. (3) We adapt the two-source extractor recipe from [CZ, Ann. of Mathematics'19] to construct our additive extractors. We do this by constructing and using a correlation breaker against small circuits. Our correlation breaker extends recent constructions of non-malleable extractors against small circuits [BDL, CRYPTO'22; BD, CRYPTO'25].

*Supported by NSF Awards CCF-2443735 and CCF-2552209.

†Supported by NSF Award CCF-2514586.

‡Supported by ISF grant 1006/23 and also co-funded by the European Union (ERC, NFITSC, 101097959). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

Contents

1 Introduction **1**
- 1.1 Background . . . 1
- 1.2 Our Results . . . 2
  - 1.2.1 Standard (Additive) Extractors . . . 3
  - 1.2.2 Multiplicative Extractors . . . 3

2 Overview of the Techniques **5**
- 2.1 Proof Overview for Multiplicative Extractors . . . 5
  - 2.1.1 Constructing robust block errorless condenser, and multiplicative extractor from errorless condenser . . . 7
  - 2.1.2 Constructing errorless condenser from robust block errorless condenser . . . 9
- 2.2 Proof Overview for Additive Extractors . . . 10
  - 2.2.1 Errorless seeded condensing to obtain many good rows. . . . 11
  - 2.2.2 Local correlation breaking . . . 12
  - 2.2.3 The one-bit construction . . . 12
  - 2.2.4 The many-bit construction . . . 13
  - 2.2.5 Construction of the correlation breaker . . . 15
- 2.3 Organization . . . 17

3 AI Use Statement **18**

4 Preliminaries **18**
- 4.1 Basic Probability . . . 18
  - 4.1.1 Standard definitions . . . 18
  - 4.1.2 Useful probability lemmas . . . 19
- 4.2 Circuits and Hardness assumptions . . . 19
  - 4.2.1 Circuit definitions . . . 19
  - 4.2.2 Hardness assumptions . . . 20
- 4.3 Extractors, Condensers, and Models of Randomness . . . 20
  - 4.3.1 Extractors and condensers for a class of distributions . . . 20
  - 4.3.2 Seeded extractors . . . 21
  - 4.3.3 Block-wise sources . . . 21
  - 4.3.4 Two-source extractors and condensers . . . 21
  - 4.3.5 Two-source non-malleable extraction . . . 22
  - 4.3.6 Samplable distributions . . . 23
  - 4.3.7 Extraction from samplable sources . . . 23
  - 4.3.8 Robust errorless block condensers . . . 24
  - 4.3.9 Output amplification . . . 25
- 4.4 Algebra and Coding theory tools . . . 25
  - 4.4.1 Low-degree extension . . . 25
  - 4.4.2 Parametrized lines . . . 25
  - 4.4.3 List recovery . . . 26
- 4.5 Interactive Protocols . . . 27
  - 4.5.1 Approximate counting and pseudorandomness . . . 27

I Multiplicative Extractors **27**

5 Additive Extractors from Multiplicative Extractors **29**

6 Errorless condensers using robust errorless block condensers **29**
6.1 Structural chopping result for any source . . . 30
6.2 Constructing basic errorless Condenser . . . 32
6.3 Constructing the Errorless Condenser . . . 34

7 Overview of the proofs of Theorem 4.43 and Theorem 4.41 **38**
7.1 The extractor of [TV00] and the analysis of [BGDM23] . . . 38
7.2 A simpler test that can be generalized to low min-entropy . . . 39
7.3 Extending the construction and test to the case of low min-entropy . . . 42
7.4 Special modifications for Theorem 4.43 . . . 44
7.5 Special modifications for Theorem 4.41 . . . 44

8 Joint preparations **44**
8.1 A nondeterministic circuit for counting the sum of maximums . . . 44
8.2 Reconstruction from a polynomial test . . . 46
8.2.1 Pinning the correct polynomial . . . 46
8.2.2 Correcting a nondeterministic approximation . . . 48

9 Proof of the robust errorless block condenser theorem **49**
9.1 Construction . . . 49
9.1.1 Components . . . 49
9.1.2 Operation . . . 50
9.2 Analysis of the construction . . . 51
9.2.1 Preparations . . . 51
9.2.2 The test . . . 53
9.2.3 Completeness . . . 54
9.2.4 Soundness . . . 55
9.2.5 Constructing Thr . . . 58
9.2.6 Reconstruction and completion of the proof . . . 58

10 Proof of the multiplicative extractor theorem **59**
10.1 Construction . . . 59
10.1.1 Components . . . 60
10.1.2 Operation . . . 62
10.2 Analysis of the construction . . . 62
10.2.1 Preparations . . . 62
10.2.2 The test . . . 64
10.2.3 Completeness . . . 65
10.2.4 Soundness . . . 66
10.2.5 Constructing Thr . . . 68
10.2.6 Reconstruction and completion of the proof . . . 69

II Additive Extractors **70**

11 Errorless seeded condensation by hashing **70**

12 Decorrelation of correlated samplable sources **72**
12.1 Section Overview . . . 72
12.2 Computational preliminaries . . . 74
12.3 Correlation Breaker Ingredients . . . 75
12.4 Correlation Breaker Construction Template . . . 76
12.5 Reduction Template . . . 77
12.6 Additive decorrelation . . . 79
12.6.1 A constant number of rows and one output bit . . . 81
12.6.2 More rows and output bits . . . 82
12.7 Multiplicative decorrelation . . . 82
12.7.1 The interactive lemma . . . 83
12.7.2 Construction and analysis . . . 85
12.8 Relativization of the generator constructions . . . 86

13 One-bit extraction from logarithmic entropy **88**
13.1 Tools . . . 88
13.2 The extractor construction . . . 89
13.3 Analysis of the construction . . . 90

14 Extraction from polylogarithmic entropy **90**
14.1 Tools . . . 91
14.2 The extractor construction . . . 92
14.3 Analysis . . . 94
14.4 Extracting almost all the entropy . . . 95

15 Polynomial-time extraction from polylogarithmic entropy **96**
15.1 Relative probabilities of bin labels . . . 96
15.2 The polynomial-time construction . . . 96

1 Introduction

1.1 Background

Randomized algorithms and cryptographic protocols typically assume access to uniform random bits. Available sources of randomness, however, may be biased and exhibit dependencies. Randomness extractors aim to convert such "weak sources of randomness" into distributions that are close to uniform, assuming that the source contains sufficient randomness (measured by min-entropy). Deterministic extraction is impossible for general sources with high min-entropy. We therefore restrict the class of sources, seeking a single extractor that works for every source in the class that has sufficient min-entropy.

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

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

Definition 1.2 (Samplable distributions)**.** *We say that a distribution* X *over* $\{0,1\}^n$ *is* sampled *by a function* $\mathsf{Samp} : \{0,1\}^r \to \{0,1\}^n$ *if* $X = \mathsf{Samp}(U_r)$. X *is* samplable *by a class* $\mathcal{C}$ *of functions, if there exists* $\mathsf{Samp} \in \mathcal{C}$ *that samples* X.

Trevisan and Vadhan made the philosophical assumption that "nature is efficient" and that therefore, weak sources of randomness that are "available to computers" are efficiently samplable. They considered extractors for distributions that are samplable by poly-size circuits, namely distributions samplable by circuits of size n^c for some constant parameter c. They showed that such extractors cannot run in time smaller than n^c, and therefore considered extractors that run in time n^d for a constant $d > c$. They showed that such extractors imply circuit lower bounds, and so, motivated by the hardness vs. randomness paradigm, they gave a conditional construction based on hardness assumptions.

Hardness assumptions against various types of nondeterministic circuits. We say that "E is hard for exponential size circuits of some type", if there exists a problem $L \in \mathsf{E} = \mathsf{DTIME}(2^{O(n)})$ and a constant $\beta > 0$, such that for every sufficiently large n, circuits of size $2^{\beta \cdot n}$ (of the specified type) fail to compute the characteristic function of L on inputs of length n. (See Section 4.2 for a more formal definition).

The assumption that E is hard for exponential size (deterministic) circuits was used by the celebrated paper of Impagliazzo and Wigderson [IW97] to imply that $\mathsf{BPP} = \mathsf{P}$. The stronger assumption that E is hard for exponential size nondeterministic circuits[2], originated in works on hardness versus randomness for AM, and is used in many results [AK02; KM02; MV05; SU05; BOV07; GW02; GST03; SU06; SU09; Dru13; AASY16; BV17; AIKS16; HNY17; DMOZ22; BDL22; CT22; BGDM23; BSS24; SS24; Sha25b]. It can be viewed as a scaled, nonuniform version of the widely believed assumption that $\mathsf{EXP} \neq \mathsf{NP}$.[3]

In their seminal paper on extractors for samplable distributions, Trevisan and Vadhan [TV00] introduced a version of the assumption for a stronger circuit class. A Σ_i-circuit, is a circuit that in addition to the standard gates, is also allowed to use a special gate (with large fan-in) that solves the canonical complete language for the class Σ_i^{P} (the i'th level of the polynomial time hierarchy).[4] The extractor of Trevisan and Vadhan [TV00] relies on the strong assumption that E is hard for exponential size Σ_6-circuits (which can be viewed as a scaled, nonuniform version of the widely believed assumption that $\mathsf{EXP} \neq \Sigma_6^{\mathrm{P}}$).

[1] See Section 4.1.1 for the standard definitions of min-entropy and statistical distance.

[2] A precise definition of nondeterministic circuits appears in Section 4.2.

[3] We remark that the recent refutation of the ETH assumption, does not refute any of these assumptions.

[4] A Σ_i-circuit is a nonuniform analogue of the class $\mathsf{P}^{\Sigma_i^{\mathrm{P}}}$ that contains Σ_i^{P}, and recall that $\mathsf{P} = \Sigma_0^{\mathrm{P}}$ and $\mathsf{NP} = \Sigma_1^{\mathrm{P}}$. See Definition 4.8 for a formal definition.

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

Theorem 1.3 ([TV00]). *If* E *is hard for exponential size* Σ_6*-circuits then for every constant* $\gamma > 0$*, and every constant* $c > 1$*, there are constants* $\eta > 0$ *and* $d > c$ *such that for every sufficiently large* n*, and every* $(1-\eta)n \leq k \leq n$ *there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *that is a* (k, ε)*-extractor for distributions samplable by circuits of size* n^c*, where* $\varepsilon = n^{-c}$ *and* $m = (1-\gamma)k$*. Furthermore,* Ext *is computable in time* n^d*.*

Recent work on extractors for samplable distributions improved Theorem 1.3 in several respects.

Weaker hardness assumptions: Ball, Goldin, Dachman-Soled and Mutreja [BGDM23] (see also Shaltiel [Sha25b]) achieved the conclusion of Theorem 1.3 under the weaker, and more standard assumption that E is hard for exponential size nondeterministic circuits. Prior to this work, this was the lowest min-entropy threshold achieved under this hardness assumption. The objective of this paper is to construct extractors for sources with substantially lower min-entropy under the same assumption.

Extractors for samplable distributions with low min-entropy: Ball, Shaltiel and Silbak [BSS25] improved the min-entropy threshold of Trevisan and Vadhan to $k = n^{1-\gamma}$ under a hardness assumption for Σ_5-circuits. Shaltiel [Sha25a] improved the min-entropy threshold to $k = n^{1/i}$ for an arbitrary constant i, at the cost of a hardness assumption for Σ_{i+3}-circuits.

Different hardness assumptions: Oh and Shaltiel [OS26] gave an extractor construction under a different and incomparable hardness assumption; the extractor of [OS26] has the advantage that it achieves very low min-entropy of $k = O(\log n)$, but it only extracts one bit, with constant error.

Samplable distributions with postselection: Ball, Goldin, Dachman-Soled and Mutreja [BGDM23] considered a generalization of samplable distributions which allows "postselection". In this model, there are two functions: a sampler $\mathsf{Samp} : \{0,1\}^r \to \{0,1\}^n$, and a post-selector, $\mathsf{Post} : \{0,1\}^r \to \{0,1\}$, and we say that X is sampled by Samp with postselection by Post, if X is the distribution $(\mathsf{Samp}(R) | \, \mathsf{Post}(R) = 1)$, for $R \leftarrow U_r$. Note that this is indeed a generalization of samplable distributions, as we can take the circuit $\mathsf{Post}(R)$ that outputs one on every input. One motivation for such sources, suggested in [BGDM23], is that this experiment captures a scenario where information about the sampled value $\mathsf{Samp}(R)$ is leaked to an adversary in the form of an efficiently computable leakage function $L(R)$. Postselection captures conditioning on any specific value of $L(R)$. This means that extraction remains possible even given the leakage, provided sufficient min-entropy remains.

With the exception of the aforementioned extractor of Oh and Shaltiel [OS26], all constructions in the literature (including the constructions of this paper) apply also to samplable distributions with postselection. In fact, this holds even if the circuit Post is allowed to be a size n^c nondeterministic circuit. Shaltiel [Sha25b] shows that extractors for this richer class of distributions imply lower bounds against nondeterministic circuits.[5]

A detailed comparison of past work and our new results appears in Table 1.

1.2 Our Results

In this paper we give constructions of both standard extractors, and of multiplicative extractors (that are defined in Section 1.2.2).

[5] It is not known whether explicit extractors for distributions samplable by deterministic circuits with post selection by nondeterministic circuits imply that E is hard for exponential size nondeterministic circuits. What is shown in [Sha25b] is that such extractors cannot be computed (even on the average) by nondeterministic circuits of size slightly smaller than n^c.

1.2.1 Standard (Additive) Extractors

We construct extractors for samplable distributions with min-entropy $k = O(\log n)$. This min-entropy threshold k matches the one achieved by the nonconstructive probabilistic argument (up to constants). Our first result assumes hardness for nondeterministic circuits, but only extracts one bit with constant error.

Theorem 1 (extracting one bit for $k = O(\log n)$ under nondeterministic hardness)**.** *If* E *is hard for exponential size nondeterministic circuits then for every constant* $\varepsilon > 0$*, and every constant* $c > 1$*, there are constants* d, a *such that for every sufficiently large* n*, and every* $a \log n \le k \le n$ *there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ *that is a* (k, ε)*-extractor for distributions samplable by circuits of size* n^c*. Furthermore,* Ext *is computable in time* n^d*.*

The parameters of the extractor of Theorem 1 match the parameters of the extractor of [OS26]. The hardness assumptions in these two results are incomparable.

The extractor of Theorem 1 works for distributions that are samplable by deterministic circuits with postselection by nondeterministic circuits (see Theorem 13.1 for a precise statement) whereas that of [OS26] is not known to allow samplable distributions with postselection, even if postselection is by deterministic circuits.

Our next result gives an extractor that outputs almost all the randomness with polynomially small error. This is achieved under hardness for nondeterministic circuits, but applies only for $k = n^\alpha$ for an arbitrary constant $\alpha > 0$.

Theorem 2 (extracting many bits for $k = n^\alpha$ under nondeterministic hardness)**.** *If* E *is hard for exponential size nondeterministic circuits then for all constants* $\alpha, \gamma > 0$*, and every constant* $c > 1$*, there is a constant* d *such that for every sufficiently large* n*, and every* $n^\alpha \le k \le n$ *there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *that is a* (k, ε)*-extractor for distributions samplable by circuits of size* n^c*, with* $\varepsilon = n^{-c}$ *and* $m = (1-\gamma)k$*. Furthermore,* Ext *is computable in time* n^d*.*

This extractor has exactly the same parameters as the extractor construction of Shaltiel [Sha25a]. However, whereas Shaltiel's extractor assumes hardness for Σ_i-circuits with $i = 3 + 1/\alpha$, Theorem 2 assumes only hardness for nondeterministic circuits. See Theorem 5.1 for the general statement.

We are also able to construct extractors that extract almost all of the randomness from sources with min-entropy $k = \mathrm{polylog}(n)$ with error $\varepsilon = n^{-c}$. However, for this result, we use a hardness assumption against Σ_2-circuits.

Theorem 3 (extracting many bits for $k = \mathrm{polylog}(n)$ under Σ_2-hardness)**.** *If* E *is hard for exponential size* Σ_2*-circuits then there exists a constant* a *such that for every constant* $\gamma > 0$*, and every constant* $c > 1$*, there is a constant* d *such that for every sufficiently large* n*, and every* $\log^a n \le k \le n$ *there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *that is a* (k, ε)*-extractor for distributions samplable by circuits of size* n^c*, with* $\varepsilon = n^{-c}$ *and* $m = (1-\gamma)k$*. Furthermore,* Ext *is computable in time* n^d*.*

Compared with previous constructions that extract many bits [BSS25; Sha25a], Theorem 3 achieves both a lower min-entropy threshold and a weaker hardness assumption. See Theorem 15.1 for the general statement.

We remark that the hardness assumption in Theorem 3 can be weakened to hardness for nondeterministic circuits, at the cost of allowing the extractor to run in time $2^{\log^2 n}$. See a more general statement in Theorem 14.2.

1.2.2 Multiplicative Extractors

A signature application of seedless extractors is choosing keys for cryptographic protocols by extracting randomness from weak random sources. (Indeed, this was the original motivation of Trevisan and Vadhan

Table 1: Constructions of extractors for samplable distributions.
Below, $\gamma > 0$ is an arbitrarily small constant, $\eta > 0$ is a sufficiently small constant, $\alpha > 0$ and $c \geq 1$ are arbitrary constants, and $i > 1$ is an arbitrary integer.

Reference	Hardness assumption	Threshold k	Output m	Error ε	Post-selection?
[TV00]	E is hard for exponential-size Σ_6-circuits	$(1-\eta)n$	$(1-\gamma)k$	n^{-c}	Yes
[BGDM23]	E is hard for exponential-size nondeterministic circuits	$(1-\eta)n$	$(1-\gamma)k$	n^{-c}	Yes
[BSS25]	E is hard for exponential-size Σ_5-circuits	$n^{1-\gamma}$	$(1-\gamma)k$	n^{-c}	Yes
[Sha25a]	E is hard for exponential-size Σ_{i+3}-circuits	$n^{1/i}$	$(1-\gamma)k$	n^{-c}	Yes
[OS26]	An incomparable hardness assumption	$O(\log n)$	1	$\Theta(1)$	No
This work Thm 1	**E is hard for exponential-size nondeterministic circuits**	$\boldsymbol{O(\log n)}$	$\mathbf{1}$	$\boldsymbol{\Theta(1)}$	**Yes**
This work Thm 2	**E is hard for exponential-size nondeterministic circuits**	$\boldsymbol{n^{\alpha}}$	$\boldsymbol{(1-\gamma)k}$	$\boldsymbol{n^{-c}}$	**Yes**
This work Thm 3	**E is hard for exponential-size Σ_2-circuits**	$\mathbf{polylog}(\boldsymbol{n})$	$\boldsymbol{(1-\gamma)k}$	$\boldsymbol{n^{-c}}$	**Yes**

[TV00] for introducing extractors for samplable distributions). More specifically, consider a cryptographic protocol which is known to be secure when the key of an honest party is chosen according to U_m. That is, the probability that an adversary can steal the honest party's money is smaller than some "negligible" $\mu > 0$.

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

Applebaum, Artemenko, Shaltiel and Yang [AASY16] showed that "current techniques" cannot yield extractors for samplable distributions with error $\varepsilon = n^{-\omega(1)}$, under a hardness assumption against Σ_i-circuits, and this holds for every constant i. (See [AASY16] for a precise formulation). This means that with current extractors for samplable distributions (which only achieve $\varepsilon = n^{-c}$) we cannot expect cryptographic protocols to maintain negligible probability of cheating, when selecting the key with an extractor.

In order to address this problem, Applebaum et al. [AASY16], and later Shaltiel [Sha25b], introduced a notion of "multiplicative extractors" which are designed to maintain negligible probability of cheating in cryptographic protocols. Below, we use the definition of [Sha25b] which is inspired by choices made in definitions in differential privacy [DMNS06].

Definition 1.4 (Multiplicative seedless extractor [Sha25b])**.** *A function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *is a* (k,ε)-multiplicative extractor *for a class* $\mathcal{D}$ *of distributions, if for every distribution* $X \in \mathcal{D}$ *over* $\{0,1\}^n$ *with* $\mathbf{H}_\infty(X) \geq k$*, and every event* $A \subseteq \{0,1\}^m$

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

It is easy to see that every multiplicative extractor is also a standard extractor.[6] However, unlike regular extractors, multiplicative extractors apply for the application of selecting keys for cryptographic protocols, even if $\varepsilon > 0$ is a small constant. This is because such extractors give that the adversary's probability to cheat is at most $e^{\varepsilon} \cdot \mu \leq (1 + O(\varepsilon)) \cdot \mu$.[7]

[6] Indeed, it is easy to see that a (k,ε) multiplicative extractor is a $(k, O(\varepsilon))$ additive extractor, see [Sha25b].

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

Table 2: Constructions of multiplicative extractors for samplable distributions.
Below, $\eta > 0$ is a sufficiently small constant, $\alpha > 0$ and $c \geq 1$ are arbitrary constants, and $i > 1$ is an arbitrary integer.

Reference	Hardness assumption	Threshold k	Output m	Error ε	Post-selection?
[AASY16]	E is hard for exponential-size Σ_4-circuits	$(1-\eta)n$	ηn	n^{-c}	Yes
[Sha25b]	E is hard for exponential-size nondeterministic circuits	$(1-\eta)n$	ηn	n^{-c}	Yes
[Sha25a]	E is hard for exponential-size $\Sigma_{i+O(1)}$-circuits	$n^{1/i}$	$k^{\Omega(1)}$	n^{-c}	Yes
This work Thm 4	**E is hard for exponential-size nondeterministic circuits**	$\boldsymbol{n^{\alpha}}$	$\boldsymbol{k^{\Omega(1)}}$	$\boldsymbol{n^{-c}}$	**Yes**

Multiplicative extractors for samplable distributions are harder to construct. Before this paper, the only construction for min-entropy $k < n/2$ was given by Shaltiel [Sha25a], who achieved $k = n^{\alpha}$ for every constant $\alpha > 0$.

In this paper, we give a multiplicative extractor with the same parameters. Shaltiel's extractor assumes hardness for Σ_i-circuits with $i = 1/\alpha + O(1)$; in contrast, our extractor assumes only hardness for nondeterministic circuits.

Theorem 4 (extracting many bits, under nondeterministic hardness)**.** *If* E *is hard for exponential size nondeterministic circuits then for every constant $\alpha > 0$, and every constant $c > 1$, there is a constant d such that for every sufficiently large n, and every $n^{\alpha} \leq k \leq n$ there is a function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *that is a (k, ε)-multiplicative extractor for distributions samplable by circuits of size n^c, with $\varepsilon = n^{-c}$ and $m = k^{\Omega(1)}$. Furthermore,* Ext *is computable in time n^d.*

As in our previous theorems, this extractor can handle distributions that are samplable by deterministic circuits with postselection by nondeterministic circuits (see Theorem 4.41).

We remark that in the intended application of selecting keys for cryptographic protocols, we typically need $m = \sigma^{\Theta(1)}$ random bits, where σ is the security parameter. As we are only allowed to run in time $\mathrm{poly}(\sigma)$, the source length n must be polynomially bounded in σ. We therefore need an extractor with output length $m = n^{\Omega(1)}$, which requires min-entropy $k = n^{\Omega(1)}$. Theorem 4 matches these parameters.

2 Overview of the Techniques

In this section, we give a detailed informal overview of the main ideas that we use to construct our multiplicative extractors (Section 2.1) and additive extractors (Section 2.2). The later technical sections contain full definitions, statements and proofs and do not build on the informal explanation of this section. The readers can skip to the technical section if they wish.

2.1 Proof Overview for Multiplicative Extractors

We provide a sketch of our multiplicative extractor construction, as guaranteed by Theorem 4, for $n^{0.01}$ entropy under nondeterministic hardness. The same construction directly yields additive extractors with the same parameters as well (see Theorem 5.1). Our construction synthesizes and builds on several ideas across prior works [BGDM23; BSS25; Sha25a].

We follow the high level strategy from [BSS25; Sha25a], of constructing errorless condensers and using them to construct multiplicative extractors.

Definition 2.1 (Errorless Condenser, Informal version of Definition 4.15)**.** *We say* $\mathsf{Cond} : \{0,1\}^n \to \{0,1\}^m$ *is a (k_{in}, k_{out})-errorless condenser for a class $\mathcal{C}$ of sources if for all $\mathbf{X} \in \mathcal{C}$ with $\mathbf{H}_{\infty}(\mathbf{X}) \geq k_{in}$, we have that $\mathbf{H}_{\infty}(\mathsf{Cond}(\mathbf{X})) \geq k_{out}$.*

Unlike traditional goals of condensers where one wants to increase the entropy rate, our goal with an errorless condenser is to make the output length as small as possible, while retaining some non-trivial entropy in the output distribution. For our applications, we require the output length of Cond to be smaller than the min-entropy of the input distribution $\mathbf{X}$. We will construct the following errorless condensers.

Theorem 2.2 (Informal version of Theorem 4.42)**.** *Assume* E *is hard for exponential-size nondeterministic circuits. Then, there exists* $\delta > 0$ *such that for all* $\gamma > 0$*, we can construct an explicit* $(n^\gamma, n^{\gamma\cdot\delta})$*-errorless condenser* $\mathsf{Cond} : \{0,1\}^n \to \{0,1\}^{n^{\gamma/2}}$ *for samplable sources.*

Equipped with this errorless condenser, we will build on ideas from [BGDM23] to construct our multiplicative extractor under nondeterministic hardness. To construct such an errorless condenser, as well as to construct our extractor from errorless condenser, we introduce a new pseudorandom object which we call a *robust errorless block condenser*. This is an errorless condenser for block sources which are subsources of a joint distribution.[8] Formally,

Definition 2.3 (Robust errorless block condenser, Informal version of Definition 4.32)**.** *We say* $\mathsf{rbCond}\colon \{0,1\}^{n_y+n_x} \longrightarrow \{0,1\}^m$ *is a* $(\lambda, k_y, k_y^{\geq}, k_x, k_{out})$*-robust errorless block condenser for class of sources* $(\mathcal{Y}, \mathcal{X})$ *if for every* $(\mathbf{Y}, \mathbf{X}) \in (\mathcal{Y}, \mathcal{X})$ *and every decomposition*

$$(\mathbf{Y}, \mathbf{X}) = \lambda(\mathbf{Y}_{good}, \mathbf{X}_{good}) + (1-\lambda)(\mathbf{Y}_{rest}, \mathbf{X}_{rest})$$

satisfying: (1) $\mathbf{H}_\infty(\mathbf{Y}_{good}) \geq k_y$*; (2)* $\Pr[\mathbf{Y}_{good} = y] \geq 2^{-k_y^{\geq}}$ *for every* $y \in \mathrm{Supp}(\mathbf{Y}_{good})$*; (3)* $\mathbf{H}_\infty(\mathbf{X}) \geq k_x$*, we have*

$$\mathbf{H}_\infty\big(\mathsf{rbCond}(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}})\big) \geq k_{\mathrm{out}}.$$

We will ensure that $k_y^{\geq} < k_x$ so that $(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}})$ form a block source.[9] We were inspired by a similar notion of robust extractors introduced by [BKMO26] to define these objects.

We will construct the following robust errorless block condensers.

Theorem 2.4 (Informal version of Theorem 4.43)**.** *Assume* E *is hard for exponential-size nondeterministic circuits. Then, for all* $\gamma > 0$ *and* $k_x > k_y > n^\gamma$ *there is an explicit polynomial-time computable function*

$$\mathsf{rbCond}\colon \{0,1\}^n \times \{0,1\}^n \longrightarrow \{0,1\}^m$$

that is a

$$\big(\lambda = 2^{-m/4},\ k_y,\ k_y^{\geq} = k_x/2,\ k_x,\ k_{\mathrm{out}} = m/4\big)$$

robust errorless block condenser for samplable sources for any $(\log(n))^2 \leq m \leq k_y^{0.01}$.

To construct such a robust errorless block condenser, we will utilize and build on ideas from [BSS25; BGDM23]. Using this, we will use an iterative win-win strategy, that has featured in several works in extractor theory such as [RSW06; Sha25a], to construct the errorless condenser of Theorem 2.2.

To summarize, our construction of multiplicative extractors has three technical steps:

1. Construct robust errorless block condensers under nondeterministic hardness.
2. Use robust errorless block condensers, combined with win-win arguments to construct errorless condensers.

[8] $(\mathbf{A}, \mathbf{B})$ form a block source if $\mathbf{A}$ has entropy, $\mathbf{B}$ has entropy conditioned on every fixing of $\mathbf{A}$. See Definition 4.19 for a formal definition.

[9] Loosely speaking, conditioning on fixing of $\mathbf{Y}_{\mathrm{good}}$, $\mathbf{X}_{\mathrm{good}}$ will output an element with probability at most $\lambda^{-1} \cdot 2^{-k_x + k_y^{\geq}}$, bringing its conditional entropy to $k_x - k_y^{\geq} - \log(1/\lambda)$.

3. Construct multiplicative extractors using errorless condensers and robust errorless block condensers under nondeterministic hardness.

The first and the third steps share many ideas and we will provide a combined technical overview of them in Section 2.1.1. We will separately provide a technical overview of the second step in Section 2.1.2.

2.1.1 Constructing robust block errorless condenser, and multiplicative extractor from errorless condenser

This part is quite technical and builds on a long line of work [STV01; TV00; BGDM23; BSS25; Sha25a], introducing several new technical insights. To save space, we only give a high level overview of them; for a more detailed exposition, see Section 7.

We sketch how we construct our multiplicative extractor for samplable source $\mathbf{X}$ with sufficient min-entropy, given the output of an errorless condenser $\mathbf{W} = \mathsf{Cond}(\mathbf{X})$, such that $|\mathbf{W}| \ll \mathbf{H}_\infty(\mathbf{X})$.

Warm-up: Assume W has high entropy We first assume that $\mathbf{W}$ has entropy rate 0.999, and sketch how we construct multiplicative extractors in this case. We will follow the same recipe as laid out by Trevisan and Vadhan [TV00]. For this, we require access to a low-error two-source extractor $\mathsf{TExt} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ which has the property that for any two independent sources $\mathbf{A}, \mathbf{B}$ with sufficient entropy, $\mathsf{SD}(\mathsf{TExt}(\mathbf{A}, \mathbf{B}), \mathbf{U}_m) \leq \varepsilon$ where $\varepsilon > 0$ is very small. Also, from our assumption that E is hard for exponential size nondeterministic circuits, we have that there exists a constant $d > 0$ and a function $f : \{0,1\}^{d \log(n)} \to \{0,1\}$ such that f is computable in $\mathrm{poly}(n)$ time, and it is hard to compute for size $n^{\gamma d}$ circuits for some constant $\gamma > 0$. With this, we define a size $2^{O(n)}$ finite field $\mathbb{F}_q$ and obtain a d-variate low degree extension of f, $\widehat{f} : \mathbb{F}_q^d \to \mathbb{F}_q$.[10] We can ensure that $\widehat{f}$ has degree at most $O(n)$.

Equipped with these, we construct our extractor by computing:

$$\mathsf{Ext}(\mathbf{X}) = \mathsf{TExt}(\widehat{f}(\mathbf{W}), \mathbf{X}).$$

To argue about correctness, we proceed by contradiction and assume there exists a heavy outcome $\sigma \in \{0,1\}^m$ such that $\Pr[\mathsf{Ext}(\mathbf{X}) = \sigma] > (1+\varepsilon) \cdot 2^{-m}$. We will show how using this, we can construct a small $\mathrm{poly}(n)$ size nondeterministic circuit C such that $C(x) = \widehat{f}(x)$ for all $x \in \mathbb{F}_q^d$, contradicting the hardness of f. In fact, by standard arguments, it suffices to find C such that $C(x) = \widehat{f}(x)$ for $15/16$ fraction of the inputs $x \in \mathbb{F}_q^d$.

We now sketch how to construct such C. On input x, the circuit C will choose a random point $z \in \mathbb{F}_q^d$ and consider the line $L : \mathbb{F}_q \to \mathbb{F}_q^d$ between x and z. The line will be such that $L(0) = x, L(1) = z$. Then, $\widehat{f} \circ L : \mathbb{F}_q \to \mathbb{F}_q$ will be a univariate degree $O(n)$ polynomial, and $\widehat{f} \circ L(0) = \widehat{f}(x)$, our desired answer. We show how C will reconstruct this polynomial $\widehat{f} \circ L$. To do so, C will nondeterministically guess a degree $O(n)$ univariate polynomial $p : \mathbb{F}_q \to \mathbb{F}_q$, and devise a test such that the following holds: 1) Completeness: If p equals $\widehat{f} \circ L$, then the test accepts. 2) Soundness: If p does not equal $\widehat{f} \circ L$, then the test rejects. In fact, it suffices to devise such a test with relaxed soundness requirement: C accepts some small ($\ll q$) number of polynomials p which do not equal $\widehat{f} \circ L$.

The original test of [TV00] reconstructed small size Σ_5 circuits (which required initial hardness assumption against the same). Later, [BGDM23] constructed such a test so that it results in a small nondeterministic circuit. As a first step, we revisit and simplify the test of [BGDM23]. This simplification is key for obtaining a generalized version of this test, that we will be able to apply to the case that the samplable source has low min-entropy; and this is a key core contribution of our work. For our test, we introduce a quantity defined for $w \in \mathbb{F}_q^d, a \in \mathbb{F}_q$ as:

$$J(w, a) = \Pr[\mathbf{W} = w, \mathsf{TExt}(a, \mathbf{X}) = \sigma].$$

[10] This low degree extension has the property that to efficiently compute $f(x)$, it suffices to efficiently compute $\widehat{f}(\widehat{x})$ for a related $\widehat{x}$.

We easily see that $J(w, \widehat{f}(w)) = \Pr[\mathbf{W} = w, \mathsf{Ext}(\mathbf{X}) = \sigma]$. From this, we obtain that

$$\mathop{\mathbb{E}}_{w \leftarrow \mathbb{F}_q^d}[J(w, \widehat{f}(w))] = \frac{1}{q^d} \sum_{w \in \mathbb{F}_q^d} \Pr[\mathbf{W} = w,\ \mathsf{Ext}(\mathbf{X}) = \sigma] = \frac{1}{q^d} \Pr[\mathsf{Ext}(\mathbf{X}) = \sigma].^{11}$$

With this, we define our test to be that it accepts polynomial p, as a guess for $\widehat{f} \circ L$ if

$$\mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q}[J(L(t), p(t))] \geq \frac{1}{q^d}(1 + \varepsilon/2)2^{-m}.$$

We can implement this test since $\mathbf{X}$ is samplable by an efficient deterministic circuit, and we can use the Goldwasser-Sipser set lower bound protocol [GS86] to execute this test by a small nondeterministic circuit (instead of cycling over all the seeds used to sample $\mathbf{X}$, which will blow up circuit size). Completeness easily follows since a random line is a good sampler; to argue about soundness, we will use properties of a random line, as well as combinatorial list-recovery bounds on Reed Solomon codes [Sud97]. For soundness, we will also use the fact that in a two-source extractor with error at most $\varepsilon \cdot 2^{-m}$, the set $B_w = \{a : \Pr[\mathsf{TExt}(a, \mathbf{X}) = \sigma \mid (\mathbf{W} = w)] > (1 + \varepsilon) \cdot 2^{-m}\}$ is such that $|B_w| < 2^{k_1}$, where k_1 is the entropy requirement of the first source of TExt.

Handling the case when $\mathbf{W}$ has low entropy It is important to note that the parameters of the low degree extension (a.k.a. Reed-Muller code) are not suitable to handle the case where $\mathbf{W}$ has min-entropy smaller than $|\mathbf{W}|/2$. As our aforementioned condenser does not achieve such output min-entropy, we will have to modify the construction and test to handle the case where $\mathbf{W}$ has low min-entropy.

Recall that our test was over the uniform distribution over sampling t. This is the reason why we required $\mathbf{W}$ to have very high entropy (entropy rate 0.99). Following an approach from [BSS25], we modify our construction as follows to handle low entropy case. We use a strong seeded extractor SExt (see Definition 4.17) with $\mathbf{W}$ as a source to obtain a list of $D = \mathrm{poly}(n)$ many sources $\mathsf{SExt}(\mathbf{W}, 1) = \mathbf{W}_1, \ldots, \mathsf{SExt}(\mathbf{W}, D) = \mathbf{W}_D$ such that there exists some i^* (dependent on $\mathbf{W}$) for which $\mathbf{W}_{i^*}$ is statistically close to the uniform distribution. With this, our construction is:

$$\mathsf{Ext}(\mathbf{X}) = \mathsf{TExt}(\widehat{f}(\mathbf{W}_1) \,\|\, \cdots \,\|\, \widehat{f}(\mathbf{W}_D), \mathbf{X}).$$

Observe that the random variable $(\mathbf{W}_1, \ldots, \mathbf{W}_D)$ itself may not have high entropy, and all we need to use is that $\mathbf{W}_{i^*}$ is statistically close to the uniform distribution.[12] When attempting to apply all the pseudorandom properties from above to $\widehat{f}(\mathbf{W}_{i^*})$, we run into a big problem: we also need to recover values of $\widehat{f}(\mathbf{W}_i)$ for all $i \neq i^*$, which are not accessible to us. To overcome these, we modify our test by defining a generalized version $J(w, a)$ for $w \in \mathbb{F}_q^d, a \in \mathbb{F}_q$:

$$J(w, a) = \sum_{y : \mathsf{SExt}(y, i^*) = w} \max_{v \in \mathbb{F}_q^{[D] \setminus i^*}} \Pr[\mathbf{W} = y,\ \mathsf{TExt}((v, a), \mathbf{X}) = \sigma].$$

Crucially in this definition, v can depend on y, but not on a sample from $\mathbf{X}$. Using this, we similarly devise an appropriately modified test on $J(L(t), p(t))$ where we nondeterministically guess the values for v; here we do use the fact that the Goldwasser-Sipser protocol can tolerate when the membership function is nondeterministic. We also carefully modify other parts of the test to ensure our reconstruction can indeed succeed.[13] A more detailed overview of the test of [BGDM23], our simplifications and generalizations, and an overview of the analysis appears in Section 7.

[11] We use $w \leftarrow \mathbb{F}_q^d$ to denote sampling w uniformly from $\mathbb{F}_q^d$.

[12] We also crucially rely on the fact that conditioned on all fixings of $(\mathbf{W}_1, \ldots, \mathbf{W}_d)$, $\mathbf{X}$ still has entropy. This is because $(\mathbf{W}_1, \ldots, \mathbf{W}_D)$ are a deterministic function of $\mathbf{W}$, and $\mathbf{W}$ has length smaller than the entropy in $\mathbf{X}$.

[13] For this step, we require low error two source extractors with very low entropy requirements. While we have constructions of two-source extractors which require low entropy, they do not attain low enough error of less than 2^{-m}, for our applications as multiplicative extractors. To overcome this, we use low error two-source condensers [BCDT19], and low error two source extractors with structural requirements [Li15].

Constructing robust errorless block condenser To extend this to constructing robust errorless block condensers, we will use a similar construction and analogously modify the analysis. The primary challenge here is that rbCond needs to provide guarantees for when the input distribution is the block source $(\mathbf{Y}_{\text{good}}, \mathbf{X}_{\text{good}})$, a distribution that may not even be samplable since all we are guaranteed is that it is a subsource of the samplable source $(\mathbf{Y}, \mathbf{X})$. We overcome this with careful conditioning arguments.

2.1.2 Constructing errorless condenser from robust block errorless condenser

We here sketch how, using a robust block errorless condenser rbCond from Theorem 2.4, we can construct errorless condensers as specified in Theorem 2.2, with output length smaller than the input entropy.

Warm-up: Errorless condenser for entropy $n^{0.51}$ We will build up towards our construction of errorless condensers for entropy $n^{0.01}$ samplable sources by first constructing an errorless condenser for samplable sources with entropy $n^{0.51}$. Given such a source $\mathbf{X} \sim \{0,1\}^n$, we will chop $\mathbf{X}$ into $\ell = n^{0.5}$ parts of size $n/\ell = n^{0.5}$ each, to obtain $\mathbf{X}_1, \ldots, \mathbf{X}_\ell$. With this, we define

$$\mathsf{Cond}(\mathbf{X}) = \mathsf{rbCond}(\mathbf{X}_1, \mathbf{X}) \parallel \mathsf{rbCond}(\mathbf{X}_{[1,2]}, \mathbf{X}) \parallel \cdots \parallel \mathsf{rbCond}(\mathbf{X}_{[1,\ell]}, \mathbf{X}).$$

We will set the output length of rbCond such that the total output length of our errorless condenser is $\ell \cdot n^{0.001} = n^{0.501}$, which is much smaller than the input entropy of $n^{0.51}$, as desired. The remaining task is for us to show that $\mathbf{H}_\infty(\mathsf{Cond}(\mathbf{X})) \geq k_{\text{out}} = n^\alpha$ for some constant $\alpha > 0$.

To do that, we use a classical technique, by [SSZ98; Ta-96], which proved properties of the source obtained by dividing $\mathbf{X}$ into these ℓ blocks. Informally, it states that $\mathbf{X}$ can be expressed as a convex combination of sources $\mathbf{X} = \sum_{i=0}^{\ell} \lambda_i \mathbf{X}^{(i)}$ with the following properties, where $\varepsilon = 2^{-m/16}$:

- The sets $\text{Supp}(\mathbf{X}^{(0)}), \ldots, \text{Supp}(\mathbf{X}^{(\ell)})$ are disjoint and form a partition of the support of $\mathbf{X}$.
- For $i > 0$, either $\lambda_i > \varepsilon/2n$, or $\lambda_i = 0$. Also, $\lambda_0 \leq \varepsilon$.
- For all $i > 0$, $\mathbf{H}_\infty\left(\mathbf{X}^{(i)}_{[1,i]}\right) \geq k/5$.
- For all $i > 0$, and $x_i \in \text{Supp}\left(\mathbf{X}^{(i)}_{[1,i]}\right)$, we have that $\Pr\left[\mathbf{X}^{(i)}_{[1,i]} = x_i\right] \geq 2^{-k/2}$.
- For all $i \geq 1$, it holds that $\mathbf{X}^{(i)}_{[1,i]} \coloneqq \mathbf{X}_{[1,i]} \mid (\mathbf{X} \in \text{Supp}(\mathbf{X}^{(i)}))$.

In words, λ_0 is the "error" part of the decomposition, and $\mathbf{X}^{(i)}$ is such that the prefix block $\mathbf{X}^{(i)}_{[1,i]}$ has min-entropy, and there is a lower bound on the output probabilities. See Lemma 6.1 for a formal statement, and a self-contained proof of this decomposition statement.

Using this decomposition lemma, we argue about the output entropy of $\mathsf{Cond}(\mathbf{X})$ as follows. We will show that for all $i > 0$, $\mathbf{H}_\infty(\mathsf{Cond}(\mathbf{X}) \mid (\mathbf{X} \in \text{Supp}(\mathbf{X}^{(i)}))) \geq k_{\text{out}} + 1$. This suffices to prove the desired claim since we can "absorb" the error λ_0 by using the fact that $\lambda_0 \leq \varepsilon = 2^{-m/16} \leq 2^{-k_{\text{out}}-1}$ (as $k_{\text{out}} \leq m/32$), bringing our output min-entropy to k_{out}.

To prove that $\mathbf{H}_\infty(\mathsf{Cond}(\mathbf{X}) \mid (\mathbf{X} \in \text{Supp}(\mathbf{X}^{(i)}))) \geq k_{\text{out}} + 1$, we will show that the i^{th} output block has the desired entropy guarantees, i.e., that $\mathbf{H}_\infty\left(\mathsf{rbCond}\left(\mathbf{X}_{[1,i]}, \mathbf{X}\right) \mid \left(\mathbf{X} \in \text{Supp}\left(\mathbf{X}^{(i)}\right)\right)\right) \geq k_{\text{out}} + 1$. By the last property above, we have that $\mathbf{X}^{(i)}_{[1,i]} \coloneqq \mathbf{X}_{[1,i]} \mid (\mathbf{X} \in \text{Supp}(\mathbf{X}^{(i)}))$. Moreover, we have lower bounds on min-entropy, lower bounds on output probability on each outcome in the support of $\mathbf{X}^{(i)}_{[1,i]}$, and lower bound that $\lambda_i \geq \varepsilon/2n \geq 2^{-m/4}$. Therefore, we satisfy all of the requirements to apply rbCond and can indeed conclude that $\mathbf{H}_\infty(\mathsf{Cond}(\mathbf{X}) \mid (\mathbf{X} \in \text{Supp}(\mathbf{X}^{(i)}))) \geq k_{\text{out}} + 1$, as desired.

Generalizing to obtain mild robust condenser For our low-entropy condenser construction, we will need to generalize the condenser from above. First, we make the condenser itself robust, i.e., for any samplable source $\mathbf{A} = p \cdot \mathbf{A}_{\text{good}} + (1-p) \cdot \mathbf{A}_{\text{rest}}$ with $\mathbf{H}_\infty(\mathbf{A}_{\text{good}}) > n^{0.51}$, we guarantee that $\mathbf{H}_\infty(\mathsf{Cond}(\mathbf{A}_{\text{good}})) \geq n^\alpha$ for some constant $\alpha > 0$ and where $p = 2^{-m/20}$.[14] Second, we generalize the parameters in the construction from above, to obtain a mildly shrinking robust errorless condenser. Concretely, for any constant $\gamma_{\text{in}} > 0$ we set $\ell = n^{1-3\gamma_{\text{in}}/4}$, and output length of the robust errorless block condenser to $n^{\gamma_{\text{in}}/4}$ in the construction above, to obtain a $(n^{\gamma_{\text{in}}}, n^{0.01\gamma_{\text{in}}})$ robust errorless condenser $\mathsf{Cond} : \{0,1\}^n \to \{0,1\}^{n^{1-\gamma_{\text{in}}/2}}$. See Lemma 6.3 for the general statement.

Constructing $n^{0.01}$ entropy errorless condenser We finally sketch how to construct errorless condensers for entropy $n^{0.01}$ sources, with output length much smaller than $n^{0.01}$. We will require a mild robust condenser Cond from above, as well as a robust block errorless condenser rbCond. We construct our desired errorless condenser, which has output length smaller than $n^{0.01}$ as follows.

Our high level strategy will be to use an iterative construction and a win-win argument, that has featured in several works in extractor theory including [NT99; RSW06; Sha25a]. Given a source $\mathbf{X}$ with $\mathbf{H}_\infty(\mathbf{X}) = n^{0.01}$, we first set parameters of Cond appropriately to obtain $\mathbf{Y}_1 = \mathsf{Cond}(\mathbf{X})$ with $\mathbf{H}_\infty(\mathbf{Y}_1) \geq n^\alpha$ for some constant $\alpha > 0$ with $|\mathbf{Y}_1| = n^{0.995}$.

Now, consider the source $(\mathbf{Y}_1, \mathbf{X})$. Ideally, we would like to show that this is a block source. We know that $\mathbf{Y}_1$ has entropy, but it is not clear whether $\mathbf{X}$ has entropy conditioned on $\mathbf{Y}_1$. This is because the length of $\mathbf{Y}_1$ is much larger than the entropy guarantees on $\mathbf{X}$, and hence it may "steal" all the entropy from $\mathbf{X}$. To overcome this, we will therefore split $(\mathbf{Y}_1, \mathbf{X})$ into a convex combination of sources, one of which will form a block source and the other, we will directly be able to condense from (note that the components need not be efficiently samplable). To obtain this convex combination, we partition $\mathrm{Supp}(\mathbf{Y}_1)$ into medium and light weight sets $M_1, L_1 \subset \mathrm{Supp}(\mathbf{Y}_1)$ and show that on the medium part, the resultant sources will form a block source where we will apply rbCond, while the light part will have non-trivial entropy guarantees, and we will be able to recursively condense by applying Cond.

Formally, we let $M_1 = \{y_1 \in \mathrm{Supp}(\mathbf{Y}_1) : \Pr[\mathbf{Y}_1 = y_1] \geq 2^{-n^{0.005}}\}$, and let $L_1 = \mathrm{Supp}(\mathbf{Y}_1) \setminus M_1$. Without loss of generality, we assume that $\Pr[\mathbf{Y}_1 \in M_1]$ is large enough to satisfy the requirements of rbCond (if $\mathbf{Y}_1$ outputs an element from M_1 with very low probability, then we will "pay" an additive factor to the final output probability, which will have next to no effect on the output min-entropy). With this, we carefully check that we satisfy all the requirements to apply our robust block errorless condenser rbCond to $(\mathbf{Y}_1, \mathbf{X})$ to conclude that $\mathbf{Z}_1 = \mathsf{rbCond}(\mathbf{Y}_1, \mathbf{X}) \mid (\mathbf{Y}_1 \in M_1)$ is such that $\mathbf{H}_\infty(\mathbf{Z}_1) > n^\beta$ for some constant $\beta > 0$. By definition of L_1, neglecting the minor factor that depends on $\Pr[\mathbf{Y}_1 \in L_1]$, we have that $\mathbf{H}_\infty(\mathbf{Y}_1 \mid (\mathbf{Y}_1 \in L_1)) \geq n^{0.005}$, and we recurse on this source.

We then repeat this process: We let $\mathbf{Y}_2 = \mathsf{Cond}(\mathbf{Y}_1)$ with $|\mathbf{Y}_2| = n^{0.99}$; we split the support of $\mathbf{Y}_2 \mid (\mathbf{Y}_1 \in L_1)$ into M_2, L_2; we apply appropriate rbCond to obtain $\mathbf{Z}_2 = \mathsf{rbCond}(\mathbf{Y}_2, \mathbf{X}) \mid (\mathbf{Y}_1 \in L_1, \mathbf{Y}_2 \in M_2)$ with $\mathbf{H}_\infty(\mathbf{Z}_2) \geq n^\beta$; and have that $\mathbf{H}_\infty(\mathbf{Y}_2 \mid (\mathbf{Y}_1 \in L_1, \mathbf{Y}_2 \in L_2)) \geq n^{0.005}$. We continue repeating this until we obtain a source $\mathbf{Y}_t$ with $|\mathbf{Y}_t| < n^{0.005}$, and then we directly apply robust block errorless condenser to the resultant source. We obtain that each of $\mathbf{Z}_1, \ldots, \mathbf{Z}_t$ has small length, and entropy at least n^β. Moreover, our overall output distribution can be written as a convex combination of t many sources, where the source j will receive its entropy from the block $\mathbf{Z}_j$. Note that $t = O(1)$, and so any of the intermediate sources are still samplable by $\mathrm{poly}(n)$ size circuits. See Section 6.3 for complete details.

2.2 Proof Overview for Additive Extractors

We provide a proof overview of the additive extractors guaranteed by Theorem 1. We will also give the proof ideas for Theorem 3, in particular, we will introduce the high level ideas to extract many bits

[14] Of course, $\mathbf{A}_{\text{good}}$ need not be samplable at all.

from polylogarithmic entropy, under nondeterministic hardness, with the caveat that the extractor runs in quasipolynomial time. Detailed discussion on how to get efficient extractors under a stronger hardness assumption is given in Section 15.

We state the informal versions of both results below.

Theorem 2.5 (One-bit, informal)**.** *If* E *is hard for exponential size nondeterministic circuits then for every constant* $\varepsilon > 0$*, and entropy* $k = O(\log n)$*, there is an efficiently computable function* $\mathsf{Ext}_{\mathrm{one}} : \{0,1\}^n \to \{0,1\}$ *that is a* (k, ε)*-extractor for distributions samplable by polynomial-size circuits.*

Theorem 2.6 (Many-bit, informal)**.** *If* E *is hard for exponential-size nondeterministic circuits, then for all constants* $b > 0$*, there is a constant* a*, such that for* $k = \operatorname{polylog} n$*, there is a* (k, n^{-b})*-extractor* $\mathsf{Ext}_{\mathrm{many}} : \{0,1\}^n \to \{0,1\}^{k^a}$ *for distributions samplable by polynomial-size circuits. Moreover,* $\mathsf{Ext}_{\mathrm{many}}$ *is computable in time* $2^{O(\log^2 n/\log k)}$.

Our strategy follows a familiar paradigm from independent-source extraction: first reduce to a *non-oblivious bit-fixing source*, in which the good coordinates are approximately t-wise independent while the remaining coordinates may depend arbitrarily on them, and then extract using a suitable resilient function. In the two-source construction of Chattopadhyay and Zuckerman [CZ19], a seeded non-malleable extractor produces a table in which a large set of rows has approximately t-wise independent entries; the second source is then used to sample a smaller table, from which a resilient function extracts a nearly uniform bit. Ben-Aroya, Doron, and Ta-Shma [BDT17] refine the sampling step so that, after combining small groups of outputs, Majority suffices.

We pursue an analogous strategy using only a single samplable source $\mathbf{X}$. Exploiting samplability and our hardness assumption, we first transform $\mathbf{X}$ into many correlated rows, most of which have almost full min-entropy. Crucially, this guarantee is *errorless*: each *good* row itself has high min-entropy, rather than merely being statistically close to a high-entropy source. We turn this entropy into approximate independence using a *local correlation breaker*, constructed under the same hardness assumption by adapting a template for *seedless non-malleable extractors for polynomial-size tampering* [BDL22; BD25]. The correlation breaker is used to process each indexed row separately, ensuring that any t good output rows are jointly close to independent uniform strings. Thus, both the high-entropy rows and the randomness used to break their correlations come from $\mathbf{X}$ itself.

For one-bit extraction, we arrange that sufficiently few output bits are bad and apply Majority. For many-bit extraction, inspired by the techniques in [Li13a; Li13b; Li15], we use the approximately independent outputs to select a small group containing a good row via Feige's lightest bin protocol [Fei99], and then apply correlation breaking again to extract a longer string. Here the source itself drives the selection, so we must preserve enough entropy and account for the correlations that selection introduces.

We first introduce errorless seeded condensing and local correlation breaking, then sketch both constructions, and finally explain how to construct the local correlation breaker.

2.2.1 Errorless seeded condensing to obtain many good rows.

In this subsection, we state the common starting point of the one-bit and the many-bit constructions: converting the source into polynomially many rows, *most* of which have almost full entropy.

Lemma 2.7. *Assume* E *is hard for exponential-size nondeterministic circuits. For* $k \geq C \log n$*, we construct a fixed family of* $R = \operatorname{poly}(n)$ *hashes* $h_i : \{0,1\}^n \to \{0,1\}^m$*, where* $m = k - \log k + O(1)$*. For every samplable source* $\mathbf{X}$ *over* $\{0,1\}^n$ *with* $\mathbf{H}_\infty(\mathbf{X}) \geq k$*, let* $\mathbf{X}_i = h_i(\mathbf{X})$*, then at least* $1 - R^{-\eta}$ *fraction of rows* $\mathbf{X}_i$ *satisfy* $\mathbf{H}_\infty(\mathbf{X}_i) \geq m - O(1)$*, for a constant* $\eta > 0$*.*

Note that the rows remain jointly samplable, and the entropy guarantee holds for each good row itself, without error.

Proof sketch. We obtain our list of condensers in three steps.

Random hashing gives errorless condensation. Fix a source $\mathbf{X}$ with $\mathbf{H}_\infty(\mathbf{X}) \geq k$. Choose a random k-wise independent hash $h : \{0,1\}^n \to \{0,1\}^m$, where $m = k - \log k + \Delta$ for a sufficiently large constant Δ. A collision bound shows that, except with probability at most n^{-2} over h, we have $\mathbf{H}_\infty(h(\mathbf{X})) \geq k - \log k = m - \Delta$. This guarantee is errorless: it bounds the largest probability in $h(\mathbf{X})$ itself. Thus we have a large family of efficiently computable candidates, most of which are good for each fixed source.

Samplability lets us test the candidates. Write $\mathbf{X} = \mathsf{Samp}(\mathbf{U}_r)$ for a sampling circuit Samp of the prescribed size. If $\mathbf{H}_\infty(h(\mathbf{X})) < k - \log k$, then there must be some very heavy output, forcing an unusually large probability that k distinct samples all hash to the same value. This collision event can be checked efficiently on independent seeds of Samp. The Goldwasser–Sipser counting protocol [GS86], followed by amplification and nonuniform fixing of randomness, therefore gives a polynomial-size nondeterministic test on hash descriptions. It accepts every h violating the entropy bound. The collision estimate also shows that it accepts at most an n^{-2} fraction of random hashes.

A PRG reduces the family to polynomial size. Under our hardness assumption, the PRG of [SU05] fools all these tests with error n^{-2} using $O(\log n)$ seed bits. Enumerating its seeds gives $R = \mathrm{poly}(n)$ hashes, at most $2n^{-2}R \leq R^{1-\eta}$ of which violate the entropy bound for any fixed source, for a sufficiently small constant $\eta > 0$. The same list works for every source because the generator fools every test of the prescribed size. □

2.2.2 Local correlation breaking

In this subsection, we state the guarantees of our local correlation breaker and explain its role in our extractor constructions. We will sketch its construction and analysis in Section 2.2.5 and give full details and proofs in Section 12.3.

After obtaining the rows from Section 2.2.1, we want to extract from a small collection containing at least one good row. A natural approach is to apply a local map to each row and XOR the outputs. This works if one output is almost uniform even given the others. Our *local correlation breaker* provides precisely this guarantee for sufficiently small collections of distinctly indexed rows, despite their correlations.

Lemma 2.8 (Local correlation breaker, informal)**.** *Assume* E *is hard for exponential-size nondeterministic circuits. For appropriate parameters, there exists an efficiently computable function* $\mathsf{CB} : \{0,1\}^L \to \{0,1\}^q$ *with the following guarantee. For every jointly samplable tuple* $(\mathbf{V}_1, \ldots, \mathbf{V}_D)$ *whose inputs contain distinct fixed indices, and every* $g \in [D]$ *such that* $\mathbf{V}_g$ *has almost full min-entropy,*

$$\big(\mathsf{CB}(\mathbf{V}_g), \mathsf{CB}(\mathbf{V}_{-g})\big) \approx_\varepsilon \big(U_q, \mathsf{CB}(\mathbf{V}_{-g})\big).$$

Here $\mathsf{CB}(\mathbf{V}_{-g})$ *denotes the tuple of all other outputs, and* U_q *is uniform and independent of that tuple. The inputs may be arbitrarily correlated within the sampling model, and only* $\mathbf{V}_g$ *is required to have high entropy.*

In our constructions, we will instantiate the local correlation breaker with different parameters including the output length q, the number D of rows it can handle, the entropy deficiency of an "almost full entropy" row, etc. We now describe our extractor constructions using the above common tools.

2.2.3 The one-bit construction

To prove Theorem 2.5, we turn the rows from Lemma 2.7 into an NOBF source with almost t-wise independent good coordinates: outside a small fixed set of bad coordinates, every t bits are jointly close to uniform. Here t is a sufficiently large constant depending on the desired error; the majority function will extract the final bit.

Grouping. The seeded errorless condenser gives R rows $\mathbf{X}_1, \ldots, \mathbf{X}_R$, at most $R^{1-\eta}$ of which are bad. Following [BDT17], we use Zuckerman's disperser [Zuc07] to form $M = \Theta(R^8)$ groups of constant size d_0. By the guarantee of the disperser, any $M^{1/4}$ groups collectively touch more than $R^{1-\eta}$ original rows, so fewer than $M^{1/4}$ groups can contain only bad rows. We call a group good if it contains at least one good row. The grouping is fixed independently of the source. Since each row $\mathbf{X}_i$ might appear in multiple groups, we give each row copy a distinct index specifying its group u and position in the group a, and denote it by $\mathbf{V}_{u,a}$.

Local correlation breaking. Instantiate CB for $D = d_0 t$ indexed rows and $q = 1$ output bit, with sufficiently small inverse-polynomial error, and for each group u define

$$\mathbf{Y}_u = \bigoplus_{a=1}^{d_0} \mathsf{CB}(\mathbf{V}_{u,a}).$$

We claim that $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_M)$ has the required NOBF guarantee that except for fewer than $M^{1/4}$ bad coordinates, all the good coordinates are almost t-wise independent. To see this, fix any t good groups and choose one good row from each. Since each group contains d_0 rows, they supply $d_0 t$ indexed inputs; note that copies of the same original row are distinct because we gave them distinct indices. Correlation breaking makes each chosen output nearly uniform given all the others in this collection. XORing within each group and applying a hybrid argument across groups shows that the t group bits are jointly almost uniform.

Extracting from the NOBF source. Finally, output

$$\mathsf{Ext}_{\text{one}}(\mathbf{X}) = \mathsf{Maj}(\mathbf{Y}).$$

By Viola's majority bound [Vio14], in the form stated in [CGL21, Lemma 4.28], majority achieves any desired constant error while tolerating a sufficiently small constant multiple of $\sqrt{M}$ bad coordinates, for constant t and sufficiently small almost-independence error. Our bound of $M^{1/4}$ bad coordinates therefore suffices.

2.2.4 The many-bit construction

We give the construction of $\mathsf{Ext}_{\text{many}}$ in Theorem 2.6. For simplicity, let's fix a sufficiently small constant $a > 0$ and given samplable source $\mathbf{X}$ over $\{0,1\}^n$ of min-entropy k, we would like to extract k^a bits from it.

Extracting from a small collection. Suppose we have a correlation breaker $\mathsf{CB}_{\text{ext}} : \{0,1\}^L \to \{0,1\}^q$, with $L = \Theta(k)$, $q = k^a$, and inverse-polynomial error ζ, that handles $D = O(k^a)$ rows. Then we can extract from these rows:

Claim 2.9. *Let $\mathbf{V}_1, \ldots, \mathbf{V}_D$ be jointly samplable, distinctly indexed rows with nondeterministic postselection. If there exists a row $\mathbf{V}_g$ such that $\mathbf{V}_g$ has almost full entropy, then XOR of the correlation breaker outputs on each row extracts $q = k^a$ almost-uniform bits:*

$$\bigoplus_{j=1}^{D} \mathsf{CB}_{\text{ext}}(\mathbf{V}_j) \approx_\zeta \mathbf{U}_q.$$

Indeed, the guarantee of the correlation breaker ensures that the good output is ζ-close to uniform even given all the other outputs so the XOR of them is also close to uniform. Such a map CB_{ext} can be obtained by instantiating Lemma 2.8:

Lemma 2.10. *Assume E is hard for exponential-size nondeterministic circuits. For $k = \text{poly}(\log n)$, $L = \Theta(k)$, there is a map $\mathsf{CB}_{\text{ext}} : \{0,1\}^L \to \{0,1\}^{k^a}$ with any desired inverse-polynomial error ζ. For*

every collection of $D = O(k^a)$ distinctly indexed rows $\mathbf{V}_1, \dots, \mathbf{V}_D$, jointly samplable with nondeterministic postselection, and every good row $\mathbf{V}_g$ with $\mathbf{H}_\infty(\mathbf{V}_g) \geq L - O(\log n)$, we have

$$\big(\mathsf{CB}_{\mathrm{ext}}(\mathbf{V}_g), (\mathsf{CB}_{\mathrm{ext}}(\mathbf{V}_i))_{i \neq g}\big) \approx_\zeta \big(\mathbf{U}_q, (\mathsf{CB}_{\mathrm{ext}}(\mathbf{V}_i))_{i \neq g}\big).$$

Consequently, $\bigoplus_{i=1}^{D} \mathsf{CB}_{\mathrm{ext}}(\mathbf{V}_i) \approx_\zeta \mathbf{U}_q$.

Theorem 2.6 will follow if we have the appropriate $\mathbf{V}_1, \dots, \mathbf{V}_D$ with the following two guarantees (i) $D = O(k^a)$; (ii) at least one row has high entropy.

Selecting a small collection. The errorless seeded condenser gives $R = \mathrm{poly}(n)$ rows $\mathbf{X}_1, \dots, \mathbf{X}_R$ of length $\Theta(k)$, all but $R^{1-\eta}$ of which have constant entropy deficiency, for a constant $\eta > 0$. Following Feige's lightest-bin protocol [Fei99], assign each row a label in $[T]$, where $T \approx R/k^a$, and select the least populated bin. By averaging, its size is always at most $D = R/T = O(k^a)$, which satisfies guarantee (i). It remains to ensure guarantee (ii)—the lightest bin contains at least one good row. To show this, it suffices to bound the probability that every bin contains at least one good row. In an ideal world, independent uniform labels on the good rows would provide this guarantee: each bin receives $\Theta(k^a)$ good rows in expectation, and the probability that there exists a bin that receives no good row is small. However, we don't have independent uniform bin labels—we instead use another correlation breaker to implement enough of the random labeling on the good rows.

Lemma 2.11. *Assume E is hard for exponential-size nondeterministic circuits. Let $T = R/k^a, d = O(\log n/\log k)$. There is a map $\mathsf{CB}_{\mathrm{lab}} : \{0,1\}^L \to [T]$ computable in time $2^{O(\log^2 n/\log k)}$, such that for any d distinctly indexed rows $\mathbf{V}_1, \dots, \mathbf{V}_d$, jointly samplable with nondeterministic postselection, if every row has entropy deficiency $O(\log n)$, then with small (at most inverse polynomial) error δ*

$$\big(\mathsf{CB}_{\mathrm{lab}}(\mathbf{V}_1), \dots, \mathsf{CB}_{\mathrm{lab}}(\mathbf{V}_d)\big) \approx_\delta \mathbf{U}_{[T]^d}.$$

Under assignment $\mathsf{CB}_{\mathrm{lab}}$, any labeling of d good rows acts like independent uniform labeling. In reality, we instantiate $\mathsf{CB}_{\mathrm{lab}}$ with sufficiently small error, and prove using a standard moment bound from [BR94] that every bin contains a good row, except with inverse-polynomial probability.

So far, the lightest-bin rule guarantees we select a small collection of size $O(k^a)$ and the independence property of $\mathsf{CB}_{\mathrm{lab}}$ guarantees a good row in that collection.

Construction: two applications of correlation breaking To recap, let $\mathbf{X}_1, \dots, \mathbf{X}_R$ be the condensed rows from Lemma 2.7. Our plan is to use $\mathsf{CB}_{\mathrm{lab}}$ and the lightest-bin rule to select at most $D = O(k^a)$ rows, including a good row with high probability, and then apply $\mathsf{CB}_{\mathrm{ext}}$ and XOR its outputs to extract k^a bits.

To implement the two applications, split each row as $\mathbf{X}_i = \mathbf{A}_i \parallel \mathbf{B}_i$ into nearly equal halves, with $|\mathbf{B}_i| = m_1$. Assign row i the bin label $\mathbf{Y}_i = \mathsf{CB}_{\mathrm{lab}}(i \parallel \mathbf{A}_i) \in [T]$, and let $i_1 < \cdots < i_s$ be the indices selected by the lightest bin. Retain their second halves and index them by their positions in the bin: set $\mathbf{W}_j = j \parallel \mathbf{B}_{i_j}$ for $j \leq s$, and pad with $\mathbf{W}_j = j \parallel 0^{m_1}$ for $s < j \leq D$. For example, suppose $D = 6$ and the lightest bin contains exactly rows $\mathbf{X}_4, \mathbf{X}_9, \mathbf{X}_{17}, \mathbf{X}_{23}$. Then $s = 4$ and $(i_1, i_2, i_3, i_4) = (4, 9, 17, 23)$, so we construct

$$\begin{aligned} \mathbf{W}_1 &= 1 \parallel \mathbf{B}_4, & \mathbf{W}_2 &= 2 \parallel \mathbf{B}_9, \\ \mathbf{W}_3 &= 3 \parallel \mathbf{B}_{17}, & \mathbf{W}_4 &= 4 \parallel \mathbf{B}_{23}, \\ \mathbf{W}_5 &= 5 \parallel 0^{m_1}, & \mathbf{W}_6 &= 6 \parallel 0^{m_1}. \end{aligned}$$

Thus, $\mathbf{W}_3$ contains the second half of original row 17, prefixed by its selected position 3. The last two rows are padding, ensuring that the tuple always contains exactly D rows. We also make sure all position indices are encoded using the same number of bits. Finally, output

$$\mathsf{Ext}_{\mathrm{many}}(\mathbf{X}) = \bigoplus_{j=1}^{D} \mathsf{CB}_{\mathrm{ext}}(\mathbf{W}_j).$$

Preserving entropy after selection. The construction as stated above has one major problem: the half of the input used for bin selection and the other half used for extraction are not independent! We split the rows into two halves, and while both halves of a good row have constant deficiency—their indexed versions have $O(\log n)$ deficiency—they remain highly correlated. Bin selection using the first halves can therefore bias the second halves. We must show that enough entropy remains after the first application of the correlation breaker. Moreover, the position of a good selected row may depend on the input. We handle both issues by conditioning on the first good row in the selected bin. For each good original row $\mathbf{X}_i$ and position j in the selected bin, let $E_{i,j}$ be the event that $\mathbf{X}_i$ is the first good row in the ordered selected list and occupies position j. On this event, we have $\mathbf{W}_j = j \parallel \mathbf{B}_i$, while the other selected rows may vary. We emphasize that these events are for the purpose of analysis only, the extractor need *not* identify the good row. In the indexing example above, if rows 4 and 9 are bad and row 17 is good, then $E_{17,3}$ occurs: the first good row selected is $\mathbf{X}_{17}$, and it is on position 3 in the bin.

These sub-events partition the event that the lightest bin contains a good row. There are at most RD such events. Fix an inverse polynomial $\gamma > 0$, and discard those events with probability below $\gamma/(RD)$; their total probability is at most γ.

Let $h = \log D$ be the length of the position index, so each indexed row has length $L = m_1 + h$. Every remaining event satisfies $\Pr[E_{i,j}] \geq \gamma/(RD)$, and hence, for every b,

$$\Pr[\mathbf{B}_i = b \mid E_{i,j}] \leq \frac{\Pr[\mathbf{B}_i = b]}{\Pr[E_{i,j}]} \leq \frac{RD}{\gamma} \Pr[\mathbf{B}_i = b].$$

Thus conditioning loses at most $\log(RD/\gamma)$ bits of min-entropy. On $E_{i,j}$, we have $\mathbf{W}_j = j \parallel \mathbf{B}_i$. Appending the fixed index j leaves min-entropy unchanged but increases the row length by h. Since $\mathbf{B}_i$ initially has constant deficiency,

$$\begin{aligned}
\mathbf{H}_\infty(\mathbf{W}_j \mid E_{i,j}) &= \mathbf{H}_\infty(\mathbf{B}_i \mid E_{i,j}) \\
&\geq m_1 - O(1) - \log(RD/\gamma) \\
&= L - \big(h + O(1) + \log(RD/\gamma)\big) \\
&= L - O(\log n),
\end{aligned}$$

where the last equality uses that R, D are polynomial in n and γ is inverse-polynomial. Therefore, the indexed row has deficiency $O(\log n)$ even after conditioning. This is the deficiency that CB_{ext} is instantiated to tolerate, so we are allowed to apply the correlation breaker after bin selection and get close to uniform bits. This concludes the proof for Theorem 2.6. The final error of the extractor will accumulate γ above, the error of CB_{lab}, and the error of CB_{ext}.

Running time The total running time is $2^{O(\log^2 n/\log k)}$, which is polynomial when $k = n^{\Omega(1)}$, and quasipolynomial when $k = \text{poly}(\log n)$. In Section 15, we obtain polynomial running time for polylogarithmic entropy under the stronger hardness assumption that E is hard for exponential-size Σ_2 circuits. There, Σ_2-hardness is used to get a two-sided multiplicative (or relative-error following [AASY16]) guarantee on the correlation breaker which can then be exploited via refined analysis of the lightest bin protocol via a (truncated) inclusion–exclusion argument.

2.2.5 Construction of the correlation breaker

In the rest of this section, we build the local correlation breaker used in our constructions. Recall that a local correlation breaker makes a good row's output nearly uniform and independent of the other outputs: for a tuple $\mathbf{V}_1, \ldots, \mathbf{V}_D$ and any good row $\mathbf{V}_g$,

$$\big(\mathsf{CB}(\mathbf{V}_g), (\mathsf{CB}(\mathbf{V}_i))_{i \neq g}\big) \approx \big(\mathbf{U}, (\mathsf{CB}(\mathbf{V}_i))_{i \neq g}\big)$$

where $\mathbf{U}$ is uniform and independent of the companion outputs. For *additive decorrelation*, we obtain a statistical-distance guarantee between LHS and RHS above. The additive guarantee suffices for our one-bit extractor and quasipolynomial-time many-bit extractors under the assumption that E is hard for exponential-size nondeterministic circuits. Here, we give the construction with additive guarantee, following the template of Ball, Dachman-Soled, and Loss [BDL22] and Ball and Dachman-Soled [BD25]. In Section 12, we additionally provide more details for *multiplicative decorrelation*.

The construction. At entropy threshold k, the errorless seeded condenser of Lemma 2.7 gives $R = \mathrm{poly}(n)$ rows. After indexing and/or splitting, the row length remains $L = \Theta(k)$. We construct $\mathsf{CB} : \{0,1\}^L \to \{0,1\}^q$ with output length q. Its guarantee concerns a tuple of D indexed rows $\mathbf{V}_1, \ldots, \mathbf{V}_D \in \{0,1\}^L$ from this collection, jointly samplable with nondeterministic postselection by circuits of total size at most S. The indices are fixed and distinct. At least one row has nearly full min-entropy; the others may have arbitrarily low entropy. We use three ingredients:

- A *preliminary extractor* $\mathsf{Ext}_{\mathrm{pre}} : \{0,1\}^L \to \{0,1\}^\ell$ for postselecting sources extracts a short nearly uniform seed of length $\ell = O(\log S)$ from a good row. Short seeds limit the entropy lost when we fix all D seeds.
- A *seed-extending PRG* $G : \{0,1\}^\ell \to \{0,1\}^L$, with $\ell = O(\log S)$ and $G(\sigma) = \sigma \parallel G'(\sigma)$. We get G by composing the seed-extending PRG of [Sha25b] with the long-stretch PRG of [DS26], with inverse-polynomial error against nondeterministic circuits of size $\mathrm{poly}(S)$.
- A *strong two-source non-malleable extractor* $\mathsf{2NMExt}$ whose output on independent high-entropy inputs stays nearly uniform given its first input, namely the expanded seed, as well as the outputs on t separately tampered pairs. We use different $\mathsf{2NMExt}$ for our two constructions. For the one-bit extractor at logarithmic entropy, we use an extractor based on Li's construction [Li23], supporting constant t and one output bit. For the many-bit extractor at polylogarithmic entropy, we use the construction of Chattopadhyay, Goyal, and Li [CGL20], allowing both t and the output length q to grow as sufficiently small fixed powers of L.

Combining the above, for a complete indexed row v, define

$$\mathsf{CB}(v) = \mathsf{2NMExt}\big(G(\mathsf{Ext}_{\mathrm{pre}}(v)), v\big).$$

Building on [BDL22], Ball and Dachman-Soled [BD25] establish non-malleability under nondeterministic hardness against bounded-complexity tamperings of sources with efficiently computable postselection predicates. Our application requires local decorrelation for distinctly indexed rows that are jointly samplable with *nondeterministic* postselection. Moreover, a companion row need not be efficiently computable from the good row. The local decorrelation viewpoint is already implicit in [BD25]. We instantiate their framework with stronger ingredients and extend the analysis to handle nondeterministic postselection under nondeterministic hardness. We outline the analysis below following the template of [BD25].

The analysis: failed CB implies distinguisher for G The central obstacle in the analysis is that the first input to $\mathsf{2NMExt}$ is a deterministic function of the second. Following [BDL22; BD25], we overcome this by reducing a failure of correlation breaking to a violation of G's pseudorandomness against nondeterministic circuits:

$$\exists g, (\mathsf{CB}(\mathbf{V}_g), (\mathsf{CB}(\mathbf{V}_i))_{i \neq g}) \not\approx (\mathbf{U}_q, (\mathsf{CB}(\mathbf{V}_i))_{i \neq g}) \implies \text{nondeterministic distinguisher } T \text{ for } G$$

Naively, a distinguisher T might try to determine whether $w = \sigma \parallel y$ is generated or uniform by simulating the correlation breaker. It samples rows v_i and computes their seeds $p_i = \mathsf{Ext}_{\mathrm{pre}}(v_i)$. If $p_g \neq \sigma$, it rejects. Otherwise, it treats w as the generator output for the good row, computing $z_g = \mathsf{2NMExt}(w, v_g)$, while computing $z_i = \mathsf{2NMExt}(G(p_i), v_i)$ for each companion. However, computing the generator will exceed T's size budget.

Interactive protocol To avoid this expensive simulation, [BDL22; BD25] followed [AASY16]'s observation that a constant-round private-coin protocol suffices to yield a (nondeterministic) distinguisher: an unbounded prover, Merlin, supplies the evaluations and performs the statistical test for a randomized verifier, Arthur. Since constant-round interactive proofs can be simulated by nondeterministic circuits, i.e. $\mathsf{IP}[O(1)] \subseteq \mathsf{NP/poly}$ [GS86; GL20; BM88; Adl78], we convert this protocol into a polynomial-size nondeterministic circuit T that guesses Merlin's reply. This circuit T is our distinguisher for G.

The protocol is as follows,

Fixed data: The joint sampler and a good index g witnessing the assumed failure of CB.

Common input: $w = \sigma \parallel y$.

(i) **Sample and match the seed.** Arthur privately samples random coins r, computes candidate rows v_i and seeds $p_i = \mathsf{Ext}_{\mathrm{pre}}(v_i)$, and rejects unless $p_g = \sigma$. He sends the companion seeds $(p_i)_{i \neq g}$ to Merlin.

(ii) **Supply generator evaluations.** Merlin returns claimed generator suffixes $(a_i)_{i \neq g}$. Arthur computes $z_i = \mathsf{2NMExt}(p_i \parallel a_i, v_i)$, for each $i \neq g$.

(iii) **Guess the hidden bit.** Arthur chooses a uniform private bit c. If $c = 1$, he sets $z = \mathsf{2NMExt}(w, v_g)$; otherwise, he samples $z \leftarrow \mathbf{U}_q$ independently. He sends $(z, (z_i)_{i \neq g})$ to Merlin, who returns a guess $\widehat{c}$.

(iv) **Certify postselection.** Only after receiving $\widehat{c}$ does Arthur reveal r and request a witness certifying successful postselection. He accepts exactly when the witness is valid and $\widehat{c} = c$.

Figure 1: The private-coin protocol used in the reduction.

We show that if correlation breaking fails at the good row g, then honest generator evaluations give Merlin a noticeable guessing advantage on a noticeable fraction of generated inputs $w = G(\sigma)$. On uniform inputs $w \leftarrow \mathbf{U}_L$, strong non-malleability bounds every prover's average guessing advantage, even when Merlin supplies arbitrary claimed evaluations. In particular, conditioning on the "seed" σ and Arthur's message to Merlin (which is short) induces independent high min-entropy distributions and a split-state tampering strategy: Merlin "tampers" the LHS, (σ, y), and Arthur "tampers" the RHS, v_g, by sampling the v_i's conditioned on v_g. This latter sampling is not efficiently computable, but we observe that by the strong multiplicative guarantee of Goldwasser–Sipser [GS86; GL20], it suffices to construct a protocol where Arthur simply gives up whenever the conditional sampling fails. After suitable amplification, this completeness–soundness gap yields a nondeterministic circuit T that distinguishes $G(\mathbf{U}_\ell)$ from $\mathbf{U}_L$, contradicting the security of G.

2.3 Organization

We provide a statement on AI use in Section 3. We will provide preliminaries for the technical sections in Section 4. The technical sections are divided into a multiplicative extractor part (Part I) and an additive extractor part (Part II).[15] Each of these will contain several sections that will present the full details of the constructions of these extractors.

[15] Our additive extractor for n^α entropy samplable sources, for any $\alpha > 0$, will follow directly from our multiplicative extractor, and is presented there.

3 AI Use Statement

The work on this project started in 2025, and some (weaker versions) of the current results were already obtained by the authors without any use of AI.

Once AI became more useful, it was used to assist with proofreading, developing proofs of technical lemmas, and in helping to write them.

The use of AI was also specifically helpful in obtaining some improvements over the preliminary results that we had. More specifically:

- A long conversation with AI (in which we discussed our improved [BGDM23] test, that we already had at that point) helped in obtaining an analysis of the improved test that applies to multiplicative extractors with large output length. This is part of the proof of Theorem 4 and specifically affected the proofs in Section 9 and Section 10. Before this improvement, we could only prove Theorem 4 under a hardness assumption for Σ_3 circuits, with a larger min-entropy threshold.
- While the main ideas behind the additive extractors came from the authors, AI (particularly OpenAI's GPT 5.6) proved immensely helpful in rapidly exploring the delicate parameter space and working out the technical details. This assistance was especially useful in developing the inclusion–exclusion analysis of Feige's protocol so that it was compatible with the parameters achievable by our relative-error correlation breaker.

4 Preliminaries

We fix the notation and source models, then state the tools used in the constructions.

Notation. Logarithms are base 2, and e is Euler's number. Write $[D] = \{1, \ldots, D\}$. Random variables are boldface, and lowercase letters denote their values; $x \leftarrow \mathbf{X}$ denotes sampling. For a finite set S, $\mathbf{U}_S$ is uniform on S, and $\mathbf{U}_m := \mathbf{U}_{\{0,1\}^m}$. Uniform random variables appearing in a product distribution are independent. We identify a random variable with its distribution when no confusion can arise. Min-entropy thresholds are denoted by k and errors by ε, with subscripts naming the object. A sampler has r input bits; $\mathbf{R} \leftarrow \mathbf{U}_r$ is its random input and $\rho \in \{0,1\}^r$ a fixed one. For $x \in (\{0,1\}^b)^\ell$, $x_{[1,i]} = (x_1, \ldots, x_i)$ denotes its first i blocks. Tuples are written $(u_1, \ldots, u_D)$. We write $\|$ for concatenation and reserve $\circ$ for composition. The size of a circuit, deterministic or nondeterministic, includes its input, output, and witness lengths.

4.1 Basic Probability

4.1.1 Standard definitions

Definition 4.1 (Statistical distance)**.** *For distributions $\mathbf{X}, \mathbf{Y}$ on a common finite set Ω, their statistical distance is*

$$\mathsf{SD}(\mathbf{X}, \mathbf{Y}) = \frac{1}{2} \sum_{a \in \Omega} \big|\Pr[\mathbf{X} = a] - \Pr[\mathbf{Y} = a]\big| = \max_{A \subseteq \Omega} \big|\Pr[\mathbf{X} \in A] - \Pr[\mathbf{Y} \in A]\big|.$$

They are ε-close if $\mathsf{SD}(\mathbf{X}, \mathbf{Y}) \le \varepsilon$.

Definition 4.2 (Min-entropy)**.** *The min-entropy of a finite random variable $\mathbf{X}$ is*

$$\mathbf{H}_\infty(\mathbf{X}) = -\log \max_x \Pr[\mathbf{X} = x].$$

An (n, k)-source is a distribution on $\{0,1\}^n$ with min-entropy at least k. Its entropy deficiency is $n - \mathbf{H}_\infty(\mathbf{X})$, and its support is $\mathrm{Supp}(\mathbf{X}) = \{x : \Pr[\mathbf{X} = x] > 0\}$.

4.1.2 Useful probability lemmas

Definition 4.3 (A subsource of density λ)**.** *Let $\mathbf{Z}$ and $\mathbf{Z}'$ be distributions on the same finite set and let $0 < \lambda \leq 1$. We say that $\mathbf{Z}'$ is a subsource of $\mathbf{Z}$ of density λ if there is a distribution $\mathbf{Z}''$ such that*

$$\mathbf{Z} = \lambda \mathbf{Z}' + (1-\lambda)\mathbf{Z}''.$$

The same definition applies to joint distributions. In particular, for every event A on their common domain, $\Pr[\mathbf{Z} \in A] \geq \lambda \cdot \Pr[\mathbf{Z}' \in A]$.

We use the following simple bounds when conditioning on the seeds extracted from the rows.

Lemma 4.4 (Conditioning on an event)**.** *If $\mathbf{H}_\infty(\mathbf{Y}) \geq k$ and E is an event jointly defined with $\mathbf{Y}$ of probability $p > 0$, then*

$$\mathbf{H}_\infty(\mathbf{Y} \mid E) \geq k - \log(1/p).$$

Proof. For every y, $\Pr[\mathbf{Y} = y \mid E] \leq \Pr[\mathbf{Y} = y]/p \leq 2^{-k}/p$. □

Lemma 4.5 (Density bounds from statistical distance)**.** *Let $\mathbf{P}$ be a distribution on a finite set Ω of size M with $\mathsf{SD}(\mathbf{P}, \mathbf{U}_\Omega) \leq \varepsilon$. Write $p(w) = \Pr[\mathbf{P} = w]$. For every $\lambda > 0$,*

$$\Pr_{w \leftarrow \mathbf{U}_\Omega}\left[p(w) < \frac{1-\lambda}{M}\right] \leq \frac{\varepsilon}{\lambda}, \qquad \Pr_{w \leftarrow \mathbf{U}_\Omega}\left[p(w) > \frac{1+\lambda}{M}\right] \leq \frac{\varepsilon}{\lambda}.$$

Proof. The sum of the positive deficits $1/M - p(w)$ is at most ε, as is the sum of the positive excesses $p(w) - 1/M$. Each point counted in either event contributes more than λ/M to its respective sum. □

We record a simple fact that follows from the hybrid argument.

Lemma 4.6. *Let $\mathcal{A}$ be a finite alphabet and $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_t) \in \mathcal{A}^t$. Suppose that, for every $i \in [t]$,*

$$\mathsf{SD}\Big((\mathbf{X}_i, (\mathbf{X}_j)_{j \neq i}), \big(\mathbf{U}_\mathcal{A}, (\mathbf{X}_j)_{j \neq i}\big)\Big) \leq \varepsilon,$$

where $\mathbf{U}_\mathcal{A}$ is uniform on $\mathcal{A}$ and independent of $\mathbf{X}$. Then $\mathsf{SD}(\mathbf{X}, \mathbf{U}_{\mathcal{A}^t}) \leq t\varepsilon$.

4.2 Circuits and Hardness assumptions

4.2.1 Circuit definitions

Definition 4.7 (Nondeterministic circuits)**.** *For $n, w \in \mathbb{N}$, a nondeterministic circuit C is a Boolean circuit that takes in two inputs $(x, \mathrm{wit}) \in \{0,1\}^{n+w}$ and computes a Boolean function $f : \{0,1\}^n \to \{0,1\}$ such that $f(x) = 1$ if and only if there exists $\mathrm{wit} \in \{0,1\}^w$ such that $C(x, \mathrm{wit}) = 1$.*

Definition 4.8 (Σ_i circuits)**.** *For $n \in \mathbb{N}$, a Σ_i circuit $C(x)$ is a Boolean circuit equipped with special oracle gates which can query a canonical Σ_i complete problem.*

We will additionally find it useful to define single valued nondeterministic circuits

Definition 4.9 (Nondeterministic outputs, PSV, and SV circuits)**.** *Let $\mathcal{O}$ be an output set with a fixed binary encoding. A multi-output nondeterministic circuit is a deterministic circuit $C(x, \mathrm{wit})$ that, on input x and witness wit, either rejects or accepts and outputs an element of $\mathcal{O}$. Define the set of accepted outputs by*

$$\mathrm{ACCEPT}_C(x) \coloneqq \{a \in \mathcal{O} : \exists \mathrm{wit} \text{ such that } C(x, \mathrm{wit}) \text{ accepts and outputs } a\}.$$

(1) C is partial single-valued *(PSV) if $|\,\mathrm{ACCEPT}_C(x)| \leq 1$ for every input x.*

(2) C is single-valued *(SV) if $|\,\mathrm{ACCEPT}_C(x)| = 1$ for every input x.*

For a function $g\colon \mathcal{X} \to \mathcal{O}$, we say that C is correct and single-valued at x *if $\mathrm{ACCEPT}_C(x) = \{g(x)\}$. If this holds for every $x \in \mathcal{X}$, then C is an SV circuit computing g. Thus PSV permits inputs with no successful witness, whereas SV requires at least one successful witness on every input.*

4.2.2 Hardness assumptions

Write $\mathsf{E} = \bigcup_{c>0} \mathsf{DTIME}(2^{cn})$.

Assumption 4.10 (Nondeterministic hardness, H_{NP})**.** *There are a Boolean function $f \in \mathsf{E}$ and a constant $\beta_{\mathrm{hard}} > 0$ such that, for every sufficiently large input length ℓ, no nondeterministic circuit of size at most $2^{\beta_{\mathrm{hard}}\ell}$ computes f on all inputs in $\{0,1\}^\ell$.*

When H_{NP} holds, we also say that E is hard for exponential-size nondeterministic circuits.

Assumption 4.11 (Σ_i hardness, H_{Σ_i})**.** *There is a Boolean function $f \in \mathsf{E}$ and a constant $\beta_{\mathrm{hard}} > 0$ such that, for every sufficiently large input length ℓ, no Σ_i circuit of size at most $2^{\beta_{\mathrm{hard}}\ell}$ computes f on all inputs in $\{0,1\}^\ell$.*

Observe for instance that the assumption H_{Σ_2} implies H_{NP}.

Remark 4.12. *Note that H_{NP} hardness assumption is against nondeterministic circuits while H_{Σ_2} hardness assumption is against Boolean circuits equipped with Σ_2 oracle gates. Moreover, to prove our later theorems stated under H_{Σ_2}, it suffices to establish them under the weaker assumption that E is hard for exponential-size nondeterministic circuits with SAT oracles, since this assumption follows from H_{Σ_2}.*

4.3 Extractors, Condensers, and Models of Randomness

4.3.1 Extractors and condensers for a class of distributions

Definition 4.13 (Extractors for a class of distributions)**.** *Let $\mathcal{C}$ be a class of distributions on $\{0,1\}^n$. A function $\mathsf{Ext}\colon \{0,1\}^n \to \{0,1\}^m$ is a (k,ε)-extractor for $\mathcal{C}$ if, for every $\mathbf{X} \in \mathcal{C}$ with $\mathbf{H}_\infty(\mathbf{X}) \geq k$,*

$$\mathsf{SD}(\mathsf{Ext}(\mathbf{X}), \mathbf{U}_m) \leq \varepsilon.$$

We also call this the additive *guarantee.*

Definition 4.14 (Multiplicative extractors for a class of distributions)**.** *Let $\mathcal{C}$ be a class of distributions on $\{0,1\}^n$. A function $\mathsf{Ext}\colon \{0,1\}^n \to \{0,1\}^m$ is a (k,ε)-multiplicative extractor for $\mathcal{C}$ if, for every $\mathbf{X} \in \mathcal{C}$ with $\mathbf{H}_\infty(\mathbf{X}) \geq k$ and every $\sigma \in \{0,1\}^m$,*

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

Equivalently, for every such $\mathbf{X}$ and every $A \subseteq \{0,1\}^m$,

$$\Pr[\mathsf{Ext}(\mathbf{X}) \in A] \leq e^\varepsilon \cdot \Pr[\mathbf{U}_m \in A].$$

Observe that $e^\varepsilon \leq 1 + 2\varepsilon$ for $\varepsilon \leq 1$.

Definition 4.15 (Errorless condensers for a class of distributions)**.** *Let $\mathcal{C}$ be a class of distributions on $\{0,1\}^n$. A function $\mathsf{Cond}\colon \{0,1\}^n \to \{0,1\}^{m_{\mathsf{Cond}}}$ is an errorless (k, k_{Cond})-condenser for $\mathcal{C}$ if, for every $\mathbf{X} \in \mathcal{C}$,*

$$\mathbf{H}_\infty(\mathbf{X}) \geq k \quad \Longrightarrow \quad \mathbf{H}_\infty(\mathsf{Cond}(\mathbf{X})) \geq k_{\mathsf{Cond}}.$$

The min-entropy guarantee holds for the output distribution itself.

Definition 4.16 (Errorless strong seeded condenser)**.** *Let $\mathcal{C}$ be a class of distributions on $\{0,1\}^n$ and $0 \leq \delta < 1$. A function $\mathsf{Cond}\colon \{0,1\}^n \times [R] \to \{0,1\}^m$ is an errorless strong seeded (k, k', δ)-condenser for $\mathcal{C}$ if, for every $\mathbf{X} \in \mathcal{C}$ with $\mathbf{H}_\infty(\mathbf{X}) \geq k$, at least a $1 - \delta$ fraction of seeds $i \in [R]$ satisfy $\mathbf{H}_\infty(\mathsf{Cond}(\mathbf{X}, i)) \geq k'$.*

4.3.2 Seeded extractors

Definition 4.17 (Seeded extractors)**.** *A function* $\mathsf{SExt}\colon \{0,1\}^n \times \{0,1\}^s \to \{0,1\}^m$ *is a seeded* (k,ε)*-extractor if, for every* (n,k)*-source* $\mathbf{X}$ *and an independent uniform seed* $\mathbf{S} \leftarrow \mathbf{U}_s$*,*

$$\mathsf{SD}(\mathsf{SExt}(\mathbf{X},\mathbf{S}),\mathbf{U}_m) \le \varepsilon.$$

We say that SExt *is* strong *if*

$$\mathsf{SD}\big((\mathbf{S},\mathsf{SExt}(\mathbf{X},\mathbf{S})),(\mathbf{S},\mathbf{U}_m)\big) \le \varepsilon.$$

We use the following strong seeded extractor from [GUV09].

Theorem 4.18. *For every constant* $\delta \in (0,1)$*, there is a constant* C *such that, for input length* n*, error* $0 < \eta < 1/2$*, and* $r \ge C(\log n + \log(1/\eta))$*, there is an explicit strong* (r,η)*-extractor* $\mathsf{SExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with seed length* $d = O(\log n + \log(1/\eta))$ *and output length* $m = (1-\delta)r$*. The extractor is computable in time polynomial in* n *and* $1/\eta$*.*

4.3.3 Block-wise sources

Definition 4.19 (Two-block source)**.** *A pair* $(\mathbf{A},\mathbf{B})$ *is a* (r_1,r_2)*-block-wise source if* $\mathbf{H}_\infty(\mathbf{A}) \ge r_1$ *and, for every* $a \in \mathrm{Supp}(\mathbf{A})$*,* $\mathbf{H}_\infty(\mathbf{B} \mid \mathbf{A} = a) \ge r_2$*. A distribution is* η*-close to such a source if its statistical distance from some such joint distribution is at most* η*.*

4.3.4 Two-source extractors and condensers

Definition 4.20 (Two-source extractors)**.** *A function* $\mathsf{TExt}\colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *is a* (k_1,k_2,ε)*-two-source extractor if, for every pair of independent sources* $\mathbf{X}_1,\mathbf{X}_2$ *with* $\mathbf{H}_\infty(\mathbf{X}_i) \ge k_i$*,*

$$\mathsf{SD}(\mathsf{TExt}(\mathbf{X}_1,\mathbf{X}_2),\mathbf{U}_m) \le \varepsilon.$$

Definition 4.21 (Two-source condensers)**.** *A function* $\mathsf{2Cond}\colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *is a* $(k_1,k_2,k_{\mathrm{out}},\eta)$*-two-source condenser if, for every pair of independent sources* $\mathbf{X}_1,\mathbf{X}_2$ *with* $\mathbf{H}_\infty(\mathbf{X}_i) \ge k_i$*, the distribution* $\mathsf{2Cond}(\mathbf{X}_1,\mathbf{X}_2)$ *is* η*-close to a distribution of min-entropy at least* k_{out}*. It is* errorless *if* $\eta = 0$*; in this case, we omit* η *from the parameters.*

Definition 4.22 (An extractor for one source and one block-wise source)**.** *A function*

$$\mathsf{IExt}\colon \{0,1\}^{n_1} \times (\{0,1\}^{n_2} \times \{0,1\}^{n_3}) \to \{0,1\}^m$$

is a $(k_1,k_2,k_3,\varepsilon)$*-extractor for one source and one block-wise source if, whenever* $\mathbf{U}$ *is independent of* $(\mathbf{A},\mathbf{B})$*,* $\mathbf{H}_\infty(\mathbf{U}) \ge k_1$*, and* $(\mathbf{A},\mathbf{B})$ *is a* (k_2,k_3)*-block-wise source,*

$$\mathsf{SD}\big(\mathsf{IExt}(\mathbf{U},(\mathbf{A},\mathbf{B})),\mathbf{U}_m\big) \le \varepsilon.$$

We also write $\mathsf{IExt}(\mathbf{U},\mathbf{A},\mathbf{B})$ *for* $\mathsf{IExt}(\mathbf{U},(\mathbf{A},\mathbf{B}))$*.*

We use a two-source condenser that guarantees high output min-entropy with no statistical error. The following is a convenient consequence of Ben-Aroya, Cohen, Doron, and Ta-Shma [BCDT19].

Theorem 4.23 (Errorless two-source condensation [BCDT19, Theorem 27])**.** *There is a universal constant* $C_{\mathsf{2Cond}} \ge 1$ *such that, for every sufficiently large* n *and all* r,m *satisfying*

$$\log n \le m, \qquad (m \log n)^{C_{\mathsf{2Cond}}} \le r \le n,$$

there is an explicit polynomial-time function

$$\mathsf{2Cond}_0\colon \{0,1\}^n \times \{0,1\}^n \longrightarrow \{0,1\}^m$$

such that, for every pair of independent sources $\mathbf{U}, \mathbf{V}$ *on* $\{0,1\}^n$ *with* $\mathbf{H}_\infty(\mathbf{U}), \mathbf{H}_\infty(\mathbf{V}) \geq r$*,*

$$\mathbf{H}_\infty\big(\mathsf{2Cond}_0(\mathbf{U},\mathbf{V})\big) \geq 3m/4.$$

To obtain this consequence, apply the cited theorem with error 2^{-m}. For a universal constant C_0, its min-entropy requirement is $r \geq [m\log(n2^m)]^{C_0}$. This follows from the displayed hypothesis after choosing the universal constant C_{2Cond} sufficiently large, since $m \geq \log n$. The theorem bounds every output point probability by

$$2^{-(m-o(m))} + 2^{-m} + 2^{-r/2} \leq 2^{-3m/4}$$

for sufficiently large n. Thus the min-entropy guarantee holds for the output distribution itself, with no statistical error.

Theorem 4.24 (One independent source and one block-wise source [Li15, Theorem 1.4])**.** *For every sufficiently large* n *and every* r *with* $\log^{12} n \leq r \leq n$*, there is an explicit polynomial-time function*

$$\mathsf{IExt}_0\colon \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n \longrightarrow \{0,1\}^{0.9r}$$

such that, whenever $\mathbf{U}$ *is independent of* $(\mathbf{A},\mathbf{B})$*,* $\mathbf{H}_\infty(\mathbf{U}) \geq r$*, and* $(\mathbf{A},\mathbf{B})$ *is an* (r,r)*-block-wise source,* $\mathsf{IExt}_0(\mathbf{U},\mathbf{A},\mathbf{B})$ *is* $2^{-r^{\Omega(1)}}$*-close to uniform. The constants in the error exponent and the polynomial running-time bound are universal. In particular, there is a universal* $c_{\mathsf{IExt}} > 1$ *such that the error is at most* $2^{-r^{1/c_{\mathsf{IExt}}}}$ *for all sufficiently large admissible parameters.*

Li proves the stronger statement in which either independent input may also be revealed; the weaker guarantee above is all we use. This is the component used in the construction of a multiplicative extractor of [Sha25a, Section 4.3].

4.3.5 Two-source non-malleable extraction

The correlation breaker uses an extractor whose output remains close to uniform given outputs on modified versions of its inputs.

Definition 4.25 (Two-source non-malleable extractor)**.** *A function* $\mathsf{2NMExt}\colon \{0,1\}^L \times \{0,1\}^L \to \{0,1\}^q$ *is* t*-non-malleable for min-entropy* k *and error* ε *if the following holds. Let* $\mathbf{X},\mathbf{Y}$ *be independent sources with min-entropy at least* k*. Let* I *be an index set with* $|I| \leq t$*, and let* $f_i, g_i\colon \{0,1\}^L \to \{0,1\}^L$*,* $i \in I$*, be arbitrary functions such that, for each* i*, at least one of* f_i, g_i *has no fixed point. Writing* $\mathbf{X}_i = f_i(\mathbf{X})$ *and* $\mathbf{Y}_i = g_i(\mathbf{Y})$*, we require*

$$\mathsf{SD}\Big(\big(\mathsf{2NMExt}(\mathbf{X},\mathbf{Y}), (\mathsf{2NMExt}(\mathbf{X}_i,\mathbf{Y}_i))_{i\in I}\big),\\ \big(\mathbf{U}_q, (\mathsf{2NMExt}(\mathbf{X}_i,\mathbf{Y}_i))_{i\in I}\big)\Big) \leq \varepsilon.$$

The extractor is strong in its first input if the same bound holds with $\mathbf{X}$ *included in both tuples, and strong in its second input if it holds with* $\mathbf{Y}$ *included. A function* f *has no fixed point if* $f(x) \neq x$ *for every* x*.*

4.3.6 Samplable distributions

We consider distributions generated by deterministic sampling circuits, possibly conditioned on an event recognized by a nondeterministic circuit.

Definition 4.26 (Samplable distributions)**.** *A distribution $\mathbf{X}$ on $\{0,1\}^n$ is samplable by a circuit of size at most s if there is a deterministic Boolean circuit $\mathsf{Samp}\colon \{0,1\}^r \to \{0,1\}^n$ of size at most s such that $\mathsf{Samp}(\mathbf{U}_r)$ has exactly the distribution $\mathbf{X}$. In particular, $r, n \le s$.*

Definition 4.27 (Nondeterministic postselection)**.** *Let $\mathcal{C}^{\mathrm{NP}}_{n,S}$ be the class of distributions on $\{0,1\}^n$ of the form*

$$\mathbf{X} = (\mathsf{Samp}(\mathbf{R}) \mid \mathsf{Post}(\mathbf{R}) = 1), \qquad \mathbf{R} \leftarrow \mathbf{U}_r,$$

where $\mathsf{Samp}\colon \{0,1\}^r \to \{0,1\}^n$ is deterministic, $\mathsf{Post}\colon \{0,1\}^r \to \{0,1\}$ is nondeterministic, their total circuit size is at most S, and $\Pr[\mathsf{Post}(\mathbf{R}) = 1] > 0$.

4.3.7 Extraction from samplable sources

We use the following two ingredients from [Sha25b, Theorem 2.3 and Lemma 4.5]. The first constructs a seed-extending generator and the second shows that its new output bits extract from samplable sources.

Lemma 4.28 (Short-stretch seed-extending PRG)**.** *Assume H_{NP}. For every sufficiently large s, integer $1 \le m \le s$, and $2^{-s} \le \eta \le 1/s$, there is a $\mathrm{poly}(s)$-time computable function $K : \{0,1\}^d \to \{0,1\}^m$, where $m \le d = O(m + \log(1/\eta))$, such that every size-$s$ nondeterministic circuit D satisfies*

$$\Pr[D(\mathbf{U}_d \,\|\, K(\mathbf{U}_d)) = 1] \le e^{1/s} \Pr[D(\mathbf{U}_{d+m}) = 1] + \eta.$$

Here both occurrences of $\mathbf{U}_d$ denote the same sample. Furthermore, if you assume H_{Σ_2}, then the same guarantee holds for nondeterministic circuits with oracle gates to SAT.

Lemma 4.29 (From pseudorandomness to extraction)**.** *There is an absolute constant $C > 1$ with the following property. Let $s \ge d \ge m$, $0 < \varepsilon \le 1$, and $\Delta > 0$. Suppose $K : \{0,1\}^d \to \{0,1\}^m$ satisfies*

$$\Pr[D(\mathbf{U}_d \,\|\, K(\mathbf{U}_d)) = 1] \le e^{\varepsilon} \Pr[D(\mathbf{U}_{d+m}) = 1] + \eta$$

for every size-s nondeterministic circuit D, where $\eta \le \varepsilon 2^{-(m+\Delta)}$. If $(S_0/\varepsilon)^C \le s$, then every $\mathbf{X} \in \mathcal{C}^{\mathrm{NP}}_{d,S_0}$ with $\mathbf{H}_\infty(\mathbf{X}) \ge d - \Delta$ satisfies

$$\Pr[K(\mathbf{X}) = z] \le e^{12\varepsilon} 2^{-m} \qquad \text{for every } z \in \{0,1\}^m.$$

Furthermore, if you assume H_{Σ_2}, then the same guarantee holds for nondeterministic circuits with oracle gates to SAT.

Combining these ingredients gives the following multiplicative extractor, with the source length and sampling circuit size as separate parameters.

Lemma 4.30. *Assume H_{NP}. There is a constant $C_0 > 0$ such that, whenever $2 \le L \le S$, $m \ge 1$, $\Delta \ge 0$, and $0 < \nu \le 1$ satisfy*

$$L \ge C_0\big(\Delta + m + \log S + \log(1/\nu)\big), \tag{1}$$

there is a function $P\colon \{0,1\}^L \to \{0,1\}^m$, computable in time $\mathrm{poly}(S, 1/\nu)$, such that every $\mathbf{V} \in \mathcal{C}^{\mathrm{NP}}_{L,S}$ with $\mathbf{H}_\infty(\mathbf{V}) \ge L - \Delta$ satisfies

$$\Pr[P(\mathbf{V}) = z] \le e^{\nu} 2^{-m} \qquad (z \in \{0,1\}^m). \tag{2}$$

Proof. We construct a generator using Lemma 4.28 and apply Lemma 4.29 to a projection of the source.

Set $\varepsilon_0 = \nu/12$ and $\Delta' = \Delta + 1$. Choose $s = \mathrm{poly}(S, 1/\nu)$ sufficiently large, and let $\eta = \min\{\varepsilon_0 2^{-(m+\Delta')}, 1/s\}$. The length hypothesis and $L \le S$ give $m + \Delta + \log(1/\nu) = O(S)$. Thus we can ensure $2^{-s} \le \eta$, $1/s \le \varepsilon_0$, and the size requirement in Lemma 4.29 for sampling circuits of size $O(S)$.

By Lemma 4.28, there is a function $K : \{0,1\}^d \to \{0,1\}^m$, computable in time $\mathrm{poly}(S, 1/\nu)$, with $d = O(m + \Delta + \log S + \log(1/\nu)) \le L$, where the final inequality follows from (1), taking C_0 sufficiently large. Define $P(v) = K(v_{[d]})$, where $v_{[d]}$ denotes the first d coordinates of v.

For any source $\mathbf{V}$ in the statement, its projection $\mathbf{V}_{[d]}$ has min-entropy at least $d - \Delta$ and belongs to $\mathcal{C}^{\mathrm{NP}}_{d,O(S)}$. Since $1/s \le \varepsilon_0$ and $\eta \le \varepsilon_0 2^{-(m+\Delta')}$, Lemma 4.29 gives

$$\Pr[P(\mathbf{V}) = z] = \Pr[K(\mathbf{V}_{[d]}) = z] \le e^{12\varepsilon_0} 2^{-m} = e^{\nu} 2^{-m},$$

as required. □

Thus P is an $(L - \Delta, \nu)$-multiplicative extractor for $\mathcal{C}^{\mathrm{NP}}_{L,S}$. The multiplicative bound also gives small statistical distance from uniform, with output length linear in the source length.

Corollary 4.31. *Assume H_{NP} and fix $a > 0$. There are constants $c_0, C_a > 0$ such that, whenever $2 \le L \le S$ and $L \ge C_a(\Delta + \log S)$, there is a function $P : \{0,1\}^L \to \{0,1\}^{\lfloor c_0 L \rfloor}$, computable in time $\mathrm{poly}(S)$, satisfying $\mathsf{SD}(P(\mathbf{V}), \mathbf{U}_{\lfloor c_0 L \rfloor}) \le S^{-a}$ for every $\mathbf{V} \in \mathcal{C}^{\mathrm{NP}}_{L,S}$ with $\mathbf{H}_\infty(\mathbf{V}) \ge L - \Delta$.*

Proof. In Lemma 4.30, take $m = \lfloor c_0 L \rfloor$, the deficiency bound L/C_a, and $\nu = S^{-a}/4$. For sufficiently small c_0 and sufficiently large C_a, the length condition holds. Summing the positive differences between the output probabilities and 2^{-m} gives $\mathsf{SD}(P(\mathbf{V}), \mathbf{U}_m) \le e^{\nu} - 1 \le 2\nu \le S^{-a}$. □

4.3.8 Robust errorless block condensers

Definition 4.32 (Robust errorless block condenser)**.** *Fix $0 < \lambda \le 1$. A function*

$$\mathsf{rbCond} \colon \{0,1\}^{n_y + n_x} \longrightarrow \{0,1\}^m$$

is a $(\lambda, k_y, k_y^{\ge}, k_x, k_{\mathrm{out}})$-robust errorless block condenser for a class $\mathcal{YX}$ of joint distributions on $\{0,1\}^{n_y+n_x}$ if the following holds. For every $(\mathbf{Y}, \mathbf{X}) \in \mathcal{YX}$ and every decomposition

$$(\mathbf{Y}, \mathbf{X}) = \lambda'(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}}) + (1 - \lambda')(\mathbf{Y}_{\mathrm{rest}}, \mathbf{X}_{\mathrm{rest}})$$

as a mixture of joint distributions with $\lambda' \ge \lambda$, satisfying

- $\mathbf{H}_\infty(\mathbf{Y}_{\mathrm{good}}) \ge k_y$;
- $\Pr[\mathbf{Y}_{\mathrm{good}} = y] \ge 2^{-k_y^{\ge}}$ *for every* $y \in \mathrm{Supp}(\mathbf{Y}_{\mathrm{good}})$;
- $\mathbf{H}_\infty(\mathbf{X}) \ge k_x$;

we have

$$\mathbf{H}_\infty\big(\mathsf{rbCond}(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}})\big) \ge k_{\mathrm{out}}.$$

4.3.9 Output amplification

We will use the results from [Sha08] to amplify the number of output bits of an extractor, in particular, we use the following formulation for sources sampled with postselection.

Lemma 4.33 (Output amplification [Sha25a, Section 4.2])**.** *For all constants $c, d > 0$ and $\eta \in (0,1)$, there are constants $C, c_0, d_0 > 0$ such that the following holds for all sufficiently large n.*

Let $t = C \log n$ and $k \geq t$. Suppose that $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^t$ is a (k, n^{-d_0})-extractor for distributions samplable by size-n^{c_0} deterministic circuits with postselection. Then, for every k' satisfying $\frac{2k}{\eta} \leq k' \leq n$, there is a (k', n^{-d})-extractor $\mathsf{Ext}' : \{0,1\}^n \to \{0,1\}^{\lfloor (1-\eta)k' \rfloor}$ for distributions samplable by size-n^c deterministic circuits with postselection. Moreover, if Ext is computable in time $T(n)$, then Ext' is computable in time $T(n) + \mathrm{poly}(n)$.

The lemma also holds for sources with nondeterministic postselection, since the conditioning used in the proof only conjoins an efficient deterministic predicate to the original postselection predicate, preserving nondeterministic postselection with polynomial size overhead.

4.4 Algebra and Coding theory tools

4.4.1 Low-degree extension

The following is the standard low-degree extension, as stated in [DS26, Lemma 3.9], with its individual-degree bound made explicit.

Theorem 4.34 (Low-degree extension)**.** *Let $f \colon \{0,1\}^\ell \to \{0,1\}$, and let $d \leq h \leq q$ be positive integers such that $h^d = 2^\ell$ and q is a power of two. Let $H \subseteq \mathbb{F}_q$ be the set of the first h field elements, and let*

$$\phi \colon \{0,1\}^\ell \to H^d$$

be an efficiently computable bijection. There is a polynomial

$$\widehat{f} \colon \mathbb{F}_q^d \to \mathbb{F}_q$$

of degree at most $h-1$ in each variable, and hence total degree at most

$$\hat{h} := d(h-1),$$

such that $\widehat{f}(\phi(x)) = f(x)$ for every $x \in \{0,1\}^\ell$. Furthermore, $\widehat{f}$ can be computed in time $\mathrm{poly}(2^\ell, \log q)$, given oracle access to f.

4.4.2 Parametrized lines

Definition 4.35 (Parametrized lines and random lines)**.**

(1) A parametrized affine line *is a map $L \colon \mathbb{F}_q \to \mathbb{F}_q^d$ of the form*

$$L(t) = u + t(v - u), \qquad u, v \in \mathbb{F}_q^d.$$

(2) We call L degenerate *if $u = v$, in which case L is constant.*

(3) The experiment $\mathbf{L} \leftarrow \mathsf{RandLine}$ samples $\mathbf{u}, \mathbf{v}$ independently and uniformly from $\mathbb{F}_q^d$, and returns $\mathbf{L}(t) = \mathbf{u} + t(\mathbf{v} - \mathbf{u})$.

In particular,

$$\Pr_{\mathbf{L}\leftarrow\mathsf{RandLine}}[\mathbf{L} \text{ is degenerate}] = q^{-d}.$$

We also write L for its image, as in $w \in L$. For nondegenerate L, let $\mathbf{U}_L$ denote the uniform distribution on its q points, and let $L^{-1}(w)$ be the unique $t \in \mathbb{F}_q$ with $L(t) = w$.

Lemma 4.36 (Basic sampling facts)**.** *Let* $\mathbf{L} \leftarrow \mathsf{RandLine}$.

(1) For each fixed $t \in \mathbb{F}_q$, $\mathbf{L}(t)$ is uniform on $\mathbb{F}_q^d$. For distinct t, t', the points $\mathbf{L}(t)$ and $\mathbf{L}(t')$ are independent.

(2) For every fixed $w \in \mathbb{F}_q^d$,

$$\Pr[w \in \mathbf{L} \mid \mathbf{L} \text{ is nondegenerate}] = q^{-(d-1)}.$$

Proof. Write $\mathbf{L}(t) = \mathbf{u} + t\mathbf{b}$, where $\mathbf{u}, \mathbf{b}$ are independent uniform points of $\mathbb{F}_q^d$. For $t \neq t'$, the map $(u, b) \mapsto (u + tb, u + t'b)$ is invertible, proving the first assertion. Conditioned on nondegeneracy, $\mathbf{b}$ is uniform on $\mathbb{F}_q^d \setminus \{0\}$ and independent of $\mathbf{u}$. Each $\mathbf{L}(t)$ is still uniform, and the q points are distinct, so the incidence probability is $\sum_{t\in\mathbb{F}_q} \Pr[\mathbf{L}(t) = w \mid \mathbf{b} \neq 0] = q \cdot q^{-d}$. □

Lemma 4.37 (Averaging sampler property of a random line)**.** *For every $A\colon \mathbb{F}_q^d \to [0, 1]$ and $\tau > 0$,*

$$\Pr_{\mathbf{L}\leftarrow\mathsf{RandLine}}\left[\left|\mathop{\mathbb{E}}_{t\leftarrow\mathbb{F}_q}[A(\mathbf{L}(t))] - \mathop{\mathbb{E}}_{w\leftarrow\mathbb{F}_q^d}[A(w)]\right| > \tau\right] \leq \frac{1}{4q\tau^2}.$$

Proof. By Lemma 4.36, the variables $A(\mathbf{L}(t))$ are pairwise independent. Each has mean $\mathbb{E}_{w\leftarrow\mathbb{F}_q^d}[A(w)]$ and variance at most $1/4$. Their average therefore has variance at most $1/(4q)$. Apply Chebyshev's inequality. □

4.4.3 List recovery

Lemma 4.38 (Johnson bound for list recovery)**.** *Suppose $S_t \subseteq \mathbb{F}_q$ and $|S_t| \leq K$ for every t. If $q\xi^2 \geq 2\hat{h}K$, the number of degree-at-most-$\hat{h}$ polynomials satisfying $p(t) \in S_t$ on at least a ξ fraction of field elements $t \in \mathbb{F}_q$ is at most $2K/\xi^2$.*

This is a standard form of the Johnson bound for list recovery; see [GKORS18, Lemma 5.2]. For completeness, we give a proof sketch.

Proof. Let the candidate family have size $N > 0$, and let $N(t, a)$ count its members taking value a at t. Distinct polynomials agree at most $\hat{h}$ times, so $\sum_{t,a} N(t, a)^2 \leq Nq + N(N - 1)\hat{h}$. There are at most qK pairs (t, a) with $a \in S_t$. Cauchy–Schwarz therefore gives

$$(Nq\xi)^2 \leq qK\big(Nq + N(N - 1)\hat{h}\big).$$

Dividing and using the hypothesis yields $N\xi^2 \leq K + NK\hat{h}/q \leq K + N\xi^2/2$, which proves the claim. □

4.5 Interactive Protocols

4.5.1 Approximate counting and pseudorandomness

The hashing construction uses approximate counting to recognize unsuitable hashes, then uses a pseudorandom generator to reduce the number of hashes.

An AM protocol has two messages: Arthur sends a uniform string u, and Merlin replies with v. A deterministic verifier V accepts according to $V(x, u, v)$, so the maximum acceptance probability on input x is $\Pr_u[\exists v : V(x, u, v) = 1]$.

We use the following nondeterministic large-set formulation of the Goldwasser–Sipser theorem [GS86], as stated in [Sha25b, Theorem 3.14 of the full version].

Theorem 4.39 (Goldwasser–Sipser for nondeterministic circuits)**.** *For every integer $s \geq 1$ and $\lambda > 0$, there is a nondeterministic circuit $A(C, \gamma)$ of size $\mathrm{poly}(s, 1/\lambda)$ with the following guarantees for every nondeterministic circuit C of size at most s and every rational $\gamma \in (0, 1]$ represented using at most s bits:*

$$\Pr_\rho[C(\rho) = 1] \geq \gamma \quad \Longrightarrow \quad A(C, \gamma) = 1,$$
$$\Pr_\rho[C(\rho) = 1] < \gamma e^{-\lambda} \quad \Longrightarrow \quad A(C, \gamma) = 0.$$

Here ρ is uniform over the ordinary input bits of C.

The Goldwasser–Sipser protocol extends to nondeterministic C by having Merlin supply a membership witness for each claimed accepted input. Amplification and fixing the public coins give the stated nondeterministic circuit, as in $\mathsf{AM} \subseteq \mathsf{NP/poly}$.

We will apply this theorem to circuits C_x obtained by hardwiring an external input x into a fixed circuit. Supplying the description of C_x and a computed threshold $\gamma(x)$ to A gives one circuit working for all x, with only polynomial overhead. The parameter s is taken large enough to include the threshold's bit length.

Lemma 4.40 (Shaltiel–Umans [SU05])**.** *Assume H_{NP}. For all constants $a, b > 0$, there is an algorithm that, for all sufficiently large n and every integer $1 \leq q \leq n^a$, computes a generator $G_{n,q} \colon \{0,1\}^d \to \{0,1\}^q$ on a given seed in time $\mathrm{poly}(n)$, where $d = O(\log n)$, such that every nondeterministic circuit $T \colon \{0,1\}^q \to \{0,1\}$ of size at most n^a satisfies*

$$|\Pr[T(G_{n,q}(\mathbf{U}_d)) = 1] - \Pr[T(\mathbf{U}_q) = 1]| \leq n^{-b}.$$

Part I: Multiplicative Extractors

We here construct multiplicative extractors for $n^{0.01}$ entropy samplable sources, under nondeterministic hardness. Formally,

Theorem 4.41 (Multiplicative extractors under nondeterministic hardness)**.** *Assume that E is hard for exponential-size nondeterministic circuits. For every constant $0 < \alpha \leq 1$ there exists a constant $\alpha' > 0$, depending only on α, such that, for all constants $c_{\mathrm{size}} > 1$ and $c_\varepsilon > 0$, the following holds for all sufficiently large n. There is an explicit polynomial-time computable function*

$$\mathsf{Ext} \colon \{0,1\}^n \longrightarrow \{0,1\}^m, \qquad m = n^{\alpha'},$$

that is an $(n^\alpha, n^{-c_\varepsilon})$-multiplicative extractor for distributions samplable by deterministic circuits with postselection by nondeterministic circuits, both of size at most $n^{c_{\mathrm{size}}}$, in the sense of Definition 4.14. One may take $\alpha' = \alpha\delta_{\mathrm{out}}^2/(8c_{\mathsf{IExt}})$, where δ_{out} and c_{IExt} are the universal constants from Theorems 4.24 and 4.42.

Using this multiplicative extractor, we will also obtain an additive extractor, with similar requirements, under the same assumptions — Theorem 5.1.

To construct such a multiplicative extractor, we will construct errorless condensers for $n^{0.01}$ entropy samplable sources under nondeterministic hardness. Formally,

Theorem 4.42 (Errorless condensers under nondeterministic hardness)**.** *There exists a universal constant $0 < \delta_{\text{out}} < 1/2$ such that the following holds. Assume that E is hard for exponential-size nondeterministic circuits. Then, for all constants $0 < \gamma_{\text{in}} \leq 1$ and $c > 1$, there is an explicit polynomial-time function*

$$\mathsf{Cond}_{\gamma_{\text{in}},c} \colon \{0,1\}^n \longrightarrow \{0,1\}^{n^{\gamma_{\text{in}}/2}}$$

with the following guarantee for all sufficiently large n. For every source $\mathbf{X}$ samplable by a deterministic circuit with postselection by a nondeterministic circuit, both of size at most n^c,

$$\mathbf{H}_\infty(\mathbf{X}) \geq n^{\gamma_{\text{in}}} \quad \Longrightarrow \quad \mathbf{H}_\infty\big(\mathsf{Cond}_{\gamma_{\text{in}},c}(\mathbf{X})\big) \geq n^{\gamma_{\text{in}}\delta_{\text{out}}}.$$

Thus the function is an errorless $(n^{\gamma_{\text{in}}}, n^{\gamma_{\text{in}}\delta_{\text{out}}})$-condenser for this class of distributions, in the sense of Definition 4.15. The polynomial running-time exponent may depend on γ_{in} and c, whereas δ_{out} is independent of both.

The construction of this errorless condenser will rely on a related object which we call robust errorless block condenser, as defined in Definition 4.32. We will construct the following condenser.

Theorem 4.43 (Robust errorless block condensers under nondeterministic hardness)**.** *There exists a universal constant $\delta > 0$ such that the following holds. Assume that E is hard for exponential-size nondeterministic circuits. Then, for every constant $0 < \gamma \leq 1$ and any $c_{\text{size}} > 1$, the following holds for every sufficiently large n. For every*

$$n^\gamma \leq k_y \leq k_x \leq n, \qquad (\log n)^2 \leq m \leq k_y^\delta,$$

there is an explicit polynomial-time computable function

$$\mathsf{rbCond} \colon \{0,1\}^n \times \{0,1\}^n \longrightarrow \{0,1\}^m$$

that is a

$$\big(\lambda = 2^{-m/4},\ k_y,\ k_y^{\geq} = k_x/2,\ k_x,\ k_{\text{out}} = m/4\big)$$

robust errorless block condenser for joint distributions samplable by deterministic circuits with postselection by nondeterministic circuits, both of size at most $n^{c_{\text{size}}}$, in the sense of Definition 4.27.

The rest of the sections are organized as follows. We will show how to obtain additive extractors from our multiplicative extractors in Section 5. We show how to construct robust errorless block condensers (Theorem 4.43), as well as multiplicative extractors from errorless condensers (Theorem 4.41) in Section 9 and Section 10 respectively. Both these proofs will follow a common theme, and will be an elaborate extension of arguments in [BGDM23] and the common components of these proofs are provided in Section 8. We will also give an overview of the arguments involved in these constructions in Section 7.

We will show how to construct errorless condenser from robust errorless block condenser and obtain Theorem 4.42 in Section 6. We will do so by extending ideas appearing in [Sha25a].

5 Additive Extractors from Multiplicative Extractors

In this section, we will show how using the multiplicative extractors from Theorem 4.41, we can obtain additive extractors with similar guarantees.

Theorem 5.1 (Additive extractors under nondeterministic hardness)**.** *Assume that* E *is hard for exponential-size nondeterministic circuits. For all constants* $0 < \alpha, \eta < 1$ *and constants* $c_{\text{size}} > 1$ *and* $c_\varepsilon > 0$*, the following holds for all sufficiently large* n *and* $k = n^\alpha$*. There is an explicit polynomial-time computable function*

$$\mathsf{Ext}\colon \{0,1\}^n \longrightarrow \{0,1\}^m, \qquad m = (1-\eta)k,$$

that is a (k, n^{-c_ε})*-extractor for distributions samplable by deterministic circuits with postselection by nondeterministic circuits, both of size at most* $n^{c_{\text{size}}}$.

Proof. We set constants $C_0, C_1, C_2 > 0$ appropriately later and let Ext_0 be an $(n^{0.99\alpha}, n^{-C_0})$-multiplicative extractor for distributions samplable by size n^{C_1} circuits, with postselection by size n^{C_1} nondeterministic circuits, as guaranteed by Theorem 4.41.

Construction On input $x \in \{0,1\}^n$ we let $y = \mathsf{Ext}_0(x)$, and let y' be the length $C_2 \log(n)$ prefix of y (we can take such a prefix since y has length at least n^δ for some constant $\delta > 0$). Using this, we instantiate Lemma 4.33 to obtain $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ which is an $(n^\alpha, n^{-c_\varepsilon})$ extractor for sources samplable by size $n^{c_{\text{size}}}$ circuits with postselection by size $n^{c_{\text{size}}}$ nondeterministic circuits.

Analysis Let $\mathbf{X}$ be the input distribution and let $\mathbf{Y}, \mathbf{Y}'$ be the distributions of y, y' on that source. To obtain the desired guarantees on Ext using Lemma 4.33, it suffices to show that $\mathsf{SD}(\mathbf{Y}', \mathbf{U}_{C_2 \log(n)}) \le n^{-C_3}$ and that circuit size parameter for Ext_0 is at least n^{C_4} and entropy parameter for Ext_0 is at most $\eta k/2 = \eta \cdot n^\alpha/2$ where $C_2, C_3, C_4 > 0$ are constants supplied by Lemma 4.33 (and depend on $\eta, c_{\text{size}}, c_\varepsilon$). Using this, we set $C_1 = C_4$ and $C_0 = C_2 + C_3 + 1$ to see that the parameter guarantees hold for Ext_0, where we used the fact that $\eta \cdot n^\alpha/2 > n^{0.99\alpha}$. It remains to show that the desired statistical distance guarantees hold for $\mathbf{Y}'$.

Let $n_y = |\mathbf{Y}|$. By guarantees of Ext_0, we have that for all $y \in \{0,1\}^{n_y}$, setting $\varepsilon = n^{-C_0}$,

$$\Pr[\mathbf{Y} = y] \le e^\varepsilon \cdot 2^{-n_y} \le (1 + 2\varepsilon) \cdot 2^{-n_y}.$$

As $\mathbf{Y}'$ is length $n_{y'} = C_2 \log(n)$ prefix of $\mathbf{Y}$, we obtain that for all $y' \in \{0,1\}^{n_{y'}}$,

$$\Pr[\mathbf{Y}' = y'] \le (1 + 2\varepsilon) \cdot 2^{-n_{y'}}.$$

From this we conclude that,

$$\mathsf{SD}(\mathbf{Y}', \mathbf{U}_{n_{y'}}) \le 2^{n_{y'}} \cdot 2\varepsilon \le n^{C_2} \cdot 2 \cdot n^{-C_0} \le n^{-C_3}$$

as desired, wherein the last inequality follows by our setting of C_0. □

6 Errorless condensers using robust errorless block condensers

We here extend the errorless condensers of [Sha25a] under fewer computational assumptions and prove Theorem 4.42.

Towards doing so, we will need two key ingredients. First, we need the notion of a 'robust' errorless block condenser which we defined in Definition 4.32. We call such condensers robust since they actually condense from sub-sources of a samplable source. Second, we need a classical structural result regarding chopping up any entropic source appropriately; we provide a self-contained statement and proof of this in our setting in Section 6.1. We will use these ingredients and carefully instantiate them to construct our errorless condensers. To do that, we will use the robust errorless condenser from Theorem 4.43 and construct a mild errorless condenser in Section 6.2. Using this, we will construct our claimed errorless condenser in Section 6.3.

6.1 Structural chopping result for any source

We next prove the following structural chopping result for any source; such a result was introduced by [SSZ98; Ta-96] and has been useful over the years in several works in extractor theory including [SSZ98; NT99; RSW06; Sha25a].

Lemma 6.1. *For all $\varepsilon > 0$ and $k, n, b, \ell \in \mathbb{N}$ such that $k \le n$ and $\ell \cdot b = n$, the following holds. For any distribution $\mathbf{X} \sim \{0,1\}^n$ with $\mathbf{H}_\infty(\mathbf{X}) \ge k$, consider the partition $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_\ell)$ where $\mathbf{X}_i \sim \{0,1\}^b$. Then for all $0 \le i \le \ell$, there exist sources $\mathbf{X}^{(i)} \sim \{0,1\}^n$ and $\lambda_i \ge 0$ such that $\sum_{i=0}^{\ell} \lambda_i = 1$, so that we can express $\mathbf{X}$ as a convex combination $\sum_{i=0}^{\ell} \lambda_i \mathbf{X}^{(i)}$ with the following properties:*

1. *For all $i \ge 0$ with $\lambda_i > 0$, each $\mathbf{X}^{(i)}$ has disjoint support, and hence partitions $\mathrm{Supp}(\mathbf{X})$. Moreover, for $i \ge 1$, and $x \in \mathrm{Supp}(\mathbf{X})$, membership of $x \in \mathrm{Supp}(\mathbf{X}^{(i)})$ is determined by the prefix $x_{[1,i]}$.*
2. *$\lambda_0 \le \varepsilon$ and for all $i \ge 1$, either $\lambda_i \ge \frac{\varepsilon}{2n}$ or $\lambda_i = 0$.*
3. *For all $i \ge 1$ with $\lambda_i > 0$, we have that $\mathbf{H}_\infty\left(\mathbf{X}^{(i)}_{[1,i]}\right) \ge \frac{k}{2} - \log(2n/\varepsilon)$.*
4. *For all $i \ge 1$ with $\lambda_i > 0$, and for all $x_i \in \mathrm{Supp}\left(\mathbf{X}^{(i)}_{[1,i]}\right)$, we have that*
$$\Pr\left[\mathbf{X}^{(i)}_{[1,i]} = x_i\right] \ge 2^{-k/2-b-\log(2n/\varepsilon)}.$$
5. *For all $i \ge 1$ with $\lambda_i > 0$, it holds that $\mathbf{X}^{(i)}_{[1,i]} \coloneqq \mathbf{X}_{[1,i]} \mid \left(\mathbf{X} \in \mathrm{Supp}\left(\mathbf{X}^{(i)}\right)\right)$.*

Proof. We will first define several helper functions and sets so that we can define λ_i and $\mathbf{X}^{(i)}$. Then, we will show how each of the properties holds for them.

Setup First, define an assignment function $\eta : \mathrm{Supp}(\mathbf{X}) \to [\ell]$ as follows. Let $\eta(x)$ be the smallest index $i \in [\ell]$ such that $\Pr[\mathbf{X}_{[1,i]} = x_{[1,i]}] \le 2^{-k/2}$ (such an index always exists since $\Pr[\mathbf{X} = x] \le 2^{-k}$).

Next, for $i \in [\ell]$, let

$$S_i = \{x \in \mathrm{Supp}(\mathbf{X}) : \eta(x) = i \wedge \Pr[\mathbf{X}_{[1,i]} = x_{[1,i]}] \ge 2^{-k/2-b-\log(2n/\varepsilon)}\}.$$

Then, for $i \in [\ell]$, let

$$T_i = \begin{cases} S_i & \Pr[\mathbf{X} \in S_i] \ge \varepsilon/2n \\ \emptyset & \text{otherwise} \end{cases}$$

And then define

$$T_0 = \mathrm{Supp}(\mathbf{X}) \setminus \left(\bigcup_{i=1}^{\ell} T_i\right).$$

Finally, for all $0 \le i \le \ell$, we define $\mathbf{X}^{(i)} = \mathbf{X} \mid (\mathbf{X} \in T_i)$ and $\lambda_i = \Pr[\mathbf{X} \in T_i]$ (if $\lambda_i = 0$, then let $\mathbf{X}^{(i)}$ be arbitrary).

Proving Properties By definition, the sets $\{S_i\}_{i \ge 1}$ are disjoint and hence, the sets $\{T_i\}_{i \ge 0}$ partition the support of $\mathbf{X}$. From this, we can indeed express $\mathbf{X}$ as the convex combination

$$\mathbf{X} = \sum_{i=0}^{\ell} \Pr[\mathbf{X} \in T_i] \cdot (\mathbf{X} \mid \mathbf{X} \in T_i) = \sum_{i=0}^{\ell} \lambda_i \mathbf{X}^{(i)}$$

as desired. Moreover, membership of $x \in T_i$ clearly depends only on $x_{[1,i]}$. This establishes the first property. We will prove each of the remaining properties one after the other.

Property 2 By definition of T_i, for each $i \geq 1$ with $\lambda_i > 0$, we have that $\lambda_i \geq \frac{\varepsilon}{2n}$ as desired.

For λ_0, we find it helpful to define

$$S_0 = \mathrm{Supp}(\mathbf{X}) \setminus \left(\bigcup_{i=1}^{\ell} S_i \right).$$

We then claim the following:

Claim 6.2. $\Pr[\mathbf{X} \in S_0] \leq \frac{\varepsilon}{2}$.

We will prove this claim later. Let us see how our desired bound on λ_0 follows from this. Let $L = \{i \in [\ell] : \Pr[\mathbf{X} \in S_i] < \frac{\varepsilon}{2n}\}$. We observe that $T_0 = S_0 \cup \bigcup_{i \in L} S_i$. From a union bound, we infer that

$$\lambda_0 = \Pr[\mathbf{X} \in T_0] \leq \Pr[\mathbf{X} \in S_0] + \sum_{i \in L} \Pr[\mathbf{X} \in S_i] \leq \frac{\varepsilon}{2} + \ell \cdot \frac{\varepsilon}{2n} \leq \varepsilon$$

as desired.

We now prove our remaining claim regarding S_0.

Proof of Claim 6.2. For $i \in [\ell]$, define $R_i = \{x \in \mathrm{Supp}(\mathbf{X}) : \eta(x) = i\}$. Then, $R_1, \ldots, R_\ell$ partition $\mathrm{Supp}(\mathbf{X})$. Also, observe that $S_i \subset R_i$ for all $i \geq 1$. We claim that for all $i \geq 1$, $\Pr[\mathbf{X} \in (R_i \setminus S_i)] \leq \frac{\varepsilon}{2n}$. This suffices to prove the claim since

$$\Pr[\mathbf{X} \in S_0] = \sum_{i=1}^{\ell} \Pr[\mathbf{X} \in (R_i \setminus S_i)] \leq \ell \cdot \frac{\varepsilon}{2n} \leq \frac{\varepsilon}{2}$$

as desired.

For the rest of this proof, we fix $j \in [\ell]$. We will show that $\Pr[\mathbf{X} \in (R_j \setminus S_j)] \leq \frac{\varepsilon}{2n}$. Fix any $x^* \in R_j \setminus S_j$. As $x^* \in R_j$, we have that $\eta(x^*) = j$ and so $\Pr\left[\mathbf{X}_{[1,j-1]} = x^*_{[1,j-1]}\right] > 2^{-k/2}$. Since $x^* \notin S_j$ even though $\eta(x^*) = j$, we must have that $\Pr\left[\mathbf{X}_{[1,j]} = x^*_{[1,j]}\right] \leq 2^{-k/2-b-\log(2n/\varepsilon)}$. Hence, we compute that

$$\Pr\left[\mathbf{X}_{[1,j]} = x^*_{[1,j]} \mid \left(\mathbf{X}_{[1,j-1]} = x^*_{[1,j-1]}\right)\right] = \frac{\Pr\left[\mathbf{X}_{[1,j]} = x^*_{[1,j]}\right]}{\Pr\left[\mathbf{X}_{[1,j-1]} = x^*_{[1,j-1]}\right]} \leq \frac{2^{-k/2-b-\log(2n/\varepsilon)}}{2^{-k/2}} \leq \frac{\varepsilon \cdot 2^{-b}}{2n}.$$

Since this is true for all $x^* \in R_j \setminus S_j$, we obtain the following: Let $V_j \subset \{0,1\}^{j \cdot b}$ be the projection of $R_j \setminus S_j$ onto the first j blocks; then for all $v \in V_j$, we have that

$$\Pr\left[\mathbf{X}_{[1,j]} = v \mid \left(\mathbf{X}_{[1,j-1]} = v_{[1,j-1]}\right)\right] \leq \frac{\varepsilon \cdot 2^{-b}}{2n}. \tag{3}$$

Using this, we compute that

$$\begin{aligned}\Pr[\mathbf{X} \in R_j \setminus S_j] &\le \Pr[\mathbf{X}_{[1,j]} \in V_j] \\ &= \sum_{\alpha \in \{0,1\}^{(j-1)b}} \Pr[\mathbf{X}_{[1,j-1]} = \alpha] \cdot \sum_{\beta \in \{0,1\}^b : (\alpha,\beta) \in V_j} \Pr\left[\mathbf{X}_{[1,j]} = (\alpha,\beta) \mid \left(\mathbf{X}_{[1,j-1]} = \alpha\right)\right] \\ &\le \sum_{\alpha \in \{0,1\}^{(j-1)b}} \Pr[\mathbf{X}_{[1,j-1]} = \alpha] \cdot \sum_{\beta \in \{0,1\}^b : (\alpha,\beta) \in V_j} \frac{\varepsilon \cdot 2^{-b}}{2n} \\ &\le \sum_{\alpha \in \{0,1\}^{(j-1)b}} \Pr[\mathbf{X}_{[1,j-1]} = \alpha] \cdot \frac{\varepsilon}{2n} \\ &= \frac{\varepsilon}{2n}\end{aligned}$$

as desired, where the first inequality follows since V_j is a projection of $R_j \setminus S_j$, and the second inequality follows by Equation (3). □

Properties 3 and 4 Fix $j \ge 1$ with $\lambda_j > 0$ (so that $T_j = S_j$) and let $v \in \mathrm{Supp}(\mathbf{X}^{(j)}_{[1,j]})$ be arbitrary. By Property 1, every $x \in \mathrm{Supp}(\mathbf{X})$ with $x_{[1,j]} = v$ lies in T_j (since v is the prefix of some element in T_j). Therefore, $\Pr[\mathbf{X}^{(j)}_{[1,j]} = v] = \Pr[\mathbf{X}_{[1,j]} = v]/\lambda_j$. By the definitions of η and S_j, we have that $2^{-k/2-b-\log(2n/\varepsilon)} \le \Pr[\mathbf{X}_{[1,j]} = v] \le 2^{-k/2}$. Moreover, by Property 2, we have that $\varepsilon/2n \le \lambda_j \le 1$. Combining these three observations, and appropriately dividing, we infer that:

$$2^{-k/2-b-\log(2n/\varepsilon)} \le \Pr\left[\mathbf{X}^{(j)}_{[1,j]} = v\right] \le 2^{-k/2} \cdot \frac{2n}{\varepsilon}.$$

This establishes Property 4, and also Property 3 after taking logarithms.

Property 5 This holds because for any $x \in T_i$, whether or not $x \in T_i$ can be decided solely by testing properties of $x_{[1,i]}$.

□

6.2 Constructing basic errorless Condenser

We here construct the following condenser that mildly shrinks and condenses from subsources of a samplable source, given access to another (potentially fully dependent) source.

Lemma 6.3. *There exists a universal constant $0 < \delta \le 1/4$ such that, assuming E is hard for nondeterministic exponential size circuits, the following holds. For any constants $0 < \alpha, \gamma \le 1$, and $c > 1$, for all sufficiently large $n \in \mathbb{N}$, and all $n_a \in \mathbb{N}$ such that $n_a \le n^\alpha$, setting $k = n^\gamma$, there exists an explicit function $\mathsf{Cond} : \{0,1\}^{n_a} \times \{0,1\}^n \to \{0,1\}^m$ where $m = n^{\alpha-(1-\delta)\gamma}$ with the following properties. Let $\mathbf{A} \sim \{0,1\}^{n_a}$ and $\mathbf{W} \sim \{0,1\}^n$ be such that the distribution $(\mathbf{A}, \mathbf{W})$ is samplable by size n^c circuits with postselection by size n^c nondeterministic circuits, and $(\mathbf{A}, \mathbf{W}) = (1-p)(\mathbf{A}^{(0)}, \mathbf{W}^{(0)}) + p(\mathbf{A}_{good}, \mathbf{W}_{good})$ for some $p \ge 2^{-n^{\delta\gamma}/32}$ where $\mathbf{H}_\infty(\mathbf{A}_{good}) \ge k$ and $\mathbf{H}_\infty(\mathbf{W}) \ge k$. Then, $\mathbf{H}_\infty(\mathsf{Cond}(\mathbf{A}_{good}, \mathbf{W}_{good})) \ge n^{\delta\gamma/4}$.*

Proof. Let $k_y = \frac{k}{16}$. Let $0 < \delta_{\textit{Theorem 4.43}} \le 1$ be the universal constant from Theorem 4.43. Then, we let $\delta = \delta_{\textit{Theorem 4.43}}/4$.

We will construct such a function Cond and then analyze it.

Construction On input $a \sim \mathbf{A}, w \sim \mathbf{W}$, we proceed as follows.

- Let $b = 6k_y, \ell = \frac{n_a}{6k_y}$ and chop a to obtain $a_1, \ldots, a_\ell \in \{0,1\}^b$.
- Let $\mathsf{rbCond} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{n_1}$ be a $(2^{-k_1}, k_y/4, k/2, k, k_1)$ robust errorless block condenser for sources samplable by size n^c circuits with postselection by size n^c nondeterministic circuits from Theorem 4.43, with parameter $\gamma_{Theorem\ 4.43} = \gamma/2$, where $n_1 = k^\delta/4$ and $k_1 = n_1/4$ (we meet the output length requirement that $n_1 \leq (k_y/4)^{\delta_{Theorem\ 4.43}}$ by using the fact that $\delta \leq \delta_{Theorem\ 4.43}/4$; the rest of the requirements are easily met).
- For $i \in [\ell]$, let $y_i = \mathsf{rbCond}(a_{[1,i]}, w)$ and output $y = (y_1, \ldots, y_\ell)$ and pad appropriately to obtain the desired output length.

Output Length First, we compute that the output length is

$$m = \ell \cdot n_1 = \frac{n_a}{6k_y} \cdot \frac{k^\delta}{4} \leq \frac{1}{24} \cdot n^{\alpha+\gamma\delta} \cdot \frac{16}{n^\gamma} = \frac{2}{3} \cdot n^{\alpha-(1-\delta)\gamma} \leq n^{\alpha-(1-\delta)\gamma}$$

as desired.

Setting up random variables and applying the structure chopping result Let $\mathbf{X} = \mathbf{A}_{\text{good}}$ so that

$$(\mathbf{A}, \mathbf{W}) = (1-p)(\mathbf{A}^{(0)}, \mathbf{W}^{(0)}) + p(\mathbf{X}, \mathbf{W}_{\text{good}}).$$

Let $\mathbf{Y} = \mathsf{Cond}(\mathbf{A}, \mathbf{W}) = (\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell)$ with $\mathbf{Y}_i = \mathsf{rbCond}(\mathbf{A}_{[1,i]}, \mathbf{W})$. Let $\mathbf{Y}' = \mathsf{Cond}(\mathbf{X}, \mathbf{W}_{\text{good}})$, and analogously define $\mathbf{Y}'_i$. Our overall claim then becomes that $\mathbf{H}_\infty(\mathbf{Y}') \geq n^{\delta\gamma/4}$.

We apply Lemma 6.1 to $\mathbf{X}$ with values of b, ℓ, k_y as above with parameter $\varepsilon = 2^{-k_1/4}$ and length parameter n_a to write $\mathbf{X}$ as a convex combination of sources $\mathbf{X} = \sum_{i=0}^{\ell} \lambda_i \mathbf{X}^{(i)}$ with the properties guaranteed there. When applying these properties below, we will repeatedly use that $n_a \leq n$.

For all $0 \leq i \leq \ell$, let $\mathbf{Y}^{(i)'} = \mathbf{Y}' \mid (\mathbf{X} \in \text{Supp}(\mathbf{X}^{(i)}))$ so that we can express $\mathbf{Y}'$ as the convex combination $\mathbf{Y}' = \sum_{i=0}^{\ell} \lambda_i \mathbf{Y}^{(i)'}$. Let $G = \{i \in [\ell] : \lambda_i > 0\}$. For $i \in G$, write $\mathbf{Y}_i^{(i)'}$ for the i-th block of $\mathbf{Y}^{(i)'}$, that is,

$$\mathbf{Y}_i^{(i)'} = \mathbf{Y}'_i \mid \left(\mathbf{X} \in \text{Supp}(\mathbf{X}^{(i)})\right).$$

We claim that for all $i \in G$, $\mathbf{H}_\infty(\mathbf{Y}_i^{(i)'}) \geq k_1$. This suffices to prove the desired claim as then for any $y \in \{0,1\}^m$, we will have that

$$\Pr[\mathbf{Y}' = y] = \sum_{i=0}^{\ell} \lambda_i \Pr[\mathbf{Y}^{(i)'} = y] \leq \lambda_0 + 2^{-k_1} \leq \varepsilon + 2^{-k_1} = 2^{-k_1/4} + 2^{-k_1} \leq 2^{-k_1/8}.$$

From this, we will be able to conclude that $\mathbf{H}_\infty(\mathbf{Y}') \geq k_1/8 = n^{\delta\gamma}/128 \geq n^{\delta\gamma/4}$ as desired.

Applying robust errorless condenser It remains to show that $\mathbf{H}_\infty(\mathbf{Y}_i^{(i)'}) \geq k_1$ for all $i \in G$. Fix any such $i \in G$. By Property 5 of Lemma 6.1, we have that $\mathbf{X}_{[1,i]} \mid (\mathbf{X} \in \text{Supp}\, \mathbf{X}^{(i)})$ is exactly the same distribution as $\mathbf{X}^{(i)}_{[1,i]}$. Therefore, we can express $\mathbf{X}_{[1,i]}$ as the convex combination

$$\mathbf{X}_{[1,i]} = \lambda_i \left(\mathbf{X}_{[1,i]} \mid \mathbf{X} \in \text{Supp}\left(\mathbf{X}^{(i)}\right)\right) + (1-\lambda_i)\left(\mathbf{X}_{[1,i]} \mid \left(\mathbf{X} \notin \text{Supp}\left(\mathbf{X}^{(i)}\right)\right)\right).$$

Combining this with the fact that $\mathbf{A} = p\mathbf{X} + (1-p)\mathbf{A}^{(0)}$, we obtain

$$\left(\mathbf{A}_{[1,i]}, \mathbf{W}\right) = p\lambda_i \left(\mathbf{X}_{[1,i]}, \mathbf{W}_{\text{good}} \mid \mathbf{X} \in \text{Supp}\left(\mathbf{X}^{(i)}\right)\right) + (1-p\lambda_i)\left(\widetilde{\mathbf{A}}_i, \widetilde{\mathbf{W}}_i\right)$$

for some distribution $(\widetilde{\mathbf{A}}_i, \widetilde{\mathbf{W}}_i)$.

We write, $\mathbf{Y}_i^{(i)'} = \mathsf{rbCond}\left(\mathbf{X}_{[1,i]}, \mathbf{W}_{\text{good}} \mid \left(\mathbf{X} \in \text{Supp}\left(\mathbf{X}^{(i)}\right)\right)\right)$ and apply Theorem 4.43 to the pair $(\mathbf{A}_{[1,i]}, \mathbf{W})$ to obtain the desired min-entropy bounds on $\mathbf{Y}_i^{(i)'}$. Note that $(\mathbf{A}_{[1,i]}, \mathbf{W})$ is samplable by size n^c circuits with postselection by size n^c nondeterministic circuits, as we can truncate the output of the sampler for $(\mathbf{A}, \mathbf{W})$ and keep the same postselection circuit.

We now show that we meet all the requirements required to apply Theorem 4.43.

1. First, we show that we meet the requirement on $p\lambda_i$, the weight with which $\mathbf{X}_{[1,i]}^{(i)}$ appears as a subsource of $\mathbf{A}_{[1,i]}$. Since $i \in G$, we have that

$$\begin{aligned} p \cdot \lambda_i &\geq p \cdot \frac{\varepsilon}{2n} \\ &\geq 2^{-n^{\gamma\delta}/32} \cdot \frac{2^{-k_1/4}}{2n} \\ &= 2^{-n^{\gamma\delta}/32} \cdot 2^{-(n^{\gamma\delta}/64)-1-\log(n)} \\ &\geq 2^{-n^{\gamma\delta}/16} \\ &= 2^{-k_1} \end{aligned}$$

   as desired.

2. We next show that $\mathbf{X}_{[1,i]}^{(i)}$ has sufficient entropy. Indeed, by Property 3 of Lemma 6.1, we have that

$$\mathbf{H}_\infty\left(\mathbf{X}_{[1,i]}^{(i)}\right) \geq \frac{k_y}{2} - \log(2n/\varepsilon) = \frac{k_y}{2} - \frac{k_1}{4} - \log(n) - 1 \geq \frac{k_y}{4}$$

   as desired.

3. We now lower bound the probability of each outcome of $\mathbf{X}_{[1,i]}^{(i)}$. Indeed, by Property 4 of Lemma 6.1, for any $x_i \in \text{Supp}\left(\mathbf{X}_{[1,i]}^{(i)}\right)$ we have that

$$\Pr\left[\mathbf{X}_{[1,i]}^{(i)} = x_i\right] \geq 2^{-k_y/2-b-\log(2n/\varepsilon)} = 2^{-(k/32)-(3k/8)-(k^\delta/64)-\log(n)-1} \geq 2^{-k/2}$$

   as desired.

4. Lastly, we check that $\mathbf{W}$ has sufficient min-entropy since by assumption, $\mathbf{H}_\infty(\mathbf{W}) \geq k$.

□

6.3 Constructing the Errorless Condenser

We finally construct our desired errorless condenser using the basic condenser from Lemma 6.3 and robust errorless block condenser from Theorem 4.43.

Proof of Theorem 4.42. We first construct the condenser and then analyze it. The construction repeatedly shrinks $\mathbf{W}$ using the basic condenser of Lemma 6.3, and at every stage applies the robust condenser of Theorem 4.43 to the resultant string, together with $\mathbf{W}$.

Construction On input $w \in \{0,1\}^n$, we proceed as follows.

- Let $\gamma = \gamma_{\text{in}}/4$ and let $0 < \delta \leq \frac{1}{4}$ be the universal constant from Lemma 6.3.
- Let $k = n^\gamma$ and for $i \geq 0$, let $n_i = \frac{n}{k^{i\cdot(1-\delta)}}$.
- Let t be the earliest index for which $n_t < k$. We immediately see that $k^\delta \leq n_t < k$.
- For $i \in [t]$, let $\mathsf{Cond}_i : \{0,1\}^{n_{i-1}} \times \{0,1\}^n \to \{0,1\}^{n_i}$ be the condenser from Lemma 6.3 for input entropy parameter k and for sources samplable by size n^{c_i} circuits with postselection by size n^{c_i} nondeterministic circuits, where we show how we recursively set c_i in the Explicitness part below. We check that $n_{i-1} \geq k$ since t is the earliest index j for which $n_j < k$; also by output length guarantees of Cond_i, we have that $n_i = n_{i-1}/k^{1-\delta}$, which by induction equals $n/k^{i\cdot(1-\delta)}$, the value that we defined n_i to be.
- Let $a_0 = w$ and for $i \in [t]$, iteratively define $a_i = \mathsf{Cond}_i(a_{i-1}, w)$.
- Let $\mathsf{rbCond}^* : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{m^*}$ be the $(2^{-m^*/4}, k_y, k_x/2, k_x, m^*/4)$ robust errorless block condenser from Theorem 4.43 for input parameter $k_y = n^{\gamma\delta/8}, k_x = n^{4\gamma}$ and for sources samplable by size n^{c^*} circuits with postselection by size n^{c^*} nondeterministic circuits, where $m^* = n^{\gamma\delta\cdot\delta'/8}$ (δ' is the universal constant from Theorem 4.43) and we set c^* to be large enough, dependent on c_t.
- For $i \in [t]$, let $y_i = \mathsf{rbCond}^*(a_i, w)$.
- Output $y = (y_1, \ldots, y_t)$, appropriately padded with zeroes to obtain the desired output length.

Explicitness We here argue that our construction is polynomial time computable.

We first show that $(\mathbf{A}_i, \mathbf{W})$ are sources that can be generated by polynomial sized circuits. More concretely, we can show that the pair $(\mathbf{A}_i, \mathbf{W})$ (defined below) is samplable by polynomial size circuits with postselection by the same size n^c nondeterministic circuit used for $\mathbf{W}$: we sample $\mathbf{W}$, and then compute $\mathbf{A}_i$ from it using $\mathsf{Cond}_1, \ldots, \mathsf{Cond}_i$. To establish this, we need to define the size parameter c_i, that is used to define Cond_i. We associate a related quantity $d_i \in \mathbb{N}$ that we also recursively define. We set $c_1 = c + 1$ and let $d_1 \in \mathbb{N}$ be such that Cond_1 can be implemented by size n^{d_1} circuits (since Cond_1 runs in polynomial time, such d_1 exists). With this, we let c_i be defined as $c_i = c_{i-1} + d_{i-1}$ so that one can generate the source $\mathbf{A}_{i-1}$ by size n^{c_i} circuits. With this, we let d_i be such that Cond_i can be implemented by size n^{d_i} circuits. We will set $c^* = c_t + d_t$ so that $\mathbf{A}_t$ can be generated by size n^{c^*} circuits.

We easily see that $t = O(1)$ is a constant, and so we iterate only a constant number of times. Using this, we have that each c_i will be a constant, and each Cond_i will be polynomial time computable. From this, we have that c^* will be a constant, and the resultant rbCond^* will be polynomial time computable, making our overall function explicit as desired.

Output length We now show that our output length is as claimed. As the length of y is $t \cdot n^{\gamma\delta\cdot\delta'/8} \leq n^{\gamma/2} \leq n^{\gamma_{\text{in}}/2}$, we can indeed pad the output to obtain the desired output length.

Setup for Entropy bound It remains to prove our output min-entropy bound and this is what we do for the rest of the proof. Towards proving this, we define the following distributions. The input distribution is $\mathbf{W}$. We let $\mathbf{A}_0 = \mathbf{W}$ and for $i \in [t]$, iteratively define $\mathbf{A}_i = \mathsf{Cond}_i(\mathbf{A}_{i-1}, \mathbf{W})$. Then, for $i \in [t]$, let $\mathbf{Y}_i = \mathsf{rbCond}^*(\mathbf{A}_i, \mathbf{W})$ and let $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_t)$. To prove our output min-entropy bound, it suffices to show that $\mathbf{H}_\infty(\mathbf{Y}) \geq n^{\gamma_{\text{in}}\cdot\delta_{\text{out}}}$ where we will set $\delta_{\text{out}} = \delta \cdot \delta'/256$.

While the above quantities correspond to the input, we also introduce the following quantities for the purpose of analysis. We will recursively partition the support of $\mathbf{W}$ into various sets. Towards that, we will introduce some sets $P_1, \ldots, P_t, Q_1, \ldots, Q_t \subset \{0,1\}^n$. They will maintain the invariant that for all $i \in [t]$,

$Q_1, \ldots, Q_i, P_i$ partition the support of $\mathbf{W}$. To construct Q_{i+1}, we will further split P_i, and then the set P_{i+1} will equal $P_i \setminus Q_{i+1}$.

We now show how to form such a partition of $\mathrm{Supp}(\mathbf{W})$. First, for any input $w \in \{0,1\}^n$, we let $a_i(w)$ be the quantity a_i when run on input w in the above algorithm. For the base case of our new definitions, let $\mathbf{X}_0 = \mathbf{A}'_0 = \mathbf{A}_0 (= \mathbf{W})$. Let $L_0 = P_0 = \mathrm{Supp}(\mathbf{W})$ and $M_0 = \emptyset$. Then, for $i \in [t]$, define the following quantities.

- Let $\mathbf{A}'_i = \mathbf{A}_i \mid (\mathbf{W} \in P_{i-1})$. This is the distribution that we want to condense from and is the 'remaining' distribution.
- Let $M_i = \{x \in \{0,1\}^{n_i} : \Pr[\mathbf{A}'_i = x] \geq 2^{-2k}\}$, the set of outcomes with medium weight under $\mathbf{A}'_i$.
- Let $L_i = \mathrm{Supp}(\mathbf{A}'_i) \setminus M_i$, the set of outcomes with light weight under $\mathbf{A}'_i$.
- Let $P_i = \{w \in P_{i-1} : a_i(w) \in L_i\}$ and let $p_i = \Pr[\mathbf{W} \in P_i]$. So, this is the part of the support of $\mathbf{W}$ that is given light weight under $\mathbf{A}'_i$.
- Let $Q_i = \{w \in P_{i-1} : a_i(w) \notin L_i\}$ and let $q_i = \Pr[\mathbf{W} \in Q_i]$. So, this is the part of the support of $\mathbf{W}$ that is given medium weight under $\mathbf{A}'_i$.
- Let $\mathbf{X}_i = \mathbf{A}'_i \mid (\mathbf{A}'_i \in L_i)$. Equivalently, $\mathbf{X}_i = \mathbf{A}_i \mid \mathbf{W} \in P_i$. This will be the good subsource of $\mathbf{A}_i$, on which Cond_{i+1} will act.
- Let $\mathbf{Y}'_i = \mathbf{Y}_i \mid (\mathbf{W} \in Q_i)$. This will be the good subsource of $\mathbf{Y}_i$, and we will apply rbCond^* on it.

Let $\tau = n^{\gamma \cdot \delta \cdot \delta'/32}$.

Defining the cutoff index j Let $j \in [t]$ be the earliest index such that $p_j \leq 2^{-\tau}$. Since $p_{i-1} \geq p_i$ for all $i \in [t]$, to show that such a j exists, it suffices to prove that $p_t \leq 2^{-\tau}$. By definition of L_t, and the fact that $n_t \leq k$, we have that $\Pr[\mathbf{A}'_t \in L_t] \leq 2^{n_t} \cdot 2^{-2k} \leq 2^{-k}$. From this, we directly compute that

$$p_t = p_{t-1} \cdot \Pr[\mathbf{A}'_t \in L_t] \leq 1 \cdot 2^{-k} \leq 2^{-\tau}$$

as desired where for the last inequality we used the fact that $k = n^\gamma \geq n^{\gamma \cdot \delta \cdot \delta'/32} = \tau$.

Overview of the remaining analysis We will prove lower bounds on entropy for $\mathbf{A}'_i, \mathbf{Y}'_i$ for $1 \leq i \leq j$, and using them obtain our lower bound on output entropy. We call M_i and L_i the medium and light parts of $\mathbf{A}'_i$; we first claim that the light part at each stage has enough entropy to satisfy the requirements of the basic condenser Cond_{i+1}, so that the conditioned sources stay entropic up to stage j. Formally,

Claim 6.4. *For every $0 \leq i < j$, we have $\mathbf{H}_\infty(\mathbf{A}'_{i+1}) \geq n^{\delta\gamma/4}$.*

Next, we will show that rbCond^* condenses from every medium part of non-negligible weight. Formally,

Claim 6.5. *For every $1 \leq i \leq j$ with $q_i \geq 2^{-\tau}$, we have $\mathbf{H}_\infty(\mathbf{Y}'_i) \geq m^*/4$.*

We will prove these claims one after the other towards the end. Let us first obtain the theorem assuming them.

Lower bounding output entropy We will lower bound the entropy of our output distribution $\mathbf{Y}$. Let $\text{Low}_q = \{i \in [j] : q_i < 2^{-\tau}\}$. As $Q_1, \ldots, Q_j, P_j$ partition the support of the input distribution $\mathbf{W}$, we have that $p_j + \sum_{i=1}^{j} q_i = 1$. Using this, and that for $i \notin \text{Low}_q$, Claim 6.5 shows $\mathbf{H}_\infty(\mathbf{Y}'_i) \geq m^*/4$, we obtain that for any outcome $y \in \{0,1\}^m$:

$$\begin{aligned}\Pr[\mathbf{Y} = y] &\leq p_j + \sum_{i \in \text{Low}_q} q_i + \sum_{i \notin \text{Low}_q} q_i \cdot 2^{-m^*/4} \\ &\leq 2^{-\tau} + |\text{Low}_q| \cdot 2^{-\tau} + 2^{-m^*/4} \\ &\leq (j+2) \cdot 2^{-\tau}\end{aligned}$$

where for the last inequality, we used the fact that $m^*/4 \geq \tau$.

From this, we conclude that

$$\mathbf{H}_\infty(\mathbf{Y}) \geq \tau - \log(j+2) \geq \tau/2 \geq n^{\gamma_{\text{in}}\delta\delta'/256}.$$

As δ, δ' are both universal constants, we set $\delta_{\text{out}} = \delta\delta'/256$, to obtain the desired result.

Lower bounding entropy of $\mathbf{A}'_i$ for $1 \leq i \leq j$ Next, we show that $\mathbf{H}_\infty(\mathbf{A}'_{i+1}) \geq n^{\delta\gamma/4}$ for $i < j$.

Proof of Claim 6.4. This holds for $i = 0$ since $\mathbf{A}'_0 = \mathbf{W}$, $p_0 = 1$, and by the guarantees of Cond_1 from Lemma 6.3. Fix $i \geq 1$. We will verify the hypotheses of the basic condenser Cond_{i+1} for $(\mathbf{A}_i, \mathbf{W})$, with good component $(\mathbf{A}_i, \mathbf{W}) \mid (\mathbf{W} \in P_i)$ of weight p_i, and show that its output guarantees yield our entropy bound. Also recall that $\mathbf{X}_i = \mathbf{A}_i \mid (\mathbf{W} \in P_i)$.

First, by assumption on j, we have that the weight of $\mathbf{X}_i$ in $\mathbf{A}_i$, p_i is such that

$$p_i \geq 2^{-\tau} = 2^{-n^{\gamma \cdot \delta \cdot \delta'/32}} \geq 2^{-n^{\delta\gamma}/32}.$$

and so we satisfy the weight requirement of Cond_{i+1}. We also easily satisfy the requirement that $\mathbf{H}_\infty(\mathbf{W}) \geq n^{\gamma_{\text{in}}} \geq n^\gamma = k$.

We now show that we satisfy the remaining requirement, that $\mathbf{H}_\infty(\mathbf{X}_i) \geq n^\gamma = k$. By definition of L_i, for all $x \in L_i$, we have that $\Pr[\mathbf{A}'_i = x] \leq 2^{-2k}$. Also, by definition of $\mathbf{A}'_i$, we see that

$$\Pr[\mathbf{A}'_i \in L_i] = \frac{p_i}{p_{i-1}} \geq p_i \geq 2^{-\tau}$$

where we used the fact that $i < j$ for the last inequality. As $\mathbf{X}_i = \mathbf{A}'_i \mid (\mathbf{A}'_i \in L_i)$, we conclude that

$$\mathbf{H}_\infty(\mathbf{X}_i) \geq 2k - \tau \geq k$$

as desired. Hence, we satisfy all the requirements to apply Lemma 6.3 and indeed conclude that $\mathbf{H}_\infty(\mathbf{A}'_{i+1}) \geq n^{\delta\gamma/4}$. □

Lower bounding entropy of $\mathbf{Y}'_i$ for $1 \leq i \leq j$ We next show entropic bounds on $\mathbf{Y}'_i$.

Proof of Claim 6.5. For that, we check that we satisfy the requirements to apply Theorem 4.43. This suffices to prove our claim since the output guarantees of rbCond^* will show that $\mathbf{H}_\infty(\mathbf{Y}'_i) \geq m^*/4$.

Recall that we apply rbCond^* to the source $(\mathbf{A}_i, \mathbf{W})$, which is samplable by size n^{c^*} circuits with postselection by size n^{c^*} nondeterministic circuits, and take $(\mathbf{A}_i, \mathbf{W}) \mid (\mathbf{W} \in Q_i)$ as the good subsource, which will have weight $\Pr[\mathbf{W} \in Q_i] = q_i$. Also recall that $\mathbf{A}_i \mid (\mathbf{W} \in Q_i) = \mathbf{A}'_i \mid (\mathbf{A}'_i \in M_i)$. The corresponding conditioned output is exactly $\mathbf{Y}'_i$, i.e., $\mathbf{Y}'_i = \mathbf{Y}_i \mid (\mathbf{W} \in Q_i)$.

To apply rbCond^*, we first require that $q_i \geq 2^{-m^*/4}$. By assumption on i, $q_i \geq 2^{-\tau}$. As $\tau = n^{\gamma \cdot \delta \cdot \delta'/32} \leq \frac{1}{4} \cdot n^{\gamma \cdot \delta \cdot \delta'/8} = m^*/4$, the desired lower bound on q_i follows.

We now show that we satisfy the min-entropy requirements for rbCond^*. We satisfy the requirement for $\mathbf{W}$, since by assumption, $\mathbf{H}_\infty(\mathbf{W}) \geq n^{4\gamma} = n^{\gamma_{\text{in}}}$. The remaining entropy requirement is that $\mathbf{H}_\infty(\mathbf{A}'_i \mid (\mathbf{A}'_i \in M_i)) \geq n^{\delta\gamma/8}$. From Claim 6.4, we have that $\mathbf{H}_\infty(\mathbf{A}'_i) \geq n^{\delta\gamma/4}$. Using the fact that $\Pr[\mathbf{A}'_i \in M_i] = \frac{q_i}{p_{i-1}} \geq q_i \geq 2^{-\tau} = 2^{-n^{\delta\delta'\gamma/32}}$, we obtain that

$$\mathbf{H}_\infty(\mathbf{A}'_i \mid (\mathbf{A}'_i \in M_i)) \geq n^{\delta\gamma/4} - n^{\delta\delta'\gamma/32} \geq n^{\delta\gamma/8}$$

as desired.

Recall that for all $x \in M_i$, we have that $\Pr[\mathbf{A}'_i = x] \geq 2^{-2\cdot n^\gamma}$. Therefore, for any such x, we infer that

$$\Pr[\mathbf{A}'_i = x \mid (\mathbf{A}'_i \in M_i)] = \frac{\Pr[\mathbf{A}'_i = x]}{\Pr[\mathbf{A}'_i \in M_i]} \geq \Pr[\mathbf{A}'_i = x] \geq 2^{-2\cdot n^\gamma} \geq 2^{-n^{4\gamma}/2}$$

satisfying the probability lower bound requirement of rbCond^*. Hence, we satisfy all requirements of rbCond^* and conclude that $\mathbf{H}_\infty(\mathbf{Y}'_i) \geq m^*/4$. □

□

7 Overview of the proofs of Theorem 4.43 and Theorem 4.41

In this section, we give an informal overview of the technique used to prove Theorems 4.41 and 4.43. This overview is not used in the formal proofs. Our approach builds on past work on extractors for samplable distributions [TV00; AASY16; BGDM23; Sha25b; BSS25; Sha25a], and it will be helpful to begin by recalling some of this work.

7.1 The extractor of [TV00] and the analysis of [BGDM23]

The extractor construction of [TV00]. In their seminal paper, Trevisan and Vadhan [TV00] constructed an extractor for samplable distributions for very high min-entropy, using the following construction.[16] For an input random variable $\mathbf{X}$ on $\{0,1\}^n$, the construction is as follows.

- The construction starts from a hard function $f\colon \{0,1\}^{d\log n} \to \{0,1\}$, for a sufficiently large constant d. It uses the low-degree extension (stated precisely in Theorem 4.34) to convert f into a d-variate polynomial $\widehat{f}\colon \mathbb{F}_q^d \to \mathbb{F}_q$ of degree roughly n, where $q = 2^{\eta n/d}$ for some very small constant $\eta > 0$. This choice gives that the input length of $\widehat{f}$ in bits is $d\log q = \eta n$.
- Let $\mathsf{Cond}(\mathbf{X})$ denote the first ηn bits of $\mathbf{X}$, identified with an element of $\mathbb{F}_q^d$.
- The construction also requires a $(k_1, k_2, \varepsilon_{\mathsf{TExt}})$-two-source extractor $\mathsf{TExt}\colon \mathbb{F}_q \times \{0,1\}^n \to \{0,1\}^m$. Although we aim for large m, for now think of $m = 1$.

The final construction is

$$\mathsf{Ext}(\mathbf{X}) = \mathsf{TExt}(\widehat{f}(\mathbf{W}), \mathbf{X}), \qquad \mathbf{W} = \mathsf{Cond}(\mathbf{X}).$$

Trevisan and Vadhan showed that if f is hard for exponential size Σ_6 circuits, then Ext is an extractor for samplable distributions with very high min-entropy, namely $k = (1-\eta')n$ for some constant $\eta' > 0$ which is sufficiently smaller than η/d.

[16] Some of the description here uses later observations by [AASY16; BSS25].

Improved analysis by [BGDM23]. Subsequent work by Ball, Dachman-Soled, Goldin and Mutreja [BGDM23] improved the analysis, and showed that this construction works, even if f is only assumed to be hard for exponential size *nondeterministic* circuits.

In this paper, we simplify and generalize the analysis of [BGDM23]. This allows us to tailor the construction and analysis to the case of low min-entropy, and is a major component in the proofs of Theorems 4.41 and 4.43. We start by reviewing the approach of [BGDM23].

High level description of the analysis of [BGDM23]. Assume that Ext is not a good extractor. Then there exists a samplable distribution $\mathbf{X}$ with $\mathbf{H}_\infty(\mathbf{X}) \geq k$, and an output $\sigma \in \{0,1\}^m$, such that $\Pr[\mathsf{Ext}(\mathbf{X}) = \sigma] > (1+\varepsilon)2^{-m}$. We will derive a contradiction by showing that there exists a small nondeterministic circuit C such that for every $x \in \mathbb{F}_q^d$, $C(x) = \widehat{f}(x)$, contradicting the hardness assumption.[17] Here, small means size $\mathrm{poly}(n)$ for a polynomial that does not depend on d. Note that q is exponential in n, and so the circuit C needs to have size that is much smaller than q.

In fact, by previous work, it is sufficient to construct such a C that only succeeds well on average. That is, $\Pr_{x \leftarrow \mathbb{F}_q^d}[C(x) = \widehat{f}(x)] \geq 15/16$, as such a circuit can be converted into a slightly larger circuit that computes f correctly on every input.[18]

Following [STV01], the circuit C will work as follows: when receiving a random point $x \leftarrow \mathbb{F}_q^d$ as input, C will choose another random point $z \leftarrow \mathbb{F}_q^d$, and consider the line $L \colon \mathbb{F}_q \to \mathbb{F}_q^d$ that passes through these two points, say with $L(0) = x$ and $L(1) = z$. Note that the univariate polynomial $\widehat{f} \circ L \colon \mathbb{F}_q \to \mathbb{F}_q$ is of degree roughly n, and $(\widehat{f} \circ L)(0) = \widehat{f}(x)$, which is our desired output.

The nondeterministic circuit C will guess a degree-n univariate polynomial $p \colon \mathbb{F}_q \to \mathbb{F}_q$. If we can devise a small nondeterministic circuit $T_L(p)$ that accepts the correct polynomial $p = \widehat{f} \circ L$ and rejects all other polynomials, then C can test whether p is correct and output $p(0)$.

In fact, following [STV01], the task of T_L can be relaxed. It is sufficient to devise a test $T_L(p)$ such that the correct polynomial is accepted, and the total number of accepted degree-n polynomials is significantly smaller than q. We do not explain this step in this overview, and the precise details appear in Section 8.2.

Following the previous discussion, we have reduced the task of constructing the circuit C to the task of devising a test $T_L(p)$ such that for a random line $\mathbf{L} \colon \mathbb{F}_q \to \mathbb{F}_q^d$, we have the following properties with high probability:

Efficiency: $T_{\mathbf{L}}$ can be computed by a small nondeterministic circuit.

Completeness: $T_{\mathbf{L}}$ accepts the correct polynomial $\widehat{f} \circ \mathbf{L}$.

Soundness: The number of degree-n polynomials that $T_{\mathbf{L}}$ accepts is significantly smaller than q.

7.2 A simpler test that can be generalized to low min-entropy

A significant technical contribution of [BGDM23] is devising such a test. In this paper, we give a simplified version of this test. (We will later need to modify the construction and generalize the test in order to handle low min-entropy.)

A simplified version of the test of [BGDM23]. We present our simplified version, starting with the setup of [BGDM23]. Recall that $\mathbf{W} = \mathsf{Cond}(\mathbf{X})$. For $w \in \mathbb{F}_q^d$ and $a \in \mathbb{F}_q$, we define:

$$J(w,a) = \Pr[\mathbf{W} = w,\ \mathsf{TExt}(a, \mathbf{X}) = \sigma].$$

[17] As $\widehat{f}$ is not a Boolean function, we consider *single-valued* nondeterministic circuits; see Definition 4.9.

[18] This is standard for deterministic circuits [GLRSW91], and there are subtleties in generalizing this argument for nondeterministic circuits. Jumping ahead, we mention that in the formal treatment in Section 8.2, we need to consider single-valued nondeterministic circuits. We ignore this technicality in this high-level overview.

This definition is made so that for every $w \in \mathbb{F}_q^d$,

$$J(w, \widehat{f}(w)) = \Pr[\mathbf{W} = w,\ \mathsf{TExt}(\widehat{f}(\mathbf{W}), \mathbf{X}) = \sigma] = \Pr[\mathbf{W} = w,\ \mathsf{Ext}(\mathbf{X}) = \sigma].$$

This gives that:

$$\mathop{\mathbb{E}}_{w \leftarrow \mathbb{F}_q^d}[J(w, \widehat{f}(w))] = \frac{1}{q^d} \sum_{w \in \mathbb{F}_q^d} \Pr[\mathbf{W} = w,\ \mathsf{Ext}(\mathbf{X}) = \sigma] = \frac{1}{q^d} \Pr[\mathsf{Ext}(\mathbf{X}) = \sigma].$$

We stress that the expectation is with respect to a *uniform* $w \leftarrow \mathbb{F}_q^d$. We then define

$$\mathrm{Test}_L(p) = \mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q}[J(L(t), p(t))], \qquad T_L(p) = 1 \iff \mathrm{Test}_L(p) \geq \frac{1}{q^d}(1 + \varepsilon/2)2^{-m}.$$

Efficiency. Efficiency follows because, writing $\mathbf{X} = \mathsf{Samp}(\mathbf{R})$, we can define a deterministic circuit $A_{L,p}(t, R)$ that accepts if $\mathsf{Cond}(\mathsf{Samp}(R)) = L(t)$ and $\mathsf{TExt}(p(t), \mathsf{Samp}(R)) = \sigma$. Its acceptance probability over uniform (t, R) is exactly $\mathrm{Test}_L(p)$. (A careful implementation of this approach appears in Section 8.1.) While C does not have time to go over all pairs (t, R) and compute $A_{L,p}(t, R)$, it can use the set-lower-bound protocol of Goldwasser and Sipser [GS86] (see Theorem 4.39 for a precise statement) to approximately count the fraction of pairs (t, R) that $A_{L,p}$ accepts. (It should be noted that the fraction is typically very small, and so it is critical to use a multiplicative approximation, as in the Goldwasser–Sipser protocol, rather than an additive approximation as given by standard random sampling.) A technicality that we ignore in this high-level overview is that the Goldwasser–Sipser protocol has a (multiplicative) approximation error.

Completeness. Completeness follows directly from the fact that a random line is a good averaging sampler. More precisely, for the correct polynomial $p(t) = (\widehat{f} \circ L)(t)$,

$$\mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q}[J(L(t), p(t))] = \mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q}[J(L(t), \widehat{f}(L(t)))],$$

and for a random line $\mathbf{L}$, with high probability this quantity is close to

$$\mathop{\mathbb{E}}_{w \leftarrow \mathbb{F}_q^d}[J(w, \widehat{f}(w))] = \frac{\Pr[\mathsf{Ext}(\mathbf{X}) = \sigma]}{q^d} > \frac{(1 + \varepsilon)2^{-m}}{q^d}.$$

We remark that we need the two quantities to be very close because the latter is tiny and can be smaller than $\frac{1}{q^d}$. To compensate for that, one must require that $\mathbf{H}_\infty(\mathbf{W})$ is very large (say $\mathbf{H}_\infty(\mathbf{W}) > (d-1)\log q + m$, with sufficient slack). To achieve this for the function $\mathbf{W} = \mathsf{Cond}(\mathbf{X})$ that is considered by [TV00] (namely, $\mathsf{Cond}(\mathbf{X})$ is a prefix of $\mathbf{X}$), it suffices to require that $\mathbf{H}_\infty(\mathbf{X}) > n - (\log q - m)$, again with sufficient slack, and so this approach can only produce extractors for samplable distributions for very high min-entropy $k \gg n/2$. The focus of this paper is constructing extractors for much lower min-entropy k, and so we will need to modify the construction and analysis.

Soundness. For every $w \in \mathbb{F}_q^d$ with $\Pr[\mathbf{W} = w] > 0$, we define

$$R_w = \{a \in \mathbb{F}_q : \Pr[\mathsf{TExt}(a, \mathbf{X}) = \sigma \mid \mathbf{W} = w] > 2^{-m} + \varepsilon_{\mathsf{TExt}}\}.$$

We say that $w \in \mathbb{F}_q^d$ is *valid* if $\Pr[\mathbf{W} = w] > 0$ and $\mathbf{H}_\infty(\mathbf{X} \mid \mathbf{W} = w) \geq k_2$. That is, if the source $(\mathbf{X} \mid \mathbf{W} = w)$ meets the requirement that TExt makes of its second source. For zero-probability points w, we set $R_w = \varnothing$ and regard w as invalid. For every valid w, we have that $|R_w| \leq 2^{k_1}$. This is because

otherwise, the two independent sources $\mathbf{X}_1$, uniform on R_w, and $\mathbf{X}_2$, distributed as $(\mathbf{X} \mid \mathbf{W} = w)$, violate the guarantee of the two-source extractor, as

$$\Pr[\mathsf{TExt}(\mathbf{X}_1, \mathbf{X}_2) = \sigma] - \Pr[\mathbf{U}_m = \sigma] > \varepsilon_{\mathsf{TExt}}.$$

In order to bound the number of degree-n polynomials p such that $\mathrm{Test}_L(p) \geq (1 + \varepsilon/2)2^{-m}/q^d$, we will show that every such p has a sufficiently large fraction of $t \in \mathbb{F}_q$ such that $L(t)$ is valid and $p(t) \in R_{L(t)}$. (This is sufficient, as by the list-recovery properties of Reed–Solomon codes (see, e.g., [Sud97]), applied to the lists at valid points, the number of such polynomials is small.)

We divide the points $t \in \mathbb{F}_q$ into three categories:

- t such that $L(t)$ is not valid.
- t such that $L(t)$ is valid and $p(t) \notin R_{L(t)}$.
- t such that $L(t)$ is valid and $p(t) \in R_{L(t)}$.

The quantity $\mathrm{Test}_L(p) = \mathbb{E}_{t \leftarrow \mathbb{F}_q}[J(L(t), p(t))]$ averages the contributions of points in these three categories. By an appropriate choice of k_2, the probability that $\mathbf{W}$ is invalid can be made sufficiently small compared with $\varepsilon 2^{-m}$. For the first category,

$$\mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q}\left[J(L(t), p(t))\mathbf{1}_{\{L(t) \text{ is invalid}\}}\right] \leq \mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q}\left[\Pr[\mathbf{W} = L(t)]\mathbf{1}_{\{L(t) \text{ is invalid}\}}\right].$$

The expectation of the right-hand side over a random line $\mathbf{L}$ is

$$\frac{\Pr[\mathbf{W} \text{ is invalid}]}{q^d} \ll \frac{\varepsilon 2^{-m}}{q^d}.$$

By Markov's inequality, with high probability over $\mathbf{L}$, the contribution from the first category is small compared with $\varepsilon 2^{-m}/q^d$.

For the second category, by definition,

$$\begin{aligned} J(L(t), p(t)) &= \Pr[\mathbf{W} = L(t),\ \mathsf{TExt}(p(t), \mathbf{X}) = \sigma] \\ &= \Pr[\mathsf{TExt}(p(t), \mathbf{X}) = \sigma \mid \mathbf{W} = L(t)] \Pr[\mathbf{W} = L(t)] \\ &\leq (2^{-m} + \varepsilon_{\mathsf{TExt}}) \Pr[\mathbf{W} = L(t)]. \end{aligned}$$

Consequently, the total contribution from the second category is bounded by

$$\mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q}\left[J(L(t), p(t))\mathbf{1}_{\{t \text{ is in the second category}\}}\right] \leq (2^{-m} + \varepsilon_{\mathsf{TExt}}) \mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q} \Pr[\mathbf{W} = L(t)].$$

The sampling property of a random line gives that, with high probability,

$$\mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q} \Pr[\mathbf{W} = \mathbf{L}(t)] \quad \text{is close to} \quad \mathop{\mathbb{E}}_{w \leftarrow \mathbb{F}_q^d} \Pr[\mathbf{W} = w] = \frac{1}{q^d}.$$

By picking $\varepsilon_{\mathsf{TExt}} \ll \varepsilon 2^{-m}$ and controlling the sampling error, the total contribution from the second category is at most the baseline $2^{-m}/q^d$ plus an amount that is small compared with $\varepsilon 2^{-m}/q^d$.

However, as $\mathrm{Test}_L(p) \geq (1 + \varepsilon/2)2^{-m}/q^d$, the third category must contribute an amount of order $\varepsilon 2^{-m}/q^d$. For every t, we have

$$J(L(t), p(t)) \leq \Pr[\mathbf{W} = L(t)] \leq 2^{-\mathbf{H}_\infty(\mathbf{W})}.$$

We choose the min-entropy deficiency $d \log q - \mathbf{H}_\infty(\mathbf{W})$, the output length m, and k_1 sufficiently small relative to $\log q$. Consequently, there must be a sufficiently large fraction of t such that $L(t)$ is valid and $p(t) \in R_{L(t)}$. With an appropriate choice of parameters, list recovery then bounds the number of accepted polynomials by a quantity significantly smaller than q.

Remark 7.1 (Comparison with the test of [BGDM23])**.** *Ball et al. [BGDM23] use a test based on the following quantity:*

$$\mathrm{Test}_L^{\mathrm{BDGM}}(p) = \Pr[\mathsf{TExt}(p(L^{-1}(\mathbf{W})), \mathbf{X}) = \sigma \mid \mathbf{W} \in L].$$

This test is more complicated than the one we consider here. For a nondegenerate L with $\Pr[\mathbf{W} \in L] > 0$, we have

$$\begin{aligned}
\mathrm{Test}_L(p) &= \frac{1}{q} \sum_{t \in \mathbb{F}_q} J(L(t), p(t)) \\
&= \frac{1}{q} \sum_{t \in \mathbb{F}_q} \Pr[\mathbf{W} = L(t),\ \mathsf{TExt}(p(t), \mathbf{X}) = \sigma] \\
&= \frac{1}{q} \Pr[\mathbf{W} \in L,\ \mathsf{TExt}(p(L^{-1}(\mathbf{W})), \mathbf{X}) = \sigma] \\
&= \frac{\Pr[\mathbf{W} \in L]}{q} \Pr[\mathsf{TExt}(p(L^{-1}(\mathbf{W})), \mathbf{X}) = \sigma \mid \mathbf{W} \in L] \\
&= \frac{\Pr[\mathbf{W} \in L]}{q} \cdot \mathrm{Test}_L^{\mathrm{BDGM}}(p).
\end{aligned}$$

Our test avoids the normalization by $\Pr[\mathbf{W} \in L]$. The analysis used by [BGDM23] is significantly more complicated than the one used here, and is difficult to generalize to the low min-entropy case.

7.3 Extending the construction and test to the case of low min-entropy

Ball, Shaltiel and Silbak [BSS25] extended the approach of [TV00] to lower min-entropy. Rather than taking $\mathsf{Cond}(\mathbf{X})$ to be a prefix of $\mathbf{X}$, they use an errorless condenser for samplable distributions (as in Definition 4.15). Unfortunately, the errorless condensers constructed in that paper and the subsequent paper by Shaltiel [Sha25a] do not guarantee that $\mathsf{Cond}(\mathbf{X})$ has small min-entropy deficiency, and are therefore not compatible with the argument above.

In order to use the approach above, we modify the construction as follows: given $\mathbf{X}$, we compute $\mathbf{Y} = \mathsf{Cond}(\mathbf{X})$ of length n_y, and then use a strong seeded extractor $\mathsf{SExt}\colon \{0,1\}^{n_y} \times [D] \to \{0,1\}^{d \log q}$, with $D = \mathrm{poly}(n)$, that is set up to extract randomness from $\mathbf{Y} = \mathsf{Cond}(\mathbf{X})$. Here, we choose $q \gg 2^{k_1}$ with $d \log q$ sufficiently smaller than the min-entropy of $\mathbf{Y}$, where k_1 is the min-entropy required by TExt in its first source. For every $i \in [D]$, we define $\mathbf{W}_i = \mathsf{SExt}(\mathbf{Y}, i)$ and then

$$\mathsf{Ext}(\mathbf{X}) = \mathsf{TExt}\big(\widehat{f}(\mathbf{W}_1) \,\|\, \cdots \,\|\, \widehat{f}(\mathbf{W}_D), \mathbf{X}\big).$$

A complication is that we currently do not have explicit two-source extractors TExt with the low-entropy and low-error parameters needed here, and so, in the proofs of Theorems 4.41 and 4.43, we use other objects. Specifically, in Theorem 4.43, we use a low-error two-source condenser by Ben-Aroya, Cohen, Doron and Ta-Shma [BCDT19], and in Theorem 4.41, we use a low-error two-source extractor by Li [Li15], which makes the additional requirement that its second source is a block-wise source. We briefly discuss these modifications below (Sections 7.4 and 7.5).

The intuition behind this construction is that the guarantee of a strong seeded extractor gives that there exists $i^* \in [D]$ such that $\mathbf{W}_{i^*} = \mathsf{SExt}(\mathbf{Y}, i^*)$ is close to uniform. For simplicity of notation, assume without loss of generality that $i^* = D$. The hope is that we can try to repeat the argument above, with $\mathbf{W}_D$ playing the role of the condenser output in the previous argument. Indeed, we now have that $\mathbf{W}_D$ is close to uniform.

A major problem is that it seems that when the test checks whether a polynomial p is $\widehat{f} \circ L$ using the last coordinate D, it will need to somehow compute $\widehat{f}$ on the remaining $D-1$ coordinates (which it cannot do). To overcome this problem, we define the following generalization of $J(w,a)$:

$$J(w,a) = \sum_{y:\mathsf{SExt}(y,D)=w} \max_{v \in \mathbb{F}_q^{D-1}} \Pr[\mathbf{Y} = y,\ \mathsf{TExt}((v,a),\mathbf{X}) = \sigma].$$

The placement of the maximum in this expression allows v to depend on y, but not on the specific $x \leftarrow \mathbf{X}$. Choosing $v = \widehat{f}(\mathbf{W}_1) \parallel \cdots \parallel \widehat{f}(\mathbf{W}_{D-1})$ in the maximum gives

$$J(w, \widehat{f}(w)) \geq \Pr[\mathbf{W}_D = w,\ \mathsf{Ext}(\mathbf{X}) = \sigma], \tag{4}$$

which was key for the previous completeness analysis.

Soundness. The main difficulty in extending the soundness argument is that the guarantee that $\mathbf{W}_D$ is close to uniform is somewhat different from the previous guarantee that $\mathbf{W}$ has high min-entropy. In particular, $\mathbf{W}_D$ may have elements with large probability. To keep the number of seeds D, and hence the running time, polynomial in n, we choose $\varepsilon_{\mathsf{SExt}} = 1/n^{O(1)}$. We aim for $m \gg \log n$, and therefore want to allow $\varepsilon_{\mathsf{SExt}}$ to be much larger than 2^{-m}. In this regime, elements t with large $\Pr[\mathbf{W}_D = L(t)]$ can completely dominate the outcome of the averaging test. Consequently, a large average need not imply agreement with the candidate lists at sufficiently many points.

In this scenario, it is better to change the test from averaging $J(L(t),p(t))$ over $t \leftarrow \mathbb{F}_q$ to a "sum of thresholds" as follows:

$$\begin{aligned}\mathrm{Test}_L(p) &= \mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q}\left[\mathbf{1}_{\left\{J(L(t),p(t)) \geq \frac{(1+\varepsilon/2)2^{-m}}{q^d}\right\}}\right] \\ &= \Pr_{t \leftarrow \mathbb{F}_q}\left[J(L(t),p(t)) \geq \frac{(1+\varepsilon/2)2^{-m}}{q^d}\right].\end{aligned}$$

The test accepts if a sufficiently large constant fraction of the points on the line pass the threshold. The advantage is that this test gives "equal weight" to all points on the line, so that points w with atypically large $\Pr[\mathbf{W}_D = w]$ do not get to significantly change the value of the test.

We can now imitate the previous soundness analysis. Whereas previously we used the fact that conditioning on the condenser output $\mathbf{W}$ fixes the first input of TExt, we will now use that conditioning on $\{\mathbf{Y} = y\}$ fixes the first input of TExt. Each valid value y gives a small list of candidate first inputs to TExt, and hence a small list of their last coordinates. Different values y that map to the same w may give different lists.

Here we use the unconditional closeness of $\mathbf{W}_D$ to uniform: for all but a small fraction of points w, measured under the uniform distribution on $\mathbb{F}_q^d$, we have $\Pr[\mathbf{W}_D = w] \approx 1/q^d$. Moreover, $J(w,a) \leq \Pr[\mathbf{W}_D = w]$. Outside a small exceptional set of points w, a value a can pass the threshold only if it belongs to the lists of sufficiently much of the probability mass mapping to w. Since each individual list is small, a counting argument shows that only a small number of values a can pass. We can then apply the list-recovery argument as before.

Completeness. We can guarantee that for the good seed $i^* = D$, $\mathbf{W}_D$ is close to uniform even conditioned on $\mathsf{Ext}(\mathbf{X}) = \sigma$. This helps the completeness argument: recall from Equation (4) that

$$\begin{aligned}J(w, \widehat{f}(w)) &\geq \Pr[\mathbf{W}_D = w,\ \mathsf{Ext}(\mathbf{X}) = \sigma] \\ &= \Pr[\mathsf{Ext}(\mathbf{X}) = \sigma] \cdot \Pr[\mathbf{W}_D = w \mid \mathsf{Ext}(\mathbf{X}) = \sigma].\end{aligned}$$

We now have a handle on both probabilities: the first is larger than $(1+\varepsilon)2^{-m}$ by assumption, and the second is close to $1/q^d$ for most w under the uniform distribution. This will be used to ensure that with high probability over $\mathbf{L}$, the correct polynomial passes the threshold at most points on the line.

Efficiency. This "sum of thresholds" test can also be implemented by a nondeterministic circuit using an additional layer of the Goldwasser–Sipser protocol. More specifically, we first apply Goldwasser–Sipser to obtain a nondeterministic circuit $\mathrm{Thr}(w, a)$ that checks whether $J(w, a) > (1 + \varepsilon/2)2^{-m}/q^d$, up to a small gap around this threshold. We then apply Goldwasser–Sipser again to check whether the fraction of $t \in \mathbb{F}_q$ for which $\mathrm{Thr}(L(t), p(t)) = 1$ is sufficiently large.

The proofs of both Theorems 4.41 and 4.43 follow this outline. However, there are some special modifications that we make for each of the two settings.

7.4 Special modifications for Theorem 4.43

In Theorem 4.43, we aim to construct a robust block condenser rather than an extractor. Such an object receives $\mathbf{Y}$ in addition to $\mathbf{X}$, and the construction is

$$\mathsf{rbCond}(\mathbf{Y}, \mathbf{X}) = \mathsf{2Cond}\big(\widehat{f}(\mathbf{W}_1) \parallel \cdots \parallel \widehat{f}(\mathbf{W}_D), \mathbf{X}\big),$$

where $\mathbf{W}_i = \mathsf{SExt}(\mathbf{Y}, i)$ for every $i \in [D]$, as before, and $\mathsf{2Cond}$ is the low-error two-source condenser of [BCDT19].

An additional complication is that, in a robust block condenser, the test has access to the distribution $(\mathbf{Y}, \mathbf{X})$, whereas some of the entropy guarantees apply to a subsource $(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}})$. This dictates some changes in the test and the argument that are explained in detail in Section 9.

7.5 Special modifications for Theorem 4.41

In Theorem 4.41, we aim to construct a multiplicative extractor for low min-entropy with large output length. For this purpose, the argument described above requires a two-source extractor TExt with error $2^{-n^{\Omega(1)}}$. Such explicit constructions are not known for low min-entropy, and we therefore replace TExt with a two-source extractor IExt by Li [Li15] which achieves these parameters, but requires the second source to be a block-wise source. We will imitate the approach of Shaltiel [Sha25a] that constructed such an extractor under stronger hardness assumptions.

More specifically, we will apply the condenser $\mathsf{Cond}(\mathbf{X})$ twice as follows: We set $\mathbf{Y} = \mathsf{Cond}(\mathsf{Cond}(\mathbf{X}))$, $\mathbf{W}_i = \mathsf{SExt}(\mathbf{Y}, i)$ for $i \in [D]$, and

$$\mathsf{Ext}(\mathbf{X}) = \mathsf{IExt}\big(\widehat{f}(\mathbf{W}_1) \parallel \cdots \parallel \widehat{f}(\mathbf{W}_D), (\mathsf{Cond}(\mathbf{X}), \mathbf{X})\big).$$

The high-level intuition is that, for most values y of $\mathbf{Y}$, the conditional distribution $(\mathsf{Cond}(\mathbf{X}), \mathbf{X}) \mid \mathbf{Y} = y$ is close to a block-wise source: $\mathbf{Y}$ is significantly shorter than the min-entropy of $\mathsf{Cond}(\mathbf{X})$, and $\mathsf{Cond}(\mathbf{X})$ is significantly shorter than the min-entropy of $\mathbf{X}$. This generalizes the construction described above, and precise details are given in Section 10.

8 Joint preparations

In this section, we set up some machinery that will be used in the proofs of both Theorems 4.41 and 4.43.

8.1 A nondeterministic circuit for counting the sum of maximums

In this subsection, we prove a general lemma that will be used to argue that the tests defined in Sections 9 and 10 can be implemented by a small nondeterministic circuit, establishing the efficiency property of these tests.

Lemma 8.1 (Counting a sum of maximums)**.** *Let $C(u, y, v, r)$ be a nondeterministic circuit, with inputs of length n_u, n_y, n_v, n_r, respectively, and define $S(u)$ as follows:*

$$S(u) \coloneqq \sum_{y \in \{0,1\}^{n_y}} \max_{v \in \{0,1\}^{n_v}} \Pr_{r \leftarrow \{0,1\}^{n_r}} [C(u, y, v, r) = 1].$$

For every rational $0 < \Theta \leq 2^{n_y}$ and $0 < \eta \leq 1/2$, there is a nondeterministic circuit $V(u)$ such that for every $u \in \{0,1\}^{n_u}$,

$$\begin{aligned} S(u) \geq \Theta \quad &\implies \quad V(u) = 1, \\ S(u) \leq (1-\eta)\Theta \quad &\implies \quad V(u) = 0. \end{aligned}$$

Its size is polynomial in the size of C, the input lengths, the bit length of Θ, and $1/\eta$.

Proof. Set $M_u(y) = \max_v \Pr_r[C(u, y, v, r) = 1]$. Choose a threshold grid

$$G = \{j 2^{-n_r} : 1 \leq j \leq 2^{n_r}\}.$$

Note that $M_u(y)$ is an integer multiple of 2^{-n_r}.

For every u, y, v, define the nondeterministic circuit $C_{u,y,v}(r) \coloneqq C(u, y, v, r)$. Apply Theorem 4.39 with $e^\lambda = 1 + \eta/8$ to obtain a nondeterministic circuit $A(u, y, v, \tau)$ such that for every u, y, v and every $\tau \in G$,

$$\begin{aligned} \Pr_r[C_{u,y,v}(r) = 1] \geq \tau \quad &\implies \quad A(u, y, v, \tau) = 1, \\ \Pr_r[C_{u,y,v}(r) = 1] < \frac{\tau}{1+\eta/8} \quad &\implies \quad A(u, y, v, \tau) = 0. \end{aligned}$$

Let $B(u, y, \tau)$ be the nondeterministic circuit that guesses v and simulates $A(u, y, v, \tau)$. For each fixed u, y, $B(u, y, \tau)$ accepts every $\tau \in G$ satisfying $\tau \leq M_u(y)$, and rejects every $\tau \in G$ satisfying $\tau > (1 + \eta/8)M_u(y)$. Consequently,

$$M_u(y) \leq \Pr_{\tau \leftarrow G}[B(u, y, \tau) = 1] \leq (1 + \eta/8)M_u(y).$$

Averaging over $y \leftarrow \{0,1\}^{n_y}$ gives, for every u,

$$2^{-n_y} S(u) \leq \Pr_{y \leftarrow \{0,1\}^{n_y},\, \tau \leftarrow G}[B(u, y, \tau) = 1] \leq (1 + \eta/8) 2^{-n_y} S(u).$$

Thus the two cases in the statement imply, respectively,

$$\begin{aligned} S(u) \geq \Theta \quad &\implies \quad \Pr_{y \leftarrow \{0,1\}^{n_y},\, \tau \leftarrow G}[B(u, y, \tau) = 1] \geq 2^{-n_y}\Theta, \\ S(u) \leq (1-\eta)\Theta \quad &\implies \quad \Pr_{y \leftarrow \{0,1\}^{n_y},\, \tau \leftarrow G}[B(u, y, \tau) = 1] \leq (1 - 7\eta/8) 2^{-n_y}\Theta. \end{aligned}$$

For every u, define the nondeterministic circuit $B_u(y, \tau) \coloneqq B(u, y, \tau)$, where $y \in \{0,1\}^{n_y}$ and $\tau \in G$ are its inputs. Apply Theorem 4.39 to B_u with threshold $2^{-n_y}\Theta$ and the same λ to obtain a nondeterministic circuit $V(u)$. It satisfies the required guarantees because

$$(1 - 7\eta/8) 2^{-n_y}\Theta < \frac{2^{-n_y}\Theta}{1 + \eta/8}.$$

Both applications have polynomial size in the stated parameters. □

8.2 Reconstruction from a polynomial test

As we explained in Section 7, the proofs of Theorems 4.41 and 4.43 will use the overall approach of [STV01; BGDM23]. More specifically, in both proofs, the construction converts the supplied hard function into its low-degree extension $\widehat{f}\colon \mathbb{F}_q^d \to \mathbb{F}_q$, which is a polynomial of total degree at most $\hat{h}$ for some parameter $\hat{h}$. The construction of the required object will use $\widehat{f}$ as a component. The correctness argument assumes that the constructed object fails, and obtains a contradiction by constructing a small nondeterministic circuit C that computes $\widehat{f}$. In both proofs, this will be done in three steps:

- Design a small nondeterministic circuit $\mathrm{Thr}(w, a)$ and use it to define a test $\mathrm{Test}_L(p)$. (In the terminology of Section 7, $\mathrm{Thr}(w, a)$ checks whether $J(w, a)$ is larger than some threshold.)
- Prove completeness and soundness with high probability over a random line $\mathbf{L}\colon \mathbb{F}_q \to \mathbb{F}_q^d$.
- Use the test to construct the desired nondeterministic circuit C that computes $\widehat{f}$.

The next lemma states the (by now standard) third step in this approach. More specifically, given a small nondeterministic circuit Thr, suppose that with high probability over a random line $\mathbf{L}$, the induced test $\mathrm{Test}_{\mathbf{L}}(p)$ satisfies completeness ($\mathrm{Test}_{\mathbf{L}}(\widehat{f} \circ \mathbf{L}) \geq 7/8$) and soundness (the number of p of degree at most $\hat{h}$ with $\mathrm{Test}_{\mathbf{L}}(p) > 3/4$ is small). The lemma gives a small single-valued nondeterministic circuit C that computes $\widehat{f}$. This will be used to contradict the hardness assumption.

Lemma 8.2 (Reconstruction from a polynomial test)**.** *Let q be a power of two, let $d \geq 1$ and $\hat{h} \geq 0$ be integers, and let $\widehat{f}\colon \mathbb{F}_q^d \to \mathbb{F}_q$ be a polynomial of total degree at most $\hat{h}$. Let $\mathrm{Thr}(w, a)$ be a nondeterministic circuit of size S, with inputs $w \in \mathbb{F}_q^d$ and $a \in \mathbb{F}_q$. For every nondegenerate line L and every $p \in \mathbb{F}_q[t]$ of degree at most $\hat{h}$, define*

$$\mathrm{Test}_L(p) = \Pr_{t \leftarrow \mathbb{F}_q}[\mathrm{Thr}(L(t), p(t)) = 1].$$

Set $\mathrm{Test}_L(p) = 0$ for degenerate L, and define

$$\mathcal{P}_L = \{p \in \mathbb{F}_q[t] : \deg p \leq \hat{h},\ \mathrm{Test}_L(p) > 3/4\}.$$

Suppose that

$$\Pr_{\mathbf{L} \leftarrow \mathsf{RandLine}}\left[\mathrm{Test}_{\mathbf{L}}(\widehat{f} \circ \mathbf{L}) \geq \frac{7}{8}, \quad |\mathcal{P}_{\mathbf{L}}| \leq \frac{q}{32(\hat{h}+1)}\right] \geq \frac{31}{32}.$$

There is a single-valued nondeterministic circuit C of size $\mathrm{poly}(S, \hat{h}, d, \log q)$ such that

$$\mathrm{ACCEPT}_C(x) = \{\widehat{f}(x)\} \qquad \text{for every } x \in \mathbb{F}_q^d.$$

Proof of Lemma 8.2

We follow the reconstruction strategy of Sudan–Trevisan–Vadhan [STV01] and Ball et al. [BGDM23]: pin a polynomial using one correct evaluation, and then correct the remaining errors.

8.2.1 Pinning the correct polynomial

For every nondegenerate L and every $p \in \mathbb{F}_q[t]$ of degree at most $\hat{h}$, define the nondeterministic circuit $D_{L,p}(t) \coloneqq \mathrm{Thr}(L(t), p(t))$. Apply Theorem 4.39 with threshold $7/8$ and $e^{\zeta} = 9/8$ to obtain a nondeterministic circuit $T(L, p)$ such that

$$\begin{aligned} \mathrm{Test}_L(p) \geq 7/8 \quad &\Longrightarrow \quad T(L, p) = 1, \\ \mathrm{Test}_L(p) \leq 3/4 \quad &\Longrightarrow \quad T(L, p) = 0, \end{aligned}$$

where $3/4 < 7/9 = (7/8)e^{-\zeta}$. The inputs encode L by its endpoints and p by its coefficients. Set $T(L,p) = 0$ for degenerate L. The size of T is polynomial in $S, \hat{h}, d, \log q$.

Definition 8.3 (The circuit $C_{z,b}$)**.** *For every $z \in \mathbb{F}_q^d$ and $b \in \mathbb{F}_q$, define the nondeterministic circuit $C_{z,b}$ as follows. On input $x \in \mathbb{F}_q^d$:*

(1) Set $L_{x,z}(t) = x + t(z - x)$.

(2) Guess a polynomial $p \in \mathbb{F}_q[t]$ of degree at most $\hat{h}$ and a witness wit_T for T.

(3) If $T(L_{x,z}, p; \mathrm{wit}_T) = 1$ and $p(1) = b$, accept and output $p(0)$; otherwise, reject.

Thus,

$$\mathrm{ACCEPT}_{C_{z,b}}(x) = \{p(0) : \deg p \le \hat{h},\ p(1) = b,\ T(L_{x,z}, p) = 1\}.$$

Lemma 8.4 (Pinning)**.** *Under the hypotheses of Lemma 8.2,*

$$\Pr_{x,z \leftarrow \mathbb{F}_q^d}\Big[\mathrm{ACCEPT}_{C_{z,\widehat{f}(z)}}(x) = \{\widehat{f}(x)\}\Big] \ge \frac{15}{16}.$$

Consequently, there is $z^* \in \mathbb{F}_q^d$ such that

$$\Pr_{x \leftarrow \mathbb{F}_q^d}\Big[\mathrm{ACCEPT}_{C_{z^*,\widehat{f}(z^*)}}(x) = \{\widehat{f}(x)\}\Big] \ge \frac{15}{16}.$$

The size of $C_{z,b}$ is polynomial in $S, \hat{h}, d, \log q$, with an absolute polynomial bound.

Proof. For every nondegenerate L and affine bijection $\psi\colon \mathbb{F}_q \to \mathbb{F}_q$,

$$\mathrm{Test}_{L\circ\psi}(p \circ \psi) = \Pr_{t \leftarrow \mathbb{F}_q}[\mathrm{Thr}(L(\psi(t)), p(\psi(t))) = 1] = \mathrm{Test}_L(p).$$

Hence

$$\mathcal{P}_{L\circ\psi} = \{p \circ \psi : p \in \mathcal{P}_L\}.$$

Fix a nondegenerate L with $|\mathcal{P}_L| \le q/(32(\hat{h}+1))$. Each $p \ne \widehat{f} \circ L$ agrees with $\widehat{f} \circ L$ at at most $\hat{h}$ points. Therefore,

$$\Pr_{t \leftarrow \mathbb{F}_q}\Big[\exists p \in \mathcal{P}_L \setminus \{\widehat{f} \circ L\} : p(t) = \widehat{f}(L(t))\Big] \le \frac{|\mathcal{P}_L|\hat{h}}{q} \le \frac{1}{32}.$$

Conditioned on $x \ne z$, we can imagine first choosing a uniform geometric line, fixing any parametrization L, and then choosing a uniformly random affine bijection $\psi\colon \mathbb{F}_q \to \mathbb{F}_q$. Set

$$x = L(\psi(0)), \qquad z = L(\psi(1)).$$

Then $L_{x,z} = L \circ \psi$. Since $\psi(1)$ is uniform in $\mathbb{F}_q$, the invariance of $\mathcal{P}_L$ under reparametrization and the preceding bound give

$$\Pr_{x,z}\begin{bmatrix} x \ne z, \quad |\mathcal{P}_{L_{x,z}}| \le q/(32(\hat{h}+1)), \\ \exists p \in \mathcal{P}_{L_{x,z}} \setminus \{\widehat{f} \circ L_{x,z}\} : p(1) = \widehat{f}(z) \end{bmatrix} \le \frac{1}{32}.$$

By the hypothesis of Lemma 8.2, with probability at least $31/32$ over x, z, the line $L_{x,z}$ is nondegenerate, $T(L_{x,z}, \widehat{f} \circ L_{x,z}) = 1$, and $|\mathcal{P}_{L_{x,z}}| \le q/(32(\hat{h}+1))$. Every polynomial accepted by T belongs to $\mathcal{P}_{L_{x,z}}$. Thus,

$$\Pr_{x,z}\Big[\mathrm{ACCEPT}_{C_{z,\widehat{f}(z)}}(x) \ne \{\widehat{f}(x)\}\Big] \le \frac{1}{32} + \frac{1}{32} = \frac{1}{16}.$$

Averaging over z gives z^*. The size bound follows from the size of T and the cost of evaluating p. □

8.2.2 Correcting a nondeterministic approximation

The next lemma converts a nondeterministic circuit that is correct and single-valued on a $15/16$-fraction of inputs into one that is correct and single-valued on every input.

Lemma 8.5 (Nondeterministic local correction)**.** *Let $g\colon \mathbb{F}_q^d \to \mathbb{F}_q$ be a polynomial of total degree at most $\hat{h}$, where $q - 1 \geq 4(\hat{h}+1)$. Suppose a nondeterministic circuit C of size S satisfies*

$$\Pr_{x \leftarrow \mathbb{F}_q^d}[\mathrm{ACCEPT}_C(x) = \{g(x)\}] \geq 15/16.$$

Then there is a single-valued nondeterministic circuit C^* satisfying $\mathrm{ACCEPT}_{C^*}(x) = \{g(x)\}$ for every $x \in \mathbb{F}_q^d$, of size polynomial in $S, \hat{h}, d, \log q$, with an absolute polynomial bound.

Proof. Let $G = \{x : \mathrm{ACCEPT}_C(x) = \{g(x)\}\}$, so $|G| \geq (15/16)q^d$.

Correction at one point. Set $s = 4(\hat{h}+1)$ and fix distinct nonzero field elements $t_1, \ldots, t_s \in \mathbb{F}_q$. For a target x, choose $y \leftarrow \mathbb{F}_q^d$ uniformly and set $L(t) = x + ty$. Guess a polynomial p of degree at most $\hat{h}$, a set of $3s/4$ distinct indices, and, for each selected index i, a successful witness certifying $p(t_i) \in \mathrm{ACCEPT}_C(L(t_i))$. Reject if any certificate fails; otherwise, accept and output $p(0)$.

Define

$$G_x = \{y \in \mathbb{F}_q^d : |\{i \in [s] : x + t_i y \in G\}| \geq 3s/4\}.$$

For each i, the point $x + t_i y$ is uniform on $\mathbb{F}_q^d$. Thus the expected number of sampled points outside G is at most $s/16$. By Markov's inequality,

$$\Pr_{y \leftarrow \mathbb{F}_q^d}[y \in G_x] \geq 3/4.$$

For every $y \in G_x$, $p^* = g \circ L$ has successful witnesses on at least $3s/4$ positions. Any other polynomial can have witnesses only at the at most $s/4$ bad positions and at the at most $\hat{h}$ field elements where it agrees with p^*. Since

$$s/4 + \hat{h} < 3s/4,$$

no wrong polynomial is accepted. Consequently, for every $y \in G_x$, the set of possible outputs is exactly $\{g(x)\}$.

Amplification and fixing $y_1, \ldots, y_R$. Set $R = 16(d \log q + 1)$ and choose $y_1, \ldots, y_R$ independently and uniformly from $\mathbb{F}_q^d$. Guess a common output $a \in \mathbb{F}_q$ and successful correction witnesses producing a for more than half of the repetitions. Reject if any selected certificate fails; otherwise, accept and output a. If $|\{j \in [R] : y_j \in G_x\}| > R/2$, this circuit has output set exactly $\{g(x)\}$: the repetitions with $y_j \in G_x$ certify $g(x)$, whereas any other value can be certified only when $y_j \notin G_x$. We require no witnesses for the remaining repetitions.

For each fixed x, a Chernoff bound gives

$$\Pr[|\{j \in [R] : y_j \in G_x\}| \leq R/2] \leq e^{-R/8}.$$

Hence

$$\Pr\Big[\exists x \in \mathbb{F}_q^d : |\{j \in [R] : y_j \in G_x\}| \leq R/2\Big] \leq q^d e^{-R/8} < 1.$$

Fix $y_1, \ldots, y_R$ such that $|\{j \in [R] : y_j \in G_x\}| > R/2$ for every x, and hardwire them. The resulting circuit is C^*.

Each repetition uses at most s copies of the verifier for C, field arithmetic, and polynomial evaluation. There are $O(d \log q + 1)$ repetitions. All guesses and their witnesses form one existential witness, and the total size is polynomial in the stated parameters. □

Proof of Lemma 8.2. The hypothesis gives some L with $1 \le |\mathcal{P}_L| \le q/(32(\hat{h}+1))$, so $q \ge 32(\hat{h}+1)$ and $q - 1 \ge 4(\hat{h}+1)$. Apply Lemma 8.5 to the circuit $C_{z^*,\widehat{f}(z^*)}$ from Lemma 8.4 to obtain the required circuit of size $\mathrm{poly}(S, \hat{h}, d, \log q)$. □

9 Proof of the robust errorless block condenser theorem

In this section we prove Theorem 4.43. The construction (that appears in Section 9.1) and analysis (that appears in Section 9.2) closely follow the high level description of Section 7.

9.1 Construction

Fix constants $c_{\mathrm{size}} > 1$ and $0 < \gamma \le 1$, and let

$$n^{\gamma} \le k_y \le k_x \le n, \qquad (\log n)^2 \le m \le k_y^{\delta}.$$

The universal constant $\delta > 0$ will be chosen sufficiently small in the two-source condenser item of Section 9.1.1. Our goal is to construct an explicit function

$$\mathsf{rbCond}\colon \{0,1\}^n \times \{0,1\}^n \longrightarrow \{0,1\}^m$$

that is computable in polynomial time and satisfies the robust errorless block condenser guarantee of Theorem 4.43 for joint distributions samplable by deterministic circuits with postselection by nondeterministic circuits, both of size at most $n^{c_{\mathrm{size}}}$. We assume that E is hard for exponential-size nondeterministic circuits.

9.1.1 Components

We rely on the following components:

- **Hard function and low-degree extension.** We assume that E is hard for exponential-size nondeterministic circuits. By Assumption 4.10, this gives a function $f \in \mathsf{E}$ and a constant $\beta_{\mathrm{hard}} > 0$ such that, for every sufficiently large input length ℓ, no nondeterministic circuit of size at most $2^{\beta_{\mathrm{hard}}\ell}$ computes f on all inputs in $\{0,1\}^{\ell}$. Let $d \ge 2$ be a sufficiently large integer that will be chosen later in Section 9.2.6, and set

$$\ell = d \log n, \qquad h = 2^{\ell/d} = n, \qquad q = 2^{k_y^{1/2}}.$$

  Let $\mathbb{F}_q$ be the field with q elements, which exists because q is a power of two. For sufficiently large n, $d \le h \le q$. Let $\phi\colon \{0,1\}^{\ell} \to H^d$ be the natural bijection, where H is as in Theorem 4.34. By Theorem 4.34, f has a low-degree extension

$$\widehat{f}\colon \mathbb{F}_q^d \longrightarrow \mathbb{F}_q$$

  with $\widehat{f}(\phi(x)) = f(x)$ for every $x \in \{0,1\}^{\ell}$, and total degree at most

$$\hat{h} \coloneqq d(h-1) < d \cdot n.$$

  Thus, for every affine line $L\colon \mathbb{F}_q \to \mathbb{F}_q^d$, $\widehat{f} \circ L$ is a univariate polynomial of degree at most $\hat{h}$. Since $f \in \mathsf{E}$, Theorem 4.34 gives that $\widehat{f}$ can be computed in time polynomial in 2^{ℓ} and $\log q$, and hence in time polynomial in n for every fixed d.

- **Seeded extractor.** We use
$$\mathsf{SExt}\colon \{0,1\}^{n_y} \times [D] \longrightarrow \mathbb{F}_q^d,$$
with
$$n_y = n, \qquad k_{\mathsf{SExt}} = k_y - m/4, \qquad \varepsilon_{\mathsf{SExt}} = 1/8192.$$
Remark 9.1 (D and d are unrelated)**.** *Note that the parameters d and D are unrelated, and in particular, it is not the case that $D = 2^d$. We use d to measure the dimension of the vector space over which $\widehat{f}$ is defined. We use D as the number of seeds for the seeded extractor* SExt.

The choice of δ in the next item will satisfy $\delta < 1/2$. Consequently, for every fixed d and sufficiently large n,
$$\tfrac{3}{4}k_{\mathsf{SExt}} - O(\log n) = \tfrac{3}{4}(k_y - m/4) - O(\log n) \geq d \cdot k_y^{1/2} = d \log q.$$
Theorem 4.18, followed by truncation, supplies the stated output length. The seed length satisfies
$$\log D = O(\log n_y + \log(1/\varepsilon_{\mathsf{SExt}})) = O(\log n).$$
We write $D = n^{c_D}$, allowing unused seed bits if needed. The constant c_D is absolute and does not depend on d or m.

- **Two-source condenser.** Set
$$k_1 = k_2 = k_y^{1/4} = (\log q)^{1/2}, \qquad \delta = \min\{1/16,\ 1/(8C_{\mathsf{2Cond}})\},$$
where C_{2Cond} is the universal constant in Theorem 4.23. By Theorem 4.23, we obtain, after padding the inputs,
$$\mathsf{2Cond}\colon \mathbb{F}_q^D \times \{0,1\}^n \longrightarrow \{0,1\}^m.$$
The function 2Cond is an errorless $(k_1, k_2, 3m/4)$-two-source condenser (Definition 4.21).

Indeed, the first input has bit length
$$D \log q = n^{c_D} \cdot k_y^{1/2} \leq n^{c_D+1/2},$$
so padding both inputs to $N = n^{c_D+1}$ suffices. For sufficiently large n, we have $m \geq (\log n)^2 \geq \log N$ and
$$(m \log N)^{C_{\mathsf{2Cond}}} \leq k_y^{1/8} \cdot (\log N)^{C_{\mathsf{2Cond}}} = o(k_y^{1/4}) = o(k_1),$$
where we used $m \leq k_y^{\delta}$ and $k_y \geq n^{\gamma}$. Thus Theorem 4.23 applies with $r = k_1 = k_2$. The choice of δ is universal and does not depend on γ or on the sampler-size exponent c_{size}.

9.1.2 Operation

For every $y \in \{0,1\}^{n_y}$ and $i \in [D]$, define
$$w_i(y) \coloneqq \mathsf{SExt}(y, i).$$
Then define
$$g(y) \coloneqq \big(\widehat{f}(w_1(y)), \ldots, \widehat{f}(w_D(y))\big) \in \mathbb{F}_q^D.$$
The robust errorless block condenser construction is the function
$$\mathsf{rbCond}\colon \{0,1\}^n \times \{0,1\}^n \longrightarrow \{0,1\}^m$$
defined by
$$\boxed{\mathsf{rbCond}(y, x) = \mathsf{2Cond}(g(y), x).}$$
Since $\ell = O(\log n)$ and $f \in \mathsf{E}$, the truth table of f can be computed in polynomial time. Evaluating g and the two-source condenser therefore takes polynomial time once d is fixed.

9.2 Analysis of the construction

9.2.1 Preparations

Assume for contradiction that the construction fails for a joint distribution

$$(\mathbf{Y}, \mathbf{X}) = (\mathsf{Samp}(\mathbf{R}) \mid \mathsf{Post}(\mathbf{R}) = 1), \qquad \mathbf{R} \leftarrow \mathbf{U}_r,$$

where Samp is deterministic, Post is nondeterministic, and both have size at most $n^{c_{\mathrm{size}}}$. Set

$$p_{\mathrm{post}} \coloneqq \Pr_{\mathbf{R}}[\mathsf{Post}(\mathbf{R}) = 1] > 0.$$

Set $\lambda = 2^{-m/4}$ and fix a decomposition

$$(\mathbf{Y}, \mathbf{X}) = \lambda'(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}}) + (1 - \lambda')(\mathbf{Y}_{\mathrm{rest}}, \mathbf{X}_{\mathrm{rest}}), \qquad \lambda' \geq \lambda,$$

satisfying the hypotheses of Definition 4.32 with $k_y^{\geq} = k_x/2$. Thus

$$\mathbf{H}_\infty(\mathbf{Y}_{\mathrm{good}}) \geq k_y, \qquad \mathbf{H}_\infty(\mathbf{X}) \geq k_x, \qquad \Pr[\mathbf{Y}_{\mathrm{good}} = y] \geq 2^{-k_x/2} \quad \text{for every } y \in \mathrm{Supp}(\mathbf{Y}_{\mathrm{good}}).$$

Set $\mu = 2^{-m/4}$. Failure of the output min-entropy guarantee gives some $\sigma \in \{0,1\}^m$ with

$$\Pr[\mathsf{rbCond}(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}}) = \sigma] > \mu.$$

Fix this joint distribution, decomposition, and output. Write

$$\mathbf{Y}_\sigma = \big(\mathbf{Y}_{\mathrm{good}} \mid \mathsf{rbCond}(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}}) = \sigma\big).$$

Lemma 9.2 (A seed good for the conditioned distribution)**.** *There is $i^* \in [D]$ such that*

$$\mathsf{SD}\big(\mathsf{SExt}(\mathbf{Y}_\sigma, i^*), \mathbf{U}_{\mathbb{F}_q^d}\big) \leq \varepsilon_{\mathsf{SExt}} = 1/8192.$$

Proof. By Lemma 4.4,

$$\begin{aligned}\mathbf{H}_\infty(\mathbf{Y}_\sigma) &\geq \mathbf{H}_\infty(\mathbf{Y}_{\mathrm{good}}) + \log \Pr[\mathsf{rbCond}(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}}) = \sigma] \\ &> k_y - m/4 = k_{\mathsf{SExt}}.\end{aligned}$$

Strong extraction bounds the average over a uniform seed of the statistical distance in the statement by $\varepsilon_{\mathsf{SExt}}$. Some seed has distance no greater than this average. □

In the analysis, in order to simplify the notation, we will assume without loss of generality that $i^* = D$. This assumption is made only for convenience in the proof. The choice of i^* can be supplied to the circuits that we construct as nonuniform advice, and we relabel the seeds and the corresponding first-input coordinates of 2Cond accordingly. Define

$$w(y) = \mathsf{SExt}(y, D), \qquad \mathbf{W} = w(\mathbf{Y}), \qquad \mathbf{W}_\sigma = w(\mathbf{Y}_\sigma),$$

$$P(w) = \Pr[\mathbf{W} = w], \qquad Q(w) = \Pr[\mathbf{W}_\sigma = w].$$

The distribution Q is within distance $1/8192$ of uniform. We do not assert that P is close to uniform.

Valid values and the set G. We call $y \in \mathrm{Supp}(\mathbf{Y})$ valid if $\mathbf{H}_\infty(\mathbf{X} \mid \mathbf{Y} = y) \geq k_2$. We will construct a set G whose membership can be verified by a small nondeterministic circuit and such that $\mathrm{Supp}(\mathbf{Y}_{\mathrm{good}}) \subseteq G \subseteq \mathcal{Y}_{\mathrm{valid}}$.

Set

$$\mathcal{Y}_{\mathrm{valid}} = \{y \in \mathrm{Supp}(\mathbf{Y}) : \mathbf{H}_\infty(\mathbf{X} \mid \mathbf{Y} = y) \geq k_2\}.$$

Remark 9.3 (Changes in the definition of valid y)**.** *In the definition above, we modify the notion of validity considered in Section 7.3. The most natural generalization would be to define validity in terms of the min-entropy of $(\mathbf{X}_{\mathrm{good}} \mid \mathbf{Y}_{\mathrm{good}} = y)$. However, we only have a sampler with postselection for the original joint distribution $(\mathbf{Y}, \mathbf{X})$, rather than for its good component. We therefore define validity in terms of $(\mathbf{X} \mid \mathbf{Y} = y)$. In the argument below, we will show that every $y \in \mathrm{Supp}(\mathbf{Y}_{\mathrm{good}})$ remains valid under this definition: the mixture decomposition gives a lower bound on $\Pr[\mathbf{Y} = y]$, which, together with the min-entropy guarantee for $\mathbf{X}$, ensures that $(\mathbf{X} \mid \mathbf{Y} = y)$ has sufficiently high min-entropy.*

For every $y \in \mathrm{Supp}(\mathbf{Y}_{\mathrm{good}})$, we have the lower bound

$$\Pr[\mathbf{Y}_{\mathrm{good}} = y] \geq 2^{-k_y^{\geq}} = 2^{-k_x/2},$$

and hence

$$\Pr[\mathbf{Y} = y] \geq \lambda \cdot 2^{-k_x/2}.$$

We can use Goldwasser–Sipser (Theorem 4.39) to recognize, up to an approximation gap, values $y \in \{0,1\}^{n_y}$ for which $\Pr[\mathbf{Y} = y]$ is large. Write $\mathsf{Samp}(R) = (\mathsf{Samp}_Y(R), \mathsf{Samp}_X(R))$, where Samp_Y and Samp_X denote the two output components of Samp. For every $y \in \{0,1\}^{n_y}$, define a nondeterministic circuit A_y with input $R \in \{0,1\}^r$ as follows. Compute $\mathsf{Samp}_Y(R)$, and guess a witness $\mathrm{wit}_{\mathrm{post}}$ for Post. Accept if and only if $\mathsf{Samp}_Y(R) = y$ and $\mathsf{Post}(R; \mathrm{wit}_{\mathrm{post}}) = 1$. The circuit has size

$$\mathrm{size}(A_y) = O(n^{c_{\mathrm{size}}}),$$

and its acceptance probability satisfies

$$\begin{aligned}
\Pr_{\mathbf{R} \leftarrow \mathbf{U}_r}[A_y(\mathbf{R}) = 1] &= \Pr_{\mathbf{R} \leftarrow \mathbf{U}_r}[\mathsf{Samp}_Y(\mathbf{R}) = y \ \wedge \ \mathsf{Post}(\mathbf{R}) = 1] \\
&= p_{\mathrm{post}} \cdot \Pr_{\mathbf{R} \leftarrow \mathbf{U}_r}[\mathsf{Samp}_Y(\mathbf{R}) = y \mid \mathsf{Post}(\mathbf{R}) = 1] \\
&= p_{\mathrm{post}} \cdot \Pr[\mathbf{Y} = y].
\end{aligned}$$

Let A_{GS} be the nondeterministic circuit supplied by Theorem 4.39, with the gap parameter λ of that theorem set to $1/2$ and size parameter large enough to accommodate A_y and the threshold $p_{\mathrm{post}} \cdot \lambda \cdot 2^{-k_x/2}$. Define the nondeterministic circuit A_{mass} by

$$A_{\mathrm{mass}}(y) \coloneqq A_{\mathrm{GS}}(A_y, p_{\mathrm{post}} \cdot \lambda \cdot 2^{-k_x/2}).$$

Here p_{post} is hardwired into A_{mass} as nonuniform advice. Note that p_{post} can be described using only $O(r+1)$ bits, where $r \leq n^{c_{\mathrm{size}}}$. The threshold has bit length $O(r + m + k_x + 1) = O(n^{c_{\mathrm{size}}})$. Consequently,

$$\mathrm{size}(A_{\mathrm{GS}}) = \mathrm{poly}(n^{c_{\mathrm{size}}}), \qquad \mathrm{size}(A_{\mathrm{mass}}) = \mathrm{poly}(n^{c_{\mathrm{size}}}),$$

with absolute polynomial exponents independent of d. Since $1/2 < e^{-1/2}$, the circuit satisfies

$$\begin{aligned}
\Pr[\mathbf{Y} = y] \geq \lambda \cdot 2^{-k_x/2} \quad &\Longrightarrow \quad A_{\mathrm{mass}}(y) = 1, \\
\Pr[\mathbf{Y} = y] \leq \lambda \cdot 2^{-k_x/2-1} \quad &\Longrightarrow \quad A_{\mathrm{mass}}(y) = 0.
\end{aligned}$$

Define

$$G \coloneqq \{y \in \{0,1\}^{n_y} : A_{\mathrm{mass}}(y) = 1\}.$$

Lemma 9.4 (The set G contains the good component and only valid values)**.** *For sufficiently large n,*

$$\mathrm{Supp}(\mathbf{Y}_{\mathrm{good}}) \subseteq G \subseteq \mathcal{Y}_{\mathrm{valid}}.$$

Proof. For every $y \in \mathrm{Supp}(\mathbf{Y}_{\mathrm{good}})$, the mixture decomposition, $\lambda' \geq \lambda$, and the lower bound on its probability in $\mathbf{Y}_{\mathrm{good}}$ give

$$\Pr[\mathbf{Y} = y] \geq \lambda \cdot \Pr[\mathbf{Y}_{\mathrm{good}} = y] \geq \lambda \cdot 2^{-k_x/2}.$$

Thus $y \in G$.

To prove the second inclusion, let $y \in G$. Then $\Pr[\mathbf{Y} = y] > \lambda \cdot 2^{-k_x/2-1} > 0$. For every x,

$$\Pr[\mathbf{X} = x \mid \mathbf{Y} = y] = \frac{\Pr[\mathbf{X} = x \ \wedge \ \mathbf{Y} = y]}{\Pr[\mathbf{Y} = y]} \leq \frac{\Pr[\mathbf{X} = x]}{\Pr[\mathbf{Y} = y]} < \frac{2^{-k_x}}{\lambda \cdot 2^{-k_x/2-1}} = 2^{-k_x/2+m/4+1}.$$

Consequently,

$$\mathbf{H}_\infty(\mathbf{X} \mid \mathbf{Y} = y) > k_x/2 - m/4 - 1 \geq k_2.$$

The last inequality holds for sufficiently large n, since $k_x \geq k_y$, $m \leq k_y^\delta = o(k_y)$, and $k_2 = k_y^{1/4} = o(k_y)$. This proves $y \in \mathcal{Y}_{\mathrm{valid}}$. □

Candidate lists. Set $\Theta := 2^{-3m/4}$, the bound on an output point probability supplied by the two-source condenser.

Definition 9.5. *For every $y \in \mathcal{Y}_{\mathrm{valid}}$, define*

$$\mathcal{V}_y = \left\{ u \in \mathbb{F}_q^D : \Pr[\mathsf{2Cond}(u, \mathbf{X}) = \sigma \mid \mathbf{Y} = y] > \Theta \right\}. \tag{5}$$

Lemma 9.6. *For every $y \in \mathcal{Y}_{\mathrm{valid}}$, $|\mathcal{V}_y| < 2^{k_1}$.*

Proof. Otherwise take a uniform vector $\mathbf{U}_y$ from $\mathcal{V}_y$, independently of a random variable $\mathbf{X}_y$ with distribution $(\mathbf{X} \mid \mathbf{Y} = y)$. Then $\mathbf{H}_\infty(\mathbf{U}_y) \geq k_1$ and $\mathbf{H}_\infty(\mathbf{X}_y) \geq k_2$. Theorem 4.23 and its instantiation in Section 9.1.1 give

$$\Pr[\mathsf{2Cond}(\mathbf{U}_y, \mathbf{X}_y) = \sigma] \leq 2^{-3m/4} = \Theta.$$

On the other hand, averaging the strict inequality in the definition of $\mathcal{V}_y$ over $\mathbf{U}_y$ gives $\Pr[\mathsf{2Cond}(\mathbf{U}_y, \mathbf{X}_y) = \sigma] > \Theta$, a contradiction. □

Definition 9.7. *For every $y \in \mathcal{Y}_{\mathrm{valid}}$, define*

$$R_y = \{u_D : u \in \mathcal{V}_y\}, \qquad |R_y| \leq 2^{k_1}. \tag{6}$$

9.2.2 The test

Given a line L, we seek a test for whether a degree-at-most-$\hat{h}$ polynomial p is the correct restriction $\widehat{f} \circ L$. We first define a joint contribution at a point, then a pointwise predicate, and finally the score of a candidate along a line.

Definition 9.8 (Joint contribution at a point)**.** *For a point $w \in \mathbb{F}_q^d$ and a field element $a \in \mathbb{F}_q$, set*

$$J_G(w, a) = \sum_{\substack{y \in G \\ w(y) = w}} \max_{v \in \mathbb{F}_q^{D-1}} \Pr_{(\mathbf{Y}, \mathbf{X})} \left[\mathbf{Y} = y \ \wedge \ \mathsf{2Cond}((v, a), \mathbf{X}) = \sigma \right]. \tag{7}$$

Remark 9.9 (Using the original joint source)**.** *In Definition 9.8, we consider a slightly different notion of joint contribution at a point from the one considered in Section 7.3. More specifically, rather than summing over all y with $w(y) = w$, we restrict the sum to those $y \in G$. Intuitively, we would like to restrict the sum to $\mathrm{Supp}(\mathbf{Y}_{\mathrm{good}})$, but we only have access to a sampler for the original joint distribution $(\mathbf{Y}, \mathbf{X})$. We instead use the set G, which is recognized by the small nondeterministic circuit A_{mass}. Lemma 9.4 guarantees that $\mathrm{Supp}(\mathbf{Y}_{\mathrm{good}}) \subseteq G$ and every $y \in G$ is valid. For completeness, the mixture decomposition allows us to lower-bound the contribution from the good component, and the additional nonnegative terms corresponding to $y \in G \setminus \mathrm{Supp}(\mathbf{Y}_{\mathrm{good}})$ cannot hurt. For soundness, we use the fact that every $y \in G$ is valid, which is the property we already used in the soundness analysis in Section 7.3.*

Set

$$\theta_+ = \frac{\lambda\mu}{2q^d}, \qquad \theta_- = \frac{\lambda\mu}{4q^d}. \tag{8}$$

Recall that $\lambda = \mu = 2^{-m/4}$, so $\lambda\mu = 2^{-m/2}$. These are thresholds for the joint contribution $J_G(w, a)$, whereas $\Theta = 2^{-3m/4}$ is the conditional-probability threshold used to define the candidate lists in Definition 9.5. Fix any Boolean predicate $\mathrm{Thr}(w, a)$ satisfying

$$J_G(w, a) \geq \theta_+ \implies \mathrm{Thr}(w, a) = 1, \tag{9}$$

$$J_G(w, a) \leq \theta_- \implies \mathrm{Thr}(w, a) = 0. \tag{10}$$

Its behavior between the thresholds is unrestricted. Section 9.2.5 constructs such a predicate as a small nondeterministic circuit. Completeness and soundness will hold for every fixed predicate with these two properties, and hence for that implementation.

Definition 9.10 (The polynomial test)**.** *For a nondegenerate line L and a polynomial $p \in \mathbb{F}_q[t]$ of degree at most $\hat{h}$, define*

$$\mathrm{Test}_L(p) = \Pr_{t \leftarrow \mathbb{F}_q}[\mathrm{Thr}(L(t), p(t)) = 1]. \tag{11}$$

Set $\mathrm{Test}_L(p) = 0$ for degenerate lines.

We prove completeness and soundness as required by Lemma 8.2, and then construct Thr.

9.2.3 Completeness

Lemma 9.11 (Completeness)**.** *For all sufficiently large n, with probability at least $1 - \frac{1}{3n}$ over $\mathbf{L} \leftarrow \mathsf{RandLine}$, the line is nondegenerate and*

$$\mathrm{Test}_{\mathbf{L}}(\widehat{f} \circ \mathbf{L}) \geq 7/8.$$

Proof. **Correct evaluations pass at almost all points.** Recall that

$$Q(w) = \Pr\big[w(\mathbf{Y}_{\mathrm{good}}) = w \mid \mathsf{rbCond}(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}}) = \sigma\big].$$

Let

$$\mathcal{A} := \{w \in \mathbb{F}_q^d : Q(w) < 1/(2q^d)\}.$$

Since $\mathsf{SD}(Q, \mathbf{U}_{\mathbb{F}_q^d}) \leq 1/8192$, Lemma 4.5 implies

$$\Pr_{w \leftarrow \mathbb{F}_q^d}[w \in \mathcal{A}] \leq \frac{1/8192}{1/2} = \frac{1}{4096} \leq \frac{1}{128}.$$

We will now show that, for every $w \notin \mathcal{A}$, $J_G(w, \widehat{f}(w)) \geq \theta_+$, which implies that $\mathrm{Thr}(w, \widehat{f}(w)) = 1$.

For every $w \notin \mathcal{A}$, in each summand of J_G, choose v to be the first $D-1$ coordinates of $g(y)$. When $w(y) = w$, its last coordinate is $\widehat{f}(w)$. Consequently,

$$\begin{aligned}
J_G(w, \widehat{f}(w)) &\geq \sum_{\substack{y \in G \\ w(y)=w}} \Pr[\mathbf{Y} = y \ \wedge \ \mathsf{rbCond}(\mathbf{Y}, \mathbf{X}) = \sigma] \\
&\geq \lambda \cdot \sum_{\substack{y \in G \\ w(y)=w}} \Pr[\mathbf{Y}_{\text{good}} = y \ \wedge \ \mathsf{rbCond}(\mathbf{Y}_{\text{good}}, \mathbf{X}_{\text{good}}) = \sigma] \\
&= \lambda \cdot \Pr[w(\mathbf{Y}_{\text{good}}) = w \ \wedge \ \mathsf{rbCond}(\mathbf{Y}_{\text{good}}, \mathbf{X}_{\text{good}}) = \sigma] \\
&= \lambda \cdot \Pr[\mathsf{rbCond}(\mathbf{Y}_{\text{good}}, \mathbf{X}_{\text{good}}) = \sigma] \cdot \Pr[w(\mathbf{Y}_{\text{good}}) = w \mid \mathsf{rbCond}(\mathbf{Y}_{\text{good}}, \mathbf{X}_{\text{good}}) = \sigma] \\
&= \lambda \cdot \Pr[\mathsf{rbCond}(\mathbf{Y}_{\text{good}}, \mathbf{X}_{\text{good}}) = \sigma] \cdot Q(w) \\
&> \frac{\lambda\mu}{2q^d} = \theta_+.
\end{aligned}$$

The second inequality uses the mixture decomposition of $(\mathbf{Y}, \mathbf{X})$ and $\lambda' \geq \lambda$. The first equality uses $\mathrm{Supp}(\mathbf{Y}_{\text{good}}) \subseteq G$, from Lemma 9.4. The final inequality uses $\Pr[\mathsf{rbCond}(\mathbf{Y}_{\text{good}}, \mathbf{X}_{\text{good}}) = \sigma] > \mu$ and $Q(w) \geq 1/(2q^d)$. This gives that, for every $w \notin \mathcal{A}$, $\mathrm{Thr}(w, \widehat{f}(w)) = 1$. Thus the correct evaluation passes at all but at most a $1/128$ fraction of the points.

Restriction to a random line. Let $B(w) = \mathbf{1}\{\mathrm{Thr}(w, \widehat{f}(w)) = 1\}$. The bound above gives $\mathbb{E}_{w \leftarrow \mathbb{F}_q^d}[B(w)] \geq 127/128$. Applying Lemma 4.37 with $127/128 - 7/8 = 15/128$ gives

$$\Pr_{\mathbf{L} \leftarrow \mathsf{RandLine}}\left[\mathop{\mathbb{E}}_{t \leftarrow \mathbb{F}_q}[B(\mathbf{L}(t))] < 7/8\right] \leq \frac{1}{4q(15/128)^2} < \frac{21}{q}.$$

For a nondegenerate line, $\mathrm{Test}_L(\widehat{f} \circ L) = \mathbb{E}_{t \leftarrow \mathbb{F}_q}[B(L(t))]$. Adding degeneracy, the correct polynomial satisfies $\mathrm{Test}_{\mathbf{L}}(\widehat{f} \circ \mathbf{L}) \geq 7/8$ except with probability at most $q^{-d} + 21/q = o(1/n) \leq \frac{1}{3n}$. □

9.2.4 Soundness

Lemma 9.12 (Soundness)**.** *Set*

$$K = \frac{1024 \cdot 2^{k_1}}{\lambda\mu} = 1024 \cdot 2^{k_1 + m/2}. \tag{12}$$

For all sufficiently large n, with probability at least $1 - \frac{1}{3n}$ over $\mathbf{L} \leftarrow \mathsf{RandLine}$, the line is nondegenerate and the list

$$\mathcal{P}_{\mathbf{L}} = \{p \in \mathbb{F}_q[t] : \deg p \leq \hat{h}, \ \mathrm{Test}_{\mathbf{L}}(p) > 3/4\}$$

has size at most $8 \cdot K = o(q/\hat{h})$.

Proof. **Good points for soundness.** Every $y \in G$ is valid, by Lemma 9.4. Define

$$\mathcal{H} := \{w \in \mathbb{F}_q^d : P(w) > 128/q^d\}.$$

Since $\sum_w P(w) = 1$, Markov's inequality gives

$$\Pr_{w \leftarrow \mathbb{F}_q^d}[w \in \mathcal{H}] \leq \frac{\mathbb{E}_{w \leftarrow \mathbb{F}_q^d}[P(w)]}{128/q^d} = \frac{1/q^d}{128/q^d} = \frac{1}{128}.$$

Thus $\mathcal{H}$ contains at most a $1/128$-fraction of the points. This bound does not require P to be close to uniform.

Few field elements pass at a good point. Set

$$\mathcal{S}_w = \{a \in \mathbb{F}_q : \mathrm{Thr}(w, a) = 1\}.$$

By Equation (10), every $a \in \mathcal{S}_w$ satisfies $J_G(w, a) > \theta_-$. We show that $|\mathcal{S}_w| \le K$ for every $w \notin \mathcal{H}$.

Fix $w \notin \mathcal{H}$. For $a \in \mathbb{F}_q$, define

$$E_w(a) = \sum_{\substack{y \in G \\ w(y)=w \\ a \in R_y}} \Pr[\mathbf{Y} = y].$$

For every $y \in G \subseteq \mathcal{Y}_{\mathrm{valid}}$, every $v \in \mathbb{F}_q^{D-1}$, and every $a \in \mathbb{F}_q \setminus R_y$, Definitions 9.5 and 9.7 give

$$\Pr[\mathsf{2Cond}((v, a), \mathbf{X}) = \sigma \mid \mathbf{Y} = y] \le \Theta = 2^{-3m/4}.$$

For every $y \in G$, every $v \in \mathbb{F}_q^{D-1}$, and every $a \in R_y$, we use the trivial bound

$$\Pr[\mathsf{2Cond}((v, a), \mathbf{X}) = \sigma \mid \mathbf{Y} = y] \le 1.$$

These two groups give

$$\begin{aligned} J_G(w, a) &= \sum_{\substack{y \in G \\ w(y)=w}} \max_{v \in \mathbb{F}_q^{D-1}} \Pr_{(\mathbf{Y}, \mathbf{X})} \left[\mathbf{Y} = y \ \wedge \ \mathsf{2Cond}((v, a), \mathbf{X}) = \sigma\right] \\ &\le \Theta \cdot \sum_{\substack{y \in G \\ w(y)=w \\ a \notin R_y}} \Pr[\mathbf{Y} = y] + \sum_{\substack{y \in G \\ w(y)=w \\ a \in R_y}} \Pr[\mathbf{Y} = y] \\ &\le \Theta \cdot P(w) + E_w(a). \end{aligned} \tag{13}$$

The last inequality bounds the mass of the first group by all of $P(w)$. Since $w \notin \mathcal{H}$, we have

$$\begin{aligned} \Theta \cdot P(w) &\le \frac{128 \cdot 2^{-3m/4}}{q^d} \\ &= \frac{128 \cdot 2^{-m/4} \cdot \lambda\mu}{q^d} \\ &\le \frac{\lambda\mu}{8q^d}. \end{aligned}$$

The last inequality holds for sufficiently large n, since $m \ge (\log n)^2$ and hence $128 \cdot 2^{-m/4} \le 1/8$. Recall that $\theta_- = \lambda\mu/(4q^d)$. For every field element $a \in \mathcal{S}_w$, we therefore have

$$\begin{aligned} E_w(a) &\ge J_G(w, a) - \Theta \cdot P(w) \\ &> \theta_- - \frac{\lambda\mu}{8q^d} \\ &= \frac{\lambda\mu}{4q^d} - \frac{\lambda\mu}{8q^d} = \frac{\lambda\mu}{8q^d}. \end{aligned}$$

On the other hand, expanding the definition of $E_w(a)$, interchanging the sums, and using $|R_y| \le 2^{k_1}$, we have

$$\begin{aligned}\sum_{a\in\mathbb{F}_q} E_w(a) &= \sum_{\substack{y\in G\\ w(y)=w}} \Pr[\mathbf{Y}=y]\cdot|R_y| \\ &\le 2^{k_1}\cdot \sum_{\substack{y\in G\\ w(y)=w}} \Pr[\mathbf{Y}=y] \\ &\le 2^{k_1}\cdot \Pr[w(\mathbf{Y})=w] \\ &= 2^{k_1}\cdot P(w) \\ &\le \frac{128\cdot 2^{k_1}}{q^d}.\end{aligned}$$

The last inequality uses $w \notin \mathcal{H}$. Therefore,

$$|\mathcal{S}_w|\cdot\frac{\lambda\mu}{8q^d} \le \sum_{a\in\mathcal{S}_w} E_w(a) \le \sum_{a\in\mathbb{F}_q} E_w(a) \le \frac{128\cdot 2^{k_1}}{q^d},$$

and consequently

$$|\mathcal{S}_w| \le \frac{1024\cdot 2^{k_1}}{\lambda\mu} = K.$$

Restriction to a random line and list recovery. Since $\mathcal{H}$ contains at most a $1/128$-fraction of the points, applying Lemma 4.37 to its indicator gives

$$\Pr_{\mathbf{L}\leftarrow\mathsf{RandLine}}\left[\Pr_{t\leftarrow\mathbb{F}_q}[\mathbf{L}(t)\in\mathcal{H}] > \frac{1}{8}\right] \le \frac{1}{4q(1/8-1/128)^2} < \frac{21}{q}.$$

Fix a nondegenerate line L satisfying

$$\Pr_{t\leftarrow\mathbb{F}_q}[L(t)\in\mathcal{H}] \le \frac{1}{8}.$$

Recall that $\mathcal{S}_{L(t)} = \{a\in\mathbb{F}_q : \mathrm{Thr}(L(t),a)=1\}$ and that $|\mathcal{S}_{L(t)}| \le K$ whenever $L(t)\notin\mathcal{H}$. Define

$$\mathcal{S}_t = \begin{cases}\mathcal{S}_{L(t)}, & L(t)\notin\mathcal{H},\\ \varnothing, & L(t)\in\mathcal{H}.\end{cases}$$

Every list has size at most K. A polynomial p such that $\mathrm{Test}_L(p) > 3/4$ agrees with these lists on more than a $3/4-1/8=5/8$ fraction of the field elements $t\in\mathbb{F}_q$, hence on at least half. Lemma 4.38 with $\xi=1/2$ bounds their number by $8\cdot K$, provided $q \ge 8\cdot\hat{h}\cdot K$. This condition holds for sufficiently large n, because

$$\log(8\cdot\hat{h}\cdot K) = k_y^{1/4} + m/2 + O(\log n) = o(k_y^{1/2}) = o(\log q).$$

Here we used $m \le k_y^{\delta}$ with $\delta \le 1/16$ and $k_y \ge n^{\gamma}$. The same estimate gives $8\cdot K = o(q/\hat{h})$. By a union bound,

$$\Pr_{\mathbf{L}\leftarrow\mathsf{RandLine}}\left[\mathbf{L}\text{ is degenerate } \vee \Pr_{t\leftarrow\mathbb{F}_q}[\mathbf{L}(t)\in\mathcal{H}] > \frac{1}{8}\right] \le q^{-d} + \frac{21}{q} = o(1/n) \le \frac{1}{3n}.$$

□

9.2.5 Constructing Thr

Lemma 9.13 (A small threshold circuit)**.** *There is a constant $C_{\mathrm{Thr}} > 0$, depending on c_{size} and the fixed component constructions but independent of d, such that the following holds for all sufficiently large n. There is a nondeterministic circuit $\mathrm{Thr}(w, a)$ of size at most $n^{C_{\mathrm{Thr}}}$ satisfying Equation (9)–Equation (10) for every w, a. This circuit may depend on the circuits Samp and Post, the probability p_{post}, the circuit A_{mass}, the output σ, and the selected seed i^* from Section 9.2.1.*

Proof. Recall that $\mathsf{Samp}(R) = (\mathsf{Samp}_Y(R), \mathsf{Samp}_X(R))$. With external input (w, a) and uniform sampler coins $R \leftarrow \mathbf{U}_r$, define a nondeterministic circuit C by

$$C((w,a),y,v,R)=1 \quad \Longleftrightarrow \quad \begin{cases} \mathsf{Post}(R)=1, \\ A_{\mathrm{mass}}(y)=1, \\ \mathsf{SExt}(y,D)=w, \\ \mathsf{Samp}_Y(R)=y, \\ \mathsf{2Cond}((v,a),\mathsf{Samp}_X(R))=\sigma. \end{cases}$$

The circuit C guesses accepting witnesses for $\mathsf{Post}(R)$ and $A_{\mathrm{mass}}(y)$. We will apply Lemma 8.1 to the circuit C, with $u = (w, a)$. By Definition 9.8 and the definition of G,

$$S((w,a)) = p_{\mathrm{post}} \cdot J_G(w,a),$$

where S is defined in Lemma 8.1. Indeed, a value $y \notin G$ contributes zero, and for $y \in G$, conditioning on $\mathsf{Post}(R) = 1$ gives the corresponding summand of J_G.

Apply Lemma 8.1 with threshold $p_{\mathrm{post}}\theta_+$ and $\eta = 1/4$. Since $\theta_- = \theta_+/2 < (1-\eta)\theta_+$, the resulting circuit Thr satisfies Equation (9)–Equation (10). The threshold can be represented using $O(r+m+d\log q+1)$ bits. Here p_{post} is the hardwired probability from Section 9.2.1. Since p_{post} is an integer multiple of 2^{-r}, it can be described using $O(r+1)$ bits.

Circuit size. The circuit C evaluates Samp, the verifier of A_{mass}, SExt, and $\mathsf{2Cond}$, and verifies Post; it does not evaluate g or $\widehat{f}$. Lemma 8.1 gives a fixed polynomial bound in these circuit sizes, the input and witness lengths, and the threshold bit length. The circuit A_{mass} already has size polynomial in n with an exponent independent of d. For every fixed d, the same is true of all the other quantities: $r \le n^{c_{\mathrm{size}}}$, $\log q = k_y^{1/2} \le n^{1/2}$, $D = n^{c_D}$, and $d\log q \le dn^{1/2}$. Absorbing factors depending on d into the sufficiently large-n threshold gives the claimed exponent C_{Thr} independent of d. □

9.2.6 Reconstruction and completion of the proof

Applying the reconstruction lemma. By Lemmas 9.11 and 9.12, with probability at least $1 - \frac{2}{3n}$ over $\mathbf{L} \leftarrow \mathsf{RandLine}$,

$$\mathrm{Test}_{\mathbf{L}}(\widehat{f} \circ \mathbf{L}) \ge 7/8, \qquad |\mathcal{P}_{\mathbf{L}}| \le 8 \cdot K,$$

where $\mathcal{P}_L$ is defined in Lemma 8.2. Since $8 \cdot K = o(q/\hat{h})$, for sufficiently large n,

$$\frac{2}{3n} \le \frac{1}{32}, \qquad 8 \cdot K \le \frac{q}{32(\hat{h}+1)}.$$

Applying Lemma 8.2 to the circuit Thr from Lemma 9.13, there is a single-valued nondeterministic circuit C of size $\mathrm{poly}(\mathrm{size}(\mathrm{Thr}), \hat{h}, d, \log q)$ such that

$$\mathrm{ACCEPT}_C(x) = \{\widehat{f}(x)\} \qquad \text{for every } x \in \mathbb{F}_q^d.$$

Restricting to the hard function. Define $C_f(x)$ to simulate $C(\phi(x))$ and accept if it outputs 1. Since $\mathrm{ACCEPT}_C(\phi(x)) = \{\widehat{f}(\phi(x))\} = \{f(x)\}$, C_f computes f on $\{0,1\}^\ell$.

The natural encoding ϕ (from Theorem 4.34) has size polynomial in ℓ and $\log q$, with an absolute polynomial bound. Since $\ell = d \log n$ and $\hat{h} = d(n-1)$, incorporating this encoding gives absolute constants $C_0, a_0 > 0$ such that

$$\mathrm{size}(C_f) \le C_0(\mathrm{size}(\mathrm{Thr}) + \hat{h} + d + \log q + 1)^{a_0}.$$

By Lemma 9.13, $\mathrm{size}(\mathrm{Thr}) \le n^{C_{\mathrm{Thr}}}$ with C_{Thr} independent of d. For each fixed d, using $\hat{h} < d \cdot n$ and $\log q \le n$, the quantity in parentheses is at most $n^{C_{\mathrm{Thr}}+3}$ for sufficiently large n. Thus

$$\mathrm{size}(C_f) \le n^{C_{\mathrm{rec}}}, \qquad C_{\mathrm{rec}} = a_0(C_{\mathrm{Thr}} + 3) + 1,$$

where C_{rec} is independent of d.

Choosing the dimension and deriving the contradiction. The remaining component parameters were fixed independently of d. We can therefore first obtain C_{Thr} and C_{rec}, and then choose a sufficiently large integer $d \ge 2$ such that

$$\beta_{\mathrm{hard}} \cdot d > C_{\mathrm{rec}} + 1.$$

With this choice of d fixed, all the preceding requirements hold for sufficiently large n. Since $\ell = d \log n$,

$$\mathrm{size}(C_f) \le n^{C_{\mathrm{rec}}} < n^{\beta_{\mathrm{hard}} \cdot d} = 2^{\beta_{\mathrm{hard}} \ell},$$

contradicting the hardness of f. Thus, for every admissible joint distribution and decomposition, no output of $\mathsf{rbCond}(\mathbf{Y}_{\mathrm{good}}, \mathbf{X}_{\mathrm{good}})$ has probability greater than $2^{-m/4}$. Its min-entropy is at least $m/4$, proving Theorem 4.43.

10 Proof of the multiplicative extractor theorem

In this section we prove Theorem 4.41. The construction (that appears in Section 10.1) and analysis (that appears in Section 10.2) closely follow the high level description of Section 7.

10.1 Construction

Fix constants $c_{\mathrm{size}} > 1$, $0 < \alpha \le 1$, and $c_\varepsilon > 0$, and set

$$k = n^\alpha, \qquad \varepsilon = n^{-c_\varepsilon}.$$

Our goal is to construct an explicit function

$$\mathsf{Ext}\colon \{0,1\}^n \longrightarrow \{0,1\}^m, \qquad m = n^{\alpha'},$$

computable in polynomial time, that is a (k, ε)-multiplicative extractor for distributions samplable by deterministic circuits with postselection by nondeterministic circuits, both of size at most $n^{c_{\mathrm{size}}}$, as required by Theorem 4.41. The constant $\alpha' > 0$, depending only on α, will be chosen sufficiently small in the two-source extractor item of Section 10.1.1. We assume that E is hard for exponential-size nondeterministic circuits.

10.1.1 Components

We rely on the following components:

- **Errorless condenser.** Theorem 4.42 supplies a universal constant $0 < \delta_{\text{out}} < 1/2$ and errorless condensers $\mathsf{Cond}_{\gamma_{\text{in}},c}$ for the required sampling class. We instantiate this theorem twice, using common input length n and padding shorter inputs with zeros. First set

$$\gamma_{\text{in},1} = \alpha, \qquad c_1 = c_{\text{size}}, \qquad \mathsf{Cond}_1 = \mathsf{Cond}_{\gamma_{\text{in},1},c_1}.$$

Fix an exponent e_1 bounding both the running time and the Boolean circuit size of Cond_1 by n^{e_1}, enlarging it if necessary. Choose

$$c_2 > \max\{c_{\text{size}}, e_1, 1\} + 1, \qquad \gamma_{\text{in},2} = \alpha\delta_{\text{out}}, \qquad \mathsf{Cond}_2 = \mathsf{Cond}_{\gamma_{\text{in},2},c_2}.$$

The size bound n^{c_2} accommodates the deterministic sampler for $\mathbf{X}$ followed by Cond_1 and zero-padding, and also the original nondeterministic postselection circuit. Indeed, if $\mathbf{X} = (\mathsf{Samp}(\mathbf{R}) \mid \mathsf{Post}(\mathbf{R}) = 1)$, then

$$\mathsf{Cond}_1(\mathbf{X}) = (\mathsf{Cond}_1(\mathsf{Samp}(\mathbf{R})) \mid \mathsf{Post}(\mathbf{R}) = 1).$$

Let n^{e_2} bound the time and circuit size of Cond_2. Both instances are fixed before the dimension d is chosen; their size bounds are independent of that dimension.

Write $m_{\mathsf{Cond},i}$ and $k_{\mathsf{Cond},i}$ for the output length and guaranteed output min-entropy of the two instances:

Instance	Output length	Output min-entropy at least
Cond_1	$m_{\mathsf{Cond},1} = n^{\alpha/2}$	$k_{\mathsf{Cond},1} = n^{\alpha\delta_{\text{out}}}$
Cond_2	$m_{\mathsf{Cond},2} = n^{\alpha\delta_{\text{out}}/2}$	$k_{\mathsf{Cond},2} = n^{\alpha\delta_{\text{out}}^2}$

We will denote the distribution obtained after the second condenser application by $\mathbf{Y} = \mathsf{Cond}_2(\mathsf{Cond}_1(\mathbf{X}))$, and set

$$n_y \coloneqq m_{\mathsf{Cond},2} = n^{\alpha\delta_{\text{out}}/2}, \qquad k_y \coloneqq k_{\mathsf{Cond},2} = n^{\alpha\delta_{\text{out}}^2}.$$

Thus n_y is the length of $\mathbf{Y}$, and k_y is its guaranteed min-entropy for every source $\mathbf{X}$ in the specified class with $\mathbf{H}_\infty(\mathbf{X}) \geq k$.

- **Hard function and low-degree extension.** We assume that E is hard for exponential-size nondeterministic circuits. By Assumption 4.10, this gives a function $f \in \mathsf{E}$ and a constant $\beta_{\text{hard}} > 0$ such that, for every sufficiently large input length ℓ, no nondeterministic circuit of size at most $2^{\beta_{\text{hard}}\ell}$ computes f on all inputs in $\{0,1\}^\ell$. Let $d \geq 2$ be a sufficiently large integer that will be chosen later in Section 10.2.6, and set

$$\ell = d \log n, \qquad h = 2^{\ell/d} = n, \qquad q = 2^{k_y^{1/2}}.$$

Let $\mathbb{F}_q$ be the field with q elements, which exists because q is a power of two. For sufficiently large n, $d \leq h \leq q$. Let $\phi\colon \{0,1\}^\ell \to H^d$ be the natural bijection, where H is as in Theorem 4.34. By Theorem 4.34, f has a low-degree extension

$$\widehat{f}\colon \mathbb{F}_q^d \longrightarrow \mathbb{F}_q$$

with $\widehat{f}(\phi(x)) = f(x)$ for every $x \in \{0,1\}^\ell$, and total degree at most

$$\hat{h} \coloneqq d(h-1) < d \cdot n.$$

Thus, for every affine line $L\colon \mathbb{F}_q \to \mathbb{F}_q^d$, $\widehat{f} \circ L$ is a univariate polynomial of degree at most $\hat{h}$. Since $f \in \mathsf{E}$, Theorem 4.34 gives that $\widehat{f}$ can be computed in time polynomial in 2^ℓ and $\log q$, and hence in time polynomial in n for every fixed d.

- **Seeded extractor.** We use
$$\mathsf{SExt}\colon \{0,1\}^{n_y} \times [D] \longrightarrow \mathbb{F}_q^d,$$
with
$$k_{\mathsf{SExt}} = k_y - m, \qquad \varepsilon_{\mathsf{SExt}} = \varepsilon/8192.$$

Remark 10.1 (D and d are unrelated)**.** *Note that the parameters d and D are unrelated, and in particular, it is not the case that $D = 2^d$. We use d to measure the dimension of the vector space over which $\widehat{f}$ is defined. We use D as the number of seeds for the seeded extractor* SExt.

Here $m = n^{\alpha'}$, where $\alpha' > 0$ will be chosen sufficiently small in the next item; that choice will satisfy $\alpha' < \alpha\delta_{\mathrm{out}}^2/2$. Consequently, for every fixed d and sufficiently large n,
$$\tfrac{3}{4}k_{\mathsf{SExt}} - O(\log n) = \tfrac{3}{4}(k_y - m) - O(\log n) \geq d \cdot k_y^{1/2} = d \log q.$$

Theorem 4.18, followed by truncation, supplies the stated output length. The seed length satisfies
$$\log D = O(\log n_y + \log(1/\varepsilon_{\mathsf{SExt}})) = O(\log n).$$

We write $D = n^{c_D}$, allowing unused seed bits if needed. The constant c_D depends on c_ε, because the number of seeds depends on the seeded extractor's error. It does not depend on d or m. The seeded-extractor error depends on ε, not on m.

- **Two-source extractor.** Following Shaltiel [Sha25a], we will generalize the approach of [BSS25] to work with a two-source extractor with a weaker guarantee. More specifically, the extractor works only when its second source is a block-wise source (Definition 4.19). The advantage is that Li [Li15] constructed such an extractor with very low error, whereas known constructions of ordinary two-source extractors in the low-min-entropy regime do not achieve the error needed here. Set
$$k_1 = k_2 = k_y^{1/4} = (\log q)^{1/2}, \qquad \alpha' = \frac{\alpha\delta_{\mathrm{out}}^2}{8c_{\mathsf{IExt}}}, \qquad m = n^{\alpha'}.$$

Theorem 4.24, followed by padding and truncation, supplies
$$\mathsf{IExt}\colon \mathbb{F}_q^D \times \{0,1\}^{m_{\mathsf{Cond},1}} \times \{0,1\}^n \longrightarrow \{0,1\}^m, \qquad \varepsilon_{\mathsf{IExt}} := \frac{\varepsilon \cdot 2^{-m}}{n^2}.$$

The function IExt is a $(k_1, k_2, k_2, \varepsilon_{\mathsf{IExt}})$-extractor for one source and one block-wise source (Definition 4.22). Indeed, the first input has bit length
$$D \log q = n^{c_D} \cdot k_y^{1/2} = n^{c_D + \alpha\delta_{\mathrm{out}}^2/2} \leq n^{c_D + 1/2},$$
so padding all three inputs to $N = n^{c_D+1}$ suffices. We have $k_1 \geq \log^{12} N$ and $m < 0.9k_1$ for large n. Moreover, $k_1^{1/c_{\mathsf{IExt}}} = m^2$, and hence
$$2^{-k_1^{1/c_{\mathsf{IExt}}}} = 2^{-m^2} \leq \varepsilon \cdot 2^{-m}/n^2.$$

This holds for every fixed c_ε. The output exponent α' is independent of c_ε and c_{size}; with the component constants fixed, it depends only on α.

10.1.2 Operation

For every $y \in \{0,1\}^{n_y}$ and $i \in [D]$, define

$$w_i(y) \coloneqq \mathsf{SExt}(y, i).$$

Then define

$$g(y) \coloneqq \big(\widehat{f}(w_1(y)), \ldots, \widehat{f}(w_D(y))\big) \in \mathbb{F}_q^D.$$

For $x \in \{0,1\}^n$, define

$$F_1(x) \coloneqq \mathsf{Cond}_1(x), \qquad F_2(x) \coloneqq \mathsf{Cond}_2(F_1(x)),$$

where the input to Cond_2 is padded with zeros to length n. Write $\mathrm{Src}(x) = (F_1(x), x)$, and use the shorthand $\mathsf{IExt}(u, \mathrm{Src}(x)) \coloneqq \mathsf{IExt}(u, F_1(x), x)$. The extractor construction is the function

$$\mathsf{Ext}\colon \{0,1\}^n \longrightarrow \{0,1\}^m$$

defined by

$$\boxed{\mathsf{Ext}(x) = \mathsf{IExt}(g(F_2(x)), F_1(x), x) = \mathsf{IExt}(g(F_2(x)), \mathrm{Src}(x)).}$$

Thus the distribution $\mathbf{Y}$ introduced above is $\mathbf{Y} = F_2(\mathbf{X})$. Since $\ell = O(\log n)$ and $f \in \mathsf{E}$, the truth table of f can be computed in polynomial time. Evaluating g and all other components therefore takes polynomial time once d is fixed.

10.2 Analysis of the construction

10.2.1 Preparations

Assume for contradiction that the construction fails for a source $\mathbf{X}$ on $\{0,1\}^n$ with $\mathbf{H}_\infty(\mathbf{X}) \geq k$, where

$$\mathbf{X} = (\mathsf{Samp}(\mathbf{R}) \mid \mathsf{Post}(\mathbf{R}) = 1), \qquad \mathbf{R} \leftarrow \mathbf{U}_r,$$

Samp is deterministic, Post is nondeterministic, and both have size at most $n^{c_{\mathrm{size}}}$. Set

$$p_{\mathrm{post}} \coloneqq \Pr_{\mathbf{R}}[\mathsf{Post}(\mathbf{R}) = 1] > 0, \qquad \mu = 2^{-m}.$$

Failure of multiplicative extraction gives some $\sigma \in \{0,1\}^m$ with

$$\Pr[\mathsf{Ext}(\mathbf{X}) = \sigma] > e^{\varepsilon}\mu \geq (1+\varepsilon)\mu.$$

Fix this source and output. Write

$$\mathbf{Y} = F_2(\mathbf{X}), \qquad \mathbf{Y}_\sigma = (\mathbf{Y} \mid \mathsf{Ext}(\mathbf{X}) = \sigma).$$

Lemma 10.2 (A seed good for both distributions)**.** *For all sufficiently large n, there is $i^* \in [D]$ such that*

$$\mathsf{SD}(\mathsf{SExt}(\mathbf{Y}, i^*), \mathbf{U}_{\mathbb{F}_q^d}) \leq \frac{\varepsilon}{4096} \qquad \text{and} \qquad \mathsf{SD}(\mathsf{SExt}(\mathbf{Y}_\sigma, i^*), \mathbf{U}_{\mathbb{F}_q^d}) \leq \frac{\varepsilon}{4096}.$$

Proof. The second errorless condenser gives $\mathbf{H}_\infty(\mathbf{Y}) \geq k_y$. By Lemma 4.4,

$$\mathbf{H}_\infty(\mathbf{Y}_\sigma) \geq k_y + \log \Pr[\mathsf{Ext}(\mathbf{X}) = \sigma] \geq k_y - m = k_{\mathsf{SExt}}.$$

Both distributions meet the seeded-extractor min-entropy threshold. Strong extraction bounds the average over a uniform seed of the sum of the two distances by $2\varepsilon_{\mathsf{SExt}} = \varepsilon/4096$. Some seed has sum no greater than this average, so each of the two distances is at most $\varepsilon/4096$. □

In the analysis, in order to simplify the notation, we will assume without loss of generality that $i^* = D$. This assumption is made only for convenience in the proof. The choice of i^* can be supplied to the circuits that we construct as nonuniform advice, and we relabel the seeds and the corresponding first-input coordinates of IExt accordingly. Define

$$w(y) = \mathsf{SExt}(y, D), \qquad \mathbf{W} = w(\mathbf{Y}), \qquad \mathbf{W}_\sigma = w(\mathbf{Y}_\sigma),$$

$$P(w) = \Pr[\mathbf{W} = w], \qquad Q(w) = \Pr[\mathbf{W}_\sigma = w].$$

Thus each of P, Q is within distance $\varepsilon/4096$ of uniform.

Valid condenser outputs and their candidate lists. We will now define a set of "valid y's". These are the values of y for which $(\mathrm{Src}(\mathbf{X}) \mid \mathbf{Y} = y)$ is close to a source meeting the requirements that IExt imposes on its second source. Set

$$\eta_{\mathrm{blk}} := \frac{\varepsilon\mu}{n^2},$$

and define

$$\mathcal{Y}_{\mathrm{valid}} = \left\{ y \in \mathrm{Supp}(\mathbf{Y}) : \begin{array}{l} (\mathrm{Src}(\mathbf{X}) \mid \mathbf{Y} = y) \text{ is } \eta_{\mathrm{blk}}\text{-close to a} \\ (k_2, k_2)\text{-block-wise source} \end{array} \right\}.$$

Remark 10.3 (Valid and useful condenser outputs)**.** *In the notation used here, the approach of [TV00; BSS25] applies a two-source extractor* $\mathsf{TExt}(g(\mathsf{Cond}(\mathbf{X})), \mathbf{X})$*, where* Cond *is an errorless condenser. We replace* TExt *with* IExt*, which expects its second input to be a block-wise source. This is the task of the function* $\mathrm{Src}(x) = (F_1(x), x)$*. In other words, we can think of [TV00; BSS25] as using the function* $\mathrm{Src}(x) = x$*.*

In this proof, we say that y *is valid if* $(\mathrm{Src}(\mathbf{X}) \mid \mathbf{Y} = y)$ *meets the requirements that the two-source extractor imposes on its second source. Here, unlike in [TV00; BSS25], we allow this conditional distribution to be statistically close to the required source type.*

Lemma 10.4 (The conditional second input is close to a block-wise source)**.** *With*

$$\beta := 2^{m_{\mathsf{Cond},2} + k_2 - k_{\mathsf{Cond},1}} + \frac{2^{m_{\mathsf{Cond},1} + k_2 - k}}{\eta_{\mathrm{blk}}},$$

we have, for sufficiently large n*,*

$$\Pr[\mathbf{Y} \notin \mathcal{Y}_{\mathrm{valid}}] \le \beta \le \frac{\varepsilon\mu}{256n} \le \frac{\varepsilon\mu}{4096}. \tag{14}$$

Proof. Write $\mathbf{A} = F_1(\mathbf{X})$; then $\mathbf{Y} = \mathsf{Cond}_2(\mathbf{A})$ is a deterministic function of $\mathbf{A}$. We have $\mathbf{H}_\infty(\mathbf{A}) \ge k_{\mathsf{Cond},1}$ and $\mathbf{H}_\infty(\mathbf{X}) \ge k$. First, any y with $\Pr[\mathbf{Y} = y] \ge 2^{k_2 - k_{\mathsf{Cond},1}}$ satisfies

$$\Pr[\mathbf{A} = a \mid \mathbf{Y} = y] \le \frac{2^{-k_{\mathsf{Cond},1}}}{\Pr[\mathbf{Y} = y]} \le 2^{-k_2},$$

so $\mathbf{H}_\infty(\mathbf{A} \mid \mathbf{Y} = y) \ge k_2$. The total probability of the remaining values of y is at most $2^{m_{\mathsf{Cond},2} + k_2 - k_{\mathsf{Cond},1}}$.

Next let

$$\mathcal{B} = \{a \in \mathrm{Supp}(\mathbf{A}) : \Pr[\mathbf{A} = a] < 2^{k_2 - k}\}.$$

Its probability is at most $2^{m_{\mathsf{Cond},1} + k_2 - k}$. For $a \notin \mathcal{B}$, conditioning gives $\mathbf{H}_\infty(\mathbf{X} \mid \mathbf{A} = a) \ge k_2$. By Markov's inequality, the set of y with $\Pr[\mathbf{A} \in \mathcal{B} \mid \mathbf{Y} = y] > \eta_{\mathrm{blk}}$ has probability at most $2^{m_{\mathsf{Cond},1} + k_2 - k}/\eta_{\mathrm{blk}}$.

Fix a y outside the two exceptional sets. Keep the marginal distribution $(\mathbf{A} \mid \mathbf{Y} = y)$ exactly as it is. For each $a \notin \mathcal{B}$, keep $(\mathbf{X} \mid \mathbf{A} = a, \mathbf{Y} = y)$ unchanged. Because $\mathbf{Y}$ is a function of $\mathbf{A}$, this conditional distribution equals $(\mathbf{X} \mid \mathbf{A} = a)$ and has min-entropy at least k_2. For $a \in \mathcal{B}$, replace the conditional second

block by a uniform n-bit string. The resulting pair is a (k_2, k_2)-block-wise source. Its statistical distance from $(\mathbf{A}, \mathbf{X}) \mid \mathbf{Y} = y$ is at most $\Pr[\mathbf{A} \in \mathcal{B} \mid \mathbf{Y} = y] \le \eta_{\mathrm{blk}}$. This proves the first inequality in Equation (14).

For the parameter bound, the choices in Section 10.1 give $m_{\mathsf{Cond},2} = o(k_{\mathsf{Cond},1})$, $k_2 = o(k_{\mathsf{Cond},1})$, and $m = o(k_{\mathsf{Cond},1})$. Thus the first term of β is at most $2^{-k_{\mathsf{Cond},1}/2} = o(\varepsilon\mu/n)$. Similarly, $m_{\mathsf{Cond},1}, k_2, m, \log(1/\varepsilon), \log n = o(k)$, so the second term is at most

$$2^{m_{\mathsf{Cond},1}+k_2-k+m+(c_\varepsilon+2)\log n} \le 2^{-k/2} = o(\varepsilon\mu/n).$$

Their sum is at most $\varepsilon\mu/(256n)$ for sufficiently large n, proving the remaining inequalities. □

Candidate lists. Set $\Theta := (1 + \varepsilon/8)\mu$.

Definition 10.5. *For every $y \in \mathcal{Y}_{\mathrm{valid}}$, define*

$$\mathcal{V}_y = \left\{ u \in \mathbb{F}_q^D : \Pr[\mathsf{IExt}(u, \mathrm{Src}(\mathbf{X})) = \sigma \mid \mathbf{Y} = y] > \Theta \right\}. \tag{15}$$

Lemma 10.6. *For every $y \in \mathcal{Y}_{\mathrm{valid}}$, $|\mathcal{V}_y| < 2^{k_1}$.*

Proof. Otherwise take a uniform vector $\mathbf{U}_y$ from $\mathcal{V}_y$, independently of $(\mathrm{Src}(\mathbf{X}) \mid \mathbf{Y} = y)$. Its min-entropy is at least k_1. The conditional second input is η_{blk}-close to a (k_2, k_2)-block-wise source, and tensoring with the independent $\mathbf{U}_y$ preserves this distance. Theorem 4.24 and contraction of statistical distance therefore give

$$\Pr[\mathsf{IExt}(\mathbf{U}_y, \mathrm{Src}(\mathbf{X})) = \sigma \mid \mathbf{Y} = y] \le \mu + \varepsilon_{\mathsf{IExt}} + \eta_{\mathrm{blk}} = \mu + 2\varepsilon\mu/n^2 < (1 + \varepsilon/8)\mu = \Theta$$

for sufficiently large n, contradicting the definition of every vector in $\mathcal{V}_y$. This is the only use of the extraction property of IExt in the polynomial-test analysis. □

Definition 10.7. *For every $y \in \mathcal{Y}_{\mathrm{valid}}$, define*

$$R_y = \{u_D : u \in \mathcal{V}_y\}, \qquad |R_y| \le 2^{k_1}. \tag{16}$$

10.2.2 The test

Given a line L, we seek a test for whether a degree-at-most-$\hat{h}$ polynomial p is the correct restriction $\widehat{f} \circ L$. We first define a joint contribution at a point, then a pointwise predicate, and finally the score of a candidate along a line.

Definition 10.8 (Joint contribution at a point)**.** *For a point $w \in \mathbb{F}_q^d$ and a field element $a \in \mathbb{F}_q$, set*

$$J(w, a) = \sum_{y : w(y) = w} \max_{v \in \mathbb{F}_q^{D-1}} \Pr_{(\mathbf{Y}, \mathbf{X})}[\mathbf{Y} = y \ \wedge \ \mathsf{IExt}((v, a), \mathrm{Src}(\mathbf{X})) = \sigma]. \tag{17}$$

Remark 10.9 (Why we sum over all condenser outputs)**.** *In Section 9.2.2, we restrict the contribution J_G to values in G, because invalid values may carry substantial mass. Here, Lemma 10.4 shows that the total probability of invalid values is very small. We can therefore use the simpler contribution J, summing over all values of y, and account for the contribution of invalid values in the soundness analysis.*

Set

$$\theta_+ = (1 + 3\varepsilon/4)\mu/q^d, \qquad \theta_- = (1 + \varepsilon/2)\mu/q^d. \tag{18}$$

Fix any Boolean predicate $\mathrm{Thr}(w, a)$ satisfying

$$J(w, a) \ge \theta_+ \implies \mathrm{Thr}(w, a) = 1, \tag{19}$$

$$J(w, a) \le \theta_- \implies \mathrm{Thr}(w, a) = 0. \tag{20}$$

Its behavior between the thresholds is unrestricted. Section 10.2.5 constructs such a predicate as a small nondeterministic circuit. Completeness and soundness below hold for every fixed predicate with these two properties, and hence for that implementation.

Definition 10.10 (The polynomial test)**.** *For a nondegenerate line L and a polynomial $p \in \mathbb{F}_q[t]$ of degree at most $\hat{h}$, define*

$$\mathrm{Test}_L(p) = \Pr_{t \leftarrow \mathbb{F}_q}[\mathrm{Thr}(L(t), p(t)) = 1]. \tag{21}$$

Set $\mathrm{Test}_L(p) = 0$ for degenerate lines.

We prove completeness and soundness as required by Lemma 8.2, and then construct Thr.

10.2.3 Completeness

Lemma 10.11 (Completeness)**.** *For all sufficiently large n, with probability at least $1 - \frac{1}{3n}$ over $\mathbf{L} \leftarrow \mathsf{RandLine}$, the line is nondegenerate and*

$$\mathrm{Test}_{\mathbf{L}}(\widehat{f} \circ \mathbf{L}) \geq 7/8.$$

Proof. **Correct evaluations pass at almost all points.** Recall that $Q(w) = \Pr[\mathbf{W} = w \mid \mathsf{Ext}(\mathbf{X}) = \sigma]$. Let

$$\mathcal{A} := \left\{ w \in \mathbb{F}_q^d : Q(w) < \frac{1 - \varepsilon/16}{q^d} \right\}.$$

Since $\mathsf{SD}(Q, \mathbf{U}) \leq \varepsilon/2048$, Lemma 4.5 implies

$$\Pr_{w \leftarrow \mathbb{F}_q^d}[w \in \mathcal{A}] \leq \frac{\varepsilon/2048}{\varepsilon/16} = \frac{1}{128}.$$

We will now show that, for every $w \notin \mathcal{A}$, $J(w, \widehat{f}(w)) \geq \theta_+$, which implies that $\mathrm{Thr}(w, \widehat{f}(w)) = 1$.

For every $w \notin \mathcal{A}$, in each summand of J, choose v to be the first $D - 1$ coordinates of $g(y)$. When $w(y) = w$, its last coordinate is $\widehat{f}(w)$. Consequently,

$$\begin{aligned}
J(w, \widehat{f}(w)) &\geq \sum_{y : w(y) = w} \Pr[\mathbf{Y} = y \ \wedge \ \mathsf{Ext}(\mathbf{X}) = \sigma] \\
&= \Pr[w(\mathbf{Y}) = w \ \wedge \ \mathsf{Ext}(\mathbf{X}) = \sigma] \\
&= \Pr[\mathsf{Ext}(\mathbf{X}) = \sigma] \cdot \Pr[w(\mathbf{Y}) = w \mid \mathsf{Ext}(\mathbf{X}) = \sigma] \\
&= \Pr[\mathsf{Ext}(\mathbf{X}) = \sigma] \cdot Q(w) \\
&> (1 + \varepsilon) \cdot (1 - \varepsilon/16)\mu/q^d \\
&\geq (1 + 7\varepsilon/8)\mu/q^d \geq \theta_+.
\end{aligned}$$

The penultimate inequality uses $0 < \varepsilon \leq 1$. This gives that, for every $w \notin \mathcal{A}$, $\mathrm{Thr}(w, \widehat{f}(w)) = 1$. Thus the correct evaluation passes at all but at most a $1/128$ fraction of the points. This is where extraction from the conditioned source $\mathbf{Y}_\sigma$ spreads the heavy-output mass.

Restriction to a random line. Let $B(w) = \mathbf{1}\{\mathrm{Thr}(w, \widehat{f}(w)) = 1\}$. The bound above gives $\mathbb{E}_{w \leftarrow \mathbb{F}_q^d}[B(w)] \geq 127/128$. Applying Lemma 4.37 with $127/128 - 7/8 = 15/128$ gives

$$\Pr_{\mathbf{L} \leftarrow \mathsf{RandLine}}\left[\mathbb{E}_{t \leftarrow \mathbb{F}_q}[B(\mathbf{L}(t))] < 7/8 \right] \leq \frac{1}{4q(15/128)^2} < \frac{21}{q}.$$

For a nondegenerate line, $\mathrm{Test}_L(\widehat{f} \circ L) = \mathbb{E}_{t \leftarrow \mathbb{F}_q}[B(L(t))]$. Adding degeneracy, the correct polynomial satisfies $\mathrm{Test}_{\mathbf{L}}(\widehat{f} \circ \mathbf{L}) \geq 7/8$ except with probability at most $q^{-d} + 21/q = o(1/n) \leq \frac{1}{3n}$. □

10.2.4 Soundness

Lemma 10.12 (Soundness)**.** *Set*

$$K = \frac{8 \cdot 2^{k_1}}{\varepsilon\mu} = \frac{8 \cdot 2^{k_1+m}}{\varepsilon}. \tag{22}$$

For all sufficiently large n, with probability at least $1 - \frac{1}{3n}$ over $\mathbf{L} \leftarrow \mathsf{RandLine}$, the line is nondegenerate and the list

$$\mathcal{P}_{\mathbf{L}} = \{p \in \mathbb{F}_q[t] : \deg p \le \hat{h},\ \mathrm{Test}_{\mathbf{L}}(p) > 3/4\}$$

has size at most $8 \cdot K = o(q/\hat{h})$.

Proof. **Good points for soundness.** We will try to identify a large subset of points $w \in \mathbb{F}_q^d$ that are "good for soundness." For this purpose, we consider how much of the weight

$$P(w) = \Pr[\mathbf{W} = w] = \Pr[w(\mathbf{Y}) = w]$$

comes from $y \notin \mathcal{Y}_{\mathrm{valid}}$. Set

$$b(w) = \Pr[\mathbf{Y} \notin \mathcal{Y}_{\mathrm{valid}} \ \wedge\ \mathbf{W} = w].$$

A point $w \in \mathbb{F}_q^d$ is good for soundness if $P(w)$ is not too large and the weight contributed by non-valid y's is small. More specifically, we require that

$$P(w) \le (1 + \varepsilon/16)/q^d, \qquad b(w) \le \varepsilon\mu/(32q^d). \tag{23}$$

Let $\mathcal{H}$ be the set of points failing either condition. The first fails on at most a $1/128$ fraction of the points, by Lemma 4.5. Since $\sum_w b(w) \le \beta$, by Markov's inequality the second fails on a fraction at most $32\beta/(\varepsilon\mu) \le 1/128$. Thus $\mathcal{H}$ contains at most a $1/64$-fraction of the points.

Few field elements pass at a good point. Set

$$\mathcal{S}_w = \{a \in \mathbb{F}_q : \mathrm{Thr}(w, a) = 1\}.$$

By Equation (20), every $a \in \mathcal{S}_w$ satisfies $J(w, a) > \theta_-$. We show that $|\mathcal{S}_w| \le K$ for every $w \notin \mathcal{H}$.

Fix $w \notin \mathcal{H}$. For $a \in \mathbb{F}_q$, define

$$E_w(a) = \sum_{\substack{y \in \mathcal{Y}_{\mathrm{valid}} \\ w(y) = w \\ a \in R_y}} \Pr[\mathbf{Y} = y].$$

For every $y \in \mathcal{Y}_{\mathrm{valid}}$, every $v \in \mathbb{F}_q^{D-1}$, and every $a \in \mathbb{F}_q \setminus R_y$, Definitions 10.5 and 10.7 give

$$\Pr[\mathsf{IExt}((v, a), \mathrm{Src}(\mathbf{X})) = \sigma \mid \mathbf{Y} = y] \le \Theta = (1 + \varepsilon/8)\mu.$$

For every $y \in \mathcal{Y}_{\mathrm{valid}}$, every $v \in \mathbb{F}_q^{D-1}$, and every $a \in R_y$, we use the trivial bound

$$\Pr[\mathsf{IExt}((v, a), \mathrm{Src}(\mathbf{X})) = \sigma \mid \mathbf{Y} = y] \le 1.$$

For every $y \in \mathrm{Supp}(\mathbf{Y}) \setminus \mathcal{Y}_{\mathrm{valid}}$, every $v \in \mathbb{F}_q^{D-1}$, and every $a \in \mathbb{F}_q$, we use the trivial bound

$$\Pr[\mathsf{IExt}((v, a), \mathrm{Src}(\mathbf{X})) = \sigma \mid \mathbf{Y} = y] \le 1.$$

These three groups give

$$\begin{aligned} J(w,a) &= \sum_{y:w(y)=w} \max_{v\in\mathbb{F}_q^{D-1}} \Pr_{(\mathbf{Y},\mathbf{X})}[\mathbf{Y}=y \ \wedge\ \mathsf{IExt}((v,a),\mathrm{Src}(\mathbf{X}))=\sigma] \\ &\le \Theta\cdot \sum_{\substack{y\in\mathcal{Y}_{\mathrm{valid}}\\ w(y)=w\\ a\notin R_y}} \Pr[\mathbf{Y}=y] + \sum_{\substack{y\in\mathcal{Y}_{\mathrm{valid}}\\ w(y)=w\\ a\in R_y}} \Pr[\mathbf{Y}=y] \\ &\quad + \sum_{\substack{y\in\mathrm{Supp}(\mathbf{Y})\setminus\mathcal{Y}_{\mathrm{valid}}\\ w(y)=w}} \Pr[\mathbf{Y}=y] \\ &\le \Theta\cdot P(w) + E_w(a) + b(w). \end{aligned} \tag{24}$$

The last inequality bounds the mass of the first group by all of $P(w)$. Since $w\notin\mathcal{H}$, we have

$$\begin{aligned} \Theta\cdot P(w)+b(w) &\le \left((1+\varepsilon/8)\cdot(1+\varepsilon/16)+\varepsilon/32\right)\mu/q^d \\ &= \left(1+7\varepsilon/32+\varepsilon^2/128\right)\mu/q^d \le (1+\varepsilon/4)\mu/q^d. \end{aligned}$$

The last inequality uses $\varepsilon\le 1$. Recall that $\theta_- = (1+\varepsilon/2)\mu/q^d$. For every field element $a\in\mathcal{S}_w$, we therefore have

$$\begin{aligned} E_w(a) &\ge J(w,a) - \left(\Theta\cdot P(w)+b(w)\right) \\ &> \theta_- - (1+\varepsilon/4)\frac{\mu}{q^d} \\ &= \left((1+\varepsilon/2)-(1+\varepsilon/4)\right)\frac{\mu}{q^d} \\ &= \frac{\varepsilon\mu}{4q^d}. \end{aligned}$$

On the other hand, expanding the definition of $E_w(a)$, interchanging the sums, and using $|R_y|\le 2^{k_1}$, we have

$$\begin{aligned} \sum_{a\in\mathbb{F}_q} E_w(a) &= \sum_{\substack{y\in\mathcal{Y}_{\mathrm{valid}}\\ w(y)=w}} \Pr[\mathbf{Y}=y]\cdot|R_y| \\ &\le 2^{k_1}\cdot \sum_{\substack{y\in\mathcal{Y}_{\mathrm{valid}}\\ w(y)=w}} \Pr[\mathbf{Y}=y] \\ &\le 2^{k_1}\cdot \Pr[w(\mathbf{Y})=w] \\ &= 2^{k_1}\cdot P(w) \\ &\le 2^{k_1}\cdot\frac{1+\varepsilon/16}{q^d} \\ &\le \frac{17\cdot 2^{k_1}}{16q^d}. \end{aligned}$$

The last two inequalities use $w\notin\mathcal{H}$ and $\varepsilon\le 1$, respectively.

Therefore,

$$|\mathcal{S}_w|\cdot\frac{\varepsilon\mu}{4q^d} \le \sum_{a\in\mathcal{S}_w} E_w(a) \le \sum_{a\in\mathbb{F}_q} E_w(a) \le \frac{17\cdot 2^{k_1}}{16q^d},$$

and consequently

$$|\mathcal{S}_w| \le \frac{17 \cdot 2^{k_1}}{4\varepsilon\mu} \le K.$$

Restriction to a random line and list recovery. Since $\mathcal{H}$ contains at most a $1/64$-fraction of the points, applying Lemma 4.37 to its indicator gives

$$\Pr_{\mathbf{L} \leftarrow \mathsf{RandLine}}\left[\Pr_{t \leftarrow \mathbb{F}_q}[\mathbf{L}(t) \in \mathcal{H}] > \frac{1}{8}\right] \le \frac{1}{4q(1/8 - 1/64)^2} < \frac{21}{q}.$$

Fix a nondegenerate line L satisfying

$$\Pr_{t \leftarrow \mathbb{F}_q}[L(t) \in \mathcal{H}] \le \frac{1}{8}.$$

Recall that $\mathcal{S}_{L(t)} = \{a \in \mathbb{F}_q : \mathrm{Thr}(L(t), a) = 1\}$ and that $|\mathcal{S}_{L(t)}| \le K$ whenever $L(t) \notin \mathcal{H}$. Define

$$\mathcal{S}_t = \begin{cases} \mathcal{S}_{L(t)}, & L(t) \notin \mathcal{H}, \\ \varnothing, & L(t) \in \mathcal{H}. \end{cases}$$

Every list has size at most K. A polynomial p such that $\mathrm{Test}_L(p) > 3/4$ agrees with these lists on more than a $3/4 - 1/8 = 5/8$ fraction of the field elements $t \in \mathbb{F}_q$, hence on at least half. Lemma 4.38 with $\xi = 1/2$ bounds their number by $8 \cdot K$, provided $q \ge 8 \cdot \hat{h} \cdot K$. This condition holds for sufficiently large n, because

$$\log(8 \cdot \hat{h} \cdot K) = k_y^{1/4} + m + O(\log n) = o(\log q).$$

The same estimate gives $8 \cdot K = o(q/\hat{h})$. By a union bound,

$$\Pr_{\mathbf{L} \leftarrow \mathsf{RandLine}}\left[\mathbf{L} \text{ is degenerate } \vee \Pr_{t \leftarrow \mathbb{F}_q}[\mathbf{L}(t) \in \mathcal{H}] > \frac{1}{8}\right] \le q^{-d} + \frac{21}{q} = o(1/n) \le \frac{1}{3n}.$$

□

10.2.5 Constructing Thr

Lemma 10.13 (A small threshold circuit)**.** *There is a constant $C_{\mathrm{Thr}} > 0$, depending on $\alpha, c_\varepsilon, c_{\mathrm{size}}, e_1, e_2$ but independent of d, such that the following holds for all sufficiently large n. There is a nondeterministic circuit $\mathrm{Thr}(w, a)$ of size at most $n^{C_{\mathrm{Thr}}}$ satisfying Equation (19)–Equation (20) for every w, a. This circuit may depend on the circuits Samp and Post, the probability p_{post}, the output σ, and the selected seed i^* from Section 10.2.1.*

Proof. With external input (w, a) and uniform sampler coins $R \leftarrow \mathbf{U}_r$, define a nondeterministic circuit C by

$$C((w,a), y, v, R) = 1 \quad \Longleftrightarrow \quad \begin{cases} \mathsf{Post}(R) = 1, \\ \mathsf{SExt}(y, D) = w, \\ F_2(\mathsf{Samp}(R)) = y, \\ \mathsf{IExt}((v,a), \mathrm{Src}(\mathsf{Samp}(R))) = \sigma. \end{cases}$$

The circuit C guesses an accepting witness for $\mathsf{Post}(R)$. We will apply Lemma 8.1 to the circuit C, with $u = (w, a)$. By Definition 10.8,

$$S((w,a)) = p_{\mathrm{post}} \cdot J(w,a),$$

where S is defined in Lemma 8.1. Indeed, conditioning on $\mathsf{Post}(R) = 1$ gives the corresponding summand of J.

Apply Lemma 8.1 with threshold $p_{\mathrm{post}}\theta_+$ and $\eta = \varepsilon/16$. For $0 < \varepsilon \le 1$,

$$(1-\eta)\theta_+ = (1-\varepsilon/16)(1+3\varepsilon/4)\mu/q^d > (1+\varepsilon/2)\mu/q^d = \theta_-.$$

The resulting circuit Thr therefore satisfies Equation (19)–Equation (20). The threshold can be represented using $O(r + m + d\log q + \log(1/\varepsilon) + 1)$ bits. Here p_{post} is the hardwired probability from Section 10.2.1. Since p_{post} is an integer multiple of 2^{-r}, it can be described using $O(r+1)$ bits.

Circuit size. The circuit C evaluates $\mathsf{Samp}, F_1, F_2, \mathsf{SExt}, \mathsf{IExt}$ and verifies Post; it does not evaluate g or $\widehat{f}$. Lemma 8.1 gives a fixed polynomial bound in these circuit sizes, the input and witness lengths, the threshold bit length, and $1/\varepsilon$. For every fixed d, these quantities are polynomial in n with exponents independent of d: $r \le n^{c_{\mathrm{size}}}$, $\log q = k_y^{1/2} \le n^{1/2}$, $D = n^{c_D}$, $d\log q \le dn^{1/2}$, and $1/\varepsilon = n^{c_\varepsilon}$. The condenser circuit sizes were fixed independently of d in Section 10.1.1. Absorbing factors depending on d into the sufficiently large-n threshold gives the claimed exponent C_{Thr} independent of d. □

10.2.6 Reconstruction and completion of the proof

The argument used here is identical to that of Section 9.2.6, and we repeat it for completeness.

Applying the reconstruction lemma. By Lemmas 10.11 and 10.12, with probability at least $1 - \frac{2}{3n}$ over $\mathbf{L} \leftarrow \mathsf{RandLine}$,

$$\mathrm{Test}_{\mathbf{L}}(\widehat{f} \circ \mathbf{L}) \ge 7/8, \qquad |\mathcal{P}_{\mathbf{L}}| \le 8 \cdot K,$$

where $\mathcal{P}_L$ is defined in Lemma 8.2. Since $8 \cdot K = o(q/\hat{h})$, for sufficiently large n,

$$\frac{2}{3n} \le \frac{1}{32}, \qquad 8 \cdot K \le \frac{q}{32(\hat{h}+1)}.$$

Applying Lemma 8.2 to the circuit Thr from Lemma 10.13, there is a single-valued nondeterministic circuit C of size $\mathrm{poly}(\mathrm{size}(\mathrm{Thr}), \hat{h}, d, \log q)$ such that

$$\mathrm{ACCEPT}_C(x) = \{\widehat{f}(x)\} \qquad \text{for every } x \in \mathbb{F}_q^d.$$

Restricting to the hard function. Define $C_f(x)$ to simulate $C(\phi(x))$ and accept if it outputs 1. Since $\mathrm{ACCEPT}_C(\phi(x)) = \{\widehat{f}(\phi(x))\} = \{f(x)\}$, C_f computes f on $\{0,1\}^\ell$.

The natural encoding ϕ (from Theorem 4.34) has size polynomial in ℓ and $\log q$, with an absolute polynomial bound. Since $\ell = d\log n$ and $\hat{h} = d(n-1)$, incorporating this encoding gives absolute constants $C_0, a_0 > 0$ such that

$$\mathrm{size}(C_f) \le C_0(\mathrm{size}(\mathrm{Thr}) + \hat{h} + d + \log q + 1)^{a_0}.$$

By Lemma 10.13, $\mathrm{size}(\mathrm{Thr}) \le n^{C_{\mathrm{Thr}}}$ with C_{Thr} independent of d. For each fixed d, using $\hat{h} < d \cdot n$ and $\log q \le n$, the quantity in parentheses is at most $n^{C_{\mathrm{Thr}}+3}$ for sufficiently large n. Thus

$$\mathrm{size}(C_f) \le n^{C_{\mathrm{rec}}}, \qquad C_{\mathrm{rec}} = a_0(C_{\mathrm{Thr}} + 3) + 1,$$

where C_{rec} is independent of d.

Choosing the dimension and deriving the contradiction. The remaining component parameters were fixed independently of d. We can therefore first obtain C_{Thr} and C_{rec}, and then choose a sufficiently large integer $d \geq 2$ such that

$$\beta_{\mathrm{hard}} \cdot d > C_{\mathrm{rec}} + 1.$$

With this choice of d fixed, all the preceding requirements hold for sufficiently large n. Since $\ell = d \log n$,

$$\mathrm{size}(C_f) \leq n^{C_{\mathrm{rec}}} < n^{\beta_{\mathrm{hard}} \cdot d} = 2^{\beta_{\mathrm{hard}} \ell},$$

contradicting the hardness of f. Thus, for every source in the specified class with min-entropy at least k and every output σ, $\Pr[\mathsf{Ext}(\mathbf{X}) = \sigma] \leq e^{\varepsilon} \cdot 2^{-m}$. This proves Theorem 4.41, using the errorless condenser supplied by Theorem 4.42.

Part II: Additive Extractors

In this part, we construct the following additive extractors for $\mathcal{C}^{\mathrm{NP}}_{n,n^c}$:

- **One-bit extraction from logarithmic entropy** (Theorem 13.1): Under H_{NP}, a polynomial-time extractor producing one bit from entropy $C \log n$, with arbitrarily small constant error.
- **Almost-all extraction from polylogarithmic entropy** (Theorem 14.2): Under H_{NP}, an extractor producing $\lfloor (1-\eta)k \rfloor$ bits from entropy $k > (\log n)^A$, with error n^{-e} and running time $2^{O(\log^2 n / \log k)}$.
- **Polynomial-time almost-all extraction** (Theorem 15.1): Under H_{Σ_2}, a polynomial-time extractor producing $\lfloor (1-\eta)k \rfloor$ bits from entropy $k > (\log n)^A$, with error n^{-e}.

Part II is organized as follows: In Section 11, we construct an errorless strong seeded condenser. Section 12 develops additive and multiplicative correlation breakers for jointly samplable sources. We use these tools in Section 13 to construct a polynomial-time one-bit extractor for sources with logarithmic min-entropy under H_{NP}. In Section 14, we use lightest-bin selection and correlation breaking to extract from sources with polylogarithmic min-entropy under the same assumption. Finally, Section 15 improves the running time of this construction to polynomial under the stronger assumption H_{Σ_2}.

11 Errorless seeded condensation by hashing

In this section, we construct an errorless strong seeded condenser (Definition 4.16). Enumerating its seeds gives polynomially many jointly samplable rows, most of which have almost full min-entropy. This is going to be the first step of all our extractor constructions.

Lemma 11.1 (Errorless strong seeded condensation)**.** *Assume H_{NP} and fix a constant $c > 0$. There are constants $C > 1$, $\eta > 0$, and an integer $\Delta \geq 1$ such that, for all sufficiently large n and every integer k satisfying $C \log n \leq k \leq n$, there is a uniformly polynomial-time computable errorless strong seeded $(k, k - \log k, R^{-\eta})$-condenser $\mathsf{Cond} : \{0,1\}^n \times [R] \to \{0,1\}^m$ for $\mathcal{C}^{\mathrm{NP}}_{n,n^c}$, where $n \leq R = \mathrm{poly}(n)$ and $m = \lfloor k - \log k \rfloor + \Delta$.*

Write $h_i(x) = \mathsf{Cond}(x, i)$. For every such source $\mathbf{X}$, the lemma gives a set $B_{\mathbf{X}} \subseteq [R]$ with $|B_{\mathbf{X}}| \leq R^{1-\eta}$ such that $\mathbf{H}_{\infty}(h_i(\mathbf{X})) \geq k - \log k \geq m - \Delta$ for every $i \in [R] \setminus B_{\mathbf{X}}$.

Proof. We first prove the entropy guarantee for a random hash, then use the counting theorem and pseudorandom generator from Preliminaries to derandomize the choice.

Step 1: Random hashing. Fix a sufficiently large integer Δ, and set $r = k$ and $m = \lfloor k - \log k \rfloor + \Delta$. Let $\mathbf{H}$ be a uniformly random hash from an r-wise independent family $\mathcal{H}$ of maps $\{0,1\}^n \to \{0,1\}^m$. Identify $\{0,1\}^n$ with a subset of $\mathbb{F}_{2^v}$, where $v = \max\{n, m\}$, and evaluate a polynomial of degree at most $r-1$ with independent uniform coefficients. Retaining m fixed binary coordinates of the output gives such a family. Its hashes are computable in polynomial time and indexed by $q = rv = O(n^2)$ random bits.

Fix $\mathbf{X} \in \mathcal{C}^{\mathrm{NP}}_{n,n^c}$ with $\mathbf{H}_\infty(\mathbf{X}) \geq k$. For independent copies $\mathbf{X}_1, \ldots, \mathbf{X}_r$, define

$$W(h) = \Pr\big[\mathbf{X}_1, \ldots, \mathbf{X}_r \text{ are distinct and } h(\mathbf{X}_1) = \cdots = h(\mathbf{X}_r)\big].$$

By r-wise independence, $\mathbb{E}_{\mathbf{H}}[W(\mathbf{H})] \leq 2^{-m(r-1)}$. Set $\tau_H = r!\, 2^{-rk}$ and $\tau_L = \tau_H/2$. Then

$$\begin{aligned}
\Pr[W(\mathbf{H}) > \tau_L] &\leq \frac{\mathbb{E}_{\mathbf{H}}[W(\mathbf{H})]}{\tau_L} \\
&\leq \frac{2^{1+rk-m(r-1)}}{r!} \\
&\leq \frac{2^{1+k^2-(k-\log k+\Delta-1)(k-1)}}{k!} \\
&\leq \frac{2^{k+1-(\Delta-1)(k-1)} k^{k-1}}{(k/e)^k} \\
&= \frac{2^\Delta}{k}\big(e\, 2^{2-\Delta}\big)^k \leq 2^{-\Omega(k)} \leq n^{-2},
\end{aligned}$$

where the last two inequalities hold for sufficiently large Δ and C.

Now fix $h \in \mathcal{H}$ and suppose $p_z = \Pr[h(\mathbf{X}) = z] > r2^{-k}$ for some $z \in \{0,1\}^m$. Given distinct values $x_1, \ldots, x_j$ with $h(x_1) = \cdots = h(x_j) = z$, the next independent sample has hash value z and avoids $x_1, \ldots, x_j$ with probability at least $p_z - j2^{-k}$. Hence

$$W(h) \geq p_z(p_z - 2^{-k}) \cdots (p_z - (r-1)2^{-k}) > r!\, 2^{-rk} = \tau_H.$$

Thus $W(h) \leq \tau_H$ implies $\mathbf{H}_\infty(h(\mathbf{X})) \geq k - \log r = k - \log k$.

Step 2: Derandomizing the construction. We use Theorem 4.39 to test whether the collision probability exceeds τ_H, with rejection guaranteed below τ_L, and then apply Lemma 4.40.

Let $(\mathsf{Samp}, \mathsf{Post})$ be a deterministic sampling circuit with nondeterministic postselection of total size at most n^c for $\mathbf{X}$. Write ℓ for its input length and $p_{\mathrm{post}} = \Pr[\mathsf{Post}(\mathbf{U}_\ell) = 1] > 0$. For each hash description $u \in \{0,1\}^q$, let C_u accept an ordered r-tuple $(\rho_1, \ldots, \rho_r) \in (\{0,1\}^\ell)^r$ exactly when

$$\begin{gathered}
\mathsf{Post}(\rho_j) = 1 \quad \text{for every } j \in [r], \\
\mathsf{Samp}(\rho_1), \ldots, \mathsf{Samp}(\rho_r) \text{ are distinct}, \\
h_u(\mathsf{Samp}(\rho_1)) = \cdots = h_u(\mathsf{Samp}(\rho_r)).
\end{gathered}$$

The circuit C_u is nondeterministic, which is allowed by Theorem 4.39. Its witnesses certify the postselection tests and are not counted as ordinary inputs. With $(\mathsf{Samp}, \mathsf{Post})$ fixed, the description of C_u is computable from u in polynomial time. Conditioning on all r inputs passing Post gives r independent copies of $\mathbf{X}$, so

$$\Pr[C_u(\mathbf{U}_{r\ell}) = 1] = p_{\mathrm{post}}^r W(h_u).$$

Apply Theorem 4.39 with $\gamma = p_{\mathrm{post}}^r \tau_H$ and $\lambda = 1/2$. The threshold γ is positive, at most one, and has polynomially many bits: p_{post} has denominator 2^ℓ, and $\tau_H = r!\, 2^{-rk}$. Since $\tau_L = \tau_H/2 < e^{-1/2}\tau_H$, the resulting nondeterministic test $T_{\mathsf{Samp},\mathsf{Post}}$ satisfies

$$\begin{aligned}
W(h_u) > \tau_H &\implies T_{\mathsf{Samp},\mathsf{Post}}(u) = 1, \\
W(h_u) \leq \tau_L &\implies T_{\mathsf{Samp},\mathsf{Post}}(u) = 0.
\end{aligned}$$

The sampling circuit and γ are fixed in the test's description; the test need not compute the postselection probability. Its size is bounded by a common polynomial in n, over all sampling circuits of size at most n^c and all $k \le n$.

Use Lemma 4.40 for this common size bound and error n^{-2}. This gives a generator $G : \{0,1\}^d \to \{0,1\}^q$ with $d = O(\log n)$. Pad the seed with unused bits if necessary so that $d \ge \lceil \log n \rceil$. Put $R = 2^d$, enumerate the seeds as $\sigma_1, \ldots, \sigma_R$, and define $\mathsf{Cond}(x, i) = h_{G(\sigma_i)}(x)$.

Since $T_{\mathsf{Samp,Post}}$ rejects whenever $W(h_u) \le \tau_L$, Step 1 gives $\Pr[T_{\mathsf{Samp,Post}}(\mathbf{U}_q) = 1] \le n^{-2}$. For $B_{\mathbf{X}} = \{i : T_{\mathsf{Samp,Post}}(G(\sigma_i)) = 1\}$, we have

$$\begin{aligned}\frac{|B_{\mathbf{X}}|}{R} &= \Pr[T_{\mathsf{Samp,Post}}(G(\mathbf{U}_d)) = 1] \\ &\le \Pr[T_{\mathsf{Samp,Post}}(\mathbf{U}_q) = 1] + n^{-2} \le 2n^{-2}.\end{aligned}$$

Choose a constant $\eta > 0$ such that $R^\eta \le n$. Then $|B_{\mathbf{X}}| \le R^{1-\eta}$ for sufficiently large n. For every $i \notin B_{\mathbf{X}}$, the test's acceptance guarantee implies $W(h_{G(\sigma_i)}) \le \tau_H$, and Step 1 gives $\mathbf{H}_\infty(\mathsf{Cond}(\mathbf{X}, i)) \ge k - \log k$. □

12 Decorrelation of correlated samplable sources

12.1 Section Overview

We construct correlation breakers for indexed sources/rows that are jointly samplable with postselection. The correlation breaker will ultimately allow us to convert indexed (high-entropy) correlated rows into something resembling a non-oblivious bit-fixing source, on which we can then invoke techniques from the two-source extractor literature. To do this, the correlation breaker can be locally applied to each individual source. The resulting outputs have the property that outputs on good sources are effectively independent of the rest.

The construction follows the precise template of Ball, Dachman-Soled, and Loss [BDL22; BD25]'s approach to constructing non-malleable extractors. Essentially, Ball and Dachman-Soled [BD25] used their non-malleable extractor to instantiate this precise object—without naming it—for their application in black-box barriers. The present construction instantiates their recipe with stronger ingredients and provides a better analysis that yields the result from better assumptions. In particular, we construct decorrelators with additive error under H_{NP}, and pointwise multiplicative error with an additive slack under H_{Σ_2}.

The recipe of Ball and Dachman-Soled is as follows:

1. Extract a short seed, σ from a row, v, using an extractor for post-selecting sources, $\mathsf{Ext}_{\mathrm{pre}}$: $\sigma = \mathsf{Ext}_{\mathrm{pre}}(v)$.

   For our application, we will promise "good" rows v have extremely high min-entropy (just small poly deficiency), so it suffices to use Shaltiel's extractor [Sha25b], or the construction from Section 4.3.7. Both of these constructions follow from the H_{NP} assumption. For the additive-error correlation breaker, an additive-error extractor suffices.

2. Expand the seed, σ, with a *seed-extending multiplicative pseudorandom generator*, G: $G(\sigma) = \sigma \parallel y$. For the additive version, we require G to be secure against nondeterministic circuits.

   A seed-extending multiplicative pseudorandom generator for nondeterministic circuits is a PRG with special properties: (a) the seed is output as a prefix, (b) fooling is with respect to one-sided multiplicative error: for any distinguisher* D,

$$\Pr[D(G(\mathbf{U}_k)) = 1] \le e^{\varepsilon} \Pr[D(\mathbf{U}_N) = 1] + \delta.$$

In this inequality, ε is inverse-polynomial and δ can be substantially smaller.[19] To get the desired properties and parameters, we compose a seed-extending short stretch PRG [Sha25b] with a long stretch PRG [DS26].

*For the additive error variant, we require security against bounded-polynomial-size nondeterministic circuits, which follows from H_{NP}. For the multiplicative variant, we require security against nondeterministic NP-circuits, which follows from H_{Σ_2}.

3. Apply a two-source t-non-malleable extractor, $\mathsf{2NMExt}$, to the expanded seed (σ, y), and the original indexed row v: $z = \mathsf{2NMExt}((\sigma, y), v)$.

   A two-source t-non-malleable extractor extracts randomness from two *independent* high min-entropy sources, even given its output on t tampered variants of this pair. Specifically, for independent high min-entropy $\mathbf{X}, \mathbf{Y}$, and arbitrary fixed-point-free[20] pairs of functions (f_i, g_i):

$$\mathsf{SD}\Big(\big(\mathsf{2NMExt}(\mathbf{X},\mathbf{Y}), (\mathsf{2NMExt}(f_i(\mathbf{X}), g_i(\mathbf{Y})))_{i=1}^{t}\big), \big(\mathbf{U}_q, (\mathsf{2NMExt}(f_i(\mathbf{X}), g_i(\mathbf{Y})))_{i=1}^{t}\big)\Big) < \varepsilon.$$

   We use two constructions for $\mathsf{2NMExt}$. For constant t and one output bit, Li's construction [Li23], combined with the reduction of Chattopadhyay and Liao [CL22], gives error $2^{-\Omega(L)}$. This is the ingredient used in our one-bit extractor at logarithmic entropy. When t or the output length q grows with L, we use Chattopadhyay, Goyal, and Li [CGL20], obtaining error $2^{-cL^{\gamma}}$ whenever $(t+1)q \le cL^{\gamma}$, for absolute constants $c, \gamma > 0$. In both cases, we apply the standard conversion to make the extractor strong in either input.

An obvious obstacle is that the two inputs to the non-malleable extractor are highly dependent: one is in fact completely determined by the other! However, following Ball and Dachman-Soled's analysis [BDL22; BD25], the computational guarantee will grant us enough "pseudo-independence" to exploit the t-non-malleability of $\mathsf{2NMExt}$. More precisely, we turn a failure of correlation breaking into a distinguisher for the pseudorandom generator.

The challenge in this setting is that any such reduction seemingly must run the tampering experiment, but that involves evaluating the PRG on adversarially correlated outputs. For the PRGs used, if an efficient reduction (within the circuit-size bound) could evaluate them then they would be insecure. Instead, to escape this contradiction, [BDL22; BD25] realized it suffices to allow an untrusted Merlin to provide the tampered PRG evaluations. In particular, it suffices to use the failure of correlation breaking to construct a constant-round, private-coin interactive proof that distinguishes the PRG from random. This is sufficient, because constant-round interactive proofs can be simulated by nondeterministic circuits: $\mathsf{IP}[O(1)] \subseteq \mathsf{NP}/\operatorname{poly}$ [GS86; BM88; Adl78].

We sketch the argument at a high level before proceeding. To distinguish a generated pair $(\sigma, y) = G(\sigma)$ from a uniform pair $(\sigma, y) \leftarrow \mathbf{U}$, we assume for contradiction that the correlation breaker doesn't work. Namely, if g is the index of a "good" (i.e. high entropy) row

$$(\mathsf{CB}(\mathbf{V}_g), (\mathsf{CB}(\mathbf{V}_i))_{i\neq g}) \not\approx (\mathbf{U}, (\mathsf{CB}(\mathbf{V}_i))_{i\neq g}).$$

To exploit this, our interactive proof reduction will use (σ, y) to simulate either the LHS (if (σ, y) is pseudorandom) or the RHS (if (σ, y) is truly random).

Concretely, the verifier samples the rows privately, giving up if the "good" pre-extracted sample misses σ.[21] Then the prover supplies the other rows' purported generator outputs. From here, we can use an additive

[19] The seed length grows as $O(\log(1/\delta))$; consequently, polynomial stretch is compatible with additive slack exponentially small in the seed length.

[20] For each i, either f_i or g_i has no fixed points.

[21] The event of hitting σ happens with very low probability, but to collapse to $\mathsf{NP}/\operatorname{poly}(n)$ we just need a multiplicative gap around a fixed threshold (due to the power of the Goldwasser–Sipser transformation [GS86; GL20]).

or multiplicative test on the distance between LHS and RHS above. Critically, Arthur does not actually evaluate the PRG himself, and so can be much more efficient (with Merlin's assistance of course).

Why does this work? Suppose that (σ, y) is pseudorandom. In this case, there is a gap between LHS and RHS, and Merlin can realize it by simply responding with the actual PRG outputs on the correlated seeds sent by Arthur. On the other hand, if (σ, y) is truly random, then the total amount of information sent by Arthur is small and thus most of the entropy of $\mathbf{V}_g$ is retained even if we condition on a given σ and Arthur message. We will thus think about $\mathbf{V}_g$ and y being the inputs to the two-source non-malleable extractor, both of which have high entropy (recall that in this case y is uniform and independent of σ). We thus can show that any Merlin strategy here that convinces Arthur to accept (conditioned on hitting σ with $\mathsf{Ext}_{\mathrm{pre}}$ and the initial message from Arthur of correlated seeds) would yield a split-state tampering strategy on the two source extractor. The left tamperings on the y portion are simply the Merlin replies. The tamperings on the $\mathbf{V}_g$ portion are induced by conditional sampling $\mathbf{V}_i$ consistent with the transcript sent (and conditioned on). Thus, because Arthur and Merlin represent a valid split-state t-tampering strategy, and the non-malleable extractor resists this, the LHS experiment (with random y, instead of pseudorandom y) and the RHS experiment (again with random y, instead of pseudorandom y) are close. Thus, the final test distinguishing LHS from RHS will only work (on average) when the input is pseudorandom and hence allows Arthur to efficiently distinguish PRG outputs from random (with Merlin's help).

12.2 Computational preliminaries

We first recall some useful background.

Nondeterminism and NP gates/oracles. A nondeterministic circuit C has an associated deterministic circuit D (with additional inputs for a witness: $C(x) = 1$ iff there exists a witness a for which $D(x, a) = 1$). An NP-circuit is allowed SAT-gates. A nondeterministic NP-circuit is a nondeterministic circuit where the underlying deterministic circuit is allowed NP-gates. Nondeterministic NP-circuits are equivalent, up to polynomial size overhead, to Σ_2 circuits.

Write $\mathcal{O} \in \{\varnothing, \mathsf{SAT}\}$ for these two verifier models. The corresponding hardness assumption is H_{NP} or H_{Σ_2} from Section 4.2: in both cases the hard function belongs to ordinary E, and the assumed lower bound is against nondeterministic circuits with access to $\mathcal{O}$. Fixing inputs, deterministic composition, and conjunction preserve the relevant circuit class. In the oracle model, existential guesses and conjunctions can be combined into one existential block and one universal block.

Seed-extending pseudorandomness. A generator is seed-extending if $G(\sigma) = \sigma \parallel G'(\sigma)$. The visible seed lets a verifier match its privately sampled row to the input without evaluating G. The constructions of [Sha25b; DS26] and their reconstruction arguments give the following two versions of the same ingredient. We verify the oracle extension and the composition in Section 12.8.

Lemma 12.1 (Seed-extending multiplicative PRG)**.** *Fix $\mathcal{O} \in \{\varnothing, \mathsf{SAT}\}$ and assume H_{NP} when $\mathcal{O} = \varnothing$, or H_{Σ_2} when $\mathcal{O} = \mathsf{SAT}$. There is a constant C such that, whenever $s \geq 2$, $0 < \delta \leq 1/s$ and $C \log(1/\delta) \leq L \leq s$, there is a uniformly $\mathrm{poly}(s)$-time computable function $G : \{0,1\}^{\ell} \to \{0,1\}^{L}$, with $\ell = O(\log(1/\delta))$ and $G(\sigma) = \sigma \parallel G'(\sigma)$, such that every size-$s$ nondeterministic $\mathcal{O}$-oracle circuit D satisfies*

$$\Pr[D(G(\mathbf{U}_\ell)) = 1] \leq e^{1/s} \Pr[D(\mathbf{U}_L) = 1] + \delta.$$

The algorithm evaluating G uses no oracle; its constants may depend on the fixed hard function in the relevant assumption. An arbitrary error bound δ is implemented using the integer precision $\lceil \log(1/\delta) \rceil$ and the smaller power-of-two error $2^{-\lceil \log(1/\delta) \rceil}$.

Corollary 12.2 (Additive guarantee)**.** *Assume H_{NP}. For every fixed $a, b > 0$, there is a constant C such that, whenever $C \log S \leq L \leq S$, there is a uniformly $\mathrm{poly}(S)$-time computable seed-extending generator*

$G : \{0,1\}^\ell \to \{0,1\}^L$ with $\ell = O(\log S)$ such that every nondeterministic circuit D of size at most S^a satisfies

$$\Pr[D(G(\mathbf{U}_\ell)) = 1] \le \Pr[D(\mathbf{U}_L) = 1] + S^{-b}.$$

Proof. Choose a fixed integer $K \ge \max\{1, a, b+3\}$, set $s = \lceil S^K \rceil$, and use Lemma 12.1 with $\mathcal{O} = \varnothing$ and $\delta = 1/s$. Since $e^{1/s} - 1 + 1/s \le 3/s \le S^{-b}$ for $S \ge 2$, the claim follows. □

Interactive proofs and round collapse. A private-coin protocol $\mathcal{P}$ has an unbounded prover and a randomized verifier. For an input w, let $\mathrm{val}_{\mathcal{P}}(w)$ be its maximum acceptance probability on w, over all prover strategies. When the protocol is clear from context, we abbreviate $\mathrm{val}_{\mathcal{P}}(w)$ to $\mathrm{val}(w)$. A public-coin verifier sends fresh uniform random bits in each message. In an AM protocol Arthur sends one random string and Merlin replies once.

The following lemma demonstrates that a constant-round private coin interactive proof with a sufficiently large constant *multiplicative* gap between acceptance probabilities in the case of yes vs no instances suffices for compilation down to a nondeterministic circuit. The resulting nondeterministic circuit is only polynomially larger than the complexity of the private coin Arthur. This process follows from composing classical reductions: Goldwasser–Sipser public-coin simulation [GS86], in the small-acceptance-probability formulation of Goldreich–Leshkowitz [GL20], Babai–Moran round reduction [BM88], and, finally, Adleman's trick [Adl78]. We state the lemma without formally proving it.

Lemma 12.3 (Collapse of a constant-round protocol)**.** *For every fixed integer $r \ge 1$, there are a constant $B > 1$ and a polynomial p_r with the following property, for either $\mathcal{O} \in \{\varnothing, \mathsf{SAT}\}$. Let $\mathcal{P}$ be a private-coin protocol with at most r rounds, whose verifier is a deterministic $\mathcal{O}$-oracle circuit once its random tape is fixed. Suppose its verifier size, input length, total randomness, and total communication are at most T. Let c_N, c_Y be rationals with descriptions of at most T bits, where $0 \le c_N < c_Y \le 1$ and $c_Y \ge B^r c_N$. There is a nondeterministic $\mathcal{O}$-oracle circuit of size $p_r(T)$ accepting every w with $\mathrm{val}_{\mathcal{P}}(w) \ge c_Y$ and rejecting every w with $\mathrm{val}_{\mathcal{P}}(w) \le c_N$. In particular, the size has no dependence on $1/c_Y$ or $1/c_N$.*

In the multiplicative reduction of Lemma 12.13, the verifier uses a SAT oracle for approximate counting. Applying Lemma 12.3 with $\mathcal{O} = \mathsf{SAT}$ therefore produces a nondeterministic SAT-oracle distinguisher, equivalently a Σ_2 circuit up to polynomial size overhead. This is why the multiplicative construction uses the $\mathcal{O} = \mathsf{SAT}$ version of Lemma 12.1, obtained under H_{Σ_2}.

12.3 Correlation Breaker Ingredients

We use the following setup for convenience.

An indexed row is a pair $\mathbf{V}_i = (\alpha_i, \mathbf{W}_i)$, identified with the concatenation $\alpha_i \,\|\, \mathbf{W}_i \in \{0,1\}^L$, where $\alpha_i \in \{0,1\}^h$ is fixed and the labels are distinct, and $\mathbf{W}_i \in \{0,1\}^{L-h}$ is the random part. Thus L denotes the total indexed-row length, including the label. Throughout, the entire tuple is sampled by a deterministic circuit Samp with a nondeterministic postselection circuit Post, of total size at most S, conditioned on $\mathsf{Post}(\mathbf{U}_{r_0}) = 1$, where $r_0 \le S$ and the conditioning event has positive probability. Write $\mathsf{Post}(R) = 1$ iff $\exists a_0\, Q(R, a_0) = 1$, where Q is deterministic and its size and witness length are at most S. There is no lower bound on that probability. A good row $\mathbf{V}_g$ has $\mathbf{H}_\infty(\mathbf{V}_g) \ge L - \Delta$; this entropy already includes the loss from its fixed index.

Preliminary extraction. We use Lemma 4.30 to construct a multiplicative extractor for postselecting sources:

$$\mathsf{Ext}_{\mathrm{pre}} : \{0,1\}^L \to \{0,1\}^\ell.$$

This map has the property that under H_{NP}, a good row can be mapped to ℓ bits with point probabilities at most $e^\nu 2^{-\ell}$, provided $L \ge C_0(\Delta + \ell + \log S + \log(1/\nu))$. For the additive case, Corollary 4.31 gives

$\ell = \lfloor c_0 L \rfloor$ bits within inverse-polynomial statistical distance of uniform. For the multiplicative case we use the pointwise bound with $\nu = 1/100$. (Both guarantees are preserved by taking an output prefix.)

Non-malleability for correlated rows. We call all the inputs besides the good input the "companion rows." The distinct indices ensure that each companion input differs from the good input. The following lemma shows how to apply ordinary two-source non-malleability even when the companion rows are not deterministic functions of the good row.

Lemma 12.4. *Let* $\mathsf{2NMExt} : \{0,1\}^L \times \{0,1\}^L \to \{0,1\}^q$ *be* t*-non-malleable for min-entropy* k_{NM} *and error* $\varepsilon_{\mathsf{NM}}$*, strong in its first input. Let* $(\mathbf{V}_i)_{i \in I}$ *be at most* $t+1$ *rows with fixed, distinct indices, fix* $g \in I$*, and let* $\mathbf{Z} \in \{0,1\}^L$ *be independent of this tuple. If* $\mathbf{H}_\infty(\mathbf{Z}), \mathbf{H}_\infty(\mathbf{V}_g) \geq k_{\mathsf{NM}}$*, then for arbitrary functions* $f_i : \{0,1\}^L \to \{0,1\}^L$*,*

$$\mathsf{SD}\Big(\big(\mathsf{2NMExt}(\mathbf{Z}, \mathbf{V}_g), \mathbf{Z}, (\mathsf{2NMExt}(f_i(\mathbf{Z}), \mathbf{V}_i))_{i \neq g}\big), \\ \big(\mathbf{U}_q, \mathbf{Z}, (\mathsf{2NMExt}(f_i(\mathbf{Z}), \mathbf{V}_i))_{i \neq g}\big)\Big) \leq \varepsilon_{\mathsf{NM}}.$$

To see this, for each possible value v of $\mathbf{V}_g$, choose a tuple with the conditional distribution of $(\mathbf{V}_i)_{i \neq g}$ given $\mathbf{V}_g = v$. Make all these choices independently of $(\mathbf{Z}, \mathbf{V}_g)$. Fixing the choices expresses each $\mathbf{V}_i$ as a deterministic function of $\mathbf{V}_g$. On the support of $\mathbf{V}_g$, this function changes its fixed index to the distinct index of $\mathbf{V}_i$, so it has no fixed point. Extend it outside the support without fixed points. Non-malleability now applies to the independent sources $\mathbf{Z}, \mathbf{V}_g$ with the fixed functions, and averaging over the choices gives the lemma.

For a constant number of rows, we use the following quantitative consequence of [CL22; Li23]. The extension to t tamperings follows from the standard advice-generator construction using the correlation breaker of [Li23, Theorem 7.12], based on the reduction of [CL22], with advice length $O(\lambda)$ and error parameter $2^{-\lambda}$, where $\lambda = \alpha L/(t+1)^3$ for a sufficiently small absolute constant $\alpha > 0$. The standard conversion of [ADNOPRS19, Lemma 54] makes the extractor strong in either input, after adjusting the absolute constants.

Theorem 12.5. *There are absolute constants* $c, C > 0$ *such that, for all integers* $t, L \geq 1$ *satisfying* $L \geq C(t+1)^3(t + \log_2 L)$*, there is an explicit function* $\mathsf{2NMExt} : \{0,1\}^L \times \{0,1\}^L \to \{0,1\}$ *that is* t*-non-malleable for min-entropy* $L - cL/(t+1)^3$ *and error* $2^{-cL/(t+1)^3}$*, strong in either input and computable in time polynomial in* L *and* t*.*

For more rows and output bits, we apply the same conversion to the extractor of [CGL20]. The restriction on Dq accounts for the factor 2^{Dq} in the error bound of the conversion.

Lemma 12.6 (Strong non-malleability for many tamperings)**.** *There are absolute constants* $\gamma, c > 0$ *such that, for all sufficiently large* L *and integers* $D, q \geq 1$ *satisfying* $Dq \leq cL^\gamma$*, there is an explicit polynomial-time function* $\mathsf{2NMExt} : \{0,1\}^L \times \{0,1\}^L \to \{0,1\}^q$ *that is* $(D-1)$*-non-malleable for min-entropy* $L - cL^\gamma$ *and error* 2^{-cL^γ}*, strong in either input.*

In particular, fix positive constants $\rho_t, \rho_o, \rho_k, \rho_e$ satisfying $\rho_t + \rho_o < \gamma$, $\rho_k < \gamma$, and $\rho_e < \gamma$. For all sufficiently large L, Lemma 12.6 applies whenever $D - 1 \leq L^{\rho_t}$ and $q \leq L^{\rho_o}$, with min-entropy requirement $L - L^{\rho_k}$ and error $2^{-L^{\rho_e}}$.

12.4 Correlation Breaker Construction Template

Given $\mathsf{Ext}_{\mathrm{pre}} : \{0,1\}^L \to \{0,1\}^\ell$, a seed-extending generator $G : \{0,1\}^\ell \to \{0,1\}^L$, and a two-source non-malleable extractor $\mathsf{2NMExt} : \{0,1\}^L \times \{0,1\}^L \to \{0,1\}^q$, define $\mathsf{CB} : \{0,1\}^h \times \{0,1\}^{L-h} \to \{0,1\}^q$ by

$$\mathsf{CB}(\alpha, v) = \mathsf{2NMExt}\big(G(\mathsf{Ext}_{\mathrm{pre}}(\alpha \parallel v)), \alpha \parallel v\big). \tag{25}$$

Here $\alpha \in \{0,1\}^h$ is the fixed row label and $v \in \{0,1\}^{L-h}$ is the random part. The label remains explicit in $\mathsf{2NMExt}$'s second input. Moving to a companion row changes this label, which supplies the no-fixed-point condition in Lemma 12.4. No such condition is needed on the left tampering, so $\mathsf{Ext}_{\mathrm{pre}}$ may process the label without preserving it in its output. Since the label is fixed, it contributes no entropy. For a complete indexed row $u = \alpha \,\|\, v$, we also write $\mathsf{CB}(u) = \mathsf{CB}(\alpha, v) = \mathsf{2NMExt}(G(\mathsf{Ext}_{\mathrm{pre}}(u)), u)$; in particular, $\mathsf{CB}(\mathbf{V}_i)$ below includes the label α_i on the right. The notation $\mathsf{CB}_{\mathrm{mult}}$ denotes this same construction with ingredients and parameter choices chosen for the multiplicative guarantee.

Parameter key. There are numerous parameters involved in this construction. We introduce them in the order in which we set them. We write D for the maximum number of rows, denoted D_* in the multiplicative case.

(i) *Source and output parameters:* L, S, h, Δ, D, q. Each row has length L, its fixed index has length h, and the joint sampling-and-postselection circuit has size at most S. A good row has min-entropy at least $L - \Delta$, already accounting for its fixed index. The correlation breaker outputs q bits per row.

(ii) *Error parameters:* θ, η *or* ϑ, ξ, ν. For additive error, fix $b > 0$, set $\theta = S^{-b}$, and take $\eta = \theta^3/2^{20}$ as the error bound for both extraction ingredients; see Lemma 12.8. For multiplicative error, the target factor e^{ϑ} and pointwise slack ξ satisfy (32). Take $\nu = 1/100$ for the preliminary extractor's pointwise cap and require non-malleable-extractor error at most $\xi/100$; see Lemma 12.13.

(iii) *Extractor parameters:* m_P *and* t. First choose a preliminary extractor P_0 with $m_P = \lfloor c_0 L \rfloor$ output bits, for a sufficiently small fixed $c_0 > 0$. Its length and accuracy constraints are given in Lemma 4.30 and Corollary 4.31. Choose $\mathsf{2NMExt}$ with $t = D - 1$ tamperings and q output bits. The available entropy, output-length, and error guarantees are Theorem 12.5 and Lemma 12.6.

(iv) *PRG parameters:* s, δ_G, ℓ. With P_0 and $\mathsf{2NMExt}$ fixed, bound the interactive verifier's size and its size after collapse. Choose G to fool circuits of that size. In the additive case, require additive error at most $\theta/4$, as in Corollary 12.9; then $\ell = O(\log S)$. In the multiplicative case, choose s to dominate the collapsed circuit size and take $\delta_G = \min\{\xi/256, 1/s\}$ in Lemma 12.1; then $\ell = O(\log S + \log(1/\xi))$. Define $\mathsf{Ext}_{\mathrm{pre}}$ to be the first ℓ output bits of P_0, requiring $\ell \le m_P$.

(v) *Remaining entropy:* $L - B_*$. After ℓ is fixed, compute the conditioning loss B_* from (26) in the additive case or from Lemma 12.13 in the multiplicative case. Check that the chosen $\mathsf{2NMExt}$ works at min-entropy $L - B_*$ with the error required above. The sufficient length bound for constant D and $q = 1$ is in Lemma 12.10; for growing D, q it is (31); for the multiplicative guarantee it is (33). Their proofs verify these constraints and $\ell \le m_P$ for the choices just described.

12.5 Reduction Template

Fix a tuple $(\mathbf{V}_i)_{i \in I}$ of indexed rows and a good index $g \in I$ as in Section 12.3. Write $J = I \setminus \{g\}$, $\mathbf{P}_i = \mathsf{Ext}_{\mathrm{pre}}(\mathbf{V}_i)$, $\mathbf{P} = (\mathbf{P}_i)_{i \in I}$, and $p_0 = \Pr[\mathsf{Post}(\mathbf{U}_{r_0}) = 1]$. We write $\mathsf{CB}(\mathbf{V}_J) = (\mathsf{CB}(\mathbf{V}_i))_{i \in J}$.

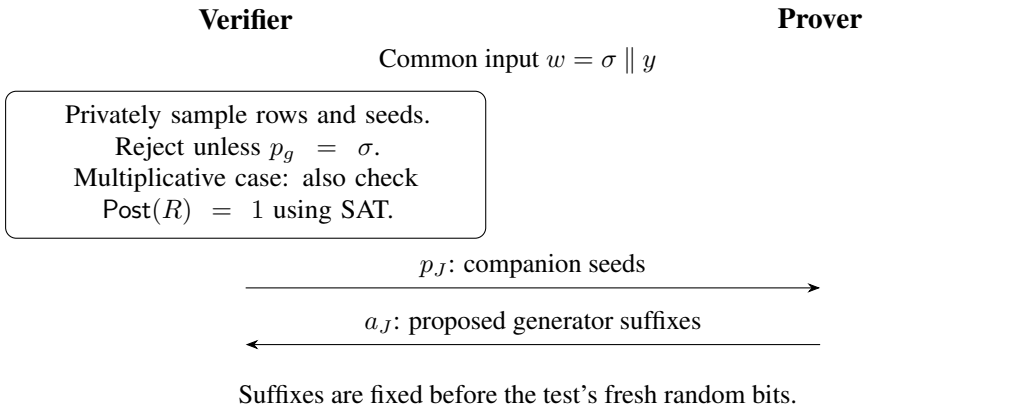
Suppose that this tuple and good index violate the desired decorrelation guarantee. In the additive case, this means

$$\mathsf{SD}\big((\mathsf{CB}(\mathbf{V}_g), \mathsf{CB}(\mathbf{V}_J)), (\mathbf{U}_q, \mathsf{CB}(\mathbf{V}_J))\big) > \theta.$$

In the multiplicative case, additionally fix an output tuple $z = (z_i)_{i \in I}$ at which one of the pointwise comparisons in (34) fails.

On common input $w = \sigma \,\|\, y \in \{0,1\}^L$, the verifier privately samples $R \leftarrow \mathbf{U}_{r_0}$, computes $(v_i)_{i \in I} = \mathsf{Samp}(R)$, and sets $p_i = \mathsf{Ext}_{\mathrm{pre}}(v_i)$ for each $i \in I$. It rejects unless $p_g = \sigma$. In the multiplicative case, it also checks $\mathsf{Post}(R) = 1$ using its SAT oracle before continuing. It sends the companion seeds p_J to the prover,

which returns suffixes $a = (a_i)_{i \in J}$. These define proposed first inputs $p_i \parallel a_i$ to $\mathsf{2NMExt}$. The verifier then runs the test specified below, using fresh randomness. In the additive case, the rows and sampling tape remain private until the prover has sent its hidden-bit guess; only then does the verifier reveal R and request a witness a_0, accepting only if $Q(R, a_0) = 1$. In the multiplicative case, the rows and sampling tape remain private throughout. Figure 2 shows the two possible continuations.



Additive test. Choose a hidden bit c. Send the companion outputs and either $\mathsf{2NMExt}(w, v_g)$ or an independent uniform string. After the prover guesses c, reveal R and request a_0. Accept iff the guess is correct and $Q(R, a_0) = 1$.

Multiplicative test. With fresh private coins and a SAT oracle, compare the counts defining the real and ideal output probabilities. Accept if the chosen comparison test accepts.

Figure 2: The common interaction and its two tests. The verifier never computes G. On a generator input, the honest prover supplies $a_i = G'(p_i)$; soundness allows arbitrary suffix replies.

On generator inputs, a successful prover can supply the actual suffixes $G'(p_i)$, so failure of decorrelation becomes detectable. For uniform inputs, fixing the seed tuple makes the uniform suffix independent of the conditioned rows. Non-malleability then limits the success of every prover. The next observation isolates the common final step, after these two acceptance estimates have been proved.

Lemma 12.7 (From an interactive reduction to a contradiction)**.** *Let G satisfy $\Pr[D(G(\mathbf{U}_\ell)) = 1] \le e^{\varepsilon_G} \Pr[D(\mathbf{U}_L) = 1] + \delta_G$ for nondeterministic $\mathcal{O}$-oracle circuits of size $p_r(T)$, where p_r is from Lemma 12.3. There cannot be an r-round private-coin protocol obeying that lemma's resource and threshold bounds such that, writing $v(w) = \mathrm{val}_{\mathcal{P}}(w)$,*

$$\Pr_{\sigma \leftarrow \mathbf{U}_\ell}[v(G(\sigma)) \ge c_Y] \ge \gamma_Y, \qquad \Pr_{w \leftarrow \mathbf{U}_L}[v(w) > c_N] \le \gamma_N,$$

with $c_Y \ge B^r c_N$ and $\gamma_Y > e^{\varepsilon_G}\gamma_N + \delta_G$.

Proof. Apply Lemma 12.3 to the promise sets $\{w : v(w) \ge c_Y\}$ and $\{w : v(w) \le c_N\}$. Its circuit accepts at least a γ_Y fraction of generator outputs and at most a γ_N fraction of uniform inputs, contradicting the stated pseudorandomness inequality. □

To apply Lemma 12.7, we establish the required acceptance bounds in Lemmas 12.8 and 12.13. Both reductions apply to any seed-extending map G, without assuming pseudorandomness or requiring the verifier to evaluate G.

12.6 Additive decorrelation

The additive test asks the prover to distinguish the good row's output from an independent uniform string, given the companion outputs. Its advantage is exactly half the statistical distance. We first establish an interactive reduction for arbitrary G; only the ensuing corollary assumes pseudorandomness.

Lemma 12.8 (Interactive reduction for additive error)**.** *For fixed integers $a, b \geq 1$, let $2 \leq L, D \leq S$, $1 \leq q \leq S$, $1 \leq \ell \leq L$, and $\Delta \geq 0$. Put $\theta = S^{-b}$, $\eta = \theta^3/2^{20}$, and*

$$B_* = \max\{\ell, \Delta + D\ell + \log(1/\eta)\}. \tag{26}$$

Suppose that $\mathsf{Ext}_{\mathrm{pre}} : \{0,1\}^L \to \{0,1\}^\ell$ and $\mathsf{2NMExt} : \{0,1\}^L \times \{0,1\}^L \to \{0,1\}^q$ have circuit size at most S^a, that $\mathsf{SD}(\mathsf{Ext}_{\mathrm{pre}}(\mathbf{X}), \mathbf{U}_\ell) \leq \eta$ for every $\mathbf{X} \in \mathcal{C}^{\mathrm{NP}}_{L,S}$ with $\mathbf{H}_\infty(\mathbf{X}) \geq L - \Delta$, and that $\mathsf{2NMExt}$ is $(D-1)$-non-malleable for min-entropy $L - B_*$ and error η, strong in its first input. Let $G : \{0,1\}^\ell \to \{0,1\}^L$ be any seed-extending map.

If some tuple of at most D indexed rows and good index g as above satisfies

$$\mathsf{SD}\big((\mathsf{CB}(\mathbf{V}_g), \mathsf{CB}(\mathbf{V}_J)), (\mathbf{U}_q, \mathsf{CB}(\mathbf{V}_J))\big) > \theta,$$

then there is a constant-round private-coin protocol with an ordinary verifier, resource bound $T = \mathrm{poly}_{a,b}(S)$, and thresholds t_Y, t_N of polynomial description length satisfying $t_Y \geq B^r t_N$, where r is its number of rounds, such that

$$\Pr_{\sigma \leftarrow \mathbf{U}_\ell}[\mathrm{val}(G(\sigma)) \geq t_Y] \geq 3\theta/4, \qquad \Pr_{w \leftarrow \mathbf{U}_L}[\mathrm{val}(w) > t_N] \leq \theta/8.$$

Proof. Fix a violating tuple $(\mathbf{V}_i)_{i \in I}$ and its good index g. Use the notation of Section 12.5, writing $p = p_0$.

For the analysis, consider the ideal protocol that also checks $\mathsf{Post}(R) = 1$ before sending the companion seeds in Section 12.5; this check need not be efficient. After its initial sampling check and the prover's suffix replies a_J, run the following additive test on input $w = \sigma \parallel y$:

(i) Choose an independent uniform bit c. If $c = 0$, choose $u \leftarrow \mathbf{U}_q$ independently; if $c = 1$, set $u = \mathsf{2NMExt}(w, v_g)$. Send $\big(u, (\mathsf{2NMExt}(p_i \parallel a_i, v_i))_{i \in J}\big)$ to the prover.

(ii) Accept if the prover correctly guesses c.

The ordinary verifier omits the initial Post check. After receiving the guess $\widehat{c}$, it reveals R, requests a_0, and accepts iff $\widehat{c} = c$ and $Q(R, a_0) = 1$. For every fixed strategy up to the guess, optimizing the final witness reply gives acceptance probability

$$\Pr[\mathsf{Post}(R) = 1,\ p_g = \sigma,\ \widehat{c} = c].$$

The two protocols allow the same suffix replies and guesses on valid tapes, and invalid tapes cannot accept. Hence their optimal values agree on every input. We analyze the ideal protocol below.

Completeness: generator inputs. For σ with $\Pr[\mathbf{P}_g = \sigma] > 0$, let d_σ be the statistical distance, conditioned on $\mathbf{P}_g = \sigma$, between

$$\big(\mathbf{P}_J, \mathsf{2NMExt}(G(\sigma), \mathbf{V}_g), (\mathsf{2NMExt}(G(\mathbf{P}_i), \mathbf{V}_i))_{i \in J}\big)$$

and the same tuple with its second entry replaced by an independent $\mathbf{U}_q$. Set $d_\sigma = 1$ when $\Pr[\mathbf{P}_g = \sigma] = 0$. Revealing $\mathbf{P}$ can only increase the distance assumed to exceed θ, so $\mathbb{E}_{\sigma \leftarrow \mathbf{P}_g} d_\sigma > \theta$. Since $\mathsf{SD}(\mathbf{P}_g, \mathbf{U}_\ell) \leq \eta$, this gives $\mathbb{E}_{\sigma \leftarrow \mathbf{U}_\ell} d_\sigma > \theta - \eta$.

Put $\beta = \theta/8$ and $\alpha = \theta/128$. Call σ bad if $|\Pr[\mathbf{P}_g = \sigma] - 2^{-\ell}| > \alpha 2^{-\ell}$, and good otherwise. By Lemma 4.5, the uniform fraction of bad seeds is at most

$$b_0 = 2\eta/\alpha = \theta^2/4096. \tag{27}$$

Using $0 \le d_\sigma \le 1$, at least a $3\theta/4$ fraction of uniform seeds are good and satisfy $d_\sigma > \beta$.

For each such seed, the input $G(\sigma)$ passes the first check with probability $p \Pr[\mathbf{P}_g = \sigma] \ge p2^{-\ell}(1-\alpha)$. The prover can reply with $a_i = G'(p_i)$ and then distinguish the two challenge distributions with success probability $(1+d_\sigma)/2$. Thus the protocol's maximum acceptance probability on $G(\sigma)$ is at least

$$c_Y = p\, 2^{-\ell-1}(1-\alpha)(1+\beta). \tag{28}$$

Soundness: uniform inputs. We condition on the extracted seeds while controlling the min-entropy of the good row. Discard values $\mathbf{p}$ of $\mathbf{P}$ with $\Pr[\mathbf{P} = \mathbf{p}] < \eta 2^{-D\ell}$. There are at most $2^{D\ell}$ such values, so their total probability is at most η. For every remaining value, Lemma 4.4 gives

$$\mathbf{H}_\infty(\mathbf{V}_g \mid \mathbf{P} = \mathbf{p}) \ge L - \Delta - D\ell - \log(1/\eta). \tag{29}$$

Fix a deterministic prover and a remaining value $\mathbf{p} = (p_i)_{i \in I}$. Its first replies are functions $a_i(w, p_J)$, so put $f_i(w) = p_i \parallel a_i(w, p_J)$. The source $\mathbf{Z} = p_g \parallel \mathbf{U}_{L-\ell}$ is independent of the rows conditioned on $\mathbf{P} = \mathbf{p}$, and has min-entropy $L - \ell$. By Equations (26) and (29), both $\mathbf{Z}$ and the conditioned good row have min-entropy at least $L - B_*$. Applying Lemma 12.4, even given $\mathbf{Z}$, the two challenge distributions have statistical distance at most η. Consequently, for this fixed $\mathbf{p}$, the prover's average probability of guessing c, over the uniform suffix of $\mathbf{Z}$, is at most $(1+\eta)/2$.

For each input w, let $d^*(w)$ be twice the maximum probability of correctly guessing c, conditioned on passing the first check, minus one. Set $d^*(w) = 1$ when that check has probability zero. A single deterministic prover can attain these maxima for every input and first message, since the prover has no computational restriction. Apply the preceding non-malleability bound to this prover.

When averaging over $w \leftarrow \mathbf{U}_L$, bad prefixes contribute at most b_0 to $\mathbb{E}_w\, d^*(w)$. For a good prefix σ, conditioning on passing the first check gives the row distribution $(\mathbf{V}_i)_{i \in I} \mid \mathbf{P}_g = \sigma$. Averaging over uniform σ multiplies the original probability of each such row tuple by $1/(2^\ell \Pr[\mathbf{P}_g = \sigma]) \le (1-\alpha)^{-1}$. Thus discarded values of $\mathbf{P}$ contribute at most $\eta/(1-\alpha)$. For every remaining $\mathbf{p}$, the average advantage over the uniform suffix is at most η. Hence

$$\mathop{\mathbb{E}}_{w \leftarrow \mathbf{U}_L} d^*(w) \le \eta + b_0 + \frac{\eta}{1-\alpha} \le \frac{\theta^2}{2048}.$$

Markov's inequality gives $d^*(w) \le \alpha$ except on uniform probability at most $\theta/16$. If the prefix σ is also good, the first check passes with probability at most $p2^{-\ell}(1+\alpha)$. Combining this with Equation (27), on at least a $1 - \theta/8$ fraction of uniform inputs every prover is accepted with probability at most

$$c_N = p\, 2^{-\ell-1}(1+\alpha)^2. \tag{30}$$

Amplifying the acceptance ratio. The acceptance bounds satisfy

$$\frac{c_Y}{c_N} = \frac{(1-\alpha)(1+\beta)}{(1+\alpha)^2} \ge 1 + \beta/2.$$

Repeat the ordinary protocol independently $N = O(1/\theta)$ times in parallel, receiving all hidden-bit guesses before revealing any sampling tapes and requesting the postselection witnesses. Accept if every copy accepts. By the same argument, its value equals that of N parallel copies of the ideal protocol, which is exactly the

Nth power of the one-copy value. For fixed first verifier messages and prover replies, the hidden rows and challenge experiments remain independent across copies. For each vector of second messages, the most likely challenge-bit vector is obtained by choosing the most likely bit in each copy. Thus the maximum probability of guessing all challenge bits is the product of the individual maxima. Optimizing the first replies separately and averaging over the independent first messages preserves this product.

Choose N so that $(c_Y/c_N)^N \geq B^{r_{\mathrm{IP}}}$, where r_{IP} is the constant number of rounds including the final witness round and B is the constant in Lemma 12.3. The repeated protocol still has r_{IP} rounds, and its verifier, messages, and random inputs have size $\mathrm{poly}(S)$. Since the sampling circuit has at most S input bits and $p > 0$, we have $\log(1/p) \leq S$. Thus c_Y^N and c_N^N have polynomial-length descriptions.

Set $t_Y = c_Y^N$ and $t_N = c_N^N$. The repeated protocol has value equal to the Nth power of the original value, so the two input-density bounds are unchanged. Its verifier uses only Samp and Q, $\mathsf{Ext}_{\mathrm{pre}}$, and $\mathsf{2NMExt}$, and its threshold descriptions have polynomial length. □

Corollary 12.9 (Additive transfer)**.** *For every fixed $a, b \geq 1$ there is a constant A such that, under the extraction and parameter hypotheses of Lemma 12.8, if G satisfies*

$$\Pr[D(G(\mathbf{U}_\ell)) = 1] \leq \Pr[D(\mathbf{U}_L) = 1] + \theta/4$$

for every nondeterministic circuit of size at most S^A, then every admissible tuple and good index satisfy

$$\mathsf{SD}\big((\mathsf{CB}(\mathbf{V}_g), \mathsf{CB}(\mathbf{V}_J)), (\mathbf{U}_q, \mathsf{CB}(\mathbf{V}_J))\big) \leq \theta.$$

Proof. Choose A to bound the collapsed circuit size for the protocol in Lemma 12.8. If the conclusion failed, that lemma and Lemma 12.7 would give $3\theta/4 \leq \theta/8 + \theta/4$, a contradiction. □

12.6.1 A constant number of rows and one output bit

We now choose the ingredients to obtain the logarithmic-length correlation breaker used in the one-bit extractor.

Lemma 12.10 (Local correlation breaker)**.** *Assume H_{NP}. For every fixed integer $D \geq 1$ and constant $b > 0$, there is a constant C such that, for all parameters satisfying $\Delta \geq 0$, $h \leq L$, and $C(\Delta + \log S) \leq L \leq S$, there is a $\mathrm{poly}(S)$-time computable function $\mathsf{CB} : \{0,1\}^L \to \{0,1\}$ with the following property.*

Let $(\mathbf{Z}_1, \ldots, \mathbf{Z}_d)$ be a tuple of at most D indexed rows with h-bit indices, jointly samplable with nondeterministic postselection by a size-S circuit. For every $g \in [d]$ with $\mathbf{H}_\infty(\mathbf{Z}_g) \geq L - \Delta$, we have

$$\mathsf{SD}\Big(\big(\mathsf{CB}(\mathbf{Z}_g), (\mathsf{CB}(\mathbf{Z}_i))_{i \neq g}\big), \big(\mathbf{U}_1, (\mathsf{CB}(\mathbf{Z}_i))_{i \neq g}\big)\Big) \leq S^{-b}.$$

Proof of Lemma 12.10. Increase b to an integer. For $D = 1$, take one bit from Corollary 4.31, increasing the length constant if necessary. Suppose $D \geq 2$, and put $\theta = S^{-b}$ and $\eta = \theta^3/2^{20}$.

Choose P from Corollary 4.31 with error at most η and $\lfloor c_0 L \rfloor$ output bits, and choose the $(D-1)$-non-malleable extractor $\mathsf{2NMExt}$ from Theorem 12.5. Their circuit sizes are bounded by S^a for a fixed integer a. With the exponent A from Corollary 12.9, choose the generator G from Corollary 12.2 with error at most $\theta/4$. Its seed length is $\ell = O(\log S)$. Define $\mathsf{Ext}_{\mathrm{pre}}$ to be the first ℓ bits of P.

Since D is fixed, B_* in Equation (26) is $O(\Delta + \log S)$. Taking C sufficiently large in $L \geq C(\Delta + \log S)$ ensures $\ell \leq \lfloor c_0 L \rfloor$, $B_* \leq \frac{cL}{D^3}$, and $2^{-cL/D^3} \leq \eta$, where c is the constant in Theorem 12.5. Thus Corollary 12.9 applies with $q = 1$ and proves the claimed local guarantee. □

12.6.2 More rows and output bits

The same interactive reduction applies when the number of rows and the output length grow. Only the parameters of 2NMExt and the entropy budget change.

Lemma 12.11 (Multi-bit local correlation breaker)**.** *Assume* H_{NP}. *There is a constant* $0 < \rho < 1$ *such that, for every constant* $b > 0$*, there is a constant* C *with the following property. Let* $D, q \geq 1$*,* $h \leq L \leq S$*, and* $\Delta \geq 0$ *satisfy*

$$\Delta + Dq + C(D+1)\log S \leq L^{\rho}. \tag{31}$$

There is a function $\mathsf{CB} : \{0,1\}^L \to \{0,1\}^q$*, computable in time* $\mathrm{poly}(S)$*, with the following guarantee.*

Let $(\mathbf{Z}_1, \ldots, \mathbf{Z}_d)$ *be a tuple of at most* D *indexed rows with* h*-bit indices, jointly samplable with nondeterministic postselection by a size-*S *circuit. For every* $g \in [d]$ *with* $\mathbf{H}_{\infty}(\mathbf{Z}_g) \geq L - \Delta$*, we have*

$$\mathsf{SD}\Big(\big(\mathsf{CB}(\mathbf{Z}_g), (\mathsf{CB}(\mathbf{Z}_i))_{i \neq g}\big), \big(\mathbf{U}_q, (\mathsf{CB}(\mathbf{Z}_i))_{i \neq g}\big)\Big) \leq S^{-b}.$$

Proof of Lemma 12.11. Fix $0 < \rho < \min\{1, \rho_t, \rho_o, \rho_k, \rho_e\}$, where the four exponents are from Lemma 12.6. Increase b to an integer if necessary, and put $\theta = S^{-b}$ and $\eta = \theta^3/2^{20}$. Use Corollary 4.31 with error exponent $3b + 20$ to obtain $P : \{0,1\}^L \to \{0,1\}^{\lfloor c_0 L \rfloor}$ with error at most $S^{-(3b+20)} \leq \eta$. Its length requirement follows from Equation (31) after increasing C: the budget gives $\Delta + \log S \leq L^{\rho}$, and $L^{\rho} = o(L)$. Increasing C also makes L sufficiently large for all the bounds below. If $D = 1$, the first q output bits of P prove the claim, since $q \leq L^{\rho} \leq \lfloor c_0 L \rfloor$.

Suppose $D \geq 2$. By Lemma 12.6, there is an extractor $\mathsf{2NMExt} : \{0,1\}^L \times \{0,1\}^L \to \{0,1\}^q$ that is $(D-1)$-non-malleable for min-entropy $L - L^{\rho_k}$ and error $2^{-L^{\rho_e}}$, strong in either input. Fix an integer a_0 bounding the circuit sizes of P and 2NMExt by S^{a_0}. The constant in Corollary 12.9 then determines the circuit size to be fooled by the generator. Choose G from Corollary 12.2 for that size and error at most $\theta/4$. Its seed length is $\ell = O(\log S)$, and it has the form $G(\sigma) = \sigma \parallel G'(\sigma)$. Increasing C ensures $\ell \leq \lfloor c_0 L \rfloor$; let $\mathsf{Ext}_{\mathrm{pre}}$ be this prefix of P.

Choose C large enough that $\ell \leq C \log S$ and $\log(1/\eta) \leq C \log S$. The budget then gives

$$B_* := \max\{\ell, \Delta + D\ell + \log(1/\eta)\} \leq L^{\rho} \leq L^{\rho_k}, \qquad \log(1/\eta) \leq L^{\rho} \leq L^{\rho_e}.$$

Thus 2NMExt works at min-entropy $L - B_*$ with error at most η. All hypotheses of Corollary 12.9 hold, so $\mathsf{CB}(v) = \mathsf{2NMExt}(G(\mathsf{Ext}_{\mathrm{pre}}(v)), v)$ has the required guarantee. □

12.7 Multiplicative decorrelation

For distributions A, B on the same finite space, we write $A \approx^{\mathrm{pt}}_{\vartheta,\xi} B$ if, for every outcome z,

$$\Pr[A = z] \leq e^{\vartheta} \Pr[B = z] + \xi, \qquad \Pr[B = z] \leq e^{\vartheta} \Pr[A = z] + \xi.$$

This expresses relative error with an additive slack for each individual outcome. It resembles the comparison used in differential privacy, but here summing over an event of size m gives slack $m\xi$. We use only the pointwise guarantee.

We use the same construction (25), now with a preliminary extractor satisfying a pointwise seed bound and a PRG against nondeterministic NP-oracle circuits. To test a failure at a fixed output tuple, the verifier compares its real and ideal probabilities using approximate counting. The following theorem states the guarantee locally; its proof follows the interactive lemma below.

Theorem 12.12. *Assume H_{Σ_2}. There is a constant $\rho_m > 0$ such that, for every constant $b > 0$, there is a constant C with the following property. Let $D_*, q \geq 1$, $h \leq L \leq S$, and $\Delta \geq 0$. Suppose*

$$0 < \xi \leq S^{-b} \leq \vartheta \leq 1/10 \tag{32}$$

and

$$\Delta + D_* q + C\big(D_* \log S + (D_*+1)\log(1/\xi)\big) \leq L^{\rho_m}. \tag{33}$$

There is a $\mathrm{poly}(S)$-time computable function $\mathsf{CB}_{\mathrm{mult}} : \{0,1\}^L \to \{0,1\}^q$ with the following guarantee.

Let $(\mathbf{V}_i)_{i\in I}$ be a tuple of at most D_* indexed rows with h-bit indices, jointly samplable with nondeterministic postselection by a size-S sampling circuit. For every $g \in I$ with $\mathbf{H}_\infty(\mathbf{V}_g) \geq L - \Delta$ and every output tuple $z = (z_i)_{i\in I}$, the probabilities

$$P_z = \Pr[\mathsf{CB}_{\mathrm{mult}}(\mathbf{V}_i) = z_i \text{ for every } i \in I],$$
$$Q_z = 2^{-q}\Pr[\mathsf{CB}_{\mathrm{mult}}(\mathbf{V}_i) = z_i \text{ for every } i \neq g]$$

satisfy

$$P_z \leq e^{\vartheta} Q_z + \xi, \qquad Q_z \leq e^{\vartheta} P_z + \xi. \tag{34}$$

Equivalently, the two output distributions in the theorem are $\approx^{\mathrm{pt}}_{\vartheta,\xi}$-close.

12.7.1 The interactive lemma

Lemma 12.13 (Interactive reduction for multiplicative error)**.** *Let $2 \leq L \leq S$, $1 \leq D_*, q \leq S$, $1 \leq \ell < L$, $\Delta \geq 0$, and $0 < \xi \leq \vartheta \leq 1/10$, where ξ and ϑ are powers of two. Set $\zeta = \varepsilon_c = \xi^8$, $\Lambda = \lceil 8\log(1/\xi)\rceil + 2$, and*

$$B_* = \max\{\ell + \Lambda, \Delta + D_*\ell + \log(1/\zeta)\}.$$

Assume $\ell + \Lambda \leq L$. Suppose $\mathsf{Ext}_{\mathrm{pre}} : \{0,1\}^L \to \{0,1\}^\ell$ and $\mathsf{2NMExt} : \{0,1\}^L \times \{0,1\}^L \to \{0,1\}^q$ have size at most S^a for a fixed a, and $\mathsf{2NMExt}$ is (D_*-1)-non-malleable for min-entropy $L - B_*$ and error at most $\xi/100$, strong in its first input. Let G be any seed-extending map and define CB by (25).

Fix an admissible tuple of at most D_* rows and a good index g such that

$$\Pr[\mathsf{Ext}_{\mathrm{pre}}(\mathbf{V}_g) = \sigma] \leq e^{1/100}2^{-\ell} \quad \text{for every } \sigma. \tag{35}$$

If either comparison in (34) fails at some output tuple z, there is a constant-round private-coin protocol with a SAT-oracle verifier and resource bound

$$T = \mathrm{poly}_a(S, D_*, q, 1/\vartheta, \log(1/\xi))$$

whose optimal acceptance probability v satisfies

$$\Pr_{\sigma\leftarrow\mathbf{U}_\ell}[v(G(\sigma)) \geq c_Y] \geq \xi/32, \qquad \Pr_{w\leftarrow\mathbf{U}_L}[v(w) > c_N] \leq 768K\xi^6.$$

Here K is a fixed integer at least B^r for the constant number r of rounds, $c_Y = p_0 2^{-\ell}\xi^2/256$, and $c_N = c_Y/K$.

Proof. Fix the violating sampler, good index, output tuple, and direction. Use the common outer interaction of Section 12.5. After receiving the suffix tuple a, the verifier runs the count-comparison test described below with fresh private random bits. We first define the probabilities it compares, then analyze its acceptance on generator and uniform inputs.

The real and ideal events. Write $d = |I| \le D_*$ and $\mathbf{P} = (\mathsf{Ext}_{\mathrm{pre}}(\mathbf{V}_i))_{i \in I}$. For a seed tuple p of positive probability, a suffix y, and companion suffixes a, put

$$A_p(y,a) = \Pr[\mathsf{2NMExt}(p_g \parallel y, \mathbf{V}_g) = z_g, \mathsf{2NMExt}(p_i \parallel a_i, \mathbf{V}_i) = z_i \ (i \neq g) \mid \mathbf{P} = p],$$
$$B_p(a) = 2^{-q} \Pr[\mathsf{2NMExt}(p_i \parallel a_i, \mathbf{V}_i) = z_i \ (i \neq g) \mid \mathbf{P} = p].$$

Let N_p count the original sampler inputs with $\mathsf{Post} = 1$ and $\mathbf{P} = p$. Let N_A count those additionally satisfying the full output event in $A_p(y,a)$, and let N_B count those satisfying the companion output event. Then $A_p(y,a) = N_A/N_p$ and $B_p(a) = 2^{-q} N_B/N_p$. Each count has a nondeterministic membership circuit that guesses a_0 and evaluates only Q, Samp, $\mathsf{Ext}_{\mathrm{pre}}$, and $\mathsf{2NMExt}$.

The comparison test. Stockmeyer's approximate counting algorithm [Sto85] uses a SAT oracle to estimate the number of accepting inputs of a Boolean circuit within $e^{\pm\tau}$, with failure probability β, in time polynomial in the circuit size, $1/\tau$, and $\log(1/\beta)$. The same algorithm counts the distinct sampling tapes accepted by our nondeterministic membership circuits: its hash constraints apply only to R, and each existence query guesses both R and a witness a_0. Thus these queries remain in NP; postselection witnesses are not counted. Zero counts are recognized exactly by SAT queries. Reject when $N_p = 0$. Otherwise estimate each of N_A, N_B, N_p with $\tau = \vartheta/100$ and failure probability at most $\varepsilon_c/3$. Choose t with $|\ln t - 5\vartheta/8| \le \vartheta/100$, and accept the forward test exactly when

$$\widehat{N}_A > t\, 2^{-q} \widehat{N}_B + (5\xi/8)\widehat{N}_p.$$

On the event that all three estimates are accurate, this test accepts if

$$A_p(y,a) > e^{3\vartheta/4} B_p(a) + 3\xi/4, \tag{36}$$

and rejects if

$$A_p(y,a) \le e^{\vartheta/2} B_p(a) + \xi/2. \tag{37}$$

Indeed, after allowing two estimation errors and threshold rounding, the multiplicative margins are at least $(1/8 - 3/100)\vartheta > 0$. The additive margins follow from $e^{2\vartheta/100} < 6/5$ and $e^{2\vartheta/100} < 5/4$. Thus each promised behavior holds with probability at least $1 - \varepsilon_c$. Interchange N_A and $2^{-q} N_B$ for the reverse test.

Completeness: generator inputs. Suppose first that $P_z > e^{\vartheta} Q_z + \xi$. On input $G(\sigma)$, the prover replies $a_i = G'(p_i)$. With these honest suffixes define

$$F(p) = A_p(G'(p_g), G'(p_J)) - e^{3\vartheta/4} B_p(G'(p_J)) - 3\xi/4.$$

Averaging over $\mathbf{P}$ gives $\mathbb{E}\, F(\mathbf{P}) > \xi/4$. Since $F \le 1$, the event $\mathcal{E} = \{p : F(p) > 0\}$ has probability greater than $\xi/4$. Let $h_\sigma = \Pr[\mathcal{E} \mid \mathbf{P}_g = \sigma]$ on seeds of positive mass. The seeds with $h_\sigma \ge \xi/8$ have total $\mathbf{P}_g$-mass greater than $\xi/8$: outside this set the contribution to $\mathbb{E}\, h_{\mathbf{P}_g}$ is at most $\xi/8$. Remove seeds with $\Pr[\mathbf{P}_g = \sigma] < (\xi/16)2^{-\ell}$; their total mass is at most $\xi/16$. The remaining set has mass greater than $\xi/16$, hence uniform seed density at least $\xi/32$ by (35).

For every remaining seed the initial sampling check succeeds with probability at least $p_0(\xi/16)2^{-\ell}$. Conditioned on that check, (36) holds with probability at least $\xi/8$, and the test then accepts with probability at least $1 - \varepsilon_c \ge 1/2$. The optimal acceptance probability is therefore at least $c_Y = p_0 2^{-\ell} \xi^2/256$. The reverse violation $Q_z > e^{\vartheta} P_z + \xi$ is identical with A_p and B_p interchanged.

Soundness: uniform inputs. Discard seed tuples with probability below $\zeta 2^{-d\ell}$; their total probability is at most ζ. For every remaining p, Lemma 4.4 gives

$$\mathbf{H}_\infty(\mathbf{V}_g \mid \mathbf{P} = p) \ge L - \Delta - d\ell - \log(1/\zeta).$$

The following consequence of non-malleability bounds the suffixes on which some prover response can cause the test to accept outside its error event.

Claim 12.14. *For every remaining p, each of the sets*

$$\mathcal{L}_p^+ = \{y : \exists a,\ A_p(y,a) > e^{\vartheta/2} B_p(a) + \xi/2\},$$
$$\mathcal{L}_p^- = \{y : \exists a,\ B_p(a) > e^{\vartheta/2} A_p(y,a) + \xi/2\}$$

has size less than $2^{L-\ell-\Lambda}$.

Proof. If either set were that large, choose $\mathbf{Y}$ uniformly on a subset of exactly that cardinality and a witnessing tuple $a(y)$ for each y. This cardinality is an integer because $\Lambda \le L - \ell$. The independent sources $p_g \parallel \mathbf{Y}$ and $\mathbf{V}_g \mid \mathbf{P} = p$ have deficiency at most B_*. Apply Lemma 12.4 with tampered first inputs $p_i \parallel a_i(\mathbf{Y})$, extended arbitrarily outside the support. Testing the output tuple z yields

$$\left| \mathop{\mathbb{E}}_{\mathbf{Y}} [A_p(\mathbf{Y}, a(\mathbf{Y})) - B_p(a(\mathbf{Y}))] \right| \le \xi/100.$$

For every y in the support of $\mathbf{Y}$, the difference $A_p(y,a(y)) - B_p(a(y))$ is greater than $\xi/2$ in the $\mathcal{L}_p^+$ case and less than $-\xi/2$ in the $\mathcal{L}_p^-$ case. Thus the absolute value of its expectation is greater than $\xi/2$, contradicting the above bound. □

Fix an optimal deterministic prover response for every input and visible seed message. Average over $w \leftarrow \mathbf{U}_L$ and the sampler tape. Passing postselection contributes p_0, and matching the independent uniform prefix contributes $2^{-\ell}$. For each retained seed tuple the fraction of suffixes in the relevant violating set is at most $2^{-\Lambda}$. Outside it, (37) holds for every suffix response, so the test accepts with probability at most ε_c. Its coins are drawn after the response is fixed. Including discarded tuples gives

$$\mathop{\mathbb{E}}_{w \leftarrow \mathbf{U}_L} v(w) \le p_0 2^{-\ell} (2^{-\Lambda} + \varepsilon_c + \zeta) \le 3 p_0 2^{-\ell} \xi^8.$$

With $c_N = c_Y / K$, Markov's inequality gives $\Pr[v(\mathbf{U}_L) > c_N] \le 768 K \xi^6$. The factor p_0 cancels in the ratio $c_Y / c_N = K$. Finally, p_0 is an integer multiple of 2^{-r_0} with $r_0 \le S$, so the thresholds have polynomial-length descriptions. This proves all the interactive lemma's claims. □

12.7.2 Construction and analysis

Proof of Theorem 12.12. Choose ρ_m strictly smaller than 1 and the four exponents in Lemma 12.6. We first assume that ξ and ϑ are powers of two. For arbitrary values, round both down to the nearest powers of two. This strengthens the required guarantee, increases $\log(1/\xi)$ by at most one, and changes the lower bound on ϑ by at most a factor of two.

We first choose the ingredients.

- Use Lemma 4.30 with $\nu = 1/100$ to obtain P_0 with $m_P = \lfloor c_0 L \rfloor$ output bits, for a sufficiently small fixed $c_0 > 0$. Use Lemma 12.6 to choose $\mathsf{2NMExt}$ for $D_* - 1$ tamperings and q output bits. Fix a_0 so that both circuit sizes are at most S^{a_0}.

- The verifier's resource bound in Lemma 12.13, and the circuit size after applying Lemma 12.3, are bounded by a fixed polynomial in S, $1/\vartheta$, and $\log(1/\xi)$, uniformly over all prefix lengths $\ell \le m_P$. Choose an integer $s = \mathrm{poly}(S, 1/\vartheta, \log(1/\xi))$ dominating this size and S. Set $\delta_G = \min\{\xi/256, 1/s\}$, choosing s large enough that $2^{-s} \le \delta_G$, and apply Lemma 12.1 with $\mathcal{O} = \mathsf{SAT}$ and output length L. The resulting seed length satisfies

  $$\ell = O(\log(1/\delta_G)) = O(\log S + \log(1/\xi)),$$

  since $1/\vartheta = O(S^b)$. Let $\mathsf{Ext}_{\mathrm{pre}}$ be the first ℓ bits of P_0, and define $\mathsf{CB}_{\mathrm{mult}}$ by (25).

We now verify the parameter conditions and prove the claimed guarantee. Set $\zeta = \varepsilon_c = \xi^8$ and $\Lambda = \lceil 8\log(1/\xi)\rceil + 2$ as in Lemma 12.13. By (33), we have

$$B_* := \max\{\ell + \Lambda, \Delta + D_*\ell + \log(1/\zeta)\} = O(L^{\rho_m}), \qquad \log(1/\xi) = O(L^{\rho_m}).$$

Since $\rho_m < \rho_k, \rho_e$, for sufficiently large L we have

$$B_* \le L^{\rho_k}, \qquad 2^{-L^{\rho_e}} \le \xi/100.$$

Thus 2NMExt meets the entropy and error requirements of Lemma 12.13. The same condition ensures $D_* - 1 \le L^{\rho_t}$, $q \le L^{\rho_o}$, $\ell \le m_P$, $\ell + \Lambda \le L$, and the length conditions for P_0 and G. Increasing C excludes the remaining bounded values of L. Taking a prefix of P_0 preserves the pointwise seed cap (35).

Suppose that (34) fails. By Lemma 12.13, there is a protocol satisfying the resource and acceptance-ratio conditions of Lemma 12.7, with

$$\gamma_Y = \xi/32, \qquad \gamma_N = 768K\xi^6.$$

We show that these bounds contradict the guarantee of G.

Since $\xi \le 1/10$, (33) implies $C \le L^{\rho_m} \le S^{\rho_m}$, and therefore

$$\xi \le S^{-b} \le C^{-b/\rho_m}.$$

As K is fixed and $e^{1/s} < 2$, taking C sufficiently large ensures $e^{1/s}768K\xi^6 \le \xi/256$. Together with $\delta_G \le \xi/256$, this gives

$$e^{1/s}\gamma_N + \delta_G \le \xi/128 < \xi/32 = \gamma_Y.$$

Thus Lemma 12.7 rules out this protocol, proving (34).

Since $1/\vartheta = O(S^b)$ and (33) gives $\log(1/\xi) = O(S^{\rho_m})$, all ingredients, and hence $\mathsf{CB}_{\mathrm{mult}}$, are computable in $\mathrm{poly}(S)$ time. □

12.8 Relativization of the generator constructions

We justify the oracle extension in Lemma 12.1. The ordinary nondeterministic versions are the published results of [Sha25b; DS26]; the nondeterministic SAT-oracle version follows by preserving the oracle in their reconstructions. This is an adaptation of those proofs, not a change to the algorithms evaluating the generators.

Lower-bound tests for acceptance probabilities. The following oracle-preserving form of Goldwasser–Sipser is the only extra counting ingredient in the reconstructions.

Lemma 12.15. *For $\mathcal{O} \in \{\varnothing, \mathsf{SAT}\}$, let C be a size-s nondeterministic $\mathcal{O}$-oracle circuit on $r \le s$ ordinary input bits. Given C and thresholds, the promise problem*

$$\Pr[C(\mathbf{U}_r) = 1] \ge \alpha \quad \textit{versus} \quad \Pr[C(\mathbf{U}_r) = 1] \le e^{-\lambda}\alpha$$

has a nonuniform nondeterministic $\mathcal{O}$-oracle circuit of size $\mathrm{poly}(s, B, 1/\lambda)$, for $0 < \alpha, \lambda \le 1$. Here B bounds the binary description lengths of the thresholds; real thresholds may be rounded within a constant fraction of the gap. The dependence on α is only through B.

Proof. For $\mathcal{O} = \varnothing$, apply Theorem 4.39 with size bound $O(s + B)$ and gap parameter $\lambda/2$. The slack accommodates the non-strict inequality in the rejection promise and threshold rounding.

For $\mathcal{O} = \mathsf{SAT}$, the same Goldwasser–Sipser argument applies. Merlin supplies a membership witness for each claimed accepted input of C, and Arthur checks it using the deterministic SAT-oracle verifier of C. The hashing analysis is unchanged. Amplification and fixing public coins proceed as in Theorem 4.39, preserving the oracle and giving the claimed size bound. □

Proof of Lemma 12.1. Write $M = L$ for the output length. For $\mathcal{O} = \varnothing$ use the cited constructions directly. We explain the adaptation for $\mathcal{O} = \mathsf{SAT}$ by checking that each reconstruction still produces a nondeterministic SAT-oracle circuit, or equivalently an Σ_2 circuit. Their evaluation algorithms are unchanged.

Shaltiel's reconstruction. For [Sha25b, Theorem 2.3], the proof from Lemma 2.5 in Section 2.3.1 guesses the coefficients of a polynomial p, checks its values at fixed reference points, and tests a lower bound on the acceptance probability of

$$(t, v) \longmapsto D(C_x(t), v, \mathrm{SExt}(p(t), v)).$$

Here C_x is the curve through the input x used in that proof, and SExt is its strong seeded extractor. These maps have deterministic polynomial-size evaluations, so their composition with D is a Σ_2 circuit. Apply Lemma 12.15 to this circuit and conjoin the result with the reference-point checks and $p(0) = 1$. The guess of p joins the existential block, giving a Σ_2 circuit for the hard language.

Dermer–Shaltiel reconstruction. For [DS26, Theorem 4.3], the XOR and hybrid steps fix inputs of the distinguisher or compose it with a fixed XOR. The test in Figure 2 is

$$(t, v) \longmapsto D_j\big(\mathrm{SExt}(o_1(t), v), \dots, \mathrm{SExt}(o_{m-1}(t), v), \mathrm{SExt}(p(t), v)\big),$$

where the o_j are restriction polynomials fixed by the reconstruction. The test checks reference values and invokes a lower-bound test of the form in Lemma 12.15. Figure 3 requires two such tests to accept. Section 4.2.9 guesses the polynomials at successive offsets and requires all verification steps to accept. If the probability tests have the form $\exists a_j \forall b_j R_j$, the final predicate is

$$\exists\big((p_j)_j, (a_j)_j\big)\, \forall (b_j)_j \quad [\text{all consistency checks hold}] \wedge \bigwedge_j R_j \wedge [p_{\mathrm{final}}(0) = 1].$$

The polynomials are guessed together; their successive verification introduces no further quantifier block. Neither reconstruction complements the distinguisher.

Where a test samples t uniformly from the nonzero elements of $\mathbb{F}_Q$, conjoin $t \neq 0$ with the test and multiply both probability thresholds by $(Q-1)/Q$. This expresses the test using uniform bits and preserves its relative gap. The curve-sampling and list-size arguments determine which reference values and thresholds may be fixed nonuniformly; the reconstructed circuit only uses those fixed values. The thresholds can be rounded within the promised gaps using $\mathrm{poly}(s)$ bits, since $\log(1/\delta) \leq s$ and the relative gaps are inverse-polynomial in s.

The size calculation in [DS26, Section 4.2.9] bounds the reconstruction by $(sd_0)^c$ for a fixed constant c, after absorbing the polynomial cost of our oracle simulation. Here d_0 is a freely chosen constant dimension. Choose $d_0 > (c+1)/\beta_{\mathrm{hard}}$ and set $u = d_0 \log s$. For sufficiently large s, $(sd_0)^c < 2^{\beta_{\mathrm{hard}} u}$, contradicting H_{Σ_2}. The algorithms use the hard language only on $u = O(\log s)$ bits, so they run in $\mathrm{poly}(s)$ time even when $\log(1/\delta) = \Theta(s)$.

We therefore obtain the generator of [DS26, Theorem 4.3] against Σ_2 circuits, with seed length $O(\log(1/\delta))$ and the stated errors. The generator of [Sha25b, Theorem 2.3], against the same class, produces r additional bits from $O(r + \log(1/\delta))$ seed bits while including its seed in the output.

Including the seed. Compose the two generators as follows. At size bound $4s$, obtain $F : \{0,1\}^r \to \{0,1\}^M$ with multiplicative exponent $1/(4s)$ and additive error $\delta/4$. For $T = s^J$ with a sufficiently large fixed $J \geq 2$, obtain $A : \{0,1\}^\ell \to \{0,1\}^r$ such that

$$\Pr[D''(\mathbf{U}_\ell, A(\mathbf{U}_\ell)) = 1] \leq e^{1/T} \Pr[D''(\mathbf{U}_\ell, \mathbf{U}_r) = 1] + \frac{\delta}{4T}$$

for tests of size T. Its seed length is $\ell = O(r + \log(4T/\delta)) = O(\log(1/\delta))$. The hypothesis $C \log(1/\delta) \leq M \leq s$, with $C \geq 1$, implies $\delta \geq 2^{-s}$. Consequently $\delta/4$ is in the permitted range at circuit size $4s$, and

$\delta/(4T)$ is in the permitted range at size $T = s^J$ for sufficiently large s. Increasing C excludes the remaining bounded parameter values and ensures $\ell \le M$. Now put $H(z) = z \parallel F(A(z))_{[M-\ell]}$.

For each fixed z, the test of F remains a nondeterministic $\mathcal{O}$-oracle circuit of size at most $4s$. The circuit $(z, v) \mapsto D(z, F(v)_{[M-\ell]})$ is also a nondeterministic $\mathcal{O}$-oracle circuit and has size at most T by the choice of J. Applying the two guarantees gives

$$\begin{aligned}\Pr[D(H(\mathbf{U}_\ell)) = 1] &\le e^{1/T+1/(4s)} \Pr[D(\mathbf{U}_M) = 1] + e^{1/T}\frac{\delta}{4} + \frac{\delta}{4T} \\ &\le e^{1/s} \Pr[D(\mathbf{U}_M) = 1] + \delta.\end{aligned}$$

□

13 One-bit extraction from logarithmic entropy

The condenser from Section 11 gives polynomially many rows that are jointly samplable with nondeterministic postselection, most of which have almost full min-entropy. We use these rows to construct a one-bit extractor for sources with logarithmic min-entropy. Following [CZ19; BDT17], we transform the rows into bits that are almost t-wise independent outside a small set of bad coordinates, and then apply a resilient function.

Our main result is the following.

Theorem 13.1. *Assume* H_{NP}. *For all constants* $c, \varepsilon > 0$, *there is a constant* C *and a polynomial-time computable* $(C \log n, \varepsilon)$*-extractor* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ *for* $\mathcal{C}^{\mathrm{NP}}_{n,n^c}$.

We collect the tools for the construction in Section 13.1. We then present the extractor construction in Section 13.2 and prove Theorem 13.1 in Section 13.3.

13.1 Tools

We use local correlation breaking and a grouping lemma to obtain a non-oblivious bit-fixing source. We then extract one bit using majority.

Local correlation breaking. We use the correlation breaker that we developed in Lemma 12.10 to make the output on a good row close to a uniform bit independent of the outputs on a constant number of other rows.

Non-oblivious bit-fixing sources. We apply the correlation breaker to indexed rows in fixed groups of constant size and then XOR its outputs within each group to obtain one bit per group. We will show that the resulting bits from any t groups outside a small fixed set are jointly close to $\mathbf{U}_t$. The following definition captures this guarantee.

Definition 13.2 (Non-oblivious bit-fixing source [CZ19; BDT17])**.** *Let* $1 \le t \le M - q$. *A distribution* $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_M) \in \{0,1\}^M$ *is a* (q, t, γ)*-non-oblivious bit-fixing (NOBF) source if there is a fixed set* $\mathcal{Q} \subseteq [M]$ *of at most* q *bad coordinates such that, for every* $T \subseteq [M] \setminus \mathcal{Q}$ *of size* t, *the tuple* $\mathbf{Y}_T = (\mathbf{Y}_u)_{u \in T}$ *satisfies* $\mathsf{SD}(\mathbf{Y}_T, \mathbf{U}_t) \le \gamma$. *The bad bits may depend arbitrarily on the good bits. When* $\gamma = 0$, *we call* $\mathbf{Y}$ *a* (q, t)*-NOBF source.*

Grouping rows. Following [BDT17], we use Zuckerman's disperser [Zuc07] to form groups, all but a small number of which contain a good row. The following lemma is a direct consequence of Zuckerman's disperser [Zuc07], in the form stated in [CGL21, Lemma 4.24].

Lemma 13.3. *For every constant* $0 < \eta \le 1/2$, *there is a constant* d_0 *such that, for every sufficiently large integer* R, *one can uniformly construct in polynomial time a map* $\Gamma : [M] \times [d_0] \to [R]$, *where* $M = \Theta(R^8)$, *with the following property. For every* $B \subseteq [R]$ *with* $|B| \le R^{1-\eta}$, *fewer than* $M^{1/4}$ *groups* $u \in [M]$ *satisfy* $\Gamma(u, a) \in B$ *for every* $a \in [d_0]$.

Extraction by majority. To extract from the resulting NOBF source, we use the following guarantee for majority under bounded independence [Vio14], in the form stated in [CGL21, Lemma 4.28]. Define $\mathsf{Maj}(y) = 1$ if $\sum_{u=1}^{M} y_u \geq \lceil M/2 \rceil$, and $\mathsf{Maj}(y) = 0$ otherwise.

Lemma 13.4. *There is a constant C_{Maj} such that every (q, t, γ)-NOBF source $\mathbf{Y} \in \{0,1\}^M$ with $2 \leq t \leq M - q$ satisfies*

$$\mathsf{SD}(\mathsf{Maj}(\mathbf{Y}), \mathbf{U}_1) \leq C_{\mathrm{Maj}} \left(\frac{\log t}{\sqrt{t}} + \frac{q}{\sqrt{M}} \right) + \gamma M^t.$$

13.2 The extractor construction

We now present the extractor construction, first specifying its ingredients and their parameters.

Fix a constant $c > 0$ and an integer $t \geq 2$. Let C_{hash}, η, and Δ be the constants in Lemma 11.1, decreasing η to at most $1/2$ if necessary. We leave the entropy constant C to be fixed below.

- **Condenser rows.** Apply Lemma 11.1 at threshold $k_* = \lfloor C \log n \rfloor$. Write $h_i(x) = \mathsf{Cond}(x, i)$ for the resulting row maps, indexed by $i \in [R]$ and of length m. Every good row has min-entropy at least $m - \Delta$.
- **Groups.** Choose d_0 and $\Gamma : [M] \times [d_0] \to [R]$ from Lemma 13.3, where $M = \Theta(R^8)$. Group u consists of the rows $h_{\Gamma(u,1)}, \ldots, h_{\Gamma(u,d_0)}$. Put $D = d_0 t$, the number of row positions in any t groups.
- **Indexed rows.** For each appearance of a row, add an index (u, a), where u is the group and a is its position within that group. Set $h = \lceil \log(M d_0) \rceil$ and $L = m + h$, and define

  $$Z_{u,a}(x) = (u, a) \parallel h_{\Gamma(u,a)}(x),$$

  encoding (u, a) in h bits. Distinct positions have distinct prefixes, even when they contain the same hash row. If $h_{\Gamma(u,a)}(\mathbf{X})$ is good, then $\mathbf{H}_\infty(Z_{u,a}(\mathbf{X})) \geq m - \Delta = L - (\Delta + h)$.
- **Sampling circuit size.** Choose a constant $s \geq 2$ and put $S = n^s$ so that any at most D indexed rows are jointly samplable with nondeterministic postselection by a size-S circuit. For a source sampled by $(\mathsf{Samp}, \mathsf{Post})$, this circuit chooses a uniform input ρ, computes the specified hashes of $\mathsf{Samp}(\rho)$, appends their indices, and conditions on $\mathsf{Post}(\rho) = 1$. The polynomial bounds in Lemma 11.1 let us choose s independently of C.
- **Local correlation breaker.** Since M is polynomially bounded in n, choose a constant $b > 0$, also independently of C, such that $tS^{-b} \leq M^{-(t+1)}$, and put $\theta = S^{-b}$. We use Lemma 12.10 with at most D rows, input length L, index length h, entropy deficiency $\Delta + h$, sampling circuit size at most S, and error θ.

  Now fix C sufficiently large that $k_* \geq C_{\mathrm{hash}} \log n$ and the correlation breaker's length requirement holds. We verify these bounds below. Thus Lemma 12.10 supplies the required function $\mathsf{CB} : \{0,1\}^L \to \{0,1\}$.

Construction 13.5. *Using these ingredients, define $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ as follows on input x.*

(i) Compute $h_i(x)$ for every $i \in [R]$ and form the indexed rows $Z_{u,a}(x)$ for every $(u, a) \in [M] \times [d_0]$.

(ii) Apply the correlation breaker and XOR within each group:

$$Y_u(x) = \bigoplus_{a=1}^{d_0} \mathsf{CB}(Z_{u,a}(x)), \qquad F(x) = (Y_1(x), \ldots, Y_M(x)). \tag{38}$$

(iii) Output $\mathsf{Ext}(x) = \mathsf{Maj}(F(x))$.

13.3 Analysis of the construction

We show that the XORs from any t groups that each contain a good row are jointly close to $\mathbf{U}_t$. The grouping lemma then makes $F(\mathbf{X})$ a NOBF source, from which majority extracts one bit.

Lemma 13.6. *Assume H_{NP}. For every constant $c > 0$ and integer $t \geq 2$, there is a constant C such that the map F defined in (38) is polynomial-time computable, has output length $n \leq M = \mathrm{poly}(n)$, and for every $\mathbf{X} \in \mathcal{C}^{\mathrm{NP}}_{n,n^c}$ with $\mathbf{H}_\infty(\mathbf{X}) \geq C \log n$, $F(\mathbf{X})$ is a $(M^{1/4}, t, M^{-(t+1)})$-NOBF source.*

Proof. We first check the length requirement in Lemma 12.10. Since $h = O(\log n)$, Δ is constant, and $S = n^s$, we have $\Delta + h + \log S = O(\log n)$, with a constant independent of C. Also $L \geq k_* - \log k_* \geq (C/2)\log n$ for sufficiently large n. We can therefore choose C so that $k_* \geq C_{\mathrm{hash}} \log n$ and $L \geq C_{\mathrm{cb}}(\Delta + h + \log S)$, where C_{cb} is the constant in Lemma 12.10. We also have $L = O(\log n) \leq S$ and $k_* \leq n$ for sufficiently large n.

Fix a source $\mathbf{X} \in \mathcal{C}^{\mathrm{NP}}_{n,n^c}$ with $\mathbf{H}_\infty(\mathbf{X}) \geq C \log n$. Apply the condenser guarantee from Lemma 11.1 at threshold k_*, and let $\mathcal{G} = [R] \setminus B_{\mathbf{X}}$ be its set of good indices. For every $i \in \mathcal{G}$, we have $\mathbf{H}_\infty(h_i(\mathbf{X})) \geq m - \Delta$. Write $\mathbf{Z}_{u,a} = Z_{u,a}(\mathbf{X})$ and $\mathbf{Y}_u = Y_u(\mathbf{X})$ for the indexed rows and group outputs. Let $\mathcal{Q} \subseteq [M]$ be the set of groups containing no row indexed by $\mathcal{G}$. By Lemma 13.3, $|\mathcal{Q}| < M^{1/4}$. For every $\Gamma(u,a) \in \mathcal{G}$, we have $\mathbf{H}_\infty(\mathbf{Z}_{u,a}) \geq m - \Delta = L - (\Delta + h)$.

Fix any t groups outside $\mathcal{Q}$ and choose one good row in each. These groups give $D = d_0 t$ indexed rows with distinct prefixes (u,a). By Lemma 12.10, the output on each chosen row is θ-close to an independent uniform bit, jointly with the outputs on all the other indexed rows in these groups. Taking the XOR within each group preserves this guarantee: each of the t group outputs is θ-close to an independent uniform bit jointly with the other $t-1$ group outputs. By Lemma 4.6, their joint distribution is $t\theta$-close to $\mathbf{U}_t$. This completes the proof, recalling that by choice $t\theta \leq M^{-(t+1)}$. □

Proof of Theorem 13.1. Choose an integer $t \geq 2$ such that $C_{\mathrm{Maj}}(\log t)/\sqrt{t} \leq \varepsilon/3$, and use Construction 13.5 with the resulting constant C. The result follows from Lemmas 13.4 and 13.6. □

14 Extraction from polylogarithmic entropy

In this section, we extend the construction of Section 13 to extract $k^{\Omega(1)}$ bits with inverse-polynomial error when $k \geq \mathrm{poly}(\log n)$. The higher entropy allows the local correlation breaker to handle a growing number of rows. We replace the resilient function with a construction based on Feige's lightest bin protocol [Fei99], using the correlation breaker twice: first to assign almost t-wise independent bin labels, and then to extract from the rows in a least populated bin. Finally, we apply Shaltiel's result [Sha08] to obtain the desired output length.

Theorem 14.1. *Assume H_{NP}. There are constants $A, a > 0$ such that, for every pair of constants $c, e > 0$, there is $n_0 = n_0(c,e)$ such that the following holds. For every $n \geq n_0$ and every $(\log n)^A \leq k$, there is a (k, n^{-e})-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{\lfloor k^a \rfloor}$ for $\mathcal{C}^{\mathrm{NP}}_{n,n^c}$, computable in time $2^{O(\log^2 n/\log k)}$.*

Theorem 14.2 (Extracting almost all entropy)**.** *Assume H_{NP}. There is a constant $A > 0$ such that, for every $\eta \in (0,1)$, and for every pair of constants $c, e > 0$, there is $n_0 = n_0(c,e,\eta)$ such that the following holds. For every $n \geq n_0$ and every $(\log n)^A \leq k$, there is a (k, n^{-e})-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{\lfloor (1-\eta)k \rfloor}$ for $\mathcal{C}^{\mathrm{NP}}_{n,n^c}$, computable in time $2^{O(\log^2 n/\log k)}$.*

We first state the correlation-breaking and selection tools, then set up the ingredients and prove Theorem 14.1. We prove Theorem 14.2 in Section 14.4. In Section 15, we improve the running time of our extractor to polynomial under H_{Σ_2}.

14.1 Tools

The first tool is local correlation breaking for more rows and output bits. The other two bound the probability that some bin contains no good row.

Local correlation breaking. We use the multi-bit local correlation breaker from Lemma 12.11, with the indexed-row convention of Section 12.3. The row length and entropy deficiency include the fixed index. We instantiate this result for bin labels and final extraction in Lemmas 14.5 and 14.6.

Committee selection. We use Feige's lightest bin protocol for committee selection on the rows: we assign each row a bin label, then select a lightest bin. The selected bin has at most $\lfloor R/T \rfloor$ rows. We prove a couple of lemmas upper bounding the probability that there's a bin with no good row. Both bounds follow from standard moment arguments. For completeness, we provide the proofs.

Lemma 14.3 (A good row in every bin: statistical distance)**.** *Let $\mathbf{Y}_1, \dots, \mathbf{Y}_R$ be jointly distributed labels in $[T]$, where $1 \le T \le R$. Fix $G \subseteq [R]$, let $g = |G|$ and $\mu = g/T$, and let $\mathbf{Z}_z = \sum_{i \in G} \mathbf{1}[\mathbf{Y}_i = z]$ for $z \in [T]$. Suppose t is even with $2 \le t \le \mu$. If every set of at most t labels indexed by G has statistical distance at most ω from independent uniform labels in $[T]$, then*

$$\Pr[\exists z \in [T] : \mathbf{Z}_z = 0] \le T \left(\frac{64t}{\mu}\right)^{t/2} + 2\omega T^{t+1}.$$

Lemma 14.4 (A good row in every bin: relative error)**.** *Let $\mathbf{Y}_1, \dots, \mathbf{Y}_R$, G, μ, and $\mathbf{Z}_z$ be as in Lemma 14.3. Suppose t is even with $2 \le t \le \mu$. If, for every $1 \le d \le t$, distinct $i_1, \dots, i_d \in G$, and $z \in [T]$,*

$$\left|\Pr[\mathbf{Y}_{i_1} = \cdots = \mathbf{Y}_{i_d} = z] - T^{-d}\right| \le \delta T^{-d},$$

then

$$\Pr[\exists z \in [T] : \mathbf{Z}_z = 0] \le T \left(\frac{64t}{\mu}\right)^{t/2} + 3^t T\delta.$$

Setup for both proofs: Fix a bin $z \in [T]$, put $p = 1/T$, and let $\xi_i = \mathbf{1}[\mathbf{Y}_i = z]$ for $i \in G$. Let ξ'_i, $i \in G$, be independent Bernoulli variables of mean p. Write $\mathbf{Z}_z = \sum_{i \in G} \xi_i$ and $\mathbf{Z}'_z = \sum_{i \in G} \xi'_i$, so $\mathbb{E}\,\mathbf{Z}'_z = \mu = gp$. The standard moment bound [BR94] gives, for even t with $2 \le t \le \mu$,

$$\mathbb{E}[(\mathbf{Z}'_z - \mu)^t] \le (64\mu t)^{t/2}. \tag{39}$$

Proof of Lemma 14.3. Expand

$$(\mathbf{Z}_z - \mu)^t = \sum_{(i_1,\dots,i_t) \in G^t} \prod_{j=1}^{t} (\xi_{i_j} - p),$$

and similarly for $\mathbf{Z}'_z$. Each product depends on at most t labels and takes values in $[-1, 1]$. The statistical distance hypothesis bounds the difference of corresponding expectations by 2ω. There are g^t products, so

$$\mathbb{E}[(\mathbf{Z}_z - \mu)^t] \le (64\mu t)^{t/2} + 2\omega g^t.$$

Since $\mathbf{Z}_z = 0$ implies $(\mathbf{Z}_z - \mu)^t = \mu^t$, Markov's inequality gives

$$\Pr[\mathbf{Z}_z = 0] \le \left(\frac{64t}{\mu}\right)^{t/2} + 2\omega T^t.$$

A union bound over the T bins proves the claim. □

Proof of Lemma 14.4. The case $T = 1$ is immediate, so suppose $T \geq 2$. For every nonempty $I \subseteq G$ with $|I| \leq t$, the hypothesis gives

$$\left| \mathbb{E}\left[\prod_{i \in I} \xi_i\right] - \mathbb{E}\left[\prod_{i \in I} \xi'_i\right] \right| \leq \delta p^{|I|}.$$

We use these bounds to compare the expansions of $(\mathbf{Z}_z - \mu)^t$ and $(\mathbf{Z}'_z - \mu)^t$.

Fix an ordered tuple $(i_1, \ldots, i_t) \in G^t$, let J be its set of distinct indices, and let m_i be the multiplicity of $i \in J$. Since $\xi_i \in \{0, 1\}$,

$$(\xi_i - p)^{m_i} = \alpha_i + \beta_i \xi_i, \qquad \alpha_i = (-p)^{m_i}, \qquad \beta_i = (1-p)^{m_i} - (-p)^{m_i}.$$

Here $|\alpha_i| \leq p$ and $|\beta_i| \leq 1$. Writing $d = |J|$, the coefficient of $\prod_{i \in I} \xi_i$ in the product over J has absolute value at most $p^{d-|I|}$. The constant terms cancel, so the corresponding product expectations differ by at most

$$\delta \sum_{\varnothing \neq I \subseteq J} p^{d-|I|} p^{|I|} \leq \delta (2p)^d.$$

There are at most $\binom{t}{d} g^d t^{t-d}$ ordered tuples with d distinct indices: choose the positions where new indices first appear, their values, and an earlier index for each remaining position. Summing the expectation differences gives

$$\begin{aligned} \left| \mathbb{E}[(\mathbf{Z}_z - \mu)^t] - \mathbb{E}[(\mathbf{Z}'_z - \mu)^t] \right| &\leq \delta \sum_{d=1}^{t} \binom{t}{d} g^d t^{t-d} (2p)^d \\ &\leq \delta (2\mu + t)^t \leq \delta (3\mu)^t. \end{aligned}$$

Together with Equation (39), this yields

$$\mathbb{E}[(\mathbf{Z}_z - \mu)^t] \leq (64\mu t)^{t/2} + \delta (3\mu)^t.$$

Markov's inequality and a union bound over the T bins now give

$$\Pr[\exists z \in [T] : \mathbf{Z}_z = 0] \leq T \left(\frac{64t}{\mu}\right)^{t/2} + 3^t T \delta.$$

□

14.2 The extractor construction

We now present the extractor construction, first setting up its ingredients. We split each hash row into two parts. The first part determines its bin label; the second part supplies the entropy for extraction after selection.

Fix $c, e > 0$. Choose a constant $0 < \rho \leq 1/2$ no larger than the exponent in Lemma 12.11, and put $a = \rho/8$. We take the entropy exponent A sufficiently large that $Aa > 4$. All bounds below hold for $(\log n)^A \leq k \leq n$. Parameters below with subscript f denote parameters of the final extractor.

- **Hash rows.** Apply Lemma 11.1 at threshold k to obtain $h_1, \ldots, h_R : \{0,1\}^n \to \{0,1\}^m$, with $n \leq R \leq n^{C_H}$. For every source under consideration, there is a fixed set $G \subseteq [R]$ of size at least $R - R^{1-\eta_H}$ such that $\mathbf{H}_\infty(h_i(\mathbf{X})) \geq m - \Delta_h$ for $i \in G$, where $\eta_H, \Delta_h > 0$ are constants. Write $h_i = h_i^{(0)} \parallel h_i^{(1)}$, with part lengths $m_0 = \lfloor m/2 \rfloor$ and $m_1 = m - m_0$. Both parts of every good row have deficiency at most Δ_h.

- **Bins and indices.** Set $s = \lceil k^a \rceil$ and $T = 2^{\lfloor \log(R/s) \rfloor}$, so $s \le R/T < 2s$. Let $D_f = \lfloor R/T \rfloor$, $h_0 = \lceil \log R \rceil$, and $h_f = \lceil \log D_f \rceil$. Fix distinct h_0-bit encodings of $[R]$ and distinct h_f-bit encodings of $[D_f]$. The indexed row lengths are $L_0 = m_0 + h_0$ and $L_f = m_1 + h_f$.

- **Labeling parameters.** Choose a sufficiently large constant K and let t be the least even integer at least $K \log n / \log k$. Choose a constant $C_s > e + 6$ such that any t indexed first parts are jointly samplable with nondeterministic postselection by a circuit of size at most n^{C_s}. Let S_0 be the least power of two at least $n^{C_s} T^{2t}$, and set $q_0 = \log T$.

- **Extraction parameters.** Set $q_1 = \lfloor k^a \rfloor$ and $\zeta = n^{-(e+3)}$. Let S_f be the least power of two at least $(nS_0)^{C_f}$, for a sufficiently large constant C_f, and put

$$\Delta_f = \Delta_h + h_f + \left\lceil \log \frac{D_f R}{\zeta} \right\rceil .$$

  The choice of C_f covers the conditional sampling circuits in Lemma 14.8 and ensures $S_f^{-1} \le \zeta$.

The next two lemmas instantiate the correlation breaker with these parameters.

Lemma 14.5 (Correlation breaking for bin labels)**.** *Assume* H_{NP}. *With the preceding parameters, there is a* $\mathrm{poly}(S_0)$*-time computable map* $\mathsf{CB}_{\mathrm{lab}} : \{0,1\}^{L_0} \to \{0,1\}^{q_0}$ *such that, for every* $\mathbf{X} \in \mathcal{C}^{\mathrm{NP}}_{n,n^c}$ *with* $\mathbf{H}_\infty(\mathbf{X}) \ge k$*, the labels*

$$\mathbf{Y}_i = \mathsf{CB}_{\mathrm{lab}}(i \parallel h_i^{(0)}(\mathbf{X})) \in [T]$$

satisfy

$$\mathsf{SD}\big((\mathbf{Y}_{i_1}, \dots, \mathbf{Y}_{i_d}), \mathbf{U}_{[T]^d}\big) \le d/S_0$$

for every $1 \le d \le t$ *and distinct* $i_1, \dots, i_d \in G$*. Here* $\mathbf{U}_{[T]^d}$ *denotes* d *independent uniform labels in* $[T]$.

Proof. Apply Lemma 12.11 with $D = t$, output length q_0, row length L_0, deficiency $\Delta_h + h_0$, and error S_0^{-1}. The joint sampling circuit computes the specified hashes of $\mathsf{Samp}(\mathbf{R})$, keeps their first parts, appends their indices, and conditions on $\mathsf{Post}(\mathbf{R}) = 1$. Its size is at most $n^{C_s} \le S_0$.

We have $L_0 \ge k/3$, $\log S_0 = O(t \log n)$, and $q_0 = O(\log n)$. The left side of Equation (31) is $O(t^2 \log n) = O(\log^3 n) = o(k^\rho)$, and $t, q_0 \le L_0^\rho$. Thus the correlation breaker can be used. By Lemma 4.6, any d good labels are jointly dS_0^{-1}-close to independent uniform labels. □

Lemma 14.6 (Correlation breaking for the selected rows)**.** *Assume* H_{NP}. *With the preceding parameters, there is a* $\mathrm{poly}(S_f)$*-time computable map* $\mathsf{CB}_{\mathrm{ext}} : \{0,1\}^{L_f} \to \{0,1\}^{q_1}$ *with the following property. For every tuple* $(\mathbf{Z}_1, \dots, \mathbf{Z}_d)$ *of at most* D_f *indexed rows with distinct* h_f*-bit prefixes, jointly samplable with nondeterministic postselection by a size-*S_f *circuit, and every* $j \in [d]$ *with* $\mathbf{H}_\infty(\mathbf{Z}_j) \ge L_f - \Delta_f$*,*

$$\mathsf{SD}\Big(\big(\mathsf{CB}_{\mathrm{ext}}(\mathbf{Z}_j), (\mathsf{CB}_{\mathrm{ext}}(\mathbf{Z}_\ell))_{\ell \ne j}\big), \big(\mathbf{U}_{q_1}, (\mathsf{CB}_{\mathrm{ext}}(\mathbf{Z}_\ell))_{\ell \ne j}\big)\Big) \le \zeta.$$

Proof. Apply Lemma 12.11 with $D = D_f$, output length q_1, row length L_f, deficiency Δ_f, and error S_f^{-1}. Here $L_f = \Theta(k)$, $\Delta_f = O(\log n)$, $D_f = O(k^a)$, and $\log S_f = O(t \log n)$. Hence the left side of Equation (31) is $O(k^a \log^2 n + k^{2a}) = o(k^\rho)$, and $D_f, q_1 \le L_f^\rho$. The resulting error is at most $S_f^{-1} \le \zeta$. □

Construction 14.7. *Define* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{q_1}$ *as follows: on input* x*,*

(i) Compute the hash rows and their bin labels

$$Y_i(x) = \mathsf{CB}_{\mathrm{lab}}(i \parallel h_i^{(0)}(x)) \in [T], \qquad i \in [R].$$

(ii) Choose a least populated bin, breaking ties by its index, and list its row indices as $i_1 < \cdots < i_d$. Then $d \le D_f$.

(iii) Index the selected second parts by their positions and pad with zeros:

$$W_j(x) = \begin{cases} j \parallel h_{i_j}^{(1)}(x), & 1 \le j \le d, \\ j \parallel 0^{m_1}, & d < j \le D_f. \end{cases}$$

(iv) Output

$$\mathsf{Ext}(x) = \bigoplus_{j=1}^{D_f} \mathsf{CB}_{\mathrm{ext}}(W_j(x)). \tag{40}$$

14.3 Analysis

The proof has two steps: we first analyze extraction from the selected rows assuming that every bin contains a good row with high probability, and then we establish this condition using the almost-independence of the good labels. We formulate the first step for arbitrary row maps h_i and an arbitrary labeling map $\mathsf{CB}_{\mathrm{lab}}$, so that it also applies in Section 15.

Lemma 14.8 (Extraction from the selected rows)**.** *There is a fixed polynomial p with the following property. Consider the procedure in Construction 14.7 with arbitrary row maps $h_i = h_i^{(0)} \parallel h_i^{(1)}$ of lengths m_0, m_1 and an arbitrary labeling map $\mathsf{CB}_{\mathrm{lab}}$. Suppose $\mathbf{X}$ is samplable with nondeterministic postselection by a size-s_X circuit, each row map has circuit size at most s_h, and the labeling map has circuit size at most s_0. Let $G \subseteq [R]$ satisfy $\mathbf{H}_\infty(h_i^{(1)}(\mathbf{X})) \ge m_1 - \Delta_h$ for every $i \in G$, and suppose every bin contains a row indexed by G with probability at least $1 - \varepsilon_0$.*

For $0 < \zeta \le 1/2$, let $D_f = \lfloor R/T \rfloor$, $h_f = \lceil \log D_f \rceil$, $L_f = m_1 + h_f$, and $\Delta_f = \Delta_h + h_f + \lceil \log(D_f R/\zeta) \rceil$. Suppose $\mathsf{CB}_{\mathrm{ext}}$ satisfies the local guarantee in Lemma 12.11 with error at most ζ for at most D_f indexed rows of length L_f, good-row deficiency at most Δ_f, and joint sampling circuit size at most S_f. If $S_f \ge p(n, s_X, s_h, s_0, R, m_0 + m_1)$, then

$$\mathsf{SD}(\mathsf{Ext}(\mathbf{X}), \mathbf{U}_{q_1}) \le \varepsilon_0 + 2\zeta.$$

Proof. Let $\mathcal{C}$ be the event that every bin contains a good row. For $j \in [D_f]$ and $i \in G$, let $E_{j,i}$ be the event that $\mathcal{C}$ occurs and the first good selected row has original index i and selected position j. These events partition $\mathcal{C}$. There are at most $D_f R$ events, so those with probability below $\zeta/(D_f R)$ have total probability at most ζ.

Fix a remaining event $E_{j,i}$. By Lemma 4.4,

$$\mathbf{H}_\infty(h_i^{(1)}(\mathbf{X}) \mid E_{j,i}) \ge m_1 - \Delta_h - \log(D_f R/\zeta).$$

On this event, $W_j(\mathbf{X}) = j \parallel h_i^{(1)}(\mathbf{X})$, so $\mathbf{H}_\infty(W_j(\mathbf{X}) \mid E_{j,i}) \ge L_f - \Delta_f$.

Write the original sampling circuit as $(\mathsf{Samp}, \mathsf{Post})$. To sample the selected tuple conditioned on $E_{j,i}$, choose a uniform input $\mathbf{R}$, compute $x = \mathsf{Samp}(\mathbf{R})$ and $(W_1(x), \ldots, W_{D_f}(x))$, and condition on $\mathsf{Post}(\mathbf{R}) = 1$ and $E_{j,i}(x)$. The circuit tests $E_{j,i}$ using the fixed membership bits of G: it checks that every

bin contains a good row and identifies the original index and position of the first good selected row. Since $E_{j,i}$ can be checked deterministically, conditioning further on $E_{j,i}$ preserves nondeterministic postselection.

The conditional sampling circuit has size at most $p(n, s_X, s_h, s_0, R, m_0 + m_1) \le S_f$. The rows have distinct position prefixes, and the jth row has deficiency at most Δ_f. The local guarantee therefore allows us to replace $\mathsf{CB}_{\mathrm{ext}}(W_j(\mathbf{X}))$ by a uniform string independent of the other outputs, changing their joint distribution by statistical distance at most ζ. Note that taking XOR cannot increase this distance. After the replacement, the XOR is uniform, so

$$\mathsf{SD}\big((\mathsf{Ext}(\mathbf{X}) \mid E_{j,i}), \mathbf{U}_{q_1}\big) \le \zeta.$$

Let β be the total probability of $\neg\mathcal{C}$ and the discarded events. We have $\beta \le \varepsilon_0 + \zeta$. On each remaining event, the conditional output is ζ-close to uniform; on the other events, we use the trivial bound of 1 on statistical distance. Averaging these bounds gives

$$\mathsf{SD}(\mathsf{Ext}(\mathbf{X}), \mathbf{U}_{q_1}) \le (1-\beta)\zeta + \beta \le \zeta + \varepsilon_0 + \zeta = \varepsilon_0 + 2\zeta.$$

□

Proof of Theorem 14.1. Use the parameters from Section 14.2, and let $\mathbf{X} \in \mathcal{C}^{\mathrm{NP}}_{n,n^c}$ have min-entropy at least k. By Lemma 14.5, every set of at most t good labels is t/S_0-close to independent uniform labels. Since $g = |G| \ge R/2$, we have $\mu = g/T \ge s/2$. For sufficiently large n, the entropy threshold gives $2 \le t \le \mu$ and $\log(s/(128t)) \ge (a/2)\log k$. Thus Lemma 14.3 gives

$$\begin{aligned}\Pr[\neg\mathcal{C}] &\le T(128t/s)^{t/2} + 2tT^{t+1}/S_0 \\ &\le n^{C_H - aK/4} + 2tn^{-C_s} \le \zeta,\end{aligned}$$

for sufficiently large K and n.

The choice of C_f ensures that S_f covers the conditional sampling circuits in Lemma 14.8: the original circuit, row maps, and number of rows have polynomial size in n, and the labeling map has size $\mathrm{poly}(S_0)$. By Lemmas 14.6 and 14.8, with $\varepsilon_0 = \zeta$, we have

$$\mathsf{SD}(\mathsf{Ext}(\mathbf{X}), \mathbf{U}_{q_1}) \le 3\zeta \le n^{-e}.$$

There are R computations of $\mathsf{CB}_{\mathrm{lab}}$ and D_f computations of $\mathsf{CB}_{\mathrm{ext}}$, with polynomial overhead for hashing and selection. The total running time is $\mathrm{poly}(n, S_0, S_f) = 2^{O(t \log n)} = 2^{O(\log^2 n/\log k)}$. □

14.4 Extracting almost all the entropy

We use Shaltiel's result [Sha08], in the form stated in Lemma 4.33 to extend the output to almost all the entropy.

Proof of Theorem 14.2. Take C, c_0, d_0 from Lemma 4.33 with $d = e$, and set $k_0 = \lfloor \eta k/2 \rfloor$. For sufficiently large A independent of c, e and sufficiently large n, Theorem 14.1 with parameters c_0, d_0 gives an extractor at threshold k_0 whose output can be truncated to $t = \lceil C \log n \rceil \le k_0$ bits. The extension of Lemma 4.33 to $\mathcal{C}^{\mathrm{NP}}$ applies since $2k_0/\eta \le k$, giving the desired output length and error in time $2^{O(\log^2 n/\log k_0)} + \mathrm{poly}(n) = 2^{O(\log^2 n/\log k)}$. □

15 Polynomial-time extraction from polylogarithmic entropy

We obtain polynomial running time under H_{Σ_2} by using the multiplicative correlation breaker from Theorem 12.12 for the bin labels in Section 14. Its pointwise guarantee controls the rare label events needed for selection with only polynomial cost. We then select $O(\log n)$ rows, extract using Lemmas 12.11 and 14.8, and amplify the output using Lemma 4.33.

Theorem 15.1. *Assume H_{Σ_2}. There is a constant $A > 0$, such that for every $\eta \in (0,1)$, and for all constants $c, e > 0$, for every $k > (\log n)^A$, there is a polynomial-time computable (k, n^{-e})-extractor $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^{\lfloor (1-\eta)k \rfloor}$ for $\mathcal{C}^{\mathrm{NP}}_{n,n^c}$.*

15.1 Relative probabilities of bin labels

Applied to any $d \le t$ good indexed rows and the output tuple $(z, \dots, z)$, Theorem 12.12 gives the comparisons in the following lemma.

Lemma 15.2 (Relative probabilities of labels)**.** *Let $\mathbf{Y}_1, \dots, \mathbf{Y}_R$ be labels in $[T]$, with good index set G. Suppose that, for every ordered tuple of $1 \le d \le t$ distinct indices in G and every $z \in [T]$, the probabilities*

$$P = \Pr[\mathbf{Y}_{i_1} = \cdots = \mathbf{Y}_{i_d} = z], \qquad Q = T^{-1} \Pr[\mathbf{Y}_{i_2} = \cdots = \mathbf{Y}_{i_d} = z]$$

satisfy $P \le e^{\vartheta} Q + \xi$ and $Q \le e^{\vartheta} P + \xi$, with the empty conjunction having probability one. If $0 < \delta \le 1/2$, $\vartheta = \delta/(8t)$, and $\xi = \vartheta T^{-t}$, then

$$\left| \Pr[\mathbf{Y}_{i_1} = \cdots = \mathbf{Y}_{i_d} = z] - T^{-d} \right| \le \delta T^{-d}.$$

Proof. Fix distinct indices $i_1, \dots, i_d \in G$ and $z \in [T]$, where $d \le t$. We replace the labels one at a time by independent uniform labels. After replacing the first j labels, the probability that all labels equal z is $p_j = T^{-j} \Pr[\mathbf{Y}_{i_{j+1}} = \cdots = \mathbf{Y}_{i_d} = z]$. Thus $p_0 = P$ and $p_d = T^{-d}$.

For $0 \le j < d$, the hypothesis gives $p_j \le e^{\vartheta} p_{j+1} + T^{-j}\xi$ and $p_{j+1} \le e^{\vartheta} p_j + T^{-j}\xi$. Iterating these inequalities and using $T^{-j} \le 1$ gives

$$e^{-d\vartheta} T^{-d} - d\xi \le P \le e^{d\vartheta} T^{-d} + d e^{d\vartheta} \xi.$$

Our choices ensure $d\vartheta \le \delta/8$ and $d\xi T^d = d\vartheta T^{d-t} \le \delta/8$. The upper bound is therefore at most $e^{\delta/8}(1 + \delta/8)T^{-d} \le (1+\delta)T^{-d}$, and the lower bound is at least $(e^{-\delta/8} - \delta/8)T^{-d} \ge (1-\delta)T^{-d}$, since $0 < \delta \le 1/2$. Thus $|P - T^{-d}| \le \delta T^{-d}$. □

15.2 The polynomial-time construction

Theorem 15.3. *Assume H_{Σ_2}. There exist constants $A, \gamma > 0$, such that for all constants $c, e > 0$, for every $k > (\log n)^A$, there is a polynomial-time computable (k, n^{-e})-extractor $B_0 : \{0,1\}^n \to \{0,1\}^{\lfloor k^{\gamma} \rfloor}$ for $\mathcal{C}^{\mathrm{NP}}_{n,n^c}$.*

Proof. Choose $0 < \rho \le 1/2$ no larger than the exponent in Lemma 12.11, and let ρ_m be the exponent in Theorem 12.12. Set $\gamma = \rho/2$ and choose A so that $A\rho, A\rho_m > 4$.

Ingredients.

- Use the hash rows $h_i = h_i^{(0)} \parallel h_i^{(1)}$ from Section 14.2, with $n \le R \le n^{C_H}$ and $m_0, m_1 = \Theta(k)$. Let $\Delta_H = O(1)$ denote the upper bound on the entropy deficiency of a good row.
- Choose an integer $e_0 > e + 3$ and set $\zeta = n^{-e_0}$. Let t be the least even integer at least $K \log n$, and let T be the largest power of two at most $R/(C_0 t)$, for sufficiently large constants C_0, K. Then $D_f = \lfloor R/T \rfloor = O(\log n)$. Set

$$\delta = \frac{\zeta}{3^{t+1} R}, \qquad \vartheta = \frac{\delta}{8t}, \qquad \xi = \vartheta T^{-t}.$$

- Let $h_0 = \lceil \log R \rceil$, $L_0 = m_0 + h_0$, $\Delta_0 = \Delta_H + h_0$, and $q_0 = \log T$. Choose $S_0 = n^{C_s}$, with $C_s \ge 2$ large enough to cover the joint sampler for any t indexed first parts. Take $\mathsf{CB}_{\mathrm{lab}}$ from Theorem 12.12 with these parameters, at most t rows, and errors ϑ, ξ. Define $Y_i(x) = \mathsf{CB}_{\mathrm{lab}}(i \parallel h_i^{(0)}(x))$.
- Use these labels in Construction 14.7, and let $W_1, \dots, W_{D_f}$ be the selected, reindexed, and padded rows. Set $q_1 = \lfloor k^\gamma \rfloor$, $h_f = \lceil \log D_f \rceil$, and $L_f = m_1 + h_f$. Let

$$\Delta_f = \Delta_H + h_f + \lceil \log(D_f R/\zeta) \rceil .$$

  Choose a polynomial bound $S_f \ge n^2$ covering the conditional sampling circuits in Lemma 14.8 and satisfying $S_f^{-1} \le \zeta$. Since H_{Σ_2} implies H_{NP}, take $\mathsf{CB}_{\mathrm{ext}}$ from Lemma 12.11 with row length L_f, at most D_f rows, deficiency Δ_f, output length q_1, and error S_f^{-1}.

Define

$$B_0(x) = \bigoplus_{j=1}^{D_f} \mathsf{CB}_{\mathrm{ext}}(W_j(x)).$$

Analysis. We first check the parameter conditions for the correlation breakers. For the labeling map, $1/\vartheta = n^{O(1)}$ and $\xi = 2^{-\Theta(\log^2 n)}$, so a fixed constant $b > 0$ satisfies $\xi \le S_0^{-b} \le \vartheta \le 1/10$. The left-hand side of (33), with $D_* = t$, $q = q_0$, and $S = S_0$, is $O(\log^3 n) = o(L_0^{\rho_m})$ because $A\rho_m > 4$. Thus Theorem 12.12 supplies the required labeling map in polynomial time.

Consequently, the conditional sampling circuits in Lemma 14.8 have polynomial size, so we can choose $S_f = \mathrm{poly}(n)$. We have $\Delta_f = O(\log n)$, $D_f = O(\log n)$, and $\log S_f = O(\log n)$. The left-hand side of (31) is therefore $O(\log^2 n + k^\gamma \log n) = o(L_f^\rho)$, since $\gamma = \rho/2$ and $A\rho > 4$. Thus Lemma 12.11 gives a polynomial-time computable $\mathsf{CB}_{\mathrm{ext}}$ with error $S_f^{-1} \le \zeta$.

Now fix $\mathbf{X} \in \mathcal{C}^{\mathrm{NP}}_{n,n^c}$ with $\mathbf{H}_\infty(\mathbf{X}) \ge k$. The hash guarantee gives a fixed set G of at least $R/2$ good rows. Applying the pointwise guarantee of $\mathsf{CB}_{\mathrm{lab}}$ and Lemma 15.2 shows that the good labels satisfy the hypothesis of Lemma 14.4. Let $\mathcal{C}$ be the event that every bin contains a good row. Since $\mu = |G|/T \ge C_0 t/2$, that lemma gives

$$\Pr[\neg \mathcal{C}] \le T(64t/\mu)^{t/2} + 3^t T\delta \le R 2^{-t/2} + \zeta/3 \le \zeta,$$

where the last two inequalities follow by choosing C_0 and then K sufficiently large.

By Lemma 14.8, with $\varepsilon_0 = \zeta$, we obtain $\mathsf{SD}(B_0(\mathbf{X}), \mathbf{U}_{q_1}) \le 3\zeta \le n^{-e}$. The construction makes polynomially many evaluations of polynomial-time maps, so its running time is polynomial. □

Proof of Theorem 15.1. Follows by combining Lemma 4.33 and Theorem 15.3. The proof is the same as in Theorem 14.2. □

Acknowledgement Part of the work was done while the authors were visiting the Simons Institute for Theory of Computing, UC Berkeley. We would like to thank the Simons Institute for the stimulating research environment and the hospitality.

References

[Adl78] Leonard M. Adleman. "Two theorems on random polynomial time". In: *Proceedings of the 19th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE, 1978, pp. 75–83 (cit. on pp. 17, 73, 75).

[ADNOPRS19] Divesh Aggarwal, Ivan Damgård, Jesper Buus Nielsen, Maciej Obremski, Erick Purwanto, João Ribeiro, and Mark Simkin. "Stronger Leakage-Resilient and Non-Malleable Secret Sharing Schemes for General Access Structures". In: *Advances in Cryptology – CRYPTO 2019, Part II*. Ed. by Alexandra Boldyreva and Daniele Micciancio. Vol. 11693. Lecture Notes in Computer Science. Full version: Cryptology ePrint Archive, Report 2018/1147. Springer, 2019, pp. 510–539 (cit. on p. 76).

[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 pp. 1, 4, 5, 15, 17, 38).

[AIKS16] S. Artemenko, R. Impagliazzo, V. Kabanets, and R. Shaltiel. "Pseudorandomness When the Odds are Against You". In: *31st Conference on Computational Complexity, CCC*. Vol. 50. 2016, 9:1–9:35 (cit. on p. 1).

[AK02] V. Arvind and J. Köbler. "New Lowness Results for ZPP^{NP} and Other Complexity Classes". In: *J. Comput. Syst. Sci.* 65.2 (2002), pp. 257–277 (cit. on p. 1).

[BM88] László Babai and Shlomo Moran. "Arthur–Merlin games: a randomized proof system, and a hierarchy of complexity classes". In: *Journal of Computer and System Sciences* 36.2 (1988), pp. 254–276 (cit. on pp. 17, 73, 75).

[BSS24] M. Ball, R. Shaltiel, and J. Silbak. "Non-malleable Codes with Optimal Rate for Poly-Size Circuits". In: *Advances in Cryptology - EUROCRYPT*. Vol. 14654. Lecture Notes in Computer Science. 2024, pp. 33–54 (cit. on p. 1).

[BD25] Marshall Ball and Dana Dachman-Soled. "Uniform black-box separations via non-malleable extractors". In: *Advances in Cryptology—CRYPTO 2025, Part I*. Full version: Cryptology ePrint Archive, Report 2025/1060. 2025, pp. 585–616 (cit. on pp. 11, 16, 17, 72, 73).

[BDL22] Marshall Ball, Dana Dachman-Soled, and Julian Loss. "(Nondeterministic) hardness vs. non-malleability". In: *Advances in Cryptology—CRYPTO 2022, Part I*. Full version: Cryptology ePrint Archive, Report 2022/070. 2022, pp. 148–177 (cit. on pp. 1, 11, 16, 17, 72, 73).

[BGDM23] Marshall Ball, Eli Goldin, Dana Dachman-Soled, and Saachi Mutreja. "Extracting randomness from samplable distributions, revisited". In: *Proceedings of the 64th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2023, pp. 1505–1514 (cit. on pp. 1, 2, 4–8, 18, 28, 38, 39, 42, 46).

[BSS25] Marshall Ball, Ronen Shaltiel, and Jad Silbak. "Extractors for samplable distributions with low min-entropy". In: *Proceedings of the 57th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*. Ed. by Michal Koucký and Nikhil Bansal. ACM, 2025, pp. 596–603 (cit. on pp. 2–8, 38, 42, 61, 63).

[BOV07] B. Barak, S. J. Ong, and S. P. Vadhan. “Derandomization in Cryptography”. In: *SIAM J. Comput.* 37.2 (2007), pp. 380–400 (cit. on p. 1).

[BR94] Mihir Bellare and John Rompel. “Randomness-efficient oblivious sampling”. In: *Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 1994, pp. 276–287 (cit. on pp. 14, 91).

[BCDT19] Avraham Ben-Aroya, Gil Cohen, Dean Doron, and Amnon Ta-Shma. “Two-source condensers with low error and small entropy gap via entropy-resilient functions”. In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*. Vol. 145. LIPIcs. Schloss Dagstuhl—Leibniz-Zentrum für Informatik, 2019, 43:1–43:20 (cit. on pp. 8, 21, 42, 44).

[BDT17] Avraham Ben-Aroya, Dean Doron, and Amnon Ta-Shma. “An efficient reduction from two-source to non-malleable extractors: achieving near-logarithmic min-entropy”. In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*. ACM, 2017, pp. 1185–1194 (cit. on pp. 11, 13, 88).

[BV17] N. Bitansky and V. Vaikuntanathan. “A Note on Perfect Correctness by Derandomization”. In: *Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Vol. 10211. 2017, pp. 592–606 (cit. on p. 1).

[BKMO26] Farzan Byramji, Daniel M. Kane, Jackson Morris, and Anthony Ostuni. “Hard-To-Sample Distributions from Robust Extractors”. In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2026)*. Ed. by Mohit Singh and Tom Gur. Vol. 392. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026, 37:1–37:22 (cit. on p. 6).

[CGL21] Eshan Chattopadhyay, Jesse Goodman, and Jyun-Jie Liao. “Affine extractors for almost logarithmic entropy”. In: *Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. Full version: ECCC Report TR21-075, June 3, 2021. 2021, pp. 622–633 (cit. on pp. 13, 88, 89).

[CGL20] Eshan Chattopadhyay, Vipul Goyal, and Xin Li. “Nonmalleable extractors and codes, with their many tampered extensions”. In: *SIAM Journal on Computing* 49.5 (2020). Journal version of the STOC 2016 paper, pp. 999–1040 (cit. on pp. 16, 73, 76).

[CL22] Eshan Chattopadhyay and Jyun-Jie Liao. “Extractors for sum of two sources”. In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. 2022, pp. 1584–1597 (cit. on pp. 73, 76).

[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. 11, 88).

[CT22] L. Chen and R. Tell. “When Arthur has Neither Random Coins nor Time to Spare: Superfast Derandomization of Proof Systems”. In: *Electron. Colloquium Comput. Complex.* TR22-057 (2022). ECCC: TR22-057 (cit. on p. 1).

[DS26] Alon Dermer and Ronen Shaltiel. “Multiplicative pseudorandom generators for nondeterministic circuits”. In: *41st Computational Complexity Conference (CCC)*. Ed. by Dana Moshkovitz. Vol. 383. LIPIcs. Full version: ECCC Report TR26-017, Revision 1, June 22, 2026. Schloss Dagstuhl—Leibniz-Zentrum für Informatik, 2026, 18:1–18:20 (cit. on pp. 16, 25, 73, 74, 86, 87).

[DY13] Y. Dodis and Y. Yu. “Overcoming Weak Expectations”. In: *Theory of Cryptography - 10th Theory of Cryptography Conference, TCC*. Vol. 7785. Lecture Notes in Computer Science. 2013, pp. 1–22 (cit. on p. 4).

[DMOZ22] D. Doron, D. Moshkovitz, J. Oh, and D. Zuckerman. “Nearly Optimal Pseudorandomness from Hardness”. In: *J. ACM* 69.6 (2022), 43:1–43:55 (cit. on p. 1).

[Dru13] Andrew Drucker. “Nondeterministic Direct Product Reductions and the Success Probability of SAT Solvers”. In: *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS*. 2013, pp. 736–745 (cit. on p. 1).

[DMNS06] C. Dwork, F. McSherry, K. Nissim, and A. D. Smith. “Calibrating Noise to Sensitivity in Private Data Analysis”. In: *Theory of Cryptography, Third Theory of Cryptography Conference, TCC*. Vol. 3876. Lecture Notes in Computer Science. Springer, 2006, pp. 265–284 (cit. on p. 4).

[Fei99] Uriel Feige. “Noncryptographic selection protocols”. In: *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 1999, pp. 142–152 (cit. on pp. 11, 14, 90).

[GLRSW91] Peter Gemmell, Richard J. Lipton, Ronitt Rubinfeld, Madhu Sudan, and Avi Wigderson. “Self-testing/correcting for polynomials and for approximate functions”. In: *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing (STOC)*. Ed. by Cris Koutsougeras and Jeffrey Scott Vitter. Vol. 91. ACM, 1991, pp. 32–42 (cit. on p. 39).

[GW02] O. Goldreich and A. Wigderson. “Derandomization That Is Rarely Wrong from Short Advice That Is Typically Good”. In: *APPROX-RANDOM*. 2002, pp. 209–223 (cit. on p. 1).

[GL20] Oded Goldreich and Maya Leshkowitz. “On emulating interactive proofs with public coins”. In: *Computational Complexity and Property Testing: On the Interplay Between Randomness and Computation*. Vol. 12050. Lecture Notes in Computer Science. 2020, pp. 178–198 (cit. on pp. 17, 73, 75).

[GS86] Shafi Goldwasser and Michael Sipser. “Private coins versus public coins in interactive proof systems”. In: *Proceedings of the 18th Annual ACM Symposium on Theory of Computing (STOC)*. 1986, pp. 59–68 (cit. on pp. 8, 12, 17, 27, 40, 73, 75).

[GKORS18] Sivakanth Gopi, Swastik Kopparty, Rafael Mendes de Oliveira, Noga Ron-Zewi, and Shubhangi Saraf. “Locally testable and locally correctable codes approaching the Gilbert–Varshamov bound”. In: *IEEE Transactions on Information Theory* 64.8 (2018), pp. 5813–5831 (cit. on p. 26).

[GUV09] Venkatesan Guruswami, Christopher Umans, and Salil P. Vadhan. “Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes”. In: *Journal of the ACM* 56.4 (2009), 20:1–20:34 (cit. on p. 21).

[GST03] Dan Gutfreund, Ronen Shaltiel, and Amnon Ta-Shma. “Uniform hardness versus randomness tradeoffs for Arthur-Merlin games”. In: *Computational Complexity* 12.3-4 (2003), pp. 85–130 (cit. on p. 1).

[HNY17] P. Hubácek, M. Naor, and E. Yogev. “The Journey from NP to TFNP Hardness”. In: *8th Innovations in Theoretical Computer Science Conference, ITCS*. Vol. 67. 2017, 60:1–60:21 (cit. on p. 1).

[IW97] Russell Impagliazzo and Avi Wigderson. “$P = \mathsf{BPP}$ if E requires exponential circuits: derandomizing the XOR lemma”. In: *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC)*. 1997, pp. 220–229 (cit. on p. 1).

[KM02] A. Klivans and D. van Melkebeek. “Graph Nonisomorphism Has Subexponential Size Proofs Unless the Polynomial-Time Hierarchy Collapses”. In: *SIAM J. Comput.* 31.5 (2002), pp. 1501–1526 (cit. on p. 1).

[Li13a] Xin Li. “Extractors for a constant number of independent sources with polylogarithmic min-entropy”. In: *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. Full version: ECCC Report TR13-025, Revision 1. IEEE, 2013, pp. 100–109 (cit. on p. 11).

[Li13b] Xin Li. “New independent source extractors with exponential improvement”. In: *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC)*. Full version: ECCC Report TR12-147. ACM, 2013, pp. 783–792 (cit. on p. 11).

[Li15] Xin Li. “Three-source extractors for polylogarithmic min-entropy”. In: *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 2015, pp. 863–882 (cit. on pp. 8, 11, 22, 42, 44, 61).

[Li23] Xin Li. “Two source extractors for asymptotically optimal entropy, and (many) more”. In: *Proceedings of the 64th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. Full version: arXiv:2303.06802v2, May 30, 2023. IEEE, 2023, pp. 1271–1281 (cit. on pp. 16, 73, 76).

[MV05] P. Bro Miltersen and N. V. Vinodchandran. “Derandomizing Arthur-Merlin Games using Hitting Sets”. In: *Computational Complexity* 14.3 (2005), pp. 256–279 (cit. on p. 1).

[NT99] Noam Nisan and Amnon Ta-Shma. “Extracting randomness: A survey and new constructions”. In: *Journal of Computer and System Sciences* 58.1 (1999), pp. 148–173 (cit. on pp. 10, 30).

[OS26] Justin Oh and Ronen Shaltiel. “Extractors for samplable distributions from the two-source extractor recipe”. In: *Proceedings of the 58th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*. ACM, 2026, pp. 1344–1352 (cit. on pp. 2–4).

[RSW06] Omer Reingold, Ronen Shaltiel, and Avi Wigderson. “Extracting randomness via repeated condensing”. In: *SIAM Journal on Computing* 35.5 (2006), pp. 1185–1209 (cit. on pp. 6, 10, 30).

[SSZ98] Michael Saks, Aravind Srinivasan, and Shiyu Zhou. “Explicit OR-dispersers with polylogarithmic degree”. In: *Journal of the ACM* 45.1 (1998), pp. 123–154 (cit. on pp. 9, 30).

[SS24] R. Shaltiel and J. Silbak. “Explicit Codes for Poly-Size Circuits and Functions That Are Hard to Sample on Low Entropy Distributions”. In: *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC*. 2024, pp. 2028–2038 (cit. on p. 1).

[SU06] R. Shaltiel and C. Umans. “Pseudorandomness for Approximate Counting and Sampling”. In: *Computational Complexity* 15.4 (2006), pp. 298–341 (cit. on p. 1).

[SU09] R. Shaltiel and C. Umans. “Low-End Uniform Hardness versus Randomness Tradeoffs for AM”. In: *SIAM J. Comput.* 39.3 (2009), pp. 1006–1037 (cit. on p. 1).

[Sha08] Ronen Shaltiel. “How to get more mileage from randomness extractors”. In: *Random Structures & Algorithms* 33.2 (2008), pp. 157–186 (cit. on pp. 25, 90, 95).

[Sha25a] Ronen Shaltiel. “Extractors for samplable distributions with polynomially small min-entropy”. In: *Proceedings of the 66th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE Computer Society, 2025, pp. 934–946 (cit. on pp. 2–7, 10, 22, 25, 28–30, 38, 42, 44, 61).

[Sha25b] Ronen Shaltiel. “Multiplicative extractors for samplable distributions”. In: *40th Computational Complexity Conference (CCC)*. Ed. by Srikanth Srinivasan. Vol. 339. LIPIcs. Full version: ECCC Report TR24-168, Revision 3, June 22, 2026. Schloss Dagstuhl—Leibniz-Zentrum für Informatik, 2025, 22:1–22:22 (cit. on pp. 1, 2, 4, 5, 16, 23, 27, 38, 72–74, 86, 87).

[SU05] Ronen Shaltiel and Christopher Umans. “Simple extractors for all min-entropies and a new pseudorandom generator”. In: *Journal of the ACM* 52.2 (2005), pp. 172–216 (cit. on pp. 1, 12, 27).

[Ta-96] Amnon Ta-Shma. “On extracting randomness from weak random sources”. In: *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC)*. 1996, pp. 276–285 (cit. on pp. 9, 30).

[Sto85] Larry Stockmeyer. “On approximation algorithms for #P”. In: *SIAM Journal on Computing* 14.4 (1985), pp. 849–861 (cit. on p. 84).

[Sud97] Madhu Sudan. “Decoding of Reed–Solomon codes beyond the error-correction bound”. In: *Journal of Complexity* 13.1 (1997), pp. 180–193 (cit. on pp. 8, 41).

[STV01] Madhu Sudan, Luca Trevisan, and Salil P. Vadhan. “Pseudorandom generators without the XOR lemma”. In: *Journal of Computer and System Sciences* 62.2 (2001), pp. 236–266 (cit. on pp. 7, 39, 46).

[TV00] Luca Trevisan and Salil Vadhan. “Extracting randomness from samplable distributions”. In: *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2000, pp. 32–42 (cit. on pp. 1, 2, 4, 7, 38, 40, 42, 63).

[Vio14] Emanuele Viola. “Extractors for circuit sources”. In: *SIAM Journal on Computing* 43.2 (2014), pp. 655–672 (cit. on pp. 13, 89).

[Zuc07] David Zuckerman. “Linear degree extractors and the inapproximability of Max Clique and Chromatic Number”. In: *Theory of Computing* 3.1 (2007), pp. 103–128 (cit. on pp. 13, 88).