Quadratic Extractors

Gautam Chandrasekaran[*†] Jesse Goodman[‡] David Zuckerman[*§]

Abstract

A quadratic source is a random variable $X \sim \mathbb{F}_2^n$ generated by applying n quadratics to a hidden uniform string. Previously, the best-known explicit extractor for this model required min-entropy roughly $k = n - \sqrt{\log n}$. In this paper, we give an explicit construction for min-entropy $k = \text{polylog}(n)$. Even for the special case of 2-local sources, where each bit of X depends on two hidden uniform bits, achieving an entropy requirement below $k = \sqrt{n}$ was an important open problem. This is because such sources provably cannot be reduced (in a black-box manner) to our most powerful general-purpose seedless extractor, known as a sumset extractor. Our new extractor bypasses this barrier, providing the first explicit construction to go beyond sumset extraction.

[*]This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing, Berkeley.

[†]University of Texas at Austin. gautamc@cs.utexas.edu. Supported in part by the NSF AI Institute for Foundations of Machine Learning (IFML) and a Simons Investigator Award (#409864, David Zuckerman).

[‡]Simons Institute for the Theory of Computing. jpmgoodman@berkeley.edu. Most of this work was done while the author was at The University of Texas at Austin, supported by a Simons Investigator Award (#409864, David Zuckerman).

[§]University of Texas at Austin. diz@cs.utexas.edu. Supported in part by NSF Grant CCF-2312573 and a Simons Investigator Award (#409864).

1 Introduction

Uniform random bits are a fundamental building block of randomized algorithms, cryptography, and distributed computing. However, the sources of randomness we encounter in nature (usually arising from physical phenomena) are imperfect, and almost never truly uniform. The goal of the field of *randomness extraction* is to design efficient deterministic algorithms that convert these imperfect (also called *weak*) sources of randomness into truly uniform random bits.

Definition 1.1 (Randomness Extractor)**.** *A randomness extractor* $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ *with error* ϵ *for a class of distributions* $\mathcal{X}$ *over* $\{0,1\}^n$ *is a deterministic function such that for all* $X \in \mathcal{X}$*,*

$$\mathrm{Ext}(X) \approx_\epsilon U_m,$$

where $\approx_\epsilon$ *denotes statistical distance and* U_m *is the uniform distribution on* m *bits. We say that the extractor is explicit if it is a deterministic polynomial time algorithm.*

We quantify the randomness present in these *weak* sources using the notion of *min-entropy*.

Definition 1.2 (Min-Entropy)**.** *The min-entropy of a random variable* X *is defined as*

$$H_\infty(X) = \min_{x \in \mathrm{supp}(X)} \log_2 \left(\frac{1}{\Pr[X = x]} \right).$$

The dream randomness extractor would be a single deterministic polynomial time algorithm that works for the class of all weak random sources that have an entropy guarantee. However, this is known to be impossible. For any candidate extractor, an adversary can design a weak random source (even with min-entropy $n-1$) on which the extractor outputs a constant string.

The community has broadly studied three approaches of extractors that get around this impossibility result. The first approach, called *seeded extraction*, assumes that the extractor is also given a short truly uniform string in addition to the weak random source [NZ96, Tre01, LRVW03, GUV09, DKSS13]. The second approach, called *independent source extraction*, assumes that the extractor is given multiple independent weak sources as input [Vaz87, CG88, BIW06, Bou05, Li15, CZ19, Li23]. The final approach, and the one we study, assumes algebraic or complexity theoretic structure in the source.

Algebraic Sources The most basic type of algebraic source is an *affine source*. These are distributions obtained by passing uniform bits through an affine map (or equivalently, uniform over the pre-images of linear maps). That is, they are sources of the form $X = \ell(U_s) + b$ where $\ell : \mathbb{F}_2^s \to \mathbb{F}_2^n$ is a linear map, U_s is the uniform distribution over $\mathbb{F}_2^s$ and $b \in \mathbb{F}_2^n$ is a fixed vector. Extractors and dispersers[1] for affine sources have a rich history, starting from [BSHR+01, BKS+10, Bou07, Rao09, Yeh11, BK12, Li11, Sha11, Li16, CGL21, CL22] culminating in constructions achieving the optimal entropy requirement $k = O(\log n)$ by Li [Li23]. Affine extractors have also been constructed over larger fields by [GR08, DG10, BDL16, BGLZ15, GVJZ23]. Affine extractors and dispersers have had various applications both within extractors [Vio14, CG22] and to circuit lower bounds [DK11, CS16, FGHK16, LY22].

There are two natural ways to generalize affine sources to polynomials. The first is the class of *variety sources*: uniform distributions over the solutions of polynomial equations. This type of source, introduced by Dvir [Dvi12], has good extractors over large fields [Dvi12], but the best constructions over $\mathbb{F}_2$ require entropy $0.99n$ [CT15, Rem16, LZ19]. Better constructions for these sources would imply new circuit lower bounds [GKW21]. The second generalization is the class of *polynomial sources*, which will be the main object of our study.

[1] A disperser is a function that is non-constant on the source.

Polynomial sources are sources obtained by passing uniform bits through polynomial maps. That is, the sources are of the form $X = p(U_s)$ where p is a low degree polynomial map. This type of source was first introduced by Dvir, Gabizon and Wigderson [DGW09], where they constructed good extractors over fields with size polynomial in n. Subsequently, Ben-Sasson and Gabizon [BSG12] improved their result, in particular constructing extractors for sources with linear entropy over constant size fields (at least 4). However, the techniques of both these papers, which use character sum estimates, all break down over $\mathbb{F}_2$. The problem of constructing *quadratic source* (polynomial sources with degree 2 polynomial maps) extractors over $\mathbb{F}_2$ remained open for a long time, even for entropy $n - 1$! Recently, Chattopadhyay, Goodman and Gurumukhani [CGG24] constructed such extractors for entropy roughly $n - \sqrt{\log n}$. In this paper, we give an exponential improvement by constructing quadratic extractors over $\mathbb{F}_2$ with entropy requirement $\operatorname{poly}\log n$.

Theorem 1.3 (Extractors for quadratic sources over $\mathbb{F}_2$, see Theorem 9.4)**.** *There exists a universal constant $C > 0$ such that for any $k \geq \log^C n$, there exists an explicit function $\mathrm{QExt} : \{0,1\}^n \to \{0,1\}^m$ such that for any quadratic source X over $\mathbb{F}_2$ with entropy at least k, it holds that*

$$\mathrm{QExt}(X) \approx_{n^{-\Omega(1)}} U_m$$

where $m = k^{\Omega(1)}$.

Our construction is the first to go beyond sumset extraction, due to known barrier results on using sumset extractors for quadratic sources [CGG24, AGMR25]. Our key technical ideas include a win-win argument, a structure theorem for quadratic sources and a new object in algebraic pseudorandomness (that we call a rank combiner). We refer to Section 3 for more details. We now discuss some implications of our result for extracting from complexity-theoretic sources.

Complexity-Theoretic Sources The other commonly studied class of structured sources involves sources motivated by complexity theory. Trevisan and Vadhan initiated this line of work in their seminal paper [TV00] where they studied sources samplable by polynomial sized circuits. Their extractor is conditional, relying on strong complexity theoretic assumptions. The pursuit of extractors for *samplable sources* under various complexity theoretic assumptions has seen remarkable progress recently [BDSGM23, Sha25b, BSS25, Sha25a, OS25]. Unconditional extractors are known for sources generated in small space [KRVZ06, CL16, CG22, CL22] .

Our focus will also be on unconditional extractors. De and Watson [DW12], and Viola [Vio14], introduce the notion of locally samplable sources: sources obtained by passing uniform bits through local maps. A d-local source has the form $X = f(U_s)$, where each output bit f_i of f only depends on d input bits. A 1-local source (also called *projections* in [PSZ97]) is a special case of affine sources and hence we know explicit extractors [Li23] even for entropy $k = O(\log n)$. The situation is much worse for 2-local sources, where the best known extractor requires entropy roughly $\sqrt{n}$ [Li16] (by a reduction to affine sources). There were barriers (similar to the quadratic source case) on going below $\sqrt{n}$ using sumset extractors [ACG+22, AGMR25]. As an immediate corollary of our main theorem Theorem 1.3, we have the following result.

Corollary 1.4 (Extractors for 2-local sources)**.** *There exists a universal constant $C > 0$ such that for any $k \geq \log^C n$, there exists an explicit function $\mathrm{QExt} : \{0,1\}^n \to \{0,1\}^m$ such that for any 2-local source X over $\mathbb{F}_2$ with entropy at least k, it holds that*

$$\mathrm{QExt}(X) \approx_{n^{-\Omega(1)}} U_m$$

where $m = k^{\Omega(1)}$.

Our final result concerns distributions samplable by polynomial size circuits of constant depth (AC^0). Viola [Vio14] had shown reductions from this class to local sources. He gave extractors with entropy requirement roughly $n^{2/3}$ and Li [Li16] improved this to $\sqrt{n}$. In this work, we give an improved reduction to 2-local sources and obtain extractors that only need entropy roughly $n^{1/3}$.

Corollary 1.5 (Extractors for AC^0 sources, see Theorem 10.6)**.** *For any constant $\alpha > 0$ and any $k \geq n^{1/3+\alpha}$, there exists an explicit function* $\mathrm{QExt} : \{0,1\}^n \to \{0,1\}^m$ *such that for any* AC^0 *samplable source X over $\mathbb{F}_2$ with entropy at least k, it holds that*

$$\mathrm{QExt}(X) \approx_{n^{-\Omega(1)}} U_m$$

where $m = k^{\Omega(1)}$.

AI Disclosure We used an explicit construction of a rank combiner (Sections 8.1 to 8.3) provided to us by ChatGPT Astra, after we defined the object. This was one of several ingredients in our proof. We humans did everything else in this paper, including the writing in Sections 8.1 to 8.3.

2 Notation

We will use capitalized letters to denote random variables. We use the notation U_t to denote a uniform random string on t bits. For a field $\mathbb{F}$, we denote $\mathbb{F}[t]$ to be the polynomial ring over $\mathbb{F}$ and $\mathbb{F}(t)$ to be the field of rational functions over $\mathbb{F}$. We use $\mathbb{N}$ to denote the natural numbers. We use $|X|$ to denote the length of the string X. For a function f with domain $\mathbb{F}_2^n$, we define $f|_S$ to be the restriction of f to a subset $S \subseteq \mathbb{F}_2^n$. We use $d_{\mathrm{TV}}(X,Y)$ to denote the total variation distance (also called statistical distance) between two random variables. For distributions X, Y, we say that $X \approx_\epsilon Y$ if $d_{\mathrm{TV}}(X,Y) \leq \epsilon$. For random variables R, S, T satisfying $(R,T) \approx_\epsilon (S,T)$, we say that $(R \approx_\epsilon S) \mid T$. We call a function f a d-junta is if it only depends on at most d of its input bits.

We use $X^{[t]}$ to denote a sequence of random variables $\{X^i\}_{i\in[t]}$. For a string $x \in \mathbb{F}_2^n$, $x_{[a,b]}$ denotes the substring of x with starting index a and ending index b. When $a = 1$, we use the shorthand $x_{[b]}$.

Given a distribution X on $\mathbb{F}_2^n$, for any set $S \subseteq [n]$, we define $X_S := \{X_i\}_{i\in S}$. We say that a source X is *flat* if X is uniform over a set. For a set S, we define U_S as the random variable that is uniform over it. We use the shorthand U_m to denote a uniform string on m bits. We denote the support of X by $\mathrm{supp}(X)$. Given a set $S \subseteq [n]$, we define the function $\chi_S : \mathbb{F}_2^n \to \{-1,1\}$ as $\chi_S(x) = (-1)^{\sum_{i\in S} x_i}$. A source X has ϵ-smooth min entropy k, denoted by $H_\infty^\epsilon(X) \geq k$ if there exists some source $\tilde{X}$ with $H_\infty(\tilde{X}) \geq k$ such that $X \approx_\epsilon \tilde{X}$. A set $S \subseteq \mathbb{N}$ is a Sidon set if $|S+S| \geq \binom{|S|}{2} + |S|$ where $S+S = \{a+b \mid a \in S, b \in S\}$.

3 Proof Overview

Throughout this section we use C and c to denote sufficiently large and small positive constants, respectively.

3.1 A First Attempt: Black Box Reductions to Affine/Sumset Sources

A natural first attempt at this problem is the following: try to find a sufficiently large affine subspace/sumset in the support of the quadratic source, and then appeal to excellent existing constructions of low entropy affine/sumset extractors. This strategy had already achieved varying levels of success for many other types of sources, including NC^0 and AC^0-sources with entropy $C\sqrt{n}$ [Vio14, Li16], and small space sources with entropy $C \log n$ [CG22, CL22].

Unfortunately, there is a fundamental barrier to this approach [ACG+22, CGG24, AGMR25]. Quadratic maps (and even 2-local maps) can sample *Sidon sets*, which are sets that achieve maximum doubling

($|S + S| = \binom{|S|}{2} + 1$). These are maximally non-affine in a sense, as affine subspaces do not grow under addition. More formally, Sidon sets do not contain affine spaces of dimension 2 (see [CGG24, Claim 5.6]). This means that any model that can sample (large) Sidon sets cannot be reduced, in a black-box way, to affine sources. In fact, the same holds for sumset sources.

Our approach begins with the questions: Can we exploit the additive structure, or the lack thereof, that makes these reductions fail? Is there a win-win approach to constructing quadratic extractors?

3.2 A Simple Affine Extractor

The above impossibility results show that a generic sumset extractor cannot work. To get around this, we attempt to analyze a specific construction, and see where things break down. We will then exploit the structure from the barrier results to overcome these failures.

A central object in almost all known affine extractors is a *strong linear seeded extractor*.

Definition 3.1 (Strong Linear Seeded Extractor)**.** *A function* $\mathrm{LExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a* (k, ϵ, δ)*-strong linear seeded extractor with seed length* d *and output length* m *if*

1. *the function* $\mathrm{LExt}_y(x) = \mathrm{LExt}(x, y)$ *is linear as a function of* x *for all seeds* y*, and*
2. *for every source* X *with* $H_\infty(X) \geq k$*, with probability at least* $1 - \epsilon$ *over* $y \sim \{0,1\}^d$*, it holds that* $\mathrm{LExt}(X, y) \approx_\delta U_m$.

When $\epsilon = \delta$*, we refer to the function as a* (k, ϵ)*-strong linear seeded extractor. If item 2 above only works for sources from a class* $\mathcal{X}$*, then we say that the function is an* (ϵ, δ)*-strong linear seeded extractor for* $\mathcal{X}$*. We refer to* ϵ *as the strength error and* δ *as the closeness error.*

We start by analyzing one of the simplest affine extractors:[2]

$$\mathrm{Ext}(X) = \mathrm{LExt}(X, X_{[d]})$$

where $X_{[d]}$ outputs the first d bits in X and d is appropriately chosen.

We first recall the proof of why $\mathrm{Ext}(X) = \mathrm{LExt}(X, X_{[0.1n]})$ is an affine extractor for affine sources with entropy $0.99n$. We choose LExt to be a $(0.1n, \epsilon)$-strong linear seeded extractor with seed-length $0.1n$, outputting 1 bit with $\epsilon = 2^{-0.05n}$ (we can use for instance an extractor from [CGL21]). The proof involves two main steps that we now discuss.

Affine Conditioning The first step in the proof is the following lemma from [Rao09], known as *affine conditioning*.

Lemma 3.2 (Affine Conditioning, [Rao09])**.** *Let* X *be an affine source with* $H_\infty(X) \geq k$ *and suppose* $L : \mathbb{F}_2^n \to \mathbb{F}_2^m$ *is a linear map. Then,* $X = A + B$ *for independent affine sources* A, B *such that*

1. $H_\infty(A) \geq k - m$,
2. $H_\infty(B) \leq m$*, and*
3. $L(a) = 0$ *for all* $a \in \mathrm{supp}(A)$.

[2] There are simpler affine extractors (like inner product), but the following extractor will be most relevant to our construction.

Before explaining why this lemma is useful, we briefly sketch its proof. Observe that an affine source with entropy k can be represented as $X = v + M(U_k)$ for a full rank linear map M and fixed vector v. Now, the map $L \circ M$ has rank at most m; hence its null space P has dimension at least $k - m$. Hence, $\mathbb{F}_2^k = P + Q$ for Q with dimension at most m. Taking $A = M(P)$ and $B = v + M(Q)$ completes the proof.

Now, going back to our candidate extractor Ext, observe that the function $S(X) = X_{[0.1n]}$ is linear. Thus, we have

$$\mathrm{Ext}(X) = \mathrm{LExt}(X, X_{[0.1n]}) = \mathrm{LExt}(A + B, S(A + B)) = \mathrm{LExt}(A + B, S(B)) \tag{1}$$

where A and B are the independent affine sources given by Lemma 3.2 satisfying (1) $H_\infty(A) \geq 0.89n$ and (2) $H_\infty(B) \leq 0.1n$.

Using Linearity From the linearity of LExt after fixing the seed, we have that $\mathrm{Ext}(X) = \mathrm{LExt}(A, S(B)) + \mathrm{LExt}(B, S(B))$. Suppose first that $S(B)$ is uniform. The guarantee of the strong seeded extractor would imply that $\mathrm{LExt}(A, S(b))$ is close to uniform with high probability over fixing $B = b$. The second term becomes constant after fixing B and thus does not affect the uniformity.

Now, we remove this restriction and consider our scenario where $S(X)$ is a prefix of X of length $0.1n$. The fact that $H_\infty(X) \geq 0.99n$ implies that $H_\infty(S(B)) \geq 0.09n$. This implies that the probability of $S(B)$ being in any set of seeds is at most $2^{0.01n}$ times the corresponding probability over the uniform distribution. Our choice of $\epsilon = 2^{-0.05n}$ implies that $S(B)$ is a good seed with high probability and completes the proof.

3.3 Extending to Quadratics

We will now try to analyze the previous construction for quadratic sources, but we will see that a few problems arise. We take $\mathrm{Ext}(X) = \mathrm{LExt}(X, X_{[d]})$, where we now choose $d = n^c$ for some small constant $c > 0$. We are again targeting the near-linear entropy regime $k \geq n - O(n^c)$ (note that the previous best known extractors from [CGG24] required entropy $n - O(\sqrt{\log n})$). Throughout most of this discussion, we assume that our quadratic source $X = f(U_k)$ satisfies $H_\infty(X) = k$. This is not without loss of generality: unlike affine sources, a quadratic source need not admit a representation where the number of generating bits matches the entropy. We make this assumption for simplicity, and in our actual proof, we use an input reduction lemma to remove it (Lemma 4.14).

Quadratic Conditioning The first part of our argument will generalize the affine conditioning step in the previous argument. Clearly, the lack of linearity in our source prevents us from using the rank based argument from earlier. Regardless, we will argue the following.

Lemma 3.3 (Informal)**.** *Let $X = f(U_k)$ be a quadratic source and let $p : \mathbb{F}_2^k \to \mathbb{F}_2^m$ be a quadratic map. Then $(f(U_k), p(U_k))$ is ϵ-close to a convex combination of sources of the form $(f(A + B), q(B))$ where A and B are independent affine sources in $\mathbb{F}_2^k$ such that*

1. *$H_\infty(A) = \Omega(k/m)$.*
2. *$H_\infty(B) = \ell \leq m + O(\log(1/\epsilon))$,*
3. *$q(B) = p(a + B)$ for all $a \in \mathrm{supp}(A)$,*
4. *$q(B) \approx_\epsilon p(U_k)$, and*

The lemma states that the source can be decomposed into sources satisfying the property that p only depends on a low dimensional affine subspace (property 2 and 3) of the generating bits. Furthermore, this decomposition preserves the distribution of p (property 4). The analogy in the affine case is taking the linear

function $\ell = L \circ M$ (where M is the matrix generating the source). It is also crucial that there is entropy in the source independent of p. This is ensured by property 1 (note the multiplicative loss this time, compared to the additive loss in the affine case). We now sketch a proof of Lemma 3.3.

Proof Sketch of Lemma 3.3. There are two main steps. First, we argue that the distribution of p is preserved even after fixing all but $m + O(\log(1/\epsilon)$ generating bits. We then argue that we can do this fixing efficiently, where we make p independent of the rest of the generating bits without trivially fixing all of them.

The key tool for the first step is the *input reduction lemma* due to Chattopadhyay, Goodman and Gurumukhani [CGG24].

Lemma 3.4 (Input Reduction: Short Output Length [CGG24])**.** *Let $p : \mathbb{F}_2^k \to \mathbb{F}_2^m$ be a quadratic map. For any $0 < \epsilon < 1/4$, for a random linear map $L : \mathbb{F}_2^k \to \mathbb{F}_2^{k-\ell}$ with $\ell = m + O(\log(1/\epsilon))$ it holds with probability at least $1 - \epsilon$ that*

$$(p(U_k), L(U_k)) \approx_\epsilon (p(U_k), U'_{k-\ell}))$$

where U', U are independent, uniform random variables.

Let $e_1, \ldots, e_\ell$ be a basis for the nullspace of L and $e_{\ell+1}, \ldots, e_k$ be a extension of the basis to all of $\mathbb{F}_2^k$. We have that $U_k = A + B$ for independent A, B where A is uniform over the span of $\{e_{\ell+1}, \ldots, e_k\}$ and B is uniform over the nullspace of L. We can rewrite the polynomial p as being a quadratic function of A, B. We have

$$p(A + B) = p_1(A) + p_2(B) + p_3(A, B)$$

where p_3 is bilinear. From Lemma 3.4, we now have that with probability $1 - \epsilon$ over $a \sim A$, it holds that $p(a + B) \approx_\epsilon p(U_k)$. This completes the proof of the first step.

We now implement the second step. We will convert p to a function q solely depending on B by fixing all dependencies on A. We want to do so while still retaining entropy in A. Note that p_3 depends on $m\ell$ linear functions of A. We will fix these $m\ell$ linear functions along with $p_1(A)$ which is a quadratic function. However, this causes an issue when compared to the affine case: A becomes a quadratic variety after fixing $p_1(A)$. The additional guarantee that p_1 only has m output bits now comes to the rescue. We use the following theorem by Cohen and Tal [CT15].[3]

Theorem 3.5 ([CT15, Theorem 3.6])**.** *Let $q_1, \ldots, q_t$ be quadratic functions from $\mathbb{F}_2^k \to \mathbb{F}_2$. There exists a universal constant C (independent of $q_1, \ldots, q_t$) such that $\mathbb{F}_2^k$ can be partitioned into affine subspaces each of dimension $d = C(k/t) - t$ where $q_1, \ldots, q_t$ are constant when restricted to each of these subspaces.*

The above theorem is saying that the input space of a small set of quadratic polynomials can be partitioned with affine subspaces such that these polynomials are constant on each of these subspaces. This is exactly what we need to fix $p_1(A)$. Theorem 3.5 allows us to fix $p_1(A)$ by breaking A into affine subspaces where p_1 is constant (we need to also fix $m\ell$ linear functions of A to make p_3 a function of B). Each such subspace A' has dimension $\Omega(k/m)$ as long as $k \gg m\ell$, which will be the case in all our applications. Renaming A' as A completes the proof. □

Since A and B are affine subspaces of $\mathbb{F}_2^k$, we can rewrite $f(A + B) = f'(Y, Z)$ where $Y = U_{\dim(A)}$, $Z = U_{\dim(B)}$, and f' is a quadratic map. Similarly, $q(B) = q'(Z)$ for some quadratic q'.

For our running example $\mathrm{Ext}(X) = \mathrm{LExt}(X, X_{[d]})$, we take $p = (f(U_k))_{[d]}$ with $d = n^c$ as the quadratic in Lemma 3.3. Assuming $k \geq n - 0.01n^c$, we have $H_\infty(p(U_k)) \geq 0.99d$. Thus, finally after applying the decomposition from Lemma 3.3, and using the renaming above, it suffices to analyze random variables of the form $\mathrm{LExt}(f(Y, Z), q(Z))$ where $Y = U_s$ for $s = \Omega(k/d)$ and $H_\infty^\epsilon(q(Z)) \geq 0.99d$.

[3] The use of Cohen-Tal to retain affine structure after fixing a few quadratics (in the context of extractors) appeared in [DG26].

Obtaining Linearity via Cauchy Schwarz Similar to the affine case, we analyze the case where $q(Z) = U_d$ is exactly uniform (the $0.99d$ entropy case can be easily handled by using the same argument as in Section 3.2 and we omit it here).

Thus, our setup is the following: we have a source $X = f(Y, Z)$ with $Y = U_s$ for $s = \Omega(k/d) = \Omega(n^{1-c})$ and $q(Z) = U_{n^c}$. We are analyzing the function $\mathrm{LExt}(X, q(Z)) = \mathrm{LExt}(f(Y, Z), q(Z))$.

It suffices to argue that all the parities are unbiased, as we can infer near-uniformity from Vazirani's XOR lemma. The affine argument clearly does not extend: f is a quadratic and hence we can't break it into two terms where the first is independent of Z. Let $b(z) = |\mathbb{E}_Y[\chi_S(\mathrm{LExt}(f(Y, z), q(z)))]|$ and $f(Y, Z) = f_1(Y) + f_2(Y, Z) + f_3(Z)$ where f_2 is bilinear. Our insight is to use the Cauchy-Schwarz inequality. This was inspired by other applications of Cauchy-Schwarz in the proofs of correlations bounds, xor lemmas and pseudorandom generators for $\mathbb{F}_2$-polynomials [BNS92, BV10, Lov08, Vio09, VW08]. We have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{Z}[(b(Z))^2] &= \mathop{\mathbb{E}}_{Z}(\mathop{\mathbb{E}}_{Y}[\chi_S(\mathrm{LExt}(f(Y, Z), q(Z)))])^2 \\
&= \mathop{\mathbb{E}}_{Y,Z}\mathop{\mathbb{E}}_{Y,A}[\chi_S(\mathrm{LExt}(f(Y, Z) + f(Y + A, Z), q(Z)))] \\
&= \mathop{\mathbb{E}}_{Z}\mathop{\mathbb{E}}_{Y,A}[\chi_S(\mathrm{LExt}(\partial_A f_1(Y) + f_2(A, Z), q(Z)))
\end{aligned}$$

where $\partial_a g(x) \coloneqq g(x) + g(x + a)$ is the direction derivative in the direction a.

Observe that for any fixed a, we have that $\partial_a f_1(.)$ is an affine source! Also, $f_2(a, Z)$ is only a function of Z. Now, we can use linearity similar to the affine proof. We have

$$\mathop{\mathbb{E}}_{Z}[(b(Z))^2] \le 2^{-s} + \max_{a\neq 0}\mathop{\mathbb{E}}_{Z}|\mathop{\mathbb{E}}_{Y}[\chi_S(\mathrm{LExt}(\partial_a f_1(Y), q(Z)))|.$$

We have reduced the problem to bounding the error of a linear seeded extractor when applied to the derivative sources $\partial_a f_1(Y)$. A sufficient condition for the error to be small is if $\min_{a\neq 0} H_\infty(\partial_a f_1(Y))$ is large (greater than the entropy requirement of LExt). We call these quantities "derivative entropies". We record this argument in the following lemma.

Lemma 3.6 (Informal, see Lemma 6.1)**.** *Let $X = f(Y, Z)$ be a source with f quadratic, $Y = U_s$, $Z = U_t$ independent of Y and such that $f(Y, Z) = f_1(Y) + f_2(Y, Z) + f_3(Z)$ where (1) $\min_{a\neq 0} H_\infty(\partial_a f_1(Y)) \ge k$ and (2) f_2 is bilinear. Let $\mathrm{LExt} : \mathbb{F}_2^n \times \mathbb{F}_2^d \to \mathbb{F}_2^m$ be a (k, ϵ) strong linear seeded extractor and $q(Z)$ be a function of Z satisfying $q = U_d$. Then, it holds with probability $1 - 2^m \cdot (2^{-s} + 2\epsilon)^{1/2}$ over $z \sim U_t$ that*

$$d_{\mathrm{TV}}(\mathrm{LExt}(f(Y, z), q(z)), U_m) \le 2^{m/2}(2^{-s} + 2\epsilon)^{1/4}.$$

It is still unclear if we have made any progress, as we do not know any bounds on these derivative entropies.

Now, recall our discussion on the barriers to reducing the whole problem to affine sources in Section 3.1. The bad examples were Sidon sets: sets which have maximum doubling. Remarkably, these are precisely the good cases for our discussion in this section: Sidon sets have maximum derivative entropy! Indeed, for a source representing a Sidon set, say $X = f(U_k)$, we have that for any $a \neq 0$, $f(x) + f(x + a)$ is distinct for every unordered pair $(x, x + a)$, which implies $H_\infty(\partial_a f(U_k)) = k - 1$. This is the starting point of our win-win analysis.

3.4 Structure Theorem for Quadratics

The main ingredient for our win-win analysis is a structural result on quadratic sources.

Theorem 3.7. *Let X be a quadratic source with an injective generating map $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$. Let $r, t \in \mathbb{N}$ be such that $rt \le k/2$. Then, X can be exactly expressed as a convex combination of sources of the following two types:*

1. *Affine sources with entropy at least t (obtained by applying f to an affine subspace V such that $f|_V$ is affine), or*

2. *Quadratic sources with an injective generating map $g : \mathbb{F}_2^{k'} \to \mathbb{F}_2^n$ satisfying (1) $\min_{a \neq 0} H_\infty(\partial_a g(U_{k'})) \geq r$, (2) $k' \geq k - rt$ and (3) $g = f \circ L$ for an injective affine map $L : \mathbb{F}_2^{k'} \to \mathbb{F}_2^k$.*

Again, for simplicity we assume that the number of generating bits matches the entropy. The above theorem says that any quadratic source is a convex combination of (1) high entropy affine sources, or (2) sources with high derivative entropy. As discussed in the previous sections, this dichotomy seems useful, as we could try employing an affine analysis for the first case, and a quadratic analysis (using quadratic conditioning and Cauchy-Schwarz) for the second. We now give a proof sketch.

Proof Sketch of Theorem 3.7. Our proof of the theorem relies on a simple observation. Consider a quadratic source $X = f(U_k)$ with $\partial_a f(.) = c$ for some non-zero direction a and constant vector c. This implies that the quadratic function looks linear when viewed along the direction a. Formally, take a basis B of dimension $k - 1$ such that $\mathbb{F}_2^k = \text{span}\{\{a\} \cup B\}$. We have that

$$f(ya + b) = y\partial_a f(b) + f(b) = yc + f(b)$$

for any bit y and $b \in B$. Thus taking a random bit Y and a random variable Z uniform over B and independent of Y, we have that original source is can be expressed as

$$g(Y, Z) = Yc + f(Z).$$

Thus, we have obtained a one dimensional affine source from our original source f: fixing Z decomposes X into a convex combination of shifts of these one dimensional affine sources.

The idea is now to do this linearizing procedure every time we find a direction a with $H_\infty(\partial_a f(.)) \leq r$. For any such direction, we condition on the value of $\partial_a f(.) = c$. This loses at most r bits of entropy by our definition of a. Now after conditioning, we can use the linearization trick from above to obtain a one dimensional affine source. We repeat this t times as long as $rt \leq k$ to obtain the decomposition stated in the theorem statement. □

We now define the notion of a "pure" quadratic source: a quadratic source that does not contain a large dimensional affine subsource. Our analysis will take different paths based on whether the source is pure or not.

Definition 3.8 (Pure Quadratic Source)**.** *A quadratic source $X = f(U_s)$ is q-pure if there exists no affine subspace $V \subseteq \mathbb{F}_2^s$ such that $f|_V$ is affine and $\dim(f(V)) \geq q$.*

We note that the above notion of *purity* depends on both the source X and the generating map f. This definition of pure quadratic gives us a natural corollary of Theorem 3.7.

Corollary 3.9 (Pure Quadratics can be decomposed into high derivative entropy sources)**.** *Let $X = f(U_k)$ be a t-pure quadratic source with $H_\infty(X) = k$. Then, X can be expressed as a convex combination of sources of the form $X' = g(U_{k'})$ satisfying (1) $\min_{a \neq 0} H_\infty(\partial_a g(.)) \geq r$ and (2) $k' \geq k - rt$.*

3.5 Warmup: Extractor for Pure Quadratics

Recall that our goal was to analyze the construction $\text{Ext}(X) = \text{LExt}(X, X_{[n^c]})$ for quadratic sources with min-entropy greater than $n - n^c$ for small constant c.

In Section 3.3, we showed via *quadratic conditioning* that it suffices to study random variables of the form

$$\text{LExt}(f(Y, Z), q(Z))$$

where $X = f(Y, Z)$ with $Y = U_s$ for $s \geq \Omega(n^{1-c})$ and $H^{\epsilon}_{\infty}(q(Z)) \geq 0.99n^c$. For simplicity, we assume $q(Z) = U_{n^c}$ (the high entropy version can be handled similar to Section 3.2).

Let $f(Y, Z) = f_1(Y) + f_2(Z) + f_3(Y, Z)$ for bilinear f_3. In Lemma 3.6, we established near-uniformity of $\mathrm{LExt}(f(Y, Z), q(Z))$ assuming $\min_{a \neq 0} H_{\infty}(\partial_a f_1(.))$ is large. We are now ready to use our techniques from Section 3.4 to obtain this structure when the source is pure. We use a slightly stronger version of Corollary 3.9.

Lemma 3.10 (Informal, see Corollary 5.2)**.** *Let $X = f(Y, Z)$ be a t-pure quadratic source with $H_{\infty}(X) = k$. Then, for any $rt \leq s/2$, X is exactly expressed as a convex combination of sources of the form*

$$X' = g(Y', Z) = g_1(Y') + g_2(Y', Z) + g_3(Z)$$

such that (1) $\min_{a \neq 0} H_{\infty}(\partial_a g_1(Y')) \geq r$, (2) $H_{\infty}(X') \geq s' + |Z|$, (3) $Y' = U_{s'}$ with $s' \geq s/2$, and (4) g_2 is bilinear. Importantly, the random variable Z is left unchanged by the decomposition.

The above theorem is proved similar to Theorem 3.7, except that we only fix functions of Y in the proof and leave Z unchanged.

Note that the first property is exactly what we wanted for Lemma 3.6. Now choosing r greater than the entropy requirement of LExt (say $r = n^{0.5}$) implies the following theorem.

Theorem 3.11 (Informal)**.** *Let X be an $n^{0.2}$-pure quadratic source with $H_{\infty}(X) \geq n - n^c$ for a small enough constant $c > 0$. Then $\mathrm{Ext}(X) = \mathrm{LExt}(X, X_{[n^c]}) \approx U_m$.*

This argument works well when the source is a pure quadratic (for a formal proof, see Lemma A.5). Unfortunately, there is an issue when the source is not pure (contains a large subspace). Note that for any useful bound on derivative entropy, we need to choose $t \leq s/(2r) = o(n)$ in Lemma 3.10. Thus, when the source is not pure, we are only guaranteed an affine subspace of entropy $o(n)$. However, this entropy is too small for the candidate construction (Ext is only an affine extractor for entropy $0.99n$ or so, see Section 3.2).

If we are happy with a 2-bit disperser (whose output is non-constant), then there is a simple construction: Output $(\mathrm{Ext}(X), \mathrm{AffineExt}(X))$ where $\mathrm{AffineExt}(X)$ is a low entropy affine extractor (say from [Li16]). When the source is pure, the first bit is uniform, and when the source is not pure, the second bit is non-constant.

It turns out that with one more idea, we can actually use this approach to construct a 1-bit disperser (for entropy $n - n^{0.1}$, which already surpasses prior work). We use the Challenge-Response framework from [BKS+10] which has proven very successful in constructing two source and affine dispersers [BRSW06, Sha11, Coh16]. We use this framework to detect if we are in the pure quadratic or affine case with a single sample. We find this application of the Challenge-Response framework quite interesting, as it different from its usual usage. We defer this construction to the appendix (see Theorem A.1).

3.6 Low Entropy Extraction: the CZ Framework and Li's Affine Extractor

We now begin our discussion of our construction for the polylogarithmic entropy regime. Our construction follows the framework introduced by Chattopadhyay and Zuckerman [CZ19] to construct polylogarithmic entropy two source extractors. We refer the reader to the wonderful survey by Chattopadhyay [Cha20] for an exposition that starts from first principles. This framework was applied to affine sources by Li [Li16]. Analogous to the previous sections, we will consider this affine extractor and then see what has to be adapted for quadratic sources.

Overview of CZ construction and (q, t)-NOBF sources At a high level, the CZ framework, including Li's construction, has two steps: (1) reduce the source X into a (q, t)-non oblivious bit fixing (NOBF) source, and (2) apply an explicit extractor for these sources.

To make sense of the above description, we need to define the notion of a (q, t)-NOBF source.

Definition 3.12 ((q,t)-non oblivious bit fixing (NOBF) sources)**.** *A distribution X on $\mathbb{F}_2^n$ is a (q,t)-NOBF source if there exists a set Q of size at most q such that $X_{[n]\setminus Q}$ is t-wise independent.*

Informally, these sources have the following property: outside q adversarial bits, the rest of the distribution is t-wise independent. One of the main technical achievements in [CZ19] was the construction of explicit extractors (with a single output bit) for these sources (also known as *resilient functions*[4]). Subsequently, there were improvements by Li [Li16] (obtaining multiple output bits) and Meka [Mek17] (obtaining almost optimal parameters matching the non-explicit bound of [AL93]).

For our purposes, we use the following theorem.

Theorem 3.13 (Explicit NOBF Extractor [CZ19, Li16])**.** *For any positive constant $\delta \in (0,1)$, there exists an explicit function* $\mathrm{NOBFExt} : \{0,1\}^n \to \{0,1\}^m$ *and constant C such that for any $(n^{1-\delta}, (m\log n)^C)$-NOBF source X, it holds that* $\mathrm{NOBFExt}(X) \approx_{n^{-\Omega(1)}} U_m$.

Reduction to (q,t)-NOBF sources: Given the above NOBF-extractor, all that is left to do is convert the original affine source X into a (q,t)-NOBF source for $q = N^{1-\delta}$ and $t = \log^C N$ for some N that we choose. Li [Li16] uses two components to achieve this: (1) a *strong linear seeded extractor* and (2) an *affine correlation breaker*.

We have already seen the definition of the first object Definition 3.1. We now define the second.

Definition 3.14 (Affine Correlation Breaker)**.** [5] *A function* $\mathrm{AffCB} : \mathbb{F}_2^n \times \mathbb{F}_2^d \times \{0,1\}^a \to \{0,1\}^m$ *is a (t,k,δ)-affine correlation breaker if it holds that for random variables $X, W, W^{[t]}$ and advice strings $\alpha, \{\alpha^i\}_{i\in[t]}$ in $\{0,1\}^a$ satisfying*

1. *$X = f(V)$ is an affine source with $H_\infty(X) \geq k$ and $V = U_k$,*
2. *$W = \ell(V)$, $W^i = \ell^i(V)$ for all $i \in [t]$ with $\ell, \ell^1, \ldots, \ell^t : \mathbb{F}_2^k \to \mathbb{F}_2^d$ are affine maps,*
3. *$W = U_d$,*
4. *$\alpha \neq \alpha_i$ for all $i \in [t]$,*

it holds that

$$(\mathrm{AffCB}(X, W, \alpha) \approx_\delta U_m) \mid \{\mathrm{AffCB}(X, W^i, \alpha^i)\}_{i\in[t]}$$

where U_m is uniform random string independent of $X, W^{[t]}$.

Li [Li16] gives the first explicit construction of this object and directly uses it to reduce affine sources to (q,t)-NOBF sources. His reduction is the following:

1. Compute $W = \{\mathrm{LExt}(X, i)\}_{i\in[N]}$
2. Output $S = \{\mathrm{AffCB}(X, W^i, i)\}_{i\in[N]}$.

Here LExt is a strong linear seeded extractor with seedlength $\log N = O(\log n)$ (we treat the seeds as integers). For an affine source, one can show that all but $q = N^{1-\delta}$ of the (W^i)'s are truly uniform. Now, given any set of t good indices, the definition of AffCB implies that S_T is almost t-wise independent with error δt. Assume δ is sufficiently small, one can argue that S is close to a (q,t)-NOBF source [AGM03]. Li's construction then outputs $\mathrm{NOBFExt}(S)$ and our argument implies that it is an affine extractor.

[4] technically, resilient functions only need to work for $(q, n-q)$-NOBF sources

[5] The standard definition of an affine correlation breaker is more general, we state it this way for ease of exposition.

3.7 Quadratic Correlation Breaking

One immediate issue when extending to quadratic sources $X = f(U_k)$ is that the random variables W^i output by the linear seeded extractor are now quadratic functions of the generating bits. These are not handled by the definition of the affine correlation breaker. This motivates the definition of a *quadratic correlation breaker*.

Definition 3.15 (Quadratic Correlation Breaker, see Definition 7.1 for a more precise version)**.** *A function* $\text{QuadCB} : \mathbb{F}_2^n \times \mathbb{F}_2^d \times \{0,1\}^a \to \{0,1\}^m$ *is a* (t, k, q, δ)*-quadratic correlation breaker if it holds that for random variables* $X, W, W^{[t]}$ *and advice strings* $\alpha, \{\alpha^i\}_{i\in[t]}$ *in* $\{0,1\}^a$ *satisfying*

1. $X = f(V)$ *is a* q*-pure quadratic source with* $H_\infty(X) \geq k$ *and* $V = U_k$*,*
2. $W = p(V)$*,* $W^i = p^i(V)$ *for all* $i \in [t]$ *with* $p, p^1, \ldots, p^t : \mathbb{F}_2^k \to \mathbb{F}_2^d$ *are quadratic maps,*
3. $W = U_d$*,*
4. $\alpha \neq \alpha_i$ *for all* $i \in [t]$*,*

it holds that

$$(\text{QuadCB}(X, W, \alpha) \approx_\delta U_m) \mid \{\text{QuadCB}(X, W^i, \alpha^i)\}_{i\in[t]} \tag{2}$$

where U_m *is uniform random string independent of* $X, W^{[t]}$*.*

Note that our definition of quadratic correlation breaker requires the underlying source to be a pure quadratic source. This will be crucial for our construction. We show that an existing affine correlation breaker from [CL22] also satisfies the quadratic correlation breaking guarantee (see Theorem 7.2). This allows us to execute the win-win analysis: if the source is not-pure, we employ the affine correlation breaker analysis, else we use our new quadratic analysis.

Our proof of Theorem 7.2 uses all the tools we developed in the previous sections: quadratic conditioning, Cauchy-Schwarz and the structure theorem. We refer the reader to Section 7 for the full proof. We now sketch an argument, assuming familiarity with standard advice correlation breakers (see Definition 4.4).

The construction we use, proposed by [CL22] uses a standard advice correlation breaker CB and a strong linear seeded extractor LExt as black-box objects. We define $\text{QuadCB}(X, W, \alpha)$ in the following way:

1. $S_1 = W_{[d_0]}$
2. $R = \text{LExt}(X, S_1)$
3. Output $S = \text{CB}(W, R, \alpha)$.

Here $d_0 \ll d$ is the seedlength of LExt and the output length of LExt matches the seedlength of CB.

We now sketch our quadratic analysis that works for pure quadratics. Let $S_1^{[t]}$, $R^{[t]}$ and $S^{[t]}$ be the tampered versions. Our goal is to argue that $S \approx U_m \mid S^{[t]}$. In the first step, we argue that R is uniform, even conditioned on S_1. To do, we first apply quadratic conditioning (from Lemma 3.3) with respect to $\{W, W^{[t]}\}$) and reduce to analyzing $X = f(Y, Z)$ and $W = p(Z), W^{[t]} = p^{[t]}(Z)$ (Y, Z are still uniform, but over fewer bits). Now, we also have that $S_1, S_1^{[t]}$ are solely functions of Z.

We now use the fact that X is a pure quadratic. We apply the structure theorem Lemma 3.10 and the Cauchy Schwarz argument Lemma 3.6 to conclude that R is uniform (even conditioned on Z, S_1).

Now, our goal is to make the sources of the form required by the final advice correlation breaker. For this, we need to make $W, W^{[t]}$ independent of $R, R^{[t]}$. Our goal is to do this while maintaing the uniformity of R (as CB requires a uniform seed). It turns out the analogous step in affine case is relatively straightforward. For us, achieving this independence needs us to again use the input reduction lemma (now with respect to R) and carefully fix information so that R remains uniform and W's entropy is not fully destroyed (this is important

as CB needs a high entropy source). We skip this argument here and refer to the proof of Lemma 7.3 (we use a stronger version of the input reduction lemma, see Lemma 7.4).

Finally, after $R, R^{[t]}$ and $W, W^{[t]}$ are made independent, R is uniform, and W has entropy, we can appeal to the guarantee of the correlation breaker CB to complete the proof.

3.8 Good Linear Seeded Extractors for Quadratic Sources: Rank Combiners

Recall that Li's construction first computed $W = \{\text{LExt}(X, i)\}_{i\in[N]}$ and then passed these through the affine correlation breaker as $S = \{\text{AffCB}(X, W^i, i)\}_{i\in[n]}$ to obtain the final (q, t)-NOBF source. Now, the natural extension for us would be to use the quadratic correlation breaker that we constructed above in the second step. This seems reasonable as each W^i is now a quadratic polynomial of the generating bits of the source (as they are linear functions of the quadratic source) and QuadCB was defined precisely to handle such quadratically correlated seeds.

However, there is a subtle issue. Note that the good entries of W are ϵ-close to uniform, where ϵ is at most the error of LExt. Passing this through the correlation breaker implies that roughly, the good bits (corresponding to near uniform entries in W) are almost t-wise independent with error ϵt. The issue is that the result of [AGM03] needs ϵ to be close to N^{-t} for completing the proof and showing that S is close to a (q, t)-NOBF source. However, $N \geq 1/\epsilon^2$ is the minimum seed length required for LExt and so such small error is not achievable anymore. This issue was completely avoided in the affine case, as one can show that the near-uniform bits of W obtained from an affine source are in fact exactly uniform (as linear functions are either constant or have zero-bias).

This is no longer the case for quadratics, as now we can have almost any bias for a non constant quadratic function (take inner product on $2b$ bits to get bias 2^{-b}). We resolve this issue by using linear seeded extractors with unusually good error on the good seeds (when applied to quadratics with high derivative entropy).

Theorem 3.16 (See Theorem 8.1 for formal statement)**.** *For any $k, d \geq \log^C n$, and $\epsilon \geq 1/n^2$, there exists an explicit function* $\text{RCombExt} : \mathbb{F}_2^n \times [N] \to \mathbb{F}_2^{d^{\Omega(1)}}$ *with $N = \text{poly}(n)$ such that* RCombExt *is a $(\epsilon, 2^{(-d^{\Omega(1)})})$-strong linear seeded extractor for the class of quadratic sources $X = f(U_k)$ satisfying $\min_{a\neq 0} H_\infty(\partial_a f(\cdot)) \geq d$.* RCombExt *is also an $(\epsilon, 0)$ strong linear seeded extactor for affine sources with entropy at least k.*

We remind the reader that an (ϵ, δ) strong linear seeded extractor is one where at most ϵ fraction of the seeds have error greater than δ. To construct this object, we introduce a new pseudorandom object that we named a *rank combiner*.

Definition 3.17 (Rank Combiner, see Definition 8.2 for a more general definition)**.** *A (d, r, ϵ)-rank combiner over a field $\mathbb{F}$ is a collection of vectors $w_1, \ldots, w_N$ in $\mathbb{F}^n$ such that for all square matrices $G_1, \ldots, G_n$ satisfying $\min_{y\neq 0} \dim\{G_i y\}_{i\in[n]} \geq d$, it holds with probability $1 - \epsilon$ over $t \sim [N]$ that*

$$\text{rank}_{\mathbb{F}}\left(\sum_{i=1}^{n} (w_t)_i G_i\right) \geq r \tag{3}$$

Note that property $\min_{y\neq 0} \dim\{G_i y\}_{i\in[n]} \geq d$ is just a matrix reformulation of the derivative entropy property[6]. Thus, constructing rank combiners over $\mathbb{F}_2$ imply Theorem 3.16[7]. ChatGPT Astra gave us such a construction (Theorem 8.11).

The notion of a rank combiner seems interesting in its own right. It adds to the rich body of work in algebraic pseudorandomness, studying objects such as rank extractors [GR08, DGW09], rank condensers [FS12, FSS14], subspace designs [GX13, GK13] and the various connections between them [FG15].

[6]Another reformulation is that the vector space of matrices $\mathcal{G} = \{[G_1 y, \ldots, G_n y]\}_y$ is a rank-metric code with distance d.

[7]The definition in Equation (3) corresponds to one output bit. For a version corresponding to multiple output bits, see Definition 8.2.

Rank combiners do not seem reducible to any of these existing objects. However, Astra's proof uses various ideas originating from this line of work. We present the proof and more details on connections to prior work in Section 8.

3.9 The Extractor

Finally, we are ready to present our explicit quadratic extractor. Given a source X, our extractor $\mathrm{QExt}(X)$ does the following:

1. $W = \{\mathrm{RCombExt}(X, i)\}_{i\in[N]}$
2. $S = \{\mathrm{QuadCB}(X, W^i, i)\}_{i\in[N]}$
3. Output $\mathrm{NOBFExt}(S)$.

Our analysis splits into two cases based on if the source is pure or not. In the pure case, we first use Theorem 3.7 to decompose our source into sources having high derivative entropy. We then use the new linear seeded extractor (Theorem 3.16), our quadratic correlation breaker (Theorem 7.2) and the NOBF extractor from Theorem 3.13 to complete the proof. When the source is not pure, we use the affine analysis and the fact that all the quadratic objects we constructed are easily shown to also work for affine sources. See Theorem 9.4 for the proof.

4 Preliminaries

This section contains useful definitions, results from prior work, and some useful elementary lemmas.

4.1 Definitions

We first define the notion of samplable sources.

Definition 4.1 (Samplable Source)**.** *Let $\mathcal{F}$ be a class of boolean functions mapping s bits to n bits. A source X is $\mathcal{F}$-samplable if $X = f(U_s)$ for some $f \in \mathcal{F}$. f is called the generating map and the inputs bits to f are called the generating bits. We use the shorthand $\mathcal{F}$-source.*

We now define *Quadratic sources*, which are the main focus of this paper.

Definition 4.2 (Quadratic sources)**.** *Quadratic sources are the class of $\mathcal{F}$-sources for $\mathcal{F}$ being the class of maps computed by quadratic polynomials over $\mathbb{F}_2$.*

A d-local map is function $f : \mathbb{F}_2^s \to \mathbb{F}_2^n$ such that f_i is a d-junta for each $i \in [n]$. A d-local source is a source whose generating map is a d-local function. NC^0 sources are the set of all d-local sources for constant d. AC^0 sources are the set of all sources with the generating map being computed by a low depth polynomial size circuit.

Definition 4.3. *Let $X = f(U_s)$ be a samplable source. A source Y is a subsource of X, denoted by $Y \subset X$, if there exists a subset $S \subseteq \{0,1\}^s$ such that $Y = f(U_S)$. Similar to purity Definition 3.8, this definition also depends on the generating map f.*

Our notion of subsource is non-standard but slightly more convenient ([BKS+10] defines subsources as obtained by conditioning the distribution X on some event, any subsources from their definition is also a subsource in ours). Throughout this paper, when we use the term *subsource*, we refer to the above notion.

We now define advice correlation breakers.

Definition 4.4 (Advice correlation breaker)**.** *A function* $\mathrm{CB} : \{0,1\}^n \times \{0,1\}^d \times \{0,1\}^a \to \{0,1\}^m$ *is a* (t, k, ϵ)*-advice correlation breaker if for random variables* $X, X^{[t]}, Y, Y^{[t]}$ *and advice strings* $\alpha, \{\alpha^i\}_{i\in[t]}$ *in* $\{0,1\}^a$ *satisfying*

1. $H_\infty(X) \geq k$,
2. $Y = U_d$,
3. $X, X^{[t]}$ *is independent of* $Y, Y^{[t]}$,
4. $\alpha \neq \alpha^i$ *for all* $i \in [t]$,

it holds that

$$(\mathrm{CB}(X, Y, \alpha) \approx_\epsilon U_m) \mid \{\mathrm{CB}(X^i, Y^i, \alpha^i)\}_{i\in[t]} \tag{4}$$

where U_m *is uniform random string independent of* $X, X^{[t]}, Y, Y^{[t]}$.

Definition 4.5 (t-wise independence)**.** *A distribution* X *on* $\mathbb{F}_2^n$ *is* t*-wise independent if* $X_S = U_{|S|}$ *for all sets* S *satisfying* $|S| \leq t$.

4.2 Results from Prior Work

We need a linear seeded extractor due to [CGL21], which has good seed length when the output is small.

Theorem 4.6 ([CGL21])**.** *For any* n, ϵ, m*, there exists an explicit* $(m + 2\log(1/\epsilon), \epsilon)$ *strong seeded extractor* $\mathrm{CGLExt} : \mathbb{F}_2^n \times \mathbb{F}_2^d \to \mathbb{F}_2^m$ *with* $d = 2m + 8\log(n/\epsilon) + O(1)$*. In particular for* $m = O(\log(n/\epsilon))$*, the seedlength is* $O(\log(n/\epsilon))$.

We also need a linear seeded extractor from [Li16], which has good seed length when the error is large.

Theorem 4.7 ([Li16])**.** *There is a constant* $c > 1$ *such that for all* n, k *such that* $c\log^8 n \leq k \leq n$*, and* $\epsilon \geq n^{-2}$*, there exists an explicit* (k, ϵ) *strong linear seeded extractor* $\mathrm{LExt} : \mathbb{F}_2^n \times \mathbb{F}_2^d \to \mathbb{F}_2^m$ *with* $d = O(\log n)$ *and* $m = \sqrt{k}$.

The following lemmas measure closeness to uniform and t-wise independent distributions respectively.

Fact 4.8 ([Vaz86])**.** *Let* X *be a random variable over* $\mathbb{F}_2^n$*. Suppose* $|\mathbb{E}[\chi_S(X)]| \leq \epsilon$ *for all non-empty* $S \subseteq [n]$*. Then,* X *is* $\epsilon \cdot 2^{n/2}$*-close to uniform.*

Fact 4.9 ([AGM03])**.** *Let* X *be a random variable over* $\mathbb{F}_2^n$*. Suppose* $|\mathbb{E}[\chi_S(X)]| \leq \epsilon$ *for all non-empty* $S \subseteq [n]$ *satisfying* $|S| \leq t$*. Then,* X *is* $2n^t\epsilon$*-close to* t*-wise independent.*

We need the following affine extractor due to Li [Li16].

Theorem 4.10 ([Li23, Theorem 1.7])**.** *There exists a constant* $C > 0$ *and function* $\mathrm{AffineExt} : \{0,1\}^n \to \{0,1\}^1$ *such that for all affine sources* X *with* $H_\infty(X) \geq C\log n$*, it holds that* $d_{\mathrm{TV}}(\mathrm{AffineExt}(X), U_1) \leq 0.1$.

We also need the following claim, which shows that a strong linear seeded extractor with overall error ϵ automatically achieves closeness error 0 and strength error 2ϵ on affine sources.

Lemma 4.11 ([Rao09])**.** *Let* $\mathrm{LExt} : \mathbb{F}_2^n \times \mathbb{F}_2^d \to \mathbb{F}_2^m$ *be a* (k, ϵ) *strong linear seeded extractor with* $\epsilon \in [0, 1]$*. Then, for any affine source* X *with entropy* k*, it holds that*

$$\Pr_{s\in\mathbb{F}_2^d}[d_{\mathrm{TV}}(\mathrm{LExt}(X, s), U_m) = 0] \geq 1 - 2\epsilon.$$

That is, LExt *is a* $(k, \epsilon, 0)$*-strong linear seeded extractor for affine sources.*

The following is a corollary of Theorem 3.5.

Corollary 4.12. *Given any quadratic polynomials $q_1, \ldots, q_t$ from $\mathbb{F}_2^n \to \{0,1\}$, there exists a universal constant C and a function* CTPartition $: \mathbb{F}_2^n \to \mathbb{F}_2^{n-d}$ *for $d = C(n/t) - t$ such that for all $c \in \mathbb{F}_2^{n-d}$, it holds that $q_1, \ldots, q_t$ is constant on the set* CTPartition$^{-1}(c)$ *and this set is an affine subspace.*

Proof. Index the subspaces in the partition given by Theorem 3.5 in arbitrarily. Note that there are exactly 2^{n-d} such (affine) subspaces. Now, the function CTPartition just outputs the index of the subspace that a given input belongs to. □

We also use the following strong t-advice correlation breaker from [Li23].

Theorem 4.13 (Explicit advice correlation breaker [Li23])**.** *For any positive integer t, there exists an explicit strong (t, k, ϵ)-advice correlation breaker* CB $: \{0,1\}^n \times \{0,1\}^d \times \{0,1\}^a \to \{0,1\}^m$ *for $k = O(tm + t\log(n/\epsilon) \cdot \frac{\log a}{\log\log a} + t^2\log(n/\epsilon))$, with $d = O(tm + t\log(n/\epsilon) \cdot \frac{\log a}{\log\log a} + t\log^3(t)\log(n/\epsilon))$.*

The following input reduction lemma is due to [CGG24].

Lemma 4.14 (Input reduction: entropy version)**.** *Let $X = f(U_s)$ be a quadratic source such that $H_\infty(X) \geq k$. For any $\epsilon > 0$, there exists a full rank linear map $L : \mathbb{F}_2^s \to \mathbb{F}_2^{s-\ell}$ with $\ell \leq k + O(\log(1/\epsilon))$ such that*

$$\Pr_{b \sim L(U_s)}[H_\infty^\epsilon(f(U_s) \mid L(U_s) = b) \geq k - 1] \geq 1 - \epsilon.$$

In particular, X is ϵ-close to a convex combination of quadratic sub-sources X' with $H_\infty^\epsilon(X') \geq k - 1$ and at most $k + O(\log(1/\epsilon))$ generating bits.

Proof. We borrow notation from the proof of Lemma 2.4 in [CGG24]. Let $\ell = k + O(\log(1/\epsilon))$. Their proof shows that there exists a full rank linear map $L : \mathbb{F}_2^s \to \mathbb{F}_2^{s-\ell}$ and a function $S : \mathbb{F}_2^n \to \mathbb{F}_2^{k+1}$ such that (1) $H_\infty(S(X)) \geq k-1$ and (2) with probability at least $1-\epsilon$ over $b \sim L(U_s)$, the source $\tilde{X} = (f(U_s) \mid L(U_s) = b)$ satisfies

$$d_{\mathrm{TV}}(S(\tilde{X}), S(X)) \leq \epsilon.$$

From the fact that L is full rank and has codimension ℓ, we have that $\tilde{X}$ is a quadratic source with at most ℓ generating bits.

Now, all that remains to prove the lemma is to show that $\tilde{X}$ is ϵ close to a source X' with $H_\infty(X') \geq k - 1$. We now construct such a X'. Define $\tilde{X}_s$ as the distribution as the conditional distribution $(\tilde{X} \mid S(\tilde{X}) = s)$ (for $s \notin \mathrm{supp}(S(\tilde{X}))$, define $\tilde{X}_s$ as an arbitrary distribution over $S^{-1}(s)$). Define X' as follows: (1) sample $s \sim S(X)$, and (2) sample $X' \sim \tilde{X}_s$. From data processing inequality, we have that

$$d_{\mathrm{TV}}(\tilde{X}, X') \leq d_{\mathrm{TV}}(S(\tilde{X}), S(X)) \leq \epsilon.$$

Now, for any string x, we have $\Pr_{X'}[X' = x] \leq \Pr_{X'}[S(X') = S(x)] = \Pr_X[S(X) = S(x)] \leq 2^{-k+1}$. This proves that $H_\infty(X') \geq k - 1$. □

4.3 Additional Lemmas

We recall that a source X has ϵ-smooth min entropy k, denoted by $H_\infty^\epsilon(X) \geq k$ if there exists some source $\tilde{X}$ with $H_\infty(\tilde{X}) \geq k$ such that $X \approx_\epsilon \tilde{X}$. We will need the following smooth min-entropy chain rule. We include the proof here for completeness.

Lemma 4.15. *Let (X, Y) be random variables such that $H_\infty^\epsilon(X) \geq k$ and $|\mathrm{supp}(Y)| \leq M$. Then, it holds with probability at least $1 - (\epsilon + 2\delta)$ over $y \sim Y$ that*

$$H_\infty^{2\epsilon/\delta}(X \mid Y = y) \geq k - \log M - \log(1/\delta).$$

Proof. Let Y_x be the conditional distribution of Y conditioned on $X = x$. Let $\tilde{X}$ be the distribution such that $H_\infty(\tilde{X}) \geq k$ and $d_{\mathrm{TV}}(X, \tilde{X}) \leq \epsilon$. Let $\tilde{Y}$ be the random variable obtained by first sampling $x \sim \tilde{X}$ and then sampling $\tilde{Y} \sim Y_x$ (define Y_x arbitrarily over $\mathrm{supp}(Y)$ for $x \notin \mathrm{supp}(X)$).

From min-entropy chain rule, we have that with probability at least $1 - \delta$ over $y \sim \tilde{Y}$, it holds that $H_\infty(\tilde{X} \mid \tilde{Y} = y) \geq k - \log M - \log(1/\delta)$, By data processing inequality, we also have that $d_{\mathrm{TV}}((X, Y), (\tilde{X}, \tilde{Y})) \leq \epsilon$. This implies that

$$\mathop{\mathbb{E}}_{y \sim Y} d_{\mathrm{TV}}((X \mid Y = y), (\tilde{X} \mid \tilde{Y} = y)) \leq d_{\mathrm{TV}}((X, Y), (\tilde{X}, \tilde{Y})) + d_{\mathrm{TV}}(Y, \tilde{Y}) \leq 2\epsilon.$$

From Markov's inequality, we have that with probability at least $1-\delta$ over $y \sim Y$, we have $d_{\mathrm{TV}}((X \mid Y = y), (\tilde{X} \mid \tilde{Y} = y)) \leq 2\epsilon/\delta$. Using the fact that $d_{\mathrm{TV}}(Y, \tilde{Y}) \leq \epsilon$ and $\Pr_{y \sim \tilde{Y}}[H_\infty(\tilde{X} \mid \tilde{Y} = y) < k - \log M - \log(1/\delta)] \leq \delta$, we have that with probability $1 - (\epsilon + 2\delta)$ over $y \sim Y$,

$$H_\infty^{2\epsilon/\delta}(X \mid Y = y) \geq k - \log M - \log(1/\delta).$$

□

Lemma 4.16. *Let X be a distribution on $\mathbb{F}_2^n$ such that $H_\infty^\epsilon(X) \geq n - \Delta$. Then, $H_\infty^\epsilon(X_{[d]}) \geq d - \Delta$.*

Proof. By definition of smooth min-entropy, there is some $\tilde{X} \approx_\epsilon X$ with $H_\infty(\tilde{X}) \geq n - \Delta$. By [CGG24, Claim 4.11], $H_\infty(\tilde{X}_{[d]}) \geq d - \Delta$. By the data-processing inequality, $X_{[d]} \approx_\epsilon \tilde{X}_{[d]}$, which completes the proof. □

The following lemma shows that for affine sources, smooth min-entropy is close to the true min-entropy.

Lemma 4.17. *Let X be a flat source satisfying $H_\infty^{1/2}(X) \geq k$. Then, it holds that $H_\infty(X) \geq k - 1$*

Proof. Let $S = \mathrm{supp}(X)$ and f be the indicator function on this set. Let $\tilde{X}$ be the distribution satisfying $H_\infty(\tilde{X}) \geq k$ and $\tilde{X} \approx_{1/2} X$ guaranteed by the premise of the lemma. We have that $\mathbb{E}[f(\tilde{X})] \geq 1/2$ (from the TV distance guarantee) and $\mathbb{E}[f(\tilde{X})] \leq |S|/2^k$ (from the min-entropy guarantee). Thus, we have that $|S| \geq 2^{k-1}$ which implies $H_\infty(X) \geq k - 1$. □

The next lemma is useful in the proof of our structure theorem Theorem 5.1.

Lemma 4.18 (Weak-injectivity preservation)**.** *Let $X = f(U_{k+\ell})$ be a random variable satisfying $H_\infty^\epsilon(X) \geq k$ for $f : \mathbb{F}_2^{k+\ell} \to \mathbb{F}_2^n$. Let $\mathcal{S}$ be a distribution over subsets of $\mathbb{F}_2^{k+\ell}$ such that (1) $\mathrm{supp}(\mathcal{S})$ forms a partition of $\mathbb{F}_2^{k+\ell}$ and (2) for every set $s \in \mathrm{supp}(\mathcal{S})$, it holds that $\Pr_{S \sim \mathcal{S}}[S = s] = |s|/2^{k+\ell}$. Then, with probability $1 - \delta$ over $S \sim \mathcal{S}$, it holds that*

$$H_\infty^{\epsilon/\delta}(f(U_S)) \geq \log(|S|) - \ell.$$

Proof. Let $A = \{x \mid |f^{-1}(x)| \geq 2^\ell, x \in \mathbb{F}_2^n\}$. By the smooth min-entropy guarantee, we have that

$$|f^{-1}(A)|/2^{k+\ell} = \Pr[X \in A] \leq |A|/2^k + \epsilon.$$

This implies that $|f^{-1}(A)| - |A|2^\ell \leq \epsilon 2^{k+\ell}$. For each $x \in A$, choose exactly 2^ℓ entries in $f^{-1}(x)$ arbitrarily. Let B be the elements in $f^{-1}(A)$ that are not chosen. We have that $|B| \leq \epsilon 2^{k+\ell}$. Define a function g that agrees with f on $\mathbb{F}_2^{k+\ell} \setminus B$ and assigns the points in B to $\mathbb{F}_2^n$ in such a way that $|g^{-1}(x)| \leq 2^\ell$ for all $x \in \mathbb{F}_2^n$.

Now, it is easy to see that $H_\infty(g(U_S)) \geq \log(|S|) - \ell$ for all sets S. Also, using the fact that $\Pr[U_{k+\ell} \in B] = \mathbb{E}_{S \sim \mathcal{S}}[\frac{|S \cap B|}{|S|}] \leq \epsilon$, we have that with probability at most $1 - \delta$ over $S \sim \mathcal{S}$, it holds that $\frac{|S \cap B|}{|S|} \leq \epsilon/\delta$. Now, since f and g agree outside B, we have $g(U_S) \approx_{\epsilon/\delta} f(U_S)$. Combining this with the fact that $H_\infty(g(U_S)) \geq \log(|S|) - \ell$ completes the proof. □

We use the following lemma that states that pure-quadratics remain pure after restricting to subsources formed by affine subspaces of the generating bits.

Lemma 4.19. *Let $X = f(U_s)$ be a q-pure quadratic source. Then $X' = (f \circ L)(U_r)$ is also q-pure for any injective affine map L.*

Proof. Let $V \subseteq \mathbb{F}_2^r$ be an affine subspace such that $(f \circ L)|_V$ is affine with $\dim((f \circ L)(V)) \geq q$. This implies that the purity of f is violated on the subspace $L(V)$. □

5 Structure Theorem

We now show a structure theorem for quadratic sources.

5.1 Injective Sources

We present the proof of Theorem 3.7 here for completeness.

Proof. We will decompose the source by iteratively restricting into affine subspaces of the generating bits. We will execute the decomposition in t stages. We maintain the following invariant: at the end of stage j, we will ensure that our original source will be exactly decomposed into a convex combination of subsources such that

1. each subsource W has the form $W = h(Y, Z) = g(Y) + \sum_{i=1}^{j'} b_i Z_i +$ where $Y = U_{k'}$, $Z = U_{j'}$ (for $k' \geq k - rj'$ and $j' \leq j$) are uniform and independent,
2. $j' \neq j$ implies that $H_\infty(\partial_a g(Y)) \geq r$ for all $a \neq 0$, and
3. h is injective.

We now describe our iterative decomposition procedure. Note that the invariant trivially holds at the start ($j = 0$). We now describe stage $j + 1$, assuming that the invariant holds after stage j.

Consider any subsource $W = g(Y) + \sum_{i=1}^{j'} b_i Z_i$ obtained after stage i. First, we leave the subsource unchanged if it holds that $H_\infty(\partial_a g(Y)) \geq r$ for all $a \neq 0$. If the previous property does not hold, then it holds from item (2) of the invariant that $j' = j$. Let $a \neq 0$ be a direction along which $H_\infty(\partial_a g(Y)) < r$. We will further decompose W by conditioning on the value of $\partial_a g(Y)$. Consider the subsource obtained from W by conditioning on the event $\{\partial_a g(Y) = b\}$.

Let $A = c + B \subseteq \mathbb{F}_2^{k'}$ be the (affine) subspace such that $\partial_a g(y) = b$ for $y \in A$ where B is a subspace and c is a fixed vector. We know that $\dim(A) \geq |Y| - (r - 1) = k - rj - (r - 1)$ as $\dim(\partial_a g(\cdot)) < r$. Observe that $a + A = A$ which implies $a \in B$. Starting with a, construct an arbitrary basis $\{a, e_1, \ldots, e_{\dim(B)-1}\}$ of B. Let $C = c + \mathrm{span}(\{e_i\}_{i \in [\dim(B)]-1})$. Let Y' be a uniform string from A. We have that $Y' = Za + U_C$ where Z is a uniform bit and U_C is independent of Z and uniform over C.

The subsource W' obtained from W by conditioning on the derivative is exactly $W' = g(Y') + \sum_{i=1}^{j'} b_i Z_i$. Note that $g(y' + a) = g(y') + b$ for all $y' \in A$. Hence, we have that $W' = g(U_C) + \sum_{i=1}^{j} b_i Z_i + Zb$. We have that $\dim(C) = \dim(A) - 1 \geq k - r(j + 1)$. Thus $E = L(U_{k''})$ for some full rank affine map L and $k'' \geq k - r(j + 1)$.

Defining $Z_{j+1} = Z$, $b_{j+1} = b$ and $g' = g \circ L$ shows that W' has the form $g'(U_{k''}) + \sum_{i=1}^{j+1} b_i Z_i$. Note that the injectivity of the map is preserved as we started out with an injective map and Y' is uniform over a subset of inputs. Thus, W' satisfies the invariant after stage $j + 1$.

We will execute the procedure for $j = t$ stages. Consider any subsource W in the decomposition obtained after t stages. Recall $W = h(Y, Z) = g(Y) + \sum_{i=1}^{j'} Z_i b_i$. There are two cases. In the first case, we have that $j' = t$. In this case, we can further fix the random variable Y to obtain an affine source $W' = \sum_{i=1}^{t} Z_i b_i + d$ where d is a constant shift. Since h was injective, this subsource must have entropy t. In the second case, we have that $H_\infty(\partial_a g(Y)) \geq r$ for all $a \neq 0$. In this case, we further fix all the variables Z to obtain a source of the second type in the theorem statement. □

5.2 General Quadratic Sources

We now extend Theorem 3.7 to general quadratic sources with a smooth min-entropy guarantee with the number of generating bits slightly larger than the entropy. Our decomposition procedure only acts on a subset of the generating bits that we choose.

Theorem 5.1. *Let $X = f(Y, Z)$ be a quadratic source with $H_\infty^{\epsilon^3}(X) \geq k$ and generating map $f : \mathbb{F}_2^{k+\ell} \to \mathbb{F}_2^n$ where $Y = U_s$ and $Z = U_{k+\ell-s}$. Let r, t be positive integers such that $rt \leq s$. Then X is ϵ close to a convex combination of subsources of the form $f(U_V)$ for affine subspace $V \subseteq \mathbb{F}_2^{k+\ell}$ such that the subsources are one of the following two types:*

1. *Affine sources W with min-entropy at least $t - \ell - 1$ ($f|_V$ is affine) or*
2. *Quadratic sources $W(Y', Z) = g_1(Y') + g_2(Y', Z) + g_3(Z)$ such that (1) $Y' = U_{s'}$ for $s' \geq s - rt$, (2) g_2 is bilinear, (3) $H_\infty(\partial_a g_1(Y')) \geq r$ for all $a \neq 0$ and (4) $H_\infty^{\epsilon^2}(W(Y', Z)) \geq s' + |Z| - \ell$.*

Furthermore, Z is the same seed random variable as that in the source X and is independent of Y'.

Proof. As in the proof of Theorem 3.7, we will iteratively decompose the source into sub-sources. A crucial difference is that we will only condition on functions of the input seed Y, leaving Z unchanged. We maintain the following invariant.

After j stages, we ensure that all the component subsources W in the convex combination satisfy the following:

1. $W = h(Y' = (A, B), Z) = g_1(B) + \sum_{i=1}^{j'} A_i \cdot (b_i(Z)) + g_2(B, Z) + g_3(Z)$ where $b_i(Z)$ are affine functions of Z for all $i \in [j']$ and g_2 is bilinear (here $j' \leq j$).
2. $A = U_{j'}$ and $B = U_{s'}$ for $s' \geq s - rj'$. A and B are independent of each other and Z.
3. Furthermore $j' \neq j$ implies that $H_\infty(\partial_a g_1(.)) \geq r$ for all $a \neq 0$.

Observe that the invariant trivially holds for $j = 0$ (with $|A| = 0$). We will now describe stage $j + 1$. Assume the invariant holds upto stage j. Take a component W obtained after j stages (satisfying the above invariant). We check if all the non-zero derivative entropies of g_1 are greater than equal to r. If so, we make no changes to W (this component is frozen in future iterations as well). Alternatively, if there is a direction a with low derivative entropy, we choose the lexicographically smallest such direction and fix the value of this derivative.

Suppose we condition on the event $\{\partial_a g_1(B) = b\}$. Observe that $\partial_a g_2(B, Z) = g_2(a, Z)$ which is a function of Z. Thus, after fixing the derivative, we can retrieve one "affine" direction (if one treats Z as a constant). That is, using the same argument as the proof of Theorem 3.7, the new subsource can be expressed as $W' = h(Y' = (A, D), Z, B') = g_1(B') + \sum_{i=1}^{j'} A_i \cdot (b_i(Z)) + D(g_2(a, Z) + b) + g_2(B', Z) + g_3(Z)$ where D is a fresh random bit, and $B' + Da = B \mid \{\partial_a g_1(B) = b\}$. Rewrite B' as $L(U_{s''})$ for a full rank affine function with $s'' \geq s' - r \geq s - r(j' + 1)$. Now, we observe that W' satisfies the invariant.

We run the decomposition for $j = t$ stages. Let the subsource obtained after the t stages be $W = h(Y' = (A, B), Z) = g_1(B) + \sum_{i=1}^{j'} A_i \cdot (b_i(Z)) + g_2(B, Z) + g_3(Z)$.

There are again two cases. In the first case, we have $j' = t$. In this case, we fix B, Z to obtain an affine source. In the second case, we have $H_\infty(\partial_a g_1(Y')) \geq r$ for all $a \neq 0$. In this case, we further fix all the variables A to obtain a source of the form described in Item 2 of the theorem statement.

We now argue that the entropy guarantees stated in both the cases are satisfied (recall that this was immediate in the injective case). To do this, we use Lemma 4.18. First, observe that all the subsources obtained by this decomposition are of the form $f(U_V)$ for some affine subspace $V \subseteq \mathbb{F}_2^{k+\ell}$. Also, the (affine) subspaces

corresponding to different subsources form a partition on $\mathbb{F}_2^{k+\ell}$. In the branches of the decomposition leading to affine subspaces (Item 1), we have $\dim(V) = t$ and in the other case, we have $\dim(V) = s' + |Z|$. Let $\mathcal{V}$ be a random variables that chooses these affine subspaces according to the above decomposition. From Lemma 4.18, we have that with probability $1 - \epsilon$ over $V \sim \mathcal{V}$, it holds that (1) $H_\infty^{\epsilon^2}(f(U_V)) \geq t - \ell$ if V came from decompositions in the first case, or (2) $H_\infty^{\epsilon^2}(f(U_V)) \geq s' + |Z| - \ell$ in the second case. Applying Lemma 4.17 to the first case implies the entropy guarantees in the theorem statement. □

We now state an immediate corollary of the above theorem that gives a sufficient condition for decomposing quadratic sources into sub-sources with high derivative entropy.

Corollary 5.2 (Pure Quadratics can be decomposed into high derivative entropy sources)**.** *Let $X = f(Y, Z)$ be a quadratic source with $H_\infty^{\epsilon^3}(X) \geq k$ and generating map $f : \mathbb{F}_2^{k+\ell} \to \mathbb{F}_2^n$ where $Y = U_s$, $Z = U_{k+\ell-s}$. Let r, t be such that $rt \leq s/2$ and suppose X is $(t - \ell - 1)$-pure. Then, X is ϵ close to a convex combination of subsources of the form*

$$X' = g(Y', Z) = g_1(Y') + g_2(Y', Z) + g_3(Z)$$

such that (1) $H_\infty(\partial_a g_1(Y')) \geq r$ for all $a \neq 0$, (2) $H_\infty^{\epsilon^2}(X') \geq s' + |Z| - \ell$, (3) $Y' = U_{s'}$ with $s' \geq s/2$, and (4) g_2 is bilinear. Again, the random variable Z is left unchanged by the decomposition and X' is obtained by restricting f to an affine subspace of the generating bits.

Proof. The proof is immediate from Theorem 5.1 as Item 1 in Theorem 5.1 is not possible anymore. □

6 Exploiting High Derivative Entropy

The below argument shows that a linear seeded extractor extracts from a high derivative entropy quadratic source with a quadratically correlated seed.

Lemma 6.1. *Let $X = f(Y, Z)$ be a source with f quadratic, $Y = U_s$, $Z \in \mathbb{F}_2^t$ independent of Y and such that $f(Y, Z) = f_1(Y) + f_2(Y, Z) + f_3(Z)$ where f_2 is bilinear and $H_\infty(\partial_a f_1(Y)) \geq k$ for all $a \neq 0$. Let $\mathrm{LExt} : \mathbb{F}_2^n \times \mathbb{F}_2^d \to \mathbb{F}_2^m$ be a (k, ϵ) strong linear seeded extractor. Then, for any random variable W (correlated with Z, independent of Y) satisfying $W \approx_\delta U_d$, it holds with probability $1 - 2^m \cdot (2^{-s} + 2\epsilon + \delta)^{1/2}$ over $(z, w) \sim (Z, W)$ that*

$$d_{\mathrm{TV}}(\mathrm{LExt}(f(Y, z), w), U_m) \leq 2^{m/2}(2^{-s} + 2\epsilon + \delta)^{1/4}.$$

Furthermore, the theorem also holds when $W \approx_\delta U_d$ is replaced by $H_\infty^\delta(W) \geq d - \Delta$ with $2\epsilon + \delta$ replaced by $2^{\Delta+1}\epsilon + \delta$ in both the probability and TV distance bounds.

Proof. We will first argue that for any $S \subseteq [m]$, it holds high probability over $(z, w) \sim (Z, W)$ that $|\mathbb{E}[\chi_S(\mathrm{LExt}(f(Y, z), w))]| \leq \epsilon$. The theorem then follows by a union bound over all subsets S and Fact 4.8.

Let $b(w, z) = |\mathbb{E}_Y[\chi_S(\mathrm{LExt}(f(Y, z), w))]|$. We have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{W,Z}[(b(W, Z))^2] &= \mathop{\mathbb{E}}_{W,Z}(\mathop{\mathbb{E}}_{Y}[\chi_S(\mathrm{LExt}(f(Y, Z), W))])^2 \\
&= \mathop{\mathbb{E}}_{W,Z}\mathop{\mathbb{E}}_{Y,A}[\chi_S(\mathrm{LExt}(f(Y, Z) + f(Y + A, Z), W))] \\
&= \mathop{\mathbb{E}}_{W,Z}\mathop{\mathbb{E}}_{Y,A}[\chi_S(\mathrm{LExt}(\partial_A f_1(Y) + f_2(A, Z), W)).
\end{aligned}$$

In the second inequality A is uniform and independent of Y. Now for every $a \neq 0$, we have $H_\infty(\partial_a f_1(Y)) \geq k$ and $\partial_a f_1(Y)$ is an affine source. From Lemma 4.11, we have that with probability at least $1 - 2\epsilon - \delta$ over

uniform string $w \sim W$, it holds that $d_{\mathrm{TV}}(\mathrm{LExt}(\partial_a f_1(Y), w), U_m) = 0$. Thus, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{W,Z}[(b(W,Z))^2] &= \mathop{\mathbb{E}}_{W,Z}\mathop{\mathbb{E}}_{Y,A}[\chi_S(\mathrm{LExt}(\partial_A f_1(Y) + f_2(A,Z), W))] \\
&\leq \Pr[A=0] + \max_{a\neq 0}\mathop{\mathbb{E}}_{Y,W,Z}[\chi_S(\mathrm{LExt}(\partial_a f_1(Y) + f_2(a,Z), W))] \\
&\leq 2^{-s} + \max_{a\neq 0}\mathop{\mathbb{E}}_{W,Z}|\mathop{\mathbb{E}}_{Y}[\chi_S(\mathrm{LExt}(\partial_a f_1(Y), W))]| \\
&\leq 2^{-s} + 2\epsilon + \delta.
\end{aligned}$$

Now, by Markov's inequality, we have that $b(w,z) \leq (2^{-s} + 2\epsilon + \delta)^{1/4}$ with probability $1 - (2^{-s} + 2\epsilon + \delta)^{1/2}$. The theorem statement now follows from a union bound and Fact 4.8. For the guarantee assuming high smooth min-entropy, the probability of sampling a bad seed is at most $2^{\Delta+1}\epsilon + \delta$ and the rest of the proof is the same.

□

7 Quadratic Correlation Breaking

In this section we construct a new type of correlation breaker that can handle quadratic sources and seeds. This can be viewed as a generalization of an affine correlation breaker. The construction we use is from [CL22], where they show how to upgrade a black box advice correlation breaker into an affine correlation breaker. We will show that this construction is also a *quadratic correlation breaker* when the source is a pure-quadratic.

We first define the notion of a quadratic correlation breaker.

Definition 7.1 (Quadratic Correlation Breaker)**.** *A function* $\mathrm{QuadCB} : \mathbb{F}_2^n \times \mathbb{F}_2^d \times \{0,1\}^a \to \{0,1\}^m$ *is a* (t,k,q,δ)*-quadratic correlation breaker if there exists an absolute constant* $c > 0$ *such that for random variables* $X, W, W^{[t]}$ *and advice strings* $\alpha, \{\alpha^i\}_{i\in[t]}$ *in* $\{0,1\}^a$ *satisfying*

1. $X = f(V)$ *is a* q*-pure quadratic source with* $H_\infty(X) \geq k$ *and* $V = U_s$ *where* $s \leq k + c\log(1/\delta)$*,*
2. $W = p(V)$*,* $W^i = p^i(V)$ *for all* $i \in [t]$ *with* $p, p^1, \ldots, p^t : \mathbb{F}_2^s \to \mathbb{F}_2^d$ *are quadratic maps,*
3. $W = U_d$*,*
4. $\alpha \neq \alpha_i$ *for all* $i \in [t]$*,*

it holds that

$$(\mathrm{QuadCB}(X, W, \alpha) \approx_\delta U_m) \mid \{\mathrm{QuadCB}(X, W^i, \alpha^i)\}_{i\in[t]} \tag{5}$$

where U_m *is uniform random string independent of* $X, W^{[t]}$*.*

We now describe the construction of our quadratic correlation breaker (as mentioned before, the construction already appears in [CL22]).

We now state the main theorem of this section.

Theorem 7.2. *For any* $\delta \in (0,1)$ *and positive integers* t, a, m*, there exists a* $(t,k,q,O(\delta))$*-quadratic correlation breaker* $\mathrm{QuadCB} : \mathbb{F}_2^n \times \mathbb{F}_2^d \times \mathbb{F}_2^a \to \mathbb{F}_2^m$ *(from Algorithm 1) whenever*

- $d \geq Ct^5(m + \log a \log(n/\delta))^2$*,*
- $k \geq Ct^2d^4$*,*
- $q \leq c\sqrt{k/(td^2)}$*,*

Algorithm 1 Quadratic Correlation Breaker (QuadCB)

Input: $x \in \{0,1\}^n$— an n-bit string, $w \in \{0,1\}^d$— the seed, $\alpha \in \{0,1\}^a$— the advice string.
Output: $s \in \{0,1\}^m$— output string.

Sub-Routines and Parameters:
Let $d_1 = O(t(m + t \log a \log(n/\delta)))$, $d_0 = O(d_1 + \log(1/\delta))$, $d = O(t^5(m + \log a \log(n/\delta))^2)$.
Let $\mathrm{LExt} : \{0,1\}^n \times \{0,1\}^{d_0} \to \{0,1\}^{d_1}$ be a $(k, 2^{-4d_1}\delta^{32})$ strong linear seeded from Theorem 4.6.
Let $\mathrm{CB} : \{0,1\}^d \times \{0,1\}^{d_1} \times \{0,1\}^a \to \{0,1\}^m$ be a $(t, d/2, \delta/2)$-advice correlation breaker from Theorem 4.13.

1. Let $s_1 = w_{[d_0]}$.
2. Let $r = \mathrm{LExt}(x, s_1)$.
3. Let $s = \mathrm{CB}(w, r, \alpha)$

for absolute constant $c, C > 0$. Moreover, the guarantee in Equation (4) also holds when the entropy assumption in Item 1 and the uniformity assumption in Item 3 of Definition 7.1 are relaxed to $H_\infty^{\delta^4}(X) \geq k$ and $W \approx_{\delta^2} U_d$ respectively.

Before proving Theorem 7.2, we establish an intermediate lemma that proves that the correlation breaker works when the source is decomposed in a specific form.

Lemma 7.3. *Let δ be the error parameter in Algorithm 1. Let $X = f(Y, Z)$ be a source with quadratic f, $Y = U_s$, Z independent of Y and such that $f(Y, Z) = f_1(Y) + f_2(Y, Z) + f_3(Z)$ where f_2 is bilinear and $H_\infty(\partial_a f_1(Y)) \geq k$ for all $a \neq 0$. Also let $W, W^{[t]}$ be functions of Z satisfying $W = U_d$. Also, let $\alpha, \alpha^{[t]}$ be such that $\alpha \neq \alpha_i$ for all $i \in [t]$. Suppose $s, k \geq C(d_1 + \log(1/\delta))$ for a large constant C. Then, it holds that*

$$(\mathrm{QuadCB}(X, W, \alpha) \approx_{O(\delta)} U_m) \mid \{\mathrm{QuadCB}(X, W^i, \alpha^i)\}_{i \in [t]}.$$

We need a stronger version of the input reduction lemma to prove the above lemma.

Lemma 7.4 (Input Reduction: Partial Input)**.** *Let $f : \mathbb{F}_2^{s+t} \to \mathbb{F}_2^n$ be a quadatic map. Furthermore, suppose it holds that $d_{\mathrm{TV}}(f(X, Y), f(X, y)) \leq \delta$ with probability at-least $1 - \delta$ over $y \sim Y$ where $X = U_s$ and Y is independent of X. For any $0 < \epsilon < 1/4$, there exists a full rank linear map $L : \mathbb{F}_2^s \to \mathbb{F}_2^{s-\ell}$ with $\ell = n + O(\log(1/\epsilon))$ such that*

$$(f(X, Y), Y, L(X)) \approx_{\epsilon + 2\delta} (f(X, Y), Y', U_{s-\ell})$$

where U, X, Y, Y' are independent and Y, Y' are identically distributed.

Proof. We choose L to be a random linear map. By triangle inequality, we have that for any L, it holds that

$$\begin{aligned} d_{\mathrm{TV}}((f(X, Y), Y, L(X)), (f(X, Y), Y', U_{s-\ell})) \leq\ & d_{\mathrm{TV}}((f(X, Y), Y, L(X)), (f(X, Y), Y, U_{s-\ell})) \\ & + d_{\mathrm{TV}}((f(X, Y), Y, U_{s-\ell}), (f(X, Y), Y', U_{s-\ell})) \end{aligned}$$

We bound the first term using Lemma 3.4. That is for every fixed $y \sim Y$, we can apply Lemma 3.4. By averaging over $y \sim Y$ and L and then using Markov's inequality, we have that $d_{\mathrm{TV}}((f(X, Y), Y, L(X)), (f(X, Y), Y, U_{s-\ell})) \leq \epsilon/2$ with probability at least $1 - \epsilon$ over L. The second term is at most 2δ from the premise of the lemma that $d_{\mathrm{TV}}(f(X, Y), f(X, y)) \leq \delta$ with probability $1 - \delta$ over $y \sim Y'$. □

We are now ready to prove Lemma 7.3.

Proof of Lemma 7.3. We use S_1, R, S to denote the intermediate random variables in Algorithm 1 when applied on (X, W, α). Similarly, we define S_1^i, R^i to denote the intermediate random variables in the algorithm when applied to (X, W^i, α^i). Let $\delta_1 = 2^{d_1}(2^{-s} + 3\delta^{32} \cdot 2^{-4d_1})^{1/4} \leq O(\delta^8)$. From Lemma 6.1 and the (1) our premise regarding s, and (2) the choice of d_1, LExt in Algorithm 1, we have that with probability at least $1 - \delta_1$ over $(s_1, z) \sim (S_1, Z)$, it holds that

$$R \approx_{\delta_1} U_{d_1}. \tag{6}$$

Note that $R, R^{[t]}$ depend on both Y and Z. In the next step of the proof, we will show that we can fix additional information to make $R, R^{[t]}$ independent of W, while retaining the uniformity of R. Once we have this, we can use the definition of the advice correlation breaker (Definition 4.4) to complete the proof.

We first fix $(S_1, S_1^{[t]})$. From the above argument, we have that with probability $1 - \sqrt{\delta_1}$ over $(S_1, S_1^{[t]})$, it holds that $(R, Z) \approx_{\sqrt{\delta_1}} (U_{d_1}, Z)$. Furthermore, $H_\infty(W \mid S_1 = s_1, S_1^{[t]} = s_1^{[t]}) \geq d - td_0 - O(\log(1/\delta)) \geq d/2$ (from the choice of d, d_0 in Algorithm 1) with probability $1 - O(\sqrt{\delta_1})$ over $(s_1, s_1^{[t]}) \sim (S_1, S_1^{[t]})$ from Lemma 4.15. Thus these two events happen with probability at least $1 - O(\sqrt{\delta_1})$ over $(S_1, S_1^{[t]}) = (s_1, s_1^{[t]})$. We have $R = L(f(Y, Z))$ for some linear function L (which depends on s_1). Similarly, we have $R^i = L^i(f(Y, Z))$ where L^i is determined by s_1^i. Note that Z is not uniform anymore.

From Lemma 7.4, we have a full rank linear map $M : \mathbb{F}_2^s \to \mathbb{F}_2^{s-\ell}$ with $\ell = d_1 + O(\log(1/\delta))$ such that

$$(R = L(f(Y, Z), Z, M(Y)) \approx_{O(\sqrt{\delta_1})} (R = L(f(Y, Z)), Z', U_{s-\ell}) \tag{7}$$

for U, Y, Z, Z' independent and Z, Z' identically distributed.

Let $e_1, \ldots, e_\ell$ be a basis for the nullspace of M and $e_{\ell+1}, \ldots, e_s$ be a completion of the basis to all of $\mathbb{F}_2^s$. We have that $Y = A + B$ for independent A, B where (1) A is uniform over the nullspace of M, and B is uniform over the span of $\{e_{\ell+1}, \ldots, e_s\}$. Overloading notation, we have $A = A(P)$ and $B = B(Q)$ for linear maps A and B of appropriate dimension and independent uniform seeds $P = U_\ell$ and $Q = U_{s-\ell}$.

Using the fact that A, B are linear and overloading notation, we henceforth represent the source as $X = f(P, Q, Z)$. Now, rewriting Equation (7), we have

$$(R, Z, M(B(Q))) \approx_{\delta_2} (R, Z', U_{s-\ell}) \tag{8}$$

where $\delta_2 = O(\sqrt{\delta_1})$. This implies that R is close to uniform even conditioned on Q and Z. Now, we will argue that we can fix Q and a few bits of information from Z to make $(R, R^{[t]})$ independent of $(W, W^{[t]})$ while retaining the uniformity of R. Let $R = g(P, Q, Z)$ and $R^i = g^i(P, Q, Z)$ for quadratic maps $g, g^{[t]}$ each having output length d_1.

We have that

$$g(P, Q, Z) = g_1(P, Q) + g_2(Z) + g_3(P, Z) + g_4(Q, Z).$$

Similarly, we have

$$g^i(P, Q, Z) = g_1^i(P, Q) + g_2^i(Z) + g_3^i(P, Z) + g_4^i(Q, Z).$$

After fixing Q, we have that that with probability at least $1 - \sqrt{\delta_2}$ over $q \sim Q$, it holds that

$$(g(P, q, Z), Z) \approx_{\sqrt{\delta_2}} (g(P, q, Z), Z').$$

Recall that $Z, W, W^{[t]}$ are independent of P, Q and hence fixing Q does not affect them. Now, $g, g^{[t]}$ are only functions of P, Z. Observe that $\{g_3, g_3^1, \ldots, g_3^t\}$ each depend on at most $d_1\ell$ one-bit linear functions of Z as P is uniform on ℓ bits. Let $(c(Z), c^{[t]}(Z))$ store these linear function in $(t+1)d_1\ell$ bits. Thus, fixing $(c(Z), c^{[t]}(Z), g_2(Z), g_2^{[t]}(Z), g_4(Z), g_4^{[t]}(Z))$ makes $R, R^{[t]}$ solely depend on P while retaining uniformity of R (for the latter guarantee, we use Equation (8)).

We now fix $C(Z) = (c(Z), c^{[t]}(Z), g_2(Z), g_2^{[t]}(Z), g_4(Z), g_4^{[t]}(Z))$. With probability $1 - O(\delta_2^{1/4})$ over $c \sim C(Z)$, we have that

$$\begin{aligned} H_\infty(W \mid (S_1, S_1^{[t]}) = (s_1, s_1^{[t]}), C(Z) = c) &\geq d - td_0 - 3(t+1)d_1\ell - O(\log(1/\delta)) \\ &\geq t^5(m + \log a \log(n/\delta))^2. \end{aligned}$$

Also, from Equation (8), we have that with probability $1 - \delta_2^{1/4}$ over $c \sim C(Z)$, it holds that

$$R \approx_{\delta_2^{1/4}} U_{d_1}.$$

Also, we now have that $(R, R^{[t]})$ is independent of $(W, W^{[t]})$ as the former is only a function of P, Q and the latter is a function only of Z (which is independent of P, Q). We are now ready to invoke the guarantee of the advice correlation breaker. From Theorem 4.13, we have that

$$(CB(W, R, \alpha) \approx_{\delta + O(\delta_2^{1/4})}, U_m) \mid \{CB(W^i, R^i, \alpha^i)\}_{i \in [t]}.$$

Recall that $\delta_2 = O(\sqrt{\delta_1})$ where $\delta_1 = O(\delta^8)$. Thus, we have that the final error $O(\delta)$. □

We are now ready to complete the proof of Theorem 7.2.

Proof of Theorem 7.2. We prove the theorem by first fixing information to make the random variables $W, W^{[t]}$ depend only on a subset of the bits generating the source. We then use the q-purity of X and Theorem 5.1 to decompose the original source into subsources with large derivative entropy. Finally, we apply Lemma 7.3.

Quadratic Conditioning To implement the first step (which we call "quadratic conditioning"), we use Lemma 3.4. We have a full rank linear map $L : \mathbb{F}_2^{k+\ell} \to \mathbb{F}_2^{k+\ell-h}$ with $h = d + O(\log(1/\delta))$ such that

$$(p(V), L(V)) \approx_{\delta^2} (p(V), U_{k+\ell-h}). \tag{9}$$

where $U_{k+\ell-p}$ is independent of V.

Let $e_1, \ldots, e_h$ be a basis for the nullspace of L and $e_{h+1}, \ldots, e_{k+\ell}$ be a completion of the basis to all of $\mathbb{F}_2^{k+\ell}$. We have that $V = A + B$ for independent A, B where (1) A is uniform over the nullspace of L, and B is uniform over the span of $\{e_{h+1}, \ldots, e_{k+\ell}\}$. Overloading notation, we have $A = A(Z)$ and $B = B(Y)$ for linear maps A and B of appropriate dimension and independent uniform seeds $Z = U_h$ and $Y = U_{k+\ell-h}$. We can represent our source as $X = f(M_1(Y) + M_2(Z))$ for linear maps M_1, M_2 of appropriate dimensions. Overloading notation, we can rewrite the polynomials $p, p^{[t]}$ as being quadratic functions of Y, Z. We have

$$p(Y, Z) = p_1(Y) + p_2(Z) + p_3(Y, Z)$$

and similarly for each of the polynomials in $p^{[t]}$ (say $p^i(Y, Z) = p_1^i(Y) + p_2^i(Z) + p_3^i(Y, Z)$ for each $i \in [t]$).

From Equation (9), we also have that with probability $1 - \delta$ over $y \sim Y$, it holds that $p(y, Z) \approx_\delta U_d$. We will now convert $p, p^{[t]}$ to functions solely depending on Z by fixing all dependencies on Y. Each of $p_3, p_3^{[t]}$ depend on dh linear functions of Y. We will fix these dht linear functions along with $p_1(Y), p_1^{[t]}(Y)$. Recall that $h = d + O(\log(1/\delta))$. We thus need to fix at most $b = O(dt(d + \log(1/\delta)))$ quadratic functions of Y.

Let $\text{CTPartition} : \mathbb{F}_2^{k+\ell-h} \to \mathbb{F}_2^{k+\ell-h-s}$ be the function given by Theorem 3.5 (when the input functions are given by the b quadratic functions we just considered) where $s = C(k + \ell - h)/b - b \geq Ck/(td^2)$ (for some constant C) and $\text{CTPartition}^{-1}(c)$ is an affine subspace for all c. We fix $c \sim \text{CTPartition}(Y)$. We have that with probability $1 - \delta$ over $c \sim \text{CTPartition}(Y)$, it holds that

1. $Y' = (Y \mid \text{CTPartition}(Y) = c)$ is uniform over an affine subspace of dimension s,

2. $W, W^{[t]}$ are independent of Y',

3. $W \approx_{O(\delta)} U_d$ and

4. $H_\infty^{\delta^3}(f(M_1(Y') + M_2(Z)) \geq k - (k + \ell - h - s) - O(\log(1/\delta)) \geq C(k/(td^2)) - td^2 + h - O(\log(1/\delta))$ for constant C (by Lemma 4.15 and the bound on s).

Observe that $f(M_1(Y') + M_2(Z))$ is a subsource of X of the form $f(U_V)$ for affine subspace V. Thus, by Lemma 4.19, this subsource is still q-pure.

We overload notation and redefine X, f, Y such that the source we have arrived at is $X = f(Y, Z)$ where $Y = U_s, Z = U_h$. We have that (1) $W, W^{[t]}$ are functions only of Z, (2) $H_\infty^{\delta^3}(X) \geq s + h - O(\log(1/\delta))$ and (3) X us q-pure. We are now ready to move to the next step where we invoke our structure theorem.

Structure Theorem We now call Corollary 5.2. From the end of the previous paragraph we have that X satisfies the premise of Corollary 5.2 (as it is q-pure for $q \leq c\sqrt{k/(td^2)}$) Thus, we have that X is 2δ-close to a convex combination of subsources of the form

$$X'(Y', Z) = g_1(Y') + g_2(Y', Z) + g_3(Z)$$

such that (1) $H_\infty(\partial_a g_1(Y')) = \Omega(\sqrt{s}) = \Omega(d\sqrt{t})$ for all $a \neq 0$, (2) $H_\infty^{2\delta^2}(X') = \Omega(td^2)$, (3) $Y' = U_{s'}$ with $s' \geq s/2$, and (4) g_2 is bilinear.

Importantly, this decomposition left Z and hence $W, W^{[t]}$ unaffected. We are now ready to finish the proof.

Quadratic Correlation Breaking Our source satisfies the premise of Lemma 7.3. Applying Lemma 7.3 implies Equation (5) (we absorb the fact that $W \approx_{O(\delta)} U_d$ into the error). □

8 Rank combiners

The main goal of this section is to construct a linear seeded extractor for quadratic sources that has unusually low error on the good seeds. Note that for affine sources this is immediate, via Lemma 4.11 (as linear functions are either constant or have zero bias). For quadratics, this is no longer immediate, as one can have non-constant quadratic polynomials achieving nearly any bias (consider for example inner product on $2b$ variables which has bias 2^{-b}).

Having unusually good error is important in the proof of the extractor as we want our good seeds to have error much smaller than the inverse of the total number of seeds. This is not possible for general sources as linear seeded extractors that work for arbitrary sources have seed length at least $\Omega(\log(1/\epsilon))$ (see Theorem 5 in [AGO+19]).

In this section, we construct such extractors for quadratic sources with high derivative entropy (with the error going to zero as the derivative entropy increases). The extractor we construct is also a linear seeded extractor for affine sources. This suffices for our applications as we can always reduce to one of these two cases. After we defined the notion of a *rank combiner* and proved that it implies such extractors, ChatGPT Astra gave us an explicit construction (Theorem 8.11) which implies the following.

Theorem 8.1 (Good linear seeded extractors for quadratics over $\mathbb{F}_2$)**.** *For any $d \geq \log^{40} n$, $\epsilon \geq 1/n^2$ and $m \leq d^{1/10}$, there exists an explicit function* RCombExt $: \mathbb{F}_2^n \times [N] \to \mathbb{F}_2^m$ *with $N = \text{poly}(n)$ such that*

1. RCombExt *is a $(\epsilon, 2^{(m-r)/2})$-strong linear seeded extractor with $r \geq (d^{1/10}/\log^{3/2} n)$ for the class of quadratic sources $X = f(U_k)$ satisfying $H_\infty(\partial_a f(\cdot)) \geq d$ for all non-zero a.*

2. RCombExt *is a* $(\epsilon, 0)$*-strong linear seeded extractor for affine sources with entropy at least* d*.*

To construct this object, we introduce a new pseudorandom object that we call a *rank combiner*.

Definition 8.2. *A* (d, r, ϵ)*-rank combiner over a field* $\mathbb{F}$ *is a collection of matrices* $W_1, \ldots, W_N$ *in* $\mathbb{F}^{m \times n}$ *such that for all square matrices* $G_1, \ldots, G_n$ *satisfying* $\dim\{G_i y\}_{i \in [n]} \geq d$ *for all* $y \neq 0$*, it holds with probability* $1 - \epsilon$ *over* $t \sim [N]$ *that*

$$\operatorname{rank}_{\mathbb{F}} \left(\sum_{i=1}^{n} (c^\top W_t)_i G_i \right) \geq r \tag{10}$$

for all non-zero $c \in \mathbb{F}^m$*. We call* m *the output length and* N *the size.*

We now argue why the above definition implies a linear seeded extractor with unusually low error for all good seeds, when applied to a source with high derivative entropy.

Lemma 8.3 (Rank combiners are good linear seeded extractors for quadratics)**.** *Let* $W_1, \ldots, W_N \in \mathbb{F}_2^{m \times n}$ *be a* (d, r, ϵ) *rank combiner. Then, the function* $L : \mathbb{F}_2^n \times [N] \to \mathbb{F}_2^m$ *defined as* $L(x, i) = W_i x$ *satisfies the following properties:*

1. L *is an* $(\epsilon, 2^{(m-r)/2})$*-strong linear seeded extractor for the class of quadratic sources* $X = f(U_k)$ *satisfying* $H_\infty(\partial_a f(\cdot)) \geq d$ *for all non-zero* a*.*
2. L *is an* $(\epsilon, 0)$*-strong linear seeded extractor for affine sources with entropy at least* d*.*

Proof. We first prove (1). Let $f(x) = (f_1(x), \ldots, f_n(x))$ for quadratic functions $f_1, \ldots f_n$. Let $f_i = x^\top M_i x + b_i^\top x$ for matrix $M \in \mathbb{F}_2^{k \times k}$ and vector $b \in \mathbb{F}_2^k$. For any non-zero a, the source $\partial_a f(U)$ is the affine source over the affine subspace $c_a + S_a$ for $c_a = ((M_i^\top a + b_i))^\top a)_{i \in [n]}$ and $\dim(S_a) = \dim(\operatorname{span}\{(M_i + M_i^\top)a\}_{i \in [n]})$. The entropy guarantee on $\partial_a f(.)$ translates to the condition $\dim(S_a) \geq d$. Define $G_1, \ldots, G_n$ as the matrices $G_i := M_i + M_i^\top$. Thus, we have that $\dim\{G_i a\}_{i \in [n]} \geq d$ for all non-zero a. Thus, these matrices satisfy the premise of the rank combiner.

The guarantee of the rank combiner gives us a set $T \subseteq [N]$ of size $(1 - \epsilon)N$ such that for all $t \in T$, it holds that

$$\operatorname{rank}_{\mathbb{F}_2} (\sum_{i=1}^{n} (c^\top W_t)_i G_i) \geq r$$

for all non-zero c. We will argue that for all $t \in T$, holds that

$$L(f(U), t) \approx_{2^{(m-r)/2}} U_m.$$

Establishing the above inequality would complete the proof of (1). Let $g_t(x) = L(f(x), t)$ be the quadratic map which we are analyzing. We will argue that for all non-zero $c \in \mathbb{F}_2^m$, it holds that $c^\top g_t(U_k)$ has bias at most $2^{-r/2}$. The final claim then follows from Vazirani's XOR lemma Fact 4.8. We have that

$$c^\top g_t(x) = x^\top (\sum_{i=1}^{n} (c^\top W_t)_i M_i) x + (\sum_{i=1}^{n} (c^\top W_t)_i b_i^\top) x$$

We have that

$$(\mathop{\mathbb{E}}_{U} [(-1)^{c^\top g_t(U)}])^2 \leq \mathop{\mathbb{E}}_{U,a} [(-1)^{\partial_a (c^\top g_t(U))}]$$

where a is uniform in $\mathbb{F}_2^k$. Observe that $\partial_a (c^\top g_t(U))$ is a linear function that is either constant or has bias zero. Thus, it suffices to bound the probability over a that $\partial_a (c^\top g_t(.))$ is a constant. It is easy to see that

this happens if and only if $(\sum_{i=1}^{n}(c^\top W_t)_i G_i)a = 0$. From our rank combiner guarantee, this happens with probability 2^{-r} over $a \sim \mathbb{F}_2^k$. Thus, we have that

$$|\mathop{\mathbb{E}}_{U}[(-1)^{c^\top g_t(U)}]| \leq 2^{-r/2}.$$

This completes the proof of (1).

We now prove the affine case (2). Let $X = f(U)$ satisfy the condition required by (1). For this step, a $(d, 1, \epsilon)$ rank combiner suffices. An affine source X of entropy $k \geq d$ can be be defined as the output of an affine map acting on k uniform bits, that is $X = AU + b$ for uniform U for $A \in \mathbb{F}_2^{n\times k}$ with $\mathrm{rank}(A) = k$. To prove the lemma statement, it suffices to show that with probability $1 - \epsilon$ over $i \sim [N]$, it holds that the random variable $W_i AU$ is uniform.

Let a_i be the i-th row of A. We also use a_i to refer to the corresponding element of $\mathbb{F}_{2^k}$ when the vector is viewed as a field element. Let M_i be the matrix such that $M_i x = a_i x$ when x is appropriately interpreted as a vector and field element respectively. We have $\mathrm{rank}_{\mathbb{F}_2}(A) = \dim_{\mathbb{F}_2}\{a_i\}_{i\in[n]} = k$. Thus, for any non-zero x, we have $\dim_{\mathbb{F}_2}\{M_i x\}_{i\in[n]} = k \geq d$. Thus, $\{M_i\}_{i\in[n]}$ satisfies the premise of the rank combiner. Thus, with probability at least $1 - \epsilon$ over $i \sim [N]$, it holds that $\mathrm{rank}(\sum_{j=1}^{n}(c^\top W_i)_j M_j) \geq 1$ for all $c \in \mathbb{F}_2^m$. Since M_i's are denoting multiplication by field elements, we have that with probability $1 - \epsilon$ over $i \sim [N]$, it holds that

$$\sum_{j=1}^{n}(c^\top W_i)_j a_j \neq 0$$

for all non-zero $c \in \mathbb{F}_2^m$. This proves that $c^\top W_i X = c^\top W_i AU + c^\top W_i Ab$ is uniform for all non-zero c. Hence we have that L is a (d, ϵ) strong linear seeded extractor for affine sources over $\mathbb{F}_2$. □

Like many pseudorandom objects, the existence of great rank combiners is easy to show via the probabilistic method.

Theorem 8.4 (Probabilistic existence of rank combiners)**.** *There exists a $(d, d - m - \log(1/\epsilon) - 1, \epsilon)$-rank combiner over $\mathbb{F}_2$ with output length m and size* $\mathrm{poly}(nk^2/\epsilon)$.

We provide a proof in the appendix (Appendix C).

8.1 Rank combiners over large finite fields

Our explicit construction of rank combiners proceeds in two steps. First, we give a construction over large fields of polynomial size. Then, we give a field size reduction argument using a linear seeded extractor in Section 8.2.

The proof of the main theorem of this section uses two additional lemmas (Lemmas 8.6 and 8.7) which we prove immediately afterward.

Theorem 8.5. *For any finite field $\mathbb{F}$ with $|\mathbb{F}| \geq 16mn^2/\epsilon$ and $m \leq \sqrt{d}/2$, there exists an explicit $(d, \sqrt{d}/(2m), \epsilon)$-rank combiner over $\mathbb{F}$ with output length m and size $|\mathbb{F}|$.*

Proof. We will first describe the construction before proving correctness. We will describe a matrix $W(\beta) \in \mathbb{F}^{m\times n}$ for $\beta \in \mathbb{F}$. The final rank combiner is $\{W(\beta)\}_{\beta\in\mathbb{F}}$.

Choose an arbitrary set of distinct non-zero elements $\lambda_1, \ldots \lambda_n \in \mathbb{F}$. Let $w_1, \ldots, w_n$ be positive integers forming a Sidon set with $\max_i\{w_i\} \leq 8n^2$ (see [ET41] for a construction). We define the entries of $W(\beta)$ as $(W(\beta))_{ji} = \lambda_i^{j-1}\beta^{w_i}$. We will now prove that that with probability at least $1 - \epsilon$ over $\beta \sim \mathbb{F}$ we have

$$\mathrm{rank}_{\mathbb{F}}\left(\sum_{i=1}^{n}(c^\top W(\beta))_i G_i\right) \geq \sqrt{d}/(2m) \tag{11}$$

for all non-zero $c \in \mathbb{F}^m$. We say that $\beta \in \mathbb{F}$ is *bad* if $W(\beta)$ does not satisfy Equation (11) for some non-zero $c \in \mathbb{F}^m$. Let B be the set of bad field elements. We want to show that $|B| \leq \epsilon|\mathbb{F}|$.

We now describe our high level plan. For any $\beta \in B$, there exists a non-zero vector $c_\beta \in \mathbb{F}^m$ such that Equation (11) is not satisfied by $W(\beta)$. First we will construct a low degree polynomial map $c : \mathbb{F} \to \mathbb{F}^m$ such that $c(\beta) \in \text{span}\{c_\beta\}$ for all $\beta \in \mathbb{F}$ and $c(0) \neq 0$. This will give us a matrix with polynomial entries $M(\beta) = \sum_{i=1}^{n}(c(\beta)^\top W(\beta)))_i G_i$. We then use Lemma 8.6 to isolate a large sub-matrix with non-zero determinant. We then show that any bad β that has low rank within this sub-matrix must have large multiplicity in the determinant polynomial. This allows us to bound $|B|$.

First, we construct a polynomial map c such that $c(\beta) \in \text{span}\{c_\beta\}$ for all $\beta \in B$ and $c(0) \neq 0$. We will show that there exists a polynomial c of degree $\ell = |B|(1 - 1/m)$ that satisfies these properties. Finding such a c is equivalent to solving the following homogeneous system of equations for the coefficients of c in $\mathbb{F}^{m(\ell+1)}$ and $(\alpha_\beta)_{\beta \in B} \in \mathbb{F}^{|B|}$:

$$c(\beta) = \alpha_\beta c_\beta, \text{ for all } \beta \in B.$$

There are $m|B|$ constraints and the number of variables is $m(\ell + 1) + |B|$ and the system is homogeneous. From our choice $\ell = |B|(1 - 1/m)$, system has more variables than constraints and is homogenous. Thus, there exists a non-zero solution. This solution must have $c \neq 0$ as $(\alpha_\beta)_{\beta \in B}$ is forced to be 0 when $c = 0$. We can make $c(0) \neq 0$ by factoring out lowest degree term if our original solution had no constant term as this preserves the rank. Thus, we have obtained the polymial map c that we desired.

Now consider the matrix $M(\beta) = \sum_{i=1}^{n}(c(\beta)^\top W(\beta)))_i G_i$. We have that $\text{rank}(M(\beta)) \leq r$ for all $\beta \in B$ from (1) the fact that $c(\beta) = c_\beta$ or $c(\beta) = 0$ for all $\beta \in B$ and (2) the definition of c_β. Expanding using the definition of $W(\beta)$, we obtain that

$$M(\beta) = \sum_{i=1}^{n} b_i(\beta)\beta^{w_i} G_i$$

where $b_i(\beta) = \sum_{j=1}^{m} c_j(\beta) \cdot \lambda_i^{j-1}$. Since $c(0)$ is non-zero, we have that $b_i(0) = \sum_{j=1}^{m} c_j(0) \cdot \lambda_i^{j-1}$ is only zero for at most m indices as the λ_i's are distinct. Thus, $M(\beta)$ satisfies the premise of Lemma 8.6 with $p = m$. Since $m \leq \sqrt{d}/2$, we have that $M(\beta)$ has a non-zero $\sqrt{d} \times \sqrt{d}$ minor.

We now restrict to the $\sqrt{d} \times \sqrt{d}$ submatrix of $M(\beta)$ that has non-zero determinant. Call this submatrix $P(\beta)$. We will argue that $\text{rank}(P(\beta)) \geq \sqrt{d}/(2m)$ with probability $1 - \epsilon$ over $\beta \sim \mathbb{F}$. For any $\beta \in B$, we have that $\text{rank}(P(\beta)) < \sqrt{d}/(2m)$ and hence from Lemma 8.7, all elements of B have multiplicity at least $\sqrt{d}(1 - 1/(2m))$ in $\det(P(\beta))$. Recall that $\deg(c) = \ell \leq |B|(1 - 1/m)$. This implies that that $\deg(\det(P(\beta))) \leq (\deg(c) + \max_i\{w_i\})\sqrt{d} \leq \sqrt{d}(|B|(1 - 1/m) + 8n^2)$ as the elements of the sidon set are at most $8n^2$. Combining the multiplicity lower bound and the degree upper bound, we obtain $|B| \leq 16n^2m$. Since $|\mathbb{F}| \geq 16mn^2/\epsilon$, we have that the probability of sampling β from B is at most ϵ. This completes the proof. □

We use the following lemma which considers constructions with polynomial entries.

Lemma 8.6. *Let $G_1, \ldots, G_n \in \mathbb{F}^{k \times k}$ such that $\dim_{\mathbb{F}}\{G_i y\} \geq d$ for all non-zero $y \in \mathbb{F}$. Let $w_1, \ldots, w_n$ be a set of positive integers that form a Sidon set. Then, for any vector of polynomials $b(t) \in (\mathbb{F}[t])^n$ with $|\{i \mid b_i(0) \neq 0\}| \geq n - p$, it holds that the matrix*

$$M(t) = \sum_{i=1}^{n} b_i(t)t^{w_i} G_i$$

has a non-zero $r \times r$ minor with $r \geq \min(\sqrt{d}, \frac{d}{2p})$.

Proof. We will argue that the rank of $M(t)$ over the field of rational functions $\mathbb{F}(t)$ is at least r. We show that there every vector $g(t) \in (\mathbb{F}(t))^k$ in its kernel has at least r non-zero entries.

For sake of contradiction assume that a sparse coefficient vector exists. Without loss of generality assume that $g(t) = (g_1(t), \ldots, g_r(t), 0, \ldots 0)$ and that all the g_i's are polynomials. Let $f_1, \ldots f_{\tilde{r}}$ be a basis for $\text{span}\{g_i\}_{i\in[r]}$ such that $\text{mindeg}(f_i) \neq \text{mindeg}(f_j)$ for $i \neq j$ (here $\text{mindeg}(f)$ is the degree of the smallest non-zero monomial in f).

Let $g(t) = \sum_{j=1}^{\tilde{r}} c_j f_j(t)$ where $c_j \in \mathbb{F}^k$. Thus, we have that

$$\sum_{j=1}^{\tilde{r}} \sum_{i=1}^{n} b_i(t) t^{w_i} f_j(t) G_i \cdot c_j = 0. \tag{12}$$

We want to use this equation to arrive at a contradiction. We will show that

$$\dim_{\mathbb{F}}\{c \in \mathbb{F}^{n\tilde{r}} \mid \sum_{i=1}^{n} \sum_{j=1}^{\tilde{r}} c_{ij} b_i(t) t^{w_i} f_j(t) = 0\} \leq pr + \binom{r}{2}. \tag{13}$$

We first show argue how Equation (13) implies a contradiction and then we will prove it. Consider the matrix B in $\mathbb{F}^{k\times n\tilde{r}}$ where the (i,j)-th column is $G_i c_j$ for $i \in [n], j \in [\tilde{r}]$. From Equation (12) and Equation (13), we have that $\text{rank}_{\mathbb{F}}(B) \leq pr + \binom{r}{2}$. Since g is non-zero, there exists some j such that c_j is non zero. The matrix $[G_i c_j]_{i\in[n]}$ is a submatrix of B and hence has rank at most $pr + \binom{r}{2}$ over $\mathbb{F}$. From the premise of the lemma, we have $d \leq (pr + \binom{r}{2}) \leq pr + r^2$. Solving for r, we get $r \geq \min(\sqrt{d}, \frac{d}{2p})$.

We now prove Equation (13). To do this, we show that $\dim_{\mathbb{F}}\{b_i(t)t^{w_i} f_j(t)\}_{i\in[n], j\in[\tilde{r}]} \geq (n-p)\tilde{r} - \binom{\tilde{r}}{2}$. Equation (13) then follows by using the fact that $\tilde{r} \leq r$. Note that the dimension is lower bounded by the number of distinct min-degrees (degree of smallest non-zero monomial) among the $n\tilde{r}$ polynomials, as any set of polynomials with distinct min-degrees are linearly independent. Let I be the set of indices i for which $b_i(0) \neq 0$. The number of distinct min degrees is lower bounded by the terms only involving indices in I. Note that for such indices i, it holds that $\text{mindeg}(b_i(t)t^{w_i} f_j(t)) = w_i + \text{mindeg}(f_j)$ for all $j \in [\tilde{r}]$.

Let $S_j = \{d \mid d = \text{mindeg}(f_j(t)t^{w_i}) \text{ for some } i \in I\}$. Clearly, the distinct min-degree count is lower bounded by $|\cup_{j=1}^{\tilde{r}} S_j|$. Furthermore, for any $i \neq j$, it holds that $|S_i \cap S_j| \leq 1$. To see this, note that for any $a, b \in [n]$, the equation $w_a + \text{mindeg}(f_i) = w_b + \text{mindeg}(f_j)$ fully determines the indices a and b (as w_i's form a Sidon set). Thus, from inclusion-exclusion, we have that

$$|\cup_j S_j| \geq (n-p)\tilde{r} - \binom{r}{2}.$$

This completes the proof of the dimension bound and hence the theorem. □

We need the following standard lemma about the determinant of matrices with polynomial entries. A similar statement and argument appears in [GK13, Claim 16].

Lemma 8.7. *Let $P(t) \in (\mathbb{F}[t])^{k\times k}$ be a matrix with polynomial entries with non-zero determinant. Then any $\beta \in \mathbb{F}$ with $\text{rank}_{\mathbb{F}}(P(\beta)) \leq r$ is a root of $\det(P(t))$ with multiplicity at least $k - r$.*

Proof. Take any β satisfying $\text{rank}(P(\beta)) \leq r$. Let $P_i(t)$ be the i-column of P. We have that $P_i(t) = P_i(\beta) + (t-\beta)q_i(t)$ where q_i is a polynomial, as β is a root for $P_i(t) - P_i(\beta)$. Define the matrices $P_S(t)$ for $S \subseteq [k]$ as the matrix for which $(P_s(t))_i = P_i(\beta)$ if $i \in S$ and $(P_S(t))_i = q_i(t)$ if $i \notin S$. By expanding the determinint we have that $\det(P(t)) = \sum_{S\subseteq[k]} (t-\beta)^{k-|S|} \det(P_S(t))$. Observe that $\det(P_S(t)) = 0$ for all $|S| \geq r+1$ as $P(\beta)$ has rank at most r and $P_S(t)$ contains more than r columns of $P(\beta)$. Thus, the only surviving terms in the expansion of $\det(P(t))$ are divided by $(t-\beta)^{k-r}$ and this completes the proof. □

8.2 Field Size Reduction

The next lemma is used to translate dimension guarantees between $\mathbb{F}_2$ and $\mathbb{F}_{2^h}$.

Lemma 8.8. *Let $G_1, \ldots, G_n \in \mathbb{F}_2^{k \times k}$ such that $\dim_{\mathbb{F}_2}\{G_i y\}_{i\in[n]} \geq d$ for all non-zero $y \in \mathbb{F}_2^k$. Then, for any positive integer h, it holds that $\dim_{\mathbb{F}_{2^h}}\{G_i z\}_{i\in[n]} \geq d/h$ for all non-zero $z \in \mathbb{F}_{2^h}^k$.*

Proof. We will prove the contrapositive. Suppose there is a non-zero $z \in \mathbb{F}_{2^h}$ such that $\dim_{\mathbb{F}_{2^h}}\{G_i z\} < d/h$. Let $e_1, \ldots, e_h$ be a basis of $\mathbb{F}_{2^h}$ over $\mathbb{F}_2$. Let $V = \mathrm{span}_{\mathbb{F}_{2^h}}\{G_i z\}$. First observe that $\dim_{\mathbb{F}_2}(V) = h \cdot \dim_{\mathbb{F}_{2^h}}(V) < d$ (the first equality holds because $\{e_i v_j\}_{i\in[j], j\in[\dim(V)]}$ is a basis of V over $\mathbb{F}_2$ whenever $v_1, \ldots, v_{\dim(V)}$ is a basis of V over $\mathbb{F}_{2^h}$). Let $z = \sum_{i=1}^h c_i e_i$ where $c_i \in \mathbb{F}_2^k$. Since $z \neq 0$, there exists j such that $c_j \neq 0$. Now consider the function π that maps $G_i z$ to $G_i c_j$. Extend this map linearly over all of V. Clearly $\pi(V)$ is a subspace and $\dim_{\mathbb{F}_2}(\pi(V)) \leq \dim_{\mathbb{F}_2}(V) < d$. This implies that $\dim_{\mathbb{F}_2}\{G_i c_j\} < d$. This completes the proof of the contrapositive. □

We are now ready to state and prove our theorem on field size reduction.

Theorem 8.9 (Field size reduction for rank combiners)**.** *Let $U_1, \ldots, U_N \in \mathbb{F}_{2^h}^{m\times n}$ be an explicit (d, r, ϵ)-rank combiner over $\mathbb{F}_{2^h}$ and let $\mathrm{LExt} : \mathbb{F}_2^{hm} \times \{0,1\}^\ell \to \mathbb{F}_2^{m'}$ be an explicit (m, δ)-strong linear seeded extractor. Then, there is an explicit $(dh, r/(hm), \epsilon + \delta)$-rank combiner $W_1, \ldots, W_{M=N\cdot 2^\ell}$ in $\mathbb{F}_2^{m'\times n}$.*

Proof. Let $G_1, \ldots, G_n$ be matrices in $\mathbb{F}_2^{k\times k}$ satisfying $\dim_{\mathbb{F}_2}\{G_i y\}_{i\in[n]} \geq dh$ for all non-zero $y \in \mathbb{F}_2^k$. Let $e_1, \ldots, e_h$ be basis vectors for $\mathbb{F}_{2^h}$ when it is treated as a vector space over $\mathbb{F}_2$. Note that for a fixed seed s, the function computed by $\mathrm{LExt}(\cdot, s)$ can be denoted by a matrix L_s in $\mathbb{F}_2^{m'\times hm}$. For $U \in \mathbb{F}_{2^h}^{m\times n}$, write $U = \sum_{j=1}^h M_j(U) e_j$ with $M_j(U) \in \mathbb{F}_2^{m\times n}$ and define

$$M(U) := \begin{pmatrix} M_1(U) \\ \vdots \\ M_h(U) \end{pmatrix}, \quad W(U, s) := L_s M(U), \quad W(i, s) := W(U_i, s).$$

We will show that $\{W_{(i,s)}\}_{i\in[N], s\in\mathbb{F}_2^\ell}$ is a $(dh, r/(hm), \epsilon+\delta)$-rank combiner over $\mathbb{F}_2$. The key property of the linear seeded extractor we use is that the set of vectors given by its seeds evade small subspaces. From Lemma 8.8, we have that $\mathrm{rank}_{\mathbb{F}_{2^h}}\{G_i z\} \geq d$ for all non-zero $z \in \mathbb{F}_{2^h}^k$. Then, from the definition of a rank combiner, we have that with probability $1-\epsilon$ over $j \sim [N]$, it holds that $\mathrm{rank}_{\mathbb{F}_{2^h}}\left(\sum_{i=1}^n (c^\top U_j)_i G_i\right) \geq r$ for all non-zero $c \in \mathbb{F}_{2^h}^m$. We set $U = U_j$ for U_j that satisfied the previous condition. We use M to denote $M(U)$ as defined above. Let $V \in \mathbb{F}_2^{hm}$ be the span (over $\mathbb{F}_2$) of vectors v satisfying

$$\mathrm{rank}_{\mathbb{F}_2}\left(\sum_{i=1}^n (v^\top M(U))_i G_i\right) < r/(hm). \tag{14}$$

Claim 8.10. $\dim_{\mathbb{F}_2}(V) \leq hm - m$.

Proof. We will argue that $\dim_{\mathbb{F}_2}(V) \leq hm - m$. Take a basis $v_1, \ldots, v_t$ of V such that each v_i satisfies Equation (14). Let $\tilde{V}$ be the subspace over $\mathbb{F}_{2^h}^{hm}$ obtained by taking the span of $v_1, \ldots v_t$ over $\mathbb{F}_{2^h}$. Clearly, $\dim_{\mathbb{F}_2}(V) = \dim_{\mathbb{F}_{2^h}}(\tilde{V})$ as linear independence of the basis vectors is preserved. Thus, it suffices to bound the latter quantity. For any $\tilde{v} \in \tilde{V}$, we have that

$$\mathrm{rank}_{\mathbb{F}_{2^h}}\left(\sum_{i=1}^n (\tilde{v}^\top M)_i G_i\right) \leq \sum_{j=1}^t \mathrm{rank}_{\mathbb{F}_{2^h}}\left(\sum_{i=1}^n (v_j^\top M)_i G_i\right) \leq \sum_{j=1}^t \mathrm{rank}_{\mathbb{F}_2}\left(\sum_{i=1}^n (v_j^\top M)_i G_i\right) < r$$

where the first inequality follows by subadditivity of rank and the second by the fact that rank only increases when looking at a subfield. The final inequality follows from Equation (14) and the fact that $t \leq hm$.

For any $0 \neq c = (c_1, \ldots, c_m) \in \mathbb{F}_{2^h}^m$, its holds that $\text{rank}_{\mathbb{F}_{2^h}} \left(\sum_{i=1}^n (c^\top U)_i G_i\right) \geq r$. Expanding using basis elements, we have that $\text{rank}_{\mathbb{F}_{2^h}} \left(\sum_{i=1}^n \sum_{j=1}^h e_j (c^\top M_j)_i G_i\right) \geq r$. Defining $v(c) = (e_i c_j)_{i \in [h], j \in [m]}$, we have that $\text{rank}_{\mathbb{F}_{2^h}} \left(\sum_{i=1}^n (v(c)^\top M)_i G_i\right) \geq r$ for all $c \neq 0$. This means that for all $c \neq 0$, it holds that $v(c) \notin \tilde{V}$. Clearly the map $c \mapsto v(c)$ is a linear injective map. Thus the subspace $\overline{V} = \{v(c) \mid c \in \mathbb{F}_{2^h}^m\}$ satisfies $\dim_{\mathbb{F}_{2^h}}(\overline{V}) = m$ and $\tilde{V} \cap \overline{V} = \{0\}$. This implies that $\dim_{\mathbb{F}_{2^h}}(\tilde{V}) = \dim_{\mathbb{F}_2}(V) \leq hm - m$. □

We are now ready to complete the proof. Define $V^\perp = \{u \in \mathbb{F}_2^{hm} \mid u^\top v = 0 \text{ for all } v \in V\}$. From Claim 8.10, we have that $\dim_{\mathbb{F}_2}(V^\perp) \geq m$. Let X be a uniform random variable over $V^\perp$. From the fact that LExt is an (m, δ)-strong linear seeded extractor and Lemma 4.11, we have with probability $1 - \delta$ over $s \in \mathbb{F}_2^\ell$, that $b^\top \text{LExt}(X, s)$ is a uniform bit for all non-zero $b \in \mathbb{F}_2^{m'}$. However, this must mean that $b^\top L_s \notin V$ for all non-zero b as otherwise $b^\top \text{LExt}(X, s) = 0$ by the definition of $V^\perp$. Since $b^\top L_s \notin V$ for all non-zero b, it implies that $W = L_s \cdot M$ satisfies

$$\text{rank}_{\mathbb{F}_2} \left(\sum_{i=1}^n (b^\top W)_i G_i\right) \geq r/(hm).$$

Thus, we have shown that for each good matrix U_i in the original rank combiner, $1 - \delta$ fraction of the seeds $s \in \mathbb{F}_2^\ell$ lead to a good matrix $W(U_i, s)$. Thus, the overall success probability is $1 - (\epsilon + \delta)$ and the proof is complete. □

8.3 Putting Things Together

Theorem 8.11 (Explicit rank combiner over $\mathbb{F}_2$)**.** *For any $d \geq \log^{40} n$ and $\epsilon \geq 1/n^2$, there exists an explicit $(d, (d^{1/10}/\log^{3/2} n), \epsilon)$-rank combiner over $\mathbb{F}_2$ with output length $d^{1/10}$ and size* $\text{poly}(n)$.

Proof. Take $h = O(\log(n))$ and $m = d^{1/5}/2$. Let $U_1, \ldots, U_N \in \mathbb{F}_{2^h}^{m \times n}$ be the $(d/h, d^{3/10}/\sqrt{h}, \epsilon/2)$ rank combiner from Theorem 8.5 with $N = \text{poly}(n)$ (we use the fact that $\epsilon \geq 1/n^2$). We will now use an appropriately chosen strong linear seeded extractor over $\mathbb{F}_2$ and apply Theorem 8.9.

We use the linear seeded extractor due to Li Theorem 4.7 with a small modification. Recall that the linear seeded extractor needed by Theorem 8.9 acts on hm bits. Directly applying Theorem 4.7 with input length hm forces the error to be at least $1/(hm)^2$. To overcome this, we modify Li's extractor as follows: (1) first pad the hm length input string with zero's to make the total length $2n$, (2) apply the extractor from Theorem 4.7 with input length $2n$. As long as $m \geq \Omega(\log^8 n)$, the output length is $\sqrt{m}$ with error $1/(4n^2) \leq \epsilon/2$. From our assumption that $d \geq \log^{40} n$, we have that $m = d^{1/5}/2 \geq \log^8 n$. Thus, Theorem 8.9 implies an explicit $(d, (d^{1/10}/\log^{3/2} n), \epsilon)$-rank combiner with output length $d^{1/10}$ and size $\text{poly}(n)$. □

This immediately gives the construction of the linear seeded extractor where the good rows have unusually good error.

Proof of Theorem 8.1. Truncate the output length of the rank combiner from Theorem 8.11 and apply Lemma 8.3. □

8.4 Provenance of the rank combiner construction

Related constructions and arguments have appeared in prior work studying rank condensers, subspace designs, and polynomial identity testing. In particular, constructions similar to the large field construction (Theorem 8.5) appear in [FS12, FSS14] (see Lemma 3.10 and Remark 3.11 in the full version of [FSS14]).

However, their constructions do not use the Sidon set weights which are used in the proof of Lemma 8.6. Other papers using similar constructions include [KS01] (see Section 3 of the full version) and [FGT16] (see Section 3 of the full version) in the context of polynomial identity testing and bipartite matching. The multiplicity argument for the determinant polynomial in Lemma 8.7 appears in [GK13] (see Claim 16). Recent work of [GRSZ26] (Lemma 5.6) studied the related problem of symbolic determinant identity testing when restricted to the setting of rank one summands.

9 Quadratic extractors

Algorithm 2 Quadratic Extractor (QExt)

Input: $x \in \{0,1\}^n$— an n-bit string.
Output: $v \in \{0,1\}^m$— output string.

Sub-Routines and Parameters:
Let C_1, C_2, C_3 be appropriately chosen constants. Let $N = n^{C_1}$.
Let RCombExt : $\mathbb{F}_2^n \times [N] \to \mathbb{F}_2^d$ be the linear seeded extractor from Theorem 8.1.
Let QuadCB : $\mathbb{F}_2^n \times \mathbb{F}_2^d \times [N] \to \{0,1\}$ be the quadratic correlation breaker from Theorem 7.2 with output length 1, $t = (m \log n)^{C_2}$, $\delta = N^{-10t}$ and $d = t^5(\log(n/\delta))^3$.
Let NOBFExt : $\{0,1\}^N \to \{0,1\}^m$ be the NOBF extractor from Theorem 3.13.

1. Let $w = \{\text{RCombExt}(x,i)\}_{i\in[N]}$.
2. Let $s = \{\text{QuadCB}(x, w_i, i)\}_{i\in[N]}$.
3. Output NOBFExt(s).

We execute a win-win argument where we analyze the extractor differently based on if it contains a large subspace or not. We use the following elementary lemma from [CT15] that states that any quadratic map that is affine on an affine subspace V is also affine on shifts of V.

Lemma 9.1 (Claim 3.2 from [CT15])**.** *Let f be a quadratic map on $\mathbb{F}_2^n$ such that $f|_V$ is affine on an affine subspace $V \subseteq \mathbb{F}_2^n$. Then $f|_{u+V}$ is also affine, for all $u \in \mathbb{F}_2^n$.*

The following lemma is now immediate.

Lemma 9.2. *Let $X = f(U)$ be a quadratic source such that $H_\infty^\epsilon(X) \geq k$ and $f : \mathbb{F}_2^{k+\ell} \to \mathbb{F}_2^n$ for $\ell = O(\log(1/\epsilon))$. Suppose X is not q-pure. Then, X can be decomposed is $O(\epsilon)$ close to a convex combination of affine sources with entropy $q - O(\log(1/\epsilon))$.*

Proof. From the fact that X is not q-pure, we have an affine subspace $V = v + S$ for subspace S such that $f|_V$ is affine and $\dim(f(V)) \geq q$. Let $\dim(S) = d \geq q$. From Lemma 9.1, we have that $f|_{u+V}$ is affine for all shifts u. Let $C : \mathbb{F}_2^{k+\ell} \to \mathbb{F}_2^{k+\ell-d}$ be the full rank linear map with kernel S. $C^{-1}(c)$ is a shift of V for all c. Thus, we have that $f(U) \mid \{C(U) = c\}$ is an affine source for all c. Also, $H_\infty^{1/2}(f(U) \mid C(U) = c) \geq q - O(\log(1/\epsilon)$ with probability $1 - O(\epsilon)$ over $c \sim C(U)$. Since affine sources are flat, Lemma 4.17 implies that $H_\infty(X \mid C(U) = c) \geq q - O(\log(1/\epsilon))$ with probability at least $1 - O(\epsilon)$ over $c \sim C(U)$. □

9.1 QExt is an affine extractor

First, we argue that the construction in Algorithm 2 is an affine extractor. Similar constructions have been used in prior work (with RCombExt replaced by a standard linear seeded extractor) and showing that QExt is an affine extractor is relatively straightforward.

Theorem 9.3. *There exists a constant $C > 0$ such that for any affine source X with $H_\infty(X) = k \geq \log^C n$, it holds that* $\mathrm{QExt}(X) \approx_{n^{-\Omega(1)}} U_{k^{\Omega(1)}}$.

We defer the proof to the appendix Appendix B.

9.2 QExt is a quadratic extractor

We now state and prove the main theorem of the paper.

Theorem 9.4. *There exists a constant $C > 0$ such that for any quadratic source X satisfying $H_\infty(X) = k \geq \log^C n$, it holds that* $\mathrm{QExt}(X) \approx_{n^{-\Omega(1)}} U_{k^{\Omega(1)}}$.

Proof. We will first instantiate various parameters that we use throughout the proof. Let m be the output length of the extractor with $m = k^c$ for some constant c that we will choose later. We instantiate the linear seeded extractor RCombExt from Theorem 8.1 with $\epsilon = 1/n^2$ and derivative entropy requirement $k^{1/4}$ (we will choose the output length soon). Let $N = n^{C_1}$ be the number of seeds required by Theorem 8.1. Let $t = (m \log n)^{C_2}$ be chosen such that Theorem 3.13 gives a $(q = N/n^2, t)$-NOBF extractor with error n^{-C_3} and output length m. We define an error parameter $\delta = N^{-10t}$ which will be the error of each almost-uniform bit in the NOBF source that we finally obtain. Let $d = C_4 t^5 (\log n \log(n/\delta))^2$ be the seed length of QuadCB when we have target error δ and t as chosen earlier. d will also be the output length of RCombExt (we will analyze it's error later).

We choose the constant c in $m = k^c$ sufficiently small and the constant C in $k \geq \log^C n$ sufficiently large such that

$$\max(t, d, \log(1/\delta)) \leq k^{1/100}$$

and $k^{1/3}$ is asymptotically greater than the entropy requirement of Theorem 9.3.

Our first step in the proof is to apply the input reduction lemma (Lemma 4.14): X is δ-close to a convex combination of quadratic sources X' with $H_\infty^{\delta^6}(X') \geq k - 1$ and $X' = f(U_{k+\ell})$ for quadratic f and $\ell = O(\log(1/\delta))$. We will argue that QExt extracts from any source X' satisfying this property. Henceforth, without loss of generality, we assume that our source $X = f(U_{l+\ell})$ satisfies (1) $H_\infty^{\delta^6}(X) \geq k - 1$ and (2) $\ell = O(\log(1/\delta))$.

We now split the proof into two cases depending on if X is $(ck^{1/3})$-pure for some small constant $c > 0$.

Not pure case Suppose X is not $(ck^{1/3})$-pure. Then, Lemma 9.2 implies that X is $O(\delta)$ close to a convex combination of affine sources with entropy $ck^{1/3}/2$ for constant $c' < c$. The theorem statement now follows from Theorem 9.3.

Pure case Suppose X is $(ck^{1/3})$-pure. We now apply the structure theorem. Corollary 5.2 (with $s = k + \ell$) implies that X is δ^2 close to a convex combination of sources X' satisfying (1) $X' = g(Y)$ for $Y = U'_s$ and $s' \geq s/2$, (2) $H_\infty(\partial_a g(.)) \geq k^{2/3}/(4c)$, (3) $H_\infty^{\delta^4}(X) \geq s' - \ell$. From Lemma 4.19, we have that each x' is still $(ck^{1/3})$-pure. We rename X' as X and g as f and henceforth consider X satisfying the above property.

Henceforth let $X = f(Y)$ where (1) $H_\infty^{\delta^4}(X) = \overline{k} = \Omega(k)$, (2) $f : \mathbb{F}_2^{\overline{k}+\ell} \to \mathbb{F}_2^n$ is quadratic, and (4) $Y = U_{\overline{k}+\ell}$, (3) $H_\infty(\partial_a f(Y)) \geq \Omega(k^{2/3})$ for all non-zero a. We will first argue that the random variable $W = (\mathrm{RCombExt}(X, i))_{i \in [N]}$ has the following property: $W_i \approx_{\delta^2} U_d$ with probability $1 - 1/n^2$ over $i \sim [N]$.

Recall the output length d of RCombExt that was chosen earlier satisfies $d \leq k^{1/100}$. Thus, from our derivative entropy guarantee of $\Omega(k^{2/3})$, we have that $W_i \approx_{2^{-k^{1/40}}} U_d$ (with probability $1 - 1/n^2$ over $i \sim [N]$) and this error is much smaller than δ^2. We also have that $W_i = p^i(Y)$ for some quadratic polynomial p^i.

Let $Q \subseteq [N]$ be the set of indices where $W_i \approx_{\delta^2} U_d$. Consider the random variable

$$S = (\mathrm{QuadCB}(X, W^i, i))_{i \in [N]}.$$

We will argue that for all sets $T \subseteq Q$ with $|T| \leq t$, we have that $|\mathbb{E}[\chi_T(S)]| \leq O(\delta)$. This is immediate from Theorem 7.2 and the fact that X, W satisfy the premise (by our choice of parameters) and the indices $i \in [N]$ form valid advice strings.

Now, applying Fact 4.9 implies that S is N^{-t}-close to a (q,t)-NOBF source with $q = N/n^2$ and $t = (m \log n)^{C_2}$ chosen earlier. The theorem statement now follows from Theorem 3.13. □

10 Applications

We now state and prove improved extractors for NC^0 and AC^0 sources. Prior constructions for these sources were obtained by reducing to 1-local sources (which are a special type of affine source). We give a more efficient reduction to 2-local sources (which are a special type of quadratic source), and use this to obtain the improved results.

Before we show our results, we define an intermediate notion that is useful.

Definition 10.1 (d-local NOBF source)**.** *A source X is d local NOBF if $X = f(U_s)$ for a d-local map f such that for any $i \in [s]$, there exist $j(i) \in [n]$ such that $f_{j(i)}(x) = x_i$.*

Informally, a d-local NOBF source is a d-local source where the generating bits are printed out somewhere in the source.

We use the following theorem (from [ACG+22]) that reduces extracting from d-local sources to extracting from d-local NOBF sources.

Theorem 10.2 (Theorem 2 of [ACG+22])**.** *There exists a constant $c > 0$ such that any d-local source X with $H_\infty(X) \geq k$ is ϵ-close to a convex combination of d-local NOBF sources with min-entropy at least k' for $k' \geq ck/(2^d d^2)$ and $\epsilon = 2^{-ck'}$.*

We first show that every d-local source of sufficient entropy is close to a convex combination of 2-local sources.

Theorem 10.3. *There exists a constant $c, C > 0$ such that any d-local source X with $H_\infty(X) \geq k$ is ϵ-close to a convex combination of 2-local sources with min-entropy at least k' for $k' \geq Ck^{3/2}/(2^{2d}\sqrt{n})$ and $\epsilon = 2^{-ck/(2^d d^2)}$.*

Proof. From Theorem 10.2 we have that X is $2^{-ck'}$ close to a convex combination of d-local NOBF sources with min-entropy $k' \geq ck/(2^d d^2)$ for some constant $c > 0$. Let $S_i \subseteq [k']$ be the set of input generating bits that the function f_i depends on. Consider the set system $\mathcal{S} = \{T \mid T \subseteq S_i \text{ for some } i \in [n] \text{ and } |T| = 3\}$. Clearly, we have $|\mathcal{S}| \leq n \cdot \binom{d}{3}$ as each S_i has size at most d. Recall Turan's theorem for hypergraphs [Spe72] which states that the independence number of a 3-uniform hypergraph with p vertices and q edges is at least $C \min(p, p^{3/2}/\sqrt{q})$ for some constant C. Interpreting $\mathcal{S}$ as a hypergraph then implies that $\mathcal{S}$ has an independent set of size at least $C'(k')^{3/2}/\sqrt{nd^3}$ for some constant C'.

Let $I \subseteq [k']$ be the independent set of $\mathcal{S}$. From the definition of independent set, we have that I does not contain S_i for any $i \in [n]$. This means that $|I \cap S_i| \leq 2$ for all $i \in [n]$. Consider the source X' obtained by conditioning on the generating bits in $[k'] \setminus I$. By the intersection property described above, this source

is 2-local NOBF. Also, the min-entropy of the source is at least $|I|$ as the it is still NOBF (the unfixed generating bits are still printed out in the source). Thus, we have that our original source is $\epsilon = 2^{-ck'}$ to a convex combination of 2-local NOBF sources with entropy at least $|I| \geq C'' k^{3/2}/(2^{2d}\sqrt{n})$ for some constant $C'' > 0$. □

Theorem 10.4 (Extractors for NC^0-sources)**.** *For any constant $\alpha > 0$ and any NC^0-source X with $H_\infty(X) = k \geq n^{1/3+\alpha}$, it holds that* $\mathrm{QExt}(X) \approx_{n^{-\Omega(1)}} U_{k^{\Omega(1)}}$

Proof. The theorem follows from applying Theorem 10.3, Theorem 9.4 and the fact that a 2-local source is a quadratic source. □

The following theorem by [Vio14] shows that AC^0-sources can be reduced to NC^0-sources (we instantiate the theorem to entropy close to $n^{1/3}$.

Theorem 10.5 ([Vio14], see Theorem 8.17 from [Goo23])**.** *For any $\alpha > 0$ and AC^0-source X satisfying $H_\infty(X) \geq n^{1/3+\alpha}$, there exists constants d, β such that X is $n^{-\Omega(1)}$-close to a convex combination of d-local sources with entropy at least $n^{1/3+\beta}$.*

Our theorem on extraction from AC^0 sources is now immedaite.

Theorem 10.6 (Extractors for AC^0-sources)**.** *For any constant $\alpha > 0$ and any AC^0-source X with $H_\infty(X) = k \geq n^{1/3+\alpha}$, it holds that* $\mathrm{QExt}(X) \approx_{n^{-\Omega(1)}} U_{k^{\Omega(1)}}$

Proof. Apply Theorem 10.5 and Theorem 10.4. □

11 Open Directions

The natural open question is to construct better extractors for higher degree sources. We now discuss the parts of our proof that extend and the parts that need new ideas.

The proof of quadratic conditioning Lemma 3.3 naturally extends to all constant degree polynomials, as the lemma from [CT15] also has constant degree extensions. The proofs using Cauchy Schwarz in Lemma 6.1 and Lemma 7.3 also extends to degree d sources if we assume the premise that the sources corresponding to $(d-1)$-th directional derivatives have high entropy.

The proofs of Theorems 3.7 and 5.1 do not seem to extend immediately to degree 3 and higher. A natural question is the following: can all degree d sources with a min-entropy guarantee be decomposed into a convex combination of sources of one of two types: (1) degree $(d-1)$ sources with high entropy, or (2) sources with high $(d-1)$-th derivative entropy? Answering this question positively for cubics would already give an improved disperser by adapting our proof in Appendix A.

Another open direction is to find tensor analogues of rank combiners (Definition 8.2). These seem useful to construct good linear seeded extractors for higher degree polynomial sources.

Finally, an interesting question is to see if any of our techniques are useful for sources of the type $f(X+Y)$, where X and Y are independent sources each having a min-entropy guarantee, and f is a quadratic map. This is the natural generalization of sum-set sources to higher degree.

References

[ACG+22] Omar Alrabiah, Eshan Chattopadhyay, Jesse Goodman, Xin Li, and João Ribeiro. Low-Degree Polynomials Extract From Local Sources. In Mikołaj Bojańczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*, volume 229 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:20, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[AGM03] Noga Alon, Oded Goldreich, and Yishay Mansour. Almost k-wise independence versus k-wise independence. *Information Processing Letters*, 88(3):107–110, 2003.

[AGMR25] Omar Alrabiah, Jesse Goodman, Jonathan Mosheiff, and João Ribeiro. Low-Degree Polynomials Are Good Extractors. In Alina Ene and Eshan Chattopadhyay, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2025)*, volume 353 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 38:1–38:25, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[AGO+19] Divesh Aggarwal, Siyao Guo, Maciej Obremski, João L. Ribeiro, and Noah Stephens-Davidowitz. Extractor lower bounds, revisited. *Electron. Colloquium Comput. Complex.*, TR19, 2019.

[AL93] Miklós Ajtai and Nathan Linial. The influence of large coalitions. *Combinatorica*, 13(2):129–145, 1993.

[BDL16] Jean Bourgain, Zeev Dvir, and Ethan Leeman. Affine extractors over large fields with exponential error. *computational complexity*, 25:921–931, 2016.

[BDSGM23] Marshall Ball, Dana Dachman-Soled, Eli Goldin, and Saachi Mutreja. Extracting randomness from samplable distributions, revisited. In *64th Annual Symposium on Foundations of Computer Science (FOCS 2023, to appear)*. IEEE, 2023.

[BGLZ15] Abhishek Bhowmick, Ariel Gabizon, Thái Hoang Lê, and David Zuckerman. Deterministic extractors for additive sources. In *6th Conference on Innovations in Theoretical Computer Science (ITCS 2015)*, pages 277–286, 2015.

[BIW06] Boaz Barak, Russell Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. *SIAM Journal on Computing*, 36(4):1095–1118, 2006. Preliminary version in FOCS 2004.

[BK12] Eli Ben-Sasson and Swastik Kopparty. Affine dispersers from subspace polynomials. *SIAM Journal on Computing*, 41(4):880–914, 2012. Preliminary version in STOC 2009.

[BKS+10] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson. Simulating independence: New constructions of condensers, ramsey graphs, dispersers, and extractors. *J. ACM*, 57(4), May 2010.

[BNS92] László Babai, Noam Nisant, and Márió Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *Journal of Computer and System Sciences*, 45(2):204–232, 1992.

[Bou05] Jean Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1(01):1–32, 2005.

[Bou07] Jean Bourgain. On the construction of affine extractors. *GAFA Geometric And Functional Analysis*, 17(1):33–57, 2007.

[BRSW06] Boaz Barak, Anup Rao, Ronen Shaltiel, and Avi Wigderson. 2-source dispersers for sub-polynomial entropy and ramsey graphs beating the frankl-wilson construction. In *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '06, page 671–680, New York, NY, USA, 2006. Association for Computing Machinery.

[BSG12] Eli Ben-Sasson and Ariel Gabizon. Extractors for polynomials sources over constant-size fields of small characteristic. In Anupam Gupta, Klaus Jansen, José Rolim, and Rocco Servedio, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 399–410, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[BSHR+01] Eli Ben-Sasson, Shlomo Hoory, Eyal Rozenman, Salil Vadhan, and Avi Wigderson. Extractors for affine sources. Unpublished manuscript, 2001.

[BSS25] Marshall Ball, Ronen Shaltiel, and Jad Silbak. Extractors for samplable distributions with low min-entropy. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 596–603, 2025.

[BV10] Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. *SIAM Journal on Computing*, 39(6):2464–2486, 2010.

[CG88] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988. Preliminary version in FOCS 1985.

[CG22] Eshan Chattopadhyay and Jesse Goodman. Improved extractors for small-space sources. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 610–621. IEEE, 2022.

[CGG24] Eshan Chattopadhyay, Jesse Goodman, and Mohit Gurumukhani. Extractors for polynomial sources over $\mathbb{F}_2$. In *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, pages 28–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2024.

[CGL21] Eshan Chattopadhyay, Jesse Goodman, and Jyun-Jie Liao. Affine extractors for almost logarithmic entropy. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 622–633. IEEE, 2021.

[Cha20] Eshan Chattopadhyay. Guest column: A recipe for constructing two-source extractors. *SIGACT News*, 51(2):38–57, June 2020.

[CL16] Eshan Chattopadhyay and Xin Li. Extractors for sumset sources. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, page 299–311, New York, NY, USA, 2016. Association for Computing Machinery.

[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*, pages 1584–1597, 2022.

[Coh16] Gil Cohen. Two-source dispersers for polylogarithmic entropy and improved ramsey graphs. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 278–284, 2016.

[CS16] Gil Cohen and Igor Shinkar. The complexity of DNF of parities. In *7th Conference on Innovations in Theoretical Computer Science (ITCS 2016)*, pages 47–58, 2016.

[CT15] Gil Cohen and Avishay Tal. Two structural results for low degree polynomials and applications. In Naveen Garg, Klaus Jansen, Anup Rao, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2015, Princeton, NJ, USA, August 24-26, 2015*, volume 40 of *LIPIcs*, pages 680–709. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015.

[CZ19] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. *Annals of Mathematics*, 189(3):653–705, 2019. Preliminary version in STOC 2016.

[DG10] Matt DeVos and Ariel Gabizon. Simple affine extractors using dimension expansion. In *IEEE 25th Annual Conference on Computational Complexity (CCC 2010)*, pages 50–57. IEEE, 2010.

[DG26] Dean Doron and Jesse Goodman. Low-error extractors. 2026.

[DGW09] Zeev Dvir, Ariel Gabizon, and Avi Wigderson. Extractors and rank extractors for polynomial sources. *computational complexity*, 18(1):1–58, 2009.

[DK11] Evgeny Demenkov and Alexander Kulikov. An elementary proof of a $3n - o(n)$ lower bound on the circuit complexity of affine dispersers. In *International Symposium on Mathematical Foundations of Computer Science*, pages 256–265. Springer, 2011.

[DKSS13] Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan. Extensions to the method of multiplicities, with applications to kakeya sets and mergers. *SIAM Journal on Computing*, 42(6):2305–2328, 2013. Preliminary version in FOCS 2009.

[Dvi12] Zeev Dvir. Extractors for varieties. *Computational complexity*, 21:515–572, 2012. Preliminary version in CCC 2009.

[DW12] Anindya De and Thomas Watson. Extractors and lower bounds for locally samplable sources. *ACM Trans. Comput. Theory*, 4(1), March 2012.

[ET41] P. Erdös and P. Turán. On a problem of sidon in additive number theory, and on some related problems. *Journal of the London Mathematical Society*, s1-16(4):212–215, 1941.

[FG15] Michael A Forbes and Venkatesan Guruswami. Dimension expanders via rank condensers. In *18th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2015, and 19th International Workshop on Randomization and Computation, RANDOM 2015*, pages 800–814. Schloss Dagstuhl-Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing, 2015.

[FGHK16] Magnus Gausdal Find, Alexander Golovnev, Edward Hirsch, and Alexander S Kulikov. A better-than-$3n$ lower bound for the circuit complexity of an explicit function. In *57th Annual Symposium on Foundations of Computer Science (FOCS 2016)*, pages 89–98. IEEE, 2016.

[FGT16] Stephen Fenner, Rohit Gurjar, and Thomas Thierauf. Bipartite perfect matching is in quasi-nc. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 754–763, 2016.

[FS12] Michael A Forbes and Amir Shpilka. On identity testing of tensors, low-rank recovery and compressed sensing. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 163–172, 2012.

[FSS14] Michael A Forbes, Ramprasad Saptharishi, and Amir Shpilka. Hitting sets for multilinear read-once algebraic branching programs, in any order. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 867–875, 2014.

[GK13] Venkatesan Guruswami and Swastik Kopparty. Explicit subspace designs. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 608–617, 2013.

[GKW21] Alexander Golovnev, Alexander Kulikov, and Ryan Williams. Circuit depth reductions. In *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021.

[Goo23] Jesse Patrick McGrenra Goodman. *Seedless Extractors*. Cornell University, 2023.

[GR08] Ariel Gabizon and Ran Raz. Deterministic extractors for affine sources over large fields. *Combinatorica*, 28(4):415–440, 2008. Preliminary version in FOCS 2005.

[GRSZ26] Zeyu Guo, Roshan Raj, Chong Shangguan, and Zihan Zhang. Explicit rank extractors and subspace designs via function fields, with applications to strong blocking sets. *arXiv preprint arXiv:2604.13431*, 2026.

[GUV09] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. Unbalanced expanders and randomness extractors from parvaresh–vardy codes. *Journal of the ACM (JACM)*, 56(4):1–34, 2009. Preliminary version in CCC 2007.

[GVJZ23] Zeyu Guo, Ben Lee Volk, Akhil Jalan, and David Zuckerman. Extractors for images of varieties. In *55th Annual ACM Symposium on Theory of Computing (STOC 2023)*, 2023.

[GX13] Venkatesan Guruswami and Chaoping Xing. List decoding reed-solomon, algebraic-geometric, and gabidulin subcodes up to the singleton bound. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 843–852, 2013.

[KRVZ06] Jesse Kamp, Anup Rao, Salil Vadhan, and David Zuckerman. Deterministic extractors for small-space sources. In *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '06, page 691–700, New York, NY, USA, 2006. Association for Computing Machinery.

[KS01] Adam R. Klivans and Daniel Spielman. Randomness efficient identity testing of multivariate polynomials. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, STOC '01, page 216–223, New York, NY, USA, 2001. Association for Computing Machinery.

[Li11] Xin Li. A new approach to affine extractors and dispersers. In *26th Annual Conference on Computational Complexity (CCC 2011)*, pages 137–147. IEEE, 2011.

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

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

[Li23] Xin Li. Two source extractors for asymptotically optimal entropy, and (many) more. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1271–1281. IEEE, 2023.

[Lov08] Shachar Lovett. Unconditional pseudorandom generators for low degree polynomials. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, page 557–562, New York, NY, USA, 2008. Association for Computing Machinery.

[LRVW03] Chi-Jen Lu, Omer Reingold, Salil Vadhan, and Avi Wigderson. Extractors: Optimal up to constant factors. In *35th Annual ACM Symposium on Theory of Computing (STOC 2003)*, pages 602–611, 2003.

[LY22] Jiatu Li and Tianqi Yang. $3.1n - o(n)$ circuit lower bounds for explicit functions. In *54th Annual ACM Symposium on Theory of Computing (STOC 2022)*, pages 1180–1193, 2022.

[LZ19] Fu Li and David Zuckerman. Improved extractors for recognizable and algebraic sources. In *23rd International Conference on Randomization and Computation (RANDOM 2019)*, 2019.

[Mek17] Raghu Meka. Explicit resilient functions matching ajtai-linial. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1132–1148. SIAM, 2017.

[NZ96] Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, 1996. Preliminary version in STOC 1993.

[OS25] Justin Oh and Ronen Shaltiel. Extractors for samplable distributions from the two-source extractor recipe. (TR25-107), July 2025. https://eccc.weizmann.ac.il/report/2025/107.

[PSZ97] Ramamohan Paturi, Michael E. Saks, and Francis Zane. Exponential lower bounds for depth 3 boolean circuits. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97, page 86–91, New York, NY, USA, 1997. Association for Computing Machinery.

[Rao09] Anup Rao. Extractors for low-weight affine sources. In *2009 24th Annual IEEE Conference on Computational Complexity*, pages 95–101. IEEE, 2009.

[Rem16] Zachary Remscrim. The Hilbert function, algebraic extractors, and recursive fourier sampling. In *57th Annual Symposium on Foundations of Computer Science (FOCS 2016)*, pages 197–208. IEEE, 2016.

[Sha11] Ronen Shaltiel. Dispersers for affine sources with sub-polynomial entropy. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 247–256, 2011.

[Sha25a] Ronen Shaltiel. Extractors for samplable distribution with polynomially small min-entropy. April 2025. ECCC TR25-054.

[Sha25b] Ronen Shaltiel. Multiplicative extractors for samplable distributions. In *40th Computational Complexity Conference (CCC 2025)*, pages 22–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2025.

[Spe72] Joel Spencer. Turán's theorem for k-graphs. *Discrete Mathematics*, 2(2):183–186, 1972.

[Tre01] Luca Trevisan. Extractors and pseudorandom generators. *Journal of the ACM*, 48(4):860–879, 2001. Preliminary version in STOC 1999.

[TV00] L. Trevisan and S. Vadhan. Extracting randomness from samplable distributions. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, FOCS '00, page 32, USA, 2000. IEEE Computer Society.

[Vaz86] Umesh Virkumar Vazirani. *Randomness, adversaries and computation (random polynomial time)*. University of California, Berkeley, 1986.

[Vaz87] Umesh V Vazirani. Strong communication complexity or generating quasi-random sequences from two communicating semi-random sources. *Combinatorica*, 7(4):375–392, 1987. Preliminary version in STOC 1985.

[Vio09] Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d. *computational complexity*, 18(2):209–217, 2009.

[Vio14] Emanuele Viola. Extractors for circuit sources. *SIAM Journal on Computing*, 43(2):655–672, 2014.

[VW08] Emanuele Viola and Avi Wigderson. Norms, xor lemmas, and lower bounds for polynomials and protocols. *Theory of Computing*, 4(7):137–168, 2008.

[Yeh11] Amir Yehudayoff. Affine extractors over prime fields. *Combinatorica*, 31(2):245, 2011.

A High-entropy dispersers using the challenge-response framework

We now present an alternative construction that gives a disperser for quadratic sources with min-entropy $n - n^{0.1}$. This is a worse result than our main theorem (which gives an extractor at polylogarithmic entropy), however we believe the usage of the challenge-response framework in the proof is interesting in its own right.

Theorem A.1. *There exists an explicit function* QDisp : $\mathbb{F}_2^n \to \mathbb{F}_2$ *(Algorithm 3) that is a disperser for quadratic sources with entropy* $n - n^{0.1}$.

Algorithm 3 Quadratic Disperser (QDisp)

Input: $x \in \{0,1\}^n$— an n-bit string.
Output: $v \in \{0,1\}$— output bit.

Sub-Routines and Parameters:
Let $\Delta = n^{0.1}$, $\ell = 0.1n^{0.1}$, $d_1 = 100\Delta$, $d_2 = O(\log n)$, $\epsilon = 2^{-10n^{0.1}}$.
Let CGLExt : $\mathbb{F}_2^n \times \mathbb{F}_2^{d_1} \to \{0,1\}^\ell$ be a (k,ϵ) linear seeded extractor with seed length $O(\log(n/\epsilon))$ (this is for the challenge/final output, we use Theorem 4.6)
Let LExt : $\mathbb{F}_2^n \times \mathbb{F}_2^{d_2} \to \{0,1\}^\ell$ be a $(k, 1/n^2)$ linear seeded extractor (this is for the response, we use Theorem 4.7)
Let AffineExt : $\mathbb{F}_2^n \to \{0,1\}$ be an affine extractor (Theorem 4.10)

1. Let $s = x_{[d_1]}$.
2. Let $a = \text{CGLExt}(x, s)$.
3. For each $i \in \mathbb{F}_2^{d_2}$, check if $a = \text{LExt}(x, i)$.
4. If there exists an $i \in \mathbb{F}_2^{d_2}$ that satisfied the previous check, then output AffineExt(x).
5. Else, output a_1.

For each variable (x, s, a) defined in Algorithm 3, we use capitalized letters $X, S(X), A(X)$ etc to denote the corresponding random variables. That is, $S(X) = X_{[d_1]}$, $A = \text{CGLExt}(X, S(X))$ and so on for random variable X.

We analyze the disperser by splitting into two cases based on if X is $n^{0.2}$-pure.

A.1 X is not Pure

Theorem A.2. *Let* $X = f(U_s)$ *be a quadratic source with an affine subsource* W *with min-entropy at least* $n^{0.2}$. *Then* QDisp *(from Algorithm 3) is a disperser for* X.

Proof. We will restrict our attention to W and show that QDisp is a disperser for W. A disperser on a subsource is also a disperser on the original source.

Since $S(W)$ has output length at most $d_1 = 100n^{0.1}$, fixing $S(W) = s$ only reduces the dimension by d_1. Thus, W has a subsource W' of dimension at least $\Omega(n^{0.2})$ where $S(W') = s$ is a constant. Now, note that $A(W')$ is a linear function of W' (as s is fixed). Thus, we can further restrict $A(W') = a$ at a loss of additional dimension $\ell = 0.1n^{0.1}$ to obtain a subsource W'' of dimension $\Omega(n^{0.2})$ where $A(W'') = a$ is a constant.

From Lemma 4.11, we have that there exists $i \in \mathbb{F}_2^{d_2}$ such that $\text{LExt}(W'', i) = U_\ell$. Condition on the event that $\text{LExt}(W'', i) = a$. This happens with probability at least $2^{-\ell}$ over W''. Furthermore, this conditioning

reduces the dimension of W'' by at most $\ell \leq 0.1n^{0.1}$. Let $\tilde{W}$ be the subsource of W'' obtained after this conditioning. This subsource has min-entropy at least $\Omega(n^{0.2})$. Restricted to the subsource $\tilde{W}$, the condition checked in Line 3 of Algorithm 3 holds with probability 1 for the seed i identified earlier. Thus, on $\tilde{W}$, the output distribution of the disperser is AffineExt$(\tilde{W})$ which is 0.1 close to U_1 and hence is not constant. This completes the proof. □

A.2 X is Pure

We are now ready to argue that QDisp is an extractor for pure quadratic sources.

Theorem A.3. *Let X be a $n^{0.2}$-pure quadratic source with $H_\infty(X) \geq n - n^{0.1}$. Then,* QDisp *is an extractor for X with error $2^{-\Omega(n^{0.1})}$.*

Proof. Let $\Delta = n^{0.1}$ and let $H_\infty(X) \geq n - \Delta$. From Lemma 4.14, we have that X is $\epsilon = 2^{-\Omega(n^{0.1})}$ close to a convex combination of quadratic subsources X' where $H^\epsilon_\infty(X') \geq n - \Delta - 1$ and the generating map f of X' has seedlength $m = n - \Delta + O(n^{0.1})$. It suffices to prove that QDisp extracts from such subsources X'. By Lemma 4.19, these sources are still $n^{0.2}$-pure. Henceforth, we will use X to refer to X'. Thusm X is a quadratic source with generating map f, $m = n - \Delta + O(n^{0.1})$ generating bits and satisfying $H^\epsilon_\infty(X) \geq n - \Delta - 1$.

We use $V = U_m$ to denote the input seed for the source. That is, $X = f(V)$. From our choice of $d_1 = 100\Delta$ and Lemma 4.16, we have that $H^\epsilon_\infty(S(X)) \geq 0.99d_1$.

Quadratic conditioning: input reduction Some additional set up will be useful. Since restricting to a subset of bits is a linear operation, we have that $S = h(V)$ for a quadratic map h. From Lemma 3.4, we have a full rank linear map $L : \mathbb{F}_2^m \to \mathbb{F}_2^{m-r}$ such that

$$d_{\mathrm{TV}}((h(V), L(V)), (h(V), U'_{m-r})) \leq \epsilon \tag{15}$$

where U' is a uniform string independent of V and $r = 100\Delta + O(\log(1/\epsilon))$. Let $e_1, \ldots, e_r$ be a basis for the nullspace of L. Let $e_{r+1}, \ldots, e_m$ be a completion of the basis to span all of $\mathbb{F}_2^m$. We have that $V = A + B$ A is uniform over the span of $e_{r+1}, \ldots, e_m$, B is uniform over the nullspace of L and A and B are independent. Observe that we can express $A = M_1(U_{m-r})$ and $B = M_2(U_r)$ for appropriate full rank linear maps M_1 and M_2. We will use $Y = U_{m-r}$ and $Z = U_r$ to denote the input random variables of M_1 and M_2. Thus, the original source can be rewritten as $X = f(M_1(Y) + M_2(Z))$. Also, $L(M_1(Y) + M_2(Z)) = L(M_1(Y))$. From our choice of M_1, M_2, Y, Z, we have that $L(M_1(Y))$ uniquely determines Y. Thus, Equation (15) now becomes

$$d_{\mathrm{TV}}(((h(Y, Z)), Y), (h(Y, Z), U_{m-r})) \leq \epsilon \tag{16}$$

where U_{m-r} is independent of (Y, Z). Henceforth, we will rename the map $(Y, Z) \mapsto f(M_1(Y) + M_2(Z))$ as $f(Y, Z)$. The above argument has established the following claim.

Claim A.4. *For Y, Z defined as above, we have that*

$$d_{\mathrm{TV}}((S(X) \mid Y = y), S(X)) \leq O(\sqrt{\epsilon})$$

with probability $1 - \sqrt{\epsilon}$ over $y \sim Y$.

We first show the following lemma.

Lemma A.5. *The distribution of $A(X)$ is $2^{-\Omega(n^{0.1})}$-close to U_ℓ.*

Proof. Recall that $S(X) = h(Y, Z)$ where h is a quadratic map with output length $d_1 = 100\Delta$. Expand h as $h(Y, Z) = h_1(Y) + h_2(Y, Z) + h_3(Z)$ where h_2 is a bilinear map. Let $\ell_1, \ldots \ell_j$ be the linear functions of Y that are used in h_2. Clearly $j \leq 100\Delta \cdot r \leq O(n^{0.2})$.

Quadratic conditioning: Cohen-Tal conditioning Our goal is to fix functions of Y so that S is only a function of Z. We want to do this fixing without affecting the distribution of S, and also retaining entropy in Y. Claim A.4 allows us to achieve the former and the fact that S is a short quadratic map allows us to achieve the latter.

First, we consider the subsource of X obtained by fixing the value $\ell_1, \dots \ell_j$. With probability $1 - \epsilon^{1/4}$ over this fixing, the subsource $X' = f(Y', Z)$ has $S(X') = h(Y', Z) = h_1(Y') + h_3(Z)$ (h has no bilinear terms that depend on Y') and $Y' = M(U_{m-r-j})$ for some full rank linear map M. Absorbing the linear map M into the source and overloading notation (redefining $f, h, Y \leftarrow M^{-1}(Y')$), we have that the subsource we reached is of the form $X = f(Y, Z)$ with $S(X) = h(Y, Z) = h_1(Y) + h_3(Z)$ and $H_\infty^{2\epsilon^{3/4}}(X) \geq n - \Delta - j - O(\log(1/\epsilon))$. Furthermore, $Y = U_{m-r-j}$.

From Corollary 4.12, we have a function $\text{CTPartition} : \mathbb{F}_2^{m-r-j} \to \mathbb{F}_2^{m-r-j-\Omega((m-r-j)/d_1)}$ such that h_1 is constant on $\text{CTPartition}^{-1}(b)$ for any b. With probability $1 - O(\epsilon^{1/4})$ over fixing $\text{CTPartition}(Y)$, we land in a subsource of the form $X' = f(Y', Z)$ with $H_\infty^{O(\sqrt{\epsilon})}(X') \geq c(m - r - j)/d_1) + r - O(\log(1/\epsilon))$ for some constant $c > 0$. On this subsource, we have $S(X) = h_3(Z)$. Let $\overline{k} = \Omega((m - r - j)/d_1)$. From our choice of parameters, we have that $\overline{k} = \Omega(n^{0.9})$. Redefining the notation, our final source has the structure $X = f(Y, Z)$ with (1) $S(X) = h_3(Z)$, (2) $Y = U_{\overline{k}}$, (3) $H_\infty^{O(\sqrt{\epsilon})}(X) \geq \overline{k} + r - O(\log(1/\epsilon))$.

Also, since we only fixed functions of Y throughout the conditioning, Claim A.4 implies that $H_\infty^{O(\sqrt{\epsilon})}(S(X)) \geq 0.99 d_1$.

Structure theorem Now, we apply Corollary 5.2 to X. Recall that X is $n^{0.2}$-pure. We set $s = \overline{k}, t = 10n^{0.2}, r = n^{0.6}$ and entropy parameter $k = \overline{k} + |Z| - O(\log(1/\epsilon))$. Thus, we have that X is $2^{-\Omega(n^{0.1})}$-close to a convex combination of subsources of the form

$$X' = g(Y', Z) = g_1(Y') + g_2(Y', Z) + g_3(Z)$$

such that (1) $\min_{a \neq 0} H_\infty(\partial_a g_1(.)) \geq n^{0.6}$ and (2) $Y' = U_{s'}$ for $s' \geq \overline{k}/2$. Importantly, Z is left unchanged.

Since $S(X)$ depends only on Z, its distribution is left unchanged. We can now apply Lemma 6.1 with $k = \Omega(n^{0.6})$, $\epsilon = 2^{-\Omega(n^{0.1})}$, $d = d_1 = O(n^{0.1})$ and $\Delta = n^{0.1}$. We use the linear seeded extractor from Theorem 4.6 (we truncate the output to $0.1n^{0.1}$ bits). Note that the entropy requirement of CGLExt is satisfied as $r \geq n^{0.6} \gg \log(n/\epsilon)$. Thus, we have that $A(X)$ is $2^{-\Omega(n^{0.1})}$-close to U_ℓ for all the subsources in the convex combination. Adding up all the TV distances due to various conditionings, we still have that that A is $2^{-\Omega(n^{0.1})}$ close to U_ℓ. This completes the proof of the lemma. □

Lemma A.6. *The probability that there exists an $i \in \mathbb{F}_2^{d_2}$ such that $A(X) = \text{LExt}(X, i)$ is at most $2^{-\Omega(n^{0.1})}$.*

Proof. Throughout the proof, we will set $\epsilon = 2^{-\Omega(n^{0.1})}$ The proof uses the same idea as the proof of Lemma A.5. We will argue that $\Pr[A(X) = \text{LExt}(X, i)] \leq 2^{-\Omega(n^{0.1})}$ for a fixed i. The final claim follows by a union bound on $\text{poly}(n)$ choices of i.

Again, we assume that $X = f(Y, Z)$ with $S(X) = h(Y, Z)$ satisfying Claim A.4. Since LExt is a linear seeded extractor, note that $\text{LExt}(X, i)$ is also given by a quadratic map $g(Y, Z)$ which has $\ell = 0.1n^{0.1}$ output bits. We first write $g(Y, Z) = g_1(Y) + g_2(Y, Z) + g_3(Z)$. Our aim is to fix functions of Y to make both $S(X) = h(Y, Z)$ and $\text{LExt}(X, i) = g(Y, Z)$ solely functions of Z.

Note that one can treat (h, g) as a quadratic map with $d_1 + \ell \leq O(n^{0.1})$ output bits and apply the quadratic conditioning argument again. After conditioning on appropriate functions of Y, the source obtains the form $X = f(Y, Z)$ with $Y = U_{\overline{k}}$, $\overline{k} \geq n^{0.9}$, $S(X) = h_3(Z)$ and $\text{LExt}(X, i) = g_3(Z)$. We now again repeat the application of the structure theorem (Corollary 5.2) as in the proof of Lemma A.5 to prepare the source to be of the form required by Lemma 6.1. The proof is then completed by applying Lemma 6.1.

Lemma 6.1 guarantees that with high probability over fixing $z \sim Z$, the function $A = \mathrm{CGLExt}(f(Y,z), h_3(z))$ looks $2^{-\Omega(n^{0.1})}$ close to U_ℓ. However, note that fixing Z fixes $\mathrm{LExt}(X,i)$ (as we had earlier made $\mathrm{LExt}(X,i)$ solely a function of Z). Thus, after fixing Z, we have that A is close to uniform and $\mathrm{LExt}(X,i)$ is constant. This implies that the probability that $A = \mathrm{LExt}(X,i)$ is at most $2^{-\ell} + 2^{-\Omega(n^{0.1})} \leq 2^{-\Omega(n^{0.1})}$. Now, applying a union bound over $i \in \mathbb{F}_2^{d_2}$ completes the proof. □

Given, the two lemmas above, the proof that QDisp is an extractor is straightforward. First, Lemma A.6 implies that with probability $1 - 2^{-\Omega(n^{0.1})}$, execute Step 5 of Algorithm 3 and output A_1. Lemma A.5 establishes that this bit is $2^{-\Omega(n^{0.1})}$ close to uniform. □

We now complete the proof of the main theorem Theorem A.1.

Proof of Theorem A.1. The theorem follows immediately by combining Theorem A.2 and Theorem A.3. □

B Affine Extraction

We first show that the construction in Algorithm 1 is also an affine correlation breaker (see Definition 3.14). This argument is not new to our paper, it was already sketched in Section 2.3 of [CL22] and we include a formal proof for completeness.

Theorem B.1. *For any $\delta \in (0,1)$ and positive integers t, a, m, the function* QuadCB *(from Algorithm 1) with parameters instantiated the same way as in Theorem 7.2 is also a $(t, k, O(\delta))$-affine correlation breaker.*

Proof. Our proof is much easier than the quadratic analogue in Theorem 7.2. We first do a step of affine conditioning.

Affine conditioning Recall that $W, W^1, \ldots, W^t$ all have exactly t output bits and are linear functions of the seed bits V. From Lemma 3.2 (applied to $\ell, \ell^{[t]}$, we have that there exists subspaces A, B such that (1) $A + B = V = \mathbb{F}_2^k$, (2) $\dim(B) \leq d(t+1)$, (3) $\dim(A) \geq k - d(t+1)$ and (4) $\ell(A), \ell^1(A), \ldots, \ell^t(A)$ is constant over the support of A.

Recall initially we have that $X = f(V) = c + M(V) = c + M(A) + M(B)$ for constant vector c and linear map M. After the conditioning above, we have that $S_1, S_1^{[t]}, W, W^{[t]}$ are solely functions of B. Thus, we have that

$$R = \mathrm{LExt}(c + M(A) + M(B), S_1(B)) = \mathrm{LExt}(c + M(A), S_1(B)) + \mathrm{LExt}(M(B), S_1(B)). \tag{17}$$

Since M is a full rank map and $\dim(A) \geq k - d(t+1)$, we have that $c + M(A)$ satisfies the entropy requirement of LExt. Also, $S_1(B)$ is uniform. Thus, we have that with probability $1 - \delta$ over fixing $S_1(B)$, it holds that R is uniform. In fact, we can also fix $S_1^{[t]}(B)$ without affecting the uniformity of R. After this fixing we have that $H_\infty(W) \geq d - (t+1)d_0 - O(\log(1/\epsilon))$ (with probability $1 - \delta$ over the fixing) from min-entropy chain rule.

Now, we make $R, R^{[t]}$, independent of $W, W^{[t]}$. To do that we have to also fix the second term in Equation (17) and the analogous terms in $R^{[t]}$. These are solely functions of B, so fixing them retains the uniformity of R. Since LExt outputs d_1 bits, fixing these functions of B makes the entropy of W now $d - (t+1)(d_0 + d_1) - O(\log(1/\delta)$ with probability $1 - O(\delta)$ over all the fixings. At this point, we have that $\{R, R^{[t]}\}$ is independent of $\{W, W^{[t]}\}$, R is uniform and W satisfies the entropy requirement of CB. Now, we apply to correlation breaking guarantee Theorem 4.13 to complete the proof. □

We are now ready to complete the proof of Theorem 9.3.

Proof of Theorem 9.3. Let our affine source be $X = f(V)$ for $V = U_k$, We will first argue that the random variable $W = (\mathrm{RCombExt}(X, i))_{i\in[N]}$ has the following property: $W_i = U_d$ with probability $1 - 1/n^2$ over $i \sim [N]$. This follows from Item 2 in Theorem 8.1 and Lemma 4.11. We also have that $W_i = \ell^i(V)$ for some affine functions ℓ^i.

Let $Q \subseteq [N]$ be the set of indices where $W_i = U_d$. Consider the random variable

$$S = (\mathrm{QuadCB}(X, W^i, i))_{i\in[N]}.$$

We will argue that for all sets $T \subseteq Q$ with $|T| \le t$, we have that $|\mathbb{E}[\chi_T(S)]| \le O(\delta)$. This is immediate from Theorem B.1 and the fact that X, W satisfy the premise (by our choice of parameters) and the indices $i \in [N]$ form valid advice strings.

Now, applying Fact 4.9 implies that S is N^{-t}-close to a (q, t)-NOBF source with $q = N/n^2$ and $t = (m \log n)^{C_2}$ chosen earlier. The theorem statement now follows from Theorem 3.13. □

C Additional Proofs

We give the proof of the probabilistic construction of the rank combiner.

Proof of Theorem 8.4. Let $G_1, \ldots, G_n \in \mathbb{F}_2^{k\times k}$ such that $\dim\{G_i a\}_{i\in[n]} \ge d$ for all $a \ne 0$. Let r be the target rank, we fix it later. We first argue that a uniform random matrix $\mathcal{W} \in \mathbb{F}_2^{m\times n}$ satisfies

$$\mathrm{rank}_{\mathbb{F}_2}\left(\sum_{i=1}^{n}(c^\top \mathcal{W})_i G_i\right) \ge r \tag{18}$$

for all non-zero $c \in \mathbb{F}_2^m$ with high probability over $\mathcal{W}$. Indeed for each $c \ne 0$, the vector $c^\top W$ is a random vector in $\mathbb{F}_2^n$. Thus, it suffices to argue that a random vector X satisfies

$$\mathrm{rank}_{\mathbb{F}_2}\left(\sum_{i=1}^{n} X_i G_i\right) \ge r$$

with high probability over X and then union bound over non-zero c.

This is what we will do. Lower bounding the rank of a matrix is equivalent to upper bounding the size of its null space. For any $a \ne 0$, we have that the matrix $G_a = \{G_i a\}_{i\in[n]}$ has rank at least d. That means with probability at most 2^{-d}, it holds that $G_a X = 0$ for uniform random vector X. Thus, we have that

$$\mathop{\mathbb{E}}_{X}[|\{a \mid (\sum_{i=1}^{n} X_i G_i)a = 0\}|] \le \sum_{a\in\mathbb{F}_2^k} \Pr[G_a X = 0] \le (2^k - 1)/2^d.$$

By Markov's inequality, with probability at least $1 - \delta$, we have $\mathrm{rank}_{\mathbb{F}_2}\left(\sum_{i=1}^{n} X_i G_i\right) \ge d - \log(1/\delta)$. Union bounding over c, we have that with probability at least $1 - \delta \cdot 2^m$, Equation (18) holds. Choosing $\delta = \epsilon/2^{m+1}$ and $r = d - (m+1) - \log(1/\epsilon)$, we have that Equation (18) holds with probability $1 - \epsilon/2$. Now sample N independent random matrices $W_1, \ldots, W_N$ for $W = O(nk^2/\epsilon)$. Note that the total number of choices of $G_1, \ldots, G_n$ is atmost 2^{nk^2}. Thus, by a Chernoff bound and the bound we just established, we have that $W_1, W_2, \ldots, W_N$ is a $(d, d - m - \log(1/\epsilon), \epsilon)$-rank combiner with $N = O(nk^2/\epsilon)$. □