Extractors for Polynomial Sources over $\mathbb{F}_2$

Eshan Chattopadhyay
Cornell University, Ithaca, NY, USA

Jesse Goodman
Cornell University, Ithaca, NY, USA

Mohit Gurumukhani
Cornell University, Ithaca, NY, USA

Abstract

We explicitly construct the first nontrivial extractors for degree $d \geq 2$ polynomial sources over $\mathbb{F}_2$. Our extractor requires min-entropy $k \geq n - \frac{\sqrt{\log n}}{(\log \log n/d)^{d/2}}$. Previously, no constructions were known, even for min-entropy $k \geq n - 1$. A key ingredient in our construction is an *input reduction lemma*, which allows us to assume that any polynomial source with min-entropy k can be generated by $O(k)$ uniformly random bits.

We also provide strong formal evidence that polynomial sources are unusually challenging to extract from, by showing that even our most powerful general purpose extractors cannot handle polynomial sources with min-entropy below $k \geq n - o(n)$. In more detail, we show that *sumset extractors* cannot even *disperse* from degree 2 polynomial sources with min-entropy $k \geq n - O(n/\log \log n)$. In fact, this impossibility result even holds for a more specialized family of sources that we introduce, called *polynomial non-oblivious bit-fixing (NOBF) sources*. Polynomial NOBF sources are a natural new family of algebraic sources that lie at the intersection of polynomial and variety sources, and thus our impossibility result applies to both of these classical settings. This is especially surprising, since we *do* have variety extractors that slightly beat this barrier - implying that sumset extractors are not a panacea in the world of seedless extraction.

2012 ACM Subject Classification Theory of computation → Expander graphs and randomness extractors

Keywords and phrases Extractors, low-degree polynomials, varieties, sumset extractors

Digital Object Identifier 10.4230/LIPIcs.ITCS.2024.28

Related Version *Full Version*: `https://eccc.weizmann.ac.il/report/2023/140/`

Funding Supported by a Sloan Research Fellowship and NSF CAREER Award 2045576.

Acknowledgements We want to thank Michael Jaber for helpful discussions.

1 Introduction

Randomness is a very important resource in computation. It is widely used in theoretical and practical implementations of algorithms, distributed computing protocols, cryptographic protocols, machine learning algorithms, and much more [28]. Unfortunately, the randomness produced in practice is not of the highest quality, and the corresponding distribution over bits is often biased and has various correlations [21]. To overcome this, an extractor is used to convert this biased distribution to a uniform distribution. The extractors used in practice are based on unproven theoretical assumptions, and so the theoretical study of constructing efficient extractors is important. Extractors usually come in two flavors: seeded and seedless extractors. We focus here on the latter, and whenever we mention *extractor* this is what we mean. Towards this end, let's formally define extractors for a class of distributions:

▶ **Definition 1** (Extractor). *A function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *is called an* ε*-extractor for a class* $\mathcal{X}$ *of distributions over* $\{0,1\}^n$ *if for all* $\mathbf{X} \in \mathcal{X}$*,*

$$|\mathsf{Ext}(\mathbf{X}) - \mathbf{U}_m| \leq \varepsilon,$$

where $|\cdot|$ *denotes statistical distance and* $\mathbf{U}_m$ *is the uniform random variable.*

In this paper and throughout, we use min-entropy as our measure for randomness: For a source $\mathbf{X}$ with support Ω, we define its min-entropy $H_\infty(\mathbf{X}) = -\log(\max_{x\in\Omega} \Pr(\mathbf{X} = x))$. Note that for $\mathbf{X} \sim \{0,1\}^n$, $0 \leq H_\infty(\mathbf{X}) \leq n$.

It is well known that there do not exist extractors for arbitrary distributions, even when they have a lot of randomness (min-entropy $= n - 1$). To overcome this, a long body of work has been dedicated to extracting randomness from distributions that not only have some min-entropy, but also exhibit structure. Two such widely studied classes of structured sources are: (1) *Samplable* sources, i.e., sources generated by low complexity classes such as AC^0 circuits, decision trees, local sources, branching programs, and more [36, 22, 12, 38, 7, 1], and (2) *Recognizable* sources, i.e., sources that are uniform over the zeroes of a function from some low complexity class mentioned above [24]. This study has provided insight into the structure of such complexity classes, and there is an argument to be made that in nature, most distributions are likely to be generated by such low complexity classes. In this paper, we study *algebraic sources* which are samplable and recognizable sources with respect to *low degree multivariate polynomials over* $\mathbb{F}_2$ (another such natural computational model).

Algebraic sources

Two different flavors of algebraic sources have been studied: variety sources and polynomial sources. The task of constructing extractors for these sources, apart from being a fundamentally important task to help us gain structural insights into polynomials, also has other nice motivations.

Extractors for polynomial sources (over $\mathbb{F}_2$) with $\mathrm{poly}(\log n)$ degree would immediately yield extractors for sources sampled by $\mathsf{AC}^0[\oplus]$ circuits, based on well-known approximations of such circuits by polynomials [31, 35].[1] To the best of our knowledge, there are no known nontrivial explicit extractors for sources sampled by such circuits.

Extractors for variety sources (over $\mathbb{F}_2$), on the other hand, have important applications in circuit lower bounds. If one can construct explicit extractors (or even dispersers - Definition 3.3) against degree 2 varieties with min-entropy $0.01n$, or against degree $n^{0.01}$ varieties with min-entropy $0.99n$, then one immediately gets new state-of-the-art circuit lower bounds [17, 18].

With these motivations in hand, let's proceed to formally define these sources. Both polynomial and variety sources are parameterized by min-entropy k, degree d and finite field $\mathbb{F}_q$. When $q = 2$, it is typical to identify $\mathbb{F}_2^n$ with $\{0,1\}^n$.

▶ **Definition 2** (Polynomial sources). *A degree* d *polynomial source* $\mathbf{X} \sim \mathbb{F}_q^n$ *is associated with a polynomial map* $P = (p_1, \ldots, p_n)$ *where each* $p_i : \mathbb{F}_q^m \to \mathbb{F}_q$ *is a polynomial of degree at most* d*. Then,* $\mathbf{X} = P(\mathbf{U}_m)$ *where* $\mathbf{U}_m$ *is the uniform distribution over* $\mathbb{F}_q^m$*.*

[1] $\mathsf{AC}^0[\oplus]$ circuits are constant depth, polynomial sized circuits with unbounded fan-in AND, OR, NOT, and PARITY gates.

▶ **Definition 3** (Variety sources). *Here, each source* $\mathbf{X} \sim \mathbb{F}_q^n$ *has associated polynomials* $p_1, \ldots, p_m : \mathbb{F}_q^n \to \mathbb{F}_q$. $\mathbf{X}$ *is uniform over the set of common zeroes of these polynomials* $V = \{x \in \mathbb{F}_q^n : \forall i \in [m] : p_i(x) = 0\}$.

In this paper, we introduce and study a natural class of sources that is a subclass of polynomial sources and the widely studied NOBF (non-oblivious bit-fixing) sources. Surprisingly, as we will soon see, it is also a subclass of variety sources (Claim 5).

▶ **Definition 4** (Polynomial NOBF sources). *A degree* d *polynomial NOBF source* $\mathbf{X} \sim \mathbb{F}_q^n$ *with* $H_\infty(\mathbf{X}) = k$ *must have* k *as an integer and have the following structure:*

1. *There exists* $G \subset [n]$ *with* $|G| = k$ *that we call the good coordinates of* $\mathbf{X}$. *These good coordinates in* $\mathbf{X}$ *are sampled uniformly and independently at random.*
2. *Each coordinate outside* G *is a deterministic function of the* k *good coordinates of* $\mathbf{X}$. *Moreover, each such deterministic function is a degree* d *polynomial.*

We make a basic observation regarding polynomial NOBF sources (this observation seems to apply to most classes of samplable, NOBF, and recognizable sources):

▷ Claim 5. If $\mathbf{X} \sim \mathbb{F}_q^n$ is a degree d polynomial NOBF source, then it is also a degree d polynomial source and a degree d variety source.

Proof. Let $H_\infty(\mathbf{X}) = k$ and the bad positions be specified by polynomials $p_1, \ldots, p_{n-k}$. As the k good positions are degree 1 polynomials over $x_1, \ldots, x_k$ and the $n-k$ bad positions are degree d polynomials over $x_1, \ldots, x_k$, $\mathbf{X}$ is indeed a polynomial source. Without loss of generality, assume that the first k positions of $\mathbf{X}$ are the good positions and last $n-k$ positions are the bad positions. Consider the following set of polynomial equations over variables $y_1, \ldots, y_n$:

$$\begin{aligned} y_{k+1} - p_1(y_1, \ldots, y_k) &= 0 \\ &\vdots \\ y_n - p_{n-k}(y_1, \ldots, y_k) &= 0 \end{aligned}$$

Note that $\mathbf{X}$ is uniform over the variety defined by these equations, and thus is also a variety source. ◁

Related work

Degree 1 polynomial / variety sources, i.e., affine sources have been widely studied both over $\mathbb{F}_2$ and other $\mathbb{F}_q$ [5, 16, 30, 2, 13, 25, 34, 39, 4, 6, 26, 8, 19]. Recently, [27] constructed affine extractors over $\mathbb{F}_2$ with asymptotically optimal dependence on min-entropy.

[15] initiated the study of extractors for polynomial sources. Their extractors worked when either when $q \geq \text{poly}(n,d)^{O(k)}$ or when field has characteristic $\geq \text{poly}(n,k,d)$. [3] used sum product estimates and constructed extractors for degree 2 polynomial sources when $q \geq O(1)$ and min-entropy is $\geq O(n)$. They also constructed dispersers for arbitrary multilinear polynomials over $\mathbb{F}_4$ with min-entropy $\geq n/2 + O(1)$. Extractors for variety sources were first constructed by [14]. They constructed extractors for when either $q \geq \exp(n)$ or $q \geq \text{poly}(d)$ and min-entropy $\geq O(n)$. Recently, [19] constructed extractors for images of varieties over $\mathbb{F}_q$ when $q \geq \text{poly}(n,d)$ with no min-entropy restrictions (they define degree parameter d differently). Over $\mathbb{F}_2$, [33] constructed extractors for degree n^{δ_1} varieties with min-entropy $\geq n - n^{\delta_2}$ for arbitrary $\delta_1 + \delta_2 < \frac{1}{2}$. Using correlation bounds against low degree polynomials, [10, 24] constructed extractors for constant degree d variety sources over $\mathbb{F}_2$ with min-entropy $\geq (1 - c_d)n$ where c_d is a tiny constant that depends on d.

We reiterate that before our work, no extractors were constructed for polynomial sources over $\mathbb{F}_2$ even for min-entropy $k \geq n - 1$!

1.1 Our results

We construct the first nontrivial extractors for polynomial sources over $\mathbb{F}_2$.

▶ **Theorem 6** (Explicit extractor for polynomial sources, informal version of Theorem 4.10)**.** *Let $\varepsilon > 0$ be an arbitrary constant. For all $d \in \mathbb{N}$, there exists an explicit ε-extractor $\mathsf{Ext} : \mathbb{F}_2^n \to \mathbb{F}_2^{\Omega(\log\log n)}$ for degree d polynomial sources over $\mathbb{F}_2$ with min-entropy $k \geq n - \Omega\left(\frac{\sqrt{\log n}}{(\log\log n/d)^{d/2}}\right)$.*

Prior to our work, there were no known constructions of extractors for polynomial sources over $\mathbb{F}_2$ that worked for degree $d > 1$ and min-entropy $k = n - 1$. Indeed, all prior constructions required the field size q to be large, or the degree to be $d = 1$.

As polynomial sources can have arbitrarily large input length, it's not clear what is the size of the class of degree d polynomial sources. Therefore, it is unclear if an extractor should even exist for this class. To get around this problem, we come up with an input reduction technique that allows us to bound the number of inputs to the polynomial source by the min-entropy of the source. We view this as our main technical contribution, and it is the key ingredient behind our explicit extractor in Theorem 6.

▶ **Lemma 7** (Input reduction, informal version of Theorem 4.1)**.** *Every degree d polynomial source with min-entropy k (and an arbitrary number of input variables) is 2^{-k}-close to a convex combination of polynomial sources of min-entropy $k - 2$ and $O(k)$ input variables.*

Recall that $\mathbf{X}$ is a convex combination of distributions $\{\mathbf{Y}_i\}$ if there exist probabilities $\{p_i\}$ summing up to 1 such that $\mathbf{X} = \sum_i p_i \mathbf{Y}_i$. It is well known that an ε-extractor for $\{\mathbf{Y}_i\}$ will also be an ε-extractor for $\mathbf{X}$. Hence, this lemma reduces the task of extracting from polynomial sources with an arbitrary number of input variables to the task of extracting from polynomial sources with $O(k)$ input variables.

We also show negative results for polynomial NOBF sources against sumset extractors. Sumset extractors are extremely powerful and can be used to extract not only from sumset sources but also, using reductions to sumset sources, from many well studied models of weak sources such as degree 1 polynomial / variety sources (affine sources), class of two independent sources, sources generated by branching programs, sources generated by AC^0 circuits and many more [9]. Let's first define sumset sources:

▶ **Definition 8.** *A source $\mathbf{X}$ is a (k, k) sumset source if it is of the form $\mathbf{A} + \mathbf{B}$, where $\mathbf{A}, \mathbf{B}$ are independent distributions on $\{0,1\}^n$ with $H_\infty(\mathbf{A}) \geq k, H_\infty(\mathbf{B}) \geq k$, and $+$ denotes bitwise xor.*

Note that $k \leq H_\infty(\mathbf{X}) \leq 2k$ and so, $H_\infty(\mathbf{X}) = \Theta(k)$. When we write $H_\infty(\mathbf{X})$ is a sumset source of min-entropy k, we actually mean $\mathbf{X} = \mathbf{A} + \mathbf{B}$ where $H_\infty(\mathbf{A}) \geq k, H_\infty(\mathbf{B}) \geq k$. Recently, sumset extractors with the smallest possible dependence on min-entropy ($O(\log n)$) were constructed [27]. A natural question is whether various algebraic sources can be reduced to sumset sources. We show here that sumset extractors cannot even disperse (see Definition 3.3) let alone extract below certain min-entropy against quadratic NOBF sources.

▶ **Theorem 9** (Sumset extractor lower bound, informal version of Theorem 5.16)**.** *Sumset extractors cannot be used to disperse from degree 2 polynomial NOBF sources with min-entropy $n - O\left(\frac{n}{\log\log n}\right)$.*

As polynomial NOBF sources are both variety *and* polynomial sources, this also implies that sumset extractors cannot be used to extract from degree 2 variety sources or degree 2 polynomial sources over $\mathbb{F}_2$ with min-entropy below $n - O\left(\frac{n}{\log\log n}\right)$. For degree 2 variety

sources over $\mathbb{F}_2$, one can use a sumset extractor to extract above min-entropy $n - O\left(\frac{n}{\log n}\right)$ [11]. Moreover, using correlation bounds, one can construct explicit extractors against degree 2 varieties with min-entropy $(1-c)n$ for very small constant c [24]. Hence, the above result shows that sumset extractors cannot be used to get better extractors than what we get using correlation bounds against low degree polynomials. We find this surprising as it implies the generalized inner product function is a better extractor for degree 2 variety sources than any optimal (blackbox) sumset extractor.

Organization

The rest of the paper is organized as follows. In Section 2, we give an overview of our proofs. In Section 3, we provide basic definitions and useful properties that we use later. In Section 4, we prove Lemma 7, our input reduction argument and using it prove Theorem 6, the construction of polynomial source extractor. In Section 5, we prove Theorem 9, limitations of the sumset extractor against quadratic NOBF sources. In Section 6, we conclude with various open problems.

2 Overview of our techniques

In this section, we sketch the proofs of all our main results.

2.1 Existential results

To warm up, it is not clear whether a random function is a good extractor for degree d polynomial sources. Usually such proofs proceed by arguing that for a fixed source of min-entropy k, a random function is an ε-extractor with probability at least $1 - 2^{-2^k\varepsilon^2}$. Then, one can do a union bound over total number of sources in the class to obtain that a random function is a good extractor. The main issue that arises for polynomial sources is that the number of input variables to the polynomials can be arbitrary and hence, it's not clear what is the size of this class. To overcome this difficulty, we use our input reduction lemma (Lemma 7). Using this, it suffices to consider degree d polynomial sources with $O(k)$ inputs. This class of polynomial sources has size $2^{O(k)^d \cdot n}$. Thus, the the earlier union bound based argument now works out:

▶ **Lemma 2.1** (Informal version of Lemma 4.6). *A random function with $O(k)$ output bits is a $2^{-\Omega(k)}$ extractor for degree d polynomial sources over $\mathbb{F}_2$ with min-entropy $k \geq d \log n$.*

2.2 Input reduction

We will know sketch the proof for the input reduction lemma that was useful above (in fact, it will also be very useful for the explicit construction).

We begin by showing that for a polynomial map $f(\mathbf{U}_m)$, there exists a linear function L (acting on the same set of variables as f) and a fixing b of L such that $f(\mathbf{U}_m) \approx_\varepsilon f(\mathbf{U}_m)|L = b$. In fact, we show the stronger claim that most such fixings b work:

▶ **Lemma 2.2** (Existence of affine white-box PRGs). *For any polynomial map $f : \mathbb{F}_2^m \to \mathbb{F}_2^n$ and $\varepsilon > 0$, there exists a linear function $L : \mathbb{F}_2^m \to \mathbb{F}_2^{m-\ell}$ with $\ell = n + 3\log(1/\varepsilon)$ such that*

$$|f(\mathbf{U}_m) \circ L(\mathbf{U}_m) - f(\mathbf{U}_m) \circ \mathbf{U}_{m-\ell}| \leq 2\varepsilon,$$

where $\mathbf{U}_m$ and $\mathbf{U}_{m-\ell}$ are independent.

Proof sketch. We show a random L works. Indeed, by definition:

$$|f(\mathbf{U}_m) \circ L(\mathbf{U}_m) - f(\mathbf{U}_m) \circ \mathbf{U}_{m-\ell}| = \mathbb{E}_{z \sim f(\mathbf{U}_m)} \left[|(L(\mathbf{U}_m) \mid f(\mathbf{U}_m) = z) - \mathbf{U}_{m-\ell}| \right]$$

We now apply the min-entropy chain rule (see Lemma 3.2 for details) to infer that with high probability over fixings of f to z, the input distribution to L will have high min-entropy. Indeed, there will exist some distribution $\mathbf{X}$ with min-entropy at least $k = m - n - \log(1/\varepsilon) = m - \ell + 2\log(1/\varepsilon)$ such that

$$\mathbb{E}_{z \sim f(\mathbf{U}_m)} \left[|(L(\mathbf{U}_m) \mid f(\mathbf{U}_m) = z) - \mathbf{U}_{m-\ell}| \right] \leq \varepsilon + |L(\mathbf{X}) - \mathbf{U}_{m-\ell}|$$

As L was initially chosen as a random function, we apply the leftover hash lemma (see Corollary 3.6) to infer that $|L(\mathbf{X}) - \mathbf{U}_{m-\ell}| \leq \varepsilon$ as desired. ◀

Note that this lemma already yields an input reduction to a single polynomial source with $O(n)$ variables. This can be done by fixing output of L to some z that preserves small distance between the two distributions. Once we fix output of L, we induce $m - \ell$ affine constraints on the variables. As polynomial sources are closed under affine restrictions, the resulting polynomial map is still a degree d polynomial map and the resulting distribution is still close enough to the original one as desired. However, we can do better.

We will first prove the following helpful claim. This claim shows there exists a way to map every min-entropy k source to a source over $O(k)$ bits and almost full min-entropy.

▷ Claim 2.3. Let $\mathbf{X} \sim \{0,1\}^n$ be a polynomial source with min-entropy at least $k > 0$. Then there exists a function $S : \{0,1\}^n \to \{0,1\}^{k+1}$ such that $S(\mathbf{X})$ has min-entropy at least $k - 1$.

This claim is actually true for arbitrary sources $\mathbf{X}$ and we prove it by a simple case analysis on probabilities of the smallest two elements in support of $\mathbf{X}$ (see Claim 4.3 for further details). Using this claim, we are ready to sketch the proof of our main lemma that will help us achieve the input reduction:

▶ **Lemma 2.4.** *For any polynomial source $f : \mathbb{F}_2^m \to \mathbb{F}_2^n$ where $f(\mathbf{U}_m)$ has min-entropy at least k, there exists a linear function $L : \mathbb{F}_2^m \to \mathbb{F}_2^{m-O(k)}$ such that*

$$\Pr_{b \sim L(\mathbf{U}_m)} \Big[H_\infty(f(\mathbf{U}_m) \mid L(\mathbf{U}_m) = b) \geq k - 2 \Big] \geq 1 - 2^{-k}.$$

Proof sketch. Let $S : \mathbb{F}_2^n \to \mathbb{F}_2^{k+1}$ be a function guaranteed to exist by above claim so that $S(f(\mathbf{U}_m))$ has min-entropy at least $k-1$. Using data processing inequality (see Claim 4.4), it suffices to show that $S(f(\mathbf{U}_m))$ has high enough min-entropy with high probability over fixing $L(\mathbf{U}_m)$. Let $\mathbf{Y} = \mathbf{U}_m$. By Lemma 2.2, there exists $L : \mathbb{F}_2^m \to \mathbb{F}_2^{m-\ell}$ with $\ell = k + 2 + 3\log(1/\varepsilon)$ such that:

$$|S(f(\mathbf{Y})) \circ L(\mathbf{Y}) - S(f(\mathbf{Y})) \circ \mathbf{U}_{m-\ell}| \leq \varepsilon$$

This implies

$$\mathbb{E}_{b \sim f(\mathbf{U}_m)} \left[|(S(f(\mathbf{Y})) \mid f(\mathbf{Y}) = b) - S(f(\mathbf{U}_m))| \right] \leq \varepsilon$$

By Markov's inequality, we infer that

$$\Pr_{b \sim f(\mathbf{U}_m)} \left[|(S(f(\mathbf{Y})) \mid f(\mathbf{Y}) = b) - S(f(\mathbf{U}_m))| \geq \sqrt{\varepsilon} \right] \leq \sqrt{\varepsilon}$$

Setting $\varepsilon = 2^{-2k}$, every element in support of the distribution $S(f(\mathbf{Y} \mid L(\mathbf{Y}) = b))$ must occur with probability at most $2^{-k+1} + \sqrt{\varepsilon} \leq 2^{-k+2}$, and thus has min-entropy at least $k - 2$. The result follows. ◀

Notice that using this lemma, most fixings of L leave f with min-entropy at least $k-2$. These good fixings form a convex combination of such sources f. As argued earlier, such L induces $m - O(k)$ linear fixings on the input variables and hence the resulting polynomial map in each of these convex combinations is over $O(k)$ variables and has degree d as desired.

2.3 Explicit construction

We sketch here the proof of a slightly weaker result that illustrates our main idea.

▶ **Theorem 2.5** (Weaker version of Theorem 6). *For all constant degree $d \in \mathbb{N}$, there exists an explicit extractor* $\mathsf{Ext} : \mathbb{F}_2^n \to \mathbb{F}_2$ *for polynomial sources with min-entropy $k \geq n - O(\log\log n)$.*

Proof sketch. Fix degree d polynomial source $\mathbf{X}$ with m inputs and n outputs with min-entropy $n-g$ where $g = O(\log\log n)$. Consider a small length $t = 2g$ prefix of the output bits, and let this source be $\mathbf{X}_{pre}$. We observe that $\mathbf{X}_{pre}$ has min-entropy at least $t - g = t/2$ (see Claim 4.11). We now use our input reduction argument: Lemma 7 over $\mathbf{X}_{pre}$ to infer there exists a source $\mathbf{X}'_{pre}$ with $O(t)$ inputs such that $|\mathbf{X}'_{pre} - \mathbf{X}_{pre}| \leq 2^{-t}$. Hence, it suffices to construct an extractor for min-entropy $t/2$ degree d polynomial sources with $O(t)$ inputs and t outputs.

By Lemma 2.1, we know a random function over t bits will be an extractor for such sources. We exhaustively try all the 2^{2^t} functions from t bits to 1 bits as our candidate extractor. We brute force search over all the $2^{O(t)^d \cdot t}$ degree d polynomial sources with $O(t)$ inputs and t outputs. Then, for each of them, we check if it has enough min-entropy. If it does, we input the source into our candidate extractor and check if the output is close to uniform. We will eventually find a candidate extractor that will work for all such sources, and we output that function as our extractor.

The time required by the above procedure is $2^{2^t + O(t)}$. As $t = O(\log\log n)$, the above procedure indeed runs in $\mathrm{poly}(n)$ time. ◀

In our actual construction, we achieve better parameters by brute forcing over all r-wise independent functions (for very large r) as our candidate extractor instead of all functions. We also take advantage of the fact that the input reduction lemma actually reduces number of input variables to $O(k)$, making the class of polynomial sources that we have to brute force over even smaller. Together, these optimizations allow us to handle smaller min-entropy. See Theorem 4.10 for further details.

2.4 Impossibility results

All our impossibility results are against polynomial NOBF sources and hence apply (via Claim 5) to both polynomial sources and variety sources. We show that sumset extractors, arguably the most powerful general purpose extractors, cannot be used to even disperse from degree 2 polynomial NOBF sources below min-entropy $n - O\left(\frac{n}{\log\log n}\right)$ (Theorem 9). These results are formally proven in Section 5.3, and Section 5.4. We will use the following useful theorem to show this. This theorem states there exists some quadratic NOBF source which does not contain any sumset source of small min-entropy.

▶ **Theorem 2.6** (Informal version of Theorem 5.11). *There exists a degree 2 polynomial NOBF source $\mathbf{X} \sim \mathbb{F}_2^n$ with $H_\infty(\mathbf{X}) = n - O\left(\frac{n}{\log\log n}\right)$ such that for all $\mathbf{A}, \mathbf{B} \sim \mathbb{F}_2^n, H_\infty(\mathbf{A}) \geq \Omega(\log\log n), H_\infty(\mathbf{B}) \geq \Omega(\log\log n)$, it holds that support$(\mathbf{A})$ + support$(\mathbf{B}) \not\subset$ support$(\mathbf{X})$.*

Proof sketch. We take the $n-k$ bad bits in $\mathbf{X}$ to be random degree 2 polynomials. Say such $\mathbf{A}, \mathbf{B}$ exist and let C, D be projections of $\text{support}(\mathbf{A}), \text{support}(\mathbf{B})$ respectively onto the good bits of $\mathbf{X}$. Let $P : \mathbb{F}_2^k \to \mathbb{F}_2^{n-k}$ be the polynomial map of the bad bits. Then, it holds that $P(C) + P(D) = P(C + D) + y$ for some $y \in \mathbb{F}_2^{n-k}$. To simplify presentation, assume for this proof sketch that $y = 0^{n-k}$. We first observe the following:

▷ Claim 2.7 (Informal version of Claim 5.14). There exist affine subspaces U, V such that $P(U) + P(V) = P(U + V)$ and $|U| \geq |C|, |V| \geq |D|$.

Hence, without loss of generality, we can assume that C and D are affine subspaces. We now use a probabilistic argument to show such large affine subspaces C and D cannot exist with high probability for a random quadratic map:

▷ Claim 2.8 (Informal version of Claim 5.13). There exists a degree 2 polynomial map $P : \mathbb{F}_2^k \to \mathbb{F}_2^{k-n}$ such that for every pair of affine subspaces U, V, both of dimension $\geq \Omega(\log \log n)$, there exist $u \in U, v \in V$ such that $P(u) + P(v) \neq P(u + v)$.

Hence, the sumset property is violated and we get a contradiction. ◀

Using these, we finally present the proof of our lower bound result:

Proof sketch of Theorem 9. Let $\mathbf{X}$ be the degree 2 polynomial NOBF source with min-entropy $n - O\left(\frac{n}{\log \log n}\right)$ that doesn't contain any sumset of min-entropy $O(\log \log n)$. We apply a bipartite Ramsey bound (Corollary 5.20), to show that if a quadratic NOBF source doesn't contain sumsets where each of the two sets has size s, then it has small intersection with sumsets where each of the two sets has size $O(2^s)$ (see Lemma 5.17 for details). This implies $\mathbf{X}$ has very small intersection with sumset sources of min-entropy $\Omega(\log n)$. From this, we infer $\mathbf{X}$ is far away from any convex combination (see Definition 3.7) of sumset sources with min-entropy $\Omega(\log n)$. As sumset extractors below min-entropy $O(\log n)$ cannot exist (every function is constant on $\Omega(\log n)$ dimensional affine subspace), this shows we cannot use sumset extractors to even disperse against quadratic NOBF sources. See Theorem 5.16 for further details. ◀

3 Preliminaries

To simplify notation, we use $\circ$ to mean concatenation. Also, all logs in this paper are base 2.

3.1 Basic probability lemmas

Given two random variables $\mathbf{X}, \mathbf{Y}$, we let $|\mathbf{X} - \mathbf{Y}|$ denote their statistical distance, defined as

$$|\mathbf{X} - \mathbf{Y}| := \max_S [\Pr[\mathbf{X} \in S] - \Pr[\mathbf{Y} \in S]] = \frac{1}{2} \sum_z |\Pr[\mathbf{X} = z] - \Pr[\mathbf{Y} = z]| .$$

We write $\mathbf{X} \approx_\varepsilon \mathbf{Y}$ and say that $\mathbf{X}, \mathbf{Y}$ are ε-*close* if $|\mathbf{X} - \mathbf{Y}| \leq \varepsilon$, and we write $\mathbf{X} \equiv \mathbf{Y}$ if $|\mathbf{X} - \mathbf{Y}| = 0$. We will often use the fact that applying a function can only decrease the distance between two distributions:

▶ **Fact 3.1** (Data-processing inequality). *For any random variables $\mathbf{X}, \mathbf{X}' \sim X$ and function $f : X \to Y$, it holds that $|\mathbf{X} - \mathbf{X}'| \geq |f(\mathbf{X}) - f(\mathbf{X}')|$.*

We will utilize the well known fact that for any two distributions $\mathbf{X}, \mathbf{Y}$, with high probability, fixings of $\mathbf{Y}$ decrease min entropy of $\mathbf{X}$ by about $\log(|\text{support}(\mathbf{Y})|)$:

▶ **Lemma 3.2** (Min-entropy chain rule). *For any random variables $\mathbf{X} \sim X$ and $\mathbf{Y} \sim Y$ and $\varepsilon > 0$, it holds that $\Pr_{y \sim \mathbf{Y}}[H_\infty(\mathbf{X} \mid \mathbf{Y} = y) \geq H_\infty(\mathbf{X}) - \log |\mathit{support}(\mathbf{Y})| - \log(1/\varepsilon)] \geq 1 - \varepsilon$.*

3.2 Extractors

We start by defining *dispersers*, which are a weaker version of extractors. While the output of an extractor must look nearly uniform, the output of a disperser only needs to be *nonconstant*.

▶ **Definition 3.3** (Disperser). *A function* $\mathsf{Disp} : \{0,1\}^n \to \{0,1\}$ *is a disperser for a class of distributions* $\mathcal{X}$ *if for all* $\mathbf{X} \in \mathcal{X}$*, the set* $\{\mathsf{Disp}(\mathbf{X})\} = \{0,1\}$.

While the main purpose of this paper is to construct seedless extractors (for polynomial sources), it turns out that *seeded* extractors will also be useful in our arguments. We define them, below.

▶ **Definition 3.4** (Seeded extractor). *We say that a deterministic function* $\mathsf{sExt} : \{0,1\}^m \times \{0,1\}^s \to \{0,1\}^r$ *is a* (k,ε)-strong seeded extractor *if for any* $\mathbf{X} \sim \{0,1\}^m$ *with min-entropy at least* k*,*

$$\mathsf{sExt}(\mathbf{X}, \mathbf{Y}) \circ \mathbf{Y} \approx_\varepsilon \mathbf{U}_r \circ \mathbf{Y},$$

where $\mathbf{Y} \sim \{0,1\}^s$ *and* $\mathbf{U}_r \sim \{0,1\}^r$ *are independent uniform random variables. We say* sExt *is* linear *if the function* $\mathsf{sExt}(\cdot, y) : \{0,1\}^m \to \{0,1\}^r$ *is a degree* 1 *function, for all* $y \in \{0,1\}^s$.

One classic way to construct seeded extractors is via the following theorem.

▶ **Theorem 3.5** (Leftover Hash Lemma [20]). *Let* $\mathcal{H} = \{H : \{0,1\}^n \to \{0,1\}^m\}$ *be a 2-universal hash family with output length* $m = k - 2\log(1/\varepsilon)$*, meaning that for any* $x \neq y$*,* $\Pr_{H\sim\mathcal{H}}[H(x) = H(y)] \leq 2^{-m}$*. Then the function* $\mathsf{sExt} : \{0,1\}^n \times \mathcal{H} \to \{0,1\}^m$ *defined as*

$$\mathsf{sExt}(x, h) := h(x)$$

is a (k,ε)*-strong seeded extractor.*

▶ **Corollary 3.6.** *For any* $\varepsilon > 0$ *and* $m = k - 2\log(1/\varepsilon)$*, the function* $\mathsf{sExt} : \{0,1\}^n \times \mathbb{F}_2^{m\times n} \to \{0,1\}^m$ *defined as*

$$\mathsf{sExt}(x, L) := Lx$$

is a linear (k,ε)*-strong seeded extractor.*

Proof. It suffices to show that the family of all linear functions $L : \mathbb{F}_2^n \to \mathbb{F}_2^m$, which correspond to matrices $\mathbb{F}_2^{m\times n}$, is a 2-universal hash family. That is, we must show that for any distinct x, y,

$$\Pr_{L\sim\mathbb{F}_2^{m\times n}}[L(x) = L(y)] = \Pr_L[L(x+y) = 0] \leq 2^{-m}.$$

This is equivalent to showing that $\Pr_L[Lx = 0] \leq 2^{-m}$ for any nonzero x. This is clearly true (in fact, equality holds), since the rows of L are exactly m independent uniform parity checks on a nonzero x. ◀

Next, we define the notion of *reductions* for extractors.

▶ **Definition 3.7** (Convex combination). *We say* $\mathbf{X}$ *is a convex combination of distributions* $\{\mathbf{Y}_i\}$ *if there exist probabilities* $\{p_i\}$ *summing up to* 1 *such that* $\mathbf{X} = \sum_i p_i \mathbf{Y}_i$.

▶ **Fact 3.8.** *Let* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}$ *be an* ε*-extractor for class* $\mathcal{X}$*. Let* $\mathbf{X} \sim \{0,1\}^n$ *be a distribution that can be written as convex combination of distributions in* $\mathcal{X}$*. Then,* Ext *is also an* ε*-extractor for* $\mathbf{X}$*.*

Finally, the following proposition shows that for a fixed source, a random function is a good extractor.

▶ **Proposition 3.9** (Implicit in [29, Theorem 2.5.1])**.** *For every* $n, m \in \mathbb{N}$*, every* $k \in [0, n]$*, every* $\varepsilon > 0$*, and every* $\mathbf{X} \sim \{0,1\}^n$ *with* $H_\infty(\mathbf{X}) = k$*, if we choose a random function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *with* $m \leq k - 2\log(1/\varepsilon) - O(1)$*, then* $\mathsf{Ext}(\mathbf{X}) \approx_\varepsilon \mathbf{U}_m$ *with probability* $1 - 2^{-\Omega(K\varepsilon^2)}$ *where* $K = 2^k$*.*

4 Constructing extractors

4.1 Input reduction

We show that it suffices to construct extractors for polynomial sources with input length linear in k:

▶ **Theorem 4.1** (Input reduction)**.** *Let* $\mathbf{X} \sim \mathbb{F}_2^n$ *be a degree* d *polynomial source with min-entropy at least* k*. Then* $\mathbf{X}$ *is* 2^{-k}*-close to a convex combination of degree* d *polynomial sources with min-entropy at least* $k - 2$ *and input length at most* $8k$*.*

We begin with the following useful lemma:

▶ **Lemma 4.2** (Existence of affine white-box PRGs)**.** *For any function* $f : \mathbb{F}_2^m \to \mathbb{F}_2^n$ *and* $\varepsilon > 0$*, there exists a linear* $L : \mathbb{F}_2^m \to \mathbb{F}_2^{m-\ell}$ *with* $\ell = n + 3\log(1/\varepsilon)$ *such that*

$$|f(\mathbf{U}_m) \circ L(\mathbf{U}_m) - f(\mathbf{U}_m) \circ \mathbf{U}_{m-\ell}| \leq 2\varepsilon,$$

where $\mathbf{U}_m$ *and* $\mathbf{U}_{m-\ell}$ *are independent.*

Proof. We show a random choice of L works. Indeed, we compute:

$$\begin{aligned} |f(\mathbf{U}_m) \circ L(\mathbf{U}_m) - f(\mathbf{U}_m) \circ \mathbf{U}_{m-\ell}| &= \mathbb{E}_{z \sim f(\mathbf{U}_m)} \left[|(L(\mathbf{U}_m) \mid f(\mathbf{U}_m) = z) - \mathbf{U}_{m-\ell}| \right] \\ &\leq \varepsilon + |L(\mathbf{X}) - \mathbf{U}_{m-\ell}| \\ &\leq 2\varepsilon. \end{aligned}$$

Above $\mathbf{X}$ has min-entropy at least $k = m - n - \log(1/\varepsilon) = m - \ell + 2\log(1/\varepsilon)$ by the min-entropy chain rule (Lemma 3.2), and the last inequality follows via the leftover hash lemma (Corollary 3.6). ◀

We will utilize the following helpful claim:

▷ Claim 4.3 (Entropy smoothing). Let $\mathbf{X} \sim \{0,1\}^n$ be a random variable with min-entropy at least $k > 0$. Then there exists a function $S : \{0,1\}^n \to \{0,1\}^{k+1}$ such that $S(\mathbf{X})$ has min-entropy at least $k - 1$.

Proof. Consider the two least probable elements $x_1, x_2 \in \mathrm{support}(\mathbf{X})$ that occur with probabilities $0 < p_1 \leq p_2$ respectively. We take cases on their values:

Case 1. $p_1 \geq 2^{-k-1}$. This implies $|\mathrm{support}(\mathbf{X})| \leq 2^{k+1}$ and we are done.

Case 2. $p_1 < 2^{-k-1} \leq p_2$. We now merge x_1, x_2 into a single element (that gets hit with probability $< 2^{-k-1} + 2^{-k} \leq 2^{-k+1}$), and end up with a support which again has size at most 2^{k+1}.

Case 3. $p_1 \le p_2 < 2^{-k-1}$. We again merge x_1, x_2 into a single support element that gets hit with probability at most 2^{-k}. Now even if the support size is still too big, we have decreased it by 1, while maintaining the invariant that $\mathbf{X}$ has min-entropy at least k. We now recurse on this same argument until we hit one of the first two good cases, which are eventually guaranteed to happen. ◁

We will also use the data processing inequality for min-entropy:

▷ Claim 4.4 ([37, Lemma 6.8]). For any random variable $\mathbf{X}$ and function f, it holds that $H_\infty(f(\mathbf{X})) \le H_\infty(\mathbf{X})$.

Equipped with these, we prove our main lemma:

▶ **Lemma 4.5** (Existence of affine white-box PEGs). *For any function $f : \mathbb{F}_2^m \to \mathbb{F}_2^n$ such that $f(\mathbf{U}_m)$ has min-entropy at least k, there exists a linear $L : \mathbb{F}_2^m \to \mathbb{F}_2^{m-8k}$ such that*

$$\Pr_{b \sim L(\mathbf{U}_m)}\Big[H_\infty(f(\mathbf{U}_m) \mid L(\mathbf{U}_m) = b) \ge k - 2\Big] \ge 1 - 2^{-k}.$$

Proof. Let $S : \mathbb{F}_2^n \to \mathbb{F}_2^{k+1}$ be a function, guaranteed to exist by Claim 4.3, such that $S(f(\mathbf{U}_m))$ has min-entropy at least $k-1$. By Claim 4.4, it suffices to show that $S(f(\mathbf{U}_m)) \sim \{0,1\}^{k+1}$ has high enough min-entropy with high probability over fixing $L(\mathbf{U}_m)$. Let $\mathbf{Y} = \mathbf{U}_m$. By Lemma 4.2, there exists $L : \mathbb{F}_2^m \to \mathbb{F}_2^{m-\ell}$ with $\ell = k + 2 + 3\log(1/\varepsilon)$ such that:

$$|S(f(\mathbf{Y})) \circ L(\mathbf{Y}) - S(f(\mathbf{Y})) \circ \mathbf{U}_{m-\ell}| \le \varepsilon.$$

This implies

$$\mathbb{E}_{b \sim f(\mathbf{U}_m)}\left[|(S(f(\mathbf{Y})) \mid f(\mathbf{Y}) = b) - S(f(\mathbf{U}_m))|\right] \le \varepsilon$$

By Markov's inequality, we infer that

$$\Pr_{b \sim f(\mathbf{U}_m)}\left[|(S(f(\mathbf{Y})) \mid f(\mathbf{Y}) = b) - S(f(\mathbf{U}_m))| \ge \sqrt{\varepsilon}\right] \le \sqrt{\varepsilon}$$

Setting $\varepsilon = 2^{-2k}$, we infer that the for every "good" fixing b, and $z \in \{0,1\}^{k+1}$, it holds that the probability $S(f(\mathbf{Y} \mid L(\mathbf{Y}) = b))$ outputs z is at most $2^{-k+1} + \sqrt{\varepsilon} \le 2^{-k+2}$. Hence, $S(f(\mathbf{Y} \mid L(\mathbf{Y}) = b))$ has min-entropy at least $k-2$. The result follows, since $\ell = k + 2 + 3\log(1/\varepsilon) = k + 2 + 6k \le 8k$. ◀

Using this main lemma, our theorem easily follows:

Proof of Theorem 4.1. We apply Lemma 4.5 and use the fact that polynomial sources are closed under affine restrictions. ◀

4.2 Existential results

We first show that with high probability, a random function is a good extractor. We will then improve upon it to show that for large enough t, a function sampled using t-wise distribution is a good enough extractor.

▶ **Lemma 4.6.** *Let n, d, k, ε be such that $d < O(n/\log\log n), k \ge \Omega(\log n + d\log\log n), 2^{-\Omega(k)} \le \varepsilon \le 1/2, m = k - 2\log(1/\varepsilon) - O(1)$. For any polynomial source $\mathbf{X} \sim \mathbb{F}_2^n$ of degree at most d with $H_\infty(\mathbf{X}) \ge k$, a random function $r : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is a ε-extractor with high probability.*

Proof. By Theorem 4.1 and Fact 3.8, it suffices to extract from degree d polynomial sources $\mathbf{X}' \sim \mathbb{F}_2^n$ with $O(k)$ inputs and $H_\infty(\mathbf{X}) \geq k-2$. By Fact 3.1, we infer that an extractor with error ε for $\mathbf{X}'$ is also an extractor for $\mathbf{X}$ with error $\varepsilon + 2^{-k}$.

By Proposition 3.9, for a fixed source $\mathbf{Y}$ with $H_\infty(\mathbf{Y}) = k$, a random function $r : \mathbb{F}_2^n \to \mathbb{F}_2^m$ satisfies $r(\mathbf{Y}) \approx_\varepsilon \mathbf{U}_m$ with probability $1 - 2^{-\Omega(2^k)\varepsilon^2}$ where $m = k - 2\log(1/\varepsilon) - O(1)$. We now do a union bound over all the $2^{\binom{\ell}{\leq d}\cdot n}$ degree d sources with ℓ inputs and n outputs. As $\varepsilon \geq 2^{-\Omega(k)}, k \geq \log n + \Omega(d \log\log n), \ell = O(k)$, the union bound indeed succeeds and we infer the claim. ◀

We will use k-wise independent hash functions to help construct extractors for polynomial sources. We will show a random function from a family of such functions will be an extractor. Lets first define them:

▶ **Definition 4.7** ([37, Definition 3.3.1]). *For $n, m, t \in \mathbb{N}$ such that $t \leq 2^n$, a family of functions $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^m\}$ is t-wise independent, if for all distinct $x_1, \ldots, x_t \in \{0,1\}^n$, the random variables $h(x_1), \ldots, h(x_t)$ are independently and uniformly distributed in $\{0,1\}^m$ when h is a randomly chosen function from $\mathcal{H}$.*

We rely on the following property of t-wise independent hash functions in our construction:

▶ **Lemma 4.8** (Implicit in [36, Proposition A.1]). *Let $\mathcal{X}$ be arbitrary class of min-entropy at least k distributions over n bits. Let $\mathcal{H}$ be class of t-wise independent hash functions from n bits to m bits where $t = 2\log(k + |\mathcal{X}|), m = k - 2\log(1/\varepsilon) - \log(t) - 2$. Then, there exists $h \in \mathcal{H}$ such that h is a (k, ε) extractor against all sources in class $\mathcal{X}$.*

Using this, we extend our existential result for t-wise independent hash functions.

▶ **Corollary 4.9.** *Let n, d, k, t, ε be such that $t = 2\log\left(k + 2^{\binom{O(k)}{d}\cdot n}\right), m = k - \log(t) - 2\log(1/\varepsilon) - O(1)$. For any polynomial source $\mathbf{X} \sim \mathbb{F}_2^n$ of degree at most d with $H_\infty(\mathbf{X}) \geq k$, a random function $\mathsf{Ext} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ from a family of t-wise independent functions is an $\varepsilon + 2^{-k}$-extractor with high probability.*

Proof. As earlier, by Theorem 4.1 and Fact 3.8, it suffices to extract from degree d polynomial sources $\mathbf{X}' \sim \mathbb{F}_2^n$ with $O(k)$ inputs and $H_\infty(\mathbf{X}) \geq k-2$. By Fact 3.1, we infer that an extractor with error ε for $\mathbf{X}'$ is also an extractor for $\mathbf{X}$ with error $\varepsilon + 2^{-k}$. Using naive bounds on the number of such polynomial sources, there are at most $2^{\binom{O(k)}{\leq d}\cdot n}$ such sources. We apply Lemma 4.8 for our choice of parameters and infer the claim. ◀

4.3 Explicit construction

We use the input reduction trick and the existential results to construct non-trivial extractors for polynomial sources and prove the formal version of our main result, i.e., Theorem 6.

▶ **Theorem 4.10.** *Let d, n, k be such that $d \leq \Theta(\log\log n / \log\log\log n)$ and $k \geq n - O\left(\frac{\sqrt{\log n}}{(\log\log n/d)^{d/2}}\right)$. Let $\mathcal{X}$ be class of degree d polynomial sources that output n bits and have min-entropy at least k. Then we can construct an extractor for $\mathcal{X}$ in time $\mathrm{poly}(n)$ that extracts $\Omega(\log\log n)$ bits and has error $2^{-\Omega(\log\log n)}$.*

Towards proving the theorem, we first need the following simple observation that the entropy gap of a source cannot get worse by projecting onto a few bits:

▷ Claim 4.11. Let $\mathbf{X} \sim \mathbb{F}_2^n$ be an arbitrary source such that $H_\infty(\mathbf{X}) = k$. Let $\mathbf{X}_0$ be projection of $\mathbf{X}$ onto arbitrary n_0 bits. Then, $H_\infty(\mathbf{X}_0) \geq n_0 - (n-k)$.

Proof. Let $x_0 \in \mathbb{F}_2^{n_0}$ be arbitrary. Let $\Pr(\mathbf{X}_0 = x_0) = p_0$. Then by an averaging argument, there exists $x \in \mathbb{F}_2^n$ such that the projection of x onto coordinates corresponding to $\mathbf{X}_0$ equals x_0 and $\Pr(\mathbf{X} = x) = p \geq p_0 \cdot 2^{-(n-n_0)}$. Hence, if $p_0 > 2^{-(n_0-(n-k))}$, then $p > 2^{-k}$, a contradiction. ◁

We use the following lemma to efficiently construct t-wise independent hash functions:

▶ **Lemma 4.12** (Follows from [37, Corollary 3.3.4]). *For every $n, m, t \in \mathbb{N}$, there exists a family of t-wise independent functions $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^m\}$ such that we can enumerate the family in $2^{t \cdot \max(n,m)} \cdot \mathrm{poly}(n,m,t)$ time and evaluate each function in $\mathrm{poly}(n,m,t)$ time.*

Here is the construction algorithm that we will utilise to construct extractors in Lemma 4.13.

Algorithm 1 Extractor from t-wise independent family.

input : degree d, input source length ℓ, output source length n_0, min-entropy $k_0 = n_0 - g$, extractor output length r, target error ε, the parameter t for t-wise independence

output : An extractor f from n_0 bits to r bits with error ε for degree d polynomial sources from ℓ bits to n_0 bits if it exists from some t-wise independent family

Let $\mathcal{F}$ be some fixed family of t-wise independent functions from n_0 bits to r bits.
for *every function $f \in \mathcal{F}$* **do**
  flag $\leftarrow$ True.
  for *every degree d polynomial map $\mathcal{P}$ from ℓ bits to n_0 bits* **do**
    Brute force over all 2^ℓ assignments to compute min-entropy of $\mathcal{P}(\mathbf{U}_\ell)$ and let it be k_p.
    if $k_p \geq k_0$ **then**
      Brute force over all 2^ℓ assignments to compute $\varepsilon_{f,\mathcal{P}} = |\mathbf{U}_m - f(\mathcal{P}(U_\ell))|$.
      if $\varepsilon_{f,\mathcal{P}} > \varepsilon$ **then**
        flag $\leftarrow$ False.
      end
    end
  end
  if *flag = True* **then**
    return f
  end
end
return *Fail*

Using these, we prove our algorithm in Algorithm 1 indeed works:

▶ **Lemma 4.13.** *Let d, g, n, k, r be such that $0 \leq g \leq n, k \geq n - g, d \leq O\left(\frac{g}{\log g}\right), \Omega(\log g + d \log\log g) \leq r \leq O(g)$. Let $\mathcal{X}$ be class of degree d polynomial sources that output n bits and have min-entropy at least k. Then we can construct an extractor for $\mathcal{X}$ in time $2^{O(\binom{\Theta(r)}{\leq d} \cdot g^2)}$ that extracts r bits and has error $2^{-\Omega(r)}$.*

Proof. Let $\mathbf{X} \in \mathcal{X}$ be arbitrary. Consider the first $n_0 = 1.01g$ bits of $\mathbf{X}$ and let this source be $\mathbf{X}_0$. Then, by Claim 4.11, it holds that $H_\infty(\mathbf{X}_0) \geq n_0 - g \geq \Omega(n_0)$. We use Theorem 4.1 with min-entropy $k_0 = \Omega(r)$ to infer that it suffices to construct extractors polynomial sources with input length $\ell = \Theta(k_0)$. By Corollary 4.9, there exists a function $f : \mathbb{F}_2^{n_0} \to \mathbb{F}_2^r$ such that for all polynomial sources $\mathbf{Y}$, $|f(\mathbf{Y}) - \mathbf{U}_r| \leq 2^{-\Theta(k_0)}$. Moreover, such f will be one of the functions in family of t-wise independent functions where $t = 2\log(\Theta(k_0) + |\mathcal{X}|)$. By setting these input parameters to Algorithm 1, we will indeed find such f.

Let's analyze the runtime of Algorithm 1. The number of degree d sources with input length ℓ and output length n_0 is $2^{\binom{\ell}{\leq d} \cdot n_0}$. The time to enumerate the t-wise independent family is $2^{tn_0}\operatorname{poly}(t, n_0) \leq 2^{2\binom{\ell}{\leq d} \cdot n_0^2}\operatorname{poly}(\ell, d, n_0)$ (Lemma 4.12). Computing entropy and checking if the function is an extractor takes $O(2^{O(\ell+n_0)} \cdot \operatorname{poly}(n_0))$. As $\ell = \Theta(r)$ and $d \leq O(n_0/\log n_0)$, the overall runtime of this algorithm is $2^{O\left(\binom{\Theta(r)}{\leq d} \cdot n_0^2\right)}$. As $n_0 = 1.01g$, the runtime is as desired. ◀

We specialize above lemma to obtain Theorem 4.10.

Proof of Theorem 4.10. Set $g = \Theta\left(\frac{\sqrt{\log n}}{(\log\log n/d)^{d/2}}\right)$ and $r = \Theta(\log g + d\log\log g)$ in Lemma 4.13. ◀

5 Impossibility results

In this section, we show various impossibility results for polynomial NOBF sources and hence, these results apply to both polynomial sources and variety sources. We first show a sampling result that demonstrates power of the quadratic NOBF sources: they can sample optimal sized Sidon sets. We then show affine dispersers cannot be used to disperse from degree d polynomial NOBF sources below certain min-entropy (this is tight). We finally will prove Theorem 9, that sumset extractors cannot be used to even disperse against quadratic NOBF sources below certain min-entropy.

5.1 A warm-up via Sidon sets

To start things off, let us recall the definition of Sidon sets.

▶ **Definition 5.1** (Sidon sets). *We say $S \subset \mathbb{F}_2^n$ is a Sidon set if for all $a, b, c, d \in S$ such that $a + b = c + d$, it holds that $\{a, b\} = \{c, d\}$.*

We show that quadratic NOBF sources can uniformly sample the largest possible Sidon sets over $\mathbb{F}_2^n$, and thus we cannot use sumset extractors below min-entropy $n/2$ to extract from polynomial NOBF sources. Later, we obtain a much stronger version of the latter claim.

We consider the correspondence between $\mathbb{F}_{2^t}$ and $\mathbb{F}_2^t$:

▶ **Definition 5.2.** *For a finite field $\mathbb{F}_{2^t}$, we define the function $\phi : \mathbb{F}_{2^t} \to \mathbb{F}_2^t$ that sends the field element to its vector representation.*

We observe that ϕ is additive:

▶ **Fact 5.3.** *For all $x, y \in \mathbb{F}_{2^t}$, it holds that $\phi(x + y) = \phi(x) + \phi(y)$. Moreover, ϕ is a bijection.*

We will use the following nice lemma involving ϕ:

▶ **Lemma 5.4** ([23, Lemma 2.3.1]). *Let $p : \mathbb{F}_{2^t} \to \mathbb{F}_{2^t}$ be a degree d polynomial and let hamming weight of d when expressed in binary be w. Then, there exists a degree w multilinear polynomial $q : \mathbb{F}_2^t \to \mathbb{F}_2^t$ such that for all $x \in F_{2^t}$, it holds that $\phi(p(x)) = q(\phi(x))$.*

Using these, we show there exists a quadratic NOBF source that uniformly sample largest possible Sidon set over $\mathbb{F}_2^n$:

▷ Claim 5.5. There exists a degree 2 polynomial NOBF source $\mathbf{Y}$ with $H_\infty(\mathbf{Y}) = n/2$ such that $\mathbf{Y}$ uniformly samples a Sidon set.

Proof. Consider the set $S = \{(x, x^3) : x \in \mathbb{F}_{2^{n/2}}\}$. It's well known that this set is a Sidon set [32]. Let $\phi : \mathbb{F}_{2^{n/2}} \to \mathbb{F}_2^{n/2}$ be the function that sends the field element to their vector representation. Using Lemma 5.4, we infer that there exists a degree 2 polynomial map $q : \mathbb{F}_2^{n/2} \to \mathbb{F}_2^{n/2}$ such that for all $x \in \mathbb{F}_2^{n/2}$, $\phi(x^3) = q(\phi(x))$. Applying Fact 5.3, we infer that $T = \{(y, q(y)) : y \in \mathbb{F}_2^{n/2}\}$ is also a Sidon set. We define $\mathbf{Y} : \mathbb{F}_2^{n/2} \to \mathbb{F}_2^n$ to be the degree 2 polynomial NOBF source that is uniform over the set T. Then, $\mathbf{Y}$ uniformly samples a Sidon set and $H_\infty(\mathbf{Y}) = n/2$ as desired. ◁

We show the following towards our impossibility result:

▷ Claim 5.6. Let $S \subset \mathbb{F}_2^n$ be a Sidon set. For all $A, B \subset \mathbb{F}_2^n, |A| \geq 2, |B| \geq 3 : A + B \not\subset S$.

Proof. Say such A, B existed. Pick $a_1, a_2 \in A$ and $b_1, b_2 \in B$ such that $a_1 + a_2 \neq b_1 + b_2$. Let $C = \{a_1+b_1, a_1+b_2, a_2+b_1, a_2+b_2\}$. Then, $|C| = 4$ and $C \subset S$. However, $(a_1+b_1)+(a_1+b_2) = (a_2 + b_1) + (a_2 + b_2)$ which contradicts the fact that S is a Sidon set. ◁

From these, we infer an impossibility result as a corollary:

▶ **Corollary 5.7.** *There exists a degree 2 polynomial NOBF source $\mathbf{Y}$ with $H_\infty(\mathbf{Y}) = n/2$ such that for all $A, B \subset \mathbb{F}_2^n, |A| \geq 2, |B| \geq 3 : A + B \not\subset$ support$(\mathbf{Y})$.*

Looking ahead, we can apply Lemma 5.17 and Lemma 5.18 to infer that we cannot use a (black box) sumset extractor to extract from polynomial NOBF sources of min-entropy $n/2$.

5.2 Affine dispersers cannot disperse from Polynomial NOBF sources

We show that we cannot use affine dispersers to disperse from degree d polynomial NOBF sources below min-entropy $n - n/(\log n)^{d-1}$. By [11, Theorem 7], we know that that an affine disperser for dimension $\Omega(\log n)$ is a disperser for degree d variety sources with min entropy $n - \frac{n}{(\log n)^{d-1}}$. As polynomial NOBF sources are also variety sources, and affine dispersers below min-entropy $\log n$ cannot exist, this result is tight.

▶ **Theorem 5.8.** *Let $c_1 > 0$ be an arbitrary constant. Then, there exists another constant $c_2 > 0$ such that the following holds: For $2 \leq d \leq \frac{\log n}{2 \cdot \log\log n}$, There exists a degree d polynomial NOBF source $\mathbf{X}$ with $H_\infty(\mathbf{X}) = n - c_2 \frac{n}{(\log n)^{d-1}}$ such that support$(\mathbf{X})$ does not contain any affine subspace of dimension $c_1 \log n$.*

As affine dispersers with min-entropy requirement $\log n$ can't exist, it indeed follows that we can't use affine dispersers to disperse from polynomial NOBF sources with the stated min-entropy bound.

We first show that a random degree d polynomial map will not become a linear map over any small affine subspace.

▷ Claim 5.9. There exists a universal constant c such that the following holds. Let d, n, t be such that $2 \leq d < n/2$ and $t < n$. Then, there exist degree d polynomials $p_1, \ldots, p_t$ such that on every affine subspace U of dimension $k \geq cd \cdot (n/t)^{1/(d-1)}$, there exists at least one i such that p_i has degree ≥ 2.

Proof. Let $p_1, \ldots, p_t : \mathbb{F}_2^n \to \mathbb{F}_2$ be random polynomials of degree d. Let U be arbitrary but fixed affine subspace of dimension k. Then, $p_1|_U, \ldots, p_t|_U$ are also uniform polynomials over k variables of degree d. Hence, it must be that:

$$\Pr_{p_1,\ldots,p_t}\left[\bigwedge_{1\leq i\leq t} \deg(f|_U) \leq 1\right] \leq 2^{-(\binom{k}{\leq d}-\binom{k}{\leq 1})t}$$

We union bound over all $\leq 2^n \binom{2^n}{k}$ affine subspaces of dimension k and see that the probability that there exists some affine subspace of dimension k over which all these polynomials have degree at most 1 is at most

$$2^{-(\binom{k}{\leq d}-\binom{k}{\leq 1})t} \cdot 2^n \cdot \binom{2^n}{k}$$

We set c to a large constant so that the above probability less than 1. ◁

We now show a random polynomial NOBF source does not contain any small affine subspace.

▷ Claim 5.10. There exists a universal constant c such that the following holds: Let d, k be such that $2 \leq d < k/2$. For any $0 < t < k$, there exists a degree d polynomial NOBF source $\mathbf{X}$ over $k + t$ bits with $H_\infty(\mathbf{X}) = k$ such that support$(\mathbf{X})$ does not contain any affine subspace of dimension $cd \cdot (k/t)^{1/(d-1)}$.

Proof. Let $s = cd \cdot (k/t)^{1/(d-1)}$. Let $(p_1, \ldots, p_t) : \mathbb{F}_2^k \to \mathbb{F}_2^t$ be the t polynomials from Claim 5.9. Let $\mathbf{X}$ be the polynomial NOBF source over $k + t$ bits where first k bits are $x_1, \ldots, x_k$ and last t bits are $p_1(x_1, \ldots, x_k), \ldots, p_t(x_1, \ldots, x_k)$.

Assume that there exists an affine subspace $U \subset$ support$(\mathbf{X})$ such that $\dim(U) = s$. Observe that once the first k bits of $\mathbf{X}$ are fixed, the last t bits are also fixed. As $U \subset$ support$(\mathbf{X})$, U must also have this property. Let $P \subset \mathbb{F}_2^k$ be the projection of U over the first k bits. Then, $\dim(P) = \dim(U) = s$. Moreover, as U is an affine subspace, the last t bits of U are linear functions of the first k bits. However, this implies that for each $1 \leq i \leq t$, $\deg(p_i|_P) \leq 1$, which is a contradiction. ◁

Proof of Theorem 5.8. Let C be a large enough constant. We apply Claim 5.10 and set $t = k\left(\frac{Cd}{\log n}\right)^{d-1}$ to infer the claim. ◀

5.3 Sumset dispersers cannot disperse from Polynomial NOBF sources

We show that we cannot use sumset dispersers to disperse from quadratic NOBF sources below min-entropy $n - n/\log n$.

▶ **Theorem 5.11.** *Let $c_1 > 0$ be an arbitrary constant. Then, there exists another constant $c_2 > 0$ such that the following holds: There exists a degree 2 polynomial NOBF source $\mathbf{X}$ with $H_\infty(\mathbf{X}) = n - c_2 \frac{n}{\log n}$ such that support$(\mathbf{X})$ does not contain any sumset $A + B$ where $|A| \geq n^{c_1}, |B| \geq n^{c_1}$.*

In fact, we will prove the following, more fine-grained, version of Theorem 5.11.

▷ Claim 5.12. There exists a universal constant c such that the following holds. For any $0 < t < k$, there exists a degree 2 polynomial NOBF source $\mathbf{X}$ over $n = k + t$ bits with $H_\infty(\mathbf{X}) = k$ such that support$(\mathbf{X})$ does not contain any sumset $A + B$ where $|A| \geq 2^{cn/t}, |B| \geq 2^{cn/t}$.

Given the above claim, it is easy to prove Theorem 5.11.

Proof of Theorem 5.11. The theorem follows by setting $t = O(n/\log n)$ in Claim 5.12. ◀

The rest of this section is devoted to proving Claim 5.12. To do so, we prove two claims.

▷ Claim 5.13. There exists a universal constant c such that the following holds. Let $t, n \in \mathbb{N}$ be such that $t < n$. There exist degree 2 polynomials $p_1, \ldots, p_t : \mathbb{F}_2^n \to \mathbb{F}_2$ such that for every pair of affine subspaces U, V of dimensions $r \geq c(n/t)$ each, and for all $y \in \mathbb{F}_2^n$, there exists at least at least one i and at least one $u \in U, v \in V$ such that $p_i(u + v) \neq p_i(u) + p_i(v) + y_i$.

▷ Claim 5.14. Let $P = (p_1, \ldots, p_t) : \mathbb{F}_2^n \to \mathbb{F}_2^t$ be a degree 2 polynomial map. Let $y \in \mathbb{F}_2^t$ be arbitrary. Let $A, B \subset \mathbb{F}_2^n$ be such that for all $a \in A, b \in B$, it holds that $P(a) + P(b) = P(a + b) + y$. Then, there exist affine subspaces $U, V \subset \mathbb{F}_2^n$ such that for all $u \in U, v \in V$, it holds that $P(u) + P(v) = P(u + v) + y$ and $|U| \geq |A|, |V| \geq |B|$.

Using these two claims, it is not too difficult to prove Claim 5.12.

Proof of Claim 5.12. Let $\mathbf{X}$ be the polynomial NOBF source where first k bits are uniform variables and last t bits are output of polynomial map P from Claim 5.13 (hence set c to the universal constant from there). We now proceed by contradiction and assume there exist $A, B \subset \mathbb{F}_2^n$ such that $|A| \geq 2^{cn/t}, |B| \geq 2^{cn/t}$, and $A + B \subset$ support$(\mathbf{X})$. Let $a_0 \in A, b_0 \in B$ be arbitrary. Let $A' = a_0 + A, B' = b_0 + B, \mathbf{X}' = \mathbf{X} + (a_0 + b_0)$. Then, $A' + B' \subset$ support$(\mathbf{X}')$. Observe that $0^n \in A'$ and $0^n \in B'$. So, $A' \subset$ support$(\mathbf{X}')$, and $B' \subset$ support$(\mathbf{X}')$. Moreover, $\mathbf{X}'$ is a degree 2 polynomial NOBF source with $H_\infty(\mathbf{X}') = H_\infty(\mathbf{X})$.

Let the last $n - k$ bits of $\mathbf{X}'$ be the output of the degree 2 polynomial map P'. Let $A'_0, B'_0 \subset \mathbb{F}_2^k$ be the projections of A', B' respectively onto the first k bits. As $A' \subset$ support$(\mathbf{X}'), B' \subset$ support$(\mathbf{X}')$, and the last $n - k$ bits are deterministic functions of the first k bits, it must be that $|A'_0| = |A'|$ and $|B'_0| = |B'|$. Similarly, as $A' + B' \subset$ support$(\mathbf{X}')$, it must be that $P'(A'_0) + P'(B'_0) = P'(A'_0 + B'_0)$. By Claim 5.14, there exist affine subspaces $U', V' \subset \mathbb{F}_2^k$ such that for all $u' \in U'^{\mathrm{Pr}}, v' \in V'^{\mathrm{Pr}}$, it holds that $P'(u') + P'(v') = P'(u' + v')$ where $|U'| \geq |A'_0| = |A'|, |V'| \geq |B'_0| = |B'|$.

Observe that $P'(x) = P(g + x) + h$ where $g \in \mathbb{F}_2^k, h \in \mathbb{F}_2^t$ are some fixed strings. Then, $P(g+U')+P(g+V') = P(g+U'+V')+h$. Let $U, V \subset \mathbb{F}_2^k$ be such that $U = g+U', V = g+V'$. Then, U, V are affine subspaces, $P(U) + P(V) = P(U + V) + h$, and $|U| = |U'| \geq |A'| = |A|, |V| = |V'| \geq |B'| = |B|$. However, this is a contradiction to the choice of P. ◁

We now prove the helpful claim that random quadratic maps P have the property that for every large affine subspaces U, V, there exist $u \in U, v \in V$ such that $P(u) + P(v) \neq P(u + v)$.

Proof of Claim 5.13. Fix $y \in \mathbb{F}_2^t$. At the end, we will union bound over these 2^t distinct y. Let $P = (p_1, \ldots, p_t) : \mathbb{F}_2^n \to \mathbb{F}_2^t$ be a random degree 2 polynomial map. Without loss of generality assume that $y = 0^t$. Indeed, we can define $P'(x) = P(x) + y$ so that for all $a \in A, b \in B$, it holds that $P'(a) + P'(b) = P'(a + b)$. Moreover, if there exist U, V such that for all $u \in U, v \in V$: $P'(u) + P'(v) = P'(u + v)$, then we will recover that for all $u \in U, v \in V$: $P(u) + P(v) = P(u + v) + y$ as desired. Moreover, P' as defined will be distributed as a uniformly random degree 2 polynomial map.

Call a degree 2 polynomial map P "bad" if there exist affine subspaces C, D of dimensions r each such that for all $c \in C, d \in D$ it holds that $P(c) + P(d) = P(c + d)$. Call such C, D the affine subspaces that "witness" the badness of P. We will go over each pair of affine subspaces C, D and show that the fraction of bad maps witnessed by the pair (C, D) are very small.

Let U, V be arbitrary dimension r subspaces. Let $u_0, v_0 \in \mathbb{F}_2^n$ be arbitrary. Then, we fix $u_0 + U, v_0 + V$ to be arbitrary but fixed affine subspaces of dimension r each. We consider two cases:

Case 1. $\dim(U \cap V) \geq r/2$.

Let $W = (U \cap V)$. Let P be a bad map witnessed by $u_0 + U, v_0 + V$, i.e., for all $u \in (u_0 + U), v \in (v_0 + V)$, it holds that $P(u) + P(v) = P(u + v)$. We claim that $P|_{(u_0+W)}$ is a degree 1 polynomial map. Indeed, above condition guarantees that $\forall w_1, w_2 \in W$: $P(u_0 + w_1) + P(v_0 + w_2) = P(u_0 + v_0 + w_1 + w_2)$. This also implies that $\forall w \in W$: $P(u_0 + w) + P(v_0 + w) = P(u_0 + v_0)$. Repeatedly applying these, we infer that:

$$\begin{aligned} P(u_0 + (w_1 + w_2)) &= P(u_0 + v_0) + P(v_0 + (w_1 + w_2)) \\ &= P(u_0 + v_0) + (P(u_0 + w_1) + P(u_0 + v_0 + w_2)) \\ &= P(u_0 + w_1) + (P(u_0 + v_0)) + (P(v_0 + (u_0 + w_2))) \\ &= P(u_0 + w_1) + (P(u_0) + P(v_0)) + (P(v_0) + P(u_0 + w_2))) \\ &= P(u_0 + w_1) + P(u_0 + w_2) + P(u_0) \end{aligned}$$

Hence, P restricted to $u_0 + W$ is indeed an affine map. We observe that $p_1|_{u_0+W}, \ldots, p_t|_{u_0+W}$ are distributed as uniform degree at most 2 polynomials over $r/2$ variables. The probability that each of these polynomials has degree at most 1 is at most $2^{-\binom{r/2}{2}t}$.

Case 2. $\dim(U \cap V) < r/2$.

Let $u_0 + S$ be the largest affine subspace inside $u_0 + U$ such that $S \cap (U \cap V) = \{0\}$. Similarly, let $v_0 + T$ be the largest affine subspace inside V such that $T \cap (U \cap V) = \{0\}$. It must be that $\dim(S), \dim(T) \geq r/2$ and $S \cap T = \emptyset$. By considering appropriate subsets of S and T, we without loss of generality assume $\dim(S) = \dim(T) = r/3, (u_0 + S) \cap (T \cup (v_0 + T)) = (v_0 + T) \cap (S \cup (u_0 + S)) = \emptyset$. Let basis vectors of S and T be $(s_1, \ldots, s_{r/3})$, and $(t_1, \ldots, t_{r/3})$ respectively. Without loss of generality, let it be that $u_0 + s_1, \ldots, u_0 + s_{r/3}$ are linearly independent and $v_0 + t_1, \ldots, v_0 + t_{r/3}$ are also linearly independent. Then, by using the various empty intersection conditions above, the vectors $u_0 + s_1, \ldots, u_0 + s_{r/3}, v_0 + t_1, \ldots, v_0 + r_{r/3}$ are also linearly independent. Let $b_1, \ldots, b_{n-r/3}$ be linearly independent vectors so that $u_0 + s_1, \ldots, u_0 + s_{r/3}, v_0 + t_1, \ldots, v_0 + r_{r/3}, b_1, \ldots, b_{n-r/3}$ are all linearly independent. Let's rename these vectors to be $c_1, \ldots, c_n$.

Now, we choose the random quadratic polynomials $p_1, \ldots, p_t$ by randomly sampling monomials of degree at most 2 over these c_i. As the c_i are linearly independent, P will still be a uniformly random quadratic map. Let P be a bad map witnessed by $u_0 + U, v_0 + V$. We claim there does not exist i such that p_i contains the monomial $c_j c_k$ where $c_j \in \{u_0 + s_1, \ldots, u_0 + s_{r/3}\}$ and $c_k \in \{v_0 + t_1, \ldots, v_0 + t_{r/3}\}$. Proceed by contradiction and assume there exists i such that p_i contains the monomial $c_j c_k$ where $c_j \in \{u_0 + s_1, \ldots, u_0 + s_{r/3}\}$ and $c_k \in \{v_0 + t_1, \ldots, v_0 + t_{r/3}\}$. Without loss of generality we assume that the singleton monomials c_j and c_k are not present in p_i and the degree 0 monomial is also absent (if any of them are present, then we can easily change assignments α_1, α_2 and their outcomes below and get the same claim). Consider the following two assignments:

1. Assignment α_1 where $c_j = 1$, and remaining variables are set to 0.
2. Assignment α_2 where $c_k = 1$, and remaining variables set to 0.

Then, $p_i(\alpha_1) = p_i(\alpha_2) = 0$. Moreover, we observe that $p_i(\alpha_1 + \alpha_2) = 1$. However, this means we found $\alpha_1 \in (u_0 + U)$ and $\alpha_2 \in (v_0 + V)$ such that $P(\alpha_1) + P(\alpha_2) \neq P(\alpha_1 + \alpha_2)$, contradicting the fact that $u_0 + U, v_0 + V$ witnessed badness of P. Hence, for this not to happen, all such "cross" monomials must not occur in any p_i. This happens with probability at most $2^{-(r/3)^2 t}$.

We union bound over all pairs of affine subspaces of dimension r and consider whether they fall into the first case or the second case. If they fall into the first case, then we only union bound over $\leq 2^n \cdot \binom{2^n}{r/2}$ affine subspaces of dimension $r/2$ and consider the probability that a bad map P that they witness becomes linear over that affine subspace. If they fall into the second case, then we union bound over all $\leq \left(2^n \cdot \binom{2^n}{r/3}\right)^2$ disjoint pairs of affine subspaces of dimension $r/3$ use consider the probability that any bad map they witness won't have such cross monomials. We finally add both these probabilities to get our final bound. For the first case, the expression will be

$$2^{-\binom{r/2}{2} t} \cdot 2^n \cdot \binom{2^n}{r/2}$$

For the second case, the expression will be:

$$2^{-(r/3)^2 t} \cdot \left(2^n \cdot \binom{2^n}{r/3}\right)^2$$

We can choose c large enough so that the sum of the above probabilities is much smaller than than 2^{-t}. Then, we union bound over all 2^t of the $y \in \mathbb{F}_2^t$ to get the desired claim. ◁

We lastly prove the useful claim that for a quadratic map P, if there exist sets A, B such that for all $a \in A, b \in B$ it holds that $P(a) + P(b) = P(a + b)$, then we can also find affine subspaces U, V with the same property and of larger sizes. We first need the notion of directional derivatives:

▶ **Definition 5.15** (Directional derivative). *For a polynomial $p : \mathbb{F}_2^n \to \mathbb{F}_2$ and $a \in \mathbb{F}_2^n$, we define its directional derivative in direction a, i.e., $D_a(p)(\cdot) : \mathbb{F}_2^n \to \mathbb{F}_2$ as*

$$D_a(p)(x) = p(x) + p(x + a)$$

Clearly $D_a(p)(\cdot)$ is a polynomial. It's well known that $\deg(D_a(p)(\cdot)) \leq \deg(p) - 1$. We also extend the definition of directional derivatives to apply to a polynomial map $P : \mathbb{F}_2^n \to \mathbb{F}_2^t$. For a fixed direction $a \in \mathbb{F}_2^n$, we define $D_a(P)(\cdot) : \mathbb{F}_2^n \to \mathbb{F}_2^m$ as $D_a(P)(x) = (D_a(p_1)(x), \ldots, D_a(p_m)(x))$. Using these, we present our proof:

Proof of Claim 5.14. Without loss of generality assume that $y = 0^t$. Indeed, let $P'(x) = P(x) + y$ so that for all $a \in A, b \in B$: $P'(a) + P'(b) = P'(a + b)$. Moreover, if there exist such affine subspaces U, V so that for all $u \in U, v \in V$: $P'(u) + P'(v) = P'(u + v)$, then we will indeed recover the fact that for all $u \in U, v \in V$: $P(u) + P(v) = P(u + v) + y$ as desired. Let C, D be such that $A \subset C$, $B \subset D$, and for all $c \in C, d \in D$ it holds that $P(c) + P(d) = P(c + d)$. Moreover, let C and D are the largest such sets. To prove the claim, it suffices to show that C and D are affine subspaces.

For $a \in \mathbb{F}_2^n$, define $D_a(P)(x) = (D_a(p_1)(x), \ldots, D_a(p_t)(x)) = P(x) + P(x+a)$, the map of directional derivatives in direction a. Let $S_a = \{z \in \mathbb{F}_2^n : D_a(P)(z) = P(a)\}$. We claim that $y \in S_a \iff P(y) + P(a) = P(a+y)$. Indeed,

$$D_a(P)(y) = P(a) \iff P(y) + P(y+a) = P(a)$$

Let $S_C = \bigcap_{c \in C} S_c$. Then, it must be that $B \subset S_C$. As D is maximal such set and for all $c \in C, s \in S_C$: $P(c) + P(s) = P(c+s)$, it must be that $D = S_C$. By a symmetric argument, $C = S_D$. Observe that for arbitrary $a \in \mathbb{F}_2^n$, S_a is an affine subspace. As intersection of affine subspaces is an affine subspace, $S_C = D$ and $S_D = C$ are affine subspaces, as desired. ◁

5.4 Sumset extractors cannot disperse from Polynomial NOBF sources

We show that we cannot use sumset extractors to disperse from quadratic NOBF sources below min-entropy $n - n/\log\log n$.

▶ **Theorem 5.16.** *Let $0 < \varepsilon < 1, 0 < c_1$ be arbitrary constants. Then, there exists another constant $c_2 > 0$ such that the following holds: There exists a degree 2 polynomial NOBF source $\mathbf{X}$ with $H_\infty(\mathbf{X}) = n - c_2 \frac{n}{\log\log n}$ such that $\mathbf{X}$ is $(1-\varepsilon)$-far from a convex combination of sumset sources of min-entropy $c_1 \log n$.*

Note that if a distribution is $\frac{1}{2}$ distance away from any convex combination of sumset sources then a sumset extractor cannot be used in a blackbox way as a disperser. Also, as no sumset extractor can exist for min-entropy below $\log n$, these results indeed show we can't use sumset extractors in a blackbox way to disperse from degree 2 polynomial NOBF sources.

We will prove a worst case to average case type reduction for sumsets.

▶ **Lemma 5.17.** *Let $0 < \delta < 1$ be a fixed constant. Let $\mathbf{X} \sim \mathbb{F}_2^n$ be such that for all flat sources $\mathbf{A}, \mathbf{B} \sim \mathbb{F}_2^n$ with $H_\infty(\mathbf{A}) = H_\infty(\mathbf{B}) = t$, it holds that $(\mathbf{A} + \mathbf{B}) \not\subset$ support$(\mathbf{X})$. Then, for all flat sources $\mathbf{R}, \mathbf{S} \sim \mathbb{F}_2^n$ such that $H_\infty(\mathbf{R}) = H_\infty(\mathbf{S}) = c \cdot 2^t$, it holds that $\Pr_{r \sim \mathbf{R}, s \sim \mathbf{S}}[r + s \in$ support$(\mathbf{X})] \leq \delta$. Here, $c > 0$ is a constant depending only on δ.*

We will show that if a source is far from all sumset sources, then it is also far from all convex combination of sumset sources:

▶ **Lemma 5.18** (Similar to [1, Theorem 14])**.** *Let $0 \leq \delta \leq 1, 0 \leq k$, and $\mathbf{X} \sim \mathbb{F}_2^n$ be such that for all flat sources $\mathbf{R}, \mathbf{S} \sim \mathbb{F}_2^n$ with $H_\infty(\mathbf{R}), H_\infty(\mathbf{S}) \geq k$, it holds that $\Pr_{r \sim \mathbf{R}, s \sim \mathbf{S}}[r + s \in$ support$(\mathbf{X})] \leq \delta$. Then, for all $\mathbf{Y}$ such that $\mathbf{Y}$ is a convex combination of sumset sources of min-entropy at least k, it holds that $|\mathbf{X} - \mathbf{Y}| \geq 1 - \delta$.*

Using these, and Claim 5.12 from Section 5.3, we show that sumset extractors cannot even *disperse* from degree 2 polynomial NOBF sources:

Proof of Theorem 5.16. Let $\mathbf{X}$ be source guaranteed by Claim 5.12 with $H_\infty(\mathbf{X}) = n - c_2 \frac{n}{\log\log n}$ such that for all $A, B \subset \mathbb{F}_2^n$ with $|A| = |B| = c \log n$, it holds that $(A + B) \not\subset$ support$(\mathbf{X})$. Using Lemma 5.17, we infer that for all flat sources $\mathbf{R}, \mathbf{S} \sim \mathbb{F}_2^n$ with $H_\infty(\mathbf{R}) = H_\infty(\mathbf{S}) = c_1 \log n$, it holds that $\Pr(\mathbf{R}+\mathbf{S}) \in$ support$(\mathbf{X}) \leq \delta$. Let $\mathbf{Y} \sim \mathbb{F}_2^n$ be arbitrary convex combination of sumset sources $\{(\mathbf{R}^{(i)} + \mathbf{S}^{(i)})\}_i$, each with min-entropy $c_1 \log n$. Applying Lemma 5.18, we infer that $|\mathbf{X} - \mathbf{Y}| \geq 1 - \delta$ as desired. ◀

We use a bipartite Ramsey bound to get a worst case to average case type reduction for sumset sources:

▶ **Lemma 5.19** ([40]). *The maximum number of edges in a bipartite graph over* $[n] \times [n]$ *without inducing a complete bipartite* $t \times t$ *subgraph is at most* $(t-1)^{1/t} \cdot n^{2-1/t} + \frac{1}{2} \cdot (t-1) \cdot n$.

We will utilize the following corollary of this statement:

▶ **Corollary 5.20.** *Fix* $0 < \delta \leq 1$. *Let* G *be a bipartite graph over* $[n] \times [n]$ *with at* $\delta \cdot n^2$ *edges. Then,* G *induces a complete bipartite subgraph* H *over* $[\varepsilon \cdot \log n] \times [\varepsilon \cdot \log n]$ *where* $0 < \varepsilon \leq 1$ *is a constant depending only on* δ.

Equipped with this, we prove our main lemma:

Proof of Lemma 5.17. Assume this is not the case and there exist such $\mathbf{R}$ and $\mathbf{S}$. Let $H_\infty(\mathbf{R}) = H_\infty(\mathbf{S}) = k$. Consider a bipartite graph G over $\text{support}(\mathbf{R}) \times \text{support}(\mathbf{S})$ with an edge between $r \in \text{support}(\mathbf{R})$ and $s \in \text{support}(\mathbf{S})$ if $r + s \in \text{support}(\mathbf{X})$. By assumption, G has at least $\delta \cdot 2^{2k}$ edges. Using Corollary 5.20, we infer that G induces a complete bipartite subgraph where where each part has size $\varepsilon \cdot k$ (ε depends only on δ). Equivalently, there exist sets $C \subset \text{support}(\mathbf{R}), D \subset \text{support}(\mathbf{S})$ such that $|C| = |D| = \varepsilon \cdot k$ and $(C + D) \subset \text{support}(\mathbf{X})$. Let $\mathbf{A}$ be the uniform distribution over C and $\mathbf{B}$ be the uniform distribution over D. Then, $H_\infty(\mathbf{A}) = H_\infty(\mathbf{B}) = \log(\varepsilon \cdot k)$. Setting $c = 1/\varepsilon$, we get a contradiction. ◀

Finally, we show that if a distribution is far from every sumset source, then it's far from every convex combination of sumset sources.

Proof of Lemma 5.18. Let $\mathbf{Y} \sim \mathbb{F}_2^n$ be arbitrary convex combination of sumset sources $\{(\mathbf{R}^{(i)} + \mathbf{S}^{(i)})\}_i$, each with min-entropy at least k. Let $T = \text{support}(\mathbf{X})$. Then,

$$\begin{aligned} |\mathbf{X} - \mathbf{Y}| &\geq \Pr[\mathbf{Y} \in \overline{T}] - \Pr[\mathbf{X} \in \overline{T}] = \Pr[\mathbf{Y} \in \overline{T}] \\ &\geq \min_i \Pr[\mathbf{Y}^{(i)} \in \overline{T}] = 1 - \max_i \Pr[\mathbf{Y}^{(i)} \in T] \\ &\geq 1 - \delta. \end{aligned}$$

◀

6 Open problems

The problem of constructing extractors for sources sampled by $\mathbb{F}_2$-polynomials is a natural one, and we view our results as initial progress on this question. We leave open a number of interesting open directions:

1. Construct extractors or dispersers for polynomial sources with better min-entropy dependence than what we constructed here. For instance, some interesting potential candidates to explore are the MAJORITY function, or the generalized inner product function.
2. It will be interesting to make progress on the easier question of extracting from constant degree polynomial NOBF sources below min-entropy $0.999n$. Extracting from constant degree variety sources below min-entropy $0.999n$ is an important open problem and here, we introduced an interesting subclass of variety sources – polynomial NOBF sources – for which we also don't have better extractors.
   An even simpler question is to construct *dispersers* for constant degree polynomial NOBF sources below min-entropy $n/2$. Note that for any NOBF source with $> n/2$ good bits, the MAJORITY function is a disperser.
3. Construct extractors or dispersers for polynomial sources with degree $\text{poly}(\log n)$. Such an extractor will also extract from sources sampled by $\mathsf{AC}^0[\oplus]$ circuits, a model for which no non-trivial extractors are known. Our constructions work for degree up to $O(\log \log n)$, and thus fall short of achieving this.

References

1 Omar Alrabiah, Eshan Chattopadhyay, Jesse Goodman, Xin Li, and João Ribeiro. Low-degree polynomials extract from local sources. In Mikolaj Bojanczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022, Paris, France*, volume 229 of *LIPIcs*, pages 10:1–10:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPICS.ICALP.2022.10.

2 Boaz Barak, Guy Kindler, Ronen Shaltiel, Benny Sudakov, and Avi Wigderson. Simulating independence: New constructions of condensers, ramsey graphs, dispersers, and extractors. *J. ACM*, 57(4):20:1–20:52, 2010. doi:10.1145/1734213.1734214.

3 Eli Ben-Sasson and Ariel Gabizon. Extractors for polynomial sources over fields of constant order and small characteristic. *Theory Comput.*, 9:665–683, 2013. doi:10.4086/toc.2013.v009a021.

4 Eli Ben-Sasson and Swastik Kopparty. Affine dispersers from subspace polynomials. *SIAM J. Comput.*, 41(4):880–914, 2012. doi:10.1137/110826254.

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

6 Jean Bourgain, Zeev Dvir, and Ethan Leeman. Affine extractors over large fields with exponential error. *Comput. Complex.*, 25(4):921–931, 2016. doi:10.1007/s00037-015-0108-5.

7 Eshan Chattopadhyay and Jesse Goodman. Improved extractors for small-space sources. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 610–621. IEEE, 2021. doi:10.1109/FOCS52979.2021.00066.

8 Eshan Chattopadhyay, Jesse Goodman, and Jyun-Jie Liao. Affine extractors for almost logarithmic entropy. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 622–633. IEEE, 2021. doi:10.1109/FOCS52979.2021.00067.

9 Eshan Chattopadhyay and Jyun-Jie Liao. Extractors for sum of two sources. In Stefano Leonardi and Anupam Gupta, editors, *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 1584–1597. ACM, 2022. doi:10.1145/3519935.3519963.

10 Eshan Chattopadhyay and Avishay Tal. Personal communication to li and zuckerman.

11 Gil Cohen and Avishay Tal. Two structural results for low degree polynomials and applications. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2015, August 24-26, 2015, Princeton, NJ, USA*, volume 40 of *LIPIcs*, pages 680–709. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015. doi:10.4230/LIPIcs.APPROX-RANDOM.2015.680.

12 Anindya De and Thomas Watson. Extractors and lower bounds for locally samplable sources. *ACM Trans. Comput. Theory*, 4(1):3:1–3:21, 2012. doi:10.1145/2141938.2141941.

13 Matt DeVos and Ariel Gabizon. Simple affine extractors using dimension expansion. In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity, CCC 2010, Cambridge, Massachusetts, USA, June 9-12, 2010*, pages 50–57. IEEE Computer Society, 2010. doi:10.1109/CCC.2010.14.

14 Zeev Dvir. Extractors for varieties. *Comput. Complex.*, 21(4):515–572, 2012. doi:10.1007/s00037-011-0023-3.

15 Zeev Dvir, Ariel Gabizon, and Avi Wigderson. Extractors and rank extractors for polynomial sources. *Comput. Complex.*, 18(1):1–58, 2009. doi:10.1007/s00037-009-0258-4.

16 Ariel Gabizon and Ran Raz. Deterministic extractors for affine sources over large fields. *Comb.*, 28(4):415–440, 2008. doi:10.1007/s00493-008-2259-3.

17 Alexander Golovnev and Alexander S. Kulikov. Weighted gate elimination: Boolean dispersers for quadratic varieties imply improved circuit lower bounds. In Madhu Sudan, editor, *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science, Cambridge, MA, USA, January 14-16, 2016*, pages 405–411. ACM, 2016. doi:10.1145/2840728.2840755.

18 Alexander Golovnev, Alexander S. Kulikov, and R. Ryan Williams. Circuit depth reductions. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference, ITCS 2021, January 6-8, 2021, Virtual Conference*, volume 185 of *LIPIcs*, pages 24:1–24:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.ITCS.2021.24.

19 Zeyu Guo, Ben Lee Volk, Akhil Jalan, and David Zuckerman. Extractors for images of varieties. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 46–59. ACM, 2023. doi:10.1145/3564246.3585109.

20 Johan Håstad, Russell Impagliazzo, Leonid A Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999. Preliminary version in STOC 1989.

21 Miguel Herrero-Collantes and Juan Carlos Garcia-Escartin. Quantum random number generators. *Reviews of Modern Physics*, 89(1):015004, 2017.

22 Jesse Kamp, Anup Rao, Salil Vadhan, and David Zuckerman. Deterministic extractors for small-space sources. *Journal of Computer and System Sciences*, 77(1):191–220, 2011. Preliminary version in STOC 2006.

23 Swastik Kopparty. *Algebraic methods in randomness and pseudorandomness*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2010. URL: https://hdl.handle.net/1721.1/62425.

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

25 Xin Li. A new approach to affine extractors and dispersers. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity, CCC 2011, San Jose, California, USA, June 8-10, 2011*, pages 137–147. IEEE Computer Society, 2011. doi:10.1109/CCC.2011.27.

26 Xin Li. Improved two-source extractors, and affine extractors for polylogarithmic entropy. In Irit Dinur, editor, *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 168–177. IEEE Computer Society, 2016. doi:10.1109/FOCS.2016.26.

27 Xin Li. Two source extractors for asymptotically optimal entropy, and (many) more. In *64th Annual Symposium on Foundations of Computer Science (FOCS 2023)*. IEEE Computer Society, 2023.

28 Rajeev Motwani and Prabhakar Raghavan. *Randomized algorithms*. Cambridge University Press, 1995.

29 Anup Rao. *Randomness extractors for independent sources and applications*. PhD thesis, The University of Texas at Austin, 2007.

30 Anup Rao. Extractors for low-weight affine sources. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity, CCC 2009, Paris, France, 15-18 July 2009*, pages 95–101. IEEE Computer Society, 2009. doi:10.1109/CCC.2009.36.

31 Razbarov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical Notes of the Academy of Sciences of the USSR*, 41(4):333–338, 1987.

32 Maximus Redman, Lauren Rose, and Raphael Walker. A small maximal sidon set in $\mathbb{Z}_2^n$. *SIAM J. Discret. Math.*, 36(3):1861–1867, 2022. doi:10.1137/21m1454663.

33 Zachary Remscrim. The hilbert function, algebraic extractors, and recursive fourier sampling. In Irit Dinur, editor, *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 197–208. IEEE Computer Society, 2016. doi:10.1109/FOCS.2016.29.

34 Ronen Shaltiel. Dispersers for affine sources with sub-polynomial entropy. In Rafail Ostrovsky, editor, *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 247–256. IEEE Computer Society, 2011. doi:10.1109/FOCS.2011.37.

35 R. Smolensky. On representations by low-degree polynomials. In *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*, pages 130–138, 1993. doi:10.1109/SFCS.1993.366874.

36 Luca Trevisan and Salil P. Vadhan. Extracting randomness from samplable distributions. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 32–42. IEEE Computer Society, 2000. doi:10.1109/SFCS.2000.892063.

37 Salil P. Vadhan. Pseudorandomness. *Found. Trends Theor. Comput. Sci.*, 7(1-3):1–336, 2012. doi:10.1561/0400000010.

38 Emanuele Viola. Extractors for circuit sources. *SIAM J. Comput.*, 43(2):655–672, 2014. doi:10.1137/11085983X.

39 Amir Yehudayoff. Affine extractors over prime fields. *Comb.*, 31(2):245–256, 2011. doi:10.1007/s00493-011-2604-9.

40 S Znám. On a combinatorical problem of k. zarankiewicz. In *Colloquium Mathematicum*, volume 11, pages 81–84. Instytut Matematyczny Polskiej Akademii Nauk, 1963.