Extractors and Rank Extractors for Polynomial Sources

Zeev Dvir
Weizmann Institute of Science
Rehovot, Israel

Ariel Gabizon
Weizmann Institute of Science
Rehovot, Israel

Avi Wigderson
Institute for Advanced Study
Princeton, NJ

Abstract

In this paper we construct explicit deterministic extractors from polynomial sources*, namely from distributions sampled by low degree multivariate polynomials over finite fields. This naturally generalizes previous work on extraction from affine sources. A direct consequence is a deterministic extractor for distributions sampled by polynomial size arithmetic circuits over exponentially large fields.*

The first step towards extraction is a construction of rank extractors*, which are polynomial mappings that "extract" the algebraic rank from any system of low degree polynomials. More precisely, for any n polynomials, k of which are algebraically independent, a rank extractor outputs k algebraically independent polynomials of slightly higher degree. A result of Wooley allows us to relate algebraic rank and min-entropy and to show that a rank extractor is also a high quality* condenser *for polynomial sources over polynomially large fields.*

Finally, to turn this condenser into an extractor, we employ a theorem of Bombieri, giving a character sum estimate for polynomials defined over curves. It allows extracting all the randomness (up to a multiplicative constant) from polynomial sources over exponentially large fields.

1 Introduction

Randomness extraction has been a major research area for nearly two decades, and requires little introduction. One important reason is that the functions studied and constructed in this theory: extractors, dispersers, condensers, samplers, etc., turn out to do far more than required. While they are designed to convert weak sources of randomness into "high quality" random bits, they end up being essential in applications where randomness is not even an issue, such as expander constructions [24], error correction [21] and metric embedding [12], to name but a few examples.

Most of this research has concentrated on the so-called "seeded" extractors, which allow the use of a short, truly random seed, and enables handling extremely general classes of weak sources. An excellent survey of this broad field is [20]. More recently there has been a burst of activity on "seedless" or "deterministic" extractors, which can use no additional random "seed". The general question is for which classes of distributions is deterministic extraction possible. The main types of sources for which progress has been made include the following (overlapping) classes.

- *Few independent sources*: the given distribution is of several, independent weak sources, as in e.g. [7, 1, 2, 18, 16, 3].
- *Computational sources*: the given distribution is the output of some (space- or time-) efficient algorithm on a uniformly random input, as in e.g. [23, 4, 22, 13].
- *Bit-fixing sources*: the given distribution is fixed in some coordinates, and independent in others, as in e.g. [14, 9]
- *Affine sources*: the given distribution is the output of some affine map, applied to a random input as in e.g. [2, 6, 8]

Since our work is best viewed as extending the last class of sources, let us describe these results in some more detail. An *affine source* over a finite field $\mathbb{F}$ is a random variable which is uniformly distributed on some k-dimensional affine subspace of $\mathbb{F}^n$. Such a distribution is usually described by a non-degenerate affine mapping $x(t) : \mathbb{F}^k \mapsto \mathbb{F}^n$ defined by n linear functions

$$x(t) = (x_1(t_1, \ldots, t_k), \ldots, x_n(t_1, \ldots, t_k)),$$

in k variables. The affine source is thought of as the output of $x(t)$ on a uniformly chosen input $t \in \mathbb{F}^k$. Clearly, the entropy (and more importantly, min-entropy) of such sources is $k \cdot \log |\mathbb{F}|$ (all logarithms in this paper are base two).

The works of Barak et. al. [2] and of Bourgain [6] deal with the case of the binary field $\mathbb{F}_2$. The first gives an explicit disperser, and the second an extractor, for the case where $k = \Omega(n)$. In particular, Bourgain [6] extracts a constant fraction the entropy with exponentially small error for such k. No explicit construction is known for smaller rank (over $\mathbb{F}_2$) despite the fact that, non explicitly, extractors exist even for logarithmic rank.

Gabizon and Raz [8] show that if the field $\mathbb{F}$ is large, then one can even handle the case of 1-dimensional affine sources (distributions on affine lines). They show how to construct a deterministic extractor that extracts almost all the entropy (with polynomial error) for any given k, for fields $\mathbb{F}$ of size polynomial in n.

1.1 Low Degree Polynomial Sources

A natural generalization of affine sources is to allow our source to be sampled by low-degree multivariate polynomials over $\mathbb{F}$. We note that while low-degree polynomials play an essential role in complexity theory, extraction from sources defined by such polynomials has apparently not been studied before.

Let $\mathbb{F}$ be a field (finite or infinite). For integers $k \leq n$ and d we consider the family of all mappings $x : \mathbb{F}^k \mapsto \mathbb{F}^n$ that are defined by polynomials of total degree at most d (we denote our mapping by x since this will represent our source). That is,

$$x(t) = (x_1(t_1, \ldots, t_k), \ldots, x_n(t_1, \ldots, t_k)),$$

where for each $1 \leq i \leq n$ the coordinate x_i of the mapping is a k-variate polynomial of total degree at most d. We denote this set of mappings by $\mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$. We will focus on the case where the field $\mathbb{F}$ is much larger than d (we will specify in each result how large the field has to be). This will allow us to refer to the elements of $\mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ as *low degree* mappings. It is important to note that any weak source can be represented as the image of *some* polynomial mapping over a finite field $\mathbb{F}$. However, in general, the polynomials representing the source will have very high degrees (this can be seen by a simple counting argument). Since it is known [15] that deterministic extraction from arbitrary sources is impossible, we see that restricting our attention to low degree mappings is essential.

For affine sources we have the requirement that the affine mapping defining the source is non-degenerate. This ensures that the source sampled by this mapping has 'enough' entropy. We would like to extend this requirement also to the case of low degree mappings in $\mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$. The way to generalize this notion is via the partial derivative matrix (sometimes called the *Jacobian*) of a mapping $x \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$. This is an $n \times k$ matrix denoted $\frac{\partial x}{\partial t}$ defined as follows:

$$\frac{\partial x}{\partial t} \triangleq \begin{pmatrix} \frac{\partial x_1}{\partial t_1} & \cdots & \frac{\partial x_1}{\partial t_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial t_1} & \cdots & \frac{\partial x_n}{\partial t_k} \end{pmatrix},$$

where the partial derivatives are defined in the standard way, as formal derivatives of polynomials. Let us define the *rank* of $x \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ to be the rank of the matrix $\frac{\partial x}{\partial t}$ when considered as a matrix over the field of rational functions in variables $t_1, \ldots, t_k$. We say that $x \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ is *non-degenerate* if its rank is k (obviously, x cannot have rank larger than k).

Definition 1.1 (Polynomial Source). *Let $\mathbb{F}$ be a finite field. A distribution X over $\mathbb{F}^n$ is an (n, k, d)-polynomial source over $\mathbb{F}$, if there exists a non-degenerate mapping $x \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ such that X is sampled by choosing t uniformly at random in $\mathbb{F}^k$ and outputting $x(t)$.*

It is easy to see that the above definition of a polynomial source is indeed a generalization of the affine case, since the partial derivative matrix of an affine mapping is simply its coefficient matrix (in some basis).

Rank and min-entropy: One reason for using the rank of the partial derivative matrix is that, over sufficiently large prime fields, it allows us to prove a lower-bound on the entropy of an (n, k, d)-polynomial source. This lower bound follows from a theorem of Wooley [25] (see Theorem 2.1). Roughly speaking, Wooley's theorem implies that a distribution sampled by a non-degenerate mapping $x \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ is close (in statistical distance) to a distribution with min-entropy at least $k \cdot \log\left(\frac{|\mathbb{F}|}{2d}\right)$.

Rank and algebraic independence. Over fields of exponential characteristic (or of characteristic zero) it can be showen that the above notion of the rank of a mapping coincides with the more intuitive notion of *algebraic independence*. Roughly speaking, over such fields, a mapping $x = (x_1, \ldots, x_n) \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ has rank k iff the set of polynomials $\{x_1(t), \ldots, x_n(t)\}$ contains k algebraically independent polynomials (we should note that the direction "rank $k \mapsto$ algebraic independence" is true over any field, regardless of its characteristic). Since we want some of our results to hold also over fields of polynomial size we opt to use the rank of the partial derivative matrix in our definition of a polynomial source. In the full version of the paper we give a detailed discussion of the connection between algebraic independence and rank.

1.2 Rank Extractors

The above discussion of polynomial sources raises the following natural question: Can we 'extract' the rank of these sources without destroying their structure? In other words, can we construct a *fixed* polynomial mapping $y : \mathbb{F}^n \mapsto \mathbb{F}^k$ such that for any non-degenerate $x \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ the composition of y with x is a non-degenerate mapping from $\mathbb{F}^k$ to $\mathbb{F}^k$? We call a non-degenerate mapping $x : \mathbb{F}^k \mapsto \mathbb{F}^k$ a *full rank* mapping and a mapping y satisfying the above condition a *rank extractor*.

Definition 1.2 (Rank Extractor). *Let $\mathbb{F}$ be some field. Let $y : \mathbb{F}^n \mapsto \mathbb{F}^k$ be a polynomial mapping defined by*

$$y(x) = (y_1(x_1, \ldots, x_n), \ldots, y_k(x_1, \ldots, x_n)),$$

where each y_i is a multivariate polynomial over $\mathbb{F}$. We say that y is an (n, k, d)-rank extractor over $\mathbb{F}$ if for every non-degenerate mapping $x \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ the composition $y \circ x : \mathbb{F}^k \mapsto \mathbb{F}^k$ has rank k. We will call such a mapping y **explicit** *if it can be computed in polynomial time.*

Clearly, a construction of a rank extractor will bring us closer to constructing an extractor for low degree polynomial sources. Using an explicit rank extractor reduces the problem of constructing an extractor for arbitrary polynomial sources into the problem of constructing an extractor for polynomial sources of full rank. This problem, as we shall see later, can be solved using tools from algebraic geometry.

Our first main result is a construction of an explicit (n, k, d)-rank extractor over $\mathbb{F}$, where $\mathbb{F}$ can be any field of characteristic zero or of characteristic at least poly(n, d). It is natural to require that the degree of the rank extractor will be as small as possible. Clearly the degree has to be larger than 1 since an affine mapping cannot be a rank extractor (we can always 'hide' a polynomial source in the kernel of such a mapping). The rank extractors we construct have degree which is bounded by a polynomial in n and in d. In Section 3 we prove the following theorem:

Theorem 1. *Let $k \leq n$ and d be integers. Let $\mathbb{F}$ be a field of characteristic zero or of characteristic larger than $8k^2d^3n$. There exists an explicit (n, k, d)-rank extractor over $\mathbb{F}$ whose degree is bounded by $8k^2d^2n$. Moreover, this rank extractor can be computed in time poly$(n, \log(d))$.*

We note that our construction of rank extractors does not depend on the underlying field. We give a single construction, defined using integers, that is a rank extractor over any field satisfying the conditions of Theorem 1.

1.3 Extractors & Condensers for Polynomial Sources

As was mentioned in the previous section, applying the rank extractor given by Theorem 1 reduces the problem of constructing an extractor for (n, k, d)-polynomials sources into the problem of constructing an extractor for (k, k, d')-polynomial sources, where d' is the degree of the source obtained **after** applying the rank extractor (Theorem 1 implies that d' is polynomial in n and d). Our second main result is a construction of such an extractor. Before stating our result we give a formal definition of an extractor for polynomial sources.

Definition 1.3 (Extractor). *Let $k \leq n$ and d be integers. Let $\mathbb{F}$ be a finite field. A function $E : \mathbb{F}^n \mapsto \{0, 1\}^m$ is a (k, d, ϵ)-extractor for polynomial sources if for every (n, k, d)-polynomial source X over $\mathbb{F}^n$, the random variable $E(X)$ is ϵ-close to uniform. We say that E is* **explicit** *if it can be computed in poly$(n, \log(d))$ time.*

The following theorem, which we prove in Section 4, shows the existence of an explicit extractor for full rank polynomial sources over sufficiently large prime fields. The output length of this extractor is $\Omega(k \cdot \log(|\mathbb{F}|))$ - within a multiplicative constant of the maximal length possible. The main tool in the proof of this theorem is a theorem of Bombieri [5] giving exponential sum estimates for polynomials defined over low degree curves.

Theorem 2. *There exists absolute constants C and c such that the following holds: Let k and $d > 1$ be integers and let $\mathbb{F}$ be a field of prime cardinality $p > d^{Ck}$. Then, there exists a function $E : \mathbb{F}^k \mapsto \{0, 1\}^m$ that is an explicit (k, d, ϵ)-extractor for polynomial sources over $\mathbb{F}^k$ with $m = \lfloor c \cdot k \cdot \log(p) \rfloor$ and $\epsilon = p^{-\Omega(1)}$.*

Combining this last theorem with Theorem 1 gives an extractor for general polynomial sources. This extractor, whose existence is stated in the following corollary, also has output length which is within a multiplicative constant of optimal.

Corollary 1.4. *There exists absolute constants C and c such that the following holds: Let $k \leq n$ and $d > 1$ be integers and let $d' = 8k^2d^3n$. Let $\mathbb{F}$ be a field of prime cardinality $p > (d')^{Ck}$. Then, there exists a function $E : \mathbb{F}^k \mapsto \{0, 1\}^m$ that is an explicit (k, d, ϵ)-extractor for polynomial sources over $\mathbb{F}^n$ with $m = \lfloor c \cdot k \cdot \log(p) \rfloor$ and $\epsilon = p^{-\Omega(1)}$.*

It is possible to improve the output length of our extractors so that it is equal to a $(1 - \alpha)$-fraction of the source min entropy, for any constant $\alpha > 0$. This improvement, which was suggested to us by Salil Vadhan is described in Section 5.

We note that in both the last corollary and in Theorem 2, the bound on the field size does not pose a computational problem. Over a finite field $\mathbb{F}$, arithmetic operations can be performed in time polynomial in $\log(|\mathbb{F}|)$, and hence all computations required by the extractor can be performed in polynomial time. However, it remains an interesting open problem whether extraction can be performed over smaller fields, say of size polynomial in n and in d.

Condensers Over Polynomially Large Fields: Over Polynomially large fields, our techniques give a deterministic *condenser* for polynomial sources. A condenser is a relaxation of an extractor and is required to output a distribution with 'high' min-entropy rather than a uniform distribution. The word 'condenser' implies that the length of the output should be smaller then the length of the input. That is, the aim of a condenser is to 'compress' the source while keeping as much of the entropy as possible. For convenience we define condensers as mappings over alphabet $\mathbb{F}$ rather than the standard definition using binary alphabet.

Definition 1.5 (Condenser). *Let $\mathcal{D}$ be a family of distributions over $\mathbb{F}^n$. A function $C : \mathbb{F}^n \mapsto \mathbb{F}^m$ is an (ϵ, k')-*condenser *for $\mathcal{D}$ if for every $X \in \mathcal{D}$ the distribution $C(X)$ is ϵ-close to having min-entropy at least k'. A condenser is* explicit *if it can be computed in polynomial time.*

From Wooley's theorem [25], mentioned earlier, it follows that if we apply a rank extractor to a polynomial source we get a source which is close to having high min-entropy. The next theorem follows immediately by combining Theorem 1 and Wooley's Theorem (Corollary 2.2).

Theorem 3. *Let $k \leq n$ and d be integers. Let $\mathbb{F}$ be a field of prime cardinality larger than $d' = 8k^2d^3n$. Let $y : \mathbb{F}^n \mapsto \mathbb{F}^k$ be the rank extractor from Theorem 1. Then y is an (ϵ, k')-condenser for the family of (n, k, d)-polynomial sources over $\mathbb{F}$, where $\epsilon = \frac{d' \cdot k}{|\mathbb{F}|}$ and $k' = k \cdot \log(|\mathbb{F}|/2d')$.*

It should be noted that this condenser is 'almost' the best one could hope for (without building an extractor, of course). To see this, suppose that $|\mathbb{F}| \approx (2d')^c$ for some constant $c > 1$. We get that the output of the condenser is close to having min-entropy

$$k' = k \cdot \log(|\mathbb{F}|/2d') \approx \left(1 - \frac{1}{c}\right) \cdot k \cdot \log(|\mathbb{F}|),$$

and so the ratio between the length of the output (in bits) and its min-entropy can be made arbitrarily close to one by choosing c to be large enough.

1.4 Rank vs. Entropy - Weak Polynomial Sources

So far we focused on extraction from sources which were defined algebraically - we were given a bound on the algebraic rank of the set of polynomials we extract from. We now switch to the more standard definition (from the extractor literature standpoint) of extraction from sources with given min-entropy. These will be called *Weak Polynomial Sources*.

Definition 1.6 (Weak Polynomial Source). *A distribution X over $\mathbb{F}^n$ is an (n, k, d)-weak polynomial source (WPS) if*

- *There exists a polynomial mapping $x \in \mathcal{M}(\mathbb{F}^n \mapsto \mathbb{F}^n, d)$ such that X is sampled by choosing t uniformly in $\mathbb{F}^n$ and outputting $x(t)$.*
- *X has min entropy at least $k \cdot \log(|\mathbb{F}|)$.*

Notice in the definition that the min-entropy threshold is $k \cdot \log(|\mathbb{F}|)$ (instead of just k). This is to hint to the connection (which we prove later) between the rank of the source and its entropy. Intuitively, a distribution sampled by a rank r mapping $x : \mathbb{F}^n \mapsto \mathbb{F}^n$ "should" have entropy roughly $r \cdot \log(|\mathbb{F}|)$ and indeed, for affine sources, this is exactly the case.

The following theorem, whose proof can be found in the full version of this paper, shows the existence of an explicit deterministic extractor for the class of weak polynomial sources.

Theorem 4. *There exists absolute constants C and c such that the following holds: Let $k \leq n$ and $d > 1$ be integers and let $d' = 8k^2d^3n$. Let $\mathbb{F}$ be a field of prime cardinality $p > (d')^{Ck}$. Then, there exists a function $E : \mathbb{F}^n \mapsto \{0, 1\}^m$ that is an explicit (k, d, ϵ)-extractor for weak polynomial sources over $\mathbb{F}^n$ with $m = \lfloor c \cdot k \cdot \log(p) \rfloor$ and $\epsilon = p^{-\Omega(1)}$.*

The parameters of the extractor given by the theorem can be seen to be roughly the same as those of the extractor for regular polynomial sources (Corollary 1.4). In fact, the extractor we use for weak polynomial sources is the same one we used for polynomial sources. The proof of Theorem 4 will follow from showing that any (n, k, d)-WPS is close (in statistical distance) to a convex combination of (n, k, d)-polynomial sources. Clearly, this will imply that any extractor that works for polynomial sources will work also for weak polynomial sources.

The Entropy of a Polynomial Mapping: We can use the methods employed in the proof of Theorem 4 to show that over sufficiently large fields, the entropy of the output of a low degree polynomial mapping $x \in \mathcal{M}(\mathbb{F}^n \mapsto \mathbb{F}^n, d)$ is always 'close' to rank$(x) \cdot \log(|\mathbb{F}|)$. This can be viewed as a generalization of the simple fact that for an *affine* mapping x, the entropy is always equal to rank$(x) \cdot \log(|\mathbb{F}|)$. The formal statement of this result is as follows:

Theorem 1.7. *Let $k \leq n$ and d be integers. Let $D = (2k+1)d^{2k}$ and let $0 < \delta < 1$ be a real number. Let $\mathbb{F}$ be a field of prime cardinality p such that $p > \max\{(2d)^{\frac{k}{\delta}}, 2^{\frac{10}{\delta}}, (2D)^{\frac{2}{\delta}}\}$. Let $x \in \mathcal{M}(\mathbb{F}^n \mapsto \mathbb{F}^n, d)$ be of rank k and let $X = x(U_n)$ be the distribution sampled by x. Then*

1. *X has min entropy $\leq (k + \delta) \cdot \log(p)$.*
2. *X is ϵ-close to having min entropy $\geq (k - \delta) \cdot \log(p)$, where $\epsilon = \frac{2 \cdot d \cdot k}{p}$.*

Extractors for Poly-Size Arithmetic Circuits: An interesting corollary of Theorem 4 is the existence of deterministic extractors for the class of distributions sampled by polynomial sized arithmetic circuits over exponentially large fields. This follows from the fact that the degrees of the polynomials computed by poly-size circuits are at most exponential, and the construction of an (n, k, d)-rank extractor is efficient even when d is exponential (since the dependence on d is poly-logarithmic).

We say that a distribution X on $\mathbb{F}^n$ is sampled by a size s arithmetic circuit if there exists an arithmetic circuit A of size s with n inputs and n outputs such that the fan-in of each gate is at most two and such that X is the distribution of the output of A on a random input, chosen uniformly from $\mathbb{F}^n$. We say that X is an (n, k, s)-*arithmetic source* if X is sampled by a size s arithmetic circuit and its min-entropy is at least $k \cdot \log(|\mathbb{F}|)$.

Corollary 1.8. *There exists absolute constants C and c such that the following holds: Let $k \leq n$ and $s > 1$ be integers. Let $d = 2^s$ and let $d' = 8k^2d^3n$. Let $\mathbb{F}$ be a field of prime cardinality $p > (d')^{Ck}$. Then, there exists an explicit function $E : \mathbb{F}^n \mapsto \{0,1\}^m$ such that for every (n, k, s)-arithmetic source X over $\mathbb{F}$, the distribution of $E(X)$ is ϵ-close to uniform, where $m = \lfloor c \cdot k \cdot \log(p) \rfloor$ and $\epsilon = p^{-\Omega(1)}$. That is, E is an extractor for the class of (n, k, s)-arithmetic sources.*

It is interesting to contrast this result to the extractors of [22] from polynomial size *boolean* circuits. Their extractors rely on complexity assumptions, and they prove that such assumptions are necessary. It is interesting that over large fields no such assumptions, nor lower bounds, are necessary.

1.5 Organization

Section 2 contains general preliminaries on probability distributions and finite field algebra. In Section 3 we describe our construction of a rank extractor and prove Theorem 1. In Section 4 we construct and analyze an extractor for full rank polynomial sources and prove Theorem 2. In Section 5 we show how to increase the output length of our extractors. In Section 6 we discuss some open problems related to our results.

2 General Preliminaries

2.1 Probability Distributions

The *statistical distance* between two distributions P and Q on Ω, denoted by $|P - Q|$, is defined as

$$|P - Q| \triangleq \max_{S \subseteq \Omega} \left| \Pr_P(S) - \Pr_Q(S) \right|$$

We say that P is ϵ-*close* to Q, denoted $P \overset{\epsilon}{\sim} Q$, if $|P - Q| \leq \epsilon$. We denote the fact that P and Q are identically distributed by $P \sim Q$. We use *min-entropy* to measure the amount of randomness in a given distribution. Let X be a distribution over a finite set Γ. The min-entropy of X is defined as

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

2.2 Polynomials Over Finite Fields

For a field $\mathbb{F}$ we denote by $\mathbb{F}[t_1, \ldots, t_k]$ the ring of polynomials in k-variables $t_1, \ldots, t_k$ with coefficients in $\mathbb{F}$. We denote by $\mathbb{F}(t_1, \ldots, t_k)$ the field of rational functions in variables $t_1, \ldots, t_k$. We denote by $\deg(f)$ the total degree of f and by $\deg_{t_j}(f)$ the degree of f as a polynomial in t_j. We write $f \equiv 0$ or $f(t) \equiv 0$ if f is the zero polynomial (all coefficients of f are zero). Note that over the finite field $\mathbb{F}$ of prime cardinality p, the polynomial $f(t) = t^p - t$ is **not** the zero polynomial, even though $f(a) = 0$ for all $a \in \mathbb{F}$.

For a polynomial $f \in \mathbb{F}[t_1, \ldots, t_k]$ we denote by $\frac{\partial f}{\partial t_j} \in \mathbb{F}[t_1, \ldots, t_k]$ the formal partial derivative of f with respect to the variable t_j. For a vector of polynomials $\bar{f} = (f_1, \ldots, f_m) \in (\mathbb{F}[t_1, \ldots, t_k])^m$ we can define the *partial derivative matrix* of $\bar{f}$ as

$$\frac{\partial \bar{f}}{\partial t} \triangleq \begin{pmatrix} \frac{\partial f_1}{\partial t_1} & \cdots & \frac{\partial f_1}{\partial t_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial t_1} & \cdots & \frac{\partial f_m}{\partial t_k} \end{pmatrix}.$$

We denote by rank$(\bar{f})$ the rank, over $\mathbb{F}(t_1, \ldots, t_k)$, of the matrix $\frac{\partial \bar{f}}{\partial t}$.

2.3 The Number of Solutions to a System of Polynomial Equations

We will use a version of Bezout's Theorem proved by Wooley [25]. This theorem, mentioned informally in the introduction, will give us a connection between algebraic rank and min entropy. We note that the formulation of Wooley's theorem stated here is weaker then the original formulation appearing in [25] (the original form of the theorem speaks of congruences modulo p^s for any s).

Theorem 2.1 (Rephrased from Theorem 1 in [25]). *Let $\mathbb{F}$ be a field of prime cardinality p. Let k and d be integers. Let $x = (x_1, \ldots, x_k) \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^k, d)$ be such that rank$(x) = k$ and denote by $J(t) \triangleq \det\left(\frac{\partial x}{\partial t}\right)(t)$. For $a \in \mathbb{F}^k$ let*

$$N_a \triangleq \left|\left\{c \in \mathbb{F}^k \quad : \quad x(c) = a \quad \textit{and} \quad J(c) \neq 0\right\}\right|.$$

Then for every $a \in \mathbb{F}^k$, $N_a \leq d^k$.

We can interpret this theorem as saying that a distribution X sampled by a non-degenerate mapping $x \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^k, d)$ is close to a distribution with high min-entropy, where the closeness is related to the number of zeros of the determinant of $\frac{\partial x}{\partial t}$. Since this determinant is a non-zero low-degree polynomial, we get that the distance from the high min-entropy distribution is small. This is stated more precisely by the following Corollary, which also extends our view to mappings in $\mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ for $k \leq n$. The (easy) proof of this Corollary is omitted due to space constraints.

Corollary 2.2. *Let $\mathbb{F}$ be a field of prime cardinality. Let $k \leq n$ and d be integers such that $|\mathbb{F}| > 2dk$. Let X be an (n, k, d)-polynomial source over $\mathbb{F}$. Then X is ϵ-close to a distribution with min-entropy at least $k \cdot \log\left(\frac{|\mathbb{F}|}{2d}\right)$, where $\epsilon = \frac{d \cdot k}{|\mathbb{F}|}$.*

3 An Explicit Rank Extractor

In this section we describe our construction of a rank extractor and prove Theorem 1.

Construction 1. *Let $k \leq n$ and d be integers. Let $s_2 = dk + 1$ and $s_1 = (2dn + 1) \cdot s_2$. Let $l_{ij} = i \cdot (s_1 + j \cdot s_2)$. Define for each $1 \leq i \leq k$*

$$y_i(x) = y_i(x_1, \ldots, x_n) \triangleq \sum_{j=1}^{n} \frac{1}{l_{ij} + 1} \cdot x_j^{l_{ij}+1}.$$

Let $y = (y_1, \ldots, y_k) : \mathbb{F}^n \mapsto \mathbb{F}^k$ be the output of the construction. Notice that $y(x)$ is defined in such a way that the partial derivative $\frac{\partial y_i}{\partial x_j}$ is exactly $x_j^{l_{ij}}$.

We prove the following theorem, which directly implies Theorem 1.

Theorem 3.1. *Let $\mathbb{F}$ be a field of characteristic zero or of characteristic larger than $d' = 8k^2d^3n$. Let $x \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ be of rank k. Let $y : \mathbb{F}^n \mapsto \mathbb{F}^k$ be as in Construction 1. Then the composition $(y \circ x)(t)$ is in $\mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^k, d')$ and has rank k.*

3.1 Preliminaries For The Proof Of Theorem 3.1

3.1.1 Sums of Powers of Polynomials

The following lemma shows how to pick integers $c_1, \ldots, c_n$ in such a way that for any set of n polynomials $x_1(t), \ldots, x_n(t)$ of bounded degree, the polynomials $x_1(t)^{c_1}, \ldots, x_n(t)^{c_n}$ will have degrees that are different by at least some fixed number. The proof of this lemma can be found in the full version of the paper.

Lemma 3.2. *Let $x_1(t), \ldots, x_n(t)$ be k-variate non-constant polynomials over some field $\mathbb{F}$. Denote by $d_i > 0$ the degree of the polynomial x_i. Let $d \geq \max_i\{d_i\}$. Let A and B be two positive integers such that $A \geq (2dn + 1) \cdot B$ and let $c_i \triangleq A + Bi$ for $i \in [n]$. Then, for every $1 \leq i < j \leq n$, we have*

$$|\deg(x_i(t)^{c_i}) - \deg(x_j(t)^{c_j})| = |d_i \cdot c_i - d_j \cdot c_j| \geq B.$$

3.1.2 The Cauchy-Binet Formula

The Cauchy-Binet formula gives the determinant of the product of a $k \times n$ matrix with an $n \times k$ matrix (for $k \leq n$). Let $k \leq n$. Let A be a $k \times n$ matrix and B an $n \times k$ matrix. For a set $I \subset [n]$ of size k we denote by A_I the $k \times k$ sub-matrix of A composed of the columns of A whose indices appear in I. Similarly, we denote by B_I the sub-matrix of B composed of the rows of B whose indices are in I. The proof of the following formula can be found in [10].

Lemma 3.3 (Cauchy-Binet). *Let $k \leq n$. Let A be a $k \times n$ matrix and B an $n \times k$ matrix over a field $\mathbb{F}$. Using the above notations we have*

$$\det(A \cdot B) = \sum_{\substack{I \subset [n] \\ |I| = k}} \det(A_I) \cdot \det(B_I).$$

3.2 Proof Of Theorem 3.1

Let $k \leq n$, d be integers. Let $\mathbb{F}$ be a field of characteristic zero or of characteristic larger than $d' = 8k^2d^3n$. Let $x = (x_1, \ldots, x_n) \in \mathcal{M}(\mathbb{F}^k \mapsto \mathbb{F}^n, d)$ be such that rank$(x) = k$. Let $y : \mathbb{F}^n \mapsto \mathbb{F}^k$ be defined as in Construction 1, that is

$$y_i(x) = y_i(x_1, \ldots, x_n) \triangleq \sum_{j=1}^{n} \frac{1}{l_{ij} + 1} \cdot x_j^{l_{ij}+1}, \quad (1)$$

where

$$l_{ij} = i \cdot (s_1 + j \cdot s_2)$$
$$s_1 = (2dn + 1) \cdot s_2 \quad , \quad s_2 = dk + 1$$

It is easy to verify that the degree of the mapping y is bounded by $8k^2d^2n$. Therefore, the degree of the composition $(y \circ x)(t)$ is bounded by $d' = 8k^2d^3n$. Therefore, since the characteristic of $\mathbb{F}$ is larger than d' (or is zero), for the rest of the proof we don't need to worry about non constant polynomials becoming zero after we take their derivative.

Our goal is to show that the composition $y \circ x$ has rank k. In order to prove this we need to show that the determinant of the partial derivatives matrix of the composition is non zero. Write $y(t)$ to denote $y(x(t))$ and let $\frac{\partial y}{\partial t}$ denote the $k \times k$ partial derivative matrix of the mapping $y(t)$. Using the chain rule we have that

$$\frac{\partial y}{\partial t} = \frac{\partial y}{\partial x} \cdot \frac{\partial x}{\partial t},$$

where $\frac{\partial y}{\partial x}$ is a $k \times n$ matrix and $\frac{\partial x}{\partial t}$ is an $n \times k$ matrix. All the elements in these two matrices are polynomials in t, since we evaluate $\frac{\partial y}{\partial x}$ at $x = x(t)$.

Consider the element at position (i, j) in the matrix $\frac{\partial y}{\partial x}$. Taking the derivative of (1) with respect to x_j we get that

$$\frac{\partial y_i}{\partial x_j} = x_j(t)^{l_{ij}} = x_j(t)^{i \cdot (s_1 + j s_2)}.$$

The Vandermonde structure of $\frac{\partial y}{\partial x}$ becomes more apparent by denoting $r_j(t) \triangleq x_j(t)^{s_1 + j s_2}$. We now have that the (i, j)'th element of $\frac{\partial y}{\partial x}$ is $r_j(t)^i$. That is

$$\frac{\partial y}{\partial x} = \begin{pmatrix} r_1(t) & r_2(t) & \cdots & \cdots & r_n(t) \\ r_1(t)^2 & r_2(t)^2 & \ddots & & r_n(t)^2 \\ \vdots & \vdots & & \ddots & \vdots \\ r_1(t)^k & r_2(t)^k & \cdots & \cdots & r_n(t)^k \end{pmatrix}.$$

To facilitate writing, let us denote by $R \triangleq \frac{\partial y}{\partial x}$ and $D \triangleq \frac{\partial x}{\partial t}$. We can also assume w.l.o.g that

$$\deg(r_1(t)) \leq \ldots \leq \deg(r_n(t)), \tag{2}$$

(we let $\deg(0) = 0$) since applying the same permutation on the rows of R and on the columns of D will not change the determinant of $R \cdot D$. Now, from Lemma 3.3 (Cauchy-Binet) and using the notations of Section 3.1.2 we have that

$$\det\left(\frac{\partial y}{\partial t}\right) = \det(R \cdot D) = \sum_{\substack{I \subset [n] \\ |I| = k}} \det(R_I) \cdot \det(D_I) \tag{3}$$

Notice that if $r_i(t)$ is constant, then $x_i(t)$ is also constant and so the i'th row of the matrix D is zero. Therefore, $\det(D_I) = 0$ for every I that contains an index i such that $r_i(t)$ is constant. In view of (3) and this last observation, we can assume w.l.o.g that for all $i \in [n]$, $r_i(t)$ is non constant. (Notice that since D has maximal rank, we have at least k indices in $[n]$ for which $x_i(t)$ is non constant and so the condition $n \geq k$ is maintained).

The next three claims will show that there exist a unique set I in the above sum for which the degree of $\det(R_I) \cdot \det(D_I)$ is maximal. This will conclude the proof, since then we will have that $\det\left(\frac{\partial y}{\partial t}\right)$ is non zero, as required.

We start with a simple claim showing that the degrees of the polynomials $r_i(t)$ have large gaps between them.

Claim 3.4. *Let $r_1(t), \ldots, r_n(t)$ be the polynomials defined above. Then for every $i \in [n-1]$ we have*

$$\deg(r_{i+1}(t)) > \deg(r_i(t)) + dk.$$

Proof. Recall that $r_i(t) = x_i(t)^{s_1 + j \cdot s_2}$ and that $s_1 \geq (2dn+1) \cdot s_2$. Using Lemma 3.2 we get that

$$|\deg(r_{i+1}(t)) - \deg(r_i(t))| \geq s_2 > dk.$$

Using (2) the claim follows. □

Let $I \subset [n]$ be such that $|I| = k$. We denote by

$$d_I \triangleq \deg\left(\det(R_I)\right).$$

The next claim gives a convenient formula for d_I.

Claim 3.5. *Let $I \subset [n]$, $I = \{i_1 < \ldots < i_k\}$. Then*

$$d_I = \deg(R_I) = \sum_{j=1}^{k} j \cdot \deg\left(r_{i_j}(t)\right).$$

Proof. Using the Vandermonde structure of the matrix R_I we get that

$$\det(R_I) = \prod_{j=1}^{k} r_{i_j}(t) \prod_{1 \leq j_1 < j_2 \leq k} \left(r_{i_{j_1}}(t) - r_{i_{j_2}}(t)\right).$$

In view of (2), the degree of the highest monomial in $\det(R_I)$ is obtained my multiplying k copies of $r_{i_k}(t)$ with $k-1$ copies of $r_{i_{k-1}}(t)$ and so on. this will give a monomial with degree $\sum_{j=1}^{k} j \cdot \deg(r_j(t))$. □

Define

$$\Gamma \triangleq \{I \subset [n] \mid |I| = k \,,\, \det(D_I) \neq 0\}$$

The next and final claim shows that there exists a **unique** $I \in \Gamma$ with maximal d_I. The proof uses standard techniques from matroid theory and can be found in the full version of the paper.

Claim 3.6. *Let $d_{\max} \triangleq \max_{I \in \Gamma}\{d_I\}$. Then there exists a unique $I^* \in \Gamma$ such that $d_{I^*} = d_{\max}$. Moreover, for every $I \neq I^*$ we have that $d_I < d_{I^*} - dk$.*

We can now use Claim 3.6 to show that the sum in (3) is not zero. Let $I^* \in \Gamma$ be the set with unique maximal d_{I^*} given by Claim 3.6. Rewrite (3) in the following form

$$\begin{aligned} \det(R \cdot D) &= \sum_{\substack{I \subset [n] \\ |I| = k}} \det(R_I) \cdot \det(D_I) \\ &= \sum_{I \in \Gamma} \det(R_I) \cdot \det(D_I) \\ &= \det(R_{I^*}) \cdot \det(D_{I^*}) + \\ &\quad \sum_{I \in \Gamma I \neq I^*} \det(R_I) \cdot \det(D_I). \end{aligned} \tag{4}$$

The degree of the first summand in (4) is at least

$$\deg\left(\det(R_{I^*}) \cdot \det(D_{I^*})\right) = d_{I^*} + \deg\left(\det(D_{I^*})\right) \geq d_{I^*}.$$

Using Claim 3.6 we can upper bound the degrees of the other summands in (4). That is, for all $I \in \Gamma$ different from I^* we have

$$\begin{aligned}\deg\left(\det(R_I)\cdot\det(D_I)\right) &= \\ d_I+\deg\left(\det(D_I)\right) &\leq d_I+dk<d_{I^*},\end{aligned}$$

(we use the fact that all the entries of D are polynomials of degree at most d). Therefore, the sum in (4) cannot be zero. This concludes the proof of Theorem 3.1. □

4 Extractors for Polynomial Sources

In this section we describe our construction of an extractor for full rank polynomial sources and prove Theorem 2. As was mentioned in the introduction, this construction, together with the rank extractor constructed in previous sections, will give an extractor for polynomial sources of any rank. In order to describe our construction we require some additional notations. Let $\mathbb{F}$ be a field of prime cardinality p. For an integer $M \leq p$, we denote by $\text{mod}_M : \mathbb{F} \mapsto \{0,\ldots,M-1\}$ the modulo-M function. For a vector $x \in \mathbb{F}^n$ we apply the function $\text{mod}_M(x)$ coordinate wise. The following theorem directly implies Theorem 2.

Theorem 4.1. *There exist absolute constants $C > 0$ and $c > 0$ such that the following holds: Let k, d be integers and let $\mathbb{F}$ be a field of prime cardinality $p > d^{Ck}$. Let $m > 0$ be an integer such that $m < c \cdot log(p)$, let $M = 2^m$ and define the function $E : \mathbb{F}^k \mapsto \{0,1\}^{km}$ as $E(y) \triangleq mod_M(y)$. Then for every (k,k,d)-polynomial source Y over $\mathbb{F}$, the distribution $E(Y)$ is ϵ-close to uniform with $\epsilon = p^{-\Omega(1)}$.*

Notice that the construction of the extractor is very simple - taking a module in each coordinate. Proving that this is an extractor is much more complicated. The main tool in the proof of Theorem 4.1 will be a theorem of Bombieri [5] giving an exponential sum estimate for low degree polynomials defined over curves (one dimensional varieties). This section uses basic notions from algebraic geometry that can be found in any elementary text on the subject.

4.1 Preliminaries for the proof of Theorem 4.1

4.1.1 Block Distributions

Our proof will rely on the following standard lemma concerning block distributions. We will use the notation that for a vector $v = (v_1,\ldots,v_n)$ and for an index $i \in [n]$ we have $v^{(-i)} = (v_1,\ldots,v_{i-1},v_{i+1},\ldots,v_n)$. In some places we will define a new vector of length $n-1$ by writing $u = u^{(-i)} \in A^{n-1}$. This means that the indices of u go from 1 to n, skipping the i'th index. That is, $u = (u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_n) \in A^{n-1}$.

Lemma 4.2. *Let A be some finite set and let $X = (X_1,\ldots,X_k)$ be a distribution on A^k. Let $0 < \epsilon < 1$ and suppose that for each $1 \leq i \leq k$ there exists a set $S_i \subset A^{k-1}$ such that*

1. *$\mathbf{Pr}[X^{(-i)} \in S_i] \geq 1-\epsilon$ and*
2. *For each $s^{(-i)} \in S_i$, the conditional distribution $(X_i|X^{(-i)} = s^{(-i)})$ is ϵ-close to uniform.*

Then X is $O(k\cdot\sqrt{\epsilon})$-close to uniform.

4.1.2 Distributions With Small Fourier Coefficients

The following lemma is an extension of the now folklore Vazirani XOR Lemma [11] and is used [6, 3] to extract randomness from distributions with bounded Fourier coefficients. What the lemma says is that if we have a distribution X with a bound of $p^{-\Omega(1)}$ on all of its Fourier coefficients then we can deterministically extract from X (using the modulo function) $\Omega(\log(p))$ bits that are $p^{-\Omega(1)}$-close to uniform. The following formulation of the lemma follows from the version proved in [17].

Lemma 4.3. *Let p be a prime number and let $0 < \alpha < 1$ be such that $\log(p) < p^{\alpha/2}$. Let X be a distribution on $\mathbb{F}$ - the field of p elements. Suppose that for every non-trivial additive character $\chi : \mathbb{F} \mapsto \mathbb{C}^*$ we have the bound $\mathbb{E}[\chi(X)] \leq p^{-\alpha}$. Let $m = \lfloor(\alpha/2)\cdot\log(p)\rfloor$, let $M = 2^m$ and let $Y = mod_M(X)$ be an m-bit random variable. Then Y is $p^{-\alpha/4}$-close to uniform.*

4.1.3 A Theorem of Bombieri

The next ingredient required for the proof of Theorem 4.1 is an exponential sum estimate due to Bombieri [5]. We quote here a weak version of Bombieri's Theorem which is sufficient for our needs. This theorem provides the necessary bounds required to apply the XOR lemma above.

Theorem 4.4 (Theorem 6 in [5]). *Let p be a prime and let $1 < d$ be an integer such that $d^n < p$. Let $\mathbb{F}$ be the field of p elements and let $\bar{\mathbb{F}}$ be its algebraic closure. Let $f_1,\ldots,f_{n-1} \in \mathbb{F}[x_1,\ldots,x_n]$ be $n-1$ polynomials of degree $\leq d$ such that the set $\hat{V} = \{x \in \bar{\mathbb{F}}^n | f_1(x) = \ldots = f_{n-1}(x) = 0\}$ is a curve. Let $g \in \mathbb{F}[x_1,\ldots,x_n]$ be a polynomial of degree $\leq d$ that is non-constant on at least one of the irreducible components of $\hat{V}$. Let $\hat{V} = \hat{V}_1 \cup \ldots \cup \hat{V}_L$ be the decomposition of $\hat{V}$ into irreducible components. Let $\hat{U}$ be the union of those irreducible components of $\hat{V}$ on which $g(x)$ is non constant and let $U = \hat{U} \cap \mathbb{F}$. Let $\chi : \mathbb{F} \mapsto \mathbb{C}^*$ be a non-trivial additive character of $\mathbb{F}$. Then*

$$\left|\sum_{x\in U}\chi(g(x))\right| \leq 4d^{2n}\cdot p^{1/2}.$$

4.1.4 Intersections of Hypersurfaces

Consider a system of $n-1$ polynomial equations in n variables. The next lemma gives a bound on the number of ’shifts’ of the system for which the set of solutions has dimension larger than one (for the precise meaning of ’shift’ see the lemma). The proof of the lemma is omitted due to space constraints.

Lemma 4.5. *Let $\mathbb{F}$ be a finite field of size p and let $\bar{\mathbb{F}}$ denote its algebraic closure. Let $f_1,\ldots,f_{n-1} \in \mathbb{F}[x_1,\ldots,x_n]$ be polynomials of degree $\leq d$. For every $a = (a_1,\ldots,a_{n-1}) \in \mathbb{F}^{n-1}$ let $\hat{V}_a = \{x \in \bar{\mathbb{F}}^n \mid f_i(x) = a_i\, , i \in [n-1]\}$ and let $A = \{a \in \mathbb{F}^{n-1} \mid \hat{V}_a \neq \emptyset$ and $\dim(\hat{V}_a) \neq 1\}$. Then $|A| \leq nd^n p^{n-2}$.*

4.2 Proof of Theorem 4.1

Let $Y : \mathbb{F}^k \mapsto \mathbb{F}^k$ be a (k,k,d)-polynomial source and let $f = (f_1,\ldots,f_k) \in \mathbb{F}[x_1,\ldots,x_k]$ be a vector of polynomials of degree at most d such that $Y(x) = f(x) = (f_1(x),\ldots,f_k(x))$. For $i \in [k]$ and $a = a^{(-i)} \in \mathbb{F}^{k-1}$, we let $V_a = \left\{x \in \mathbb{F}^k \mid f^{(-i)}(x) = a\right\}$ and also $\hat{V}_a = \left\{x \in \bar{\mathbb{F}}^k \mid f^{(-i)}(x) = a\right\}$, where $\bar{\mathbb{F}}$ denotes the algebraic closure of $\mathbb{F}$. For a non trivial additive character $\chi : \mathbb{F} \mapsto \mathbb{C}^*$, such that $V_a \neq \emptyset$ we define the exponential sum

$$\Upsilon_\chi(a) = \frac{1}{|V_a|}\sum_{x \in V_a} \chi(f_i(x)).$$

In view of Lemma 4.2 and Lemma 4.3 the theorem will follow from the following lemma.

Lemma 4.6. *Using the above notations, there exists $0 < \alpha < 1$ such that for every $i \in [k]$ there exists a set $S_i \subset \mathbb{F}^{k-1}$ such that*

1. *$f^{(-i)}(x)$ lands in S_i with probability at least $1 - p^{-\alpha}$, when x is chosen uniformly in $\mathbb{F}^k$.*
2. *For every $a = a^{(-i)} \in S_i$ and for every non trivial χ, $|\Upsilon_\chi(a)| \leq p^{-\alpha}$.*

Before proving the lemma we proceed to show how it is used to complete the proof of Theorem 4.1. Let us denote by

$$Z_i = \mathrm{mod}_M(f_i(x))$$

the random variable representing the i’th block of $E(Y)$. Let $0 < \alpha < 1$ be the constant given by Lemma 4.6. Let $i \in [k]$ and let $S_i \subset \mathbb{F}^{k-1}$ be the set given by Lemma 4.6. We define the set $S'_i = \mathrm{mod}_M(S_i)$ to be the image of S_i under the function $\mathrm{mod}_M(\cdot)$. From part (1) of Lemma 4.6 we get that $Z^{(-i)}$ lands in S'_i with probability at least $1 - p^{-\Omega(1)}$. For $b = b^{(-i)} \in [M]^{k-1}$ let $Z_i(b)$ be the random variable distributed according to the conditional distribution $(Z_i|Z^{(-i)} = b)$. The random variable $Z_i(b)$ is a convex combination of distributions $W_i(a) = (Z_i|f^{(-i)}(x) = a)$ taken over all $a = a^{(-i)}$ such that $\mathrm{mod}_M(a) = b$. Since, by the definition of S'_i, these a’s are all in S_i we can use part (2) of Lemma 4.6 together with Lemma 4.3 to get that each $W_i(a)$ in the convex combination of $Z_i(b)$ is $p^{-\Omega(1)}$-close to uniform. This, of course, holds then also for $Z_i(b)$. We finish the proof by observing that $Z = (Z_1,\ldots,Z_k)$ satisfies all the conditions of Lemma 4.2 with $\epsilon = p^{-\Omega(1)}$ and so we are done since $O(k \cdot \sqrt{p^{-\Omega(1)}}) = p^{-\Omega(1)}$ when $p > d^{Ck}$ and C is sufficiently large.

4.2.1 Proof of Lemma 4.6

Let $i \in [k]$. We would like to distinguish between ”good” and ”bad” fixings of $f^{(-i)}(x)$. The ”good” fixings will be those values $a = a^{(-i)} \in \mathbb{F}^{k-1}$ for which we can bound the exponential sum $\Upsilon_\chi(a)$. Before proving the Lemma formally let us describe briefly the intuition behind the proof. Each fixing $f^{(-i)}(x) = a^{(-i)}$ defines a variety V. We would like to apply Bombieri’s Theorem to bound the exponential sum of $f_i(x)$ over this variety. In order to do so we need to make sure that V is a curve and that $f_i(x)$ is not constant on ’enough’ of the components of the curve V (where the word ’enough’ takes into account the number of points in $\mathbb{F}$ in each component). The fact that most fixings satisfy the first condition, that V is a curve, will follow from a counting argument, based on a version of Bezout’s theorem. The second condition will follow from Wooley’s Theorem (Theorem 2.1). Intuitively, Wooley’s theorem tells us that the image of f is close to having high min-entropy. Clearly, this should allow us to bound the size of those components on which $f_i(x)$ is constant (for ’most’ fixings of $f^{(-i)}(x)$).

In order to be able to define these ”good” fixings of $f^{(-i)}(x)$ we need to consider the singular points of the mapping $f(x)$, namely the zeros of its Jacobian. Let $J(x) = \det\left(\frac{\partial f}{\partial x}\right)$ be the determinant of the Jacobian of $f(x)$, which is a non zero polynomial since the source Y has full rank. Let $\mathrm{Sing} = \{x \in \mathbb{F}^k \mid J(x) = 0\}$ be the set of singular points and for each $a = a^{(-i)} \in \mathbb{F}^{k-1}$ let $\mathrm{Sing}_a = \mathrm{Sing} \cap V_a$.

Definition 4.7. *We say that $a = a^{(-i)} \in \mathbb{F}^{k-1}$ is ”good” if it satisfies the following three conditions:*

1. *$|V_a| \geq p^{5/6}$.*
2. *$|\mathit{Sing}_a| \leq p^{1/6}$.*
3. *$\hat{V}_a$ is a curve. That is, $\dim(\hat{V}_a) = 1$.*

We define the set $S_i \subset \mathbb{F}^{k-1}$ to be the set of all ”good” a’s.

The next claim shows that most a’s are ”good”. Thus proving part (1) of Lemma 4.6.

Claim 4.8. *Let S_i be as above. Then $\mathbf{Pr}[f^{(-i)} \in S_i] \geq 1 - p^{-\Omega(1)}$, where the probability is over uniformly chosen $x \in \mathbb{F}^k$.*

Proof. Let $a = a^{(-i)} \in \mathbb{F}^{k-1}$ be the random variable sampled by $a = f^{(-i)}(x)$, x uniform. For $1 \leq j \leq 3$ let E_j denote the event that a satisfies condition j in Definition 4.7. We can write

$$\begin{aligned}&\mathbf{Pr}[a \text{ is ''bad''}] \leq \\ &\mathbf{Pr}[E_1^c] + \mathbf{Pr}[E_2^c] + \mathbf{Pr}[E_1 \wedge E_2 \wedge E_3^c]. \qquad (5)\end{aligned}$$

We will bound each of these three probabilities independently by $p^{-\Omega(1)}$, which will prove the claim. The first probability can be seen to be bounded by $p^{-1/6}$ by a simple union bound on all a's with small $|V_a|$.

To bound the second probability we first observe that $|\text{Sing}| \leq \deg(J(x)) \cdot p^{k-1} \leq dk \cdot p^{k-1}$. Therefore, the number of different a's not satisfying condition (2) is at most $dk \cdot p^{k-7/6}$. From Theorem 2.1 we have that for every $a = a^{(-i)} \in \mathbb{F}^{k-1}$ the set V_a contains at most $d^k \cdot p$ non-singular points. Therefore, the size of the union of all V_a's for which condition (2) is not satisfied is bounded by

$$kd \cdot p^{k-1} + (kd \cdot p^{k-7/6})(d^k \cdot p) \leq p^{k-\Omega(1)}$$

(the first term counts all singular points and the second term counts all non singular points), where the inequality holds for $p > d^{Ck}$ for sufficiently large constant C. Therefore the second probability in Eq. 5 is also bounded by $p^{-\Omega(1)}$.

We now bound the third probability in Eq. 5. Let $A \subset \mathbb{F}^{k-1}$ be the set of a's satisfying conditions (1) and (2) but not (3) in the definition of a "good" a. We first observe that Lemma 4.5 gives us the bound $|A| \leq kd^k \cdot p^{k-2}$ on the size of A. Now, For each $a \in A$ the size of V_a is bounded by $p^{1/6} + d^k \cdot p$ (V_a does not contain many singular points since a satisfies condition (2)). Therefore, we have that

$$\begin{aligned}\sum_{a \in A} |V_a| &\leq |A| \cdot (p^{1/6} + d^k \cdot p) \\ &\leq kd^k \cdot p^{k-2} \cdot (p^{1/6} + d^k \cdot p) \\ &\leq p^{k-\Omega(1)},\end{aligned}$$

(when $p > d^{Ck}$ and C is sufficiently large). This completes the proof of the claim. □

We now move to proving part (2) of Lemma 4.6. We will show that for every $a = a^{(-i)} \in S_i$ and for every non trivial character χ the sum $|\Upsilon_\chi(a)|$ is bounded by $p^{-\Omega(1)}$.

Claim 4.9. *Let $a = a^{(-i)} \in S_i$. Then we have the bound $|\Upsilon_\chi(a)| \leq p^{-\Omega(1)}$.*

Proof. Let $\hat{V}_a = \hat{C}_1 \cup \ldots \cup \hat{C}_L$ be the decomposition of the curve $\hat{V}_a$ into irreducible components and let $C_j = \hat{C}_j \cap \mathbb{F}^k$ for $j \in [L]$. Using standard facts about the number of irreducible components (see full version for details) we can show that $L \leq d^k$ (this can be shown, for example, using Bezout's Theorem). We wish to use Theorem 4.4 to bound $|\Upsilon_\chi(a)|$. Our first step will be to show that the polynomial $f_i(x)$ can be constant only on those irreducible components $\hat{C}_j$ that have few points in $\mathbb{F}_p$. To show this, notice that if the polynomial $f_i(x)$ is constant on one of the irreducible components $\hat{C}_j$ then , using Theorem 2.1 and part (2) of the definition of "good" a's, we get that $|C_j| \leq p^{1/6} + d^k$.

We now consider the modified curve $\hat{U}_a$ constructed by taking the union of those components $\hat{C}_j$ of $\hat{V}_a$ for which $|C_j| > p^{1/6} + d^k$ and let $U_a = \hat{U}_a \cap \mathbb{F}^k$. We can now use Theorem 4.4 to get the bound

$$\left| \sum_{x \in U_a} \chi(f_i(x)) \right| \leq 4d^{2k} \cdot p^{1/2},$$

which translates into the bound

$$\left| \sum_{x \in V_a} \chi(f_i(x)) \right| \leq d^k \cdot (p^{1/6} + d^k) + 4d^{2k} p^{1/2} \leq p^{2/3}$$

(separating the sum into points in the small components and in the large components) where the inequality hold when $p > d^{Ck}$, C sufficiently large. Dividing this sum by $|V_a| > p^{5/6}$ we get the required bound of $p^{-\Omega(1)}$ on $|\Upsilon_\chi(a)|$. □

Combining the above two claims concludes the proof of Lemma 4.6. □

5 Improving the Output Length

The extractor constructed in Section 4 can extract a constant fraction of the min-entropy of the source. It was suggested to us by Salil Vadhan that we can extract almost all of the min-entropy by using special properties of the source. This indeed works, and in this section we explain how. Due to space constraints we only give a rough sketch of the argument. A complete treatment can be found in the full version.

Roughly speaking the method to extract many bits from a full rank source Y is as follows: Let $E_1 : \mathbb{F} \mapsto \{0,1\}^{m_1}$ be the extractor for distributions with small Fourier coefficients given by Lemma 4.3 (namely the $\mod 2^{m_1}$ function) and let $E_2 : \mathbb{F}^{k-1} \times \{0,1\}^s \mapsto \{0,1\}^{m_2}$ be any seeded extractor with seed length s and output length m_2. Consider the composition of these two extractors given by $E(Y) = E_2(Y^{(-k)}, E_1(Y_k))$ (recall that $Y^{(-k)} = (Y_1, \ldots, Y_{k-1})$) in which the role of the uniform seed is taken by $E_1(Y_k)$. We would like to claim that $E(Y)$ is close to uniform. The first thing to observe is that m_1 has to be larger than s. This requirement will be easy to satisfy since in our setting, when $p \geq d^{O(k)}$, the output of E_1 will be larger then the seed length of standard seeded extractors. The more important thing to justify is the fact that we can replace the uniform seed of E_2 with a seed that is correlated with the source - $Y^{(-k)}$. This can be done since for 'most' fixings of $Y^{(-k)}$, the random variable $E_1(Y_k)$ is close to uniform (this follows from Bombieri's Theorem and the analysis of Section 4).

6. Open Problems

Our paper invites further work in several directions:

- The extractors we give in this paper work when the field size is $d^{\Omega(k)}$. Extending our results to the case where the field size is polynomial in k is an interesting open problem. Building on the results of this paper it is enough to construct such an extractor for polynomial sources of full rank.
- An affine source may be viewed in two dual ways: as the image of an affine map, or as the kernel of one. Our extension here to low degree sources takes the first view. An interesting problem is extending the second view: extracting from low degree algebraic varieties. We note that The case of one dimensional varieties is already covered by Bombieri's Theorem (See Section 4).

7 Acknowledgments

We would like to thank Ran Raz and Amir Shpilka for numerous helpful conversations. We thank Jean Bourgain for bringing to our attention the results of Wooley [25] and Bombieri [5]. We thank Salil Vadhan for the idea presented in Section 5. We thank Andrej Bogdanov and Gil Alon for helpful conversations.

References

[1] B. Barak, R. Impagliazzo, and A. Wigderson. Extracting randomness using few independent sources. *FOCS '04*, pages 384–393, Washington, DC, USA, 2004.

[2] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson. Simulating independence: new constructions of condensers, ramsey graphs, dispersers, and extractors. *STOC '05*, pages 1–10, New York, NY, USA, 2005.

[3] B. Barak, A. Rao, R. Shaltiel, and A. Wigderson. 2-source dispersers for sub-polynomial entropy and ramsey graphs beating the frankl-wilson construction. *STOC '06*, pages 671–680, New York, NY, USA, 2006.

[4] M. Blum. Independent unbiased coin flips from a correlated biased source: a finite state markov chain. *Combinatorica*, 6(2):97–108, 1986.

[5] E. Bombieri. On exponential sums in finite fields. *American Journal of Mathematics*, 88:71–105, 1966.

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

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

[8] A. Gabizon and R. Raz. Deterministic extractors for affine sources over large fields. *FOCS '05*, pages 407–418, Washington, DC, USA, 2005.

[9] A. Gabizon, R. Raz, and R. Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. *FOCS '04*, pages 394–403, Washington, DC, USA, 2004.

[10] F. R. Gantmacher. *The Theory of Matrices*, volume 1. New York, NY, USA, 1959.

[11] O. Goldreich. Three XOR-lemmas - an exposition. *ECCC*, 2(056), 1995.

[12] P. Indyk. Uncertainty principles, extractors, and explicit embeddings of l2 into l1. *STOC '07*, pages 615–620, 2007.

[13] J. Kamp, A. Rao, S. Vadhan, and D. Zuckerman. Deterministic extractors for small-space sources. *STOC '06*, pages 691–700, New York, NY, USA, 2006. ACM Press.

[14] J. Kamp and D. Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. *FOCS '03*, 2003.

[15] N. Nisan and D. Zuckerman. More deterministic simulation in logspace. *STOC '93*, pages 235–244, New York, NY, USA, 1993.

[16] A. Rao. Extractors for a constant number of polynomially small min-entropy independent sources. *STOC '06*, pages 497–506, New York, NY, USA, 2006.

[17] A. Rao. An exposition of bourgain's 2-source extractor. Technical Report TR07-034, ECCC, 2007.

[18] R. Raz. Extractors with weak random seeds. *STOC '05*, pages 11–20, New York, NY, USA, 2005.

[19] R. Raz, O. Reingold, and S. Vadhan. Error reduction for extractors. *FOCS '99*, page 191, Washington, DC, USA, 1999.

[20] R. Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, 77:67–95, 2002.

[21] A. Ta-Shma and D. Zucherman. Extractor codes. *STOC '01*, pages 193–199, New York, NY, USA, 2001.

[22] L. Trevisan and S. Vadhan. Extracting randomness from samplable distributions. *FOCS '00*, page 32, Washington, DC, USA, 2000.

[23] J. von Neumann. Various techniques used in connection with random digits. *Applied Math Series*, 12:36–38, 1951.

[24] A. Wigderson and D. Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. *Combinatorica*, 19(1):125–138, 1999.

[25] T. Wooley. A note on simultaneous congruences. *J. Number Theory*, 58:288–297, 1996.