Extractors for Varieties [*]

Zeev Dvir [†]

Abstract

We study the task of randomness extraction from sources which are distributed uniformly on an unknown algebraic variety. In other words, we are interested in constructing a function (an extractor) whose output is close to uniform even if the input is drawn uniformly from the set of solutions of an unknown system of low degree polynomials. This problem generalizes the problem of extraction from affine sources which has drawn a considerable amount of interest lately.

We present two constructions of explicit extractors for varieties. The first works for varieties of any size (including one dimensional varieties, or curves) and requires field size which is exponential in the overall dimension of the space. Our second extractor allows the field size to be polynomial in the degree of the equations defining the variety, but works only for varieties whose size is at least the square root of the total size of the space.

1 Introduction

A deterministic extractor (or simply an extractor) for a class of distributions $\mathcal{C}$, over a common finite domain Ω, is a function $E : \Omega \mapsto \{0,1\}^m$ such that for every random variable X, distributed according to some distribution in $\mathcal{C}$, we have that $E(X)$ is close to the uniform distribution (in statistical distance). That is, an extractor gives us a way of producing close to uniform bits from a *single* sample taken from some unknown distribution which satisfies the condition of being in $\mathcal{C}$. It is not hard to see that, when the size of $\mathcal{C}$ is not too large, picking the function E at random will give a deterministic extractor with the best possible parameters with probability close to one. The main difficulty is, therefore, to find *explicit* constructions of extractors, where explicit means computable in polynomial time by a deterministic Turing Machine.

Requiring the unknown distribution (sometimes called *source*) to have only high entropy (or even min entropy) is known to not be enough to allow for deterministic extractors (in this case one must allow the extractor to use a short random seed). So, if we want to have a deterministic extractor we have to assume some additional structure on the distributions in $\mathcal{C}$.

A natural way to limit the structure of a distribution is to assume it has 'low complexity' in some natural measure. The pioneering work of Trevisan and Vadhan [TV00] considered the class of distributions for which there exists an efficient sampling algorithm. They showed that the problem of constructing deterministic extractors for this class is closely related to proving complexity lower

[*] An earlier manuscript of this work, containing the same results, was titled *'deterministic extractors for algebraic sources'*.

[†] Department of Computer Science, Weizmann institute of science, Rehovot, Israel. `zeev.dvir@weizmann.ac.il`. Research supported by Binational Science Foundation (BSF) grant.

bounds and proved conditional results that give deterministic extractors for samplable distributions based on lower bounds assumptions. Later, in [KRVZ06], the class of distributions sampled by small *space* machines was tackled and an (unconditional) construction of extractors for this class was given, assuming only that the entropy of the distribution is larger than the space of the machine which generates it.

Another natural way to limit the structure of the distribution is to impose some algebraic structure on it. This way is less general than assuming only an efficient sampling algorithm but it may allow us to prove *unconditional* results (these algebraic constraints cannot be described using small space machines and so the results of [KRVZ06] do not apply). The simplest algebraic family of sources is of *affine sources* or distributions sampled uniformly from an unknown affine subspace. Over small fields the result of [Bou07] gives extractors for subspaces whose dimension is linear in the total dimension of the space. Over larger fields (polynomial in the dimension) [GR05] gave an extractor which works for subspaces of *any* dimension. These results on affine sources generalize earlier works on *bit fixing* sources which are a special case of affine sources which arise in cryptographic applications [CGH$^+$85, KZ03, GRS04].

An extractor for affine sources can be viewed as a function which is hard to approximate using subspaces (similarly to the point of view taken in [TV00]). More precisely, there is no affine subspace of large dimension on which the function can be 'predicted' with some non trivial advantage (since it is close to uniform on any subspace). This point of view motivates the study of extractors for distributions defined using low degree polynomials since low degree polynomials play an important role in approximation theory and computational complexity.

The first step in this direction was taken in [DGW07] were the class of *polynomial sources* was defined. This is the class of distributions which are *sampled* by low degree polynomials (that is, by evaluating some unknown polynomial map on a uniform input). This class generalizes the class of affine sources in a natural way since we can view an affine source as the image of an affine (or degree one) mapping. The extractor given in [DGW07] works when the field size is at least $d^{O(n)}$, where d is the degree of the polynomials defining the source and n is the dimension of the space.

Another way to view an affine source is as the *kernel*, or set of zeros, of an affine mapping. Generalizing this definition to allow low degree polynomials brings up the class of sources sampled uniformly from *varieties* or sets of common zeros of one or more polynomials. We name this class of sources *algebraic sources*. Our main results (which are given in more detail below) are the following:

- We give an explicit extractor for arbitrary varieties defined using degree d equations. This extractor works when $|\mathbb{F}| > d^{O(n^2)}$.
- We give an explicit extractor which works for varieties in $\mathbb{F}^n$ whose number of points is bigger than $|\mathbb{F}|^{n/2}$. The advantage of this extractor over the first one is that it only reuires $\mathbb{F}$ to be larger than $d^{O(1)}$, where d is the bound on the degrees of the polynomials defining the variety.

To the best of our knowledge, the only previous work on extraction from varieties dealt exclusively with the special case of sources in $\mathbb{F}^2$ that are defined using a single bivariate equation of the form $y^m = f(x)$ [Gur05, CFGP06]. Extraction from these sources comes up naturally in the implementation of certain cryptographic protocols (e.g Key Exchange) that involve Elliptic Curves. In this scenario, two parties agree on a secret key which is distributed uniformly on some curve.

In order to use this key for encryption they must produce uniform (or close to uniform) bits. This brings up the task of deterministic extraction from varieties in a natural way.

1.1 Formal statement of our results

We will denote by $\mathbb{F}$ the finite field of p elements, where p is a prime number. An algebraic source is defined to be a random variable distributed uniformly over the set of common zeros of one or more polynomials in n variables defined over $\mathbb{F}$. As is always the case with extractors we will assume that the number of such points is larger than some fixed threshold. This guarantees that the entropy of the source is sufficiently high and is a prerequisite to any form of extraction (deterministic or randomized).

Dimension: Before stating our results we will need to say something about the notion of *dimension.* This notion is quite simple to define for affine sources (and even for polynomial sources) but is more tricky to describe for algebraic sources. Let $V \subset \mathbb{F}^n$ be the set of common zeros of the polynomials $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$. A set V defined in this way is called an algebraic set. We could view the polynomials $f_1, \ldots, f_s$ also as polynomials over the algebraic closure of $\mathbb{F}$, denoted $\mathbb{E}$, and to define the set $\hat{V} \subset \mathbb{E}^n$ as the set of common zeros (now taken in $\mathbb{E}$) of the same system of equations. The set $\hat{V}$ is called a variety and the reason we want to look at the algebraic closure is that over finite fields some polynomials behave in a 'strange' way (the number of roots they have can be much smaller than the 'typical' number). Varieties are the main object of study in Algebraic Geometry and share many properties with subspaces (which are of course also varieties). One such property is the existence of an integer $0 \leq k \leq n$ called the dimension of $\hat{V}$ and denoted $\dim(\hat{V})$. We will not go into the formal definition of the dimension of a variety here (formal definitions will be given in Section 2) but rather just say that it is indeed a generalization of the same notion for subspaces and has similar behavior. For example, if we take the number of polynomials s to be $n - k$ then for almost all choices of $f_1, \ldots, f_{n-k}$ the variety $\hat{V}$ will have dimension k. Also, taking the intersection of a k-dimensional variety with a hypersurface (the set of zeros of *one* polynomial) can reduce its dimension by at most one (and this will indeed be the case for 'almost all' hypersurfaces). For the purpose of this introduction we define the dimension of the algebraic set V to be the dimension of the corresponding variety $\hat{V}$ (in later sections we will adopt a more common notation – starting with a k-dimensional variety and then considering its intersection with $\mathbb{F}^n$).

We now state our first result. For convenience we limit s (the number of polynomials) to be at most n. We could allow s to grow independently of n at the cost of introducing more parameters into the theorem.

Theorem 1 (Extractor for small varieties over large fields). *There exists a family of functions $\mathbf{Ext}_{n,k,d} : \mathbb{F}^n \mapsto \{0,1\}^m$ with $m = \Omega(k \cdot \log(p))$ that satisfies the following: Let*

$$V = \{x \in \mathbb{F}^n | f_1(x) = \ldots = f_n(x) = 0\}$$

be an algebraic set of dimension k such that $\deg(f_i) \leq d$ for all $i \in [n]$. Suppose that $|V| > p^{k-1/48}$ and that $p > d^{\Omega(n^2)}$. Let X_V be a random variable distributed uniformly on V. Then $\mathbf{Ext}_{n,k,d}(X_V)$ is $p^{-\Omega(1)}$-close to uniform. Moreover, there exists a deterministic algorithm that on input (n, k, d, p) runs in time polynomial in $n, \log(d)$ and $\log(p)$ and returns a circuit computing $\mathbf{Ext}_{n,k,d}$.

A natural question is whether the bound of $p^{k-1/48}$ (the constant $1/48$ is arbitrary, any small enough constant would do) is necessary. An affine subspace of dimension k always has p^k points. What can we say about the number of points in a k-dimensional algebraic set? Take for example the algebraic set $U_a \subset \mathbb{F}^n$ defined as the set of zeros of the quadratic polynomial $x_1^2 - a$ for $a \in \mathbb{F}$. The corresponding variety $\hat{U}_a \subset \mathbb{E}^n$ has dimension $n-1$ regardless of the choice of a. However, the set U_a will be empty for roughly half of the values of a and will have size $2 \cdot p^{n-1}$ for the other half. This example shows that Theorem 1 must impose some bound on the size of V. The next natural question is whether the bound of $p^{k-1/48}$ is actually obtained for some algebraic set (the theorem would be meaningless otherwise). To see why this is the case consider a randomly chosen system of $n-k$ equations of degree d (with coefficients in $\mathbb{F}$). The algebraic set defined by these equations will have dimension k with probability close to one (assuming $\mathbb{F}$ is sufficiently large). A simple counting argument shows that the average number of solutions to such a system is exactly p^k. Therefore, with probability $1 - p^{-\Omega(1)}$, the algebraic set defined by these equations will satisfy the conditions of the theorem.[1]

Our second extractor is considerably more simple than the one given in Theorem 1 and has the additional advantage of working over fields as small as $d^{O(1)}$ (for constant d this means that the field size is also constant). The main drawback is that it works only for algebraic sets that have size larger than $p^{n/2}$.

Theorem 2 (Extractor for large varieties over small fields). *There exists a family of functions* $\mathbf{Ext2}_{n,d} : \mathbb{F}^n \mapsto \{0,1\}^m$ *with* $m = \Omega(\log(p))$ *that satisfies the following: Let*

$$V = \{x \in \mathbb{F}^n | f_1(x) = \ldots = f_n(x) = 0\}$$

be an algebraic set such that $\deg(f_i) \leq d$ *for all* $i \in [n]$*. Suppose that* $|V| > p^{n\cdot(1/2+\delta)}$ *and that* $p > d^{2/\delta}$ *for some* $\delta > 0$*. Let* X_V *be a random variable distributed uniformly on* V*. Then* $\mathbf{Ext2}_{n,d}(X_V)$ *is* $p^{-\Omega(1)}$*-close to uniform. Moreover, there exists a deterministic algorithm that on input* (n, d, p) *runs in time polynomial in* $n, \log(d)$ *and* $\log(p)$ *and returns a circuit computing* $\mathbf{Ext2}_{n,d}$*.*

The extractor of Theorem 2 is in fact given by the simple formula

$$\mathbf{Ext2}_{n,d}(x_1, \ldots, x_n) = x_1^{d+1} + \ldots + x_n^{d+1} \mathrm{mod} 2^m.$$

We believe that the fact that this simple function is an extractor for varieties is independently interesting and could have more applications.

1.2 Overview of the techniques

The proof of Theorem 1 follows the same framework used in [GR05] for extraction from affine sources. In fact, there is a part of our construction (a seeded extractor for algebraic sets) that is taken 'as is' from [GR05] (proving that it works also for varieties, however, requires new ideas). We can divide the proof of Theorem 1 into three steps:

[1] It is possible, using more advances methods, to determine exactly those k-dimensional algebraic sets for which the conditions of Theorem 1 do not hold. All of these varieties are in fact 'similar' to the degenerate example of $x_1^2 - a$ for $a \in \mathbb{F}$ which is not a quadratic residue. We refer the interested reader to [Sch76] for a discussion of this subject.

1. Constructing an extractor with short output($\sim \log(p)$) that works for algebraic sets of arbitrary dimension ≥ 1. This construction will follow from a construction of a polynomial that is not constant on any curve (one dimensional variety) of bounded degree.
2. Constructing a seeded extractor (an extractor that uses an auxiliary short random input called a 'seed') for k-dimensional algebraic sets, whose output length is $\Omega(k \cdot \log(p))$. This seeded extractor will have the additional property of being 'alomst' linear for every fixed seed.
3. Combining the above two constructions (using a theorem from [Sha06]) to give a deterministic extractor with output length $\Omega(k \cdot \log(p))$. The composition is done by first applying the extractor from step 1 (with short output) and then using the result as the seed for the seeded extractor of step 2 (applied again on the same source). The output of this composition is then shown to be close to uniform (even though the 'seed' was chosen as a function of the source).

The proof of Theorem 2 relies on an exponential sum estimate due to Deligne [Del74]. This powerful theorem of Deligne is one of the landmark results of modern algebraic geometry and generalizes Weil's theorem from curves to varieties. The trick here is to choose the 'right' polynomial (in this case $x_1^{d+1} + \cdots + x_n^{d+1}$) and then to show that the exponential sum of this polynomial is bounded over any variety with many points, that is defined only using degree d polynomials.

1.3 Organization

The paper is organized as follows:

- In Section 2 we give general preliminaries required by later sections. These include definitions and basic results related to varieties.
- In Section 3 we construct a deterministic extractor whose output length is $O(\log(p))$.
- In Section 4 we generalize the 'extractor for full rank source' of [DGW07] to the case of a polynomial mapping defined over a variety. This generalization is required in order to adapt the seeded extractor of [GR05] from affine sources to algebraic sources.
- In Section 5 we show that a slightly modified version of the seeded extractor for affine sources of [GR05] works also for algebraic sources.
- In Section 6 we combine the above two constructions to give a deterministic extractor with long output for *irreducible* varieties (see Section 2 for the definition of irreducibility).
- In Section 7, we show how to extend the results of the previous section from irreducible varieties to general varieties and so prove Theorem 1 (which is restated as Theorem 7.1).
- In Section 8 we prove Theorem 2 (which is restated as Theorem 8.2).

2 General Preliminaries

Throughout the paper $\mathbb{F}$ will denote a finite prime field of p elements and $\mathbb{E}$ will denote its algebraic closure. For a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ we denote by $\deg(f)$ the total degree of f. We call f a linear polynomial if f is homogenous and of degree 1. All logarithms are taken to the base 2 unless otherwise noted.

2.1 Randomness Extractors

The statistical distance between two distributions P and Q on a finite domain Ω is defined as

$$\max_{S \subseteq \Omega} |P(S) - Q(S)| .$$

We say that P is ϵ-close to Q if the statistical distance between P and Q is at most ϵ, otherwise we say that P and Q are ϵ-far. If P and Q are ϵ-close we write $P \overset{\epsilon}{\sim} Q$. For a set A, we denote the uniform distribution on A by $\mathrm{Uni}(A)$. We say that P is a convex-combination of the distributions $P_1, \ldots, P_m$ if there exist real numbers $q_1, \ldots, q_m \geq 0$ such that $\sum_{i \in [m]} q_i = 1$ for which

$$P(E) = \sum_{i \in [m]} q_i \cdot P_i(E),$$

for any event $E \subset \Omega$.

The min-entropy of a random variable X is defined as

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

Another useful measure of entropy is the collision probability:

$$\mathrm{cp}(X) \triangleq \sum_{x \in \mathrm{supp}(X)} \mathbf{Pr}[X = x]^2 .$$

Let $\mathcal{C}$ be a class of distributions over a finite set Ω and let $E : \Omega \mapsto \{0,1\}^m$ be some function. We call E an ϵ-deterministic-extractor (or simply an ϵ-extractor) for the class $\mathcal{C}$ if for every distribution $X \in \mathcal{C}$ the random variable $E(X)$ is ϵ-close to the uniform distribution on $\{0,1\}^m$. An ϵ-seeded extractor for $\mathcal{C}$ is a function $S : \Omega \times \{0,1\}^d \mapsto \{0,1\}^m$ such that for every $X \in \mathcal{C}$ and for every uniformly distributed random variable Y, independent of X, we have that $S(X, Y)$ is ϵ-close to uniform. The number d is called the seed-length of S. If

2.2 Affine Varieties

An affine variety (or simply variety) in $\mathbb{E}^n$ is the set of common zeros of one or more polynomials in n variables. More formally, let $S \subset \mathbb{E}[x_1, \ldots, x_n]$, we define

$$\mathbf{V}(S) = \{x \in \mathbb{E}^n | f(x) = 0 \, , \forall f \in S\},$$

to be the variety generated by the set S. It is a well known fact (see [Sha94]) that every variety is generated by a finite set of polynomials, so $\mathbf{V}(S) = \mathbf{V}(f_1, \ldots, f_r)$ for some $f_1, \ldots, f_r \in \mathbb{E}[x_1, \ldots, x_n]$. In this paper, unless otherwise stated, we will deal only with varieties that are defined using polynomial with coefficients in the finite field $\mathbb{F}$. We say that a point $x \in V$ in a variety is $\mathbb{F}$-rational (or simply rational) if all of its coordinates are in $\mathbb{F}$.

It is easy to verify that the intersection and the union of varieties is also a variety. We call a variety V irreducible if it cannot be written as the union of two varieties different than V, even if these varieties are defined using elements of $\mathbb{E}$ (these varieties are sometimes called absolutely irreducible). Another well known fact is that every variety can be decomposed in a unique way into irreducible varieties called the components of the variety. For a variety V we denote by

$$\mathbf{I}(V) = \{f \in \mathbb{E}[x_1, \ldots, x_n] \mid f(x) = 0 \ \forall x \in V\}$$

the ideal of V.

The dimension of a variety V, denoted $\dim(V)$, is defined as the largest integer k, such that there exists a chain

$$V_0 \subsetneqq V_1 \subsetneqq \ldots \subsetneqq V_k \subset V$$

such that V_i is an irreducible variety for every $0 \leq i \leq k$. One can show that, for a variety $V \subset \mathbb{E}^n$, we have $0 \leq \dim(V) \leq n$ and that this definition of dimension agrees with the definition for linear subspaces (which are also varieties). We define the co-dimension of V, denoted $\operatorname{codim}(V)$, to be $n - \dim(V)$. We say that a variety V has pure dimension if all of its irreducible components have the same dimension. A variety of pure dimension $n-1$ is called a hypersurface and a variety of pure dimension 1 is called a curve. One can verify that the only irreducible zero-dimensional varieties are points $P \in \mathbb{E}^n$. Therefore, a variety is zero-dimensional iff it is a finite subset of $\mathbb{E}^n$.

Let $V \subset \mathbb{E}^n$ be an irreducible variety of dimension k. The degree of V, denoted $\deg(V)$, is defined as the largest integer D such that there exists an affine subspace $U \subset \mathbb{E}^n$ of dimension $n - k$ such that $V \cap U$ is finite and of size D (one can show that this number is always positive and finite). For a reducible variety V we define the degree of V to be the sum of the degrees of its irreducible components.[2]

The following lemma describes the simplest kind of varieties – hypersurfaces. A proof of this lemma can be found in most texts on algebraic geometry (e.g [Sha94, Har77, Har92]).

Lemma 2.1. *A variety $H \subset \mathbb{E}^n$ is a hypersurface of degree D iff there exists a polynomial $h \in \mathbb{E}[x_1, \ldots, x_n]$ of degree D such that $H = \mathbf{V}(h)$. Furthermore, H is irreducible iff h is irreducible.*

We will often use the following special case of Bezout's theorem to upper bound the degree of a variety (this theorem is usually stated for projective varieties, but the affine version can be derived by moving to the projective closure, see [DGW07]).

Theorem 2.2 (Bezout, see [Dan94]). *Let $V \subset \mathbb{E}^n$ be a variety of pure dimension k and let $H \subset \mathbb{E}^n$ be a hypersurface that does not contain any of the irreducible components of V and such that $V \cap H$ is non empty. Then, $V \cap H$ has pure dimension $k-1$ and*

$$\deg(V \cap H) \leq \deg(V) \cdot \deg(H).$$

The following Corollary of Bezout's Theorem bounds the degrees of the irreducible components of a variety $V = \mathbf{V}(f_1, \ldots, f_s)$ defined by degree d polynomials.

Corollary 2.3. *Let $V = \mathbf{V}(f_1, \ldots, f_s)$ be such that $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$ are of degree $\leq d$. Then $deg(V) \leq d^s$.*

[2] The notion of degree is usually defined only for projective varieties. However, all the results we will use regarding the degree can be derived from their projective analogs (see the Appendix of [DGW07] for a more detailed discussion of projective vs. affine varieties).

Proof. The proof is by induction on s. The case $s = 1$ follows from Lemma 2.1. Suppose the corollary holds for the variety $U = \mathbf{V}(f_1, \ldots, f_{s-1})$. The Corollary now follows by observing that, from Theorem 2.2, the sum of degrees of all components can be multiplied by at most d when we intersect them with $\mathbf{V}(f_s)$. □

2.3 Bombieri's Theorem

We will use the following exponential sum estimate due to Bombieri [Bom66]. We quote here a weak version of Bombieri's Theorem which is sufficient for our needs (a derivation of this form of the theorem appears in the appendix of [DGW07]).

Theorem 2.4 (Theorem 6 in [Bom66]). *Let $\hat{C} \subset \mathbb{E}^n$ be a curve of degree $\leq D$ and 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 $\hat{C}$. Let $\hat{C} = \hat{C}_1 \cup \ldots \cup \hat{C}_L$ be the decomposition of $\hat{C}$ into irreducible components. Let $\hat{U}$ be the union of those irreducible components of $\hat{C}$ on which $g(x)$ is non constant and let $U = \hat{U} \cap \mathbb{F}$. Let $\chi : \mathbb{F} \to \mathbb{C}^*$ be a non-trivial additive character of $\mathbb{F}$. Then, if $p > d \cdot D$, we have the bound*

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

Combining Bombieri's Theorem with the next lemma allows us to extract random bits from distributions of polynomials whose input is chosen uniformly from a curve and such that the polynomial is non constant on many of the irreducible components of the curve. This lemma is an extension of the now folklore Vazirani XOR Lemma [Gol95] and is used [Bou07, BRSW06, DGW07] to extract randomness from distributions that satisfy certain exponential sum estimates.

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 [Rao07].

Lemma 2.5. *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 $\mathbb{F}$. Suppose that for every non-trivial additive character $\chi : \mathbb{F} \to \mathbb{C}^*$ 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.*

2.4 The Closure Theorem

We will need the following theorem regarding the projection of an affine variety on a subset of the coordinates. We say that a variety $V \subset \mathbb{E}^n$ is defined over a subfield $\mathbb{E}' \subset \mathbb{E}$ if V can be defined as the common zeros of polynomials with coefficients in the subfield $\mathbb{E}'$.

Theorem 2.6 (Closure Theorem). *Let $V \subset \mathbb{E}^n$ be an (irreducible) variety of dimension k and degree D that is defined over a subfield $\mathbb{E}'$. Let $\pi_m : \mathbb{E}^n \mapsto \mathbb{E}^m$ be the projection map onto the first $m \leq n$ coordinates. Let I_m be the intersection of $\mathbf{I}(V)$ and $\mathbb{E}[x_1, \ldots, x_m]$ and let $V_m = \mathbf{V}(I_m)$. Then*

1. *V_m is (irreducible) of dimension $k' \leq k$ and degree $\leq D$ and is defined over $\mathbb{E}'$.*

2. *V_m is the intersection of all varieties containing $\pi_m(V)$ (that is, V_m is the* closure *of $\pi_m(V)$, hence the name of the theorem).*

3. *There exists a variety $W \subsetneq V_m$ such that $V_m - W \subset \pi_m(V)$.*

Proof. Parts 2. and 3. of the theorem are given in [CLO92], Chapter 3, §2, Theorem 3. The fact that V_m is irreducible (if V is) follows from the fact that the ideal I_m is prime (this follows from the fact that $\mathbf{I}(V)$ is prime). The bound on the dimension of V_m is trivial (for example using the equivalent definition given in [CLO92], Chapter 9, §5, Proposition 5, of the dimension as the maximal k for which there is a projection on a k-dim subspace that is dense). The bound on the degree of V_m can be derived from the same argument for projective varieties (see [Dan94], Chapter 3, §2, 2.4). Showing that $\mathbf{V}_m$ is defined over $\mathbb{E}'$ can be done, for example, using Gröbner Bases (see [CLO92]). □

The following simple lemma can be derived from the Closure Theorem.

Lemma 2.7. *Let $V \subset \mathbb{E}^n$ be a variety of dimension $k < n$ and degree $\leq D$. Then V is contained in a hypersurface H of degree at most D. Moreover, if V is defined over $\mathbb{F}$ then H is defined over $\mathbb{F}$. That is, H is defined by a degree $\leq D$ polynomial with coefficients in $\mathbb{F}$.*

Proof. We prove the lemma by induction on n. If $n = 1, 2$ then the lemma is trivial. Suppose $n > 2$. If $k = n - 1$ then we take $V = H$, therefore suppose $k < n - 1$. Let $\pi : \mathbb{E}^n \mapsto \mathbb{E}^{n-1}$ be the projection onto the first $n - 1$ coordinates. Using Theorem 2.6, there exists a variety $U \subset \mathbb{E}^{n-1}$ of dimension at most $k < n - 1$ and degree at most D that contains $\pi(V)$. From the inductive hypothesis there exists a hypersurface $H' = \mathbf{V}(h(x_1, \ldots, x_{n-1}))$ of degree at most D containing U. Let $H \subset \mathbb{E}^n$ be defined as $H = \mathbf{V}(h)$, treating h as a polynomial in n variables. Then H is also a hypersurface of degree at most D and H contains V. The 'Moreover' part of the lemma follows by observing that if V is defined over $\mathbb{F}$ then Theorem 2.6 guarantees that U will also be defined over $\mathbb{F}$. Therefore, when we apply the inductive hypothesis, we get that H' (and so also H) is also defined over $\mathbb{F}$. □

Suppose $V \subset \mathbb{E}^n$ is a variety and $S \subset [n]$ is a set of size larger than $\dim(V)$. Intuitively, the restriction of V to the coordinates in S cannot 'fill' the entire $|S|$-dimensional space spanned by those coordinates. Therefore, we expect that there will be some non-zero polynomial $g(x)$ that depends *only* on the variables $\{x_i | i \in S\}$ and vanishes identically on V. The next lemma proves this fact and gives a bound on the degree of g.

Lemma 2.8. *Let $V \subset \mathbb{E}^n$ be a variety of dimension $k < n$ and degree $\leq D$ that is defined over $\mathbb{F}$. Let $S \subset [n]$ be such that $|S| > k$. Then there exist a polynomial $g \in \mathbb{F}[\mathbf{x}_S] \cap \mathbf{I}(V)$ of degree $\leq D$.*

Proof. We project V onto the coordinates of S. Using Theorem 2.6 the image is contained in a variety U of dimension at most $k < |S|$ and degree at most D that is defined over $\mathbb{F}$. Using Lemma 2.7 there exists a hypersurface $H \subset \mathbb{E}^{|S|}$, defined over $\mathbb{F}$, of degree at most D that contains U. This means that there is a polynomial $g \in \mathbb{F}[\mathbf{x}_S] \cap \mathbf{I}(V)$ of degree $\leq D$ that vanished on U and so also on V. □

2.5 The number of solutions to a system of polynomial equations

Bounding the number of rational solutions to a system of polynomial equations in n variables is one of the fundamental questions of Algebraic Geometry and of Number Theory. In this section we describe two variants of such bounds: one bound on the number of $\mathbb{F}$-rational points in a k-dimensional variety and one on the number of non-singular solutions to a system of n polynomials in n variables.

We will want to give an upper bound on the number of $\mathbb{F}$-rational points of a variety V in term of the dimension of V and the degree of V. Since we are only interested in an upper bound we will be able to derive our results from the following elementary result of Schwartz and Zippel [Sch80, Zip79].

Theorem 2.9 (Schwartz-Zippel). *Let $f \in \mathbb{E}[x_1, \ldots, x_n]$ be a non zero polynomial with* $\deg(f) \leq d$. *Then, for any finite subset $S \subset \mathbb{E}$ we have*

$$|\mathbf{V}(f) \cap S^n| \leq d \cdot p^{n-1}.$$

We now extend this bound for varieties of arbitrary dimension, when the field $\mathbb{F}$ is sufficiently large (this is a very crude bound that will suffice for our purposes. Much stronger bounds can be obtained using more sophisticated methods).

Theorem 2.10. *Let $V \subset \mathbb{E}^n$ be an irreducible variety of dimension $k < n$ and degree $\leq D$. Suppose $p > 2^n D$. Then*

$$|V \cap \mathbb{F}^n| \leq 2^n D \cdot p^k.$$

Proof. Let $V_p \triangleq V \cap \mathbb{F}^n$. We will prove the lemma by induction on n. If $n = 1$ then the lemma is trivial since a 0-dimensional variety has size $\leq$ its degree. If $n = 2$ then the proof is also easy using Theorem 2.9. Suppose that $n > 2$. If $k = n - 1$ then there exists a polynomial $h \in \mathbb{E}[x_1, \ldots, x_n]$ of degree $\leq D$ such that $V = \mathbf{V}(h)$. We can thus use Theorem 2.9 to bound the size of V_p by $D \cdot p^{n-1} \leq 2^n D \cdot p^k$. Suppose next that $k < n - 1$. We will use the next claim to map V into a space of dimension $< n$ in a way that will not decrease the size of V_p by much.

Claim 2.11. *There exists a linear mapping $\ell : \mathbb{F}^n \mapsto \mathbb{F}^{n-1}$, such that*

$$|\ell(V_p)| > \frac{1}{2} \cdot \min\{|V_p|, p^{n-1}\}.$$

Proof. Let $\mathcal{L}$ denote the set of all linear mappings from $\mathbb{F}^n$ to $\mathbb{F}^{n-1}$ and let L be a random variable uniformly distributed over $\mathcal{L}$. Let X denote a random variable uniform on V_p. Let us observe the average collision probability of the distribution $\ell(X)$ when we average over all $\ell \in \mathcal{L}$.

$$\begin{aligned}
\frac{1}{|\mathcal{L}|} \sum_{\ell \in \mathcal{L}} \mathrm{cp}(\ell(X)) &= \sum_{\ell \in \mathcal{L}} \Pr[L = \ell] \cdot \Pr_{v_1, v_2 \leftarrow X}[L(v_1) = L(v_2) \,|\, L = \ell] \\
&= \Pr_{v_1, v_2 \leftarrow X}[L(v_1) = L(v_2)] \\
&\leq \Pr_{v_1, v_2 \leftarrow X}[v_1 = v_2] + \Pr_{v_1, v_2 \leftarrow X}[L(v_1) = L(v_2) \,|\, v_1 \neq v_2] \\
&\leq |V_p|^{-1} \quad + \quad p^{-n+1} \\
&\leq 2 \cdot \max\{|V_p|^{-1}, p^{-n+1}\}.
\end{aligned}$$

Therefore, there exists ℓ such that $\mathrm{cp}(\ell(X)) \leq 2 \cdot \max\{|V_p|^{-1}, p^{-n+1}\}$. Now, since the collision probability is minimized for the uniform distribution, we get that the support of the random variable $\ell(X)$ must have size at least $\frac{1}{2} \cdot \min\{|V_p|, p^{n-1}\}$, as was required. □

Let $\ell : \mathbb{F}^n \mapsto \mathbb{F}^k$ be as in Claim 2.11. We can extend ℓ naturally to a linear mapping from $\mathbb{E}^n$ to $\mathbb{E}^{n-1}$. Using the Closure Theorem (Theorem 2.6) we get that there exists an irreducible variety $U \subset \mathbb{E}^{n-1}$ of dimension $k' \leq k < n-1$ and degree $\leq D$ such that $\ell(V) \subset U$. Let $U_p \triangleq U \cap \mathbb{F}^n$. Using the inductive hypothesis we have

$$|U_p| \leq 2^{n-1} D \cdot p^k.$$

Observe also that, since ℓ is defined using constants from $\mathbb{F}$ we have that

$$\ell(V_p) \subset U_p.$$

We separate our analysis into two cases. The first case is when $|V_p| > p^{n-1}$. In this case we have $|\ell(V_p)| \geq p^{n-1}/2$ and so we get that

$$|U_p| \geq |\ell(V_p)| \geq p^{n-1}/2.$$

combining this with the upper bound on $|U_p|$ we get that

$$p^{n-1}/2 \leq 2^{n-1} D \cdot p^k,$$

which is a contradiction since $k < n-1$ and $p > 2^n D$. The second case is when $|V_p| \leq p^{n-1}$. Then we have $|\ell(V_p)| \geq |V_p|/2$ which translates into the bound

$$|V_p| \leq 2|\ell(V_p)| \leq 2|U_p| \leq 2^n D \cdot p^k,$$

as was required. □

We now describe the second result we will need concerning the number of solutions to a system of polynomial equations. Before we can do so we need the following definition: The Jacobian $J(x)$ of a polynomial mapping $f(x) : \mathbb{F}^n \mapsto \mathbb{F}^n$, given by n polynomials $f_1, \ldots, f_n \in \mathbb{F}[x_1, \ldots, x_n]$, is defined as the $n \times n$ matrix whose (i,j)'th element is the polynomial $\frac{\partial f_i}{\partial x_j}(x)$, where the partial derivatives are defined in the same way as for polynomials over the field of real numbers (using the same syntactic rules of derivation).

Theorem 2.12 (Wooley [Woo96]). *Let k and d be integers. Let $f(x) : \mathbb{F}^n \mapsto \mathbb{F}^n$ be a polynomial mapping given by n polynomials $f_1, \ldots, f_n \in \mathbb{F}[x_1, \ldots, x_n]$ od degree $\leq d$. Let $J(x)$ be its Jacobian. For $a \in \mathbb{F}^n$ let*

$$N_a \triangleq |\{c \in \mathbb{F}^n \quad : \quad f(c) = a \quad \text{and} \quad \det(J(c)) \neq 0\}|.$$

Then for every $a \in \mathbb{F}^n$ we have $N_a \leq d^n$.

3 An extractor with short output

In this section we will construct a deterministic extractor for varieties, whose output is of length $\Omega(\log(p))$ bits. In later sections we will use this extractor to construct an extractor with longer output length. Our construction is based on the following theorem that gives a polynomial $R_{n,D}$ that is non-constant on any curve in $\mathbb{E}^n$ whose degree is at most D.

Theorem 3.1. *For every integers n and D there exists a polynomial $R_{n,D} \in \mathbb{F}[x_1, \ldots, x_n]$ of degree $\leq D^{O(n)}$ such that for every irreducible curve $C \subset \mathbb{E}^n$ of degree at most D, $R_{n,D}$ is not constant on C. Moreover, there exists a deterministic algorithm that, on input (n, D), runs in time poly$(n, \log(D))$ and outputs an arithmetic circuit computing $R_{n,D}(x)$.*

We defer the proof of Theorem 3.1 to Section 3.1 and continue now with the theorem that uses $R_{n,D}$ to get an extractor with short output.

Theorem 3.2. *There exist constants $C, c > 0$ such that the following holds: Suppose k, n, D, m are such that $k \leq n$, $p > D^{Cn}$ and $m \leq c \cdot \log(p)$. Let $V \subset \mathbb{E}^n$ be a variety of pure dimension k and degree $\leq D$. Suppose $|V \cap \mathbb{F}^n| \geq p^{k-1/24}$ and let X_V denote a random variable uniformly distributed over $V \cap \mathbb{F}^n$. Then the random variable*

$$\text{mod}\ _{2^m}(R_{n,D}(X_V))$$

is $p^{-\Omega(1)}$-close to uniform.

Proof of Theorem 3.2. If we could partition V into disjoint curves (of degree at most D) then, using Bombieri's Theorem, we could bound the exponential sum of $R_{n,D}$ over V by writing it as an average over the exponential sums on each individual curve. The following lemma shows how this could be achieved.

Lemma 3.3. *Let $V \subset \mathbb{E}^n$ be a variety of pure dimension $k \geq 1$ and degree at most $D < p$. Then there exist a set A of size p^{k-1} and a family of curves $\{U_a\}_{a \in A}$ of degree at most D such that $V \cap \mathbb{F}^n$ is the disjoint union of the sets $U_a \cap \mathbb{F}^n$, $a \in A$.*

Proof. The proof is by induction on k. If $k = 1$ then there is nothing to prove since $|A| = 1$ and V is already a curve. Suppose $k > 1$ and let $V_1, \ldots, V_t$ be the irreducible components of V. We know that $t \leq D < p$ since $\deg(V) \leq D$. Each component V_i is of dimension $k > 1$ and so is infinite. Let L denote a random variable uniformly distributed on the set of linear polynomials in $\mathbb{F}[x_1, \ldots, x_n]$. The probability that L is constant on a component V_i is at most $1/p$ (consider the value of L on two distinct points in V_i). Therefore, by a union bound, the probability that L is fixed on any of the irreducible components of V is at most $t/p < 1$. Therefore, there exists a linear polynomial $\ell \in \mathbb{F}[x_1, \ldots, x_n]$ such that ℓ is not constant on any of the V_i's. Using Theorem 2.2 (Bezout), we have that for each $a \in \mathbb{F}$ the variety

$$V_a \triangleq V \cap \mathbf{V}(\ell(x) - a)$$

has pure dimension $k - 1$ and degree $\leq D$. The p varieties $\{V_a\}_{a \in \mathbb{F}}$ are disjoint and we also have

$$V \cap \mathbb{F}^n = \bigcup_{a \in \mathbb{F}} (V_a \cap \mathbb{F}^n),$$

since $\ell(x) \in \mathbb{F}$ for all $x \in V \cap \mathbb{F}^n$.

From the inductive hypothesis we know that for each V_a there exists a family of p^{k-2} curves $\{V_{a,b}\}_{b \in B_a}$ such that $V_a \cap \mathbb{F}^n$ is the disjoint union of the sets $V_{a,b} \cap \mathbb{F}^n$. Therefore, we get the required disjoint partition of $V \cap \mathbb{F}^n$ into the p^{k-1} sets $V_{a,b} \cap \mathbb{F}^n$ where $a \in \mathbb{F}$ and $b \in B_a$. □

We continue with the proof of Theorem 3.2. Let

$$V' \triangleq V \cap \mathbb{F}^n.$$

By the lemma we just proved, there exists a family of p^{k-1} curves $\{U_a\}_{a \in A}$ of degree at most D such that V' is the disjoint union:

$$V' = \bigcup_{a \in A} (U_a \cap \mathbb{F}^n).$$

Let $\chi : \mathbb{F} \mapsto \mathbb{C}^*$ be a non-trivial additive character. Let us denote by

$$R(x) \triangleq R_{n,D}(x).$$

We can now use Bombieri's Theorem (Theorem 2.4), together with the fact that $R(x)$ is of degree $D^{O(n)}$ and not constant on any of the irreducible components of U_a (for every $a \in A$) to get the following exponential sum estimate.

$$\begin{aligned} \left| \frac{1}{|V'|} \sum_{x \in V'} \chi(R(x)) \right| &\leq \frac{1}{|V'|} \cdot \sum_{a \in A} \left| \sum_{x \in U_a \cap \mathbb{F}^n} \chi(R(x)) \right| \\ &\leq \frac{p^{k-1}}{p^{k-1/24}} \cdot 4 \cdot D^{O(n)} \cdot D^2 \cdot p^{1/2} \\ &\leq p^{-\Omega(1)}, \end{aligned}$$

for $p > D^{Cn}$ and C sufficiently large constant. Applying Lemma 2.5 we get that reducing $R(x)$ module 2^m for $m \leq c \cdot \log(p)$ gives a distribution which is $p^{-\Omega(1)}$-close to uniform. This concludes the proof of Theorem 3.2. □

3.1 proof of Theorem 3.1

3.1.1 Preliminaries

The next lemma deals with two irreducible curves and states that their intersection can be infinite iff they are identical.

Lemma 3.4. *Let $C_1, C_2 \subset \mathbb{E}^n$ be two irreducible curves. If $C_1 \cap C_2$ is infinite then $C_1 = C_2$.*

Proof. If $U = C_1 \cap C_2$ is infinite then it is a one dimensional variety. Now, if $U \neq C_1$ we could construct a chain of proper inclusions that would show that $\dim(C_1) > 1$. □

In the proof of Theorem 3.1 we will need to use a bivariate polynomial $F(x, y)$ such that all 'shifts' of this polynomial are irreducible. The next lemma describes such a polynomial.

Lemma 3.5. *Let $m \geq 0$ be an integer. let $a \in \mathbb{E}$ and let*

$$F(x, y) = x^m + y + a.$$

Then $F(x, y)$ is irreducible.

Proof. Suppose in contradiction that $F(x, y) = g(x, y) \cdot h(x, y)$ with $\deg(g), \deg(h) > 0$. Then wither g or h must be a polynomial only in x (otherwise we would have powers of y larger than one in the product). W.l.o.g suppose $g(x, y) = g(x)$. We thus have that the coefficient of y in F is divisible by $g(x)$ and so $g(x)$ is constant, a contradiction. □

The following lemma is derived from the Closure Theorem (Theorem 2.6) and describes the image of a curve under a polynomial mapping.

Lemma 3.6. *Let $C \subset \mathbb{E}^n$ be an irreducible curve of degree D. Let $g_1, g_2 \in \mathbb{E}[x_1, \ldots, x_n]$ and let $d = \max\{\deg(g_1), \deg(g_2)\}$. Let*

$$U = \{(g_1(x), g_2(x)) \,|\, x \in C\}$$

and suppose that U is infinite. Then, there exists an irreducible curve $\tilde{C}$ and a finite set $W \subset \tilde{C}$ such that $U = \tilde{C} - W$ and $\deg(\tilde{C}) \leq d \cdot D$.

In order to prove the lemma we will need the following claim.

Claim 3.7. *Let $h \in \mathbb{E}[x, y]$ be of degree d and let $W \subset \mathbb{E}^2$ be a finite set. Then, there exists a line $L \subset \mathbb{E}^2$, passing through $(0, 0)$, that intersects $\mathbf{V}(h) - W$ in d distinct points.*

Proof. For every $a, b \in \mathbb{E}$ consider the line

$$L_{a,b} \triangleq \{(at, bt) | t \in \mathbb{E}\}$$

and let $h_{a,b}(t) = h(at, bt)$. The intersection of $L_{a,b}$ with $\mathbf{V}(h)$ is given by the solutions of $h_{a,b}(t) = 0$. We would thus like to show that there exists a pair (a, b) such that $h_{a,b}(t)$ has d distinct zeros $t_1, \ldots, t_d$ such that $\{(at_i, bt_i) | i \in [d]\} \cap W = \emptyset$. This follows by showing that the set of pairs $(a, b) \in \mathbb{E}^2$ for which this condition does not hold can be described using a finite number of (non trivial) polynomial equations in a and b, and is therefore a variety of dimension < 2. Therefore, there exists a pair, outside this variety that gives the required intersection. We leave the details to the reader. □

Proof of Lemma 3.6. Let $f_1, \ldots, f_r \in \mathbb{E}[x_1, \ldots, x_n]$ be such that $C = \mathbf{V}(f_1, \ldots, f_r)$. Consider the set

$$V = \{(x_1, \ldots, x_n, g_1(x), g_n(x)) \,|\, x \in C\} \subset \mathbb{E}^{n+1}.$$

Let $h_1, h_2 \in \mathbb{E}[x_1, \ldots, x_n, y_1, y_2]$ be given by (we write $(\mathbf{x}, \mathbf{y})$ to denote $(x_1, \ldots, x_n, y_1, y_2)$)

$$h_1(\mathbf{x}, \mathbf{y}) = y_1 - g_1(\mathbf{x}) \, , \, h_2(\mathbf{x}, \mathbf{y}) = y_2 - g_2(\mathbf{x}).$$

Then

$$V = \mathbf{V}(f_1, \ldots, f_n, h_1, h_2)$$

and is therefore a variety. The dimension of V is at least one since V is infinite and can be seen to be at most one by considering the fact that V is the image of a curve under a polynomial mapping. The same reasoning shows also that V must be irreducible.

Let

$$I = I(V) \triangleq \{f \in \mathbb{E}[\mathbf{x}, \mathbf{y}] \,|\, f(\mathbf{x}, \mathbf{y}) = 0,\, \forall (\mathbf{x}, \mathbf{y}) \in V\},$$

be the ideal corresponding to the variety V and let

$$\tilde{I} = I \cap \mathbb{E}[y_1, y_2].$$

Observe that $\tilde{I}$ is also an ideal and that it is in fact a prime ideal (this is because I is prime). Let

$$\tilde{V} = \mathbf{V}(\tilde{I}) \subset \mathbb{E}^2$$

be the (irreducible) variety corresponding to $\tilde{I}$ (since $\tilde{I}$ is prime, it is also radical, and so, by Hilbert Nullstelensatz, it is the ideal of $\tilde{V}$).

It is easy to see that $U \subset \tilde{V}$. The Closure Theorem (Theorem 2.6) asserts that there exists a variety $W \subsetneqq \tilde{V}$ such that $\tilde{V} - W \subset U$. To prove the lemma we need to show the following four things

1. $\tilde{V}$ is an irreducible curve.

2. W is finite.

3. $U = \tilde{V} - W$.

4. $\deg(\tilde{V}) \le d \cdot D$.

(1) We know that $\tilde{V}$ is infinite (since $U \subset \tilde{V}$) and so $\dim(\tilde{V}) > 0$. Assume towards a contradiction that $\tilde{V}$ is not a curve. Then $\dim(\tilde{V}) = 2$ and so $\tilde{V} = \mathbb{E}^2$ and $\tilde{I} = \{0\}$. We can further assume that $\dim(W) < 2$ for otherwise we would have $W = \tilde{V}$. Therefore, we can find a non zero polynomial $h(y_1, y_2)$ such that $\mathbf{V}(h) \cap W$ is finite. Now, since $\mathbf{V}(h)$ is infinite and since $\mathbb{E}^2 - W \subset U$ we have that $V(h) \cap U$ must be infinite. Let

$$A \triangleq V \cap \{(x_1, \ldots, x_n, y_1, y_2) \,|\, h(y_1, y_2) = 0\}.$$

Then, by the previous arguments, we have that A is infinite and so has dimension at least one. Since $A \subset V$ and V is an irreducible curve, we have $A = V$ (see Lemma 3.4) and so h has to vanish identically on V. This means that $h \in I$ and so, since $h \in \mathbb{E}[y_1, y_2]$ we also have $h \in \tilde{I}$, contradicting the assumption that $\tilde{I} = \{0\}$.

(2) If W is infinite then W has dimension one, and since it is contained in the irreducible curve $\tilde{V}$ we have $W = \tilde{V}$, a contradiction.

(3) This follows immediately from (2) and from the fact that $\tilde{V} - W \subset U$ (we can replace W with a smaller finite subset if neccessary).

(4) Let

$$\tilde{D} = \deg(\tilde{V}).$$

Lemma 2.1 implies that there exists an irreducible polynomial $\tilde{h}(y_1, y_2)$ of degree $\tilde{D}$ such that

$$\tilde{V} = \mathbf{V}(\tilde{h})$$

(we can use the lemma on $\tilde{V}$ since, in the plane, a curve is also a hypersurface). Now, from Claim 3.7, we get that there exists a line through the origin

$$L = \{(y_1, y_2) \mid \alpha \cdot y_1 + \beta \cdot y_2 = 0\}$$

that intersects $\tilde{V} - W$ (which is equal to U by the previous items) in $\tilde{D}$ distinct points. Let

$$s(x) = \alpha \cdot g_1(x) + \beta \cdot g_2(x) \in \mathbb{E}[x_1, \ldots, x_n].$$

We thus have that the hypersurface

$$H_s = \{x \in \mathbb{E}^n \mid s(x) = 0\}$$

intersects the curve C in at least $\tilde{D}$ different points. Recall that C is irreducible and so has pure dimension one. We also know that H_s does not contain C (if it did than the line L would contain U) and that $\deg(H_s) \leq d$ (Lemma 2.1). We can thus apply Theorem 2.2 (Bezout) and get that

$$\deg(C \cap H_s) \leq d \cdot D.$$

Item (4) now follows since the degree of $C \cap H_s$ is equal to its size, which is $\tilde{D}$. □

3.1.2 Proof of Theorem 3.1

Let n, D be integers. We will describe an algorithm to compute the polynomial $R_{n,d}$ and then prove that it is not constant on any curve of degree D. We may assume w.l.o.g that n is a power of two (otherwise we can add $O(n)$ zero coordinates to the space). Our input variables are $x_1, \ldots, x_n$. We treat them as elements of $\mathbb{E}[x_1, \ldots, x_n]$ and define the set

$$\mathcal{P}_0 = \{x_1, \ldots, x_n\} \subset \mathbb{E}[x_1, \ldots, x_n].$$

We will inductively define sets of polynomials $\mathcal{P}_1, \mathcal{P}_2, \ldots$ with the property that $|\mathcal{P}_k| = |\mathcal{P}_{k-1}|/2$. Since n is a power of two we will have $\mathcal{P}_{\log(n)} = \{R(x)\}$ for some polynomial R. This R will be our required $R_{n,D}$. A building block in our construction will be the following bivariate polynomial

$$T_M(y, z) = y^M + z.$$

In order to define these sets we first define a sequence of integers $D_0, D_1, \ldots$ as follows:

$$D_0 = D \quad , \quad D_k = D_{k-1} \cdot (D_{k-1} + 1).$$

Suppose that

$$\mathcal{P}_{k-1} = \{g_1(x), g_2(x), \ldots, g_{2\ell-1}(x), g_{2\ell}(x)\}$$

($|\mathcal{P}_{k-1}|$ is even since n is a power of two). We define

$$\mathcal{P}_k = \left\{T_{D_{k-1}+1}(g_1(x), g_2(x)),\, T_{D_{k-1}+1}(g_3(x), g_4(x)), \ldots, T_{D_{k-1}+1}(g_{2\ell-1}(x), g_{2\ell}(x))\right\}.$$

Let $\deg(\mathcal{P}_k)$ denote the maximal degree of a polynomial in $\mathcal{P}_k$. We thus have

$$\deg(\mathcal{P}_0) = 1 \quad , \quad \deg(\mathcal{P}_k) = \deg(\mathcal{P}_{k-1}) \cdot (D_{k-1} + 1).$$

Observing the definition of the numbers D_k we see that

$$\deg(R_{n,D}) \leq D^{O(n)},$$

as required. We also have that an arithmetic circuit for $R_{n,D}$ can be generated in time polynomial in n and in $\log(D)$, since raising a variable to a power M can be done by a circuit of size $O(\log(M))$ using repeated squaring. It is also clear that the coefficients of $R_{n,d}$ are in fact in the prime field $\mathbb{F}$ and so $R_{n,D} \in \mathbb{F}[x_1, \ldots, x_n]$ as was stated in the theorem.

We now turn to the analysis of the construction. Let

$$C = \mathbf{V}(f_1, \ldots, f_r)$$

be a curve in $\mathbb{E}^n$ of degree $\leq D$. We will call a polynomial $g \in \mathbb{E}[x_1, \ldots, x_n]$ active on C (or just active, since C is fixed) if the set $\{g(x) \,|\, x \in C\}$ is infinite. We will show by induction on $k = 1, 2, \ldots, \log(n)$ that each set $\mathcal{P}_k$ contains at least one active polynomial. This will prove the theorem since it will show that $R_{n,D}$ is active (and therefore non constant). Clearly we have that there exists $i \in [n]$ such that x_i is active (otherwise C has only a finite number of points) and so the base case for the induction is proved. Suppose now that the set $\mathcal{P}_{k-1} = \{g_1, \ldots, g_{2\ell}\}$ contains an active polynomial. W.l.o.g we assume it is $g_1(x)$ (the proof will be identical if it is $g_2(x)$, even though $T_M(y,z)$ is not symmetrical).

We will show that the first element of $\mathcal{P}_k$

$$G(x) \triangleq T_{D_{k-1}+1}(g_1(x), g_2(x))$$

is active. Let

$$U \triangleq \{(g_1(x), g_2(x)) \,|\, x \in C\} \subset \mathbb{E}^2.$$

We know that U is infinite (since $g_1(x)$ is active) and so, by Lemma 3.6, we have that there exists an irreducible curve $\tilde{C} \subset \mathbb{E}^2$ with

$$\deg(\tilde{C}) \leq \deg(\mathcal{P}_{k-1}) \cdot D$$

and a finite subset $W \subset \tilde{C}$ such that $U = \tilde{C} - W$. Using the recursive formulas for D_k and for $\deg(\mathcal{P}_k)$ one can show that

$$\deg(\mathcal{P}_{k-1}) \cdot D \leq D_{k-1}$$

and so we have

$$\deg(\tilde{C}) \leq D_{k-1}. \tag{1}$$

Suppose, in contradiction, that G is constant on C. That is, that there exists $a \in \mathbb{E}$ such that $G(x) = a$ for all $x \in C$. This means that $T_{D_{k-1}+1}(y,z) = a$ for all $(y,z) \in U$. Consider the curve

$$A \triangleq \{(y,z) \,|\, T_{D_{k-1}}(y,z) = a\} \subset \mathbb{E}^2.$$

Using Lemma 3.5 we have that A is irreducible of degree $D_{k-1} + 1$. But, since A contains U, and since $U \cap \tilde{C}$ is infinite, we have that $A \cap \tilde{C}$ is also infinite. Now, using Lemma 3.4, we have that $\tilde{C} = A$ which is impossible since Eq. 1 tells us that they have different degrees. □

4 A generalized 'extractor for full rank sources'

One of the main results of [DGW07] is a theorem which shows how to extract many random bits from the output of a full rank polynomial mapping $f : \mathbb{F}^k \mapsto \mathbb{F}^k$ (that is, a polynomial mapping whose Jacobian has full rank) when the input is chosen uniformly at random from $\mathbb{F}^k$. In fact, [DGW07] shows that it is enough to apply the $\mathrm{mod}_M(\cdot)$ function on each of the k coordinates of $f(x)$, for suitably chosen M, to get an output which is close to uniform.

We will need to generalize this theorem to the case where the input is chosen at random from the $\mathbb{F}$-rational points of a k-dimensional variety $V \subset \mathbb{E}^n$. Naturally, we will need to impose some 'niceness' conditions on the variety V and in particular on the size of its intersection with the set of singular points of the mapping f. It will be easier to state our results using the following definition.

Definition 4.1. [(δ, D)**-bounded**] *Let R_1, R_2, R_3 be finite sets such that $R_2 \subset R_1$ and let $f : R_1 \mapsto R_3$ be some function. We say that f is (δ, D)-bounded on R_2 if there exists a set $R_2' \subset R_2$ such that $|R_2'| \leq \delta \cdot |R_2|$ and such that for all $c \in R_3$ we have*

$$|\{b \in R_2 - R_2' \mid f(b) = c\}| \leq D.$$

That is, each element $c \in R_3$ has at most D pre-images in R_2 that lie outside of some set R_2' that has density at most δ in R_2.

Observing Theorem 2.12 (Wooley), we see for example that any polynomial mapping $f : \mathbb{F}^n \mapsto \mathbb{F}^n$, given by n degree d polynomials, is $(dn/p, d^n)$-bounded on $\mathbb{F}^n$, provided that its Jacobian has a non-zero determinant. Therefore, if we could construct an extractor for (δ, D)-bounded polynomial mappings over a variety, we will indeed generalize the result of [DGW07].

For our purposes it is enough to state and prove the theorem for a *linear* mapping defined over a variety and for a fixed error parameter of $\delta = p^{-11/12}$. This will slightly simplify the proof and save us some notations. One can easily modify the proof to hold also for polynomial mappings of higher degree and for arbitrary δ (within reasonable bounds). We extend the definition of the $\mathrm{mod}_M(\cdot)$ function for vectors $a = (a_1, \ldots, a_k) \in \mathbb{F}^k$ as follows:

$$\mathrm{mod}_M(a) = (\mathrm{mod}_M(a_1), \ldots, \mathrm{mod}_M(a_k)).$$

Theorem 4.2. *There exist constants C, c such that the following holds: Let D_1, D_2, k, n, m be integers such that $k \leq n$, $p > (n \cdot D_1 \cdot D_2)^C$ and $m \leq c \cdot \log(p)$. Let $\hat{V} \subset \mathbb{E}^n$ be a variety of pure dimension k and degree $\leq D_2$ and suppose that the set $V = \hat{V} \cap \mathbb{F}^n$ satisfies $|V| \geq p^{k-1/24}$. Let $L : \mathbb{F}^n \mapsto \mathbb{F}^k$ be a linear mapping which is $(p^{-11/12}, D_1)$-bounded on V. Let X_V denote a random variable uniformly distributed over V. Then, $mod_{2^m}(L(X_V))$ is $p^{-\Omega(1)}$-close to the uniform distribution on $\mathbb{F}^k$.*

Proof: The proof of the theorem is very similar to the proof of Theorem 5.1 in [DGW07] with small modifications. We will use the notation of [DGW07] that for a vector $v = (v_1, \ldots, v_n)$ and for an index $i \in [n]$ we have

$$v^{(-i)} \triangleq (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n).$$

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, $u = (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n) \in A^{n-1}$.

We denote by $L_1, \ldots, L_k : \mathbb{F}^n \mapsto \mathbb{F}$ the linear functions corresponding to the k coordinates of L. For $i \in [k]$ and $a = a^{(-i)} \in \mathbb{F}^{k-1}$, we let

$$\hat{C}_a \triangleq \mathbf{V}(L^{(-i)} - a) \cap \hat{V} = \{x \in \hat{V} | L^{(-i)}(x) = a\}$$

and let

$$C_a \triangleq \hat{C}_a \cap \mathbb{F}^n.$$

For $a = a^{(-i)} \in \mathbb{F}^{k-1}$ such that $C_a \neq \emptyset$ and for a non trivial additive character $\chi : \mathbb{F} \to \mathbb{C}^*$ we define the exponential sum

$$\mathcal{S}_i(a, \chi) \triangleq \frac{1}{|C_a|} \sum_{x \in C_a} \chi(L_i(x)).$$

As in [DGW07], the theorem will follow from the following lemma and Lemma 2.5 (we will not repeat here the derivation of the theorem from the lemma, since it is identical to the one appearing in [DGW07]).

Lemma 4.3. *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. *$L^{(-i)}(x)$ lands in S_i with probability at least $1 - p^{-\alpha}$, when x is chosen uniformly in V.*
2. *For every $a = a^{(-i)} \in S_i$ and for every non trivial χ, $|\mathcal{S}_i(a, \chi)| \leq p^{-\alpha}$.*

Proof of Lemma 4.3. Let $i \in [k]$. We would like to distinguish between "good" and "bad" fixings of $L^{(-i)}(x)$. The "good" fixings will be those values $a = a^{(-i)} \in \mathbb{F}^{k-1}$ for which we can bound the exponential sum $\mathcal{S}_i(a, \chi)$.

Let $V' \subset V$ be such that for all $a \in \mathbb{F}^k$

$$|\{x \in V - V' \,|\, L(x) = a\}| \leq D_1$$

and such that

$$|V'| \leq p^{-11/12} \cdot |V|$$

(such V' exists since L is $(p^{-11/12}, D_1)$-bounded on V). We will refer to the elements of V' as the *singular* points of L (this will make the connection to the proof in [DGW07] easier to see). We also define, for every $a = a^{(-i)} \in \mathbb{F}^{k-1}$, the set

$$V'_a \triangleq C_a \cap V',$$

to be the set of singular points that map to a under $L^{(-i)}$.

Definition 4.4. *We say that $a = a^{(-i)} \in \mathbb{F}^{k-1}$ is "good" if it satisfies the following three conditions:*

1. *$|C_a| \geq p^{5/6}$ (C_a has many points).*
2. *$|V'_a| \leq p^{1/6}$ (C_a doesn't have many singular points).*
3. *$\hat{C}_a$ is a curve. That is $\hat{C}_a$ has pure dimension one.*

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

We will use the next lemma to show that most a's are "good". Thus proving part (1) of Lemma 4.3.

Claim 4.5. *Let S_i be as above. Then*

$$\mathbf{Pr}[L^{(-i)}(x) \in S_i] \geq 1 - p^{-\Omega(1)},$$

where the probability is over uniformly chosen $x \in V$.

Proof. Let $a = a^{(-i)} \in \mathbb{F}^{k-1}$ be the random variable sampled by $a = L^{(-i)}(x)$, x uniform in V. For $1 \leq j \leq 3$ let E_j denote the event that a satisfies condition j in Definition 4.4. We can write

$$\mathbf{Pr}[a \text{ is "bad"}] \leq \mathbf{Pr}[E_1^c] + \mathbf{Pr}[E_2^c] + \mathbf{Pr}[E_1 \wedge E_2 \wedge E_3^c]. \tag{2}$$

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 $p^{-1/8}$ by a simple union bound on all a's with $|C_a| \leq p^{5/6}$ and using the fact that $|V| \geq p^{k-1/24}$.

To bound the second probability we first observe that, since $|V'| \leq p^{-11/12} \cdot |V|$, the number of different a's not satisfying condition (2) is at most $|V| \cdot p^{-11/12-1/6} = |V| \cdot p^{-13/12}$. Now, for every $a = a^{(-i)} \in \mathbb{F}^{k-1}$ the set C_a contains at most $D_1 \cdot p$ points outside of V'. This follows by counting the number of solutions to the k equations

$$L^{(-i)}(x) = a \ , L_i(x) = b$$

and summing over all $b \in \mathbb{F}$. Each b can give at most D_1 non singular solutions and so the total number of non-singular points in C_a is at most $D_1 \cdot p$. Therefore, the size of the union of all C_a's for which condition (2) is not satisfied is bounded by

$$|V'| + |V| \cdot p^{-13/12} \cdot (D_1 \cdot p) \leq p^{-\Omega(1)} \cdot |V|$$

(the first term counts all singular points and the second term counts all non singular points), where the inequality holds for $p > D_1^C$ for sufficiently large constant C. Therefore the second probability in Eq. 2 is also bounded by $p^{-\Omega(1)}$.

We now bound the third probability in Eq. 2. We start by proving the following lemma.

Lemma 4.6. *Let $W \subset \mathbb{E}^n$ be a variety of pure dimension k and degree $\leq D$. Let $\ell_1, \ldots, \ell_{k-1} \in \mathbb{F}[x_1, \ldots, x_n]$ be linear polynomials. For each $\xi = (\xi_1, \ldots, \xi_{k-1})$ let*

$$W_\xi \triangleq W \cap \mathbf{V}(\ell_1(x) - \xi_1, \ldots, \ell_{k-1}(x) - \xi_{k-1}).$$

and let

$$\Pi_W \triangleq \{\xi \in \mathbb{F}^{k-1} \,|\, W_\xi \neq \emptyset \text{ and } \dim(W_\xi) \neq 1\}.$$

Then $|\Pi_W| \leq k \cdot D \cdot p^{k-2}$.

Proof. In order to bound $|\Pi_W|$ we will describe an injective mapping from Π_W to some small set. Fix some $\xi = (\xi_1, \ldots, \xi_{k-1}) \in \Pi_W$. For $i \in [k-1]$ let

$$H_i \triangleq \mathbf{V}(\ell_i(x) - \xi_i)$$

and let

$$U_i \triangleq W \cap H_1 \cap \ldots \cap H_i$$

so that $U_0 = W$ and $U_{k-1} = W_\xi$. If W_ξ is not empty and $\dim(W_\xi) \neq 1$ then, using Theorem 2.2, there must be some $1 \leq i \leq k-1$ such that H_i contains one of the irreducible components of U_{i-1}. Let i' be the smallest i satisfying this condition and let $0 < \gamma \leq D$ be the index of the corresponding irreducible component of $U_{i'-1}$ (using some arbitrary ordering of the components), where the bound of D on γ follows from Theorem 2.2 and the fact that the number of irreducible components of a variety is bounded by its degree. Observe that if we are given the set

$$\{\xi_1, \ldots, \xi_{i'-1}, \xi_{i'+1}, \ldots, \xi_{k-1}, i', \gamma\}$$

we can determine $\xi_{i'}$ and so recover ξ. Therefore, there exists an injective mapping from Π_W into the set $\mathbb{F}^{k-2} \times [k] \times [D]$ and so we get the required bound. □

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, using Lemma 4.6, we have the bound

$$|A| \leq k \cdot D_2 \cdot p^{k-2}.$$

Now, For each $a \in A$ the size of C_a is bounded by $p^{1/6} + D_1 \cdot p$ (C_a does not contain many singular points since a satisfies condition (2)). Therefore, we have that

$$\begin{aligned}\sum_{a \in A} |C_a| &\leq |A| \cdot (p^{1/6} + D_1 \cdot p) \\ &\leq (k \cdot D_2 \cdot p^{k-2}) \cdot (p^{1/6} + D_1 \cdot p) \\ &\leq p^{k-\Omega(1)}\end{aligned}$$

(when $p > (n \cdot D_1 \cdot D_2)^C$ and C is sufficiently large). This completes the proof of Claim 4.5. □

We now move to proving part (2) of Lemma 4.3.

Claim 4.7. *Let $a = a^{(-i)} \in S_i$. Then we have the bound $|\mathcal{S}_i(a, \chi)| \leq p^{-\Omega(1)}$.*

Proof. Let $\hat{C}_a = \hat{C}^1 \cup \ldots \cup \hat{C}^t$ be the decomposition of the curve $\hat{C}_a$ into t irreducible components and let $C^j = \hat{C}^j \cap \mathbb{F}^n$ for $j \in [t]$. From Bezout's Theorem (Theorem 2.2), and since the L_i's are linear, we have that

$$\deg(\hat{C}_a) \leq \deg(\hat{V}) = D_2$$

and so we have also $t \leq D_2$ (since the number of components is at most the degree of the curve). We wish to use Theorem 2.4 to bound $|\mathcal{S}_i(a, \chi)|$. Our first step will be to show that the polynomial $L_i(x)$ can be constant only on those irreducible components $\hat{C}^j$ that have few points in $\mathbb{F}^n$. To show this, notice that if the polynomial $L_i(x)$ is constant, say $L_i(x) = b$, on one of the irreducible components $\hat{C}^j$ then the system of equations

$$L^{(-i)}(x) = a \ , L_i(x) = b$$

has at least $|C^j|$ solutions. Therefore, using part (2) of the definition of "good" a's, we get that

$$|C^j| \leq p^{1/6} + D_1$$

(since we know that there are at most $p^{1/6}$ singular solutions in $|C^j|$).

We now consider the modified curve $\hat{B}_a$ constructed by taking the union of those components $\hat{C}^j$ of $\hat{C}_a$ for which $|C^j| > p^{1/6} + D_1$ and let $B_a = \hat{B}_a \cap \mathbb{F}^n$. Notice that

$$\deg(\hat{B}_a) \leq \deg(\hat{C}_a) \leq D_2.$$

We are now in a position to apply Theorem 2.4 (Bombieri's Theorem) to get the bound

$$\left| \sum_{x \in B_a} \chi(L_i(x)) \right| \leq 4 \cdot D_2^2 \cdot p^{1/2},$$

which translates into the bound

$$\left| \sum_{x \in C_a} \chi(L_i(x)) \right| \leq D_2 \cdot (p^{1/6} + D_1) + 4 \cdot D_2^2 \cdot p^{1/2} \leq p^{2/3}$$

(separating the sum into points in the small components and in the large components) where the inequality holds when $p > (D_1 \cdot D_2)^C$, and C sufficiently large. Dividing this sum by $|C_a| > p^{5/6}$ we get the required bound of $p^{-\Omega(1)}$ on $|S_i(a, \chi)|$. □

Combining the above two claims concludes the proof of Lemma 4.3 and of Theorem 4.2 □

5 A seeded extractor for varieties

The goal of this section is to describe a construction of a *seeded* extractor for irreducible varieties. This will be achieved by constructing a relatively small family of linear mappings from $\mathbb{F}^n$ to $\mathbb{F}^k$ and then showing that for *any* irreducible k-dimensional variety $V \subset \mathbb{E}^n$, with 'enough' rational points, the output of *most* mappings in this family on inputs randomly chosen from $V \cap \mathbb{F}^n$ is close to uniform, when reduced modulo some integer of size $\sim p^c$ for some constant $c > 0$. The family of linear mappings we will construct is the same one used by Gabizon and Raz [GR05] to extract randomness from *linear* varieties (that is, from subspaces of $\mathbb{F}^n$) and is given in the following definition.

Definition 5.1. *Let $u \in \mathbb{E}$. For $k \leq n$ we define the linear mapping*

$$\mathcal{E}_u^{(k)} : \mathbb{E}^n \mapsto \mathbb{E}^k$$

to be

$$\mathcal{E}_u^{(k)}(x) \triangleq \begin{pmatrix} u^{1\cdot 1} & \dots & u^{1\cdot n} \\ \vdots & \ddots & \vdots \\ u^{k\cdot 1} & \dots & u^{k\cdot n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix}.$$

In [GR05] the following theorem was proven.

Theorem 5.2 ([GR05]). *Let $M \subset \mathbb{E}^n$ be an affine subspace of dimension k. For $\mathcal{E}_u^{(k)} : \mathbb{E}^n \mapsto \mathbb{E}^k$, given by Definition 5.1, let*

$$\Omega_M \triangleq \left\{ u \in \mathbb{E} \mid \mathcal{E}_u^{(k)}(M) \neq \mathbb{E}^k \right\}.$$

Then $|\Omega_M| \leq n^3$. Furthermore, the set Ω_M depends only on the linear part of M. That is, $\Omega_M = \Omega_{M+b}$ for every vector $b \in \mathbb{E}^n$.

Another way to phrase this result is given by the following corollary, which we will use later on.

Corollary 5.3. *Let A be an $(n-k) \times n$ matrix with elements in $\mathbb{E}$. Let $B(u)$ denote the $k \times n$ matrix whose (i,j)'th element is $u^{i \cdot j}$ (this is the coefficient matrix of $\mathcal{E}_u^{(k)}$) and let $C(u)$ denote the $n \times n$ matrix obtained by putting A on top of $B(u)$ (that is, taking the first $n-k$ rows from A and the last k rows from $B(u)$). Suppose A has rank $n-k$. Then*

$$|\{u \in \mathbb{E} \mid \det(C(u)) = 0\}| \leq n^3$$

Proof. If $\det(C(u)) = 0$ then the image of the linear mapping $\Phi_u : \mathbb{E}^n \mapsto \mathbb{E}^n$, given by multiplication by $C(u)$, is contained in a subspace of dimension $< n$ and in particular does not fill the entire space. This means that there exists a vector $a = (a_1, \ldots, a_{n-k}) \in \mathbb{E}^{n-k}$ such that the last k coordinates of the mapping Φ_u do not fill the entire k-dimensional space, when we restrict the first $n-k$ coordinates to a. This, in turn, implies that $u \in \Omega_M$ (as defined in Theorem 5.2), where M is the kernel of A (which is k dimensional). Using Theorem 5.2 we get that the number of such u's is at most n^3. $\square$

The generalization of Theorem 5.2 to varieties will follow from the following theorem, which we prove in Section 5.1, and Theorem 4.2. To simplify things we will prove the theorem only for an irreducible component of a given variety. As we shall see in Section 7, this will be enough in order to deduce results also for reducible varieties.

Theorem 5.4. *There exists constants $C, c > 0$ such that the following holds: Suppose k, n, d, m, s are such that $s, k \leq n$, $p > d^{Cn^2}$ and $m \leq c \cdot \log(p)$. Let $\hat{V} = \mathbf{V}(f_1, \ldots, f_s)$ be a variety with $\dim(V) = k$ and such that $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$ are of degree $\leq d$. Let V be an irreducible k-dimensional component of $\hat{V}$ such that $|V \cap \mathbb{F}^n| \geq p^{k-1/24}$. Let $B(V) \subset \mathbb{F}$ be defined as follows:*

$$B(V) \triangleq \left\{ u \in \mathbb{F} \;\middle|\; \mathcal{E}_u^{(k)} \textit{ is not } (p^{-11/12}, d^{n^2})\textit{-bounded on } |V \cap \mathbb{F}^n| \right\}.$$

Then $|B(V)| \leq n^3$.

Combining Theorem 5.4 with Theorem 4.2 we get the following Corollary:

Corollary 5.5. *There exists constants $C, c, \alpha > 0$ such that the following holds: Suppose k, n, d, m, s are such that $s, k \leq n$, $p > d^{Cn^2}$ and $m \leq c \cdot \log(p)$. Let $\hat{V}, V$ and $B(V)$ be as in Theorem 5.4 and let X_V denote a random variable uniformly distributed over $V \cap \mathbb{F}^n$. Then, for all $u \notin B(V)$ we have that*

$$\bmod_{2^m} \left(\mathcal{E}_u^{(k)}(X_V) \right) \tag{3}$$

is $p^{-\alpha}$-close to uniform. In particular, there are at most n^3 values of $u \in \mathbb{F}$ for which the random variable in (3) is $p^{-\alpha}$-far from uniform.

Proof. Notice that, using Corollary 2.3, we have $\deg(V) \leq d^n$. Now, if $u \notin B(V)$ (as defined in Theorem 5.4) we can apply Theorem 4.2 to get that

$$\operatorname{mod}_{2^m}\left(\mathcal{E}_u^{(k)}(X_V)\right)$$

is $p^{-\alpha}$-close to uniform for some constant α (notice that α does not depend on u). We should also check that p satisfies the bound required in Theorem 4.2, which is $p > (n \cdot D_1 \cdot D_2)^C$. This holds in our case since we have $D_1 = d^{n^2}$, $D_2 = d^n$ and $p > d^{Cn^2}$ (for large enough C). □

5.1 Proof of Theorem 5.4

We start with a high level description of the proof. Assume for a moment that $s = n - k$ and that $\hat{V} = V$. A possible strategy for proving the theorem would be to consider, for each $u \in \mathbb{F}$, the polynomial mapping $\phi_u : \mathbb{F}^n \mapsto \mathbb{F}^n$ given by

$$\phi_u(x) = (f_1(x), \ldots, f_{n-k}(x), L_1(x), \ldots, L_k(x)),$$

where $L_1, \ldots, L_k$ are the linear functions corresponding to the coordinates of $\mathcal{E}_u^{(k)}$. Let $J_u(x)$ denote the Jacobian (the $n \times n$ partial derivative matrix) of the mapping ϕ_u. If we could show that the set

$$\{x \in \mathbb{F}^n | \det(J_u(x)) = 0\}$$

has a small intersection with $V \cap \mathbb{F}^n$ for almost all values of $u \in \mathbb{F}$, then it would follow from Theorem 2.12 (Wooley) that $\mathcal{E}_u^{(k)}$ is $(p^{-11/12}, d^n)$-bounded on $V \cap \mathbb{F}^n$ for almost all values of u. This strategy, however, cannot work as described as the following example demonstrates: Suppose there exists a polynomial f_i that is a square of another polynomial. It is easy to verify that, for such a polynomial, all partial derivatives belong to the ideal of $\mathbf{V}(f_i)$, and so are identically zero for every $x \in V$. Therefore, the matrix $J_u(x)$ will have rank at most $n-1$ for *every* x in V (one could also come up with less obvious examples involving more than one polynomial).

In order to overcome this difficulty we will show that there exist $n-k$ polynomials $g_1, \ldots, g_{n-k}$, whose variety $V(g) = \mathbf{V}(g_1, \ldots, g_{n-k})$ contains V, and such that their partial derivative matrix has full rank for almost all $x \in V \cap \mathbb{F}^n$. This will enable us to show that for almost all values of $u \in \mathbb{F}$, the Jacobian $J_u(x)$ (now taken with the derivatives of the g_i's in the first $n-k$ rows) has a determinant which is non zero almost everywhere in $V \cap \mathbb{F}^n$ (this will follow from Corollary 5.3 above). We will than use Theorem 2.12 (Wooley) to show that $\mathcal{E}_u^{(k)}$ is bounded on $V \cap \mathbb{F}^n$ (using the fact that $V \subset V(g)$). The reason why we get d^{n^2} pre-images instead of d^n is that the degrees of the polynomials $g_1, \ldots, g_{n-k}$ can grow to be as large as d^n.

To show the existence of polynomials $g_1, \ldots, g_{n-k}$ as above we will use the following auxiliary lemma, which we prove in Section 5.1.1.

Lemma 5.6. *Let $\hat{V} = \mathbf{V}(f_1, \ldots, f_s)$ be a variety with $\dim(V) = k$ and such that $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$ are of degree $\leq d$. Let V be an irreducible k-dimensional component of $\hat{V}$. Then, there exist polynomials $g_1, \ldots, g_{n-k} \in \mathbb{F}[x_1, \ldots, x_n] \cap \mathbf{I}(V)$ of degree at most d^n such that the matrix*

$$D_g(x) \triangleq \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \cdots & \frac{\partial g_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{n-k}}{\partial x_1}(x) & \cdots & \frac{\partial g_{n-k}}{\partial x_n}(x) \end{pmatrix}$$

contains an $(n-k)\times(n-k)$ *minor* $M(x)$ *such that* $\det(M(x)) \notin \mathbf{I}(V)$.

Proof of Theorem 5.4. Let $g_1,\dots,g_{n-k} \in \mathbb{F}[x_1,\dots,x_n]$ be given by Lemma 5.6. For each $u \in \mathbb{F}$ let

$$\phi_u : \mathbb{F}^n \mapsto \mathbb{F}^n$$

be the polynomial mapping given by

$$\phi_u(x) = (g_1(x),\dots,g_{n-k}(x), L_1(x),\dots,L_k(x)),$$

where $L_1,\dots,L_k$ are the linear functions corresponding to the k coordinates of $\mathcal{E}_u^{(k)}$. The $n\times n$ matrix

$$J_u(x) \triangleq \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \dots & \frac{\partial g_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{n-k}}{\partial x_1}(x) & \dots & \frac{\partial g_{n-k}}{\partial x_n}(x) \\ u^{1\cdot 1} & \dots & u^{1\cdot n} \\ \vdots & \ddots & \vdots \\ u^{k\cdot 1} & \dots & u^{k\cdot n} \end{pmatrix}$$

is therefore the Jacobian of $\phi_u(x)$. We denote by

$$r(x,u) = \det(J_u(x))$$

the determinant of $J_u(x)$ and treat $r(x,u)$ as a polynomial in the $n+1$ variables $x_1,\dots,x_n,u$. Notice that the degree of r in the variable u is at most n^3 and the total degree of r in the variables $x_1,\dots,x_n$ is at most $n\cdot d^n$.

The next claim shows that in order for an element $u_0 \in \mathbb{F}$ to belong to $B(V)$ (the set of “bad” u’s) the restriction of the polynomial $r(x,u)$ to $u=u_0$ must belong to $\mathbf{I}(V)$. Later on we will show that this can happen only to at most n^3 values of u (thus proving the theorem).

Claim 5.7. *Let* $u_0 \in \mathbb{F}$. *If* $r(x,u_0) \notin \mathbf{I}(V)$ *then* $\mathcal{E}_{u_0}^{(k)}$ *is* $(p^{-11/12}, d^{n^2})$*-bounded on* $V\cap\mathbb{F}^n$.

Proof. If $r(x,u_0) \notin \mathbf{I}(V)$ then there exists a point $x_0 \in V$ such that $r(x_0,u_0) \neq 0$. Applying Theorem 2.2 (Bezout) we have that the variety

$$U_0 = V \cap \mathbf{V}(r(x,u_0))$$

has pure dimension $k-1$ and degree at most $\deg(V)\cdot\deg(r)$ which is at most $n\cdot d^{2n}$ (we can bound $\deg(V)$ by d^n using Corollary 2.3). Therefore, applying Theorem 2.10 (on each of the components of U_0), we get that

$$|U_0\cap\mathbb{F}^n| \leq 2^n\cdot(n\cdot d^{2n})^2\cdot p^{k-1} \leq p^{k-23/24} \leq |V\cap\mathbb{F}^n|\cdot p^{-11/12}$$

where the inequality holds for $p > d^{Cn^2}$, and C is sufficiently large (notice that p is also large enough to satisfy the bound required by Theorem 2.10).

Applying Theorem 2.12 (Wooley), we get that for every $a \in \mathbb{F}^k$, the number of $\mathbb{F}$-rational solutions to the system of n equations

$$g_1(x) = 0,\dots,g_{n-k}(x) = 0,\ \mathcal{E}_{u_0}^{(k)}(x) = a$$

that lie outside of U_0, is at most d^{n^2}. This means that the number of points in

$$(\mathbf{V}(g_1,\dots,g_{n-k})\cap\mathbb{F}^n)-U_0$$

for which $\mathcal{E}^{(k)}_{u_0}(x)=a$ is at most d^{n^2}. Using the fact that $V\subset\mathbf{V}(g_1,\dots,g_{n-k})$ we get that

$$\left|\left\{x\in(V\cap\mathbb{F}^n)-U_0 \;\middle|\; \mathcal{E}^{(k)}_{u_0}(x)=a\right\}\right|\le d^{n^2},$$

which, combined with the bound on $|U_0\cap\mathbb{F}^n|$, proves the claim. □

We proceed with the proof of Theorem 5.4. Claim 5.7 tells us that

$$|B(V)|\le|\{u_0 \mid r(x,u_0)\in\mathbf{I}(V)\}.$$

We will now show that the size of the set on the right-hand-side is bounded by n^3. Write $r(x,u)$ as a polynomial in u with coefficients in $\mathbb{F}[x_1,\dots,x_n]$. As was observed before, the degree of this polynomial is at most n^3 and so we can write

$$r(x,u)=\sum_{j=0}^{n^3}Q_j(x)\cdot u^j.$$

Let

$$\Gamma\triangleq\left\{(z_0,\dots,z_{n^3})\in\mathbb{E}^{n^3+1} \;\middle|\; \sum_{j=0}^{n^3}Q_j(x)\cdot z_j\in\mathbf{I}(V)\right\}.$$

It is easy to verify that Γ is a subspace of $\mathbb{E}^{n^3+1}$ (it is closed under linear operations and contains the zero vector). The following claim shows that Γ has co-dimension ≥ 1.

Claim 5.8. *Let Γ be as above. Then* $\dim(\Gamma)<n^3+1$

Proof. We will prove the claim by showing that there exists $x_0\in V$ and $u_0\in\mathbb{E}$ such that $r(x_0,u_0)\neq 0$. This will be enough since we will then have that $r(x,u_0)\notin\mathbf{I}(V)$ and so

$$\left(1,u_0,u_0^2,\dots,u_0^{n^3}\right)\notin\Gamma.$$

Lemma 5.6 guarantees that there exists a point $x_0\in V$ such that the first $n-k$ rows of $J_u(x_0)$ are linearly independent (this follows from the fact that there is an $(n-k)\times(n-k)$ minor not in $\mathbf{I}(V)$). Corollary 5.3 now tells us that there there can be at most n^3 elements $u\in\mathbb{E}$ for which the determinant of $J_u(x_0)$ is zero. Hence, there exists $u_0\in\mathbb{E}$ such that $r(x_0,u_0)=\det(J_{u_0}(x_0))\neq 0$. □

Using Claim 5.8 we have that there exists some non zero vector $z'=(z'_0,\dots,z'_{n^3})\in\mathbb{E}^{n^3+1}$ which is orthogonal to all vectors in the subspace Γ. Therefore, if $r(x,u_0)\in\mathbf{I}(V)$ then u_0 must satisfy the equation

$$\sum_{j=0}^{n^3}z'_j\cdot u_0^j=0$$

and this equation can have at most n^3 solutions. This concludes the proof of Theorem 5.4. □

5.1.1 Proof of Lemma 5.6

We will construct the required set of polynomials $G = \{g_1, \ldots, g_{n-k}\} \subset \mathbf{I}(V)$ iteratively – starting with an empty set G_0 and, at each step, adding to the set

$$G_\ell = \{g_1, \ldots, g_\ell\}$$

a single polynomial $g_{\ell+1} \in \mathbf{I}(V)$ until we reach $G_{n-k} = G$. To aid us in the construction of the sets $G_0, G_1, \ldots$ we will also construct a sequence of $n - k$ *distinct* integers

$$s_1, \ldots, s_{n-k} \in [n]$$

(this sequence will also be constructed iteratively). We will make sure that at at each iteration the polynomials in the set G_ℓ, in addition to being in $\mathbf{I}(V)$, satisfy the following two conditions:

$$\forall j \in [\ell] \ , \ \frac{\partial g_j}{\partial x_{s_j}} \notin \mathbf{I}(V). \tag{4}$$

$$\forall j \in [\ell] \ , \forall j < i \leq \ell \ , \ \frac{\partial g_i}{\partial x_{s_j}} = 0. \tag{5}$$

Before proceeding with the construction of the sets G_ℓ, notice that if $G = G_{n-k}$ satisfies the two conditions above, then the $(n-k) \times (n-k)$ minor $M(x)$ of the matrix $D_g(x)$, that comprises of the columns indexed by $s_1, \ldots, s_{n-k}$, is upper diagonal (after permuting the columns) with determinant

$$\det(M(x)) = \prod_{j \in [n-k]} \frac{\partial g_j}{\partial x_{s_j}}$$

and this determinant is not in $\mathbf{I}(V)$ since it is a product of polynomials not in $\mathbf{I}(V)$ and since $\mathbf{I}(V)$ is a prime ideal (the ideal of an irreducible variety is always prime).

One extra tool that we will need for the construction is the following simple claim.

Claim 5.9. *Let $I \subsetneqq \mathbb{E}[x_1, \ldots, x_n]$ be an ideal. Let $h \in \mathbb{F}[x_1, \ldots, x_n]$ be such that $h \in I$ and $0 < \deg(h) < p$. Then, there exists a polynomial $g \in \mathbb{F}[x_1, \ldots, x_n]$ and an integer $j \in [n]$ such that $\deg(g) \leq \deg(h)$, $g \in I$ and $\frac{\partial g}{\partial x_j} \notin I$. Furthermore, if a variable x_i does not appear in h then it also does not appear in g.*

Proof. As long as all partial derivatives of h are in I and $\deg(h) > 0$ we can replace h with one of its non-zero partial derivatives (if all the derivatives are zero then $\deg(h) = 0 \mod p$). This process will end if we find a partial derivative that is not in I or if all partial derivatives are zero. In the later case we get that I contains a constant (degree zero) polynomial and so $I = \mathbb{E}[x_1, \ldots, x_n]$, in contradiction. Otherwise we find a polynomial g (obtained from h by successively taking partial derivatives) such that one of its partial derivatives is not in I. The way we derived g guarantees that g will have a degree bounded by the degree of h and will satisfy the 'Furthermore' part of the claim. □

We proceed with the construction of the sets G_ℓ. Let $h_1 \in \mathbf{I}(V)$ be some non zero polynomial. Notice that $\mathbf{I}(V) \neq \mathbb{E}[x_1, \ldots, x_n]$ since otherwise we would have $V = \emptyset$. We can assume $\deg(h_1) \leq d$ since we can always take $h_1 = f_i$ for some non zero f_i (and use the fact that $\mathbf{I}(\hat{V}) \subset \mathbf{I}(V)$). Using

Claim 5.9 we have that there exists a polynomial $g_1(x) \in \mathbf{I}(V)$ of degree $\leq d$ and an index s_1 such that

$$\frac{\partial g_1}{\partial x_{s_1}} \notin \mathbf{I}(V).$$

Therefore, the set $G_1 = \{g_1\}$ satisfies the two conditions (4) and (5) above (condition (5) is trivial in this first iteration).

Next, let $1 \leq \ell < n - k$ and suppose that we have a set $G_\ell = \{g_1, \ldots, g_\ell\}$ and a list of ℓ *distinct* integers $s_1, \ldots, s_\ell \in [n]$, such that G_ℓ satisfies conditions (4) and (5). We define the set

$$S = [n] - \{s_1, \ldots, s_\ell\}.$$

Since $|S| > k = \dim(\hat{V})$ we can apply Lemma 2.8 on the variety $\hat{V}$ to get that there exists a polynomial

$$h_{\ell+1} \in \mathbb{F}[\mathbf{x}_S] \cap \mathbf{I}(\hat{V}),$$

that is, a polynomial in $\mathbf{I}(\hat{V})$ that depends only on variables from S, with degree $\leq d^n$. Since $\mathbf{I}(\hat{V}) \subset \mathbf{I}(V)$ we also have that $h_{\ell+1} \in \mathbf{I}(V)$. We now apply Claim 5.9 on $h_{\ell+1}$ to get a polynomial $g_{\ell+1} \in \mathbb{F}[\mathbf{x}_S] \cap \mathbf{I}(V)$ and an integer $s_{\ell+1} \in S$ such that

$$\frac{\partial g_{\ell+1}}{\partial x_{s_{\ell+1}}} \notin \mathbf{I}(V).$$

This shows that $G_{\ell+1} = G_\ell \cup \{g_{\ell+1}\}$ satisfies condition (4). The fact that $G_{\ell+1}$ satisfies also condition (5) follows from the fact that the variables $x_{s_1}, \ldots, x_{s_\ell}$ do not appear in $g_{\ell+1}$ and so its derivative w.r.t each one of them is zero. This concludes the proof of Lemma 5.6. □

6 A deterministic extractor for irreducible varieties

In this section we derandomize the construction of the seeded extractor of Section 5 to give a deterministic extractor with long output. This will be done via a composition argument that was developed in [GRS04, GR05, Sha06]. This argument shows that, under certain conditions, one can replace the random seed, used by a seeded extractor, with a function computed from the source, without changing the output distribution by much.

We start by describing the construction of the 'derandomized' extractor. In the following we assume that the set $\{0,1\}^m$, where $m \leq \log(p)$, is embedded into $\mathbb{F}$ using some arbitrary fixed mapping (for example, as the first 2^m integers mod p).

Definition 6.1. *Let k, n, m, D be integers such that $k \leq n$ and $m \leq \log(p)$. We define the function*

$$S_{m,D} : \mathbb{F}^n \times \mathbb{F} \mapsto \{0,1\}^{m(k-1)}$$

as follows:

$$S_{m,D}(x, u) \triangleq mod_{2^m}\left(\mathcal{E}_u^{(k-1)}(x)\right),$$

where $\mathcal{E}_u^{(k-1)} : \mathbb{F}^n \mapsto \mathbb{F}^{k-1}$ is given by Definition 5.1. Let $R_{n,D} \in \mathbb{F}[x_1, \ldots, x_n]$ be the polynomial given by Theorem 3.1 and define the function

$$E_{m,D} : \mathbb{F}^n \mapsto \{0,1\}^{m(k-1)}$$

as follows:

$$E_{m,D}(x) \triangleq S_{m,D}\left(x, \mathit{mod}_{2^m}(R_{n,D}(x))\right).$$

The main result of this section is the following theorem.

Theorem 6.2. *There exist constants $C, \tilde{c}$ such that the following holds. Let k, n, m, d, s be integers such that $s, k \le n$, $p > d^{Cn^2}$ and $m = \lfloor \tilde{c} \cdot \log(p) \rfloor$ and set $D = d^n$. Let $\hat{V} = \mathbf{V}(f_1, \ldots, f_s)$ be a k-dimensional variety such that $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$ are of degree $\le d$. Let V be an irreducible k-dimensional component of $\hat{V}$ such that $|V \cap \mathbb{F}^n| \ge p^{k-1/24}$ and let X_V be a random variable uniformly distributed over $|V \cap \mathbb{F}^n|$. Let $E_{m,D} : \mathbb{F}^n \mapsto \{0,1\}^{m(k-1)}$ be given by Definition 6.1. Then*

$$E_{m,D}(X_V) \quad \textit{is} \quad p^{-\Omega(1)}\textit{-close to uniform.}$$

Let X_V be a random variable as in the above theorem. Observing Corollary 5.5 we see that $S_{m,D}(X_V, Y)$ is close to uniform whenever Y is independent of X_V and uniformly distributed over some sufficiently large ($>> n^3$) subset of $\mathbb{F}$. In order to prove Theorem 6.2 we will need to argue that replacing Y with the distribution $T(X_V) \triangleq \mod_{2^m}(R_{n,D}(X_V))$ does not change the output of $S_{m,D}$ by much. It is not hard to see that $T(X_V)$ is close to being uniformly distributed over $\{0,1\}^m$. This, however, is not enough for the composition to go through, since $T(X_V)$ is highly correlated with X_V. In order to argue about this situation we will use the following theorem proved in [Sha06].

Theorem 6.3 (Composition Theorem [Sha06]). *Let X be a random variable taking values in a finite set Ω. Let $T : \Omega \mapsto M$ and $S : \Omega \times M \mapsto \{0,1\}^\ell$ be two functions, where M is another finite set. For every $y \in M$ and $a \in \{0,1\}^\ell$, such that $\Pr[S(X,y) = a] > 0$, let $X_{y,a}$ denote a random variable distributed according to the conditional distribution of X given $S(X,y) = a$. Suppose that the following two conditions hold*

1. *$T(X)$ is ϵ-close to the uniform distribution on M.*
2. *For every $y \in M$ there exists a set $G_y \subset \{0,1\}^\ell$ such that $\Pr[S(X,y) \notin G_y] \le \epsilon$ and for every $a \in G_y$, such that $\Pr[S(X,y) = a] > 0$, the random variable $T(X_{y,a})$ is ϵ-close to uniform.*

Then

$$S(X, T(X)) \quad \textit{is } O(\epsilon \cdot |M|)\textit{-close to} \quad S(X, Y),$$

where Y is a random variable independent of X and uniform on the set M.

Overview of the proof: We would like to apply Theorem 6.3 with $T = \mod_{2^m}(R_{n,D}(\cdot))$, $S = S_{m,D}(\cdot,\cdot)$ and $X = X_V$, as defined in Theorem 6.2. The first condition of Theorem 6.3 will be satisfied using Theorem 3.2 (extractor with short output). More interesting is the second condition, namely that, conditioned on the event $S_{m,D}(X,u) = a$, the distribution of $T(X)$ is still close to uniform. This will follow by observing that fixing $S_{m,D}(X_V, u) = a$ induces a fixing (or more accurately, a convex combination of fixings) of the form $\mathcal{E}_u^{(k-1)}(X_V) = a'$ with $a' \in \mathbb{F}^{k-1}$. Now, since $\mathcal{E}_u^{(k-1)}$ is linear, every fixing of its output causes X_V to be distributed uniformly on the intersection of $V \cap \mathbb{F}^n$ with some $(k-1)$-dimensional affine subspace. We will show that for most fixings $a' \in \mathbb{F}^{k-1}$ this intersection will define a curve U of degree at most D. Now, since $R_{n,D}$ is not constant on this curve (or on any other degree D curve), and using Bombieri's Theorem, we get that $\mod_{2^m}(R_{n,D}(X_U))$ is close to uniform, where X_U is uniformly distributed on the rational points of U. A complete proof is given below.

6.1 Proof of Theorem 6.2

Let $C, c > 0$ be a constants such that Theorem 5.4 and Corollary 5.5 hold for $m \leq c \cdot \log(p)$ and $p > d^{Cn^2}$ and such that Theorem 3.2, holds for $p > D^{Cn}$ and $m \leq c \cdot \log(p)$. Later on we will set $m = \lfloor \tilde{c} \cdot \log(p) \rfloor$ for another constant $0 < \tilde{c} \leq c$ (and so it is safe to use the above three results). To simplify the notations, let us denote by

$$S(x,u) \triangleq S_{m,D}(x,u),$$

$$R(x) \triangleq R_{n,D}(x),$$

$$T(x) \triangleq \mod_{2^m}(R_{n,D}(x)),$$

$$E(x) \triangleq S(x, T(x)).$$

Recall also that $\mathrm{Uni}(A)$ denotes the uniform distribution on a set A. Using Theorem 3.2 (and observing that $\deg(V) \leq D = d^n$ by Corollary 2.3) we have that

$$T(X_V) \overset{\epsilon_1}{\approx} \mathrm{Uni}(\{0,1\}^m), \tag{6}$$

where $\epsilon_1 = p^{-\Omega(1)}$. We will not be able to apply Theorem 6.3 directly on the composition $E(x) = S(x, T(x))$ since the second condition of the theorem is hard to satisfy for the seeds $u \in B(V)$, where $B(V)$ is defined as in Theorem 5.4 as

$$B(V) \triangleq \left\{ u \in \mathbb{F} \;\middle|\; \mathcal{E}_u^{(k)} \text{ is not } (p^{-11/12}, d^{n^2})\text{-bounded on } |V \cap \mathbb{F}^n| \right\}$$

(notice that $\mathcal{E}_u^{(k-1)}$ is the projection of $\mathcal{E}_u^{(k)}$ on the first $k-1$ coordinates). To overcome this difficulty, we will define a modified distribution on the seeds that will be close to the original distribution $T(X_V)$ and will avoid hitting the seeds that are in $B(V)$. We will use the bound of

$$|B(V)| \leq n^3, \tag{7}$$

provided by Theorem 5.4, and the fact that $2^m >> n^3$, to bound the distance between $T(X_V)$ and the 'modified' distribution. Let

$$M' \triangleq \{0,1\}^m - B(V)$$

and let $\varphi : \{0,1\}^m \mapsto M'$ be a function whose restriction to M' is the identity function. From (7) we have that φ 'moves' at most n^3 elements of $\{0,1\}^m$. Therefore, if we define

$$T'(x) \triangleq \varphi(T(x))$$

we have that, for some $\epsilon_2 = p^{-\Omega(1)}$,

$$T'(X_V) \overset{\epsilon_2}{\approx} T(X_V), \tag{8}$$

and, using (6), that

$$T'(X_V) \overset{\epsilon_2}{\approx} \mathrm{Uni}(M').$$

Our goal is now to apply Theorem 6.3 on the function

$$E'(x) \triangleq S(x, T'(x))$$

(later we will translate this back to $E(x)$ using (8)). This will be possible using the following claim, whose proof we defer to the end of this section, showing that the second condition of Theorem 6.3 is satisfied by E'.

For every $u \in \mathbb{F}$ and $a \in \{0,1\}^{m(k-1)}$ we define $X_{u,a}$ to be a random variable distributed according to the conditional distribution of X_V given $S(X_V, u) = a$.

Claim 6.4. *There exists $\beta > 0$ such that the following holds: For every $u \in M'$ there exists a set $G_u \subset \{0,1\}^{m(k-1)}$ such that*

1. $\Pr[S(X_V, u) \notin G_u] \leq p^{-\beta}$.
2. *For all $a \in G_u$ such that* $\Pr[S(X_V, u) = a] > 0$, $T'(X_{u,a})$ *is $p^{-\beta}$-close to uniform.*

We continue with the proof of Theorem 6.2. Let

$$\epsilon \triangleq \max\{\epsilon_2, p^{-\beta}\},$$

where β is given by Claim 6.4. Applying Theorem 6.3 we get that

$$E'(X_V) = S(X_V, T'(X_V)) \overset{\epsilon'}{\sim} S(X_V, Y'), \tag{9}$$

where

$$\epsilon' = O(\epsilon \cdot |M'|) \leq O(\epsilon \cdot 2^m)$$

and Y' is a random variable independent of X_V and uniformly distributed over M'. We can now set

$$\tilde{c} \triangleq \min\left\{c, \frac{1}{2} \cdot \log_p(1/\epsilon)\right\}$$

and so get that

$$\epsilon' = p^{-\Omega(1)}.$$

We now combine (9) and (8) to claim that

$$E(X_V) = S(X_V, T(X_V)) \overset{\epsilon_2}{\approx} S(X_V, T'(X_V)) \overset{\epsilon'}{\sim} S(X_V, Y'),$$

and so

$$E(X_V) \overset{p^{-\Omega(1)}}{\sim} S(X_V, Y').$$

Using Corollary 5.5, and the fact that $\mathcal{E}_u^{(k-1)}$ is the projection of $\mathcal{E}_u^{(k)}$ on the first $k-1$ coordinates, we get that

$$S(X_V, Y') \overset{p^{-\Omega(1)}}{\sim} \mathrm{Uni}(\{0,1\}^{m(k-1)}).$$

Combining all of the above we have

$$E(X_V) \overset{p^{-\Omega(1)}}{\sim} \mathrm{Uni}(\{0,1\}^{m(k-1)}),$$

as was required. □

6.1.1 Proof of Claim 6.4

Fix $u \in M'$. The function $S(x,u)$ is computed in two steps. In the first, we apply $\mathcal{E}_u^{(k-1)}$ on x to get a value $b \in \mathbb{F}^{k-1}$ and in the second step we take $\mod_{2^m}(b)$. In order to define the set $G_u \subset \{0,1\}^{m(k-1)}$, required by the claim, we will first define a set $G'_u \subset \mathbb{F}^{k-1}$, satisfying similar conditions to the ones we want from G_u, and then define G_u to be the set of elements $a \in \{0,1\}^{m(k-1)}$ such that, conditioned on $\mod_{2^m}(b) = a$, the probability that $b \in G'_u$ is higher than some fixed threshold.

For each $b \in \mathbb{F}^{k-1}$ let

$$V_b \triangleq V \cap \mathbf{V}\left(\mathcal{E}_u^{(k-1)}(x) - b\right)$$

and let Y_b denote a random variable uniformly distributed over $V_b \cap \mathbb{F}^n$. We define

$$G'_u \triangleq \left\{ b \in \mathbb{F}^{k-1} \mid |V_b \cap \mathbb{F}^n| \geq p^{5/6} \text{ and } \dim(V_b) = 1 \right\}.$$

Claim 6.5. *Let G'_u be as above. Then, there exists $\beta' > 0$ such that*

1. $\Pr\left[\mathcal{E}_u^{(k-1)}(X_V) \notin G'_u\right] \leq p^{-\beta'}$.

2. *For every $b \in G'_u$, the random variable $T'(Y_b)$ is $p^{-\beta'}$-close to uniform.*

Proof. We start with item 1. Let us denote by

$$Z \triangleq \mathcal{E}_u^{(k-1)}(X_V)$$

the random variable obtained from X_V in the first step of the computation of $S(x,u)$. The probability that $Z \notin G'_u$ is bounded by the sum of probabilities of the two bad events described in the definition of G'_u. The probability of the first bad event (that is, that V_Z contains less than $p^{5/6}$ rational points) is bounded from above by $p^{-1/8}$, since $|V \cap \mathbb{F}^n| \geq p^{k-1/24}$.

We will now bound the probability of the second bad event, namely that $\dim(V_Z) \neq 1$. Lemma 4.6 tells us that there is a set $A \subset \mathbb{F}^{k-1}$ of size at most $k \cdot D \cdot p^{k-2}$ such that for all $b \notin A$ we have that $\dim(V_b) = 1$ (or that V_b is empty – a case we can safely ignore). Using the fact that $\mathcal{E}_u^{(k)}$ is $(p^{-11/12}, d^{n^2})$-bounded on $V \cap \mathbb{F}^n$ and the observation that Z is the projection of $\mathcal{E}_u^{(k)}(X_V)$ on the first $k-1$ coordinates, we get that

$$\Pr[Z \in A] \leq p^{-11/12} + (k \cdot D \cdot p^{k-2}) \cdot p \cdot d^{n^2} \cdot p^{-k+1/24} \leq p^{-\Omega(1)},$$

(for every $a \in A$ there are at most $p \cdot d^{n^2}$ pre-images outside of some set of weight $p^{-11/12}$). This concludes the proof of item 1. of the claim.

We now prove item 2. Fix some $b \in G'_u$. Recall that $T(x) = \mod_{2^m}(R(x))$ and that T' and T are 'almost' the same. We know from Theorem 3.1 that $R(x)$ is not constant on any of the irreducible components of the curve V_b (notice that $\deg(V_b) \leq D$ since we got it from intersecting V with linear varieties). We can therefore apply Bombieri's Theorem (Theorem 2.4), together with the fact that $|V_b \cap \mathbb{F}^n| \geq p^{5/6}$, to get a bound (for any non trivial character χ) of

$$\frac{1}{|V_b \cap \mathbb{F}^n|}\left|\sum_{x \in V_b \cap \mathbb{F}^n} \chi(R(x))\right| \leq p^{-5/6} \cdot 4 \cdot D^{O(n)} \cdot D^2 \cdot p^{1/2} \leq p^{-\Omega(1)}.$$

Using Lemma 2.5 we get that $T(Y_b)$ is $p^{-\Omega(1)}$-close to uniform, and from the relation between T and T' we also have that $T'(Y_b)$ is $p^{-\Omega(1)}$-close to uniform. □

We can now define the set G_u that will satisfy the requirements of Claim 6.4. Let

$$\beta = \beta'/2,$$

where β' is given by Claim 6.5. For each $a \in \{0,1\}^{m(k-1)}$ let

$$q_a \triangleq \Pr\left[\, Z \in G'_u \quad | \quad \bmod{}_{2^m}(Z) = a\right].$$

We define

$$G_u \triangleq \left\{ a \in \{0,1\}^{m(k-1)} \mid q_a \le p^{-\beta} \right\}.$$

A simple averaging argument now shows that

$$\Pr[S(X_V, u) \notin G_u] \le p^{-\beta}.$$

Now, for every $a \in G_u$, the distribution $T'(X_{u,a})$ is $p^{-\beta}$-close to a convex combination of distributions of the form $T'(Y_b)$ with $b \in G'_u$. Each one of these distributions is $p^{-\Omega(1)}$-close to uniform and, therefore, so is $T'(X_{u,a})$. □

7 Extraction from reducible varieties

In Section 6 we constructed an extractor that works when its input is distributed over one of the irreducible components of a variety $\mathbf{V}(f_1, \ldots, f_s)$, defined using degree d polynomials. In this section we show that the same construction works also for distributions over reducible varieties. This will follow simply from the fact that the intersections of the k-dimensional irreducible components can have dimension at most $k-1$ and thus, by Theorem 2.10, have only few rational points. In fact, our analysis shows that *any* extractor for irreducible varieties will work also for reducible ones. The main result of this section is the following:

Theorem 7.1. *There exist constants $C, \tilde{c}$ such that the following holds. Let k, n, m, d, s be integers such that $s, k \le n$, $p > d^{Cn^2}$ and $m = \lfloor \tilde{c} \cdot \log(p) \rfloor$ and set $D = d^n$. Let $V = \mathbf{V}(f_1, \ldots, f_s)$ be a k-dimensional variety such that $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$ are of degree $\le d$. Suppose $|V \cap \mathbb{F}^n| \ge p^{k-1/48}$ and let X_V be a random variable uniformly distributed over $|V \cap \mathbb{F}^n|$. Let $E_{m,D} : \mathbb{F}^n \mapsto \{0,1\}^{m(k-1)}$ be given by Definition 6.1. Then*

$$E_{m,D}(X_V) \quad \textit{is} \quad p^{-\Omega(1)}\textit{-close to uniform.}$$

Before proving this theorem we prove a lemma about general distributions over sets that have small intersections.

Lemma 7.2. *Let $A_1, \ldots, A_t$ be finite sets and let $A = \cup_i A_i$ be their union. Let*

$$B \triangleq \bigcup_{i \ne j} (A_i \cap A_j)$$

and suppose that $|B| \le \epsilon \cdot |A|$ for some $\epsilon > 0$. For each $i \in [t]$ let X_i denote a random variable uniform over A_i and let $\mu_i = |A_i|/|A|$. Let X denote a random variable uniform on A and let Y denote a random variable distributed according to the convex combination $\sum_i \mu_i \cdot X_i$. Then, X and Y are $O(\epsilon \cdot t)$-close to each other.

Proof. First, observe that $\Pr[X \in B] \le \epsilon$. We also have that

$$\begin{aligned}
\Pr[Y \in B] &= \sum_{i=1}^{t} \mu_i \cdot \Pr[X_i \in B] \\
&= \sum_{i=1}^{t} \frac{|A_i|}{|A|} \cdot \frac{|A_i \cap B|}{|A_i|} \le t \cdot \frac{|B|}{|A|} \le \epsilon \cdot t.
\end{aligned}$$

Another observation is that, for every set $E \subset A$, if $E \cap B = \emptyset$ then $\Pr[X \in E] = \Pr[Y \in E]$. Combining all of the above we get that for any set $E \subset A$ we have

$$\begin{aligned}
|\Pr[X \in E] - \Pr[Y \in E]| &= |\Pr[X \in E \cap B] - \Pr[Y \in E \cap B]| \\
&\le \Pr[X \in B] + \Pr[Y \in B] \\
&\le (t+1) \cdot \epsilon,
\end{aligned}$$

and so X is $(t+1) \cdot \epsilon$-close to Y. □

Proof of Theorem 7.1: Let $V_1, \ldots, V_t$ be the irreducible components of V. We denote by $V' \triangleq V \cap \mathbb{F}^n$ and $V_i' \triangleq V_i \cap \mathbb{F}^n$. Notice that, by Corollary 2.3, $t \le D$. We define the set

$$J \triangleq \left\{ i \in [t] \mid |V_i'| < p^{k-1/24} \right\},$$

and the variety

$$V_J \triangleq \bigcup_{i \in J} V_i.$$

Let us also denote by $V_J' \triangleq V_J \cap \mathbb{F}^n$. From the definition of J we have

$$|V_J'| \le t \cdot p^{k-1/24} \le |V'| \cdot p^{-\Omega(1)}$$

(using the bound of $p^{k-1/48}$ on the size of V' and the fact that p is large enough). Notice also that, from Theorem 2.10, we have that all components V_i with dimension less than k are included in V_J. We now define the variety

$$U = \bigcup_{i \notin J} V_i.$$

and let $U' \triangleq U \cap \mathbb{F}^n$. Let X_U be a random variable uniformly distributed over U'. Since we threw away only components with weight $p^{-\Omega(1)}$ in the distribution of X_V, we have that X_U and X_V are $p^{-\Omega(1)}$-close. Observe that the intersection of any two components in U is of dimension $< k$ and degree $\le D^2$ (using Bezout). Therefore, from Theorem 2.10 we have that the probability that X_U lands on a point that belongs to two or more components is at most $p^{-\Omega(1)}$. We can now apply Lemma 7.2 on X_U to claim that X_U is $p^{-\Omega(1)}$-close to a convex combination of distributions X_i, where X_i is uniform on $V_i \cap \mathbb{F}^n$ and $i \notin J$. We now apply Theorem 6.2 to get that for all $i \notin J$ the random variable $E_{m,D}(X_i)$ is $p^{-\Omega(1)}$-close to uniform. By the above discussion we now get that $E_{m,D}(X_U)$, and therefore also $E_{m,D}(X_V)$, are $p^{-\Omega(1)}$-close to uniform. □

8 Large varieties over small fields

In this section we describe a simple construction of an extractor for varieties V for which $|V \cap \mathbb{F}^n| \geq p^{n \cdot \theta}$ and $\theta > 1/2$. The advantage of this extractor will be the fact that it works when $p = d^{O(1)}$, where d is the degree of the equations defining the variety. Therefore, p can be even constant, when d is constant (that is, independent of n).

The main tool in the construction will be an exponential sum estimate due to Deligne [Del74] (see [MK93] for a statement of the theorem in the form we use here). Before stating the theorem we will need the following definition: Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a homogenous polynomial. We say that f is smooth if the only common zero of the (homogenous) n partial derivatives $\frac{\partial f}{\partial x_i}(x)$, $i \in [n]$, is the all zero vector.

Theorem 8.1 (Deligne). *Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a polynomial of degree d and let f_d denote its homogenous part of degree d. Suppose f_d is smooth. Then, for every non trivial additive character $\chi : \mathbb{F} \mapsto \mathbb{C}^*$ we have*

$$\left| \sum_{x \in \mathbb{F}^n} \chi(f(x)) \right| \leq (d-1)^n \cdot p^{n/2}.$$

Theorem 8.2. *Let n, d, s, m, δ be such that $p > d^{2/\delta}$ and $m = \lfloor (1/4) \cdot \log(p) \rfloor$. Let $V = \mathbf{V}(f_1, \ldots, f_s)$, where $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$ are of degree $\leq d$. Suppose that*

$$|V \cap \mathbb{F}^n| \geq p^{n \cdot (1/2 + \delta)}$$

and let X_V denote a random variable distributed uniformly over $V \cap \mathbb{F}^n$. Let

$$E(x) \triangleq x_1^{d+1} + \ldots + x_n^{d+1}.$$

Then the random variable $\mod_{2^m}(E(X_V))$ *is $p^{-\Omega(1)}$-close to uniform.*

Proof. Let $V' \triangleq V \cap \mathbb{F}^n$ and let χ be a non trivial additive character. We will use the following identity:

$$\frac{1}{p} \sum_{t \in \mathbb{F}} \chi(t \cdot A) = \begin{cases} 0, & A \neq 0 \\ 1, & A = 0 \end{cases}.$$

Using the above identity we can write the exponential sum for $E(x)$ over $x \in V'$ as follows

$$\begin{aligned} \left| \sum_{x \in V'} \chi(E(x)) \right| &= \left| \sum_{x \in \mathbb{F}^n} \chi(E(x)) \cdot \prod_{i \in [s]} \left(\frac{1}{p} \sum_{t_i \in \mathbb{F}} \chi(t_i \cdot f_i(x)) \right) \right| \\ &\leq \frac{1}{p^s} \cdot \sum_{t_1, \ldots, t_s \in \mathbb{F}} \left| \sum_{x \in \mathbb{F}^n} \chi\left(E(x) + \sum_{i \in [s]} t_i \cdot f_i(x) \right) \right|. \end{aligned}$$

Notice that for every fixing of $t_1, \ldots, t_s \in \mathbb{F}$, the homogenous part of highest degree in the polynomial

$$E(x) + \sum_{i \in [s]} t_i \cdot f_i(x)$$

is $E(x)$, and is therefore smooth of degree $d+1$. Using Theorem 8.1 (for every $t \in \mathbb{F}^s$), we get that

$$\left|\sum_{x \in V'} \chi(E(x))\right| \quad \leq \quad d^n \cdot p^{n/2} \leq |V'| \cdot p^{-\frac{\delta \cdot n}{2}},$$

where the last inequality used the bounds on $|V'|$ and p supplied by the conditions of the theorem. We can now apply Lemma 2.5 on the distribution $E(X_V)$ to get that $\mod_{2^m}(E(X_V))$ is close to uniform. □

Notice that we could have used the fact that the exponential sum of $E(x)$ over V' is bounded with error $p^{-\Omega(n)}$ (instead of just $p^{-\Omega(1)}$) in order to show that our extractor in fact has exponentially small error (and to, possibly, extract more bits). We omit the details of these improvements since this scenario (of high entropy) is not the main focus of the paper.

9 Acknowledgments

I am grateful to Ariel Gabizon and Avi Wigderson for helpful conversations on the topic of this work and to my advisors Ran Raz and Amir Shpilka for their continuous support. I thank Dmitry Gourevitch for helpful comments.

References

[Bom66] E. Bombieri. On exponential sums in finite fields. *American Journal of Mathematics*, 88:71–105, 1966.

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

[BRSW06] B. Barak, A. Rao, R. Shaltiel, and A. Wigderson. 2-source dispersers for sub-polynomial entropy and ramsey graphs beating the frankl-wilson construction. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 671–680, New York, NY, USA, 2006. ACM Press.

[CFGP06] Olivier Chevassut, Pierre-Alain Fouque, Pierrick Gaudry, and David Pointcheval. The twist-augmented technique for key exchange. In Moti Yung, Yevgeniy Dodis, Aggelos Kiayias, and Tal Malkin, editors, *Public Key Cryptography*, volume 3958 of *Lecture Notes in Computer Science*, pages 410–426. Springer, 2006.

[CGH$^+$85] Benny Chor, Oded Goldreich, Johan Håstad, Joel Friedman, Steven Rudich, and Roman Smolensky. The bit extraction problem of t-resilient functions (preliminary version). In *Proceedings of the 26th Annual Symposium on Foundations of Computer Science*, pages 396–407, 1985.

[CLO92] D. Cox, J. Little, and D. O'shea. *Ideals, Varieties and Algorithms.* Springer, 1992.

[Dan94] V. Danilov. Algebraic varieties and schemes. In *I. Shafarevich (Ed.), Algebraic Geometry I, Encyclopedia of Mathematical Sciences, Vol. 23,*, pages 167–297, Berlin, 1994. Springer.

[Del74] P. Deligne. La conjecture de weil. *I , Inst. Hautes Etudes Sci. Publ. Math.*, 43:273–307, 1974.

[DGW07] Z. Dvir, A. Gabizon, and A. Wigderson. Extractors and rank extractors for polynomial sources. In *FOCS '07*, 2007.

[Gol95] O. Goldreich. Three XOR-lemmas - an exposition. *Electronic Colloquium on Computational Complexity (ECCC)*, 2(056), 1995.

[GR05] A. Gabizon and R. Raz. Deterministic extractors for affine sources over large fields. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 407–418, Washington, DC, USA, 2005. IEEE Computer Society.

[GRS04] A. Gabizon, R. Raz, and R. Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 394–403, Washington, DC, USA, 2004. IEEE Computer Society.

[Gur05] N. Gurel. Extracting bits from coordinates of a point of an elliptic curve, 2005.

[Har77] Robin Hartshorne. *Algebraic Geometry.* Number 52 in Graduate Texts in Mathematics. Springer, 1977.

[Har92] J. Harris. *Algebraic Geometry - A First Course.* Springer, 1992.

[KRVZ06] J. Kamp, A. Rao, S. Vadhan, and D. Zuckerman. Deterministic extractors for small-space sources. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 691–700, New York, NY, USA, 2006. ACM Press.

[KZ03] J. Kamp and D. Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, 2003.

[MK93] Oscar Moreno and P. Vijay Kumar. Minimum distance bounds for cyclic codes and deligne's theorem. *IEEE Transactions on Information Theory*, 39(5):1524–1534, 1993.

[Rao07] A. Rao. An exposition of bourgain's 2-source extractor. Technical Report TR07-034, ECCC, 2007.

[Sch76] W. M. Schmidt. *Equations over Finite Fields: An Elementary Approach*, volume 536. Springer-Verlag, Lecture Notes in Mathematics, 1976.

[Sch80] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, 1980.

[Sha94] I. R. Shafarevich. *Basic algebraic geometry.* Springer-Verlag New York, Inc., New York, NY, USA, 1994.

[Sha06] R. Shaltiel. How to get more mileage from randomness extractors. In *CCC '06: Proceedings of the 21st Annual IEEE Conference on Computational Complexity*, pages 46–60, Washington, DC, USA, 2006. IEEE Computer Society.

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

[Woo96] T. Wooley. A note on simultaneous congruences. *J. Number Theory*, 58:288–297, 1996.

[Zip79] R. Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of the International Symposiumon on Symbolic and Algebraic Computation*, pages 216–226. Springer-Verlag, 1979.