Appendix A

Sampling and Partitioning

In this appendix we give constructions of samplers and prove Lemmas 2.16, 2.17 and 2.18.

A.1 Sampling Using ℓ-wise Independence

Bellare and Rompel [6] gave a sampler construction based on ℓ-wise independent variables. We use a twist on their method: Suppose we are aiming to hit k/r bits when given a subset S of size k. We generate ℓ-wise independent variables $Z_1, \ldots, Z_n \in [r]$ and define $T = \{i | Z_i = 1\}$. It follows that with high probability $S \cap T$ is of size approximately k/r. This is stated formally in the following lemma. (We explain the difference between this method and that of [6] in Remark A.2.)

Lemma A.1. *For every integers n, k, r, t such that $r \leq k \leq n$ and $6 \log n \leq t \leq \frac{k \log n}{20r}$ there is an explicit $(n, k, \frac{1}{2} \cdot \frac{k}{r}, 3 \cdot \frac{k}{r}, 2^{-\Omega(t/\log n)})$-sampler which uses a seed of t random bits.*

Before proving this lemma we show that Lemma 2.16 is a special case.

Proof. (of Lemma 2.16) We use Lemma A.1 with the parameters n, k and $r = \frac{3k}{n^{1/2+\gamma}}$, $t = \alpha \cdot n^{2\gamma}$. We need to check that $6 \log n \leq t \leq \frac{k \log n}{20r}$. Clearly, $t \geq 6 \log n$ (for a large enough n depending on α and γ). On the other hand,

$$\frac{k \log n}{20r} = \frac{n^{1/2+\gamma} \log n}{60} \geq \alpha \cdot n^{2\gamma} = t$$

(for a large enough n depending on α and γ). Thus, applying Lemma A.1, we get an $(n, k, k/2r, 3k/r, \delta)$-sampler $Samp : \{0,1\}^t \to P([n])$ where

$$\delta = 2^{-\Omega(t/\log n)} = 2^{-\Omega(\alpha \cdot n^{2\gamma}/\log n)} = 2^{-\Omega(\alpha \cdot n^{\gamma})}$$

(for a large enough n depending on α and γ). □

A. Gabizon, *Deterministic Extraction from Weak Random Sources*,
Monographs in Theoretical Computer Science. An EATCS Series,
DOI 10.1007/978-3-642-14903-0, © Springer-Verlag Berlin Heidelberg 2011

We need the following tail inequality for ℓ-wise independent variables due to Bellare and Rompel [6].

Theorem A.1 ([6]). *Let $\ell \geq 6$ be an even integer. Suppose that $X_1, \ldots, X_n$ are ℓ-wise independent random variables taking values in $[0,1]$. Let $X = \sum_{1 \leq i \leq n} X_i$ and $\mu = E(X)$, and let $A > 0$. Then*

$$\Pr[|X - \mu| \geq A] \leq 8 \left(\frac{\ell\mu + \ell^2}{A^2} \right)^{\ell/2}.$$

We now prove Lemma A.1.

Proof. (of Lemma A.1) Let ℓ be the largest even integer such that $\ell \log n \leq t$ and let $q = \lfloor \log r \rfloor \leq \log n$. There are constructions which use $\ell \log n \leq t$ random bits to generate n random variables $Z_1 \ldots, Z_n \in \{0,1\}^q$ that are ℓ-wise independent [13]. The sampler generates such random variables. Let $a \in \{0,1\}^q$ be some fixed value. We define a random variable $T = \{i | Z_i = a\}$. Let $S \subseteq [n]$ be some subset of size k. For $1 \leq i \leq n$ we define a boolean random variable X_i such that $X_i = 1$ if $Z_i = a$. Let $X = |S \cap T| = \sum_{i \in S} X_i$. Note that $\mu = E(X) = k/2^q$ and that the random variables $X_1, \ldots, X_n$ are ℓ-wise independent. Applying Theorem A.1 with $A = k/2r$ we get that

$$\Pr[|X - \mu| \geq A] \leq 8 \left(\frac{\ell k/2^q + \ell^2}{A^2} \right)^{\ell/2}.$$

Note that

$$\{|X - \mu| < A\} \subseteq \left\{ \frac{k}{2^q} - A < X < \frac{k}{2^q} + A \right\} \subseteq \left\{ \frac{k}{r} - A < X < \frac{2k}{r} + A \right\}$$

$$\subseteq \{k_{min} \leq X \leq k_{max}\}$$

for $k_{min} = k/2r$ and $k_{max} = 3k/r$. Note that $\ell \leq \frac{t}{\log n} \leq \frac{k}{20r}$. We conclude that

$$\Pr[k_{min} \leq |S \cap T| \leq k_{max}] \geq 1 - 8 \left(\frac{\ell \frac{k}{2^q} + \ell^2}{(\frac{k}{2r})^2} \right)^{\ell/2} \geq 1 - 8 \left(\frac{4r^2 (\frac{2\ell k}{r} + \frac{\ell k}{20r})}{k^2} \right)^{\ell/2}$$

$$\geq 1 - 8 \left(\frac{10\ell r}{k} \right)^{\ell/2} \geq 1 - 2^{-(\ell/2+3)} \geq 1 - 2^{-\Omega(t/\log n)}.$$

□

Remark A.2. *We remark that this construction is different from the common way of using ℓ-wise independence for sampling [6]. The more common way is to take n/r random variables $V_1, \ldots, V_{n/r} \in [n]$ which are ℓ-wise independent and sample the multi-set $T = \{V_1, \ldots, V_{n/r}\}$. The expected size of the multi-set $|S \cap T|$ is k/r and one gets the same probability of success*

$\delta = 2^{-\Omega(\ell)}$ by the tail inequality of [6]. The two methods require roughly the same number of random bits. Nevertheless, the method of Lemma A.1 has the following advantages:

- *It can also be used for partitioning.*
- *The method used in Lemma A.1 guarantees that T is a set whereas the standard method may produce a multi-set.*
- *The method used in Lemma A.1 can be derandomized and use much fewer bits (at least for small r and large δ). More precisely, suppose that $r \leq \log n$ and say $\ell = 2$. In this range of parameters, one can use $O(\log \log n)$ random bits to generate n variables $Z_1, \ldots, Z_n \in \{0,1\}^{\log r}$ which are $(1/\log n)$-close to being pairwise independent. Thus, the same technique can be used to construct more randomness efficient samplers (at the cost of having a larger error parameter δ.) We use this idea in Section A.2. We remark that in the case of the standard method no savings can be made as it requires variables Z_i over $\{0,1\}^{\log n}$ and even sampling one such variable requires $\log n$ random bits.*

A.2 Sampling and Partitioning Using Fewer Bits

We now derandomize the construction of Lemma A.1 to give schemes which use only $O(\log k)$ bits and prove Lemmas 2.17 and 2.18. These two lemmas follow from the following more general lemma.

Lemma A.3. *Fix any integer $n \geq 16$. Let k be an integer such that $k \leq n$. Let r satisfy $r \leq k$. Let r' be the power of 2 that satisfies $(1/2)r < r' \leq r$. Let $\epsilon > 0$ satisfy $1/kr \leq \epsilon \leq 1/8r$. We can use $7 \log r + 3(\log \log n + \log(1/\epsilon))$ random bits to explicitly partition $[n]$ into r' sets $T_1, \ldots, T_{r'}$ such that for any $S \subseteq [n]$ where $|S| = k$*

$$\Pr(\forall i, \quad k/2r \leq |T_i \cap S| \leq 3k/r) \geq 1 - O(\epsilon \cdot r^3).$$

We prove Lemma A.3 in the next section. We now explain how the two lemmas follow from Lemma A.3.

Proof. (of Lemma 2.18) Set $b = \alpha/38$. Use Lemma A.3 with the parameters $r = k^b$ and $\epsilon = k^{-4b}$ to obtain a partition $T_1, \ldots, T_{r'}$ of $[n]$ where $(1/2)r < r' \leq r$ is a power of 2.

To use Lemma A.3 with these parameters we need $7 \log r + 3(\log \log n + \log(1/\epsilon)) = 7 \log k^b + 3(\log \log n + \log k^{4b})$ random bits. We want to use at most $\alpha \cdot \log k$ bits.
Set $c = 6/\alpha$. Since we assume that $k \geq \log^c n$,

$$(\alpha/2) \log k \geq (\alpha/2)(6/\alpha) \log \log n = 3 \log \log n.$$

So now we need

$$(\alpha/2)\log k \geq 7\log k^b + 3\log k^{4b} = b(7+12)\log k.$$

Or, equivalently,

$$b \leq \alpha/38.$$

Set $e = 1 - b$. So $k/2r = k^e/2$ and $3k/r = 3 \cdot k^e$. Note that $e > 1/2$ as required.
Using Lemma A.3,

$$\Pr(\forall i, \quad k^e/2 \leq |T_i \cap S| \leq 3 \cdot k^e) \geq 1 - O(\epsilon \cdot r^3) = 1 - O(k^{-b}).$$

□

Lemma 2.17 easily follows from Lemma 2.18.

Proof. (of Lemma 2.17) Use Lemma 2.18 with the parameters n, k and α to obtain a partition of $[n]$ $T_1, \ldots, T_m$ and take T_1 as the sample. It is immediate that the required parameters are achieved. □

Proof of Lemma A.3

The sampler construction in Lemma A.1 relied on random variables $Z_1, \ldots, Z_n \in [r]$, which are ℓ-wise independent. We now show that we can derandomize this construction and get a (weaker) sampler by using $Z_1, \ldots, Z_n$ which are only *pairwise ϵ-dependent*. Naor and Naor [44] (and later Alon et al.[2]) gave constructions of such variables using very few random bits. This allows us to reduce the number of random bits required for sampling and partitioning.

The following definition formalizes a notion of limited independence, slightly more general than the one discussed above:

Definition A.4 (ℓ-wise ϵ-dependent variables)**.** *Let D be a distribution. We say that the random variables $Z_1, \ldots, Z_n$ are ℓ-wise ϵ-dependent according to D if for every $M \subseteq [n]$ such that $|M| \leq \ell$, the distribution Z_M (that is, the joint distribution of the Z_is such that $i \in M$) is ϵ-close to the distribution $D^{\otimes |M|}$, i.e., the distribution of $|M|$ independent random variables chosen according to D. We sometimes omit D when it is the uniform distribution. Random bit variables $B_1, \ldots, B_n$ are ℓ-wise ϵ-dependent with mean p if they are ℓ-wise ϵ-dependent according to the distribution $D = (1-p, p)$ on $\{0, 1\}$.*

We need two properties about ℓ-wise ϵ-dependent variables: That they can be generated using very few random bits and that their sum is concentrated around the expectation. The first property is proven in Lemma A.5 and the second in Lemma A.6.

The following theorem states that ℓ-wise ϵ-dependent bit variables can be generated by very few random bits.

Theorem A.2 ([2]). [1] *For any $n \geq 16$, $\ell \geq 1$ and $0 < \epsilon < 1/2$, ℓ-wise ϵ-dependent bits $B_1, \ldots, B_n$ can be generated using $3(\ell + \log\log n + \log(1/\epsilon))$ truly random bits.*

We can generate pairwise ϵ-dependent variables in larger domains using ℓ-wise ϵ-dependent bit variables.[2]

Lemma A.5. *Let $r < n$ be a power of 2. For any $n \geq 16$ and $0 < \epsilon < 1/2$, we can generate pairwise ϵ-dependent variables $Z_1, \ldots, Z_n \in [r]$ using $7\log r + 3(\log\log n + \log(1/\epsilon))$ truly random bits.*

Proof. Using Theorem A.2, we generate $2\log r$-wise ϵ-dependent bit variables $B_1, \ldots, B_{n\log r}$ using $3(2\log r + \log\log(n\log r) + \log(1/\epsilon)) \leq 7\log r + 3(\log\log n + \log(1/\epsilon))$ bits. We partition the B_is into n blocks of size $\log r$ and interpret the ith block as a value $Z_i \in [r]$.

The joint distribution of the bits in any block or 2 blocks is ϵ-close to uniform. Therefore, the Z_is are pairwise ϵ-dependent. □

In the following lemma, we use Chebychev's inequality to show that the sum of pairwise ϵ-dependent bit variables is close to its expectation with high probability.

Lemma A.6. *Let p satisfy $0 < p < 1$. Let $\epsilon > 0$ satisfy $p/k \leq \epsilon \leq p/4$. Let $B_1, \ldots, B_k$ be pairwise ϵ-dependent bit variables with mean p. Let $B = \sum_{i=1}^{k} B_i$.*
Then

$$\Pr(|B - pk| > pk/2) = O(\epsilon/p^2).$$

Proof. Using linearity of expectation we get $|E(B) - pk| \leq \epsilon k$.
Therefore,

$$\Pr(|B - pk| > pk/2) \leq \Pr(|B - E(B)| > pk/2 - \epsilon k).$$

So it's enough to bound

$$\Pr(|B - E(B)| > pk/2 - \epsilon k).$$

Fix any $i, j \in [k]$ where $i \neq j$. The covariance of B_i and B_j will be small since they are almost independent:

$$cov(B_i, B_j) = E(B_i \cdot B_j) - E(B_i)E(B_j)$$

$$= \Pr(B_i = 1; B_j = 1) - \Pr(B_i = 1)\Pr(B_j = 1)$$

[1] The theorem is stated a bit differently and only for odd ℓ in ([2]), but this form is easily deduced from Theorem 3 in that paper by observing that $(\ell + 1)$-wise ϵ-dependence implies ℓ-wise ϵ-dependence.

[2] Actually, a construction of such (and more general types of) variables already appears in [23].

$$\leq (p^2 + \epsilon) - (p - \epsilon)^2 = (1 + 2p - \epsilon)\epsilon \leq 3\epsilon$$

(where the second equality is because B_i and B_j are bit variables)

Therefore, the variance of B won't be too large:

$$Var(B) = \sum_i Var(B_i) + \sum_{i \neq j} cov(B_i, B_j) \leq (p + \epsilon)k + 3\epsilon k^2 \leq pk + 4\epsilon k^2.$$

Therefore, by Chebychev's inequality,

$$\Pr(|B - E(B)| > pk/2 - \epsilon k) < \frac{pk + 4\epsilon k^2}{(pk/2 - \epsilon k)^2}.$$

We required that $\epsilon \leq p/4$, and therefore

$$\leq \frac{pk + 4\epsilon k^2}{(pk/4)^2} = O(1/pk) + O(\epsilon/p^2) = O(\epsilon/p^2)$$

(where the last equality follows by the requirement that $\epsilon \geq p/k$). □

Now we can easily prove Lemma A.3.

Proof. (of Lemma A.3) Let r' be the power of 2 in the statement of the lemma. Using Lemma A.5, we generate pairwise ϵ-dependent $Z_1, \ldots, Z_n \in [r']$. For $1 \leq i \leq r'$, we define $T_i = \{j | Z_j = i\}$.

Assume, w.l.o.g., that $S = \{1, \ldots, k\}$. Given $i \in [r']$, define the bit variables $B_1, \ldots, B_k$ by $B_j = 1 \Leftrightarrow Z_j = i$. It is easy to see that the B_js are pairwise 2ϵ-dependent with mean $1/r'$. Define $C_i = \sum_{j=1}^{k} B_j$.

Note that $C_i = |T_i \cap S|$. Notice that $1/r'$ and 2ϵ satisfy the requirements in Lemma A.6.

Using Lemma A.6,

$$\Pr(|C_i - k/r'| > k/2r') = O(\epsilon \cdot (r')^2) = O(\epsilon \cdot r^2).$$

Using the union bound,

$$\Pr(\exists i \; s.t \; |C_i - k/r'| > k/2r') = O(\epsilon \cdot r^3).$$

Thus, we can obtain a partition $T_1, \ldots, T_{r'}$ of $[n]$ such that, with probability at least $1 - O(\epsilon \cdot r^3)$,

$$\forall i \; k/2r' \leq |T_i \cap S| \leq 3k/2r',$$

which implies that with at least the same probability,

$$\forall i \; k/2r \leq |T_i \cap S| \leq 3k/r.$$

□

Appendix B

Basic Notions from Algebraic Geometry

In Section 4.5 we use a theorem of Bombieri [8] regarding character sums over curves. The very statement, let alone the applicability of Bombieri's theorem, requires some basic notions from algebraic geometry. In this appendix, we give some basic background necessary for stating the theorem and applying it as done in Section 4.5. The main issue in Section 4.5 is to show that the varieties that come up there are suitable for the theorem. Specifically, we need to show that these varieties are indeed curves, i.e., have dimension 1, and that their 'degree' is not too large. (All these terms will be defined formally). For this purpose, we need some lemmas regarding the dimension and degree of intersections of varieties. Another issue is that Bombieri's theorem is stated for projective curves while we want to apply it on affine curves. For this purpose, we need some lemmas on the relations between affine and projective varieties. We note that all these issues are standard. We stress that this section is far from a full introduction to basic algebraic geometry. For a very accessible introduction we recommend [17], of which most the definitions and notations in this section follow.
Throughout this section $\mathbb{F}$ will always denote an algebraically closed field.

B.1 Affine and Projective Varieties

The basic objects of study in algebraic geometry are the sets of solutions to a system of polynomial equations. Such a set is called a *variety*. We now formally define affine space and affine varieties.

Definition B.1 (affine space)**.** *We define n-dimensional affine space over $\mathbb{F}$ as*[1]

$$\mathbb{F}^n \triangleq \{(a_1, \ldots, a_n) \mid a_i \in \mathbb{F}\}.$$

[1] In most textbooks in algebraic geometry the notation $\mathbb{A}^n$ is used rather than $\mathbb{F}^n$. However, in [17], which we are following, $\mathbb{F}^n$ is used.

Definition B.2 (affine variety)**.** *Let $f_1, \ldots, f_s$ be polynomials in $\mathbb{F}[x_1, \ldots, x_n]$. We set*

$$\mathbf{V}(f_1, \ldots, f_s) = \{(a_1, \ldots, a_n) \in \mathbb{F}^n \mid \forall\, 1 \leq i \leq s\ f_i(a_1, \ldots, a_n) = 0\}.$$

We call $\mathbf{V}(f_1, \ldots, f_s)$ the affine variety *defined by $f_1, \ldots, f_s$. A subset $V \subseteq \mathbb{F}^n$ is an* affine variety *if $V = \mathbf{V}(f_1, \ldots, f_s)$ for some set of polynomials $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$. We say that V is* reducible *if it can be written as $V = V_1 \cup V_2$ where the V_is are affine varieties such that $V \neq V_1, V_2$. Otherwise, we say that V is* irreducible.[2]

As a simple example of an affine variety, take $V = \mathbf{V}(x_1 \cdot x_2) \subseteq \mathbb{F}^2$. Note that V is reducible as it is the union of the varieties $V_1 = \mathbf{V}(x_1)$ and $V_2 = \mathbf{V}(x_2)$, i.e., the sets $\{(0, x_2) | x_2 \in \mathbb{F}\}, \{(x_1, 0) | x_1 \in \mathbb{F}\} \subseteq \mathbb{F}^2$. It can be shown that V_1 and V_2 are irreducible. Note that this is not a disjoint union as $V_1 \cap V_2 = (0, 0)$.

Though affine space and affine varieties seem to be the natural objects we want to investigate, it turns out to be very useful to work in *projective space*. Intuitively, projective space is affine space extended with additional 'extra points'. This intuition may not be clear from the following definition but will become clearer later on.

Definition B.3 (projective space)**.** *We define an equivalence relation $\sim$ over $\mathbb{F}^{n+1} \backslash \{0\}$ by setting*

$$(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n)$$

if and only if there exists a nonzero $\lambda \in \mathbb{F}$ such that $(x_0, \ldots, x_n) = (\lambda \cdot y_0, \ldots, \lambda \cdot y_n)$. We define n-dimensional projective space $\mathbb{P}^n$ over $\mathbb{F}$ to be the set of all equivalence classes of $\sim$. Thus,

$$\mathbb{P}^n = (\mathbb{F}^{n+1} - \{0\}) / \sim .$$

Each non-zero $(n+1)$-tuple $(x_0, \ldots, x_n) \in \mathbb{F}^n$ defines a point $p \in \mathbb{P}^n$. We say that $(x_0, \ldots, x_n)$ are homogenous coordinates *of p.*

We say that a polynomial $f \in \mathbb{F}[x_0, \ldots, x_n]$ is *homogenous* if all of its monomials have the same total degree. It is easy to see that for a homogenous polynomial f of total degree d and any nonzero $\lambda \in \mathbb{F}$

$$f(\lambda \cdot a_0, \ldots, \lambda \cdot a_n) = \lambda^d f(a_0, \ldots, a_n).$$

In particular, $f(\lambda \cdot a_0, \ldots, \lambda \cdot a_n) = 0$ if and only if $f(a_0, \ldots, a_n) = 0$. Thus, the set of 'zeros' of f is a well-defined object in $\mathbb{P}^n$.

This leads to the following definition.

[2] In many textbooks, the term variety always means an irreducible variety and general varieties are called *algebraic sets*.

Definition B.4 (projective variety)**.** *Let* $f_1, \ldots, f_s \in \mathbb{F}[x_0, \ldots, x_n]$ *be homogenous polynomials. We set*

$$\mathbf{V}(f_1, \ldots, f_s) = \{(a_0, \ldots, a_n) \in \mathbb{P}^n \mid \forall\, 1 \leq i \leq s\ f_i(a_0, \ldots, a_n) = 0\}.$$

A subset $V \subseteq \mathbb{P}^n$ *is a* projective variety *if* $V = \mathbf{V}(f_1, \ldots, f_s)$ *for some set of homogenous polynomials* $f_1, \ldots, f_s \in \mathbb{F}[x_0, \ldots, x_n]$*. We say that* V *is* reducible *if it can be written as* $V = V_1 \cup V_2$*, where the* V_i*s are projective varieties such that* $V \neq V_1, V_2$*. Otherwise, we say that* V *is* irreducible*.*

An important basic property of (affine and projective) varieties is that they decompose into irreducible varieties in a unique way. Thus, we can speak unambiguously about the irreducible components of a variety.

Proposition B.1. *-[[17], Chapter 4, §6, Theorem 4, and Chapter 8, §3, Theorem 6] We say that* $V = V_1 \cup \ldots \cup V_m$ *is a* minimal decomposition *of* V *if* $V_i \nsubseteq V_j$ *for every* $i \neq j$*. Let* V *be an affine (projective) variety. Then* V *has a minimal decomposition*

$$V = V_1 \cup \ldots \cup V_m$$

where the V_i*s are irreducible affine (projective) varieties. Furthermore, this minimal decomposition is unique up to the order in which* $V_1, \ldots, V_m$ *are written.*

B.2 Varieties and Ideals

An affine variety is essentially a geometric object — a set of points in the space $\mathbb{F}^n$. A fundamental idea in algebraic geometry is to relate a variety to an algebraic object. This algebraic object will be the set of all polynomials that vanish on the variety. It is easy to see that this set of polynomials forms an ideal in the ring $\mathbb{F}[x_1, \ldots, x_n]$. First we recall some basic facts and notation regarding ideals in $\mathbb{F}[x_1, \ldots, x_n]$. For $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$ we denote by $< f_1, \ldots, f_s >$ the ideal generated by $f_1, \ldots, f_s$. That is,

$$< f_1, \ldots, f_s > \triangleq \left\{ \sum_{i=1}^{s} g_i \cdot f_i \mid \forall\, 1 \leq i \leq s\ g_i \in \mathbb{F}[x_1, \ldots, x_n] \right\}.$$

By the Hilbert Basis Theorem (see [17], Chapter 2, §5) every ideal $I \subset \mathbb{F}[x_1, \ldots, x_n]$ is *finitely generated*, i.e., $I = < f_1, \ldots, f_s >$ for some $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n]$. For an ideal $I = < f_1, \ldots, f_s >$, it is easy to see that a point $(a_1, \ldots, a_n) \in \mathbb{F}^n$ is a zero of every $f \in I$ if and only if it is a zero of $f_1, \ldots, f_s$.

Definition B.5 (affine varieties and ideals)**.** *For an affine variety* $V \subseteq \mathbb{F}^n$ *we define* $\mathbf{I}(V)$ *to be the ideal of all polynomials* f *such that* $f(a_1, \ldots, a_n) = 0$ *for every* $(a_1, \ldots, a_n) \in V$*. For an ideal* $I = < f_1, \ldots, f_s > \subseteq \mathbb{F}[x_1, \ldots, x_n]$ *we define* $\mathbf{V}(I) \subseteq \mathbb{F}^n$ *to be the affine variety* $\mathbf{V}(I) = \{(a_1, \ldots, a_n) \mid f(a_1, \ldots, a_n) = 0,\ \forall f \in I\} = \mathbf{V}(f_1, \ldots, f_s)$.

Before making the corresponding definitions for projective varieties we will need some terminology. We remarked above that it makes sense to ask whether a homogenous polynomial $f \in \mathbb{F}[x_0, \ldots, x_n]$ vanishes at a point $p \in \mathbb{P}^n$. For a non-homogenous polynomial $f \in \mathbb{F}[x_0, \ldots, x_n]$ we say that $f(p) = 0$ for $p \in \mathbb{P}^n$ if $f(a_0, \ldots, a_n) = 0$ for all representatives $(a_0, \ldots, a_n)$ of p.

We say that an ideal $I \subseteq \mathbb{F}[x_0, \ldots, x_n]$ is *homogenous* if it is generated by a set of homogenous polynomials, i.e., $I =< f_1, \ldots, f_s >$ where $f_1, \ldots, f_s$ are homogenous. We can now make the following definitions.

Definition B.6 (projective varieties and homogenous ideals)**.** *For a projective variety $X \subseteq \mathbb{P}^n$ we define $\mathbf{I}(X)$ to be the ideal of all polynomials f with $f(p) = 0$ for every $p \in X$. It can be shown that $\mathbf{I}(X)$ is always a homogenous ideal. For a homogenous ideal $I \subseteq \mathbb{F}[x_0, \ldots, x_n]$ we define $\mathbf{V}(I) \subseteq \mathbb{P}^n$ to be the projective variety of all points $p \in \mathbb{P}^n$ that are zeros of all polynomials $f \in I$. If $I =< f_1, \ldots, f_s >$ for homogenous polynomials $f_1, \ldots, f_s$ then it can be shown that $\mathbf{V}(I) = \mathbf{V}(f_1, \ldots, f_s)$.*

One reason the correspondence between ideals and varieties is useful is that operations on ideals have simple corollaries in terms of the corresponding varieties. We need the following fact about intersections of ideals.

Proposition B.2 ([17], Chapter 4, §3, Theorem 15, and Chapter 8, §3, Exercise 7)**.** *Let I_1, I_2 be ideals in $\mathbb{F}[x_1, \ldots, x_n]$ or homogenous ideals in $\mathbb{F}[x_0, \ldots, x_n]$. Then*

$$\mathbf{V}(I_1 \cap I_2) = \mathbf{V}(I_1) \cup \mathbf{V}(I_2).$$

B.3 The Dimension and Degree of a Variety

There are several equivalent definitions of the dimension and degree of a variety (degree is defined only for projective varieties). Here we define dimension and degree in terms of the Hilbert polynomial of a variety. First we need to define the Hilbert function and Hilbert polynomial of an ideal. The definitions are taken from [17].

We say that an ideal I is a *monomial ideal* if it is generated by a set of monomials.[3]. For example, $I =< x_1, x_2^2 >$ is a monomial ideal. We first define the Hilbert function for monomial ideals.

Definition B.7 (Hilbert function of a monomial ideal)**.** *Let I be a monomial ideal in $\mathbb{F}[x_1, \ldots, x_n]$. The* affine Hilbert function *of I, denoted by ${}^aHF_I(s)$, is a function on non-negative integers defined by ${}^aHF_I(s) =$ number of monic monomials in $\mathbb{F}[x_1, \ldots, x_n]$ of degree at most s not contained in I. Similarly, let I be a homogenous monomial ideal in $\mathbb{F}[x_0, \ldots, x_n]$. The* Hilbert function

[3] By Dickson's Lemma ([17], Chapter 2, §4, Theorem 5), if I is a monomial ideal it can always be generated by a finite set of monomials.

of I, denoted by $HF_I(s)$, is a function on non-negative integers defined by $HF_I(s)$ = number of monic monomials in $\mathbb{F}[x_0, \ldots, x_n]$ of degree exactly *s not contained in I.*

Roughly speaking, for a monomial ideal I the monomials not in I are a basis for the space of polynomials that are 'different modulo I'. Thus, ${}^aHF_I(s)$ is the dimension of the space of such polynomials of degree at most s. This is the idea behind the definition of the Hilbert function for general ideals. First we need some notation. For a subset of polynomials $V \subseteq \mathbb{F}[x_1, \ldots, x_n]$ and a non-negative integer s, we denote by $V_{\leq s} \subseteq \mathbb{F}[x_1, \ldots, x_n]$ the set of polynomials in V of (total) degree at most s. For example, $\mathbb{F}[x_1, \ldots, x_n]_{\leq s}$ is the set of all polynomials of degree at most s. Similarly, for a subset $V \subseteq \mathbb{F}[x_0, \ldots, x_n]$ we denote by $V_s \subseteq \mathbb{F}[x_0, \ldots, x_n]$ the set of all polynomials in V of degree *exactly* s. Note that if $V \subseteq \mathbb{F}[x_1, \ldots, x_n]$ is a linear subspace, then so are $V_{\leq s}$ and V_s. In particular if $I \subseteq \mathbb{F}[x_1, \ldots, x_n]$ is an ideal, then it is also a linear subspace, and so is $V_{\leq s}$. We recall a basic notion for linear algebra: For subspaces $W \subseteq V \subseteq \mathbb{F}[x_1, \ldots, x_n]$ we denote by V/W the *quotient space* of equivalence classes of V over W. That is, we define an equivalence relation $\sim$ over V by $v \sim v' \leftrightarrow v - v' \in W$ and let V/W be the space of these equivalence classes. We can now make the following definition.

Definition B.8 (Hilbert function of a general ideal)**.** *Let I be an ideal in $\mathbb{F}[x_1, \ldots, x_n]$. The* affine Hilbert function *of I, denoted by ${}^aHF_I(s)$, is defined as*
${}^aHF_I(s) \triangleq dim\,(\mathbb{F}[x_1, \ldots, x_n]_{\leq s}/I_{\leq s})$.
Let I be a homogenous ideal in $\mathbb{F}[x_0, \ldots, x_n]$; the Hilbert function *of I, denoted $HF_I(s)$, is defined as $HF_I(s) \triangleq dim\,(\mathbb{F}[x_0, \ldots, x_n]_s/I_s)$.*

It can be shown that for large enough input s, the Hilbert Function coincides with a polynomial.

Theorem B.1 (see [17] Chapter 9, §3)**.**

1. *Let I be an ideal in $\mathbb{F}[x_1, \ldots, x_n]$. There exists a polynomial ${}^aHP_I(s)$ such that for large enough s, ${}^aHP_I(s) = {}^aHF_I(s)$. We call ${}^aHP_I(s)$ the* affine Hilbert polynomial *of I.*

2. *Let I be a homogenous ideal in $\mathbb{F}[x_0, \ldots, x_n]$. There exists a polynomial $HP_I(s)$ such that for large enough s, $HP_I(s) = HF_I(s)$. We call $HP_I(s)$ the* Hilbert polynomial *of I.*

Let $V \subseteq \mathbb{F}^n$ be an affine variety with $I = \mathbf{I}(V)$. Let's try to see why it could make sense to define the dimension of a variety in terms of the affine Hilbert polynomial of I. Since I is exactly the set of polynomials that vanish on V, polynomials $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ are identical on V if and only if $f - g \in I$. It follows that $\mathbb{F}[x_1, \ldots, x_n]/I$ is exactly the space of polynomial functions from V to $\mathbb{F}$. Now recall that for a linear subspace $A \subseteq \mathbb{F}^n$, the

dimension of A can be defined as the dimension of the space of linear functions from A to $\mathbb{F}$. Similarly, we could try to define the dimension of V as the dimension of the space of *polynomial* functions from V to $\mathbb{F}$, i.e., the dimension of $\mathbb{F}[x_1, \ldots, x_n]/I$. However, since the polynomials in this space have unbounded degree, $\mathbb{F}[x_1, \ldots, x_n]/I$ has infinite dimension. Instead, we can take an 'asymptotic' approach and define the dimension of V by how fast this space grows as we increase the degree of the polynomials. More accurately, we can define $dim(V)$ by how fast ${}^aHP_I(s) = dim(\mathbb{F}[x_1, \ldots, x_n]_{\leq s}/I_{\leq s})$ grows as s increases. This corresponds to the degree of ${}^aHP_I(s)$.

Definition B.9 (dimension of a variety)**.** *Let $V \subseteq \mathbb{F}^n$ be an affine variety and let $I = \mathbf{I}(V)$. The* dimension *of V, denoted $dim(V)$, is defined to be the degree of ${}^aHP_I(s)$. Let $V \subseteq \mathbb{P}^n$ be a projective variety and let $I = \mathbf{I}(V)$. The* dimension *of V is defined to be the degree of $HP_I(s)$.*

To gain intuition on the above definition, it is helpful to see how it coincides with the definition of dimension for a linear subspace. Take for example the subspace $V \subseteq \mathbb{F}^n$ defined by the constraints $\{x_1 = 0,\ x_2 = 0\}$. Then $I \triangleq \mathbf{I}(V) =< x_1, x_2 >$ and the monomials *not* in I are exactly the monomials $x_3^{a_3} \cdots x_n^{a_n}$, where $a_3, \ldots, a_n$ are non-negative integers. In particular, the number of such monomials of degree at most s is $\binom{n-2+s}{n-2}$, which is a degree $n-2$ polynomial in s. Therefore, since I is a monomial ideal by the definition above, $dim(V) = deg(HP_I(s)) = n - 2$.

The following property of the dimension of a variety will be very useful for us later on.

Proposition B.3 ([17], Chapter 9, §4 Corollary 9)**.** *Let V be an affine or projective variety. The dimension of V is equal to the maximum of the dimensions of its irreducible components.*

We now define the *degree* of a projective variety (degree is not defined for affine varieties).

Definition B.10 (degree of a variety)**.** *The* degree *of V, denoted by $deg(V)$, is defined to be the leading coefficient of $HP_I(s)$ multiplied by $dim(V)$!.*

Though not immediate from the definition, it can be shown that the degree is always a non-negative integer. To gain intuition on the above definition, let us see how it coincides with the definition of degree for a univariate polynomial. For simplicity of the example we'll assume degree is defined for an affine variety V in a similar way to projective varieties. That is, $deg(V)$ is the leading coefficient of the affine Hilbert polynomial of $\mathbf{I}(V)$ times $dim(V)$!. Let $I \subseteq \mathbb{F}[x_1]$ be the ideal $< x_1^3 - 1 >$. It can be shown that $I = \mathbf{I}(V)$ where $V = \mathbf{V}(x_1^3 - 1) \in \mathbb{F}$, i.e., V is simply the roots of $x_1^3 - 1$ and $|V| = 3$ (since $\mathbb{F}$ is algebraically closed). Furthermore, it can be seen that $\{1, x_1, x_1^2\}$ is a basis for $k[x_1]/I$. Hence, $HP_I(s)$ is simply the constant 3, and therefore $dim(V) = deg(HP_I(s)) = 0$ and $deg(V) = 3 \cdot 0! = 3$. Thus $deg(V)$ bounds the size of V. It can be shown that $deg(V)$ always bounds $|V|$ when V is a projective variety of finite size.

B.4 The Projective Closure of an Affine Variety

We call an affine (projective) variety of dimension 1 an affine (projective) curve. As mentioned above, in Section 4.5 we use a theorem of Bombieri[8] for affine curves while in [8] the theorem is stated for projective curves. The transition between the cases, presented in subsection B.7, is completely standard. For this purpose, the following definitions enable us to relate an affine variety with its 'corresponding' projective variety. First we need the following definitions.

Definition B.11 (homogenization)**.**

- *For a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ of degree d, we define the homogenized version $f^h \in \mathbb{F}[x_0, \ldots, x_n]$ by*

$$f^h(x_0, x_1, \ldots, x_n) = x_0^d \cdot f(x_1/x_0, \ldots, x_n/x_0).$$

- *Similarly, for an ideal $I = < f_1, \ldots, f_s >$ we define the ideal $I^h = < f^h | f \in I >$. Note that I^h is always a homogenous. In particular, it is easy to see that $I^h =< f_1^h, \ldots, f_s^h >$.*

We can now define the projective closure of an affine variety.

Definition B.12 (projective closure)**.** *Let $V \subseteq \mathbb{F}^n$ be an affine variety with ideal $I = \mathbf{I}(V)$. We define the* projective closure $\overline{V} \subseteq \mathbb{P}^n$ *to be the projective variety $\mathbf{V}(I^h)$. Let $U_0 \subseteq \mathbb{P}^n$ be defined as $U_0 = \{(a_0, a_1, \ldots, a_n) \in \mathbb{P}^n | a_0 = 1\}$. Note that U_0 can be identified with $\mathbb{F}^n$. Thus, we can think of an affine variety $V \subseteq \mathbb{F}^n$ as being contained in U_0. For a projective variety $V \subseteq \mathbb{P}^n$, we denote $V^a \triangleq V \cap U_0$. Intuitively, this is "the affine part of V".*

The following propositions show various connections between an affine variety and its projective closure.

Proposition B.4 ([17] Chapter 8, §4, Proposition 7 and Exercise 9)**.** *Let $V \subseteq \mathbb{F}^n$ be an affine variety. Then*

1. *$\overline{V} \cap U_0 = V$.*
2. *V is irreducible if and only if $\overline{V}$ is irreducible.*

Proposition B.5 ([17] Chapter 9, §3, Theorem 12)**.** *Let $V \subseteq \mathbb{F}^n$ be an affine variety. Then*

$$dim(V) = dim(\overline{V}).$$

Proposition B.6 ([17] Chapter 8, §4, Theorem 8)**.** *Let $f_1, \ldots, f_r \in \mathbb{F}[x_1, \ldots, x_n]$ be polynomials such that $V = \mathbf{V}(f_1, \ldots, f_r) \subseteq \mathbb{F}^n$ is non-empty. Then*

$$\overline{V} = \mathbf{V}(f_1^h, \ldots, f_r^h).$$

Claim B.12.1. *Let $V_1, \ldots, V_r \subseteq \mathbb{F}^n$ be affine varieties. Then $\overline{V_1 \cup \ldots \cup V_r} = \overline{V}_1 \cup \ldots \cup \overline{V}_r$.*

Proof. We prove the claim for $r = 2$. The statement for general r follows by induction.

Let I_1, I_2 be the ideals $\mathbf{I}(V_1), \mathbf{I}(V_2)$ respectively. It can be shown that $\mathbf{V}(I_1^h \cap I_2^h) = \mathbf{V}((I_1 \cap I_2)^h)$. We have

$$\overline{V_1 \cup V_2} = \mathbf{V}((I_1 \cap I_2)^h) = \mathbf{V}(I_1^h \cap I_2^h) = \mathbf{V}(I_1^h) \cup \mathbf{V}(I_2^h) = \overline{V}_1 \cup \overline{V}_2,$$

where we used Proposition B.2 in the first and second to last equalities. □

Corollary B.1. *Let $V \subseteq \mathbb{F}^n$ be an affine variety with irreducible components $V_1, \ldots, V_r$. Then, the irreducible components of $\overline{V} \subseteq \mathbb{P}^n$ are $\overline{V}_1, \ldots, \overline{V}_r$.*

Proof. Follows from Proposition B.4 and Claim B.12.1. □

Claim B.12.2. *Let $V \subseteq \mathbb{F}^n$ be an affine variety. If $f \in \mathbb{F}[x_1, \ldots, x_n]$ does not vanish identically on V then f^h does not vanish identically on $\overline{V} \subseteq \mathbb{P}^n$.*

Proof. For any $a \in \mathbb{F}^n$, $f(a) = f^h(1, a)$. Therefore, if $f(a) \neq 0$ for $a \in V$, then $f^h(1, a) \neq 0$, where $(1, a) \in \overline{V}$ by Proposition B.4. □

B.5 The Dimension of Intersections of Hypersurfaces

We say that an affine (projective) variety V is a *hypersurface* if $V = \mathbf{V}(f)$ for a (homogenous) polynomial f. In this subsection we state and prove standard results regarding the dimension of intersections of hypersurfaces. The following definition will be important.

Definition B.13. *We say that an affine or projective variety V has* pure dimension *if all its irreducible components have the same dimension.*

We need the following propositions about the intersection of a hypersurface with a variety.

Proposition B.7 ([17] Chapter 9, §4, Proposition 7)**.** *Let $V \subseteq \mathbb{P}^n$ be a projective variety with $dim(V) \geq 1$. Then for any non-constant homogenous polynomial $f \in \mathbb{F}[x_0, \ldots, x_n]$, $V \cap \mathbf{V}(f) \neq \emptyset$.*

Proposition B.8 ([61], Chapter I, §6, Corollary 1 of Theorem 5)**.** *Let $V \subseteq \mathbb{P}^n$ be an irreducible projective variety. Let $f \in \mathbb{F}[x_0, \ldots, x_n]$ be a homogenous polynomial that does not vanish identically on V and denote $H = \mathbf{V}(f)$. If $V \cap H \neq \emptyset$, then $V \cap H$ has pure dimension $dim(V) - 1$.*

Claim B.13.1. *Let $V \subseteq \mathbb{P}^n$ be a projective variety of pure dimension $dim(V) \geq 1$. Let $f \in \mathbb{F}[x_0, \ldots, x_n]$ be a non-constant homogenous polynomial and let $H = \mathbf{V}(f) \subseteq \mathbb{P}^n$. Assume that f does not vanish identically on any of the irreducible components of V. Then $V \cap H$ has pure dimension $dim(V) - 1$.*

Proof. Let $V = Z_1 \cup \ldots \cup Z_k$ be the decomposition of V into irreducible components. Fix any $j \in [k]$. By Proposition B.7, $Z_j \cap H$ is non-empty, and since f does not vanish on Z_j, by Proposition B.8 all irreducible components of $Z_j \cap H$ have dimension $dim(V) - 1$. To conclude, note that the union of the irreducible components of $Z_j \cap H$ over all $j \in [k]$ is $V \cap H$, and therefore the irreducible components of $V \cap H$ are just a subset of these components (excluding any component that is contained in another). Hence, all irreducible components of $V \cap H$ have dimension $dim(V) - 1$ and the claim follows. □

As a special case we get the following.

Corollary B.2. *Let $f \in \mathbb{F}[x_0, \ldots, x_n]$ be a non-constant homogenous polynomial. Then the hypersurface $H = \mathbf{V}(f) \subseteq \mathbb{P}^n$ has pure dimension $n - 1$.*

Proof. $\mathbb{P}^n$ can be shown to be irreducible and in particular has pure dimension. Thus, using Claim B.13.1 with $V = \mathbb{P}^n$ we get the desired result. □

We can now state and prove the main lemma we use regarding the dimension of intersections of hypersurfaces.

Lemma B.9. *Let $0 < r < n$ be integers and let $f_1, \ldots, f_r \in \mathbb{F}[x_0, \ldots, x_n]$ be non-constant homogenous polynomials. For each $i \in [r]$, let $H_i = \mathbf{V}(f_i) \subseteq \mathbb{P}^n$ and $V_i = \mathbf{V}(f_1, \ldots, f_i) = H_1 \cap \ldots \cap H_i$. Then*

1. *All irreducible components of the projective variety V_r have dimension at least $n - r$.*
2. *Suppose furthermore that for each $2 \leq i \leq r$, f_i does not vanish identically on any of the irreducible components of V_{i-1}. Then V_r is a projective variety of pure dimension $n - r$.*

Proof. We prove the first item by induction on r. For $r = 1$ this follows from Corollary B.2. Assume the claim for $r - 1$. Let $V_{r-1} = Z_1 \cup \ldots \cup Z_k$ be the decomposition of V_{r-1} into irreducible components. Fix any $j \in [k]$. Similarly to the proof of Claim B.13.1, we will show that all the irreducible components of $Z_j \cap H_r$ have dimension at least $n - r$, and since the irreducible components of V_r are a subset of these, the claim follows. From the induction hypothesis we have $dim(Z_j) \geq n - (r - 1)$. If f_r vanishes on Z_j then $Z_j \cap H_r = Z_j$ (which is the only irreducible component) and we are done. Otherwise, by Claim B.13.1 all components of $Z_j \cap H_r$ have dimension at least $n - r$.
We now prove the second item by induction on r. For $r = 1$ this is exactly Corollary B.2. Assume the claim for $r - 1$. Then by the induction hypothesis, V_{r-1} has pure dimension $n - r + 1$. Therefore, by Claim B.13.1 $V_r = V_{r-1} \cap H_r$ has pure dimension $n - r$. □

We also need the corresponding lemma in affine space.

Lemma B.10. *Let $0 < r < n$ be integers and let $f_1, \ldots, f_r \in \mathbb{F}[x_1, \ldots, x_n]$ be non-constant polynomials. For each $i \in [r]$, let $H_i = \mathbf{V}(f_i) \subseteq \mathbb{F}^n$ and let $V_i = \mathbf{V}(f_1, \ldots, f_i) = H_1 \cap \ldots \cap H_i$. Suppose that for each $2 \leq i \leq r$, f_i does not vanish identically on any of the irreducible components of the affine variety V_{i-1}. Then, if V_r is non-empty it is an affine variety of pure dimension $n - r$.*

Proof. For $1 \leq i \leq r$, let $X_i = \mathbf{V}(f_1^h, \ldots, f_i^h)$. By Proposition B.6, for every $1 \leq i \leq r$ $X_i = \overline{V}_i$. Therefore, by Corollary B.1 the irreducible components of X_{i-1} are simply the projective closures of the irreducible components of V_{i-1}. By Claim B.12.2 it follows that f_i^h does not vanish identically on any of the irreducible components of X_{i-1}. Hence, we can use Lemma B.9, and X_r is a projective variety of pure dimension $n - r$; and since $X_r = \overline{V}_r$, using Proposition B.5 V_r is an affine variety of pure dimension $n - r$. □

B.6 The Degree of Intersections of Hypersurfaces

We now discuss degree. The main result we prove is the following corollary of Bezout's theorem.

Lemma B.11. *Let $f_1, \ldots, f_r \in \mathbb{F}[x_0, \ldots, x_n]$ be non-constant homogenous polynomials of degrees $d_1, \ldots, d_r$ respectively, and let $D = d_1 \cdot \cdots d_r$. Let $X = \mathbf{V}(f_1, \ldots, f_r) \subseteq \mathbb{P}^n$. Assume that $dim(X) = n - r$. Then*

1. *$deg(X) \leq D$.*
2. *The number of irreducible components of X is at most D.*

Using this Lemma, we immediately get a bound on the number of irreducible components of an affine variety.

Lemma B.12. *Let $f_1, \ldots, f_r \in \mathbb{F}[x_1, \ldots, x_n]$ be non-constant polynomials of degrees $d_1, \ldots, d_r$, respectively, and let $D = d_1 \cdots d_r$. Let $V = \mathbf{V}(f_1, \ldots, f_r) \subseteq \mathbb{F}^n$. Assume that V is non-empty and $dim(V) = n - r$. Then the number of irreducible components of V is at most D.*

Proof. Let $X = \overline{V}$. By Proposition B.6, $X = \mathbf{V}(f_1^h, \ldots, f_r^h)$. Therefore, by Lemma B.11, X has at most D irreducible components, and by Corollary B.1 V has at most D irreducible components. □

The following proposition states that a degree of a hypersurface is at most the degree of any polynomial defining it.

Proposition B.13 ([17], Chapter 9, §4, Exercise 12)**.** *Let f be a non-constant homogenous polynomial. Let $H = \mathbf{V}(f_1)$. Then $deg(H) \leq deg(f)$.*

We will need the following definitions taken from [36].

Definition B.14. *Let $X, Y \subseteq \mathbb{P}^n$ be projective varieties. We say that X and Y* intersect properly *if*

$$dim(X \cap Y) = dim(X) + dim(Y) - n.$$

We quote (a corollary of) Bezout's theorem.

Theorem B.2 (Bezout-[36] Chapter 18, Theorem 18.4 and Corollary 18.5)**.** *Let $X, Y \subseteq \mathbb{P}^n$ be projective varieties of pure dimension intersecting properly. Then*

1. *$deg(X \cap Y) \leq deg(X) \cdot deg(Y)$.*
2. *The number of irreducible components of $X \cap Y$ is at most $deg(X) \cdot deg(Y)$.*

Claim B.14.1. *Let $X = \mathbf{V}(f_1, \ldots, f_r) \subseteq \mathbb{P}^n$ where $f_1, \ldots, f_r \in \mathbb{F}[x_0, \ldots, x_n]$ are non-constant homogenous polynomials. Assume that $dim(X) = n - r$. For $i = 1, \ldots, r$ let $H_i = \mathbf{V}(f_i)$ and $X_i = \mathbf{V}(f_1, \ldots, f_i) = H_1 \cap \ldots \cap H_i$. Then for all $i \in [r]$, X_i has pure dimension $n - i$.*

Proof. By Lemma B.9, all irreducible components of $\mathbf{V}(f_1, \ldots, f_i)$ have dimension at least $n - i$. Thus, it is enough to prove that $\mathbf{V}(f_1, \ldots, f_i)$ has (not necessarily pure) dimension $n - i$. We prove this by backwards induction on i. For $i = r$ it is already given that $dim(X) = dim(X_r) = n - r$. Assume the claim for $i + 1$ holds and assume for contradiction that $dim(X_i) \neq n - i$. Using Lemma B.9 it follows that $dim(X_i) > n - i$. Therefore, by Claim B.13.1 $dim(X_{i+1}) = dim(X_i \cap \mathbf{V}(f_{i+1})) > n - (i + 1)$, and this contradicts the induction hypothesis. □

We can now prove Lemma B.11.

Proof. (of Lemma B.11). We prove the claim by induction on r. For $r = 1$, it follows from Proposition B.13 that $deg(X) \leq deg(f_1) = d_1$. Assume the claim for $r - 1$. For $i = 1, \ldots, r$ denote $H_i = \mathbf{V}(f_i)$. Given $H_1, \ldots, H_r$, let $X_{r-1} = H_1 \cap \ldots \cap H_{r-1}$. We know from the induction hypothesis that

$$deg(X_{r-1}) \leq d_1 \cdots d_{r-1}.$$

From Claim B.14.1, X_{r-1} has pure dimension $n - (r - 1)$ and it follows that X_{r-1} and H_r intersect properly. Therefore, we can use Theorem B.2 and get

$$\deg(X) = deg(X_{r-1} \cap H_r) \leq deg(X_{r-1}) \cdot \deg(H_r) \leq d_1 \cdots d_r = D.$$

Similarly, from Theorem B.2 we get that the number of irreducible components of X is at most $deg(X_{r-1}) \cdot deg(H_r) \leq D$. □

B.7 Bombieri's Theorem

We quote an estimate of Bombieri [8] for character sums over projective curves and show that the estimate can be used also for affine curves. (Recall that a curve is a variety of dimension 1.) First we introduce some notation. Let $X \subseteq \mathbb{P}^n$ be a projective curve of degree D. Let $\mathbb{F}$ denote the algebraic closure of $\mathbb{F}_p$ for some prime p. Let $R \in \mathbb{F}_p(x_0, \ldots, x_n)$ be a homogenous rational function whose numerator and denominator both have degree d. Then, for any $x \in \mathbb{F}^{n+1}$ and non-zero $\lambda \in \mathbb{F}$ we have

$$R(\lambda \cdot x) = \frac{p(\lambda \cdot x)}{q(\lambda \cdot x)} = \frac{\lambda^d p(x)}{\lambda^d q(x)}$$

$$= \frac{p(x)}{q(x)} = R(x).$$

Therefore R is a well-defined function on points of $\mathbb{P}^n$ that are not poles of R, i.e., points $x \in \mathbb{P}^n$ such that $q(x) \neq 0$. We define

$$S_m(R, X) \triangleq \sum_{x \in X_m, q(x) \neq 0} e_p(\sigma R(x))$$

where X_m is the set of points of X with coordinates in $\mathbb{F}_{p^m}$, σ denotes the trace[4] from $\mathbb{F}_{p^m}$ to $\mathbb{F}_p$ and $e_p(x)$ is the function $e^{2\pi i x/p}$. Note that we sum only over non-poles of R.

Theorem B.3 (Theorem 6 in [8])**.** *Let R and X be as above. Let $\Gamma_1, \ldots, \Gamma_L$ be the irreducible components of X. Assume R is non-constant on Γ_i for $i = 1, \ldots, L$. If $d \cdot D < p$ then*

$$|S_m(R, X)| \leq 4dD^2 \cdot p^{m/2}.$$

For an affine curve $C \subseteq \mathbb{F}^n$ and a polynomial $g \in \mathbb{F}_p[x_1, \ldots, x_n]$ we define

$$S_m(g, C) \triangleq \sum_{(a_1, \ldots, a_m) \in C_m} e_p(\sigma g(a_1, \ldots, a_m))$$

where C_m denotes the set of points of C with coordinates in $\mathbb{F}_{p^m}$. We also denote $S(g, C) \triangleq S_1(g, C)$. We can now state and prove a version of Theorem B.3 for affine curves.

Theorem B.4. *Let $V \subseteq \mathbb{F}^n$ be an affine curve such that $V = \mathbf{V}(f_1, \ldots, f_{n-1})$ for polynomials $f_i \in \mathbb{F}[x_1, \ldots, x_n]$. Let $D = deg(f_1) \cdot \cdots \cdot deg(f_{n-1})$. Let $V_1, \ldots, V_L$ be the irreducible components of V. Let $g \in \mathbb{F}_p[x_1, \ldots, x_n]$ be a polynomial of degree d that is non-constant on some V_i. Let C be the union*

[4] See [39] for a definition of the trace function. For the case $m = 1$, which is the only one we will use, the trace is simply the identity function.

of the irreducible components V_i such that g is non-constant on V_i. Assume that $d \cdot D < p$. We have

$$S_m(g, C) \leq 4dD^2 \cdot p^{m/2}.$$

In particular,

$$S(g, C) \leq 4dD^2 \cdot p^{1/2}.$$

Proof. We identify g with a homogenous rational function R defined as

$$R(x_0, x) = \frac{g^h(x_0, x)}{x_0^d}.$$

Note that for every $a \in \mathbb{F}^n$ $R(1, a) = g(a)$.
Denote $X = \overline{C}$.

Claim B.14.2.

$$S_m(g, C) = S_m(R, X).$$

Proof. Using Proposition B.4 X consists precisely of the points $(1, a)$ where $a \in C$ and, possibly, some 'points at infinity', i.e., points of the form $(0, a)$ for $a \in \mathbb{F}^n$. Since R has poles on all points of the form $(0, a)$ and $R(1, a) = g(a)$ for all $x \in \mathbb{F}^n$, we get that summing R over all non-poles in X is exactly the same as summing g over all of C. In particular, summing R over all non-poles in X_m is exactly the same as summing g over all of C_m. That is,

$$S_m(g, C) = S_m(R, X).$$

□

We now want to bound $S_m(R, X)$ using Theorem B.3. Note that both the numerator and the denominator of R are homogenous of degree exactly d, so R is suitable for the theorem. We need to show that X is a projective variety of dimension 1 such that R is non-constant on any of its irreducible components: Recall that the irreducible components of C are simply a subset of $V_1, \ldots, V_L$. Assume w.l.o.g. that $C = V_1 \cup \ldots \cup V_r$. Using Corollary B.1, it is clear that if g is non-constant on the irreducible components $V_1, \ldots, V_r$ of C, then R is non-constant on the irreducible components $\overline{V}_1, \ldots, \overline{V}_r$ of X. By Proposition B.5 and Corollary B.1 $dim(\overline{V}) = 1$ and $\overline{V}_1, \ldots, \overline{V}_L$ are the irreducible components of $\overline{V}$. By Proposition B.6, $\overline{V} = \mathbf{V}(f_1^h, \ldots, f_{n-1}^h)$, and therefore by Claim B.14.1 for every i $\overline{V}_i$ has dimension 1. It follows that $X = \overline{V}_1 \cup \ldots \cup \overline{V}_r$ has dimension 1.

Finally, we need to bound the degree of X. By Lemma B.11 $deg(\overline{V}) \leq D$. Since the degree of a projective variety is the sum of degrees of its irreducible components (see [36], Chapter 18), $deg(X) \leq D$.

Therefore, we can use Theorem B.3. We get

$$|S_m(g, C)| = |S_m(R, X)| \leq 4dD^2 \cdot p^{m/2}.$$

□

Bibliography

[1] N. Alon. Tools from higher algebra. In *R. L. Graham & M. Grotschel & L. Lovasz (eds.), Handbook of Combinatorics, Elsevier and The MIT Press*, volume 2. 1995.

[2] N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple constructions of almost k-wise independent random variables. In *Proceedings of the 31st Annual IEEE Symposium on Foundations of Computer Science*, volume II, pages 544–553, 1990.

[3] B. Barak, R. Impagliazzo, and A. Wigderson. Extracting randomness using few independent sources. *SIAM J. Comput*, 36(4):1095–1118, 2006.

[4] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson. Simulating independence: New constructions of condensers, Ramsay graphs, dispersers, and extractors. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 1–10, 2005.

[5] 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 38th Annual ACM Symposium on Theory of Computing*, pages 671–680, 2006.

[6] M. Bellare and J. Rompel. Randomness-efficient oblivious sampling. FOCS 1994.

[7] M. Ben-Or and N. Linial. Collective coin flipping. *ADVCR: Advances in Computing Research*, 5:91–115, 1989.

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

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

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

[11] V. Boyko. On the security properties of OAEP as an all-or-nothing transform. In *Proc. 19th International Advances in Cryptology Conference – CRYPTO '99*, pages 503–518, 1999.

[12] R. Canetti, Y. Dodis, S. Halevi, E. Kushilevitz, and A. Sahai. Exposure-resilient functions and all-or-nothing transforms. *Lecture Notes in Computer Science*, 1807, 2000.

[13] I. L. Carter and M. N. Wegman. Universal classes of hash functions. In *Proceedings of the 9th Annual ACM Symposium on Theory of Computing*, pages 106–112, 1977.

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

[15] B. Chor, O. Goldreich, J. Hastad, J. Friedman, S. Rudich, and R. Smolensky. The bit extraction problem or t-resilient functions. In *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science*, pages 396–407, 1985.

[16] A. Cohen and A. Wigderson. Dispersers, deterministic amplification and weak random sources. In *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, pages 14–25, 1989.

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

[18] Y. Dodis. *Exposure-Resilient Cryptography.* PhD thesis, Department of Electrical Engineering and Computer Science, MIT, August 2000.

[19] Y. Dodis, A. Elbaz, R. Oliveira, and R. Raz. Improved randomness extraction from two independent sources. In *RANDOM: International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 334–344, 2004.

[20] Y. Dodis, A. Sahai, and A. Smith. On perfect and adaptive security in exposure-resilient cryptography. *Lecture Notes in Computer Science*, 2045, 2001.

[21] R. Ehrenborg and G. Rota. Apolarity and canonical forms for homogeneous polynomials. *Europ. J. Combinatorics*, 14:157–181, 1993.

[22] A. Elbaz. Improved constructions for extracting quasi-random bits from sources of weak randomness. *M.Sc. Thesis, Weizmann Institute*, 2003.

[23] S. Even, O. Goldreich, M. Luby, N. Nisan, and B. Velickovic. Efficient approximation of product distributions. *RSA: Random Structures & Algorithms*, 13, 1998.

[24] A. Fiat and M. Naor. Implicit O(1) probe search. *SICOMP: SIAM Journal on Computing*, 22, 1993.

[25] 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, 2005.

[26] A. Gabizon, R. Raz, and R. Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. *SICOMP: SIAM Journal on Computing*, 36(4):1072–1094, 2006.

[27] A. Gabizon and R. Shaltiel. Increasing the output length of zero-error dispersers. In *APPROX-RANDOM*, pages 430–443, 2008.

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

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

[30] O. Goldreich. A sample of samplers – A computational perspective on sampling (survey). In *ECCCTR: Electronic Colloquium on Computational Complexity, technical reports*, 1997a.

[31] O. Goldreich. A sample of samplers – A computational perspective on sampling (survey). In *ECCCTR: Electronic Colloquium on Computational Complexity, Technical Reports*, 1997b.

[32] R. L. Graham, B. L. Rothschild, and J. H. Spencer. *Ramsey Theory*. Wiley, 1980.

[33] R. L. Graham and J. H. Spencer. A constructive solution to a tournament problem. *Canad. Math. Bull.*, 14:45–48, 1971.

[34] V. Guruswami, C. Umans, and S. P. Vadhan. Unbalanced expanders and randomness extractors from Parvaresh-Vardy codes. In *Proceedings of the 22nd Annual IEEE Conference on Computational Complexity*, pages 96–108, 2007.

[35] A. Hales and R. Jewett. Regularity and positional games. *Trans. Amer. Math. Soc.*, 106:222–229, 1963.

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

[37] J. Kamp and D. Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. *SIAM J. Comput*, 36(5):1231–1247, 2007.

[38] N. Kayal. The complexity of the annihilating polynomial. *Manuscript*, 2007.

[39] R. Lide and H. Niederreiter. *Finite fields*. Cambridge University Press, New York, NY, USA, 1997.

[40] R. Lipton and N. Vishnoi. Manuscript. 2004.

[41] L. Lovasz. *Combinatorial Problems and Exercises*. North-Holland, Amsterdam, 1979.

[42] C. Lu, O. Reingold, S. Vadhan, and A. Wigderson. Extractors: Optimal up to constant factors. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pages 602–611, 2003.

[43] M. S. L'vov. Calculation of invariants of programs interpreted over an integrality domain. *Kibernetika*, (4):23–28, 1984.

[44] J. Naor and M. Naor. Small-bias probability spaces: Efficient constructions and applications. In *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing*, pages 213–223, 1990.

[45] M. Naor, A. Nussboim, and E. Tromer. Efficiently constructible huge graphs that preserve first order properties of random graphs. In *TCC*, pages 66–85, 2005.

[46] N. Nisan and A. Ta-Shma. Extracting randomness: A survey and new constructions. *Journal of Computer and System Sciences*, 58, 1999.

[47] N. Nisan and D. Zuckerman. More deterministic simulation in logspace. In *Proceedings of the Twenty-Fifth Annual ACM Symposium on Theory of Computing*, pages 235–244, New York, NY, USA, 1993. ACM Press.

[48] N. Nisan and D. Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, 1996.

[49] J. Radhakrishnan and A. Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13(1):2–24, 2000.

[50] A. Rao. Extractors for a constant number of polynomially small min-entropy independent sources. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 497–506, 2006.

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

[52] A. Rao. Extractors for low weight affine sources. *Manuscript*, 2008.

[53] R. Raz. Extractors with weak random seeds. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 11–20, 2005.

[54] R. Raz, O. Reingold, and S. Vadhan. Error reduction for extractors. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, 1999.

[55] R. Raz, O. Reingold, and S. Vadhan. Extracting all the randomness and reducing the error in Trevisan's extractors. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 149–158, 1999.

[56] O. Reingold, R. Shaltiel, and A. Wigderson. Extracting randomness via repeated condensing. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, 2000.

[57] R. Rivest. All-or-nothing encryption and the package transform. In *Fast Software Encryption: 4th International Workshop, FSE*, volume 1267 of *Lecture Notes in Computer Science*, 1997.

[58] M. Santha and U. V. Vazirani. Generating quasi-random sequences from semi-random sources. *Journal of Computer and System Sciences*, 33:75–87, 1986.

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

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

[61] I. R. Shafarevich. *Basic algebraic geometry.* Springer, 1994.

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

[63] 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, 2006.

[64] R. Shaltiel and C. Umans. Simple extractors for all min-entropies and a new pseudo-random generator. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, 2001.

[65] A. Ta-Shma, D. Zuckerman, and S. Safra. Extractors from Reed-Muller codes. In IEEE, editor, *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 638–647, 2001. IEEE Computer Society Press.

[66] L. Trevisan. Construction of extractors using pseudorandom generators. In *Proceedings of the 31st ACM Symposium on Theory of Computing*, 1999.

[67] L. Trevisan and S. Vadhan. Extracting randomness from samplable distributions. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, 2000a.

[68] L. Trevisan and S. Vadhan. Extracting randomness from samplable distributions. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, pages 32–42, 2000b.

[69] S. Vadhan. On constructing locally computable extractors and cryptosystems in the bounded storage model, November 01 2002.

[70] U. Vazirani. Strong communication complexity or generating quasi-random sequences from two communicating semi-random sources. *Combinatorica*, 7:375–392, 1987.

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

[72] A. Weil. On some exponential sums. In *Proc. Nat. Acad. Sci. USA*, volume 34, pages 204–207, 1948.

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

[74] M. Prabhakaran Y. Dodis, S. J. Ong and A. Sahai. On the (im)possibility of cryptography with imperfect randomness. In *FOCS 2004*, 2004.

[75] A. C.-C. Yao. Should tables be sorted? *J. ACM*, 28(3):615–628, 1981.

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