Sub-Constant Error Low Degree Test of Almost Linear Size

Dana Moshkovitz * Ran Raz †

October 1, 2007

Abstract

Given (the table of) a function $f : \mathbb{F}^m \to \mathbb{F}$ over a finite field $\mathbb{F}$, a *low degree tester* tests its agreement with an m-variate polynomial of total degree at most d over $\mathbb{F}$. The tester is usually given access to an oracle $\mathcal{A}$ providing the *supposed* restrictions of f to affine subspaces of constant dimension (e.g., lines, planes, etc.). The tester makes very few (probabilistic) queries to f and to $\mathcal{A}$ (say, one query to f and one query to $\mathcal{A}$), and decides whether to accept or reject based on the replies.

We wish to minimize two parameters of the tester: its *error* and its *size*. The *error* bounds the probability that the tester accepts although the function is far from a low degree polynomial. The *size* is the number of bits required to write the oracle replies on all possible tester's queries.

Low degree testing is a central ingredient in most constructions of probabilistically checkable proofs (PCPs). The error of the low degree tester is related to the error of the PCP and its size is related to the size of the PCP.

We design and analyze new low degree testers that have both *sub-constant error* $o(1)$ and *almost-linear size* $n^{1+o(1)}$ (where $n = |\mathbb{F}|^m$). Previous constructions of *sub-constant error* testers had *polynomial size* (works by Arora and Sudan [3] and by Raz and Safra [17]). These testers enabled the construction of PCPs with *sub-constant error*, but *polynomial size* (see the work by Dinur *et al* [9]). Previous constructions of *almost-linear size* testers obtained only *constant error* (Ben-Sasson, Sudan, Vadhan and Wigderson [7]). These testers were used to construct *almost-linear size* PCPs with *constant error* (see Ben-Sasson *et al* [5]). The testers we present in this work enabled the construction of PCPs with both *sub-constant error* and *almost-linear size* (Moshkovitz and Raz [15]).

*dana.moshkovitz@weizmann.ac.il. Department of Computer Science and Applied Mathematics, The Weizmann Institute, Rehovot, Israel. Research supported by an Adams Fellowship of the Israel Academy of Sciences and Humanities and by an ISF grant.

†ran.raz@weizmann.ac.il. Department of Computer Science and Applied Mathematics, The Weizmann Institute, Rehovot, Israel. Research supported by ISF and BSF grants.

1 Introduction

1.1 Low Degree Testing

Let $\mathbb{F}$ be a finite field, let m and d be two positive integers. [A particular setting of parameters to have in mind is the one used in constructions of Probabilistically Checkable Proofs: a large field $\mathbb{F}$, a smaller m and a fairly large d that satisfy $m^{O(1)}d \leq o(|\mathbb{F}|)$].

Define $\mathcal{P}$ to be the set of all m-variate polynomials of total degree at most d over $\mathbb{F}$. The *agreement* of a function $f : \mathbb{F}^m \to \mathbb{F}$ with a low degree polynomial is

$$\mathsf{agr}(f, \mathcal{P}) \stackrel{def}{=} \max_{Q \in \mathcal{P}} \left\{ \Pr_{\vec{x} \in \mathbb{F}^m} [f(\vec{x}) = Q(\vec{x})] \right\}$$

Note that $\mathsf{agr}(f, \mathcal{P})$ is simply $1 - \Delta(f, \mathcal{P})$, where Δ denotes the (normalized) Hamming distance between functions that are given by their tables.

A *low degree tester* is a probabilistic procedure M that is meant to check the agreement of a function f with a low degree polynomial by making as few *queries* to f as possible. If $f \in \mathcal{P}$, M should always accept, while if f is far from $\mathcal{P}$ (i.e., $\mathsf{agr}(f, \mathcal{P})$ is small) M should reject with significant probability.

It is easy to see that when having oracle access only to f, any low degree tester must make more than d queries. To break this degree barrier, the low degree tester is usually given access to an additional oracle $\mathcal{A}$ providing the *supposed* restrictions of f to affine subspaces of constant dimension (e.g., lines, planes, etc.). We assume, without loss of generality, that these restrictions in themselves are polynomials of total degree at most d over the subspaces.

The tester is required to satisfy:

- *Completeness:* If $f \in \mathcal{P}$, then there is an oracle $\mathcal{A}$ that makes the tester accept with probability 1.
- *Soundness:* If $\mathsf{agr}(f, \mathcal{P})$ is small, then for any oracle $\mathcal{A}$, the tester may accept only with a small probability.

Rubinfeld and Sudan [18] designed the Line vs. Point tester that makes only two probabilistic queries. This tester picks independently at random a line l in $\mathbb{F}^m$ and a point $\vec{x} \in l$, queries the oracle $\mathcal{A}$ for the (supposed) restriction of f to l (which is simply a univariate polynomial of degree at most d over $\mathbb{F}$), queries f at $\vec{x}$, and checks whether the two restrictions are consistent on $\vec{x}$, i.e., $\mathcal{A}(l)(\vec{x}) = f(\vec{x})$.

The importance of low degree testers comes from the key role they play in the construction of *Probabilistically Checkable Proofs* (PCPs), which are proofs for NP statements that can be probabilistically verified by making only a constant number of queries to the proof [4, 10, 2, 1]. This motivated further improvements to low degree testing.

Specifically, the following parameters were of interest:

1. **Queries:** How many *queries* does the tester make?
2. **Error:** How sound is the tester?
3. **Size:** How many bits are needed to write the oracle replies on all possible queries?

Henceforth, the number of queries will always be 2. The two other parameters are discussed next.

1.1.1 Error

To prove that a low degree tester is sound, most results address contrapositive arguments of the following type: assume that the tester accepts with probability $\gamma \geq \gamma_0$ and show the existence of a low degree polynomial that agrees with f on at least $\approx \gamma$ of the points. In this case, we say that γ_0 bounds the *error* of the tester, since the probability that the tester accepts although the function is very far from a low degree polynomial is at most γ_0.

The first analyses of the Line vs. Point tester [18, 2, 12] only showed that the error of the tester is bounded away from 1. The error can be amplified to any *constant*, by a constant number of repetitions. Nevertheless, to keep the total number of queries constant, one cannot perform more than a constant number of repetitions.

Only a later, more careful, inspection [3, 17] revealed that there are low degree testers with a *sub-constant error*. Specifically, [3, 17] proved claims of the following type for various low degree testers: there exist (large enough) constants $C \geq 1$, $a, b \geq 0$, and a (small enough) constant $0 < c \leq 1$, such that the error is at most $Cm^a d^b/|\mathbb{F}|^c$. In other words, the error can be made arbitrarily small by taking m and d to be small enough with respect to $|\mathbb{F}|$. The number of queries remains 2.

Arora and Sudan [3] proved that the error of the Line vs. Point tester is in fact sub-constant. Their proof was algebraic in nature. Raz and Safra [17] proved a sub-constant error for a slightly different tester, considering planes that intersect on a line, or a plane and a point within it. Their proof was more combinatorial in nature. The two proofs led to the construction of PCPs with *sub-constant error* [3, 17, 9].

1.1.2 Size

Let us represent the set of honest oracles by a code. That is, for every polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ of degree at most d, we have a codeword. The codeword has an entry for every affine subspace s that the tester may query. This entry contains the oracle's reply when it is queried regarding s, i.e., the restriction of Q to s. The *size* of a tester is the length (in bits) of a codeword.

For instance, the size of Rubinfeld and Sudan's Line vs. Point tester [18] is roughly $|\mathbb{F}|^{2m}(d+1)\log|\mathbb{F}|$: For every line (defined by two points), the oracle should provide a univariate polynomial of degree at most d over $\mathbb{F}$. The size of a tester is measured with respect to $n = |\mathbb{F}^m|$. The size of the Line vs. Point tester [18] is quadratic $n^{2+o(1)}$.

Alternatively, we refer to the *randomness* of the tester, which is the amount of random bits that the tester requires. Note that the size of a tester that uses r random bits to make q queries to a proof over alphabet Σ is bounded by $2^r \cdot q \log|\Sigma|$. Thus, when the number of queries q is constant and the alphabet Σ is relatively small, 2^r is a good estimate on the size.

For instance, to pick a random line and a random point within it, we merely have to pick a random point $\vec{x} \in \mathbb{F}^m$ and a random direction $\vec{y} \in \mathbb{F}^m$. The line is $\vec{x} + t \cdot \vec{y}$ for $t \in \mathbb{F}$. Hence, the randomness of the Line vs. Point tester [18] is $2m \log|\mathbb{F}| = \log(|\mathbb{F}|^{2m})$.

The size of a tester is related to the size of probabilistically checkable proofs and locally testable codes constructed using it. Hence, Goldreich and Sudan [13] suggested to improve the Line vs. Point tester by considering a relatively small subset of lines (instead of all lines). Goldreich and Sudan achieved *non-explicit* constant error tester of *almost-linear* size $n^{1+o(1)}$, instead of quadratic size $n^{2+o(1)}$.

Shortly afterwards, Ben-Sasson, Sudan, Vadhan and Wigderson [7] gave an explicit construction of a constant error Line vs. Point tester of almost-linear size. Their idea was to choose a

line by picking a uniformly distributed point over $\mathbb{F}^m$ (as before), and a direction that is uniformly distributed over a small ϵ-biased set $S \subseteq \mathbb{F}^m$. They showed that the error of this tester is bounded away from 1. Unfortunately, their analysis is inherently only able to show error larger than $\frac{1}{2}$. It is possible that their tester has smaller error, but proving it would require a substantially different analysis.

The work of [13, 7] gave rise to explicit constructions of *almost-linear size PCP*s with *constant error* [13, 7, 5]. The recent work of Dinur [8] also constructs *almost-linear size PCP*s with *constant error*, based on the *PCP* theorem of [2, 1] and the work of Ben-Sasson and Sudan [6]. Both use low degree testers with *constant error*. Dinur's work [8] also gives new constructions of *PCP*s without low degree testers. However, at this point, these constructions achieve neither sub-constant error nor almost-linear size.

1.2 Our Contribution: Randomness-Efficient Sub-Constant Error Testers

We design and analyze two low degree testers that have both *sub-constant error* and *almost-linear size*. Subsequent to this work and using it, we showed a construction of a *PCP* with both *sub-constant error* and *almost-linear size* [15].

Before we present our testers, let us revisit the construction of Ben-Sasson, Sudan, Vadhan and Wigderson [7] for constant error, and point out the most severe difficulty one encounters when trying to argue it has error smaller than $\frac{1}{2}$. The reader who is not familiar with the work of Ben-Sasson *et al* may skip this, and move directly to the text after Remark 1.1.

Assume a Line vs. Point tester that only inspects lines whose directions are taken from a small *random* set $S \subseteq \mathbb{F}^m$. Recall that Ben-Sasson *et al* used a small ϵ-biased set because of its *pseudo-random* properties [7].

Consider two linearly independent directions $\vec{y}_1, \vec{y}_2 \in S$. With high probability, the set S does not contain any additional vector from the linear span of $\vec{y}_1$ and $\vec{y}_2$ (since the fractional size of the linear span is merely $|\mathbb{F}|^2 / |\mathbb{F}|^m$). Thus, the only lines inside this two-dimensional linear subspace that get inspected by the tester are those that are parallel to either $\vec{y}_1$ or $\vec{y}_2$. It is known by a lemma of Polishchuk and Spielman [16] that if the acceptance probability of the Line vs. Point test in this setting approaches 1, then there exists a low degree polynomial for the entire subspace that agrees with almost all lines. However, it may be the case that the acceptance probability is as large as $\frac{1}{2}$, although the agreement of those lines with any low degree polynomial is very small.

Let us demonstrate this.

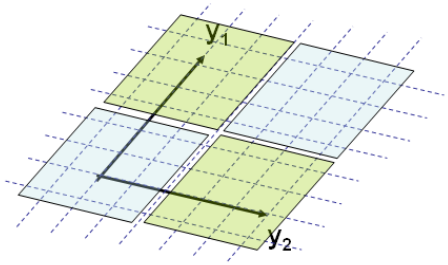

Figure 1: Acceptance probability $\frac{1}{2}$ inside a plane does not necessarily imply agreement with a low degree polynomial.

Note that each of the $|\mathbb{F}|^2$ points on the plane spanned by $\vec{y}_1$ and $\vec{y}_2$ can be uniquely repre-

sented as $\alpha_1\vec{y}_1 + \alpha_2\vec{y}_2$, where $\vec{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{F}^2$.

Let $d \ll d' \ll \frac{|\mathbb{F}|}{4}$. Define polynomials $C, R \in \mathbb{F}[\alpha_1, \alpha_2]$ (for *columns* and *rows* respectively) as follows. The degree of C in α_1 is d and the degree in α_2 is d'. The degree of R in α_1 is d' and the degree in α_2 is d. Let the $|\mathbb{F}|$ *columns*, i.e., lines parallel to $\vec{y}_1$, agree with C. Let the $|\mathbb{F}|$ *rows*, i.e., lines parallel to $\vec{y}_2$, agree with R. Note that all lines, columns and rows, are assigned (univariate) polynomials of degree at most d.

Let points in the dark region agree with columns and points in the bright region agree with rows. Both R and C are polynomials of degree *at least* $d' \gg d$ for the plane that agree with at least $\frac{1}{2}$ of the points. However, no polynomial of degree *at most* d for the plane agrees with a fraction of more than $\frac{4d'}{|\mathbb{F}|} = o(1)$ of the points. On the other hand, the acceptance probability of the Line vs. Point tester on this plane (where all lines are assigned polynomials of degree at most d) is at least $\frac{1}{2}$.

Remark 1.1. *One may consider other low degree testers, like the* Line vs. Line *tester, in order to solve the problem we described. However, it is not known whether or not the Line vs. Line tester (on the plane) with lines parallel to the axes, gives a probability of error lower than* $\frac{1}{2}$.

We manage to overcome this difficulty by considering sets that are *not pseudo-random*. Our key idea is to consider a subfield $\mathbb{H} \subseteq \mathbb{F}$, and generate subspaces by picking directions uniformly over $\mathbb{H}^m$, instead of over $\mathbb{F}^m$. Note that this eliminates the problem we described: for every two $\vec{y}_1, \vec{y}_2 \in \mathbb{H}^m$, for every two scalars $\alpha_1, \alpha_2 \in \mathbb{H}$, we have $\alpha_1\vec{y}_1 + \alpha_2\vec{y}_2 \in \mathbb{H}^m$.

Moreover, the field structure of $\mathbb{H}$ allows us to use the combinatorial approach of Raz and Safra [17], and, more importantly, it allows us to use induction: the structure of the problem when restricted to affine subspaces of dimension $k \leq m$ is the same as its structure in $\mathbb{F}^m$.

As in the analysis of Raz and Safra [17], we abandon the Line vs. Point test, and address subspaces of dimension larger than 1, rather than lines. Specifically, given access to f and to an oracle $\mathcal{A}$, our *Randomness-Efficient Plane vs. Point* tester chooses a plane and a point within it and checks that they are consistent:

1. Pick uniformly and independently at random $\vec{z} \in \mathbb{F}^m$, $\vec{y}_1, \vec{y}_2 \in \mathbb{H}^m$.
2. Accept if either $\vec{y}_1, \vec{y}_2$ are linearly dependent, or if the plane p through $\vec{z}$ in directions $\vec{y}_1, \vec{y}_2$ satisfies $\mathcal{A}(p)(\vec{z}) = f(\vec{z})$.

Figure 2: Randomness-Efficient Plane vs. Point Tester

Note that the same plane p goes through many points $\vec{z} \in \mathbb{F}^m$ and in many directions $\vec{y}_1, \vec{y}_2 \in \mathbb{H}^m$. However, the oracle's reply $\mathcal{A}(p)$ depends on the plane p, and not on its representation given by $\vec{z}$ and $\vec{y}_1, \vec{y}_2$.

For $\mathbb{H} = \mathbb{F}$, the Randomness-Efficient Plane vs. Point Tester is exactly the Plane vs. Point tester of Raz and Safra [17]. However, in our work the more interesting case is $|\mathbb{H}| \leq |\mathbb{F}|^{o(1)}$. In this case, the tester requires only $m \log |\mathbb{F}| + 2m \log |\mathbb{H}| = m \log |\mathbb{F}| \, (1 + o(1))$ bits of randomness. This corresponds to an almost linear size $n^{1+o(1)}$ (recall that $n = |\mathbb{F}^m|$). The tester is randomness efficient in comparison to all known testers with *sub-constant error*, such as the tester of Arora and Sudan [3] that requires $2m \log |\mathbb{F}|$ bits of randomness and the tester of Raz and Safra [17] that requires $3m \log |\mathbb{F}|$ bits of randomness. As to testers with *constant error*: that of Ben-Sasson, Sudan, Vadhan and Wigderson [7] requires $m \log |\mathbb{F}| + O(\log \log |\mathbb{F}^m|)$ bits of randomness, which is (usually) less than the randomness of our tester, but the difference is only in the dependence of the *low order term* in m.

The tester is clearly *complete*, namely, if there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ of degree at most d, such that for every $\vec{x} \in \mathbb{F}^m$, $f(\vec{x}) = Q(\vec{x})$, and for every affine subspace s, the oracle $\mathcal{A}$ replies $\mathcal{A}(s) = Q_{|s}$, then the tester accepts with probability 1. We show that the tester is also *sound*: if the tester accepts with probability γ then f agrees with a polynomial of total degree at most md on a fraction of at least $\gamma - \varepsilon$ of the points in $\mathbb{F}^m$, where $\varepsilon \leq const \cdot m \left(\sqrt[8]{\frac{1}{|\mathbb{H}|}} + \sqrt[4]{\frac{md}{|\mathbb{F}|}} \right)$. Note that the analysis works for any acceptance probability γ. In particular, this means that when γ is significantly larger than ε, say $\gamma \geq 100\varepsilon$, f agrees with a polynomial of total degree at most md on at least $\approx \gamma$ of the points. [Even if $\mathbb{H} = \mathbb{F}$, the constants 4 and 8 in the error expression appear to improve on the results of [3, 17], where unspecified constants were given].

The downside of the Randomness-Efficient Plane vs. Point tester is that it only allows us to argue something about the agreement of the oracle with a polynomial of degree md, rather than d. Hence, we design another tester that has essentially the same parameters, but ensures agreement with a polynomial of degree at most d.

The additional consideration that comes into play when designing the new tester is *degree preservation*. We want the total degree of a polynomial not to decrease when restricted to most of the subspaces queried by the tester. We achieve this by picking one of the *directions* for the subspace (rather than the base-point) uniformly from $\mathbb{F}^m$. In order to keep the size almost linear, this tester considers linear subspaces (i.e., affine subspaces through the origin), rather than general affine subspaces. A related technique was previously used by [7].

Specifically, given access to f and to an oracle $\mathcal{A}$, the *Randomness-Efficient Subspace vs. Point* tester chooses a three dimensional subspace and a point within it and checks that they are consistent:

1. Pick uniformly and independently at random $\vec{z} \in \mathbb{F}^m$, $\vec{y}_1, \vec{y}_2 \in \mathbb{H}^m$.
2. Accept if either $\vec{z}, \vec{y}_1, \vec{y}_2$ are linearly dependent, or if the linear subspace s spanned by $\vec{z}, \vec{y}_1, \vec{y}_2$ satisfies $\mathcal{A}(s)(\vec{z}) = f(\vec{z})$.

Figure 3: Randomness-Efficient Subspace vs. Point Tester

This tester uses the same number of random bits as the Randomness-Efficient Plane vs. Point tester $m \log |\mathbb{F}| + 2m \log |\mathbb{H}|$, and its size is only slightly larger (as the answer size is larger: the oracle should provide polynomials over three-dimensional subspaces rather than two-dimensional subspaces). For this small price, we manage to prove a stronger soundness claim: if the Randomness-Efficient Subspace vs. Point tester accepts with probability γ, then f agrees with a polynomial of total degree at most d (rather than md) on a fraction of at least $\gamma - \varepsilon$ of the points in $\mathbb{F}^m$, where $\varepsilon \leq const \cdot m \left(\sqrt[8]{\frac{1}{|\mathbb{H}|}} + \sqrt[4]{\frac{md}{|\mathbb{F}|}} \right)$. This follows rather easily from the soundness of the Randomness-Efficient Plane vs. Point tester together with an argument showing that the degree of the recovered polynomials must in fact be at most d.

There is a tradeoff between the size of the testers and their error. To make the size as small as possible, one wishes to minimize $|\mathbb{H}|$. In particular, to get an almost-linear size, one needs to take $|\mathbb{H}| \leq |\mathbb{F}|^{o(1)}$. On the other hand, to make the error as small as possible, one wishes to maximize $|\mathbb{H}|$. In particular, to get a sub-constant error, one needs to take $|\mathbb{H}| \geq \omega(m^8)$.

All finite fields are isomorphic to $GF(p^k)$ for a prime p and a natural number k. All subfields of $GF(p^k)$ are isomorphic to $GF(p^r)$ for $r|k$. For a wide family of finite fields $GF(p^k)$ there are subfields of suitable sizes (see [14, 11] for analysis of the distribution of k's with suitable divisors). Though, indeed, not every finite field is such. We wish to emphasize that in the

settings that interest us (e.g., construction of PCPs), *we get to choose the field*. For instance, we can take $\mathbb{F} = GF(2^{r_1 \cdot r_2})$ for appropriate r_1, r_2.

1.3 Sampling

A basic step in our proof is the analysis of the sampling properties of affine subspaces with directions over a subfield. This analysis may be of independent interest.

By *sampling* we refer to assertions of the following nature: if one colors a large enough fraction of the points in $\mathbb{F}^m$ green then a subspace (e.g., a line) picked at random is likely to hit the green points in almost their true fraction.

First, let us consider the non-randomness-efficient setting. For instance, consider choosing a line by picking a point and a direction independently at random from $\mathbb{F}^m$. The indicator variables "is the i'th point on the line green?" for $i = 1, \ldots, |\mathbb{F}|$ are *pairwise independent*. Thus, one can easily bound the variance of the number of green points on a line. This yields a sampling property by Chebyshev's inequality (see, e.g., [3]).

In the randomness-efficient setting, more subtle arguments are needed. For instance, consider the work of Ben-Sasson, Sudan, Vadhan and Wigderson [7]. They use an ϵ-biased set $S \subseteq \mathbb{F}^m$, and choose a line by independently picking a uniformly distributed base-point in $\mathbb{F}^m$ and a uniformly distributed direction in S. They show that *almost pairwise independence* still holds, and this allows them to bound the variance, by bounding the covariances.

Our set of directions is $\mathbb{H}^m$, which does not have a small bias (when $\mathbb{H} \subsetneq \mathbb{F}$). Nevertheless, we are still able to prove a sampling property. We observe that we can directly bound the variance of the number of green points on a line by analyzing the convolution of two relatively simple functions. We do this by means of Fourier analysis. The difference between the previous approaches and our approach is that instead of giving one bound for the probability that two points $i \neq j$ on a line are green *for every* $i \neq j$, we directly bound the *average* probability over all pairs $i \neq j$.

The extension to higher dimensional subspaces is a relatively simple consequence of the analysis for lines.

1.4 More Randomness-Efficient Line Samplers

Ariel Gabizon has noted that our analysis implies numerous randomness-efficient line samplers.

Recall that the set $\mathbb{H}^m$ for a subfield $\mathbb{H} \subseteq \mathbb{F}$ – in addition to implying a sampling property – also has an algebraic structure that is essential for our analysis. However, if one is only interested in the sampling property, more randomness-efficient constructions may be obtained.

Jointly with Ariel we arrived at the following corollaries to our analysis.

Direct product construction. Our sampling lemma holds for any field $\mathbb{F} = GF(p^k)$ and a subset of it $H \subseteq \mathbb{F}$ (not necessarily a subfield). Formally:

Corollary 1.2. *For any subset $A \subseteq \mathbb{F}^m$ of density $\mu = |A| / |\mathbb{F}^m|$, for any $\varepsilon > 0$,*

$$\Pr_{\vec{x} \in \mathbb{F}^m, \vec{y} \in H^m} \left[\left| \frac{\left| l_{\vec{x},\vec{y}} \cap A \right|}{\left| l_{\vec{x},\vec{y}} \right|} - \mu \right| \geq \varepsilon \right] \leq \frac{1}{|H|} \cdot \frac{\mu}{\varepsilon^2}$$

where $l_{\vec{x},\vec{y}} \stackrel{def}{=} \{\vec{x} + t \cdot \vec{y} \mid t \in \mathbb{F}\}$.

Linear code construction. Assume a linear code of length k, dimension m and relative distance $1-\delta$ over alphabet $\mathbb{F} = GF(p)$, given by its generating matrix

$$\begin{pmatrix} - & \vec{y}_1 & - \\ & \vdots & \\ - & \vec{y}_k & - \end{pmatrix}$$

Let $S = \{\vec{y}_1, \ldots, \vec{y}_k\} \subseteq \mathbb{F}^m$ be the set of rows of the generating matrix. Then, our analysis actually implies the following:

Corollary 1.3. *For any subset $A \subseteq \mathbb{F}^m$ of density $\mu = |A| / |\mathbb{F}^m|$, for any $\varepsilon > 0$,*

$$\Pr_{\vec{x}\in\mathbb{F}^m, \vec{y}\in S}\left[\left|\frac{\left|l_{\vec{x},\vec{y}} \cap A\right|}{\left|l_{\vec{x},\vec{y}}\right|} - \mu\right| \geq \varepsilon\right] \leq \delta \cdot \frac{\mu}{\varepsilon^2}$$

where $l_{\vec{x},\vec{y}} \stackrel{def}{=} \{\vec{x} + t \cdot \vec{y} \mid t \in \mathbb{F}\}$.

Note that every $S \subseteq \mathbb{F}^m$ that is ϵ-biased forms a generating matrix of a linear code with distance $1 - (\frac{1}{|\mathbb{F}|} + \epsilon \cdot \frac{|\mathbb{F}|-1}{|\mathbb{F}|})$. Yet, the converse does not necessarily hold, and the corollary is a strengthening of the sampling lemma of [7] for the case $\mathbb{F} = GF(p)$.

A randomness-efficient line sampler can be constructed by using an efficient linear code. For instance, we can use the Reed-Solomon code that corresponds to $S = \left\{(1, t, t^2, \ldots, t^{m-1}) \mid t \in \mathbb{F}\right\}$. This code has relative distance $1-\delta$ for $\delta = \frac{m-1}{|\mathbb{F}|}$. It gives a line sampler that has randomness complexity $(m+1)\log|\mathbb{F}|$ and query complexity $|\mathbb{F}|$.

1.5 Proof Outline

We first prove the soundness of the Randomness-Efficient Plane vs. Point tester, and then deduce the soundness of the Randomness-Efficient Subspace vs. Point tester from it. For the purpose of this outline we only consider the first. Assume that the Randomness-Efficient Plane vs. Point tester, given access to an input function $f : \mathbb{F}^m \to \mathbb{F}$ and oracle $\mathcal{A}$, accepts with probability γ. Let us prove the existence of a polynomial over $\mathbb{F}^m$ of degree at most md that agrees with f on at least $\gamma - \varepsilon$ fraction of the points, for $\varepsilon \leq const \cdot m \left(\sqrt[8]{\frac{1}{|\mathbb{H}|}} + \sqrt[4]{\frac{md}{|\mathbb{F}|}}\right)$.

1.5.1 Reformulating our goal

First, let us reformulate the problem in a more convenient manner. For dimensions k, m, where $k \leq m$, let $\mathcal{S}_k^m$ be the family of all affine subspaces of dimension k in $\mathbb{F}^m$ that are of the type we are interested in. Namely, a k-dimensional affine subspace $s \subseteq \mathbb{F}^m$ is in $\mathcal{S}_k^m$ if it can be written as $s = \left\{\vec{z} + \sum_{i=1}^k \alpha_i \vec{y}_i \;\middle|\; (\alpha_1, \ldots, \alpha_k) \in \mathbb{F}^k\right\}$ for some point $\vec{z} \in \mathbb{F}^m$ and some linearly independent directions $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$ (where the linear independence is over $\mathbb{F}$).

We can express (up to very small additive errors) the acceptance probability of the tester given access to $f : \mathbb{F}^m \to \mathbb{F}$ and $\mathcal{A}$ as follows:

$$\begin{aligned} \Pr\left[\text{tester accepts}\right] &\approx \Pr_{s\in\mathcal{S}_2^m, \vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) = f(\vec{x})\right] \\ &= \mathop{\mathbf{E}}_{s\in\mathcal{S}_2^m}\left[\Pr_{\vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) = f(\vec{x})\right]\right] \end{aligned}$$

For an affine subspace s and a degree d, let $\mathcal{Q}_{s,d}$ be the set of polynomials of degree at most d over s. It is evident from the last expression that an oracle $\mathcal{A}$ that optimizes the acceptance probability of the tester on input f assigns each subspace $s \in \mathcal{S}_2^m$ a polynomial $Q \in \mathcal{Q}_{s,d}$ that maximizes the agreement $Q(\vec{x}) = f(\vec{x})$ on points $\vec{x} \in s$. Hence, for every dimension m, function $f : \mathbb{F}^m \to \mathbb{F}$, dimension k and degree d, consider the *average agreement* of f with polynomial of degree at most d over subspaces $s \in \mathcal{S}_k^m$,

$$\mathsf{agr}_d^{k,m}(f) \stackrel{def}{=} \mathop{\mathbf{E}}_{s \in \mathcal{S}_k^m} \left[\max_{Q \in \mathcal{Q}_{s,d}} \left\{ \Pr_{\vec{x} \in s} \left[Q(\vec{x}) = f(\vec{x}) \right] \right\} \right]$$

Then,

$$\gamma = \Pr\left[\text{tester accepts}\right] \lesssim \mathsf{agr}_d^{2,m}(f)$$

For every m, the space $\mathbb{F}^m$ is the only affine subspace of dimension m in $\mathbb{F}^m$, and $\mathbb{H}^m$ contains a basis for $\mathbb{F}^m$, so $\mathcal{S}_m^m = \{\mathbb{F}^m\}$. Thus, for every dimension m, function $f : \mathbb{F}^m \to \mathbb{F}$, degree d and fraction γ, $\mathsf{agr}_d^{m,m}(f) \geq \gamma$ means that there exists $Q : \mathbb{F}^m \to \mathbb{F}$ of degree at most d, such that $\Pr_{\vec{x} \in \mathbb{F}^m}\left[Q(\vec{x}) = f(\vec{x})\right] \geq \gamma$.

We conclude that our goal can be reformulated as showing that large average agreement over planes implies large average agreement over $\mathbb{F}^m$. More accurately, for every function $f : \mathbb{F}^m \to \mathbb{F}$ and fraction $0 \leq \gamma \leq 1$,

$$\mathsf{agr}_d^{2,m}(f) \geq \gamma \quad \Rightarrow \quad \mathsf{agr}_{md}^{m,m}(f) \geq \gamma - \varepsilon$$

1.5.2 Main idea

We fix a dimension m, and our proof is by induction on the dimension k of the affine subspaces within $\mathbb{F}^m$. We assume that $\mathsf{agr}_d^{2,m}(f) \geq \gamma$, and show that for every dimension $2 \leq k \leq m$,

$$\mathsf{agr}_{kd}^{k,m}(f) \geq \gamma - \frac{k}{m} \cdot \varepsilon$$

Fix a dimension k such that $\mathsf{agr}_{(k-1)d}^{k-1,m}(f) \geq \gamma - \frac{k-1}{m} \cdot \varepsilon$, and let us outline how the induction step is done.

Consider *any* affine subspace $s \in \mathcal{S}_k^m$. Assume s contains the point $\vec{z} \in \mathbb{F}^m$ and is in directions $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$, where $\vec{y}_1, \ldots, \vec{y}_k$ are linearly independent over $\mathbb{F}$. The directions within s, $\{\vec{x}_1 - \vec{x}_2 \mid \vec{x}_1, \vec{x}_2 \in s\}$, are precisely $\sum_{i=1}^k \alpha_i \vec{y}_i$ for $\vec{\alpha} = (\alpha_1, \ldots, \alpha_k) \in \mathbb{F}^k$. Moreover, *since $\mathbb{H}$ is a subfield of $\mathbb{F}$*,

$$\vec{\alpha} \in \mathbb{H}^k \quad \Leftrightarrow \quad \sum_{i=1}^k \alpha_i \vec{y}_i \in \mathbb{H}^m$$

Therefore (unlike the construction of [7] via ϵ-biased sets), the families of affine subspaces we consider preserve the following two properties enabling induction:

1. **Self-similarity:** Every affine subspace $s \in \mathcal{S}_k^m$ is mapped onto $\mathbb{F}^k$ (via the natural bijection $\vec{z} + \sum_{i=1}^k \alpha_i \vec{y}_i \in s \;\leftrightarrow\; \vec{\alpha} \in \mathbb{F}^k$), such that the directions the tester considers (namely, the vectors in $\mathbb{H}^m$) that are also in s are mapped onto $\mathbb{H}^k$.

2. **Uniformity:** For every dimension $k' \leq k$, each subspace $s \in \mathcal{S}_k^m$ contains exactly the same number of subspaces $s' \in \mathcal{S}_{k'}^m$, and each subspace $s' \in \mathcal{S}_{k'}^m$ is contained in exactly the same number of subspaces $s \in \mathcal{S}_k^m$.

Let $f_{|s} : \mathbb{F}^k \to \mathbb{F}$ denote the restriction of f to s; namely, for every $(\alpha_1, \ldots, \alpha_k) \in \mathbb{F}^k$, let $f_{|s}(\alpha_1, \ldots, \alpha_k) = f(\vec{z} + \sum_{i=1}^{k} \alpha_i \vec{y}_i)$.

Consider some degree d' and dimension $k' \leq k$. By *self-similarity* and *uniformity*,

$$\mathsf{agr}_{d'}^{k',m}(f) = \mathop{\mathbf{E}}_{s \in \mathcal{S}_k^m} \left[\mathsf{agr}_{d'}^{k',k}(f_{|s}) \right] \tag{1}$$

Thus, it is sufficient (as we see shortly) to show that for every function $f : \mathbb{F}^k \to \mathbb{F}$ and every fraction $0 \leq \gamma \leq 1$,

$$\mathsf{agr}_{(k-1)d}^{k-1,k}(f) \geq \gamma \quad \Rightarrow \quad \mathsf{agr}_{kd}^{k,k}(f) \geq \gamma - \frac{\varepsilon}{m} \tag{2}$$

The inductive step is then completed applying the induction hypothesis as well as 1 and 2 above:

$$\begin{aligned}
\mathsf{agr}_{kd}^{k,m}(f) &= \mathop{\mathbf{E}}_{s \in \mathcal{S}_k^m} \left[\mathsf{agr}_{kd}^{k,k}(f_{|s}) \right] \\
&\geq \mathop{\mathbf{E}}_{s \in \mathcal{S}_k^m} \left[\mathsf{agr}_{(k-1)d}^{k-1,k}(f_{|s}) - \frac{\varepsilon}{m} \right] \\
&= \mathsf{agr}_{(k-1)d}^{k-1,m}(f) - \frac{\varepsilon}{m} \\
&\geq \gamma - \frac{k}{m} \cdot \varepsilon
\end{aligned}$$

1.5.3 Proving (2)

By an adaptation of an idea by Raz and Safra [17], we can prove that there exists a small error $\delta \ll \varepsilon/m$, such that for every function $f : \mathbb{F}^k \to \mathbb{F}$ and every fraction $0 \leq \gamma \leq 1$,

$$\mathsf{agr}_{(k-1)d}^{k-1,k}(f) \geq \gamma \quad \Rightarrow \quad \mathsf{agr}_{2(k-1)d}^{k,k}(f) \geq \gamma^2 - \delta$$

The idea of Raz and Safra [17] centers around a construction of a *consistency graph*. The vertices of the graph are the affine subspaces of dimension $(k-1)$ within $\mathbb{F}^k$ (namely, *hyperplanes*). The edges of the graph indicate whether there is an agreement between assignments of degree $(k-1)d$ polynomials to the hyperplanes. Due to its algebraic structure, the graph has a combinatorial property called *almost-transitivity*. It allows us to use a graph-theoretic lemma originally proven in [17], and go up from dimension $(k-1)$ to dimension k.

The reduction to the graph-theoretic setting introduces a certain deterioration of the degree and agreement parameters. The degree doubles (from $(k-1)d$ to $2(k-1)d$, rather than to kd) and the agreement is raised to the power of two (from γ to $\gamma^2 - \delta$, rather than to $\gamma - \varepsilon/m$). We cannot tolerate either deterioration, since they ultimately cause an exponential decay in k. Hence, we apply steps of what we call *consolidation* to retain the desired parameters. Similar techniques were already used in previous works, and they rely on the sampling properties we discussed above.

1.6 Organization

We state the main theorems regarding the soundness of our testers in section 2. The rest of the paper is devoted to proving these theorems. We start with some preliminary definitions and propositions in section 3. We discuss basic properties of affine subspaces with directions

over a subfield in section 4. We prove sampling properties in section 5. This allows us to prove consolidation claims in section 6. We present and analyze the consistency graph in section 7 and use it for going up one dimension in section 8. The soundness of the Randomness-Efficient Plane vs. Point tester is proven via induction in section 9. We show that the soundness of the Randomness-Efficient Subspace vs. Point tester follows in section 10. We give the proof of the combinatorial lemma of [17] in the appendix.

2 Our Results

2.1 Notation

In all that follows, we consider a finite field $\mathbb{F}$, a subfield $\mathbb{H} \subseteq \mathbb{F}$, a dimension m, and a degree d.

Given vectors $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{F}^m$, we define the *linear subspace* they span by $span\{\vec{y}_1, \ldots, \vec{y}_k\} \stackrel{def}{=} \{a_1\vec{y}_1 + \ldots + a_k\vec{y}_k \mid a_1, \ldots, a_k \in \mathbb{F}\}$. We say that $\vec{y}_1, \ldots, \vec{y}_k$ are *linearly independent*, and denote $\mathsf{ind}(\vec{y}_1, \ldots, \vec{y}_k)$, if for every $a_1, \ldots, a_k \in \mathbb{F}$, if $\sum_{i=1}^{k} a_i\vec{y}_i = 0$ then $a_1 = \cdots = a_k = 0$. Throughout the paper we will refer to span over $\mathbb{F}$ (and not over a subfield, even if the vectors are over a subfield). Note that vectors $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$ are linearly independent over $\mathbb{H}$ if and only if $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$ are linearly independent over $\mathbb{F}$.

Given two sets $A, B \subseteq \mathbb{F}^m$, we define $A + B \stackrel{def}{=} \{\vec{x} + \vec{y} \mid \vec{x} \in A, \vec{y} \in B\}$. Given a point $\vec{x} \in \mathbb{F}^m$ and a set $A \subseteq \mathbb{F}^m$, define $\vec{x} + A \stackrel{def}{=} \{\vec{x}\} + A$. A k-dimensional *affine subspace* in the vector space $\mathbb{F}^m$ is defined by a base-point $\vec{x} \in \mathbb{F}^m$ and k linearly independent directions, $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{F}^m$, as

$$\text{affine}(\vec{x}; \vec{y}_1, \ldots, \vec{y}_k) \stackrel{def}{=} \vec{x} + span\{\vec{y}_1, \ldots, \vec{y}_k\}$$

Points are 0-dimensional affine subspaces. *Lines* are 1-dimensional affine subspaces. *Planes* are 2-dimensional affine subspaces. Every affine subspace can be equivalently represented by many choices of vectors $\vec{x}; \vec{y}_1, \ldots, \vec{y}_k$, but, clearly, there is an affine transformation between every two representations of the same affine subspace.

An *m-variate polynomial* over a field $\mathbb{F}$ is a function $Q : \mathbb{F}^m \to \mathbb{F}$ of the form

$$Q(x_1, \ldots, x_m) = \sum_{i_1, \ldots, i_m} a_{i_1, \ldots, i_m} x_1^{i_1} \cdots x_m^{i_m}$$

where all the *coefficients* $a_{i_1,\ldots,i_m}$ are in $\mathbb{F}$. The *degree* of Q is $\deg Q \stackrel{def}{=} \max\left\{\sum_{j=1}^{m} i_j \mid a_{i_1,\ldots,i_m} \neq 0\right\}$, where the degree of the *identically zero* polynomial is defined to be 0.

The restriction of a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ to an affine subspace s represented as $s = \text{affine}(\vec{x}; \vec{y}_1, \ldots, \vec{y}_k)$ is a polynomial in k variables, $Q_{|s}(\alpha_1, \ldots, \alpha_k) \stackrel{def}{=} Q(\vec{x} + \alpha_1\vec{y}_1 + \ldots + \alpha_k\vec{y}_k)$. We will sometimes wish to refer to a polynomial Q defined over an affine subspace s without specifying the subspace's representation, in which case we will use the notation $Q(\vec{x})$ for a point $\vec{x} \in s$. Note that the *degree* of the polynomial does not depend on the representation of s.

2.2 Oracles

We assume an oracle $\mathcal{A}$ that given any affine subspace s in $\mathbb{F}^m$, provides a polynomial $\mathcal{A}(s)$ of degree at most d defined over s. For the sake of simplicity, we do not refer to both an oracle $\mathcal{A}$ and a function $f : \mathbb{F}^m \to \mathbb{F}$ as in the introduction. Instead, we assume that f's values on points $\vec{x}$ are given by $\mathcal{A}(\vec{x})$. Our testers query $\mathcal{A}$ only on affine subspaces of constant dimension.

However, for the analysis, it will be convenient to consider oracles queried regarding higher dimensional affine subspaces as well. Hence, an oracle $\mathcal{A}$ is defined to provide a value for any affine subspace.

For a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$, we will use the notation $(Q \equiv \mathcal{A})(s)$ to indicate that Q and $\mathcal{A}$ agree on a subspace s, i.e., for every $\vec{x} \in s$, $Q(\vec{x}) = \mathcal{A}(s)(\vec{x})$.

2.3 Low Degree Testers

Define two predicates for our two testers: for $\vec{z} \in \mathbb{F}^m$ and $\vec{y}_1, \vec{y}_2 \in \mathbb{H}^m$, let

1. $PlanePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2)$: $\vec{y}_1, \vec{y}_2$ are linearly dependent or $\mathcal{A}(\text{affine}(\vec{z}; \vec{y}_1, \vec{y}_2))(\vec{z}) = \mathcal{A}(\vec{z})$
2. $SpacePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2)$: $\vec{z}, \vec{y}_1, \vec{y}_2$ are linearly dependent or $\mathcal{A}(\text{affine}(\vec{0}; \vec{z}, \vec{y}_1, \vec{y}_2))(\vec{z}) = \mathcal{A}(\vec{z})$

2.4 Soundness

To prove that a tester is sound we assume that it accepts with probability γ when given access to an oracle $\mathcal{A}$ and show the agreement of $\mathcal{A}$ with a low degree polynomial. Specifically, for a sub-constant ε, we prove two claims, which we argue to be essentially equivalent:

1. *(decoding)* There exists a low degree polynomial that is consistent with the oracle $\mathcal{A}$ on at least $\gamma - \varepsilon$ fraction of the points.
2. *(list decoding)* For every $0 < \delta < 1$, there exists a short list of $t = t(\delta)$ low degree polynomials that *explains* all the tester's acceptance, but $\delta + \varepsilon$ fraction of the probability (explanation follows).

When saying that a list of polynomials *explains* almost all the success, we mean that with high probability over the random bits of the tester (i.e., over the choice of a subspace and a point within it), either the tester rejects or one of the polynomials agrees with the oracle on the subspace and on the point. There is a tradeoff between the amount of success explained and the length of the list: the more one wishes to explain – the longer the list is.

We wish ε to be as small as possible. The parameter ε we achieve depends on $\frac{md}{|\mathbb{F}|}$. This comes from the use of the Schwartz-Zippel Lemma. It also depends on $\frac{1}{|\mathbb{H}|}$ which is the price we pay for considering the subfield $\mathbb{H}$ instead of the entire field $\mathbb{F}$.

The statement for the Randomness-Efficient Plane vs. Point tester is as follows. Note that we make no effort to optimize the constants.

Theorem 1 (Plane vs. Point Soundness). *Fix a dimension $m \geq 2$, a field $\mathbb{F}$, a subfield $\mathbb{H} \subseteq \mathbb{F}$ and a degree d. Denote $\varepsilon \stackrel{def}{=} 2^7 m \left(\sqrt[8]{\frac{1}{|\mathbb{H}|}} + \sqrt[4]{\frac{md}{|\mathbb{F}|}} \right)$. For every oracle $\mathcal{A}$,*

1. ***(Decoding)*** *There exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq md$, such that*

$$\Pr_{\vec{x} \in \mathbb{F}^m} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \quad \geq \quad \Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m} \left[PlanePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2) \right] - \varepsilon$$

2. ***(List decoding)*** *For every $\delta > 2\varepsilon$, there exist $t \leq 2/\delta$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q_i \leq md$, such that*

$$\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m} \left[\neg PlanePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2) \ \vee \ \exists i \ (Q_i \equiv \mathcal{A})(\text{affine}(\vec{z}; \vec{y}_1, \vec{y}_2)) \right] \quad \geq \quad 1 - \delta - 2\varepsilon$$

We prove a similar theorem for the Randomness-Efficient Subspace vs. Point tester. Note that for this tester we manage to show agreement with polynomials of degree at most d, rather than md.

Theorem 2 (Subspace vs. Point Soundness). *Fix a dimension $m \geq 3$, a field $\mathbb{F}$, a subfield $\mathbb{H} \subseteq \mathbb{F}$ and a degree d. Denote $\varepsilon \stackrel{def}{=} 2^7 m \left(\sqrt[8]{\frac{1}{|\mathbb{H}|}} + \sqrt[4]{\frac{md}{|\mathbb{F}|}} \right)$. For every oracle $\mathcal{A}$,*

1. ***(Decoding)** There exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq d$, such that*

$$\Pr_{\vec{x} \in \mathbb{F}^m} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \quad \geq \quad \Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m} \left[SpacePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2) \right] - 3\varepsilon$$

2. ***(List decoding)** For every $\delta > 3\varepsilon$, there exist $t \leq 2/\delta$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q_i \leq d$, such that*

$$\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m} \left[\neg SpacePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2) \ \vee \ \exists i \ (Q_i \equiv \mathcal{A})(\text{affine}(\vec{0}; \vec{z}, \vec{y}_1, \vec{y}_2)) \right] \quad \geq \quad 1 - \delta - 3\varepsilon$$

It is interesting to note that our sampling arguments also imply a converse to the above theorems: for any polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq d$, there exists an oracle $\mathcal{A}'$ agreeing with $\mathcal{A}$ on the points and assigning affine subspaces polynomials of degree at most d, such that both our testers accept with probability at least $\Pr_{\vec{x} \in \mathbb{F}^m} [Q(\vec{x}) = \mathcal{A}(\vec{x})] - \varepsilon$ when given access to $\mathcal{A}'$.

3 Preliminaries

3.1 Orthogonality and Vector Spaces

Given a vector $\vec{y} \in \mathbb{F}^m$, we write $\vec{y} = (y_1, \ldots, y_m)$. For a sequence of vectors $\vec{y}_1, \ldots, \vec{y}_k$, we write for every $1 \leq i \leq k$, $\vec{y}_i = (y_{i,1}, \ldots, y_{i,m})$.

We define an *inner-product* between two vectors $\vec{x}, \vec{y} \in \mathbb{F}^m$ as $(\vec{x}, \vec{y}) \stackrel{def}{=} \sum_{i=1}^m x_i \cdot y_i$. We say that $\vec{x}, \vec{y}$ are *orthogonal* if $(\vec{x}, \vec{y}) = 0$.

Proposition 3.1. *For every $\vec{y} \neq \vec{0} \in \mathbb{F}^m$, for every $c \in \mathbb{F}$,*

$$\Pr_{\vec{z} \in \mathbb{H}^m} [(\vec{z}, \vec{y}) = c] \leq \frac{1}{|\mathbb{H}|}$$

Proof. As $\vec{y} \neq \vec{0} \in \mathbb{F}^m$, there exists $1 \leq i \leq m$ such that $y_i \neq 0$. For every fixing of all $\vec{z}$'s coordinates but the i'th, the condition $(\vec{z}, \vec{y}) = c$ uniquely determines z_i to some scalar in $\mathbb{F}$. This scalar may or may not be in the subfield $\mathbb{H}$, but, in any case, there exists at most one possibility for $z_i \in \mathbb{H}$. ∎

Proposition 3.2. *For every $\vec{y} \neq \vec{0} \in \mathbb{F}^m$, for every $k < m$,*

$$\Pr_{\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m} [\vec{y} \in span\{\vec{y}_1, \ldots, \vec{y}_k\} \mid \mathit{ind}(\vec{y}_1, \ldots, \vec{y}_k)] \leq \frac{1}{|\mathbb{H}|}$$

Proof. Consider uniformly distributed linearly independent $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$. Pick uniformly and independently at random a vector $\vec{z} \neq \vec{0} \in \mathbb{H}^m$ that is orthogonal to $\vec{y}_1, \ldots, \vec{y}_k$ (there exist such vectors since $k < m$). Note that for every $\vec{y} \in span\{\vec{y}_1, \ldots, \vec{y}_k\}$ it holds that $(\vec{z}, \vec{y}) = 0$. By Proposition 3.1, since $\vec{z}$ is uniformly distributed over $\mathbb{H}^m \setminus \{\vec{0}\}$, this happens with probability at most $\frac{1}{|\mathbb{H}|}$. ∎

Proposition 3.3. *For every subset $A \subseteq \mathbb{F}^m$ with $|A| > |\mathbb{F}|^{m-1}$, there exist linearly independent $\vec{y}_1, \ldots, \vec{y}_m \in \mathbb{F}^m$, such that for every $1 \leq i \leq m$, $\vec{y}_i \in A$.*

Proof. We have

$$|span(A)| \geq |A| > |\mathbb{F}|^{m-1}$$

Since $span(A)$ is a linear subspace in $\mathbb{F}^m$, we must have $|span(A)| = |\mathbb{F}|^m$. Thus, $span(A) = \mathbb{F}^m$, and so A contains a basis for $\mathbb{F}^m$. ∎

3.2 Polynomials

The Schwartz-Zippel Lemma shows that different low degree polynomials differ on most points,

Proposition 3.4 (Schwartz-Zippel). *For two different polynomials $Q, P : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q, \deg P \leq d$,*

$$\Pr_{\vec{x} \in \mathbb{F}^m} [Q(\vec{x}) = P(\vec{x})] \leq \frac{d}{|\mathbb{F}|}$$

The Schwartz-Zippel Lemma can be viewed as showing the unique-decoding property of the Reed-Muller code. This immediately implies a list decoding property, namely, that only few polynomials can agree with a function on many of the points.

We include a simple proof of this property.

Proposition 3.5 (list decoding). *Fix a finite field $\mathbb{F}$ and a dimension m. Let $f : \mathbb{F}^m \to \mathbb{F}$ be some function and consider some degree $d \leq |\mathbb{F}|$. Then, for any $\delta \geq 2\sqrt{\frac{d}{|\mathbb{F}|}}$, if $Q_1, \ldots, Q_l : \mathbb{F}^m \to \mathbb{F}$ are* different *polynomials of degree at most d, and for every $1 \leq i \leq l$, the polynomial Q_i agrees with f on at least δ fraction of the points, i.e., $\Pr_{\vec{x} \in \mathbb{F}^m} [Q_i(\vec{x}) = f(\vec{x})] \geq \delta$, then $l \leq \frac{2}{\delta}$.*

Proof. Let $\delta \geq 2\sqrt{\frac{d}{|\mathbb{F}|}}$, and assume by way of contradiction that there exist $l = \lfloor \frac{2}{\delta} \rfloor + 1$ different polynomials $Q_1, \ldots, Q_l : \mathbb{F}^m \to \mathbb{F}$ as stated.

For every $1 \leq i \leq l$, let $A_i \stackrel{def}{=} \{\vec{x} \in \mathbb{F}^m \mid Q_i(\vec{x}) = f(\vec{x})\}$. By inclusion-exclusion,

$$|\mathbb{F}^m| \geq \left| \bigcup_{i=1}^{l} A_i \right| \geq \sum_{i=1}^{l} |A_i| - \sum_{i \neq j} |A_i \cap A_j|$$

By Schwartz-Zippel, for every $1 \leq i \neq j \leq l$, $|A_i \cap A_j| \leq \frac{d}{|\mathbb{F}|} \cdot |\mathbb{F}^m|$. Therefore, by the premise,

$$|\mathbb{F}^m| \geq l\delta \, |\mathbb{F}^m| - \binom{l}{2} \frac{d}{|\mathbb{F}|} \, |\mathbb{F}^m|$$

On one hand, since $l > \frac{2}{\delta}$, we get $l\delta > 2$. On the other hand, since $\frac{2}{\delta} \leq \sqrt{\frac{|\mathbb{F}|}{d}}$ and $d \leq |\mathbb{F}|$, we get $\binom{l}{2} \leq \frac{|\mathbb{F}|}{d}$. This results in a contradiction. ∎

4 Affine Subspaces With Directions Over a Subfield

In this section we prove basic facts regarding affine subspaces in $\mathbb{F}^m$ that are spanned by directions over a subfield $\mathbb{H} \subseteq \mathbb{F}$. All the properties we prove for such subspaces are well known when $\mathbb{H} = \mathbb{F}$.

For $0 \leq k \leq m$, consider the set of representations of affine subspaces with directions over a subfield,

$$\mathcal{R}_k^m \stackrel{def}{=} \{(\vec{z}; \vec{y}_1, \ldots, \vec{y}_k) \mid \vec{z} \in \mathbb{F}^m, \vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m, \mathsf{ind}(\vec{y}_1, \ldots, \vec{y}_k)\}$$

The corresponding set of affine subspaces is

$$\mathcal{S}_k^m \stackrel{def}{=} \{\text{affine}(r) \mid r \in \mathcal{R}_k^m\}$$

First we would like to assert that every subspace in $\mathcal{S}_k^m$ is associated with the same number of tuples in $\mathcal{R}_k^m$, and that every subspace in $\mathcal{S}_k^m$ contains the same number of subspaces in $\mathcal{S}_{k'}^m$ for $k' \leq k$,

Proposition 4.1 (uniformity). *For every dimension k, there is a number $T = T(k)$, such that for every $s \in \mathcal{S}_k^m$, $|\{r \in \mathcal{R}_k^m \mid s = \text{affine}(r)\}| = T$.*

Proposition 4.2 (uniformity downwards). *For every dimensions $k' \leq k$, there is a number $T = T(k, k')$, such that for every $s \in \mathcal{S}_k^m$, $\left|\left\{s' \in \mathcal{S}_{k'}^m \mid s' \subseteq s\right\}\right| = T$.*

To prove both assertions we introduce an additional notation allowing us to refer to affine subspaces in $\mathcal{S}_k^m$ as isomorphic copies of $\mathbb{F}^k$. Fix an affine subspace together with a representation for it, $s = \text{affine}(\vec{z}; \vec{y}_1, \ldots, \vec{y}_k)$. For a representation $r = (\vec{\alpha}_0; \vec{\alpha}_1, \ldots, \vec{\alpha}_{k'})$ of a k'-dimensional affine subspace within $\mathbb{F}^k$, we define the representation r *relative to* (the representation of) the space s by

$$r_s \stackrel{def}{=} \left(\vec{z} + \sum_{i=1}^{k} \vec{\alpha}_{0,i}\vec{y}_i \;;\; \sum_{i=1}^{k} \vec{\alpha}_{1,i}\vec{y}_i, \ldots, \sum_{i=1}^{k} \vec{\alpha}_{k',i}\vec{y}_i\right)$$

Note that since $\vec{y}_1, \ldots, \vec{y}_k$ are linearly independent, if two representations r, r' are the same relative to a subspace s, $r_s = r'_s$, then they are the same representation $r = r'$.

Denote the corresponding relative affine subspace:

$$\text{affine}_s(r) \stackrel{def}{=} \text{affine}(r_s)$$

Note that for every r, $\text{affine}_s(r) \subseteq s$. Moreover, if $\text{affine}(r) = \text{affine}(r')$ then $\text{affine}_s(r) = \text{affine}_s(r')$. Now, the above two propositions follow from the following proposition:

Proposition 4.3. *For every subspace $s \in \mathcal{S}_k^m$, for every dimension $k' \leq k$,*

$$S_1 \stackrel{def}{=} |\{r \in \mathcal{R}_{k'}^m \mid \text{affine}(r) \subseteq s\}| = \left|\mathcal{R}_{k'}^k\right| \stackrel{def}{=} S_2$$

Proof. Fix a subspace $s \in \mathcal{S}_k^m$ and fix a tuple $(\vec{z}; \vec{y}_1, \ldots, \vec{y}_k) \in \mathcal{R}_k^m$ with $s = \text{affine}(\vec{z}; \vec{y}_1, \ldots, \vec{y}_k)$.

1. $S_1 \geq S_2$: for every tuple $r = (\vec{\alpha}_0; \vec{\alpha}_1, \ldots, \vec{\alpha}_{k'}) \in \mathcal{R}_{k'}^k$, the tuple r_s satisfies $r_s \in \mathcal{R}_{k'}^m$ and $\text{affine}(r_s) \subseteq s$.

2. $S_1 \leq S_2$: for every tuple $r \in \mathcal{R}_{k'}^m$ satisfying $\text{affine}(r) \subseteq s$, there exists exactly one $\alpha = (\vec{\alpha}_0; \vec{\alpha}_1, \ldots, \vec{\alpha}_{k'})$, $\vec{\alpha}_0, \vec{\alpha}_1, \ldots, \vec{\alpha}_{k'} \in \mathbb{F}^k$, $\mathsf{ind}(\vec{\alpha}_1, \ldots, \vec{\alpha}_{k'})$, such that $r = \alpha_s$. Since $r \in \mathcal{R}_{k'}^m$ and $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$, also $\vec{\alpha}_1, \ldots, \vec{\alpha}_{k'} \in \mathbb{H}^k$.

■

Every subspace in $\mathcal{S}_k^m$ is contained in the same number of subspaces in $\mathcal{S}_{k'}^m$ for $k' \geq k$,

Proposition 4.4 (uniformity upwards). *For every dimensions $k \leq k' \leq m$, there is a number $T = T(m,k,k')$, such that for every subspace $s \in \mathcal{S}_k^m$,*

$$\left|\left\{ s' \in \mathcal{S}_{k'}^m \;\middle|\; s' \supseteq s \right\}\right| = T$$

Proof. Let us introduce an additional piece of notation: $\mathcal{L}_{k'}^m$ is the set of all linear subspaces of dimension k' in $\mathbb{F}^m$ spanned by vectors from $\mathbb{H}^m$.

Fix $s = \text{affine}(\vec{z}; \vec{y}_1, \ldots, \vec{y}_k) \in \mathcal{S}_k^m$. Since $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$ are linearly independent, the proposition will clearly follow if we prove the following:

$$S_1 \stackrel{def}{=} \left|\left\{ s' \in \mathcal{S}_{k'}^m \;\middle|\; s' \supseteq s \right\}\right| = \left|\left\{ Y' \in \mathcal{L}_{k'}^m \;\middle|\; Y' \supseteq \{\vec{y}_1, \ldots, \vec{y}_k\} \right\}\right| \stackrel{def}{=} S_2$$

1. $S_1 \leq S_2$: Let $s' = \text{affine}(\vec{z}'; \vec{y}'_1, \ldots, \vec{y}'_{k'}) \in \mathcal{S}_{k'}^m$, $(\vec{z}'; \vec{y}'_1, \ldots, \vec{y}'_{k'}) \in \mathcal{R}_{k'}^m$, $s' \supseteq s$. Let $Y' = span\{\vec{y}'_1, \ldots, \vec{y}'_{k'}\}$. Clearly, Y' is in $\mathcal{L}_{k'}^m$ and Y' is uniquely defined by s', $s' = \vec{z}' + Y'$. It holds that $\vec{z} \in s \subseteq s' = \vec{z}' + Y'$, thus $\vec{z}' \in \vec{z} + Y'$, and, hence, $s' = \vec{z} + Y'$. Let $1 \leq i \leq k$. It holds that $\vec{z} + \vec{y}_i \in s \subseteq s'$. This implies that $\vec{z} + \vec{y}_i \in \vec{z} + Y'$. Hence, $\vec{y}_i \in Y'$. Therefore, $\{\vec{y}_1, \ldots, \vec{y}_k\} \subseteq Y'$.
2. $S_1 \geq S_2$: Let $Y' \in \mathcal{L}_{k'}^m$, $Y' \supseteq \{\vec{y}_1, \ldots, \vec{y}_k\}$. Clearly, $\vec{z} + Y' \in \mathcal{S}_{k'}^m$ and $s \subseteq \vec{z} + Y'$.

■

Uniformity is so important because it allows us to count in several ways. A simple argument of this nature is that the fraction of affine subspaces $s \in \mathcal{S}_k^m$ satisfying some condition is exactly the same as the fraction of $r \in \mathcal{R}_k^m$ such that affine(r) satisfies the condition. Let us demonstrate a more sophisticated argument of this nature. Fix $k' \leq k$. Suppose that we have a predicate R indicating whether an affine subspace $s \in \mathcal{S}_k^m$ and an affine subspace $s' \in \mathcal{S}_{k'}^m$ *contained in it*, $s' \subseteq s$, satisfy some relation. Then,

$$\mathop{\mathbf{E}}_{s}\left[\Pr_{s' \subseteq s}\left[R(s,s')\right]\right] = \mathop{\mathbf{E}}_{s'}\left[\Pr_{s \supseteq s'}\left[R(s,s')\right]\right]$$

The intersection between two subspaces $s_1 \in \mathcal{S}_{k^{(1)}}^m$ and $s_2 \in \mathcal{S}_{k^{(2)}}^m$ is again (provided it is not empty) a subspace in $\mathcal{S}_{k^{(3)}}^m$ for some $k^{(3)}$.

Proposition 4.5 (closure under intersection). *If $s_1 \in \mathcal{S}_{k^{(1)}}^m$ and $s_2 \in \mathcal{S}_{k^{(2)}}^m$ where $s_1 \cap s_2 \neq \phi$, then there exists $k^{(3)}$ such that $s_1 \cap s_2 \in \mathcal{S}_{k^{(3)}}^m$.*

Proof. Write $s_1 = \vec{z}_1 + V_1$ and $s_2 = \vec{z}_2 + V_2$ where $\vec{z}_1, \vec{z}_2 \in \mathbb{F}^m$ and $V_1, V_2 \subseteq \mathbb{F}^m$ are linear subspaces spanned by vectors in $\mathbb{H}^m$. Assume $\vec{x} \in s_1 \cap s_2$. Then, we can alternatively write $s_1 = \vec{x} + V_1$ and $s_2 = \vec{x} + V_2$. Thus, $s_1 \cap s_2 = \vec{x} + (V_1 \cap V_2)$. The proposition follows noticing that $V_1 \cap V_2$ can be spanned by vectors in $\mathbb{H}^m$. ■

A useful representation of affine subspaces is given in the following proposition,

Proposition 4.6 (affine subspaces as solutions of linear equations). *Let $s = \text{affine}(\vec{z}; \vec{y}_1, \ldots, \vec{y}_k) \in \mathcal{S}_k^m$, let $\vec{\alpha}_1, \ldots, \vec{\alpha}_{m-k} \in \mathbb{H}^m$ be $(m-k)$ linearly independent vectors orthogonal to $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$. Then,*

$$s = \{\vec{x} \in \mathbb{F}^m \mid \forall 1 \leq j \leq m-k,\ (\vec{x}, \vec{\alpha}_j) = (\vec{z}, \vec{\alpha}_j)\}$$

Proof. Fix $\vec{x} \in s$. Hence, there exists $\vec{c} \in \mathbb{F}^k$ such that $\vec{x} = \vec{z} + \sum_{i=1}^k c_i \vec{y}_i$. For every $1 \leq j \leq m-k$,

$$(\vec{x}, \vec{\alpha}_j) = \left(\vec{z} + \sum_{i=1}^k c_i \vec{y}_i, \vec{\alpha}_j\right) = (\vec{z}, \vec{\alpha}_j) + \sum_{i=1}^k c_i \cdot (\vec{y}_i, \vec{\alpha}_j) = (\vec{z}, \vec{\alpha}_j)$$

Thus, $s \subseteq \{\vec{x} \in \mathbb{F}^m \mid \forall 1 \leq j \leq m-k,\ (\vec{x}, \vec{\alpha}_j) = (\vec{z}, \vec{\alpha}_j)\}$. The proposition follows noticing that, in addition, the two sets are of size $\left|\mathbb{F}^k\right|$. ■

5 Affine Subspaces With Directions Over A Subfield Sample Well

We say that an affine subspace s in $\mathbb{F}^m$ samples a set $A \subseteq \mathbb{F}^m$ well if the fraction of points from A contained in it, i.e., $\frac{|s \cap A|}{|s|}$, is approximately $\frac{|A|}{|\mathbb{F}^m|}$. We say that a distribution $\mathcal{D}$ on affine subspaces in $\mathbb{F}^m$ samples well, if no matter how one fixes a large enough subset $A \subseteq \mathbb{F}^m$, a random subspace $s \sim \mathcal{D}$ samples A well with high probability. In this section we use Fourier analysis to show that the distributions induced by our testers sample well.

5.1 Fourier Transform

Let $(G, +)$ be a finite Abelian group. Consider functions from the group to the complex numbers $f : G \to \mathbb{C}$. One example for such a function is the indicator function of a multi-set $A \subseteq G$, i.e., the function $\mathcal{I}_A$ that assigns every $\vec{x} \in G$ its multiplicity in A.

We define an *inner-product* between functions $f, g : G \to \mathbb{C}$ as

$$\langle f, g \rangle \stackrel{def}{=} \frac{1}{|G|} \sum_{x \in G} f(x)\overline{g(x)}$$

A *character* of G is a homomorphism $\chi : G \to \mathbb{C}^*$, where $\mathbb{C}^*$ is the multiplicative group of the complex numbers. Namely, for every $x, y \in G$,

$$\chi(x + y) = \chi(x) \cdot \chi(y)$$

Every group G trivially has the identically 1 function as a character.

It can be shown that the set of all characters of G forms an orthonormal basis for the space of all functions $f : G \to \mathbb{C}$ under the inner-product defined above. Hence, every function $f : G \to \mathbb{C}$ can be equivalently represented as $f(x) = \sum_\chi \hat{f}(\chi) \cdot \chi(x)$, where $\hat{f}(\chi) \stackrel{def}{=} \langle f, \chi \rangle$ is called the *Fourier coefficient* of f corresponding to the character χ. The linear transformation from f to $\hat{f}$ is called the *Fourier transform* of f.

We will need two basic facts regarding the Fourier transform:

Proposition 5.1 (Parseval's identity). *For two functions $f, g : G \to \mathbb{C}$, $\langle f, g \rangle = |G| \cdot \langle \hat{f}, \hat{g} \rangle = \sum_\chi \hat{f}(\chi)\overline{\hat{g}(\chi)}$.*

Define the *convolution* of two functions, $f, g : G \to \mathbb{C}$, denoted $(f * g) : G \to \mathbb{C}$, as $(f * g)(x) \stackrel{def}{=} \frac{1}{|G|} \sum_{y \in G} f(y)g(x - y)$.

Proposition 5.2 (convolution formula). *Fix two functions, $f, g : G \to \mathbb{C}$. For every character χ of G, $\widehat{(f * g)}(\chi) = \hat{f}(\chi) \cdot \hat{g}(\chi)$.*

We focus on the additive group $G = \mathbb{F}^m$ for some finite field $\mathbb{F} = GF(p^k)$. The field $\mathbb{F}$ is also viewed as a vector space of dimension k over the field $GF(p)$.

Denote $\omega_p = e^{2\pi i/p}$ the p'th *primitive root of unity* in $\mathbb{C}$. For every $\alpha \in \mathbb{F}^m$, there is a character $\chi_\alpha : \mathbb{F}^m \to \mathbb{C}$,

$$\chi_\alpha(x) \stackrel{def}{=} \omega_p^{\sum_{i=1}^m (\alpha_i, x_i)}$$

Note that we view α_i, x_i as vectors in $GF(p)^k$. Their inner product is in $GF(p)$ and so is the sum in the above expression.

For a function $f : \mathbb{F}^m \to \mathbb{C}$, we denote its Fourier coefficient corresponding to the character χ_α by $\hat{f}(\alpha)$.

5.2 Sampling Lemma

In this subsection we prove our basic lemma via Fourier analysis. Given $z, y \in \mathbb{F}^m$ and a subset $A \subseteq \mathbb{F}^m$, define $X_{z,y}$ to be the number of $c \in \mathbb{F}$ satisfying $z + c \cdot y \in A$. Clearly, the expectation of $X_{z,y}$ when picking independently at random $z \in \mathbb{F}^m$ and $y \in \mathbb{H}^m$ is $|\mathbb{F}| \cdot \frac{|A|}{|\mathbb{F}^m|}$. We bound the variance of $X_{z,y}$, implying that it is concentrated around its expectation.

Lemma 5.3. *For any subset $A \subseteq \mathbb{F}^m$ of density $\mu = |A|/|\mathbb{F}^m|$,*

$$\mathop{\mathbf{Var}}_{z \in \mathbb{F}^m, y \in \mathbb{H}^m}[X_{z,y}] \leq |\mathbb{F}|^2 \frac{\mu}{|\mathbb{H}|}$$

Proof. If we denote the indicator function of A by $\mathcal{I}_A$, and the indicator function of the multi-set $\{c \cdot y \mid c \in \mathbb{F}\}$ by $\mathcal{I}_{\mathbb{F}y}$, we can express:

$$X_{z,y} = \sum_{x \in \mathbb{F}^m} \mathcal{I}_A(x)\mathcal{I}_{\mathbb{F}y}(z-x) = |\mathbb{F}^m| \cdot (\mathcal{I}_A * \mathcal{I}_{\mathbb{F}y})(z)$$

Hence, by Parseval's identity and the convolution formula,

$$\begin{aligned}
\mathop{\mathbf{E}}_{z \in \mathbb{F}^m, y \in \mathbb{H}^m}\left[X_{z,y}^2\right] &= \frac{1}{|\mathbb{F}^m|\,|\mathbb{H}^m|} \cdot \sum_{y \in \mathbb{H}^m} \sum_{z \in \mathbb{F}^m} \left(|\mathbb{F}^m|\,(\mathcal{I}_A * \mathcal{I}_{\mathbb{F}y})(z)\right)^2 \\
&= \frac{|\mathbb{F}^m|^2}{|\mathbb{H}^m|} \cdot \sum_{y \in \mathbb{H}^m} \sum_{\alpha \in \mathbb{F}^m} \left|\widehat{(\mathcal{I}_A * \mathcal{I}_{\mathbb{F}y})}(\alpha)\right|^2 \\
&= \frac{|\mathbb{F}^m|^2}{|\mathbb{H}^m|} \cdot \sum_{y \in \mathbb{H}^m} \sum_{\alpha \in \mathbb{F}^m} \left|\hat{\mathcal{I}}_A(\alpha)\right|^2 \cdot \left|\hat{\mathcal{I}}_{\mathbb{F}y}(\alpha)\right|^2
\end{aligned}$$

By definition, for any multi-set $S \subseteq \mathbb{F}^m$, $\hat{\mathcal{I}}_S(\vec{0}) = \frac{|S|}{|\mathbb{F}^m|}$ (where $|S| = \sum_{\vec{x} \in \mathbb{F}^m} \mathcal{I}_S(\vec{x})$), hence,

$$\begin{aligned}
\mathop{\mathbf{E}}_{z \in \mathbb{F}^m, y \in \mathbb{H}^m}\left[X_{z,y}^2\right] &= \frac{|\mathbb{F}^m|^2}{|\mathbb{H}^m|} \cdot \sum_{y \in \mathbb{H}^m} \left(\left|\hat{\mathcal{I}}_A(\vec{0})\right|^2 \cdot \left|\hat{\mathcal{I}}_{\mathbb{F}y}(\vec{0})\right|^2 + \sum_{\alpha \neq \vec{0} \in \mathbb{F}^m} \left|\hat{\mathcal{I}}_A(\alpha)\right|^2 \cdot \left|\hat{\mathcal{I}}_{\mathbb{F}y}(\alpha)\right|^2 \right) \\
&= \left(\frac{|\mathbb{F}|\,|A|}{|\mathbb{F}^m|}\right)^2 + \sum_{\alpha \neq \vec{0} \in \mathbb{F}^m} \left(\left|\hat{\mathcal{I}}_A(\alpha)\right|^2 \cdot \frac{|\mathbb{F}^m|^2}{|\mathbb{H}^m|} \sum_{y \in \mathbb{H}^m} \left|\hat{\mathcal{I}}_{\mathbb{F}y}(\alpha)\right|^2 \right)
\end{aligned}$$

We will show that $\frac{|\mathbb{F}^m|^2}{|\mathbb{H}^m|} \cdot \sum_{y \in \mathbb{H}^m} \left|\hat{\mathcal{I}}_{\mathbb{F}y}(\alpha)\right|^2 \leq \frac{|\mathbb{F}|^2}{|\mathbb{H}|}$. Let us see how the lemma follows. Using this bound and applying Parseval's identity again we get,

$$\begin{aligned}
\mathop{\mathbf{E}}_{z \in \mathbb{F}^m, y \in \mathbb{H}^m}\left[X_{z,y}^2\right] &\leq \left(\frac{|\mathbb{F}|\,|A|}{|\mathbb{F}^m|}\right)^2 + \frac{|\mathbb{F}|^2}{|\mathbb{H}|} \cdot \sum_{\alpha \neq \vec{0} \in \mathbb{F}^m} \left|\hat{\mathcal{I}}_A(\alpha)\right|^2 \\
&\leq \left(\frac{|\mathbb{F}|\,|A|}{|\mathbb{F}^m|}\right)^2 + \frac{|\mathbb{F}|^2}{|\mathbb{H}|} \cdot \frac{1}{|\mathbb{F}^m|} \cdot \sum_{z \in \mathbb{F}^m} |\mathcal{I}_A(z)|^2 \\
&= \left(\frac{|\mathbb{F}|\,|A|}{|\mathbb{F}^m|}\right)^2 + \frac{|\mathbb{F}|^2}{|\mathbb{H}|} \cdot \frac{|A|}{|\mathbb{F}^m|}
\end{aligned}$$

By linearity of expectations,

$$\mathop{\mathbf{E}}_{z\in\mathbb{F}^m, y\in\mathbb{H}^m}[X_{z,y}] = \frac{|\mathbb{F}|\,|A|}{|\mathbb{F}^m|}$$

Therefore,

$$\begin{aligned}
\mathop{\mathbf{Var}}_{z\in\mathbb{F}^m, y\in\mathbb{H}^m}[X_{z,y}] &= \mathop{\mathbf{E}}_{z,y}\left[X_{z,y}^2\right] - \mathop{\mathbf{E}}_{z,y}[X_{z,y}]^2 \\
&\leq \left(\frac{|\mathbb{F}|\,|A|}{|\mathbb{F}^m|}\right)^2 + \frac{|\mathbb{F}|^2}{|\mathbb{H}|}\cdot\frac{|A|}{|\mathbb{F}^m|} - \left(\frac{|\mathbb{F}|\,|A|}{|\mathbb{F}^m|}\right)^2 \\
&= |\mathbb{F}|^2\,\frac{\mu}{|\mathbb{H}|}
\end{aligned}$$

We conclude that proving the lemma boils down to proving the following:

Claim 5.3.1. *For every $\alpha \neq \vec{0} \in \mathbb{F}^m$,*

$$\frac{1}{|\mathbb{H}^m|}\cdot\sum_{y\in\mathbb{H}^m}\left|\hat{\mathcal{I}}_{\mathbb{F}y}(\alpha)\right|^2 \leq \frac{|\mathbb{F}|^2}{|\mathbb{F}^m|^2}\cdot\frac{1}{|\mathbb{H}|}$$

Proof. Assume $\mathbb{F} = GF(p^k)$. Fix some $\alpha \neq \vec{0} \in \mathbb{F}^m$.

$$\left|\hat{\mathcal{I}}_{\mathbb{F}y}(\alpha)\right| = |\langle \mathcal{I}_{\mathbb{F}y}, \chi_\alpha\rangle| = \left|\frac{1}{|\mathbb{F}^m|}\cdot\sum_{z\in\mathbb{F}^m}\mathcal{I}_{\mathbb{F}y}(z)\omega_p^{-\sum_{i=1}^m(\alpha_i,z_i)}\right| = \left|\frac{1}{|\mathbb{F}^m|}\cdot\sum_{c\in\mathbb{F}}\omega_p^{-\sum_{i=1}^m(\alpha_i,c\cdot y_i)}\right|$$

Multiplication by a field element $a \in \mathbb{F}$ in the field $\mathbb{F} = GF(p^k)$ corresponds to a linear transformation in the vector space $GF(p)^k$. That is, for every $a \in \mathbb{F}$, there exists a $k \times k$ matrix M_a over $GF(p)$, such that for every $b \in \mathbb{F} = GF(p)^k$, $a \cdot b = M_a b$. Hence,

$$\begin{aligned}
\sum_{i=1}^m(\alpha_i, c\cdot y_i) &= \sum_{i=1}^m(\alpha_i, M_{y_i}c) \\
&= \sum_{i=1}^m(M_{y_i}^T\alpha_i, c) \\
&= (\sum_{i=1}^m M_{y_i}^T\alpha_i, c)
\end{aligned}$$

Thus, for every $y \in \mathbb{H}^m$,

$$\left|\hat{\mathcal{I}}_{\mathbb{F}y}(\alpha)\right| = \begin{cases} 0 & \sum_{i=1}^m M_{y_i}^T\alpha_i \neq \vec{0} \\ \frac{|\mathbb{F}|}{|\mathbb{F}^m|} & otherwise \end{cases}$$

Assume $1 \leq i \leq m$ is such that $\alpha_i \neq \vec{0} \in GF(p)^k$. Note that for every $a_1 \neq a_2 \in \mathbb{F}$, we know that $M_{a_1}^T\alpha_i \neq M_{a_2}^T\alpha_i$ (For every $b \neq 0 \in \mathbb{F}$, $a_1 \cdot b \neq a_2 \cdot b$. Thus, for every $b \neq \vec{0} \in GF(p)^k$, $(M_{a_1} - M_{a_2})b \neq \vec{0}$ and for every $b \neq \vec{0} \in GF(p)^k$, $M_{a_1}^T b - M_{a_2}^T b = (M_{a_1} - M_{a_2})^T b \neq \vec{0}$). Hence, for every $v \in GF(p)^k$, there exists at most one $a \in \mathbb{H}$ for which $M_a^T\alpha_i = v$. In particular,

$$\Pr_{\vec{y}\in\mathbb{H}^m}\left[M_{y_i}^T\alpha_i = -\sum_{j\neq i}M_{y_j}^T\alpha_j\right] \leq \frac{1}{|\mathbb{H}|}$$

The claim follows. ■(of claim 5.3.1)

■

Remark 5.4. *To get the corollaries stated in the introduction, note the following.*

1. *Only claim 5.3.1 uses the nature/structure of $\mathbb{H}^m$.*
2. *Claim 5.3.1 does not require $\mathbb{H}$ to be a subfield of $\mathbb{F}$. Its proof holds for any subset $H \subseteq \mathbb{F}$.*
3. *If $\mathbb{F} = GF(p)$, then for any subset $S \subseteq \mathbb{F}^m$, for any $\vec{\alpha} \neq \vec{0} \in \mathbb{F}^m$,*

$$\Pr_{\vec{y} \in S}\left[\sum_{i=1}^{m} M_{y_i}^T \alpha_i \neq \vec{0}\right] = \Pr_{\vec{y} \in S}\left[\sum_{i=1}^{m} y_i \cdot \alpha_i \neq 0\right] = \Pr_{\vec{y} \in S}\left[(\vec{y}, \vec{\alpha}) \neq 0\right]$$

We have that $\min_{\vec{\alpha} \neq \vec{0} \in \mathbb{F}^m} \Pr_{\vec{y} \in S}\left[(\vec{y}, \vec{\alpha}) \neq 0\right]$ is the relative distance of the linear code obtained when using the vectors of S as the rows of a generating matrix.

5.3 Affine Subspaces Sample Well

Using the sampling lemma (Lemma 5.3), we can prove that the uniform distribution over lines in $\mathcal{S}_1^m$ samples well. Note that the sampling lemma does not show exactly this, as it considers y uniformly distributed over $\mathbb{H}^m$, instead of over $\mathbb{H}^m \setminus \{\vec{0}\}$.

Lemma 5.5. *For any $A \subseteq \mathbb{F}^m$ of density $\mu = |A|/|\mathbb{F}^m|$,*

$$\mathop{\mathbf{Var}}_{l \in \mathcal{S}_1^m}\left[\frac{|l \cap A|}{|l|}\right] \leq \frac{\mu}{|\mathbb{H}|}$$

Proof. Note that the probability that a random point in $\mathbb{F}^m$ is in A is the same as the expected fraction of points in A on a random line in $\mathcal{S}_1^m$,

$$\mathop{\mathbf{E}}_{p \in \mathcal{S}_0^m}\left[\frac{|p \cap A|}{|p|}\right] = \mathop{\mathbf{E}}_{l \in \mathcal{S}_1^m}\left[\frac{|l \cap A|}{|l|}\right] = \mu$$

but the variance may only decrease when considering lines rather than points,

$$\mathop{\mathbf{Var}}_{p \in \mathcal{S}_0^m}\left[\frac{|p \cap A|}{|p|}\right] \geq \mathop{\mathbf{Var}}_{l \in \mathcal{S}_1^m}\left[\frac{|l \cap A|}{|l|}\right]$$

Hence, since $\mathbf{Var}[X] = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right]$ and expectations satisfy that for every random variable Y and set A, $\mathbf{E}[Y] = \Pr[Y \in A] \cdot \mathbf{E}[Y|Y \in A] + \Pr[Y \notin A] \cdot \mathbf{E}[Y|Y \notin A]$,

$$\begin{aligned}
\mathop{\mathbf{Var}}_{z \in \mathbb{F}^m, y \in \mathbb{H}^m}\left[\frac{1}{|\mathbb{F}|} \cdot X_{z,y}\right] &= \frac{1}{|\mathbb{H}|^m} \cdot \mathop{\mathbf{Var}}_{p \in \mathcal{S}_0^m}\left[\frac{|p \cap A|}{|p|}\right] + \left(1 - \frac{1}{|\mathbb{H}|^m}\right) \cdot \mathop{\mathbf{Var}}_{l \in \mathcal{S}_1^m}\left[\frac{|l \cap A|}{|l|}\right] \\
&\geq \frac{1}{|\mathbb{H}|^m} \cdot \mathop{\mathbf{Var}}_{l \in \mathcal{S}_1^m}\left[\frac{|l \cap A|}{|l|}\right] + \left(1 - \frac{1}{|\mathbb{H}|^m}\right) \cdot \mathop{\mathbf{Var}}_{l \in \mathcal{S}_1^m}\left[\frac{|l \cap A|}{|l|}\right] \\
&= \mathop{\mathbf{Var}}_{l \in \mathcal{S}_1^m}\left[\frac{|l \cap A|}{|l|}\right]
\end{aligned}$$

The lemma follows from Lemma 5.3. ■

Using the analysis for dimension 1, we can bound the variance of the hitting rate for any larger dimension,

Lemma 5.6. *Fix dimensions k and m, $1 \le k \le m$. For any $A \subseteq \mathbb{F}^m$ of density $\mu = |A|/|\mathbb{F}^m|$,*

$$\operatorname*{\mathbf{Var}}_{s \in \mathcal{S}_k^m}\left[\frac{|s \cap A|}{|s|}\right] \le \frac{\mu}{|\mathbb{H}|}$$

Proof. Pick $s \in \mathcal{S}_k^m$ and additional $r \in \mathcal{R}_1^k$ independently at random. Denote by $l = \text{affine}_s(r)$ the line within s corresponding to r (the notation affine_s was introduced in section 4). By uniformity, l is uniformly distributed in $\mathcal{S}_1^m$. Hence, by Lemma 5.5 and uniformity,

$$\begin{aligned}
\operatorname*{\mathbf{Var}}_{s}\left[\frac{|s \cap A|}{|s|}\right] &= \operatorname*{\mathbf{Var}}_{s}\left[\operatorname*{\mathbf{E}}_{r}\left[\frac{|l \cap A|}{|l|}\right]\right] \\
&\le \operatorname*{\mathbf{Var}}_{s,r}\left[\frac{|l \cap A|}{|l|}\right] \\
&\le \frac{\mu}{|\mathbb{H}|}
\end{aligned}$$

■

We can now bound the deviation of the hitting rate from its expected value,

Corollary 5.7 (sampling). *Fix dimensions k and m, $1 \le k \le m$. Fix $A \subseteq \mathbb{F}^m$ of density $\mu = |A| / |\mathbb{F}^m|$. Then, for any $\varepsilon > 0$,*

$$\Pr_{s \in \mathcal{S}_k^m}\left[\left|\frac{|s \cap A|}{|s|} - \mu\right| \ge \varepsilon\right] \le \frac{\mu}{\varepsilon^2 |\mathbb{H}|}$$

Proof. Apply Lemma 5.6 and then Chebyshev's inequality. ■

5.4 Linear Subspaces Sample Well

We can similarly prove that linear subspaces with one direction chosen from $\mathbb{F}^m$ and all other directions chosen from $\mathbb{H}^m$ sample well. We will need this lemma to analyze the Randomness-Efficient Subspace vs. Point tester.

Lemma 5.8. *Fix dimensions k and m, $1 \le k < m$. Fix a set $A \subseteq \mathbb{F}^m$ of density $\mu = |A|/|\mathbb{F}^m|$. Pick uniformly $\vec{z} \in \mathbb{F}^m, \vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$, such that $\vec{z}, \vec{y}_1, \ldots, \vec{y}_k$ are linearly independent. Denote $s = \text{affine}(\vec{0}; \vec{z}, \vec{y}_1, \ldots, \vec{y}_k)$. Then,*

$$\operatorname*{\mathbf{E}}_{s}\left[\left(\frac{|s \cap A|}{|s|} - \mu\right)^2\right] \le \frac{\mu}{|\mathbb{H}|} + \frac{1}{|\mathbb{F}|}$$

Proof. Pick an additional scalar $\alpha \in \mathbb{F}$ independently at random. Let $s_\alpha = \text{affine}(\alpha\vec{z}; \vec{y}_1, \ldots, \vec{y}_k)$. Note that s_α is distributed in $\mathcal{S}_k^m$ as follows: with probability $\frac{1}{|\mathbb{F}|}$, s_α is uniformly distributed in the set of affine subspaces in $\mathcal{S}_k^m$ through the origin; with probability $1 - \frac{1}{|\mathbb{F}|}$, s_α is uniformly distributed in the set of affine subspaces in $\mathcal{S}_k^m$ that do not contain the origin. Therefore,

$$\begin{aligned}
\operatorname*{\mathbf{E}}_{s,\alpha}\left[\left(\frac{|s_\alpha \cap A|}{|s_\alpha|} - \mu\right)^2\right] &\le 1 \cdot \operatorname*{\mathbf{E}}_{s' \in \mathcal{S}_k^m}\left[\left(\frac{|s' \cap A|}{|s'|} - \mu\right)^2\right] + \frac{1}{|\mathbb{F}|} \cdot 1 \\
&= \operatorname*{\mathbf{Var}}_{s' \in \mathcal{S}_k^m}\left[\frac{|s' \cap A|}{|s'|}\right] + \frac{1}{|\mathbb{F}|}
\end{aligned}$$

By Lemma 5.6,

$$\mathop{\mathbf{E}}_{s,\alpha}\left[\left(\frac{|s_\alpha \cap A|}{|s_\alpha|} - \mu\right)^2\right] \quad \leq \quad \frac{\mu}{|\mathbb{H}|} + \frac{1}{|\mathbb{F}|}$$

By Jensen inequality,

$$\mathop{\mathbf{E}}_{s}\left[\left(\frac{|s \cap A|}{|s|} - \mu\right)^2\right] \quad \leq \quad \mathop{\mathbf{E}}_{s,\alpha}\left[\left(\frac{|s_\alpha \cap A|}{|s_\alpha|} - \mu\right)^2\right]$$

The lemma follows. ∎

We can now bound the deviation of the hitting rate from its expected value,

Corollary 5.9 (sampling). *Fix dimensions k and m, $1 \leq k < m$. Fix a set $A \subseteq \mathbb{F}^m$ of density $\mu = |A|/|\mathbb{F}^m|$. Pick uniformly $\vec{z} \in \mathbb{F}^m, \vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$, such that $\vec{z}, \vec{y}_1, \ldots, \vec{y}_k$ are linearly independent. Denote $s = \text{affine}(\vec{0}; \vec{z}, \vec{y}_1, \ldots, \vec{y}_k)$. Then, for any $\varepsilon > 0$,*

$$\Pr_s\left[\left|\frac{|s \cap A|}{|s|} - \mu\right| \geq \varepsilon\right] \leq \frac{1}{\varepsilon^2} \cdot \left(\frac{\mu}{|\mathbb{H}|} + \frac{1}{|\mathbb{F}|}\right)$$

Proof. Apply Markov inequality on Lemma 5.8. ∎

6 Consolidation

In this section we show that *weak* low degree testing claims imply *strong* low degree testing claims. Specifically, we are interested in the following (for exact definitions, see the next subsections):

1. *decoding/list decoding*: by *decoding* we refer to finding a single polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ agreeing with the oracle on many of the points. By *list-decoding* we refer to finding a short list of polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ explaining almost all the acceptance probability of a tester.

2. *consistency consolidation*: we are able to construct polynomials $Q : \mathbb{F}^m \to \mathbb{F}$ agreeing with the oracle on some fraction of the points, and wish to find polynomials agreeing with the oracle on a larger fraction of the points.

3. *degree consolidation*: we are able to construct polynomials $Q : \mathbb{F}^m \to \mathbb{F}$ of degree at most $d' \geq d$, and wish to find polynomials of degree at most d.

We call such arguments *consolidating arguments*. They are standard in the low degree testing literature (see, e.g., [3, 17, 9]), however, they require some adaptation to our new setting. In the following subsections we provide the statements and the proofs of the exact claims we need.

6.1 From Decoding to List-Decoding

If we have a way to decode, then we can list-decode by repeatedly applying decoding. In our setting, it is easy to force the decoding process to output a polynomial that differs from existing polynomials, by modifying the oracle.

Lemma 6.1 (from decoding to list-decoding). *Assume $|\mathbb{F}| \geq 4$. Fix a distribution $\mathcal{D}$ over affine subspaces of dimension $k > 0$ in $\mathbb{F}^m$. Fix a function $f : \mathbb{R} \to \mathbb{R}$, and a degree d' such that $d \leq d' \leq |\mathbb{F}| - 3$. If*
(decoding:)
for every success probability $0 < \gamma \leq 1$ *and oracle $\mathcal{A}$,*

(much consistency)

$$\mathop{\mathbf{E}}_{s\sim\mathcal{D}}\left[\Pr_{\vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right] \geq \gamma$$

implies

(a relatively-low degree polynomial that slightly agrees with the oracle)
There exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$, with $\deg Q \leq d'$, such that

$$\Pr_{\vec{x}\in\mathbb{F}^m}\left[Q(\vec{x}) = \mathcal{A}(\vec{x})\right] \geq f(\gamma)$$

Then
(list-decoding:)
for every oracle $\mathcal{A}$,

(almost all consistency is explained by a relatively short list),
Fix $\epsilon_0 \stackrel{def}{=} \sqrt{\frac{d'}{|\mathbb{F}|}}$. For every $\epsilon_0 < \delta < 1$, such that $\delta' \stackrel{def}{=} f\left(\delta - \epsilon_0\right) - \epsilon_0 \geq 2\epsilon_0$, there exists a list of $t \leq 2/\delta'$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q_i \leq d'$, such that

$$\mathop{\mathbf{E}}_{s\sim\mathcal{D}}\left[\Pr_{\vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \ \vee \ \exists i \ (Q_i \equiv \mathcal{A})(s)\right]\right] \geq 1 - \delta$$

Proof. Assume by way of contradiction that decoding holds and there exists an oracle $\mathcal{A}$ for which there exists $\epsilon_0 < \delta < 1$ satisfying $f\left(\delta - \epsilon_0\right) - \epsilon_0 \geq 2\epsilon_0$, such that there is no list-decoding for δ.

Let $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ be all polynomials of degree at most d' for which

$$\Pr_{\vec{x}\in\mathbb{F}^m}\left[Q_i(\vec{x}) = \mathcal{A}(\vec{x})\right] \geq \delta'$$

By proposition 3.5, $t \leq 2/\delta'$. By our assumption, $Q_1, \ldots, Q_t$ is not a list-decoding for δ. Note that $t \leq 1/\epsilon_0$.

When picking a subspace $s \sim \mathcal{D}$ and a point $\vec{x}$ uniformly distributed in s, define the following events:

1. C : $\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})$ *(consistent).*
2. P : $\exists i \in [t], \ \mathcal{A}(\vec{x}) = Q_i(\vec{x})$ *(point explained).*
3. S : $\exists i \in [t], \ (Q_i \equiv \mathcal{A})(s)$ *(subspace explained).*

Using this notation, the contradicting assumption implies that there is much consistency within unexplained subspaces,

$$\Pr_{s,\vec{x}}\left[C \wedge \neg S\right] = 1 - \Pr_{s,\vec{x}}\left[\neg C \vee S\right] = 1 - \mathop{\mathbf{E}}_{s}\left[\Pr_{\vec{x}}\left[\neg C \vee S\right]\right] > \delta$$

When C and P both happen, the polynomial $\mathcal{A}(s)$ agrees with a polynomial Q_i for some $i \in [t]$ on the point $\vec{x}$. Hence, by a union bound over the $i \in [t]$ and by the Schwartz-Zippel Lemma, an unexplained subspace is rarely consistent with explained points,

$$\Pr_{s,\vec{x}}[C \wedge P | \neg S] \leq \frac{td'}{|\mathbb{F}|} \leq \frac{1}{\epsilon_0} \cdot \epsilon_0^2 = \epsilon_0$$

Thus, there is much consistency on unexplained points,

$$\begin{aligned}
\Pr_{s,\vec{x}}[C \wedge \neg P] &\geq \Pr_{s,\vec{x}}[C \wedge \neg P \wedge \neg S] \\
&= \Pr_{s,\vec{x}}[C \wedge \neg S] - \Pr_{s,\vec{x}}[C \wedge P \wedge \neg S] \\
&\geq \Pr_{s,\vec{x}}[C \wedge \neg S] - \Pr_{s,\vec{x}}[C \wedge P | \neg S] \\
&> \delta - \epsilon_0
\end{aligned}$$

Pick an arbitrary polynomial $Q' : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q' = d' + 1$. Define a new oracle $\mathcal{A}'$ as follows: $\mathcal{A}'$ assigns $Q'(\vec{x})$ to all explained points $\vec{x}$, and agrees with $\mathcal{A}$ on all other affine subspaces (recall that points are affine subspaces of dimension 0). Hence,

$$\begin{aligned}
\mathop{\mathbf{E}}_{s \sim \mathcal{D}}\left[\Pr_{\vec{x} \in s}\left[\mathcal{A}'(s)(\vec{x}) = \mathcal{A}'(\vec{x})\right]\right] &\geq \Pr_{s \sim \mathcal{D}, \vec{x} \in s}\left[\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x}) \wedge \mathcal{A}(\vec{x}) = \mathcal{A}'(\vec{x})\right] \\
&\geq \Pr_{s,\vec{x}}[C \wedge \neg P] \\
&> \delta - \epsilon_0
\end{aligned}$$

Thus, by decoding, there exists a polynomial Q, $\deg Q \leq d'$, agreeing with $\mathcal{A}'$ on many of the points

$$\Pr_{\vec{x} \in \mathbb{F}^m}\left[\mathcal{A}'(\vec{x}) = Q(\vec{x})\right] \geq f\left(\delta - \epsilon_0\right)$$

The polynomials Q and Q' are necessarily distinct (they do not have the same degree). Thus, by the Schwartz-Zippel Lemma,

$$\Pr_{\vec{x} \in \mathbb{F}^m}\left[\mathcal{A}'(\vec{x}) = Q(\vec{x}) \wedge \mathcal{A}'(\vec{x}) \neq \mathcal{A}(\vec{x})\right] \leq \Pr_{\vec{x} \in \mathbb{F}^m}\left[Q'(\vec{x}) = Q(\vec{x})\right] \leq \frac{d'+1}{|\mathbb{F}|} \leq \epsilon_0$$

Hence,

$$\begin{aligned}
\Pr_{\vec{x} \in \mathbb{F}^m}\left[\mathcal{A}(\vec{x}) = Q(\vec{x}) = \mathcal{A}'(\vec{x})\right] &= \Pr_{\vec{x} \in \mathbb{F}^m}\left[\mathcal{A}'(\vec{x}) = Q(\vec{x})\right] - \Pr_{\vec{x} \in \mathbb{F}^m}\left[\mathcal{A}'(\vec{x}) = Q(\vec{x}) \wedge \mathcal{A}'(\vec{x}) \neq \mathcal{A}(\vec{x})\right] \\
&\geq f(\delta - \epsilon_0) - \epsilon_0 \\
&= \delta'
\end{aligned}$$

Therefore,

$$\Pr_{\vec{x} \in \mathbb{F}^m}\left[\mathcal{A}(\vec{x}) = Q(\vec{x})\right] \geq \Pr_{\vec{x} \in \mathbb{F}^m}\left[\mathcal{A}(\vec{x}) = Q(\vec{x}) = \mathcal{A}'(\vec{x})\right] \geq \delta'$$

Hence, there exists $i \in [t]$ such that $Q = Q_i$. However, if this is the case, by definition of $\mathcal{A}'$,

$$\delta' \leq \Pr_{\vec{x} \in \mathbb{F}^m}\left[\mathcal{A}(\vec{x}) = Q_i(\vec{x}) = \mathcal{A}'(\vec{x})\right] \leq \Pr_{\vec{x} \in \mathbb{F}^m}\left[Q'(\vec{x}) = Q(\vec{x})\right] \leq \epsilon_0$$

Contradiction. ∎

We can additionally demand that each member of the list decoding agrees with the oracle on many of the subspaces, i.e., there are no non-useful members in the list,

Lemma 6.2 (pruning the list). *Fix a distribution $\mathcal{D}$ over affine subspaces in $\mathbb{F}^m$. For every $0 < \epsilon < 1$ and oracle $\mathcal{A}$, if $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ are $t > 0$ polynomials satisfying*

(almost all consistency is explained by the list)

$$\mathop{\mathbf{E}}_{s\sim\mathcal{D}}\left[\Pr_{\vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \ \vee \ \exists i \ (Q_i \equiv \mathcal{A})(s)\right]\right] \geq 1-\delta$$

then there exists a sublist $T \subseteq [t]$, such that

1. (each polynomial agrees with the oracle on many of the subspaces)
   for every $i \in T$,
   $$\Pr_{s\sim\mathcal{D}}\left[(Q_i \equiv \mathcal{A})(s)\right] > \frac{\epsilon}{t}$$
2. (still almost all consistency is explained by the list)
   $$\mathop{\mathbf{E}}_{s\sim\mathcal{D}}\left[\Pr_{\vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \ \vee \ \exists i \in T \ (Q_i \equiv \mathcal{A})(s)\right]\right] \geq 1-\delta-\epsilon$$

Proof. We prun the given list $Q_1, \ldots, Q_t$ by throwing away any polynomial Q_i, for which the first item does not hold. In other words,

$$T \stackrel{def}{=} \left\{ i \in [t] \ \middle| \ \Pr_{s\sim\mathcal{D}}\left[(Q_i \equiv \mathcal{A})(s)\right] > \frac{\epsilon}{t} \right\}$$

By the union bound, $\Pr_{s\sim\mathcal{D}}\left[\exists i \in [t] \setminus T, \ (Q_i \equiv \mathcal{A})(s)\right] \leq t \cdot \frac{\epsilon}{t} = \epsilon$. For a subspace $s \sim \mathcal{D}$ and a point $\vec{x}$ uniformly distributed in s, define the following events:

1. C : $\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})$ *(consistent).*
2. B : $\exists i \in [t], \ (Q_i \equiv \mathcal{A})(s)$ *(explained before).*
3. N : $\exists i \in T, \ (Q_i \equiv \mathcal{A})(s)$ *(explained now).*

Using this notation, we have (e.g., by observing the appropriate Venn diagram),

$$\begin{aligned}
\mathop{\mathbf{E}}_{s}\left[\Pr_{\vec{x}}\left[\neg C \vee N\right]\right] &= \Pr_{s,\vec{x}}\left[\neg C \vee N\right] \\
&\geq \Pr_{s,\vec{x}}\left[\neg C \vee B\right] - \Pr_{s,\vec{x}}\left[B \wedge \neg N\right] \\
&\geq 1-\delta-\epsilon
\end{aligned}$$

■

6.2 Consistency Consolidation

In this subsection, we prove a lemma allowing us to deduce that a *significant consistency* γ together with a *list-decoding* for it imply that at least one of the polynomials in the list agrees with the oracle on almost γ fraction of the points. The lemma requires that the distribution over affine subspaces samples well (see section 5). Together with Lemma 6.1 that transforms decoding into list decoding, this lemma allows us to improve the consistency we manage to recover.

We phrase a rather general lemma addressing *distributional oracles*, instead of oracles. We say that $\widetilde{\mathcal{A}}$ is a *distributional oracle*, if it assigns each affine subspace s a *distribution* over functions $s \to \mathbb{F}$ (not necessarily a *single* polynomial of degree at most d over s). Our semantic even permits the distribution to produce a *null* function $\perp$ with some probability. The null function satisfies that for every subspace s, point $\vec{x} \in s$ and scalar $a \in \mathbb{F}$, the probability (over $\widetilde{\mathcal{A}}$, namely, over the randomness in choosing $\widetilde{\mathcal{A}}(s)$) that $\widetilde{\mathcal{A}}(s)(\vec{x}) = a$, when $\widetilde{\mathcal{A}}(s)$ evaluates to $\perp$, is 0.

Lemma 6.3 (from list-decoding to decoding). *Fix a distribution $\mathcal{D}$ over affine subspaces that samples well, i.e., there exists $\Delta : [0,1] \to [0,1]$, such that for every set $A \subseteq \mathbb{F}^m$, for every $0 < \varepsilon < 1$,*

$$\Pr_{s \sim \mathcal{D}}\left[\left|\frac{|s \cap A|}{|s|} - \frac{|A|}{|\mathbb{F}^m|}\right| \geq \varepsilon\right] \leq \Delta(\varepsilon)$$

Let $\mathcal{A}$ denote an oracle, and let $\widetilde{\mathcal{A}}$ denote a distributional oracle. Assume

1. (the oracles are γ-consistent)

$$\mathop{\mathbf{E}}_{\widetilde{\mathcal{A}}}\left[\mathop{\mathbf{E}}_{s \sim \mathcal{D}}\left[\Pr_{\vec{x} \in s}\left[\widetilde{\mathcal{A}}(s)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right]\right] \geq \gamma$$

2. (most consistency is explained by a relatively short list)
   There exist t functions $f_1, \ldots, f_t : \mathbb{F}^m \to \mathbb{F}$, such that,

$$\mathop{\mathbf{E}}_{\widetilde{\mathcal{A}}}\left[\mathop{\mathbf{E}}_{s \sim \mathcal{D}}\left[\Pr_{\vec{x} \in s}\left[\widetilde{\mathcal{A}}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \ \vee \ \exists i \ (f_i \equiv \widetilde{\mathcal{A}})(s)\right]\right]\right] \geq 1 - \delta$$

Then, for any $0 < \varepsilon < 1$ such that $\varepsilon \geq t \cdot \Delta(\varepsilon)$, there exists $1 \leq i \leq t$, such that

$$\Pr_{\vec{x} \in \mathbb{F}^m}\left[f_i(\vec{x}) = \mathcal{A}(\vec{x})\right] \geq \gamma - \delta - 2\varepsilon$$

Proof. Assume, by way of contradiction, that for every $1 \leq i \leq t$, $\Pr_{\vec{x} \in \mathbb{F}^m}\left[f_i(\vec{x}) = \mathcal{A}(\vec{x})\right] < \gamma - \delta - 2\varepsilon$. Let us bound the consistency towards a contradiction to the first item of the premise. For every $1 \leq i \leq t$, define the set of points explained by f_i,

$$A_i \stackrel{def}{=} \{\vec{x} \in \mathbb{F}^m \mid f_i(\vec{x}) = \mathcal{A}(\vec{x})\}$$

For every $1 \leq i \leq t$, note that $\mu_i \stackrel{def}{=} \frac{|A_i|}{|\mathbb{F}^m|} < \gamma - \delta - 2\varepsilon$.

As $\mathcal{D}$ samples well, for every $1 \leq i \leq t$, a random subspace $s \sim \mathcal{D}$ is not likely to hit A_i much more than it is expected,

$$\Pr_{s \sim \mathcal{D}}\left[\frac{|s \cap A_i|}{|s|} \geq \mu_i + \varepsilon\right] \leq \Delta(\varepsilon) \leq \frac{\varepsilon}{t}$$

By the union bound,

$$\Pr_{s \sim \mathcal{D}}\left[\exists i \in [t], \ \frac{|s \cap A_i|}{|s|} \geq \gamma - \delta - \varepsilon\right] \leq \varepsilon$$

For a random oracle assignment $\widetilde{\mathcal{A}}$, a subspace $s \sim \mathcal{D}$ and a uniformly distributed point $\vec{x} \in s$ chosen independently at random, define the following events:

1. B : $\exists i \in [t],\ |s \cap A_i| \geq (\gamma - \delta - \varepsilon) \cdot |s|$ *(bad subspace).*
2. C : $\widetilde{\mathcal{A}}(s)(\vec{x}) = \mathcal{A}(\vec{x})$ *(consistent).*
3. E : $\exists i \in [t],\ (f_i \equiv \widetilde{\mathcal{A}})(s)$ *(explained).*

Using this notation, we have established that

$$\begin{aligned}
\Pr_{\widetilde{\mathcal{A}},s,\vec{x}}[C \wedge E] &= \Pr_{\widetilde{\mathcal{A}},s,\vec{x}}[C \wedge E \wedge \neg B] + \Pr_{\widetilde{\mathcal{A}},s,\vec{x}}[C \wedge E \wedge B] \\
&\leq \Pr_{\widetilde{\mathcal{A}},s,\vec{x}}[C | E \wedge \neg B] + \Pr_{s}[B] \\
&< (\gamma - \delta - \varepsilon) + \varepsilon \\
&= \gamma - \delta
\end{aligned}$$

The second item of the premise implies

$$\begin{aligned}
\Pr_{\widetilde{\mathcal{A}},s,\vec{x}}[C] &= \Pr_{\widetilde{\mathcal{A}},s,\vec{x}}[C \wedge \neg E] + \Pr_{\widetilde{\mathcal{A}},s,\vec{x}}[C \wedge E] \\
&< \delta + (\gamma - \delta) \\
&= \gamma
\end{aligned}$$

This contradicts the first item of the premise. ∎

6.3 Degree Consolidation

Degree consolidation shows that if one reconstructs a polynomial of not too large degree that agrees with the oracle on many of our subspaces then the polynomial's true degree is, in fact, low. The reason is that the polynomial's degree does not decrease much when restricted to almost all our subspaces.

First we prove a lemma allowing us to deduce degree d if one of the *directions* of our subspaces is distributed over $\mathbb{F}^m$ (rather than $\mathbb{H}^m$). This is used only in the analysis of the Randomness-Efficient Subspace vs. Point tester.

Lemma 6.4 (degree d consolidation). *Fix dimensions k and m, $0 \leq k < m$. Fix an oracle $\mathcal{A}$ assigning polynomials of degree at most d to all affine subspaces. Suppose that a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ satisfies the following for some $0 \leq \delta \leq 1$:*

1. $\deg Q \leq \delta\,|\mathbb{F}|$.
2. *Q and $\mathcal{A}$ agree on a linear subspace chosen at random,*

$$\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m} \left[(Q \equiv \mathcal{A})(\mathrm{affine}(\vec{0}; \vec{z}, \vec{y}_1, \ldots, \vec{y}_k)) \,|\, \mathsf{ind}(\vec{z}, \vec{y}_1, \ldots, \vec{y}_k) \right] > \delta + \frac{1}{|\mathbb{F}|}$$

Then, $\deg Q \leq d$.

Proof. Assume by way of contradiction that $\deg Q > d$. Consider linearly independent $\vec{z} \in \mathbb{F}^m$ and $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$. Denote $s = \mathrm{affine}(\vec{0}; \vec{z}, \vec{y}_1, \ldots, \vec{y}_k)$, and observe the polynomial

$$Q_{|s}(\alpha_0, \alpha_1, \ldots, \alpha_k) = Q(\alpha_0 \vec{z} + \alpha_1 \vec{y}_1 + \ldots + \alpha_k \vec{y}_k)$$

Note that each of the coefficients of this polynomial can be viewed as a polynomial in $z_1, \ldots, z_m$ and $y_{1,1}, \ldots, y_{1,m}, \cdots, y_{k,1}, \ldots, y_{k,m}$ of total degree at most $\deg Q$. In particular, consider the coefficient of the degree $\deg Q$ monomial $\alpha_0^{\deg Q}$ in $Q_{|s}$. Note that it depends solely on $z_1, \ldots, z_m$ (and not on $y_{1,1}, \ldots, y_{1,m}, \cdots, y_{k,1}, \ldots, y_{k,m}$). Hence, let us denote it by $P(z_1, \ldots, z_m)$.

To analyze P we will need more notation. Denote $Q(x_1, \ldots, x_m) = \sum_{i_1 \cdots i_m} a_{i_1 \cdots i_m} x_1^{i_1} \cdots x_m^{i_m}$. Define $I \stackrel{def}{=} \left\{ (i_1, \ldots, i_m) \;\middle|\; \sum_j i_j = \deg Q \right\}$. Now, $P(z_1, \ldots, z_m) = \sum_{(i_1 \ldots i_m) \in I} a_{i_1 \cdots i_m} z_1^{i_1} \cdots z_m^{i_m}$. Thus, by definition, $\deg P = \deg Q$ and P is not identically zero.

Clearly,

$$\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m} \left[\deg Q_{|\mathrm{affine}(\vec{0}; \vec{z}, \vec{y}_1, \ldots, \vec{y}_k)} > d \,|\mathsf{ind}(\vec{z}, \vec{y}_1, \ldots, \vec{y}_k) \right] \geq \Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m} \left[P(\vec{z}) \neq 0 \,|\mathsf{ind}(\vec{z}, \vec{y}_1, \ldots, \vec{y}_k) \right]$$

By the Schwartz-Zippel Lemma, we have

$$\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m} \left[P(\vec{z}) \neq 0 \right] \geq 1 - \frac{\deg Q}{|\mathbb{F}|} \geq 1 - \delta$$

For any linearly independent $\vec{y}_1, \ldots, \vec{y}_k$, the probability that a uniformly distributed $\vec{z} \in \mathbb{F}^m$ satisfies: $\neg\mathsf{ind}(\vec{z}, \vec{y}_1, \ldots, \vec{y}_k)$ is at most $\frac{1}{|\mathbb{F}|}$. Therefore,

$$\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m} \left[\deg Q_{|\mathrm{affine}(\vec{0}; \vec{z}, \vec{y}_1, \ldots, \vec{y}_k)} > d \;\mid \mathsf{ind}(\vec{z}, \vec{y}_1, \ldots, \vec{y}_k) \right] \geq 1 - \delta - \frac{1}{|\mathbb{F}|}$$

However, $\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m} \left[\deg Q_{|\mathrm{affine}(\vec{0}; \vec{z}, \vec{y}_1, \ldots, \vec{y}_k)} \leq d \;\mid \mathsf{ind}(\vec{z}, \vec{y}_1, \ldots, \vec{y}_k) \right] > \delta + \frac{1}{|\mathbb{F}|}$. ∎

Next we prove a lemma allowing us to deduce degree md (rather than d), even if we only observe affine subspaces in $\mathcal{S}_k^m$. This lemma will be used in the analysis of the Randomness-Efficient Plane vs. Point tester.

Lemma 6.5 (degree md consolidation). *Fix dimensions k and m, $1 \leq k \leq m$. Fix an oracle $\mathcal{A}$ assigning polynomials of degree at most d to all affine subspaces. Suppose that a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ satisfies the following for some $0 \leq \delta \leq 1$:*

1. $\deg Q \leq \delta |\mathbb{F}|$.
2. $\Pr_{s \in \mathcal{S}_k^m} [(Q \equiv \mathcal{A})(s)] > \delta + \frac{1}{|\mathbb{H}|}$.

Then, $\deg Q \leq md$.

Proof. By the premise and uniformity,

$$\Pr_{\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m} \left[\Pr_{\vec{z} \in \mathbb{F}^m} [(Q \equiv \mathcal{A})(\mathrm{affine}(\vec{z}; \vec{y}_1, \ldots, \vec{y}_k))] > \delta \;\mid \mathsf{ind}(\vec{y}_1, \ldots, \vec{y}_k) \right] > \frac{1}{|\mathbb{H}|}$$

Thus,

$$\Pr_{\vec{y} \neq \vec{0} \in \mathbb{H}^m} \left[\Pr_{\vec{z} \in \mathbb{F}^m} \left[\deg Q_{|\mathrm{affine}(\vec{z}; \vec{y})} \leq d \right] > \delta \right] > \frac{1}{|\mathbb{H}|}$$

By Proposition 3.3, there exist linearly independent $\vec{y}_1, \ldots, \vec{y}_m \in \mathbb{H}^m$, such that for every $1 \leq i \leq m$,

$$\Pr_{\vec{z} \in \mathbb{F}^m} \left[\deg Q_{|\mathrm{affine}(\vec{z}; \vec{y}_i)} \leq d \right] > \delta$$

$\vec{y}_1, \ldots, \vec{y}_m$ is a basis for $\mathbb{F}^m$. Thus, every point $\vec{x} \in \mathbb{F}^m$ can be represented as $\vec{x} = \sum_{i=1}^m \alpha_i \vec{y}_i$ for some $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$. Hence, view Q as a polynomial in variables $\alpha_1, \ldots, \alpha_m$. Assume by way of contradiction that $\deg Q > md$. Hence, there exists $1 \leq i \leq m$ such that the degree of Q in the variable α_i, which we will denote by D, is larger than d. The coefficient of α_i^D in the polynomial $Q_{|\text{affine}(\vec{z};\vec{y}_i)}$ is a non-zero polynomial $P(z_1, \ldots, z_m)$ of degree at most $\deg Q$. Hence, by the Schwartz-Zippel Lemma,

$$\Pr_{\vec{z} \in \mathbb{F}^m} [P(z_1, \ldots, z_m) = 0] \leq \frac{\deg Q}{|\mathbb{F}|} \leq \delta$$

Thus, $\Pr_{\vec{z} \in \mathbb{F}^m} \left[\deg Q_{|\text{affine}(\vec{z};\vec{y}_i)} \leq d\right] \leq \delta$, which is a contradiction. ∎

7 Consistency Graph

Fix a dimension $k \geq 3$. In this section we define and analyze a *graph* that captures the consistency among *hyperplanes* in $\mathbb{F}^k$, i.e., affine subspaces of dimension $(k-1)$. Using the graph we prove a list decoding lemma (Lemma 7.4). This lemma is used in the analysis of the Randomness-Efficient Plane vs. Point tester to go up one dimension (see section 8). Lemma 7.4 is also the only lemma in this section that is used outside it.

The idea is a variation of the analysis of Raz and Safra for the non-randomness-efficient setting [17]. Our crucial observation is that we can essentially still apply their analysis *when considering only directions with coordinates in a subfield* $\mathbb{H} \subseteq \mathbb{F}$, instead of the entire field $\mathbb{F}$.

7.1 Graph Construction

Given an oracle $\mathcal{A}$ assigning affine subspaces polynomials of degree at most d, define a simple undirected graph $G_{\mathcal{A}} = (V, E_{\mathcal{A}})$ that captures the consistency among affine subspaces in $\mathcal{S}_{k-1}^k$ as follows. Let the vertices be the set of affine subspaces $V \stackrel{def}{=} \mathcal{S}_{k-1}^k$. Let the edges indicate whether two affine subspaces are assigned polynomials that are consistent on the intersection of the subspaces,

$$E_{\mathcal{A}} \stackrel{def}{=} \{(s_1, s_2) \mid \forall \vec{x} \in s_1 \cap s_2,\ \mathcal{A}(s_1)(\vec{x}) = \mathcal{A}(s_2)(\vec{x})\}$$

Note that every two subspaces in $\mathcal{S}_{k-1}^k$ are either parallel (i.e., identify or do not intersect) or intersect by an affine subspace from $\mathcal{S}_{k-2}^k$ (see closedness under intersection; Proposition 4.5).

7.2 Graph is Almost-Transitive

A graph $G = (V, E)$ is said to be *transitive*, if for every three vertices $u, v, w \in V$, if $(u, v) \in E$ and $(v, w) \in E$, then $(u, w) \in E$. In other words, a graph is transitive if and only if for every two vertices $u, w \in V$, $u \neq w$, the are not neighbors, namely, $(u, w) \notin E$, no vertex $v \in V$ neighbors both u and w, i.e., for every $v \in V$, either $(u, v) \notin E$ or $(v, w) \notin E$.

We first wish to establish that the graph is *almost-transitive* in the sense that every two vertices that are not neighbors do not have too many common neighbors (whereas, if the graph had been transitive, they would not have had common neighbors at all):

Lemma 7.1 (almost transitivity). *Fix an oracle $\mathcal{A}$ assigning affine subspaces polynomials of degree at most d. Let $G_{\mathcal{A}} = (V, E_{\mathcal{A}})$ be its corresponding consistency graph for affine subspaces of dimension $k \geq 3$. Then, for every two different affine subspaces $s_1, s_2 \in V$,*

$$(s_1, s_2) \notin E_{\mathcal{A}} \quad \Rightarrow \quad \Pr_{s_3 \in V}\left[(s_1, s_3) \in E_{\mathcal{A}} \ \wedge \ (s_3, s_2) \in E_{\mathcal{A}}\right] \leq \frac{1}{|\mathbb{H}|} + \frac{d}{|\mathbb{F}|}$$

Proof. Assume $(s_1, s_2) \notin E_{\mathcal{A}}$. By definition, there exists $\vec{x} \in s_1 \cap s_2$, for which $\mathcal{A}(s_1)(\vec{x}) \neq \mathcal{A}(s_2)(\vec{x})$. Hence, $a \stackrel{def}{=} s_1 \cap s_2 \in \mathcal{S}^k_{k-2}$ and $\mathcal{A}(s_1)$ and $\mathcal{A}(s_2)$ induce two different polynomials of degree at most d on a. Let us denote these polynomials by P_1 and P_2. Fix a representation in $\mathcal{R}^k_{k-2}$ for a. We say that a vertex $s_3 \in V$ *spots inconsistency*, if there exists $\vec{x} \in s_3 \cap a$, such that $P_1(\vec{x}) \neq P_2(\vec{x})$. We wish to argue that a random vertex $s_3 \in V$ is likely to spot inconsistency.

Pick uniformly $r = (\vec{z}; \vec{y}_1, \ldots, \vec{y}_{k-1}) \in \mathcal{R}^k_{k-1}$. Let us say that $s_3 = \text{affine}(r)$ is *bad*, if s_3 either contains a or does not intersect it. Since $(k-2) + (k-1) \geq k$, for s_3 to be bad, a's directions must be linearly dependent on $\vec{y}_1, \ldots, \vec{y}_{k-1}$. Hence, by uniformity and by Proposition 3.2,

$$\Pr_{s_3 \in V}\left[s_3 \text{ is bad}\right] \quad \leq \quad \frac{1}{|\mathbb{H}|} \tag{3}$$

By the Schwartz-Zippel Lemma, $\Pr_{\vec{x} \in a}\left[P_1(\vec{x}) \neq P_2(\vec{x})\right] \geq 1 - \frac{d}{|\mathbb{F}|}$. For all the hyperplanes s that do not contain a but do intersect it, the dimension of their intersection with a is $(k-1) + (k-2) - k = k - 3$. Let $I \stackrel{def}{=} \{s \cap a \mid s \in V; a \nsubseteq s \wedge s \cap a \neq \phi\}$. By closedness under intersection and uniformity, $\mathbf{E}_{a' \in I}\left[\Pr_{\vec{x} \in a'}\left[P_1(\vec{x}) \neq P_2(\vec{x})\right]\right] = \Pr_{\vec{x} \in a}\left[P_1(\vec{x}) \neq P_2(\vec{x})\right] \geq 1 - \frac{d}{|\mathbb{F}|}$. By uniformity,

$$\Pr_{s_3 \in V}\left[s_3 \text{ spots inconsistency} \,|\, s_3 \text{ is not bad}\right] \quad \geq \quad 1 - \frac{d}{|\mathbb{F}|} \tag{4}$$

Combining inequalities 3 and 4, we get

$$\Pr_{s_3}\left[s_3 \text{ spots inconsistency}\right] \quad \geq \quad 1 - \frac{1}{|\mathbb{H}|} - \frac{d}{|\mathbb{F}|}$$

If s_3 spots inconsistency then either $(s_1, s_3) \notin E_{\mathcal{A}}$ or $(s_3, s_2) \notin E_{\mathcal{A}}$. Thus, $(s_1, s_3) \in E_{\mathcal{A}}$ and $(s_3, s_2) \in E_{\mathcal{A}}$ with probability at most $\frac{1}{|\mathbb{H}|} + \frac{d}{|\mathbb{F}|}$. ∎

7.3 Graph-Based List Decoding

The almost-transitivity of the graph $G_{\mathcal{A}}$ can be used to prove that, other than relatively few edges, the graph is truly transitive, i.e., composed of disjoint cliques. Moreover, these cliques are relatively large. This was shown by Raz and Safra [17],

Lemma 7.2 (graph partition). *Fix $\epsilon = \frac{1}{|\mathbb{H}|} + \frac{d}{|\mathbb{F}|}$. Fix an oracle $\mathcal{A}$ assigning affine subspaces polynomials of degree at most d. Let $G_{\mathcal{A}} = (V, E_{\mathcal{A}})$ be its corresponding consistency graph for affine subspaces of dimension $k \geq 3$. Then, there exists a partition of the vertices of $G_{\mathcal{A}}$ into cliques, $V = \biguplus_{i=1}^{t} V_i$, such that*

1. *(all non-trivial cliques are large) For every $1 \leq i \leq t$, either $|V_i| = 1$, or $|V_i| > 2\sqrt{\epsilon}\,|V|$.*
2. *(almost all edges are within cliques)*

$$\Pr_{s_1, s_2 \in V}\left[(s_1, s_2) \notin E_{\mathcal{A}} \ \vee \ \exists i \ s_1, s_2 \in V_i\right] \geq 1 - 5\sqrt{\epsilon}$$

Proof. By Lemma 7.1 and the combinatorial lemma of Raz and Safra [17] (for completeness we include a proof for this lemma; see Lemma A.1 in the appendix). ∎

A large clique in $G_{\mathcal{A}}$ corresponds to a single relatively-low degree polynomial agreeing with the oracle $\mathcal{A}$ on all affine subspaces associated with the vertices in the clique,

Lemma 7.3 (from large clique to polynomial). *Fix an oracle $\mathcal{A}$ assigning affine subspaces polynomials of degree at most d. Let $G_{\mathcal{A}} = (V, E_{\mathcal{A}})$ be its corresponding consistency graph for affine subspaces of dimension $k \geq 3$. Then, for every large clique $U \subseteq V$, $|U| > \left(\frac{2d}{|\mathbb{F}|} + \frac{1}{|\mathbb{H}|}\right) \cdot |V|$, there exists a polynomial $Q : \mathbb{F}^k \to \mathbb{F}$ with $\deg Q \leq 2d$, such that for every $s \in U$, $(Q \equiv \mathcal{A})(s)$.*

Proof. For linearly independent $\vec{y}_1, \ldots, \vec{y}_{k-1} \in \mathbb{F}^k$, there are exactly $|\mathbb{F}|$ different hyperplanes of the form $\vec{z} + span\{\vec{y}_1, \ldots, \vec{y}_{k-1}\}$ for some $\vec{z} \in \mathbb{F}^k$. Let us denote their set by $\mathcal{H}[\vec{y}_1, \ldots, \vec{y}_{k-1}]$.

Pick uniformly at random linearly independent $\vec{y}_1, \ldots, \vec{y}_{k-1} \in \mathbb{H}^k$ and consider the random variable X denoting the fraction of hyperplanes in $\mathcal{H}[\vec{y}_1, \ldots, \vec{y}_{k-1}]$ that land in U.

By linearity of expectations,

$$\mathop{\mathbf{E}}_{\vec{y}_1,\ldots,\vec{y}_{k-1} \in \mathbb{H}^k \,:\, \mathsf{ind}(\vec{y}_1,\ldots,\vec{y}_{k-1})}[X] = \frac{|U|}{|V|} > \frac{2d}{|\mathbb{F}|} + \frac{1}{|\mathbb{H}|}$$

Hence, since $0 \leq X \leq 1$,

$$\Pr_{\vec{y}_1,\ldots,\vec{y}_{k-1} \in \mathbb{H}^k}\left[X > \frac{2d}{|\mathbb{F}|} \;\middle|\; \mathsf{ind}(\vec{y}_1, \ldots, \vec{y}_{k-1})\right] > \frac{1}{|\mathbb{H}|}$$

Let us say that linearly independent directions $\vec{y}_1, \ldots, \vec{y}_{k-1} \in \mathbb{H}^k$ are *good*, if the number of hyperplanes in $\mathcal{H}[\vec{y}_1, \ldots, \vec{y}_{k-1}]$ that land in U is more than $2d$.

It follows from our calculations that there are good linearly independent directions $\vec{y}^1_1, \ldots, \vec{y}^1_{k-1} \in \mathbb{H}^k$. Fix any $\vec{y}^1_k \in \mathbb{H}^k$ that is not spanned by $\vec{y}^1_1, \ldots, \vec{y}^1_{k-1}$ (such necessarily exists since $k-1 < k$). Then, there exist at least $(2d+1)$ scalars $c_0, \ldots, c_{2d} \in \mathbb{F}$ such that for every $0 \leq i \leq 2d$, we have $\text{affine}(c_i\vec{y}^1_k; \vec{y}^1_1, \ldots, \vec{y}^1_{k-1}) \in U$.

But recall that we in fact established that for uniformly distributed linearly independent $\vec{y}_1, \ldots, \vec{y}_{k-1} \in \mathbb{H}^k$, the probability that $\vec{y}_1, \ldots, \vec{y}_{k-1}$ are good is larger than $\frac{1}{|\mathbb{H}|}$. Thus (using Proposition 3.2), for uniformly distributed linearly independent $\vec{y}^2_1, \ldots, \vec{y}^2_{k-1} \in \mathbb{H}^k$, the probability that $\vec{y}^2_1, \ldots, \vec{y}^2_{k-1}$ are good and $\vec{y}^1_1 \notin span\{\vec{y}^2_1, \ldots, \vec{y}^2_{k-1}\}$, is also positive.

Therefore, there necessarily exists a basis $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^k$ for $\mathbb{F}^k$ as well as $2 \cdot (2d+1)$ scalars $c_0, \ldots, c_{2d}, c'_0, \ldots, c'_{2d} \in \mathbb{F}$ such that

$$\begin{aligned}
s_0 = \text{affine}(c_0\vec{y}_k; \vec{y}_1, \ldots, \vec{y}_{k-1}) &\in U \\
&\vdots \\
s_{2d} = \text{affine}(c_{2d}\vec{y}_k; \vec{y}_1, \ldots, \vec{y}_{k-1}) &\in U \\
s'_0 = \text{affine}(c'_0\vec{y}_1; \vec{y}_2, \ldots, \vec{y}_k) &\in U \\
&\vdots \\
s'_{2d} = \text{affine}(c'_{2d}\vec{y}_1; \vec{y}_2, \ldots, \vec{y}_k) &\in U
\end{aligned}$$

Let us define a polynomial $Q : \mathbb{F}^k \to \mathbb{F}$ such that for every $0 \leq i \leq d$, $(Q \equiv \mathcal{A})(s_i)$. This is done using Lagrange's interpolation formula:

$$Q\left(\sum_{i=1}^{k} \alpha_i \vec{y}_i\right) = \sum_{i=0}^{d} \frac{\prod_{j \in \{0,\ldots,d\}-\{i\}}(\alpha_k - c_j)}{\prod_{j \in \{0,\ldots,d\}-\{i\}}(c_i - c_j)} \cdot \mathcal{A}(s_i)\left(c_i\vec{y}_k + \sum_{j=1}^{k-1} \alpha_j \vec{y}_j\right)$$

The degree of Q in α_k is at most d and its total degree is $\deg Q \leq 2d$.

We would like to argue that for every $s \in U$, $(Q \equiv \mathcal{A})(s)$. Let $0 \leq j \leq 2d$. For every line of the form $l = \text{affine}\left(\sum_{i=1}^{k-1} a_i \vec{y}_i; \vec{y}_k\right)$ contained in s'_j, the polynomial $Q_{|l}$ has degree at most d. Moreover, for every $0 \leq i \leq d$, $Q_{|l}$ and $\mathcal{A}(s'_j)$ identify on $l \cap s_i$. By the Schwartz-Zippel Lemma, $Q_{|l}$ and $\mathcal{A}(s'_j)$ identify on the entire line l. Thus, for every $0 \leq j \leq 2d$, Q and $\mathcal{A}$ identify on s'_j. Hence, by the Schwartz-Zippel Lemma, for every $0 \leq j \leq 2d$ (and not only for every $0 \leq j \leq d$), the polynomial Q (of degree at most $2d$) and $\mathcal{A}$ identify on s_j.

Let $s \in U$. Necessarily, s intersects the s_j's or the s'_j's (or both). Hence, $Q_{|s}$ and $\mathcal{A}(s)$ identify on more than $\frac{2d}{|\mathbb{F}|}$ of the points on s. $Q_{|s}$ is of degree at most $2d$. Thus, by the Schwartz-Zippel Lemma, Q and $\mathcal{A}$ identify on s. ∎

The partition of $G_{\mathcal{A}}$ into cliques yields list decoding,

Lemma 7.4 (hyperplane vs. hyperplane). *Fix an oracle $\mathcal{A}$ assigning affine subspaces polynomials of degree at most d. Let $G_{\mathcal{A}} = (V, E_{\mathcal{A}})$ be its corresponding consistency graph for affine subspaces of dimension $k \geq 3$. Then, for any $\delta \geq 8\sqrt{\frac{d}{|\mathbb{F}|} + \frac{1}{|\mathbb{H}|}}$ there exists a list of polynomials $Q_1, \ldots, Q_t : \mathbb{F}^k \to \mathbb{F}$, $t \leq \frac{4}{\delta}$, with $\deg Q_i \leq 2d$, such that*

$$\Pr_{s_1, s_2 \in V}\left[(s_1, s_2) \notin E_{\mathcal{A}} \ \vee \ \exists i, \ (Q_i \equiv \mathcal{A})(s_1) \wedge (Q_i \equiv \mathcal{A})(s_2)\right] > 1 - \delta$$

Proof. Consider the partition of Lemma 7.2. Let $S_1, \ldots, S_l$ denote the small cliques in this partition, i.e., cliques whose size is $|S_i| < \frac{\delta}{4}|V|$. Clearly,

$$\sum_{i=1}^{l} |S_i|^2 < \frac{\delta}{4}|V| \cdot \sum_{i=1}^{l} |S_i| \leq \frac{\delta}{4}|V|^2$$

Hence, $\Pr_{s_1, s_2 \in V}\left[\exists i, \ s_1, s_2 \in S_i\right] < \frac{\delta}{4}$. Let $L_1, \ldots, L_t$ be the set of all large cliques $|L_i| \geq \frac{\delta}{4}|V|$. We have $t \leq \frac{4}{\delta}$. Moreover,

$$\Pr_{s_1, s_2 \in V}\left[(s_1, s_2) \notin E_{\mathcal{A}} \ \vee \ \exists i, \ s_1, s_2 \in L_i\right] > 1 - \frac{5}{8}\delta - \frac{1}{4}\delta > 1 - \delta$$

For every $1 \leq i \leq t$, let $Q_i : \mathbb{F}^k \to \mathbb{F}$ be the polynomial associated with L_i according to Lemma 7.3. We have $\deg Q_i \leq 2d$ and

$$\Pr_{s_1, s_2 \in V}\left[(s_1, s_2) \notin E_{\mathcal{A}} \ \vee \ \exists i, \ (Q_i \equiv \mathcal{A})(s_1) \wedge (Q_i \equiv \mathcal{A})(s_2)\right] > 1 - \delta$$

∎

Note that the lemma is meaningful only when the *density* of the graph, $|E_{\mathcal{A}}|/|V|^2$, is large enough with respect to δ, otherwise, the list might be empty. This corresponds to the fact that the oracle must assign the affine subspaces somewhat consistent polynomials if we wish to (list) decode.

8 Going Up One Dimension

Let $\mathcal{A}$ be an oracle assigning polynomials of degree at most d to affine subspaces in $\mathbb{F}^m$. Let us say that $\mathcal{A}$ is γ-*consistent* over subspaces of dimension k (we usually omit the dimensions when they are clear from the context), if

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}_k^m}\left[\Pr_{\vec{x} \in s}\left[\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right] \geq \gamma$$

Fix a dimension $k \geq 3$. Let $\mathcal{A}$ be an oracle assigning polynomials of degree at most d to affine subspaces in $\mathbb{F}^k$. In this section we prove that if $\mathcal{A}$ is γ-consistent over affine subspaces of dimension $(k-1)$ in $\mathbb{F}^k$, then there exists a polynomial $Q : \mathbb{F}^k \to \mathbb{F}$ of degree at most $2d$ that agrees with the oracle on almost γ of the points. This is done in three steps:

1. We use an argument of counting in several ways to transform our setting to one that resembles that of the consistency graph of section 7.
2. We use the analysis of the consistency graph to prove the claim we want, while losing in the consistency parameter.
3. We fix the consistency parameter via the consistency consolidation of section 6.

The final result of this section is given in Lemma 8.3. This is also the only lemma in this section that is used outside it. Note that the degree parameter grows from d to $2d$, and we need to take care of that when we use this lemma.

8.1 From Hyperplane vs. Point to Hyperplane vs. Hyperplane

We start by showing that γ-consistency over hyperplanes implies that for an average pair (s_1, s_2) of intersecting hyperplanes, $\mathcal{A}(s_1)$ and $\mathcal{A}(s_2)$ agree (with each other and with $\mathcal{A}$) on at least γ^2-fraction of the points in the intersection of s_1 and s_2.

The proof uses repeatedly the trick of counting in several ways, which is made possible due to uniformity considerations (see section 4).

For an affine subspace $a \in \mathcal{S}^k_{k-2}$, denote the set of hyperplane pairs that intersect in a by $S_a \stackrel{def}{=} \left\{(s_1, s_2) \;\middle|\; s_1, s_2 \in \mathcal{S}^k_{k-1}, s_1 \cap s_2 = a\right\}$.

Lemma 8.1 (counting in several ways). *If for an oracle $\mathcal{A}$,*

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}^k_{k-1}} \left[\Pr_{\vec{x} \in s} \left[\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x}) \right] \right] \geq \gamma$$

Then,

$$\mathop{\mathbf{E}}_{a \in \mathcal{S}^k_{k-2}} \left[\mathop{\mathbf{E}}_{(s_1,s_2) \in S_a} \left[\Pr_{\vec{x} \in a} \left[\mathcal{A}(s_1)(\vec{x}) = \mathcal{A}(\vec{x}) = \mathcal{A}(s_2)(\vec{x}) \right] \right] \right] \geq \gamma^2 - \frac{1}{|\mathbb{H}|}$$

Proof. For an affine subspace $s \in \mathcal{S}^k_{k-1}$, an affine sub-space of it $a \subset s$, $a \in \mathcal{S}^k_{k-2}$, and a point $\vec{x} \in a$, let $I_{s,a,\vec{x}}$ be the indicator variable of the event $\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})$.
By the premise and uniformity,

$$\mathop{\mathbf{E}}_{s} \left[\mathop{\mathbf{E}}_{a \subset s} \left[\mathop{\mathbf{E}}_{\vec{x} \in a} \left[I_{s,a,\vec{x}} \right] \right] \right] \geq \gamma$$

By uniformity, we can also count in a different way and obtain:

$$\mathop{\mathbf{E}}_{a} \left[\mathop{\mathbf{E}}_{\vec{x} \in a} \left[\mathop{\mathbf{E}}_{s \supset a} \left[I_{s,a,\vec{x}} \right] \right] \right] \geq \gamma$$

By convexity considerations,

$$\mathop{\mathbf{E}}_{a} \left[\mathop{\mathbf{E}}_{\vec{x} \in a} \left[\left(\mathop{\mathbf{E}}_{s \supset a} \left[I_{s,a,\vec{x}} \right] \right)^2 \right] \right] \geq \gamma^2$$

Or, in other words,

$$\mathop{\mathbf{E}}_{a}\left[\mathop{\mathbf{E}}_{\vec{x}\in a}\left[\mathop{\mathbf{E}}_{s_1,s_2\supset a}\left[I_{s_1,a,\vec{x}}I_{s_2,a,\vec{x}}\right]\right]\right] \geq \gamma^2$$

We can change the order of summation once again, and get:

$$\mathop{\mathbf{E}}_{a}\left[\mathop{\mathbf{E}}_{s_1,s_2\supset a}\left[\mathop{\mathbf{E}}_{\vec{x}\in a}\left[I_{s_1,a,\vec{x}}I_{s_2,a,\vec{x}}\right]\right]\right] \geq \gamma^2$$

The lemma follows using uniformity and noticing that the probability that $s_1 = s_2$ given that $s_1, s_2 \supset a$ is at most $\frac{1}{|\mathbb{H}|}$. ∎

8.2 Hyperplane vs. Point Lemma

Next, we show that considerable consistency between $(k-1)$-dimensional affine subspaces and points implies a significant correspondence of the values assigned to points with a relatively low degree polynomial over $\mathbb{F}^k$. The heart of the proof is the analysis of the consistency graph (Lemma 7.4).

Lemma 8.2 (hyperplane vs. point). *Assume $\mathcal{A}$ assigns polynomials of degree at most d to affine subspaces. Fix $\delta \stackrel{def}{=} 16\max\left\{\sqrt{\frac{d}{|\mathbb{F}|}}, \sqrt[4]{\frac{1}{|\mathbb{H}|}}\right\}$. Assume that*

$$\mathop{\mathbf{E}}_{s\in\mathcal{S}^k_{k-1}}\left[\Pr_{\vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right] \geq \gamma$$

Then there exists a polynomial $Q : \mathbb{F}^k \to \mathbb{F}$, with $\deg Q \leq 2d$, such that

$$\Pr_{\vec{x}\in\mathbb{F}^k}\left[Q(\vec{x}) = \mathcal{A}(\vec{x})\right] \geq \gamma^2 - 3\delta$$

Proof. Lemma 8.1 allows us to translate the consistency given in this lemma to consistency between pairs of hyperplanes on points,

$$\mathop{\mathbf{E}}_{a\in\mathcal{S}^k_{k-2}}\left[\mathop{\mathbf{E}}_{(s_1,s_2)\in S_a}\left[\Pr_{\vec{x}\in a}\left[\mathcal{A}(s_1)(\vec{x}) = \mathcal{A}(\vec{x}) = \mathcal{A}(s_2)(\vec{x})\right]\right]\right] \geq \gamma^2 - \frac{1}{|\mathbb{H}|}$$

Lemma 7.4 gives list decoding $Q_1, \ldots, Q_t : \mathbb{F}^k \to \mathbb{F}$, $\deg Q_i \leq 2d$, $t \leq \frac{4}{\delta}$, for consistency among pairs of hyperplanes. We wish to argue that at least one of these polynomials also agrees with the oracle on many of the points.

Let us define an appropriate notation. Choose independently and uniformly at random a subspace $a \in \mathcal{S}^k_{k-2}$, hyperplanes that intersect on a, $(s_1, s_2) \in S_a$, and a point $\vec{x} \in a$. Define the following events:

1. X : $\mathcal{A}(s_1)(\vec{x}) = \mathcal{A}(\vec{x}) = \mathcal{A}(s_2)(\vec{x})$ *(hyperplanes are consistent on (and with) a point).*
2. C : $(s_1, s_2) \in E_{\mathcal{A}}$ *(hyperplanes are consistent).*
3. E : $\exists i\ (Q_i \equiv \mathcal{A})(s_1) \wedge (Q_i \equiv \mathcal{A})(s_2)$ *(hyperplanes are explained).*

Using this notation we have $\Pr_{a,s_1,s_2,\vec{x}}[X] \geq \gamma^2 - \frac{1}{|\mathbb{H}|}$. By uniformity, s_1, s_2 are uniformly distributed over the set of all pairs with $s_1 \cap s_2 \in \mathcal{S}^k_{k-2}$. Since for a uniformly distributed pair $s_1, s_2 \in V$, the probability that $s_1 \cap s_2 \notin \mathcal{S}^k_{k-2}$ is bounded by $\frac{1}{|\mathbb{H}|}$ (see Proposition 3.2), the list decoding translates into

$$\Pr_{s_1,s_2}[\neg C \vee E] \geq 1 - \delta - \frac{1}{|\mathbb{H}|}$$

$\vec{x}$ is uniformly distributed within $s_1 \cap s_2$. Hence, by the Schwartz-Zippel Lemma, $\Pr_{a,s_1,s_2,\vec{x}}[X|\neg C] \leq \frac{d}{|\mathbb{F}|}$. Therefore, the probability that s_1, s_2 are consistent on $\vec{x}$ but not explained is small,

$$\begin{aligned}
\Pr_{a,s_1,s_2,\vec{x}}[X \wedge \neg E] &= \Pr_{a,s_1,s_2,\vec{x}}[C \wedge X \wedge \neg E] + \Pr_{a,s_1,s_2,\vec{x}}[\neg C \wedge X \wedge \neg E] \\
&\leq \Pr_{s_1,s_2}[C \wedge \neg E] + \Pr_{a,s_1,s_2,\vec{x}}[\neg C \wedge X] \\
&\leq 1 - \Pr_{s_1,s_2}[\neg C \vee E] + \Pr_{a,s_1,s_2,\vec{x}}[X|\neg C] \\
&\leq \delta + \frac{1}{|\mathbb{H}|} + \frac{d}{|\mathbb{F}|}
\end{aligned}$$

Thus, the probability that s_1, s_2 are consistent on $\vec{x}$ and are explained is large

$$\begin{aligned}
\Pr_{a,s_1,s_2,\vec{x}}[X \wedge E] &\geq \Pr_{a,s_1,s_2,\vec{x}}[X] - \Pr_{a,s_1,s_2,\vec{x}}[X \wedge \neg E] \\
&\geq \gamma^2 - \frac{1}{|\mathbb{H}|} - \delta - \frac{1}{|\mathbb{H}|} - \frac{d}{|\mathbb{F}|} \\
&\geq \gamma^2 - 2\delta \qquad (5)
\end{aligned}$$

Let us define a *distributional oracle* $\widetilde{\mathcal{A}}$, assigning each affine subspace $a \in \mathcal{S}^k_{k-2}$, a distribution over polynomials of degree at most d over a (for clarification of our notion of distributional oracles, see the discussion before Lemma 6.3). To define the distribution $\widetilde{\mathcal{A}}(a)$, we indicate how to sample a polynomial accordingly:

- Pick uniformly at random hyperplanes that intersect on a, $(s_1, s_2) \in S_a$.
- If there is i such that $(Q_i \equiv \mathcal{A})(s_1)$ and $(Q_i \equiv \mathcal{A})(s_2)$, output the restriction of Q_i to a (note that if there are two (or more) such polynomials, they must identify on a).
- Otherwise, output a *null* polynomial.

If $\widetilde{\mathcal{A}}(a)$ is not *null*, then there exists i such that $(Q_i \equiv \widetilde{\mathcal{A}})(a)$, while if $\widetilde{\mathcal{A}}(a)$ is *null*, $\widetilde{\mathcal{A}}(a)(\vec{x}) \neq \mathcal{A}(\vec{x})$ for every $\vec{x} \in a$. Thus,

$$\mathop{\mathbf{E}}_{\widetilde{\mathcal{A}}}\left[\mathop{\mathbf{E}}_{a \in \mathcal{S}^k_{k-2}}\left[\Pr_{\vec{x} \in a}\left[\widetilde{\mathcal{A}}(a)(\vec{x}) \neq \mathcal{A}(\vec{x}) \vee \ \exists i \ (Q_i \equiv \widetilde{\mathcal{A}})(a)\right]\right]\right] = 1$$

By the construction of $\widetilde{\mathcal{A}}$ and inequality 5, $\widetilde{\mathcal{A}}$ is $(\gamma^2 - 2\delta)$-consistent with $\mathcal{A}$,

$$\mathop{\mathbf{E}}_{\widetilde{\mathcal{A}}}\left[\mathop{\mathbf{E}}_{a \in \mathcal{S}^k_{k-2}}\left[\Pr_{\vec{x} \in a}\left[\widetilde{\mathcal{A}}(a)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right]\right] \geq \gamma^2 - 2\delta$$

By Corollary 5.7, the uniform distribution on $\mathcal{S}^k_{k-2}$ samples well: for every set $A \subseteq \mathbb{F}^k$, for every $0 < \varepsilon < 1$,

$$\Pr_{a \in \mathcal{S}^k_{k-2}}\left[\left|\frac{|a \cap A|}{|a|} - \frac{|A|}{|\mathbb{F}^k|}\right| \geq \varepsilon\right] \leq \frac{1}{\varepsilon^2 \, |\mathbb{H}|}$$

Thus, by Lemma 6.3, since $\frac{\delta}{2} \geq t \cdot \frac{4}{\delta^2 |\mathbb{H}|}$, there exists $1 \leq i \leq t$ such that

$$\Pr_{\vec{x} \in \mathbb{F}^k}\left[Q_i(\vec{x}) = \mathcal{A}(\vec{x})\right] \geq \gamma^2 - 3\delta$$

■

8.3 Consolidating

We can apply consistency consolidation to improve the result of the last subsection. The following summarizes what we establish in this section:

Lemma 8.3 (consistency consolidated). *Denote* $\theta_0 \stackrel{def}{=} 2^4 \cdot \left(\sqrt[8]{\frac{1}{|\mathbb{H}|}} + \sqrt[4]{\frac{d}{|\mathbb{F}|}}\right)$. *Fix* $k \geq 3$. *Fix an oracle* $\mathcal{A}$ *assigning polynomials of degree at most* d *to all affine subspaces. Assume that*

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}^k_{k-1}}\left[\Pr_{\vec{x} \in s}\left[\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right] \geq \gamma$$

Then there exists a polynomial $Q : \mathbb{F}^k \to \mathbb{F}$, *with* $\deg Q \leq 2d$, *such that*

$$\Pr_{\vec{x} \in \mathbb{F}^k}\left[Q(\vec{x}) = \mathcal{A}(\vec{x})\right] \geq \gamma - 2\theta_0$$

Proof. Assume $\theta_0 \leq 1$ (otherwise, the claim trivially holds). Denote $\epsilon_0 = \sqrt{\frac{2d}{|\mathbb{F}|}}$ and $\delta_0 = 16 \max\left\{\sqrt{\frac{d}{|\mathbb{F}|}}, \sqrt[4]{\frac{1}{|\mathbb{H}|}}\right\}$. Define $f(\gamma) \stackrel{def}{=} \gamma^2 - 3\delta_0$. It holds that

$$f(\theta_0 - \epsilon_0) - \epsilon_0 = (\theta_0 - \epsilon_0)^2 - 3\delta_0 - \epsilon_0 \geq \theta_0^2/2$$

where $\theta_0^2/2 \geq 2\epsilon_0$. Apply Lemma 6.1 on Lemma 8.2 to deduce the existence of $t \leq 4/\theta_0^2$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^k \to \mathbb{F}$, with $\deg Q_i \leq 2d$, such that

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}^k_{k-1}}\left[\Pr_{\vec{x} \in s}\left[\mathcal{A}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \ \vee \ \exists i (Q_i \equiv \mathcal{A})(s)\right]\right] \geq 1 - \theta_0$$

For $\varepsilon = \frac{\theta_0}{2}$, we have $\varepsilon \geq \frac{t}{\varepsilon^2 |\mathbb{H}|}$. Thus, by Lemma 6.3 (using sampling Corollary 5.7), there exists $1 \leq i \leq t$, such that

$$\Pr_{\vec{x} \in \mathbb{F}^k}\left[Q_i(\vec{x}) = \mathcal{A}(\vec{x})\right] \geq \gamma - 2\theta_0$$

■

9 The Randomness-Efficient Plane vs. Point Tester is Sound

We wish to show that if the average consistency between planes and points is large then the oracle assigns points values that are close to a low degree polynomial. Theorem 1 will follow.

Lemma 9.1 (from dimension 2 to dimension k). *Denote* $\theta_k \stackrel{def}{=} 2^4 \left(\sqrt[4]{\frac{kd}{|\mathbb{F}|}} + \sqrt[8]{\frac{1}{|\mathbb{H}|}} \right)$. *For every dimension* $k \geq 2$, *for every* $0 < \gamma \leq 1$ *and oracle* $\mathcal{A}$, *if*

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}_2^k} \left[\Pr_{\vec{x} \in s} [\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})] \right] \geq \gamma$$

then there exists a polynomial $Q : \mathbb{F}^k \to \mathbb{F}$ *with* $\deg Q \leq kd$ *such that*

$$\Pr_{\vec{x} \in \mathbb{F}^k} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - (8k - 10)\theta_k$$

Proof. We prove the lemma by induction on k. Let us formulate two inductive claims. The second argues what we wish to show. The first argues slightly better consistency, but worse degree:

$Claim_1[k]$**:** For every $0 < \gamma \leq 1$ and oracle $\mathcal{A}$, if

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}_2^k} \left[\Pr_{\vec{x} \in s} [\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})] \right] \geq \gamma$$

then there exists a polynomial $Q : \mathbb{F}^k \to \mathbb{F}$ with $\deg Q \leq 2(k-1)d$ such that

$$\Pr_{\vec{x} \in \mathbb{F}^k} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - (8k - 16)\theta_k$$

$Claim_2[k]$**:** For every $0 < \gamma \leq 1$ and oracle $\mathcal{A}$, if

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}_2^k} \left[\Pr_{\vec{x} \in s} [\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})] \right] \geq \gamma$$

then there exists a polynomial $Q : \mathbb{F}^k \to \mathbb{F}$ with $\deg Q \leq kd$ such that

$$\Pr_{\vec{x} \in \mathbb{F}^k} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - (8k - 10)\theta_k$$

$Claim_1[2]$ holds by taking Q to be $\mathcal{A}(s)$ for the only plane s. Hence, the lemma will follow if we prove that for every $k \geq 2$,

$$Claim_1[k] \;\Rightarrow\; Claim_2[k] \;\Rightarrow\; Claim_1[k+1]$$

Claim 9.1.1. $Claim_1[k] \;\Rightarrow\; Claim_2[k]$

Proof. Fix $0 < \gamma \leq 1$ and oracle $\mathcal{A}$ such that

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}_2^k} \left[\Pr_{\vec{x} \in s} [\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})] \right] \geq \gamma$$

Assume that $(8k-10)\theta_k \leq 1$ (otherwise, we are done). Denote $\epsilon_0 = \sqrt{2(k-1)d/|\mathbb{F}|}$. Define $f(\gamma) \stackrel{def}{=} \gamma - (8k-16)\theta_k$. Let $\delta = (8k-14)\theta_k$. It holds that

$$f(\delta - \epsilon_0) - \epsilon_0 = (8k-14)\theta_k - \epsilon_0 - (8k-16)\theta_k - \epsilon_0 \geq \theta_k$$

where $\theta_k \geq 2\epsilon_0$. By lemmata 6.1 and 6.2 applied on $Claim_1[k]$, there exist $t \leq 2/\theta_k$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^k \to \mathbb{F}$, $\deg Q_i \leq 2(k-1)d$, such that

1. *(each agrees with many planes)* for every $1 \leq i \leq t$,

$$\Pr_{s \in \mathcal{S}_2^k}[(Q_i \equiv \mathcal{A})(s)] \geq \frac{\theta_k}{t} > \frac{2(k-1)d}{|\mathbb{F}|} + \frac{1}{|\mathbb{H}|}$$

2. *(all explain almost all the consistency)*

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}_2^k}\left[\Pr_{\vec{x} \in s}[\mathcal{A}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \vee \exists i \ (Q_i \equiv \mathcal{A})(s)]\right] \geq 1 - \delta - \theta_k$$

By Lemma 6.5, for every $1 \leq i \leq t$, $\deg Q_i \leq kd$. By Lemma 6.3 (using sampling Corollary 5.7), there exists $1 \leq i \leq t$ such that

$$\Pr_{\vec{x} \in \mathbb{F}^k}[Q_i(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - (8k-10)\theta_k$$

■(of claim 9.1.1)

Claim 9.1.2. $Claim_2[k] \ \Rightarrow \ Claim_1[k+1]$

Proof. Fix $0 < \gamma \leq 1$ and oracle $\mathcal{A}$ such that

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}_2^{k+1}}\left[\Pr_{\vec{x} \in s}[\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})]\right] \geq \gamma$$

Let $s \in \mathcal{S}_k^{k+1}$. Define an oracle relative to s, $\mathcal{A}_s$, as follows: for every affine subspace $s' = \text{affine}(r)$ in $\mathbb{F}^k$ (this includes the points in $\mathbb{F}^k$), let $\mathcal{A}_s(s') \stackrel{def}{=} \mathcal{A}(\text{affine}_s(r))$ (the notation affine_s was introduced in section 4). Let the *consistency within* s be

$$\gamma_s \stackrel{def}{=} \mathop{\mathbf{E}}_{s' \in \mathcal{S}_2^k}\left[\Pr_{\vec{x} \in s'}\left[\mathcal{A}_s(s')(\vec{x}) = \mathcal{A}_s(\vec{x})\right]\right]$$

By uniformity, the *average* consistency within $s \in \mathcal{S}_k^{k+1}$ is large,

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}_k^{k+1}}[\gamma_s] = \mathop{\mathbf{E}}_{s' \in \mathcal{S}_2^{k+1}}\left[\Pr_{\vec{x} \in s'}\left[\mathcal{A}(s')(\vec{x}) = \mathcal{A}(\vec{x})\right]\right] \geq \gamma$$

$Claim_2[k]$ implies the existence of a new oracle $\mathcal{A}'$ that assigns each hyperplane $s \in \mathcal{S}_k^{k+1}$ a polynomial of degree at most kd that agrees with $\mathcal{A}$ on at least $\gamma_s - (8k-10)\theta_k$ of its points. It holds that

$$\mathop{\mathbf{E}}_{s \in \mathcal{S}_k^{k+1}}\left[\Pr_{\vec{x} \in s}\left[\mathcal{A}'(s)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right] \geq \mathop{\mathbf{E}}_{s \in \mathcal{S}_k^{k+1}}[\gamma_s - (8k-10)\theta_k] \geq \gamma - (8k-10)\theta_k$$

By Lemma 8.3, there exists a polynomial $Q : \mathbb{F}^{k+1} \to \mathbb{F}$ with $\deg Q \leq 2kd$ such that

$$\Pr_{\vec{x}\in\mathbb{F}^{k+1}} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - (8(k+1) - 16)\theta_{k+1}$$

■(of claim 9.1.2)

Lemma 9.1 follows by induction. ■

The soundness of the Randomness-Efficient Plane vs. Point tester easily follows:

Proof. (of Theorem 1) Assume that

$$\Pr_{\vec{z}\in\mathbb{F}^m,\vec{y}_1,\vec{y}_2\in\mathbb{H}^m} \left[PlanePoint^{\mathcal{A}}(\vec{z},\vec{y}_1,\vec{y}_2)\right] = \gamma$$

The probability that $\vec{y}_1, \vec{y}_2$ are linearly dependent is at most $\frac{1}{|\mathbb{H}|^m} + \frac{1}{|\mathbb{H}|^{m-1}} \leq \frac{2}{|\mathbb{H}|}$. Thus,

$$\mathop{\mathbf{E}}_{s\in\mathcal{S}_2^m}\left[\Pr_{\vec{x}\in s} [\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})]\right] \geq \gamma - \frac{2}{|\mathbb{H}|}$$

By Lemma 9.1, we have decoding: there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq md$, such that

$$\Pr_{\vec{x}\in\mathbb{F}^m} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - \varepsilon$$

By Lemma 6.1, we have list-decoding: for every δ, $\delta > 2\varepsilon$, there exist $t \leq 2/\delta$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q_i \leq md$, such that

$$\mathop{\mathbf{E}}_{s\in\mathcal{S}_2^m}\left[\Pr_{\vec{x}\in s} [\mathcal{A}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \ \vee \ \exists i \ (Q_i \equiv \mathcal{A})(s)]\right] \geq 1 - \delta - 2\varepsilon + \frac{2}{|\mathbb{H}|}$$

Therefore,

$$\Pr_{\vec{z}\in\mathbb{F}^m,\vec{y}_1,\vec{y}_2\in\mathbb{H}^m} \left[\neg PlanePoint^{\mathcal{A}}(\vec{z},\vec{y}_1,\vec{y}_2) \ \vee \ \exists i \ (Q_i \equiv \mathcal{A})(\text{affine}(\vec{z};\vec{y}_1,\vec{y}_2))\right] \geq 1 - \delta - 2\varepsilon$$

■

10 The Randomness-Efficient Subspace vs. Point Tester is Sound

In this section we use the result from the previous section, namely, the soundness of the Randomness-Efficient Plane vs. Point tester, to prove the soundness of the Subspace vs. Point tester.

Consider the distribution $\mathcal{D}$ over three-dimensional affine subspaces induced by the tester: pick uniformly $\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m$, such that $\vec{z}, \vec{y}_1, \vec{y}_2$ are linearly independent, and output $\text{affine}(\vec{0}; \vec{z}, \vec{y}_1, \vec{y}_2)$.

Lemma 10.1 (from Plane vs. Point to Subspace vs. Point). *Fix dimension $m \geq 3$. Fix $\varepsilon \stackrel{def}{=} 2^7 m \left(\sqrt[4]{\frac{md}{|\mathbb{F}|}} + \sqrt[8]{\frac{1}{|\mathbb{H}|}}\right)$. If an oracle $\mathcal{A}$ assigning polynomials of degree at most d to affine subspaces satisfies*

$$\mathop{\mathbf{E}}_{s\sim\mathcal{D}}\left[\Pr_{\vec{x}\in s} [\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})]\right] \geq \gamma$$

then there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq md$ such that

$$\Pr_{\vec{x}\in\mathbb{F}^m} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - \varepsilon$$

Proof. Let us construct a new oracle $\mathcal{A}'$. For every plane $p \in \mathcal{S}_2^m$ that does not contain the origin, let $\mathcal{A}'(p)$ be the restriction of $\mathcal{A}(s)$ to p, where s is the unique three-dimensional linear subspace that contains p. Let $\mathcal{A}'$ identify with $\mathcal{A}$ on all other affine subspaces.

For a subspace $s \sim \mathcal{D}$, $s = \text{affine}(\vec{0}; \vec{z}, \vec{y}_1, \vec{y}_2)$, and a random scalar $\alpha \in \mathbb{F}$, let $s_\alpha = \text{affine}(\alpha\vec{z}; \vec{y}_1, \vec{y}_2)$. Clearly, the premise implies that

$$\mathop{\mathbf{E}}_{s,\alpha}\left[\Pr_{\vec{x}\in s_\alpha}\left[\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right] \geq \gamma$$

The plane s_α is distributed as follows: with probability $\frac{1}{|\mathbb{F}|}$, s_α is uniformly distributed within the planes in $\mathcal{S}_2^m$ that contain the origin; with probability $1 - \frac{1}{|\mathbb{F}|}$, s_α is uniformly distributed within the planes in $\mathcal{S}_2^m$ that do not contain the origin.

Hence, noticing that a uniformly distributed plane in $\mathcal{S}_2^m$ contains the origin with probability $\frac{|\mathbb{F}|^2}{|\mathbb{F}|^m} \leq \frac{1}{|\mathbb{F}|}$,

$$\mathop{\mathbf{E}}_{p\in\mathcal{S}_2^m}\left[\Pr_{\vec{x}\in p}\left[\mathcal{A}'(p)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right] \geq \gamma - \frac{1}{|\mathbb{F}|}$$

By Lemma 9.1, there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq md$ such that

$$\Pr_{\vec{x}\in\mathbb{F}^m}\left[Q(\vec{x}) = \mathcal{A}(\vec{x})\right] \geq \gamma - \varepsilon$$

■

Now we can apply degree consolidation and get

Lemma 10.2 (degree consolidated). *Fix dimension $m \geq 3$. Fix $\varepsilon \stackrel{def}{=} 2^7 m\left(\sqrt[4]{\frac{md}{|\mathbb{F}|}} + \sqrt[8]{\frac{1}{|\mathbb{H}|}}\right)$. If an oracle $\mathcal{A}$ assigning polynomials of degree at most d to affine subspaces satisfies*

$$\mathop{\mathbf{E}}_{s\sim\mathcal{D}}\left[\Pr_{\vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})\right]\right] \geq \gamma$$

then there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq d$ such that

$$\Pr_{\vec{x}\in\mathbb{F}^m}\left[Q(\vec{x}) = \mathcal{A}(\vec{x})\right] \geq \gamma - 2\varepsilon$$

Proof. Assume $\varepsilon \leq \frac{1}{2}$ (otherwise, we are done). Denote $\epsilon_0 = \sqrt{\frac{md}{|\mathbb{F}|}}$, $\delta = 1.5\varepsilon - \epsilon_0$.

Applying Lemma 6.1 and Lemma 6.2 on Lemma 10.1, we know that there exist $t \leq 8/\varepsilon$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q_i \leq md$, such that

1. *(each agrees with many planes)* for every $1 \leq i \leq t$,

$$\Pr_{s\sim\mathcal{D}}\left[(Q_i \equiv \mathcal{A})(s)\right] > \frac{\epsilon_0}{t} \geq \frac{md}{|\mathbb{F}|} + \frac{1}{|\mathbb{F}|}$$

2. *(all explain almost all the consistency)*

$$\mathop{\mathbf{E}}_{s\sim\mathcal{D}}\left[\Pr_{\vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \vee \exists i(Q_i \equiv \mathcal{A})(s)\right]\right] \geq 1 - \delta - \epsilon_0$$

By Lemma 6.4, for every $1 \leq i \leq t$, $\deg Q_i \leq d$. By Corollary 5.9, $\mathcal{D}$ samples well: for every set $A \subseteq \mathbb{F}^m$, for every $0 < \epsilon < 1$,

$$\Pr_{s \sim \mathcal{D}} \left[\left| \frac{|s \cap A|}{|s|} - \frac{|A|}{|\mathbb{F}^m|} \right| \geq \epsilon \right] \leq \frac{1}{\epsilon^2} \cdot \left(\frac{1}{|\mathbb{H}|} + \frac{1}{|\mathbb{F}|} \right)$$

Hence, by Lemma 6.3, there exists $1 \leq i \leq t$, such that

$$\Pr_{\vec{x} \in \mathbb{F}^m} [Q_i(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - 2\varepsilon$$

■

Our main theorem stating the soundness of the Randomness Efficient Subspace vs. Point tester follows:

Proof. (of Theorem 2) Assume that

$$\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m} \left[SpacePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2) \right] = \gamma$$

The probability that $\vec{z}, \vec{y}_1, \vec{y}_2$ are linearly dependent is very small,

$$\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m} [\neg \mathsf{ind}(\vec{z}, \vec{y}_1, \vec{y}_2)] \leq \frac{1}{|\mathbb{H}|^m} + \frac{1}{|\mathbb{H}|^{m-1}} + \frac{1}{|\mathbb{F}|^{m-2}} \leq \frac{2}{|\mathbb{H}|} + \frac{1}{|\mathbb{F}|}$$

Hence,

$$\mathop{\mathbf{E}}_{s \sim \mathcal{D}} \left[\Pr_{\vec{x} \in s} [\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})] \right] \geq \gamma - \frac{2}{|\mathbb{H}|} - \frac{2}{|\mathbb{F}|}$$

By Lemma 10.2 we have decoding: there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq d$ such that

$$\Pr_{\vec{x} \in \mathbb{F}^m} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - 2.5\varepsilon$$

Lemma 6.1 applied on Lemma 10.2 gives list-decoding: there exist $t \leq 2/\delta$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q_i \leq d$ such that

$$\mathop{\mathbf{E}}_{s \sim \mathcal{D}} \left[\Pr_{\vec{x} \in s} [\mathcal{A}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \ \vee \ \exists i \ (Q_i \equiv \mathcal{A})(s)] \right] \geq 1 - \delta - 2.75\varepsilon$$

Therefore,

$$\Pr_{\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m} \left[\neg SpacePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2) \ \vee \ \exists i \ (Q_i \equiv \mathcal{A})(\mathrm{affine}(\vec{0}; \vec{z}, \vec{y}_1, \vec{y}_2)) \right] \geq 1 - \delta - 3\varepsilon$$

■

Acknowledgements

We are grateful to Muli Safra for many discussions. We would also like to thank Ariel Gabizon, Amir Yehudayoff and Igor Shparlinski for their suggestions. We thank Ariel Gabizon for allowing us to include his observations regarding our sampling lemma. We also wish to thank the anonymous referees for faithfully doing their job.

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998.

[2] S. Arora and S. Safra. Probabilistic checking of proofs: a new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998.

[3] S. Arora and M. Sudan. Improved low-degree testing and its applications. *Combinatorica*, 23(3):365–426, 2003.

[4] L. Babai, L. Fortnow, and C. Lund. Nondeterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1:3–40, 1991.

[5] E. Ben-Sasson, O. Goldreich, P. Harsha, M. Sudan, and S. Vadhan. Robust PCPs of proximity, shorter pcps and applications to coding. In *Proc. 36th ACM Symp. on Theory of Computing*, pages 1–10, 2004.

[6] E. Ben-Sasson and M. Sudan. Simple PCPs with poly-log rate and query complexity. In *Proc. 37th ACM Symp. on Theory of Computing*, pages 266–275, 2005.

[7] E. Ben-Sasson, M. Sudan, S. P. Vadhan, and A. Wigderson. Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. In *Proc. 34th ACM Symp. on Theory of Computing*, pages 612–621, 2003.

[8] I. Dinur. The PCP theorem by gap amplification. In *Proc. 38th ACM Symp. on Theory of Computing*, 2006.

[9] I. Dinur, E. Fischer, G. Kindler, R. Raz, and S. Safra. PCP characterizations of NP: Towards a polynomially-small error-probability. In *Proc. 31st ACM Symp. on Theory of Computing*, pages 29–40, 1999.

[10] U. Feige, S. Goldwasser, L. Lovasz, S. Safra, and M. Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM*, 43(2):268–292, 1996.

[11] K. Ford. The distribution of integers with a divisor in a given interval. Technical Report arXiv:math/0401223, 2004.

[12] K. Friedl and M. Sudan. Some improvements to low degree tests. In *3rd Israel symp. on Theory and Computing Systems*, 1995.

[13] O. Goldreich and M. Sudan. Locally testable codes and PCPs of almost-linear length. In *Proc. 43rd IEEE Symp. on Foundations of Computer Science*, pages 13–22, 2002.

[14] R. Hall and G. Tenenbaum. *Divisors*, volume 90 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 1988.

[15] D. Moshkovitz and R. Raz. Sub-constant error probabilistically checkable proof of almost-linear size. Technical Report TR07-026, Electronic Colloquium on Computational Complexity, 2007.

[16] A. Polishchuk and D. A. Spielman. Nearly-linear size holographic proofs. In *Proc. 26th ACM Symp. on Theory of Computing*, pages 194–203, 1994.

[17] R. Raz and S. Safra. A sub-constant error-probability low-degree test and a sub-constant error-probability PCP characterization of NP. In *Proc. 29th ACM Symp. on Theory of Computing*, pages 475–484, 1997.

[18] R. Rubinfeld and M. Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.

A Combinatorial Lemma

Let us prove the lemma of Raz and Safra [17] that we use. First, let us introduce several notations. Given a graph $G = (V, E)$ and a vertex $v \in V$, the *neighbors* of v are $\mathcal{N}_G(v) \stackrel{def}{=} \{u \in V \mid (v, u) \in E\}$. The *degree* of v is $d_G(v) \stackrel{def}{=} |\mathcal{N}_G(v)|$. The *connected component* of v is $C_G(v) \stackrel{def}{=} \{u \in V \mid u \text{ is reachable from } v\}$. The non-neighbors of v within its connected component are denoted $\mathcal{D}_G(v) \stackrel{def}{=} C_G(v) \setminus (\{v\} \cup \mathcal{N}_G(v))$.

Lemma A.1 (graph partition, [17]). *Let $G = (V, E)$ be an undirected graph in which every two non-neighbors $u, v \in V$, $(u, v) \notin E$, have at most $\epsilon |V|$ common neighbors. Then, there exists a partition of the vertices into cliques, $V = \biguplus_{i=1}^{t} V_i$, such that*

1. *(all non-trivial cliques are large) For every $1 \leq i \leq t$, either $|V_i| = 1$, or $|V_i| > 2\sqrt{\epsilon}\,|V|$.*
2. *(almost all edges are within cliques)*

$$\Pr_{u,v \in V}[(u, v) \notin E \ \vee \ \exists i \ u, v \in V_i] \geq 1 - 5\sqrt{\epsilon}$$

Proof. Consider the following operation on graphs, meant to improve transitivity by removing some edges:

Pick a vertex $v \in V$.

1. If $d_G(v) \leq 2\sqrt{\epsilon}\,|V|$, remove all the edges that touch v.
2. If $d_G(v) > 2\sqrt{\epsilon}\,|V|$, remove all edges between neighbors of v and non-neighbors of v (these edges are necessarily within v's connected component).

If there is no vertex for which this operation causes removal of edges, then the graph is necessarily transitive, and, moreover, all its non-trivial cliques are of size more than $2\sqrt{\epsilon}\,|V|$.

Hence, iteratively perform this operation, picking each time an arbitrary vertex for which edges would be removed, until this is no longer possible. Let $v_1, v_2, \ldots, v_l$ denote the picked (not necessarily distinct) vertices. Let $G_1, G_2, \ldots, G_l$ denote the subgraphs obtained in the l iterations. Let I_1 be the set of all indices $1 \leq i \leq l$ in which step 1 was performed. Let I_2 be the set of all indices $1 \leq i \leq l$ in which step 2 was performed.

We will bound the total number of edges removed. Observe that if step 1 is performed for a vertex v_i, then its connected component becomes a singleton. Thus, $|I_1| \leq |V|$, and we have

$$\sum_{i \in I_1} |\mathcal{N}_{G_i}(v_i)| \leq \sum_{i \in I_1} 2\sqrt{\epsilon}\,|V| = |I_1| \cdot 2\sqrt{\epsilon}\,|V| \leq 2\sqrt{\epsilon}\,|V|^2$$

Observe that if step 2 is performed for a vertex v_i, then after the i'th operation, the vertices of $\mathcal{N}_{G_i}(v_i)$ and the vertices of $\mathcal{D}_{G_i}(v_i)$ do not belong to the same connected component. Thus, $\sum_{i\in I_2} |\mathcal{D}_{G_i}(v_i)| \cdot |\mathcal{N}_{G_i}(v_i)| \leq |V|^2$ (no pair of vertices appears twice in this sum). By the almost-transitivity, for every $i \in I_2$, every vertex $u \in \mathcal{D}_{G_i}(v_i)$ has at most $\epsilon |V|$ neighbors in $\mathcal{N}_{G_i}(v_i)$ (each is a common neighbor of u and v_i). Therefore, we can bound the total number of edges removed in step 2 by

$$\sum_{i\in I_2} |\mathcal{D}_{G_i}(v_i)| \cdot \epsilon |V| < \sum_{i\in I_2} |\mathcal{D}_{G_i}(v_i)| \cdot \epsilon \cdot \frac{|\mathcal{N}_{G_i}(v_i)|}{2\sqrt{\epsilon}} \leq \frac{\sqrt{\epsilon}}{2} \cdot \sum_{i\in I_2} |\mathcal{D}_{G_i}(v_i)| \cdot |\mathcal{N}_{G_i}(v_i)| \leq \frac{\sqrt{\epsilon}}{2} |V|^2$$

Therefore, the total number of edges removed is at most $2.5\sqrt{\epsilon}|V|^2$ and the total number of pairs $u, v \in V$ for which $(u, v) \in E$ but u and v are not in the same clique is at most $5\sqrt{\epsilon}|V|^2$. ■