Sub-Constant Error Low Degree Test of Almost-Linear Size

Dana Moshkovitz*
Department of Computer Science
and Applied Mathematics
The Weizmann Institute, Rehovot, Israel
dana.moshkovitz@weizmann.ac.il

Ran Raz†
Department of Computer Science
and Applied Mathematics
The Weizmann Institute, Rehovot, Israel
ran.raz@weizmann.ac.il

ABSTRACT

Given 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 a 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) and locally testable codes (LTCs). The error of the low degree tester is related to the soundness of the PCP and its size is related to the size of the PCP (or the length of the LTC).

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* [3, 16]. These testers enabled the construction of PCPs with *sub-constant soundness*, but *polynomial size* [3, 16, 9]. Previous constructions of *almost-linear size* testers obtained only *constant error* [13, 7]. These testers were used to construct *almost-linear size* LTCs and *almost-linear size* PCPs with *constant soundness* [13, 7, 5, 6, 8].

Categories and Subject Descriptors

F.2.2 [**Nonnumerical Algorithms and Problems**]: Complexity of proof procedures; E.4 [**Coding and Information Theory**]; F.2.1 [**Numerical Algorithms and Problems**]: Computations on polynomials, Computations in finite fields.

General Terms

Theory.

Keywords

Low degree testing, Plane vs. Point test, Probabilistically Checkable Proofs, Locally Testable Codes.

1. INTRODUCTION

1.1 Low Degree Testing

Let $\mathbb{F}$ be a finite field, let m be a dimension and let d be a degree. [A particular setting of parameters to have in mind is the one used in construction of PCPs and LTCs: a large field $\mathbb{F}$, and a fairly large degree d, which is, nonetheless, considerably smaller than $|\mathbb{F}|$; specifically, $\frac{md}{|\mathbb{F}|} \leq o(1)$, but $md \geq |\mathbb{F}|^{1-o(1)}$].

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.

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 be likely to reject.

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.). The convention is that these restrictions in themselves are polynomials of total degree at most d over the subspaces. The tester is required to satisfy:

- *Completeness:* If $f \in \mathcal{P}$, there is an oracle $\mathcal{A}$ that makes the tester accept with prob. 1.
- *Soundness:* If $\mathsf{agr}(f, \mathcal{P})$ is small, then for every oracle $\mathcal{A}$, the tester is not likely to accept.

*Supported by grant 263/02 of the Israel Science Foundation (ISF) and a grant of the Binational Science Foundation (BSF).
†Supported by a grant of the Israel Science Foundation (ISF) and a grant of the Binational Science Foundation (BSF).

STOC'06, May 21–23, 2006, Seattle, Washington, USA.

Rubinfeld and Sudan [17] 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., $\mathcal{A}(l)(\vec{x}) = f(\vec{x})$.

Low degree testers enabled the construction of *Probabilistically Checkable Proofs* (PCPs) [4, 10, 2, 1] and *Locally Testable Codes* (LTCs) [12, 13], hence their great importance. These applications motivated further improvements to low degree testing.

Specifically, the following parameters were of interest:

1. **Queries:** How many *queries* does the tester make? (should be a constant; preferably 2).

   and two more parameters to be thoroughly discussed in the next subsections:

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

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 [17, 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, 16] revealed that there are low degree testers with a *sub-constant error*. Specifically, [3, 16] 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 very algebraic in nature. Raz and Safra [16] proved a sub-constant error for a slightly different tester, considering planes that intersect by 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 soundness* [3, 16, 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 Sudans' Line vs. Point tester [17] 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 $\mathbb{F}$.

Alternatively, we refer to the *randomness* of the tester, which is the amount of random bits that the tester requires. For instance, to pick a random line and a random point within it, we merely have to pick a random point and a random direction in $\mathbb{F}^m$. Hence, the randomness of the Line vs. Point tester [17] is $2m \log |\mathbb{F}|$.

The size of a tester is measured with respect to $n = |\mathbb{F}|^m$. For instance, the size of the Line vs. Point tester [17] is quadratic $n^{2+o(1)}$. 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 within $\mathbb{F}^m$. They showed that the error of this tester is bounded away from 1. Unfortunately, their elegant analysis is inherently applicable only for acceptance probability $\gamma > \frac{1}{2}$.

The work of [13, 7] gave rise to explicit constructions of *almost-linear size* LTCs and PCPs with *constant soundness* [13, 7, 5]. The recent work of Dinur [8] also depicts *almost-linear size* LTCs and PCPs with *constant soundness*, 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 PCPs and LTCs 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*. Potential applications of our constructions are constructions of locally testable codes and PCPs with *sub-constant soundness* and *almost-linear size* (and a constant number of queries).

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 $\mathbb{F}^m$. The field structure of $\mathbb{H}$ allows us to use the combinatorial approach of Raz and Safra [16], 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 [16], 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 1: Randomness-Efficient Plane vs. Point

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 [16]. 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 [16] 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 $m \log |\mathbb{F}| + polylog(m, \log |\mathbb{F}|)$ 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 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, 16], 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 a *relatively low* 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 2: Randomness-Efficient Subspace vs. Point

This tester uses the same number of random bits as the Randomness-Efficient Plane vs. Point tester, namely, it uses $m \log |\mathbb{F}| + 2m \log |\mathbb{H}|$ bits, 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 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. 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 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 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. Thereof, we only consider the first. Assume that the Randomness-Efficient Plane vs. Point tester, given access to 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 $\gamma - \varepsilon$ of the points, for $\varepsilon \leq const \cdot m \left(\sqrt[8]{\frac{1}{|\mathbb{H}|}} + \sqrt[4]{\frac{md}{|\mathbb{F}|}} \right)$.

1.4.1 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 degree 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 $0 \leq \gamma \leq 1$,

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

1.4.2 Main idea

Our proof is by induction on the dimension k. 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 $\vec{\alpha} \in \mathbb{F}^k \leftrightarrow \vec{z} + \sum_{i=1}^k \alpha_i \vec{y}_i \in s$), 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:

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

1.4.3 Proving (2)

By an adaptation of an idea by Raz and Safra [16], 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 $0\leq\gamma\leq 1$,

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

The idea of Raz and Safra [16] 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 [16], 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.

1.5 Organization

We state the main theorems regarding the completeness and 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. Our main theorems are proved in section 9.

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 a linear 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 a polynomial does not depend on the representation.

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 $\vec{y}_1,\vec{y}_2\in\mathbb{H}^m$ chosen uniformly at random, 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 $\gamma - \varepsilon$ 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 success, but $\delta + \varepsilon$ (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 $\mathcal{A}$ and every* success probability $0 < \gamma \leq 1$*, satisfying*

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

The following hold:

1. ***(Decoding)*** *There exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq md$, such that*
$$\Pr_{\vec{x} \in \mathbb{F}^m} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - \varepsilon$$
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 with probability at least $1 - \delta - 2\varepsilon$ over the choice of $\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m$, either the test $PlanePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2)$ fails, or there exists $i \in [t]$, such that $(Q_i \equiv \mathcal{A})(\text{affine}(\vec{z}; \vec{y}_1, \vec{y}_2))$.*

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 $\mathcal{A}$ and every* success probability $0 < \gamma \leq 1$*, satisfying*

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

The following hold:

1. ***(Decoding)*** *There exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq d$, such that*
$$\Pr_{\vec{x} \in \mathbb{F}^m} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma - 3\varepsilon$$
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 with probability at least $1 - \delta - 3\varepsilon$ over the choice of $\vec{z} \in \mathbb{F}^m, \vec{y}_1, \vec{y}_2 \in \mathbb{H}^m$, either the test $SpacePoint^{\mathcal{A}}(\vec{z}, \vec{y}_1, \vec{y}_2)$ fails, or there exists $i \in [t]$, such that $(Q_i \equiv \mathcal{A})(\text{affine}(\vec{0}; \vec{z}, \vec{y}_1, \vec{y}_2))$.*

It is interesting to note that our sampling arguments also imply a converse to the above theorems: if there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$ with $\deg Q \leq d$, such that $\Pr_{\vec{x} \in \mathbb{F}^m} [Q(\vec{x}) = \mathcal{A}(\vec{x})] \geq \gamma$, then 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 $\gamma - \varepsilon$ 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 series 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 $\vec{y} \neq \vec{0} \in \mathbb{F}^m$ and $c \in \mathbb{F}$,*

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

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

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

PROPOSITION 3.3. *For every subset $A \subseteq \mathbb{F}^m$, if*

$$\Pr_{\vec{y} \in \mathbb{F}^m} [\vec{y} \in A] > \frac{1}{|\mathbb{F}|}$$

then there exist linearly independent $\vec{y}_1, \ldots, \vec{y}_m \in \mathbb{F}^m$, such that for every $1 \leq i \leq m$, $\vec{y}_i \in A$.

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.

PROPOSITION 3.5 (LIST DECODING OF REED-MULLER). *For every function $f : \mathbb{F}^m \to \mathbb{F}$, if there are t different polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ such that for every $1 \leq i \leq t$, $\deg Q_i \leq d$ and $\Pr_{\vec{x} \in \mathbb{F}^m} [f(\vec{x}) = Q_i(\vec{x})] \geq \rho$ for $\rho \geq 2\sqrt{\frac{d}{|\mathbb{F}|}}$, then $t \leq \frac{2}{\rho}$.*

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 $s \in \mathcal{S}_k^m$, $|\{s' \in \mathcal{S}_{k'}^m \mid s' \subseteq s\}| = T$.*

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 an 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 $k' \leq k$,*

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

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 $s \in \mathcal{S}_k^m$,*

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

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

PROPOSITION 4.5 (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,*

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

Using this dual representation, we can easily conclude closure under intersection,

PROPOSITION 4.6 (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 $s_1 \cap s_2 \in \mathcal{S}_{k^{(3)}}^m$.*

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

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

5.2 Affine Subspaces Sample Well

We can now bound the deviation of the hitting rate of an affine subspace $s \in \mathcal{S}_k^m$ from its expected value,

COROLLARY 5.2 (SAMPLING). *Fix dimensions k and m, $1 \leq k \leq m$. Fix $A \subseteq \mathbb{F}^m$ of density $\mu = |A| / |\mathbb{F}^m|$. Then, for any $\varepsilon > 0$,*

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

5.3 Linear Subspaces Sample Well

We can similarly bound the deviation of the hitting rate of a linear subspace from its expected value,

COROLLARY 5.3 (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 $\varepsilon > 0$,*

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

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*: 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*: 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, 16, 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 (DECODING ⇒ LIST-DECODING). *Assume that $|\mathbb{F}| \geq 4$. 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
(relatively-low degree polynomial slightly agrees with $\mathcal{A}$)
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 list)*,*
Fix $\epsilon_0 \stackrel{def}{=} \sqrt{\frac{d'}{|\mathbb{F}|}}$. For every $\epsilon_0 < \delta < 1$, such that $\delta' \stackrel{def}{=} f(\delta - \epsilon_0) - \epsilon_0 \geq 2\epsilon_0$, 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

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

We can slightly enhance lemma 6.1 by requiring each member Q_i of the list to agree with the oracle $\mathcal{A}$ on more than ϵ/t weight of the subspaces $s \sim \mathcal{D}$. This can be done at the expense of lowering the weight of subspaces $s \sim \mathcal{D}$ explained by the list by another ϵ. In other words, instead of the above *list decoding* condition, we can have the following *two* conditions:

1. For every $i \in [t]$, $\Pr_{s\sim\mathcal{D}}\left[(Q_i \equiv \mathcal{A})(s)\right] > \frac{\epsilon}{t}$
2. $\mathbf{E}_{s\sim\mathcal{D}}\left[\Pr_{\vec{x}\in s}\left[\mathcal{A}(s)(\vec{x}) \neq \mathcal{A}(\vec{x}) \ \vee \ \exists i \ (Q_i \equiv \mathcal{A})(s)\right]\right] \geq 1 - \delta - \epsilon$

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 would sample 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 $f : s \to \mathbb{F}$ (not necessarily a *single* polynomial of degree at most d over s). Our semantic even permits the distribution to produce a *null* function with some probability. We interpret a null function as one that does not satisfy any property of the form "the function evaluates to..." (and hence satisfies every property of the form "the function does *not* evaluate to...").

LEMMA 6.2 (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 significantly 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 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

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

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 6.3 (DEGREE d CONSOLIDATION). *Fix dimensions k and m, $0 \le k < m$. 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 $0 \le \delta \le 1$: $\deg Q \le \delta |\mathbb{F}|$, and, when picking independently at random $\vec{z} \in \mathbb{F}^m$ and $\vec{y}_1, \ldots, \vec{y}_k \in \mathbb{H}^m$ such that $\vec{z}, \vec{y}_1, \ldots, \vec{y}_k$ are linearly independent,*

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

Then, $\deg Q \le d$.

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

LEMMA 6.4 (DEGREE md CONSOLIDATION). *Fix dimensions k and m, $1 \le k \le m$. Fix an oracle $\mathcal{A}$ assigning polynomials of degree at most d to all affine subspaces. Suppose that for some $0 \le \delta \le 1$, there exists a polynomial $Q : \mathbb{F}^m \to \mathbb{F}$, such that $\deg Q \le \delta |\mathbb{F}|$ and $\Pr_{s \in \mathcal{S}^m_k}[(Q \equiv \mathcal{A})(s)] > \delta + \frac{1}{|\mathbb{H}|}$. Then, $\deg Q \le md$.*

7. CONSISTENCY GRAPH

Fix a dimension $k \ge 3$. 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 that is smaller by 1 than k. 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 is used outside it.

The idea is a variation of the analysis of Raz and Safra for the non-randomness-efficient setting [16]. 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 $\mathcal{S}^k_{k-1}$ as follows. Let the vertices be all those subspaces, $V \stackrel{def}{=} \mathcal{S}^k_{k-1}$. 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 $\mathcal{S}^k_{k-1}$ are either parallel (i.e., identify or do not intersect) or intersect by an affine subspace from $\mathcal{S}^k_{k-2}$ (see closedness under intersection; proposition 4.6).

7.2 Graph is Almost-Transitive

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). *If $(s_1, s_2) \notin E_{\mathcal{A}}$,*

$$\Pr_{s_3 \in V}[(s_1, s_3) \in E_{\mathcal{A}} \ \wedge\ (s_3, s_2) \in E_{\mathcal{A}}] \le \frac{1}{|\mathbb{H}|} + \frac{d}{|\mathbb{F}|}$$

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 [16],

LEMMA 7.2 (GRAPH PARTITION). *Fix $\epsilon = \frac{1}{|\mathbb{H}|} + \frac{d}{|\mathbb{F}|}$. 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 $1 \le i \le t$, either $|V_i| = 1$, or $|V_i| > 2\sqrt{\epsilon}|V|$.*
2. *(almost all edges are within cliques)*

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

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). *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 $\deg Q \le 2d$, such that for every $s \in U$, $(Q \equiv \mathcal{A})(s)$.*

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

LEMMA 7.4 (HYPERPLANE VS. HYPERPLANE). *Assume $\mathcal{A}$ assigns polynomials of degree at most d to affine subspaces. For any $\delta \ge 8\sqrt{\frac{d}{|\mathbb{F}|} + \frac{1}{|\mathbb{H}|}}$ there exists a list of polynomials $Q_1, \ldots, Q_t : \mathbb{F}^k \to \mathbb{F}$, $t \le \frac{4}{\delta}$, with $\deg Q_i \le 2d$, such that*

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

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

Fix dimension $k \ge 3$. Let $\mathcal{A}$ be an oracle assigning polynomials of degree at most d to affine subspaces. In this section we prove that if there is γ consistency between affine subspaces of dimension $(k-1)$ in $\mathbb{F}^k$ and points within them, 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 several steps:

1. We use an argument of counting in several manners 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 but with not as good 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 used outside it. Note that the degree parameter grows from d to $2d$, and we indeed 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 between hyperplanes and points within them implies that for an average pair (s_1, s_2) of intersecting hyperplanes, $\mathcal{A}(s_1)$ and $\mathcal{A}(s_2)$ identify (with each other and with $\mathcal{A}$) on at least γ^2 of the points in the intersection of s_1 and s_2.

The proof uses repeatedly the trick of counting in several manners, 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 on 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 MANNERS). *If for oracle $\mathcal{A}$, holds $\mathbf{E}_{s \in \mathcal{S}^k_{k-1}} [\Pr_{\vec{x} \in s} [\mathcal{A}(s)(\vec{x}) = \mathcal{A}(\vec{x})]] \geq \gamma$, then,*

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

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*

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

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

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

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*

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

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

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

9. PROVING OUR MAIN THEOREMS

We prove Theorem 1 by induction using lemma 8.3 and the consolidation machinery we developed in section 6. We then conclude Theorem 2. Further details can be found in the full version of this paper [15].

10. 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.

11. REFERENCES

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

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

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

[4] L. Babai, L. Fortnow, and C. Lund. Non-deterministic exponential time has two-prover interactive protocols. In *Proc. 31st IEEE FOCS*, pages 16–25, 1990.

[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 STOC*, pages 1–10, 2004.

[6] E. Ben-Sasson and M. Sudan. Simple PCPs with poly-log rate and query complexity. In *Proc. 37th ACM STOC*, 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 STOC*, pages 612–621, 2003.

[8] I. Dinur. The PCP theorem by gap amplification. In *Proc. 38th ACM STOC*, 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 STOC*, pages 29–40, 1999.

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

[11] K. Ford. The distribution of integers with a divisor in a given interval. 2004.

[12] K. Friedl and M. Sudan. Some improvements to low degree tests. In *3rd ISTCS*, 1995.

[13] O. Goldreich and M. Sudan. Locally testable codes and PCPs of almost-linear length. In *Proc. 43rd IEEE FOCS*, 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 low degree test of almost-linear size. Technical Report TR05-086, ECCC, 2005.

[16] 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 STOC*, pages 475–484, 1997.

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