Electronic Colloquium on Computational Complexity, Report No. 90 (2018)

Worst-case to average case reductions for the distance to a code

Eli Ben-Sasson Swastik Kopparty Shubhangi Saraf

May 4, 2018

Abstract

Algebraic proof systems reduce computational problems to problems about estimating the distance of a sequence of functions $\vec{u} = (u_1, \ldots, u_k)$, given as oracles, from a linear error correcting code V. The soundness of such systems relies on methods that act "locally" on $\vec{u}$ and map it to a single function u^* that is, roughly, as far from V as are $u_1, \ldots, u_k$.

Motivated by these applications to efficient proof systems, we study a natural worst-case to average-case reduction of distance for linear spaces, and show several general cases in which the following statement holds: If some member of a linear space $U = \mathsf{span}(u_1, \ldots, u_k)$ is δ-far from (all elements) of V in relative Hamming distance, then nearly all elements of U are $(1-\epsilon)\delta$-far from V; the value of ϵ depends only on the distance of the code V and approaches 0 as that distance approaches 1. Our results improve on the previous state-of-the-art which showed that nearly all elements of U are $\frac{1}{2}\delta$-far from V [Rothblum, Vadhan and Wigderson, STOC 2013].

When V is a Reed-Solomon (RS) code, as is often the case for algebraic proof systems, we show how to *boost* distance via a new "local" transformation that may be useful elsewhere. Relying on the affine-invariance of V, we map a vector u to a random linear combination of affine transformations of u, and show this process amplifies distance from V. Assuming V is an RS code with sufficiently large distance, this amplification process converts a function u that is somewhat far from V to one that is $(1-\epsilon)$-far from V; as above, ϵ depends only on the distance of V and approaches 0 as the distance of V approaches 1.

We give two concrete application of these techniques. First, we revisit the axis-parallel low-degree test for bivariate polynomials of [Polischuk-Spielman, STOC 1994] and prove a "list-decoding" type result for it, when the degree of one axis is extremely small. This result is similar to the recent list-decoding-regime result of [Chiesa, Manohar and Shinkar, RANDOM 2017] but is proved using different techniques, and allows the degree in one axis to be arbitrarily large. Second, we improve the soundness analysis of the recent RS proximity testing protocol of [Ben-Sasson et al., ICALP 2018] and extend it to the "list-decoding" regime, bringing it closer to the Johnson bound.

1 Introduction

Proof systems that involve *interaction* between a *randomized* verifier and a prover have revolutionized computational complexity and cryptography [BM88, GMR89]. A question of paramount importance here is *soundness* — the minimal probability of the verifier rejecting a falsity. Transformations that maintain or increase soundness, while improving other aspects of the proof system (like proof length, or query complexity), are few and hard to obtain. Here, we study certain soundness-preserving techniques for the special case of *linear spaces*, improving on the prior state-of-the-art which was due to Rothblum, Vadhan and Wigderson [RVW13]; see Section 1.2. Then, in Section 1.4, we introduce a soundness-amplifying technique for the special case of Reed-Solomon codes; these codes are used in constructions of efficient proof systems. Before presenting the results we explain their relevance to the general study of proof systems.

ISSN 1433-8092

1.1 Motivation — improving concrete soundness and communication complexity of interactive protocols

Arithmetization is a technique that was introduced to the construction of interactive proof (IP) systems by [LFKN92], and later applied to other systems including multi-prover interactive proof (MIP) [BFL90], probabilistically checkable proof (PCP) [BFLS91, AS98, ALM+98] and zero knowledge (ZK) systems [GMR89], to name a few notable examples. Arithmetization refers to a family of *reductions* from languages (like 3SAT) to promise problems involving *algebraic codes* like Reed-Solomon (RS), Reed-Muller (RM), or their generalization to algebraic geometry (AG) codes; all are, in particular, *linear* codes.

An arithmetization reduction maps an instance x (like a 3SAT formula) to a sequence of algebraic codes $V_1, \ldots, V_k$, along with a set of "local" constraints, meaning that each constraint depends only on a small number of entries from k purported codewords. The reduction implies that $x \in L$ if and only if there exists a sequence $\vec{u} = (u_1, \ldots, u_k) \in V_1 \times \ldots \times V_k$ that satisfies all local constraints[1]. The locality of the constraints, along with the distance property of the codes $V_1, \ldots, V_k$ also implies that when $x \notin L$, *every* sequence $\vec{u}$ falsifies a large fraction of local constraints, as long as each member u_i of the sequence is *sufficiently close* to the code V_i in relative Hamming distance. Therefore, a major problem in the construction of such proof systems is to build protocols that efficiently ensure each u_i is in close proximity to V_i, and reject with non-negligible probability $s = s(\delta)$ a purported codeword u_i that is δ-far in relative Hamming distance from V_i. This problem is known as *proximity testing*; the study of the reliance of the soundness parameter s on the query complexity q and proximity parameter δ is referred to as *soundness analysis*.

Suffice it to say that protocols that solve the proximity testing problem are often a bottleneck in the construction of efficient proof systems, and the quality of their soundness analysis determines concrete efficiency and applicability (see, e.g., [AHIV17, BBHR18a] for recent instances). Therefore, it is desirable to construct transformations that minimize the number of proximity testing problems that are needed to be addressed by a proof system, and boost and maintain the distance of $\vec{u}$ from $V_1 \times \ldots \times V_k$ when $x \notin L$.

Certain proof systems use several instances of the same proximity problem, i.e., $V_1 = \ldots = V_k = V$ for a single linear code V. In this case, a natural optimization arises: instead of having the prover and verifier interact to solve k independent proximity problem, let the verifier sample $r_1, \ldots, r_k \in \mathbb{F}$, send them to the prover, and then interact to solve the *single* proximity problem that refers to $\sum_i r_i u_i$. The cost of an extra round of interaction (and extra randomness) are often well-worth the benefit of reducing the number of proximity testing problems. The linearity of C implies that this transformation does not harm (perfect) completeness, because when $\vec{u} \subset V$ then $\Pr[(\sum r_i u_i) \in V] = 1$.

The more interesting question, discussed next, is to understand what happens to the "typical" distance of $\sum r_i u_i$ as a function of the maximal distance, defined as $\delta_{\max} = \max_i \Delta(u_i, V)$.

1.2 Soundness transference results for linear spaces and error correcting codes

Our question is a special case of the "worst-case to average-case" problem: Suppose that a member u^* of a linear space $U \subseteq \mathbb{F}^n$ is $\delta_{\max}$-far in relative Hamming distance from all members of another linear space $V \subseteq \mathbb{F}^n$ (this is the "worst-case" assumption), what can be said about the *median*[2] distance δ_{med} from V, where this median is computed among the members of U? We address this question first for the case of V be a general space, then for V being an error correcting code.

1.2.1 General spaces

The basic question above was first raised by Rothblum, Vadhan and Wigderson, as part of their construction of efficient interactive proofs of proximity (IPPs) [RVW13]. They also showed that nearly all members of U — all but a $\frac{1}{|\mathbb{F}|-1}$-fraction of them — are $\delta/2$-far from V (Lemma 2.4). Thus, $\delta_{\mathsf{med}} \geq \delta_{\max}/2$. On the other

[1] The exact nature of these constraints is not relevant to our study here. The interested reader is referred, e.g., to [HS00, Section 3.1] and [BCGV16, Section 5] for examples and more information.

[2] All our results hold with high probability, i.e., with respect to the average and 99.9th percentile but we stick to using "median" for simplicity.

hand, $\delta_{\max} \geq \delta_{\mathsf{med}}$ for certain spaces U (including all 1-dimensional ones). We are interested in closing the gap between these two bounds.

Our first result (Theorem 4.1) looks at δ_{med} as a function of $\delta_{\max}$ and says

$$\delta_{\mathsf{med}}(\delta_{\max}) \geq 1 - \sqrt{1 - \delta_{\max}} - o(1)$$

Here and henceforth, $o(1)$ denotes negligible terms that approach 0 as $|\mathbb{F}| \to \infty$. In words, the median distance scales roughly like the *Johnson list-decoding function* of $\delta_{\max}$, denoted $J(\delta_{\max})$, where $J(x) \triangleq 1 - \sqrt{1-x}$. Thus, the median distance δ_{med} is strictly greater than $\delta_{\max}/2$ for all $\delta_{\max} > 0$, and approaches 1 as $\delta_{\max}$ approaches 1; the prior state-of-the-art approached $1/2$ in this case. For small values of $\delta_{\max}$, our bound approaches $\delta_{\max}/2$, as in prior works, but for special (and natural) cases we obtain better bounds on δ_{med}, even when it is arbitrarily small, as discussed next.

1.2.2 Linear error correcting codes

Most of the applications to interactive proof systems use a space V that is an *error correcting code*, i.e., the members of V are pair-wise far. Letting $\Delta(V)$ denote the relative distance of V, our second result (Theorem 4.3) states

$$\forall \delta_{\max} \leq J(J(\Delta(V)) - o(1), \quad \delta_{\mathsf{med}} \geq \delta_{\max} - o(1).$$

In simple words, $\delta_{\mathsf{med}} \approx \delta_{\max}$ for sufficiently small values of $\delta_{\max}$, where "sufficiently small" depends on $\Delta(V)$ and approaches 1 as $\Delta(V) \to 1$. Combining Theorems 4.1 and 4.3, one sees that for any $\epsilon > 0$ there exists a code-distance parameter δ_ϵ, such that for every V with $\Delta(V) > \delta_\epsilon$ and all spaces U, we have $\delta_{\mathsf{med}} \geq (1-\epsilon)\delta_{\max}$.

1.3 Applications to low-degree testing

We now present two different applications of our results. First, we extend the soundness analysis of the ubiquitous bivariate low-degree test of Polischuck and Spielman to the high-error regime for polynomials that have constant degree in one variable. Then we improve the soundness bounds on the recently suggested "fast RS interactive oracle proof of proximity" (FRI) protocol to beyond the unique-decoding radius.

1.3.1 High error bivariate testing

The bivariate axis-parallel test theorem of Polischuck and Spielman [PS94] is a fundamental component in many efficient PCP constructions. Roughly, the theorem says that if a function $f : \mathbb{F} \times \mathbb{F}$ has the property that its restriction to *most* columns is *very close* to a degree d_Y polynomial, and the restriction to *most* rows is a function that is *very close* to a degree d_X polynomial, then f is *very close* to being the evaluation of a bivariate polynomial of degree d_X in X and degree d_Y in Y.

As stated there, the result works for degrees d_X, d_Y as large as $\approx |\mathbb{F}|/2$ but requires the columns and rows to have *large agreement* with univariate low-degree polynomials, and this setting is known as the *low error regime*. An intriguing question is whether a similar result holds in the high-error regime, when only a non-trivial fraction of rows/columns exhibit non-trivial agreement with degree d polynomials.

This question has been given a positive answer by Arora and Sudan for a richer class of tests that includes the restriction of f to all lines (not just axis-parallel ones), and when $d < |\mathbb{F}|^{1/3}$ [AS03]. Recently, Chiesa, Manohar and Shinkar have proven the high-error case of the axis parallel test for small degree, i.e., when both d_X and d_Y are less than $\log |\mathbb{F}|$ [CMS17].

As the first application of our results, we improve on [PS94] and present a high-error analysis of the axis-parallel test. Our result, stated in Theorem 6.1 works when one of the degrees is constant ($d_X = O(1)$) and the other can be linear ($d_Y = \Omega(|\mathbb{F}|)$). Another setting of parameters for our result is when one of the degrees is $O(\log \log |\mathbb{F}|)$ and the other is $|\mathbb{F}|^{1-o(1)}$. Thus, our result is incomparable to that of [CMS17], because of the different requirements on d_X, d_Y; the proof techniques are also quite different.

1.3.2 Improved soundness analysis of the Fast Reed-Solomon interactive oracle proof of proximity (IOPP)

The *fast RS IOPP (FRI)* protocol [BBHR18b] is an interactive oracle proof of proximity (IOPP) for the *RS proximity testing* (RPT) problem (cf. [BCS16, BCF+16] for a definition and discussion of the IOPP model). For RS-codes of message length N over a field $\mathbb{F}$, prover arithmetic complexity is $O(N)$ and verifier arithmetic complexity for each test[3] is $O(\log N)$; this also bounds the query complexity of a single test. The efficiency of the FRI protocol is important for proof systems realized in code, like the recent zero knowledge proof system of [BBHR18a], called a "zk-STARK" there.

The soundness of a proximity testing protocol is described by a *soundness function* $s(\cdot)$ that takes as input a *proximity parameter* δ, and outputs the minimum rejection probability of the verifier, where this minimum is taken over all words that are δ-far from the code. In the case of FRI soundness for a single test, an upper bound $s(\delta) \leq \delta$ is easy to establish. The analysis in [BBHR18b] showed a nearly matching lower bound for *sufficiently small* values of δ. In particular, the bound obtained there gives

$$s(\delta) \geq \min\{\delta, \delta_0\} - o(1) \tag{1}$$

where δ_0 is a *soundness threshold* constant that depends on the code rate ρ as follows $\delta_0 \approx \frac{1-3\rho}{4}$ (see red line in Figure 1). For codes of rate $\geq 1/3$ this bound is meaningless, and even when $\rho \to 0$ it holds that $\delta_0 \to 1/4$; this rather low soundness means that many tests must be applied in order to reach a target soundness error; for soundness error $2^{-\lambda}$ and maximal proximity parameter $1-\rho$, the number of tests must still be greater than $\frac{\lambda}{-\log_2 \frac{3}{4}} \approx 2.4 \cdot \lambda$.

Using the results described in Sections 1.2.1 and 4.2 we improve on this state of affairs, and show that FRI soundness (for a single test) behaves as in Equation (1) but for a larger value of δ_0, namely, $\delta_0 \approx 1 - \sqrt[4]{\rho}$ (see blue line in Figure 1). Consequently, to reach soundness error $2^{-\lambda}$ as before, the number of tests is reduced to $\approx \frac{4\lambda}{-\log \rho}$ which is always smaller than $2.4 \cdot \lambda$ and approaches 0 as $\rho \to 0$. We end by pointing out that [BBHR18b] conjecture that the trivial soundness upper bound (green line in Figure 1) is nearly tight, i.e., that $s(\delta) \approx \delta$ for all values of δ. Reducing further the gap between soundness upper bounds (green line) and lower bounds (blue line) remains an interesting open problem that is relevant to realized proof systems like the zk-STARK of [BBHR18a]. In follow-up work we shall report qualitative improvements to the bounds posted here that are optimal and tight for at least a certain range of code parameters.

1.4 Soundness amplification for Reed-Solomon codes

So far we tried to minimize the loss in distance incurred by sampling an element of U. Next, we suggest a way to boost distance via a family of "locally-computable" transformations acting on a *single* purported codeword u. A *q-locally computable transformation* is a mapping $M : \mathbb{F}^n \to \mathbb{F}^n$ for which the ith entry of $M(u)$ can be computed by querying at most q entries of u. To preserve completeness, we require the mappings M to preserve the space V, and this leads to a natural suggestion. Let $\mathsf{Aut}(V)$ be the *automorphism group* of V. Sample $M_1 \ldots, M_{q-1} \in \mathsf{Aut}(V)$ and $r_1, \ldots, r_{q-1}$ and let $u^* = M(u) \triangleq u + \sum_{i<q} r_i M_i(u)$. By definition, this mapping is q-local and it preserves (perfect) completeness: if u belongs to V then so does each $M_i(u)$, so by linearity $M(u) \in V$. It now stands to reason that if $\mathsf{Aut}(V)$ is sufficiently "pseudo-random", say, a doubly-transitive group, then the median distance of $M(u)$ should be even greater than $\Delta(u, V)$ (the distance of u from V).

For example, consider the family of Reed-Solomon codes $\mathsf{RS}[\mathbb{F}, \rho]$, which are comprised of all functions $f : \mathbb{F} \to \mathbb{F}$ such that $\deg(f) < \rho|\mathbb{F}|$ where $\deg(f)$ is the degree of the interpolating polynomial of (the function) f. It is well known that $\mathsf{Aut}(RS[\mathbb{F}, \rho])$ is the 1-dimensional affine group of $\mathbb{F}$, denoted $\mathsf{Aff}_1(\mathbb{F})$, whose members are all invertible affine transformations $\mathsf{Aff}_1(\mathbb{F}) = \{M(X) = aX + b \mid a \in \mathbb{F}^*, b \in \mathbb{F}\}$; this group is indeed doubly-transitive.

Our final set of results studies the effect of taking random linear combinations of random automorphisms for Reed-Solomon codes. Suppose we start with a function u, and then take random linear combinations

[3] In [BBHR18b], a single test means a single invocation of the QUERY protocol.,

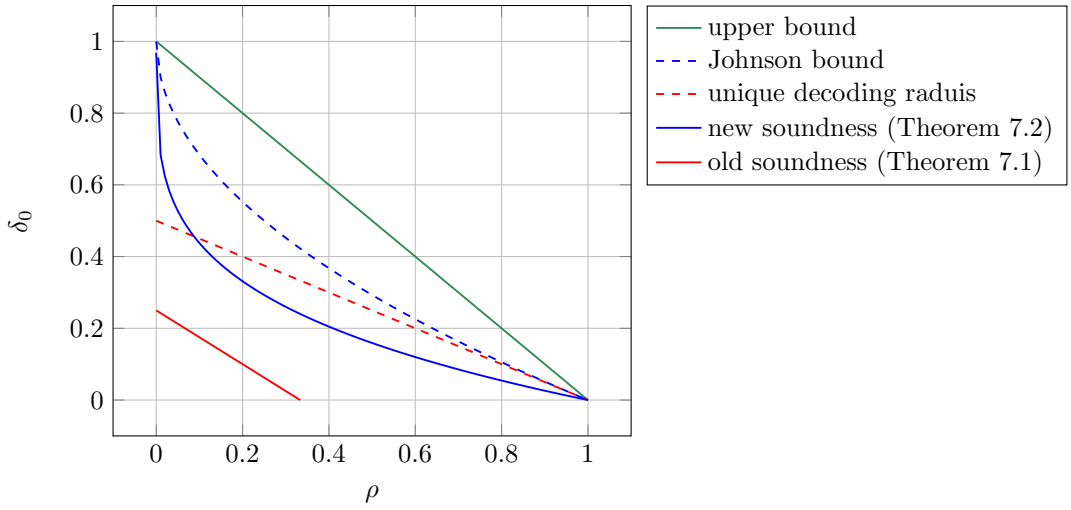

Figure 1: FRI soundness threshold δ_0 as a function of RS code rate ρ, for a single invocation of the FRI QUERY phase (see Equation (1) and explanation in text there for the meaning of the constant δ_0). Higher lines are better. The top line is the trivial upper bound on soundness; the bottom line is the soundness of the original analysis of [BBHR18b] (cf. Theorem 7.1). The middle line is the new and improved analysis given by Theorem 7.2. This analysis presents non-trivial soundness bounds for all code rates, and these bounds are better than the prior state of the art.

of a few random affine shifts of u to produce a function u^*. From the discussion above, if u is in a Reed-Solomon code, then so is u^*. We show in Theorem 5.1 that if u is far from a Reed-Solomon code, then with high probability u^* is very far from that Reed-Solomon code. The main strength of this result is that this process can then amplify the distance to V all the way to $1 - o(1)$ (while more direct analyses, related to the Rothblum-Vadhan-Wigderson [RVW13] lemma, cannot amplify beyond distance $1/2$).

2 Preliminaries

We use Δ to denote normalized Hamming distance, and $\mathbf{0} = 0^n$ denotes the identity element of an n-dimensional vector space, viewed as an additive group.

In what follows Σ is a finite alphabet. For $S \subset \Sigma^n$ let $\Delta(S) = \min\{\Delta(w, w') \mid w, w' \in S, w \neq w'\}$ denote the relative Hamming distance of S. For $w \in \Sigma^n$ let $B(w, \delta)$ denote the Hamming ball in Σ^n of normalized radius δ centered at w,

$$B(w, \delta) = \{r \in \Sigma^n \mid \Delta(w, r) < \delta\}$$

Definition 2.1 (List decodability)**.** *For $\rho \in [0, 1]$ and $L \geq 1$, we say a set $S \subseteq \Sigma^n$ is (ρ, L)-*list-decodable *if for all $w \in \Sigma^n$,*

$$|B(w, \rho) \cap S| \leq L.$$

We have the fundamental Johnson bound, which says that sets with large minimum distance have non-trivial list-decodability. See, e.g., [Gur07, Corollary 3.2] for a proof.

Theorem 2.2 (Johnson bound)**.** *For every $\epsilon \in (0, 1]$, Let $J_\epsilon : [0, 1] \to [0, 1]$ be the function*

$$J_\epsilon(\delta) = 1 - \sqrt{1 - \delta(1 - \epsilon)}.$$

Let Σ be a finite alphabet, n an integer and $S \subseteq \Sigma^n$. Then S is $(J_\epsilon(\Delta(S)), 1/\epsilon)$-list-decodable for every $\epsilon \in (0, 1]$.

An *affine space* U is an additive coset of a vector space U', i.e., for some fixed $a \in \mathbb{F}^n, U = a + U' \triangleq \{a + u \mid u \in U'\}$. We introduce the following definition.

Definition 2.3 (Divergence)**.** *For $U, V \subseteq \Sigma^n$, the* divergence *of U from V is $D(U,V) = \max_{u \in U} \Delta(u, V)$.*

Divergence is not a distance measure because it is not symmetric. This is witnessed by $U = \left\{\mathbf{0}, 10^{n-1}\right\}, V = \{\mathbf{0}, 1^n\} \subset \{0,1\}^n$, which gives $D(U,V) = \frac{1}{n} \neq \frac{n-1}{n} = D(V,U)$.

The next lemma, due to Rothblum-Vadhan-Wigderson, says that if some vector in a linear space U is δ-far from a space V, then nearly all elements of U are $\delta/2$-far from V.

Lemma 2.4 ([RVW13, Lemma 1.6])**.** *For any pair of linear spaces U, V over a finite field $\mathbb{F}$,*

$$\Pr_{u \in U}\left[\Delta(u,V) < \frac{D(U,V)}{2}\right] \leq \frac{1}{|\mathbb{F}| - 1}. \tag{2}$$

3 Preserving distances for general subspaces

In this section, we prove our first strengthening of the Rothblum-Vadhan-Wigderson lemma Lemma 2.4 from above. The main new qualitative feature is that if $D(U,V) = 1 - o(1)$, then the lemma concludes that most elements of u are at distance $1 - o(1)$ from V.

Theorem 3.1. *For a pair of affine spaces U, V over a finite field $\mathbb{F}$, and for all $\epsilon \in (0,1]$,*

$$\Pr_{u \in U}\left[\Delta(u,V) < J_\epsilon(D(U,V))\right] < \frac{1}{\epsilon\left(|\mathbb{F}| - 1\right)}.$$

Theorem 3.1 is a consequence of the following lemma, which says that if u^* is δ-far from V, then for any line passing through u^* in direction u, most points are $J_\epsilon(\delta)$ from V. We state the Lemma, prove Theorem 3.1 and then prove the lemma.

Lemma 3.2. *Let $V \subseteq \mathbb{F}^n$ be a linear space over a finite field $\mathbb{F}$; suppose $u^* \in \mathbb{F}^n$ satisfies $\Delta(u^*, V) \geq \delta$. For any $u \in \mathbb{F}^n$ and $\epsilon \in (0,1]$ let $A = A_{u,\epsilon} = \{\alpha \in \mathbb{F} \setminus \{0\} \mid \Delta(u^* + \alpha u, V) < J_\epsilon(\delta)\}$. Then $|A| \leq 1/\epsilon$.*

Proof of Theorem 3.1. It suffices to prove the Theorem for the case that V is a linear space and U is an affine space (which may be linear as well), because Hamming distance is invariant under shifting both U and V by the same vector v. Let $u^* \in U$ be some element for which $\Delta(u^*, V) = D(U,V)$. We may assume $u^* \neq \mathbf{0}$, otherwise $D(U,V) = 0$ because $\mathbf{0} \in V$ so the claim trivially holds. If $\dim(U) = 0$ the claim also trivially holds because $|U| = 1$. Therefore, we assume $U = u^* + U'$ for some linear space U' of positive dimension d (which may include u^*). There exist $k = |\mathbb{F}|^{d-1}$ vectors $u_1, \ldots, u_k$ such that $U \setminus \{u^*\}$ can be partitioned into equi-sized sets, the ith set being the line $\{u^* + \alpha u_i \mid \alpha \in \mathbb{F} \setminus \{0\}\}$. Theorem 3.1 follows by applying Lemma 3.2 to each of the sets in this partition. □

Proof of Lemma 3.2. For $\alpha \in A$, let $v^\alpha \in V$ be such that $\Delta(u^* + \alpha u, v^\alpha) < J_\epsilon(\delta)$. Rewriting, we have that for each $\alpha \in A$,

$$\Delta\left(u, \frac{v^\alpha - u^*}{\alpha}\right) < J_\epsilon(\delta).$$

Assume by way of contradiction that $|A| > 1/\epsilon$. Thus, a set (or possibly multi-set) of more than $1/\epsilon$ vectors are all $J(\delta,\epsilon)$-close to u. By the Johnson bound, two of the vectors must be δ-close to one another. Let α, α' be these distinct members of A for which

$$\Delta\left(\frac{v^\alpha - u^*}{\alpha}, \frac{v^{\alpha'} - u^*}{\alpha'}\right) < \delta.$$

Recalling $\Delta(u,v) = \Pr_{i\in[n]}[u_i \neq v_i]$ where u_i, v_i denote the ith entry of u, v, respectively, we have

$$\begin{aligned}
\delta &> \Pr_{i\in[n]}\left[\left(\frac{v^\alpha - u^*}{\alpha}\right)_i \neq \left(\frac{v^{\alpha'} - u^*}{\alpha'}\right)_i\right] = \Pr_{i\in[n]}\left[\left(\frac{v^\alpha - u^*}{\alpha} - \frac{v^{\alpha'} - u^*}{\alpha'}\right)_i \neq 0\right] \\
&= \Pr_{i\in[n]}\left[\left((\alpha-\alpha')u^* - (\alpha v^{\alpha'} - \alpha' v^\alpha)\right)_i \neq 0\right] \\
&= \Pr_{i\in[n]}\left[u_i^* \neq \left(\frac{\alpha v^{\alpha'} - \alpha' v^\alpha}{\alpha - \alpha'}\right)_i\right].
\end{aligned}$$

Setting $v' = \frac{\alpha v^{\alpha'} - \alpha' v^\alpha}{\alpha - \alpha'}$ and noticing $v' \in V$ we conclude

$$\Delta(u^*, V) \leq \Delta(u^*, v') < \delta$$

which is false and which contradicts our hypothesis on the size of A. We conclude $|A| \leq 1/\epsilon$, as claimed. □

4 Preserving distances for good error correcting code

In this section we prove another strengthening of the Rothblum-Vadhan-Wigderson lemma. This strengthening only works when the subspace V is a code of good distance. Assume for now that V is a code with minimum distance $1-o(1)$. Then the strengthened theorem gives a stronger guarantee: they show that most elements of U are at distance $\min(D(U,V) - o(1)$ from V. Thus the maximum distance of an element of U from V is also the typical distance of an element of U from V.

We begin with a warm-up: we show a "unique-decoding" version which only works up to $1/3$ of the minimum distance of the code V. The next "list-decoding" version works up to a much larger distance, and in particular for V having distance $1 - o(1)$, it works up to a distance of $1 - o(1)$.

4.1 Unique-Decoding version

Theorem 4.1. *Let $V \subseteq \mathbb{F}^n$ be a linear space over a finite field $\mathbb{F}$ with $\Delta(V) = \lambda$. Let U be an affine space and suppose $D(U,V) > \delta$. For any $\epsilon > 0$ such that $\delta - \epsilon < \lambda/3$,*

$$\Pr_{u\in U}\left[\Delta(u,V) < \delta - \epsilon\right] \leq \frac{1}{\epsilon|\mathbb{F}|}$$

Theorem 4.1 is a consequence of the following lemma. As in Section 3, we state the lemma, prove Theorem 4.1 and then prove the lemma.

Lemma 4.2. *Let $V \subseteq \mathbb{F}^n$ be a linear space over a finite field $\mathbb{F}$ with $\Delta(V) = \lambda$. Suppose $u^* \in \mathbb{F}^n$ satisfies $\Delta(u^*, V) > \delta$ and fix arbitrary $u \in \mathbb{F}^n$. For $\epsilon > 0$ satisfying $\delta - \epsilon < \lambda/3$ let $A = \{\alpha \in \mathbb{F} \mid \Delta(u^* + \alpha u, V) < \delta - \epsilon\}$. If $|A| > 1/\epsilon$ then there exist $v, v^* \in V$ such that:*

$$|\{i \in [n] \mid (u_i = v_i) \wedge (u_i^* = v_i^*)\}| \geq (1-\delta)\cdot n.$$

Proof of Theorem 4.1. We prove the contra-positive: If the assumptions on $\epsilon, \delta, \lambda$ hold and

$$\Pr_{u\in U}\left[\Delta(u,V) < \delta - \epsilon\right] > \frac{1}{\epsilon|\mathbb{F}|}, \tag{3}$$

then $D(U,V) \leq \delta$.

Let $u^* \in U$ satisfy $\Delta(u^*, V) = D(U,V)$. We may assume V is a linear space and $\dim(U) > 0$, as argued in the proof of Theorem 3.1. As there, partition $U \setminus \{u^*\}$ into equi-size sets, each of the form $\{u^* + \alpha u_i \mid \alpha \in \mathbb{F} \setminus \{0\}\}$ for some set $u_1, \ldots, u_k \in \mathbb{F}^n$ of vectors. By our assumption in Equation (3) there exists u_i such that the set $A = \{\alpha \in \mathbb{F} \mid \Delta(u^* + \alpha u_i, V) < \delta - \epsilon\}$ is of size greater than $1/\epsilon$. Apply Lemma 4.2 to this set, and conclude $\Delta(u^*, v^*) \leq \delta$, as claimed. □

Proof of Lemma 4.2. For $\alpha \in A$, let $v^\alpha \in V$ be such that $\Delta(u^* + \alpha u, v^\alpha) < \delta - \epsilon$.

We first show that for all $\alpha \in A$, the points (α, v^α) are all collinear. To see this, let $\alpha_1, \alpha_2, \alpha_3 \in A$ be distinct. We have $\Delta(u^* + \alpha_3 u, v^{\alpha_3}) \leq \delta - \epsilon$. On the other hand, if $\beta = \frac{\alpha_3 - \alpha_2}{\alpha_1 - \alpha_2}$, we have:

$$u^* + \alpha_3 u = \beta(u^* + \alpha_1 u) + (1 - \beta)(u^* + \alpha_2 u),$$

and so:

$$\begin{aligned}\Delta(u^* + \alpha_3 u, \beta v^{\alpha_1} + (1-\beta)v^{\alpha_2}) &\leq \Delta(u^* + \alpha_1 u, v^{\alpha_1}) + \Delta(u^* + \alpha_2 u, v^{\alpha_2}) \\ &\leq (\delta - \epsilon) + (\delta - \epsilon) \\ &= 2(\delta - \epsilon).\end{aligned}$$

Thus $\Delta(\beta v^{\alpha_1} + (1-\beta)v^{\alpha_2}, v^{\alpha_3}) \leq 3(\delta - \epsilon) < \lambda$. By the minimum distance hypothesis on V, we conclude that

$$\beta v^{\alpha_1} + (1 - \beta)v^{\alpha_2} = v^{\alpha_3},$$

which implies the desired collinearity.

Thus there exist $v, v^* \in V$ such that for all $\alpha \in A$,

$$v^\alpha = v^* + \alpha v.$$

Taking this information back to the definition of v^α, we have that for all $\alpha \in A$,

$$\Delta(u^* + \alpha u, v^* + \alpha v) < \delta - \epsilon.$$

Rewriting,

$$\Delta(u^* - v^*, \alpha(v - u)) < \delta - \epsilon.$$

for all $\alpha \in A$.

Now for any coordinate $i \in [n]$ where $u_i \neq v_i$ or $u_i^* \neq v_i^*$, there can be at most one value of $\alpha \in \mathbb{F}$ for which $u_i^* - v_i^* = \alpha(v_i - u_i)$. Let $t = |A|$. Thus there is an $\alpha \in A$ such that:

$$\delta - \epsilon > \Delta(u^* - v^*, \alpha(v - u)) \geq 1 - \frac{|\{i \in [n] \mid (u_i = v_i) \wedge (u_i^* = v_i^*)\}|}{n} - \frac{1}{t}.$$

Putting everything together, we get that:

$$\frac{|\{i \in [n] \mid (u_i = v_i) \wedge (u_i^* = v_i^*)\}|}{n} \geq 1 - \delta + \epsilon - \frac{1}{t}.$$

Thus if $t > \frac{1}{\epsilon}$, we have:

$$\frac{|\{i \in [n] \mid (u_i = v_i) \wedge (u_i^* = v_i^*)\}|}{n} \geq 1 - \delta,$$

as claimed. □

4.2 List-Decoding version

Theorem 4.3. *Let $V \subseteq \mathbb{F}_q^n$ be a subspace with minimum distance λ. Let $\epsilon, \delta > 0$ with $\delta < J_\epsilon(J_\epsilon(\lambda))$.*

Suppose $u^* \in \mathbb{F}_q^m$ is such that $\Delta(u^*, V) > \delta$. Then for all $u \in \mathbb{F}_q^n$, there are at most $2/\epsilon^3$ values of $\alpha \in \mathbb{F}_q$ such that $\Delta(u^* + \alpha u, V) < \delta - \epsilon$.

This is a consequence of the following theorem.

Theorem 4.4. *Let $V \subseteq \mathbb{F}^n$ be a linear space over a finite field $\mathbb{F}$ with $\Delta(V) = \lambda$. Let $u^* \in \mathbb{F}^n$ and let $\epsilon > 0$ satisfy $\delta < J_\epsilon(J_\epsilon(\lambda))$. For $u \in \mathbb{F}^n$ let $A = A_{u,\epsilon} = \{\alpha \in \mathbb{F} \setminus \{0\} \mid \Delta(u^* + \alpha u, V) < \delta - \epsilon\}$. If $|A| > 2/\epsilon^3$ then there exist $v^*, v \in V$ such that*

$$|\{i \in [n] \mid (u_i = v_i) \wedge (u_i^* = v_i^*)\}| \geq (1-\delta)n.$$

In particular,

$$\Delta(u^*, v^*) \leq \delta.$$

Proof. Let $t = |A|$. For $\alpha \in A$, let $v^\alpha \in V$ be such that $\Delta(u^* + \alpha u, v^\alpha) < \delta - \epsilon$. Thus $\Delta(u^*, v^\alpha - \alpha u) < \delta - \epsilon$.

Now consider the following graph with vertex set A: α and α' are adjacent if $\Delta(v^\alpha - \alpha u, v^{\alpha'} - \alpha' u) < J_\epsilon^{-1}(\delta)$. The Johnson bound implies that this graph has no independent set of size $c' = 1/\epsilon$. Thus by Turan's theorem, there is a vertex α_0 of degree at least $\epsilon|A| - 1$.

Concretely, this means that there is a set $B \subseteq A$, with $|B| \geq \epsilon|A| - 1$, such that for all $\alpha \in B$,

$$\Delta(v^{\alpha_0} - \alpha_0 u, v^\alpha - \alpha u) < J_\epsilon^{-1}(\delta).$$

Rewriting, we have:

$$\Delta(u, \frac{1}{\alpha - \alpha_0} \cdot (v^\alpha - v^{\alpha_0})) < J_\epsilon^{-1}(\delta), \tag{4}$$

for every $\alpha \in B$.

Now we apply the Johnson bound again. Since V has distance λ, and $J_\epsilon(\lambda) > J_\epsilon^{-1}(\delta)$, there can be at most $1/\epsilon$ distinct vectors $v \in V$ such that $\Delta(u, v) < J_\epsilon^{-1}(\delta)$.

The only way this can be consistent with Equation (4) is if many of the $\frac{1}{\alpha - \alpha_0} \cdot (v^\alpha - v^{\alpha_0})$ are identical. Specifically, by the pigeonhole principle we get that there is a $v \in V$ and a set $C \subseteq B$, with $|C| \geq \epsilon|B|$, such that for all $\alpha \in C$,

$$v = \frac{1}{\alpha - \alpha_0} \cdot (v^\alpha - v^{\alpha_0}).$$

So for all $\alpha \in C$,

$$v^\alpha = (v^{\alpha_0} - \alpha_0 v) + \alpha \cdot v.$$

Let us denote this by $v^\alpha = v^* + \alpha v$, where $v, v^* \in V$.

Taking this information back to the definition of v^α, we have that for all $\alpha \in C$,

$$\Delta(u^*, v^* + \alpha(v - u)) < \delta - \epsilon.$$

Rewriting,

$$\Delta(u^* - v^*, \alpha(v - u)) < \delta - \epsilon.$$

for all $\alpha \in C$.

Now for any coordinate $i \in [n]$ where $u_i \neq v_i$ or $u_i^* \neq v_i^*$, there can be at most one value of $\alpha \in \mathbb{F}$ for which $u_i^* - v_i^* = \alpha(v_i - u_i)$. Thus there is an $\alpha \in C$ such that

$$\Delta(u^* - v^*, \alpha(v - u)) \geq 1 - \frac{|\{i \in [n] \mid (u_i = v_i) \wedge (u_i^* = v_i^*)\}|}{n} - \frac{1}{|C|}.$$

Combining this with our upper bound on $\Delta(u^* - v^*, \alpha(v - u))$, we get that:

$$\frac{|\{i \in [n] \mid (u_i = v_i) \wedge (u_i^* = v_i^*)\}|}{n} \geq 1 - \delta + \epsilon - \frac{1}{|C|}. \tag{5}$$

Since $|C| \geq \epsilon|B| \geq \epsilon(\epsilon|A| - 1)$, and since $A \geq 2/\epsilon^3$, we get that

$$|C| > 1/\epsilon,$$

and the desired conclusion follows from Equation (5). □

We now state a generalization of the above theorem from lines to higher degree curves, which we prove by induction on degree. We include a stronger conclusion in this generalization because it is useful for induction purposes.

Define $J_\epsilon^{[k]}(\lambda) = J_\epsilon(J_\epsilon(\dots(J_\epsilon(\lambda))))$, where there are k iterations of the function J_ϵ.

Theorem 4.5. *Let $V \subseteq \mathbb{F}^n$ be a linear space over a finite field $\mathbb{F}$ with $\Delta(V) = \lambda$. Let $u^* \in \mathbb{F}^n$ and let $\epsilon > 0$ satisfy $\delta < J_\epsilon^{[\ell+1]}(\lambda)$. For $u_1, u_2, \cdots, u_\ell \in \mathbb{F}^n$ let $A = A_{u_1,u_2,\dots,u_\ell,\epsilon} = \left\{\alpha \in \mathbb{F} \setminus \{0\} \mid \Delta(u^* + \alpha u_1 + \alpha^2 u_2 + \alpha^3 u_3 + \cdots + \alpha^\ell u_\ell, V$ If $|A| > \ell \cdot \left(\frac{2}{\epsilon}\right)^{\ell+2}$, then the following two statements hold:*

1. *There exist $v^*, v_1, v_2, \dots, v_\ell \in V$ such that*

$$|\{i \in [n] \mid (u_i^* = v_i^*) \wedge ((u_1)_i = (v_1)_i) \wedge \cdots \wedge ((u_\ell)_i = (v_\ell)_i)\}| \geq (1 - \delta - \epsilon)n.$$

   In particular,

$$\Delta(u^*, v) \leq \delta + \epsilon.$$

2. *For $\alpha \in A$, fix $v^\alpha \in V$ such that $\Delta(u^* + \alpha u_1 + \alpha^2 u_2 + \alpha^3 u_3 + \cdots + \alpha^\ell u_\ell, v^\alpha) < \delta$. Then there is a subset $C \subseteq A$ with $|C| \geq \left(\frac{\epsilon}{2}\right)^{\ell+1} |A|$ such that for all $\alpha \in C$,*

$$v^* + \alpha v_1 + \alpha^2 v_2 + \alpha^3 v_3 + \cdots + \alpha^\ell v_\ell = v^\alpha.$$

Proof. The proof of the above theorem is by induction on ℓ (and shares many ideas with the proof of Theorem 4.4). For $\ell = 0$ this theorem follows from the Johnson bound: all the v^α are within distance $J_\epsilon(\delta)$ of u^*, and thus the number of distinct v_α is at most $\frac{1}{\epsilon}$. This easily implies the $\ell = 0$ case.

Let us assume the result is true for $\ell \leq k-1$, and show that the result is true for $\ell = k$.

Suppose $|A| > k \cdot \left(\frac{2}{\epsilon}\right)^{k+2}$. For $\alpha \in A$, let $v^\alpha \in V$ be such that $\Delta(u^* + \alpha u_1 + \alpha^2 u_2 + \alpha^3 u_3 + \cdots + \alpha^k u_k, v^\alpha) < \delta$. Thus $\Delta(u^*, v^\alpha - (\alpha u_1 + \alpha^2 u_2 + \alpha^3 u_3 + \cdots + \alpha^k u_k)) < \delta$.

Now consider the following graph with vertex set A: α and α' are adjacent if $\Delta(v^\alpha - (\alpha u_1 + \alpha^2 u_2 + \alpha^3 u_3 + \cdots + \alpha^k u_k), v^{\alpha'} - (\alpha' u_1 + \alpha'^2 u_2 + \alpha'^3 u_3 + \cdots + \alpha'^k u_k)) < J_\epsilon^{-1}(\delta)$. The Johnson bound implies that this graph has no independent set of size $c' = 1/\epsilon$. Thus by Turan's theorem, there is a vertex α_0 of degree at least $\epsilon|A| - 1 \geq \epsilon|A|/2$.

Concretely, this means that there is a set $B \subseteq A$, with $|B| \geq \epsilon|A|/2$, such that for all $\alpha \in B$,

$$\Delta(v^{\alpha_0} - (\alpha_0 u_1 + \alpha_0^2 u_2 + \alpha_0^3 u_3 + \cdots + \alpha_0^k u_k), v^\alpha - (\alpha u_1 + \alpha^2 u_2 + \alpha^3 u_3 + \cdots + \alpha^k u_k)) < J_\epsilon^{-1}(\delta).$$

Rewriting, we have:

$$\Delta\left(\frac{v^{\alpha_0} - v^\alpha}{\alpha_0 - \alpha}, \frac{1}{\alpha_0 - \alpha} \cdot \sum_{i=1}^{k} (\alpha_0^i - \alpha^i) u_i\right) < J_\epsilon^{-1}(\delta), \tag{6}$$

for every $\alpha \in B$.

Now, let $w^\alpha \in V$ be the vector

$$\frac{v^{\alpha_0} - v^\alpha}{\alpha_0 - \alpha},$$

and let $u_1', u_2', \dots, u_k' \in \mathbb{F}^n$ be such that u_i' is the coeffecient of α^{i-1} in

$$\frac{1}{\alpha_0 - \alpha} \cdot \sum_{i=1}^{k} (\alpha_0^i - \alpha^i) u_i.$$

Thus,

$$\frac{1}{\alpha_0 - \alpha} \cdot \sum_{i=1}^{k} (\alpha_0^i - \alpha^i) u_i = \sum_{i=1}^{k} u_i' \cdot \alpha^{i-1},$$

and for all $\alpha \in B$,

$$\Delta(w^\alpha, u'_1 + \alpha u'_2 + \cdots + \alpha^{k-1} u'_k) < J_\epsilon^{-1}(\delta)\delta'.$$

Since $\delta < J_\epsilon^{[k+1]}(\lambda)$, we have that $J_\epsilon^{-1}(\delta) < J_\epsilon^{[k]}(\lambda)$. Further, we have that $|B| \geq \epsilon|A|/2 > (k-1)\cdot\left(\frac{2}{\epsilon}\right)^{k+1}$. This means that we may apply the induction hypothesis (for the setting $\ell = k-1$). We get that there is a large subset $C \subseteq B$ and there exist $v'_1, v'_2, \ldots, v'_k \in V$ such that

$$|C| \geq \left(\frac{\epsilon}{2}\right)^k |B| \geq \left(\frac{\epsilon}{2}\right)^{k+1} |A|,$$

and for all $\alpha \in C$,

$$v'_1 + \alpha v'_2 + \alpha^2 v'_3 + \cdots + \alpha^k v'_k = w^\alpha.$$

This proves the second part of the theorem.

Thus

$$\frac{v^{\alpha_0} - v^\alpha}{\alpha_0 - \alpha} = v'_1 + \alpha v'_2 + \alpha^2 v'_3 + \cdots + \alpha^k v'_k,$$

where $v'_1, v'_2, \ldots, v'_k \in V$.

Rearranging, this shows that for all $\alpha \in C$, we can express v^α as $v^* + \alpha v_1 + \alpha^2 v_2 + \alpha^3 v_3 + \cdots + \alpha^k v_k$, where $v^*, v_1, v_2, \ldots, v_k \in V$.

Taking this back to the definition of v^α, we have that for all $\alpha \in C$,

$$\Delta(u^*, v^* + (\alpha(v_1 - u_1) + \alpha^2(v_2 - u_2) + \alpha^3(v_3 - u_3) + \cdots + \alpha^k(v_k - u_k))) < \delta.$$

Rewriting, we have that for all $\alpha \in C$,

$$\Delta((u^* - v^*) + \alpha(u_1 - v_1) + \alpha^2(u_2 - v_2) + \alpha^3(u_3 - v_3) + \cdots + \alpha^k(u_k - v_k), 0) < \delta.$$

Now for any coordinate $i \in [n]$ where $u_i^* \neq v_i^*$ or $(u_j)_i \neq (v_j)_i$ for any $j \in [k]$, the restriction to the ith coordinate gives us a nonzero degree k polynomial in α, and so there are at most k values of $\alpha \in \mathbb{F}$ for which $(u^* - v^*)_i + \alpha \cdot (u_1 - v_1)_i + \alpha^2 \cdot (u_2 - v_2)_i + \alpha^3 \cdot (u_3 - v_3)_i + \cdots + \alpha^k \cdot (u_k - v_k)_i = 0$.

Thus there is an $\alpha \in C$ such that

$$\Delta((u^* - v^*) + \alpha(u_1 - v_1) + \alpha^2(u_2 - v_2) + \alpha^3(u_3 - v_3) + \cdots + \alpha^k(u_k - v_k), 0) \geq 1 - \frac{|\{i \in [n] \mid (u_i^* = v_i^*) \wedge ((u_1)_i = (v_1)_i) \wedge \cdots \wedge ((u_k)_i = (v_k)_i)\}|}{n} - \frac{k}{|C|}.$$

Combining this with our upper bound on

$$\Delta(u^* - v^* + \alpha(u_1 - v_1) + \alpha^2(u_2 - v_2) + \alpha^3(u_3 - v_3) + \cdots + \alpha^k(u_k - v_k), 0),$$

we get that:

$$\frac{|\{i \in [n] \mid (u_i^* = v_i^*) \wedge ((u_1)_i = (v_1)_i) \wedge \cdots \wedge ((u_k)_i = (v_k)_i)\}|}{n} \geq 1 - \delta - \frac{k}{|C|}. \tag{7}$$

Since $|C| \geq \left(\frac{\epsilon}{2}\right)^{k+1} |A|$, our assumption about $|A|$ implies that

$$|C| > k/\epsilon,$$

and the first part of the theorem then follows from Equation (7). □

5 Distance Amplification for Reed-Solomon codes

In this section, we show how to use the results of the previous section to show that some simple transformations *amplify* the distance of a function from the space of low-degree polynomials (i.e., Reed-Solomon codes). In the previous section, we saw results with the flavor: if u^* is δ-far from the subspace V, then there are many other functions (related to u^*) that are also almost δ-far from the subspace V. Now we will get more: we will find many functions related to u^* that are δ'-far from V for some δ' bigger than δ. The main strength of this result is that this process can then amplify the distance to V all the way to $1 - o(1)$ (while more direct analyses, related to the Rothblum-Vadhan-Wigderson [RVW13] lemma, cannot amplify beyond distance $1/2$).

For a function u^* we consider taking random linear combinations of a few random affine shifts of u^*. Notice that if u^* was actually a low-degree polynomial, then the resulting function would also be a low-degree polynomial (since low-degree polynomials are closed under taking affine shifts and taking linear combinations). We show that if u^* is far from low-degree polynomials, this operation amplifies distance to low-degree polynomials noticeably. More precisely, suppose V is the space of polynomials of degree at most ρq, let $\delta > 0$, and suppose $\rho > 0$ is small enough as a function of δ. We show that if u^* is δ-far from V, then the function $u(x) = u^*(x) + c \cdot u^*(ax + b)$ (where a, b, c are picked uniformly at random from $\mathbb{F}_q$) is with high probability $\approx (2\delta - \delta^2)$ far from V. This final distance matches what one would expect if we took the sum of two random functions that were δ-far from V - thus the random affine shift of u^* behaves nearly independently of u^* (subject to the trivial constraint that the random affine shift is also δ-far from V).

To state the theorem, we begin with some notation. For a function $f : \mathbb{F}_q \to \mathbb{F}_q$, we denote by $T_{a,b}(f)$ the function $g : \mathbb{F}_q \to \mathbb{F}_q$ given by:

$$g(\beta) = f(a\beta + b),$$

for each $\beta \in \mathbb{F}_q$.

Theorem 5.1. *Let $V = \mathsf{RS}(\mathbb{F}_q, (1-\lambda)q) \subseteq \mathbb{F}_q^q$ be the Reed-Solomon code over $\mathbb{F}_q$ with minimum distance λ. Let $u', u'' : \mathbb{F}_q \to \mathbb{F}_q$ be functions with $\Delta(u', V) \geq \delta'$ and $\Delta(u'', V) \geq \delta''$. Let $\epsilon > 0$, and let*

$$\delta = \min(J_\epsilon(J_\epsilon(\lambda)) - \epsilon, \delta' + \delta'' - \delta'\delta'' - 2\epsilon).$$

Then:

$$\Pr_{a,b,c \in \mathbb{F}_q}[\Delta(u' + c \cdot T_{a,b}(u''), V) < \delta] \leq \frac{K}{q}, \tag{8}$$

where $K = 8/\epsilon^4$.

Proof. Set $\overline{\delta} = \delta + \epsilon$. Note that

$$\overline{\delta} = \delta + \epsilon < J_\epsilon(J_\epsilon(\lambda)) < J_\epsilon(\lambda).$$

Suppose Equation (8) did not hold. Thus:

$$\Pr_{a,b,c \in \mathbb{F}_q}[\Delta(u' + c \cdot T_{a,b}(u''), V) < \overline{\delta} - \epsilon] > \frac{K}{q}.$$

Then with probability at least $\frac{K}{2q}$ over the choice of (a, b), we have that:

$$\Pr_{c \in \mathbb{F}_q}[\Delta(u' + c \cdot T_{a,b}(u''), V) < \overline{\delta} - \epsilon] > \frac{K}{2q}.$$

Fix such an $(a, b) \in \mathbb{F}_q^2$. Since $K > 4/\epsilon^3$and $\overline{\delta} < J_\epsilon(J_\epsilon(\lambda))$, we may apply Theorem 4.4 to u' and $T_{a,b}(u'')$. It tells us that there exist $y, y^* \in V$ such that:

$$|\{\beta \in \mathbb{F}_q \mid u'(\beta) = y(\beta) \wedge u''(a\beta + b) = y^*(\beta)\}| \geq (1 - \overline{\delta})q, \tag{9}$$

which, after letting $y^{**}(T) = y^*((\beta - b)/a))$, can be rewritten as:

$$|\{\beta \in \mathbb{F}_q \mid u'(\beta) = y(\beta) \wedge u''(a\beta + b) = y^{**}(a\beta + b)\}| \geq (1 - \overline{\delta})q. \tag{10}$$

It is thus natural to consider the collection of polynomials close to u', u'':

$$\mathcal{L}' = \{f \in V \mid \Delta(u', f) \leq \overline{\delta}\},$$
$$\mathcal{L}'' = \{f \in V \mid \Delta(u'', f) \leq \overline{\delta}\},$$

as well as the collection of agreement sets:

$$\mathcal{F}' = \{A \subseteq \mathbb{F}_q \mid \text{ for some } f \in \mathcal{L}' \text{ we have } A = \{\beta \in \mathbb{F}_q \mid f(\beta) = u'(\beta)\}\}.$$
$$\mathcal{F}' = \{A \subseteq \mathbb{F}_q \mid \text{ for some } f \in \mathcal{L}'' \text{ we have } A = \{\beta \in \mathbb{F}_q \mid f(\beta) = u''(\beta)\}\}.$$

By the Johnson bound, Theorem 2.2, (and since $\overline{\delta} < J_\epsilon(\lambda)$), we have that

$$|\mathcal{L}'|, |\mathcal{L}''|, |\mathcal{F}'|, |\mathcal{F}''| < 1/\epsilon.$$

Equation (10) and the discussion before it tells us that with probability at least $\frac{K}{2q}$ over the choice of $(a, b) \in \mathbb{F}_q^2$, there exists some $A' \in \mathcal{F}'$ and some $A'' \in \mathcal{F}''$ such that

$$|A' \cap \frac{1}{a}(A'' - b)| \geq (1 - \overline{\delta})q.$$

By averaging, this means that there must exist some $A' \in \mathcal{F}'$ and $A'' \in \mathcal{F}''$ such that with probability at least $\frac{\epsilon^2 K}{2q}$ over the choice of $(a, b) \in \mathbb{F}_q^2$,

$$|A' \cap \frac{1}{a}(A'' - b)| \geq (1 - \overline{\delta})q. \tag{11}$$

We will use this to deduce that either A' or A'' must be big. For each $r \in A''$, let X_r denote the indicator random variable for the event that $(r - b)/a \in A'$. Let $X = \sum_{r \in A''} X_r$. Note that

$$X = |A' \cap \frac{1}{a}(A'' - b)|.$$

It is easy to see that $\mathbf{E}[X_r] = |A'|/q$, and so:

$$\mathbf{E}[X] = \frac{|A'| \cdot |A''|}{q} = \mu.$$

Furthermore, the X_r are pairwise independent, and thus the variance of X is bounded by:

$$\mathbf{Var}[X] \leq 4\frac{|A''||A'|}{q} \leq 4q.$$

Thus:

$$\Pr[X \geq \mu + 2t\sqrt{q}] \leq \frac{1}{t^2}.$$

If $|A'|, |A''|$ are such that $|A'| \cdot |A''| \leq (1 - \overline{\delta} - \epsilon) \cdot q^2$, then $\mu \leq (1 - \overline{\delta} - \epsilon)q$, and the above equation with $t = \frac{\epsilon}{2}\sqrt{q}$ gives us that:

$$\Pr[X \geq (1 - \overline{\delta})q] \leq \frac{4}{\epsilon^2 q}.$$

This is a contradiction to Equation (11), since by the choice of K,

$$\frac{\epsilon^2 K}{2q} < \frac{4}{\epsilon^2 q}.$$

Thus we must have that:

$$|A'| \cdot |A''| > (1 - \overline{\delta} - \epsilon)q^2.$$

Recalling that $A' \in \mathcal{F}'$ and $A'' \in \mathcal{F}''$, we conclude that

$$(1 - \delta')(1 - \delta'') > (1 - \delta - 2\epsilon),$$

a contradiction to our assumption on δ', δ''. □

6 A low-agreement analysis of the Polischuck–Spielman axis-parallel test

In this section, we use the tools we developed above to give improved versions of the Polischuk-Spielman robust low-degree test [PS94] in certain settings. Their result gives a way to robustly test proximity of a 2-dimensional function $f : \mathbb{F}_q^2 \to \mathbb{F}_q$ to bivariate polynomials with individual degrees (d, ℓ). Our result shows that for $\ell = O(1)$, and for $d = O(q)$, the Polischuk-Spielman low-degree test works even in the presence of high noise: even if the test passes with some tiny probability η, it means that the underlying bivariate function has nontrivial agreement with some low degree bivariate polynomial.

The original Polischuk-Spielman analysis (improving on Arora-Safra [AS98]) allows d, ℓ to both be $\Omega(q)$, but could only conclude something if the passing probability η was at least $1/2$. The very recent analysis of the Polischuk-Spielman test due to Chiesa et al. [CMS17] allows η to be small, as in the result we obtain below, but the two results are incomparable (neither implies the other). The result of [CMS17] works for d, ℓ as large as $O(\log q)$ whereas ours requires $\ell = O(1)$ but allows d to be as large as $\Omega(q)$.

Theorem 6.1 (High-error soundness analysis of the Polischuk–Spielman test)**.** *Let $d = \rho q$. Suppose for each $x \in \mathbb{F}_q$, we have a degree ℓ polynomial $Q_x(Y)$, and for each $y \in \mathbb{F}_q$ we have a degree d polynomial $P_y(X)$. Suppose that for some* agreement parameter $\eta > 1 - J_\epsilon^{[\ell+1]}(1-\rho) + \epsilon$ *all these polynomials meet the following* non-trivial agreement condition*:*

$$\Pr_{x,y \in F_q}[Q_x(y) = P_y(x)] \geq \eta. \tag{12}$$

Then there exists a bivariate polynomial $R(X, Y)$ of individual degree (d, ℓ) such that

$$\Pr_{x,y \in \mathbb{F}_q}[Q_x(y) = R(x, y)] \geq \eta - 2\epsilon,$$

$$\Pr_{x,y \in \mathbb{F}_q}[Q_x(y) = P_y(x) = R(x, y)] \geq \left(\frac{\epsilon}{2}\right)^{\ell+2} \cdot \eta.$$

Proof. Our plan is to use Theorem 4.5 to deduce some information about Q_x and P_y. Let $V \subseteq \mathbb{F}_q^q$ be the Reed-Solomon code of polynomials of degree at most d. Let $\lambda = \Delta(V) = 1 - \rho$.

Let $u^*, u_1, \ldots, u_\ell$ be functions from $\mathbb{F}_q$ to $\mathbb{F}_q$ such that:

$$Q_x(Y) = u^*(x) + u_1(x)Y + u_2(x)Y^2 + \ldots + u_\ell(x)Y^\ell.$$

For each $\alpha \in \mathbb{F}_q$, define $v^\alpha(X) = P_\alpha(X)$.

The non-trivial agreement hypothesis of Equation (12) tells us that:

$$\Pr_{\alpha,x \in \mathbb{F}_q}[u^*(x) + u_1(x)\alpha + \ldots + u_\ell(x)\alpha^\ell = v^\alpha(x)] \geq \eta.$$

Equivalently:

$$\mathbb{E}_{\alpha \in \mathbb{F}_q}[\Delta(u^* + \alpha u_1 + \alpha^2 u_2 + \ldots + \alpha^\ell u_\ell, v^\alpha)] \leq 1 - \eta.$$

Set $\delta = 1 - \eta + \epsilon$. By an averaging argument, we get:

$$\Pr_{\alpha \in \mathbb{F}_q}[\Delta(u^* + \alpha u_1 + \alpha^2 u_2 + \ldots + \alpha^\ell u_\ell, v^\alpha) < \delta] \geq \epsilon.$$

Let A be the set of α for which the above event happens: thus $|A| \geq \epsilon \cdot q$.

Note that $\delta < J_\epsilon^{[\ell+1]}(\lambda)$, and thus we can apply Theorem 4.5. We get that there exist $v^*, v_1, \ldots, v_\ell \in V$ and a subset $G \subseteq \mathbb{F}_q$ with $|G| \geq (1-\delta-\epsilon)q$ for all $x \in G$,

$$v^*(x) = u^*(x), v_1(x) = u_1(x), \ldots, v_\ell(x) = u_\ell(x).$$

Since v^* and the v_i are all in V, they are polynomials of degree at most d. Define

$$R(X,Y) = v^*(X) + v_1(X)Y + \ldots + v_\ell(X)Y^\ell.$$

Rephrasing what we just concluded in terms of R, we get that for all $x \in G$:

$$R(x,Y) = Q_x(Y),$$

and thus:

$$\Pr_{x \in \mathbb{F}_q, y \in \mathbb{F}_q}[R(x,y) = Q_x(y)] \geq 1 - \delta - \epsilon = \eta - 2\epsilon.$$

Moreover, we conclude from Item 2 of Theorem 4.5 that for at least $\left(\frac{\epsilon}{2}\right)^{\ell+1}$ fraction of the $\alpha \in A$, we have:

$$v^\alpha = v^* + \alpha v_1 + \ldots + \alpha^\ell v_\ell.$$

For any such α where this identity holds, we get that:

$$R(x,\alpha) = v^\alpha(x) = P_\alpha(x),$$

and thus

$$\Pr_{x \in \mathbb{F}_q}[R(x,\alpha) = P_\alpha(x) = Q_x(\alpha)] \geq 1 - \delta = \eta - \epsilon > \eta/2.$$

Overall:

$$\Pr_{\alpha \in \mathbb{F}_q, x \in \mathbb{F}_q}[R(x,\alpha) = P_\alpha(x) = Q_x(\alpha)] > \left(\frac{\epsilon}{2}\right)^{\ell+2} \cdot \eta.$$

This completes the proof of the theorem. □

As a sample application, if we want to take $\eta = 0.01$, then for any ℓ we can take[4] $\rho = \epsilon = (0.001)^{2^{\ell+1}}$. Thus for $\ell = O(1)$ we can take $\rho = \Omega(1)$ (and thus $d = \Omega(q)$), and for $\ell < C \log\log q$ we can take $\rho = q^{-o(1)}$ (and thus $d = q^{1-o(1)}$).

7 Improved soundness for the Fast RS IOPP (FRI) protocol

In this section we describe how our prior results lead to a better analysis of the soundness of the FRI protocol of [BBHR18b]. We start by recalling the notation introduced in [BBHR18b, Sections 3.4, 4.2.1].

Our starting point is a function $f^{(0)} : L^{(0)} \to \mathbb{F}$ where $\mathbb{F}$ is a finite field of characteristic 2 and size 2^n, the evaluation domain $L^{(0)} \subset \mathbb{F}$ is an affine space over the two element field $\mathbb{F}_2$, i.e., $L^{(0)}$ is a coset of an additive subgroup of $\mathbb{F}$, and $|L^{(0)}| = 2^{k^{(0)}}$ which means that $k^{(0)} = \dim\left(L^{(0)}\right)$. We assume the target rate is $\rho = 2^{-\mathcal{R}}$ for some positive integer $\mathcal{R}$. The FRI protocol of [BBHR18b] is a two-phase protocol (whose two phases are called COMMIT and QUERY) that convinces a verifier that $f^{(0)}$ is close to the Reed-Solomon code $\mathsf{RS}[\mathbb{F}, L^{(0)}, \rho]$. We now state the previously known result about FRI and our improvement to it.

[4]Here we use the crude upper bound $J_\epsilon(1-\gamma) \geq 1 - \sqrt{\epsilon} - \sqrt{\gamma}$. This implies that $J_\epsilon^{[\ell]}(1-\gamma) \geq 1 - 2\epsilon^{1/2^\ell} - \gamma^{1/2^\ell}$ provided $\epsilon < 2^{-2^\ell}$.

7.1 Statement of results

The following is the main theorem from [BBHR18b], and we improve its soundness in Theorem 7.2, stated after it.

Theorem 7.1 (FRI — main properties)**.** *The following properties hold when the* FRI *protocol is invoked on oracle* $f^{(0)} : L^{(0)} \to \mathbb{F}$ *with rate* $\rho = 2^{-\mathcal{R}}$ *for* $\mathcal{R} \in \mathbb{N}^+$ *such that* $\rho|L^{(0)}| > 16$*:*

1. **Completeness** *If* $f^{(0)} \in \mathsf{RS}^{(0)} \triangleq \mathsf{RS}[\mathbb{F}, L^{(0)}, \rho = 2^{-\mathcal{R}}]$ *and* $f^{(1)}, \ldots, f^{(\mathsf{r})}$ *are computed by the prover specified in the* COMMIT *phase, then the* FRI *verifier outputs* accept *with probability* 1*.*

2. **Soundness** *Suppose* $\delta^{(0)} \triangleq \Delta^{(0)}\left(f^{(0)}, \mathsf{RS}^{(0)}\right) > 0$*. Then with probability at least*

$$1 - \frac{3|L^{(0)}|}{|\mathbb{F}|} \tag{13}$$

over the randomness of the verifier during the COMMIT *phase, and for any (adaptively chosen) prover oracles* $f^{(1)}, \ldots, f^{(\mathsf{r})}$*, the* QUERY *protocol with repetition parameter* ℓ *outputs* accept *with probability at most*

$$\left(1 - \min\left\{\delta^{(0)}, \frac{1 - 3\rho - 2/\sqrt{|L^{(0)}|}}{4}\right\}\right)^{\ell} \tag{14}$$

Consequently, the soundness of FRI *is at least*

$$\mathbf{s}^{-}\left(\delta^{(0)}\right) \triangleq 1 - \left(\frac{3|L^{(0)}|}{|\mathbb{F}|} + \left(1 - \min\left\{\delta^{(0)}, \frac{1 - 3\rho - 2/\sqrt{|L^{(0)}|}}{4}\right\}\right)^{\ell}\right). \tag{15}$$

3. **Prover complexity** *is* $O(|L^{(0)}|)$ *arithmetic operations over* $\mathbb{F}$

4. **Verifier complexity** *is* $O(\log |L^{(0)}|)$ *arithmetic operations over* $\mathbb{F}$ *for a single invocation of the* QUERY *phase; this also bounds communication and query complexity (measured in field elements).*

We improve FRI soundness as follows:

Theorem 7.2 (FRI with improved soundness)**.** *The following properties hold when the* FRI *protocol is invoked on oracle* $f^{(0)} : L^{(0)} \to \mathbb{F}$*, with rate* $\rho = 2^{-\mathcal{R}}, \mathcal{R} \in \mathbb{N}^+$ *such that* $\rho|L^{(0)}| > 16$*:*

1. **Soundness** *Suppose* $\delta^{(0)} \triangleq \Delta^{(0)}\left(f^{(0)}, \mathsf{RS}^{(0)}\right) > 0$*. Then for any* $\epsilon > 0$*, with probability at least*

$$1 - \frac{2\log|L^{(0)}|}{\epsilon^3|\mathbb{F}|} \tag{16}$$

over the randomness of the verifier during the COMMIT *phase, and for any (adaptively chosen) prover oracles* $f^{(1)}, \ldots, f^{(\mathsf{r})}$*, the* QUERY *protocol with repetition parameter* ℓ *outputs* accept *with probability at most*

$$\left(1 - \min\left\{\delta^{(0)}, J_\epsilon(J_\epsilon(1-\rho))\right\} + \epsilon\log|L^{(0)}|\right)^{\ell} \tag{17}$$

Consequently, the soundness of FRI *is at least*

$$\mathbf{s}^{-}\left(\delta^{(0)}\right) \triangleq 1 - \left(\frac{2\log|L^{(0)}|}{\epsilon^3|\mathbb{F}|} + \left(1 - \min\left\{\delta^{(0)}, J_\epsilon(J_\epsilon(1-\rho))\right\} + \epsilon\log|L^{(0)}|\right)^{\ell}\right). \tag{18}$$

7.2 Proof of Theorem 7.2

Before presenting the proof of our main theorem for this section, we briefly recall the FRI protocol and some notation. Our presentation essentially follows the original presentation of [BBHR18b], but slightly reformulates some steps in terms of an "algebraic hash function" H_x.

7.2.1 Outline

The COMMIT phase of the FRI protocol involves $\mathsf{r} = k^{(0)} - \mathcal{R}$ rounds. Before any communication, the prover and verifier agree on some special $\mathbb{F}_2$-subspaces $L^{(i)}$, where $\dim(L^{(i)}) = k^{(0)} - i$. Let $\mathsf{RS}^{(i)}$ denote the Reed-Solomon code $\mathsf{RS}[\mathbb{F}, L^{(i)}, \rho]$.

The main ingredient of the FRI protocol is a special algebraic hash function H_x, which takes a seed $x \in \mathbb{F}$, and given as input a function $f : L^{(i)} \to \mathbb{F}$, it produces as output a hash whose length is $1/2$ as long as f. More concretely, $H_x[f]$ is a function

$$H_x[f] : L^{(i+1)} \to \mathbb{F}$$

with the following properties:

1. For any $s \in L^{(i+1)}$, $H_x[f](s)$ can be computed by querying f at just two points in its domain (these two points are $(q^{(i)})^{-1}(s)$).
2. If $f \in \mathsf{RS}^{(i)}$, then for all $x \in \mathbb{F}$, we have that $H_x[f] \in \mathsf{RS}^{(i+1)}$.
3. Finally, Theorem 4.4 implies that if $\Delta(f, \mathsf{RS}^{(i)}) \geq \delta$ for some $\delta \leq J_\epsilon(J_\epsilon(1-\rho))$, then

$$\Pr_{x \in \mathbb{F}}[\Delta(H_x[f], \mathsf{RS}^{(i)}) < \delta - \epsilon] < \frac{1}{\epsilon^3 |\mathbb{F}|}.$$

   In words, we have that with high probability over the choice of seed x, $H_x[f]$ is as far from $\mathsf{RS}^{(i+1)}$ as f is from $\mathsf{RS}^{(i)}$.

These last two properties roughly show that for random x, H_x preserves distance to Reed-Solomon codes.

The high-level idea of the FRI protocol can then be described as follows. First we are in the COMMIT phase of the protocol. The verifier picks a random $x^{(0)} \in \mathbb{F}$ and asks the prover to write down the hash $H_{x^{(0)}}[f^{(0)}] : L^{(1)} \to \mathbb{F}$. By Properties 2 and 3 above, our original problem of estimating the distance of $f^{(0)}$ to $\mathsf{RS}^{(0)}$ reduces to estimating the distance of $H_{x^{(0)}}[f^{(0)}]$ to $\mathsf{RS}^{(1)}$ (which is a problem of $1/2$ the size). This process is then repeated: the verifier picks a random $x^{(1)} \in \mathbb{F}$ and asks the prover to write down $H_{x^{(1)}}[H_{x^{(0)}}[f^{(0)}]]$, and so on. After r rounds of this, we are reduced to a constant sized problem which can be solved in a trivial manner. However, the verifier cannot blindly trust that the functions $f^{(1)}, \ldots$ that were written down by the prover truly are obtained by repeatedly hashing $f^{(0)}$. This has to be checked, and the verifier does this in the QUERY phase of the protocol, using Property 1 above.

We describe the phases of the protocol below.

COMMIT **Phase:**

1. For $i = 0$ to $\mathsf{r} - 1$:
   (a) The verifier picks uniformly random $x^{(i)} \in \mathbb{F}$ and sends it to the prover.
   (b) The prover writes down a function $f^{(i+1)} : L^{(i+1)} \to \mathbb{F}$. (In the case of an honest prover, $f^{(i+1)} = H_{x^{(i)}}[f^{(i)}]$.)
2. The prover writes down a constant $C \in \mathbb{F}$. (In the case where $f^{(0)} \in \mathsf{RS}^{(0)}$, then $f^{(\mathsf{r})}$ should be a constant function, and in this case we expect the honest prover to make $C =$ that constant.)

QUERY **Phase:** (executed by the Verifier)

1. Repeat ℓ times:
   (a) Pick $s^{(0)} \in L^{(0)}$ uniformly at random.
   (b) For $i = 0$ to $\mathsf{r} - 1$:
      i. Define $s^{(i+1)} \in L^{(i+1)}$ by $s^{(i+1)} = q^{(i)}(s^{(i)})$.
      ii. Compute $H_{x^{(i)}}[f^{(i)}](s^{(i+1)})$ by making 2 queries to $f^{(i)}$.
      iii. If $i < \mathsf{r} - 1$ and $f^{(i+1)}(s^{(i+1)}) \neq H_{x^{(i)}}[f^{(i)}](s^{(i+1)})$, then REJECT
      iv. If $i = \mathsf{r} - 1$ and $C \neq H_{x^{(i)}}[f^{(i)}](s^{(i+1)})$, then REJECT
2. ACCEPT

7.2.2 The algebraic hash function

We now describe the algebraic hash function H_x.

The description of the hash function requires fixing some choices of certain subspaces. For each $i \in [0, \mathsf{r}]$ we choose $\mathbb{F}_2$-subspaces $L_0^{(i)}$ and $L^{(i)}$, satisfying the following properties.

1. $L_0^{(i)} \subseteq L^{(i)}$ with $\dim(L_0^{(i)}) = 1$,
2. $L^{(i+1)} = q^{(i)}(L^{(i)})$, where $q^{(i)}(X)$ is the *subspace polynomial* of $L_0^{(i)}$ (i.e., $q^{(i)}(X) = \prod_{\alpha \in L_0^{(i)}}(X - \alpha)$, thus this is an $\mathbb{F}_2$-linear map with kernel $L_0^{(i)}$). In particular, $\dim(L^{(i+1)}) = \dim(L^{(i)}) - 1$.

Let $\mathcal{S}^{(i)}$ denote the set of cosets of $L_0^{(i)}$ contained in $L^{(i)}$.

Given $x \in \mathbb{F}$ and $f : L^{(i)} \to \mathbb{F}$, the hash of f with seed x is defined to be the function $H_x[f] : L^{(i+1)} \to \mathbb{F}$ as follows. For $s \in L^{(i+1)}$, let $s_0, s_1 \in L^{(i)}$ be the two roots of $q^{(i)}(X) - s$. Let $P_{f,s}(X) \in \mathbb{F}[X]$ be the unique degree ≤ 1 polynomial satisfying

$$P_{f,s}(s_0) = f(s_0),$$
$$P_{f,s}(s_1) = f(s_1).$$

Then we define

$$H_x[f](s) = P_{f,s}(x). \tag{19}$$

Observe that $H_x[f](s)$ can be computed by querying f on the set $\{s_0, s_1\}$ (this set is a coset of $L_0^{(i)}$, and we denote it by $S_s^{(i)}$).

To understand H_x better, it is instructive to see what it does to $\mathsf{RS}^{(i)}$. Let $f \in \mathsf{RS}^{(i)}$. The underlying polynomial $f(X)$ thus has degree at most $\rho|L^{(i)}|$. We may write $f(X)$ in base $q^{(i)}(X)$ as:

$$f(X) = a_0(X) + a_1(X)q^{(i)}(X) + \ldots + a_t(X)(q^{(i)}(X))^t, \tag{20}$$

where each $a_i(X)$ has degree at most 1, and $t \leq \rho|L^{(i)}|/2$. Since the polynomials $f(X)$ and $P_{f,s}(X)$ agree on the roots of $q(X) - s$, we get that $f(X) \equiv P_{f,s}(X) \mod (q^{(i)}(X) - s)$. From Equation (20), we get that

$$P_{f,s}(X) = a_0(X) + a_1(X)s + \ldots + a_t(X)s^t.$$

In particular, for all $x \in \mathbb{F}$,

$$H_x[f](s) = P_{f,s}(x) = a_0(x) + a_1(x)s + \ldots + a_t(x)s^t,$$

and thus

$$H_x[f] \in \mathsf{RS}^{(i+1)}.$$

Next we will show that for f far away from $\mathsf{RS}^{(i)}$, for a random x we have that $H_x[f]$ is far from $\mathsf{RS}^{(i+1)}$ with high probability.

7.2.3 A corollary of Theorem 4.4

For $f, g : L^{(i)} \to \mathbb{F}$ let $\Delta^{(i)}(f,g)$ be the block-wise distance between f, g (cf. [BBHR18b, Definition 3.2]), defined as the fraction of cosets of $L_0^{(i)}$ on which f and g differ,

$$\Delta^{(i)}(f,g) \triangleq \Pr_{S \in \mathcal{S}^{(i)}}[f|_S \neq g|_S]$$

where $f|_S$ is the restriction of f to S (and $g|_S$ is similarly defined) and equality above is in the space $\mathbb{F}^S$. Notice $\Delta^{(i)}(f,g) \geq \Delta(f,g)$. For a set of functions $V \subset \mathbb{F}^{L^{(i)}}$ let $\Delta^{(i)}(f,V) = \min\left\{\Delta^{(i)}(f,v) \mid v \in V\right\}$.

The following statement is a corollary of Theorem 4.4.

Corollary 7.3. *Let $i < \mathsf{r}$, and let $f : L^{(i)} \to \mathbb{F}$ be an arbitrary function. Let $\delta \triangleq \min(\Delta^{(i)}\left(f, \mathsf{RS}^{(i)}\right), J_\epsilon(J_\epsilon(1-\rho)))$. Then*

$$\Pr_{x \in \mathbb{F}}\left[\Delta\left(H_x[f], \mathsf{RS}^{(i+1)}\right) \leq \delta - \epsilon\right] \leq \frac{2}{\epsilon^3|\mathbb{F}|}. \tag{21}$$

Proof. Consider the space of functions $U = \{H_x[f] \mid x \in \mathbb{F}\} \subset \mathbb{F}^{L^{(i+1)}}$. Let

$$u^* = H_0[f], \quad u = H_1[f] - H_0[f].$$

Since $\deg(P_{f,s}) \leq 1$ for every $s \in L^{(i+1)}$ it follows that every $u' = H_x[f] \in U$ can be written as a linear combination of u^*, u; specifically, $H_x[f] = u^* + x \cdot u$. Let $\bar{U} \subseteq U$ be the set of elements in U that have distance less than $\delta - \epsilon$ to $\mathsf{RS}^{(i+1)}$.

Assume by way of contradiction that $|\bar{U}| > \frac{2}{\epsilon^3}$. Then Theorem 4.4 implies the existence of $v^*, v \in \mathsf{RS}^{(i+1)}$ and a subset $T \subset L^{(i+1)}$, $\frac{|T|}{|L^{(i+1)}|} \geq 1 - \delta$, such that $v^*|_T = u^*|_T$ and $v|_T = u|_T$. Let $Q^*(Y), Q(Y)$ be the polynomials interpolating v^* and v respectively. We have $\deg(Q^*), \deg(Q) < \rho|L^{(i+1)}|$ because $v^*, v \in \mathsf{RS}^{(i+1)}$. Let

$$\hat{Q}(X,Y) \triangleq Q^*(Y) + X \cdot Q(Y)$$

and notice that (i) $\deg_X(\hat{Q}) < 2$, $\deg_Y(\hat{Q}) < \rho|L^{(i+1)}|$ (ii) $\hat{Q}(0,Y) = Q^*(Y)$, (iii) $\hat{Q}(1,Y) = Q(Y)$.

Consider the polynomial $R(X) \triangleq \hat{Q}(X, q^{(i)}(X))$. We have

$$\deg(R) \leq 2 \cdot \deg_Y(\hat{Q}) - 1 < 2|L^{(i+1)}| = \rho|L^{(i)}|.$$

We claim that R agrees with f on $\left\{S_s^{(i)} \mid s \in T\right\}$. Indeed, for each $s \in T$ let $S_s^{(i)} = \{s_0, s_1\} \in \mathcal{S}^{(i)}$ be the pair of roots of the polynomial $q^{(i)}(X) - s$. By our assumption on T,

$$\hat{Q}(0,s) = H_0[f](s) = P_{f,s}(0) \text{ and } \hat{Q}(1,s) = H_1[f](s) = P_{f,s}(1).$$

The polynomials $\hat{Q}(X,s)$ and $P_{f,s}(X)$ are both of degree less than 2 and they agree on the two points $\{0,1\}$, hence they agree everywhere. It follows that

$$f(s_0) = \hat{Q}(s_0,s) = \hat{Q}\left(s_0, q^{(i)}(s_0)\right) = R(s_0)$$

and similarly $f(s_1) = R(s_1)$. Therefore, R and f agree on T, as claimed.

We have established $\Delta^{(i)}\left(f, \mathsf{RS}^{(i)}\right) \leq 1 - \frac{|T|}{|L^{(i+1)}|} \leq \delta$ and this contradicts our assumption. Therefore $|\bar{U}| \leq \frac{2}{\epsilon^3}$, as claimed. □

7.2.4 Proof of improved soundness

Armed with Corollary 7.3 we move on to the proof of the main theorem of this section.

Proof of Theorem 7.2. Let $\delta^{(i)} = \min(\Delta(f^{(i)}, \mathsf{RS}^{(i)}), J_\epsilon(J_\epsilon(1-\rho)))$. Let $\delta_\star^{(i)} = \min(\Delta^{(i)}(f^{(i)}, \mathsf{RS}^{(i)}), J_\epsilon(J_\epsilon(1-\rho)))$. Observe that $\delta^{(i)} \leq \delta_\star^{(i)}$.

Let $E^{(i)}$ be the "bad" event that $\Delta\left(H_{x^{(i)}}[f^{(i)}], \mathsf{RS}^{(i+1)}\right) \leq \delta^{(i)} - \epsilon$. Let $E_\star^{(i)}$ be the "bad" event that $\Delta\left(H_{x^{(i)}}[f^{(i)}], \mathsf{RS}^{(i+1)}\right) \leq \delta_\star^{(i)} - \epsilon$. Corollary 7.3 implies that $\Pr\left[E_\star^{(i)}\right] \leq \frac{2}{\epsilon^3|\mathbb{F}|}$. Thus[5]

$$\Pr\left[E^{(i)}\right] \leq \frac{2}{\epsilon^3|\mathbb{F}|}.$$

By the union bound

$$\Pr\left[\bigvee_{i=0}^{\mathsf{r}-1} E^{(i)}\right] \leq \frac{2\mathsf{r}}{\epsilon^3|\mathbb{F}|} \leq \frac{2k^{(0)}}{\epsilon^3|\mathbb{F}|} \tag{22}$$

We continue our analysis assuming no such event holds. Let $f^{(0)}, \ldots, f^{(\mathsf{r})}$ be the sequence of functions sent by the prover, which is not necessarily honest. For simplicity of notation, we only describe the analysis when the repetition parameter ℓ equals 1; the case of general ℓ follows triviall from this.

Recall that during the QUERY phase of the FRI protocol, the verifier selects a random $s^{(0)} \in L^{(0)}$ and this defines a sequence $s^{(0)}, \ldots, s^{(\mathsf{r})}$ inductively by using the rule $s^{(i+1)} = q^{(i)}(s^{(i)})$ for $i \in \{0, \ldots, \mathsf{r}-1\}$; Recall $S^{(i)}_{s^{(i+1)}} \in \mathcal{S}^{(i)}$ is the coset containing the two roots of the polynomial $q^{(i)}(X) - s^{(i+1)}$, and one of them is $s^{(i)}$. The test associated with $s^{(0)}$ accepts iff

$$f^{(i+1)}(s^{(i+1)}) = H_{x^{(i)}}[f^{(i)}](s^{(i+1)}) (= P_{f^{(i)}, s^{(i+1)}}(x^{(i)})). \tag{23}$$

holds for all $i \in \{0, \ldots, \mathsf{r}-1\}$ and additionally $f^{(\mathsf{r})}$ is a constant function; we assume it by associating the constant function with the first entry of $f^{(\mathsf{r})}$.

For the sake of analysis, consider the directed graph with vertex set $L^{(0)} \sqcup L^{(1)} \sqcup \ldots L^{(\mathsf{r})}$, in which an edge appears from $s^{(i)} \in L^{(i)}$ to $s^{(i+1)} \in L^{(i+1)}$ if and only if $s^{(i)} \in S^{(i)}_{s^{(i+1)}}$. This graph has $\mathsf{r}+1$ layers, and the vertices in the ith layer are the elements of $L^{(i)}$. Since the value of $f^{(\mathsf{r})} : L^{(\mathsf{r})} \to \mathbb{F}$ is enforced to be a constant function, it will be easier in our analysis to use a single node to represent $L^{(\mathsf{r})}$. Under this simplification, the resulting graph is a directed tree (we direct edges from leaves to root). For all nodes but for the leaves and root, the in-degree is 2; all non-root nodes have out-degree 1. A single invocation of the QUERY phase involves selecting a leaf $s^{(0)}$ and performing the sequence of tests along the path from $s^{(0)}$ to the root (which corresponds to $L^{(\mathsf{r})}$).

Call a vertex $s^{(i+1)} \in L^{(i+1)}$ *bad* if Equation (23) fails to hold for $s^{(i+1)}$. All other vertices are called *good*. Observe that a QUERY test rejects if and only if the path examined by it contains a bad vertex. To analyze the rejection probability of the test, it will be simpler to consider only the last such bad vertex along a path. To this end, we shall modify the sequence of functions $f^{(1)}, \ldots, f^{(\mathsf{r}-1)}$ (but not $f^{(0)}$ and $f^{(\mathsf{r})}$) in a way that may change some bad vertices into good ones, but will not make any good vertex bad. We will then prove a lower bound on the rejection probability of a QUERY test applied to the modified set of functions; this will give a lower bound on the rejection probability of a QUERY test applied to the original set of functions.

Working top down with $i = \mathsf{r}, \ldots, 2$ in decreasing order, for each bad vertex $s^{(i)} \in L^{(i)}$, we modify the entries in the sub-tree whose root is $s^{(i)}$, as follows. Let $L^{(j)}$ be the set of vertices in layer j that have a path to $s^{(i)}$. For $j \in \{0, \ldots, i-2\}$, in increasing order, set

$$f^{(j+1)}|_{L^{(j+1)}} = H_{x^{(j)}}[f^{(j)}]|_{L^{(j+1)}}.$$

[5]When $\delta^{(i)} < \frac{1-\rho}{2}$ the stronger bound $\Pr\left[E^{(i)}\right] \leq \frac{1}{\epsilon|\mathbb{F}|}$ holds.

Equivalently, for each $j \in \{0, \ldots, i-2\}$ and $s \in L^{(j+1)}$, we set:

$$f^{(j+1)}(s) = H_{x^{(j)}}[H_{x^{(j-1)}}[\ldots[H_{x^{(0)}}[f^{(0)}]]\ldots]](s).$$

This modification process may change the entries of $f^{(1)}, \ldots, f^{(\mathsf{r}-1)}$ but does not change neither $f^{(0)}$, nor $f^{(\mathsf{r})}$ because $0 \leq j \leq \mathsf{r}-2$ and we only modify entries in layer $j+1$, so we may only affect vertices in layers $1, \ldots, \mathsf{r}-1$. Crucially, the probability of rejecting during the QUERY phase does not increase as a result of this modification, because the modification does not turn a good vertex into a bad one and hence the set of post-modification bad vertices is a subset of the pre-modification bad vertices.

Consider the sequence of modified functions $f^{(0)}, \ldots, f^{(\mathsf{r})}$. Let $\beta^{(i)}$ denote the fraction of bad vertices in $L^{(i)}$. As said earlier, the probability of rejection during a single QUERY invocation is precisely the probability that a path originating in a random leaf passes through a bad vertex. After our modification process, for each pair of distinct bad vertices v_i, v_j, the set of leaves that have a path to v_i is distinct from the set of leaves that have a path to v_j; furthermore, along any path from leaf to root there is at most one bad vertex. Hence, the sets of leaves of bad vertices are pairwise disjoint. Thus, the probability that the FRI verifier rejects on a single invocation of the QUERY protocol is precisely $\sum_{=1}^{\mathsf{r}} \beta^{(i)}$. All that remains is to bound this sum from below, as done next.

Claim 7.4. *If $E^{(i)}$ does not hold, then*

$$\beta^{(i+1)} \geq \delta^{(i)} - \delta^{(i+1)} - \epsilon.$$

Proof. We may assume that $\delta^{(i+1)} \leq \delta^{(i)} - \epsilon$ (since otherwise the right hand side of the claimed inequality is negative, and the result is clear). This implies, by definition of $\delta^{(i)}$ and $\delta^{(i+1)}$, that

$$\delta^{(i+1)} < \delta^{(i)} \leq J_\epsilon(J_\epsilon(1-\rho)),$$

and thus $\delta^{(i+1)} = \Delta\left(f^{(i+1)}, \mathsf{RS}^{(i+1)}\right)$.

Assuming $E^{(i)}$ does not hold, we have that:

$$\Delta\left(H_{x^{(i)}}[f^{(i)}], \mathsf{RS}^{(i+1)}\right) \geq \delta^{(i)} - \epsilon$$

By the properties of the modification process, $f^{(i+1)}(s^{(i+1)}) = H_{x^{(i)}}[f^{(i)}](s^{(i+1)})$ for every vertex $s^{(i+1)} \in L^{(i+1)}$ that is not bad. By the triangle inequality:

$$\delta^{(i+1)} = \Delta\left(f^{(i+1)}, \mathsf{RS}^{(i+1)}\right) \geq \Delta\left(H_{x^{(i)}}[f^{(i)}], \mathsf{RS}^{(i+1)}\right) - \Delta(f^{(i+1)}, H_{x^{(i)}}[f^{(i)}]) \geq \delta^{(i)} - \epsilon - \beta^{(i+1)}$$

or, rearranging,

$$\beta^{(i+1)} \geq \delta^{(i)} - \delta^{(i+1)} - \epsilon,$$

as claimed. □

We continue with the proof of Theorem 7.2. By assumption we have $\delta^{(\mathsf{r})} = 0$, and $f^{(0)}$ is unchanged by the modification process, so

$$\delta^{(0)} = \delta^{(0)} - \delta^{(\mathsf{r})} = \sum_{i=0}^{\mathsf{r}-1} \delta^{(i)} - \delta^{(i+1)}$$

Applying Claim 7.4 to the rightmost term above we conclude that whenever no event $E^{(i)}$ holds (cf. Equation (22)), then the probability of the verifier rejecting during a single invocation of the QUERY phase is at least $\sum_{i=1}^{\mathsf{r}} \beta^{(i)} \geq \delta^{(0)} - \mathsf{r}\epsilon$. This completes the proof. □

Acknowledgements

Work supported by the USA–Israel binational science fund, grant # 2014359. SK and SS would like to thank the CS department at the Technion for the warm hospitality during this research.

References

[AHIV17] Scott Ames, Carmit Hazay, Yuval Ishai, and Muthuramakrishnan Venkitasubramaniam. Ligero: Lightweight sublinear arguments without a trusted setup. In *Proceedings of the 24th ACM Conference on Computer and Communications Security*, October 2017.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998. Preliminary version in FOCS '92.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: a new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998. Preliminary version in FOCS '92.

[AS03] Sanjeev Arora and Madhu Sudan. Improved low-degree testing and its applications. *Combinatorica*, 23(3):365–426, 2003. Preliminary version appeared in STOC '97.

[BBHR18a] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Scalable, transparent, and post-quantum secure computational integrity. Cryptology ePrint Archive, Report 2018/046, 2018. Available at `https://eprint.iacr.org/2018/046`.

[BBHR18b] Eli Ben-Sasson, Iddo Bentov, Ynon Horesh, and Michael Riabzev. Fast Reed-Solomon Interactive Oracle Proofs of Proximity. In *Proceedings of the 45th International Colloquium on Automata, Languages, and Programming (ICALP)*, 2018.

[BCF+16] Eli Ben-Sasson, Alessandro Chiesa, Michael A. Forbes, Ariel Gabizon, Michael Riabzev, and Nicholas Spooner. On probabilistic checking in perfect zero knowledge. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:156, 2016.

[BCGV16] Eli Ben-Sasson, Alessandro Chiesa, Ariel Gabizon, and Madars Virza. Quasilinear-size zero knowledge from linear-algebraic PCPs. In *Proceedings of the 13th Theory of Cryptography Conference*, TCC '16, pages 33–64, 2016.

[BCS16] Eli Ben-Sasson, Alessandro Chiesa, and Nicholas Spooner. Interactive oracle proofs. In *Theory of Cryptography - 14th International Conference, TCC 2016-B, Beijing, China, October 31 - November 3, 2016, Proceedings, Part II*, pages 31–60, 2016.

[BFL90] László Babai, Lance Fortnow, and Carsten Lund. Nondeterministic exponential time has two-prover interactive protocols. In *Proceedings of the 31st Annual Symposium on Foundations of Computer Science*, SFCS '90, pages 16–25, 1990.

[BFLS91] László Babai, Lance Fortnow, Leonid A. Levin, and Mario Szegedy. Checking computations in polylogarithmic time. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, STOC '91, pages 21–32, 1991.

[BM88] László Babai and Shlomo Moran. Arthur-merlin games: A randomized proof system, and a hierarchy of complexity class. *J. Comput. Syst. Sci.*, 36(2):254–276, April 1988.

[CMS17] Alessandro Chiesa, Peter Manohar, and Igor Shinkar. On axis-parallel tests for tensor product codes. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2017, August 16-18, 2017, Berkeley, CA, USA*, pages 39:1–39:22, 2017.

[GMR89] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. *SIAM Journal on Computing*, 18(1):186–208, 1989. Preliminary version appeared in STOC '85.

[Gur07] Venkatesan Guruswami. Algorithmic results in list decoding. *Foundations and Trends in Theoretical Computer Science*, 2(2):107–195, 2007.

[HS00] Prahladh Harsha and Madhu Sudan. Small PCPs with low query complexity. *Computational Complexity*, 9(3–4):157–201, Dec 2000. Preliminary version in STACS '01.

[LFKN92] Carsten Lund, Lance Fortnow, Howard J. Karloff, and Noam Nisan. Algebraic methods for interactive proof systems. *Journal of the ACM*, 39(4):859–868, 1992.

[PS94] Alexander Polishchuk and Daniel A. Spielman. Nearly-linear size holographic proofs. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing*, STOC '94, pages 194–203, 1994.

[RVW13] Guy N. Rothblum, Salil Vadhan, and Avi Wigderson. Interactive proofs of proximity: delegating computation in sublinear time. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 793–802. ACM, 2013.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il