Electronic Colloquium on Computational Complexity, Report No. 166 (2025)

Optimal Proximity Gaps for Subspace-Design Codes and (Random) Reed-Solomon Codes

Rohan Goyal[*] Venkatesan Guruswami[†]

October 2025

Abstract

Reed-Solomon (RS) codes were recently shown to exhibit an intriguing *proximity gap* phenomenon. Specifically, given a collection of strings with some algebraic structure (such as belonging to a line or affine space), either all of them are δ-close to RS codewords, or most of them are δ-far from the code. Here δ is the proximity parameter which can be taken to be the Johnson radius $1 - \sqrt{R}$ of the RS code (R being the code rate), matching its best known list-decodability. Proximity gaps play a crucial role in the soundness analysis of Interactive Oracle Proof (IOP) protocols used in Succinct Non-Interactive Arguments of Knowledge (SNARKs) and the resulting proof sizes.

Proving proximity gaps beyond the Johnson radius, and in particular approaching $1 - R$ (which is best possible), has been posed multiple times as a challenge with significant practical consequences to the efficiency of SNARKs. Here we prove that variants of RS codes, such as folded RS codes and univariate multiplicity codes, indeed have proximity gaps for δ approaching $1 - R$. The result applies more generally to codes with a certain subspace-design property. Our proof hinges on a clean property we abstract called line (or more generally curve) decodability, which we establish leveraging and adapting techniques from recent progress on list-decoding such codes. Importantly, our analysis avoids the heavy algebraic machinery used in previous works, and requires a field size only linear in the block length.

The behavior of subspace-design codes w.r.t "local properties" has recently been shown to be similar to random linear codes and random RS codes (where the evaluation points are chosen at random from the underlying field). We identify a local property that implies curve decodability, and thus also proximity gaps, and thereby conclude that random linear and random RS codes also exhibit proximity gaps up to the $1 - R$ bound. Our results also establish the stronger (mutual) correlated agreement property which implies proximity gaps. Additionally, we also a show a *slacked* proximity gap theorem for constant-sized fields using AEL-based constructions and local property techniques.

[*] Massachusetts Institute of Technology, Cambridge rohan_g@mit.edu.

[†] Simons Institute for the Theory of Computing, and the University of California, Berkeley. venkatg@berkeley.edu.

ISSN 1433-8092

Contents

1 Introduction

The concept of *proximity gaps* is a fascinating one that sprung out of work on ultra-efficient Interactive Oracle Proof (IOP) protocols for certifying the proximity of functions to low-degree polynomials (i.e., codewords of a corresponding Reed-Solomon code). Such IOPs are in turn used to construct SNARKs (succinct non-interactive arguments of knowledge), which are protocols for parties to generate a short, quick-to-verify proof that they know a piece of information or have performed a computation, with natural applications in blockchain and related technologies. See [BSCS16, BSBHR18, BCI$^+$20] for detailed discussions of IOPs, SNARKs, and their role in blockchains.

In one of its key steps, the verifier in such an IOP protocol, for example the famous FRI protocol [BBHR18] and its extensions [BSGKS19, ZCF23, ACFY24a, ACFY24b], test the claim that all points, say on a line or curve, are close to Reed-Solomon codewords by checking such proximity, for efficiency purposes, only for a few random choices among those points. The soundness analysis of the protocol then naturally hinges on so-called proximity gaps for Reed-Solomon codes.

An error-correcting code $\mathcal{C} \subset \Sigma^n$ is said to exhibit *proximity gaps* with respect to a family $\mathcal{F}$ of subsets of Σ^n, if for every $S \in \mathcal{F}$, either *all* elements of S are within (relative Hamming) distance δ of some codeword of $\mathcal{C}$, or *most* elements of S are δ-far from the code $\mathcal{C}$. In other words, the fraction of elements of S that are close $\mathcal{C}$ exhibits a *gap*—either it equals 1 or is at most err for some negligible $\mathsf{err} > 0$. Therefore, checking the proximity to the code $\mathcal{C}$ of a *random* element of S gives good confidence that, in fact, *every* element of S is δ-close to $\mathcal{C}$. The quantity δ is called the proximity parameter, and ε is the soundness. Both these parameters contribute to the failure of an IOP employing a proximity test, and it is important to have a large proximity parameter and small soundness. Of these, the proximity gap is the more significant constraint, as the soundness can often be reduced by working with larger fields. Typical families $\mathcal{F}$ of interest in protocols are algebraic structures such as lines, curves, or affine spaces.

We now review known proximity gap results focusing on the proximity parameter. For every linear code $\mathcal{C}$ of relative distance $\delta_{\mathcal{C}}$, proximity gaps of subspaces up to proximity parameter $\delta < \delta_{\mathcal{C}}/3$ was established in [BKS18, AHIV22]. Going beyond the unique-decoding radius, proximity gaps up to $\delta < 1 - \sqrt[3]{1-\delta_{\mathcal{C}}}$ (the so-called "1.5 Johnson radius") were established in [BSGKS19]. They also showed that this is tight for certain Reed-Solomon codes with code length equal to the field size.

In a fascinating work, Ben-Sasson, Carmon, Ishai, Kopparty, and Saraf [BSCI$^+$23] showed stronger proximity gaps for Reed-Solomon (RS) codes over large enough fields. Specifically, they showed proximity gaps of affine spaces for $\delta < (1-R)/2$ when the field size $q \gg n/\mathsf{err}$, and for $\delta < 1-\sqrt{R}$ when $q \gg n^2/\operatorname{poly}(\mathsf{err})$, where R is the rate of the RS code. Recall that the relative distance of rate R RS codes equals $\delta_{\mathrm{RS}} = 1 - R$, so these bounds correspond to $\delta_{\mathrm{RS}}/2$ and $1-\sqrt{1-\delta_{\mathrm{RS}}}$, which coincide with the unique-decoding radius and the Johnson radius (which is the best known list-decoding radius) respectively of Reed-Solomon codes. Indeed, BCIKS proved these results by extending ideas from algebraic unique and list decoding algorithms in a highly non-trivial way.

Until this work, there was no code known with proximity gaps exceeding the $1-\sqrt{R}$ bound. Improving the proximity gap from the current state-of-the-art $1-\sqrt{R}$ to a bound approaching $1-R$ would reduce proof sizes in SNARKs by 50% and therefore have significant practical consequences. (We note that $1-R$ is the best possible bound, as for a rate R code, every point is $1-R$ close to some codeword.) In light of this, the challenge of improving the proximity gaps to approach $1-R$ has been explicitly highlighted in the literature, e.g., see Conjecture 8.4 of [BCI$^+$20]. Such a result has significant and direct relevance to the concrete security guarantees of practical implementations such as ethSTARK codebase (see Section 7 of [BS21]). In practice, when designing IOPPs for hash-based SNARKs, it is common to assume that Reed–Solomon codes have proximity gaps with low soundness err for all $\delta \in (0, 1 - R)$ (see, e.g., [BSGKS19, ACFY24a, ACFY24b]). The Ethereum Foundation recently announced the proximity prize initiative [Eth25, Hex25] related to the conjectures on whether Reed-Solomon codes have optimal proximity gaps.

1.1 Our results

With the above backdrop, we now turn to stating our results. Informally, we show that most codes in the literature that have been shown to achieve list-decoding capacity, i.e., are *list-decodable* up to a fraction of errors approaching $1 - R$, exhibit proximity gaps for $\delta < 1 - R$. This includes folded Reed-Solomon codes [GR08], univariate multiplicity codes [GW13a, Kop14], random linear codes [AGG$^+$25], and random Reed-Solomon codes [BGM22, AGG$^+$25] (which are sampled by choosing random evaluation points from the underlying field).

The specific property of codes we tie to proximity gaps is that of being a subspace design. Subspace-design codes with near-optimal parameters are known to achieve list-decoding capacity. We prove that they also have proximity gaps for proximity parameters approaching the optimal $1 - R$ bound. This yields as direct corollaries the results for folded RS and multiplicity codes, as these codes are known to have good subspace design properties. By carefully tapping into recent results [BCDZ25b] that reveal tight connections between local properties of subspace-design codes and random linear codes, as well as random Reed-Solomon codes [LMS25], we establish that random RS codes indeed have proximity gaps for $\delta \to 1 - R$. The previous best known results were upto Johnson bound due to [GCXK25] for all the aforementioned families of codes. This, however, requires some technical work as proximity gaps cannot directly be cast as local properties. At the core of our approach to establish proximity gaps is a clean definition we introduce called line-decodability (or more generally curve-decodability). We will discuss the key ideas driving our results in Section 1.2.

Before stating our results, let us define proximity gaps formally. Our results establish proximity gaps for lines, curves, and affine spaces, but for simplicity we focus on lines in the introduction. A code $\mathcal{C} \subseteq \Sigma^n$ is said to be $\mathbb{F}$-additive if its alphabet Σ is a vector space over a finite field $\mathbb{F}$ and $\mathcal{C}$ is linear over $\mathbb{F}$. For a vector $u \in \Sigma^n$, $\Delta(u, \mathcal{C})$ denotes relative Hamming distance from u to its closest codeword in $\mathcal{C}$.

Definition 1.1. *An $\mathbb{F}$-additive code $\mathcal{C} \subseteq \Sigma^n$ has (δ, err)-proximity gap (for lines) if for all $u_0, u_1 \in \Sigma^n$, if $\Pr_{\alpha \in \mathbb{F}}[\Delta(u_0 + \alpha u_1, \mathcal{C}) \leq \delta] > \mathsf{err}$ then $\Delta(u_0 + \alpha u_1, \mathcal{C}) \leq \delta$ for all $\alpha \in \mathbb{F}$.* ⌟

A stronger property called *correlated agreement* [BKS18, BSCI$^+$23] guarantees that in fact there are codewords $c_0, c_1 \in \mathcal{C}$ such that on at least $(1-\delta)n$ positions $i \in [n]$, we have *both* the agreements $u_{0,i} = c_{0,i}$ and $u_{0,i} = c_{0,i}$. Our results also establish correlated agreement, and in fact also an even stronger property called mutual correlated agreement defined in [ACFY24b], but again for simplicity we restrict to proximity gaps for the introduction.

Theorem 1.2 (Optimal proximity gaps for folded RS code)**.** *For all $R, \eta, \mathsf{err} \in (0, 1)$, a rate R folded Reed-Solomon code over alphabet $\mathbb{F}^s$ and block length n has $(1 - R - \eta, \mathsf{err})$-proximity gap (in fact, mutual correlated agreement) for lines, if $s \gtrsim 1/\eta^2$ and $\mathsf{err} \cdot |\mathbb{F}| \gtrsim \frac{n}{\eta} + \frac{1}{\eta^3}$. The same holds for order-s univariate multiplicity codes.*

We want to draw attention to the field size in Theorem 1.2 which only needs to linear in the block length. In contrast, the proof of proximity gaps up to the Johnson radius in [BCI$^+$20] required a quadratic field size.

We next state our result for normal Reed-Solomon codes, albeit with random evaluation points.

Theorem 1.3 (Random RS codes have optimal proximity gaps)**.** *For all $R, \eta \in (0, 1)$ and large enough n, for any $\mathsf{err} > 0$, a rate R Reed-Solomon code with n random evaluation points from $\mathbb{F}_q$ has $(1 - R - \eta, \mathsf{err})$-proximity gap (in fact, mutual correlated agreement) for lines with high probability, if $\mathsf{err} \cdot q \gtrsim \frac{n}{\eta} + \frac{1}{\eta^5}$.*

A similar result holds for random linear codes; see Section 5.5. Again note that the result applies for field sizes linear in the block length. We can also get results over *constant-sized* fields when a small slack is allowed in the proximity parameter, for example when we guarantee that either most points on a line are δ-far from the code, or all points on the line are within, say, 1.01δ of some codeword. We show this both for random linear codes and explicit expander based codes (using the Alon-Edmonds-Luby transform to translate from random to explicit following [JS25]). See Sections 4.4 and 5.5 for formal statements.

1.2 Overview of approach and main ideas

We now discuss the main ideas underlying our proofs of Theorems 1.2 and 1.3.

As mentioned earlier, the claim of Theorem 1.2 holds more generally for any subspace-design code with optimal parameters, of which folded Reed-Solomon and multiplicity codes are prominent examples [GK16]. Informally, a rate R code $\mathcal{C} \subseteq (\mathbb{F}^s)^n$ is a good subspace-design code if for every low-dimensional subspace A of $\mathcal{C}$, the average dimension of $A \cap \mathcal{C}_i$ for random $i \in [n]$ drops to $\approx R \cdot \dim(A)$, where $\mathcal{C}_i = \{c \in \mathcal{C} \mid c_i = 0\}$.

Such subspace design codes are known to achieve list-decoding capacity, i.e., are list-decodable up to radius $1 - R - \eta$ for any constant $\eta > 0$. In fact they even do so with optimal list-size of $O(1/\eta)$ [Sri25, CZ25].

However, we do not know how to show proximity gaps directly from list-decoding. One of the main ingredients in our approach is a clean definition of *line-decodability* (and more generally, curve-decodability). Informally, we say a code is (δ, a, b) line-decodable if for every line in the universe of the form $u_0 + \alpha u_1$, and an arbitrary set of codewords $c_1, \cdots, c_a$ which are δ-close to distinct points on the line, at least b of these codewords must also lie on a corresponding line.

We prove that line-decodability implies proximity gaps, and in fact (mutual) correlated agreement. This is since if there are many close codewords on a single line, then the entire line is close to the line of codewords. We also show that the parameters can be significantly improved if one further assumes the list-decodablity of the code (which we know holds for subspace-design codes). These proofs are short and modular, and appear in Section 3.2.

To prove line-decodability of subspace-design codes, we first prove a result on list-recovering a subspace-design code into a small subspace of solutions given as input low-dimensional (as opposed to small) input spaces in each position.

Once, we have a low dimensional linear space $\mathcal{H}$ of dimension r, we can simply randomly sample a set $S = \{s_1, \cdots, s_r\}$ of r coordinates of $[n]$. If $\mathcal{H} \cap (\bigcap_{j=1}^{r} \mathcal{C}_{s_j}) = \{0\}$ then, for any pair $y \in \Sigma^n$ and $c \in \mathcal{H}$ if $\Delta(y, c) \leq \delta_C - \eta$, then $y_{|S} = c_{|S}$ holds with probability at least η^r. This is precisely the idea behind [KRSW23, Tam24] for the task of list decoding as they gave the first constant size list bounds for FRS codes.

We observe that one particularly nice fact about this algorithm for *pruning* an affine space is that it can be run to just get a set S of coordinates independent of y, c. Thus, we can run the algorithm for all y, c pairs with the same shared randomness! In particular, this implies the existence of a set S on which many such pairs will agree. Complete agreement on a single set can immediately be leveraged to give us line decodability and consequently proximity gaps for FRS codes, albeit with rather poor parameters.

One major drawback of the above pruning algorithm is the low success probability. Recently, [AHS25] gave an efficient list-decoding algorithm through a step-by-step pruning with a more judicious distribution to choose the next coordinate at every step. A technically simple but conceptually powerful novelty that we introduce here is to pin positions without fixing codeword values at those positions. The algorithm of [AHS25] was proposed for the task of list decoding, i.e., to find elements close to a codeword y in a low dimensional affine space $\mathcal{H}$. They repeatedly sample coordinates $i \in [n]$ from a clever distribution closely tied to the subspace-design property and restrict to the subspace of $\mathcal{H}$ which is equal to y_i on the ith coordinate. They show that during this pruning process, any codeword within $\mathcal{H}$ close to y survives with good probability. This immediately implies a list-size upper bound.

We modify the above algorithm to always fix coordinates to 0, and only at the end look for codewords that agree with y on the set of fixed coordinates. Using this idea, we can run the pruning algorithm along the entire line in one go with the same randomness. We therefore obtain a pruning approach that will give us a set S on which many points on the line completely agree with their close-by codewords. As before, we can now argue that all such codewords must lie on a line! This implies line-decodability and consequently proximity gaps.

Our main contribution here is to abstract the right, flexible notions around proximity gaps, which then allows us to leverage and adapt recent progress on list-decoding and properties of random linear and random RS codes. As a result, our techniques are quite modular and general, and we hope will be applicable more broadly.

We now turn to Theorem 1.3.

A recent body of results has shown that folded Reed-Solomon codes, random linear codes, and random RS codes have similar behaviors w.r.t all "local" properties [LMS25, BCDZ25b], formally called local coordinate-wise linear (LCL) properties, which capture list-decoding and list-recovery. This understanding has been the culmination of a long series of recent works beginning with [MRRZ$^+$21, BGM22] and continuing and converging along various directions in [GLM$^+$22, GMR$^+$22, GM24, AGG$^+$25, GXYZ24, BDGZ25, BCDZ25b, JS25].

Informally, a b-LCL property $\mathcal{P}$ is a collection of so-called "local profiles" $\mathcal{V} = (\mathcal{V}_1, \cdots, \mathcal{V}_n)$ where $\mathcal{V}_i$ are linear subspaces of $\mathbb{F}_q^b$. A code $\mathcal{C} \subseteq \mathbb{F}_q^n$ does not have the property $\mathcal{P}$ if there is no matrix $N \in \mathbb{F}_q^{n \times b}$ and profile $\mathcal{V} \in \mathcal{P}$ such that each column is a distinct codeword, and the ith row is in $\mathcal{V}_i$.

It is easy to see that one cast having a list of $L + 1$ distinct codewords which are all close to some codeword or product space as a local property. Thus, if a code does not contain this property, then it would be list-decodable or list-recoverable with list-size L respectively.

In [LMS25], it was established that every local profile $\mathcal{V}$ has a threshold $R_{\mathcal{V}}$ such that a random linear code (RLC) of rate R random linear code is extremely unlikely to contain this profile when $R \leq R_{\mathcal{V}} - \varepsilon$ and extremely likely to contain the profile when $R \geq R_{\mathcal{V}} + \varepsilon$. They further showed that the same threshold applies to random Reed-Solomon (RRS) codes. Thus, any LCL property is likely to start occurring at roughly the same threshold rate for RLCs and RRS codes.

A recent work [BCDZ25b] showed that the thresholds for any LCL properties of RLCs are the same as the thresholds for subspace-design codes. In a companion paper [BCDZ25a], they proved near-optimal list-recovery guarantees for subspace-design codes. They then applied the results of [BCDZ25b] to translate the list recovery guarantees to random linear and random RS codes. This is also the approach we take toward proximity gaps for RLCs and RRS codes in this paper, though as explained below, proximity gaps do not directly fit the local property framework. In another recent work relating properties of random and explicit codes, [JS25] showed that all *reasonable* local properties of RLCs also hold for explicit expander (AEL) based codes with constant alphabet size.

Recall that to to establish proximity gaps using our approach, we need to establish line-decodability. However, line-decodability cannot be directly captured as a LCL property and so does not fit the framework of the above-mentioned general theories relating subspace-design codes to RRS codes via RLCs. The reason is a bit technical, as it concerns a local collection of codewords no three of which are the same, whereas the local profiles of [LMS25] only consider collections of pairwise distinct codewords. If the definition allowed two codewords to be equal but no three equal or similar more permissive conditions, then the framework would apply directly.

It is quite likely that the theory in [LMS25, BCDZ25b] can be reworked with this slightly more general notion of local profiles. However, to avoid reworking a lot of technical details and to use their theory as a black-box, we instead manage to cast a notion related to line-decodability, which we call V-decodability, directly as an LCL property. This V-decodability property is equivalent to line-decodability up to some modest loss in parameters, and therefore holds for subspace-design codes. Applying the theory in [LMS25, BCDZ25b], we establish V-decodabilty and hence also line-decodability and optimal proximity gaps for random linear codes and random Reed-Solomon codes, yielding Theorem 1.3. The details appear in Section 5.

1.3 Concurrent Works

In a concurrent work, [DG25] proved that the general conjecture of [BCI+20] cannot hold up to $1-R-\eta$ when R,η are $o(1/\sqrt{n})$ for any MDS codes. They then refine the conjecture to state that it works when R,η are constants bounded away from 0. Our work does not deal with the case when η is $o(1/\sqrt{n})$ so this impossibility does not apply for us.

Additionally, [CS25] also showed that the aforementioned conjecture cannot hold when list-decoding with small lists is not possible. Again this does not apply to us since our work addresses regimes where optimal list-size bounds are already known.

[BSCH+25] show improved error bounds for general Reed-Solomon codes and that correlated agreement cannot hold up to capacity for some specific Reed-Solomon codes. Our work complements their results by showing that correlated agreement does hold up to capacity for *random* Reed-Solomon codes.

2 Preliminaries

We begin by introducing the basic coding-theoretic definitions we will be using.

Definition 2.1 (Fractional Hamming Distance)**.** *For any two vectors x,y in Σ^n where Σ is some alphabet, we define $\Delta(x,y)=\frac{|\{i\in[n]:x_i\neq y_i\}|}{n}$ to be the fraction of coordinates where they differ.*

For a set $S\subseteq\Sigma^n$, we define $\Delta(x,S)=\min_{y\in S}\frac{|\{i\in[n]:x_i\neq y_i\}|}{n}$ to be the minimum fractional Hamming distance of x to its closest vector in S. ⌟

The two fundamental quantities associated with a code are its rate and distance.

Definition 2.2 (Distance of a code)**.** *For code $\mathcal{C} \subseteq \Sigma^n$, we define its (relative) distance as* $\Delta(\mathcal{C}) = \min_{x,y \in \mathcal{C}, x \neq y} \Delta(x, y)$ ⌟

Definition 2.3 (Rate of a code)**.** *For a code $\mathcal{C} \subseteq \Sigma^n$, its rate $R(\mathcal{C})$ is defined as* $R(\mathcal{C}) = \frac{\log_{|\Sigma|} |\mathcal{C}|}{n}$ ⌟

We will focus on linear/additive codes in this paper, defined as follows.

Definition 2.4 (Additive codes)**.** *Let $\mathbb{F}$ be a finite field and let $\Sigma = \mathbb{F}^s$ for some positive integer s. A code $\mathcal{C} \subseteq \Sigma^n$ is said to be $\mathbb{F}$-additive (or just additive when the field $\mathbb{F}$ is clear from context or not relevant to the discussion) if $\mathcal{C}$ is an $\mathbb{F}$-linear subspace of Σ^n. When $s = 1$, the code is simply called a linear code.* ⌟

2.1 Proximity Gaps and Correlated Agreement

The study of proximity gaps was initiated in [RVW13] as a property of interest of general linear codes. They proved a slacked proximity gap result nothing about the distance of the code space at all. Since then, proximity testing was connected to important problems in the area of building SNARKs [AHIV22, RZ18, BKS18]. In subsequent works, proximity gaps for general linear codes [BKS18], and then for Reed-Solomon codes [BCI+20] became focuses of study.

In our definitions, we add back the *slack* parameter that was present in the previous works [RVW13, BKS18].

There are three definitions related to proximity gaps that we could desire from a code increasing strength. We will present all three definitions and note that our results achieve all three definitions.

Definition 2.5 (Proximity Gaps)**.** *An $\mathbb{F}$-additive code $\mathcal{C} \subseteq \Sigma^n$ is said to have $(\ell, \delta, \mathsf{err}, \gamma)$ proximity gap if for all $u_0, u_1, \cdots, u_\ell \in \Sigma^n$, $\delta' \leq \delta$*

$$\Pr_{\alpha \in \mathbb{F}}\left[\Delta\Big(\sum_{j=0}^{\ell} u_j \cdot \alpha^j, \mathcal{C}\Big) \leq \delta'\right] > \mathsf{err} \implies \forall \alpha \in \mathbb{F},\ \Delta\Big(\sum_{j=0}^{\ell} u_j \cdot \alpha^j, \mathcal{C}\Big) \leq \frac{\delta'}{1-\gamma}\ .$$

When $\ell = 1$, we also use the terminology $(\delta, \mathsf{err}, \gamma)$ line proximity gap for this notion.[1] ⌟

The following stronger notion was defined in [BCI+20] but [BKS18] already proved correlated agreement theorems to get their proximity gaps results as well.

Definition 2.6 (Correlated Agreement)**.** *An $\mathbb{F}$-additive code $\mathcal{C} \subseteq \Sigma^n$ is said to have $(\ell, \delta, \mathsf{err}, \gamma)$ correlated agreement if for all $u_0, u_1, \cdots, u_\ell \in \Sigma^n$, $\delta' \leq \delta$*

$$\Pr_{\alpha \in \mathbb{F}}\left[\Delta\Big(\sum_{j=0}^{\ell} u_j \cdot \alpha^j, \mathcal{C}\Big) \leq \delta'\right] > \mathsf{err} \implies \exists c_0, c_1, \cdots c_\ell \in \mathcal{C} \ s.t.\ |\{i \mid \bigvee_{j=0}^{\ell} c_{j,i} \neq u_{j,i}\}| \leq \frac{\delta' n}{1-\gamma}\ .$$

In particular, this implies that for all $\alpha \in \mathbb{F}$, we have that $\Delta\left(\sum_{j=0}^{\ell} u_j \cdot \alpha^j, \sum_{j=0}^{\ell} c_j \cdot \alpha^j\right) \leq \frac{\delta'}{1-\gamma}$. When $\ell = 1$, we also use the terminology $(\delta, \mathsf{err}, \gamma)$ line correlated agreement for this notion. ⌟

2.1.1 Mutual correlated agreement

[ACFY24b] defined a further strengthening of correlated agreement called *Mutual Correlated Agreement* that they use within the soundness analysis of their protocol. The authors also express a belief

[1] Note that setting $\gamma \leq 1/n$ has the same effect as setting $\gamma = 0$.

that their definition would be useful in soundness analysis of other protocols as well. We modify their definition slightly into a form that is easier to work with. Our definition implies theirs.

[ACFY24b, GKL24, Zei24] established mutual correlated agreement for various linear codes and regimes as well matching some of the previous best known results for correlated agreement and proximity gaps.

Informally, the mutual correlated agreement definition says that even if there is a special subclass of codewords that the curve is often close to, then there is a curve of close by codewords passing through at least one of the codewords in this special subclass.

Definition 2.7 (Mutual Correlated Agreement)**.** *An $\mathbb{F}$-additive code $\mathcal{C} \subseteq \Sigma^n$ is said to have $(\ell, \delta, \mathsf{err}, \gamma)$ mutual correlated agreement if for all $u_0, u_1, \cdots, u_\ell \in \Sigma^n$, $\delta' \leq \delta$, and $f : \mathbb{F}_q \to \mathcal{C}$, if we let $u_\alpha = \sum_{j=0}^{\ell} u_j \cdot \alpha^j$, then*

$$\Pr_{\alpha \in \mathbb{F}} \left[\Delta \left(u_\alpha, f(\alpha) \right) \leq \delta' \right] > \mathsf{err} \implies \exists c_0, c_1, \cdots c_\ell \in \mathcal{C} \ s.t. \ |\{i \mid \bigvee_{j=0}^{\ell} c_{j,i} \neq u_{j,i}\}| \leq \frac{\delta' n}{1-\gamma} \ ,$$

and there exists some $\beta \in \mathbb{F}$ such that $f(\beta) = \sum_{j=0}^{\ell} c_j \beta^j$.

In particular, this implies that for all $\alpha \in \mathbb{F}$, we have that $\Delta \left(\sum_{j=0}^{\ell} u_j \cdot \alpha^j, \sum_{j=0}^{\ell} c_j \cdot \alpha^j \right) \leq \frac{\delta'}{1-\gamma}$. When $\ell = 1$, we also use the terminology $(\delta, \mathsf{err}, \gamma)$ line mutual correlated agreement for this notion. ⌟

Remark 2.8. *We have currently defined $f : \mathbb{F}_q \mapsto \mathcal{C}$ since the randomness for the linear combination is defined by a single element of $\mathbb{F}_q$ i.e. α in the above definition. For a more generic definition, one can simply define $f : \mathsf{Supp}(Randomness\ of\ Linear\ Combination) \mapsto \mathcal{C}$.* ⌟

Remark 2.9. *Mutual Correlated Agreement was originally defined in [ACFY24b] to stipulate the following property:*

$$\Pr_{\alpha}[\exists S \subseteq [n] \mid |S| \geq (1-\delta)n, \ \exists c \in \mathcal{C} : c_{|S} = \sum_{j=0}^{\ell} \alpha^j u_{j|S}, \ \exists j \in [\ell] \ s.t. \ \forall c' \in \mathcal{C}, c'_{|S} \neq u_{j|S}] \leq \mathsf{err} \ .$$

We note that our definition implies theirs. For any set S of size at least $(1-\delta)n$ such that $\forall j \in [\ell], \exists c'_j$ s.t. $c'_{j|S} = u_{j|S}$, do not define $f(\alpha)$ to agree with u_α on S. ⌟

Remark 2.10. *We also introduce an extra parameter in the correlated agreement i.e. γ which allows for some slack in the final conclusion. In many applications, it is sufficient even if γ is a small constant and in these cases, we do not require that $|\mathbb{F}| \geq n$. In particular, the field size can simply be a constant while allowing for any constant slack.*

As part of our main results, we demonstrate that for any constant slack γ, constant error probability, constant dimension ℓ and $\delta \leq 1 - R - \varepsilon$ where ε is constant, random linear codes with constant field size achieve this definition. Additionally, explicit codes constructed through AEL based methods also achieve constant field size while achieving this definition. ⌟

Lemma 2.11. *Let $\mathcal{C}$ be an $\mathbb{F}$-additive code. Then*

- *If $\mathcal{C}$ has $(\ell, \delta, \mathsf{err}, \gamma)$ correlated agreement then it has $(\ell, \delta, \mathsf{err}, \gamma)$ proximity gap.*
- *If $\mathcal{C}$ has $(\ell, \delta, \mathsf{err}, \gamma)$ mutual correlated agreement then it has $(\ell, \delta, \mathsf{err}, \gamma)$ correlated agreement.*

Proof. The notions are defined to be stronger than one another. □

2.1.2 Proximity Gaps for Lines imply proximity gaps for Affine Spaces

The following proofs are implicit in [BSCI+23].

Definition 2.12 (Proximity Gaps for Spaces)**.** *An $\mathbb{F}$-additive code $\mathcal{C} \subseteq \Sigma^n$ is said to have $(\delta, \mathsf{err}, \gamma)$ space proximity gap if for all affine spaces $U \subseteq \Sigma^n$, $\delta' \leq \delta$*

$$\Pr_{u \in U}[\Delta(u, \mathcal{C}) \leq \delta'] > \mathsf{err} \implies \forall u \in U : \Delta(u, \mathcal{C}) \leq \frac{\delta'}{1-\gamma} \qquad \lrcorner$$

Lemma 2.13. *If an $\mathbb{F}_q$-additive code $\mathcal{C}$ has $(\delta, \mathsf{err}, \gamma)$ line proximity gap then it has $(\delta, \mathsf{err} \cdot \frac{q}{q-1}, \gamma)$ space proximity gap.*

Proof. Let $\delta' \leq \delta$ and $U \subseteq \Sigma^n$ be an affine space such that $\Pr_{u \in U}[\Delta(u, \mathcal{C}) \leq \delta'] > \mathsf{err} \cdot \frac{q}{q-1}$. We would like to prove that $\Delta(u, \mathcal{C}) \leq \frac{\delta'}{1-\gamma}$ for all $u \in U$.

Now, let $u_0 \in U$ be arbitrary. If $\Delta(u_0, \mathcal{C}) \leq \frac{\delta'}{1-\gamma}$ then there is nothing to prove.

Else, consider a uniformly random $u' \in U \setminus \{u_0\}$ and the line $u_0 + \alpha(u' - u_0)$ for $\alpha \in \mathbb{F}$. Observe that this line contains $q - 1$ distinct points other than u_0 and each point from $U \setminus \{u_0\}$ lies on it with equal probability.

This means that there must be a $u' \in U \setminus \{u_0\}$ for which

$$\Pr_{\alpha \in \mathbb{F} \setminus \{0\}}[\Delta(u_0 + \alpha(u_0 - u'), \mathcal{C}) \leq \delta'] > \mathsf{err} \cdot \frac{q}{q-1} \ .$$

Thus, $\Pr_{\alpha \in \mathbb{F}}[\Delta(u_0 + \alpha(u_0 - u'), \mathcal{C}) \leq \delta'] > \mathsf{err}$ and we can apply the line proximity gap hypothesis. In particular, this means that $\forall \alpha \in \mathbb{F}$, we have $\Delta(u_0 + \alpha(u_0 - u'), \mathcal{C}) \leq \frac{\delta'}{1-\gamma}$. Now, plugging in $\alpha = 0$, tells us that $\Delta(u_0, \mathcal{C}) \leq \frac{\delta'}{1-\gamma}$ as desired. □

Definition 2.14 (Correlated Agreement for spaces)**.** *An $\mathbb{F}$-additive code $\mathcal{C} \subseteq \Sigma^n$ is said to have $(\delta, \mathsf{err}, \gamma)$ space correlated agreement if for all affine spaces, $U = u_0 + \mathsf{span}(u_1, \cdots, u_\ell) \subseteq \Sigma^n$, $\delta' \leq \delta$: if*

$$\Pr_{u \in U}[\Delta(u, \mathcal{C}) \leq \delta'] > \mathsf{err}$$

then there exists an affine space $C = c_0 + \mathsf{span}(c_1, \cdots, c_\ell) \subseteq \mathcal{C}$ such that

$$\{i \mid \bigvee_{j=0}^{\ell} c_{j,i} \neq u_{j,i}\}| \leq \frac{\delta' n}{1-\gamma} \ .$$

In particular, this implies that $\Delta(c_0 + \alpha_1 c_1 + \cdots + \alpha_\ell c_\ell, u_0 + \alpha_1 u_1 + \cdots \alpha_\ell u_\ell) \leq \delta'/(1-\gamma)$ for all $\alpha_1, \cdots, \alpha_\ell \in \mathbb{F}$. ⌟

Lemma 2.15. *If an $\mathbb{F}_q$ additive code $\mathcal{C}$ has the following properties:*

- $(\delta, \mathsf{err}, 0)$ *line-correlated agreement.*
- *For all $x \in \Sigma^n$, $|\{c \in \mathcal{C} \mid \Delta(y, c) \leq \delta\}| < q$.*
- $\delta < \Delta(\mathcal{C})$.

then $\mathcal{C}$ has $(\delta, \mathsf{err} \cdot \frac{q}{q-1}, 0)$ space correlated agreement.

Proof. Let $\delta' \leq \delta$ and $U \subseteq \Sigma^n$ be an affine space such that $\Pr_{u \in U}[\Delta(u, \mathcal{C}) \leq \delta'] > \mathsf{err} \cdot \frac{q}{q-1}$. By, Lemma 2.13, we have that $\Delta(u, \mathcal{C}) \leq \delta'$ for all $u \in U$.

Let $\delta^* = \max_{u \in U} \Delta(u, \mathcal{C})$ and $u^* = \arg\max_{u \in U} \Delta(u, \mathcal{C})$.

Additionally, let $c_1, \cdots c_L = \{c \in \mathcal{C} \mid \Delta(u^*, c) \leq \delta^*\}$. We know that $L < q$ and that $\Delta(u^*, c_j) = \delta^*$ for all $j \in [L]$.

Let $T_j = \{i \in [n] \mid c_{j,i} = u_i^*\}$. We have that $|T_j| = (1 - \delta^*) \cdot n$ for all $j \in [L]$. Let $\mathcal{U}_i = \{u \in U \mid \exists c \in \mathcal{C} \text{ s.t. } T_j \subseteq \{i \mid u_i = c_i\}\}$.

Consider any u' such that $u' + u^* \in U$. By applying the line correlated agreement to u^*, u', we get that there exists codewords c^* and c' such that if $T = \{i \mid c_i^* = u_i^* \wedge u_i' = c_i'\}$ then $|T| \geq (1 - \delta^*)\, n$. Thus, $T = T_j$ for some $j \in [n]$, and thus $u' + u^* \in \mathcal{U}_j$ for some $j \in [L]$. Thus, $\mathcal{U} = \bigcup_{j=1}^{L} \mathcal{U}_j)$. Thus, there exists a $\mathcal{U}_j$ such that $|\mathcal{U}_j| > |\mathcal{U}|/q$.

Now, if we consider $U_{|T_j}$ then we have that it is a subspace and so is $\mathcal{C}_{|T_j}$. We know that $|U_{|T_j} \cap \mathcal{C}_{|T_j}| > |U_{|T_j}|/q$ but then we must have that $U_{|T_j} \subseteq \mathcal{C}_{|T_j}$. Now, we can just find the codewords corresponding to this space and we are done. □

Remark 2.16. *We note that we are only aware of reductions from lines to affine spaces for Correlated Agreement in this perfect agreement setting, and no reduction for Mutual Correlated Agreement from lines or curves.* ⌟

2.2 Subspace-Design Codes

We now review the important concept of subspace designs which drives our work. Subspace designs were introduced in [GX13] as a way to pre-code certain Reed-Solomon and algebraic-geometric codes so that they could then be list-decoded with small lists. Informally, a subspace design consists of a collection of subspaces $\mathcal{H}_i$ of some ambient space $\mathbb{F}_q^m$ such that every low-dimensional space has non-trivial intersection with few of them. While [GX13] used random constructions of subspace designs, explicit constructions of subspace designs were given in [GK16], curiously using list-decodable codes such as folded RS and univariate multiplicity codes to define the subspaces.

Below, we isolate this subspace design defining property of such codes abstractly, and call them subspace-design codes. A similar concept was also defined in [CZ25] where they called them subspace-designable codes.

Definition 2.17 (Subspace-Design Property)**.** *For any function $\tau : \mathbb{N} \to \mathbb{R}_{\leq 1}$, an $\mathbb{F}_q$-additive code $\mathcal{C} \subseteq (\mathbb{F}_q^s)^n$ is said to be a τ-subspace design code if for every $r \in \mathbb{N}$, and every $\mathbb{F}_q$-linear subspace $\mathcal{A}$ of $\mathcal{C}$ of dimension at most r, the following holds:*

$$\frac{\sum_{i=1}^{n} \dim \mathcal{A}_i}{n} \leq \dim(\mathcal{A}) \cdot \tau(r)$$

where $\mathcal{A}_i = \{a \in \mathcal{A} \mid a_i = 0\}$. ⌟

Lemma 2.18. *For any τ-subspace-design $\mathbb{F}_q$-additive code of rate R, we must have $\tau(r) \geq R - \frac{1}{n}$ for all $r \in \mathbb{N}$.*

Proof. We just need to prove the result for $r = 1$ since we can just take $\mathcal{A}$ of dimension at 1. Pick any non-zero codeword c which is zero on at least $Rn - 1$ coordinates. Such a codeword exists since the code is additive and there are $|\Sigma|^{Rn}$ different codewords. Now, define the 1-dimensional subspace $\mathcal{A} = \{\alpha \cdot c \mid \alpha \in \mathbb{F}_q\}$. Clearly, $\dim \mathcal{A}_i = 1$ for at least $R - \frac{1}{n}$ fraction of coordinates and so $\tau(1) \geq R - 1/n$. □

We also isolate a special case of subspace-design codes which have $\tau(r)$ extremely close to R for dimensions r that are not too large.

Definition 2.19 (Strong Subspace Design Code)**.** *A code $\mathcal{C} \subseteq \Sigma^n$ is called a d-strong subspace-design code, if for all $d' \leq d$, and any linear subspace $\mathcal{A}$ of dimension d' of $\mathcal{C}$, we have*

$$\sum_{i=1}^{n} \dim \mathcal{A}_i \leq R \cdot d' \cdot n + 1 \ . \qquad \lrcorner$$

The subspace design property has proven to be extremely useful and surprisingly powerful, and is at the heart of the recent breakthrough list size-bounds for list-decoding of folded Reed-Solomon codes (defined below) and univariate multiplicity codes established in [Sri25, CZ25, AHS25]. All of these results in fact apply more generally to the class of subspace-design codes. We also prove some of our key results in the generic language of subspace-design codes.

Definition 2.20 (s-Folded Reed-Solomon Codes [GR08])**.** *Let n, k, s be positive integers and $\mathbb{F}_q$ be a field with $q > sn$ and γ be an element of the field. Additionally, let $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}_q$ be distinct elements such that $\alpha_i \gamma^t \neq \alpha_j$ for all $i \neq j$ and $t < s$.*

The s-Folded Reed-Solomon Code $\mathcal{C} \subseteq (\mathbb{F}_q^s)^n$ with message length k on $(\alpha_1, \ldots, \alpha_n)$ is given by:

$$\mathsf{FRS}_{q,n,k,s}(\bar{\alpha}) = \left\{ \left(\begin{bmatrix} f(\alpha_1) \\ f(\alpha_1\gamma) \\ \vdots \\ f(\alpha_1\gamma^{s-1}) \end{bmatrix}, \begin{bmatrix} f(\alpha_2) \\ f(\alpha_2\gamma) \\ \vdots \\ f(\alpha_2\gamma^{s-1}) \end{bmatrix}, \cdots, \begin{bmatrix} f(\alpha_n) \\ f(\alpha_n\gamma) \\ \vdots \\ f(\alpha_n\gamma^{s-1}) \end{bmatrix} \right) \middle| f \in \mathbb{F}_q[x], \deg f < k \right\} . \lrcorner$$

Theorem 2.21 ([GK16])**.** *Folded Reed-Solomon (*FRS*) codes of order s and rate R are τ-subspace-design codes for $\tau(r) = \frac{sR}{s-r+1}$ for all $1 \leq r \leq s$ and $\tau(r) = 1$ otherwise.*

Corollary 2.22. *For any fixed r, n, there exists a large enough s_0 such that for all $s > s_0$, all rate R order* FRS *codes are r-strong subspace design codes.*

2.3 List-size bounds for list decoding

We now collect state-of-the-art results on the list-size bounds for list-decoding subspace-design, random linear, and random Reed-Solomon codes.

Recent breakthroughs have led to significant improvements in the list-size for list decoding folded RS codes up to capacity, i.e., a decoding radius of $1 - R - \varepsilon$ for rate, improving the previous exponential bounds [KRSW23, Tam24] to polynomial in $1/\varepsilon$, specifically $O(1/\varepsilon^2)$ in [Sri25] and concurrently the optimal $O(1/\varepsilon)$ in [CZ25]. These works exploit the subspace-design property of folded RS codes. We state the optimal result of [CZ25] in the setting of general subspace-design codes.

Theorem 2.23 ([CZ25])**.** *For every τ-subspace design code additive code $\mathcal{C} \subset \Sigma^n$, all $\varepsilon > 0$, and every $y \in \Sigma^n$ we have*

$$|\{c \in \mathcal{C} \mid \Delta(y, c) < 1 - \tau(1/\varepsilon) - \varepsilon\}| \leq (1 - \tau(1/\varepsilon))/\varepsilon \ .$$

We next state results for list-decoding random linear codes and random RS codes up to capacity with optimal list-sizes. These build upon the framework developed in the breakthrough work [BGM22] which showed that random Reed-Solomon codes over exponential-sized fields meet the generalized Singleton bound, which is the precise optimal trade-off for list-decoding with any specific fixed list-size. If one settles for a trade-off slightly off from optimal, then a linear field size suffices, as the following result showed. The same work also proved the following result for random linear codes over sufficiently large constant sized alphabets.

Theorem 2.24 ([AGG+25]). *For all $\varepsilon > 0$, $R \in (0,1)$, prime powers $q \geq 2^{30/\varepsilon^2}$ and sufficiently large n, a random linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of rate R has the following list-decodability property with high probability:*

$$\forall y \in \mathbb{F}_q^n,\ |\{c \in \mathcal{C} \mid \Delta(y,c) < 1 - R - \tfrac{3\varepsilon}{2}\}| \leq \frac{1-R}{\varepsilon}\ .$$

Theorem 2.25 ([AGG+25]). *There exists a constant C, such that all integers, $\varepsilon > 0$, $R \in (0,1)$, $n \geq C/\varepsilon^3$, prime powers $q \geq n \cdot 2^{30/\varepsilon^2}$, a random Reed-Solomon code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of rate R has the following list-decodability property with high probability:*

$$\forall y \in \mathbb{F}_q^n,\ |\{c \in \mathcal{C} \mid \Delta(y,c) < 1 - R - \tfrac{3\varepsilon}{2}\}| \leq \frac{1-R}{\varepsilon}\ .$$

2.4 Local Views

We will now review the terminology and notation concerning local properties, following [LMS25, BCDZ25b].

Definition 2.26 (Linear subspaces). *For a vector space V, we define $\mathcal{L}(V)$ to be the set of all linear subspaces of V.* ⌟

Definition 2.27 (b-Local Profile). *Let $\mathbb{F}_q$ be a finite field of q elements and let n be a positive integer. A tuple $\mathcal{V} = (\mathcal{V}_1, \cdots, \mathcal{V}_n) \in \mathcal{L}(\mathbb{F}_q^b)^n$ is called a b-local profile.* ⌟

Definition 2.28 (Containing a local profile). *A code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is said to contain a local profile $\mathcal{V}$ if there exists an $M \in \mathbb{F}_q^{n \times b}$ such that:*

- *M has pairwise distinct columns and each column is a codeword in $\mathcal{C}$.*
- *Row i of M is in $\mathcal{V}_i$.*

This is to say that there exist pairwise distinct codewords $c_1, \cdots, c_b \in \mathcal{C}$ such that for all $i \in [n]$, we have that $(c_1(i), \cdots, c_b(i)) \in \mathcal{V}_i$. ⌟

The following definition is due to [BCDZ25b].

Definition 2.29 (Additive codes containing local profiles). *An $\mathbb{F}_q$-additive code $\mathcal{C} \subseteq (\mathbb{F}_q^s)^n$ is said to contain a local profile $\mathcal{V} \subseteq \mathcal{L}(\mathbb{F}_q^b)^n$ if the unfolded code $\mathcal{C}' \subseteq \mathbb{F}_q^{sn}$ contains the s-duplicated b-local profile defined as* $\mathcal{V}' := (\underbrace{\mathcal{V}_1, \cdots, \mathcal{V}_1}_{s\ times}, \underbrace{\mathcal{V}_2, \cdots, \mathcal{V}_2}_{s\ times}, \cdots, \underbrace{\mathcal{V}_n, \cdots, \mathcal{V}_n}_{s\ times}) \in \mathcal{L}(\mathbb{F}_q^b)^{sn}$. ⌟

Theorem 2.30 (Random Linear Code Threshold [LMS25]). *For any b-local profile $\mathcal{V}$, there exists a threshold $R_{\mathcal{V}}$ such that if $\mathcal{C} \subseteq \mathbb{F}_q^n$ is a random linear code over $\mathbb{F}_q$ of rate R, then if $R \leq R_{\mathcal{V}} - \varepsilon$ then*

$$\Pr[\mathcal{C}\ contains\ \mathcal{V}] \leq q^{-\varepsilon n + b^2}$$

and if $R \geq R_{\mathcal{V}} + \varepsilon$, then

$$\Pr[\mathcal{C} \text{ contains } \mathcal{V}] \geq 1 - q^{-\varepsilon n + b^2}$$

Theorem 2.31 (Threshold for RLCs from subspace design codes [BCDZ25b])**.** *Let n, q, b, d be fixed constants. If for all large enough s, all d-strong subspace design codes $\mathcal{C} \subseteq (\mathbb{F}_q^s)^n$ do not contain a profile $\mathcal{V}$, then it follows that $R_{\mathcal{V}} \geq R - (b^2+1)/n$.*

Theorem 2.32 (Random Reed-Solomon Threshold [LMS25])**.** *Let $\mathcal{V}$ be a b-local profile. Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a random[2] rate R Reed-Solomon code then for any $\varepsilon > 0$ such that $\varepsilon n > 2b(b+1)$ and $q > Rnb$, we have that if $R \leq R_{\mathcal{V}} - \varepsilon$ then*

$$\Pr[\mathcal{C} \text{ contains } \mathcal{V}] \leq (2^b - 1) \cdot \left(\frac{(4b)^{4b} \cdot Rn}{q} \right)^{\frac{\varepsilon n}{2b}}$$

Definition 2.33. *A b-local property $\mathcal{P}$ is a family of b-local profiles $\mathcal{V}$. We say that a code $\mathcal{C}$ does not have the property $\mathcal{P}$ if it does not contain any $\mathcal{V} \in \mathcal{P}$.* ⌟

Definition 2.34. *A family of b-local properties $\mathcal{P}_{n,q} \subseteq \mathcal{L}(\mathbb{F}_q^b)^n$ is reasonable if for any $\varepsilon > 0$, there exist positive integers C_1, C_2 such that for all $n > C_1$ and $q > C_2$, the number of local profiles in $\mathcal{P}_{n,q}$ is at most $q^{\varepsilon n}$.* ⌟

Theorem 2.35 ([JS25])**.** *Let $R \in \{0, 1\}$, and $\mathcal{P}_{n,q}$ be a reasonable family of b-local properties such that:*

- *$(\mathcal{V}_1, \cdots, \mathcal{V}_n) \in \mathcal{P}_{n,q}$ and $\pi : [n] \to [n]$ is a permutation, we must also have $(\mathcal{V}_{\pi(1)}, \cdots, \mathcal{V}_{\pi(n)}) \in \mathcal{P}_{n,q}$.*
- *$R_{\mathcal{V}} \geq R$ for all large enough n, q and $\mathcal{V} \in \mathcal{P}_{n,q}$.*

Then for any $\varepsilon > 0$, there exist constants C_1, C_2 such that for any prime power $q > C_1$ and $n > C_2$, there exists an explicit code with alphabet $\mathbb{F}_q$, block length n and rate $R - \varepsilon$ which does not contain $\mathcal{P}_{n,q}$.

2.5 Vandermonde Matrix and Inversion definition

For any $\alpha_1, \cdots, \alpha_k$ in a field $\mathbb{F}$, the $k \times k$ Vandermonde Matrix $V(\alpha_1, \cdots, \alpha_k)$ is given with $V_{i,j} = \alpha_i^{j-1}$.

For all pairwise distinct $\alpha_1, \cdots, \alpha_k$, we have that $V(\alpha_1, \cdots, \alpha_k)$ is invertible.

Definition 2.36. *Let $M_{\alpha_1, \cdots, \alpha_k}(\beta)$ be the (row) vector such that*

$$M_{\alpha_1, \cdots, \alpha_k}(\beta) V(\alpha_1, \cdots, \alpha_k) = [1 \ \beta \ \cdots \ \beta^{k-1}] \ .$$ ⌟

Lemma 2.37. *Let V be a linear space over $\mathbb{F}$. Let $\alpha_1, \cdots, \alpha_k \in \mathbb{F}$ be distinct elements and let $v_{\alpha_1}, \cdots, v_{\alpha_k}$ be elements of V.*

Then, there exist vectors $w_0, \cdots, w_{k-1} \in V$ such that

$$v_{\alpha_i} = \sum_{j=0}^{k-1} w_j \alpha_i^j \ .$$

[2] Sample the evaluation points from the field uniformly and independently

Additionally, for any $\beta \in \mathbb{F}$, we have that

$$M_{\alpha_1,\cdots,\alpha_k}(\beta)[v_{\alpha_1}\ v_{\alpha_2}\ \cdots\ v_{\alpha_k}]^T = \sum_{j=0}^{k-1} w_j \beta^j$$

Proof. We consider $\alpha_1, \cdots, \alpha_k$ to be fixed and will thus drop them from subscripts for this proof.

For any $j \in \{0, \cdots, k-1\}$, let N_j be a row vector such that $N_j V(\alpha_1, \cdots, \alpha_k) = e_j$ where $e_j \in \mathbb{F}_q^{1\times k}$ with the $j+1$th entry as 1 and all other entries equal to 0.

Observe that for any $\beta \in \mathbb{F}_q$, we have that $(\sum_{j=0}^{k-1} \beta^j \cdot N_j)V(\alpha_1, \cdots, \alpha_k) = [1\ \beta\ \cdots\ \beta^{k-1}]$. Thus, we have that $M(\beta) = \sum_{j=0}^{k-1} \beta^j \cdot N_j$.

In particular, for all $i \in [k]$, we have that $M(\alpha_i)V(\alpha_1, \cdots, \alpha_k) = [1\ \alpha_i\ \cdots\ \alpha_i^{k-1}]$. In particular, we must have that $M(\alpha_i) = e_{i-1}$.

In particular, this means that if let $w_j := N_j[v_{\alpha_1}, \cdots, v_{\alpha_k}]^T$ for each $j \in \{0, 1, \cdots, k-1\}$ then we have

$$M(\beta)[v_{\alpha_1}, \cdots, v_{\alpha_k}]^T = \sum_{j=0}^{k-1} \beta^j w_j\ .$$

Additionally since $M(\alpha_i) = e_{i-1}$, we have

$$v_{\alpha_i} = M(\alpha_i)[v_{\alpha_1}, \cdots, v_{\alpha_k}]^T = \sum_{j=0}^{k-1} \alpha_i^j w_j\ . \qquad \square$$

3 Proximity gaps and correlated agreement from curve decodability

In this section, we define the concept of curve decodability which is our unified, novel approach to establish proximity gaps and (mutual) correlated agreement.

3.1 Curve-decodability

We begin with defining curve-decodability. Informally, it says that if there is a curve of points many of which have a close-by codeword, then many of those codewords themselves lie on a curve with the same parametrization.

Definition 3.1 (Curve-decodability)**.** *An additive code $\mathcal{C} \subseteq \Sigma^n$ is (ℓ, δ, a, b) curve-decodable if for every $u_0, u_1, \cdots, u_\ell \in \Sigma^n$, all functions $f : \mathbb{F}_q \to \mathcal{C}$, whenever the set*

$$A = \Big\{\alpha \in \mathbb{F} \mid \Delta\big(\sum_{j=0}^{\ell} u_j \alpha^j, f(\alpha)\big) \le \delta\Big\}$$

has at least a elements, there exist $c_0, c_1, \cdots, c_\ell \in \mathcal{C}$ such that

$$\Big|\Big\{\alpha \in A \mid f(\alpha) = \sum_{j=0}^{\ell} c_j \alpha^j\Big\}\Big| \ge b\ .$$

When $\ell = 1$, we use the phrase line decodable. ⌟

3.2 Curve Decodability to Mutual Correlated Agreement

We now prove that if there are many points on a curve close to codewords that lie on a curve, then in fact *all* points of the two curves are close-by, within a slightly larger proximity.

Lemma 3.2. *For an alphabet Σ that is an $\mathbb{F}_q$-linear space, for all $u_0, u_1, \cdots, u_\ell, c_0, c_1, \cdots, c_\ell \in \Sigma^n$ and all $t \in \mathbb{N}$, the following implication holds:*

$$\left|\left\{\alpha \mid \Delta\left(\sum_{j=0}^{\ell} u_j\alpha^j, \sum_{j=0}^{\ell} c_j\alpha^j\right) \leq \delta\right\}\right| \geq t \implies \Delta\left(\sum_{j=0}^{\ell} u_j\alpha^j, \sum_{j=0}^{\ell} c_j\alpha^j\right) \leq \delta(1+\tfrac{\ell}{t-\ell}) \quad \forall \alpha \in \mathbb{F}_q$$

Additionally, $|\{i \mid \bigvee_{j=0}^{\ell}(u_{j,i} \neq c_{j,i})\}| \leq (1+\frac{\ell}{t-\ell})\delta n$

Proof. Define $T = \{i \mid \bigvee_{j=0}^{\ell} u_{j,i} \neq c_{j,i}\}$. We will prove $|T| \leq \delta n \cdot (1+\frac{\ell}{t-\ell})$ which implies the lemma.

For each $\alpha \in \mathbb{F}_q$, define $T_\alpha = \left\{i \in T \mid \sum\limits_{j=0}^{\ell} (c_{j,i} - u_{j,i}) \cdot \alpha^j = 0\right\}$. Note that by the degree bound, each $i \in T$, can belong to at most ℓ different T_α's. Now, for each $\alpha \in \mathbb{F}_q$, we have

$$\Delta\left(\sum_{j=0}^{\ell} u_j\alpha^j, \sum_{j=0}^{\ell} c_j\alpha^j\right) = \Delta\left(\sum_{j=0}^{\ell} (c_j - u_j)\cdot\alpha^j, 0\right) = \frac{|T| - |T_\alpha|}{n}$$

Defining $A = \left\{\alpha \in \mathbb{F}_q \mid \Delta\left(\sum_{j=0}^{\ell} u_j\alpha^j, \sum_{j=0}^{\ell} c_j\alpha^j\right) \leq \delta\right\}$, by hypothesis $|A| \geq t$. Observe that since each i belongs to at most ℓ different sets T_α, there must exists an $\alpha_0 \in A$, such that $|T_{\alpha_0}| \leq \ell \cdot |T|/t$. Therefore,

$$\Delta\left(\sum_{j=0}^{\ell} u_j{\alpha_0}^j, \sum_{j=0}^{\ell} c_j{\alpha_0}^j\right) \leq \delta \implies |T|(1-\tfrac{\ell}{t}) \leq \delta n \implies |T| \leq \delta n \cdot (1+\tfrac{\ell}{t-\ell})\ . \qquad \square$$

Theorem 3.3 (Mutual Correlated agreement from Curve-Decodability)**.** *Let $\mathcal{C} \subseteq \Sigma^n$ be an $\mathbb{F}_q$-additive code. Let $t, \ell, a \in \mathbb{N}$ and $\delta \in (0,1)$. Suppose that $\mathcal{C}$ is (ℓ, δ, a, t) curve-decodable. Then $\mathcal{C}$ has $(\ell, \delta, \frac{a}{|\mathbb{F}_q|}, \ell/t)$ mutual correlated agreement.*

Proof. Let $u_0, \cdots, u_\ell \in \Sigma^n$, $f : \mathbb{F}_q \mapsto \mathcal{C}$, $u_\alpha = \sum_{j=0}^{\ell} u_j\alpha^j$ and

$$A := \{\alpha \in \mathbb{F}_q \mid \Delta(u_\alpha, f(\alpha)) \leq \delta\}\ ,$$

with $|A| \geq a$. By the hypothesis about line decodability of $\mathcal{C}$, we know there exist $c_0, c_1, \cdots, c_\ell \in \mathcal{C}$ such that

$$\left|\{\alpha \in A \ \mid \ f(\alpha) = \sum_{j=0}^{\ell} c_j\alpha^j\}\right| \geq t.$$

In particular, we surely have the existence of $\beta \in \mathbb{F}$ with $f(\beta) = \sum_{j=0}^{\ell} c_j\beta^j$. Applying Lemma 3.2, we can conclude

$$|\{i \mid \bigvee_{j=0}^{\ell} u_{j,i} \neq c_{j,i}\}| \leq (1+\tfrac{\ell}{t-\ell})\delta n = \tfrac{\delta n}{1-\ell/t}\ ,$$

which establishes the claimed mutual correlated agreement. $\square$

The above argument is extremely strong but it remains to understand what kind of a would need to be chosen if t has be made as large as $n\ell$ in the above theorem to get perfect agreement. It's conceivable that we might need to pick a which is extremely large in these situations.

We thus ask, can we exploit any other property of the code $\mathcal{C}$ to get mutual correlated agreement even from curve decodability for small t?

We show that list decodability of the code is one such property. We now improve the parameters of the mutual correlated agreement, specifically related to the increase in proximity, if we further assume that $\mathcal{C}$ has good list-decoding properties.

Theorem 3.4 (Mutual Correlated agreement from Curve and List-Decodability). *Let $\mathcal{C} \subseteq \Sigma^n$ be an $\mathbb{F}_q$-linear code. Let $\ell, t_1, t_2 \in \mathbb{N}$ and $\delta \in (0,1)$. Suppose the following two conditions hold:*

1. *$\mathcal{C}$ is $(\delta(1+\frac{\ell}{t_1-\ell}), L)$ list-decodable.*
2. *$\mathcal{C}$ is (ℓ, δ, a, t_1) curve-decodable.*

Then, $\mathcal{C}$ has $(\ell, \delta, \frac{(t_2-1)\cdot L+a}{q}, \frac{\ell}{t_2})$ mutual correlated agreement as long as $q > (L+1)^2 \cdot \ell$.

Proof. Let $\delta' \leq \delta$, $u_0, \cdots, u_\ell \in \Sigma^n$, $f_0 : \mathbb{F}_q \mapsto \mathcal{C}$, $u_\alpha = \sum_{j=0}^{\ell} u_j \alpha^j$. Suppose $A_0 = \{\alpha \in \mathbb{F}_q \mid \Delta(u_\alpha, f_0(\alpha)) \leq \delta'\}$ satisfies $|A_0| \geq (t_2-1)\cdot L + a$.

We will repeatedly apply the curve-decodability with different choices of f and A, starting with $f = f_0$ and $A = A_0$.

By the assumed curve-decodability of the code $\mathcal{C}$, we know there are codewords $c_0^{(0)}, c_1^{(0)}, \cdots, c_\ell^{(0)}$ such that if

$$B_0 = \left\{\alpha \in A_0 \mid f_0(\alpha) = \sum_{j=0}^{\ell} c_j^{(0)} \alpha^j \right\}$$

then $|B_0| \geq t_1$.

Observe that if $|B_0| \geq t_2$ then Lemma 3.2 implies the theorem statement since for all $\alpha \in B_0$, we already have that the curve passes through $f_0(\alpha)$. So assume that $|B_0| \leq t_2 - 1$. Regardless, we have for all $\alpha \in \mathbb{F}_q$,

$$\Delta(\sum_{j=0}^{\ell} u_j \alpha^j, \sum_{j=0}^{\ell} c_j \alpha^j) \leq \delta \cdot (1 + \tfrac{\ell}{t_1-\ell}).$$

Now we will update $A = A_1 := A \setminus A_0$, and f to f_1 which equals f_0 for $\alpha \in A_0 \setminus B_0$, and for $\alpha \in B_0$ it will satisfy $\Delta(f_1(\alpha), u_\alpha) > \delta$ (so the curve points associated to $\alpha \in B_0$ are no longer in proximity to their associated codewords). We can now repeat this procedure applying the curve-decodability to obtain a sequence of A_i, f_i, resulting in sets B_i and codewords $c_0^{(i)}, c_1^{(i)}, \cdots, c_\ell^{(i)}$, for $i = 0, 1, \ldots$, for up to L iterations.

If we ever find a B_i with $|B_i| \geq t_2$, we are once again done by applying Lemma 3.2. Otherwise we will have, for $i = 0, 1, \ldots, L$, sets $B_i \subset \mathbb{F}_q$ with each $|B_i| \leq t_2 - 1$ and codewords $c_0^{(i)}, c_1^{(i)}, \cdots, c_\ell^{(i)}$ such for each $i \in \{0, 1, \cdots L\}$ and every $\alpha \in \mathbb{F}_q$,

$$\Delta\left(\sum_{j=0}^{\ell} u_j \alpha^j, \sum_{j=0}^{\ell} c_j^{(i)} \alpha^j\right) \leq \delta(1 + \tfrac{\ell}{t_1-\ell}) \ . \tag{3.5}$$

Finally, we argue that this contradicts the assumed list-decodability of $\mathcal{C}$. We have $(L+1)$ distinct degree-ℓ curves of codewords, $\sum_{j=0}^{\ell} c_j^{(i)} \alpha^j$, and each pair can intersect on at most ℓ values of $\alpha \in \mathbb{F}_q$. When $q > (L+1)^2 \ell$, there must exist some $\zeta \in \mathbb{F}_q$ such that all the codewords $\sum_{j=0}^{\ell} c_j^{(i)} \zeta^j$ are distinct. But then by (3.5) the center $\sum_{j=0}^{\ell} u_j \zeta^j$ has more than L codewords within relative distance $\delta(1 + \frac{\ell}{t_1 - \ell})$ from it, a contradiction to the list-decodability hypothesis. □

Remark 3.6. *This is the first instance where we have used the strength of the curve decodability definition. It holds for any function f and thus we can remove curves and curate for whatever properties we like. Not doing the above would have resulted in worse error parameters throughout and worsened the linear factor dependencies with n.*

We thus ask whether this trick of combining list decoding with the ability to pick f arbitrarily can be used in other proximity gaps results to improve the error parameters. In particular, the approach strongly suggests that the combination of adding a slack along with mutual correlated agreement is extremely useful in improving our understanding of the error bounds even when we may not have optimal curve decodability bounds. ⌟

4 Proximity Gaps for Subspace-Design Codes

4.1 Dimension Bounds for Subspace-Design Codes

Throughout this section, all the codes we use are going to be $\mathbb{F}$-additive. We will consistently use $\mathbb{F}$ (or $\mathbb{F}_q$ when we want to make the field size q explicit) as the underlying field over which our codes are additive.

We begin with a dimension bound for the solutions to a "list-recovery" problem for subspace-design codes. The key difference in this version is that we allow each of the input lists L_i to be a low-dimensional *subspace* as opposed to a small *subset* of the alphabet Σ.

This observation that output lists are contained in low dimensional spaces for the list recovery problem on subspace design codes even for *low dimensional inputs* and not just *small inputs* is crucial to us being able to leverage techniques from the list decoding literature for proximity gaps.

Lemma 4.1 (Dimension bounds for decoding of subspace design codes)**.** *Let $\mathcal{C} \subseteq \Sigma^n$ be a τ-subspace-design additive code. Then, for every input $L_1, \cdots, L_n \subseteq \Sigma$ with each L_i being a linear space of dimension ℓ_i, we have that output list*

$$\left\{ c \in \mathcal{C} \;\mid \Delta(c, L_1 \times L_2 \times \cdots \times L_n) < 1 - \tau(r) - \frac{\sum\limits_{i=1}^{n} \ell_i}{nr} \right\}$$

is contained within a linear space of dimension less than r.

Proof. If $\dim \mathcal{C} < r$, then there is nothing to prove. Else, consider any r linearly independent codewords $c^{(1)}, \cdots, c^{(r)} \in \mathcal{C}$. For each $j \in [n]$, defined $\mathcal{B}_j = \{c^{(i)} \mid i \in [r], c_j^{(i)} \in L_j\}$.

Let $\mathcal{A} = \mathsf{span}(c^{(1)}, \cdots, c^{(r)})$, and for each $j \in [n]$, let $\mathcal{A}_j = \{c \in \mathcal{A} \mid c_j = 0\}$. We have $\ell_j + \dim \mathcal{A}_j \geq |\mathcal{B}_j|$. Let $\ell n = \sum_{j=1}^{n} \ell_j$.

Combining these with the subspace design property, we can conclude that

$$n \cdot r \cdot \tau(r) + \sum_{j=1}^{n} \ell_j \geq \sum_{j=1}^{n} (\dim \mathcal{A}_j + \ell_j) \geq \sum_{j=1}^{n} |\mathcal{B}_j|.$$

The first inequality is again the subspace design property, and the last inequality follows from the proximity of each $c^{(i)}$ to the lists L_j.

Thus, there exists an $i \in [r]$ such that $\Delta(c, L_1 \times \cdots \times L_n) \geq 1 - \tau(r) - \frac{\ell}{r}$ as desired. □

Remark 4.2. *We note that the above proof natively establishes the stronger* average-radius list-recoverability *property, namely that the average-radius of any r linearly independent codewords to any product of ℓ-dimensional linear spaces is at least $1 - \frac{\ell}{r} - \tau(r)$.* ⌟

Remark 4.3. *There has been exciting progress on the list-recovery problem via a striking connection to discrete Brascamp-Lieb inequalities in [BCDZ25a] to get near optimal list sizes for subspace-design codes. They then use the framework of [LMS25, BCDZ25b] to carry these optimal list bounds even to random linear and random Reed-Solomon codes.* ⌟

4.2 New pruning methods for subspaces

The pruning algorithm we present is a variation of the algorithm in [AHS25]. Theirs was an algorithm for the task of list-decoding and we alter their algorithm to work more generally. Their algorithm for list-decoding took as input a vector y and a low-dimensional space $\mathcal{H}$ of codewords and proceeded as follows towards finding all close-by codewords to y in $\mathcal{H}$.

Algorithm 1: $\mathsf{PRUNE}_{\mathcal{D}}(\mathcal{H}, y)$: Pruning a linear space

```
Input: An $\mathbb{F}_q$ affine space $\mathcal{H}$, a received word $y \in \mathbb{F}_q^n$. Let $\mathcal{H}_i = \{h \in \mathcal{H} \mid h_i = y_i\}$.
If $\dim \mathcal{H} = 0$, output the only element in $\mathcal{H}$.
Else, sample $i \sim \mathcal{D}(\mathcal{H}, y)$ where $\mathcal{D}(\mathcal{H}, y)$ is a distribution on $[n]$ and return $\mathsf{PRUNE}_{\mathcal{D}}(\mathcal{H}_i, y)$.
```

We make a key change in the above algorithms to work only on linear spaces $\mathcal{H}$ and making the algorithm oblivious of the codeword y and just to output a set $S \subseteq [n]$.

Algorithm 2: $\mathsf{PRUNE}_{\mathcal{D}}(\mathcal{H})$: Pruning a linear space

```
Input: An $\mathbb{F}_q$ affine space $\mathcal{H}$. Let $\mathcal{H}_i = \{h \in \mathcal{H} \mid h_i = 0\}$.
If $\dim \mathcal{H} = 0$, output $\emptyset$.
Else, sample $i \sim \mathcal{D}(\mathcal{H})$ where $\mathcal{D}(\mathcal{H})$ is a distribution on $[n]$ and return $\{i\} \cup \mathsf{PRUNE}_{\mathcal{D}}(\mathcal{H}_i)$.
```

Pruning algorithms for list decoding were also earlier considered in [KRSW23, Tam24]. These work for more general codes as well and not only subspace-design codes. Their algorithm has the same structure as above, with uniform distribution used to sample the coordinates i. Since we only use pruning algorithms to prune subspace-design codes, we do not discuss the [KRSW23, Tam24] pruning argument here.

Definition 4.4 (Pinned Subspaces)**.** *For any linear space $\mathcal{H} \subseteq \Sigma^n$, and any set $S \subseteq [n]$, let $\mathcal{H}_S = \{h \in \mathcal{H} \mid h_i = 0 \ \ \forall i \in S\}$. For singletons, we simply use $\mathcal{H}_i$ to denote $\mathcal{H}_{\{i\}}$.* ⌟

Theorem 4.5 ([AHS25] based pruning)**.** *Let $\mathcal{C}$ be a τ-subspace design additive code. Let $\mathcal{H} \subseteq C$ be a linear space of dimension r, and let $y \in \Sigma^n$. Fix any $c \in \mathcal{C} \cap \mathcal{H}$ satisfying $\Delta(c, y) \leq 1 - \tau(r) - \varepsilon$. Then the set S output by $\mathsf{PRUNE}_{\mathcal{D}_\varepsilon}$ satisfies the following guarantee:*

$$\Pr_S[c_i = y_i \ \forall i \in S \textit{ and } \dim \mathcal{H}_S = 0] \geq \frac{\varepsilon}{r + \varepsilon} .$$

Here $\mathcal{D}_\varepsilon(\mathcal{H})$ samples coordinate i with probability $p_i = \frac{w_i}{\sum_{j=1}^n w_j}$ where the weights w_i are defined by

$$w_i = \begin{cases} 0 & \mathcal{H}_i = \mathcal{H} \\ \dim \mathcal{H}_i + \varepsilon & \textit{otherwise} \end{cases} .$$

Proof. We provide a different proof than [AHS25] which works in the same spirit.

Let $f_\varepsilon(\mathcal{H}) = \frac{1}{\dim \mathcal{H} + \varepsilon}$ for any linear space $\mathcal{H}$. Define the indicator variable $X_S = 1$ if $c_i = y_i \ \forall i \in S$ and 0 otherwise.

Claim 4.6. *Let $\mathcal{H}$ be any linear space of dimension at most r, $S \subseteq [n]$ be any set such that $\mathcal{H}_S \neq \{0\}$ then*

$$\mathbb{E}_{i \sim \mathcal{D}_\varepsilon(\mathcal{H}_S)}[X_{S \cup \{i\}} \cdot f_\varepsilon(\mathcal{H}_{S \cup \{i\}})] \geq X_S \cdot f_\varepsilon(\mathcal{H}_S)$$

Proof. Let $T = \{i \in [n] \mid \mathcal{H}_{S \cup \{i\}} = \mathcal{H}_S\}$. Let $w = \dim \mathcal{H}_S + \varepsilon \geq 1$.

Observe that if $X_S = 0$ then there is nothing to prove. Else we have

$$\begin{aligned} \mathbb{E}_{i \sim \mathcal{D}_\varepsilon(\mathcal{H}_S)}[X_{S \cup \{i\}} \cdot f_\varepsilon(\mathcal{H}_{S \cup \{i\}})] &= \sum_{i \notin T:\ c_i = y_i} \frac{w_i}{\sum\limits_{j=1}^{n} w_j} \cdot \frac{1}{w_i} \\ &\geq \frac{|\{i \mid c_i = y_i\}| - |T|}{w \cdot \tau(r) \cdot n + \varepsilon n - |T| w} \\ &\geq \frac{\tau(r) + \varepsilon - |T|/n}{w(\tau(r) + \varepsilon - |T|/n)} = \frac{1}{w} , \end{aligned}$$

where the second step uses the subspace-design property Definition 2.17. □

Now, observe that initially $f_\varepsilon(\mathcal{H}) \cdot X_\emptyset = \frac{1}{r+\varepsilon}$. Thus, we get that if S is output by the algorithm then

$$\mathbb{E}_{S \sim \mathsf{PRUNE}_{\mathcal{D}_\varepsilon}(\mathcal{H})}[X_S \cdot f_\varepsilon(\mathcal{H}_S)] \geq \frac{1}{r + \varepsilon} \implies \mathbb{E}[X_S] \geq \frac{\varepsilon}{r + \varepsilon}$$

since we have that $f_\varepsilon(\mathcal{H}_S) = \frac{1}{\varepsilon}$ for the algorithm's output always. □

4.3 Curve Decoding for subspace design codes

The $\mathsf{PRUNE}_{\mathcal{D}_\varepsilon}$ algorithm described above being oblivious to the codeword is extremely powerful. In particular, this means that if our space of close-by codewords all belong to the same low dimensional space, then we could run the algorithm using shared randomness. We know that this is indeed true due to Lemma 4.1.

This observation now gets us a single set on which many of the codewords close to the curve will agree with their respective close-by elements. This is quite helpful as complete agreement implies that the remaining close-by codewords must have been on a single curve as well!

The following theorem is thus one of our main technical contributions, and is the central ingredient that enables the proximity gaps results for subspace-design codes.

Theorem 4.7. *For arbitrary positive integers r, ℓ, a and any $\varepsilon \geq \frac{\ell+1}{r}$, every τ-subspace design code $\mathcal{C} \subseteq \Sigma^n$ is $(\ell, 1 - \tau(r) - \varepsilon, a, \frac{\varepsilon}{r+\varepsilon} \cdot a)$ curve-decodable.*

Proof. Let $\delta \leq 1 - \tau(r) - \varepsilon$, $u_0, \cdots, u_\ell \in \Sigma^n$, and $f : \mathbb{F}_q \to \mathcal{C}$. Define $u_\alpha = \sum_{j=0}^{\ell} u_j \alpha^j$. Let $U = \mathsf{span}(u_0, \cdots, u_\ell)$. Define

$$H = \{c \in \mathcal{C} \mid \Delta(c, U) \leq 1 - \tau(r) - \varepsilon\} \ .$$

By Lemma 4.1, H is contained in a linear space of dimension at most r. Denote this space by $\mathcal{H}$.

Let $A = \{\alpha \in \mathbb{F} \mid \Delta(u_\alpha, f(\alpha)) \leq \delta\}$. If $|A| < a$ then there is nothing to prove. So, assume that $|A| \geq a$. By definition, for each $\alpha \in A$, we know that $f(\alpha) \in H \subseteq \mathcal{H}$.

Now, by Theorem 4.5,

$$\begin{aligned}\mathbb{E}_{S \sim \mathsf{PRUNE}_{\mathcal{D}_\varepsilon}(\mathcal{H})}[|\{\alpha \in A \mid u_{\alpha,i} = f(\alpha)_i \quad \forall i \in S\}] &= \sum_{\alpha \in A} \Pr_{S \sim \mathsf{PRUNE}_{\mathcal{D}_\varepsilon}(\mathcal{H})}[u_{\alpha,i} = f(\alpha)_i \ \forall i \in S] \\ &\geq |A| \cdot \frac{\varepsilon}{r + \varepsilon} \ \geq a \cdot \frac{\varepsilon}{r + \varepsilon} \ .\end{aligned}$$

Therefore, there exists a subset $S \subseteq [n]$ such that $\mathcal{H}_S = \{0\}$ and

$$B := \{\alpha \in A \mid f(\alpha)_i = u_{\alpha,i} \ \forall i \in S\}$$

satisfies $b := |B| \geq a \cdot \frac{\varepsilon}{r+\varepsilon}$.

Now, for each $\alpha \in B$, denote $c^\alpha = f(\alpha)$. Our goal is now to show that these c^α all lie on a degree ℓ curve with the correct parametrization.

For convenience, let $B = \{\alpha_1, \cdots, \alpha_b\}$. Let $d = \min(b-1, \ell)$. Now, we know that the $(d+1) \times (d+1)$ Vandermonde matrix $V(\alpha_1, \cdots, \alpha_{d+1})$ is invertible. Thus, for any $\beta \in \mathbb{F}_q$, there is a row vector $M_{\alpha_1, \cdots, \alpha_{d+1}}(\beta)$ such that $M_{\alpha_1, \cdots, \alpha_{d+1}}(\beta) V(\alpha_1, \cdots, \alpha_{d+1}) = [1 \ \beta \ \cdots \ \beta^d]$. (Definition 2.36)

Now, observe that for any $\beta \in B$, we have that

$$M_{\alpha_1, \cdots, \alpha_{d+1}}(\beta)[c^{\alpha_1}_{|S} \ \cdots \ c^{\alpha_{d+1}}_{|S}]^T = M_{\alpha_1, \cdots, \alpha_{d+1}}(\beta)[u_{\alpha_1|S} \ \cdots \ u_{\alpha_{d+1}|S}]^T = u_{\beta|S} = c^{\beta}_{|S}$$

In particular, we get that

$$(M_{\alpha_1, \cdots, \alpha_{d+1}}(\beta)[c^{\alpha_1} \ \cdots \ c^{\alpha_{d+1}}]^T - c^\beta)_{|S} = 0 \ .$$

We know that $c_{\alpha_1}, \cdots, c_{\alpha_{d+1}}, c_\beta \in \mathcal{H}$. Thus, so is the above described linear combination. By hypothesis, $\mathcal{H}_S = \{0\}$. So, we must have

$$M_{\alpha_1, \cdots, \alpha_{d+1}}(\beta)[c^{\alpha_1} \ \cdots \ c^{\alpha_{d+1}}]^T = c^\beta \ .$$

Thus, all c^α lie on a single degree d curve with the right parametrization due to Lemma 2.37.

Since this holds for each $\beta \in B$, we conclude that there exist $c_0, c_1, \cdots, c_\ell \in \mathcal{C}$ such that $c^\beta = \sum_{j=0}^{\ell} c_j \beta^j$ for all $\beta \in B$. This is exactly the result we desired. □

4.4 Main results for Subspace Design Codes

Theorem 4.8 (Mutual Correlated Agreement for Subspace Design codes)**.** *For any integers $t, m, \ell \in \mathbb{N}$, any τ-subspace design code $\mathcal{C} \subseteq \Sigma^n$ has mutual correlated agreement*

$$\left(\ell, 1 - \tau(t\ell + t) - \tfrac{3}{2t}, t \cdot \ell \cdot \left(\frac{m(1 - \tau(t\ell + t)) + 2t^2(\ell + 1)}{|\mathbb{F}|}\right), \frac{1}{m}\right) .$$

Proof. We set parameters as follows:

$$r = (\ell+1)t,\ t_1 = 2\ell\cdot t,\ t_2 = m\cdot\ell \text{ and } \delta = (1-\tau(r)-1/t)(1-\ell/t_1)\ .$$

By Theorem 2.23, $\mathcal{C}$ is $(\delta\cdot\frac{t_1}{t_1-\ell}, t(1-\tau(r)))$ list decodable. Note that for our choice of parameters $\delta \le 1-\tau(r)-\frac{3}{2t}$, and $\frac{3}{2t} \ge \frac{\ell+1}{r}$. So by Theorem 4.7, $\mathcal{C}$ is $(\ell,\delta,t_1(2rt/3+1),t_1)$ line-decodable, and thus also $(\ell,\delta,2\ell rt^2,t_1)$ line-decodable.

Combining these results using Theorem 3.4, we get that $\mathcal{C}$ has

$$\left(\ell,\delta,\frac{m\ell t(1-\tau(r))+2\ell rt^2}{|\mathbb{F}|},\frac{1}{m}\right)$$

mutual correlated agreement. Recalling that $\delta \le 1-\tau(r)-\frac{3}{2t}$ and plugging back the value of $r=(\ell+1)t$ gives us the claimed bound. □

Corollary 4.9 (Perfect Mutual Correlated Agreement for Subspace Design codes)**.** *For any integers $t,\ell\in\mathbb{N}$ and $\varepsilon\in\mathbb{R}_{\ge 0}$, then any $\mathbb{F}$-additive τ-subspace design code $\mathcal{C}$ has mutual correlated agreement $\left(\ell, 1-\tau(t\cdot\ell+\ell)-3/2t, t\cdot\ell\cdot\frac{n+2t^2(\ell+1)}{|\mathbb{F}|}, 0\right)$ mutual correlated agreement*

Proof. Follows from Theorem 4.8 with the choice $m=n$. □

Corollary 4.10 (Mutual Correlated agreement for Folded Reed-Solomon Codes)**.** *Let t,m,ℓ be any integers, then for $s>4t^2\ell$, we have that every s-Folded Reed Solomon Code (Definition 2.20) over $\mathbb{F}$ has*

- *(MCA) $(\ell, 1-R-2/t, \frac{mt\ell+3t^3\ell^2}{|\mathbb{F}|}, 1/m)$ mutual correlated agreement,*
- *(Perfect MCA) $(\ell, 1-R-2/t, \frac{nt\ell+3t^3\ell^2}{|\mathbb{F}|}, 0)$ mutual correlated agreement, and*
- *(Perfect line MCA) $(1-R-2/t, \frac{nt+O(t^3)}{|\mathbb{F}|}, 0)$ line mutual correlated agreement (this only requires $s>4t^2$).*

Proof. Observe that by Theorem 2.21, $\tau(t\cdot\ell+t) \le (1+\frac{r-1}{s-r+1})\cdot R \le (1+\frac{1}{2t})\cdot R \le R+\frac{1}{2t}$. Setting this in Theorem 4.8, we get that any subspace design code has mutual correlated agreement $(\ell, 1-R-2/t, \frac{mt\cdot\ell+3t^3\ell^2}{|\mathbb{F}|}, 1/m)$ mutual correlated agreement as desired.

For the perfect mutual correlated agreement, we just set $m=n$. And for the perfect line mutual correlated agreement, we further set $\ell=1$. □

We next state corollaries for proximity gaps and correlated agreement for affine spaces.

Corollary 4.11 (Affine space proximity gaps for FRS codes)**.** *Let t be an integer, then for $s>4t^2$, we have that every s-Folded Reed Solomon code over $\mathbb{F}$ has*

- $(1-R-2/t, \frac{mt+O(t^3)}{|\mathbb{F}|-1}, 1/m)$ *affine space proximity gap.*
- $(1-R-2/t, \frac{nt+O(t^3)}{|\mathbb{F}|-1}, 0)$ *perfect affine space correlated agreement.*

Proof. The first item follows from Lemma 2.13 and Corollary 4.10 with $\ell=1$. The second item follows from Lemma 2.15 and Corollary 4.10. □

Remark 4.12 (Univariate Multiplicity Codes). *The above results also extend to Univariate Multiplicity Codes as they also possess the subspace design property due to [GK16].* ⌟

Remark 4.13 (MCA for affine spaces). *We note that even though we do not have a blackbox conversion for line or curve mutual correlated agreement to affine space mutual correlated agreement, the above established techniques work perfectly well even for affine spaces as well for subspace design codes. We haven't written this out as it is not the focus of the work but the above techniques give affine space mutual correlated agreement with parameters* $(1-\tau(\ell(t+1))-2/t, \frac{mt+O(t^3\ell^2)}{|\mathbb{F}|}, 1/m)$ *for affine spaces of dimension at most* ℓ*. This also leads to affine space mutual correlated agreement for Folded Reed-Solomon codes.* ⌟

Remark 4.14 (Efficient decoding and knowledge extraction). *For explicit subspace design codes like Folded Reed-Solomon and Univariate Multiplicity Codes, we have algorithms such as [GW13b, GHKS24, GHKS25] which efficiently output the space* $\mathcal{H}$ *as well. Thus, once we have* $\mathcal{H}$*, the above way of pruning through [AHS25] style methods is also algorithmic. In particular, this means that there are* $\widetilde{O}(\mathrm{poly}(n, \ell/\varepsilon))$ *algorithms which can indeed find the curves for curve-decodability as long as there is some slack. Thus, for the mutual correlated agreement, by slacking the probability of ending up close a little bit, it is also possible to efficiently find the desired codewords* $c_0, c_1, \cdots, c_\ell$*.*

The key application of correlated agreement theorems is in designing SNARKs. SNARK designers tend to make the assumption that if there exists a cheating dishonest prover succeeeds with non-negligible property, then by rewinding and rerunning the prover a few times, a false witness can be extracted i.e. it can be shown that the prover must have had knowledge of a false witness. Having an algorithm as above to efficiently extract the close by codewords from lying close to the codespace serves this purpose and helps prove knowledge soundness of such constructions. ⌟

5 Proximity Gaps for Random Reed-Solomon and Linear Codes

Since we already know that all local profiles that are avoided by subspace-design codes are also avoided by random linear and random Reed-Solomon codes, it is enticing to just cast the definition of curve-decodability as a local profile. Unfortunately, as far as we can tell, it does not fit the current framework of local profiles. In particular, current framework of local profiles Definition 2.28 require that all codewords are pairwise distinct. A direct casting as local profiles for us would require the property of containment as "no $\ell + 2$" are all equal i.e., we seem to need to use a more general definition of local properties.

In this work, we avoid reworking the local properties literature to fit this model but consider a different property that we call V decoding which can indeed be cast as a local property. We show that any code which has V decoding, with some loss of parameters, also has the property of curve decoding. This conversion is quite lossy for large ℓ and we view this as good motivation to rework the local properties framework in an appropriately more general language. But for simplicity, and because it is not related to our main focus of presenting a novel route to proximity gaps via curve decodability, we have not attempted such a generalization in this version.

5.1 $\vee$ decoding

Definition 5.1. *An additive code $\mathcal{C} \subseteq \Sigma^n$ is (ℓ, δ, a) $\vee$-decodable if for all $u_0, u_1, \cdots, u_l \in \Sigma^n$ and all functions $f : \mathbb{F}_q \to \mathcal{C}$, if the set*

$$A = \left\{ \alpha \in \mathbb{F} \mid \Delta\Big(\sum_{j=0}^{\ell} u_j \alpha^j, f(\alpha)\Big) \leq \delta \right\}$$

has size at least a, then the following holds for every $\beta \in \mathbb{F}$: there exist codewords $c_0, \cdots, c_\ell \in \mathcal{C}$ and $c'_0, \cdots, c'_\ell \in \mathcal{C}$ such that the two curves of codewords defined by $g_0(\alpha) = \sum_{j=0}^{\ell} c_j \alpha^j$ and $g_1(\alpha) = \sum_{j=0}^{\ell} c'_j \alpha^j$ satisfy the following conditions:

1. *$g_0(\beta) = g_1(\beta)$*
2. *The sets $S_b := \{\alpha \in A \mid f(\alpha) = g_b(\alpha)\}$ for $b = 0, 1$ satisfy $|S_0|, |S_1| \geq \ell + 1$ and $|S_0 \cup S_1| \geq \ell + 2$.* ⌟

Observe that if there is a degree ℓ curve $g : \mathbb{F}_q \to \mathcal{C}$ such that $f(\alpha) = g(\alpha)$ for $\ell + 2$ different elements of A then the conditions are satisfied by simply letting the c_i and c'_is be equal to the coefficients of the polynomial. This point is highlighted more in Proposition 5.2.

The more interesting case is when no $\ell + 2$ are on a single curve, then this definition says that for any $\beta \in \mathbb{F}$, there are two distinct degree ℓ curves passing through $\ell + 1$ points in $f(A)$ and also intersecting at β. Since, this property holds for all β, we can use this to get a degenerate curve as well if the field is large. This point is highlighted in Proposition 5.3.

5.2 $\vee$ decoding implies Curve Decoding and vice versa

Proposition 5.2 (Curve decoding to $\vee$ decoding)**.** *Let $\mathcal{C} \subseteq \Sigma^n$ be an $\mathbb{F}$-additive $(\ell, \delta, a, \ell + 2)$ curve-decodable code then $\mathcal{C}$ is (ℓ, δ, a) $\vee$-decodable.*

Proof. Let $u_0, \cdots, u_\ell$ be codewords, $f : \mathbb{F}_q \to \mathcal{C}$ be a function such that

$$A = \left\{ \alpha \mid \Delta\Big(\sum_{j=0}^{\ell} u_j \alpha^j, f(\alpha)\Big) \leq \delta \right\} \text{ satisfies } |A| \geq a \ .$$

By the curve-decodability, there exist $c''_0, \cdots, c''_\ell$ such that if $g(\alpha) = \sum_{j=0}^{\ell} c''_j \alpha^j$ then

$$\left| \left\{ \alpha \in A \mid f(\alpha) = g(\alpha), \right\} \right| \geq \ell + 2 \ .$$

Thus, setting $c_i = c'_i = c''_i$ in the definition of $\vee$-decoding clearly suffices. □

Proposition 5.3 ($\vee$ decoding to curve decoding)**.** *Let $\mathcal{C} \subseteq \Sigma^n$ be an $\mathbb{F}$ additive (ℓ, δ, a) $\vee$-decodable code then $\mathcal{C}$ is $(\ell, \delta, a, \ell + 2)$ curve-decodable if $|\mathbb{F}| > a^{2\ell+2}$.*

Proof. Let $u_0, \cdots, u_\ell$ be codewords, $f : \mathbb{F}_q \to \mathcal{C}$ be a function be such that

$$A = \left\{ \alpha \mid \Delta\Big(\sum_{j=0}^{\ell} u_j \alpha^j, f(\alpha)\Big) \leq \delta \right\} \text{ satisfies } |A| \geq a \ .$$

For simplicity, we assume $|A| = a$ since we can just replace A with any subset. Now, if there are at least $\ell+2$ points on a single curve then there is nothing to prove. Now, observe that in the definition of curve decodability, the set S_b actually determines g_b as the $(\ell+1)\times(\ell+1)$ Vandermonde matrix is invertible. Thus, we consider all possible sets $S \subseteq A$ of size $\ell+1$. There are $\binom{a}{\ell+1}$ such subsets and thus $\binom{a}{\ell+1}$ different possible functions g. If any two different sets give the same function g then we are done. Else, every pair of functions agrees in at most ℓ distinct points but since V-decodability holds, we need to have $|\mathbb{F}| \leq \ell \cdot \binom{\binom{a}{\ell+1}}{2} \leq a^{2\ell+2}$, which is a contradiction. Thus, there are two functions that are the same and thus there are $\ell+2$ points on a single curve as desired. □

Via V decodability we can ensure that there are at least $\ell+2$ codewords on a degree ℓ curve amongst $f(\alpha)$, $\alpha \in A$. We can amplify this guarantee by assuming many more points of proximity on the curve.

Lemma 5.4. *Let $\mathcal{C} \subseteq \Sigma^n$ be an $\mathbb{F}$-additive $(\ell, \delta, a, \ell+2)$ curve-decodable code, then it is also a $(\ell, \delta, (m-1)\binom{a-1}{\ell+1}+1, m)$ curve decodable code.*

Proof. Let $u_0, \cdots, u_\ell$ be codewords, $f : \mathbb{F}_q \mapsto \mathcal{C}$ be a function.

Consider a maximal set $A \subseteq \mathbb{F}_q$ such that no $\ell+2$ points in $(\alpha, f(\alpha))$ lie on a degree $\ell+2$ curve. We know that $|A| \leq a-1$. Now consider all degree ℓ curves using points of A, there are at most $\binom{a-1}{\ell+1}$ such distinct curves. We know that all points in the space lie on at least one of these curves as otherwise we would contradict the maximality of A.

Thus, by the pigeonhole principle, one of the $\binom{a-1}{\ell+1}$ curves must contain m points as desired. □

Remark 5.5. *The above argument is quite lossy and extremely weak combinatorial bounds have been used. It is conceivable that significantly better bounds hold generically which would translate into better error terms in our final results.* ⌟

5.3 Not V-decoding is a local property

We will now create a family of local profiles whose presence in a code precisely witnesses not being V-decodable. Then, we will use the fact that subspace design codes are V-decodable together with the local property threshhold theorems to show that indeed random Reed-Solomon and Random Linear-Codes are also $(\ell, \delta - o(1), a)$ V-decodable, by virtue of not containing the relevant local profiles.

Lemma 5.6. *An $\mathbb{F}$-additive code $\mathcal{C} \subseteq \Sigma^n$ is (ℓ, δ, a) V-decodable, if for all $u_0, u_2, \cdots, u_\ell \in \Sigma^n$, $A = \{\alpha_1, \cdots \alpha_a\} \subseteq \mathbb{F}$, all $\beta \in \mathbb{F}$, and all codewords c_α, $\alpha \in A$, that satisfy*

$$\Delta(\sum_{j=0}^{\ell} u_j \alpha^j, c_\alpha) \leq \delta \quad \forall \alpha \in A$$

the following holds: There exist $S_0 = \{\beta_1, \cdots, \beta_{\ell+1}\}$ and $S_1 = \{\beta'_1, \cdots, \beta'_{\ell+1}\} \subseteq A$ with $S_0 \neq S_1$ and

$$M_{\beta_1,\cdots,\beta_{\ell+1}}(\beta)[c_{\beta_1} \ \cdots \ c_{\beta_{\ell+1}}]^T = M_{\beta'_1,\cdots,\beta'_{\ell+1}}(\beta)[c_{\beta'_1} \ \cdots \ c_{\beta'_{\ell+1}}]^T \ .$$

Here, we recall that $M_{\beta_1,\cdots,\beta_{\ell+1}}(\beta)$ is a row vector defined in Definition 2.36.

Proof. We just rewrote the definition of V-decoding by fixing $A, \beta \notin A, f(A)$ along with using the fact that any $\ell + 1$ sized set $S = \{\beta_1, \cdots, \beta_{\ell+1}\} \subset \mathbb{F}$ fixes g, and $g(\beta)$ is just given by the linear map $M_{\beta_1, \cdots, \beta_{\ell+1}}(\beta)$. □

Theorem 5.7. *The violation of (ℓ, δ, a) V-decodability is an $\binom{a}{\ell+1}$-LCL property containing at most $\binom{n}{\delta n}^a \cdot q^{a+1}$ profiles.*

Proof. Observe that in the above definition, we can assume that $c_{\alpha_1}, \cdots, c_{\alpha_{\ell+1}}$ are all the 0 codewords.

This is true since if there exist $u_0, \cdots, u_\ell$, $A = \{\alpha_1, \cdots, \alpha_a\}$, c_α for each $\alpha \in A$ such that $\Delta(\sum_{j=0}^{\ell} u_j \alpha^j, c_\alpha) \leq \delta \quad \forall \alpha \in A$

Now, let $\hat{c}_0, \cdots, \hat{c}_\ell$ be codewords such that $\sum_{j=0}^{\ell} \hat{c}_j \alpha_i^j = c_{\alpha_i}$ for $i \in [\ell + 1]$. These exist by Lemma 2.37.

Now, let $u'_i = u_i - \hat{c}_i$ for all $i \in \{0, \cdots, \ell\}$, $c'_\alpha = c_\alpha - \sum_{j=0}^{\ell} \hat{c}_j \alpha^j$. Thus, we still have $\Delta(\sum_{j=0}^{\ell} u'_j \cdot \alpha^j, c'_\alpha) = \Delta(\sum_{j=0}^{\ell} u_j \alpha^j, c_\alpha) \leq \delta$ for all $\alpha \in A$. This lets us assume that $c'_{\alpha_i} = 0$ for $i \in [\ell + 1]$. Observe that since we only shifted by a degree ℓ curve, all other curves also stay preserved under translation by $\sum_{j=0}^{\ell} \hat{c}_j \alpha^j$.

Thus, we assume from now on that $c_{\alpha_1}, \cdots, c_{\alpha_{\ell+1}}$ are all all 0 vectors.

Now, we cast (ℓ, δ, a) V-decodability as not containing a carefully defined local profile.

Fix arbitrary sets $\mathcal{I} = \{I_1, , \cdots, I_a\} \in (\binom{n}{(1-\delta)n})^a$, $\vec{\alpha} = \alpha_1, \cdots, \alpha_a \in \mathbb{F}_q$, and $\beta \in \mathbb{F} \setminus \{\alpha_1, \cdots, \alpha_a\}$. Now define $\mathcal{V}_{\mathcal{I}, \vec{\alpha}, \beta}$ as follows.

First, we define an a-local profile $\mathcal{W}_{\mathcal{I}, \vec{\alpha}} = (\mathcal{W}_1, \cdots, \mathcal{W}_n) \in \mathcal{L}(\mathbb{F}_q^a)^n$.

We define each $\mathcal{W}_i$ separately by describing linear constraints on it. Let w_{ij} refer to the jth coordinate of the b coordinates inside $\mathcal{W}_i$.

- $\mathcal{W}_i \in \mathcal{L}(\mathbb{F}_q^a)$
- $w_{i,j} = 0 \;\; \forall j \in [\ell + 1]$
- Let $T_i = \{j \mid j \in I_i\}$ then for all $t_1, \cdots, t_{\ell+2} \in T_i$: we have the linear constraint that

$$M_{\alpha_{t_1}, \cdots, \alpha_{t_{\ell+1}}}(\beta)[w_{i,t_1} \;\; \cdots \;\; w_{i,t_{\ell+1}}]^T = w_{i,t_{\ell+2}} \; .$$

If there exists a matrix $M \in \mathbb{F}^{n \times b}$ such that each column of M is a codeword and each row is in W_i then we can recover $u_0, \cdots, u_\ell$ using the points of agreement i.e., $\mathcal{I}$ such that $\Delta\left(\sum_{j=0}^{\ell} u_j \alpha_i^j, c_{\alpha_i}\right) \leq 1 - |I_i|/n = \delta n$. Now, we need to capture the condition of sets S_0, S_1 existing as described above.

Now, consider the linear map

$$M_{\beta, \ell, \vec{\alpha}}(\alpha_1, \cdots, \alpha_a) : \mathbb{F}^a \to \mathbb{F}^{\binom{a}{\ell+1}}$$

where we index the entries of the vector in the range naturally using tuples $(i_1, \cdots, i_{\ell+1})$ with $1 \leq i_1 < \cdots i_{\ell+1} \leq a$. The entry corresponding to $i_1, \cdots, i_{\ell+1}$ of $M_{\beta, \ell, \vec{\alpha}}(v)$ is given by product $M_{\alpha_{i_1}, \cdots, \alpha_{i_{\ell+1}}}(\beta)[v_{i_1}, v_{i_2}, \cdots, v_{i_{\ell+1}}]^T$.

Finally we define $\mathcal{V}_{\mathcal{I}, \vec{\alpha}, \beta, i} = M_{\beta, \ell, \vec{\alpha}} \mathcal{W}_i$ for all $i \in [n]$.[3].

[3] This map is rank preserving as we know that the first $\ell + 1$ coordinates are 0 so if the result is the all 0 vector then we know that the curve passing through $\beta, \alpha_1, \cdots, \alpha_\ell, \alpha_{i'}$ is the 0 curve for all $i' \in [a]$.

Observe that if a code $\mathcal{C}$ contains $\mathcal{V}_{\mathcal{I},\vec{\alpha},\beta}$ then by the previous argument, we know that there exist $u_0, \cdots, u_\ell \in \Sigma^n, c_{\alpha_1}, \cdots, c_{\alpha_a} \in \mathcal{C}$, and $c_{\alpha_i,e} = \sum_{j=0}^{\ell} u_j \alpha_{i,e}^j$ for all $e \in I_i$. Additionally, $c_{\alpha_1}, \cdots, c_{\alpha_{\ell+1}}$ are all 0, and there are no sets $S_0 = \{\beta_1, \cdots, \beta_{\ell+1}\}, S_1 = \{\beta_1', \cdots, \beta_{\ell+1}'\}$ such that $S_0 \neq S_1$ but $M_{\beta_1,\cdots,\beta_{\ell+1}}(\beta)[c_{\beta_1}\ c_{\beta_2}\ \cdots\ c_{\beta_{\ell+1}}] = M_{\beta_1',\cdots,\beta_{\ell+1}'}(\beta)[c_{\beta_1'}\ c_{\beta_2'}\ \cdots\ c_{\beta_{\ell+1}'}]$.

Now, if $\mathcal{C}$ does not contain any $\mathcal{V}_{\mathcal{I},\vec{\alpha},\beta}$, then we know that the code is indeed $\vee$ decodable. Additionally, if a code $\mathcal{C}$ is $\vee$ decodable, then it cannot contain any such profile.

Thus, $\mathcal{C}$ is (ℓ, δ, a) V-decodable iff it does not contain any of the above described $\leq \binom{n}{\delta n}^a \cdot q^{a+1}$ local profiles with locality $\binom{a}{\ell+1}$. □

We will denote the union of all local profiles from Theorem 5.7 to be $\mathcal{P}_{a,\ell,\delta}$ (we will suppress the dependence on $n.q$ for simplicity).

5.4 V (and curve) decodability of random linear and RS codes

Lemma 5.8. *For all integers $\ell \geq 1$ and $R, \varepsilon \in (0,1)$, all $\lceil(\ell+1)/\varepsilon\rceil$-strong subspace-design codes of rate R are $(\ell, 1-R-\varepsilon, (\ell+2)^2/\varepsilon^2)$ V-decodable.*

Proof. Follows from Theorem 4.7 since we know that $\ell + 2$ curve decoding implies V-decoding by Proposition 5.2. □

Thus, all $\lceil(\ell+1)/\varepsilon\rceil$-strong subspace-design codes do not satisfy the property $\mathcal{P}_{a,\ell,\delta}$.

Lemma 5.9. *For all $\mathcal{V} \in \mathcal{P}_{\ell,\delta}$, we have that $R_{\mathcal{V}} \geq 1 - \delta - \varepsilon - \frac{O_{\ell,\varepsilon}(1)}{n}$.*

Proof. We begin by setting $\delta = 1 - R - \varepsilon$.

Then, by [BCDZ25b], Theorem 2.31, we get that for all $\mathcal{V} \in \mathcal{P}_{\ell,1-R-\varepsilon}$, we have that $\mathcal{R}_{\mathcal{V}} \geq R - \frac{\binom{(\ell+2)^2/\varepsilon^2+1}{\ell+1}^2}{n}$. In particular, for all $\mathcal{V} \in \mathcal{P}_{\ell,1-R-\varepsilon}$, we have that $R_{\mathcal{V}} \geq R - \frac{O_{\ell,\varepsilon}(1)}{n}$ as desired. □

Lemma 5.10 (Curve and V decodability of RLCs)**.** *For all integers ℓ and $R, \varepsilon \in (0,1)$, there exist positive integers C_1, C_2 such that for all $q > C_1$, $n > C_2$, we have that a random linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of rate R is*

- *$(\ell, 1-R-2\varepsilon, (\ell+2)^2/\varepsilon^2)$ V-decodable*
- *$(\ell, 1-R-2\varepsilon, m \cdot \mathrm{poly}(\ell, 1/\varepsilon)^\ell, m)$ curve decodable for all positive integers m.*

with probability 5/6. [4]

Proof. Let $a = \lceil(\ell+2)^2/\varepsilon^2\rceil$.

Observe that if we let $\delta = 1 - R - 2\varepsilon$, then observe that for all $\mathcal{V} \in \mathcal{P}_{\ell,\delta}$, we have that $R_{\mathcal{V}} \geq R + \varepsilon - \frac{O_{\ell,\varepsilon}(1)}{n}$. Thus, the probability that a rate R code contains at least one of the profiles in $\mathcal{P}_{a,\ell,\delta}$ is at most $q^{-\varepsilon n + O_{\ell,\varepsilon}(1)} \cdot q^{a+1} \cdot 2^{na}$ by Theorem 2.30 [LMS25]. Thus, the V decodability follows for all large enough q and n.

Now, once we have V decoding then curve decoding simply follows by Proposition 5.3 and Lemma 5.4. □

[4] Can be made arbitrarily close to 1.

Lemma 5.11 (RRS are V and curve-decodable)**.** *There exists an absolute constant C_1 such that for all constant positive integers ℓ, and $R, \varepsilon \in (0,1)$, there exists a C_2, such that for all $n > C_1 \cdot (e\ell/\varepsilon^2)^{2\ell+2}/\varepsilon$ and prime powers $q > nC_2$, random Reed-Solomon codes $\mathcal{C} \subseteq \mathbb{F}_q^n$ of rate R are*

- *$(\ell, 1-R-2\varepsilon, (\ell+2)^2/\varepsilon^2)$ V-decodable*
- *$(\ell, 1-R-2\varepsilon, m \cdot O((e \cdot \ell/\varepsilon^2)^{\ell+1}), m)$ curve decodable for all positive integers m*

with probability atleast $5/6$.[5]

Proof. Let $a = \lceil (\ell+2)^2/\varepsilon^2 \rceil$ and $b = \binom{a}{\ell+1}$. Observe that now, we have that $b = O\left(\frac{a^{\ell+1}}{(\ell+1)!}\right) = O\left(\frac{(ae)^{\ell+1}}{(\ell+1)^{\ell+1}}\right) = O((e\ell/\varepsilon^2)^{\ell+1})$. Thus, $b^2+1 < \varepsilon n/2$ for the condition on n that we have.

Observe that if we let $\delta = 1-R-2\varepsilon$, then observe that for all $\mathcal{V} \in \mathcal{P}_{a,\ell,\delta}$, we have that $R_{\mathcal{V}} \geq R + \varepsilon - \frac{b^2+1}{n} \geq R + \varepsilon/2$. Thus, the probability that a rate R code contains at least one of the profiles in $\mathcal{P}$ is at most $(2^b - 1) \cdot \left(\frac{(4b)^{4b} \cdot Rn}{q}\right)^{\varepsilon n/4b} \cdot \binom{n}{\delta n} \cdot q^{a+1}$ by Theorem 2.30 [LMS25].

Now, we have that $\binom{n}{\delta n} < 2^n = \left(2^{4b/\varepsilon}\right)^{\varepsilon n/4b}$. Similarly, $2^b \leq 4^{\varepsilon n/4b}$. Letting $d_1 = (4b)^{4b} \cdot 4^{4b} \cdot 2^{4b/\varepsilon}$ and $d_2 = 4b(a+1)/\varepsilon$, the probability that a rate R random Reed-Solomon codes contains at least one of the profiles is at most $\left(\frac{d_1 n}{q}\right)^n \cdot q^{d_2}$. Since we can now treat d_1 and d_2 as constants and we know that $n > d_2$, the result follows.

Now, once we have V decoding then curve decoding simply follows by Proposition 5.3 and Lemma 5.4. □

Lemma 5.12 (Explicit V decodable codes with constant sized alphabet)**.** *For all positive integers ℓ, and $R, \varepsilon \in (0,1)$ there exist constants C_1, C_2, C_3 such that for all prime powers $q > C_1$, $s > C_2$, and $n > C_3$ there exists an explicit AEL based additive code $\mathcal{C} \subseteq ((\mathbb{F}_q)^s)^n$ of rate R which is*

- *$(\ell, 1-R-2\varepsilon, (\ell+2)^2/\varepsilon^2)$ V-decodable.*
- *$(\ell, 1-R-2\varepsilon, m \cdot \mathrm{poly}(\ell, 1/\varepsilon)^\ell, m)$ curve decodable for all positive integers m.*

Proof. Observe that as before if we let $\delta = 1-R-2\varepsilon$, we have that for all $\mathcal{V} \in \mathcal{P}_{\ell,\delta}$, we have that $R_{\mathcal{V}} \geq R + \varepsilon - \frac{O_{\ell,\varepsilon}(1)}{n}$, and the number of local profiles in $\mathcal{P}_{\ell,\delta}$ is at most $2^n \cdot q^{a+1}$ and since the constraints are given by the choices of sets, we must have that for all $\mathcal{V} \in \mathcal{P}_{a,\ell,\delta}$ and permutation $\pi : [n] \mapsto [n]$, we have that $\pi(\mathcal{V}) \in \mathcal{P}_{a,\ell,\delta}$ as well. Thus, the result follows from Theorem 2.35. □

5.5 Final Results on Correlated Agreement of random linear and RS codes

Theorem 5.13 (Curve Based Proximity Gaps for Random Linear Codes)**.** *For all positive integers ℓ and $R, \varepsilon \in (0,1)$, there exist positive integers C_1, C_2 such that for all $q > C_1$, $n > C_2$, it holds with probability $2/3$ that a random linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of rate R has*

- *$(\ell, 1-R-2\varepsilon, \frac{m\ell(1-R)+\mathrm{poly}(\ell,1/\varepsilon)^\ell}{\varepsilon|\mathbb{F}|}, 1/m)$ mutual correlated agreement for all positive integers m.*
- *$(\ell, 1-R-2\varepsilon, \frac{n\ell(1-R)+\mathrm{poly}(\ell,1/\varepsilon)^\ell}{\varepsilon|\mathbb{F}|}, 0)$ perfect mutual correlated agreement.*

[5] This probability can be made arbitrarily close to 1.

- $(1-R-2\varepsilon, \frac{n(1-R)+\mathrm{poly}(1/\varepsilon)}{\varepsilon|\mathbb{F}|}, 0)$ *perfect line mutual correlated agreement.*

Proof. Observe that by Lemma 5.10, we know that with probability 5/6, the code $\mathcal{C}$ is $(\ell, 1-R-2\varepsilon, \mathrm{poly}(\ell, 1/\varepsilon)^\ell, \ell/\varepsilon)$ curve decodable. Now, combining with Theorem 3.4, we just need that with probability at least 5/6, a rate R random linear code $\mathcal{C}$ is $(1-R-3\varepsilon/2, 1/\varepsilon)$ list-decodable. This holds by Theorem 2.24 as desired.

The second result follows by simply setting $m = n$ and the third by setting $\ell = 1$. □

Theorem 5.14 (Curve Based Proximity Gaps for Random Reed-Solomon Codes)**.** *There exists a constant C_1 such that for all constant positive integer ℓ, and reals $\varepsilon \in (0,1)$ there exists a C_2, such that for all $n > C_1 \cdot (e/\varepsilon)^{4\ell+5} \cdot \ell^{2\ell+2}$ and $q > nC_2$, for any $R \in (0,1)$, a random Reed-Solomon code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of rate R has*

- $(\ell, 1-R-2\varepsilon, \frac{m\ell(1-R)+O((e\ell)^{\ell+2}/\varepsilon^{2\ell+2})}{\varepsilon|\mathbb{F}|}, 1/m)$ *mutual correlated agreement for all positive integers m*
- $(\ell, 1-R-2\varepsilon, \frac{n\ell(1-R)+O(\ell^{\ell+2}\cdot e^\ell/\varepsilon^{2\ell+2})}{\varepsilon|\mathbb{F}|}, 0)$ *perfect mutual correlated agreement.*
- $(1-R-2\varepsilon, \frac{n(1-R)+O(1/\varepsilon^4)}{\varepsilon|\mathbb{F}|}, 0)$ *perfect line mutual correlated agreement*

with probability 2/3*.*

Proof. Observe that by Lemma 5.11, we know that with probability 5/6, the code $\mathcal{C}$ is $(\ell, 1-R-2\varepsilon, O((e\ell)^{\ell+2}/\varepsilon^{2\ell+3}), \ell/2\varepsilon)$ curve decodable. Now, combining with Theorem 3.4, we just need that with probability at least 5/6, a rate R random Reed-Solomon code $\mathcal{C}$ is $(1-R-3\varepsilon/2, 1/\varepsilon)$ list-decodable. This holds by Theorem 2.25 as desired.

The second result follows by simply setting $m = n$ and the third by setting $\ell = 1$. □

Theorem 5.15 (Explicit Constant Alphabet Codes with Curve Based MCA)**.** *For all positive integers ℓ, and $R, \varepsilon \in (0,1)$ there exist constants C_1, C_2, C_3 such that for all prime powers $q > C_1$, $s > C_2$, and $n > C_3$ there exists an explicit AEL based additive code $\mathcal{C} \subseteq ((\mathbb{F}_q)^s)^n$ of rate R which has $(\ell, 1-R-2\varepsilon, \frac{m\ell(1-R)+\mathrm{poly}(\ell/\varepsilon)^\ell}{\varepsilon|\mathbb{F}|}, 1/m)$ mutual correlated agreement for all positive integers m.*

Proof. By Lemma 5.12, there is an explicit code with the required properties which is also curve decodable as desired. Now, since the same code could also be made to have optimal list sizes by adding the list decoding properties of RLCs as reasonable local properties to be avoided as well, we are done using Theorem 3.4. □

Theorem 5.16 (Affine Space Proximity Gaps for RLCs)**.** *For all $R, \varepsilon \in (0,1)$, there exist positive integers C_1, C_2 such that for all $q > C_1$, $n > C_2$, with probability at least 2/3 a random linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of rate R has:*

- $(1-R-2\varepsilon, \frac{n(1-R)+\mathrm{poly}(1/\varepsilon)}{\varepsilon(|\mathbb{F}|-1)}, 0)$ *affine space perfect correlated agreement*
- $(1-R-2\varepsilon, \frac{m(1-R)+\mathrm{poly}(1/\varepsilon)}{\varepsilon(|\mathbb{F}|-1)}, 1/m)$ *affine space proximity gap for all positive integers m.*

Proof. The result follows by considering the $\ell = 1$ case in Theorem 5.13 and applying Lemma 2.15 and Lemma 2.13 respectively. □

Theorem 5.17 (Affine Space Proximity Gaps for RRSs)**.** *There exists a constant C_1 such that for all $R, \varepsilon \in (0,1)$ there exists a constant C_2, such that for all $n > C_1 \cdot (1/\varepsilon)^9$ and prime powers $q > nC_2$, it holds with probability at least $2/3$ that a random Reed-Solomon code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of rate R has:*

- $(1 - R - 2\varepsilon, \frac{n(1-R)+O(1/\varepsilon^4)}{\varepsilon(|\mathbb{F}|-1)}, 0)$ *affine space perfect correlated agreement*
- $(1 - R - 2\varepsilon, \frac{m(1-R)+O(1/\varepsilon^4)}{\varepsilon(|\mathbb{F}|-1)}, 1/m)$ *proximity gaps for all positive integers m.*

Proof. The result follows by considering the $\ell = 1$ case in Theorem 5.14 and applying Lemma 2.15 and Lemma 2.13 respectively. □

Theorem 5.18 (Explicit constant alphabet codes with affine space proximity gaps)**.** *For all $R, \varepsilon \in (0,1)$ there exist constants C_1, C_2, C_3 such that for all prime powers $q > C_1$, $s > C_2$, and $n > C_3$ there exists an explicit AEL based additive code $\mathcal{C} \subseteq ((\mathbb{F}_q)^s)^n$ of rate R which has $(1 - R - 2\varepsilon, \frac{m(1-R)+\mathrm{poly}(1/\varepsilon)}{\varepsilon|\mathbb{F}|}, 1/m)$ affine space proximity gap for all positive integers m.*

Proof. The result follows by considering the $\ell = 1$ case in Theorem 5.15 and then applying Lemma 2.13. □

6 Conclusions and Open Questions

In this paper, we presented the first results establishing optimal proximity gaps answering the conjecture of [BCI+20] for various families of codes such as folded Reed-Solomon, random Reed-Solomon, expander based codes, and random linear codes.

There are various natural follow up questions that come up as extensions of our result. We mention a few here:

- **Error bound ℓ dependencies:** Can the dependence on ℓ be improved? Our algorithms result in a $\mathrm{poly}(\ell/\varepsilon)^\ell$ dependence for mutual correlated agreement for curve based proximity generators. This is due to the weakness of the local transform going through V decoding (Definition 5.1) and the combinatorial problem (Lemma 5.4). Improving either would lead to improvements in this dependency.
- **Generic transforms for affine space proximity generators:** Our current results are mostly written for curve based proximity generators. This is since no generic transforms are known from slacked correlated agreement to slacked affine space correlated agreement. We do not know of any generic transform for curve based MCA to affine space MCA either even in perfect agreement regimes. Are there generic transforms to get these extensions?

  It is also interesting to ask whether it is possible at all to use line or affine space based proximity gaps to get any sort of generic result for curve based proximity gaps. See Lemma 2.13 and Lemma 2.15. The known results only work for proximity gaps and for correlated agreement in the perfect agreement from lines to affine spaces.
- **Generalizing the LCL literature:** Is it possible to generalize local profiles literature to handle more general constraints other than just no pairwise-equal (as stipulated in Definition 2.27 and Definition 2.28). This might allow us to directly translate curve decodability from subspace-designs to random RS codes, rather than going through the intermediate techical notion of V decodability.

7 Acknowledgements

We thank Joshua Brakensiek for useful feedback and valuable discussions on the relationship between subspace-designs and random linear codes. We also thank Dan Boneh for comments and suggestions on a draft of this paper. R.G. thanks Yael Kalai and Anastasiia Struss for useful discussions on proximity gaps.

V.G. is supported by a Simons Investigator award, NSF grant CCF-2211972, and ONR grant N00014-24-1-2491.

R.G. is supported by (Yael Tauman Kalai's) grant from Defense Advanced Research Projects Agency (DARPA) under Contract No. HR0011-25-C-0300. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Defense Advanced Research Projects Agency (DARPA).

References

[ACFY24a] Gal Arnon, Alessandro Chiesa, Giacomo Fenzi, and Eylon Yogev. STIR: Reed–solomon proximity testing with fewer queries. Cryptology ePrint Archive, Paper 2024/390, 2024. 1, 2

[ACFY24b] Gal Arnon, Alessandro Chiesa, Giacomo Fenzi, and Eylon Yogev. WHIR: Reed–solomon proximity testing with super-fast verification. Cryptology ePrint Archive, Paper 2024/1586, 2024. 1, 2, 6, 7

[AGG$^+$25] Omar Alrabiah, Zeyu Guo, Venkatesan Guruswami, Ray Li, and Zihan Zhang. Random Reed-Solomon codes achieve list-decoding capacity with linear-sized alphabets. *Advances in Combinatorics*, October 2025. 2, 4, 11

[AHIV22] Scott Ames, Carmit Hazay, Yuval Ishai, and Muthuramakrishnan Venkitasubramaniam. Ligero: Lightweight sublinear arguments without a trusted setup. Cryptology ePrint Archive, Paper 2022/1608, 2022. 1, 6

[AHS25] Vikrant Ashvinkumar, Mursalin Habib, and Shashank Srivastava. Algorithmic improvements to list decoding of folded Reed-Solomon codes. *arXiv preprint arXiv:2508.12548*, 2025. 4, 10, 17, 18, 21

[BBHR18] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Fast Reed-Solomon interactive oracle proofs of proximity. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPIcs*, pages 14:1–14:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. 1

[BCDZ25a] Joshua Brakensiek, Yeyuan Chen, Manik Dhar, and Zihan Zhang. Combinatorial bounds for list recovery via discrete Brascamp–Lieb inequalities. *arXiv preprint arXiv:2510.13775*, 2025. 5, 17

[BCDZ25b] Joshua Brakensiek, Yeyuan Chen, Manik Dhar, and Zihan Zhang. From random to explicit via subspace designs with applications to local properties and matroids. *arXiv preprint arXiv:2510.13777*, 2025. 2, 4, 5, 11, 12, 17, 25

[BCI$^+$20] Eli Ben-Sasson, Dan Carmon, Yuval Ishai, Swastik Kopparty, and Shubhangi Saraf. Proximity gaps for Reed-Solomon codes. In Sandy Irani, editor, *61st IEEE Annual Symposium on Foundations of Computer Science*, pages 900–909. IEEE, 2020. 1, 3, 5, 6, 28

[BDGZ25] Joshua Brakensiek, Manik Dhar, Sivakanth Gopi, and Zihan Zhang. AG codes achieve list-decoding capacity over constant-sized fields. *IEEE Trans. Inf. Theory*, 71(8):5935–5956, 2025. 4

[BGM22] Joshua Brakensiek, Sivakanth Gopi, and Visu Makam. Generic Reed-Solomon codes achieve list-decoding capacity. *CoRR*, abs/2206.05256, 2022. 2, 4, 11

[BKS18] Eli Ben-Sasson, Swastik Kopparty, and Shubhangi Saraf. Worst-case to average case reductions for the distance to a code. In Rocco A. Servedio, editor, *33rd Computational Complexity Conference, CCC 2018, June 22-24, 2018, San Diego, CA, USA*, volume 102 of *LIPIcs*, pages 24:1–24:23. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. 1, 2, 6

[BS21] Eli Ben-Sasson. ethSTARK documentation. Cryptology ePrint Archive, Paper 2021/582, 2021. 2

[BSBHR18] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Fast Reed-Solomon Interactive Oracle Proofs of Proximity. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 14:1–14:17, Dagstuhl, Germany, 2018. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 1

[BSCH$^+$25] Eli Ben-Sasson, Dan Carmon, Ulrich Haböck, Swastik Kopparty, and Shubhangi Saraf. On proximity gaps for reed–solomon codes, 2025. Manuscript. 5

[BSCI$^+$23] Eli Ben-Sasson, Dan Carmon, Yuval Ishai, Swastik Kopparty, and Shubhangi Saraf. Proximity gaps for reed–solomon codes. *J. ACM*, 70(5), October 2023. 1, 2, 8

[BSCS16] Eli Ben-Sasson, Alessandro Chiesa, and Nicholas Spooner. Interactive oracle proofs. In *Proceedings, Part II, of the 14th International Conference on Theory of Cryptography - Volume 9986*, page 31–60, Berlin, Heidelberg, 2016. Springer-Verlag. 1

[BSGKS19] Eli Ben-Sasson, Lior Goldberg, Swastik Kopparty, and Shubhangi Saraf. DEEP-FRI: Sampling outside the box improves soundness. Cryptology ePrint Archive, Paper 2019/336, 2019. 1, 2

[CS25] Elizabeth Crites and Alistair Stewart. On reed–solomon proximity gaps conjectures. Cryptology ePrint Archive, Paper 2025/2046, 2025. 5

[CZ25] Yeyuan Chen and Zihan Zhang. Explicit folded Reed-Solomon and multiplicity codes achieve relaxed generalized singleton bound. In Michal Koucký and Nikhil Bansal, editors, *Proc. 57th ACM Symp. on Theory of Computing (STOC)*, pages 1–12, 2025. 3, 9, 10

[DG25] Benjamin E. Diamond and Angus Gruen. On the distribution of the distances of random words. Cryptology ePrint Archive, Paper 2025/2010, 2025. 5

[Eth25] Ethereum Foundation. The proximity prize – $1m proximity gaps challenge, 2025. https://proximityprize.org/, Accessed: 2025-11-02. 2

[GCXK25] Yiwen Gao, Dongliang Cai, Yang Xu, and Haibin Kan. From list-decodability to proximity gaps. Cryptology ePrint Archive, Paper 2025/870, 2025. 2

[GHKS24] Rohan Goyal, Prahladh Harsha, Mrinal Kumar, and Ashutosh Shankar. Fast list-decoding of univariate multiplicity and folded Reed-Solomon codes. In Santosh Vempala, editor, *Proc. 65th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 328–343, 2024. 21

[GHKS25] Rohan Goyal, Prahladh Harsha, Mrinal Kumar, and Ashutosh Shankar. Fast list-recovery of univariate multiplicity and folded Reed-Solomon codes. 2025. Manuscript. 21

[GK16] Venkatesan Guruswami and Swastik Kopparty. Explicit subspace designs. *Combinatorica*, 36(2):161–185, 2016. 3, 9, 10, 21

[GKL24] Yiwen Gao, Haibin Kan, and Yuan Li. Linear proximity gap for linear codes within the 1.5 johnson bound. Cryptology ePrint Archive, Paper 2024/1810, 2024. 7

[GLM$^+$22] Venkatesan Guruswami, Ray Li, Jonathan Mosheiff, Nicolas Resch, Shashwat Silas, and Mary Wootters. Bounds for list-decoding and list-recovery of random linear codes. *IEEE Transactions on Information Theory*, 68(2):923–939, February 2022. 4

[GM24] Venkatesan Guruswami and Jonathan Mosheiff. Punctured low-bias codes behave like random linear codes. *Discrete Analysis*, June 2024. Preliminary version in FOCS'22. 4

[GMR$^+$22] Venkatesan Guruswami, Jonathan Moshieff, Nicolas Resch, Shashwat Silas, and Mary Wootters. Threshold rates for properties of random codes. *IEEE Trans. Inf. Theory*, 68(2):905–922, 2022. 4

[GR08] Venkatesan Guruswami and Atri Rudra. Explicit codes achieving list decoding capacity: Error-correction with optimal redundancy. *IEEE Transactions on Information Theory*, 54(1):135–150, 2008. 2, 10

[GW13a] Venkatesan Guruswami and Carol Wang. Linear-algebraic list decoding for variants of Reed-Solomon codes. *IEEE Transactions on Information Theory*, 59(6):3257–3268, 2013. 2

[GW13b] Venkatesan Guruswami and Carol Wang. Linear-algebraic list decoding for variants of Reed-Solomon codes. *IEEE Trans. Inform. Theory*, 59(6):3257–3268, 2013. (Preliminary version in *26th IEEE Conference on Computational Complexity*, 2011 and *15th RANDOM*, 2011). 21

[GX13] Venkatesan Guruswami and Chaoping Xing. List decoding Reed-Solomon, Algebraic-Geometric, and Gabidulin subcodes up to the Singleton bound. In *Proceedings of the 45th ACM Symposium on Theory of Computing*, pages 843–852, June 2013. 9

[GXYZ24] Zeyu Guo, Chaoping Xing, Chen Yuan, and Zihan Zhang. Random gabidulin codes achieve list decoding capacity in the rank metric. In *65th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2024, Chicago, IL, USA, October 27-30, 2024*, pages 1846–1873. IEEE, 2024. 4

[Hex25] Hexens. The ethereum proximity prize: Reed–solomon codes and MCA, 2025. https://hexens.io/blog/proximity-gaps, Accessed: 2025-11-02. 2

[JS25] Fernando Granha Jeronimo and Nikhil Shagrithaya. Probabilistic guarantees to explicit constructions: Local properties of linear codes. *arXiv preprint arXiv:2510.06185*, 2025. 3, 4, 5, 12

[Kop14] Swastik Kopparty. Some remarks on multiplicity codes. In Alexander Barg and Oleg R. Musin, editors, *Discrete Geometry and Algebraic Combinatorics*, volume 625 of *Contemporary Mathematics*, pages 155–176. AMS, 2014. 2

[KRSW23] Swastik Kopparty, Noga Ron-Zewi, Shubhangi Saraf, and Mary Wootters. Improved list decoding of Folded Reed-Solomon and Multiplicity codes. *SIAM J. Comput.*, 52(3):794–840, 2023. (Preliminary version in *59th FOCS*, 2018). 3, 10, 17

[LMS25] Matan Levi, Jonathan Mosheiff, and Nikhil Shagrithaya. Random Reed-Solomon codes and random linear codes are locally equivalent. *arXiv preprint arXiv:2406.02238*, 2025. 2, 4, 5, 11, 12, 17, 25, 26

[MRRZ$^+$21] Jonathan Mosheiff, Nicolas Resch, Noga Ron-Zewi, Shashwat Silas, and Mary Wootters. Low-density parity-check codes achieve list-decoding capacity. *SIAM Journal on Computing*, 53(6):FOCS20–38–FOCS20–73, November 2021. 4

[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*, STOC '13, page 793–802, New York, NY, USA, 2013. Association for Computing Machinery. 6

[RZ18] Ron Roth and Gilles Zémor. Personal communication to the authors of [ahiv17]. Personal communication, 2018. Cited in Ligero[AHIV17], `https://eprint.iacr.org/2022/1608`. 6

[Sri25] Shashank Srivastava. Improved list size for folded Reed-Solomon codes. In Yossi Azar and Debmalya Panigrahi, editors, *Proc. 36th Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 2040–2050, 2025. 3, 10

[Tam24] Itzhak Tamo. Tighter list-size bounds for list-decoding and recovery of folded Reed-Solomon and multiplicity codes. *IEEE Trans. Inform. Theory*, 70(12):8659–8668, 2024. 3, 10, 17

[ZCF23] Hadas Zeilberger, Binyi Chen, and Ben Fisch. BaseFold: Efficient field-agnostic polynomial commitment schemes from foldable codes. Cryptology ePrint Archive, Paper 2023/1705, 2023. 1

[Zei24] Hadas Zeilberger. Khatam: Reducing the communication complexity of code-based SNARKs. Cryptology ePrint Archive, Paper 2024/1843, 2024. 7