Linear Proximity Gap for Reed-Solomon Codes within the 1.5 Johnson Bound

Yiwen Gao*, Haibin Kan†and Yuan Li‡

November 5, 2024

Abstract

We establish a linear proximity gap for Reed-Solomon (RS) codes within the one-and-a-half Johnson bound. Specifically, we investigate the *proximity gap* for RS codes, revealing that any affine subspace is either entirely δ-close to an RS code or nearly all its members are δ-far from it. When δ is within the one-and-a-half Johnson bound, we prove an upper bound on the number of members (in the affine subspace) that are δ-close to the RS code for the latter case. Our bound is linear in the length of codewords. In comparison, Ben-Sasson, Carmon, Ishai, Kopparty and Saraf [FOCS 2020] prove a linear bound when δ is within the unique decoding bound and a quadratic bound when δ is within the Johnson bound. Note that when the rate of the RS code is smaller than 0.23, the one-and-a-half Johnson bound is larger than the unique decoding bound.

Proximity gaps for Reed-Solomon (RS) codes have implications in various RS code-based protocols. In many cases, a stronger property than individual distance—known as *correlated agreement*—is required, i.e., functions in the affine subspace are not only δ-close to an RS code, but also agree on the same evaluation domain. Our results support this stronger property.

*School of Computer Science, Fudan University, Shanghai 200433, China. Email: ywgao21@m.fudan.edu.cn
†School of Computer Science, Fudan University, Shanghai 200433, China. Email: hbkan@fudan.edu.cn
‡School of Computer Science, Fudan University, Shanghai 200433, China. Email: yuan_li@fudan.edu.cn

Contents

1 Introduction

Reed-Solomon (RS) codes [RS60] are a class of error-correcting codes. They are fundamental objects of study in algebraic coding theory and theoretical computer science. Let $\mathbb{F}_q$ be a finite field with q elements, and let $L \subseteq \mathbb{F}_q$ be the evaluation domain. Let $\rho \in (0, 1]$ be the code rate and $n = |L|$ be the code length. Let $\mathrm{RS}[\mathbb{F}_q, L, \rho]$ denote the set of functions $f : L \to \mathbb{F}_q$ that are evaluation results of polynomials of degree strictly less than $\rho|L|$. Reed-Solomon codes have a wide range of applications. For example, many protocols in areas such as blockchain, distributed storage, and cryptography utilize Reed-Solomon codes as essential building blocks. In some protocols, the soundness of relies on the existence of a series of vectors that are close to the Reed-Solomon (RS) code (in relative Hamming distance). Consequently, it is critical to efficiently identify vectors that are far from the RS code.

The *RS Proximity Testing* (RPT) problem involves a verifier determining whether a given function $f : L \to \mathbb{F}_q$ is a member of $\mathrm{RS}[\mathbb{F}_q, L, \rho]$ or is far from all codewords in $\mathrm{RS}[\mathbb{F}_q, L, \rho]$. The verifier has limited query access to f, and an untrusted prover may assist the verifier. We consider this problem under the interactive oracle proofs of proximity (IOPP) model [BCS16] (also called probabilistically checkable interactive proofs of proximity in [RRR16]). This model combines aspects of probabilistically checkable proof (PCPs) and interactive proofs (IPs). The prover provides the verifier with auxiliary proofs, and the verifier has oracle access to the messages from the prover.

For a batch of vectors $\mathbf{u} = \{u_0, \ldots, u_l\} \subset \mathbb{F}_q^n$, one can implement a protocol for the RPT problem on each vector to ensure that they are all close to $\mathrm{RS}[\mathbb{F}_q, L, \rho]$. However, this approach is inefficient. [RVW13] provides an approach: randomly choose a vector u' in the span of $\mathbf{u}$ (denoted by $\mathrm{span}(\mathbf{u})$) and check if u' is close to $\mathrm{RS}[\mathbb{F}_q, L, \rho]$. The soundness proof of this method raises an important question: *If* $\exists u_i \in \mathbf{u}$ *that is far from all the members of* $\mathrm{RS}[\mathbb{F}_q, L, \rho]$*, can we prove* u' *is far from* $\mathrm{RS}[\mathbb{F}_q, L, \rho]$ *with high probability?*

Many previous works have explored this question and provided positive answers. This property is referred to as the *proximity gap* for Reed-Solomon codes, as formally defined by Ben-Sasson, Carmon, Ishai, Kopparty, and Saraf in [Ben+20b]. In a more general setting, let $V \subset \mathbb{F}_q^n$ be any linear code and $\delta_V \in [0, 1]$ be its minimal relative distance. Suppose $u_i \in \mathbf{u}$ is δ-far from V, denoted by $\delta(u_i, V) > \delta$. In [Ame+17], Ames, Hazay, Ishai, and Venkitasubramaniam proved that when $\delta < \delta_V/4$, with high probability $u' \in \mathrm{span}(\mathbf{u})$ is δ-far from V. When $u' \in \mathrm{span}(\mathbf{u})$ is on a line, i.e., $\mathbf{u} = \{u_0, u_1\}$, Ben-Sasson, Kopparty, and Saraf [BKS18] demonstrated that when $\delta < 1 - \sqrt[4]{1 - \delta_V}$ (the double Johnson bound), with high probability (related to a small constant ϵ) $u' \in \mathrm{span}(\mathbf{u})$ is $(\delta - \epsilon)$-far from V. Later, Ben-Sasson, Goldberg, Kopparty, and Saraf improved the bound to $1 - \sqrt[3]{1 - \delta_V}$ (the 1.5 Johnson bound) in [Ben+20a]. Furthermore, they showed their result is tight for certain RS codes. Especially, when $V = \mathrm{RS}[\mathbb{F}_q, L, \rho]$, [Ben+20b] bounded the probability that u' is δ-close to V for the unique decoding bound $\delta_V/2$ and the Johnson bound $\sqrt{1 - \delta_V}$ respectively. See Table 1 for details.

1.1 Our results

We present proximity gaps for Reed-Solomon codes within the one-and-a-half Johnson bound. Our result is linear in the length of the code. We begin by considering a simplified case where $\mathbf{u} = \{u_0, u_1\}$. Here, $u' = u_0 + zu_1, z \in \mathbb{F}_q$ is over a line. We have the following result.

Theorem 1 (Informal)**.** *Let* L *be a subset of* $\mathbb{F}_q^\times$*. Let* $u_0, u_1 : L \to \mathbb{F}_q$ *be two functions on*

L. *Let* $\delta, \eta, \rho > 0$ *satisfy* $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ *and* $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. *If there exists* $i \in \{0, 1\}$ *such that* $\delta(u_i, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta$, *then*

$$\mathbb{P}_{z \in \mathbb{F}_q}(\Delta(u_0 + zu_1, \mathrm{RS}[\mathbb{F}_q, L, \rho]) \leq \delta) < \frac{2(1-\rho)|L|}{9\rho\eta^3|\mathbb{F}_q|}. \tag{1}$$

We prove that if either u_0 or u_1 is δ-far from the RS code, then in probability that is linear in the code length, u' is δ-far from the RS code. The formal statement of this theorem is presented in Theorem 4, which describes the result in its contrapositive form. Furthermore, the formal theorem is stronger. We utilize the concept of *correlated agreement*, as defined in [Ben+20b]. A series of functions $u_0, \ldots, u_l$ have δ-correlated agreement with $\mathrm{RS}[\mathbb{F}_q, L, \rho]$ if there exists a sufficiently large subdomain $L' \subseteq L$ and $v_0, \ldots, v_l \in \mathrm{RS}[\mathbb{F}_q, L, \rho]$ such that

$$|L'| \geq (1-\delta)|L| \text{ and } u_i|_{L'} = v_i|_{L'}, 1 \leq i \leq l.$$

The definition of correlated agreement is relevant in the context of real-world protocol applications. Notice that even if all of u_i are δ-close to $\mathrm{RS}[\mathbb{F}_q, L, \rho]$, they may not have δ-correlated agreement. Our formal theorem supports this stronger notion of agreement; specifically, if u_0 and u_1 do not have δ-correlated agreement, (1) holds.

	δ bound	u' distance	Probability	Code
[Ame+17]	$\delta_V/4$	δ	$(\delta+1)/\lvert\mathbb{F}_q\rvert$	Linear code
[BKS18]	$J_\epsilon(J_\epsilon(\delta_V))$	$\delta - \epsilon$	$2/(\epsilon^3\lvert\mathbb{F}_q\rvert)$	Linear code
[Ben+20a]	$1 - \sqrt[3]{1-\delta_V+\epsilon}$	$\delta - \epsilon$	$2/(\epsilon^2\lvert\mathbb{F}_q\rvert)$	Linear code
[Ben+20b]	$1 - \sqrt{\rho} - \epsilon$	δ	$(\rho^2 n^2)/((2\epsilon)^7\lvert\mathbb{F}_q\rvert)$	RS
[Ben+20b]	$(1-\rho)/2$	δ	$n/\lvert\mathbb{F}_q\rvert$	RS
This work	$1 - \sqrt[3]{\rho} - \epsilon$	δ	$2(1-\rho)n/(9\rho\epsilon^3\lvert\mathbb{F}_q\rvert)$	RS

Table 1: When $\mathbf{u} = \{u_0, u_1\}$ and $u_0(u_1)$ is δ-far from the code V, the provable probability that randomly chosen $u' \in \mathrm{span}(\mathbf{u})$ is δ (or $\delta - \epsilon$) close to V. δ bound is the upper bound. δ_V is the minimal relative distance of V. The latter three rows focus on RS codes with rate ρ and $\delta_V = 1 - \rho$. $J_\epsilon(\delta_V) = 1 - \sqrt{1 - \delta_V(1-\epsilon)}$ is the Johnson bound.

Table 1 compares our result with previous results. [Ben+20b] provides the linear proximity gap under the unique decoding bound $(1-\rho)/2$ and the quadratic proximity gap under the Johnson bound $1 - \sqrt{\rho}$. When the one-and-a-half Johnson bound $1 - \sqrt[3]{\rho}$ is better than the unique decoding bound (related to ρ), we improve the provable proximity gap to linear. Additionally, [Ben+20b] conjectures that we can prove the proximity gap when $\delta \leq 1 - \rho$. We will briefly introduce this conjecture in Section 5.3. Figure 1 compares various bounds. When $\rho \leq 1/8$, the one-and-a-half Johnson bound is better than the unique decoding bound. And we make improvements in this case.

We prove our result under the generalized case that $\mathbf{u} = \{u_0, \ldots, u_l\}$.

Theorem 2 (Informal)**.** *Let* L *be a subset of* $\mathbb{F}_q^\times$. *Let* $u_0, \cdots, u_l : L \to \mathbb{F}_q, l \geq 1$ *be a sequence of functions on* L. *Let* $\delta, \eta, \rho > 0$ *satisfy* $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ *and* $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. *Denote* $V = \mathrm{RS}[\mathbb{F}_q, L, \rho]$ *and*

$$S = \{\mathbf{z}_l = \langle z_1, \cdots, z_l\rangle \in \mathbb{F}_q^l : \Delta(u_0 + z_1u_1 + \cdots + z_lu_l, V) \leq \delta|L|\}.$$

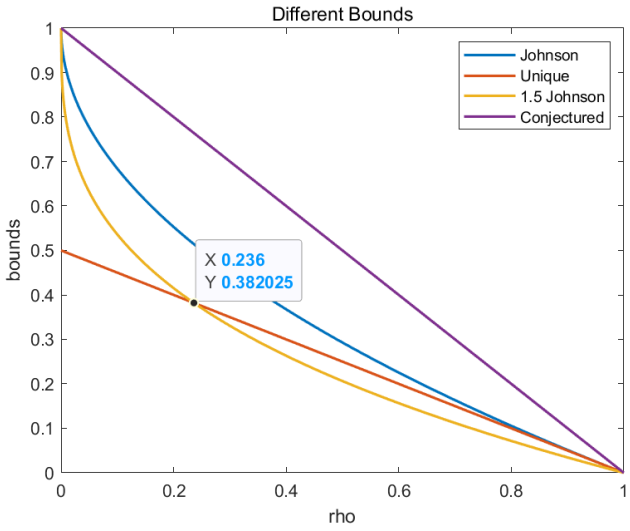

Figure 1: Bounds for RS codes

If $\exists u_i \in \mathbf{u}$ *such that* $\delta(u_i, V) > \delta$*, then*

$$\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S) < \left(\frac{2(1-\rho)|L|}{9\rho\eta^3|\mathbb{F}_q|}\right) \cdot l.$$

This result also supports the correlated agreement version. The formal theorem can be found in Theorem 6.

1.2 Applications

Proximity gaps for RS codes provide provable soundness for a variety of protocols. For example, the Fast RS IOPP (known as FRI) [Ben+18] is a widely used IOPP for RS codes due to its high efficiency. FRI is implemented as a subprotocol in many recent (zk)SNARKs and real-world systems [Ben+19][KPV22][Sta23][Pol][Zha+20][Xie+22].

Our result can be implied to prove the soundness of FRI. Previously, Ben-Sasson, Carmon, Ishai, Kopparty, and Saraf established the best provable soundness of FRI in [Ben+20b], utilizing elegant mathematical techniques. They proved the soundness error bound of FRI is

$$\epsilon_{\text{FRI}} \leq \max\left\{O\left(\frac{\rho^2}{\eta^7} \cdot \frac{|L|^2}{|\mathbb{F}_q|}\right), (1-\delta)^t\right\} \tag{2}$$

when $\delta \leq 1 - \sqrt{\rho} - \eta$. Let t represent the iteration time during the QUERY phase in FRI. It is important to note that the first term is a constant dependent on the parameters. For small values of t, the second term dominates the inequality. Furthermore, this term decreases as t increases. Consequently, when t becomes sufficiently large, the first term establishes a provable upper bound on the soundness error of FRI. We will introduce the protocol in detail in Section 6. We provide an alternative soundness error bound of FRI:

$$\epsilon_{\text{FRI}} \leq \max\left\{O\left(\frac{1}{\rho\eta^3} \cdot \frac{|L|}{|\mathbb{F}_q|}\right), (1-\delta)^t\right\} \tag{3}$$

when $\delta \leq 1 - \sqrt[3]{\rho} - \eta$. When t is large, the first term dominates the soundness error bound. In practical applications, $|L|$ is large and significantly influences the soundness error bound. Consequently, our bound indicates that FRI can provide enhanced security. However, when t is small, the previous bound is more advantageous. In practical applications, we can select the minimum of these two bounds.

Our result and the provable soundness in [Ben+20b] fit the correlated agreement condition. The correlated agreement of FRI is used to prove the (knowledge) soundness of protocols that use FRI as a sub-protocol [Sta23], as well as to prove the round-by-round soundness of FRI [Sta23][Blo+23]. There are also a variety of generalized protocols of FRI [ZCF23][Arn+24]. These works build upon the soundness of FRI, and improvements in the soundness of FRI can also be applied to these subsequent works.

2 Technical overview

In this section, we outline the overall idea behind our proof of the main result, Theorem 3. We take a new approach to address this problem through combinatorial methods. This section introduces the tools we employ and explains how this can be done. The main theorem is articulated in the context of the folding operation within the FRI framework. Folding and proximity gaps can be transformed into each other. We will introduce the transformation in Section 5. For now, we will concentrate on the concept of folding.

2.1 Polynomial folding and function folding

Let $\mathbb{F}_q$ be a finite field and $L \subseteq \mathbb{F}_q$ be an evaluation domain, with $|L| = n$. Denote $\rho \in (0, 1]$ as the rate. The notation $\mathrm{RS}[\mathbb{F}_q, L, \rho]$ represents the set of code words $p : L \to \mathbb{F}_q$ that are evaluation results of polynomials of degree strictly less than $\rho|L|$. For a given code word $f : L \to \mathbb{F}_q$, we want to know whether f is a member of $\mathrm{RS}[\mathbb{F}_q, L, \rho]$. This verification is called the *low degree test(LDT)* of f.

We can use FFT to check directly. However, the time complexity of FFT is $O(n \log n)$, which is unacceptable in many real-world applications. The FRI protocol[Ben+18] is an interactive oracle proof(IOP) to achieve this goal with time complexity $O(\log n)$. An untrusted prover may help us to complete the verification and the time complexity of an honest prover is $O(n)$.

The *folding* operation plays an important role in the FRI protocol. It can fold a polynomial into half its degree. So we can fold a polynomial of degree n into a constant after $\log_2(n)$ rounds of folding. More precisely, let $p(X) \in \mathbb{F}_q[X]$ be a polynomial with degree d and suppose

$$p(X) = c_0 + c_1 X + c_2 X^2 + \ldots + c_d X^d.$$

We can divide $p(X)$ into even and odd parts, i.e.,

$$p^{(\mathrm{even})}(X) = \sum_{\text{even } i} c_i X^i \text{ and } p^{(\mathrm{odd})}(X) = \sum_{\text{odd } i} c_i X^{i-1},$$

and we have

$$p(X) = p^{(\mathrm{even})}(X) + X p^{(\mathrm{odd})}(X).$$

Let $\alpha \in \mathbb{F}_q$ be a randomly chosen folding point, then define the folding result of $p(X)$ at folding point α, denoted by $\text{PolyFold}_\alpha(p)$, to be the following polynomial

$$\text{PolyFold}_\alpha(p) = \begin{cases} (c_0 + \alpha c_1) + (c_2 + \alpha c_3)X + \ldots + (c_{d-1} + \alpha c_d)X^{\frac{d-1}{2}} & \text{if } d \text{ is odd} \\ (c_0 + \alpha c_1) + (c_2 + \alpha c_3)X + \ldots + (c_{d-2} + \alpha c_{d-1})X^{\frac{d-2}{2}} + c_d X^{\frac{d}{2}} & \text{if } d \text{ is even} \end{cases}$$

of degree $\lfloor \frac{d}{2} \rfloor$. Notice that we have

$$\text{PolyFold}_\alpha(p)(X^2) = p^{(\text{even})}(X) + \alpha p^{(\text{odd})}(X).$$

Furthermore, for any $x^2 \in L^2$, we can prove

$$\text{PolyFold}_\alpha(p)(x^2) = \frac{p(x) + p(-x)}{2} + \alpha \cdot \frac{p(x) - p(-x)}{2x}$$

by calculating directly.

Recall that our goal is to verify whether a given function $f : L \to \mathbb{F}_q$ is a low-degree polynomial. As a result, the folding operation needs to be defined on the functions. Especially, the folding operation works on functions over evaluation domains with pairing elements, i.e. if $x \in L$, we have $-x \in L$. This is because of the folding structure, which will be explained later.

Suppose $L = \{x_1, -x_1, x_2, -x_2, \ldots, x_{\frac{n}{2}}, -x_{\frac{n}{2}}\}$ is an evaluation domain with pairing elements and $L^2 = \{x_1^2, x_2^2, \ldots, x_{\frac{n}{2}}^2\}$. Let $\alpha \in \mathbb{F}_q$ be a randomly chosen folding point. The folding result of f, denoted as $\text{FuncFold}_\alpha(f)$, is a function on L^2. More precisely, suppose we have a function

$$f = \{f(x_1), f(-x_1), f(x_2), f(-x_2), \ldots, f(x_{\frac{n}{2}}), f(-x_{\frac{n}{2}})\}.$$

Define $\text{FuncFold}_\alpha(f) : L^2 \to \mathbb{F}_q$ as follows:

$$\text{FuncFold}_\alpha(f)(x^2) = \frac{f(x) + f(-x)}{2} + \alpha \cdot \frac{f(x) - f(-x)}{2x}$$

for any $x \in L$. For example, in the finite field $\mathbb{F}_{17}$, let $L = \{1, -1, 4, -4, 2, -2, 8, -8\}$ be an evaluation domain with eight elements. Then we have $L^2 = \{1, -1, 4, -4\}$. Let $\alpha \in \mathbb{F}_q$ be the folding point. For a given function $f = \{0, -1, 4, -4, 2, -2, 8, -8\}$, the folding result of f at point α is $\text{FuncFold}_\alpha(f) = \{-9 + 9\alpha, \alpha, \alpha, \alpha\}$. Figure 2 shows the folding result.

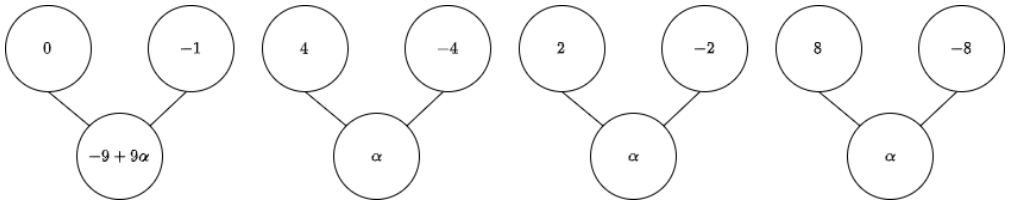

Figure 2: Function folding result of f at point α. f is on the first layer and $\text{FuncFold}_\alpha(f)$ is on the second layer. Each element on the second layer is calculated from its parents.

Furthermore, we have the following properties:

1. For any $\alpha \in \mathbb{F}_q$, the length of $\text{FuncFold}_\alpha(f)$ is half of f.

2. When $f \in \mathrm{RS}[\mathbb{F}_q, L, \rho]$, i.e., f agrees with some low-degree polynomial p, we have

$$\mathrm{FuncFold}_\alpha(f) = \mathrm{PolyFold}_\alpha(p)$$

for any $\alpha \in \mathbb{F}_q$. This will be proved in Proposition 2.

3. For any $\alpha \in \mathbb{F}_q$ and $x^2 \in L^2$, $\mathrm{FuncFold}_\alpha(f)(x^2)$ only depends on $f(x)$ and $f(-x)$.

Condition 2 implies when $f \in \mathrm{RS}[\mathbb{F}_q, L, \rho]$, we have $\mathrm{FuncFold}_\alpha(f) \in \mathrm{RS}[\mathbb{F}_q, L^2, \rho]$. Suppose L^2 is still an evaluation domain with pairing elements, we can do function folding on $\mathrm{FuncFold}_\alpha(f)$ again. We say L is smooth if $L, L^2, L^4, \ldots$ are all with pairing elements. Thus, we can recursively fold a function on L until it becomes a constant. Condition 3 says the function folding operation is local. If we want to check the accuracy of the folding on any location $x^2 \in L^2$, we only need to know $f(x)$ and $f(-x)$. Furthermore, for some $f \in \mathrm{RS}[\mathbb{F}_q, L, \rho]$, we can do the folding operation on it even without knowing the polynomial coefficients.

Since the folding result of a low-degree polynomial is still a low-degree polynomial, we can fold a polynomial of degree d into a constant after $\lceil \log_2 d \rceil$ rounds. It is easy to check a constant in an evaluation domain. However, we can not claim a function is a low-degree polynomial even if it can be folded into a constant after some limited number of rounds. This is because there exist some *bad* folding points that will disclose the relative distance between the given function f and $\mathrm{RS}[\mathbb{F}_q, L, \rho]$. Set $\rho = \frac{1}{4}$ in the above example. Then $\mathrm{RS}[\mathbb{F}_{17}, L, \frac{1}{4}]$ is the set of polynomials with degree $< \frac{1}{4} \cdot 8 = 2$. It is easy to verify that f is not a member of $\mathrm{RS}[\mathbb{F}_{17}, L, \frac{1}{4}]$. The closest code word of f in $\mathrm{RS}[\mathbb{F}_{17}, L, \frac{1}{4}]$ is $p(X) = X$ and the relative distance between f and $\mathrm{RS}[\mathbb{F}_{17}, L, \frac{1}{4}]$ is $\frac{1}{8}$. When $\alpha = -1$, the folding result is $\mathrm{FuncFold}_{-1}(f) = \{-1, -1, -1, -1\}$. This is a constant and is a member of $\mathrm{RS}[\mathbb{F}_{17}, L^2, \frac{1}{4}]$. Then $\alpha = -1$ is a bad folding point in this example. For a given relative distance $\delta > 0$ and a function f that is δ-far from the RS code. Define the bad folding points as folding points whose folding results are δ-close to the RS code. Our goal is to prove that the number of bad folding points is limited. This is called the proximity gap for Reed-Solomon codes, defined in [Ben+20b].

2.2 Partition the bad folding points into blocks

For a fixed $0 < \delta < 1$ and a function f that is δ-far from the RS code, define the set of bad folding points to be

$$\mathrm{Bad}(f) = \{\alpha \in \mathbb{F}_q \mid \mathrm{FuncFold}_\alpha(f) \text{ is } \delta\text{-close to the RS code}\}.$$

Many previous studies focus on the folding results and use the list-decoding skill to restrict the number of bad folding points. We provide a new approach to deal with this problem. Instead of the folding result, we focus on the origin function f. We pay attention to some sub-evaluation domains of f and transform the problem into a combinatorics problem.

More precisely, let $\alpha \in \mathrm{Bad}(f)$ be a bad folding point and $p_\alpha \in \mathrm{RS}[\mathbb{F}_q, L, \rho]$ is the closest code word of $\mathrm{FuncFold}_\alpha(f)$. The following sub-evaluation domains are related to α :

$$C_\alpha^* = \{x^2 \in L^2 \mid \mathrm{FuncFold}_\alpha(f)(x^2) = p_\alpha(x^2)\}$$

and

$$P_\alpha^* = \{x \in L \mid x^2 \in C_\alpha^*\} \text{ i.e., } P_\alpha^* \text{ is the parent set of } C_\alpha^*.$$

For two distinct bad folding points $\alpha, \beta \in \mathrm{Bad}(f)$, we prove the following lemma:

Lemma 1. *[Informal] For any distinct $\alpha, \beta \in \mathrm{Bad}(f)$, there exists a polynomial $p(X) \in \mathbb{F}_q[X]$ of degree strictly less than $\rho|L|$ such that*

$$f|_{P^*_\alpha \cap P^*_\beta} = p|_{P^*_\alpha \cap P^*_\beta},$$

that is, $f(x) = p(x)$ for any $x \in P^*_\alpha \cap P^*_\beta$.

Figure 3 shows a simple example. In the finite field $\mathbb{F}_{17}$, let $L = \{1, -1, 4, -4, 2, -2, 8, -8\}$ and $L^2 = \{1, -1, 4, -4\}$. Let $\rho = \frac{1}{4}$, so the member of $\mathrm{RS}[\mathbb{F}_{17}, L^2, \frac{1}{4}]$ are constants. For a given function $f = \{0, 2, 4, -4, 2, -2, 5, 7\}$. The folding result of f at point $\alpha = 9$ is $\{9, 9, 9, 7\}$. Thus, we have $C^*_\alpha = \{1, -1, 4\}$ and $P^*_\alpha = \{1, -1, 4, -4, 2, -2\}$. Similarly, the folding result of f at point $\beta = 11$ is $\{7, 11, 11, 11\}$ and $C^*_\beta = \{-1, 4, -4\}, P^*_\beta = \{4, -4, 2, -2, 8, -8\}$. As a result, $P^*_\alpha \cap P^*_\beta = \{4, -4, 2, -2\}$. Lemma 4 says $f|_{P^*_\alpha \cap P^*_\beta}$ agrees with a low-degree polynomial p of degree $< \rho|L| = 2$. It is easy to find that $p(X) = X$ in our example.

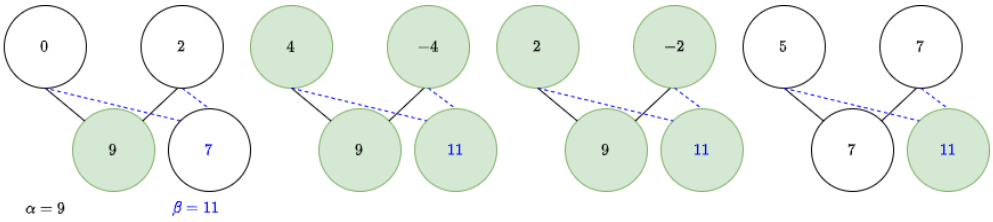

Figure 3: Folding results of f at folding point $\alpha = 9$(black solid line) and $\beta = 11$(blue dotted line). C^*_α and C^*_β are filled in green on the second layer. $P^*_\alpha \cap P^*_\beta$ is filled in green on the first layer.

Because of the use of Corrádi's lemma, which will be introduced later, we extract a series of subsections of $P^*_\alpha, C^*_\alpha, \alpha \in \mathrm{Bad}(f)$ such that these subsections have the same size. For a given $0 < \delta < 1$, notice that $|P^*_\alpha| \geq (1-\delta)|L|$ and $|C^*_\alpha| \geq (1-\delta)|L^2|$ for $\alpha \in \mathrm{Bad}(f)$. We define

- C_α: The set of first $(1-\delta)|L^2|$ elements of C^*_α, i.e., $|C_\alpha| = (1-\delta)|L^2|$.
- $P_\alpha = \{x : x^2 \in C_\alpha\}$. That is, P_α is the "parent set" of C_α.

The above definitions make sense since the elements in the evaluation domain L are in order. Notice that $P_\alpha \cap P_\beta \subseteq P^*_\alpha \cap P^*_\beta$, Lemma 1 still holds on $P_\alpha \cap P_\beta$. Using these definitions, we can partition $\mathrm{Bad}(f)$ into blocks.

Our partition is based on some *long* low-degree polynomials in f. More precisely, let $0 < \xi \leq 1$ and $D \subseteq L$ satisfying $|D| \geq \xi|L|$. If $f|_D$ agrees with some low-degree polynomial p, then we say p is a long low-degree polynomial contained in f. There may be many such long low-degree polynomials $p_1, \ldots, p_s$, denote by $D_i, 1 \leq i \leq s$ the maximal agree domains of $p_i, 1 \leq i \leq s$ and f, i.e., $f|_{D_i} = p_i|_{D_i}$. We partition the set of bad folding points $\mathrm{Bad}(f)$ based on Algorithm 1.

Let $\{A_1, \ldots, A_r\}$ be the output blocks of the algorithm and $\{\alpha_1, \ldots, \alpha_r\}$ be the corresponding represent elements. For a block $A \in \{A_1, \ldots, A_r\}$ and its represent element α, we have $|P_\beta \cap P_\alpha| \geq \xi|L|$. Lemma 1 implies a long low-degree polynomial in f is contained in both P_α and P_β. On the other hand, for distinct represent elements $\alpha_i, \alpha_j \in \{\alpha_1, \ldots, \alpha_r\}$, we have $|P_{\alpha_i} \cap P_{\alpha_j}| < \xi|L|$.

The number of blocks is limited. Let $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\xi = (1-\delta)^2 - \eta'$, where $\eta' = \frac{3\sqrt[3]{\rho}\eta^2}{2}$, we prove $r \leq \frac{1-\rho}{\eta'}$. Focusing on the represent elements $\{\alpha_1, \ldots, \alpha_r\}$, we have

Algorithm 1 Partition Bad Folding Points(Informal)

input: $\mathrm{Bad}(f)$
initialization: $r = 0$
set $X^* = \mathrm{Bad}(f)$
while $X^* \neq \emptyset$ **do**
  $r = r + 1$
  pick an arbitrary $x \in X^*$ and let $\alpha_i = x$
  let $A_i = \{\beta \in \mathrm{Bad}(f) : |P_\beta \cap P_{\alpha_i}| \geq \xi|L|\}$
  $X^* = X^* - A_i$
end while
return $A_1, \ldots, A_r$ and $\alpha_1, \ldots, \alpha_r$

- $|P_{\alpha_i}| = (1-\delta)|L|, 1 \leq i \leq r$;
- $|P_{\alpha_i} \cap P_{\alpha_j}| < \xi|L| = \left((1-\delta)^2 - \eta'\right)|L|, 1 \leq i < j \leq r$.

Notice that $\bigcup_{i=1}^{r} P_{\alpha_i} \subseteq L$. Using the following lemma by Corrádi, we can restrict the number of blocks.

Lemma 2 (Corrádi 1969 [Juk11]). *Let $P_1, \ldots, P_r$ be s-element sets. If $|P_i \cap P_j| \leq k$ for any distinct $i, j \in \{1, 2, \ldots, r\}$, then*

$$\left|\bigcup_{i=1}^{r} P_i\right| \geq \frac{s^2 r}{s + (r-1)k}.$$

Calculating directly, we have

$$r < \frac{1-\rho}{\eta'}.$$

The upper bound of the number of blocks is a constant and is independent of $|L|$, the length of the code word.

2.3 Partition the blocks into equivalence classes

We further restrict the number of elements in each block. Let $A \in \{A_1, \ldots, A_r\}$ be a block and α be the corresponding represent element, i.e., for all $\beta \in A$, we have $|P_\alpha \cap P_\beta| \geq \xi|L|$. Lemma 1 implies a long low-degree polynomial in f is contained in both P_α and P_β. Denote by $p_1, \ldots, p_s$ all the long low-degree polynomials contained in P_α, i.e., $\exists D_i \subseteq P_\alpha$ such that $|D_i| \geq \xi|L|$ and $f|_{D_i} = p_i|_{D_i}$. Furthermore, for all $\beta \in A \setminus \{\alpha\}$, we can find one and only one low-degree polynomial $p \in \{p_1, \ldots, p_s\}$ such that $f|_{P_\alpha \cap P_\beta} = p|_{P_\alpha \cap P_\beta}$. We define an equivalence relation $\mathcal{R}$ on $A \setminus \{\alpha\}$:

$$(\beta_1, \beta_2) \in \mathcal{R} \iff \text{The low-degree polynomials decided by } P_\alpha \cap P_{\beta_1} \text{ and } P_\alpha \cap P_{\beta_2} \text{ are the same.}$$

The number of equivalence classes is limited. We restrict the number of long low-degree polynomials $p_1, \ldots, p_s$ contained in P_α. We have

- $|D_i| \geq \xi|L|, 1 \leq i \leq s$ according to the definition of p_i.
- $|D_i \cap D_j| \leq \rho|L|, 1 \leq i < j \leq s$ because p_i, p_j are distinct polynomials with degree $< \rho|L|$.

- $|D_i| \leq (1-\delta)|L|$ since f is δ-far from $\mathrm{RS}[\mathbb{F}_q, L, \rho]$.

The first two conditions are similar to the condition of Corrádi's lemma(Lemma 2). But we have $|D_i| \geq \xi|L|$ instead of $|D_i| = \xi|L|$ in this case. As a result, we use the third condition and follow the proof of Corrádi's lemma to prove

$$s \leq \frac{1}{3\eta\rho^{\frac{2}{3}}}$$

when $\delta \leq 1 - \sqrt[3]{\rho} - \eta, \xi = (1-\delta)^2 - \eta'$ and $\eta' = \frac{3\sqrt[3]{\rho}\eta^2}{2}$. The upper bound of the number of equivalence classes is a constant and is independent of $|L|$, the length of the code word.

The size of each equivalence class is bounded. For any equivalence class $\{\beta_1, \ldots, \beta_t\}$. Suppose p is the low-degree polynomial related to the equivalence class and D is the maximal agree domain of p, i.e., $f|_D = p|_D$. Since f is δ-far from $\mathrm{RS}[\mathbb{F}_q, L, \rho]$, $|D| < (1-\delta)|L|$. On the other hand, for any $\beta_i, 1 \leq i \leq t$ in the equivalence class, we have $|P_{\beta_i}| = (1-\delta)|L|$ according to the definition of P_{β_i}. As a result, $|P_{\beta_i} \setminus D| \geq 1$. We prove that

$$(P_{\beta_i} \setminus D) \cap \left(P_{\beta_j} \setminus D\right) = \emptyset, i \neq j.$$

For any $x \in P_{\beta_i} \setminus D$, we have $f(x) \neq p(x)$. (Remember that $x \neq 0$ since $P_{\beta_i} \subseteq L$ and L is with pairing elements.) According to the definition of P_{β_i}, we have

$$\mathrm{FuncFold}_{\beta_i}(f)(x) = \mathrm{PolyFold}_{\beta_i}(p)(x).$$

The above equation is equivalent to

$$\frac{f(x)+f(-x)}{2} + \beta_i \cdot \frac{f(x)-f(-x)}{2x} = \frac{p(x)+p(-x)}{2} + \beta_i \cdot \frac{p(x)-p(-x)}{2x}$$

according to the definition and property of folding. If $\frac{f(x)-f(-x)}{2x} = \frac{p(x)-p(-x)}{2x}$, then we have $\frac{f(x)+f(-x)}{2} = \frac{p(x)+p(-x)}{2}$ to make the above equation hold. However, this implies $f(x) = p(x)$. A contradiction. As a result, we have $\frac{f(x)-f(-x)}{2x} \neq \frac{p(x)-p(-x)}{2x}$ and we can transform the equation into

$$\beta_i = x \cdot \frac{(p(x)+p(-x)) - (f(x)+f(-x))}{(f(x)-f(-x)) - (p(x)-p(-x))}.$$

Notice that once x is fixed, the right side of the above equation is fixed. Since β_j, β_i are distinct, $x \notin P_{\beta_j} \setminus D$.

Thus,

$$|L \setminus D| \geq \left|\bigcup_{i=1}^{t} (P_{\beta_i} \setminus D)\right| = \sum_{i=1}^{t} |P_{\beta_i} \setminus D| \geq t.$$

The size of each equivalence class is bounded by $|L \setminus D|$. We can further optimize the bound to be $\frac{|L \setminus D|}{2}$, see Corollary 5 for details.

Combining all these above, we finish our proof of the main theorem, i.e., the number of bad folding points is linear in the length of the code word.

3 Preliminaries

We present the preliminaries and notations. We use $\mathbb{F}_q$ to denote a finite field with q elements. Denote by $\mathbb{F}_q^\times = \mathbb{F}_q \setminus \{0\}$ the cyclic group.

3.1 Reed-Solomon codes

Let $L \subseteq \mathbb{F}_q$ be a subset. Let $\mathrm{RS}[\mathbb{F}_q, L, \rho] : \mathbb{F}_q^{d+1} \to \mathbb{F}_q^{|L|}$ denote the Reed-Solomon code of degree strictly less than $\rho|L|$ evaluated on L, where $\mathrm{RS}[\mathbb{F}_q, L, \rho]$ maps $(c_0, c_1, \ldots, c_d) \in \mathbb{F}_q^{d+1}$ to $\left(\sum_{i=0}^{k} c_i x^i\right)_{x \in L} \in \mathbb{F}_q^{|L|}$, and $d = \lceil \rho|L| \rceil - 1$. **Throughout the paper, we assume $\rho|L|/2$ is an integer.** Then $\mathrm{RS}[\mathbb{F}_q, L, \rho]$ has code rate ρ.

Let $f, g : L \to \mathbb{F}$ be two codewords. The distance between f and g is defined as

$$\Delta(f, g) = |\{x \in L : f(x) \neq g(x)\}|.$$

The distance between f and $\mathrm{RS}[\mathbb{F}_q, L, \rho]$ is defined as

$$\Delta(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) = \min_{g \in \mathrm{RS}[\mathbb{F}_q, L, \rho]} \Delta(f, g).$$

3.2 Polynomial folding and function folding

We define an operation called *folding* in this subsection, which is used in FRI.

Definition 1 (polynomial folding)**.** *Let $p(X) \in \mathbb{F}_q[X]$ be a polynomial of degree d, that is,*

$$p(X) = c_0 + c_1 X + \ldots + c_d X^d.$$

Define the folding of polynomial $p(X)$ at point $\alpha \in \mathbb{F}_q$*, denoted by* $\mathrm{PolyFold}_\alpha(p)$*, to be the following polynomial*

$$\begin{cases} (c_0 + \alpha c_1) + (c_2 + \alpha c_3)X + \ldots + (c_{d-1} + \alpha c_d)X^{\frac{d-1}{2}} & \text{if } d \text{ is odd} \\ (c_0 + \alpha c_1) + (c_2 + \alpha c_3)X + \ldots + (c_{d-2} + \alpha c_{d-1})X^{\frac{d-2}{2}} + c_d X^{\frac{d}{2}} & \text{if } d \text{ is even} \end{cases}$$

of degree $\lfloor \frac{d}{2} \rfloor$.

The following proposition says PolyFold is a linear operator, which is straightforward to prove.

Proposition 1. *For any $p(X), q(X) \in \mathbb{F}_q[X]$, and any $\alpha, \beta, \gamma \in \mathbb{F}_q$ we have*

$$\mathrm{PolyFold}_\alpha(\beta p + \gamma q) = \beta \mathrm{PolyFold}_\alpha(p) + \gamma \mathrm{PolyFold}_\alpha(q).$$

Definition 2. *Let $p(X) \in \mathbb{F}_q[X]$ be a polynomial, where $p(X) = \sum_{i=0}^{d} c_i X^i$. Let*

$$p^{(\mathrm{even})}(X) = \sum_{even\ i} c_i X^i \text{ and } p^{(\mathrm{odd})}(X) = \sum_{odd\ i} c_i X^{i-1}.$$

Both $p^{(\mathrm{even})}(X)$ and $p^{(\mathrm{odd})}(X)$ are even functions, and $p(X) = p^{(\mathrm{even})}(X) + X p^{(\mathrm{odd})}(X)$. One can easily verify that

$$\mathrm{PolyFold}_\alpha(p)(x^2) = p^{(\mathrm{even})}(x) + \alpha p^{(\mathrm{odd})}(x) \tag{4}$$

for any $x \in \mathbb{F}_q$.

Codeword folding is defined on the codewords over evaluation domains *with pairing elements* defined as follows:

Definition 3. *Let $L \subseteq \mathbb{F}_q^\times$ be a subset.*

- *Say L is* with pairing elements *if for all $x \in L$, we have $-x \in L$.*
- *Say L is* smooth *if L is a coset of a multiplicative group whose order is a power of 2.*

Notice that a smooth set must be with pairing elements. Suppose L is a set with pairing elements, we define $L^2 = \{x^2 : x \in L\}$. And we have $|L^2| = |L|/2$.

Definition 4 (codeword folding)**.** *Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $f : L \to \mathbb{F}_q$ be any codeword (function). We define the* folding of f at point α*, denoted by* $\text{FuncFold}_\alpha(f) : L^2 \to \mathbb{F}_q$*, as follows*

$$\text{FuncFold}_\alpha(f)(x^2) = \frac{f(x) + f(-x)}{2} + \alpha \cdot \frac{f(x) - f(-x)}{2x}$$

for any $x \in L$.

Proposition 2. *Let $p(X) \in \mathbb{F}_q[X]$ be a polynomial, and let $\alpha \in \mathbb{F}_q$. We have*

$$\text{PolyFold}_\alpha(p)(X^2) = \frac{p(X) + p(-X)}{2} + \alpha \cdot \frac{p(X) - p(-X)}{2X}. \tag{5}$$

Proof. By the linearity of PolyFold operator (Prop 1), it suffices to prove (5) when p is a monomial.

If $p(X) = X^d$, where d is even, we have $\text{PolyFold}_\alpha(p) = X^{d/2}$. Thus,

$$\text{PolyFold}_\alpha(p)(X^2) = X^d.$$

The right-hand side of (5) is also X^d.

If $p(X) = X^d$, where d is odd, we have $\text{PolyFold}_\alpha(p) = \alpha X^{(d-1)/2}$. Thus, $\text{PolyFold}_\alpha(p)(X^2) = \alpha X^{d-1}$. The right-hand side of (5) is

$$\alpha \cdot \frac{X^d - (-X)^d}{2X} = \alpha X^{d-1} = \text{PolyFold}_\alpha(p)(X^2).$$

□

Remark 1. *When the characteristic of the finite field $\text{char}(\mathbb{F}_q) = 2$, the polynomial and function folding structures are different from Definition 1 and Definition 4. This is because $q(X) = X^2$ is no longer a 2 to 1 map. But we can still prove the corresponding Proposition 2 under this situation. Further details can be found in [Ben+18] and [BKS18].*

3.3 Correlated agreement

For a series of functions $f_0, \ldots, f_t$, we not only require them to be close to a codeword set V individually but also to share a common large agreement domain. The property that such a domain exists is called *correlated agreement.*

Definition 5. *[Ben+20b] Let $L \subseteq \mathbb{F}_q$ be a subset. Let $f_0, \ldots, f_t : L \to \mathbb{F}_q$ be a sequence of functions. Let V be a set of codewords. Let $0 < \delta \leq 1$. If there exists a subdomain $L' \subseteq L$ and $v_0, \ldots, v_t \in V$ satisfying*

- ***Density:*** *$|L'|/|L| \geq 1 - \delta$, and*
- ***Agreement:*** *for all $i \in \{0, \ldots, t\}$, the functions f_i and v_i agree on L'.*

Then we say $f_0^{(0)}, \ldots, f_t^{(0)}$ have correlated agreement with V on L

4 Main proof

Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. $f : L \to \mathbb{F}_q$ is a codeword δ-far from RS codes, where δ is under the one-and-a-half Johnson bound $1 - \sqrt[3]{\rho}$. Let $\eta > 0$ be the gap between δ and the double Johnson bound, i.e., $\delta \leq 1 - \sqrt[3]{\rho} - \eta$. Our goal is to prove that with high probability, the folded codeword $\text{FuncFold}_\alpha(f)$ is still δ-far from RS codes, where $\alpha \in \mathbb{F}_q$ is chosen uniformly at random. Our motivation is to analyze the soundness of the FRI protocol.

Definition 6 (bad folding points)**.** *Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $0 < \delta < 1 - \rho$. Let $f : L \to \mathbb{F}_q$ be a codeword such that $\Delta(f, \text{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$. Define the* bad folding points *to be*

$$\text{Bad}(f) = \left\{\alpha \in \mathbb{F}_q : \Delta\left(\text{FuncFold}_\alpha(f), \text{RS}[\mathbb{F}_q, L^2, \rho]\right) \leq \delta|L^2|)\right\}.$$

Our main theorem limits the number of bad folding points when δ is within the Johnson bound.

Theorem 3. *Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt[3]{\rho} - \eta$. Let $f : L \to \mathbb{F}_q$ be a codeword such that $\Delta(f, \text{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$. Then,*

$$|\text{Bad}(f)| < \frac{(1 - \sqrt[3]{\rho})(1 - \rho)}{9\rho\eta^3}|L| + \frac{2(1 - \rho)}{3\rho^{\frac{1}{3}}\eta^2}.$$

Corollary 1. *Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. Let $f : L \to \mathbb{F}_q$ be a codeword such that $\Delta(f, \text{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$. Then,*

$$|\text{Bad}(f)| < \frac{(1 - \rho)|L|}{9\rho\eta^3}.$$

Proof. Since L is a set with pairing elements, we have $|L| \geq 2$. Thus, $|L| \geq 6\rho^{\frac{1}{3}}\eta$. As a result, we have

$$\begin{aligned}
|\text{Bad}(f)| &< \frac{(1 - \sqrt[3]{\rho})(1 - \rho)}{9\rho\eta^3}|L| + \frac{2(1 - \rho)}{3\rho^{\frac{1}{3}}\eta^2} && \text{by Theorem 3} \\
&= \frac{(1 - \rho)\left(|L| - \rho^{\frac{1}{3}}(|L| - 6\rho^{\frac{1}{3}}\eta)\right)}{9\rho\eta^3} \\
&\leq \frac{(1 - \rho)|L|}{9\rho\eta^3}.
\end{aligned}$$

□

To prove the main theorem, we first introduce a few definitions.

Definition 7. *Fix an arbitrary order for all the elements in the finite field $\mathbb{F}_q$. Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $f : L \to \mathbb{F}_q$ be a codeword. Let $\alpha \in \text{Bad}(f)$.*

- *Let $\text{Closest}(f, \text{RS}[\mathbb{F}_q, L, \rho]) \in \text{RS}[\mathbb{F}_q, L, \rho]$ denote the closest codeword (polynomial). If there are more than one codewords with the same minimal distance, choose the one with the smallest lexicographical order.*

- $C_\alpha^* = \{x \in L^2 : \text{FuncFold}_\alpha(f)(x) = \text{Closest}(\text{FuncFold}_\alpha(f), \text{RS}[\mathbb{F}_q, L, \rho])(x)\}$. *That is, C_α is the set of evaluations points where* $\text{FuncFold}_\alpha(f)$ *agrees with* $\text{Closest}(f, \text{RS}[\mathbb{F}_q, L, \rho])$.
- $P_\alpha^* = \{x : x^2 \in C_\alpha^*\}$. *That is, P_α^* is the "parent set" of C_α^*.*
- C_α*: The set of first* $(1-\delta)|L^2|$ *elements of* C_α^**, i.e.,* $|C_\alpha| = (1-\delta)|L^2|$.
- $P_\alpha = \{x : x^2 \in C_\alpha\}$. *That is, P_α is the "parent set" of C_α.*

Poof overview: The overall strategy is to partition the bad points $\text{Bad}(f) \subseteq \mathbb{F}_q$ into r subsets, denoted by $A_1, A_2, \ldots, A_r$, with representatives $\alpha_1 \in A_1, \alpha_2 \in A_2, \ldots, \alpha_r \in A_r$ such that r is bounded. To put restriction on $|A_i|$, we provide an equivalent relation on $A_i \setminus \{\alpha_i\}$. Denote by s the number of equivalence classes and t the number of elements in an equivalence class. We put upper bounds on s and t respectively. Thus, we restrict the number of elements in the set of bad folding points.

We use a greedy algorithm to find the partition of $\text{Bad}(f)$ and the representatives.

Algorithm 2 Partition Bad Folding Points

```
input: Bad(f)
initialization: r = 0
set X* = Bad(f)
while X* ≠ ∅ do
  r = r + 1
  pick an arbitrary x ∈ X* and let α_i = x
  let A_i = {β ∈ Bad(f) : |P_β ∩ P_{α_i}| ≥ ((1−δ)^2 − η') |L|}
  X* = X* − A_i
end while
return A_1, ..., A_r and α_1, ..., α_r
```

4.1 Upper bound on the number of blocks

In this subsection, we limit the number of blocks. We prove that $r \leq \frac{1-\rho}{\eta'}$ when $\delta \leq 1 - \sqrt[3]{\rho} - \eta$, where $\eta' = \frac{3\sqrt[3]{\rho}\eta^2}{2}$.

Without loss of generality, we can set $\eta < 1$ since $0 < \delta \leq 1 - \sqrt[3]{\rho} - \eta$. Then we have

$$\begin{aligned}(1-\delta)^2 - \eta' &\geq (\rho^{\frac{1}{3}} + \eta)^2 - \frac{3\rho^{\frac{1}{3}}\eta^2}{2} \\ &= \rho^{\frac{2}{3}} + (2\rho^{\frac{1}{3}}\eta - \frac{3\rho^{\frac{1}{3}}\eta^2}{2}) + \eta^2 \\ &> \rho^{\frac{2}{3}} \geq \rho. \end{aligned} \tag{6}$$

Using Corrádi's lemma(Lemma 2), we can restrict the number of blocks.

Lemma 3. *Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $f : L \to \mathbb{F}_q$ be any codeword. We have*

$$r < \frac{1-\rho}{\eta'}.$$

Proof. We associate a set P_{α_i} for each representative α_i. From Algorithm 2, we know

- $|P_{\alpha_i} \cap P_{\alpha_j}| < \left((1-\delta)^2 - \eta'\right)|L|$ for any distinct i, j.
- $|P_{\alpha_i}| = 2|C_{\alpha_i}| = (1-\delta)|L|$ for any i.

By Lemma 2, we have

$$\left|\bigcup_{i=1}^{r} P_{\alpha_i}\right| \geq \frac{r(1-\delta)^2|L|^2}{(1-\delta)|L| + (r-1)\left((1-\delta)^2 - \eta'\right)|L|}.$$

On the other hand, $\bigcup_{i=1}^{r} P_{\alpha_i}$ is a subset of L. Thus,

$$|L| \geq \frac{r(1-\delta)^2|L|}{(1-\delta) + (r-1)\left((1-\delta)^2 - \eta'\right)},$$

that is

$$(1-\delta) + (r-1)\left((1-\delta)^2 - \eta'\right) \geq r(1-\delta)^2,$$

which implies

$$\begin{aligned} r &\leq \frac{1-\delta-\left((1-\delta)^2-\eta'\right)}{(1-\delta)^2-\left((1-\delta)^2-\eta'\right)} \\ &< \frac{1-\delta-\rho}{\eta'} && \text{by (6)} \\ &< \frac{1-\rho}{\eta'}. \end{aligned}$$

□

4.2 Upper bound on the size of each block

In this subsection, we aim to bound the number of elements in each block A_i. Combined with Lemma 3, we can limit the number of bad folding points and finish our proof of the main theorem.

Lemma 4. *For any distinct $\alpha_1, \alpha_2 \in \mathrm{Bad}(f)$, there exists a polynomial $p(X) \in \mathbb{F}_q[X]$ of degree strictly less than $\rho|L|$ such that*

$$f|_{P_{\alpha_1} \cap P_{\alpha_2}} = p|_{P_{\alpha_1} \cap P_{\alpha_2}}, \tag{7}$$

and

$$f|_{P^*_{\alpha_1} \cap P^*_{\alpha_2}} = p|_{P^*_{\alpha_1} \cap P^*_{\alpha_2}}. \tag{8}$$

Moreover, if $|P_{\alpha_1} \cap P_{\alpha_2}| \geq \rho|L|$, polynomial $p(X)$ is uniquely determined.

Proof. Notice that (8) implies (7) because $P_{\alpha_1} \cap P_{\alpha_2} \subseteq P^*_{\alpha_1} \cap P^*_{\alpha_2}$. So we only prove (8). By Definition 7, there exists a polynomial $p_1(X) \in \mathbb{F}_q[X]$ of degree strictly less than $\rho|L|/2$ such that $\mathrm{FuncFold}_{\alpha_1}(f)|_{C^*_{\alpha_1}} = p_1|_{C^*_{\alpha_1}}$. By Definition 4, we have

$$p_1(x^2) = \frac{f(x) + f(-x)}{2} + \alpha_1 \cdot \frac{f(x) - f(-x)}{2x} \tag{9}$$

for any $x \in P^*_{\alpha_1}$. Similarly, there exists a polynomial $p_2(X) \in \mathbb{F}_q[X]$ of degree strictly less than $\rho|L|/2$ such that $\text{FuncFold}_{\alpha_2}(f)|_{C^*_{\alpha_2}} = p_2|_{C^*_{\alpha_2}}$. And we have

$$p_2(x^2) = \frac{f(x)+f(-x)}{2} + \alpha_2 \cdot \frac{f(x)-f(-x)}{2x} \tag{10}$$

for any $x \in P^*_{\alpha_2}$.

From (9) and (10), we have

$$\begin{cases} p_1(x^2) - p_2(x^2) = \dfrac{\alpha_1-\alpha_2}{2x} \cdot (f(x)-f(-x)) \\ \alpha_2 p_1(x^2) - \alpha_1 p_2(x^2) = \dfrac{\alpha_2-\alpha_1}{2} \cdot (f(x)+f(-x)) \end{cases}$$

for any $x \in P^*_{\alpha_1} \cap P^*_{\alpha_2}$. That is,

$$\begin{cases} f(x)-f(-x) = \dfrac{2x}{\alpha_1-\alpha_2} \cdot \big(p_1(x^2) - p_2(x^2)\big) \\ f(x)+f(-x) = \dfrac{2}{\alpha_2-\alpha_1} \cdot \big(\alpha_2 p_1(x^2) - \alpha_1 p_2(x^2)\big) \end{cases}$$

Therefore, for any $x \in P^*_{\alpha_1} \cap P^*_{\alpha_2}$, we have

$$f(x) = \frac{x}{\alpha_1-\alpha_2} \cdot \big(p_1(x^2) - p_2(x^2)\big) + \frac{1}{\alpha_2-\alpha_1} \cdot \big(\alpha_2 p_1(x^2) - \alpha_1 p_2(x^2)\big). \tag{11}$$

Note that $\deg(p_1), \deg(p_2) \le \frac{1}{2} \cdot \rho|L| - 1$. From (11), we have

$$\begin{aligned} \deg(f) &\le 1 + 2\max(\deg(p_1), \deg(p_2)) \\ &\le 1 + \rho|L| - 2 \\ &= \rho|L| - 1. \end{aligned}$$

If $|P^*_{\alpha_1} \cap P^*_{\alpha_2}| \ge |P_{\alpha_1} \cap P_{\alpha_2}| \ge \rho|L|$, polynomial $p(X)$ is unique, since $\rho|L|$ points uniquely determine a polynomial of degree at most $\rho|L| - 1$. □

Proposition 3. *Let $\alpha, \beta \in \text{Bad}(f)$ be different points such that $|P_\alpha \cap P_\beta| \ge \rho|L|$. Let $p(X)$ be the unique polynomial of degree $< \rho|L|$ such that $p|_{P_\alpha \cap P_\beta} = f|_{P_\alpha \cap P_\beta}$. Then*

$$\text{Closest}(\text{FuncFold}_\beta(f), \text{RS}[\mathbb{F}_q, L^2, \rho]) = \text{PolyFold}_\beta(p). \tag{12}$$

Proof. Since both sides of (12) are polynomials of degree at most $\rho|L^2| - 1$, it suffices to find $\rho|L^2|$ points on which both sides of (12) are equal. We have $|C_\alpha \cap C_\beta| = \frac{1}{2} \cdot |P_\alpha \cap P_\beta| \ge \rho|L^2|$.

Let $x^2 \in C_\alpha \cap C_\beta$. By the definition of folding, we have

$$\begin{aligned} \text{PolyFold}_\beta(p)(x^2) &= \frac{p(x)+p(-x)}{2} + \beta \cdot \frac{p(x)-p(-x)}{2x} \\ &= \frac{f(x)+f(-x)}{2} + \beta \cdot \frac{f(x)-f(-x)}{2x} && \text{By Definition 7} \\ &= \text{FuncFold}_\beta(f)(x^2). \end{aligned}$$

Thus, we have exhibited at least $\rho|L^2|$ points where the evaluations of both sides of (12) are equal. □

Corollary 2. *Let $\alpha, \beta \in \mathrm{Bad}(f)$ be distinct and $|P^*_\alpha \cap P^*_\beta| \geq |P_\alpha \cap P_\beta| \geq \rho|L|$. Let $p(X)$ be the unique polynomial such that $f|_{P^*_\alpha \cap P^*_\beta} = p|_{P^*_\alpha \cap P^*_\beta}$ (by Lemma 4). For any $x \in L$, $x \in P^*_\alpha \cap P^*_\beta$ if and only if $f(x) = p(x)$ and $f(-x) = p(-x)$.*

Proof. The "only if" direction is trivial. Let $x \in P^*_\alpha \cap P^*_\beta$. Then $-x \in P^*_\alpha \cap P^*_\beta$ by Definition 7. Since $f|_{P^*_\alpha \cap P^*_\beta} = p|_{P^*_\alpha \cap P^*_\beta}$, we have $f(x) = p(x)$ and $f(-x) = p(-x)$.

For the "if" direction, assuming $f(x) = p(x)$ and $f(-x) = p(-x)$, our goal is to prove $x \in P^*_\alpha \cap P^*_\beta$. We prove $x \in P^*_\alpha$; $x \in P^*_\beta$ is similar to prove.

By Proposition 3, we have

$$\mathrm{PolyFold}_\alpha(p) = \mathrm{Closest}(\mathrm{FuncFold}_\alpha(f), \mathrm{RS}[\mathbb{F}_q, L^2, \rho]). \tag{13}$$

Since $f(x) = p(x)$ and $f(-x) = p(-x)$, by the definition of folding, we have

$$\mathrm{FuncFold}_\alpha(f)(x^2) = \mathrm{PolyFold}_\alpha(p)(x^2). \tag{14}$$

Combining (13) with (14), we have

$$\mathrm{Closest}(\mathrm{FuncFold}_\alpha(f), \mathrm{RS}[\mathbb{F}_q, L^2, \rho])(x^2) = \mathrm{FuncFold}_\alpha(f)(x^2),$$

which implies $x^2 \in C^*_\alpha$, i.e., $x \in P^*_\alpha$. □

Lemma 5. *Let $\alpha, \beta_1, \beta_2 \in \mathrm{Bad}(f)$ be different such that $|P_\alpha \cap P_{\beta_1}| \geq \rho|L|$ and $|P_\alpha \cap P_{\beta_2}| \geq \rho|L|$. Denote by $p_1(X)$ and $p_2(X)$ the polynomials of degree at most $\rho|L| - 1$ decided by $f|_{P_\alpha \cap P_{\beta_1}}$ and $f|_{P_\alpha \cap P_{\beta_2}}$ (Lemma 4). Then exactly one of the followings holds:*

- *$p_1 = p_2$ and $P^*_\alpha \cap P^*_{\beta_1} = P^*_\alpha \cap P^*_{\beta_2}$.*
- *$p_1 \neq p_2$ and $|P_\alpha \cap P_{\beta_1} \cap P_{\beta_2}| \leq |P^*_\alpha \cap P^*_{\beta_1} \cap P^*_{\beta_2}| \leq \rho|L| - 1$.*

Proof. If $p_1 = p_2$, we claim $P^*_\alpha \cap P^*_{\beta_1} = P^*_\alpha \cap P^*_{\beta_2}$. Let us prove $P^*_\alpha \cap P^*_{\beta_1} \subseteq P^*_\alpha \cap P^*_{\beta_2}$ first. Let $x \in P^*_\alpha \cap P^*_{\beta_1}$. By Corollary 2, we have $f(x) = p_1(x)$ and $f(-x) = p_1(-x)$. Since $p_1 = p_2$, we have $f(x) = p_2(x)$ and $f(-x) = p_2(-x)$. By Corollary 2, $x \in P^*_\alpha \cap P^*_{\beta_2}$. Thus, we have shown that $P^*_\alpha \cap P^*_{\beta_1} \subseteq P^*_\alpha \cap P^*_{\beta_2}$. The other direction, $P^*_\alpha \cap P^*_{\beta_2} \subseteq P^*_\alpha \cap P^*_{\beta_1}$, is similar to prove.

If $p_1 \neq p_2$, we want to prove $|P^*_\alpha \cap P^*_{\beta_1} \cap P^*_{\beta_2}| \leq \rho|L| - 1$. Note that, for any $x \in P^*_\alpha \cap P^*_{\beta_1} \cap P^*_{\beta_2}$, by Corollary 2, $f(x) = p_1(x)$ and $f(x) = p_2(x)$, which implies $p_1(x) = p_2(x)$. Since p_1, p_2 are different polynomials of degree at most $\rho|L| - 1$, and two different polynomials of degree at most $\rho|L| - 1$ agree on at most $\rho|L| - 1$ points, we have $|P^*_\alpha \cap P^*_{\beta_1} \cap P^*_{\beta_2}| \leq \rho|L| - 1$. □

Proposition 4. *Let $\alpha, \beta_1, \beta_2 \in \mathrm{Bad}(f)$ be different points such that*

- *$|P_\alpha \cap P_{\beta_1}|$, $|P_\alpha \cap P_{\beta_2}| \geq \rho|L|$*
- *$p_1 \neq p_2$, where p_1 and p_2 are the polynomials of degree at most $\rho|L| - 1$ determined by $f|_{P_\alpha \cap P_{\beta_1}}$ and $f|_{P_\alpha \cap P_{\beta_2}}$ respectively.*

Then $\mathrm{PolyFold}_\alpha(p_1) = \mathrm{PolyFold}_\alpha(p_2)$.

Proof. By our condition, we have $p_1|_{P_\alpha \cap P_{\beta_1}} = f|_{P_\alpha \cap P_{\beta_1}}$ and $p_2|_{P_\alpha \cap P_{\beta_2}} = f|_{P_\alpha \cap P_{\beta_2}}$, which implies that

$$\text{PolyFold}_\alpha(p_i)|_{C_\alpha \cap C_{\beta_i}} = \text{FuncFold}_\alpha(f)|_{C_\alpha \cap C_{\beta_i}}, i = 1, 2.$$

By Definition 7, $\text{FuncFold}_\alpha(f)$ coincides with a polynomial of degree $< \rho|L^2|$ on C_α, denoted by $p(X)$. Notice that $|C_\alpha \cap C_{\beta_1}| \geq \frac{1}{2} \cdot |P_\alpha \cap P_{\beta_1}| \geq \rho \cdot \frac{|L|}{2} = \rho|L^2|$; thus, $\text{PolyFold}_\alpha(p_1)$ is uniquely determined, and equals p. A similar argument shows $\text{PolyFold}_\alpha(p_1) = p$. Therefore, $\text{PolyFold}_\alpha(p_1) = \text{PolyFold}_\alpha(p_2)$. □

Remark 2. *Definition 7 can be extended to general points in $\mathbb{F}_q$, i.e., we can define the corresponding C^*_α, P^*_α for any $\alpha \in \mathbb{F}_q$. Furthermore, if $|P^*_\alpha \cap P^*_\beta| \geq \rho|L|$, the above results still hold.*

Let $A \in \{A_1, \ldots, A_r\}$ be a set of folding points defined in Algorithm 2 and α be the corresponding folding point, i.e., for any $\beta \in A$, we have $|P_\beta \cap P_\alpha| \geq \left((1-\delta)^2 - \eta'\right)|L|$. Since $\left((1-\delta)^2 - \eta'\right) \geq \rho$ according to (6), we can define an equivalence relation $\mathcal{R}$ on $A \setminus \{\alpha\}$ as follows:

$$(\beta_1, \beta_2) \in \mathcal{R} \iff p_1 = p_2,$$

where p_1 and p_2 are the low-degree polynomials determined by $f|_{P_\alpha \cap P_{\beta_1}}$ and $f|_{P_\alpha \cap P_{\beta_2}}$ respectively. $\mathcal{R}$ gives a partition on set $A \setminus \{\alpha\}$. Let s be the number of equivalence classes. For convenience, let

$$[\beta_i] = \{\beta_{i,1}, \ldots, \beta_{i,t_i}\},$$

$1 \leq i \leq s$ denote the equivalence classes.

From the definitions, we know

- (Lemma 5) Different equivalence classes correspond to different low-degree polynomials, denoted by $p_1(X), \ldots, p_s(X)$.
- (Proposition 4) The foldings of the polynomials $p_1, \ldots, p_s$ at point α are the same. That is,

$$\text{PolyFold}_\alpha(p_1) = \text{PolyFold}_\alpha(p_2) = \ldots = \text{PolyFold}_\alpha(p_s).$$

- (Proposition 3) For any $i \in \{1, \ldots, s\}$, and for any $j \in \{1, \ldots, t_i\}$, we have

$$\text{Closest}(\text{FuncFold}_{\beta_{i,j}}(f), \text{RS}_q[\mathbb{F}, L^2, \rho]) = \text{PolyFold}_{\beta_{i,j}}(p_i).$$

On the one hand, we can bound the number of equivalence classes by using the following generalization of Corrádi's lemma, whose proof is almost the same as Corrádi's.

Corollary 3 (Corollary of Corrádi's lemma)**.** *Let $P_1, \ldots, P_r$ be r sets satisfying*

$$s_1 \leq |P_i| < s_2, 1 \leq i \leq r.$$

If $|P_i \cap P_j| \leq k$ for any distinct $i, j \in \{1, 2, \ldots, r\}$, then

$$\left|\bigcup_{i=1}^{r} P_i\right| > \frac{s_1^2 r}{s_2 + (r-1)k}.$$

Proof. Follow the proof of Corrádi's Lemma. For any $x \in \bigcup_{i=1}^{r} P_i$, denote by $d(x)$ the count of x, i.e., the number of P_i containing x, we have

$$\begin{aligned}\sum_{x\in P_i} d(x) &= \sum_{j=1}^{r} |P_i \cap P_j| = |P_i| + \sum_{j\neq i} |P_i \cap P_j| \\ &< s_2 + (r-1)k. \end{aligned} \tag{15}$$

Summing over all sets P_i, we have

$$\begin{aligned}\sum_{i=1}^{r}\sum_{x\in P_i} d(x) &= \sum_{x\in\bigcup_{i=1}^{r} P_i} d(x)^2 \\ &\geq \frac{1}{|\bigcup_{i=1}^{r} P_i|}\left(\sum_{x\in\bigcup_{i=1}^{r} P_i} d(x)\right)^2 \qquad \text{by Cauchy–Schwarz inequality} \\ &= \frac{1}{|\bigcup_{i=1}^{r} P_i|}\left(\sum_{i=1}^{r} |P_i|\right)^2 \geq \frac{s_1^2 r^2}{|\bigcup_{i=1}^{r} P_i|}. \end{aligned} \tag{16}$$

Combining (15) and (16), we have

$$r\left(s_2 + (r-1)k\right) > \frac{s_1^2 r^2}{|\bigcup_{i=1}^{r} P_i|} \Rightarrow \left|\bigcup_{i=1}^{r} P_i\right| > \frac{s_1^2 r}{s_2 + (r-1)k}.$$

□

Corollary 4. *When $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\eta' = \frac{3\sqrt[3]{\rho}\eta^2}{2}$. Let $A \in \{A_1, \ldots, A_r\}$ be a set of folding points defined in Algorithm 2 and α be the corresponding folding point. Let s be the number of equivalence classes decided by $\mathcal{R}$ on $A \setminus \{\alpha\}$. We have $s \leq \frac{1}{3\eta\rho^{\frac{2}{3}}}$.*

Proof. Let

$$[\beta_i] = \{\beta_{i,1}, \ldots, \beta_{i,t_i}\}, 1 \leq i \leq s$$

be the equivalence classes. Notice that we have

- $|P_\alpha \cap P_{\beta_{i,1}}| < (1-\delta)|L|, 1 \leq i \leq s$ since $\Delta(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$ and $f|_{P_\alpha \cap P_{\beta_{i,1}}}$ agrees with a low-degree polynomial according to Lemma 4.
- $|P_\alpha \cap P_{\beta_{i,1}}| \geq \left((1-\delta)^2 - \eta'\right)|L|, 1 \leq i \leq s$ according to Algorithm 2.
- $|P_\alpha \cap P_{\beta_{i,1}} \cap P_{\beta_{j,1}}| < \rho|L|, 1 \leq i < j \leq s$ according to Lemma 5.
- $|\bigcup_{i=1}^{s} (P_\alpha \cap P_{\beta_i,1})| \leq |P_\alpha| = (1-\delta)|L|$.

By Corollary 3, we have

$$(1-\delta)|L| > \frac{\left(\left((1-\delta)^2 - \eta'\right)|L|\right)^2 s}{(1-\delta)|L| + (s-1)\rho|L|}.$$

Thus,

$$\begin{aligned} s &< \frac{(1-\delta-\rho)(1-\delta)}{((1-\delta)^2-\eta')^2-\rho(1-\delta)} \\ &= \frac{(1-\delta-\rho)(1-\delta)}{((1-\delta)^3-\rho)(1-\delta)-2\eta'(1-\delta)^2+\eta'^2} \\ &\leq \frac{(1-\delta-\rho)(1-\delta)}{\left(3\eta^2\rho^{\frac{1}{3}}+3\eta\rho^{\frac{2}{3}}+\eta^3\right)(1-\delta)-2\eta'(1-\delta)^2+\eta'^2}. \end{aligned}$$

Notice that $\eta' = \frac{3\rho^{\frac{1}{3}}\eta^2}{2}$, then we have

$$3\eta^2\rho^{\frac{1}{3}}(1-\delta)-2\eta'(1-\delta)^2 \geq 3\eta^2\rho^{\frac{1}{3}}(1-\delta)-2\eta'(1-\delta) = 0.$$

Thus, we have

$$s < \frac{(1-\delta-\rho)(1-\delta)}{3\eta\rho^{\frac{2}{3}}(1-\delta)} = \frac{1-\delta-\rho}{3\eta\rho^{\frac{2}{3}}} \leq \frac{1}{3\eta\rho^{\frac{2}{3}}}.$$

□

On the other hand, we can bound the number of elements in each equivalence class based on the following lemma.

Lemma 6. *Let $\alpha, \beta_1, \ldots, \beta_t \in \mathrm{Bad}(f)$ be distinct such that*

- $P^*_\alpha \cap P^*_{\beta_1} = P^*_\alpha \cap P^*_{\beta_2} = \ldots = P^*_\alpha \cap P^*_{\beta_t}$, *and*
- $|P^*_\alpha \cap P^*_{\beta_1}| \geq |P_\alpha \cap P_{\beta_1}| \geq \rho|L|$.

Then $P^*_{\beta_1}, \ldots, P^*_{\beta_t}$ form a sunflower with core $P^*_\alpha \cap P^*_{\beta_1}$. That is, for any distinct $i, j \in \{1, \ldots, t\}$, we have

$$P^*_{\beta_i} \cap P^*_{\beta_j} = P^*_\alpha \cap P^*_{\beta_1}.$$

Proof. Let $i, j \in \{1, \ldots, t\}$ be any two distinct numbers. Let $p(X)$ denote the unique polynomial of degree at most $\rho|L|-1$ determined by $f|_{P^*_\alpha \cap P^*_{\beta_i}}$ (Lemma 4).

Observe that $P^*_\alpha \cap P^*_{\beta_i} = P^*_\alpha \cap P^*_{\beta_i} \cap P^*_{\beta_j} \subseteq P^*_{\beta_i} \cap P^*_{\beta_j}$. By Lemma 4, there exists a unique polynomial $q(X) \in \mathbb{F}_q[X]$ of degree strictly less than $\rho|L|$ determined by $f|_{P^*_{\beta_i} \cap P^*_{\beta_j}}$. Since $p|_{P^*_\alpha \cap P^*_{\beta_i}} = f|_{P^*_\alpha \cap P^*_{\beta_i}}$, and polynomial p is uniquely determined, we have $q = p$.

For any $x \in P^*_{\beta_i} \cap P^*_{\beta_j}$, we have $f(x) = q(x)$ and $f(-x) = q(-x)$ by Corollary 2. Since $f(x) = p(x)$ and $f(-x) = p(-x)$, using Corollary 2 again, we have $x \in P^*_\alpha \cap P^*_{\beta_i} = P^*_\alpha \cap P^*_{\beta_i} \cap P^*_{\beta_j}$. Thus,

$$P^*_{\beta_i} \cap P^*_{\beta_j} = P^*_\alpha \cap P^*_{\beta_i} \cap P^*_{\beta_j} = P^*_\alpha \cap P^*_{\beta_1}.$$

Since i, j are arbitrary, we have completed the proof. □

Corollary 5. *Denote by t the upper bound of the number of elements in each equivalence class. Then we have $t \leq \frac{\delta}{2}|L|$.*

Proof. Let

$$[\beta_i] = \{\beta_{i,1}, \ldots, \beta_{i,t_i}\}$$

be an equivalence class. According to the definition of equivalent relation, we have $P^*_\alpha \cap P^*_{\beta_{i,1}} = \ldots = P^*_\alpha \cap P^*_{\beta_{i,t_i}}$. According to Algorithm 2, we have $|P^*_\alpha \cap P^*_{\beta_{i,1}}| \geq |P_\alpha \cap P_{\beta_{i,1}}| \geq \left((1-\delta)^2 - \eta'\right)|L| \geq \rho|L|$. Lemma 6 tells us

$$P^*_{\beta_{i,j}} \cap P^*_{\beta_{i,k}} = P^*_\alpha \cap P^*_{\beta_{i,1}}, 1 \leq j < k \leq t_i.$$

This implies

$$\left(P^*_{\beta_{i,j}} \setminus P^*_\alpha\right) \cap \left(P^*_{\beta_{i,k}} \setminus P^*_\alpha\right) = \emptyset, 1 \leq j < k \leq t_i. \tag{17}$$

Thus,

$$\delta|L| \geq |L \setminus P^*_\alpha| \geq \left|\bigcup_{j=1}^{t_i} \left(P^*_{\beta_{i,j}} \setminus P^*_\alpha\right)\right| = \sum_{j=1}^{t_i} \left|P^*_{\beta_{i,j}} \setminus P^*_\alpha\right| \tag{18}$$

For each $\beta_{i,j} \in [\beta_i]$, we have $|P^*_{\beta_{i,j}}| \geq (1-\delta)|L|$ according to the definition of bad folding points (Definition 6). On the other hand, Lemma 4 tells us f agrees with some low-degree polynomial on $P^*_\alpha \cap P^*_{\beta_{i,j}}$. Since $\Delta\left(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]\right) > \delta|L|$, we have $|P^*_\alpha \cap P^*_{\beta_{i,j}}| < (1-\delta)|L|$. Furthermore, according to Definition 7, the number of elements in $P^*_{\beta_{i,j}} \setminus P^*_\alpha$ is even. As a result, we have

$$\left|P^*_{\beta_{i,j}} \setminus P^*_\alpha\right| \geq 2. \tag{19}$$

Plug (19) into (18), we have

$$\delta|L| \geq 2t_i \Rightarrow t_i \leq \frac{\delta}{2}|L|.$$

□

4.3 Proof of Theorem 3

Algorithm 2 gives a partition of $\mathrm{Bad}(f)$, denoted by $A_1, \ldots, A_r \subseteq \mathrm{Bad}(f)$. By Lemma 3, the number of blocks is at most $r < \frac{1-\rho}{\eta'}$. The size of each block A_i is less than $st+1$, which is bounded by Corollary 4 and Corollary 5. So

$$\begin{aligned}
|\mathrm{Bad}(f)| &= \sum_{i=1}^{r} |A_i| \leq r \cdot (st+1) \\
&< \frac{1-\rho}{\eta'} \cdot \left(\frac{1}{3\eta\rho^{\frac{2}{3}}} \cdot \frac{\delta}{2}|L| + 1\right) && \text{Corollary 4 and Corollary 5} \\
&\leq \frac{1-\rho}{\eta'} \cdot \left(\frac{1}{3\eta\rho^{\frac{2}{3}}} \cdot \frac{1-\rho^{\frac{1}{3}}-\eta}{2}|L| + 1\right) \\
&\leq \frac{(1-\rho)\left(1-\rho^{\frac{1}{3}}\right)}{\frac{3}{2}\rho^{\frac{1}{3}}\eta^2 \cdot 3\eta\rho^{\frac{2}{3}} \cdot 2}|L| + \frac{1-\rho}{\frac{3}{2}\rho^{\frac{1}{3}}\eta^2} \\
&= \frac{(1-\rho)\left(1-\rho^{\frac{1}{3}}\right)}{9\rho\eta^3}|L| + \frac{2(1-\rho)}{3\rho^{\frac{1}{3}}\eta^2}
\end{aligned}$$

5 Proximity gaps for Reed-Solomon codes

A property displays a proximity gap (Definition 1.1 in [Ben+20b]) for Reed-Solomon codes if either all the members of a Reed-Solomon code are δ-close to the property or only a tiny fraction of members are δ-close this property. [Ben+20b] proposes this notion to improve the analysis of the soundness of (batched) FRI. Furthermore, their analysis covers the power of *correlated agreement*, which is used in the proof of (knowledge) soundness of protocols that use FRI as a sub-protocol.

In this section, we analyze the proximity gaps for Reed-Solomon codes using our main theorem. The correlated agreement property is also held in our improved results. More precisely, we prove correlated agreement over lines (Theorem 4) and correlated agreement over affine spaces (Theorem 6). These theorems lead to the soundness analysis of (batched) FRI in Section 6.

5.1 Correlated agreement over lines

Theorem 4. *Let L be a subset of $\mathbb{F}_q^\times$. Let $u_0, u_1 : L \to \mathbb{F}_q$. Let $\delta, \eta, \rho > 0$ satisfy $\delta \le 1 - \sqrt[3]{\rho} - \eta$ and $\eta \le \frac{1}{3}\rho^{-\frac{1}{3}}$, and suppose*

$$\mathbb{P}_{z\in\mathbb{F}_q}(\Delta(u_0 + zu_1, \mathrm{RS}[\mathbb{F}_q, L, \rho]) \le \delta) \ge \frac{2(1-\rho)|L|}{9\rho\eta^3|\mathbb{F}_q|}.$$

Then u_0, u_1 are simultaneously δ-close to $\mathrm{RS}[\mathbb{F}_q, L, \rho]$, i.e., $\exists v_0, v_1 \in \mathrm{RS}[\mathbb{F}_q, L, \rho]$ such that

$$|\{x \in L : (u_0(x), u_1(x)) = (v_0(x), v_1(x))\}| \ge (1-\delta)|L|.$$

Proof of Theorem 4 We use Corollary 1 to prove Theorem 4. In order to apply Theorem 3, we construct a function f whose folding result at point z is exactly $u_0 + zu_1$.

We can find an extension field of $\mathbb{F}_q$, denoted by $\mathbb{K}$, such that $\sqrt{L} \triangleq \{x \mid x^2 \in L\}$ is in this field. We define a function $f : \sqrt{L} \to \mathbb{K}$ as follows

$$f(x) = u_0(x^2) + xu_1(x^2), x \in \sqrt{L}.$$

Then for any $z \in \mathbb{F}_q$, for any $x^2 \in L$,

$$\mathrm{FuncFold}_z(f)(x^2) = \frac{f(x)+f(-x)}{2} + z\cdot\frac{f(x)-f(-x)}{2x} = u_0(x^2) + zu_1(x^2) \in \mathbb{F}_q.$$

Thus we have

$$\begin{aligned}&\{z \in \mathbb{K} \mid \Delta(\mathrm{FuncFold}_z(f), \mathrm{RS}[\mathbb{K}, L, \rho]) \le \delta|L|\}\\ \supseteq &\{z \in \mathbb{F}_q \mid \Delta(u_0 + zu_1, \mathrm{RS}[\mathbb{F}_q, L, \rho]) \le \delta|L|\}.\end{aligned}$$

Since

$$\mathbb{P}_{z\in\mathbb{F}_q}(\Delta(u_0 + zu_1, \mathrm{RS}[\mathbb{F}_q, L, \rho]) \le \delta) \ge \frac{2(1-\rho)|L|}{9\rho\eta^3|\mathbb{F}_q|},$$

we have

$$|\{z \in \mathbb{K} \mid \delta(\mathrm{FuncFold}_z(f), \mathrm{RS}[\mathbb{K}, L, \rho]) \le \delta\}| \ge \frac{2(1-\rho)|L|}{9\rho\eta^3} = \frac{(1-\rho)|\sqrt{L}|}{9\rho\eta^3}. \tag{20}$$

Corollary 1 implies that f is δ-close to $\mathrm{RS}[\mathbb{K}, \sqrt{L}, \rho]$. However, this is insufficient for our purposes, as we aim to demonstrate the correlated agreement of u_0 and u_1. For this purpose, we define a new distance called *pairing distance* and slightly improve Theorem 3 and Corollary 1 under this new distance.

Definition 8 (pairing distance)**.** *Let $L \subseteq \mathbb{F}_q^\times$ be a set with pairing elements. Let $f, f' : L \to \mathbb{F}_q$ be two functions. Define the pairing distance between f and f' to be*

$$\Delta_P(f, f') = |\{x \in L \mid f(x) \neq f'(x) \text{ or } f(-x) \neq f'(-x)\}|.$$

Let V be a set of codewords on L. Define the pairing distance between f and V to be

$$\Delta_P(f, V) \triangleq \min_{f' \in V} \Delta_P(f, f').$$

Let $\text{Closest}_P(f, V)$ denote the closest codeword (polynomial) under the pairing distance. If there are more than one codeword with the same minimal pairing distance, choose the one with the smallest lexicographical order.

The following lemma improves the result of Corollary 5 by using the pairing distance. $C_\alpha, P_\alpha, C_\alpha^*$ and P_α^* are defined as in Definition 7, with the distance is replaced by the pairing distance.

Lemma 7. *Let $L \subseteq \mathbb{F}_q^\times$ be a set with pairing elements. Let $f : L \to \mathbb{F}_q$ be a codeword such that $\Delta_P(f, \text{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$. $\alpha, \beta_1, \ldots, \beta_t \in \text{Bad}(f)$ are distinct and C, P, C^*, P^* are defined as in Definition 7 satisfying*

1. $P_\alpha^* \cap P_{\beta_1}^* = P_\alpha^* \cap P_{\beta_2}^* = \ldots = P_\alpha^* \cap P_{\beta_t}^*$.
2. $|P_\alpha^* \cap P_{\beta_1}^*| \geq |P_\alpha \cap P_{\beta_1}| \geq \rho|L|$;

Then we have $t \leq \frac{\delta}{2}|L|$.

Proof. Since $\alpha, \beta_1, \ldots, \beta_t$ satisfy the first two conditions, Lemma 6 tells us

$$P_{\beta_i}^* \cap P_{\beta_j}^* = P_\alpha^* \cap P_{\beta_1}^*, 1 \leq i < j \leq t.$$

This implies

$$\left(P_{\beta_i}^* \setminus P_\alpha^*\right) \cap \left(P_{\beta_j}^* \setminus P_\alpha^*\right) = \emptyset, 1 \leq i < j \leq t.$$

We have $|P_\alpha^*| \geq (1-\delta)|L|$ since $\alpha \in \text{Bad}(f)$. Thus,

$$\delta|L| \geq |L \setminus P_\alpha^*| \geq \left|\bigcup_{i=1}^{t} \left(P_{\beta_i}^* \setminus P_\alpha^*\right)\right| = \sum_{i=1}^{t} \left|P_{\beta_i}^* \setminus P_\alpha^*\right| \tag{21}$$

Since $\Delta_P(f, \text{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$, we have $|P_\alpha^* \cap P_{\beta_i}^*| < (1-\delta)|L|$. Otherwise, if $|P_\alpha^* \cap P_{\beta_i}^*| \geq (1-\delta)|L|$, by Lemma 4, let p be the low degree polynomial that agrees with f on $P_\alpha^* \cap P_{\beta_i}^*$. By Corollary 2, for any $x \in P_\alpha^* \cap P_{\beta_i}^*$, $f(x) = p(x)$. According to Definition 7, $-x \in P_\alpha^* \cap P_{\beta_i}^*$ and $f(-x) = p(-x)$. Thus, $\Delta_P(f, p) \leq |L \setminus (P_\alpha^* \cap P_{\beta_i}^*)| \leq \delta|L|$. A contradiction. On the other hand, we have $|P_{\beta_i}^*| \geq (1-\delta)|L|$ since $\beta_i \in \text{Bad}(f)$. As a result, $\left|P_{\beta_i}^* \setminus P_\alpha^*\right| > 0$. Furthermore, according to Definition 7, the number of elements in $P_{\beta_i}^* \setminus P_\alpha^*$ is even. Thus, we have

$$\left|P_{\beta_i}^* \setminus P_\alpha^*\right| \geq 2. \tag{22}$$

Plug (22) into (21), we have

$$\delta|L| \geq 2t \Rightarrow t \leq \frac{\delta}{2}|L|.$$

□

Lemma 7 limits the number of elements in the same equivalence class in this case. Following a similar argument as that of Theorem 3 (Corollary 1), we can prove the following result, which strengthens Theorem 3 (Corollary 1).

Theorem 5. *Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. Let $f : L \to \mathbb{F}_q$ be a codeword such that $\delta_P(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta$. Then,*

$$\left|\left\{\alpha \in \mathbb{F}_q : \Delta\left(\mathrm{FuncFold}_\alpha(f), \mathrm{RS}[\mathbb{F}_q, L^2, \rho]\right) \leq \delta|L^2|\right\}\right| < \frac{(1-\rho)|L|}{9\rho\eta^3}.$$

Notice that $\sqrt{L}$ is with pairing elements. According to Theorem 5, (20) implies $\delta_P(f, \mathrm{RS}[\mathbb{K}, \sqrt{L}, \rho]) \leq \delta$. Let $p = \mathrm{Closest}_P(f, \mathrm{RS}[\mathbb{K}, \sqrt{L}, \rho])$ be the closest codeword (in pairing distance), then we have

$$\begin{aligned} \left|\sqrt{L}\right| - \Delta_P(f, p) &= \left|\{x \in \sqrt{L} : (f(x), f(-x)) = (p(x), p(-x))\}\right| \\ &\geq (1-\delta)\left|\sqrt{L}\right|. \end{aligned} \tag{23}$$

Construct $v_0, v_1 \in \mathrm{RS}[\mathbb{K}, L, \rho]$ as follows:

$$\begin{cases} v_0(x^2) = \frac{1}{2}\left(p(x) + p(-x)\right) \\ v_1(x^2) = \frac{1}{2x}\left(p(x) - p(-x)\right) \end{cases}, \quad \forall x \in \sqrt{L}.$$

Then

$$\begin{aligned} &\left|\{x^2 \in L : \left(u_0(x^2), u_1(x^2)\right) = \left(v_0(x^2), v_1(x^2)\right)\}\right| \\ =\ &\frac{1}{2}\left|\{x \in \sqrt{L} : (f(x), f(-x)) = (p(x), p(-x))\}\right| \\ \geq\ &(1-\delta)\frac{\left|\sqrt{L}\right|}{2} = (1-\delta)|L|. \qquad \text{By (23)} \end{aligned}$$

By the definition of v_0, v_1, we know v_0 and v_1 are low-degree polynomials in $\mathbb{K}[X]$. Furthermore, since $(1-\delta)|L| \geq \rho|L|$, we can interpolate on the agree points to get the unique low-degree polynomials v_0, v_1. And v_0, v_1 takes the values in $\mathbb{F}_q$ on these points. Thus, $v_0, v_1 \in \mathrm{RS}[\mathbb{F}_q, L, \rho]$. Therefore, we have completed the proof of Theorem 4.

5.2 Correlated agreement over affine spaces

Theorem 6. *Let L be a subset of $\mathbb{F}_q^\times$. Let $u_0, \cdots, u_l : L \to \mathbb{F}_q, l \geq 1$ be a sequence of functions. Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. For convenience, denote $V = \mathrm{RS}[\mathbb{F}_q, L, \rho]$. Define*

$$S = \{\mathbf{z}_l = \langle z_1, \cdots, z_l\rangle \in \mathbb{F}_q^l : \Delta(u_0 + z_1u_1 + \cdots + z_lu_l, V) \leq \delta|L|\}$$

and suppose

$$\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S) \geq \left(\frac{2(1-\rho)|L|}{9\rho\eta^3|\mathbb{F}_q|}\right) \cdot l.$$

Then $u_0, \cdots, u_l$ are simultaneously δ-close to V, i.e., $\exists v_0, \cdots, v_l \in V$ such that

$$|\{x \in L \mid u_i(x) = v_i(x), i = 1, \ldots, l\}| \geq (1-\delta)|L|.$$

To prove the above theorem, we consider the list decoding of a given function. We bound the number of folding points that *enlarge* the agree domains. More precisely, we consider the following domains:

Definition 9 (Maximal δ-pairing-agree domain)**.** *Let $0 \leq \delta \leq 1$. Let $L \subseteq \mathbb{F}_q^\times$ be a set with pairing elements. Let $f : L \to \mathbb{F}_q$ be a function. Let V be a set of codewords on L. Let $D \subseteq L$ be a domain satisfying:*

- ***Density:*** *$|D| \geq (1-\delta)|L|$;*
- ***Pairing:*** *$x \in D \iff -x \in D$;*
- ***Agreement:*** *$\exists v \in V$ such that $f|_D = v|_D$;*
- ***Maximal:*** *If $D \subsetneq D'$, then $\nexists v \in V$ such that $f|_{D'} = v|_{D'}$.*

We define such domain as a maximal δ-pairing-agree domain between f and V. Denote the set of all of the maximal δ-pairing-agree domains between f and V as

$$\mathcal{D}_{\delta,f,V} \triangleq \{D_1, \ldots, D_m\}.$$

Notice that $\mathcal{D}_{\delta,f,V}$ is unique and when $\delta_P(f,V) > \delta$, $\mathcal{D}_{\delta,f,V}$ is empty.

Definition 10. *Let $L \subseteq \mathbb{F}_q$ be an evaluation domain. Let $f_1, f_2 : L \to \mathbb{F}_q$ be to functions on L. Function $\mathrm{Agree}(f_1, f_2)$ returns the locations where f_1 agrees with f_2, i.e.,*

$$\mathrm{Agree}(f_1, f_2) = \{x \in L \mid f_1(x) = f_2(x)\}.$$

Based on the above definitions, we can define the set of bad folding points that *enlarge* the list-decoding agree domains.

Definition 11. *Let $L \subseteq \mathbb{F}_q^\times$ be a set with pairing elements. Let $0 < \delta \leq 1$. Let $f : L \to \mathbb{F}_q$ be a function. Suppose $\mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{F},L,\rho]} = \{D_1, \ldots, D_m\}$. Define the set of list-bad folding points to be*

$$\begin{aligned}\mathrm{Bad_L}(f) \triangleq \{\alpha \in \mathbb{F}_q \mid &\exists v \in \mathrm{RS}[\mathbb{F}_q, L^2, \rho], \text{ such that } \delta(\mathrm{FuncFold}_\alpha(f), v) \leq \delta \\ &\text{and } \mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \neq D_i^2, i = 1, \ldots, m\}.\end{aligned}$$

Theorem 7. *Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. Let $L \subseteq \mathbb{F}_q^\times$ be a set with pairing elements. Let $f : L \to \mathbb{F}_q$ be a function. Then we have*

$$|\mathrm{Bad_L}(f)| < \frac{(1-\rho)|L|}{9\rho\eta^3}.$$

The proof of Theorem 7 follows the proof of Theorem 3. We only outline the modifications here. The proof details can be found in Appendix A.

Proof. We first modify the definition of some sets corresponding to Definition 7. For any $\alpha \in \mathrm{Bad_L}(f)$, $\exists v \in \mathrm{RS}[\mathbb{F}_q, L^2, \rho]$ such that $\delta(\mathrm{FuncFold}_\alpha(f), v) \leq \delta$ and $\mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \neq D_i^2, i = 1, \ldots, m$. If more than one codewords satisfy these conditions, choose the one with the smallest lexicographical order. Define

$$C^*_{L,\alpha} \triangleq \mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \subseteq L^2, \tag{24}$$

and $P^*_{L,\alpha} \triangleq \{x \mid x^2 \in C^*_{L,\alpha}\} \subseteq L$ is the parent set of $C^*_{L,\alpha}$. For each $C^*_{L,\alpha}$, we prove that there are two possible cases:

1. $|C^*_{L,\alpha} \cap D_i^2| < \rho|L^2|, i \in \{1, \ldots, m\}$, or
2. $\exists D_i \in \mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{F}_q,L,\rho]}$, such that $C^*_{L,\alpha} \supsetneq D_i^2$.

This is because if $\exists D_i, |C^*_{L,\alpha} \cap D_i^2| \geq \rho|L^2|$, denote by p the low-degree polynomial f agrees on D_i. According to Definition 9, $x \in D_i \iff -x \in D_i$, thus

$$\mathrm{FuncFold}_\alpha(f)|_{D_i^2} = \mathrm{PolyFold}_\alpha(p)|_{D_i^2}.$$

Notice that

$$\mathrm{PolyFold}_\alpha(p)|_{C^*_{L,\alpha} \cap D_i^2} = v|_{C^*_{L,\alpha} \cap D_i^2}$$

because of (24). Since $|C^*_{L,\alpha} \cap D_i^2| \geq \rho|L^2|$, we have $\mathrm{PolyFold}_\alpha(p) = v$. Then

$$\mathrm{FuncFold}_\alpha(f)|_{D_i^2} = \mathrm{PolyFold}_\alpha(p)|_{D_i^2} = v|_{D_i^2}.$$

As a result, we have $C^*_{L,\alpha} = \mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \supseteq D_i^2$. Definition 11 tells us $\mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \neq D_i^2$, thus,

$$\mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \supsetneq D_i^2.$$

Run Algorithm 2 on $\mathrm{Bad}_{\mathrm{L}}(f)$ and the corresponding parent sets to give a partition on $\mathrm{Bad}_{\mathrm{L}}(f)$. Denote the output as $A'_1, \ldots, A'_r$ and $\alpha_1, \ldots, \alpha_r$. Lemma 3 still holds because our partition strategy is unchanged and we have $|C_{L,\alpha}| = (1-\delta)|L^2|$ for any $\alpha \in \mathrm{Bad}_{\mathrm{L}}(f)$. Thus, we have $r < \frac{1-\rho}{\eta'}$. If we can bound the size of each block A_i, we finish our proof of the theorem.

For any $\alpha \in \{\alpha_1, \ldots, \alpha_r\}$, denote by A the block α is in.

For the first case that $|C^*_{L,\alpha} \cap D_i^2| < \rho|L^2|, i = 1, \ldots, m$, we claim that $A = \{\alpha\}$. Otherwise, if $\exists \beta \in A$ and $\beta \neq \alpha$, we have f agree with some low-degree polynomial p on $P^*_{L,\alpha} \cap P^*_{L,\beta}$ by using the same skill as Lemma 4 (see Lemma 9 in Appendix A for details). Since $|P^*_{L,\alpha} \cap P^*_{L,\beta}| \geq |P_{L,\alpha} \cap P_{L,\beta}| \geq \rho|L|$ according to the partition in Algorithm 2 and $|C^*_{L,\alpha} \cap D_i^2| < \rho|L^2|, i = 1, \ldots, m$ according to our assumption, $P^*_{L,\alpha} \cap P^*_{L,\beta} \nsubseteq D_i, i = 1, \ldots, m$. A contradiction to Definition 9. So $|A| = 1$.

For the second case, if $|A| = 1$, we finish our proof of the theorem. Otherwise, we prove $|A| \leq \frac{\delta|L|}{2\eta'} + 1$ in this case, we outline the proof sketch here, details can be found in Appendix A. We still have the equivalent relation on $A \setminus \{\alpha\}$ (see Lemma 10 in Appendix A). Notice that Corollary 4 still holds in this case, we can bound the number of equivalence classes (Corollary 7), i.e., $s \leq \frac{1}{3\eta\rho^{\frac{2}{3}}}$. On the other hand, Lemma 6 still holds (Lemma 11), so we can limit the number of elements in each equivalence class (Corollary 8). Let

$$[\beta] = \{\beta_1, \ldots, \beta_t\}$$

be an equivalence class. Based on Lemma 11, we can prove

$$\delta|L| \geq |L \setminus P^*_{L,\alpha}| \geq \left|\bigcup_{i=1}^{t} \left(P^*_{L,\beta_i} \setminus P^*_{L,\alpha}\right)\right| = \sum_{i=1}^{t} \left|P^*_{L,\beta_i} \setminus P^*_{L,\alpha}\right| \tag{25}$$

According to Lemma 9, f agrees with a low-degree polynomial p on $P^*_{L,\alpha} \cap P^*_{L,\beta_i}$. Since $x \in P^*_{L,\alpha} \cap P^*_{L,\beta_i} \iff -x \in P^*_{L,\alpha} \cap P^*_{L,\beta_i}$, $\exists D_i \in \mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{F}_q,L,\rho]}$, such that $P^*_{L,\alpha} \cap P^*_{L,\beta_i} \subseteq D_i$. According

to the definition of $\mathrm{Bad}_{\mathrm{L}}(f)$ and P^*_{L,β_i}, $P^*_{L,\beta_i} \neq D_i$. Furthermore, the number of elements in $P^*_{L,\beta_i} \setminus P^*_{L,\alpha}$ is even. As a result,

$$|P^*_{L,\beta_i} \setminus P^*_{L,\alpha}| = |P^*_{L,\beta_i} \setminus (P^*_{L,\alpha} \cap P^*_{L,\beta_i})| \geq 2. \tag{26}$$

Plugging (26) into (25), we have

$$\delta|L| \geq 2t \Rightarrow t \leq \frac{\delta}{2}|L|.$$

So we have $|A| \leq st + 1 \leq \frac{\delta|L|}{6\eta\rho^{\frac{2}{3}}} + 1$ in this case. □

Definition 12 (maximal δ-correlated-agree domains)**.** *Let $0 \leq \delta \leq 1$. Let $L \subseteq \mathbb{F}_q$ be an evaluation domain. Let $l \geq 2$, $u_j : L \to \mathbb{F}_q, 1 \leq i \leq l$ be a series of functions. Let V be a set of code words on L. Let $D \subseteq L$ be a domain satisfying:*

- ***Density:*** *$|D| \geq (1-\delta)|L|$;*
- ***Correlated agreement:*** *For $i \in \{1, \ldots, l\}$, $\exists v_i \in V$, such that $u_i|_D = v_i|_D$;*
- ***Maximal:*** *If $D \subsetneq D'$, then $\exists i \in \{1, \ldots, l\}$, for all $v \in V$, $u_i|_{D'} \neq v|_{D'}$.*

We define such domain as a maximal δ-correlated-agree domain between u and V. Denote the set of all of the maximal δ-correlated-agree domains between u and V as $\mathcal{A}_{\delta,\{u_1,\ldots,u_l\},V} \triangleq \{D_1, \ldots, D_m\}$. Notice that $\mathcal{A}_{\delta,\{u_1,\ldots,u_l\},V}$ is unique and can be empty.

We have the corresponding theorem in the proximity gap version:

Theorem 8. *Let L be a subset of $\mathbb{F}_q^\times$. Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. Let $u_0, u_1 : L \to \mathbb{F}_q$. Let*

$$\begin{aligned} S_1 \triangleq \{z \in \mathbb{F}_q \mid &\exists v \in \mathrm{RS}[\mathbb{F}_q, L, \rho], \text{ such that } \delta(u_0 + zu_1, v) \leq \delta \\ &\text{and } \mathrm{Agree}(u_0 + zu_1, v) \notin \mathcal{A}_{\delta,\{u_0,u_1\},\mathrm{RS}[\mathbb{F}_q,L,\rho]}\}. \end{aligned}$$

Then

$$|S_1| < \frac{2(1-\rho)|L|}{9\rho\eta^3}.$$

Proof. Let $\sqrt{L} \triangleq \{x \in \mathbb{K} \mid x^2 \in L\}$ be the parent set of L, where $\mathbb{K}$ is an extension field of $\mathbb{F}_q$ whose characteristic is the same as $\mathbb{F}_q$. Construct a function $f : \sqrt{L} \to \mathbb{K}$ as follows:

$$f(x) = u_0(x^2) + xu_1(x^2), x \in \sqrt{L}.$$

Suppose $\mathcal{A}_{\delta,\{u_0,u_1\},\mathrm{RS}[\mathbb{F}_q,L,\rho]} = \{D_1, \cdots, D_m\}$. Let $\sqrt{D_i} \triangleq \{x \mid x^2 \in D_i\} \subseteq \sqrt{L}, i = 1, \ldots, m$ be the parent sets of $D_i, i = 1, \ldots, m$. $\mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{K},\sqrt{L},\rho]}$ is the set of the maximal δ-pairing-agree domains between f and $\mathrm{RS}[\mathbb{F}_q, \sqrt{L}, \rho]$. We want to prove that

$$\mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{K},\sqrt{L},\rho]} = \{\sqrt{D_1}, \ldots, \sqrt{D_m}\}. \tag{27}$$

On one hand, for any $D_i \in \mathcal{A}_{\delta,\{u_0,u_1\},\mathrm{RS}[\mathbb{F}_q,L,\rho]}$, let $v_0, v_1 \in \mathrm{RS}[\mathbb{F}_q, L, \rho]$ be the corresponding codewords satisfying

$$u_0|_{D_i} = v_0|_{D_i}, u_1|_{D_i} = v_1|_{D_i}.$$

Construct $v(x) = v_0(x^2) + xv_1(x^2), x \in \sqrt{L}$. Then $v \in \mathrm{RS}[\mathbb{K}, \sqrt{L}, \rho]$ and $f|_{\sqrt{D_i}} = v|_{\sqrt{D_i}}$. So $\exists \sqrt{D} \in \mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{K},\sqrt{L},\rho]}$ such that $\sqrt{D_i} \subseteq \sqrt{D}$. For any $x \in \sqrt{D}$, we have $-x \in \sqrt{D}$ according to Definition 9. Then we have $f(x) = v(x)$ and $f(-x) = v(-x)$, which implies the following equations:

$$\begin{cases} u_0(x^2) + xu_1(x^2) = v_0(x^2) + xv_1(x^2) \\ u_0(x^2) - xu_1(x^2) = v_0(x^2) - xv_1(x^2) \end{cases}.$$

Since $x \neq 0$, we have $u_0(x^2) = v_0(x^2), u_1(x^2) = v_1(x^2)$. As a result, $x \in \sqrt{D_i}$. So we have $\sqrt{D} \subseteq \sqrt{D_i}$. Thus $\sqrt{D_i} = \sqrt{D} \in \mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{K},\sqrt{L},\rho]}$.

On the other hand, for any $\sqrt{D} \in \mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{K},\sqrt{L},\rho]}$, let $v \in \mathrm{RS}[\mathbb{K}, \sqrt{L}, \rho]$ be the corresponding codeword such that

$$f|_{\sqrt{D}} = v|_{\sqrt{D}}.$$

Since $x \in \sqrt{D} \iff -x \in \sqrt{D}$ according to Definition 9, $\mathrm{FuncFold}_z(f)|_D = \mathrm{PolyFold}_z(v)|_D$, $\forall z \in \mathbb{F}_q$. Since $|D| \geq (1-\delta)|L|$, $\exists D_i \in \mathcal{A}_{\delta,\{u_0,u_1\},\mathrm{RS}[\mathbb{F}_q,L,\rho]}$ such that $D \subseteq D_i$. Let the v_0, v_1 be the code words such that

$$u_0|_{D_i} = v_0|_{D_i}, u_1|_{D_i} = v_1|_{D_i}.$$

Then we have

$$f(x) = u_0(x^2) + xu_1(x^2) = v_0(x^2) + xv_1(x^2) = v(x), x \in \sqrt{D_i}.$$

This implies $\sqrt{D_i} \subseteq \sqrt{D}$. We have proved $D = D_i$.

Thus, (27) holds. Furthermore, $\forall z \in \mathbb{F}_q, \forall x^2 \in L$, we have

$$\mathrm{FuncFold}_z(f)(x^2) = \frac{f(x) + f(-x)}{2} + z \cdot \frac{f(x) - f(-x)}{2x} = u_0(x^2) + zu_1(x^2) \in \mathbb{F}_q.$$

Then

$$\begin{aligned} S_1 = \{z \in \mathbb{F}_q \,|&\exists v \in \mathrm{RS}[\mathbb{F}_q, L, \rho], \text{ such that } \delta(u_0 + zu_1, v) \leq \delta \\ &\text{and } \mathrm{Agree}(u_0 + zu_1, v) \neq D_i, i = 1, \ldots, m\} \\ \subseteq \{z \in \mathbb{K} \,|&\exists v \in \mathrm{RS}[\mathbb{K}, L, \rho], \text{ such that } \delta(\mathrm{FuncFold}_z(f), v) \leq \delta \\ &\text{and } \mathrm{Agree}(\mathrm{FuncFold}_z(f), v) \neq D_i, i = 1, \ldots, m\}. \end{aligned}$$

We have proved that (27) holds, so Theorem 7 tells us

$$|S| < \frac{(1-\rho)|\sqrt{L}|}{9\rho\eta^3} = \frac{2(1-\rho)|L|}{9\rho\eta^3}.$$

□

Corollary 6. *Let L be a subset of $\mathbb{F}_q^\times$. Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. $l \geq 1$, let $u_0, \cdots u_l : L \to \mathbb{F}_q$ be a series of functions, denoted as $\mathbf{u}_l \triangleq \{u_0, \cdots, u_l\}$. $V \triangleq \mathrm{RS}[\mathbb{F}_q, L, \rho]$. Suppose $\mathcal{A}_{\delta,\mathbf{u}_l,V} = \{D_1, \cdots, D_{m_l}\}$. Let*

$$\begin{aligned} S_l \triangleq \{\mathbf{z}_l \in \mathbb{F}_q^l \,|&\exists v \in V, \text{ such that } \Delta(u_0 + \cdots + z_l u_l, v) \leq \delta \\ &\text{and } \mathrm{Agree}(u_0 + \cdots + z_l u_l, v) \notin \mathcal{A}_{\delta,\mathbf{u}_l,V}\}. \end{aligned}$$

Then

$$\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S_l) < \frac{2(1-\rho)|L|}{9\rho\eta^3|\mathbb{F}_q|} \cdot l.$$

Proof. We use induction to prove the result. When $l = 1$, the problem is reduced to Theorem 8. Suppose the result holds for $l-1$ functions. Denote $\{u_0, \cdots, u_{l-1}\}$ as $\mathbf{u}_{l-1}$. Define

$$S_{l-1} \triangleq \{\mathbf{z}_{l-1} \in \mathbb{F}_q^{l-1} \,|\exists v \in V, \text{ such that } \Delta(u_0 + \cdots + z_{l-1}u_{l-1}, v) \leq \delta \\ \text{and Agree}(u_0 + \cdots + z_l u_l, v) \notin \mathcal{A}_{\delta,\mathbf{u}_{l-1},V}\}.$$

Then we have

$$\begin{aligned} &\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S_l) \\ =&\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S_l \mid \mathbf{z}_{l-1} \in S_{l-1}) \cdot \mathbb{P}_{\mathbf{z}_{l-1} \in \mathbb{F}_q^{l-1}}(\mathbf{z}_{l-1} \in S_{l-1}) + \\ &\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S_l \mid \mathbf{z}_{l-1} \notin S_{l-1}) \cdot \mathbb{P}_{\mathbf{z}_{l-1} \in \mathbb{F}_q^{l-1}}(\mathbf{z}_{l-1} \notin S_{l-1}) \\ \leq&\mathbb{P}_{\mathbf{z}_{l-1} \in \mathbb{F}_q^{l-1}}(\mathbf{z}_{l-1} \in S_{l-1}) + \mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S_l \mid \mathbf{z}_{l-1} \notin S_{l-1}). \end{aligned} \tag{28}$$

According to our assumption, the first item satisfies

$$\mathbb{P}_{\mathbf{z}_{l-1} \in \mathbb{F}_q^{l-1}}(\mathbf{z}_{l-1} \in S_{l-1}) < \frac{2(1-\rho)|L|}{9\rho\eta^3|\mathbb{F}_q|} \cdot (l-1). \tag{29}$$

For the second item, for any fixed $\mathbf{z}_{l-1} \notin S_{l-1}$, define the set

$$S_{\mathbf{z}_{l-1}} \triangleq \{z_l \in \mathbb{F}_q : \langle \mathbf{z}_{l-1}, z_l \rangle \in S_l\}.$$

For convenience, use u' to denote $u_0 + \cdots + z_{l-1}u_{l-1}$. Consider the set

$$S_{\delta,\{u',u_l\},V} = \{z_l \in \mathbb{F}_q \,|\exists v \in V, \text{ such that } \delta(u' + z_l u_l, v) \leq \delta \\ \text{and Agree}(u' + z_l u_l, v) \notin \mathcal{A}_{\delta,\{u',u_l\},V}\}.$$

We want to prove

$$S_{\mathbf{z}_{l-1}} \subseteq S_{\delta,\{u',u_l\},V}. \tag{30}$$

Notice that $|S_{\delta,\{u',u_l\},V}| < \frac{2(1-\rho)|L|}{9\rho\eta^3}$ by Theorem 8. So if (30) holds, we have

$$|S_{\mathbf{z}_{l-1}}| < \frac{2(1-\rho)|L|}{9\rho\eta^3}. \tag{31}$$

Suppose $\mathcal{A}_{\delta,\{u',u_l\},V} = \{D'_1, \ldots, D'_{m'_1}\}$. Since $\mathbf{z}_{l-1} \notin S_{\delta,\mathbf{u}_{l-1},V}$, $\{D'_1, \ldots, D'_{m'_1}\}$ are δ-correlated-agree domains of $u_0, \cdots, u_{l-1}$ (may be not maximal). According to the definition of $\mathcal{A}_{\delta,\{u',u_l\},V}$, $\{D'_1, \ldots, D'_{m'_1}\}$ are δ-correlated-agree domains of $u_0, \cdots, u_l$. For any $D'_i \in \mathcal{A}_{\delta,\{u',u_l\},V}$, $\exists D \in \mathcal{A}_{\delta,\mathbf{u}_l,V}$, such that $D'_i \subseteq D$. If $D'_i \neq D$, it is obvious that D is a correlated-agreement between u', u_l and V. This is a contradiction to $D'_i \in \mathcal{A}_{\delta,\{u',u_l\},V}$. So we have

$$\mathcal{A}_{\delta,\{u',u_l\},V} \subseteq \mathcal{A}_{\delta,\mathbf{u}_l,V}. \tag{32}$$

Fix a $z \in S_{\mathbf{z}_{l-1}}$, $\langle \mathbf{z}_{l-1}, z \rangle \in S_l$. According to the definition of S_l, $\exists v \in V$ such that

$$\text{Agree}(u' + zu_l, v) \notin \mathcal{A}_{\delta,\mathbf{u}_l,V} \text{ and } |\text{Agree}(u' + zu_l, v)| \geq (1-\delta)|L|.$$

(32) tells us $\text{Agree}(u' + zu_l, v) \notin \mathcal{A}_{\delta,\{u',u_l\},V}$. Thus,

$$z \in S_{\delta,\{u',u_l\},V}.$$

We have proved (30).

Plugging (29) and (31) into (28), we have

$$\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S_l) < \frac{2(1-\rho)|L|}{9\rho\eta^3|\mathbb{F}_q|} \cdot l.$$

□

Proof of Theorem 6

We have $S_l \subseteq S$ according to their definitions. Furthermore, since

$$\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S) \geq \frac{2(1-\rho)|L|}{9\rho\eta^3|\mathbb{F}_q|} \cdot l > \mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^l}(\mathbf{z}_l \in S_l) \qquad \text{by Corollary 6}$$

we have $S_l \subsetneq S$. This means there exists a δ-correlated-agree domain between $u_0, \cdots, u_l$ and $\text{RS}[\mathbb{F}_q, L, \rho]$.

5.3 Conjectured proximity gaps

Besides the provable proximity gaps of the RS codes, [Ben+20b] also provides a conjecture. Many implementations of FRI are based on this conjecture.

Conjecture 1 (Conjecture 8.4 in [Ben+20b]). *There exist universal constants $c_1, c_2 > 0$ such that the following holds. Let $u_0, u_1 : L \to \mathbb{F}_q$. Let $\delta, \eta > 0$ and $\delta \leq 1 - \rho - \eta$, and suppose*

$$\mathbb{P}_{z \in \mathbb{F}_q}(\Delta(u_0 + zu_1, \text{RS}[\mathbb{F}_q, L, \rho]) \leq \delta) > \frac{1}{(\eta\rho)^{c_1}} \cdot \frac{|L|^{c_2}}{|\mathbb{F}_q|}.$$

Then u_0, u_1 are simultaneously δ-close to $\text{RS}[\mathbb{F}_q, L, \rho]$, i.e. $\exists v_0, v_1 \in \text{RS}[\mathbb{F}_q, L, \rho]$ such that

$$|\{x \in L : (u_0(x), u_1(x)) = (v_0(x), v_1(x))\}| \geq (1-\delta)|L|.$$

[Ben+20b] proofs the conjecture when $c_2 = 2$ and δ is under the Johnson bound, i.e., $\delta \leq 1 - \sqrt{\rho} - \eta$. They state that "To the best of our knowledge, nothing contradicts setting $c_1 = c_2 = 2$" and "When limiting the scope to fields of characteristic greater than k (degree of the RS code), we are not aware of anything contradicting $c_1 = c_2 = 1$".

Theorem 4 provides proof for a part of the conjecture when setting $c_2 = 1$. The parameter δ needs to be under the double Johnson bound in our setting, i.e., $\delta \leq 1 - \sqrt[3]{\rho} - \eta$.

The proof of the remaining part of the conjecture is still open.

6 Soundness of batched FRI

FRI[Ben+18] is an IOPP for testing proximity to the RS codes. It is used to help the verifier to check whether a given function $f : L^{(0)} \to \mathbb{F}_q$ belongs to $\text{RS}[\mathbb{F}_q, L^{(0)}, \rho]$ or is far from the code. In particular, FRI works for *smooth* evaluation sets defined in Definition 9. We apply our results to prove the soundness of FRI.

6.1 The batched FRI protocol

Let $L^{(0)}$ be a smooth domain, $0 < \rho < 1$. For a given function $f^{(0)} : L^{(0)} \to \mathbb{F}_q$, the verifier wants to know whether it is a member of $\mathrm{RS}[\mathbb{F}_q, L^{(0)}, \rho]$. An untrusted prover may help the verifier while the verifier has query accesses to $f^{(0)}$. The verifier and the prover agree on a series of smooth evaluation domains $L^{(0)}, L^{(1)}, \cdots, L^{(n_r)}$, where n_r is the number of interactive rounds. For convenience, we will focus on a specific case of folding where $L^{(k+1)} = (L^{(k)})^2$. For general cases, the definition of folding can be found in [Ben+18], and we will not discuss those details here.

The FRI protocol has two phases, called COMMIT and QUERY.

In the COMMIT phase, the prover and the verifier work together round by round to *fold* the target function $f^{(0)}$ into a field element (or a short vector). Thus, the verifier can check the element easily. In the k^{th} round, the prover sends the oracle of a function $f^{(k)}$ to the verifier. The verifier randomly selects a folding parameter $\alpha^{(k)} \in \mathbb{F}_q$ and sends it to the prover. In this context, we assume that the folding parameter cannot be zero. Upon receiving $\alpha^{(k)}$, the prover folds $f^{(k)}$ using this parameter to obtain a new function $f^{(k+1)} : L^{(k+1)} \to \mathbb{F}_q$. If the prover is honest, the folding result is supposed to be

$$f^{(k+1)} = \mathrm{FuncFold}_{\alpha^{(k)}}(f^{(k)}).$$

If $f^{(k)}$ is a member of $\mathrm{RS}[\mathbb{F}_q, L^{(k)}, \rho]$, then the degree of $f^{(k+1)}$ is expected to be halved. Consequently, any member of $\mathrm{RS}[\mathbb{F}_q, L^{(0)}, \rho]$ will be folded into a single element after $\log(\rho|L^{(0)}|)$ rounds.

In the QUERY phase, the verifier queries some random locations in $L^{(0)}$, and the prover responds with the queried elements as well as those involved in the folding path. The verifier then calculates the folding results to verify the correctness of the folding process.

Batching Batched FRI is a generalization of the FRI protocol. Instead of checking only one function $f^{(0)}$ is near $\mathrm{RS}[\mathbb{F}_q, L^{(0)}, \rho]$, the prover is now required to prove a series of functions $f_0^{(0)}, \cdots, f_l^{(0)} : L^{(0)} \to \mathbb{F}_q$ are near $\mathrm{RS}[\mathbb{F}_q, L^{(0)}, \rho]$. A trivial strategy is checking each function individually; however, this approach becomes inefficient when the number of functions is large. Batched FRI provides a way to do the verifications at one time.

Suppose the prover has a series of functions $f_0^{(0)}, \cdots, f_l^{(0)} \in \mathbb{F}_q^{L^{(0)}}$, and the verifier has oracle access to these functions. Before executing the FRI protocol, the verifier randomly selects $z_1, \cdots, z_l \in \mathbb{F}_q$ and sends them to the prover. The prover and the verifier then run the FRI protocol on the combined function defined as $f^{(0)} \triangleq f_0^{(0)} + z_1 f_1^{(0)} + \cdots + z_l f_l^{(0)}$. The COMMIT phase of the batched FRI protocol is the same as the basic FRI protocol, while the QUERY phase includes additional checks to verify that the combination is correct. More precisely, the batched FRI protocol works as follows:

BATCH **Phase:**

1. The verifier picks uniformly random $z_1, \ldots, z_l \in \mathbb{F}_q^{\times}$.
2. Set $f^{(0)} \triangleq f_0^{(0)} + z_1 f_1^{(0)} + \cdots + z_l f_l^{(0)}$.

COMMIT **Phase:**

1. For each $k \in [0, n_r - 1]$:
   (a) The verifier picks a uniformly random $\alpha^{(k)} \in \mathbb{F}_q^{\times}$.

(b) The prover writes down a function

$$f^{(k+1)} : L^{(k+1)} \to \mathbb{F}_q$$

and sends the oracle of $f^{(k+1)}$ to the verifier. For an honest prover, we have $f^{(k+1)} = \text{FuncFold}_{\alpha^{(k)}}(f^{(k)})$.

2. The prover writes down a value $C \in \mathbb{F}_q$.

QUERY **Phase:** Repeat t times:

1. The verifier picks a uniformly random $s^{(0)} \in L^{(0)}$.
2. If $f^{(0)}(s^{(0)}) \neq f_0^{(0)}(s^{(0)}) + z_1 f_1^{(0)}(s^{(0)}) + \cdots + z_l f_l^{(0)}(s^{(0)})$, REJECT.
3. For each $i \in [0, n_r - 1]$:
   (a) Define $s^{(k+1)} \in L^{(k+1)}$ by $s^{(k+1)} = (s^{(k)})^2$.
   (b) Compute $\text{FuncFold}_{\alpha^{(k)}}(f^{(k)})(s^{(k+1)})$ by making queries to $f^{(k)}(s^{(k)})$ and $f^{(k)}(-s^{(k)})$.
   (c) If $\text{FuncFold}_{\alpha^{(k)}}(f^{(k)})(s^{(k+1)}) \neq f^{(k+1)}(s^{(k+1)})$, REJECT.
4. If $f^{(n_r)}(s^{(n_r)}) \neq C$, REJECT.
5. ACCEPT.

6.2 Soundness of batched FRI

The soundness error of batched FRI consists of bad batching, bad folding, and prover's cheating.

The bad batching is restricted by Theorem 6 directly. We propose an analysis of the possibility of bad folding based on Corollary 1. Let $f^{(0)} : L^{(0)} \to \mathbb{F}_q$ be the initial function. Without loss of generality, we state the case $q(X) = X^2$ here. The prover and the verifier have agreed on a series of "smooth" evaluation domains, $L^{(0)}, L^{(1)}, \ldots$. Suppose there are n_r rounds in the FRI protocol. Let

$$\delta^{(k)} \triangleq \delta(f^{(k)}, \text{RS}[\mathbb{F}_q, L^{(k)}, \rho])$$

be the relative distance.

Let $B^{(k)} = \min\{\delta^{(k)} - \frac{1}{|L^{(k+1)}|}, 1 - \sqrt[3]{\rho} - \eta\}$. Let $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. Define the k^{th} *Bad Event* $E^{(k)}, 0 \leq k \leq n_r - 1$ as the event:

$$E^{(k)} = \{\alpha^{(k)} \in \mathbb{F}_q : \delta(\text{FuncFold}_{\alpha^{(k)}}(f^{(k)}), \text{RS}[\mathbb{F}_q, L^{(k+1)}, \rho]) \leq B^{(k)}\}$$

where $\alpha^{(k)}$ is the random folding point chosen by the verifier in the k^{th} round. According to Corollary 1, we have

$$\mathbb{P}(E^{(k)}) < \frac{(1-\rho)|L^{(k)}|}{9\rho\eta^3|\mathbb{F}_q|}.$$

Notice that $|L^{(k+1)}| = \frac{|L^{(k)}|}{2}$. Then the possibility that in all of the n_r rounds, the bad events do not happen satisfies:

$$\begin{aligned}\mathbb{P}(\bigwedge_{k=0}^{n_r-1} \neg E^{(k)}) &\geq 1 - \sum_{k=0}^{n_r-1} \mathbb{P}(E^{(k)}) \\ &> 1 - \frac{2(1-\rho)|L^{(0)}|}{9\rho\eta^3|\mathbb{F}_q|}.\end{aligned}$$

Suppose bad batching and bad folding do not happen in all the n_r rounds. A dishonest prover may modify some locations of the codeword to pass the verification. However, modifications will be checked during the QUERY phase and can not increase the possibility of passing. As a result, we have the following soundness error bound of batched FRI.

Suppose that bad batching and bad folding do not occur in any of the n_r rounds. A dishonest prover may alter some positions of the codeword to pass verification; however, these modifications will be scrutinized during the QUERY phase and cannot increase the likelihood of passing. Consequently, we derive the following soundness error bound for batched FRI. Further details regarding the soundness of FRI can be found in [Ben+18], which offers a comprehensive soundness analysis.

Theorem 9 (Batched FRI soundness)**.** *Let $\mathbb{F}_q$ be a finite field. Let $L^{(0)} \subseteq \mathbb{F}_q$ be a smooth evaluation domain.*

Let $f_0^{(0)}, \cdots, f_l^{(0)} : L^{(0)} \to \mathbb{F}_q, 1 \leq l$ be a sequence of functions and let $V^{(0)} = \mathrm{RS}[\mathbb{F}_q, L^{(0)}, \rho]$ and ρ satisfies $\rho = 2^{-R}$ for a positive integer R. Let $\delta, \eta > 0$ satisfy $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\eta \leq \frac{1}{3}\rho^{-\frac{1}{3}}$. Furthermore, let t denote the number of invocations of the FRI QUERY step.

Suppose there exists a batched FRI prover P^* that interacts with the batched FRI verifier and causes it to output "accept" with a probability greater than

$$\epsilon_{\textit{Batched-FRI}} = \left(\frac{2(1-\rho)|L^{(0)}|}{9\rho\eta^3|\mathbb{F}_q|}\right) \cdot (l+1) + (1-\delta)^t \tag{33}$$

Then $f_0^{(0)}, \cdots, f_l^{(0)}$ have correlated agreement with $V^{(0)}$ on a domain $D \subseteq L^{(0)}$ of density at least $1-\delta$.

Remark 3. *For general cases that $q(X) = X^{2^k}, k \in \mathbb{N}^*$, this error bound also holds. One folding in this case can be seen as k foldings of the special case with the same folding parameter.*

6.3 Numerical Example

We provide a numerical example to show the improvement in the provable soundness of FRI. Set $q = |\mathbb{F}_q| > 2^{183}$ (the extension field used in [Sta23]), $\rho = \frac{1}{8}$, $m = 3$, $\eta = 2^{-6}$, $|L^{(0)}| = 2^{24}$ and $l = 28$. $n_r = \log_2(|L^{(0)}|) = 24$ is the number of rounds. t is the number of QUERY times.

Let

$$\epsilon_c \triangleq \frac{(m+\frac{1}{2})^7 \cdot |L^{(0)}|^2}{2\rho^{3/2}q} + \frac{(2m+1)\cdot(|L^{(0)}|+1)}{\sqrt{\rho}} \cdot \frac{\sum_{i=0}^{n_r-1} l^{(i)}}{q},$$

where $l^{(i)} = \frac{|L^{(i)}|}{|L^{(i+1)}|} = 2$ in our example. And we have

$$2^{-122} < \epsilon_c < 2^{-121}.$$

The soundness error bound provided in [Ben+20b] is

$$\epsilon_{\text{Batched-FRI}} = \epsilon_c + \left(\sqrt{\rho}(1+\frac{1}{2m})\right)^t. \tag{34}$$

This can reach 121 bits of security when $t \geq 97$, and can not reach 128 bits of security. For higher security levels, we can apply the FRI protocol in a bigger extension field. However, this will increase

the cost of operations. Our soundness error bound is provided in (33). And we have

$$2^{-136} < \left(\frac{|L^{(0)}|}{\sqrt{\rho}\eta^2 |\mathbb{F}_q|} \right) \cdot (l+1) < 2^{-135}.$$

We prove FRI can reach 128 bits of security in the current field when $t \geq 134$.

Acknowledgment

We thank Swastik Kopparty for pointing out a critical mistake in the earlier version of this work, and for his invaluable discussions. His insights have greatly assisted us in completing this work.

References

[Ame+17] Scott Ames et al. "Ligero: Lightweight sublinear arguments without a trusted setup". In: *Proceedings of the 2017 acm sigsac conference on computer and communications security*. 2017, pp. 2087–2104.

[Arn+24] Gal Arnon et al. "STIR: Reed–Solomon Proximity Testing with Fewer Queries". In: *Cryptology ePrint Archive* (2024).

[Ben+20a] E Ben-Sasson et al. "DEEP-FRI: Sampling Outside the Box Improves Soundness". In: *11th Innovations in Theoretical Computer Science Conference, ITCS 2020*. LIPIcs Dagstuhl. 2020.

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

[BKS18] Eli Ben-Sasson, Swastik Kopparty, and Shubhangi Saraf. "Worst-case to average case reductions for the distance to a code". In: *33rd Computational Complexity Conference (CCC 2018)*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik. 2018.

[Ben+18] Eli Ben-Sasson et al. "Fast reed-solomon interactive oracle proofs of proximity". In: *45th international colloquium on automata, languages, and programming (icalp 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. 2018.

[Ben+19] Eli Ben-Sasson et al. "Aurora: Transparent succinct arguments for R1CS". In: *Advances in Cryptology–EUROCRYPT 2019: 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt, Germany, May 19–23, 2019, Proceedings, Part I 38*. Springer. 2019, pp. 103–128.

[Ben+20b] Eli Ben-Sasson et al. "Proximity gaps for Reed–Solomon codes". In: *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE. 2020, pp. 900–909.

[Blo+23] Alexander R Block et al. "Fiat-Shamir security of FRI and related snarks". In: *International Conference on the Theory and Application of Cryptology and Information Security*. Springer. 2023, pp. 3–40.

[Juk11] Stasys Jukna. *Extremal combinatorics: with applications in computer science*. Vol. 571. Springer, 2011.

[KPV22] Assimakis A Kattis, Konstantin Panarin, and Alexander Vlasov. "RedShift: transparent SNARKs from list polynomial commitments". In: *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*. 2022, pp. 1725–1737.

[Pol] Polygon. *Plonky2: Fast recursive arguments with plonk and fri*. https://github.com/mir-protocol/plonky2/tree/main/plonky2. URL: https://github.com/mir-protocol/plonky2/tree/main/plonky2.

[RS60] Irving S Reed and Gustave Solomon. "Polynomial codes over certain finite fields". In: *Journal of the society for industrial and applied mathematics* 8.2 (1960), pp. 300–304.

[RRR16] Omer Reingold, Guy N Rothblum, and Ron D Rothblum. "Constant-round interactive proofs for delegating computation". In: *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 2016, pp. 49–62.

[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*. 2013, pp. 793–802.

[Sta23] StarkWare. *ethSTARK Documentation v1.2*. Cryptology ePrint Archive, Paper 2021/582. https://eprint.iacr.org/2021/582. 2023. URL: https://eprint.iacr.org/2021/582.

[Xie+22] Tiancheng Xie et al. "zkbridge: Trustless cross-chain bridges made practical". In: *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*. 2022, pp. 3003–3017.

[ZCF23] Hadas Zeilberger, Binyi Chen, and Ben Fisch. "BaseFold: Efficient Field-Agnostic Polynomial Commitment Schemes from Foldable Codes". In: *Cryptology ePrint Archive* (2023).

[Zha+20] Jiaheng Zhang et al. "Transparent polynomial delegation and its applications to zero knowledge proof". In: *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2020, pp. 859–876.

A Proof of Theorem 7

Definition 13. *Let $L \in \mathbb{F}_q^\times$ be a set with pairing elements. Let $f : L \to \mathbb{F}_q$ be a function. Let $0 < \delta \leq 1$ and $\mathrm{Bad}_{\mathrm{L}}(f)$ is defined in Definition 11. For any $\alpha \in \mathrm{Bad}_{\mathrm{L}}(f)$, $\exists v \in \mathrm{RS}[\mathbb{F}_q, L^2, \rho]$ such that $\delta(\mathrm{FuncFold}_\alpha(f), v) \leq \delta$ and $\mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \neq D_i^2, i = 1, \ldots, m$. If more than one codewords satisfy these conditions, choose the one with the smallest lexicographical order. Define*

- $\mathrm{Closest}_L(\delta, \mathrm{FuncFold}_\alpha(f), \mathrm{RS}[\mathbb{F}_q, L^2, \rho]) = v$.
- $C^*_{L,\alpha} = \mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \subseteq L^2$.
- $P^*_{L,\alpha} = \{x \mid x^2 \in C^*_{L,\alpha}\} \subseteq L$ *is the parent set of* $C^*_{L,\alpha}$.
- $C_{L,\alpha}$*: The set of the first* $(1-\delta)|L^2|$ *elements of* $C^*_{L,\alpha}$*, i.e.,* $|C_{L,\alpha}| = (1-\delta)|L^2|$.

- $P_{L,\alpha} = \{x \mid x^2 \in C_{L,\alpha}\} \subseteq L$ *is the parent set of* $C_{L,\alpha}$.

Lemma 8. *Let* $L \in \mathbb{F}_q^\times$ *be a set with pairing elements. Let* $f : L \to \mathbb{F}_q$ *be a function. Let* $0 < \delta \leq 1$ *and* $\mathrm{Bad}_{\mathrm{L}}(f)$ *is defined in Definition 11. For each* $C^*_{L,\alpha}$*, we prove that there are two possible cases:*

1. $|C^*_{L,\alpha} \cap D_i^2| < \rho|L^2|, i \in \{1, \ldots, m\}$*, or*
2. $\exists D_i \in \mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{F}_q,L,\rho]}$*, such that* $C^*_{L,\alpha} \supsetneq D_i^2$.

Proof. If $\exists D_i, |C^*_{L,\alpha} \cap D_i^2| \geq \rho|L^2|$, denote by p the low-degree polynomial f agrees on D_i. According to Definition 9, $x \in D_i \iff -x \in D_i$, thus

$$\mathrm{FuncFold}_\alpha(f)|_{D_i^2} = \mathrm{PolyFold}_\alpha(p)|_{D_i^2}.$$

Notice that

$$\mathrm{PolyFold}_\alpha(p)|_{C^*_{L,\alpha} \cap D_i^2} = v|_{C^*_{L,\alpha} \cap D_i^2}$$

because of (24). Since $|C^*_{L,\alpha} \cap D_i^2| \geq \rho|L^2|$, we have $\mathrm{PolyFold}_\alpha(p) = v$. Then

$$\mathrm{FuncFold}_\alpha(f)|_{D_i^2} = \mathrm{PolyFold}_\alpha(p)|_{D_i^2} = v|_{D_i^2}.$$

As a result, we have $C^*_{L,\alpha} = \mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \supseteq D_i^2$. Definition 11 tells us $\mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \neq D_i^2$, thus,

$$\mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \supsetneq D_i^2.$$

□

Lemma 9. *For any distinct* $\alpha_1, \alpha_2 \in \mathrm{Bad}_{\mathrm{L}}(f)$*, there exists a polynomial* $p(X) \in \mathbb{F}_q[X]$ *of degree strictly less than* $\rho|L|$ *such that*

$$f|_{P_{L,\alpha_1} \cap P_{L,\alpha_2}} = p|_{P_{L,\alpha_1} \cap P_{L,\alpha_2}}, \tag{35}$$

that is, $f(x) = p(x)$ *for any* $x \in P_{L,\alpha_1} \cap P_{L,\alpha_2}$.

Moreover, if $|P_{L,\alpha_1} \cap P_{L,\alpha_2}| \geq \rho|L|$*, polynomial* $p(X)$ *is uniquely determined. (If* $|P_{L,\alpha_1} \cap P_{L,\alpha_2}| < \rho|L|$*, the existence of* $p(X)$ *is obvious.)*

Furthermore, we have

$$f|_{P^*_{\alpha_1} \cap P^*_{\alpha_2}} = p|_{P^*_{\alpha_1} \cap P^*_{\alpha_2}}. \tag{36}$$

Proof. Notice that if (36) holds, then we have (35) because $P_{L,\alpha_1} \cap P_{L,\alpha_2} \subseteq P^*_{L,\alpha_1} \cap P^*_{L,\alpha_2}$. So we only prove (36). By Definition 13, there exists a polynomial

$$v_1(X) = \mathrm{Closest}_L(\delta, \mathrm{FuncFold}_{\alpha_1}(f), \mathrm{RS}[\mathbb{F}_q, L^2, \rho]) \in \mathbb{F}_q[X]$$

of degree strictly less than $\rho|L|/2$ such that $\mathrm{FuncFold}_{\alpha_1}(f)|_{C^*_{L,\alpha_1}} = v_1|_{C^*_{L,\alpha_1}}$. By Definition 4, we have

$$v_1(x^2) = \frac{f(x) + f(-x)}{2} + \alpha_1 \cdot \frac{f(x) - f(-x)}{2x} \tag{37}$$

for any $x \in P^*_{L,\alpha_1}$. Similarly, there exists a polynomial $v_2(X) \in \mathbb{F}_q[X]$ of degree strictly less than $\rho|L|/2$ such that $\text{FuncFold}_{\alpha_2}(f)|_{C^*_{L,\alpha_2}} = v_2|_{C^*_{L,\alpha_2}}$. And we have

$$v_2(x^2) = \frac{f(x)+f(-x)}{2} + \alpha_2 \cdot \frac{f(x)-f(-x)}{2x} \tag{38}$$

for any $x \in P^*_{L,\alpha_2}$.

From (37) and (38), we have

$$\begin{cases} v_1(x^2) - v_2(x^2) = \dfrac{\alpha_1-\alpha_2}{2x} \cdot (f(x)-f(-x)) \\ \alpha_2 v_1(x^2) - \alpha_1 v_2(x^2) = \dfrac{\alpha_2-\alpha_1}{2} \cdot (f(x)+f(-x)) \end{cases}$$

for any $x \in P^*_{L,\alpha_1} \cap P^*_{L,\alpha_2}$. That is,

$$\begin{cases} f(x)-f(-x) = \dfrac{2x}{\alpha_1-\alpha_2} \cdot \big(v_1(x^2) - v_2(x^2)\big) \\ f(x)+f(-x) = \dfrac{2}{\alpha_2-\alpha_1} \cdot \big(\alpha_2 v_1(x^2) - \alpha_1 v_2(x^2)\big) \end{cases}$$

Therefore, we have

$$f(x) = \frac{x}{\alpha_1-\alpha_2} \cdot \big(v_1(x^2) - v_2(x^2)\big) + \frac{1}{\alpha_2-\alpha_1} \cdot \big(\alpha_2 v_1(x^2) - \alpha_1 v_2(x^2)\big). \tag{39}$$

for any $x \in P^*_{L,\alpha_1} \cap P^*_{L,\alpha_2}$.

Note that $\deg(v_1), \deg(v_2) \le \frac{1}{2} \cdot \rho|L| - 1$. From (39), we have

$$\begin{aligned} \deg(f) &\le 1 + 2\max(\deg(v_1), \deg(v_2)) \\ &\le 1 + \rho|L| - 2 \\ &= \rho|L| - 1. \end{aligned}$$

If $|P^*_{L,\alpha_1} \cap P^*_{L,\alpha_2}| \ge \rho|L|$, polynomial $p(X)$ is unique, since $\rho|L|$ points uniquely determine a polynomial of degree at most $\rho|L| - 1$. □

Lemma 10. *Let $\alpha, \beta_1, \beta_2 \in \text{Bad}_{\text{L}}(f)$ be different such that $|P^*_{L,\alpha} \cap P^*_{L,\beta_1}| \ge \rho|L|$ and $|P^*_{L,\alpha} \cap P^*_{L,\beta_2}| \ge \rho|L|$. Denote by $p_1(X)$ and $p_2(X)$ the polynomials of degree at most $\rho|L|-1$ decided by $f|_{P^*_{L,\alpha} \cap P^*_{L,\beta_1}}$ and $f|_{P^*_{L,\alpha} \cap P^*_{L,\beta_2}}$ (Lemma 9). Then exactly one of the followings holds:*

- *$p_1 = p_2$ and $P^*_{L,\alpha} \cap P^*_{L,\beta_1} = P^*_{L,\alpha} \cap P^*_{L,\beta_2}$.*
- *$p_1 \ne p_2$ and $|P_{L,\alpha} \cap P_{L,\beta_1} \cap P_{L,\beta_2}| \le |P^*_{L,\alpha} \cap P^*_{L,\beta_1} \cap P^*_{L,\beta_2}| \le \rho|L| - 1$.*

Proof. The proof is the same as Lemma 5. □

Run Algorithm 2 on $\text{Bad}_{\text{L}}(f)$ and the corresponding parent sets to give a partition on $\text{Bad}_{\text{L}}(f)$. Denote the output as $A'_1, \ldots, A'_r$ and $\alpha_1, \ldots, \alpha_r$. Lemma 3 still holds because our partition strategy is unchanged and we have $|C_{L,\alpha}| = (1-\delta)|L^2|$ for any $\alpha \in \text{Bad}_{\text{L}}(f)$. So if we can bound the size of each block A_i, we finish our proof of the theorem.

Let $A \in \{A'_1, \ldots, A'_r\}$ and α be the corresponding folding point, i.e., for any $\beta \in A$, we have $|P_{L,\beta} \cap P_{L,\alpha}| \geq \left((1-\delta)^2 - \eta'\right)|L| \geq \rho|L|$. Since Lemma 10 holds, The relation $\mathcal{R}$ on $A \setminus \{\alpha\}$:

$$(\beta_1, \beta_2) \in \mathcal{R} \iff p_1 = p_2,$$

where p_1 and p_2 are the low-degree polynomials determined by $f|_{P_{L,\alpha} \cap P_{L,\beta_1}}$ and $f|_{P_{L,\alpha} \cap P_{L,\beta_2}}$ respectively. $\mathcal{R}$ is still an equivalence relation and gives a partition on set $A \setminus \{\alpha\}$. Let s be the number of equivalence classes.

Corollary 7. *When $\delta \leq 1 - \sqrt[3]{\rho} - \eta$ and $\eta' = \frac{3\sqrt[3]{\rho}\eta^2}{2}$. Let $A \in \{A'_1, \ldots, A'_r\}$ be a set of folding points defined in Algorithm 2 and α be the corresponding folding point. Let s be the number of equivalence classes decided by $\mathcal{R}$ on $A \setminus \{\alpha\}$. We have $s \leq \frac{1}{3\eta\rho^{\frac{2}{3}}}$.*

Proof. The proof is the same as Corollary 4. □

We modify Lemma 6 into the following lemma.

Lemma 11. *Let $\alpha, \beta_1, \ldots, \beta_t \in \mathrm{Bad_L}(f)$ be distinct such that*

- $P^*_{L,\alpha} \cap P^*_{L,\beta_1} = P^*_{L,\alpha} \cap P^*_{L,\beta_2} = \ldots = P^*_{L,\alpha} \cap P^*_{L,\beta_t}$, *and*
- $|P^*_{L,\alpha} \cap P^*_{L,\beta_1}| \geq |P_{L,\alpha} \cap P_{L,\beta_1}| \geq \rho|L|$.

Then $P^*_{L,\beta_1}, \ldots, P^*_{L,\beta_t}$ form a sunflower with core $P^*_{L,\alpha} \cap P^*_{L,\beta_1}$. That is, for any distinct $i, j \in \{1, \ldots, t\}$, we have

$$P^*_{L,\beta_i} \cap P^*_{L,\beta_j} = P^*_{L,\alpha} \cap P^*_{L,\beta_1}.$$

The proof of Lemma 11 is the same as Lemma 6. We omit the details here. Now we can bound the number of elements in each equivalence class. The following lemma has the same result as Corollary 5, but the proof is slightly different.

Corollary 8. *Denote by t the upper bound of the number of elements in each equivalence class. Then we have $t \leq \frac{\delta}{2}|L|$.*

Proof. Let

$$[\beta] = \{\beta_1, \ldots, \beta_t\}$$

be an equivalence class. According to the definition of equivalent relation, we have $P^*_{L,\alpha} \cap P^*_{L,\beta_1} = \ldots = P^*_{L,\alpha} \cap P^*_{L,\beta_t}$. According to Algorithm 2, we have $|P^*_{L,\alpha} \cap P^*_{L,\beta_1}| \geq (\sqrt{\rho} + \eta')|L| > \rho|L|$. Lemma 11 tells us

$$P^*_{L,\beta_i} \cap P^*_{L,\beta_j} = P^*_{L,\alpha} \cap P^*_{L,\beta_1}, 1 \leq i < j \leq t.$$

This implies

$$\left(P^*_{L,\beta_i} \setminus P^*_{L,\alpha}\right) \cap \left(P^*_{L,\beta_j} \setminus P^*_{L,\alpha}\right) = \emptyset, 1 \leq i < j \leq t.$$

Thus,

$$\delta|L| \geq |L \setminus P^*_{L,\alpha}| \geq \left|\bigcup_{i=1}^{t} \left(P^*_{L,\beta_i} \setminus P^*_{L,\alpha}\right)\right| = \sum_{i=1}^{t} \left|P^*_{L,\beta_i} \setminus P^*_{L,\alpha}\right| \tag{40}$$

According to Lemma 9, f agrees with a low-degree polynomial p on $P^*_{L,\alpha} \cap P^*_{L,\beta_i}$. Since $x \in P^*_{L,\alpha} \cap P^*_{L,\beta_i} \iff -x \in P^*_{L,\alpha} \cap P^*_{L,\beta_i}$, $\exists D_i \in \mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{F}_q,L,\rho]}$, such that $P^*_{L,\alpha} \cap P^*_{L,\beta_i} \subseteq D_i$. According to the definition of $\mathrm{Bad}_{\mathrm{L}}(f)$ and P^*_{L,β_i}, $P^*_{L,\beta_i} \neq D_i$. Furthermore, the number of elements in $P^*_{L,\beta_i} \setminus P^*_{L,\alpha}$ is even. As a result,

$$|P^*_{L,\beta_i} \setminus P^*_{L,\alpha}| = |P^*_{L,\beta_i} \setminus (P^*_{L,\alpha} \cap P^*_{L,\beta_i})| \geq 2. \tag{41}$$

Plugging (41) into (40), we have

$$\delta|L| \geq 2t \Rightarrow t \leq \frac{\delta}{2}|L|.$$

□