Improved Soundness Analysis of the FRI Protocol

Yiwen Gao, Haibin Kan, and Yuan Li

Fudan University, Shanghai, China
ywgao21@m.fudan.edu.cn, hbkan@fudan.edu.cn, yuan_li@fudan.edu.cn

Abstract. We enhance the provable soundness of FRI, an interactive oracle proof of proximity (IOPP) for Reed-Solomon codes introduced by Ben-Sasson et al. in ICALP 2018. More precisely, we prove the soundness error of FRI is less than $\max\left\{O\left(\frac{1}{\eta}\cdot\frac{n}{|\mathbb{F}_q|}\right),(1-\delta)^t\right\}$, where $\delta\leq 1-\sqrt{\rho}-\eta$ is within the Johnson bound and $\mathbb{F}_q$ is a finite field with characteristic greater than 2. Previously, the best-known provable soundness error for FRI was $\max\left\{O\left(\frac{\rho^2}{\eta^7}\cdot\frac{n^2}{|\mathbb{F}_q|}\right),(1-\delta)^t\right\}$, as established by Ben-Sasson et al. in FOCS 2020.

We prove the number of *bad* folding points in FRI is linear in the length n of codeword when it is δ-far from the Reed-Solomon code. This implies the linear proximity gaps for Reed-Solomon codes and improves the provable soundness of batched FRI. Our results indicate that the FRI protocol can be implemented over a smaller field, thereby enhancing its efficiency. Furthermore, for a fixed finite field $\mathbb{F}_q$, we prove that FRI can achieve improved security.

Keywords: FRI protocol · Reed-Solomon proximity testing.

1 Introduction

Reed-Solomon (RS) codes [15] are a class of error-correcting codes. They are fundamental objects of study in algebraic coding theory and theoretical computer science. Recently, they have been widely used in the design of zero-knowledge proofs. Let $\mathbb{F}_q$ be a finite field with q elements, and let $L\subseteq\mathbb{F}_q$ be the evaluation domain. Denote $\rho\in(0,1]$ as the rate. The notation $\mathrm{RS}[\mathbb{F}_q,L,\rho]$ represents the set of functions $f:L\to\mathbb{F}_q$ that are evaluation results of polynomials of degree strictly less than $\rho|L|$.

The RS proximity testing problem asks the verifier to determine 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 few query accesses to f. An untrusted prover may help the verifier. We consider this problem under the interactive oracle proofs of proximity (IOPP) model [6] (also called probabilistically checkable interactive proofs of proximity in [16]). 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. The RS proximity testing problem can be used to construct

zk-STARKs (zero-knowledge scalable and transparent arguments of knowledge) [3][7].

In particular, the Fast RS IOPP (known as FRI)[2] 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 [5][13][17][14][20][18]. Previously, the best provable soundness of FRI was established by Ben-Sasson, Carmon, Ishai, Kopparty, and Saraf in [4], utilizing elegant mathematical techniques. Furthermore, they demonstrate a stronger property of the FRI protocol known as correlated agreement, which offers an enhanced version of soundness. The correlated agreement of FRI is used to prove the (knowledge) soundness of protocols that use FRI as a sub-protocol [17], as well as to prove the round-by-round soundness of FRI [17][9]. There are also a variety of generalized protocols of FRI [19][1]. These works build upon the soundness of FRI, and improvements to the results in [4] can also be applied to these related works.

Ben-Sasson et al. also propose a conjecture on the soundness of FRI in [4] (Conjecture 1 in Section 5.1). Many real-world deployments of FRI rely on this conjecture. However, there is a gap between the conjectured soundness and the provable soundness of FRI [10]. It remains open to prove or disprove this conjecture.

1.1 Our results

We present linear proximity gaps for Reed-Solomon codes (Section 5.1). We also improve the provable soundness for the (batched) FRI protocol over any finite field with characteristics greater than 2. Our results provide a part of the proof of the conjecture. Furthermore, our results cover the power of correlated agreement.

Theorem 1 (Informal). *Let $\mathbb{F}_q$ be a finite field with characteristic* $\mathrm{char}(\mathbb{F}_q) > 2$ *Let n be the code length. Let $0 < \rho \leq 1$ be the rate of the Reed-Solomon codes. Let $\delta, \eta > 0$ satisfy $\delta \leq \delta_0(\rho, \eta) := 1 - \sqrt{\rho} - \eta$ and $t \in \mathbb{Z}^+$ is the repeating time of the query phase in FRI. Then the soundness of FRI, denoted as ϵ_{FRI}, satisfies*

$$\epsilon_{FRI} = \max\left\{O\left(\frac{1}{\eta} \cdot \frac{n}{|\mathbb{F}_q|}\right), (1-\delta)^t\right\}.$$

The formal version of the above theorem is presented in Section 5.2, which establishes the provable soundness of *batched FRI*. In the FRI protocol, the verifier chooses a set of random points and sends them to the prover. Our main theorem in Section 4 describes a property of the Reed-Solomon codes, which limits the number of points that will cause the loss of soundness of FRI. Based on this property, we improve the provable soundness of FRI.

Comparison to prior works. The previous best provable soundness of FRI is provided by Ben-Sasson et al. in [4]. They prove that for some $\delta \in (0, 1 - \sqrt{\rho} - \eta]$ with $\eta \leq \frac{\sqrt{\rho}}{20}$, the soundness error of FRI satisfies

$$\epsilon_{\mathrm{FRI}} = \max\left\{O\left(\frac{\rho^2}{\eta^7} \cdot \frac{n^2}{|\mathbb{F}_q|}\right), (1-\delta)^t\right\}.$$

Linear in code length: We improve the soundness of FRI from quadratic to linear in the code length. This new provable soundness guarantees that FRI and its variants can be safely implemented in a much smaller finite field, thereby reducing operational costs.
Under the Johnson bound: Both of our results and the results in [4] require the distance δ to be under the Johnson bound. In [4], this bound is required due to the use of the Guruswami-Sudan list decoding algorithm[11]. Our proofs do not rely on this algorithm. Instead, we use a combinatorial lemma by Corrádi [12], which coincides with the Johnson bound.
Improvement in coefficient: Our result also improves the coefficient of the first item: from $\frac{\rho^2}{\eta^7}$ to $\frac{1}{\eta}$. This coefficient plays a role when δ is close to the Johnson bound.
Characteristic of the field: We focus on finite fields with characteristics greater than 2. This is because our proof relies on the folding structure in the FRI protocol. When using the characteristic 2 finite field, the corresponding evaluation domain is expressed in an additive version. Our result does not work in the additive version; see Appendix B for details.

2 Proof overview

In this section, we outline the basic idea behind our proof of the main result, Theorem 2.

Our proof is based on the property of Reed-Solomon codes and the *folding* operation in FRI. Over a finite field $\mathbb{F}_q$ and evaluation domain L, where $|L| = n$. The folding operation folds a codeword $f : L \to \mathbb{F}_q$ of length n into a codeword $\text{FuncFold}_\alpha(f)$ of length $n/2$, where $\alpha \in \mathbb{F}_q^\times$ is a folding parameter. Let $0 < \rho \leq$ be the rate of the RS code. If f is a member of $\text{RS}[\mathbb{F}_q, L, \rho]$, then $\text{FuncFold}_\alpha(f)$ is a member of $\text{RS}[\mathbb{F}_q, L^2, \rho]$, where $L^2 = \{x^2 \mid x \in L\}$. However, if f is not a member of $\text{RS}[\mathbb{F}_q, L, \rho]$, for some α, $\text{FuncFold}_\alpha(f)$ can be a member of $\text{RS}[\mathbb{F}_q, L^2, \rho]$. We observe that if a codeword f can be folded into $\text{RS}[\mathbb{F}_q, L^2, \rho]$ through more than one folding points, then $f \in \text{RS}[\mathbb{F}_q, L, \rho]$.

Based on this observation, we restrict the number of *bad* folding points that *close* the distance between f and $\text{RS}[\mathbb{F}_q, L, \rho]$. More precisely, suppose f is δ-far from $\text{RS}[\mathbb{F}_q, L, \rho]$, if the relative distance between $\text{FuncFold}_\alpha(f)$ and $\text{RS}[\mathbb{F}_q, L^2, \rho]$ is smaller than δ, then we say α is a bad folding point.

Denote by $\text{Closest}(\text{FuncFold}_\alpha(f)) \in \text{RS}[\mathbb{F}_q, L^2, \rho]$ the closest codeword of $\text{FuncFold}_\alpha(f)$. Use C_α to denote the sub-evaluation domain that $\text{FuncFold}_\alpha(f)$ agrees with $\text{Closest}(\text{FuncFold}_\alpha(f))$. Let P_α be the parent set of C_α, i.e., $P_\alpha = \{x \mid x^2 \in C_\alpha\}$. Suppose α and β are two distinct bad folding points. Our observation says that f agrees with some low-degree polynomial p on $P_\alpha \cap P_\beta$ with $\deg(p) < \rho|L|$. If $|P_\alpha \cap P_\beta| \geq \rho|L|$ then p is uniquely determined.

For a fixed folding point α, let $A_\alpha = [\beta_1, \ldots, \beta_t]$ be the set of bad folding points that $|P_\alpha \cap P_{\beta_i}| \geq \rho|L|$. We prove the number of such bad folding points is linear in $|L|$. More precisely, we prove that $t \leq \frac{\delta|L|}{2}$.

Using a greedy algorithm, we can select a series of representatives $\alpha_1, \ldots, \alpha_r$ such that

1. $\bigcup_{i=1}^{r}(\{\alpha_i\} \cup A_{\alpha_i})$ covers all the bad folding points;
2. $|P_{\alpha_i}| \geq (1-\delta)|L|$ for $i \in \{1, \ldots, r\}$;
3. $|P_{\alpha_i} \cap P_{\alpha_j}| < \rho|L|$ for $i \neq j$.

Notice that for $i \in \{1, \ldots, r\}$, P_{α_i} is a subset of L. Combining this with 2 and 3, Corrádi's lemma tells us when $\delta < 1 - \sqrt{\rho} - \eta$, we have $r \leq \frac{1}{\eta}$, where η is the gap between δ and the Johnson bound.

Therefore, the number of bad folding points is less than $\frac{\delta}{2\eta}|L| + \frac{1}{\eta}$.

3 Preliminaries

We use $\mathbb{F}_q$ to denote a finite field with q elements. We use $\mathrm{char}(\mathbb{F}_q)$ to denote the characteristic of $\mathbb{F}_q$. Throughout the paper, we assume $\mathrm{char}(\mathbb{F}_q) > 2$. 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 smooth set. 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*

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

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*

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

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

$$\text{PolyFold}_\alpha(p)(x^2) = p^{(\text{even})}(x) + \alpha p^{(\text{odd})}(x) \tag{1}$$

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*

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

Proof. By the linearity of PolyFold operator (Prop 1), it suffices to prove (2) 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 (2) 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 (2) is

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

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. *[4] 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 Johnson bound $1 - \sqrt{\rho}$. Let η be the gap between δ and the Johnson bound, i.e., $\delta \leq 1 - \sqrt{\rho} - \eta$. Our goal is to prove that with probability $1 - \frac{\delta|L|+2}{2\eta|\mathbb{F}_q^\times|}$, the folded codeword $\mathrm{FuncFold}_\alpha(f)$ is still δ-far from RS codes, where $\alpha \in \mathbb{F}_q^\times$ 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 $0 < \delta \leq 1$. Let $f : L \to \mathbb{F}_q$ be a codeword such that $\Delta(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$. Define the* bad folding points *to be*

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

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

Theorem 2. *Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt{\rho} - \eta$. Let $f : L \to \mathbb{F}_q$ be a codeword such that $\Delta(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$. Then,*

$$\mathrm{Bad}(f) \leq \frac{\delta}{2\eta}|L| + \frac{1}{\eta}.$$

First, let us introduce a few definitions.

Definition 7. *Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $f : L \to \mathbb{F}_q$ be a codeword.*

- *Let $\mathrm{Closest}(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) \in \mathrm{RS}[\mathbb{F}_q, L, \rho]$ 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.*

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 $r \leq \frac{1}{\eta}$, and $|A_i| \leq \frac{1}{2} \cdot \delta|L| + 1$ for any $i = 1, 2, \ldots, r$. Therefore, $|\text{Bad}(f)| \leq \frac{\delta}{2\eta}|L| + \frac{1}{\eta}$.

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

Algorithm 1 Partition Bad Folding Points

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

4.1 Upper bound on the number of blocks

In this subsection, we limit the number of blocks. We prove that $r \leq \frac{1}{\eta}$ when $\delta \leq 1 - \sqrt{\rho} - \eta$.

Lemma 1 (Corrádi 1969 [12]). *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*

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

Using the above combinatorial result, we can show the number of partitions $r \leq 1/\eta$.

Lemma 2. *Let $L \subseteq \mathbb{F}_q$ be a set with pairing elements. Let $f : L \to \mathbb{F}_q$ be any codeword. For any $\delta \in (0, 1 - \sqrt{\rho} - \eta]$, we have*

$$r \leq \frac{\sqrt{\rho}(1 - \sqrt{\rho}) + \eta}{2\sqrt{\rho}\eta + \eta^2} \leq \frac{1}{\eta}.$$

Proof. We associate a set P_{α_i} for each representative α_i. From Algorithm 1, we know

- $|P_{\alpha_i} \cap P_{\alpha_j}| < \rho|L|$ for any distinct i, j.
- $|P_{\alpha_i}| = 2|C_{\alpha_i}| \geq (1-\delta)|L|$ for any i.

By Lemma 1, we have

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

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

$$|L| \geq \frac{r(1-\delta)^2|L|}{(1-\delta) + (r-1)\rho},$$

that is

$$(1-\delta) + (r-1)\rho \geq r(1-\delta)^2,$$

which implies

$$r \leq \frac{1-\delta-\rho}{(1-\delta)^2 - \rho}. \tag{3}$$

The right-hand side of (3) is increasing in δ when $\delta \in (0, 1-\sqrt{\rho}-\eta]$. So we have

$$\begin{aligned} r &\leq \frac{\sqrt{\rho}+\eta-\rho}{(\sqrt{\rho}+\eta)^2-\rho} = \frac{\sqrt{\rho}(1-\sqrt{\rho})+\eta}{2\sqrt{\rho}\eta+\eta^2} \\ &\leq \frac{\sqrt{\rho}+\eta}{\sqrt{\rho}\eta+\eta^2} = \frac{1}{\eta}. \end{aligned}$$

4.2 Upper bound on the size of each block

In this subsection, our goal is to prove $|A_i| \leq \frac{1}{2} \cdot \delta|L|$. Combined with Lemma 2, we can conclude $|\mathrm{Bad}(f)| \leq \frac{\delta|L|}{2\eta} + \frac{1}{\eta}$.

Lemma 3. *For any distinct $\alpha_1, \alpha_2 \in \mathbb{F}_q^\times$, there exists a polynomial $p(X) \in \mathbb{F}_q[X]$ of degree strictly less than $\rho|L|$ such that*

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

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

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

Proof. 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

$$p_1(x^2) = \frac{f(x)+f(-x)}{2} + \alpha_1 \cdot \frac{f(x)-f(-x)}{2x} \tag{4}$$

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

$$p_2(x^2) = \frac{f(x) + f(-x)}{2} + \alpha_2 \cdot \frac{f(x) - f(-x)}{2x} \tag{5}$$

for any $x \in P_{\alpha_2}$.

From (4) and (5), 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

$$f(x) = \frac{x}{\alpha_1 - \alpha_2} \cdot \big(p_1(x^2) - p_2(x^2)\big) + \frac{1}{\alpha_2 - \alpha_1} \cdot \big(\alpha_2 p_1(x^2) - \alpha_1 p_2(x^2)\big). \tag{6}$$

Note that $\deg(p_1), \deg(p_2) \le \frac{1}{2} \cdot \rho|L| - 1$. From (6), 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 $|P_{\alpha_1} \cap P_{\alpha_2}| \ge \rho|L|$, polynomial $p(X)$ is unique, since $\rho|L|$ points uniquely determines a polynomial of degree at most $\rho|L| - 1$.

Proposition 3. *Let $\alpha, \beta \in \mathbb{F}_q^\times$ 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*

$$\text{Closest}(\text{FuncFold}_\beta(f), \text{RS}[\mathbb{F}_q, L^2, \rho]) = \text{PolyFold}_\beta(p). \tag{7}$$

Proof. Since both sides of (7) are polynomials of degree at most $\rho|L^2| - 1$, it suffices to find $\rho|L^2|$ points on which both sides of (7) 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

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

Thus, we have exhibited at least $\rho|L^2|$ points where the evaluations of both sides of (7) are equal.

Corollary 1. *Let $\alpha, \beta \in \mathbb{F}_q^\times$ be distinct and $|P_\alpha \cap P_\beta| \geq \rho|L|$. Let $p(X)$ be the unique polynomial such that $f|_{P_\alpha \cap P_\beta} = p|_{P_\alpha \cap P_\beta}$ (by Lemma 3). 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

$$\text{PolyFold}_\alpha(p) = \text{Closest}(\text{FuncFold}_\alpha(f), \text{RS}[\mathbb{F}_q, L^2, \rho]). \tag{8}$$

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

$$\text{FuncFold}_\alpha(f)(x^2) = \text{PolyFold}_\alpha(p)(x^2). \tag{9}$$

Combining (8) with (9), we have

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

which implies $x^2 \in C_\alpha$, i.e., $x \in P_\alpha$.

Lemma 4. *Let $\alpha, \beta_1, \beta_2 \in \mathbb{F}_q^\times$ 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 3). 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 $|P_\alpha \cap P_{\beta_1} \cap P_{\beta_2}| \leq \rho|L| - 1$.*

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 1, 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 1, $x \in P_\alpha \cap P_{\beta_2}$. 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 1, $f(x) = p_1(x)$ 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 $\alpha, \beta_1, \beta_2 \in \mathbb{F}_q^\times$ be different points such that*

- *$|P_\alpha \cap P_{\beta_1}|$, $|P_\alpha \cap P_{\beta_2}| \geq \rho|L|$*
- *$P_\alpha \cap P_{\beta_1} \neq P_\alpha \cap P_{\beta_2}$*

Let p_1 and p_2 be 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 $\text{PolyFold}_\alpha(p_1) = \text{PolyFold}_\alpha(p_2)$.

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

Let $A \in \{A_1, \ldots, A_r\}$ be a set of folding points defined in Algorithm 1 and α be the corresponding folding point, i.e., for any $\beta \in A$, we have $|P_\beta \cap P_\alpha| \geq \rho|L|$. We define an equivalence relation $\mathcal{R}$ on $A \setminus \{\alpha\}$ as follows:

$$(\beta_1, \beta_2) \in \mathcal{R} \iff P_\alpha \cap P_{\beta_1} = P_\alpha \cap P_{\beta_2}.$$

$\mathcal{R}$ 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 4) 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

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

Lemma 5. *Let $\alpha, \beta_1, \ldots, \beta_t \in \mathbb{F}_q^\times$ be distinct such that*

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

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 3).

Observe that $P_\alpha \cap P_{\beta_i} \cap P_{\beta_j} = P_\alpha \cap P_{\beta_i} \subseteq P_{\beta_i} \cap P_{\beta_j}$. By Lemma 3, 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 1. Since $f(x) = p(x)$ and $f(-x) = p(-x)$, using Corollary 1 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.

Lemma 6. *Let $\alpha, \beta_1, \beta_2 \in \mathbb{F}_q^\times$ be three different points such that*

- $|P_\alpha \cap P_{\beta_1}|, |P_\alpha \cap P_{\beta_2}| \geq \rho|L|$, *and*
- $|P_\alpha \cap P_{\beta_1} \cap P_{\beta_2}| < \rho|L|$.

Denote by p_1 and p_2 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, for any $x \in P_\alpha \cap P_{\beta_1} \cap P_{\beta_2}$, we have ${p_1}^{(\mathrm{even})}(x) = {p_2}^{(\mathrm{even})}(x)$ and ${p_1}^{(\mathrm{odd})}(x) = {p_2}^{(\mathrm{odd})}(x)$.

Proof. According to the assumption $|P_\alpha \cap P_{\beta_1} \cap P_{\beta_2}| < \rho|L|$, we have $p_1 \neq p_2$ by Lemma 4. By Proposition 4, we have

$$\mathrm{PolyFold}_\alpha(p_1)(X) = \mathrm{PolyFold}_\alpha(p_2)(X). \tag{10}$$

Combining (10) with (1), we have

$$\begin{aligned} \mathrm{PolyFold}_\alpha(p_1)(X^2) &= p_1^{(\mathrm{even})}(X) + \alpha p_1^{(\mathrm{odd})}(X) \\ =\mathrm{PolyFold}_\alpha(p_2)(X^2) &= p_2^{(\mathrm{even})}(X) + \alpha p_2^{(\mathrm{odd})}(X). \end{aligned} \tag{11}$$

We claim $p_1^{(\mathrm{odd})}(X) \neq p_2^{(\mathrm{odd})}(X)$. Otherwise, by (11), we have $p_1^{(\mathrm{even})}(X) = p_2^{(\mathrm{even})}(X)$. Thus, $p_1(X) = p_2(X)$. A contradiction.

Suppose $x \in P_\alpha \cap P_{\beta_1} \cap P_{\beta_2}$. By Corollary 1, $f(x) = p_1(x) = p_2(x)$, that is,

$$p_1^{(\mathrm{even})}(x) + x p_1^{(\mathrm{odd})}(x) = p_2^{(\mathrm{even})}(x) + x p_2^{(\mathrm{odd})}(x).$$

Equivalently, we have

$$x\left(p_2^{(\mathrm{odd})}(x) - p_1^{(\mathrm{odd})}(x)\right) = p_1^{(\mathrm{even})}(x) - p_2^{(\mathrm{even})}(x). \tag{12}$$

Assume for contradiction that ${p_1}^{(\mathrm{odd})}(x) \neq {p_2}^{(\mathrm{odd})}(x)$. By (12), we have

$$x = \frac{p_1^{(\mathrm{even})}(x) - p_2^{(\mathrm{even})}(x)}{p_2^{(\mathrm{odd})}(x) - p_1^{(\mathrm{odd})}(x)}.$$

By identity (11), we have

$$\alpha = \frac{p_1^{(\mathrm{even})}(x) - p_2^{(\mathrm{even})}(x)}{p_2^{(\mathrm{odd})}(x) - p_1^{(\mathrm{odd})}(x)}.$$

Thus, $x = \alpha$. Since $-x \in P_\alpha \cap P_{\beta_1} \cap P_{\beta_2}$, the same argument shows that $-x = \alpha$, which is impossible. So we have ${p_1}^{(\mathrm{odd})}(x) = {p_2}^{(\mathrm{odd})}(x)$, and (11) implies ${p_1}^{(\mathrm{even})}(x) = {p_2}^{(\mathrm{even})}(x)$.

Corollary 2. *Let $\alpha, \beta_1, \beta_2 \in \mathbb{F}_q^\times$ be different such that*

- $|P_\alpha \cap P_{\beta_1}|, |P_\alpha \cap P_{\beta_2}| \geq \rho|L|$, *and*
- $P_\alpha \cap P_{\beta_1} \neq P_\alpha \cap P_{\beta_2}$.

Denote by p_1 and p_2 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. We have

$$\begin{cases} p_2^{(\text{even})}(X) = p_1^{(\text{even})}(X) + \alpha d(X), \\ p_2^{(\text{odd})}(X) = p_1^{(\text{odd})}(X) - d(X) \end{cases},$$

where $d(X)$ is an even polynomial of degree less than $\rho|L|$. Furthermore,

$$\prod_{\gamma \in P_\alpha \cap P_{\beta_1} \cap P_{\beta_2}} (X - \gamma) \mid d(X). \tag{13}$$

Proof. Let $d(X) = p_1^{(\text{odd})}(X) - p_2^{(\text{odd})}(X)$. Since $\deg(p_1^{(\text{odd})}(X)), \deg(p_2^{(\text{odd})}(X)) < \rho|L|$, we have $\deg(d(X)) < \rho|L|$. Since both $p_1^{(\text{odd})}(X)$ and $p_2^{(\text{odd})}(X)$ are even polynomials, their difference $d(X)$ is also an even polynomial.

By Proposition 4, we have

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

That is

$$p_1^{(\text{even})}(X) + \alpha p_1^{(\text{odd})}(X) = p_2^{(\text{even})}(X) + \alpha p_2^{(\text{odd})}(X),$$

which implies that

$$p_2^{(\text{even})}(X) - p_1^{(\text{even})}(X) = \alpha p_1^{(\text{odd})}(X) - \alpha p_2^{(\text{odd})} = \alpha d(X),$$

as desired.

By Lemma 6, for any $\gamma \in P_\alpha \cap P_{\beta_1} \cap P_{\beta_2}$, we have $p_1^{(\text{odd})}(\gamma) = p_2^{(\text{odd})}(\gamma)$, that is, $d(\gamma) = 0$. This proves (13).

Lemma 7. *Let $\alpha, \beta_1, \beta_2 \in \mathbb{F}_q^\times$ be different such that*

- $|P_\alpha \cap P_{\beta_1}|, |P_\alpha \cap P_{\beta_2}| \geq \rho|L|$, *and*
- $P_\alpha \cap P_{\beta_1} \neq P_\alpha \cap P_{\beta_2}$.

Then $P_{\beta_1} \cap P_{\beta_2} \subseteq P_\alpha$.

Proof. Denote by p_1 and p_2 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. For any $x \in P_{\beta_1} \cap P_{\beta_2} \subseteq P_{\beta_1}$, we have

$$\begin{aligned} \text{FuncFold}_{\beta_1}(f)(x^2) &= \frac{1}{2}(f(x) + f(-x)) + \frac{\beta_1}{2x}(f(x) - f(-x)) && \text{By Definition 4} \\ &= \text{PolyFold}_{\beta_1}(p_1)(x^2) && \text{By Definition 4} \\ &= p_1^{(\text{even})}(x) + \beta_1 p_1^{(\text{odd})}(x). && (14) \end{aligned}$$

Similarly, we have

$$\mathrm{FuncFold}_{\beta_2}(f)(x^2) = p_2^{(\mathrm{even})}(x) + \beta_2 p_2^{(\mathrm{odd})}(x). \tag{15}$$

Combining (14) and (15), we have

$$\begin{aligned} f(x) + f(-x) = &\frac{2}{\beta_1 - \beta_2}\left(\beta_1 p_2^{(\mathrm{even})}(x) - \beta_2 p_1^{(\mathrm{even})}(x)\right) \\ &+ \frac{2\beta_1\beta_2}{\beta_1 - \beta_2}\left(p_2^{(\mathrm{odd})}(x) - p_1^{(\mathrm{odd})}(x)\right) x. \end{aligned} \tag{16}$$

Notice that $-x \in P_{\beta_1} \cap P_{\beta_2}$ and $p_1^{(\mathrm{even})}, p_1^{(\mathrm{odd})}, p_2^{(\mathrm{even})}, p_2^{(\mathrm{odd})}$ are all even polynomials. Thus,

$$\begin{aligned} f(-x) + f(x) = &\frac{2}{\beta_1 - \beta_2}\left(\beta_1 p_2^{(\mathrm{even})}(x) - \beta_2 p_1^{(\mathrm{even})}(x)\right) \\ &- \frac{2\beta_1\beta_2}{\beta_1 - \beta_2}\left(p_2^{(\mathrm{odd})}(x) - p_1^{(\mathrm{odd})}(x)\right) x. \end{aligned} \tag{17}$$

Subtracting (17) from (16), we have

$$\frac{4\beta_1\beta_2}{\beta_1 - \beta_2}\left(p_2^{(\mathrm{odd})}(x) - p_1^{(\mathrm{odd})}(x)\right) x = 0,$$

which implies $p_2^{(\mathrm{odd})}(x) - p_1^{(\mathrm{odd})}(x) = 0$. By Corollary 2, we have $p_1^{(\mathrm{even})}(x) = p_2^{(\mathrm{even})}(x)$. From (16), we have $f(x) + f(-x) = 2p_1^{(\mathrm{even})}(x)$. From (14) and (15), we have $f(x) - f(-x) = 2xp_1^{(\mathrm{odd})}(x)$. Thus, $f(x) = p_1(x)$ and $f(-x) = p_1(-x)$. Corollary 1 tells us $x \in P_\alpha \cap P_{\beta_1} \subseteq P_\alpha$.

Now we are ready to bound the size of each block.

Lemma 8. *Let $f : L \to \mathbb{F}_q$ be a codeword such that $\Delta(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$, and C_α, P_α are defined as in Definition 7. Let $\alpha, \beta_1, \ldots, \beta_t \in \mathbb{F}_q^\times$ be different such that, for any $i \in \{1, \ldots, t\}$, $|P_\alpha \cap P_{\beta_i}| \geq \rho|L|$. We have*

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

Proof. Let s be the number of equivalence classes. Denote by $\{\beta_{i,1}, \ldots, \beta_{i,t_i}\}$ the ith equivalence class, where $i = 1, \ldots, s$. That is,

$$P_\alpha \cap P_{\beta_{i,1}} = \ldots = P_\alpha \cap P_{\beta_{i,i_t}},$$

and for any distinct i, j, $P_\alpha \cap P_{\beta_{i,1}} \neq P_\alpha \cap P_{\beta_{j,1}}$. Clearly, $t = \sum_{i=1}^{s} t_i$.

We have

$$\begin{aligned} |L \setminus P_\alpha| &\geq \left| \left(\bigcup_{i=1}^{s} \bigcup_{j=1}^{t_i} P_{\beta_{i,j}} \right) \cap (L \setminus P_\alpha) \right| \\ &\geq \left| \bigcup_{i=1}^{s} \bigcup_{j=1}^{t_i} \left(P_{\beta_{i,j}} \setminus P_\alpha\right) \right| \qquad \text{By the distributive law.} \end{aligned}$$

By Lemma 7, for distinct $i_1 \neq i_2$, for any j_1, j_2, we have $\left(P_{\beta_{i_1,j_1}} \setminus P_\alpha\right) \cap \left(P_{\beta_{i_2,j_2}} \setminus P_\alpha\right) = \emptyset$. By Lemma 5, for any i, for any distinct $j_1 \neq j_2$, we have $\left(P_{\beta_{i,j_1}} \setminus P_\alpha\right) \cap \left(P_{\beta_{i,j_2}} \setminus P_\alpha\right) = \emptyset$. Thus, we have

$$\bigcup_{i=1}^{s} \bigcup_{j=1}^{t_i} \left(P_{\beta_{i,j}} \setminus P_\alpha\right) = \sum_{i=1}^{s} \sum_{j=1}^{t_i} \left|P_{\beta_{i,j}} \setminus P_\alpha\right| \tag{18}$$

Since $\beta_{i,j}$ is a bad folding point, the distance between $\mathrm{FuncFold}_{\beta_{i,j}}$ and $\mathrm{RS}[\mathbb{F}_q, L^2, \rho]$ is at most $\rho|L^2| - 1$. By the definition of $C_{\beta_{i,j}}$ (Definition 7), we have $|C_{\beta_{i,j}}| \geq (1-\delta)|L^2|$, that is, $|P_{\beta_{i,j}}| \geq (1-\delta)|L|$. On the other hand, f is δ-far from a low degree polynomial, we have $|P_\alpha \cap P_{\beta_{i,j}}| < (1-\delta)|L|$. Otherwise, by Lemma 3, there exists a low degree polynomial which agrees with f on more than $(1-\delta)|L|$ points. Furthermore, according to Definition 7, $\left|P_{\beta_{i,j}} \setminus P_\alpha\right|$ is even. Thus,

$$\left|P_{\beta_{i,j}} \setminus P_\alpha\right| \geq 2. \tag{19}$$

Plugging (19) into (18), we have

$$|L \setminus P_\alpha| \geq \sum_{i=1}^{s} \sum_{j=1}^{t_i} \left|P_{\beta_{i,j}} \setminus P_\alpha\right| \geq 2 \sum_{i=1}^{s} t_i.$$

On the other hand, $|P_\alpha| \geq (1-\delta)|L|$, since α is a bad folding point. So we have

$$\sum_{i=1}^{s} t_i \leq \frac{\delta}{2}|L|.$$

4.3 Proof of Theorem 2

Algorithm 1 gives a partition of $\mathrm{Bad}(f)$, denoted by $A_1, \ldots, A_r \subseteq \mathbb{F}_q^\times$. By Lemma 2, the number of blocks is at most $r \leq 1/\eta$. By Lemma 8, the size of each block A_i is at most $\delta|L|/2 + 1$. So $\sum_{i=1}^{r} |A_i| \leq \frac{1}{\eta} \cdot (\frac{\delta|L|}{2} + 1)$.

5 Applications

In Section 5.1, we use our main theorem to prove the proximity gaps for Reed-Solomon codes. The definition of proximity gaps was introduced in [4]. In Section 5.2, we improve the soundness analysis of (batched) FRI, based on the results in Section 5.1.

5.1 Proximity gaps for Reed-Solomon codes

A property displays a proximity gap (Definition 1.1 in [4]) 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. [4] 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.

This subsection, 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 3) and correlated agreement over affine spaces (Theorem 5). These theorems lead to the soundness analysis of (batched) FRI in Section 5.2.

Correlated agreement over lines

Theorem 3. *Let L be a subset of $\mathbb{F}_q^\times$. Let $u_0, u_1 : L \to \mathbb{F}_q$. Let $\delta, \eta > 0$ satisfy $\delta \le \delta_0(\rho, \eta) := 1 - \sqrt{\rho} - \eta$, and suppose*

$$\mathbb{P}_{z \in \mathbb{F}_q^\times}(\Delta(u_0 + z u_1, \mathrm{RS}[\mathbb{F}_q, L, \rho]) \le \delta) > \frac{\delta|L| + 1}{\eta |\mathbb{F}^\times|}$$

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 3 We use Theorem 2 to prove Theorem 3. In order to apply Theorem 2, we construct a function f whose folding result at point z is exactly $u_0 + z u_1$.

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

$$f(x) = u_0(x^2) + x u_1(x^2), x \in \sqrt{L}.$$

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

$$\mathrm{FuncFold}_z(f)(x^2) = \frac{f(x) + f(-x)}{2} + z \cdot \frac{f(x) - f(-x)}{2x} = u_0(x^2) + z u_1(x^2) \in \mathbb{F}_q.$$

Thus we have

$$\begin{aligned} &\left\{z \in \mathbb{K}^\times \mid \Delta(\mathrm{FuncFold}_z(f), \mathrm{RS}[\mathbb{K}, L, \rho]) \le \delta|L|\right\} \\ \supseteq &\left\{z \in \mathbb{F}_q^\times \mid \Delta(u_0 + z u_1, \mathrm{RS}[\mathbb{F}_q, L, \rho]) \le \delta|L|\right\}. \end{aligned}$$

Since

$$\mathbb{P}_{z \in \mathbb{F}_q^\times}(\Delta(u_0 + z u_1, \mathrm{RS}[\mathbb{F}_q, L, \rho]) \le \delta) > \frac{\delta|L| + 1}{\eta |\mathbb{F}_q^\times|},$$

we have

$$\left|\{z \in \mathbb{K}^\times \mid \delta(\mathrm{FuncFold}_z(f), \mathrm{RS}[\mathbb{K}, L, \rho]) \le \delta\}\right| > \frac{\delta|L| + 1}{\eta} = \frac{\delta\left|\sqrt{L}\right|}{2\eta} + \frac{1}{\eta}. \quad (20)$$

Theorem 2 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 2 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 $\mathrm{Closest}_P(f, V)$ 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 Lemma 8 by using the pairing distance. C_α and P_α are defined as in Definition 7, with the distance is replaced by the pairing distance.

Lemma 9. *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 $\Delta_P(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$, and C_α, P_α are defined as in Definition 7. Let $\alpha, \beta_1, \ldots, \beta_t \in \mathbb{F}_q^\times$ be different points such that, for any $i \in \{1, \ldots, t\}$, $|P_\alpha \cap P_{\beta_i}| \geq \rho|L|$. We have*

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

Proof. Since the definitions of P_α and P_{β_i} are unchanged, (18) still holds. It remains to prove (19) holds in this case.

Since $\beta_{i,j}$ is a bad folding point, the distance between $\mathrm{FuncFold}_{\beta_{i,j}}(f)$ and $\mathrm{RS}[\mathbb{F}_q, L^2, \rho]$ is at most $\rho|L^2| - 1$. By Definition 7, we have $|C_{\beta_{i,j}}| \geq (1-\delta)|L^2|$, that is, $|P_{\beta_{i,j}}| \geq (1-\delta)|L|$. On the other hand, since $\Delta_P(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$, we have $|P_\alpha \cap P_{\beta_{i,j}}| < (1-\delta)|L|$. Otherwise, if $|P_\alpha \cap P_{\beta_{i,j}}| \geq (1-\delta)|L|$, by Lemma 3, let p be the low degree polynomial that agrees with f on $P_\alpha \cap P_{\beta_{i,j}}$. By Corollary 1, for any $x \in P_\alpha \cap P_{\beta_{i,j}}$, $f(x) = p(x)$. According to Definition 7, $-x \in P_\alpha \cap P_{\beta_{i,j}}$ and $f(-x) = p(-x)$. Thus, $\Delta_p(f, p) \leq |L \setminus (P_\alpha \cap P_{\beta_{i,j}})| \leq \delta|L|$. A contradiction. Furthermore, according to Definition 7, the number of elements in $P_{\beta_{i,j}} \setminus P_\alpha$ is even. Thus, (19) holds.

Using Lemma 9, following a similar argument as that of Theorem 2, we can prove the following result, which strengthens Theorem 2.

Theorem 4. *Let $L \subseteq \mathbb{F}_q^\times$ be a set with pairing elements. Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt{\rho} - \eta$. Let $f : L \to \mathbb{F}_q$ be a codeword such that $\delta_P(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta$. Then,*

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

Notice that $\sqrt{L}$ is with pairing elements and $\mathrm{char}(\mathbb{K}) \neq 2$. According to Theorem 4, (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

$$\begin{aligned}\left|\sqrt{L}\right| - \Delta_P(f,p) &= \left|\{x \in \sqrt{L} : (f(x), f(-x)) = (p(x), p(-x))\}\right| \\ &\geq (1-\delta)\left|\sqrt{L}\right|. \end{aligned} \tag{21}$$

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

$$\begin{aligned} &\left|\{x^2 \in L : \big(u_0(x^2), u_1(x^2)\big) = \big(v_0(x^2), v_1(x^2)\big)\}\right| \\ =\ &\frac{1}{2}\left|\{x \in \sqrt{L} : (f(x), f(-x)) = (p(x), p(-x))\}\right| \\ \geq\ &(1-\delta)\frac{\left|\sqrt{L}\right|}{2} = (1-\delta)|L|. \qquad \text{By (21)} \end{aligned}$$

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 3.

Correlated agreement over affine spaces

Theorem 5. *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 $\delta, \eta > 0$ satisfy $\delta \leq \delta_0(\rho, \eta) := 1 - \sqrt{\rho} - \eta$. For convenience, denote $V = \mathrm{RS}[\mathbb{F}_q, L, \rho]$. Define*

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

and suppose

$$\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^{\times l}}(\mathbf{z}_l \in S) > \left(\frac{\delta|L|+1}{\eta|\mathbb{F}^\times|}\right) \cdot l.$$

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 worse folding points to be*

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

Theorem 6. *Let $\delta, \eta, \rho > 0$ satisfy $\delta \leq 1 - \sqrt{\rho} - \eta$. 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*

$$|\mathrm{Bad_L}(f)| \leq \frac{\delta|L|}{2\eta} + \frac{1}{\eta}.$$

The proof of Theorem 6 follows the proof of Theorem 2. Due to the space limitation, we only outline the modifications here. Details of the proof 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

$$C'_\alpha \triangleq \mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \subseteq L^2, \tag{22}$$

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

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

This is because if $\exists D_i, |C'_\alpha \cap D_i^2| \geq \rho|L^2|$, 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

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

because of (22). Since $|C'_\alpha \cap D_i^2| \geq \rho|L^2|$, 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 $C'_\alpha = \mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \supseteq D_i^2$. 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 1 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 2 still holds because our partition strategy is unchanged and we have $|C'_\alpha| \geq (1-\delta)|L^2|$ for any $\alpha \in \mathrm{Bad}_\mathrm{L}(f)$. So 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 $|C'_\alpha \cap D_i^2| < \rho|L^2|, i = 1, \ldots, m$, 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 $P'_\alpha \cap P'_\beta$ by using the same skill as Lemma 3 (see Lemma 11 in Appendix A for details). Since $|P'_\alpha \cap P'_\beta| \geq \rho|L|$ according to the partition in Algorithm 1 and $|C'_\alpha \cap D_i^2| < \rho|L^2|, i = 1, \ldots, m$ according to our assumption, $P'_\alpha \cap P'_\beta \nsubseteq D_i, i = 1, \ldots, m$. 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 $|A| \leq \frac{\delta|L|}{2\eta} + \frac{1}{\eta}$ in this case, we outline the proof sketch here, details can be found in Lemma 15 in Appendix A. Notice that Lemma 5 and Lemma 7 still hold in this case (see Lemma 13 and Lemma 14 in Appendix A for details), so our goal is to prove (19) in this case. Using the same skill as Lemma 3, f agrees with a low-degree polynomial p on $P'_\alpha \cap P'_\beta$ (Lemma 11, Appendix A). Since $x \in P'_\alpha \cap P'_\beta \iff -x \in P'_\alpha \cap P'_\beta$, $\exists D_i \in \mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{F}_q,L,\rho]}$, such that $P'_\alpha \cap P'_\beta \subseteq D_i$. According to the definition of $\mathrm{Bad}_\mathrm{L}(f)$ and P'_β, $P'_\beta \neq D_i$. Furthermore, the number of elements in $P'_\beta \setminus P'_\alpha$ is even. As a result,

$$|P'_\beta \setminus P'_\alpha| = |P'_\beta \setminus (P'_\alpha \cap P'_\beta)| \geq 2.$$

Plugging the above inequality into (18), we have

$$|L \setminus P'_\alpha| \geq \sum_{i=1}^{s} \sum_{j=1}^{t_i} \left| P'_{\beta_{i,j}} \setminus P'_\alpha \right| \geq 2 \sum_{i=1}^{s} t_i.$$

On the other hand, $|P'_\alpha| \geq (1-\delta)|L|$, since α is a bad folding point. So we have

$$\sum_{i=1}^{s} t_i \leq \frac{\delta}{2}|L|.$$

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 7. *Let L be a subset of $\mathbb{F}_q^\times$. Let $\delta, \eta > 0$ satisfy $\delta \leq \delta_0(\rho, \eta) := 1 - \sqrt{\rho} - \eta$. Let $u_0, u_1 : L \to \mathbb{F}_q$. Let*

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

Then

$$|S_1| \leq \frac{\delta|L| + 1}{\eta}.$$

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

$$\mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{K},\sqrt{L},\rho]} = \{\sqrt{D_1}, \ldots, \sqrt{D_m}\}. \tag{23}$$

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, $\mathrm{FuncFold}_z(f)|_D = \mathrm{PolyFold}_z(v)|_D, \forall z \in \mathbb{F}_q^\times$. 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, (23) holds. Furthermore, $\forall z \in \mathbb{F}_q^\times, \forall x^2 \in L$, 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

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

We have proved that (23) holds, so Theorem 6 tells us

$$|S| \leq \frac{\delta}{2\eta}|\sqrt{L}| + \frac{1}{\eta} = \frac{\delta|L|+1}{\eta}.$$

Corollary 3. *Let L be a subset of $\mathbb{F}_q^\times$. Let $\delta, \eta > 0$ satisfy $\delta \leq \delta_0(\rho, \eta) := 1 - \sqrt{\rho} - \eta$. $l \geq 1$, 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*

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

Then

$$\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^{\times l}}(\mathbf{z}_l \in S_l) \leq \frac{\delta|L|+1}{\eta|\mathbb{F}_q^\times|} \cdot l.$$

Proof. We use induction to prove the result. When $l = 1$, the problem is reduced to Theorem 7. Suppose the result holds for $l-1$ functions. Denote $\{u_0, \cdots, u_{l-1}\}$ as $\mathbf{u}_{l-1}$. Define

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

Then we have

$$\begin{aligned} &\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^{\times l}}(\mathbf{z}_l \in S_l) \\ =&\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^{\times l}}(\mathbf{z}_l \in S_l \mid \mathbf{z}_{l-1} \in S_{l-1}) \cdot \mathbb{P}_{\mathbf{z}_{l-1} \in \mathbb{F}_q^{\times l-1}}(\mathbf{z}_{l-1} \in S_{l-1}) + \\ &\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^{\times l}}(\mathbf{z}_l \in S_l \mid \mathbf{z}_{l-1} \notin S_{l-1}) \cdot \mathbb{P}_{\mathbf{z}_{l-1} \in \mathbb{F}_q^{\times l-1}}(\mathbf{z}_{l-1} \notin S_{l-1}) \\ \leq&\mathbb{P}_{\mathbf{z}_{l-1} \in \mathbb{F}_q^{\times l-1}}(\mathbf{z}_{l-1} \in S_{l-1}) + \mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^{\times l}}(\mathbf{z}_l \in S_l \mid \mathbf{z}_{l-1} \notin S_{l-1}). \end{aligned} \tag{24}$$

According to our assumption, the first item satisfies

$$\mathbb{P}_{\mathbf{z}_{l-1} \in \mathbb{F}_q^{\times l-1}}(\mathbf{z}_{l-1} \in S_{l-1}) \leq \frac{\delta|L| + 1}{\eta|\mathbb{F}_q^{\times}|} \cdot (l-1). \tag{25}$$

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

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

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

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

We want to prove

$$S_{\mathbf{z}_{l-1}} \subseteq S_{\delta,\{u',u_l\},V}. \tag{26}$$

Notice that $|S_{\delta,\{u',u_l\},V}| \leq \frac{\delta|L|+1}{\eta}$ by Theorem 7. So if (26) holds, we have

$$|S_{\mathbf{z}_{l-1}}| \leq \frac{\delta|L| + 1}{\eta}. \tag{27}$$

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

$$\mathcal{A}_{\delta,\{u',u_l\},V} \subseteq \mathcal{A}_{\delta,\mathbf{u}_l,V}. \tag{28}$$

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

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

(28) 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 (26).

Plugging (25) and (27) into (24), we have

$$\mathbb{P}_{\mathbf{z}_l \in \mathbb{F}_q^{\times l}}(\mathbf{z}_l \in S_l) \leq \frac{\delta|L|+1}{\eta|\mathbb{F}_q^\times|} \cdot l.$$

Proof of Theorem 5

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

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

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

Conjectured proximity gaps Besides the provable proximity gaps of the RS codes, [4] also provides a conjecture. Many implementations of FRI are based on this conjecture.

Conjecture 1 (Conjecture 8.4 in [4]). 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

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

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 3 provides a proof for a part of the conjecture when setting $c_1 = c_2 = 1$:

- The characteristic of the finite field should be greater than 2. An example in [8] shows that this conjecture does not hold for finite fields of characteristic 2 when $c_1 = c_2 = 1$. Furthermore, Appendix B provides additional insights into why the conjecture cannot be true in our proof.
- The parameter δ still needs to be under the Johnson bound because of the use of Corrádi Lemma (Lemma 1). The proof of the conjecture beyond the Johnson bound remains open.

5.2 Soundness of batched FRI

The batched FRI protocol FRI[2] 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 $\mathrm{RS}[\mathbb{F}_q, L^{(0)}, \rho]$ or is far from the code. In particular, FRI works for *smooth* evaluation sets defined in Definition 9.

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 [2], 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 kth round, the prover sends the oracle of a function $f^{(k)}$ to the verifier. The verifier randomly selects a folding parameter $\alpha^{(k)} \in \mathbb{F}_q^\times$ 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. For those

who may be unfamiliar with it, Appendix C offers a detailed description of the FRI protocol.

Soundness of batched FRI The soundness error of batched FRI consists of bad batching, bad folding, and prover's cheating. Theorem 5 and Theorem 2 restrict the happening of the first and the second items respectively. The prover's cheating will be checked during the query phase. Appendix D proposes an analysis of the possibility of bad folding.

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 [2], which offers a comprehensive soundness analysis.

Theorem 8 (Batched FRI soundness). *Let $\mathbb{F}_q$ be a field with $\mathrm{char}(\mathbb{F}_q) \neq 2$. Let $L^{(0)} \subseteq \mathbb{F}_q^\times$ 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 $\delta \leq \delta_0(\rho, \eta) := 1 - \sqrt{\rho} - \eta$. 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

$$\epsilon_{Batched\text{-}FRI} = \left(\frac{\left(1 - \sqrt{\rho} - \eta\right) |L^{(0)}|}{\eta |\mathbb{F}_q^\times|} \right) \cdot (l+1) + (1-\delta)^t$$

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 1. In the theorem, we assume that the folding point α chosen by the verifier is non-zero. Notice that allowing α to be zero would add at most one bad point.

Remark 2. For general cases that $q(X) = 2^k, k \in \mathbb{N}^+$, 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.

Numerical Example We provide a numerical example to show the progress of the provable soundness of ethSTARK[17] using the suggested parameters in [17] (Section 5.10.2). Let λ be the number of bits of security, the protocol has the following formula:

$$\lambda \geq \min \left\{ \begin{array}{l} -\log_2 \left(\begin{array}{l} \frac{s}{|\mathbb{K}|} + \frac{(d_{\max} + 2^h + a) \cdot s^2}{|\mathbb{K}| - a \cdot |D| + |H_0|} \\ + \frac{(m + \frac{1}{2})^7 \cdot |L^{(0)}|^2}{2\rho^{3/2} q} + \frac{(2m+1) \cdot (|L^{(0)}| + 1)}{\sqrt{\rho}} \cdot \frac{\sum_{i=0}^{n_r - 1} l^{(i)}}{q} \end{array} \right), \\ \zeta - t \cdot \log_2 \left(\sqrt{\rho} \cdot (1 + \frac{1}{2m}) \right) \end{array} \right\} - 1,$$

where $\mathbb{K}$ is an extension of $\mathbb{F}_q$ with extension degree e and FRI protocol applies on $\mathbb{K}$. In their example, they set $|\mathbb{K}| > 2^{183}$; $\rho = 1/4$; $m = 3$, so that $s = 12 \leq 2^4$. $|H_0| = 2^h$ for $h = 20$ and $|D| = |H_0|/\rho = 2^{22}$. Furthermore, $d_{\max} = 2^{22}$ and $a = 4$ and the number of batching functions $l = a + |I| = 4 + 24 < 2^5$. Then the sum in logarithm on the top row is bounded by

$$\frac{2^4}{2^{183}} + \frac{2^{23} \cdot 2^8}{2^{182}} + \frac{4^7 \cdot (2^{22})^2}{2/8 \cdot 2^{183}} + 2^{26} \cdot \frac{2^{20}}{2^{183}} \leq 2^{-122},$$

where the dominating term is the third summand above.

Set $\eta = \frac{\sqrt{\rho}}{2m} = 1/12$, we have improved the corresponding formula to be:

$$\lambda \geq \min \left\{ \begin{array}{l} -\log_2 \left(\frac{s}{|\mathbb{K}|} + \frac{(d_{\max}+2^h+a)\cdot s^2}{|\mathbb{K}|-a\cdot|D|+|H_0|} + \frac{n}{\eta(|\mathbb{K}|-1)} \cdot (l+1) \right), \\ \zeta - t \cdot \log_2 \left(\sqrt{\rho} \cdot (1 + \frac{1}{2m}) \right) \end{array} \right\} - 1,$$

and the numerical result of the sum in logarithm on the top row is bounded by

$$\frac{2^4}{2^{183}} + \frac{2^{23} \cdot 2^8}{2^{182}} + \frac{2^{22}}{1/12 \cdot 2^{183}} \cdot 2^5 < 2^{-150},$$

the dominating term is the second term and the third term is close to it.

Fixing the number of bits of grinding to $\zeta = 20$ as in [17], for provable soundness level of $\lambda = 80$ bits, we may use $t = 79$ invocations; for $\lambda = 100$ bits, $t = 104$. So far, the dominating term in λ is the term in the second row, so our results remain the same. However, when $\lambda = 128$ bits, we can set $t = 140$ while the former calculation result can not reach this security level under this parameter setting.

References

1. Arnon, G., Chiesa, A., Fenzi, G., Yogev, E.: Stir: Reed–solomon proximity testing with fewer queries. Cryptology ePrint Archive (2024)
2. Ben-Sasson, E., Bentov, I., Horesh, Y., Riabzev, M.: 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)
3. Ben-Sasson, E., Bentov, I., Horesh, Y., Riabzev, M.: Scalable, transparent, and post-quantum secure computational integrity. Cryptology ePrint Archive (2018)
4. Ben-Sasson, E., Carmon, D., Ishai, Y., Kopparty, S., Saraf, S.: Proximity gaps for reed–solomon codes. In: 2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS). pp. 900–909. IEEE (2020)
5. Ben-Sasson, E., Chiesa, A., Riabzev, M., Spooner, N., Virza, M., Ward, N.P.: 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. pp. 103–128. Springer (2019)
6. Ben-Sasson, E., Chiesa, A., Spooner, N.: Interactive oracle proofs. In: Theory of Cryptography: 14th International Conference, TCC 2016-B, Beijing, China, October 31-November 3, 2016, Proceedings, Part II 14. pp. 31–60. Springer (2016)

7. Ben-Sasson, E., Goldberg, L., Kopparty, S., Saraf, S.: Deep-fri: sampling outside the box improves soundness. arXiv preprint arXiv:1903.12243 (2019)
8. Ben-Sasson, E., Kopparty, S., Saraf, S.: 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)
9. Block, A.R., Garreta, A., Katz, J., Thaler, J., Tiwari, P.R., Zając, M.: Fiat-shamir security of fri and related snarks. In: International Conference on the Theory and Application of Cryptology and Information Security. pp. 3–40. Springer (2023)
10. Block, A.R., Tiwari, P.R.: On the concrete security of non-interactive fri. Cryptology ePrint Archive (2024)
11. Guruswami, V., Sudan, M.: Improved decoding of reed-solomon and algebraic-geometric codes. In: Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No. 98CB36280). pp. 28–37. IEEE (1998)
12. Jukna, S.: Extremal combinatorics: with applications in computer science, vol. 571. Springer (2011)
13. Kattis, A.A., Panarin, K., Vlasov, A.: Redshift: transparent snarks from list polynomial commitments. In: Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security. pp. 1725–1737 (2022)
14. Polygon: Plonky2: Fast recursive arguments with plonk and fri, https://github.com/mir-protocol/plonky2/tree/main/plonky2, https://github.com/mir-protocol/plonky2/tree/main/plonky2
15. Reed, I.S., Solomon, G.: Polynomial codes over certain finite fields. Journal of the society for industrial and applied mathematics **8**(2), 300–304 (1960)
16. Reingold, O., Rothblum, G.N., Rothblum, R.D.: Constant-round interactive proofs for delegating computation. In: Proceedings of the forty-eighth annual ACM symposium on Theory of Computing. pp. 49–62 (2016)
17. StarkWare: ethSTARK documentation v1.2. Cryptology ePrint Archive, Paper 2021/582 (2023), https://eprint.iacr.org/2021/582, https://eprint.iacr.org/2021/582
18. Xie, T., Zhang, J., Cheng, Z., Zhang, F., Zhang, Y., Jia, Y., Boneh, D., Song, D.: zkbridge: Trustless cross-chain bridges made practical. In: Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security. pp. 3003–3017 (2022)
19. Zeilberger, H., Chen, B., Fisch, B.: Basefold: Efficient field-agnostic polynomial commitment schemes from foldable codes. Cryptology ePrint Archive (2023)
20. Zhang, J., Xie, T., Zhang, Y., Song, D.: Transparent polynomial delegation and its applications to zero knowledge proof. In: 2020 IEEE Symposium on Security and Privacy (SP). pp. 859–876. IEEE (2020)

A Proof of Theorem 6

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*

- $\mathrm{Closest}_L(\delta, \mathrm{FuncFold}_\alpha(f), \mathrm{RS}[\mathbb{F}_q, L^2, \rho]) = v$.
- $C'_\alpha \triangleq \mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \subseteq L^2$.

– $P'_\alpha \triangleq \{x \mid x^2 \in C'_\alpha\} \subseteq L$ *is the parent set of* C'_α.

Lemma 10. *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'_α*, we prove that there are two possible cases:*

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

Proof. If $\exists D_i, |C'_\alpha \cap D_i^2| \geq \rho|L^2|$, 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

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

because of (22). Since $|C'_\alpha \cap D_i^2| \geq \rho|L^2|$, 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 $C'_\alpha = \mathrm{Agree}(\mathrm{FuncFold}_\alpha(f), v) \supseteq D_i^2$. 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 11. *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*

$$f|_{P'_{\alpha_1} \cap P'_{\alpha_2}} = p|_{P'_{\alpha_1} \cap P'_{\alpha_2}},$$

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

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

Proof. By Definition 13, there exists a polynomial

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

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

$$v_1(x^2) = \frac{f(x) + f(-x)}{2} + \alpha_1 \cdot \frac{f(x) - f(-x)}{2x} \tag{29}$$

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

$$v_2(x^2) = \frac{f(x) + f(-x)}{2} + \alpha_2 \cdot \frac{f(x) - f(-x)}{2x} \tag{30}$$

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

From (29) and (30), 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 $x \in P'_{\alpha_1} \cap P'_{\alpha_2}$. 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

$$f(x) = \frac{x}{\alpha_1 - \alpha_2} \cdot \big(v_1(x^2) - v_2(x^2)\big) + \frac{1}{\alpha_2 - \alpha_1} \cdot \big(\alpha_2 v_1(x^2) - \alpha_1 v_2(x^2)\big). \quad (31)$$

for any $x \in P'_{\alpha_1} \cap P'_{\alpha_2}$.

Note that $\deg(v_1), \deg(v_2) \le \frac{1}{2} \cdot \rho|L| - 1$. From (31), 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 $|P'_{\alpha_1} \cap P'_{\alpha_2}| \ge \rho|L|$, polynomial $p(X)$ is unique, since $\rho|L|$ points uniquely determines a polynomial of degree at most $\rho|L| - 1$.

Lemma 12. *Let $\alpha, \beta_1, \beta_2 \in \mathrm{Bad}_{\mathrm{L}}(f)$ be different such that $|P'_\alpha \cap P'_{\beta_1}| \ge \rho|L|$ and $|P'_\alpha \cap P'_{\beta_2}| \ge \rho|L|$. Denote by $p_1(X)$ and $p_2(X)$ the polynomials of degree at most $\rho|L| - 1$ decided by $f|_{P'_\alpha \cap P'_{\beta_1}}$ and $f|_{P'_\alpha \cap P'_{\beta_2}}$ (Lemma 11). Then exactly one of the followings holds:*

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

Proof. The proof is the same as Lemma 4.

Run Algorithm 1 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 2 still holds because our partition strategy is unchanged and we have $|C'_\alpha| \ge (1 - \delta)|L^2|$ for any $\alpha \in \mathrm{Bad}_{\mathrm{L}}(f)$. 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 $|P_\beta \cap P_\alpha| \ge \rho|L|$. Since Lemma 12 holds, The relation $\mathcal{R}$ on $A \setminus \{\alpha\}$:

$$(\beta_1, \beta_2) \in \mathcal{R} \iff P_\alpha \cap P_{\beta_1} = P_\alpha \cap P_{\beta_2}.$$

$\mathcal{R}$ is still an equivalence relation and gives a partition on set $A \setminus \{\alpha\}$. Let s be the number of equivalence classes.

We modify Lemma 5 and Lemma 7 into the following two lemmas.

Lemma 13. *Let $\alpha, \beta_1, \ldots, \beta_t \in \mathrm{Bad}_{\mathrm{L}}(f)$ be distinct such that*

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

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

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

Lemma 14. *Let $\alpha, \beta_1, \beta_2 \in \mathrm{Bad}_{\mathrm{L}}(f)$ be different such that*

- $|P'_\alpha \cap P'_{\beta_1}|, |P'_\alpha \cap P'_{\beta_2}| \geq \rho|L|$, *and*
- $P'_\alpha \cap P'_{\beta_1} \neq P'_\alpha \cap P'_{\beta_2}$.

Then $P'_{\beta_1} \cap P'_{\beta_2} \subseteq P'_\alpha$.

The proofs of the above two lemmas are the same as Lemma 5 and Lemma 7 respectively. We omit the details here. Based on these two lemmas, we can bound the number of elements in A. The following lemma has the same result as Lemma 8, but the proof is slightly different.

Lemma 15. *Let $f : L \to \mathbb{F}_q$ be a codeword such that $\Delta(f, \mathrm{RS}[\mathbb{F}_q, L, \rho]) > \delta|L|$, and C'_α, P'_α are defined as in Definition 13. Let $\alpha, \beta_1, \ldots, \beta_t \in \mathrm{Bad}_{\mathrm{L}}(f)$ be different such that, for any $i \in \{1, \ldots, t\}$, $|P'_\alpha \cap P'_{\beta_i}| \geq \rho|L|$. We have*

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

Proof. Let s be the number of equivalence classes. Denote by $\{\beta_{i,1}, \ldots, \beta_{i,t_i}\}$ the ith equivalence class, where $i = 1, \ldots, s$. That is,

$$P'_\alpha \cap P'_{\beta_{i,1}} = \ldots = P'_\alpha \cap P'_{\beta_{i,i_t}},$$

and for any distinct i, j, $P'_\alpha \cap P'_{\beta_{i,1}} \neq P'_\alpha \cap P'_{\beta_{j,1}}$. Clearly, $t = \sum_{i=1}^{s} t_i$.

We have

$$\begin{aligned} |L \setminus P'_\alpha| &\geq \left| \left(\bigcup_{i=1}^{s} \bigcup_{j=1}^{t_i} P'_{\beta_{i,j}} \right) \cap (L \setminus P'_\alpha) \right| \\ &\geq \left| \bigcup_{i=1}^{s} \bigcup_{j=1}^{t_i} \left(P'_{\beta_{i,j}} \setminus P'_\alpha \right) \right| \qquad \text{By the distributive law.} \end{aligned}$$

By Lemma 14, for distinct $i_1 \neq i_2$, for any j_1, j_2, we have $\left(P'_{\beta_{i_1,j_1}} \setminus P'_\alpha\right) \cap \left(P'_{\beta_{i_2,j_2}} \setminus P'_\alpha\right) = \emptyset$. By Lemma 13, for any i, for any distinct $j_1 \neq j_2$, we have $\left(P'_{\beta_{i,j_1}} \setminus P'_\alpha\right) \cap \left(P'_{\beta_{i,j_2}} \setminus P'_\alpha\right) = \emptyset$. Thus, we have

$$\bigcup_{i=1}^{s} \bigcup_{j=1}^{t_i} \left(P'_{\beta_{i,j}} \setminus P'_\alpha \right) = \sum_{i=1}^{s} \sum_{j=1}^{t_i} \left| P'_{\beta_{i,j}} \setminus P'_\alpha \right| \tag{32}$$

According to Lemma 11, f agrees with a low-degree polynomial p on $P'_\alpha \cap P'_\beta$. Since $x \in P'_\alpha \cap P'_\beta \iff -x \in P'_\alpha \cap P'_\beta$, $\exists D_i \in \mathcal{D}_{\delta,f,\mathrm{RS}[\mathbb{F}_q,L,\rho]}$, such that $P'_\alpha \cap P'_\beta \subseteq D_i$. According to the definition of $\mathrm{Bad}_{\mathrm{L}}(f)$ and P'_β, $P'_\beta \neq D_i$. Furthermore, the number of elements in $P'_\beta \setminus P'_\alpha$ is even. As a result,

$$|P'_\beta \setminus P'_\alpha| = |P'_\beta \setminus (P'_\alpha \cap P'_\beta)| \geq 2. \tag{33}$$

Plugging (33) into (32), we have

$$|L \setminus P'_\alpha| \geq \sum_{i=1}^{s} \sum_{j=1}^{t_i} \left| P'_{\beta_{i,j}} \setminus P'_\alpha \right| \geq 2 \sum_{i=1}^{s} t_i.$$

On the other hand, $|P'_\alpha| \geq (1-\delta)|L|$, since α is a bad folding point. So we have

$$\sum_{i=1}^{s} t_i \leq \frac{\delta}{2}|L|.$$

B Finite field with characteristic 2

Consider the characteristic p of the finite field $\mathbb{F}_q = \mathbb{F}_{p^n}$. When $p = 2$, the FRI protocol is designed on an additive domain and the folding function FuncFold is different. This is because $q(X) : X \to X^2$ is not a 2 to 1 function in this case.

Lemma 3, Lemma 4, and Lemma 5 still hold when $\mathrm{char}(\mathbb{F}_q) = 2$. However, in this case, we do not have Lemma 6 and Lemma 7. This is because these two lemmas are based on the structure of folding. When using additive cosets as the evaluation domain, the folding function FuncFold has a more complicated structure. We briefly discuss this case.

Using the same parameters as the example in [8]. Let $\mathbb{F}_{2^n}$ be the finite field. Suppose r is the number of rounds of the FRI protocol. For each $i \in [0, r]$, choose subspaces $L_0^{(k)}$ and $L^{(k)}$ satisfying the following properties:

- $L_0^{(k)} \subseteq L^{(k)}$ with $\dim(L_0^{(k)}) = 1$,
- $L^{(i+1)} = q^{(k)}(L^{(k)})$, where $q^{(k)}(X)$ is the subspace polynomial of $L_0^{(k)}$, i.e.,

$$q^{(k)}(X) = \prod_{\alpha \in L_0^{(k)}} (X - \alpha).$$

Suppose $f^{(k)}$ agrees with a low degree polynomial $p^{(k)}$ on the subdomain $\mathcal{D}$. For convenience, we suppose $\mathcal{D}$ is a collection of cosets of $L_0^{(k)}$. $q^{(k)}(X)$ is the vanishing polynomial corresponding to $L_0^{(k)}$. Then we can write $p^{(k)}(X)$ in base $q^{(k)}(X)$ as:

$$p^{(k)}(X) = a_0(X) + a_1(X)q^{(k)}(X) + \cdots + a_t(X)(q^{(k)}(X))^t,$$

where $a_i(X)$ is a polynomial with degree at most 1 and $t < \rho|L^{(k)}|/2$. Let α be the folding points, the corresponding folding function $\text{PolyFold}_\alpha(p^{(k)})$ agrees with

$$a_0(\alpha) + a_1(\alpha)X + \cdots + a_t(\alpha)X^t$$

on the subdomain $q^{(k)}(\mathcal{D})$.

In the proof of Lemma 7, let $\alpha, \beta_1, \beta_2 \in \mathbb{F}_q^\times$ be different such that

- $|P_\alpha \cap P_{\beta_1}|, |P_\alpha \cap P_{\beta_2}| \geq \rho|L|$, and
- $P_\alpha \cap P_{\beta_1} \neq P_\alpha \cap P_{\beta_2}$.

Suppose $p^{(k)}, p'^{(k)}$ are the low-degree polynomials related to $P_\alpha \cap P_{\beta_1}$ and $P_\alpha \cap P_{\beta_2}$ respectively, i.e.,

$$p^{(k)}(X) = a_0(X) + a_1(X)q^{(k)}(X) + \cdots + a_t(X)(q^{(k)}(X))^t;$$
$$p'^{(k)}(X) = a'_0(X) + a'_1(X)q^{(k)}(X) + \cdots + a'_t(X)(q^{(k)}(X))^t.$$

For any $y \in P_{\beta_1} \cap P_{\beta_2}$, we have $a_i(\alpha) = a'_i(\alpha), 0 \leq i \leq t$ and

$$\text{FuncFold}_{\beta_1}(f^{(k)})(q^{(k)}(y)) = \text{PolyFold}_{\beta_1}(p^{(k)})(q^{(k)}(y));$$
$$\text{FuncFold}_{\beta_2}(f^{(k)})(q^{(k)}(y)) = \text{PolyFold}_{\beta_2}(p^{(k)})(q^{(k)}(y)).$$

However, there exist examples that satisfy the above conditions but $f^{(k)}(y) \neq p^{(k)}(y)$. Then by Corollary 1 in additive version, $y \notin P_\alpha$. We provide a concrete example

Example 1. Consider the finite field $\mathbb{F}_{2^3}$. Let g be a root of $X^3 + X + 1$, then the finite field

$$\mathbb{F}_{2^3} \cong \{0, 1, g, g^2, g+1, g^2+g, g^2+g+1, g^2+1\}.$$

Set $L^{(0)} = \{0, 1, g, g^2, g+1, g^2+g, g^2+g+1, g^2+1\}$, when $|L_0^{(0)}| = 2$, we have $L_0^{(0)} = \{0, 1\}$. At this time, $q^{(0)}(X) = X(X-1)$.

$L^{(0)}$ can be seen as a collection of cosets:

$$L^{(0)} = \{0, 1\} \cup \{g, g+1\} \cup \{g^2, g^2+1\} \cup \{g^2+g, g^2+g+1\}.$$

We can calculate

$$L^{(1)} = q^{(0)}(L^{(0)}) = \{0, g^2+g, g, g^2\} \text{ and } L_0^{(1)} = \{0, g\}.$$

$\forall y \in L^{(1)}$, define $S_y \triangleq \{y_0, y_1 \mid q^{(0)}(y_0) = q^{(0)}(y_1) = y\}$. In our example, we have $|y_1 - y_0| = 1$, without loss of generality, we set $y_1 = y_0 + 1$.

$f^{(1)}(y)$ can be calculated through S_y with folding parameter $x^{(0)}$:

$$f^{(1)}(y) = f^{(0)}(y_1) + \frac{x^{(0)} - y_1}{y_0 - y_1} f^{(0)}(y_0) = f^{(0)}(y_1) + (x^{(0)} - y_1) f^{(0)}(y_0).$$

In Lemma 7, we want to prove that for two different given low-degree polynomials

$$p^{(0)}(X) = a_0(X) + a_1(X)q^{(0)}(X) + \cdots + a_t(X)(q^{(0)}(X))^t;$$
$$p'^{(0)}(X) = a'_0(X) + a'_1(X)q^{(0)}(X) + \cdots + a'_t(X)(q^{(0)}(X))^t,$$

if their folding results with the folding parameter α are the same, i.e.,

$$\text{PolyFold}_\alpha(p^{(0)}(X)) = \text{PolyFold}_\alpha(p'^{(0)}(X))$$

and for two different folding parameters β_1, β_2 :

$$\text{FuncFold}_{\beta_1}(f^{(0)})(q^{(0)}(y_0)) = \text{PolyFold}_{\beta_1}(p^{(0)}(y_0)),$$
$$\text{FuncFold}_{\beta_2}(f^{(0)})(q^{(0)}(y_0)) = \text{PolyFold}_{\beta_2}(p'^{(0)}(y_0))$$

then we have $f^{(0)}(y_0) = p^{(0)}(y_0)$ and $f^{(0)}(y_1) = p^{(0)}(y_1)$. However, this is not necessarily true in this case. We have

$$\begin{aligned}
&f^{(0)}(y_1) + (\beta_1 - y_1)f^{(0)}(y_0) \\
=&a_0(\beta_1) + a_1(\beta_1)\left(y_0(y_0-1)\right) + \cdots + a_t(\beta_1)\left(y_0(y_0-1)\right)^t; \\
&f^{(0)}(y_1) + (\beta_2 - y_1)f^{(0)}(y_0) \\
=&a'_0(\beta_2) + a'_1(\beta_2)\left(y_0(y_0-1)\right) + \cdots + a'_t(\beta_2)\left(y_0(y_0-1)\right)^t.
\end{aligned}$$

Then we can calculate

$$\begin{aligned}
(\beta_1 - \beta_2)f^{(0)}(y_0) =& \left(a_0(\beta_1) - a'_0(\beta_2)\right) + \left(a_1(\beta_1) - a'_1(\beta_2)\right)\left(y_0(y_0-1)\right) + \cdots \\
&+ \left(a_t(\beta_1) - a'_t(\beta_2)\right)\left(y_0(y_0-1)\right)^t.
\end{aligned}$$

We can set

$$\begin{cases} a_0(X) = x^{(0)}; \\ a'_0(X) = X; \\ a_i(X) = a'_i(X) = 0, 1 \le i \le t. \end{cases}$$

This satisfies our condition that $a_i(x^{(0)}) = a'_i(x^{(0)}), 0 \le i \le t$. But $f^{(0)}(y_0) = \frac{x^{(0)} - \beta_2}{\beta_1 - \beta_2}$ while $p^{(0)}(y_0) = x^{(0)}$. This is a counterexample.

C The batched FRI protocol

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.

D Analysis of bad folding

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

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

where $\alpha^{(k)}$ is the random folding point chosen by the verifier in the kth round. According to Theorem 2, we have

$$\mathbb{P}(E^{(k)}) \leq \frac{\left(1 - \sqrt{\rho} - \eta - \frac{1}{|L^{(k+1)}|}\right)|L^{(k)}| + 2}{2\eta|\mathbb{F}_q^{\times}|} = \frac{(1 - \sqrt{\rho} - \eta)|L^{(k)}|}{2\eta|\mathbb{F}_q^{\times}|}.$$

The possibility that in all of the n_r rounds, the bad events do not happen satisfies:

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

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.