Dive into BCIKS20-FRI Soundness

- Jade Xie jade@secbit.io
- Yu Guo yu.guo@secbit.io

The paper [BCIKS20] improves the soundness of the FRI protocol in [BBHR18], mainly analyzing the case of batched FRI. This article will provide a detailed analysis of the content related to batched FRI soundness in the [BCIKS20] paper.

Introduction

In the context of interactive proofs, distributed storage, and cryptography, various protocols have emerged that raise questions about the proximity of a linear code $V \subset \mathbb{F}_q^n$, where $\mathbb{F}_q$ is a finite field and V has minimum relative distance δ_V. These protocols assume access to oracles for a batch of vectors $\mathbf{u} = \{u_0, \cdots, u_l\} \subset \mathbb{F}_q^n$, and their soundness requires each vector u_i to be close to V in relative Hamming distance. Furthermore, soundness deteriorates as a function of the largest distance between some vector u_i and the code V. Therefore, we aim to find protocols that minimize the number of queries to elements of $\mathbf{u}$ while maximizing the probability of identifying when some vector u_i is far from V.

> ❓ **Questions**
>
> - [ ] How to understand the following sentence? Furthermore, soundness deteriorates as a function of the largest distance between some vector ui and the code V. Does soundness decrease as the maximum distance between some vectors u_i and code V increases, meaning that the probability of the Verifier rejecting decreases?
> - [ ] How to clearly explain the decrease in soundness?

Due to the linearity of V, a natural approach ([RVW13]) is to randomly sample a vector u' uniformly from span($\mathbf{u}$) (i.e., linear combinations of elements in $\mathbf{u}$), and treat the distance $\Delta(u', V)$ between u' and V as a proxy for the maximum distance between some elements of $\mathbf{u}$ and V. To prove soundness, we want that even if only one u_i is δ-far from all elements in V, the randomly chosen u' is also far from V.

In the following, Δ represents the relative Hamming distance. When $\Delta(u, v) \leq \delta$ holds for some $v \in V$, we say "u is δ-close to V", denoted as $\Delta(u, V) \leq \delta$; otherwise, we say "u is δ-far from V", denoted as $\Delta(u, V) > \delta$.

Regarding this problem, some research results are:

1. [AHIV17] If $\delta < \delta_V/4$, almost all $u' \in \mathrm{span}(\mathbf{u})$ are δ-far from V.
2. [RZ18] Improved the above result to $\delta < \delta_V/3$.
3. [BKS18] Improved to $\delta < 1 - \sqrt[4]{1 - \delta_V}$.
4. [BGKS20] Improved to $\delta < 1 - \sqrt[3]{1 - \delta_V}$, but this bound is tight for RS codes, as it can be achieved when $n = q$.

> 🤔 **Thoughts**
>
> - [ ] Why is the focus of research on increasing the upper bound of this δ? Regarding this question, my current thoughts are: The upper bound of δ here is related to δ_V, and for RS code, $\delta_V = 1 - \rho$, which is essentially related to the code rate. So increasing the upper bound means lowering the code rate, which implies more redundancy. If with the same security or the same high probability of rejecting errors, fewer queries are needed. Or to put it another way, if for the same protocol, the number of queries is fixed, the larger δ is, the higher the probability of rejection, thus improving soundness.
> - [ ] The second point of the above analysis seems to contradict "Furthermore, soundness deteriorates as a function of the largest distance between some vector u_i and the code V." This sentence says that the larger δ is, the smaller the soundness? How should this be understood?

One question we are currently concerned with is: For which codes and what range of δ does the following statement hold?

If some $u^* \in \mathrm{span}(\mathbf{u})$ is δ-far from V, then for almost all $u' \in \mathrm{span}(\mathbf{u})$, u' is also δ-far from V.

One of the main conclusions of the [BCIKS20] paper shows that when V is an RS code over a sufficiently large field (the field size is polynomially related to the block length of the code) and δ is less than the Johnson/Guruswami-Sudan list decoding bound, the above statement holds. Next, we call this a proximity gap.

Proximity Gaps

First, let's give the definition of Proximity Gaps.

Definition 1.1 [BCIKS20, Definition 1.1] (Proximity gap). Let $P \subset \Sigma^n$ be a property and $C \subset 2^{\Sigma^n}$ be a collection of sets. Let Δ be a distance measure on Σ^n. We say that C displays a (δ, ϵ)-proximity gap with respect to P under Δ if every $S \in C$ satisfies exactly one of the following:

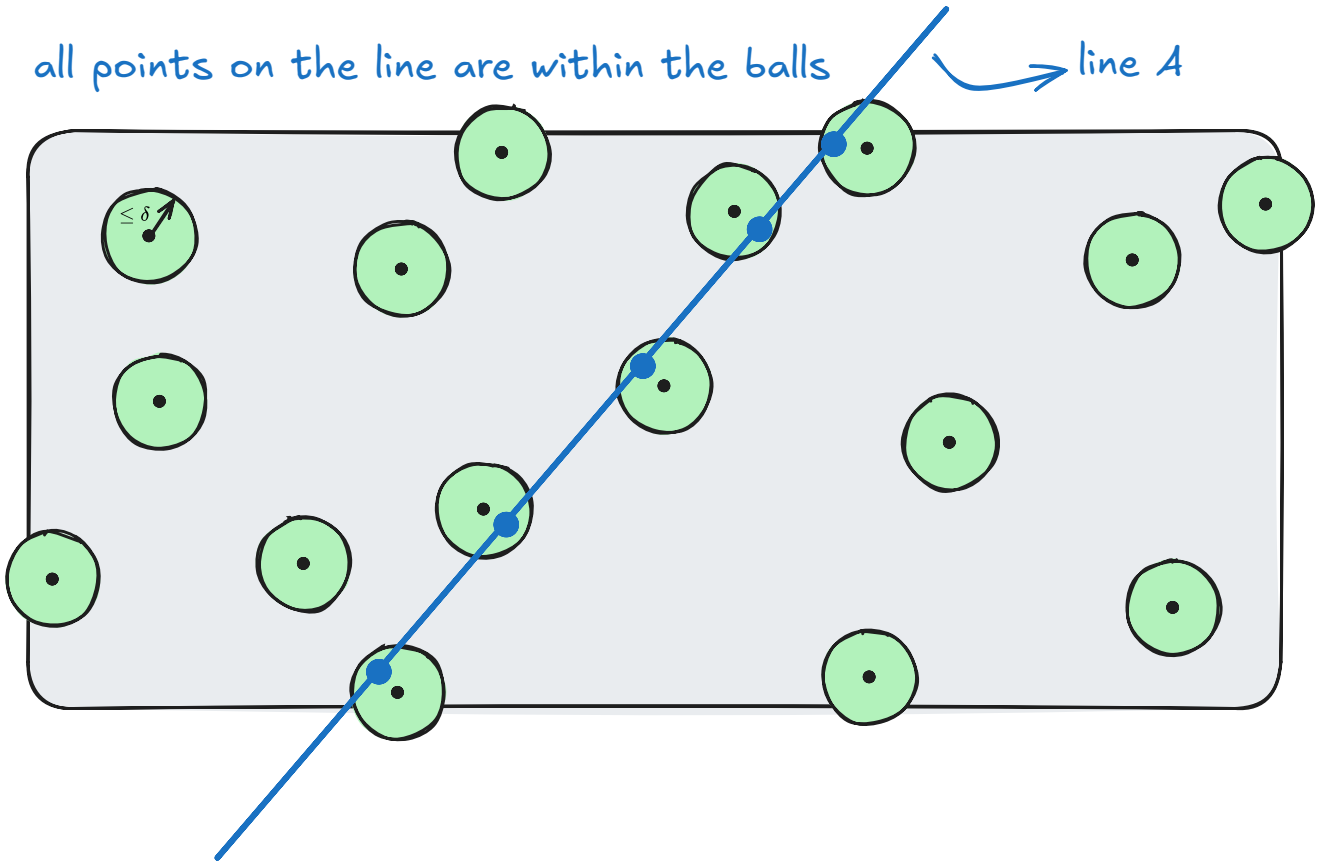
1. $\Pr_{s \in S}[\Delta(s, P) \leq \delta] = 1$.
2. $\Pr_{s \in S}[\Delta(s, P) \leq \delta] \leq \epsilon$.

We call δ the proximity parameter and ϵ is the error parameter. By default, Δ denotes the relative Hamming distance measure.

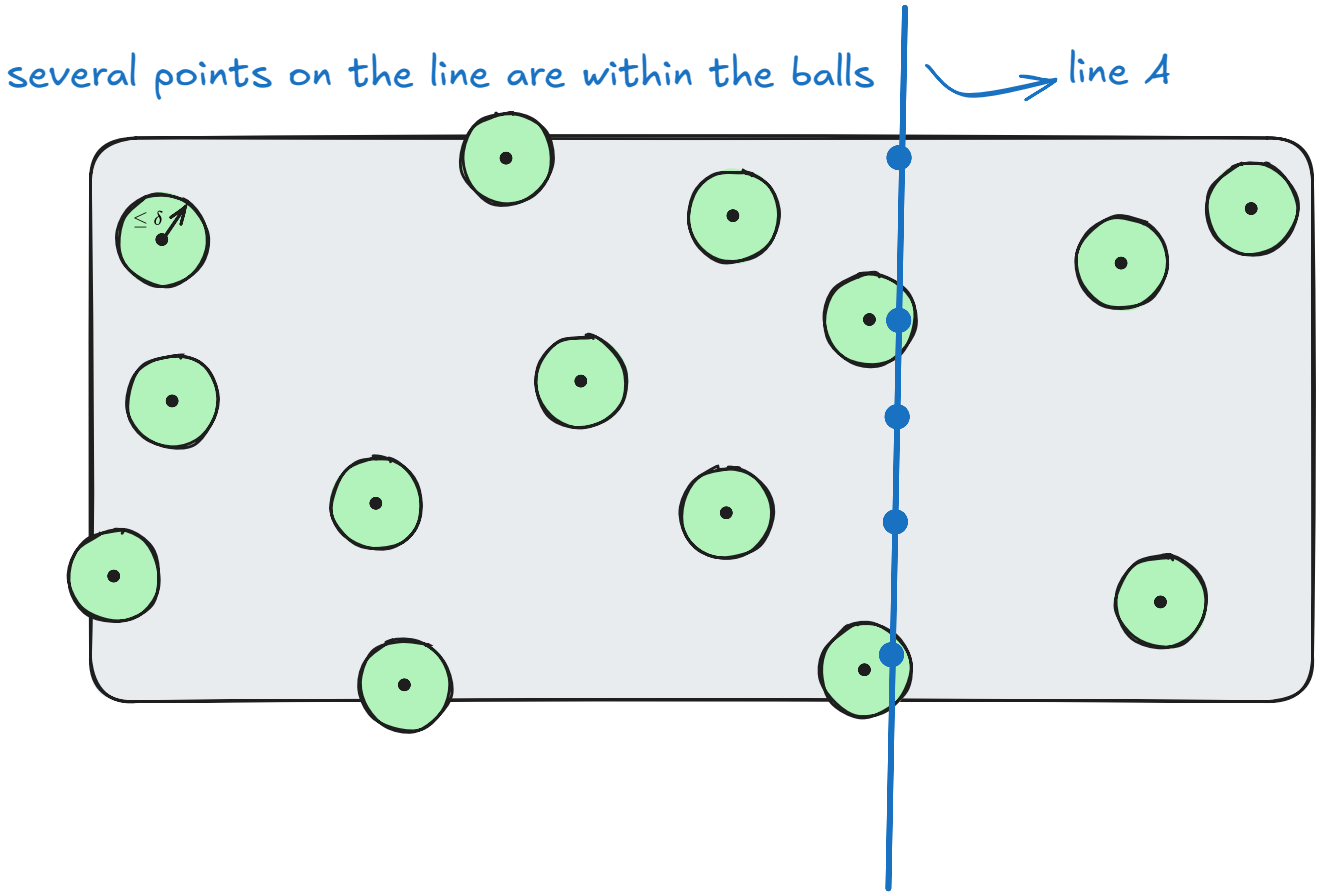
For RS code, if $V \subset \mathbb{F}^n$ is an RS encoding, corresponding to P in the above definition, and $A \subset \mathbb{F}^n$ is an affine space, corresponding to S in the above definition, then either all elements in A are δ-close to V, or almost all elements in A are δ-far from V. In other words, there is no such affine space A where about half of the elements are close to V, but at the same time, the other half are far from V.

As shown in the figure below, A is an affine space, represented here by a line, and elements in the encoding space V are represented by black dots. Circles are drawn with these points as centers and δ as the radius. Then there are only two situations:

1. All elements on line A fall within the green circular area.



2. Only a few elements on the line fall within the green circular area.



The elements in A cannot be half inside the circular area and half outside, which is also the meaning of gap. It divides all the elements in A into exactly two cases, and these two cases form a huge gap based on the relative Hamming distance.

In the following, we use $\mathbb{F}_q$ to represent a finite field of size q, and $\mathrm{RS}[\mathbb{F}_q, \mathcal{D}, k]$ to represent an RS code with dimension $k+1$ and blocklength $n = |\mathcal{D}|$, whose codewords are evaluated on $\mathcal{D}$ and are polynomials of degree $\leq k$. We use ρ to represent the code rate, so $\rho = \frac{k+1}{n}$. δ represents the relative Hamming distance relative to the RS code, and ϵ represents the error parameter, which is the probability of a "bad event" occurring.

Below is the Proximity gaps theorem for RS code.

Theorem 1.2 [BCIKS20, Theorem 1.2] (Proximity gap for RS codes). The collection C_{Affine} of affine spaces in $\mathbb{F}_q^n$ displays a (δ, ϵ)-proximity gap with respect to the RS code $V := \mathrm{RS}[\mathbb{F}_q, \mathcal{D}, k]$ of blocklength n and rate $\rho = \frac{k+1}{n}$, for any $\delta \in (0, 1-\sqrt{\rho})$, and $\epsilon = \epsilon(q, n, \rho, \delta)$ defined as the following piecewise function:

- Unique decoding bound: For $\delta \in [0, \frac{1-\rho}{2})$, the error parameter ϵ is

$$\epsilon = \epsilon_{\mathrm{U}} = \epsilon_{\mathrm{U}}(q, n) := \frac{n}{q} \tag{1.1}$$

- List decoding bound: For $\delta \in (\frac{1-\rho}{2}, 1-\sqrt{\rho})$, setting $\eta := 1 - \sqrt{\rho} - \delta$, the error parameter ϵ is

$$\epsilon = \epsilon_{\mathrm{J}} = \epsilon_{\mathrm{J}}(q, n, \rho, \delta) := \frac{(k+1)^2}{\left(2\min\left(\eta, \frac{\sqrt{\rho}}{20}\right)^7\right) q} = O\left(\frac{1}{(\eta\rho)^{O(1)}} \cdot \frac{n^2}{q}\right) \tag{1.2}$$

> 🤔 **Question**
>
> ☐ The larger δ is, the more elements may fall into the circular area, so ϵ_{J} is larger than ϵ_{U}. Is this the reason?

Correlated agreements

The main theorem proved in the paper is correlated agreement. For two vectors $u_0, u_1 \in \mathbb{F}^{\mathcal{D}}$ in $\mathbb{F}^{\mathcal{D}}$, we choose a random number z in $\mathbb{F}$, and we are concerned with the distance between the space formed by $u_0 + zu_1$ after linear combination with z and V, that is, the one-dimensional affine space $A = \{u_0 + zu_1 : z \in \mathbb{F}\}$. The correlated agreement conclusion states that if there are enough elements in A that are close enough to the RS code space V (δ-close), then there must exist a non-trivial subdomain $\mathcal{D}' \subset \mathcal{D}$, whose size is at least $1 - \delta$ times the size of $\mathcal{D}$, such that restricting u_0, u_1 to $\mathcal{D}'$, there are valid RS codes v_0, v_1 that agree with u_0, u_1 respectively on $\mathcal{D}'$. We say that such a $\mathcal{D}'$ has the correlated agreement property, meaning that u_0, u_1 and elements in A not only have a large agreement with RS code respectively, but also share a common large agreement set. This result has two parameter ranges, one is the proximity parameter within the unique decoding range, and the other is the proximity parameter within the list decoding range.

The following presents correlated agreements for three situations. Combined with other conclusions about correlated agreement in the paper, they are shown in the table below.

	Space U	$\Delta_u(u, V)$	$\Delta_u(u, V)$ **unique decoding**	$\Delta_u(u, V)$ **list decoding**	$\text{agree}_\mu(u, V)$
lines	$\{u_0 + zu_1 : z \in \mathbb{F}\}$	Theorem 1.4	Theorem 4.1	Theorem 5.1 & Theorem 5.2	
low-degree parameterized curves	$\text{curve}(\mathbf{u}) = \left\{u_z := \sum_{i=0}^{l} z^i \cdot u_i \| z \in \mathbb{F}_q\right\}$	Theorem 1.5	Theorem 6.1	Theorem 6.2	Theorem 7.1 & Theorem 7.2 (More precise version of the Johnson bound)
affine spaces	$u_0 + \text{span}\{u_1, \cdots, u_l\}$	Theorem 1.6			Theorem 7.3 & Theorem 7.4 (More precise version of the Johnson bound)

The following three theorems correspond to the correlated agreement theorems for lines, low-degree parameterized curves, and affine spaces, respectively.

Theorem 1.4 [BSCIK20, Theorem 1.4] (Main Theorem - Correlated agreement over lines). Let V, q, n, k and ρ be as defined in Theorem 1.2. For $u_0, u_1 \in \mathbb{F}_q^{\mathcal{D}}$, if $\delta \in (0, 1 - \sqrt{\rho})$ and

$$\Pr_{z \in \mathbb{F}_q}[\Delta(u_0 + z \cdot u_1, V) \leq \delta] > \epsilon, \tag{1}$$

where ϵ is as defined in Theorem 1.2, then there exist $D' \subset D$ and $v_0, v_1 \in V$ satisfying

- **Density**: $|D'|/|D| \geq 1 - \delta$, and
- **Agreement**: v_0 agrees with u_0 and v_1 agrees with u_1 on all of D'.

Let $\mathbf{u} = \{u_0, \ldots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$, then a parameterized curve of degree l is the set of vectors in $\mathbb{F}_q^{\mathcal{D}}$ generated by $\mathbf{u}$ as follows,

$$\text{curve}(\mathbf{u}) := \left\{ u_z := \sum_{i=0}^{l} z^i \cdot u_i \middle| z \in \mathbb{F}_q \right\} \tag{2}$$

Theorem 1.5 [BSCIK20, Theorem 1.5] (Correlated agreement for low-degree parameterized curves). Let V, q, n, k and ρ be as defined in Theorem 1.2. Let $\mathbf{u} = \{u_0, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$. If $\delta \in (0, 1 - \sqrt{\rho})$ and

$$\Pr_{u \in \text{curve}(u)}[\Delta(\mathbf{u}, V) \leq \delta] > l \cdot \epsilon, \tag{3}$$

where ϵ is as defined in Theorem 1.2, then there exist $\mathcal{D}' \subset \mathcal{D}$ and $v_0, \cdots, v_l \in V$ satisfying

- **Density**: $|\mathcal{D}'|/|\mathcal{D}| \geq 1 - \delta$, and
- **Agreement**: for all $i \in \{0, \cdots, l\}$, the functions u_i and v_i agree on all of $\mathcal{D}'$.

Theorem 1.6 [BSCIK20, Theorem 1.6] (Correlated agreement over affine spaces). Let V, q, n, k and ρ be as defined in Theorem 1.2. For $u_0, u_1, \cdots, u_l \in \mathbb{F}_q^{\mathcal{D}}$ let $U = u_0 + \text{span}\{u_1, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$ be an affine subspace. If $\delta \in (0, 1 - \sqrt{\rho})$ and

$$\Pr_{u \in U}[\Delta(u, V) \leq \delta] > \epsilon, \tag{4}$$

where ϵ is as defined in Theorem 1.2, then there exist $\mathcal{D}' \subset \mathcal{D}$ and $v_0, \cdots, v_l \in V$ satisfying

- **Density**: $|\mathcal{D}'|/|\mathcal{D}| \geq 1 - \delta$, and
- **Agreement**: for all $i \in \{0, \cdots, l\}$, the functions u_i and v_i agree on all of $\mathcal{D}'$.

Furthermore, in the unique decoding regime $\delta \in \left(0, \frac{1-\rho}{2}\right]$, there exists a unique maximal $\mathcal{D}'$ satisfying the above, with unique v_i.

Correlated Weighted Agreement

To analyze the soundness of the FRI protocol, we need a weighted version of Theorem 1.5.

For a given weight vector $\mu : \mathcal{D} \to [0,1]$, the (relative) μ-agreement between u and v is defined as

$$\operatorname{agree}_\mu(u,v) := \frac{1}{|\mathcal{D}|} \sum_{x : u(x)=v(x)} \mu(x). \tag{5}$$

That is, it looks at the proportion of agreement between u and v on $\mathcal{D}$ under the weight μ. If we let $\mu \equiv 1$, then

$$\operatorname{agree}_\mu(u,v) = \frac{1}{|\mathcal{D}|} \sum_{x : u(x)=v(x)} 1 = 1 - \frac{1}{|\mathcal{D}|} \sum_{x : u(x)\neq v(x)} 1 = 1 - \Delta(u,v). \tag{6}$$

The agreement between a word u and a linear code V is the maximum agreement between u and a codeword in V,

$$\operatorname{agree}_\mu(u,V) := \max_{v \in V} \operatorname{agree}_\mu(u,v). \tag{7}$$

The weighted size of a subdomain $\mathcal{D}' \subset \mathcal{D}$ is defined as

$$\mu(\mathcal{D}') := \frac{1}{|\mathcal{D}|} \sum_{x \in \mathcal{D}'} \mu(x). \tag{8}$$

If we define $\mathcal{D}'$ in the above definition as $\{x \in \mathcal{D} : u(x) = v(x)\}$, then the agreement satisfies $\operatorname{agree}_\mu(u,v) = \mu(\{x \in \mathcal{D} : u(x) = v(x)\})$.

Finally, for $\mathbf{u} = \{u_0, \cdots, u_l\}$, where $u_i \in \mathbb{F}_q^{\mathcal{D}}$ is a group of words, the μ-weighted correlated agreement is the maximum μ-weighted size of a subdomain $\mathcal{D}' \subset \mathcal{D}$, such that the restriction of $\mathbf{u}$ to $\mathcal{D}'$ belongs to $V|_{\mathcal{D}'}$, i.e., for each $i = 0, \cdots, l$, there exists $v_i \in V$ such that $u_i|_{\mathcal{D}'} = v_i|_{\mathcal{D}'}$. When μ is not specified, it is set to the constant weight function 1, which recovers the notion of correlated agreement metric discussed earlier.

Next, we assume that the weight function μ has some structure, specifically, all weights $\mu(x)$ are of the form $\mu(x) = \frac{a_x}{M}$, where a_x are varying integers with a common denominator M. For the special case of FRI soundness (where M equals the blocklength of the RS code to which the FRI protocol is applied), this assumption indeed holds. The following is a weighted generalization of Theorem 1.5.

Theorem 7.1 [BSCIK20, Theorem 7.1] (Weighted correlated agreement over curves – Version I). Let V, q, n, k and ρ be as defined in Theorem 1.2. Let $\mathbf{u} = \{u_0, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$. Let $\alpha \in (\sqrt{\rho}, 1)$ and let $\mu : \mathcal{D} \to [0,1]$ be a vector of weights, whose values all have denominator M. Suppose

$$\Pr_{u \in \operatorname{curve}(\mathbf{u})} [\operatorname{agree}_\mu(u,V) \geq \alpha] > l \cdot \epsilon, \tag{9}$$

where ϵ is as defined in Theorem 1.2 (with $\eta = \min(\alpha - \sqrt{\rho}, \frac{\sqrt{\rho}}{20})$), and additionally suppose

$$\Pr_{u \in \operatorname{curve}(\mathbf{u})} [\operatorname{agree}_\mu(u,V) \geq \alpha] \geq \frac{l(M|\mathcal{D}|+1)}{q} \left(\frac{1}{\eta} + \frac{3}{\sqrt{\rho}} \right). \tag{10}$$

Then there exists $\mathcal{D}' \subset \mathcal{D}$ and $v_0, \cdots, v_l \in V$ satisfying

- **Density**: $\mu(\mathcal{D}') \geq \alpha$, and
- **Agreement**: for all $i \in \{0, \cdots, l\}$, the functions u_i and v_i agree on all of $\mathcal{D}'$.

A more precise form that only applies to the Johnson bound range is as follows.

Theorem 7.2 [BSCIK20, Theorem 7.2] (Weighted correlated agreement over curves – Version II). Let V, q, n, k and ρ be as defined in Theorem 1.2. Let $\mathbf{u} = \{u_0, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$. Let $\mu : \mathcal{D} \to [0,1]$ be a vector of weights, whose values all have denominator M. Let $m \geq 3$ and let

$$\alpha \geq \alpha_0(\rho, m) := \sqrt{\rho} + \frac{\rho}{2m}. \tag{11}$$

Let

$$S = \{z \in \mathbb{F}_q : \operatorname{agree}_\mu(u_0 + z u_1 + \cdots + z^l u_l, V) \geq \alpha\} \tag{12}$$

and suppose

$$|S| > \max\left(\frac{(1+\frac{1}{2m})^7 m^7}{3\rho^{3/2}} n^2 l, \frac{2m+1}{\sqrt{\rho}}(M\cdot n+1)l\right). \tag{7.1}$$

Then $u_0, \ldots, u_l$ have at least α correlated μ-agreement with V, i.e. $\exists v_0, \cdots, v_l \in V$ such that

$$\mu(\{x \in \mathcal{D} : \forall 0 \le i \le l, u_i(x) = v_i(x)\}) \ge \alpha. \tag{13}$$

Theorem 7.3 [BSCIK20, Theorem 7.3] (Weighted correlated agreement over affine spaces). Let V, q, n, k and ρ be as defined in Theorem 1.2. Let $\mathbf{u} = \{u_0, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$ and let $U = u_0 + \mathrm{span}\{u_1, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$ be an affine subspace. Let $\alpha \in (\sqrt{\rho}, 1)$ and let $\mu : \mathcal{D} \to [0,1]$ be a vector of weights, whose values all have denominator M. Suppose

$$\Pr_{u\in U}[\mathrm{agree}_\mu(u, V) \ge \alpha] > \epsilon, \tag{14}$$

where ϵ is as defined in Theorem 1.2 (with $\eta = \min(\alpha - \sqrt{\rho}, \frac{\sqrt{\rho}}{20})$), and additionally suppose

$$\Pr_{u\in U}[\mathrm{agree}_\mu(u, V) \ge \alpha] \ge \frac{M|\mathcal{D}|+1}{q}\left(\frac{1}{\eta} + \frac{3}{\sqrt{\rho}}\right). \tag{15}$$

Then there exist $\mathcal{D}' \subset \mathcal{D}$ and $v_0, \cdots, v_l \in V$ satisfying

- μ-**Density**: $\mu(\mathcal{D}') \ge \alpha$, and
- **Agreement**: for all $i \in \{0, \cdots, l\}$, the functions u_i and v_i agree on all of $\mathcal{D}'$.

Similarly, there is a more precise form for Theorem 7.3 regarding the Johnson bound.

Theorem 7.4 [BSCIK20, Theorem 7.4] (Weighted correlated agreement over affine spaces – Version II). Let V, q, n, k and ρ be as defined in Theorem 1.2. Let $\mathbf{u} = \{u_0, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$ and let $U = u_0 + \mathrm{span}\{u_1, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$ be an affine subspace. Let $\mu : \mathcal{D} \to [0,1]$ be a vector of weights, whose values all have denominator M. Let $m \ge 3$ and let

$$\alpha \ge \alpha_0(\rho, m) := \sqrt{\rho} + \frac{\sqrt{\rho}}{2m}. \tag{16}$$

Suppose

$$\Pr_{u\in U}[\mathrm{agree}_\mu(u, V) \ge \alpha] > \max\left(\frac{(1+\frac{1}{2m})^7 m^7}{3\rho^{3/2}} \cdot \frac{n^2}{q}, \frac{2m+1}{\sqrt{\rho}} \cdot \frac{M\cdot n+1}{q}\right). \tag{7.2}$$

Then $u_0, \ldots, u_l$ have at least α correlated μ-agreement with V, i.e. $\exists v_0, \cdots, v_l \in V$ such that

$$\mu\left(\{x \in \mathcal{D} : \forall 0 \le i \le l, u_i(x) = v_i(x)\}\right) \ge \alpha. \tag{17}$$

FRI Protocol

The purpose of the FRI protocol is to solve the Reed-Solomon proximity testing problem in the IOP model, that is, for a received word $f^{(0)} : \mathcal{D}^{(0)} \to \mathbb{F}$, verify its proximity to $V^{(0)} := \mathrm{RS}[\mathbb{F}, \mathcal{D}^{(0)}, k^{(0)}]$. If $f^{(0)}$ belongs to $V^{(0)}$, accept; if it is δ-far from $V^{(0)}$, reject. The FRI protocol applies to any case where the evaluation domain $\mathcal{D}^{(0)}$ is a coset of a *2-smooth group*, i.e., for any $\mathcal{D}^{(0)}$, it is a coset of an (additive or multiplicative) group of size 2^s, where s is an integer. Therefore, for simplicity, we assume that the group $\mathcal{D}^{(0)}$ is multiplicative. The FRI protocol has two phases: the COMMIT phase and the QUERY phase.

In the COMMIT phase, after a finite number of r rounds of interaction, a series of functions $f^{(1)} : \mathcal{D}^{(1)} \to \mathbb{F}, f^{(2)} : \mathcal{D}^{(2)} \to \mathbb{F}, \cdots, f^{(r)} : \mathcal{D}^{(r)} \to \mathbb{F}$ will be generated. In each iteration, the size of the domain $|\mathcal{D}^{(i)}|$ decreases. Assuming that for an honest prover, $f^{(0)}$ is low-degree, then for each $f^{(i)}$, they will all be low-degree (see Proposition 1). At the beginning of the i-th round, the prover's message $f^{(i)} : \mathcal{D}^{(i)} \to \mathbb{F}$ has been generated, and the verifier can access the oracle of this message. The Verifier now sends a uniformly random $z^{(i)} \in \mathbb{F}$, and then the prover replies with a new function $f^{(i+1)} : \mathcal{D}^{(i+1)} \to \mathbb{F}$, where $\mathcal{D}^{(i+1)}$ is a (2-smooth) strict subgroup of $\mathcal{D}^{(i)}$, meaning that $\mathcal{D}^{(i)}$ is not only a subgroup of $\mathcal{D}^{(i+1)}$, but also its proper subset.

$\mathcal{D}^{(i+1)}$ divides $\mathcal{D}^{(i)}$ into cosets of size $l^{(i)} := |\mathcal{D}^{(i)}|/|\mathcal{D}^{(i+1)}|$. Let $C_g^{(i)}$ represent the coset corresponding to $g \in \mathcal{D}^{(i+1)}$, i.e.,

$$C_g^{(i)} := \{g' \in \mathcal{D}^{(i)} \mid (g')^{l^{(i)}} = g\}. \tag{8.1}$$

This means selecting those elements in $\mathcal{D}^{(i)}$ that can be mapped to g in $\mathcal{D}^{(i+1)}$ through the mapping $q(x) = x^{l^{(i)}}$, and these elements form the set $C_g^{(i)}$, which is also a coset.

For each coset $C_g^{(i)}$, the *interpolation map* $M_g^{(i)}$ is an invertible linear map $M_g^{(i)} : \mathbb{F}^{C_g^{(i)}} \to \mathbb{F}^{l^{(i)}}$, which maps $f^{(i)}|_{C_g^{(i)}} : C_g^{(i)} \to \mathbb{F}$ (i.e., restricting $f^{(i)}$ to the domain $C_g^{(i)} \subset \mathcal{D}^{(i)}$) to the coefficient vector $\mathbf{u}^{(i)}(g) = (u_0^{(i)}(g), \ldots, u_{l^{(i)}-1}^{(i)}(g))^\intercal$ of the polynomial $P^{(i)}_{\mathbf{u}^{(i)}(g)}(Z) = \sum_{j<l^{(i)}} u_j^{(i)}(g) Z^j$, where $P^{(i)}_{\mathbf{u}^{(i)}(g)}(Z)$ is the polynomial interpolating $f^{(i)}|_{C_g^{(i)}}$. In other words, $M_g^{(i)}$ is the inverse of the Vandermonde matrix generated by $C_g^{(i)}$, which means that $\left(M_g^{(i)}\right)^{-1} \cdot (u_0, \ldots, u_{l^{(i)}-1})^\intercal$ is the evaluation of the polynomial $P_{\mathbf{u}}(X) = \sum_{i<l^{(i)}} u_i X^i$ on the coset $C_g^{(i)}$.

> 👀 **Notice** To maintain consistency throughout this article, we use $(x_0, \ldots, x_n)$ to represent a row vector, and $(x_0, \ldots, x_n)^\intercal$ to represent a column vector, which can also be written as:
>
> $$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{18}$$

Let's explain the description of the interpolation map above in more detail. According to the definition of $C_g^{(i)}$, we know that it contains $l^{(i)}$ elements. Let $C_g^{(i)} = \{g'_1, \cdots, g'_{l^{(i)}}\}$, we can write out the Vandermonde matrix generated by $C_g^{(i)}$:

$$V_{C_g^{(i)}} = \begin{bmatrix} 1 & g'_1 & (g'_1)^2 & \cdots & (g'_1)^{l^{(i)}-1} \\ 1 & g'_2 & (g'_2)^2 & \cdots & (g'_2)^{l^{(i)}-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & g'_{l^{(i)}} & (g'_{l^{(i)}})^2 & \cdots & (g'_{l^{(i)}})^{l^{(i)}-1} \end{bmatrix} \tag{19}$$

Then $M_g^{(i)} = V^{-1}_{C_g^{(i)}}$, which is the inverse of the Vandermonde matrix generated by $C_g^{(i)}$, therefore

$$\begin{aligned} \left(M_g^{(i)}\right)^{-1} \cdot (u_0, \ldots, u_{l^{(i)}-1})^\intercal &= \left(V^{-1}_{C_g^{(i)}}\right)^{-1} \cdot (u_0, \ldots, u_{l^{(i)}-1})^\intercal \\ &= V_{C_g^{(i)}} \cdot (u_0, \ldots, u_{l^{(i)}-1})^\intercal \\ &= \begin{bmatrix} 1 & g'_1 & (g'_1)^2 & \cdots & (g'_1)^{l^{(i)}-1} \\ 1 & g'_2 & (g'_2)^2 & \cdots & (g'_2)^{l^{(i)}-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & g'_{l^{(i)}} & (g'_{l^{(i)}})^2 & \cdots & (g'_{l^{(i)}})^{l^{(i)}-1} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{l^{(i)}-1} \end{bmatrix} \\ &= \begin{bmatrix} u_0 + u_1 g'_1 + u_2 (g'_1)^2 + \cdots + u_{l^{(i)}-1} (g'_1)^{l^{(i)}-1} \\ u_0 + u_1 g'_2 + u_2 (g'_2)^2 + \cdots + u_{l^{(i)}-1} (g'_2)^{l^{(i)}-1} \\ \vdots \\ u_0 + u_1 g'_{l^{(i)}} + u_2 (g'_{l^{(i)}})^2 + \cdots + u_{l^{(i)}-1} (g'_{l^{(i)}})^{l^{(i)}-1} \end{bmatrix} \\ &= \begin{bmatrix} P_{\mathbf{u}}(g'_1) \\ P_{\mathbf{u}}(g'_2) \\ \vdots \\ P_{\mathbf{u}}(g'_{l^{(i)}}) \end{bmatrix} \end{aligned} \tag{20}$$

From the above derivation, we can see that $\left(P_{\mathbf{u}}(g'_1), P_{\mathbf{u}}(g'_2) \cdots, P_{\mathbf{u}}(g'_{l^{(i)}})\right)^\intercal$ is the evaluation of the polynomial $P_{\mathbf{u}}(X) = \sum_{i<l^{(i)}} u_i X^i$ on the coset $C_g^{(i)}$, so $\left(M_g^{(i)}\right)^{-1} \cdot (u_0, \ldots, u_{l^{(i)}-1})^\intercal$ is the evaluation of the polynomial $P_{\mathbf{u}}(X) = \sum_{i<l^{(i)}} u_i X^i$ on the coset $C_g^{(i)}$.

The following proposition uses the above notation and restates [BBHR18, Section 4.1], differing from [BBHR18, Section 4.1] in that it is done over a multiplicative group rather than an additive group. This proposition describes the property of maintaining low-degree.

Claim 1 [BCIKS20, Claim 8.1]. Suppose that $f^{(i)} \in \mathrm{RS}[\mathbb{F}, \mathcal{D}^{(i)}, k^{(i)}]$ where $k^{(i)}+1$ is an integral power of 2. Then, for any $z^{(i)} \in \mathbb{F}$, letting $\mathbf{z}^{(i)} = \left(\left(z^{(i)}\right)^0, \left(z^{(i)}\right)^1, \ldots, \left(z^{(i)}\right)^{l^{(i)}-1}\right)^{\mathsf{T}}$, the function $f^{(i+1)}_{f^{(i)},z^{(i)}} : \mathcal{D}^{(i+1)} \to \mathbb{F}$ defined on $g \in \mathcal{D}^{(i+1)}$ by

$$f^{(i+1)}_{f^{(i)},z^{(i)}}(g) := \left(\mathbf{z}^{(i)}\right)^{\mathsf{T}} \cdot \mathbf{u}^{(i)}(g) = \left(\mathbf{z}^{(i)}\right)^{\mathsf{T}} \cdot M^{(i)}_g \cdot f^{(i)}|_{C^{(i)}_g} \tag{2}$$

is a valid codeword of $V^{(i+1)} := \mathrm{RS}[\mathbb{F}, \mathcal{D}^{(i+1)}, k^{(i+1)}]$ where $k^{(i+1)} := \frac{k^{(i)}+1}{l^{(i)}} - 1$.

According to [BBHR18] and the above notation, in the COMMIT phase of the FRI protocol, fixing a $g \in \mathcal{D}^{(i+1)}$, the next step constructs $f^{(i+1)}_{f^{(i)},z^{(i)}}(g) := P^{(i)}_{\mathbf{u}^{(i)}(g)}(z^{(i)})$. Let's understand the construction formula above.

$$\begin{aligned}
f^{(i+1)}_{f^{(i)},z^{(i)}}(g) &= P^{(i)}_{\mathbf{u}^{(i)}(g)}(z^{(i)}) \\
&= \sum_{j<l^{(i)}} u^{(i)}_j(g) \cdot (z^{(i)})^j \\
&= \left(z^{(i)}\right)^0 \cdot u^{(i)}_0(g) + \left(z^{(i)}\right)^1 \cdot u^{(i)}_1(g) + \cdots + \left(z^{(i)}\right)^{l^{(i)}-1} \cdot u^{(i)}_{l^{(i)}-1}(g) \\
&= \begin{bmatrix} \left(z^{(i)}\right)^0 & \left(z^{(i)}\right)^1 & \cdots & \left(z^{(i)}\right)^{l^{(i)}-1} \end{bmatrix} \cdot \begin{bmatrix} u^{(i)}_0(g) \\ u^{(i)}_1(g) \\ \vdots \\ u^{(i)}_{l^{(i)}-1}(g) \end{bmatrix} \\
&= \left(\mathbf{z}^{(i)}\right)^{\mathsf{T}} \cdot \mathbf{u}^{(i)}(g)
\end{aligned} \tag{21}$$

Let's explain the second equation given in Proposition 1, i.e., $\mathbf{u}^{(i)}(g) = M^{(i)}_g \cdot f^{(i)}|_{C^{(i)}_g}$. According to the previous analysis, for the Vandermonde matrix generated by $C^{(i)}_g$, we have

$$V_{C^{(i)}_g} \cdot \begin{bmatrix} u^{(i)}_0 \\ u^{(i)}_1 \\ \vdots \\ u^{(i)}_{l^{(i)}-1} \end{bmatrix} = \begin{bmatrix} P_{\mathbf{u}}(g'_1) \\ P_{\mathbf{u}}(g'_2) \\ \vdots \\ P_{\mathbf{u}}(g'_{l^{(i)}}) \end{bmatrix} = \begin{bmatrix} f^{(i)}|_{C^{(i)}_g}(g'_1) \\ f^{(i)}|_{C^{(i)}_g}(g'_2) \\ \vdots \\ f^{(i)}|_{C^{(i)}_g}(g'_{l^{(i)}}) \end{bmatrix} \tag{22}$$

Therefore

$$\begin{bmatrix} u^{(i)}_0 \\ u^{(i)}_1 \\ \vdots \\ u^{(i)}_{l^{(i)}-1} \end{bmatrix} = \left(V_{C^{(i)}_g}\right)^{-1} \cdot \begin{bmatrix} f^{(i)}|_{C^{(i)}_g}(g'_1) \\ f^{(i)}|_{C^{(i)}_g}(g'_2) \\ \vdots \\ f^{(i)}|_{C^{(i)}_g}(g'_{l^{(i)}}) \end{bmatrix} = M^{(i)}_g \cdot f^{(i)}|_{C^{(i)}_g} \tag{23}$$

This gives us $\left(\mathbf{u}^{(i)}(g)\right)^{\mathsf{T}} = M^{(i)}_g \cdot f^{(i)}|_{C^{(i)}_g}$.

Batching

In some cases, the first prover's oracle $f^{(0)}$ is sampled from functions in an affine space $F \subset \mathbb{F}^{\mathcal{D}^{(0)}}$, which serves as our input,

$$F = \left\{ f^{(0)}_0 + \sum_{i=1}^{t} x_i \cdot f^{(0)}_i \mid x_i \in \mathbb{F}, f_i : \mathcal{D}^{(0)} \to \mathbb{F} \right\} \tag{3}$$

When using the FRI protocol to "batch" multiple instances of different low degree testing problems, we combine them all together through random linear combinations, i.e., $f^{(0)}_0 + x_1 f^{(0)}_1 + \cdots + x_t f^{(0)}_t$ in the above formula. In this batching setting, we assume that the prover has already committed to $f^{(0)}_1, \ldots, f^{(0)}_t$ (note that in this case we set $f^{(0)}_0 = 0$), and the verifier of the batched FRI uniformly randomly samples $x_1, \ldots, x_t \in \mathbb{F}$, the prover replies with $f^{(0)}$, which should equal $f^{(0)}_0 + \sum_{i=1}^{t} x_i \cdot f^{(0)}_i$, and now the FRI protocol is applied to $f^{(0)}$. Correspondingly, the QUERY phase of the batched FRI is also extended, so that each time a query for $f^{(0)}(g)$ is requested, the verifier also queries $f^{(0)}_0(g), \ldots, f^{(0)}_t(g)$ and verifies $f^{(0)}(g) = f^{(0)}_0(g) + \sum_{i=1}^{t} x_i \cdot f^{(0)}_i(g)$.

> 🐞 **Fix**
>
> - [ ] The formula in the paper here is $f^{(0)}(g) = f_0^{(0)}(g) + \sum_{i=1}^{t} f_i^{(0)}(g)$. I think it's missing the coefficient x_i from before, and should be $f^{(0)}(g) = f_0^{(0)}(g) + \sum_{i=1}^{t} x_i \cdot f_i^{(0)}(g)$.

The (batched) FRI QUERY phase

Proposition 1 shows that for an honest prover, for any value $z^{(i)}$ chosen by the verifier, for each $y \in D^{(i+1)}$, the prover can construct a new codeword $f^{(i+1)} \in V^{(i+1)}$ from a codeword $f^{(i)} \in V^{(i)}$ by calculating equation (2). Therefore, we will always assume that $f^{(r)} \in V^{(r)}$, for example, by assuming that the verifier always queries the first $k^{(r)}$ elements of $f^{(r)}$ (in some canonical order) and compares $f^{(r)}$ with the interpolation polynomial of this function.

Proposition 1 provides a very natural testing method to check the consistency between $f^{(i)}$ and $f^{(i+1)}$, and the query phase of FRI follows this process by iteratively applying this natural test from the "top" ($f^{(r)}$) to the "bottom" ($f^{(0)}$).

> 🤔 **Question**
>
> - [ ] How to better explain this natural testing method here?

A single invocation of the FRI QUERY phase

1. Choose $g^{(r)}$ uniformly randomly from $\mathcal{D}^{(r)}$. For $i = r, \cdots, 1$, choose $g^{(i-1)}$ uniformly randomly from the coset $C_{g^{(i)}}^{(i-1)}$.
2. If $f^{(0)}(g^{(0)}) \neq f_0^{(0)}(g^{(0)}) + \sum_{i=1}^{t} x_i \cdot f_i^{(0)}(g^{(0)})$, then reject.
3. If, for any $i \in \{0, \cdots, r-1\}$, $f^{(i+1)}(g^{(i+1)}) \neq \left(\mathbf{z}^{(i)}\right)^{\mathsf{T}} \cdot M_g^{(i)} \cdot f^{(i)}|_{C_g^{(i)}}$, then reject.
4. Otherwise — if all equations in the above conditions hold, then accept.

The above QUERY process differs from the QUERY process of FRI in [BBHR18] in that the random number selection starts from the last $\mathcal{D}^{(r)}$ instead of from the initial $\mathcal{D}^{(0)}$. Compared to the QUERY phase in [BBHR18], here we also want to verify whether the batch is correct at step 0, that is, $f^{(0)}(g^{(0)}) \neq f_0^{(0)}(g^{(0)}) + \sum_{i=1}^{t} x_i \cdot f_i^{(0)}(g^{(0)})$.

Summary of the batched FRI protocol

Let's summarize the important properties mentioned so far, which will be used in the following soundness analysis.

1. At the end of the COMMIT phase of the protocol, the verifier can access through oracles a series of functions $f^{(0)} : \mathcal{D}^{(0)} \to \mathbb{F}, \ldots, f^{(r)} : \mathcal{D}^{(r)} \to \mathbb{F}$, where $\mathcal{D}^{(0)} \supsetneq \ldots \supsetneq \mathcal{D}^{(r)}$ is a series of 2-smooth groups, and $f^{(i)}$ arbitrarily depends on $z^{(0)}, \ldots, z^{(i)}$ (and $f^{(0)}, \ldots, f^{(i-1)}$). We assume $f^{(r)} \in V^{(r)}$.
2. There exists a set of $l^{(i)} \times l^{(i)}$ invertible matrices $\{M_{g^{(i+1)}}^{(i)} : g^{(i+1)} \in D^{(i+1)}\}$, so that applying $M_{g^{(i+1)}}^{(i)}$ to $f^{(i)}|_{C_{g(i+1)}^{(i)}}$ maps $f^{(i)}$ to a sequence of vectors $\mathbf{u} = \mathbf{u}^{(i)} = \{u_0^{(i)}, \ldots, u_{l(i)}^{(i)}\} \subset \mathbb{F}^{D^{(i+1)}}$, where

$$\mathbf{u}^{(i)}\left(g^{(i+1)}\right) = \left(u_0^{(i)}\left(g^{(i+1)}\right), \cdots, u_{l^{(i)}-1}^{(i)}\left(g^{(i+1)}\right)\right) = M_{g^{(i+1)}}^{(i)} \cdot f^{(i)}|_{C_{g(i+1)}^{(i)}}. \tag{4}$$

Certainly. I'll translate the text into English while maintaining all formulas and markdown formats. Here's the translation:

Moreover, if $f^{(i)}$ is a valid RS codeword with rate ρ on $D^{(i)}$, then each vector on the parametric curve through $\mathbf{u}^{(i)}$ is also a valid RS codeword with rate ρ on $D^{(i+1)}$.

1. In each iteration of the QUERY phase, it checks whether $f^{(i+1)}$ is constructed from $f^{(i)}$ through equation (2), and (in the batched case) checks whether $f^{(0)}$ is correctly calculated through equation (3).

Soundness

Lemma 8.2 [BSCIK20, Lemma 8.2] (batched FRI error bound). Let $V^{(0)} = \mathrm{RS}[\mathbb{F}, \mathcal{D}^{(0)}, k^{(0)}]$ where $\mathcal{D}^{(0)}$ is a coset of a 2-smooth multiplicative group, and $k^{(0)} + 1$ is a power of 2; set $\rho = (k^{(0)} + 1)/|\mathcal{D}^{(0)}|$.

Let $F \subseteq \mathbb{F}^{\mathcal{D}^{(0)}}$ be a space of functions as defined in Eq. (3) whose correlated agreement density with $V^{(0)}$ is α. For integer $m \geq 3$, let

$$\alpha^{(0)}(\rho, m) = \max\{\alpha, \sqrt{\rho}(1 + 1/2m)\}. \tag{24}$$

Assume the FRI protocol is used with r rounds, and let $l^{(i)} = |\mathcal{D}^{(i)}|/|\mathcal{D}^{(i+1)}|$ denote the ratio between prover messages (oracles) i and $i+1$. Let ϵ_Q denote the probability that the verifier accepts a single FRI QUERY invocation. Then,

$$\Pr_{x_1,\ldots,x_t,z^{(0)},\ldots,z^{(r-1)}}\left[\epsilon_Q > \alpha^{(0)}(\rho, m)\right] \leq \epsilon_{\mathrm{C}}, \tag{8.5}$$

where

$$\epsilon_{\mathrm{C}} = \frac{\left(m+\frac{1}{2}\right)^7 \cdot |\mathcal{D}^{(0)}|^2}{2\rho^{3/2}|\mathbb{F}|} + \frac{(2m+1)\cdot(|\mathcal{D}^{(0)}|+1)}{\sqrt{\rho}} \cdot \frac{\sum_{i=0}^{r-1} l^{(i)}}{|\mathbb{F}|}. \tag{25}$$

In words: For any interactive FRI prover P^*, the probability that the oracles $f^{(0)},\ldots,f^{(r)}$ sent by P^* will pass a single invocation of the batched FRI QUERY test with probability greater than $\alpha^{(0)}(\rho, m)$, is smaller than ϵ_{C}. The probability is over the random variables $x_1,\ldots,x_t$ used to sample $f^{(0)}$ from F and over the random messages $z^{(0)},\ldots,z^{(r-1)}$ sent by the verifier during the COMMIT phase.

Theorem 8.3 [BSCIK20, Theorem 8.3] (Batched FRI Soundness). Let $f_0^{(0)},\ldots,f_t^{(r)} : \mathcal{D}^{(0)} \to \mathbb{F}$ be a sequence of functions and let $V^{(0)} = \mathrm{RS}[\mathbb{F}, \mathcal{D}^{(0)}, k^{(0)}]$ where $\mathcal{D}^{(0)}$ is a coset of a 2-smooth group of size $n^{(0)} = |\mathcal{D}^{(0)}|$, and $\rho = \frac{k^{(0)}+1}{n^{(0)}}$ satisfies $\rho = 2^{-\mathrm{R}}$ for positive integer R. Let $\alpha = \sqrt{\rho}(1+1/2m)$ for integer $m \geq 3$ and ϵ_{C} be as defined in Lemma 8.2.

Assume the FRI protocol is used with r rounds. Let $l^{(i)} = |\mathcal{D}^{(i)}|/|\mathcal{D}^{(i+1)}|$ denote the ratio between prover messages (oracles) i and $i+1$. Assume furthermore that s is 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 probability greater than

$$\epsilon_{\mathrm{FRI}} := \epsilon_{\mathrm{C}} + \alpha^s = \frac{\left(m+\frac{1}{2}\right)^7 \cdot |\mathcal{D}^{(0)}|^2}{2\rho^{3/2}|\mathbb{F}|} + \frac{(2m+1)\cdot(|\mathcal{D}^{(0)}|+1)}{\sqrt{\rho}} \cdot \frac{\sum_{i=0}^{r-1} l^{(i)}}{|\mathbb{F}|} + \left(\sqrt{\rho}\cdot\left(1+\frac{1}{2m}\right)\right)^s. \tag{26}$$

Then $f_0^{(0)},\ldots,f_t^{(0)}$ have correlated agreement with $V^{(0)}$ on a domain $\mathcal{D}' \subset \mathcal{D}^{(0)}$ of density at least α.

Proof of Theorem 8.3: By contradiction, then directly prove through Lemma 8.2. Assume that the maximum correlated agreement of $f_0^{(0)},\ldots,f_t^{(0)}$ with $V^{(0)}$ is less than $\alpha^{(0)}(\rho, m) = \sqrt{\rho}(1+1/2m)$, but at the same time, the acceptance probability is greater than $\epsilon_{\mathrm{C}} + (\alpha^{(0)}(\rho, m))^s$.

Let E be the event that the acceptance probability in each FRI QUERY phase is greater than $\alpha^{(0)}(\rho, m)$. This event depends on $x_1,\ldots,x_t, f^{(0)}, z^{(0)},\ldots,z^{(r-1)}, f^{(r)}$, where each $f^{(i)}$ is generated by P^* based on the previous messages from the Verifier. According to Lemma 8.2, for any Prover P^*, the probability of event E occurring does not exceed ϵ_{C}. When event E does not hold, the probability that s independent invocations of FRI QUERY all return "accept" does not exceed $(\alpha^{(0)}(\rho, m))^s$.

Therefore, the probability that the FRI Verifier accepts does not exceed $\epsilon_{\mathrm{C}} + (\alpha^{(0)}(\rho, m))^s$, which contradicts the assumption. □

> ❓ **Question**
>
> Here, for the event E where the acceptance probability is greater than $\alpha^{(0)}(\rho, m)$ in each FRI QUERY phase, how is it still not exceeding ϵ_{C} when called s times? Why isn't it $(\epsilon_{\mathrm{C}})^s$?

Proof of Lemma 8.2

Before proving Lemma 8.2, let's introduce a method to track whether the verifier's consistency check passes. Specifically, the Prover will construct function $f^{(i+1)}$ based on the random number $z^{(i)}$ sent by the Verifier, and then respond to the Verifier with function $f^{(i+1)}$. In the QUERY phase of FRI, the Verifier will check the consistency between function $f^{(i+1)}$ and function $f^{(i)}$.

Define a series of weight functions, $\mu^{(i)} : \mathcal{D}^{(i)} \to [0,1]$ and $\nu^{(i)} : \mathcal{D}^{(i)} \to [0,1]$, where $i = 0,\ldots,r$. These weight functions are defined by induction. When $i = 0$, use $\{0,1\}$ weights to indicate whether $f^{(0)}(g)$ is calculated correctly:

$$\mu^{(0)}(g) = \begin{cases} 1 & f^{(0)}(g) = f_0^{(0)}(g) + \sum_{i=1}^{t} x_i f_i^{(0)}(g) \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

Now, $\mu^{(i)}$ obtained by induction can be used to define an auxiliary weight function $\nu^{(i+1)} : \mathcal{D}^{(i+1)} \to [0,1]$. By taking an element g in $\mathcal{D}^{(i+1)}$, we can get a coset $C_g^{(i)} \subset \mathcal{D}^{(i)}$, which is composed of all elements in $\mathcal{D}^{(i)}$ that can be mapped to g through the mapping $q^{(i)}(x) = x^{l^{(i)}}$, as shown in (8.1),

$$C_g^{(i)} := \{g' \in \mathcal{D}^{(i)} \mid (g')^{l^{(i)}} = g\}. \tag{28}$$

Then the definition of $\nu^{(i+1)}$ is

$$\nu^{(i+1)}(g) = \mathbb{E}_{g' \in C_g^{(i)}} \left[\mu^{(i)}(g') \right]. \tag{8.6}$$

In other words, $\nu^{(i+1)}(g)$ is the expected value of the $\mu^{(i)}$ weights of all elements in the coset $C_g^{(i)}$. Finally, define the function $\mu^{(i+1)}$, for each $g \in \mathcal{D}^{(i+1)}$:

$$\mu^{(i+1)}(g) = \begin{cases} \nu^{(i+1)}(g) & f^{(i+1)}(g) = f^{(i+1)}_{f^{(i)},z^{(i)}}(g) \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

Regarding the definition of $\mu^{(i)}$, an important property is that $\mu^{(i)}(g)$ is a measure of the probability of success in the FRI QUERY phase, conditional on querying g from $f^{(i)}$, which is an important reason for the following proposition to hold.

Claim 8.5. The probability ϵ_{Q} that a single invocation of the batched FRI QUERY accepts $f^{(0)}, \ldots, f^{(r)}$, where $f^{(r)} \in \mathrm{RS}[\mathbb{F}, \mathcal{D}^{(r)}, k^{(r)}]$, satisfies

$$\epsilon_{\mathrm{Q}} = \mathbb{E}_{g^{(r)} \in \mathcal{D}^{(r)}} \left[\mu^{(r)}(g^{(r)}) \right]. \tag{30}$$

Proof: Recall the invocation of FRI QUERY, a series of random $g^{(r)}, \ldots, g^{(0)}$ will be selected, where $g^{(i-1)}$ is uniformly randomly selected from the coset $C_{g^{(i)}}^{(i-1)}$. We will prove by induction that for $i = 0, \ldots, r$

$$\mathbb{E}_{g^{(i)} \in \mathcal{D}^{(i)}} \left[\mu^{(i)}(g^{(i)}) \right] \tag{31}$$

equals the probability that when $g^{(i)}$ is uniformly randomly selected, and it is generated from a random sequence $g^{(i-1)} \in C_{g^{(i)}}^{(i-1)}, \ldots, g^{(0)} \in C_{g^{(1)}}^{(0)}$, all tests related to $g^{(i)}$ and its induced tests pass.

The idea of the induction proof is as follows:

1. Prove that $\mu^{(0)}$ holds for the most basic case when $i = 0$
2. Assume $\mu^{(i-1)}$ holds for $i-1$, prove that $\mu^{(i)}$ holds for i

This will prove the proposition.

When $i = 0$, by the definition of $\mu^{(0)}$,

$$\mu^{(0)}(g^{(0)}) = \begin{cases} 1 & f^{(0)}(g^{(0)}) = f_0^{(0)}(g^{(0)}) + \sum_{i=1}^{t} x_i f_i^{(0)}(g^{(0)}) \\ 0 & \text{otherwise} \end{cases} \tag{32}$$

The probability of FRI QUERY passing naturally equals $\mathbb{E}_{g^{(0)} \in \mathcal{D}^{(0)}} \left[\mu^{(0)}(g^{(0)}) \right]$.

Assuming $\mu^{(i-1)}$ holds for $i-1$, now analyze $\mu^{(i)}(g^{(i)})$. If $f^{(i)}(g^{(i)})$ is not calculated correctly according to equation (8.2), then $\mu^{(i)}(g^{(i)}) = 0$, otherwise, according to the definition

$$\mu^{(i)}(g^{(i)}) = \nu^{(i)}(g^{(i)}) = \mathbb{E}_{g^{(i-1)} \in C_{g^{(i)}}^{(i-1)}} \left[\mu^{(i-1)}(g^{(i-1)}) \right]. \tag{33}$$

This indicates that $\mu^{(i)}(g^{(i)})$ is the average of the values of $\mu^{(i-1)}$ on the coset $C_{g^{(i)}}^{(i-1)} \subseteq \mathcal{D}^{(i-1)}$. By the induction hypothesis, it is the probability that all tests related to $g^{(i-1)}, \ldots, g^{(0)}$ pass, therefore $\mu^{(i)}$ holds for i. □

Lemma 8.2 needs to estimate the probability in the FRI QUERY phase. Recall the protocol of the batched FRI QUERY phase, there are two places involving random numbers:

1. In step 2 of the protocol, use t random numbers $x_1, \ldots, x_t$ to batch $f_1^{(0)}, f_2^{(0)}, \ldots, f_t^{(0)}$, which corresponds to the case of affine space, and will use the conclusion corresponding to Theorem 7.4.
2. In step 3 of the protocol, use $\mathbf{z}^{(i)} = \left(\left(z^{(i)}\right)^0, \left(z^{(i)}\right)^1, \ldots, \left(z^{(i)}\right)^{l^{(i)}-1} \right)$ for batching, corresponding to the case of curves, and will use the conclusion of Theorem 7.2.

Proof of Lemma 8.2: Now we need to prove Lemma 8.2. By Proposition 8.5, we only need to prove that in the verifier's random selection, with probability greater than $1 - \epsilon_C$,

$$\mathbb{E}_{g \in \mathcal{D}^{(r)}} \left[\mu^{(r)}(g) \right] \leq \alpha^{(0)}(\rho, m). \tag{8.7}$$

If we prove the above holds, it means that when selecting random numbers in $\mathbb{F}_q$, if $\epsilon_Q > \alpha^{(0)}(\rho, m)$, then its probability is less than or equal to ϵ_C, which proves Lemma 8.2.

> 🤔 **Question**
>
> - [ ] Why isn't it "with probability greater than $1 - \epsilon_C$" here?

The proof idea is to first define a series of bad events $E^{(0)}, \ldots, E^{(r)}$, where the probability of some events occurring is the sum of the probabilities of each event occurring, proving that this probability is less than or equal to ϵ_C. Then assuming that no bad events occur, prove that equation (8.7) holds.

Let $E^{(0)}$ be the event

$$\operatorname{agree}_{\mu^{(0)}}\left(f^{(0)}, V^{(0)}\right) > \alpha^{(0)}(\rho, m). \tag{34}$$

By the definition of $\mu^{(0)}$, event $E^{(0)}$ is

$$\operatorname{agree}\left(f_0^{(0)} + \sum_{i=1}^{t} x_i f_i^{(0)}, V^{(0)}\right) > \alpha^{(0)}(\rho, m) = \max\{\alpha, \sqrt{\rho}(1 + 1/2m)\}. \tag{35}$$

> 🤔 **Question**
>
> - [ ] What exactly does agree mean here? What's the difference between it and $\operatorname{agree}_{\mu^{(0)}}$? Does it represent the constant 1?

Therefore, this event $E^{(0)}$ mainly depends on the random numbers $x_1, \ldots, x_t$. According to the assumption in the lemma, the maximum correlated agreement density of $(f_0^{(0)}, \ldots, f_t^{(0)})$ with $V^{(0)}$ does not exceed α.

> Recall Theorem 7.4: **Theorem 7.4** (Weighted correlated agreement over affine spaces – Version II). Let V, q, n, k and ρ be as defined in Theorem 1.2. Let $\mathbf{u} = \{u_0, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$ and let $U = u_0 + \operatorname{span}\{u_1, \cdots, u_l\} \subset \mathbb{F}_q^{\mathcal{D}}$ be an affine subspace. Let $\mu : \mathcal{D} \to [0, 1]$ be a vector of weights, whose values all have denominator M . Let $m \geq 3$ and let
>
> $$\alpha \geq \alpha_0(\rho, m) := \sqrt{\rho} + \frac{\sqrt{\rho}}{2m}. \tag{36}$$
>
> Suppose
>
> $$\Pr_{u \in U}[\operatorname{agree}_\mu(u, V) \geq \alpha] > \max\left(\frac{(1 + \frac{1}{2m})^7 m^7}{3\rho^{3/2}} \cdot \frac{n^2}{q}, \frac{2m + 1}{\sqrt{\rho}} \cdot \frac{M \cdot n + 1}{q}\right). \tag{7.2}$$
>
> Then $u_0, \ldots, u_l$ have at least α correlated μ -agreement with V , i.e. $\exists v_0, \cdots, v_l \in V$ such that
>
> $$\mu\left(\{x \in \mathcal{D} : \forall 0 \leq i \leq l, u_i(x) = v_i(x)\}\right) \geq \alpha. \tag{37}$$
>
> Its contrapositive is: If $u_0, \ldots, u_l$ have at most α correlated μ -agreement with V, then
>
> $$\Pr_{u \in U}[\operatorname{agree}_\mu(u, V) \geq \alpha] \leq \max\left(\frac{(1 + \frac{1}{2m})^7 m^7}{3\rho^{3/2}} \cdot \frac{n^2}{q}, \frac{2m + 1}{\sqrt{\rho}} \cdot \frac{M \cdot n + 1}{q}\right). \tag{38}$$

By the contrapositive of Theorem 7.4, taking $\alpha = \alpha^{(0)}(\rho, m)$, $\mu \equiv 1$ and $M = 1$, we have

$$\begin{aligned}
\Pr_{x_1, \ldots, x_t}[E^{(0)}] &= \Pr_{u \in U}[\operatorname{agree}_\mu(u, V) > \alpha^{(0)}(\rho, m)] \\
&\quad \text{(Why can the parentheses here be directly changed to strictly } >?) \\
&\leq \max\left(\frac{(1 + \frac{1}{2m})^7 m^7}{3\rho^{3/2}} \cdot \frac{n^2}{q}, \frac{2m + 1}{\sqrt{\rho}} \cdot \frac{M \cdot n + 1}{q}\right) \\
&= \max\left(\frac{(m + \frac{1}{2})^7}{3\rho^{3/2}} \cdot \frac{n^2}{q}, \frac{2m + 1}{\sqrt{\rho}} \cdot \frac{n + 1}{q}\right)
\end{aligned} \tag{39}$$

Note that according to Theorem 7.4 and Theorem 1.2, $V = \mathsf{RS}[\mathbb{F}_q, \mathcal{D}^{(0)}, k^{(0)}]$, $n = |\mathcal{D}^{(0)}|$, $\rho = \frac{k^{(0)}+1}{n}$.

Let's derive

$$\frac{(m + \frac{1}{2})^7}{3\rho^{3/2}} \cdot \frac{n^2}{q} > \frac{2m + 1}{\sqrt{\rho}} \cdot \frac{n + 1}{q} \tag{40}$$

Since

$$\begin{aligned}&\frac{(m+\frac{1}{2})^7}{3} \geq 2m+1\\ \Rightarrow &\frac{(2m+1)^7}{3\times 2^7} \geq 2m+1\\ \Rightarrow &(2m+1)^6 \geq 3\times 2^7\end{aligned} \tag{41}$$

By the condition $m \geq 3$ in the theorem, $(2m+1)^6$ is an increasing function when $m \geq 3$, so $(2m+1)^6 \geq (2\times 3+1)^6 = 7^6 = 117649$, while the right side of the above equation $3\times 2^7 = 384$, satisfying $(2m+1)^6 \geq 117649 > 3\times 2^7$. From this, we get that $\frac{(m+\frac{1}{2})^7}{3} > 2m+1$ (not equal) holds. Then

$$\begin{aligned}\frac{(m+\frac{1}{2})^7}{3\rho^{3/2}}\cdot\frac{n^2}{q} &> \frac{2m+1}{\rho^{3/2}}\cdot\frac{n^2}{q}\\ &= \frac{2m+1}{\rho\cdot\rho^{1/2}}\cdot\frac{n^2}{q}\\ &\text{(Since } \rho<1\text{)}\\ &> \frac{2m+1}{\sqrt{\rho}}\cdot\frac{n^2}{q}\\ &\text{(Since } n^2 > n+1 \text{ when } n\geq 2\text{)}\\ &> \frac{2m+1}{\sqrt{\rho}}\cdot\frac{n+1}{q}\end{aligned} \tag{42}$$

Thus

$$\begin{aligned}\Pr_{x_1,\ldots,x_t}[E^{(0)}] &\leq \max\left(\frac{(m+\frac{1}{2})^7}{3\rho^{3/2}}\cdot\frac{n^2}{q}, \frac{2m+1}{\sqrt{\rho}}\cdot\frac{n+1}{q}\right)\\ &= \frac{(m+\frac{1}{2})^7}{3\rho^{3/2}}\cdot\frac{n^2}{q}\end{aligned} \tag{43}$$

Let

$$\epsilon = \frac{(m+\frac{1}{2})^7}{3\rho^{3/2}}\cdot\frac{n^2}{q} \quad \text{(Note that here } n = |\mathcal{D}^{(0)}|\text{)} \tag{44}$$

We get

$$\Pr_{x_1,\ldots,x_t}[E^{(0)}] \leq \epsilon \tag{8.8}$$

Now fix $i \in \{0,\ldots,r-1\}$. Define event $E^{(i+1)}$ as

$$\text{agree}_{\nu^{(i+1)}}\left(f^{(i+1)}_{f^{(i)},z^{(i)}}, V^{(i+1)}\right) > \max\left(\text{agree}_{\mu^{(i)}}\left(f^{(i)}, V^{(i)}\right), \sqrt{\rho}(1+1/2m)\right). \tag{8.9}$$

> 📖 **Notes** Understand event $E^{(i+1)}$. According to the definition
>
> $$\begin{aligned}\text{agree}_{\nu^{(i+1)}}\left(f^{(i+1)}_{f^{(i)},z^{(i)}}, V^{(i+1)}\right) &= \max_{g^{(i+1)}\in V^{(i+1)}} \text{agree}_{\nu^{(i+1)}}\left(f^{(i+1)}_{f^{(i)},z^{(i)}}, g^{(i+1)}\right)\\ &= \max_{g^{(i+1)}\in V^{(i+1)}} \frac{1}{|\mathcal{D}^{(i+1)}|}\sum_{x:f^{(i+1)}_{f^{(i)},z^{(i)}}(x)=g^{(i+1)}(x)} \nu^{(i+1)}(x)\\ &= \max_{g^{(i+1)}\in V^{(i+1)}} \frac{1}{|\mathcal{D}^{(i+1)}|}\sum_{x:f^{(i+1)}_{f^{(i)},z^{(i)}}(x)=g^{(i+1)}(x)} \mathbb{E}_{g'\in C^{(i)}_x}\left[\mu^{(i)}(g')\right]\end{aligned} \tag{45}$$
>
> It measures, after constructing $f^{(i+1)}_{f^{(i)},z^{(i)}}$ from $f^{(i)}$ and random number $z^{(i)}$, finding x in $\mathcal{D}^{(i+1)}$ that make $f^{(i+1)}_{f^{(i)},z^{(i)}}$ consistent with a polynomial $g^{(i+1)}$ in $V^{(i+1)}$, then calculating the sum of the expected values of the $\mu^{(i)}$ weights of the elements in the corresponding cosets in $\mathcal{D}^{(i)}$ for these x.
>
> On the right side of equation (8.9)

$$\operatorname{agree}_{\mu^{(i)}}\left(f^{(i)}, V^{(i)}\right)=\max_{g^{(i)}\in V^{(i)}}\frac{1}{|\mathcal{D}^{(i)}|}\sum_{x:f^{(i)}(x)=g^{(i)}(x)}\mu^{(i)}(x) \tag{46}$$

$\mu^{(i)}(x)$ measures the probability of passing in the FRI QUERY phase when querying x from $f^{(i)}$.

Event $E^{(i+1)}$ aims to define such events: for $f^{(i+1)}_{f^{(i)},z^{(i)}}$ constructed from $f^{(i)}$ and $z^{(i)}$, for a polynomial $g^{(i+1)}$ in $V^{(i+1)}$, take out the set of points x that make their values equal, calculate the sum of the expected $\mu^{(i)}$ weights of the cosets corresponding to these points, and the ratio to the size of $\mathcal{D}^{(i+1)}$.

Fix $f^{(i)}$ and $\mu^{(i)}$, then event $E^{(i+1)}$ is determined by the random number $z^{(i)}$. According to the definition of $\mu^{(i+1)}$, we have

$$\mu^{(i+1)}(g)=\begin{cases}\nu^{(i+1)}(g) & f^{(i+1)}(g)=f^{(i+1)}_{f^{(i)},z^{(i)}}(g)\\ 0 & \text{otherwise}\end{cases} \tag{47}$$

Only when the condition $f^{(i+1)}(g)=f^{(i+1)}_{f^{(i)},z^{(i)}}(g)$ is satisfied, $\mu^{(i+1)}(g)$ will be equal to $\nu^{(i+1)}(g)$. Naturally, we can get

$$\operatorname{agree}_{\mu^{(i+1)}}\left(f^{(i+1)}, V^{(i+1)}\right)\le\operatorname{agree}_{\nu^{(i+1)}}\left(f^{(i+1)}_{f^{(i)},z^{(i)}}, V^{(i+1)}\right) \tag{48}$$

Therefore, if event E^{i+1} does not occur, then according to equation (8.9), we can get

$$\operatorname{agree}_{\mu^{(i+1)}}\left(f^{(i+1)}, V^{(i+1)}\right)\le\operatorname{agree}_{\nu^{(i+1)}}\left(f^{(i+1)}_{f^{(i)},z^{(i)}}, V^{(i+1)}\right)\le\max\left(\operatorname{agree}_{\mu^{(i)}}\left(f^{(i)}, V^{(i)}\right), \sqrt{\rho}(1+1/2m)\right) \tag{49}$$

Then

$$\operatorname{agree}_{\mu^{(i+1)}}\left(f^{(i+1)}, V^{(i+1)}\right)\le\max\left(\operatorname{agree}_{\mu^{(i)}}\left(f^{(i)}, V^{(i)}\right), \sqrt{\rho}(1+1/2m)\right) \tag{8.10}$$

Let $\alpha=\max\left(\operatorname{agree}_{\mu^{(i)}}\left(f^{(i)}, V^{(i)}\right), \sqrt{\rho}(1+1/2m)\right)$. According to the definition, expanding $f^{(i+1)}_{f^{(i)},z^{(i)}}$, we get that event $E^{(i+1)}$ is

$$\operatorname{agree}_{\nu^{(i+1)}}\left(u_0+z^{(i)}u_1+\ldots+(z^{(i)})^{l^{(i)}-1}u_{l^{(i)}-1}, V^{(i+1)}\right)>\alpha, \tag{50}$$

where $u_0,\ldots,u_{l^{(i)}-1}:\mathcal{D}^{(i+1)}\to\mathbb{F}$ are the functions obtained from $f^{(i)}$ as defined in the FRI protocol (see Proposition 8.1). This is exactly the case handled by Theorem 7.2.

> 📖 **Recall Theorem 7.2**
>
> **Theorem 7.2** (Weighted correlated agreement over curves – Version II). Let V, q, n, k and ρ be as defined in Theorem 1.2. Let $\mathbf{u}=\{u_0,\cdots,u_l\}\subset\mathbb{F}_q^{\mathcal{D}}$. Let $\mu:\mathcal{D}\to[0,1]$ be a vector of weights, whose values all have denominator M. Let $m\ge 3$ and let
>
> $$\alpha\ge\alpha_0(\rho,m):=\sqrt{\rho}+\frac{\rho}{2m}. \tag{51}$$
>
> Let
>
> $$S=\{z\in\mathbb{F}_q:\operatorname{agree}_{\mu}(u_0+zu_1+\cdots+z^lu_l, V)\ge\alpha\} \tag{52}$$
>
> and suppose
>
> $$|S|>\max\left(\frac{(1+\frac{1}{2m})^7m^7}{3\rho^{3/2}}n^2l, \frac{2m+1}{\sqrt{\rho}}(M\cdot n+1)l\right). \tag{7.1}$$
>
> Then $u_0,\ldots,u_l$ have at least α correlated μ-agreement with V, i.e. $\exists v_0,\cdots,v_l\in V$ such that
>
> $$\mu(\{x\in\mathcal{D}:\forall 0\le i\le l, u_i(x)=v_i(x)\})\ge\alpha. \tag{53}$$

In Theorem 7.2, take $M=|\mathcal{D}^{(0)}|/|\mathcal{D}^{(i+1)}|$. At this time, we are analyzing the case of $i+1$, so $n=|\mathcal{D}^{i+1}|$ in the theorem, then $M\cdot n=|\mathcal{D}^{(0)}|$. Since we are analyzing $\mathbf{u}=\{u_0,\cdots,u_{l^{(i)}-1}\}$, $l=l^{(i)}-1$ in equation (7.1). According to Theorem 7.2, if

$$\Pr_{z^{(i)}}\left[E^{(i+1)}\right]\ge(l^{(i)}-1)\cdot\left(\epsilon^{(i)}+\frac{2m+1}{\sqrt{\rho}}\cdot\frac{|\mathcal{D}^{(0)}|+1}{|\mathbb{F}|}\right) \tag{54}$$

where,

$$\epsilon^{(i)} = \frac{|\mathcal{D}^{(i+1)}|^2}{|\mathcal{D}^{(0)}|^2}\epsilon = \frac{\epsilon}{(l^{(0)}\cdots l^{(i)})^2} = \frac{(m+\frac{1}{2})^7}{3\rho^{3/2}} \cdot \frac{|\mathcal{D}^{(0)}|^2}{q} \cdot \frac{1}{(l^{(0)}\cdots l^{(i)})^2} \tag{55}$$

If the above condition is satisfied, referring to Theorem 7.2, we have

$$\begin{aligned}
|S| &\geq |\mathbb{F}| \cdot (l^{(i)}-1) \cdot \left(\epsilon^{(i)} + \frac{2m+1}{\sqrt{\rho}} \cdot \frac{|\mathcal{D}^{(0)}|+1}{\mathbb{F}}\right) \\
&= \frac{(m+\frac{1}{2})^7}{3\rho^{3/2}} \cdot \frac{|\mathcal{D}^{(0)}|^2}{q} \cdot \frac{|\mathcal{D}^{(i+1)}|^2}{|\mathcal{D}^{(0)}|^2} \cdot |\mathbb{F}| \cdot (l^{(i)}-1) + \frac{2m+1}{\sqrt{\rho}} \cdot (|\mathcal{D}^{(0)}|+1) \cdot (l^{(i)}-1) \\
&= \frac{(1+\frac{1}{2m})^7 m^7}{3\rho^{3/2}} \cdot |\mathcal{D}^{(i+1)}|^2 \cdot (l^{(i)}-1) + \frac{2m+1}{\sqrt{\rho}} \cdot (M\cdot n+1) \cdot (l^{(i)}-1) \\
&= \frac{(1+\frac{1}{2m})^7 m^7}{3\rho^{3/2}} \cdot n^2 \cdot (l^{(i)}-1) + \frac{2m+1}{\sqrt{\rho}} \cdot (M\cdot n+1) \cdot (l^{(i)}-1) \\
&> \max\left(\frac{(1+\frac{1}{2m})^7 m^7}{3\rho^{3/2}} n^2 l, \frac{2m+1}{\sqrt{\rho}}(M\cdot n+1)l\right)
\end{aligned} \tag{56}$$

Satisfying equation (7.1), therefore by Theorem 7.2, there exists a set $S \subseteq \mathcal{D}^{(i+1)}$, and codewords $v_0, \ldots, v_{l^{(i)}-1} \in V$, such that u_i and v_i are consistent on S, and $\nu^{(i+1)}(S) > \alpha$. Recalling equation (8.4), we know

$$\mathbf{u}^{(i)}\left(g^{(i+1)}\right) = M^{(i)}_{g^{(i+1)}} \cdot f^{(i)}|_{C^{(i)}_{g^{(i+1)}}} \tag{57}$$

The invertible interpolation mapping $M^{(i)}_{g^{(i+1)}}$ maps $f^{(i)}|_{C^{(i)}_{g^{(i+1)}}}$ to $\mathbf{u}^{(i)}\left(g^{(i+1)}\right)$. Using its inverse mapping, i.e., the evaluation mapping, for each $g^{(i+1)} \in \mathcal{D}^{(i+1)}$, apply this inverse mapping to $v_0(g^{(i+1)}), \ldots, v_{l^{(i)}-1}(g^{(i+1)})$. Let $C^{(i)}_{g^{(i+1)}} = \{g'_0, \cdots, g'_{l^{(i)}-1}\}$, then the result after application is

$$\begin{aligned}
\left(M^{(i)}_{g^{(i+1)}}\right)^{-1} \cdot \begin{bmatrix} v_0(g^{(i+1)}) \\ v_1(g^{(i+1)}) \\ \vdots \\ v_{l^{(i)}-1}(g^{(i+1)}) \end{bmatrix} &= V_{C^{(i)}_{g^{(i+1)}}} \cdot \begin{bmatrix} v_0(g^{(i+1)}) \\ v_1(g^{(i+1)}) \\ \vdots \\ v_{l^{(i)}-1}(g^{(i+1)}) \end{bmatrix} \\
&= \begin{bmatrix} 1 & g'_0 & (g'_0)^2 & \cdots & (g'_0)^{l^{(i)}-1} \\ 1 & g'_1 & (g'_1)^2 & \cdots & (g'_1)^{l^{(i)}-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & g'_{l^{(i)}-1} & (g'_{l^{(i)}-1})^2 & \cdots & (g'_{l^{(i)}-1})^{l^{(i)}-1} \end{bmatrix} \begin{bmatrix} v_0(g^{(i+1)}) \\ v_1(g^{(i+1)}) \\ \vdots \\ v_{l^{(i)}-1}(g^{(i+1)}) \end{bmatrix} \\
&= \begin{bmatrix} v_0(g^{(i+1)}) + v_1(g^{(i+1)})g'_0 + u_2(g'_0)^2 + \cdots + v_{l^{(i)}-1}(g^{(i+1)})(g'_0)^{l^{(i)}-1} \\ v_0(g^{(i+1)}) + v_1(g^{(i+1)})g'_1 + u_2(g'_1)^2 + \cdots + v_{l^{(i)}-1}(g^{(i+1)})(g'_1)^{l^{(i)}-1} \\ \vdots \\ v_0(g^{(i+1)}) + v_1(g^{(i+1)})g'_{l^{(i)}-1} + u_2(g'_{l^{(i)}-1})^2 + \cdots + v_{l^{(i)}-1}(g^{(i+1)})(g'_{l^{(i)}-1})^{l^{(i)}-1} \end{bmatrix} \\
&= \begin{bmatrix} h^{(i)}(g'_0) \\ h^{(i)}(g'_1) \\ \vdots \\ h^{(i)}(g'_{l^{(i)}-1}) \end{bmatrix}
\end{aligned} \tag{58}$$

We can get function $h^{(i)} : \mathcal{D}^{(i)} \to \mathbb{F}$, for each $g^{(i)} \in C^{(i)}_{g^{(i+1)}}$ we have

$$h^{(i)}(g^{(i)}) = \sum_{j=0}^{l^{(i)}-1} \left(g^{(i)}\right)^j \cdot v_j\left(g^{(i+1)}\right) = \sum_{j=0}^{l^{(i)}-1} \left(g^{(i)}\right)^j \cdot v_j\left(\left(g^{(i)}\right)^{l^{(i)}}\right). \tag{59}$$

Therefore, since $v_j \in V^{(i+1)}$, we have $h^{(i)} \in V^{(i)}$. Moreover, according to the definition

$$
\begin{aligned}
\operatorname{agree}_{\mu^{(i)}}\left(f^{(i)}, V^{(i)}\right) &= \max_{v \in V^{(i)}} \operatorname{agree}_{\mu^{(i)}}\left(f^{(i)}, v\right) \\
&\geq \operatorname{agree}_{\mu^{(i)}}\left(f^{(i)}, h^{(i)}\right) \\
&= \frac{1}{|\mathcal{D}^{(i)}|} \sum_{x: f^{(i)}(x)=h^{(i)}(x)} \mu^{(i)}(x) \\
&= \nu^{(i+1)}(S) \\
&> \alpha,
\end{aligned} \tag{60}
$$

This contradicts the definition of α: $\alpha = \max\left(\operatorname{agree}_{\mu^{(i)}}\left(f^{(i)}, V^{(i)}\right), \sqrt{\rho}(1+1/2m)\right)$. This means that the assumption we made when applying Theorem 7.2 does not hold, that is, the following equation holds:

$$
\Pr_{z^{(i)}}\left[E^{(i+1)}\right] < (l^{(i)}-1) \cdot \left(\epsilon^{(i)} + \frac{2m+1}{\sqrt{\rho}} \cdot \frac{|\mathcal{D}^{(0)}|+1}{|\mathbb{F}|}\right). \tag{61}
$$

Therefore, if event $E^{(i+1)}$ does not occur, according to equation (8.10), for all $i \in 0, 1, \ldots, r-1$ we have:

$$
\operatorname{agree}_{\mu^{(i+1)}}\left(f^{(i+1)}, V^{(i+1)}\right) \leq \max\left(\operatorname{agree}_{\mu^{(i)}}\left(f^{(i)}, V^{(i)}\right), \sqrt{\rho}(1+1/2m)\right) \tag{62}
$$

According to equation (8.8), we get

$$
\Pr_{x_1,\ldots,x_t}[E^{(0)}] \leq \epsilon, \quad \text{where} \epsilon = \frac{(m+\frac{1}{2})^7}{3\rho^{3/2}} \cdot \frac{|\mathcal{D}^{(0)}|^2}{q}. \tag{63}
$$

If the probability of event $E^{(0)}$ or some $E^{(i+1)}$ occurring is estimated as

$$
\begin{aligned}
\Pr_{x_1,\ldots,x_t}\left[E^{(0)}\right] + \sum_{i=0}^{r-1} \Pr_{z^{(i)}}\left[E^{(i+1)}\right] &\leq \epsilon + \sum_{i=0}^{r-1} \left((l^{(i)}-1) \cdot \left(\epsilon^{(i)} + \frac{2m+1}{\sqrt{\rho}} \cdot \frac{|\mathcal{D}^{(0)}|+1}{|\mathbb{F}|}\right)\right) \\
&= \epsilon + \sum_{i=0}^{r-1} (l^{(i)}-1) \cdot \epsilon^{(i)} + \sum_{i=0}^{r-1} \left((l^{(i)}-1) \cdot \frac{2m+1}{\sqrt{\rho}} \cdot \frac{|\mathcal{D}^{(0)}|+1}{|\mathbb{F}|}\right) \\
&= \left(1 + \sum_{i=0}^{r-1} \frac{l^{(i)}-1}{(l^{(0)} \cdots l^{(i)})^2}\right)\epsilon + \frac{2m+1}{\sqrt{\rho}} \cdot \frac{|\mathcal{D}^{(0)}|+1}{|\mathbb{F}|} \cdot \sum_{i=0}^{r-1}(l^{(i)}-1)
\end{aligned} \tag{64}
$$

Let's estimate $\sum_{i=0}^{r-1} \frac{l^{(i)}-1}{(l^{(0)} \cdots l^{(i)})^2}$. Since for $i \in \{0, \ldots, r-1\}$ we have $l^{(i)} \geq 2$, therefore

$$
\begin{aligned}
\sum_{i=0}^{r-1} \frac{l^{(i)}-1}{(l^{(0)} \cdots l^{(i)})^2} &= \sum_{i=0}^{r-1} \left(\frac{l^{(i)}}{(l^{(0)} \cdots l^{(i)})^2} - \frac{1}{(l^{(0)} \cdots l^{(i)})^2}\right) \\
&= \sum_{i=0}^{r-1} \left(\frac{1}{(l^{(0)} \cdots l^{(i-1)})^2 l^{(i)}} - \frac{1}{(l^{(0)} \cdots l^{(i)})^2}\right) \\
&= \frac{1}{l^{(0)}} + \left(-\frac{1}{(l^{(0)})^2} + \frac{1}{(l^{(0)})^2 l^{(1)}}\right) + \left(-\frac{1}{(l^{(0)}l^{(1)})^2} + \frac{1}{(l^{(0)}l^{(1)})^2 l^{(2)}}\right) \\
&\quad + \ldots + \left(-\frac{1}{(l^{(0)} \cdots l^{(r-2)})^2} + \frac{1}{(l^{(0)} \cdots l^{(r-2)})^2 l^{(r-1)}}\right) - \frac{1}{(l^{(0)} \cdots l^{(r-1)})^2} \\
&\leq \frac{1}{l^{(0)}} + \left(-\frac{1}{(l^{(0)})^2} + \frac{1}{(l^{(0)})^2 \cdot 2}\right) + \left(-\frac{1}{(l^{(0)}l^{(1)})^2} + \frac{1}{(l^{(0)}l^{(1)})^2 \cdot 2}\right) \\
&\quad + \ldots + \left(-\frac{1}{(l^{(0)} \cdots l^{(r-2)})^2} + \frac{1}{(l^{(0)} \cdots l^{(r-2)})^2 \cdot 2}\right) - \frac{1}{(l^{(0)} \cdots l^{(r-1)})^2} \\
&= \frac{1}{l^{(0)}} - \frac{1}{(l^{(0)})^2 \cdot 2} - \frac{1}{(l^{(0)}l^{(1)})^2 \cdot 2} - \cdots - \frac{1}{(l^{(0)} \cdots l^{(r-2)})^2 \cdot 2} \\
&\quad - \frac{1}{(l^{(0)} \cdots l^{(r-1)})^2} \\
&< \frac{1}{l^{(0)}} \\
&< \frac{1}{2}
\end{aligned} \tag{65}
$$

> 📖 **Another proof method: using geometric series summation**

$$
\begin{aligned}
\sum_{i=0}^{r-1} \frac{l^{(i)}-1}{(l^{(0)}\cdots l^{(i)})^2} &= \sum_{i=0}^{r-1}\left(\frac{l^{(i)}}{(l^{(0)}\cdots l^{(i)})^2} - \frac{1}{(l^{(0)}\cdots l^{(i)})^2}\right) \\
&= \sum_{i=0}^{r-1}\left(\frac{1}{(l^{(0)}\cdots l^{(i-1)})^2 l^{(i)}} - \frac{1}{(l^{(0)}\cdots l^{(i)})^2}\right) \\
&< \sum_{i=0}^{r-1} \frac{1}{(l^{(0)}\cdots l^{(i-1)})^2 l^{(i)}} \\
&\text{(Because } \frac{1}{(l^{(0)}\cdots l^{(i)})^2} > 0\text{)} \\
&< \sum_{i=0}^{r-1} \frac{1}{l^{(0)}\cdots l^{(i-1)} l^{(i)}} \\
&\text{(Because } l^{(i)} \geq 2\text{, so } {l^{(i)}}^2 > l^{(i)}\text{)} \\
&< \sum_{i=0}^{r-1} \left(\frac{1}{2}\right)^{i+1} \\
&= \frac{1}{2}\sum_{i=0}^{r-1} \left(\frac{1}{2}\right)^{i} \\
&< \frac{1}{2}\cdot\frac{1}{2} \\
&< \frac{1}{2}
\end{aligned} \tag{66}
$$

Question

- [ ] Is there a more concise way for the above proof?

Therefore

$$
\begin{aligned}
\Pr_{x_1,\ldots,x_t}\left[E^{(0)}\right] + \sum_{i=0}^{r-1}\Pr_{z^{(i)}}\left[E^{(i+1)}\right] &\leq \left(1 + \sum_{i=0}^{r-1}\frac{l^{(i)}-1}{(l^{(0)}\cdots l^{(i)})^2}\right)\epsilon + \frac{2m+1}{\sqrt{\rho}}\cdot\frac{|\mathcal{D}^{(0)}|+1}{|\mathbb{F}|}\cdot\sum_{i=0}^{r-1}(l^{(i)}-1) \\
&< \left(1+\frac{1}{2}\right)\epsilon + \frac{2m+1}{\sqrt{\rho}}\cdot\frac{|\mathcal{D}^{(0)}|+1}{|\mathbb{F}|}\cdot\sum_{i=0}^{r-1}(l^{(i)}-1) \\
&< \frac{3}{2}\epsilon + \frac{2m+1}{\sqrt{\rho}}\cdot\frac{|\mathcal{D}^{(0)}|+1}{|\mathbb{F}|}\cdot\sum_{i=0}^{r-1} l^{(i)}.
\end{aligned} \tag{67}
$$

In summary, we have obtained that when some bad events $E^{(i)}$ occur, their probability is strictly less than

$$
\frac{3}{2}\epsilon + \frac{2m+1}{\sqrt{\rho}}\cdot\frac{|\mathcal{D}^{(0)}|+1}{|\mathbb{F}|}\cdot\sum_{i=0}^{r-1} l^{(i)} = \epsilon_C, \tag{68}
$$

When no bad events occur, the following equation holds

$$
\begin{aligned}
\text{agree}_{\mu^{(r)}}\left(f^{(r)}, V^{(r)}\right) &= \mathbb{E}_{g^{(r)}\in\mathcal{D}^{(r)}}\left[\mu^{(r)}(g^{(r)})\right] \\
&\text{(Because } f^{(r)} \in V^{(r)}\text{)} \\
&\leq \max\left(\text{agree}_{\mu^{(r-1)}}\left(f^{(r-1)}, V^{(r-1)}\right), \sqrt{\rho}(1+1/2m)\right) \\
&\text{(According to the definition of event } E^{(i+1)}\text{, see equation (8.9))} \\
&\leq \max\left(\text{agree}_{\mu^{(r-2)}}\left(f^{(r-2)}, V^{(r-2)}\right), \sqrt{\rho}(1+1/2m)\right) \\
&\leq \ldots \\
&\leq \max\left(\text{agree}_{\mu^{(0)}}\left(f^{(0)}, V^{(0)}\right), \sqrt{\rho}(1+1/2m)\right) \\
&= \max\left(\alpha, \sqrt{\rho}(1+1/2m)\right) \\
&= \alpha^{(0)}(\rho, m)
\end{aligned} \tag{69}
$$

At this point, we have proved that equation (8.7) holds, thus proving Lemma 8.2. □

References

- [BBHR18] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. "Fast Reed–Solomon Interactive Oracle Proofs of Proximity". In: *Proceedings of the 45th International Colloquium on Automata, Languages and Programming (ICALP)*, 2018.
- [BCIKS20] Eli Ben-Sasson, Dan Carmon, Yuval Ishai, Swastik Kopparty, and Shubhangi Saraf. Proximity Gaps for Reed–Solomon Codes. In *Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science*, pages 900–909, 2020.
- [RVW13] Guy N. Rothblum, Salil Vadhan, and Avi Wigderson. Interactive proofs of proximity: delegating computation in sublinear time. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 793–802. ACM, 2013.