[BBHR18] FRI 论文 soundness 解析

作者： Yu Guo(yu.guo@secbit.io) Jade Xie(jade@secbit.io)

本篇文章主要讲解 Eli Ben-Sasson 等人在 2018 年发表的论文 [BBHR18b]，重点放在对 FRI 协议的 completeness 和 soundness 证明上。他们在这篇论文中针对 Reed-Solomon (RS) 编码提出了一种新的 IOPP (Interactive Oracle Proof of Proximity, IOPP)，称之为 FRI (Fast RS IOPP, FRI)。随后，在 [BBHR18a] 中使用 FRI 协议构建了一个实用的 ZK 系统，即我们熟知的 STARK。

首要问题

对于在有限域 $\mathbb{F}$ 中的求值 (evaluation) 集合 S，假设 S 中的元素个数为 N，给定一个码率参数 $\rho \in (0,1]$，编码 $\mathrm{RS}[\mathbb{F},S,\rho]$ 表示的是所有函数 $f:S\to\mathbb{F}$ 的集合，其中 f 是次数 $d<\rho N$ 的多项式的求值 (evaluations)，即存在次数 $d<\rho N$ 的多项式 $\hat{f}$ 使得 f 与 $\hat{f}$ 在 S 上的值是一致的。

论文主要关注的就是 *RS proximity problem*：假设我们能获得关于函数 $f:S\to\mathbb{F}$ 的 oracle，需要 Verifier 用较少的查询复杂度，同时有很高的把握能辨别出 f 属于下面哪一种情况：

1. $f\in\mathrm{RS}[\mathbb{F},S,\rho]$
2. $\Delta(f,\mathrm{RS}[\mathbb{F},S,\rho])>\delta$

也就是要么 f 是 RS 编码 $\mathrm{RS}[\mathbb{F},S,\rho]$ 中的一个码字，要么距离所有 $\mathrm{RS}[\mathbb{F},S,\rho]$ 中的码字的相对 Hamming 距离都大于接近参数 δ。一个自然的想法是 verifier 可以轮讯 $d+1$ 次，然后判断 f 属于上述哪一种情况，如果属于第一种，则接受，如果属于第二种，则拒绝。此时的轮讯复杂度为 $d+1=\rho N$。在计算 Testing 方法的复杂度时，没有额外的信息提供给 verifier，那么说 prover 尝试让 verifier 相信 $f\in\mathrm{RS}[\mathbb{F},S,\rho]$ 所消耗的计算复杂度为 0，交互的轮数为 0，以及产生的证明长度为 0。对比此方法 (Testing, [RS92]) 与FRI 的复杂度，如下表所示([BBHR18b])。

	prover 计算复杂度	证明长度	verifier计算复杂度	查询复杂度	轮询复杂度
Testing [RS92]	0	0	$\rho N\cdot\log^{O(1)}$	ρN	0
FRI [BBHR18b]	$<6\cdot N$	$<\frac{N}{3}$	$\le 21\cdot\log N$	$2\log N$	$\frac{\log N}{2}$

可以看出，FRI 中 prover 的计算复杂度是严格线性的并且 verfier 的计算复杂度是严格对数的，而查询复杂度是对数级别的([BBHR18b])。

FRI 性质

上文提到 FRI 是一种 IOPP，下面给出 IOPP 的定义。

Definition 1 [BBHR18b, Definition 1.1] (Interactive Oracle Proof of Proximity (IOPP)). An $\mathbf{r}$-round Interactive Oracle Proof of Proximity (IOPP) $\mathbf{S}=(\mathrm{P},\mathrm{V})$ is a $(r+1)$-round IOP. We say $\mathbf{S}$ is an ($\mathbf{r}$-round) IOPP for the error correcting code $C=\{f:S\to\Sigma\}$ with soundness $\mathbf{s}^-:(0,1]\to[0,1]$ with respect to distance measure Δ, if the following conditions hold:

- **First message format:** the first prover message, denote $f^{(0)}$, is a purported codeword of C, i.e., $f^{(0)}:S\to\Sigma$
- **Completeness:** $\Pr[\langle \mathrm{P}\leftrightarrow\mathrm{V}\rangle=\mathrm{accept}|\Delta(f^{(0)},C)=0]=1$
- **Soundness:** For any P^*, $\Pr[\langle \mathrm{P}^*\leftrightarrow\mathrm{V}\rangle=\mathrm{reject}|\Delta(f^{(0)},C)=\delta]\ge\mathbf{s}^-(\delta)$

意思是 Prover 和 Verifier 会进行 $\mathbf{r}$ -轮的交互，需要满足三个条件。

1. 第一个消息 $f^{(0)}$ 是 Prover 初始声称的在 C 中的码字。
2. 完备性：说的是对于诚实的 Prover，如果 $f^{(0)}$ 在 C 中，那么 Verifier 一定会输出 accept。
3. Soundness：分析的是作恶的 Prover，经过交互之后 Verifier 拒绝的概率是多少。定义中的 soundness $\mathbf{s}^-:(0,1]\to[0,1]$ 是一个函数，自变量 $\delta\in(0,1]$，这也表示在分析 soundness 时，我们考虑的是作恶的 Prover，也就是初始的 $\Delta(f^{(0)},C)=\delta>0$，在这种情况下 Prover 和 Verifier 进行交互，来算拒绝的概率是多少，这个概率的下界就是 $\mathbf{s}^-(\delta)\in[0,1]$，由于这里表示的是概率，自然 $\mathbf{s}^-(\delta)$ 函数值在闭区间 $[0,1]$ 中。

FRI 协议

下面摘录下论文 [BBHR18b] 中对 FRI 协议的描述。

定义和记号

Interpolant For a function $f:S\to\mathbb{F}$, $S\subset\mathbb{F}$, let $\mathrm{interpolant}^f$ denote the *interpolant* of f, defined as the unique polynomial $P(X)=\sum_{i=0}^{|S|-1}a_iX^i$ of degree less than $|S|$ whose evaluation on S equals $f|_S$, i.e., $\forall x\in S, f(x)=P(x)$. We assume the interpolant $P(X)$ is represented as a formal sum, i.e., by the sequence of monomial coefficients $a_0,\cdots,a_{|S|-1}$.

Subspace polynomials Given a set $L_0\subset\mathbb{F}$, let $\mathrm{Zero}_{L_0}\triangleq\prod_{x\in L_0}(X-x)$ be the unique non-zero monic polynomial of degree $|L_0|$ that vanishes on L_0. When L_0 is an additive coset contained in a binary field, the polynomial $\mathrm{Zero}_{L_0}(X)$ is an *affine subspace polynomial*, a special type of a linearized polynomial. We shall use the following properties of such polynomials, referring the interested reader to [LN97, Chapter 3.4] for proofs and additional background:

1. The map $x\mapsto\mathrm{Zero}_{L_0}(x)$ maps each additive coset S of L_0 to a single field element, which will be denoted by y_S.
2. If $L\supset L_0$ are additive cosets, then $\mathrm{Zero}_{L_0}(L)\triangleq\{\mathrm{Zero}_{L_0}(z)|z\in L\}$ is an additive coset and $\dim(\mathrm{Zero}_{L_0}(L))=\dim(L)-\dim(L_0)$.

Subspace specification Henceforth, the letter L always denotes an additive coset in a binary field $\mathbb{F}$, we assume all mentioned additive cosets are specified by an additive shift $\alpha\in\mathbb{F}$ and a basis $\beta_1,\cdots,\beta_k\in\mathbb{F}^k$ so that $L=\left\{\alpha+\sum_{i=1}^k b_i\beta_i|b_1,\cdots,b_k\in\mathbb{F}_2\right\}$; we assume α and $\vec{\beta}=(\beta_1,\cdots,\beta_k)$ are agreed upon by prover and verifier.

COMMIT 阶段

协议的轮数为 $r \triangleq \left\lfloor \frac{k^{(0)}-\mathcal{R}}{\eta} \right\rfloor$，其中 $\mathcal{R} = \log(1/\rho)$，$\rho$ 表示码率。在 COMMIT 阶段的第 i 轮，$i \in \{0, \cdots, r-1\}$，Verifier 可以访问一个由 Prover 提交的函数 $f^{(i)} : L^{(i)} \to \mathbb{F}$ 的 oracle，其中 $\dim(L^{(i)}) = k^{(i)} = k^{(0)} - \eta \cdot i$，并且空间 $L^{(i)}$ 是预先固定的，特别地，它们不依赖于 Verifier 的消息。

> **FRI-COMMIT:** Common input:
>
> - Parameters $\mathcal{R}, \eta, i$, all are positive integers: – rate parameter $\mathcal{R}$: logarithm of RS code rate ($\rho = 2^{-\mathcal{R}}$) – localization parameter η : dimension of $L_0^{(i)}$ (i.e., $|L_0^{(i)}| = 2^\eta$); let $r \triangleq \left\lfloor \frac{k^{(0)}-\mathcal{R}}{\eta} \right\rfloor$ denote round complexity – $i \in \{0, \cdots, r\}$: round counter
> - A parametrization of $\mathrm{RS}^{(i)} \triangleq \mathrm{RS}[\mathbb{F}, L^{(i)}, \rho = 2^{-\mathcal{R}}]$, denote $k^{(i)} = \log_2 |L^{(i)}|$ (notice $k^{(i)} = \dim(L^{(i)})$);
> - $L_0^{(i)} \subset L^{(i)}$, $\dim(L_0^{(i)}) = \eta$; let $q^{(i)}(X) = \mathrm{Zero}_{L_0^{(i)}}(X)$ and denote $L^{(i+1)} = q^{(i)}(L^{(i)})$
>
> Prover input: $f^{(i)} : L^{(i)} \to \mathbb{F}$, a purported codeword of $\mathrm{RS}^{(i)}$
>
> Loop: While $i \leq r$:
>
> 1. Verifier sends a uniformly random $x^{(i)} \in \mathbb{F}$
> 2. Prover defines the function $f^{(i+1)}_{f^{(i)},x^{(i)}}$ with domain $L^{(i+1)}$ thus, for each $y \in L^{(i+1)}$:
>   - Let $S_y = \{x \in L^{(i)} | q^{(i)}(x) = y\}$ be the coset of $L_0^{(i)}$ mapped by $q^{(i)}$ to $\{y\}$;
>   - $P_y^{(i)}(X) \triangleq \mathrm{interpolant}^{f^{(i)}|_{S_y}}$;
>   - $f^{(i+1)}_{f^{(i)},x^{(i)}}(y) \triangleq P_y^{(i)}(x^{(i)})$;
> 3. If $i = r$ then:
>   - let $f^{(r)} = f^{(r)}_{f^{(r-1)},x^{(r-1)}}$ for $f^{(r)} = f^{(r)}_{f^{(r-1)},x^{(r-1)}}$ defined in step 2 above;
>   - let $P^{(r)}(X) = \sum_{j \geq 0} a_j^{(r)} X^j \triangleq \mathrm{interpolant}^{f^{(r)}}(X)$;
>   - let $d = \rho \cdot |L^{(r)}| - 1$;
>   - prover commits to first $d+1$ coefficients of $P^{(r)}(X)$, namely, to $\langle a_0^{(r)}, \cdots, a_d^{(r)} \rangle$
>   - COMMIT phase terminates;
> 4. Else ($i < r$):
>   - let $f^{(i+1)} = f^{(i+1)}_{f^{(i)},x^{(i)}}$ for $f^{(i+1)}_{f^{(i)},x^{(i)}}$ defined in step 2 above;
>   - prover commits to oracle $f^{(i+1)}$
>   - both parties repeat the COMMIT protocol with common input
>   - parameters $(\mathcal{R}, \eta, i+1)$
>   - a parametrization of $\mathrm{RS}^{(i+1)} \triangleq \mathrm{RS}[\mathbb{F}, L^{(i+1)}, \rho = 2^{-\mathcal{R}}]$ and $L_0^{(i+1)} \subset L^{(i+1)}$, $\dim(L_0^{(i+1)}) = \eta$ and prover input $f^{(i+1)}$ defined at the beginning of this step;

QUERY 阶段

> **FRI-QUERY:** verifier input:
>
> - parameters $\mathcal{R}, \eta$ as defined in the COMMIT phase
> - repetition parameter l
> - sequence of rate-ρ RS-codes $\mathrm{RS}^{(0)}, \cdots, \mathrm{RS}^{(r)}$, where $\mathrm{RS}^{(i)} \triangleq \mathrm{RS}[\mathbb{F}, L^{(i)}, \rho]$ and $\log_2 |L^{(i)}| = k^{(i)} = k^{(0)} - \eta$; (notice $k^{(i)} = \dim(L^{(i)})$);
> - sequence of affine spaces $L_0^{(0)}, \cdots, L_0^{(r-1)}$, each $L_0^{(i)}$ is of dimension η and contained in $L^{(i)}$;
> - transcript of verifier messages $x^{(0)}, \cdots, x^{(r-1)} \in \mathbb{F}$
> - access to oracles $f^{(0)}, \cdots, f^{(r-1)}$
> - access to last oracle $P^{(r)}(X) = \sum_{j \geq 0} a_j^{(r)} X^j$ for $d = \rho \cdot |L^{(r)}| - 1$;
>
> Terminal function reconstruction:
>
> - query $a_0^{(r)}, \cdots, a_d^{(r)}$;(a total of $d+1 \leq 2^\eta$ queries)
> - let $P'(X) \triangleq \sum_{j \geq 0} a_j^{(r)} X^j$;
> - let $f^{(r)}$ be the evaluation of $P'(X)$ on $L^{(r)}$; (notice $f^{(r)} \in \mathrm{RS}^{(r)}$)
>
> Repeat l times: {
>
> 1. Sample uniformly random $s^{(0)} \in L^{(0)}$ and for $i = 0, \cdots, r-1$ let
>   - $s^{i+1} = q^{(i)}(s^{(i)})$
>   - $S^{(i)}$ be the coset of $L_0^{(i)}$ in $L^{(i)}$ that contains $s^{(i)}$
> 2. For $i = 0, \cdots, r-1$,
>   - query $f^{(i)}$ on all of $S^{(i)}$; (a total of 2η queries)
>   - compute $P^{(i)}(X) \triangleq \mathrm{interpolant}^{f^{(i)}|_{S^{(i)}}}$; (notice $\deg(P^{(i)}) < 2^\eta$)
> 3. **round consistency**: If for some $i \in \{0, \cdots, r-1\}$ it holds that
>
> $$f^{(i+1)}(s^{(i+1)}) \neq P^{(i)}(x^{(i)}) \tag{1}$$

> then reject and abort;
>
> }
>
> Return accept

FRI 协议的主要性质

下面的定理给出了 FRI 协议的主要性质，包括完备性(Completeness)、Soundness、Prover 复杂度以及 Verifier 复杂度。其实论文中还给出了一个稍微简略的版本，见论文 [BBHR18b] Theorem 1.3，该定理可以通过在下述定理中设置 $\eta = 2$ 与 $l = 1$ 证明得到的，这里就主要阐述这个更加复杂的版本。

Theorem 1 [BBHR18b, Theorem3.3] (Main properties of the **FRI** protocol). The following properties hold when the **FRI** protocol is invoked on oracle $f^{(0)} : L^{(0)} \to \mathbb{F}$ with localization parameter η and rate parameter $\mathcal{R}$ (and rate $\rho = 2^{-\mathcal{R}}$) such that $\rho|L^{(0)}| > 16$:

1. **Completeness** If $f^{(0)} \in \mathrm{RS}^{(0)} \triangleq \mathrm{RS}[\mathbb{F}, L^{(0)}, \rho = 2^{-\mathcal{R}}]$ and $f^{(1)}, \cdots, f^{(r)}$ are computed by the prover specified in the COMMIT phase, then the **FRI** verifier outputs **accept** with probability 1.
2. **Soundness** Suppose $\delta^{(0)} \triangleq \Delta^{(0)}(f^0, \mathrm{RS}^{(0)}) > 0$. Then with probability at least

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

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

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

Consequently, the soundness of **FRI** is at least

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

3. **Prover complexity** The i^{th} step of commit phase can be computed by a parallel random access machine (PRAM) with concurrent read and exclusive write (CREW) in $2\eta + 3$ cycles — each cycle involves a single arithmetic operation in $\mathbb{F}$ — using $2|L^{(i)}| + \eta$ processors and a total of $4|L^{(i)}|$ arithmetic operations over $\mathbb{F}$. Consequently, the total prover complexity is at most $6|L^{(0)}|$ arithmetic operations, which can be carried out in at most $4|L^{(0)}|$ cycles on a PRAM-CREW with $2n + 3$ processors.
4. **Verifier complexity** Verifier communication during the COMMIT phase equals $\mathbf{r}$ field elements; query complexity (during QUERY phase) equals $l2^{\eta}\mathbf{r} = l2^{\eta}\left(1 + \left\lfloor \frac{\log|L^{(0)}| - \mathcal{R}}{\eta} \right\rfloor\right)$. On a PRAM with exclusive read and exclusive write (EREW) with $l\mathbf{r} \cdot 2\eta$ processors, the verifier's decision is obtained after $2\eta + 3 + \log l$ cycles and a total of $l \cdot \mathbf{r} \cdot (6 \cdot 2\eta + 6\eta + 6)$ arithmetic operations in $\mathbb{F}$.

在第 2 项，Soundness 结论中，先给了一个参数 $\delta^{(0)} \triangleq \Delta^{(0)}(f^0, \mathrm{RS}^{(0)}) > 0$，这里的 $\Delta^{(0)}(f^0, \mathrm{RS}^{(0)})$ 其实并不是常见的相对 Hamming 距离，下面给出此测度的定义，同时说明它与相对 Hamming 距离之间的关系。

Block-wise 距离测度

Definition 2 [BBHR18b, Definition3.2] (Block-wise distance measure). Let $\mathcal{S} = \{S_1, \cdots, S_m\}$ be a partition of a set S and Σ be an alphabet. The relative $\mathcal{S}$-Hamming distance measure on Σ^S is defined for $f, g \in \Sigma^S$ as the relative Hamming distance over $\Sigma^{S_1} \times \cdots \times \Sigma^{S_m}$,

$$\Delta^{\mathcal{S}}(f, g) \triangleq \Pr_{i \in [m]}[f|_{S_i} \neq g|_{S_i}] = \frac{|\{i \in [m] | f|_{S_i} \neq g|_{S_i}\}|}{m}. \tag{5}$$

Thus, for $\mathcal{F} \subset \Sigma^S$ let $\Delta^{\mathcal{S}}(g, \mathcal{F}) = \min\{\Delta^{\mathcal{S}}(g, f) | f \in \mathcal{F}\}$.

为了更好的理解这个定义，在 FRI 协议中，考虑在 $\mathbb{F}^{L^{(i)}}$ 上的 block-wise 距离，即用 FRI 协议中在第 i 步的 $L^{(i)}$ 来替代上述定义中的集合 S，用 $\mathbb{F}$ 替换上述定义中的字母表 Σ。在第 i 步，我们能够确定集合 $L_0^{(i)}$ 。$L_0^{(i)}$ 其实可以设为映射 $q^{(i)}$ 的核，也就是在 $L^{(i)}$ 集合中那些被 q^i 映射为 $L^{(i+1)}$ 中单位元 e 的元素的集合，用数学符号表示出来即

$$L_0^{(i)} = \{x \in L^{(i)} | q^{(i)}(x) = e\}. \tag{6}$$

那么通过 $L_0^{(i)}$ 的陪集可以对集合 $L^{(i)}$ 进行划分，假设划分成 m 个集合，则对 $L^{(i)}$ 的划分可记为 $\mathcal{S}^{(i)} = \{L_0^{(i)}, \cdots, L_{m-1}^{(i)}\}$ 。那么简记

$$\Delta^{(i)}(f, g) \triangleq \Delta^{\mathcal{S}^{(i)}}(f, g) \tag{7}$$

对于两个函数 $f, g : L^{(i)} \to \mathbb{F}$ ，定义域均为 $L^{(i)}$ ，值域均为 $\mathbb{F}$ ，现在这个 Block-wise 距离说的是这两个函数在 $\mathcal{S}^{(i)}$ 中这些陪集中不完全一致的陪集个数的比值。例如在 $\mathcal{S}^{(i)} = \{L_0^{(i)}, \cdots, L_{m-1}^{(i)}\}$ 中 (假设 $m \geq 2$)，只有在 $L_0^{(i)}$ 与 $L_1^{(i)}$ 这两个集合上函数 f 与 g 对应的函数值不完全相同，即 $f|_{L_0^{(i)}} \neq g|_{L_0^{(i)}}$ 且 $f|_{L_1^{(i)}} \neq g|_{L_1^{(i)}}$ ，在其余的陪集上函数 f 与 g 完全一致，那么可以计算出 $\Delta^{(i)}(f, g) = \frac{2}{m}$ 。

上面的 $\Delta^{(i)}(f, g)$ 说的是 $\mathbb{F}^{L^{(i)}}$ 中两个元素的测度，下面解释下定义中关于集合中一个元素 $f^{(i)} \in \mathbb{F}^{L^{(i)}}$ 与一个子集 $\mathrm{RS}^{(i)} \subset \mathbb{F}^{L^{(i)}}$ ($\mathrm{RS}^{(i)} = RS[\mathbb{F}, L^{(i)}, \rho]$ 自然是 $\mathbb{F}^{L^{(i)}}$ 的子集)对应的 block-wise 距离测度，表示成

$$\Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)}) \triangleq \Delta^{\mathcal{S}^{(i)}}(f^{(i)}, \mathrm{RS}^{(i)}) = \min\{\Delta^{\mathcal{S}^{(i)}}(f^{(i)}, g^{(i)}) | g^{(i)} \in \mathrm{RS}^{(i)}\}, \tag{8}$$

其含义是取遍集合 $\mathrm{RS}^{(i)}$ 中所有的码字 $g^{(i)}$ ，算出这些 $\Delta^{\mathcal{S}^{(i)}}(f^{(i)}, g^{(i)})$ ，其中最小的那个值就是 $\Delta^{\mathcal{S}^{(i)}}(f^{(i)}, \mathrm{RS}^{(i)})$ 。关于该 Block-wise 距离测度，一个重要的不等式是

$$1 - \rho \geq \Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)}) \geq \Delta_H(f^{(i)}, \mathrm{RS}^{(i)}) \tag{4}$$

该等式会在 FRI 的 Soundness 证明中反复用到，比较重要，这里给出其证明。

证明：先证明不等式的左半边，即 $1-\rho \geq \Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)})$ 。总是存在这样一个多项式 $g^{(i)} \in \mathrm{RS}^{(i)}$ ，其次数 $deg(g^{(i)}) < \rho|L^{(i)}|$ ，同时 $\Delta^{(i)}(f^{(i)}, g^{(i)}) = 1-\rho$ 。下面说明 $g^{(i)}$ 的存在性。我们进行如下的构造： 在划分集合 $\mathcal{S}^{(i)} = \{L_0^{(i)}, \cdots, L_{m-1}^{(i)}\}$ 中，按顺序可得到集合序列 $\{L_0^{(i)}, \cdots, L_{m-1}^{(i)}\} = \{x_0, x_1, \cdots, x_{|L^{(i)}|-1}\}$ ，连续选择前 $\rho|L^{(i)}|$ 个点 $\{x_0, x_1, \cdots, x_{\rho|L^{(i)}|-1}\}$，得到这些点对应的 $f^{(i)}$ 的值 $\{f^{(i)}(x_0), f^{(i)}(x_1), \cdots, f^{(i)}(x_{\rho|L^{(i)}|-1})\}$ ，拿到这些点值对可以进行 Lagrange 插值，得到一个次数 $< \rho|L^{(i)}|$ 的多项式 $g^{(i)}$ ，同时易得这样构造的 $g^{(i)} \in \mathrm{RS}^{(i)} = RS[\mathbb{F}, L^{(i)}, \rho]$ 。同时根据前面的构造发现在集合 $\{L_0^{(i)}, \cdots, L_{\rho m-1}^{(i)}\} = \{x_0, x_1, \cdots, x_{\rho|L^{(i)}|-1}\}$ 上函数 $f^{(i)}$ 与 $g^{(i)}$ 的函数值是完全相等的 (这里 $\rho|L^{(i)}|$ 个点刚好完全占满在 ρm 个集合中，不会出现最后一些点只占最后一个集合的一部分的情况，这是由于选取 ρ 、$|L^{(i)}|$ 都是 2 的幂次形式，能够整除)，那么可计算出

$$\Delta^{(i)}(f^{(i)}, g^{(i)}) = \frac{|\{j \in [m] | f^{(i)}|_{L_j^{(i)}} \neq g^{(i)}|_{L_j^{(i)}}\}|}{m} = 1-\rho. \tag{9}$$

因此 $\Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)})$ 计算的 $\mathrm{RS}^{(i)}$ 中元素与 $f^{(i)}$ 在测度 $\Delta^{(i)}$ 下的最小值，那肯定不会超过找到的 $g^{(i)} \in \mathrm{RS}^{(i)}$ 的距离，也就证明了不等式的左半边 $1-\rho \geq \Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)})$ 。接下来证明不等式的右半边 $\Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)}) \geq \Delta_H(f^{(i)}, \mathrm{RS}^{(i)})$ 。假设 $\Delta^{(i)}(f^{(i)}, g^{(i)} \in \mathrm{RS}^{(i)}) = \delta$，不失一般性，假设 $f^{(i)}$ 与 $g^{(i)}$ 在陪集 $\{L_0^{(i)}, \cdots, L_{\delta m-1}^{(i)}\} = \{x_0, \cdots, x_{\delta|L^{(i)}|-1}\}$ 上不完全一致，在剩余的集合 $\{L_0^{(i)}, \cdots, L_{m-1}^{(i)}\} \backslash \{L_0^{(i)}, \cdots, L_{\delta m-1}^{(i)}\}$ 上是完全一致的。那么考虑在 $L^{(i)}$ 上的所有点时，$g^{(i)}$ 最多在 $\{L_0^{(i)}, \cdots, L_{\delta m-1}^{(i)}\} = \{x_0, \cdots, x_{\delta|L^{(i)}|-1}\}$ 这 $\delta|L^{(i)}|$ 点上都与 $f^{(i)}$ 不一致，因此也就说明了 $\Delta_H(f^{(i)}, g^{(i)}) \leq \delta$ ，进而如果设 $\Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)}) = \delta^*$ 可得出 $\Delta_H(f^{(i)}, \mathrm{RS}^{(i)}) \leq \delta^* = \Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)})$ 。 □

定理 1 完备性证明

下面说明定理 1 中完备性证明的思路，复述下完备性：

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

完备性说的是对于诚实的 Prover ，初始的函数 $f^{(0)}$ 是在 $\mathrm{RS}^{(0)}$ 编码空间中的，那么通过 FRI 的 COMMIT 阶段会产生一些列的函数 $f^{(1)}, \cdots, f^{(r)}$ ，那么 Verifier 在 QUERY 阶段结束后肯定会输出 **accept** 。

首先给出了一个递归的引理，再用该引理来证明完备性，引理表述的是在第 i 步如果 $f^{(i)} \in \mathrm{RS}^{(i)}$ ，那么在 COMMIT 阶段，Verifier 会从 $\mathbb{F}$ 中随机选取 $x^{(i)}$ 发给 Prover ，Prover 用该随机数来构造下一步的函数 $f_{f^{(i)}, x^{(i)}}^{(i+1)}$ ，那么对于 $\mathbb{F}$ 中任意一个 $x^{(i)}$ ，都有构造出来的 $f_{f^{(i)}, x^{(i)}}^{(i+1)}$ 都在 $\mathrm{RS}^{(i+1)}$ 空间中。递归的引理正式表述如下，关于该引理的证明留在后面进行说明。

Lemma 1 [BBHR18b, Lemma 4.1] (Inductive argument). If $f^{(i)} \in \mathrm{RS}^{(i)}$ then for all $x^{(i)} \in \mathbb{F}$ it holds that $f_{f^{(i)}, x^{(i)}}^{(i+1)} \in \mathrm{RS}^{(i+1)}$.

完备性证明的思路是，在 QUERY 阶段，Verifier 主要是在检查第 3 步的 round consistency 是否成立，一旦某一步的 $i \in \{0, \cdots, r-1\}$ 不成立就会直接输出 reject，直到对所有的 i 的检测都通过，最终才会输出 accept 。那么对于 $i < r-1$ ，根据 COMMIT 阶段 $f^{(i+1)}$ 的构造过程，round consistency 都会通过。对于 $i = r-1$，根据根据完备性的初始条件 $f^{(0)} \in \mathrm{RS}^{(0)}$ ，该定理递归的说明了 $f^{(r)} \in RS^{(r)}$ ，最后根据该结论说明在 QUERY 阶段也会检测通过 round consistency，最终 Verifier 也就一定会输出 accept 了。具体的完备性证明如下。

定理 1 第一项完备性证明：对于诚实的 Prover，对于任意的一个函数 $f^{(i)}$ ，在 COMMIT 阶段的第 2 步中，对于任意的 $i < r-1$ ，构造出

$$f_{f^{(i)}, x^{(i)}}^{(i+1)}(y) \triangleq P_y^{(i)}(x^{(i)}). \tag{10}$$

根据该构造，那么一定能在 QUERY 阶段的第 3 步一定会通过 round consistency ，即

$$f^{(i+1)}(s^{(i+1)}) = P^{(i)}(x^{(i)}) \tag{11}$$

成立。

下面只需证明对于 $i = r-1$ 时，round consistency 也能通过。根据完备性的假设知 $f^{(0)} \in \mathrm{RS}^{(0)}$ ，由 Lemma 1 递归可得 $f^{(r)} \in RS^{(r)}$ ，那么一定存在一个次数 $< \rho|L^{(r)}|$ 的多项式 $P^{(r)}(X)$ 使得 $f^{(r)}(X)$ 与 $P^{(r)}(X)$ 在 $L^{(r)}$ 上是完全一致的。因此 Prover 会在 COMMIT 阶段的第 3 步发送 $P^{(r)}(X)$ 的 $d+1 = \rho|L^{(r)}|$ 个系数 $\langle a_0^{(r)}, \cdots, a_d^{(r)} \rangle$ ，Verifier 在 QUERY 阶段的 "Terminal function reconstruction" 阶段会根据发送过来的 $d+1$ 个系数构造出 $P'(X) \triangleq \sum_{j \leq d} a_j^{(r)} X^j$ ，再根据 $P'(X)$ 得到函数 $f'^{(r)}$ ，函数 $f'^{(r)}$ 是 $P'(X)$ 在 $L^{(r)}$ 上的估计 (evaluation) 。那么可以推断出 $f'^{(r)}|_{L^{(r)}} = P'(X) = P^{(r)}(X) = f^{(r)}|_{L^{(r)}}$ 。自然会通过第 $i = r-1$ 轮的 round consistency ，即

$$f'^{(r)}(s^{(i+1)}) = P^{(r-1)}(x^i) \tag{12}$$

从而得证 Verifier 最后一定会输出 accept 。 □

命题 1 的引入

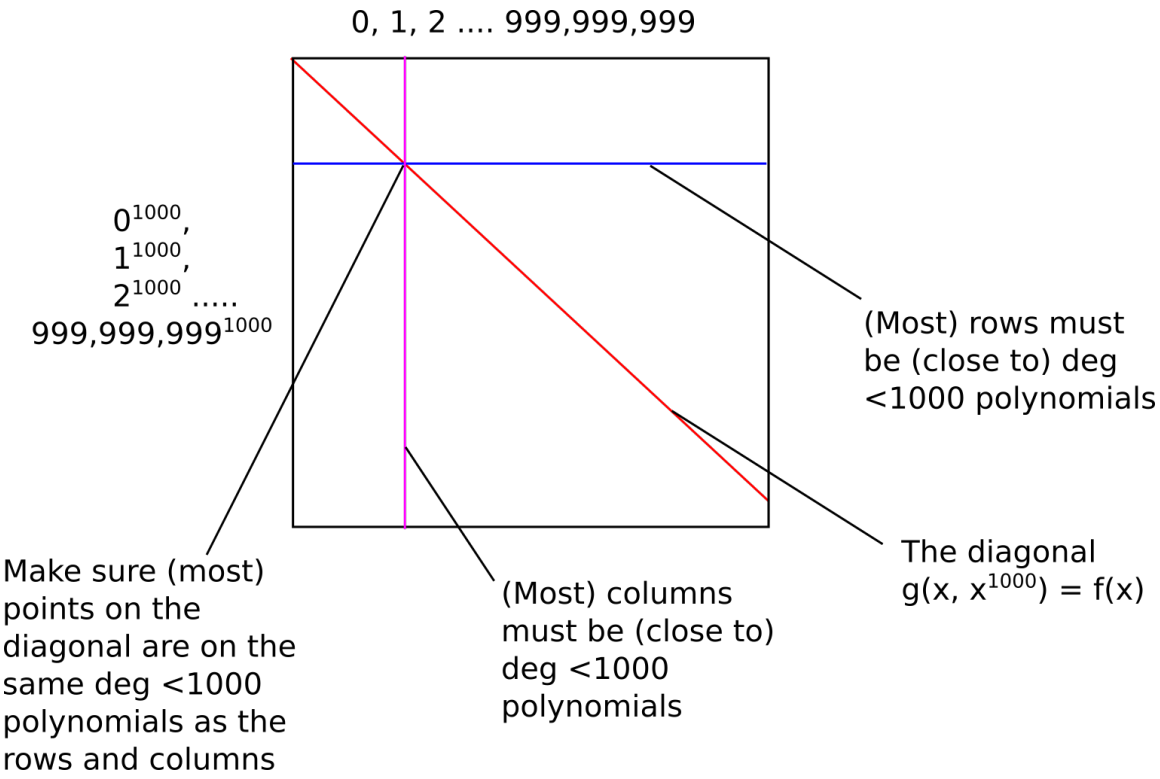
在证明引理 1 前先给出一个重要的命题，再用该命题来证明引理 1 。在下述命题中，用小写字母 x, y 来表示域中的元素，用大写字母 X, Y 来表示自变量。

Claim 1 [BBHR18b, Claim 4.2]. For every $f^{(i)} : L^{(i)} \to \mathbb{F}$ there exists $Q^{(i)}(X, Y) \in \mathbb{F}[X, Y]$ satisfying

1. $f^{(i)}(x) = Q^{(i)}(x, q^{(i)}(x))$ for all $x \in L^{(i)}$
2. $\deg_X(Q^{(i)}) < |L_0^{(i)}|$
3. If $f^{(i)} \in RS[\mathbb{F}, L^{(i)}, \rho]$ then $\deg_Y(Q^{(i)}) < \rho|L^{(i+1)}|$

该命题对于理解 FRI 协议是比较重要的。Vitalik 在其博客文章 STARKs, Part II: Thank Goodness It's FRI-day 的 A First Look at Sublinearity 小节中给出了一个具体的例子，其协议过程已初具 FRI 协议的雏形，我们在这里用命题 1 的视角来重新看看这个例子。 假设有限域 L 的大小为 $N = 10^9$， 设多项式 $f(X) : L \to \mathbb{F}$ ，且其次数 $< 10^6$ ，那么有 $f \in RS[\mathbb{F}, L, \rho = 10^{-3}]$ 。根据命题 1 可得，一定存在一个二元多项式 $g(X, Y) \in \mathbb{F}[X, Y]$ 满足：

1. 对于 $\forall x \in L$ 都有 $g(x, q(x)) = f(x)$ ，其中 $q(x) = x^{1000}$
2. $\deg_X(g) < |L_0| = 10^3$
3. 由于 $f \in RS[\mathbb{F}, L, \rho = 10^{-3}]$ ，则 $\deg_Y(g) < \rho|L^{(1)}| = 10^{-3} \times 10^6 = 10^3$ 现在 Prover 想向 Verifier 证明 $f(x)$ 的次数确实是小于 10^6 的。在文章中用了直观的几何图形来说明证明的过程。



在图中，正方形的横向表示的是自变量 X，取值范围就是 L，总共有 10^9 个，而纵向表示的是自变量 Y，其取值范围是 $\{x^{1000}|x \in L\}$。正方形中的一个点 (x, y) 对应的值表示的就是计算出的 $g(x, y)$ 的值。对于在正方形的对角线上的点 (x, y)，满足 $x = y$，那么 $g(x, y) = g(x, x^{1000}) = f(x)$。

证明的过程如下：

1. Prover 承诺上述正方形中关于 $g(X, Y)$ 的所有点的估计，例如使用 Merkle 树来进行承诺。
2. Verifier 随机选取大约几十行和列，对于选择的每一行或列，Verifier 会要求例如 1010 个点的样本，确保在每种情况下所需的点之一位于对角线上。比如 Verifier 选取第 5 列，那么此时 $x = x_4$，此时需要选取 1010 个样本点，那么这些点的横坐标已经确定了，只需随机纵坐标就行，在纵坐标中选取 $y = x_4^{1000}$ 就确保了该点 (x_4, x_4^{1000}) 在对角线上了。
3. Prover 回复 Verifier 要求的点对应的值 $g(x, y)$，并带上对应的 Merkle 分支，证明它们是 Prover 原来承诺的数据的一部分。
4. Verifier 检查 Merkle 分支是否匹配，同时对于每一行或每一列，Verifier 验证 Prover 提供的这些点是否真的对应一个次数 < 1000 的多项式。Verifier 可以通过对这些点进行插值来验证这一点。

原文提到：

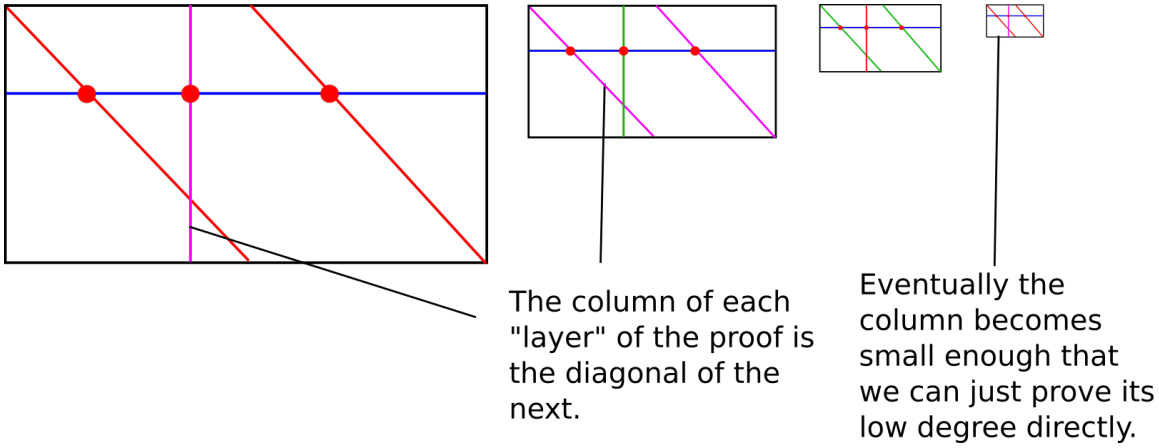
> This gives the verifier a statistical proof that (i) most rows are populated mostly by points on degree < 1000 polynomials, (ii) most columns are populated mostly by points on degree < 1000 polynomials, and (iii) the diagonal line is mostly on these polynomials. This thus convinces the verifier that most points on the diagonal actually do correspond to a degree $< 1,000,000$ polynomial.

这几点与结论可以联系命题 1 给出的那三项：

1. 对于大多数行，对应的是次数 < 1000 的多项式，也就是说明 $\deg_X(g) < 1000$。
2. 对于大多数列，对应的是次数 < 1000 的多项式，也就是说明 $\deg_Y(g) < 1000$。
3. 对角线主要由这些多项式上的点组成，也就是说明这些点的值满足 $g(x, x^{1000})$。

这也就能说明对角线上的大多数点 (x, x^{1000}) 对应一个次数 $< 10^6$ 的多项式，又因为 $f(x) = g(x, x^{1000})$，也就让 Verifier 相信多项式 $f(X)$ 的次数是 $< 10^6$ 的了。

综上，如果我们想要证明多项式 $f(X)$ 的次数小于某个值，根据命题 1，一定存在一个二元多项式 $g(X, Y)$ 能与 $f(X)$ 产生联系，首先就是 $f(x) = g(x, q(x))$，剩下两个结论是关于 $g(X, Y)$ 的次数 $\deg_X(g)$ 与 $\deg_Y(g)$ 的两个结论，这就分别对应着图中横线与竖线所表示的多项式的次数。其实可以就上述步骤进行递归，这部分对应文章中 And Even More Efficiency 小节，描述的也就是 FRI 协议的过程。



下面给出命题 1 的证明。

命题 1 证明：令 $P^{(i)} = \text{interpolant}^{f^{(i)}}$，即将函数 $f^{(i)}$ 在 $L^{(i)}$ 进行插值，得到多项式 $P^{(i)}$。用 $\mathbb{F}[X,Y]$ 表示在有限域 $\mathbb{F}$ 上的二元多项式环；先按照多项式的总次数对其中的单项式进行排序，再按照 X-次数进行排序。令

$$Q^{(i)}(X,Y) = P^{(i)}(X) \qquad \bmod Y - q^{(i)}(X) \tag{13}$$

为 $P^{(i)}(X)$ 除以 $Y - q^{(i)}(X)$ 的余式。通过该定义，可以得出一定存在一个商式 $R(X,Y) \in \mathbb{F}[X,Y]$ 使得

$$P^{(i)}(X) = Q^{(i)}(X,Y) + (Y - q^{(i)}(X)) \cdot R(X,Y). \tag{14}$$

对于 $\forall x \in L^{(i)}$ 以及 $y = q^{(i)}(x)$，带入上式中的最右边一项，可以得到 $(Y - q^{(i)}(X)) \cdot R(X,Y) = (y - q^{(i)}(x)) \cdot R(x,y) = 0$。因此 $P^{(i)}(x) = Q^{(i)}(x,y) = Q^{(i)}(x, q^{(i)}(x))$，而 $P^{(i)}(X)$ 是由 $f^{(i)}(X)$ 在 $L^{(i)}$ 上插值得到的，那么 $f^{(i)}(x) = P^{(i)}(x) = Q^{(i)}(x, q^{i}(x))$，也就证明了命题中的第 1 项。由单项式的排序，可得定义的余式 Q 满足

$$\deg_X(Q^{(i)}(X,Y)) < \deg(q^{(i)}) = |L_0^{(i)}|, \tag{15}$$

因此命题 1 的第 2 项成立。

最后证明命题 1 的第 3 项。由条件 $f^{(i)} \in RS[\mathbb{F}, L^{(i)}, \rho]$ 可得 $\deg(P^{(i)}) < \rho|L^{(i)}|$。根据除法法则以及单项式排序规则，得

$$\deg_Y(Q^{(i)}) = \left\lfloor \frac{\deg(P^{(i)})}{\deg(q^{(i)})} \right\rfloor = \left\lfloor \frac{\deg(P^{(i)})}{|L_0^{(i)}|} \right\rfloor < \left\lfloor \frac{\rho|L^{(i)}|}{|L_0^{(i)}|} \right\rfloor = \left\lfloor \rho|L^{(i+1)}| \right\rfloor \le \rho|L^{(i+1)}|. \tag{16}$$

因此得证命题 1 第 3 项。 □

引理 1 的证明

使用命题 1 的记号。由命题的第 3 项得，对于任意的 $x^{(i)}$ 有 $\deg_Y(Q^{(i)}) < \rho \cdot |L^{(i+1)}|$。下面证明

$$\forall y \in L^{(i+1)}, f^{(i+1)}(y) = Q^{(i)}(x^{(i)}, y) \tag{17}$$

上式如果成立就证明了 $\deg(f^{(i+1)}) \le \deg_Y(Q^{(i)}) < \rho \cdot |L^{(i+1)}|$，这就证明了 $\deg(f^{(i+1)}) \in \text{RS}^{(i+1)}$。

为了证明上式，先固定 $y \in L^{(i+1)}$，令 $S_y \in \mathcal{S}^{(i)}$ 是满足 $q^{(i)}(S_y) = \{y\}$ 的集合，它也是在 $L^{(i)}$ 中 $L_0^{(i)}$ 的陪集。由 $f^{(i+1)}$ 的构造可知

$$f^{(i+1)}(y) = \text{interpolant}^{f^{(i)}|_{S_y}}(x^{(i)}). \tag{18}$$

由命题 1 的第 1 项得

$$\forall x \in S_y, \quad f^{(i)}(x) = P^{(i)} = Q^{(i)}(x,y) \tag{19}$$

由命题 1 的第 2 项，可知 $\deg_X(Q^{(i)}) < |L_0^{(i)}| = |S_y|$，因此可以将 X 当作一个形式自变量，得到

$$\text{interpolant}^{f^{(i)}|_{S_y}}(X) = Q^{(i)}(X,y) \tag{20}$$

再令 $X = x^{(i)}$，左右两边的多项式 $x^{(i)}$ 上的估计肯定是相同的。从而得到

$$f^{(i+1)}(y) = \text{interpolant}^{f^{(i)}|_{S_y}}(x^{(i)}) = Q^{(i)}(x^{(i)}, y) \tag{21}$$

自然，当对于任意的 $y \in L^{(i+1)}$，有

$$\forall y \in L^{(i+1)}, f^{(i+1)}(y) = Q^{(i)}(x^{(i)}, y) \tag{22}$$

因此得证。 □

定理 1 Soundness 证明分析

本节主要说明定理 1 中 soundness 的证明思路。首先给出几个在证明中用到的定义，接着说明两个重要的引理，最后根据这两个引理来证明 soundness。

round consistency 与 失真集

soundness 分析的难点就在于怎么准确的估计出对于任意作恶的 prover， 通过和 verifier 交互，最终通过该协议的概率。想要准确的进行估计，我们就需要考虑在协议的过程中，哪些地方可能会产生误差，如果我们将这些误差过程都毫无遗失的都估计出出错的概率，最后再综合来分析，就能得到 soundness 了。在这个过程中，为了对这些可能出现误差的情况进行概率估计分析，我们需要准确地描述出这些估计，也就是我们需要对其进行量化，下面就给出在这个过程中必要的一些定义。

在第 i 步，给出关于 $f^{(i)}$ 与 $f^{(i+1)}$ 的 oracle，以及 Verifier 给出的随机数 $x^{(i)}$ 。

> ❓ 疑问
>
> ☐ 这里论文是否写错，改为 $f^{(i-1)}$？

- **inner-layer distance** 第 ith 的 *inner-layer distance* 就是 $f^{(i)}$ 距离 $\mathrm{RS}^{(i)}$ 的 $\Delta^{(i)}$-距离。

$$\delta^{(i)} \triangleq \Delta^{(i)}(f^i, \mathrm{RS}^{(i)}) \tag{23}$$

该定义就是前文提到的第 i 步的 block-wise 距离。

- **round error** 对于 $i > 0$，第 ith *round 误差集* (*round error set*) 是 $L^{(i)}$ 的一个子集，定义如下

$$A_{\mathrm{err}}^{(i)}\left(f^{(i)}, f^{(i-1)}, x^{(i-1)}\right) \triangleq \left\{y_S^{(i)} \in L^{(i)} | \mathrm{interpolant}^{f^{(i-1)}|_S}\left(x^{(i-1)}\right) \neq f^{(i)}\left(y_S^{(i)}\right)\right\} \tag{24}$$

round error set 描述的就是在第 i 轮中 Verifier 会在检查 round consistency 测试失败的那些 $L^{(i)}$ 中的元素。相应的概率就是 ith round error $\mathrm{err}^{(i)}$ 。

$$\mathrm{err}^{(i)}\left(f^{(i)}, f^{(i-1)}, x^{(i-1)}\right) \triangleq \frac{|A_{\mathrm{err}}^{(i)}|}{|L^{(i)}|} \tag{25}$$

- **closest codeword** 令 $\bar{f}^{(i)}$ 表示在 $\Delta^{(i)}(\cdot)$-测度下在 $\mathrm{RS}^{(i)}$ 中距离 $f^{(i)}$ 最近的码字。我们知道$\Delta^{(i)}(\cdot)$-测度是在 $L^{(i)}$ 的陪集划分集合 $\mathcal{S}^{(i)}$ 中的一种度量，令 $\mathcal{S}_B^{(i)} \subset \mathcal{S}^{(i)}$ 表示 $f^{(i)}$ 与码字 $\bar{f}^{(i)}$ 在划分 $\mathcal{S}^{(i)}$ 中不一致的“坏” ("bad") 的陪集，即

$$\mathcal{S}_B^{(i)} = \left\{S \in \mathcal{S}^i | f^{(i)}|_S \neq \bar{f}^{(i)}|_S\right\} \tag{26}$$

将这些在 $\mathcal{S}^{(i)}$ 中“坏”的陪集放在一起组成集合为 $D^{(i)} = \cup_{S \in \mathcal{S}_B^{(i)}} S$，可以发现 $D^{(i)}$ 是 $L^{(i)}$ 的子集，其中每一个元素是一个“坏”的陪集。

如果 $\delta^{(i)} < (1-\rho)/2$，那么根据上文关于 block-wise 距离 $\Delta^{(i)}$ 的不等式，可得

$$\Delta_H^{(i)} \leq \delta^{(i)} < (1-\rho)/2, \tag{27}$$

根据相对 Hamming 距离的界，此时可以唯一解码，根据 $f^{(i)}$ 可以解码出唯一的 $\bar{f}^{(i)}$，那么此时自然 $\mathcal{S}_B^{(i)}$ 是唯一的，进而 $\Delta_H^{(i)}$ 也就能唯一确定了。

- **失真集** 对于 $\epsilon > 0$，$f^{(i)}$ 的*失真集* (*distortion set*) 为

$$B\left[f^{(i)}; \epsilon\right] \triangleq \left\{x^{(i)} \in \mathbb{F} | \Delta_H\left(f^{(i+1)}_{f^{(i)}, x^{(i)}}, \mathrm{RS}^{(i+1)}\right) < \epsilon\right\} \tag{28}$$

注意上述使用的测度是相对 Hamming 距离。可以这样来理解这个失真集，我们知道 Verifier 会从有限域 $\mathcal{F}$ 中选取随机数 $x^{(i)}$ 发送给 Prover，Prover 根据 Verifier 发送的 $x^{(i)}$ 以及 $f^{(i)}$ 去构造下一步的 $f^{(i+1)}$，接着我们看构造的下一步的 $f^{(i+1)}$ 与 $\mathrm{RS}^{(i+1)}$ 之间的相对 Hamming 距离，如果我们给定一个值 ϵ，我们看 $\mathbb{F}$ 中哪些 $x^{(i)}$ 会导致构造的 $f^{(i+1)}_{f^{(i)}, x^{(i)}}$ 距离编码空间 $\mathrm{RS}^{(i+1)}$ 的最小相对 Hamming 距离小于给定的参数 ϵ 。进一步理解，那就是考虑域 $\mathbb{F}$ 上所有的 $x^{(i)}$，看看哪些 $f^{(i+1)}_{f^{(i)}, x^{(i)}}$ 会距离全体编码空间 $\mathrm{RS}^{(i+1)}$ 有一定距离，这个距离参数最大就是 ϵ，根据 $\epsilon > 0$ 的条件，我们知道 $f^{(i+1)}_{f^{(i)}, x^{(i)}}$ 到编码空间至少就有一个正数的距离，肯定不在 $\mathrm{RS}^{(i+1)}$ 的空间中。

那么失真集考虑的是哪些可能出现误差的情况呢？它是从 Verifier 的行为的角度出发，考虑的是 Verifier 在挑选随机数的过程中可能由于随机数的选择导致的不再在编码空间的情况。

soundness 证明思路

刚刚讲了失真集考虑是从 Verifier 选取随机数过程中可能造成的误差，那么另一个角度就是 Prover 在构造过程或者说 COMMIT 承诺阶段产生的误差。也就是当我们要估计 soundness 时，考虑以下两种会发生误差的情况：

1. Verifier 从 $\mathbb{F}$ 中选取随机数 $x^{(i)}$ 导致的误差。
2. Prover 在 COMMIT 阶段导致的误差。

由此有了 soundness 分析的大致思路，先估计第 1 种情况发生的概率，再假设第 1 种情况不会发生，发生第 2 种情况的概率。最后再来分析两种情况都同时发生的概率，也就得到了我们想要的 soundness 。

为了估计出第 1 种情况的概率，首先给出关于失真集的一对引理，这两对引理考虑的是不同的 ϵ。我们知道在对 code 解码的过程中，会先有一个相对 Hamming 距离的参数 δ，对 δ 的值分两种情况：

1. 如果 $\delta \leq (1-\rho)/2$，则解码是唯一的，即 unique decoding 。
2. 如果 $\delta > (1-\rho)/2$，此时解码出来是一个列表，是 List decoding。

> 📖 **Notes**
>
> 为了更好地理解 List Decoding，这里给出其定义：
>
> **Definition 2** [Essential Coding Theory, Definition 7.2.1] Given $0 \leq \rho \leq 1$, $L \geq 1$, a code $C \subseteq \Sigma^n$ is (ρ, L)-list decodable if for every received word $\vec{y} \in \Sigma^n$,
>
> $$|\{c \in C | \Delta(\vec{y}, c) \leq \rho n\}| \leq L. \tag{29}$$
>
> 意思就是提前给定一个相对 Hamming 距离参数 δ，以及列表的长度上限 L，对于每一个接收到的消息 $\vec{y}$，在编码空间 C 中，只要码字 c 与消息 $\vec{y}$ 之间的相对 Hamming 距离小于等于 ρn，我们就认为 c 是有效的解码。同时要求符合该距离条件的有效编码 c 的个数不能超过 L，我们就说这个编码是 $(\rho n, L)$-list decodable。
>
> 根据 Hamming 距离，有这样一个性质：

> **Proposition 1** [Essential Coding Theory, Proposition 1.4.2] Given a code C, the following are equivalent:
>
> 1. C has minimum distance $d \geq 2$,
> 2. If d is odd, C can correct $(d-1)/2$ errors.
> 3. C can detect $d-1$ errors.
> 4. C can correct $d-1$ erasures.

假设 C 的相对 Hamming 距离为 δ，那么 $\delta = d/n$。根据上述的性质，知道对于 C，可以纠正最坏情况的错误的编码的比例为 $\leq \frac{\delta}{2}$。又由 Singleton bound 知，

$$\delta \leq 1 - \rho \tag{30}$$

因此，当错误的编码比例 $\leq \frac{1-\rho}{2}$ 时，此时这些错误是可以纠正的，也就是可以唯一编码。

下面正式给出这一对引理。Lemma 3 描述的是解码半径超过唯一解码界 $(1-\rho)/2$ 的情况，而 Lemma 4 说的是解码半径小于 $(1-\rho)/2$ 的情况，即唯一解码。

Lemma 3 [BBHR18b, Lemma 4.3] (Soundness above unique decoding radius). For any $\epsilon \leq \frac{2^\eta}{|\mathbb{F}|}$ and $f^{(i)}$ such that $\delta^{(i)} > 0$

$$\Pr_{x^{(i)} \in \mathbb{F}} \left[x^{(i)} \in B\left[f^{(i)}; \frac{1}{2} \cdot \left(\delta^{(i)}(1-\epsilon) - \rho \right) \right] \right] \leq \frac{2^\eta}{\epsilon |\mathbb{F}|} \tag{31}$$

Lemma 4 [BBHR18b, Lemma 4.4] (Soundness within unique decoding radius). If $\delta^{(i)} < (1-\rho)/2$ then

$$\Pr_{x^{(i)} \in \mathbb{F}} \left[x^{(i)} \in B\left[f^{(i)}; \delta^{(i)} \right] \right] \leq \frac{|L^{(i)}|}{|\mathbb{F}|}. \tag{32}$$

Moreover, suppose that for $i < r$ the sequences $\vec{f} = (f^{(i)}, \cdots, f^{(r)})$ and $\vec{x} = (x^{(i)}, \cdots, x^{(r-1)})$ satisfy

1. for all $j \in \{i, \cdots, r\}$ we have $\delta^{(j)} < \frac{1-\rho}{2}$
2. for all $j \in \{i, \cdots, r-1\}$ we have $\bar{f}^{(j+1)} = f^{(j+1)}_{\bar{f}^{(j)}, x^{(j)}}$
3. for all $j \in \{i, \cdots, r\}$ we have $x^{(j)} \notin B[f^{(i)}; \delta^{(j)}]$

then

$$\Pr_{s^{(i)} \in D^{(i)}} \left[\mathrm{QUERY}(\vec{f}, \vec{x}) = \mathrm{reject} \right] = 1 \tag{33}$$

and consequently

$$\Pr_{s^{(i)} \in L^{(i)}} \left[\mathrm{QUERY}(\vec{f}, \vec{x}) = \mathrm{reject} \right] \geq \frac{|D^{(i)}|}{|L^{(i)}|} = \delta^{(i)} \tag{34}$$

根据失真集的定义，这两个引理说得是在不同的解码半径 ϵ 下 Verifier 选取随机数 $x^{(i)}$ 进入失真集的概率。

Lemma 4 后面的 moreover 跟着的结论说的是如果满足如下的条件：

1. 对于所有的 $j \in \{i, \cdots, r\}$，满足唯一解码，也就是 $\delta^{(j)} < \frac{1-\rho}{2}$.
2. 对于所有的 $j \in \{i, \cdots, r-1\}$，在 $\mathrm{RS}^{(j)}$ 中，选取距离 $f^{(j)}$ 最近的码字 $\bar{f}^{(j)}$，与随机数 $x^{(i)}$ 构造的下一步的函数为 $f^{(j+1)}_{\bar{f}^{(j)}, x^{(j)}}$，假设其等于在 $\mathrm{RS}^{(j+1)}$ 中距离 $f^{(j+1)}$ 最近的码字，即满足 $\bar{f}^{(j+1)} = f^{(j+1)}_{\bar{f}^{(j)}, x^{(j)}}$.
3. 对于所有的 $j \in \{i, \cdots, r-1\}$，满足随机数 $x^{(j)}$ 没有进入失真集，即 $x^{(j)} \notin B[f^{(i)}; \delta^{(j)}]$.

那么得到的结论就是在 QUERY 阶段，如果从“坏”的陪集 $D^{(i)}$ 里去选择 s^i，那么 Veriifer 一定会在 QUERY 阶段拒绝，即

$$\Pr_{s^{(i)} \in D^{(i)}} \left[\mathrm{QUERY}(\vec{f}, \vec{x}) = \mathrm{reject} \right] = 1 \tag{35}$$

从而可以得到如果 s^i 是从整个 $L^{(i)}$ 中选取的，QUERY 阶段 Verifier 拒绝的概率至少为 $\frac{|D^{(i)}|}{|L^{(i)}|}$，即

$$\Pr_{s^{(i)} \in L^{(i)}} \left[\mathrm{QUERY}(\vec{f}, \vec{x}) = \mathrm{reject} \right] \geq \frac{|D^{(i)}|}{|L^{(i)}|} = \delta^{(i)}. \tag{36}$$

现在已经做好准备工作了，开始证明协议的 soundness。到目前为止，考虑之前提到可能发生误差的情况，soundness 证明思路如下。

1. 在 COMMIT 阶段，Verifier 可能选到失真集中的随机数。 现在 Lemma 3 和 Lemma 4 的结论可以帮助我们去估计发生这种情况的概率。我们称 Verifier 选到失真集中的随机数 $x^{(i)}$ 为发生了“坏”的事件，Verifier 总共会选择 r 个随机数，记为 $x^{(0)}, \cdots, x^{(r-1)}$，每一轮将随机数选到了失真集中的事件分别记为 $E^{(0)}, \cdots, E^{(r-1)}$，我们估计发生了一些“坏”的事件的概率的界，其概率最多为

$$\frac{3|L^{(0)}|}{|\mathbb{F}|}. \tag{37}$$

2. 在 QUERY 阶段，Verifier 可能会拒绝。 假设情况 1 不会发生，在这种条件下，估计 QUERY 阶段 Verifier 拒绝的概率的界，只进行完整的一轮的拒绝概率至少为

$$\min \left\{ \delta^{(0)}, \frac{1 - 3\rho - 2^\eta / \sqrt{|L^{(0)}|}}{4} \right\}. \tag{38}$$

3. 同时考虑情况 1 和情况 2 都会发生，同时考虑 Verifier 在 QUERY 阶段重复了 l 次，那么可以得到 FRI 协议的 soundness 至少为

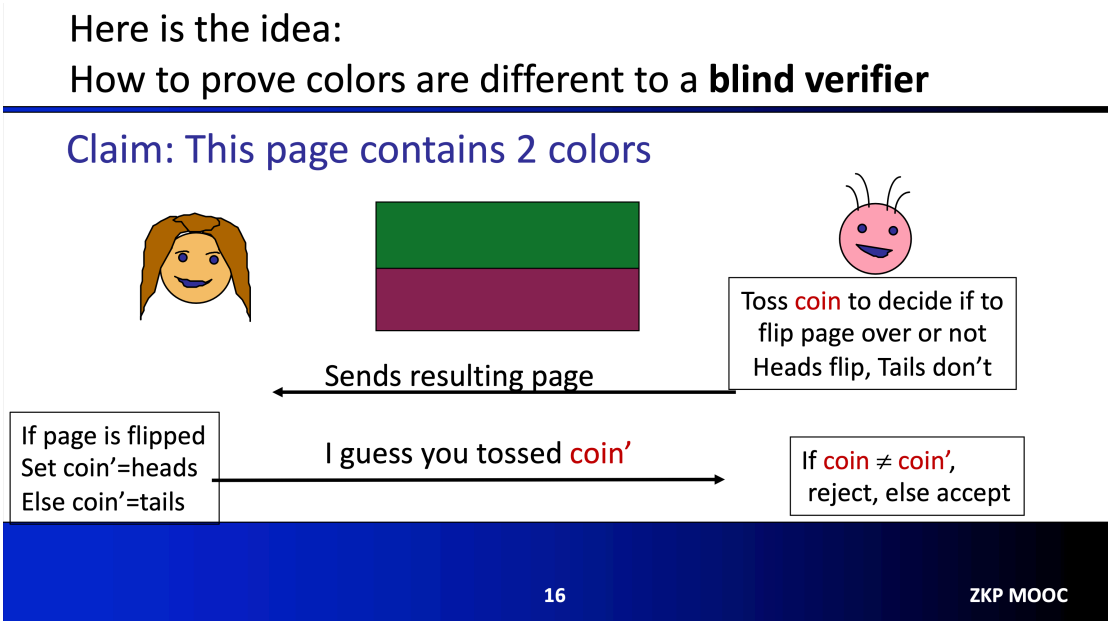
$$\mathbf{s}^{-}(\delta^{(0)}) \triangleq 1-\left(\frac{3|L^{(0)}|}{|\mathbb{F}|}+\left(1-\min\left\{\delta^{(0)}, \frac{1-3\rho-2^{\eta}/\sqrt{|L^{(0)}|}}{4}\right\}\right)^{l}\right). \tag{39}$$

🤔 **Thoughts**

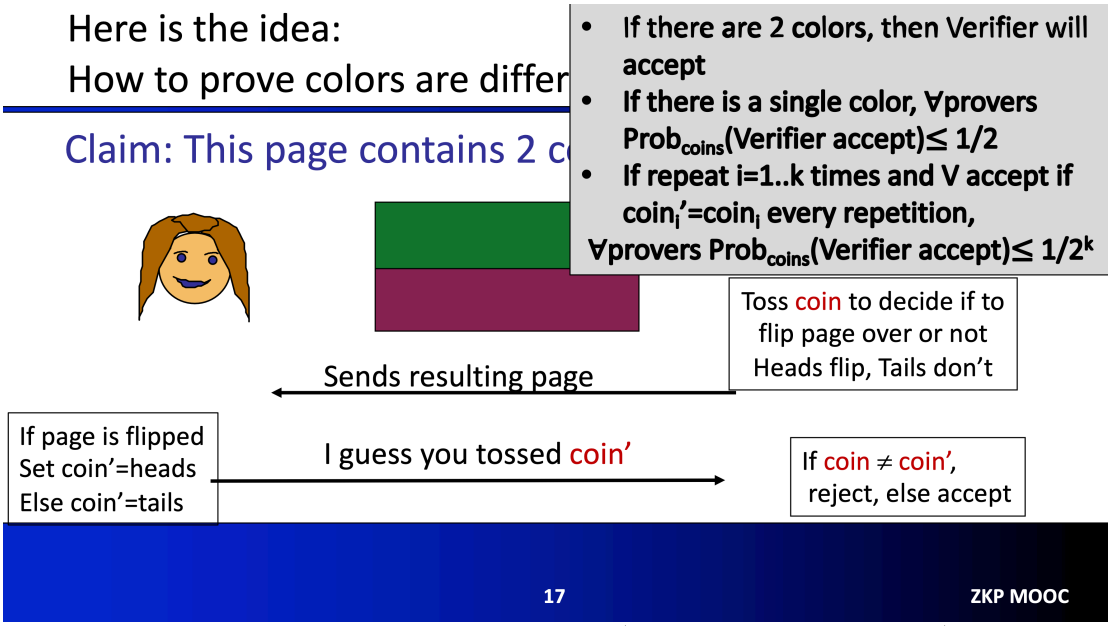
真的会发生一种情况，那就是 Verifier 选取了一些随机数 $x^{(i)}$，进入了失真集中，然后由一个距离 RS code 比较远（假设 ϵ 远）的 $f^{(i)}$ 以及 $x^{(i)}$ 构造出的 $f^{(i+1)}_{f^{(i)},x^{(i)}}$ 的这个距离没有保持，比原来更小，也就是失真了，这个时候如果我们运行 QUERY 步骤，我们没有能力能够辨别这种情况，也就是如果是一个多项式 $f^{(i+1)}_{f^{(i)},x^{(i)}}$ 它本身没有在 $\mathrm{RS}^{(i+1)}$ 中，同时呢它又距离 $\mathrm{RS}^{(i+1)}$ 小于 ϵ，Veriifier 具备的能力是能够辨别出一个多项式它距离 RS code 空间有 ϵ 那么远，现在它困惑了，迷失了，它认为 Prover 没有作弊，因为这个时候确实小于给的一个参数 ϵ，最后它输出了 accept。

😎 **关于整体 soundness 概率推导的想法**

首先考虑一个最简单的 ZK 协议 (该例子与图片来自 Zero Knowledge Proofs - Introduction and History of ZKP)



我们现在考虑 soundness 分析，说的是在 Prover 给出一个不是有两种颜色的纸的情况下，计算 Verifier 拒绝的概率。这里假设 Prover 用的是一张只有一个颜色的纸来和 Veriifier 进行交互，那么每次 Prover 最多有 $1/2$ 的概率能够通过，也就是 Veriifer 能够输出 accpet.最终得到的概率如下图所示。



如果我们来分析 soundness，那就是 Verifier 拒绝的概率，接受的概率最多为 $1/2$，那么一次交互拒绝的概率就至少为 $1/2$。如果要重复 k 次，那么 soundness 为，对于任意的 P^*，有

$$\Pr[\langle \mathrm{P}^* \leftrightarrow \mathrm{V}\rangle = \text{reject} | \text{This page only contains 1 color}] \geq 1-\left(\frac{1}{2}\right)^k \tag{40}$$

类似于这个简单的例子分析 soundness 的过程，我们来看看 FRI 的 soundness。简单例子的概率考虑的是在输入错误的知识的情况下，从 Verifier 抛随机硬币中我们有概率能使得 Verifier 最后接受。对于 FRI 协议来说，就是在我们输入一个 $f^{(0)} \notin \mathrm{RS}^{(0)}$，它不在 $\mathrm{RS}^{(0)}$ 中，那么如何衡量呢，我们衡量它在 block-wise 测度下距离 $\mathrm{RS}^{(0)}$ 有多远，即$\delta^{(0)} \triangleq \Delta^{(0)}(f^0, \mathrm{RS}^{(0)}) > 0$。接着我们类似地考虑 Veriifier 抛随机数能让 Prover 有空子可钻。由于 Verifier 抛了一些随机数 $x^{(i)}$ 使得 Prover 能够用错误的 $f^{(0)} \notin \mathrm{RS}^{(0)}$ 通过协议。也就是一些“坏”的事件发生了，使得

> 选到的随机数进入了失真集，那么 Verifier 通过的概率最多为 $\frac{3|L^{(0)}|}{|\mathbb{F}|}$.
>
> 还有一个是发生在 QUERY 阶段 Verifier 会拒绝的概率，上面那个例子 Verifier 直接判断 Prover 发的 coin' 与 Verifier 自己手里有的 coin 是否相等，是直接的，也没有引入什么随机性，如果计算不相等，就会直接拒绝，不会说还有钻空子的机会。那么我们现在审视下 FRI 协议中的 QUERY 阶段是否有什么会是包含随机性的呢？我们会发现在 QUERY 阶段，Verifier 会从 $L^{(0)}$ 中选取随机数 $s^{(0)}$，然后再进行计算检查 round consistency 是否能够通过，这个 $s^{(0)}$ 引入的随机性的过程就是我们去估计在 QUERY 阶段 Verifier 会拒绝的概率的关键。
>
> 为了能够更加清晰地分析清楚，假设 COMMIT 阶段 Verifier 选取的随机数 $x^{(i)}$ 都没有落入失真集。接着我们看看 QUERY 阶段引入的随机，也就是 $s^{(0)}$ 的选取。可以用 Lemma 4 的 moreover 的结论来看，如果其中的三个条件都成立，给出了一个拒绝的可能性，那就是至少为 $\delta^{(0)}$，然后再来考虑这三个条件不同时满足的情况下 Verifier 会拒绝的概率至少是多少。这时在证明的过程中会用到集合 $A_{\text{err}}^{(i)}$ 和 $D^{(i)}$。

下面正式给出 Soundness 证明。

定理 1 Soundness 证明：设 $\epsilon = \frac{2^\eta}{|L^{(r/2)}|}$；为简单起见，假设 r 是偶数（使用 $\epsilon = \frac{2^\eta}{|L^{\lfloor r/2 \rfloor}|}$ 会得到同样的界，但是其分析会有一点复杂）。

Part I - 一系列的坏事件 第 i 个坏事件 $E^{(i)}$ 定义如下：

- **large distance:** 如果 $\delta^{(i)} \geq \frac{1-\rho}{2}$，那么 $E^{(i)}$ 就是事件

$$x^{(i)} \in B\left[f^{(i)}; \frac{1}{2} \cdot \left(\delta^{(i)}(1-\epsilon) - \rho\right)\right] \tag{41}$$

- **small distance:** 如果 $\delta^{(i)} < \frac{1-\rho}{2}$，那么 $E^{(i)}$ 就是事件

$$x^{(i)} \in B\left[f^{(i)}; \delta^{(i)}\right] \tag{42}$$

假设事件 $E^{(i)}$ 没有发生，

1. 如果 $\delta^{(i)} < \frac{1-\rho}{2}$，那么根据事件 $E^{(i)}$ 以及失真集的定义，可以得到

$$x^{(i)} \notin B\left[f^{(i)}; \delta^{(i)}\right], \tag{43}$$

即

$$x^{(i)} \notin \left\{x^{(i)} \in \mathbb{F} | \Delta_H\left(f^{(i+1)}_{f^{(i)},x^{(i)}}, \text{RS}^{(i+1)}\right) < \delta^{(i)}\right\}, \tag{44}$$

因此可得

$$\Delta_H\left(f^{(i+1)}_{f^{(i)},x^{(i)}}, \text{RS}^{(i+1)}\right) \geq \delta^{(i)} \tag{45}$$

又根据 Block-wise 距离不等式得

$$\Delta^{(i+1)}\left(f^{(i+1)}_{f^{(i)},x^{(i)}}, \text{RS}^{(i+1)}\right) \geq \Delta_H\left(f^{(i+1)}_{f^{(i)},x^{(i)}}, \text{RS}^{(i+1)}\right) \geq \delta^{(i)} \tag{46}$$

2. 如果 $\delta^{(i)} \geq \frac{1-\rho}{2}$，那么根据事件 $E^{(i)}$ 以及失真集的定义，可以得到

$$\begin{aligned}
\Delta_H\left(f^{(i+1)}_{f^{(i)},x^{(i)}}, \text{RS}^{(i+1)}\right) &\geq \frac{1}{2} \cdot \left(\delta^{(i)}(1-\epsilon) - \rho\right) \\
&\geq \frac{1}{2} \cdot \left(\frac{(1-\rho)}{2}(1-\epsilon) - \rho\right) \\
&= \frac{(1-\rho)(1-\epsilon)}{4} - \frac{\rho}{2} \\
&= \frac{1-3\rho-\epsilon+\rho\epsilon}{4} \\
&\geq \frac{1-3\rho-\epsilon}{4}
\end{aligned} \tag{47}$$

根据 Block-wise 距离不等式得

$$\Delta^{(i+1)}\left(f^{(i+1)}_{f^{(i)},x^{(i)}}, \text{RS}^{(i+1)}\right) \geq \Delta_H\left(f^{(i+1)}_{f^{(i)},x^{(i)}}, \text{RS}^{(i+1)}\right) \geq \frac{1-3\rho-\epsilon}{4} \tag{48}$$

记 $\delta_0 = \frac{1-3\rho-\epsilon}{4}$，则总结上述两种情况，如果没有事件 $E^{(i)}$ 没有发生，则有

$$\Delta^{(i+1)}\left(f^{(i+1)}_{f^{(i)},x^{(i)}}, \text{RS}^{(i+1)}\right) \geq \min\left\{\delta^{(i)}, \delta_0\right\} \tag{49}$$

Part II - 界定一个坏的事件发生的概率 通过 Lemma 3 和 Lemma 4，以及我们对参数 ϵ 的选择，有

$$\Pr\left[E^{(i)}\right] \leq \max\left\{\frac{2^\eta}{\epsilon|\mathbb{F}|}, \frac{|L^{(i)}|}{|\mathbb{F}|}\right\} = \max\left\{\frac{|L^{(r/2)}|}{|\mathbb{F}|}, \frac{|L^{(i)}|}{|\mathbb{F}|}\right\} \tag{50}$$

由于 $|L^i|$ 是递减的，因此，当 $i \leq r/2$ 时，

$$\max\left\{\frac{|L^{(r/2)}|}{|\mathbb{F}|}, \frac{|L^{(i)}|}{|\mathbb{F}|}\right\} \leq \frac{|L^{(i)}|}{|\mathbb{F}|} \tag{51}$$

当 $i > r/2$ 时，

$$\max\left\{\frac{|L^{(r/2)}|}{|\mathbb{F}|}, \frac{|L^{(i)}|}{|\mathbb{F}|}\right\} \leq \frac{|L^{(r/2)}|}{|\mathbb{F}|} \tag{52}$$

综上得

$$\max\left\{\frac{|L^{(r/2)}|}{|\mathbb{F}|},\frac{|L^{(i)}|}{|\mathbb{F}|}\right\}\leq\begin{cases}\frac{|L^{(i)}|}{|\mathbb{F}|} & i\leq r/2\\ \frac{|L^{(r/2)}|}{|\mathbb{F}|} & i>r/2\end{cases} \tag{53}$$

因此对于事件 $E^{(0)},\cdots,E^{(r-1)}$，都不发生的概率至少是

$$\Pr\left[\bigwedge_{i=1}^{r-1}\neg E^{(i)}\right]\geq 1-\left(\sum_{i\leq r/2}\frac{|L^{(i)}|}{|\mathbb{F}|}+\frac{r}{2}\frac{|L^{(r/2)}|}{|\mathbb{F}|}\right) \tag{54}$$

由于 $\dim(L^{(i)})=\dim(L^{(0)})-i\eta$，因此

$$|L^{(i)}|=2^{\dim(L^{(i)})}=2^{\dim(L^{(0)})-i\eta}=2^{\dim(L^{(0)})}\cdot\left(\frac{1}{2^\eta}\right)^i=|L^{(0)}|\left(\frac{1}{2^\eta}\right)^i \tag{55}$$

根据 r 的定义

$$r\triangleq\lfloor\frac{k^{(0)}-\mathcal{R}}{\eta}\rfloor \tag{56}$$

而 $k^{(0)}=\log|L^{(0)}|$，可得

$$r=\lfloor\frac{k^{(0)}-\mathcal{R}}{\eta}\rfloor\leq\frac{k^{(0)}-\mathcal{R}}{\eta}=\frac{\log|L^{(0)}|-\mathcal{R}}{\eta} \tag{57}$$

则概率不等式为

$$\begin{aligned}
\Pr\left[\bigwedge_{i=1}^{r-1}\neg E^{(i)}\right]&\geq 1-\left(\sum_{i\leq r/2}\frac{|L^{(i)}|}{|\mathbb{F}|}+\frac{r}{2}\frac{|L^{(r/2)}|}{|\mathbb{F}|}\right)\\
&\geq 1-\left(\sum_{i\leq r/2}|L^{(0)}|\left(\frac{1}{2^\eta}\right)^i\frac{1}{|\mathbb{F}|}+\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}\frac{|L^{(r/2)}|}{|\mathbb{F}|}\right)\\
&(\text{代入}|L^{(i)}|=|L^{(0)}|\left(\frac{1}{2^\eta}\right)^i,r\leq\frac{\log|L^{(0)}|-\mathcal{R}}{\eta})\\
&\geq 1-\left(\frac{|L^{(0)}|}{|\mathbb{F}|}\sum_{i\leq r/2}\left(\frac{1}{2^\eta}\right)^i+\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}\cdot\frac{|L^{(0)}|}{2^{\eta\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}}}\frac{1}{|\mathbb{F}|}\right)\\
&(\text{代入}|L^{(r/2)}|=|L^{(0)}|\left(\frac{1}{2^\eta}\right)^{r/2}\leq|L^{(0)}|\left(\frac{1}{2^\eta}\right)^{\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}}=\frac{|L^{(0)}|}{2^{\eta\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}}})\\
&\geq 1-\left(\frac{|L^{(0)}|}{|\mathbb{F}|}\sum_{i\leq r/2}\left(\frac{1}{2^\eta}\right)^i+\frac{\log|L^{(0)}|-\mathcal{R}}{\eta}\cdot\frac{|L^{(0)}|}{2^{\eta\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}}}\frac{1}{|\mathbb{F}|}\right)\\
&(\text{因为}\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}\leq\frac{\log|L^{(0)}|-\mathcal{R}}{\eta},\text{而前面还有一个负号，因此整体缩小了})\\
&\geq 1-\left(\frac{|L^{(0)}|}{|\mathbb{F}|}\sum_{i\leq r/2}\left(\frac{1}{2}\right)^i+\frac{\log|L^{(0)}|-\mathcal{R}}{\eta}\cdot\frac{|L^{(0)}|}{2^{\eta\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}}}\frac{1}{|\mathbb{F}|}\right)\\
&(\text{因为}\eta\geq 1\Rightarrow 2^\eta\geq 2\Rightarrow\frac{1}{2^\eta}\leq\frac{1}{2}\Rightarrow\sum_{i\leq r/2}\left(\frac{1}{2^\eta}\right)^i\leq\sum_{i\leq r/2}\left(\frac{1}{2}\right)^i)\\
&\geq 1-\frac{1}{|\mathbb{F}|}\left(2|L^{(0)}|+\frac{\log|L^{(0)}|-\mathcal{R}}{\eta}\cdot\frac{|L^{(0)}|}{2^{\eta\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}}}\right)\\
&(\text{因为利用等差数列求和公式可得}\sum_{i\leq r/2}\left(\frac{1}{2}\right)^i=\frac{1\left(1-(\frac{1}{2})^{r/2+1}\right)}{1-\frac{1}{2}}\leq\frac{1}{2})\\
&\geq 1-\frac{1}{|\mathbb{F}|}\left(2|L^{(0)}|+\log(\rho|L^{(0)}|)\cdot\frac{|L^{(0)}|}{2^{\eta\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}}}\right)\\
&(\text{因为}\frac{\log|L^{(0)}|-\mathcal{R}}{\eta}\leq\log|L^{(0)}|-\mathcal{R}=\log|L^{(0)}|-\log(1/\rho)=\log(\rho|L^{(0)}|))\\
&=1-\frac{1}{|\mathbb{F}|}\left(2|L^{(0)}|+\log(\rho|L^{(0)}|)\cdot\sqrt{|L^{(0)}|/\rho}\right)\\
&(\text{因为}\frac{|L^{(0)}|}{2^{\eta\frac{\log|L^{(0)}|-\mathcal{R}}{2\eta}}}=\frac{|L^{(0)}|}{2^{\frac{\log|L^{(0)}|-\mathcal{R}}{2}}}=\frac{|L^{(0)}|}{2^{\frac{\log(\rho|L^{(0)}|)}{2}}}=\frac{|L^{(0)}|}{2^{\log(\sqrt{\rho|L^{(0)}|})}}=\frac{|L^{(0)}|}{\sqrt{\rho|L^{(0)}|}}=\sqrt{|L^{(0)}|/\rho})
\end{aligned} \tag{58}$$

<!-- 错误的思路：

根据定理条件知 $\rho|L^{(0)}|>16$，则

$$\rho|L^{(0)}|>16\Rightarrow\frac{1}{\rho|L^{(0)}|}<\frac{1}{16}\Rightarrow\frac{1}{\rho}<\frac{|L^{(0)}|}{16} \tag{59}$$

因此

Error: You can't use 'macro parameter character #' in math mode

\begin{aligned} \rho |L^{(0)}| > 16 & \Rightarrow \log(\rho |L^{(0)}|) < \sqrt{\rho |L^{(0)}|} \ & \Rightarrow \log(\rho |L^{(0)}|) \cdot \sqrt{|L^{(0)}| / \rho} < \sqrt{\rho |L^{(0)}|} \cdot \sqrt{|L^{(0)}|/\rho} \ & {\color{blue} (\text{由于} \rho < 1, \text{因此} \sqrt{|L^{(0)}| / \rho} > \sqrt{|L^{(0)}|} > 1, \text{两边同乘以一个大于1的数不会改变不等式的符号})} \ & \Rightarrow \log(\rho |L^{(0)}|) \cdot \sqrt{|L^{(0)}| \rho} < |L^{(0)}| \ \end{aligned}

将上述不等式代入概率不等式中得 (60)

\begin{aligned} \Pr \left[\bigwedge_{i=1}^{r-1} \neg E^{(i)} \right] & \ge 1 - \frac{1}{|\mathbb{F}|}\left(2|L^{(0)}| + \log(\rho|L^{(0)}|) \cdot \sqrt{|L^{(0)}|/\rho} \right) \ & > 1 - \frac{1}{|\mathbb{F}|}\left(2|L^{(0)}| + |L^{(0)}| \right) \ & = 1 - 3\frac{|L^{(0)}|}{|\mathbb{F}|}. \end{aligned}

下面我们假设没有事件$E^{(i)}$会发生，继续$soundness$的证明。 **$PartIII$ 当没有坏的事件发生时界定$soundness$**先回顾下$Lemma4$中对于序列$

\delta^{(j)} \ge \frac{1-\rho}{2}

最后是考虑序号3的情况，此时条件为 (62)

\delta^{(j)} < \frac{1-\rho}{2} \text{ 且 }\bar{f}^{(j+1)} \neq f_{\bar{f}^{(j)},x^{(j)}}^{(j+1)}

综上，存在一些$i \in \{0, \cdots, r-1\}$有下列两种情况之一成立$1.$$\delta^{(i)} \geq \frac{1-\rho}{2}$$2.$$\delta^{(i)} < \frac{1-\rho}{2}$ 且 $\bar{f}^{(i+1)} \neq f_{\bar{f}^{(i)},x^{(i)}}^{(i+1)}$忽略记号，令$i < r$表示满足上述

\Delta_H(\bar{f}^{(i+1)}, f_{f^{(i)},x^{(i)}}^{(i+1)}) \ge \delta_0

$Prover$如果得知$f^{(i)}$和$x^{(i)}$，按照$COMMIT$阶段的方法去诚实的执行，就能构造得到$f_{f^{(i)},x^{(i)}}^{(i+1)}$，而$\bar{f}^{(i+1)}$表示在$\mathrm{RS}^{i+1}$中距离$f^{(i+1)}$在$\Delta^{(i-

```
\Delta_H \left( f_{f^{(i)},x^{(i)}}^{(i+1)},\text{RS}^{(i+1)}\right) \ge \frac{1 - 3\rho - \epsilon }{4} = \delta_0
$$

由于 $\bar{f}^{(i+1)}$ 表示的是在 $\text{RS}^{(i+1)}$ 中距离 $f_{f^{(i)},x^{(i)}}^{(i+1)}$ 最近的码字，因此

$$
\Delta_H(\bar{f}^{(i+1)}, f_{f^{(i)},x^{(i)}}^{(i+1)}) = \Delta_H \left(
f_{f^{(i)},x^{(i)}}^{(i+1)},\text{RS}^{(i+1)}\right) \ge \delta_0
$$

因此命题成立。
```

2. $\delta^{(i)} < \frac{1-\rho}{2}$ 且 $\bar{f}^{(i+1)} \neq f_{\bar{f}^{(i)},x^{(i)}}^{(i+1)}$ 成立。为了简化描述，记 $g = f_{\bar{f}^{(i)},x^{(i)}}^{(i+1)}$。因为 $\bar{f}^{(i)} \in \mathrm{RS}^{(i)}$，那么由 Lemma 1 可得 $g = f_{\bar{f}^{(i)},x^{(i)}}^{(i+1)} \in \mathrm{RS}^{(i+1)}$。同时显然 $\bar{f}^{(i+1)} \in \mathrm{RS}^{(i+1)}$。由 $\mathrm{RS}^{(i+1)} = \mathrm{RS}^{(i+1)}[\mathbb{F}, L^{(i+1)}, \rho]$，那么由 RS code 的 MDS 性质(相对 Hamming 距离等于 $1-\rho$)可得其相对 Hamming 距离 $\Delta_H(\mathrm{RS}^{(i+1)}[\mathbb{F}, L^{(i+1)}, \rho]) = 1-\rho$，那么对于 $\mathrm{RS}^{(i+1)}$ 中的两个 code $\bar{f}^{(i+1)}$ 与 g 有，它们之间的相对 Hamming 距离至少为 $1-\rho$。由三角不等式得

$$1-\rho \le \Delta_H(\bar{f}^{(i+1)}, g) \le \Delta_H(\bar{f}^{(i+1)}, f_{f^{(i)},x^{(i)}}^{(i+1)}) + \Delta_H(f_{f^{(i)},x^{(i)}}^{(i+1)}, g) \tag{65}$$

由假设 $\delta^{(i)} < \frac{1-\rho}{2}$ 以及前面证明过的 block-wise 测度与相对 Hamming 距离之间的不等式得

$$\Delta_H(f_{f^{(i)},x^{(i)}}^{(i+1)}, g) \le \Delta^{(i)}(f_{f^{(i)},x^{(i)}}^{(i+1)}, g) = \delta^{(i)} < \frac{1-\rho}{2} \tag{66}$$

将上面的三角不等式进行移项可得

$$\begin{aligned} \Delta_H(\bar{f}^{(i+1)}, f_{f^{(i)},x^{(i)}}^{(i+1)}) &\ge \Delta_H(\bar{f}^{(i+1)}, g) - \Delta_H(f_{f^{(i)},x^{(i)}}^{(i+1)}, g) \\ &> (1-\rho) - \frac{1-\rho}{2} \\ &= \frac{1-\rho}{2} \\ &= \frac{2-2\rho}{4} \\ &> \frac{2-2\rho-(1+\rho+\epsilon)}{4} \\ &= \frac{1-3\rho-\epsilon}{4} \\ &= \delta_0 \end{aligned} \tag{67}$$

因此命题成立。

综上所述，命题得证。 □

下一个命题是

Claim 6 [BBHR18b, Claim 4.6].

$$\frac{|A_{\mathrm{err}}^{(i+1)} \cup D^{(i+1)}|}{|L^{(i+1)}|} \ge \Delta_H(\bar{f}^{(i+1)}, f_{f^{(i)},x^{(i)}}^{(i+1)}). \tag{68}$$

证明：由 $D^{(i+1)}$ 的定义可得，对于所有 $x \notin D^{(i+1)}$，有

$$\bar{f}^{(i+1)}(x) = f^{(i+1)}(x) \tag{69}$$

又由 $A_{\mathrm{err}}^{(i+1)}$ 的定义可得，对于所有 $x \notin A_{\mathrm{err}}^{(i+1)}$，有

$$f^{(i+1)}(x) = f_{f^{(i)},x^{(i)}}^{(i+1)}(x) \tag{70}$$

因此对所有 $x \notin A_{\text{err}}^{(i+1)} \cup D^{(i+1)}$ ，有

$$\bar{f}^{(i+1)}(x) = f^{(i+1)}(x) = f_{f^{(i)},x^{(i)}}^{(i+1)}(x) \tag{71}$$

根据相对 Hamming 距离的定义可得

$$\Pr_{x \in L^{(i+1)}}[\bar{f}^{(i+1)}(x) \neq f_{f^{(i)},x^{(i)}}^{(i+1)}(x)] = \Delta_H(f_{f^{(i)},x^{(i)}}^{(i+1)}, \bar{f}^{(i+1)}) \tag{72}$$

那么

$$\Pr_{x \in L^{(i+1)}}[\bar{f}^{(i+1)}(x) = f_{f^{(i)},x^{(i)}}^{(i+1)}(x)] = 1 - \Delta_H(f_{f^{(i)},x^{(i)}}^{(i+1)}, \bar{f}^{(i+1)}) \tag{73}$$

因此对于所有 $x \notin A_{\text{err}}^{(i+1)} \cup D^{(i+1)}$ ，要求以下两个等式同时成立：

1. $\bar{f}^{(i+1)}(x) = f_{f^{(i)},x^{(i)}}^{(i+1)}(x)$
2. $\bar{f}^{(i+1)}(x) = f^{(i+1)}(x)$

现在已经得到第一个等式成立的概率为 $1 - \Delta_H(f_{f^{(i)},x^{(i)}}^{(i+1)}, \bar{f}^{(i+1)})$ ，那么上述两个等式同时成立的概率肯定不会超过只要求第一个等式成立的概率，即

$$\begin{aligned}
\Pr_{x \in L^{(i+1)}}[x \notin A_{\text{err}}^{(i+1)} \cup D^{(i+1)}] &= \Pr_{x \in L^{(i+1)}}[\bar{f}^{(i+1)}(x) = f_{f^{(i)},x^{(i)}}^{(i+1)}(x) = f^{(i+1)}(x)] \\
&\leq \Pr_{x \in L^{(i+1)}}[\bar{f}^{(i+1)}(x) = f_{f^{(i)},x^{(i)}}^{(i+1)}(x)] \\
&= 1 - \Delta_H(f_{f^{(i)},x^{(i)}}^{(i+1)}, \bar{f}^{(i+1)})
\end{aligned} \tag{74}$$

因此

$$\begin{aligned}
\frac{|A_{\text{err}}^{(i+1)} \cup D^{(i+1)}|}{|L^{(i+1)}|} &= \Pr_{x \in L^{(i+1)}}[x \in A_{\text{err}}^{(i+1)} \cup D^{(i+1)}] \\
&= 1 - \Pr_{x \in L^{(i+1)}}[x \notin A_{\text{err}}^{(i+1)} \cup D^{(i+1)}] \\
&\geq 1 - (1 - \Delta_H(f_{f^{(i)},x^{(i)}}^{(i+1)}, \bar{f}^{(i+1)})) \\
&= \Delta_H(f_{f^{(i)},x^{(i)}}^{(i+1)}, \bar{f}^{(i+1)})
\end{aligned} \tag{75}$$

由此命题得证。 □

结合 Claim 5 和 Claim 6 的结论得

$$\frac{|A_{\text{err}}^{(i+1)} \cup D^{(i+1)}|}{|L^{(i+1)}|} \geq \Delta_H(\bar{f}^{(i+1)}, f_{f^{(i)},x^{(i)}}^{(i+1)}) \geq \delta_0 \tag{76}$$

即

$$\frac{|A_{\text{err}}^{(i+1)} \cup D^{(i+1)}|}{|L^{(i+1)}|} \geq \delta_0. \tag{77}$$

现在考虑在 QUERY 阶段使用的随机数 $s^{(i+1)}$ 。首先根据 $A_{\text{err}}^{(i+1)}$ 的定义，我们知道如果 $s^{(i+1)} \in A_{\text{err}}^{(i+1)}$ ，那么在 QUERY 阶段 Verifer 一定会拒绝。接着我们根据 i 的不同，分两种情况来考虑 Verifier 拒绝的概率。

如果 $i+1 = r$ ，那么由于 $f^{(r)} \in RS^{(r)}$ ，根据 $D^{(i+1)}$ 的定义，此时 $D^{(i+1)} = \emptyset$ ，在这种情况下如果 $s^{(i+1)} \in A_{\text{err}}^{(i+1)}$ ，Verifier 一定会拒绝，又

$$\frac{|A_{\text{err}}^{(i+1)} \cup D^{(i+1)}|}{|L^{(i+1)}|} = \frac{|A_{\text{err}}^{(i+1)}|}{|L^{(i+1)}|} \geq \delta_0. \tag{78}$$

此种情况下 Verifier 拒绝的概率至少为 δ_0 。

如果 $i+1 < r$ ，通过前面我们对 i 的选择，选取的 i 表示满足以下两个条件

1. $\delta^{(i)} \geq \frac{1-\rho}{2}$
2. $\delta^{(i)} < \frac{1-\rho}{2}$ 且 $\bar{f}^{(i+1)} \neq f_{\bar{f}^{(i)},x^{(i)}}^{(i+1)}$

之一的最大整数，也就说明在 i 后面的序列 $\vec{f} = (f^{(i+1)}, \cdots, f^{(r)})$ 和 $\vec{x} = (x^{(i+1)}, \cdots, x^{(r-1)})$ 都不为空，且均满足 Lemma 4 的三个条件。根据 Lemma 4 的结论，如果 $s^{(i+1)} \in D^{(i+1)}$ 那么在 QUERY 阶段就一定会拒绝。如果 $s^{(i+1)} \in A_{\text{err}}^{(i+1)}$ ，Verifier 也一定会拒绝，那么这个拒绝概率就是看这两个集合的并集的大小相比 $L^{(i+1)}$ 的大小有多大，已经证明

$$\frac{|A_{\text{err}}^{(i+1)} \cup D^{(i+1)}|}{|L^{(i+1)}|} \geq \delta_0. \tag{79}$$

因此在这种情况下拒绝的概率也至少为 δ_0 。

综合上述两种情况，拒绝的概率至少为 δ_0 。

再结合之前分析满足 Lemma 4 三种情况的拒绝概率，可以得到在没有坏的事件发生的情况下，也就是 Lemma 4 的第三个条件一定成立的前提下，有

1. Lemma 4 的前两个条件均成立，Verifier 的拒绝概率至少为 $\delta^{(0)}$ 。
2. Lemma 4 的前两个条件不完全成立，Verifier 的拒绝概率至少为 δ_0 。

由于

$$
\begin{aligned}
\frac{|L^{(r/2)}|}{\sqrt{|L^{(0)}|}} &= \frac{2^{k^{(0)}-\eta\cdot(r/2)}}{\left(2^{k^{(0)}-\eta\cdot 0}\right)^{\frac{1}{2}}} \\
&= 2^{k^{(0)}-\eta\cdot(r/2)-\frac{k^{(0)}}{2}} \\
&= 2^{\frac{k^{(0)}-\eta\cdot r}{2}} \\
&\text{（由于}k^{(0)} \geq \eta\cdot r\text{，则}k^{(0)}-\eta\cdot r \geq 0\text{）} \\
&\geq 1
\end{aligned} \tag{80}
$$

因此

$$
|L^{(r/2)}| \geq \sqrt{|L^{(0)}|} \tag{81}
$$

从而有

$$
\epsilon = \frac{2^{\eta}}{|L^{(r/2)}|} \leq 2^{\eta}/\sqrt{|L^{(0)}|} \tag{82}
$$

现在估计 δ_0，得

$$
\delta_0 = \frac{1-3\rho-\epsilon}{4} \geq \frac{1-3\rho-2^{\eta}/\sqrt{|L^{(0)}|}}{4} \tag{83}
$$

因此，如果没有坏的事件发生，Verifier 的拒绝概率至少为

$$
\min\{\delta^{(0)}, \delta_0\} \geq \min\left\{\delta^{(0)}, \frac{1-3\rho-2^{\eta}/\sqrt{|L^{(0)}|}}{4}\right\} \tag{84}
$$

结合 Part II 的分析，在 COMMIT 阶段 Verifier 选取随机数的概率至少为

$$
1-\frac{3|L^{(0)}|}{\mathbb{F}}, \tag{85}
$$

那么对于任意的 Prover 的 oracle $f^{(1)}, \cdots, f^{(r)}$，在 QUERY 协议中的重复参数为 l，Verifier 输出 accept 的概率最多为

$$
\left(1-\min\left\{\delta^{(0)}, \frac{1-3\rho-2^{\eta}/\sqrt{|L^{(0)}|}}{4}\right\}\right)^{l} \tag{86}
$$

下面分析下如何得到 FRI 的 soundness 。根据 soundness 的定义：

对于任意的 P^*， $\Pr[\langle \mathrm{P}^* \leftrightarrow \mathrm{V}\rangle = \text{reject} | \Delta^{(0)}(f^{(0)}, RS^{(0)}) = \delta^{(0)}] \geq \mathbf{s}^{-}(\delta^{(0)})$.

soundness 分析主要是要得到拒绝概率的下界 $\mathbf{s}^{-}(\delta^{(0)})$ 。先考虑对于任意的 P^*，计算最后 Verifier 输出 accept 的概率最多为多少。通过上述分析，我们可以分两种情况考虑：

1. 如果有坏的事件 $E^{(i)}(i=1,\cdots,r-1)$ 发生，那么 Verifier 输出 accept 的概率最多为

$$
\frac{3|L^{(0)}|}{\mathbb{F}} \tag{87}
$$

2. 如果没有坏的事件 $E^{(i)}(i=1,\cdots,r-1)$ 发生，Verifier 输出 accept 的概率最多为

$$
\left(1-\min\left\{\delta^{(0)}, \frac{1-3\rho-2^{\eta}/\sqrt{|L^{(0)}|}}{4}\right\}\right)^{l} \tag{88}
$$

因此，对于任意的 P^*，可以得到 Verifier 输出 accept 的概率的上界，即

$$
\Pr[\langle \mathrm{P}^* \leftrightarrow \mathrm{V}\rangle = \text{accept} | \Delta^{(0)}(f^{(0)}, RS^{(0)}) = \delta^{(0)}] \leq \frac{3|L^{(0)}|}{\mathbb{F}} + \left(1-\min\left\{\delta^{(0)}, \frac{1-3\rho-2^{\eta}/\sqrt{|L^{(0)}|}}{4}\right\}\right)^{l} \tag{89}
$$

从而对于任意的 P^*，有

$$
\begin{aligned}
\Pr[\langle \mathrm{P}^* \leftrightarrow \mathrm{V}\rangle = \text{reject} | \Delta^{(0)}(f^{(0)}, RS^{(0)}) = \delta^{(0)}] &= 1 - \Pr[\langle \mathrm{P}^* \leftrightarrow \mathrm{V}\rangle = \text{accept} | \Delta^{(0)}(f^{(0)}, RS^{(0)}) = \delta^{(0)}] \\
&\geq 1 - \left(\frac{3|L^{(0)}|}{\mathbb{F}} + \left(1-\min\left\{\delta^{(0)}, \frac{1-3\rho-2^{\eta}/\sqrt{|L^{(0)}|}}{4}\right\}\right)^{l}\right)
\end{aligned} \tag{90}
$$

从而，得到 FRI 的 soundness 至少为

$$
\mathbf{s}^{-}(\delta^{(0)}) \triangleq 1 - \left(\frac{3|L^{(0)}|}{\mathbb{F}} + \left(1-\min\left\{\delta^{(0)}, \frac{1-3\rho-2^{\eta}/\sqrt{|L^{(0)}|}}{4}\right\}\right)^{l}\right) \tag{91}
$$

至此完成 soundness 证明。 □

唯一解码半径 —— Lemma 4 的证明

证明：由于 $\delta^{(i)} < \frac{1-\rho}{2}$ ，前面介绍 closet codeword 定义中的分析已经提到 $\bar{f}$ 与 $\mathcal{S}_B(f^{(i)})$ 是唯一的。对于集合 $\mathcal{S}_B(f^{(i)})$ 中的一个“坏”的陪集 S ，即 $S \in \mathcal{S}_B(f^{(i)})$ ，令

$$X_S^{(i)} = \left\{x^{(i)} \in \mathbb{F} | \text{interpolant}^{f^{(i)}|_S}(x^{(i)}) = \text{interpolant}^{\bar{f}^{(i)}|_S}(x^{(i)})\right\} \tag{92}$$

集合 $X_S^{(i)}$ 表示的是那些在 $\mathbb{F}$ 中“误导(misleading)”的 $x^{(i)}$，意思是插值多项式 $\text{interpolant}^{f^{(i)}|_S}(x^{(i)}) = \text{interpolant}^{\bar{f}^{(i)}|_S}(x^{(i)})$ 是一致的，但是由于 S 来自于“坏”的陪集，实际上它们是不同的 low-degree 多项式，即 $f^{(i)}|_S \neq \bar{f}^{(i)}|_S$ 。换句话说，这些 $x^{(i)}$ “误导”了我们，明明不是相同的多项式，用 $x^{(i)}$ 在 S 上插值出来的多项式却是一致的。在下面我们要证明

$$B\left[f^{(i)}, \delta^{(i)}\right] = \bigcup_{S \in \mathcal{S}_B(f^{(i)})} X_S^{(i)} \tag{93}$$

由于 $\bar{f}^{(i)} \in \text{RS}^{(i)}$ ，那么由 Lemma 1 得 $f^{(i+1)}_{\bar{f}^{(i)}, x^{(i)}} \in \text{RS}^{(i+1)}$ 。对于所有的 $S \notin \mathcal{S}_B(f^{(i)})$ ，并且 $y_S = q^{(i)}(S)$ ，那么由于 $\mathcal{S}_B(f^{(i)}) = \left\{S \in \mathcal{S}^{(i)} | f^{(i)}|_S \neq \bar{f}^{(i)}|_S\right\}$，因此对于 $\forall S \notin \mathcal{S}_B(f^{(i)})$ ，有 $f^{(i)}|_S = \bar{f}^{(i)}|_S$ ，自然 $\text{interpolant}^{f^{(i)}|_S} = \text{interpolant}^{\bar{f}^{(i)}|_S}$ ，向插值多项式中代入 $x^{(i)}$ 可得 $\text{interpolant}^{f^{(i)}|_S}(x^{(i)}) = \text{interpolant}^{\bar{f}^{(i)}|_S}(x^{(i)})$ ，则 $f^{(i+1)}_{f^{(i)}, x^{(i)}}(y_S) = f^{(i+1)}_{\bar{f}^{(i)}, x^{(i)}}(y_S)$ 。由于 $\delta^{(i)}$ 小于唯一解码半径 $\frac{1-\rho}{2}$ ，结合 $f^{(i+1)}_{\bar{f}^{(i)}, x^{(i)}} \in \text{RS}^{(i+1)}$ 与 $\forall S \notin \mathcal{S}_B(f^{(i)})$ ，有 $f^{(i+1)}_{f^{(i)}, x^{(i)}}(y_S) = f^{(i+1)}_{\bar{f}^{(i)}, x^{(i)}}(y_S)$ 可得 $f^{(i+1)}_{\bar{f}^{(i)}, x^{(i)}}$ 是在 Hammming 距离下距离 $f^{(i+1)}_{f^{(i)}, x^{(i)}}$ 最近的 $RS^{(i+1)}$ 中的码字(codeword)。因此 $\Delta_H(f^{(i+1)}_{f^{(i)}, x^{(i)}}, \text{RS}^{(i+1)}) = \Delta_H(f^{(i+1)}_{f^{(i)}, x^{(i)}}, f^{(i+1)}_{\bar{f}^{(i)}, x^{(i)}})$ 。

> 这里描述 $f^{(i+1)}_{\bar{f}^{(i)}, x^{(i)}}$ 是在 Hammming 距离下距离 $f^{(i+1)}_{f^{(i)}, x^{(i)}}$ 最近的 $RS^{(i+1)}$ 中的码字(codeword) 的理由是否正确呢？感觉解释得还不是足够清晰。

同时，这两个函数在 y_S 上的值相同当且仅当以下两个条件之一成立：

1. $S \notin \mathcal{S}_B(f^{(i)})$
2. $S \in \mathcal{S}_B(f^{(i)})$ 且 $x^{(i)} \in X_S^{(i)}$

由此可得，这两个函数在 y_S 上的值不同当且仅当以下两个条件同时成立：

1. $S \in \mathcal{S}_B(f^{(i)})$
2. $S \notin \mathcal{S}_B(f^{(i)})$ 或 $x^{(i)} \notin X_S^{(i)}$

当条件 1 成立时，条件 2 中的第一种情况 $S \notin \mathcal{S}_B(f^{(i)})$ 显然就不满足了，自然 $x^{(i)} \notin X_S^{(i)}$ 成立，那么可以得到这两个函数在 y_S 上的值不同当且仅当

$$S \in \mathcal{S}_B(f^{(i)}) \text{ 且 } x^{(i)} \notin X_S^{(i)} \tag{94}$$

即

$$x^{(i)} \notin \cup_{S \in \mathcal{S}_B(f^{(i)})} X_S^{(i)}. \tag{95}$$

这表明 $f^{(i+1)}_{f^{(i)}, x^{(i)}}$ 与（唯一）最近的 $\text{RS}^{(i+1)}$ -码字 $\bar{f}^{(i+1)}_{f^{(i)}, x^{(i)}}$ 在所有的 $\{y_S | S \in \mathcal{S}_B(f^{(i)})\}$ 上不一致当且仅当 $x^{(i)} \notin \cup_{S \in \mathcal{S}_B(f^{(i)})} X_S^{(i)}$ 。那么

$$B\left[f^{(i)}, \delta^{(i)}\right] = \left\{x^{(i)} \in \mathbb{F} | \Delta_H\left(f^{(i+1)}_{f^{(i)}, x^{(i)}}, \text{RS}^{(i+1)}\right) < \delta^{(i)}\right\} = \left\{x^{(i)} \in \mathbb{F} | \Delta_H\left(f^{(i+1)}_{f^{(i)}, x^{(i)}}, \bar{f}^{(i+1)}_{f^{(i)}, x^{(i)}}\right) < \delta^{(i)}\right\} \tag{96}$$

而 $\delta^{(i)}$ 表示的正是 $|\mathcal{S}_B(f^{(i)})|$ 比上 $L_0^{(i)}$ 的陪集个数，$\Delta_H\left(f^{(i+1)}_{f^{(i)}, x^{(i)}}, \bar{f}^{(i+1)}_{f^{(i)}, x^{(i)}}\right) < \delta^{(i)}$ 表示的含义就是 $f^{(i+1)}_{f^{(i)}, x^{(i)}}$ 与 $\bar{f}^{(i+1)}_{f^{(i)}, x^{(i)}}$ 能在某些 $\{y_S | S \in \mathcal{S}_B(f^{(i)})\}$ 上一致，这样自然小于 $\delta^{(i)}$ 。而 $f^{(i+1)}_{f^{(i)}, x^{(i)}}$ 与（唯一）最近的 $\text{RS}^{(i+1)}$ -码字 $\bar{f}^{(i+1)}_{f^{(i)}, x^{(i)}}$ 在某些的 $\{y_S | S \in \mathcal{S}_B(f^{(i)})\}$ 上一致当且仅当 $x^{(i)} \in \cup_{S \in \mathcal{S}_B(f^{(i)})} X_S^{(i)}$ 。因此可得

$$B\left[f^{(i)}, \delta^{(i)}\right] = \left\{x^{(i)} \in \mathbb{F} | \Delta_H\left(f^{(i+1)}_{f^{(i)}, x^{(i)}}, \bar{f}^{(i+1)}_{f^{(i)}, x^{(i)}}\right) < \delta^{(i)}\right\} = \bigcup_{S \in \mathcal{S}_B(f^{(i)})} X_S^{(i)} \tag{97}$$

至此得证上面要证的等式，即

$$B\left[f^{(i)}, \delta^{(i)}\right] = \bigcup_{S \in \mathcal{S}_B(f^{(i)})} X_S^{(i)} \tag{98}$$

有了这个等式之后，现在来估计等式右边的界。事实上，$\text{interpolant}^{f^{(i)}|_S}$ 与 $\text{interpolant}^{\bar{f}^{(i)}|_S}$ 是次数小于 $|S|$ 的两个不同的多项式，因此 $|X_S| < |S|$ ，否则如果 $|X_S| \geq |S|$ ，那么根据 X_S 的定义， $\text{interpolant}^{f^{(i)}|_S}$ 与 $\text{interpolant}^{\bar{f}^{(i)}|_S}$ 会在超过 $|S|$ 个点上一致，这时两个插值多项式就会相同了，这与 $\text{interpolant}^{f^{(i)}|_S}$ 和 $\text{interpolant}^{\bar{f}^{(i)}|_S}$ 是两个不同的多项式是矛盾的。因此

$$\left|B\left[f^{(i)}, \delta^{(i)}\right]\right| = \left|\bigcup_{S \in \mathcal{S}_B(f^{(i)})} X_S^{(i)}\right| < |S| \cdot \left|\mathcal{S}_B\left(f^{(i)}\right)\right| \leq |L^{(i)}|, \tag{99}$$

至此证得了 Lemma 4 的第一个不等式

$$\Pr_{x^{(i)} \in \mathbb{F}}\left[x^{(i)} \in B\left[f^{(i)}; \delta^{(i)}\right]\right] = \frac{\left|B\left[f^{(i)}, \delta^{(i)}\right]\right|}{|\mathbb{F}|} \leq \frac{|L^{(i)}|}{|\mathbb{F}|}. \tag{100}$$

下面考虑在 Lemma 中假设的序列 $\vec{f}$ 与 $\vec{x}$ 。为简单起见，我们假设 $\bar{f}^{(i)}$ 在 $L^{(i)}$ 上的求值得到是零函数，记这个函数为 $\mathbf{0}|_{L^{(i)}}$ 。如果不是这种情况，我们可以通过 $f^{(i)} - \bar{f}^{(i)}$ 来得到零函数。那么

$$f^{(i+1)}_{\bar{f}^{(i)}, x^{(i)}} = f^{(i+1)}_{\mathbf{0}|_{L^{(i)}}, x^{(i)}} = \mathbf{0}|_{L^{(i+1)}} \tag{101}$$

由引理的第 2 个假设得 $\bar{f}^{(i+1)} = f^{(i+1)}_{\bar{f}^{(i)}, x^{(i)}} = \mathbf{0}|_{L^{(i+1)}}$ ，那么

$$f^{(i+2)}_{\bar{f}^{(i+1)}, x^{(i)}} = f^{(i+2)}_{\mathbf{0}|_{L^{(i+1)}}, x^{(i)}} = \mathbf{0}|_{L^{(i+2)}} \tag{102}$$

同样由引理的第 2 个假设得 $\bar{f}^{(i+2)} = f^{(i+2)}_{\bar{f}^{(i+1)},x^{(i)}} = \mathbf{0}|_{L^{(i+2)}}$，以此类推，通过归纳法，可得对于所有的 $j \in \{i, \cdots, r\}$ 有 $\bar{f}^{(j)} = \mathbf{0}|_{L^{(j)}}$。特别地，$f^{(r)} = \mathbf{0}|_{L^{(r)}}$。

考虑在 QUERY 阶段的序列 $(s^{(i)}, \cdots, s^{(r)})$，其中 $s^{(i)} \in D^{(i)}$。令 j 表示使得 $s^{(j)} \in D^{(j)}$ 成立的最大的整数。由于 $s^{(i)} \in D^{(i)}$，因此定义的这个 j 是能得到的。由定义 $f^{(r)} = \mathbf{0}|_{L^{(r)}}$ 可得 $D^{(r)} = \emptyset$，因此 $j < r$。结合引理的第 3 个假设对所有的 $j \in \{i, \cdots, r\}$ 有 $x^{(j)} \notin B[f^{(i)}; \delta^{(j)}]$ 以及前面证得的等式

$$B\left[f^{(i)}, \delta^{(i)}\right] = \bigcup_{S \in \mathcal{S}_B(f^{(i)})} X_S^{(i)} \tag{103}$$

可得 $x^{(j)} \notin \bigcup_{S \in \mathcal{S}^{(j)}} X_S^{(j)}$，因此 $f^{(j+1)}_{f^{(j)},x^{(i)}}(s^{(j+1)}) \neq 0$。但是通过 j 的定义知 j 是使得 $s^{(j)} \in D^{(j)}$ 成立的最大的整数，那么对于比 j 大的 $j+1$有 $s^{(j+1)} \notin D^{(j+1)}$，根据 $D^{(j+1)}$ 的定义，$D^{(j+1)} = \cup_{S \in \mathcal{S}_B^{(j+1)}} S$，其中 S 表示那些 “坏” 的陪集，即

$$\mathcal{S}_B^{(j+1)} = \left\{ S \in \mathcal{S}^{(j+1)} | f^{(j+1)}|_S \neq \bar{f}^{(j+1)}|_S \right\} \tag{104}$$

而 $s^{(j+1)} \notin D^{(j+1)}$，因此有 $f^{(j+1)}(s^{(j+1)}) = \bar{f}^{(j+1)}(s^{(j+1)}) = 0$。至此我们得到

1. $f^{(j+1)}_{f^{(j)},x^{(i)}}(s^{(j+1)}) \neq 0$
2. $f^{(j+1)}(s^{(j+1)}) = 0$

因此

$$f^{(j+1)}_{f^{(j)},x^{(i)}}(s^{(j+1)}) \neq f^{(j+1)}(s^{(j+1)}) \tag{105}$$

这表示在 QUERY 阶段不会通过 round consistency 检查，也就是 Verifier 在 QUERY 阶段一定会拒绝序列 $(s^{(i)}, \cdots, s^{(r)})$。这证明了

$$\Pr_{s^{(i)} \in D^{(i)}}\left[\mathrm{QUERY}\left(\vec{f}, \vec{x}\right) = \mathrm{reject}\right] = 1 \tag{106}$$

由 $\delta^{(i)}$ 以及集合 $D^{(i)}$ 的定义可知

$$\delta^{(i)} = \frac{|D^{(i)}|}{|L^{(i)}|} \tag{107}$$

因此

$$\Pr_{s^{(i)} \in L^{(i)}}\left[\mathrm{QUERY}(\vec{f}, \vec{x}) = \mathrm{reject}\right] \geq \frac{|D^{(i)}|}{|L^{(i)}|} = \delta^{(i)} \tag{108}$$

至此证得了 Lemma 4。 □

超过唯一解码半径 —— Lemma 3 的证明

为了证明 Lemma 3，我们需要 [Spi95] 中引理 4.2.18 的以下改进版本。

Lemma 7 [BBHR18b, Lemma 4.7] Let $E(X,Y)$ be a polynomial of degree $(\alpha m, \delta n)$ and $P(X,Y)$ a polynomial of degree $((\alpha+\epsilon)m, (\delta+\rho)n)$. If there exist distinct $x_1, \cdots, x_m$ such that $E(x_i, Y) | P(x_i, Y)$ and $y_1, \cdots, y_n$ such that $E(X, y_i) | P(X, y_i)$ and

$$1 > \max\left\{\delta + \rho, 2\alpha + \epsilon + \frac{\rho}{\delta}\right\} \tag{109}$$

then $E(X,Y) | P(X,Y)$.

Lemma 3 的证明: 下面证明 Lemma 3 的逆否命题。先回顾下 Lemma 3，其说的是，对于任意的 $\epsilon \leq \frac{2^\eta}{|\mathbb{F}|}$ 以及 $f^{(i)}$，有 $\delta^{(i)} = \Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)}) > 0$, 那么

$$\frac{\left|B\left[f^{(i)}; \frac{1}{2} \cdot \left(\delta^{(i)}(1-\epsilon) - \rho\right)\right]\right|}{|\mathbb{F}|} \leq \frac{2^\eta}{\epsilon |\mathbb{F}|}. \tag{110}$$

现在对其取逆否命题，得到

命题8 如果对于某些 $\epsilon \geq \frac{2^\eta}{|\mathbb{F}|}$，如果有

$$\frac{\left|B\left[f^{(i)}; \frac{1}{2} \cdot \left(\delta^{(i)}(1-\epsilon) - \rho\right)\right]\right|}{|\mathbb{F}|} > \frac{2^\eta}{\epsilon |\mathbb{F}|} \tag{111}$$

那么

$$\delta^{(i)} = \Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)}) \leq 0 \tag{112}$$

其等价于下面的命题

命题9 对于某些 $\epsilon \geq \frac{2^\eta}{|\mathbb{F}|}$，如果有

$$\frac{\left|B\left[f^{(i)}; \frac{1}{2} \cdot \left(\delta(1-\epsilon) - \rho\right)\right]\right|}{|\mathbb{F}|} > \frac{2^\eta}{\epsilon |\mathbb{F}|} \tag{113}$$

那么

$$\delta^{(i)} < \delta \tag{114}$$

下面证明下命题 8 与命题 9 等价 ***证明***: $\Rightarrow$) 反证法，假设命题 9 结论不成立，那么

$$\delta^{(i)} \geq \delta \tag{115}$$

此时由命题 9 条件可得

$$\frac{\left|B\left[f^{(i)};\frac{1}{2}\cdot\left(\delta^{(i)}(1-\epsilon)-\rho\right)\right]\right|}{|\mathbb{F}|}\geq\frac{\left|B\left[f^{(i)};\frac{1}{2}\cdot(\delta(1-\epsilon)-\rho)\right]\right|}{|\mathbb{F}|}>\frac{2^{\eta}}{\epsilon|\mathbb{F}|}\tag{116}$$

那么满足命题 8 的条件，因此

$$\delta^{(i)}\leq 0\tag{117}$$

这与假设 $\delta^{(i)}\geq\delta$ 是矛盾的。因此命题 9 的结论成立。

⇐) 反证法，假设命题 8 结论不成立，那么

$$\delta^{(i)}>0\tag{118}$$

也就是存在这样一个 $\delta>0$，使得下面的式子成立

$$\delta^{(i)}\geq\delta>0\tag{119}$$

则

$$\frac{\left|B\left[f^{(i)};\frac{1}{2}\cdot\left(\delta^{(i)}(1-\epsilon)-\rho\right)\right]\right|}{|\mathbb{F}|}\geq\frac{\left|B\left[f^{(i)};\frac{1}{2}\cdot(\delta(1-\epsilon)-\rho)\right]\right|}{|\mathbb{F}|}\tag{120}$$

又由命题 8 的条件可得

$$\frac{\left|B\left[f^{(i)};\frac{1}{2}\cdot\left(\delta^{(i)}(1-\epsilon)-\rho\right)\right]\right|}{|\mathbb{F}|}>\frac{2^{\eta}}{\epsilon|\mathbb{F}|}\tag{121}$$

通过上述两个不等式，得到 $\frac{2^{\eta}}{\epsilon|\mathbb{F}|}$ 已经是

$$\frac{\left|B\left[f^{(i)};\frac{1}{2}\cdot\left(\delta^{(i)}(1-\epsilon)-\rho\right)\right]\right|}{|\mathbb{F}|}\tag{122}$$

的一个下界，那么可以得出

$$\frac{\left|B\left[f^{(i)};\frac{1}{2}\cdot(\delta(1-\epsilon)-\rho)\right]\right|}{|\mathbb{F}|}>\frac{2^{\eta}}{\epsilon|\mathbb{F}|}\tag{123}$$

满足命题 9 的条件，因此可以得到

$$\delta^{(i)}<\delta\tag{124}$$

这与假设是矛盾的，因此命题 8 的结论成立。 □

> 🤔 **Question**
>
> ☐ 上述证明两个命题等价是否有更简洁的证明方法?

现在已经证明了命题 9 与 Lemma 3 等价，下面证明命题 9。先固定一些常数：令 $n=|L^{(i+1)}|$，$\alpha=\frac{1}{2}(1-\epsilon-\frac{\rho}{\delta})$，$\delta'=\delta\alpha$，$B=B[f^{(i)};\delta']$ 以及 $m=|B|$。由 B 的定义可得，对任意的 $x\in B$，都有 $\Delta_H\left(f^{(i+1)}_{f^{(i)},x},\mathrm{RS}^{(i+1)}\right)<\delta'$。回顾下 cloest codeword 的定义，我们知道 $\bar{f}^{(i+1)}_{f^{(i)},x}\in\mathrm{RS}^{(i+1)}$ 是距离 $f^{(i+1)}_{f^{(i)},x}$ 最近的码字，由于我们考虑的解码半径超过唯一解码半径，可能有多个码字都距离 $f^{(i+1)}_{f^{(i)},x}$ 是最近的，这里我们任取其一。

令 $C(X,Y)$ 表示一个多项式，满足 $\deg_X(C)<m$，$\deg_Y(C)<\rho n$，并且对每一个 $x\in B$，多项式 $C(x,Y)$ 与 $\bar{f}^{(i+1)}_{f^{(i)},x}(Y)$ 都是一致的。多项式 $C(X,Y)$ 是存在的，因为根据定义，$\bar{f}^{(i+1)}_{f^{(i)},x}$ 是一个次数小于 ρn 的多项式的 evaluation 。通过命题 1，令 $Q^{(i)}$ 表示与 $f^{(i)}$ 相关的多项式，即

$$Q^{(i)}(X,Y)=P^{(i)}(X)\qquad \bmod Y-q^{(i)}(X)\tag{125}$$

通过命题 1 的第 2 项可得，$\deg_X(Q^{(i)})<|L^{(i)}_0|$，通过定义可知 $|L^{(i)}_0|=2^{\eta}$。由于 $\left|B\left[f^{(i)};\frac{1}{2}\cdot\left(\delta^{(i)}(1-\epsilon)-\rho\right)\right]\right|=\left|B\left[f^{(i)};\delta'\right]\right|=m$，则由命题 9 的条件

$$\frac{\left|B\left[f^{(i)};\frac{1}{2}\cdot\left(\delta^{(i)}(1-\epsilon)-\rho\right)\right]\right|}{|\mathbb{F}|}>\frac{2^{\eta}}{\epsilon|\mathbb{F}|}\tag{126}$$

可得

$$\frac{m}{|\mathbb{F}|}>\frac{2^{\eta}}{\epsilon|\mathbb{F}|}\tag{127}$$

因此 $2^{\eta}<\epsilon m$，因此 $\deg_X(Q^{(i)})<|L^{(i)}_0|=2^{\eta}<\epsilon m$。通过命题 1 的第 1 项可以得到对任意的 $x\in L^{(i)}$ 有

$$f^{(i)}(x)=Q^{(i)}(x,q^{(i)}(x))\tag{128}$$

根据 COMMIT 阶段的定义，有对于每一个 $y\in L^{(i+1)}$，

- 令 $S_y=\{x\in L^{(i)}|q^{(i)}(x)=y\}$ 是 $L^{(i)}_0$ 的陪集，并且通过映射 $q^{(i)}$ 将 x 映射到 y；
- $P^{(i)}_y(X)\triangleq\mathrm{interpolant}^{f^{(i)}|_{S_y}}$
- $f^{(i+1)}_{f^{(i)},x^{(i)}}(y)\triangleq P^{(i)}_y(x^{(i)})$.

那么可以得到

$$P^{(i)}_y(X)\triangleq\mathrm{interpolant}^{f^{(i)}|_{S_y}}=Q^{(i)}(X,y)\tag{129}$$

因此

$$f^{(i+1)}_{f^{(i)},x^{(i)}}(y) \triangleq P^{(i)}_y(x^{(i)}) = Q^{(i)}(x^{(i)}, y) \tag{130}$$

注意上面的 $x^{(i)} \in \mathbb{F}$，将随机数 $x^{(i)}$ 改记为 x，那么可以得到对于任意的 $x \in \mathbb{F}$ 以及任意的 $y \in L^{(i+1)}$ 有

$$f^{(i+1)}_{f^{(i)},x}(y) = Q^{(i)}(x, y). \tag{131}$$

通过 distortion set 的定义，得到

$$B\left[f^{(i)}; \delta'\right] = \left\{x \in \mathbb{F} | \Delta_H\left(f^{(i+1)}_{f^{(i)},x}, \mathrm{RS}^{(i+1)}\right) < \delta'\right\} \tag{132}$$

以及命题 9 的条件

$$\frac{\left|B\left[f^{(i)}; \delta'\right]\right|}{|\mathbb{F}|} > \frac{2^\eta}{\epsilon|\mathbb{F}|} \tag{133}$$

我们通过上述分析已经得到

1. $\forall x \in B$，有 $C(x, Y)$ 与 $\bar{f}^{(i+1)}_{f^{(i)},x}(Y)$ 都是一致的。
2. 对于任意的 $x \in \mathbb{F}$ 以及任意的 $y \in L^{(i+1)}$ 有 $f^{(i+1)}_{f^{(i)},x}(y) = Q^{(i)}(x, y)$ 。

那么对于任意的 $x \in B$ 以及任意的 $y \in L^{(i+1)}$ 有

1. $C(x, y) = \bar{f}^{(i+1)}_{f^{(i)},x}(y)$
2. $f^{(i+1)}_{f^{(i)},x}(y) = Q^{(i)}(x, y)$

由于 $\bar{f}^{(i+1)}_{f^{(i)},x}(y)$ 是在 $\mathrm{RS}^{(i+1)}$ 中距离 $f^{(i+1)}_{f^{(i)},x}(y)$ 最近的码字，根据 B 的定义可得

$$\Delta_H\left(f^{(i+1)}_{f^{(i)},x}(y), \bar{f}^{(i+1)}_{f^{(i)},x}(y)\right) < \delta' \tag{134}$$

而相对 Hamming 距离考虑的就是 $f^{(i+1)}_{f^{(i)},x}(y)$ 与 $\bar{f}^{(i+1)}_{f^{(i)},x}(y)$ 不一致的比例，因此

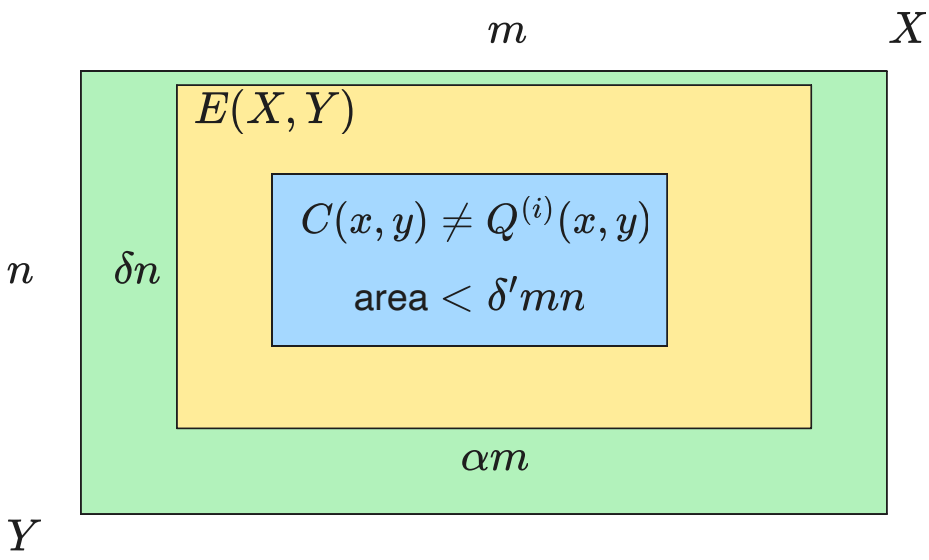
$$\Pr_{x \in B, y \in L^{(i+1)}}\left[C(x, y) \neq Q^{(i)}(x, y)\right] = \Pr_{x \in B, y \in L^{(i+1)}}\left[\bar{f}^{(i+1)}_{f^{(i)},x}(y) \neq f^{(i+1)}_{f^{(i)},x}(y)\right] < \delta'. \tag{135}$$

> 🤔 **Why?**
>
> - [ ] 下面这个非零多项式，如何证明其存在性？如何得出的？

通过构造 $\alpha\delta \geq \delta'$，因此存在一个非零多项式

$$E(X, Y), \qquad \deg_X(E) \leq \alpha m, \deg_Y(E) \leq \delta n \tag{136}$$

使得在所有的点 (x, y) 处有 $E(x, y) = 0$，其中 $x \in B, y \in L^{(i+1)}$ 且 $C(x, y) \neq Q^{(i)}(x, y)$。

> 📖 **Notes** 关于非零多项式 $E(X, Y)$ 的存在性，我是这样理解的。我们已经得到
>
> $$\Pr_{x \in B, y \in L^{(i+1)}}\left[C(x, y) \neq Q^{(i)}(x, y)\right] < \delta' \tag{137}$$
>
> 如下图所示，由于 $\alpha\delta \geq \delta'$，存在这样的一个非零多项式 $E(X, Y)$ 是合理的，其在图中蓝色的这些点的值为 0 。
>
> 
>

多项式 E 也被称为 *error locator polynomial* [Sud92]，因为它的根涵盖了错误位置的集合，其中 Q 是通过一个 low-degree 多项式得到的。

由于 $\deg_Y(C) < \rho|L^{(i+1)}|$ 以及 $\deg_X(Q^{(i)}) < 2^\eta < \epsilon m$，由 [Spi95, Chapter 4] 得存在一个多项式 $P(X, Y)$ 满足

$$\deg_X(P) < (\epsilon + \alpha)m \quad \text{且} \quad \deg_Y(P) < (\delta + \rho)n \tag{19}$$

使得

$$\forall x \in B, y \in L^{(i+1)}, \quad P(x, y) = C(x, y) \cdot E(x, y) = Q^{(i)}(x, y) \cdot E(x, y) \tag{20}$$

成立。

> 📖 **Notes** 关于 $P(X,Y)$ 多项式的存在性，我的理解目前是这样的。通过多项式的次数，来看看其存在性的合理性：先考虑自变量 X，由于 $\deg_X(E) \leq \alpha m$ 以及 $\deg_X(Q^{(i)}) < \epsilon m$，那么存在的 $P(X,Y)$ 其次数满足 $\deg_X(P) < (\epsilon+\alpha)m$，且有
>
> $$\forall x \in B, y \in L^{(i+1)}, \quad P(x,y) = Q^{(i)}(x,y) \cdot E(x,y) \tag{138}$$
>
> 是比较合理的。
>
> 同理，对于自变量 Y，由于 $\deg_Y(E) \leq \delta n$ 以及 $\deg_Y(C) < \rho n$，那么存在的 $P(X,Y)$ 其次数满足 $\deg_Y(P) < (\delta+\rho)n$，且有
>
> $$\forall x \in B, y \in L^{(i+1)}, \quad P(x,y) = C(x,y) \cdot E(x,y) \tag{139}$$
>
> 是比较合理的。

> 🙇 **TODO**
>
> - [ ] 参考 [Spi95, Chapter 4]，为什么会存在这样一个多项式。
> - [ ] 为什么 $P(x,y) = C(x,y) \cdot E(x,y) = Q^{(i)}(x,y) \cdot E(x,y)$ 成立，不是 $x \in B, y \in L^{(i+1)}$ 且 $C(x,y) \neq Q^{(i)}(x,y)$ 吗?

令 $\alpha' \triangleq \frac{\deg_X(P)}{m} - \epsilon$ 以及 $\rho' \triangleq \frac{\deg_Y(P)}{n} - \delta$，那么由公式 (19) 得

> 🐞 **Fix** 我认为论文中此处 $\alpha \triangleq \frac{\deg_X(P)}{m} - \epsilon$ 中的 α 应该改为 α'。

$$\deg_X(P) = (\epsilon+\alpha')m < (\epsilon+\alpha)m \tag{140}$$

以及

$$\deg_Y(P) = (\delta+\rho')n < (\delta+\rho)n \tag{141}$$

由此可得 $\alpha' < \alpha$ 以及 $\rho' < \rho$。

从 (19) 以及 (20) 可以得到对于任意一行 $y \in L^{(i+1)}$ 都有 $E(X,y)|P(X,y)$，类似地，对于任意一列 $x \in B$ 都有 $E(x,Y)|P(x,Y)$。换句话说，即存在不同的 $y_1, \cdots, y_n \in L^{(i+1)}$ 使得 $E(X,y_i)|P(x_i,y_i)$ 以及存在不同的 $x_1, \cdots, x_m \in B$ 使得 $E(x_i,Y)|P(x_i,Y)$。

由 (5) 式

$$1 - \rho \geq \Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)}) \geq \Delta_H(f^{(i)}, \mathrm{RS}^{(i)}) \tag{142}$$

可得 $\delta + \rho < 1$。

> 🤔 **Why?**
>
> - [ ] 这里的 $\delta + \rho < 1$ 是怎么得到的? 难道由于 $1 - \rho \geq \Delta^{(i)}(f^{(i)}, \mathrm{RS}^{(i)}) \geq \Delta_H(f^{(i)}, \mathrm{RS}^{(i)}) > \delta$?
> - [ ] 现在已知 $\Delta_H\left(f^{(i+1)}_{f^{(i)},x}, \mathrm{RS}^{(i+1)}\right) < \delta'$ 以及 $\alpha\delta = \delta'$，能否从这里推导得出呢?

通过前面推导得出的 $\alpha' < \alpha$ 以及 $\rho' < \rho$ 和 α 的定义可得

$$2\alpha' + \epsilon + \frac{\rho'}{\delta} < 2\alpha + \epsilon + \frac{\rho}{\delta} = 2 \cdot \frac{1}{2}(1 - \epsilon - \frac{\rho}{\delta}) + \epsilon + \frac{\rho}{\delta} = 1. \tag{143}$$

综合上面的推导可得

1. $\delta + \rho' < \delta + \rho < 1$
2. $2\alpha' + \epsilon + \frac{\rho'}{\delta} < 1$

则有

$$1 > \max\left\{\delta + \rho', 2\alpha' + \epsilon + \frac{\rho'}{\delta}\right\} \tag{144}$$

至此，综合上述分析，*多项式 $E(X,Y)$ 的次数为* $(\alpha' m, \delta n)$，多项式 $P(X,Y)$ 的次数为 $((\alpha'+\epsilon)m, (\delta+\rho')m)$，并且存在不同的 $x_1, \cdots, x_m \in B$ 使得 $E(x_i,Y)|P(x_i,Y)$ 以及存在不同的 $y_1, \cdots, y_n \in L^{(i+1)}$ 使得 $E(X,y_i)|P(x_i,y_i)$，同时

$$1 > \max\left\{\delta + \rho', 2\alpha' + \epsilon + \frac{\rho'}{\delta}\right\} \tag{145}$$

> 🤔 **Question**
>
> - [ ] $E(X,Y)$ 的次数为 $(\alpha' m, \delta n)$ 如何得到? 前面得到的是 $\deg_X(E) \leq \alpha m, \deg_Y(E) \leq \delta n$。

因此引理 7 的条件与假设都满足。通过引理的结论可得 $E(X,Y)|P(X,Y)$，其是环 $\mathbb{F}[X,Y]$ 中的多项式。令 $Q \equiv P/E$。我们可以得到对于每一行 $y \in L^{(i+1)}$ 且 $E(X,y)$ 非零的情况，都有 $Q(X,y) = Q^{(i)}(X,y)$。由于 $\deg_Y(E) < \delta n$，那么 $E(X,y)$ 在少于 δn 行为零，因此非零的行数的比例至少为 $1-\delta$，那么满足 $Q(X,y) = Q^{(i)}(X,y)$ 的行数比例至少为 $1-\delta$。

由于通过命题 1，我们知道 $Q^{(i)}(X,y)$ 为

$$Q^{(i)}(X,y) = P^{(i)}(X) \qquad \mod y - q^{(i)}(X) \tag{146}$$

其中

$$P^{(i)} = \mathrm{interpolant}^{f^{(i)}} \tag{147}$$

那么，$f^{(i)}$ 与次数为 $\rho|L^{(i)}|$ 的多项式 $P^{(i)}$ 是一致的。令 $S_y = \{x \in L^{(i)} | q^{(i)}(x) = y\}$，表示 $L_0^{(i)}$ 的陪集。若在陪集 S_y 上满足 $Q^{(i)}(X,y) = Q(X,y) = P^{(i)}(X)|_{S_y}$，那么根据满足 $Q(X,y) = Q^{(i)}(X,y)$ 的行数比例至少为 $1-\delta$，则 $f^{(i)}$ 与多项式 $P^{(i)}$ 在超过 $1-\delta$ 的比例的陪集 S_y 上是一致的。

> 👀 **TODO**

> - [ ] 这里理解得不是很透彻，上述解释方式不够清晰，待完善。原文
>
> > In other words $f^{(i)}$ agrees with some polynomial of degree $\rho|L^{(i)}|$ on more than a $(1-\delta)$-fraction of cosets of $L^{(i)}$ in $L^{(i)}$.

$f^{(i)}$ 在 $L^{(i)}$ 中超过 $1-\delta$ 比例的 $L_0^{(i)}$ 的陪集上和某个次数为 $\rho|L^{(i)}|$ 的多项式一致。根据定义， $\delta^{(i)}$ 说的是不一致的陪集的比例，那么自然可以得出 $\delta^{(i)} < 1-(1-\delta) = \delta$，至此完成了引理的证明。 □

参考文献

- [BBHR18a] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Scalable, transparent, and post-quantum secure computational integrity. Cryptology ePrint Archive, Report 2018/046, 2018. Available at https://eprint.iacr.org/2018/046.
- [BBHR18b] Eli Ben-Sasson, Iddo Bentov, 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.
- [RS92] Ronitt Rubinfeld and Madhu Sudan. Self-testing polynomial functions efficiently and over rational domains. In *Proceedings of the Third Annual ACM/SIGACT-SIAM Symposium on Discrete Algorithms, 27-29 January 1992, Orlando, Florida.*, pages 23–32, 1992.
- [Spi95] Daniel A. Spielman. *Computationally Efficient Error-Correcting Codes and Holographic Proofs*. PhD thesis, MIT, 1995.
- [Sud92] Madhu Sudan. *Efficient Checking of Polynomials and Proofs and the Hardness of Approximation Problems*. PhD thesis, UC Berkeley, Berkeley, CA, USA, 1992. UMI Order No. GAX93-30747.
- Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. Essential Coding Theory. https://cse.buffalo.edu/faculty/atri/courses/coding-theory/book/, 2023.
- Vitalik Buterin. STARKs, Part II: Thank Goodness It's FRI-day. https://vitalik.eth.limo/general/2017/11/22/starks_part_2.html, 2017.