Foundations and Frontiers of Probabilistic Proofs (Summer 2021)
Worksheet B.5: Analysis of FRI Protocol
Date: 2021.08.02

In this worksheet we prove *distortion lemmas*, which underlie the soundness analysis of the FRI protocol. We begin with definitions:

- if $S \subseteq \mathbb{F}^n$, we write $S^{[m]}$ for the set of $m \times n$ matrices whose rows belong to S;
- if V is an $m \times n$ matrix, we call the smallest fraction of columns of V that should be changed to obtain an element of $S^{[m]}$ the *column distance* between V and $S^{[m]}$.

A "template" for the type of lemma we wish to prove is as follows (for values of $\delta, \delta^*, \varepsilon$ that we wish to optimize):

Template Lemma. *If $S \subseteq \mathbb{F}^n$ and $v_1, \dots, v_m \in \mathbb{F}^n$ are the rows of a matrix V such that the column distance between V and $S^{[m]}$ is at least δ, then*

$$\Pr_{\alpha_1,\dots,\alpha_m}[\Delta(\alpha_1 v + \cdots + \alpha_m v_m, S) < \delta^*] \leq \varepsilon \ .$$

We start with similar lemmas that use stronger hypotheses.

Problem 1. (Distortion lemma with half distance) Fix $v_1, \dots, v_m \in \mathbb{F}^n$ and a subspace $S \subseteq \mathbb{F}^n$ such that $\Delta(v_i, S) \geq \delta$ for some $i \in [m]$. Prove that $\Pr_{\alpha_1,\dots,\alpha_m}[\Delta(\alpha_1 v + \cdots + \alpha_m v_m, S) < \delta/2] \leq \frac{1}{|\mathbb{F}|}$. (That is, prove that the statement obtained by filling in the template with $\delta^* = \delta/2$ and $\varepsilon = 1/|\mathbb{F}|$ holds under a stronger assumption.)

Problem 2. (Distortion lemma with distance preservation) Fix $v_1, \dots, v_m \in \mathbb{F}^n$ and a linear code $S \subseteq \mathbb{F}^n$ with distance $\delta(S)$. Prove that, for any $\delta < \delta(S)/4$ (i.e., for any δ at most half the unique decoding radius), if $\Delta(v_i, S) \geq \delta$ for some $i \in [m]$, then $\Pr_{\alpha_1,\dots,\alpha_m}[\Delta(\alpha_1 v_1 + \cdots + \alpha_m v_m, S) < \delta] \leq \frac{\lceil \delta n \rceil}{|\mathbb{F}|}$. (That is, prove that the statement obtained by filling in the template with $\delta^* = \delta$ and $\varepsilon \leq (\delta n + 1)/|\mathbb{F}|$ under a yet stronger assumption.) *Hint: fix a coordinate j where v_i and its closest codeword disagree. Show that the following event happens with small probability: "there exists $w \in S$ that is δ-close to the linear combination and agrees at j with it." It's also OK to prove the inequality with ε larger by a factor of 2.*

We now prove the *FRI distortion lemma* (i.e., the distortion lemma used in our analysis of the FRI protocol):

Lemma. *Define (the rate) $\rho := d/|L|$. If $f \colon L \to \mathbb{F}$ and $\delta := \mathbf{\Delta}(f, \mathrm{RS}[\mathbb{F}, L, d])$, then*

- *(large distance case) if $\delta \geq \frac{1-\rho}{2}$, then $\Pr_\alpha[\alpha \in \mathsf{Drop}(f, \delta^*)] \leq \frac{1}{|\mathbb{F}|}$, where $\delta^* := \frac{1-\rho}{8}$; and*
- *(small distance case) if $\delta < \frac{1-\rho}{2}$, then $\Pr_\alpha[\alpha \in \mathsf{Drop}(f, \delta)] \leq \frac{|L|}{|\mathbb{F}|}$.*

Recall that $\mathbf{\Delta}(f, \mathrm{RS}[\mathbb{F}, L, d])$ is the *blockwise* (or *coset*) distance between f and $\mathrm{RS}[\mathbb{F}, L, d]$ and $\mathsf{Drop}(f, \delta)$ is the set of $\alpha \in \mathbb{F}$ such that the (usual) distance between $\mathsf{Fold}(f, \alpha)$ and $\mathrm{RS}[\mathbb{F}, L^2, d/2]$ is less than δ.

Problem 3. (**FRI distortion lemma**) We first relate the template with the lemma we aim to prove, then proceed to prove it.

1. Show how the FRI distortion lemma follows from the template, and the parameters δ^*, ε thus obtained. *Hint: set* $m = 2$ *and consider the matrix whose rows are* $v_1(a^2) = \frac{f(a)+f(-a)}{2}$ *and* $v_2(a^2) = \frac{f(a)-f(-a)}{2a}$.

2. Prove the large distance case of the FRI distortion lemma using Problem 1 and the previous item.

3. Prove that, if f is within the unique decoding radius of $\mathrm{RS}[\mathbb{F}, L, d]$, then

$$\mathsf{Drop}(f, \delta) = \bigcup_{\substack{b^2 \in L^2 \\ f(b) \neq \widehat{f}(b) \text{ or} \\ f(-b) \neq \widehat{f}(-b)}} \{\, \alpha \in \mathbb{F} : \mathsf{Fold}(f, \alpha)(b^2) = \mathsf{Fold}(\widehat{f}, \alpha)(b^2) \,\},$$

where $\widehat{f} \in \mathrm{RS}[\mathbb{F}, L, d]$ is the codeword closest to f. Conclude the small distance case of the FRI distortion lemma.