Lecture B.4

FRI Protocol

(Fast Reed-Solomon IOP)

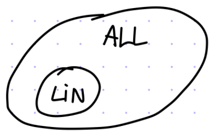
Tom Gur

Summer Graduate School on
Foundations and Frontiers of Probabilistic Proofs
2021.07.29

Recap

Linearity Testing

A function $f: \mathbb{F}^n \to \mathbb{F}$ is linear if $\exists c \in \mathbb{F}^n$ s.t. $f(x) = \sum_{i=1}^n c_i x_i$.
Equivalently, if $\forall x, y \in \mathbb{F}^n$ $f(x) + f(y) = f(x+y)$.

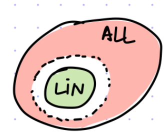


$ALL = \{f : \mathbb{F}^n \to \mathbb{F}\}$ $\quad |ALL| = |\mathbb{F}|^{|\mathbb{F}|^n}$

$LIN = \{f : \mathbb{F}^n \to \mathbb{F} \text{ is linear}\}$ $\quad |LIN| = |\mathbb{F}|^n$

We want a $O(1)$-query test that, given $f \in ALL$, says YES if $f \in LIN$ and NO if $f \notin LIN$.
But this is impossible: if f differs in 1 location from $\bar{f} \in LIN$ then $f \notin LIN$
but we cannot detect this with constant soundness error.

So we relax the question: given oracle access to $f \in ALL$,
say YES if $f \in LIN$ and NO if f is far from LIN



an instance of a problem in Property Testing

We count in Hamming distance:

$$\Delta(f,g) := \Pr_{x \in \mathbb{F}^n}[f(x) \neq g(x)] \text{ and } \Delta(f,S) := \min_{g \in S} \Delta(f,g).$$

Q1: can we solve the relaxed problem?

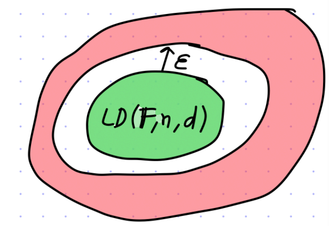
Low-Degree Testing

Recall the goal of linearity testing:

input: $\mathbb{F}, n$
oracle: $f: \mathbb{F}^n \to \mathbb{F}$
requirement: YES w.p. 1 if $f \in LIN(\mathbb{F}, n)$
YES w.p. $\frac{1}{2}$ if f is $\frac{1}{10}$-far from $LIN(\mathbb{F}, n)$

The goal of low-degree testing is:

input: $\mathbb{F}, n, d$
oracle: $f: \mathbb{F}^n \to \mathbb{F}$
requirement: YES w.p. 1 if $f \in LD(\mathbb{F}, n, d)$
YES w.p. $\frac{1}{2}$ if f is $\frac{1}{10}$-far from $LD(\mathbb{F}, n, d)$



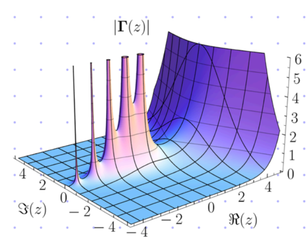
What does degree d mean?
- total degree (e.g. in this case $LD(\mathbb{F}, n, tot \leq 1) = LIN(\mathbb{F}, n)$)
- individual degree (e.g. in this case $LD(\mathbb{F}, n, ind \leq 1)$ is multilinear polys)

A test for individual degree can be derived from a test for total degree.
Either way in most applications to PCPs the difference does not matter.
Today we study total degree:

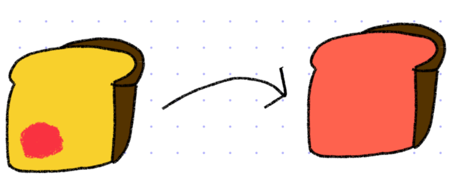
Step 1: understand $n=1$ (univariate polys) Step 2: extend to $n>1$ (multivariate polys)

Local-to-global phenomena

Idea: endow proof with a rich structure that allows checking global properties via local constraints!



aka, the "Jam principle"

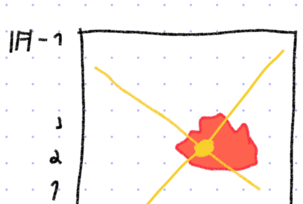


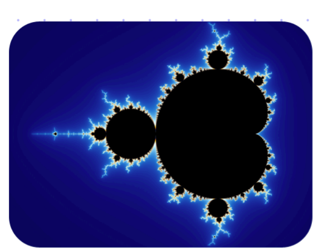
Informal example: low-degree polynomials

$p \in \mathbb{F}[X,Y]$, $deg(p) \leq d$

$p_\ell(t) := p(\bar{x} + t\bar{y}) \in \mathbb{F}[X]$

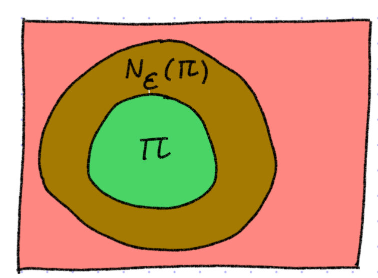
$deg(p_\ell) \leq d$





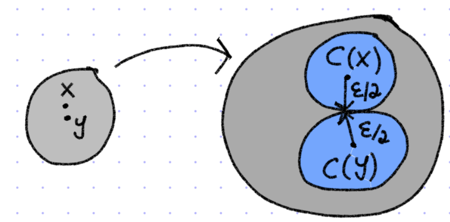
Conceptual perspectives

Property testing



Is $f \in \Pi$ (e.g. $\Pi = \mathbb{F}^{\leq d}[X]$)
or $\delta(f, \Pi) > \varepsilon$?

Coding theory



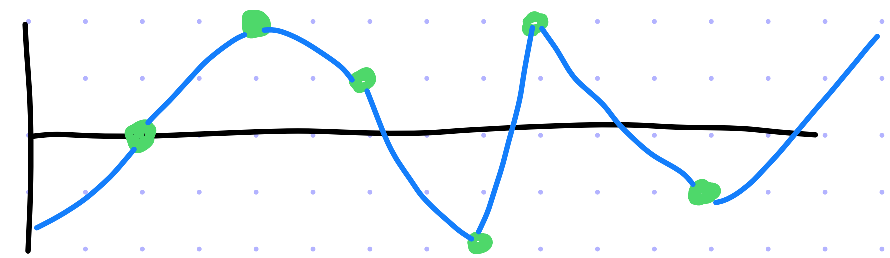
For every $x \neq y$, we have

$$\delta(C(x), C(y)) > \varepsilon$$

e.g., univariate polynomials
low-degree polynomials
Linear func. on hypercube

The Reed--Solomon Code

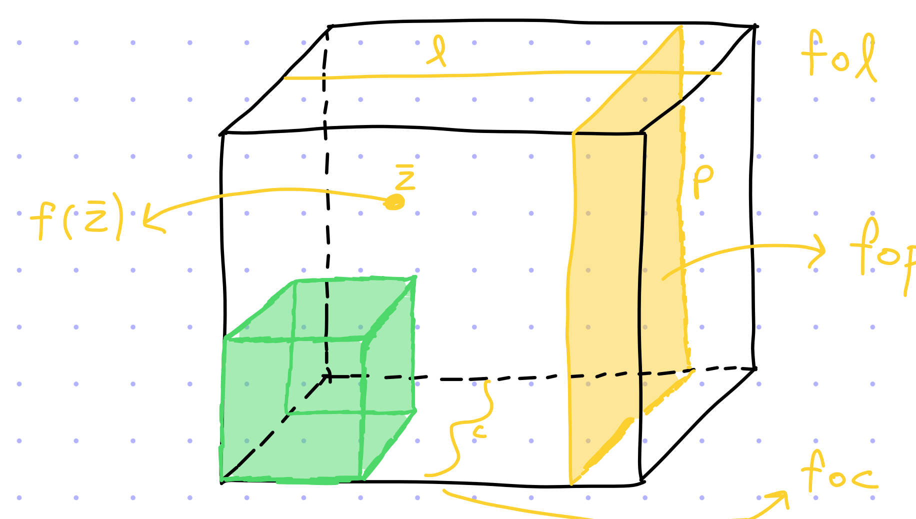
$RS[\mathbb{F}, L, d] = \{ f : L \to \mathbb{F} \text{ s.t. } \deg(\hat{f}) \leq d \}$:



locality $O(d)$ w.r.t. dim $d+1$

why wouldn't we want to use Reed-Muller instead?

$RM[\mathbb{F}, L, d, n] := \{ f : L^n \to \mathbb{F} \text{ s.t. } \deg(\hat{f}) \leq d \}$



locality $O(d)$ w.r.t. dim $O(d^n)$

Key consideration: redundancy

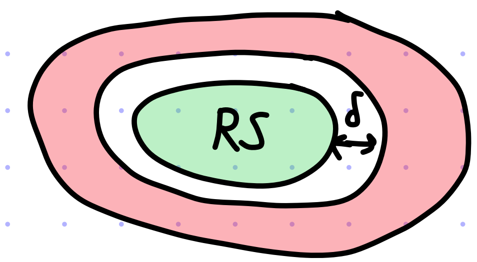
Hadamard: exp, Reed-Muller: poly, Reed-Solomon: linear!

Proximity Testing to the Reed--Solomon Code

We seek a proximity test for $\mathrm{RS}[\mathbb{F}, L, d] = \{ f : L \to \mathbb{F} \text{ s.t. } \deg(\hat{f}) \le d \}$:

① **completeness:** if $f \in \mathrm{RS}[\mathbb{F}, L, d]$ then the test accepts w.p. 1

② **soundness:** if f is δ-far from $\mathrm{RS}[\mathbb{F}, L, d]$ then the test accepts w.p. $\le \varepsilon(\delta)$ (with $\delta = \Omega(1)$ $\to \varepsilon(\delta) = O(1)$)



We have seen that:

- $d+2$ queries suffice to achieve $\varepsilon(\delta) = 1 - \delta$
- $d+2$ queries are necessary to achieve $\varepsilon(\delta) < 1$

This is ok when $d \ll |L|$ but in our case $d = \Theta(n)$ and $|L| = \Theta(d)$,
so we need query complexity that is much less than d (ideally, $\mathrm{poly}(\log d)$ or $O(1)$).

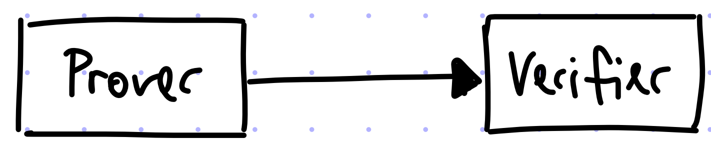
What do we do?

The above considerations are about proximity tests only.

We have the option of asking the prover's help, which leads us to a proximity proof.

Interactive Oracle Proofs

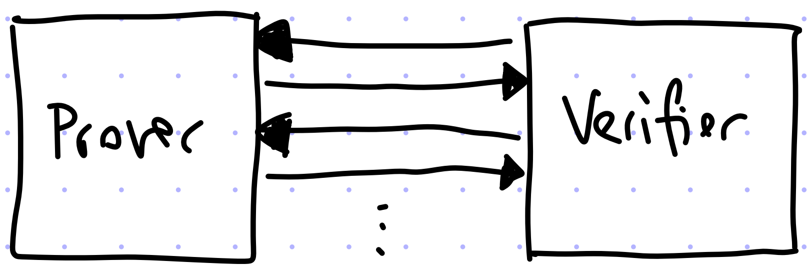
Recall that NP is the model for traditional mathematical proofs:

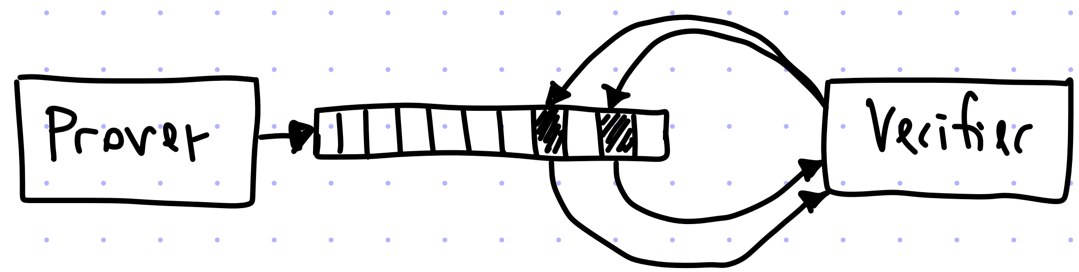


We have studied two different extensions:

IP: add randomness & interaction

PCP: add randomness & oracle access to proof

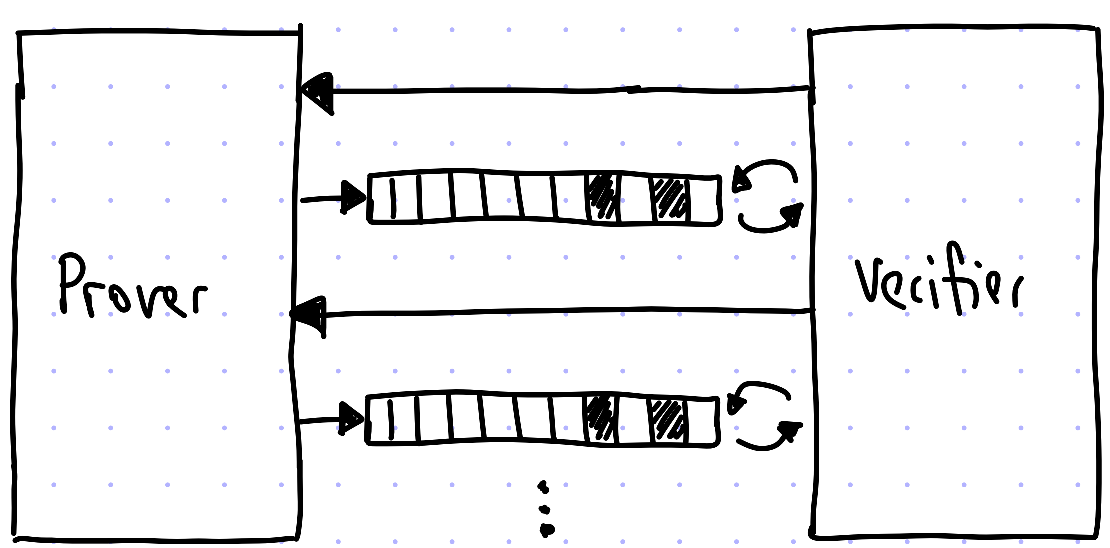




Today we consider the common extension between the two:

Interactive Oracle Proof (IOP)

add randomness, interaction, and oracle access to proof



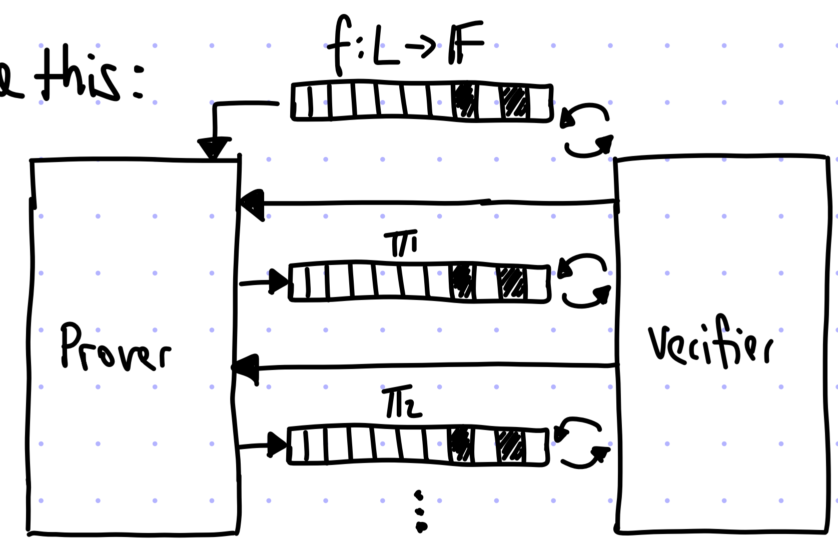
Proximity Proofs for the Reed--Solomon Code

We say that (P,V) is an IOP of proximity (IOPP) for $RS[\mathbb{F},L,d]$ if:

① completeness: if $f \in RS[\mathbb{F},L,d]$ then $\Pr[\langle P(f), V^f\rangle = 1] = 1$

② soundness: if f is δ-far from $RS[\mathbb{F},L,d]$ then $\forall \tilde{P}$ $\Pr[\langle \tilde{P}, V^f\rangle = 1] \le \varepsilon(\delta)$

An IOPP for RS look like this:



$L = \langle \omega \rangle$ with $\mathrm{ord}(\omega) = 2^k$
as a subgroup of $\mathbb{F}^*$

theorem: For every $\mathbb{F}$, smooth domain $L \subseteq \mathbb{F}$, and $d < |L|$,

$$RS[\mathbb{F},L,d] \in IOPP\left[\begin{array}{l} \varepsilon_c = 0,\ k = O(\log d),\ \ell = O(|L|),\ pt = O(|L|) \\ \varepsilon_s(\delta) = \text{“}1-\delta\text{”},\ q = O(\log d),\ vt = O(\log|L|),\ r = O(\log d) \end{array}\right]$$

this is called FRI protocol
(Fast Reed-Solomon IOPP)

This IOPP for RS is important in practice and raises many elegant questions in coding theory.

Inspiration from the Fast Fourier Transform

We can write any polynomial $\hat{f}(x) \in \mathbb{F}[x]$ as $\hat{g}(x^2) + x\,\hat{h}(x^2)$, where $\hat{g}$ are the even coefficients and $\hat{h}$ are the odd coefficients.

The (radix-2) FFT is based on the following divide-and-conquer approach:

Evaluate $\hat{f}(x)$ on $L = \langle \omega \rangle$:

1. Evaluate $\hat{g} := \mathrm{even}(\hat{f})$ on $L^2 = \langle \omega^2 \rangle$
2. Evaluate $\hat{h} := \mathrm{odd}(\hat{f})$ on $L^2 = \langle \omega^2 \rangle$
3. For $i = 0, 1, \ldots, \frac{|L|}{2} - 1$: $\hat{f}(\omega^i) := \hat{g}(\omega^{2i}) + \omega^i \hat{h}(\omega^{2i})$, $\hat{f}(-\omega^i) := \hat{g}(\omega^{2i}) - \omega^i \hat{h}(\omega^{2i})$

$-\omega^i = \omega^{i + |L|/2}$

The nested structure $L \geq L^2 \geq L^4 \geq \cdots$ enables recursion.

Each of the two subproblems have half the size, and the recursion depth is $r = \log d$.

The total number of operations is $T(|L|) = 2 \cdot T(|L|/2) + O(|L|) = O(|L| \log |L|)$.

Back to low-degree testing: $f : L \to \mathbb{F}$, $\deg(\hat{f})$ ⊛ $< d$ iff $g, h : L^2 \to \mathbb{F}$, $\deg(\hat{g}), \deg(\hat{h})$ ⊛ $< d/2$ [for $\hat{g} := \mathrm{even}(\hat{f})$ & $\hat{h} := \mathrm{odd}(\hat{f})$]

Can we devise a divide-and-conquer approach to low-degree testing?

⊛ for the rest of today we use strictly less than d

Attempt 1: Recurse on Each Subproblem

$P((\mathbb{F},L,d),f)$ | $f: L \to \mathbb{F}$ | $V((\mathbb{F},L,d))$

Compute $\hat{g} := \mathrm{even}(\hat{f})$ and $\hat{h} := \mathrm{odd}(\hat{f})$ and set $g := \hat{g}|_{L^2}$ and $h := \hat{h}|_{L^2}$

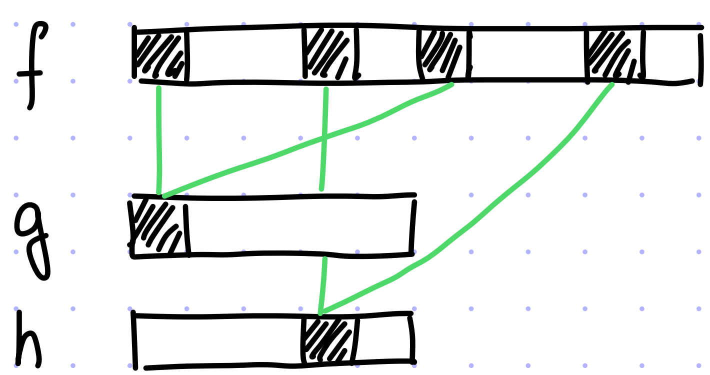
$\xrightarrow{g,h: L^2 \to \mathbb{F}}$

sample $\mu \in L$ and check $f(\mu) \stackrel{?}{=} g(\mu^2) + \mu h(\mu^2)$

recurse to test that
$g \in RS[\mathbb{F}, L^2, d/2]$
$h \in RS[\mathbb{F}, L^2, d/2]$

Problem: linear number of queries $\left(q(d) = 3 + 2q(d/2) = \Theta(d) \right)$

Problem: it's not even a test because distance decays in each recursion

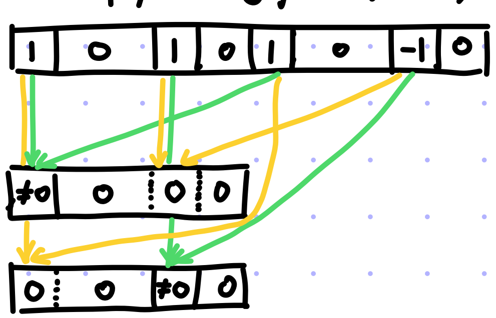


f 40 errors out of 100 locations $\delta = 40\%$

g 10 errors out of 50 locations $\delta = 20\%$

h 10 errors out of 50 locations $\delta = 20\%$

Such an example exists even if $\forall \mu \in L \;\; f(\mu) = g(\mu^2) + \mu h(\mu^2)$!



The distance could drop as $\delta \to \delta/2 \to \delta/4 \to \ldots \to \delta/2^r$.

We cannot sustain $r = \omega(1)$ rounds of interaction.

Attempt 2: Fold and Recurse

$P((\mathbb{F},L,d),f)$	$f: L \to \mathbb{F}$	$V((\mathbb{F},L,d))$
Compute $\hat{g} := \mathrm{even}(\hat{f})$ and $\hat{h} := \mathrm{odd}(\hat{f})$ and set $g := \hat{g}\|_{L^2}$ and $h := \hat{h}\|_{L^2}$	$\xrightarrow{g,h: L^2 \to \mathbb{F}}$	sample $\mu \in L$ and check $f(\mu) \stackrel{?}{=} g(\mu^2) + \mu h(\mu^2)$
	$\xleftarrow{\alpha}$	sample $\alpha \in \mathbb{F}$
Set $f_\alpha := g + \alpha h$	$\xrightarrow{f_\alpha: L^2 \to \mathbb{F}}$	check $f_\alpha(\mu^2) = g(\mu^2) + \alpha h(\mu^2)$

recurse to test that $f_\alpha \in RS[\mathbb{F}, L^2, d/2]$

The number of queries is now $q(d) = 4 + q(d/2) = O(\log d)$. This is good.

But does random folding make sense?

Let's consider the noise-free case first:

- completeness: if $\deg(\hat{f}) < d$ then $\deg(\hat{h}), \deg(\hat{g}) < d/2$ so $\forall \alpha \in \mathbb{F}$ $\deg(\hat{g} + \alpha\hat{h}) < d/2$
- "soundness": if $\deg(\hat{f}) \geq d$ then either $\deg(\hat{h}) \geq d/2$ or $\deg(\hat{g}) \geq d/2$,

  in which case $\Pr_\alpha[\deg(\hat{g} + \alpha\hat{h}) \geq d/2] \geq 1 - \frac{1}{|\mathbb{F}|}$

  as there is 1 choice of α for which the highest-degree monomial is not in $\hat{g} + \alpha\hat{f}$

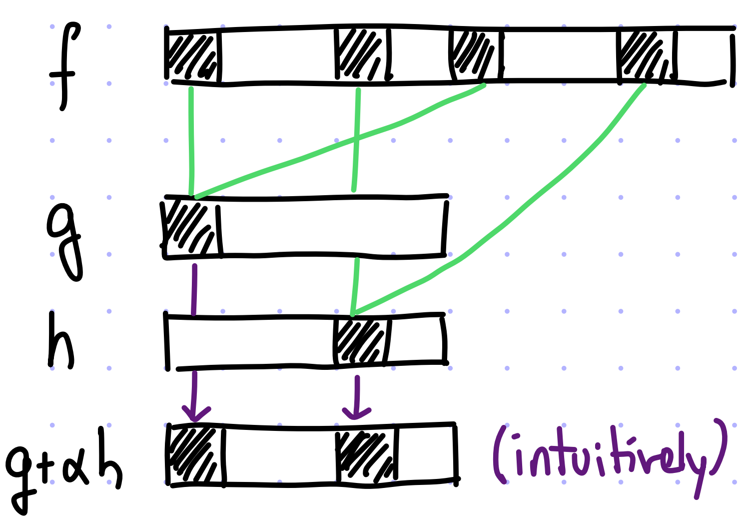
Attempt 2: Fold and Recurse

$P((\mathbb{F},L,d),f)$	$f: L \to \mathbb{F}$	$V((\mathbb{F},L,d))$
Compute $\hat{g} := \text{even}(\hat{f})$ and $\hat{h} := \text{odd}(\hat{f})$ and set $g := \hat{g}\|_{L^2}$ and $h := \hat{h}\|_{L^2}$	$\xrightarrow{g,h : L^2 \to \mathbb{F}}$	sample $\mu \in L$ and check $f(\mu) \stackrel{?}{=} g(\mu^2) + \mu h(\mu^2)$
	$\xleftarrow{\alpha}$	sample $\alpha \in \mathbb{F}$
Set $f_\alpha := g + \alpha h$	$\xrightarrow{f_\alpha : L^2 \to \mathbb{F}}$	check $f_\alpha(\mu^2) = g(\mu^2) + \alpha h(\mu^2)$

recurse to test that $f_\alpha \in RS[\mathbb{F}, L^2, d/2]$

Now consider the noisy case:

suppose f is δ-far from $RS[\mathbb{F},L,d]$.



$\delta = 20\%$

$\delta = 10\%$

$\delta = 10\%$

$\delta = 20\%$

Folding seems to address the prior problem by preserving distance!

What if the cheating prover decreases distance by sending functions g, h, f_α that are inconsistent?

We do have consistency checks in each round for this.

So, informally, we have to (at least) pay an error of

$$r \cdot \Pr[\text{a round's consistency check fails}]$$

Since $r = \Theta(\log d)$ we have two options:

(i) make $\omega(1)$ queries/round (leads to $\omega(\log d)$ queries overall)

(ii) change the protocol

The FRI Protocol

Two changes from prior protocol: drop g, h as they are not needed; do a single multi-round consistency check.

Given $f: L \to \mathbb{F}$ and $\alpha \in \mathbb{F}$, define $\mathrm{Fold}(f,\alpha): L^2 \to \mathbb{F}$ as $\mathrm{Fold}(f,\alpha)(\gamma^2) := \frac{f(\gamma)+f(-\gamma)}{2} + \alpha \cdot \frac{f(\gamma)-f(-\gamma)}{2\gamma}$.

<u>lemma</u>: $\widehat{\mathrm{Fold}(f,\alpha)}(X) \equiv \mathrm{even}(\hat{f})(X) + \alpha\, \mathrm{odd}(\hat{f})(X)$

<u>proof</u>: For every $\gamma^2 \in L^2$, $\mathrm{even}(\hat{f})(\gamma^2) + \alpha \cdot \mathrm{odd}(\hat{f})(\gamma^2) = \frac{\hat{f}(\gamma)+\hat{f}(-\gamma)}{2} + \alpha \frac{\hat{f}(\gamma)-\hat{f}(-\gamma)}{2\gamma} = \widehat{\mathrm{Fold}(f,\alpha)}(\gamma^2)$.

These changes lead to the FRI protocol:

$P((\mathbb{F},L,d),f_0)$	$f_0: L \to \mathbb{F}$	$V((\mathbb{F},L,d))$
	$\xleftarrow{\alpha_0}$	• interaction randomness: $\alpha_0, \alpha_1, \ldots, \alpha_{r-1} \leftarrow \mathbb{F}$
$f_1 := \mathrm{Fold}(f_0,\alpha_0)$	$\xrightarrow{f_1: L^2 \to \mathbb{F}}$	• consistency check randomness: $\mu \leftarrow L$ [can repeat t times]
	$\xleftarrow{\alpha_1}$	• consistency check on $f_0, f_1, \ldots, f_r$:
$f_2 := \mathrm{Fold}(f_1,\alpha_1)$	$\xrightarrow{f_2: L^4 \to \mathbb{F}}$	$f_1(\mu^2) \overset{?}{=} \frac{f_0(\mu)+f_0(-\mu)}{2} + \alpha_0 \cdot \frac{f_0(\mu)-f_0(-\mu)}{2\mu}$
	$\vdots$	$f_2(\mu^4) \overset{?}{=} \frac{f_1(\mu^2)+f_1(-\mu^2)}{2} + \alpha_1 \cdot \frac{f_1(\mu^2)-f_1(-\mu^2)}{2\mu^2}$
	$\xleftarrow{\alpha_{r-1}}$	$\vdots$
$f_r := \mathrm{Fold}(f_{r-1},\alpha_{r-1})$	$\xrightarrow{f_r: L^{2^r} \to \mathbb{F}}$	$f_r(\mu^{2^r}) \overset{?}{=} \frac{f_{r-1}(\mu^{2^{r-1}})+f_{r-1}(-\mu^{2^{r-1}})}{2} + \alpha_{r-1} \cdot \frac{f_{r-1}(\mu^{2^{r-1}})-f_{r-1}(-\mu^{2^{r-1}})}{2\mu^{2^{r-1}}}$
		• last function is low degree: $f_r \overset{?}{\in} \mathrm{RS}[\mathbb{F}, L^{2^r}, d/2^r]$

query pattern:

$f_0(\mu)$ $f_0(-\mu)$

$f_1(\mu^2)$ $f_1(-\mu^2)$

$f_2(\mu^4)$ $f_2(-\mu^4)$

$\vdots$

$f_{r-1}(\mu^{2^{r-1}})$ $f_{r-1}(-\mu^{2^{r-1}})$

$f_r(\mu^{2^r})$

Completeness

claim: FRI has perfect completeness

$P((\mathbb{F},L,d),f)$	$f_0 : L \to \mathbb{F}$	$V((\mathbb{F},L,d))$
	$\xleftarrow{\alpha_0}$	• interaction randomness: $\alpha_0, \alpha_1, \ldots, \alpha_{r-1} \leftarrow \mathbb{F}$
$f_1 := \mathrm{Fold}(f_0, \alpha_0)$	$\xrightarrow{f_1 : L^2 \to \mathbb{F}}$	• consistency check randomness: $\mu \leftarrow L$ [can repeat t times]
	$\xleftarrow{\alpha_1}$	• consistency check on $f_0, f_1, \ldots, f_r$:
$f_2 := \mathrm{Fold}(f_1, \alpha_1)$	$\xrightarrow{f_2 : L^4 \to \mathbb{F}}$	$f_1(\mu^2) \stackrel{?}{=} \frac{f_0(\mu)+f_0(-\mu)}{2} + \alpha_0 \cdot \frac{f_0(\mu)-f_0(-\mu)}{2\mu}$
	$\vdots$	$f_2(\mu^4) \stackrel{?}{=} \frac{f_1(\mu^2)+f_1(-\mu^2)}{2} + \alpha_1 \cdot \frac{f_1(\mu^2)-f_1(-\mu^2)}{2\mu^2}$
	$\xleftarrow{\alpha_{r-1}}$	$\vdots$
$f_r := \mathrm{Fold}(f_{r-1}, \alpha_{r-1})$	$\xrightarrow{f_r : L^{2^r} \to \mathbb{F}}$	$f_r(\mu^{2^r}) \stackrel{?}{=} \frac{f_{r-1}(\mu^{2^{r-1}})+f_{r-1}(-\mu^{2^{r-1}})}{2} + \alpha_{r-1} \cdot \frac{f_{r-1}(\mu^{2^{r-1}})-f_{r-1}(-\mu^{2^{r-1}})}{2\mu^{2^{r-1}}}$
		• last function is low degree: $f_r \stackrel{?}{\in} RS[\mathbb{F}, L^{2^r}, d/2^r]$

proof: Suppose that $f_0 \in RS[\mathbb{F}, L, d]$, so that $\deg(\hat{f}_0) < d$.

Fix any choice of interaction randomness:

$\alpha_0, \alpha_1, \ldots, \alpha_{r-1} \in \mathbb{F}$.

For every $i = 1, \ldots, r$, define $\hat{f}_i(X) := \mathrm{even}(\hat{f}_{i-1})(X) + \alpha_{i-1} \cdot \mathrm{odd}(\hat{f}_{i-1})(X)$.

Since $\deg(\hat{f}_0) < d$ we know that $\deg(\hat{f}_i) < d/2^i$ and thus $f_i := \hat{f}_i|_{L^{2^i}} \in RS[\mathbb{F}, L^{2^i}, d/2^i]$.

Observe that $f_i = \mathrm{Fold}(f_{i-1}, \alpha_{i-1})$ because

$$\forall \gamma \in L^{2^{i-1}} \quad f_i(\gamma^2) = \mathrm{even}(\hat{f}_{i-1})(\gamma^2) + \alpha_{i-1} \cdot \mathrm{odd}(\hat{f}_{i-1})(\gamma^2) = \frac{f_{i-1}(\gamma)+f_{i-1}(-\gamma)}{2} + \alpha_{i-1} \cdot \frac{f_{i-1}(\gamma)-f_{i-1}(-\gamma)}{2\gamma}.$$

Hence for every $\mu \in L$ all the verifier consistency checks pass.

Finally, $f_r \in RS[\mathbb{F}, L^{2^r}, d/2^r]$ as argued above, so the verifier's degree check also passes. ∎

Moreover:
- prover time is $O(|L| + |L|/2 + |L|/4 + \cdots + |L|/2^{r-1}) = O(|L|)$
- verifier time is $O(r + |L|/2^r) = O(\log d)$ when $r = \log d$ and $|L| = \Theta(d)$
- query complexity is $O(r + |L|/2^r) = O(\log d)$ when $r = \log d$ and $|L| = \Theta(d)$

Soundness

$f_0 : L \to \mathbb{F}$

$\xleftarrow{\alpha_0}$

$f_1 : L^2 \to \mathbb{F}$

$\xleftarrow{\alpha_1}$

$f_2 : L^4 \to \mathbb{F}$

$\vdots$

$\xleftarrow{\alpha_{r-1}}$

$f_r : L^{2^r} \to \mathbb{F}$

$V((\mathbb{F}, L, d))$

- interaction randomness: $\alpha_0, \alpha_1, \dots, \alpha_{r-1} \leftarrow \mathbb{F}$
- consistency check randomness: $\mu \leftarrow L$ [can repeat t times]
- consistency check on $f_0, f_1, \dots, f_r$:

$$f_1(\mu^2) \stackrel{?}{=} \frac{f_0(\mu)+f_0(-\mu)}{2} + \alpha_0 \cdot \frac{f_0(\mu)-f_0(-\mu)}{2\mu}$$

$$f_2(\mu^4) \stackrel{?}{=} \frac{f_1(\mu^2)+f_1(-\mu^2)}{2} + \alpha_1 \cdot \frac{f_1(\mu^2)-f_1(-\mu^2)}{2\mu^2}$$

$$\vdots$$

$$f_r(\mu^{2^r}) \stackrel{?}{=} \frac{f_{r-1}(\mu^{2^{r-1}})+f_{r-1}(-\mu^{2^{r-1}})}{2} + \alpha_{r-1} \cdot \frac{f_{r-1}(\mu^{2^{r-1}})-f_{r-1}(-\mu^{2^{r-1}})}{2\mu^{2^{r-1}}}$$

- last function is low degree: $f_r \stackrel{?}{\in} RS[\mathbb{F}, L^{2^r}, d/2^r]$

We will see this lower bound on soundness error:

claim: there is a prover strategy to make the verifier accept some δ-far f_0 w.p. $\geq \max\left\{\frac{1}{|\mathbb{F}|}, (1-\delta)^t\right\}$

The upper bound is, to a first order, very close:

$$O\left(\frac{|L|}{|\mathbb{F}|}\right) + \left(1 - \min\left\{\delta, c\left(\frac{d}{|L|}\right)\right\}\right)^t.$$

Here $c\left(\frac{d}{|L|}\right)$ is a universal constant with a dependence on the rate $d/|L|$.

We prove the theorem in the next lecture.

The proof relies on fundamental statements about worst-case vs average-case distances to subspaces.

Tighter upper bounds are known (which rely on tools from algebraic geometry and algebraic function fields), which lead to more efficiency in practice.

A tight soundness analysis remains an exciting open problem!