Lecture 16

Foundations of Probabilistic Proofs
Fall 2020
Alessandro Chiesa

Linear-Size IOPs for Arithmetic Computations

We have seen how to trivially adapt the basic PCP for NTIME(T) into an IOP with proof length $T^{1+O(\varepsilon)}$ and query complexity $(\log T)^{O(1/\varepsilon)}$.

Today we see how to achieve linear proof length for computations over large fields.

Recall the following NP-complete language:

def: $\mathrm{R1CS}(\mathbb{F}) = \{ (u, \underbrace{A,B,C}_{m\times n \text{ matrices}}) \mid \exists z \in \mathbb{F}^n \text{ s.t. } Az \circ Bz = Cz \ \& \ z = (u,w) \text{ for some } w \}.$

$$\begin{bmatrix} - a_1 - \\ - a_2 - \\ \vdots \\ - a_m - \end{bmatrix} \begin{bmatrix} | \\ z \\ | \end{bmatrix} \circ \begin{bmatrix} - b_1 - \\ - b_2 - \\ \vdots \\ - b_m - \end{bmatrix} \begin{bmatrix} | \\ z \\ | \end{bmatrix} = \begin{bmatrix} - c_1 - \\ - c_2 - \\ \vdots \\ - c_m - \end{bmatrix} \begin{bmatrix} | \\ z \\ | \end{bmatrix} \quad \text{i.e.} \quad \{ \langle a_i, z\rangle \langle b_i, z\rangle = \langle c_i, z\rangle \}_{i\in[m]}$$

theorem: For "large smooth" $\mathbb{F}$,

$\mathrm{R1CS}(\mathbb{F}) \in \mathrm{IOP}[\varepsilon_c = 0, \varepsilon_s = 0.5, k = O(\log m), \Sigma = \mathbb{F}, \ell = O(m), q = O(\log m), r = O(\log m)]$

This achieves linear-size IOPs for arithmetic computations!

Note: we cannot conclude that all of NP has linear-size proofs because reductions introduce overheads.

Today we assume for simplicity that $m = n$ (# equations = # variables).

Prior Choices of Encoding

Our recipe to construct PCPs so far has been to set $\Pi = (\Pi_a, \Pi_{sat})$ where

① Π_a is (allegedly) the encoding of a candidate assignment [belongs to $S := \{Enc(z)\}_z$]

② if Π_a is close to $Enc(a)$ for some a, Π_{sat} facilitates checking that a is satisfying

① What encodings did we use for an assignment $a : [n] \to \mathbb{F}$?

ⓐ for exp-size PCPs we used linear extensions (aka Hadamard code)

$Enc(a) : \mathbb{F}^n \to \mathbb{F}$ where $Enc(a) := (\langle a, c \rangle)_{c \in \mathbb{F}^n}$

exponential $|Enc| = |\mathbb{F}|^n$

ⓑ for poly-size PCPs we used multivariate low-degree extensions (aka Reed-Muller code)

$Enc(a) : \mathbb{F}^{\log n} \to \mathbb{F}$ where $Enc(a) :=$ "$(\mathbb{F}, \{0,1\}, \log n)$-extension of a" multilinear

almost polynomial $|Enc| = n^{\log|\mathbb{F}|} = n^{O(\log\log n)}$

$Enc(a) : \mathbb{F}^{\frac{\log n}{\log|H|}} \to \mathbb{F}$ where $Enc(a) :=$ "$(\mathbb{F}, H, \frac{\log n}{\log|H|})$-extension of a"

polynomial $|Enc| = n^{\frac{\log|\mathbb{F}|}{\log|H|}} = n^{1+O(\epsilon)}$

Crucially, for ⓐ we have linearity test and for ⓑ we have (multivariate) low-degree test.

② How to test satisfiability? For ⓐ, random combination & tensor test. For ⓑ, use sumcheck for everything.

A New Choice of Encoding

We seek an encoding with $\begin{cases} \bullet \text{ constant rate: } |\mathrm{Enc}(a)| = O(|a|) \\ \bullet \text{ constant relative distance: } a \neq a' \rightarrow \Delta(\mathrm{Enc}(a), \mathrm{Enc}(a')) \geq \Omega(1) \end{cases}$

that lets us execute our recipe of $\Pi = (\Pi_a, \Pi_{sat})$, which in turn means that we need

- a proximity test: "Π_a close to $\{\mathrm{Enc}(z)\}_z$" in few queries
- an approach for testing satisfiability (eg. a replacement for sumcheck protocol)

Satisfying the rate & distance alone is easy (pick any good code over $\mathbb{F}$).

Additionally satisfying the other requirements is hard.

The new encoding that we use is: univariate low-degree extensions

continue to place our hopes in polynomials!

$\mathrm{Enc}(a) : \overset{L}{\cancel{\mathbb{F}}} \rightarrow \mathbb{F}$ where $\mathrm{Enc}(a) :=$ "Univariate extension of $a : H \rightarrow \mathbb{F}$" = "evaluation of $\sum_{i \in H} a(i) L_{i,H}(x)$ on $\overset{L}{\cancel{\mathbb{F}}}$"

Actually we will evaluate on $L = \Theta(|H|)$ rather than $\mathbb{F}$ for more flexibility.

This encoding is also known as the Reed–Solomon code:

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

relative distance is $1 - \frac{d}{|L|} = \Omega(1)$ if $|L| = \Omega(d)$

Today:

we temporarily assume that we have a proximity test for univariate extensions, and show how to use this code to construct linear-size IOPs

Univariate Sumcheck [1/3]

The verifier has oracle access to $f: L \to \mathbb{F}$ s.t. $\deg(\hat{f}) \leq d$ and input $(\overset{\text{field}}{\mathbb{F}}, \underbrace{\overset{\text{evaluation domain}}{L}, \overset{\text{degree}}{d}, \overset{\text{summation domain}}{H}}_{\text{each of these has size } \Omega(n)}, \overset{\text{claimed sum}}{\gamma})$, and wants to check the claim "$\sum_{a \in H} \hat{f}(a) = \gamma$".

Attempt 1: query f at every $a \in H$ and add up the answers

What if $H \cap L = \emptyset$?

Deriving $f(a)$ for a single $a \in H$ requires $d+1 = \Omega(n)$ queries for interpolation.

Even if $H \subseteq L$, $|H| = \Omega(n)$ queries is too many.

[And even if H were small, in the noisy case we would use self-correction, which we don't have.]

Attempt 2: run sumcheck protocol for "$\sum_{a \in H^n} \hat{f}(a) = \gamma$" with $n=1$ (e.g. as IP)

The first (and only) message is the $d+1 = \Omega(n)$ coefficients of $\hat{f}$:

$\xrightarrow{(c_0, c_1, \ldots, c_d)}$ V^f: set $\tilde{f}(x) := \sum_{i=0}^{d} c_i x^i$ and check: $\sum_{a \in H} \tilde{f}(a) = \gamma$ & $\tilde{f}(s) = f(s)$ for random $s \in L$

This is tantamount to reading 1 (huge) symbol from the alphabet $\Sigma = \mathbb{F}^{d+1}$.

We need new ideas!

Univariate Sumcheck [2/3]

The verifier has oracle access to $f : L \to \mathbb{F}$ s.t. $\deg(\hat{f}) \leq d$ and input $(\overset{\text{field}}{\mathbb{F}}, \overset{\text{evaluation domain}}{L}, \overset{\text{degree}}{d}, \overset{\text{summation domain}}{H}, \overset{\text{claimed sum}}{\gamma})$, and wants to check the claim "$\sum_{a\in H} \hat{f}(a) = \gamma$".

Step 1: reduce the problem to the case $d < |H|$

Let $v_H(x) := \prod_{a\in H}(x-a)$ be the vanishing polynomial of the set H.

Divide $\hat{f}(x)$ by $v_H(x)$: $\hat{f}(x) = \hat{h}(x)v_H(x) + \hat{g}(x)$ with $\deg(\hat{g}) < |H|$ & $\deg(\hat{h}) = \deg(\hat{f}) - |H|$

Observe that $\sum_{a\in H} \hat{f}(a) = \sum_{a\in H} \hat{g}(a)$.

Step 2: assume that H is nice and use algebra ← similar to how multivariate sumcheck works for product sets in $\mathbb{F}^n$ rather than all sets

<u>lemma:</u> if H is a subgroup of $\mathbb{F}^*$ then $\sum_{a\in H} \hat{g}(a) = |H|\hat{g}(0)$

<u>proof:</u> First consider a monomial: $\sum_{a\in H} a^i = \sum_{j=0}^{|H|-1} (\omega^j)^i = \sum_{j=0}^{|H|-1} (\omega^i)^j = \begin{cases} 0 & \text{if } i \not\equiv 0 \bmod |H| \\ |H| & \text{if } i \equiv 0 \bmod |H| \end{cases}$.
Hence all monomials $\{x^i\}_{0<i<|H|}$ in $\hat{g}(x)$ sum to zero, and are left with $|H|$ times $g(0)$. ∎

Hence $\sum_{a\in H} \hat{g}(a) = \gamma$ iff $|H|\hat{g}(0) = \gamma$.

[Here we saw the case of multiplicative subgroups.
A similar statement holds for additive subgroups.]

Univariate Sumcheck [3/3]

field, evaluation domain, degree, summation domain, claimed sum

The verifier has oracle access to $f: L \to \mathbb{F}$ s.t. $\deg(\hat{f}) \le d$ and input $(\mathbb{F}, L, d, H, \gamma)$, and wants to check the claim "$\sum_{a \in H} \hat{f}(a) = \gamma$".

$P((\mathbb{F}, L, d, H, \gamma), f)$

Compute $\hat{h}(x)$ with $\deg(\hat{h}) = \deg(\hat{f}) - |H|$ and $\hat{p}(x)$ with $\deg(\hat{p}) < |H| - 1$ s.t.

$$\hat{f}(x) \equiv \hat{h}(x) v_H(x) + \left(x\hat{p}(x) + \gamma/|H|\right)$$

$$\xrightarrow[p: L \to \mathbb{F}]{h: L \to \mathbb{F}}$$

$V^{f: L \to \mathbb{F}}((\mathbb{F}, L, d, H, \gamma))$

- test that h is δ-close to degree $d - |H|$ and that p is δ-close to degree $|H| - 1$
- sample $s \leftarrow L$ and check that

$$f(s) = h(s) \cdot v_H(s) + \left(s\, p(s) + \gamma/|H|\right)$$

Analysis: If $\sum_{a \in H} \hat{f}(a) = \gamma$ then verifier accepts w.p. 1. If $\sum_{a \in H} \hat{f}(a) \neq \gamma$ then distinguish between:

① $\tilde{h}$ or $\tilde{p}$ is δ-far from (respective) low-degree sets → low-degree test accepts w.p. $\le \varepsilon_{LDT}(\delta)$

② $\tilde{h}$ and $\tilde{p}$ both δ-close to (unique) h and p

→ $\hat{f}(x) \not\equiv \hat{h}(x) v_H(x) + \left(x\hat{p}(x) + \gamma/|H|\right)$ so identity test accept w.p. $\le \frac{d}{|L|} + 2\delta$.

[or else $\hat{f}$ would sum to γ]

Checking Linear Equations

field evaluation domain degree summation domain

The verifier has oracle access to $f,g : L \to \mathbb{F}$ of degree $\le d$ and input $(\mathbb{F}, L, d, H, M)$,

H×H matrix over $\mathbb{F}$

and wants to check the claim "$\hat{g}|_H \equiv M \cdot \hat{f}|_H$".

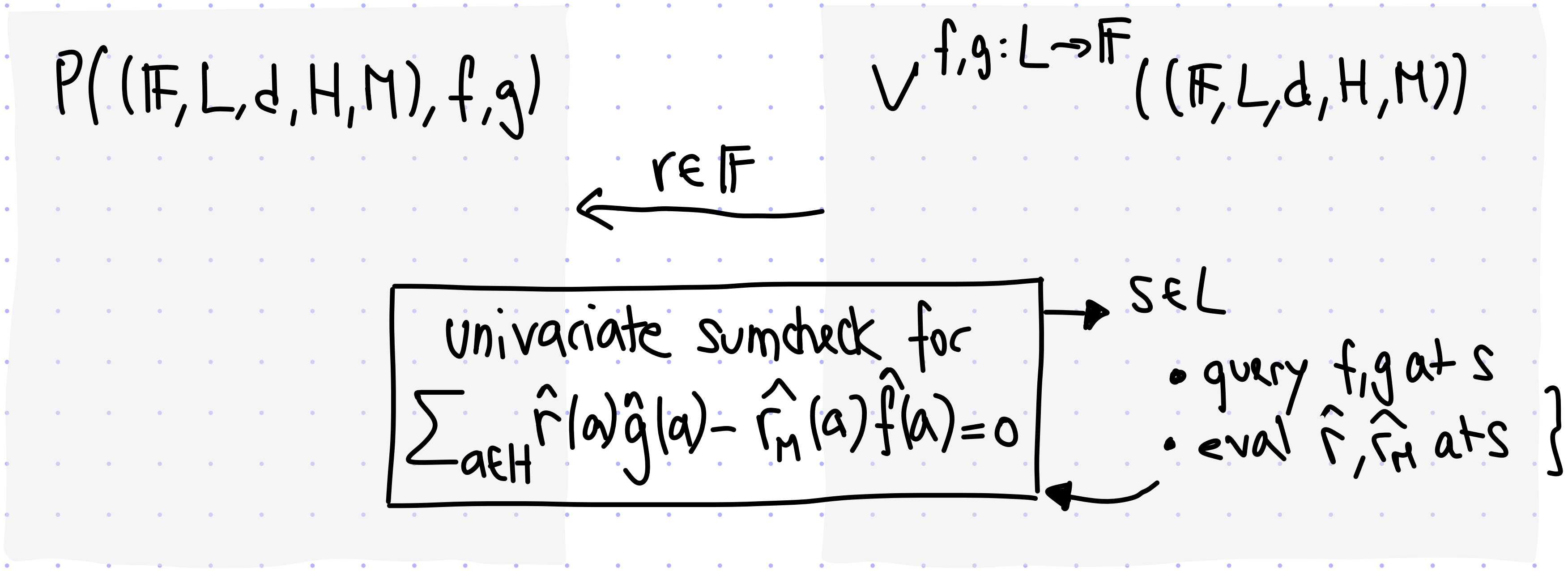
Idea: reduce to a univariate sumcheck claim

bijection from H to $\{0,1,\dots,|H|-1\}$

$$\left\{ \hat{g}(a) = \sum_{b \in H} M[a,b] \cdot \hat{f}(b) \right\}_{a \in H} \quad \text{iff} \quad \sum_{a \in H} \Big(\hat{g}(a) - \sum_{b \in H} M[a,b] \hat{f}(b) \Big) X^{\mathrm{int}(a)} \equiv 0$$

For any $r \in H$, the evaluation at r can be written as:

$$\sum_{a \in H} r^{\mathrm{int}(a)} \hat{g}(a) - \Big(\sum_{b \in H} M[b,a] r^{\mathrm{int}(b)} \Big) \hat{f}(a) \quad , \text{ equivalently } \quad \sum_{a \in H} \underbrace{\hat{r}(a)\hat{g}(a) - \hat{r}_M(a)\hat{f}(a)}_{\text{poly of deg} \le d + |H| - 1}$$



$P((\mathbb{F}, L, d, H, M), f, g)$

$V^{f,g : L \to \mathbb{F}}((\mathbb{F}, L, d, H, M))$

$r \in \mathbb{F}$

univariate sumcheck for $\sum_{a \in H} \hat{r}(a)\hat{g}(a) - \hat{r}_M(a)\hat{f}(a) = 0$

$s \in L$

- query f,g at s
- eval $\hat{r}, \hat{r}_M$ at s

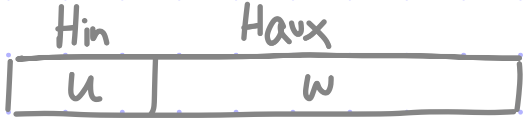
} can be done in $O(|H| + \|M\|)$ ops

The soundness error is

$$\frac{|H|-1}{|\mathbb{F}|} + \varepsilon_{sc}$$

IOP for R1CS: Construction

View H in 2 parts:



$P((u,A,B,C),w)$

Set $z := (u,w) \in \mathbb{F}^n$.

Shift w as follows:

$\forall a \in H_{aux} \quad w'(a) = \frac{w(a) - \hat{u}(a)}{v_{H_{in}}(a)}$.

Compute $f_w := \widehat{w'}|_L$.

For each $M \in \{A,B,C\}$:

compute $f_M := \widehat{Mz}|_L$

Compute

$$\hat{h}(x) := \frac{\widehat{Az}(x)\widehat{Bz}(x) - \widehat{Cz}(x)}{v_H(x)}$$

For each $M \in \{A,B,C\}$:

compute $\hat{g}_M(x)$ and $\hat{h}_M(x)$ s.t.

$$\hat{r}(x)\widehat{Mz}(x) - \hat{r}_M(x)\hat{z}(x) = \hat{h}_M(x)v_H(x) + x\hat{p}_M(x)$$

$$\xrightarrow{f_w,\, f_A,\, f_B,\, f_C,\, h \,:\, L \to \mathbb{F}}$$

$f : L \to \mathbb{F}$ is defined as

$f(a) := f_w(a)\, v_{H_{in}}(a) + \hat{u}(a)$

$$\xleftarrow{r \in \mathbb{F}}$$

For each $M \in \{A,B,C\}$:

univariate sumcheck for

$$\sum_{a \in H} \hat{r}(a)\hat{f}_M(a) - \hat{r}_M(a)\hat{f}(a) = 0$$

$$\xrightarrow{p_M,\, h_M \,:\, L \to \mathbb{F}}$$

$V((u,A,B,C))$

- Sample $s \in L$ at random.
- $f_A(s)f_B(s) - f_C(s) = h(s)\, v_H(s)$
- For each $M \in \{A,B,C\}$:

  $\hat{r}(s)f_M(s) - \hat{r}_M(s)f(s) = h_M(s) \cdot v_H(s) + s \cdot p_M(s)$
- Test that:
  - f_A, f_B, f_C are δ-close to degree $|H|-1$
  - h is δ-close to degree $|H|-2$
  - h_A, h_B, h_C are δ-close to degree $|H|-2$
  - g_A, g_B, g_C are δ-close to degree $|H|-2$

[actually the three sumchecks can be merged into one via random coeffs]

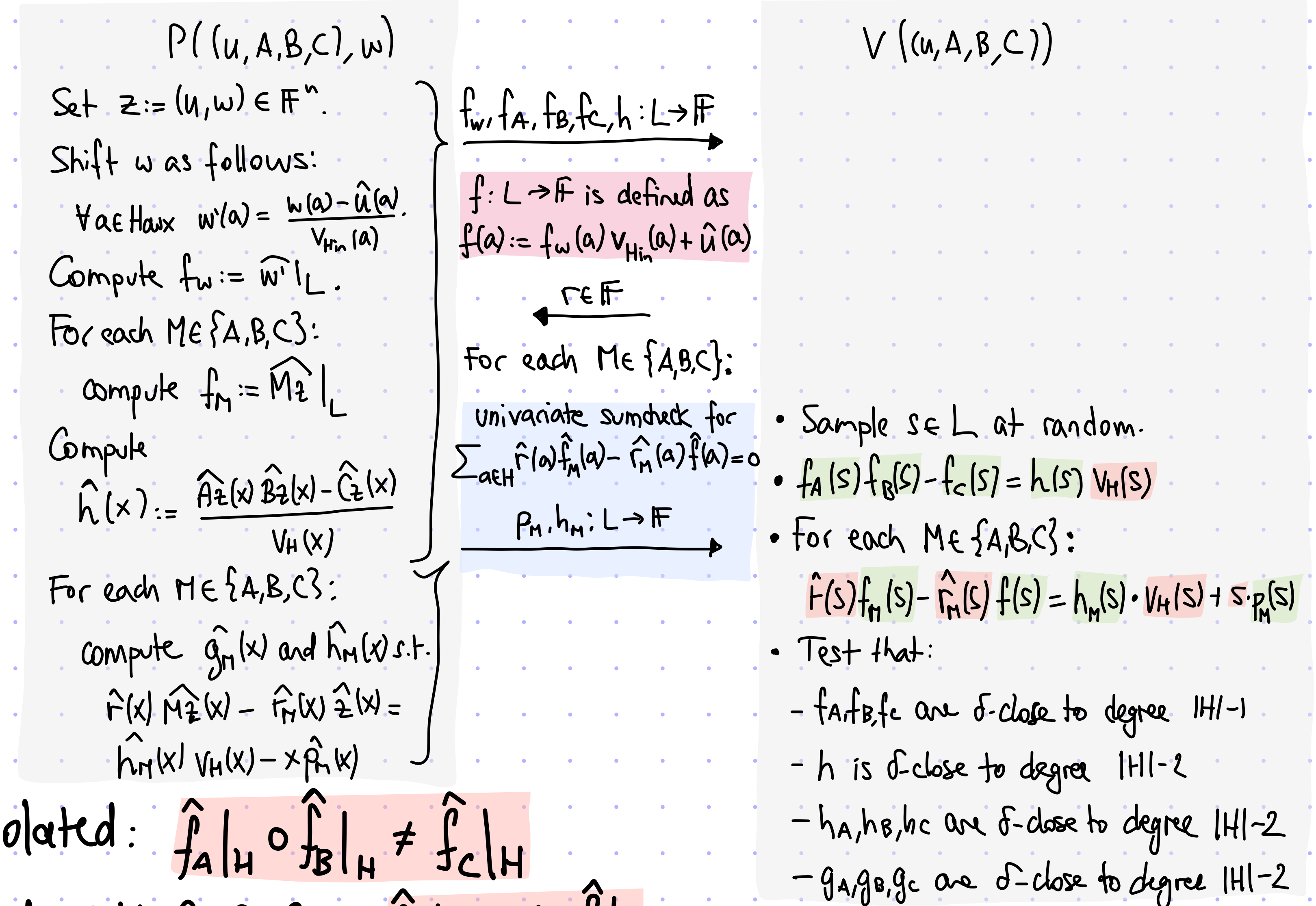
IOP for R1CS: Soundness

Suppose that $(u,A,B,C)\notin$ R1CS.

If any of the sent functions is δ-far then we are done. So suppose all are δ-close.

Let $\hat f_w, \hat f_A, \hat f_B, \hat f_C, \hat h, \hat p_A, \hat h_A, \hat p_B, \hat h_B, \hat p_C, \hat h_C$ be the (unique) closest low-degree polynomials.

One of the following must be true.



$P((u,A,B,C),w)$

Set $z:=(u,w)\in\mathbb{F}^n$.

Shift w as follows:

$\forall a\in H_{aux}\ \ w'(a)=\frac{w(a)-\hat u(a)}{V_{H_{in}}(a)}$.

Compute $f_w:=\widehat{w'}|_L$.

For each $M\in\{A,B,C\}$:

compute $f_M:=\widehat{Mz}|_L$

Compute $\hat h(x):=\frac{\widehat{Az}(x)\widehat{Bz}(x)-\widehat{Cz}(x)}{V_H(x)}$

For each $M\in\{A,B,C\}$:

compute $\hat g_M(x)$ and $\hat h_M(x)$ s.t. $\hat r(x)\widehat{Mz}(x)-\hat r_M(x)\hat z(x)=\hat h_M(x)V_H(x)-x\hat p_M(x)$

$f_w, f_A, f_B, f_C, h: L\to\mathbb{F}$

$f: L\to\mathbb{F}$ is defined as $f(a):=f_w(a)v_{H_{in}}(a)+\hat u(a)$

$r\in\mathbb{F}$

For each $M\in\{A,B,C\}$:

univariate sumcheck for $\sum_{a\in H}\hat r(a)\hat f_M(a)-\hat r_M(a)\hat f(a)=0$

$p_M, h_M: L\to\mathbb{F}$

$V((u,A,B,C))$

- Sample $s\in L$ at random.
- $f_A(s)f_B(s)-f_C(s)=h(s)\,V_H(s)$
- For each $M\in\{A,B,C\}$:

  $\hat r(s)f_M(s)-\hat r_M(s)\,f(s)=h_M(s)\cdot V_H(s)+s\cdot p_M(s)$
- Test that:
  - f_A, f_B, f_C are δ-close to degree $|H|-1$
  - h is δ-close to degree $|H|-2$
  - h_A, h_B, h_C are δ-close to degree $|H|-2$
  - g_A, g_B, g_C are δ-close to degree $|H|-2$

① the Hadamard product condition is violated: $\hat f_A|_H\circ\hat f_B|_H\neq\hat f_C|_H$

② one of the linear conditions is violated: $\exists M\in\{A,B,C\}$ s.t. $\hat f_M|_H\neq M\cdot\hat f|_H$

In case ①: $\hat f_A(X)\cdot\hat f_B(X)-\hat f_C(X)\not\equiv\hat h(X)V_H(X)$ so the verifier accepts w.p. $\leq\frac{2|H|-2}{|L|}+4\delta$

degree in polynomial equation

1 query each to 4 functions that are δ-far from LD

In case ②: except w.p. $\frac{|H|-1}{|\mathbb{F}|}$ over $r\in\mathbb{F}$, $\hat r(X)\hat f_M(X)-\hat r_M(X)\hat f(X)\not\equiv\hat h_M(X)V_H(X)+X\hat p_M(X)$

in which case the verifier accepts w.p. $\leq\frac{2|H|-2}{|L|}+4\delta$.

[Note that input consistency is accounted for: $\hat f(X):=\hat f_w(X)V_{H_{in}}(X)+\hat u(X)$]

IOP for R1CS: Efficiency

- proof complexity (in field elts):

  $O(|L| + \ell_{LDT}) = O(n + \ell_{LDT}) = O(n)$

- query complexity:

  $O(1) + q_{LDT} = O(\log n)$

- round complexity:

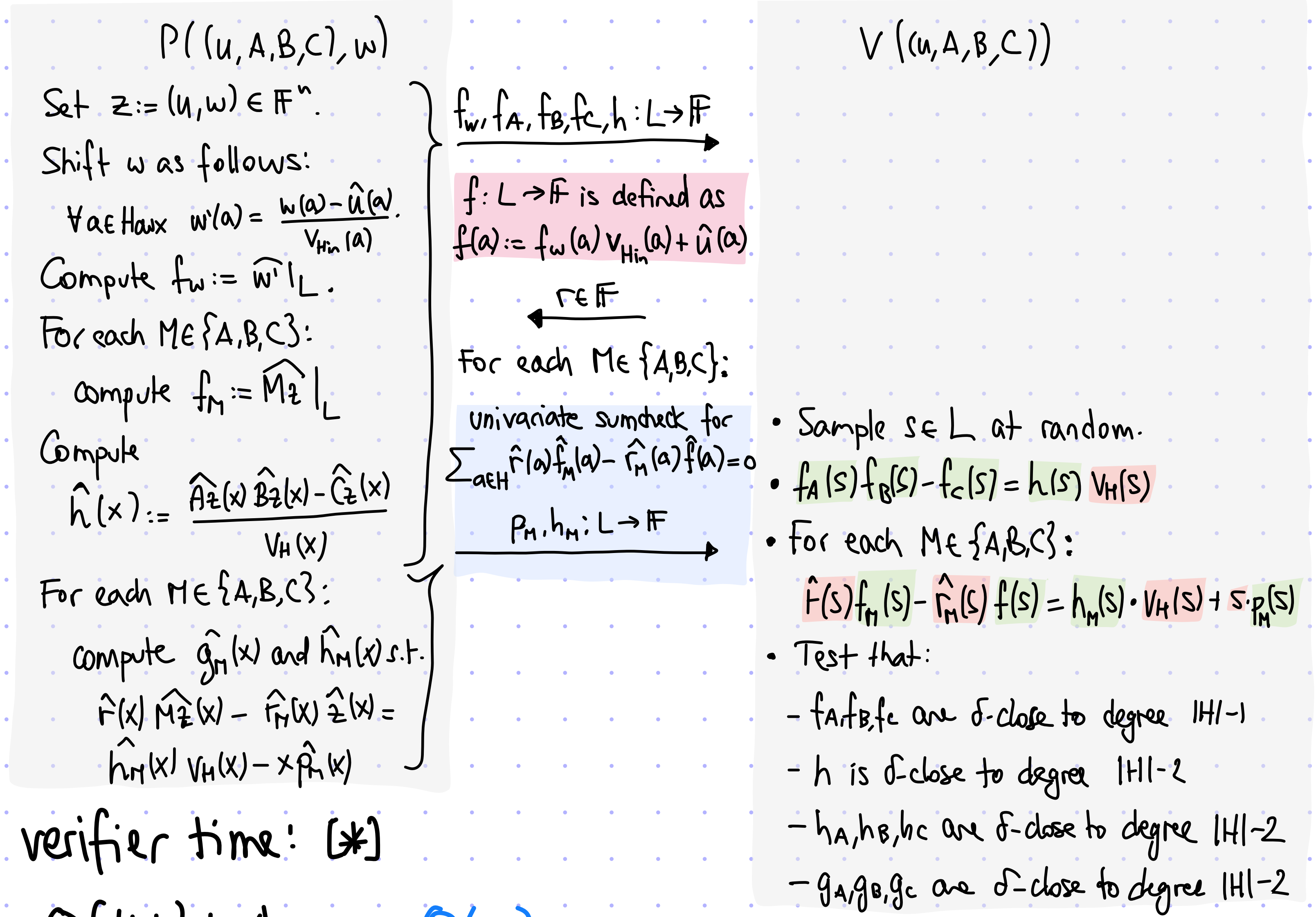
  $O(1) + k_{LDT} = O(\log n)$

- randomness complexity:

  $O(1) + r_{LDT} = O(\log n)$

- prover time: [*]

  $O(|L| \log |L|) + pt_{LDT} = O(n \log n)$



$P((u, A, B, C), w)$

Set $z := (u, w) \in \mathbb{F}^n$.

Shift w as follows:

$\forall a \in H_{aux} \quad w'(a) = \dfrac{w(a) - \hat{u}(a)}{v_{H_{in}}(a)}$.

Compute $f_w := \widehat{w'}|_L$.

For each $M \in \{A, B, C\}$:

compute $f_M := \widehat{Mz}|_L$

Compute

$\hat{h}(x) := \dfrac{\widehat{Az}(x)\widehat{Bz}(x) - \widehat{Cz}(x)}{v_H(x)}$

For each $M \in \{A, B, C\}$:

compute $\hat{g}_M(x)$ and $\hat{h}_M(x)$ s.t.

$\hat{r}(x)\widehat{Mz}(x) - \hat{r}_M(x)\hat{z}(x) = \hat{h}_M(x)v_H(x) - x\hat{p}_M(x)$

$f_w, f_A, f_B, f_C, h : L \to \mathbb{F}$

$f : L \to \mathbb{F}$ is defined as $f(a) := f_w(a) v_{H_{in}}(a) + \hat{u}(a)$

$r \in \mathbb{F}$

For each $M \in \{A, B, C\}$:

univariate sumcheck for $\sum_{a \in H} \hat{r}(a)\hat{f}_M(a) - \hat{r}_M(a)\hat{f}(a) = 0$

$p_M, h_M : L \to \mathbb{F}$

$V((u, A, B, C))$

- Sample $s \in L$ at random.
- $f_A(s) f_B(s) - f_C(s) = h(s)\, v_H(s)$
- For each $M \in \{A, B, C\}$:

  $\hat{r}(s) f_M(s) - \hat{r}_M(s) f(s) = h_M(s) \cdot v_H(s) + s \cdot p_M(s)$
- Test that:
  - f_A, f_B, f_C are δ-close to degree $|H| - 1$
  - h is δ-close to degree $|H| - 2$
  - h_A, h_B, h_C are δ-close to degree $|H| - 2$
  - g_A, g_B, g_C are δ-close to degree $|H| - 2$

- verifier time: [*]

  $O(|L|) + vt_{LDT} = O(n)$

[*]: both pt & vt also include the term $O(\|A\| + \|B\| + \|C\|)$ to multiply vectors by A, B, C

We have constructed IOPs of linear size for R1CS:

theorem: For every field $\mathbb{F}$ of size $\Omega(n)$ that is smooth ← for LDT

$$R1CS(\mathbb{F}) \le IOP\left[\begin{array}{l} \varepsilon_c = 0,\ \varepsilon_s = 0.5,\ \Sigma = \mathbb{F},\ pt = O(n \log n),\ vt = O(n) \\ k = O(\log n),\ r = O(\log n),\ \ell = O(n),\ q = O(\log n) \end{array}\right]$$

We are left to construct a univariate LDT with logarithmically-many queries.