Foundations of Probabilistic Proofs

A course by **Alessandro Chiesa**

Lecture 18

Holographic Proofs

Sublinear Verification for Every Computation

We saw how to achieve sublinear verification via PCPs/IOPs for machine computations.

More generally: sublinear verification is not impossible $\iff$ the computation description is shorter than the computation (the computation is "structured")

The verifier must at minimum read the description of the computation!

Q: How can we achieve sublinear verification for EVERY computation?

(Including ones whose shortest description is the computation itself, like a random circuit.)

An approach: HOLOGRAPHIC PROOFS (a cool-sounding but not so descriptive historical term)

Consider an offline/online model where:

- OFFLINE PHASE: the description of the computation is "encoded" into an oracle.
- ONLINE PHASE: the verifier has query access to this oracle, and may check multiple statements w.r.t. different inputs to this computation.

This model is meaningful for IPs, PCPs, IOPs (as well as variants for robustness, proximity, ...).

TODAY: we formalize this idea and study constructions for it

Indexed Languages and Relations

An indexed language is a set $L = \{(\mathbb{i}, x) \mid \cdots\}$ where $\mathbb{i}$ is an index and x an instance.

EXAMPLE: $\mathrm{CEVAL}(\mathbb{F}) = \{(C, (z_{in}, z_{out})) \mid C : \mathbb{F}^{n_{in}} \to \mathbb{F}^{n_{out}}$ is a circuit and $C(z_{in}) = z_{out}\}$

An indexed relation is a set $R = \{(\mathbb{i}, x, w) \mid \cdots\}$ where $\mathbb{i}$ is an index, x an instance, and w a witness.

The corresponding indexed language is $L(R) = \{(\mathbb{i}, x) \mid \exists\, w \text{ s.t. } (\mathbb{i}, x, w) \in R\}$.

The valid witnesses for the index-instance pair $(\mathbb{i}, x)$ are $R[(\mathbb{i}, x)] = \{w \mid (\mathbb{i}, x, w) \in R\}$.

Why the term "indexed"?
An indexed relation R can be viewed as a collection $\{R_{\mathbb{i}}\}_{\mathbb{i}}$ of standard relations $R_{\mathbb{i}} = \{(x,w) : (\mathbb{i}, x, w) \in R\}$.
Similarly, an indexed language L can be viewed as a collection $\{L_{\mathbb{i}}\}_{\mathbb{i}}$ where $L_{\mathbb{i}} = \{x : (\mathbb{i}, x) \in L\}$.
Hence $\mathbb{i}$ plays the role of an index to elements in the collection.

EXAMPLES:

- circuit satisfiability over $\mathbb{F}$

  $\mathrm{CSAT}(\mathbb{F}) = \{(C, u, w) \mid C : \mathbb{F}^n \to \mathbb{F}$ is a circuit and $C(u, w) = 0\}$

- quadratic equations over $\mathbb{F}$

  $\mathrm{QESAT}(\mathbb{F}) = \{((p_1, \ldots, p_m), u, w) \mid p_1, \ldots, p_m \in \mathbb{F}^{\leq 2}[x_1, \ldots, x_n]$ and $p_1(u, w) = \cdots = p_m(u, w) = 0\}$

- rank-1 constraints over $\mathbb{F}$

  $\mathrm{RICS}(\mathbb{F}) = \{((A, B, C), u, w) \mid A, B, C \in \mathbb{F}^{m \times n}$ and $A \cdot \begin{bmatrix} u \\ w \end{bmatrix} \circ B \cdot \begin{bmatrix} u \\ w \end{bmatrix} = C \cdot \begin{bmatrix} u \\ w \end{bmatrix}\}$

The index $\mathbb{i}$ is to be interpreted as the "large" description of a computation.

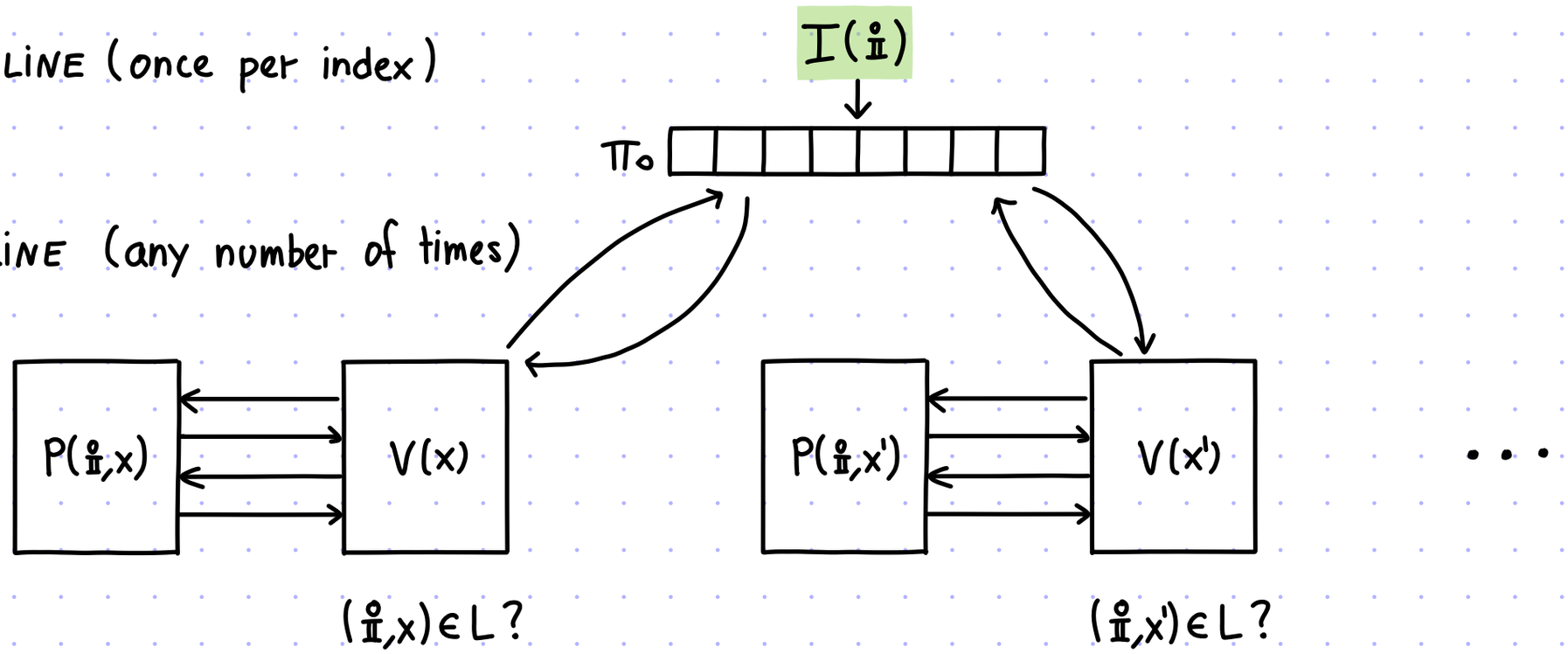
Holographic IPs

A **holographic IP** for an indexed language L is a tuple (I, P, V) s.t.

① **Completeness:** $\forall (\mathbb{i}, x) \in L$, for $\pi_0 := I(\mathbb{i})$, $\Pr[\langle P(\mathbb{i}, x), V^{\pi_0}(x) \rangle = 1] \geq 1 - \varepsilon_c$.

② **Soundness:** $\forall (\mathbb{i}, x) \notin L$, for $\pi_0 := I(\mathbb{i})$, $\forall \tilde{P}$ $\Pr[\langle \tilde{P}, V^{\pi_0}(x) \rangle = 1] \leq \varepsilon_s$.

- OFFLINE (once per index)
- ONLINE (any number of times)



EFFICIENCY MEASURES:

As in an IP except we also study

- proof length ℓ_0 (length of π_0 over an alphabet Σ)
- query complexity q_0 (number of queries by verifier to π_0)

Holographic IPs

The (doubly-efficient) IP for circuit evaluation that we saw can be made holographic.

def: $\mathrm{CEVAL}(\mathbb{F}) = \{ (C, (z_{in}, z_{out})) \mid C: \mathbb{F}^{n_{in}} \to \mathbb{F}^{n_{out}}$ is a circuit and $C(z_{in}) = z_{out} \}$

def: $\mathrm{LCEVAL}(\mathbb{F})$ is the restriction of $\mathrm{CEVAL}(\mathbb{F})$ to layered circuits $C = \{(\mathrm{add}_i, \mathrm{mul}_i)_{i \in [D]}\}$.

Recall that D denotes circuit depth and W circuit width.

theorem:

$$\mathrm{LCEVAL}(\mathbb{F}) \in \mathrm{HIP}\left[\begin{array}{lll} \varepsilon_c = 0 & \kappa = O\left(D \cdot \frac{\log W}{\log |H|}\right) & \mathrm{it} = D \cdot \mathrm{poly}\left(W^{\frac{\log |\mathbb{F}|}{\log |H|}}\right) \\ \varepsilon_s = O\left(D \cdot \frac{\log W \cdot |H|}{\log |H| \cdot |\mathbb{F}|}\right) & cc = O\left(D \cdot \frac{\log W}{\log |H|} \cdot |H|\right) & pt = D \cdot \mathrm{poly}(W) \\ & & vt = D \cdot \mathrm{poly}\left(\frac{\log W}{\log |H|}, |H|\right) + \mathrm{poly}(n_{in}, n_{out}) \end{array}\right]$$

In particular, $\mathrm{it} = D \cdot \mathrm{poly}(W)$ if $|H| = \Theta(D \cdot \log W)$ and $|\mathbb{F}| = |H|^2$.

proof: Consider the GKR bare bones protocol, an IP (P_{GKR}, V_{GKR}) for $\mathrm{LCEVAL}(\mathbb{F})$ where V_{GKR} has query access to the $(\mathbb{F}, H, \frac{\log W}{\log |H|})$-extension $\hat{C} := \{(\widehat{\mathrm{add}}_i, \widehat{\mathrm{mul}}_i)_{i \in [D]}\}$ of $C = \{(\mathrm{add}_i, \mathrm{mul}_i)_{i \in [D]}\}$.

Consider the holographic IP (I, P, V) where $P := P_{GKR}$, $V := V_{GKR}$, and

$$I\left(\mathring{\Pi} = \{(\mathrm{add}_i, \mathrm{mul}_i)_{i \in [D]}\}\right) := \text{output } \pi_0 := \{(\widehat{\mathrm{add}}_i, \widehat{\mathrm{mul}}_i)_{i \in [D]}\}$$

Each LDE is a function over $\mathbb{F}^{\frac{\log W}{\log |H|}}$ and so takes $\mathrm{poly}(|\mathbb{F}|^{\frac{\log W}{\log |H|}}) = \mathrm{poly}(W^{\frac{\log |\mathbb{F}|}{\log |H|}})$ to write down.

In $IP = (P_{GKR}, V_{GKR})$ we set $H = \{0,1\}$. Here we need $\frac{\log |\mathbb{F}|}{\log |H|} = O(1)$ and $D \cdot \frac{\log W \cdot |H|}{\log |H| \cdot |\mathbb{F}|} = O(1)$.

E.g. $\forall \mu > 0$ if $|H| = (D \cdot \log W)^{\frac{1}{\mu}}$ and $|\mathbb{F}| = |H|^{1+\mu}$ then $\frac{\log |\mathbb{F}|}{\log |H|} = 1 + \mu$ and $D \cdot \frac{\log W \cdot |H|}{\log |H| \cdot |\mathbb{F}|} = \frac{|H|^{\mu} \cdot |H|}{\log |H| \cdot |H|^{1+\mu}} = \frac{1}{\log |H|}$. ■

Holographic PCPs

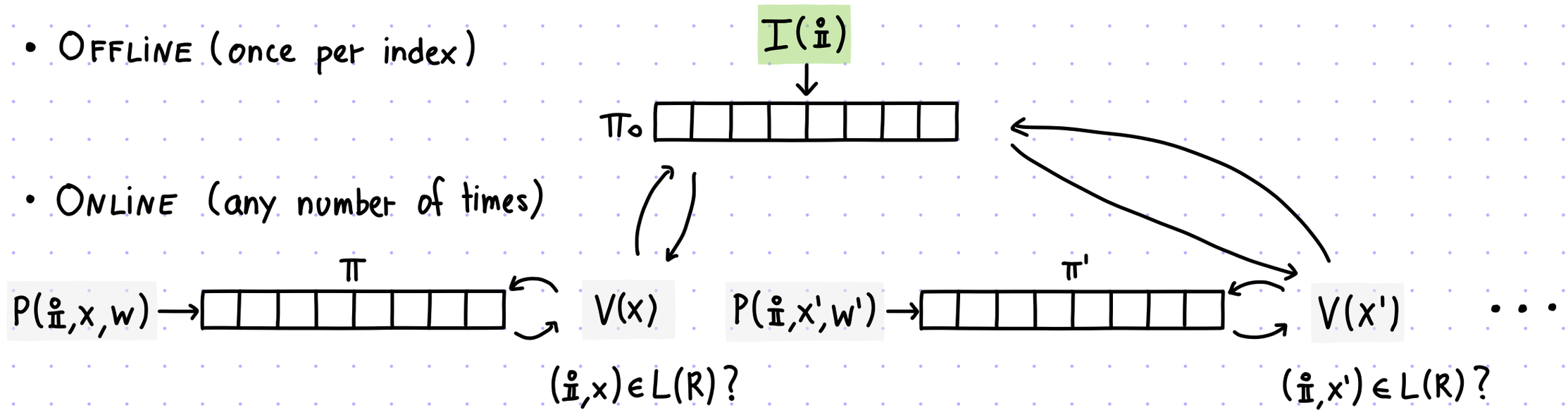
[The definition for an indexed language L is a special case.]

A **holographic PCP** for an indexed relation R is a tuple (I,P,V) s.t.

① **Completeness:** $\forall (\mathbb{i},x,w)\in R$, for $\pi_0 := I(\mathbb{i})$ and $\pi := P(\mathbb{i},x,w)$, $\Pr_\rho[V^{\pi_0,\pi}(x;\rho)=1] \geq 1-\varepsilon_c$.

② **Soundness:** $\forall (\mathbb{i},x)\notin L(R)$, for $\pi_0 := I(\mathbb{i})$, $\forall \tilde{\pi}$ $\Pr_\rho[V^{\pi_0,\tilde{\pi}}(x;\rho)=1] \leq \varepsilon_s$.

- OFFLINE (once per index)
- ONLINE (any number of times)



EFFICIENCY MEASURES:

As in a PCP except that

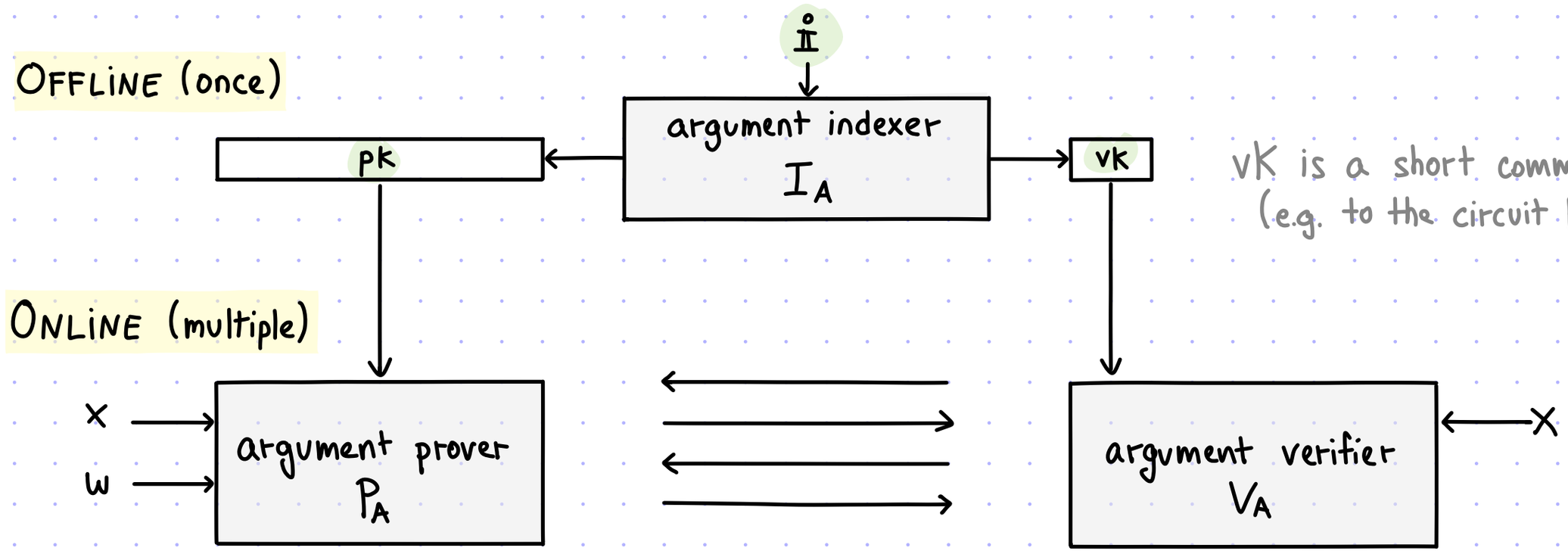
- proof length is $|\pi_0|+|\pi|$ (over an alphabet Σ)
- query complexity is q_0+q (number of queries by verifier to π_0 and π)

From Holography to Preprocessing [1/2]

A motivation to study holography is that it leads to **PREPROCESSING ARGUMENTS**.

These enable **sublinear verification for ANY computation**, given a **one-time (public) preprocessing step**.

A preprocessing argument system for an indexed relation R works as follows:



In the past we saw: PCP (or IOP) + CRH ⟶ succinct argument (e.g. Kilian's protocol)

Now we see: PCP (or IOP) + CRH ⟶ succinct argument (an extension of Kilian's protocol)

holographic

preprocessing

From Holography to Preprocessing [2/2]

SETUP: Everyone has access to a collision-resistant function h (sampled from a family H_λ).

OFFLINE: Anyone can compute the key pair for an index $\mathbb{I}$ (re-usable any number of times):

$\mathcal{I}_A(h, \mathbb{I})$: 1. Compute the encoded index: $\Pi_0 := \mathcal{I}(\mathbb{I})$.
2. Commit to encoded index: $(rt_0, aux_0) := MT[h].Commit(\Pi_0)$.
3. Output key pair $(pk, vk) := ((h, \mathbb{I}, \Pi_0), (h, rt_0))$.

ONLINE: Anyone can use the key pair to prove/verify statements of the form $(\mathbb{I}, x) \in L$:

$P_A(pk, x, w)$

Compute PCP string: $\pi := P(\mathbb{I}, x, w)$

Commit to PCP string: $(rt, aux) := MT[h].Commit(\pi)$.

$\xrightarrow{rt}$

$\xleftarrow{\rho}$

Deduce query sets $Q_0 \subseteq [\ell_0]$, $Q \subseteq [\ell]$ for $V^{\Pi_0, \Pi}(x; \rho)$.

Set answers: $a_0 := \Pi_0[Q_0] \in \Sigma^{Q_0}$, $a := \pi[Q] \in \Sigma^{Q}$.

Authenticate answers: $pf_0 := MT[h].Open(aux_0, Q_0)$

$pf := MT[h].Open(aux, Q)$

$\xrightarrow{Q_0, a_0, pf_0 \;\; Q, a, pf}$

$V_A(vk, x)$

Sample PCP randomness $\rho \leftarrow \{0,1\}^r$.

$V_{PCP}^{[Q_0, a_0], [Q, a]}(x; \rho) \stackrel{?}{=} 1$

$MT[h].Check(rt_0, Q_0, a_0, pf_0) \stackrel{?}{=} 1$

$MT[h].Check(rt, Q, a, pf) \stackrel{?}{=} 1$

$time(P_A) = time(P) + O_\lambda(\ell)$

$O_\lambda(q \cdot \log \ell)$

$time(V_A) = time(V) + O_\lambda(q \cdot \log \ell)$

Holographic PCP for NP [1/2]

We proved that NP has PCPs with polynomial proof length and polylogarithmic query complexity. The PCP verifier does not (and cannot) run in sublinear time because it reads the NP statement being proved (in that case, the list of quadratic equations).

We show how to achieve sublinear verification time via an HPCP (i.e., with the help of an indexer):

def: $QESAT(\mathbb{F}) = \{ ((p_1,\dots,p_m),u,w) \mid p_1,\dots,p_m \in \mathbb{F}^{\leq 2}[X_1,\dots,X_n] \text{ and } p_1(u,w)=\dots=p_m(u,w)=0 \}$

theorem: $QESAT(\mathbb{F}) \in HPCP\left[\begin{array}{lll} \varepsilon_c = 0 & \Sigma = \mathbb{F} & vt = poly(|u|,\log n) \\ \varepsilon_s = O(1) + O\left(\frac{\log^2 n}{\log\log n}\cdot\frac{1}{|\mathbb{F}|}\right) & \ell = |\mathbb{F}|^{O\left(\frac{\log n}{\log\log n}\right)} & q = poly(\log n) \end{array}\right]$

Via the HOLOGRAPHY → PREPROCESSING connection, we get a preprocessing succinct argument for NP where: $time(I_A) = poly(\lambda,|\mathbb{I}|)$, $time(P_A) = poly(\lambda,|\mathbb{I}|,|x|,|w|)$, $time(V_A) = poly(\lambda,|x|)$.

The ability to verify in sublinear time ANY (even unstructured) NP computation is useful in applications (e.g. it simplifies the recursive use of succinct arguments).

We prove the theorem by modifying the (non-holographic) PCP for $QESAT(\mathbb{F})$.

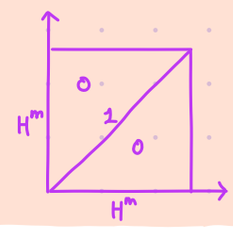
Holographic PCP for NP

Fix $H_v, H_e \subseteq \mathbb{F}$ of sizes $O(\log n), O(\log m)$ and set $s_v := \frac{\log n}{\log |H_v|}$ and $s_e := \frac{\log m}{\log |H_e|}$.

$\mathcal{I}(\mathring{\Pi} = (p_1, \dots, p_m))$:

1. $\hat{p}(X_1, \dots, X_{s_e}) := \sum_{0 \le j_1, \dots, j_{s_e} < |H_e|} X_1^{j_1} \cdots X_{s_e}^{j_{s_e}} \cdot p_{j_1 \cdots j_{s_e}}$.
2. $\hat{c}(X_1, \dots, X_{s_e}, Y_1, \dots, Y_{s_v}, Z_1, \dots, Z_{s_v}) := \sum_{a,b \in H_v^{s_v}} \hat{p}(X_1, \dots, X_{s_v})[a,b] \cdot \mathcal{I}(a,Y) \cdot \mathcal{I}(b,Z)$.
3. Output $\pi_0 : \mathbb{F}^{s_e + 2s_v} \to \mathbb{F}$ where $\pi_0 :=$ "evaluation of $\hat{c}$".

$\mathcal{I}_{H^m}(X,Y) := \prod_{i \in [m]} \sum_{\alpha \in H} L_{H,\alpha}(X_i) \cdot L_{H,\alpha}(Y_i)$.



$P(\mathring{\Pi} = (p_1, \dots, p_m), u, w)$

1. Set $a := (u, w) \in \mathbb{F}^n$.
2. For every $\sigma \in \mathbb{F}^{s_e}$:
   - $p_\sigma := \sum_{0 \le j_1, \dots, j_{s_e} < |H_e|} \sigma_1^{j_1} \cdots \sigma_{s_e}^{j_{s_e}} \cdot p_{j_1 \cdots j_{s_e}}$
   - $\pi_{sc}[\sigma] :=$ eval table for sumcheck claim "$p_\sigma(a) = 0$"
   - output $\pi_{sc}[\sigma]$
3. Output $\pi_a : \mathbb{F}^{s_v} \to \mathbb{F}$ that is the $(\mathbb{F}, H_v, s_v)$-extension of a.
4. Output π_{Ic} that proves that π_a is consistent with u.

$V(\mathring{\Pi} = (p_1, \dots, p_m), u)$

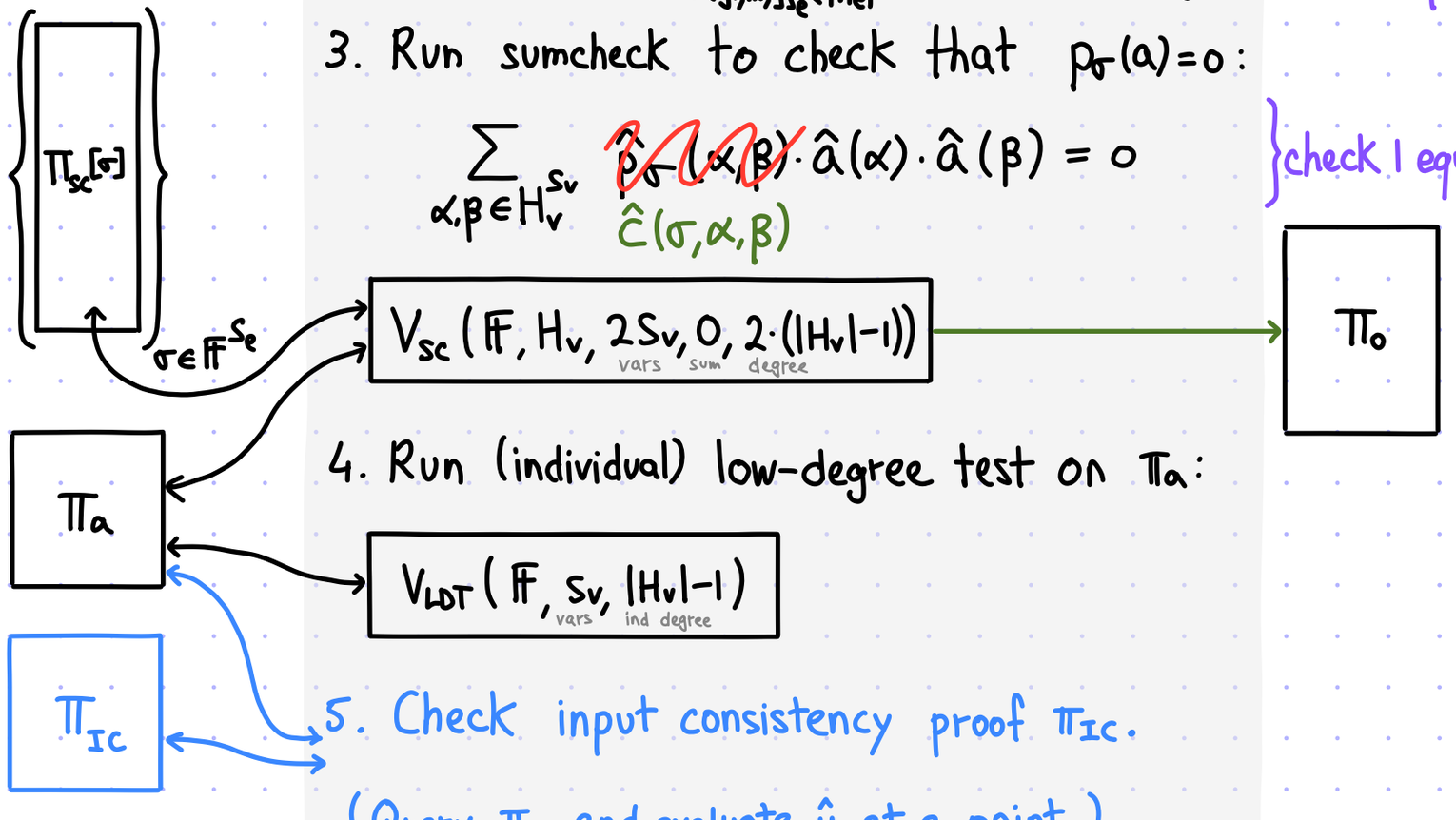
1. Sample $\sigma \in \mathbb{F}^{s_e}$.
2. Compute $p_\sigma := \sum_{0 \le j_1, \dots, j_{s_e} < |H_e|} \sigma_1^{j_1} \cdots \sigma_{s_e}^{j_{s_e}} \cdot p_{j_1 \cdots j_{s_e}}$. } m equations to 1 equation
3. Run sumcheck to check that $p_\sigma(a) = 0$:

$$\sum_{\alpha, \beta \in H_v^{s_v}} \underbrace{\hat{p}_\sigma(\alpha, \beta)}_{\hat{c}(\sigma, \alpha, \beta)} \cdot \hat{a}(\alpha) \cdot \hat{a}(\beta) = 0$$

} check 1 equation

4. Run (individual) low-degree test on π_a:
5. Check input consistency proof π_{Ic}.

(Query π_a and evaluate $\hat{u}$ at a point.)



Holographic IOP for NP

We obtained holographic PCPs for NP with polynomial proof length and polylogarithmic time (and query) complexity.

Q: can we improve the (offline and online) proof length via holographic IOPs?

YES, but we will need some new ideas!

def: $\mathrm{RICS}(\mathbb{F}) = \{((A,B,C),u,w) \mid A,B,C \in \mathbb{F}^{m \times n}$ and $A \cdot \begin{bmatrix} u \\ w \end{bmatrix} \circ B \cdot \begin{bmatrix} u \\ w \end{bmatrix} = C \cdot \begin{bmatrix} u \\ w \end{bmatrix}\}$

def: s := # of non-zero entries in A,B,C.

theorem: For every field $\mathbb{F}$ of size $\Omega(s)$ that is smooth (for FRI),

$$\mathrm{RICS}(\mathbb{F}) \in \mathrm{HIOP}\left[\begin{array}{lllll} \varepsilon_c = 0 & k = O(\log s) & \Sigma = \mathbb{F} & r = O(\log s) & it, pt = O(s \log s) \\ \varepsilon_s = 1/2 & & \ell = O(s) & q = O(\log s) & vt = O(|u| + \log s) \end{array}\right]$$

- STARTING POINT of the proof: the IOP for RICS that we constructed.

$$\forall \text{ smooth field } \mathbb{F} \text{ of size } \Omega(n),\ \mathrm{RICS}(\mathbb{F}) \in \mathrm{IOP}\left[\begin{array}{lllll} \varepsilon_c = 0 & k = O(\log n) & \Sigma = \mathbb{F} & r = O(\log n) & pt = O(s + n \log n) \\ \varepsilon_s = 1/2 & & \ell = O(n) & q = O(\log n) & vt = O(s) \end{array}\right]$$

- THEN: replace a computation of the verifier that involves A,B,C with a holographic subprotocol where the indexer and prover help the verifier.

Recall: IOP for R1CS

View H in 2 parts:

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

- $\forall M \in \{A,B,C\}$: $\hat{f}_M(x) := \widehat{M \cdot \begin{bmatrix} u \\ w \end{bmatrix}}(x)$
- $\hat{h}(x) := \dfrac{\hat{f}_A(x) \cdot \hat{f}_B(x) - \hat{f}_C(x)}{V_H(x)}$
- shift the witness:

  $\hat{f}_w(x) :=$ "LDE of $w_* : H_{aux} \to \mathbb{F}$"

  where $w_*(a) := \dfrac{w(a) - \hat{u}(a)}{V_{H_{in}}(a)}$

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

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

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

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

univariate sumcheck for $\sum_{a \in H} \widehat{pow(\sigma)}(a) \cdot \hat{f}_M(a) - (\widehat{M^T pow(\sigma)})(a) \cdot \hat{f}(a) = 0$

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

- $\forall M \in \{A,B,C\}$: compute $\hat{p}_M \& \hat{h}_M$

  $\widehat{pow(\sigma)}(x) \cdot \hat{f}_M(x) - (\widehat{M^T pow(\sigma)})(x) \cdot \hat{f}(x) \equiv \hat{h}_M(x) \cdot V_H(x) + x \cdot \hat{p}_M(x)$

TODO!

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

holographic lincheck to show $\hat{f}_M|_H \equiv M \cdot \hat{f}|_H$

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

- Sample $\sigma \leftarrow \mathbb{F}$.
- Sample $s \leftarrow L$ and check that:

  $f_A(s) \cdot f_B(s) - f_C(s) \stackrel{?}{=} h(s) \cdot V_H(s)$

  $\forall M \in \{A,B,C\}$:

  $\widehat{pow(\sigma)}(s) \cdot f_M(s) - (\widehat{M^T pow(\sigma)})(s) \cdot f(s) \stackrel{?}{=} h_M(s) \cdot V_H(s) + s \cdot p_M(s)$

- Low-degree tests:

  $V_{LDT}^{f_w}(\mathbb{F}, L, |H|-|u|) \stackrel{?}{=} 1$ $V_{LDT}^{h}(\mathbb{F}, L, |H|-1) \stackrel{?}{=} 1$

  $\forall M \in \{A,B,C\}$: $V_{LDT}^{f_M}(\mathbb{F}, L, |H|) \stackrel{?}{=} 1$

  $V_{LDT}^{h_M}(\mathbb{F}, L, |H|-1) \stackrel{?}{=} 1$

  $V_{LDT}^{p_M}(\mathbb{F}, L, |H|-1) \stackrel{?}{=} 1$

Recall: Non-Holographic Lincheck

Suppose that $f,g: L \to \mathbb{F}$ are δ-close to $\hat{f},\hat{g}$ of degree $< d$.

Check that $\hat{g}|_H \equiv M \cdot \hat{f}|_H$. Define $s :=$ "number of non-zero entries in M".

Let $R(X,Y) \in \mathbb{F}[X,Y]$ have individual degree $< |H|$ s.t. $\{R(X,a)\}_{a \in H}$ are linearly independent.

Let $\hat{M}(X,Y)$ be the (bivariate) low-degree extension of the matrix $M: H \times H \to \mathbb{F}$ (viewed as a bivariate function).

Define $R_M(X,Y) := \sum_{a \in H} R(X,a) \cdot \hat{M}(a,Y)$.

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

Compute $\hat{h},\hat{g}$ s.t.

$$R(\alpha,Y)\cdot\hat{g}(Y) - R_M(\alpha,Y)\cdot\hat{f}(Y) \equiv \hat{h}(Y)\cdot V_H(Y) + Y\cdot\hat{p}(Y)$$

$\sum_{a \in H} R(\alpha,a)\cdot\hat{g}(a) - R_M(\alpha,a)\cdot\hat{f}(a) = 0$

$\xleftarrow{\alpha}$

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

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

Sample $\alpha \leftarrow \mathbb{F}$.

Test that h,p are close to low degree:

$V^h_{LDT}(\mathbb{F},L,d-|H|) \stackrel{?}{=} 1$, $V^p_{LDT}(\mathbb{F},L,|H|-1) \stackrel{?}{=} 1$.

Sample $\beta \in L$ and check that

$R(\alpha,\beta)\cdot g(\beta) - R_M(\alpha,\beta)\cdot f(\beta) = h(\beta)\cdot V_H(\beta) + \beta\cdot p(\beta)$.

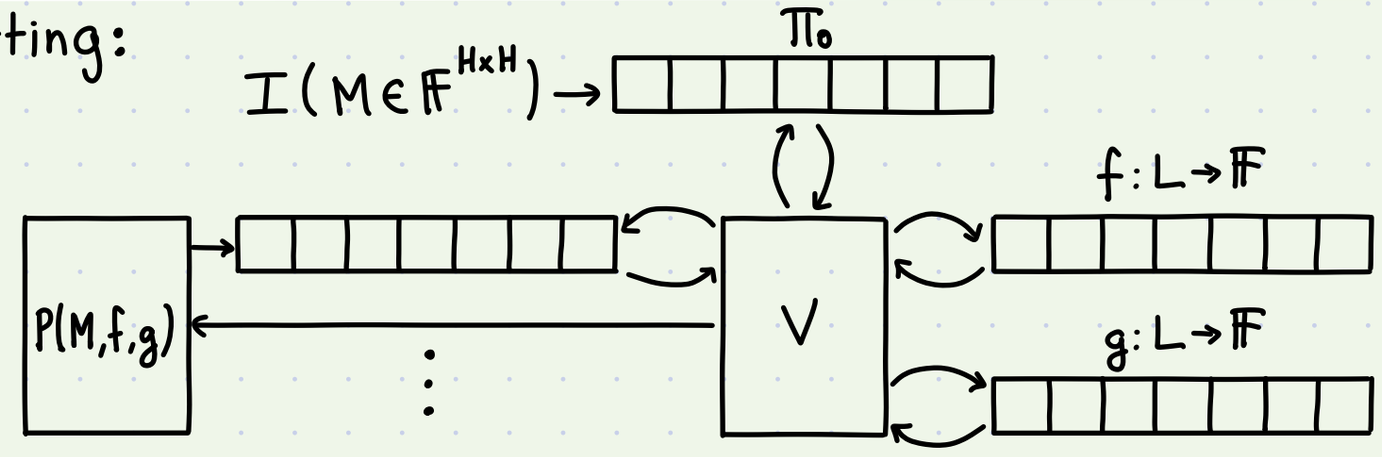
EXPENSIVE: evaluating R and R_M at (α,β) $\longrightarrow$

Fact (it underlies the univariate sumcheck protocol): Let $S \subseteq \mathbb{F}$ be a multiplicative subgroup.

Then $\sum_{a \in S} \hat{f}(a) = \sigma \leftrightarrow \exists$ $\hat{h}(x)$ of degree $\deg(\hat{f})-|S|$, $\hat{p}(x)$ of degree $|S|-2$ s.t. $\hat{f}(x) \equiv \hat{h}(x)\cdot V_S(x) + x\cdot\hat{p}(x) + \sigma/|S|$.

Holographic Lincheck

Setting:



Goal:

Suppose that f, g are δ-close to $\hat{f}, \hat{g}$ of degree $< d$.

Check that $\hat{g}|_H \equiv M \cdot \hat{f}|_H$.

(M: s non-zero entries)

Approach: we choose $R(X,Y)$ such that

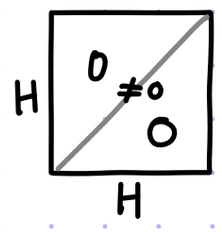
① $R(X,Y)$ is cheap to evaluate at any $(\alpha, \beta) \in \mathbb{F}^2$,

② design a holographic IOP for claims of the form "$R_M(\alpha,\beta) = \tau$".

Previously: $R(X,Y) = \sum_{a \in H} X^a \cdot L_{H,a}(Y)$. Today: $R(X,Y) = \frac{V_H(X) - V_H(Y)}{X - Y}$.

Observe that $\{R(X,a)\}_{a \in H} = \left\{\frac{V_H(X)}{X-a}\right\}_{a \in H}$ equal $\{L_{H,a}(X)\}_{a \in H}$ up to constants.

Hence:
- $\{R(X,a)\}_{a \in H}$ are linearly independent
- $\forall\, a,b \in H$: if $a \neq b$ then $R(a,b) = 0$ else if $a = b$ then $R(a,b) \neq 0$



Moreover, if H is a multiplicative subgroup: $R(X,Y) = \frac{X^{|H|} - Y^{|H|}}{X - Y}$, $R(x,x) = |H| \cdot X^{|H|-1}$.

In this case R is cheap to evaluate!

Holographic Lincheck

A first attempt:

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

1. Sumcheck for $\sum_{a\in H} R(\alpha,a)\cdot\hat{g}(a)-R_M(\alpha,a)\cdot\hat{f}(a)=0$:

   compute $\hat{h},\hat{g}$ s.t. $R(\alpha,y)\cdot\hat{g}(y)-R_M(\alpha,y)\cdot\hat{f}(y) \equiv \hat{h}(y)\cdot V_H(y)+y\cdot\hat{p}(y)$

2. $\tau := R_M(\alpha,\beta)=\sum_{a\in H} R(\alpha,a)\cdot\hat{M}(a,\beta)$.

3. Sumcheck for $\sum_{a\in H} R(\alpha,a)\cdot\hat{M}(a,\beta)=\tau$:

   compute $\hat{h}_2,\hat{p}_2$ s.t. $R(\alpha,x)\cdot\hat{M}(x,\beta) \equiv \hat{h}_2(x)\cdot V_H(x)+x\cdot\hat{p}_2(x)+\tau/|H|$.

α ←

$h,p: L\to\mathbb{F}$ →

β ←

τ →

$h_2,p_2: L\to\mathbb{F}$ →

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

1. Sample $\alpha\leftarrow\mathbb{F}$.
2. $V^{h}_{LDT}(\mathbb{F},L,d-|H|)\stackrel{?}{=}1$, $V^{p}_{LDT}(\mathbb{F},L,|H|-1)\stackrel{?}{=}1$.
3. Sample $\beta\leftarrow L$.
4. Check that

   $R(\alpha,\beta)\cdot g(\beta)-\tau\cdot f(\beta)=h(\beta)\cdot V_H(\beta)+\beta\cdot p(\beta)$.

5. $V^{h_2}_{LDT}(\mathbb{F},L,|H|-1)\stackrel{?}{=}1$, $V^{p_2}_{LDT}(\mathbb{F},L,|H|-1)\stackrel{?}{=}1$.
6. Sample $\gamma\in L$ and check that

   $R(\alpha,\gamma)\cdot\hat{M}(\gamma,\beta)=h_2(\gamma)\cdot V_H(\gamma)+\gamma\cdot p_2(\gamma)+\tau/|H|$.

We define the indexer as:

$\mathcal{I}(M)$:= output the LDE $\hat{M}: L\times L\to\mathbb{F}$

The online proof length is $O(|L|)$ but the offline proof length is $O(|L|^2)$ regardless of M's sparsity.

Fact (it underlies the univariate sumcheck protocol): Let $S\subseteq\mathbb{F}$ be a multiplicative subgroup.

Then $\sum_{a\in S}\hat{f}(a)=\sigma \leftrightarrow \exists\ \hat{h}(x)$ of degree $\deg(\hat{f})-|S|$, $\hat{p}(x)$ of degree $|S|-2$ s.t. $\hat{f}(x)\equiv\hat{h}(x)\cdot V_S(x)+x\cdot\hat{p}(x)+\sigma/|S|$.

Holographic Lincheck [3/4]

We describe the bivariate LDE $\hat{M}(X,Y)$ in terms of the s non-zero entries of M.

Let $K \subseteq \mathbb{F}$ be of size s.

def: The sparse representation of M is $(\text{row}, \text{col}, \text{val}: K \to \mathbb{F})$ where

$$\forall a \in K, \quad \text{val}(a) = M[\text{row}(a), \text{col}(a)].$$

Note that row, col output values in $H \subseteq \mathbb{F}$ because $M: H \times H \to \mathbb{F}$.

We can write $\hat{M}(X,Y)$ in terms of $\hat{\text{row}}(x), \hat{\text{col}}(x), \hat{\text{val}}(x)$:

$$\hat{M}(X,Y) \equiv \sum_{a \in K} R(X, \hat{\text{row}}(a)) \cdot R(Y, \hat{\text{col}}(a)) \cdot \frac{\hat{\text{val}}(a)}{R(\text{row}(a), \text{row}(a)) \cdot R(\text{col}(a), \text{col}(a))}.$$

$\hat{\text{val}}(a)$ → $\widehat{\text{val}^*}(a)$

Idea: The indexer outputs $\hat{\text{row}}, \hat{\text{col}}, \widehat{\text{val}^*}$. Then check the claim "$\hat{M}(\gamma,\beta) = \sigma$" via a univariate sumcheck.

Problem: the degree of the addend is $\Omega(|H| \cdot |K|) = \Omega(n \cdot s)$ (it is quadratic).

Solution: Using $R(X,Y) = \dfrac{V_H(X) - V_H(Y)}{X - Y}$, we get

$$\hat{M}(X,Y) \equiv \sum_{a \in K} \frac{V_H(X)}{X - \hat{\text{row}}(a)} \cdot \frac{V_H(Y)}{Y - \hat{\text{col}}(a)} \cdot \widehat{\text{val}^*}(a).$$

← The numerator and denominator have degree $< |K| = s$.

We need a sumcheck for univariate RATIONAL functions!

Sumcheck for Univariate Rational Functions

Given $f, g: L \to \mathbb{F}$ that are δ-close to $\hat{f}, \hat{g}$ of degree $< d$ and given $\sigma \in \mathbb{F}$, check that $\sum_{a \in H} \hat{f}(a)/\hat{g}(a) = \sigma$. (And assume that $\hat{g}(a) \neq 0 \ \forall a \in H$.)

We extend the sumcheck protocol for univariate polynomials.

Define the function $u: H \to \mathbb{F}$ as $u(a) := \hat{f}(a)/\hat{g}(a)$. Note that $\deg(\hat{u}) < |H|$.

Observe that:
- $\hat{u}$ agrees with $\hat{f}/\hat{g}$ on $H \leftrightarrow \exists\, \hat{h}$ s.t. $\hat{f}(x) - \hat{g}(x) \cdot \hat{u}(x) \equiv \hat{h}(x) \cdot V_H(x)$
- if H is a multiplicative subgroup, by the univariate sumcheck:
  $\sum_{a \in H} \hat{u}(a) = \sigma \leftrightarrow \exists\, \hat{p}$ of degree $< |H|-1$ s.t. $\hat{u}(x) \equiv x \cdot \hat{p}(x) + \sigma/|H|$

We deduce that:

lemma: If H is a multiplicative subgroup,

$$\sum_{a \in H} \frac{\hat{f}(a)}{\hat{g}(a)} = \sigma \leftrightarrow \exists \begin{matrix} \hat{h} \text{ of degree } < d-1 \\ \hat{p} \text{ of degree } < |H|-1 \end{matrix} \text{ s.t. } \hat{f}(x) - \hat{g}(x) \cdot (x \cdot \hat{p}(x) + \sigma/|H|) \equiv \hat{h}(x) \cdot V_H(x)$$

The lemma immediately leads to a protocol (which tests this polynomial identity).

Holographic Lincheck [4/4]

$\mathcal{I}(M=(\text{row},\text{col},\text{val})) :=$ output the LDEs $\hat{\text{row}}, \hat{\text{col}}, \hat{\text{val}} : L \to \mathbb{F}$

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

Compute $\hat{h},\hat{p}$ s.t. $R(\alpha,\gamma)\cdot\hat{g}(\gamma) - R_M(\alpha,\gamma)\hat{f}(\gamma) \equiv \hat{h}(\gamma)\cdot v_H(\gamma) + \gamma\cdot\hat{p}(\gamma)$

$\Sigma_{a\in H} R(\alpha,a)\cdot\hat{g}(a) - R_M(\alpha,a)\cdot\hat{f}(a) = 0$

Compute $\tau := R_M(\alpha,\beta) = \Sigma_{a\in H} R(\alpha,a)\cdot\hat{M}(a,\beta)$

Compute $\hat{h}_2,\hat{p}_2$ s.t. $R(\alpha,\chi)\cdot\hat{M}(\chi,\beta) \equiv \hat{h}_2(\chi)\cdot v_H(\chi) + \chi\cdot\hat{p}_2(\chi) + \tau/|H|$

$\Sigma_{a\in H} R(\alpha,a)\cdot\hat{M}(a,\beta) = \tau$

Compute $\sigma := \hat{M}(\gamma,\beta)$

Compute $\hat{h}_3,\hat{p}_3$ s.t.

$v_H(\gamma)\cdot v_H(\beta)\cdot\hat{\text{val}}^*(\chi) - (\gamma-\hat{\text{row}}(\chi))\cdot(\beta-\hat{\text{col}}(\chi))\cdot(\chi\cdot\hat{p}_3(\chi)+\sigma/|K|) \equiv \hat{h}_3(\chi)\cdot v_K(\chi)$

$\Sigma_{a\in K} \frac{v_H(\gamma)}{\gamma-\hat{\text{row}}(a)}\cdot\frac{v_H(\beta)}{\beta-\hat{\text{col}}(a)}\cdot\hat{\text{val}}^*(a) = \sigma$

$\xleftarrow{\alpha}$

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

$\xleftarrow{\beta}$

$\xrightarrow{\tau}$

$\xrightarrow{h_2,p_2 : L\to\mathbb{F}}$

$\xleftarrow{\gamma}$

$\xrightarrow{\sigma}$

$\xrightarrow{h_3,p_3 : L\to\mathbb{F}}$

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

Sample $\alpha\leftarrow\mathbb{F}$.

Test that h is δ-close to degree $d-|H|$.
p is δ-close to degree $|H|-2$.

Sample $\beta\in L$ and check that

$R(\alpha,\beta)\cdot g(\beta) - \tau\cdot f(\beta) = h(\beta)\cdot v_H(\beta) + \beta\cdot p(\beta)$.

Test that h_2 is δ-close to degree $|H|-2$.
p_2 is δ-close to degree $|H|-2$.

Sample $\gamma\in L$ and check that

$R(\alpha,\gamma)\cdot\sigma = h_2(\gamma)\cdot v_H(\gamma) + \gamma\cdot p_2(\gamma) + \tau/|H|$.

Test that h_3 is δ-close to degree $2|K|-3$.
p_3 is δ-close to degree $|K|-2$.

Sample $\mu\in L$ and check that

$v_H(\gamma)\cdot v_H(\beta)\cdot\hat{\text{val}}^*(\mu) - (\gamma-\hat{\text{row}}(\mu))\cdot(\beta-\hat{\text{col}}(\mu))\cdot(\mu\cdot\hat{p}_3(\mu)+\sigma/|K|) = \hat{h}_3(\mu)\cdot v_K(\mu)$

Holographic Lincheck [4/4]

$\mathcal{I}(M=(\text{row},\text{col},\text{val})) :=$ output the LDEs $\hat{\text{row}}, \hat{\text{col}}, \hat{\text{val}} : L \to \mathbb{F}$

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

1. Sumcheck for $\sum_{a\in H} R(\alpha,a)\cdot\hat{g}(a) - R_M(\alpha,a)\cdot\hat{f}(a) = 0$:

   compute $\hat{h},\hat{g}$ s.t. $R(\alpha,\gamma)\cdot\hat{g}(\gamma) - R_M(\alpha,\gamma)\cdot\hat{f}(\gamma) \equiv \hat{h}(\gamma)\cdot V_H(\gamma) + \gamma\cdot\hat{p}(\gamma)$

2. $\tau := R_M(\alpha,\beta) = \sum_{a\in H} R(\alpha,a)\cdot\hat{M}(a,\beta)$.

3. Sumcheck for $\sum_{a\in H} R(\alpha,a)\cdot\hat{M}(a,\beta) = \tau$:

   compute $\hat{h}_2,\hat{p}_2$ s.t. $R(\alpha,x)\cdot\hat{M}(x,\beta) \equiv \hat{h}_2(x)\cdot V_H(x) + x\cdot\hat{p}_2(x) + \tau/|H|$.

4. $\sigma := \hat{M}(\gamma,\beta)$.

5. Sumcheck for $\sum_{a\in K} \frac{V_H(\gamma)}{\gamma-\hat{\text{row}}(a)}\cdot\frac{V_H(\beta)}{\beta-\hat{\text{col}}(a)}\cdot\hat{\text{val}}^*(a) = \sigma$:

   compute $\hat{h}_3,\hat{p}_3$ s.t.

$$V_H(\gamma)\cdot V_H(\beta)\cdot\hat{\text{val}}^*(x) - (\gamma-\hat{\text{row}}(x))\cdot(\beta-\hat{\text{col}}(x))\cdot(x\cdot\hat{p}_3(x)+\sigma/|K|) \equiv \hat{h}_3(x)\cdot V_K(x)$$

$\xleftarrow{\alpha}$

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

$\xleftarrow{\beta}$

$\xrightarrow{\tau}$

$\xrightarrow{h_2,p_2:L\to\mathbb{F}}$

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

1. Sample $\alpha\leftarrow\mathbb{F}$.
2. $V^h_{LDT}(\mathbb{F},L,d-|H|) \stackrel{?}{=} 1$, $V^p_{LDT}(\mathbb{F},L,|H|-1) \stackrel{?}{=} 1$.
3. Sample $\beta\leftarrow L$.
4. Check that

   $R(\alpha,\beta)\cdot g(\beta) - \tau\cdot f(\beta) = h(\beta)\cdot V_H(\beta) + \beta\cdot p(\beta)$.

5. $V^{h_2}_{LDT}(\mathbb{F},L,|H|-1) \stackrel{?}{=} 1$, $V^{p_2}_{LDT}(\mathbb{F},L,|H|-1) \stackrel{?}{=} 1$.
6. Sample $\gamma\in L$ and check that

   $R(\alpha,\gamma)\cdot\hat{M}(\gamma,\beta) = h_2(\gamma)\cdot V_H(\gamma) + \gamma\cdot p_2(\gamma) + \tau/|H|$.

Bibliography

Holographic proofs

- [CHMMVW 2019]: Marlin: preprocessing zkSNARKs with universal and updatable SRS, by Alessandro Chiesa, Yuncong Hu, Mary Maller, Pratyush Mishra, Psi Vesely, Nicholas Ward.
- [COS 2019]: Fractal: post-quantum and transparent recursive proofs from holography, by Alessandro Chiesa, Dev Ojha, Nick Spooner.
- [CY 2024]: Building cryptographic proofs from hash functions, by Alessandro Chiesa, Eylon Yogev. (▸Video)

Holographic proofs to construct proof-carrying-data

Chapter on holography to preprocessing

More holography

- [GWC 2019]: PLONK: permutations over lagrange-bases for oecumenical noninteractive arguments of knowledge, by Ariel Gabizon, Zachary J. Williamson, Oana Ciobotaru.
- [Setty 2019]: Spartan: efficient and general-purpose zkSNARKs without trusted setup, by Srinath Setty.

Preprocessing for CSAT with custom gates

Preprocessing for R1CS with multilinear polynomoials