Lecture 25

Parallel Repetition

Foundations of Probabilistic Proofs
Alessandro Chiesa

Parallel Repetition

Basic Question: How to reduce the soundness error of a probabilistic proof?

Basic Answer: Run the probabilistic proof t times in sequence.

A straightforward analysis shows that the soundness error decreases from ε to ε^t.

Problem: Efficiency measures increase by a multiplicative factor of t.

Parallel Repetition loosely refers to ideas for reducing the soundness error of a probabilistic proof while **Preserving** certain efficiency measures.

Parallel repetition is a **Transformation** seperately defined for each proof model.

Today we discuss these cases:

- IPs (interactive proofs)
- MIPs (multi-prover interactive proofs)
- PCPs (probabilistically checkable proofs)
- IOPs (interactive oracle proofs)

While there are similarities and connections across models,
parallel repetition behaves surprisingly different in each case.

Parallel Repetition for IPs

Let (P,V) be an IP. Fix a repetition parameter t. Define $(P_t,V_t) := \mathrm{PR}_{\mathrm{IP}}(P,V,t)$ as follows:

$P_t(x)$

Sample prover randomness $\sigma_1,\dots,\sigma_t \leftarrow \{0,1\}^{pr}$

For $i\in[t]$: $a_{i,j} := P(b_{i,j-1};\sigma_i)$

$V_t(x)$

Sample verifier randomness $\rho_1,\dots,\rho_t \leftarrow \{0,1\}^{vr}$

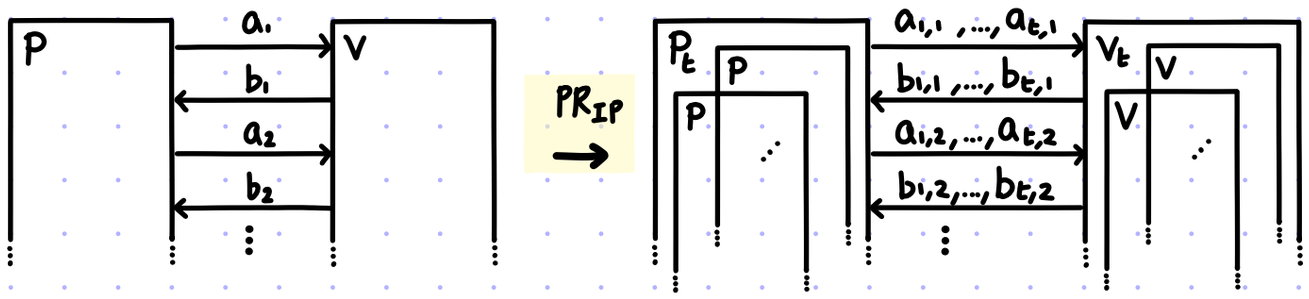
For $j=1,\dots,k$:

$\xrightarrow{(a_{i,j})_{i\in[t]}}$

$\xleftarrow{(b_{i,j})_{i\in[t]}}$

For $i\in[t]$: $b_{i,j} := V(a_{i,j};\rho_i)$

Check that all t executions of V accept.



- round complexity: $k \mapsto k$
- communication complexity: $(pc,vc) \mapsto (pc',vc') = (t\cdot pc, t\cdot vc)$
- completeness error: $\varepsilon_c \mapsto \varepsilon_c' = 1-(1-\varepsilon_c)^t \le t\cdot\varepsilon_c$ (completeness error increases slightly)
- soundness error: $\varepsilon_s \mapsto \varepsilon_s' = \varepsilon_s^t$ ← This seems intuitive but let's prove it.

Multi-Prover Interactive Proofs

A **multi-prover interactive proof (MIP)** is a probabilistic proof where a single (honest) verifier interacts with multiple NON-COMMUNICATING (possibly malicious) provers.

We say that (P,V) is an **MIP system** for a language L with p provers, completeness error ε_c, and soundness error ε_s if the following holds:

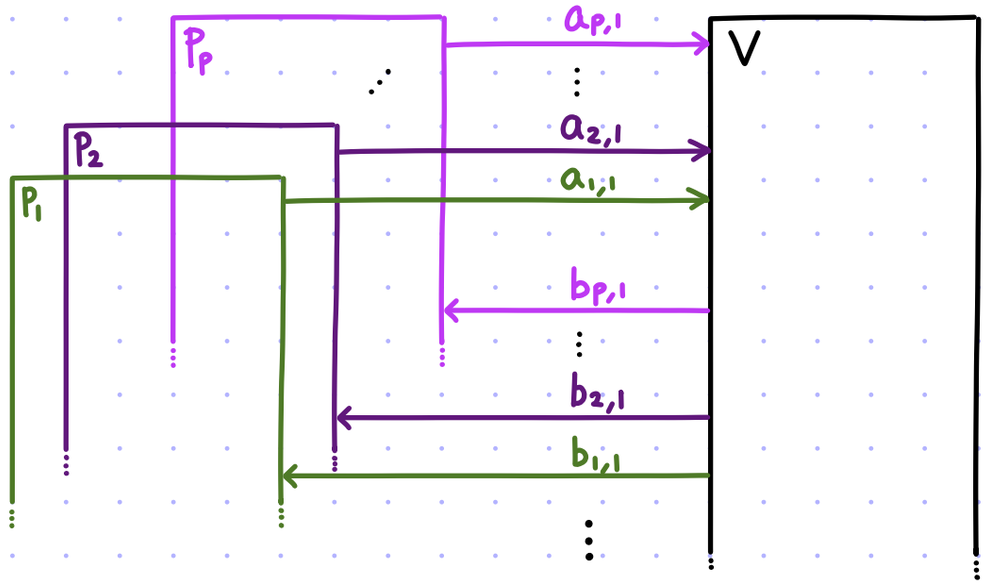
① COMPLETENESS: $\forall x \in L \quad \Pr_{\alpha,(\sigma_i)_{i\in[p]},\beta}\left[\langle (P(i,x,\alpha;\sigma_i))_{i\in[p]}, V(x;\beta)\rangle = 1\right] \geq 1-\varepsilon_c.$

② SOUNDNESS: $\forall x \notin L \quad \forall (\tilde{P}_i)_{i\in[p]} \quad \Pr_{\beta}\left[\langle (\tilde{P}_i)_{i\in[p]}, V(x;\beta)\rangle = 1\right] \leq \varepsilon_s.$

Each round $j \in [k]$ of interaction works as follows:

$\forall i \in [p]$ prover i sends a message $a_{i,j}$ to the verifier,

$\forall i \in [p]$ the verifier sends a message $b_{i,j}$ to prover i.



Efficiency measures:

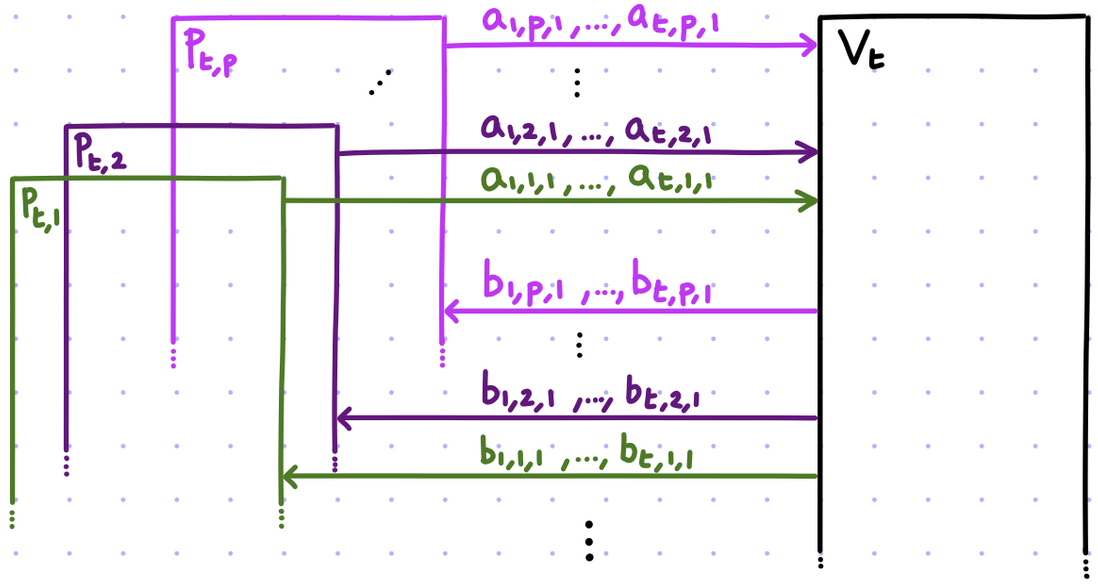
- p: number of provers
- k: round complexity
- pc: (total) prover-to-verifier communication
- vc: (total) verifier-to-prover communication

↖ could also separate by prover and round ↙

Parallel Repetition for MIPs

Let (P,V) be an MIP. Fix a repetition parameter t.

Define $(P_t,V_t) := PR_{MIP}(P,V,t)$ as follows:

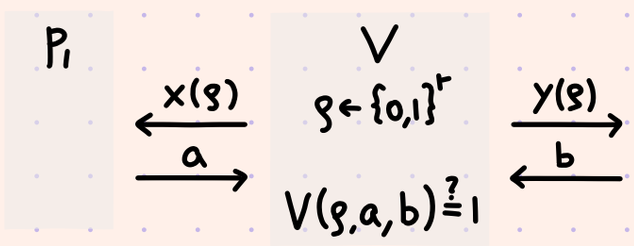


- number of provers: $p \mapsto p$
- round complexity: $k \mapsto k$
- communication complexity: $(pc, vc) \mapsto (pc', vc') = (t \cdot pc, t \cdot vc)$
- completeness error: $\varepsilon_c \mapsto \varepsilon_c' = 1-(1-\varepsilon_c)^t \leq t \cdot \varepsilon_c$ (completeness error increases slightly)
- soundness error: $\varepsilon_s \mapsto \varepsilon_s' = ?$

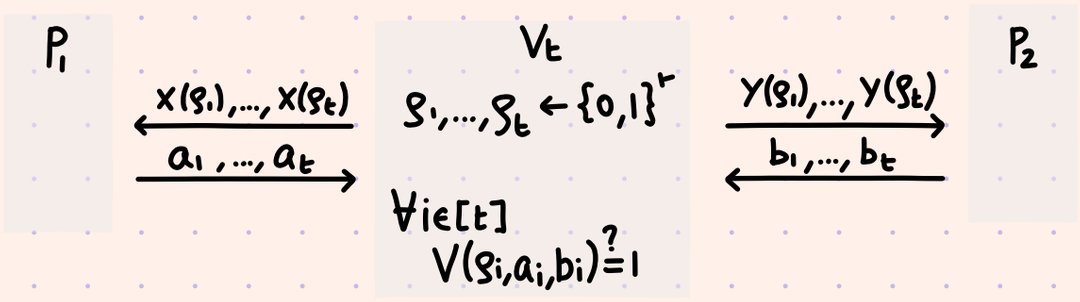
Q: $\varepsilon_s' = \varepsilon_s^t$?

Notable special case:

MIP with 2 provers, 1 round (verifier first)



Its t-wise PR has the same format:



claim: $\varepsilon_s^t \leq \varepsilon' \leq \varepsilon_s$

- ε_s^t: best prover in each repetition
- ε': winning all repetitions
- ε_s: at least as hard as winning one

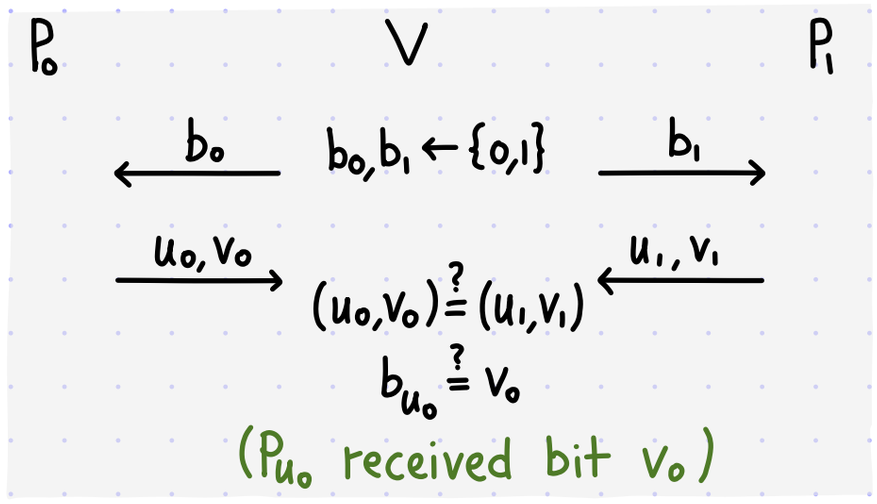
Refuting Expectation [1/2]

CONJECTURE [Fortnow, Rompel, Sipser 1988]: PR for MIPs reduces soundness error $\varepsilon \mapsto \varepsilon^k$.

[Fortnow 1989]: counterexample to conjecture.

Here we see a simpler counterample from [Feige 1991]:

MIP for NON-INTERACTIVE AGREEMENT (which has $p=2$ provers and $k=1$ rounds).



claim: $\varepsilon = 1/2$ where $\varepsilon := \max_{\tilde{P}_0, \tilde{P}_1} \Pr[\langle (\tilde{P}_0, \tilde{P}_1), V \rangle = 1]$

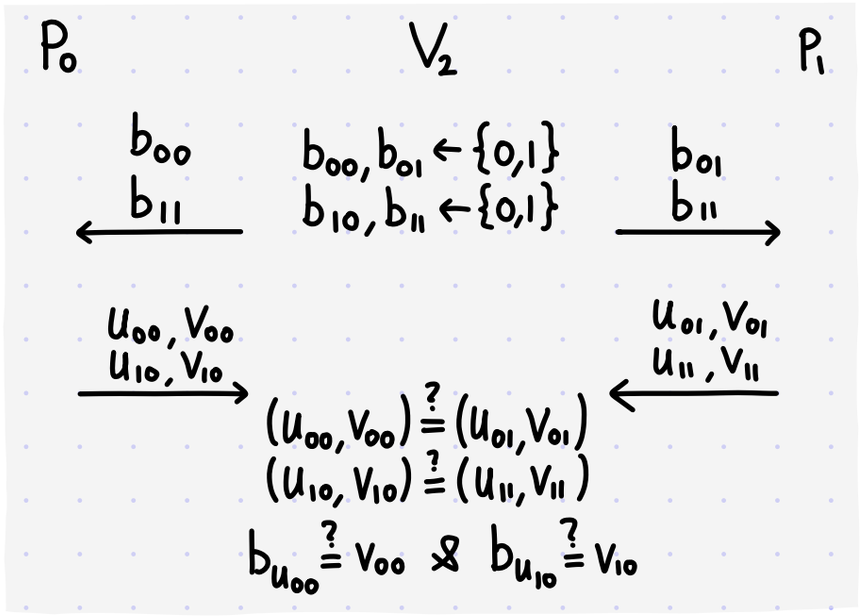
proof:

- $\varepsilon \geq 1/2$: P_0 answers $(0, b_0)$ and P_1 answers $(0, \text{random bit})$
- $\varepsilon \leq 1/2$: WLOG $\tilde{P}_0, \tilde{P}_1$ agree to guess P_0's bit b_0 ;

  hence $\tilde{P}_1$ must guess b_0 but has no information about b_0 ■

Refuting Expectation [2/2]

Consider the 2-wise PR of the MIP for non-interactive agreement:



!!!

claim: $\varepsilon_2 = 1/2$ where $\varepsilon_2 := \max_{\tilde{P}_0, \tilde{P}_1} \Pr[\langle (\tilde{P}_0, \tilde{P}_1), V_2 \rangle = 1]$

proof:

- $\varepsilon_2 \leq 1/2$: $\varepsilon_2 \leq \varepsilon = 1/2$
- $\varepsilon_2 \geq 1/2$: $\tilde{P}_0$ sends $\begin{matrix} 0, b_{00} \\ 1, b_{00} \end{matrix}$ and $\tilde{P}_1$ sends $\begin{matrix} 0, b_{11} \\ 1, b_{11} \end{matrix}$.

  If $b_{00} = b_{11}$ then $\tilde{P}_0$ and $\tilde{P}_1$ win BOTH iterations. ■

Another view: $\Pr[\text{WIN1} \wedge \text{WIN2}] = \Pr[\text{WIN1}] \cdot \Pr[\text{WIN2} \mid \text{WIN1}] = \frac{1}{2} \cdot 1 = \frac{1}{2}$.

In the second iteration, conditioning creates implicit communication between provers.

Verbitsky's Theorem

The counterexample shows that $\varepsilon_t \neq \varepsilon^t$, even in the minimal setting of $p=2$ provers and $k=1$ rounds.

But PR for the counterexample still "works", just slower than expected.

lemma [Feige 1991]: For the MIP for non-interactive agreement, $\varepsilon_t \leq \left(\frac{1}{2}\right)^{t/2} = \left(\frac{1}{\sqrt{2}}\right)^t$.

More precisely:
- t even → $\varepsilon_t = \left(\frac{1}{2}\right)^{t/2}$
- t odd → $\left(\frac{1}{2}\right)^{\frac{t+1}{2}} < \varepsilon_t < \left(\frac{1}{2}\right)^{t/2}$

Verbitsky proved that PR for 1-round MIPs always works:

theorem: Let ε be the soundness error of a 1-round MIP verifier V, and ε_t the soundness error of its t-wise parallel repetition V_t. For every instance $x \notin L$, if $\varepsilon(x) < 1$ then $\lim_{t \to \infty} \varepsilon_t(x) = 0$.

Verbitsky proved the theorem for $p=2$ provers.

The proof approach extends to work for any p.

study of combinatorial objects where "order" appears if large enough

The proof is a direct application of a deep result in RAMSEY THEORY,

and only shows that the soundness error ε_t decreases VERY SLOWLY.

A Result On Combinatorial Lines

Let A be a finite alphabet and $\Delta \notin A$ a special symbol.

A word is a string in A^* and a root is a string in $(A \cup \{\Delta\})^* \setminus A^*$.

For a root rt and $a \in A$, $rt(a)$ is the word (in A^*) obtained by replacing each Δ with a.

Ex: $A = \{1,2,3\}$ $rt = 3\,1\,\Delta\,1\,2\,\Delta$ $rt(1) = 3\,1\,1\,1\,2\,1$ $rt(3) = 3\,1\,3\,1\,2\,3$

A combinatorial line in A^t is a subset $L \subseteq A^t$ of the form $\{rt(a)\}_{a \in A}$ for a root rt.

Ex: $A = \{1,2,3\}$ $rt = 3\,1\,\Delta\,1\,2\,\Delta$ $L_{rt} = \begin{pmatrix} 3 & 1 & 1 & 1 & 2 & 1 \\ 3 & 1 & 2 & 1 & 2 & 2 \\ 3 & 1 & 3 & 1 & 2 & 3 \end{pmatrix}$

Combinatorial lines are in correspondence with roots, of which there are $(|A|+1)^t - |A|^t$.

Define $N(A,t) := \max\{|W| \mid W \subseteq A^t \text{ contains NO combinatorial lines}\} \in \{0,1,\dots,|A|^t\}$.

theorem [Furstenberg and Katznelson, 1991] $\forall A\ \forall \varepsilon > 0\ \exists T\ \forall t \geq T\ \ \frac{N(A,t)}{|A|^t} < \varepsilon$

This is a density version of the Hales–Jewett theorem ($\forall A\ \forall r\ \exists T\ \forall t \geq T$ every r-coloring of A^t has a monochromatic line).

In 2010 the Polymath project gave a quantitative bound: $T \sim Ack_{|A|}(1/\varepsilon)$.

Ackermann function:
$Ack_m(1) = 2$, $Ack_1(n) = 2n$, $Ack_m(n) = Ack_{m-1}(Ack_m(n-1))$

Proof of Verbitsky's Theorem

Let $A := \{0,1\}^r$ be the set of random strings of the MIP verifier $V := V(x)$.

We argue that if $\varepsilon < 1$ then $\varepsilon_t \leqslant \frac{N(A,t)}{|A|^t}$. This concludes the proof via [FK91].

Fix optimal $(\tilde{P}_{t,j})_{j\in[p]}$ against the t-wise PR verifier V_t.

Define the winning set W of V_t: $W := \{(s_1,\dots,s_t) \in A^t \mid \langle (\tilde{P}_{t,j})_{j\in[p]}, V_t((s_i)_{i\in[t]})\rangle = 1\}$.

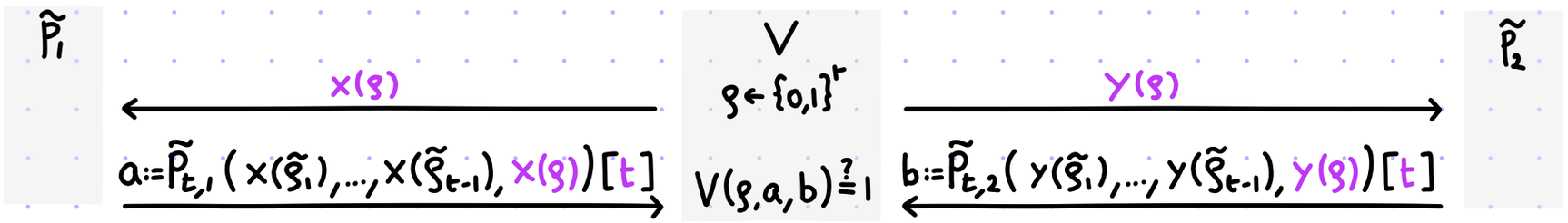
By definition, $\varepsilon_t = \frac{|W|}{|A|^t}$. It suffices to show that W contains no combinatorial line.

Suppose by way of contradiction that $\exists$ root rt whose combinatorial line L_{rt} is in W.

For simplicity $rt = \tilde{s}_1 \cdots \tilde{s}_{t-1} \Delta$. (The analysis below easily adapts to roots where Δ appears elsewhere.)

We construct $(\tilde{P}_j)_{j\in[p]}$ that convince V w.p. 1, contradicting the fact that $\varepsilon < 1$.

Here is the case of $p=2$ provers. The case $p>2$ is analogous.



Since $L_{rt} \in W$, we know that $\forall \tilde{s}_t \in \{0,1\}^r \ \wedge_{i\in[t]} V(\tilde{s}_i, \tilde{P}_{t,1}(x(\tilde{s}_1),\dots,x(\tilde{s}_t))[i], \tilde{P}_{t,2}(y(\tilde{s}_1),\dots,y(\tilde{s}_t))[i]) = 1$,

so $V(\tilde{s}_t, \tilde{P}_{t,1}(x(\tilde{s}_1),\dots,x(\tilde{s}_t))[t], \tilde{P}_{t,2}(y(\tilde{s}_1),\dots,y(\tilde{s}_t))[t]) = 1$.

Raz's Theorem

Much better rate of decay is known for a minimal special case:

theorem: [Raz 1995] There exists $c>0$ s.t. the following holds.

Let (P,V) be an MIP for a language L with soundness error ε, $p=2$ provers, and $k=1$ rounds.

Then $\forall x \notin L \quad \varepsilon(x) \leq 1-\delta \rightarrow \varepsilon_t(x) \leq (1-\delta^c)^{\Omega(t/\log|\Sigma|)}$, where Σ is the alphabet for prover answers.

- [Feige Verbitsky 1996]: the dependence on $\log|\Sigma|$ is necessary
- [Holenstein 2010]: $c \leq 3$ (vs. $c \leq 32$ in Raz's proof)
- cannot expect $c \leq 1$ in general (strong parallel repetition is the study of when $c \approx 1$)

For $p=3$ provers and $k=1$ rounds some recent progress on rate of decay (in special cases).

Understanding $p>2$ provers or $k>1$ rounds remains a CHALLENGING OPEN PROBLEM.

Main Lemma Behind Raz's Theorem

Fix strategies $\tilde{P}_1, \tilde{P}_2$ against the t-wise parallel repeated MIP verifier $V_t(x)$.

For $i \in [t]$, $W_i := V(x, g_i, \tilde{P}_1(x(g_1),\ldots,x(g_t))[i], \tilde{P}_2(y(g_1),\ldots,y(g_t))[i])$. For $S \subseteq [t]$, $W_S := \bigwedge_{i \in S} W_i$.

By assumption, $\forall i \in [t]$ $\Pr[W_i] \leq 1-\delta$.

GOAL: upper bound $\Pr[\bigwedge_{i \in [t]} W_i]$.

Main Lemma: $\exists \gamma$ (that depends on $V(x)$) $\forall S \subseteq [t]$ with $|S| \leq \gamma \cdot t$ if $\Pr[W_S] \geq 2^{-\gamma \cdot t}$ then $\exists i \in [t] \setminus S$ s.t. $\Pr[W_i | W_S] \leq 1 - \frac{\delta}{2}$.

↑ proved via sophisticated analysis based on Information Theory

This implies the theorem as explained below.

Initialize $S := \emptyset$ and do the following while $|S| \leq \gamma \cdot t$:

① If $\Pr[W_S] < 2^{-\gamma \cdot t}$ then exit loop.

② If $\Pr[W_S] \geq 2^{-\gamma \cdot t}$ then add to S some $i \in [t] \setminus S$ s.t. $\Pr[W_i | W_S] \leq 1 - \frac{\delta}{2}$. (exists by Main Lemma)

If the first condition is met at some iteration then $\Pr[\bigwedge_{i \in [t]} W_i] \leq \Pr[W_S] < 2^{-\gamma \cdot t}$.

If the first condition is never met, then we obtain $S = \{i_1, i_2, \ldots, i_{\gamma t}\}$ s.t.

$$\Pr[\bigwedge_{i \in [t]} W_i] \leq \Pr[W_S] = \Pr[W_{i_1}] \cdot \Pr[W_{i_2} | W_{\{i_1\}}] \cdot \Pr[W_{i_3} | W_{\{i_1,i_2\}}] \cdot \cdots \leq (1 - \tfrac{\delta}{2})^{\gamma \cdot t}.$$

We conclude that $\Pr[\bigwedge_{i \in [t]} W_i] \leq \max\{2^{-\gamma \cdot t}, (1 - \frac{\delta}{2})^{\gamma \cdot t}\}$.

Repetition Applied to the PCP Theorem

Recall the PCP Theorem: $NP \subseteq PCP\left[\varepsilon_c = 0, \varepsilon_s = \frac{1}{2}, \Sigma = \{0,1\}, \ell = \text{poly}(n), q = O(1), r = O(\log n)\right]$.

Rerunning the PCP verifier gives arbitrarily small ε_s with large-enough q:

corollary: $\forall \varepsilon_s > 0 \; \exists q$ s.t. $NP \subseteq PCP\left[\varepsilon_c = 0, \varepsilon_s, \Sigma = \{0,1\}, \ell = \text{poly}(n), q, r = O(\log n)\right]$

Repetition via PR of MIPs gives arbitrarily small ε_s with large-enough Σ (and with $q=2$):

theorem: $\forall \varepsilon_s > 0 \quad NP \subseteq PCP\left[\varepsilon_c = 0, \varepsilon_s, \Sigma = \{0,1\}^{O(\log \frac{1}{\varepsilon_s})}, \ell = n^{O(\log \frac{1}{\varepsilon_s})}, q = 2, r = O(\log \tfrac{1}{\varepsilon_s} \cdot \log n)\right]$

proof: PCP $\xrightarrow{①}$ MIP $\xrightarrow{②}$ repeated MIP $\xrightarrow{③}$ PCP

① From PCP to 2-prover 1-round MIP (via trivial query bundling)

$PCP[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, r] \subseteq MIP[\varepsilon_c, \varepsilon_s = 1 - \frac{1-\varepsilon_s}{q}, p = 2, k = 1, (\Sigma_v, \Sigma_p) = ([\ell]^q, \Sigma^q), r' = r + \log q]$

yields $NP \subseteq MIP\left[\varepsilon_c = 0, \varepsilon_s = O(1), p = 2, k = 1, (\Sigma_v, \Sigma_p) = (\{0,1\}^{O(\log n)}, \{0,1\}^{O(1)}), r = O(\log n)\right]$

② Apply PR for MIPs: $NP \subseteq MIP[\varepsilon_c = 0, \varepsilon_s = \varepsilon_t, p = 2, k = 1, (\Sigma_v, \Sigma_p) = (\{0,1\}^{O(t \cdot \log n)}, \{0,1\}^{O(t)}), r = O(t \cdot \log n)]$

③ Evaluate the MIP as a PCP: $NP \subseteq PCP\left[\varepsilon_c = 0, \varepsilon_s = \varepsilon_t, \Sigma = \{0,1\}^{O(t)}, \ell = n^{O(t)}, q = 2, r = O(t \cdot \log n)\right]$

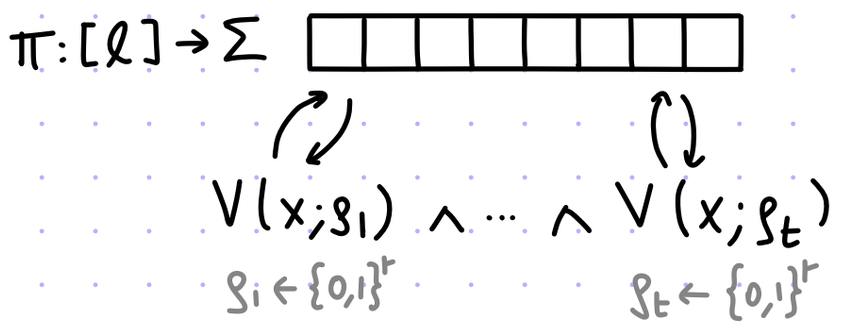
By Raz's Theorem, $\varepsilon_t = \exp(-t)$, so can set $t = O(\log 1/\varepsilon_s)$. ■

The main limitation of parallel repetition is that if we want $\ell = \text{poly}(n)$ then $\varepsilon_s = \Omega(1)$.

Reducing Soundness Error for PCPs

Q: How to reduce the soundness error of a PCP?

SIMPLE: repeat the PCP verifier multiple times



If the honest PCP prover uses randomness (e.g. for zk) one may have to sample t PCP strings $\pi_1,\ldots,\pi_t$ (one per repetition)

For every $t\in\mathbb{N}$, the t-wise repetition does the following:

- $\Sigma \mapsto \Sigma' = \Sigma$ alphabet does not change
- $\ell \mapsto \ell' = \ell$ proof length does not change
- $q \mapsto q' = t\cdot q$ query complexity increases
- $r \mapsto r' = t\cdot r$ randomness complexity increases
- $\varepsilon_c \mapsto \varepsilon_c' = 1-(1-\varepsilon_c)^t \leqslant t\cdot\varepsilon_c$ completeness error increases slightly
- $\varepsilon_s \mapsto \varepsilon_s' = \varepsilon_s^t$ soundness error decreases exponentially

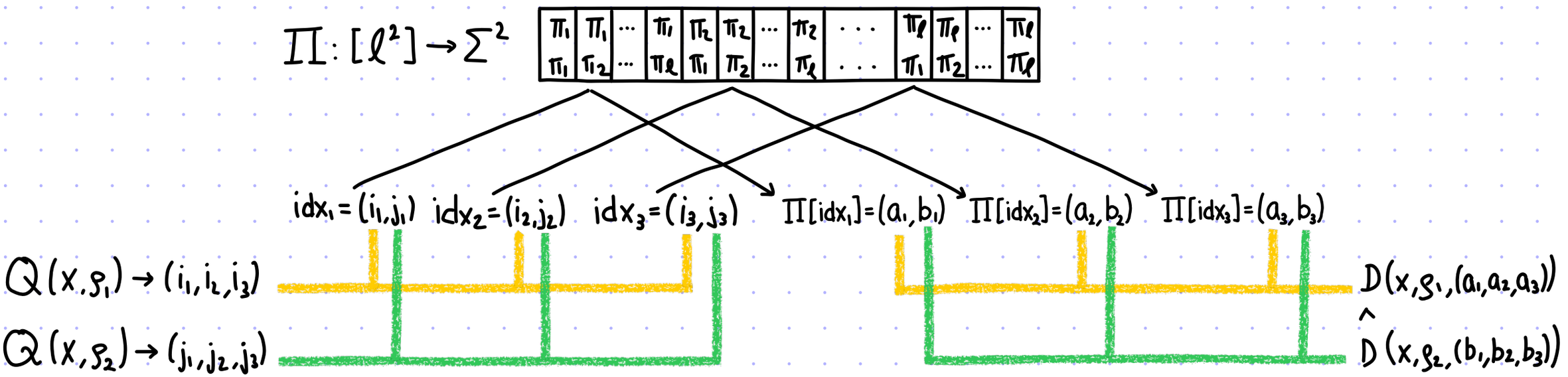
randomness efficient error-reduction (e.g. via expanders) has no better query complexity

How to reduce soundness error while preserving query complexity?

Parallel Repetition for PCPs [1/2]

IDEA: bundle queries across multiple repetitions ("BUNDLED REPETITION")

Example with $t = 2$ repetitions for a PCP with $q = 3$ queries.



The query complexity did NOT change.

Each query consists of two indices and is answered via two symbols.

→ The proof length and alphabet size square.

Parallel Repetition for PCPs [2/2]

The t-wise parallel repetition of a (non-adaptive) PCP:

$P_t(x)$

1. Compute $\pi := P(x) \in \Sigma^\ell$.
2. Set $\Pi := \left((\pi[i_1], \ldots, \pi[i_t])\right)_{i_1,\ldots,i_t \in [\ell]}$.
3. Output $\Pi : [\ell^t] \to \Sigma^t$.

$V_t(x)$

1. Sample $s_1, \ldots, s_t \in \{0,1\}^r$.
2. Deduce query sets: $\forall i \in [t]$, $Q_i := Q(x, s_i) \subseteq [\ell]$.
3. Construct tuples: $\forall j \in [q]$ $idx_j := (Q_1[j], \ldots, Q_t[j])$.
4. Check that $\wedge_{i \in [t]} D(x, s_i, \Pi[idx_1]_i \cdots \Pi[idx_q]_i) = 1$.

- $\Sigma \mapsto \Sigma' = \Sigma^t$ alphabet increases exponentially
- $\ell \mapsto \ell' = \ell^t$ proof length increases exponentially
- $q \mapsto q' = q$ query complexity does NOT increase
- $r \mapsto r' = t \cdot r$ randomness complexity increases
- $\varepsilon_c \mapsto \varepsilon_c' = 1 - (1-\varepsilon_c)^t \leqslant t \cdot \varepsilon_c$ completeness error increases slightly
- $\varepsilon_s \mapsto \varepsilon_s' = ?$ How does PR of a PCP affect soundness error?

Refuting Expectation, Again

Parallel repetition for PCPs fails to work.

theorem: $\exists$ 2-query PCP for NP-complete language L with soundness error ε s.t.

$\forall x \notin L \quad \varepsilon(x) < 1$ and $\lim_{t \to \infty} \varepsilon_t(x) = 1$ (In fact, for infinitely many $x \notin L$, $\varepsilon_{t+1}(x) > \varepsilon_t(x)$ $\forall t \in \mathbb{N}$.)

In particular, NOT true that $\varepsilon(x)^t \leq \varepsilon_t(x) \leq \varepsilon(x)$.

Here is a CRITERION of when PR for PCPs works.

def: The MIP projection of a PCP verifier V is the MIP verifier V_{MIP} that works as follows:

$V_{MIP}(x)$: 1. Sample PCP randomness $\rho \leftarrow \{0,1\}^r$.

2. Deduce query set $Q := V_q(x, \rho) \subseteq [\ell]$.

3. For every $i \in [q]$, send $Q[i]$ to prover i, and get response $a_i \in \Sigma$.

4. Check that $V_D(x, \rho, (a_i)_{i \in [q]}) = 1$.

lemma: Let (P,V) be a PCP for a language L with soundness error ϵ. Let ε_t be the soundness error of its t-wise parallel repetition. Let V_{MIP} be the MIP projection V, with soundness error ε_{MIP}.

Then $\forall x \notin L \quad \lim_{t \to \infty} \varepsilon_t(x) = 0 \iff \varepsilon_{MIP}(x) < 1$.

Consistent Parallel Repetition

A simple variant of PR for PCPs does always work.

The t-wise consistent parallel repetition (CPR) of a (non-adaptive) PCP:

$P_t(x)$

1. Compute $\pi := P(x) \in \Sigma^\ell$.
2. Set $\Pi := \left((\pi[i_1],\dots,\pi[i_t])\right)_{i_1,\dots,i_t \in [\ell]}$.
3. Output $\Pi : [\ell^t] \to \Sigma^t$.

$V_t(x)$

1. Sample $\rho_1,\dots,\rho_t \in \{0,1\}^r$.
2. Deduce query sets: $\forall i \in [t]$, $Q_i := Q(x,\rho_i) \subseteq [\ell]$.
3. Construct tuples: $\forall j \in [q]$ $idx_j := (Q_1[j],\dots,Q_t[j])$.
4. Check that $\wedge_{i\in[t]} D(x,\rho_i,\Pi[idx_1]_i \cdots \Pi[idx_q]_i) = 1$.
5. If $\exists i,i' \in [t], j,j' \in [q]$ s.t. $Q_i[j] = Q_{i'}[j']$ and $\Pi[idx_j]_i \neq \Pi[idx_{j'}]_{i'}$, reject.

<u>theorem:</u> Let (P,V) be a PCP for a language L with soundness error ε.

Let ε_t be the soundness error of its t-wise CONSISTENT parallel repetition.

Then $\forall x \notin L$ $\quad \varepsilon(x) < 1 \to \varepsilon_t(x) < \binom{2^r}{\varepsilon(x)\cdot 2^r} \cdot \varepsilon(x)^t$ $\quad$ (in particular, $\lim_{t\to\infty} \varepsilon_t(x) = 0$).

The proof is an elementary counting argument to upper bound the winning set of V_t.