Lecture 23

Foundations of Probabilistic Proofs
Fall 2020
Alessandro Chiesa

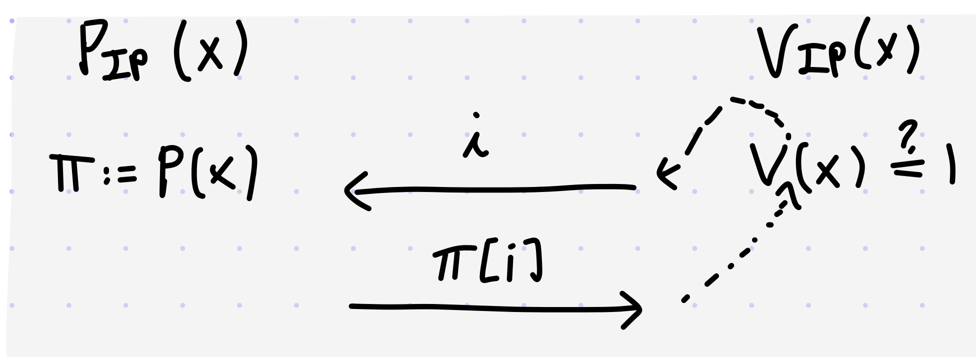
Limits on Query Complexity

We have proved the PCP Theorem: $NP \subseteq PCP[\varepsilon_c = 0, \varepsilon_s = 1/2, \Sigma = \{0,1\}, \ell = \mathrm{poly}(n), q = O(1), r = O(\log n)]$.

Q: How small can query complexity be?

- We do not expect $q=1$ for hard languages:

  Suppose that L has a PCP (P,V) with proof length ℓ over alphabet Σ, and with query complexity $q=1$. Then L has a 1-round IP as follows:



  The prover-to-verifier communication complexity is $\log|\Sigma|$.
  By the limitations on laconic IPs that we saw earlier, we cannot expect $\log|\Sigma| = o(n)$ for NP-hard languages (e.g. 3SAT).

- The situation with $q=2$ is quite different.

Two-Query PCPs

Are there two-query PCPs?

- No, if over the binary alphabet $\Sigma=\{0,1\}$ (and the PCP is non-adaptive):

<u>lemma</u>: $PCP[\varepsilon_c=0, \varepsilon_s<1, \Sigma=\{0,1\}, \ell=\text{poly}(n), q=2, r=O(\log n)] \subseteq P$

<u>proof</u>: We view a candidate PCP string as ℓ variables $z_1,\dots,z_\ell$.
For every choice of randomness $\rho\in\{0,1\}^r$, the decision algorithm of $V(x;\rho)$ is a function $\phi_{x,\rho}(z_1,\dots,z_\ell)$ that depends on two variables among the ℓ variables.
If $x\in L$ then there is an assignment $a_1,\dots,a_\ell$ s.t. $\bigwedge_\rho \phi_{x,\rho}(a_1,\dots,a_\ell)=1$
If $x\notin L$ then there is no assignment that satisfies more than an ε_s-fraction of $\{\phi_{x,\rho}\}_\rho$.
Deciding between these two is an instance of 2SAT, which is in P. ■

- Yes, if over larger alphabets Σ:

<u>lemma</u>: $\exists c\in\mathbb{N}$ $NP \subseteq PCP[\varepsilon_c=0, \varepsilon_s=1-\frac{1}{c}, \Sigma=\{0,1\}^c, \ell=\text{poly}(n), q=2, r=O(\log n)]$

<u>proof</u>: Apply the trivial query bundling to the PCP Theorem. ■

$PCP[\varepsilon_s,\Sigma,\ell,q,r] \subseteq PCP[\varepsilon_s'=1-(1-\varepsilon_s)\frac{1}{q}, \Sigma'=\Sigma^q, \ell'=O(\ell+2^r), q'=2, r'=r+\log q]$.

Small Query Complexity <u>and</u> Small Soundness Error?

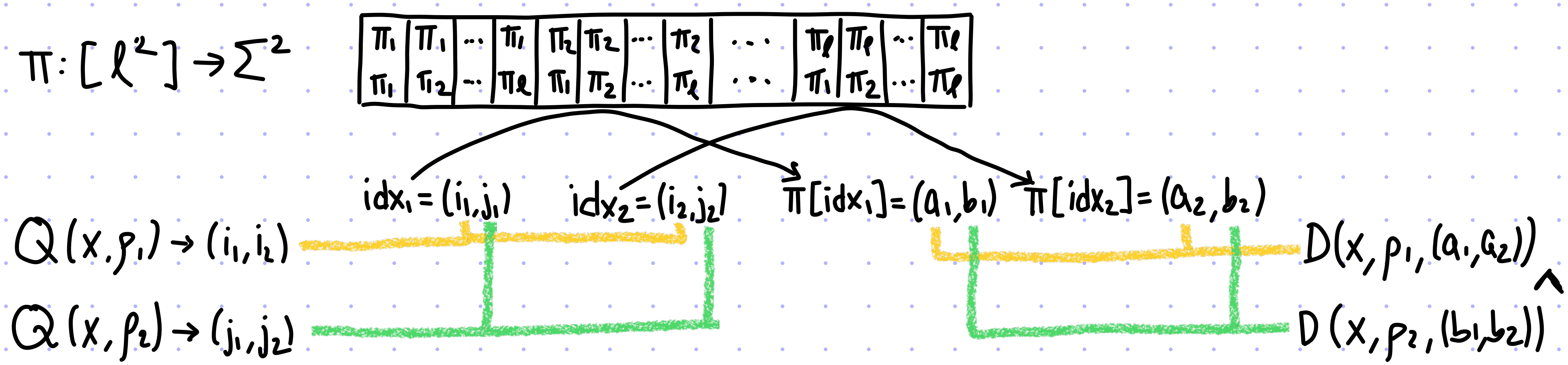
Repeating the PCP verifier reduces soundness error but also increases query complexity:

$$\forall\, t,\ \mathrm{PCP}[\varepsilon_c=1, \varepsilon_s, \Sigma, \ell, q, r] \subseteq \mathrm{PCP}[\varepsilon_c'=1, \varepsilon_s'=\varepsilon_s^t, \Sigma'=\Sigma, \ell'=\ell, q'=t\cdot q, r'=t\cdot r]$$

And randomness-efficient error reduction (e.g. via expanders) does not help for this.

Idea: bundle queries across multiple repetitions

E.g. $\pi : [\ell^2] \to \Sigma^2$



The proof length <u>and</u> the alphabet size squares.

Each query consists of one symbol per repetition.

The soundness error did not increase as winning is at least as hard as winning one instance.

The intuition is that the soundness error should be smaller, ideally quadratically so.

Parallel Repetition

More generally, this leads to the t-wise parallel repetition of a given (non-adaptive) PCP:

$P_t(x)$

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

$V_t(x)$

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 $\bigwedge_{i\in[t]} D(x,\rho_i,\Pi[idx_1]_i \dots \Pi[idx_q]_i) = 1$.

The proof length and alphabet size increase exponentially in t.

The number of queries remains the same, and each query is a tuple of t indices.

The new soundness error ε_s' must satisfy $\varepsilon_s^t \leq \varepsilon_s' \leq \varepsilon_s$ (ε_s is the old soundness error).

The intuition is that ε_s' should be equal to ε_s^t (exponentially smaller than ε_s).

In sum, we expect the t-wise parallel repetition to yield this inclusion:

$$\mathrm{PCP}[\varepsilon_c=1,\varepsilon_s,\Sigma,\ell,q,r] \subseteq \mathrm{PCP}[\varepsilon_c'=1,\varepsilon_s'=\varepsilon_s^t,\Sigma'=\Sigma^t,\ell'=\ell^t,q'=q,r'=t\cdot r].$$

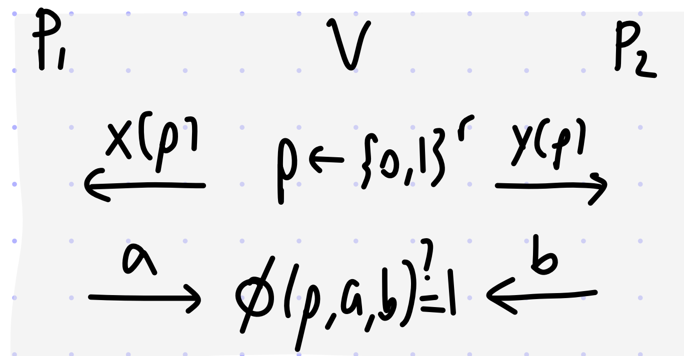
We will see that this is false in general, though the intuition is qualitatively true when $q=2$.

2-Player 1-Round Games

We focus on $q=2$ and move to a different view to discuss parallel repetition:

<u>def</u>: A 2P1R game is a tuple (x,y,ϕ) where $x:\{0,1\}^r \to S_1$ and $y:\{0,1\}^r \to S_2$ are the verifier's message functions and $\phi:\{0,1\}^r \times \Sigma_1 \times \Sigma_2 \to \{0,1\}$ is the verifier's decision predicate.

The game is played as follows:



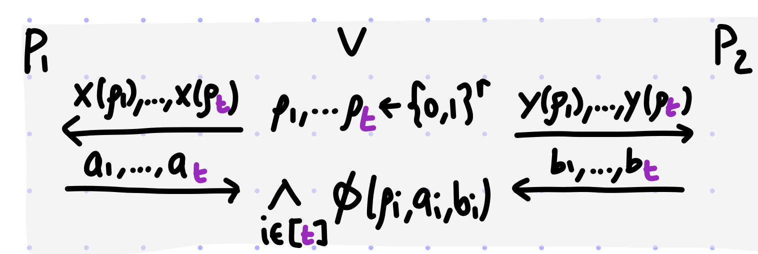
The players can share randomness but that does not change a games value.

The value of the game is $\mathrm{val}(G) := \max\limits_{f,g} \Pr\limits_{\rho}[\phi(\rho, f(x(\rho)), g(y(\rho)))=1]$.

This view is essentially equivalent to 2-query PCPs:

- $PCP[\varepsilon_c, \varepsilon_s, \Sigma, \ell, 2, r] \to 2P1R[\varepsilon_c, \varepsilon_s' = 1 - \frac{1-\varepsilon_s}{2}, (\Sigma^2, \Sigma), r]$
- $2P1R[\varepsilon_c, \varepsilon_s, (\Sigma_1, \Sigma_2), r] \to PCP[\varepsilon_c, \varepsilon_s, \Sigma = \Sigma_1 \cup \Sigma_2, \ell = 2 \cdot 2^r, q=2, r]$

The t-wise parallel repetition $pr(G,t)$ of G is the game:



That is, playing with strategies f,g means: $\bigwedge_{i\in[t]} \phi(\rho_i, f_i(x(\rho_1),\dots,x(\rho_t)), g_i(y(\rho_1),\dots,y(\rho_t)))$

It is straightforward to see that $\mathrm{val}(G)^t \le \mathrm{val}(pr(G,t)) \le \mathrm{val}(G)$.

Question (that encapsulates parallel repetition for 2-query PCPs): $\mathrm{val}(pr(G,t)) = \mathrm{val}(G)^t$?

Refuting Expectation

Fortnow, Rompel, Sipser (1988) conjectured that parallel repetition decays exactly exponentially.

Then in 1989 Fortnow found a counterexample to this conjecture.

We see a simpler counterexample due to Feige in 1991, known as "non-interactive agreement".

Consider the following 2P1R game $\tilde{G}$:

P_0 $\qquad$ V $\qquad$ P_1

$\xleftarrow{b_0}$ $\quad b_0, b_1 \leftarrow \{0,1\}$ $\quad \xrightarrow{b_1}$

$\xrightarrow{u_0, v_0}$ $\quad (u_0, v_0) \stackrel{?}{=} (u_1, v_1)$ $\quad \xleftarrow{u_1, v_1}$

$b_{u_0} \stackrel{?}{=} v_0$

("P_{u_0} received bit v_0")

claim: $\mathrm{val}(\tilde{G}) = 1/2$

proof:

- $\mathrm{val}(\tilde{G}) \geq 1/2$: P_0 answers $(0, b_0)$ and P_1 answers $(0, \$)$
- $\mathrm{val}(\tilde{G}) \leq 1/2$: WLOG both players agree on whose player's bit they guess, and one player has to guess a random bit that it knows nothing about. ∎

Refuting Expectation

Now consider the 2-wise repetition of the 2PIR game $\tilde{G}$:

P_0 $\qquad$ V $\qquad$ P_1

$$b_{00}, b_{01} \leftarrow \{0,1\}$$
$$b_{10}, b_{11} \leftarrow \{0,1\}$$

$$P_0 \xleftarrow{b_{00},\, b_{10}} V \xrightarrow{b_{01},\, b_{11}} P_1$$

$$P_0 \xrightarrow{u_{00}, v_{00};\ u_{10}, v_{10}} V \xleftarrow{u_{01}, v_{01};\ u_{11}, v_{11}} P_1$$

$$(u_{00}, v_{00}) \stackrel{?}{=} (u_{01}, v_{01})$$
$$(u_{10}, v_{10}) \stackrel{?}{=} (u_{11}, v_{11})$$

$$b_{u_{00}} \stackrel{?}{=} v_{00} \ \&\ b_{u_{10}} = v_{10}$$

<u>claim</u>: $\mathrm{val}(\mathrm{pr}(\tilde{G}, 2)) = 1/2$ $\qquad$ (!!!!)

<u>proof</u>:

- $\mathrm{val}(\mathrm{pr}(\tilde{G},2)) \leq 1/2$: $\quad \mathrm{val}(\mathrm{pr}(\tilde{G},2)) \leq \mathrm{val}(\tilde{G}) = 1/2$
- $\mathrm{val}(\mathrm{pr}(\tilde{G},2)) \geq 1/2$:

$$P_0 \xrightarrow{0, b_{00};\ 1, b_{00}} V \xleftarrow{0, b_{11};\ 1, b_{11}} P_1$$

If $b_{00} = b_{11}$ then players win <u>both</u> games.
This happens w.p. $1/2$. ■

Another view: $\Pr[\text{win } \{1,2\}] = \Pr[\text{win } 1] \cdot \Pr[\text{win } 2 \mid \text{win } 1] = \frac{1}{2} \cdot 1$

In game 2, the conditioning creates implicit communication between provers

Verbitsky's Theorem

We have learned that the following is false: $\mathrm{val}(\mathrm{pr}(G,t)) = \mathrm{val}(G)^t$.

That said for the counterexample $\tilde{G}$ parallel repetition does work, just slower than expected:

Feige in 1991 also proved that $\mathrm{val}(\mathrm{pr}(\tilde{G},t)) = \left(\frac{1}{2}\right)^{t/2} = \left(\frac{1}{\sqrt{2}}\right)^{t}$.

More generally, Verbitsky proved that parallel repetition decreases value for every game:

<u>theorem:</u> $\forall$ 2P1R game G if $\mathrm{val}(G) < 1$ then $\lim_{t\to\infty} \mathrm{val}(\mathrm{pr}(G,t)) = 0$.

The proof generalizes to any number of players.

The proof is a direct application of a deep result in Ramsey theory (which studies conditions under which "order" must appear), and so the value decays SLOWLY.

Via the PCP $\rightleftarrows$ 2P1R-game connection and the PCP Theorem we learn that:

<u>corollary:</u> $\forall \varepsilon > 0 \ \exists \Sigma$ s.t. $NP \subseteq PCP[\varepsilon_c = 0, \varepsilon_s = \varepsilon, \Sigma, \ell = \mathrm{poly}_\varepsilon(n), q = 2, r = O_\varepsilon(\log n)]$

That is, with 2 queries we can make the error as small as we want, for a large enough alphabet.
Yet at this point we don't know if the number of repetitions can be a reasonable function of ε.

A Result On Combinatorial Lines

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

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

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

Ex: $A = \{1,2,3\}$ $r = 31\triangle12\triangle$ $r(1) = 311121$ $r(3) = 313123$

A **combinatorial line** in A^t is a subset $L \subseteq A^t$ that looks like $\{r(a)\}_{a\in A}$ for a root r.

Ex: $A = \{1,2,3\}$ $r = 31\triangle12\triangle$ $L_r = \begin{pmatrix} 3 & 1 & 1 & 1 & 2 & 1 \\ 3 & 1 & 2 & 1 & 2 & 2 \\ 3 & 1 & 3 & 1 & 2 & 3 \end{pmatrix}$

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

Let $N(A,t)$ be the maximum size of any set $W \subseteq A^t$ containing NO combinatorial lines.

Note that $N(A,t)$ is an integer in $\{0,1,\dots,|A|^t\}$.

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

A density version of the Hales–Jewett Theorem ($\forall A\ \forall r \in \mathbb{N}\ \exists T\ \forall t \geq T$ every r-coloring of A^t has a monochromatic line)

The Polymath project gave in 2010 quantitative bounds: $T \sim \mathrm{Ack}_{|A|}\left(\frac{1}{\varepsilon}\right)$.

(The Ackermann function is: $\mathrm{Ack}_m(1) = 2$, $\mathrm{Ack}_1(n) = 2n$, $\mathrm{Ack}_m(n) = \mathrm{Ack}_{m-1}(\mathrm{Ack}_m(n-1))$.)

Proof of Verbitsky's Theorem

Note: the argument below works similarly for any number of players.

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

We argue that if $\mathrm{val}(G) < 1$ then $\mathrm{val}(\mathrm{pr}(G,t)) \le \frac{N(A,t)}{|A|^t}$, which concludes the proof (via [FK91]).

Let $x(\rho), y(\rho)$ be the verifier's messages to P_1, P_2 when the randomness is ρ.

Fix optimal strategies f, g for P_1, P_2 and define the winning set:

$$W = \left\{ (\rho_1, \dots, \rho_t) \in A^t \;\middle|\; \wedge_{i \in [t]} \; \phi\big(\rho_i, f_i(x(\rho_1), \dots, x(\rho_t)), g_i(y(\rho_1), \dots, y(\rho_t))\big) \right\}.$$

By definition $\mathrm{val}(\mathrm{pr}(G,t)) = \frac{|W|}{|A|^t}$. It suffices to show that W contains no combinatorial line.

Suppose by way of contradiction that there is a root rt whose combinatorial line L_{rt} is in W.

For simplicity $rt = \tilde{\rho}_1 \dots \tilde{\rho}_{t-1} \Delta$ and consider the following strategies to play G:

$$\begin{array}{ccccc}
P_1 & & V & & P_2 \\
 & \xleftarrow{x(\rho)} & \rho \leftarrow A & \xrightarrow{y(\rho)} & \\
f_t(x(\tilde{\rho}_1), \dots, x(\tilde{\rho}_{t-1}), x(\rho)) & \longrightarrow & & \longleftarrow & g_t(y(\tilde{\rho}_1), \dots, y(\tilde{\rho}_{t-1}), y(\rho))
\end{array}$$

These strategies win w.p. 1 because the entire combinatorial line is in W.

This contradicts the assumption that $\mathrm{val}(G) < 1$.