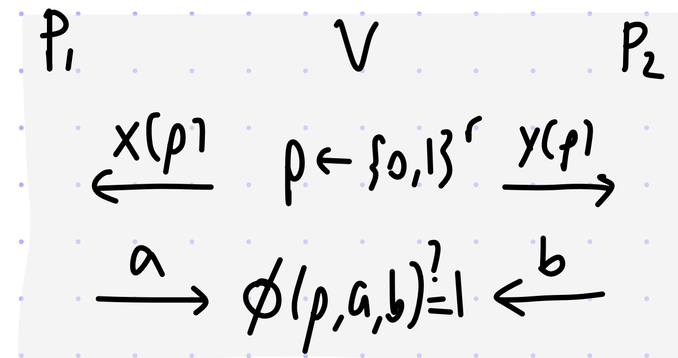
# Lecture 24

Foundations of Probabilistic Proofs
Fall 2020
Alessandro Chiesa

# Parallel Repetition of 2P1R Games

Recall the notion of a 2-prover 1-round game:

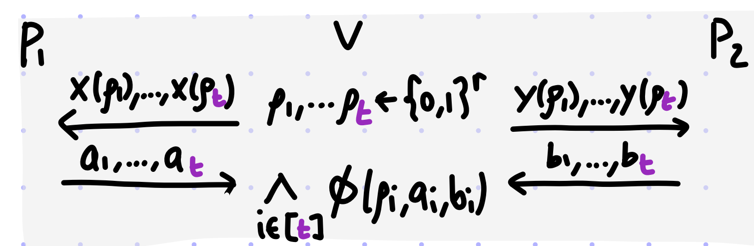



<u>def:</u> A 2P1R game is a tuple $(x, y, \phi)$ where

- $x: \{0,1\}^r \to S_1$ and $y: \{0,1\}^r \to S_2$ are the verifier's message functions;
- $\phi: \{0,1\}^r \times \Sigma_1 \times \Sigma_2 \to \{0,1\}$ is the verifier's decision predicate.

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

The view of 2P1R games is essentially equivalent to 2-query PCPs.




The $t$-wise parallel repetition $\mathrm{pr}(G,t)$ 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 $\mathrm{val}(G)^t \le \mathrm{val}(\mathrm{pr}(G,t)) \le \mathrm{val}(G)$.

Last time we proved Verbitsky's theorem: $\lim_{t\to\infty} \mathrm{val}(\mathrm{pr}(G,t)) = 0$ if $\mathrm{val}(G) < 1$.

Today we briefly discuss what is known about the rate of decay of $\mathrm{val}(\mathrm{pr}(G,t))$.

# Raz's Theorem

In 1995 Raz proved that parallel repetition makes the value decrease exponentially:

<u>theorem</u>: $\forall$ 2P1R-game $G$ $\exists \mu = \mu(G)$ s.t. $\mathrm{val}(\mathrm{pr}(G,t)) \le \mu(G)^t$.

In more detail the theorem states that there is a universal constant $c>0$ s.t.
if answers in $G$ are over alphabet $\Sigma$ and $\mathrm{val}(G) \le 1-\varepsilon$ then $\mathrm{val}(\mathrm{pr}(G,t)) \le (1-\varepsilon^c)^{\Omega(t/\log|\Sigma|)}$.

Remarks:
- [Feige Verbitsky 1996]: the dependence on $\log|\Sigma|$ is necessary
- [Holenstein 2010]: can take $c \le 3$ (vs. $c \le 32$ in Raz's proof)
- cannot expect $c \le 1$ for all games [& the study of when $c \approx 1$ is <u>strong</u> parallel repetition]

<u>corollary</u>: $\forall \varepsilon > 0$ $NP \subseteq PCP[\varepsilon_c = 0, \varepsilon_s = \varepsilon, \Sigma = \{0,1\}^{O(\log\frac{1}{\varepsilon})}, \ell = n^{O(\log\frac{1}{\varepsilon})}, q = 2, r = O(\log\frac{1}{\varepsilon}\cdot\log n)]$

<u>proof</u>: In three steps:

① Go from PCP Theorem to 2P1R-game $G$ with $\mathrm{val}(G) < 1$, $r = O(\log n)$, and $\Sigma = \{0,1\}^{O(1)}$.

② Parallel repeat game with $t = (\log\frac{1}{\varepsilon})/(\log\mu(G)) = O(\log\varepsilon)$. By Raz's Theorem $\mathrm{val}(\mathrm{pr}(G,t)) \le \varepsilon$.

③ Go from (repeated) 2P1R-game back to a 2-query PCP (with $\Sigma = \{0,1\}^{O(t)}$, $\ell = 2^{O(t\cdot r)}$).

# Main Lemma Behind Raz's Theorem

Fix strategies $f, g$ for the $t$-wise parallel repetition $pr(G,t)$.

Define the indicator $W_i =$ "$\phi(p_i, f_i(x(p_1),\dots,x(p_t)), g_i(y(p_1),\dots,y(p_t)))$" and, more generally for $S \subseteq [t]$, $W_S = \bigwedge_{i\in S} W_i$.

By assumption we know that $\Pr[W_1],\dots,\Pr[W_t] \leq 1-\varepsilon$.

The goal is to bound $\Pr[W_1 \wedge \dots \wedge W_t]$.

<u>Main Lemma</u>: $\exists \gamma = \gamma(G)$ $\forall S \subseteq [t]$ with $|S| \leq \gamma \cdot t$ if $\Pr[W_S] \geq 2^{-\gamma t}$ then $\exists i \in [t] \setminus S$ $\Pr[W_i | W_S] \leq 1 - \frac{\varepsilon}{2}$.

This implies the theorem as explained below.

Start with $S = \emptyset$ and do the following while $|S| \leq \gamma t$:

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

② If $\Pr[W_S] \geq 2^{-\gamma t}$ then add to $S$ a new index $i$ s.t. $\Pr[W_i | W_S] \leq 1 - \frac{\varepsilon}{2}$ (guaranteed by Main Lemma).

If the first condition is met at some iteration then $\Pr[W_1 \wedge \dots \wedge W_t] \leq \Pr[W_S] \leq 2^{-\gamma t}$.

If the first condition is never met, then we obtain $S = \{i_1, i_2, \dots, i_{\gamma t}\}$ such that

$$\Pr[W_1 \wedge \dots \wedge W_t] \leq \Pr[W_S] = \Pr[W_{i_1}]\Pr[W_{i_2} | W_{\{i_1\}}]\Pr[W_{i_3} | W_{\{i_1,i_2\}}] \cdots \leq \left(1 - \frac{\varepsilon}{2}\right)^{\gamma t}.$$

We conclude that $\Pr[W_1 \wedge \dots \wedge W_t] \leq \max\{2^{-\gamma t}, (1-\varepsilon/2)^{\gamma t}\} = \exp(-c(G)\, t)$.

# PCPs with Sub-Constant Soundness Error

Parallel repetition gives, for $q=2$, soundness error $\varepsilon$ over an alphabet of size $|\Sigma| = \mathrm{poly}(\frac{1}{\varepsilon})$.
The main limitation is that proof length becomes $\ell = n^{O(\frac{1}{\varepsilon})}$ so that if we want $\ell = \mathrm{poly}(n)$ then parallel repetition does not tell us anything for $\varepsilon = o(1)$.

Q: Can one achieve sub-constant soundness error over a super-constant alphabet size?
[While keeping $q=2$, or at most $q=O(1)$, and $\ell = \mathrm{poly}(n)$.]

Several constructions:

| | | | | |
|---|---|---|---|---|
| [AS97], [RS97] | $\varepsilon \geq \exp(-(\log n)^{\frac{1}{10}})$ | $\lvert\Sigma\rvert = \mathrm{poly}(\frac{1}{\varepsilon})$ | $\ell = \mathrm{poly}(n)$ | $q = 2$ |
| [DFKRS99] | $\varepsilon \geq \exp(-(\log n)^{1-\delta})$ | $\lvert\Sigma\rvert = \mathrm{poly}(\frac{1}{\varepsilon})$ | $\ell = \mathrm{poly}(n)$ | $q = O(\frac{1}{\delta})$ |
| [MR08][DH09] | $\varepsilon$ | $\lvert\Sigma\rvert = \exp(\frac{1}{\varepsilon})$ | $\ell = \mathrm{poly}(n)$ | $q = 2$ |
| [DHK15] | $\varepsilon \geq \mathrm{poly}(\frac{1}{n})$ | $\lvert\Sigma\rvert = n^{\frac{1}{\mathrm{polyloglog}\,n}}$ | $\ell = \mathrm{poly}(n)$ | $q = \mathrm{polyloglog}\,n$ |

Several of these are achieved via high-soundness composition & high-soundness ingredients.

# Sliding Scale Conjecture

The prevailing belief is that soundness error $\varepsilon$ is achievable via an alphabet of size $\text{poly}(\frac{1}{\varepsilon})$.
This was formulated in a conjecture by Bellare, Goldwasser, Lund, Russell in 1993:

<u>Sliding Scale Conjecture</u> $\exists$ constant $q_0 \in \mathbb{N}$ $\forall\, \varepsilon \geq \frac{1}{\text{poly}(n)}$

$$NP \subseteq PCP\left[\varepsilon_c = 0,\ \varepsilon_s = \varepsilon,\ \Sigma = \{0,1\}^{O(\log \frac{1}{\varepsilon})},\ \ell = \text{poly}(n),\ q = q_0,\ r = O(\log n)\right]$$

Leads to asymptotically shorter succinct arguments (fewer queries for same security level).
Implies optimal hardness of approximation results for several problems of interest (such as directed sparsest cut, directed multicut and more if PCP is a "projection" game).

The "sliding" refers to the parameter $\varepsilon$ that can move anywhere in the interval $\left[\frac{1}{\text{poly}(n)}, 1\right)$.

Next we build intuition for why the conjecture looks like this.
E.g., why can't we expect $\varepsilon = 2^{-\sqrt{n}}$ with a large enough alphabet ($\sim 2^{\sqrt{n}}$)?

# Intuition for Formulation of Conjecture

<u>Sliding Scale Conjecture</u> $\exists$ constant $q_0 \in \mathbb{N}$ $\forall\, \varepsilon \geq \frac{1}{\mathrm{poly}(n)}$

$$NP \subseteq PCP[\varepsilon_c = 0, \varepsilon_s = \varepsilon, \Sigma = \{0,1\}^{O(\log \frac{1}{\varepsilon})}, \ell = \mathrm{poly}(n), q = q_0, r = O(\log n)]$$

Why does the conjecture look like this?

Suppose that $L \in PCP[\varepsilon_c = 0, \varepsilon_s = \varepsilon, \Sigma, \ell, q, r]$ via a PCP system $(P,V)$.

Observation:

- if $\exists\, x \notin L, \rho \in \{0,1\}^r, \pi \in \Sigma^\ell$ s.t. $V^\pi(x;\rho) = 1$ then $\varepsilon \geq 2^{-r}$
- if $\exists\, x \notin L\ \forall \rho \in \{0,1\}^r\ \exists\, \pi \in \Sigma^\ell$ s.t. $V^\pi(x;\rho) = 1$ then $\varepsilon \geq |\Sigma|^{-q}$ (pick a random local view)

Moreover we may assume that $\exists\, x \notin L\ \forall \rho \in \{0,1\}^r\ \exists\, \pi \in \Sigma^\ell$ s.t. $V^\pi(x;\rho) = 1$, because if not:

<u>lemma</u>: If $\forall x \notin L\ \exists\, \rho \in \{0,1\}^r\ \forall \pi \in \Sigma^\ell\ V^\pi(x;\rho) = 0$ then $L \in \mathrm{DTime}(\exp(r + q \log|\Sigma|))$.

<u>proof</u>: By perfect completeness, $\forall x \in L\ \exists\, \pi \in \Sigma^\ell\ \forall \rho \in \{0,1\}^r\ V^\pi(x;\rho) = 1$. Hence the decider works as follows:

$D(x) :=$ For $\rho \in \{0,1\}^r$ : {if all local views in $\Sigma^q$ reject then output 0}. Else output 1. ■

We deduce that $\varepsilon \geq \max\{2^{-r}, |\Sigma|^{-q}\}$ (and hence $|\Sigma| \geq (\frac{1}{\varepsilon})^{\frac{1}{q}}$), so that $\frac{1}{\mathrm{poly}(n)} \leq \varepsilon \leq 1$

when $r = O(\log n)$, $q = O(1)$, $|\Sigma| = \mathrm{poly}(\frac{1}{\varepsilon}) = 2^{O(\log \frac{1}{\varepsilon})}$.

But what if $r = \omega(\log n)$, $|\Sigma| = \omega(\log n)$, or $\varepsilon_c > 0$?

# Limitations for High-Soundness PCPs

The amount of information read by a PCP verifier is $q \cdot \log|\Sigma|$ bits.

This is interesting for NP languages when $q \cdot \log|\Sigma| \ll n$ (as reading an n-bit witness has no soundness error).

In this regime the soundness error must be $\Omega(2^{-q \log \ell})$:

<u>theorem:</u> Assuming the (randomized) exponential-time hypothesis,
3SAT does not have PCPs where $q \cdot (\log \ell + \log|\Sigma|) = o(n)$ and $\varepsilon = o(2^{-q \log \ell})$.

In particular, for $\ell = \mathrm{poly}(n)$ and $q = O(1)$ we get $\varepsilon \geq \mathrm{poly}(\frac{1}{n})$.

In other words in this regime we cannot expect exponentially-small error, regardless of alphabet size.

The theorem follows from a generic lemma that gives "algorithms for PCPs":

<u>lemma:</u> Suppose that $L \in \mathrm{PCP}[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, r]$. If $\varepsilon_s < (1-\varepsilon_c) \cdot 2^{-q \cdot \log \ell}$ then

$$L \in \mathrm{BPTime}\left[\exp\left(q \cdot (\log \ell + \log|\Sigma|) + \log \frac{1}{(1-\varepsilon_c) 2^{-q \log \ell} - \varepsilon_s}\right)\right].$$

Proof has two steps: ① from PCP to laconic MA protocol
② from laconic MA protocol to BP algorithm

# Step 1: from PCP to Laconic MA

can improve to $2^{-h}$ where $h$ is "query entropy"

**Lemma:** Suppose that $L \in PCP[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, r]$. If $\varepsilon_s < (1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell}$ then $L$ has an MA proof with $\varepsilon_c' = 1-(1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell}$, $\varepsilon_s' = \varepsilon_s$, and $pc = q\cdot(\log\ell + \log|\Sigma|)$.

**proof:** Let $(P_{PCP}, V_{PCP})$ be the PCP for $L$. We construct the MA protocol $(P_{MA}, V_{MA})$ as follows:

$P_{MA}(x)$

1. Compute $\Pi := P_{PCP}(x)$.
2. Guess query set $Q \subseteq [\ell]$.
3. Send $\pi = (Q, \Pi[Q])$.

$V_{MA}(x, \tilde{\pi} = (\tilde{Q}, \tilde{\Pi}[\tilde{Q}]))$

1. Sample $\rho \in \{0,1\}^r$.
2. Run $V_{PCP}(x;\rho)$ and answer query $i \in \tilde{Q}$ with $\tilde{\Pi}[\tilde{Q}]$. (If any query is outside $\tilde{Q}$ then reject.)

**Completeness:** If $x \in L$ then, for $\Pi := P_{PCP}(x)$, $\Pr_\rho[V_{PCP}^{\Pi}(x;\rho)] \geq 1-\varepsilon_c$. With probability $\geq \binom{\ell}{q}^{-1} \geq 2^{-q\log\ell}$ $P_{MA}$ guesses the correct query set. Hence $\Pr_{Q,\rho}[V_{MA}(x,(Q,\Pi[Q]))=1] \geq (1-\varepsilon_c)\cdot 2^{-q\log\ell}$.

**Soundness:** Suppose that for $x \notin L$ there is $\tilde{\pi} = (\tilde{Q}, \tilde{\Pi}[\tilde{Q}])$ s.t. $\Pr_\rho[V_{MA}(x,\tilde{\pi})=1] > \varepsilon_s$. Then for $\tilde{\Pi} :=$ "equal to $\tilde{\Pi}[\tilde{Q}]$ on $\tilde{Q}$ and arbitrary outside of $\tilde{Q}$" it holds that $\Pr_\rho[V_{PCP}^{\tilde{\Pi}}(x)=1] > \varepsilon_s$ (contradiction).

**Prover communication:** $|\pi| = |Q| + |\Pi[Q]| = q\cdot\log\ell + q\cdot\log|\Sigma|$.

$\blacksquare$

# Step 2: from Laconic MA to Algorithm

**lemma:** If $L$ has an MA protocol with completeness error $\varepsilon_c$, soundness error $\varepsilon_s$, and prover communication $pc$ then $L \in \mathrm{BPTime}\left[2^{O(pc)} \mathrm{poly}\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)\right]$.

**proof:** Estimate the acceptance probability for every possible MA proof.

$A(x) :=$
1. For every possible MA proof $\tilde{\pi}$:
   1.1. Sample $\rho_1, \ldots, \rho_t \in \{0,1\}^r$ and compute $N(\tilde{\pi}) := |\{i \in [t] \mid V_{MA}(x, \tilde{\pi}; \rho_i) = 1\}|$.
   1.2. If $N(\tilde{\pi})/t > (1-\varepsilon_c) - \frac{1-\varepsilon_c-\varepsilon_s}{2}$ then output 1.
2. Output 0.

For $\tilde{\pi}$ and $\rho$ let $Z(\tilde{\pi}, \rho)$ be the indicator that $V_{MA}(x, \tilde{\pi}, \rho) = 1$.

Note that $Z(\tilde{\pi}, \rho_1), \ldots, Z(\tilde{\pi}, \rho_t)$ are i.i.d. samples from Bernoulli distribution with bias $p(\tilde{\pi}) := \Pr_\rho[V_{MA}(x, \tilde{\pi}) = 1]$.

By an additive Chernoff bound $\Pr_{\rho_1, \ldots, \rho_t}\left[\left|\frac{1}{t}\sum_{i=1}^{t} Z(\tilde{\pi}, \rho_i) - p(\tilde{\pi})\right| > \alpha\right] \leq \exp(-t\alpha^2)$.

If $x \in L$ then $\exists \pi$ s.t. $p(\pi) \geq 1-\varepsilon_c$.
If $x \notin L$ then $\forall \tilde{\pi}$ $p(\tilde{\pi}) \leq \varepsilon_s$.

To distinguish between these we need $\alpha < \frac{1}{2}((1-\varepsilon_c) - \varepsilon_s)$ and $t = O(\frac{1}{\alpha^2} \cdot pc)$ so the error is $O(\frac{1}{2^{pc}})$ for a union bound on all $\tilde{\pi}$.

We conclude that for $t = O\left(\frac{1}{(1-\varepsilon_c-\varepsilon_s)^2} \cdot pc\right)$ the algorithm $A$ has constant 2-sided error. ■

# Limitations for High-Soundness IOPs

Can we hope for significantly better soundness error via IOPs instead of PCPs?

The answer is, to a first order, NO.

The reason is that one can design similarly efficient "algorithms for IOPs".

In more detail, similarly to a PCP, the amount of information read by an IOP verifier is $q \cdot \log|\Sigma|$ bits.

This is interesting for NP languages when $q \cdot \log|\Sigma| \ll n$ (as reading an n-bit witness has no soundness error).

And, similarly to before, in this regime the soundness error must be $\Omega(2^{-q \log \ell})$.

The technical lemma is as follows:

Lemma: Suppose that $L \in \mathrm{IOP}[\varepsilon_c, \varepsilon_s, k, \Sigma, \ell, q, r]$ (with public coins). If $\varepsilon_s < (1-\varepsilon_c) \cdot 2^{-q \cdot \log \ell}$ then

$$L \in \mathrm{BPTime}\left[\exp\left(q \cdot (\log \ell + \log|\Sigma|) + k \cdot \log \frac{k}{(1-\varepsilon_c) 2^{-q \log \ell} - \varepsilon_s}\right)\right].$$

Proof has two steps:

1. from (public-coin) IOP to laconic (public-coin) IP protocol
2. from laconic (public-coin) IP protocol to BP algorithm