Lecture A.10

Limitations of PCPs and IOPs

Summer Graduate School on
Foundations and Frontiers of Probabilistic Proofs
2021.08.06

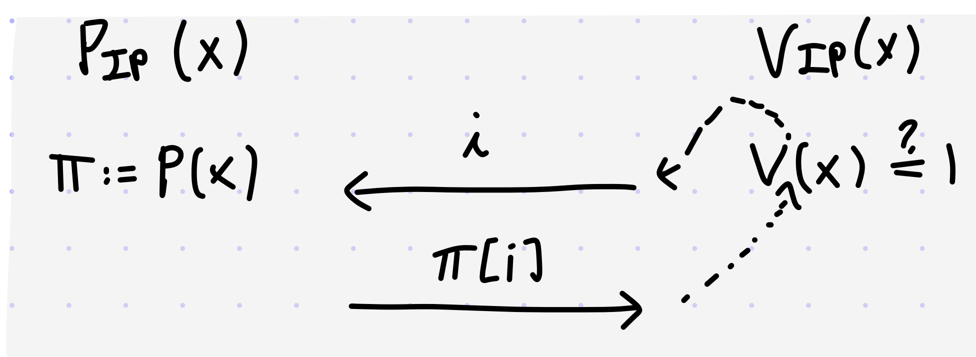
Limits on Query Complexity

We almost 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):

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

proof: 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:

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

proof: Apply the trivial query bundling to the PCP Theorem. ■

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

Small Query Complexity and Small Soundness Error?

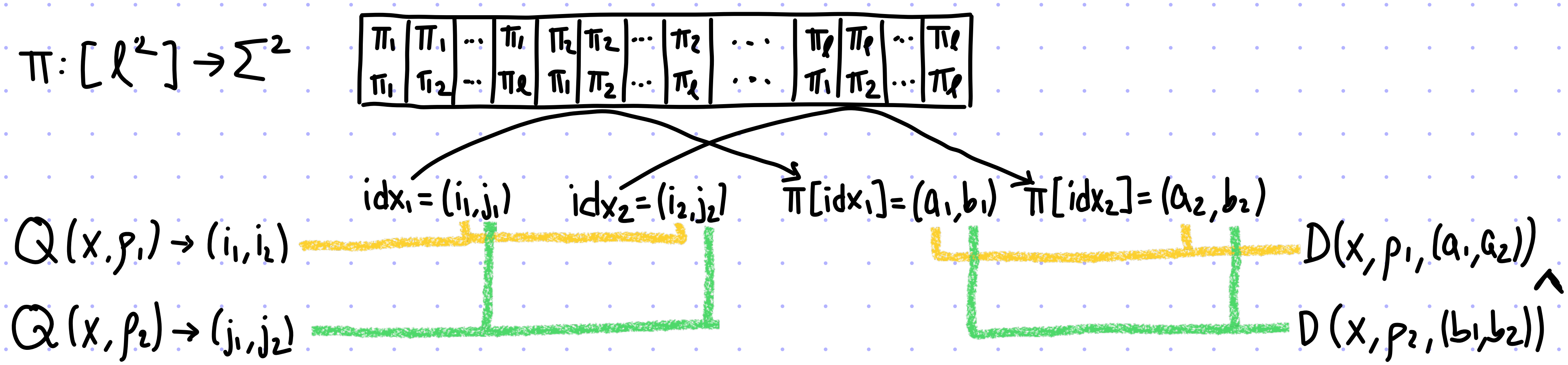
Repeating the PCP verifier reduces soundness error but also increases query complexity:

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

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 and 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.
We expect the t-wise parallel repetition to yield this inclusion:

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

BUT: the conjecture that $\varepsilon_s' = \varepsilon_s^t$ is false in general.

We know that ε_s' tends to 0 as t tends to infinity. (Proof via Ramsey Theory!)
And, if $q=2$, that $\varepsilon_s' = \varepsilon_s^{\Omega(t/\log|\Sigma|)}$. (Elaborate proof via Information Theory.)
Applied to the PCP Theorem, this yields soundness error ε over an alphabet of size $|\Sigma| = \mathrm{poly}(\frac{1}{\varepsilon})$:

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

The main limitation is that proof length becomes $\ell = n^{O(\log\frac{1}{\varepsilon})}$ so that if we want $\ell = \mathrm{poly}(n)$ then parallel repetition does not tell us anything for $\varepsilon = o(1)$.

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

Sliding Scale Conjecture

The prevailing belief is that soundness error ε 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 $q_0 \in \mathbb{N}$ $\forall \varepsilon \geq \frac{1}{\text{poly}(n)}$

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

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 ε that can move anywhere in the interval $[\frac{1}{\text{poly}(n)}, 1)$.

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 Σ^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 $\Omega(2^{-q\log \ell})$:

In worksheet B.1 we saw a weaker result: given perfect completeness and little randomness

theorem: 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 $\ell = \text{poly}(n)$ and $q = O(1)$ we get $\varepsilon \geq \text{poly}(\frac{1}{n})$.

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":

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

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

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

■

Step 2: from Laconic MA to Algorithm

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

<u>proof</u>: 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 ρ 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