# Foundations of Probabilistic Proofs

A course by **Alessandro Chiesa**

## Lecture 24

# Limitations of PCPs and IOPs

# Limits on Proof Length

A PCP verifier can treat the PCP string as an MA proof string (read it in full).

In particular, $PCP[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, r, vt] \subseteq MA[\varepsilon_c, \varepsilon_s, pc = \ell \cdot \log|\Sigma|, r, vt' = vt + \ell \cdot \log|\Sigma|]$.

Hence, PCP strings are at least as long as MA proof strings:

they inherit the limitations on proof length of MA proof strings.

We proved that $MA[\varepsilon_c, \varepsilon_s, pc, vt] \subseteq BPTIME\left(2^{O(pc)} \cdot poly\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, vt\right)\right)$.

Hence $PCP[\varepsilon_c, \varepsilon_s, \Sigma, \ell, vt] \subseteq BPTIME\left(2^{O(\ell \cdot \log|\Sigma|)} \cdot poly\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, vt\right)\right)$.

We deduce, e.g., that a PCP for 3SAT with $\ell \cdot \log|\Sigma| = o(\#\text{variables})$ violates RETH.

PCPs may have ADDITIONAL limitations on proof length.

Example: Every NP relation $R = \{(x,w): \dots\}$ has a PCP with $\ell = poly(|x|)$.

Can we achieve PCPs with $\ell = poly(|w|)$? (We expect $\ell \geq |w|$ by the above arguments.)

theorem: If $SAT \in PCP[\varepsilon_c = 0, \varepsilon_s = \frac{1}{2}, \Sigma = \{0,1\}, \ell = poly(\#\text{variables}), q = O(1)]$

then $NP \subseteq coNP/_{poly}$ (and so PH collapses).

This question is related to INSTANCE COMPRESSION for NP.

# Limits on Query Complexity [1/2]

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

Q: How small can query complexity be?

Hard languages are unlikely to have one-query PCPs:

**lemma:** $PCP[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q=1, r] \subseteq BPTIME\left(2^{O(\log|\Sigma| + \log \ell)} \cdot \text{poly}\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)\right)$

**proof:** We prove that $PCP[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q=1, r] \subseteq IP[\varepsilon_c, \varepsilon_s, K=1, pc = \log|\Sigma|, vc = \log \ell]$.

Consider the following (private-coin 1-round) IP protocol.

| $P_{IP}(x)$ | | $V_{IP}(x)$ |
|---|---|---|
| | | Sample $\rho \in \{0,1\}^r$ |
| $\pi := P_{PCP}(x) \in \Sigma^\ell$ | $\xleftarrow{i}$ | Compute $i := Q_{PCP}(x, \rho) \in [\ell]$. |
| $a := \pi[i] \in \Sigma$ | $\xrightarrow{a}$ | Check that $D_{PCP}(x, \rho, a) = 1$. |

The lemma then follows from $IP[\varepsilon_c, \varepsilon_s, pc, vc] \subseteq BPTIME\left(2^{O(pc+vc)} \cdot \text{poly}\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)\right)$. ■

EXAMPLE: a one-query PCP for 3SAT with $\log|\Sigma| = o(\#\text{variables})$ contradicts RETH.

if $\log|\Sigma| = \#$ variables then 1 symbol can encode a candidate assignment

# Limits on Query Complexity [2/2]

The situation for TWO-query PCPs is different.

- Non-adaptive two-query PCPs over the binary alphabet are unlikely.

**lemma:** $PCP[\varepsilon_c = 0, \varepsilon_s < 1, \Sigma = \{0,1\}, \ell, \overset{\text{non-adaptive}}{q=2}, r] \subseteq DTIME(2^r \cdot vt + poly(\ell))$ ← in P if $\ell = poly(n)$ and $r = O(\log n)$

**proof:** View a candidate PCP string as $\ell$ variables $z_1, \ldots, z_\ell$.

The decision of $V(x;\rho)$ is a function $\varphi_{x,\rho}(z_1, \ldots, z_\ell)$ that depends on 2 variables.

- If $x \in L$ then $\exists a_1, \ldots, a_\ell \in \{0,1\}$ $\bigwedge_{\rho \in \{0,1\}^r} \varphi_{x,\rho}(a_1, \ldots, a_\ell) = 1$.

- If $x \notin L$ then $\forall a_1, \ldots, a_\ell \in \{0,1\}$ $\frac{1}{2^r} \cdot |\{\rho \in \{0,1\}^r : \varphi_{x,\rho}(a_1, \ldots, a_\ell) = 1\}| \leq \varepsilon_s < 1$.

Deciding between these is a 2SAT instance with $\ell$ variables and $2^r$ clauses.

There is a poly(n,m)-time algorithm to decide n-variable m-clause 2SAT. ■

- There are (non-adaptive) two-query PCPs over larger alphabets.

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

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

$$PCP[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, r] \subseteq PCP[\varepsilon_c, \varepsilon_s' = 1 - \tfrac{1-\varepsilon_s}{q}, \Sigma' = \Sigma^q, \ell' = \ell + 2^r, q' = 2, r' = r + \log q]$$

# Limits on Soundness: the Bit Barrier [1/2]

A PCP verifier reads $q \cdot \log|\Sigma|$ bits from the PCP string.

For NP languages this is interesting when $q \cdot \log|\Sigma| \ll |\text{witness}|$.

In this regime the soundness error must be $\Omega(|\Sigma|^{-q})$.

Note: reading the witness achieves soundness error $\varepsilon_s = 0$.

**theorem:** Assuming RETH (the randomized exponential-time hypothesis), 3SAT does not have a PCP where $q \cdot \log|\Sigma| = o(n)$ and $\varepsilon = o(|\Sigma|^{-q})$.

The theorem is based on a counting argument for "always rejecting" randomness.

**def:** $\text{Ans}(x, \pi, r) \in \Sigma^q$ are the $q$ symbols in $\Sigma$ read by $V^{\pi}(x; r)$ for a given $\pi \in \Sigma^\ell$ and $r \in \{0,1\}^{vr}$.

**def:** For $a \in \Sigma^q$, $V^{[a]}(x; r)$ is defined as $V^{\pi}(x; r)$ if $\exists \pi \in \Sigma^\ell$ s.t. $\text{Ans}(x, \pi, r) = a$ (and 0 otherwise).

(This is well-defined because $\forall\, \pi_1, \pi_2 \in \Sigma^\ell \;\; \forall r \in \{0,1\}^{vr} \;\; \text{Ans}(x, \pi_1, r) = \text{Ans}(x, \pi_2, r) \rightarrow V^{\pi_1}(x; r) = V^{\pi_2}(x; r)$.)

The set of ALWAYS REJECTING randomness is:

$$R(x) := \{ r \in \{0,1\}^{vr} : \forall a \in \Sigma^q \;\; V^{[a]}(x; r) = 0 \}.$$

If $\varepsilon_s$ is small enough then $R(x)$ is non-empty:

**claim:** $\forall x \notin L, \;\; |R(x)| \geq (1 - |\Sigma|^q \cdot \varepsilon_s) \cdot 2^{vr}$

# Limits on Soundness: the Bit Barrier [2/2]

**claim:** $\forall x \notin L,\ |R(x)| \geq (1 - |\Sigma|^q \cdot \varepsilon_s) \cdot 2^{vr}$

**proof:** Define $S(x) := \{(\pi, \rho) \in \Sigma^\ell \times \{0,1\}^{vr} : V^\pi(x;\rho) = 1\}$.

By definition of soundness error, $|S(x)| \leq |\Sigma|^\ell \cdot \varepsilon_s \cdot 2^{vr}$.

We argue that $|S(x)| \geq |\{0,1\}^{vr} \setminus R(x)| \cdot |\Sigma|^{\ell - q}$:

$$S(x) = \bigsqcup_{\rho \in \{0,1\}^{vr}} \bigcup_{a \in \Sigma^q} \{(\pi,\rho) : V^{[a]}(x;\rho) = 1 \wedge \mathrm{Ans}(x,\pi,\rho) = a\}$$

so

$$|S(x)| = \sum_{\rho \in \{0,1\}^{vr}} \left| \bigcup_{a \in \Sigma^q} S(x,\rho,a) \right|$$

$$= \sum_{\rho \in \{0,1\}^{vr} \setminus R(x)} \left| \bigcup_{a \in \Sigma^q} S(x,\rho,a) \right| \quad (\rho \in R(x) \to \forall a\ S(x,\rho,a) = \varnothing)$$

$$\geq \sum_{\rho \in \{0,1\}^{vr} \setminus R(x)} \min_{a \in \Sigma^q} |S(x,\rho,a)| \quad (\text{no guarantees that } S(x,\rho,a_1) \text{ and } S(x,\rho,a_2) \text{ differ})$$

$$\geq \sum_{\rho \in \{0,1\}^{vr} \setminus R(x)} |\Sigma|^{\ell - q}. \quad (\text{at most } q \text{ answers of a PCP string are constrained})$$

■

**corollary:** $\varepsilon_c + |\Sigma|^q \varepsilon_s < 1 \to \mathrm{PCP}[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, vr] \subseteq \mathrm{BPTIME}[\varepsilon_c' = \varepsilon_c, \varepsilon_s' = |\Sigma|^q \cdot \varepsilon_s, \mathrm{time} = |\Sigma|^q \cdot vt, \mathrm{rand} = vr]$.

**proof:** Consider the probabilistic decider $A(x) := \bigvee_{a \in \Sigma^q} V^{[a]}(x;\rho)$ (for a random $\rho$). ■

Hence $3\mathrm{SAT} \in \mathrm{PCP}[\varepsilon_c < 1, \varepsilon_s = o(|\Sigma|^{-q}), \Sigma, \ell, q, vr]$ implies $3\mathrm{SAT} \in \mathrm{BPTIME}(2^{o(n)})$, violating RETH.

# Limits on Soundness: the Guessing Barrier

The BIT BARRIER suggests that one may arbitrarily reduce soundness error by increasing alphabet size.

E.g. could one hope to show that $3SAT \in PCP[\varepsilon_c = 0, \varepsilon_s = 2^{-\sqrt{n}}, \Sigma = \{0,1\}^{\sqrt{n}}, \ell = \text{poly}(n), q = O(1)]$?

No, we show a GUESSING BARRIER: the soundness error must be $\Omega(2^{-q \cdot \log \ell})$.

**theorem:** Assuming RETH (the randomized exponential-time hypothesis),

3SAT does not have a PCP where $q \cdot (\log \ell + \log|\Sigma|) = o(n)$ and $\varepsilon = o(2^{-q \cdot \log \ell})$.

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

In this regime we cannot achieve exponentially-small error (regardless of alphabet size).

The theorem follows from a generic lemma about ALGORITHMS FOR PCPs:

**lemma:** Let $L \in PCP[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, r]$. Then

$$\varepsilon_s < (1-\varepsilon_c) \cdot 2^{-q \cdot \log \ell} \longrightarrow L \in \text{BPTIME}\left(2^{O(q \cdot (\log \ell + \log|\Sigma|))} \cdot \text{poly}\left(\frac{1}{(1-\varepsilon_c) \cdot 2^{-q \cdot \log \ell} - \varepsilon_s}, n\right)\right).$$

The 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

**lemma:** Let $L \in \mathrm{PCP}[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, r]$. Then

$$\varepsilon_s < (1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell} \longrightarrow L \in \mathrm{MA}\left[\varepsilon_c' = 1-(1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell},\ \varepsilon_s' = \varepsilon_s,\ pc = q\cdot(\log\ell + \log|\Sigma|),\ vr = r\right]$$

**proof:** The MA protocol is as follows:

$P_{MA}(x)$

1. Compute $\Pi := P_{PCP}(x)$.
2. Sample random $Q \leftarrow \binom{[\ell]}{q}$.
3. Send $\pi := (Q, \Pi[Q])$.

$V_{MA}(x, \pi)$

1. Sample $\rho \in \{0,1\}^r$ and parse $\pi$ as $(Q, a \in \Sigma^Q)$.
2. Run $V_{PCP}(x;\rho)$ and answer query $i \in Q$ with $a[i]$.
   (Reject if any query outside of $Q$.)

**Completeness:** If $x \in L$ then, for $\Pi := P_{PCP}(x)$, $\Pr_\rho[V_{PCP}^{\Pi}(x;\rho)=1] \geq 1-\varepsilon_c$.

With probability $\geq \binom{\ell}{q}^{-1} \geq 2^{-q\cdot\log\ell}$, $P_{MA}$ samples the correct query set.

So $\Pr_{Q,\rho}[V_{MA}(x,(Q,\Pi[Q]))=1] \geq (1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell}$.

**Soundness:** Suppose that for $x \notin L$ there is $\pi = (Q, a \in \Sigma^Q)$ s.t. $\Pr[V_{MA}(x,\pi)=1] > \varepsilon_s$.

For $\Pi =$ "equal to $a$ and arbitrary on $[\ell]\setminus Q$", we have $\Pr[V_{PCP}^{\Pi}(x)=1] > \varepsilon_s$ (a contradiction).

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

# Concluding the Proof of the Lemma

**lemma:** Let $L \in \mathrm{PCP}[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, r]$. Then

$$\varepsilon_s < (1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell} \longrightarrow L \in \mathrm{BPTIME}\left(2^{O(q\cdot(\log\ell+\log|\Sigma|))}\cdot \mathrm{poly}\left(\frac{1}{(1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell}-\varepsilon_s}, n\right)\right).$$

**proof:** The lemma is a direct consequence of the two inclusions below.

① Prior slide: from PCP to laconic MA protocol

$$\varepsilon_s < (1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell} \longrightarrow \mathrm{PCP}[\varepsilon_c, \varepsilon_s, \Sigma, \ell, q, r]$$
$$\subseteq \mathrm{MA}\left[\varepsilon_c' = 1-(1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell}, \varepsilon_s' = \varepsilon_s, pc = q\cdot(\log\ell+\log|\Sigma|), vr = r\right]$$

② Prior lecture: from laconic MA protocol to BP algorithm

$$\mathrm{MA}[\varepsilon_c, \varepsilon_s, pc, vr] \subseteq \mathrm{BPTIME}\left(2^{O(pc)}\cdot \mathrm{poly}\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)\right)$$

■

Remark: The bound $\varepsilon_s < (1-\varepsilon_c)\cdot 2^{-q\cdot\log\ell}$ can straightforwardly be improved to $\varepsilon_s < (1-\varepsilon_c)\cdot 2^{-h}$ where $h$ is the query entropy of the PCP.

The query entropy of $(P,V)$ for $R$ is $h := \max_{(x,w)\in R} \min_{S\subseteq[\ell]} h(x,w,Q) \geq \log\binom{\ell}{q} \geq q\log\ell$ where $h(x,w,Q) := -\log \Pr[V^{\pi}(x) \text{ reads exactly } Q \mid \pi \leftarrow P(x,w)]$ is the entropy of $Q \subseteq [\ell]$ on $(x,w) \in R$.

# Towards The Best Soundness

The BIT BARRIER and GUESSING BARRIER together tell us that for a PCP we expect

$$\varepsilon_s = \Omega\left(\max\{|\Sigma|^{-q}, 2^{-q\log \ell}\}\right).$$

A major open question is achieving this best soundness for polynomial-size PCPs.

The prevailing belief is that for $q=O(1)$ every soundness error $\varepsilon \geq 2^{-q\cdot\log \ell} = 2^{-O(\log \ell)} = \frac{1}{\text{poly}(n)}$ is achievable with alphabet size $|\Sigma| \leq \left(\frac{1}{\varepsilon}\right)^{\frac{1}{q}} = \text{poly}\left(\frac{1}{\varepsilon}\right)$.

SLIDING SCALE CONJECTURE

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

sliding parameter

Such PCPs have applications:

- shorter succinct arguments (fewer PCP queries for the same security level)
- improved hardness of approximation (especially if the PCP is a "projection game")

Applying a technique known as PARALLEL REPETITION to the PCP Theorem achieves the BIT BARRIER but not the GUESSING BARRIER (as proof length blows up):

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

# The Case of IOPs: Proof Length

An IOP verifier can treat an IOP interaction as an IP interaction (read each prover message in full).

In particular, $\mathrm{IOP}[\varepsilon_c, \varepsilon_s, k, \Sigma, \ell, q, r, vt] \subseteq \mathrm{IP}[\varepsilon_c, \varepsilon_s, k, pc = \ell \cdot \log|\Sigma|, r, vt' = vt + \ell \cdot \log|\Sigma|]$.

Hence, proof length of a private-coin/public-coin IOP $\geqslant$ prover-to-verifier communication of the corresponding private-coin/public-coin IP.

IOPs inherit the limitations on communication complexity of IPs.

IOPs may have ADDITIONAL limitations on proof length.

However there is a key difference with PCPs: SAT has an IOP with $\ell = \mathrm{poly}(\#\text{variables})$.

That said $\ell = \mathrm{poly}(|w|)$ is not achievable by IOPs for all NP relations.

**theorem:** If $\mathrm{CSAT} \in \mathrm{IOP}[\varepsilon_c = 0, \varepsilon_s = 1/2, k = O(1), \Sigma = \{0,1\}, \ell = \mathrm{poly}(\#\text{inputs}), q = O(\log \#\text{gates})]$ then "some plausible conjecture about CSAT is false".

# The Case of IOPs: Soundness

Can we hope for significantly smaller soundness error via IOPs compared to PCPs?

The answer is NO (to a first order) because similar soundness barriers hold.

An IOP verifier reads $q \cdot \log|\Sigma|$ bits from the IOP strings.

For NP languages this is interesting when $q \cdot \log|\Sigma| \ll |\text{witness}|$.

Note: reading the witness achieves soundness error $\varepsilon_s = 0$.

In this regime the soundness error must satisfy $\varepsilon_s = \Omega(\max\{|\Sigma|^{-q}, 2^{-q\log\ell}\})$.

First we establish a BIT BARRIER.

**lemma:** $\varepsilon_c + |\Sigma|^q \varepsilon_s < 1 \rightarrow \text{IOP}[\varepsilon_c, \varepsilon_s, k, \Sigma, \ell, q, vr] \subseteq \text{BPTIME}[\varepsilon_c' = \varepsilon_c, \varepsilon_s' = |\Sigma|^q \cdot \varepsilon_s, \text{time} = |\Sigma|^q \cdot vt, \text{rand} = vr]$.

**proof:** Unroll the IOP into a (very long!) PCP.

This preserves all parameters but for proof length. Invoke the BIT BARRIER for PCPs. ■

Hence $3SAT \in \text{IOP}[\varepsilon_c < 1, \varepsilon_s = o(|\Sigma|^{-q}), k, \Sigma, \ell, q, vr]$ implies $3SAT \in \text{BPTIME}(2^{o(n)})$, violating RETH.

The takeaway is that the round complexity $k$ does not affect the BIT BARRIER.

# The Case of IOPs: Soundness

We can also establish a GUESSING BARRIER, but this requires more care.

By designing suitable ALGORITHMS FOR IOPs we can show that $\varepsilon_s = \Omega(2^{-q \cdot \log \ell})$.

The generic technical lemma is as follows.

**lemma:** Let $L \in \mathrm{IOP}[\varepsilon_c, \varepsilon_s, k, \Sigma, \ell, q, r, \text{public-coin}]$. Then

$$\varepsilon_s < (1-\varepsilon_c) \cdot 2^{-q \cdot \log \ell} \longrightarrow L \in \mathrm{BPTIME}\left( 2^{O(q \cdot (\log \ell + \log |\Sigma|))} \cdot \mathrm{poly}\left( \frac{\min\{k,q\}}{(1-\varepsilon_c) \cdot 2^{-q \cdot \log \ell} - \varepsilon_s}, n \right)^{\min\{k,q\}} \right).$$

The 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

# Bibliography


**Limitations of PCPs and IOPs**

- [BGLR 1993]: Efficient probabilistically checkable proofs and applications to approximation, by Mihir Bellare, Shafi Goldwasser, Carsten Lund, Alexander Russell.
- [CY 2020]: Barriers for succinct arguments in the random oracle model, by Alessandro Chiesa, Eylon Yogev.
- [ACY 2021]: A PCP theorem for interactive proofs and applications, by Gal Arnon, Alessandro Chiesa, Eylon Yogev.
- [ABCY 2022]: A toolbox for barriers on interactive oracle proofs, by Gal Arnon, Amey Bhangale, Alessandro Chiesa, Eylon Yogev.

**High soundness PCPs and IOPs**

- [Moshkovitz 2019]: Sliding scale conjectures in PCP, by Dana Moshkovitz.
- [DFKRS 1999]: PCP characterizations of NP: towards a polynomially small error probability, by Irit Dinur, Eldar Fischer, Guy Kindler, Ran Raz, Shmuel Safra.
- [Moshkovitz 2016]: Low degree test with polynomially small error, by Dana Moshkovitz.
- [DHK 2015]: Polynomially low error PCPs with polyloglog n queries via modular composition, by Irit Dinur, Prahladh Harsha, Guy Kindler.
- [ACY 2023]: IOPs with inverse polynomial soundness error, by Gal Arnon, Alessandro Chiesa, Eylon Yogev.