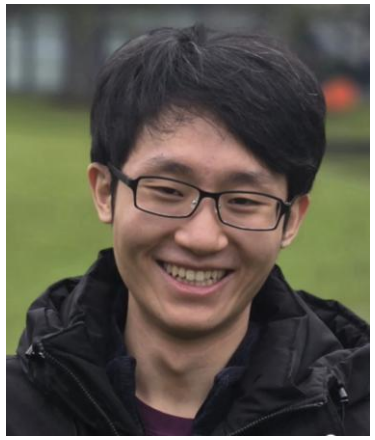
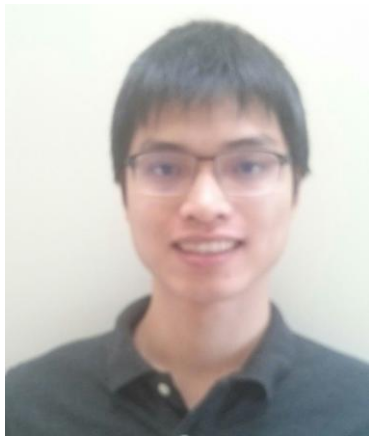
Hardness of Computing Nondeterministic Kolmogorov Complexity

Jinqiao Hu
University of Warwick



Zhenjian Lu
University of Victoria



Igor Oliveira
University of Warwick

Background: Meta-complexity

- Meta-complexity: studying the problems that are themselves about complexity.

MCSP: Estimating the circuit complexity of a given function.

MINKT: Estimating the time-bounded Kolmogorov complexity of a given string.

$$\mathrm{K}^t(x) := \min\{|\Pi| : \Pi \text{ prints } x \text{ within } t \text{ steps.}\}$$

- Meta-complexity has found its connection to **learning theory**, **cryptography** and **average-case complexity** over the last decade.
  - NP-hardness of GapMINKT would rule out Heuristica (where NP is worst-case hard, but average-case easy)! [Hir'18, Hir'20]
  - Hardness of certain meta-complexity problem (agnostic learning) characterizes Pessiland (where NP is average-case easy, but one-way functions don't exist) [HN'26]

Recent Hardness Results in Meta-complexity

[Mas'79] DNF-MCSP is NP-hard.

[Ila'20] AC0-MCSP is NP-hard.

[Hir'20] MINcKT$^{\mathrm{SAT}}$ is NP-hard.

[Hir'22] Partial versions of MCSP/MINKT are NP-hard.

[HI'25] MCSP is NP-hard under some cryptographic assumptions.

However, the central questions, NP-hardness of MCSP/MINKT remains open!

Main Result: Hardness of non-deterministic Kolmogorov complexity

$$\mathrm{nK}^t(x) := \min\left\{ |\Pi| \;\middle|\; \begin{array}{l} \exists w \in \{0,1\}^t, \Pi(w) \text{ prints } x \text{ in } t \text{ steps, and} \\ \forall w \in \{0,1\}^t, \Pi(w) \text{ prints } x \text{ or } \perp \text{ in } t \text{ steps.} \end{array} \right\}$$

- MINnKT: given string x and parameters s, t (t is given in unary), decide whether $\mathrm{nK}^t(x) \leq s$ or $\mathrm{nK}^t(x) > s$. Then:

MINnKT ∈ BPP ⇔ NP ⊆ BPP.

Other results

- We define new notions pnK^t, which is central in the proof of the main theorem, and prove it has optimal language compression and conditional coding, and ℓ-nK^t, which admits efficient weak symmetry of information.
- We prove the average-case hardness of MINnKT.

$$(\text{MINnKT}, \mathcal{U}) \in \text{AvgBPP} \Longrightarrow \text{DistNP} \subseteq \text{AvgBPP}$$

- We prove that if MINnKT is average-case easy, then NP can be solved in subexponential time in the worst-case.

$$(\text{MINnKT}, \mathcal{U}) \in \text{AvgBPP} \Longrightarrow \text{NP} \subseteq \text{RTIME}\left[2^{O(n/\log n)}\right]$$

  - This improves the previous result of [Hir'21, GKLO'22], where the condition on the left is Dist$\Sigma_2^{\text{P}} \subseteq$ AvgBPP.

Proof overview of main result

The easy direction: Easiness of NP implies easiness of MINnKT

- By [Ko'82], NP $\subseteq$ BPP implies PH $\subseteq$ BPP.
- Not hard to show MINnKT $\in \Sigma_2^{\mathrm{P}}$ from definition.

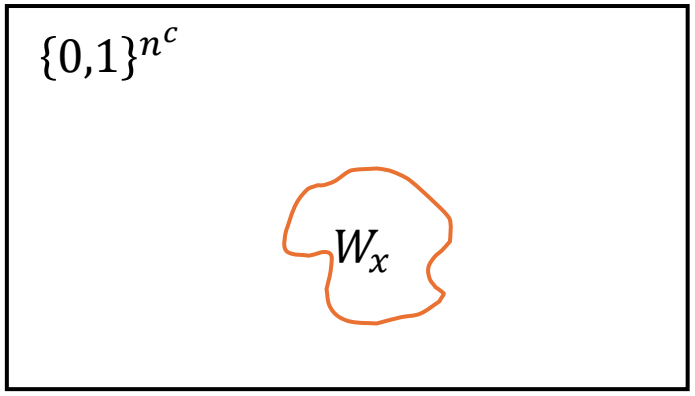
$$\mathsf{nK}^t(x) \leq s \iff \exists |\Pi| \leq s, (\exists |w| \leq t, \Pi^t(w) = x) \wedge (\forall |w| \leq t, \Pi^t(w) \in \{x, \perp\})$$

The hard direction: Easiness of MINnKT implies easiness of NP

- How to show NP $\subseteq$ BPP, assuming that MINnKT $\in$ BPP?
- Fix an NP language L, and an input x. Deciding whether $x \in L$?

First attempt: Try to guess a witness by sampling uniformly!

But the witness might be exponentially sparse!



Let W_x be the set of witnesses for input x

Possible that $|W_x| \leq \frac{2^{n^c}}{2^n}$!

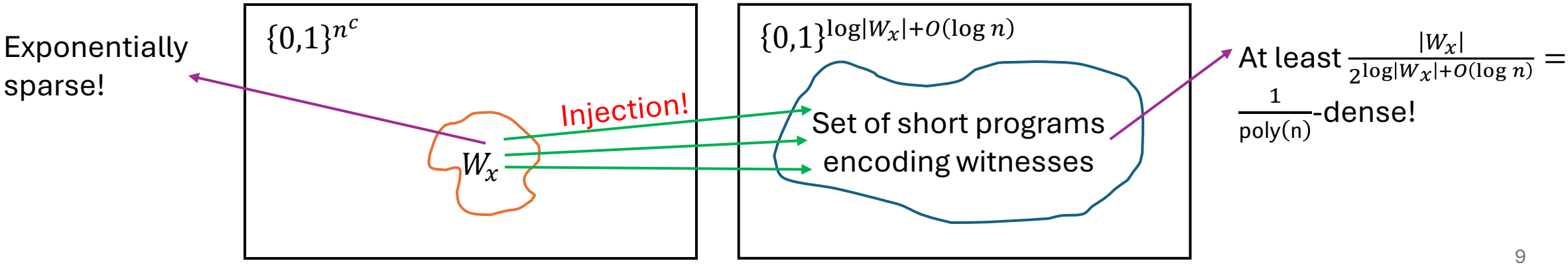
Sampling the compressed witnesses

$K^t(y \mid x)$
$= \min\{|\Pi| : \Pi(x) \text{ prints } y \text{ in } t \text{ steps.}\}$

- Let W_x be the set of witnesses for input x
- If, magically 🪄, for any $y \in W_x$, we have $K^{\text{poly}(n)}(y \mid x) \leq \log|W_x| + O(\log n)$, then:

Sampling Π uniformly from all programs of size at most $\log|W_x| + O(\log n)$, obtain $\hat{y}$ as the output of $\Pi(x)$ running for $\text{poly}(n)$ steps, then check if $\hat{y}$ is a witness for x

finds a witness with probability at least $1/\text{poly}(n)$ (When $W_x \neq \emptyset$), and therefore by error reduction, NP ⊆ BPP!

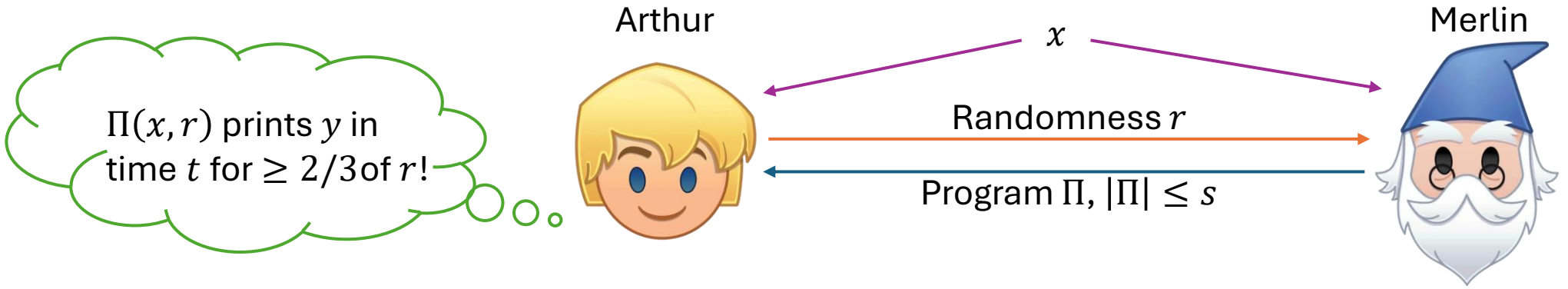


The probabilistic Kolmogorov complexity: pK

$$\mathrm{pK}^t(y \mid x) := \min\left\{ s \;\middle|\; \Pr_r\left[\exists |\Pi| \le s, \Pi(x,r) \text{ prints } y \text{ in } t \text{ steps}\right] \ge \frac{2}{3} \right\}$$

- Or alternatively,

$$\mathrm{pK}^t(y \mid x) := \min\left\{ s \;\middle|\; \Pr_r\left[\mathrm{K}^t(y \mid r, x) \le s\right] \ge \frac{2}{3} \right\}$$



- In fact, the witness sampling in the previous slide would still work, even if we replace K^t by pK^t! ($\mathrm{pK}^{\mathrm{poly}(n)}(y \mid x) \le \log|W_x| + O(\log n)$ for every $y \in W_x$)
  - The sampler samples a program Π, as well as the randomness r, runs $\Pi(x,r)$ for t steps to obtain $\hat{y}$.
  - The (Π, r) pair encoding a witness is dense!

Goal: Show every witness has small conditional pK complexity

For any language $L \in \text{NP}$, for any $x \in L, y \in W_x$, we have $\text{pK}^{\text{poly}(n)}(y \mid x) \leq \log|W_x| + O(\log n)$.

- What we have shown so far:

If NP admits pK^t witness compression, then NP $\subseteq$ BPP.

- What we need to show:

If MINnKT $\in$ BPP, then NP admits pK^t witness compression.

- This looks exactly like the **language compression** property of Kolmogorov complexity!

Non-deterministic Komogorov complexity and language compression

Π has oracle access to membership of S

[BLvM'05] For any set $S \subseteq \{0,1\}^n$ and string $y \in S$, we have

$$\mathrm{nK}^{S\,\mathrm{poly}(n)}(y) \leq \log|S| + O\left(\left(\sqrt{\log|S|} + \log n\right)\log n\right).$$

- When S is equal to some NP witness set W_x, membership of W_x can be tested in poly time given x. Therefore we have

$$\mathrm{nK}^{\mathrm{poly}(n)}(y \mid x) \leq \log|W_x| + O\left(\left(\sqrt{\log|W_x|} + \log n\right)\log n\right).$$

But what we want is:

$$\mathrm{pK}^{\mathrm{poly}(n)}(y \mid x) \leq \log|W_x| + O(\log n).$$

Qualitative gap!

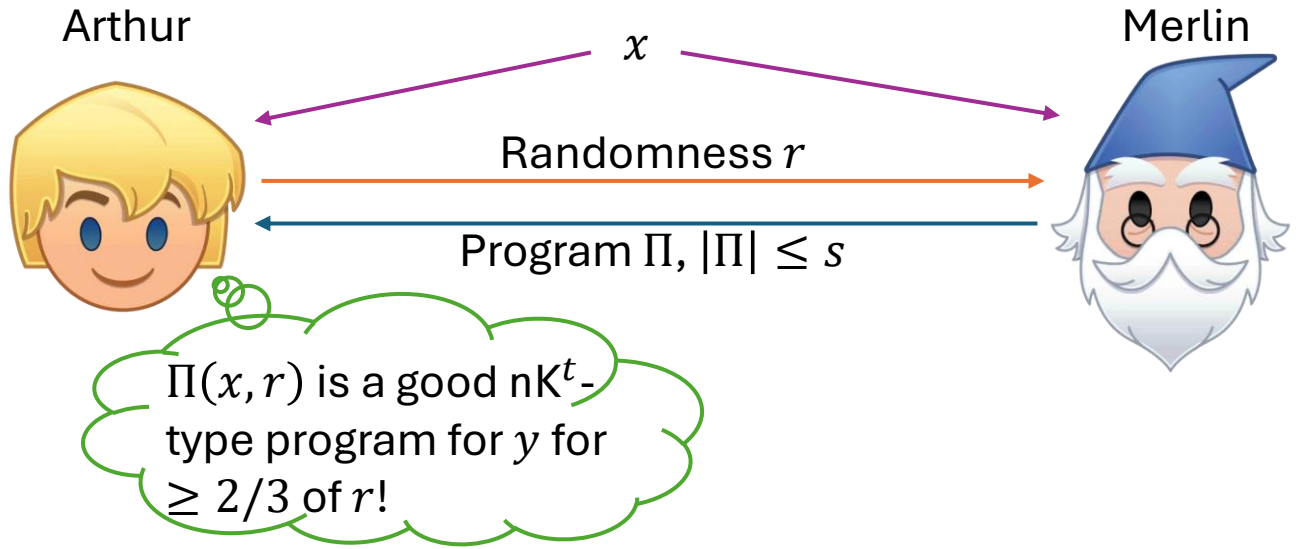
Quantitative gap!

Tool: Probabilistic nondeterministic Kolmogorov complexity pnK

$$\mathsf{pnK}^t(y \mid x) := \min\left\{ s \;\middle|\; \Pr_r\left[\exists |\Pi| \le s, \begin{cases} \exists w, \Pi(x,r,w) \text{ prints } y \text{ in } t \text{ steps}, \\ \forall w, \Pi(x,r,w) \text{ prints } y \text{ or } \perp \text{ in } t \text{ steps}. \end{cases} \right] \ge \frac{2}{3} \right\}$$

- Or equivalently,

$$\mathsf{pnK}^t(y \mid x) := \min\left\{ s \;\middle|\; \Pr_r\left[\mathsf{nK}^t(y \mid x, r) \le s \right] \ge \frac{2}{3} \right\}$$



Optimal language compression for pnK

- Using tools from 2-universal hashing, we are able to prove:

For any set $S \subseteq \{0,1\}^n$ and string $y \in S$, we have

$$\mathrm{pnK}^{S,\mathrm{poly}(n)}(y) \leq \log|S| + O(\log n).$$

- When S is equal to some NP witness set W_x, membership of W_x can be tested in poly time given x. Therefore we have

$$\mathrm{pnK}^{\mathrm{poly}(n)}(y \mid x) \leq \log|W_x| + O(\log n).$$

Optimal!

Only a qualitative gap remains! (We want pK instead of pnK.)

Goal: Collapse pnK to pK

- What we have shown so far:

If NP admits pK^t witness compression, then NP $\subseteq$ BPP.

NP admits pnK^t witness compression unconditionally.

- What we need to show:

If MINnKT $\in$ BPP, then $\mathrm{pK}^{\mathrm{poly}(t)}(y \mid x) \lesssim \mathrm{pnK}^t(y \mid x)$.

“Strong” symmetry of information for pnK

- Using the direct product generator [Hir’20], we are able to prove:

If MINnKT $\in$ BPP, then for every string x, y and time bound t, we have $\mathrm{pnK}^{t}(x,y) \gtrsim \mathrm{pK}^{\mathrm{poly}(t)}(x) + \mathrm{pK}^{\mathrm{poly}(t)}(y \mid x)$.

Rearranging the inequality:

$$\mathrm{pK}^{\mathrm{poly}(t)}(y \mid x) \lesssim \mathrm{pnK}^{t}(x,y) - \mathrm{pK}^{\mathrm{poly}(t)}(x)$$

$$\lesssim \left(\mathrm{pnK}^{t/2}(y \mid x) + \mathrm{pnK}^{t/2}(x)\right) - \mathrm{pK}^{\mathrm{poly}(t)}(x)$$

$$\lesssim \mathrm{pnK}^{t/2}(y \mid x) + \mathrm{pK}^{t/2}(x) - \mathrm{pK}^{\mathrm{poly}(t)}(x)$$

First print x in time $t/2$, then print y given x in time $t/2$.

Almost what we want!

The time bounded computation depth of x

How to eliminate the computational depth?

- We already have:

(The computational depth)

$$\mathrm{pK}^{\mathrm{poly}(t)}(y \mid x) \lesssim \mathrm{pnK}^{t}(y \mid x) + \left(\mathrm{pK}^{t}(x) - \mathrm{pK}^{\mathrm{poly}(t)}(x)\right).$$

By counting argument, for $1 - n^{-c}$ fraction of x, $\mathrm{pK}^{t}(x) - \mathrm{pK}^{\mathrm{poly}(t)}(x) \leq c \log n$. Thus for most x, the witness sampling is time efficient. This implies the average-case easiness of NP:

$$\mathrm{MINnKT} \in \mathrm{BPP} \Longrightarrow \mathrm{DistNP} \subseteq \mathrm{AvgBPP}.$$

But we want the worst-case easiness of NP! How can we eliminate the computational depth from the inequality?

Eliminating computational depth by telescoping

Notice how this is smaller than $\mathrm{poly}(t)$!

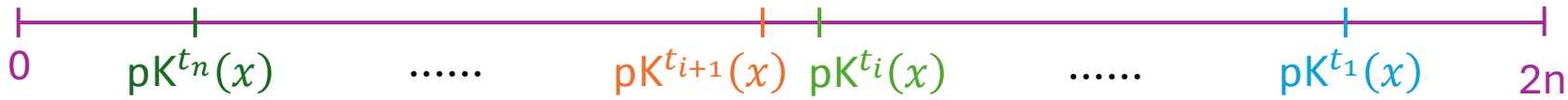
- If, magically 🪄, for every $t > t_0$, we have

$$\mathrm{pK}^{\mathrm{poly}(t)}(y \mid x) \lesssim \mathrm{pnK}^{t_0}(y \mid x) + \left(\mathrm{pK}^{t}(x) - \mathrm{pK}^{t+\mathrm{poly}(t_0)}(x)\right).$$

Then defining $t_1 = \mathrm{poly}(t_0), t_2 = t_1 + \mathrm{poly}(t_0), \ldots, t_{n+1} = t_n + \mathrm{poly}(t_0)$, we have:

$$\mathrm{pK}^{\mathrm{poly}(t_1)}(y \mid x) \lesssim \mathrm{pnK}^{t_0}(y \mid x) + \left(\mathrm{pK}^{t_1}(x) - \mathrm{pK}^{t_2}(x)\right),$$
$$\mathrm{pK}^{\mathrm{poly}(t_2)}(y \mid x) \lesssim \mathrm{pnK}^{t_0}(y \mid x) + \left(\mathrm{pK}^{t_2}(x) - \mathrm{pK}^{t_3}(x)\right),$$
$$,\ldots,$$
$$\mathrm{pK}^{\mathrm{poly}(t_n)}(y \mid x) \lesssim \mathrm{pnK}^{t_0}(y \mid x) + \left(\mathrm{pK}^{t_n}(x) - \mathrm{pK}^{t_{n+1}}(x)\right).$$

Since $\sum_{i=1}^{n}\left(\mathrm{pK}^{t_i}(x) - \mathrm{pK}^{t_{i+1}}(x)\right) = \mathrm{pK}^{t_1}(x) - \mathrm{pK}^{t_{n+1}}(x) \le \mathrm{pK}^{t_1}(x) \le 2n$, by averaging, there must exist i such that $\mathrm{pK}^{t_i}(x) - \mathrm{pK}^{t_{i+1}}(x) \le 2$.



Eliminating computational depth by telescoping

Notice how this is smaller than $\mathrm{poly}(t)$!

- If, magically 🪄, for every $t > t_0$, we have

$$\mathrm{pK}^{\mathrm{poly}(t)}(y \mid x) \lesssim \mathrm{pnK}^{t_0}(y \mid x) + \left(\mathrm{pK}^{t}(x) - \mathrm{pK}^{t+\mathrm{poly}(t_0)}(x)\right).$$

Then defining $t_1 = \mathrm{poly}(t_0), t_2 = t_1 + \mathrm{poly}(t_0), \ldots, t_{n+1} = t_n + \mathrm{poly}(t_0)$, there exists i such that $\mathrm{pK}^{t_i}(x) - \mathrm{pK}^{t_{i+1}}(x) \leq 2$. Thus we have

$$\mathrm{pK}^{\mathrm{poly}(t_i)}(y \mid x) \lesssim \mathrm{pnK}^{t_0}(y \mid x) + \mathrm{pK}^{t_i}(x) - \mathrm{pK}^{t_{i+1}}(x) \lesssim \mathrm{pnK}^{t_0}(y \mid x).$$

But t_i is also upper bounded by $\mathrm{poly}(t_0)$. Hence we get

$$\mathrm{pK}^{\mathrm{poly}(t_0)}(y \mid x) \lesssim \mathrm{pnK}^{t_0}(y \mid x).$$

Which is exactly what we need!

The “magic”?

- Bad news: we don’t know how to prove this inequality exactly!

$$\mathsf{pK}^{\mathrm{poly}(t)}(y \mid x) \lesssim \mathsf{pnK}^{t_0}(y \mid x) + \left(\mathsf{pK}^{t}(x) - \mathsf{pK}^{t+\mathrm{poly}(t_0)}(x)\right).$$

- Good news: we can prove a slightly different one, with the same consequences!

$$\mathsf{pK}^{\mathrm{poly}(t)}(y \mid x) \lesssim \mathsf{pnK}^{t_0}(y \mid x) + \left(\ell\text{-}\mathsf{nK}^{t}(x) - \ell\text{-}\mathsf{nK}^{t+\mathrm{poly}(t_0)}(x)\right).$$

$$\ell\text{-}\mathsf{nK}^{t}(x) := \min\left\{ s \;\middle|\; \Pr_{r\sim\{0,1\}^{\ell}}\left[\mathsf{nK}^{t}(x \circ r) \le s + \ell\right] \ge \frac{2}{3} \right\}.$$

Can be eliminated by telescoping!

Technical caveat: to obtain the small gap between time bounds, the underlying computation model need to allow efficient composition of programs. Roughly speaking, if f can be computed in time t_1 and g can be computed in time t_2, then $g \circ f$ can be computed in time $t_1 + \mathrm{poly}(t_2)$.

The technical notion ℓ-nK

$$\ell\text{-nK}^t(x) := \min\left\{ s \;\middle|\; \Pr_{r\sim\{0,1\}^\ell}\left[\text{nK}^t(x\circ r) \le s+\ell\right] \ge \frac{2}{3} \right\}.$$

- Comparing to pnK:

$$\text{pnK}^t(x) := \min\left\{ s \;\middle|\; \Pr_{r}\left[\text{nK}^t(x \mid r) \le s\right] \ge \frac{2}{3} \right\}.$$

- Advantage of ℓ-nK: efficient weak symmetry of information!

$$\Pr_{r\sim\{0,1\}^m}\left[\ell\text{-nK}^t(x\circ r) \gtrsim (\ell+m)\text{-nK}^t(x)+m\right] > \frac{1}{\text{poly}(n)}.$$

- For pnK, the time bound blows up by poly(t)!

$$\Pr_{r\sim\{0,1\}^m}\left[\text{pnK}^t(x\circ r) \gtrsim \text{pnK}^{\text{poly}(t)}(x)+m\right] > \frac{1}{\text{poly}(n)}.$$

Summarizing the proof of the main result

1. Sampling the compressed witnesses:

   If NP admits pK^t witness compression, then NP $\subseteq$ BPP.

2. Unconditional optimal language compression for pnK:

   NP admits pnK^t witness compression unconditionally.

3. Collapsing conditional pnK to pK, and eliminating the computational depth by telescoping:

   Assuming MINnKT $\in$ BPP, then $\mathrm{pK}^{\mathrm{poly}(t)}(y \mid x) \lesssim \mathrm{pnK}^t(y \mid x)$.

Summary of our results

- We prove that MINnKT is hard for NP.

$$\text{MINnKT} \in \text{BPP} \iff \text{NP} \in \text{BPP}.$$

- We define new notions pnK^t, and prove it has optimal language compression and conditional coding, and ℓ-nK^t, which admits efficient weak symmetry of information.
- We prove the average-case hardness of MINnKT.

$$(\text{MINnKT}, \mathcal{U}) \in \text{AvgBPP} \implies \text{DistNP} \subseteq \text{AvgBPP}$$

- We prove that if MINnKT is average-case easy, then NP can be solved in subexponential time in the worst-case.

$$(\text{MINnKT}, \mathcal{U}) \in \text{AvgBPP} \implies \text{NP} \subseteq \text{RTIME}\left[2^{O(n/\log n)}\right]$$

  - This improves the previous result of [Hir'21, GKLO'22], where the condition on the left is Dist$\Sigma_2^{\text{P}} \subseteq$ AvgBPP.

Open problems

- Worst-case hardness of GapMINnKT?
  - Accept if $\mathrm{nK}^{t}(x) \leq s$, reject if $\mathrm{nK}^{\mathrm{poly}(t)}(x) \geq s + O(\log n)$.
  - Cannot use telescoping anymore, because of the time gap is t v.s. $\mathrm{poly}(t)$.
  - Worst-case hardness of GapMINpnKT would rule out PH-heuristica.
    (Where PH is average-case easy, but worst-case hard.)
- NP-hardness of MINKT/MCSP?