Electronic Colloquium on Computational Complexity, Report No. 7 (2022)

A Simpler Proof of the Worst-Case to Average-Case Reduction for Polynomial Hierarchy via Symmetry of Information

Halley Goldberg* Valentine Kabanets†

January 13, 2022

Abstract

We give a simplified proof of Hirahara's result showing that $\mathsf{DistPH} \subseteq \mathsf{AvgP}$ would imply $\mathsf{PH} \subseteq \mathsf{DTIME}[2^{O(n/\log n)}]$ [Hir21a]. The argument relies on a proof of the new result: *Symmetry of Information* for time-bounded Kolmogorov complexity under the assumption that NP is easy on average, which is interesting in its own right and generalizes the "weak" Symmetry of Information theorem from the original [Hir21a].

1 Introduction

What kind of worst-case consequences do we get if we assume that NP is easy on average? It was shown by [Ben+92] that this would imply that $\mathsf{NE} = \mathsf{E}$. Later, [BFP03] strengthened this result by showing that for every NE language with a polynomial-time verifier $V(x, y)$ (where y, $|y| \leq 2^{O(|x|)}$, is a candidate witness for x being a member of L) has a deterministic algorithm that, given $x \in \{0,1\}^n$, will find some witness y such that $V(x, y)$ accepts, in time $2^{O(n)}$, if $x \in L$.

These results imply that $\mathsf{NP} \subseteq \mathsf{E}$ under the assumption that $\mathsf{DistNP} \subseteq \mathsf{AvgP}$. But is it possible to strengthen the upper bound for NP, ideally showing that $\mathsf{NP} = \mathsf{P}$ under some average-case easiness assumption? Recently, Hirahara [Hir21a] showed that if Σ_2^{P} is easy on average, then $\mathsf{NP} \subseteq \mathsf{TIME}[2^{O(n/\log n)}]$. Note that the conclusion is stronger than that of [Ben+92; BFP03], but also the assumption is stronger: we assume $\mathsf{Dist}\Sigma_2^{\mathsf{P}} \subseteq \mathsf{AvgP}$ rather than $\mathsf{DistNP} \subseteq \mathsf{AvgP}$.

The proof of [Hir21a] uses a combination of ideas from pseudorandomness and time-bounded Kolmogorov complexity. Essentially, it shows that, if $\mathsf{Dist}\Sigma_2^{\mathsf{P}} \subseteq \mathsf{AvgP}$, then for every language $L \in \mathsf{NP}$, with a polynomial-time verifier $V(x, y)$ (for candidate witnesses y, $|y| \leq \mathsf{poly}(|x|)$), the lexicographically first witness y_x for $x \in L$ is compressible to $O(n/\log n)$ bits, so that it can be reconstructed (decompressed) from its compressed image of size $O(n/\log n)$, and a given $x \in \{0,1\}^n$, in deterministic time $2^{O(n/\log n)}$ (i.e., the conditional time-bounded Kolmogorov complexity of y_x, given x, is small: $\mathsf{K}^{2^{O(n/\log n)}}(y_x \mid x) \leq O(n/\log n)$). Then, by trying all possible compressed images, and checking if any one of them decompresses to a valid witness y such that $V(x, y)$ accepts, one gets a deterministic algorithm for L, with the running time $2^{O(n/\log n)}$.[1]

*Simon Fraser University; halley_goldberg@sfu.ca

†Simon Fraser University; kabanets@sfu.ca

[1] Since finding the lexicographically first witness y_x for a given input x so that a polytime verifier $V(x, y_x)$ accepts is P^{NP}-hard [Kre88], it follows that $\mathsf{P}^{\mathsf{NP}} \subseteq \mathsf{TIME}[2^{O(n/\log n)}]$ under the same assumption that $\mathsf{Dist}\Sigma_2^{\mathsf{P}} \subseteq \mathsf{AvgP}$.

ISSN 1433-8092

In this paper, we give a simple self-contained proof of this result by Hirahara and some generalizations. Our proof argument follows in three steps.

Step 1: The main new ingredient behind our simplified proof is the Symmetry of Information result (see Lemma 10 below), proved under the assumption that $\mathsf{DistNP} \subseteq \mathsf{AvgP}$. Roughly, it says that there exists a constant $c \geq 1$, such that, for any binary strings x and y, and for any sufficiently large parameter t,

$$\mathsf{K}^{t^c}(y \mid x) \leq \mathsf{K}^t(x, y) - \mathsf{K}^{t^c}(x) + O(\log t).$$

Step 2: The next step is to show, under the assumption that $\mathsf{Dist}\Sigma_2^{\mathsf{P}} \subseteq \mathsf{AvgP}$, that for every language $L \in \mathsf{NP}$, there is a constant $d \geq 1$ such that

$$\mathsf{K}^t(x, y_x) \leq \mathsf{K}^{t^{1/d}}(x) + O(\log t),$$

where y_x is the lexicographically first L-witness of a given $x \in L$, and t is any sufficiently large parameter; see Lemma 11.

Step 3: Combining Steps 1 and 2, we get for any given language $L \in \mathsf{NP}$ that

$$\mathsf{K}^{t^c}(y_x \mid x) \leq \mathsf{K}^{t^{1/d}}(x) - \mathsf{K}^{t^c}(x) + O(\log t),$$

for any sufficiently large time bound t. The difference between two polynomially related time-bounded Kolmogorov complexity measures of a given string $x \in \{0,1\}^n$ (the *computational depth* of x) can be as large as n in the worst case. However, via a simple averaging argument (as in [Hir21a]), one can show that, for every $x \in \{0,1\}^n$ and constants $c, d \geq 1$, there must exist a time bound $t \leq 2^{O(n/\log n)}$ such that

$$\mathsf{K}^{t^{1/d}}(x) - \mathsf{K}^{t^c}(x) \leq O(n/\log n);$$

see Lemma 12. This implies the promised upper bound on the Kolmogorov complexity of y_x:

$$\mathsf{K}^{2^{O(n/\log n)}}(y_x \mid x) \leq O(n/\log n).$$

Recently, in independent work, Hirahara [Hir21b] also proved the Symmetry of Information result (Lemma 10), and used it to give a simplified proof of his original result from [Hir21a] along the similar lines as our proof sketch above.

2 Preliminaries

Definition 1. (Time-bounded Kolmogorov-complexity) Fix an efficient universal Turing Machine U. For strings $x, y \in \{0,1\}^*$, an oracle $A \subseteq \{0,1\}^*$, and a time bound $t \in \mathbb{N} \cup \{\infty\}$ such that $t \geq \max\{|x|, |y|\}$, the *A-oracle t-time-bounded Kolmogorov-complexity of x given y* is defined as

$$\mathsf{K}^{t,A}(x \mid y) = \min\{s \in \mathbb{N} \mid \text{for some } d \in \{0,1\}^s,\ U^A(d, y) \text{ outputs } x \text{ in time } t\}.$$

We omit "$\mid y$" if y is the empty string, the superscript "A" if $A = \varnothing$, and the superscript "t" if $t = \infty$.

Definition 2. (Time-Bounded Computational Depth) For a string $x \in \{0,1\}^*$ and a time bound $t \in \mathbb{N} \cup \{\infty\}$ such that $t \geq |x|$, the *(s,t)-time-bounded computational depth of x* is defined as

$$\mathsf{cd}^{s,t}(x) = \mathsf{K}^s(x) - \mathsf{K}^t(x).$$

Definition 3. A distributional problem (L, D) is in *average-case polynomial-time*, denoted AvgP, if there exists an algorithm A that decides L correctly on all inputs in the support of D and a constant $\epsilon > 0$ such that, for all $n \in \mathbb{N}$, $\mathbb{E}_{x \sim D_n}[t_A(x)^\epsilon] \leq n^{O(1)}$, where $t_A(x)$ is an upper bound on the running time of A on input x.

Definition 4. A distributional problem (L, D) admits a *one-sided-error heuristic algorithm* with failure probability $\delta : \mathbb{N} \to (0,1)$ if there exists an algorithm A such that for all $n \in \mathbb{N}$, for all $x \in \mathrm{supp}(D_n)$ with $L(x) = 1$, $A(x, 1^n) = 1$, and moreover, $\Pr_{x \sim D_n}[A(x, 1^n) \neq L(x)] < \delta(n)$. Denote the class of problems admitting polynomial-time one-sided-error heuristic algorithms $\mathsf{Avg}^1_\delta\mathsf{P}$.

Lemma 5. *For a complexity class $\mathfrak{C}$, if $\mathsf{Dist}\mathfrak{C} \subseteq \mathsf{AvgP}$, then for any constant $c \in \mathbb{N}$, $\mathsf{Dist}\mathfrak{C} \subseteq \mathsf{Avg}^1_{n^{-c}}\mathsf{P}$.*

Proof. Assume $\mathsf{Dist}\mathfrak{C} \subseteq \mathsf{AvgP}$, and let $c \in \mathbb{N}$ be given. Let $L \in \mathfrak{C}$, and let $D = \{D_n\}_{n \in \mathbb{N}}$ be a polytime samplable distribution, so $(L, D) \in \mathsf{Dist}\mathfrak{C}$. Let A be the algorithm from the definition of AvgP; that is, $A(x, 1^n) = L(x)$ for all $x \in \mathrm{supp}(D)$, and there exist constants ϵ and d such that for all $n \in \mathbb{N}$, $\mathbb{E}_{x \sim D_n}[t_A(x)^\epsilon] \leq n^d$.

Let A' be the following polytime algorithm. On input $(x, 1^n)$, run $A(x)$ for n^b steps, where $b := (c+d)/\epsilon$. If A halts and returns an output within that period, return that output. Otherwise, output 1.

By correctness of A, whenever $L(x) = 1$, it holds that $A'(x, 1^n) = 1$. We claim that the probability of A' erring on no-instances of L is also small. In particular, by Markov's Inequality,

$$\begin{aligned}
\Pr_{x \sim D_n}[t_A(x) > n^b] &= \Pr_{x \sim D_n}[t_A(x)^\epsilon > n^{b\epsilon}] \\
&\leq \mathbb{E}_{x \sim D_n}[t_A(x)^\epsilon] \cdot n^{-b\epsilon} \\
&\leq n^{d - b\epsilon} \\
&= n^{-c},
\end{aligned}$$

as required. □

Lemma 6 ([BFP03])**.** *If $\mathsf{DistNP} \subseteq \mathsf{AvgP}$, then there is a PRG G such that, for every $\epsilon > 0$ and all sufficiently large $n \in \mathbb{N}$, G: $\{0,1\}^\ell \to \{0,1\}^n$ ϵ-fools circuits of size n, using the seed size $\ell = O(\log(n/\epsilon))$, and an n-bit output of G on a given seed is computable in time $\mathsf{poly}(n/\epsilon)$. In particular, $\mathsf{BPP} = \mathsf{P}$.*

Definition 7. For $n, k \in \mathbb{N}$, the *k-wise direct product generator* $\mathsf{DP}_k : \{0,1\}^n \times \{0,1\}^{nk} \to \{0,1\}^{nk+k}$ is the function defined by

$$\mathsf{DP}_k(x; z^1, ..., z^k) = (z^1, ..., z^k; \langle x, z^1 \rangle, ..., \langle x, z^k \rangle),$$

where $\langle -, - \rangle$ denotes the inner product $\langle x, y \rangle = \left(\sum_{i=1}^{|x|} x_i y_i \right) \mod 2$.

Lemma 8 (Goldreich-Levin Local List-Decoding [GL89; GSR95])**.** *There is a probabilistic algorithm A with the following property. For an arbitrary $x \in \{0,1\}^n$, let $f : \{0,1\}^n \to \{0,1\}$ be such that $\Pr_y[f(y) = \langle x, y\rangle] \geq 1/2 + \epsilon$ for some $\epsilon > 0$. Then, given oracle access to f and inputs $1/\epsilon$ and n, A outputs x in time $\mathsf{poly}(n/\epsilon)$, with probability at least $\Omega(\epsilon^2)$.*[2]

Lemma 9 (DP_k Reconstruction [Hir21a])**.** *Suppose there exists a pseudorandom generator as described in Lemma 6. Let D be a randomized circuit of size s that ϵ-distinguishes $\mathsf{DP}_k(x;\mathcal{U}_{nk})$ and the uniform distribution $\mathcal{U}_{nk+k}$, where $x \in \{0,1\}^n$ and $k \leq \mathsf{poly}(n)$. Then there is a deterministic algorithm that, given oracle access to D and an advice string of length at most $k + O(\log(ns/\epsilon))$, outputs x in time $p_{\mathsf{DP}}(ns/\epsilon)$, where p_{DP} is a fixed polynomial independent of D. In particular,*

$$\mathsf{K}^{p_{\mathsf{DP}}(ns/\epsilon)}(x \mid D) \leq k + \log p_{\mathsf{DP}}(ns/\epsilon).$$

Proof. Let D be a randomized circuit of size s satisfying

$$\Pr_{z,r}[D(\mathsf{DP}_k(x;z), r) = 1] - \Pr_{w,r}[D(w, r) = 1] \geq \epsilon,$$

where r denotes the internal randomness of D. Let G_s be the pseudorandom generator from Lemma 6 that $(\epsilon/4)$-fools every circuit of size s. It follows that

$$\Pr_{z,\sigma}[D(\mathsf{DP}_k(x;z), G_s(\sigma)) = 1] - \Pr_{w,\sigma}[D(w, G_s(\sigma)) = 1] \geq \epsilon/2.$$

We define a function $D_\sigma : \{0,1\}^{nk+k} \to \{0,1\}$ so that $D_\sigma(w) = D(w, G_s(\sigma))$ and note that there must exist a "good" seed $\sigma \in \{0,1\}^{O(\log(s/\epsilon))}$ such that

$$\Pr_{z}[D_\sigma(\mathsf{DP}_k(x;z)) = 1] - \Pr_{w}[D_\sigma(w) = 1] \geq \epsilon/2.$$

That is,

$$\Pr_{\overline{z}}[D_\sigma(z^1, ..., z^k; \langle x, z^1\rangle, ..., \langle x, z^k\rangle) = 1] - \Pr_{\overline{z},b}[D_\sigma(z^1, ..., z^k; b_1, ..., b_k) = 1] \geq \epsilon/2,$$

where $\overline{z} = (z^1, ..., z^k) \sim (\{0,1\}^n)^k$ and $b = (b_1, ..., b_k) \sim \{0,1\}^k$.

Let P^{D_σ} be the following next-bit-predictor algorithm: on input

$$u = (\overline{z}; \langle x, z^1\rangle, ..., \langle x, z^{i-1}\rangle, b_i, b_{i+1}, ..., b_k),$$

output b_i if $D_\sigma(u) = 1$ and $1 - b_i$ otherwise. By a standard hybrid argument (see for example [Vad12] Proposition 7.16),

$$\Pr_{z^i}[P^{D_\sigma}(\overline{z}; \langle x, z^1\rangle, ..., \langle x, z^{i-1}\rangle, b_i, b_{i+1}, ..., b_k) = \langle x, z^i\rangle] \geq \frac{1}{2} + \frac{\epsilon}{4k}, \tag{1}$$

with probability at least $\epsilon/4k$ over the random choice of $i \in [k], z^{[k]\setminus\{i\}}$ and b.

Now we define a randomized reconstruction procedure R^{D_σ}, which takes randomness r' of length $\mathsf{poly}(n/\epsilon)$ along with an advice string $\alpha \in \{0,1\}^k$. This procedure interprets r' as containing

[2] The GL algorithm runs in time $O(n^2\epsilon^{-4}\log n)$, makes $O(n\epsilon^{-4}\log n)$ oracle queries to f, and outputs a list L of $O(\epsilon^{-2})$ candidate n-bit strings such that, with probability at least $1/2$, the list L contains x. Outputting a uniformly random element of L yields x with probability at least $1/(2|L|)$.

random choices of $i, \overline{z}, b$, and some $r_0 \sim \{0,1\}^{\mathsf{poly}(n/\epsilon)}$. Given r' and α, it defines a function $f : \{0,1\}^n \to \{0,1\}$ as

$$f(y) = P^{D_\sigma}(z^1, ..., z^{i-1}, y, z^{i+1}, ..., z^k; \alpha)$$

for $y \in \{0,1\}^n$, and then it runs the list-decoding algorithm of Lemma 8 on f using randomness r_0.

Define an advice function $A : \{0,1\}^n \times \{0,1\}^{\mathsf{poly}(n/\epsilon)} \to \{0,1\}^k$ as

$$A(x, r') = (\langle x, z^1\rangle, ..., \langle x, z^{i-1}\rangle, b_i, ..., b_k),$$

where $i, \overline{z}$, and b are as specified in r'. Note that A is computable by a circuit of size $\mathsf{poly}(n/\epsilon)$. By equation (1), $\Pr_y[f(y) = \langle x, y\rangle] \geq \frac{1}{2} + \frac{\epsilon}{4k}$ with probability at least $\epsilon/4k$ over $i, \overline{z}$, and b. When this occurs, the list-decoding algorithm will output x with probability at least $\Omega(k^2/\epsilon^2) \geq 1/\mathsf{poly}(n/\epsilon)$ over r_0. Therefore, given the correct advice $\alpha = A(x, r')$,

$$\Pr_{r'}[R^{D_\sigma}(\alpha, r') = x] \geq \frac{\epsilon}{4k} \cdot \frac{1}{\mathsf{poly}(n/\epsilon)} = \frac{1}{\mathsf{poly}(n/\epsilon)} =: \epsilon'.$$

Observe that the condition $R^{D_\sigma}(A(x, r'), r') = x$ can be checked by a circuit of size $s' \in \mathsf{poly}(ns/\epsilon)$ given r' as input. Thus, applying the pseudorandom generator $G_{s'}$ of Lemma 6 that $(\epsilon'/2)$-fools all circuits of size s', we have that

$$\Pr_{\sigma'}[R^{D_\sigma}(A(x, G_{s'}(\sigma')), G_{s'}(\sigma')) = x] \geq \epsilon'/2,$$

and hence there is a good seed $\sigma' \in \{0,1\}^{O(\log(s'/\epsilon'))}$ such that $R^{D_\sigma}(A(x, G_{s'}(\sigma')), G_{s'}(\sigma')) = x$.

We are now in a position to describe a deterministic algorithm M that outputs x. Given an advice string $\alpha = A(x, G_{s'}(\sigma')) \in \{0,1\}^k$ along with the seeds $\sigma \in \{0,1\}^{O(\log(s/\epsilon))}$ and $\sigma' \in \{0,1\}^{O(\log(s'/\epsilon'))}$ defined above, M computes and outputs $R^{D_\sigma}(\alpha, G_{s'}(\sigma'))$ in time $\mathsf{poly}(ns/\epsilon)$. Observe that the total advice required is of length $k + O(\log(ns/\epsilon))$. □

3 Main proof ingredients

3.1 Symmetry of information

The following lemma generalizes the "weak" symmetry of information result of [Hir21a, Theorem 5.2]. The same lemma was also independently proved by Hirahara [Hir21b]. For some previous work on time-bounded symmetry of information, see [LW95; LR05].

Lemma 10 (Symmetry of Information)**.** *If* $\mathsf{DistNP} \subseteq \mathsf{AvgP}$*, then there exist polynomials* p *and* p_0 *such that for all sufficiently large* $x, y \in \{0,1\}^*$ *and every* $t \geq p_0(|x| + |y|)$*,*

$$\mathsf{K}^t(x, y) > \mathsf{K}^{p(t)}(x) + \mathsf{K}^{p(t)}(y \mid x) - \log p(t).$$

Proof. Let $x \in \{0,1\}^n$, $y \in \{0,1\}^m$, $k, k' \in \mathbb{N}$ to be defined later. Observe that there exists a polynomial p_0 and a constant $d \in \mathbb{N}$ such that for any $t \geq p_0(n+m)$ and any choice of z and z',

$$\mathsf{K}^{2t}(\mathsf{DP}_k(x; z), \mathsf{DP}_{k'}(y; z')) \leq \mathsf{K}^t(x, y) + |z| + |z'| + d \log t. \tag{2}$$

In particular, $p_0(n+m)$ reflects the time required to compute $(\mathsf{DP}_k(x; z), \mathsf{DP}_{k'}(y; z'))$ given xy, z, z', and $d \log t$ bits of information to delineate x from y. In what follows, we will give a lower bound on $\mathsf{K}^{2t}(\mathsf{DP}_k(x; z), \mathsf{DP}_{k'}(y; z'))$ and thereby a lower bound on $\mathsf{K}^t(x, y)$.

Define a language

$$L := \{(u, w, 1^t, 1^s) \mid \mathsf{K}^{2t}(u, w) \leq s\},$$

and note that $L \in \mathsf{NP}$. Define a distribution family $D = \{D_{\langle n,m.k,k',t,s\rangle}\}$, each member of which does the following: sample $u \sim \{0,1\}^{nk+k}$ and $w \sim \{0,1\}^{mk'+k'}$, and then output $(u, w, 1^t, 1^s)$. By assumption, $(L, D) \in \mathsf{AvgP}$. Let B be a one-sided error heuristic algorithm for (L, D) with failure probability at most $1/t$.

Let $t \geq p_0(n, m)$, and define $s := |z| + k + |z'| + k' - \log t - 1$. By a counting argument, for randomly selected u and w,

$$\Pr_{u,w}[(u, w, 1^t, 1^s) \in L] \leq \frac{2^{s+1}}{2^{|z|+k+|z'|+k'}} = \frac{1}{t}.$$

Then by definition of B,

$$\Pr_{u,w}[B(u, w, 1^t, 1^s) = 1] \leq \frac{2}{t}. \tag{3}$$

Next we show, using a hybrid argument, that $B(-, 1^t, 1^s)$ *cannot* distinguish between the uniform distribution and the distribution $(\mathsf{DP}_k(x; z), \mathsf{DP}_{k'}(y; z'))$, for random independent z, z', where $k \approx \mathsf{K}^{p_{\mathsf{DP}}(t)}(x)$ and $k' \approx \mathsf{K}^{p_{\mathsf{DP}}(t)}(y \mid x)$. This will imply that $B(\mathsf{DP}_k(x; z), \mathsf{DP}_{k'}(y; z')) = 0$ for some z, z', yielding the desired lower bound on $\mathsf{K}^{2t}(\mathsf{DP}_k(x; z), \mathsf{DP}_{k'}(y; z'))$. We give the details of the hybrid argument next.

Toward a contradiction, suppose $\Pr_{z,w}[B(\mathsf{DP}_k(x; z), w, 1^t, 1^s) = 1] > 1/2$. In this case, comparing with Eq. (3), we get a randomized distinguisher for $\mathsf{DP}_k(x; \mathcal{U}_{nk})$ defined by sampling $w \sim \{0,1\}^{mk'+k'}$ and outputting $B(-, w, 1^t, 1^s)$. By Lemma 9,

$$\mathsf{K}^{p'(t)}(x) \leq k + \log p'(t) \tag{4}$$

for some polynomial p' such that $p'(t) \geq p_{\mathsf{DP}}(3 \cdot n \cdot s_B)$, where p_{DP} is the polynomial from Lemma 9 and s_B denotes the size of a circuit computing $B(-, w, 1^t, 1^s)$.

We now choose $k := \mathsf{K}^{p'(t)}(x) - \log p'(t) - 1$ so that Eq. (4) *does not hold*. Assume for now that $k > 0$. Hence,

$$\Pr_{z,w}[B(\mathsf{DP}_k(x; z), w, 1^t, 1^s) = 1] \leq 1/2. \tag{5}$$

Again, toward a contradiction, suppose $\Pr_{z,z'}[(\mathsf{DP}_k(x; z), \mathsf{DP}_{k'}(y; z'), 1^t, 1^s) \in L] = 1$, which implies that $\Pr_{z,z'}[B(\mathsf{DP}_k(x; z), \mathsf{DP}_{k'}(y; z')), 1^t, 1^s) = 1] = 1$, since B never errs on yes-instances of L. In this case, comparing with equation (5), we get a randomized distinguisher B' for $\mathsf{DP}_{k'}(y; \mathcal{U}_{mk'})$ defined by sampling $z \sim \{0,1\}^{nk}$ and outputting $B(\mathsf{DP}_k(x; z), -, 1^t, 1^s)$. By Lemma 9,

$$\mathsf{K}^{p''(t)}(y \mid x) \leq k' + \log p''(t) \tag{6}$$

for some polynomial p'' with $p''(t) \geq p_{\mathsf{DP}}(2 \cdot m \cdot s_{B'})$, where $s_{B'}$ denotes the size of a circuit computing B'.

We now choose $k' := \mathsf{K}^{p''(t)}(y \mid x) - \log p''(t) - 1$ so that Eq. (6) does not hold. Assume for now that $k' > 0$. Hence, there exist z and z' such that

$$\mathsf{K}^{2t}(\mathsf{DP}_k(x; z), \mathsf{DP}_{k'}(y; z')) > s.$$

For these z, z', by definition of s, $\mathsf{K}^{2t}(\mathsf{DP}_k(x;z), \mathsf{DP}_{k'}(y;z')) > |z|+k+|z'|+k'-\log t-1$. Combining this inequality with Eq. (2), we get

$$\begin{aligned}\mathsf{K}^t(x,y) &\geq \mathsf{K}^{2t}(\mathsf{DP}_k(x;z), \mathsf{DP}_{k'}(y;z')) - |z| - |z'| - d\log t \\ &> k + k' - d\log t - \log t - 1 \\ &= \mathsf{K}^{p'(t)}(x) + \mathsf{K}^{p''(t)}(y \mid x) - \log p'(t) - \log p''(t) - d\log t - \log t - 3.\end{aligned}$$

As desired, for the polynomial $p(t) := 8 \cdot (t^{d+1}) \cdot p'(t) \cdot p''(t)$,

$$\mathsf{K}^t(x,y) > \mathsf{K}^{p(t)}(x) + \mathsf{K}^{p(t)}(y \mid x) - \log p(t).$$

Finally, consider the case that $k \leq 0$ or $k' \leq 0$. If $k \leq 0$, then $\mathsf{K}^{p'(t)}(x) \leq \log p'(t) + 1$, implying that $\mathsf{K}^{p(t)}(x) < \log p(t)$. But then the lemma simply follows from the fact that $\mathsf{K}^t(x,y) \geq \mathsf{K}^{p(t)}(y \mid x)$. Similarly, if $k' \leq 0$, then $\mathsf{K}^{p(t)}(y \mid x) < \log p(t)$, and the lemma follows from the fact that $\mathsf{K}^t(x,y) \geq \mathsf{K}^{p(t)}(x)$. □

3.2 Oracle elimination

Here we show how to bound the Kolmogorov complexity of the lexicographically first witness y_x for an input x, for any given efficient verifier $V(x,y)$, using the search-to-decision reduction. For a language $L \in \mathsf{NP}$ and $x \in L$, let y_x be the lexicographically first witness for x being in L. It is well-known that y_x can be computed from x via an efficient search-to-decision reduction, given an NP oracle. In other words, the time-bounded Kolmogorov complexity of (x, y_x), given an NP oracle, is essentially at most the oracle-free Kolmogorov complexity of x. The lemma below shows that, under the assumption that PH is easy on average, the NP oracle can be *eliminated*: the oracle-free time-bounded Kolmogorov complexity of (x, y_x) (for a slightly bigger time bound) is essentially at most that of x.

The next lemma is stated in a more general form so that it applies to any pair of (x,y) such that the oracle Kolmogorov complexity of (x,y) is bounded (which is needed for our argument in Corollary 18 below to handle the case of languages in PH; see also Lemma 16).

Lemma 11 (Oracle Elimination)**.** *Let $\ell \in \mathbb{N}$ and suppose $\mathsf{DistNP}^{\Sigma_\ell^\mathsf{P}} \subseteq \mathsf{AvgP}$. Let $x, y \in \{0,1\}^*$ be arbitrary strings. Suppose, for some function $t_0 : \mathbb{N} \to \mathbb{N}$, it holds for all $t \geq t_0(|x|+|y|)$ that*

$$\mathsf{K}^{2t,\Sigma_\ell^\mathsf{P}}(x,y) \leq \mathsf{K}^t(x) + O(\log t).$$

Then there exists a polynomial q such that for all $t \geq t_0(|x|+|y|)$,

$$\mathsf{K}^{q(t)}(x,y) \leq \mathsf{K}^t(x) + \log q(t).$$

Proof. Define a language

$$L' := \{(\mathsf{DP}_k(x,y;z), 1^t, 1^s) \mid \mathsf{K}^{2t,\Sigma_\ell^\mathsf{P}}(x,y) \leq s \text{ and } |x|+|y| \leq t\},$$

for some k to be chosen later. Note that $L' \in \mathsf{NP}^{\Sigma_\ell^\mathsf{P}}$. Define a distribution family $D = \{D_{\langle n,m,t,s\rangle}\}$, each member of which samples $w \sim \{0,1\}^{(n+m)k+k}$ and outputs $(w, 1^t, 1^s)$. By assumption, $(L', D) \in \mathsf{AvgP}$. Let B be a one-sided error heuristic algorithm for (L', D) with failure probability at most t^{-1}, as in Lemma 5.

Let $x \in \{0,1\}^n$ and $y \in \{0,1\}^m$ be such that, for some function t_0 and $d \in \mathbb{N}$, for all $t \geq t_0(n+m)$, $\mathsf{K}^{2t,\Sigma_\ell^\mathsf{P}}(x,y) \leq \mathsf{K}^t(x) + d\log t$. Defining $t \geq t_0(n+m)$ and $s := \mathsf{K}^t(x) + d\log t$, we have

$$\Pr_z[B(\mathsf{DP}_k(x,y;z), 1^t, 1^s) = 1] = 1, \tag{7}$$

since B never errs on yes-instances of L'.

On the other hand, for w selected uniformly at random,

$$\begin{aligned}\Pr_w[(w,1^t,1^s) \in L'] &= \Pr_w[\exists u \in \{0,1\}^{n+m}, \exists z \in \{0,1\}^{(n+m)k}, w = \mathsf{DP}_k(u;z) \wedge \mathsf{K}^{2t,\Sigma_\ell^\mathsf{P}}(u) \leq s] \\ &\leq \frac{2^{s+1} \cdot 2^{|z|}}{2^{|z|+k}} \\ &= t^{-1},\end{aligned}$$

where the second line follows from a union bound and a counting argument, and the third line by defining $k := s + 1 + \log t$. Then

$$\begin{aligned}\Pr_w[B(w,1^t,1^s) = 1] &\leq \Pr_w[(w,1^t,1^s) \in L'] + \Pr_w[B(w,1^t,1^s) \neq L'(w,1^t,1^s)] \\ &\leq 2 \cdot t^{-1} \in o(1).\end{aligned}$$

Comparing with Eq. (7), it is clear that $B(-,1^t,1^s)$ distinguishes $\mathsf{DP}_k(x,y;\mathcal{U}_{|z|})$ from uniform. Lemma 9 implies that

$$\begin{aligned}\mathsf{K}^{p'(t)}(x,y) &\leq k + \log p'(t) \\ &= s + 1 + \log t + \log p'(t) \\ &= \mathsf{K}^t(x) + d\log t + 1 + \log t + \log p'(t),\end{aligned}$$

for some polynomial p' with $p'(t) \geq p_{\mathsf{DP}}(2 \cdot (n+m) \cdot s_B)$, where p_{DP} is the polynomial from Lemma 9 and s_B denotes the size of a circuit computing $B(-,1^t,1^s)$. It follows that that $\mathsf{K}^{q(t)}(x,y) \leq \mathsf{K}^t(x) + \log q(t)$ for the polynomial $q(t) = 2 \cdot t^{d+1} \cdot p'(t)$. □

3.3 Computational depth

Here we show how to bound the computational depth of any input string x.

Lemma 12 (Computational Depth)**.** *There is a constant $d \geq 0$ such that the following holds. For an arbitrary string $x \in \{0,1\}^n$, constants $c \geq 1$ and $0 < \epsilon \leq 1$, and functions $m, I : \mathbb{N} \to \mathbb{N}$, there exists a time bound t such that*

- $m(n)^{c^{I(n)\cdot(1-\epsilon)}} \leq t \leq m(n)^{c^{I(n)}}$, *and*
- $\mathsf{cd}^{t,t^c}(x) \leq \frac{n+d}{\epsilon \cdot I(n)}$.

Proof. Let $m := m(n)$ and $I := I(n)$. For $p(n) := n^c$, consider the following telescoping sum:

$$\begin{aligned}\mathsf{K}^{p(m)}(x) - \mathsf{K}^{p^{I+1}(m)}(x) = &\left(\mathsf{K}^{p(m)}(x) - \mathsf{K}^{p^2(m)}(x)\right) + \left(\mathsf{K}^{p^2(m)}(x) - \mathsf{K}^{p^3(m)}(x)\right) + \\ &\ldots + \left(\mathsf{K}^{p^I(m)}(x) - \mathsf{K}^{p^{I+1}(m)}(x)\right),\end{aligned}$$

where $p^i(m)$ denotes the composition of p with itself i times. For any choice of $x \in \{0,1\}^n$, p, and m, $\mathsf{K}^{p(m)}(x) \leq n + d$, for some universal constant $d \geq 0$ (dependent on the choice of the universal TM U in the definition of Kolmogorov complexity); hence, the above sum is at most $n' := n+d$. By averaging, for a uniformly randomly chosen index $1 \leq i \leq I$, the expected value of the ith bracketed difference is at most n'/I. By Markov's inequality, the fraction of i's where the ith difference is greater than $(1/\epsilon) \cdot n'/I$ is at most ϵ. We get that there is an index i_0 such that $(1-\epsilon) \cdot I \leq i_0 \leq I$ and

$$\mathsf{K}^{p^{i_0}(m)}(x) - \mathsf{K}^{p^{i_0+1}(m)}(x) \leq (1/\epsilon) \cdot n'/I. \tag{8}$$

For this i_0, define $t := p^{i_0}(m)$. Then by (8), we get $\mathsf{cd}^{t,p(t)}(x) \leq \frac{n+d}{\epsilon \cdot I}$. Since $p^{i_0}(m) = m^{c^{i_0}}$, we also get that

$$m^{c^{I\cdot(1-\epsilon)}} \leq t \leq m^{c^I},$$

as desired. □

Corollary 13. *For an arbitrary string $x \in \{0,1\}^n$ (for large enough $n \geq 1$), a constant $c \geq 1$, and a function $t_0 : \mathbb{N} \to \mathbb{N}$ such that $n \leq t_0(n) \leq 2^{bn^{1-a}}$ for some constants $a, b > 0$, there exists a time bound t such that $t_0(n) \leq t \leq 2^{n/\log n}$ and $\mathsf{cd}^{t,t^c}(x) \leq O(n/\log n)$.*

Proof. Apply Lemma 12 with $m(n) = n$, $I = \lfloor \log_c(n/(\log_2 n)^2) \rfloor$ and $\epsilon = a/3$. □

Corollary 14. *For an arbitrary string $x \in \{0,1\}^n$ (for large enough $n \geq 1$), constants $c \geq 1$ and $\delta > 0$, and a function $t_0 : \mathbb{N} \to \mathbb{N}$ such that $t_0(n) \geq n$, there exists a time bound t such that $t_0(n) \leq t \leq t_0(n)^{O(1)}$ and $\mathsf{cd}^{t,t^c}(x) \leq \delta \cdot n$.*

Proof. Since $d \leq n$ for a sufficiently large n, it suffices to apply Lemma 12 for any parameters ϵ and I such that $(2n)/(\epsilon I) \leq \delta n$. Setting $m(n) = t_0(n)$, $\epsilon = 1$, and $I = \lceil 2/\delta \rceil$ concludes the proof. □

4 Putting everything together

4.1 Case of PH

We will give a simplified proof of a generalization of the following theorem.

Theorem 15 ([Hir21a])**.** *If $\mathsf{DistPH} \subseteq \mathsf{AvgP}$, then $\mathsf{PH} \subseteq \mathsf{DTIME}[2^{O(n/\log n)}]$.*

We shall use the following auxiliary lemma. For intuition, in the statement below, set $\ell = 1$ and think of y as the lexicographically first witness for $x \in L$, for some $L \in \mathsf{NP}$; this y can be efficiently reconstructed from x, via search-to-decision, given an NP oracle, and hence Eq. (9) is satisfied. Then Item 1 of the lemma immediately implies a special case of Theorem 15, concluding that $\mathsf{NP} \subseteq \mathsf{DTIME}[2^{O(n/\log n)}]$. However, in order to extend the argument to PH and to conclude that $\mathsf{PH} \subseteq \mathsf{DTIME}[2^{O(n/\log n)}]$ (as we do in Corollary 18 below), it is important that we can apply this lemma to any y satisfying Eq. (9).

Lemma 16 (implicit in [Hir21a])**.** *Let $\ell \in \mathbb{N}$ and suppose $\mathsf{DistNP}^{\Sigma^{\mathsf{P}}_\ell} \subseteq \mathsf{AvgP}$. Let $x \in \{0,1\}^n$ and $y \in \{0,1\}^m$. Let the function $t_0 : \mathbb{N} \to \mathbb{N}$ be such that $t_0(n+m) \leq 2^{O(n^{1-\epsilon})}$ for some constant $0 < \epsilon < 1$, and suppose that for all $t \geq t_0(n+m)$,*

$$\mathsf{K}^{2t,\Sigma^{\mathsf{P}}_\ell}(x,y) \leq \mathsf{K}^t(x) + O(\log t). \tag{9}$$

Then,

1. *there exists a constant $c > 0$ such that*

$$\mathsf{K}^{2^{cn/\log n}}(y \mid x) \leq cn/\log n;$$

2. *for every constant $\delta > 0$, there exists a constant $c > 0$ such that*

$$\mathsf{K}^{t_0(n+m)^c}(y \mid x) \leq \delta \cdot n.$$

Proof. Item 1. Let $x \in \{0,1\}^n$, $y \in \{0,1\}^m$, and t_0 be as described above. Let p_0, p and q be the polynomials from Lemmas 10 and 11 respectively. By Corollary 13, there exists $t \leq 2^{n/\log n}$ such that $t \geq \max\{p_0(n+m), t_0(n+m)\}$ and

$$\mathsf{K}^t(x) - \mathsf{K}^{p(q(t))}(x) \leq O(n/\log n).$$

For such a t, by Lemma 11,

$$\mathsf{K}^{q(t)}(x,y) \leq \mathsf{K}^t(x) + \log q(t),$$

and by Lemma 10,

$$\mathsf{K}^{q(t)}(x,y) > \mathsf{K}^{p(q(t))}(x) + \mathsf{K}^{p(q(t))}(y \mid x) - \log p(q(t)).$$

Combining the previous three inequalities,

$$\begin{aligned}
\mathsf{K}^{p(q(t))}(y \mid x) &< \mathsf{K}^{q(t)}(x,y) - \mathsf{K}^{p(q(t))}(x) + \log p(q(t)) \\
&\leq \left(\mathsf{K}^t(x) - \mathsf{K}^{p(q(t))}(x)\right) + \log p(q(t)) + \log q(t) \\
&\leq O(n/\log n).
\end{aligned}$$

Item 2. Argue as in Item 1 above, invoking Corollary 14 instead of Corollary 13. □

Lemma 16 above shows that if PH is easy on average, every language in PH is verifiable with arbitrarily small linear-length witnesses. More precisely, we have the following.

Corollary 17. *Let $\ell \in \mathbb{N}$ and suppose $\mathsf{DistNP}^{\Sigma_\ell^\mathsf{P}} \subseteq \mathsf{AvgP}$. Then for every $L \in \Sigma_\ell^\mathsf{P}$ and constant $\epsilon > 0$, there is a polytime relation R such that for every sufficiently large $x \in \{0,1\}^n$,*

$$x \in L \iff \exists z_1\, \forall z_2\, \ldots Q z_\ell \;\; R(x, z_1, \ldots, z_\ell),$$

where $|z_i| \leq \epsilon n$ for all $1 \leq i \leq \ell$.[3]

Proof Sketch. Suppose there is some polytime relation R_0 such that

$$x \in L \iff \exists y_1\, \forall y_2\, \ldots Q y_l \;\; R_0(x, y_1, \ldots, y_l),$$

where the lengths of all witnesses y_i are polynomial in $n = |x|$. Using a search-to-decision procedure with a a Σ_ℓ^P-oracle, y_1 may be obtained from x in time n^b for some $b > 0$, which implies that for all $t \geq n^b$,

$$\mathsf{K}^{2t,\Sigma_\ell^\mathsf{P}}(x, y_1) \leq \mathsf{K}^t(x) + O(\log t).$$

[3] Chen, Hirahara, and Vafa [CHV21] prove a version of this statement for $\ell = 1$. In their terminology, our remark states that for all $\ell \geq 1$ and $\epsilon > 0$, if $\mathsf{Dist}\Sigma_{\ell+1}^\mathsf{P} \subseteq \mathsf{AvgP}$, then $\Sigma_\ell^\mathsf{P} \subseteq \Sigma_\ell \mathsf{TIMEGUESS}[\mathsf{poly}(n), \epsilon n]$.

By Item 2 of Lemma 16, the first witness y_1 can be compressed to a witness z_1 of length ϵn, so that the membership of x in L can be decided using the new polytime relation R_1:

$$x \in L \;\Leftrightarrow\; \exists z_1\, \forall y_2\, \ldots Q y_l \;\; R_1(x, z_1, y_2, \ldots, y_l),$$

where R_1 first decodes y_1 from z_1 in polynomial time given x, and then computes $R_0(x, y_1, \ldots, y_l)$. This observation implies the required claim by induction. □

We now prove the following is a generalization of Theorem 15.

Corollary 18 ([Hir21a])**.** *For all $\ell \in \mathbb{N}$, if $\mathsf{DistNP}^{\Sigma_\ell^\mathsf{P}} \subseteq \mathsf{AvgP}$, then for any function $\tau(n) \leq 2^{O(n^{1-\epsilon})}$ for constant $0 < \epsilon < 1$, $\mathsf{DTIME}^{\Sigma_\ell^\mathsf{P}}[\tau(n)] \subseteq \mathsf{DTIME}[2^{O(n/\log n)}]$.*[4]

Proof. We will give a proof by induction on ℓ. The base case of $\ell = 0$ is trivially true. Suppose the claim holds for some $\ell - 1 \in \mathbb{N}$. Let $\tau(n)$ be as described above, and let $L \in \Sigma_\ell^{\tau(n)}$ with verifier V of complexity $\Pi_{\ell-1}^{\tau(n)}$. Let $y_x \in \{0,1\}^{\tau(n)}$ be the lexicographically first L-witness for x under V. We will show that this y_x can be found in deterministic time $2^{O(n/\log n)}$; this will imply the required result as finding such a witness is known to be a $\mathsf{DTIME}^{\Sigma_\ell^\mathsf{P}}[\tau(n)]$-hard problem [Kre88].

First, note that for all $t \geq t_0(\tau(n))$,

$$\mathsf{K}^{2t,\Sigma_\ell^\mathsf{P}}(x, y_x) \leq \mathsf{K}^t(x) + O(\log t),$$

where the polynomial $t_0(\tau(n))$ reflects the time required to compute y_x given x by a search-to-decision procedure for L using a Σ_ℓ^P-oracle. Applying Item 2 of Lemma 16 to x and y_x, we get that $\mathsf{K}^{\tau(n)^b}(y_x \mid x) \leq n$, for some constant $b \in \mathbb{N}$. This implies a new verifier V' for L, of complexity $\Pi_{\ell-1}^{\mathsf{poly}(\tau(n))}$, such that

$$x \in L_n \;\Leftrightarrow\; \exists z\, (|z| \leq 2n) \;\; V'(x, z),$$

where V' decodes $z \in \{0,1\}^n$ to $y \in \{0,1\}^{\tau(n)}$ in time $\tau(n)^b$, and then simulates $V(x, y)$. Since $\mathsf{poly}(\tau(n)) \leq 2^{O(n^{1-\epsilon})}$, we get by the inductive hypothesis that

$$\{(x, z) \mid V'(x, z)\} \in \Pi_{\ell-1}^{\mathsf{poly}(\tau(n))} \subseteq \mathsf{DTIME}[2^{dn'/\log n'}]$$

for some constant $d \in \mathbb{N}$ and $n' = |x| + |z| \leq 2n$.

Let $z_0 \in \{0,1\}^n$ be such that the universal TM on input (x, z_0) outputs y_x within $\tau(n)^b$ steps. While z_0 may not be the lexicografically first n-bit string to satisfy $V'(x, -)$, it is still possible to compute z_0 from x efficiently, given oracle access to Σ_ℓ^P. Namely, first compute y_x in time $t_0(\tau(n))$ via search-to-decision, and then use NP-oracle queries to produce, via binary search, a string $z \in \{0,1\}^n$ such that the universal TM on input (x, z) outputs y_x within $\tau(n)^b$ steps. It follows that

$$\mathsf{K}^{2t,\Sigma_\ell^\mathsf{P}}(x, z_0) \leq \mathsf{K}^t(x) + O(\log t)$$

for every $t \geq t_0'(\tau(n))$, for some polynomial t_0'. Applying Item 1 of Lemma 16 to x and z_0, there exists a constant $c \in \mathbb{N}$ such that

$$\mathsf{K}^{2^{cn/\log n}}(z_0 \mid x) \leq cn/\log n.$$

[4] In proving a version of this statement for super-polynomial time-bounds τ, the original [Hir21a] uses a padding argument after first proving a version for polynomial time-bounds (see [Hir21a] Theorem 1.8). We could have taken a similar approach here but have opted instead to present an alternative, which relies on the application of Markov's inequality in Lemma 12.

Finally, to find the lexicographically first L-witness y_x for x, we enumerate all strings $\alpha \in \{0,1\}^{cn/\log n}$, running a universal TM on input (x,α) for $2^{cn/\log n}$ steps to obtain some output z_α. For each of these strings z_α, we run a universal TM on input (x, z_α) for $\tau(n)^b$ steps to obtain some output y_α. We then check if $V'(x, z_\alpha)$, using the $2^{O(n/\log n)}$-time deterministic algorithm for V' shown to exist earlier. In the end, we return the lexicographically first y_α obtained such that $V'(x, z_\alpha)$ holds. Clearly, the returned y_α is correct since z_0 is among the z_α's. The total running time of this procedure is at most $2^{cn/\log n} \cdot (2^{cn/\log n} + \tau(n)^b + 2^{O(n/\log n)}) \leq 2^{O(n/\log n)}$. □

4.2 Case of UP

In this section we give a similarly simplified proof of another result from [Hir21a]: if $\mathsf{DistNP} \subseteq \mathsf{AvgP}$ then $\mathsf{UP} \subseteq \mathsf{DTIME}[2^{O(n/\log n)}]$. The following Lemma is analogous to Lemma 11, here for the case of unique witnesses.

Lemma 19. *Suppose* $\mathsf{DistNP} \subseteq \mathsf{AvgP}$. *Let* $\tau(n) \leq 2^{O(n^{1-\epsilon})}$ *for some constant* $0 < \epsilon < 1$. *Then for every* $L \in \mathsf{UTIME}[\tau(n)]$, *there exists a polynomial* q *such that for all sufficiently large* $n \in \mathbb{N}$, *all* $x \in L_n$, *and all* $t \geq n + \tau(n)$,

$$\mathsf{K}^{q(t)}(x, y_x) \leq \mathsf{K}^t(x) + \log q(t),$$

where $y_x \in \{0,1\}^{\tau(n)}$ *is the unique* L*-witness for* x.

Proof. Let $\tau(n)$ be as described above. Let $L \in \mathsf{UTIME}[\tau(n)]$ with verifier V running in deterministic time $\tau(n)$. Let

$$L' := \{(\mathsf{DP}_k(x, y; z), 1^t, 1^s) \mid \mathsf{K}^t(x) \leq s \text{ and } V(x, y) = 1\}$$

for some k to be defined later. Note that $L' \in \mathsf{NP}$. Define a distribution family $D = \{D_{\langle n,t,s\rangle}\}$, each member of which samples $w \sim \{0,1\}^{(n+\tau(n))k+k}$ and outputs $(w, 1^t, 1^s)$. By assumption, $(L', D) \in \mathsf{AvgP}$. Let B be a one-sided error heuristic algorithm for (L', D) with failure probability at most t^{-c}, as in Lemma 5.

Let $x \in \{0,1\}^n$ be sufficiently large, and let $y_x \in \{0,1\}^{\tau(n)}$ be the unique L-witness for x; that is, y_x is the only string y such that $V(x, y)$ holds. Let $t \geq n + \tau(n)$ and $s := \mathsf{K}^t(x)$. Since B never errs on yes-instances of L',

$$\Pr_z[B(\mathsf{DP}_k(x, y_x; z), 1^t, 1^s) = 1] = 1. \tag{10}$$

On the other hand, for w selected uniformly at random,

$$\begin{aligned}
\Pr_w[(w, 1^t, 1^s) \in L'] &= \Pr_w[\exists (x, y) \in \{0,1\}^{n+\tau(n)}, \exists z \in \{0,1\}^{(n+\tau(n))k}, \\
&\qquad\qquad w = \mathsf{DP}_k(u; z) \ \wedge \ \mathsf{K}^t(x) \leq s \ \wedge \ V(x, y) = 1] \\
&\leq \frac{2^{s+1} \cdot 2^{|z|}}{2^{|z|+k}} \\
&= t^{-1},
\end{aligned}$$

where the second line follows from a union bound and a counting argument, and the third line by defining $k := s + 1 + \log t$. Note in particular that for each x, there can only be one y such that $V(x, y)$. Then

$$\begin{aligned}
\Pr_w[B(w, 1^t, 1^s) = 1] &\leq \Pr_w[(w, 1^t, 1^s) \in L'] + \Pr_w[B(w, 1^t, 1^s) \neq L'(w, 1^t, 1^s)] \\
&\leq t^{-c} + t^{-1} \in o(1).
\end{aligned}$$

Comparing with Eq. (10), it is clear that $B(-, 1^t, 1^s)$ is a distinguisher for $\mathsf{DP}_k(x, y_x; \mathcal{U}_{|z|})$. Lemma 9 implies that

$$\begin{aligned}\mathsf{K}^{p'(t)}(x, y_x) &\leq k + \log p'(t) \\ &= s + 1 + \log t + \log p'(t) \\ &= \mathsf{K}^t(x) + 1 + \log t + \log p'(t),\end{aligned}$$

for some polynomial p' with $p'(t) \geq p_{\mathsf{DP}}(2 \cdot (n + \tau(n)) \cdot s_B)$, where p_{DP} is the polynomial from Lemma 9 and s_B denotes the size of a circuit computing $B(-, 1^t, 1^s)$. It follows that $\mathsf{K}^{q(t)}(x, y_x) \leq \mathsf{K}^t(x) + \log q(t)$ for the polynomial $q(t) := 2 \cdot t \cdot p'(t)$. □

Corollary 20 (implicit in [Hir21a])**.** *Suppose* $\mathsf{DistNP} \subseteq \mathsf{AvgP}$. *Then for every function* $\tau(n) \leq 2^{O(n^{1-\epsilon})}$ *for constant* $0 < \epsilon < 1$,

$$\mathsf{UTIME}[\tau(n)] \subseteq \mathsf{DTIME}[2^{O(n/\log n)}].$$

Proof Sketch. Use Lemma 19 with Lemmas 12 and 10 as in the proof of Lemma 16, Item 1. □

5 Open questions

Can one show that $\mathsf{Dist}\Sigma_2^{\mathsf{P}} \subseteq \mathsf{AvgP}$ implies $\mathsf{NTIME}[2^{O(n/\log n)}] = \mathsf{DTIME}[2^{O(n/\log n)}]$? The matching time-bounds in the conclusion would make this analogous in some sense to the [Ben+92] result that $\mathsf{DistNP} \subseteq \mathsf{AvgP}$ implies $\mathsf{NE} = \mathsf{E}$, with the stronger assumption in the former case potentially allowing for a stronger conclusion. [CHV21] provide a version of this statement for the fine-grained setting; namely, $\mathsf{Dist}\Sigma_2\mathsf{TIME}[n] \subseteq \mathsf{AvgTIME}[\tilde{O}(n)]$ implies $\mathsf{NTIME}[2^{O(\sqrt{n \log n})}] = \mathsf{DTIME}[2^{O(\sqrt{n \log n})}]$.

Acknowledgements. We thank Igor C. Oliveira and Zhenjian Lu for their comments on an early version of this manuscript. We also thank Shuichi Hirahara for letting us know about his independent work [Hir21b]. This research was partially supported by NSERC Discovery and NSERC CGS M programs.

References

[Ben+92] Shai Ben-David, Benny Chor, Oded Goldreich, and Michael Luby. "On the Theory of Average Case Complexity". In: *J. Comput. Syst. Sci.* 44.2 (1992), pp. 193–219. DOI: 10.1016/0022-0000(92)90019-F. URL: https://doi.org/10.1016/0022-0000(92)90019-F.

[BFP03] Harry Buhrman, Lance Fortnow, and Aduri Pavan. "Some Results on Derandomization". In: *Proceedings of the 20th Annual Symposium on Theoretical Aspects of Computer Science*. STACS '03. Berlin, Heidelberg: Springer-Verlag, 2003, pp. 212–222. ISBN: 3540006230.

[CHV21] Lijie Chen, Shuichi Hirahara, and Neekon Vafa. "Average-case Hardness of NP and PH from Worst-case Fine-grained Assumptions". In: *Electron. Colloquium Comput. Complex.* (2021), p. 166. URL: https://eccc.weizmann.ac.il/report/2021/166.

[GL89] Oded Goldreich and Leonid A. Levin. "A Hard-Core Predicate for All One-Way Functions". In: *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*. STOC '89. Seattle, Washington, USA: Association for Computing Machinery, 1989, pp. 25–32. ISBN: 0897913078. DOI: 10.1145/73007.73010. URL: https://doi.org/10.1145/73007.73010.

[GSR95] Oded Goldreich, Madhu Sudan, and Ronitt Rubinfeld. "Learning polynomials with queries: The highly noisy case". In: *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*. Los Alamitos, CA, USA: IEEE Computer Society, Oct. 1995, p. 294. DOI: 10.1109/SFCS.1995.492485. URL: https://doi.ieeecomputersociety.org/10.1109/SFCS.1995.492485.

[Hir21a] Shuichi Hirahara. "Average-Case Hardness of NP from Exponential Worst-Case Hardness Assumptions". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2021. Virtual, Italy: Association for Computing Machinery, 2021, pp. 292–302. ISBN: 9781450380539. DOI: 10.1145/3406325.3451065. URL: https://doi.org/10.1145/3406325.3451065.

[Hir21b] Shuichi Hirahara. "Symmetry of Information in Heuristica". manuscript. 2021.

[Kre88] Mark W. Krentel. "The Complexity of Optimization Problems". In: *J. Comput. Syst. Sci.* 36.3 (1988), pp. 490–509. DOI: 10.1016/0022-0000(88)90039-6. URL: https://doi.org/10.1016/0022-0000(88)90039-6.

[LR05] Troy Lee and Andrei E. Romashchenko. "Resource bounded symmetry of information revisited". In: *Theor. Comput. Sci.* 345.2-3 (2005), pp. 386–405. DOI: 10.1016/j.tcs.2005.07.017. URL: https://doi.org/10.1016/j.tcs.2005.07.017.

[LW95] Luc Longpré and Osamu Watanabe. "On Symmetry of Information and Polynomial Time Invertibility". In: *Inf. Comput.* 121.1 (1995), pp. 14–22. DOI: 10.1006/inco.1995.1120. URL: https://doi.org/10.1006/inco.1995.1120.

[Vad12] Salil P. Vadhan. *Pseudorandomness*. Hanover, MA, USA: Now Publishers Inc., 2012. ISBN: 1601985940.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il