

# Proving that $prBPP = prP$ is as hard as proving that “almost $NP$ ” is not contained in $P/poly^*$

Roei Tell <sup>†</sup>

January 10, 2019

## Abstract

What circuit lower bounds are *necessary* in order to prove that  $promise-BPP = promise-P$ ? We show that the recent breakthrough result of Murray and Williams (STOC 2018) can be used to show a dramatic strengthening of the previously-known answer to this question. Specifically, we show that if  $promise-BPP = promise-P$ , then  $NTIME[n^{f(n)}] \not\subseteq P/poly$ , for essentially any  $f(n) = \omega(1)$ .

We also prove a technical strengthening of this result. Specifically, we show that if  $promise-BPP = promise-P$ , then for essentially any  $s : \mathbb{N} \rightarrow \mathbb{N}$  it holds that  $NTIME[s^{O(1)} \circ s^{O(1)}] \not\subseteq SIZE[s]$ . Moreover, we show that size- $s$  circuits fail to compute the “hard” function in any interval of length  $\text{poly}(s(\text{poly}(n)))$ . The proof of this result uses tools of Murray and Williams, but relies on a different proof strategy. (Their proof strategy yields three compositions of  $s$  instead of two, and does not yield the guarantee of failure in any small interval.)

Lastly, we present an alternative proof of the main result, which only relies on a generalization of the well-known lower bound of Santhanam (SICOMP, 2009).

---

\*The current paper is a revised version of a technical report that appeared online under a slightly different title (ECCC, TR18-003). The current version includes additional results and a revised exposition.

<sup>†</sup>Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel. Email: roei.tell@weizmann.ac.il

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	The main new result . . . . .	1
1.2	A technical strengthening of Theorem 2 . . . . .	3
1.3	The meaning of the results in this paper . . . . .	3
1.4	Organization . . . . .	5
<b>2</b>	<b>Overviews of the proofs</b>	<b>5</b>
2.1	Proof overview for Theorem 2 . . . . .	5
2.2	Proof overview for Theorem 3 . . . . .	7
<b>3</b>	<b>Preliminaries</b>	<b>8</b>
<b>4</b>	<b>Proof of Theorems 1 and 2</b>	<b>11</b>
4.1	A parametrized “derandomization implies lower bounds” theorem . . . . .	11
4.2	Theorems 1 and 2 as corollaries . . . . .	13
<b>5</b>	<b>Proof of Theorem 3</b>	<b>14</b>
	<b>Acknowledgements</b>	<b>18</b>
	<b>Appendix A An alternative proof of Theorem 2</b>	<b>20</b>
	<b>Appendix B Sufficient conditions for admissibility</b>	<b>23</b>

# 1 Introduction

The  $BPP = P$  conjecture asserts that any decision problem that can be efficiently solved using randomness (while allowing for a small error) can also be efficiently solved deterministically. In other words, the conjecture asserts that randomness is not needed to efficiently solve decision problems. This conjecture is central to the complexity-theoretic study of the role of randomness in computation.

The  $BPP = P$  conjecture is often interpreted as an *algorithmic* problem, namely the problem of explicitly constructing efficient deterministic algorithms that simulate randomized algorithms. In fact, a version of the conjecture is *equivalent* to the conjectured existence of an algorithm for a single, specific problem (i.e., the *circuit acceptance probability problem*; see Proposition 6). However, as will be discussed next, it has also been known for at least two decades that the conjecture is actually intimately related to *circuit lower bounds*; that is, to lower bounds for non-uniform models of computation.

Informally, following a very recent breakthrough by Murray and Williams [MW18], the main result in this paper considerably strengthens the known connection between the  $BPP = P$  conjecture and circuit lower bounds. To present the new result, let us first spell out the previously-known connections:

- On the one hand, *any proof of sufficiently strong circuit lower bounds would also prove the  $BPP = P$  conjecture*. Specifically, if there is a function in  $\mathcal{E}$  that requires exponential-sized circuits, then  $BPP = P$  (and even  $prBPP = prP$ , i.e. the promise-problem versions of  $BPP$  and of  $P$  are equal; see [IW99], which relies on the hardness-randomness paradigm [Yao82; BM84; NW94]).
- On the other hand, *any proof that  $prP = prBPP$  implies long-sought circuit lower bounds*. As a prominent example, any proof that  $prBPP = prP$  implies that there exists a function in  $\mathcal{NEXPTIME}$  that cannot be computed by any polynomial-sized circuit family [BFT98].<sup>1</sup> In fact, the latter circuit lower bound follows even from much weaker hypotheses (e.g., it follows from the hypothesis that  $MA \neq \mathcal{NEXPTIME}$ ; see, e.g., [IKW02; Wil13]).

## 1.1 The main new result

The starting point of the current work is the observation that an immediate corollary of a result from the recent work of Murray and Williams [MW18, Thm 1.2] is the following: *If  $prBPP = prP$ , then there exists a function in  $NTIME[n^{\text{poly} \log(n)}]$  (rather than  $\mathcal{NEXPTIME}$ ) that cannot be computed by any polynomial-sized circuit family*. This is a dramatic (almost exponential) strengthening of previously-known results (i.e., of [BFT98; IKW02]), and we believe that it is a fundamental result that is worth spelling out and highlighting. Furthermore, this result can even be further strengthened. In particular,

---

<sup>1</sup>In [BFT98] it is shown, unconditionally, that  $MA \notin \mathcal{NEXPTIME} / \text{poly}$ . Thus, under the hypothesis  $prBPP \subseteq prP$  we have that  $MA \notin \mathcal{NEXPTIME} / \text{poly}$  (see [IKW02, Rmk. 26]).

by using the proof approach of [MW18] while instantiating their technical tools with different parameters, we get the following:

**Theorem 1** (*main theorem; informal*). *If  $\text{pr}\mathcal{BPP} = \text{pr}\mathcal{P}$ , then, for essentially any super-constant function  $f(n) = \omega(1)$ , there exists a set in  $\text{NTIME}[n^{f(n)}] \setminus \mathcal{P}/\text{poly}$ .*

One might a-priori hope to strengthen the conclusion of Theorem 1 by improving the time bound in the non-deterministic class; that is, to prove that “if  $\text{pr}\mathcal{BPP} = \text{pr}\mathcal{P}$ , then  $\mathcal{NP} \not\subseteq \mathcal{P}/\text{poly}$  (and  $\mathcal{P} \neq \mathcal{NP}$ )”. However, such a result cannot be proved without *unconditionally* proving that  $\mathcal{P} \neq \mathcal{NP}$ , since any proof of the conditional statement “ $\text{pr}\mathcal{BPP} = \text{pr}\mathcal{P} \implies \mathcal{P} \neq \mathcal{NP}$ ” would unconditionally imply that  $\mathcal{P} \neq \mathcal{NP}$  (see Proposition 12). Therefore, the conclusion of Theorem 1 is optimal in this sense.

Theorem 1 is a special case of a more general “derandomization implies lower bounds” result that follows using the tools of Murray and Williams [MW18]. In this general result, the circuit lower bound in the conclusion can be parameterized:

**Theorem 2** (*a generalized version of Theorem 1; informal, see Corollary 15*). *There exists  $\epsilon > 0$  such that for any time-computable  $s : \mathbb{N} \rightarrow \mathbb{N}$  satisfying  $n < s(n) < 2^{\epsilon \cdot n}$  it holds that*

$$\text{pr}\mathcal{BPP} = \text{pr}\mathcal{P} \implies \text{NTIME}[s' \circ s' \circ s'] \not\subseteq \text{SIZE}[s],$$

where  $s' = \text{poly}(s(O(n)))$ .

Indeed, Theorem 1 follows as a special case of Theorem 2 by using  $s(n) = n^{\omega(1)}$  (in which case  $s' \circ s' \circ s' = n^{\omega(1)}$  and  $\text{SIZE}[s] \supset \mathcal{P}/\text{poly}$ ; see Corollary 17). The hypothesis of Theorem 2 can also be significantly relaxed, since its proof relies on Williams’ [Wil13] celebrated proof strategy (which is well-known to support such relaxations). For example, the theorem holds under the hypothesis  $\text{pr-co}\mathcal{RP} \subseteq \text{pr}\mathcal{NP}$ , and also under the hypothesis that there exists a “non-trivial” algorithm for the Circuit Acceptance Probability Problem (CAPP) (i.e., an algorithm that approximates the acceptance probability of a circuit of size  $m$  with  $v$  variables in time  $2^{99 \cdot v} \cdot \text{poly}(m)$ ). See Section 4 for precise details of these relaxations.

Theorem 2 may be compared to the following result of Kinne, van Melkebeek, and Shaltiel [KMS12, Thm. 9]: If  $\mathcal{BPP} = \mathcal{P}$  (i.e., the “non-promise” version of the  $\mathcal{BPP} = \mathcal{P}$  conjecture holds),<sup>2</sup> then either the *permanent function* of  $\{0,1\}$ -matrices over  $\mathbb{Z}$  does not have polynomial-sized arithmetic circuits, or for essentially any  $s : \mathbb{N} \rightarrow \mathbb{N}$  it holds that  $\text{NTIME}[s^{O(1)}] \not\subseteq \text{SIZE}[s]$ . Indeed, in this result from [KMS12] the hypothesis is weaker than in Theorem 2 (since it only refers to the “non-promise” version of  $\mathcal{BPP} = \mathcal{P}$ ), and the conclusion is not a circuit lower bound as in Theorem 2, but rather a disjunction of two circuit lower bounds. However, note that in the lower bound  $\text{NTIME}[s^{O(1)}] \not\subseteq \text{SIZE}[s]$ , the time-complexity of the “hard” function is just  $s^{O(1)}$ , rather than  $s^{O(1)} \circ s^{O(1)} \circ s^{O(1)}$  as in Theorem 2.

<sup>2</sup>Similarly to Theorem 2, the result in [KMS12] also follows from the weaker hypothesis  $\text{co}\mathcal{RP} \subseteq \mathcal{NP}$ .

## 1.2 A technical strengthening of Theorem 2

In this section we present a technical strengthening of Theorem 2. Recall that the “hard” function in the conclusion of Theorem 2 was in  $NTIME[s^{O(1)} \circ s^{O(1)} \circ s^{O(1)}]$ . In particular, since there are three compositions of  $s$ , the concluded lower bound becomes trivial when  $s$  is half-exponential or larger (i.e., when  $s(s(n)) \geq 2^n$ ).<sup>3</sup> The current improvement removes this limitation: We prove that if  $pr\mathcal{BPP} = pr\mathcal{P}$ , then there exists a function in  $NTIME[s^{O(1)} \circ s^{O(1)}]$  that cannot be computed by circuits of size  $s$ .

Moreover, we also improve the concluded lower bound by showing that size- $s$  circuits fail to compute the “hard” function on a “dense” set of input lengths (the conclusion in Theorem 2 only guarantees failure on infinitely-many input lengths). Specifically, for  $s_I(n) = \text{poly}(s(\text{poly}(n)))$ , we conclude that size- $s$  circuits fail to compute the “hard” function on an input length in any interval of the form  $[n, s_I(n)]$ .

Similarly to Theorem 2, the foregoing (stronger) conclusions follow also from a hypothesis that is weaker than  $pr\mathcal{BPP} = pr\mathcal{P}$ . Specifically, the conclusions follow from the hypothesis that  $pr\mathcal{BPP} \subseteq pr\mathcal{NP}$ :

**Theorem 3** (strengthening the conclusion of Theorem 2; informal, see Theorem 23). *There exists  $\epsilon > 0$  such that for any time-computable  $s : \mathbb{N} \rightarrow \mathbb{N}$  satisfying  $n < s(n) < 2^{\epsilon \cdot n}$  it holds that*

$$pr\mathcal{BPP} \subseteq pr\mathcal{NP} \implies NTIME[s' \circ s'] \not\subseteq \text{i.o.}_{[s_I]}\text{-SIZE}[s],$$

where  $s' = \text{poly}(s)$ , and  $\text{i.o.}_{[s_I]}\text{-SIZE}[s]$  is the class of problems such that there exists a size- $s$  circuit that, for infinitely-many intervals of length  $s_I(n) = \text{poly}(s(n^2))$ , solves the problem on some input length in the interval.

The proof of Theorem 3 does not follow the proof approach of Murray and Williams, and in particular does not use their new “easy witness lemma”. Nevertheless, the proof crucially relies on one of their technical results, namely their strengthening of Santhanam’s circuit lower bound [San09]. See Section 2.2 for further details.

In Appendix A we present an alternative and relatively simple proof of a weaker form of Theorem 3 (the weaker form does not include the guarantee of failure in every “small” interval). This proof does not use the results of Murray and Williams, but is based only on (a generalization of) the well-known circuit lower bound of Santhanam [San09]. The idea for this alternative proof was suggested to us by Igor Oliveira (after a preliminary version of this paper appeared online).

## 1.3 The meaning of the results in this paper

What is the meaning of the results in this paper? Let us first focus on Theorem 1; that is, on the statement “ $pr\mathcal{BPP} = pr\mathcal{P} \implies NTIME[n^{\omega(1)}] \not\subseteq \mathcal{P}/\text{poly}$ ”. Note that the

---

<sup>3</sup>This is since when  $s(s(n)) \geq 2^n$  we have that  $NTIME[s^{O(1)} \circ s^{O(1)} \circ s^{O(1)}] \supseteq NTIME[2^{\text{poly}(s(n))}]$ , whereas  $DTIME[2^{\text{poly}(s(n))}] \not\subseteq \text{SIZE}[s]$  holds unconditionally (by a diagonalization argument).

lower bound  $NTIME[n^{\omega(1)}] \not\subseteq \mathcal{P}/\text{poly}$  asserts that polynomial-sized circuits cannot simulate both “slightly” super-polynomial running time and non-determinism.<sup>4</sup>

Thus, on the one hand, one may view the lower bound  $NTIME[n^{\omega(1)}] \not\subseteq \mathcal{P}/\text{poly}$  as a weaker form of  $\mathcal{NP} \not\subseteq \mathcal{P}/\text{poly}$ . From this perspective, Theorem 1 can be interpreted as saying that proving that  $pr\mathcal{BPP} = pr\mathcal{P}$  is as hard as proving a lower bound that is essentially a precursor of  $\mathcal{NP} \subseteq \mathcal{P}/\text{poly}$ . On the other hand, as pointed out by Ryan Williams, one can alternatively view the lower bound  $NTIME[n^{\omega(1)}] \not\subseteq \mathcal{P}/\text{poly}$  as a weaker form of the statement  $DTIME[n^{\omega(1)}] \not\subseteq \mathcal{P}/\text{poly}$ . The latter statement asserts that polynomial-sized circuits cannot simulate algorithms with superpolynomial running time. From this perspective, Theorem 1 implies that proving that  $pr\mathcal{BPP} = pr\mathcal{P}$  is as hard as proving a weak form of a “strengthened time-hierarchy theorem” (in which we compare uniform algorithms to non-uniform circuits).

In fact, continuing the latter view, it seems instructive to compare the lower bounds implied by  $pr\mathcal{BPP} = pr\mathcal{P}$  to the lower bounds that are known to imply  $pr\mathcal{BPP} = pr\mathcal{P}$  (using the results of Impagliazzo and Wigderson [IW99]). Specifically, being slightly informal,<sup>5</sup> we have that:

$$\begin{aligned} \forall s(n) < 2^{\epsilon \cdot n}, \text{DTIME}[\text{poly}(s)] &\not\subseteq \text{i.o.}\text{-SIZE}[s] && (1) \\ \Downarrow &&& \text{(by [IW99])} \\ pr\mathcal{BPP} &= pr\mathcal{P} \\ \Downarrow &&& \text{(by Thm 3)} \\ \forall s(n) < 2^{\epsilon \cdot n}, \text{NTIME}[s' \circ s'] &\not\subseteq \text{i.o.}_{[s_I]}\text{-SIZE}[s] && (2) \end{aligned}$$

where  $\epsilon > 0$ ,  $s'$ , and  $s_I$  are defined as in Theorem 3.

Indeed, the lower bounds in Eq. (1) are stronger than the lower bounds in Eq. (2): This is both since the “hard” function in Eq. (2) lies in a complexity class that is larger than that of the “hard” function in Eq. (1) (due to the use of non-determinism and to the two compositions of  $s$ ); and since in Eq. (2) failure of size- $s$  circuits to compute the function is guaranteed only in any “small” interval, whereas in Eq. (1) this failure is guaranteed almost everywhere. This comparative perspective suggests the following interpretation of the results in this paper:

The lower bounds implied by  $pr\mathcal{BPP} = pr\mathcal{P}$  are now significantly stronger (compared to the lower bounds that were previously known to be implied by this conjecture); but they are nevertheless still weaker than the lower bounds that are known to imply that  $pr\mathcal{BPP} = pr\mathcal{P}$ .

**Is  $pr\mathcal{BPP} = pr\mathcal{P}$  equivalent to a specific circuit lower bound?** The circuit lower bounds implied in Eq. (2) hold not only when  $pr\mathcal{BPP} = pr\mathcal{P}$ , but also under the

---

<sup>4</sup>This lower bound can be viewed as a significant strengthening of the (unconditionally-known) lower bound  $\Sigma_3[n^{\omega(1)}] \not\subseteq \mathcal{P}/\text{poly}$ , which asserts that polynomial-sized circuits cannot simulate both super-polynomial running time and “several levels” of non-determinism/alterations. (The proof of the lower bound is a diagonalization argument a-la Kannan’s theorem; see, e.g., [Juk12, Lem. 20.12].)

<sup>5</sup>The informality is by ignoring time-computability constraints on  $s$ .

(intuitively) weaker hypothesis  $pr\mathcal{BPP} \subseteq pr\mathcal{NP}$ . Therefore, one might suspect that the conclusion in Theorem 2 can be strengthened. Moreover, recall that the question of whether specific derandomization results are *equivalent* to *specific* circuit lower bounds has been raised several times in the past (see, e.g., [IKW02, Beginning of the Introduction] and [TV07, Sec. 1.1]). We thus propose the following natural conjecture (we have found no explicit prior mentions of this conjecture in the literature):

**Conjecture 4** ( $pr\mathcal{BPP} = pr\mathcal{P}$  is equivalent to the [IW99] lower bounds). *The statement that  $pr\mathcal{BPP} = pr\mathcal{P}$  is equivalent to the statement that for some  $\epsilon > 0$  and every  $s(n) < 2^{\epsilon n}$  it holds that  $DTIME[\text{poly}(s)] \not\subseteq \text{i.o.-SIZE}[s]$ .*

The most important gap between Theorem 3 and Conjecture 4 is that in Theorem 3, the lower bounds implied by  $pr\mathcal{BPP} = pr\mathcal{P}$  are against *non-deterministic* classes. Note that even a modest first step towards proving Conjecture 4, namely proving that  $pr\mathcal{BPP} = pr\mathcal{P} \implies \mathcal{EXPTIME} \not\subseteq \mathcal{P}/\text{poly}$ , already implies that any polynomial-time derandomization of  $pr\mathcal{BPP}$  requires pseudorandom generators (see [BFN+93]).

## 1.4 Organization

In Section 2 we present high-level overviews of the proofs of our main theorems. In Section 3 we present preliminary definitions. In Section 4 we prove Theorems 1 and 2, and in Section 5 we prove Theorem 3.

## 2 Overviews of the proofs

In Section 2.1 we present an overview of the proof of Theorem 2, and in Section 2.2 we present an overview of the proof of Theorem 3. Since the proof approaches for the two theorems are very different, one may read Section 2.2 without first reading Section 2.1.

### 2.1 Proof overview for Theorem 2

The proof of Theorem 2 follows the approach used by Murray and Williams [MW18], which is based on the celebrated proof strategy of Williams [Wil13]. The main new component in [MW18] is a new “easy witness lemma”, which allows for flexible scaling of the parameters in the original proof strategy of Williams (see below; this new lemma improves the original easy witness lemma of [IKW02]). Murray and Williams stated consequences with two *specific parameter settings*. We extend their result by stating a *general* (parametrized) “derandomization implies lower bounds” result that uses this proof approach with the new easy witness lemma (see Theorem 14), and deduce Theorems 1 and 2 as special cases.

Let us now overview the proof of Theorem 2. The point of the overview is to describe how the (well-known) proof strategy of Williams can be instantiated with the new easy witness lemma for general parameters in order to deduce Theorem 2. The starting point for the proof is the Circuit Acceptance Probability Problem (or CAPP,

in short): Given as input the description of a Boolean circuit  $C$ , the problem is to distinguish between the case that the acceptance probability of  $C$  is at least  $2/3$  and the case that the acceptance probability of  $C$  is at most  $1/3$ . It is well-known that a deterministic polynomial-time algorithm for CAPP exists if and only if  $pr\mathcal{BPP} = pr\mathcal{P}$  (see Proposition 6). The current argument relies on the much weaker hypothesis that CAPP for circuits of size  $m$  with  $v$  input variables can be solved in time  $2^{99 \cdot v} \cdot \text{poly}(m)$ ; for simplicity, let us assume that the CAPP algorithm runs in time  $2^{99 \cdot v} \cdot m^2$ .

Fix any time-computable function  $n < s(n) < 2^{\epsilon \cdot n}$ , where  $\epsilon > 0$  is a universal constant. Denoting  $t = s^{O(1)} \circ s^{O(1)} \circ s^{O(1)}$ , our goal is to prove that  $NTIME[t] \not\subseteq SIZE[s]$ . (The definition of  $t$  in this high-level overview is slightly informal; see Definition 9 and Corollary 15 for precise details.) To do so, assume towards a contradiction that  $NTIME[t] \subseteq SIZE[s]$ , and let  $t_0(n) = t(n)^\delta$ , where  $\delta > 0$  is sufficiently small. We will construct, for any  $L \in NTIME[t_0]$ , a non-deterministic machine that decides  $L$  in time  $t_0^{1-\Omega(1)}$ ; this will contradict the non-deterministic time hierarchy [Coo72].

Using the new easy witness lemma, if  $NTIME[t] \subseteq SIZE[s]$  where  $t = t_0^{1/\delta}$ , then for every  $L' \in NTIME[(t_0)^2]$ , every  $(t_0)^2$ -time verifier  $V$  for  $L'$  and every  $x \in L'$ , there exists a circuit  $P_x \in SIZE[t_0^{001}]$  that encodes a witness  $\pi_x$  such that  $V(x, \pi_x)$  accepts.<sup>6</sup> (Again, our parameters in the overview are informal; see Lemma 10 for a statement that uses precise parameters.) The point is that witnesses for the verifier  $V$  are a-priori of size  $(t_0)^2$ , but the lemma asserts that (under the hypothesis) every  $x \in L'$  has a witness that can be concisely represented by a circuit of much smaller size  $t_0^{001}$ . We note that the main “bottleneck” in the proof that requires using  $t = s^{O(1)} \circ s^{O(1)} \circ s^{O(1)}$  (rather than, say,  $t = \text{poly}(s)$ ) is the new “easy witness lemma”.

Let us now construct the non-deterministic machine for  $L \in NTIME[t_0]$ , relying on the existence of the foregoing “compressible” witnesses. We first fix a PCP system for  $L$  with a verifier  $V$  that runs in time  $t_V = \text{poly}(n, \log(t_0))$  and uses  $\ell = \log(t_0) + O(\log \log(t_0))$  random bits. (For concreteness, we use the PCP of Ben-Sasson and Viola [BV14], but previous ones such as [BGH+05] also suffice for the proof.) Using the new easy witness lemma, for every  $x \in L$  there exists a circuit of size  $t_0(|x|)^{001}$  that encodes a valid proof for  $x$  in this PCP system.<sup>7</sup>

Now, given input  $x \in \{0,1\}^n$ , the non-deterministic machine  $M$  first guesses a circuit  $P_x$  of size  $t_0(n)^{001}$ , in the hope that such a circuit encodes a valid proof for  $x$ . Then, the machine constructs a circuit  $C_x^{P_x}$  that, when given  $r \in \{0,1\}^\ell$  as input, simulates the execution of  $V$  on  $x$  using randomness  $r$  when  $V$  is given oracle access to the witness  $P_x$  (i.e.,  $C_x^{P_x}(r) = V^{P_x}(x, r)$ ). Finally, the machine  $M$  uses the CAPP algorithm on the circuit  $C_x^{P_x}$  to determine whether the verifier accepts  $x$  with high probability over  $r$  or rejects  $x$  with high probability over  $r$ .

Note that if  $x \in L$ , then for *some* guess of  $P_x$  it holds that  $C_x^{P_x}$  has acceptance

<sup>6</sup>A circuit  $P_x : \{0,1\}^{\log(|\pi_x|)} \rightarrow \{0,1\}$  encodes a string  $\pi_x$  if for every  $i \in [|\pi_x|]$  it holds that  $P_x(i)$  is the  $i^{\text{th}}$  bit of  $\pi_x$  (equivalently,  $\pi_x$  is the truth-table of  $P_x$ ).

<sup>7</sup>To apply the easy witness lemma, consider the deterministic verifier  $V'$  that, when given input and a proof, enumerates the random coins of  $V$  and decides by a majority vote. This verifier runs in time  $2^\ell \cdot t_V < (t_0)^2$ , so we can apply the lemma to  $L$  with this verifier.

probability one, and thus the machine  $M$  will accept  $x$ . On the other hand, if  $x \notin L$ , then for *any* guess of  $P_x$  it holds that  $C_x^{P_x}$  has low acceptance probability (corresponding to the soundness of the PCP verifier), and thus the machine  $M$  will reject  $x$ .

The point is that all the operations of the machine happened in time much shorter than  $t_0(n)$ . Specifically, the size of  $P_x$  is  $t_0(n)^{.001}$ , and the size of  $C_x^{P_x}$  is  $m < t_V(n) \cdot t_0(n)^{.001} < t_0(n)^{.002}$ ; thus, guessing  $P_x$  and constructing  $C_x^{P_x}$  can be done in time  $\text{poly}(m) \ll \sqrt{t_0(n)}$ . Now, note that  $C_x^{P_x}$  has  $\ell = \log(t_0) + O(\log \log(t_0))$  variables; thus, when the CAPP algorithm is given  $C_x^{P_x}$  it runs in time

$$2^{99 \cdot \ell} \cdot m^2 < t_0(n)^{.995} \cdot (t_0(n)^{.002})^2 = (t_0(n))^{1-\Omega(1)},$$

and we get a contradiction.

As mentioned in the introduction, the hypothesis in this proof strategy can be further relaxed in various (known) ways. For details of these relaxations, see the statement of Theorem 14 and the remark following the theorem's proof.

## 2.2 Proof overview for Theorem 3

The proof of Theorem 3 is very different than the proof of Theorem 2, and in particular does *not* rely on the proof strategy of Williams [Wil13] or on an “easy witness lemma”. Recall that we assume that  $\text{prBPP} \subseteq \text{prNP}$ , and want to deduce that for essentially any  $s : \mathbb{N} \rightarrow \mathbb{N}$  it holds that  $\text{NTIME}[t] \not\subseteq \text{i.o.}_{[s_I]}\text{-SIZE}[s]$ , where  $t = s^{O(1)} \circ s^{O(1)}$  and the prefix  $\text{i.o.}_{[s_I]}$  means that in any interval of length  $s_I(n) = \text{poly}(s(n^2))$  there exists an input length in which size- $s$  circuits fail to compute the “hard” function.<sup>8</sup>

The starting point of the proof is Murray and Williams' strengthening of Santhanam's circuit lower bound [San09]. The strengthening asserts (unconditionally) that there exists a set  $S \subseteq \{0, 1\}^*$  that can be decided by Merlin-Arthur protocols running in time  $t$  with  $\ell = O(\log(s))$  bits of non-uniform advice (i.e.,  $S \in \text{MATIME}[t]/\ell$ ) such that  $S \not\subseteq \text{i.o.}_{[\text{poly}(s)]}\text{-SIZE}[s]$  (see Theorem 19).

The first observation in the proof of Theorem 3 is that if  $\text{prBPP} \subseteq \text{prNP}$ , then the Merlin-Arthur protocol that decides  $S$  with non-uniform advice can be derandomized, in a straightforward way (see Proposition 20). Hence, under our hypothesis, the set  $S$  is in  $\text{NTIME}[\text{poly}(t)]/\ell$ , which implies that  $\text{NTIME}[\text{poly}(t)]/\ell \not\subseteq \text{i.o.}_{[\text{poly}(s)]}\text{-SIZE}[s]$ .

The second observation is that if  $\text{NTIME}[\text{poly}(t)]/\ell \not\subseteq \text{i.o.}_{[\text{poly}(s)]}\text{-SIZE}[s]$ , then  $\text{NTIME}[\text{poly}(t)]$  (without the advice) is not contained in  $\text{i.o.}_{[\text{poly}(s)]}\text{-SIZE}[s']$ , for  $s'$  that is moderately smaller than  $s$ . Let us first prove this statement while ignoring the issue of failure in almost all intervals for a moment. Assuming towards a contradiction that  $\text{NTIME}[\text{poly}(t)] \subseteq \text{SIZE}[s']$ , for any  $S \in \text{NTIME}[\text{poly}(t)]/\ell$  we construct a family of size- $s$  circuits that decides  $S$ . To do so, consider a non-deterministic machine  $M$  that decides  $S$  with advice  $\{a_n\}$ , and let  $S^{\text{adv}}$  be the set of pairs  $(x, \sigma)$  such that  $|\sigma| = \ell(|x|)$  and  $M$  (non-deterministically) accepts  $x$  when given advice  $\sigma$ . Note that

<sup>8</sup>The actual lower bound is even slightly stronger, since it asserts that the circuit family fails to compute  $S$  on (at least) one of the “end-points” of the interval. For further details see Section 5.

$S^{\text{adv}}$  can be decided by a non-deterministic machine that simulates  $M$  (and requires no advice), and thus, by our hypothesis,  $S^{\text{adv}}$  can be solved by a circuit family  $\{C_n\}$  of size  $s'$ .<sup>9</sup> By hard-wiring the “good” advice  $a_n$  into each  $C_n$ , we obtain a circuit family  $\{C'_n\}$  of size  $s'$  that decides  $S$ . Note that the size of the circuit is still  $s'$ , but it is now a function of a smaller input length, since we “hard-wired” the advice in place of input bits; however, since the advice is relatively short (i.e.,  $\ell = O(\log(s))$ ), the new size function, denoted  $s$ , is not much larger than  $s'$  (see Proposition 22).

This “elimination of advice” argument extends to the setting where failure is guaranteed in any “small” interval, with a bit of care. The source of trouble is that now our “towards-a-contradiction” hypothesis only implies that  $S^{\text{adv}} \in \text{i.o.}_{[\text{poly}(s)]}\text{-SIZE}[s]$ , which only guarantees the existence of an infinite “dense” set  $I \subseteq \mathbb{N}$  of input lengths for which  $S^{\text{adv}}$  has small circuits. In particular, we have no guarantee that every  $n \in I$  is of the form  $m + \ell(m)$ , which is what we need to deduce that  $S_m = S \cap \{0,1\}^m$  has small circuits. To overcome this problem, we “embed” all pairs  $(x, \sigma)$  such that  $|\sigma| = \ell(|x|)$  and  $|x| + |\sigma| < n$  into  $\{0,1\}^n$ , and define  $S_n^{\text{adv}} = S^{\text{adv}} \cap \{0,1\}^n$  such that deciding  $S_n^{\text{adv}}$  allows to determine the output of  $M$  on  $(x, \sigma)$  for all pairs satisfying  $|x| + |\sigma| < n$ . Thus, for any  $n \in I$ , a circuit of size  $s(n)$  that decides  $S_n^{\text{adv}}$  allows us to solve  $S_m$  where  $m + \ell(m) < n$ . And similarly to above, since the advice is relatively small (i.e.,  $\ell(m) = O(\log(s(m))) < m$ ), both the size  $s(n)$  of the circuit and the interval length  $\text{poly}(s(n))$  in which failure is guaranteed are not too large as a function of  $m$ . For precise details see the proof of Proposition 22.

### 3 Preliminaries

We assume familiarity with basic notions of complexity theory; for background see, e.g., [Gol08; AB09]. Throughout the paper, fix any standard model of a Turing machine (we need a fixed model since we discuss time-constructible functions).

Whenever we refer to circuits (without qualifying which type), we mean non-uniform circuit families over the De-Morgan basis (i.e., AND/OR/NOT gates) with fan-in at most two and unlimited fan-out, and without any specific structural restrictions (e.g., without any limitation on their depth). The size of a circuit is the number of its gates. Moreover, we consider some fixed standard form of representation for such circuits, where the representation size is polynomial in the size of the circuit.

We use the standard notation  $\text{i.o.}\text{-SIZE}[s]$  to denote the class of sets that can be decided by some size- $s$  circuit on infinitely-many input lengths. Extending this notation, we denote by  $\text{i.o.}_{[q]}\text{-SIZE}[s]$  the class of problems such that there exists a size- $s$  circuit that, for infinitely-many intervals of the form  $[n, q(n)]$ , solves the problem on some input length in the interval. We warn in advance that in Section 5 we slightly abuse this notation, by writing  $S \notin \text{i.o.}_{[q]}\text{-SIZE}[s]$  to deduce a slightly stronger conclusion

---

<sup>9</sup>Note that the foregoing argument only follows through after the “derandomization” (i.e., for  $NTIME$  and not for  $MATIME$ ). This is the case since when dealing with probabilistic machines, it is not clear how to define  $S^{\text{adv}}$  in a way that will allow a probabilistic machine without advice to decide it (since a probabilistic machine that is given a “wrong” advice might not “distinctly” accept or reject some inputs).

than failure on *some* input in almost all intervals of length  $q$  (i.e., we use it to denote failure on one of the end-points in almost all such intervals; see Definition 18).

### 3.1 Circuit acceptance probability problem

We now formally define the circuit acceptance probability problem (or CAPP, in short); this well-known problem is also sometimes called Circuit Derandomization, Approx Circuit Average, and GAP-SAT or GAP-UNSAT.

**Definition 5** (CAPP). *The circuit acceptance probability problem with parameters  $\alpha, \beta \in [0, 1]$  such that  $\alpha > \beta$  (or  $(\alpha, \beta)$ -CAPP, in short) is the following promise problem:*

- *The YES instances are (representations of) circuits that accept at least  $\alpha$  of their inputs.*
- *The NO instances are (representations of) circuits that accept at most  $\beta$  of their inputs.*

*We define the CAPP problem (i.e., omitting  $\alpha$  and  $\beta$ ) as the  $(2/3, 1/3)$ -CAPP problem.*

It is well-known that CAPP is complete for  $pr\mathcal{BPP}$  under deterministic polynomial-time reductions; in particular, CAPP can be solved in deterministic polynomial time if and only if  $pr\mathcal{BPP} = pr\mathcal{P}$ .

**Proposition 6** (CAPP is equivalent to  $pr\mathcal{BPP} = pr\mathcal{P}$ ). *The circuit acceptance probability problem can be solved in deterministic polynomial time if and only if  $pr\mathcal{BPP} = pr\mathcal{P}$ .*

For a proof of Proposition 6 see any standard textbook on the subject (e.g. [Vad12, Cor. 2.31], [Gol08, Exer. 6.14]). In Proposition 6 we considered the complexity of CAPP as a function of the input size, which is the size of the (description of the) circuit. However, following [Wil13], it can also be helpful to consider the complexity of CAPP as a function of both the circuit size  $m$  (which corresponds to the input size) and of the number  $v$  of input variables to the circuit. In this case, a naive deterministic algorithm can solve the problem in time  $2^v \cdot \text{poly}(m)$ , whereas the naive probabilistic algorithm solves the problem in time  $v \cdot \text{poly}(m) \leq \text{poly}(m)$ .

### 3.2 Witness circuits and the new easy witness lemma of [MW18]

We now recall the definition of witness circuits for a proof system.

**Definition 7** (verifiers and witnesses). *Let  $t : \mathbb{N} \rightarrow \mathbb{N}$  be a time-constructible, non-decreasing function, and let  $L \subseteq \{0, 1\}^*$ . An algorithm  $V(x, y)$  is a  $t$ -time verifier for  $L$  if  $V$  runs in time at most  $t(|x|)$  and satisfies the following: For all strings  $x$  it holds that  $x \in L$  if and only if there exists a witness  $y$  such that  $V(x, y)$  accepts.*

**Definition 8** (witness circuits). *Let  $t : \mathbb{N} \rightarrow \mathbb{N}$  be a time-constructible, non-decreasing function, let  $w : \mathbb{N} \rightarrow \mathbb{N}$ , and let  $L \subseteq \{0, 1\}^*$ . We say that a  $t$ -time verifier  $V$  has witness circuits of size  $w$  if for every  $x \in L$  there exists a witness  $y_x$  such that  $V(x, y_x)$  accepts and there exists a circuit  $C_{y_x} : \{0, 1\}^{\log(|y_x|)} \rightarrow \{0, 1\}$  of size  $w(|x|)$  such that  $C_{y_x}(i)$  is the  $i^{\text{th}}$  bit of  $y_x$ . We say that  $NTIME[t]$  has witness circuits of size  $w$  if for every  $L \in NTIME[t]$ , every  $t$ -time verifier for  $L$  has witness circuits of size  $w$ .*

In Definitions 7 and 8 we considered verifiers that are deterministic algorithms that get the witness as an explicit input. As outlined in Section 2, in the proof we will consider PCP verifiers (which are probabilistic algorithms, and only get oracle access to their witness). However, we will not consider witness circuits for these PCP verifiers, but rather for deterministic verifiers (with explicit inputs) that are derived from the PCP verifiers (see the proof of Theorem 14 for precise details).

Let us now state the new easy witness lemma of [MW18]. Loosely speaking, the lemma asserts that for any two functions  $t(n) \gg s(n)$  with sufficient “gap” between them, if  $\text{NTIME}[\text{poly}(t)] \subseteq \text{SIZE}[s]$ , then  $\text{NTIME}[t]$  has witness circuits of size  $\hat{s}$ , where  $\hat{s}(n) > s(n)$  is the function  $s$  with some “overhead”. To more conveniently account for the exact parameters, we introduce some auxiliary technical notation:

**Definition 9** (*sufficiently gapped functions*). Let  $\gamma, \gamma', \gamma'' \in \mathbb{N}$  be universal constants.<sup>10</sup> For any function  $s : \mathbb{N} \rightarrow \mathbb{N}$ , let  $s' : \mathbb{N} \rightarrow \mathbb{N}$  be the function  $s'(n) = (s(\gamma \cdot n))^\gamma$ , and let  $\hat{s} : \mathbb{N} \rightarrow \mathbb{N}$  be the function  $\hat{s}(n) = (s'(s'(s'(n))))^{\gamma'}$ . We say that two functions  $s, t : \mathbb{N} \rightarrow \mathbb{N}$  are sufficiently gapped if both functions are increasing and time-constructible, and  $s'$  is also time-constructible, and  $s(n) < 2^{n/\gamma}/n$ , and  $t(n) \geq (\hat{s}(n))^{\gamma''}$ .

**Lemma 10** (*easy witnesses for low nondeterministic time* [MW17, Lem. 4.1]). Let  $s, t : \mathbb{N} \rightarrow \mathbb{N}$  be sufficiently gapped functions, and assume that  $\text{NTIME}[O(t(n))^\gamma] \subset \text{SIZE}[s]$ , where  $\gamma$  is the constant from Definition 9. Then,  $\text{NTIME}[t]$  has witness circuits of size  $\hat{s}$ .

### 3.3 Merlin-Arthur protocols

We recall the standard definition of Merlin-Arthur protocols (i.e., MA verifiers) that receive non-uniform advice.

**Definition 11** (*MA verifiers with non-uniform advice*). For  $t, \ell : \mathbb{N} \rightarrow \mathbb{N}$ , a set  $S \subseteq \{0, 1\}^*$  is in  $\text{MATIME}[t]/\ell$  if there exists a probabilistic machine  $V$ , called a verifier, such that the following holds: The verifier  $V$  gets input  $x \in \{0, 1\}^*$ , and a witness  $w \in \{0, 1\}^*$ , and an advice string  $a \in \{0, 1\}^*$ , and runs in time  $t(|x|)$ ; and there exists a sequence  $\{a_n\}_{n \in \mathbb{N}}$  of advice such that  $|a_n| = \ell(n)$  and:

1. For every  $x \in S$  there exists  $w \in \{0, 1\}^{t(|x|)}$  such that  $\Pr[V(x, w, a_{|x|}) = 1] \geq 2/3$ .
2. For every  $x \notin S$  and every  $w \in \{0, 1\}^{t(|x|)}$  it holds that  $\Pr[V(x, w, a_{|x|}) = 1] \leq 1/3$ .

It is common to denote by  $\text{MATIME}[t]$  the class  $\text{MATIME}[t]/0$  (i.e., when the verifier receives no non-uniform advice). Note that  $\text{MA} = \bigcup_{c \in \mathbb{N}} \text{MATIME}[n^c]$ , and also note that the definition of MA does not change if we insist on perfect completeness (see, e.g., [Gol08, Exer. 6.12(2)]).

<sup>10</sup>Specifically, the values of these constants are  $\gamma = e$  and  $\gamma' = 2g$  and  $\gamma'' = d$ , where  $e, g$ , and  $d$  are the universal constants from Lemma 4.1 in [MW17].

### 3.4 A barrier for proving “ $pr\mathcal{BPP} = pr\mathcal{P} \implies \mathcal{P} \neq \mathcal{NP}$ ”

We note that it is impossible to prove the statement “if  $pr\mathcal{P} = pr\mathcal{BPP}$  then  $\mathcal{P} \neq \mathcal{NP}$ ” without *unconditionally* proving that  $\mathcal{P} \neq \mathcal{NP}$ .

**Proposition 12** (a barrier for “derandomization implies lower bounds” statements). *If the conditional statement “ $pr\mathcal{BPP} = pr\mathcal{P} \implies \mathcal{P} \neq \mathcal{NP}$ ” holds, then  $\mathcal{P} \neq \mathcal{NP}$ .*

**Proof.** Assume towards a contradiction that  $\mathcal{P} = \mathcal{NP}$ . Then, the polynomial-time hierarchy collapses to  $\mathcal{P}$ , and similarly the promise-problem version of the polynomial-time hierarchy collapses to  $pr\mathcal{P}$ .<sup>11</sup> Now, since  $pr\mathcal{BPP}$  is contained in the promise-problem version of the polynomial-time hierarchy (e.g., by adapting the well-known argument of Lautemann [Lau83]), it follows that  $pr\mathcal{BPP} = pr\mathcal{P}$ . Finally, we can use the hypothesized conditional statement to deduce that  $\mathcal{P} \neq \mathcal{NP}$ , which is a contradiction. ■

## 4 Proof of Theorems 1 and 2

We will first prove a general and parametrized “derandomization implies lower bounds” theorem. This theorem is obtained by using the proof strategy of Williams [Wil13] with general parameters, while leveraging the new easy witness lemma of Murray and Williams [MW18]. We then prove Theorems 1 and 2 as corollaries. Towards presenting the proofs, we first need the following auxiliary definition:

**Definition 13** (non-deterministically solving CAPP). *We say that  $(1, 1/3)$ -CAPP can be solved in non-deterministic time  $T : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$  if there exists a non-deterministic machine that, when given as input a circuit  $C$  of size  $m$  over  $v$  variables, runs in time  $T(m, v)$  and satisfies the following: If  $C$  has acceptance probability one, then for some non-deterministic choice the machine accepts; and if  $C$  has acceptance probability at most  $1/3$ , then the machine always rejects (regardless of the non-deterministic choices).*

### 4.1 A parametrized “derandomization implies lower bounds” theorem

Loosely speaking, in the following theorem statement we assume that CAPP can be solved in non-deterministic time  $T(m, v)$ , and deduce that for any two functions  $t(n) \gg s(n)$  such that  $T(\text{poly}(n, \hat{s}(n)), \log(t(n))) \ll t(n)$  it holds that  $NTIME[\text{poly}(t(n))]$  does not have circuits of size  $s(n)$ .

<sup>11</sup> To see that this is the case, let  $\Pi = (Y, N) \subseteq \{0, 1\}^* \times \{0, 1\}^*$  be a promise problem in  $pr\Sigma_k$ , for some  $k \in \mathbb{N}$ . Then, there exists a polynomial-time algorithm  $A$  such that for every  $x \in Y$  it holds that  $\exists y_1, \forall y_2, \dots, y_k : A(x, y_1, \dots, y_k) = 1$ , and for every  $x \in N$  it does not hold that  $\exists y_1, \forall y_2, \dots, y_k : A(x, y_1, \dots, y_k) = 1$ . We define a set  $S = S_A$  that consists of all strings  $x$  such that  $\exists y_1, \forall y_2, \dots, y_k : A(x, y_1, \dots, y_k) = 1$ . Note that  $S \supseteq Y$ , and that  $S \cap N = \emptyset$ , and that  $S \in \Sigma_k$  (using the algorithm  $A$ ). By our assumption that the polynomial-time hierarchy collapses, there exists a polynomial-time algorithm  $A'$  that decides  $S$ . It follows that  $A'$  solves the problem  $\Pi$ .

**Theorem 14** (*derandomization implies lower bounds, with general parameters*). *There exist constants  $c, c' \in \mathbb{N}$  and  $\alpha < 1$  such that the following holds. For  $T : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ , assume that  $(1, 1/3)$ -CAPP on circuits of size  $m$  with at most  $v$  input variables can be solved in non-deterministic time  $T(m, v)$ . Let  $s, t : \mathbb{N} \rightarrow \mathbb{N}$  be sufficiently gapped functions such that  $s(n) > n$  and for some constant  $\epsilon > 0$  and any constant  $\alpha > 0$  it holds that*

$$T\left((n \cdot \hat{s}(n) \cdot \log(t(n)))^c, \alpha \cdot \log(t(n))\right) \leq t(n)^{(1-\epsilon) \cdot \alpha},$$

where  $\hat{s}$  is defined as in Definition 9. Then,  $\text{NTIME}[t(n)^{c'}] \not\subseteq \text{SIZE}[s(n)]$ .

**Proof.** The starting point of the proof is the non-deterministic time hierarchy [Coo72]: For an appropriate function  $t' = t'(n)$  (that will be determined in a moment), there exists a set  $L \in \text{NTIME}[t']$  that cannot be decided by non-deterministic machines running in time  $(t')^{1-\Omega(1)}$ . Specifically, for a sufficiently small constant  $\alpha > 0$ , let  $t'(n) = (t(n))^{(1-\epsilon/2) \cdot \alpha}$ , and let  $L \in \text{NTIME}[t'] \setminus \text{NTIME}\left[(t')^{\frac{1-\epsilon}{1-\epsilon/2}}\right]$ .<sup>12</sup> Now, for a sufficiently large constant  $c'$ , assume towards a contradiction that  $\text{NTIME}[t(n)^{c'}] \subseteq \text{SIZE}[s(n)]$ . Our goal is to construct a non-deterministic machine that decides  $L$  in time  $(t')^{\frac{1-\epsilon}{1-\epsilon/2}}$ , which will yield a contradiction.

To do so, consider the PCP verifier of [BV14] for  $L$ , denoted by  $V$ . On inputs of length  $n$ , the verifier  $V$  runs in time  $\text{poly}(n, \log(t'(n)))$ , uses  $\ell = \log(t'(n)) + O(\log \log(t'(n)))$  bits of randomness, and has perfect completeness and soundness (much) lower than  $1/3$ .<sup>13</sup> Furthermore, using the hypothesis that  $\text{NTIME}[t(n)^{c'}] \subseteq \text{SIZE}[s(n)]$  and the “easy witness lemma” (i.e., Lemma 10), for every  $x \in L$  there exists a circuit  $P_x \in \text{SIZE}[\hat{s}(n)]$  such that  $\Pr_r[V^{P_x}(x, r) \text{ accepts}] = 1$ . (We actually apply Lemma 10 to the deterministic verifier  $V'$  that enumerates the random coins of  $V$ , which runs in time  $2^\ell \cdot \text{poly}(n, \log(t')) = \text{poly}(t') = \text{poly}(t)$ . We can use the lemma since we assumed that  $\text{NTIME}[t(n)^{c'}] \subseteq \text{SIZE}[s(n)]$ , for a sufficiently large  $c'$ .)

Given input  $x \in \{0, 1\}^n$ , the non-deterministic machine  $M$  acts as follows. The machine non-deterministically guesses a (description of a) circuit  $P_x$  of size  $\hat{s}(n)$ , and constructs a circuit  $C_x^{P_x} : \{0, 1\}^\ell \rightarrow \{0, 1\}$  such that  $C_x^{P_x}(r) = V^{P_x}(x, r)$ . Then, the machine feeds the description of  $C_x^{P_x}$  as input to the machine  $M_{\text{CAPP}}$  that solves CAPP in non-deterministic time  $T$  and exists by the hypothesis, and outputs the decision of  $M_{\text{CAPP}}$ . By the properties of the PCP verifier and of  $M_{\text{CAPP}}$ , if  $x \in L$  then for some guess of  $P_x$  and for some non-deterministic choices of  $M_{\text{CAPP}}$ , the machine  $M$  will accept  $x$ ; and if  $x \notin L$ , then for any guess of  $P_x$  and any non-deterministic choices of  $M_{\text{CAPP}}$ , the machine  $M$  will reject  $x$ .

To conclude let us upper-bound the running-time of the machine  $M$ . The circuit  $C_x^{P_x}$  has  $\ell = \log(t') + O(\log \log(t')) < \alpha \cdot \log(t)$  input bits, and its size is  $m(n) = \text{poly}(n, \log(t')) \cdot \hat{s}(n)$ ; thus, its representation size is  $\text{poly}(m(n))$ . Therefore, the circuit

<sup>12</sup>Such a function exists by standard non-deterministic time hierarchy theorems (e.g., [Coo72]), since  $t'(n) > n^{\Omega(1)}$ , which implies that the gap between  $t'$  and  $(t')^{1-\Omega(1)}$  is sufficiently large.

<sup>13</sup>Note that the only upper-bound that we need on the number of oracle queries issued by  $V$  is the trivial bound given by the running time of  $V$ .

$C_x^{P_x}$  can be constructed in time  $\text{poly}(m(n))$ , and the CAPP algorithm runs in time  $T(m(n), \ell)$ . The total running-time of the non-deterministic machine  $M$  is thus at most  $T((n \cdot \hat{s}(n) \cdot \log(t(n)))^c, \alpha \cdot \log(t))$ , for some constant  $c$ . By our hypothesized upper-bound on  $T$ , the running time of  $M$  is at most  $t(n)^{(1-\epsilon)\alpha} = (t')^{\frac{1-\epsilon}{1-\epsilon/2}}$ , which yields a contradiction. ■

**Additional relaxations of the hypothesis in Theorem 14.** Since the proof of Theorem 14 relies on the strategy of [Wil13], it is well-known that the hypothesis of the theorem can be further relaxed. First, we do not have to *unconditionally* assume that the non-deterministic machine for CAPP exists, and it suffices to assume that the machine exists *under the hypothesis that*  $\text{NTIME}[t(n)^{c'}] \subseteq \text{SIZE}[s(n)]$  (this is the case since we are only using the existence of the machine to contradict the latter hypothesis). And secondly, the non-deterministic machine that solves CAPP can use (sub-linearly many) bits of non-uniform advice; this follows by using a strengthened non-deterministic time hierarchy theorem, which was proved by Fortnow and Santhanam [FS16] (see [MW18, Remark 1] for details).

## 4.2 Theorems 1 and 2 as corollaries

We now prove Theorem 2 as a corollary of Theorem 14. As detailed in Section 2.1, we start from the hypothesis that  $(1, 1/3)$ -CAPP can be solved in non-deterministic time  $T(m, v) = 2^{99 \cdot v} \cdot \text{poly}(m)$  (which is weaker than the hypothesis  $\text{prBPP} = \text{prP}$ ). The proof amounts to verifying that, given such a CAPP algorithm, the hypothesis of Theorem 14 holds for essentially any  $s$  and  $t \approx s^{O(1)} \circ s^{O(1)} \circ s^{O(1)}$ .

**Corollary 15** (Theorem 2, restated). *Assume that  $(1, 1/3)$ -CAPP can be solved in non-deterministic time  $T(m, v) \leq 2^{(1-\epsilon) \cdot v} \cdot \text{poly}(m)$ , for some constant  $\epsilon > 0$ . Then, there exists a constant  $k \in \mathbb{N}$  such that for any two sufficiently gapped functions  $s : \mathbb{N} \rightarrow \mathbb{N}$  and  $t : \mathbb{N} \rightarrow \mathbb{N}$  it holds that  $\text{NTIME}[t(n)^k] \not\subseteq \text{SIZE}[s]$ .*

**Proof.** Let  $k' > 1$  be such that  $T(m, v) \leq 2^{(1-\epsilon) \cdot v} \cdot m^{k'}$ . We invoke Theorem 14 with the sufficiently gapped functions  $s$  and  $t_1(n) = t(n)^{k'}$ , where  $k'' > 1$  is a sufficiently large constant that depends on  $k'$ . Note that for any  $\alpha > 0$  it holds that

$$\begin{aligned} T\left((n \cdot \hat{s}(n) \cdot \log(t_1(n)))^c, \alpha \cdot \log(t_1(n))\right) \\ \leq (n \cdot \log(t_1(n)))^{c \cdot k'} \cdot (t_1(n))^{\epsilon/2} \cdot t_1(n)^{(1-\epsilon)\alpha} & \quad (\hat{s}(n)^{c \cdot k'} < t_1(n)^{\epsilon/2}) \\ \leq (t_1(n))^{1-\epsilon/3}, & \quad (n^{c \cdot k'} < s(n)^{c \cdot k'} < t_1(n)^{\epsilon/12}) \end{aligned}$$

where both inequalities relied on the hypothesis that  $k''$  is sufficiently large. Thus, we conclude that  $\text{NTIME}[t_2] \not\subseteq \text{SIZE}[s]$ , where  $t_2(n) = t_1(n)^{c'} = t(n)^{c' \cdot k''}$ . ■

Finally, we prove Theorem 1 as a corollary of Corollary 15. Recall that the conclusion in Theorem 1 is that  $\text{NTIME}[n^{f(n)}] \not\subseteq \mathcal{P}/\text{poly}$  for “essentially” any super-constant function  $f$ . We now specify exactly what this means. Our goal is to deduce

that  $NTIME[n^{f(n)}] \not\subseteq SIZE[n^{g(n)}]$ , where  $g(n) \ll f(n)$  and  $g(n) = \omega(1)$ . Therefore, the proof works for any  $f$  such that a suitable  $g$  exists. We note in advance that this minor technical detail imposes no meaningful restrictions on  $f$  (see next).

**Definition 16** (*admissible functions*). We say that a function  $f : \mathbb{N} \rightarrow \mathbb{N}$  is admissible if  $f$  is super-constant (i.e.  $f(n) = \omega(1)$ ), and if there exists another super-constant function  $g : \mathbb{N} \rightarrow \mathbb{N}$  that satisfies the following: The function  $g$  is super-constant, and  $t(n) = n^{f(n)}$  and  $s(n) = n^{g(n)}$  are sufficiently gapped, and  $\hat{s}(n) = n^{o(f(n))}$ .

Essentially any increasing function  $f(n) = \omega(1)$  such that  $f(n) \leq n$  is admissible, where the only additional constraints that the admissibility condition imposes are time-constructibility of various auxiliary functions (we require  $t$  and  $s$  to be sufficiently gapped, which enforces time-constructibility constraints); for a precise (and tedious) discussion, see Appendix B. We can now formally state Theorem 1 and prove it:

**Corollary 17** (*Theorem 1, restated*). Assume that  $(1, 1/3)$ -CAPP can be solved in non-deterministic time  $T(m, v) \leq 2^{(1-\epsilon) \cdot v} \cdot \text{poly}(m)$ , for some constant  $\epsilon > 0$ . Then, for every admissible function  $f$  there exists a set in  $NTIME[n^{O(f(n))}] \setminus \mathcal{P}/\text{poly}$ .

**Proof.** Since  $f$  is admissible, there exists a function  $g$  that satisfies the requirements of Definition 16. We thus invoke Corollary 15 with the functions  $t(n) = n^{f(n)}$  and  $s(n) = n^{g(n)}$ , and conclude that there exists a set in  $NTIME[n^{O(f(n))}] \setminus SIZE[n^{g(n)}]$ . Since  $g(n) = \omega(1)$ , the latter set does not belong to  $\mathcal{P}/\text{poly}$ . ■

## 5 Proof of Theorem 3

In this section we prove Theorem 3. Throughout the section, for a set  $S \subseteq \{0, 1\}^*$  and  $n \in \mathbb{N}$ , we denote  $S_n = S \cap \{0, 1\}^n$ .

Recall that the conclusion in Theorem 3 is that there exists a set  $S$  such that for every polynomial-sized circuit family and sufficiently large  $n \in \mathbb{N}$ , the family fails to decide  $S$  on some input length in the interval  $[n, s_I(n)]$ . Our actual conclusion will be slightly stronger: We will conclude that for every sufficiently large  $n \in \mathbb{N}$ , the circuit family fails on at least one of the “end-points” of the interval; that is, either on input length  $n$ , or on input length  $s_I(n)$  (or on both). This is equivalent to saying that there does not exist an infinite set of pairs  $(n, s_I(n))$  such that the circuit family correctly decides  $S$  on both input lengths in the pair. This leads to the following definition:

**Definition 18** (*a stronger notion of infinitely-often computation*). For  $s, q : \mathbb{N} \rightarrow \mathbb{N}$  and  $S \subseteq \{0, 1\}^*$ , we say that  $S \in \text{i.o.}_{[q]}\text{-SIZE}[s]$  if there exists an infinite set  $I \subseteq \mathbb{N}$  and a circuit family  $\{C_n\}_{n \in \mathbb{N}}$  of size at most  $s$  such that for every  $n \in I$ , it holds that:

1. The circuit  $C_n : \{0, 1\}^n \rightarrow \{0, 1\}$  computes  $S_n$ .
2. The circuit  $C_{q(n)} : \{0, 1\}^{q(n)} \rightarrow \{0, 1\}$  computes  $S_{q(n)}$ .

Indeed, Definition 18 implies the following: If  $S \notin \text{i.o.}_{[q]}\text{-SIZE}[s]$ , then every circuit family  $\{C_n\}$  of size  $s$  that tries to decide  $S$  fails, for every sufficiently large  $n \in \mathbb{N}$ , either on inputs on size  $n$  or on inputs of size  $q(n)$  (or on both).

The starting point of the proof of Theorem 3 is Murray and Williams' [MW17, Thm 3.1] strengthening of Santhanam's [San09] circuit lower bound. Following [MW18], we say that a function  $s : \mathbb{N} \rightarrow \mathbb{N}$  is a circuit-size function if  $s$  is increasing, time-constructible, and for all sufficiently large  $n \in \mathbb{N}$  it holds that  $s(n) < 2^n / (2n)$ .

**Theorem 19** (Murray and Williams' [MW17, Thm 3.1] strengthening of Santhanam's [San09] lower bound). *Let  $s$  be a super-linear circuit-size function, and let  $t = \text{poly}(s(\text{poly}(s)))$  (for sufficiently large polynomials that do not depend on  $s$ ). Then, there exists a set  $S \in \text{MATIME}[t]/O(\log(s))$  such that  $S \notin \text{i.o.}_{[\text{poly}(s)]}\text{-SIZE}[s]$ .*

As mentioned in Section 2.2, the first observation in the proof is that if  $\text{prBPP} = \text{prP}$  then we can derandomize MA verifiers that receive non-uniform advice. In fact, we can do so also under the weaker hypothesis that  $\text{prBPP} \subseteq \text{prNP}$ .

**Proposition 20** (derandomization of MA with advice). *If  $\text{prBPP} \subseteq \text{prNP}$ , then for any  $t, \ell : \mathbb{N} \rightarrow \mathbb{N}$  such that  $t$  is time-constructible it holds that  $\text{MATIME}[t]/\ell \subseteq \text{NTIME}[\text{poly}(t)]/\ell$ .*

**Proof.** Since CAPP is in  $\text{prBPP}$ , and we assume that  $\text{prBPP} \subseteq \text{prNP}$ , there exists a non-deterministic polynomial-time machine  $M_{\text{CAPP}}$  that gets as input a Boolean circuit  $C$  and satisfies the following: If the acceptance probability of  $C$  is at most  $1/3$  then  $M_{\text{CAPP}}$  rejects, regardless of the non-deterministic choices; and if the acceptance probability of  $C$  is at least  $2/3$ , then for some non-deterministic choices  $M_{\text{CAPP}}$  accepts.

Let  $S$  be a set in  $\text{MATIME}[t]/\ell$ , let  $V$  be the  $\text{MATIME}[t]/\ell$  verifier for  $S$ , and let  $\{a_n\}_{n \in \mathbb{N}}$  be a sequence of "good" advice that allows  $V$  to decide  $S$ . We want to construct a non-deterministic machine  $M$  that runs in time  $\text{poly}(t)$  and decides  $S$  with  $\ell$  bits of non-uniform advice. Given input  $x \in \{0, 1\}^n$  and advice  $a_n$ , the machine  $M$  guesses a witness  $w \in \{0, 1\}^{t(n)}$ , and constructs a circuit  $C = C_{V,x,w,a_n} : \{0, 1\}^{t(n)} \rightarrow \{0, 1\}$  that gets as input  $r \in \{0, 1\}^{t(n)}$  and computes  $V(x, w, r, a_n)$ . Finally, the machine  $M$  feeds  $C$  to the machine  $M_{\text{CAPP}}$ , and outputs the decision of  $M_{\text{CAPP}}$ .

Since  $V$  is a verifier for  $S$  and  $a_n$  is the "good" advice for  $V$ , if  $x \in S$  then there exists  $w$  such that the acceptance probability of  $C$  is at least  $2/3$ , which means that there exist non-deterministic choices for  $M_{\text{CAPP}}$  such that  $M_{\text{CAPP}}$  will accept  $C$  (in which case  $M$  will accept  $x$ ). On the other hand, if  $x \notin S$  then for any  $w$  the acceptance probability of  $C$  is at most  $1/3$ , which means that for any non-deterministic choices for  $M_{\text{CAPP}}$  it holds that  $M_{\text{CAPP}}$  rejects  $C$  (in which case  $M$  rejects  $x$ ). The running time of the machine  $M$  is dominated by the running time of  $M_{\text{CAPP}}$ , which is at most  $\text{poly}(t(n))$ . ■

The second observation in the proof is that if  $\text{NTIME}[t]/\ell$  is not contained in a non-uniform class of circuits, then  $\text{NTIME}[O(t)]$  (i.e., without non-uniform advice) is also not contained in a (related) non-uniform class of circuits. Moreover, this assertion still holds if the "separation" between the classes is in the sense of Definition 18.

We first prove a simpler form of this statement, which showcases the main idea but is much less cumbersome. In the following statement, we only consider a single bit of advice, and do not refer to separations in the sense of Definition 18.

**Proposition 21** (*eliminating the advice*). *Let  $s_0, s, t : \mathbb{N} \rightarrow \mathbb{N}$  such that  $t$  is increasing, and for all sufficiently large  $n \in \mathbb{N}$  it holds that  $s_0(n) \geq s(n+1)$ . If  $NTIME[t]/1 \not\subseteq SIZE[s_0]$ , then  $NTIME[O(t)] \not\subseteq SIZE[s]$ .*

**Proof.** We prove the contrapositive statement: If  $NTIME[O(t)] \subseteq SIZE[s]$ , then  $NTIME[t]/1 \subseteq SIZE[s_0]$ . To do so, fix any  $S \in NTIME[t]/1$ , and let us construct a circuit family of size  $s_0$  that decides  $S$ .

To construct the circuit family we consider an auxiliary set  $S^{\text{adv}}$ , which is defined as follows. Let  $M$  be a  $t$ -time non-deterministic machine and let  $\{a_n\}$  be a sequence of advice bits such that  $M$  correctly decides  $S$  when given advice  $\{a_n\}$ . Let  $S^{\text{adv}}$  be the set of pairs  $(x, \sigma)$ , where  $x \in \{0, 1\}^*$  and  $\sigma \in \{0, 1\}$ , such that  $M$  (non-deterministically) accepts  $x$  when given advice  $\sigma$ . Note that  $S^{\text{adv}} \in NTIME[O(t)]$ , because a non-deterministic machine that gets input  $(x, \sigma)$  simulate the machine  $M$  on input  $x$  with advice  $\sigma$  and decide according to the output of  $M$ .

Relying on the hypothesis that  $NTIME[O(t)] \subseteq SIZE[s]$ , there exists a circuit family  $\{C_n\}$  of size  $s$  such that each  $C_n$  decides  $S_n^{\text{adv}}$ . By hard-wiring the “correct” advice bit  $a_n$  in place of the last input bit into every  $C_n$ , we obtain a circuit family  $\{C'_n\}$  such that each  $C'_n$  decides  $S_n$ , and its size is at most  $s(n+1) \leq s_0(n)$ . ■

The following proposition is a stronger form of Proposition 21, which considers possibly long advice strings, and refers to separations in the sense of Definition 18.

**Proposition 22** (*eliminating the advice*). *Let  $s_0, s, \ell, t, q : \mathbb{N} \rightarrow \mathbb{N}$  be functions such that  $t$  is super-linear and increasing, and  $q, s_0$  and  $s$  are increasing, and the mapping  $1^n \mapsto 1^{\ell(n)}$  is computable in time  $O(n + \ell(n))$ . Assume that for every sufficiently large  $n \in \mathbb{N}$  it holds that  $\ell(n) < n/2$  and  $s_0(n) \geq s(2n)$  and  $s_0(q(n)) \geq s(2q(2n))$ . Further assume that  $NTIME[t]/\ell \not\subseteq \text{i.o.}_{[q]}\text{-SIZE}[s_0]$ . Then,  $NTIME[O(t)] \not\subseteq \text{i.o.}_{[2q]}\text{-SIZE}[s]$ .*

We comment that a statement that is more general than the one in Proposition 22 can be proved, foregoing some of the requirements (e.g., on  $\ell$ ) while allowing potential degradation in the parameters of the conclusion. Since the statement of Proposition 22 suffices for our parameter setting, and for simplicity, we avoid such generalizations.

**Proof of Proposition 22.** Assuming that  $NTIME[O(t)] \subseteq \text{i.o.}_{[2q]}\text{-SIZE}[s]$ , we prove that  $NTIME[t]/\ell \subseteq \text{i.o.}_{[q]}\text{-SIZE}[s_0]$ . Fixing any  $S \in NTIME[t]/\ell$ , let us construct a circuit family of size  $s_0$  that decides  $S$  infinitely-often on inputs of length  $n$  and  $q(n)$ .

We first define a set  $S^{\text{adv}}$  as follows. Let  $M$  be a  $t$ -time non-deterministic machine and let  $\{a_n\}$  be a sequence of “good” advice strings of length  $|a_n| = \ell(n)$  such that  $M$  correctly decides  $S$  when given advice  $\{a_n\}$ . For every  $n \in \mathbb{N}$ , the set  $S_n^{\text{adv}}$  will include representations of all pairs  $(x, \sigma)$ , where  $|\sigma| = \ell(|x|)$  and  $|x| + 2|\sigma| < n$ , such that  $M$  accepts  $x$  when given advice  $\sigma$ . Specifically, we define  $S_n^{\text{adv}}$  to be the set of all  $n$ -bit

strings of the form  $1^t 0^{|\sigma|} 1x\sigma$ , where  $t = n - (|x| + 2|\sigma| + 1)$ , such that  $M$  accepts  $x$  when given advice  $\sigma$ .<sup>14</sup>

Note that  $S^{\text{adv}} \in \text{NTIME}[O(t)]$ . This is the case since a non-deterministic machine that gets input  $z \in \{0,1\}^n$  can first verify that  $z$  can be parsed as  $z = 1^t 0^{|\sigma|} 1x\sigma$  such that  $|\sigma| = \ell(|x|)$  (and reject  $z$  if the parsing fails); and then simulate the machine  $M$  on input  $x$  with advice  $\sigma$ , in time  $O(t(|x|)) = O(t(n))$ , and decide according to the output of  $M$ . Now, since we assume that  $\text{NTIME}[O(t)] \subseteq \text{i.o.}_{[2q]}\text{-SIZE}[s]$ , there exists an infinite set  $I \subseteq \mathbb{N}$  and a circuit family  $\{C_n\}$  of size  $s$  such that for every  $n \in I$ :

1.  $C_n : \{0,1\}^n \rightarrow \{0,1\}$  correctly computes  $S_n^{\text{adv}}$ ; and
2.  $C_{2q(n)} : \{0,1\}^{2q(n)} \rightarrow \{0,1\}$  correctly computes  $S_{2q(n)}^{\text{adv}}$ .

We transform  $\{C_n\}$  into a circuit family of size  $s_0$  that decides  $S$  infinitely-often on inputs of length both  $n$  and  $q(n)$ . To do so, we rely on the following simple claim:

**Claim 22.1.** *Let  $n, m \in \mathbb{N}$  such that  $m + 2\ell(m) < n$ . Assume that there exists a circuit of size  $s(n)$  that decides  $S_n^{\text{adv}}$ . Then, there exists a circuit of size  $s(n)$  that decides  $S_m$ .*

*Proof.* Let  $C_n$  be the circuit of size  $s(n)$  for  $S_n^{\text{adv}}$ . The circuit  $C_m$  for  $S_m$  is obtained by hard-wiring into  $C_n$  the “correct” advice  $a_m$  instead of the last  $\ell(m)$  input bits, and the correct initial padding  $1^{n-m-2\ell(m)-1}0^{\ell(m)}1$  instead of the first  $n - m - \ell(m)$  input bits.  $\square$

For every  $n \in I$ , let  $m = m(n)$  be the largest integer such that  $m + 2\ell(m) + 1 \leq n$ . Let  $I' = \{m(n)\}_{n \in \mathbb{N}}$ , and note that  $I'$  is infinite. For every sufficiently large  $m \in I'$ , relying on the fact that  $m = m(n)$  for some  $n \in I$  and on Claim 22.1, we have that:

1. There exists a circuit  $C_m : \{0,1\}^m \rightarrow \{0,1\}$  of size  $s(n) \leq s_0(\lceil n/2 \rceil) \leq s_0(m)$  that decides  $S_m$ . (We relied on the fact that  $m \geq n/2$ , since  $\ell(m) < m/2$ .)
2. There exists a circuit  $C_{q(m)} : \{0,1\}^{q(m)} \rightarrow \{0,1\}$  of size  $s_0(q(m))$  that decides  $S_{q(m)}$ . To see this, recall that there exists a circuit  $C_{2q(n)}$  of size  $s(2q(n))$  that decides  $S_{2q(n)}^{\text{adv}}$ . We can invoke Claim 22.1 because  $q(m) + 2\ell(q(m)) < 2q(m) < 2q(n)$ . Also, relying on the fact that  $m \geq n/2$  and on the hypotheses regarding  $s_0$ ,  $s$  and  $q$ , we have that  $s(2q(n)) \leq s(2q(2m)) \leq s_0(q(m))$ .

It follows that  $S \in \text{i.o.}_{[q]}\text{-SIZE}[s_0]$ .  $\blacksquare$

We now combine the foregoing ingredients into a proof of Theorem 3. The structure of the theorem that we obtain is similar to the structure of Theorem 14: Assuming a sufficiently strong derandomization hypothesis (in this case, that  $\text{prBPP} = \text{prP}$ ), and taking two functions  $t$  and  $s$  with sufficient “gap” between them, we deduce that  $\text{NTIME}[t] \not\subseteq \text{i.o.}_{[s_0]}\text{-SIZE}[s]$ , where  $s_0$  is a function moderately larger than  $s$ . (The fact that  $s_0 > s$  is no coincidence; see discussion after the proof of Theorem 23.)

<sup>14</sup>The  $0^{|\sigma|}$  term facilitates the parsing of the suffix of the  $n$ -bit string as a pair  $x\sigma$ .

**Theorem 23** (Theorem 3, restated). *There exists a constant  $\epsilon > 0$  such that the following holds.*

- *Let  $s : \mathbb{N} \rightarrow \mathbb{N}$  be an increasing, super-linear and time-constructible function such that for all sufficiently large  $n \in \mathbb{N}$  it holds that  $s(n) \leq 2^{\epsilon \cdot n}$  and that  $s(2n) \leq s(n)^2$ .*
- *Let  $t = \text{poly}(s(\text{poly}(s)))$ , for sufficiently large polynomials (that do not depend on  $s$ ).*
- *Let  $s_0 : \mathbb{N} \rightarrow \mathbb{N}$  be an increasing and time-constructible function such that for all sufficiently large  $n \in \mathbb{N}$  it holds that  $s_0(n) \geq s(n^2)^2$  and that  $s_0(2n) \leq s_0(n)^2$ .*

*Assume that  $\text{prBPP} \subseteq \text{prNP}$ . Then,  $\text{NTIME}[t] \not\subseteq \text{i.o.}_{[\text{poly}(s_0)]}\text{-SIZE}[s]$ .*

Note that if the function  $s$  in Theorem 23 satisfies  $s(n^2) < s(n)^k$ , for a sufficiently large constant  $k \in \mathbb{N}$ , then we can use the function  $s_0(n) = s(n)^{2k}$ , and deduce that  $\text{NTIME}[t] \not\subseteq \text{i.o.}_{[\text{poly}(s)]}\text{-SIZE}[s]$ .

**Proof of Theorem 23.** Let  $t_0 = \text{poly}(s_0(\text{poly}(s_0)))$ , for sufficiently large polynomials, and let  $\ell = O(\log(s_0))$  (the universal constant hidden in the  $O$ -notation is the one from Theorem 19). By Theorem 19, there exists  $S \in \text{MATIME}[t_0]/\ell \setminus \text{i.o.}_{[(s_0)^c]}\text{-SIZE}[s_0]$ , for a sufficiently large constant  $c \in \mathbb{N}$ . By Proposition 20, and relying on the hypothesis that  $\text{prBPP} \subseteq \text{prNP}$ , it holds that  $S \in \text{NTIME}[\text{poly}(t_0)]/\ell \setminus \text{i.o.}_{[(s_0)^c]}\text{-SIZE}[s_0]$ .

We now want to use Proposition 22 to deduce that  $\text{NTIME}[\text{poly}(t_0)]$  is not contained in  $\text{i.o.}_{[\text{poly}(s_0)]}\text{-SIZE}[s]$ , and thus we need to verify that the functions  $\ell$ ,  $s$ ,  $s_0$ , and  $(s_0)^c$  satisfy the hypothesis of Proposition 22. This is indeed the case since for all sufficiently large  $n \in \mathbb{N}$  it holds that  $\ell(n) < n/2$  (assuming that  $\epsilon$  is sufficiently small); and since  $s_0(n) > s(n^2)^2 \geq s(2n)$ , and  $s(2s_0(2n)^c) \leq s(s_0(n)^{2c})^2 \leq s_0(s_0(n)^c)$ . ■

As mentioned before the statement of Theorem 23, the “gap” between the input lengths  $n$  and  $q(n) = \text{poly}(s_0(n))$  (on which any size- $s$  circuit family is guaranteed to fail) in Theorem 23 is larger than the function  $s$  that bounds the size of the circuits. This is no coincidence: If the gap function  $q$  would have been *significantly smaller* than the bound  $s$  on the circuit size, then we would have obtained an “almost-everywhere” lower bound (for circuits of size about  $s(q^{-1})$ ).<sup>15</sup>

## Acknowledgements

The author thanks his advisor, Oded Goldreich, for his close guidance in the research and writing process, and for very useful comments on drafts of the paper. The author thanks Ryan Williams for several very helpful email exchanges and conversations, and

<sup>15</sup>To see this, assume that  $S \notin \text{i.o.}_{[q]}\text{-SIZE}[s]$ , for  $q \ll s$ . We define a set  $S^{\text{emb}}$  by “embedding” all strings in  $S$  of length  $n-1$  and  $q^{-1}(n-1)$  into  $\{0,1\}^n$ : For each  $n \in \mathbb{N}$ , let  $S_n^{\text{emb}}$  consist of all  $n$ -bit strings  $0^{n-|x|}1x$  such that  $x \in S$ . Since  $S \notin \text{i.o.}_{[q]}\text{-SIZE}[s]$ , for every sufficiently large  $n \in \mathbb{N}$  the circuit complexity of  $S_n^{\text{emb}}$  is larger either than  $s(n-1)$  or than  $s(q^{-1}(n-1))$ . In natural cases where  $s(q^{-1}(n-1)) < s(n-1)$ , we obtain an “almost-everywhere” lower bound for circuits of size about  $s(q^{-1})$ .

in particular for stressing the point that the conclusion in the main theorem can be interpreted as a weak form of a “strengthened time-hierarchy” theorem (that contrasts circuits and uniform algorithms), and for suggesting to try and prove an “almost-everywhere” lower bound as a consequence of  $pr\mathcal{BPP} = pr\mathcal{P}$ . The author is very grateful to Igor Oliveira for proposing the alternative proof of Theorem 1 (i.e., the one presented in Appendix A), and for his permission to include the alternative proof in this paper. The author is partially supported by Irit Dinur’s ERC-CoG grant 772839.

## References

- [AB09] Sanjeev Arora and Boaz Barak. *Computational complexity: A modern approach*. Cambridge University Press, Cambridge, 2009.
- [BFN+93] László Babai, Lance Fortnow, Noam Nisan, and Avi Wigderson. “BPP has subexponential time simulations unless EXPTIME has publishable proofs”. In: *Computational Complexity* 3.4 (1993), pp. 307–318.
- [BFT98] Harry Buhrman, Lance Fortnow, and Thomas Thierauf. “Nonrelativizing separations”. In: *Proc. 13th Annual IEEE Conference on Computational Complexity (CCC)*. 1998, pp. 8–12.
- [BGH+05] Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil Vadhan. “Short PCPs Verifiable in Polylogarithmic Time”. In: *Proc. 20th Annual IEEE Conference on Computational Complexity (CCC)*. 2005, pp. 120–134.
- [BM84] Manuel Blum and Silvio Micali. “How to Generate Cryptographically Strong Sequences of Pseudo-random Bits”. In: *SIAM Journal of Computing* 13.4 (1984), pp. 850–864.
- [BV14] Eli Ben-Sasson and Emanuele Viola. “Short PCPs with projection queries”. In: *Proc. 41st International Colloquium on Automata, Languages and Programming (ICALP)*. 2014, pp. 163–173.
- [Coo72] Stephen A. Cook. “A Hierarchy for Nondeterministic Time Complexity”. In: *Proc. 4th Annual ACM Symposium on Theory of Computing (STOC)*. 1972, pp. 187–192.
- [FS16] Lance Fortnow and Rahul Santhanam. “New non-uniform lower bounds for uniform classes”. In: *Proc. 31st Annual IEEE Conference on Computational Complexity (CCC)*. 2016, Art. No. 19, 14.
- [Gol08] Oded Goldreich. *Computational Complexity: A Conceptual Perspective*. New York, NY, USA: Cambridge University Press, 2008.
- [IKW02] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. “In search of an easy witness: exponential time vs. probabilistic polynomial time”. In: *Journal of Computer and System Sciences* 65.4 (2002), pp. 672–694.

- [IW99] Russell Impagliazzo and Avi Wigderson. “P = BPP if E requires exponential circuits: derandomizing the XOR lemma”. In: *Proc. 29th Annual ACM Symposium on Theory of Computing (STOC)*. 1999, pp. 220–229.
- [Juk12] Stasys Jukna. *Boolean function complexity*. Springer, Heidelberg, 2012.
- [KMS12] Jeff Kinne, Dieter van Melkebeek, and Ronen Shaltiel. “Pseudorandom generators, typically-correct derandomization, and circuit lower bounds”. In: *Computational Complexity* 21.1 (2012), pp. 3–61.
- [Lau83] Clemens Lautemann. “BPP and the polynomial hierarchy”. In: *Information Processing Letters* 17.4 (1983), pp. 215–217.
- [MW17] Cody Murray and Ryan Williams. “Circuit Lower Bounds for Nondeterministic Quasi-Polytime: An Easy Witness Lemma for NP and NQP”. In: *Electronic Colloquium on Computational Complexity: ECCC 24* (2017), p. 188.
- [MW18] Cody Murray and Ryan Williams. “Circuit Lower Bounds for Nondeterministic Quasi-Polytime: An Easy Witness Lemma for NP and NQP”. In: *Proc. 50th Annual ACM Symposium on Theory of Computing (STOC)*. 2018.
- [NW94] Noam Nisan and Avi Wigderson. “Hardness vs. randomness”. In: *Journal of Computer and System Sciences* 49.2 (1994), pp. 149–167.
- [San09] Rahul Santhanam. “Circuit lower bounds for Merlin-Arthur classes”. In: *SIAM Journal of Computing* 39.3 (2009), pp. 1038–1061.
- [TV07] Luca Trevisan and Salil P. Vadhan. “Pseudorandomness and Average-Case Complexity Via Uniform Reductions”. In: *Computational Complexity* 16.4 (2007), pp. 331–364.
- [Vad12] Salil P. Vadhan. *Pseudorandomness*. Foundations and Trends in Theoretical Computer Science. Now Publishers, 2012.
- [Wil13] Ryan Williams. “Improving Exhaustive Search Implies Superpolynomial Lower Bounds”. In: *SIAM Journal of Computing* 42.3 (2013), pp. 1218–1244.
- [Yao82] Andrew C. Yao. “Theory and Application of Trapdoor Functions”. In: *Proc. 23rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 1982, pp. 80–91.

## Appendix A An alternative proof of Theorem 2

In this section we present an alternative proof of Theorem 1, which does not rely on the work of Murray and Williams [MW18], but rather on the work of Santhanam [San09]. The idea for this alternative proof was suggested to us by Igor Oliveira (after a preliminary version of this paper appeared online).

The structure of this alternative proof is very similar to the proof of Theorem 3 (which was described in Section 2.2), but uses as a starting point a generalization of the circuit lower bound proved by Santhanam [San09], instead of its subsequent

strengthening by Murray and Williams [MW18]. Specifically, the starting point of the proof is the following:

**Theorem 24** (a generalization of [San09, Thm. 1]). *Let  $s : \mathbb{N} \rightarrow \mathbb{N}$  be an increasing, super-linear and time-computable function such that for all sufficiently large  $n \in \mathbb{N}$  it holds that  $s(3n) \leq s(n)^3$ . Then, for  $t : \mathbb{N} \rightarrow \mathbb{N}$  such that  $t(n) = \text{poly}(s(\text{poly}(s(n))))$  it holds that  $\text{MATIME}[t]/1 \not\subseteq \text{SIZE}[s]$ .*

The proof of Theorem 24 imitates the original argument from [San09], but uses more general parameters. We include the full proof for completeness, but since it requires no new significant ideas, we defer its presentation to the end of the appendix. The alternative proof of Theorem 1 follows by combining Theorem 24, Proposition 20 (instantiated with the value  $\ell = 1$ ), and Proposition 21.

**Theorem 25** (Theorem 1, an alternative technical statement). *Let  $s : \mathbb{N} \rightarrow \mathbb{N}$  be an increasing, super-linear and time-computable function such that for all sufficiently large  $n \in \mathbb{N}$  it holds that  $s(3n) \leq s(n)^3$ , and let  $t : \mathbb{N} \rightarrow \mathbb{N}$  such that  $t(n) = \text{poly}(s(\text{poly}(s(n))))$ , for sufficiently large polynomials. Assume that  $\text{prBPP} = \text{prP}$ . Then,  $\text{NTIME}[t] \not\subseteq \text{SIZE}[s]$ .*

**Proof.** Let  $s_0 = s^3$ , and let  $t_0 = \text{poly}(s_0(\text{poly}(s_0)))$ , for sufficiently large polynomials. According to Theorem 24, there exists a set  $S$  in  $\text{MATIME}[t_0]/1$  such that  $S \notin \text{SIZE}[s_0]$ . By Proposition 20, and relying on the hypothesis that  $\text{prBPP} = \text{prP}$ , it holds that  $S \in \text{NTIME}[t_1]/1 \setminus \text{SIZE}[s_0]$ , where  $t_1 = \text{poly}(t_0)$ . Using Proposition 21, it holds that  $\text{NTIME}[t] \not\subseteq \text{SIZE}[s_1]$ , where  $t = O(t_1) = \text{poly}(s(\text{poly}(s)))$  and  $s_1(n) = s_0(n-1)$ . Finally, since  $s$  is increasing and  $s(n) \leq s(\lceil n/3 \rceil)^3$ , we have that  $s_1(n) = s_0(n-1) \geq s_0(\lceil n/3 \rceil) \geq s(n)$ , and hence  $\text{NTIME}[t] \not\subseteq \text{SIZE}[s]$ . ■

It is just left to detail the proof of Theorem 24. The first technical ingredient in the proof is the  $\mathcal{PSPACE}$ -complete set of Trevisan and Vadhan [TV07]. We use this set, but instead of relying on the fact that the set is  $\mathcal{PSPACE}$ -complete, we will use padding to claim that the set is complete for  $\text{DSPACE}[n^{\omega(1)}]$  under  $n^{\omega(1)}$ -time reductions.

**Lemma 26** (scaling the  $\mathcal{PSPACE}$ -complete set of [TV07]). *There exists a set  $L^{\text{TV}} \subseteq \{0,1\}^*$  and a probabilistic polynomial-time oracle Turing machine  $M$  that satisfy the following:*

1. *Let  $t : \mathbb{N} \rightarrow \mathbb{N}$  be a super-linear, time-computable function. Then, for every set  $L \in \text{DSPACE}[t]$  there exists a deterministic Turing machine  $R_L$  that runs in time  $\text{poly}(t)$  such that for every  $x \in \{0,1\}^*$  it holds that  $x \in L \iff R_L(x) \in L^{\text{TV}}$ .*
2. *On input  $x \in \{0,1\}^*$ , the machine  $M$  only issues queries of length  $|x|$ .*
3. *For any  $x \in L^{\text{TV}}$  it holds that  $\Pr[M^{\mathbf{1}_{L^{\text{TV}}}}(x) = 1] = 1$ , where  $\mathbf{1}_{L^{\text{TV}}} : \{0,1\}^n \rightarrow \{0,1\}$  is the indicator function of  $L^{\text{TV}} \cap \{0,1\}^n$ .*
4. *For any  $x \notin L^{\text{TV}}$  and any  $f : \{0,1\}^n \rightarrow \{0,1\}$  it holds that  $\Pr[M^f(x) = 0] \geq 2/3$ .*

**Proof.** We take  $L^{\text{TV}}$  to be the  $\mathcal{PSPACE}$ -complete set from [San09, Lem. 12], which is the same set constructed in [TV07]. Items (2) – (4) follow immediately from the original statement in [San09].<sup>16</sup> Item (1) follows since  $L^{\text{TV}}$  is  $\mathcal{PSPACE}$ -complete, and using a padding argument. Specifically, for any  $t$  and  $L$ , consider the machine  $R_L$  that combines a reduction of  $L$  to  $L' = \{(x, 1^t) : x \in L\}$  with a reduction of  $L'$  to  $L^{\text{TV}}$ . The first reduction maps  $x \mapsto (x, 1^t)$ , and since  $L' \in \mathcal{PSPACE}$ , there exists a second reduction of  $L'$  to  $L^{\text{TV}}$  that can be computed in time  $\text{poly}(t + |x|) < \text{poly}(t)$  (the inequality is since  $t$  is super-linear). ■

**Proof of Theorem 24.** Let  $t_0 : \mathbb{N} \rightarrow \mathbb{N}$  such that  $t_0(n) = s^4(n)$ , and let  $t_1 = \text{poly}(t_0)$  and  $t = t_2 = \text{poly}(t_0(\text{poly}(t_0)))$ , for sufficiently large polynomials. Let  $L^{\text{TV}}$  be the set from Lemma 26. Our goal is to prove that there exists a set in  $\text{MATIME}[t_2]/1$  that is not in  $\text{SIZE}[t_0^{1/4}]$ . The proof proceeds by a case analysis.

**Case 1:**  $L^{\text{TV}} \in \text{SIZE}[t_0]$ . By a standard diagonalization argument, there exists a set  $L^{\text{diag}} \in \text{DSPACE}[t_1] \setminus \text{SIZE}[t_0]$ .<sup>17</sup> Our main goal now will be to prove that  $\text{DSPACE}[t_1] \subseteq \text{MATIME}[t_2]$ , which will imply that  $L^{\text{diag}} \in \text{MATIME}[t_2] \setminus \text{SIZE}[t_0]$ . (Indeed, in this case we are proving a stronger result, since the  $MA$  verifiers do not need advice, and since the circuits are of size  $t_0$  rather than  $s = t_0^{1/4}$ .)

To do so, let  $L \in \text{DSPACE}[t_1]$ , and consider the following  $MA$  verifier for  $L$ . On input  $x \in \{0, 1\}^n$ , the verifier computes  $x' = R_L(x)$ , where  $R_L$  is the machine from Lemma 26. Note that  $n' = |x'| \leq \text{poly}(t_1(n))$ , and that  $x \in L \iff x' \in L^{\text{TV}}$ . Now, the verifier parses the witness  $w \in \{0, 1\}^{\text{poly}(t_0(n'))}$  as a description of a circuit  $C : \{0, 1\}^{n'} \rightarrow \{0, 1\}$  of size  $t_0(n')$ , and runs the machine  $M$  from Lemma 26 on input  $x'$ , while answering each oracle query of  $M$  using the circuit  $C$ .

Note that, since  $L^{\text{TV}} \in \text{SIZE}[t_0]$ , there exists a circuit  $C$  over  $n'$  input bits of size  $t_0(n')$  that correctly computes  $L^{\text{TV}}$  on inputs of length  $n'$ . Therefore, by Lemma 26, when  $x \in L$  there exists a witness such that the verifier accepts  $x$  with probability one, whereas the verifier rejects any  $x \notin L$  with probability at least  $2/3$ , regardless of the witness. The total running time of the verifier is dominated by the time it takes to simulate  $M$  using the circuit  $C$ , which is at most  $\text{poly}(n') \cdot \text{poly}(t_0(n')) \leq t_2(n)$ .

**Case 2:**  $L^{\text{TV}} \notin \text{SIZE}[t_0]$ . In this case we show an explicit set  $L^{\text{pad}}$ , which will be a padded version of  $L^{\text{TV}}$ , such that  $L^{\text{pad}}$  can be decided in  $\text{MATIME}[t_2]$  with one bit of advice, but cannot be decided by circuits of size  $s = t_0^{1/4}$ . To do so, let  $\text{sz}_{\text{TV}} : \mathbb{N} \rightarrow \mathbb{N}$  be such that  $\text{sz}_{\text{TV}}(n)$  is the minimum circuit size for  $L_n^{\text{TV}} = L^{\text{TV}} \cap \{0, 1\}^n$ . Also, for any integer  $m$ , let  $p(m) = 2^{\lfloor \log(m) \rfloor}$  be the largest power of two that is not larger than  $m$ ,

<sup>16</sup>The original statement asserts that any  $x \notin L^{\text{TV}}$  is rejected with probability at least  $1/2$  (rather than  $2/3$ ), but this probability can be amplified to  $2/3$  using standard error-reduction.

<sup>17</sup>For example,  $L^{\text{diag}} = \{x : C_{|x|}(x) = 1\}$ , where  $C_n$  is the lexicographically-first circuit over  $n$  bits of size at most  $t_0^2(n)$  that decides a set whose circuit complexity is more than  $t_0(n)$ . The proof that  $L^{\text{diag}} \in \text{DSPACE}[t_1]$  follows the well-known idea used in Kannan's theorem (see, e.g., [Juk12, Lem. 20.12]).

and let  $n(m) = m - p(m)$ . We think of  $n(m)$  as the “effective input length” indicated by  $m$ , and on  $p(m)$  as the length of padding. We define the set  $L^{\text{pad}}$  as follows:

$$L^{\text{pad}} = \left\{ (x, 1^p) : x \in L^{\text{TV}}, \text{ and } |x| = n(|x| + p), \right. \\ \left. \text{and } t_0(|x| + p) \leq \text{sz}_{\text{TV}}(|x|)^3 < t_0(|x| + 2p) \right\}.$$

Let us first see that  $L^{\text{pad}}$  cannot be decided by circuits of size  $t_0^{1/4}$ . Assume towards a contradiction that there exists a circuit family  $\{C_m\}$  of size  $t_0^{1/4}$  that decides  $L_m^{\text{pad}}$  correctly for every  $m$ . Since  $L^{\text{TV}} \notin \text{SIZE}[t_0]$ , there exists an infinite set  $I \subseteq \mathbb{N}$  such that for every  $n \in I$  it holds that  $\text{sz}_{\text{TV}}(n) > t_0(n)$ . For a sufficiently large  $n \in I$ , we will construct a circuit  $C'_n : \{0, 1\}^n \rightarrow \{0, 1\}$  of size less than  $\text{sz}_{\text{TV}}(n)$  that computes  $L_n^{\text{TV}}$ , which yields a contradiction to the definition of  $\text{sz}_{\text{TV}}$ .

Specifically, consider the circuit  $C'_n : \{0, 1\}^n \rightarrow \{0, 1\}$  that acts as follows. Let  $p$  be a power of two such that  $t_0(n + p) \leq \text{sz}_{\text{TV}}^3(n) < t_0(n + 2p)$ ; there exists such a  $p$  since  $t_0(n + 2^{\lceil \log(n) \rceil}) \leq t_0(n)^3 < \text{sz}_{\text{TV}}^3(n)$ . The value of this  $p$  is hard-coded into  $C'_n$ . Given  $x \in \{0, 1\}^n$ , the circuit  $C'_n$  pads  $x$  with  $1^p$ , simulates the circuit  $C_m$  on  $(x, 1^p)$  (where  $m = n + p$ ), and outputs  $C_m(x, 1^p)$ . By the definition of  $L^{\text{pad}}$  it holds that  $C'_n$  correctly computes  $L_n^{\text{TV}}$ . The size of  $C'_n$  is dominated by the size of  $C_m$ , and is thus at most  $O(t_0(n + p)^{1/4}) = o(t_0(n + p)^{1/3})$ . Since  $t_0(n + p)^{1/3} \leq \text{sz}_{\text{TV}}(n)$  and  $n$  is sufficiently large, the size of  $C'_n$  is less than  $\text{sz}_{\text{TV}}(n)$ , which yields a contradiction.

Let us now see that  $L^{\text{pad}}$  can be decided by an MA verifier that runs in time  $t_2$  and uses one bit of advice. Given an input  $z$  of length  $m$ , the advice bit is set to one if and only if  $L_m^{\text{pad}} \neq \emptyset$ ; if the advice is zero, the verifier immediately rejects. Otherwise, the verifier computes  $n = n(m)$  and  $p = p(m)$ , and parses the input  $z$  as  $(x, 1^p)$  where  $|x| = n$  (if the verifier fails to parse the input, it immediately rejects). The verifier parses the witness  $w \in \{0, 1\}^{\text{poly}(t_0(n+2p))}$  as a circuit  $C : \{0, 1\}^n \rightarrow \{0, 1\}$  of size at most  $t_0(n + 2p)^{1/3}$ , and emulates the machine  $M$  from Lemma 26 on input  $x$ , answering each oracle query of  $M$  using the circuit  $C$ . The verifier outputs the decision of  $M$ .

Since  $\text{sz}_{\text{TV}}(n) < t_0(n + 2p)^{1/3}$ , there exists a circuit  $C$  of size at most  $t_0(n + 2p)^{1/3}$  that computes  $L_n^{\text{TV}}$ . For any  $z \in L^{\text{pad}}$ , when the witness represents this circuit, the verifier accepts  $z$  with probability one. Also, for any  $z \notin L^{\text{TV}}$ , the verifier rejects  $x$  with probability  $2/3$ , regardless of the witness. Finally, note that the running time of the verifier is dominated by the time that it takes to run the machine  $M$  while simulating the oracle answers, which is at most  $\text{poly}(n) \cdot \text{poly}(t_0(2m)) \leq t_2(m)$ . ■

## Appendix B Sufficient conditions for admissibility

The point of the current appendix is to show that essentially any increasing function  $f(n) = \omega(1)$  such that  $f(n) \leq n$  is admissible (in the sense of Definition 16).

**Claim 27.** *Let  $f(n) = \omega(1)$  be any increasing function such that  $f(n) \leq n$  for all  $n$ , and*

$t(n) = n^{f(n)}$  is time-constructible, and  $s(n) = n^{\log(f(\log(n)))}$  is time-constructible, and  $s'(n)$  is time-constructible. Then,  $f$  is admissible.

**Proof.** Let  $g(n) = \log(f(\log(n)))$  and let  $s(n) = n^{g(n)}$ . We need to verify that  $g$  is super-constant (which holds because  $f$  is super-constant), and that  $t$  and  $s$  are sufficiently gapped, and that  $\hat{s}(n) = n^{o(f(n))}$ . To see that  $t$  and  $s$  are sufficiently gapped, first note that both functions are increasing (since  $f$  is increasing, and hence  $g$  is also increasing) and are time-constructible, as is  $s'$  (we assumed time-constructibility in the hypothesis). Also note that  $s(n) \leq n^{\log \log(n)} < 2^{n/\gamma}/n$ .

Thus, it is left to verify that  $\hat{s}(n) = n^{o(f(n))}$ . The proof of this fact amounts to the following elementary calculation. First note that

$$s'(n) = (s(\gamma \cdot n))^\gamma = (\gamma \cdot n)^{\gamma \cdot \log(f(\log(\gamma \cdot n)))} < n^{\log^2(f(\log^2(n)))}.$$

Thus, for any function  $k = k(n)$  and constant  $c \geq 2$  such that  $k(n) \leq \log^c(f(\log^{3c}(n)))$  (which in particular implies that  $k(n) \leq \log^c(n)$ ), we have that

$$s'(n^k) < n^{k \cdot \log^2(f(\log^2(n^k)))} \leq n^{\log^{2c}(f(\log^{3c}(n)))}. \quad (1)$$

In particular, using Eq. (1) with  $k(n) = \log^2(f(\log^2(n)))$  and  $c = 2$ , we deduce that  $s'(s'(n)) < n^{\log^4(f(\log^6(n)))}$ . Then, using Eq. (1) again with  $k(n) = \log^4(f(\log^6(n)))$  and  $c = 4$ , we deduce that  $s'(s'(s'(n))) < n^{\log^8(f(\log^{12}(n)))}$ . Therefore, we have that  $\hat{s}(n) < n^{\gamma' \cdot \log^8(f(\log^{12}(n)))} < n^{\gamma' \cdot \text{poly} \log(f(n))} = n^{o(f(n))}$ . ■