

Relations between Average-case and Worst-case Complexity

A. Pavan* N. V. Vinodchandran†

May 9, 2006

Abstract

The consequences of the worst-case assumption $NP = P$ are very well understood. On the other hand, we only know a few consequences of the analogous average-case assumption “NP is easy on average.” In this paper we establish several new results on the *worst-case complexity* of Arthur-Merlin games (the class AM) under the *average-case complexity* assumption “NP is easy on average.”

- We first consider a stronger notion of “NP is easy on average” namely NP is easy on average with respect to distributions that are computable by polynomial size circuit families. Under this assumption we show that AM can be derandomized to nondeterministic subexponential time.
- Under the assumption that NP is easy on average with respect to polynomial-time computable distributions, we show (a) $AME = E$ where AME is the exponential version of AM. This improves an earlier known result that if NP is easy on average then $NE = E$ (b) For every $c > 0$, $AM \subseteq [io\text{-pseudo}_{NTIME(nc)}]\text{-}NP$. Roughly this means that for any language L in AM there is a language L' in NP so that it is computationally infeasible to distinguish L from L' .

We use recent results from the area of derandomization for establishing our results.

1 Introduction

Can an average-case complexity collapse lead to the collapse of worst-case complexity classes? In this paper we explore this question in the average-case complexity framework introduced by Levin [Lev86].

In Levin’s framework, an average-case complexity class consists of *distributional problems*. A distributional problem is a pair consisting of a decision problem, say A , and a probability

*Department of Computer Science, Iowa State University, Ames, IA 50011. pavan@cs.iastate.edu. Research supported by NSF grants CCR-0344817 and CCF-0430807.

†Department of Computer Science and Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588. vinod@cse.unl.edu. Research supported by NSF grant CCF-0430991, University of Nebraska Layman Award, and Big 12 Fellowship

distribution μ on the instances of A . Such a distributional problem is in average polynomial time (in notation AvgP) if there is a deterministic algorithm that solves A so that the running time of the algorithm is polynomial on average with respect to μ (the exact definition of polynomial on the average is given later). The average-case complexity class DistNP is the set of pairs (A, μ) where $A \in \text{NP}$ and μ is a polynomial-time computable distribution. The average-case analog of $\text{NP} \stackrel{?}{=} \text{P}$ question is whether $\text{DistNP} \subseteq \text{AvgP}$. Intuitively $\text{DistNP} \subseteq \text{AvgP}$ means NP is easy on average. Levin [Lev86] showed that there are distributional problems that are complete for DistNP in the sense that NP is easy on average if and only if the complete problem is solvable in AvgP. We refer the reader to [Gur91, Wan97] for pointers to many nice results in this area.

Understanding relations between the average-case complexity world and the worst-case complexity world is an important problem in complexity theory. In general we would like to answer the following question: if NP is easy on average, what happens in the worst-case complexity world?

Several results that connect the average-case complexity of NP to the worst-case complexity of classes such as NE, BPP, and MA are known. Ben-David, Chor, Goldreich, and Luby [BDCGL92] is the first to give an unlikely worst-case collapse under the assumption that NP is easy on average. They showed that if NP is easy on average, then the nondeterministic exponential-time (NE) collapses to the deterministic exponential-time (E). Ben-David et al. observed that if a machine runs in average polynomial time on unary strings, then it should in fact run in worst-case polynomial time. The collapse now follows using the fact that $E = \text{NE}$ if and only if all the unary languages in NP are in P [Boo74]. The assumption that NP is easy on average also implies derandomization of certain randomized classes. Impagliazzo [Imp95], using the results of Nisan and Wigderson [NW94], showed that if $\text{DistNP} \subseteq \text{AvgP}$ then $\text{BPP} = \text{ZPP}$. Köbler and Schuler [KS04] explored this connection further and showed that the assumption also derandomizes Babai's Merlin-Arthur (MA) games to NP. Later Arvind and Köbler [AK02] showed that in fact the assumption implies $\text{AM} \cap \text{co-AM} = \text{NP} \cap \text{co-NP}$ where AM is the class of languages accepted by one round Arthur-Merlin games. Recently, Buhrman, Fortnow, and Pavan [BFP03] showed that if NP is easy on average, then pseudorandom generators against sub-exponential circuits exists and hence $\text{BPP} = \text{P}$.

Thus the average-case easiness of NP implies derandomization of BPP and MA. It is natural to ask whether AM also can be derandomized if NP is easy on average. However, this question is still open. In this paper we explore this possibility.

Before we explain the results of this paper, we like to remark that exploring the connections between average-case and worst-case complexity under different notions of average-case efficiency is a very active research topic with many applications in complexity theory. For example, it is known that if the complexity class E can be approximated by small circuits, then E indeed has small circuits [BFNW91, IW97]. Problems such as the Permanent which are *random self-reducible* have their worst-case complexity equal to their average-case complexity [Lip91, GS92, FL92]. Such average-case to worst-case results based on random self-reducibility are strong enough to give average-case to worst-case collapse in Levin's setting. For example random self-reducibility of PSPACE-complete problems can be directly used to show that if PSPACE is easy on average (in Levin's sense), then $\text{PSPACE} = \text{P}$ [KS04].

We first consider a stronger notion of NP being easy on average. We show that if NP is easy on average with respect to distributions that are computable by polynomial-size circuits, then AM can be derandomized to nondeterministic subexponential time. Under the assumption $\text{DistNP} \subseteq \text{AvgP}$, we show that tally languages in AM can be derandomized to NP and hence $\text{AME} = \text{E}$ where AME is the exponential version of AM. This improves the earlier mentioned collapse result $\text{DistNP} \subseteq \text{AvgP} \Rightarrow \text{NE} = \text{E}$ [BDCGL92]. We also show that $\text{DistNP} \subseteq \text{AvgP}$ implies AM is almost NP: for any language $L \in \text{AM}$ there is a language L' in NP so that L and L' are computationally indistinguishable infinitely often. We use recent results from the area of derandomization to prove our results.

2 Preliminaries

2.1 Worst-case complexity classes

We assume familiarity with definitions of standard complexity classes such as P, NP, BPP, P/poly, NP/poly, E, and NE. Refer to [BDG88, Pap94] for definitions of standard complexity classes which are not defined in this paper.

Babai defined Arthur-Merlin games as a combinatorial game, played by Arthur—a probabilistic polynomial-time machine (with public coins), and Merlin—a computationally unbounded Turing machine [Bab85, BM88]. We will use the following operator based definition of the Arthur-Merlin classes. Let $\text{NLIN} = \text{NTIME}(O(n))$.

A language $L \in \text{AMTIME}(t(n))$ if there exist a language $A \in \text{NLIN}$ so that for all inputs x ,

$$\begin{aligned} x \in L &\Rightarrow \Pr_{y \in \{0,1\}^{t(|x|)}}[\langle x, y \rangle \in A] \geq 2/3 \\ x \notin L &\Rightarrow \Pr_{y \in \{0,1\}^{t(|x|)}}[\langle x, y \rangle \in A] \leq 1/3 \end{aligned}$$

A language $L \in \text{MATIME}(t(n))$ if there exist a probabilistic polynomial-time machine M running in linear time such that for all inputs x ,

$$\begin{aligned} x \in L &\Rightarrow \exists y \in \{0,1\}^{t(|x|)} \Pr[M(x, y) \text{ accepts}] \geq 2/3 \\ x \notin L &\Rightarrow \forall y \in \{0,1\}^{t(|x|)} \Pr[M(x, y) \text{ accepts}] \leq 1/3 \end{aligned}$$

$\text{AM} = \cup_{k>0} \text{AMTIME}(n^k)$ and $\text{MA} = \cup_{k>0} \text{MATIME}(n^k)$.

We need the exponential versions of certain complexity classes: $\text{E} = \text{DTIME}(2^{O(n)})$, $\text{EXP} = \text{DTIME}(2^{n^{O(1)}})$, $\text{NEXP} = \text{DTIME}(2^{2^{n^{O(1)}}})$, $\text{AME} = \text{AMTIME}(2^{O(n)})$. For the nondeterministic case we will also need the sub-exponential version $\text{NSUBEXP} = \cap_{\epsilon>0} \text{NTIME}(2^{n^\epsilon})$.

Let $\Sigma = \{0,1\}$ and $A^n = \Sigma^n \cap A$. For any complexity class \mathcal{C} , the class ${}_{\text{i}}\mathcal{C}$ is the class of languages $\{A \mid \exists B \in \mathcal{C} \text{ such that for infinitely many } n, A^n = B^n\}$.

We will use the following standard hierarchy results when needed without referring to them.

Theorem 1 (Deterministic time hierarchy [GHS91]). *For every $k > 0$, $\text{E} \not\subseteq {}_{\text{i}}\text{DTIME}(2^{kn})$.*

Theorem 2 (Nondeterministic time hierarchy [SFM78, Ž83]). *Let t and T be two time constructible functions so that $t(n+1) = o(T(n))$. Then $\text{NTIME}(T(n)) \not\subseteq \text{NTIME}(t(n))$.*

2.2 Average-case complexity

We review some basics of the theory of average-case complexity. The notation that we follow is mostly in line with Wang’s survey [Wan97].

A *distributional problem* is a pair (A, μ) so that $A \subseteq \Sigma^*$ and μ is a probability distribution on Σ^* . Given a distribution μ , μ^* denotes the *distribution function* of μ . Recall that $\mu^*(x) = \sum_{y \leq x} \mu(y)$. A distribution μ is *polynomial-time computable* if its distribution function is polynomial-time computable. A distribution is *P/poly-computable* if there exists a polynomial-size circuit family (C_1, C_2, \dots) such that for every x , $\mu^*(x) = C_{|x|}(x)$.

DistNP is the class of distributional problems (A, μ) so that $A \in \text{NP}$ and μ is a polynomial time computable distribution. Similarly $\text{DistNP}_{\text{P/poly}}$ is the class of distributional problems (A, μ) so that $A \in \text{NP}$ and μ is a P/poly-computable distribution.

Definition 1. Let μ be a distribution on Σ^* .

- A function $f : \Sigma^* \rightarrow \mathbf{N}$ is said to be n^k on μ -average [Lev86] if

$$\sum_{x \in \Sigma^*} \frac{f^{1/k}(x) \mu(x)}{|x|} < \infty^1.$$

- A distributional problem (A, μ) is in $\text{AvgTIME}(n^k)$ if there exists a deterministic Turing machine M that accepts A so that the running time of M is n^k on μ -average. (A, μ) is in AvgP if there exists a k so that (A, μ) is in $\text{AvgTIME}(n^k)$.

By “NP is easy on average” we mean $\text{DistNP} \subseteq \text{AvgP}$. Thus if $\text{DistNP} \subseteq \text{AvgP}$, then for every language L in NP, and every polynomial-time computable distribution μ , $(L, \mu) \in \text{AvgP}$.

Levin [Lev86] introduced the notion of reductions between distributional problems and defined DistNP-completeness. There are distributional problems that are DistNP-complete in the sense that NP is easy on average if and only if the complete problem is in AvgP. Proving distributional completeness is much more challenging than proving usual NP-completeness since the reductions must satisfy certain additional *domination* properties.

Definition 2. ([Lev86]) Let μ and ν be two distributions on Σ^* .

- We say that ν *dominates* μ *within* n^k if for all x , $\mu(x) \leq |x|^k \nu(x)$.
- Let $f : \Sigma^* \rightarrow \Sigma^*$. Then we say that ν *dominates* μ *within* n^k *via* f if there is a distribution μ_1 such that μ_1 dominates μ within n^k , and for all y in the range of f $\nu(y) = \sum_{x=f(x)} \mu_1(x)$.
- (A, μ) *reduces to* (B, ν) if there is a polynomial time computable many-one reduction f from A to B so that for some k , ν dominates μ within n^k via f

¹Technically $x \neq \lambda$. We will assume this through out the paper

Gurevich [Gur91] showed that the *distributional halting problem* (K, μ_K) is complete for DistNP, where $K = \{\langle i, x, 0^n \rangle \mid N_i \text{ accepts } x \text{ in } n \text{ steps}\}$, and $\mu_K(\langle i, x, 0^n \rangle) = \frac{1}{2^{|i||x|^2}} \frac{1}{2^{|x||x|^2}} \frac{1}{n^2}$. Here N_i denotes the i^{th} nondeterministic Turing machine in some fixed enumeration. We denote distributional halting problem by DH in this paper.

Let U denote the standard uniform distribution which puts a weight of $\frac{1}{n^2 \times 2^n}$ on a string of length n . Any machine accepting a problem (L, U) in AvgTIME(n^k) should solve all but a polynomial fraction of the inputs in polynomial time. In particular, a standard averaging argument shows the following proposition.

Proposition 1. *If $(L, U) \in \text{AvgTIME}(n^k)$, then there is a deterministic Turing machine M that decides L and for all but finitely many n , there exist at most $2^n/n^2$ strings of length n , on which M takes more than n^{4k} time. Here U is the standard uniform distribution on Σ^* .*

We need two more definitions to describe what we mean by closeness of complexity classes.

Definition 3. ([Kab01]) Two languages L and L' are ae-NTIME(n^c)-distinguishable, if there is a n^c -time bounded nondeterministic machine N such that for all but finitely many n , $M(0^n)$ outputs a string from $L\Delta L'$ along every accepting path. We say L and L' are io-NTIME(n^c)-indistinguishable, if they are not ae-NTIME(n^c)-distinguishable.

Thus if L and L' are io-NTIME(n^c)-indistinguishable, then no NTIME(n^c) machine can detect the places at which L and L' differ at all input lengths.

Definition 4. We say $\text{AM} \subseteq [\text{io-pseudo}_{\text{NTIME}(n^c)}]\text{-NP}$, if for every language L in AM there exists a language L' in NP such that L and L' are io-NTIME(n^c)-indistinguishable.

3 Derandomizing Arthur-Merlin Games

We show several results on the worst-case complexity of Arthur-Merlin games under the assumption that NP is easy on average. First we consider a stronger notion of NP being easy on average— $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$. We prove that under this assumption $\text{AM} \subseteq \text{NSUBEXP}$.

Before we present our results, we like to compare the assumptions “ $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$ ” and “ $\text{DistNP} \subseteq \text{AvgP}$ ”. The former states that every language in NP is easy on average with respect to every P/poly-computable distribution, whereas the later says that every language in NP is easy on average with respect to every p -computable distribution. It is reasonable to conjecture that the former assumption is stronger than the later one. Consider a language that is not in P. It is conceivable that a small circuit can store “hard instances” of this language, but no polynomial-time algorithm might be able to detect those hard instances. Then this language would be hard on average with respect to a P/poly-computable distribution, but is easy on average with respect to p -computable distributions. We can show that such languages exist in EXP. Thus the hypothesis “ $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$ ” seems to be stronger than the hypothesis “ $\text{DistNP} \subseteq \text{AvgP}$ ”. However, we do not know of any candidate language in NP that is likely to be easy on average with P/poly-computable distributions, but not easy on average with p -computable distributions. We also point that, prior to this paper,

we do not know of any consequence that is known to follow from the former hypothesis, but is not known to follow from the later hypothesis.

We show as one of our main results that if $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$, then $\text{AM} \subseteq \text{NSUBEXP}$. We first show that the hypothesis implies $E \not\subseteq_{\text{i.o.}}(\text{NP/poly})$. The result then follows from the following theorem due to Shaltiel and Umans [SU01].

Theorem 3 ([SU01]). *If $E \not\subseteq_{\text{i.o.}}(\text{NP/poly})$, then $\text{AM} \subseteq \text{NSUBEXP}$.*

We use the following chain of arguments. If $E \subseteq \text{NP/poly}$ and NP is easy on average then $E \subseteq \text{AvgP/poly}$ with respect to certain distribution (which is nonuniformly computable). Since there are random-self-reducible complete problems for E, we will actually get $E \subseteq \text{P/poly}$. From [BFNW91] we have that $E \subseteq \text{P/poly} \Rightarrow E = \text{MA}$. Since NP is easy on average we have $\text{MA} = \text{NP}$ [KS04] and $\text{NE} = E$ [BDCGL92]. So we will finally get $\text{NE} = \text{NP}$ which is a contradiction to the nondeterministic time hierarchy theorem. We now present a more formal argument. We start with the definition of a non-uniform version of AvgP.

Definition 5. A distributional problem (L, μ) is in AvgP/poly , if there is a Turing machine M and a polynomial-bounded function $a : \mathbb{N} \rightarrow \Sigma^*$, such that $x \in L \Leftrightarrow M$ accepts $\langle x, a(|x|) \rangle$, and there exists a constant $k > 0$, such that

$$\sum_x \frac{T_M^{1/k}(\langle x, a(|x|) \rangle)}{|x|} \mu(x) < \infty.$$

where $T_M(y)$ denotes the running time of M on input y .

Note that the above definition requires the running time of M to be polynomial on average only on correct advice. The following proposition is easy to prove.

Proposition 2. *Let (L, U) is in AvgP/poly , where U is the uniform distribution. Let M be a machine and $a : \mathbb{N} \rightarrow \Sigma^*$ be a function that witness that (L, U) is in AvgP/poly . Then there is a constant $k > 0$ such that for every n , $M(x, a|x|)$ halts within n^k steps on more than $(1 - 1/n^2)$ fraction of strings from Σ^n .*

We need the following lemma regarding random-self-reducibility of E complete problems.

Lemma 1 ([BFNW91]). *There exists a random-self-reducible complete language $L \in E$ so that if there is a polynomial-size circuit C such that*

$$\forall n, \Pr_x[C(x) = L(x)] \geq 1 - 1/n^2,$$

then $L \in \text{P/poly}$.

We now prove the result we need to derandomize AM (actually we need an i.o. version of the following theorem. But we state and prove a cleaner non-i.o. version first).

Theorem 4. *If $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$ then $E \not\subseteq \text{NP/poly}$.*

Proof. Assume $E \subseteq \text{NP/poly}$. Let L be a complete language for E that is random self-reducible (provided by Lemma 1). Since $L \in \text{NP/poly}$, there is a language L' in NP and a polynomial-bounded advice function $a : \mathbb{N} \rightarrow \Sigma^*$, such that

$$\forall x, x \in L \Leftrightarrow \langle x, a(|x|) \rangle \in L'$$

Consider the following distribution μ

$$\mu(\langle x, y \rangle) = \begin{cases} U(x) & \text{if } y = a(|x|) \\ 0 & \text{otherwise} \end{cases}$$

Here U is the uniform distribution on Σ^* , i.e., $U(x) = \frac{1}{n^2} \frac{1}{2^n}$, where $n = |x|$. It is clear that μ is P/poly -computable. Since $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$, $(L', \mu) \in \text{AvgP}$. Consider the following reduction from (L, U) to (L', μ) : $f(x) = \langle x, a(|x|) \rangle$. It is clear that f is P/poly -computable and satisfies the dominance condition. Thus $(L, U) \in \text{AvgP/poly}$.

Thus there is a Turing machine M and an advice function $a : \mathbb{N} \rightarrow \Sigma^*$, such that $x \in L \Leftrightarrow M$ accepts $\langle x, a(|x|) \rangle$, and there exists a constant l , such that

$$\sum_x \frac{T_M^{1/l}(\langle x, a(|x|) \rangle)}{|x|} \mu(x) < \infty.$$

By Proposition 2, there is a constant k , such that for every n , on at least $(1 - 1/n^2)$ fraction of strings of the form $\langle x, a(|x|) \rangle$, M halts within n^k steps, and M accepts $\langle x, a(|x|) \rangle$ if and only if $x \in L$.

We now claim that this implies $L \in \text{P/poly}$. Define a new machine M' as follows: M' on input $\langle x, y \rangle$ runs M on $\langle x, y \rangle$ for n^k steps. If M does not halt within n^k steps, then M' rejects its input. If M halts within n^k steps, then M' accepts $\langle x, y \rangle$ if and only if M accepts $\langle x, y \rangle$.

By converting M' into a circuit and hardwiring $a(|x|)$ into it, we obtain a polynomial-size circuit C such that

$$\forall n, \Pr[C(x) = L(x)] \geq 1 - 1/n^2.$$

Since L is random self-reducible, by Lemma 1, $L \in \text{P/poly}$. Since L is complete for E , $E \subseteq \text{P/poly}$. By [BFNW91], if $E \subseteq \text{P/poly}$, then $E \subseteq \text{MA}$. If $\text{DistNP} \subseteq \text{AvgP}$, then $\text{MA} = \text{NP}$ [KS04] and also $\text{NE} = \text{E}$ [BDCGL92]. Thus we have $\text{NE} = \text{NP}$. A contradiction follows from the nondeterministic time hierarchy theorem. \square

We can actually obtain a stronger version of the above theorem.

Theorem 5. *If $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$, then $E \not\subseteq_{\text{io}} \text{NP/poly}$.*

Proof. Essentially the same proof will show that if $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$ then $E \not\subseteq_{\text{io}} (\text{NP/poly})$. To see this, if $E \subseteq_{\text{io}} (\text{NP/poly})$ then we will get that E can be approximated by a polynomial size circuit on infinitely many input lengths. As argued in [BFNW91] we then get that $E \subseteq_{\text{io}} (\text{P/poly}) \subseteq_{\text{io}} \text{MA}$. If $\text{DistNP} \subseteq \text{AvgP}$, then $\text{MA} = \text{NP}$ [KS04], Thus we get $E \subseteq_{\text{io}} \text{NP}$. Since the nondeterministic time hierarchy theorem is not sufficient to separate NE from ioNP we need to argue slightly differently. By [BDCGL92], if $\text{DistNP} \subseteq \text{AvgP}$,

then $E = NE$. By a result of Impagliazzo, Kabanets, and Wigderson [IKW02], if $E = NE$, then there is a fixed constant k such that $\text{NTIME}(2^n) \subseteq \text{DTIME}(2^{kn})$. Thus we obtain $E \subseteq {}_{\text{io}}\text{NP} \subseteq {}_{\text{io}}\text{NTIME}(2^n) \subseteq {}_{\text{io}}\text{DTIME}(2^{kn})$. This contradicts the deterministic time hierarchy theorem. \square

Theorem 6. *If $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$, then $\text{AM} \subseteq \text{NSUBEXP}$.*

Proof. By Theorem 5, the assumption implies that E is not in NP/poly almost everywhere. By Theorem 3, $\text{AM} \subseteq \text{NSUBEXP}$. \square

Before we continue with further results on AM , we quickly discuss a result on the worst-case complexity of computing satisfying assignments for a satisfiable formula. Ben-David, Chor, Goldreich, and Luby [BDCGL92] showed that if $\text{DistNP} \subseteq \text{AvgP}$, then $E = NE$. Buhrman, Fortnow, and Pavan showed that in fact a stronger conclusion follows namely if $\text{DistNP} \subseteq \text{AvgP}$, then for every NE -predicate $R(x, y)$, there is an E machine M such that $M(x)$ outputs a y such that $R(x, y)$ holds. Buhrman [Buh93] showed that if $E = NE$, and $\text{FSAT} \in \text{PF}_{tt}^{\text{NP}}$, then witnesses of NE predicates can be computed in E -time. Here FSAT denotes the problem of computing a satisfying assignment of a satisfiable propositional formula, and $\text{PF}_{tt}^{\text{NP}}$ is the class of functions that can be computed by polynomial-time machines which make nonadaptive queries to an NP -oracle. This raises the following question: “If $\text{DistNP} \subseteq \text{AvgP}$, is $\text{FSAT} \in \text{PF}_{tt}^{\text{NP}}$?” We have a partial answer.

Theorem 7. *If $\text{DistNP}_{\text{P/poly}} \subseteq \text{AvgP}$, then $\text{FSAT} \in \text{SUBEXP}_{tt}^{\text{NP}}$.*

Proof. By Theorem 5, the hypothesis implies that $E \not\subseteq {}_{\text{io}}(\text{NP/poly})$. Very recently, Shaltiel and Umans [SU05] showed that $E \subseteq {}_{\text{io}}\text{NP/poly}$ if and only if $E \subseteq {}_{\text{io}}\text{P}_{tt}^{\text{NP}}/\text{poly}$. Thus the hypothesis implies the existence of a language in E that does not have $\text{P}_{tt}^{\text{NP}}$ -circuits almost everywhere. Klivans and van Melkebeek [KvM02] showed that such languages can be used to construct pseudo-random generators with polynomial stretch that are secure against $\text{P}_{tt}^{\text{NP}}$ -circuits. These pseudo-random generators can be used to derandomize $\text{BPP}_{tt}^{\text{NP}}$ to $\text{SUBEXP}_{tt}^{\text{NP}}$. Since $\text{FSAT} \in \text{BPP}_{tt}^{\text{NP}}$ [VV85], we have the conclusion. \square

Now we show that, under the assumption $\text{DistNP} \subseteq \text{AvgP}$, $\text{AME} = E$, and $\text{AM} \subseteq [{}_{\text{io}}\text{-pseudo}_{\text{NTIME}(n^c)}]\text{-NP}$. We also give an alternate proof of the result due to Arvind and Köbler [AK02] that under the assumption $\text{AM} \cap \text{co-AM} = \text{NP} \cap \text{co-NP}$. We use uniform derandomization results for AM and $\text{AM} \cap \text{co-AM}$ due to Gutfreud, Shaltiel, and Ta-Shma [GSTS03], and the fact that pseudorandom generators exist if NP is easy on average [BFP03].

Theorem 8 ([GSTS03]). *If $E \not\subseteq \text{AMTIME}(2^{\beta n})$ for some constant β then for every $c > 0$, $\text{AM} \subseteq [{}_{\text{io}}\text{-pseudo}_{\text{NTIME}(n^c)}]\text{-NP}$.*

Theorem 9 ([BFP03]). *If $\text{DistNP} \subseteq \text{AvgP}$ then there is an algorithm that on input 1^n outputs a pseudorandom set for circuits of size n . The algorithm runs in time n^c for a fixed c .*

Recall that a set S is a pseudorandom set for circuits of size $s(n)$, if for every circuit C of size $s(n)$ if $\Pr_{x \in \Sigma^n}[C(x) = 1]$ is close to $\Pr_{x \in S}[C(x) = 1]$. Efficiently computable pseudorandom generators implies that for any k there exists a k' so that $\text{BPTIME}(n^k) \subseteq \text{DTIME}(n^{k'})$.

We use the following line of argument. Under the assumption $\text{DistNP} \subseteq \text{AvgP}$, we show that $\text{AMTIME}(2^n)$ is a subset of $\text{DTIME}(2^{kn})$ for a fixed k . The same argument will also show that ${}_{\text{io}}\text{AMTIME}(2^n) \subseteq {}_{\text{io}}\text{DTIME}(2^{kn})$. Since for any fixed k , $\text{E} \not\subseteq {}_{\text{io}}\text{DTIME}(2^{kn})$ we get that $\text{E} \not\subseteq {}_{\text{io}}\text{AMTIME}(2^n)$ and therefore $\text{AM} \subseteq \text{Pseudo}_{\text{NTIME}(n^c)}\text{-NP}$ from Theorem 8.

For showing that $\text{AMTIME}(2^n) \subseteq \text{DTIME}(2^{kn})$ for a fixed k , we first observe a result that if $\text{DistNP} \subseteq \text{AvgP}$ then DistNLIN is in $\text{AvgTIME}(n^k)$ for a fixed k . Here $\text{DistNLIN} = \{(A, \mu) \mid A \in \text{NLIN} \text{ and } \mu^* \text{ is linear-time computable}\}$. We use this result to first show that every tally language in $\text{AMTIME}(n)$ is in $\text{BPTIME}(n^k)$ for a fixed k . Finally we use the pseudorandom generator from Theorem 9 to get $\text{BPTIME}(n^k) \subseteq \text{DTIME}(n^l)$ for some l . A standard padding gives the collapse in the exponential level.

We know that if $\text{NP} = \text{P}$ then for any k there is an l so that $\text{NTIME}(n^k) \subseteq \text{DTIME}(n^l)$. One way to see this is to observe that any problem in $\text{NTIME}(T(n))$ is $T^2(n)$ time reducible to SAT and the result follows since under the assumption SAT is in $\text{DTIME}(n^{k'})$ for a fixed k' . We observe that such a theorem exists in the average-case setting also. We use the completeness of the Distributional Halting problem [Gur91] to prove an analogous result in the average-case setting. A proof can be found in [Wan97].

Definition 6. $\text{DistNLIN} = \{(A, \mu) \mid A \in \text{NLIN} \text{ and } \mu^* \text{ is linear-time computable}\}$ where μ^* is the distribution function of the density function μ .

Lemma 2 ([Lev86, Gur91]). *Let (A, μ) and (B, ν) be two distributional problems and let f be a reduction from A to B so that:*

- f is computable in time n^l
- ν dominates μ within n^r via f

Then if $(B, \nu) \in \text{AvgTIME}(n^k)$ then $(A, \mu) \in \text{AvgTIME}(n^{2klr})$.

Lemma 3 ([Lev86, Gur91]). *Every distributional problem $(A, \mu) \in \text{DistNLIN}$ is reducible to the DistNP complete problem $\text{DH} = (K, \mu_K)$ via a reduction f_A so that:*

- f_A is computable in time n^3
- μ_K dominates μ within n^3 via f_A

Using the above two lemmas we get the following theorem.

Theorem 10. *If $\text{DistNP} \subseteq \text{AvgP}$ then there exists a k so that $\text{DistNLIN} \subseteq \text{AvgTIME}(n^k)$.*

Now we prove our main Lemma.

Lemma 4. *If $\text{DistNP} \subseteq \text{AvgP}$ then there exists a k so that $\text{AMTIME}(2^n) \subseteq \text{DTIME}(2^{kn})$.*

Proof. We show that under the assumption any tally set in $\text{AMTIME}(n)$ is in $\text{DTIME}(n^l)$ for a fixed l . The Lemma follows from the fact that for any language $L \in \text{AMTIME}(2^n)$, the tally version is in $\text{AMTIME}(n)$ and similarly if the tally version of L is in $\text{DTIME}(n^k)$ then $L \in \text{DTIME}(2^{ln})$.

Let L be a tally language in $\text{AMTIME}(n)$. Then there is a language $B \in \text{NLIN}$ so that

$$\begin{aligned} 0^n \in L &\Rightarrow \Pr_{y \in \Sigma^n}[\langle 0^n, y \rangle \in B] \geq 2/3 \\ 0^n \notin L &\Rightarrow \Pr_{y \in \Sigma^n}[\langle 0^n, y \rangle \in B] < 1/3 \end{aligned}$$

Consider the linear-time computable distribution μ which on a string $\langle 0^n, y \rangle$ has a probability of $\frac{1}{n^{2|y|}}$. The distributional problem (B, μ) is in DistNLIN and hence is in $\text{AvgTIME}(n^k)$ for a fixed k . Let M be the deterministic machine that witnesses this fact. Now consider a n^{5k} time deterministic machine M' that simulates M for n^{5k} steps and if decides according to M if M stops else rejects. An easy counting argument shows that for every n , M' correctly decides B on at least $1 - 1/n^2$ fraction of strings of the form $\langle 0^n, y \rangle$. From this it easily follows that $L \in \text{BPTIME}(n^{5k})$.

By Theorem 9, if $\text{DistNP} \subseteq \text{AvgP}$, there is an n^c time-bounded algorithm that outputs a pseudorandom set for circuits of size n . Using this we can derandomize $\text{BPTIME}(n^{5k})$ to $\text{DTIME}(n^{10k^2c})$. Thus L is in $\text{DTIME}(n^l)$ for $l = 10k^2c$. \square

Our theorems follow easily from the above lemma.

Theorem 11. *If $\text{DistNP} \subseteq \text{AvgP}$, then $\text{AME} = \text{E}$.*

Theorem 12. *If $\text{DistNP} \subseteq \text{AvgP}$ then, $\text{AM} \subseteq [\text{io-pseudo}_{\text{NTIME}(n^c)}] - \text{NP}$ for every $c > 0$.*

Proof. By Lemma 4, the hypothesis implies that $\text{AMTIME}(2^n) \subseteq \text{DTIME}(2^{kn})$ for a fixed $k > 0$. If $\text{E} \subseteq \text{AMTIME}(2^n)$, then it follows that $\text{E} \subseteq \text{DTIME}(2^{kn})$ for a fixed $k > 0$. This contradicts the time-hierarchy theorem. Thus $\text{E} \not\subseteq \text{AMTIME}(2^n)$. Thus by Theorem 8 $\text{AM} \subseteq [\text{io-pseudo}_{\text{NTIME}(n^c)}] - \text{NP}$ for every $c > 0$. \square

Our approach gives a different proof of the following result due to Arvind and Köbler [AK02].

Theorem 13 ([AK02]). *If $\text{DistNP} \subseteq \text{AvgP}$, then $\text{AM} \cap \text{co-AM} = \text{NP} \cap \text{co-NP}$.*

Proof. The same argument as in Lemma 4 shows that if $\text{DistNP} \subseteq \text{AvgP}$, then ${}_{\text{io}}\text{AMTIME}(2^n) \subseteq {}_{\text{io}}\text{DTIME}(2^{kn})$ for a fixed k . By Lemma 1, $\text{E} \not\subseteq {}_{\text{io}}\text{DTIME}(2^{kn})$ for any fixed k . Therefore $\text{E} \not\subseteq {}_{\text{io}}\text{AMTIME}(2^n)$. Gutfreud, Shaltiel, and Ta-shma [GSTS03] showed that this implies $\text{AM} \cap \text{co-AM} = \text{NP} \cap \text{co-NP}$. \square

Acknowledgments

We thank V. Arvind and Johannes Köbler for helpful discussions.

References

- [AK02] V. Arvind and J. Köbler. New lowness results for ZPP^{NP} and other complexity classes. *Journal of Computer and System Sciences*, 65(2):257–277, 2002.
- [Bab85] L. Babai. Trading group theory for randomness. In *Proc. 17th Annual ACM Symp. on Theory of Computing*, pages 421–429, 1985.
- [BDCGL92] S. Ben-David, B. Chor, O. Goldreich, and M. Luby. On the theory of average case complexity. *Journal of Computer and System Sciences*, 44(2):193–219, 1992.
- [BDG88] J. Balcázar, J. Diaz, and J. Gabarró. *Structural Complexity I*. Springer-Verlag, Berlin, 1988.
- [BFNW91] L. Babai, L. Fortnow, N. Nisan, and A. Wigderson. BPP has subexponential time simulations unless *exptime* has publishable proofs. In *Proceedings of the 6th Annual Conference on Structure in Complexity Theory, 1991*, pages 213–219, 1991.
- [BFP03] H. Buhrman, L. Fortnow, and A. Pavan. Some results on derandomization. In *Proceedings of the 20th Annual Symposium on Theoretical Aspects of Computer Science*, volume LNCS 2607, pages 212–222, 2003.
- [BM88] L. Babai and S. Moran. Arthur-merlin games: a randomized proof system, and a hierarchy of complexity class. *Journal of Computer and System Sciences*, 36(2):254–276, 1988.
- [Boo74] R. Book. Tally languages and complexity classes. *Information and Control*, 26:186–193, 1974.
- [Buh93] H. Buhrman. *Resource bounded reductions*. PhD thesis, University of Amsterdam, 1993.
- [FL92] U. Feige and C. Lund. On the hardness of computing permanent of random matrices. In *Proceedings of 24th Annual ACM Symposium on Theory of Computing*, pages 643–654, 1992.
- [GHS91] J. Geske, D. Huynh, and J. Seiferas. A note on almost-everywhere-complex sets and separating deterministic-time-complexity classes. *Information and Computation*, 92(1):97–104, 1991.
- [GS92] P. Gemmel and M. Sudan. Highly resilient correctors for polynomials. *Information Processing Letters*, 43:169–174, May 1992.
- [GSTS03] D. Gutfreund, R. Shaltiel, and A. Ta-Shma. Uniform hardness vs. randomness tradeoffs for Arthur-Merlin games. *Computational Complexity*, 12:85–130, 2003.

- [Gur91] Y. Gurevich. Average case completeness. *Journal of Computer and System Sciences*, 42:346–398, 1991.
- [IKW02] R. Impagliazzo, V. Kabanets, and A. Wigderson. In search of an easy witness: exponential time vs. probabilistic polynomial time. *Journal of Computer and System Sciences*, 65:672–694, 2002.
- [Imp95] R. Impagliazzo. A personal view of average-case complexity theory. In *Proceedings of the 10th Annual Conference on Structure in Complexity Theory*, pages 134–147. IEEE Computer Society Press, 1995.
- [IW97] R. Impagliazzo and A. Wigderson. $P = BPP$ if E requires exponential circuits: Derandomizing the XOR lemma. In *Proceedings of the 29th ACM Symposium on Theory of Computing*, pages 220–229, 1997.
- [Kab01] V. Kabanets. Easiness assumptions and hardness tests: trading time for zero error. *Journal of Computer and System Sciences*, 63(2):236–252, 2001.
- [KS04] J. Köbler and R. Schuler. Average-case intractability vs. worst-case intractability. *Information and Computation*, 190(1):1–17, 2004.
- [KvM02] A. Klivans and D. van Melkebeek. Graph nonisomorphism has subexponential size proofs unless the polynomial-time hierarchy collapses. *SIAM Journal on Computing*, 31:1501–1526, 2002.
- [Lev86] L. Levin. Average case complete problems. *SIAM Journal of Computing*, 15:285–286, 1986.
- [Lip91] R. Lipton. New directions in testing. In *Distributed Computing and Cryptography*, volume 2 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 191–202. American Mathematics Society, 1991.
- [NW94] N. Nisan and A. Wigderson. Hardness vs Randomness. *Journal of Computer and System Sciences*, 49(2):149–167, 1994.
- [Pap94] C. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.
- [SFM78] J. Seiferas, M. Fischer, and A. Meyer. Separating nondeterministic time complexity classes. *J. of the ACM*, 25(1):146–147, 1978.
- [SU01] R. Shaltiel and C. Umans. Simple extractors for all min-entropies and a new pseudo-random generator. In *42nd IEEE Symposium on Foundations of Computer Science*, pages 648–657, 2001.
- [SU05] R. Shaltiel and C. Umans. Pseudorandomness for approximate counting and sampling. In *20th IEEE Conference on Computational Complexity*, pages 212–226, 2005.

- [VV85] L. Valiant and V. Vazirani. NP is as easy as detecting unique solutions. In *Proc. 17th ACM Symp. Theory of Computing*, pages 458–463, 1985.
- [Ž83] S. Žák. A Turing machine time hierarchy. *Theor. Computer Science*, 26:327–333, 1983.
- [Wan97] J. Wang. Average-case computational complexity theory. In L. Hemaspaandra and A. Selman, editors, *Complexity Theory Retrospective II*, pages 295–328. Springer-Verlag, 1997.