Reductions Do Not Preserve Fast Convergence Rates in Average Time

Jay Belanger* A. Pavan† Jie Wang‡

Abstract

Cai and Selman [CS96] proposed a general definition of average computation time that, when applied to polynomials, results in a modification of Levin's [Lev86] notion of average-polynomial-time. The effect of the modification is to control the rate of convergence of the expressions that define average computation time. With this modification, they proved a hierarchy theorem for average-time complexity that is as tight as the Hartmanis-Stearns [HS65] hierarchy theorem for worst-case deterministic time. They also proved that under a fairly reasonable condition on distributions, called condition W, a distributional problem is solvable in average-polynomial-time under the modification exactly when it is solvable in average-polynomial-time under Levin's definition.

Various notions of reductions, as defined by Levin [Lev86] and others, play a central role in the study of average-case complexity. However, the class of distributional problems that are solvable in average-polynomial-time under the modification is not closed under

*Division of Mathematics and Computer Science, Truman State University, Kirksville, MO 63501, USA. Email: belanger@mathax.truman.edu. Supported in part by NSF under grant CCR-9503601.

†Department of Computer Science, State University of New York at Buffalo, Buffalo, NY 14260, USA. Email: aduri@cs.buffalo.edu Supported in part by NSF under grant CCR-9400229.

‡Department of Mathematical Sciences, University of North Carolina at Greensboro, Greensboro, NC 27412, USA. Email: wang@uncg.edu. Supported in part by NSF under grant CCR-9424164.

the standard reductions. In particular, we prove that there is a distributional problem that is not solvable in average-polynomial-time under the modification but is reducible, by the identity function, to a distributional problem that is, and whose distribution even satisfies condition W.

1 Introduction

Average-case analyses of algorithms often provide more useful information than worst-case analyses. For example, quick-sort is a useful algorithm because it runs in $O(n \log n)$ time on average under a uniform distribution, even though its worse-case complexity is $\Theta(n^2)$. More to the point of this paper, there are NP-hard problems whose expected running time is polynomial. For example, the Hamiltonian path problem can be solved in expected linear time under commonly used distributions on random graphs [GS87].

A distributional problem is a decision problem paired with a probability distribution on instances. Given a distributional problem, it is an important issue to either find an expected polynomial-time algorithm that solves the problem or prove that such an algorithm does not exist. Levin [Lev86] provided two central notions for studying this issue. One is analogous to the class P, and provides an *easiness* notion; the other is analogous to the class of NP-complete sets, and provides a *hardness* notion. For the first, Levin defined a robust notion on what it means for the running time of an algorithm to be polynomial on average. (In this case, we say that the problem is in AP.) For the second, Levin defined reductions between distributional problems. These reductions are transitive and if a distributional problem is reducible to a second distributional problem that is in AP, then the original distributional problem is also in AP. With this machinery in place, a distributional problem is *average-case* NP-complete if the decision problem component belongs to NP and every distributional problem consisting of an NP problem and a reasonable distribution [1] is reducible to it. Levin showed that distributional tiling with a simple distribution is average-case NP-complete and since then, several additional average-case NP-complete problems have been found [BG95, Gur91, VL88, VR92, WB95, Wan95].

[1] By reasonable distribution we mean a distribution that is polynomial-time computable or is dominated by a distribution that is. See section 2 for definition.

Levin's definitions concern only the distinction between polynomial on average and super-polynomial on average. Ben-David et al. [BCGL92] proposed a straightforward generalization and gave a definition of T on average for an arbitrary time bound T. Under their definition, when a distribution puts most of the weight on a few instances, so the weights on the rest of the instances are negligible, a function larger than T everywhere could still be T on average. To avoid this, Cai and Selman [CS96] formulated a definition of T on average that requires, for every n, that the expectation over the set $A_n = \{x : |x| \geq n\}$, with respect to the conditional distribution over A_n, be less than or equal to 1. The effect of this is to restrict the rate of convergence of the average sum. It follows from their definition that if $t(x) > T(x)$ for almost all (i.e., all but finitely many) instances x, then t cannot be T on average. This allows them to obtain an hierarchy theorem for arbitrary average-case time-bounds that is independent of distributions and is as tight as the Hartmanis-Stearns hierarchy theorem for worst-case deterministic time [HS65].

For convenience, if the time-bound T is a polynomial, then we will say that a function is *(modified) polynomial on average* if it is T on average in accordance with the definition of Cai and Selman. In this manner we distinguish their formulation from Levin's original notion of p on average. If a distributional problem can be solved by a deterministic algorithm whose running time is (modified) p on average for some polynomial p, we say that the problem is in AVP. (Recall that AP denotes the class of problems that are polynomial on average using Levin's definition.) It is straightforward that AVP is a subset of AP, and Cai and Selman showed that the inclusion is proper. They also proved that if a distributional problem has a distribution that satisfies a reasonable condition that they called *condition W*, then the problem is in AVP exactly when it is in AP.

Now consider reductions on distributional problems. The simplest of these is Levin's deterministic polynomial-time many-one reduction, which satisfies the important property that if distributional problem (A, μ_A) is reducible to distributional problem (B, μ_B) and (B, μ_B) belongs to AP, then (A, μ_A) belongs to AP. It follows from this closure property of the reductions that a complete problem is in AP if and only if every NP problem under every reasonable distribution is in AP.

Observe that if μ_A and μ_B both satisfy condition W, (A, μ_A) is reducible to (B, μ_B), and (B, μ_B) belongs to AVP, then (A, μ_A) belongs to AVP—

because in this case membership in AVP and AP are identical. The main result of this paper shows that AVP fails to satisfy this closure property unless such a restriction is made on the class of admissible distributions. Specifically, we demonstrate distributional problems (A, μ_1) and (A, μ_2) such that the identity function reduces (A, μ_1) to (A, μ_2), (A, μ_2) belongs to AVP, μ_2 even satisfies condition W, but (A, μ_1) does not belong to AVP. Since AVP is closed under standard reductions when distributions satisfy condition W, it would be interesting to know to what extent the lack of a general closure property for AVP matters.

2 Definitions and Background

Let $\Sigma = \{0, 1\}$. We assume that all languages are subsets of Σ^*. Let μ denote a probability distribution (distribution, in short) over Σ^*, i.e., for all x, $\mu(x) \geq 0$ and $\sum_x \mu(x) = 1$. Its distribution function $\mu^*(x) = \sum_{y \leq x} \mu(y)$ is the cumulative distribution of μ up to x, where $\leq$ is the standard lexicographical order. The following definitions are due to Levin [Lev86]. A function ℓ is *linear on μ-average* if the expected value of $\ell(x)/|x|$ with respect to μ is bounded, i.e. if $\sum_x \ell(x)|x|^{-1}\mu(x) < \infty$. A function t is *polynomial on μ-average* if t is a polynomial of a linear on μ-average function. This can be rephrased as follows.

Definition 1 ([Lev86]) Let μ be a distribution on Σ^*, and let $t : \Sigma^* \to \mathbb{N}$. Then t is *polynomial on μ-average* if there exists a positive integer k such that $\sum_x t^{1/k}(x)|x|^{-1}\mu(x) < \infty$.

Several authors have explained the motivations and justifications of this definition, and have explained why seemingly more obvious formulations fail [Lev86, Gur89, Gur91, Ven91, Imp95, Wan96].

Let AP denote the class of all distributional decision problems (A, μ) such that A can be solved by a deterministic algorithm whose running time is polynomial on μ-average.

Let μ and ν be two distributions. Then μ is *dominated* by ν, denoted by $\mu \preceq \nu$, if there is a polynomial p such that for all x, $\mu(x) \leq p(|x|)\nu(x)$.

Let μ_A and μ_B be two distributions and let $f : \Sigma^* \to \Sigma^*$. Recall, for every distribution ν on Σ^*, that f induces a distribution $f(\nu)$ on Σ^* that is defined by $f(\nu)(y) = \sum_{f(x)=y} \nu(x)$, for all $y \in \Sigma^*$. Then, we say that μ_A is *dominated by* μ_B *with respect to* f, denoted by $\mu_A \preceq_f \mu_B$, if there exists a distribution ν such that $\mu_A \preceq \nu$ and for all $y \in \mathrm{range}(f)$, $\mu_B(y) = f(\nu)(y)$.

Levin [Lev86] first defined polynomial-time many–one reducibility, we will use the form given by Gurevich [Gur91]. In what it follows, we will use "p-time" to denote "polynomial-time".

Definition 2 Let (A, μ_A) and (B, μ_B) be two distributional problems. Then (A, μ_A) is *p-time many–one reducible to* (B, μ_B), denoted by $(A, \mu_A) \leq_m^p (B, \mu_B)$, if there exists a p-time computable function $f : \Sigma^* \to \Sigma^*$ such that A is many–one reducible to B via f and $\mu_A \preceq_f \mu_B$.

The following properties are important. Gurevich [Gur91] and Wang [Wan96] provide proofs.

Lemma 1 *1. Let (A, μ_A) and (B, μ_B) be two distributional problems such that $(A, \mu_A) \leq_m^p (B, \mu_B)$. If $(B, \mu_B) \in \mathrm{AP}$, then $(A, \mu_A) \in \mathrm{AP}$.*

2. Polynomial-time many-one reductions are transitive.

Other reductions [VL88, Gur91, BCGL92] between distributional problems have been defined and used to prove average-case completeness results. These reductions are weaker than p-time many-one reductions in that p-time reductions imply the other types of reductions. We will restrict our attention in this paper to p-time many-one reductions.

A real-valued function $r : \Sigma^* \to [0, 1]$ is p-time computable if there exists a deterministic algorithm $\mathcal{A}$ such that for every string x and every positive integer k, $\mathcal{A}$ outputs a finite binary fraction y such that $|r(x) - y| \leq 2^{-k}$ and the running time of $\mathcal{A}$ is polynomially bounded in $|x|$ and k [Ko83]. If μ^* is p-time computable, then so is μ, but Blass showed that the converse is not true unless P = NP [Gur91]. With this fact in mind, we assume throughout that when we say that μ is p-time computable, both μ and μ^* are p-time computable.

Levin [Joh84] hypothesized that any natural distribution μ satisfies the following property: μ is either p-time computable or is dominated by a distribution that is. Let DistNP be the class of distributional problems (A, μ) such that $A \in \mathrm{NP}$ and μ satisfies this property. Then DistNP should include all natural distributional NP problems. A distributional problem is *average-case many-one complete for* NP if it is in DistNP and every other problem in

DistNP is $\leq_m^p$-reducible to it. In this case, we will say, more simply, that the problem is *average-case* NP-*complete*. Levin [Lev86] showed that a distributional tiling problem with a simple distribution on instances is average-case NP-complete. Since then, several other average-case NP-complete problems have been found.

Levin's definition intends only to distinguish between "hard on average" and "easy on average" distributional problems. To make finer distinctions, Ben-David et al. [BCGL92] suggested a straightforward generalization and defined a function t to be T on average if $t(x) = T(\ell(x))$ for some linear on average function ℓ. Letting $T^{-1}(n) = \min\{m : T(m) \geq n\}$, this can be rephrased as follows:

Definition 3 ([BCGL92]) Let μ be a distribution on Σ^*, and let $T : \mathbb{N} \to \mathbb{N}$. Then function $t : \Sigma^* \to \mathbb{N}$ is T *on* μ-*average* if $\sum_x T^{-1}(t(x))|x|^{-1}\mu(x) < \infty$.

Denote by ADTime$(T(n))$ the class of all distributional decision problems that can be decided by a deterministic algorithm in time T on average.

The definition of linear on average does not distinguish between n on average and cn on average for any constant c, so $T(n)$ on average is the same as $T(cn)$ on average for any function T. This sets a limit on how fine a separation of the classes ADTime$(T(n))$ can be achieved, although hierarchy theorems for ADTime$(T(n))$ [BW95, SY96] can still be obtained. Such hierarchy results are bound to be weaker than Hartmanis-Stearns' hierarchy for worst-case deterministic time. On the other hand, it is possible to get a hierarchy theorem as strong as Hartmanis-Stearns' when restricted to certain types of polynomially bounded functions. In particular, Cai and Selman [CS96] showed the following result. Let $t, T : \mathbb{N} \to \mathbb{N}$ be *logarithmico-exponential* functions[2] such that t is bounded above by a polynomial and T is fully time constructible. If $t(n) \log t(n) = o(T(n))$, then ADTime$(t(n)) \subsetneq$ ADTime$(T(n))$.

[2]Hardy [Har11] first defined and studied logarithmico-exponential functions to provide a "scale of infinities". He showed, among other things, that a logarithmico-exponential function cannot increase more slowly than every iterated logarithm function, nor faster than every iterated exponential function, and moreover any logarithmico-exponential function is either eventually positive or eventually negative or identically zero. Despite these restrictions, they include many standard functions.

However, as noted by Cai and Selman [CS96], ADTime(4^n) = ADTime(2^n). We also note that by the almost everywhere hierarchy theorem [GHS87], there exists a language L such that L is not in DTIME(2^n) almost everywhere, but $L \in$ DTIME(4^n), which implies that $(L, \mu) \in$ ADTime(4^n) for every distribution μ. Hence, $(L, \mu) \in$ ADTime(2^n). A language that requires more than $T(n)$ time to compute almost everywhere should also require $T(n)$ time to compute on average. In order to ensure that a function that is greater than $T(n)$ almost everywhere is not considered to be T on average, we would like to distinguish between cn on average for different values of c. We might attempt to do so by defining a function ℓ to be n on average if $\sum_x \ell(x)|x|^{-1}\mu(x) \leq 1$, and so we would be comparing ℓ to $|x|$ rather than $c|x|$ for arbitrary c. A function t could then be said to be T on average if $t(x) = T(\ell(x))$ for some n on average function ℓ, i.e. if $\sum_x T^{-1}(t(x))|x|^{-1}\mu(x) \leq 1$. However, since an algorithm that solves a decision problem can be given arbitrarily large look-up tables, we would again have ADTime(2^n) = ADTime(4^n). So one would like to remove dependency on any finite number of inputs. One way of doing that is to require that the expectation over the set $A_n = \{x : |x| \geq n\}$, with respect to the conditional distribution over A_n, be less than or equal to 1. This gives rise naturally to the following definition of Cai and Selman [CS96].

Definition 4 Let μ be a distribution on Σ^*, and let $W_n = \mu(\{x : |x| \geq n\})$. Let T be a function from $\mathbb{N}$ to $\mathbb{N}$. Then a function $t : \Sigma^* \to \mathbb{N}$ is *(modified) T on μ-average* if for all $n \geq 1$, $\sum_{|x| \geq n} T^{-1}(t(x))|x|^{-1}\mu(x) \leq W_n$.

For convenience, we denote by AVDTime($T(n)$) the class of all distributional decision problems solvable by a deterministic algorithm whose running time is (modified) T on average. The union of AVDTime($p(n)$) over all polynomials p is called AVP.

Since the inclusion of a problem in a time complexity class does not depend on any finite number of instances, the following lemma is obvious.

Lemma 2 *Suppose that (A, μ) is solvable by a deterministic algorithm in time t such that for sufficiently large n, $\sum_{|x| \geq n} T^{-1}(t(x))|x|^{-1}\mu(x) \leq W_n$. Then $(A, \mu) \in$ AVDTime($T(n)$).*

The following result is an easy consequence of definition 4 [CS96]. Let T_1, T_2 be fully time constructible, and (A, μ) a distributional decision problem. If A is in DTime($T_1(n)$), then (A, μ) is in AVDTime($T_1(n)$). If every algorithm that solves A requires more than T_2 time for all but finitely many instances, then (A, μ) is not in AVDTime($T_2(n)$). Cai and Selman then obtained the following hierarchy result.

Theorem 1 ([CS96]) *Let T be fully time-constructible and $t(x) \log t(x) = o(T(x))$. Then there is a language L such that for any distribution μ, $(L, \mu) \in$ AVDTime($T(n)$) $-$ AVDTime($t(n)$).*

3 Main Theorems

Clearly, AVP is contained in AP. While the converse is not true, there is a partial converse. Under a fairly reasonable condition on the distribution μ, a distributional problem is in AP if and only if it is in AVP. A distribution μ is said to satisfy *condition W* if there exists $s > 0$ such that $W_n = \Omega(1/n^s)$.

Lemma 3 ([CS96]) *Let μ be a distribution that satisfies condition W. Then a distributional decision problem (A, μ) is in* AP *iff it is in* AVP*.*

Suppose $(A, \mu_A) \leq_m^p (B, \mu_B)$ and (B, μ_B) is in AVP. Then (B, μ_B) will be in AP, and so by Theorem 1, (A, μ_A) will also be in AP. As long as μ_A satisfies condition W, (A, μ_A) is also in AVP. Unfortunately, if μ_A does not satisfy condition W, then (A, μ_A) does not have to be in AVP.

Theorem 2 *There exists a language A and p-time computable distributions μ, ν such that $(A, \mu) \leq_m^p (A, \nu)$, (A, ν) is in* AVP*, and ν satisfies condition W, but (A, μ) is not in* AVP*.*

Proof. From the almost everywhere hierarchy theorem [GHS87], there exists a language L such that $L \in$ DTIME(2^{2n}), but is not in DTIME(2^n) almost everywhere. Let $A = L \cup \{1^n : n \geq 1\}$. Notice, then, that if M is a Turing machine that accepts A, then for all but finitely many x in $\Sigma^* - \{1^n\}$, $T_M(x) > 2^{|x|}$.

Define the distributions μ and ν by

$$\mu(x) = \frac{1}{4^n},$$

$$\nu(x) = \begin{cases} 1/(n^2 2^n (2^n - 1)) & \text{if } x \neq 1^n, \\ (2^n - 1)/(n^2 2^n) & \text{if } x = 1^n, \end{cases}$$

where $n = |x|$. Then both μ and ν are p-time computable. Since $\nu(\{x : |x| = n\}) = 1/n^2$, ν satisfies condition W.

Next, we show that (A, ν) is in AVP. Since ν satisfies condition W, it is sufficient, by Lemma 3, to show that (A, ν) is in AP. Let M' be a Turing machine that accepts A in time 2^{2n}, and let M be a machine that, on input x, will accept if $x = 1^{|x|}$, and will simulate M' otherwise. Then

$$\begin{aligned}
\sum_x \frac{\sqrt{T_M(x)}}{|x|}\nu(x) &= \sum_{n=1}^{\infty}\left(\sum_{|x|=n, x\neq 1^n} \frac{\sqrt{T_M(x)}}{|x|}\nu(x) + \frac{\sqrt{T_M(1^n)}}{|1^n|}\nu(1^n)\right) \\
&\leq \sum_{n=1}^{\infty} \frac{\sqrt{2^{2n}}}{n} \cdot \frac{1}{n^2 2^n(2^n-1)} \cdot (2^n - 1) + \sum_{n=1}^{\infty} \frac{\sqrt{n}}{n} \cdot \frac{2^n-1}{n^2 2^n} \\
&\leq \sum_{n=1}^{\infty} \frac{1}{n^3} + \sum_{n=1}^{\infty} \frac{1}{n^{5/2}} < \infty
\end{aligned}$$

This proves that (A, ν) is in AVP.

We now show that (A, μ) is not in AVP. Notice, first of all, that $W_n = \mu(\{x : |x| \geq n\}) = \sum_{|x|\geq n} 1/4^{|x|} = \sum_{k=n}^{\infty} 1/2^k$. Let M be a Turing machine that accepts A. Then, by construction of A, for sufficiently large n, $T_M(x) > 2^{|x|}$ if $|x| \geq n$ and $x \neq 1^{|x|}$. So, for any $\epsilon > 0$, if n is sufficiently large,

$$\begin{aligned}
\sum_{|x|\geq n} \frac{T_M^{\epsilon}(x)}{|x|} \cdot \mu(x) &\geq \sum_{k=n}^{\infty}\left(\sum_{|x|=k, x\neq 1^k} \frac{T_M^{\epsilon}(x)}{|x|} \cdot \mu(x)\right) \\
&\geq \sum_{k=n}^{\infty} \frac{2^{\epsilon k}}{k} \cdot \frac{1}{4^k} \cdot (2^k - 1) \\
&> \sum_{k=n}^{\infty} \frac{1}{2^k} = W_n.
\end{aligned}$$

Hence, (A, μ) is not in AVP.

To complete the proof, we show that $(A, \mu) \leq_m^p (A, \nu)$ via the identity function. It suffices to show that $\mu \preceq \nu$. In particular, we show that for all x, $\mu(x) \leq |x|^2\nu(x)$. To see this, we note that if $x \neq 1^{|x|}$, we have

$$\mu(x) = \frac{1}{4^{|x|}} \leq \frac{1}{4^{|x|} - 2^{|x|}} = |x|^2\nu(x),$$

and if $x = 1^{|x|}$, we have

$$\mu(x) = \frac{1}{4^{|x|}} \leq \frac{2^{|x|}-1}{2^{|x|}} = |x|^2\nu(x).$$

This completes the proof. ∎

Corollary 1 ([CS96]) AP *strictly includes* AVP.

One may ask if we do not insist on condition W, or if we do not insist on any specific condition on distributions, what conditions on reductions could ensure that AVP is closed. By Theorem 2, no condition on reductions alone can ensure this. However, it is possible to ensure that AVP is closed if we also impose stronger conditions on what happens to the distributions under the reduction. One such example follows.

Definition 5 Let (B, μ) be a distributional problem, and let Y be a subset of Σ^*. Then the *restriction* of (B, μ) to Y is the distributional problem (B, μ_Y), where

$$\mu_Y(x) = \begin{cases} \mu(x)/\mu(Y) & \text{if } x \in Y, \\ 0 & \text{if } x \notin Y. \end{cases}$$

Theorem 3 *Suppose $f : \Sigma^* \to \Sigma^*$ is one-one and p-time computable, and that $|f|$ is non-decreasing for strings of the same length, and strictly increasing for strings of different length. Suppose $(A, \mu_A) \leq_m^p (B, \mu_B)$ via f with $\mu_B = f(\mu_A)$. Then if the restriction of (B, μ_B) to $Y = \mathrm{range}(f)$ is in* AVP*, (A, μ_A) is also in* AVP.

Proof. Since the restriction of (B, μ_B) to Y is in AVP, there exists a Turing machine that accepts B in time T, and a positive integer k such that for all $n \in \mathbb{Z}^+$,

$$\sum_{\substack{|y| \geq n, \\ y \in Y}} \frac{T^{1/k}(y)}{|y|} \mu_Y(y) \quad \leq \quad \sum_{\substack{|y| \geq n, \\ y \in Y}} \mu_Y(y). \tag{1}$$

This implies that

$$\sum_{\substack{|y| \geq n, \\ y \in Y}} \frac{T^{1/k}(y)}{|y|} \mu_B(y) \quad \leq \quad \sum_{\substack{|y| \geq n, \\ y \in Y}} \mu_B(y). \tag{2}$$

Without loss of generality, we assume that $n = o(T(n))$. Since f is p-time computable, there exists $m > 0$ such that for all x, $|f(x)|^{1/m} \leq |x|$. By the linear speed-up theorem [HS65], we know that there exists a Turing machine accepting B in time $T' = T/2^{km}$. Since f is one-one, $\mu_B(f(x)) = \mu_A(x)$. Since $A \leq_m^p B$ via f, there exists a Turing machine accepting A in time $h' = T' \circ f$.

Let $h = T \circ f$. From (2), we get, for all $n \in \mathbb{Z}^+$,

$$\sum_{|f(x)| \geq n} \frac{h^{1/k}(x)}{|f(x)|} \mu_A(x) \quad \leq \quad \sum_{|f(x)| \geq n} \mu_A(x).$$

If $\frac{h^{1/k}(x)}{|f(x)|} \geq 1$, then $(\frac{h^{1/k}(x)}{|f(x)|})^{1/m} \leq \frac{h^{1/k}(x)}{|f(x)|}$. If $\frac{h^{1/k}(x)}{|f(x)|} < 1$, then $(\frac{h^{1/k}(x)}{|f(x)|})^{1/m} < 1$. So for all positive n, we have,

$$\sum_{|f(x)| \geq n} \frac{h^{1/km}(x)}{|x|} \mu_A(x) \quad \leq \quad 2 \sum_{|f(x)| \geq n} \mu_A(x).$$

This means for all $n > 0$:

$$\sum_{|f(x)| \geq n} \frac{(h')^{1/km}(x)}{|x|} \mu_A(x) \quad \leq \quad \sum_{|f(x)| \geq n} \mu_A(x).$$

Now, because of the monotonicity condition on $|f|$, for any $\ell > 0$, there exists $n_\ell > 0$ such that $\{x : |x| \geq \ell\} = \{x : |f(x)| \geq n_\ell\}$. (Indeed, $n_\ell = |f(0^\ell)|$.) So, we then get, for all $\ell > 0$,

$$\begin{aligned}
\sum_{|x| \geq \ell} \frac{(h')^{1/km}(x)}{|x|} \mu_A(x) &= \sum_{|f(x)| \geq n_\ell} \frac{(h')^{1/km}(x)}{|x|} \mu_A(x) \\
&\leq \sum_{|f(x)| \geq n_\ell} \mu_A(x) \\
&= \sum_{|x| \geq \ell} \mu_A(x).
\end{aligned}$$

This shows that (A, μ_A) is in AVP. ∎

Remark. Notice that if a function $t : \Sigma^* \to \mathbb{R}^+$ is T on average (under Ben-David et al.'s definition), then the restriction of t to any subset $S \subset \Sigma^*$ will also be T on average, i.e. $\sum_{x\in S} T^{-1}(t(x))|x|^{-1}\mu_S(x) < \infty$. This property is important in showing that reductions are closed for AP. However, this property is not true in general for (modified) T on average. So in the above theorem, it is not even enough to assume that (B, μ_B) is in AVP.

4 Concluding Remarks

Levin has provided a useful and robust framework for studying NP problems that are difficult on average. This framework has been enhanced by a number of researchers. The reader is referred to [Wan96] for a survey of this theory. The definition of T on average proposed by Cai and Selman, which modifies an earlier definition given by Ben-David et al., provides a fine separation of average-time complexity classes. In order to use this definition to study algorithmic properties of average-case complexity, the class of admissible distributions needs to be restricted. Such a restriction, e.g., condition W, is often acceptable in practice.

One may perhaps wonder whether there is a feasible way to measure computation time on average that satisfies the needs of studying difficult-on-average NP problems and also provides all the desirable structural properties without restricting the class of admissible distributions to condition W or similar conditions. Ideally, one would like a definition to have the following properties:

1. It would provide the same AP as Levin's definition when it is applied to polynomials.

2. If any algorithm that solves a problem A takes time greater than $T(|x|)$ for almost all x, then for every "well-behaved" distribution μ, (A, μ) would not belong to the class of distributional problems that are solvable in time T on μ-average. (Note: for some rather peculiar distribution μ, (A, μ) could still be in that class.)

3. It would provide a tight hierarchy, at least as tight as the Hartmanis-Stearns' hierarchy for worst-case deterministic time.

In any case, it is important to work on a definition of average-case time that is suitable for studying difficult-on-average NP problems.

Final Remark. This paper is a journal version of a conference paper of Belanger and Wang [BW96]. Pavan subsequently found a simpler proof of Theorem 2, which has the added advantage of using a reduction as simple as the identity function, rather than the more complicated reduction used in the original proof. We would like to thank Alan Selman for telling the authors of each other's work.

Acknowledgment. We are grateful to Alan Selman for a number of constructive comments and suggestions, which helped improve the exposition of the paper.

References

[BW95] J. Belanger and J. Wang. Rankable distributions do not provide harder instances than uniform distributions. In *Proceedings of the 1st Annual International Computing and Combinatorics Conference*, vol. 959 of *Lecture Notes in Computer Science*, Springer-Verlag, pages 410–419, 1995.

[BW96] J. Belanger and J. Wang. Reductions and convergence rates of average time. In *Proceedings of the 2nd Annual International Computing and Combinatorics Conference*, vol. 1090 of *Lecture Notes in Computer Science*, Springer-Verlag, pages 300–309, 1996.

[BCGL92] S. Ben-David, B. Chor, O. Goldreich, and M. Luby. On the theory of average case complexity. *Journal of Computer and System Sciences*, 44:193–219, 1992. (First appeared in In *Proceedings of the 21st Annual Symposium on Theory of Computing*, ACM Press, pages 204–216, 1989.)

[BG95] A. Blass and Y. Gurevich. Matrix transformation is complete for the average case. *SIAM Journal on Computing*, 24:3–29, 1995.

[CS96] J.-Y. Cai and A. Selman. Fine separation of average time complexity classes. In *Proceedings of the 13th Annual Symposium on*

Theoretical Aspects of Computer Science, vol 1046 of *Lecture Notes in Computer Science*, Springer-Verlag, pages 331–343, 1996.

[GHS91] J. Geske, D. Huynh and J. Seiferas. A note on almost-everywhere complex sets with application to polynomial complexity degrees. *Information and Computation*, 92(1):97-104,1991.

[GHS87] J. Geske, D. Huynh and A. Selman A hierarchy theorem for almost everywhere complex sets with application to polynomial complexity degrees. In *Proceedings of the 4th Annual Symposium on Theoretical Aspects of Computer Science*, vol. 247 of *Lecture Notes in Computer Science*, Springer-Verlag, pages 125–135, 1987.

[Gur89] Y. Gurevich. The challenger-solver game: variations on the theme of P =? NP. *Bulletin of the European Association for Theoretical Computer Science*, pages 112–121, 1989. Reprinted in G. Rozenberg and A. Salomaa, editors, *Current Trends in Theoretical Computer Science*, World Scientific, pages 245–253, 1993.

[Gur91] Y. Gurevich. Average case completeness. *Journal of Computer and System Sciences*, 42:346–398, 1991.

[GS87] Y. Gurevich and S. Shelah. Expected Computation Time for Hamiltonian Path Problem. *SIAM Journal on Computing*, 16:486-502, 1987.

[Har11] G. Hardy. Properties of logarithmico-exponential functions. *Proc. London Math. Soc.*, 10:54–90, 1911.

[HS65] J. Hartmanis and R. Stearns. On the computational complexity of algorithms. *Trans. Amer. Math. Soc.*, 117:285–306, 1965.

[Joh84] D. Johnson. The NP-completeness column: an ongoing guide. *Journal of Algorithms*, 5:284–299, 1984.

[Imp95] R. Impagliazzo. A personal view of average-case complexity. In *Proceedings of the 10th Conference on Structure in Complexity Theory*, IEEE Computer Society Press, pages 134–147, 1995.

[Ko83] K. Ko. On the definition of some complexity classes of real numbers. *Mathematical Systems Theory*, 16:95–109, 1983.

[Lev86] L. Levin. Average case complete problems. *SIAM Journal on Computing*, 15:285–286, 1986. (First appeared in *Proceedings of the 16th Symposium on Theory of Computing*, ACM Press, page 465, 1984.)

[SY96] R. Schuler and T. Yamakami. Structural average case complexity. *Journal of Computer and System Sciences*, 52: 308–327, 1996. (First appeared in *Proceedings of the 12th Conference on the Foundations of Software Technology and Theoretical Computer Science*, vol. 652 of *Lecture Notes in Computer Science*, Springer-Verlag, pages 128–139, 1992.)

[Ven91] R. Venkatesan. *Average-Case Intractability*. Ph.D. Thesis, Boston University, 1991.

[VL88] R. Venkatesan and L. Levin. Random instances of a graph coloring problem are hard. In *Proceedings of the 20th Annual Symposium on Theory of Computing*, ACM Press, pages 217–222, 1988.

[VR92] R. Venkatesan and S. Rajagopalan. Average case intractability of Diophantine and matrix problems. In *Proceedings of the 24th Annual Symposium on Theory of Computing*, ACM Press, pages 632–642, 1992.

[Wan95] J. Wang. Average-case completeness of a word problem for groups. In *Proceedings of the 27th Annual Symposium on Theory of Computing*, ACM Press, pages 325–334, 1995.

[Wan96] J. Wang. Average-case computational complexity theory. In L. Hemaspaandra and A. Selman , editors, *Complexity Theory Retrospective II*, Springer-Verlag, to appear in 1996.

[WB93] J. Wang and J. Belanger. On average-P vs. average-NP. In K. Ambos-Spies, S. Homer, and U. Schönings, editors, *Complexity Theory—Current Research*, pages 47–67. Cambridge University Press, 1993. (First appeared in *Proceedings of the 7th Annual*

Conference on Structure in Complexity Theory, IEEE Computer Society Press, pages 318–326, 1992.)

[WB95] J. Wang and J. Belanger. On the NP-isomorphism problem with respect to random instances. *Journal of Computer and System Sciences*, 50:151–164, 1995.