Probabilistic Lower Bounds
For Average Case Complexity

by
A. Sharell and J. Makowsky

Technical Report #746
July 1992
Probabilistic Lower Bounds for Average Case Complexity

Abraham Sharell and Johann Makowsky
Faculty of Computer Science
Technion—Israel Institute of Technology
Haifa, Israel

June 1992

Abstract

We develop a framework to study the definitions of upper and lower bounds commonly employed in the probabilistic analysis of algorithms and the theory of average case complexity. Our goal is to find a suitable definition of average lower bounds which fits both the framework and real applications. For this purpose we exhibit several properties such a definition must satisfy. Among these properties we have: Various transitivity properties, honesty with respect to worst case bounds, consistency with upper average bounds, dichotomy with respect to bounding functions and robustness under small changes of the underlying probability function. We then examine in detail three candidate definitions, bounds on the expectation, bounds with probability 1 and our newly introduced average lower bounds. All of them satisfy some of these properties, but our main result is that only our average lower bounds satisfy all these properties.
# Contents

1 Introduction .................................................. 1

2 Background and definitions ................................. 4
   2.1 Local and global randomization ....................... 4
   2.2 Weight-function ........................................ 5
   2.3 Bounding functions ...................................... 7

3 Probabilistic upper bounds .................................. 8
   3.1 Definitions of probabilistic upper bounds ............ 8
   3.2 Connections between local and global probabilistic upper bounds 10

4 Probabilistic lower bounds .................................. 12
   4.1 Properties of probabilistic lower bounds ............ 12
   4.2 Definitions for probabilistic lower bounds .......... 13

5 Connections between local and global probabilistic lower bound .... 18
   5.1 Transfer theorems where the weight-function is fixed .... 18
   5.2 Transfer theorems, where the weight-function is free ... 22

6 Connections between upper and lower average bounds ............ 24
   6.1 Consistency ............................................. 25
   6.2 Diagonalization over a set of bounding functions ......... 27

7 Robustness of our framework .................................. 29
   7.1 Domination .............................................. 29
   7.2 Choice of weight function and local domination ....... 30
   7.3 Robustness of average bounds .......................... 32
   7.4 Robustness of bounds with probability 1 ............... 36
   7.5 Robustness of bounds on the expectation ............... 38

8 A concluding example ........................................ 38
   8.1 Deriving an average lower bound ....................... 39
   8.2 Changing the bounding function ......................... 39
   8.3 Changing the randomization ............................. 40
   8.4 Summary .................................................. 40

9 Conclusions and further research ............................ 40

A Appendix ................................................................ 42
   A.1 History of the definition of polynomial on the average .. 42
   A.2 Some remarks on regular weight functions ............... 43
   A.3 Some remarks on ‘linear on the average’ ................. 44
   A.4 Alternative definitions for average lower bounds ....... 44
1 Introduction

Central to any complexity theory that interprets the input set as a probability space are the various definitions of probabilistic bounds, i.e., bounds with respect to a randomization of the input-set. We develop a framework to study the definitions of upper and lower bounds commonly employed in the probabilistic analysis of algorithms (exposed in [PB85, Hof87, Kem84] and the theory of Average Case Complexity. Probabilistic upper bounds come in three flavours, bounds on the expectation, bounds with probability 1 and, as developed in Average Case Complexity, average bounds.

Average Case Complexity theory was initiated by the seminal paper of Levin ([Lev86]) and further developed in [Gur91, Gol88, BG90, BCGL92, IL90] among others. The purpose of Average Case Complexity theory is to establish a framework analogous to the theory of NP-completeness for the complexity of problems with respect to a probability function on the input set. As in Worst Case Complexity, easy problems are those which have a polynomial average upper bound, hard problems are those which are complete for some non-polynomial randomized complexity class, say DistNP, i.e., problems such that every problem in DistNP is polynomially reducible to them. In Worst Case Complexity theory, very hard problems are those which have a non-polynomial lower bound, and hierarchies of such problems are studied. In Average Case Complexity theory the issue of average lower bounds was so far not developed, and even for average upper bounds the discussion is mostly restricted to polynomial bounds.

We propose a suitable definition for average lower bounds and show that it complements exactly the existing definition for corresponding average upper bounds. To discuss the notion of probabilistic lower bounds in more depth, we propose several properties which allow us to evaluate candidate definitions. The properties for probabilistic lower bounds we discuss are: Various transitivity properties, honesty with respect to worst case bounds, consistency with probabilistic upper bounds, dichotomy with respect to bounding functions and robustness under small changes of the underlying probability function. To formalize the notion of small changes we use the notion of domination between probability functions as used in Average Case Complexity theory. We then examine in detail three candidate definitions of lower bounds: bounds on the expectation, bounds with probability 1 and our newly introduced average lower bounds. All of them satisfy some of these properties, but our main result is that only our average lower bounds satisfy all these properties. We also study the interrelationship between the various
definitions. In particular, we show which lower bounds with probability 1 imply which average lower bounds.

In detail, the paper is organized as follows: In section 2 we introduce local and global randomizations and the notion of a weight-function, which establishes the link between them. We also introduce the notion of bounding functions needed in the sequel. In section 3, we present the definitions of various probabilistic upper bounds and the connections between them. The main result in this section is theorem 1, which states that an upper bound on the expectation implies an average upper bound. The result is not really new, but is stated in a more general way and its proof is considerably shortened by the use of Jensen's inequality. In section 4 we first introduce the properties of probabilistic lower bounds which serve as our guideline. We then define various probabilistic lower bounds and discuss these properties. The main results in this section are theorems 2 and 3, which state that the main defect of lower bounds on the expectation and lower bounds with probability 1 is the absence of the dichotomy property, but that both are honest. Here we also state our main result concerning average lower bounds (Theorem 4), namely that they satisfy all the properties mentioned. The proof is split into several parts, the last of which being the Consistency Theorem (theorem 7) in section 6. In section 5 we analyze the connection between the various lower bounds. Here, the main results are theorem 6 and proposition 5.4. They state the conditions on lower bounds with probability 1, which allow us to infer average lower bounds. As a corollary we get that average lower bounds are honest (proposition 5.5). In section 6 we analyze the connections between lower and upper bounds. Theorem 7 states that average lower bounds satisfy the consistency property. This completes the proof that only our average lower bounds satisfy all the properties we had postulated. The last result on this section is theorem 8 shows how to pass from the non-existence of average upper bounds to the existence of sharp average lower bounds. In section 7, we review the notion of domination and use it to analyze robustness under the choice of weight-functions and robustness of probabilistic bounds. Theorems 10 and 11 show exactly in which way each bound is preserved under domination. The former extends a result previously known for polynomial bounds. The latter gives a sufficient condition for bounds with probability 1 to be preserved under domination and is new. In section 9 we draw conclusions and discuss further research. The appendices contain various remarks and elaborations, whose inclusion in the main text would have been distracting.
Acknowledgments: We are indebted to S. Ben-David, O. Goldreich, Y. Gurevich and M. Luby for fruitful discussions. The results of this paper are part of the M.Sc. Thesis [Sha92] of the first author written under the guidance of the second author. The first author also expresses his gratitude to B. Chor, S. Ben-David and O. Goldreich from whose lectures he profited much, and who were also willing to listen. Thanks also to A. Tal, Y. Bar-Guri and G. Even for long talks and discussions.

The second author was supported by the Fund for the Promotion of Research of the Technion-Israel Institute of Technology.

We would like to thank M. Hofri for his challenging bewilderment about the average behaviour of SAT, which he expressed personally and in chapter 1 of [Hof87]. This paper might be an lengthy answer to his skeptical questions.
2 Background and definitions

Passing from worst case to average case analysis of algorithms and problem we have to define a probability function over the inputs. We shall present two approaches: local randomizations and global randomizations. The instrument to connect local and global randomizations is the weight-function, and this is discussed in subsection 2.2. The last subsection (2.3) formulates some restrictions on the sets of functions, we consider admissible as bounding functions.

2.1 Local and global randomization

What we call local randomizations is a formalization of the standard approach employed in the probabilistic analysis of algorithms. For every size \( n \) a probability function on the (finite) set of inputs of size \( n \) is defined separately. The local randomization consists of the sequence of this functions. Global randomizations are simpler — they consist of a single probability function defined on all the inputs. This is the approach taken in the theory of average case complexity.

A side issue of this paper is to provide for ‘translations’ of results between those contexts, so we have to consider both approaches. In many cases the difference between those two approaches of randomizing the input-set is rather technical. For details see the next subsection and subsection A.2 in the appendix. Some of the techniques for translating have been used in previous work but we present here a rigorous and detailed presentation of (hopefully most of) the relevant issues.

As preliminaries we need the notion of an input set, a size associated with each input and a function that assigns to each input a probability. Throughout this work we will usually denote the set of inputs by \( S \). Formally let \( S \) be any countable (or finite) set. A size function on \( S \) is a function \( |\cdot| : S \to N^+ \) such that the set \( S_n = \{ x \in S : |x| = n \} \) is finite. The set of inputs together with the size function gives:

**Definition 2.1 (Input set)** An input set \( S \) is a pair \( < S, |\cdot| > \), where \( S \) is a countable set and \( |\cdot| \) a size function on \( S \).

A probability function (pf) on \( S \) is a function \( \mu : S \to [0,1] \) so that \( \Sigma_{x \in S} \mu(x) = 1 \). We will freely make use of functions that sum to some constant other than 1 as pf's. Assuming implicit normalization whenever necessary. Now we are ready to define local and global randomizations.
Definitions 2.2 (Local and global randomization)

(i) Let $S$ be an input set and let $S_n \subseteq S$ be the set of inputs of size $n$. The local approach gives a probability function $\mu_n$ for inputs of each fixed size $n$. The sequence $<S_n, \mu_n>$ is called a local randomization of $S$.

(ii) Let $S$ be an input set and $\mu$ a pf on $S$. The pair $<S, \mu>$ is called a global randomization of $S$.

Traditional probabilistic analysis of algorithms uses local randomizations $<S_n, \mu_n>$. Statements are usually of the form that for all (sufficiently large) numbers $n <S_n, \mu_n>$ satisfies some conditions.

In the theory of average case complexity distributional problems are defined as decision (or search) problems paired with a global randomization on the inputs. The main concern is to explore classes of distributional problems, their hierarchies and interreductions.

Definition 2.3 (Distributional problems) Let $<S, \mu>$ be a global randomization and $D$ be a subset of $S$. The pair $<D, \mu>$ is called a distributional problem. We think of the subset of inputs defined by $D$ as the set of positive instances of some problem.

2.2 Weight-function

A global randomization can be decomposed into:

(i) a local randomization $<S_n, \mu_n>$, where $\mu_n(x) = Pr\{x | x \in S_n\}$

and

(ii) a weight-function $w(n) = Pr\{S_n\}$.

Given a local randomization $<S_n, \mu_n>$ there is no unique global randomization $<S_n, \mu_n>$ that decomposes into $<S_n, \mu_n>$. To get a global randomization from a local one we have to supply the weight-function which is a pf on $N^+$ according to which the input-length is chosen. For the development of this work it will make sense to restrict somewhat the weight-functions that appear implicit from a global randomization or used in combination with a local randomization to produce a global one. We now define those restrictions and then give examples for weight functions and explain the intuitive reasons behind the restrictions.

Definitions 2.4 (Weight-functions)
(i) We call a pf $w$ on $N^+$ a weight function.

(ii) If for some constant $c > 0$ and for every $n \in N^+$ $w(n) \geq n^{-c}$ then we say that $w$ is a regular weight function.

(iii) If for every $\varepsilon > 0$:
$$\sum w(n)n^\varepsilon = \infty$$
then we say that $w$ is a strongly regular weight-function.

Examples 2.1 The most commonly used regular weight-function is $w(n) = n^{-2}$. As an example for a strongly regular weight-function take $w(n) = n^{-1}(\log n)^{-2}$.

Definition 2.5 (Induced Global Randomization) Let $S$ be an input set, $<S_n, \mu_n>$ a local randomization of $S$ and $w$ a weight function. The unique global randomization induced by $w$ and $<S_n, \mu_n>$ on $S$ is $<S, \mu>$, where $\mu$ is defined by $\mu(x) = w(|x|)\mu_{|x|}(x)$.

We call $<S, \mu>$ (strongly) regular if $w$ is (strongly) regular.

We have three reasons to restrict ourselves to regular weight-functions. First of all in most contexts one would not want short problem instances to influence the average case analysis significantly. If the weight function decreases too fast however (to be interpreted as non-regular for the case of polynomial average bounds) too much probability is concentrated on short instances and this might happen. Secondly we would like that for any global randomization $<S, \mu>$ a function $T$ on $S$, that depends only on the input-size and is not bounded by a polynomial, is not polynomial on the average with respect to $<S, \mu>$. We show in proposition A.1 in the appendix, that this does not hold for non-regular global randomizations The last reason is that the restriction to regular weight functions allows us to state results in a way that does not depend on the specific weight function. Proposition 3.2 for example, defines polynomial on the average directly on local randomizations, i.e. independent from any weight function. This definition is equivalent to the one on global randomizations if those are restricted to be regular.

The reasons for strongly regular weight function are somewhat less compelling, but still significant. The strongly regular weight function $w(n) = n^{-1}(\log n)^{-2}$ has been proposed in [Gur90]. This choice can be justified as follows. Often the pf $\mu_0(x) = 2^{-|x|/|x|^2}$ is taken as the standard pf on binary strings. If one considers the natural numbers $N^+$ coded as
binary strings then \( \mu_0 \) defines a (also standard) pf on \( \mathbb{N}^+ \), that is, a weight function \( w_0 \) which is given (up to a constant) by:

\[
   w_0(n) = \frac{\mu_0(\text{str}(n))}{n(\log n)^2}
\]

since the binary representation of \( n \), denoted by \( \text{str}(n) \), has length (approximately) \( \log n \). In [Gur90] the choice of this weight function (instead of \( n^{-2} \)) is crucial. In our paper restriction to strongly regular weight functions leads sometimes to an easier formulation of results, especially at the end of subsection 4.

2.3 Bounding functions

Before we go on to define various ways in which a function \( f \) may be a bound in some probabilistic sense, we will slightly constrain the set of functions acceptable as a bounding function \( f \).

Definition 2.6 (Bounding Function) Let \( \mathcal{B}_f \subseteq \{ f: \mathbb{R}^+ \rightarrow \mathbb{R}^+ \} \). We say that \( \mathcal{B}_f \) is a set of bounding functions if it fulfills the following conditions:

(i) Every \( f \in \mathcal{B}_f \) is a continuous strictly increasing function.

(ii) No \( f \in \mathcal{B}_f \) is bounded. Together with the first condition this gives \( \lim_{a \to \infty} f(a) = \infty \).

(iii) For every \( f, g \in \mathcal{B}_f \):

\[
   \lim_{a \to \infty} \frac{f(a)}{g(a)} = \begin{cases} 
   \infty & \text{denoted by } f \succ g \\
   c & 0 < c < \infty & \text{denoted by } f \asymp g \\
   0 & \text{denoted by } f \prec g 
   \end{cases}
\]

And no other possibility exists. \( f \preceq g \) means that either \( f \asymp g \) or \( f \prec g \).

Remark 2.2 (Notation) We denote by \( \mathcal{O}(\mathcal{B}_f) \) the class of bounding functions \( g \) for which there exists \( f \in \mathcal{B}_f \) so that \( g = \mathcal{O}(f) \).

Examples 2.3 (Examples of sets of bounding functions) A natural example of a set of bounding functions are the logarithmic-exponential functions presented in [Har24]. Those are all finite compositions of the functions \( +, -, \times, /, \log, \exp \) and real constants, applied only to real values. The sets of bounding functions that appear in this work are:
(i) \( \text{Log} = \{ f(a) = c \log a : c \in \mathbb{R}^+ \} \)

(ii) \( \text{Poly} = \{ f(a) = a^k : k \in \mathbb{N} \} \)

(iii) \( \text{Rat} = \{ f(a) = a^c : c \in \mathbb{R}^+ \} \)

(iv) \( \text{SubExp} = \{ f(a) = 2^a : 0 < \varepsilon < 1 \} \)

(v) \( \text{Exp} = \{ f(a) = 2^{ca} : c \in \mathbb{R}^+ \} \)

(vi) \( \text{GExp} = \{ f(a) = 2^{ac} : c \in \mathbb{R}^+ \} \)

The restrictions imposed are quite severe. Nevertheless, any standard expression usually employed as a bounding function is included. For most of this work it would have been sufficient to have \( f : \mathbb{N}^+ \rightarrow \mathbb{N}^+ \), increasing and for two bounding functions \( f \) and \( g \) either \( f = O(g) \) or \( g = O(f) \). But the additional restrictions allow sometimes for more concise statements as well as the explicit use of the inverse function.

3 Probabilistic upper bounds

In this section we will define various types of probabilistic upper bounds on a function \( T \), measuring some resource like run-time or storage-size used by an algorithm on an input-set \( S \). For the purpose of this work we will ignore the source of the function and regard it formally as a function of type \( T : S \rightarrow \mathbb{R}^+ \).

3.1 Definitions of probabilistic upper bounds

With respect to a local randomization we may define the expectation of \( T \) on inputs of size \( n \) by

\[
E_{\mu_n}(T) = \sum_{x \in S_n} T(x) \mu_n(x),
\]

giving a sequence \( E_{\mu_n} : \mathbb{N}^+ \rightarrow \mathbb{R}^+ \). One possibility to define local probabilistic bounds is to define bounds on this sequence. The other is to request that a bound on \( T \) holds with asymptotic probability one.

Definitions 3.1 (Local Probabilistic Upper Bounds) Let \( < S_n, \mu_n > \) be a local randomization on \( S \) and let \( f : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \).
(i) We say that f is an upper bound on the expectation of T if
\[ \forall n \in N^+ : E_{\mu_n}(T) \leq f(n) \]

(ii) f is a (local) upper bound with probability 1 on T if
\[ \lim_{n \to \infty} Pr_{\mu_n}\{T(x) \leq f(n)\} = 1. \]

Results in probabilistic analysis of algorithms are usually expressed with this two types of local bounds on T.

In average case complexity theory upper bounds with respect to global randomizations are defined as follows:

Definitions 3.2 (Average Upper Bounds (AUB)) Let \(<S, \mu>\) be a global randomization on S and T : S → R^+.

(i) We say that T is linear on the average with respect to the global randomization if \(E_{\mu}\left(\frac{T(x)}{|x|}\right)\) is finite. The set of all linear on the average functions T will be denoted by \(\mathcal{L}(<S, \mu>)\) or by \(\mathcal{L}\) when the randomization is understood from context.

(ii) For f : R^+ → R^+ we say that T is at most f on the average w.r.t. the global randomization if there exists \(\ell \in \mathcal{L}\) so that:
\[ \forall x \in S : T(x) \leq f(\ell(x)). \]

The set of all at most f on the average functions T will be denoted by \(\text{AUB}(<S, \mu>, f)\) or by \(\text{AUB}(f)\) when the randomization is understood from context.

(iii) If for some set of functions \(\mathcal{F} \subset \{R^+ \to R^+\}\) there is f ∈ \(\mathcal{F}\) such that T is at most f on the average we say that T is (at most) \(\mathcal{F}\) on the average and denote this by \(T \in \text{AUB}(<S, \mu>, \mathcal{F})\). In the case of \(\mathcal{F} = \text{Poly}\) we say simply that T is polynomial on the average and write \(T \in \text{AUB}(<S, \mu>, \text{Poly})\). In all this cases we will feel free to omit the randomization if it is evident from the context.

For some remarks about the historical development of this definition and further references see the appendix. Here we point only point out that this definition is robust under composition with polynomials, whereas the
more natural seeming definition, formulated by the condition $E_{\mu_n}(T(x)) = O(\text{Poly})$, is not.

In this work we will often employ a more convenient condition for $T$ to be at most $f$ on the average, which is equivalent to definition 3.2 for strictly increasing functions $f$.

Lemma 3.1 (AUB(f) for $f$ strictly increasing.) Let $f : R^+ \rightarrow R^+$ be strictly increasing, then $T \in \text{AUB}(f)$ iff $f^{-1}(T(x))$ is linear on the average.

Proof: If $T \in \text{AUB}(f)$ then for some linear on the average function $\ell$ and for all $x \in S$ we have $T(x) \leq f(\ell(x))$. $f$ is strictly increasing so $f^{-1}$ exists and is strictly increasing too, and for all $x \in S$ $f^{-1}(T(x)) \leq \ell(x)$, which implies that $f^{-1}(T(x))$ is linear on the average. In the other direction assume that $f^{-1}(T(x))$ is linear on the average and take $\ell(x) = f^{-1}(T(x))$. Then for all $x \in S$, $T(x) \leq f(\ell(x))$ holds of course. 

Note that for bounding functions as defined in subsection 2.3 the assumption that $f$ is strictly increasing is true and we may freely use the above condition as a necessary and sufficient one.

3.2 Connections between local and global probabilistic upper bounds

Another interesting case for which there is a necessary and sufficient condition is when the randomization is regular and the bounding function is polynomially bounded.

Proposition 3.2 (Gurevich [Gur90]) Let $<S_n, \mu_n>$ be a local randomization, $w$ a regular weight-function. Let $<S, \mu>$ be induced by $w$ and $<S_n, \mu_n>$. Then $T : S \rightarrow N^+$ is AUB($<S, \mu>$, Poly) iff for some constant $\varepsilon > 0$

$$E_{\mu_n}(T(x)^\varepsilon) = O(n).$$

Proof: By assumption on $w$ there is $c > 0$ such that $w(n) \geq n^{-c}$. If $T \in \text{AUB}(<S, \mu>, \text{Poly})$ then for some constant $\alpha > 0$ the expectation of $(T(x)\alpha/|x|)$ is finite. Writing this expectation as

$$\sum_{n \in N^+} n^{-1}w(n)E_{\mu_n}(T(x)^\alpha)$$

10
and substituting \( w(n) \geq n^{-c} \) gives
\[
E_{\mu_n}(T(x)\alpha) \leq n^{c+1} \quad \text{(almost always)}.
\]

For \( \varepsilon = \frac{\alpha}{c+1} \) Jensen’s inequality (which states that for any \( \text{r.v. } X \) with finite expectation and for a convex function \( u \) \( u(E(X)) \leq E(u(X)) \); cf. [Bre68, Man86]) gives us
\[
E_{\mu_n}(T(x)^\varepsilon(n+1)) \leq E_{\mu_n}(T(x)\alpha).
\]

Combining the last two inequalities we get
\[
E_{\mu_n}(T(x)^\varepsilon(n+1)) \leq n \quad \text{(almost always)},
\]
which finishes the proof of the first direction.

The second direction is always true (also without the restriction on \( w \)) and follows immediately from the definitions.

In works on this subject it is usually stated that upper bounds on \( E_{\mu_n}(T) \) imply average upper bounds on \( T \). We refer to statements that connect local and upper bounds as transfer theorems.

Theorem 1 (Transfer theorem for upper bounds) For any strictly increasing convex function \( f : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \) if for every sufficiently large \( n \in \mathbb{N}^+ : \)
\[
E_{\mu_n}(T(x)) \leq f(n)
\]
thен \( T \in \text{AUB}(f) \).

Proof: Observe that by applying first Jensen’s inequality (cf. [Bre68, Man86]) and then the hypothesis for every sufficiently large \( n \):
\[
f(E_{\mu_n}(f^{-1}(T(x)))) \leq E_{\mu_n}(T(x)) \leq f(n).
\]
Since \( f \) is strictly increasing this gives for almost all \( n \):
\[
E_{\mu_n}(f^{-1}(T(x))) \leq n.
\]
4 Probabilistic lower bounds

In this section we define lower bounds in a similar way as in subsection 3. The lower bounds counterparts of the probabilistic (local) upper bounds are quite obvious and we collect them in definition 4.2.

In average case complexity theory there is so far no definition of lower bounds. Average upper bounds (3.2) can be thought of as an average analog to (non-strict) upper bounds as denoted usually by the big $O$ notation. We propose a definition (4.3) for average lower bounds that formalizes the notion of 'T is greater than f on the average'. With other words an average analog to 'T = $\omega(f)$'. To compare and evaluate candidate definitions, we propose some properties probabilistic lower bounds should posers.

4.1 Properties of probabilistic lower bounds

Intuitively we would like probabilistic lower bounds to preserve the transitivity properties of (worst case) strict lower bounds, as adapted to the probabilistic context (Transitivity-1,2 in the list below). Also a worst case strict lower bound should imply an probabilistic lower bound, that is probabilistic bounds should be honest with regard to worst case bounds (Honesty). Probabilistic lower bounds should be consistent with their analog definition for probabilistic upper bounds (Consistency). Finally probabilistic lower bounds should be as weak as possible while satisfying all other conditions. Ideal would be that for each function $T$ and bounding function $f$, $f$ is either a probabilistic lower bound or a probabilistic upper bound on $T$ (Dichotomy).

To state this relations succinctly let us denote with $T >_{plb} f$ that $f$ is a probabilistic lower bound on $T$. Similarly denote with $T \leq_{plb} f$ that $f$ is a probabilistic upper bound on $T$. Remember that $<$ is our order relation on bounding functions. In the following let $S$ be an input set and $BF$ a set of bounding functions as defined in subsection 2.3.

Definition 4.1 (Properties of probabilistic lower bounds.)

**Transitivity-1** For every $T : S \rightarrow R^+$ and bounding functions $f, g \in BF$, if $T >_{plb} f$ and $f \geq g$ then $T >_{plb} g$.

**Weak Transitivity-1** For every $T : S \rightarrow R^+$ and bounding functions $f, g \in BF$, if $T >_{plb} f$ and $f > g$ then $T >_{plb} g$. 

12
Transitivity—2 For every $T_1, T_2 : S \rightarrow R^+$ and $f \in BF$, if $T_1 >_{plb} f$ and $T_2(x) \geq T_1(x)$ for all but a finite number of $x$ then $T_2 >_{plb} f$.

Honesty For every $T : S \rightarrow R^+$ and $f \in BF$, if for every $k \in N^+$ and for all but a finite number of $x$ $T(x) > f(k|x)$ then $T >_{plb} f$ with respect to any local randomization on $S$ (excluding trivial randomizations that accord positive probability to only a finite number of inputs).

Consistency For every $T : S \rightarrow R^+$ and $f \in BF$, if $T >_{plb} f$ and $T \leq_{plb} g$ then $f < g$.

Weak Consistency For every $T : S \rightarrow R^+$ and $f \in BF$, if $T >_{plb} f$ and $T \leq_{plb} g$ then $f \leq g$.

Dichotomy For every $T : S \rightarrow R^+$ and $f \in BF$, either $T >_{plb} f$ or $T \leq_{plb} f$.

4.2 Definitions for probabilistic lower bounds

We now give the definitions of the various probabilistic lower bounds and check which of the properties they satisfy.

Definitions 4.2 (Local Probabilistic Lower Bounds) Let $S$ be an input set with a local randomization $<S_n, \mu_n>$ defined on it. Let $T : S \rightarrow N^+$ and $f : N^+ \rightarrow N^+$ be two functions.

(i) We say that $f$ is an (local) lower bound on the expectation of $T$ if

$$E_{\mu_n}(T) > f(n)$$

(ii) $f$ is a (local) lower bound with probability 1 on $T$ if

$$\lim_{n \rightarrow \infty} Pr_{\mu_n}\{T(x) > f(n)\} = 1.$$ 

Lower bounds with probability 1 appear in the literature of probabilistic analysis. Lower bounds on the expectation are not commonly used — they just convey too little information. See the comment preceding proposition 5.1.

The following definition for average lower bounds is new, and most of this article is motivated by justifying this definition and integrating it with the already existing definitions for probabilistic upper and lower bounds.
Definition 4.3 (Average Lower Bounds (ALB)) Let \( <\mathcal{S}, \mu> \) be a global randomization and let \( T : S \rightarrow N^+ \) and \( f : N^+ \rightarrow N^+ \) be two functions.

(i) We say that \( T \) is greater than \( f \) on the average (or equivalently that \( f \) is an average lower bound for \( T \)) if there exists a function \( u : S \rightarrow N^+ \) that is not linear on the average so that

\[
\forall x \in S \colon T(x) \geq f(u(x)).
\]

Similar to the upper bound case we denote this by \( \text{ALB}(<\mathcal{S}, \mu>, f) \) or simply \( \text{ALB}(f) \).

(ii) If for some some set of functions \( \mathcal{F} \subseteq \{R^+ \rightarrow R^+\} \) there is \( f \in \mathcal{F} \) so that \( T \in \text{ALB}(<\mathcal{S}, \mu>, f) \) we write \( T \in \text{ALB}(<\mathcal{S}, \mu>, \mathcal{F}) \).

We now derive a convenient condition for \( T \) to be greater than \( f \) on the average, which is equivalent for strictly increasing functions (lemma 4.1).

For the rest of this subsection let \( <\mathcal{S}, \mu> \) be the global randomization induced by \( <\mathcal{S}_n, \mu_n> \) and \( w \).

Lemma 4.1 (ALB(\( f \)) for \( f \) strictly increasing.) Let \( f : R^+ \rightarrow R^+ \) be strictly increasing. Then \( T \in \text{ALB}(f) \) iff \( f^{-1}(T(x)) \) is not linear on the average.

Proof: Exactly symmetric to the proof of the corresponding lemma (3.1) for upper bounds.

Note that for bounding functions as defined in subsection 2.3 the assumption is true and we may freely use the above condition as a necessary and sufficient one. In the following let \( \text{BF} \) be a set of bounding functions, \( <\mathcal{S}_n, \mu_n> \) be a local randomization and \( <\mathcal{S}, \mu> \) a global randomization.

Theorem 2 (Properties of lower bounds on the expectation.)

For \( T : S \rightarrow R^+ \) and a bounding function \( f \in \text{BF} \) interpret \( T \leq_{plb} f \) as

\[
E_{\mu_n}(T(x)) = O(f)
\]

and interpret \( T >_{plb} f \) as

\[
E_{\mu_n}(T(x)) = \Omega(f).
\]

Then properties (Weak) Transitivity–1, Transitivity–2, Honesty and Weak Consistency are satisfied, and properties Consistency and Dichotomy are not.
Proposition 4.2 (Transitivity of average lower bounds) For $T : S \rightarrow R^+$ and a bounding function $f$ interpret $T \leq_{plb} f$ as $T \in AUB(f)$ and interpret $T >_{plb} f$ as $T \in ALB(f)$. Then (Weak) Transitivity-1, and Transitivity-2 are satisfied.

Proof: (Weak) Transitivity-1 We have to prove that if $T$ is greater than $f$ on the average and $f \geq g$ then $T$ is greater than $g$ on the average. It is sufficient to prove the property for the linear case, where $f$ is the identity function. Assume that $T$ is greater than linear on the average and that $c, n_0$ are positive constants so that for every $n \geq n_0$ the inequality $g(n) < cn$ holds. Define

\[
u(x) = \begin{cases} 
    T(x)/c & \text{if } T(x) \geq cn_0 \\
    g^{-1}(T(x)/2) & \text{else}
\end{cases}
\]

It can be easily verified that for all $x$ $T(x) > g(\nu(x))$ and that $\nu$ is not linear on the average.

Transitivity-2 We have to prove that if $T_1$ is greater than $f$ on the average and for almost all but a finite number of $x$ $T_2(x) \geq T_1(x)$ then $T_2$ is greater than $f$ on the average.

The proof is immediate and we omit it.

Remark 4.3 Note that if we replace in Transitivity-2 the condition that $T_2 \geq T_1$ for almost all inputs with the condition

\[
\lim_{n \to \infty} Pr_{\mu_n} \{T_2(x) \geq T_1(x)\} = 1
\]

then the definition no longer satisfies Transitivity-2. To see this note that for a function to be greater than linear on the average it is sufficient for it to have very large values on a subset of inputs $A$ so that $Pr_{\mu_n}\{A\}$ converges to zero.
Proposition 4.4 (Dichotomy of average bounds) Let $<S, \mu>$ be a randomization and $f \in BF$. Then $\text{AUB}(<S, \mu>, f)$ and $\text{ALB}(<S, \mu>, f)$ partition the set of functions $\{T : S \to R^+\}$ into two disjoint subsets.

Proof: This follows immediately from lemmata 3.1 and 4.1 since for every $T : S \to R^+$ the expectation

$$E_\mu \left( \frac{f^{-1}(T(x))}{|x|} \right)$$

either converges in which case $T \in \text{AUB}(<S, \mu>, f)$ or diverges in which case $T \in \text{ALB}(<S, \mu>, f)$.

The proof that the remaining two properties are satisfied by average lower bounds is delegated to subsections 5 and 6. For property Honesty this is done in proposition 5.5, and for (Weak) Consistency in the Consistency Theorem (theorem 7).

5 Connections between local and global probabilistic lower bound

As in 3 we call statements that connect local and global probabilistic bounds transfer theorems. Let $BF$ be a set of bounding functions as defined in section 2.3.

5.1 Transfer theorems where the weight-function is fixed

First of all similar to Theorem 1 we have a transfer theorem for concave functions $f$.

Theorem 5 If $f \in BF$ is concave and $T$ is greater than $f$ on the average then $E_{\mu_n}(T) \geq f(n)$.

Proof: Exactly like for theorem 1, taking into account that $f^{-1}$ is convex.

This theorem might be useful in the context of average-log-space (see [BCGL92]). Since in the context of average-log-space the concern is with upper and not with lower bounds it is actually the theorem's contranegation that might be useful. Imagine a situation where it is easily shown that
$E_{\mu_n}(T) \geq \log n$ is not true. Then $\log n$ can not be an average lower bound on $T$ and, as we will see in section 6, this implies that $T$ is at most logarithmic on the average.

A lower bound on $E_{\mu_n}(T)$ does not imply any other kind of lower bound (of those we have defined in this subsection) on $T$, since a set of inputs with vanishing measure but on which $T$ is very large can cause $E_{\mu_n}(T)$ to become arbitrary large. The opposite direction as formulated in the next proposition is of course true and is a direct consequence of Markov's inequality. (We omit the proof.)

**Proposition 5.1** If $f : N^+ \to N^+$ is a lower bound with probability 1 on $T$ then for every $0 < \varepsilon < 1$ there is a constant $c$ so that $E_{\mu_n}(T) > \varepsilon f(n) + c$.

A similar but much more interesting result holds for average lower bounds:

**Theorem 6 (Transfer theorem for lower bounds)** Let $<S, \mu>$ be induced by $<S_n, \mu_n>$ and $w$ and let $f, g$ be two bounding functions in $BF$. If $f$ is a lower bound with probability 1 on $T$ w.r.t. $<S_n, \mu_n>$ and $g$ is sufficiently small for

$$\sum_{n=1}^{\infty} \frac{w(n)}{n} g^{-1}(f(n)) = \infty$$

to hold then

$T \in ALB(g)$.

**Proof:** Since $g$ is strictly increasing is suffices to show that

$$E_{\mu_n}(\frac{g^{-1}(T(x))}{|x|}) = \infty.$$ 

Applying Markov's inequality to the (strictly positive) random variable $g^{-1}(T(x))$ we derive for every $n \in N^+$:

$$E_{\mu_n}(g^{-1}(T(x))) > g^{-1}(f(n)) \Pr_{\mu_n}\{g^{-1}(T(x)) > g^{-1}(f(n))\}$$

$$= g^{-1}(f(n)) \Pr_{\mu_n}\{T(x) > f(n)\}.$$

Observing that

$$E_{\mu_n}(\frac{g^{-1}(T(x))}{|x|}) = \sum_{n=1}^{\infty} \frac{w(n)}{n} E_{\mu_n}(g^{-1}(T(x)))$$

$$> \sum_{n=1}^{\infty} \frac{w(n)}{n} g^{-1}(f(n)) \Pr_{\mu_n}\{T(x) > f(n)\}$$
together with the assumptions in the hypothesis gives the desired result.

Note that the condition that \( f \) is a lower bound with probability 1 on \( T \), i.e.

\[
\lim_{n \to \infty} \Pr_{\mu_n} \{ T(x) > f(n) \} = 1
\]

in theorem 6 is actually too stringent. It can be replaced with the weaker condition that \( \Pr_{\mu_n} \{ T(x) > f(n) \} \) is bounded away from zero.

The condition on \( g \) implied by the divergence of the sum can not be relaxed. To see this let \( g, f \in BF \) such that

\[
\sum_{n=1}^{\infty} \frac{w(n)}{n} g^{-1}(f(n)) < \infty.
\]

Then there exists a function \( T : S \to R^+ \) so that \( f \) is a lower bound with probability 1 on \( T \) but \( g \) is not an average lower bound on \( T \). As a concrete example take

\[
T(x) = g(2g^{-1}(f(|x|)))
\]

(where 2 can be replaced by any constant greater than 1).

The last theorem gives us a way to obtain average lower bounds for \( T \) that are very close to, but strictly smaller than \( f \). Of course it is possible that larger functions as those obtained through the theorem (\( f \) itself for example) are average lower bounds for \( T \). But this does not mean that the condition in the theorem is poor, but that \( f \) is not a sharp lower bound with probability 1 for \( T \) to begin with. How close to \( f \) we can obtain an average lower bound depends on \( w \): the larger \( w \) is, the closer. The following corollary makes this explicit. We consider only regular and strongly regular weight functions since for them the results can be formulated in a more natural manner then for general weight functions. Also this is general enough for our purpose.

**Corollary 5.2 (Corollary to the transfer theorem)** Let \( <S, \mu> \) be induced by \( <S_n, \mu_n> \) and a weight function \( w \). Let \( f : R^+ \to R^+ \) be a bounding function and assume that \( f \) is a lower bound with probability 1 on \( T \) w.r.t. \( <S_n, \mu_n> \).

1. If \( w \) is regular then there exists \( 0 < \varepsilon < 1 \) so that \( T \in \text{ALB}(f(n^\varepsilon)) \).
2. If \( w \) is strongly regular then for every \( 0 < \varepsilon < 1 \) we have \( T \in \text{ALB}(f(n^\varepsilon)) \).
**Proof:** For regular \( w \) let \( c > 1 \) be a constant so that for all \( n \in N^+ \):
\[
w(n) \geq n^{-c}.
\]
Set \( \epsilon = \frac{1}{c} \) and \( g(n) = f(n^\epsilon) \). Then \( g^{-1}(f(n)) = n^\epsilon \) and
\[
\sum_{n=1}^{\infty} \frac{w(n)}{n} g^{-1}(f(n)) = \sum_{n=1}^{\infty} \frac{1}{n^\epsilon} = \infty.
\]
By theorem 6 we conclude that \( T \in \text{ALB}(g) \).

For strongly regular weight functions let \( 0 < \epsilon < 1 \) and \( g(n) = f(n^\epsilon) \). Then the general term in the above sum evaluates to \( w(n)n^{(\frac{1}{\epsilon}-1)} \) and by the definition of strongly regular the sum diverges.

For strongly regular function it can be shown that the statement in the corollary is optimal. That is if \( h > n^\epsilon \) for every \( 0 < \epsilon < 1 \) then \( f(h(n)) \) is not an average lower bound on \( T \) any more. For regular weight functions it is possible to derive slightly larger average lower bounds. For example, with everything as in the first part of the proof, \( f(n^\epsilon \log n) \) is still an average lower bound for \( T \). But the real limitation in increasing the average lower bound is the \( n^\epsilon \) term, which is impossible to increase and which depends only on the weight function. (If \( w(n) = n^{-c} \) then \( \epsilon \) has to be smaller or equal to \( \frac{1}{2} \), otherwise the sum will converge.) So the basic result, that it is sufficient to take a root of the argument before applying \( f \) to derive an average lower bound, remains the same.

We formulate now the results of this subsection in terms of classes of bounding functions. The following list also serves as a set of examples how the transfer theorem and its corollary can be employed.

**Examples 5.3 (Application of the transfer theorem)** Let \( <S, \mu> \) be induced by \( <S_n, \mu_n> \) and a weight function \( w \). In the context of this proposition all lower bounds with probability 1 are w.r.t \( <S_n, \mu_n> \) and all average lower bounds are w.r.t. \( <S, \mu> \).

(i) Let BF be any set of bounding functions (Log, Poly, Rat, SubExp, Exp, GExp or other) that does not have a maximal member. If \( w \) is regular and every \( f \in \text{BF} \) is a bound with probability 1 on \( T \) then every \( f \in \text{BF} \) is also an average lower bound on \( T \).

(ii) If \( w \) is regular and \( T \) has a lower bound with probability 1 in one of the classes Log, Rat, SubExp or GExp w.r.t. \( <S_n, \mu_n> \) then \( T \) has an
average lower bound with probability 1 in the same class of bounding functions.

(iii) If \( w \) is strongly regular and \( T \) has a lower bound with probability 1 that is in \( \text{Poly} \) with degree greater or equal to 2 then \( T \) has an average lower bound in \( \text{Poly} \).

(iv) If \( w \) is regular and \( T \) has a lower bound with probability 1 in \( \text{Exp} \) then \( T \) has an average lower bound in \( \text{SubExp} \). If \( w \) is strongly regular then every function in \( \text{SubExp} \) is an average lower bound on \( T \).

5.2 Transfer theorems, where the weight-function is free

If we are free to choose the weight function we can derive average lower bounds that are very close to the bound with probability 1. To prove this all we have to do is to construct a sufficiently large weight function, which is exactly what we do in the following proposition.

Proposition 5.4 (Close average-lower bounds.) Let \( f, g \in \text{BF} \) be two bounding functions so that:

\[
\lim_{n \to \infty} \frac{g^{-1}(f(n))}{n} = \infty.
\]

If \( f \) is a lower bound with probability 1 on \( T \) w.r.t. \( <S_n, \mu_n> \) then there exists a weight function \( w \) so that \( g \) is an average lower bound on \( T \) w.r.t. the global randomization \( <S, \mu> \) that is induced by \( <S_n, \mu_n> \) and \( w \).

Proof: Define a new function \( \delta \) on \( a \in R^+ \) by:

\[
\delta(a) = \frac{g^{-1}(f(a))}{a}.
\]

By theorem 6 it is sufficient to construct a weight function \( w \) so that

\[
\sum_{n=1}^{\infty} w(n)\delta(n) = \infty.
\]

We assume without loss of generality that \( \delta \) is strictly positive and differentiable in \( R^+ \). Denote the derivative by \( \delta' \) and define \( w \) by:

\[
w(n) = \delta'(n)\delta(n)^{-2}.
\]
To verify that \( w \) is indeed a weight function we have to show that its sum converges. This can be done by bounding the sum with an appropriate integral as follows:

\[
\sum_{n=2}^{\infty} w(n) \leq \int_{1}^{\infty} \delta'(a)\delta(a)^{-2} \, da = \frac{1}{\delta(1)}.
\]

On the other hand the sum over \( w(n)\delta(n) \) diverges:

\[
\sum_{n=1}^{\infty} w(n)\delta(n) \geq \int_{1}^{\infty} \delta'(a)\delta(a)^{-1} \, da = [\ln \delta(a)]_{1}^{\infty} = \infty.
\]

For the method of bounding a sum from above and below by integrals to be applicable, some further conditions are needed. For example it is sufficient that the derivative of the function \( \delta'(a)(\delta(a))^{-2} \) exists and is negative. The construction in the proof can be modified so as to guarantee this, but we do not present the more elaborate construction.

The Honesty property for average lower bounds can be proved with a similar construction.

**Proposition 5.5 (Honesty of average lower bounds)** For \( T : S \rightarrow R^+ \) and a bounding function \( f \) interpret \( T \leq_{plb} f \) as \( T \in AUB(f) \) and interpret \( T >_{plb} f \) as \( T \in ALB(f) \). Assume that for every \( k \in N^+ \) and for all but a finite number of inputs \( x \in S \) \( T(x) > f(k|x|) \) holds. Then there exists a weight function \( w \) so that for every global non-trivial randomization \( \langle S, \mu \rangle \) of \( S \) with weight-function \( w \) \( T \in ALB(\langle S, \mu \rangle, f) \).

**Proof:** We will prove the proposition only for \( f \) strictly increasing. In this case \( f^{-1} \) exists. Define

\[
T_n = \min \{ T(x) : x \in S_n \} \quad \text{and} \quad \delta_n = \frac{f^{-1}(T_n)}{n}.
\]

By the hypothesis \( \lim_{n \to \infty} \delta_n = \infty \). It is possible to interpolate \( \delta_n \) to a function \( \delta : R^+ \rightarrow R^+ \) that is twice differentiable and is appropriate for the construction of the weight function as in the proof of proposition 5.4. That is \( w \) defined by:

\[
w(n) = \delta'(n)\delta(n)^{-2}
\]
is indeed a weight function and

\[ \sum_{n=1}^{\infty} w(n)\delta(n) = \infty. \]

Now let \( \langle S_n, \mu_n \rangle \) be any local randomization on \( S \) and \( \langle S, \mu \rangle \) the global randomization induced by \( \langle S_n, \mu_n \rangle \) and \( w \). To complete the proof note that:

\[
E_{\mu} \left( \frac{f^{-1}(T(x))}{|x|} \right) = \sum_{n=1}^{\infty} w(n)E_{\mu_n} \left( \frac{f^{-1}(T(x))}{n} \right)
\]

\[
\geq \sum_{n=1}^{\infty} w(n)\frac{f^{-1}(T_n)}{n} = \sum_{n=1}^{\infty} w(n)\delta(n) = \infty
\]

and therefore \( T \in ALB(\langle S, \mu \rangle, f) \).

Proposition 5.4 can be used to derive statements similar to the one in the list of examples at the end of the last subsection.

**Examples 5.6** By proposition 5.4 the following conditions on the bounding functions \( f, g \in BF \) imply that if \( f \) is a lower bound with probability 1 on \( T \) w.r.t. the local randomization \( \langle S_n, \mu_n \rangle \) then there exists a global randomization \( \langle S, \mu \rangle \) induced by \( \langle S_n, \mu_n \rangle \) (and some weight function) so that \( g \) is a average lower bound on \( T \).

(i) \( g < f \) and \( f, g \in O(Poly) \).

(ii) \( f \in Exp \) and \( g(n) = 2^{\Omega(n)} \).

Note that if \( g < f \) but \( f, g \notin O(Poly) \) then \( g^{-1}(f(n))/n \) might be bounded.

**6 Connections between upper and lower average bounds**

In this subsection we show some connections between upper and lower average bounds.
6.1 Consistency

Intuitively we would like to make statements of the form: If \( f \) is a probabilistic upper bound for \( T \) and \( g \) is a probabilistic lower bound for \( T \) then there should be a way in which \( f \) can be said to be larger than \( g \). This is exactly the statement formalized by property \((\text{Weak})\) Consistency in definition 4.1. We have already treated the case of bounds on the expectation (theorem 2 and of bounds with probability 1 (theorem 3). Here we treat the remaining case of average bounds.

In proposition 4.4 we showed that average bounds satisfy the Dichotomy property. This shows that 3.2 and 4.3 are complementary in the sense that for a bounding function \( f \in \text{AUB}(f) \) and \( \text{ALB}(f) \) are complements with respect to the set of all functions \( T : S \to R^+ \).

From this it follows that if \( f > g \) then \( \text{AUB}(<S,\mu>,f) \cap \text{ALB}(<S,\mu>,g) = \emptyset \). On the other hand if for some functions this intersection is not empty then \( f \) has to increase faster than \( g \). This is the content of the next lemma and theorem.

Using only the fact that \( f \) and \( g \) are strictly increasing we can prove the following:

**Lemma 6.1** Let \( f, g \) be strictly increasing functions of type \( N^+ \to N^+ \) so that:

\[
\text{AUB}(f) \cap \text{ALB}(g) \neq \emptyset.
\]

(i) There are \( u \not\in \mathcal{L} \), \( \ell \in \mathcal{L} \) so that

\[
\forall x \in S \quad g(u(x)) \leq f(|x|).
\]

(ii) Let

\[
A = \{ n \in N^+ : f(n) < g(n) \} = \{ n_1, n_2, n_3, ... \} \quad n_1 < n_2 < n_3 ... .
\]

Then either \( A \) is finite or \( \frac{n_{i+1}}{n_i} \) is not bounded.

(iii) For every \( k \in N^+ \) infinitely often

\[
f(n) > g(kn).
\]

**Proof:** (i) If \( T \in \text{AUB}(f) \cap \text{ALB}(g) \) then there are \( u \not\in \mathcal{L} \), \( \ell \in \mathcal{L} \) so that

\[
\forall x \in S : \quad g(u(x)) \leq T(x) \leq f(\ell(x)).
\]
(ii) Assume by contradiction that \( A \) is not finite and that for some \( c > 0 \):

\[
\forall i \in \mathbb{N}^+ : \frac{n_i + 1}{n_i} \leq c .
\]

Observe that for the above \( \ell \) and \( u \):

\[
\forall x \in S : [\ell(x), u(x)] \cap A = \emptyset ,
\]

since otherwise for some \( x \in S \) and some \( n \in A \):

\[
f(\ell(x)) \leq f(n) < g(n) \leq g(u(x)) ,
\]

which violates the choice of \( \ell \) and \( u \). Now denote for \( x \in S \):

\[
i(x) = \max\{i \in \mathbb{N}^+ : n_i \leq \ell(x)\} .
\]

Since \( A \) is not finite \( n_{i(x)+1} \) always exists. For every \( x \in S \) for which \( \ell(x) \leq u(x) \) we get the following ordering:

\[
n_{i(x)} \leq \ell(x) \leq u(x) \leq n_{i(x)+1} ,
\]

because \( n_{i(x)} \leq \ell(x) < n_{i(x)+1} \) by the definition of \( i(x) \) and \( u(x) \leq n_{i(x)+1} \) because a member of \( A \) cannot be between \( \ell(x) \) and \( u(x) \). The above implies

\[
\forall x \in S : \frac{u(x)}{\ell(x)} \leq c ,
\]

a contradiction.

(iii) Assume by contradiction that for some \( k \in \mathbb{N}^+ \) and all \( n \in \mathbb{N}^+ \)

\[
f(n) \leq g(kn) .
\]

There has to be \( x \in S \) so that \( u(x) > k\ell(x) \) and so we derive:

\[
g(u(x)) > g(k\ell(x)) \geq f(\ell(x)) ,
\]

a contradiction.

The above lemma by itself is not strong enough to show that \( f \) is larger than \( g \) in some significant sense. For functions that increase faster than \( \log \) the last part shows at least that \( f \not\in O(g) \). In any case we get \( f \not\in g \) from the lemma which is sufficient for our purpose if we restrict ourselves to bounding functions as defined in subsection 2.3, since they obey the well-ordering condition.
Theorem 7 (Consistency of \( AUB(f) \) and \( ALB(g) \)) Let \( f, g \in BF \) so that
\[
AUB(f) \cap ALB(g) \neq \emptyset.
\]
Then \( f > g \).

Proof: By lemma 6.1 \( f \nless g \). If \( f \less g \) then it is easy to see that \( AUB(f) = AUB(g) \) and by proposition 4.4 this implies \( AUB(f) \cap ALB(g) = \emptyset \) which contradicts the assumption. By the total-ordering condition on \( BF \) the only possibility left is \( f > g \).

This theorem completes our task of showing that average lower bounds satisfy all the properties listed in 4.1 as stated in theorem 4.

6.2 Diagonalization over a set of bounding functions

Until now we have tried to use the existence of some lower bound to prove the non-existence of upper bounds (or vice versa). One could reverse this and ask if the non-existence of some upper bounds imply the existence of a lower bound. The next theorem is an example of such a statement. We formulate and proof the theorem for polynomial bounds, but it can be generalized quite easily as stated in remark 6.2.

From the Dichotomy property of average bounds (4.4), we can conclude that if a function \( T \) is not in \( AUB(Poly) \) then every polynomial is an average lower bound for \( T \). Using Diagonalization over Poly we can construct a function that increases faster than any polynomial, and still is an average lower bound on \( T \).

Theorem 8 If \( T \notin AUB(Poly) \) then there is a monotone increasing sequence \( g : N^+ \rightarrow N^+ \) so that

(i) \( \forall f \in Poly : f < g \)

(ii) \( T \in ALB(g) \).

Proof: Define for \( k \in N^+ \):
\[
u_k(x) = \lfloor T(x)^k \rfloor.
\]
By assumption for every \( k \in N^+ \) the function \( u_k \) is not linear on the average, which implies that for every \( m_0 \in N^+ \):
\[
\lim_{m \to \infty} \sum_{n=m_0}^{m} w(n)n^{-1} E_{\mu_n}(u_k(x)) = \infty.
\]
So the following inductive sequence is well defined:

\[ n_0 = 0 \]
\[ n_k = \min\{m \in \mathbb{N}^+ : \sum_{n_{(k-1)} < n \leq m} w(n)n^{-1}E_{\mu_n}(u_k(x)) \geq 1\}. \]

(The choice of 1 is not critical. Any positive constant will do, or even a value that depends on \( k \), like \( 1/k \), as long as it’s sum diverges.) Note that the sequence is strictly increasing and that the segments \( \{(n_{(k-1)}, n_k) : k \in \mathbb{N}^+\} \) partition \( \mathbb{N}^+ \).

For \( x \in S \) let \( \text{seg}(x) \) be the unique number \( k \in \mathbb{N}^+ \) so that \( |x| \in (n_{(k-1)}, n_k] \). Clearly \( \text{seg} \) is a function \( S \rightarrow \mathbb{N}^+ \).

For the last part of our construction define \( u : S \rightarrow \mathbb{R}^+ \) and a sequence \( k(m) \):

\[ u(x) = u_{\text{seg}(x)}(x) \]
\[ k(m) = \min\{\text{seg}(x) : u(x) \geq m\}. \]

Since \( \text{seg} \) is a function \( u(x) \) and \( k(m) \) are well defined. Now we can define \( g \) by:

\[ g(m) = mk(m). \]

Note that by definition of \( k(m) \) both, \( k(m) \) and \( g(m) \), are monotone increasing. To complete the proof we prove the following three claims:

Claim 1: \( u \notin \mathcal{L} \)

Claim 2: \( \forall k \in \mathbb{N}^+ : \lim_{m \to \infty} \frac{g(m)}{m^k} = \infty \)

Claim 3: \( \forall x \in S : T(x) \geq g(u(x)) \).

**Proof of claim 1:** By the definition of \( u \) and the choice of \( n_k \):

\[ E_{\mu_n} \left( \frac{u(x)}{|x|} \right) = \sum_{k=1}^{\infty} \sum_{n_{k-1} < n \leq n_k} \frac{w(n)}{n} E_{\mu_n}(u_k(x)) \geq \sum_{k=1}^{\infty} 1 = \infty. \]

**Proof of claim 2:** Since we now already that \( k(m) \) is monotone increasing it suffices to show that \( k(m) \) is not bounded. Let \( k \in \mathbb{N}^+ \). By the definition of size-functions there are only finitely many inputs \( x \in S \) with \( |x| \leq n_k \) so we may define \( M \) by

\[ M = \max\{u(x) : |x| \leq n_k\} + 1. \]

If \( u(x) \geq M \) then \( |x| > n_k \) which is true if and only if \( \text{seg}(x) > k \) and this implies that \( k(M) > k \).
Proof of claim 3: Let $x \in S$ and let $k = \text{seg}(x)$. Then

$$u(x) = u_k(x) = \lfloor T(x)^{\frac{1}{k}} \rfloor.$$ 

Also by definition $k(u(x)) \leq k$. So:

$$g(u(x)) = u(x)^{k(u(x))} = (\lfloor T(x)^{\frac{1}{k}} \rfloor)^{k(u(x))} \leq T(x)^{\frac{k(u(x))}{k}} \leq T(x).$$

Remark 6.2 To construct such a proof for another class $BF$ of bounding functions all we need is a set $A \subset BF$ so that

(i) $A = \{f_k : k \in \mathbb{N}^+\}, f_1 < f_2 < f_3 < ...$

(ii) $\forall g \in A \exists f \in BF : g \preceq f$

(iii) $\forall f \in BF \exists g \in A : g \succeq f$

In this way we may derive similar statements for Exp or $GExp$.

7 Robustness of our framework

7.1 Domination

In average case complexity theory various types of reductions are used to study closure properties of classes of distributional problems. This encompasses many-to-one and Turing reductions, deterministic or randomizing (in this context randomizing means that the algorithm has access to a random source of bits, independent of the input). In average case complexity, as in worst case complexity, the conditions on those reductions are chosen to guarantee that if problem $\Pi_1$ is reducible to problem $\Pi_2$, then problem $\Pi_1$ is not harder to solve than problem $\Pi_2$. In addition to the efficiency and correctness conditions that are placed on reductions in worst case complexity, we have to demand that the reduction does not map problem-instances with a high probability (in the randomization of $\Pi_1$) to those with a much lower probability (in the randomization of $\Pi_2$). Otherwise an average upper bound for $\Pi_2$ does not necessarily imply a similar bound for $\Pi_1$. This restriction is usually referred to as domination. To be more precise domination is a relation between global randomizations on the same input set (see...
7.1. The domination condition for reductions is formulated by associating with the reduction two randomizations on the same input set (either the domain or the range of the reduction) and requesting domination to hold for those two (see definition 7.3, the comment following it and remark 7.5). We present here the definition for domination developed in [BG90], where it is shown that this definition is a sufficient and necessary condition for polynomial average bounds to be preserved by reductions (this is generalized in 10).

Definitions 7.1 (Domination) Let $S$ be an input set and $<S, \mu>$ and $<S, \nu>$ be two global randomizations on $S$.

(i) We say that $<S, \mu>$ dominates $<S, \nu>$ and write $<S, \mu> \leq <S, \nu>$, if the function $r(x) = \frac{\mu(x)}{\nu(x)}$ is polynomial on the average w.r.t. $<S, \mu>$. If the input-set is evident then we will abbreviate and write $\mu \leq \nu$.

(ii) We say that $<S, \mu>$ is $d$-similar to $<S, \nu>$ and write $<S, \mu> \sim <S, \nu>$ if they dominate each other. As before we may abbreviate this by $\mu \sim \nu$.

Proposition 7.1 ([BCGL92, Gur91]) Domination is a partial pre-order on the pf's on $S$. Hence $d$-similarity is an equivalence relation.

7.2 Choice of weight function and local domination

In the context of polynomial probabilistic bounds regular weight functions are especially appropriate as shown by proposition A.1. As we have seen in proposition 3.2 they also allow for a convenient equivalent condition for a function to be polynomial on the average. We now show that if we restrict ourselves to regular weight functions then all the global randomizations induced from the same local randomization are $d$-similar. It is not surprising that the weight-function of the randomization does not really matter (as long as it is regular) and $d$-similarity gives us a way to formulate this rigorously.

Theorem 9 (Uniqueness Theorem) Let $S$ be a input set and let $<S_n, \mu_n>$ be a local randomization. Furthermore, let $w_1, w_2$ be two regular weight functions and $\nu_1, \nu_2$ be the pf's induced by $w_1$ and $w_2$ respectively. Then $\nu_1 \sim \nu_2$. 

30
Proof: We will show that \( \nu_1 \leq \nu_2 \), the other direction is symmetric. Let \( c > 0 \) so that \( w_2(n) \geq n^{-c} \) and let \( \nu_1 \) and \( \nu_2 \) be as defined above. By definition
\[
\frac{\nu_1(x)}{\nu_2(x)} = \frac{w_1(|x|)}{w_2(|x|)} < |x|^c.
\]
Which implies that \( \nu_1 \leq \nu_2 \).

It is natural to look for a notion of domination which applies to local randomizations. The key to such a definition is proposition 3.2 that gives a characterization of polynomial on the average for local randomizations.

Definition 7.2 (Domination for local randomizations)
Let \( <S_n, \mu_n> \) and let \( <S_n, \nu_n> \) be two local randomizations on the same input set \( S \). We say that \( <S_n, \nu_n> \) dominates \( <S_n, \mu_n> \) if there exists a constant \( \varepsilon > 0 \) so that:
\[
E_{\mu_n} \left( \left( \frac{\mu_n(x)}{\nu_n(x)} \right)^\varepsilon \right) = O(n).
\]

As can be expected our 'local' definition of domination coincides with the 'global' one for regular weight functions.

Proposition 7.2 (i) Let \( <S, \nu> \) and \( <S, \mu> \) be induced by \( <S_n, \nu_n> \) and \( <S_n, \mu_n> \) respectively and a weight function \( w \). If \( w \) is regular then \( <S_n, \nu_n> \) dominates \( <S_n, \mu_n> \) iff \( <S, \nu> \) dominates \( <S, \mu> \).

(ii) Let \( <S_n, \nu_n> \) \( <S_n, \mu_n> \) be two local randomizations. \( <S_n, \nu_n> \) dominates \( <S_n, \mu_n> \) iff for every regular weight function \( w \) \( <S, \nu> \) dominates \( <S, \mu> \), where \( <S, \nu> \) and \( <S, \mu> \) are induced by \( w \) and \( <S_n, \nu_n> \) and \( <S_n, \mu_n> \) respectively.

Proof: (i) Denote
\[
r(x) = \frac{\mu(x)}{\nu(x)} = \frac{\mu(|x|)}{\nu(|x|)}.
\]
The 'only if' direction is true also without the restriction to regular global randomizations. Assume that \( <S_n, \nu_n> \) dominates \( <S_n, \mu_n> \). Then there exists \( \varepsilon > 0 \)
\[
E_{\mu_n} (r(x)^\varepsilon) = O(n),
\]

31
which implies
\[ E_\mu(r(x)^\varepsilon) < \infty \]
and so \(<S, \nu>\) dominates \(<S, \mu>\).

For the 'if' direction assume that \(<S, \nu>\) dominates \(<S, \mu>\). This means that \(r(x)\) is at most polynomial on the average w.r.t. \(<S, \mu>\).

Since \(<S, \mu>\) is regular proposition 3.2 implies that there exists \(\varepsilon > 0\) so that:
\[ E_{\mu_n}(r(x)^\varepsilon) = O(n) , \]
which means that \(<S_n, \nu_n>\) dominates \(<S_n, \mu_n>\).

(ii) Use the first part, taking into account that all the global randomizations induced by regular weight functions and the same local randomization are d-similar \((9\) and that d-similarity is an equivalence relation \((7.1)\).

7.3 Robustness of average bounds

We now explore the relations between probabilistic bounds and domination. The main question we ask is what conditions are necessary and sufficient to guarantee that a bound is preserved under changes of the randomization. Since domination was defined in the framework of average case complexity especially to preserve probabilistic polynomial (global) upper bounds we begin with a generalization of this case.

The proof for the sufficient condition to preserve average-upper-bounds under domination is given in lemma 7.3 and is essentially a generalization of the corresponding one in [Gur91].

Lemma 7.3 (Preservation of average bounds under domination)

Let \(<S, \mu>, <S, \nu>\) be global randomizations and let \(f, g : R^+ \rightarrow R^+\) be strictly increasing. If \(\varepsilon > 0\) is a constant so that:
\[ E_\mu \left( \left( \frac{\mu(x)}{\nu(x)} \right)^\varepsilon \right) < \infty \]
and for all \(a > 0\):
\[ g(a) \geq f(a^{1+\varepsilon}) \]

then:
\[ \text{AUB}(<S, \nu>, f) \subseteq \text{AUB}(<S, \mu>, g). \]
Proof: To simplify notation we set:

\[ r(x) = \frac{\mu(x)}{\nu(x)} \quad h(a) = \frac{f^{-1}(a)}{g^{-1}(a)} \quad \text{and} \quad X = \{ x \in S : r(x) \leq h(T(x)) \}. \]

For \( x \in X \) we observe that

\[ g^{-1}(T(x)) \mu(x) \leq f^{-1}(T(x)) \nu(x) \]

and use \( T \in \text{AUB}(\mathcal{S}, \nu, f) \) as follows (ignoring constant factors that arise from the conditional probabilities):

\[
E_{\mu} \left( \frac{g^{-1}(T(x))}{|x|} \middle| x \in X \right) = \sum_{x \in X} g^{-1}(T(x))|x|^{-1} \mu(x) \\
\leq \sum_{x \in X} f^{-1}(T(x))|x|^{-1} \nu(x) \leq E_{\nu} \left( \frac{f^{-1}(T(x))}{|x|} \right) < \infty.
\]

For \( x \not\in X \) by definition:

\[ f^{-1}(T(x)) < g^{-1}(T(x)) r(x) \]

Furthermore the choice of \( g \) implies :

\[ \left( g^{-1}(a) \right)^{1+\epsilon} \leq \left( f^{-1}(a) \right)^{\epsilon}. \]

Substituting \( a = T(x) \) and combining the last two inequalities we obtain:

\[ \left( g^{-1}(T(x)) \right)^{1+\epsilon} < \left( g^{-1}(T(x)) \right)^{\epsilon} r(x)^{\epsilon} \Rightarrow g^{-1}(T(x)) < r(x)^{\epsilon}. \]

Now we can use domination to finish the proof (again ignoring constant factors):

\[ E_{\mu} \left( \frac{g^{-1}(T(x))}{|x|} \middle| x \not\in X \right) = \sum_{x \not\in X} g^{-1}(T(x))|x|^{-1} \mu(x) \leq \sum_{x \not\in X} r(x)^{\epsilon}|x|^{-1} \mu(x) < \infty. \]

If we restrict the bounding functions as defined in subsection 2.3 we can give a simpler formulation:
Corollary 7.4 (Preservation of average bounds under domination)

Let $<S,\mu>, <S,\nu>$ be global randomizations and $f,g$ two bounding functions so that $f,g \leq \text{Poly}$. If $<S,\mu> \leq <S,\nu>$ and $f < g$ then

$$\text{AUB}(<S,\nu>, f) \subseteq \text{AUB}(<S,\mu>, g).$$

Proof: Note that for $f < g \leq \text{Poly}$ there exists $\epsilon_0 > 0$ so that for every $0 < \epsilon \leq \epsilon_0$

$$g(a) \geq f(a^{1+\epsilon}).$$

holds for all sufficiently large $a$. Choosing $\epsilon \leq \epsilon_0$ small enough for

$$E_\mu \left( \left( \frac{\mu(x)}{\nu(x)} \right)^\epsilon \right) < \infty$$

to hold, together with lemma 7.3 gives the desired result.

We formulate and prove now the necessary condition. (In essence it has been observed by S. Ben-David and is also stated in [BG90] for the polynomial case.)

Lemma 7.5 Let $BF$ be a set of bounding-functions so that for all $f,g \in BF (g^{-1} \circ f) \leq \text{Poly}$. Let $<S,\mu>$ and $<S,\nu>$ be two global randomizations on the same input-set. If

$$\text{AUB}(<S,\nu>, BF) \subseteq \text{AUB}(<S,\nu>, BF)$$

then $<S,\nu>$ dominates $<S,\mu>$.  

Proof: As before we set $r(x) = \frac{\mu(x)}{\nu(x)}$. Clearly $r(x)$ is linear on the average w.r.t. $<S,\nu>$. Let $g \in BF$ then by definition $g(r(x)) \in \text{AUB}(<S,\nu>, BF)$ and by the inclusion we assumed as hypothesis we derive:

$$g(r(x)) \in \text{AUB}(<S,\mu>, BF).$$

This means that there exists $f \in BF$ so that $f^{-1}(g(r(x)))$ is linear on the average w.r.t. $<S,\mu>$. By the hypothesis on $BF$

$$(f^{-1} \circ g)^{-1} = g^{-1} \circ f \leq \text{Poly}$$

and we can conclude that $r(x) \in \text{AUB}(<S,\mu>, \text{Poly})$.
Now we are ready to pack everything together and state the main theorem concerning domination for (global) average bounds.

**Theorem 10 (Preservation of average bounds under domination)**

Let \(<S, \mu>\), \(<S, \nu>\) be global randomizations then the following conditions are equivalent:

(i) \(<S, \nu>\) dominates \(<S, \mu>\)

(ii) \(\text{AUB}(<S, \nu>, \text{Poly}) \subseteq \text{AUB}(<S, \mu>, \text{Poly})\)

(iii) \(\text{AUB}(<S, \nu>, \text{GExp}) \subseteq \text{AUB}(<S, \mu>, \text{GExp})\)

(iv) \(\text{AUB}(<S, \nu>, \text{Log}) \subseteq \text{AUB}(<S, \mu>, \text{Log})\)

(v) \(\text{ALB}(<S, \nu>, \text{Poly}) \supseteq \text{ALB}(<S, \mu>, \text{Poly})\)

(vi) \(\text{ALB}(<S, \nu>, \text{GExp}) \supseteq \text{ALB}(<S, \mu>, \text{GExp})\)

(vii) \(\text{ALB}(<S, \nu>, \text{Log}) \supseteq \text{ALB}(<S, \mu>, \text{Log})\)

**Proof:** Let \(\mathcal{F}\) be either one of Poly, GExp, Log. To prove inclusion note that \(\mathcal{F}\) is closed under composition with polynomials. That is for every \(n \in \mathbb{N}^+\) and \(f \in \mathcal{F}\) also \(f(a^n) \in \mathcal{F}\). So we can use lemma 7.3. For the average lower bounds we simply use the set-complement of the corresponding statement for average upper bounds, as justified by the Dichotomy property (cf. proposition 4.4). To prove domination we note that \(\mathcal{F}\) fulfills the condition of lemma 7.5.

We present here the domination condition for reductions for immediate reference but do not pursue this direction in more detail since just now we are only interested in the behavior of bounds for a particular run-time (or other resource measure) function \(T\), and not of the run-time function that takes the time for the reduction into account. The definition we give here is taken from [BG90].

**Definition 7.3** Let \(<A, \mu>\) and \(<B, \nu>\) be two global randomizations on the input sets \(A\) and \(B\) respectively. Let \(f : A \rightarrow B\). We say that \(<B, \nu>\)

\(f\)-dominates \(<A, \mu>\) if the function on \(A\) defined by the ratio

\[
\text{Pr}_\mu\{f^{-1}(f(x))\} \quad \text{Pr}_\nu\{f(x)\}
\]
is polynomial on the average w.r.t $\langle A, \mu \rangle$. This relation is denoted by $\langle A, \mu \rangle \leq_f \langle B, \nu \rangle$.

To paraphrase this in terms of domination define two randomizations on $A$ by

$$
\mu_f(x) = \Pr_\mu \{ f^{-1}(f(x)) \} \\
\nu_f(x) = \Pr_\nu \{ f(x) \} = \nu(f(x)).
$$

Then definition 7.3 amounts to $\langle A, \mu \rangle \leq_f \langle B, \nu \rangle$ iff $\langle A, \mu_f \rangle \leq \langle B, \nu_f \rangle$. It can be shown ([BG90]) that for functions $f \in \text{AUB}(\langle A, \mu \rangle, \text{Poly})$ the above is a necessary and sufficient condition to guarantee that for every function $T$ in $\text{AUB}(\langle B, \nu \rangle, \text{Poly})$ the composition $T(f(x))$ is in $\text{AUB}(\langle A, \mu \rangle, \text{Poly})$. For a comparison of this definition to the one in [Lev86] the reader is referred to [BG90, Gur91].

### 7.4 Robustness of bounds with probability 1

In the proof of lemma 7.3 we compensated the change in the randomization by enlarging the bounding function. This method cannot work for bounds with probability 1, since all that is required there is that the probability that $T$ exceeds the upper bound, condition on the input-size, converges to zero. There is no information about how much $T$ exceeds the bounds. (This is not strictly true, since normally one considers problems in, say NP, but this does not seem to be always helpful for the problem at hand.)

It turns out that if the convergence rate is fast enough then the same bound holds w.r.t. the dominated randomization (for average bounds we had to adjust the bound). This is true for upper bounds with probability 1 as well as for lower bounds! The condition on the convergence rate naturally depends on the ratio of the pf's in the randomization. In our case this ratio is polynomial on the average (since domination holds) and super-polynomial convergence is sufficient.

To be more precise say that the sequence $a_n$ vanishes superpolynomially fast if for every polynomial $p$ and for all sufficiently large $n \in \mathbb{N}^+$ $a_n \leq p(n)^{-1}$.

We define now bounds with probability 1 with super-polynomial convergence and proof that they are preserved under domination.

**Definitions 7.4 (Strong upper and lower bounds with probability 1)**

Let $\langle S_n, \mu_n \rangle$ be a local randomization and $T$ a run-time function.
(i) We say that \( f \) is a strong upper bound with probability 1 on \( T \) (with respect to \( \mu_n \)) if the sequence \( \Pr_{\mu_n} \{ T(x) > f(n) \} \) vanishes superpolynomially fast.

(ii) We say that \( f \) is a strong lower bound with probability 1 on \( T \) (with respect to \( \mu_n \)) if the sequence \( \Pr_{\mu_n} \{ T(x) \leq f(n) \} \) vanishes superpolynomially fast.

(iii) For both of these bounds we say strong bounds with probability 1.

Theorem 11 (Robustness of strong bounds with probability 1)

Let \( <S_n, \mu_n> \), \( <S_n, \nu_n> \) be local randomizations such that \( \nu \) dominates \( \mu \). Also let \( f : R^+ \to R^+ \) and \( T : S \to R^+ \). If \( f \) is a strong upper (lower) bound with probability 1 on \( T \) with respect to \( <S_n, \nu_n> \) then \( f \) is a strong upper (lower) bound with probability 1 on \( T \) with respect to \( <S_n, \mu_n> \).

Proof: We will prove that for any event \( C \subseteq S \) for which \( \Pr_{\nu_n} \{ C \} \) vanishes superpolynomially fast also \( \Pr_{\mu_n} \{ C \} \) vanishes superpolynomially fast. Define

\[
r(x) = \frac{\mu|x|}{\nu|x|} \]

and let \( \varepsilon > 0 \) such that

\[
E^{\mu_n} (r(x)^\varepsilon) \leq n.
\]

Let \( c \in N^+ \) be any constant and for \( n \in N^+ \) define:

\[
A_n = \{ x \in C \cap S_n : r(x) \leq n^{\frac{\varepsilon + 1}{\varepsilon}} \}
\]

\[
B_n = \{ x \in C \cap S_n : r(x) > n^{\frac{\varepsilon + 1}{\varepsilon}} \}
\]

We will show that \( \Pr_{\mu_n} \{ A_n \} \) and \( \Pr_{\mu_n} \{ B_n \} \) are bounded by \( n^{-\varepsilon} \) which is sufficient to show that \( \Pr_{\mu_n} \{ C \} \) vanishes superpolynomially fast. To see this for \( A_n \) note that \( \Pr_{\nu_n} \{ A_n \} \) vanishes superpolynomially fast and that

\[
\Pr_{\mu_n} \{ A_n \} = \sum_{x \in A_n} \mu_n(x) \leq \sum_{x \in A_n} n^{\frac{\varepsilon + 1}{\varepsilon}} \nu_n(x)
\]

\[
= n^{\frac{\varepsilon + 1}{\varepsilon}} \Pr_{\nu_n} \{ A_n \}.
\]

For the second part we recall domination and write:

\[
n \geq \sum_{x \in S_n} r(x)^\varepsilon \mu_n(x) \geq \sum_{x \in B_n} r(x)^\varepsilon \mu_n(x)
\]

\[
> \sum_{x \in B_n} n^{\varepsilon + 1} \mu_n(x).
\]
Therefore
\[ n^{-c} > \sum_{x \in B_n} \mu_n(x) = Pr_{\mu_n} \{ B_n \} . \]

The condition of super-polynomial convergence to zero in the theorem is necessary and counterexamples for the case of only polynomially fast convergence are easy to construct.

7.5 Robustness of bounds on the expectation

It should be noted that bounds on the expectation \( E_{\nu_n}(T) \) are not preserved if \( \nu \) is replaced by a pf \( \mu \) so that \( < S_n, \nu_n > \) dominates \( < S_n, \mu_n > \).

In the original work of Levin([Lev86]) as well as in [BCGL92] and [Gol88] domination is defined by requesting the ratio \( \frac{\mu(x)}{\nu(x)} \) to be bounded by a polynomial in \( |x| \).

Clearly if for some polynomial \( p \) and all \( x \in S_n (\mu_n(x)/\nu_n(x)) \leq p(n) \) then \( E_{\mu_n}(T) \leq p(n)E_{\nu_n}(T) \). This implies that if \( < S_n, \nu_n > \) dominates \( < S_n, \mu_n > \) with this (Levin's) definition, and \( E_{\nu_n}(T) \) is bounded by a polynomial then \( E_{\mu_n}(T) \) is also bounded by a polynomial. So for this definition a polynomial bound on the expected run time on inputs of size \( n \) is conserved.

8 A concluding example

Assume that you prove in the course of a probabilistic analysis of your favorite algorithm the following statement about its run-time \( T \) and a local randomization \( < S_n, \mu_n > \) on the inputs:

\[ \lim_{n \to \infty} Pr_{\mu_n} \{ T(x) > 2^m \} = 1. \] (1)

Here \( c \) is a fixed positive real number. This is a lower bound with probability 1. Such a result was proved in [CS88] about resolution and a uniform randomization on \( k\)--CNF, but the details are irrelevant for our discussion here. A detailed analysis of resolution in our framework may be found in [MS]. In this section we want to exemplify our results by discussing the meaning of a result like 1 in terms of average bounds.
8.1 Deriving an average lower bound

From theorem 3 we know that lower bounds with probability 1 do not have the **Dichotomy property**.

Let $<S, \mu>$ be the global randomization induced by some strongly regular weight function and $<S_n, \mu_n>$. Theorem 6 allows us to get from 1 that for every $b \in R^+$ we get

$$\forall \epsilon \in (0, 1) \ T \in ALB(<S, \mu>, 2^{bn\epsilon}). \quad (2)$$

8.2 Changing the bounding function

Although 2 is slightly weaker than 1, it is more manageable in the following sense. Let $d \in R^+$ and $d > c$. It could happen that for such a $d$

$$Pr_{\mu_n}\{T(x) > 2^{dn}\} \quad (3)$$

does not converge at all. For the average, the Dichotomy property and Consistency (theorem 4) still give us that

either $T \in ALB(<S, \mu>, 2^{dn})$ or $T \in AUB(<S, \mu>, 2^{dn})$. \quad (4)

Assume now that $T \in AUB(<S, \mu>, 2^{dn})$. From equation 2 we still get

$$\forall \epsilon \in (0, 1) \ T \in ALB(<S, \mu>, 2^{dn}) \quad (5)$$

but statement 4 shows that this is optimal in the sense that $\epsilon = 1$ is excluded.

From theorem 3 we also know that lower bounds with probability 1 do not have the Transitivity-1 property. This can be illustrated as follows. If all we know about our favorite algorithm is given by the statement 1 it is conceivable that for some constant $a \in R^+$

$$Pr_{\mu_n}\{T(x) > 2^{cn+a}\} \quad (6)$$

does not converge. But for the average we still have

$$\forall \epsilon \in (0, 1) \ T \in ALB(<S, \mu>, 2^{bn\epsilon+a}) \quad (7)$$

and that this is optimal as before. This is clearly so, because the $\epsilon < 1$ in the exponent absorbs the other changes.
8.3 Changing the randomization

Let \( <S, \nu> \) be a global randomization on the inputs such that \( <S, \nu> \) dominates \( <S, \mu> \). Theorem 10 allows us to replace \( \mu \) by \( \nu \) in statement 2, but not necessarily in statement 1.

Let \( <S_n, \eta_n> \) be a local randomization on the inputs such that \( \mu \) dominates \( \eta \). Now, theorem 11 allows us to replace \( \mu \) by \( \eta \) in statement 1, provided the convergence is superpolynomially fast. In this case we have statement 2 for all corresponding global randomizations \( <S, \eta> \leq <S, \mu> \leq <S, \nu> \) induced by some strongly regular weight function.

In average case complexity theory \( \nu \) dominates \( \mu \) is usually interpreted as saying that a problem with input randomization \( \mu \) is 'easier' than the same problem with input randomization \( \nu \). This interpretation stems from looking at average upper bounds (cf. theorem 10). Our average lower bounds behave consistently with this interpretation, whereas lower bounds with probability 1 behave counter-intuitively even if superpolynomial convergence is assumed.

8.4 Summary

The conclusions are, that statement 1 is a very strong result in the probabilistic analysis of a particular algorithm, but it may be too sensitive to minor changes, if it were to be used in the context of average case complexity theory.

9 Conclusions and further research

Our main goal was to discuss the advantages and disadvantages of various definitions of probabilistic lower bounds. Lower bounds on the expectation carry very little information. Such a lower bound may be caused by a small set of ill-behaving inputs, without reflecting the average behavior. Lower bounds with probability 1, on the other side, carry too much information, in the sense that it gets lost, when small changes in the bounding functions or in the underlying randomization are allowed. Our newly introduced lower bound on the average seems to fit best into the general framework of average upper bounds and seems to offer the best balance between robustness under changes and precision in describing the probabilistic behavior of a run-time function \( T \).

We have also extended the definitions of average upper bounds such as to include other than polynomial bounds. This can be used to define
average complexity classes $\text{AverDTIME}(f)$ and $\text{AverNTIME}(f)$, analogous to $\text{DTIME}(f)$ and $\text{NTIME}(f)$ in worst case complexity. Moreover, our definition of average lower bounds and the Dichotomy property allow the passage from the non-existence of certain average upper bounds to the existence of closely related average lower bounds, and vice versa, which facilitates the formulation of separating average complexity classes. Future work should develop Average Case Complexity Theory further for non-polynomial classes by identifying the right notion of randomized reductions and by finding natural problems which are complete for the respective average case complexity classes. Ultimately, the study of these generalized average case complexity classes should shed new light on our understanding of feasible and non-feasible problems.

As a side issue we discussed the advantages and disadvantages of local and global randomizations. For average complexity global randomization are more appropriate since they average also the input-size and so make it possible to define averages over the whole input-set. However results 3.2 and 9 show that if one restricts the global randomization to be regular then one may also use local randomizations in the average context and the difference between local and global randomizations becomes mainly technical.
A Appendix

A.1 History of the definition of polynomial on the average.

The original condition for a run-time function $T$ to be polynomial on the average with respect to a probability function $\mu$ as given by Levin in [Lev86], is that for some $\varepsilon > 0$ the expectation of $\frac{T(x)}{|x|}^\varepsilon$ (w.r.t. $\mu$) is finite.

This definition does not seem intuitively natural. Consequently the authors of [GoI88],[BCGL92] and [Gur91] have invested some effort in pointing out shortcomings of alternative, more natural seeming, definitions. To summarize consider to define that $T$ is polynomial on the average if $E_{\mu_n}(T(x)) \in O(Poly)$. The major drawback of this definition is that it is not robust under application of polynomials, for example $T$ might be polynomial on the average but $T^2$ not. This leads to (unacceptable) problems in the definition of any complexity class based on polynomial on the average bounded time.

Such a class will not be closed under functional composition of algorithms since the run-time of a compound algorithm composed of two polynomial on the average time bounded algorithm is in general bound only by a polynomial of a function that is polynomial on the average. (For example consider what happens when the first algorithm produces output with length a square of the input-length.) Such a class will also not be machine-independent, since simulating an algorithm model on another might square the run-time. The same effect leads to coding-dependency — that is an algorithm might be polynomial on the average for one coding of the inputs, but the algorithm that accepts a different coding of the inputs and retrieves first the original coding and then calls the first algorithm might not be polynomial on the average. For much more concrete examples see [GoI88, BCGL92].

In [BCGL92, GoI88] it is also pointed out that this definition deals separately with inputs of different length. This is true for every type of bound defined with respect to a local randomization. In proposition 3.2 we present a definition (from [Gur90]) for a function to be at most average-polynomial that is defined directly for local randomizations. This 'local' definition of polynomial on the average is equivalent to Levin's original 'global' definition if the weight function of the global randomization is restricted to be regular. So if we restrict ourselves to regular randomizations, dealing separately with inputs of different sizes does not present any problems. The
more technical problem that arises if we would like to consider only a subset of the input-sizes, for example only inputs of even length, can be avoided easily by redefining the run-time or size function appropriately.

In Levin's original definition the bounding function appears only implicitly through its inverse. The formulation given here for at most \( f \) on the average is taken from [BCGL92] where the bounding function is made explicit. It is equivalent to Levin's original definition for polynomial bounding functions. But the definition that makes the bounding function explicit gives a much easier intuitive access to the nature of average upper bounds and also provides for a natural generalization to general bounding functions. Historically the definition in [BCGL92] was developed to motivate Levin's original definition but it stands by its own merit. So it is used in the same paper ([BCGL92]) to define average-log-space (i.e. problems solvable in space bounded by \( s(n) \) with \( s \in AUB(\log) \)). This definition of average-log-space has several desired properties namely that average-log-space is included in \( \text{AverP} \) (defined below) and closed under average-log-space (many-to-one) reductions. (It is also shown that if one replaces the condition \( s \in AUB(\log) \) with \( E(s(x) / \log |x|) < \infty \) this properties do not hold; this would allow much too large functions.)

A.2 Some remarks on regular weight functions

For the rest of this subsection let \( S \) be an input set, \( < S_n, \mu_n > \) be a local randomization, \( w \) be a weight function and \( < S, \mu > \) be the global randomization induced by \( w \). Furthermore, let \( T : S \rightarrow R^+ \).

One of the reasons to restrict the (weight–function of the) global randomization to be regular, is given by proposition 3.2, which offers a definition of polynomial on the average which is independent of the weight–function. Another reason we present here. Intuitively one would expect that if \( T \) depends only on \( |x| \) and is not bounded by a polynomial then \( T \) should not be polynomial on the average w.r.t. any randomization (excluding trivial randomizations that give positive probability to only a finite number of inputs). This is true only if randomizations are restricted to be regular and is stated formally in the following proposition A.1.

**Proposition A.1** Let \( S \) be an input set and \( T : S \rightarrow R^+ \). If \( w \) is not regular then there is \( T \) so that \( T \) depends only on the size of the input and is not bounded by any polynomial but is polynomial on the average w.r.t. any
randomization with weight-function \( w \).

**Proof:** Let \( w \) be the weight-function of \( <S, \mu> \) and assume that \( w(n)^{-1} \) is not bounded by any polynomial in \( n \), then the function \( T(x) = w(|x|)^{-1} \) is not bounded from above by any polynomial. On the other hand \( T \) is polynomial on the average with respect to \( <S, \mu> \).

### A.3 Some remarks on 'linear on the average'

One should be careful not to take the name *linear on the average* too literally. For example for any \( 0 < \varepsilon < 1 \) the function \( \ell(x) = |x|^{1+\varepsilon} \) is linear on the average with respect to *any* randomization that has \( w(n) = n^{-2} \) as weight function. More general for any \( <S, \mu> \) \( L(<S, \mu>) \) contains functions of the form \( \ell(x) = h(|x|) \) with \( h(n)/n \to \infty \). This hinges on the fact that there is no largest weight function, i.e. there is no largest sequence \( w(n) \) with converging sum. We emphasize this fact about \( L \) since it is crucial for the understanding of the notion of *linear on the average*, which in turn is the basis for all the global bounds we will define in this work.

In the rest of this remark we show how to construct functions \( h \) as mentioned above from the weight function \( w \). Without loss of generality assume that there exists a differentiable function \( f : \mathbb{R}^+ \to \mathbb{R}^+ \) so that \( f(n) \to \infty \) and for every \( n \in \mathbb{N}^+ \):

\[
    w(n) \leq - \left( \frac{1}{f(n)} \right)' = f'(n) f(n)^{-2}.
\]

(Such an \( f \) can be constructed by interpolating the sequence \( s_n = \sum_{n=1}^m w(n) \) to a positive differentiable function \( s \) and taking \( f(a) = s(a)^{-1} \).) For every \( 0 < \varepsilon < 1 \) the function \( h \) defined by

\[
    h(n) = nf(n)\varepsilon
\]

will have the desired property.

### A.4 Alternative definitions for average lower bounds

The form of the definition of average lower bounds was inspired by the negation of the definition of average upper bounds (3.2). We will now discuss alternative definitions that can be obtained by modifying the negation of definition 3.2. For example:
Definition A.1 (1rst attempt for an alternative definition of $\text{ALB}$)

Let $\text{ALB}'(f)$ be the set of all functions $T : S \rightarrow R^+$ so that for every linear on the average function $\ell \in \mathcal{L}$ and for all but a finite number of inputs $x$ $T(x) \geq f(\ell(x))$.

The problem with this definition is that a function has to be very large in order to satisfy it. This is so because $\mathcal{L}$ contains functions that are very large infinitely often (but with vanishing probability). As an example consider the randomization of binary strings with a uniform pf (i.e. induced by $\mu_n(x) = 2^{-|x|}$ and some $w$). Let $\ell(x) = 2|x|$ if $x$ is a string of 1's only and otherwise $\ell(x) = |x|$. $\ell \in \mathcal{L}$ but any $T \in \text{ALB}'(n)$ has to be greater or equal to $2^n$ infinitely often.

We might try to relax the definition a little and define:

Definition A.2 (2nd attempt for an alternative definition of $\text{ALB}$)

Let $\text{ALB}^*(f)$ be the set of all functions $T : S \rightarrow R^+$ so that for every linear on the average function $\ell \in \mathcal{L}$:

$$\lim_{n \rightarrow \infty} \Pr_{\mu_n}(T(x) \geq f(\ell(x))) = 1.$$  

This definition is quite a bit more satisfactory (but maybe a bit cumbersome). It can be shown that it satisfies properties (Weak) Transitivity–1, Transitivity–2 and (Weak) Consistency but does not satisfy properties Honesty and Dichotomy. For Honesty the problem is caused by the fact that for every fixed global randomization $\mathcal{L}$ contains functions that increase strictly faster than any linear function. For example if the global randomization has weight function $w(n) = n^{-2}$ then $\ell(x) = |x|^{3/2} \in \mathcal{L}$ (See the remark in A.3). So $T(x) = |x|^{5/4}$ certainly fits the condition in Honesty for $f(n) = n$ but $T$ is always smaller than $\ell$. As a counterexample for Dichotomy consider again a uniform randomization on binary strings and let $T(x) = 2^{|x|}$ if $x$ is a string of 1's only and $T(x) = \sqrt{|x|}$ otherwise. $T$ is not in $\text{ALB}(n)$ (= $\mathcal{L}$) and also not in $\text{ALB}^*(n)$, since it is smaller than the linear on the average function $\ell(x) = |x|$ with conditional probability ($Pr_{\mu_n}$) converging to 0.

To complete our discussion of definition $D^*$ we will show that for any strictly increasing function $f$ $\text{ALB}^*(f) \subset \text{ALB}(f)$. Since $\text{ALB}(f)$ does satisfy condition Wc and $\text{ALB}^*(f)$ does not, this means that our definition of greater than $f$ on the average is strictly weaker than $D^*$. To show that $\text{ALB}^*(f) \subset$
ALB(f) for f strictly increasing it is sufficient to show that \( \text{ALB}^*(n) \subseteq \text{ALB}(n) \). Assume now that \( T \not\in \text{ALB}(n) \). By proposition 4.4 (which shows that average bounds satisfy the Dichotomy property) this means that \( T \) is at most linear on the average, that is \( T \in \mathcal{L} \). Since there is no largest function in \( \mathcal{L} \) there exist a function \( \delta : R^+ \rightarrow R^+ \) so that

1. \( \lim_{n \to \infty} \frac{\delta(n)}{n} = \infty \) and
2. \( -1z\ell(x) = T(x)\delta(|z|) \in \mathcal{L} \).

(See the remark in A.3 and the proof of proposition 5.4 for a method to construct \( \delta \).) But \( T(x) \geq \ell(x) \) only for some finite set of \( x \) and so \( T \not\in \text{ALB}^*(n) \).
References


