Equivalence of Measures of Complexity Classes [1]

Josef M. Breutzmann
Department of Mathematics
and Computer Science
Wartburg College
Waverly, Iowa 50677
U.S.A.

Jack H. Lutz
Department of Computer Science
Iowa State University
Ames, Iowa 50011
U.S.A.

Abstract. The resource-bounded measures of complexity classes are shown to be robust with respect to certain changes in the underlying probability measure. Specifically, for any real number $\delta > 0$, any uniformly polynomial-time computable sequence $\vec{\beta} = (\beta_0, \beta_1, \beta_2, \dots)$ of real numbers (biases) $\beta_i \in [\delta, 1-\delta]$, and any complexity class $\mathcal{C}$ (such as P, NP, BPP, P/Poly, PH, PSPACE, etc.) that is closed under positive, polynomial-time, truth-table reductions with queries of at most linear length, it is shown that the following two conditions are equivalent.

(1) $\mathcal{C}$ has p-measure 0 (respectively, measure 0 in E, measure 0 in E_2) relative to the coin-toss probability measure given by the sequence $\vec{\beta}$.

(2) $\mathcal{C}$ has p-measure 0 (respectively, measure 0 in E, measure 0 in E_2) relative to the uniform probability measure.

The proof introduces three techniques that may be useful in other contexts, namely, (i) the transformation of an efficient martingale for one probability measure into an efficient martingale for a "nearby" probability measure; (ii) the construction of a *positive bias reduction*, a truth-table reduction that encodes a positive, efficient, approximate simulation of one bias sequence by another; and (iii) the use of such a reduction to *dilate* an efficient martingale for the simulated probability measure into an efficient martingale for the simulating probability measure.

1. Introduction. In the 1990's, the measure-theoretic study of complexity classes has yielded a growing body of new, quantitative insights into various much-studied aspects of computational complexity. Benefits of this study to date include improved bounds on the densities of hard languages [15]; newly discovered relationships among circuit-size complexity, pseudorandom generators, and natural proofs [21]; strong new hypotheses that may have sufficient explanatory power (in terms of provable, plausible consequences) to help unify our present plethora of unsolved fundamental problems [18, 15, 7, 17, 14]; and a new generalization of the completeness phenomenon that dramatically enlarges the set of computational problems that are provably strongly intractable [13, 6, 2, 7, 8, 1]. See [12] for a survey of these and related developments.

Intuitively, suppose that a language $A \subseteq \{0,1\}^*$ is chosen according to a random experiment in which an independent toss of a fair coin is used to decide whether each string is in A. Then *classical* Lebesgue measure theory (described in [5, 20], for example) identifies certain *measure 0* sets X of languages, for which the probability that $A \in X$ in this experiment is 0. *Effective* measure theory, which says what it means for a set of decidable languages to have measure 0 as a subset of the set of all such languages, has been investigated by Freidzon [4], Mehlhorn [19], and others. The *resource-bounded* measure theory introduced by Lutz [11] is a powerful generalization of Lebesgue measure. Special cases of resource-bounded measure include classical Lebesgue measure; a strengthened version of effective measure; and most importantly, measures in $\mathrm{E} = \mathrm{DTIME}(2^{\mathrm{linear}})$, $\mathrm{E}_2 = \mathrm{DTIME}(2^{\mathrm{polynomial}})$, and other complexity classes. The *small* subsets of such a complexity class are then the measure 0 sets; the *large* subsets are the measure 1 sets (complements of measure 0 sets). We say that *almost every* language in a complexity class $\mathcal{C}$ has a given property if the set of

[1] This research was supported in part by National Science Foundation Grants CCR-9157382 (with matching funds from Rockwell, Microware Systems Corporation, and Amoco Foundation) and CCR-9610461.

languages in $\mathcal{C}$ that exhibit the property has measure 1 in $\mathcal{C}$.

All work to date on the measure-theoretic structure of complexity classes has employed the resource-bounded measure that is described briefly and intuitively above. This resource-bounded measure is based on the *uniform* probability measure, corresponding to the fact that the coin tosses are fair and independent in the above-described random experiment. The uniform probability measure has been a natural and fruitful starting point for the investigation of resource-bounded measure (just as it was for the investigation of classical measure), but there are good reasons to also investigate resource bounded measures that are based on other probability measures. For example, the study of such alternative resource-bounded measures may be expected to have the following benefits.

(i) The study will enable us to determine which results of resource-bounded measure are particular to the uniform probability measure and which are not. This, in turn, will provide some criteria for identifying contexts in which the uniform probability measure is, or is not, the natural choice.

(ii) The study is likely to help us understand how the complexity of the underlying probability measure interacts with other complexity parameters, especially in such areas as algorithmic information theory, average case complexity, cryptography, and computational learning, where the variety of probability measures already plays a major role.

(iii) The study will provide new tools for proving results concerning resource-bounded measure based on the uniform probability measure.

The present paper initiates the study of resource-bounded measures that are based on nonuniform probability measures.

Let $\mathbf{C}$ be the set of all languages $A \subseteq \{0,1\}^*$. (The set $\mathbf{C}$ is often called *Cantor space.*) Given a probability measure ν on $\mathbf{C}$ (a term defined precisely below), section 3 of this paper describes the basic ideas of resource-bounded ν-measure, generalizing definitions and results from [11, 13, 12] to ν in a natural way. In particular, section 3 specifies what it means for a set $X \subseteq \mathbf{C}$ to have p-ν-measure 0 (written $\nu_{\mathrm{p}}(X) = 0$), p-ν-measure 1, ν-measure 0 in E (written $\nu(X|\mathrm{E}) = 0$), ν-measure 1 in E, ν-measure 0 in E_2, or ν-measure 1 in E_2. Intuitively, for example, a set $X \subseteq \mathbf{C}$ has p-ν-measure 0 if, when a language $A \in \mathbf{C}$ is chosen probabilistically according to the probability measure ν, the event "$A \in X$" has probability 0 in a strong sense that can be exploited by a polynomial-time betting algorithm.

Most of the results in the present paper concern a restricted (but broad) class of probability measures on $\mathbf{C}$, namely, coin-toss probability measures that are given by P-computable, strongly positive sequences of biases. These probability measures are described intuitively in the following paragraphs (and precisely in section 3).

Given a sequence $\vec{\beta} = (\beta_0, \beta_1, \beta_2, \dots)$ of real numbers (biases) $\beta_i \in [0,1]$, the *coin-toss probability measure* (also call the *product probability measure*) *given by* $\vec{\beta}$ is the probability measure $\mu^{\vec{\beta}}$ on $\mathbf{C}$ that corresponds to the random experiment in which a language $A \in \mathbf{C}$ is chosen probabilistically as follows. For each string s_i in the standard enumeration $s_0, s_1, s_2, \dots$ of $\{0,1\}^*$, we toss a special coin, whose probability is β_i of coming up heads, in which case $s_i \in A$, and $1 - \beta_i$ of coming up tails, in which case $s_i \notin A$. The coin tosses are independent of one another.

In the special case where $\vec{\beta} = (\beta, \beta, \beta, \dots)$, i.e., the biases in the sequence $\vec{\beta}$ are all β, we write μ^{β} for $\mu^{\vec{\beta}}$. In particular, $\mu^{\frac{1}{2}}$ is the uniform probability measure, which, in the literature of resource-bounded measure, is denoted simply by μ.

A sequence $\vec{\beta} = (\beta_0, \beta_1, \beta_2, \dots)$ of biases is *strongly positive* if there is a real number $\delta > 0$ such that each $\beta_i \in [\delta, 1-\delta]$. The sequence $\vec{\beta}$ is P-*computable* (and we call it a P-*sequences of biases*) if there is a polynomial-time algorithm that, on input $(s_i, 0^r)$, computes a rational approximation of β_i to within 2^{-r}.

In section 4, we prove the Summable Equivalence Theorem, which implies that, if $\vec{\alpha}$ and $\vec{\beta}$ are strongly positive P-sequences of biases that are "close" to one another, in the sense that $\sum_{i=0}^{\infty} |\alpha_i - \beta_i| < \infty$, then for every set $X \subseteq \mathbf{C}$,

$$\mu_{\mathrm{p}}^{\vec{\alpha}}(X) = 0 \Longleftrightarrow \mu_{\mathrm{p}}^{\vec{\beta}}(X) = 0.$$

That is, the p-measure based on $\vec{\alpha}$ and the p-measure based on $\vec{\beta}$ are in absolute agreement as to which sets of languages are small.

In general, if $\vec{\alpha}$ and $\vec{\beta}$ are not in some sense close to one another, then the p-measures based on $\vec{\alpha}$ and $\vec{\beta}$ need not agree in the above manner. For example, if $\alpha, \beta \in [0,1]$, $\alpha \neq \beta$, and

$$X_\alpha = \left\{ A \in \mathbf{C} \middle| \lim_{n \to \infty} 2^{-n} |A \cap \{0,1\}^n| = \alpha \right\},$$

then a routine extension of the Weak Stochasticity Theorem of [15] shows that $\mu_{\mathrm{p}}^{\alpha}(X_\alpha) = 1$, while $\mu_{\mathrm{p}}^{\beta}(X_\alpha) = 0$.

Notwithstanding this example, many applications of resource-bounded measure do not involve *arbitrary* sets $X \subseteq \mathbf{C}$, but rather are concerned with the measures of *complexity classes* and other closely related classes of languages. Many such classes of interest, including P, NP, co-NP, R, BPP, AM, P/Poly, PH, PSPACE, etc., are closed under positive, polynomial-time truth-table reductions ($\leq_{\mathrm{pos-tt}}^{\mathrm{P}}$-reductions), and their intersections with E are closed under $\leq_{\mathrm{pos-tt}}^{\mathrm{P}}$-reductions with linear bounds on the lengths of the queries ($\leq_{\mathrm{pos-tt}}^{\mathrm{P,lin}}$-reductions).

The main theorem of this paper is the Bias Equivalence Theorem. This result, proven in section 8, says that, for every class $\mathcal{C}$ of languages that is closed under $\leq_{\mathrm{pos-tt}}^{\mathrm{P,lin}}$-reductions, the p-measure of $\mathcal{C}$ is somewhat robust with respect to changes in the underlying probability measure. Specifically, if $\vec{\alpha}$ and $\vec{\beta}$ are strongly positive P-sequences of biases and $\mathcal{C}$ is a class of languages that is closed under $\leq_{\mathrm{pos-tt}}^{\mathrm{P,lin}}$-reductions, then the Bias Equivalence Theorem says that

$$\mu_{\mathrm{p}}^{\vec{\alpha}}(\mathcal{C}) = 0 \Longleftrightarrow \mu_{\mathrm{p}}^{\vec{\beta}}(\mathcal{C}) = 0.$$

To put the matter differently, for every strongly positive P-sequence $\vec{\beta}$ of biases and every class $\mathcal{C}$ that is closed under $\leq_{\mathrm{pos-tt}}^{\mathrm{P,lin}}$-reductions,

$$\mu_{\mathrm{p}}^{\vec{\beta}}(\mathcal{C}) = 0 \Longleftrightarrow \mu_{\mathrm{p}}(\mathcal{C}) = 0.$$

This result implies that most applications of resource-bounded measure to date can be immediately generalized from the uniform probability measure (in which they were developed) to arbitrary coin-toss probability measures given by strongly positive P-sequences of biases.

The Bias Equivalence Theorem also offers the following new technique for proving resource-bounded measure results. If $\mathcal{C}$ is a class that is closed under $\leq_{\mathrm{pos-tt}}^{\mathrm{P,lin}}$-reductions, then in order to prove that $\mu_{\mathrm{p}}(\mathcal{C}) = 0$, it suffices to prove that $\mu_{\mathrm{p}}^{\vec{\beta}}(\mathcal{C}) = 0$ for some conveniently chosen strongly positive P-sequence $\vec{\beta}$ of biases. (The Bias Equivalence Theorem has already been put to this use in the forthcoming paper [16].)

The plausibility and consequences of the hypothesis $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$ are subjects of recent and ongoing research [18, 15, 7, 17, 14, 3, 16]. The Bias Equivalence Theorem immediately implies that the following three statements are equivalent.

(H1) $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$.

(H2) For every strongly positive P-sequence $\vec{\beta}$ of biases, $\mu_{\mathrm{p}}^{\vec{\beta}}(\mathrm{NP}) \neq 0$.

(H3) There exists a strongly positive P-sequence $\vec{\beta}$ of biases such that $\mu_{\mathrm{p}}^{\vec{\beta}}(\mathrm{NP}) \neq 0$.

The statements (H2) and (H3) are thus new, equivalent formulations of the hypothesis (H1).

The proof of the Bias Equivalence Theorem uses three main tools. The first is the Summable Equivalence Theorem, which we have already discussed. The second is the Martingale Dilation Theorem, which is proven in section 6. This result concerns martingales (defined in section 3), which are the betting algorithms on which resource-bounded measure is based. Roughly speaking, the Martingale Dilation Theorem gives a method of transforming ("dilating") a martingale for one coin-toss probability measure into a martingale for another, perhaps very different, coin-toss probability measure, provided that the former measure is obtained from the latter via an "orderly" truth-table reduction.

The third tool used in the proof of our main theorem is the Positive Bias Reduction Theorem, which is presented in section 7. If $\vec{\alpha}$ and $\vec{\beta}$ are two strongly positive sequences of biases that are exactly P-computable (with no approximation), then the *positive bias reduction* of $\vec{\alpha}$ to $\vec{\beta}$ is a truth-table reduction (in fact, an orderly $\leq_{\mathrm{pos-tt}}^{\mathrm{P,lin}}$-reduction) that uses the sequence $\vec{\beta}$ to "approximately simulate" the sequence $\vec{\alpha}$. It is especially crucial for our main result that this reduction is efficient and positive. (The circuits constructed by the truth-table reduction contain AND gates and OR gates, but no NOT gates.)

The Summable Equivalence Theorem, the Martingale Dilation Theorem, and the Positive Bias Reduction Theorem are only developed and used here as tools to prove our main result. Nevertheless, these three results are of independent interest, and are likely to be useful in future investigations.

2. Preliminaries. In this paper, $\mathbb{N}$ denotes the set of all nonnegative integers, $\mathbb{Z}$ denotes the set of all integers, $\mathbb{Z}^+$ denotes the set of all positive integers, $\mathbb{Q}$ denotes the set of all rational numbers, and $\mathbb{R}$ denotes the set of all real numbers.

We write $\{0,1\}^*$ for the set of all (finite, binary) *strings*, and we write $|x|$ for the length of a string x. The empty string, λ, is the unique string of length 0. The *standard enumeration* of $\{0,1\}^*$ is the sequence $s_0 = \lambda, s_1 = 0, s_2 = 1, s_3 = 00, \ldots$, ordered first by length and then lexicographically. For $x, y \in \{0,1\}^*$, we write $x < y$ if x precedes y in this standard enumeration. For $n \in \mathbb{N}$, $\{0,1\}^n$ denotes the set of all strings of length n, and $\{0,1\}^{\leq n}$ denotes the set of all strings of length at most n.

If x is a string or an (infinite, binary) *sequence*, and if $0 \leq i \leq j < |x|$, then $x[i..j]$ is the string consisting of the i^{th} through j^{th} bits of x. In particular, $x[0..i-1]$ is the *i-bit prefix* of x. We write $x[i]$ for $x[i..i]$, the i^{th} bit of x. (Note that the leftmost bit of x is $x[0]$, the 0^{th} bit of x.)

If w is a string and x is a string or sequence, then we write $w \sqsubseteq x$ if w is a prefix of x, i.e., if there is a string or sequence y such that $x = wy$.

The *Boolean value* of a condition ϕ is $[\![\phi]\!] = \textbf{if } \phi \textbf{ then } 1 \textbf{ else } 0$.

In this paper we use both the binary logarithm $\log \alpha = \log_2 \alpha$ and the natural logarithm $\ln \alpha = \log_e \alpha$.

Many of the functions in this paper are real-valued functions on discrete domains. These typically have the form

$$f : \mathbb{N}^m \times \{0,1\}^* \longrightarrow \mathbb{R}, \tag{2.1}$$

where $m \in \mathbb{N}$. (If $m = 0$, we interpret this to mean that $f : \{0,1\}^* \longrightarrow \mathbb{R}$.) Such a function f is defined to be p-*computable* if there is a function

$$\hat{f} : \mathbb{N} \times \mathbb{N}^m \times \{0,1\}^* \longrightarrow \mathbb{Q} \tag{2.2}$$

with the following two properties.

(i) For all $r, k_1, \ldots, k_m \in \mathbb{N}$ and $w \in \{0,1\}^*$,

$$|\hat{f}(r, k_1, \ldots, k_m, w) - f(k_1, \ldots, k_m, w)| \leq 2^{-r}.$$

(ii) There is an algorithm that, on input $(r, k_1, \ldots, k_m, w)$, computes the value $\hat{f}(r, k_1, \ldots, k_m, w)$ in $(r + k_1 + \ldots + k_m + |w|)^{O(1)}$ time.

Similarly, f is defined to be p_2-*computable* if there is a function $\hat{f}$ as in (2.2) that satisfies condition (i) above and the following condition.

(ii$'$) There is an algorithm that, on input $(r, k_1, \ldots, k_m, w)$, computes the value $\hat{f}(r, k_1, \ldots, k_m, w)$ in $2^{\log(r+k_1+\ldots+k_m+|w|)^{O(1)}}$ time.

In this paper, functions of the form (2.1) always have the form

$$f : \mathbb{N}^m \times \{0,1\}^* \longrightarrow [0, \infty)$$

or the form

$$f : \mathbb{N}^m \times \{0,1\}^* \longrightarrow [0, 1].$$

If such a function is p-computable or p_2-computable, then we assume without loss of generality that the approximating function $\hat{f}$ of (2.2) actually has the form

$$\hat{f} : \mathbb{N} \times \mathbb{N}^m \times \{0,1\}^* \longrightarrow \mathbb{Q} \cap [0, \infty)$$

or the form

$$\hat{f} : \mathbb{N} \times \mathbb{N}^m \times \{0,1\}^* \longrightarrow \mathbb{Q} \cap [0, 1],$$

respectively.

3. Resource-Bounded ν-Measure. In this section, we develop basic elements of resource-bounded measure based on an arbitrary (Borel) probability measure ν. The ideas here generalize the corresponding ideas of "ordinary" resource-bounded measure (based on the uniform probability measure μ) in a straightforward and natural way, so our presentation is relatively brief. The reader is referred to [11, 12] for additional discussion.

We work in the *Cantor space* $\mathbf{C}$, consisting of all languages $A \subseteq \{0,1\}^*$. We identify each language A with its *characteristic sequence*, which is the infinite binary sequence χ_A defined by

$$\chi_A[n] = [\![s_n \in A]\!]$$

for each $n \in \mathbb{N}$. Relying on this identification, we also consider $\mathbf{C}$ to be the set of all infinite binary sequences.

For each string $w \in \{0,1\}^*$, the *cylinder generated by* w is the set

$$\mathbf{C}_w = \{A \in \mathbf{C} \mid w \sqsubseteq \chi_A\}.$$

Note that $\mathbf{C}_\lambda = \mathbf{C}$.

We first review the well-known notion of a (Borel) probability measure on $\mathbf{C}$.

DEFINITION 3.1. *A probability measure on* $\mathbf{C}$ *is a function*

$$\nu : \{0,1\}^* \longrightarrow [0,1]$$

such that $\nu(\lambda) = 1$, *and for all* $w \in \{0,1\}^*$,

$$\nu(w) = \nu(w0) + \nu(w1).$$

Intuitively, $\nu(w)$ is the probability that $A \in \mathbf{C}_w$ when we "choose a language $A \in \mathbf{C}$ according to the probability measure ν." We sometimes write $\nu(\mathbf{C}_w)$ for $\nu(w)$.

Examples.

1. The *uniform probability measure* μ is defined by

   $$\mu(w) = 2^{-|w|}$$

   for all $w \in \{0,1\}^*$.
2. A *sequence of biases* is a sequence $\vec{\beta} = (\beta_0, \beta_1, \beta_2, \ldots)$, where each $\beta_i \in [0,1]$. Given a sequence of biases $\vec{\beta}$, the $\vec{\beta}$*-coin-toss probability measure* (also called the $\vec{\beta}$*-product probability measure*) is the probability measure $\mu^{\vec{\beta}}$ defined by

   $$\mu^{\vec{\beta}}(w) = \prod_{i=0}^{|w|-1} ((1-\beta_i)\cdot(1-w[i]) + \beta_i \cdot w[i])$$

   for all $w \in \{0,1\}^*$.
3. If $\beta = \beta_0 = \beta_1 = \beta_2 = \ldots$, then we write μ^β for $\mu^{\vec{\beta}}$. In this case, we have the simpler formula

   $$\mu^\beta(w) = (1-\beta)^{\#(0,w)} \cdot \beta^{\#(1,w)},$$

   where $\#(b,w)$ denotes the number of b's in w. Note that $\mu^{\frac{1}{2}} = \mu$.

Intuitively, $\mu^{\vec{\beta}}(w)$ is the probability that $w \sqsubseteq A$ when the language $A \subseteq \{0,1\}^*$ is chosen probabilistically according to the following random experiment. For each string s_i in the standard enumeration $s_0, s_1, s_2, \ldots$ of $\{0,1\}^*$, we (independently of all other strings) toss a special coin, whose probability is β_i of coming up heads, in which case $s_i \in A$, and $1-\beta_i$ of coming up tails, in which case $s_i \notin A$.

DEFINITION 3.2. *A probability measure* ν *on* $\mathbf{C}$ *is positive if, for all* $w \in \{0,1\}^*$, $\nu(w) > 0$.

DEFINITION 3.3. *If* ν *is a positive probability measure and* $u, v \in \{0,1\}^*$, *then the conditional* ν*-measure of* u *given* v *is*

$$\nu(u|v) = \begin{cases} 1 & \textit{if } u \sqsubseteq v \\ \frac{\nu(u)}{\nu(v)} & \textit{if } v \sqsubseteq u \\ 0 & \textit{otherwise.} \end{cases}$$

Note that $\nu(u|v)$ is the conditional probability that $A \in \mathbf{C}_u$, given that $A \in \mathbf{C}_v$, when $A \in \mathbf{C}$ is chosen according to the probability measure ν.

Most of this paper concerns the following special type of probability measure.

DEFINITION 3.4. *A probability measure ν on $\mathbf{C}$ is strongly positive if (ν is positive and) there is a constant $\delta > 0$ such that, for all $w \in \{0,1\}^*$ and $b \in \{0,1\}$, $\nu(wb|w) \geq \delta$.*

DEFINITION 3.5. *A sequence of biases $\vec{\beta} = (\beta_0, \beta_1, \beta_2, \dots)$ is strongly positive if there is a constant $\delta > 0$ such that, for all $i \in \mathbb{N}$, $\beta_i \in [\delta, 1-\delta]$.*

If $\vec{\beta}$ is a sequence of biases, then the following two observations are clear.

1. $\mu^{\vec{\beta}}$ is positive if and only if $\beta_i \in (0,1)$ for all $i \in \mathbb{N}$.
2. If $\mu^{\vec{\beta}}$ is positive, then for each $w \in \{0,1\}^*$,

$$\mu^{\vec{\beta}}(w0|w) = 1 - \beta_{|w|}$$

and

$$\mu^{\vec{\beta}}(w1|w) = \beta_{|w|}.$$

It follows immediately from these two things that the probability measure $\mu^{\vec{\beta}}$ is strongly positive if and only if the sequence of biases $\vec{\beta}$ is strongly positive.

In this paper, we are primarily interested in strongly positive probability measures ν that are p-computable in the sense defined in section 2.

We next review the well-known notion of a martingale over a probability measure ν. Computable martingales were used by Schnorr [23, 24, 25, 26] in his investigations of randomness, and have more recently been used by Lutz [11] in the development of resource-bounded measure.

DEFINITION 3.6. *Let ν be a probability measure on $\mathbf{C}$. Then a ν-martingale is a function $d : \{0,1\}^* \longrightarrow [0,\infty)$ such that, for all $w \in \{0,1\}^*$,*

$$d(w)\nu(w) = d(w0)\nu(w0) + d(w1)\nu(w1). \tag{3.1}$$

If $\vec{\beta}$ is a sequence of biases, then a $\mu^{\vec{\beta}}$-martingale is simply called a $\vec{\beta}$-martingale. A μ-martingale is even more simply called a martingale. (That is, when the probability measure is not specified, it is assumed to be the uniform probability measure μ.)

Intuitively, a ν-martingale d is a "strategy for betting" on the successive bits of (the characteristic sequence of) a language $A \in \mathbf{C}$. The real number $d(\lambda)$ is regarded as the amount of money that the strategy starts with. The real number $d(w)$ is the amount of money that the strategy has after betting on a prefix w of χ_A. The identity (3.1) ensures that the betting is "fair" in the sense that, if A is chosen according to the probability measure ν, then the expected amount of money is constant as the betting proceeds. (See [23, 24, 25, 26, 27, 11, 13, 12] for further discussion.) Of course, the "objective" of a strategy is to win a lot of money.

DEFINITION 3.7. *A ν-martingale d succeeds on a language $A \in \mathbf{C}$ if*

$$\limsup_{n \longrightarrow \infty} d(\chi_A[0..n-1]) = \infty.$$

The success set of a ν-martingale d is the set

$$S^\infty[d] = \{A \in \mathbf{C} \mid d \textit{ succeeds on } A\}.$$

We are especially interested in martingales that are computable within some resource bound. (Recall that the p-computability and p_2-computability of real valued functions were defined in section 2.)

DEFINITION 3.8. *Let ν be a probability measure on* $\mathbf{C}$.

1. *A p-ν-martingale is a ν-martingale that is p-computable.*
2. *A p_2-ν-martingale is a ν-martingale that is p_2-computable.*

A p-$\mu^{\vec{\beta}}$-martingale is called a p-$\vec{\beta}$-martingale, a p-μ-martingale is called a p-martingale, and similarly for p_2.

We now come to the fundamental ideas of resource-bounded ν-measure.

DEFINITION 3.9. *Let ν be a probability measure on* $\mathbf{C}$, *and let $X \subseteq \mathbf{C}$.*

1. *X has p-ν-measure 0, and we write $\nu_{\mathrm{p}}(X) = 0$, if there is a p-ν-martingale d such that $X \subseteq S^\infty[d]$.*
2. *X has p-ν-measure 1, and we write $\nu_{\mathrm{p}}(X) = 1$, if $\nu_{\mathrm{p}}(X^c) = 0$, where $X^c = \mathbf{C} - X$.*

The conditions $\nu_{\mathrm{p}_2}(X) = 0$ and $\nu_{\mathrm{p}_2}(X) = 1$ are defined analogously.

DEFINITION 3.10. *Let ν be a probability measure on* $\mathbf{C}$, *and let $X \subseteq \mathbf{C}$.*

1. *X has ν-measure 0 in E, and we write $\nu(X|\mathrm{E}) = 0$, if $\nu_{\mathrm{p}}(X \cap E) = 0$.*
2. *X has ν-measure 1 in E, and we write $\nu(X|\mathrm{E}) = 1$, if $\nu(X^c|\mathrm{E}) = 0$.*
3. *X has ν-measure 0 in E_2, and we write $\nu(X|\mathrm{E}_2) = 0$, if $\nu_{\mathrm{p}_2}(X \cap \mathrm{E}_2) = 0$.*
4. *X has ν-measure 1 in E_2, and we write $\nu(X|\mathrm{E}_2) = 1$, if $\nu(X^c|\mathrm{E}_2) = 0$.*

Just as in the uniform case [11], the resource bounds p and p_2 of the above definitions are only two possible values of a very general parameter. Other choices of this parameter yield classical ν-measure [5], constructive ν-measure (as used in algorithmic information theory [29, 27]), ν-measure in the set REC, consisting of all decidable languages, ν-measure in ESPACE, etc.

The rest of this section is devoted to a very brief presentation of some of the fundamental theorems of resource-bounded ν-measure. One of the main objectives of these results is to justify the intuition that *a set with ν-measure 0 in* E *contains only a "negligibly small" part of* E (with respect to ν). For the purpose of this paper, it suffices to present these results for p-ν-measure and ν-measure in E. We note, however, that all these results hold *a fortiori* for p_2-ν-measure, rec-ν-measure, classical ν-measure, ν-measure in E_2, ν-measure in ESPACE, etc.

We first note that ν-measure 0 sets exhibit the set-theoretic behavior of small sets.

DEFINITION 3.11. *Let $X, X_0, X_1, X_2, \ldots \subseteq \mathbf{C}$.*

1. *X is a p-union of the p-ν-measure 0 sets $X_0, X_1, X_2, \ldots$ if $X = \cup_{k=0}^{\infty} X_k$ and there is a sequence $d_0, d_1, d_2, \ldots$ of ν-martingales with the following two properties.*
   (i) For each $k \in \mathbb{N}$, $X_k \subseteq S^\infty[d_k]$.
   (ii) The function $(k, w) \mapsto d_k(w)$ is p-computable.
2. *X is a p-union of the sets $X_0, X_1, X_2, \ldots$ of ν-measure 0 in E if $X = \cup_{k=0}^{\infty} X_k$ and there is a sequence $d_0, d_1, d_2, \ldots$ of ν-martingales with the following two properties.*
   (i) For each $k \in \mathbb{N}$, $X_k \cap \mathrm{E} \subseteq S^\infty[d_k]$.
   (ii) The function $(k, w) \mapsto d_k(w)$ is p-computable.

LEMMA 3.12. *Let ν be a probability measure on* $\mathbf{C}$, *and let $\mathcal{I}$ be either the collection of all p-ν-measure 0 subsets of* $\mathbf{C}$, *or the collection of all subsets of* $\mathbf{C}$ *that have ν-measure 0 in E. Then $\mathcal{I}$ has the following three closure properties.*

1. *If $X \subseteq Y \in \mathcal{I}$, then $X \in \mathcal{I}$.*

2. If X is a finite union of elements of $\mathcal{I}$, then $X \in \mathcal{I}$

3. If X is a p-union of elements of $\mathcal{I}$, then $X \in \mathcal{I}$.

Proof (sketch). Assume that X is a p-union of the p-ν-measure 0 sets $X_0, X_1, X_2, \ldots$, and let $d_0, d_1, d_2, \ldots$ be as in the definition of this condition. It suffices to show that $\nu_{\mathrm{p}}(X) = 0$. (The remaining parts of the lemma are obvious or follow directly from this.) The p-computability of the function $(k, w) \mapsto d_k(w)$ implies that there is a nondecreasing polynomial q such that, for all $k \in \mathbb{N}$ and $w \in \{0,1\}^*$, $d_k(w) \leq 2^{q(k+|w|)}$. Define

$$d : \{0,1\}^* \longrightarrow [0, \infty)$$

$$d(w) = \sum_{k=0}^{\infty} 2^{-q(2k)-k} d_k(w).$$

It is easily checked that d is a p-ν-martingale and that $X \subseteq S^\infty[d]$, so $\nu_{\mathrm{p}}(X) = 0$. $\square$

We next note that, if ν is strongly positive and p-computable, then every singleton subset of E has p-ν-measure 0.

LEMMA 3.13. *If ν is a strongly positive, p-computable probability measure on $\mathbf{C}$, then for every $A \in \mathrm{E}$,*

$$\nu_{\mathrm{p}}(\{A\}) = \nu(\{A\}|\mathrm{E}) = 0.$$

Proof (sketch). Assume the hypothesis, and fix $\delta > 0$ such that, for all $w \in \{0,1\}^*$ and $b \in \{0,1\}$, $\nu(wb|w) \geq \delta$. Define

$$d : \{0,1\}^* \longrightarrow [0, \infty)$$

$$d(\lambda) = 1$$

$$d(wb) = \frac{d(w)}{\nu(wb|w)} \cdot [\![b = [\![s_{|w|} \in A]\!]]\!].$$

It is easily checked that d is a p-ν-martingale and that, for all $n \in \mathbb{N}$, $d(\chi_A[0..n-1]) \geq (1-\delta)^{-n}$, whence $A \in S^\infty[d]$. $\square$.

Note that, for $A \in \mathrm{E}$, the "point-mass" probability measure

$$\pi_A : \{0,1\}^* \longrightarrow [0,1]$$

$$\pi_A(w) = \begin{cases} 1 & \text{if } w \sqsubseteq \chi_A \\ 0 & \text{if } w \not\sqsubseteq \chi_A \end{cases}$$

is p-computable, and $\{A\}$ does *not* have p-π_A-measure 0. Thus, the strong positivity hypothesis cannot be removed from Lemma 3.13.

We now come to the most crucial issue in the development of resource-bounded measure. If a set X has ν-measure 0 in E, then we want to say that X contains only a "negligible small" part of E. In particular, then, it is critical that E itself not have ν-measure 0 in E. The following theorem establishes this and more.

THEOREM 3.14. *Let ν be a probability measure on* $\mathbf{C}$*, and let* $w \in \{0,1\}^*$*. If* $\nu(w) > 0$*, then* $\mathbf{C}_w$ *does not have* ν*-measure 0 in* E*.*

Proof (sketch). Assume the hypothesis, and let d be a p-ν-martingale. It suffices to show that $\mathbf{C}_w \cap \mathrm{E} \not\subseteq S^\infty[d]$.

Since d is p-computable, there is a function $\hat{d} : \mathbb{N} \times \{0,1\}^* \longrightarrow \mathbb{Q} \cap [0,\infty)$ with the following two properties.

(i) For all $r \in \mathbb{N}$ and $w \in \{0,1\}^*$, $|\hat{d}(r,w) - d(w)| \leq 2^{-r}$.

(ii) There is an algorithm that computes $\hat{d}(r,w)$ in time polynomial in $r + |w|$.

Define a language A recursively as follows. First, for $0 \leq i < |w|$, $[\![s_i \in A]\!] = w[i]$. Next assume that the string $x_i = \chi_A[0..i-1]$ has been defined, where $i \geq |w|$. Then

$$[\![s_i \in A]\!] = [\![\hat{d}(i+1, x_i 1) \leq \hat{d}(i+1, x_i 0)]\!].$$

With the language A so defined, it is easy to check that $A \in \mathbf{C}_w \cap \mathrm{E}$. It is also routine to check that, for all $i \geq |w|$,

$$\begin{aligned} d(x_{i+1}) &\leq \hat{d}(i+1, x_{i+1}) + 2^{-(i+1)} \\ &= \min\left\{\hat{d}(i+1, x_i 0), \hat{d}(i+1, x_i 1)\right\} + 2^{-(i+1)} \\ &\leq \min\{d(x_i 0), d(x_i 1)\} + 2^{-i} \\ &\leq d(x_i) + 2^{-i}. \end{aligned}$$

It follows inductively that, for all $n \geq |w|$,

$$\begin{aligned} d(x_n) &\leq d(w) + \sum_{i=|w|}^{n-1} 2^{-i} \\ &< d(w) + \sum_{i=|w|}^{\infty} 2^{-i} = d(w) + 2^{1-|w|}. \end{aligned}$$

This implies that

$$\limsup_{n \longrightarrow \infty} d(\chi_A[0..n-1]) \leq d(w) + 2^{1-|w|} < \infty,$$

whence $A \notin S^\infty[d]$. □

As in the case of the uniform probability measure [11], more quantitative results on resource-bounded ν-measure can be obtained by considering the *unitary success set*

$$S^1[d] = \bigcup_{\substack{w \\ d(w) \geq 1}} \mathbf{C}_w$$

and the *initial value* $d(\lambda)$ of a p-ν-martingale d. For example, generalizing the arguments in [11] in a straightforward manner, this approach yields a Measure Conservation Theorem for ν-measure (a quantitative extension of Theorem 3.14) and a uniform, resource-bounded extension of the classical first Borel-Cantelli lemma. As these results are not used in the present paper, we refrain from elaborating here.

4. Summable Equivalence . If two probability measures on $\mathbf{C}$ are sufficiently "close" to one another, then the Summable Equivalence Theorem says that the two probability measures are in absolute agreement as to which sets of languages have p-measure 0 and which do not. In this section, we define this notion of "close" and prove this result.

DEFINITION 4.1. *Let ν be a positive probability measure on $\mathbf{C}$, let $A \subseteq \{0,1\}^*$, and let $i \in \mathbb{N}$. Then the i^{th} conditional ν-probability along A is*

$$\nu_A(i+1|i) = \nu(\chi_A[0..i] \mid \chi_A[0..i-1]).$$

DEFINITION 4.2. *Two positive probability measures ν and ν' on $\mathbf{C}$ are summably equivalent, and we write $\nu \approx \nu'$, if for every $A \subseteq \{0,1\}^*$,*

$$\sum_{i=0}^{\infty} |\nu_A(i+1|i) - \nu'_A(i+1|i)| < \infty.$$

It is clear that summable equivalence is an equivalence relation on the collection of all positive probability measures on $\mathbf{C}$. The following fact is also easily verified.

LEMMA 4.3. *Let ν and ν' be positive probability measures on $\mathbf{C}$. If $\nu \approx \nu'$, then ν is strongly positive if and only if ν' is strongly positive.*

The following definition gives the most obvious way to transform a martingale for one probability measure into a martingale for another.

DEFINITION 4.4. *Let ν and ν' be probability measures on $\mathbf{C}$ with ν' positive, and let d be a ν-martingale. Then the canonical adjustment of d to ν' is the ν'-martingale d' defined by*

$$d'(w) = \frac{\nu(w)}{\nu'(w)} d(w)$$

for all $w \in \{0,1\}^*$.

It is trivial to check that the above function d' is indeed a ν'-martingale. The following lemma shows that, for strongly positive probability measures, summable equivalence is a sufficient condition for d' to succeed whenever d succeeds.

LEMMA 4.5. *Let ν and ν' be strongly positive probability measures on $\mathbf{C}$, let d be a ν-martingale, and let d' be the canonical adjustment of d to ν'. If $\nu \approx \nu'$, then $S^\infty[d] \subseteq S^\infty[d']$.*

Proof. Assume the hypothesis, and let $A \in S^\infty[d]$. For each $i \in \mathbb{N}$, let

$$\nu_i = \nu_A(i+1|i), \quad \nu'_i = \nu'_A(i+1|i), \quad \tau_i = \nu_i - \nu'_i.$$

The hypothesis $\nu \approx \nu'$ says that $\sum_{i=0}^{\infty} |\tau_i| < \infty$. In particular, this implies that $\tau_i \longrightarrow 0$ as $i \longrightarrow \infty$, so we have the Taylor approximation

$$\ln \frac{\nu_i}{\nu'_i} = \ln(1 + \frac{\tau_i}{\nu'_i}) = \frac{\tau_i}{\nu'_i} + \mathrm{o}(\frac{\tau_i}{\nu'_i})$$

as $i \longrightarrow \infty$. Thus $|\ln \frac{\nu_i}{\nu'_i}|$ is asymptotically equivalent to $\frac{|\tau_i|}{\nu'_i}$ as $i \longrightarrow \infty$. Since ν' is strongly positive, it follows that $\sum_{i=0}^{\infty} |\ln \frac{\nu_i}{\nu'_i}| < \infty$. Thus, if we let $w_k = \chi_A[0..k-1]$, then there is a positive constant c such that, for all $k \in \mathbb{N}$,

$$c \geq \sum_{i=0}^{k-1} (-\ln \frac{\nu_i}{\nu'_i}) = -\ln \prod_{i=0}^{k-1} \frac{\nu_i}{\nu'_i} = -\ln \frac{\nu(w_k)}{\nu'(w_k)},$$

whence

$$d'(w_k) = \frac{\nu(w_k)}{\nu'(w_k)} d(w_k) \geq e^{-c} d(w_k).$$

Since $A \in S^\infty[d]$, we thus have

$$\limsup_{k \longrightarrow \infty} d'(w_k) \geq \limsup_{k \longrightarrow \infty} e^{-c} d(w_k) = \infty,$$

so $A \in S^\infty[d']$. □

The following useful result is now easily established.

Theorem 4.1 (Summable Equivalence Theorem). If ν and ν' are strongly positive, p-computable probability measures on $\mathbf{C}$ such that $\nu \approx \nu'$, then for every set $X \subseteq \mathbf{C}$,

$$\nu_{\mathrm{p}}(X) = 0 \Longleftrightarrow \nu'_{\mathrm{p}}(X) = 0.$$

Proof. Assume the hypothesis, and assume that $\nu_{\mathrm{p}}(X) = 0$. By symmetry, it suffices to show that $\nu'_{\mathrm{p}}(X) = 0$. Since $\nu_{\mathrm{p}}(X) = 0$, there is a p-computable ν-martingale d such that $X \subseteq S^\infty[d]$. Let d' be the canonical adjustment of d to ν'. Since $d, \nu,$ and ν' are all p-computable, it is easy to see that d' is p-computable. Since $\nu \approx \nu'$, Lemma 4.5 tells us that

$$X \subseteq S^\infty[d] \subseteq S^\infty[d'].$$

Thus $\nu'_{\mathrm{p}}(X) = 0$. □

5. Exact Computation. It is sometimes useful or convenient to work with probability measures that are rational-valued and efficiently computable in an exact sense, with no approximation. This section presents two very easy results identifying situations in which such probability measures are available.

DEFINITION 5.1. *A probability measure ν on $\mathbf{C}$ is exactly* p*-computable if $\nu : \{0,1\}^* \longrightarrow \mathbb{Q} \cap [0,1]$ and there is an algorithm that computes $\nu(w)$ in time polynomial in $|w|$.*

LEMMA 5.2. *For every strongly positive, p-computable probability measure ν on $\mathbf{C}$, there is an exactly p-computable probability measure ν' on $\mathbf{C}$ such that $\nu \approx \nu'$.*

Proof. Let ν be a p-computable probability measure on $\mathbf{C}$, and fix a function $\hat{\nu} : \mathbb{N} \times \{0,1\}^* \longrightarrow \mathbb{Q} \cap [0,1]$ that testifies to the p-computability of ν. Since ν is strongly positive, there is a constant $c \in \mathbb{N}$ such that, for all $w \in \{0,1\}^*$, $2^{-c|w|} \leq \nu(w) \leq 1 - 2^{-c|w|}$. Fix such a c and, for all $w \in \{0,1\}^*$, define

$$\nu'(w0|w) = \min\left\{1, \frac{\hat{\nu}((2c+1)|w| + 3, w0)}{\hat{\nu}((2c+1)|w| + 3, w)}\right\},$$

$$\nu'(w1|w) = 1 - \nu'(w0|w),$$

$$\nu'(w) = \prod_{i=0}^{|w|-1} \nu'(w[0..i]|w[0..i-1]).$$

It is clear that ν' is an exactly p-computable probability measure on $\mathbf{C}$.

Now let $w \in \{0,1\}^*$ and $b \in \{0,1\}$. For convenience, let

$$\begin{aligned}
\delta &= 2^{-(1+c|w|)}, \\
\epsilon &= 2^{-(2c+1)|w|-3}, \\
a_1 &= \nu(wb), \\
a_2 &= \nu(w).
\end{aligned}$$

Note that

$$\hat{\nu}((2c+1)|w|+3, w) \geq \nu(w) - \epsilon > \nu(w) - \delta \geq \delta.$$

It is clear by inspection that $\nu'(wb|w)$ can be written in the form

$$\nu'(wb|w) = \frac{a_1'}{a_2'},$$

where

$$|a_1' - a_1| \leq \epsilon \text{ and } |a_2' - a_2| \leq \epsilon.$$

We thus have

$$\begin{aligned}
|a_1' a_2 - a_1 a_2'| &\leq |a_1' a_2 - a_1 a_2| + |a_1 a_2 - a_1 a_2'| \\
&\leq |a_1' - a_1| + |a_2' - a_2| \\
&\leq 2\epsilon,
\end{aligned}$$

whence

$$\begin{aligned}
|\nu'(wb|w) - \nu(wb|w)| &= \left| \frac{a_1'}{a_2'} - \frac{a_1}{a_2} \right| \\
&= \frac{|a_1' a_2 - a_1 a_2'|}{a_2 a_2'} \\
&\leq 2\epsilon\delta^{-2} \\
&= 2^{-|w|}.
\end{aligned}$$

For all $A \subseteq \{0,1\}^*$, then, we have

$$\sum_{i=0}^{\infty} \left|\nu_A(i+1 \big| i) - \nu_A'(i+1 \big| i)\right| \;\leq\; \sum_{i=0}^{\infty} 2^{-i} \;=\; 2,$$

so $\nu \approx \nu'$. □

For some purposes (including those of this paper), the requirement of p-computability is too weak, because it allows $\nu(w)$ to be computed (or approximated) in time polynomial in $|w|$, which is exponential in the length of the last string decided by w when we regard w as a prefix of a language A. In such situations, the following sort of requirement is often more useful. (We only give the definitions for sequences of biases, i.e., coin-toss probability measures, because this suffices for our purposes in this paper. It is clearly a routine matter to generalize further.)

Definition 5.3.

1. *A P-sequence of biases is a sequence $\vec{\beta} = (\beta_0, \beta_1, \beta_2, \dots)$ of biases $\beta_i \in [0,1]$ for which there is a function*

$$\hat{\beta} : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{Q} \cap [0,1]$$

with the following two properties.
   (i) *For all $i, r \in \mathbb{N}$, $|\hat{\beta}(i,r) - \beta_i| \leq 2^{-r}$.*
   (ii) *There is an algorithm that, for all $i, r \in \mathbb{N}$, computes $\hat{\beta}(i,r)$ in time polynomial in $|s_i| + r$ (i.e., in time polynomial in $\log(i+1) + r$).*
2. *A P-exact sequence of biases is a sequence $\vec{\beta} = (\beta_0, \beta_1, \beta_2, \dots)$ of (rational) biases $\beta_i \in \mathbb{Q} \cap [0,1]$ such that the function $i \longmapsto \beta_i$ is computable in time polynomial in $|s_i|$.*

DEFINITION 5.4. *If $\vec{\alpha}$ and $\vec{\beta}$ are sequences of biases, then $\vec{\alpha}$ and $\vec{\beta}$ are summably equivalent, and we write $\vec{\alpha} \approx \vec{\beta}$, if $\sum_{i=0}^{\infty} |\alpha_i - \beta_i| < \infty$.*

It is clear that $\vec{\alpha} \approx \vec{\beta}$ if and only if $\mu^{\vec{\alpha}} \approx \mu^{\vec{\beta}}$.

LEMMA 5.5. *For every P-sequence of biases $\vec{\beta}$, there is a P-exact sequence of biases $\vec{\beta}'$ such that $\vec{\beta} \approx \vec{\beta}'$.*

Proof. Let $\vec{\beta}$ be a strongly positive P-sequence of biases, and let $\hat{\beta} : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{Q} \cap [0,1]$ be a function that testifies to this fact. For each $i \in \mathbb{N}$, let

$$\beta_i' = \hat{\beta}(i, 2|s_i|),$$

and let $\vec{\beta}' = (\beta_0', \beta_1', \beta_2', \cdots)$. Then $\vec{\beta}'$ is a P-exact sequence of biases, and

$$\begin{aligned}
\sum_{i=0}^{\infty} |\beta_i - \beta_i'| &\leq \sum_{i=0}^{\infty} 2^{-2|s_i|} \\
&\leq \sum_{i=0}^{\infty} 2^{-2\log(i+1)} \\
&= \sum_{i=0}^{\infty} \frac{1}{(i+1)^2} < \infty,
\end{aligned}$$

so $\vec{\beta} \approx \vec{\beta}'$. □

6. Martingale Dilation. In this section we show that certain truth-table reductions can be used to *dilate* martingales for one probability measure into martingales for another, perhaps dissimilar, probability measure on **C**. We first present some terminology and notation on truth-table reductions. (Most of this notation is standard [22], but some is specialized to our purposes.)

A truth-table reduction is a pair (f,g) of algorithms that, given a language A, is used to decide a language $B = F_{(f,g)}(A)$. Given a particular string x, the algorithm f specifies a list of membership queries regarding A, and the algorithm g specifies how the answers to these queries are to be used in deciding whether $x \in B$. The list of queries given by f is nonadaptive,*i.e.*, the entire list is specified prior to the knowing of any of the answers to the queries. More formally, a *truth-table reduction* (briefly, a $\leq_{\mathrm{tt}}$-reduction) is an ordered pair (f,g) of total recursive functions such that for each $x \in \{0,1\}^*$, there exists $n(x) \in \mathbb{Z}^+$ such that the following two conditions hold.

(i) $f(x)$ is (the standard encoding of) an $n(x)$-tuple $(f_1(x), \dots, f_{n(x)}(x))$ of strings $f_i(x) \in \{0,1\}^*$, which are called the *queries* of the reduction (f,g) on input x. We use the notation $Q_{(f,g)}(x) = \{f_1(x), \dots, f_{n(x)}(x)\}$ for the set of such queries.

(ii) $g(x)$ is (the standard encoding of) an $n(x)$-input, 1-output Boolean circuit, called the *truth table* of the reduction (f,g) on input x. We identify $g(x)$ with the Boolean function computed by this circuit, i.e.,

$$g(x) : \{0,1\}^{n(x)} \longrightarrow \{0,1\}.$$

A truth-table reduction (f,g) *induces* the function

$$F_{(f,g)} : \mathbf{C} \longrightarrow \mathbf{C}$$

$$F_{(f,g)}(A) = \left\{x \in \{0,1\}^* \mid g(x)\left([\![f_1(x) \in A]\!] \cdots [\![f_{n(x)}(x) \in A]\!]\right) = 1\right\}.$$

If A and B are languages and (f,g) is a $\leq_{\mathrm{tt}}$-reduction, then (f,g) *reduces* B *to* A, and we write

$$B \leq_{\mathrm{tt}} A \text{ via } (f,g),$$

if $B = F_{(f,g)}(A)$. More generally, if A and B are languages, then B is *truth-table reducible* (briefly, $\leq_{\mathrm{tt}}$-reducible) to A, and we write $B \leq_{\mathrm{tt}} A$, if there exists a $\leq_{\mathrm{tt}}$-reduction (f,g) such that $B \leq_{\mathrm{tt}} A$ via (f,g).

If (f,g) is a $\leq_{\mathrm{tt}}$-reduction, then the function $F_{(f,g)} : \mathbf{C} \longrightarrow \mathbf{C}$ defined above induces a corresponding function

$$F_{(f,g)} : \{0,1\}^* \longrightarrow \{0,1\}^* \cup \mathbf{C}$$

defined as follows. (It is standard practice to use the same notation for these two functions, and no confusion will result from this practice here.) Intuitively, if $A \in \mathbf{C}$ and $w \sqsubseteq A$, then $F_{(f,g)}(w)$ is the largest prefix of $F_{(f,g)}(A)$ such that w answers all queries in this prefix. Formally, let $w \in \{0,1\}^*$, and let

$$A_w = \left\{s_i \;\middle|\; 0 \leq i < |w| \text{ and } w[i] = 1\right\}.$$

If $Q_{(f,g)}(x) \subseteq \{s_0, \ldots s_{|w|-1}\}$ for all $x \in \{0,1\}^*$, then

$$F_{(f,g)}(w) = F_{(f,g)}(A_w).$$

Otherwise,

$$F_{(f,g)}(w) = \chi_{F_{(f,g)}(A_w)}[0..m-1],$$

where m is the greatest nonnegative integer such that

$$\bigcup_{i=0}^{m-1} Q_{(f,g)}(s_i) \subseteq \left\{s_0, \ldots, s_{|w|-1}\right\}$$

Now let (f,g) be a $\leq_{\mathrm{tt}}$-reduction, and let $z \in \{0,1\}^*$. Then the *inverse image* of the cylinder $\mathbf{C}_z$ under the reduction (f,g) is

$$\begin{aligned} F^{-1}_{(f,g)}(\mathbf{C}_z) &= \left\{A \in \mathbf{C} \mid F_{(f,g)}(A) \in \mathbf{C}_z\right\} \\ &= \left\{A \in \mathbf{C} \mid z \sqsubseteq F_{(f,g)}(A)\right\}. \end{aligned}$$

We can write this set in the form

$$F^{-1}_{(f,g)}(\mathbf{C}_z) = \bigcup_{w \in I} \mathbf{C}_w,$$

where I is the set of all strings $w \in \{0,1\}^*$ with the following properties.

(i) $z \sqsubseteq F_{(f,g)}(w)$.

(ii) If w' is a proper prefix of w, then $z \not\sqsubseteq F_{(f,g)}(w')$.

Moreover, the cylinders $\mathbf{C}_w$ in this union are disjoint, so if ν is a probability measure on $\mathbf{C}$, then

$$\nu(F^{-1}_{(f,g)}(\mathbf{C}_z)) = \sum_{w \in I} \nu(w).$$

The following well-known fact is easily verified.

LEMMA 6.1. *If ν is a probability measure on $\mathbf{C}$ and (f,g) is a $\leq_{\mathrm{tt}}$-reduction, then the function*

$$\nu^{(f,g)} : \{0,1\}^* \longrightarrow [0,1]$$

$$\nu^{(f,g)}(z) = \nu(F^{-1}_{(f,g)}(\mathbf{C}_z))$$

is also a probability measure on $\mathbf{C}$.

The probability measure $\nu^{(f,g)}$ of Lemma 6.1 is called the *probability measure induced by ν and (f,g).*

In this paper, we only use the following special type of $\leq_{\mathrm{tt}}$-reduction.

DEFINITION 6.2. *A $\leq_{\mathrm{tt}}$-reduction (f,g) is orderly if, for all $x, y, u, v \in \{0,1\}^*$, if $x < y$, $u \in Q_{(f,g)}(x)$, and $v \in Q_{(f,g)}(y)$, then $u < v$. That is, if x precedes y (in the standard ordering of $\{0,1\}^*$), then every query of (f,g) on input x precedes every query of (f,g) on input y.*

The following is an obvious property of orderly $\leq_{\mathrm{tt}}$-reductions.

LEMMA 6.3. *If ν is a coin-toss probability measure on $\mathbf{C}$ and (f,g) is an orderly $\leq_{\mathrm{tt}}$-reduction, then $\nu^{(f,g)}$ is also a coin-toss probability measure on $\mathbf{C}$.*

Note that, if (f,g) is an orderly $\leq_{\mathrm{tt}}$-reduction, then $F_{(f,g)}(w) \in \{0,1\}^*$ for all $w \in \{0,1\}^*$. Note also that the length of $F_{(f,g)}(w)$ depends only upon the length of w (i.e., $|w| = |w'|$ implies that $|F_{(f,g)}(w)| = |F_{(f,g)}(w')|$). Finally, note that for each $m \in \mathbb{N}$ there exists $l \in \mathbb{N}$ such that $|F_{(f,g)}(0^l)| = m$.

DEFINITION 6.4. *Let (f,g) be an orderly $\leq_{\mathrm{tt}}$-reduction.*

1. *An (f,g)-step is a positive integer l such that $F_{(f,g)}(0^{l-1}) \neq F_{(f,g)}(0^l)$.*
2. *For $k \in \mathbb{N}$, we let step(k) be the least (f,g)-step l such that $l \geq k$.*

The following construction is crucial to the proof of our main theorem.

DEFINITION 6.5. *Let ν be a positive probability measure on $\mathbf{C}$, let (f,g) be an orderly $\leq_{\mathrm{tt}}$-reduction, and let d be a $\nu^{(f,g)}$-martingale. Then the (f,g)-dilation of d is the function*

$$(f,g)\widehat{\ } d : \{0,1\}^* \longrightarrow [0,\infty)$$

$$(f,g)\widehat{\ } d(w) = \sum_{u \in \{0,1\}^{l-k}} d(F_{(f,g)}(wu))\nu(wu|w),$$

where $k = |w|$ and $l = step(k)$.

In other words, $(f,g)\widehat{\ } d(w)$ is the conditional ν-expected value of $d(F_{(f,g)}(w'))$, given that $w \sqsubseteq w'$ and $|w'| = step(|w|)$. We do not include the probability measure ν in the notation $(f,g)\widehat{\ } d$ because ν (being positive) is implicit in d.

Intuitively, the function $(f,g)\widehat{\ } d$ is a strategy for betting on a language A, assuming that d itself is a strategy for betting on the language $F_{(f,g)}(A)$. The following theorem makes this intuition precise.

Theorem 6.1 (Martingale Dilation Theorem). Assume that ν is a positive coin-toss probability measure on $\mathbf{C}$, (f,g) is an orderly $\leq_{\mathrm{tt}}$-reduction, and d is a $\nu^{(f,g)}$-martingale. Then $(f,g)\widehat{\ }d$ is a ν-martingale. Moreover, for every language $A \subseteq \{0,1\}^*$, if d succeeds on $F_{(f,g)}(A)$, then $(f,g)\widehat{\ }d$ succeeds on A.

A very special case of the above result (for strictly increasing $\leq_{\mathrm{m}}^{\mathrm{P}}$-reductions under the uniform probability measure) was developed by Ambos-Spies, Terwijn, and Zheng [2], and made explicit by Juedes and Lutz [8]. Our use of martingale dilation in the present paper is very different from the simple padding arguments of [2, 8].

The following two technical lemmas are used in the proof of Theorem 6.1.

Lemma 6.6. *Assume that ν is a positive coin-toss probability measure on $\mathbf{C}$ and (f,g) is an orderly $\leq_{\mathrm{tt}}$-reduction. Let $F = F_{(f,g)}$, let $w \in \{0,1\}^*$, and assume that $k = |w|$ is an (f,g)-step. Let $l = step(k+1)$. Then, for $b \in \{0,1\}$,*

$$\nu^{(f,g)}(F(w)b|F(w)) = \sum_{\substack{u \in \{0,1\}^{l-k} \\ F(wu) = F(w)b}} \nu(wu|w).$$

Proof. Assume the hypothesis. Then

$$\begin{aligned}
\nu^{(f,g)}(F(w)b) &= \sum_{\substack{w' \in \{0,1\}^k \\ F(w') = F(w)}} \sum_{\substack{u \in \{0,1\}^{l-k} \\ F(w'u) = F(w')b}} \nu(w'u) \\
&= \sum_{\substack{w' \in \{0,1\}^k \\ F(w') = F(w)}} \nu(w') \sum_{\substack{u \in \{0,1\}^{l-k} \\ F(w'u) = F(w')b}} \nu(w'u|w').
\end{aligned}$$

Now, since ν is a coin-toss probability measure, we have $\nu(w'u|w') = \nu(wu|w)$ for each $w' \in \{0,1\}^k$ such that $F(w') = F(w)$. Also, since (f,g) is orderly, the conditions $F(w'u) = F(w')b$ and $F(wu) = F(w)b$ are equivalent for each $u \in \{0,1\}^{l-k}$. Hence,

$$\begin{aligned}
\nu^{(f,g)}(F(w)b) &= \sum_{\substack{w' \in \{0,1\}^k \\ F(w') = F(w)}} \nu(w') \sum_{\substack{u \in \{0,1\}^{l-k} \\ F(wu) = F(w)b}} \nu(wu|w) \\
&= \nu^{(f,g)}(F(w)) \sum_{\substack{u \in \{0,1\}^{l-k} \\ F(wu) = F(w)b}} \nu(wu|w).
\end{aligned}$$

□

Lemma 6.7. *Assume that ν is a positive coin-toss probability measure on $\mathbf{C}$ and (f,g) is an orderly $\leq_{\mathrm{tt}}$-reduction. Let $F = F_{(f,g)}$, and assume that d is a $\nu^{(f,g)}$-martingale. Let $w \in \{0,1\}^*$, assume that $k = |w|$ is an (f,g)-step, and let $l = step(k+1)$. Then*

$$d(F(w)) = \sum_{u \in \{0,1\}^{l-k}} d(F(wu))\nu(wu|w).$$

Proof. Assume the hypothesis. Since d is a $\nu^{(f,g)}$-martingale and $\nu^{(f,g)}(F(w))$ is positive, we have

$$d(F(w)) = \sum_{b=0}^{1} d(F(w)b)\nu^{(f,g)}(F(w)b|F(w)).$$

It follows by Lemma 6.6 that

$$\begin{aligned}
d(F(w)) &= \sum_{b=0}^{1} d(F(w)b) \sum_{\substack{u\,\in\,\{0,1\}^{l-k} \\ F(wu)\,=\,F(w)b}} \nu(wu|w) \\
&= \sum_{b=0}^{1} \sum_{\substack{u\,\in\,\{0,1\}^{l-k} \\ F(wu)\,=\,F(w)b}} d(F(wu))\nu(wu|w) \\
&= \sum_{u\in\{0,1\}^{l-k}} d(F(wu))\nu(wu|w).
\end{aligned}$$

□

Proof of Theorem 6.1. Assume the hypothesis, and let $F = F_{(f,g)}$.

To see that $(f,g)\widehat{\ }d$ is a ν-martingale, let $w \in \{0,1\}^*$, let $k = |w|$, and let $l = \mathit{step}(k+1)$. We have two cases.

CASE I. $\mathit{step}(k) = l$. Then

$$\begin{aligned}
\sum_{b=0}^{1}(f,g)\widehat{\ }d(wb)\nu(wb) &= \sum_{b=0}^{1} \sum_{u\in\{0,1\}^{l-k-1}} d(F(wbu))\nu(wbu|wb)\nu(wb) \\
&= \sum_{b=0}^{1} \sum_{u\in\{0,1\}^{l-k-1}} d(F(wbu))\nu(wbu) \\
&= \sum_{u\in\{0,1\}^{l-k}} d(F(wu))\nu(wu) \\
&= (f,g)\widehat{\ }d(w)\nu(w).
\end{aligned}$$

CASE II. $\mathit{step}(k) < l$. Then k is an (f,g)-step, so $(f,g)\widehat{\ }d(w) = d(F(w))$, whence by Lemma 6.7

$$(f,g)\widehat{\ }d(w)\nu(w) = \sum_{u\in\{0,1\}^{l-k}} d(F(wu))\nu(wu).$$

Calculating as in Case I, it follows that

$$(f,g)\widehat{\ }d(w)\nu(w) = \sum_{b=0}^{1}(f,g)\widehat{\ }d(wb)\nu(wb).$$

This completes the proof that $(f,g)\widehat{\ }d$ is a ν-martingale.

To complete the proof, let $A \subseteq \{0,1\}^*$, and assume that d succeeds on $F(A)$. For each $n \in \mathbb{N}$, let $w_n = \chi_A[0..l_n - 1]$, where l_n is the unique (f,g)-step such that $|F(0^{l_n})| = n$. Then, for all $n \in \mathbb{N}$,

$$(f,g)\widehat{\ }d(w_n) = d(F(w_n)) = d(\chi_{F(A)}[0..n-1]),$$

so

$$\begin{aligned}
\limsup_{k\longrightarrow\infty}(f,g)\widehat{\ }d(\chi_A[0..k-1]) &\geq \limsup_{n\longrightarrow\infty}(f,g)\widehat{\ }d(w_n) \\
&= \limsup_{n\longrightarrow\infty} d(\chi_{F(A)}[0..n-1]) \\
&= \infty.
\end{aligned}$$

Thus $(f,g)\widehat{\ }d$ succeeds on A. □

7. Positive Bias Reduction. In this section, we define and analyze a positive truth-table reduction that encodes an efficient, approximate simulation of one sequence of biases by another.

Intuitively, if $\vec{\alpha}$ and $\vec{\beta}$ are strongly positive sequences of biases, then the positive bias reduction of $\vec{\alpha}$ to $\vec{\beta}$ is a $\leq_{\mathrm{tt}}$-reduction (f, g) that "tries to simulate" the sequence $\vec{\alpha}$ with the sequence $\vec{\beta}$ by causing $\mu^{\vec{\alpha}}$ to be the probability distribution induced by $\mu^{\vec{\beta}}$ and (f, g). In general, this objective will only be approximately achieved, in the sense that the probability distribution induced by $\mu^{\vec{\beta}}$ and (f, g) will actually be a probability distribution $\mu^{\vec{\alpha}'}$, where $\vec{\alpha}'$ is a sequence of biases such that $\vec{\alpha}' \approx \vec{\alpha}$. This situation is depicted schematically in Figure 7.1, where the broken arrow indicates that (f, g) "tries" to reduce $\vec{\alpha}$ to $\vec{\beta}$, while the solid arrow indicates that (f, g) actually reduces $\vec{\alpha}'$ to $\vec{\beta}$.

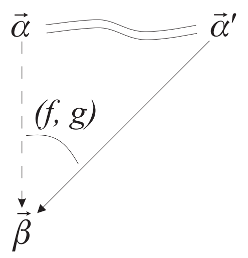

FIG. 7.1. *Schematic depiction of positive bias reduction*

The reduction (f, g) is constructed precisely as follows.

Construction 7.1 (Positive Bias Reduction). Let $\vec{\alpha}$ and $\vec{\beta}$ be strongly positive sequences of biases. Let

$$\delta = \inf \ \{\alpha_i, 1 - \alpha_i, \beta_i, 1 - \beta_i \mid i \in \mathbb{N}\} ,$$

$$c = \lceil \frac{-4 \log e}{\log(1 - \delta^2)} \rceil .$$

For each $x \in \{0, 1\}^*$ and $0 \leq n < 2^{c|x|}$, let $q(x, n) = xy$, where y is the n^{th} element of $\{0, 1\}^{c|x|}$, and let $j(x, n)$ be the index of the string $q(x, n)$, i.e., $s_{j(x,n)} = q(x, n)$. Then the *positive bias reduction of* $\vec{\alpha}$ *to* $\vec{\beta}$ is the ordered pair (f, g) of functions defined by the procedure in Figure 2. (For convenience, the procedure defines additional parameters that are useful in the subsequent analysis.)

The following general remarks will be helpful in understanding Construction 7.1.

(a) The boldface variables $\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots$ denote Boolean inputs to the Boolean function $g(x)$ being constructed. The Boolean function $g(x)$ is an OR of $k(x)$ Boolean functions $h(x, k)$, i.e.,

$$g(x) = \bigvee_{k=0}^{k(x)-1} h(x, k).$$

begin
input $x = s_i$;
$n := 0$;
$g(x,0) := 0$; $\alpha'_i(0) = 0$;
$k := 0$;
while $\alpha'_i(k) < \alpha_i - (i+1)^{-2}$ **do**
    begin
    $h(x,k,0) := 1$; $\gamma_{i,k}(0) := 1$;
    $l := 0$;
    while $\alpha'_i(k) + \gamma_{i,k}(l) - \alpha'_i(k) \cdot \gamma_{i,k}(l) > \alpha_i$ **do**
        begin
        $h(x,k,l+1) := h(x,k,l)$ AND v_n;
        $\gamma_{i,k}(l+1) := \gamma_{i,k}(l) \cdot \beta_{j(x,n)}$;
        $l := l+1$;
        $n := n+1$;
        end ;
    $l(x,k) := l$;
    $h(x,k) := h(x,k,l(x,k))$;
    $\gamma_{i,k} := \gamma_{i,k}(l(x,k))$;
    $g(x,k+1) := g(x,k)$ OR $h(x,k)$;
    $\alpha'_i(k+1) := \alpha'_i(k) + \gamma_{i,k} - \alpha'_i(k) \cdot \gamma_{i,k}$;
    $k := k+1$
    end ;
$k(x) := k$;
$n(x) := n$;
$f(x) := (q(x,0), \ldots, q(x,n(x)-1))$;
$g(x) := g(x,k(x))$;
$\alpha'_i := \alpha'_i(k(x))$
end .

FIG. 7.2. *Construction of positive bias reduction*

The Boolean functions $g(x,0), g(x,1), \ldots$ are preliminary approximations of the Boolean function $g(x)$. In particular,

$$g(x,k) = \bigvee_{j=0}^{k-1} h(x,j)$$

for all $0 \leq k \leq k(x)$. Thus $g(x,0)$ is the constant-0 Boolean function.

(b) The Boolean function $h(x,k)$ is an AND of $l(x,k)$ consecutive input variables. The subscript n is incremented globally so that no input variable appears more than once in $g(x)$. Just as $g(x,k)$ is the k^{th} "partial OR" of $g(x)$, $h(x,k,l)$ is the l^{th} "partial AND" of $h(x,k)$. Thus $h(x,k,0)$ is the constant-1 Boolean function.

(c) The input variables $\boldsymbol{v}_0$, $\boldsymbol{v}_1, \ldots$ of g correspond to the respective queries $q(x,0), q(x,1), \ldots$ of f. If $A = F_{(f,g)}(B)$, then we have $[\![x \in A]\!] = g(x)(\boldsymbol{v}_0 \cdots \boldsymbol{v}_{n(x)-1})$, where each $\boldsymbol{v}_n = [\![q(x,n) \in B]\!]$. If B is chosen according to the sequence of biases $\vec{\beta}$, then $\beta_{j(x,n)}$ is the probability that $\boldsymbol{v}_n = 1$, $\gamma_{i,k}$ is the probability that $h(x,k) = 1$, and α'_i is the probability that $g(x) = 1$. The while-loops ensure that $\alpha_i - (i+1)^{-2} \leq \alpha'_i \leq \alpha_i$.

The following lemmas provide some quantitative analysis of the behavior of Construction 7.1.

LEMMA 7.1. *In Construction 7.1, for all* $x \in \{0,1\}^*$ *and* $0 \le k \le k(x)$,

$$l(x,k) \le 1 + \frac{c|x|}{2\log e}.$$

Proof. Fix such x and k, and let $l^* = l(x,k)$. If $l^* = 0$, the result is trivial, so assume that $l^* > 0$. Then, by the minimality of l^*,

$$\alpha_i'(k) + \gamma_{i,k}(l^* - 1) > \alpha_i,$$

so

$$\gamma_{i,k}(l^* - 1) > \alpha_i - \alpha_i'(k) > (i+1)^{-2},$$

so

$$(i+1)^{-2} < \gamma_{i,k}(l^* - 1) \le (1-\delta)^{l^*-1}.$$

It follows that

$$-2\log(i+1) \le (l^* - 1)\log(1-\delta),$$

whence

$$\begin{aligned} l^* &\le 1 - \frac{2\log(i+1)}{\log(1-\delta)} \\ &\le 1 - \frac{2|x|}{\log(1-\delta^2)} \\ &\le 1 + \frac{c|x|}{2\log e}. \end{aligned}$$

□

LEMMA 7.2. *In the Construction 7.1, for all* $x \in \{0,1\}^*$, *and* $0 \le k \le k(x) - 1$,

$$\alpha_i - \alpha_i'(k) \le (1-\delta^2)^k.$$

Proof. Fix such x and k with $k < k(x) - 1$, and let $l^* = l(x,k)$. Then $\gamma_{i,k}(l^*-1) > \alpha_i - \alpha_i'(k)$, so $\gamma_{i,k} \ge \delta \cdot \gamma_{i,k}(l^* - 1) > \delta \cdot (\alpha_i - \alpha_i'(k))$, whence

$$\begin{aligned} \frac{\alpha_i - \alpha_i'(k+1)}{\alpha_i - \alpha_i'(k)} &= \frac{\alpha_i - (\alpha_i'(k) + \gamma_{i,k} - \alpha_i'(k)\cdot\gamma_{i,k})}{\alpha_i - \alpha_i'(k)} \\ &= \frac{\alpha_i - \alpha_i'(k) - \gamma_{i,k}(1 - \alpha_i'(k))}{\alpha_i - \alpha_i'(k)} \\ &< 1 - \delta \cdot (1 - \alpha_i'(k)) \\ &\le 1 - \delta \cdot (1 - \alpha_i) \\ &\le 1 - \delta^2. \end{aligned}$$

The lemma now follows immediately by induction. □

LEMMA 7.3. *In Construction 7.1, for all $x \in \{0,1\}^*$,*

$$k(x) \leq 1 + \frac{c|x|}{2\log e}$$

Proof. Fix $x \in \{0,1\}^*$. By Lemma 7.2 and the minimality of $k(x)$,

$$\alpha_i - (1-\delta^2)^{k(x)-1} \leq \alpha_i'(k(x)-1) < \alpha_i - (i+1)^{-2},$$

so

$$(1-\delta^2)^{k(x)-1} > (i+1)^{-2},$$

so

$$k(x) < 1 - \frac{2\log(i+1)}{\log(1-\delta^2)} \leq 1 + \frac{c|x|}{2\log e}.$$

□

LEMMA 7.4. *In Construction 7.1, for all $x \in \{0,1\}^*$,*

$$n(x) \leq \left(1 + \frac{c|x|}{2\log e}\right)^2 \leq 2^{c|x|}.$$

Proof. Let $x \in \{0,1\}^*$. Then

$$n(x) = \sum_{k=0}^{k(x)-1} l(x,k),$$

so by Lemmas 7.1, 7.3, and the bound $1 + t \leq e^t$,

$$n(x) \leq \left(1 + \frac{c|x|}{2\log e}\right)^2 \leq e^{\frac{c|x|}{\log e}} = 2^{c|x|}.$$

□

DEFINITION 7.5. *Let (f,g) be a $\leq_{\mathrm{tt}}$-reduction.*

1. *(f,g) is positive (briefly, a $\leq_{\mathrm{pos-tt}}$-reduction) if, for all $A, B \subseteq \{0,1\}^*$, $A \subseteq B$ implies$F_{(f,g)}(A) \subseteq F_{(f,g)}(B)$.*
2. *(f,g) is polynomial-time computable (briefly, a $\leq_{\mathrm{tt}}^{\mathrm{P}}$-reduction) if the functions f and g are computable in polynomial time.*
3. *(f,g) is polynomial-time computable with linear-bounded queries (briefly, a $\leq_{\mathrm{tt}}^{\mathrm{P,lin}}$-reduction) if (f,g) is a $\leq_{\mathrm{tt}}^{\mathrm{P}}$-reduction and there is a constant $c \in \mathbb{N}$ such that, for all $x \in \{0,1\}^*$, $Q_{(f,g)}(x) \subseteq \{0,1\}^{\leq c(1+|x|)}$.*

Of course, a $\leq_{\mathrm{pos-tt}}^{\mathrm{P,lin}}$-reduction is a $\leq_{\mathrm{tt}}$-reduction with all the above properties.

The following result presents the properties of the positive bias reduction that are used in the proof of our main theorem.

Theorem 7.2 (Positive Bias Reduction Theorem). Let $\vec{\alpha}$ and $\vec{\beta}$ be strongly positive, P-exact sequences of biases, and let (f,g) be the positive bias reduction of $\vec{\alpha}$ to $\vec{\beta}$. Then (f,g) is an orderly $\leq_{\mathrm{pos-tt}}^{\mathrm{P,lin}}$-reduction, and the probability measure induced by $\mu^{\vec{\beta}}$ and (f,g) is a coin-toss probability measure $\mu^{\vec{\alpha'}}$, where $\vec{\alpha} \approx \vec{\alpha'}$.

Proof. Assume the hypothesis. By inspection and Lemma 7.4, the pair (f,g) is an orderly $\leq^{\mathrm{P,lin}}_{\mathrm{pos-tt}}$-reduction. (Lemma 7.4 also ensures that $f(x)$ is well-defined.) The reduction is also positive, since only AND's and OR's are used in the construction of $g(x)$. Thus (f,g) is an orderly $\leq^{\mathrm{P,lin}}_{\mathrm{pos-tt}}$-reduction.

By remark (c) following Construction 7.1, the probability measure induced by $\mu^{\vec{\beta}}$ and (f,g) is the coin-toss probability measure $\mu^{\vec{\alpha}'}$, where $\vec{\alpha}' = (\alpha'_0, \alpha'_1, \dots)$ is defined in the construction. Moreover,

$$\sum_{i=0}^{\infty} |\alpha_i - \alpha'_i| \leq \sum_{i=0}^{\infty} (i+1)^{-2} < \infty,$$

so $\vec{\alpha} \approx \vec{\alpha}'$. □

8. Equivalence for Complexity Classes. Many important complexity classes, including P, NP, co-NP, R, BPP, AM, P/Poly, PH, PSPACE, etc., are known to be closed under $\leq^{\mathrm{P}}_{\mathrm{pos-tt}}$-reductions, hence certainly under $\leq^{\mathrm{P,lin}}_{\mathrm{pos-tt}}$-reductions. The following theorem, which is the main result of this paper, says that the p-measure of such a class is somewhat insensitive to certain changes in the underlying probability measure. The proof is now easy, given the machinery of the preceding sections.

Theorem 8.1 (Bias Equivalence Theorem). Assume that $\vec{\alpha}$ and $\vec{\beta}$ are strongly positive P-sequences of biases, and let $\mathcal{C}$ be a class of languages that is closed under $\leq^{\mathrm{P,lin}}_{\mathrm{pos-tt}}$-reductions. Then

$$\mu^{\vec{\alpha}}_{\mathrm{p}}(\mathcal{C}) = 0 \Longleftrightarrow \mu^{\vec{\beta}}_{\mathrm{p}}(\mathcal{C}) = 0.$$

Proof. Assume the hypothesis, and assume that $\mu^{\vec{\alpha}}_{\mathrm{p}}(\mathcal{C}) = 0$. By symmetry, it suffices to show that $\mu^{\vec{\beta}}_{\mathrm{p}}(\mathcal{C}) = 0$.

The proof follows the scheme depicted in Figure 3. By Lemma 5.5, there exist P-exact sequences $\vec{\alpha}'$ and $\vec{\beta}'$ such that $\vec{\alpha} \approx \vec{\alpha}'$ and $\vec{\beta} \approx \vec{\beta}'$. Let (f,g) be the positive bias reduction of $\vec{\alpha}'$ to $\vec{\beta}'$. Then, by the Positive Bias Reduction Theorem (Theorem 7.2), (f,g) is an orderly $\leq^{\mathrm{P,lin}}_{\mathrm{pos-tt}}$-reduction, and the probability measure induced by $\mu^{\vec{\beta}}$ and (f,g) is $\mu^{\vec{\alpha}''}$, where $\vec{\alpha}' \approx \vec{\alpha}''$.

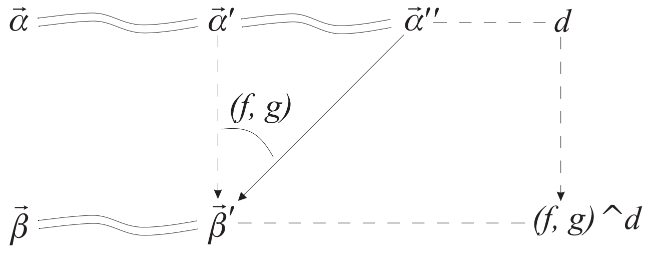

FIG. 8.1. *Scheme of proof of Bias Equivalence Theorem*

Since $\vec{\alpha} \approx \vec{\alpha}' \approx \vec{\alpha}''$ and $\mu^{\vec{\alpha}}_{\mathrm{p}}(\mathcal{C}) = 0$, the Summable Equivalence Theorem (Theorem 4.1) tells us that there is a p-$\vec{\alpha}''$-martingale d such that $\mathcal{C} \subseteq S^{\infty}[d]$. By the Martingale Dilation Theorem (Theorem 6.1), the function $(f,g)\hat{\ }d$ is then a $\vec{\beta}'$-martingale. In fact, it easily checked that $(f,g)\hat{\ }d$ is a p-$\vec{\beta}'$-martingale.

Now let $A \in \mathcal{C}$. Then, since $\mathcal{C}$ is closed under $\leq^{\mathrm{P,lin}}_{\mathrm{pos-tt}}$-reductions, $F_{(f,g)}(A) \in \mathcal{C} \subseteq S^\infty[d]$. It follows by the Martingale Dilation Theorem that $A \in S^\infty[(f,g)\hat{\ }d]$. Thus $\mathcal{C} \subseteq S^\infty[(f,g)\hat{\ }d]$. Since $(f,g)\hat{\ }d$ is a p-$\vec{\beta}'$-martingale, this shows that $\mu_{\mathrm{p}}^{\vec{\beta}'}(\mathcal{C}) = 0$. Finally, since $\vec{\beta} \approx \vec{\beta}'$, it follows by the Summable Equivalence Theorem that $\mu_{\mathrm{p}}^{\vec{\beta}}(X) = 0$.
□

It is clear that the Bias Equivalence Theorem remains true if the resource bound on the measure is relaxed. That is, the analogs of Theorem 8.1 for p_2-measure, pspace-measure, rec-measure, constructive measure, and classical measure all immediately follow. We conclude by noting that the analogs of Theorem 8.1 for measure in E and measure in E_2 also immediately follow.

COROLLARY 8.1. *Under the hypothesis of Theorem 8.1,*

$$\mu^{\vec{\alpha}}(\mathcal{C}|\mathrm{E}) = 0 \Longleftrightarrow \mu^{\vec{\beta}}(\mathcal{C}|\mathrm{E}) = 0$$

and

$$\mu^{\vec{\alpha}}(\mathcal{C}|\mathrm{E}_2) = 0 \Longleftrightarrow \mu^{\vec{\beta}}(\mathcal{C}|\mathrm{E}_2) = 0.$$

Proof. If $\mathcal{C}$ is closed under $\leq^{\mathrm{P,lin}}_{\mathrm{pos-tt}}$-reductions, then so are the classes $\mathcal{C} \cap \mathrm{E}$ and $\mathcal{C} \cap \mathrm{E}_2$. □

9. Conclusion. Our main result, the Bias Equivalence Theorem, says that every strongly positive, P-computable, coin-toss probability measure ν is *equivalent* to the uniform probability measure μ, in the sense that

$$\nu_{\mathrm{p}}(\mathcal{C}) = 0 \Longleftrightarrow \mu_{\mathrm{p}}(\mathcal{C}) = 0$$

for all classes $\mathcal{C} \in \Gamma$, where Γ is a family that contains P, NP, co-NP, R, BPP, P/Poly, PH and many other classes of interest. It would be illuminating to learn more about which probability measures are, and which probability measures are not, equivalent to μ in this sense.

It would also be of interest to know whether the Summable Equivalence Theorem can be strengthened. Specifically, say that two sequences of biases $\vec{\alpha}$ and $\vec{\beta}$ are *square-summably equivalent*, and write $\vec{\alpha} \approx^2 \vec{\beta}$, if $\sum_{i=0}^{\infty}(\alpha_i - \beta_i)^2 < \infty$. A classical theorem of Kakutani [9] says that, if $\vec{\alpha}$ and $\vec{\beta}$ are strongly positive sequences of biases such that $\vec{\alpha} \approx^2 \vec{\beta}$, then for every set $C \subseteq \mathbf{C}$, X has (classical) $\vec{\alpha}$-measure 0 if and only if X has $\vec{\beta}$-measure 0. A constructive improvement of this theorem by Vovk [28] says that, if $\vec{\alpha}$ and $\vec{\beta}$ are strongly positive, computable sequences of biases such that $\vec{\alpha} \approx^2 \vec{\beta}$, then for every set $X \subseteq \mathbf{C}$, X has constructive $\vec{\alpha}$-measure 0 if and only if X has constructive $\vec{\beta}$-measure 0. (The Kakutani and Vovk theorems are more general than this, but for the sake of brevity, we restrict the present discussion to coin-toss probability measures.) The Summable Equivalence Theorem is stronger than these results in one sense, but weaker in another. It is stronger in that it holds for p-measure, but it is weaker in that it requires the stronger hypothesis that $\vec{\alpha} \approx \vec{\beta}$. We thus ask whether there is a "square-summable equivalence theorem" for p-measure. That is, if $\vec{\alpha}$ and $\vec{\beta}$ are strongly positive, p-computable sequences of biases such that $\vec{\alpha} \approx^2 \vec{\beta}$, is it necessarily the case that, for every set $X \subseteq \mathbf{C}$, X has p-$\vec{\alpha}$-measure 0 if and only if X has p-$\vec{\beta}$-measure 0? (Note: Kautz [10] has very recently answered this question affirmatively.)

Acknowledgments. We thank Giora Slutzki, Martin Strauss, and other participants in the ISU Information and Complexity Seminar for useful remarks and suggestions. We especially thank Giora Slutzki for suggesting a simplified presentation of Lemma 4.5. We also thank the referee for a very careful reading of our manuscript.

REFERENCES

[1] K. Ambos-Spies, E. Mayordomo, and X. Zheng. A comparison of weak completeness notions. In *Proceedings of the Eleventh IEEE Conference on Computational Complexity*. IEEE Computer Society Press, 1996, pp. 171–178.

[2] K. Ambos-Spies, S. A. Terwijn, and X. Zheng. Resource bounded randomness and weakly complete problems. *Theoretical Computer Science* 172(1977):195–207.

[3] J. Cai and A. L. Selman. Fine separation of average time complexity classes. In *Proceedings of the Thirteenth Symposium on Theoretical Aspects of Computer Science*. Springer-Verlag, 1996. pp. 331–343.

[4] R. I. Freidzon. Families of recursive predicates of measure zero. translated in *Journal of Soviet Mathematics*, 6(1976):449–455, 1972.

[5] P. R. Halmos. *Measure Theory*. Springer-Verlag, 1950.

[6] D. W. Juedes. Weakly complete problems are not rare. *Computational Complexity* 5(1995):267–283.

[7] D. W. Juedes and J. H. Lutz. The complexity and distribution of hard problems. *SIAM Journal on Computing*, 24(2):279–295, 1995.

[8] D. W. Juedes and J. H. Lutz. Weak completeness in E and E_2. *Theoretical Computer Science*, 143:149–158, 1995.

[9] S. Kakutani. On the equivalence of infinite product measures. *Annals of Mathematics*, 49:214–224, 1948.

[10] S. M. Kautz. Resource-bounded randomness and compressibility with respect to nonuniform measures. *Proceedings of the International Workshop on Randomization and Approximation Techniques in Computer Science*. Springer-Verlag, 1997, pp. 197–211.

[11] J. H. Lutz. Almost everywhere high nonuniform complexity. *Journal of Computer and System Sciences*, 44:220–258, 1992.

[12] J. H. Lutz. The quantitative structure of exponential time. In L.A. Hemaspaandra and A.L. Selman (eds.), *Complexity Theory Retrospective II*, Springer-Verlag, 1997, pp. 225–254.

[13] J. H. Lutz. Weakly hard problems. *SIAM Journal on Computing*, 24:1170–1189, 1995.

[14] J. H. Lutz. Observations on measure and lowness for Δ_2^{P}. *Theory of Computing Systems* 30(1997):429–442.

[15] J. H. Lutz and E. Mayordomo. Measure, stochasticity, and the density of hard languages. *SIAM Journal on Computing*, 23:762–779, 1994.

[16] J. H. Lutz and E. Mayordomo. Genericity, measure, and inseparable pairs, 1996. In preparation.

[17] J. H. Lutz and E. Mayordomo. Cook versus Karp-Levin: Separating completeness notions if NP is not small. *Theoretical Computer Science* 164(1996):141–163.

[18] E. Mayordomo. Almost every set in exponential time is P-bi-immune. *Theoretical Computer Science*, 136(2):487–506, 1994.

[19] K. Mehlhorn. The "almost all" theory of subrecursive degrees is decidable. In *Proceedings of the Second Colloquium on Automata, Languages, and Programming*, pages 317–325. Springer Lecture Notes in Computer Science, vol. 14, 1974.

[20] J. C. Oxtoby. *Measure and Category*. Springer-Verlag, second edition, 1980.

[21] K. W. Regan, D. Sivakumar, and J. Cai. Pseudorandom generators, measure theory, and natural proofs. In *36th IEEE Symposium on Foundations of Computer Science*, pages 26–35. IEEE Computer Society Press, 1995.

[22] H. Rogers, Jr. *Theory of Recursive Functions and Effective Computability*. McGraw - Hill, New York, 1967.

[23] C. P. Schnorr. Klassifikation der Zufallsgesetze nach Komplexität und Ordnung. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 16:1–21, 1970.

[24] C. P. Schnorr. A unified approach to the definition of random sequences. *Mathematical Systems Theory*, 5:246–258, 1971.

[25] C. P. Schnorr. Zufälligkeit und Wahrscheinlichkeit. *Lecture Notes in Mathematics*, 218, 1971.

[26] C. P. Schnorr. Process complexity and effective random tests. *Journal of Computer and System Sciences*, 7:376–388, 1973.

[27] M. van Lambalgen. *Random Sequences.* PhD thesis, Department of Mathematics, University of Amsterdam, 1987.

[28] V. G. Vovk. On a randomness criterion. *Soviet Mathematics Doklady*, 35:656–660, 1987.

[29] A. K. Zvonkin and L. A. Levin. The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematical Surveys*, 25:83–124, 1970.