Bias Invariance of Small Upper Spans[1]

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

Martin Strauss
AT&T Labs
180 Park Ave., P.O. Box 971
Florham Park, NJ 07932
U.S.A.

Abstract

The resource-bounded measures of certain classes of languages are shown to be invariant under certain changes in the underlying probability measure. Specifically, for any real number $\delta > 0$, any polynomial-time computable sequence $\vec{\beta} = (\beta_0, \beta_1, \dots)$ of biases $\beta_i \in [\delta, 1-\delta]$, and any class $\mathcal{C}$ of languages that is closed *upwards or downwards* under positive, polynomial-time truth-table resuctions with linear bounds on number and length of queries, it is shown that the following two conditions are equivalent.

(1) $\mathcal{C}$ has p-measure 0 relative to the probability measure given by $\vec{\beta}$.

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

The analogous equivalences are established for measure in E and measure in E_2. (Breutzmann and Lutz [5] established this invariance for classes $\mathcal{C}$ that are closed downwards under slightly more powerful reductions, but nothing was known about invariance for classes that are closed upwards.) The proof introduces two new techniques, namely, the *contraction* of a martingale for one probability measure to a martingale for an induced probability measure, and a new, improved *positive bias reduction* of one bias sequence to another. Consequences for the BPP versus E problem and small span theorems are derived.

1 Introduction

Until recently, all research on the measure-theoretic structure of complexity classes has been restricted to the uniform probability measure. This is the probability measure μ that intuitively corresponds to a random experiment in which a language $A \subseteq \{0,1\}^*$ is chosen probabilistically, using an independent toss of a fair coin to decide whether each string is in A. When effectivized by the methods of resource-bounded measure [15], μ induces measure-theoretic structure on $\mathrm{E} = \mathrm{DTIME}(2^{\mathrm{linear}})$, $\mathrm{E}_2 = \mathrm{DTIME}(2^{\mathrm{polynomial}})$, and other

[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.

complexity classes. Investigations of this structure by a number of researchers have yielded many new insights over the past seven years. The recent surveys [3, 16, 6] describe much of this work.

There are several reasons for extending our investigation of resource-bounded measure to a wider variety of probability measures. First, such variety is essential in cryptography, computational learning, algorithmic information theory, average-case complexity, and other potential application areas. Second, applications of the probabilistic method [2] often require use of non-uniform probability measures, and this is likely to hold for the resource-bounded probabilistic method [18, 16] as well. Third, resource-bounded measure based on non-uniform probability measures provides new methods for proving results about resource-bounded measure based on the uniform probability measure [5].

Motivated by such considerations, Breutzmann and Lutz [5] initiated the study of resource-bounded measure based on an arbitrary (Borel) probability measure ν on the Cantor space $\mathbf{C}$ (the set of all languages). (Precise definitions of these and other terms appear in Appendix A.) Kautz [13] and Lutz [17] have furthered this study in different directions, and the present paper is another contribution.

The principal focus of the paper [5] is the circumstances under which the ν-measure of a complexity class $\mathcal{C}$ is invariant when the probability measure ν is replaced by some other probability measure ν'. For an *arbitrary* class $\mathcal{C}$ of languages, such invariance can only occur if ν and ν' are fairly close to one another: Extending results of Kakutani [12], Vovk [24], and Breutzmann and Lutz [5], Kautz [13] has shown that the "square-summable equivalence" of ν and ν' is sufficient to ensure $\nu_{\mathrm{p}}(\mathcal{C}) = 0 \iff \nu'_{\mathrm{p}}(\mathcal{C}) = 0$, but very little more can be said when $\mathcal{C}$ is arbitrary.

Fortunately, complexity classes have more structure than arbitrary classes. Most complexity classes of interest, including P, NP, coNP, R, BPP, AM, P/Poly, PH, etc., are closed downwards under positive, polynomial-time truth-table reductions ($\leq^{\mathrm{P}}_{\mathrm{pos-tt}}$-reductions), and their intersections with E are closed downward under $\leq^{\mathrm{P}}_{\mathrm{pos-tt}}$-reductions with linear bounds on the length of queries ($\leq^{\mathrm{P,lin}}_{\mathrm{pos-tt}}$-reductions). Breutzmann and Lutz [5] proved that every class $\mathcal{C}$ with these closure properties enjoys a substantial amount of invariance in its measure. Specifically, if $\mathcal{C}$ is any such class and $\vec{\beta}$ and $\vec{\beta}'$ are strongly positive, P-sequences of biases, then the equivalences

$$\begin{aligned}
\mu^{\vec{\beta}}_{\mathrm{p}}(\mathcal{C}) = 0 &\iff \mu^{\vec{\beta}'}_{\mathrm{p}}(\mathcal{C}) = 0, \\
\mu^{\vec{\beta}}(\mathcal{C}|\mathrm{E}) = 0 &\iff \mu^{\vec{\beta}'}(\mathcal{C}|\mathrm{E}) = 0, \\
\mu^{\vec{\beta}}(\mathcal{C}|\mathrm{E}_2) = 0 &\iff \mu^{\vec{\beta}'}(\mathcal{C}|\mathrm{E}_2) = 0
\end{aligned} \tag{1}$$

hold, where $\mu^{\vec{\beta}}$ and $\mu^{\vec{\beta}'}$ are the probability measures corresponding to the bias sequences $\vec{\beta}$ and $\vec{\beta}'$, respectively.

Our primary concern in the present paper is to extend this bias invariance to classes that are closed *upwards* under some type $\leq_r^{\mathrm{P}}$ of polynomial reductions. We have two reasons for interest in this question. First and foremost, many recent investigations in complexity theory focus on the resource-bounded measure of the *upper* $\leq_r^{\mathrm{P}}$*-span*

$$\mathrm{P}_r^{-1}(A) = \{B | A \leq_r^{\mathrm{P}} B\}$$

of a language A. Such investigations include work on small span theorems [9, 14, 4, 11, 7] and work on the BPP versus E question [1, 7, 8]. In general, the upper $\leq_r^{\mathrm{P}}$-span of a language is closed upwards, but not downwards, under $\leq_r^{\mathrm{P}}$-reductions.

Our second reason for interest in upward closure conditions is that the above-mentioned results of Breutzmann and Lutz [5] do *not* fully establish the invariance of measures of complexity classes under the indicated changes of bias sequences. For example, if $\vec{\beta}$ is an arbitrary strongly positive P-sequence of biases, the results of [5] show that

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

but they do *not* show that

$$\mu^{\vec{\beta}}(\mathcal{C}|\mathrm{E}) = 1 \iff \mu(\mathcal{C}|\mathrm{E}) = 1.$$

In general, the condition $\nu(\mathcal{C}|\mathrm{E}) = 1$ is equivalent to $\nu(\mathcal{C}^c|\mathrm{E}) = 0$, where $\mathcal{C}^c$ is the complement of $\mathcal{C}$. Since $\mathcal{C}$ is closed downwards under $\leq_r^{\mathrm{P}}$-reductions if and only if $\mathcal{C}^c$ is closed upwards under $\leq_r^{\mathrm{P}}$-reductions, we are again led to consider upward closure conditions.

Our main theorem, the Bias Invariance Theorem, states that, if $\mathcal{C}$ is any class of languages that is closed *upwards or downwards* under positive, polynomial-time, truth-table reductions with linear bounds on number and length of queries ($\leq_{\mathrm{pos-lin-tt}}^{\mathrm{P,lin}}$-reductions), and if $\vec{\beta}$ and $\vec{\beta}'$ are strongly positive P-sequences of biases, then the equivalences (1) above hold. The proof introduces two new techniques, namely, the *contraction* of a martingale for one probability measure to a martingale for an induced probability measure (dual to the martingale *dilation* technique introduced in [5]) and a new, improved *positive bias reduction* of one bias sequence to another.

We also note three easy consequences of our Bias Invariance Theorem. First, in combination with work of Allender and Strauss [1] and Buhrman, van Melkebeek, Regan, Sivakumar, and Strauss [8], it implies that, if there is *any* strongly positive P-sequence of biases $\vec{\beta}$ such that the complete $\leq_{\mathrm{T}}^{\mathrm{P}}$-degree for E_2 does not have $\mu^{\vec{\beta}}$-measure 1 in E_2, then $\mathrm{E} \not\subseteq \mathrm{BPP}$. Second, in combination with the work of Regan, Sivakumar, and Cai [19], it implies that, for any reasonable complexity class $\mathcal{C}$, if there exists a strongly positive P-sequence of biases $\vec{\beta}$ such that $\mathcal{C}$ has $\mu^{\vec{\beta}}$-measure 1 in E, then $\mathrm{E} \subseteq \mathcal{C}$ (and similarly for E_2). Third, if $\leq_r^{\mathrm{P}}$ is any polynomial reducibility such that $A \leq_{\mathrm{pos-lin-tt}}^{\mathrm{P,lin}} B$ implies $A \leq_r^{\mathrm{P}} B$, and if $\vec{\beta}$ is a strongly positive P-sequence of biases, then the small span theorem for $\leq_r^{\mathrm{P}}$-reductions holds with respect to $\mu^{\vec{\beta}}$ if and only if it holds with respect to μ. Tantalizingly, this hypothesis places $\leq_r^{\mathrm{P}}$ "just beyond" the small span theorem of Buhrman and van Melkebeek [7], which is the strongest small span theorem proven to date for exponential time.

2 Preliminaries

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 $[\![\phi]\!] =$ **if** ϕ **then** 1 **else** 0.

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. The *complement* of a set X of languages is $X^c = \mathbf{C} - X$.

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

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

3 Martingale Contraction

Given a positive coin-toss probability measure ν, an orderly truth-table reduction (f,g), and a $\nu^{(f,g)}$-martingale d (where $\nu^{(f,g)}$ is the probability measure induced by ν and (f,g)), Breutzmann and Lutz [5] showed how to construct a ν-martingale $(f,g)\widehat{\ }d$, called the (f,g)-*dilation* of d, such that $(f,g)\widehat{\ }d$ succeeds on A whenever d succeeds on $F_{(f,g)}(A)$. (See [5] or Appendix B for notation and terminology involving truth-table reductions.) In this section we present a dual of this construction. Given ν and (f,g) as above and a ν-martingale d, we show how to construct a $\nu^{(f,g)}$-supermartingale $(f,g)\smile d$, called the (f,g)-*contraction* of d, such that $(f,g)\smile d$ succeeds on A whenever d succeeds strongly on every element of $F^{-1}_{(f,g)}(\{A\})$.

The notion of an (f,g)-step, introduced in [5], will also be useful here.

Definition. 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 $\mathit{step}(k)$ be the least (f,g)-step l such that $l \geq k$.
3. For $v, w \in \{0,1\}^*$, we write $v \succ w$ to indicate that $w \sqsubseteq v$ and $|v| = \mathit{step}(|w|+1)$. (That is, $v \succ w$ means that v is a proper extension of w to the next step.)

Our construction makes use of a special-purpose inverse of $F_{(f,g)}$ that depends on both (f,g) and d.

Definition. Let (f,g) be an orderly $\leq_{\mathrm{tt}}$-reduction, let ν be a positive probability measure on $\mathbf{C}$, and let d be a ν-martingale. Then the partial function

$$F^{-1}_{(f,g),d} : \{0,1\}^* \longrightarrow \{0,1\}^*$$

is defined recursively as follows.

(i) $F^{-1}_{(f,g),d}(\lambda) = \lambda$.

(ii) For $w \in \{0,1\}^*$ and $b \in \{0,1\}$, $F^{-1}_{(f,g),d}(wb)$ is the lexicographically first string $v \succ F^{-1}_{(f,g),d}(w)$ such that $F_{(f,g)}(v) = wb$ and, for all $v' \succ F^{-1}_{(f,g),d}(w)$ such that $F_{(f,g)}(v') = wb$, we have $d(v) \leq d(v')$. (That is, v minimizes $d(v)$ on the set of all $v \succ F^{-1}_{(f,g),d}(w)$ satisfying $F_{(f,g)}(v) = wb$.)

Note that the function $F^{-1}_{(f,g),d}$ is strictly monotone (i.e., $w \sqsubsetneq w'$ implies that $F^{-1}_{(f,g),d}(w) \sqsubsetneq F^{-1}_{(f,g),d}(w')$, provided that these values exist), whence it extends naturally to a partial function

$$F^{-1}_{(f,g),d} : \mathbf{C} \longrightarrow \mathbf{C}.$$

It is easily verified that $F^{-1}_{(f,g),d}$ inverts $F_{(f,g)}$ in the sense that, for all $x \in \{0,1\}^* \cup \mathbf{C}$, $F^{-1}_{(f,g),d}$ finds a preimage of $F_{(f,g)}(x)$, i.e.,

$$F_{(f,g)}(F^{-1}_{(f,g),d}(F_{(f,g)}(x))) = F_{(f,g)}(x).$$

We now define the (f,g)-contraction of a ν-martingale d.

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

$$(f,g)_\smile d : \{0,1\}^* \longrightarrow \{0,1\}^*$$

defined as follows.

(i) $(f,g)_\smile d(\lambda) = d(\lambda)$.

(ii) For $w \in \{0,1\}^*$ and $b \in \{0,1\}$,

$$(f,g)_\smile d(wb) = \begin{cases} d(F^{-1}_{(f,g),d}(wb)) & \text{if } d(F^{-1}_{(f,g),d}(wb)) \text{ is defined} \\ 2 \cdot (f,g)_\smile d(w) & \text{otherwise.} \end{cases}$$

Theorem 3.1 (Martingale Contraction Theorem). Assume that ν is a positive probability measure on $\mathbf{C}$, (f,g) is an orderly $\leq_{\mathrm{tt}}$-reduction, and d is a ν-martingale. Then $(f,g)_\smile d$ is a $\nu^{(f,g)}$-supermartingale. Moreover, for every language $A \subseteq \{0,1\}^*$, if $F^{-1}_{(f,g)}(\{A\}) \subseteq S^\infty_{\mathrm{str}}[d]$, then $A \in S^\infty[(f,g)_\smile d]$.

4 Bias Invariance

In this section we present our main results.

Definition. Let (f, g) be a $\leq_{\text{tt}}$-reduction.

1. (f, g) is *positive* (briefly, a $\leq_{\text{pos-tt}}$*-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 $\leq_{\text{tt}}^{\text{P}}$-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 $\leq_{\text{tt}}^{\text{P,lin}}$-*reduction*) if (f, g) is a $\leq_{\text{tt}}^{\text{P}}$-reduction and there is a constant $c \in \mathbb{N}$ such that, for all $x \in \{0,1\}^*$, $Q_{(f,g)}(x) \subseteq \{0,1\}^{\leq c(1+|x|)}$.
4. (f, g) is *polynomial-time computable with a linear number of queries* (briefly, a $\leq_{\text{lin-tt}}^{\text{P}}$-*reduction*) if (f, g) is a $\leq_{\text{tt}}^{\text{P}}$-reduction and there is a constant $c \in \mathbb{N}$ such that, for all $x \in \{0,1\}^*$, $|Q_{(f,g)}(x)| \leq c(1 + |x|)$.

Of course, a $\leq_{\text{pos-tt}}^{\text{P,lin}}$-reduction is a $\leq_{\text{tt}}$-reduction with properties 1–3, and a $\leq_{\text{pos-lin-tt}}^{\text{P,lin}}$-reduction is a $\leq_{\text{tt}}$-reduction with properties 1–4.

We now present the Positive Bias Reduction Theorem. This strengthens the identically-named result of Breutzmann and Lutz [5] by giving a $\leq_{\text{pos-lin-tt}}^{\text{P,lin}}$-reduction in place of a $\leq_{\text{pos-tt}}^{\text{P,lin}}$-reduction. This technical improvement, which is essential for our purposes here, requires a substantially different construction. Details appear in Appendix D.

Theorem 4.1 (Positive Bias Reduction Theorem). Let $\vec{\beta}$ and $\vec{\beta}'$ be strongly positive, P-exact sequences of biases, and let (f, g) be the reduction defined in Appendix D. Then (f, g) is an orderly $\leq_{\text{pos-lin-tt}}^{\text{P,lin}}$-reduction, and the probability measure induced by $\mu^{\vec{\beta}}$ and (f, g) is a coin-toss probability measure $\mu^{\vec{\beta}''}$, where $\vec{\beta}'' \approx \vec{\beta}'$.

The following result is our main theorem.

Theorem 4.2 (Bias Invariance Theorem). Assume that $\vec{\beta}$ and $\vec{\beta}'$ are strongly positive P-sequences of biases, and let $\mathcal{C}$ be a class of languages that is closed upwards or downwards under $\leq^{\mathrm{P,lin}}_{\mathrm{pos-lin-tt}}$-reductions. Then

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

The "downwards" part of Theorem 4.2 is a technical improvement of the Bias Equivalence Theorem of [5] from $\leq^{\mathrm{P,lin}}_{\mathrm{pos-tt}}$-reductions to $\leq^{\mathrm{P,lin}}_{\mathrm{pos-lin-tt}}$-reductions. The proof of this improvement is simply the proof in [5] with Theorem 4.1 used in place of its predecessor in [5].

The "upwards" part of Theorem 4.2 is entirely new. The proof of this result is similar to the proof of the Bias Equivalence Theorem in [5], but now in addition to using our improved Positive Bias Reduction Theorem, we use the Martingale Contraction Theorem of section 3 in place of the Martingale Dilation Theorem of [5]. We also note that the linear bound on number of queries in Theorem 4.1 is essential for the "upwards" direction.

If $\leq^{\mathrm{P}}_{r}$ is a polynomial reducibility, then a class $\mathcal{C}$ is closed upwards under $\leq^{\mathrm{P}}_{r}$-reductions if and only if $\mathcal{C}^c$ is closed downwards under $\leq^{\mathrm{P}}_{r}$-reductions. We thus have the following immediate consequence of Theorem 4.2.

Corollary 4.3. Assume that $\vec{\beta}$ and $\vec{\beta}'$ are strongly positive P-sequences of biases, and let $\mathcal{C}$ be a class of languages that is closed upwards or downwards under $\leq^{\mathrm{P,lin}}_{\mathrm{pos-lin-tt}}$-reductions. Then

$$\mu^{\vec{\beta}}_{\mathrm{p}}(\mathcal{C}) = 1 \Longleftrightarrow \mu^{\vec{\beta}'}_{\mathrm{p}}(\mathcal{C}) = 1.$$

We now mention some consequences of Theorem 4.2, beginning with a discussion of the measure of the complete $\leq^{\mathrm{P}}_{\mathrm{T}}$-degree for exponential time, and its consequences for the BPP versus E problem.

For each class $\mathcal{D}$ of languages, we use the notations

$$\begin{aligned} \mathcal{H}_{\mathrm{T}}(\mathcal{D}) &= \{A \mid A \text{ is } \leq^{\mathrm{P}}_{\mathrm{T}}\text{-hard for } \mathcal{D}\}, \\ \mathcal{C}_{\mathrm{T}}(\mathcal{D}) &= \{A \mid A \text{ is } \leq^{\mathrm{P}}_{\mathrm{T}}\text{-complete for } \mathcal{D}\}, \end{aligned}$$

and similarly for other reducibilities. The following easy observation shows that every consequence of $\mu(\mathcal{C}_{\mathrm{T}}(\mathrm{E}_2)|\mathrm{E}_2) \neq 1$ is also a consequence of $\mu(\mathcal{C}_{\mathrm{T}}(\mathrm{E})|\mathrm{E}) \neq 1$.

Lemma 4.4. $\mu(\mathcal{C}_{\rm T}({\rm E})|{\rm E}) \neq 1 \Longrightarrow \mu(\mathcal{C}_{\rm T}({\rm E}_2)|{\rm E}_2) \neq 1.$

Proof. Juedes and Lutz [10] have shown that, if X is a set of languages that is closed downwards under $\leq_{\rm m}^{\rm P}$-reductions, then $\mu(X|{\rm E}_2) = 0 \Longrightarrow \mu(X|{\rm E}) = 0$. Applying this result with $X = \mathcal{H}_{\rm T}({\rm E})^c = \mathcal{H}_{\rm T}({\rm E}_2)^c$ yields the lemma. □

Allender and Strauss [1] have proven that $\mu_{\rm p}(\mathcal{H}_{\rm T}({\rm BPP})) = 1$. Buhrman, van Melkebeek, Regan, Sivakumar, and Strauss [8] have noted that this implies that $\mu(\mathcal{C}_{\rm T}({\rm E}_2)|{\rm E}_2) \neq 1 \Longrightarrow {\rm E} \not\subseteq {\rm BPP}$. Combining this argument with Corollary 4.3 yields the following extension.

Corollary 4.5. If there exists a strongly positive P-sequence of biases $\vec{\beta}$ such that $\mu^{\vec{\beta}}(\mathcal{C}_{\rm T}({\rm E}_2)|{\rm E}_2) \neq 1$, then ${\rm E} \not\subseteq {\rm BPP}$.

Regan, Sivakumar, and Cai [19] have proven a "most is all" lemma, stating that if $\mathcal{C}$ is any class of languages that is either closed under finite unions and intersections or closed under symmetric difference, then $\mu(\mathcal{C}|{\rm E}) = 1 \Longrightarrow {\rm E} \subseteq \mathcal{C}$. Combining this with Corollary 4.3 gives the following extended "most is all" result.

Corollary 4.6. Let $\mathcal{C}$ be a class of languages that is closed upwards or downwards under $\leq_{\rm pos-lin-tt}^{\rm P,lin}$-reductions, and is also closed under either finite unions and intersections or symmetric difference. If there is any strongly positive, P-sequence of biases $\vec{\beta}$ such that $\mu^{\vec{\beta}}(\mathcal{C}|{\rm E}) = 1$, then ${\rm E} \subseteq \mathcal{C}$.

Of course, the analagous result holds for ${\rm E}_2$.

We conclude with a brief discussion of small span theorems. Given a polynomial reducibility $\leq_r^{\rm P}$, the *lower* $\leq_r^{\rm P}$*-span* of a language A is

$$ {\rm P}_r(A) = \{B | B \leq_r^{\rm P} A\}, $$

and the *upper* $\leq_r^{\rm P}$*-span* of A is

$$ {\rm P}_r^{-1}(A) = \{B | A \leq_r^{\rm P} B\}. $$

We will use the following compact notation.

Definition. Let $\leq_r^{\mathrm{P}}$ be a polynomial reducibility type, and let ν be a probability measure on **C**. Then the *small span theorem for* $\leq_r^{\mathrm{P}}$*-reductions in the class* E *over the probability measure* ν is the assertion

$$\mathrm{SST}_\nu(\leq_r^{\mathrm{P}}, \mathrm{E})$$

stating that, for every $A \in \mathrm{E}$, $\nu(\mathrm{P}_r(A)|\mathrm{E}) = 0$ or $\nu_{\mathrm{p}}(\mathrm{P}_r^{-1}(A)) = \nu(\mathrm{P}_r^{-1}(A)|\mathrm{E}) = 0$. When the probability measure is μ, we omit it from the notation, writing $\mathrm{SST}(\leq_r^{\mathrm{P}}, \mathrm{E})$ for $\mathrm{SST}_\mu(\leq_r^{\mathrm{P}}, \mathrm{E})$. Similar assertions for other classes, e.g., $\mathrm{SST}_\nu(\leq_r^{\mathrm{P}}, \mathrm{E}_2)$, are defined in the now-obvious manner.

Juedes and Lutz [9] proved the first small span theorems, $\mathrm{SST}(\leq_{\mathrm{m}}^{\mathrm{P}}, \mathrm{E})$ and $\mathrm{SST}(\leq_{\mathrm{m}}^{\mathrm{P}}, \mathrm{E}_2)$, and noted that extending either to $\leq_{\mathrm{T}}^{\mathrm{P}}$ would establish $\mathrm{E} \not\subseteq \mathrm{BPP}$. Lindner [14] established $\mathrm{SST}(\leq_{1-\mathrm{tt}}^{\mathrm{P}}, \mathrm{E})$ and $\mathrm{SST}(\leq_{1-\mathrm{tt}}^{\mathrm{P}}, \mathrm{E}_2)$, and Ambos-Spies, Neis, and Terwijn [4] proved $\mathrm{SST}(\leq_{k-\mathrm{tt}}^{\mathrm{P}}, \mathrm{E})$ and $\mathrm{SST}(\leq_{k-\mathrm{tt}}^{\mathrm{P}}, \mathrm{E}_2)$ for all fixed $k \in \mathbb{N}$. Very recently, Buhrman and van Melkebeek [7] have taken a major step forward by proving $\mathrm{SST}(\leq_{g(n)-\mathrm{tt}}^{\mathrm{P}}, \mathrm{E}_2)$ for every function $g(n)$ satisfying $g(n) = n^{o(1)}$. We note that the Bias Invariance Theorem implies that small span theorems lying "just beyond" this latter result are somewhat robust with respect to changes of biases.

Theorem 4.7. If $\leq_r^{\mathrm{P}}$ is a polynomial reducibility such that $A \leq_{\mathrm{pos-lin-tt}}^{\mathrm{P,lin}} B$ implies $A \leq_r^{\mathrm{P}} B$, then for every strongly positive P-sequence of biases $\vec{\beta}$,

$$\mathrm{SST}_{\mu^{\vec{\beta}}}(\leq_r^{\mathrm{P}}, \mathrm{E}) \Longleftrightarrow \mathrm{SST}(\leq_r^{\mathrm{P}}, \mathrm{E}),$$

and similarly for E_2.

Acknowledgment. The first author thanks Steve Kautz for a very useful discussion.

References

[1] E. Allender and M. Strauss. Measure on small complexity classes with applications for BPP. In *Proceedings of the 35th Symposium on Foundations of Computer Science*, pages 807–818, Piscataway, NJ, 1994. IEEE Computer Society Press.

[2] N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley, 1992.

[3] K. Ambos-Spies and E. Mayordomo. Resource-bounded measure and randomness. In A. Sorbi, editor, *Complexity, Logic and Recursion Theory*, Lecture Notes in Pure and Applied Mathematics, pages 1–47. Marcel Dekker, New York, N.Y., 1997.

[4] K. Ambos-Spies, H.-C. Neis, and S. A. Terwijn. Genericity and measure for exponential time. *Theoretical Computer Science*, 168:3–19, 1996.

[5] J. M. Breutzmann and J. H. Lutz. Equivalence of measures of complexity classes. *SIAM Journal on Computing*. To appear. See also *Proceedings of the 14th Symposium on Theoretical Aspects of Computer Science*, Springer-Verlag, 1997, pp. 535–545.

[6] H. Buhrman and L. Torenvliet. Complete sets and structure in subrecursive classes. In *Proceedings of Logic Colloquium '96*, pages 45–78. Springer-Verlag, 1998.

[7] H. Buhrman and D. van Melkebeek. Hard sets are hard to find. In *Proceedings of the 13th IEEE Conference on Computational Complexity*, pages 170–181, New York, 1998. IEEE.

[8] H. Buhrman, D. van Melkebeek, K. Regan, D. Sivakumar, and M. Strauss. A generalization of resource-bounded measure, with an application. In *Proceedings of the 15th Annual Symposium on Theoretical Aspects of Computer Science*, pages 161–171, Berlin, 1998. Springer-Verlag.

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

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

[11] D. W. Juedes and J. H. Lutz. Completeness and weak completeness under polynomial-size circuits. *Information and Computation*, 125:13–31, 1996.

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

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

[14] W. Lindner. On the polynomial time bounded measure of one-truth-table degrees and p-selectivity, 1993. Diplomarbeit, Technische Universität Berlin.

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

[16] J. H. Lutz. The quantitative structure of exponential time. In L.A. Hemaspaandra and A.L. Selman, editors, *Complexity Theory Retrospective II*, pages 225–254. Springer-Verlag, 1997.

[17] J. H. Lutz. Resource-bounded measure. In *Proceedings of the 13th IEEE Conference on Computational Complexity*, pages 236–248, New York, 1998. IEEE.

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

[19] 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.

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

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

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

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

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

OPTIONAL TECHNICAL APPENDICES

Appendix A. Resource-Bounded ν-Measure

In this appendix, we present the basic elements of resource-bounded measure based on an arbitrary probability measure ν on $\mathbf{C}$. The material in Appendices A and B is taken, with permission, from [5].

Definition. 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 $\vec{\beta} = (\beta_0, \beta_1, \beta_2, \dots)$, 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 $w \in \{0,1\}^*$.

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 $s_0, s_1, s_2, \dots$ 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 $s_i \notin A$.

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

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

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

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.

<u>Definition.</u> 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 $b \in \{0,1\}$, $\nu(wb|w) \geq \delta$.

<u>Definition.</u> 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 $i \in \mathbb{N}$, $\beta_i \in [\delta, 1 - \delta]$.

We next review the well-known notion of a martingale over a probability measure ν. Computable martingales were used by Schnorr [20, 21, 22, 23] in his investigations of randomness, and have more recently been used by Lutz [15] in the development of resource-bounded measure.

<u>Definition.</u> 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 $w \in \{0,1\}^*$,

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

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.)

Definition. 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

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

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

$$S^\infty_{\rm str}[d] = \left\{ A \in \mathbf{C} \mid \liminf_{n \to \infty} d(A[0..n-1]) = \infty \right\}.$$

Definition. 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. 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 $\nu_{\rm p}(X) = 0$, 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 $\nu_{\rm p}(X) = 1$, if $\nu_{\rm p}(X^c) = 0$, 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. 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 $\nu_{\rm p}(X \cap E) = 0$.

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 $\nu(X^c|\mathrm{E}_2) = 0$.

Definition. Let ν be a positive probability measure on **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. Two positive probability measures ν and ν' on **C** are *summably equivalent*, and we write $\nu \approx \nu'$, if for every $A \subseteq \{0,1\}^*$,

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

Definition.

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. 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 A.1 (Breutzmann and Lutz [5]). For every P-sequence of biases $\vec{\beta}$, there is a P-exact sequence of biases $\vec{\beta}'$ such that $\vec{\beta} \approx \vec{\beta}'$.

Appendix B. Truth-Table Reductions

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 $(f_1(x), \ldots, f_{n(x)}(x))$ 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 $Q_{(f,g)}(x) = \{f_1(x), \ldots, f_{n(x)}(x)\}$ 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 $\leq_{\rm tt}$-reduction, and let $z \in \{0,1\}^*$. Then the *inverse image* of the cylinder $\mathbf{C}_z$ under the reduction (f,g) is

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

The following well-known fact is easily verified.

Lemma B.1. If ν is a probability measure on $\mathbf{C}$ and (f,g) is a $\leq_{\rm tt}$-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 B.1 is called the *probability measure induced by* ν *and* (f,g).

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

Definition. A $\leq_{\rm tt}$-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.

Appendix C. Proof of Martingale Contraction Theorem

Let $w \in \{0,1\}^*$, and let $y = F^{-1}_{(f,g),d}(w)$ Note that for any $v \succ y$, $|v| = \mathit{step}(|y|)$ and either $F(v) = w0$ or $F(v) = w1$. Let $l = \mathit{step}(|y|) - |y|$. We have

$$\begin{aligned}
(f,g)_{\smile} d(w) &= d(y) \\
&= \sum_{v \succ y} d(v)\nu(v|y) \\
&= \sum_{\substack{v \succ y \\ F(v)=w0}} d(v)\nu(v|y) \;+\; \sum_{\substack{v \succ y \\ F(v)=w1}} d(v)\nu(v|y) \\
&\geq \sum_{\substack{v \succ y \\ F(v)=w0}} [\min_v d(v)]\nu(v|y) \;+\; \sum_{\substack{v \succ y \\ F(v)=w1}} [min_v d(v)]\nu(v|y) \\
&= [(f,g)_{\smile} d(w0)] \sum_{\substack{v \succ y \\ F(v)=w0}} \nu(v|y) \;+\; [(f,g)_{\smile} d(w0)] \sum_{\substack{v \succ y \\ F(v)=w1}} \nu(v|y) \\
&= [(f,g)_{\smile} d(w0)] \sum_{\substack{x \in \{0,1\}^l \\ F(yx)=w0}} \nu(yx|y) \;+\; [(f,g)_{\smile} d(w0)] \sum_{\substack{x \in \{0,1\}^l \\ F(yx)=w1}} \nu(yx|y) \\
&= (f,g)_{\smile} d(w0)\nu^{(f,g)}(w0|w) \;+\; (f,g)_{\smile} d(w1)\nu^{(f,g)}(w1|w).
\end{aligned}$$

The penultimate step follows from the fact that (f,g) is an *orderly* $\leq_{\mathrm{tt}}$-reduction, and the last step is Lemma 6.4 of [5]. This shows that $(f,g)_{\smile} d$ is a $\nu^(f,g)$-supermartingale.

To see that $(f,g)_{\smile} d$ satisfies the desired success condition, let A, be a language such that $F^{-1}_{(f,g)}(\{A\}) \subseteq S^{\infty}_{\mathrm{str}}[d]$. If $A \notin$ range $F_{(f,g)}$, then $F^{-1}_{(f,g),d}(w)$ is undefined for all sufficiently long prefixes w of A, whence it is clear that $A \in S^{\infty}[(f,g)_{\smile} d]$. If $A \in$ range $F_{(f,g)}$, then $F^{-1}_{(f,g),d}(A[0..n-1])$ is defined for all n and $F^{-1}_{(f,g),d}(A) \in F^{-1}_{(f,g)}(\{A\})$, so

$$\begin{aligned}
\limsup_{n\to\infty} (f,g)_{\smile} d(A[0..n-1]) &= \limsup_{n\to\infty} F^{-1}_{(f,g),d}(A[0..n-1]) \\
&\geq \liminf_{n\to\infty} F^{-1}_{(f,g),d}(A[0..n-1]) \\
&\geq \liminf_{n\to\infty} F^{-1}_{(f,g),d}(A)[0..n-1] \\
&= \infty,
\end{aligned}$$

whence we again have $A \in S^{\infty}[(f,g)_{\smile} d]$.

Appendix D. Proof of Positive Bias Reduction Theorem

Let ν be the coin toss distribution specified by biases $\beta_0, \beta_1, \ldots \in [\delta, 1-\delta]$, and let $\beta' \in [\delta, 1-\delta]$ and $\epsilon > 0$ be given. We want to construct a formula of the form

$$C_{\beta'} = \left(\bigwedge_{j_1=1}^{a_1} x_{1j_1}\right) \wedge \left(\left(\bigvee_{k_1=1}^{b_1} y_{k_1}\right) \vee \left\{\left(\bigwedge_{j_2=1}^{a_2} x_{2j_2}\right) \wedge \left[\left(\bigvee_{k_2=1}^{b_1} y_{k_2}\right) \vee \cdots\right]\right\}\right). \tag{2}$$

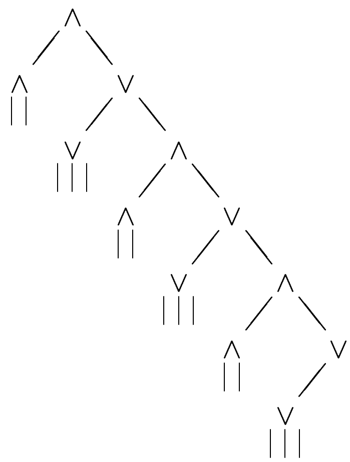
We suppose that the inputs to this circuit are random and independent, and that $\Pr(z = 1) = \beta_i, i = 0, 1, 2, \ldots,$ if z, ranging over all x's and y's, appears i^{th} in the formula above. Under this hypothesis, we want $|\Pr(C_{\beta'} = 1) - \beta'| < \epsilon$ and that the number of inputs to $C_{\beta'}$ be at most $O(\lg(1/\epsilon))$.

For example, if $a_1 = a_2 = \cdots = 2$ and $b_1 = b_2 = \cdots = 3$, we have:

```
(and
   (and
      z0 z1)
   (or
      (or
         z2 z3 z4)
      (and
         (and
            z5 z6)
         (or
            (or
               z7 z8 z9)
            (and
               (and
                  z10 z11)
               (or
                  (or
                     z12 z13 z14)))))))
```

and $\Pr(z_i = 1) = \beta_i$.

In pictures, we'd have



For real numbers $x, y \in [0,1]$, let $x \oplus y$ denote $1-(1-x)(1-y)$. Thus, for independent A and B, $\Pr(A) \oplus \Pr(B) = \Pr(A \vee B)$. Note that $\oplus$ is monotonically increasing in its arguments, that $\bigoplus_{k=1}^{n} x_k$ is monotonically increasing in n, and that the empty $\oplus$, $\bigoplus_{k=1}^{0} x_k$, is 0.

We need to determine the a's and b's in Formula (2). The algorithm, on input β', $\beta_0, \beta_1, \beta_2, \ldots \in [\delta, 1-\delta]$, and tolerance ϵ, is as follows:

- If $\epsilon > 1$ return the constant false circuit. Also do the right thing if β' is 0 or 1. Otherwise continue...
- Determine a so that
$$\prod_{j=1}^{a+1} \beta_j < \beta' \leq \prod_{j=1}^{a} \beta_j.$$
Put $A = \prod_{j=1}^{a} \beta_j$.
- Determine b so that
$$A \cdot \bigoplus_{k=a+1}^{a+b} \beta_k < \beta' \leq A \cdot \bigoplus_{k=a+1}^{a+b+1} \beta_k.$$
Put $B = \bigoplus_{k=a+1}^{a+b} \beta_k$.
- Determine β'' so that $\beta' = A(B \oplus \beta'')$, i.e., $\beta'' = \frac{\beta' - AB}{A(1-B)}$. Inductively find a formula $C_{\beta''}$ of the top-level shape whose probability of acceptance is β''. Use tolerance $\epsilon/(A(1-$

$B)$).

- Put

$$C_{\beta'} = \left(\bigwedge_{j_1=1}^{a} x_{1j_1} \right) \wedge \left(\left(\bigvee_{k_1=1}^{b} y_{k_1} \right) \vee C_{\beta''} \right).$$

Now we analyze the algorithm. First, the formula generated has at most $O(\lg(1/\epsilon_0))$ inputs, where ϵ_0 is the initial value of ϵ. Note that each recursive call increases the tolerance ϵ by at least the factor $1/(A(1-B)) = 1/(1-\delta)^{a+b}$; it follows that ϵ will grow to be at least 1 for $\sum(a_i + b_i) \le \frac{\lg \epsilon_0}{\lg(1-\delta)}$.

Next, the algorithm is correct, i.e., produces a circuit with probability of acceptance in the range $\beta' \pm \epsilon$. Clearly this is the case if the algorithm returns immediately (when $\epsilon > 1$). Otherwise, suppose inductively that $C_{\beta''}$ has probability $\beta'' \pm \epsilon/(A(1-B))$. It follows that C_β has acceptance probability in

$$\begin{aligned} A\left(B \oplus \left(\beta'' \pm \frac{\epsilon}{A(1-B)}\right)\right) &= A\left[1 - (1-B)\left(1 - \beta'' \pm \frac{\epsilon}{A(1-B)}\right)\right] \\ &= A\left[1 - (1-B)\left(1-\beta''\right)\right] \pm A(1-B)\frac{\epsilon}{A(1-B)} \\ &= A(B \oplus \beta'') \pm \epsilon. \end{aligned}$$

□