# Resource-bounded randomness and compressibility with respect to nonuniform measures

Steven M. Kautz[*]

September 3, 2004

**Abstract**

Most research on resource-bounded measure and randomness has focused on the uniform probability density, or Lebesgue measure, on $\{0,1\}^\infty$; the study of resource-bounded measure theory with respect to a *non*uniform underlying measure was recently initiated by Breutzmann and Lutz [1]. In this paper we prove a series of fundamental results on the role of nonuniform measures in resource-bounded measure theory. These results provide new tools for analyzing and constructing martingales and, in particular, offer new insight into the compressibility characterization of randomness given recently by Buhrman and Longpré [2].

We give several new characterizations of resource-bounded randomness with respect to an underlying measure $\mu$: the first identifies those martingales whose rate of success is asymptotically *optimal* on the given sequence; the second identifies martingales which induce a *maximal compression* of the sequence; the third is a (nontrivial) extension of the compressibility characterization to the nonuniform case. In addition we prove several technical results of independent interest, including an extension to resource-bounded measure of the classical theorem of Kakutani on the equivalence of product measures; this answers an open question in [1].

## 1 Introduction

Resource-bounded measure theory, as developed by Lutz in the late 1980's, provides a means for quantitatively analyzing the structure of complexity classes as well as for characterizing random or pseudorandom elements within them. Results in this area published over the last seven years include a growing body of new insights into familiar problems in computational complexity; see [8] for a recently updated survey.

Most work in the areas of resource-bounded measure and randomness has focused on the uniform distribution on $\{0,1\}^\infty$, or Lebesgue measure, in which the probabilities associated with the bits of an infinite sequence are *independent* and *uniform*; i.e., the probability that a given bit is "1" is always one-half. The uniform measure unquestionably provides the most natural starting point for developing a theory of measure and for investigating randomized complexity. Moreover, for example, Breutzmann and Lutz have recently shown that the resource-bounded measure of most complexity classes of interest is to a certain extent invariant with respect to the underlying distribution. Nonetheless there are several reasons for considering nonuniform measures.

One reason, of course, is that for modeling and analyzing computation relative to a physical source of randomness both the assumptions of independence and uniformity are likely to be unrealistic. A number of authors have studied feasible computation relative to random sequences

[*] Department of Mathematics, Randolph-Macon Woman's College, 2500 Rivermont Avenue, Lynchburg, VA 24503 (skautz@rmwc.edu).

with nonuniform distributions, such as quasi-random and slightly-random sources for BPP [13], [18], and more recently "extractors" and "dispersers" (see the survey [12]).

However, what particularly concerns us here is that results on nonuniform measures are essential to an understanding of the subject of randomness as a whole, and consequently yield techniques which can be fruitfully used to obtain results even in the uniform case. Examples of such methods can easily be found in the somewhat more mature area of constructive measure (i.e., algorithmic randomness in the sense of Martin-Löf); to cite just a few examples, techniques involving nonuniform distributions form the basis of van Lambalgen's new proof of Ville's theorem [17] and of Shen's proof that there are Kolmogorov-Loveland stochastic sequences which are not algorithmically random [15]. Indeed, any truth-table reduction induces a computable measure which is, in general, nonuniform; see [6] for details and applications. Although the study of *resource-bounded* measure in the nonuniform case was only initiated recently in [1], related techniques are already in use by Lutz and Mayordomo in [9].

In this paper we present a number of results on nonuniform distributions in the context of resource-bounded measure and the corresponding notions of resource-bounded randomness, beginning with the observation that *any* nontrivial martingale is implicitly constructed from a nonuniform measure representing its "betting strategy." (Martingales and other fundamental notions in resource-bounded measure are discussed in Section 2.) Thus, even if one is ultimately interested only in the uniform case, at the very least the analysis of and construction of martingales can be facilitated by an understanding of the role of this implicit, nonuniform measure. With this understanding in hand, for example, we show that a martingale succeeds *optimally* on a given sequence (i.e., dominates any other martingale on initial segments of the given sequence) just when the sequence is random with respect to the martingale's betting strategy (see Section 3.1).

One of our main contributions is to offer new insight into the "compressibility" characterization of resource-bounded randomness. During the 1960's Martin-Löf proposed a definition of infinite random sequences based on a form of constructive measure; at the same time, Kolmogorov, Levin, and others proposed a definition based on incompressibility of initial segments, and it was later shown that the two approaches yield exactly the same class of random sequences, the so-called algorithmically random or 1-random sequences. (See [7] for historical details and references.) Recently Buhrman and Longpré [2] have shown that resource-bounded randomness can also be characterized in terms of a kind of compressibility, providing a beautiful unification of the two approaches to randomness in the resource-bounded setting. Very briefly, they show that if a martingale succeeds on a sequence $A$, it is possible to define a "compressed" sequence $B$ along with algorithms for transforming $A$ to $B$ and vice versa (the "compression" and "decompression", respectively) which are induced in a uniform way by the martingale. In this paper we give a natural description of the mapping between $A$ and $B$ as a kind of measure transformation and prove that the compression of $A$ is *maximal*—that is, the compressed sequence $B$ is itself incompressible—if and only if the martingale which induces the compression is *optimal* for $A$, i.e., $A$ is random with respect to the measure corresponding to the martingale's betting strategy. Intuitively, this suggests that a bit stream can be maximally compressed if and only if it is random with respect to *some* computable measure.

We subsequently give a definition of compressibility with respect to nonuniform measures, which turns out to be a nontrivial extension of the original definition given in [2], and prove that the characterization of resource-bounded randomness still holds in the nonuniform case.

It is convenient to summarize our main results in the form of the following new characterizations of resource-bounded randomness. Here $A \in \{0,1\}^\infty$, $\Delta$ is a class of functions (such as $p$ or rec), $\mu$ and $\nu$ represent measures which are $\Delta$-computable, and $d$ is a martingale with underlying measure $\nu$ and betting strategy determined by $\mu$ (see Section 2 for definitions and see Theorem 6.3 for a precise statement of hypotheses). The following are equivalent :

(i) $A$ is $\Delta$-random with respect to $\mu$.

(ii) The martingale $d$ is optimal for $A$.

(iii) The martingale $d$ induces a maximal compression of $A$.

(iv) $A$ is $\Delta$-incompressible with respect to $\mu$.

In addition, we prove several results incidental to the above which are themselves of independent interest. In Section 3 we discuss the equivalence of measures and provide an extension to resource-bounded measure of the classical theorem of Kakutani on the equivalence of product measures, answering an open question in [1]. In Section 4 we show that $\Delta$-randomness is preserved by certain kinds of truth-table reductions, namely those representing the kinds of measure transformations implicit in the compressibility characterization of randomness (Section 5). While it is known that algorithmic randomness is preserved by arbitrary truth-table reductions, the resource-bounded case is much more difficult; indeed, one of the main results in [1] is that a certain restricted class of truth-table reductions preserves $\Delta$-randomness. Our result extends to $\Delta$-randomness an *invariance* property of algorithmic randomness discussed in [19] and in [6]; loosely, the binary expansion of a given real number $x$ is $\Delta$-random if and only if *every* representation of $x$ as a binary sequence is $\Delta$-random with respect to an underlying measure appropriate for the representation.

Except where other references are given, proofs of all nontrivial results are in the appendices.

# 2 Preliminaries

## 2.1 Notation

Let $\mathbb{N}$ denote the natural numbers and let $\mathbb{D}$ denote the dyadic rationals (real numbers whose binary expansion is a finite string). A *string* is an element of $\{0,1\}^*$. The concatenation of strings $x$ and $y$ is denoted $xy$. For any string $x$, $|x|$ denotes the length of $x$, and $\lambda$ is the unique string of length 0. If $x \in \{0,1\}^*$ and $j,k \in \mathbb{N}$ with $0 \le j \le k < |x|$, $x[k]$ is the $k$th bit (symbol) of $x$ and $x[j..k]$ is the string consisting of the $j$th through $k$th bits of $x$. Note that the "first" bit of $x$ is $x[0]$; it is convenient to let $x[0..-1]$ denote the empty string. For $A \in \{0,1\}^\infty$, the notations $A[k]$ and $A[j..k]$ are defined analogously. For any $x, y \in \{0,1\}^*$, $x \sqsubseteq y$ means that if $x[k]$ is defined, then $y[k]$ is also defined and $x[k] = y[k]$; we say that $x$ is an *initial segment*, or *prefix*, of $y$ or that $y$ is an *extension* of $x$. Likewise for $A \in \{0,1\}^\infty$, $x \sqsubseteq A$ means $x[k] = A[k]$ whenever bit $x[k]$ is defined. Strings $x$ and $y$ are said to be *incompatible*, or *disjoint*, if there is no string $z$ which is an extension of both $x$ and $y$, i.e., $x \not\sqsubseteq y$ and $y \not\sqsubseteq x$.

Fix a standard enumeration of $\{0,1\}^*$, $s_0 = \lambda, s_1 = 0, s_2 = 1, s_3 = 00, s_4 = 01, \ldots$. A *language* is a subset of $\{0,1\}^*$; a language $A$ will be identified with its characteristic sequence $\chi_A \in \{0,1\}^\infty$, defined by $s_y \in A \iff \chi_A[y] = 1$ for $y \in \mathbb{N}$. We will consistently write $A$ for $\chi_A$. $X^c$ denotes the complement of $X$ in $\{0,1\}^\infty$. Strings used to represent partially defined languages (initial segments) will typically be represented by lower-case greek letters.

Throughout this paper the symbol $\Delta$ represents a class of functions, i.e., the resource bounds. We are generally interested in either rec, the class of all recursive functions, or $p$, the class of functions $f : \{0,1\}^* \to \{0,1\}^*$ such that $f(x)$ is computable in time polynomial in $|x|$, though the results are easily extended to many other resource bounds. A real-valued function $f$ on

$\{0,1\}^*$ is said to be $\Delta$-*computable* if there is a function $\hat{f} : \mathbb{N} \times \{0,1\}^* \to \mathbb{D}$ in $\Delta$ such that for all $n \in \mathbb{N}$ and $x \in \{0,1\}^*$, $|\hat{f}(n,x) - f(x)| \leq 2^{-n}$. A function $f$ is *exactly* $\Delta$-*computable* if $f$ itself is in $\Delta$ and is dyadic-valued.

A string $\sigma \in \{0,1\}^*$ defines the subset $\mathrm{Ext}(\sigma) = \{A \in \{0,1\}^\infty : \sigma \sqsubseteq A\}$ of $\{0,1\}^\infty$, called a *cylinder* or *interval*. Likewise if $S$ is a subset of $\{0,1\}^*$, $\mathrm{Ext}(S)$ denotes $\bigcup_{\sigma \in S} \mathrm{Ext}(\sigma)$.

## 2.2 Measure

**Definition 2.1** *A* measure *is a function $\mu$ on $\{0,1\}^*$, taking values in $[0,1]$, such that $\mu(\lambda) = 1$ and for every $\sigma \in \{0,1\}^*$, $\mu(\sigma) = \mu(\sigma 0) + \mu(\sigma 1)$.*

The function $\mu$ specifies a probability density on $\{0,1\}^\infty$, where $\mu(\sigma)$ represents the probability associated with the interval $\mathrm{Ext}(\sigma)$. Standard results of measure theory (see [3]) show that such a function can always be extended uniquely to subsets $\mathcal{E} \subseteq \{0,1\}^\infty$ which are built up from intervals by some finite iteration of countable union and complementation operations (the *Borel* sets); we continue to write $\mu(\sigma)$ for $\mu(\mathrm{Ext}(\sigma))$ and $\mu(S)$ for $\mu(\mathrm{Ext}(S))$, where $S \subseteq \{0,1\}^*$. We may also regard a measure as a function on the unit interval $[0,1]$ via the usual correspondence between binary sequences and real numbers.

Let $\sigma \in \{0,1\}^*$, $|\sigma| = n$, and $b \in \{0,1\}$. If $\mu(\sigma) \neq 0$, the *bit probability*

$$\mu(\sigma b | \sigma) = \frac{\mu(\sigma b)}{\mu(\sigma)}$$

is defined, representing the conditional probability that bit $n$ is equal to $b$, given the initial segment $\sigma$. By the multiplication rule for conditional probabilities, we always have

$$\mu(\sigma) = \prod_{i=0}^{n-1} \mu(\sigma[0..i] | \sigma[0..i-1])$$

when $\mu(\sigma) \neq 0$. Note that $\mu(\sigma 0|\sigma) + \mu(\sigma 1|\sigma) = 1$, and any function assigning a number $p_\sigma \in (0,1)$ to each string $\sigma$ uniquely determines via the multiplication rule a measure $\mu$ whose bit probabilities are $\mu(\sigma 0|\sigma) = 1 - p_\sigma$ and $\mu(\sigma 1|\sigma) = p_\sigma$.

Most of our results will apply to measures for which all the bit probabilities are either positive or are bounded away from zero. The terminology "strongly positive" was suggested in [1].

**Definition 2.2** *A measure $\mu$ is* positive *if $\mu(\sigma) > 0$ for every string $\sigma$. A measure $\mu$ is* strongly positive *if there is a constant $\delta > 0$ such that for every string $\sigma$ and $b \in \{0,1\}$, $\delta \leq \mu(\sigma b|\sigma) \leq 1 - \delta$.*

If the bit probabilities are *independent*, that is, $\mu(\sigma b|\sigma)$ depends only on the position $n = |\sigma|$ but not on $\sigma$ itself, the measure $\mu$ is then referred to as a *product measure*; the measure is determined by a sequence of pairs $\{(1-p_n, p_n)\}_{n=0}^\infty$, i.e., $\mu(\sigma 1|\sigma) = p_n$ for all $\sigma$ of length $n$. The sequence $\{(1-p_n, p_n)\}$ may be viewed as sequence of coins, fixed in advance, such that the $n$th coin has probability $p_n$ of coming up heads. If all the coins are fair ($p_n = \frac{1}{2}$), then $\mu$ is the usual Lebesgue measure on the Borel subsets of $\{0,1\}^\infty$. We use the symbol $\lambda$ for Lebesgue measure (distinguished from the empty string, hopefully, in context).

If $\mu$ is not a product measure, $\mu$ is said to contain *dependencies*. We may still use the intuitive picture of a sequences of coin tosses, but with the following difference. Instead of the sequence of coins being fixed in advance of the experiment, there is a daemon which examines the outcomes $\sigma$ of the first $n$ tosses and then selects a coin whose probability of coming up heads is $\mu(\sigma 1|\sigma)$. Measures with dependencies are probably the appropriate models for imperfect, physical sources of randomness where it generally cannot be assumed that the bit probabilities are independent. Examples of such measures include the "adversary sources" of [13], where in analyzing the behavior of probabilistic algorithms using such a source the authors assume the source to be adversarial, that is, the daemon determining the function $\mu$ is aware of the probabilistic algorithm being used and attempts to fix the bias of successive coins, within some bounds $[\delta, 1-\delta]$, in such a way as to make the algorithm fail.

### 2.3 Martingales

**Definition 2.3** *Let $\nu$ be a positive measure. A $\nu$-*martingale *is a function $d$ on $\{0,1\}^*$ taking values in $[0,\infty)$ such that*

$$d(\sigma) = \nu(\sigma 0|\sigma)d(\sigma 0) + \nu(\sigma 1|\sigma)d(\sigma 1). \tag{1}$$

*A martingale $d$* succeeds *on a sequence $A \in \{0,1\}^\infty$ if $\limsup_{n\to\infty} d(A[0..n]) = \infty$.*

A martingale may be regarded as a strategy for betting on successive bits of an infinite sequence. The value $d(\sigma)$ represents the gambler's accumulated capital after sequence of outcomes $\sigma$. The equality (1) asserts that the game is *fair*, that is, at each node $\sigma$ the (conditional) expected value of the capital after the next bit always equal to its present value.

It follows from the standard result below, known as *Kolmogorov's inequality for martingales*, that a set $X \subseteq \{0,1\}^\infty$ has *measure zero* with respect to measure $\nu$ if and only if there is a $\nu$-martingale succeeding on every $A \in X$: that is, a martingale may be viewed as an *orderly* demonstration that a set $X$ has measure zero. It is this fact which makes martingales useful for defining measure with resource bounds.

**Lemma 2.4** *Let $d$ be a $\nu$-martingale and $t$ a positive real number; then*

$$\nu\{\sigma \in \{0,1\}^* : d(\sigma) > t\} < \frac{d(\lambda)}{t}.$$

### 2.4 Resource-bounded measure

Resource-bounded measure theory, as developed by Lutz, is a form of effective or constructive measure theory which provides a means of defining the measure or probability of sets of languages within many standard complexity classes, allowing for the first time a *quantitative* analysis of such well-studied questions as, for example, whether P is a proper subclass of NP. In addition, a definition of measure within a complexity class provides a means of defining the "random" languages for the class. See [10], [8], or [11] for a more thorough introduction.

**Definition 2.5** *Let $\Delta$ be a class of functions, and let $\nu$ be a $\Delta$-computable measure. A class $X \subseteq \{0,1\}^\infty$ has* $\Delta$-measure zero *with respect to $\nu$, written $\nu_\Delta(X) = 0$, if there is a $\Delta$-computable $\nu$-martingale which succeeds on every $A$ in $X$. $X$ has* $\Delta$-measure one w.r.t. *$\nu$ if $\nu_\Delta(X^c) = 0$. A sequence $A \in \{0,1\}^\infty$ is* $\Delta$-random w.r.t *$\nu$ if there is no $\Delta$-computable $\nu$-martingale which succeeds on $A$*

Lutz has proved a number of results showing that $\Delta$-measure behaves like a measure within an appropriate corresponding complexity class; for example, $p$-measure is appropriate as a measure in the class E = DTIME($2^{\text{linear}}$).

## 3 The decomposition of a martingale using likelihood ratios

A martingale $d$ may be fruitfully viewed as composed of two parts, which correspond respectively to the *strategy* used by the gambler and to *odds* paid on her wins. At each node $\sigma$, the gambler selects some proportion $p_\sigma$ of her capital $d(\sigma)$ to bet on $\sigma 1$ and the remaining proportion $1 - p_\sigma$ to bet on $\sigma 0$. The *strategy* is the unique measure $\mu$ determined by the bit probabilities $\mu(\sigma 1|\sigma) = p_\sigma$.

It is the underlying measure $\nu$ which determines the odds. Recall that in a fair game, if a bet of amount $B$ is placed on an event $\mathcal{E}$, and $\mathcal{E}$ occurs, the gambler receives her original $B$ plus $\frac{N}{D}B$ more, where the ratio $N/D$ is equal to the probability of **not** $\mathcal{E}$ divided by the probability of $\mathcal{E}$, the so-called "odds against $\mathcal{E}$." Let $q_\sigma$ denote the bit probability $\nu(\sigma 1|\sigma)$ and let $p_\sigma = \mu(\sigma 1|\sigma)$, the proportion of $d(\sigma)$ the gambler bets on $\sigma 1$. The odds paid on the occurrence of a "1" are $(1 - q_\sigma)/q_\sigma$, so we can compute the capital at $\sigma 1$ as

$$\begin{aligned} d(\sigma 1) &= p_\sigma d(\sigma)\left(1 + \frac{1 - q_\sigma}{q_\sigma}\right) = \frac{p_\sigma}{q_\sigma} \cdot d(\sigma), \\ \text{and similarly, } d(\sigma 0) &= \frac{1 - p_\sigma}{1 - q_\sigma} \cdot d(\sigma). \end{aligned}$$

Using the multiplication rule inductively we have, for all $\sigma$,

$$d(\sigma) = \frac{\mu(\sigma)}{\nu(\sigma)} \cdot d(\lambda). \tag{2}$$

We may refer to $d$ as the *$\nu$-martingale with strategy $\mu$*. The coefficient $\mu(\sigma)/\nu(\sigma)$ is known as a *likelihood ratio* (see [16]). In the next two sections we investigate how the relationship between $\nu$ and $\mu$ affects the success of the martingale.

Most of our results will apply only to martingales for which the underlying measure $\nu$ is strongly positive, and in many cases an assumption is required that the measure $\mu$ determining the strategy be strongly positive as well. The next lemma shows that very little generality is lost by the latter assumption. In addition we will need the fact shown below that "limsup" can be replaced by "lim" in Definition 2.3.

**Lemma 3.1** *Let $d$ be a $\Delta$-computable $\nu$-martingale, where $\nu$ is a strongly positive, $\Delta$-computable measure. There exists a strongly positive, $\Delta$-computable measure $\mu$ such that, if $\tilde{d}$ denotes the $\nu$-martingale with strategy $\mu$, then for every $A$ on which $d$ succeeds,* $\lim_{n\to\infty} \tilde{d}(A[0..n]) = \infty$.

## 3.1 Optimal martingales

Here we present a simple application of the decomposition via likelihood ratios given above. This notion of "optimality" will also be used in the proof of Theorem 4.3. Note first that it is clear from the definitions that if a sequence $A$ is $\Delta$-random w.r.t. a measure $\nu$, then for any measure $\mu$ the ratio $\mu(A[0..n])/\nu(A[0..n])$ is bounded, since no $\nu$-martingale can succeed on $A$. On the other hand, if $A$ is random with respect to $\mu$ but not $\nu$, the martingale $d(\sigma) = \mu(\sigma)/\nu(\sigma)$ not only succeeds on $A$, but does so at an asymptotically optimal rate. In fact, the converse holds as well, providing the first of the characterizations of $\Delta$-randomness promised in the introduction.

**Definition 3.2** *Fix a measure $\nu$; a $\nu$-martingale $d$ is* optimal *for $A \in \{0,1\}^\infty$ if for every other $\nu$-martingale $\tilde{d}$, there is a constant $C$ such that for all $n$, $\tilde{d}(A[0..n]) < C \cdot d(A[0..n])$.*

**Theorem 3.3** *Let $\mu$ and $\nu$ be strongly positive, $\Delta$-computable measures, let $d$ be the $\nu$-martingale with strategy $\mu$, and let $A \in \{0,1\}^\infty$. Then $A$ is $\Delta$-random with respect to $\mu$ if and only if $d$ is optimal for $A$.*

## 3.2 Equivalence of measures

It is clear from the form of (2) that if the strategy $\mu$ of a martingale is nearly the same as the underlying measure $\nu$, the martingale will not succeed on any sequence $A$. In this section we characterize just how different $\mu$ and $\nu$ must be in order for the martingale to succeed. It turns out that there are two useful forms of equivalence we might consider, which we refer to here as a "strong" form and a "weak" form. What we refer to as strongly equivalent measures are those which are completely interchangeable in the construction of martingales as in (2).

**Definition 3.4** *Let $\mu$ and $\nu$ be positive measures and let $A \in \{0,1\}^\infty$. We say that $\mu$ and $\nu$ are* strongly equivalent *if there are constants $C > c > 0$ such that for all $A \in \{0,1\}^\infty$ and all $n$, $c < \mu(A[0..n])/\nu(A[0..n]) < C$.*

On the other hand, what we call "weak" equivalence corresponds to the classical notion of *absolute continuity*, i.e., the measure zero sets for both measures are the same.

**Definition 3.5** *Let $\mu$ and $\nu$ be $\Delta$-computable measures. We say that $\mu$ and $\nu$ are* weakly equivalent *if for every $X \subseteq \{0,1\}^\infty$, $\mu_\Delta(X) = 0$ if and only if $\nu_\Delta(X) = 0$.*

There is a simple and useful characterization of strong equivalence in terms of bit probabilities. The following is found in [1].

**Lemma 3.6** *Let $\mu$ and $\nu$ be strongly positive measures; for $A \in \{0,1\}^\infty$ and $i \in \mathbb{N}$ let $u_i = \mu(A[0..i]|A[0..i-1])$ and $v_i = \nu(A[0..i]|A[0..i-1])$. Then $\mu$ and $\nu$ are strongly equivalent if and only if for every $A \in \{0,1\}^\infty$, $\sum_{i=0}^{\infty} |u_i - v_i| < \infty$.*

A fact about strongly equivalent measures which we will use repeatedly is the following "exact computation lemma" for measures, which is found in [1].

**Lemma 3.7** *Let $\mu$ be a strongly positive, $\Delta$-computable measure. Then there is an exactly $\Delta$-computable measure $\nu$ which is strongly equivalent to $\mu$.*

We now turn to the characterization of weak equivalence. A well-known theorem of Kakutani [5] on infinite product measures characterizes the (weak) equivalence of two measures in a manner similar to Lemma 3.6, but in terms of the weaker condition that the *squares* of the differences between bit probabilities are summable. It is also known that Kakutani's theorem holds for constructive measure in the sense of Martin-Löf (see [14]). What we show below is that Kakutani's theorem holds for resource-bounded measure, answering an open question in [1].

**Theorem 3.8** *Let $\mu$ and $\nu$ be strongly positive, $\Delta$-computable measures, and for $A \in \{0,1\}^\infty$ and $i \in \mathbb{N}$ let $u_i = \mu(A[0..i]|A[0..i-1])$ and $v_i = \nu(A[0..i]|A[0..i-1])$. Suppose that for each $A \in \{0,1\}^\infty$, $\sum_{i=0}^{\infty}(u_i - v_i)^2 < \infty$. Then $\mu$ and $\nu$ are weakly equivalent.*

Just as for Kakutani's original result, the converse of the Theorem 3.8 holds if $\mu$ and $\nu$ are *product* measures.

## 4 An invariance property of $\Delta$-randomness

In this section we prove a fundamental new technical result which shows that $\Delta$-randomness is invariant under certain kinds of transformations on $\{0,1\}^\infty$. This result is of independent interest since it extends to $\Delta$-randomness some known results on invariance properties of algorithmic randomness; however, we are primarily concerned here because this result is the key to our understanding of the notion of compressibility in [2] (see Definition 5.1) and the subsequent characterization of $\Delta$-randomness in terms of "maximal compressions."

We begin with a simple idea, namely, the representation of a real number with respect to a given measure. While this topic may at first appear to be a digression, it is surely the most intuitive way to introduce the kinds of transformations we need to discuss. Some of the results here come from the detailed treatment in [6], and as implied above this kind of transformation appears implicitly in [2]; however, the idea can really be traced back to the "isomorphism theorem" for Lebesgue measure (see [4]).

The following is a natural way, given a measure $\mu$, to associate with a given string $\sigma$ a subinterval of $[0,1]$ of width $\mu(\sigma)$.

**Definition 4.1** *Let $\mu$ be a measure and $\sigma \in \{0,1\}^*$. Let $S$ be a maximal, disjoint set of strings which lexicographically precede $\sigma$ (e.g., the lexicographic predecessors of length $|\sigma|$). The* basic $\mu$-interval *$(\sigma)_\mu$ is the interval $[p,q] \subseteq [0,1]$ defined by $p = \sum_{\tau \in S} \mu(\tau)$, $q = p + \mu(\sigma)$. More generally, for any subinterval $[x,y] \subseteq [0,1]$, let $([x,y])_\mu = [\mu([0,x]), \mu([0,y])]$.*

For example, $(\sigma)_\lambda$ is what we would normally call a dyadic interval (remember that $\lambda$ here denotes Lebesgue measure). It is not difficult to see that in the case of Lebesgue measure, the definition below gives the usual interpretation of the binary expansion of a real number in $[0,1]$.

**Definition 4.2** *Let $\mu$ be a strongly positive measure and $A \in \{0,1\}^\infty$. The* real number $\mu$-represented by $A$, *which we denote $(A)_\mu$, is the unique real number $x$ such that*

$$x \in \bigcap_i (A[0..i])_\mu.$$

The main result of this section asserts that the transformation between representations preserves $\Delta$-randomness.

**Theorem 4.3** *Let $\mu$ and $\nu$ be strongly positive, $\Delta$-computable measures. Let $A$ and $B$ be sequences such that $(A)_\mu = (B)_\nu$. Then $A$ is $\Delta$-random w.r.t. $\mu$ if and only if $B$ is $\Delta$-random w.r.t. $\nu$.*

Theorem 4.3 may be interpreted as an assertion that $\Delta$-randomness is an invariant property of real numbers, that is, a property which is independent of the scheme used to represent real numbers as binary sequences. The analog of Theorem 4.3 for algorithmic randomness is well-known; Theorem 4.3 is unusual in that its proof does not seem to depend on computational properties of the transformation between $A$ and $B$. Note that given the $\nu$-representation $B$ of some real number $x$, the $\mu$-representation $A$ of $x$ can be computed as follows: Having determined $A[0..i]$, choose $b = A[i+1]$ so that some interval $(B[0..j])_\nu$ is completely contained in $(A[0..i]b)_\mu$. Although it is not obvious, in most cases of interest it turns out that $A \leq_{tt} B$, and so the analog of Theorem 4.3 for algorithmic randomness follows easily from the fact that algorithmic randomness is always preserved by $tt$-reductions; see [19] or [6]. It is unknown whether $\Delta$-randomness is always preserved by $tt$-reductions, even in the case $\Delta = \text{rec}$, although Breutzmann and Lutz [1] have shown that $\Delta$-randomness is preserved by a certain restricted class of $tt$-reductions.

# 5 Compressibility

Buhrman and Longpré [2] gave the following definition of "compressibility" and the subsequent characterization of $\Delta$-measure zero.

**Definition 5.1** *$A \in \{0,1\}^\infty$ is $f$-compressible if there exists a sequence $B \in \{0,1\}^\infty$ and algorithms $C$ and $D$ such that*

*(i) The algorithm $C$ ("compression"), given an initial segment $A[0..i]$, produces in $f(i)$ steps a finite number of strings, one of which is an initial segment $B[0..j]$ such that $D$, on input $B[0..j]$, produces a prefix of $A$ which properly extends $A[0..i]$.*

*(ii) The algorithm $D$ ("decompression"), given $B[0..j]$, produces in $f(i+j)$ steps a prefix $A[0..i]$; moreover, the value $i - j$ is unbounded.*

**Theorem 5.2** *A set $X \subseteq \{0,1\}^\infty$ has $\Delta$-measure zero if and only if for some $f$ in $\Delta$, every $A$ in $X$ is $f$-compressible.*

In the proof of "measure zero implies compressibility" given in [2], a given martingale succeeding on a sequence $A$ is used to define the "compressed" sequence $B$. It is not difficult to see from the proof in [2] that the sequence $B$ is precisely the standard representation of the real number $x$ whose $\mu$-representation is the original sequence $A$, where $\mu$ is the betting strategy of the martingale (see the proof of Lemma 5.3 for details).

We have seen that if $A$ is $\Delta$-random with respect to $\mu$, then the martingale with strategy $\mu$ is optimal for $A$. Intuitively it would make sense that an "optimal" martingale should yield a "maximal" compression, that is, a compressed sequence which is itself incompressible. To make this precise (and to show that it is true!) we need to first show that a number of the restrictions on martingales imposed in the proof in [2] can be relaxed, in order to show that the optimal martingale of Theorem 3.3 does indeed yield a compression according to Definition 5.1 above.

Although we generalize this result in Theorem 6.2 the following is worth mentioning since it is vastly simpler and uses Definition 5.1 without modification.

**Lemma 5.3** *Let $\mu$ be a strongly positive, exactly $\Delta$-computable measure and let $d$ be the $\lambda$-martingale with strategy $\mu$. Suppose that $\lim_n d(A[0..n]) = \infty$; then $d$ defines a compression of $A$ in the sense of Definition 5.1, where the compressed sequence $B$ is the standard representation of the real number $(A)_\mu$.*

Now suppose $A$ is compressible and is $\Delta$-random with respect to a strongly positive, $\Delta$-computable measure $\mu$. By Lemma 3.7 we can assume $\mu$ is exactly $\Delta$-computable, and by Theorem 3.3 we know that the martingale $d(\sigma) = \mu(\sigma)/\lambda(\sigma)$ is optimal for $A$, and hence it follows from Lemma 3.1 that $\lim_n d(A[0..n]) = \infty$. Thus this optimal $d$ satisfies the hypotheses of Theorem 5.2. Since $A$ is $\Delta$-random w.r.t. $\mu$, Theorem 4.3 shows that the compressed sequence $B$ is $\Delta$-random (with respect to $\lambda$), i.e., $A$ is "maximally compressed" by $d$. Conversely, if $A$ is not $\Delta$-random w.r.t. $\mu$, the compressed sequence $B$ is not $\Delta$-random w.r.t. $\lambda$. Thus we have the following characterization:

**Theorem 5.4** *Let $\mu$ be a strongly positive, $\Delta$-computable measure, and suppose $A \in \{0,1\}^\infty$ is compressible. Let $d$ be the $\lambda$-martingale with strategy $\mu$. Then $A$ is $\Delta$-random w.r.t. $\mu$ if and only if $d$ induces a maximal compression of $A$.*

In the next section we will be able to replace the $\lambda$ in the above characterization with an arbitrary measure $\nu$.

# 6 Compressibility with respect to nonuniform measures

It is not entirely obvious how to extend Definition 5.1 to an arbitrary measure $\nu$, that is, to do so in such a way as to continue to characterize the notion of $\Delta$-measure zero with respect to $\nu$. The "easy" part is to replace the condition "$i - j$ is unbounded" in Definition 5.1(ii) with the condition (3) below (see the proof of Lemma 5.3 for an intuitive justification). The peculiar feature of Lebesgue measure which is used in an essential way in Theorem 5.2 is that every dyadic number $x$ in $[0,1]$ occurs as the endpoint of some basic dyadic interval $(\sigma)_\lambda$, where $|\sigma|$ is no longer than the representation of $x$; thus, any interval $[x, y]$ with dyadic endpoints can be *exactly* and *efficiently* covered with basic $\lambda$-intervals $(\sigma)_\lambda$ (see Lemma A.3).

Very loosely, the proof that measure zero implies compressibility uses the following idea. Given an exact martingale $d$ with strategy $\mu$ succeeding on $A$, the compression algorithm takes an initial segment $\sigma = A[0..i]$, finds the endpoints of the intervals $(\sigma 0)_\mu$ and $(\sigma 1)_\mu$, and produces a list of strings $\tau$ (the "candidates") representing a set of dyadic intervals $(\tau)_\lambda$ which exactly cover $(\sigma 0)_\mu$ and $(\sigma 1)_\mu$. These candidates can be divided into two groups $G_0$ and $G_1$, those which "decompress" into extensions of $\sigma 0$ and $\sigma 1$, respectively. Then in the other direction of the proof ("compressibility implies measure zero") it is the relative measures of $G_0$ and $G_1$ which are used to define $d(\sigma 0)$ and $d(\sigma 1)$. The difficulties arise from the fact that since $(\sigma 0)_\mu$ and $(\sigma 1)_\mu$ cannot be *exactly* covered by basic $\nu$-intervals, the sets $G_0$ and $G_1$ must overlap. The condition (4) below is imposed to ensure that the amount of overlap is controlled. These and related issues are discussed at greater length in Appendix A.5.

**Definition 6.1** *Let $A \in \{0,1\}^\infty$, let $\nu$ be a $\Delta$-computable measure, and let $f$ be a function in $\Delta$. $A$ is* $f$-compressible with respect to $\nu$ *if there exists $B \in \{0,1\}^\infty$, algorithms $C$ and $D$ with running time bounded by $f$, and a summable sequence $\{\epsilon_i\}$, $0 < \epsilon_i < 1$, satisfying the following conditions:*

*(i) The function $C$ ("compression") takes a string $\sigma$ as input and produces a pair $C(\sigma) = \{G_0(\sigma), G_1(\sigma)\}$ of finite sets of strings, called* candidates*, such that each string $\tau$ appearing in $G_0(\sigma b)$ or $G_1(\sigma b)$ extends some $\tau'$ in $G_b(\sigma)$. For every initial segment $\sigma b \sqsubseteq A$, there exists a string $\tau \in G_b(\sigma)$ such that $\tau \sqsubseteq B$ and $\sigma b \sqsubseteq D(\tau) \sqsubseteq A$.*

*(ii) The function $D$ ("decompression") takes strings to strings; whenever any string $\tau \sqsubseteq B$ appears as a candidate in $G_b(\sigma)$ for some $\sigma b \sqsubseteq A$, then $\sigma b \sqsubseteq D(\tau) \sqsubseteq A$. Moreover, given any constant $k$ there is an initial segment $A[0..i] = \sigma b$ such that $G_b(\sigma)$ contains a candidate $B[0..j]$ for which*

$$\frac{\nu(B[0..j])}{\nu(A[0..i])} > k. \tag{3}$$

*(iii) For every $\sigma \sqsubseteq A$, $i = |\sigma|$,*

$$\nu(G_0(\sigma) \cup G_1(\sigma)) \geq \nu(G_0(\sigma)) + \nu(G_1(\sigma)) - \epsilon_i \nu(G_1(\sigma)). \tag{4}$$

*We may say that $A$ is $\Delta$-incompressible if $A$ is not $f$-compressible for any $f \in \Delta$.*

The following theorem then provides the analog to Theorem 5.2.

**Theorem 6.2** *Let $\nu$ be a strongly positive, $\Delta$-computable measure, and let $X \subseteq \{0,1\}^\infty$. Then $X$ has $\Delta$-measure zero with respect to $\nu$ if and only if there is a function $f \in \Delta$ such that every $A \in X$ is $f$-compressible with respect to $\nu$.*

In the course of proving Theorem 6.2 we show that the analog of Lemma 5.3 holds for compressibility with respect to an arbitrary strongly positive measure $\nu$, and thus the analog of Theorem 5.4 holds for compressibility with respect to $\nu$ via similar reasoning.

We now have all the characterizations promised:

**Theorem 6.3** *Let $\mu$ and $\nu$ be strongly positive, $\Delta$-computable measures, and let $d$ be the $\nu$-martingale with strategy $\mu$. Suppose $A \in \{0,1\}^\infty$ is not $\Delta$-random w.r.t. $\nu$. The following are equivalent.*

*(i) $A$ is $\Delta$-random with respect to $\mu$.*

*(ii) $d$ is optimal for $A$.*

*(iii) $d$ induces a maximal compression of $A$ with respect to $\nu$.*

*(iv) $A$ is $\Delta$-incompressible with respect to $\mu$.*

# A Technical appendices

## A.1 Proof of Theorem 3.3

**Theorem 3.3** *Let $\mu$ and $\nu$ be strongly positive, $\Delta$-computable measures, let $d$ be the $\nu$-martingale with strategy $\mu$, and let $A \in \{0,1\}^\infty$. Then $A$ is $\Delta$-random with respect to $\mu$ if and only if $d$ is optimal for $A$.*

*Proof.* Assume that $A$ is $\Delta$-random w.r.t. $\mu$. Let $d$ be the $\nu$-martingale with strategy $\mu$, and let $\tilde{d}$ be any other $\Delta$-computable $\nu$-martingale; suppose its strategy is $\mu'$. Suppose that for every constant $C$ there is an $n$ such that

$$\frac{\tilde{d}(A[0..n])}{d(A[0..n])} \geq C;$$

then evidently the ratio $\mu'(A[0..n])/\mu(A[0..n])$ is unbounded, contradicting the fact that $A$ is $\Delta$-random w.r.t. $\mu$.

Conversely, assume that $A$ is not $\Delta$-random w.r.t. $\mu$. Then there is a $\Delta$-computable $\mu$-martingale $\tilde{d}$ succeeding on $A$; by Lemma 3.1, we can assume that

$$\lim_{n\to\infty} \frac{\mu'(A[0..n])}{\mu(A[0..n])} = \infty,$$

where $\tilde{d}$ has strategy $\mu'$. Thus for every $C$, there is some $n$ such that

$$\frac{\mu'(A[0..n])}{\nu(A[0..n])} = \frac{\mu'(A[0..n])}{\mu(A[0..n])}\frac{\mu(A[0..n])}{\nu(A[0..n])} > C \cdot \frac{\mu(A[0..n])}{\nu(A[0..n])},$$

that is, the $\nu$-martingale with strategy $\mu$ is not optimal for $A$. □

## A.2 Proof of Theorem 3.8

**Theorem 3.8** *Let $\mu$ and $\nu$ be strongly positive, $\Delta$-computable measures, and for $A \in \{0,1\}^\infty$ and $i \in \mathbb{N}$ let $u_i = \mu(A[0..i]|A[0..i-1])$ and $v_i = \nu(A[0..i]|A[0..i-1])$. Suppose that for each $A \in \{0,1\}^\infty$, $\sum_{i=0}^{\infty}(u_i - v_i)^2 < \infty$. Then $\mu$ and $\nu$ are weakly equivalent.*

The following lemma, which is based on a similar argument in [14], is where the square-summability is used.

**Lemma A.1** *Let $\{s_i\}$, $\{t_i\}$, and $\{\delta_i\}$ be sequences of positive real numbers such that $\delta_i \leq s_i, t_i \leq 1 - \delta_i$ and*

$$\sum_{i=0}^{\infty}\left(\frac{(s_i - t_i)}{\delta_i}\right)^2 < \infty.$$

*Then there is a constant $C$ such that for all $n$,*

$$\prod_{i=0}^{n}\sqrt{s_i t_i} \leq \prod_{i=0}^{n}\frac{s_i + t_i}{2} \leq C\prod_{i=0}^{n}\sqrt{s_i t_i}.$$

*Proof of Lemma A.1.* By familiar properties of the logarithm function we know that if $a$ and $b$ are any two real numbers in an interval $[\epsilon, 1-\epsilon]$, with $\epsilon > 0$, then

$$|b - a| \leq |\ln b - \ln a| \leq \frac{1}{\epsilon}|b - a|.$$

We take $a = s_i t_i$ and $b = (s_i + t_i)^2/4$; note that $a, b$ lie in the interval $[\delta_i^2, 1 - \delta_i^2]$ and that

$$b - a = \frac{(s_i + t_i)^2}{4} - s_i t_i = \frac{1}{4}[(s_i + t_i)^2 - 4 s_i t_i] = \frac{1}{4}(s_i - t_i)^2 \geq 0.$$

(The observation that $b > a$ is, of course, the arithmetic-geometric mean inequality of antiquity.) At any rate we may write

$$0 \leq \ln \frac{(s_i + t_i)^2}{4} - \ln s_i t_i \leq \frac{1}{4\delta_i^2}[(s_i + t_i)^2 - 4 s_i t_i] \leq \frac{1}{4\delta_i^2}(s_i - t_i)^2.$$

Since the right-hand side is summable by hypothesis, there is a constant $C_0$ such that for any $n$,

$$0 \leq \sum_{i=0}^{n} \ln \frac{(s_i + t_i)^2}{4} - \sum_{i=0}^{n} \ln s_i t_i \leq C_0,$$

and hence,

$$\prod_{i=0}^{n} s_i t_i \leq \prod_{i=0}^{n} \frac{(s_i + t_i)^2}{4} \leq e^{C_0} \prod_{i=0}^{n} s_i t_i.$$

□

*Proof of Theorem 3.8.* Let $X \subseteq \{0,1\}^\infty$, and suppose that $\mu_\Delta(X) = 0$. It will suffice to show that $\nu_\Delta(X) = 0$, since we may interchange the roles of $\mu$ and $\nu$ without otherwise modifying the proof. We may regard $X$ as the union $X_0 \cup X_1$, where

$$X_0 = \left\{ A \in X : \frac{\nu(A[0..n])}{\mu(A[0..n])} \text{ is bounded} \right\}$$

and $X_1 = X - X_0$. Let $d$ be a $\mu$-martingale witnessing that $\mu_\Delta(X) = 0$. Then the function $d_0$ defined for all $\sigma \in \{0,1\}^*$ by

$$d_0(\sigma) = \frac{\mu(\sigma)}{\nu(\sigma)} d(\sigma)$$

is a $\nu$-martingale succeeding on every $A \in X_0$.

It will take a bit more work to deal with $X_1$. First define a measure $\eta$ by specifying bit probabilities

$$\eta(\sigma b|\sigma) = \begin{cases} 2\nu(\sigma b|\sigma) - \mu(\sigma b|\sigma) & \text{if this value is in } [\delta/3, 1 - \delta/3] \\ \nu(\sigma b|\sigma) & \text{otherwise.} \end{cases}$$

(Here $\delta$ is the constant appearing in the hypothesis of the theorem.) Let $d_1$ be the $\nu$-martingale with strategy $\eta$. Fix $A \in X_1$ and let

$$\begin{aligned} u_i &= \mu(A[0..i]|A[0..i-1]), \\ v_i &= \nu(A[0..i]|A[0..i-1]), \\ \text{and } w_i &= \eta(A[0..i]|A[0..i-1]). \end{aligned}$$

For all sufficiently large $i$, $|v_i - u_i| < \delta/3$, and hence $2v_i - u_i$ is in $[\delta/3, 1 - \delta/3]$; there is a constant $J$ such that for all $i \geq J$, $w_i = 2v_i - u_i$ and hence

$$v_i = \frac{u_i + w_i}{2}.$$

Moreover, since $|u_i - w_i| \leq 2|u_i - v_i|$, we have

$$\sum_{i=0}^{\infty}(u_i - w_i)^2 < \infty.$$

Using Lemma A.1 twice, there is a constant $C$ such that, for any $n > J$,

$$\frac{\prod_{i=J}^{n} w_i}{\prod_{i=J}^{n} v_i} \geq \frac{\prod_{i=J}^{n} w_i}{C\prod_{i=J}^{n} \sqrt{w_i u_i}} = \frac{\prod_{i=J}^{n} \sqrt{w_i u_i}}{C\prod_{i=J}^{n} u_i} \geq \frac{\prod_{i=J}^{n} v_i}{C^2\prod_{i=J}^{n} u_i}.$$

It follows that for some positive constant $K$,

$$d_1(A[0..n]) = \frac{\prod_{i=0}^{n} w_i}{\prod_{i=0}^{n} v_i} \geq K\frac{\prod_{i=J}^{0} v_i}{\prod_{i=0}^{n} u_i} = K\frac{\nu(A[0..n])}{\mu(A[0..n])}.$$

Since the right-hand side is unbounded, $d_1$ succeeds on $A$. It follows that $d_1$ succeeds on every $A \in X_1$, and hence the $\nu$-martingale $d_0 + d_1$ succeeds on every $A \in X$.

It is not difficult to check that if $\mu$, $\nu$, and $d$ are $\Delta$-computable, so are $\eta$, $d_0$, and $d_1$. □

## A.3 Proof of Theorem 4.3

**Theorem 4.3** *Let $\mu$ and $\nu$ be strongly positive, $\Delta$-computable measures. Let $A$ and $B$ be sequences such that $(A)_\mu = (B)_\nu$. Then $A$ is $\Delta$-random w.r.t. $\mu$ if and only if $B$ is $\Delta$-random w.r.t. $\nu$.*

The following technical lemma is used in several places in this paper.

**Lemma A.2** *Let $\mu$ and $\nu$ be strongly positive, $\Delta$-computable measures, and let $\delta > 0$ such that all bit probabilities for $\mu$ and $\nu$ lie in $[\delta, 1-\delta]$. Let $A$ and $B$ be the $\mu$- and $\nu$-representations, respectively, of some real number; i.e., $(A)_\mu = (B)_\nu$. For each $j$, let $A[0..i_j]$ denote the largest initial segment of $A$ for which $(B[0..j])_\nu \subseteq (A[0..i_j])_\mu$. Then for infinitely many $j$, $\nu(B[0..j]) \geq \delta^2\mu(A[0..i_j])$.*

*Proof of Lemma A.2.* Let $k \in \mathbb{N}$ and consider the intervals $(B[0..k])_\nu$ and $(A[0..i_k])_\mu$. Let $z$ be the boundary between $(A[0..i_k]0)_\mu$ and $(A[0..i_k]1)_\mu$. Note that $z$ is in the interior of $(B[0..k])_\nu$, by the definition of $i_k$. Let $j$ be the least integer such that the interior of $(B[0..j])_\nu$ does not contain $z$, so the interval $(B[0..j-1])_\nu$ does contain $z$. We consider the case that $z$ lies to the left of $(B[0..j])_\nu$; the other case is similar. Note that then $z$ is also to the left of $(A[0..i_j])_\mu$, and by the definition of $i_j$, the right endpoint $z'$ of $(A[0..i_j]0)_\mu$ lies in the interior of $(B[0..j])_\nu$, so $(A[0..i_j]0)_\mu \subseteq (B[0..j-1])_\nu$. Thus

$$\delta\mu(A[0..i_j]) \leq \mu(A[0..i_j]0) \leq \nu(B[0..j-1]) \leq \frac{1}{\delta}\nu(B[0..j]).$$

□

*Proof of Theorem 4.3.* Since the roles of $\mu$ and $\nu$ are interchangeable, we will prove just the "only if" direction. Let $\delta > 0$ such that all bit probabilities for $\mu$ and $\nu$ are in $[\delta, 1-\delta]$. Suppose that $B$ is not $\Delta$-random with respect to $\nu$; then some $\Delta$-computable $\nu$-martingale $\tilde{d}$ succeeds on $B$. Let $\eta$ denote the strategy of $\tilde{d}$; by Lemmas 3.1 and 3.7 we may assume w.l.o.g. that $\eta$ is strongly positive, $\Delta$-computable, and that

$$\lim_j \frac{\eta(B[0..j])}{\nu(B[0..j])} = \infty. \tag{5}$$

The idea of the proof is simple: Let $d$ denote the $\nu$-martingale with strategy $\mu$; we want to somehow combine the strategies for $d$ and $\tilde{d}$ to obtain a $\nu$-martingale for $A$ which dominates $d$. If $A$ were $\Delta$-random with respect to $\mu$, by Theorem 3.3 the martingale $d$ would be optimal, so this will imply that $A$ is not $\Delta$-random w.r.t. $\mu$.

Although we have defined measures as functions on strings to facilitate discussing computability, at this point it may be helpful to think of a measure as a function on subintervals of $[0,1]$; thus $\mu(\sigma)$ could be interpreted as an abbreviation for $\mu((\sigma)_\lambda)$.

First define a measure $\nu'$ as follows: Given any string $\sigma$, let $\{\tau_i\}$ denote a sequence of disjoint basic $\nu$-intervals covering $(\sigma)_\lambda$, that is,

$$\bigcup_i (\tau)_\nu = (\sigma)_\lambda,$$

and let

$$\nu'(\sigma) = \sum_i \lambda(\tau_i).$$

The measure $\nu'$ has the one desirable property that for any $\sigma$,

$$((\sigma)_\nu)_{\nu'} = (\sigma)_\lambda.$$

Note that in general, an infinite sequence of $\nu$-intervals may be required to exactly cover a given dyadic interval $\sigma$, so $\nu'$ may not be dyadic-valued even if $\nu$ is. However, $\nu'$ is $\Delta$-computable.

Now define a measure $\mu'$ by

$$\mu'(\sigma) = \eta(((\sigma)_\mu)_{\nu'}).$$

It is tedious, but more or less straightforward, to show that $\mu'$ is $\Delta$-computable. (It is fairly simple to show that $\mu'$ is $\Delta$-computable if $\mu$, $\nu$, and $\eta$ are *exactly* computable, which would be sufficient for all the other results in this paper.) Note that in the case $\nu = \lambda$, $\nu' = \lambda$ also, so $\mu'(\sigma)$ reduces to $\eta((\sigma)_\mu)$.

We know from Lemma A.2 that for infinitely many $j \in \mathbb{N}$, $\nu(B[0..j]) \geq \delta^2 \mu(A[0..i_j]$, where $i_j$ is the largest integer for which $(B[0..j])_\nu \subseteq (A[0..i_j])_\mu$. Then for any such $j$ we have

$$\begin{aligned}
\frac{\mu'(A[0..i_j])}{\nu(A[0..i_j])} &= \frac{\eta(((A[0..i_j])_\mu)_{\nu'})}{\nu(A[0..i_j])} \\
&\geq \frac{\eta(((B[0..j])_\nu)_{\nu'})}{\nu(A[0..i_j])} \\
&= \frac{\eta(B[0..j])}{\nu(A[0..i_j])} \\
&= \frac{\eta(B[0..j])}{\nu(B[0..j])} \cdot \frac{\nu(B[0..j])}{\nu(A[0..i_j])} \\
&\geq \frac{\eta(B[0..j])}{\nu(B[0..j])} \cdot \frac{\delta^2 \mu(A[0..i_j])}{\nu(A[0..i_j])}.
\end{aligned}$$

It follows from (5) that given any constant $C$ there is a $j$ such that

$$\frac{\mu'(A[0..i_j])}{\nu(A[0..i_j])} \geq \frac{\mu(A[0..i_j])}{\nu(A[0..i_j])} \cdot C,$$

that is, the strategy $\mu$ is not optimal for $A$. It follows from Theorem 3.3 that $A$ is not $\Delta$-random w.r.t. $\mu$. □

## A.4 Proof of Lemma 5.3

**Lemma 5.3** *Let $\mu$ be a strongly positive, exactly $\Delta$-computable measure and let $d$ be the $\lambda$-martingale with strategy $\mu$. Suppose that $\lim_n d(A[0..n]) = \infty$; then $d$ defines a compression of $A$ in the sense of Definition 5.1, where the compressed sequence $B$ is the standard representation of the real number $(A)_\mu$.*

*Proof.* The idea is that each initial segment $A[0..i]$ determines a subinterval $(A[0..i])_\mu$ of $[0,1]$, and the sequence $B$ is the standard binary representation of the real number $(A)_\mu = \bigcap_i (A[0..i])_\mu$. The "decompression" algorithm is essentially the reduction from $B$ to $A$ described following Theorem 4.3; given an initial segment $B[0..j]$, find the longest string $\sigma$ such that $(B[0..j])_\lambda \subseteq (\sigma)_\mu$. It necessarily follows that $\sigma = A[0..i]$ is the longest initial segment of $A$ such that $(B[0..j])_\lambda \subseteq (A[0..i])_\mu$. Intuitively the intervals $(A[0..i])_\mu$ and $(B[0..j])_\lambda$ should represent about the same amount of information, i.e., the width of $(A[0..i])_\mu$ should be of the same order as the width of $(B[0..j])_\lambda$, or about $2^{-j}$, so we would expect

$$\frac{2^{-j}}{2^{-i}} \approx \frac{\mu(A[0..i])}{\lambda(A[0..i])}. \tag{6}$$

Since the right-hand side is unbounded, it would then follow that $i - j$ is unbounded.

Of course, (6) is not literally true, but by Lemma A.2 there are infinitely many $j \in \mathbb{N}$ such that the decompression algorithm produces an initial segment $A[0..i_j]$ satisfying $\lambda(B[0..j]) \geq \delta^2 \mu(A[0..i_j])$. Since $\lim_n d(A[0..n]) = \infty$, this is sufficient to conclude that $i - j$ is unbounded.

The "compression" algorithm is obtained by applying Lemma A.3 to each of the two subintervals $(A[0..i]0)_\mu$ and $(A[0..i]1)_\mu$, to produce the list of strings required by the definition. One of these strings must be an initial segment $B[0..j]$; since $(B[0..j])_\lambda$ is completely contained in either $(A[0..i]0)_\mu$ or $(A[0..i]1)_\mu$, the decompression algorithm produces a proper extension of $A[0..i]$. □

## A.5 Proof of Theorem 6.2.

We repeat the definition here for convenience.

**Definition 6.1** *Let $A \in \{0,1\}^\infty$, let $\nu$ be a $\Delta$-computable measure, and let $f$ be a function in $\Delta$. $A$ is $f$-*compressible with respect to $\nu$ *if there exists $B \in \{0,1\}^\infty$, algorithms $C$ and $D$ with running time bounded by $f$, and a summable sequence $\{\epsilon_i\}$, $0 < \epsilon_i < 1$, satisfying the following conditions:*

*(i) The function $C$ ("compression") takes a string $\sigma$ as input and produces a pair $C(\sigma) = \{G_0(\sigma), G_1(\sigma)\}$ of finite sets of strings, called* candidates*, such that each string $\tau$ appearing in $G_0(\sigma b)$ or $G_1(\sigma b)$ extends some $\tau'$ in $G_b(\sigma)$. For every initial segment $\sigma b \sqsubseteq A$, there exists a string $\tau \in G_b(\sigma)$ such that $\tau \sqsubseteq B$ and $\sigma b \sqsubseteq D(\tau) \sqsubseteq A$.*

*(ii) The function $D$ ("decompression") takes strings to strings; whenever any string $\tau \sqsubseteq B$ appears as a candidate in $G_b(\sigma)$ for some $\sigma b \sqsubseteq A$, then $\sigma b \sqsubseteq D(\tau) \sqsubseteq A$. Moreover, given any constant $k$ there is an initial segment $A[0..i] = \sigma b$ such that $G_b(\sigma)$ contains a candidate $B[0..j]$ for which*

$$\frac{\nu(B[0..j])}{\nu(A[0..i])} > k. \tag{7}$$

*(iii) For every $\sigma \sqsubseteq A$, $i = |\sigma|$,*

$$\nu(G_0(\sigma) \cup G_1(\sigma)) \geq \nu(G_0(\sigma)) + \nu(G_1(\sigma)) - \epsilon_i \nu(G_1(\sigma)). \tag{8}$$

*We may say that $A$ is $\Delta$-incompressible if $A$ is not $f$-compressible for any $f \in \Delta$.*

**Theorem 6.2** *Let $\nu$ be a strongly positive, $\Delta$-computable measure, and let $X \subseteq \{0,1\}^\infty$. Then $X$ has $\Delta$-measure zero with respect to $\nu$ if and only if there is a function $f \in \Delta$ such that every $A \in X$ is $f$-compressible with respect to $\nu$.*

**Remarks** The ideas behind the changes in the definition are as follows. Suppose we have an exactly computable martingale with strategy $\mu$ succeding on $A$. The "compression" algorithm as originally given in [2] essentially takes a string $\sigma \sqsubseteq A$, finds the endpoints of the $\mu$-intervals $(\sigma 0)_\mu$ and $(\sigma 1)_\mu$ (which are dyadic numbers), and outputs a list of strings $\tau$ representing a collection of dyadic intervals $(\tau)_\lambda$ which exactly cover each of the two $\mu$-intervals. Since the compressed sequence $B$ is the unique member of the intersection of all the intervals $(A[0..i])_\mu$, one of these candidates $\tau$ must be an initial segment of $B$. On the other hand to "decompress" an initial segment $\tau \sqsubseteq B$, the algorithm finds the smallest $\mu$-interval $(\sigma)_\mu$ completely containing $(\tau)_\lambda$; it must be the case that $\sigma \sqsubseteq A$.

To go the other direction ("compressibility implies measure zero") the strategy for determining the values of $d(\sigma 0)$ and $d(\sigma 1)$, $d(\sigma)$ already having been defined, is as follows: First take the list of candidates produced on input $\sigma$ and eliminate those which don't extend a candidate from the immediate predecessor of $\sigma$. Next use the decompression algorithm $D$ to divide them into two groups $G_0$ and $G_1$, those which decompress into an extension of $\sigma 0$ and $\sigma 1$, respectively. The relative weights (in terms of Lebesgue measure) of the sets $G_0$ and $G_1$ are used to define the proportions of the capital allocated to betting on 0 and 1 respectively, i.e., the betting strategy for the martingale.

The problem that arises when we replace Lebesgue measure $\lambda$ with an arbitrary measure $\nu$ is that it may not be possible to exactly cover $(\sigma 0)_\mu$ and $(\sigma 1)_\mu$ with a finite number of $\nu$-intervals, since not every dyadic number necessarily occurs as an endpoint of some $\nu$-interval. It is essential that *every* set of candidates produced for a valid initial segment of $A$ include a true initial segment of $B$; therefore, the candidates covering $(\sigma 0)_\mu$ and $(\sigma 1)_\mu$, and thus the sets $G_0$ and $G_1$, must overlap. This is not a problem in itself as long as the amount of overlap can be controlled, and the condition (8) is imposed in order to do so.

The more subtle difficulty is that if the candidates (the $\nu$-intervals determined by the candidates, that is) covering the interval $(\sigma 0)_\mu$ overlap the interval $(\sigma 1)_\mu$ (and vice versa), how can the decompression algorithm determine which ones should be counted in the sets $G_0$ and $G_1$ for the purpose of determining the relative weights? The trick is to take advantage of the fact that the decompression algorithm can operate with full knowledge of how the compression algorithm works; that is, the decompression algorithm, given a string $\tau$, can check whether $\tau$ was given by the compression algorithm as a candidate to cover $(\sigma 0)_\mu$ or $(\sigma 1)_\mu$ and give an answer which is in agreement. (The sets of *strings* given as candidates to cover the two intervals must be disjoint, of course; but the $\nu$-intervals the candidates define have to overlap.) This is why in condition (ii) we can only insist that the decompression algorithm produce a correct initial segment $\sigma b$ of $A$ for those initial segments $\tau \sqsubseteq B$ which are actually given as candidates to cover the correct interval $(\sigma b)_\mu$.

Note that since the compression and decompression algorithms can simulate one another, there would be no real loss of generality in the original definition if one were to require the compression algorithm $C$ to produce the sets $G_0$ and $G_1$ directly, as we have done in (i). We have also formalized in (i) an assumption, made tacitly in [2], that when $\sigma' \sqsubseteq \sigma \sqsubseteq A$, the initial segment of $B$ produced by $C(\sigma)$ is an extension of the initial segment of $B$ produced by $C$ on $\sigma'$; otherwise, the sets $G_i$ could not be assumed to always include an initial segment of $B$.

Note that when $\nu$ is Lebesgue measure, condition (7) reduces to "$i - j$ is unbounded" as given in the original definition.

We now proceed with the proof of Theorem 6.2.

### A.5.1 Measure zero implies compressibility

Let $d$ be a martingale witnessing that $X \subseteq \{0,1\}^\infty$ has $\Delta$-measure zero with respect to $\nu$. Let $\mu$ be the strategy of the martingale $d$. By Lemmas 3.1 and 3.7 we may assume that $\mu$ and $\nu$ are exactly $\Delta$-computable, that $\mu$ is strongly positive ($\nu$ is strongly positive by hypothesis) and that for each $A \in X$, $\lim_n d(A[0..n]) = \infty$. Let $\delta > 0$ be a real number such that all the bit probabilities for $\mu$ and $\nu$ are in $[\delta, 1-\delta]$. Let $A$ be any sequence in $X$ and let $B$ denote the sequence such that $(A)_\mu = (B)_\nu$. We construct a compression of $A$ according to Definition 6.1.

The following two simple results will be needed in order to define the compression and decompression algorithms.

**Lemma A.3** *Let $x$, $y \in \mathbb{D}$ have representations at most $n$ bits long. Then $[x, y]$ can be exactly covered, in a unique way, with fewer than $2n$ nonoverlapping dyadic intervals of maximal width, each of which has a representation of at most $n$ bits. More generally, for any measure $\nu$, suppose $x$ is the left endpoint of some interval $(\sigma)_\nu \subseteq [x, y]$ and $y$ is the right endpoint of an interval $(\sigma')_\nu \subseteq [x, y]$, and that $|\sigma|, |\sigma'| \leq n$. Then $[x, y]$ can be exactly covered, in a unique way, with fewer than $2n$ nonoverlapping basic $\nu$-intervals of maximal width, each of which has a representation of at most $n$ bits.*

**Lemma A.4** *There is a constant $M$, depending on $\delta$, such that for any $\sigma, \tau \in \{0,1\}^*$, and $i > 0$, if $|\tau| \geq Mi$, then*

$$\nu(\sigma\tau) \leq \frac{\delta}{2^{i+1}}\nu(\sigma).$$

We first define, inductively, the function $C$. For any initial segment $\sigma$ and $b \in \{0,1\}$, $G_b(\sigma)$ is a set of strings representing adjacent basic $\nu$-intervals; with a slight abuse of notation we can define a subinterval $(G_b(\sigma))_\nu$ of $[0,1]$ by

$$(G_b(\sigma))_\nu = \bigcup\{(\tau)_\nu : \tau \in G_b(\sigma)\}.$$

We shall see that $(G_b(\sigma))_\nu$ covers the interval $(\sigma b)_\mu$ and exceeds it in total width by at most $2^{i-1}\mu(\sigma b)$, where $i = |\sigma|$. Moreover, the intersection of $(G_0(\sigma))_\nu$ and $(G_1(\sigma))_\nu$ will be no more than $2^{-i}\mu(\sigma 1)$ in width. It is also important to note that although the intervals $(G_0(\sigma))_\nu$ and $(G_1(\sigma))_\nu$ may overlap, $G_0(\sigma)$ and $G_0(\sigma)$, as sets of strings, will always be disjoint. Although the idea is quite simple, we go into some detail to define the sets $G_b(\sigma)$ in order to be able to verify the computability requirements.

Given an initial segment $\sigma$ define $x$, $y$, and $z$ in $\mathbb{D}$ by

$$\begin{aligned}(\sigma)_\mu &= [x, y]\\ (\sigma 0)_\mu &= [x, z]\\ (\sigma 1)_\mu &= [z, y].\end{aligned}$$

We next identify four intervals $I_x$, $I_y$, $I_{z,\text{left}}$ and $I_{z,\text{right}}$, containing $x$, $y$, $z$, and $z$, respectively. Each of these intervals is either a basic $\nu$-interval or else is a single point $\{x\}$, $\{y\}$, or $\{z\}$. In a moment we shall give a careful construction and show that each interval has width no more than $2^{-i}\mu(\sigma b)$, where $i = |\sigma|$ and $b \in \{0,1\}$. However, we now define $C(\sigma) = \{G_0(\sigma), G_1(\sigma)\}$; first let

$$\begin{aligned} x_0 &= \text{right endpoint of } I_x, \\ z_0 &= \text{left endpoint of } I_{z,\text{left}}, \\ z_1 &= \text{right endpoint of } I_{z,\text{right}}, \text{ and} \\ y_1 &= \text{left endpoint of } I_y. \end{aligned}$$

Then $G_0(\sigma)$ consists of:

> the unique covering of $[x_0, z_0]$ provided by Lemma A.3, *and*
> the string $\tau_x$ such that $(\tau_x)_\nu = I_x$, if $x_0 \neq x$, *and*
> the string $\tau_{z,\text{left}}$ such that $(\tau_{z,\text{left}})_\nu = I_{z,\text{left}}$, if $z_0 \neq z$.

Likewise, $G_1(\sigma)$ consists of

> the unique covering of $[z_1, y_1]$ provided by Lemma A.3, *and*
> the string $\tau_{z,\text{right}}$ such that $(\tau_{z,\text{right}})_\nu = I_{z,\text{right}}$, if $z_1 \neq z$, *and*
> the string $\tau_y$ such that $(\tau_y)_\nu = I_y$, if $y_1 \neq y$.

In the initial stage where $\sigma$ is the empty string, so $x = 0$ and $y = 1$, let $I_x = \{0\}$ and $I_y = \{1\}$; thus $x_0 = 0$ and $y_0 = 1$. Let $\tau'_z = \lambda$ and proceed to find $I_{z,\text{left}}$ and $I_{z,\text{right}}$ just as described below, using $i = 1$.

If $\sigma$ is not the empty string, let $i = |\sigma|$, let $b \in \{0,1\}$ and $\sigma' \in \{0,1\}^*$ such that $\sigma' b = \sigma$, and assume inductively that $(\sigma)_\mu \subseteq (G_b(\sigma'))_\nu$. Find the string $\tau'_x \in G_b(\sigma')$ such that $(\tau'_x)_\nu$ contains $x$, and extend $\tau'_x$ by $Mi$ bits to obtain a string $\tau_x$ for which $x \in (\tau_x)_\nu$, where $M$ is the constant defined in Lemma A.4. If any string $\tau$ with $\tau'_x \sqsubseteq \tau \sqsubseteq \tau_x$ has $x$ as an endpoint, then let $I_x = \{x\}$; otherwise, let $I_x = (\tau_x)_\nu$. The definition of $I_y$ is completely analogous. For $z$, begin in the same way: let $\tau'_z$ be the string in $G_b(\sigma')$ such that $(\tau'_z)_\nu$ contains $z$, and let $\tau_z$ be its extension by $Mi$ bits with $z \in (\tau_z)_\nu$. If any interval $\tau$, $\tau'_z \sqsubseteq \tau \sqsubseteq \tau_z$ has $z$ for an endpoint, let $I_{z,\text{left}} = I_{z,\text{right}} = \{z\}$. If not, let $\tau_{z,\text{left}} = \tau_z$, and consider the extensions $\tau_z 0$ and $\tau_z 1$ of $\tau_z$, one of which contains $z$. If $(\tau_z 0)_\nu$ contains $z$, let $\tau_{z,\text{right}} = \tau_z 0$; otherwise let $\tau_{z,\text{right}} = \tau_z 1$. Let $I_{z,\text{left}} = (\tau_{z,\text{left}})_\nu$ and $I_{z,\text{right}} = (\tau_{z,\text{right}})_\nu$.

It is clear from the construction that

$$\begin{aligned} \nu(G_0(\sigma)) &\leq \mu(\sigma 0) + \lambda(I_x) + \lambda(I_{z,\text{left}}) \qquad (9) \\ \text{and} \quad \nu(G_1(\sigma)) &\leq \mu(\sigma 1) + \lambda(I_{z,\text{right}}) + \lambda(I_y), \end{aligned}$$

where $\lambda(I)$ is the width (Lebesgue measure) of an interval $I \subseteq [0,1]$ (and likewise $\mu(\sigma b)$ is the width of the interval $(\sigma b)_\mu$). Claim A.7 below shows that the width of any of the intervals $I_x$, $I_y$, $I_{z,\text{left}}$ and $I_{z,\text{right}}$, and hence the width of the intersection of $(G_0(\sigma))_\nu$ and $(G_1(\sigma))_\nu$, is no more than $2^{-i}\mu(\sigma 1)$.

The function $D$ is defined as follows: Given any string $\tau$, first find the longest string $\sigma$ such that $(\tau)_\nu \subseteq (\sigma)_\mu$. If $\tau \in G_b(\sigma')$ for some extension $\sigma'$ of $\sigma$, let $D(\tau) = \sigma'$ for the shortest such $\sigma'$; otherwise let $D(\tau) = \sigma$.

We now need to verify that the functions $C$ and $D$ indeed have all the properties required by conditions (i), (ii), and (iii) of Definition 6.1. We isolate some technical necessities in the following claims.

**Claim A.5** *Suppose $(\tau)_\nu \subseteq (\sigma)_\mu$ and that $(\tau)_\nu$ properly contains the boundary point $z$ between $(\sigma 0)_\mu$ and $(\sigma 1)_\mu$. Then there is at most one extension $\sigma'$ of $\sigma$ and one $b' \in \{0,1\}$ such that $\tau \in C_{b'}(\sigma')$.*

*Proof of Claim A.5.* Let $\hat{\tau}$ be the shortest prefix of $\tau$ such that $(\hat{\tau})_\nu \subseteq (\sigma)_\mu$. It follows from the maximality condition of Lemma A.3 that $\hat{\tau} \in G_b(\hat{\sigma})$, where $\hat{\sigma}b = \sigma$. By construction there is an extension of $\hat{\tau}$ by $Mi$ bits in $G_0(\sigma)$ and an extension of $\hat{\tau}$ by $Mi + 1$ bits in $G_1(\sigma)$, where $i = |\sigma|$ Since $\tau$ extends $\hat{\tau}$ and includes the endpoint $z$ between $(\sigma 0)_\mu$ and $(\sigma 1)_\mu$, $\tau$ can only occur in classes of the form $G_1(\sigma 01^k)$ or $G_0(\sigma 10^k)$ for some $k$ (since $z$ is an endpoint of these intervals only); $\tau$ is in the former only if its length $|\tau|$ is equal to

$$|\hat{\tau}| + Mi + M(i+1) + \cdots + M(i+k-1) \tag{10}$$

and $\tau$ is in the latter only if its length is exactly the quantity in (10), plus 1. □

**Claim A.6** *For any strings $\sigma$, $\tau$ and $b \in \{0,1\}$, let $\sigma'$ be the longest string for which $(\tau)_\nu \subseteq (\sigma')_\mu$. If $\tau \in G_b(\sigma)$, then either $D(\tau) = \sigma b$, or $D(\tau) = \sigma'$ and $\sigma b \sqsubseteq \sigma'$.*

*Proof of Claim A.6.* If $\sigma'$ is the longest string such that $(\tau)_\nu \subseteq (\sigma')_\mu$, it necessarily follows that $(\tau)_\nu$ contains the boundary point $z$ between $(\sigma' 0)_\mu$ and $(\sigma' 1)_\mu$. Now if $\sigma' \sqsubseteq \sigma$, we have $D(\tau) = \sigma b$ by Claim A.5 and the definition of $D$.

Suppose on the other hand that $\sigma'$ is a proper extension of $\sigma$. Note first that this implies that $(\tau)_\nu$ does not contain the boundary point between $(\sigma 0)_\mu$ and $(\sigma 1)_\mu$, that is, $(\tau)_\nu \subseteq (\sigma b)_\mu$, so we know $\sigma b \sqsubseteq \sigma'$. If, as in the proof of Claim A.5, we let $\hat{\tau}$ denote the shortest prefix of $\tau$ such that $(\hat{\tau})_\nu \subseteq (\sigma')_\mu$, then in fact $\hat{\tau} = \tau$, so $\tau$ is in $C_{b'}(\hat{\sigma}')$, where $\sigma' = \hat{\sigma}' b'$. Thus, since $\tau$ contains the endpoint $z$, only extensions of $\tau$ by at least $Mi$ bits can appear in sets $C_d(\rho)$ for extensions $\rho$ of $\sigma'$ and $d \in \{0,1\}$, but not $\tau$ itself. Hence by the definition of $D$ we have $D(\tau) = \sigma'$. □

**Claim A.7** *For any $\sigma$, let $i = |\sigma|$ and let $I$ denote the widest of the intervals $I_x$, $I_y$, $I_{z,\mathrm{left}}$ and $I_{z,\mathrm{right}}$ defined above, and let $b \in \{0,1\}$. Then*

*(i) $\lambda(I) \leq 2^{-i}\mu(\sigma b)$, and*

*(ii) $\nu(G_b(\sigma)) \leq 2\mu(\sigma b)$.*

*Proof of Claim A.7.* We induct on the length of $\sigma$, using Lemma A.4, the inequalities (9), and the fact that for any string $\gamma$, $\mu(\gamma b) \geq \delta\mu(\gamma)$. In the base case $\sigma = \lambda$, then

$$\lambda(I) \leq \frac{\delta}{2} \leq \frac{1}{2}\mu(b),$$

and therefore

$$\nu(G_b(\lambda)) \leq \mu(b) + \lambda(I) \leq \mu(b) + \frac{1}{2}\mu(b) \leq 2\mu(b).$$

(Here we have used the fact that $\lambda(I_x) = \lambda(I_y) = 0$ in the case that $\sigma$ is the empty string.)

For the induction step let $\sigma = \sigma' b$, $i = |\sigma|$, and assume that

$$\frac{1}{2}\nu(G_b(\sigma')) \leq \mu(\sigma).$$

Then

$$\begin{aligned} \lambda(I) &\leq \frac{\delta}{2^{i+1}}\nu(G_b(\sigma')) \text{ by Lemma A.4} \\ &\leq \frac{\delta}{2^i}\mu(\sigma) \text{ by the induction hypothesis} \\ &\leq \frac{1}{2^i}\mu(\sigma b), \end{aligned}$$

and so

$$\nu(G_b(\sigma)) \leq \mu(\sigma b) + 2\lambda(I) \leq \mu(\sigma b)\left(1 + \frac{2}{2^i}\right) \leq \mu(\sigma b) \cdot 2,$$

completing the induction. □

To verify (i), note that by construction and by the maximality of the covering given by Lemma A.3, it is always the case that every string in $G_0(\sigma b)$ or in $G_1(\sigma b)$ extends some string in $G_b(\sigma)$. Given any initial segment $\sigma b \sqsubseteq A$, the real number $(B)_\nu$ is in $(\sigma b)_\mu$ and hence is in the interval $(G_b(\sigma))_\nu$, and so some $\tau \in (G_b(\sigma))_\nu$ is an initial segment of $B$. Let $\sigma'$ be the longest string for which $(\tau)_\nu \subseteq (\sigma')_\mu$; note that $\sigma' \sqsubseteq A$. By Claim A.6, either $D(\tau) = \sigma b \sqsubseteq A$ or else $D(\tau) = \sigma' \sqsubseteq A$.

For (ii), Let $\tau \sqsubseteq B$, and suppose $\tau \in G_b(\sigma)$ for some $\sigma b \sqsubseteq A$. By the argument in (i), $D(\tau) \sqsubseteq A$. By Lemma A.2, there are infinitely many $j \in \mathbb{N}$ such that $(B[0..j])_\nu \subseteq (A[0..i_j])_\mu$ and $\nu(B[0..j]) \geq \delta^2 \mu(A[0..i_j])$. Let $\sigma b = A[0..i_j]$; then by the maximality condition of Lemma A.3 some initial segment $\tau \sqsubseteq B[0..j]$ is in $G_b(\sigma)$. Thus we have

$$\frac{\nu(\tau)}{\nu(A[0..i_j])} \geq \frac{\nu(B[0..j])}{\nu(A[0..i_j])} \geq \frac{\delta^2 \mu(A[0..i_j])}{\nu(A[0..i_j])}.$$

Since the right-hand term approaches $\infty$, the condition (7) is satisfied.

For (iii), by Claim A.7 the intersection $(G_0(\sigma))_\nu \cap (G_1(\sigma))_\nu$ has width less than $2^{-i}\mu(\sigma 1)$, which is less than $2^{-i}\nu(G_1(\sigma))$, where $i = |\sigma|$. Thus

$$\nu(G_0(\sigma) \cup G_1(\sigma)) \geq \nu(G_0(\sigma)) + \nu(G_1(\sigma)) - 2^i \nu(G_1(\sigma)),$$

whence (iii) follows once we take $\epsilon_0 = \frac{1}{2}$ and $\epsilon_i = 2^{-i}$.

Verifying that the time bounds can be satisfied uniformly by a function $f$ in $\Delta$ is not difficult...

### A.5.2 Compressibility implies measure zero

Let $f \in \Delta$, and assume that each $A \in X$ is $f$-compressible. For each $A \in X$, we will construct a $\Delta$-computable $\nu$-martingale. Since (as will be apparent) each such martingale will satisfy the same resource bound in $\Delta$, the set $X$ will be seen to be the $\Delta$-union of the $\Delta$-measure zero sets $\{A\}$ (see [10] or [11]), and hence $\nu_\Delta(X) = 0$.

Let $A \in X$, and let $B$ denote the "compressed" sequence of Definition 6.1. Let $G(\sigma) = \nu(G_0(\sigma) \cup G_1(\sigma))$. To define a martingale $d$ succeeding on $A$, let $d(\lambda) = 1$. Having defined $d(\sigma)$ we find $\nu(G_0(\sigma))$ and $\nu(G_1(\sigma))$ and use their relative proportions in $\nu(G(\sigma))$ to define the betting strategy at $\sigma$. If $G(\sigma)$ is empty, let $d(\sigma 0) = d(\sigma 1) = d(\sigma)$. Otherwise, let

$$\begin{aligned} r_\sigma &= \frac{\nu(G_0(\sigma))}{\nu(G(\sigma))}, \\ d(\sigma 0) &= r_\sigma d(\sigma) \frac{\nu(\sigma)}{\nu(\sigma 0)}, \text{ and} \\ d(\sigma 1) &= (1 - r_\sigma) d(\sigma) \frac{\nu(\sigma)}{\nu(\sigma 1)}. \end{aligned}$$

It is clear that $d$ is a $\Delta$-computable $\nu$-martingale. We need to show that $d$ succeeds on $A$.

Note first that for every initial segment $\sigma b \sqsubseteq A$, there is an initial segment $\tau \sqsubseteq B$ in $G_b(\sigma)$, and hence in $G(\sigma)$. In particular, $G(\sigma)$ is nonempty.

**Claim A.8** *For each initial segment $\sigma = A[0..n-1]$,*

$$d(\sigma) \geq \frac{\nu(G(\sigma))}{\nu(\sigma)} \prod_{i=0}^{n-1} (1 - \epsilon_i). \tag{11}$$

*Proof of Claim A.8.* We induct on $|\sigma|$. Clearly (11) holds for $d(\lambda)$; assume (11) for $\sigma = A[0..n-1]$, $n > 0$. Let

$$r = \frac{\nu(G_0(\sigma))}{\nu(G(\sigma))}.$$

We know by property (i) of Definition 6.1 that $\nu(G(\sigma b)) \leq \nu(G_b(\sigma))$. Thus

$$\begin{array}{rcl} r \cdot \nu(G(\sigma)) & = & \nu(G_0(\sigma)) \\ & \geq & \nu(G(\sigma 0)) \\ \text{and } \ (1-r)\cdot \nu(G(\sigma)) & = & \left(1 - \dfrac{\nu(G_0(\sigma))}{\nu(G(\sigma))}\right)\nu(G(\sigma)) \\ & = & \nu(G(\sigma)) - \nu(G_0(\sigma)) \\ & \geq & \nu(G_1(\sigma)) - \epsilon_n \cdot \nu(G_1(\sigma)) \text{ by condition (8)} \\ & \geq & \nu(G(\sigma 1))(1-\epsilon_n). \end{array}$$

Therefore

$$\begin{array}{rcl} d(\sigma 0) & = & r \cdot d(\sigma)\dfrac{\nu(\sigma)}{\nu(\sigma 0)} \\ & \geq & r \cdot \dfrac{\nu(G(\sigma))}{\nu(\sigma 0)} \displaystyle\prod_{i=0}^{n-1}(1-\epsilon_i) \\ & \geq & \dfrac{\nu(G(\sigma 0))}{\nu(\sigma 0)} \displaystyle\prod_{i=0}^{n}(1-\epsilon_i), \\ \text{and } \ d(\sigma 1) & = & (1-r)d(\sigma)\dfrac{\nu(\sigma)}{\nu(\sigma 1)} \\ & \geq & (1-r)\dfrac{\nu(G(\sigma))}{\nu(\sigma 1)} \displaystyle\prod_{i=0}^{n-1}(1-\epsilon_i) \\ & \geq & \dfrac{\nu(G(\sigma 1))}{\nu(\sigma 1)} \displaystyle\prod_{i=0}^{n}(1-\epsilon_i). \end{array}$$

Thus in either case the induction is complete. □

Now since $\sum \epsilon_i < \infty$, $\prod(1-\epsilon_i) > c$ for some constant $c > 0$, and hence we have, for any $\sigma \sqsubseteq A$,

$$d(\sigma) \geq \frac{\nu(G(\sigma))}{\nu(\sigma)} \cdot c.$$

Now by condition (ii) of Definition 6.1, for any constant $k$ there is an initial segment $\sigma b = A[0..i]$ and a corresponding $\tau = B[0..j]$ in $G_b(\sigma)$ with $\nu(\tau)/\nu(\sigma b) > k$. Since $\tau$ is in $G(\sigma)$ also, we have $\nu(G(\sigma)) \geq \nu(\tau)$, and since $\nu(\sigma) \leq \nu(\sigma b)/\delta$, we have

$$d(\sigma) \geq \frac{\nu(G(\sigma))}{\nu(\sigma)} \cdot c \geq \frac{\nu(\tau)\cdot\delta}{\nu(\sigma b)} \cdot c > ck\delta.$$

This completes the proof of Theorem 6.2.

# References

[1] J.M. Breutzmann and J. H. Lutz. Equivalence of measures of complexity classes. 1996. To appear.

[2] H. Buhrman and L. Longpré. Compressibility and resource bounded measure. 1995 STACS.

[3] W. Feller. *An Introduction to Probability Theory and its Applications.* Volume 2, John Wiley and Sons, Inc., 1971.

[4] P.R. Halmos. *Measure Theory.* Springer-Verlag, 1974.

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

[6] S. M. Kautz. *Degrees of Random Sets.* PhD thesis, Cornell University, 1991.

[7] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications.* Springer-Verlag, 1993.

[8] J. H. Lutz. The quantitative structure of exponential time. In *Proceedings of the Eighth Annual Structure in Complexity Theory Conference*, pages 158–175, 1993. Updated version to appear in L. A. Hemaspaandra and A. L. Selman (eds.), *Complexity Theory Retrospective II*, Springer-Verlag, 1996.

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

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

[11] E. Mayordomo. *Contributions to the Study of Resource-Bounded Measure.* PhD thesis, Universitat Politècnica de Catalunya, 1994.

[12] Noam Nisan. Extracting randomness: how and why. A survey. In *Proceedings of the 11th IEEE Conference on Computational Complexity*, 1996.

[13] M. Santha and U.V. Vazirani. Generating quasi-random sequences from slightly-random sources. In *Proc. 25th Ann. Symp. on the Theory of Computing*, 1984.

[14] A. Kh. Shen´. Algorithmic complexity and randomness: recent developments. *Theory Probab. Appl.*, 37(3):92–97, 1993.

[15] A. Kh. Shen´. On relations between different algorithmic definitions of randomness. *Soviet Math. Dokl.*, 38(2):316–319, 1989.

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

[17] M. van Lambalgen. Von Mises' definition of random sequences reconsidered. *Journal of Symbolic Logic*, 52(3):725–755, 1987.

[18] U.V. Vazirani and V.V. Vazirani. Random polynomial time is equal to slightly-random polynomial time. In *Proceedings of the 26th IEEE Symposium on Foundations of Computer Science*, 1985.

[19] 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*, 25:83–123, 1970.