Improved Resource-Bounded Borel-Cantelli and Stochasticity Theorems

(Research Note)

Kenneth W. Regan
regan@cs.buffalo.edu

D. Sivakumar *
sivak-d@cs.buffalo.edu

Abstract

This note strengthens and simplifies Lutz's resource-bounded version of the Borel-Cantelli lemma for density systems and martingales. We observe that the technique can be used to construct martingales that are "additively honest," and also martingales that are "multiplicatively honest." We use this to improve the "Weak Stochasticity Theorem" of Lutz and Mayordomo: their result does not address the issue of how rapidly the bias away from 1/2 converges toward zero in a "stochastic" language, while we show that the bias must vanish exponentially.

1. Introduction

Lutz [15] developed a resource-bounded version of the classical first Borel-Cantelli lemma. Lutz's formulation, and its subsequent use in [16, 17, 19, 21], is in terms of the "density systems" that he originally used to define his resource-bounded measure theory in [15]. The above-cited papers all note the equivalent formulation of the measure theory in terms of *martingales*, along lines pioneered for complexity theory by Schnorr [27, 28, 29], and full details of the equivalence may be found in the first chapter of the dissertation by Mayordomo [21].

This note presents a martingale version of Lutz's lemma that is considerably simpler and easier to apply than the original. It is also stronger in several respects: it establishes significant "honesty" properties of the martingales that are constructed, and it replaces Lutz's condition that certain sequences have "polynomially bounded modulus of convergence" by the simpler condition that values of the "payoff sequence" used in the lemma can be written down (or approximated) in polynomial time. Our second proof of one direction of the lemma, which avoids the need to compute partial sums of the payoff sequence in building the "multiplicatively honest" martingale, appears to be new.

In the second part of this paper, we apply our lemma to establish a stronger form of the "Weak Stochasticity Theorem" of Lutz and Mayordomo [17, 19]. Informally speaking, Lutz and Mayordomo defined a language $L \subseteq \Sigma^*$ to be "weakly stochastic" over a (uniform or nonuniform) complexity class $\mathcal{C}$ if for every $\mathcal{C}$-bounded recognizer R, the limit as $n \to \infty$ of the proportion of strings $x \in \Sigma^n$ on which R correctly decides whether $x \in L$ converges to 1/2. Their definitions, theorem, and proof do not address the question of the *speed* of the convergence to 1/2, and in particular whether the *bias* away from 1/2 is *negligible* or not. Negligible bias, meaning that for

*Address for both authors: Department of Computer Science, State University of New York at Buffalo, 226 Bell Hall, Buffalo, NY 14260-2000 USA. Both authors were supported in part by NSF Grant CCR-9409104

every polynomial $p(n)$ the bias multiplied by $p(n)$ still converges to zero, is an important concept in the study of pseudorandom generators and one-way functions (see [13, 11, 12, 24, 25]). We show, in fact, that for a measure-one class of languages $L \in \mathrm{E}$, not only is L weakly stochastic over the same nonuniform classes as in [17, 19], but also the bias vanishes as $1/2^{an}$, where a can be chosen as any constant less than $1/2$.

Section 2 defines the needed concepts from resource-bounded measure theory, and defines our notions of "additively honest" and "multiplicatively honest" for martingales. Section 3 and Section 4 present our main theorems. Section 5 discusses possible further applications of our results.

2. Definitions

The notation and conventions we use are essentially standard. All languages and functions are assumed to be defined over the finite alphabet $\Sigma = \{0, 1\}$. The empty string is denoted by λ. The relation $w \sqsubseteq x$ means that strong w is a prefix of string x. We identify a language A with its characteristic function χ_A. Using the standard lexigraphic ordering of Σ^*, we also regard χ_A as a member of the set $\{0, 1\}^\omega$ of infinite binary strings. Then $w \sqsubseteq A$ abbreviates $w \sqsubseteq \chi_A$, and we call w a *characteristic prefix* of A. For any 0-1 string w, the *cylinder* C_w is defined to be the class of languages A such that $w \sqsubseteq A$. (Note that C_w is uncountable.)

For all $n \geq 0$ we also identify $A^{=n}$ with the segment u_n of χ_A of length 2^n that represents the membership or nonmembership in A of all strings of length n, and likewise identify $A^{\leq n}$ with $u_0 u_1 \cdots u_n$. Note that each u_n belongs to the set F_n of Boolean functions on $\{0, 1\}^n$.

We use the same notation for language and function classes when the context is clear. Among complexity classes whose notations differ between sources, we write QP for $\mathrm{DTIME}[2^{\mathrm{polylog}\, n}]$ (this is often called *quasipolynomial time*), E for $\mathrm{DTIME}[2^{O(n)}]$, and EXP for the class $\mathrm{DTIME}[2^{\mathrm{poly}(n)}]$. P/poly denotes the class of languages accepted by some family of polynomial size circuits. Given a language A and bounds $t(n)$, $u(n)$, if there is a deterministic Turing machine M and a sequence $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots)$ of "advice strings," each α_n of length at most $u(n)$, such that for all n and $x \in \{0, 1\}^n$, M on input $x\#\alpha_n$ correctly decides whether $x \in L$ within $t(n)$ steps, then we say L is acceptable in *time* $t(n)$ *and advice* $u(n)$, and write $A \in \mathrm{DTIME}[t(n)]/\mathrm{ADV}[u(n)]$. Then P/poly is the same as the class of languages acceptable in polynomial time with polynomial advice. For more details about these and related complexity classes, see [8].

All logarithms in this paper are to the base 2. For readability we often write N for 2^n. Then the cardinality of $\{0, 1\}^n$ is N, and that of F_n is 2^N. In probability terms of the form $\Pr_{x \in S}[\ldots]$, it is assumed that x is selected under the uniform distribution on the set S.

2.1. Resource-bounded measure

Lutz's resource-bounded measure theory is patterned along the lines of classical measure theory (see [22, 9, 23]). Complexity classes correspond to point sets in the topological space whose basic open sets are the cylinders C_w. The general form of Lutz's theory, expounded recently by Mayordomo [21], defines conditions for a class $\mathcal{C}$ to be *measurable* by a function class Δ, and to have *measure* e, written $\mu_\Delta(\mathcal{C}) = e$, where $0 \leq e \leq 1$. The restriction to Δ removes the problem that uniform complexity classes are countable, and in classical measure theory, all countable point sets have measure zero. Since all complexity classes we discuss are closed under finite variations, and by a form of the *Kolmogorov zero-one law* proved in [21] have measure zero or one, we need only discuss conditions for classes to have measure zero.

The Δ-measures of Lutz were originally defined in terms of "density systems" [15], but subsequent papers mentioned above give equivalent formulations in terms of "martingales." Here, a *martingale* is a function d from $\{0,1\}^*$ into the nonnegative reals that satisfies the following "exact average law": for all $w \in \{0,1\}^*$,

$$d(w) = \frac{d(w0) + d(w1)}{2}. \tag{1}$$

(See Lutz [15, 16] and references therein.) Let $\mathbf{D}$ stand for the nonnegative dyadic rationals; i.e., those numbers of the form $n/2^r$ for integers $n, r \geq 0$.

Definition 1 (compare [15, 21]). Let Δ be a complexity class of functions. A class $\mathcal{C}$ of languages is *Δ-measurable and has Δ-measure zero*, written $\mu_\Delta(\mathcal{C}) = 0$, if there is a martingale $d : \{0,1\}^* \to \mathbf{D}$ computable in Δ such that $\mathcal{C} \subseteq S^\infty[d]$, where

$$S^\infty[d] = \{A : \lim_{w \sqsubseteq A} d(w) = +\infty\}. \tag{2}$$

One also says that $\mathcal{C}$ is *Δ-null* if $\mu_\Delta(\mathcal{C}) = 0$. In general, we say that *d succeeds on A* if $A \in S^\infty[d]$. Put another way, the *success class* $S^\infty[d]$ is the class of languages A that satisfy

$$(\forall K > 0)(\exists N > 0)(\forall w \sqsubseteq A)[|w| \geq N \Rightarrow d(w) \geq K]. \tag{3}$$

Intuitively, the martingale d is a "betting strategy" that starts with a capital sum $d(\lambda) > 0$ and makes infinite profit along the characteristic prefixes of every $A \in S^\infty[d]$. By (1), for all $w \in \{0,1\}^*$, $|d(w1) - d(w)| = |d(w0) - d(w)|$, and this represents the amount $b(w)$ "bet" on the membership of the string x indexed by the bit a in wa. Here $b(w) = 0$ is allowed, but usually we will have $b(w) > 0$. If $x \in A$ and $d(w1) = d(w) + b(w)$, then the wagerer successfully "bet" on x belonging to A, while if $d(w1) = d(w) - b(w)$ then the wagerer along A figuratively bet wrong on the xth round. If $x \notin A$ then the wagerer along A "predicted" x's membership correctly iff $d(w0) = d(w) + b(w)$.

Given a language class $\mathcal{D}$ whose structure we wish to investigate, there are rules for selecting an appropriate function class Δ to define a measure *on* $\mathcal{D}$. Let $\mathcal{D}$ be defined by a collection $\mathcal{F}$ of resource bounds such that

$$\text{For all } r \in \mathcal{F}, \text{ the function } r' \text{ defined by } r'(n) = (r(n))^2 \text{ is also in } \mathcal{F}. \tag{4}$$

Then $\Delta(\mathcal{D})$ is the class of functions defined by corresponding resource bounds of the form $r(\log n)$ for $r \in \mathcal{F}$. For example, we have:

$\mathcal{D} = \mathrm{E}$,	$\Delta = \mathrm{P}$,
$\mathcal{D} = \mathrm{EXP}$,	$\Delta = \mathrm{QP}$.

Lutz writes $\mu(\mathcal{C}|\mathcal{D}) = 0$ if $\mu_{\Delta(\mathcal{D})}(\mathcal{C} \cap \mathcal{D}) = 0$, saying that $\mathcal{C}$ has measure zero *in* $\mathcal{D}$. This is an intuitive and technically interesting way of saying that $\mathcal{C}$ (or $\mathcal{C} \cap \mathcal{D}$) is "small" as a subclass of $\mathcal{D}$. Likewise, $\mathcal{C}$ is said to have *measure one*, *in* $\mathcal{D}$, if $\mu(\mathcal{D} \setminus \mathcal{C} \mid \mathcal{D}) = 0$.

Mayordomo [21] gives a full demonstration that the martingale definition of "Δ-measure zero" given above is equivalent to the original one, so long as (4) holds and Δ contains P. She showed that the definition is robust under several changes, most notably: replacing the limit in (2) by a lim-sup, and replacing (1) by the "inexact average law"

$$d(w) = \frac{d(w0) + d(w1)}{2}. \tag{5}$$

One can also show that for every Δ-martingale d, there is another Δ-martingale d' such that $S^\infty[d'] = S^\infty[d]$ and d' has no zero values. (Measures defined on classes $\mathcal{D}$ smaller than E are not so robust, and we refer the reader to Allender and Strauss [1, 2] for a full treatment.[1])

2.2. Honest martingales

To connect our concept to standard notions of "honesty" in complexity theory, we redefine the above formalism so that bounds are expressed in terms of n rather than N. The new formalism is essentially the same as that for "holographic proofs" in [7, 5, 4, 30, 6], with w playing the role of the "proof." Namely, define a *query machine* M to have a standard TM input tape, any number of standard worktapes, and a *query tape* that provides "random-access" to bits of a string w given as an auxiliary input. M is given as input the length N of w in standard dyadic notation, and is allowed to write integers $i \leq N$ on its query tape, receiving in answer the bit w_i. The string N is the same as the string x_N whose membership or non-membership in languages with initial segment w is indexed by the last bit of w. Recall the discussion of "betting strategies" in Section 2.1. For all w, let n_w stand for the length of the string indexed by the bit c in wc; i.e., $n_w = \lfloor \log_2(|w|{+}1) \rfloor$. (In the following, we assume, without loss of generality by above remarks, that $d(w)$ is always nonzero.)

Definition 2. Given a martingale $d : \{0,1\}^* \to \mathbf{R}^+$, call the function $b(w) := d(w1) - d(w)$ the associated *betting strategy*, and call the function $b'(w) := b(w)/d(w)$ the associated *proportional betting strategy*.

(a) The martingale d is *additively honest* if the function $b(w)$ is computable by a query machine that on input $N = |w|{+}1$ queries only those parts of w that index strings of length n_w.

(b) The martingale d is *multiplicatively honest* if the property in (a) is true of the function $b'(w)$ instead.

Recall that we pictured a query machine computing $d(w)$ as receiving input N, but a machine M computing $b(w)$, which is trying to predict the result of the next string after those indexed by w, receives input $N{+}1$. Thus the honesty restriction is that M may only write down bit-probe addresses of the same length as $N{+}1$. Since these addresses must also be $\leq N$, it follows in particular that when $|w| = 2^n - 1$, meaning that w has just completed segment n–1 of a language L, the query machine must select the amount $b(w)$ [respectively, the proportion $b'(w)$ of whatever the current capital is] to bet on '$0^n \in L$' without querying any bits of w at all. Note also that $b(w)$ and $b'(w)$ are signed quantities; a negative bet on, say, '$0^n \in L$' is the same as a positive bet on '$0^n \notin L$.' One can relax the restriction, expressed in terms of n, to accord with familiar notions of "linearly honest" and "polynomially honest," but the strong length-n-only restriction suffices for our purposes.

[1] In a personal communication, 12/94, the second author of [2] reports that our results in Section 3 carry over to their measures on P.

3. An improved Borel-Cantelli lemma for martingales

For sake of intuition, consider a betting game played in countably many rounds $r = 1, 2, 3, \ldots$ against a Leprechaun. [2] The Leprechaun fixes in advance, for all r, the factor $k_r \geq 1$ by which he will multiply the player's bet b_r if the player is Lucky in round r. The Leprechaun promises that the player will be Lucky infinitely often, but he can postpone any individual stroke of luck as long as he likes. Specifically, in each round r the player first selects an amount b_r up to her current capital C_{r-1}. The Leprechaun looks at the player's bet and then decides whether the player wins or loses.' If he says "Lose!" then $C_r = C_{r-1} - b_r$. If he says "Lucky day!" then $C_r = C_{r-1} - b_r + k_r \cdot b_r$.

The *question* is: Given the sequence $[k_r]_{r=1}^{\infty}$, can one devise a "betting strategy" B that will assure infinite gain over the initial capital $C_0 > 0$, no matter what the Leprechaun does? In our applications:

- "Round r" will correspond to a batched series of bets on $\{0,1\}^n$ (so we will have $r = n$).
- The promise is that for each language L in the Leprechaun's class, there is a feasible TM that takes a small amount of advice, and which, for infinitely many n, predicts $L^{=n}$ reasonably well.
- The multiplier $k_r = k_n$ that comes into play when the TM wins will be estimated from below via bound on the tail of binomial distributions.

By the standard notion of *strategy* in the theory of infinite games, the bet money b_r may depend on entire history of the game in previous rounds, as well as on complete information about the sequence $[k_r]_{r=1}^{\infty}$. But with reference to Definition 2, we shall be interested in strategies in which b_r, or the ratio $b'_r = b_r / C_{r-1}$, depends only on r and k_r. For this reason, we give two proofs of the standard Borel-Cantelli lemma, worded from the point of view of martingales.

Lemma 1 (Borel-Cantelli lemma for martingales).

(a) If $\sum_{r=1}^{\infty} 1/k_r$ diverges, then for every betting strategy, the Leprechaun has a way to hold $C_r \leq C_0$, for all r.

(b) If $\sum_{r=1}^{\infty} 1/k_r$ converges, then the player has both an additively honest and a multiplicatively honest winning strategy.

Proof. (a) The Leprechaun waits for the first positive bet $b_r = \varepsilon$, and then says "Lose!" He continues saying "Lose!" until the player makes a bet b_r such that $k_r b_r < \varepsilon$. Then he says "Win!", but this still leaves the player with less than the initial capital C_0. The player *must* make such a bet b_r because if she maintains $b_r \geq \varepsilon / k_r$, she goes bankrupt since $\sum_{r=1}^{\infty} 1/k_r$ diverges. The Leprechaun repeats the same strategy, and holds the player to her initial capital.

(b) Let $K > \sum_{r=1}^{\infty} 1/k_r$. For the additively honest strategy, at each stage r bet $C_0 / K k_r$. Then it is straightforward to show that after each round r,

$$C_r = C_0 \left(1 - \frac{1}{K} \sum_{i=1}^{r} \frac{1}{k_i} \right) + N_r \frac{C_0}{K},$$

[2] In Irish folklore, a *leprechaun* is a mischievous elf who has a treasure that he promises to give you, but who tries all manner of delaying tactics to frustrate you.

where N_r is the number of times the Leprechaun has said "Win!" up to stage r. The first term is always positive, so the player never goes bankrupt, and the second term goes monotonically to ∞.

For the multiplicatively honest strategy, the player *waits* until a fixed stage r_0 such that $\sum_{r>r_0} 1/k_r < 1/3$. Such r_0 exists since the sum converges. At all stages $r \geq r_0$, take $b'(r) := 1/k_r$; that is, the amount $b(r)$ bet is C_{r-1}/k_r, whatever the current capital C_{r-1} happens to be. Now let r be the first round in which the Leprechaun says "Win!" Then we have

$$C_{r-1} = C_0 \prod_{i=r_0}^{r-1} (1 - \frac{1}{k_i}).$$

A simple induction (basically "inclusion-exclusion") shows that $\prod_{i=r_0}^{r-1}(1 - \frac{1}{k_i}) \geq 1 - \sum_{i=r_0}^{r-1} \frac{1}{k_i} \geq \frac{2}{3}$. Then we have

$$C_r = C_{r-1} - C_{r-1}/k_r + C_{r-1} \geq \frac{5}{3}C_{r-1}. \geq \frac{10}{9}C_0.$$

Thus when the Leprechaun first says "Win!", the player would have multiplied her capital by at least a factor of 10/9. Now she continues the same strategy (with no need to "wait" anymore), and after N wins she has at least $C_0(10/9)^N$. $\square$

Note that in either case, computing C_r (or a lower bound for C_r) involves computing the sum $\sum_{i=1}^{r} 1/k_r$. In Lutz's proof via density systems this corresponds to a requirement that certain sums have "polynomial modulus of convergence." However, the quantities $b(r)$ and $b'(r)$ in our respective honest strategies are free of such sums; all they need is that k_r itself can be written down in polynomial (or whatever) time. Both strategies work even if they are based on values k'_r that are less than the actual payoffs, so long as $\sum_r k'_r$ converges. Going back to the quantities $b(w)$ and $b'(w)$ in Definition 2, the martingale value $d(w)$ can be recovered by computing $b(v)$ or $b'(v)$ for all $v \sqsubseteq w$, at an extra factor of at most N in running time (even without honesty). This frees us to concentrate on the complexity of the payoff sequence $[k_r]$ itself, or on some lower bound for it, without worrying about any other quantities.

4. An improved stochasticity theorem

In this section, we show that for any fixed constant c, almost every problem in E is extremely hard for Turing Machines that run in time 2^{cn}. Our results improve the "Weak Stochasticity Theorem" of Lutz and Mayordomo [17, 19]. The following incorporates several forms of their notion of a "weakly stochastic" language that have appeared in these papers and in [16, 21].

Definition 3 (see [17, 19]). (a) A language A is *weakly stochastic for time $t(n)$ and advice $q(n)$* if for all $t(n)$-time bounded Turing machines M that take $q(n)$ advice, and all advice sequences $\{\alpha_n : n \in \mathbf{N}, \alpha_n \in \{0,1\}^{q(n)}\}$,

$$\lim_{n\to\infty} \Pr_{x\in\{0,1\}^n}[(M/\alpha_n)(x) = A(x)] = \frac{1}{2}.$$

(b) The language A is weakly stochastic for time $t(n)$ and advice $q(n)$ *over domains of size $S(n)$ that are set down in time $u(n)$ with advice $r(n)$* if for all languages $D \in$ DTIME$[u(n)]$/ADV$[r(n)]$ such that $\|D^{=n}\| \geq S(n)$ for all n, and all M, $\{\alpha_n\}$ as in (a),

$$\lim_{n\to\infty} \Pr_{x\in D^{=n}}[(M/\alpha_n)(x) = A(x)] = \frac{1}{2}.$$

Here uniform distribution on the domains $\{0,1\}^n$ or $D^{=n}$ is assumed.

Lutz and Mayordomo focus on the case where for some fixed $c, \gamma > 0$, $t(n) = u(n) = 2^{cn}$, $q(n) = r(n) = cn$, and $S(n) = 2^{\gamma n}$, calling the class of languages in E so defined WS$[2^{cn}, cn, 2^{\gamma n}]$. (In fact, [18] has $r(n) = 0$.) They prove that for any fixed c and γ, $\mu(\mathrm{WS}[2^{cn}, cn, 2^{\gamma n}] \mid \mathrm{E}) = 1$.

Our *improvement* has to do with the speed of convergence toward 1/2, which their results and proofs do not address. This opens up a connection to the important notion of *hardness* used in research on pseudorandom generators (PSRGs) and one-way functions [13, 11, 12, 24, 25]. Here we define an appropriate notion of a *language* being *hard* (as compared with a PSRG or one-way function being hard) for time/advice bounded machines to gain *bias* $\varepsilon(n)$.

Definition 4. (a) A language A is *hard for time-$t(n)$ machines with advice $q(n)$ to achieve bias $\varepsilon(n)$* if for every $t(n)$ time bounded deterministic TM M that takes length-$q(n)$ advice, advice sequence $\{\alpha_n\}$, and all sufficiently large n,

$$\left| \Pr_{x\in\{0,1\}^n}[(M/\alpha_n)(x) = A(x)] - \frac{1}{2} \right| < \varepsilon(n).$$

(b) The language is hard for time-$t(n)$ machines with advice $q(n)$ to achieve bias $\varepsilon(n)$ *over domains of size $S(n)$ that are set down in time $u(n)$ with advice $r(n)$* if for all languages $D \in$ DTIME$[u(n)]$/ADV$[r(n)]$ such that $\|D^{=n}\| \geq S(n)$ for all n, and all M, $\{\alpha_n\}$ as in (a), and sufficiently large n,

$$\left| \Pr_{x\in D^{=n}}[(M/\alpha_n)(x) = A(x)] - \frac{1}{2} \right| < \varepsilon(n).$$

For simplicity we state and prove our improvement of the theorem of Lutz and Mayordomo first for case (a) with all of $\{0,1\}^n$ as prediction domain, describing the adjustments for the general case (b) at the end.

Theorem 2. *Let c be any constant, and let $a < 1/2$. Let $\mathcal{H}$ denote the class of languages A that are hard for time 2^{cn} machines with cn advice to achieve bias $1/2^{an}$. Then $\mu(\mathcal{H}|\mathrm{E}) = 1$. Moreover, this is achieved both by additively honest and multiplicatively honest martingales.*

The significance of $a < 1/2$ is that $2^{\frac{1}{2}n} = \sqrt{N}$ is the standard deviation of N uniform Bernoulli trials from the mean ($N = 2^n$). Thus a "random" language A would expect to yield this bias to the simple machine M that always accepts (or that always rejects), for many n. Having $a < 1/2$ makes bias $1/2^{an}$ asymptotically a much larger deviation, enough for the tail $F_{N,k} := \sum_{i=k}^{N} \binom{N}{i}/2^N$ of the binomial distribution with $k = 2^{n-1} + 2^{n-an}$ to vanish rapidly. However, this is still a much smaller deviation than the constant bias used in the proof by Lutz and Mayordomo (starting from the supposition that limit of 1/2 does not exist). Our proof actually makes it clear that the bias can be pushed even closer to $2^{\frac{1}{2}n}$. In the general version for domains of size $2^{\gamma n}$, the corresponding bound is $a < \gamma/2$.

Proof. Let $\mathcal{H}'$ be the complement of $\mathcal{H}$ in E. We show that $\mu(\mathcal{H}'|\mathrm{E}) = 0$. That is, we construct a martingale d computable in time $N^{O(1)} = 2^{O(n)}$ that succeeds on $\mathcal{H}'$.

For all languages $A \in \mathcal{H}'$, the "Leprechaun's promise" is that there exists a TM M that runs in time 2^{cn} such that for infinitely many n, stage n is "lucky" in the sense that

$$(\exists \alpha \in \{0,1\}^{cn}) \Pr_{x \in \{0,1\}^n}[(M/\alpha)(x) = A(x)] \geq \frac{1}{2} + \frac{1}{2^{an}}.$$

Note that for infinitely many of these n, M is among the first n Turing machines $M_1, \ldots, M_n$. For all n, we calculate a lower bound on the multiplier k_n that applies in case stage n is "lucky." From this we can calculate the amounts b_n that should be "bet" at stages n, using either of the two strategies in Lemma 1. All other calculation needed to compute the martingale d over this stage will involve only the strings in $\{0,1\}^n$, so d will be additively honest or multiplicatively honest, accordingly.

To determine the multiplier, let b_n be given. The Player divides b_n into $n2^{cn}$ equal pieces, one for each TM $M \in \{M_1, \ldots, M_n\}$ and $\alpha \in \{0,1\}^{cn}$. So let $B := b_n/n2^{cn}$ and fix some M and α. Let u be the 2^n-length characteristic segment of M/α on $\{0,1\}^n$. Define

$$E_n := \{w \in \{0,1\}^{2^n} : u \text{ and } w \text{ agree in at least } 2^n(\tfrac{1}{2} + \tfrac{1}{2^{an}}) \text{ places}\}.$$

Then stage n is "lucky" if $A^{=n}$ is among the segments w in E_n.

Now for any subset E_n of the leaves of the full binary tree of depth 2^n, there is a martingale d_n on the tree (with initial capital B at the root), such that for all leaves w,

$$d_n(w) = \begin{cases} B \cdot \frac{1}{\Pr[E_n]} & \text{if } w \in E_n \\ 0 & \text{otherwise.} \end{cases}$$

The idea, which is basic to the equivalence of Lutz's *density systems* and martingales, is to define for each node v in the tree:

$$d_n(v) := B \cdot \frac{\Pr_{f_n \in F_n}[f_n \in E_n | v \sqsubseteq f_n]}{\Pr[E_n]}.$$

To calculate this, first calculate the number k of agreements between M/α and v on the strings indexed by v. Then $d_n(v)$ equals B times the quantity

$$\sum_{i=2^{n-1}+2^{n-an}-k}^{2^n-|v|} \binom{2^n - |v|}{i} \qquad \text{divided by} \qquad \sum_{i=2^{n-1}+2^{n-an}}^{2^n} \binom{2^n}{i}.$$

Without even trying to be clever, one can compute these sums exactly in time polynomial in N: There are fewer than N terms. Each term is a binomial coefficient whose value is at most 2^N, so can be written down with N bits, and which can be calculated in time $O(N^2)$ by crude methods. The martingale d itself is defined as the sum of the d_n over all M and α. Since there are only

$n2^{cn}$ terms to sum, it only remains to justify that b_n and hence B can be written down in time polynomial in N.

Note that no ability to compute the languages called "A" is assumed. Given a particular $A \in \mathcal{H}'$, if none of the M/α succeed in predicting A at length n, i.e. succeed in putting $A^{=n}$ into their corresponding E_n, then all of the capital b_n devoted to stage n is lost, and this is like the Leprechaun saying "Lose!" But if even one of them succeeds, the reward is

$$k_n = \frac{B}{\Pr[E_n]} = \left(\frac{1}{n2^{cn}}\right) \cdot \left(\frac{2^N}{\sum_{i=2^{n-1}+2^{n-an}}^{2^n} \binom{2^n}{i}}\right).$$

By the above remarks, this can be calculated by brute force in time polynomial in N. The required time bound for the whole martingale follows from this and Lemma 1. It remains only to show that $\sum_{n=1}^{\infty} 1/k_n$ converges.

Let D stand for the sum in the denominator, and let $k = 2^{n-1} + 2^{n-an}$. Then D equals $\sum_{i=0}^{N-k} \binom{N}{i}$. Rather than use Chernoff bounds (cf. [19, 3]), we use methods from Graham, Knuth, and Patashnik [10] and MacWilliams and Sloane [20]. Let $\ell = N - k$ and let $\delta = \ell/N$. Then $D < \binom{N}{\ell}\frac{1-\delta}{1-2\delta}$ (see "'Exercise 9.42" on page 572 of [10]). Here $\frac{1-\delta}{1-2\delta}$ equals $(N^a + 2)/4$. Then via Stirling's approximation as used in [20], we obtain

$$\binom{N}{\ell} \leq \frac{2^{NH_2(\delta)}}{\sqrt{2\pi N\delta(1-\delta)}},$$

where $H_2(\delta)$ denotes the binary entropy function $\delta \lg(1/\delta) + (1-\delta)\lg(1/(1-\delta))$. It follows by elementary calculations that for some constant $c_a > 0$ depending only on a,

$$\begin{aligned} \frac{2^N}{D} &> c_a \left(\frac{2^{\frac{1}{2}\lg N}}{2 + 2^{a \lg N}}\right) 2^{N(1-H_2(\delta))} \\ &> c_a N^{\frac{1}{2}-a} \cdot 2^{N^{1-2a}}. \end{aligned}$$

Since $2^{N^{1-2a}}/n2^{cn} > 2^{N^{1-2a}-2cn} > 2^{N^\delta} = 2^{2^{\delta n}}$ for some $\delta > 0$, it is abundantly clear that $\sum_n 1/k_n$ converges. □

Now it is easy to locate the adjustments for the more-general case, without needing to re-do any calculations.

Theorem 3. *Let c, γ be any constants. Suppose $a < \gamma/2$. Let $\mathcal{H}$ denote the class of languages A that are hard for time 2^{cn} machines with cn advice to achieve bias $1/2^{an}$, over domains of size $2^{\gamma n}$ determined by machines that run in time 2^{cn} with cn advice. Then $\mu(\mathcal{H}|\mathrm{E}) = 1$, likewise achievable by additively and multiplicatively honest martingales.*

Proof. Now the capital b_n for each stage is divided into $n^2 \cdot 2^{2cn}$ pieces, one for each pair of TMs $M_i, M_j \in \{M_1, \ldots, M_n\}$ and pair of advice strings α, β they take. In calculating $d_n(v)$, one lets k be the number of agreements between M_i/α and v on strings accepted by M_j/β. If at a leaf w it is found that M_j/β didn't accept at least $2^{\gamma n}$ strings as promised, the capital $d_n(w)$ for those two machines may be considered to be zero for the analysis, though for the computation of the overall martingale it doesn't matter what it is. The estimate of the multiplier in case even one pair of machines and advices "gets lucky" works in much the same way, with $N = 2^{\gamma n}$ in the inequalities. $\square$

Almost identical proofs work for the case of hard languages in EXP:

Theorem 4. *Let q be any polynomial, and let γ be any constant. Suppose $a < \gamma/2$. Let $\mathcal{H}$ denote the class of languages A that are hard for time $2^{q(n)}$ machines with $q(n)$ advice to achieve bias $1/2^{an}$, over domains of size $2^{\gamma n}$ determined by machines that run in time $2^{q(n)}$ with $q(n)$ advice. Then $\mu(\mathcal{H}|\mathrm{EXP}) = 1$, likewise achievable by additively and multiplicatively honest martingales.*

5. Concluding Remarks

One virtue of our results, in comparison with the proof in [15] of what has been called the "Borel-Cantelli-Lutz Lemma," and with the applications in [17, 19], is that they explicitly show the martingales that are being constructed. This makes it possible to analyze the resulting martingales for other properties, for instance honesty. Moreover, the simple but tight conditions of convergence we have presented might prove fruitful in strengthening known results (see, for instance, [2]) and/or obtaining new results. It will be a worthwhile task to re-work other theorems of Lutz et al. to show the construction of martingales. One example is the theorem that a countable "Δ-union" of Δ-null classes is Δ-null. It stands to reason that the *unions* of the classes in Theorem 2 over all $a < 1/2$, or those in Theorems 3 and 4 over all $a < \gamma/2$, should also be null.

This approach is especially important with the recent interest in measure on classes smaller than E, and the restrictions on martingales used to define them in [21, 1, 2]. Note that our proof of Theorem 2 does nothing clever at all in computing the martingale, and there appears to be useful "slack" for sharper results.

The concepts of additive and multiplicative honesty are also interesting in themselves. Which constructions in the measure theory preserve these restrictions? Examples of constructions that build martingales that seem to be inherently dishonest are the "incompressibility theorem" of Juedes and Lutz [14], and other theorems that build martingales that "look back" in the input for specific properties. The notion of multiplicative honesty is important in the connection between martingales and the "natural proofs" of Razborov and Rudich [24, 25], which we have demonstrated in [26].

References

[1] E. Allender and M. Strauss. Measure on small complexity classes, with applications for BPP. In *Proc. 35th FOCS*, 1994.

[2] E. Allender and M. Strauss. Measure on P: Robustness of the notion. Manuscript, 1995.

[3] N. Alon and J. Spencer. *The Probabilistic Method.* Wiley, 1992. With an appendix by P. Erdős.

[4] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and hardness of approximation problems. In *Proc. 33rd FOCS*, pages 14–23, 1992.

[5] S. Arora and S. Safra. Probabilistic checking of proofs. In *Proc. 33rd FOCS*, pages 2–13, 1992.

[6] L. Babai. Transparent (holographic) proofs. In *Proc. 10th STACS*, volume 665 of *Lect. Notes in Comp. Sci.*, pages 525–534. Springer Verlag, 1993.

[7] L. Babai, L. Fortnow, L. Levin, and M. Szegedy. Checking computations in polylogarithmic time. In *Proc. 23rd STOC*, pages 21–31, 1991.

[8] J. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity Theory.* Springer Verlag, 1988.

[9] J. Doob. *Measure Theory.* Springer Verlag, New York, 1991.

[10] R. Graham, D. Knuth, and O. Patashnik. *Concrete Mathematics.* Addison-Wesley, Reading, Massachusetts, 1989.

[11] J. Hastad. Pseudorandom generation under uniform assumptions. In *Proc. 22nd STOC*, pages 395–404, 1990.

[12] J. Hastad, R. Impagliazzo, L. Levin, and M. Luby. Construction of a pseudo-random generator from any one-way function. Technical Report 91–68, International Computer Science Institute, Berkeley, 1991.

[13] R. Impagliazzo, L. Levin, and M. Luby. Pseudo-random generation from one-way functions (extended abstract. In *Proc. 21st STOC*, pages 12–24, 1989.

[14] D. Juedes and J. Lutz. The complexity and distribution of hard problems. In *Proc. 34th FOCS*, pages 177–185, 1993. SIAM J. Comput., to appear.

[15] J. Lutz. Almost everywhere high nonuniform complexity. *J. Comp. Sys. Sci.*, 44:220–258, 1992.

[16] J. Lutz. The quantitative structure of exponential time. In *Proc. 8th Structures*, pages 158–175, 1993.

[17] J. Lutz and E. Mayordomo. Measure, stochasticity, and the density of hard languages. In *Proc. 10th STACS*, volume 665 of *Lect. Notes in Comp. Sci.*, pages 38–47. Springer Verlag, 1993.

[18] J. Lutz and E. Mayordomo. Cook versus Karp-Levin: Separating completeness notions if NP is not small. In *Proc. 11th STACS*, volume 775 of *Lect. Notes in Comp. Sci.*, pages 415–426. Springer Verlag, 1994.

[19] J. Lutz and E. Mayordomo. Measure, stochasticity, and the density of hard languages. *SIAM J. Comput.*, 23:762–779, 1994.

[20] F. MacWilliams and N. Sloane. *The Theory of Error-Correcting Codes.* North-Holland, Amsterdam, 1977.

[21] E. Mayordomo. *Contributions to the Study of Resource-Bounded Measure.* PhD thesis, Universidad Polytécnica de Catalunya, Barcelona, April 1994.

[22] J. Oxtoby. *Measure and Category.* Springer Verlag, New York, 2nd edition, 1980.

[23] K.R. Parthasarathy. *Introduction to Probability and Measure.* The Macmillan Company of India, Ltd., Madras, 1977.

[24] A. Razborov and S. Rudich. Natural proofs. In *Proc. 26th STOC*, pages 204–213, 1994.

[25] A. Razborov and S. Rudich. Natural proofs, 1994. Update of STOC paper, November 1994.

[26] K. Regan, D. Sivakumar and J. Cai. Pseudorandom generators, measure theory, and natural proofs. Technical Report UB-CS-TR 95-2, Computer Science Dept., University at Buffalo, January 1995.

[27] C.P. Schnorr. A unified approach to the definition of random sequences. *Math. Sys. Thy.*, 5:246–258, 1971.

[28] C.P. Schnorr. *Zufälligkeit und Wahrscheinlichkeit*, volume 218 of *Lect. Notes in Math.* Springer Verlag, 1971.

[29] C.P. Schnorr. Process complexity and effective random tests. *J. Comp. Sys. Sci.*, 7:376–388, 1973.

[30] M. Sudan. *Efficient checking of polynomials and proofs and the hardness of approximation problems.* PhD thesis, University of California, Berkeley, 1992.