Measure, stochasticity and the density of hard languages

Jack H. Lutz
Elvira Mayordomo

Report LSI-92-9-R

Measure, Stochasticity, and the Density of Hard Languages

Jack H. Lutz[1]
Department of Computer Science
Iowa State University
Ames, Iowa 50011
U.S.A.

Elvira Mayordomo[2]
Dept. Llenguatges i Sistemes Informàtics
Universitat Politècnica de Catalunya
Pau Gargallo 5
08028 Barcelona, Spain

Abstract

Ogiwara and Watanabe have recently shown that the hypothesis $P \neq NP$ implies that no (polynomially) sparse language is $\leq^P_{btt}$-hard for NP. Their technique does not appear to allow significant relaxation of either the query bound or the sparseness criterion. It is shown here that a stronger hypothesis — namely, that NP does not have measure 0 in exponential time — implies the stronger conclusion that, for every real $\alpha < 1$, every $\leq^P_{n^\alpha-tt}$-hard language for NP is (exponentially) dense.

The proof of this fact also yields two absolute results (not involving unproven hypotheses) concerning the structure of exponential time: First, almost every language decidable in exponential time has a *stochasticity* property, ensuring that it is statistically unpredictable by feasible deterministic algorithms, even with linear nonuniform advice. Second, for $\alpha < 1$, only a measure 0 subset of the languages decidable in exponential time are $\leq^P_{n^\alpha-tt}$-reducible to languages that are not exponentially dense.

[1]This author's research was supported in part by National Science Foundation Grant CCR-9157382, in part by Rockwell International, and in part by DIMACS, where he was a visitor during the first phase of this work.

[2]This author's research, performed while visiting Iowa State University, was supported by a Spanish Government grant, FPI PN90.

1 Introduction

How dense must a language $A \subseteq \{0,1\}^*$ be in order to be hard for a complexity class $\mathcal{C}$? The ongoing investigation of this question, especially important when $\mathcal{C} = \mathrm{NP}$, has yielded several significant results over the past 15 years.

Any formalization of this question must specify the class $\mathcal{C}$ and give precise meanings to "hard" and "how dense." The results of this paper concern the classes $\mathrm{E} = \mathrm{DTIME}(2^{\mathrm{linear}})$, $\mathrm{E}_2 = \mathrm{DTIME}(2^{\mathrm{polynomial}})$, and all subclasses $\mathcal{C}$ of these classes, though we are particularly interested in the case $\mathcal{C} = \mathrm{NP}$.

We say that A is $\leq_{\mathrm{r}}^{\mathrm{P}}$-*hard* for a class $\mathcal{C}$ of languages if $\mathcal{C} \subseteq \mathrm{P}_{\mathrm{r}}(A)$, where $\mathrm{P}_{\mathrm{r}}(A) = \{B \subseteq \{0,1\}^* | B \leq_{\mathrm{r}}^{\mathrm{P}} A\}$. Here the polynomial time-bounded reducibility $\leq_{\mathrm{r}}^{\mathrm{P}}$ may be $\leq_{\mathrm{m}}^{\mathrm{P}}$(*many-one* reducibility), $\leq_{\mathrm{T}}^{\mathrm{P}}$(*Turing* reducibility), $\leq_{\mathrm{btt}}^{\mathrm{P}}$(*bounded truth-table* reducibility), or $\leq_{q\text{-tt}}^{\mathrm{P}}$(truth-table reducibility with $q(n)$ queries on inputs of length n, where $q : \mathbf{N} \to \mathbf{Z}^+$).

Two criteria for "how dense" a language A is have been widely used. A language A is (*polynomially*) *sparse*, and we write $A \in \mathrm{SPARSE}$, if there is a polynomial p such that $|A_{\leq n}| \leq p(n)$ for all $n \in \mathbf{N}$, where $A_{\leq n} = A \cap \{0,1\}^{\leq n}$. A language A is (*exponentially*) *dense*, and we write $A \in \mathrm{DENSE}$, if there is a real number $\epsilon > 0$ such that $|A_{\leq n}| \geq 2^{n^\epsilon}$ for all sufficiently large $n \in \mathbf{N}$. It is clear that no sparse language is dense.

For any of the above choices of the reducibility $\leq_{\mathrm{r}}^{\mathrm{P}}$, *all known* $\leq_{\mathrm{r}}^{\mathrm{P}}$-*hard languages for NP are dense.* Efforts to explain this observation (and similar observations for other classes and reducibilities) have yielded many results. We mention four that are particularly relevant to the work presented here:

1. Meyer [16] proved that every $\leq_{\mathrm{m}}^{\mathrm{P}}$-hard language for E (or any larger class) is dense. That is,

   $$\mathrm{E} \not\subseteq \mathrm{P}_{\mathrm{m}}(\mathrm{DENSE}^{\mathrm{c}}),$$

   where $\mathrm{DENSE}^{\mathrm{c}}$ is the complement of DENSE and we write $\mathrm{P}_{\mathrm{r}}(\mathcal{S}) = \bigcup_{A \in \mathcal{S}} \mathrm{P}_{\mathrm{r}}(A)$.

2. Watanabe [18, 19] extended Meyer's result by proving that every $\leq_{\mathrm{btt}}^{\mathrm{P}}$-hard language for E is dense. That is,

   $$\mathrm{E} \not\subseteq \mathrm{P}_{\mathrm{btt}}(\mathrm{DENSE}^{\mathrm{c}}).$$

In fact, Watanabe's argument also works for $\leq^{\mathrm{P}}_{O(\log n)-\mathrm{tt}}$-reducibility, i.e.,

$$\mathrm{E} \not\subseteq \mathrm{P}_{O(\log n)-\mathrm{tt}}(\mathrm{DENSE}^{c}).$$

3. Mahaney [14], proving a conjecture of Berman and Hartmanis [1], showed that, unless P = NP, no sparse language is $\leq^{\mathrm{P}}_{\mathrm{m}}$-hard for NP. That is,

$$\mathrm{P} \neq \mathrm{NP} \Rightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{\mathrm{m}}(\mathrm{SPARSE}).$$

4. Ogiwara and Watanabe [17] extended Mahaney's result by proving that, unless P = NP, no sparse language is $\leq^{\mathrm{P}}_{\mathrm{btt}}$-hard for NP. That is ,

$$\mathrm{P} \neq \mathrm{NP} \Rightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{\mathrm{btt}}(\mathrm{SPARSE}).$$

The main result of this paper, Theorem 4.2, extends results 1 and 2 above by showing that, for every real $\alpha < 1$, only a measure 0 subset of the languages in E are $\leq^{\mathrm{P}}_{n^{\alpha}-\mathrm{tt}}$-reducible to non-dense languages. "Measure 0 subset" here refers to the resource-bounded measure theory of Lutz [10, 11]. In the notation of this theory, our main result says that, for every real $\alpha < 1$,

$$\mu(\mathrm{P}_{n^{\alpha}-\mathrm{tt}}(\mathrm{DENSE}^{c})|\mathrm{E}) = 0. \tag{1.1}$$

This means that $\mathrm{P}_{n^{\alpha}-\mathrm{tt}}(\mathrm{DENSE}^{c}) \cap \mathrm{E}$ is a *negligibly small* subset of E [10, 11]. This result, which requires a completely different technique from Watanabe's result 2 above, extends result 2, both by imposing the measure 0 condition and by extending the truth table reducibility from $O(\log n)$ queries to n^{α} queries ($\alpha < 1$). We also prove that this holds for E_2, i.e., for every real $\alpha < 1$,

$$\mu(\mathrm{P}_{n^{\alpha}-\mathrm{tt}}(\mathrm{DENSE}^{c})|\mathrm{E}_2) = 0. \tag{1.2}$$

Note that there is an enormous gap between polynomial and $2^{n^{\epsilon}}$ growth rates. (Consider, for example the G_i-hierarchy of [8].) Thus, a conclusion that every $\leq^{\mathrm{P}}_{\mathrm{r}}$-hard language for $\mathcal{C}$ is dense is much stronger than a conclusion that no sparse language is $\leq^{\mathrm{P}}_{\mathrm{r}}$-hard for $\mathcal{C}$.

Much of our interest in (1.1) and (1.2) concerns the class NP and results 3 and 4 above. It is well known that $\mathrm{P} \subseteq \mathrm{NP} \subseteq \mathrm{E}_2$. In fact, E_2 is the smallest deterministic time complexity class known to contain NP. It is easy to see [10] that $\mu(\mathrm{P}|\mathrm{E}) = \mu(\mathrm{P}|\mathrm{E}_2) = 0$, i.e., P has measure 0 in E and E_2.

Ogiwara and Watanabe's proof of result 4 above does not appear to allow significant relaxation of either the query bound or the sparseness criterion. However, Lutz has proposed investigation of the (apparently) stronger hypotheses $\mu(\mathrm{NP}|\mathrm{E}) \neq 0$ and $\mu(\mathrm{NP}|\mathrm{E}_2) \neq 0$. (These expressions mean that NP does not have measure 0 in E and that NP does not have measure 0 in E_2, respectively. By the resource-bounded generalization of the Kolmogorov zero-one law [11], "$\mu(\mathrm{NP}|\mathrm{E}_2) \neq 0$" is equivalent to "$\mu(\mathrm{NP}|\mathrm{E}_2) = 1$ or NP is not measurable in E_2", and similarly for E.) It follows immediately from (1.1) and (1.2) above that, for all $\alpha < 1$,

$$\mu(\mathrm{NP}|\mathrm{E}) \neq 0 \Rightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{n^\alpha\text{-}tt}(\mathrm{DENSE}^c) \tag{1.3}$$

and

$$\mu(\mathrm{NP}|\mathrm{E}_2) \neq 0 \Rightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{n^\alpha\text{-}tt}(\mathrm{DENSE}^c) \tag{1.4}$$

That is, unless NP is negligibly small in exponential time, every $\leq^P_{n^\alpha\text{-}tt}$-hard language for NP is dense. Comparing (1.3) and (1.4) with Mahaney and Ogiwara and Watanabe's results 3 and 4 above, we have obtained a stronger conclusion from stronger hypotheses. Note that this stronger conclusion is consistent with our observations to date.

It should be noted that (1.3) and (1.4) follow immediately from (1.1) and (1.2) without using any property of NP. Thus (1.3) and (1.4) hold with NP replaced by PH, PP, PSPACE, *or any other class whatsoever.*

When proving results of the form

$$\mu(X|\mathcal{C}) = 0,$$

where $\mathcal{C}$ is a complexity class, it often simplifies matters to have available some general purpose randomness properties of languages in $\mathcal{C}$. The term "general purpose randomness property" here is heuristic, meaning a set Z of languages with the following two properties.

(i) Almost every language in $\mathcal{C}$ has the property (of membership in) Z. (This condition, written $\mu(Z|\mathcal{C}) = 1$, means that $\mu(Z^c|\mathcal{C}) = 0$.)

(ii) It is often the case that, when one wants to prove a result of the form $\mu(X|\mathcal{C}) = 0$, it is easier to prove that $X \cap Z = \emptyset$.

For example, in ESPACE=DSPACE(2^{linear}), it is known [10, 4] that almost every language has very high space-bounded Kolmogorov complexity.

A variety of sets X have been shown to have measure 0 in ESPACE, simply by proving that every element of X has low space-bounded Kolmogorov complexity [10, 4, 13, 9]. Thus high space-bounded Kolmogorov complexity is a "general purpose randomness property" of languages in ESPACE.

In §3 below, after reviewing some fundamentals of measure in complexity classes, we prove a Weak Stochasticity Theorem, stating that almost every language in E, and almost every language in E_2, is statistically unpredictable by feasible deterministic algorithms, even with linear nonuniform advice. Specifically, for every $c \in \mathrm{N}$ and every real number $\gamma > 0$, almost every language in E has the following property: For all languages $C, D \in \mathrm{DTIME}(2^{cn})$ and for all advice functions $h : \mathrm{N} \to \{0,1\}^*$ satisfying $|h(n)| \leq cn$, suppose that we try to use $B = D/h = \{x | \langle x, h(|x|)\rangle \in D\}$ to predict A on the set C. If $|C_{=n}| \geq 2^{\gamma n}$ for all sufficiently large n, then our prediction scheme will be asymptotically no better than random coin-tossing, i.e.,

$$\lim_{n\to\infty} \frac{|(A \triangle B) \cap C_{=n}|}{|C_{=n}|} = \frac{1}{2}.$$

Following the terminology of Kolmogorov [7], we call such a property a *stochasticity* property of the language A. To be precise, the above result says that almost every language $A \in \mathrm{E}$ is *weakly* $(2^{cn}, cn, 2^{\gamma n})$-*stochastic*. The adverb "weakly" here defers to a stronger stochasticity property to be proven in [12], but weak stochasticity is a powerful and convenient tool. For example, in §4 below we prove (1.1) by a combinatorial construction showing that *no* language in $\mathrm{P}_{n^{\alpha}\text{-tt}}(\mathrm{DENSE}^c)$ is weakly $(2^{3n}, 3n, 2^{\frac{n}{2}})$-stochastic. We then appeal directly to the Weak Stochasticity Theorem of §3. It appears likely that the Weak Stochasticity Theorem will be useful for a variety of such applications in the future.

2 Preliminaries

In this paper, $[\![\psi]\!]$ denotes the *Boolean value* of the condition ψ, i.e.,

$$[\![\psi]\!] \begin{cases} 1 & \text{if } \psi \\ 0 & \text{if not } \psi \end{cases}$$

All *languages* here are sets of binary strings, i.e., sets $A \subseteq \{0,1\}^*$. We

identify each language A with its *characteristic sequence* $\chi_A \in \{0,1\}^\infty$ defined by

$$\chi_A = [\![s_0 \in A]\!][\![s_1 \in A]\!][\![s_2 \in A]\!]...,$$

where $s_0 = \lambda$, $s_1 = 0$, $s_2 = 1$, $s_3 = 00, ...$ is the standard enumeration of $\{0,1\}^*$. Relying on this identification, the set $\{0,1\}^\infty$, consisting of all infinite binary sequences, will be regarded as the set of all languages.

If $w \in \{0,1\}^*$ and $x \in \{0,1\}^* \cup \{0,1\}^\infty$, we say that w is a *prefix* of x, and write $w \sqsubseteq x$, if $x = wy$ for some $y \in \{0,1\}^* \cup \{0,1\}^\infty$. The *cylinder generated by* a string $w \in \{0,1\}^*$ is

$$C_w = \{x \in \{0,1\}^\infty \mid w \sqsubseteq x\}.$$

Note that C_w is a set of languages. Note also that $C_\lambda = \{0,1\}^\infty$, where λ denotes the empty string.

As noted in §1, we work with the exponential time complexity classes $\mathrm{E} = \mathrm{DTIME}(2^{\mathrm{linear}})$ and $\mathrm{E}_2 = \mathrm{DTIME}(2^{\mathrm{polynomial}})$. It is well-known that $\mathrm{P} \subsetneq \mathrm{E} \subsetneq \mathrm{E}_2$, that $\mathrm{P} \subseteq \mathrm{NP} \subseteq \mathrm{E}_2$ and that $\mathrm{NP} \neq \mathrm{E}$.

We let $\mathbf{D} = \{m2^{-n} \mid m \in \mathbf{Z}, n \in \mathbf{N}\}$ be the set of *dyadic rationals*. We also fix a one-to-one pairing function $\langle , \rangle$ from $\{0,1\}^* \times \{0,1\}^*$ onto $\{0,1\}^*$ such that the pairing function and its associated projections, $\langle x, y\rangle \mapsto x$ and $\langle x, y\rangle \mapsto y$ are computable in polynomial time.

Several functions in this paper are of the form $d : \mathbf{N}^k \times \{0,1\}^* \to Y$, where Y is $\mathbf{D}$ or $[0, \infty)$, the set of nonnegative real numbers. Formally, in order to have uniform criteria for their computational complexities, we regard all such functions as having domain $\{0,1\}^*$, and codomain $\{0,1\}^*$ if $Y = \mathbf{D}$. For example, a function $d : \mathbf{N}^2 \times \{0,1\}^* \to \mathbf{D}$ is formally interpreted as a function $\tilde{d} : \{0,1\}^* \to \{0,1\}^*$. Under this interpretation, $d(i,j,w) = r$ means that $\tilde{d}(\langle 0^i, \langle 0^j, w\rangle\rangle) = u$, where u is a suitable binary encoding of the dyadic rational r.

For a function $d : \mathbf{N} \times X \to Y$ and $k \in \mathbf{N}$, we define the function $d_k : X \to Y$ by $d_k(x) = d(k,x) = d(\langle 0^k, x\rangle)$. We then regard d as a "uniform enumeration" of the functions $d_0, d_1, d_2, ...$ For a function $d : \mathbf{N}^n \times X \to Y$ $(n \geq 2)$, we write $d_{k,l} = (d_k)_l$, etc.

For a function $\delta : \{0,1\}^* \to \{0,1\}^*$ and $n \in \mathbf{N}$, we write δ^n for the n-fold composition of δ with itself.

Our proof of the Weak Stochasticity Theorem uses the following form of the Chernoff bound.

Lemma 2.1.[2, 3]. IF $X_1, ..., X_N$ are independent 0-1-valued random variables with the uniform distribution, $S = X_1 + ... + X_N$, and $\epsilon > 0$, then

$$Pr[|S - \frac{N}{2}| \geq \frac{\epsilon N}{2}] \leq 2e^{-\frac{\epsilon^2 N}{6}}.$$

In particular, taking $\epsilon = \frac{2}{j+1}$, where $j \in \mathbf{N}$,

$$Pr[|S - \frac{N}{2}| \geq \frac{N}{j+1}] \leq 2e^{-\frac{N}{2(j+1)^2}}.$$

Proof. See [3]. □

3 Measure and Weak Stochasticity

In this section, after reviewing some fundamentals of measure in exponential time complexity classes, we prove the Weak Stochasticity Theorem. This theorem will be useful in the proof of our main result in §4. We also expect it to be useful in future investigations of the measure structure of E and E_2.

Resource-bounded measure [10, 11] is a very general theory whose special cases include classical Lebesgue measure, the measure structure of the class REC of all recursive languages, and measure in various complexity classes. In this paper we are interested only in measure in E and E_2, so our discussion of measure is specific to thoses classes.

Throughout this section, we identify every language $A \subseteq \{0,1\}^*$ with its characteristic sequence $\chi_A \in \{0,1\}^\infty$, defined as in §2.

A *constructor* is a function $\delta : \{0,1\}^* \to \{0,1\}^*$ such that $x \sqsubsetneq \delta(x)$ for all $x \in \{0,1\}^*$. The *result* of a constructor δ (i.e., the *language constructed by* δ) is the unique language $R(\delta)$ such that $\delta^n(\lambda) \sqsubseteq R(\delta)$ for all $n \in \mathbf{N}$. Intuitively, δ constructs $R(\delta)$ by starting with λ and their iteratively generating successively longer prefixes of $R(\delta)$. Given a set Δ of functions from $\{0,1\}^*$ into $\{0,1\}^*$, we write $R(\Delta)$ for the set of all languages $R(\delta)$ such that $\delta \in \Delta$ and δ is a constructor.

We first note that the exponential time complexity classes E and E_2 can be characterized in terms of constructors.

Notation. The classes $\mathrm{p}_1 = \mathrm{p}$ and p_2, both consisting of functions $f : \{0,1\}^* \to \{0,1\}^*$, are defined as follows.

$$\begin{aligned} \mathrm{p}_1 &= \mathrm{p} = \{f | f \text{ is computable is polynomial time}\} \\ \mathrm{p}_2 &= \{f | f \text{ is computable is } n^{(\log n)^{O(1)}} \text{ time}\} \end{aligned}$$

Lemma 3.1.[8]

1. $\mathrm{R}(\mathrm{p}) = \mathrm{E}$.
2. $\mathrm{R}(\mathrm{p}_2) = \mathrm{E}_2$.

Using Lemma 3.1, the measure structures of E and E_2 are now developed in terms of the classes p_i, for $i = 1, 2$.

Definition A *density function* is a function $d : \{0,1\}^* \to [0, \infty)$ satisfying

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

for all $w \in \{0,1\}^*$. The *global value* of a density function d is $d(\lambda)$. The *set covered by* a density function d is

$$S[d] = \bigcup_{\substack{w \in \{0,1\}^* \\ d(w) \geq 1}} C_w. \tag{2}$$

(Recall that $C_w = \{x \in \{0,1\}^\infty \mid w \sqsubseteq x\}$ is the cylinder generated by w.) A density function d *covers* a set $X \subseteq \{0,1\}^\infty$ if $X \subseteq S[d]$.

For all density functions in this paper, equality actually holds in (3.1) above, but this is not required.

Consider the random experiment in which a sequence $x \in \{0,1\}^\infty$ is chosen by using an independent toss of a fair coin to decide each bit of x. Taken together, (3.1) and (3.2) imply that $\Pr[x \in S[d]] \leq d(\lambda)$ in this experiment. Intuitively, we regard a density function d as a "detailed verification" that $\Pr[x \in X] \leq d(\lambda)$ for all sets $X \subseteq S[d]$.

More generally, we will be interested in "uniform systems" of density functions that are computable within some resource bound.

Definition An n-dimensional *density system* (n-*DS*) is a function

$$d : \mathbf{N}^n \times \{0,1\}^* \to [0,\infty)$$

such that $d_{\vec{k}}$ is a density function for every $\vec{k} \in \mathbf{N}^n$. It is sometimes convenient to regard a density function as a 0-DS.

Definition A *computation* of an n-DS d is a function $\hat{d} : \mathbf{N}^{n+1} \times \{0,1\}^* \to \mathbf{D}$ such that

$$| \hat{d}_{\vec{k},r}(w) - d_{\vec{k}}(w) | \leq 2^{-r}$$

for all $\vec{k} \in \mathbf{N}^n$, $r \in \mathbf{N}$, and $w \in \{0,1\}^*$. For $i = 1,2$, a p_i-*computation* of an n-DS d is a computation $\hat{d}$ of d such that $\hat{d} \in \mathrm{p}_i$. An n-DS d is p_i-*computable* if there exists a p_i-computation $\hat{d}$ of d.

If d is an n-DS such that $d : \mathbf{N}^n \times \{0,1\}^* \to \mathbf{D}$ and $d \in \mathrm{p}_i$, then d is trivially p_i-computable. This fortunate circumstance, in which there is no need to compute approximations, occurs frequently in practice. In any case, we will sometimes abuse notation by writing d for $\hat{d}$, relying on context and subscripts to distinguish an n-DS d from a computation d of d.

We now come to the key idea of resource-bounded measure theory.

Definition A *null cover* of a set $X \subseteq \{0,1\}^\infty$ is a 1-DS d such that, for all $k \in \mathbf{N}$, d_k covers X with global value $d_k(\lambda) \leq 2^{-k}$. For $i = 1,2$, a p_i-*null cover* of X is a null cover of X that is p_i-computable.

In other words, a null cover of X is a uniform system of density functions that cover X with rapidly vanishing global value. It is easy to show that a set $X \subseteq \{0,1\}^\infty$ has classical Lebesgue measure 0 (i.e., probability 0 in the above coin-tossing experiment) if and only if there exists a null cover of X.

Definition A set X has p_i-*measure 0*, and we write $\mu_{\mathrm{p}_i}(X) = 0$, if there exists a p_i-null cover of X. A set X has p_i-*measure 1*, and we write $\mu_{\mathrm{p}_i}(X) = 1$, if $\mu_{\mathrm{p}_i}(X^c) = 0$.

Thus a set X has p_i-measure 0 if p_i provides sufficient computational resources to compute uniformly good approximations to a system of density functions that cover X with rapidly vanishing global value.

We now turn to the internal measure structures of $E = R(p_1)$ and $E_2 = R(p_2)$.

Definition A set X has *measure 0 in* $R(p_i)$, and we write $\mu(X \mid R(p_i)) = 0$, if $\mu_{p_i}(X \cap R(p_i)) = 0$. A set X has *measure 1 in* $R(p_i)$, and we write $\mu(X \mid R(p_i)) = 1$, if $\mu(X^c \mid R(p_i)) = 0$. If $\mu(X \mid R(p_i)) = 1$, we say that *almost every* language in $R(p_i)$ is in X.

The following lemma is obvious but useful.

Lemma 3.2. For every set $X \subseteq \{0,1\}^\infty$,

$$\begin{array}{ccccc} \mu_p(X) = 0 & \Longrightarrow & \mu_{p_2}(X) = 0 & \Longrightarrow & \Pr[x \in X] = 0 \\ \Downarrow & & \Downarrow & & \\ \mu(X|E) = 0 & & \mu(X|E_2) = 0, & & \end{array}$$

where the probability $\Pr[x \in X]$ is computed according to the random experiment in which a sequence $x \in \{0,1\}^\infty$ is chosen probabilistically, using an independent toss of a fair coin to decide each bit of x.

Thus a proof that a set X has p-measure 0 gives information about the size of X in E, in E_2, and in $\{0,1\}^\infty$.

It was noted in Lemma 3.2 that $\mu_p(X) = 0$ implies $\mu_{p_2}(X) = 0$. In fact, more is true.

Lemma 3.3. [12] Let Z be the union of all sets X such that $\mu_p(X) = 0$. Then $\mu_{p_2}(Z) = 0$.

(The proof of Lemma 3.3 makes essential use of the fact that p_2 contains a universal function for p. It is *not* the case that $\mu_p(Z) = 0$.)

It is shown in [10] that these definitions endow E and E_2 with internal measure structure. Specifically, for $i = 1, 2$, if $\mathcal{I}$ is either the collection $\mathcal{I}_{p_i}$ of all p_i-measure 0 sets or the collection $\mathcal{I}_{R(p_i)}$ of all sets of measure 0 in $R(p_i)$, then $\mathcal{I}$ is a "p_i-ideal", i.e., is closed under subsets, finite unions, and "p_i-unions" (countable unions that can be generated with the resources of p_i). More importantly, it is shown that the ideal $\mathcal{I}_{R(p_i)}$ is a *proper* ideal, i.e., that E does *not* have measure 0 in E and E_2 does *not* have measure 0 in E_2. Taken together, these facts justify the intuition that, if $\mu(X|E) = 0$, then $X \cap E$ is a *negligibly small* subset of E (and similarly for E_2).

Our proof of the Weak Stochasticity Theorem does not directly use the above definitions. Instead we use a sufficient condition, proved in [10], for a set to have measure 0. To state this condition we need a polynomial notion of convergence for infinite series. All our series here consist of nonnegative terms. A *modulus* for a series $\sum_{n=0}^{\infty} a_n$ is a function $m : \mathbf{N} \to \mathbf{N}$ such that

$$\sum_{n=m(j)}^{\infty} a_n \leq 2^{-j}$$

for all $j \in \mathbf{N}$. A series is p-*convergent* if it has a modulus that is a polynomial. A sequence

$$\sum_{k=0}^{\infty} a_{j,k} \qquad (j = 0, 1, 2, \ldots)$$

of series is *uniformly* p-*convergent* if there exists a polynomial $m : \mathbf{N}^2 \to \mathbf{N}$ such that, for each $j \in \mathbf{N}$, m_j is a modulus for the series $\sum_{k=0}^{\infty} a_{j,k}$. We will use the following sufficient condition for uniform p-convergence. (This well-known lemma is easily verified by routine calculus.)

Lemma 3.4. Let $a_{j,k} \in [0, \infty)$ for all $j, k \in \mathbf{N}$. If there exist a real $\varepsilon > 0$ and a polynomial $g : \mathbf{N} \to \mathbf{N}$ such that $a_{j,k} \leq e^{-k^{\varepsilon}}$ for all $j, k \in \mathbf{N}$ with $k \geq g(j)$, then the series

$$\sum_{k=0}^{\infty} a_{j,k} \qquad (j = 0, 1, 2, \ldots)$$

are uniformly p-convergent. □

The proof of the Weak Stochasticity Theorem is greatly simplified by using the following special case (for p) of a uniform, resource-bounded generalization of the classical first Borel-Cantelli lemma.

Lemma 3.5.[10]. If d is a p-computable 2-DS such that the series

$$\sum_{k=0}^{\infty} d_{j,k}(\lambda) \qquad (j = 0, 1, 2, \ldots)$$

are uniformly p-convergent, then

$$\mu_{\mathrm{p}}\left(\bigcup_{j=0}^{\infty} \bigcap_{t=0}^{\infty} \bigcup_{k=t}^{\infty} S[d_{j,k}]\right) = 0.$$

□

If we write $S_j = \bigcap_{t=0}^{\infty} \bigcup_{k=t}^{\infty} S[d_{j,k}]$ and $S = \bigcup_{j=0}^{\infty} S_j$, then Lemma 3.5 gives a sufficient condition for concluding that S has p-measure 0. Note that each S_j consists of those languages A that are in infinitely many of the sets $S[d_{j,k}]$.

We now formulate our notion of weak stochasticity. For this we need a few definitions. Our notion of advice classes is standard [6]. An *advice function* is a function $h : \mathrm{N} \to \{0,1\}^*$. Given a function $q : \mathrm{N} \to \mathrm{N}$, we write $\mathrm{ADV}(q)$ for the set of all advice functions h such that $|h(n)| \leq q(n)$ for all $n \in \mathrm{N}$. Given a language $A \subseteq \{0,1\}^*$ and an advice function h, we define the language A/h ("A with advice h") by

$$A/h = \{x \in \{0,1\}^* \mid \langle x, h(|x|)\rangle \in A\}.$$

Given functions $t, q : \mathrm{N} \to \mathrm{N}$, we define the *advice class*

$$\mathrm{DTIME}(t)/\mathrm{ADV}(q) = \{A/h \mid A \in \mathrm{DTIME}(t), h \in \mathrm{ADV}(q)\}.$$

Definition Let $t, q, \nu : \mathrm{N} \to \mathrm{N}$ and let $A \subseteq \{0,1\}^*$. Then A is *weakly* (t,q,ν)*-stochastic* if, for all $B \in \mathrm{DTIME}(t)/\mathrm{ADV}(q)$ and all $C \in \mathrm{DTIME}(t)$ such that $|C_{=n}| \geq \nu(n)$ for all sufficiently large n,

$$\lim_{n\to\infty} \frac{|(A \triangle B) \cap C_{=n}|}{|C_{=n}|} = \frac{1}{2}.$$

Intuitively, B and C together form a "prediction scheme" in which B tries to guess the behavior of A on the set C. A is weakly (t,q,ν)-stochastic if no such scheme is better in the limit than guessing by random tosses of a fair coin.

The following lemma captures the main technical content of this section.

Lemma 3.6. Fix $c \in \mathrm{N}$ and $0 < \gamma \in \mathrm{R}$ and let

$$WS_{c,\gamma} = \{A \subseteq \{0,1\}^* \mid A \text{ is weakly } (2^{cn}, cn, 2^{\gamma n})\text{-stochastic}\}.$$

Then $\mu_{\mathrm{p}}(WS_{c,\gamma}) = 1$.

Proof. Assume the hypothesis. Let $U \in \mathrm{DTIME}(2^{(c+1)n})$ be a language that is universal for $\mathrm{DTIME}(2^{cn}) \times \mathrm{DTIME}(2^{cn})$ in the following sense: For each $i \in \mathrm{N}$, let

$$C_i = \{x \in \{0,1\}^* \mid \langle 0^i, 0x\rangle \in U\},$$

$$D_i = \{x \in \{0,1\}^* | \langle 0^i, 1x\rangle \in U\}.$$

Then $\mathrm{DTIME}(2^{cn}) \times \mathrm{DTIME}(2^{cn}) = \{(C_i, D_i) | i \in \mathbf{N}\}$.

For all $i, j, k \in \mathbf{N}$, define the set $Y_{i,j,k}$ of languages as follows. If k is not a power of 2, then $Y_{i,j,k} = \emptyset$. Otherwise, if $k = 2^n$, where $n \in \mathbf{N}$, then

$$Y_{i,j,k} = \bigcup_{z \in \{0,1\}^{\leq cn}} Y_{i,j,k,z},$$

where each

$$Y_{i,j,k,z} = \left\{ A \subseteq \{0,1\}^* \;\middle|\; |(C_i)_{=n}| \geq 2^{\gamma n} \right.$$
$$\left. \text{and} \left| \frac{|(A \triangle (D_i/z)) \cap (C_i)_{=n}|}{|(C_i)_{=n}|} - \frac{1}{2} \right| \geq \frac{1}{j+1} \right\}.$$

It is immediate from the definition of weak stochasticity that the complement $WS^c_{c,\gamma}$ of $WS_{c,\gamma}$ satisfies

$$WS^c_{c,\gamma} \subseteq \bigcup_{i=0}^{\infty} \bigcup_{j=0}^{\infty} \bigcap_{m=0}^{\infty} \bigcup_{k=m}^{\infty} Y_{i,j,k}.$$

It follows by Lemma 3.5 that it suffices to exhibit a p-computable 3-DS d with the following two properties.

(I) The series $\sum_{k=0}^{\infty} d_{i,j,k}(\lambda)$, for $i, j \in \mathbf{N}$, are uniformly p-convergent.

(II) For all $i, j, k \in \mathbf{N}$, $Y_{i,j,k} \subseteq S[d_{i,j,k}]$.

Define the function $d : \mathbf{N}^3 \times \{0,1\}^* \to [0, \infty)$ as follows. If k is not a power of 2, then $d_{i,j,k}(w) = 0$. Otherwise, if $k = 2^n$, where $n \in \mathbf{N}$, then

$$d_{i,j,k}(w) = \sum_{z \in \{0,1\}^{\leq cn}} \Pr(Y_{i,j,k,z} | C_w),$$

where the conditional probabilities $\Pr(Y_{i,j,k,z} | C_w) = \Pr[A \in Y_{i,j,k,z} | A \in C_w]$ are computed according to the random experiment in which the language $A \subseteq \{0,1\}^*$ is chosen probabilistically, using an independent toss of a fair coin to decide membership of each string in A.

It follows immediately from the definition of conditional probability that d is a 3-DS. Since $U \in \mathrm{DTIME}(2^{(c+1)n})$ and c is fixed, we can use binomial coefficients to (exactly) compute $d_{i,j,k}(w)$ in time polynomial in $i+j+k+|w|$. Thus d is p-computable.

To see that d has property (I), note first that the Chernoff bound, Lemma 2.1, tells us that, for all $i,j,k \in \mathrm{N}$ and $z \in \{0,1\}^{cn}$ (writing $k = 2^n$ and $N = k^\gamma = 2^{\gamma n}$),

$$\Pr(Y_{i,j,k,z}) \leq 2e^{-\frac{N}{2(j+1)^2}},$$

whence

$$\begin{aligned} d_{i,j,k}(\lambda) &= \sum_{z \in \{0,1\}^{\leq cn}} \Pr(Y_{i,j,k,z}) \\ &\leq 2^{cn+1} \cdot 2e^{-\frac{N}{2(j+1)^2}} \\ &< e^{cn+2-\frac{N}{2(j+1)^2}}. \end{aligned}$$

Let $a = \left\lceil \frac{1}{\gamma} \right\rceil$, let $\delta = \frac{\gamma}{4}$, and fix $k_0 \in \mathrm{N}$ such that

$$k^{2\delta} \geq k^\delta + c \log k + 2$$

for all $k \geq k_0$. Define $g : \mathrm{N} \to \mathrm{N}$ by

$$g(j) = 4^a(j+1)^{4a} + k_0.$$

Then $g \in \mathrm{p}$ and, for all $i,j,n \in \mathrm{N}$ (writing $k = 2^n$ and $N = k^\gamma = k^{4\delta}$), we have that

$$\begin{aligned} k \geq g(j) \implies N &= k^{2\delta}k^{2\delta} \\ &\geq [4^a(j+1)^{4a}]^{2\delta}(k^\delta + c \log k + 2) \\ &\geq 2(j+1)^2(k^\delta + cn + 2) \\ \implies d_{i,j,k}(\lambda) &< e^{-k^\delta}. \end{aligned}$$

Thus $d_{i,j,k}(\lambda) < e^{-k^\delta}$ for all $i,j,k \in \mathrm{N}$ such that $k \geq g(j)$. Since $\delta > 0$, it follows by Lemma 3.4 that (I) holds.

Finally, to see that (II) holds, fix $i,j,k \in \mathrm{N}$. If k is not a power of 2, then (II) is trivially affirmed, so assume that $k = 2^n$, where $n \in \mathrm{N}$. Let $A \in Y_{i,j,k}$.

Fix $z \in \{0,1\}^{\leq cn}$ such that $A \in Y_{i,j,k,z}$ and let w be the $(2^{n+1}-1)$-bit characteristic string of $A_{\leq n}$. Then

$$d_{i,j,k}(w) \geq \Pr(Y_{i,j,k,z}|C_w) = 1,$$

so $A \in C_w \subseteq S[d_{i,j,k}]$. This completes the proof of Lemma 3.6. □

We now have the Weak Stochasticity Theorem.

Theorem 3.7.

(1) For all $c \in \mathbf{N}$ and $\gamma > 0$, almost every language $A \in \mathrm{E}$ is weakly $(2^{cn}, cn, 2^{\gamma n})$-stochastic.

(2) Almost every language $A \in \mathrm{E}_2$ is, for all $c \in \mathbf{N}$ and $\gamma > 0$, weakly $(2^{cn}, cn, 2^{\gamma n})$-stochastic.

Proof. Part (1) follows immediately from Lemma 3.6 via Lemma 3.2. Part (2) follows from Lemma 3.6 via Lemma 3.2. Part (2) follows from Lemma 3.6 via Lemmas 3.3 and 3.2. □

4 The Density of Hard Languages

In this section we prove our main result, that for every real $\alpha < 1$, the set $\mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c)$ has measure 0 in E and in E_2. Some terminology and notation will be useful.

Given a query-counting function $q : \mathbf{N} \to \mathbf{Z}^+$, a *q-query function* is a function f with domain $\{0,1\}^*$ such that, for all $x \in \{0,1\}^*$,

$$f(x) = (f_1(x), ..., f_{q(|x|)}(x)) \in (\{0,1\}^*)^{q(|x|)}.$$

Each $f_i(x)$ is called a *query* of f on input x. A *q-truth table function* is a function g with domain $\{0,1\}^*$ such that, for each $x \in \{0,1\}^*$, $g(x)$ is the encoding of a $q(|x|)$-input, 1-output Boolean circuit. We write $g(x)(w)$ for the output of this circuit on input $w \in \{0,1\}^{q(|x|)}$. A $\leq^{\mathrm{P}}_{q-tt}$*-reduction* is an ordered pair (f,g) such that f is a q-query function, g is a q-truth table function, and f and g are computable in polynomial time.

Let $A, B \subseteq \{0,1\}^*$. A $\leq^{\mathrm{P}}_{q-tt}$-*reduction of* A *to* B is a $\leq^{\mathrm{P}}_{q-tt}$-reduction (f,g) such that, for all $x \in \{0,1\}^*$,

$$[\![x \in A]\!] = g(x)([\![f_1(x) \in B]\!] ... [\![f_{q(|x|)}(x) \in B]\!]).$$

(Recall that $[\![\psi]\!]$ denotes the Boolean value of the condition ψ.) In this case we say that $A \leq^{\mathrm{P}}_{q-tt} B$ via g. We say that A is $\leq^{\mathrm{P}}_{q-tt}$-*reducible to* B, and write $A \leq^{\mathrm{P}}_{q-tt} B$, if there exists (f,g) such that $A \leq^{\mathrm{P}}_{q-tt} B$ via (f,g).

The proof of our main result makes essential use of the following construction.

Given an n^α-query function f and $n \in \mathrm{N}$, the *sequentially most frequent query selection (smfq selection)* for f on inputs of length n is the sequence

$$(S_0, Q_0, y_0), (S_1, Q_1, y_1), ..., (S_{n^\alpha}, Q_{n^\alpha}, y_{n^\alpha})$$

defined as follows. Each $S_k \subseteq \{0,1\}^n$. Each Q_k is an $|S_k| \times n^\alpha$ matrix of strings, with each string in Q_k colored either green or red. The rows of Q_k are indexed lexicographically by the elements of S_k. For $x \in S_k$, row x of Q_k is the sequence $f_1(x), ..., f_{n^\alpha}(x)$ of queries of f on input x. If Q_k contains at least one green string, then y_k is the green string occurring in the greatest number of rows of Q_k. (Ties are broken lexicographically.) If Q_k is entirely red, then $y_k = \top$ ("top," i.e., undefined). The sets S_k and the coloring are specified recursively. We set $S_0 = \{0,1\}^n$ and color all strings in Q_0 green. Assume that S_k, Q_k, and y_k have been defined, where $0 \leq k < n^\alpha$. If $y_k = \top$, then $(S_{k+1}, Q_{k+1}, y_{k+1}) = (S_k, Q_k, y_k)$. If $y_k \neq \top$, then S_{k+1} is the set of all $x \in S_k$ such that y_k appears in row x of Q_k. The strings in Q_{k+1} are then colored exactly as they were in Q_k, except that all y_k's are now colored red. This completes the definition of the smfq selection.

For $0 \leq k \leq n^\alpha$, it is clear that every row of Q_k contains at least k red strings. In particular, the matrix Q_{n^α} is entirely red.

Our main results follow from the following lemma. Recall that $WS_{c,\gamma}$ is the set of all weakly $(2^{cn}, cn, 2^{\gamma n})$-stochastic languages.

<u>Lemma 4.1.</u> For every real $\alpha < 1$, $\mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c) \cap WS_{3,\frac{1}{2}} = \emptyset$.

<u>Proof.</u> Let $\alpha < 1$ and assume that $A \leq^{\mathrm{P}}_{n^\alpha-tt} L$ via (f,g), where $L \notin \mathrm{DENSE}$. It suffices to show that $A \notin WS_{3,\frac{1}{2}}$. Fix a polynomial p such that $|f_i(x)| \leq p(|x|)$ for all $x \in \{0,1\}^*$ and $1 \leq i \leq |x|^\alpha$. Let $\epsilon = \frac{1-\alpha}{4}$ and fix $n_0 \in \mathrm{N}$ such that the following conditions hold for all $n \geq n_0$.

(i) $n \geq 2 \cdot n^{1-2\epsilon}$.

(ii) $n^{2\epsilon} - n^{\epsilon} \geq 2$.

Define languages B, C, D and an advice function $h : \mathrm{N} \to \{0,1\}^*$ as follows. For all n, $C_{=n}$, $D_{=n}$, and $h(n)$ are defined from the smfq selection for f on inputs of length n as follows: Let $k = k(n)$ be the greatest integer such that $0 \leq k \leq n^{\alpha}$ and $|S_k| \geq 2^{n-kn^{2\epsilon}}$. (Note that k exists because $|S_0| = 2^n$.) We then define

$$C_{=n} = S_k,$$
$$h(n) = [\![y_0 \in L]\!] \ldots [\![y_{k-1} \in L]\!],$$

and we let $D_{=n}$ be the set of all coded pairs $\langle x, z \rangle$ such that $x \in S_k$, $z \in \{0,1\}^k$, and $g(x)(b_1 \ldots b_{n^{\alpha}}) = 1$, where each

$$b_i = \begin{cases} z[j] & \text{if } f_i(x) = y_j,\ 0 \leq j < k, \\ 0 & \text{if } f_i(x) \notin \{y_0, \ldots, y_{k-1}\}\ . \end{cases}$$

Finally, we let $B = D/h$. Intuitively here, B tries to predict A on C. Specifically, for each n and each $x \in C_{=n} = S_k$, the bit $[\![x \in B]\!]$ is a "guessed value" of the bit $[\![x \in A]\!]$. The actual value, given by the reduction (f, g) to L, is

$$[\![x \in A]\!] = g(x)([\![w_i \in L]\!] \ldots [\![w_{n^{\alpha}} \in L]\!]),$$

where $w_1, \ldots, w_{n^{\alpha}}$ are the entries in row x of the matrix Q_k. The guessed value $[\![x \in B]\!] = g(x)(b_1 \ldots b_{n^{\alpha}})$ uses the advice function h to get the *correct* bit $b_i = [\![w_i \in L]\!]$ when the string w_i is red in Q_k, and *guesses* that $w_i \notin L$ when the string w_i is green in Q_k.

It is easy to see that $C, D \in \mathrm{DTIME}(2^{3n})$ and $B \in \mathrm{DTIME}(2^{3n})/\mathrm{ADV}(3n)$. (The bound $3n$ is generous here.) Also, by condition (i) in our choice of n_0,

$$|C_{=n}| \geq 2^{n-n^{\alpha}n^{2\epsilon}} \geq 2^{\frac{n}{2}}$$

for all n.

Let

$$K = \{n \in \mathrm{N} \,|\, n \geq n_0 \text{ and } |L_{\leq p(n)}| < 2^{n^{\epsilon}}\}.$$

Note that K is infinite because L is not dense.

We now show that B does a good job of predicting A on $C_{=n}$, for all $n \in K$. Let $n \in K$. We have two cases.

(I) If $k = k(n) = n^\alpha$, then all strings in Q_k are red, so *all* the guesses made by B are correct, so

$$|(A \triangle B) \cap C_{=n}| = 0.$$

(II) If $k = k(n) < n^\alpha$, let r be the number of rows in Q_k, i.e., $r = |S_k| = |C_{=n}|$. By our choice of k, we have

$$|S_{k+1}| \leq 2^{n-(k+1)n^{2\epsilon}} \leq 2^{-n^{2\epsilon}} r.$$

That is, no green string appears in more than $2^{-n^{2\epsilon}} r$ of the rows of Q_k. Moreover, since $|L_{\leq p(n)}| \leq 2^{n^\epsilon}$, there are at most 2^{n^ϵ} green strings w in Q_k such that $w \in L$. Thus there are at most $2^{n^\epsilon} \cdot 2^{-n^{2\epsilon}} r = 2^{n^\epsilon - n^{2\epsilon}} r$ rows of Q_k in which B makes an incorrect guess that a green string is not in L; the guesses made by B are correct in all other rows! By condition (ii) in our choice of n_0, then, B is incorrect in at most $\frac{1}{4}r$ rows of Q_k. That is,

$$|(A \triangle B) \cap C_{=n}| \leq \frac{1}{4} r.$$

In either case, (I) or (II), we have

$$|(A \triangle B) \cap C_{=n}| \leq \frac{1}{4} |C_{=n}|.$$

Since this holds for all $n \in K$, and since K is infinite,

$$\frac{|(A \triangle B) \cap C_{=n}|}{|C_{=n}|} \not\to \frac{1}{2}.$$

Thus B and C testify that A is not weakly $(2^{3n}, 3n, 2^{\frac{n}{2}})$-stochastic, i.e., that $A \notin WS_{3,\frac{1}{2}}$. □

Our main results are now clear.

Theorem 4.2. For every real $\alpha < 1$,

$$\mu(\mathrm{P}_{n^\alpha\text{-tt}}(\mathrm{DENSE}^c)|\mathrm{E}) = \mu(\mathrm{P}_{n^\alpha\text{-tt}}(\mathrm{DENSE}^c)|\mathrm{E}_2) = 0.$$

Proof. This follows immediately from Theorem 3.7 and Lemma 4.1. □

As noted in §1, Theorem 4.2 extends Watanabe's result [18, 19] that every $\leq^{\mathrm{P}}_{\mathrm{btt}}$-hard language for E is dense, both by relaxing the query bound and by imposing the measure 0 condition: If a language A is even *weakly* $\leq^{\mathrm{P}}_{n^\alpha\text{-tt}}$-*hard* for E, in the sense that $\mathrm{P}_{n^\alpha\text{-tt}}(A)$ does not have measure 0 in E, then Theorem 4.2 tells us that A must be dense.

Finally, we note the consequence for NP.

Theorem 4.3. If $\mu(\mathrm{NP}|\mathrm{E}) \neq 0$ or $\mu(\mathrm{NP}|\mathrm{E}_2) \neq 0$, then for all $\alpha < 1$, every $\leq^{\mathrm{P}}_{n^\alpha\text{-tt}}$-hard language for NP is dense, i.e., $\mathrm{NP} \not\subseteq \mathrm{P}_{n^\alpha\text{-tt}}(\mathrm{DENSE}^c)$.

Proof. If NP has a $\leq^{\mathrm{P}}_{n^\alpha\text{-tt}}$-hard language H that is not dense then Theorem 4.2 tells us that $\mu(\mathrm{NP}|\mathrm{E}) = \mu(\mathrm{P}_{n^\alpha\text{-tt}}(H)|\mathrm{E}) = 0$ and $\mu(\mathrm{NP}|\mathrm{E}_2) = \mu(\mathrm{P}_{n^\alpha\text{-tt}}(H)|\mathrm{E}_2) = 0$. □

Note that the hypothesis and conclusion of Theorem 4.3 are both stronger than their counterparts in Ogiwara and Watanabe's result that

$$\mathrm{P} \neq \mathrm{NP} \Rightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{\mathrm{btt}}(\mathrm{SPARSE}).$$

Note also that Theorem 4.3 holds with NP replaced by PH, PP, PSPACE, or any other class.

5 Conclusion

The density criterion in Theorem 4.2 cannot be improved, since for every $\epsilon > 0$ there is a language $A \in \mathrm{E}$ that is $\leq^{\mathrm{P}}_{\mathrm{m}}$-hard for E_2 and satisfies $|A_{\leq n}| < 2^{n^\epsilon}$ for all n. It is an open question whether the query bound n^α can be significantly relaxed. A construction of Wilson [20] shows that there is an oracle B such that $\mathrm{E}^B \subseteq \mathrm{P}^B_{O(n)\text{-tt}}(\mathrm{SPARSE})$, so progress in this direction will require nonrelativizable techniques.

The hypothesis that $\mu(\mathrm{NP}|\mathrm{E}_2) \neq 0$, i.e., that NP is not a negligibly small subset of E_2, has recently been shown to have a number of plausible consequences: If $\mu(\mathrm{NP}|\mathrm{E}_2) \neq 0$, then NP contains p-random languages [12]; NP contains E-bi-immune languages [15]; every $\leq^{\mathrm{P}}_{\mathrm{m}}$-hard language for NP has an exponentially dense, exponentially hard complexity core [5]; and, by Theorem 4.3 above, every $\leq^{\mathrm{P}}_{n^\alpha\text{-tt}}$-hard language for NP ($\alpha < 1$) is exponentially dense. Further investigation of the consequences and plausiblity of $\mu(bNP|\mathrm{E}_2) \neq 0$ and related strong, measure-theoretic hypotheses is clearly indicated.

References

[1] L. Berman and J. Hartmanis, On isomorphism and density of NP and other complete sets, *SIAM Journal on Computing* **6** (1977), pp. 305–322.

[2] H. Chernoff, A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations, *Ann. of Math. Stat.*, **23** (1952) pp. 493–507.

[3] T. Hagerup and C. Rüb, A guided tour of Chernoff bounds, *Information Processing Letters* **33** (1990), pp. 305–308.

[4] D. W. Juedes and J. H. Lutz, Kolmogorov complexity, complexity cores, and the distribution of hardness, in O. Watanabe (ed.), *Kolmogorov Complexity: Theory and Relations to Computational Complexity*, Springer–Verlag, to appear.

[5] D. W. Juedes and J. H. Lutz, The complexity and distribution of hard problems, in preparation.

[6] R. M. Karp and R. J. Lipton, Some connections between nonuniform and uniform complexity classes, *Proceedings of the 12th ACM Symposium on Theory of Computing* (1980), pp. 302–309. Also published as Turing machines that take advice, *L'Enseignement Mathematique* **28** (1982), pp. 191–209.

[7] A. N. Kolmogorov and V. A. Uspenskii, Algorithms and randomness, translated in *Theory of Probability and Its Applications* **32** (1987), pp. 389–412.

[8] J. H. Lutz, Category and measure in complexity classes, *SIAM Journal on Computing* **19** (1990), pp. 1100–1131.

[9] J. H. Lutz, An upward measure separation theorem, *Theoretical Computer Science*, **81** (1991), pp. 127–135.

[10] J. H. Lutz, Almost everywhere high nonuniform complexity, *Journal of Computer and System Sciences* **44** (1992), to appear.

[11] J. H. Lutz, Resource-bounded measure, in preparation.

[12] J. H. Lutz, Intrinsically pseudorandom sequences, in preparation.

[13] J. H. Lutz and W. J. Schmidt, Circuit size relative to pseudorandom oracles, *Theoretical Computer Science*, to appear.

[14] S. R. Mahaney, Sparse complete sets for NP: Solution of a conjecture of Berman and Hartmanis, *Journal of Computer and System Sciences* **25** (1982), pp. 130–143.

[15] E. Mayordomo, Almost every set in exponential time is P-bi-immune, Technical Report 91-19, Department of Computer Science, Iowa State University, 1991, submitted.

[16] A. R. Meyer, reported in [1].

[17] M. Ogiwara and O. Watanabe, On polynomial–time bounded truth–table reducibility of NP sets to sparse sets, *SIAM Journal on Computing* **20** (1991), pp. 471–483.

[18] O. Watanabe, *On the Structure of Intractable Complexity Classes*, Ph. D. thesis, Department of Computer Science, Tokyo Institute of Technology, 1987.

[19] O. Watanabe, Polynomial time reducibility to a set of small density, *Proceedings of the Second Annual Structure in Complexity Theory Conference*, IEEE Press, 1987, pp. 138–146.

[20] C. B. Wilson, Relativized circuit complexity, *Journal of Computer and System Sciences* **31** (1985), pp. 169–181.