Measure, Stochasticity, and the Density of Hard Languages
(Preliminary Version)

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

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

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

†This author's research, performed while visiting Iowa State University, was supported in part by Spanish Government grant FPI PN90 and in part by National Science Foundation Grant CCR-9157382.

Abstract

The main theorem of this paper is that, for every real number $\alpha < 1$ (e.g., $\alpha = 0.99$), only a measure 0 subset of the languages decidable in exponential time are $\leq^{\mathrm{P}}_{n^\alpha-tt}$-reducible to languages that are not exponentially dense. *Thus every $\leq^{\mathrm{P}}_{n^\alpha-tt}$-hard language for* E *is exponentially dense.* This strengthens Watanabe's 1987 result, that every $\leq^{\mathrm{P}}_{O(\log n)-tt}$-hard language for E is exponentially dense. The combinatorial technique used here, the *sequentially most frequent query selection*, also gives a new, simpler proof of Watanabe's result.

The main theorem also has implications for the structure of NP under strong hypotheses. Ogiwara and Watanabe (1991) have shown that the hypothesis P $\neq$ NP implies that every $\leq^{\mathrm{P}}_{btt}$-hard language for NP is non-sparse (i.e., not polynomially sparse). 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 $\leq^{\mathrm{P}}_{n^\alpha-tt}$-hard language for NP is exponentially dense. Evidence is presented that this stronger hypothesis is reasonable.

The proof of the main theorem uses a new, very general *weak stochasticity theorem*, ensuring that almost every language in E is statistically unpredictable by feasible deterministic algorithms, even with linear nonuniform advice.

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 [3, 11, 19, 21, 22, 29, 30] 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 will consider the polynomial-time reducibilities $\leq_m^{\mathrm{P}}$ (*many-one* reducibility), $\leq_{\mathrm{T}}^{\mathrm{P}}$ (*Turing* reducibility), $\leq_{btt}^{\mathrm{P}}$ (*bounded truth-table* reducibility), and $\leq_{q-tt}^{\mathrm{P}}$ (truth-table reducibility with $q(n)$ queries on inputs of length n, where $q : \mathbf{N} \to \mathbf{Z}^+$). If $\leq_r^{\mathrm{P}}$ is any of these reducibilities, we say that a language A is $\leq_r^{\mathrm{P}}$-*hard* for a class $\mathcal{C}$ of languages if $\mathcal{C} \subseteq \mathrm{P}_r(A)$, where $\mathrm{P}_r(A) = \{\, B \subseteq \{0,1\}^* \mid B \leq_r^{\mathrm{P}} A\}$.

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 $\leq_r^{\mathrm{P}}$, *all known* $\leq_r^{\mathrm{P}}$-*hard languages for NP are dense.* Efforts to explain this observation (and similar observations for other classes and reducibilities) have yielded many results. (See [8] for a thorough survey.) We mention four such results that are particularly relevant to the work presented here.

Let DENSE^c denote the complement of DENSE, i.e., the set of all languages A such that, for all $\epsilon > 0$, there exist infinitely many n such that $|A_{\leq n}| < 2^{n^\epsilon}$. For each reducibility $\leq_r^{\mathrm{P}}$ and set $\mathcal{S}$ of languages, we write

$$\mathrm{P}_r(\mathcal{S}) = \bigcup_{A \in \mathcal{S}} \mathrm{P}_r(A).$$

The first result on the density of hard languages was the following.

Theorem 1.1. (Meyer [21]). Every $\leq_m^{\mathrm{P}}$-hard language for E (or any larger class) is dense. That is,

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

$\square$

Theorem 1.1 was subsequently improved to truth-table reducibility with $O(\log n)$ queries:

Theorem 1.2. (Watanabe [30, 29]). Every $\leq_{O(\log n)-tt}^{\mathrm{P}}$-hard language for E is dense. That is,

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

$\square$

Regarding NP, Berman and Hartmanis [3] conjectured that no sparse language is $\leq_m^{\mathrm{P}}$-hard for NP, unless P = NP. This conjecture was subsequently proven correct:

Theorem 1.3. (Mahaney [19]). If P $\neq$ NP, then no sparse language is $\leq_m^{\mathrm{P}}$-hard for NP. That is,

$$\mathrm{P} \neq \mathrm{NP} \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P}_m(\mathrm{SPARSE}).$$

$\square$

Theorem 1.3 has recently been extended to truth-table reducibility with a bounded number of queries:

Theorem 1.4. (Ogiwara and Watanabe [22]). If P $\neq$ NP, then no sparse language is $\leq_{btt}^{\mathrm{P}}$-hard for NP. That is,

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

$\square$

The Main Theorem of this paper, Theorem 4.2, extends Theorems 1.1 and 1.2 above by showing that, for every real $\alpha < 1$ (e.g., $\alpha = 0.99$), only a measure 0 subset of the languages in E are $\leq_{n^\alpha-tt}^{\mathrm{P}}$-reducible to non-dense languages. "Measure 0 subset" here refers to the resource-bounded measure theory of Lutz [15, 16] (also explained in section 3 below). In the notation of this theory, our Main Theorem 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 [15, 16].

In particular, our Main Theorem implies that

$$\mathrm{E} \not\subseteq \mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c), \tag{1.2}$$

i.e., that every $\leq^{\mathrm{P}}_{n^\alpha-tt}$-hard language for E is dense. This strengthens Theorem 1.2 above by extending the truth table reducibility from $O(\log n)$ queries to n^α queries ($\alpha < 1$). It is also worth noting that the combinatorial technique used to prove (1.1) and (1.2)—the *sequentially most frequent query selection*—is simpler than Watanabe's direct proof of Theorem 1.2. This is not surprising, once one considers that our proof of (1.2) via (1.1) is a resource-bounded instance of the *probabilistic method* [5, 24, 25, 6, 26, 1], which exploits the fact that it is often easier to prove the *abundance* of objects of a given type than to construct a *specific* object of that type.

Our proof of (1.1) also shows that, for every real $\alpha < 1$,

$$\mu(\mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c) \mid \mathrm{E}_2) = 0. \tag{1.3}$$

Much of our interest in the Main Theorem concerns the class NP and Theorems 1.3 and 1.4 above. As already noted, for all reducibilities $\leq^{\mathrm{P}}_r$ discussed in this paper, all known $\leq^{\mathrm{P}}_r$-hard languages for NP are dense. One is thus led to ask whether there is a reasonable hypothesis θ such that we can prove results of the form

$$\theta \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P}_r(\mathrm{DENSE}^c), \tag{1.4}$$

for various choices of the reducibility $\leq^{\mathrm{P}}_r$. (Such a result is much stronger than the corresponding result

$$\theta \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P}_r(\mathrm{SPARSE}),$$

because there is an enormous gap between polynomial and 2^{n^ϵ} growth rates.)

Ogiwara and Watanabe's proof of Theorem 1.4 does not appear to allow significant relaxation of either the query bound or the sparseness criterion.

In fact, it appears to be beyond current understanding to prove results of the form (1.4) if θ is "P $\neq$ NP." Karp and Lipton [11] have proven that

$$\Sigma_2^p \neq \Pi_2^p \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P(SPARSE)}.$$

That is, the stronger hypothesis $\Sigma_2^p \neq \Pi_2^p$ gives a stronger conclusion than those of Theorems 1.3 and 1.4. However, Karp and Lipton's proof does not appear to allow relaxation of the sparseness criterion, and results of the form (1.4) do not appear to be achievable at this time if θ is taken to be "$\Sigma_2^p \neq \Pi_2^p$."

To make progress on matters of this type, Lutz has proposed investigation of the measure-theoretic hypotheses $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$ and $\mu(\mathrm{NP} \mid \mathrm{E}) \neq 0$. These expressions say that NP does not have measure 0 in E_2 ("NP is not a negligible subset of E_2") and that NP does not have measure 0 in E ("$\mathrm{NP} \cap \mathrm{E}$ is not a negligible subset of E"), respectively. We now explain the meaning of these hypotheses. Both are best understood in terms of their negations.

The condition $\mu(\mathrm{NP} \mid \mathrm{E}_2) = 0$ means that there exist a *fixed* polynomial q, a *fixed* positive quantity c_0 of capital (money), and a *fixed* betting strategy (algorithm) σ with the following properties. Given any language A, the strategy σ *bets* on the membership or nonmembership of the successive strings $\lambda, 0, 1, 00, 01, 10, \cdots$ in A. Before the betting begins, σ has capital (money) c_0. When betting on a string $w \in \{0,1\}^*$, the strategy σ is given as input the string consisting of the successive bits $[\![v \in A]\!]$ for all strings v that precede w in the standard ordering of $\{0,1\}^*$. On this input, the strategy σ computes, in $\leq 2^{q(|w|)}$ steps, a fraction $r \in [-1,1]$ of its current capital to bet that $w \in A$. If σ's capital prior to this bet is c, then σ's capital after the bet is $c(1+r)$ if $w \in A$, and $c(1-r)$ if $w \notin A$. (That is, the betting is fair.) Finally, the strategy σ is *successful*, in the sense that, for all $A \in \mathrm{NP}$, σ's capital diverges to $+\infty$ as the betting progresses through the successive strings $w \in \{0,1\}^*$.

Thus, *the condition* $\mu(\mathrm{NP} \mid \mathrm{E}_2) = 0$ *asserts the existence of a fixed* $2^{q(n)}$*-time-bounded algorithm* for betting successfully on membership of strings in all languages in NP. If $\mathrm{NP} \subseteq \mathrm{DTIME}(2^{r(n)})$ for some fixed polynomial r, it is easy to devise such a strategy, so $\mu(\mathrm{NP} \mid \mathrm{E}_2) = 0$. Conversely, if $\mu(\mathrm{NP} \mid \mathrm{E}_2) = 0$, then NP is "nearly contained in some fixed $\mathrm{DTIME}(2^{q(n)})$," in the sense that there is a fixed $2^{q(n)}$-time-bounded algorithm σ for successfully betting on all languages in NP.

There does not appear to be any *a priori* reason for believing that such a strategy σ exists, i.e., there does not appear to be any *a priori* reason for believing that $\mu(\mathrm{NP} \mid \mathrm{E}_2) = 0$. Similarly, there does not appear to be any *a priori* reason for believing that $\mu(\mathrm{NP} \mid \mathrm{E}) = 0$. The hypotheses $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$ and $\mu(\mathrm{NP} \mid \mathrm{E}) \neq 0$ are thus *reasonable relative to our current knowledge*. (The hypothesis that the polynomial-time hierarchy separates into infinitely many levels enjoys a similar status. It *may* be false, but if it *is* false, then a very remarkable algorithm exists.) In fact, Lutz has conjectured that the conditions $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$ and $\mu(\mathrm{NP} \mid \mathrm{E}) \neq 0$ may be true.

At this time, we are unable to prove or disprove the widely-believed conjectures $\mathrm{P} \neq \mathrm{NP}$, $\mathrm{NP} \neq \mathrm{E}_2$, and $\mathrm{E} \not\subseteq \mathrm{NP}$. This, together with the known implications

$$\begin{aligned} \mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0 &\implies \mathrm{P} \neq \mathrm{NP}, \\ \mu(\mathrm{NP} \mid \mathrm{E}) \neq 0 &\implies \mathrm{P} \neq \mathrm{NP}, \\ \mu(\mathrm{NP} \mid \mathrm{E}_2) = 0 &\implies \mathrm{NP} \neq \mathrm{E}_2, \\ \mu(\mathrm{NP} \mid \mathrm{E}) = 0 &\implies \mathrm{E} \not\subseteq \mathrm{NP}, \end{aligned}$$

means that we are currently unable to prove or disprove the statements $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$ and $\mu(\mathrm{NP} \mid \mathrm{E}) \neq 0$.

Thus, at present, we are interested in the conditions $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$ and $\mu(\mathrm{NP} \mid \mathrm{E}) \neq 0$, not as conjectures, but rather as *scientific hypotheses*, which may have more *explanatory power* than traditional complexity-theoretic hypotheses such as $\mathrm{P} \neq \mathrm{NP}$ or the separation of the polynomial-time hierarchy. Until such time as a mathematical proof or refutation is available, the reasonableness (or unreasonableness) of such hypotheses can be illuminated only by investigation of their *consequences*. Such investigation may indicate, for example, that the consequences of $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$ form, *en masse*, a credible state of affairs, thereby increasing the reasonableness of this hypothesis. On the other hand, such investigation may uncover implausible consequences of $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$, or even a proof that $\mu(\mathrm{NP} \mid \mathrm{E}_2) = 0$. Either outcome would contribute to our understanding of NP.

Our Main Theorem implies that, for all $\alpha < 1$,

$$\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0 \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c) \tag{1.5}$$

and

$$\mu(\mathrm{NP} \mid \mathrm{E}) \neq 0 \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c). \tag{1.6}$$

(This is Theorem 4.4 below.) That is, each of the hypotheses $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$ and $\mu(\mathrm{NP} \mid \mathrm{E}) \neq 0$ implies that every $\leq^{\mathrm{P}}_{n^\alpha-tt}$-hard language for NP is dense.

This conclusion, which is credible and consistent with all observations to date, is not known to follow from P $\neq$ NP or other traditional complexity-theoretic hypotheses.

Recent investigation has also shown that the hypotheses $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$ and $\mu(\mathrm{NP} \mid \mathrm{E}) \neq 0$ imply that NP contains P-bi-immune languages [20] and that every $\leq_m^{\mathrm{P}}$-hard language for NP has an exponentially dense, exponentially hard complexity core [9]. Taken together, such results appear to indicate that these are reasonable hypotheses which may have considerable explanatory power.

The proof of our Main Theorem is based on a very general result on the "weak stochasticity" of languages in E and E_2. This result, proven in section 3 below, is a useful tool that is of independent interest, as we now explain.

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 $\mu(Z^c|\mathcal{C}) = 0$, where Z^c is the complement of Z.)

(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 [15, 10] 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 [15, 10, 18, 14]. Thus high space-bounded Kolmogorov complexity is a "general-purpose randomness property" of languages in ESPACE.

In section 3 below, after reviewing some fundamentals of measure in complexity classes, we prove the Weak Stochasticity Theorem, stating that almost every language in E, and almost every language in E_2, is "weakly stochastic," i.e., is statistically unpredictable by feasible deterministic algorithms, even with linear nonuniform advice. (See section 3 for precise definitions.) In section 4, then, we give a simple combinatorial proof that *no* language in $\mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c)$ is weakly stochastic, thereby proving the Main Theorem. It appears that weak stochasticity is, in the above sense, a general-purpose randomness property of languages in E and E_2 that will be useful in future investigations.

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

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

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

As noted in section 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 $\mathrm{P} \subsetneqq \mathrm{E} \subsetneqq \mathrm{E}_2$, 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.[4, 7]. 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 [7]. □

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 section 4. We also expect it to be useful in future investigations of the measure structure of E and E_2.

Resource-bounded measure [15, 16] 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 these classes. The interested reader may consult section 3 of [15] for more discussion and examples.

Throughout this section, we identify every language $A \subseteq \{0,1\}^*$ with its characteristic sequence $\chi_A \in \{0,1\}^\infty$, defined as in section 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 $n \in \mathbf{N}$. (Recall that this means that each string $\delta^n(\lambda)$ is a prefix of the characteristic sequence of $R(\delta)$.) Intuitively, δ constructs $R(\delta)$ by starting with λ and then 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 $p_1 = p$ and p_2, both consisting of functions $f : \{0,1\}^* \to \{0,1\}^*$, are defined as follows.

$$\begin{aligned} p_1 &= p = \{f \mid f \text{ is computable is polynomial time}\} \\ p_2 &= \{f \mid f \text{ is computable is } n^{(\log n)^{O(1)}} \text{ time}\} \end{aligned}$$

Lemma 3.1.[13]

1. $R(p) = E$.
2. $R(p_2) = E_2$.

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

$$d(w) \geq \frac{d(w0) + d(w1)}{2} \tag{3.1}$$

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

$$S[d] = \bigcup_{\substack{w \in \{0,1\}^* \\ d(w) \geq 1}} \mathbf{C}_w. \tag{3.2}$$

(Recall that $\mathbf{C}_w = \{x \in \{0,1\}^\infty \mid w \sqsubseteq x\}$ 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, parts (3.1) and (3.2) of the above definition 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 $\widehat{d} : \mathbf{N}^{n+1} \times \{0,1\}^* \to \mathbf{D}$ such that

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

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 $\widehat{d}$ of d such that $\widehat{d} \in \mathrm{p}_i$. An n-DS d is p_i-*computable* if there exists a p_i-computation $\widehat{d}$ 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. (Such applications typically do involve approximations, but these are "hidden" by invoking fundamental theorems whose proofs involve approximations.)

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 $\mathrm{E} = R(\mathrm{p}_1)$ and $\mathrm{E}_2 = R(\mathrm{p}_2)$.

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

The following lemma is obvious but useful.

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

$$\begin{array}{ccccc} \mu_{\mathrm{p}}(X) = 0 & \Longrightarrow & \mu_{\mathrm{p}_2}(X) = 0 & \Longrightarrow & \Pr[x \in X] = 0 \\ \Downarrow & & \Downarrow & & \\ \mu(X|\mathrm{E}) = 0 & & \mu(X|\mathrm{E}_2) = 0, & & \end{array}$$

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 $\mu_{\mathrm{p}}(X) = 0$ implies $\mu_{\mathrm{p}_2}(X) = 0$. In fact, more is true.

Lemma 3.3. [17] Let Z be the union of all sets X such that $\mu_{\mathrm{p}}(X) = 0$. Then $\mu_{\mathrm{p}_2}(Z) = \mu(Z \mid \mathrm{E}_2) = 0$.

Lemma 3.3 is also called the Abundance Theorem, because it implies that almost every language $A \in \mathrm{E}_2$ is p-*random*, i.e., has the property that the singleton set $\{A\}$ does *not* have p-measure 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 $\mu_{\mathrm{p}}(Z) = 0$.

It is shown in [15] that these definitions endow E and E_2 with internal measure structure. Specifically, for $i = 1, 2$, if $\mathcal{I}$ is either the collection $\mathcal{I}_{\mathrm{p}_i}$ of all p_i-measure 0 sets or the collection $\mathcal{I}_{R(\mathrm{p}_i)}$ of all sets of measure 0 in $R(\mathrm{p}_i)$, 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, the Measure Conservation Theorem of [15] says that the ideal $\mathcal{I}_{R(\mathrm{p}_i)}$ 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 $\mu(X|\mathrm{E}) = 0$, then $X \cap \mathrm{E}$ 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 [15], 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 $\sum\limits_{n=0}^{\infty} a_n$ 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 $\sum\limits_{k=0}^{\infty} a_{j,k}$. 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.[15]. 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 $S_j = \bigcap\limits_{t=0}^{\infty} \bigcup\limits_{k=t}^{\infty} S[d_{j,k}]$ and $S = \bigcup\limits_{j=0}^{\infty} S_j$, 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 [11]. An *advice function* is a function $h : \mathbf{N} \to \{0,1\}^*$. Given a function $q : \mathbf{N} \to \mathbf{N}$, we write $\mathrm{ADV}(q)$ for the set of all advice functions h such that $|h(n)| \leq q(n)$ for all $n \in \mathbf{N}$. 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 $t, q : \mathbf{N} \to \mathbf{N}$, 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 $t, q, \nu : \mathbf{N} \to \mathbf{N}$ 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,

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

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.

Our use of the term "stochastic" follows Kolmogorov's terminology [12, 28] for properties defined in terms of limiting frequencies of failure of prediction schemes. The adverb "weakly" distinguishes our notion from a stronger stochasticity property considered in [17], but weak stochasticity is a powerful and convenient tool.

The following lemma captures the main technical content of the Weak Stochasticity Theorem.

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

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

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 $i \in \mathbf{N}$, let

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

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 $\sum\limits_{k=0}^{\infty} d_{i,j,k}(\lambda)$, 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

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

where the conditional probabilities $\Pr(Y_{i,j,k,z}|\mathbf{C}_w) = \Pr[A \in Y_{i,j,k,z}|A \in \mathbf{C}_w]$ 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 $i, j, k \in \mathbf{N}$ and $z \in \{0,1\}^{\leq cn}$ (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 $k_0 \in \mathbf{N}$ such that

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

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

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

Then g is a polynomial and, for all $i, j, n \in \mathbf{N}$ (writing $k = 2^n$ and $N = k^{\gamma} = k^{4\delta}$),

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

Thus $d_{i,j,k}(\lambda) < e^{-k^{\delta}}$ for all $i, j, k \in \mathbf{N}$ 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 $i, j, k \in \mathbf{N}$. If k is not a power of 2, then (II) is trivially affirmed, so assume that $k = 2^n$, where $n \in \mathbf{N}$. 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

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

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

We now have the main result of this section.

Theorem 3.7 (Weak Stochasticity Theorem).

(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 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 $\mathrm{P}_{n^\alpha-\mathrm{tt}}(\mathrm{DENSE}^c)$ has measure 0 in E and in E_2. We then derive some consequences of this result. 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 $n \in \mathbf{N}$, 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.

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

Proof. 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 $n_0 \in \mathbf{N}$ such that the following conditions hold for all $n \geq n_0$.

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

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

Let

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

Note that K is infinite because L is not dense.

Define languages B, C, D and an advice function $h : \mathbf{N} \to \{0,1\}^*$ as follows. For all $n < n_0$, $C_{=n} = D_{=n} = \{0,1\}^n$ and $h(n) = \lambda$. For all $n \geq n_0$, $C_{=n}$, $D_{=n}$, 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

$$\begin{aligned} &C_{=n} = S_k, \\ &h(n) = [\![y_0 \in L]\!] ... [\![y_{k-1} \in L]\!], \end{aligned}$$

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 $g(x)(b_1...b_{n^\alpha}) = 1$, where each

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

Finally, we let $B = D/h$. Intuitively here, B tries to predict A on C. Specifically, for each $n \geq n_0$ 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

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

where $w_1, ..., w_{n^\alpha}$ are the entries in row x of the matrix Q_k. The guessed value $[\![x \in B]\!] = g(x)(b_1...b_{n^\alpha})$ 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 $n \geq n_0$, whence $|C_{=n}| \geq 2^{\frac{n}{2}}$ for all $n \in \mathbf{N}$.

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

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

That is, no green string appears in more than $2^{-n^{2\epsilon}} r$ of the rows of Q_k. Moreover, since $|L_{\le p(n)}| \le 2^{n^\epsilon}$, 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,

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

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

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

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 easily derived. We start with the fact that most languages decidable in exponential time are not $\le^{\mathrm{P}}_{n^\alpha-tt}$-reducible to non-dense languages.

Theorem 4.2 (Main Theorem). For every real number $\alpha < 1$,

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

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

The Main Theorem yields the following separation result.

Theorem 4.3. For every real $\alpha < 1$,

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

That is, every $\mathrm{P}_{n^\alpha - tt}$-hard language for E is dense.

Proof. By the Measure Conservation Theorem [15], $\mu(\mathrm{E} \mid \mathrm{E}) \neq 0$, so this follows immediately from Theorem 4.2. □

Note that Theorem 4.3 strengthens Theorem 1.2 by extending the number of queries from $O(\log n)$ to n^α, where $\alpha < 1$ (e.g., $\alpha = 0.99$).

It is worthwhile to examine the roles played by various methods. Theorem 4.2, a measure-theoretic result concerning the *quantitative* structure of E and E_2, yields the *qualitative* separation result Theorem 4.3. From a technical standpoint, this proof of Theorem 4.3 has the following three components.

(i) The sequentially most frequent query selection (Lemma 4.1). This is used to prove that every language in $\mathrm{P}_{n^\alpha - tt}(\mathrm{DENSE}^c)$ is predictable, i.e., fails to be weakly stochastic (with suitable parameters).

(ii) The Weak Stochasticity Theorem (Theorem 3.7). This shows that only a measure 0 subset of the languages in E are predictable.

(iii) The Measure Conservation Theorem [15]. This shows that E is not a measure 0 subset of itself.

Of these three components, (ii) and (iii) are general theorems concerning measure in E. Only component (i) is specific to the issue of the densities of $\mathrm{P}_{n^\alpha - tt}$-hard languages. That is, *given the general principles* (ii) and (iii), the proof of Theorem 4.3 is just the sequentially most frequent query selection, i.e., the proof of Lemma 4.1. The latter proof is combinatorially much simpler than Watanabe's direct proof of Theorem 1.2. This is not surprising, once it is noted that our proof of Theorem 4.3 is an application of (a resource-bounded generalization of) the *probabilistic method* [5, 24, 25, 6, 26, 1], which exploits the fact that it is often easier to establish the *abundance* of objects of a given type than to construct a *specific* object of that type. Much of our proof of Theorem 4.3 is "hidden" in the power of this method (i.e., in the proofs of the Measure Conservation and Weak Stochasticity Theorems), freeing us to apply the sequentially most frequent query selection to the problem at hand.

An important feature of this general method is that it is *uniformly constructive* in the following sense. Taken together, the proofs of the Measure Conservation and Weak Stochasticity Theorems give a straightforward, "automatic" construction of a language $A \in \mathrm{E} \cap \mathrm{WS}_{3,\frac{1}{2}}$. By Lemma 4.1, it follows immediately that $A \in \mathrm{E} \backslash \mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c)$. Thus one can apply this complexity-theoretic version of the probabilistic method with complete assurance that the resulting existence proof will automatically translate into a construction.

The primary objective of resource-bounded measure theory is to give a detailed account of the *quantitative structure* of E, E_2, and other complexity classes. The derivation of *qualitative* separation results, such as Theorems 4.3 and 1.2, is only a by-product of this quantitative objective. (By analogy, the value of classical Lebesgue measure and probability far surpasses their role as tools for existence proofs.) In the case of E, for example, the quantitative content of Theorem 4.2 is that the set $\mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c) \cap \mathrm{E}$ is a *negligibly small* subset of E.

As noted in the introduction to this paper, we are interested in the consequences of the hypothesis that NP is *not* a negligibly small subset of exponential time. In this regard, our main theorem yields the following result.

Theorem 4.4. If $\mu(\mathrm{NP}|\mathrm{E}) \neq 0$ or $\mu(\mathrm{NP}|\mathrm{E}_2) \neq 0$, then for all $\alpha < 1$, every $\leq^{\mathrm{P}}_{n^\alpha-tt}$-hard language for NP is dense, i.e., $\mathrm{NP} \not\subseteq \mathrm{P}_{n^\alpha-\mathrm{tt}}(\mathrm{DENSE}^c)$.

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

Note that the hypothesis and conclusion of Theorem 4.4 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 our proof of Theorem 4.4 actually shows that

$$\mathrm{NP} \cap \mathrm{WS}_{3,\frac{1}{2}} \neq \emptyset \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c).$$

In fact, this implication and Theorem 4.4 both hold 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 $\leq_m^{\mathrm{P}}$-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 [31] shows that there is an oracle B such that $\mathrm{E}^B \subseteq \mathrm{P}^B_{O(n)-\mathrm{tt}}(\mathrm{SPARSE})$, so progress in this direction will require nonrelativizable techniques.

There are several open questions involving special reducibilities. We mention just one example. Very recently, Arvind, Köbler, and Mundhenk [2] have proven that

$$\mathrm{P} \neq \mathrm{NP} \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{btt}(\mathrm{P}_{ctt}(\mathrm{SPARSE})),$$

where P_{ctt} refers to polynomial-time *conjunctive* reducibility. (This strengthens Theorem 1.4.) Does the class $\mathrm{P}_{btt}(\mathrm{P}_{ctt}(\mathrm{DENSE}^c))$ have measure 0 in E?

As noted in the introduction, all known $\leq_{\mathrm{T}}^{\mathrm{P}}$-hard languages for NP are dense, i.e., our experience suggests that $\mathrm{NP} \not\subseteq \mathrm{P}(\mathrm{DENSE}^c)$. This suggests two open questions. (See Figure 1.) Karp and Lipton [11] have shown that

$$\Sigma_2^p \neq \Pi_2^p \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P}(\mathrm{SPARSE}).$$

Theorem 4.4 of the present paper shows that

$$\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0 \Longrightarrow \mathrm{NP} \not\subseteq \mathrm{P}_{n^\alpha-tt}(\mathrm{DENSE}^c)$$

for $\alpha < 1$. The first question, posed by Selman [23], is whether the strong hypothesis $\mu(\Sigma_2^p \backslash \Pi_2^p \mid \mathrm{E}_2) \neq 0$ can be used to combine these ideas to get a conclusion that $\mathrm{NP} \not\subseteq \mathrm{P}(\mathrm{DENSE}^c)$. The second, more fundamental, question is suggested by the first. A well-known downward separation principle [27] says that, if the polynomial time hierarchy separates at some level, then it separates at all lower levels. Thus, for example, $\Sigma_2^p \neq \Pi_2^p$ implies that $\mathrm{P} \neq \mathrm{NP}$. Is there a "downward measure separation principle," stating that $\mu(\Sigma_{k+1}^p \backslash \Pi_{k+1}^p \mid \mathrm{E}_2) \neq 0 \Longrightarrow \mu(\Sigma_k^p \backslash \Pi_k^p \mid \mathrm{E}_2) \neq 0$? In particular, does $\mu(\Sigma_2^p \backslash \Pi_2^p \mid \mathrm{E}_2) \neq 0$ imply that $\mu(\mathrm{NP} \mid \mathrm{E}_2) \neq 0$?

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 credible consequences: If $\mu(\mathrm{NP}|\mathrm{E}_2) \neq 0$, then NP contains p-random languages [17]; NP contains E-bi-immune languages [20]; every $\leq_m^{\mathrm{P}}$-hard language for NP has an exponentially dense, exponentially hard complexity core [9]; and now, by

Figure 1: Insert the MSDH-diagram.ps output here

Theorem 4.3 above, every $\leq^{\mathrm{P}}_{n^\alpha-tt}$-hard language for NP ($\alpha < 1$) is exponentially dense. Further investigation of the consequences and reasonableness of $\mu(\mathrm{NP}|\mathrm{E}_2) \neq 0$ and related strong, measure-theoretic hypotheses is clearly indicated.

Acknowledgements

The first author thanks Bill Gasarch and Alan Selman for helpful discussions.

References

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

[2] V. Arvind, J. Köbler, and M. Mundhenk, Bounded truth-table and conjunctive reductions to sparse and tally sets, Technical report, University of Ulm, 1992, Technical Report Ulmer Informatik-Berichte 92–01.

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

[4] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Annals of Mathematical Statistics* **23** (1952), pp. 493–509.

[5] P. Erdös, Some remarks on the theory of graphs, *Bulletin of the American Mathematical Society* **53** (1947), pp. 292–294.

[6] P. Erdös and J. Spencer, *Probabilistic Methods in Combinatorics*, Academic Press, New York, 1974.

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

[8] L. A. Hemachandra, M. Ogiwara, and O. Watanabe, How hard are sparse sets?, *Proceedings of the Seventh Annual Structure in Complexity Theory Conference*, 1992, pp. 222–238. IEEE Press.

[9] D. W. Juedes and J. H. Lutz, The complexity and distribution of hard problems, Technical Report 92-23, Department of Computer Science, Iowa State University, 1992, submitted.

[10] D. W. Juedes and J. H. Lutz, Kolmogorov complexity, complexity cores, and the distribution of hardness, In O. Watanabe, editor, *Kolmogorov Complexity: Theory and Relations to Computational Complexity.* Springer-Verlag, to appear.

[11] 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.

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

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

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

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

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

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

[18] J. H. Lutz and W. J. Schmidt, Circuit size relative to pseudorandom oracles, *Theoretical Computer Science A* **107** (1993), to appear.

[19] 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.

[20] E. Mayordomo, Almost every set in exponential time is P-bi-immune, *Seventeenth International Symposium on Mathematical Foundations of Computer Science*, 1992, pp. 392–400. Springer-Verlag.

[21] A. R. Meyer, 1977, reported in [3].

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

[23] A. L. Selman, 1992, personal communication.

[24] C. E. Shannon, A mathematical theory of communication, *Bell System Technical Journal* **27** (1948), pp. 379–423, 623–656.

[25] C. E. Shannon, The synthesis of two-terminal switching circuits, *Bell System Technical Journal* **28** (1949), pp. 59–98.

[26] J. H. Spencer, *Ten Lectures on the Probabilistic Method*, SIAM, 1987.

[27] L. J. Stockmeyer, The polynomial-time hierarchy, *Theoretical Computer Science* **3** (1977), pp. 1–22.

[28] V. A. Uspenskii, A. L. Semenov, and A. Kh. Shen', Can an individual sequence of zeros and ones be random?, *Russian Mathematical Surveys* **45** (1990), pp. 121–189.

[29] O. Watanabe, *On the Structure of Intractable Complexity Classes*, PhD thesis, Tokyo Institute of Technology, 1987.

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

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