Electronic Colloquium on Computational Complexity, Report No. 22 (2009)

Inseparability and Strong Hypotheses for Disjoint NP Pairs

Jack H. Lutz* Elvira Mayordomo†

Abstract

This paper investigates the existence of inseparable disjoint pairs of NP languages and related strong hypotheses in computational complexity. Our main theorem says that, if NP does not have measure 0 in EXP, then there exist disjoint pairs of NP languages that are P-inseparable, in fact TIME(2^{n^k})-inseparable. We also relate these conditions to strong hypotheses concerning randomness and genericity of disjoint pairs.

1 Introduction

The main objective of complexity theory is to assess the intrinsic difficulties of naturally arising computational problems. It is often the case that a problem of interest can be formulated as a decision problem, or else associated with a decision problem of the same complexity, so much of complexity theory is focused on decision problems. Nevertheless, other types of problems also require investigation.

This paper concerns *promise problems*, a natural generalization of decision problems introduced by Even, Selman, and Yacobi [6]. A decision

*Department of Computer Science, Iowa State University, Ames, IA 50011 USA. lutz@cs.iastate.edu. Research supported in part by National Science Foundation Grants 0344187, 0652569, and 0728806.

†Departamento de Informática e Ingeniería de Sistemas, Instituto de Investigación en Ingeniería de Aragón, María de Luna 1, Universidad de Zaragoza, 50018 Zaragoza, SPAIN. elvira at unizar.es. Research supported in part by Spanish Government MICINN Project TIN2008-06582-C03-02.

ISSN 1433-8092

problem can be formulated as a set $A \subseteq \{0,1\}^*$, where a solution of this problem is an algorithm, circuit, or other device that *decides* A, i.e., tells whether or not an arbitrary input $x \in \{0,1\}^*$ is an element of A. In contrast, a promise problem is formulated as an ordered pair (A, B) of disjoint sets $A, B \subseteq \{0,1\}^*$, where a solution is an algorithm or other device that decides *any* set $S \subseteq \{0,1\}^*$ such that $A \subseteq S$ and $B \cap S = \emptyset$. Such a set S is called a *separator* of the disjoint pair (A, B). Intuitively, if we are promised that every input will be an element of $A \cup B$, then a separator of (A, B) enables us to distinguish inputs in A from inputs in B. Since each decision problem A is clearly equivalent to the promise problem (A, A^c), where $A^c = \{0,1\}^* - A$ is the complement of A, promise problems are, indeed, a generalization of decision problems.

A *disjoint NP pair* is a promise problem (A, B) in which $A, B \in \mathrm{NP}$. Disjoint NP pairs were first investigated by Selman and others in connection with public key cryptosystems [6, 13, 22, 15]. They were later investigated by Razborov [21] as a setting in which to prove the independence of complexity-theoretic conjectures from theories of bounded arithmetic. In this same paper, Razborov established a fundamental connection between disjoint NP pairs and propositional proof systems. Propositional proof systems had been used by Cook and Reckhow [5] to characterize the NP versus co-NP problem. Razborov [21] showed that each propositional proof system has associated with it a canonical disjoint NP pair and that important questions about propositional proof systems are thereby closely related to natural questions about disjoint NP pairs. This connection with propositional proof systems has motivated more recent work on disjoint NP pairs by Glasser, Selman, Sengupta, and Zhang [8, 7, 10, 11]. It is now known that the degree structure of propositional proof systems under the natural notion of proof simulation is identical to the degree structure of disjoint NP pairs under reducibility of separators [10]. Much of this recent work is surveyed in [9]. Goldreich [12] gives a recent survey of promise problems in general.

Our specific interest in this paper is the existence of disjoint NP pairs that are P-inseparable, or even $\mathrm{TIME}(2^{n^k})$-inseparable. As the terminology suggests, if $\mathcal{C}$ is a class of decision problems, then a disjoint pair is *$\mathcal{C}$-inseparable* if it has no separator in $\mathcal{C}$. The existence of P-inseparable disjoint NP pairs is a strong hypothesis in the sense that (1) it clearly implies $\mathrm{P} \neq \mathrm{NP}$, and (2) the converse implication is not known (and fails relative to some oracles [15]). It is clear that $\mathrm{P} \neq \mathrm{NP} \cap \mathrm{coNP}$ implies the existence of P-inseparable disjoint NP pairs, and Grollman and Selman [13] proved that $\mathrm{P} \neq \mathrm{UP}$ also implies the existence of P-inseparable disjoint NP pairs.

The hypothesis that NP is a non-measure 0 subset of EXP, written $\mu(\mathrm{NP} \mid \mathrm{EXP}) \neq 0$, is a strong hypothesis in the above sense. This hypothesis has been shown to have many consequences not known to follow from more traditional hypotheses such as $\mathrm{P} \neq \mathrm{NP}$ or the separation of the polynomial-time hierarchy into infinitely many levels. Each of these known consequences has resolved some pre-existing complexity-theoretic question in the way that agreed with the conjecture of most experts. This explanatory power of the $\mu(\mathrm{NP} \mid \mathrm{EXP}) \neq 0$ hypothesis is discussed in the early survey papers [19, 2, 20] and is further substantiated by more recent papers listed at [14] (and too numerous to discuss here). In several instances, the discovery that $\mu(\mathrm{NP} \mid \mathrm{EXP}) \neq 0$ implies some plausible conclusion has led to subsequent work deriving the same conclusion from some weaker hypothesis, thereby further illuminating the relationships among strong hypotheses.

Our main theorem states that, if NP does not have measure zero in EXP, then, for every positive integer k, there exist disjoint NP pairs that are $\mathrm{TIME}(2^{n^k})$-inseparable. Such pairs are *a fortiori* P-inseparable, but the conclusion of our main theorem actually gives *exponential* lower bounds on the inseparability of some disjoint NP pairs. These are the lower bounds that most experts conjecture to be true, even though an unconditional proof of such bounds may be long in coming.

The proof of our main theorem combines known closure properties of NP with the randomness that the $\mu(\mathrm{NP} \mid \mathrm{EXP}) \neq 0$ hypothesis implies must be present in NP to give an explicit construction of a disjoint NP pair that is $\mathrm{TIME}(2^{n^k})$-inseparable. (Technically, this is an overstatement. The last step of the "construction" is the removal of a finite set whose existence we prove, but which we do not construct.) The details are perhaps involved, but we preface the proof with an intuitive motivation for the approach.

We also investigate the relationships between the two strong hypotheses in our main theorem (i.e., its hypothesis and its conclusion) and strong hypotheses involving the existence of disjoint NP pairs with randomness and genericity properties. Roughly speaking (i.e., omitting quantitative parameters), we show that the existence of disjoint NP pairs that are random implies both the $\mu(\mathrm{NP} \mid \mathrm{EXP}) \neq 0$ hypothesis and the existence of disjoint NP pairs that are generic in the sense of Ambos-Spies, Fleischhack, and Huwig [1]. We also show that the existence of such generic pairs implies the existence of disjoint NP pairs that are $\mathrm{TIME}(2^{n^k})$-inseparable.

2 Preliminaries

We write $\mathbb{N}$ for the set of nonnegative integers and $\mathbb{Z}^+$ for the set of (strictly) positive integers. The *Boolean value* of an assertion ϕ is $[\![\phi]\!] =$ *if* ϕ *then 1 else 0*. All logarithms here are base-2.

We write λ for the empty string, $|w|$ for the length of a string w, and $s_0, s_1, s_2, \ldots$ for the standard enumeration of $\{0,1\}^*$. The index of a string x is the value $\mathrm{ind}(x) \in \mathbb{N}$ such that $s_{\mathrm{ind}(x)} = x$. We write $\mathrm{next}(x)$ for the string following x in the standard enumeration, i.e., $\mathrm{next}(s_n) = s_{n+1}$. More generally, for $k \in \mathbb{N}$, we write next^k for the k-fold composition of next with itself, so that $\mathrm{next}^k(s_n) = s_{n+k}$.

A *Boolean function* is a function $f : \{0,1\}^m \to \{0,1\}$ for some $m \in \mathbb{N}$. The *support* of such a function f is $\mathrm{supp}(f) = \left\{ x \in \{0,1\}^m \;\middle|\; f(x) = 1 \right\}$.

We write $w[i]$ for the i^{th} symbol in a string w and $w[i..j]$ for the string consisting of the i^{th} through j^{th} symbols. The leftmost symbol of w is $w[0]$, so that $w = w[0..|w|-1]$. For (infinite) sequences $S \in \Sigma^\infty$, the notations $S[i]$ and $S[i..j]$ are defined similarly. A string $w \in \Sigma^*$ is a prefix of a string or sequence $x \in \Sigma^* \cup \Sigma^\infty$, and we write $w \sqsubseteq x$, if there is a string or sequence $y \in \Sigma^* \cup \Sigma^\infty$ such that $wy = x$. A *language*, or *decision problem*, is a set $A \subseteq \{0,1\}^*$. We identify each language A with the sequence $A \in \{0,1\}^\infty$ defined by $A[n] = [\![s_n \in A]\!]$ for all $n \in \mathbb{N}$. If A is a language, then expressions like $\lim_{w \to A} f(w)$ refer to prefixes $w \sqsubseteq A$, e.g., $\lim_{w \to A} f(w) = \lim_{n \to \infty} f(A[0..n-1])$.

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

$$d(w) = \frac{d(w0) + d(w1)}{2} \tag{2.1}$$

for all $w \in \{0,1\}^*$. Intuitively, d is a *strategy for betting* on the successive bits of a sequence $S \in \{0,1\}^\infty$: The quantity $d(w)$ is the amount of money that the gambler using this strategy has after $|w|$ bets if $w \sqsubseteq S$. Condition (2.1) says that the payoffs are fair.

A martingale d *succeeds* on a language $A \subseteq \{0,1\}^*$, and we write $A \in S^\infty[d]$, if $\limsup_{w \to A} d(w) = \infty$. If $t : \mathbb{N} \to \mathbb{N}$, then a martingale d is (*exactly*) $t(n)$-*computable* if its values are rational and there is an algorithm that computes each $d(w)$ in $t(|w|)$ time. A martingale is p-*computable* if it is n^k-computable for some $k \in \mathbb{N}$, and it is p_2-computable if it is $2^{(\log n)^k}$-computable for some $k \in \mathbb{N}$.

Definition. [18] Let X be a set of languages, and let R be a language.

1. X has p-*measure 0*, and we write $\mu_{\mathrm{p}}(X) = 0$, if there is a p-computable martingale d such that $X \subseteq S^\infty[d]$. The condition $\mu_{\mathrm{p}_2}(X) = 0$ is defined analogously.

2. X has *measure 0 in* EXP, and we write $\mu(X \mid \mathrm{EXP}) = 0$, if $\mu_{\mathrm{p}_2}(X \cap \mathrm{EXP}) = 0$.

3. R is p-*random* if $\mu_{\mathrm{p}}(\{R\}) \neq 0$, i.e., if there is no p-computable martingale that succeeds on R. Similarly, R is $t(n)$-*random* if no $t(n)$-computable martingale succeeds on R.

It is well known that these definitions impose a nontrivial measure structure on EXP [18]. For example, $\mu(\mathrm{EXP} \mid \mathrm{EXP}) \neq 0$.

We use the following fact in our arguments.

Lemma 2.1 *[3, 16] The following five conditions are equivalent.*

1. $\mu(\mathrm{NP} \mid \mathrm{EXP}) \neq 0$.

2. $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$.

3. $\mu_{\mathrm{p}_2}(\mathrm{NP}) \neq 0$.

4. *There exists a p-random language* $R \in \mathrm{NP}$.

5. *For every* $k \geq 2$*, there exists an* n^k*-random language* $R \in \mathrm{NP}$.

Finally, we note that $\mu(\mathrm{P} \mid \mathrm{EXP}) = 0$ [18], so $\mu(\mathrm{NP} \mid \mathrm{EXP}) \neq 0$ implies $\mathrm{P} \neq \mathrm{NP}$.

3 Interval Martingales

This section presents a method for constructing a class of useful martingales that are easily controlled and analyzed.

We let $\leq$ and $<$ denote the standard ordering of $\{0,1\}^*$. An *interval* in $\{0,1\}^*$ is a set of the form

$$[u, v) = \left\{ z \in \{0,1\}^* \;\middle|\; u \leq z < v \right\},$$

where $u, v \in \{0,1\}^*$.
Definition.

1. An *interval condition* is an ordered pair (I, f), where I is an interval in $\{0,1\}^*$ and $f : \{0,1\}^{|I|} \to \{0,1\}$ is a Boolean function with $\mathrm{supp}(f) \neq \emptyset$.

2. A string $w \in \{0,1\}^*$ *satisfies* an interval condition $([u,v), f)$ if $|w| \geq \mathrm{ind}(v)$ and $f(w[\mathrm{ind}(u)..\mathrm{ind}(v)-1]) = 1$.

3. A language $R \subseteq \{0,1\}^*$ *satisfies* an interval condition $([u,v), f)$ if there is a prefix $w \sqsubseteq R$ that satisfies (I, f).

We define the *wager set* of a martingale d to be the set

$$W(d) = \left\{ w \in \{0,1\}^* \;\middle|\; d(w0) \neq d(w1) \right\},$$

i.e., the set of all strings w at which d actually bets on the next bit. It is clear that $d(w) > 0$ for all $w \in W(d)$ and that a martingale d is completely determined by its initial value $d(\lambda)$, its wager set $W(d)$, and the ratios $d(w0)/d(w)$ for $w \in W(d)$.

Lemma 3.1 *For each interval condition $(I, f) = ([u,v), f)$, there is a unique martingale d_I^f with the following three properties.*

(i) $d_I^f(\lambda) = 1$.

(ii) $W(d_I^f) \subseteq \left\{ w \in \{0,1\}^* \mid s_{|w|} \in I \right\}$.

(iii) *For all $w \in \{0,1\}^{\mathrm{ind}(v)}$,*

$$d_I^f(w) = \begin{cases} \frac{2^{|I|}}{|\mathrm{supp}(f)|} & \text{if } f(w[\mathrm{ind}(u)..\mathrm{ind}(v)-1]) = 1 \\ 0 & \text{if } f(w[\mathrm{ind}(u)..\mathrm{ind}(v)-1]) = 0. \end{cases}$$

Proof (sketch). These conditions clearly specify $d_I^f(w)$ for all w satisfying $|w| < \mathrm{ind}(u)$ or $|w| \geq \mathrm{ind}(v)$. Propagate the values in (iii) “backwards” by inductively defining

$$d_I^f(w) = \frac{d_I^f(w0) + d_I^f(w1)}{2} \tag{3.1}$$

for all w satisfying $\mathrm{ind}(u) \leq |w| < \mathrm{ind}(v)$. The crucial things to note are that the arithmetic mean of the values in (iii) is 1 and that (3.1) preserves this condition at each length, so that $d_I^f(w) = 1$ holds for all $w \in \{0,1\}^{\mathrm{ind}(u)}$. These are exactly the values dictated by (i) and (ii). □

Definition. The *interval martingale* given by an interval condition (I, f) is the martingale $d_I^f(w)$ of Lemma 3.1.

Observation 3.2 *If* $(I, f) = ([u, v), f)$ *is an interval condition and* $R \subseteq \{0,1\}^*$ *satisfies* (I, f)*, then* $d_I^f(w) = 2^{|I|}/|\mathrm{supp}(f)|$ *holds for every prefix* $w \sqsubseteq R$ *with* $|w| \geq \mathrm{ind}(v)$*.*

We often construct a martingale from many component martingales. Here is one way of doing this. We say that a family $\mathcal{M}$ of martingales *has disjoint wagers* if, for all $d, d' \in \mathcal{M}$,

$$d \neq d' \implies W(d) \cap W(d') = \emptyset.$$

Definition. If $\mathcal{M}$ is a family of martingales with disjoint wagers, then the *product martingale* $\otimes_{d \in \mathcal{M}} d$ is the unique martingale d' with the following three properties.

(i) $d'(\lambda) = 1$.

(ii) $W(d') = \cup_{d \in \mathcal{M}} W(d)$.

(iii) For all $d \in \mathcal{M}$, $w \in W(d)$, and $b \in \{0,1\}$,

$$d'(wb) = d'(w)\frac{d(wb)}{d(w)}.$$

Notation. If $[u, v)$ and $[u', v')$ are intervals in $\{0,1\}^*$, then the condition $[u, v) < [u', v')$ means that $v \leq u'$.

Lemma 3.3 *Assume that* $(I_0, f_0), (I_1, f_1), \ldots$ *are interval conditions with* $I_0 < I_1 < \ldots$*. Let* $d = \otimes_{k=0}^{\infty} d_{I_k}^{f_k}$*. If* $R \subseteq \{0,1\}^*$ *satisfies* (I_k, f_k) *for all* $k \in \mathbb{N}$*, then*

$$\limsup_{w \to R} d(w) \geq \prod_{k=0}^{\infty} \frac{2^{|I_k|}}{|\mathrm{supp}(f_k)|}.$$

Proof. This follows inductively from Observation 3.2. □

It is clear that Lemma 3.3 can be generalized to a much wider class of conditions and products, but interval conditions are sufficient for our purposes here.

4 Inseparable Disjoint NP Pairs and the Measure of NP

This section presents our main theorem, which says that, if NP does not have measure 0 in EXP, then there are disjoint NP pairs that are P-inseparable. In fact, for each $k \in \mathbb{N}$, there is a disjoint NP pair that is $\mathrm{TIME}(2^{n^k})$-inseparable.

It is convenient for our arguments to use a slight variant of the separability notion.

Definition. Let (A, B) be a pair of (not necessarily disjoint) languages, and let $\mathcal{C}$ be a class of languages.

1. A language $S \subseteq \{0,1\}^*$ *almost separates* (A, B) if there is a finite set $D \subseteq \{0,1\}^*$ such that S separates $(A - D, B - D)$.

2. We say that (A, B) is *$\mathcal{C}$-almost separable* if there is a language $S \in \mathcal{C}$ that almost separates (A, B).

Observation 4.1 *If a pair (A, B) is not $\mathcal{C}$-almost separable, then $(A - D, B - D)$ is $\mathcal{C}$-inseparable for every finite set D.*

Before proving our main theorem, we sketch the intuitive idea of the proof. We want to construct a disjoint NP pair (A, B) that is P-inseparable. Our hypothesis, that NP does not have measure 0 in EXP, implies that NP contains a language R that is p-random. Since we are being intuitive, we ignore the subtleties of p-randomness and regard R as a sequence of independent, fair coin tosses (with the n^{th} toss heads iff $s_n \in R$) that just happens to be in NP. If we use these coins to randomly put strings in A or B but not both, we can count on the randomness to thwart any would-be separator in P.

The challenge here is that, if we are to deduce $A, B \in \mathrm{NP}$ from $R \in \mathrm{NP}$, we must make the conditions "$s_n \in A$" and "$s_n \in B$" depend on the coin tosses in a *monotone* way; i.e., adding a string to R must not move a string out of A or out of B.

This monotonicity restriction might at first seem to prevent us from ensuring that A and B are disjoint. However, this is not the case. Suppose that we decide membership on the n^{th} string s_n in A and B in the following manner. We toss $2 \log n$ independent coins. If the first $\log n$ tosses all come up heads, we put s_n in A. If the second $\log n$ tosses all come up heads, we put s_n in B. If our coin tosses are taken from R, which is in NP, then A and B will be in NP. Each string s_n will be in A with probability $\frac{1}{n}$, in B with probability $\frac{1}{n}$, and in $A \cap B$ with probability $\frac{1}{n^2}$. Since $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges and $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges, the first and second Borel-Cantelli lemmas tell us that A and B are infinite and $A \cap B$ is finite. Since $A \cap B$ is finite, we can subtract it from A and B, leaving two disjoint NP languages that are, by the randomness of the construction, P-inseparable.

What prevents this intuitive argument from being a proof sketch is the fact that the language R is not truly random, but only p-random. The proof that $A \cap B$ is finite thus becomes problematic. There is a resource-bounded extension of the first Borel-Cantelli lemma [18] that works for p-random sequences, but this extension requires the relevant sum of probabilities to be p-convergent, i.e., to converge much more quickly than $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Fortunately, in this particular instance, we can achieve our objective without p-convergence or the (classical or resource-bounded) Borel-Cantelli lemmas. We do this by modifying the above construction. Instead of putting the n^{th} string into each language with probability $\frac{1}{n}$, we put each string x into each of A and B with probability $2^{-|x|}$ so that x is in $A \cap B$ with probability $2^{-2|x|}$. By the Cauchy condensation test, the relevant series have the same convergence behavior as those in our intuitive argument, but we can now replace slow approximations of tails of $\sum_{n=1}^{\infty} \frac{1}{n^2}$ with fast and exact computations of geometric series.

We now turn to the details.

Construction 4.2 1. *Define the functions $u, v : \{0,1\}^* \to \{0,1\}^*$ by the recursion*

$$
\begin{aligned}
&u(\lambda) = \lambda, \\
&v(x) = \text{next}^{|x|}(u(x)), \\
&u(\text{next}(x) = \text{next}^{|x|}(v(x)).
\end{aligned}
$$

2. *For each $x \in \{0,1\}^*$, define the intervals*

$$I_x = [u(x), v(x)), \;\; J_x = [v(x), u(next(x))).$$

3. *For each $R \subseteq \{0,1\}^*$, define the languages*

$$
\begin{aligned}
&A^+(R) = \left\{ x \;\middle|\; I_x \subseteq R \right\}, \;\; B^+(R) = \left\{ x \;\middle|\; J_x \subseteq R \right\}, \\
&A(R) = A^+(R) - B^+(R), \;\; B(R) = B^+(R) - A^+(R).
\end{aligned}
$$

Note that each $|I_x| = |J_x| = |x|$. Also, $I_\lambda = J_\lambda = \emptyset$ (so $\lambda \in A^+(R) \cap B^+(R)$), and

$$I_0 < J_0 < I_1 < J_1 < I_{00} < J_{00} < I_{01} < \dots,$$

with these intervals covering all of $\{0,1\}^*$.

A routine witness argument gives the following.

Observation 4.3 *1. If $R \in \mathrm{NP}$, then $A^+(R), B^+(R) \in \mathrm{NP}$.*

2. If $R \in \mathrm{NP}$ and $|A^+(R) \cap B^+(R)| < \infty$, then $(A(R), B(R))$ is a disjoint NP pair.

We now prove two lemmas about Construction 4.2.

Lemma 4.4 *Let $k \in \mathbb{N}$. If $R \subseteq \{0,1\}^*$ is $2^{(\log n)^{k+2}}$- random, then $(A^+(R), B^+(R))$ is not $\mathrm{TIME}(2^{n^k})$-almost separable.*

Proof. Let $k \in \mathbb{N}$, and assume that $(A^+(R), B^+(R))$ is $\mathrm{TIME}(2^{n^k})$-almost separable. It suffices to show that R is not $2^{(\log n)^{k+2}}$- random.

By our assumption, there is a language $S \in \mathrm{TIME}(2^{n^k})$ that almost separates $(A^+(R), B^+(R))$. Then there exists a positive integer l such that S separates $(A^+(R) - \{0,1\}^{<l}, B^+(R) - \{0,1\}^{<l})$.

For each string $x \in \{0,1\}^*$, define the interval condition $(K_x, \mathrm{NAND}_{|x|})$, where

$$K_x = \begin{cases} J_x & \text{if } x \in S \\ I_x & \text{if } x \notin S \end{cases}$$

and $\mathrm{NAND}_m : \{0,1\}^m \to \{0,1\}$ is defined by

$$\mathrm{NAND}_m(u) = \begin{cases} 1 & \text{if } u \neq 1^m \\ 0 & \text{if } u = 1^m \end{cases}$$

for all $u \in \{0,1\}^m$. We generally omit the subscript from NAND, writing this interval condition as (K_x, NAND).

Now let $x \in \{0,1\}^*$ with $|x| \geq l$. Since S separates $(A^+(R) - \{0,1\}^{<l}, B^+(R) - \{0,1\}^{<l})$,

$$\begin{aligned} x \in S \quad &\Longrightarrow \quad x \notin B^+(R) \Longrightarrow J_x \not\subseteq R \\ &\Longrightarrow \quad K_x \not\subseteq R \Longrightarrow R \text{ satisfies } (K_x, \mathrm{NAND}), \end{aligned}$$

and

$$\begin{aligned} x \notin S &\implies x \notin A^+(R) \implies I_x \not\subseteq R \\ &\implies K_x \not\subseteq R \implies R \text{ satisfies } (K_x, \mathrm{NAND}). \end{aligned}$$

This shows that R satisfies the interval condition (K_x, NAND) for all $x \in \{0,1\}^*$ with $|x| \geq l$.

Now define the martingale

$$d = \bigotimes_{\substack{x \in \{0,1\}^* \\ |x| \geq l}} d_{K_k}^{\mathrm{NAND}},$$

where $d_{K_k}^{\mathrm{NAND}}$ is the interval martingales given by (K_x, NAND). Then, by Lemma 3.3,

$$\begin{aligned} \limsup_{w \to R} d(w) &\geq \prod_{\substack{x \in \{0,1\}^* \\ |x| \geq l}} \frac{2^{|K_x|}}{|\mathrm{supp}(\mathrm{NAND})|} \\ &= \prod_{k=l}^{\infty} \prod_{x \in \{0,1\}^k} \frac{2^{|x|}}{2^{|x|}-1} \\ &= \prod_{k=l}^{\infty} \left(\frac{2^k}{2^k - 1} \right)^{2^k}. \end{aligned}$$

Since

$$\begin{aligned} \left(\frac{2^k}{2^k - 1} \right)^{2^k} &= \left(\frac{1}{1 - 2^{-k}} \right)^{2^k} > \left(\frac{1 - 2^{-2k}}{1 - 2^{-k}} \right)^{2^k} \\ &= (1 + 2^{-k})^{2^k} \to e \end{aligned}$$

as $k \to \infty$, it follows that

$$\limsup_{w \to R} d(w) = \infty.$$

Hence d succeeds on R.

The algorithm in Figure 1 computes the martingale d. To estimate the running time of this algorithm, note the following.

(i) Since $S \in \mathrm{TIME}(2^{n^k})$, each test "$x \in S$" or "$x \notin S$" takes at most $2^{|x|^k} \leq 2^{(\log|w|)^k}$ steps.

(ii) Each ratio $d_{I_k}^{\mathrm{NAND}}(w)/d_{I_k}^{\mathrm{NAND}}(w')$ or $d_{J_k}^{\mathrm{NAND}}(w)/d_{J_k}^{\mathrm{NAND}}(w')$ can be computed by calculating (from leaves to root) a tree of $2^{|x|+1} - 1 = O(|w|)$ values.

(iii) The while loop executes fewer than $|w|$ times.

It follows from these and simpler considerations that $d(w)$ is computable in $O(2^{(\log|w|)^{k+2}})$ time. Since d succeeds on R, it follows that R is not $2^{(\log n)^{k+2}}$-random.

□

Lemma 4.5 *If $R \subseteq \{0,1\}^*$ is p-random, then $|A^+(R) \cap B^+(R)| < \infty$.*

Proof. Assume that $|A^+(R) \cap B^+(R)| = \infty$. It suffices to show that R is not p-random.

For each string $x \in \{0,1\}^*$, define the interval condition $(L_x, \mathrm{AND}_{2|x|})$, where

$$L_x = I_x \cup J_x$$

and $\mathrm{AND}_m : \{0,1\}^m \to \{0,1\}$ is defined by

$$\mathrm{AND}_m(u) = \begin{cases} 1 & \text{if } u = 1^m \\ 0 & \text{if } u \neq 1^m \end{cases}$$

for all $u \in \{0,1\}^m$. As before, we generally omit the subscript from AND, writing this interval condition as (L_x, AND).

Note that, for each $x \in \{0,1\}^*$,

$$x \in A^+(R) \cap B^+(R) \iff R \text{ satisfies } (L_x, \mathrm{AND}).$$

It follows by Lemma 3.1 that, if $x \in A^+(R) \cap B^+(R)$, then $d_{L_x}^{\mathrm{AND}} = 2^{2|x|}$ holds for every prefix $w \sqsubseteq R$ with $|w| \geq \mathrm{ind}(v(x))$. Hence, if we define the martingale

$$d = \sum_{x \in \{0,1\}^*} 2^{-2|x|} d_{L_x}^{\mathrm{AND}}$$

(which is well defined because $d(\lambda) = 2 < \infty$), then

$$\lim_{w \to R} d(w) = |A^+(R) \cap B^+(R)| = \infty.$$

input $w \in \{0,1\}^*$;
if $|w| < \mathrm{ind}(u(0^l))$ **then** output 1 **else**
$x := 0^l$; $w' := w[0..\mathrm{ind}(u(x)) - 1]$; $d := 1$;
while $|w'| \leq |w| - 2|x|$ **do**
  $w' := w[0..|w| + 2|x| - 1]$;
  if $x \in S$ **then**
    $d := d \cdot \frac{2^{|x|}}{2^{|x|}-1} \cdot [\![w' \text{ } satisfies \text{ } (J_x, \mathrm{NAND})]\!]$
  else
    $d := d \cdot \frac{2^{|x|}}{2^{|x|}-1} \cdot [\![w' \text{ } satisfies \text{ } (I_x, \mathrm{NAND})]\!]$
  endif;
  $x := \mathrm{next}(x)$
endwhile;
if $|w'| < |w| - |x|$ **then**
  if $x \notin S$ **then** $d := d \cdot \frac{d_{I_x}^{\mathrm{NAND}}(w)}{d_{I_x}^{\mathrm{NAND}}(w')}$ **endif**
else
  $w' := w[0..|w'| + |x| - 1]$;
  if $x \notin S$ **then**
    $d := d \cdot \frac{2^{|x|}}{2^{|x|}-1} \cdot \frac{d_{I_x}^{\mathrm{NAND}}(w)}{d_{I_x}^{\mathrm{NAND}}(w')} \cdot [\![w' \text{ } satisfies \text{ } (I_x, \mathrm{NAND})]\!]$
  else
    $d := d \cdot \frac{d_{J_x}^{\mathrm{NAND}}(w)}{d_{J_x}^{\mathrm{NAND}}(w')}$
  endif
endif;
output d
endif.

Figure 1: Algorithm for martingale d in proof of Lemma 4.4.

Hence d succeeds on R.

We now consider the complexity of computing $d(w)$ for a given string $w \in \{0,1\}^*$. We can in time polynomial in $|w|$ compute the least y such that $u_y \geq s_{|w|}$ and compute the sum

$$\sigma_1 = \sum_{x<y} 2^{-2|x|} d_{L_x}^{\mathrm{AND}}(w).$$

We can then compute the quantities

$$\sigma_2 = 2^{-2|y|} |[y, 0^{|y|+1})|$$

and

$$\sigma_3 = 2^{-|y|},$$

also in time polynomial in $|w|$. Since

$$\begin{aligned} d(w) &= \sigma_1 + \sum_{\substack{x \geq y \\ |x|=|y|}} 2^{-2|x|} d_{L_x}^{\mathrm{AND}}(w) \\ &\quad + \sum_{\substack{x \\ |x|>|y|}} 2^{-2|x|} d_{L_x}^{\mathrm{AND}}(w) \\ &= \sigma_1 + \sigma_2 + \sigma_3, \end{aligned}$$

this shows that d is computable in polynomial time. Since d succeeds on R, it follows that R is not p-random. □

We now have what we need to prove our main result.

Theorem 4.6 *(main theorem) If NP does not have measure 0 in EXP, then, for every $k \in \mathbb{Z}^+$, there is a disjoint NP pair that is* $\mathrm{TIME}(2^{n^k})$*-inseparable, hence certainly P-inseparable.*

Proof. Assume that $\mu(\mathrm{NP} \mid \mathrm{EXP}) \neq 0$, and let $k \in \mathbb{N}$. Then, by Lemma 2.1, there is a $2^{(\log n)^{k+2}}$-random language $R \in \mathrm{NP}$. By Lemma 4.4, the pair $(A^+(R), B^+(R))$ is not $\mathrm{TIME}(2^{n^k})$-almost separable. Since R is certainly p-random, Lemma 4.5 tells us that $|A^+(R) \cap B^+(R)| < \infty$. It follows by Observation 4.3 that $(A(R), B(R))$ is a disjoint NP pair, and it follows by Observation 4.1 that $(A(R), B(R))$ is $\mathrm{TIME}(2^{n^k})$-inseparable. □

5 Genericity and Measure of Disjoint NP Pairs

In this section we introduce the natural notions of resource-bounded measure and genericity for disjoint pairs and relate them to the existence of P-inseparable pairs in NP. We compare the different strength hypothesis on the measure and genericity of NP and disjNP establishing all the relations in Figure 2.

Notation. Each disjoint pair (A, B) will be coded as an infinite sequence $T \in \{-1, 0, 1\}^\infty$ defined by

$$T[n] = \begin{cases} 1 & \text{if } s_n \in A \\ -1 & \text{if } s_n \in B \\ 0 & \text{if } s_n \notin A \cup B \end{cases}$$

We identify each disjoint pair with the corresponding sequence.

Resource-bounded genericity for disjoint pairs is the natural extension of the concept introduced for languages by Ambos-Spies, Fleischhack and Huwig [1].

Definition. A *condition* C is a set $C \subseteq \{-1, 0, 1\}^*$. A $t(n)$-condition is a condition $C \in \mathrm{DTIME}(t(n))$. A condition C is *dense along a pair* (A, B) if there are infinitely many $n \in \mathbb{N}$ such that $(A, B)[0..n-1]i \in C$ for some $i \in \{-1, 0, 1\}$. A pair (A, B) *meets* a condition C if $(A, B)[0..n-1] \in C$ for some n. A pair (A, B) is $t(n)$-*generic* if (A, B) meets every $t(n)$-condition that is dense along (A, B).

We first prove that generic pairs are inseparable.

Theorem 5.1 *Every $t(\log n)$-generic disjoint pair is $\mathrm{TIME}(t(n))$-inseparable.*

Proof. Let (A, B) be $\mathrm{TIME}(t(n))$-separable with separator S. We define the condition

$$C = \left\{ wb \;\middle|\; b = 1 \text{ if } s_{|w|} \notin S, \text{ and } b = -1 \text{ if } s_{|w|} \in S \right\}.$$

Then $C \in \mathrm{DTIME}(t(\log n))$, C is dense along any pair, and (A, B) does not meet C, so (A, B) is not $t(\log n)$-generic. □

We can now relate genericity in disjNP and inseparable pairs as follows.

Corollary 5.2 *If disjNP contains a $2^{(\log n)^k}$-generic pair for every $k \in \mathbb{N}$, then disjNP contains a $\mathrm{TIME}(2^{n^k})$-inseparable pair for every $k \in \mathbb{N}$.*

Resource-bounded measure on classes of disjoint pairs is the natural extension of the concept introduced for languages by Lutz [18], and is defined by using martingales on a three-symbol alphabet as follows.

Definition.

1. A *pair martingale* is a function $d : \{-1, 0, 1\}^* \to [0, \infty)$ such that for every $w \in \{-1, 0, 1\}^*$
$$d(w) = \frac{1}{4}d(w0) + \frac{3}{8}d(w1) + \frac{3}{8}d(w(-1)).$$

2. A pair martingale d *succeeds on a pair* (A, B) if $\limsup_{w \to (A,B)} d(w) = \infty$.

3. A pair martingale d *succeeds on a class of pairs* $X \subseteq \{-1, 0, 1\}^\infty$ if it succeeds on each $(A, B) \in X$.

Our intuitive rationale for the coefficients in part 1 of this definition is the following. We toss one fair coin to decide whether $s_{|w|} \in A$ and another to decide whether $s_{|w|} \in B$. If both coins come up heads, we toss a third coin to break the tie. The reader may feel that some other coefficients, such as $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$ are more natural here. Fortunately, a routine extension of the main theorem of [4] shows that the value of $\mu(\text{disjNP} \mid \text{disjEXP})$ will be the same for *any* choice of three positive coefficients summing to 1.

When restricting martingales to those computable within a certain resource bound, we obtain a resource-bounded measure that is useful within a complexity class. Here we are interested in the class of disjoint EXP pairs, disjEXP.

Definition.

1. Let p_2 be the class of functions that can be computed in time $2^{(\log n)^{O(1)}}$.

2. A class of pairs $X \subseteq \{-1, 0, 1\}^\infty$ has p_2*-measure 0* if there is a martingale $d \in \mathrm{p}_2$ that succeeds on X.

3. $X \subseteq \{-1, 0, 1\}^\infty$ has p_2*-measure 1* if X^c has p_2-measure 0.

4. A class of pairs $X \subseteq \{-1, 0, 1\}^\infty$ has *measure 0 in disjEXP*, denoted $\mu(X \mid \text{disjEXP}) = 0$, if $X \cap \text{disjEXP}$ has p_2-measure 0.

5. $X \subseteq \{-1, 0, 1\}^\infty$ has *measure 1 in disjEXP* if X^c has measure 0 in disjEXP.

It is easy to verify that p_2-measure is nontrivial on disjEXP (as proven for languages in [18]).

In the following we consider the hypothesis that disjNP does not have measure 0 in disjEXP (written $\mu(\text{disjNP} \mid \text{disjEXP}) \neq 0$). We start by proving that this hypothesis is at least as strong as the well studied $\mu(\text{NP} \mid \text{EXP}) \neq 0$ hypothesis.

Theorem 5.3 *If* $\mu(\text{disjNP} \mid \text{disjEXP}) \neq 0$ *then* $\mu(\text{NP} \mid \text{EXP}) \neq 0$.

Proof. We show that $\mu(\text{NP} \mid \text{EXP}) = 0$ implies $\mu(\text{disjNP} \mid \text{disjEXP}) = 0$. The hypothesis $\mu(\text{NP} \mid \text{EXP}) = 0$ has been proven to be equivalent to $\mu_{\text{p}}(\text{NP}) = 0$ [3]. Breutzmann and Lutz [4] have proven that the p-measure of NP is robust with respect to certain changes in the underlying probability distribution, for instance, $\mu_{\text{p}}(\text{NP}) = 0$ if for every (polynomial-time computable) $\beta \in (0,1)$ there is an $(\beta, 1-\beta)$-p-martingale succeeding on NP, that is, a function d in p such that for every $w \in \{0,1\}^*$

$$d(w) = \beta d(w0) + (1-\beta) d(w1).$$

So we assume that $\mu(\text{NP} \mid \text{EXP}) = 0$. Then, taking $\beta = 1/4$, there is a $(1/4, 3/4)$-p-martingale d that succeeds on NP. We define the pair martingale $D : \{-1,0,1\}^* \to [0,\infty)$ by

$$\begin{aligned}
D(\lambda) &= d(\lambda), \\
D(w0) &= D(w)\frac{d(\overline{w}0)}{d(\overline{w})}, \\
D(w1) &= D(w)\frac{d(\overline{w}1)}{d(\overline{w})}, \\
D(w-1) &= D(w)\frac{d(\overline{w}1)}{d(\overline{w})},
\end{aligned}$$

where $\overline{w} \in \{0,1\}^*$ is defined by $\overline{w}[i] = w[i]$ if $w[i] \in \{0,1\}$ and $\overline{w}[i] = 1$ if $w[i] = -1$.

By definition, D is a p-computable pair martingale. Notice that, for a disjoint pair (A, B), if d succeeds on $A \cup B$ then D succeeds on (A, B). Since NP is closed under union, d does succeed on $A \cup B$, so D succeeds on disjNP, whence $\mu(\text{disjNP} \mid \text{disjEXP}) = 0$. □

We finish by relating measure and genericity for disjoint pairs.

$$
\begin{tikzcd}
 & \mu(\text{disjNP} \mid \text{disjEXP}) \neq 0 \arrow[dl, Rightarrow, "\text{Theorem 5.3}"'] \arrow[dr, Rightarrow, "\text{Theorem 5.4}"] & \\
\mu(\text{NP} \mid \text{EXP}) \neq 0 \arrow[dr, Rightarrow, "\text{Theorem 4.6}"'] & & \begin{array}{c}(\forall k)\text{disjNP contains}\\ \text{a } 2^{(\log n)^k}\text{-generic pair}\end{array} \arrow[dl, Rightarrow, "\text{Corollay 5.2}"] \\
 & \begin{array}{c}(\forall k)\text{ disjNP contains}\\ \text{a TIME}(2^{n^k})\text{-inseparable}\\ \text{pair}\end{array} \arrow[d, Rightarrow, "\text{trivial}"] & \\
\text{P} \neq \text{NP} \cap \text{coNP} \arrow[r, Rightarrow, "\text{trivial}"] & \begin{array}{c}\text{disjNP contains}\\ \text{P-inseparable pairs}\end{array} \arrow[d, Rightarrow, "\text{trivial}"] & \text{P} \neq \text{UP} \arrow[l, Rightarrow, "\text{Grollman \& Selman [88]}"'] \\
 & \text{P} \neq \text{NP} &
\end{tikzcd}
$$

Figure 2: Relations among some strong hypotheses.

Theorem 5.4 *If* $\mu(\text{disjNP} \mid \text{disjEXP}) \neq 0$*, then disjNP contains a* $2^{(\log n)^k}$*-generic pair for every* $k \in \mathbb{N}$*.*

Proof. Lorentz and Lutz prove in [17] that, for every $k \in \mathbb{N}$, for every p_2-exact probability measure γ, the set of $2^{(\log n)^k}$-generic languages has p_2-γ-measure 1. The concept of p_2-exact probability measure includes cases such as a single (polynomial-time computable) bias $\beta \in (0, 1)$.

An easy extension of [17] to a three letter alphabet implies that for every $k \in \mathbb{N}$, the set of $2^{(\log n)^k}$-generic pairs has p_2-measure 1. This implies the theorem. $\square$

References

[1] K. Ambos-Spies, H. Fleischhack, and H. Huwig. Diagonalizations over polynomial time computable sets. *Theoretical Computer Science*, 51:177–204, 1987.

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

[3] K. Ambos-Spies, S. A. Terwijn, and X. Zheng. Resource bounded randomness and weakly complete problems. *Theoretical Computer Science*, 172:195–207, 1997.

[4] JM Breutzmann and JH Lutz. Equivalence of measures of complexity classes. *SIAM Journal on Computing*, 29(1):302–326, 1999.

[5] S. Cook and R. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44:36–50, 1979.

[6] S. Even, A. L. Selman, and Y. Yacobi. The complexity of promise problems with applications to public-key cryptography. *Information and Control*, 61(2):159–173, 1984.

[7] C. Glasser, A. L. Selman, and S. Sengupta. Reductions between disjoint NP-pairs. *Information and Computation*, 200:247–267, 2005.

[8] C. Glasser, A. L. Selman, S. Sengupta, and L. Zhang. Disjoint NP-pairs. *SIAM Journal on Computing*, 33:1369–1416, 2004.

[9] C. Glasser, A. L. Selman, and L. Zhang. *Survey of Disjoint NP-pairs and Relations to Propositional Proof Systems*, pages 241–253. 2006.

[10] C. Glasser, A. L. Selman, and L. Zhang. Canonical disjoint NP-pairs of propositional proof systems. *Theoretical Computer Science*, 370:60–73, 2007.

[11] C. Glasser, A. L. Selman, and L. Zhang. The informational content of canonical disjoint NP-pairs. In *COCOON*, LNCS, pages 307–317. Springer, 2007.

[12] O. Goldreich. *On Promise Problems: A Survey*, pages 254–290. 2006.

[13] J. Grollman and A. Selman. Complexity measures for public-key cryptosystems. *SIAM Journal on Computing*, 11:309–335, 1988.

[14] J.M. Hitchcock. Resource-bounded measure bibliography. http://www.cs.uwyo.edu/~jhitchco/bib/rbm.shtml.

[15] S. Homer and A. L. Selman. Oracles for structural properties: The isomorphism problem and public-key cryptography. *Journal of Computer and System Sciences*, 44:287–301, 1992.

[16] DW Juedes and JH Lutz. Weak completeness in E and E2, 1995.

[17] AK Lorentz and JH Lutz. Genericity and randomness over feasible probability measures. *Theoretical Computer Science*, 207(1):245–259, 1998.

[18] J. H. Lutz. Almost everywhere high nonuniform complexity. *Journal of Computer and System Sciences*, 44(2):220–258, 1992.

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

[20] J. H. Lutz and E. Mayordomo. Twelve problems in resource-bounded measure. In G. Păun, G. Rozenberg, and A. Salomaa, editors, *Current Trends in Theoretical Computer Science, entering the 21st century*, pages 83–101. World Scientific Publishing, 2001.

[21] A. Razborov. On provably disjoint np pairs. Technical Report 94-006, ECCC, 1994.

[22] A.L. Selman. Complexity issues in cryptography. In *Computational complexity theory (Atlanta, GA, 1988)*, volume 38 of *Proc. Sympos. Appl. Math.*, pages 92–107. Amer. Math. Soc., 1989.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/