ELSEVIER
Theoretical Computer Science 289 (2002) 861–869
Theoretical Computer Science
www.elsevier.com/locate/tcs

Note

MAX3SAT is exponentially hard to approximate if NP has positive dimension ☆

John M. Hitchcock

Department of Computer Science, Iowa State University, 226 Atanasoff Hall, Ames, IA 50011-1040, USA

Received May 2001; received in revised form July 2001; accepted July 2001
Communicated by J. Díaz

Abstract

Under the hypothesis that NP has positive p-dimension, we prove that any approximation algorithm $\mathscr{A}$ for MAX3SAT must satisfy at least one of the following:

1. For some $\delta > 0$, $\mathscr{A}$ uses at least 2^{n^δ} time.
2. For all $\varepsilon > 0$, $\mathscr{A}$ has performance ratio less than $\frac{7}{8} + \varepsilon$ on an exponentially dense set of satisfiable instances.

As a corollary, this solves one of Lutz and Mayordomo's "Twelve problems on resource-bounded measure" (Bull. European Assoc. Theoret. Comput. Sci. 68 (1999) 64–80).

Keywords: Resource-bounded measure; Resource-bounded dimension; Inapproximability; MAX3SAT

1. Introduction

MAX3SAT is a well-studied optimization problem. Tight bounds on its polynomial-time approximability are known:

(1) There exists a polynomial-time $\frac{7}{8}$-approximation algorithm [5, 3].[1]
(2) If $P \neq NP$, then for all $\varepsilon > 0$, there does not exist a polynomial-time $(\frac{7}{8} + \varepsilon)$-approximation algorithm [4].

Recently, there has been some investigation of approximating MAX3SAT in exponential time. For example, for any $\varepsilon \in (0, \frac{1}{8}]$, Dantsin et al. [2] give a $(\frac{7}{8} + \varepsilon)$-approximation

☆ This research was supported in part by National Science Foundation Grant 9988483.
E-mail address: jhitchco@cs.iastate.edu (John M. Hitchcock).
[1] An algorithm with conjectured performance ratio $\frac{7}{8}$ was given in Ref. [5], and this conjecture has since been proved according to Ref. [3].

PII: S0304-3975(01)00340-1

algorithm for MAX3SAT running in time $2^{8\varepsilon k}$ where k is the number of clauses in a formula.

Given these results, it is natural to ask for stronger lower bounds on computation time for MAX3SAT approximation algorithms that have performance ratio greater than $\frac{7}{8}$. Such lower bounds are not known to follow from the hypothesis $\mathrm{P} \neq \mathrm{NP}$. In this note we address this question using a stronger hypothesis involving resource-bounded dimension.

About a decade ago, Lutz [6] presented resource-bounded measure as an analogue for classical Lebesgue measure in complexity theory. Resource-bounded measure provides strong, reasonable hypotheses which seem to have more explanatory power than weaker, traditional complexity-theoretic hypotheses. The hypothesis that NP does not have p-measure 0, $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$, implies $\mathrm{P} \neq \mathrm{NP}$ and is known to have many plausible consequences that are not known to follow from $\mathrm{P} \neq \mathrm{NP}$.

Resource-bounded dimension was recently introduced by Lutz [7] as an analogue of classical Hausdorff dimension for complexity theory. Resource-bounded dimension refines resource-bounded measure by providing a spectrum of weaker, but still strong, hypotheses. We will use the hypothesis that NP has positive p-dimension, $\dim_{\mathrm{p}}(\mathrm{NP}) > 0$. This hypothesis is implied by $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$ and implies $\mathrm{P} \neq \mathrm{NP}$.

Under the hypothesis $\dim_{\mathrm{p}}(\mathrm{NP}) > 0$, we give an exponential-time lower bound for approximating MAX3SAT beyond the known polynomial-time achievable ratio of $\frac{7}{8}$ on all but a subexponentially-dense set of satisfiable instances. Put another way, we prove:

If $\dim_{\mathrm{p}}(\mathrm{NP}) > 0$, then any approximation algorithm $\mathscr{A}$ for MAX3SAT must satisfy at least one of the following:

(1) For some $\delta > 0$, $\mathscr{A}$ uses at least $2^{n^{\delta}}$ time.

(2) For all $\varepsilon > 0$, $\mathscr{A}$ has performance ratio less than $\frac{7}{8} + \varepsilon$ on an exponentially dense set of satisfiable instances.

Lutz and Mayordomo asked whether the hypothesis $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$ implies an exponential-time lower bound on approximation schemes for MAXSAT [8]. Our main theorem gives a strong affirmative answer to this question: we obtain a stronger conclusion from the weaker $\dim_{\mathrm{p}}(\mathrm{NP}) > 0$ hypothesis. In fact, after we present the theorem, we give an easy proposition that achieves an exponential-time lower bound from a hypothesis even weaker than $\dim_{\mathrm{p}}(\mathrm{NP}) > 0$.

In Section 2 we give our notation and basic definitions. Resource-bounded measure and dimension are briefly reviewed in Section 3. Section 4 contains a dimension result used in proving our main theorem. The main theorem is proved in Section 5. Section 6 concludes by summarizing the inapproximability results for MAX3SAT under strong hypotheses.

2. Preliminaries

The set of all finite binary strings is $\{0,1\}^*$. We use the standard enumeration of binary strings $s_0 = \lambda$, $s_1 = 0$, $s_2 = 1$, $s_3 = 00, \ldots$. The length of a string $x \in \{0,1\}^*$ is denoted by $|x|$.

All *languages* (decision problems) in this paper are encoded as subsets of $\{0,1\}^*$. For a language $A \subseteq \{0,1\}^*$, we define $A_{\leqslant n} = \{x \in A \mid |x| \leqslant n\}$. We write $A[0..n-1]$ for the n-bit prefix of the characteristic sequence of A according to the standard enumeration of strings.

We say that a language A is (*exponentially*) *dense* if there is an $\alpha > 0$ such that $|A_{\leqslant n}| > 2^{n^\alpha}$ holds for all but finitely many n. We write DENSE for the class of all dense languages.

For any classes $\mathscr{C}$ and $\mathscr{D}$ of languages we define the classes

$$\mathscr{C} \uplus \mathscr{D} = \{A \cup B \mid A \in \mathscr{C}, B \in \mathscr{D}\}$$

and

$$\mathrm{P_m}(\mathscr{C}) = \{A \subseteq \{0,1\}^* \mid (\exists B \in \mathscr{C}) A \leqslant_{\mathrm{m}}^{\mathrm{p}} B\}.$$

A real-valued function $f : \{0,1\}^* \to [0,\infty)$ is polynomial-time computable if there exists a polynomial-time computable function $g : \mathbb{N} \times \{0,1\}^* \to [0,\infty) \cap \mathbb{Q}$ such that

$$|f(x) - g(n,x)| \leqslant 2^{-n}$$

for all $x \in \{0,1\}^*$ and $n \in \mathbb{N}$ where n is represented in unary.

For an instance x of 3SAT we write MAX3SAT(x) for the maximum fraction of clauses of x that can be satisfied by a single assignment.

An *approximation algorithm* $\mathscr{A}$ for MAX3SAT outputs an assignment of the variables for each instance of 3SAT. For each instance x we write $\mathscr{A}(x)$ for the fraction of clauses satisfied by the assignment produced by $\mathscr{A}$ for x.

An approximation algorithm $\mathscr{A}$ has *performance ratio* α on x if $\mathscr{A}(x) \geqslant \alpha \cdot \mathrm{MAX3SAT}(x)$. If $\mathscr{A}$ has performance ratio α on all instances, then $\mathscr{A}$ is an α-*approximation algorithm*.

Håstad proved the following in order to show that satisfiable instances of 3SAT cannot be distinguished from instances x with $\mathrm{MAX3SAT}(x) < \frac{7}{8} + \varepsilon$ in polynomial-time unless P = NP.

Theorem 2.1 (Håstad [4]). *For each $\varepsilon > 0$, there exists a polynomial-time computable function f_ε such that for all $x \in \{0,1\}^*$,*

$$x \in \mathrm{SAT} \Rightarrow \mathrm{MAX3SAT}(f_\varepsilon(x)) = 1$$

$$x \notin \mathrm{SAT} \Rightarrow \mathrm{MAX3SAT}(f_\varepsilon(x)) < \tfrac{7}{8} + \varepsilon.$$

We will use the functions f_ε from Theorem 2.1 later in the paper.

3. Resource-bounded measure and dimension

In this section we review enough resource-bounded measure and dimension to present our result. Full details of these theories are available in Lutz's introductory papers [6, 7].

Definition 3.1. Let $s \in [0, \infty)$.
(1) A function $d : \{0,1\}^* \to [0, \infty)$ is an *s-gale* if for all $w \in \{0,1\}^*$,

$$d(w) = \frac{d(w0) + d(w1)}{2^s}.$$

(2) A *martingale* is a 1-gale.

Intuitively, a gale is viewed as a function betting on an unknown binary sequence. If w is a prefix of the sequence, then the capital of the gale after placing its first $|w|$ bets is given by $d(w)$. Assuming that w is a prefix of the sequence, the gale places bets on $w0$ and $w1$ also being prefixes. The parameter s determines the fairness of the betting; as s decreases the betting is less fair. The goal of a gale is to bet successfully on languages.

Definition 3.2. Let $s \in [0, \infty)$ and let d be an s-gale.
(1) We say d *succeeds on* a language A if

$$\limsup_{n \to \infty} d(A[0..n-1]) = \infty.$$

(2) The *success set* of d is

$$S^\infty[d] = \{A \subseteq \{0,1\}^* \mid d \text{ succeeds on } A\}.$$

Measure and dimension are defined in terms of succeeding martingales and gales, respectively.

Definition 3.3. Let $\mathscr{C}$ be a class of languages.
(1) $\mathscr{C}$ has p-measure 0, written $\mu_p(\mathscr{C}) = 0$, if there exists a polynomial-time martingale d with $\mathscr{C} \subseteq S^\infty[d]$.
(2) The p-dimension of $\mathscr{C}$ is

$$\dim_p(\mathscr{C}) = \inf \left\{ s \,\middle|\, \begin{array}{l} \text{there exists a polynomial-time} \\ s\text{-gale } d \text{ for which } \mathscr{C} \subseteq S^\infty[d] \end{array} \right\}.$$

For any class $\mathscr{C}$, $\dim_p(\mathscr{C}) \in [0, 1]$. We are interested in hypotheses on the p-dimension and p-measure of NP. The following implications are easy to verify.

$$\begin{aligned} \mu_p(\mathrm{NP}) \neq 0 &\Rightarrow \dim_p(\mathrm{NP}) = 1 \\ &\Rightarrow \dim_p(\mathrm{NP}) > 0 \\ &\Rightarrow \mathrm{P} \neq \mathrm{NP}. \end{aligned}$$

The following simple lemma will be useful in proving our main result.

Lemma 3.4. *Let $\mathscr{C}$ be a class of languages and $c \in \mathbb{N}$.*
(1) *If $\mu_p(\mathscr{C}) = 0$, then $\mu_p(\mathscr{C} \uplus \mathrm{DTIME}(2^{cn})) = 0$.*
(2) $\dim_p(\mathscr{C} \uplus \mathrm{DTIME}(2^{cn})) = \dim_p(\mathscr{C})$.

Proof. Let $s \in [0,1]$ be rational and assume that there is a polynomial-time s-gale d succeeding on $\mathscr{C}$. It suffices to give a polynomial-time s-gale succeeding on $\mathscr{C} \uplus \mathrm{DTIME}(2^{cn})$. By the Exact Computing Lemma of [7], we may assume that d is exactly computable in polynomial-time. Let $M_0, M_1, \ldots$ be a standard enumeration of all Turing machines running in time 2^{cn}. Define for each $i \in \mathbb{N}$ and $w \in \{0,1\}^*$,

$$d_i(w1) = \begin{cases} 2^s d_i(w) & \text{if } M_i \text{ accepts } s_{|w|}, \\ \frac{d(w1)}{d(w)} d_i(w) & \text{if } d(w) \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

$$d_i(w0) = 2^s d_i(w) - d_i(w1).$$

Let $d' = \sum_{i=0}^{\infty} 2^{-i} d_i$. Then d' is a polynomial-time computable s-gale. Let $A \in \mathscr{C}$ and $B = L(M_i) \in \mathrm{DTIME}(2^{cn})$. Then for all $n \in \mathbb{N}$, $d_i((A \cup B)[0..n-1]) \geqslant 2^{-i} d(A[0..n-1])$. Because $A \in S^\infty[d]$, $A \cup B \in S^\infty[d_i] \subseteq S^\infty[d']$. □

4. Dimension of $\mathrm{P_m}(\mathrm{DENSE}^c)$

Lutz and Mayordomo [9] proved that a superclass of $\mathrm{P_m}(\mathrm{DENSE}^c)$ has p-measure 0, so $\mu_p(\mathrm{P_m}(\mathrm{DENSE}^c)) = 0$. In this section we prove the stronger result that $\dim_p(\mathrm{P_m}(\mathrm{DENSE}^c)) = 0$.

We use the binary entropy function $\mathscr{H} : [0,1] \to [0,1]$ defined by

$$\mathscr{H}(x) = \begin{cases} -x \log x - (1-x)\log(1-x) & \text{if } x \in (0,1), \\ 0 & \text{if } x \in \{0,1\}. \end{cases}$$

Lemma 4.1. *For all $n \in \mathbb{N}$ and $0 \leqslant k \leqslant n$,*

$$\binom{n}{k} \leqslant \frac{n^n}{k^k (n-k)^{(n-k)}} = 2^{\mathscr{H}(k/n)n}.$$

Lemma 4.1 appears as an exercise in [1]. The following lemma is also easy to verify.

Lemma 4.2. *For all $\varepsilon \in (0,1)$,*

$$\mathscr{H}(2^{n^\varepsilon - n})2^n = o(2^{\varepsilon n}).$$

We now show that only a p-dimension 0 set of languages are $\leqslant_m^p$-reducible to non-dense languages.

Theorem 4.3.

$$\dim_p(\mathrm{P_m}(\mathrm{DENSE}^c)) = 0.$$

Proof. Let $s>0$ be rational. It suffices to show that $\dim_{\rm p}({\rm P_m}({\rm DENSE}^c))\leqslant s$.

Let $\{(f_m,\varepsilon_m)\}_{m\in\mathbb{N}}$ be a standard enumeration of all pairs of polynomial-time computable functions $f_m:\{0,1\}^*\to\{0,1\}^*$ and rationals $\varepsilon_m\in(0,1)$. Define

$$A_{m,n}=\left\{u\in\{0,1\}^{2^{n+1}-1}\;\middle|\;\begin{array}{l}(\forall i,j\geqslant 2^{n/2})(f_m(s_i)=f_m(s_j)\Rightarrow u[i]=u[j])\\ \text{and } |\{f_m(s_i)\,|\,i\geqslant 2^{n/2}\text{ and } u[i]=1\}|\leqslant 2^{n^{\varepsilon_m}}\end{array}\right\}.$$

For each string u with $2^{n/2}\leqslant |u|\leqslant 2^{n+1}-1$, define the integers

$$\text{collision}_{m,n}(u)=|\{(i,j)|2^{n/2}\leqslant i<j<|u|,\ f_m(s_i)=f_m(s_j),\text{ and } u[i]\neq u[j]\}|,$$
$$\text{committed}_{m,n}(u)=\{f_m(s_i)|2^{n/2}\leqslant i<|u|\text{ and } u[i]=1\}|$$

and

$$\text{free}_{m,n}(u)=\{f_m(s_i)|\,|u|\leqslant i<2^{n+1}-1\}-\{f_m(s_i)|2^{n/2}\leqslant i<|u|\}|.$$

Then for each u with $|u|\geqslant 2^{n/2}$ there are

$$\text{count}_{m,n}(u)=\begin{cases}\displaystyle\sum_{i=0}^{2^{n^{\varepsilon_m}}-\text{committed}_{m,n}(u)}\binom{\text{free}_{m,n}(u)}{i} & \text{if collision}_{m,n}(u)=0,\\ 0 & \text{otherwise,}\end{cases}$$

strings v for which $uv\in A_{m,n}$.

Define for each $m,n\in\mathbb{N}$ a function $d_{m,n}:\{0,1\}^*\to[0,\infty)$ by

$$d_{m,n}(u)=\begin{cases}2^{(2-1)|u|} & \text{if } |u|<2^{n/2}\\ \frac{\text{count}_{m,n}(u)}{\text{count}_{m,n}(u[0..2^{n/2}-1])}2^{s|u|-2^{n/2}} & \text{if } 2^{n/2}\leqslant|u|\leqslant 2^{n+1}-1\\ 2^{(s-1)(|u|-2^{n+1}+1)}d(u[0..2^{n+1}-2]) & \text{otherwise.}\end{cases}$$

Then each $d_{m,n}$ is a well-defined s-gale because $\text{count}_{m,n}(u)=\text{count}_{m,n}(u0)+\text{count}_{m,n}(u1)$ for all u. Define a polynomial-time computable s-gale

$$d=\sum_{m=0}^{\infty}2^{-m}\sum_{n=0}^{\infty}2^{-n}d_{m,n}.$$

Let $A\leqslant_{\rm m}^{\rm P}D\in{\rm DENSE}^c$ by a reduction f running in time n^l. Let ε be a positive rational such that for infinitely many n, $|D_{\leqslant n^l}|<2^{n^\varepsilon}$. Let $m\in\mathbb{N}$ be such that $f_m=f$ and $\varepsilon_m=\varepsilon$. Using Lemmas 4.1 and 4.2, for each $u\in\{0,1\}^{2^{n/2}}$, we have

$$\begin{aligned}\text{count}_{m,n}(u)&\leqslant\sum_{i=0}^{2^{n^\varepsilon}}\binom{|f(\{0,1\}^{\leqslant n})|}{i}\\ &\leqslant(2^{n^\varepsilon}+1)\binom{2^{n+1}-1}{2^{n^\varepsilon}}\\ &\leqslant(2^{n^\varepsilon}+1)2^{\mathscr{H}(2^{n^\varepsilon-n})2^n}\end{aligned}$$

$$\leqslant 2^{2^{\varepsilon n}}$$
$$\leqslant 2^{s2^n - 2^{n/2} - 2n}$$

for all sufficiently large n. Whenever $|D_{\leqslant n^l}| < 2^{n^\varepsilon}$, we have $A[0..2^{n+1}-2] \in A_{m,n}$. Therefore for infinitely many n,

$$\begin{aligned} d(A[0..2^{n+1}-2]) &\geqslant 2^{-(m+n)} d_{m,n}(A[0..2^{n+1}-2]) \\ &= 2^{-(m+n)} \frac{\text{count}_{m,n}(A[0..2^{n+1}-2])}{\text{count}_{m,n}(A[0..2^{n/2}-1])} 2^{s(2^{n+1}-1)} - 2^{n/2} \\ &\geqslant 2^{-(m+n)} \frac{2^{s(2^{n+1}-1)-2^{n/2}}}{2^{s2^n - 2^{n/2} - 2n}} \\ &\geqslant 2^{n-m}. \end{aligned}$$

Therefore $A \in S^\infty[d]$. This shows that $\text{P}_\text{m}(\text{DENSE}^c) \subseteq S^\infty[d]$, from which it follows that $\dim_\text{p}(\text{P}_\text{m}(\text{DENSE}^c)) = 0$. □

5. Main theorem

Theorem 5.1. *If* $\dim_\text{p}(\text{NP}) > 0$*, then for all* $\varepsilon > 0$ *there exists a* $\delta > 0$ *such that any* 2^{n^δ}*-time approximation algorithm for* MAX3SAT *has performance ratio less than* $\frac{7}{8} + \varepsilon$ *on a dense set of satisfiable instances.*

Proof. We prove the contrapositive. Let $\varepsilon > 0$ be rational. For any MAX3SAT approximation algorithm $\mathscr{A}$, define the set

$$F_{\mathscr{A}} = \{x \in 3\text{SAT} \mid \mathscr{A}(x) < \tfrac{7}{8} + \varepsilon\}.$$

Assume that for each $\delta > 0$, there exists a 2^{n^δ}-time approximation algorithm $\mathscr{A}_\delta$ for MAX3SAT with $F_{\mathscr{A}_\delta} \in \text{DENSE}^c$. By Theorem 4.3 and Lemma 3.4, it is sufficient to show that $\text{NP} \subseteq \text{P}_\text{m}(\text{DENSE}^c) \uplus \text{DTIME}(2^n)$.

Let $B \in \text{NP}$ and let r be a $\leqslant_\text{m}^\text{p}$-reduction of B to SAT. Let n^k be an almost-everywhere time bound for computing $f_\varepsilon \circ r$ where f_ε is as in Theorem 2.1. Then

$$\begin{aligned} x \in B &\Leftrightarrow r(x) \in \text{SAT} \\ &\Leftrightarrow \text{MAX3SAT}((f_\varepsilon \circ r)(x)) = 1 \\ &\Leftrightarrow \mathscr{A}_{1/k}((f_\varepsilon \circ r)(x)) \geqslant \tfrac{7}{8} + \varepsilon \text{ or } (f_\varepsilon \circ r)(x) \in F_{\mathscr{A}_{1/k}}. \end{aligned}$$

Define the languages

$$C = \{x \mid (f_\varepsilon \circ r)(x) \in F_{\mathscr{A}_{1/k}}\} \quad \text{and} \quad D = \{x \mid \mathscr{A}_{1/k}((f_\varepsilon \circ r)(x)) \geqslant \tfrac{7}{8} + \varepsilon\}.$$

Then $B = C \cup D$, $C \leqslant_{\rm m}^{\rm p} F_{\mathscr{A}_{1/k}} \in \mathrm{DENSE}^c$, and D can be decided in time $2^{(n^k)^{1/k}} = 2^n$ for all sufficiently large n, so $B \in \mathrm{P_m}(\mathrm{DENSE}^c) \uplus \mathrm{DTIME}(2^n)$. □

Theorem 5.1 provides a strong positive answer to Problem 8 of Lutz and Mayordomo [8]:

Does $\mu_{\rm p}(\mathrm{NP}) \neq 0$ imply an exponential lower bound on approximation schemes for MAXSAT?

We observe that a weaker positive answer can be more easily obtained by using a simplified version of our argument to prove the following result.

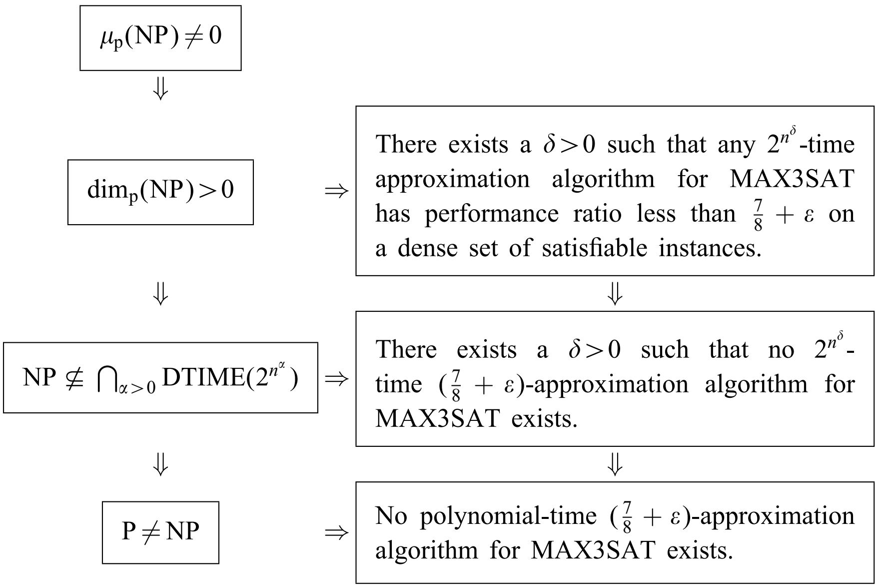
Proposition 5.2. *If*

$$\mathrm{NP} \not\subseteq \bigcap_{\alpha > 0} \mathrm{DTIME}(2^{n^\alpha}),$$

then for all $\varepsilon > 0$ *there exists a* $\delta > 0$ *such that there does not exist a* 2^{n^δ}*-time* $(\frac{7}{8} + \varepsilon)$*-approximation algorithm for* MAX3SAT.

6. Conclusion

We close by summarizing the inapproximability results for MAX3SAT derivable from various strong hypotheses in the following figure:



Acknowledgements

I thank Jack Lutz for some helpful suggestions.

References

[1] T.H. Cormen, C.E. Leiserson, R.L. Rivest, Introduction to Algorithms, MIT Press, McGraw-Hill, Cambridge, MA, New York, 1990.

[2] E. Dantsin, M. Gavrilovich, E.A. Hirsch, B. Konev, MAX SAT approximation beyond the limits of polynomial-time approximation, Ann. Pure Appl. Logic, to appear.

[3] E. Halperin, U. Zwick, Approximation algorithms for MAX 4-SAT and rounding procedures for semidefinite programs, IPCO: 7th Integer Programming and Combinatorial Optimization Conference, Graz, Austria, 1999.

[4] J. Håstad, Some optimal inapproximability results, Proc. 29th Ann. ACM Symp. on Theory of Computing, 1997, pp. 1–10.

[5] H. Karloff, U. Zwick, A 7/8-approximation algorithm for MAX 3SAT? Proc. 38th Ann. Symp. on Foundations of Computer Science, 1997, pp. 406–415.

[6] J.H. Lutz, Almost everywhere high nonuniform complexity, J. Comput. System Sci. 44 (1992) 220–258.

[7] J.H. Lutz, Dimension in complexity classes, in: Proc. 15th Ann. IEEE Conf. Computational Complexity, IEEE Computer Society Press, Los Alamitos, CA, 2000, pp. 158–169.

[8] J.H. Lutz, E. Mayordomo, Twelve problems in resource-bounded measure, Bull. European Assoc. Theoret. Comput. Sci. 68 (1999) 64–80.

[9] J.H. Lutz, E. Mayordomo, Measure, stochasticity, and the density of hard languages, SIAM J. Comput. 23 (4) (1994) 762–779.