Reliable Reductions, High Sets and Low Sets

V. Arvind*
Department of Computer Science and Engineering
Indian Institute of Technology, Delhi
New Delhi 110016, India

J. Köbler† and M. Mundhenk†
Abteilung für Theoretische Informatik
Universität Ulm
Oberer Eselsberg
D-W-7900 Ulm, Germany

Abstract

Measuring the information content of a set by the space-bounded Kolmogorov complexity of its characteristic sequence, we investigate the (non-uniform) complexity of sets A in EXPSPACE/poly that reduce to some set having very high information content. Specifically, we show that if the reducibility used has a certain property, called "reliability," then A in fact is reducible to a sparse set (under the same reducibility). As a consequence, the existence of hard sets (under "reliable" reducibilities) of very high information content is unlikely for various complexity classes as for example NP, PP, and PSPACE.

1 Introduction

An important subject of research in structural complexity theory is the study of reductions to sets of low information content (as for example sparse sets). Historically, this study originated in the Berman-Hartmanis conjecture that all NP-complete sets are polynomial-time isomorphic, a hypothesis that is empirically supported by all natural examples of NP-complete sets. An intuitive interpretation of the conjecture is that all NP-complete sets are only different encodings of the same information. Since polynomial-time isomorphic sets must have similar densities, and since there are NP-complete sets of exponential density, the Berman-Hartmanis conjecture implies that sparse sets cannot be NP-complete.

Since then, several results concerning reductions of NP-complete sets to sparse sets have been established. The basic results along this line of research were Mahaney's theorem that if any NP-complete set many-one reduces to a sparse set then P = NP [Mah82], and the result of Karp, Lipton, and Sipser, that if NP has sparse Turing-hard

*Work done while visiting Universität Ulm. Supported in part by an Alexander von Humboldt research fellowship.

†Work supported in part by the DAAD through Acciones Integradas 1992, 313-AI-e-es/zk

sets then the polynomial hierarchy collapses to Σ_2^p [KL80]. The more recent result of Ogiwara and Watanabe [OW91], that if NP has sparse hard sets under polynomial time bounded truth-table reductions then P = NP, has been followed up by analogous results for more general reductions in [AHH+92, RR92, AKM92a].

Recently, Book and Lutz [BL92] have considered reductions to sets whose characteristic sequences are of very high space-bounded Kolmogorov complexity (they call the class of such sets HIGH). Book and Lutz obtained the surprising result that every set in ESPACE that is (polynomial-time) bounded truth-table reducible to a set in HIGH is actually bounded truth-table reducible to some sparse set. Thus, oracles of very high space-bounded Kolmogorov complexity are only as useful as sparse sets provided that the oracle access is restricted to a constant number of queries to retrieve the information from the oracle set. As a consequence, if SAT bounded truth-table reduces to a set in HIGH, then SAT bounded truth-table reduces to a sparse set, and thus, using the result of Ogiwara and Watanabe, it follows that P = NP.

Of course, one would expect similar results for a wider class of reducibility types, not only for the bounded truth-table case. In the present paper, we prove analogous results for a wide range of reducibilities which we call "reliable reducibilities". Intuitively, A reduces to B via a reliable reducibility if not only A is determined by B (and the specific reduction used), but also information about B can be retrieved from the knowledge of A. More precisely, in the case of a reliable reduction, given A and the reduction procedure, it is possible to compute a *minimum* set B' of strings to which A reduces via the given reduction, that is, B' is a subset of every set to which A reduces, or equivalently, $B' \subseteq B$.

For example, deterministic many-one reductions f are reliable, since for every $x \in A$ the membership of the image $f(x)$ in B is guaranteed, and A correctly reduces to the set $B' = \{f(x) \mid x \in A\}$ of all these images (a similar consideration applies for conjunctive reductions). In contrast, Turing reductions are not considered to be reliable since in general the information about B that can be derived from the knowledge of A and the oracle machine performing the reduction is rather vague. On the other hand, as we show in this paper, the Hausdorff reducibility [Wag90] is reliable. In spite of its reliability, the Hausdorff reducibility is much more powerful than the many-one reducibility, and in some cases it has been shown that the Hausdorff reducibility has the same power as the truth-table reducibility. For example, Wagner [Wag90] proved that the closure of NP under the Hausdorff reducibility coincides with the truth-table closure of NP. Furthermore, for every set ring $\mathcal{C}$ it is known (cf. [WW85]) that the closure of $\mathcal{C}$ under bounded Hausdorff reducibility is the same as its closure under bounded truth-table reducibility (as an example we mention the equality $R^p_{bhd}(R^p_c(\text{SPARSE})) = R^p_{btt}(R^p_c(\text{SPARSE}))$ [AKM92b], where $R^p_c(\text{SPARSE})$ is the closure of sparse sets under conjunctive reducibility). In fact, it turns out that even an extension of the Hausdorff reducibility (called non-monotone Hausdorff reducibility) is reliable.

The main result of the paper states that if a set A in EXPSPACE/poly reduces via a reliable reducibility to some set in HIGH, then A reduces (via the same reducibility) to a sparse set. As a consequence of this result, it is unlikely that there exist hard sets (under the non-monotone Hausdorff and other reliable reducibilities) for complexity classes like UP, NP, PP, or PSPACE.

The paper is organized as follows. Section 2 introduces notation and gives basic definitions. In Section 3 we introduce the concept of a reliable reducibility and prove the reliability of various reducibilities, as for example the (non-monotone) Hausdorff reducibil-

ities, the composed bounded Hausdorff and conjunctive reducibilities (henceforth called *bhd-c* reducibility), the *co-np* many-one reducibility, and the *co-rp* many-one reducibility. Finally, in Section 4, we prove that no set A in EXPSPACE/poly reduces via a reliable reduction to a set B in HIGH unless A is reducible via a reduction of the same type to a sparse set, and use this result to demonstrate the unlikeliness of the existence of hard HIGH sets for various complexity classes under the reducibilities considered in Section 3.

2 Preliminaries

Let $\Sigma = \{0,1\}$ be the standard alphabet, and $A \subseteq \Sigma^*$ be a set. $A^{=n}$ ($A^{\leq n}$) denotes the set of all strings in A of length n (up to length n, respectively). χ_A denotes the characteristic function of A. $\chi^A_{\leq n}$ denotes the characteristic sequence of A for all strings up to length n, i.e. the i-th bit of $\chi^A_{\leq n}$ equals $\chi_A(s_i)$ where s_i is the i-th string in Σ^* in lexicographical order. The cardinality of A is denoted by $|A|$. The census function of a set A is $census_A(1^n) = |A^{\leq n}|$. A set S is called sparse if its census function is bounded above by a polynomial. A set T is called a tally set if $T \subseteq 0^*$. We use TALLY and SPARSE to represent the classes of tally and sparse sets, respectively. The empty string is denoted by ε. $\langle \cdot, \cdot \rangle$ denotes a standard polynomial time computable pairing function whose inverses are also computable in polynomial time.

For any reducibility type α, let $R_\alpha(\mathcal{C}) = \{A \mid \exists B \in \mathcal{C} : A \leq_\alpha B\}$ denote the class of sets $\leq_\alpha$-reducible to some set in $\mathcal{C}$ [BK88, AHOW92]. The reducibilities discussed in this paper are the standard polynomial-time reducibilities defined by Ladner, Lynch, and Selman [LLS75] and the following Hausdorff reducibility introduced by Wagner [Wag87].

Definition 2.1 *A is Hausdorff reducible to B (in symbols: $A \leq^p_{hd} B$), if there exists a polynomial-time computable function f mapping every string x to a sequence of queries, such that for all $x \in \Sigma^*$, if $f(x) = \langle y_1, \ldots, y_k \rangle$ then*

- *$y_{i+1} \in B$ implies $y_i \in B$ for all $i = 1, \ldots, k-1$, and*
- *$x \in A \iff \max\{j \mid 0 \leq j \leq k$ and for all $i = 1, \ldots, j : y_i \in B\}$ is odd.*

We call f a bounded Hausdorff reduction ($A \leq^p_{bhd} B$) if the number $k(x)$ of queries produced by f on x is bounded by a constant for all x.

In this context, the i-th query y_i computed by $f(x)$, $1 \leq i \leq k(x)$, is also denoted by $f(x,i)$. The following theorem gives two examples where the (bounded) Hausdorff reducibilities lead to the same reduction classes as the (bounded) truth-table reducibilities.

Theorem 2.2

i) [AKM92b] $R^p_{bhd}(R^p_c(\text{SPARSE})) = R^p_{btt}(R^p_c(\text{SPARSE}))$,

ii) [Wag90] $R^p_{hd}(\text{NP}) = R^p_{tt}(\text{NP})$ $(= \Theta^p_2)$.

By removing in the definition of the Hausdorff reducibility the monotony condition we obtain the following generalization of the Hausdorff reducibility which, as we will see in Proposition 2.4, is equivalent in power to the composition of the Hausdorff and the conjunctive reducibilities.

Definition 2.3 *A is non-monotone Hausdorff reducible to B ($A \leq^p_{nhd} B$), if there exists a polynomial-time computable function f mapping every string x to a sequence of queries, such that for all $x \in \Sigma^*$, if $f(x) = \langle y_1, \ldots, y_k \rangle$ then*

$$x \in A \iff \max\{j \mid 0 \leq j \leq k \text{ and for all } i = 1, \ldots, j : y_i \in B\} \text{ is odd.}$$

Proposition 2.4 *For every class $\mathcal{C}$ of languages, $R^p_{nhd}(\mathcal{C}) = R^p_{hd}(R^p_c(\mathcal{C}))$.*

Proof Assume that $A \leq^p_{nhd} B$ via a function $f \in \mathrm{FP}$. Then let B' be the set defined as

$$B' = \{\langle y_1, \ldots, y_i \rangle \mid y_j \in B \text{ for all } j = 1, \ldots, i\}$$

which clearly is in $R^p_c(B)$. Furthermore, $A \leq^p_{hd} B'$ via the FP function f' that produces the same number of queries as f, and whose i-th query consists of the sequence $f'(x, i) = \langle f(x, 1), \ldots, f(x, i) \rangle$ of the first i many queries of $f(x)$.

Conversely, assume that $A \leq^p_{hd} B$ via some function $f \in \mathrm{FP}$, and that $B \leq^p_c C$ via an FP function g, i.e., for all x, g on x computes a set (appropriately encoded) such that $x \in B$ if and only if $g(x) \subseteq C$. Consider the FP function h defined as follows. For all x, if $f(x) = \langle y_1, \ldots, y_k \rangle$, and $g(y_i) = \{y^i_1, \ldots, y^i_{m_i}\}$ for $i = 1, \ldots, k$, then

$$h(x) = \langle y^1_1, y^1_1, \ldots, y^1_{m_1 - 1}, y^1_{m_1 - 1}, y^1_{m_1}, \ \ldots \ , y^k_1, y^k_1, \ldots, y^k_{m_k - 1}, y^k_{m_k - 1}, y^k_{m_k} \rangle$$

It is easy to see that the parity of the maximum index i such that all queries $f(x, j)$, $1 \leq j \leq i$, are answered positively by B and the parity of the maximum index l such that all queries $h(x, j)$, $1 \leq j \leq l$, are answered positively by C are equal, and thus $A \leq^p_{nhd} C$ via h. ∎

As an immediate consequence of Proposition 2.4 we have the following characterizations (observing that NP is closed under conjunctive reductions).

Proposition 2.5

i) $R^p_{nhd}(\mathrm{SPARSE}) = R^p_{hd}(R^p_c(\mathrm{SPARSE}))$.

ii) $R^p_{nhd}(\mathrm{NP}) = R^p_{tt}(\mathrm{NP})$ $(= \Theta^p_2)$.

Very recently, Buhrman, Longpré and Spaan [BLS92], proved that every sparse set is conjunctively reducible to a tally set, that is, $R^p_c(\mathrm{SPARSE}) = R^p_c(\mathrm{TALLY})$. Therefore we can state the following equality.

Proposition 2.6 $R^p_{nhd}(\mathrm{TALLY}) = R^p_{nhd}(\mathrm{SPARSE})$.

We now define nondeterministic and random many-one reductions. The notion of $\leq^{co\text{-}np}_m$ reducibility (cf. [AKM92b]) can be seen as a generalization of the deterministic polynomial-time conjunctive reducibility.

Definition 2.7 *A set A is co-np many-one reducible to a set B (denoted $A \leq^{co\text{-}np}_m B$) if there exists a polynomial-time nondeterministic Turing transducer M such that for every $x \in \Sigma^*$, $x \in A$ if and only if all outputs of M on input x are members of B.*

Note that $A \leq_m^{co\text{-}np} B$ if and only if $\overline{A} \leq_m^{np} \overline{B}$ where $\leq_m^{np}$ is the more familiar polynomial-time nondeterministic many-one reducibility [LLS75]. Let q be a polynomial bounding the running time of the transducer in the above definition. Then every computation path of M on input x can be described by a string w of length $q(|x|)$, and it is easy to see that the output generated on that path can be determined in polynomial time from x and w. Thus, $A \leq_m^{co\text{-}np} B$ if and only if there exist a polynomial-time computable function h and a polynomial q such that for all x,

$$x \in A \Leftrightarrow \forall w \in \Sigma^{q(|x|)} : h(x,w) \in B$$

A special case of the $\leq_m^{co\text{-}np}$ reducibility is the *co-rp* many-one reducibility (cf. [AM77, VV86, Ro92, AKM92a]).

Definition 2.8 *A set A is co-rp many-one reducible to a set B (denoted $A \leq_m^{co\text{-}rp} B$) if there exist a polynomial-time computable function h and a polynomial q such that*

$$x \in A \Rightarrow \mathrm{Prob}_{w \in \Sigma^{q(|x|)}}[h(x,w) \in B] = 1$$

$$x \notin A \Rightarrow \mathrm{Prob}_{w \in \Sigma^{q(|x|)}}[h(x,w) \notin B] \geq 3/4$$

Here, the string w is chosen uniformly at random from the set $\Sigma^{q(|x|)}$.

Let M be a Turing machine, z be a string and let d, s be natural numbers. We say that $z \in KS_M[d,s]$, if M on some input of length at most d outputs z using space at most s. In other words, $KS_M[d,s]$ is the set of strings whose *s-space-bounded Kolmogorov complexity* relative to M is bounded by d. Similarly, for a string y, $KS_M[d,s|y]$ is the set of all strings z for which there is a string x of length at most d such that M on input $\langle x,y \rangle$ outputs z using space at most s. Well known simulation-techniques (see [LV92]) show that there is a Universal Turing machine U such that for every machine M there is a constant c such that for all d, s: $KS_M[d,s] \subseteq KS_U[d+c, cs+c]$. Henceforth, we fix such a Universal Turing machine and omit the subscript. Note that there is a constant c such that for every set A and for all n, the characteristic sequence $\chi^A_{\leq n}$ of A restricted to the set of strings up to length n is in $KS[2^{n+1}+c, 2^{cn}]$.

Finally, we define the class HIGH containing only sets of very high information content. (We say that a property holds "for almost every n" if it holds for all but finitely many n.)

Definition 2.9 *A set B is in* HIGH*, if for every constant $c > 0$ there exists a polynomial q such that for almost every n, $\chi^B_{\leq n} \notin KS[2^{n+1} - q(n), 2^{cn}]$.*

Observe that our definition of the class HIGH is an extension of the original one given by Book and Lutz [BL92] who required for q the fixed polynomial $2n$. It is easy to see that for every set A in the class EXPSPACE/poly there is a polynomial p such that for every n, $\chi^A_{\leq n} \in KS[p(n), 2^{p(n)}]$, where EXPSPACE = DSPACE($2^{n^{O(1)}}$).

For further definitions used in this paper we refer the reader to standard books on structural complexity theory (for example [BDG]).

3 Reliable reducibilities

In this section, we introduce the concept of a reliable reducibility and prove the reliability of various reducibilities, as for example the (non-monotone) Hausdorff reducibility, the composition of the bounded Hausdorff and the conjunctive reducibilities (henceforth called *bhd-c* reducibility), and the *co-np* as well as the *co-rp* many-one reducibilities.

Definition 3.1 *A reducibility $\leq_\alpha$ is reliable if for all sets A, B such that $A \leq_\alpha B$ there is a sequence $B_n \subseteq B$, $n \geq 0$, of sets fulfilling the following properties:*

- *All instances for A up to length n are correctly reduced (via the given reduction) to B_n,*
- *There exists a polynomial p such that for every $n \geq 0$, $\chi^{B_n}_{\leq n} \in KS\left[p(n), 2^{p(n)} \,\middle|\, \chi^{A}_{\leq n}\right]$.*

We start by proving the reliability of the composed Hausdorff and conjunctive reducibilities.

Theorem 3.2 *The composition of the Hausdorff reducibility and the conjunctive reducibility is reliable.*

Proof Let $A \leq^p_{hd} B$ via a polynomial time computable function f, and let B conjunctively reduce to C via a polynomial time computable function g. Let $k(x)$ be the number of queries in the list $f(x)$. The following algorithm computes on input $\chi^A_{\leq n}$ the characteristic string $\chi^{C_n}_{\leq n}$ of a finite set $C_n \subseteq C$ such that A reduces to C_n via f and g for all instances up to length n. In fact, C_n is the smallest set with this property.

input $\chi^A_{\leq n}$
$x := \varepsilon$
$C_n := \emptyset$
repeat
  $l := \max(\{0\} \cup \{j \mid 1 \leq j \leq k(|x|) \text{ and } g(f(x,j)) \subseteq C_n\})$
  if $\bigcup\{g(f(x,j)) \mid 1 \leq j \leq l\} \not\subseteq C_n$ **then**
    $C_n := C_n \cup \bigcup\{g(f(x,j)) \mid 1 \leq j \leq l\}$
    $x := \varepsilon$
  elsif $x \in A \iff l$ is even **then**
    $C_n := C_n \cup g(f(x, l+1))$
    $x := \varepsilon$
  else $x := succ(x)$ (* in lexicographical order *)
  end
until $x = 0^{n+1}$
output $\chi^{C_n}_{\leq n}$

We have to show that C_n is a subset of C and that C_n can be used in the composed reduction from A to C instead of C for all instances up to length n:

i) C_n is a subset of C for all $n \geq 0$.

ii) For all $n \geq 0$, A reduces to C_n via f and g for all instances $x \in \Sigma^{\leq n}$.

To prove i) we proceed by induction on the number of iterations of the repeat-loop. Before the repeat-loop starts, the set C_n is empty, and therefore $C_n \subseteq C$. Assume that $C_n \subseteq C$ after the $(m-1)$-st iteration, and that C_n is extended by some set $g(f(x,j))$ during the m-th iteration. There are two cases. In the first case there exists an index $l > j$ such that $g(f(x,l)) \subseteq C_n$, and since $C_n \subseteq C$, this implies $g(f(x,j)) \subseteq C$ by the monotony of the Hausdorff reduction.

In the second case it holds that $j = l+1$ where $l = \max\{j \mid 0 \leq j \leq k(|x|)$ and for all $i = 1,\ldots,j : g(f(x,i)) \subseteq C_n\}$ is even if and only if x is in A. It is clear that $l < k(x)$ since $C_n \subseteq C$. By way of a contradiction, assume $g(f(l+1,x)) \not\subseteq C$. Because $C_n \subseteq C$, this contradicts the fact that A reduces to C via f and g since it would imply that $x \in A \iff \max\{j \mid 0 \leq j \leq k(|x|)$ and for all $i = 1,\ldots,j : g(f(x,i)) \subseteq B\}$ is even.

To see ii), observe that every time when x is reset to the empty string inside the repeat-loop then the set C_n is extended by at least one new query whose length is polynomially bounded in n. Thus, it follows that the algorithm terminates and outputs some set C_n. But at the time when the algorithm stops, there is no $x \in \Sigma^{\leq n}$ left for which one of the two if-conditions is fulfilled. From this observation the validity of ii) follows immediately.

To conclude the proof of the theorem observe that $|C_n| = 2^{O(n)}$, and therefore the algorithm can be performed in time $2^{O(n)}$. As a consequence, it is easy to see that there exists a constant c such that for every n, $\chi^{C_n}_{\leq n} \in KS\left[c, 2^{cn} \,\middle|\, \chi^{A}_{\leq n}\right]$. ∎

We note that the reliability of the non-monotone Hausdorff reducibility can be obtained by a minor simplification of the above proof. Further, since the many-one, conjunctive, (bounded) Hausdorff, and *bhd-c* reducibilities all are special cases of the composed Hausdorff and conjunctive reducibilities, we can state the following corollary.

Corollary 3.3 *The many-one, conjunctive, (bounded) Hausdorff, and bhd-c reducibilities as well as the non-monotone Hausdorff reducibility are reliable.*

We proceed by proving the reliability of the *co-np* many-one and the *co-rp* many-one reducibilities.

Theorem 3.4 *The co-np many-one and the co-rp many-one reducibilities are reliable.*

Proof Let $A \leq_m^{co\text{-}np} B$ via a polynomial time nondeterministic Turing transducer M, that is, for every $x \in \Sigma^*$, $x \in A$ if and only if all outputs of M on input x are in B. Let B_n be the set defined as

$$B_n = \{y \mid \exists x \in A^{\leq n} : M(x) \text{ outputs } y\}$$

Then B_n is a subset of B to which A reduces via M for all instances up to length n. Furthermore, it is easy to see that there exists a constant c and a polynomial p such that for every n, $\chi^{B_n}_{\leq n} \in KS\left[c, 2^{p(n)} \,\middle|\, \chi^{A}_{\leq n}\right]$. The proof for the *co-rp* many-one reducibility proceeds analogously. ∎

At the end of this section we show that also the composed Hausdorff, *co-rp* many-one, and conjunctive reducibilities are reliable.

Theorem 3.5 *The reduction obtained by composing the Hausdorff, the co-rp many-one, and the conjunctive reducibilities is reliable.*

Proof Let $A \leq^p_{hd} B \leq^{co\text{-}rp}_m C \leq^p_c D$ via polynomial time computable functions f, h, and g, respectively. Let $k(x)$ be the number of queries in the list $f(x)$, and let q be the polynomial associated with h. Then for all $x \in \Sigma^*$,

- $f(x, i+1) \in B$ implies $f(x,i) \in B$ for all $i = 1, \ldots, k-1$, and
- $x \in A \iff \max\{j \mid 0 \leq j \leq k(x) \text{ and for all } i = 1, \ldots, j : f(x,i) \in B\}$ is odd,

and for all $y \in \Sigma^*$,

$$y \in B \Rightarrow \text{Prob}[h(x,w) \in C] = 1,$$

$$y \notin B \Rightarrow \text{Prob}[h(x,w) \notin C] \geq 3/4.$$

where w is chosen uniformly at random from $\Sigma^{q(|y|)}$. Third, for all $z \in \Sigma^*$, $z \in C$ if and only if $g(z) \subseteq D$. Consider the following algorithm.

input $\chi^A_{\leq n}$
$x := \varepsilon$
$D_n := \emptyset$
repeat
  $I := \{j \mid 1 \leq j \leq k(|x|) \text{ and } 1/4 < \text{Prob}[h(f(x,j),w) \in C] < 1\}$
  if $I \neq \emptyset$ **then**
    $D_n := D_n \cup \bigcup\{g(h(f(x,j),w)) \mid j \in I, |w| = q(|f(x,j)|)\}$
    $x := \varepsilon$
  else
    $l := \max(\{0\} \cup \{j \mid 1 \leq j \leq k(|x|) \text{ and } \forall w, g(h(f(x,j),w)) \subseteq D_n\})$
    if $\bigcup\{g(h(f(x,j),w)) \mid 1 \leq j \leq l, |w| = q(|f(x,j)|)\} \not\subseteq D_n$ **then**
      $D_n := D_n \cup \bigcup\{g(h(f(x,j),w)) \mid 1 \leq j \leq l, |w| = q(|f(x,j)|)\}$
      $x := \varepsilon$
    elsif $x \in A \iff l$ is even **then**
      $D_n := D_n \cup \bigcup\{g(h(f(x,l+1),w)) \mid |w| = q(|f(x,l+1)|)\}$
      $x := \varepsilon$
    else $x := succ(x)$ (* in lexicographical order *)
    end
  end
until $x = 0^{n+1}$
output $\chi^{D_n}_{\leq n}$

Analogously to the proof of Theorem 3.2 it can be seen that D_n is a subset of D that can be used instead of D for all instances up to length n in the composed Hausdorff, *co-rp* many-one, and conjunctive reduction of A to D, and that there exist a constant c and a polynomial p such that for every n, $\chi^{D_n}_{\leq n} \in KS\left[c, 2^{p(n)} \,\middle|\, \chi^A_{\leq n}\right]$. ∎

4 On reliable reductions to HIGH sets

In this section we prove that no set A in EXPSPACE/poly reduces via a reliable reduction to a set B in HIGH unless A is reducible via a reduction of the same type to a sparse set. In order to give an intuitive explanation how the proof works, consider the case that A reduces to B via a many-one reduction function f. Since the set A is of relatively low space-bounded Kolmogorov complexity, membership in B can be decided for all the instances in the range of f using only a relatively small amount of resources (assuming that f is honest). Therefore, the range of f cannot be too large since otherwise a substantial part of the characteristic sequence of B would contain only little information, contradicting the fact that B is in HIGH.

For the proof we need two lemmas that are of independent interest. Intuitively spoken, the following lemma shows that the Kolmogorov complexity of a string cannot be very high, if "many" 1's of the string are easily computable. For $b \in \{0,1\}$ let $\#_b(y)$ be the number of bits equal to b in the string y. Further, let $\preceq$ be the following partial ordering on strings defined by $a_1 \ldots a_k \preceq b_1 \ldots b_l$, if $k \leq l$ and $a_i \leq b_i$, for $i = 1, \ldots, k$.

Lemma 4.1 *There exists a constant c^* such that for all $c, d \in \mathbb{N}$ and for all $x, y \in \Sigma^*$, if $x \preceq y$ and $x \in KS[d, c]$, then $y \in KS[2\log|d| + d + |y| - \#_1(x) + c^*, \max(c, c^* \cdot |y|)]$.*

Proof Consider the following machine M that on input $\langle v, w\rangle$ where w is a string of length $|y| - \#_1(x)$ and v is a description of x, at first produces the string x (from its description v), and then outputs the string y obtained from x by replacing the i-th 0 in x by the i-th bit of w.

input $\langle v, w\rangle$
compute x from the description v
(* the following loop computes the output string y *)
$j := 1$; $y := \varepsilon$
for $i := 1$ **to** $|w| + \#_1(x)$ **do**
  if $i \leq |x|$ **and** the i-th bit of x is 1 **then**
    $y := y1$
  else
    $y := yb$ where b is the j-th bit of w
    $j := j + 1$
  end
end
output y

Letting $\langle \cdot, \cdot \rangle$ be the self-delimiting encoding scheme (cf. [LV92]), we have that $|\langle v, w\rangle| = 2\log|v| + 2 + |v| + |w|$ for all $v, w \in \Sigma^*$. Furthermore, since for every pair of strings x, y with $x \preceq y$ there exists a string w of length $|y| - \#_1(x)$ such that for an appropriate description v of x, $|v| \leq d$, M on input $\langle v, w\rangle$ outputs y, the claim follows easily by the properties of the self-delimiting encoding scheme and the simulation properties of our fixed Universal Turing machine. ∎

The second lemma that we need to establish our main result states that if in the reduction of a set A to some set B the number of positively answered queries on instances of length n is polynomially bounded, then the given reduction can be modified (by padding the queries) to yield a reduction from A to a sparse set.

Lemma 4.2 *Assume that a set A reduces via a given reduction to some set. If there is a polynomial r such that for every $n \geq 0$ there is a set S_n of cardinality $|S_n| \leq r(n)$ such that all length n instances for A are correctly reduced to S_n, then A reduces via the same type of reduction (that is, only the queries need to be padded) to a sparse set $\hat{S}$.*

Proof Assume that there is a (non-decreasing) polynomial r such that for all n, $|S_n| \leq r(n)$. Define the set $\hat{S} = \{y10^n \mid y \in S_n\}$. Since for every $n \geq 0$ and all length n instances, A reduces correctly to the set S_n, it follows that A reduces to $\hat{S}$ via the reduction obtained by replacing in the given reduction every query y on input x by the query $y10^{|x|}$. Since $census_{\hat{S}}(1^n) = \Sigma_{i=0}^{n-1} census_{S_i}(1^{n-i-1}) \leq n \cdot r(n)$ for all n, it follows that $\hat{S} \in$ SPARSE. ■

Now we are ready to prove the main result of this section. It states that in order to decide a set in EXPSPACE/poly, all the information that a reliable reduction is able to extract out of a set containing a lot of information can be provided just as well by a sparse set (which is of very low information content).

Theorem 4.3 *If a set $A \in$ EXPSPACE/poly reduces via a reliable reduction to a set $B \in$ HIGH, then A reduces via the same type of reduction (that is, only the queries need to be padded) to a sparse set.*

Proof Let $A \in$ EXPSPACE/poly, $B \in$ HIGH and assume that A reduces to B via a reliable reduction. Intuitively, using the reliability of the reduction and the fact that $A \in$ EXPSPACE/poly, it is possible to compute within the resource bounds provided by the complexity class EXPSPACE/poly a sequence of sets $B_n \subseteq B$ such that for all instances up to length n, A reduces correctly to B_n. Since $B \in$ HIGH, the number of 1's in the the characteristic sequences of the sets B_n cannot be large, and therefore, as we will see, A reduces to a sparse set. In the sequel we give the formal proof.

Since $A \in$ EXPSPACE/poly, there exists a polynomial p such that for every n, $\chi^A_{\leq n} \in KS[p(n), 2^{p(n)}]$. Also, since A reduces to B via a reliable reduction, there are a polynomial q and a sequence $B_n \subseteq B$, $n \geq 0$, of sets such that for every $n \geq 0$,

- all instances for A up to length n are correctly reduced (via the given reduction) to B_n, and
- $\chi^{B_n}_{\leq n} \in KS\left[q(n), 2^{q(n)} \,\middle|\, \chi^A_{\leq n}\right]$.

As an immediate consequence of the reliability of the reduction from A to B and the fact that $A \in$ EXPSPACE/poly we have

$$\chi^{B_n}_{\leq n} \in KS[p(n), 2^{p(n)}] \tag{1}$$

for some polynomial p and every $n \geq 0$. Next we show that there is a polynomial q such that

$$census_{B_n}(1^n) \leq q(n) \tag{2}$$

for all n. Assume otherwise, then there exist for every polynomial q infinitely many n for which $census_{B_n}(1^n) > q(n)$. But since $B_n \subseteq B$ it follows by (1) above and by Lemma 4.1 (letting $x = \chi^{B_n}_{\leq n}$ and $y = \chi^B_{\leq p(n)}$) that there is a constant c^* such that for every polynomial q there are infinitely many n for which

$$\chi^B_{\leq p(n)} \in KS[2\log|p(n)| + p(n) + 2^{p(n)} - q(n) + c^*, c^* \cdot 2^{p(n)}]$$

contradicting the fact that $B \in \text{HIGH}$.

To complete the proof of the theorem let r be a (non-decreasing) polynomial bounding the length of the queries of the given reliable reduction from A to B. Then for all length n instances, A reduces correctly to the set $B^{\leq r(n)}_{r(n)}$. By (2) above it follows that $census_{B_{r(n)}}(1^{r(n)}) \leq q(r(n))$ for all n, and thus the theorem follows by Lemma 4.2. ∎

As an easy consequence of the preceding theorem and the reliability results of Section 3 we can state the following corollary.

Theorem 4.4 *Let A be in* EXPSPACE/poly *and let α be one of the following reducibility types:*

- *non-monotone Hausdorff,*
- *bhd-c,*
- *co-np many-one,*
- *composed bhd, co-rp many-one, and conjunctive reductions.*

Then A is in R_α(HIGH) if and only if A is in R_α(SPARSE).

Proof Since for every tally set T there exists a set $B \in \text{HIGH}$ such that $B \cap 0^* = T$, it follows that $\text{TALLY} \subseteq R^p_m(\text{HIGH})$. Furthermore, using the fact that $R^p_c(\text{SPARSE}) = R^p_c(\text{TALLY})$ [BLS92], it follows that $R_\alpha(\text{SPARSE}) \subseteq R_\alpha(R^p_c(\text{TALLY})) \subseteq R_\alpha(R^p_c(\text{HIGH})) \subseteq R_\alpha(\text{HIGH})$ for every reducibility type α considered here. The backward direction follows by Theorem 4.3 and the reliability of the considered reducibilities proved in the preceding section. ∎

By Theorem 4.3 we know that the existence of hard sets in HIGH for any complexity class $\mathcal{C} \subseteq \text{EXPSPACE/poly}$ with respect to a reliable reducibility implies the existence of a sparse hard set for $\mathcal{C}$ with respect to that reducibility. Therefore the existence of hard sets in HIGH leads to the same consequences as the existence of sparse hard sets.

Corollary 4.5 *Let $\mathcal{C}$ be any of the complexity classes from* $\{\text{UP}, \text{NP}, \text{C}_=\text{P}, \text{PP}\}$. *If $\mathcal{C} \subseteq R^p_{bhd}(R^p_c(\text{HIGH}))$, then $\mathcal{C} = \text{P}$.*

Proof This is a direct consequence of Theorem 4.4 and the result that $\mathcal{C} \subseteq R^p_{bhd}(R^p_c(\text{SPARSE}))$ implies $\mathcal{C} = \text{P}$ [AKM92a]. ∎

Also, since the existence of a sparse hard set for NP (PSPACE) with respect to the non-monotone Hausdorff reducibility implies the collapse of the polynomial hierarchy to Δ^p_2 ($\text{PSPACE} = \Delta^p_2$, respectively) [AKM], we have the following corollary.

Corollary 4.6 *For $\mathcal{C} \in \{\text{NP}, \text{PSPACE}\}$, if $\mathcal{C} \subseteq R^p_{nhd}(\text{HIGH})$, then $\mathcal{C}$ is low for Δ^p_2.*

Finally, applying the result that NP is not contained in $R^p_{bhd}(R^{co\text{-}rp}_m(R^p_c(\text{SPARSE})))$ unless $\text{NP} = \text{RP}$ [AKM92a], we get

Corollary 4.7 *If* NP *is contained in $R^p_{bhd}(R^{co\text{-}rp}_m(R^p_c(\text{HIGH})))$, then* $\text{NP} = \text{RP}$.

Acknowledgements

The authors thank Montse Hermo and Elvira Mayordomo for helpful comments.

References

[AM77] L. Adleman and K. Manders. Reducibility, randomness, and intractability. *Proc. 9th ACM Symp. on Theory of Computing*, 1977, 151-163.

[AHH+92] V. Arvind, Y. Han, L. Hemachandra, J. Köbler, A. Lozano, M. Mundhenk, M. Ogiwara, U. Schöning, R. Silvestri, and T. Thierauf. Reductions to sets of low information content. *Proceedings of the 19th International Colloquium on Automata, Languages, and Programming*, Lecture Notes in Computer Science #623:162-173, Springer Verlag, 1992.

[AHOW92] E. Allender, L. Hemachandra, M. Ogiwara, and O. Watanabe. Relating equivalence and reducibility to sparse sets. *SIAM Journal on Computing*, 21(3):521–539, 1992.

[AKM92a] V. Arvind, J. Köbler, and M. Mundhenk. On bounded truth-table, conjunctive, and randomized reductions to sparse sets. In *Proceedings 12th Conference on the Foundations of Software Technology & Theoretical Computer Science,* 1992.

[AKM92b] V. Arvind, J. Köbler, and M. Mundhenk. Lowness and the complexity of sparse and tally descriptions. In *Proceedings Third International Symposium on Algorithms and Computation,* 1992.

[AKM] V. Arvind, J. Köbler, and M. Mundhenk. Self-reducibility and reductions to sparse sets. In preparation.

[Ba92] J. Balcazár. Self-reducibility. *Journal of Computer and System Sciences*, to appear.

[BDG] J.L. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity I.* EATCS Monographs on Theoretical Computer Science, Springer-Verlag, 1988.

[BK88] R. Book and K. Ko. On sets truth-table reducible to sparse sets. *SIAM Journal on Computing*, 17(5):903–919, 1988.

[BL92] R. Book and J. Lutz. On languages with very high information content. *Proceedings of the 7th Structure in Complexity Theory Conference*, 255-259, IEEE Computer Society Press, 1992.

[BLS92] H. Buhrman, L. Longpré, and E. Spaan. Sparse reduces conjunctively to tally. *Technical Report NU-CCS-92-8,* Northeastern University, Boston, 1992.

[KL80] R. Karp and R. Lipton. Some connections between nonuniform and uniform complexity classes. *Proceedings of the 12th ACM Symposium on Theory of Computing,* 302-309, April 1980.

[Ko83] K. Ko. On self-reducibility and weak p-selectivity. *Journal of Computer and System Sciences*, 26:209-221, 198.

[KOSW86] K. Ko, P. Orponen, U. Schöning, and O. Watanabe. What is a hard instance of a computational problem. In *Proceedings of the 1st Structure in Complexity Theory Conference*, pages 197–217. *Lecture Notes in Computer Science #223*, Springer-Verlag, June 1986.

[LLS75] R. Ladner, N. Lynch, and A. Selman. A comparison of polynomial time reducibilities. *Theoretical Computer Science*, 1(2):103-124, 1975.

[LV92] M. Li and P. Vitanyi. *An introduction to Kolmogorov complexity and its application.* Addison-Wesley, 1992.

[LT91] A. Lozano and J. Torán. Self-reducible sets of small density. *Mathematical Systems Theory*, 24:83-100, 1991.

[Mah82] S. Mahaney. Sparse complete sets for NP: solution of a conjecture of Berman and Hartmanis. *Journal of Computer and System Sciences*, 25(2):130-143, 1982.

[OW91] M. Ogiwara and O. Watanabe. On polynomial-time bounded truth-table reducibility of NP sets to sparse sets. *SIAM Journal on Computing*, 20(3):471-483, 1991.

[RR92] D. Ranjan and P. Rohatgi. Randomized reductions to sparse sets. *Proceedings of the 7th Structure in Complexity Theory Conference*, IEEE Computer Society Press, 239-242, 1992.

[Ro92] P. Rohatgi. Saving queries with randomness. *Proceedings of the 7th Structure in Complexity Theory Conference*, IEEE Computer Society Press, 71-83, 1992.

[VV86] L.G. Valiant and V.V. Vazirani. NP is as easy as detecting unique solutions. *Theoretical Computer Science* 47, 85–93, 1986.

[Wag87] K.W. Wagner. More complicated questions about maxima and minima, and some closures of NP. *Theoretical Computer Science*, 51:53-80, 1987.

[Wag90] K.W. Wagner. Bounded query classes. *SIAM Journal on Computing*, 19(5):83-846, 1990.

[WW85] G. Wechsung and K.W. Wagner. On the boolean closure of NP. Manuscript. (Extended abstract by: G. Wechsung, On the boolean closure of NP, in *Proc. 1985 International Conference on Fundamentals of Computation Theory*, pages 485–493. Lecture Notes in Computer Science, Springer-Verlag, 1985.)