Electronic Colloquium on Computational Complexity, Report No. 31 (2004)

On Polynomially Time Bounded Symmetry of Information

Troy Lee
CWI and University of Amsterdam
Troy.Lee@cwi.nl

Andrei Romashchenko
Institute for Information Transmission Problems
anromash@mccme.ru

Abstract

The information contained in a string x about a string y is defined as the difference between the Kolmogorov complexity of y and the conditional Kolmogorov complexity of y given x, i.e., $I(x:y) = \mathrm{C}(y) - \mathrm{C}(y|x)$. From the well-known Kolmogorov–Levin Theorem it follows that $I(x:y)$ is symmetric up to a small additive term $O(\log \mathrm{C}(x,y))$. We are interested in if this property can hold for several versions of polynomial time bounded Kolmogorov complexity.

It is proven in papers by L. Longpré and S. Mocas [LM93] and L. Longpré and O. Watanabe [LW95] that, under a natural assumption, symmetry of information does not hold for the polynomial time bounded printing version of Kolmogorov complexity. In this paper, we investigate symmetry of information for some variants of distinguishing complexity CD, where $\mathrm{CD}(x)$ is the length of a shortest program which accepts x and only x. We show relativized worlds where symmetry of information does not hold for deterministic and nondeterministic polynomial time distinguishing complexities $\mathrm{CD}^{\mathrm{poly}}$ and $\mathrm{CND}^{\mathrm{poly}}$. For nondeterministic polynomial time distinguishing with randomness, $\mathrm{CAMD}^{\mathrm{poly}}$, we prove that symmetry of information holds for most pairs of strings in any set in NP. In proving this last statement we extend a recent result of Buhrman et al. [BLvM04], which may be of independent utility.

1 Introduction

One of the most beautiful theorems in Kolmogorov Complexity is the principle of "Symmetry of Information", independently proven by Kolmogorov and Levin [ZL70]. Roughly speaking, symmetry of information states that for any two strings x and y, the information contained in x about y is equal to the information contained in y about x, up to logarithmic factors. More formally, letting $\mathrm{C}(x)$ be the length of a shortest program which prints x, and $\mathrm{C}(y|x)$ be the length of a shortest program which prints y when given input x, symmetry of information can be stated as $\mathrm{C}(y) - \mathrm{C}(y|x) \approx \mathrm{C}(x) - \mathrm{C}(x|y)$. Besides its inherent attractiveness, this principle has also seen applications in diverse areas of theoretical computer science, for example in [JSV97, VV02, ABK+02].

In this paper, we investigate the principal of symmetry of information when resource bounds are placed on the program to describe y given x. While the argument of [ZL70] can be used without modification to show that symmetry of information holds for programs using exponential time or polynomial space, things become trickier with polynomial time bounds. Though this question has been around for some time, few definite answers are known, see Section 7.1 of [LV97] for a survey. The series of works [LM93, LW95] give evidence that symmetry of information does not hold for polynomial time printing complexity: in particular, they show that the existence of certain kinds of one-way functions implies that polynomial time symmetry of information does not hold for printing complexity. Intuitively, if f is a polynomial time computable one-way function, and $f(x) = y$, then y is simple given x. On the other hand, if x is simple in polynomial time given y then this would provide a way to invert the function, by cycling through all small programs.

En route to showing that BPP is in the polynomial hierarchy, Sipser [Sip83] introduced a relaxation of printing complexity called distinguishing complexity, denoted CD. For a string x, $\mathrm{CD}(x)$ is the length of a shortest program which accepts x and only x. Note that while printing and distinguishing complexity are equivalent up to a constant additive factor without resource bounds, this is not believed to be the case in the polynomial time setting [FK96, BFL02].

ISSN 1433-8092

The arguments of [LM93, LW95] leave open the question if symmetry of information can hold for distinguishing complexity. Now if f is a polynomial time computable one-way function and $f(x) = y$, then $\mathrm{CD}^{\mathrm{poly}}(x|y)$ is constant, as with a description of f, on input z we accept if and only if $f(z) = y$.

We investigate this possibility in considering symmetry of information for the following distinguishing complexity measures:

- Polynomial time distinguishing complexity, denoted $\mathrm{CD}^{\mathrm{poly}}(x)$.
- Nondeterministic polynomial time distinguishing complexity, denoted $\mathrm{CND}^{\mathrm{poly}}$.
- Nondeterministic polynomial time distinguishing with randomness, denoted $\mathrm{CAMD}^{\mathrm{poly}}$.

We show relativized worlds where symmetry of information fails in a strong way for $\mathrm{CD}^{\mathrm{poly}}$ and $\mathrm{CND}^{\mathrm{poly}}$. On the other hand, we show that for any set $A \in \mathrm{NP}$ symmetry of information holds for most pairs of strings $\langle x, y\rangle \in A$ with respect to the measure $\mathrm{CAMD}^{\mathrm{poly}}$. To prove this last statement we extend a recent result of [BLvM04], which may be of independent interest.

1.1 Background

Denote by $\mathcal{C}^{\mathrm{poly}}$ a version of polynomial time-bounded Kolmogorov complexity, which can be $\mathrm{C}^{\mathrm{poly}}$, $\mathrm{CD}^{\mathrm{poly}}$, $\mathrm{CND}^{\mathrm{poly}}$, or $\mathrm{CAMD}^{\mathrm{poly}}$. To formulate the problem of symmetry of information more precisely, we isolate three associated properties. The first is the *Easy Direction of Symmetry of Information*:

$$\text{For any polynomial } p \text{ there exists a polynomial } q \text{ such that for all } x, y \quad \mathcal{C}^q(x, y) \leq \mathcal{C}^p(x) + \mathcal{C}^p(y|x) + O(\log(n)). \tag{EDSI}$$

Next is the *Hard Direction of Symmetry of Information*:

$$\text{For any polynomial } p \text{ there exists a polynomial } q \text{ such that for all } x, y \quad \mathcal{C}^q(x) + \mathcal{C}^q(y|x) \lesssim \mathcal{C}^p(x, y), \tag{HDSI}$$

where the inequality should hold up to an additive logarithmic term, or at least up to $o(|x| + |y|)$.

Finally we also consider the property of *Symmetry of Mutual Information*:

$$\text{For any polynomial } p \text{ there exists a polynomial } q \text{ such that for all } x, y \quad \mathcal{C}^q(x) + \mathcal{C}^q(y|x) \lesssim \mathcal{C}^p(y) + \mathcal{C}^p(x|y) \tag{SMI}$$

Notice that if both (EDSI) and (HDSI) hold for a complexity measure $\mathcal{C}$, then also (SMI) holds for $\mathcal{C}$. The property (EDSI) is quite easy to prove for $\mathrm{C}^{\mathrm{poly}}$, $\mathrm{CD}^{\mathrm{poly}}$, $\mathrm{CND}^{\mathrm{poly}}$, or $\mathrm{CAMD}^{\mathrm{poly}}$. For the sake of completeness, in Section 3 we present a proof of this statement for distinguishing complexities.

It was shown in [LM93, LW95] that (HDSI) is not true for $\mathrm{C}^{\mathrm{poly}}$ if certain kinds of one-way functions exist (using a relativization from [IR89] we get a relativized world where (HDSI) is not true for $\mathrm{C}^{\mathrm{poly}}$). Are (HDSI) and (SMI) true for $\mathrm{CD}^{\mathrm{poly}}$ and $\mathrm{CND}^{\mathrm{poly}}$? It was claimed (without a complete proof) in [BF95] that these properties do not hold in some relativized world. In this paper we give the first published proof of this fact. Let us note that non-relativized variants of the properties above are probably very hard to prove, as the answer depends on such computational complexity problems as "P = NP?" and "P = PSPACE?".

1.2 Our Results

We show a relativized world where property (HDSI) fails in a very strong way for $\mathrm{CD}^{\mathrm{poly}}$ and $\mathrm{CND}^{\mathrm{poly}}$. In fact, we show that there is a set A and a polynomial p such that for any constant k, any polynomial q, for infinitely many $\langle x, y\rangle \in A$

$$k \cdot \mathrm{CD}^{p(n),A}(x, y) < \mathrm{CD}^{q(n),A}(x) + \mathrm{CD}^{A,q(n)}(y|x),$$

where $n = |x| + |y|$. Analogously, for any $\epsilon > 0$ there is a set A and a polynomial p such that for any polynomial q, for infinitely many $\langle x, y\rangle \in A$

$$(2 - \epsilon)\mathrm{CND}^{p(n),A}(x, y) < \mathrm{CND}^{q(n),A}(x) + \mathrm{CND}^{q(n),A}(y|x),$$

$n = |x| + |y|$. The bound $(2 - \epsilon)$ in the last inequality is tight as $2\mathrm{CND}^{A,\mathrm{poly}}(x,y) > \mathrm{CND}^{A,\mathrm{poly}}(x) + \mathrm{CND}^{A,\mathrm{poly}}(y|x) - O(1)$. We also find relativizations where property (SMI) fails in a strong way for $\mathrm{CD}^{\mathrm{poly}}$ and $\mathrm{CND}^{\mathrm{poly}}$.

We do not know if property (HDSI) holds absolutely for $\mathrm{CAMD}^{\mathrm{poly}}$ complexity. We are able to prove, however, that it holds for most pairs $\langle x, y\rangle$ in any set $A \in \mathrm{NP}$. More generally, the following property holds:

$$\begin{aligned}&\text{For any set } A \text{ and any polynomial } p \text{ there is a polynomial } q \text{ such that for most } \langle x,y\rangle \in A \cap \{0,1\}^n \\ &\qquad \mathrm{CAMD}^{p,A}(x,y) \geq \mathrm{CAMD}^{q,A}(x) + \mathrm{CAMD}^{q,A}(y|x) - O(\log^3 n).\end{aligned} \tag{1}$$

We also unconditionally prove the following weaker form of property (HDSI):

$$\begin{aligned}&\text{For polynomial } p(n) \text{ there exists a polynomial } q(n) \text{ such that } \forall x,y \\ &\qquad \mathrm{C}^p(x,y) \geq \mathrm{CAMD}^q(x) + \mathrm{CAMD}^q(y|x) - O(\log^3(|x|+|y|)).\end{aligned} \tag{2}$$

1.3 Organization

The rest of the paper is organized as follows. In section 3 we shortly prove a few simple properties of $\mathrm{CD}^{\mathrm{poly}}$ complexity. In particular, we prove that (HDSI) and (SMI) do not for $\mathrm{CD}^{\mathrm{poly}}$ complexity with some oracles. In section 4 we present a relativization such that (HDSI) and (SMI) are false for $\mathrm{CND}^{\mathrm{poly}}$ complexity in a strong way. Also we show that in the constructed relativized world $\mathrm{BPP} \not\subset \mathrm{NP}$ (note that other relativizations such that $\mathrm{BPP} \not\subset \mathrm{NP}$ are well known, see [Hel86]). In section 5 we prove (1) and (2); technically it is the most difficult part of the paper. The techniques in each section are quite different, and so the sections may be read independently.

2 Preliminaries

We use the following notation:

- denote by $\mathbb{B}$ the set $\{0,1\}$; similarly, $\mathbb{B}^n$ is the set of all binary strings of length n;
- denote by $|x|$ the length of a binary string x;
- denote by $\|A\|$ the cardinality of a finite set A;
- for a set $A \subset \mathbb{B}^*$ denote by $A^{=n}$ the set $\{x \ : \ x \in A \text{ and } |x| = n\}$.
- for a set of pairs of strings $A \subset \mathbb{B}^* \times \mathbb{B}^*$ denote by $A^{=n}$ the set $\{\langle x,y\rangle \in A \ : \ |x| + |y| = n\}$.

2.1 Kolmogorov Complexity Measures

We use notation for Kolmogorov complexity from the book of Li and Vitányi [LV97]:

Definition 1 *The Kolmogorov complexity function $C_F(y|x)$ is defined as*

$$C_F(y|x) = \min_p\{|p| \ : \ F(p,x) = y\}$$

where F is a partial recursive function.

By the Kolmogorov theorem [Kol65] there exists a universal machine U such that for any other F

$$\exists c \ \forall x,y \ C_U(y|x) \leq C_F(y|x) + c.$$

We fix such a function U and denote $\mathrm{C}(y|x) = C_U(y|x)$. We define unconditional Kolmogorov complexity by $\mathrm{C}(z) = \mathrm{C}(z|\lambda)$ (λ is the empty word). The choice of U affects the Kolmogorov complexity by at most an additive constant.

We also use a time bounded versions of Kolmogorov complexity.

Definition 2 *We define* $\mathrm{C}^t(y|x)$ *as*

$$\mathrm{C}^t(y|x) = \min_p\{|p| \ : \ U(p,x) = y \text{ and } U(p,x) \text{ runs in at most } t(|x|+|y|) \text{ steps}\}$$

for a universal machine U. *Also* $\mathrm{C}^t(z) = \mathrm{C}^t(z|\lambda)$.

The choice of universal machine U affects $\mathrm{C}^t(x|y)$ by at most an additive constant and the time bound t by at most a $\log(t)$ multiplicative factor.

In the time bounded case, the distinction between a program which *prints* a string x and one which *distinguishes* a string x, that is accepts x and only x, becomes important. Sipser [Sip83] defined time bounded *distinguishing* complexity as follows.

Definition 3 *Distinguishing complexity of* y *conditional to* x *is defined as*

$$\mathrm{CD}^t(y|x) = \min_p \left\{ |p| \ : \ \begin{array}{ll} 1) & U(p,x,y) \text{ accepts} \\ 2) & U(p,x,z) \text{ rejects for } \forall z \neq y \\ 3) & U(p,x,z) \text{ runs in at most } t(|x|+|z|) \text{ steps } \forall z \end{array} \right\}$$

for a universal machine U. *Once again,* $\mathrm{CD}^t(z) = \mathrm{CD}^t(z|\lambda)$.

In [BFL02] a nondeterministic variant of distinguishing complexity is defined.

Definition 4 *Let* U_n *be a* nondeterministic *universal machine.*

$$\mathrm{CND}^t(y|x) = \min_p \left\{ |p| \ : \ \begin{array}{ll} 1) & U_n(p,x,y) \text{ accepts} \\ 2) & U_n(p,x,z) \text{ rejects for } \forall z \neq y \\ 3) & U_n(p,x,z) \text{ runs in at most } t(|x|+|z|) \text{ steps } \forall z \end{array} \right\}.$$

As usual, we let $\mathrm{CND}^t(z) = \mathrm{CND}^t(z|\lambda)$.

Further, following [BLvM04] we define a complexity based on the complexity class AM:

Definition 5 *Let* U_n *be a* nondeterministic *universal machine.*

$$\mathrm{CAMD}^t(y|x) = \min_p \left\{ |p| \ : \ \begin{array}{ll} 1) & \mathrm{Prob}_{r\in\mathbb{B}^t}[U_n(p,x,y,r) \text{ accepts}] > 2/3 \\ 2) & \mathrm{Prob}_{r\in\mathbb{B}^t}[U_n(p,z,y,r) \text{ accepts}] < 1/3 \ \forall z \neq z \\ 3) & U_n(p,z,x,r) \text{ runs in at most } t(|x|+|z|+|r|) \text{ steps} \end{array} \right\},$$

and $\mathrm{CAMD}^t(z) = \mathrm{CAMD}^t(z|\lambda)$.

We also use *relativized* version of Kolmogorov complexities. The complexity $\mathrm{C}^{t,A}$ is defined as complexity C^t above except that the universal machine U has access to the set A as an oracle. The complexities $\mathrm{CD}^{t,A}$, $\mathrm{CAMD}^{t,A}$, and $\mathrm{CND}^{t,A}$ can be defined similarly.

2.2 Language Compression Theorems

A fundamental theorem of Kolmogorov complexity, and one that is very useful in applications, is the following:

Theorem 6 (Language Compression Theorem) *For any recursively enumerable set* A, *and all* $x \in A^{=n}$,

$$\mathrm{C}(x) \leq \log \|A^{=n}\| + O(\log n).$$

This is as x can be described by its index in the enumeration of $A^{=n}$.

In particular, this theorem is essentially used in the proof of (HDSI) in the resource unbounded case given in [ZL70]. Similarly, our results about resource bounded symmetry of information (both positive and negative) crucially rely on recent resource bounded language compression theorems.

In [BLvM04] the following analogue of the Language Compression Theorem is shown for CND complexity.

Theorem 7 ([BLvM04]) *There is a polynomial $p(n)$ such that for any set $A \subset \mathbb{B}^*$ and for all $x \in A^{=n}$*

$$\mathrm{CND}^{p,A^{=n}}(x) \le \log \|A^{=n}\| + O(\sqrt{\log \|A^{=n}\|}\log(n)).$$

Further [BLvM04] show that with the power of Arthur-Merlin protocols a Language Compression Theorem holds which is optimal up to an additive $\log^3 n$ term:

Theorem 8 ([BLvM04]) *There is a polynomial $p(n)$ such that for any set $A \subset \mathbb{B}^*$ and for all $x \in A^{=n}$*

$$\mathrm{CAMD}^{p,A^{=n}}(x) \le \log \|A^{=n}\| + O(\log^3(n)).$$

For comparison we remark that for CD complexity the situation is somewhat different. In [BFL02] it is shown that there is a polynomial $p(n)$ such that for any set A and for all $x \in A^{=n}$

$$\mathrm{CD}^{p(n),A^{=n}}(x) \le 2\log \|A^{=n}\| + O(\log n).$$

Furthermore, [BLM00] show that there is a set A where this bound is tight up to $O(\log n)$ terms. That is, the factor of 2 in general cannot be improved.

3 On CD complexity

In this section we present a few simple propositions concerning CD complexity. At first we prove a positive fact: the inequality (EDSI) is true for $\mathrm{CD}^{\mathrm{poly}}$ complexity:

Proposition 9 *For any polynomial $p(n)$ there exists a polynomial $q(n)$ such that for any oracle A and any x, y*

$$\mathrm{CD}^{q(n),A}(x,y) \le \mathrm{CD}^{p(n),A}(x) + \mathrm{CD}^{q(n),A}(y|x) + O(\log n),$$

where $n = |x| + |y|$.

Proof: Let s_1 be a shortest distinguishing program for x and s_2 be a shortest distinguishing program for y conditional to x, and both programs have access to the oracle A and run in time $p(n)$. Let us show that a pair $\langle s_1, s_2\rangle$ can be used as a polynomial time distinguishing program for the pair $\langle x, y\rangle$.

Given a pair $\langle a, b\rangle$ we do as follows:

1. run $s_1(b)$; if s_1 rejects the string a, then reject $\langle a, b\rangle$;
2. otherwise run s_2 on the string b given a; accept $\langle a, b\rangle$ if s_2 accepts.

Obviously, this algorithm works in time poly(n) and accepts only the pair $\langle x, y\rangle$. Extra $O(\log n)$ bits are required to provide a prefix free encoding of the pair, so $\mathrm{CD}^{q(n),A}(x,y) \le |s_1| + |s_2| + O(\log n)$. □

The proof above obviously works for $\mathrm{CND}^{\mathrm{poly}}$ and $\mathrm{CAMD}^{\mathrm{poly}}$ as well.

In contrast to Proposition 9, we show that inequality (HDSI) does not hold for $\mathrm{CD}^{\mathrm{poly}}$ complexity in some relativized worlds. Note that for any polynomial $q(n)$

$$\mathrm{CD}^{2^{\epsilon n},A}(y|x) \le \mathrm{CD}^{p(n),A}(y|x)$$

for large enough n. Hence, it is enough to find an oracle such that $\mathrm{CD}^{p(n),A}(x,y) \ll \mathrm{CD}^{2^{\epsilon n},A}(y) + \mathrm{CD}^{2^{\epsilon n},A}(x|y)$.

Proposition 10 *There exists an oracle A and a polynomial $p(n)$ satisfying the following condition. For any $\epsilon > 0$ and large enough n there exists a pair $\langle x, y\rangle \in A^{=2n}$ such that*

- $\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(y) > (1-\epsilon)n - O(1)$,
- $\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(x|y) > (1-\epsilon)n - O(\log n)$,

- $\mathrm{CD}^{p(n),A^{=2n}}(x,y) = O(1)$,

i.e., $\mathrm{CD}^{p(n),A^{=2n}}(x,y) \ll \mathrm{CD}^{2^{\epsilon n},A^{=2n}}(y) + \mathrm{CD}^{2^{\epsilon n},A^{=2n}}(x|y)$.

Proof: We shall construct an extremely simple oracle $A^{=2n} \subset \mathbb{B}^n \times \mathbb{B}^n$. Namely, there will be only one point in the whole set. We choose $\langle x_0, y_0\rangle \in \mathbb{B}^n \times \mathbb{B}^n$ such that $C(x_0) \geq n$ and $C(y_0|x_0) \geq n - O(1)$. Now set $A^{=2n} = \{\langle x_0, y_0\rangle\}$, and the construction is finished. Obviously, $\mathrm{CD}^{p(n),A^{=2n}}(x_0,y_0) = O(1)$. It remains to show that $\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(x_0)$ and $\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(y_0|x_0)$ are large enough.

Fix a shortest CD-program s for x_0. If this program does not query the point $\langle x_0, y_0\rangle$ from the oracle, then it could work with a trivial oracle (which returns 0 for all queries) as well. In this case we get

$$\mathrm{C}(x_0) \leq |s| + O(1).$$

Conversely, if at some step $t \leq 2^{\epsilon n}$ the program s does query the point $\langle x_0, y_0\rangle$ from the oracle, we get

$$\mathrm{C}(x_0) \leq |s| + \log t \leq |s| + \epsilon n + O(\log n).$$

Thus in both cases

$$\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(x_0) > (1-\epsilon)n - O(\log n).$$

The same arguments imply $\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(y_0|x_0) > (1-\epsilon)n - O(\log n)$, and we are done. □

The proof above provides a relativized world where (HDSI) is false. In fact in this world even the expected property $\mathrm{CD}^p(x,y) \geq \mathrm{CD}^{2^{\epsilon n}}(x)$ is false. To find a relativization where also (SMI) is false for $\mathrm{CD}^{\mathrm{poly}}$ we need a bit more complex construction. The proof of the next proposition follows the idea outlined in [BF95].

Proposition 11 *There exists an oracle A and a polynomial $p(n)$ satisfying the following condition. For any $\epsilon > 0$ and large enough n there exists a pair $\langle x, y\rangle \in \mathbb{B}^n \times \mathbb{B}^n$ such that*

- $\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(y) > (1-\epsilon)n - O(1)$,
- $\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(x|y) > (1-\epsilon)n - O(\log n)$,
- $\mathrm{CD}^{p(n),A^{=2n}}(x) = O(1)$,
- $\mathrm{CD}^{p(n),A^{=2n}}(y|x) = O(1)$ *and even* $\mathrm{C}^{p(n),A^{=2n}}(y|x) = O(1)$,

i.e., $\mathrm{CD}^{p(n),A^{=2n}}(x) + \mathrm{CD}^{p(n),A^{=2n}}(y|x) \ll \mathrm{CD}^{2^{\epsilon n},A^{=2n}}(y) + \mathrm{CD}^{2^{\epsilon n},A^{=2n}}(x|y)$. *Thus, (SMI) doe not hold with the oracle A.*

Proof: Fix n and choose a random pair $\langle x_n, y_n\rangle \in \mathbb{B}^n \times \mathbb{B}^n$. Let f_n be a random permutation of strings of length n such that $f_n(x_n) = y_n$.

Now we define $A^{=2n}$. At first we define two auxiliary oracles B_n and C_n: let B_n contain the graph of the permutation f_n (on input $\langle x, i\rangle$ the oracle B_n returns the i-th bit of $y = f_n(x)$) and C_n contain a single string x_n (on input $x \in \mathbb{B}^n$ the oracle C_n returns 1 if and only if $x = x_n$). A query to B_n consists of $(n + \log n)$ bits, and a query to C_n consists of n bits. So a query to $B_n \oplus C_n$ can be encoded as a strings of length $(n + \log n + 1)$, which is less than $2n$. Thus, we may set $A^{=2n} = B_n \oplus C_n$.

Obviously, for some polynomial $p(n)$ we have $\mathrm{CD}^{p(n)}(x) = O(1)$ (it is enough to query C_n to distinguish x from any other stings) and $\mathrm{C}^{p(n)}(y|x) = O(1)$ (it is enough to query from B_n the value $f_n(x_n)$).

The same arguments as in the proof of Proposition 9 imply that for randomly chosen f_n we have $\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(y_n) \geq (1-\epsilon)n - O(\log n)$ and $\mathrm{CD}^{2^{\epsilon n},A^{=2n}}(x_n|y_n) > (1-\epsilon)n - O(\log n)$. □

In Proposition 9 and Proposition 11 the universal machine accesses the oracle $A^{=n}$, which obviously depends on the length of input. These results hold also for a more usual in complexity theory 'uniform' relativization:

Corollary 12 *There exists an oracle A and a polynomial $p(n)$ satisfying the following condition. For any $\epsilon > 0$ and large enough n there exists a pair $\langle x, y\rangle \in \mathbb{B}^n \times \mathbb{B}^n$ such that $\mathrm{CD}^{p(n),A}(x) + \mathrm{CD}^{p(n),A}(y|x) \ll \mathrm{CD}^{2^{\epsilon n},A}(y) + \mathrm{CD}^{2^{\epsilon n},A}(x|y)$.*

Proof: Let A be the oracle from Proposition 11. Denote by A' the set $\bigcup_{i=1}^{\infty} A^{=2^{2^i}}$. Then we may assume that on input of length $n = 2^{2^i}$ the universal machine, which runs in time less than $2^{\epsilon n}$, can query from the oracle only strings of length not greater than n. Further, the list of all positive answers of the oracles for queries of length *less* than n contains only $O(\log n)$ bits. Hence, for any $\langle x, y\rangle \in A^{=n}$ the difference between $\mathrm{CD}^{A'}(y|x)$ and $\mathrm{CD}^{A^{=n}}(y|x)$ is $O(\log n)$. Thus, it follows from Proposition 11 that $\mathrm{CD}^{p(n),A'}(x) + \mathrm{CD}^{p(n),A'}(y|x) \ll \mathrm{CD}^{2^{\epsilon n},A'}(y) + \mathrm{CD}^{2^{\epsilon n},A'}(x|y)$.

□

4 On CND complexity

In this section we prove that (HDSI) and (SMI) are not true for a relativized version of polynomial time bounded CND complexity. Our proof is based on the Language Compression Theorem 7.

Theorem 13 *Let $m = m(n), s = s(n), t = t(n)$ be functions such that*

$$2^{s(n)} + 2^{m(n)} \leq 2^{3n}$$

and

$$t(n)2^{m(n)} \leq 2^{3n-3}.$$

Then there is a polynomial $p(n)$, and sets A, X such that

- $X^{=3n} \subset \mathbb{B}^{3n}$, $\|X^{=3n}\| = 2^{s(n)}$,
- $A^{=6n} \subset \mathbb{B}^{3n} \times \mathbb{B}^{3n}$,
- $\|\{y \ : \ (x, y) \in A^{=6n}\}\| \geq 7/8 \cdot 2^{3n}$ *for any* $x \in X^{=3n}$,
- $\| \bigcup_{x \notin X} \{y \ : \ (x, y) \in A^{=6n}\}\| \leq 1/8 \cdot 2^{3n}$,

and for large enough n, for all $x \in X^{=3n}$, for at least $3/4 \cdot 2^{3n}$ strings $y \in \mathbb{B}^{3n}$ the following conditions hold:

$$\begin{array}{lcl} \langle x, y\rangle \in A^{=6n}, & & \\ \mathrm{CND}^{p,A^{=6n}}(x|y) & \leq & s(n) + O(\delta(n)), \\ \mathrm{CND}^{p,A^{=6n}}(x, y) & \leq & (3n + s(n)) + O(\delta(n)), \\ \mathrm{CND}^{t(n),A^{=6n}}(y|x) & \geq & 3n - O(1), \\ \mathrm{CND}^{t(n),A^{=6n}}(x) & \geq & m(n) - O(1), \end{array}$$

where $\delta(n) = \sqrt{n}\log^3(n)$.

Note that the term $\delta(n) = \sqrt{n}\log^3(n)$ comes from Theorem 7.

Corollary 14 *There exists an oracle A such that a $\mathrm{CND}^{\mathrm{poly}}$ version of (HDSI) does not hold. Moreover, for any $\varepsilon > 0$ there exists a polynomial p such that for any polynomial q for large enough n*

$$(2 - \varepsilon)\mathrm{CND}^{p,A^{=6n}}(x, y) \ll \mathrm{CND}^{q,A^{=6n}}(x) + \mathrm{CND}^{q,A^{=6n}}(y|x)$$

for most $\langle x, y\rangle \in A^{=6n}$.

Proof: It follows from Theorem 13 for $s(n) = \varepsilon n$, $m(n) = (3-\varepsilon)n$, $t(n) = 2^{\varepsilon n/2}$. □

The bound $(2-\varepsilon)$ in Corollary 14 is tight. This can be easily seen as, $\mathrm{CND}^{\mathrm{poly},A^{=6n}}(x,y) > \mathrm{CND}^{\mathrm{poly},A^{=6n}}(x) - O(1)$ and $\mathrm{CND}^{\mathrm{poly},A^{=6n}}(x,y) \geq \mathrm{CND}^{\mathrm{poly},A^{=6n}}(y|x) - O(1)$. Hence for any oracle A

$$2\mathrm{CND}^{p,A^{=6n}}(x,y) \geq \mathrm{CND}^{q,A^{=6n}}(x) + \mathrm{CND}^{q,A^{=6n}}(y|x) - O(1).$$

Corollary 15 *There exists an oracle A such that a $\mathrm{CND}^{\mathrm{poly}}$ version of (SMI) does not hold, i.e., for some polynomial p, any polynomial q, and large enough n*

$$\mathrm{CND}^{p,A^{=6n}}(y) + \mathrm{CND}^{p,A^{=6n}}(x|y) \ll \mathrm{CND}^{q,A^{=6n}}(x) + \mathrm{CND}^{q,A^{=6n}}(y|x)$$

for most $\langle x,y\rangle \in A^{=6n}$.

Proof: Again, we can apply Theorem 13. For example, set $s(n) = n$, $m(n) = 2n$, $t(n) = 2^{\varepsilon n}$ (for small enough $\varepsilon > 0$). Than for most $\langle x,y\rangle \in A^{=6n}$ we have $\mathrm{CND}^{p,A^{=6n}}(y) + \mathrm{CND}^{p,A^{=6n}}(x|y) \leq 4n + O(1)$ and $\mathrm{CND}^{q,A^{=6n}}(x) + \mathrm{CND}^{q,A^{=6n}}(y|x) \geq 5n - O(1)$. □

Remark: *Using the same trick as in Corollary 12, we can get an analog of Theorem 13 with a 'uniform' relativization, i.e., we may assume that the oracle does not depend on n. So, Corollary 14 and Corollary 15 also can be formulated for a uniform relativization: there exists an oracle A and polynomial p such that for any polynomial q and large enough n*

$$\mathrm{CND}^{p,A}(y) + \mathrm{CND}^{p,A}(x|y) \ll \mathrm{CND}^{q,A}(x) + \mathrm{CND}^{q,A}(y|x)$$

for most $\langle x,y\rangle \in A \cap \mathbb{B}^n$.

Proof:(Theorem 13) Fix an integer $n > 0$. We denote by F the characteristic function of $A^{=6n}$, i.e., $F(\langle x,y\rangle) = 1$ if $\langle x,y\rangle \in A^{=6n}$ and $F(x,y) = 0$ otherwise. Our goal is to define a function F so that the statement of the theorem hold. We define this function in a few stages: construct a sequence of functions $F_0, F_1, \ldots, F_{2^{m(n)}}$,

$$F_i : \mathbb{B}^{3n} \times \mathbb{B}^{3n} \to \{0, 1, \mathbf{undef}\}.$$

For $i < j$ the function F_j should be an extension of F_i, i.e.,

$$\forall\langle a,b\rangle \text{ if } F_i(a,b) \neq \mathbf{undef} \text{ then } F_j(a,b) = F_i(a,b).$$

The initial function is trivial: $F_0(a,b) = \mathbf{undef}$ for all $\langle a,b\rangle$; the last function $F_{2^{m(n)}}$ should range over $\mathbb{B}$, i.e., $F_{2^{m(n)}}(a,b) \neq \mathbf{undef}$ for any a,b. We set $F = F_{2^{2^{m(n)}}}$.

Let us introduce some notation. We say that a set $B \subset \mathbb{B}^{3n} \times \mathbb{B}^{3n}$ *respects* a function F_i if

$$\begin{cases} F_i(a,b) = 1 & \Rightarrow \quad \langle a,b\rangle \in B, \\ F_i(a,b) = 0 & \Rightarrow \quad \langle a,b\rangle \notin B. \end{cases}$$

Let $s_1, \ldots, s_{2^{m(n)}-1}$ be the list of all CND-programs of length less than $m(n)$. We suppose each program s_j can access an oracle O (the oracle is not fixed in advance). Also we suppose that each s_j is clocked and runs in time less than $t(n)$. We say that s_j is a *well defined* CND program for an oracle O if s_j^O accepts exactly one string x.

Further define F_i by induction. Let the functions $F_0, \ldots, F_{k-1}$ be already defined. We must construct a function

$$F_k : \mathbb{B}^{3n} \times \mathbb{B}^{3n} \to \{0, 1, \mathbf{undef}\}$$

which is an extension of F_{k-1}. Consider the program s_k. There are two possibilities:

1. for any $B \subset \mathbb{B}^{3n} \times \mathbb{B}^{3n}$ that respects F_{k-1}, the program s_k is not well defined for the oracle B;
2. there exists at least one set $B \subset \mathbb{B}^{3n} \times \mathbb{B}^{3n}$ that respects F_{k-1}, and the program s_k is well defined for the oracle B.

The first case is trivial: we set $F_k(x,y) = F_{k-1}(x,y)$ for all $\langle x,y\rangle$. In the second case there exists a set B and a string x such that s_k^B accepts x in time $T(B,x) < t(n)$ and rejects all other strings. If there are more than one B as above, we choose a set B that provides minimum to the value $T(B,x)$. Denote by x_k the fixed string x. Let the list of all queries of the program $s_k^B(x_k)$ to the oracle (for one of the accepting paths) be

$$\langle a_0,b_0\rangle, \langle a_1,b_1\rangle, \ldots, \langle a_r,b_r\rangle$$

($r < t(n)$). We include all these pairs in the oracle. More precisely, define F_k as follows:

$$\begin{array}{lcll} F_k(a,b) & = & F_{k-1}(a,b) & \text{if } F_{k-1}(a,b) \neq \textbf{undef}, \\ F_k(a_j,b_j) & = & 1 & \text{if } \langle a_j,b_j\rangle \in B,\ j=0,\ldots,r, \\ F_k(a_j,b_j) & = & 0 & \text{if } \langle a_j,b_j\rangle \notin B,\ j=0,\ldots,r, \\ F_k(a,b) & = & \textbf{undef} & \text{if } F_{k-1}(a,b) = \textbf{undef} \text{ and } \langle a,b\rangle \neq \langle a_j,b_j\rangle,\ j=0,\ldots,r. \end{array}$$

For any set R that respects F_k, the program s_k^R accepts the string x_k. This means that for any time bound $T \leq t(n)$ the program s_k^R cannot distinguish any string except for x_k.

Thus we have described an inductive procedure, which defines the functions $F_0, \ldots, F_{2^{m(n)}-1}$. At each step i we set $F_i(a,b) \neq F_{i-1}(a,b)$ for at most $t(n)$ values $\langle a,b\rangle$. Hence the function $F_{2^{2n}-1}$ is equal to **undef** for all values in $\mathbb{B}^{3n} \times \mathbb{B}^{3n}$ except for at most $t(n)2^{m(n)}$ values.

Besides we get the list L of strings x_i which can be accepted by distinguishing programs s_i^R if a set R respects $F_{2^{2^{m(n)}}-1}$. This set is rather small: $\|L\| < 2^{m(n)}$.

Further we choose an arbitrary set

$$X^{=3n} \subset \mathbb{B}^{3n} \setminus L$$

of size $2^{s(n)}$. Now define the function $F_{2^{2^{m(n)}}}$ as follows:

$$\begin{array}{lcll} F_{2^{2^{m(n)}}}(x,y) & = & F_{2^{2n}-1}(x,y) & \text{if } F_{2^{2^{m(n)}}-1}(x,y) \neq \textbf{undef}, \\ F_{2^{2^{m(n)}}}(x,y) & = & 1 & \text{if } F_{2^{2^{m(n)}}-1}(x,y) = \textbf{undef} \text{ and } x \in X, \\ F_{2^{2^{m(n)}}}(x,y) & = & 0 & \text{if } F_{2^{2^{m(n)}}-1}(x,y) = \textbf{undef} \text{ and } x \notin X. \end{array}$$

The characteristic function $F_{2^{2^{m(n)}}}$ defines the oracle $A^{=6n}$ and the construction is finished. Note that

$$\|\{y \ : \ (x,y) \in A^{=6n}\}\| \geq 7/8 \cdot 2^{3n}$$

for any $x \in X^{=3n}$, and

$$\Big\| \bigcup_{x \notin X^{=3n}} \{y \ : \ (x,y) \in A^{=6n}\} \Big\| \leq 1/8 \cdot 2^{3n}.$$

Remark: *If $x \in X^{=3n}$ then for at least $7/8$ of all $y \in \mathbb{B}^{3n}$ we have $\langle x,y\rangle \in A$; If $x \notin X^{=3n}$ then for at most $1/8$ of strings $y \in \mathbb{B}^{3n}$ we have $\langle x,y\rangle \in A$. We shall use this observation in Corollary 16 below.*

Now fix any string $x_0 \in X$. Obviously, $\mathrm{CND}^{t(n),A^{=6n}}(x_0) \geq m(n)$ because $x \notin L$. Further, there are at least

$$2^{3n} - 2^{m(n)}t(n) - 2^{3n-3} > 3/4 \cdot 2^{3n}$$

strings y such that

- $(x_0,y) \in A^{=6n}$,
- $(x,y) \notin A^{=6n}$ for any $x \notin X^{=3n}$, and
- $\mathrm{C}^{A^{=6n}}(y|x_0) \geq 3n-3$.

Denote by y_0 any of these strings. From the conditions above it follows that

- $\mathrm{CND}^{t(n),A^{=6n}}(y_0|x_0) > 3n - O(1)$ since resource bounded complexity is not less than plain complexity;
- $\mathrm{CND}^{p(n),A^{=6n}}(x_0|y_0) \leq \log\|\{x \ : (x,y_0) \in A^{=6n}\}\| + O(\delta(n)) \leq s(n) + O(\delta(n))$ (from Theorem 7);

- $\mathrm{CND}^{p(n),A^{=6n}}(x,y) \leq \log \|A^{=6n}\| + O(\delta(n)) = (3n + s(n)) + O(\delta(n))$ (also from Theorem 7).

□

Corollary 16 $\exists M \ : \ \mathrm{BPP}^M \not\subset \mathrm{NP}^M$.

Proof: Apply Theorem 13 for $s(n) = n$, $m(n) = 2n$, and $t(n) = 2^{0.1n}$. Denote $X' = \bigcup_{i=1}^{\infty} X^{=3\cdot 2^{2^i}}$ and $A' = \bigcup_{i=1}^{\infty} A^{=6\cdot 2^{2^i}}$, where A and X are sets from Theorem 13. It is easy to check that X' belongs to the class $\mathrm{BPP}^{A'}$. Let x be a string of length $3n$ ($n = 2^{2^i}$), and say we want to solve if $x \in X'$. Choose at random a string $y \in \mathbb{B}^{3n}$ and check whether $(x,y) \in A^{=6n}$; if $x \in X^{=3n}$, then probability of this event is at least $7/8$, otherwise probability is at most $1/8$. Obviously, the complement $\bar{X}'$ of the set X' also belongs to $\mathrm{BPP}^{A'}$.

The same time $\bar{X}'$ or X' does not belong to $\mathrm{NP}^{A'}$. Assume the converse, i.e., $\bar{X}'$ and X' are both in $\mathrm{NP}^{A'}$. Then a non-deterministic polynomial machine with the oracle A' can solve if $x \in X'$.

Hence, from Theorem 7 we have for all $x \in X^{=3n}$

$$\mathrm{CND}^{\mathrm{poly},A'}(x) \leq \log \|X^{=3n}\| + O(\delta(n)) = n + O(\delta(n)).$$

On the other hand, $\mathrm{CND}^{\mathrm{poly},A'}(x) \geq \mathrm{CND}^{\mathrm{poly},X^{=3n}}(x) - O(\log n)$, and from Theorem 13

$$\mathrm{CND}^{\mathrm{poly},X^{=3n}}(x) \geq 2n.$$

Thus, we get a contradiction. □

5 On CAMD complexity

In this section we study symmetry of information under the CAMD complexity measure. In contrast to the case of CND complexity, with the power of nondeterminism and randomness we can prove some positive results, showing that some weaker versions of (HDSI) hold for CAMD.

Our proof will follow the proof in the resource unbounded case as given in [ZL70]. We now recall this proof to highlight how it can be adapted for our purposes. Let α, β be two strings such that $|\alpha| + |\beta| = n$, and suppose that $\mathrm{C}(\alpha,\beta) = m$. We define the set $A_{x,m} = \{y : \mathrm{C}(x,y) \leq m\}$. Notice that $\|A_{x,m}\| \leq 2^{m+1}$ and that given x and m the set $A_{x,m}$ is recursively enumerable. Thus as $\beta \in A_{\alpha,m}$ by the Language Compression Theorem (Theorem 6), $\mathrm{C}(\beta|\alpha) \leq \log \|A_{\alpha,m}\| + O(\log n)$. Let k^* be such that $2^{k^*} \leq \|A_{\alpha,m}\| < 2^{k^*+1}$. Then the above says that $\mathrm{C}(\beta|\alpha) \leq k^* + O(\log n)$.

Now consider the set $B_{m,k} = \{x : \|A_{x,m}\| \geq 2^k\}$. Notice that the size of $B_{m,k}$ is less than 2^{m-k}, and that $\alpha \in B_{m,k^*}$. The set $B_{m,k}$ is recursively enumerable given m, k, thus by the Language Compression Theorem, $\mathrm{C}(\alpha) \leq m - k^* + O(\log n)$. And so

$$\begin{aligned} \mathrm{C}(\alpha) + \mathrm{C}(\beta|\alpha) &\leq m - k^* + k^* + O(\log n) \\ &\leq \mathrm{C}(\alpha,\beta) + O(\log n) \end{aligned}$$

Let us see what happens when we try to apply this argument to CAMD complexity. We no longer know if the set $A_m = \{(x,y) : \mathrm{CAMD}^p(x,y) \leq m\}$ is decidable in AM. We can, however, decide if (x,y) satisfies $\mathrm{C}^p(x,y) \leq m$ in nondeterministic polynomial time, by guessing the polynomial time printing program and running it. Thus we redirect our attempts at the weaker statement: for all $x, y \in \mathbb{B}^n$,

$$\mathrm{C}^p(x,y) \geq \mathrm{CAMD}^q(x) + \mathrm{CAMD}^q(y|x) - O(\log^3 n).$$

Now the first half of the argument works as in the resource unbounded case, and, using the Language Compression Theorem for AM, Theorem 8, we have $\mathrm{CAMD}^q(\beta|\alpha) \leq k^* + O(\log^3 n)$.

The set $B_{m,k}$ is more tricky to decide, as we need to count the number of y such that $\mathrm{C}^p(x,y) \leq m$. In AM, however, we can approximately lower bound the size of NP relations as shown in Theorem 5.2 of [Bab85], using Sipser's Coding Lemma [Sip83]. In our case this means there is a polynomial time predicate Q such that

- if $x \in B_{m,k}$ then $\Pr_r[\exists y\ Q_A(x,y,r) = 1] \geq 2/3$
- if $x \notin B_{m,k-1}$ then $\Pr_r[\exists y\ Q_A(x,y,r) = 1] \leq 1/3$

We would now like to apply the language compression theorem for AM, but there remains a small problem: the set $B_{m,k}$ is *not actually in* AM. Note that if x is such that there are between 2^{k-1} and 2^k strings y such that $\mathrm{C}(x,y) \leq m$, then we have no guarantee about the behavior of the above algorithm: we cannot decide if $x \in B_{m,k}$ or not. This obstacle, however, is the only problem with the above argument going through.

In the next theorem, we extend the language compression results of [BLvM04] to also work for AM gap sets of this type, thus allowing the above argument to go through.

Theorem 17 *Let $A \subseteq \mathbb{B}^*$. Suppose there is a polynomial time bound $q(n)$, and predicate Q such that*

- *for all $u \in A^{=n}$, $\Pr_{r \in \mathbb{B}^{q(n)}}[\exists v\ Q(u,v,r) = 1] \geq 2/3$*
- *$\|L = \{u \in \mathbb{B}^n : \Pr_{r \in \mathbb{B}^{q(n)}}[\exists v\ Q(u,v,r) = 1] \geq 1/3\}\| \leq 2^k$,*

and for all u, v, r the predicate $Q(u,v,r)$ can be computed in polynomial time. Then there is a polynomial time bound $p(n)$ such that for all $u \in A^{=n}$, we have $\mathrm{CAMD}^p(u) \leq k + O(\log^3 n)$.

Before going into the proof, we briefly recall the technique of [BLvM04]. Let $\mathrm{TR} : \mathbb{B}^n \times \mathbb{B}^d \to \mathbb{B}^m$ be the function underlying Trevisan's extractor [Tre99], that is the composition of a good error correcting code with the Nisan-Wigderson generator [NW94]. The output of $\mathrm{TR}(u,e)$ is the evaluation of the Nisan-Wigderson generator on seed e when using $\hat{u}$ as the 'hard' function supplied to the generator, where $\hat{u}$ is the image of u under an error correcting code. The key property of this function, what makes it a good extractor and compressor, is that if $\mathrm{TR}(u,e)$ is not close to uniform over choice of $e \in \mathbb{B}^d$ on some set $B \subseteq \mathbb{B}^m$, then u has a short description given oracle access to B. In [BLvM04] it is shown that u can be printed in polynomial time from this description and oracle access to B.

To give the elements of a set $A \in \mathbb{B}^n$ short descriptions, we let the set B be the image of $A \times \mathbb{B}^d$ under the the function TR. That is, $B = \cup_{x \in A} \cup_{e \in \mathbb{B}^d} \mathrm{TR}(x,e)$. Notice that for any $x \in A$, $\Pr_e[\mathrm{TR}(x,e) \in B] = 1$. On the other hand if we take m to be $\log \|A\| + d + 1$ then the probability that a uniformly chosen $y \in \mathbb{B}^m$ is in B is less than $1/2$. Thus all the elements of A have a short description relative to B. Now notice that with nondeterminism and an oracle for A, we can decide membership in B, thus all the elements of A have a short CND^A description. The elements of A can be given an even more succinct CAMD^A description by using the randomness in the AM protocol to simulate part of the probabilistic argument in [NW94, Tre99].

Proof:(Theorem 17) By amplification and the results of [FGM+89], we can transform the predicate Q into a predicate Q' taking random strings of length a polynomial $q'(n)$ and with the property

- if $u \in A^{=n}$ then $\Pr_r[\exists v\ Q'(u,v,r) = 1] = 1$
- $\|L' = \{u : \Pr_r[\exists v\ Q'(u,v,r) = 1] \geq 2^{-n-2}\}\| \leq 2^k$

for r chosen uniformly over $\mathbb{B}^{q'(n)}$.

For each $r \in \mathbb{B}^{q'(n)}$ we define a set

$$B_r = \{w : \exists u \in \mathbb{B}^n, \exists v, e\ \mathrm{TR}(u,e) = w \wedge Q'(u,v,r) = 1\}$$

Clearly if $u \in A^{=n}$, then $\Pr_e[B_r(\mathrm{TR}(u,e)) = 1] = 1$, for any $r \in \mathbb{B}^{q'(n)}$. We now calculate the probability that for a randomly chosen $w \in \mathbb{B}^m$ and randomly chosen $r \in \mathbb{B}^{q'(n)}$, that $w \in B_r$. As for a 0/1 variable the probability of being 1 is equal to the expectation of the variable, we have

$$\Pr_{r,w}[w \in B_r] = E_{r,w}[B_r(w)].$$

By linearity of expectation, we can divide the latter into two contributions, that from elements w for which $\exists u \in L'$ and seed e such that $\mathrm{TR}(u,e) = w$, and those w for which this is not the case.

$$E_{r,w}[B_r(w)] = \sum_{\substack{w=\mathrm{TR}(u,e)\\ u\in L'}} E[B_r(w)] + \sum_{\substack{w\neq\mathrm{TR}(u,e)\\ u\in L'}} E[B_r(w)]$$

By taking $m = k+d+2$ the first term can be bounded by $1/4$. The second term is bounded by $2^m 2^{-n-2} \leq 1/4$. Going back to probability notation, we have for any $u \in A^{=n}$

$$\Pr_{r,e}[B_r(\mathrm{TR}(u,e)) = 1] - \Pr_{r,w}[B_r(w) = 1] \geq 1/2.$$

It follows by the hybrid argument that there is an $i \in [m]$ such that

$$\Pr_{x,r,r'}[B_r(\hat{u}_1(x)\ldots\hat{u}_{i-1}(x)\hat{u}(x)r]' - \Pr_{x,r,r',b}[B_r(\hat{u}_1(x)\ldots\hat{u}_{i-1}(x)br')] \geq 1/2m \tag{3}$$

Let $F(x,b,r') = \hat{u}_1(x)\ldots\hat{u}_{i-1}(x)br'$. Our algorithm to approximate $\hat{u}$ will do the following: on input x, choose uniformly at random b, r, r' and evaluate $B_r(F(x,b,r'))$; if this evaluates to 1, then output b, otherwise output $1-b$. Call the output of this algorithm $g_b(x,r,r')$. It follows from equation (3) that

$$\Pr_{x,b,r,r'}[\hat{u}(x) = g_b(x,r,r')] \geq 1/2 + 1/2m$$

The rest of the argument now proceeds as in the proof for relativized language compression in AM (Theorem 3 in [BLvM04]), to show that the computation of $g_b(x,r,r')$ can be approximated by an AM algorithm. □

Note that the proof of Theorem 17 relativizes. We will make use of this fact in observing that the argument of [ZL70] outlined above can be used with respect to any set A of pairs of strings, not just the set of pairs with complexity at most m.

Theorem 18 *There is a polynomial $p(n)$ such that for any set $A \subset \mathbb{B}^* \times \mathbb{B}^*$ and all $\langle x,y\rangle \in A^{=n}$*

$$\log\|A^{=n}\| \geq \mathrm{CAMD}^{p,A^{=n}}(x) + \mathrm{CAMD}^{p,A^{=n}}(y|x) - O(\log^3 n).$$

Furthermore, if $A \in$ NP then then there is a polynomial $q(n)$ such that

$$\log\|A^{=n}\| \geq \mathrm{CAMD}^{q}(x) + \mathrm{CAMD}^{q}(y|x) - O(\log^3 n).$$

Proof: We follow the proof of symmetry of information in the resource unbounded case, as outlined above. Now the set A takes the place of the set $\{\langle x,y\rangle : \mathrm{C}(x,y) \leq m\}$ used before. Fix n and $\langle\alpha,\beta\rangle \in A^{=n}$. Denote $m = \log\|A^{=n}\|$ and $A_x = \{y : (x,y) \in A^{=n}\}$. Membership in the set A_x can be decided in polynomial time given x and the oracle $A^{=n}$. As $\beta \in A_\alpha$ it follows from Theorem 8 that $\mathrm{CAMD}^{q,A^{=n}}(\beta|\alpha) \leq \log\|A_\alpha\| + O(\log^3 n)$.

Now consider the set $B_k = \{x : \|A_x\| \geq 2^k\}$. Let k^* be such that $2^{k^*} \leq \|A_\alpha\| < 2^{k^*+1}$. Then $\alpha \in B_{k^*}$. Again by the approximate lower bound counting property of AM, as shown in [Bab85], there is a predicate Q (computable in polynomial time given the oracle $A^{=n}$) such that

- If $x \in B_k$ then $\Pr_r[\exists y Q(x,y,r) = 1] \geq 2/3$
- If $x \notin B_{k-1}$ then $\Pr_r[\exists y Q(x,y,r) = 1] \leq 1/3$

Thus if $\Pr_r[\exists y\ Q(x,y,r) = 1] > 1/3$ then $x \in B_{k-1}$. However $\|A^{=n}\| = 2^m$ implies that $\|B_{k-1}\| \leq 2^{m-k+1}$. Now by Theorem 8 we obtain $\mathrm{CAMD}^{q,A^{=n}}(\alpha) \leq m - k^* + O(\log^3 n)$.

Putting the above together we have

$$\mathrm{CAMD}^{q,A^{=n}}(\alpha) + \mathrm{CAMD}^{q,A^{=n}}(\beta|\alpha) \leq m - k^* + k^* + O(\log^3 n) \leq m + O(\log^3 n)$$

which gives the first statement of the theorem.

To prove the "furthermore", note that approximate lower bound counting of NP sets can be done in AM [Bab85], and apply Theorem 17 to give the bound on (unrelativized) CAMD complexity of NP sets. □

Corollary 19 *For any set $A \subset \mathbb{B}^* \times \mathbb{B}^*$ and any polynomial $p(n)$ there is a polynomial q such that for all but at most a $1/n$ fraction of $\langle x, y\rangle \in A^{=n}$,*

$$\mathrm{CAMD}^{p(n),A^{=n}}(x,y) \geq \mathrm{CAMD}^{q(x),A^{=n}}(x) + \mathrm{CAMD}^{q,A^{=n}}(y|x) - O(\log^3 n).$$

Furthermore, if $A \in$ NP then

$$\mathrm{CAMD}^{p(n)}(x,y) \geq \mathrm{CAMD}^{q}(x) + \mathrm{CAMD}^{q}(y|x) - O(\log^3 n).$$

Proof: For all but at most a $1/n$ fraction of $\langle x, y\rangle \in A^{=n}$ we have $\mathrm{CAMD}^{p(n),A^{=n}}(x,y) \geq \log \|A^{=n}\| - O(\log n)$. Applying Theorem 18 we get the first statement of the corollary. Applying the "furthermore" of Theorem 18 gives the furthermore here. □

Theorem 20 *For any strings $x, y \in \mathbb{B}^n$, and polynomial $p(n)$ there is a polynomial $q(n)$ such that*

$$\mathrm{C}^{p}(x,y) \geq \mathrm{CAMD}^{q}(x) + \mathrm{CAMD}^{q}(y|x) - O(\log^3 n).$$

Proof: Fix a pair of strings $\langle \alpha, \beta\rangle$. Let $n = |\alpha| + |\beta|$, and suppose that $\mathrm{C}^p(\alpha,\beta) = m$. Consider the set $A = \{\langle x, y\rangle : C^p(x,y) \leq m\}$. As membership in A can be decided in nondeterministic polynomial time, we may invoke the "furthermore" of Theorem 18 to give

$$\log \|A\| \geq \mathrm{CAMD}^{q}(\alpha) + \mathrm{CAMD}^{q}(\beta|\alpha) - O(\log^3 n)$$

for some polynomial q.

On the other hand, $\|A\| \leq 2^{m+1}$, and the theorem is proven. □

From Theorem 20 we obtain as a corollary a result of [LW95], up to an additive $O(\log^3(n))$ factor.

Corollary 21 *If* P = NP *then for any polynomial $p = p(n)$ there is a polynomial $q = q(n)$ such that for all $x, y \in \mathbb{B}^n$*

$$\mathrm{C}^{p}(x,y) \geq \mathrm{C}^{q}(x) + \mathrm{C}^{q}(y|x) - O(\log^3 n)$$

It remains an interesting open problem if polynomial time symmetry of information for printing complexity holds under a weaker assumption than P = NP.

Acknowledgments AR thanks Harry Buhrman and Lance Fortnow for helpful comments on [BF95] and [BFL02] and on the history of the problems under consideration. TL would like to thank Harry Buhrman and Dieter van Melkebeek for helpful comments and conversations about Section 5.

References

[ABK+02] E. Allender, H. Buhrman, M. Koucky, D. van Melkebeek, and D. Ronneburger. Power from random strings. In *Proceedings of the 47th IEEE Symposium on Foundations of Computer Science*, pages 669–678. IEEE, 2002.

[Bab85] L. Babai. Trading group theory for randomness. In *Proceedings of the 17th ACM Symposium on the Theory of Computing*, pages 421–429. ACM, 1985.

[BF95] H. Buhrman and L. Fortnow. Distinguishing complexity and symmetry of information. Technical Report TR-95-11, Department of Computer Science, The University of Chicago, 1995.

[BFL02] H. Buhrman, L. Fortnow, and S. Laplante. Resource bounded Kolmogorov complexity revisited. *SIAM Journal on Computing*, 31(3):887–905, 2002.

[BLM00] H. Buhrman, S. Laplante, and P.B. Miltersen. New bounds for the language compression problem. In *Proceedings of the 15th IEEE Conference on Computational Complexity*, pages 126–130. IEEE, 2000.

[BLvM04] H. Buhrman, T. Lee, and D. van Melkebeek. Language compression and pseudorandom generators. To appear in 19th IEEE Conference on Computational Complexity, 2004.

[FGM+89] M. Fürer, O. Goldreich, Y. Mansour, M. Sipser, and S. Zachos. On completeness and soundness in interactive proof systems. In S. Micali, editor, *Randomness and Computation*, volume 5 of *Advances in Computing Research*, pages 429–442. JAI Press, Greenwich, 1989.

[FK96] L. Fortnow and M. Kummer. On resource-bounded instance complexity. *Theoretical Computer Science A*, 161:123–140, 1996.

[Hel86] H. Heller. On relativized exponential and probabilistic complexity classes. *Information and Compuation*, 71:231–243, 1986.

[IR89] R. Impagliazzo and S. Rudich. Limits on the provable consequences of one-way functions. In *Proceedings of the 21st ACM Symposium on the Theory of Computing*, pages 41–61. ACM, 1989.

[JSV97] T. Jiang, J. Seiferas, and P. Vitányi. Two heads are better than two tapes. *Journal of the ACM*, 44(2):237–256, 1997.

[Kol65] A.N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems Information Transmission*, 1(1):1–7, 1965.

[LM93] L. Longpré and S. Mocas. Symmetry of information and one-way functions. *Information Processing Letters*, 46(2):95–100, 1993.

[LV97] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and its Applications*. Springer-Verlag, New York, second edition, 1997.

[LW95] L. Longpré and O. Watanabe. On symmetry of information and polynomial time invertibility. *Information and Compuation*, 121(1):14–22, 1995.

[NW94] N. Nisan and A. Wigderson. Hardness vs. randomness. *Journal of Computer and System Sciences*, 49:149–167, 1994.

[Sha02] R. Shaltiel. Recent developments in explicit construction of extractors. *Bulletin of the European Association for Theoretical Computer Science*, 77:67–95, 2002.

[Sip83] M. Sipser. A complexity theoretic approach to randomness. In *Proceedings of the 15th ACM Symposium on the Theory of Computing*, pages 330–335. ACM, 1983.

[Tre99] L. Trevisan. Construction of extractors using pseudo-random generators. In *Proceedings of the 31st ACM Symposium on the Theory of Computing*, pages 141–148. ACM, 1999.

[VV02] N. Vereshchagin and P. Vitányi. Kolmogorov's structure function with an application to the foundations of model selection. In *Proceedings of the 47th IEEE Symposium on Foundations of Computer Science*, pages 751–760. IEEE, 2002.

[ZL70] A. Zvonkin and L. Levin. The complexity of finite objects and the algorithmic concepts of information and randomness. *Russian Mathematical Surveys*, 25:83–124, 1970.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'