Electronic Colloquium on Computational Complexity - Reports Series 1997 - available via:
ECCC
TR97-030
FTP: ftp.eccc.uni-trier.de:/pub/eccc/
WWW: http://www.eccc.uni-trier.de/eccc/
Email: ftpmail@ftp.eccc.uni-trier.de with subject 'help eccc'

On Nondeterminism versus Randomness for Read-Once Branching Programs

Martin Sauerhoff*
FB Informatik, LS II, Univ. Dortmund, 44221 Dortmund, Germany
sauerhoff@ls2.informatik.uni-dortmund.de

Abstract

Randomized branching programs are a probabilistic model of computation defined in analogy to the well-known probabilistic Turing machines. In this paper, we present complexity theoretic results for randomized read-once branching programs.

Our main result shows that nondeterminism can be more powerful than randomness for read-once branching programs. We present a function which is computable by nondeterministic read-once branching programs of polynomial size, while on the other hand randomized read-once branching programs for this function with two-sided error at most $21/256$ have exponential size.

The same function exhibits an exponential gap between the randomized read-once branching program sizes for different constant worst-case errors, which shows that there is no "probability amplification" technique for read-once branching programs which allows to decrease the error to an arbitrarily small constant by iterating probabilistic computations.

Keywords: Branching programs, read-once branching programs, nondeterminism, randomness, lower bounds.

*This work has been supported by DFG grant We 1066/8-1.

1 Introduction

Branching programs are a theoretically and practically interesting data structure for the representation of Boolean functions. In complexity theory, among other problems, lower bounds for the size of branching programs for explicitly defined functions and the relations of the various branching program models are investigated.

A branching program (BP) on the variable set $\{x_1, \ldots, x_n\}$ is a directed acyclic graph with one source and two sinks, the latter labelled by the constants 0 and 1. Each non-sink node is labelled by a variable x_i and has exactly two outgoing edges labelled by 0 or 1. This graph represents a Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ in the following way. To compute $f(a)$ for some input $a \in \{0,1\}^n$, start at the source node. For a non-sink node labelled by x_i, check the value of this variable and follow the edge which is labelled by this value (this is called a "test" of the variable x_i). Iterate this until a sink node is reached. The value of f on input a is the value of the reached sink. The *size* of a branching program G is the number of its non-sink nodes and is denoted by $|G|$.

We can also assign a Boolean function to every node of a branching program, not only to the source. Furthermore, note that an edge of a branching program can be regarded as an assignment of a variable, and each path corresponds to a sequence of assignments of variables.

Branching programs are a sequential model of computation. Sequences of functions which can be computed by polynomial size branching programs can also be computed within logarithmic space on non-uniform Turing machines and vice versa (Pudlák and Zák, [16], see also [20]). Hence, a non-linear lower bound on the size of branching programs would amount to a major breakthrough in complexity theory.

Since the lower bound techniques presently known are too weak to prove such bounds, one has turned to restricted variants of branching programs. Read-k-times branching programs are branching programs with the restriction that on each path from the source to a sink each variable is allowed to be tested at most k times. This model is sometimes termed *syntactic* read-k-times BP, in contrast to the "non-syntactic" variant with the restriction that only on each *consistent* path from the source to a sink each variable is allowed to be tested at most k times (a path is called consistent if all assignments of variables on it are consistent). Exponential lower bounds on the size of syntactic read-k-times BPs have been independently proved by Okolnishnikova [14] for $k \leq c \log n / \log\log n$, $c < 1$ arbitrarily chosen, and by Borodin, Razborov and Smolensky [6] even for nondeterministic syntactic read-k-times BPs and $k \leq c \log n$, for an appropriate constant c.

Here we focus on the case $k = 1$, i. e. read-once branching programs. This is the variant of branching programs for which the first exponential lower bound could be established ([21], [23]). By now, the theory of deterministic read-once branching programs is well understood, and there is a large collection of interesting lower bound results (Razborov [17] gives an overview, for a summary of proof techniques for lower bounds, see [19]).

We mention another variant of restricted branching programs which will turn up in the sequel. OBDDs (ordered binary decision diagrams), introduced by Bryant [7], are even further restricted than read-once-branching programs, but have nevertheless turned out to be extremely

useful in practice. An OBDD is a read-once branching program with an additional ordering of variables. On each path from the source to a sink, the variables have to be tested according to this ordering. Lower bounds for OBDDs have been proved, e. g. , by Bryant [8], Hosaka, Takenaga and Yajima [11] and Bollig, Sauerhoff, Sieling and Wegener [5].

In this paper we are concerned with randomized branching programs, i. e. branching programs with additional "coin-tossing nodes". We will give a formal definition of this model in the next section. In the context of Turing machines, randomized models have been studied since the introductory work of Gill [10]. But to clarify the relations of the respective complexity classes among each other and to the polynomial hierarchy belongs to the famous open problems in complexity theory. In spite of this, these questions could be solved for some restricted computation models, most important perhaps communication protocols (see [4], [15]). By the analysis of these restricted models we hope to be able to improve our tools for proving lower bounds and thus also to gain deeper insights into the structure of the more general models.

It is therefore natural to ask what can be done for randomized variants of restricted branching programs. Ablayev and Karpinski [2] have made the first step by presenting a function which is computable by randomized OBDDs of polynomial size, but for which deterministic OBDDs have exponential size. In [3], they used a modified version of this function and showed that for it even the size of *nondeterministic* OBDDs is exponentially larger than the size of randomized OBDDs. On the other hand, Ablayev [1] and the author [18] managed to prove exponential lower bounds on the size of randomized OBDDs for certain functions representable by nondeterministic OBDDs of polynomial size. Altogether, it follows that the analogues of the classes NP and BPP for OBDDs are incomparable.

For read-once branching programs, the relation between nondeterminism and randomness has been open so far, as noted in the paper of Jukna, Razborov, Savický and Wegener [12].

In the technical report [18] the author has already shown that randomness can be more powerful than nondeterminism for read-once branching programs. More precisely, we have an example of a function with exponential nondeterministic read-once BP size on the one hand and polynomial randomized read-once BP size on the other. The cited paper also introduces a lower bound technique for randomized read-k-times BPs. By this technique, an exponential lower bound on the size of randomized read-k-times BPs for $k \leq c \log n$, c an appropriate constant, could be established. We note that the function considered in this case also has exponential nondeterministic read-k-times BP size (as proved by Borodin, Razborov and Smolensky [6]). It is not hard to show that the same holds for the complement of the function.

In the present work we exhibit a function that is "simple enough" to be computable by nondeterministic read-once branching programs of polynomial size, but nevertheless can be proved to have exponential size for randomized read-once branching programs by the lower bound method from [18]. As a consequence, we obtain that nondeterminism and randomness are incomparable for read-once branching programs if the error allowed for the randomized programs is not too large.

The rest of the paper is organized as follows. In Section 2, we formally define randomized branching programs. In Section 3 we give a summary of the lower bound technique which we use. Our main result is proved in Section 4.

2 Definitions and Basic Facts

In this section, we give the definitions of nondeterministic and randomized variants of general branching programs. It is easy to derive appropriate variants for the various restricted branching program models, especially for read-k-times branching programs.

For the introduction of non-deterministic branching programs we follow Meinel [13].

Definition 1: Let Ω be a set of binary Boolean operators. An Ω-*branching program* is a branching program which may contain nodes labelled by a function $\omega \in \Omega$ and which have two unlabelled outgoing edges. We define the semantics of such an Ω-branching program by inductively assigning a function to each node. The 0- and the 1-sink compute the respective constant functions. Let v be a non-sink node labelled by a function $\omega \in \Omega$ with successors v_1 and v_2, and let f_1 and f_2, resp. , be the functions represented by the successors. Then v represents the function $\omega(f_1, f_2)$. Now let v be a non-sink node labelled by a variable x_i, where the functions f_0 and f_1 are represented at the nodes reached via the 0- and 1-edge, resp. Then v represents the function $\neg x_i \cdot f_0 \vee x_i \cdot f_1$ as in a usual branching program. The size of an Ω-branching program is the number of all its non-sink nodes.

Nondeterministic branching programs are $\{\vee\}$-branching programs in the sense of this definition, ordinary branching programs are obtained by choosing $\Omega = \emptyset$. The class of sequences of Boolean functions which are computable by polynomial size nondeterministic read-once branching programs is denoted by NP-BP1.

Definition 2: A *randomized branching program* G syntactically is a branching program with two disjoint sets of variables $x_1, \ldots, x_n$ and $z_1, \ldots, z_r$. We will call the latter "stochastic" variables. Let g be the function on $n+r$ variables represented by G as a deterministic branching programm.

We say that G as a randomized branching program represents a function $f\colon \{0,1\}^n \to \{0,1\}$ with

- *one-sided error* at most ε, $0 \leq \varepsilon < 1$, if for all $x \in \{0,1\}^n$ it holds that
  $\Pr\{g(x,z) = 0\} = 1, \quad$ if $f(x) = 0$;
  $\Pr\{g(x,z) = 1\} \geq 1 - \varepsilon, \quad$ if $f(x) = 1$;
- *two-sided error* at most ε, $0 \leq \varepsilon < 1/2$, if for all $x \in \{0,1\}^n$ it holds that
  $\Pr\{g(x,z) = f(x)\} \geq 1 - \varepsilon$.

In these expressions, z is an assignment to the stochastic variables which is chosen according to the uniform distribution from $\{0,1\}^r$.

A *randomized read-once BP* is a randomized branching program with the restriction that on each path from the source to a sink, each variable x_i and each variable z_i is tested at most once. For a *randomized OBDD,* an ordering on the variables $x_1, \ldots, x_n$ and $z_1, \ldots, z_r$ is given.

In analogy to the well-known complexity classes for Turing machines, let RP$_\varepsilon$-BP1 be the class of sequences of functions computable by polynomial size randomized read-k-times branching programs with one-sided error at most ε, $\varepsilon < 1$. Let BPP$_\varepsilon$-BP1 be the class of sequences of functions computable by polynomial size randomized read-once branching programs with two-sided error at most ε, $\varepsilon < 1/2$. Furthermore, let

$$\text{RP-BP1} := \bigcup_{\varepsilon \in [0,1)} \text{RP}_\varepsilon\text{-BP1},$$
$$\text{BPP-BP1} := \bigcup_{\varepsilon \in [0,\frac{1}{2})} \text{BPP}_\varepsilon\text{-BP1}.$$

Analogous classes can be defined for general branching programs and OBDDs (we append suffixes "-BP" and "-OBDD", resp., to the names instead of "-BP1"). Finally, for each of the considered complexity classes $\mathcal{C}$ let co-$\mathcal{C}$ be the class of sequences of functions (f_n) for which $(\neg f_n) \in \mathcal{C}$.

We can adapt the well-known technique of iterating probabilistic computations (called "probability amplification") to improve the error probability of randomized branching programs and randomized OBDDs. We obtain, e. g., that for all constant ε and ε' with $0 < \varepsilon \leq \varepsilon' < 1$ it holds that

$$\text{RP}_\varepsilon\text{-BP} = \text{RP}_{\varepsilon'}\text{-BP} \quad \text{and} \quad \text{RP}_\varepsilon\text{-OBDD} = \text{RP}_{\varepsilon'}\text{-OBDD}\,.$$

This has been proved in [18]. We will see in Section 4 that an analogous assertion for read-once BPs does not hold.

As for Turing machines, we have RP-BP1 $\subseteq$ NP-BP1. It is an open problem if this inclusion is proper. It has been shown in [18] that coRP-OBDD $\setminus$ NP-BP1 $\neq \emptyset$, and thus BPP-BP1 $\not\subseteq$ NP-BP1.

For the sake of completeness, we restate the respective theorem. The function considered is called PERM and is defined on an $n \times n$-matrix $X = (x_{ij})_{1 \leq i,j \leq n}$ of Boolean variables. Let PERM$(X) = 1$ if and only if X is a permutation matrix, i. e. if each row and each column contains exactly one entry equal to 1.

Theorem 1 (Sauerhoff 1997):

(1) PERM $\in$ coRP$_{\varepsilon(n)}$-OBDD *for all* $\varepsilon(n) \in [0,1)$ *with* $\varepsilon(n)^{-1} = O(\text{poly}(n))$, *but*

(2) PERM $\notin$ NP-BP1.

It is easy to improve this result to show that BPP-BP1 $\not\subseteq$ (NP-BP1 $\cup$ coNP-BP1).
The function 2PERM: $\{0,1\}^{2n^2} \to \{0,1\}$, defined on two Boolean $n \times n$-matrices X and Y by 2PERM(X,Y) := PERM$(X) \wedge \neg$PERM(Y) obviously is contained in the class BPP-BP1 but neither in NP-BP1 nor in coNP-BP1.

3 A Lower Bound Technique for Randomized Read-Once BPs

As mentioned in the introduction, we are going to apply the technique for proving lower bounds on the size of randomized read-k-times BPs with two-sided error developped in [18]. In this section, we restate the necessary definitions and the main result for the special case $k = 1$.

The proof technique is an extension of techniques of Borodin, Razborov and Smolensky [6] and Okolnishnikova [14]. It relates the number of nodes of a read-once branching program to the number of rectangles for which the considered function computes constant values. The definition of rectangles is given below.

Definition 3 (Rectangle): Let X be a set of variables, $n := |X|$. Let (X_1, X_2) be a balanced partition of X, i. e. $X = X_1 \cup X_2$, $X_1 \cap X_2 = \emptyset$ and $||X_1| - |X_2|| \leq 1$. Then a set $R \subseteq 2^{X_1} \times 2^{X_2}$ of assignments is called *rectangle in 2^X with respect to the partition (X_1, X_2)*.

This is the type of rectangles considered also in communication complexity theory. The definition coincides with the definition of (k, p)-rectangles of Borodin, Razborov and Smolensky used in [18] if we let $k = 1$ and $p = 2$.

Notation: We occasionally identify assignments to variables and Boolean vectors if the set (and order) of variables is clear from the context.

To show a lower bound on the size of randomized read-once BPs we will establish the following two properties of the considered function f:

(i) The number of 1-inputs for f is bounded from below by a positive (non-zero) constant.

(ii) For an arbitrary balanced rectangle, the number of 0-inputs for f in this rectangle is always at least a constant fraction of the number of 1-inputs in the rectangle.

The theorem below (which is proved in [18]) makes this more precise.

Theorem 2: *Let $f : \{0,1\}^n \to \{0,1\}$ be defined on the variable set X, $|X| = n$. Assume that there is a probability distribution μ on 2^X such that for every rectangle R which belongs to a balanced partition of X it holds that*

$$\mu(R \cap f^{-1}(0)) \geq \alpha \cdot \mu(R \cap f^{-1}(1)) - \delta(n),$$

where α is a constant and δ a real-valued function.

Then it holds for every randomized read-once BP G for f with two-sided error at most ε that

$$|G| \geq \frac{1}{2n} \left(\frac{\alpha \cdot \mu(f^{-1}(1)) - (1+\alpha) \cdot \varepsilon}{\delta(n)} \right)^{1/2}.$$

In the applications of this theorem, $\delta(n)$ will be exponentially small in n.

4 The Main Result

In this section, we prove an exponential lower bound on the randomized read-once BP size of a function which is computable by nondeterministic read-once BPs of polynomial size. We obtain that NP-BP1 $\not\subseteq$ BPP$_\varepsilon$-BP1 for "small" error ε. The proof of this fact turns out to be much harder than the proof of the contrary result that BPP-BP1 $\not\subseteq$ NP-BP1 mentioned above (Theorem 1).

We consider the following function.

Definition 4: Define $\mathrm{ModSum}\colon \{0,1\}^{n^2} \to \{0,1\}$ on the $n \times n$-matrix $X = (x_{i,j})_{1 \le i,j \le n}$ of Boolean variables. Let

$$\mathrm{ModSum}(X) := \mathrm{RowTest}(X) \wedge \mathrm{RowTest}(X^T),$$

where $\mathrm{RowTest}\colon \{0,1\}^{n^2} \to \{0,1\}$ is defined by

$$\mathrm{RowTest}(X) := \left[\sum_{i=1}^{n} [x_{i,1} + \cdots + x_{i,n} \equiv 0 \bmod 3] \equiv 0 \bmod 2 \right].$$

(By the expression $[P]$, P a predicate, we denote the Boolean function which is equal to 1 iff P is true.)

In order to apply Theorem 2 from the last section, we show in Lemma 1 below that the function ModSum has "many" 1-inputs. Furthermore, we have to verify that each rectangle with respect to a balanced partition of the input variables contains at least a "certain fraction" of 0-inputs for ModSum. Lemma 2 and Lemma 3 prepare the proof of this fact.

Lemma 1: $|\mathrm{ModSum}^{-1}(1)| \ge \rho \cdot 2^{n^2}$, $\rho := \frac{21}{128} > 0.164$.

Proof: Consider an arbitrary partial assignment a to the variables in $X = (x_{ij})_{1 \le i,j \le n}$, which fixes all variables with the exception of a 3×3-submatrix, e. g. in the upper left corner of X. Then it holds that

$$\begin{aligned} \mathrm{ModSum}_a(X_a) \;=\; & \left[\sum_{i=1}^{3} [x_{i,1} + x_{i,2} + x_{i,3} + r_i \equiv 0 \bmod 3] \equiv p_r \bmod 2 \right] \wedge \\ & \left[\sum_{i=1}^{3} [x_{1,i} + x_{2,i} + x_{3,i} + c_i \equiv 0 \bmod 3] \equiv p_c \bmod 2 \right], \end{aligned}$$

where ModSum_a denotes the subfunction of ModSum obtained by substituting the variables according to a, $X_a = (x_{ij})_{1 \le i,j \le n}$ is the matrix of remaining free variables, and the constants $p_r, p_c \in \{0,1\}$ and $r_1, r_2, r_3, c_1, c_2, c_3 \in \mathbb{Z}_3$ depend on the assignment a.

For all possible values of $p_r, p_c \in \{0,1\}$ and $r_1, r_2, r_3, c_1, c_2, c_3 \in \mathbb{Z}_3$ we can by means of a computer count the number of Boolean 3×3-matrices for which ModSum_a, a the assignment

belonging to the constants, outputs 1. We obtain a minimum number of 84 1-inputs if

$$\begin{aligned}
&p_r = 0,\ p_c = 1 \quad \text{and} \quad r_i = c_i = 1, \quad i = 1,2,3 \quad \textit{or}\\
&p_r = 0,\ p_c = 1 \quad \text{and} \quad r_i = c_i = 2, \quad i = 1,2,3 \quad \textit{or}\\
&p_r = 1,\ p_c = 0 \quad \text{and} \quad r_i = c_i = 1, \quad i = 1,2,3 \quad \textit{or}\\
&p_r = 1,\ p_c = 0 \quad \text{and} \quad r_i = c_i = 2, \quad i = 1,2,3.
\end{aligned}$$

Since there are $2^{3^2} = 512$ choices for the values of x_{ij}, $1 \leq i,j \leq 3$, altogether, the claim follows. □

Lemma 2: *Let (X_1, X_2) be an arbitrary balanced partition of the variable set X. Then it holds that there is a set $I \subseteq \{1,\ldots,n\}$ with $|I| \geq n/4$ such that*

$$2 \leq |\{x_{i,1},\ldots,x_{i,n}\} \cap X_1| \leq n-2 \quad \textit{for all } i \in I;$$

or

$$2 \leq |\{x_{1,i},\ldots,x_{n,i}\} \cap X_1| \leq n-2 \quad \textit{for all } i \in I.$$

Proof: Define $s_i := |\{j \mid j \in \{1,\ldots,n\} \wedge x_{i,j} \in X_1\}|$, $1 \leq i \leq n$, and

$$\begin{aligned}
L &:= \{i \mid s_i < 2\},\\
H &:= \{i \mid s_i > n-2\}.
\end{aligned}$$

Since the given partition is balanced, it holds that

$$\lfloor n^2/2 \rfloor \leq \sum_{i=1}^{n} s_i \leq \lceil n^2/2 \rceil.$$

The union $L \cup H$ contains exactly the indices of rows of the matrix X which do not fulfill the first assertion in the claim above. We show that if the first assertion (for the rows) is not fulfilled, then the second (for the columns) holds.

Assume in the following that the first assertion does not hold, i. e. $|L \cup H| \geq \frac{3}{4}n + 1$. We have

$$|H| \leq \frac{\lceil n^2/2 \rceil}{n-1} \leq \frac{n^2/2 + 1/2}{n-1} \leq n/2 + 1,$$

for n large enough. If we swap the roles of X_1 and X_2, we also obtain $|L| \leq n/2 + 1$. Taking the assumption into account, it follows that $|L| \geq n/4$ and $|H| \geq n/4$. Hence, there is a set $I \subseteq \{1,\ldots,n\}$ with $|I| \geq n/2$ such that for each $i \in I$ the column i of X contains $n/4$ variables from X_1 and $n/4$ variables from X_2. Therefore, the second assertion of the claim is fulfilled for this set I. □

Definition 5: Let $V := \{0,1\}^2$. Let $\varphi\colon V \to \mathbb{Z}_3$ be defined by $\varphi(x,y) := x + y$, where $x, y \in \{0,1\}$ (thus we have $\varphi^{-1}(0) = \{(0,0)\}$, $\varphi^{-1}(1) = \{(0,1),(1,0)\}$ and $\varphi^{-1}(2) = \{(1,1)\}$).

For arbitrary $c_0 \in \{0,1\}$ and $c_1, \ldots, c_n \in \mathbb{Z}_3$ define $\mathrm{RowTestComm}_{c_0,c_1,\ldots,c_n} \colon V^n \times V^n \to \{0,1\}$ by

$$\mathrm{RowTestComm}_{c_0,c_1,\ldots,c_n}(x,y) := \left[\sum_{i=1}^{n} [\varphi(x_i) + \varphi(y_i) + c_i \equiv 0 \bmod 3] \equiv c_0 \bmod 2 \right],$$

where $x, y \in V^n$.

Definition 6 (Discrepancy): Let finite sets X and Y and a function $f \colon X \times Y \to \{0,1\}$ be given. Then we define the *discrepancy of f with respect to a rectangle R, $R = A \times B$ and $A \subseteq X$, $B \subseteq Y$,* by

$$\mathrm{Disc}(f,R) := \frac{1}{|X||Y|} \cdot \left| |f^{-1}(1) \cap R| - |f^{-1}(0) \cap R| \right|.$$

By $\mathrm{Disc}(f)$ we denote the maximum of $\mathrm{Disc}(f,R)$ taken over all choices of rectangles R in $X \times Y$.

Lemma 3: *For arbitrary $c_0 \in \{0,1\}$ and $c_1, \ldots, c_n \in \mathbb{Z}_3$, it holds that*

$$\mathrm{Disc}(\mathrm{RowTestComm}_{c_0,c_1,\ldots,c_n}) \leq \left(\sqrt{14}/4\right)^n.$$

Proof: The technique used for this proof is the same as in the well-known proof of the lower bound on the probabilistic communication complexity of the inner-product function (see, e. g. , [9]). Define the $4^n \times 4^n$-matrix M, $M = (m(x,y))_{x,y \in V^n}$, by

$$m(x,y) := \begin{cases} 1, & \text{if } \mathrm{RowTestComm}_{c_0,c_1,\ldots,c_n}(x,y) = 1; \\ -1, & \text{otherwise.} \end{cases}$$

Let $R = S \times T$, with $S, T \subseteq V^n$, be an arbitrary rectangle. We show that

$$\begin{aligned} \mathrm{Disc}(\mathrm{RowTestComm}_{c_0,c_1,\ldots,c_n}, R) &= \frac{1}{|V^{2n}|} \left| \sum_{(x,y) \in R} m(x,y) \right| \\ &= \frac{1}{4^{2n}} \cdot |1_S^T \cdot M \cdot 1_T| \leq (\sqrt{14}/4)^n \end{aligned}$$

where 1_S and 1_T are the characteristic vectors of S and T, respectively. To establish this upper bound, we show that $\|M\|_2$, the spectral norm of M, is small compared to 4^{2n}. The first step in the proof is to compute the entries of $\tilde{M} = (\tilde{m}(x,y))_{x,y \in V^n}$, defined by $\tilde{M} := M^T M$. It holds that $\|M\|_2 = \sqrt{\lambda_{\max}}$, where $\lambda_{\max}$ is the largest eigenvalue of $\tilde{M}$ (see, e. g. , [22]). Note that all eigenvalues of $\tilde{M}$ are real and non-negative. The second step will be to derive an upper bound on $\lambda_{\max}$.

First step: Let $m(x)$ be the column of M with index $x \in V^n$. It holds that

$$\tilde{m}(x,y) = m(x)^T m(y) = \sum_{z \in V^n} m(x,z) m(y,z).$$

We evaluate this sum by counting the number of 1's and (-1)'s, i. e. we compute

$$N_1(x,y) := |\{z \in V^n \mid m(x,z)m(y,z) = 1\}|, \quad \text{and}$$
$$N_{-1}(x,y) := |\{z \in V^n \mid m(x,z)m(y,z) = -1\}|.$$

It is sufficient to determine $N_1(x,y)$, since $N_{-1}(x,y) = 4^n - N_1(x,y)$. It holds that

$$\begin{aligned} m(x,z)m(y,z) = 1 \Leftrightarrow & \sum_{i=1}^{n} [\varphi(x_i) + \varphi(z_i) + c_i \equiv 0 \bmod 3] \equiv \\ & \sum_{i=1}^{n} [\varphi(y_i) + \varphi(z_i) + c_i \equiv 0 \bmod 3] \bmod 2 \\ \Leftrightarrow & \sum_{i=1}^{n} ([\varphi(x_i) + \varphi(z_i) + c_i \equiv 0 \bmod 3] - \\ & [\varphi(y_i) + \varphi(z_i) + c_i \equiv 0 \bmod 3]) \equiv 0 \bmod 2 \end{aligned}$$

For $x, y \in V^n$, $i \in \{1, \ldots, n\}$ and $z' \in V$ define

$$S_i(z') := ([\varphi(x_i) + \varphi(z') + c_i \equiv 0 \bmod 3] - [\varphi(y_i) + \varphi(z') + c_i \equiv 0 \bmod 3]) \bmod 2.$$

We have to compute the number of vectors $z \in V^n$ with

$$S_1(z_1) + \cdots + S_n(z_n) \equiv 0 \bmod 2.$$

Let $D := \{i \mid \varphi(x_i) \neq \varphi(y_i)\}$ and $d := |D|$. For $i \notin D$, it holds that $S_i(z') = 0$ for arbitrary $z' \in V$, which leads to $|V| = 4$ possible choices for z'. Hence, for all z_i with $i \notin D$ we have 4^{n-d} choices altogether.

Now we consider the case $i \in D$, i. e. we have $\varphi(x_i) \neq \varphi(y_i)$. It holds that $S_i(z') = 0$ if and only if

$$(\varphi(x_i) + \varphi(z') + c_i) \bmod 3 \in \{1,2\} \ \wedge \ (\varphi(y_i) + \varphi(z') + c_i) \bmod 3 \in \{1,2\}.$$

We count the number of z' satifying this condition:

$(\varphi(x_i) + c_i) \bmod 3$	$(\varphi(y_i) + c_i) \bmod 3$	possible $\varphi(z')$	number of $z' \in V$
0	1	$\{1,2\} \cap \{0,1\} = \{1\}$	2
0	2	$\{1,2\} \cap \{0,2\} = \{2\}$	1
1	0	$\{0,1\} \cap \{1,2\} = \{1\}$	2
1	2	$\{0,1\} \cap \{0,2\} = \{0\}$	1
2	0	$\{0,2\} \cap \{1,2\} = \{2\}$	1
2	1	$\{0,2\} \cap \{0,1\} = \{0\}$	1

Let $P := \{i \mid \{(\varphi(x_i) + c_i) \bmod 3, (\varphi(y_i) + c_i) \bmod 3\} = \{0,1\}\} \subseteq D$ and $p := |P|$. From the table above we see that

$$|S_i^{-1}(0)| = \begin{cases} 2, & \text{if } i \in P \text{ and} \\ 1, & \text{if } i \in D \backslash P, \end{cases}$$

we also have

$$|S_i^{-1}(1)| = \begin{cases} 2, & \text{if } i \in P \text{ and} \\ 3, & \text{if } i \in D \backslash P. \end{cases}$$

Now we calculate the number of choices for the z_i, $i \in D$, under the assumption that exactly k of the $S_i(z_i)$ for $i \in P$ and exactly l of the $S_i(z_i)$ for $i \in D \backslash P$ are equal to 1. By our considerations above, there are

$$\binom{p}{k} \cdot 2^k \cdot 2^{p-k} \cdot \binom{d-p}{l} \cdot 3^l \cdot 1^{d-p-l}$$

possible values for all z_i with $i \in D$. Summation of these expressions over *all* choices for $k \in \{0, \ldots, p\}$ and $l \in \{0, \ldots, d-p\}$ obviously yields 4^d. But we only need the number of z_i for which $k + l \equiv 0 \mod 2$. Hence, we have to compute

$$\sum_{k=0}^{p} \sum_{l=0}^{d-p} \binom{p}{k} \binom{d-p}{l} \cdot 2^p \cdot 3^l \cdot [k + l \equiv 0 \mod 2].$$

Substituting $(1 + (-1)^{k+l})/2$ for $[k + l \equiv 0 \mod 2]$ this sum can easily be evaluated by application of the binomial theorem, leading to the result

$$\frac{1}{2} \cdot 4^d + (-1)^d \cdot 2^{d-1} \cdot [p = 0].$$

Putting the results together, we obtain

$$N_1(x, y) = 4^{n-d} \left(\frac{1}{2} \cdot 4^d + (-1)^d \cdot 2^{d-1} \cdot [p = 0] \right),$$
$$N_{-1}(x, y) = 4^{n-d} \left(\frac{1}{2} \cdot 4^d - (-1)^d \cdot 2^{d-1} \cdot [p = 0] \right).$$

Since $m(x)^T m(y) = N_1(x, y) - N_{-1}(x, y)$, we get

$$\tilde{m}(x, y) = m(x)^T m(y) = 4^{n-d} \cdot (-1)^d \cdot 2^d \cdot [p = 0], \quad \text{where}$$
$$d := |\{i \mid \varphi(x_i) \neq \varphi(y_i)\}| \text{ and}$$
$$p := |\{i \mid \{(\varphi(x_i) + c_i) \bmod 3, (\varphi(y_i) + c_i) \bmod 3\} = \{0, 1\}\}|.$$

Second step: Having obtained a closed form for the entries of $\tilde{M}$, we are now going to derive an upper bound on the value of the largest eigenvalue $\lambda_{\max}$ of $\tilde{M}$. For the estimation of this value, we use the following simple fact from linear algebra.

Let $\|\cdot\|$ denote a vector norm on $\mathbb{C}^n$ as well as a matrix norm which is compatible with this vector norm, i. e. it holds that $\|Ax\| \leq \|A\| \cdot \|x\|$ for an arbitrary complex-valued $n \times n$-matrix A and $x \in \mathbb{C}^n$. Let A be an arbitrary complex-valued $n \times n$-matrix, λ an eigenvalue of A and x

($x \neq 0$) an eigenvector belonging to λ. Then it holds that $\|A\|\|x\| \geq \|Ax\| = \|\lambda x\| = |\lambda|\|x\|$, hence, $|\lambda| \leq \|A\|$ (where $|\cdot|$ is the absolute value in $\mathbb{C}$). For our purpose, it turns out to be useful to choose the norm defined by

$$\|A\|_\infty := \max\{\sum_{j=1}^{n} |a_{ij}| \mid i = 1, \ldots, n\},$$

where $A = (a_{ij})_{1 \leq i,j \leq n}$ is a complex-valued $n \times n$-matrix. This norm is compatible with the vector norm $\|x\|_\infty := \max\{x_i \mid 1 \leq i \leq n\}$, where $x \in \mathbb{C}^n$. (Obviously, summing column-wise instead of row-wise works as well.)

For $x, y \in V^n$ define $d(x,y) := |\{i \mid \varphi(x_i) \neq \varphi(y_i)\}|$ and $p(x,y) := |\{i \mid \{(\varphi(x_i)+c_i) \bmod 3, (\varphi(y_i)+c_i) \bmod 3\} = \{0,1\}\}|$. We calculate the sum of the absolute values of the entries in an arbitrary row $x \in V^n$ of $\tilde{M}$:

$$\sum_{y \in V^n} |\tilde{m}(x,y)| = \sum_{y \in V^n} 4^n \cdot 2^{-d(x,y)} \cdot [p(x,y) = 0] \leq \sum_{y \in V^n} 4^n \cdot 2^{-d(x,y)}$$

To get rid of the function d, we count for fixed $k \in \{0, \ldots, n\}$ the number of $y \in V^n$ for which $d(x,y) = k$. For each i there are at most 3 values y_i for which $\varphi(x_i) \neq \varphi(y_i)$, and at most 2 values y_i for which $\varphi(x_i) = \varphi(y_i)$. Hence, the number of $y \in V^n$ with $d(x,y) = k$ is at most

$$\binom{n}{k} \cdot 3^k \cdot 2^{n-k}.$$

With this estimation, we get

$$\sum_{y \in V^n} 4^n \cdot 2^{-d(x,y)} \leq \sum_{k=0}^{n} \binom{n}{k} \cdot 3^k \cdot 2^{n-k} \cdot 4^n \cdot 2^{-k} = 14^n$$

It follows that $|\lambda_{\max}| \leq \|\tilde{M}\|_\infty \leq 14^n$ and thus $\|M\|_2 = \sqrt{\lambda_{\max}} \leq \sqrt{14}^n$.

Finally, we use these results to estimate the discrepancy of $\mathrm{RowTestComm}$ with reference to the rectangle $R = S \times T$. It holds that

$$\begin{aligned}
\mathrm{Disc}(\mathrm{RowTestComm}_{c_0,c_1,\ldots,c_n}, R) &\leq 4^{-2n} \cdot |1_S^T \cdot M \cdot 1_T| \\
&\leq 4^{-2n} \cdot \|1_S\|_2 \cdot \|M \cdot 1_T\|_2 \\
&\leq 4^{-2n} \cdot \|1_S\|_2 \cdot \|M\|_2 \cdot \|1_T\|_2 \\
&\leq 4^{-2n} \cdot \sqrt{|S||T|} \cdot \sqrt{14}^n \\
&\leq 4^{-2n} \cdot \sqrt{4^{2n}} \cdot \sqrt{14}^n = \left(\sqrt{14}/4\right)^n.
\end{aligned}$$

In the second line, we have applied Cauchy-Schwartz's Theorem, and in the last line we have used the trivial upper bounds $|S|, |T| \leq 4^n$. □

Theorem 3:

(1) $\mathrm{ModSum} \in \mathrm{coRP}_{1/2}\text{-BP1}$;

(2) $\mathrm{ModSum} \notin \mathrm{BPP}_\varepsilon\text{-BP1}$, *for* $\varepsilon < \frac{21}{256} < 0.083$.

Proof: *Part (1):* We obtain a nondeterministic read-once BP for $\neg\,\mathrm{ModSum}$ in the following way. We use two polynomial size OBDDs whose variables are ordered "row-wise" and "column-wise", respectively, to compute $\neg\,\mathrm{RowTest}(X)$ and $\neg\,\mathrm{RowTest}(X^T)$. These two graphs are combined by an $\vee$-node. To obtain a randomized read-once BP with one-sided error at most $1/2$, replace the $\vee$-node by a single stochastic variable.

Part (2): We are going to apply the technique described in Section 3. We choose μ as the uniform distribution on $\{0,1\}^{n^2}$ and show that ModSum has the following two properties:

(i) There is a constant $\rho > 0$ such that $\mu(\mathrm{ModSum}^{-1}(1)) \geq \rho$ for all n.

(ii) For an arbitrary rectangle R belonging to a balanced partition of the variables of ModSum it holds that

$$(*) \quad \mu(\mathrm{ModSum}^{-1}(0) \cap R) \geq \mu(\mathrm{ModSum}^{-1}(1) \cap R) - \delta(n),$$

where δ is a real-valued function and $\delta(n)$ is exponentially small in n.

We have shown property (i) in Lemma 1. It remains to establish property (ii). Let (X_1, X_2) be an arbitrary balanced partition of the input variables $X = (x_{ij})_{1\leq i,j\leq n}$ of ModSum, and let $R = A \times B$, $A \subseteq 2^{X_1}$, $B \subseteq 2^{X_2}$, be an arbitrary rectangle with respect to this partition.

We first apply Lemma 2. W. l. o. g. let the first assertion of the lemma hold. Then we can fix sets $X_1' \subseteq X_1$ and $X_2' \subseteq X_2$ such that there are $m := n/4$ rows i for which exactly two variables x_{ij} are in X_1' and two in X_2', and we have $|X_1'| = |X_2'| = 2m$.

We prove that for an arbitrary assignment a to all variables which are not in $X_1' \cup X_2'$

$$2^{-|X_1'|-|X_2'|} \cdot |\,\mathrm{ModSum}_a^{-1}(0) \cap R_a| \geq 2^{-|X_1'|-|X_2'|} \cdot |\,\mathrm{ModSum}_a^{-1}(1) \cap R_a| - \delta(n).$$

(For an arbitrary function f and a (partitial) assignment a, we write f_a for the subfunction (restriction) of f obtained by substituting variables by constants according to a. R_a is the restriction of R by a if we regard R as a characteristic function.) The claim $(*)$ follows from the above inequality by the law of total probability: for $c \in \{0,1\}$ it holds that

$$\begin{aligned}
&\sum_{a \in 2^{\overline{X_1' \cup X_2'}}} 2^{-|X_1'|-|X_2'|} \cdot |\,\mathrm{ModSum}_a^{-1}(c) \cap R_a| \cdot 2^{-|\overline{X_1' \cup X_2'}|} \\
&\quad = 2^{-|X|} \cdot \sum_{a \in 2^{\overline{X_1' \cup X_2'}}} |\,\mathrm{ModSum}_a^{-1}(c) \cap R_a| \\
&\quad = 2^{-|X|} \cdot |\,\mathrm{ModSum}^{-1}(c) \cap R|.
\end{aligned}$$

For the rest of the proof let a be a fixed assignment to the variables not in $X_1' \cup X_2'$. It holds that $R_a = A' \times B'$, where $A' \subseteq 2^{X_1'}$ and $B' \subseteq 2^{X_2'}$. Let us call the remaining free variables $x_{i,1}^1$, $x_{i,2}^1$ and $x_{i,1}^2$, $x_{i,2}^2$, where $i \in \{1,\ldots,m\}$ and the variables with upper index j are from the set X_j, $j = 1, 2$. Then the function $\mathrm{RowTest}_a$ can be written as

$$\mathrm{RowTest}_a(X_a) = \left[\sum_{i=1}^{m} [x_{i,1}^1 + x_{i,2}^1 + x_{i,1}^2 + x_{i,2}^2 + c_i \equiv 0 \bmod 3] \equiv c_0 \bmod 2\right],$$

with appropriate constants $c_0 \in \{0,1\}$ and $c_1, \ldots, c_m \in \mathbb{Z}_3$ depending only on a. By the definitions it follows that

$$\mathrm{ModSum}_a(X_a) = 1 \;\Rightarrow\; \mathrm{RowTestComm}_{c_0,c_1,\ldots,c_m}(x(X_a), y(X_a)) = 1,$$

where $x(X_a) := ((x^1_{1,1}, x^1_{1,2}), \ldots, (x^1_{m,1}, x^1_{m,2}))$ and $y(X_a) := ((x^2_{1,1}, x^2_{1,2}), \ldots, (x^2_{m,1}, x^2_{m,2}))$.

Now we can apply Lemma 3. Since $\mathrm{Disc}(\mathrm{RowTestComm}_{c_0,c_1,\ldots,c_m}) \leq (\sqrt{14}/4)^m$, we have

$$2^{-4m} \cdot |\,\mathrm{RowTestComm}^{-1}_{c_0,c_1,\ldots,c_m}(0) \cap R_a| \geq$$
$$2^{-4m} \cdot |\,\mathrm{RowTestComm}^{-1}_{c_0,c_1,\ldots,c_m}(1) \cap R_a| - (\sqrt{14}/4)^m,$$

and thus also

$$2^{-4m} \cdot |\,\mathrm{ModSum}^{-1}_a(0) \cap R_a| \geq 2^{-4m} \cdot |\,\mathrm{ModSum}^{-1}_a(1) \cap R_a| - (\sqrt{14}/4)^m,$$

therefore, inequality $(*)$ holds with $\delta(n) := (\sqrt{14}/4)^{n/4}$.

It only remains to apply Theorem 2 from Section 3. Let G be an arbitrary randomized read-once BP for ModSum with two-sided error at most ε. Let $\rho = 21/128$ as in Lemma 1. Then we obtain that

$$|G| \geq \frac{1}{2n}\left(\frac{\rho - 2\varepsilon}{(\sqrt{14}/4)^{n/4}}\right)^{1/2} = 2^{cn - O(\log n)}, \quad c := (1/8) \cdot \log_2(4/\sqrt{14}) \approx 0.012,$$

for $\varepsilon < \rho/2 = \frac{21}{256}$. □

Main Theorem: For ε with $0 \leq \varepsilon < \frac{21}{256}$, it holds that

(1) BPP_ε-BP1 $\not\subseteq$ NP-BP1;

(2) RP_ε-BP1 $\subsetneq$ $\mathrm{RP}_{1/2}$-BP1 $\subseteq$ NP-BP1.

The second part of this theorem shows that there is no "probability amplification" technique for read-once BPs that decreases the error below an arbitrary small positive constant. As we have already mentioned, this is contrary to the situation for OBDDs or general branching programs.

Conclusion and Open Problems

We have shown that BPP_ε-BP1 is incomparable to NP-BP1 if the error ε is not too large. This partially solves the open problem raised in [12] to separate the classes BPP-BP1 and NP-BP1.

We even have obtained an exponential gap between the randomized read-once BP sizes for different constant worst-case errors. In this respect, read-once branching programs turn out to behave rather "pathological" compared to the well-known probabilistic computation models and to randomized general branching programs or randomized OBDDs.

Some interesting problems concerning randomized read-once BPs still remain open, e. g. :

(1) Find a function f with $f \in$ NP-BP1, but $f \notin \mathrm{BPP}_{1/2-\varepsilon}$-BP1 for arbitrarily small $\varepsilon > 0$, showing that BPP-BP1 $\not\subseteq$ NP-BP1.

(3) Show that for arbitrary ε and ε' with $0 \leq \varepsilon < \varepsilon' < 1$ it holds that RP_ε-BP1 $\subsetneq$ $\mathrm{RP}_{\varepsilon'}$-BP1.

Acknowledgement

I would like to thank Ingo Wegener for helpful discussions on the subject of this paper.

References

[1] F. Ablayev. Randomization and nondeterminism are incomparable for polynomial ordered binary decision diagrams. In *Proc. of ICALP '97*.

[2] F. Ablayev and M. Karpinski. On the power of randomized branching programs. In *Proc. of ICALP '96*, LNCS 1099, 348–356. Springer, 1996.

[3] F. Ablayev and M. Karpinski. On the power of randomized ordered branching programs. *Manuscript*, Dec. 1996.

[4] L. Babai, P. Frankl, and J. Simon. Complexity classes in communication complexity theory. In *Proc. of the 27th IEEE Symp. on Foundations of Computer Science*, 337–347, 1986.

[5] B. Bollig, M. Sauerhoff, D. Sieling, and I. Wegener. Hierarchy theorems for kOBDDs and kIBDDs. To appear in *Theoretical Computer Science*, 1996.

[6] A. Borodin, A. A. Razborov, and R. Smolensky. On lower bounds for read-k-times-branching programs. *Computational Complexity*, 3:1–18, 1993.

[7] R. E. Bryant. Graph-based algorithms for Boolean function manipulation. *IEEE Trans. Computers*, C-35(8):677–691, Aug. 1986.

[8] R. E. Bryant. On the complexity of VLSI implementations and graph representations of Boolean functions with application to integer multiplication. *IEEE Trans. Computers*, C-40(2):205–213, Feb. 1991.

[9] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.*, 17(2):230–261, Apr. 1988.

[10] J. Gill. *Probabilistic Turing Machines and Complexity of Computations*. Ph. D. dissertation, U. C. Berkeley, 1972.

[11] K. Hosaka, Y. Takenaga, and S. Yajima. Size of ordered binary decision diagrams representing threshold functions. In *Proc. of the 5th Int. Symp. on Algorithms and Computation*, LNCS 834, 584–592. Springer, 1994.

[12] S. Jukna, A. Razborov, P. Savický, and I. Wegener. On P versus NP ∩ coNP for decision diagrams and read-once branching programs. Technical Report 647, Universität Dortmund, 1997. Submitted to *Computational Complexity*.

[13] C. Meinel. The power of polynomial size Ω-branching programs. In *Proc. of the 5th Ann. ACM Symp. on Theoretical Aspects of Computer Science*, LNCS 294, 81–90. Springer, 1988.

[14] E. A. Okolnishnikova. On lower bounds for branching programs. *Siberian Advances in Mathematics*, 3(1):152–166, 1993.

[15] C. H. Papadimitriou and M. Sipser. Communication complexity. In *Proc. of the 14th Ann. ACM Symp. on Theory of Computing*, 196–200, 1982.

[16] P. Pudlák and S. Zák. Space complexity of computations. Technical report, Univ. Prague, 1983.

[17] A. A. Razborov. Lower bounds for deterministic and nondeterministic branching programs. In *Proc. of Fundamentals of Computation Theory*, LNCS 529, 47–60. Springer, 1991.

[18] M. Sauerhoff. A lower bound for randomized read-k-times branching programs. Technical Report TR97-019, Electronic Colloquium on Computational Complexity, 1997. Available via WWW from `http://www.eccc.uni-trier.de/`.

[19] J. Simon and M. Szegedy. A new lower bound theorem for read-only-once branching programs and its applications. In J.-J. Cai, editor, *Advances in Computational Complexity Theory*, volume 13 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*. American Mathematical Society, 1993.

[20] I. Wegener. *The Complexity of Boolean Functions*. Wiley-Teubner Series in Computer Science. Teubner, Stuttgart; Wiley, Chichester; 1987.

[21] I. Wegener. On the complexity of branching programs and decision trees for clique functions. *J. ACM*, 35(2):461–471, Apr. 1988.

[22] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford, 1965.

[23] S. Žák. An exponential lower bound for one-time-only branching programs. In *Proc. of MFCS*, LNCS 176, 562–566. Springer, 1984.