A Lower Bound for Randomized Read-k-Times Branching Programs

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

Abstract

In this paper, we are concerned with randomized OBDDs and randomized read-k-times branching programs. We present an example of a Boolean function which has polynomial size randomized OBDDs with small, one-sided error, but only nondeterministic read-once branching programs of exponential size. Furthermore, we discuss a lower bound technique for randomized OBDDs with two-sided error and prove an exponential lower bound of this type. Our main result is an exponential lower bound for randomized read-k-times branching programs with two-sided error.

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 inner 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 an inner node labelled by x_i, check the value of this variable and follow the edge which is labelled by this value. 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 inner 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 every path of a branching program can be regarded as an assignment of variables.

*This work has been supported by DFG grants We 1066/7-3 and We 1066/8-1.

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 its assignment of variables is consistent). In this paper, we only consider syntactic read-k-times BPs.

The first exponential lower bounds for a branching program model have been proved for read-once branching programs (i. e. the case $k = 1$ in the above definition). The theory of this model is especially well understood, and there is a large collection of interesting lower bound results for it (Razborov (1991) gives an overview, for a summary of proof techniques, see Simon and Szegedy (1993)). In the context of this paper, lower bounds for general k are more interesting. For syntactic read-k-times BPs, exponential lower bounds have been independently proved by Okolnishnikova (1993) for $k \leq \varepsilon(\log n/\log\log n)$, $\varepsilon < 1$, and by Borodin, Razborov, and Smolensky (1993) even for non-deterministic syntactic read-k-times BPs and $k \leq \varepsilon \log n$, for appropriate ε. Recently, Jukna (1995) extended these results by showing an exponential gap between the size of non-deterministic read-k-times BPs for the function of Okolnishnikova and its complement. Up to now, proving lower bounds for non-syntactic branching programs is an open problem.

We mention another type of branching program which will be important in the sequel. OBDDs (ordered binary decision diagrams), introduced by Bryant (1986), are a restricted form of read-once-branching programs. An OBDD is a branching program with a prescribed ordering of variables. On each path from the source to the sinks, the variables have to be tested according to this ordering. OBDDs are the most important branching program type for applications (see the survey of Bryant (1992) for further information). Lower bounds for OBDDs have been proved, e. g., by Bryant (1991), Hosaka, Takenaga, and Yajima (1994) and Bollig, Sauerhoff, Sieling, and Wegener (1996).

In this paper we are concerned with randomized branching programs, i. e. branching programs with additional "coin-tossing nodes". Probabilistic complexity classes can be defined as for probabilistic Turing machines, which have been intensively studied since their introduction by Gill (1972). There are very few published results on randomized branching programs so far.

Ablayev and Karpinski (1996a) analyzed randomized OBDDs defined in analogy to randomized circuits from switching theory (Ajtai and Ben-Or 1984). They used a string-comparison function to show that there are functions that have randomized OBDDs which are exponentially smaller than deterministic kOBDDs for $k = o(n/\log n)$. (A kOBDD with variable ordering π is a read-k-times BP which can be partitioned into k layers such that each layer is an OBDD ordered according to π. This model has been studied by Bollig, Sauerhoff, Sieling, and Wegener (1996).) Recently, Ablayev and Karpinski were able to extend the lower bound even to non-deterministic kOBDDs for a modified version of the function (unpublished manuscript, Ablayev and Karpinski 1996b).

Ablayev and Karpinski have also shown that a nonboolean variant of their string-comparison function has exponential read-once BP size. But it has been an open problem so far to find a Boolean function exhibiting an exponential gap between deterministic and randomized complexity for read-once BPs.

What is known about lower bounds for randomized branching programs? Of course, all lower bounds for non-deterministic branching programs are also lower bounds for randomized branching programs with one-sided error (as for Turing machines, one-sided error is a weaker concept than non-determinism). Lower bounds for non-deterministic OBDDs can be obtained by the well-known fooling set technique (see Krause (1991), Bryant (1991) and also Kushilevitz and Nisan (1997) for the related technique for covers by 1-rectangles). Thus, we know e. g. for the multiplication function (see Bryant 1991) or for the functions analyzed in the paper of Krause (1991) that they are not computable by polynomial size randomized OBDDs with one-sided error, since it follows from the known fooling set constructions that these functions have exponential non-deterministic OBDD size.

But in order to get lower bounds for randomized branching programs with two-sided error, the known proof techniques cannot be used. Lower bounds for randomized OBDDs with two-sided error have been proved independently already in 1996 by Ablayev (manuscript, Ablayev 1996, accepted at ICALP '97) and the author.

We contribute to the theory of randomized branching programs in the following way. We first present an example of a Boolean function which has polynomial size randomized OBDDs with small one-sided error, but which has exponential size even for non-deterministic read-once branching programs (this extends the gap between deterministic and randomized OBDDs shown by Ablayev and Karpinski (1996a)). On the other hand, we show how exponential lower bounds on the size of randomized OBDDs with two-sided error can be established using tools from communication complexity theory.

Our main result is an exponential lower bound on the size of randomized read-k-times branching programs. We use a function introduced by Borodin, Razborov, and Smolensky (1993) that decides if a special kind of inner product of the input vectors is zero. We show that randomized read-k-times branching programs with two-sided error for this function have exponential size.

The rest of the paper is structured as follows. In Section 2, we first state some further definitions. Section 3 contains results on randomized OBDDs. The last two sections are devoted to our main result. In Section 4, we introduce our proof technique for randomized read-k-times BPs, and Section 5 contains the proof of the lower bound.

2 Definitions and Basic Facts

All the following definitions deal with variants of general branching programs, but it is easy to see that analogous notions can also be defined for the restricted models, especially read-k-times BPs and OBDDs.

For non-deterministic branching programs we use the definition of Meinel (1988).

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.

The semantics of an Ω-branching program is inductively defined. We assign a Boolean function f_v to each node v as follows. For a sink with value c, let $f_v \equiv c$. If v is an inner node labelled by a variable x_i, and v_0, v_1 are the nodes reached from v by the 0- and 1-edge, resp., then define

$$f_v := \overline{x_i} \cdot f_{v_0} \vee x_i \cdot f_{v_1},$$

as for an ordinary branching program. If v is an inner node labelled by $\omega \in \Omega$ and has successors v_1 and v_2, let

$$f_v := \omega(v_1, v_2).$$

The size of an Ω-branching program is the number of all its inner nodes.

Non-deterministic branching programs are $\{\vee\}$-branching programs in the sense of this definition, ordinary branching programs are obtained by choosing $\Omega = \emptyset$. We denote the class of sequences of Boolean functions which are computable by polynomial size non-deterministic branching programs by NP-BP. The class coNP-BP contains all sequences of Boolean functions computable by $\{\wedge\}$-branching programs of polynomial size. (Analogously define the classes NP-BPk and coNP-BPk for read-k-times BPs, NP-OBDD and coNP-OBDD for OBDDs.)

Definition 2: A *randomized branching program* G syntactically is a (deterministic) 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 $g\colon \{0,1\}^{n+r} \to \{0,1\}$ be the function computed by G as a deterministic branching program.

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, $0 \le \varepsilon < 1$, if for all $x \in \{0,1\}^n$ it holds that
  $$\begin{array}{lll} \Pr\{g(x,z)=0\} & = 1, & \text{if } f(x)=0; \\ \Pr\{g(x,z)=1\} & \ge 1-\varepsilon, & \text{if } f(x)=1; \end{array}$$
- *two-sided error* at most ε, $0 \le \varepsilon < 1/2$, if for all $x \in \{0,1\}^n$ it holds that
  $$\Pr\{g(x,z)=f(x)\} \ge 1-\varepsilon.$$

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-k-times 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 k times. For a *randomized OBDD*, a variable ordering on all 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$_\varepsilon$-BP be the class of sequences of functions computable by polynomial size randomized branching programs with one-sided error at most ε, $\varepsilon < 1/2$. Let BPP$_\varepsilon$-BP be the class of sequences of functions computable by polynomial size randomized branching programs with two-sided error at most ε, $\varepsilon < 1/2$. Furthermore, let

$$\begin{aligned} \text{RP-BP} &:= \bigcup_{0 \le \varepsilon < 1/2} \text{RP}_\varepsilon\text{-BP}, \\ \text{BPP-BP} &:= \bigcup_{0 \le \varepsilon < 1/2} \text{BPP}_\varepsilon\text{-BP}. \end{aligned}$$

Analogous classes for read-k-times BPs and OBDDs obtain "-BPk" and "-OBDD", resp., as a suffix to their name.

As for Turing machines, it holds that RP-BP $\subseteq$ NP-BP (analogously for read-k-times BPs, OBDDs). We can also adapt the well-known technique of iterating probabilistic computations to improve the error probability of randomized branching programs.

Lemma 1 (Probability amplification):

(1) Let G be a randomized read-k-times BP representing $f\colon \{0,1\}^n \to \{0,1\}$ with one-sided error at most $\varepsilon \in [0,1)$. Then a randomized read-(mk)-times BP G' for f with one-sided error at most ε^m and size $|G'| = O(m|G|)$ can be constructed.

(2) Let G be a randomized read-k-times BP representing $f\colon \{0,1\}^n \to \{0,1\}$ with two-sided error at most $\varepsilon \in [0,\frac{1}{2})$. Let $0 \le \varepsilon' \le \varepsilon$. Then a randomized read-(mk)-times BP G' for f with two-sided error less than ε' can be constructed which has size $|G'| = O(m^2|G|)$, with $m = O\left(\log((\varepsilon')^{-1})\left(\frac{1}{2}-\varepsilon\right)^{-2}\right)$.

Proof: We use standard construction techniques for branching programs to simulate the proof for Turing machines. We only sketch the construction, the estimation of the error works in the same way as for Turing machines.

Part (1): Use m copies of G with disjoint sets of stochastic variables, and compute the conjunction of these copies simply by identifying the 1-sink of the ith copy with the source of the $(i+1)$-th copy.

Part (2): Use a branching program consisting of m levels, where the ith level contains i copies of G, to count the number of 1's occuring during an m-fold iteration of the computation of G. The 1-sink is reached for all paths for which at least half of the copies of G computed 1. □

Corollary 1: RP-BP = BPP-BP = P-BP.

Proof: Decrease the error probability of a given randomized branching program for an n-variable function with two-sided error $\varepsilon < 1/2$ to less than 2^{-n} by Lemma 1. As for randomized circuits, it can be shown that the resulting branching program can be made deterministic by an appropriate choice of the stochastic variables (see the proof of Ajtai and Ben-Or (1984)). □

There is also an OBDD-version of Lemma 1.

Lemma 2 (Probability amplification for OBDDs):

(1) Let G be a randomized OBDD with variable ordering π representing $f\colon \{0,1\}^n \to \{0,1\}$ with one-sided error at most $\varepsilon \in [0,1)$. Then a randomized OBDD G' for f with the same variable ordering π can be constructed, which has one-sided error at most ε^m and size $|G'| = O(|G|^m)$.

(2) Let G be a randomized OBDD with variable ordering π representing $f\colon \{0,1\}^n \to \{0,1\}$ with two-sided error at most $\varepsilon \in [0,\frac{1}{2})$, and let $0 \le \varepsilon' \le \varepsilon$. Then a randomized OBDD G' for f with the same variable ordering π can be constructed, which has two-sided error less than ε' and size $|G'| = O(|G|^m)$, where $m = O\left(\log((\varepsilon')^{-1})\left(\frac{1}{2}-\varepsilon\right)^{-2}\right)$.

Proof, part (1): As for part (1) of Lemma 1, but here we use the "apply" algorithm for OBDDs (see Bryant 1986) with the operator "$\wedge$" to compute the graph G'. This is an OBDD with the same variable ordering as G and size $O(|G|^m)$.

Part (2): We again use the "apply" algorithm, but in a generalized version for multi-terminal OBDDs with sink values in $\{0, \ldots, m\}$. We apply the addition of integers as operation to sum up m copies of the OBDD G. The resulting graph is an OBDD with the same variable ordering as G and with sink values in $\{0, \ldots, m\}$. Replace the sinks with values greater or equal to $m/2$ by the 1-sink and all others by the 0-sink. □

It is not clear if the error probability of (general) read-k-times BPs can be decreased without increasing the number of tests of variables as done for OBDDs above.

3 Upper and Lower Bounds for Randomized OBDDs

We present two results on randomized OBDDs in this section. The first one is an example for an exponential gap between the size of randomized OBDDs and non-deterministic read-once BPs. Furthermore, we prove a lower bound result for randomized OBDDs with two-sided error.

We consider the function PERM, which is defined on an $n \times n$-matrix $X = (x_{ij})_{1 \leq i,j \leq n}$ of Boolean variables. Let $\mathrm{PERM}(X) = 1$ if and only if X is a permutation matrix, i. e. if each row and each column contains exactly one entry equal to 1.

Jukna (1989) and Krause, Meinel, and Waack (1988) have shown independently that non-deterministic read-once BPs for PERM have exponential size. They also showed that PERM $\in$ coNP-BP1. We extend this result by showing that PERM can be represented by randomized OBDDs with small one-sided error.

Theorem 1:

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

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

Proof: It only remains to improve the upper bound. The basic construction principle is the same as in the paper of Ablayev and Karpinski (1996a).

For the randomized computation of PERM, we use the following idea. Let $x_i = (x_{i,1}, \ldots, x_{i,n})$ be the i-th row of X. Let $|x|$ be the value of x interpreted as a binary representation. Then it holds that

$$\mathrm{PERM}(X) = 1 \;\Leftrightarrow\; \sum_{i=0}^{n} |x_i| = 2^n - 1 \;\wedge\; \text{all } x_i \text{ contain exactly one entry equal to 1.}$$

We choose the "row-wise" ordering $\pi = (x_{1,1}, x_{1,2}, \ldots, x_{1,n}, \ldots, x_{n,1}, \ldots, x_{n,n})$ of the variables for the randomized OBDD. Let $p(i)$ be the ith prime number. We start by choosing a prime number p at random from $\{p(1), \ldots, p(M)\}$, where M is fixed below. This choice of p is done by a tree at the top of the OBDD. After that, we compute the sum of all $|x_i|$ modulo p. In the following, we describe the subgraph of the OBDD which will do this.

Construct a "sub-module" for a row x_i and fixed p as follows. The module is an OBDD with p sources and $p + 1$ sinks, p sinks labelled by $0, \ldots, p-1$ and an "error" sink. We imagine this graph to be divided into p "columns" which correspond to intermediate results $0, \ldots, p-1$. Starting at source j, we check $x_{i,1}, x_{i,2}, \ldots, x_{i,n}$ successively and look for the first "1". If $x_{i,k}$ is the first variable which has the value 1, jump to column $(j + 2^k) \mod p$ and check in this column if all the remaining variables $x_{i,k+1},\dots,x_{i,n-1}$ of row x_i have the value 0. If not, or if no entry equal to 1 is found in row x_i, jump to the "error" sink. For each source, we need $O(n^2)$ nodes, hence $O(n^2p)$ nodes for a complete module.

Now form a sequence of the modules for all rows $x_1,\dots,x_n$ by identifying the sinks with values $0,\dots,p-1$ of the ith module with the appropriate sources of the $(i+1)$-th module. The sinks of the last module with values different from $(2^n-1) \bmod p$ are substituted by the 0-sink, the sink with value $(2^n-1) \bmod p$ by the 1-sink. All "error" sinks are also substituted by the 0-sink.

The complete OBDD (including the tree for the choice of p) has size

$$O\left(M + \sum_{p\in\{p(1),\dots,p(M)\}} n^3p\right).$$

Obviously, if $\mathrm{PERM}(X)=1$, we always reach the 1-sink for all p. The randomized OBDD can make an error only if $\mathrm{PERM}(X)=0$ and the matrix X has exactly one entry equal to 1 in each row. For such an input X, the OBDD computes the wrong output if the sum of all $|x_i|$ modulo p is equal to 2^n-1 for a prime p. Thus, we get for the error probability:

$$p_{\mathrm{err}} = \Pr\{\sum_{i=0}^{n-1}|x_i| \equiv 2^n-1 \bmod p\} = \Pr\{p \mid \left|\sum_{i=0}^{n-1}|x_i| - 2^n + 1\right|\},$$

where each x_i has exactly one entry equal to 1. It holds that

$$\left|\sum_{i=0}^{n-1}|x_i| - 2^n + 1\right| \le n\cdot 2^{n-1},$$

hence, there are fewer than $(n-1)\lceil\log n\rceil$ primes dividing this number. Choose $M := \varepsilon(n)^{-1}n\log n$, then $p_{\mathrm{err}} < \varepsilon(n)$. From number theory we know that $p(M) = O(M\log M)$, hence, we obtain the following upper bound for the size of the randomized OBDD:

$$O(\varepsilon(n)^{-2}n^5\log^3 n).$$

□

Corollary 2: RP-BP1 $\neq$ coRP-BP1.

Next, we discuss our lower bound results for randomized OBDDs with two-sided error. We will use tools from communication complexity theory for the proof. For definitions and a thorough introduction of communication complexity theory, see, e. g., the monographs of Hromkovič (1997) or Kushilevitz and Nisan (1997).

Our proof technique is an extension of a well-known lower bound technique for deterministic OBDDs, which appears, e. g., in the paper of Bollig, Sauerhoff, Sieling, and Wegener (1996) and in Theorem 12.12 of the monograph of Kushilevitz and Nisan. The main idea is to "reduce" a function which is known to be "hard" for one-way communication protocols to the function considered for OBDDs. We describe the randomized variant of the technique in form of a definition and a lemma.

Definition 3 (CC-OBDD Reduction): Let X, Y be arbitrary finite sets. Let a function $f\colon X \times Y \to \{0,1\}$, a function $g\colon \{0,1\}^n \to \{0,1\}$ defined on the variables $\{x_1, \ldots, x_n\}$ and a permutation $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ be given. Let $1 \leq k \leq n$.

We call two functions

$$\begin{aligned} &\varphi^A_{\pi,k}\colon X \to \{0,1\}^L, \quad L := \{x_{\pi(1)}, \ldots, x_{\pi(k)}\}, \quad \text{and} \\ &\varphi^B_{\pi,k}\colon Y \to \{0,1\}^R, \quad R := \{x_{\pi(k+1)}, \ldots, x_{\pi(n)}\}, \end{aligned}$$

a *CC-OBDD reduction from f to g*, if for all $(x,y) \in X \times Y$ it holds that

$$f(x,y) = g(\varphi^A_{\pi,k}(x) \cup \varphi^B_{\pi,k}(y)).$$

Here, $\varphi^A_{\pi,k}(x) \cup \varphi^B_{\pi,k}(y)$ denotes the assignment to $\{x_1, \ldots, x_n\}$ obtained by assigning the variables in L according to $\varphi^A_{\pi,k}(x)$ and the variables in R according to $\varphi^B_{\pi,k}(y)$.

Lemma 3: *Let f, g, π and k be as in Definition 3 and let $\varphi^A_{\pi,k}$, $\varphi^B_{\pi,k}$ be a CC-OBDD reduction from f to g. Let G be a randomized OBDD with non-stochastic variables $\{x_1, \ldots, x_n\}$ ordered according to π, and let G represent the function g with one-sided or two-sided error. Then there is a randomized one-way communication protocol for f of length $O(\log |G|)$ with the same error as G (also one-sided or two-sided).*

Proof: We describe an appropriate communication protocol for the function f. Let player A obtain the inputs in X and player B the inputs in Y. Furthermore, both players use a copy of G and their respective functions $\varphi^A_{\pi,k}$ or $\varphi^B_{\pi,k}$. Let $z_1, \ldots, z_i$ be the stochastic variables tested before $x_{\pi(k)}$ in G.

To compute f for $(x,y) \in X \times Y$, player A first chooses random values for $z_1, \ldots, z_i$ and then evaluates G for the partial assignment $\varphi^A_{\pi,k}(x)$. She sends the node v of G thus reached to player B. In the same manner, player B chooses his stochastic variables $z_{i+1}, \ldots, z_r$, follows the path for the assignment $\varphi^B_{\pi,k}(y)$ in G from v to a sink, and outputs the value of this sink.

This protocol has length $\log w$, where w is the number of nodes of G reached by assignments to $x_{\pi(1)}, \ldots, x_{\pi(k)}$ and $z_1, \ldots, z_i$. It is easy to verify that it indeed computes f with the error probability of G. □

We apply this technique to the function ISA ("indirect storage access") from the OBDD literature. Breitbart, Hunt III, and Rosenkrantz (1995) have proved that this function has exponential deterministic OBDD size.

Definition 4: The function ISA is defined on the $n = 2^p + p$ variables $x_0, \ldots, x_{2^p-1}$ and $y_0, \ldots, y_{p-1}$. Let $m := \lceil 2^p/p \rceil$. To compute the value of the function, first interpret y as a binary representation. Let i be the obtained value. If $i \geq m$, the value of ISA is zero. Otherwise, evaluate the group $(x_{ip}, \ldots, x_{(i+1)p-1})$ of the x-variables, let j be the value of this vector interpreted as a binary representation. Then the value of ISA is defined to be x_j.

Theorem 2: ISA $\notin$ BPP-OBDD.

Proof: Let G be a randomized OBDD for ISA with two-sided error ε, which has stochastic variables $z_1, \ldots, z_r$, and let $\pi\colon \{0, \ldots, 2^p - 1\} \to \{0, \ldots, 2^p - 1\}$ be the sub-ordering of the x-variables with respect to the variable ordering of G.

Consider the situation where the first $m-1$ of the x-variables, $x_{\pi(0)}, \ldots, x_{\pi(m-2)}$, have been tested. There is a group $x_{i_0 p}, \ldots, x_{(i_0+1)p-1}$ of x-variables (with index i_0) which has none of its variables tested under these first $m-1$ ones.

Now choose an assignment a to y which evaluates to i_0 (when interpreted as a binary representation). Apply the ordinary OBDD algorithm for the subsititution of variables by constants to compute from G a randomized OBDD with the same error as G for the restriction $\mathrm{ISA}|_{y=a}$. In the following, we only consider the graph thus computed (we again call it G). Note that G does not contain any y-variables and that it is no larger than the original graph.

We now present a communication problem which can be reduced to the function computed by G as described in Lemma 3. Let $X := \{0,1\}^M$, $Y := \{1, \ldots, M\}$ and $\mathrm{INDEX}\colon X \times Y \rightarrow \{0,1\}$ defined by $\mathrm{INDEX}(x,y) = x_y$ ("output the bit of x addressed by y").

We describe the mappings $\varphi^A_{\pi,k}$ and $\varphi^B_{\pi,k}$ for the application of Lemma 3. Define a partition of the x-variables by $L := \{x_{\pi(0)}, \ldots, x_{\pi(m-2)}\}$ and $R := \{x_{\pi(m-1)}, \ldots, x_{\pi(2^p-1)}\}$. Choose $k := m-1$ and also $M := m-1$. The mapping $\varphi^A_{\pi,k}$ is trivial, simply choose the identity. For $\varphi^B_{\pi,k}$ choose assignments $b_1, \ldots, b_{m-1} \in \{0,1\}^{2^p-m+1}$ to R such that for b_i the ith variable from L is addressed, i. e. for an assignment $c = (c_1, \ldots, c_{m-1}) \in \{0,1\}^{m-1}$ we have

$$\mathrm{ISA}|_{y=a, x_{\pi(0)}=c_1, \ldots, x_{\pi(m-2)}=c_{m-1}}(b_i) = c_i.$$

This can be done since no variable of group $x_{i_0 p}, \ldots, x_{(i_0+1)p-1}$ has been tested in L. Now define $\varphi^B_{\pi,k}(i) := b_i$ for $i \in \{1, \ldots, m-1\}$.

It is a well-known fact that the function INDEX is hard for one-way communication protocols where the player with the "memory" has to start. Kremer, Nisan, and Ron (1994) proved that every randomized one-way communication protocol for INDEX with error probability smaller than 1/8 has length $\Omega(n)$. From this, the claim follows by Lemma 3, using Lemma 2 to decrease the error probability. □

It is easy to see that $\mathrm{ISA} \in$ NP-OBDD (use a representation of ISA as a disjunctive form). Hence, we also get:

Corollary 3: RP-OBDD $\subsetneqq$ NP-OBDD.

By a different CC-OBDD reduction from INDEX we can also show that the well-known "hidden-weighted-bit" function HWB (Bryant 1991) is not contained in BPP-OBDD. Another example of a function which is hard for randomized OBDDs is the following extension of PERM.

Definition 5: The function XPERM is defined on a Boolean $n \times n$-matrix $X = (x_{i,j})_{1 \le i,j \le n}$. Let $\mathrm{XPERM}(X) = 1$ if and only if each column and each row of X contains *at most one* entry equal to one.

Using the known lower bound for the set-disjointness function from communication complexity theory (Razborov 1992), we can show that XPERM $\notin$ BPP-OBDD.

4 A Lower Bound Technique for Randomized Read-k-Times Branching Programs

In the remainder of this paper, we consider randomized read-k-times branching programs. Our goal for the next two sections is to prove an exponential lower bound for this model.

We start by describing a new proof technique for randomized read-k-times branching programs. For this technique as well as for the actual proof of the lower bound in the next section, we will apply some of the results of Borodin, Razborov, and Smolensky (1993), who have presented a lower bound technique for non-deterministic read-k-times BPs.

Before we start, we need some further definitions. As Borodin, Razborov and Smolensky, we first prove results for a generalized type of branching program, called s-way branching program, which uses s-valued variables instead of Boolean ones. For the whole section, let S be a finite set and $s := |S|$,

Definition 6: An *s-way branching program* on the variable set $\{x_1, \ldots, x_n\}$ is a directed acyclic graph which has one source and two sinks, the latter labelled by the constants 0 and 1. Each inner node is labelled by a variable x_i and has exactly s outgoing edges labelled by "1" to "s".

The semantics of an s-way BP is an obvious generalization of the semantics of 2-way BPs. Such a branching program computes a function $f\colon S^n \to \{0,1\}$, and the value for a given assignment of variables is obtained by following a path from the source to a sink. At each inner node, follow the edge which is labelled by the value of the variable belonging to that node.

We omit explicit definitions of read-k-times s-way BPs and randomized s-way BPs, since these are analogous to the 2-way case. Also note that Lemma 1 holds in an analogous form for s-way BPs. The following lemma connects randomized and deterministic s-way BPs.

Lemma and Definition 4: Let G be a randomized read-k-times s-way BP for $f\colon S^n \to \{0,1\}$ with error probabilty at most ε. Let μ be an arbitrary probability distribution on S^n. Then there is a *deterministic* read-k-times s-way BP G' with $|G'| \le |G|$ which computes a function $f'\colon S^n \to \{0,1\}$ with

$$\Pr_\mu\{f'(x) = f(x)\} \ge 1 - \varepsilon.$$

We call G' a *(μ,ε)-distributional read-k-times s-way BP for f.*

This is an adaption of an analogous fact for the so-called (μ,ε)-distributional communication complexity, which has been observed by Yao (1983). (The proof of this lemma is based on a simple counting argument.)

By Lemma 4, we are able to make distributional BPs the focus of our attention. In the following, we describe a lower bound technique for this model. The notion of rectangles used here is from Borodin, Razborov, and Smolensky (1993).

Definition 7 ((k,p)-Rectangle): Let X be a set of variables, $n := |X|$. Let k be an integer and $1 \le p \le n$. Let sets $X_1, \ldots, X_{kp} \subseteq X$ be given with

(1) $|X_i| \le \lceil n/p \rceil$, for $i = 1, \ldots, kp$;

(2) each variable from X appears in at most k sets X_i.

Furthermore, let $R \subseteq S^n$ be given. If there are functions $f_i\colon S^n \to \{0,1\}$ depending only on the variables from X_i such that for the characteristic function $f_R\colon S^n \to \{0,1\}$ of R (with $f_R(x) = 1$ iff $x \in R$) it holds that

$$f_R = \bigwedge_{1 \le i \le kp} f_i,$$

then we call R a (k,p)*-rectangle in* S^n *(with respect to the sets* $X_1, \ldots, X_{kp}$*).*

Notation: We will regard rectangles as sets or as characteristic functions, depending on what is more convenient, and we will use the same name for the set as well as for its characteristic function.

Lemma 5: *Let G be a deterministic read-k-times s-way BP for a function $f\colon S^n \to \{0,1\}$. Let $1 \le p \le n$. Then the following holds:*

(1) The branching program G defines a partition of S^n into (k,p)-rectangles such that f assumes a constant value within each of these rectangles.

(2) For the number r of these rectangles, it holds that

$$r \le (skn|G|)^{kp}.$$

Proof: This connection between (k,p)-rectangles and read-k-times BPs is a modified version of an analogous result of Borodin, Razborov, and Smolensky (1993). We additionally use ideas of Okolnishnikova (1993) here. Since Borodin, Razborov and Smolensky work with nondeterministic branching programs, they only obtain a *cover* of the 1-inputs of the function by rectangles. We need a (disjoint) partition of S^n here.

Let $X := \{x_1, \ldots, x_n\}$ be set of all variables of G. The first step of the proof is to modify G into an *uniform* read-k-times BP G'. A branching program is called uniform if it holds for each node v that on all paths from the source to v the same set of variables is tested. We obtain a uniform read-k-times BP G' by inserting dummy tests of missing variables into G, and it is easy to ensure that $|G'| \le skn|G|$.

Consider an arbitrary path P in G' from the source to one of the sinks. We are going to partition this path into l segments, where the ith segment is described by its first node w_i and its last node w'_i, and $w_i = w'_{i-1}$ for $i = 2, \ldots, l$.

For two nodes v and w in G define the set $X(v,w)$ of all variables tested on paths between v and w, including the variable of v and excluding the variable of w. Choose w_i and w'_i inductively as follows. The node w_1 is the source of G'. If w_i is fixed, choose w'_i as the first node on P (starting from w_i) such that $X(w_i, w'_i) \ge n/p$. Since G' is uniform, we also obtain that for this choice of w_i and w'_i it holds that $|X(w_i, w'_i)| \le \lceil n/p \rceil$, for $i = 1, \ldots, l$. It is easy to prove that $l \le kp$.

Now let L_i be the set of the last nodes on the ith segments of all paths in G', $i = 1, \ldots, kp$. If a path has less than kp segments, insert dummy segments at the end, consisting of the sink of the path as its first and last node. The set L_{kp} contains the sinks of G'. Let $L := L_1 \times \ldots \times L_{kp}$. For a sequence of nodes $v = (v_1, \ldots, v_{kp}) \in L$ let $R(v)$ be set of all $x \in S^n$ such that the path for x in G' starting at the source runs through $v_1, \ldots, v_{kp}$. Note that this set may be empty. By construction, the sets $R(v)$, $v \in L$, are (k,p)-rectangles with respect to sets $X_1(v), \ldots, X_{kp}(v)$, where $X_i(v) := X(v_{i-1}, v_i)$ ($i = 1, \ldots, kp$, let v_0 be the source). They even form a partition of the set of all inputs.

To prove the claim on the number of rectangles r, we only need to observe that $r \leq |L_1| \cdot \ldots \cdot |L_{kp}|$, and that $|L_i| \leq |G'|$ for all i. Taking into account that $|G'| \leq skn|G|$ the upper bound follows. □

We are now ready to state the theorem which summarizes our lower bound technique.

Theorem 3: *Let G be a (μ, ε)-distributional read-k-times s-way BP for a function $f \colon S^n \to \{0,1\}$. If for every (k,p)-rectangle R belonging to a partition of S^n, which is induced by G as described in the lemma above, it holds that*

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

where δ is a real valued function of n, then

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

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

Proof: The proof works in the same way as proofs of lower bounds on the (μ, ε)-distributional communication complexity (see, e. g., Kushilevitz and Nisan (1997)).

We estimate the number of (k,p)-rectangles for which f computes the value 1. Let $R_1, \ldots, R_r$ be these rectangles. Then it holds that

$$\mu\Big(\bigcup_{1 \leq i \leq r} R_i \Big) \geq \mu(f^{-1}(1)) - \varepsilon,$$

since G has error probability at most ε. On the other hand, by the assumption of the theorem a rectangle R_i which is not “very small” contains a “large” fraction of inputs for which f assumes the value zero, and for these inputs G makes an error. Hence, also

$$\begin{aligned}
\varepsilon &\geq \mu\Big(\bigcup_{1 \leq i \leq r} (R_i \cap f^{-1}(0)) \Big) \\
&= \sum_{1 \leq i \leq r} \mu(R_i \cap f^{-1}(0)) \\
&\geq \sum_{1 \leq i \leq r} \big(\alpha \cdot \mu(R_i \cap f^{-1}(1)) - \delta(n) \big) \\
&\geq \alpha(\mu(f^{-1}(1)) - \varepsilon) - r \cdot \delta(n).
\end{aligned}$$

For the second line we have used the fact that the rectangles form a disjoint partition of the inputs. Hence, we get the lower bound

$$r \geq \frac{\alpha \cdot \mu(f^{-1}(1)) - (1+\alpha) \cdot \varepsilon}{\delta(n)}$$

on the number of rectangles for which g computes the result one. We get the claimed lower bound on $|G|$ by the second part of Lemma 5. □

5 The Main Result

Now we are going to prove our main result, an exponential lower bound on the size of randomized read-k-times BPs. We apply the technique from the last section.

We consider the function SIP: $\mathbb{Z}_3^n \times \mathbb{Z}_3^n \rightarrow \{0,1\}$ ("Sylvester inner product"), $n = 2^d$, with

$$\mathrm{SIP}(x,y) = 1 \quad \Leftrightarrow \quad x^T A y \equiv 0,$$

where $A = (a_{i,j})_{1 \le i,j \le 2^d}$ is the Sylvester matrix of dimension $2^d \times 2^d$, i.e.

$$a_{i+1,j+1} := (-1)^{<\mathrm{bin}(i),\mathrm{bin}(j)>},$$

for $0 \le i,j \le 2^d - 1$, where bin(i) is the binary representation of i and $<\cdot\,,\,\cdot>$ the inner product in $\mathbb{Z}_2^d$. For the whole section, let $X := \{x_1, \ldots, x_n\}$ and $Y := \{y_1, \ldots, y_n\}$ be the sets of variables on which SIP is defined.

Borodin, Razborov, and Smolensky (1993) have proved that this function has no polynomial size non-deterministic read-k-times BP for $k \le c \log n$ for appropriate c. Our main theorem shows that SIP has no polynomial size $(\mathrm{uni}_{\mathbb{Z}_3^{2n}}, \varepsilon)$-distributional read-$k$-times 3-way BP, where $\mathrm{uni}_{\mathbb{Z}_3^{2n}}$ is the uniform distribution on $\mathbb{Z}_3^{2n}$. From this, we immediately get a lower on the size of randomized read-k-times BPs.

Before we start with the proof, we state the facts which we use from the paper of Borodin, Razborov and Smolensky. Furthermore, we present two lemmas which we will need later on.

5.1 Facts from the Paper of Borodin, Razborov and Smolensky

Borodin, Razborov and Smolensky show that a large number of (k,p)-rectangles is needed to cover all 1-inputs of SIP. As a first step in their proof, they consider a restriction of the (k,p)-rectangles and the function itself which reduces the original rectangles to much simpler 2-dimensional rectangles. We describe this step in the lemma below.

Lemma 6: *Let k be an integer, $1 \le p \le n$. Let $X_i \subseteq X$, $Y_i \subseteq Y$, $i = 1, \ldots, kp$, and let R be a (k,p)-rectangle in $\mathbb{Z}_3^{2n}$ with respect to the sets $X_i \cup Y_i$, $i = 1, \ldots, kp$ (especially, let $X_i \cup Y_i$ fulfill the conditions of Definition 7).*

Then there are sets $X_0 \subseteq X_1 \cup \ldots \cup X_{kp}$, $Y_0 \subseteq Y_1 \cup \ldots \cup Y_{kp}$ such that

$$R \;=\; R' \;\wedge\; R'',$$

where R', R'': $S^n \rightarrow \{0,1\}$ are functions that only depend on the variables from $X_0 \cup Y$ and $X \cup Y_0$, resp., and $|\overline{X_0} \times \overline{Y_0}| \ge 2kn^2/(p \cdot 4^k)$.

Corollary 4: *Let R be a (k,p)-rectangle as described above. Then there are sets $X_0 \subseteq X$ and $Y_0 \subseteq Y$ such that for each assignment a to $X_0 \cup Y_0$ the restriction R_a, which is obtained by considering R as a function and substituting variables according to a, can be represented as $R_a = T \times U$, for sets $T \subseteq \mathbb{Z}_3^t$, $U \subseteq \mathbb{Z}_3^u$, where $t,u \le n$ and $t \cdot u \ge 2kn^2/(p \cdot 4^k)$.*

Sets which can be written as a 2-dimensional cartesian product as R_a above are also termed *rectangles* in communication complexity theory. We call such sets *2-dimensional rectangles* here to distinguish them from (k, p)-rectangles.

The main part of the proof of Borodin, Razborov and Smolensky essentially is a generalization of a well-known proof of a linear lower bound on the deterministic communication complexity of the inner product in $\mathbb{Z}_2^n$. The key property of the function SIP which they use is that not only full Sylvester matrices, but also their submatrices have "large" rank. To be more precise, they have proved the following lemma.

Lemma 7: *For an arbitrary matrix X let $\alpha_s(X)$ be the minimal rank of a submatrix of X with at least s entries. Let A be the Sylvester matrix of dimension $n = 2^d$. Then*

$$\alpha_s(A) \geq \frac{s}{2n(\ln(2n) - (1/2) \cdot \ln s)}.$$

In our proof, we need a different approach for the main part. Borodin, Razborov and Smolensky show that (k, p)-rectangles for which the function SIP computes the result 1 are exponentially small. Following our technique from the last section, we prove the stronger fact that each (k, p)-rectangle which is "not too small" contains a "large fraction" of 0-inputs for SIP. Essentially, our proof will be a generalization of the technique used to prove a lower bound on the *randomized* communication complexity of the inner product in $\mathbb{Z}_2^n$.

5.2 Preparations for the Proof

In the following, we state the most important building blocks of our proof in form of two lemmas. One important step is to see that an arbitrary subfunction of SIP can be written as the transformation of an appropriate bilinear form. This is described in the following lemma.

Lemma 8: *Let a be an assignment to the variables from $X_0 \cup Y_0$, where $X_0 \subseteq X$ and $Y_0 \subseteq Y$. Define $t := |X \backslash X_0|$, $u := |Y \backslash Y_0|$. Let R be an arbitrary 2-dimensional rectangle in $\mathbb{Z}_3^t \times \mathbb{Z}_3^u$, i. e. $R = T \times U$ with $T \subseteq \mathbb{Z}_3^t$ and $U \subseteq \mathbb{Z}_3^u$.*

Then there is a one-to-one function φ: $\mathbb{Z}_3^t \times \mathbb{Z}_3^u \to \mathbb{Z}_3^{t+1} \times \mathbb{Z}_3^{u+1}$ and a bilinear form F: $\mathbb{Z}_3^{t+1} \times \mathbb{Z}_3^{u+1} \to \mathbb{Z}_3$ defined by

$$F(x, y) := x^T B y,$$

where $x \in \mathbb{Z}_3^{t+1}$, $y \in \mathbb{Z}_3^{u+1}$ and B is a $(t + 1) \times (u + 1)$-matrix over $\mathbb{Z}_3$, such that the following holds:

(1) $\mathrm{SIP}_a(x, y) = 1 \Leftrightarrow F(\varphi(x, y)) \equiv 0 \pmod 3$, *and* $\mathrm{SIP}_a(x, y) = 0 \Leftrightarrow F(\varphi(x, y)) \in \mathbb{Z}_3 \backslash \{0\}$*;*

(2) $|R \cap \mathrm{SIP}_a^{-1}(1)| = |\varphi(R) \cap F^{-1}(0)|$, *and* $|R \cap \mathrm{SIP}_a^{-1}(0)| = |\varphi(R) \cap F^{-1}(\mathbb{Z}_3 \backslash \{0\})|$*;*

(3) $\mathrm{rank}(B) \geq \alpha_{t \cdot u}(A)$ *(where A is the Sylvester matrix of dimension $n \times n$ and $\alpha_{t \cdot u}(A)$ as defined in Lemma 7).*

In these expressions, SIP_a *denotes the restrictions of* SIP *resulting from the substitution of variables according to a.*

Proof: As in the definition of SIP, let A be the Sylvester matrix of dimension $n \times n$. Let A' be the $t \times u$-submatrix of A which is obtained by deleting the rows and columns of A corresponding to X_0 and Y_0, resp. There exist $v \in \mathbb{Z}_3^t$, $w \in \mathbb{Z}_3^u$ and $\gamma \in \mathbb{Z}_3$ such that $\mathrm{SIP}_a(x, y) = 1$ if and only if

$$x^T A' y + x^T v + w^T y + \gamma \equiv 0 \pmod 3,$$

where $x \in \mathbb{Z}_3^{X'}$, $y \in \mathbb{Z}_3^{Y'}$.
We can write

$$x^T A' y + x^T v + w^T y + \gamma = x'^T B y'$$

where $x' \in \mathbb{Z}_3^{t+1}$ and $y' \in \mathbb{Z}_3^{u+1}$ are defined by

$$x'_i := \begin{cases} 1, & \text{if } i = 1; \\ x_{i-1}, & \text{if } i \in \{2, \ldots, t+1\}; \end{cases} \quad \text{and} \quad y'_i := \begin{cases} 1, & \text{if } i = 1; \\ y_{i-1}, & \text{if } i \in \{2, \ldots, u+1\}; \end{cases}$$

and the matrix $B = (b_{i,j})_{1 \le i \le t+1, 1 \le j \le u+1}$ is defined by

$$B := \left(\begin{array}{c|c} \gamma & w^T \\ \hline & \\ v & A' \\ & \end{array} \right).$$

Obviously, $\mathrm{rank}(B) \ge \mathrm{rank}(A')$. Define the bilinear form $F\colon \mathbb{Z}_3^{t+1} \times \mathbb{Z}_3^{u+1} \to \mathbb{Z}_3$ by

$$F(x, y) := x^T B y,$$

and $\varphi\colon \mathbb{Z}_3^t \times \mathbb{Z}_3^u \to \mathbb{Z}_3^{t+1} \times \mathbb{Z}_3^{u+1}$ by

$$\varphi(x, y) := (x', y'), \quad x' \text{ and } y' \text{ as above.}$$

Then φ, F and B fulfill the conditions (1) to (3). (Statement (2) follows from the fact that φ is one-to-one.) □

The second building block is a generalization of a lemma attributed to Lindsey (see, e. g., Chor and Goldreich 1988). In its familiar form, this lemma states that in every submatrix of a Hadamard matrix which is not too small the number of 1's and (-1)'s is nearly balanced. (A Hadamard matrix is an orthogonal matrix with entries equal to -1 or 1. Sylvester matrices, defined by the inner product in $\mathbb{Z}_2^n$ as seen above, are a special type of Hadamard matrices.)

For our generalization of Lindsey's lemma, we consider a matrix defined by a bilinear form with values in $\mathbb{Z}_3$. Consider the $3^t \times 3^u$-matrix $M = (m(x, y))_{x \in \mathbb{Z}_3^t, y \in \mathbb{Z}_3^u}$, defined by $m(x, y) := x^T A y$, where $x \in \mathbb{Z}_3^t$, $y \in \mathbb{Z}_3^u$, and A is an arbitrary $t \times u$-matrix over $\mathbb{Z}_3$. We show that in every submatrix of a M which is not too small the number of entries 0, 1 and -1 is nearly balanced, i. e. amounts to approximately one third of all entries.

This is done by the following indirect approach. For each pair $\{i,j\}$, $i,j \in \mathbb{Z}_3$, $i \neq j$, we define a separate $3^t \times 3^u$ matrix $M_{i,j}$ by

$$M_{i,j}(x,y) := \begin{cases} 1, & \text{if } x^T A y \equiv i \pmod 3; \\ -1, & \text{if } x^T A y \equiv j \pmod 3; \\ 0, & \text{otherwise;} \end{cases}$$

where $x \in \mathbb{Z}_3^t$ and $y \in \mathbb{Z}_3^u$. We show that the sum of 1's and (-1)'s in submatrices of these $M_{i,j}$ is small if the matrix A has large rank.

Lemma 9: *Let A be an arbitrary $t \times u$-matrix over $\mathbb{Z}_3$. Define the matrices $M_{i,j}$ as described above. Furthermore, let R be an arbitrary 2-dimensional rectangle in $\mathbb{Z}_3^t \times \mathbb{Z}_3^u$, $R = S \times T$ with $S \subseteq \mathbb{Z}_3^t$, $T \subseteq \mathbb{Z}_3^u$. Let $d_{i,j}(S,T)$ denote the sum of 1's and (-1)'s of $M_{i,j}$ in this rectangle, i. e.*

$$d_{i,j}(S,T) := \sum_{x \in S} \sum_{y \in T} M_{i,j}(x,y).$$

Then it holds that

(1) $|d_{1,-1}(S,T)| \leq (2/\sqrt{3}) \cdot 3^{t+u} \cdot 3^{-\operatorname{rank}(A)/2}$.

(2) $|d_{1,0}(S,T)| \leq 3^{t+u} \cdot 3^{-\operatorname{rank}(A)/2}$.

(3) $|d_{-1,0}(S,T)| \leq 3^{t+u} \cdot 3^{-\operatorname{rank}(A)/2}$.

We defer the lengthy and technical proof of this lemma to the appendix.

5.3 The Proof of the Main Result

We are now ready to prove the following.

Theorem 4: *Let G be a $(\mathrm{uni}_{\mathbb{Z}_3^{2n}}, \varepsilon)$-distributional 3-way read-k-times BP for* SIP, *where $\mathrm{uni}_{\mathbb{Z}_3^{2n}}$ is the uniform distribution on $\mathbb{Z}_3^{2n}$ and $\varepsilon < 1/9$. Then*

$$|G| = \exp\left(\Omega\left(\frac{n}{k^3 \cdot 4^k}\right)\right).$$

Proof: We apply Theorem 3 with $p := 4k$ and $\mu = \mathrm{uni}_{\mathbb{Z}_3^{2n}}$. Consider a partition of $\mathbb{Z}_3^{2n}$ into (k,p)-rectangles induced by G as described in Lemma 5 from Section 4.

Let R be an arbitrary (k,p)-rectangle from this partition. The main work of the proof will be to derive a relation between the number of 0-inputs and 1-inputs for SIP in R. This will have the form

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

where α and $\delta(n)$ are defined later on. The proof of this fact consists of three parts.

Part (1): Let R be a (k,p)-rectangle with respect to the sets $X_i \cup Y_i$, $i = 1, \ldots, kp$, obtained from the partition of $\mathbb{Z}_3^{2n}$ induced by G.

From Corollary 4, we get sets $X_0 \subseteq X$ and $Y_0 \subseteq Y$ such that for each assignment a to $X_0 \cup Y_0$ the restriction R_a is a 2-dimensional rectangle in $\mathbb{Z}_3^t \times \mathbb{Z}_3^u$, where $t := |X \backslash X_0|$, $u := |Y \backslash Y_0|$. We also know that this rectangle is "not too small", it holds that $t \cdot u \geq n^2/(2 \cdot 4^k) =: s$. In the following, we only consider the rectangle R_a and the subfunction SIP_a (resulting from the substitution of variables according to a) for such an assignment a.

Part (2): While the restriction by a has transformed the (k,p)-rectangle R into a simple 2-dimensional rectangle, the function SIP_a is too complicated to be used itself. This problem is solved by Lemma 8. From this lemma, we obtain a one-to-one function φ and a bilinear form F in $\mathbb{Z}_3^{t+1} \times \mathbb{Z}_3^{u+1}$ such that

$$\mathrm{SIP}_a(x,y) = 1 \Leftrightarrow F(\varphi(x,y)) \equiv 0 \pmod{3}.$$

The 2-dimensional rectangle R_a is transformed into the 2-dimensional rectangle $\varphi(R_a)$ in $\mathbb{Z}_3^{t+1} \times \mathbb{Z}_3^{u+1}$. Moreover, we know that for the matrix B of F it holds that $\mathrm{rank}(B) \geq \alpha_s(A)$ by statement (3) of Lemma 8 (s as defined above). By Lemma 7, we get a lower bound on the rank of B.

We have now managed to transform our original task into the problem to show that a 2-dimensional rectangle for the bilinear form F which is not "too small" contains a "large fraction" of inputs from $F^{-1}(0)$ as well as from $F^{-1}(\mathbb{Z}_3 \backslash \{0\})$.

Part (3): At this point we apply our generalized form of Lindsey's lemma. Let R' be an arbitrary 2-dimensional rectangle in $\mathbb{Z}_3^{t+1} \times \mathbb{Z}_3^{u+1}$. In Lemma 9, substitute the matrix B for A, R' for the rectangle and $t+1$, $u+1$ for t and u, resp. Then it follows that

$$\begin{aligned}
\big||R' \cap F^{-1}(1)| - |R' \cap F^{-1}(-1)|\big| &\leq (2/\sqrt{3}) \cdot 3^{t+u+2} \cdot 3^{-\mathrm{rank}(B)/2} =: b, \\
\big||R' \cap F^{-1}(1)| - |R' \cap F^{-1}(0)|\big| &\leq 3^{t+u+2} \cdot 3^{-\mathrm{rank}(B)/2} \leq b, \quad \text{and} \\
\big||R' \cap F^{-1}(-1)| - |R' \cap F^{-1}(0)|\big| &\leq 3^{t+u+2} \cdot 3^{-\mathrm{rank}(B)/2} \leq b.
\end{aligned}$$

We conclude that

$$|R' \cap F^{-1}(\mathbb{Z}_3 \backslash \{0\})| \geq 2 \cdot |R' \cap F^{-1}(0)| - 2b.$$

To see this, let $x := |R' \cap F^{-1}(0)|$, $y := |R' \cap F^{-1}(1)|$, $z := |R' \cap F^{-1}(-1)|$. Then $y+z$ is minimized under the constraints

$$|y-z| \leq b, \quad |y-x| \leq b, \quad |z-x| \leq b,$$

if $y = z = x - b$. Hence, $y + z \geq 2x - 2b$.

Now we apply these results to our original problem. Substituting $R' := \varphi(R_a)$ and applying statement (2) from Lemma 8 we obtain

$$|R_a \cap \mathrm{SIP}_a^{-1}(0)| \geq 2 \cdot |R_a \cap \mathrm{SIP}_a^{-1}(1)| - 2b.$$

Applying $\mu = \mathrm{uni}_{\mathbb{Z}_3^{2n}}$ on both sides gives

$$\mu(R_a \cap \mathrm{SIP}_a^{-1}(0)) \geq 2 \cdot \mu(R_a \cap \mathrm{SIP}_a^{-1}(1)) - 2b \cdot 3^{-t-u}.$$

This inequality holds for all assignments a to $X_0 \cup Y_0$, and hence, by the law of total probability, it carries over to R and SIP. We have thus obtained the desired relationship between 0-inputs and 1-inputs in R,

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

with $\alpha := 2$ and

$$\delta(n) := 2b \cdot 3^{-t-u} = 12\sqrt{3} \cdot 3^{-\operatorname{rank}(B)/2}.$$

The last information which we need in order to apply Theorem 3 is a lower bound for $\mu(\mathrm{SIP}^{-1}(1))$. As in the definition of SIP, let A be the Sylvester matrix of dimension $n \times n$. It is easy to verify that $x^T A y = c$, $c \in \mathbb{Z}_3$, for approximately one third of all $x, y \in \mathbb{Z}_3^n$. More precisely, we get

$$\mu(\mathrm{SIP}^{-1}(1)) = 1/3 - o(1).$$

Now we are ready to apply Theorem 3. As mentioned above, it holds that $\operatorname{rank}(B) \geq \alpha_s(A)$, where $s = n^2/(2 \cdot 4^k)$. By Lemma 7, we have

$$\alpha_s(A) = \Omega\left(\frac{n}{k \cdot 4^k}\right).$$

Altogether, we obtain:

$$\begin{aligned} |G| &\geq \frac{1}{3kn} \cdot \left(\frac{2 \cdot \left(\frac{1}{3} - o(1) - 3\varepsilon\right)}{\delta(n)}\right)^{1/(4k^2)} \\ &= \exp\left(\Omega\left(\frac{n}{k^3 \cdot 4^k}\right)\right), \end{aligned}$$

for $\varepsilon < 1/9$. □

Corollary 5: *Let G be a randomized read-k-times 3-way BP for* SIP *with two-sided error at most ε, where $\varepsilon \in [0, 1/2)$ is an arbitrary constant. Then there is a constant c with*

$$|G| = \exp\left(\Omega\left(\frac{n}{k^3 \cdot c^k}\right)\right).$$

Proof: For $\varepsilon < 1/9$, this follows immediately from Theorem 4 and Lemma 4. For a larger $\varepsilon < 1/2$, ε constant, let a randomized read-k-times 3-way BP G with such a maximal error probability be given. By applying the probability amplification technique described in Lemma 1 (in the version for s-way BPs) with $\varepsilon' := 1/9$, this BP is transformed into a read-(mk)-times 3-way BP G'. It holds that $m = O(1)$ and $|G'| = O(m^2|G|) = O(|G|)$. From this, we get the claimed lower bound with $c := 4^m$ by applying Theorem 4. □

Finally, we are going to use Theorem 4 to derive a lower bound for a Boolean variant of SIP which is obtained by encoding the values from $\mathbb{Z}_3$ by Boolean values.

First, define the incompletely specified function $c_0\colon \{0,1\}^2 \to \mathbb{Z}_3 \cup \{\perp\}$ ($\perp$ = "undefined") by

$$c_0(0,0) := 0, \quad c_0(0,1) := 1, \quad c_0(1,0) := -1 \quad \text{and} \quad c_0(1,1) := \perp.$$

Let $c\colon \{0,1\}^{4n} \to \mathbb{Z}_3^{2n} \cup \{\perp\}$ be the incompletely specified coding function defined as follows. For $u_i := (x_i^0, x_i^1)$ and $v_i := (y_i^0, y_i^1)$, $i = 1, \ldots, n$, let

$$c(u_1, \ldots, u_n, v_1, \ldots, v_n) := (c_0(u_1), \ldots, c_0(u_n), c_0(v_1), \ldots, c_0(v_n)),$$

if $c_0(u_i), c_0(v_i) \in \mathbb{Z}_3$ for all i, and let

$$c(u_1, \ldots, u_n, v_1, \ldots, v_n) := \perp,$$

if $u_i = \perp$ or $v_i = \perp$ for at least for one $i \in \{1, \ldots, n\}$.

Now define $\widetilde{\mathrm{SIP}}\colon \{0,1\}^{4n} \to \{0,1\}$ by

$$\widetilde{\mathrm{SIP}}(x,y) := \begin{cases} \mathrm{SIP}(c(x,y)), & \text{if } c(x,y) \in \mathbb{Z}_3^{2n}; \\ 0, & \text{if } c(x,y) = \perp. \end{cases}$$

Furthermore, define the probability distribution $\mu\colon \{0,1\}^{4n} \to [0,1]$ by

$$\mu(u,v) := \begin{cases} 3^{-2n}, & \text{if } c(u,v) \in \mathbb{Z}_3^{2n}; \\ 0, & \text{if } c(u,v) = \perp. \end{cases}$$

Theorem 5: *Let G be a (μ, ε)-distributional read-k-times 2-way BP for $\widetilde{\mathrm{SIP}}$, $\varepsilon < 1/9$. Then*

$$|G| = \exp\left(\Omega\left(\frac{n}{k^3 \cdot 16^k}\right)\right).$$

Proof: Let x_i^j, y_i^j, for $i \in \{1, \ldots, n\}$, $j \in \{0,1\}$, be the variables of G. We construct a read-k-times 3-way BP G' with variables $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ from G as follows. First consider a node in G which is labelled by a variable x_i^0 or y_i^0. Replace the variable by x_i or y_i, resp. Replace the 0-edge by two edges labelled by "0" and by "-1", resp., and the 1-edge by an edge labelled by "1". Next, consider a node labelled by x_i^1 or y_i^1. The variable is again replaced by x_1 or y_1. Replace the 0-edge by two edges labelled by "0" and by "1", resp., and the 1-edge by an edge labelled by "-1".

We claim that G' is a $(\mathrm{uni}_{\mathbb{Z}_3^{2n}}, \varepsilon)$-distributional read-$(2k)$-times 3-way BP for the function SIP, where $\mathrm{uni}_{\mathbb{Z}_3^{2n}}$ again is the uniform distribution on $\mathbb{Z}_3^{2n}$. G' obviously is of the same size as G.

Let g' be the function computed by G' and g the function computed by G. Let $c'\colon \mathbb{Z}_3^{2n} \to c^{-1}(\mathbb{Z}_3^{2n})$ be the one-to-one and onto mapping with $c(c'(x,y)) = (x,y)$. By the definition of $\widetilde{\mathrm{SIP}}$, it holds that $\widetilde{\mathrm{SIP}}(c'(x,y)) = \mathrm{SIP}(x,y)$ for all $(x,y) \in \mathbb{Z}_3^{2n}$. By the contruction of G', it follows that $g(c'(x,y)) = g'(x,y)$ for all $(x,y) \in \mathbb{Z}_3^{2n}$. Hence,

$$\begin{aligned} |\{(x,y) \in \mathbb{Z}_3^{2n} \mid g'(x,y) = \mathrm{SIP}(x,y)\}| &= |\{(x,y) \in \mathbb{Z}_3^{2n} \mid g(c'(x,y)) = \widetilde{\mathrm{SIP}}(c'(x,y))\}| \\ &= |\{(u,v) \in c^{-1}(\mathbb{Z}_3^{2n}) \mid g(u,v) = \widetilde{\mathrm{SIP}}(u,v)\}|. \end{aligned}$$

Since μ is zero outside $c^{-1}(\mathbb{Z}_3^{2n})$, we finally get

$$\mathrm{uni}_{\mathbb{Z}_3^{2n}}\{(x,y) \in \mathbb{Z}_3^{2n} \mid g'(x,y) = \mathrm{SIP}(x,y)\} = \mu\{(u,v) \in \{0,1\}^{4n} \mid g(u,v) = \widetilde{\mathrm{SIP}}(u,v)\}.$$

By Theorem 4 it follows that

$$|G| = \exp\left(\Omega\left(\frac{n}{(2k)^3 \cdot 4^{2k}}\right)\right)$$

and thus the claimed lower bound. □

Corollary 6: *Let G be a randomized read-k-times 2-way BP for $\widetilde{\mathrm{SIP}}$ with two-sided error at most ε, for arbitrary constant $\varepsilon \in [0, 1/2)$. Then there is a constant c with*

$$|G| = \exp\left(\Omega\left(\frac{n}{k^3 \cdot c^k}\right)\right).$$

Proof: Follows from Lemma 1 and Lemma 4. □

Acknowledgement

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

Appendix

In this appendix, we supply the proof of Lemma 9 which we have omitted above. We split the proof of the three statements in Lemma 9 into three separate lemmas here. The first statement can be easily proved even for *arbitrary* fields $\mathbb{Z}_p$, p an odd prime, not only for $\mathbb{Z}_3$.

First, we state some general definitions. For the whole appendix, let $p \neq 2$ be a prime. Let $r_1, \ldots, r_{(p-1)/2}$ be the quadratic residues modulo p and $\tilde{r}_1, \ldots, \tilde{r}_{(p-1)/2}$ the non-residues (we consider 0 neither as a residue nor as a non-residue). Let

$$\left(\frac{a}{p}\right) := \begin{cases} 1, & \text{if } a \in \{r_1, \ldots, r_{(p-1)/2}\}; \\ 0, & \text{if } a = 0; \\ -1, & \text{if } a \in \{\tilde{r}_1, \ldots, \tilde{r}_{(p-1)/2}\}; \end{cases}$$

the Legendre symbol modulo p. Furthermore, let A be an arbitrary $t \times u$-matrix over $\mathbb{Z}_p$.

Lemma 9, Part (1): Define the $p^t \times p^u$-Matrix $M = (m(x,y))_{x \in \mathbb{Z}_p^t, y \in \mathbb{Z}_p^u}$ by $m(x,y) := \left(\frac{x^T A y}{p}\right)$. Let an arbitrary 2-dimensional rectangle be given by the sets $S \subseteq \mathbb{Z}_p^t$, $T \subseteq \mathbb{Z}_p^u$. Let d denote the sum of ones and minus ones of M in this rectangle, i. e.

$$d := \sum_{x \in S} \sum_{y \in T} m(x,y).$$

Then we have

$$|d| \leq (p-1) p^{t+u} p^{-(\mathrm{rank}(A)+1)/2}.$$

Proof: In the following, we consider M as a real-valued matrix with entries 0, 1 and -1. If we do not state something different explicitly, all calculations are done in real-valued vector spaces. Furthermore, all congruences "$\equiv$" are modulo p.

We consider the matrix $\tilde{M} = (\tilde{m}(x,y))_{x,y\in\mathbb{Z}_p^u}$, $\tilde{M} := M^TM$ (calculated in $\mathbb{R}$). It holds that $\tilde{m}(x,y) = m(x)^Tm(y)$, where $m(x)$ is the column of M with number $x \in \mathbb{Z}_p^u$. Notice that $\tilde{M}$ is symmetric. In the following, our main goal is to calculate the maximal eigenvalue of $\tilde{M}$. The whole proof is divided into three steps.

Step 1: We first calculate the entries of M. Let $x, y \in \mathbb{Z}_p^u$. We claim that

$$m(x)^Tm(y) = \begin{cases} \left(\frac{\alpha}{p}\right)(p-1)p^{t-1}, & \text{if } Ay \equiv \alpha Ax,\ Ax, Ay \not\equiv 0,\ \alpha \in \mathbb{Z}_p\backslash\{0\}; \\ 0, & \text{otherwise.} \end{cases}$$

Proof of the claim: Obviously, $m(x)^Tm(y) = 0$ if $Ax \equiv 0$ or $Ay \equiv 0$. Therefore, let $Ax \not\equiv 0$, $Ay \not\equiv 0$. We have

$$\begin{aligned} m(x)^T \cdot m(y) &= \sum_{z\in\mathbb{Z}_p^t} \left(\frac{z^TAx}{p}\right)\left(\frac{z^TAy}{p}\right) \\ &= \sum_{(a_z,b_z)=(1,1)} 1 \;+ \sum_{(a_z,b_z)=(-1,1)} (-1) + \sum_{(a_z,b_z)=(1,-1)} (-1) + \sum_{(a_z,b_z)=(-1,-1)} 1, \end{aligned} \tag{1}$$

where $a_z := \left(\frac{z^TAx}{p}\right)$, $b_z := \left(\frac{z^TAy}{p}\right)$ and the summation is done over all $z \in \mathbb{Z}_p^t$ which fulfill the given restrictions.

First of all, we count the number of $z \in \mathbb{Z}_p^t$ with $(a_z, b_z) = (1,1)$. To do this, we apply linear algebra in $\mathbb{Z}_p^t$. It holds that

$$\left(\frac{z^TAx}{p}\right) = 1 \;\wedge\; \left(\frac{z^TAy}{p}\right) = 1 \;\Leftrightarrow\; \exists\, i,j\colon z^TAx \equiv r_i \;\wedge\; z^TAy \equiv r_j. \tag{2}$$

For fixed i and j we are looking for the number of solutions for the system of linear equations

$$\begin{aligned} & a_1z_1 + \ldots + a_tz_t \equiv r_i \\ \wedge\;\; & b_1z_1 + \ldots + b_tz_t \equiv r_j \end{aligned}$$

in the variables $z_1, \ldots, z_t$, where $a := Ax$, $a = (a_i)_{1\le i\le t}$, and $b := Ay$, $b = (b_i)_{1\le i\le t}$.

If Ax and Ay are linearly independent in $\mathbb{Z}_p^t$, this system has exactly p^{t-2} solutions. Hence, there are $(\frac{p-1}{2})^2p^{t-2}$ vectors $z \in \mathbb{Z}_p^t$ which fulfill (2).

If Ax and Ay are linearly dependent, $Ay \equiv \alpha Ax$ for an $\alpha \not\equiv 0$, then this system has either no solution or exactly p^{t-1} solutions. The latter is the case if and only if

$$r_j \equiv \alpha r_i.$$

If $\left(\frac{\alpha}{p}\right) = 1$, then for each $i \in \{1, \ldots, (p-1)/2\}$ there is exactly one j with $r_j \equiv \alpha r_i$, and hence the total number of vectors fulfilling (2) is $(\frac{p-1}{2})p^{t-1}$. If $\left(\frac{\alpha}{p}\right) = -1$, then $r_j \equiv \alpha r_i$ is always false and the total number of z-vectors is 0.

By an analogous argumentation, we get the same number of z-vectors with $(a_z, b_z) = (-1, -1)$. It remains to calculate the number of $z \in \mathbb{Z}_p^t$ with $(a_z, b_z) = (1, -1)$ (analogously for $(a_z, b_z) = (-1, 1)$). Again, we count the solutions of linear equations in $\mathbb{Z}_p^t$. Here we are looking for solutions for the following system:

$$\begin{aligned} & a_1 z_1 + \ldots + a_t z_t \equiv r_i \\ \wedge\ & b_1 z_1 + \ldots + b_t z_t \equiv \tilde{r}_j, \end{aligned}$$

where the a_i and b_i are defined as above. If Ax and Ay are linearly independent, the number of solutions for each i, j is again p^{t-2}. If $Ay \equiv \alpha Ax$, $\alpha \not\equiv 0$, it is required that $\tilde{r}_j = \alpha r_j$ for solutions to exist. Hence, we get a total number of $(\frac{p-1}{2})p^{t-1}$ vectors $z \in \mathbb{Z}_p^t$ with $(a_z, b_z) = (1, -1)$, if $\left(\frac{\alpha}{p}\right) = -1$, and 0 solutions, otherwise.

By substituting our results into (1) the claim follows. □

Step 2: We calculate the maximal eigenvalue of $\tilde{M} = M^T M$. We claim that all columns $\tilde{m}(x)$ of $\tilde{M}$, where $x \in \mathbb{Z}_p^u$, already are eigenvectors of $\tilde{M}$. To see this, we compute the inner product of two arbitrary columns with numbers $x, y \in \mathbb{Z}_p^u$. We claim that

$$\tilde{m}(x)^T \tilde{m}(y) = \begin{cases} \left(\frac{\alpha}{p}\right)(p-1)^3 p^{u-\mathrm{rank}(A)} p^{2(t-1)}, & \text{if } Ay \equiv \alpha Ax,\ Ax, Ay \not\equiv 0,\ \alpha \in \\ & \mathbb{Z}_p \backslash \{0\}; \\ 0, & \text{otherwise.} \end{cases}$$

Proof of the claim: Let $Ax, Ay \not\equiv 0$. It holds that

$$\tilde{m}(x)^T \tilde{m}(y) = \sum_{z \in \mathbb{Z}_p^u} \tilde{m}(z, x) \tilde{m}(z, y).$$

We apply our first claim and get:

$$\tilde{m}(z, x)\tilde{m}(z, y) = \begin{cases} \left(\frac{\beta_z^1}{p}\right)\left(\frac{\beta_z^2}{p}\right)(p-1)^2 p^{2(t-1)}, & \text{if } Az \equiv \beta_z^1 Ax,\ Az \equiv \beta_z^2 Ay, \text{ for} \\ & \beta_z^1, \beta_z^2 \not\equiv 0,\ Ax, Ay \not\equiv 0; \\ 0, & \text{otherwise.} \end{cases}$$

We consider the first case, let $Az \equiv \beta_z^1 Ax$, $Az \equiv \beta_z^2 Ay$, $\beta_z^1, \beta_z^2 \not\equiv 0$. Then $Ay \equiv \beta_z^1 (\beta_z^2)^{-1} Ax$, and Ax and Ay are linearly dependent. Therefore, there is an $\alpha \not\equiv 0$ with $Ay \equiv \alpha Ax$, and $\beta_z^1 (\beta_z^2)^{-1} \equiv \alpha$ for all z. Especially, it holds that

$$\left(\frac{\beta_z^1}{p}\right)\left(\frac{\beta_z^2}{p}\right) = \left(\frac{\beta_z^1}{p}\right)\left(\frac{(\beta_z^2)^{-1}}{p}\right) = \left(\frac{\beta_z^1 (\beta_z^2)^{-1}}{p}\right) = \left(\frac{\alpha}{p}\right)$$

for all $z \in \mathbb{Z}_p^u$. How many z-vectors are there for which the first case applies? Their number is obviously equal to the number of solutions of

$$Az \equiv \beta_z^1 Ax,$$

for $\beta_z^1 \not\equiv 0$. There are $p^{u-\mathrm{rank}(A)}$ solutions for fixed β_z^1, and $p-1$ values $\beta_z^1 \not\equiv 0$. Hence, the total number of z-vectors for which the first case applies is $(p-1)p^{u-\mathrm{rank}(A)}$. Putting all the results together, we get the claimed value for $\tilde{m}(x)^T \tilde{m}(y)$. □

From the above claim it follows that

$$\tilde{M} \cdot \tilde{m}(x) = (p-1)^2 p^{t+u-\mathrm{rank}(A)-1} \cdot \tilde{m}(x),$$

for $x \in \mathbb{Z}_3^u$ with $Ax \not\equiv 0$ (to see this, compare the entries with index z, $z \in \mathbb{Z}_p^u$, on both sides). Hence, $\tilde{m}(x)$ is an eigenvector for the eigenvalue $\lambda := (p-1)^2 p^{t+u-\mathrm{rank}(A)-1}$ of $\tilde{M}$. Furthermore, 0 and λ are all the eigenvalues of $\tilde{M}$, since all columns of $\tilde{M}$ are eigenvectors for these eigenvalues. Therefore, λ is the maximal eigenvalue of $\tilde{M} = M^T M$.

Step 3: Now we are ready to estimate the sum of 1's and (-1)'s in the given rectangle of the matrix M. It is a well-known fact from linear algebra that

$$\max\{x^T M^T M x \mid x \in \mathbb{R}^{p^u}, \|x\|_2 \leq 1\} = \lambda,$$

where $\|\cdot\|_2$ denotes the Euclidean norm of real-valued vectors. On the other hand,

$$\begin{aligned}\left(\max\{\|Mx\|_2 \mid x \in \mathbb{R}^{p^u}, \|x\|_2 \leq 1\}\right)^2 &= \max\{\|Mx\|_2^2 \mid x \in \mathbb{R}^{p^u}, \|x\|_2 \leq 1\} \\ &= \max\{x^T M^T M x \mid x \in \mathbb{R}^{p^u}, \|x\|_2 \leq 1\}.\end{aligned}$$

Hence,

$$\max\{\|Mx\|_2 \mid x \in \mathbb{R}^{p^u}, \|x\|_2 \leq 1\} = \sqrt{\lambda} = (p-1)p^{(t+u-\mathrm{rank}(A)-1)/2}.$$

Let 1_S and 1_T be the characteristic vectors of S and T, resp., then we get by the inequality of Cauchy-Schwartz and the trivial bounds $|S| \leq p^t$, $|T| \leq p^u$:

$$\begin{aligned}|d| &= |1_S \cdot M \cdot 1_T| \leq \|1_S\|_2 \cdot \|M \cdot 1_T\|_2 \\ &\leq p^{(t+u)/2} \cdot (p-1)p^{(t+u-\mathrm{rank}(A)-1)/2} = (p-1)p^{t+u}p^{-(\mathrm{rank}(A)+1)/2}.\end{aligned}$$

□

For the next part of Lemma 9, we only consider the case $p = 3$.

Lemma 9, Part (2): Define the $3^t \times 3^u$-Matrix $M = (m(x,y))_{x \in \mathbb{Z}_3^t, y \in \mathbb{Z}_3^u}$ by

$$m(x,y) := \begin{cases} 1, & \text{if } x^T A y \equiv 1; \\ -1, & \text{if } x^T A y \equiv 0; \\ 0, & \text{otherwise.} \end{cases}$$

Let S, T as in part (1) of Lemma 9. Let d be the sum of 1's and (-1)'s of M in the rectangle $S \times T$. Then it holds that

$$|d| \leq 3^{t+u} \cdot 3^{-\mathrm{rank}(A)/2}.$$

Proof: Along the same lines as for part (1). The first step can again be easily done for general p, we substitute $p = 3$ later on to simplify the calculations.

Step 1: Again, let $m(x)$ be the column with number $x \in \mathbb{Z}_p^u$ of M. Then claim that

$$m(x)^T m(y) = \begin{cases} (\frac{3-p}{2})^2 p^{t-2}, & \text{if } Ax \text{ and } Ay \text{ are linearly independent in } \mathbb{Z}_p; \\ (\frac{p+1}{2}) p^{t-1}, & \text{if } Ay \equiv \alpha Ay,\ Ax, Ay \not\equiv 0,\ \left(\frac{\alpha}{p}\right) = 1; \\ p^{t-1}, & \text{if } Ay \equiv \alpha Ay,\ Ax, Ay \not\equiv 0,\ \left(\frac{\alpha}{p}\right) = -1; \\ (\frac{3-p}{2}) p^{t-1}, & \text{if either } Ax \equiv 0 \text{ or } Ay \equiv 0; \\ p^t, & \text{if } Ax \equiv Ay \equiv 0. \end{cases}$$

Proof of the claim: We start with the calculation of the number of $z \in \mathbb{Z}_p^t$ with $m(z,x) = 1$ and $m(z,y) = 1$. We see that we have done this already in the proof of part (1) of Lemma 9,

$$\sum_{(a_z,b_z)=(1,1)} 1 \;=\; \begin{cases} (\frac{p-1}{2})^2 \cdot p^{t-2}, & \text{if } Ax \text{ and } Ay \text{ are linearly independent;} \\ (\frac{p-1}{2}) \cdot p^{t-1}, & \text{if } Ay \equiv \alpha Ax,\ Ax, Ay \not\equiv 0,\ \left(\frac{\alpha}{p}\right) = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Next, we compute the number of $z \in \mathbb{Z}_p^t$ with $m(z,x) = -1$ and $m(z,y) = -1$. We again apply linear algebra in $\mathbb{Z}_p$-vector spaces to do this. The two given conditions are equivalent to

$$z^T Ax \equiv 0 \ \wedge\ z^T Ay \equiv 0,$$

and this system of linear equations in the variables $z_1, \ldots, z_t$ has exactly $p^{t-\text{rank}(Ax,Ay)}$ solutions. More explicitly, we have

$$\sum_{(a_z,b_z)=(-1,-1)} 1 \;=\; \begin{cases} p^{t-2}, & \text{if } Ax \text{ and } Ay \text{ are linearly independent;} \\ p^t, & \text{if } Ax \equiv Ay \equiv 0; \\ p^{t-1}, & \text{otherwise.} \end{cases}$$

Finally, we need the number of $z \in \mathbb{Z}_p^t$ with $m(z,x) = 1$ and $m(z,y) = -1$ (analogously for $m(z,x) = -1$ and $m(z,x) = 1$). The condition $m(z,x) = 1$ is equivalent to

$$\exists\, i \in \{1, \ldots, (p-1)/2\} \colon z^T Ax \equiv r_i \ \wedge\ z^T Ay \equiv 0.$$

Let us consider the number of solutions for fixed i. If Ax and Ay are linearly independent, there are p^{t-2} solutions. If $Ax \equiv 0$, there are no solutions (since $r_i \not\equiv 0$); and if $Ay \equiv 0$ and $Ax \not\equiv 0$, we have exactly p^{t-1} solutions. Finally, for $Ax, Ay \not\equiv 0$ and Ax, Ay linearly dependent, the number of solutions is again 0 ($r_i \not\equiv 0$). Hence,

$$\sum_{(a_z,b_z)=(1,-1)} 1 \;=\; \begin{cases} (\frac{p-1}{2}) \cdot p^{t-2}, & \text{if } Ax \text{ and } Ay \text{ are linearly independent;} \\ (\frac{p-1}{2}) \cdot p^{t-1}, & \text{if } Ax \not\equiv 0,\ Ay \equiv 0; \\ 0, & \text{otherwise.} \end{cases}$$

By summing up our results, we get

$$\begin{aligned}
&m(x)^T \cdot m(y)\\
&= \sum_{(a_z,b_z)=(1,1)} 1 \;+ \sum_{(a_z,b_z)=(-1,1)} (-1) + \sum_{(a_z,b_z)=(1,-1)} (-1) + \sum_{(a_z,b_z)=(-1,-1)} 1\\
&= \begin{cases} (\frac{p-1}{2})^2 p^{t-2} + p^{t-2} - (p-1)p^{t-2}, & \text{if } Ax \text{ and } Ay \text{ are linearly independent;}\\ (\frac{p-1}{2})p^{t-1} + p^{t-1}, & \text{if } Ay \equiv \alpha Ay,\ Ax, Ay \not\equiv 0,\ \left(\frac{\alpha}{p}\right) = 1;\\ p^{t-1}, & \text{if } Ay \equiv \alpha Ay,\ Ax, Ay \not\equiv 0,\ \left(\frac{\alpha}{p}\right) = -1;\\ p^{t-1} - (\frac{p-1}{2})p^{t-1}, & \text{if either } Ax \equiv 0 \text{ or } Ay \equiv 0;\\ p^t, & \text{if } Ax \equiv Ay \equiv 0. \end{cases}
\end{aligned}$$

□

Now let $p = 3$. All the following congruences are modulo 3. By the above claim, we get

$$m(x)^T m(y) = \begin{cases} 2 \cdot 3^{t-1}, & \text{if } Ay \equiv \alpha Ay,\ Ax, Ay \not\equiv 0,\ \alpha \equiv 1;\\ 3^{t-1}, & \text{if } Ay \equiv \alpha Ay,\ Ax, Ay \not\equiv 0,\ \alpha \equiv -1;\\ 3^t, & \text{if } Ax \equiv Ay \equiv 0;\\ 0, & \text{otherwise.} \end{cases}$$

Step 2: As in the proof of part (1) of Lemma 9, we compute the largest eigenvalue of $\tilde{M}$, where $\tilde{M} = M^T M$, $\tilde{m}(x, y) = m(x)^T m(y)$. This is a little bit more complicated here, since the columns of $\tilde{M}$ are no longer eigenvectors. But by the first claim, we will obtain that this matrix has a simple block structure.

The matrix $\tilde{M}$ has only entries 0, $a := 2 \cdot 3^{t-1}$, $b := 3^{t-1}$ and 3^t. Let

$$\begin{aligned} \ker(A) &:= \{x \in \mathbb{Z}_3^u \mid Ax = 0\} \quad \text{and}\\ \operatorname{im}(A) &:= \{y \in \mathbb{Z}_3^t \mid \exists\, x \in \mathbb{Z}_3^u \colon Ax = y\} \end{aligned}$$

Obviously, $\tilde{m}(x, y) = 3^t$, if $x, y \in \ker(A)$ and $\tilde{m}(x, y) = 0$, if $x \in \ker(A)$, but $y \notin \ker(A)$ or vice versa. Next, we consider the vectors $x, y \notin \ker(A)$. For $v \in \operatorname{im}(A)\backslash\{0\}$ define the following subspaces of $\mathbb{Z}_3^u$:

$$U^+(v) := \{x \in \mathbb{Z}_3^u \mid Ax = v\} \quad \text{and} \quad U^-(v) := \{x \in \mathbb{Z}_3^u \mid Ax = -v\}.$$

These sets are either disjoint or equal for different v. There are $\frac{1}{2}(3^{\operatorname{rank}(A)} - 1) =: r$ vectors $v_1, \ldots, v_r$ such that the sets $U^+(v_i), U^-(v_i)$ form a partition of $\operatorname{im}(A)\backslash\{0\}$.

We have shown above that for $v \in \operatorname{im}(A)\backslash\{0\}$ it holds that

$$\tilde{m}(x, y) = \begin{cases} a, & x, y \in U^+(v)\\ b, & x \in U^+(v) \text{ and } y \in U^-(v) \text{ or vice versa;}\\ 0, & x \in (U^+(v) \cup U^-(v)) \text{ and } y \notin (U^+(v) \cup U^-(v)) \text{ or vice versa.} \end{cases}$$

It holds that $|\ker(A)| = 3^{u-\mathrm{rank}(A)} =: k$ and also $|U^+(v)| = |U^-(v)| = k$ for all $v \in \mathrm{im}(A)\backslash\{0\}$. Let P be a $p^u \times p^u$-permutation matrix such that after application of the respective permutation the order of vectors of $\mathbb{Z}_3^u$ is consistent with the following order of subspaces:

$$\ker(A), U^+(v_1), U^-(v_1), U^+(v_2), U^-(v_2), \ldots, U^+(v_r), U^-(v_r).$$

(The order of the vectors within each of these subspaces does not matter.) By the considerations above, we obtain that $\tilde{M}' := P^{-1}\tilde{M}P$ is a block diagonal matrix of the form

$$\tilde{M}' = \mathrm{diag}(B_0, B_1, \ldots, B_r),$$

where the block B_0 is a $k \times k$-matrix with all entries equal to 3^t and the blocks B_i with $i \geq 1$ are $(2k) \times (2k)$-matrices of the form

$$B_i = \left(\begin{array}{c|c} a \ldots a & b \ldots b \\ \vdots \;\; \vdots & \vdots \;\; \vdots \\ a \ldots a & b \ldots b \\ \hline b \ldots b & a \ldots a \\ \vdots \;\; \vdots & \vdots \;\; \vdots \\ b \ldots b & a \ldots a \end{array}\right),$$

each of the four constant submatrices has dimension $k \times k$.
The matrix B_0 has the eigenvalues 0 and $k \cdot 3^t = 3^{u+t-\mathrm{rank}(A)}$, and the matrices B_i, $i \geq 1$, have the eigenvalues $k \cdot (a-b) = 3^{t+u-\mathrm{rank}(A)-1}$, $k \cdot (a+b) = 3^{t+u-\mathrm{rank}(A)}$ and, if $k \geq 2$, also 0. It follows that $\tilde{M}$ altogether has the eigenvalues 0 (if $\mathrm{rank}(A) \leq u-1$), 3^t, $3^{t+u-\mathrm{rank}(A)-1}$ and $3^{t+u-\mathrm{rank}(A)}$, and thus $3^{t+u-\mathrm{rank}(A)}$ is the maximal eigenvalue.

Step 3: This step is analogous to the proof of part (1) of Lemma 9. □

Lemma 9, Part (3): Again, $p = 3$. Define the $3^t \times 3^u$-Matrix $M = (m(x,y))_{x \in \mathbb{Z}_3^t, y \in \mathbb{Z}_3^u}$ by

$$m(x,y) := \begin{cases} 1, & \text{if } x^TAy \equiv -1; \\ -1, & \text{if } x^TAy \equiv 0; \\ 0, & \text{otherwise.} \end{cases}$$

Let S, T be as above. Let d be the sum of 1's and (-1)'s of M in the rectangle $S \times T$. Then it holds that

$$|d| \leq 3^{t+u} \cdot 3^{-\mathrm{rank}(A)/2}.$$

Proof: Analogous to the proof of part (2) of Lemma 9 due to the "duality" of the values -1 and 1. □

References

Ablayev, F. (1996). Randomization and nondeterminism are incomparable for polynomial ordered binary decision diagrams. *Accepted at ICALP '97.*

Ablayev, F. and M. Karpinski (1996a). On the power of randomized branching programs. In *Proc. of ICALP '96*, *LNCS 1099*, 348 – 356. Springer-Verlag.

Ablayev, F. and M. Karpinski (1996b, December). On the power of randomized ordered branching programs. *Manuscript.*

Ajtai, M. and M. Ben-Or (1984). A theorem on probabilistic constant depth computations. In *Proc. of the 16th Ann. ACM Symp. on Theory of Computing*, 471 – 474.

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

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

Breitbart, Y., H. Hunt III, and D. Rosenkrantz (1995). On the size of binary decision diagrams representing Boolean functions. *Theoretical Computer Science 145*, 45 – 69.

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

Bryant, R. E. (1991). 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.

Bryant, R. E. (1992). Symbolic Boolean manipulation with ordered binary-decision diagrams. *ACM Computing Surveys 24*(3), 293–318.

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

Gill, J. (1972). *Probabilistic Turing Machines and Complexity of Computations.* Ph. D. dissertation, U. C. Berkeley.

Hosaka, K., Y. Takenaga, and S. Yajima (1994). 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-Verlag.

Hromkovič, J. (1997). *Communication Complexity and Parallel Computing.* Springer-Verlag.

Jukna, S. P. (1989). On the effect of null-chains on the complexity of contact schemes. In *Proc. of Fundamentals of Computation Theory*, *LNCS 380*, 246–256. Springer-Verlag.

Jukna, S. P. (1995). A note on read-k-times branching programs. *Theoretical Informatics and Applications 29*(1), 75 – 83.

Krause, M. (1991). Lower bounds for depth-restricted branching programs. *Information and Computation 91*(1), 1–14.

Krause, M., C. Meinel, and S. Waack (1988). Separating the eraser turing machine classes L_e, NL_e, co-NL_e and P_e. In *Proc. of MFCS*, *LNCS 324*, 405–413. Springer-Verlag.

Kremer, I., N. Nisan, and D. Ron (1994, November). On randomized one-round communication complexity. *Manuscript.*

Kushilevitz, E. and N. Nisan (1997). *Communication Complexity.* Cambridge University Press.

Meinel, C. (1988). *Modified Branching Programms and Their Computational Power.* Habilitationsschrift, Humboldt-Universität Berlin.

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

Razborov, A. A. (1991). Lower bounds for deterministic and nondeterministic branching programs. In *Proc. of Fundamentals of Computation Theory, LNCS 529*, 47–60. Springer-Verlag.

Razborov, A. A. (1992). On the distributional complexity of disjointness. *Theoretical Computer Science 106*, 385 – 390.

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

Yao, A. C. (1983). Lower bounds by probabilistic arguments. In *Proc. of the IEEE Symp. on Foundations of Computer Science*, Volume 27, 420 – 428.