

Classical vs. Quantum Read-Once Branching Programs

Martin Sauerhoff

LS 2, FB Informatik, Univ. Dortmund
44221 Dortmund, Germany
sauerhoff@ls2.cs.uni-dortmund.de

Abstract. A simple, explicit boolean function on $2n$ input bits is presented that is computable by errorfree quantum read-once branching programs of size $O(n^3)$, while each classical randomized read-once branching program and each quantum OBDD for this function with bounded two-sided error requires size $2^{\Omega(n)}$.

Keywords. Quantum branching program, randomized branching program, read-once.

1 Introduction

This paper deals with the comparison of the space complexity of sequential randomized and quantum algorithms in the nonuniform setting, modeled by restricted quantum branching programs. Since the general model of branching programs is still only very poorly understood when it comes to proving lower bounds, it is natural to consider reasonably restricted variants of the model first.

Lower bounds and separation results generally come in two main flavors: results for multi-output-bit functions and for single-output-bit functions or decision problems. Of the former type are recent time-space tradeoffs for quantum circuits computing some practically important functions, including sorting [11,1,13] and boolean matrix-vector and matrix-matrix multiplication [13,12].

Here we are concerned with separation results for decision problems, for which proving lower bounds is usually much harder than for multi-output-bit problems in the same model. Such separation results have been proved for the uniform model of quantum finite automata (QFAs, see, e. g., [14,17,4]). On the nonuniform side, general quantum branching programs and quantum OBDDs (ordered binary decision diagrams) have been considered (see the next section for an introduction of these models). Extending independently obtained results by Špalek [25], it has been shown in [23] that the logarithm of the size of general quantum branching programs captures the space complexity of nonuniform quantum Turing machines. Ablayev, Moore, and Pollett [3] have proved that NC^1 is included in the class of functions that can be exactly computed by quantum oblivious width-2 branching programs of polynomial size, in contrast to the classical case where width 5 is necessary unless $\text{NC}^1 = \text{ACC}$. Furthermore,

exponential gaps have been established between the width of quantum OBDDs and classical deterministic OBDDs (Ablayev, Gainutdinova, and Karpinski [2]) and classical randomized OBDDs, resp. (Nakanishi, Hamaguchi, and Kashiwabara [18]). Finally, it has been shown in [23] that the classes of functions with polynomial size quantum OBDDs and deterministic OBDDs are incomparable and an example of a partially defined function for which quantum OBDDs are exponentially smaller than classical randomized ones has been presented.

Previous results about quantum variants of branching programs for decision problems have been limited to models that are oblivious, i. e., are required to read their input bits in a fixed order. Here we consider the non-oblivious model of quantum read-once branching programs, which are quantum branching programs that during each computation may access each input bit at most once. The logarithm of the size of quantum read-once branching programs is a lower bound on the space-complexity of (uniform or nonuniform) quantum read-once Turing machines. This follows by an easy adaptation of the proof in [23] for general quantum branching programs. On the other hand, the upper bound presented here in terms of quantum read-once branching programs can easily be modified to work also for (uniform or nonuniform) quantum read-once Turing machines.

As the main result of this paper, we present a simple function for which quantum read-once branching programs are exponentially smaller than classical randomized ones. This result is even for a total function (compare this to the fact that analogous results for quantum OBDDs [23] and quantum one-way communication complexity [5] known so far are only for partially defined functions). We use the *weighted sum function* due to Savický and Žák [24] as a building block. For a positive integer n and $x = (x_1, \dots, x_n) \in \{0, 1\}^n$, let $p(n)$ be the smallest prime larger than n and let $s_n(x) = (\sum_{i=1}^n i \cdot x_i) \bmod p(n)$. Define the weighted sum function by $WS_n(x) = x_{s_n(x)}$ if $s_n(x) \in \{1, \dots, n\}$ and 0 otherwise. For a further input vector $y = (y_1, \dots, y_n) \in \{0, 1\}^n$ define the *mixed weighted sum function* by $MWS_n(x, y) = x_i \oplus y_i$ if $i = s_n(x) = s_n(y) \in \{1, \dots, n\}$ and 0 otherwise.

Theorem 1: *Each randomized read-once branching program and each quantum OBDD computing MWS_n with two-sided error bounded by an arbitrary constant smaller than $1/2$ requires size $2^{\Omega(n)}$, while MWS_n can be computed by an error-free quantum read-once branching program of size $O(n^3)$.*

The above result shows that being able to choose different variable orders for different inputs may help a lot for quantum read-once algorithms, even compared to classical randomized read-once algorithms that are allowed the same option.

The rest of the paper is organized as follows: In the next section, we define the variants of quantum branching programs considered here. In the section following that, we present the proof of the main result.

2 Preliminaries

We assume a general background on quantum computing (as provided, e. g., by the textbook of Nielsen and Chuang [19]) and on classical branching pro-

grams (BPs) (see, e. g., the textbook of Wegener [28]). We start with the definition of general quantum branching programs.

Definition 1: A quantum branching program (QBP) over the variable set $X = \{x_1, \dots, x_n\}$ is a directed multigraph $G = (V, E)$ with a start node $s \in V$ and a set $F \subseteq V$ of sinks. Each node $v \in V - F$ is labeled by a variable $x_i \in X$ and we define $\text{var}(v) = i$. Each node $v \in F$ carries a label from $\{0, 1\}$, denoted by $\text{label}(v)$. Each edge $(v, w) \in E$ is labeled by a boolean constant $b \in \{0, 1\}$ and a (transition) amplitude $\delta(v, w, b) \in \mathbb{C}$. We assume that there is at most one edge carrying the same boolean label between a pair of nodes and set $\delta(v, w, b) = 0$ for all $(v, w) \notin E$ and $b \in \{0, 1\}$.

The graph G is required to satisfy the following two constraints. First, it has to be well-formed, meaning that for each pair of nodes $u, v \in V - F$ and all assignments $a = (a_1, \dots, a_n)$ to the variables in X ,

$$\sum_{w \in V} \delta^*(u, w, a_{\text{var}(u)}) \delta(v, w, a_{\text{var}(v)}) = \begin{cases} 1, & \text{if } u = v; \\ 0, & \text{otherwise.} \end{cases}$$

Second, G has to be unidirectional, which means that for each $w \in V$, all nodes $v \in V$ such that $\delta(v, w, b) \neq 0$ for some $b \in \{0, 1\}$ are labeled by the same variable.

A computational state of the QBP is a pure quantum state over the Hilbert space $\mathcal{H} = \mathbb{C}^{|V|}$ spanned by an ON-basis $(|v\rangle)_{v \in V}$. The computation for an input $a = (a_1, \dots, a_n)$ starts with the computational state $|s\rangle$, called initial state. Let the QBP be in the computational state $|\psi\rangle = \sum_{v \in V} \alpha_v |v\rangle \in \mathcal{H}$ at the beginning of a computation step. Then the QBP first carries out a projective measurement of the output label at the nodes in $|\psi\rangle$. This yields the result $r \in \{0, 1\}$ with probability $\sum_{v \in F, \text{label}(v)=r} |\alpha_v|^2$. If one of these events occurs, the respective output is produced and the computation stops. The computation carries on for the non-sink nodes with nonzero amplitude in $|\psi\rangle$. Let $|\psi'\rangle = \sum_{v \in V-F} \alpha'_v |v\rangle$ be the state obtained by projecting $|\psi\rangle$ to the subspace spanned by the non-sink nodes and renormalizing. Then the next computational state is defined as $|\psi''\rangle = \sum_{v \in V-F} \alpha'_v \sum_{w \in V} \delta(v, w, a_{\text{var}(v)}) |w\rangle$.

The probability that G outputs $r \in \{0, 1\}$ on input $a \in \{0, 1\}^n$ is defined as the sum of the probabilities of obtaining the output r after any finite number of steps. Let $G(a)$ be the random variable describing the output of G on input a , called the output random variable of G for a . We say that the function $f: \{0, 1\}^n \rightarrow \{0, 1\}$ defined on X is computed by G

- with two-sided error at most ε , $0 \leq \varepsilon < 1/2$, if for each $a \in \{0, 1\}^n$, $\Pr\{G(a) \neq f(a)\} \leq \varepsilon$; and it is computed
- exactly (or G is an error-free QBP for f), if for each $a \in \{0, 1\}^n$, $\Pr\{G(a) \neq f(a)\} = 0$.

Furthermore, by bounded two-sided error we mean two-sided error with some unspecified constant bound ε . (Other modes of acceptance may be defined as usual for other quantum models of computation.)

The size of a QBP G is the number of its nodes and is denoted by $|G|$. Its width is the maximum number of nodes with the same distance from the start node.

The definition of QBPs is similar to that of the uniform models of quantum finite automata (QFAs) and quantum Turing machines (QTMs), whose relationships to the respective classical models have already been studied to a considerable extent (see, e. g., [14,17,4,6,26,27]). A strong motivation why QBPs are a natural model is provided by the fact that the logarithm of their size and the space complexity for nonuniform QTMs are polynomially related [25,23]. For the scenario of sublinear space bounds, it has turned out to be useful to work with *unidirectional* QTMs, i. e., QTMs whose directions of head movements depend only on the entered state of the finite control. This is the standard model in the papers of Watrous [26,27] and also that used for the simulation between QBPs and QTMs in [25,23]. The unidirectionality constraint for QBPs (called *parental condition* in [25]) turns up as a natural counterpart of that for QTMs required to make the simulations work. In order to prevent QBPs from being unreasonably powerful, it is further realistic to restrict the set of allowed amplitudes, see also [23]. This is no issue here, since the upper bound for quantum read-once branching programs in the paper only uses amplitudes from $\{0, 1, \pm 1/2\}$.

For the construction of QBPs it is sometimes convenient to use *unlabeled nodes* with an arbitrary number of outgoing edges carrying only amplitude labels. An unlabeled node v can be regarded as an abbreviation for a node according to the standard definition labeled by a dummy variable on which the considered function does not depend. Each edge leading from the unlabeled node v to a successor w with amplitude α is then regarded as a pair of edges from the node labeled by the dummy variable to w that carry the boolean labels 0 and 1, resp., and that both have amplitude α .

A special case of QBPs are *reversible* classical BPs, where each node is reachable from at most one node v by a 0-edge and from at most one node w by a 1-edge and v and w are labeled by the same variable. It has been proved by Špalek [25] that each sequence of (possibly non-reversible) classical BPs with at least linear size can be simulated by a sequence of reversible ones with at most polynomial larger size. Since randomized (general) BPs can be derandomized while maintaining polynomial size analogously to probabilistic circuits (see [21] for details), the same is true in the randomized case.

We consider the following variants of quantum BPs defined analogously to their classical counterparts.

Definition 2:

- A quantum read-once BP is a QBP where each variable may appear at most once on each path.
- A quantum OBDD (quantum ordered binary decision diagram) is a quantum read-once BP with an order π of the variables such that for each path in the graph the order in which the variables appear is consistent with π .

3 Proof of the Main Theorem

For the whole section, let $p = p(n)$ be the smallest prime larger than n for a fixed positive integer n . We first deal with the easier upper bound. Our goal is to show that MWS_n can be computed by polynomially small error-free quantum read-once BPs.

Proof of Theorem 1 – Upper Bound: The essence of the proof is to apply the Deutsch-Jozsa algorithm, evaluating the sums $s_n(x)$ and $s_n(y)$ in parallel and computing the output $x_i \oplus y_i$ if $i = s_n(x) = s_n(y)$. We first describe the algorithm by a quantum circuit. We use a four-part quantum register consisting of two qubits for the Deutsch-Jozsa algorithm and two further parts whose basis states are indexed by $\{0, \dots, p-1\}$. The oracle gate for the Deutsch-Jozsa algorithm unitarily extends the mapping S specified for $a, b \in \{0, 1\}$ by $|a\rangle|b\rangle|0\rangle|0\rangle \mapsto |a\rangle|b \oplus (1-a)y_i \oplus ax_j\rangle|i\rangle|j\rangle$, where $i = s_n(x)$ and $j = s_n(y)$. This gate is applied to the initial state $(1/2)(|0\rangle + |1\rangle)(|0\rangle - |1\rangle)|0\rangle|0\rangle$, giving the final state $(1/2)((-1)^{y_i}|0\rangle + (-1)^{x_j}|1\rangle)(|0\rangle - |1\rangle)|i\rangle|j\rangle$. If a measurement of the last two parts of the quantum register yields that $i \neq j$, the output of the circuit is 0 with probability 1. Otherwise, $i = j$ and measuring the first two qubits in the Hadamard basis yields the output $x_i \oplus y_i = \text{MWS}_n(x, y)$ for the first qubit with probability 1.

Next we describe the implementation of the obtained quantum circuit as a quantum read-once BP. For an easier exposition, we first use unlabeled nodes. We start with the construction of a subgraph G_S realizing the mapping S . The nodes of G_S are laid out on a grid with $2n+1$ rows and $4p^2$ columns, the latter labeled by (a, b, i, j) with $a, b \in \{0, 1\}$ and $i, j \in \{0, \dots, p-1\}$. Each row represents an intermediate state of the four-part quantum register used for the above algorithm. The graph G_S consists of two disjoint classical reversible OBDDs G_0 and G_1 on the subsets of nodes in the columns with $a = 0$ and $a = 1$, resp. We first describe how G_0 works. The computation starts at a node in row 1 and column $(0, b, 0, 0)$ with $b \in \{0, 1\}$. The variable vector x is read (the order of the variables within the vector does not matter) and the node in row $n+1$ and column $(0, b, s_n(x), 0)$ is reached. Then the variable vector y is read (again, the order of the individual variables is arbitrary) and the sink in row $2n+1$ and column $(0, b \oplus y_{s_n(x)}, s_n(x), s_n(y))$ is reached. It is easy to see how the described computation can be implemented by a reversible OBDD with nodes on the prescribed grid. The OBDD G_1 works in the same way, but with exchanged roles of x and y and exchanged roles of the last two column indices. Altogether, we obtain a classical reversible read-once BP for G_S with at most $(2n+1) \cdot 4p^2$ nodes, which is of order $O(n^3)$ due to the prime number theorem.

We add a new, unlabeled source that for $(a, b) \in \{0, 1\}^2$ is connected to the node in row 1 and column $(a, b, 0, 0)$ of G_S by an edge with amplitude $(-1)^b(1/2)$. The sinks of G_S in row $2n+1$ and in columns (a, b, i, j) with $i \neq j$ are replaced with 0-sinks. All other sinks of G_S are replaced with unlabeled nodes connected to a new level of sinks with boolean output labels. The outgoing edges of these unlabeled nodes are labeled by amplitudes such that, together with the sinks,

a measurement in the Hadamard basis is realized. The whole graph still has size $O(n^3)$.

Finally, we remove the unlabeled nodes. For this, we first ensure that all nodes on the first level of G_S are labeled by the same variable and the same for all nodes on the last level of G_S with variable labels. We rearrange (e.g.) the variable order of the OBDD G_1 and update the OBDD accordingly. W.l.o.g., let x_1 be the first variable read in G_0 and let y_n be the last. We move the variable x_1 to the front of the variable order of G_1 and y_n to the end. It is not hard to see that we can modify G_1 in such a way that it complies to the new variable order while increasing its size by at most a constant factor and maintaining reversibility. After this transformation, we merge the unlabeled nodes with their successors (in the case of the source) or with their predecessors (in the case of the nodes on the level directly above the sinks). It is obvious how the edges should be relabeled such that the resulting graph still computes the same final state as a quantum read-once BP. We observe that after the reordering process also the unidirectionality requirement for quantum BPs is satisfied. Altogether, we have obtained the desired quantum read-once BP for MWS_n of size $O(n^3)$. \square

Next we prove the lower bound on the size of randomized read-once BPs for MWS_n with bounded error. We reuse main ideas from the proof an analogous lower bound for WS_n in [22]. However, the result for MWS_n is no obvious consequence of that for WS_n . We have to carefully argue why, different from the quantum case, having two input vectors present that play the same roles does not help in the randomized case.

The proof employs a variant of the rectangle bound method from communication complexity theory (see, e.g., the textbook of Kushilevitz and Nisan [16]) suitable for read-once BPs, which we first describe. For this, we introduce some notation. We consider boolean functions defined on the union of the disjoint sets of variables $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_n\}$. For a set of variables $Z \subseteq X \cup Y$, let 2^Z denote the set of all assignments to Z , i.e., mappings from Z to $\{0, 1\}$ that we usually identify with vectors in $\{0, 1\}^{|Z|}$. A (*combinatorial*) *rectangle* with respect to a partition $\Pi = (\Pi_1, \Pi_2)$ of $X \cup Y$ is a set of assignments $R = A \times B$ with $A \subseteq 2^{\Pi_1}$ and $B \subseteq 2^{\Pi_2}$. For $\ell \in \{1, \dots, n-1\}$ call R an ℓ -*rectangle* if Π_1 contains exactly ℓ variables from X and at most $\ell-1$ variables from Y or the same with exchanged roles of X and Y . Call R a *one-way rectangle* if $B = 2^{\Pi_2}$. Given a function g on $X \cup Y$, R is said to be *g-uniform* if for all $a, a' \in A$ and $b \in B$, $g(a, b) = g(a', b)$.

For the following, let a function f on $X \cup Y$ and a distribution \mathcal{D} on the inputs of f be given. Let $0 \leq \varepsilon < 1/2$. We describe how to prove lower bounds for deterministic read-once BPs whose output is allowed to differ from f on at most an ε -fraction of the inputs with respect to \mathcal{D} . By a well-known averaging argument due to Yao [29], this also gives lower bounds of the same size for randomized read-once BPs computing f with the same error probability.

The essence of the proof technique is to show that, on the one hand, any small deterministic read-once BP that correctly computes f on a large fraction of the inputs with respect to \mathcal{D} would give a rectangle with large \mathcal{D} -measure on which f

is well approximated, while on the other hand, using the specific properties of f , the \mathcal{D} -measure of any such rectangle necessarily has to be small. We now make this more precise. Let $R = A \times B$ be a rectangle and let $0 \leq \varepsilon < 1/2$. A function g on $X \cup Y$ is said to *uniformly approximate f on R with error ε with respect to \mathcal{D}* , if for all $a \in A$, g differs from f for at most an ε -fraction of the inputs in $\{a\} \times B$ with respect to \mathcal{D} . The following main lemma of the proof technique is a variant of a similar statement from [22], where the uniform distribution and functions on a single set of variables have been considered.

Lemma 1: *Let $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_n\}$. Let f be a boolean function on $X \cup Y$ and let \mathcal{D} be a distribution on the inputs of f . Let $\ell \in \{1, \dots, n-1\}$ and $0 \leq \varepsilon < \varepsilon' < 1/2$. Then for every deterministic read-once BP G computing a function g that differs from f on at most an ε -fraction of the inputs with respect to \mathcal{D} there is a one-way ℓ -rectangle R that is g -uniform, on which g uniformly approximates f with error at most ε' with respect to \mathcal{D} , and which satisfies $\mathcal{D}(R) \geq (1 - \varepsilon/\varepsilon')/(2n|G|)$.*

Proof: By an easy adaptation of the well-known proof technique of Borodin, Razborov, and Smolensky [7] (see also [28], Section 7.6), we get a partition of the input space into at most $k \leq 2n|G|$ one-way ℓ -rectangles $R_1 = A_1 \times B_1, \dots, R_k = A_k \times B_k$ that are all g -uniform. We claim that there is an $i \in \{1, \dots, k\}$ and a subset $A'_i \subseteq A_i$ such that for $R = A'_i \times B_i$, $\mathcal{D}(R) \geq (1 - \varepsilon/\varepsilon')/k$ and g uniformly approximates f on R with error ε' with respect to \mathcal{D} . This obviously suffices to prove the claim.

Let $A^* = A_1 \cup \dots \cup A_k$. For each $x \in A^*$, let $(\Pi_1(x), \Pi_2(x))$ be the partition of the input variables used by the rectangle to which x belongs, and let $S_x = \{x\} \times 2^{\Pi_2(x)}$. Let $A = \{x \in A^* \mid \mathcal{D}(S_x) > 0\}$. For each $x \in A$ let $\varepsilon(x)$ be the \mathcal{D} -fraction of inputs from S_x for which g differs from f . Due to the definitions, the sets S_x , $x \in A$, are disjoint and their union has \mathcal{D} -measure 1. Hence, by the law of total probability, $\sum_{x \in A} \varepsilon(x) \mathcal{D}(S_x) \leq \varepsilon$. Let $A' = \{x \in A \mid \varepsilon(x) \leq \varepsilon'\}$ and let S be the union of all S_x for $x \in A'$. By Markov's inequality, $\mathcal{D}(S) \geq 1 - \varepsilon/\varepsilon'$. By averaging, there is a set $A'' \subseteq A'$ such that for the union S' of all S_x with $x \in A''$, we have $\mathcal{D}(S') \geq \mathcal{D}(S)/k$ and all inputs from A'' belong to the same rectangle. Let (Π_1, Π_2) be the partition of input variables of this rectangle. It is now obvious that the set $R = A'' \times 2^{\Pi_2}$ with $A'' \subseteq 2^{\Pi_1}$ and $\mathcal{D}(R) \geq (1 - \varepsilon/\varepsilon')/k$ is a one-way ℓ -rectangle with the desired properties. \square

Next we cite technical lemmas also used in [22] that build the common core of the lower bounds both for the mixed weighted sum function MWS_n and the usual weighted sum function WS_n . The first lemma, due to Erdős and Heilbronn [8] (reproved independently in [22]), allows us to argue that partial weighted sums of enough random bits are essentially uniformly distributed over the whole range of possible values.

Lemma 2 ([8]): *Let $q = q(n)$ be a sequence of primes and let $n \leq q-1$ and $n = \Omega(q^{2/3+\delta})$ for any constant $\delta > 0$. Let $a_1, \dots, a_n, b \in \mathbb{Z}_q^* = \mathbb{Z}_q - \{0\}$ where the numbers a_1, \dots, a_n are pairwise different. Then for $(x_1, \dots, x_n) \in \{0, 1\}^n$ chosen uniformly at random, $|\Pr\{a_1x_1 + \dots + a_nx_n \equiv b \pmod{q}\} - 1/q| = 2^{-\Omega(q^{3\delta})}$.*

Furthermore, we need the following classical fact about the distribution of primes, which follows from a more precise bound due to Hoheisel [9] (see, e. g., the monograph of Ribenboim [20], pp. 252–254, for more details and improved bounds):

Lemma 3 ([9]): $p(n) = n + o(n)$.

In the final technical lemma, we consider the *index function* $\text{IND}_n: \{0, 1\}^n \times \{1, \dots, n\} \rightarrow \{0, 1\}$ from communication complexity theory defined for $u \in \{0, 1\}^n$ and $v \in \{1, \dots, n\}$ by $\text{IND}_n(u, v) = u_v$. We state an upper bound on the size of one-way rectangles on which IND_n is well approximated that is implicit in a couple of papers, the earliest one being probably that of Kremer, Nisan, and Ron [15]. For the sake of completeness, we include the easy proof. Here and in the following, U denotes the uniform distribution on the domain implied by its respective argument.

Lemma 4 ([15]): *Let ε be a constant with $0 \leq \varepsilon < 1/2$. Let $R = A \times \{1, \dots, n\}$ with $A \subseteq \{0, 1\}^n$ be a one-way rectangle for which a function g exists such that R is g -uniform and g uniformly approximates IND_n on R with error ε with respect to U . Then $U(R) = 2^{-\Omega(n)}$.*

Proof: Since R is g -uniform, there is a vector $r \in \{0, 1\}^n$ such that, for each $a \in A$, $(g(a, 1), \dots, g(a, n)) = r$. Since g uniformly approximates IND_n on R with error at most ε with respect to the uniform distribution, r has Hamming distance at most $\lfloor \varepsilon n \rfloor$ to each vector in A . It follows that $|A|$ is upper bounded by the size of Hamming balls of radius $\lfloor \varepsilon n \rfloor$, which is known to be at most $2^{H(\varepsilon)n}$, where $H(x) = -(x \log x + (1-x) \log(1-x))$ for $x \in [0, 1]$ is the binary entropy function. Thus, $U(R) = |R|/(n \cdot 2^n) = |A|/2^n \leq 2^{-(1-H(\varepsilon))n} = 2^{-\Omega(n)}$. \square

Now we describe the details that are particular to the function MWS_n . For the rest of the section, let $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_n\}$ be the sets of variables on which MWS_n is defined. Recall that $p = p(n)$ is the smallest prime larger than n . We concentrate on the set of difficult inputs $D = \{(x, y) \mid s_n(x) = s_n(y)\}$ by working with the distribution \mathcal{D} with $\mathcal{D}(x, y) = 1/|D|$ if $(x, y) \in D$ and $\mathcal{D}(x, y) = 0$ otherwise.

As a preparation of the proof of the lower bound for randomized read-once BPs computing MWS_n , we derive some basic facts about the considered one-way rectangles. We use the following notation. For a set $S \subseteq X$ (or $S \subseteq Y$) of variables and a partial assignment a that fixes at least all variables in S , let $\sigma_S(a) = (\sum_{v \in S} i(v) \cdot a(v)) \bmod p$, where $i(v) \in \{1, \dots, n\}$ denotes the index of the variable v in X (or Y , resp.), and $a(v)$ is the value that it obtains by the assignment a .

Lemma 5: *Let $\ell = n - \Theta(p^{2/3+\delta})$ for some constant δ with $0 < \delta < 1/3$. Let $\Pi = (\Pi_1, \Pi_2)$ be a partition of $X \cup Y$ with $|\Pi_1 \cap X| = \ell$ and $|\Pi_1 \cap Y| \leq \ell - 1$. Let $R = A \times 2^{\Pi_2}$ with $A \subseteq 2^{\Pi_1}$ and suppose there are $i_x, i_y \in \{0, \dots, p-1\}$ such that for all $a \in A$, $\sigma_{\Pi_1 \cap X}(a) = i_x$ and $\sigma_{\Pi_1 \cap Y}(a) = i_y$. For each $k \in \{0, \dots, p-1\}$ define B_k as the set of all assignments $b \in 2^{\Pi_2}$ with $\sigma_{\Pi_2 \cap X}(b) \equiv (k - i_x) \bmod p$ and $\sigma_{\Pi_2 \cap Y}(b) \equiv (k - i_y) \bmod p$. Then we have the following.*

- (i) For each $k \in \{0, \dots, p-1\}$ and $(a, b) \in A \times B_k$, $\sigma_X(a, b) = \sigma_Y(a, b) = k$. Furthermore, $U(B_k) = (1/p^2) \cdot (1 \pm o(1))$ and $\mathcal{D}(A \times B_k) = (1/p) \cdot U(R) \cdot (1 \pm o(1))$.
- (ii) $\mathcal{D}(R) = U(R) \cdot (1 \pm o(1))$.

Proof: Part (i): The first part of the statement is obvious. It remains to prove the claims about $U(B_k)$ and $\mathcal{D}(A \times B_k)$. Let b denote an assignment from B_k chosen uniformly at random. Then, using that disjoint parts of b are independent of each other and applying Lemma 2, we get

$$\begin{aligned} U(B_k) &= \Pr\{\sigma_{\Pi_2 \cap X}(b) \equiv k - i_x \wedge \sigma_{\Pi_2 \cap Y}(b) \equiv k - i_y\} \\ &= \Pr\{\sigma_{\Pi_2 \cap X}(b) \equiv k - i_x\} \cdot \Pr\{\sigma_{\Pi_2 \cap Y}(b) \equiv k - i_y\} = \frac{1}{p^2} \cdot (1 \pm o(1)). \end{aligned}$$

Furthermore, also by Lemma 2, $U(D) = (1/p) \cdot (1 \pm o(1))$. Again by the independence of disjoint parts of uniformly random assignments and by observing that $A \times B_k \subseteq D$ and $U(A) = U(R)$, we obtain

$$\mathcal{D}(A \times B_k) = \frac{U((A \times B_k) \cap D)}{U(D)} = \frac{U(A) \cdot U(B_k)}{U(D)} = \frac{1}{p} \cdot U(R) \cdot (1 \pm o(1)).$$

Part (ii): This follows from the first part, since $R \cap D$ is the disjoint union of the sets $A \times B_k$ over all $k = 0, \dots, p-1$. \square

Finally, we are ready to prove the desired lower bound on the size of randomized read-once BPs for MWS_n .

Proof of Theorem 1 – Lower bound for randomized read-once BPs: Following the outline above, we prove the lower bound for deterministic read-once BPs that correctly compute MWS_n on a large fraction of the inputs. Let $0 \leq \varepsilon_G < 1/2$ be any constant and let G be a deterministic read-once BP computing a function g that differs from MWS_n on at most an ε_G -fraction of the inputs with respect to \mathcal{D} . Choose $\ell = n - \Theta(p^{2/3+\delta})$ for a some constant δ with $0 < \delta < 1/3$. Let ε be a constant with $\varepsilon_G < \varepsilon < 1/2$. Let R be a one-way ℓ -rectangle that is g -uniform and on which MWS_n is uniformly approximated by g with error at most ε . We prove that $\mathcal{D}(R) = 2^{-\Omega(n)}$. By Lemma 1, this yields the desired lower bound $|G| = 2^{\Omega(n)}$.

Let $\Pi = (\Pi_1, \Pi_2)$ be the partition of the input variables used by R , where w.l.o.g. $|\Pi_1 \cap X| = \ell$ and $|\Pi_1 \cap Y| \leq \ell - 1$. Let $R = A_R \times 2^{\Pi_2}$ with $A_R \subseteq 2^{\Pi_1}$. Using averaging, we fix an assignment $a \in 2^{\Pi_1 \cap Y}$ and an $i_x \in \{0, \dots, p-1\}$ such that for the set A of all assignments $a' \in A_R$ that are consistent with a and satisfy $\sigma_{\Pi_1 \cap X}(a') = i_x$, we have $\mathcal{D}(A \times 2^{\Pi_2}) \geq \mathcal{D}(R)/(p \cdot 2^{|\Pi_1 \cap Y|})$. Let $i_y = \sigma_{\Pi_1 \cap Y}(a)$. Let $R' = \{x \in A \times 2^{\Pi_2} \mid \mathcal{D}(x) > 0\}$. Since g approximates MWS_n uniformly on R with error at most ε with respect to \mathcal{D} , we know that g differs from MWS_n for at most an ε -fraction of the inputs in R' with respect to \mathcal{D} .

Let $\Pi_1 \cap X = \{x_{j_1}, \dots, x_{j_\ell}\}$. We observe that, due to Lemma 3, $p \leq n + o(n)$ and thus $\ell \geq n - o(n)$ and $\ell/p \geq 1 - o(1)$. Let $B_0, \dots, B_{p-1} \subseteq 2^{\Pi_2}$ be the sets of assignments according to Lemma 5 for R' and i_x, i_y . Let $B = B_{j_1} \cup \dots \cup B_{j_\ell}$. Then we have the following.

Claim 1: *The function g differs from MWS_n on at most a fraction of $\varepsilon \cdot (1 + o(1))$ of the inputs in $A \times B$ with respect to the uniform distribution.*

Proof of Claim 1: Due to part (i) of Lemma 5, $\mathcal{D}(A \times B) \geq (\ell/p) \cdot U(R') \cdot (1 - o(1)) \geq U(R') \cdot (1 - o(1))$. On the other hand, by part (ii) of Lemma 5, $\mathcal{D}(R') \leq U(R') \cdot (1 + o(1))$. Thus, the inputs in $A \times B$ cover at least a $(1 - o(1))$ -fraction of the rectangle R' with respect to \mathcal{D} . It follows that g differs from MWS_n on at most a fraction of $\varepsilon \cdot (1 + o(1))$ of the inputs in $A \times B$ with respect to \mathcal{D} . Since $A \times B \subseteq D$, the same is true for the uniform distribution. \square

Next we further reduce the obtained set $A \times B$ by picking appropriate representatives of each of the subsets $B_{j_1}, \dots, B_{j_\ell}$ of B .

Claim 2: *There are $b_1 \in B_{j_1}, \dots, b_\ell \in B_{j_\ell}$ such that g differs from MWS_n on at most a fraction of $\varepsilon \cdot (1 + o(1))$ of the inputs in $R'' = A \times \{b_1, \dots, b_\ell\}$ with respect to the uniform distribution.*

Proof of Claim 2: We choose a collection of disjoint subsets $\{b_1, \dots, b_\ell\}$ of B with $b_1 \in B_{j_1}, \dots, b_\ell \in B_{j_\ell}$ whose union B' is as large as possible. Since $U(B_k) \geq (1/p^2) \cdot (1 - o(1))$ for each $k = 0, \dots, p-1$ by part (i) of Lemma 5, we can ensure that $U(B') \geq (\ell/p^2) \cdot (1 - o(1)) \geq (1/p) \cdot (1 - o(1))$. On the other hand, also by Lemma 5, $U(B) \leq (1/p) \cdot (1 + o(1))$. Hence, the set $A \times B'$ covers at least a $(1 - o(1))$ -fraction of the inputs in $A \times B$. It follows that the relative error of g on $A \times B'$ with respect to the uniform distribution is bounded by some ε' with $\varepsilon' \leq \varepsilon \cdot (1 + o(1))$. By averaging, there is thus at least one subset $\{b_1, \dots, b_\ell\}$ in B' such that $A \times \{b_1, \dots, b_\ell\}$ has relative error ε' with respect to the uniform distribution. \square

Let $R'' = A \times \{b_1, \dots, b_\ell\}$ be a rectangle according to the above claim. Now we apply the result for the index function from Lemma 4. For simplicity, we assume that $j_1 = 1, \dots, j_\ell = \ell$ such that the set of all restrictions of the assignments in A to the variables in $\Pi_1 \cap X$ can be identified in the obvious way with a subset $A_{\text{IND}} \subseteq \{0, 1\}^\ell$ of the same size. Recall that for each assignment in A , the variables in $\Pi_1 \cap Y$ are fixed according to the assignment a chosen above. We regard $R_{\text{IND}} = A_{\text{IND}} \times \{1, \dots, \ell\}$ as a one-way rectangle for the index function IND_ℓ . Define the function h on inputs $u \in \{0, 1\}^\ell$ and $v \in \{1, \dots, \ell\}$ by

$$h(u, v) = \begin{cases} g((u, a), b_v) \oplus a(y_v), & \text{if } y_v \in \Pi_1; \text{ and} \\ g((u, a), b_v) \oplus b_v(y_v), & \text{if } y_v \in \Pi_2; \end{cases}$$

where we regard u as an assignment to $\Pi_1 \cap X$ in the argument of g . Since $b_v \in B_v$ and for each $a' \in A$, $\sigma_X(a', b_v) = \sigma_Y(a', b_v) = v$,

$$\text{MWS}_n((u, a), b_v) = \begin{cases} u(x_v) \oplus a(y_v), & \text{if } y_v \in \Pi_1; \text{ and} \\ u(x_v) \oplus b_v(y_v), & \text{if } y_v \in \Pi_2; \end{cases}$$

and $h(u, v) = u_v = \text{IND}_\ell(u, v)$ if $g((u, a), b_v) = \text{MWS}_n((u, a), b_v)$.

The rectangle R_{IND} is h -uniform since R'' is g -uniform and the values $a(y_v)$ and $b_v(y_v)$, resp., added to the output of g depend only on the second part v of the input. Since g differs from MWS_n on at most a fraction of $\varepsilon' = \varepsilon \cdot (1 + o(1))$ of the inputs of R'' with respect to the uniform distribution, h differs from IND_ℓ on at most an ε' -fraction of R_{IND} with respect to the uniform distribution. By Lemma 4, it follows that $U(R_{\text{IND}}) = 2^{-\Omega(\ell)}$. Furthermore,

$$U(R') = |A|/2^{|\Pi_1|} = 2^{-|\Pi_1 \cap Y|} \cdot |A_{\text{IND}}|/2^\ell = 2^{-|\Pi_1 \cap Y|} \cdot U(R_{\text{IND}})$$

and, by part (ii) of Lemma 5, $\mathcal{D}(R') \leq U(R') \cdot (1 + o(1))$. Finally, $\mathcal{D}(R) \leq p \cdot 2^{|\Pi_1 \cap Y|} \cdot \mathcal{D}(R')$. Putting everything together, we have shown that $\mathcal{D}(R) = p \cdot 2^{-\Omega(\ell)}$. Since $p \leq n + o(n)$ and $\ell \geq n - o(n)$, this bound is of the desired size. \square

The lower bound for quantum OBDDs stated in Theorem 1 follows by standard communication complexity arguments and the properties of MWS_n already used above.

Proof of Theorem 1 – Lower bound for quantum read-once BPs: Let G be a quantum OBDD computing MWS_n with error bounded by a constant ε , $0 \leq \varepsilon < 1/2$. Let $\ell = n - \Theta(p^{2/3+\delta})$ for some constant δ with $0 < \delta < 1/3$. Appropriately cutting the list of variables used as the variable order for G in two parts gives a partition $\Pi = (\Pi_1, \Pi_2)$ of the set of variables $X \cup Y$ that, w.l.o.g., satisfies $|\Pi_1 \cap X| = \ell$ and $|\Pi_1 \cap Y| \leq \ell - 1$. Choose $a \in 2^{\Pi_1 \cap Y}$ somehow arbitrarily and let $i_y = \sigma_{\Pi_1 \cap Y}(a)$. Furthermore, again w.l.o.g., suppose that $\Pi_1 \cap X = \{1, \dots, \ell\}$. For any $i_x \in \{0, \dots, p-1\}$, Lemma 2 yields the existence of assignments $b_{i_x,1}, \dots, b_{i_x,\ell} \in 2^{\Pi_2}$ such that $\sigma_{\Pi_2 \cap X}(b_{i_x,j}) \equiv (j - i_x) \pmod p$ and $\sigma_{\Pi_2 \cap Y}(b_{i_x,j}) \equiv (j - i_y) \pmod p$ for $j = 1, \dots, \ell$.

The given quantum OBDD G can now be used by the two players Alice and Bob in a quantum one-way communication protocol for IND_ℓ as follows. Let $u \in \{0, 1\}^\ell$ and $v \in \{1, \dots, \ell\}$ be the inputs for IND_ℓ . Alice follows the computation in G for the partial input (u, a) , regarding u as an assignment to the variables in $\Pi_1 \cap X$, and sends the reached superposition as well as the partial weighted sum $\sigma_{\Pi_1 \cap X}(u)$ to Bob. Bob finishes the computation of G using the partial input $b_{i_x,v}$ and outputs the XOR of output bit of G with $a(y_v)$, if $y_v \in \Pi_1 \cap Y$, or with $b_{i_x,v}(y_v)$, otherwise. It is easy to see that, analogously to the end of the proof of the lower bound for randomized read-once BPs, this gives a protocol for IND_ℓ that has the same error probability as G . As proved by Klauck [10], the complexity of quantum one-way communication protocols for IND_ℓ with bounded error is lower bounded by $\Omega(\ell)$, which together with the facts that only $O(\log p) = O(\log n)$ bits are required to communicate i_x and that $\ell \geq n - o(n)$ implies $|G| = 2^{\Omega(n)}$, as claimed. \square

Acknowledgment

Thanks to Igor Shparlinski for supplying me with the original source [8] of Lemma 2 and a pointer to the literature about upper bounds for $p(n)$, the smallest prime larger than n .

References

1. S. Aaronson. Limitations of quantum advice and one-way communication. *Theory of Computing*, 1:1–28, 2005. <http://arxiv.org/quant-ph/0402095>.
2. F. M. Ablayev, A. Gainutdinova, and M. Karpinski. On computational power of quantum branching programs. In *Proc. of 13th Fundamentals of Computation Theory (FCT), LNCS 2138*, 59–70. Springer-Verlag, 2001. <http://arxiv.org/quant-ph/0302022>.
3. F. M. Ablayev, C. Moore, and C. Pollett. Quantum and stochastic branching programs of bounded width. In *Proc. of 29th ICALP, LNCS 2380*, 343–354. Springer-Verlag, 2002. <http://arxiv.org/quant-ph/0201139>.
4. A. Ambainis and R. Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. In *Proc. of 39th FOCS*, 332–341, 1998. <http://arxiv.org/quant-ph/9802062>.
5. Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. Exponential separation of quantum and classical one-way communication complexity. In *Proc. of 36th STOC*, 128–137, 2004.
6. E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM J. Comp.*, 26(5):1411–1473, 1997.
7. A. Borodin, A. A. Razborov, and R. Smolensky. On lower bounds for read- k -times branching programs. *Computational Complexity*, 3:1–18, 1993.
8. P. Erdős and H. Heilbronn. On the addition of residue classes mod p . *Acta Arithmetica*, 9:149–159, 1964.
9. G. Hoheisel. Primzahlprobleme in der Analysis. In *Sitzungsberichte Berliner Akademie der Wissenschaften*, 580–588, 1930.
10. H. Klauck. On quantum and probabilistic communication: Las Vegas and one-way protocols. In *Proc. of 32nd STOC*, 644–651, 2000.
11. H. Klauck. Quantum time-space tradeoffs for sorting. In *Proc. of 35th STOC*, 69–76, 2003. <http://arxiv.org/quant-ph/0211174>.
12. H. Klauck. Quantum and classical communication-space tradeoffs from rectangle bounds. In *Proc. of 24th Conf. on the Foundations of Software Technology and Theoretical Computer Science, LNCS 3328*, 384–395. Springer-Verlag, 2004. <http://arxiv.org/quant-ph/0412088>.
13. H. Klauck, R. Špalek, and R. de Wolf. Quantum and classical strong direct product theorems and optimal time-space tradeoffs. In *Proc. of 45th FOCS*, 12–21, 2004. <http://arxiv.org/quant-ph/0402123>.
14. A. Kondacs and J. Watrous. On the power of quantum finite state automata. In *Proc. of 38th FOCS*, 66–75, 1997.
15. I. Kremer, N. Nisan, and D. Ron. On randomized one-round communication complexity. *Computational Complexity*, 8(1):21–49, 1999.
16. E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, Cambridge, 1997.
17. C. Moore and J. P. Crutchfield. Quantum automata and quantum grammars. *Theoretical Computer Science*, 237:275–306, 2000.
18. M. Nakanishi, K. Hamaguchi, and T. Kashiwabara. Ordered quantum branching programs are more powerful than ordered probabilistic branching programs under a bounded-width restriction. In *Proc. of 6th COCOON, LNCS 1858*, 467–476. Springer-Verlag, 2000.
19. M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

20. P. Ribenboim. *The New Book of Prime Number Records*. Springer-Verlag, New York, 3rd edition, 1996.
21. M. Sauerhoff. *Complexity Theoretical Results for Randomized Branching Programs*. PhD thesis, Univ. Dortmund. Shaker, Aachen, 1999.
22. M. Sauerhoff. Randomness versus nondeterminism for read-once and read- k branching programs. In *Proc. of 20th STACS, LNCS 2607*, 307–318. Springer, 2003.
23. M. Sauerhoff and D. Sieling. Quantum branching programs and space-bounded nonuniform quantum complexity. *Theoretical Computer Science*, 334:177–225, 2005. <http://arxiv.org/quant-ph/0403164>.
24. P. Savický and S. Žák. A read-once lower bound and a $(1, +k)$ -hierarchy for branching programs. *Theoretical Computer Science*, 238(1-2):347–362, 2000.
25. R. Špalek. *Space Complexity of Quantum Computation*. Master's thesis, Karl's University Prague, 2002.
26. J. Watrous. Space-bounded quantum complexity. *Journal of Computer and System Sciences*, 59:281–326, 1999.
27. J. Watrous. On the complexity of simulating space-bounded quantum computations. *Computational Complexity*, 12:48–84, 2004.
28. I. Wegener. *Branching Programs and Binary Decision Diagrams—Theory and Applications*. Monographs on Discrete and Applied Mathematics. SIAM, Philadelphia, PA, 2000.
29. A. C. Yao. Probabilistic complexity: towards a unified measure of complexity. In *Proc. of 18th FOCS*, 222–227, 1977.