

# Optimal Direct Sum Results for Deterministic and Randomized Decision Tree Complexity \*

Rahul Jain<sup>†</sup> Hartmut Klauck<sup>‡</sup> Miklos Santha<sup>§</sup>

## Abstract

A *Direct Sum Theorem* holds in a model of computation, when for every problem solving some  $k$  input instances together is  $k$  times as expensive as solving one. We show that Direct Sum Theorems hold in the models of deterministic and randomized decision trees for all relations. We also note that a near optimal Direct Sum Theorem holds for quantum decision trees for boolean functions.

Keywords: Computational Complexity, Direct Sum, Decision Tree Complexity.

## 1 Introduction

One of the goals of complexity theory is to understand the structural properties of different models of computation. A fundamental question that can be asked in every model of computation is how well different computations may be combined. Can we achieve substantial savings when solving the same problem on  $k$  (independent) inputs together? Or is the straightforward approach, namely running the same algorithm  $k$  times, optimal? This question is known as the *direct sum problem*, and has been studied in many different settings and variations.

We say that a *Direct Sum Theorem* holds for a measure of complexity, when solving  $k$  input instances together is roughly as costly as  $k$  times solving one instance according to that measure. Since we are often interested in bounded error computations, we also need to specify how the error on  $k$  input instances relates to the error on one instance. The direct sum question in a narrower sense relates to solving  $k$  instances with constant error, while a *Strong Direct Product Theorem* holds when even using roughly  $k$  times the resources required to solve one instance with constant error, the success probability goes down exponentially in  $k$ . This happens when we solve the  $k$  instances independently, and a Strong Direct Product Theorem states that this is optimal with respect to resources and error. In this paper we only consider the direct sum problem in the narrower sense: we compare solving one instance with constant (resp. no) error to solving  $k$  instances with constant (resp. no) error.

The decision tree model (see [BW02]) is perhaps the simplest model of computation, measuring the number of input positions that need to be accessed in order to compute a function/solve a

---

\*Research supported in part by the French ANR QRAC project under contract ANR-08-EMER-012. Research at the Centre for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation.

<sup>†</sup>Centre for Quantum Technologies and Department of Computer Science, National University of Singapore. Email: rahul@comp.nus.edu.sg

<sup>‡</sup>Centre for Quantum Technologies (NUS) and School of Physical and Mathematical Sciences, Nanyang Technological University. Email: hklauck@gmail.com

<sup>§</sup>CNRS - LRI, Université Paris-Sud, Orsay, France and Centre for Quantum Technologies, National University of Singapore. Email: santha@lri.fr

relation. Still many questions about this model remain open. In this paper we show that the direct sum property holds for deterministic and randomized decision trees.

Previously, a Strong Direct Product Theorem for decision trees was established by Nisan et al. [NRS94]. However, their result does not imply a Direct Sum Theorem in our sense, because it is only shown in a weaker setting. Instead of analyzing a single algorithm that has access to all  $k$  inputs and produces all  $k$  outputs, Nisan et al. consider a setting where  $k$  algorithms (that can access all inputs), each making at most  $d$  queries, compute one of the  $k$  outputs each, where  $d$  is the query complexity of computing one instance (with bounded error). Hence this does not establish a Direct Sum Theorem in the above sense.

Previous papers [NRS94, BN95] have dismissed the direct sum problem for decision trees as either very simple, or uninteresting. To quote [NRS94]: “While it is an easy exercise to see that ‘direct-sum holds for decision tree depth, the other two problems (direct product and help bits) are more difficult.” The paper does not make it clear, what kind of decision tree is meant (the setting considered there is distributional complexity). In the distributional setting a general counterexample by Shaltiel [Sha01] makes it clear that some very tight direct sum statements are not even true for the model where there is one decision tree that has to solve all  $k$  input instances together.

Ben-Asher and Newman claim in [BN95]: “In the standard decision tree model the question is quite uninteresting as queries do not involve variables of more than one of the problem instances at a time.” This does not seem to be a valid assessment of the problem, because with the same argument the strong direct product question for decision trees could be dismissed, which has only very recently been resolved (in the setting where one algorithm makes all outputs) by Drucker [D10].

The following partial results are known about the direct sum question for randomized and quantum decision trees. Klauck et al. [KSW07] show a Strong Direct Product Theorem (and hence a Direct Sum Theorem) for the randomized complexity of all functions, for which the block-sensitivity bound is tight, as well as for the quantum complexity of all functions, for which the square root of block-sensitivity is a tight bound. Ambainis et al. [ASW09] show a Strong Direct Product Theorem for the quantum complexity of all symmetric functions. Spalek [S01] shows a Strong Direct Product Theorem for all functions for which the “multiplicative quantum adversary bound” is tight.

We give proofs of Direct Sum Theorems for the case of deterministic and randomized decision trees. In the deterministic case the main problem is to construct a more efficient decision tree for one instance from a given tree for two instances. In the randomized case the proof is along the lines of some proofs of Direct Sum Theorems in communication complexity e.g. [JRS05]. After a preprint of the present paper was circulated, Drucker [D10] showed various Strong Direct Product Theorems for randomized query complexity. His Theorem 12.3 implies a Direct Sum Theorem for the randomized query complexity of all relations with somewhat worse parameters (and a more complicated proof) than ours.

One may ask if a similar direct sum result is true for quantum decision trees, also known as quantum query algorithms. A somewhat weaker statement than for the randomized and deterministic cases can be derived from recent results by Reichardt. In [R09, R10] he shows that the general quantum adversary bound is tight for the quantum decision tree complexity of every boolean function (his Theorem 1.3 in [R10]). He also shows that the general adversary bound has a direct sum behavior for boolean functions (see Theorem 7.2 in the long version of [R09]. Note that one has to choose a good “connection” function  $f$  like the parity function). The direct sum for the general adversary bound has also been shown previously in Ambainis, Childs, Le Gall and Tani [ACGT09]. Hence we can conclude that in the quantum case, at least for boolean functions, a Direct Sum Theorem also holds.

## 2 Preliminaries

A *deterministic decision tree* on  $m$  variables is a rooted binary tree  $T$  whose internal vertices are labeled by the boolean variables  $x_1, \dots, x_m$ , and whose leaves are labeled by the output values from a set  $\mathcal{Y}$ . For every vertex  $v$  in  $T$ , we denote by  $v_0$  (respectively  $v_1$ ) the left son (respectively the right son) of  $v$ , and by  $T(v)$  the subtree of  $T$  rooted at  $v$ . We set  $T_b = T(v_b)$ , for  $b \in \{0, 1\}$ , where  $v$  is the root of  $T$ . The depth  $d_T(v)$  of vertex  $v$  in tree  $T$ , is defined recursively: it is 0 if  $v$  is a leaf, otherwise  $d_T(v) = \max\{d_T(v_0), d_T(v_1)\} + 1$ . The depth  $d(T)$  of  $T$  is simply the depth of its root. Every tree naturally computes a function  $f_T$  on  $m$  variables, whose value at an assignment  $x = (x_1, \dots, x_m) \in \{0, 1\}^m$  is defined recursively as follows: If the root of  $T$  is a leaf, then  $f_T(x)$  is the value of its label. Otherwise, if  $x_i$  is the label of the root and  $x_i = b$ , then  $f_T(x) = f_{T_b}(x)$ .

Clearly, more than one decision tree computes the same function  $f$ . The *deterministic decision tree complexity* of  $f$ , denoted  $D(f)$ , is the minimal depth of a decision tree  $T$  such that  $f_T = f$ .

The above definitions naturally extend to trees whose leaves are labeled by elements of  $\mathcal{Y}^k$ , for some positive integer  $k$ . We call these trees *k-output deterministic decision trees*, they compute *k-output* functions whose range is by definition  $\mathcal{Y}^k$ . We will use the notation  $f = (f^{(1)}, \dots, f^{(k)})$  for *k-output* functions, where  $f^{(i)}$  is the function computing the  $i$ th output of  $f$ . In particular, we are interested here in the case when  $m = kn$  and the functions do not share common input variables. More precisely, let  $f : \{0, 1\}^{kn} \rightarrow \mathcal{Y}^k$  be a *k-output* function whose input variables are  $x_{1,1}, \dots, x_{1,n}, \dots, x_{k,1}, \dots, x_{k,n}$ . We set  $\bar{x}_i = (x_{i,1}, \dots, x_{i,n})$ , and say that  $f$  is *k-independent* if the value of  $f^{(i)}$  depends only on  $\bar{x}_i$ .

One can also extend the definition of deterministic decision trees and *k-independence* to relations  $f \subseteq \{0, 1\}^m \times \mathcal{Y}$  instead of functions in a straightforward way (decision trees are required to find an output  $y$  for each input  $x \in \{0, 1\}^m$  so that  $(x, y) \in f$ ).

In particular, for a relation  $f \subseteq \{0, 1\}^m \times \mathcal{Y}^k$ , the relation  $f^{(i)} \subseteq \{0, 1\}^m \times \mathcal{Y}$  consists of all  $(x, y)$ , such that  $(x, y_1, \dots, y_{i-1}, y, y_{i+1}, \dots, y_k) \in f$  for some  $y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_k$ .

Note that for inputs  $x$  for which there is no  $y$  with  $(x, y) \in f$  no requirement on the output is made, and hence we can assume that all relations are total without loss of generality. Since for each input only one output can be produced, each deterministic tree automatically computes a function that is consistent with the relation in question.

For a relation  $f \subseteq \{0, 1\}^n \times \mathcal{Y}$ , we define the *kth tensor power* of  $f$  as the relation  $f^{\otimes k} \subseteq \{0, 1\}^{kn} \times \mathcal{Y}^k$  by  $f^{\otimes k} = \{((\bar{x}_1, \dots, \bar{x}_k), (y_1, \dots, y_k)) : \forall i : (\bar{x}_i, y_i) \in f\}$ . Note that  $f^{\otimes k}$  is *k-independent*.

A *randomized decision tree* on  $m$  variables with error  $\epsilon$  is a probability distribution on deterministic  $m$ -variable decision trees, such that for each input  $x$  a correct output is computed with probability  $1 - \epsilon$  by a tree chosen from the distribution. If not mentioned otherwise  $\epsilon = 1/3$ . Note that for *k-output* relations  $f$  an output  $(y_1, \dots, y_k)$  is considered erroneous, if  $(\bar{x}_i, y_i) \notin f^{(i)}$  for some  $i$ , i.e., all  $k$  outputs are required to be correct simultaneously. The *depth* (or maximum number of queries) of a randomized decision tree is the maximal depth of any tree occurring in the distribution with positive probability.

For  $\epsilon > 0$ , let  $R_\epsilon(f)$  denote the  $\epsilon$ -error randomized query complexity of  $f$ , which is the maximum number of queries made by the minimal depth randomized decision tree with error at most  $\epsilon$ . Let  $\mu$  be a distribution on  $\{0, 1\}^n$ . Let  $R_\epsilon^\mu(f)$  represent the  $\epsilon$ -error distributional query complexity of  $f$  under  $\mu$ , which is the maximum number of queries made by the minimal depth randomized decision tree that has *average* error at most  $\epsilon$  under  $\mu$  (note that such a tree can be assumed to be deterministic w.l.o.g., but sometimes it is simpler to give a randomized tree).

We also consider randomized decision trees that make no errors. In this case we are interested in the expected number of queries of a randomized tree, i.e., the maximum (over inputs) of the expected (over the distribution on trees) length of the path that the input determines in the tree. Let  $R_0(f)$

denote the 0-error randomized query complexity of  $f$  which is the minimum over all randomized decision trees  $T$  that make no error in computing  $f$  of the maximum over inputs  $x$  of the expected number of queries that  $T$  makes on  $x$ . Let  $R_0^\mu(f)$  represent the 0-error distributional query complexity of  $f$  under  $\mu$ , which is the expected number (now over inputs as well) of queries made by the best randomized (w.l.o.g deterministic) decision tree with no error.

We have the following fact from [Y83].

**Fact 1 (Yao's Principle)** *Let  $\epsilon > 0$ , then  $R_\epsilon(f) = \max_\mu R_\epsilon^\mu(f)$ . Also  $R_0(f) = \max_\mu R_0^\mu(f)$ .*

### 3 Direct Sum for Deterministic Complexity

Let  $f \subseteq \{0, 1\}^{kn} \times \mathcal{Y}^k$  be a  $k$ -output relation. Obviously  $D(f) \leq \sum_{i=1}^k D_{f^{(i)}}$  since the values  $f^{(i)}$  can be evaluated sequentially. We prove that for  $k$ -independent relations this is in fact the least expensive way to evaluate  $f$ , that is the opposite inequality also holds.

**Theorem 1 (Deterministic Direct Sum)** *For every  $k$ -independent relation  $f \subseteq \{0, 1\}^{kn} \times \mathcal{Y}^k$ , we have  $D(f) \geq \sum_{i=1}^k D(f^{(i)})$ .*

**Proof** Let  $T$  be a  $k$ -output deterministic decision tree on variables  $\{x_{1,1}, \dots, x_{k,n}\}$ . For  $i = 1, \dots, k$ , we refer to  $\{x_{i,1}, \dots, x_{i,n}\}$  as the  $i$ th group of variables. For every vertex  $v$  of  $T$ , we define recursively  $k$  single output decision trees  $T_1(v), \dots, T_k(v)$ , where the vertices of  $T_i(v)$  are labeled by the variables from the  $i$ th group. If  $v$  is a leaf with label  $(b_1, \dots, b_k)$ , then  $T_i(v)$  is a single node tree (a leaf), with label  $b_i$ . Otherwise, let  $v$  be an internal node and let's suppose that its label is from the  $j$ th group of variables. The root of  $T_j(v)$  is by definition  $v$  with the same label as in  $T$ , its left subtree is  $T_j(v_0)$  and its right subtree is  $T_j(v_1)$ . For all  $i \neq j$ , the tree  $T_i(v)$  is defined as the shallower (smaller depth) tree between  $T_i(v_0)$  and  $T_i(v_1)$ .

**Claim 1** *For every vertex  $v$  of  $T$ , we have  $\sum_{i=1}^k d(T_i(v)) \leq d_T(v)$ .*

**Proof** The proof is by induction on the depth of  $v$ , and the statement is obviously true when  $v$  is a leaf. We assume that the label of  $v$  is from the  $j$ th group. Let  $b \in \{0, 1\}$  such that  $d(T_j(v)) = d(T_j(v_b)) + 1$ . By definition, for all  $i \neq j$ , we have  $d(T_i(v)) = \min\{d(T_i(v_0)), d(T_i(v_1))\}$ , and therefore  $d(T_i(v)) \leq d(T_i(v_b))$ . Thus

$$\begin{aligned} \sum_{i=1}^k d(T_i(v)) &\leq d(T_j(v_b)) + 1 + \sum_{i \neq j} d(T_i(v_b)) \\ &\leq d_T(v_b) + 1 \\ &\leq d_T(v), \end{aligned}$$

where the second inequality follows from the inductive hypothesis, and the third one from the definition of the depth.  $\square$

We say that  $T$  is *parsimonious* if no variable appears twice on the same root-leaf simple path.

**Claim 2** *Let  $T$  be parsimonious. Then for every vertex  $v$  in  $T$ , for every  $1 \leq i \leq k$ , and for every assignment  $\bar{x}_i \in \{0, 1\}^n$  for the variables in the  $i$ th group, there exists, for all  $j \neq i$ , an assignment  $\bar{x}_j \in \{0, 1\}^n$  for the variables in the  $j$ th group such that*

$$f_{T_i(v)}(\bar{x}_i) = f_{T(v)}^{(i)}(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_k).$$

The intuition behind the inductive proof is that the inductive hypothesis for  $v_b, i$  and  $\bar{x}_i$ , where  $b \in \{0, 1\}$  is such that  $f_{T_i(v)}(\bar{x}_i) = f_{T_i(v_b)}(\bar{x}_i)$ , can be directly used except maybe for the label of  $v$  which should be consistent with this equality. This potential change will not create a problem since  $T$  is parsimonious, and therefore the same label will not be encountered again.

**Proof** The proof is again by induction on the depth of  $v$ . Fix  $1 \leq i \leq k$ . If  $v$  is a leaf, we can choose for every  $\bar{x}_i \in \{0, 1\}^n$  an arbitrary  $\bar{x}_j \in \{0, 1\}^n$ , for  $j \neq i$ . Otherwise, we distinguish two cases, according to the label of  $v$ .

Case 1: The label of  $v$  is  $x_{i,p}$  from the  $i$ th group of variables, for some  $1 \leq p \leq n$ . Let  $\bar{x}_i \in \{0, 1\}^n$  be an assignment for the variables in the  $i$ th group, and let  $x_{i,p} = b$ . By the inductive hypothesis there exists  $\bar{x}'_j$ , for  $j \neq i$ , such that

$$f_{T_i(v_b)}(\bar{x}_i) = f_{T(v_b)}^{(i)}(\bar{x}'_1, \dots, \bar{x}_i, \dots, \bar{x}'_k).$$

We set  $\bar{x}_j = \bar{x}'_j$ , for  $j \neq i$ . Then we have

$$\begin{aligned} f_{T_i(v)}(\bar{x}_i) &= f_{T_i(v_b)}(\bar{x}_i) \\ &= f_{T(v_b)}^{(i)}(\bar{x}'_1, \dots, \bar{x}_i, \dots, \bar{x}'_k) \\ &= f_{T(v)}^{(i)}(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_k). \end{aligned}$$

The first equality follows from the definition of  $f_{T_i(v_b)}(\bar{x}_i)$  since  $x_{i,p} = b$ . The third equality also holds because by definition  $f_{T(v)}(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_k) = f_{T(v_b)}(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_k)$ .

Case 2: The label of  $v$  is  $x_{j,p}$  from the  $j$ th set of variables for some  $j \neq i$  and  $1 \leq p \leq n$ . Let  $b$  be such that  $T_i(v) = T_i(v_b)$ . Then again by the inductive hypothesis, for every  $\bar{x}_i \in \{0, 1\}^n$ , there exists  $\bar{x}'_j$ , for  $j \neq i$ , that satisfy

$$f_{T_i(v_b)}(\bar{x}_i) = f_{T(v_b)}^{(i)}(\bar{x}'_1, \dots, \bar{x}_i, \dots, \bar{x}'_k).$$

We define  $x_{l,q}$  for  $l \neq i$  and  $q = 1, \dots, n$  by

$$x_{l,q} = \begin{cases} b & \text{if } (l, q) = (j, p), \\ x'_{l,q} & \text{otherwise.} \end{cases}$$

Then, similarly to Case 1, we have the following series of equalities:

$$\begin{aligned} f_{T_i(v)}(\bar{x}_i) &= f_{T_i(v_b)}(\bar{x}_i) \\ &= f_{T(v_b)}^{(i)}(\bar{x}'_1, \dots, \bar{x}_i, \dots, \bar{x}'_k) \\ &= f_{T(v)}^{(i)}(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_k). \end{aligned}$$

The first equality is true because  $T_i(v) = T_i(v_b)$ . The path followed on input  $(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_k)$  in  $T(v)$  goes from  $v$  to  $v_b$  since  $x_{j,p} = b$ , and then it is identical to the path followed on input  $(\bar{x}'_1, \dots, \bar{x}_i, \dots, \bar{x}'_k)$  in  $T(v_b)$  because  $T$  is parsimonious. Therefore  $f_{T(v_b)}(\bar{x}'_1, \dots, \bar{x}_i, \dots, \bar{x}'_k) = f_{T(v)}(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_k)$ , and the last equality also holds.  $\square$

We now prove Theorem 1 by contradiction. Let us suppose that  $D(f) < \sum_{i=1}^n D(f^{(i)})$ . Let  $T$  be a deterministic decision tree of depth  $D(f)$  which computes a function  $\tilde{f}$  that is consistent with the relation  $f$ . Since  $T$  is a minimal depth decision tree computing  $\tilde{f}$ , we can suppose without loss of generality that  $T$  is parsimonious. Let  $r$  be the root of  $T$ , then  $d(r) = D(f)$ . For  $i = 1, \dots, k$ , let  $T_i = T_i(r)$ . By Claim 2 and  $k$ -independence,  $T_i$  computes an  $\tilde{f}^{(i)}$ , which is consistent with  $f^{(i)}$ , and therefore  $D(f^{(i)}) \leq d(T_i)$ . Thus  $d(r) < \sum_{i=1}^k d(T_i)$ , contradicting Claim 1.  $\square$

**Corollary 1** For every relation  $f \subseteq \{0, 1\}^n \times \mathcal{Y}$  and for every integer  $k$ , we have  $D(f^{\otimes k}) = k \cdot D(f)$ .

## 4 Direct Sum for Randomized Query Complexity

**Theorem 2 (Randomized Direct Sum)** *Let  $f \subseteq \{0, 1\}^n \times \mathcal{Y}$  be a relation. Let  $k$  be a positive integer and let  $\epsilon, \delta > 0$ . Then*

1.  $R_\epsilon(f^{\otimes k}) \geq \delta \cdot k \cdot R_{\epsilon'}(f)$ , where  $\epsilon' = \epsilon + \delta$ .
2.  $R_0(f^{\otimes k}) = k \cdot R_0(f)$ .

### Proof

1. Let  $c = R_\epsilon(f^{\otimes k})$ . Let  $\mathcal{P}$  be a randomized protocol for  $f^{\otimes k}$  with  $c$  queries and worst case error at most  $\epsilon$ . Let  $\mu$  be a distribution on  $\{0, 1\}^n$ . Let  $\mu^{\otimes k}$  represent the distribution on  $\{0, 1\}^{kn}$  which consists of  $k$  independent copies of  $\mu$ . Now let us consider the situation when we provide inputs to  $\mathcal{P}$  distributed according to  $\mu^{\otimes k}$ . In such a situation we can fix the random coins of  $\mathcal{P}$  in a suitable manner to get another protocol  $\mathcal{P}_1$  such that  $\mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} [e(x_1 \dots x_k)] \leq \epsilon$ , where  $e(x_1 \dots x_k)$  represents the error made by  $\mathcal{P}_1$ , which is now a deterministic protocol, on input  $(x_1 \dots x_k)$  (each  $x_i \in \{0, 1\}^n$  represents the input for the  $i$ th instance of  $f$ ; for notional convenience we use  $x_i$  here instead of  $\bar{x}_i$  as used in the previous section). Let  $q(x_1 \dots x_k)$  represent the number of queries made by  $\mathcal{P}_1$  on input  $(x_1 \dots x_k)$ . For each  $1 \leq i \leq k$ , let  $q_i(x_1 \dots x_k)$  represent the number of queries made by  $\mathcal{P}_1$  on  $x_i$  on input  $(x_1 \dots x_k)$ . Since  $q(x_1 \dots x_k) = \sum_{i=1}^k q_i(x_1 \dots x_k)$ , we have,

$$\begin{aligned} c &\geq \mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} [q(x_1 \dots x_k)] \\ &= \mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} \left[ \sum_{i=1}^k q_i(x_1 \dots x_k) \right] \\ &= \sum_{i=1}^k \mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} [q_i(x_1 \dots x_k)] . \end{aligned}$$

Therefore there exists  $1 \leq j \leq k$  such that  $\mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} [q_j(x_1 \dots x_k)] \leq \frac{c}{k}$ . Now  $\mathcal{P}_1$  naturally gives rise to a protocol  $\mathcal{P}_2$  for  $f$  which imitates  $\mathcal{P}_1$  as follows. In  $\mathcal{P}_2$ , the actual input  $x \in \{0, 1\}^n$  is considered to be the input  $x_j$  of  $\mathcal{P}_1$  and the inputs  $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k$  of  $\mathcal{P}_1$  are generated according to  $\mu^{k-1}$ , using random coins. The protocol  $\mathcal{P}_2$  then imitates  $\mathcal{P}_1$  and outputs the answer corresponding to the  $j$ th coordinate. The expected number of queries of  $\mathcal{P}_2$  under  $\mu$  is at most  $\frac{c}{k}$  and expected error under  $\mu$  is at most  $\epsilon$ . Now let us consider a protocol  $\mathcal{P}_3$  which proceeds exactly as  $\mathcal{P}_2$  but terminates whenever the number of queries exceeds  $\frac{c}{\delta k}$ . Using Markov's inequality we can argue that the expected error of  $\mathcal{P}_3$  under  $\mu$  is at most  $\epsilon' = \epsilon + \delta$  and of course the maximum queries made by  $\mathcal{P}_3$  is at most  $\frac{c}{\delta k}$ . Hence by definition  $R_{\epsilon'}^\mu(f) \leq \frac{c}{\delta k}$ . Since this is true for every distribution  $\mu$  on  $\{0, 1\}^n$ , we get from Yao's Principle, Fact 1, the desired result as follows.

$$R_{\epsilon'}(f) = \max_{\mu} R_{\epsilon'}^\mu(f) \leq \frac{c}{\delta k} = \frac{R_\epsilon(f^{\otimes k})}{\delta k} .$$

2. The inequality  $R_0(f^{\otimes k}) \leq k \cdot R_0(f)$  is easily seen, indeed  $k$  independent executions of the best protocol for  $f$  give a protocol for  $f^{\otimes k}$ . Consider now the reverse inequality. Let  $c = R_0(f^{\otimes k})$ . Let  $\mathcal{P}$  be a randomized protocol for  $f^{\otimes k}$  with expected queries on any input at most  $c$  and no error. Let  $\mu$  be a distribution on  $\{0, 1\}^n$ . Let  $\mu^{\otimes k}$  be as before. Now let us consider the situation when we provide input to  $\mathcal{P}$  distributed according to  $\mu^{\otimes k}$ . In such a situation we can

fix the random coins of  $\mathcal{P}$  in a suitable manner to get a deterministic 0-error protocol  $\mathcal{P}_1$  such that  $\mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} [q(x_1 \dots x_k)] \leq c$ , where  $q(x_1 \dots x_k)$  represents the number of queries made by  $\mathcal{P}_1$ , on input  $(x_1 \dots x_k)$  (each  $x_i \in \{0, 1\}^n$  represents the input for the  $i$ th instance of  $f$ ). For each  $1 \leq i \leq k$ , let  $q_i(x_1 \dots x_k)$  represent the number of queries made by  $\mathcal{P}_1$  on  $x_i$  on input  $(x_1 \dots x_k)$ . Since  $q(x_1 \dots x_k) = \sum_{i=1}^k q_i(x_1 \dots x_k)$ , we have,

$$\begin{aligned} c &\geq \mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} [q(x_1 \dots x_k)] \\ &= \mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} \left[ \sum_{i=1}^k q_i(x_1 \dots x_k) \right] \\ &= \sum_{i=1}^k \mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} [q_i(x_1 \dots x_k)] . \end{aligned}$$

Therefore there exists  $1 \leq j \leq k$  such that  $\mathbb{E}_{(x_1 \dots x_k) \leftarrow \mu^{\otimes k}} [q_j(x_1 \dots x_k)] \leq \frac{c}{k}$ . Without loss of generality assume  $j = 1$ . Using this we can argue that there exist  $x'_2 \dots x'_k \in \{0, 1\}^{kn-n}$  such that  $\mathbb{E}_{x_1 \leftarrow \mu} [q_1(x_1 x'_2 \dots x'_k)] \leq \frac{c}{k}$ . Therefore fixing  $x'_2 \dots x'_k$  in  $\mathcal{P}_1$  naturally gives rise to a protocol  $\mathcal{P}_2$  for  $f$  with expected number of queries under  $\mu$  being at most  $\frac{c}{k}$  and zero error. Hence by definition  $R_0^\mu(f) \leq \frac{c}{k}$ . Since this is true for every distribution  $\mu$  on  $\{0, 1\}^n$ , we get from Yao's Principle, Fact 1, the desired result as follows.

$$R_0(f) = \max_{\mu} R_0^\mu(f) \leq \frac{c}{k} = \frac{R_0(f^{\otimes k})}{k} .$$

□

## 5 Open Problems

We proved direct sum theorems for deterministic and randomized query complexity. Note that it is also very easy to establish the direct sum property for nondeterministic query complexity (also known as certificate complexity  $C(f)$ , see [BW02]):  $C(f^{\otimes k}) = k \cdot C(f)$ . However, several related open problems remain:

1. The direct sum theorem in the randomized case loses a factor of  $\delta$  in the lower bound, as well as an additive  $\delta$  in the error bound. While at least the factor in the lower bound is unavoidable in the setting of distributional complexity according to a result by Shaltiel [Sha01], this might not be necessary in the worst case complexity setting.
2. In the quantum case the result following from Reichardt's work holds only for boolean functions. Can a tight result be established, even for all relations?
3. In the randomized setting the obvious way to compute  $k$  instances with constant error requires a success amplification step that costs a factor of  $\log k$ . This is known to be unnecessary in the quantum case [BNRW07], but unknown to be tight in the randomized case.

## 6 Acknowledgements

The authors would like to thank the anonymous referees for helpful comments.

## References

- [ACGT09] A. Ambainis, A. M. Childs, F. Le Gall, S. Tani. The quantum query complexity of certification. Preprint, arXiv:0903.1291.
- [ASW09] A. Ambainis, R. palek, and R. de Wolf. A New Quantum Lower Bound Method, with Applications to Direct Product Theorems and Time-Space Tradeoffs. *Algorithmica*, 55(3):422-461, 2009.
- [BN95] Y. Ben-Asher and I. Newman. Decision Trees with AND, OR Queries. In *Proceedings of 10th IEEE Conference Structure in Complexity Theory*, pp. 74–81, 1995.
- [BW02] H. Buhrman and R. de Wolf. Complexity measures and decision tree complexity: A survey. *Theoretical Computer Science*, 288(1):21–43, 2002.
- [BNRW07] H. Buhrman, I. Newman, H. Rhrig, and R. de Wolf. Robust Quantum Algorithms and Polynomials. *Theory of Computing Systems* 40(4):379-395, 2007.
- [D10] Andrew Drucker. Improved Direct Product Theorems for Randomized Query Complexity. Preprint: arXiv:1005.0644, 2010.
- [JRS05] R. Jain, J. Radhakrishnan and P. Sen. Prior entanglement, message compression and privacy in quantum communication. In *Proceedings of 20th IEEE Conference on Computational Complexity*, pp. 285–296, 2005.
- [KSW07] H. Klauck, R. palek, and R. de Wolf. Quantum and Classical Strong Direct Product Theorems and Optimal Time-Space Tradeoffs. In *SIAM Journal on Computing*, 36(5):1472-1493, 2007.
- [NRS94] N. Nisan, S. Rudich and M.E. Saks. Products and Help Bits in Decision Trees. In *Proceedings of 35th IEEE Symposium on Foundations of Computer Science*, pp. 318–329, 1994.
- [R09] B. W. Reichardt. Span programs and quantum query complexity: The general adversary bound is nearly tight for every boolean function. In *Proceedings of 50th IEEE Symposium on Foundations of Computer Science*, pp. 544–551, 2009. Long version under quant-ph/0904.2759
- [R10] B. W. Reichardt. Reflections for quantum query algorithms. Preprint: arXiv:1005.1601, 2010.
- [Sha01] R. Shaltiel. Towards proving strong direct product theorems. In *Proceedings of 16th IEEE Conference on Computational Complexity*, pp. 107–119, 2001.
- [S01] Robert Spalek. The Multiplicative Quantum Adversary. In *Proceedings of 23rd IEEE Conference on Computational Complexity*, pp.237–248, 2008
- [Y83] A.C.C. Yao. Lower Bounds by Probabilistic Arguments. In *Proceedings of 24th IEEE Symposium on Foundations of Computer Science*, pp. 420–428, 1983.