SIAM J. COMPUT.
Vol. 32, No. 3, pp. 681–699

NONDETERMINISTIC QUANTUM QUERY AND COMMUNICATION COMPLEXITIES*

RONALD DE WOLF†

Abstract. We study nondeterministic quantum algorithms for Boolean functions f. Such algorithms have positive acceptance probability on input x iff $f(x) = 1$. In the setting of query complexity, we show that the nondeterministic quantum complexity of a Boolean function is equal to its "nondeterministic polynomial" degree. We also prove a quantum-vs.-classical gap of 1 vs. n for nondeterministic query complexity for a total function. In the setting of communication complexity, we show that the nondeterministic quantum complexity of a two-party function is equal to the logarithm of the rank of a nondeterministic version of the communication matrix. This implies that the quantum communication complexities of the equality and disjointness functions are $n + 1$ if we do not allow any error probability. We also exhibit a total function in which the nondeterministic quantum communication complexity is exponentially smaller than its classical counterpart.

Key words. quantum computing, query complexity, communication complexity, nondeterminism

AMS subject classification. 68Q10

PII. S0097539702407345

1. Introduction.

1.1. Motivation.

In classical computing, *nondeterministic* computation has a prominent place in many different models and for many good reasons. For example, in Turing machine complexity, the study of nondeterminism leads naturally to the class of NP-complete problems, which contains some of the most important and practically relevant computer science problems—as well as some of the hardest theoretical open questions. In fields like query complexity and communication complexity, there is a tight relation between deterministic complexity and nondeterministic complexity, but it is often much easier to analyze upper and lower bounds for the latter than for the former.

Suppose we want to compute a Boolean function f in some algorithmic setting, such as that of Turing machines, decision trees, or communication protocols. Consider the following two ways of viewing a nondeterministic algorithm. The first and most common way is to think of it as a "certificate verifier": a deterministic algorithm A that receives, apart from the input x, a "certificate" y whose validity it needs to verify. For all inputs x, if $f(x) = 1$, then there is a certificate y such that $A(x, y) = 1$; if $f(x) = 0$, then $A(x, y) = 0$ for all y. Second, we may view A as a *randomized* algorithm whose acceptance probability is positive if $f(x) = 1$ and whose acceptance probability is zero if $f(x) = 0$. It is easy to see that these two views are equivalent in the classical case. To turn an algorithm A of the first kind into one of the second kind, we can just guess a certificate y at random and output $A(x, y)$. This will have positive acceptance probability iff $f(x) = 1$. For the other direction, we can consider

*Received by the editors May 8, 2002; accepted for publication (in revised form) December 12, 2002; published electronically April 17, 2003. This paper combines results from the conference papers [50, 29] with some new results.

http://www.siam.org/journals/sicomp/32-3/40734.html

†CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands (rdewolf@cwi.nl). This author was partially supported by the EU fifth framework project QAIP, IST–1999–11234. Part of this paper was written when the author was a postdoc at UC Berkeley, supported by Talent grant S 62–565 from the Netherlands Organization for Scientific Research (NWO).

the sequence of coin flips used by an algorithm of the second kind as a certificate. Clearly, there will be a certificate leading to output 1 iff $f(x) = 1$, which gives us an algorithm of the first kind.

Both views may be generalized to the quantum case, yielding *three* potential definitions of nondeterministic quantum algorithms, possibly nonequivalent. The quantum algorithm may be required to output the right answer $f(x)$ when given an appropriate certificate, which we can take to be either quantum or classical. Or, third, the quantum algorithm may be required to have positive acceptance probability iff $f(x) = 1$. An example is given by two alternative definitions of quantum nondeterminism in the case of quantum Turing machine complexity. Kitaev defines the class "bounded-error quantum-NP" (BNQP) as the set of languages accepted by polynomial-time bounded-error quantum algorithms that are given a polynomial-size quantum certificate (e.g., [32, 31] and [30, Chapter 14]). On the other hand, Adleman, Demarrais, and Huang [2] and Fenner et al. [24] define quantum-NP as the set of languages L for which there is a polynomial-time quantum algorithm whose acceptance probability is positive iff $x \in L$. This quantum class was shown to be equal to the classical counting class co-$\mathrm{C_{=}P}$ [24, 52] using tools from Fortnow and Rogers [25].

In this paper, we adopt the latter view: a nondeterministic quantum algorithm for f is defined to be a quantum algorithm that outputs 1 with positive probability if $f(x) = 1$ and that always outputs 0 if $f(x) = 0$. This definition contrasts with the more traditional view of classical determinism as "certificate verification." The motivation for our choice of definition of quantum nondeterminism is twofold. First, in the appendix, we show that this definition is strictly more powerful than the other two possible definitions in the sense of being able to simulate the other definitions efficiently, while the reverse is not true. Second, it turns out that this definition lends itself to very crisp results. Rather than in the quantum Turing machine setting of Kitaev, Adleman, etc., we study the complexity of nondeterministic algorithms in the query complexity and communication complexity settings. Our main results are exact characterizations of these nondeterministic quantum complexities in algebraic terms and large gaps between quantum and classical complexities in both settings. Our algebraic characterizations can be extended to nontotal functions in the obvious way, but we will stick to total functions in our presentation.

1.2. Query complexity. We first consider the model of query complexity, also known as decision tree complexity or black box complexity. Here the goal is to compute some function $f : \{0,1\}^n \to \{0,1\}$, making as few queries to input bits as possible. Most existing quantum algorithms can naturally be expressed in this model and achieve provable speed-ups over the best classical algorithms. Examples can be found, e.g., in [22, 48, 26, 12, 13, 14] and also include the order-finding problem on which Shor's celebrated factoring algorithm is based [47].

Let $D(f)$ and $Q_E(f)$ denote the query complexities of optimal deterministic and quantum algorithms that compute f exactly. Let $deg(f)$ denote the minimal degree among all multilinear polynomials that represent f. (A polynomial p represents f if $f(x) = p(x)$ for all $x \in \{0,1\}^n$.) The following relations are known. The first inequality is due to Beals et al. [6], the second inequality is obvious, and the last is due to Nisan and Smolensky—unpublished, but described in the survey paper [20].

$$\frac{deg(f)}{2} \leq Q_E(f) \leq D(f) \leq O(deg(f)^4).$$

Thus $deg(f)$, $Q_E(f)$, and $D(f)$ are polynomially related for all total f. (The situation

is very different for partial f [22, 48, 47, 7].) Nisan and Szegedy [42] exhibit a function with a large gap between $D(f) = n$ and $deg(f) = n^{0.6\cdots}$, but no function is known where $Q_E(f)$ is significantly larger than $deg(f)$, and it may in fact be true that $Q_E(f)$ and $deg(f)$ are linearly related. In section 2, we show that the *nondeterministic* versions of $Q_E(f)$ and $deg(f)$ are in fact *equal*:

$$NQ(f) = ndeg(f).$$

Here $NQ(f)$ denotes the query complexity of an optimal nondeterministic quantum algorithm for f, which has nonzero acceptance probability iff $f(x) = 1$. The nondeterministic degree $ndeg(f)$ is the minimal degree of a so-called *nondeterministic* polynomial for f, which is required to be nonzero iff $f(x) = 1$. A note on terminology: the name "nondeterministic polynomial" is based only on analogy with the acceptance probability of a nondeterministic algorithm. This name is less than ideal, since such polynomials have little to do with the traditional view of nondeterminism as certificate verification. Nevertheless, we use this name because any alternatives that we could think of were worse (too verbose or confusing).

Apart from the algebraic characterization of the nondeterministic quantum query complexity $NQ(f)$, we also show that $NQ(f)$ may be much smaller than its classical analogue $N(f)$: we exhibit an f where $NQ(f) = 1$ and $N(f) = n$, which is the biggest possible gap allowed by this model. Accordingly, while the case of exact (or, for that matter, bounded-error) computation allows at most polynomial quantum-classical query complexity gaps for total functions, the nondeterministic case allows *unbounded* gaps.

1.3. Communication complexity. In the case of communication complexity, the goal is for two distributed parties, Alice and Bob, to compute some function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. Alice receives an $x \in \{0,1\}^n$, and Bob receives a $y \in \{0,1\}^n$, and they want to compute $f(x,y)$, exchanging as few bits of communication as possible. This model was introduced by Yao [53] and is fairly well understood for the case in which Alice and Bob are classical players exchanging classical bits [36]. Much less is known about *quantum* communication complexity, where Alice and Bob have a quantum computer and can exchange qubits. This was first studied by Yao [54], and it was shown later that quantum communication complexity can be significantly smaller than classical communication complexity [21, 17, 5, 44, 16].

Let $Dcc(f)$ and $Qcc_E(f)$ denote the communication required for optimal deterministic classical and exact quantum protocols for computing f, respectively.[1] Here we assume Alice and Bob do not share any randomness or prior entanglement. Let $rank(f)$ be the rank of the $2^n \times 2^n$ communication matrix M_f, which is defined by $M_f(x,y) = f(x,y)$. The following relations are known:

$$\frac{\log rank(f)}{2} \leq Qcc_E(f) \leq Dcc(f).$$

The first inequality follows from work of Kremer [35] and Yao [54], as first noted by Buhrman, Cleve, and Wigderson [17]. (In [19] it is shown that this lower bound also holds if the quantum protocol can make use of unlimited prior entanglement between Alice and Bob.) It is an open question whether $Dcc(f)$ can in turn be upper bounded

[1] The notation $D(f)$ is used for deterministic complexity in decision tree complexity as well as in communication complexity. To avoid confusion, we will consistently add "cc" to indicate communication complexity.

by some polynomial in $\log rank(f)$. The conjecture that it can is known as the *log-rank conjecture*. If this conjecture holds, then $Dcc(f)$ and $Qcc_E(f)$ are polynomially related for all total f (which may well be true). It is known that $\log rank(f)$ and $Dcc(f)$ are not linearly related [43]. In section 3, we show that the *nondeterministic* version of $\log rank(f)$ in fact fully determines the nondeterministic version of $Qcc_E(f)$:

$$NQcc(f) = \lceil \log nrank(f) \rceil + 1.$$

Here $nrank(f)$ denotes the minimal rank of a matrix whose (x, y)-entry is nonzero iff $f(x, y) = 1$. Thus we can characterize the nondeterministic quantum communication complexity fully by the logarithm of the rank of its nondeterministic matrix. As far as we know, only two other log-rank-style characterizations of certain variants of communication complexity are known: the communication complexity of quantum sampling due to Ambainis et al. [5] and the so-called modular communication complexity due to Meinel and Waack [38].

Equality and disjointness both have nondeterministic rank 2^n, so their nondeterministic complexities are maximal: $NQcc(\text{EQ}) = NQcc(\text{DISJ}) = n+1$. Since $NQcc(f)$ lower bounds $Qcc_E(f)$, we also obtain optimal bounds for the exact quantum communication complexity of equality and disjointness. In particular, for the equality function, we get $Qcc_E(\text{EQ}) = n + 1$, which answers a question posed by Gilles Brassard in a personal communication [10]. Surprisingly, no proof of this fact seems to be known that avoids our detour via nondeterministic computation. Thus our methods also give new lower bounds for regular quantum communication complexity.

Finally, analogous to the query complexity case, we also show an exponential gap between quantum and classical nondeterministic communication complexity: we exhibit an f where $NQcc(f) \leq \log(n + 1) + 1$ and $Ncc(f) \in \Omega(n)$. Massar et al. [37] earlier found another gap that is unbounded, yet in some sense smaller: $NQcc(\text{NE}) = 2$ versus $Ncc(\text{NE}) = \log n + 1$, where NE is the nonequality function.

2. Nondeterministic quantum query complexity.

2.1. Functions and polynomials. For $x \in \{0,1\}^n$, we use $|x|$ for the Hamming weight (number of 1's) of x, and x_i for its ith bit, $i \in [n] = \{1, \ldots, n\}$. We use $\vec{0}$ for a string of n zeros. If $B \subseteq [n]$ is a set of (indices of) variables, then x^B denotes the input obtained from x by complementing all variables in B. If $x, y \in \{0,1\}^n$, then $x \wedge y$ denotes the n-bit string obtained by bitwise ANDing x and y. Let $f : \{0,1\}^n \to \{0,1\}$ be a total Boolean function. For example, $\text{OR}(x) = 1$ iff $|x| \geq 1$, $\text{AND}(x) = 1$ iff $|x| = n$, $\text{PARITY}(x) = 1$ iff $|x|$ is odd. We use $\overline{f}$ for the function $1 - f$.

For $b \in \{0,1\}$, a *b-certificate* for f is an assignment $C : S \to \{0,1\}$ to some set S of variables, such that $f(x) = b$ whenever x is consistent with C. The *size* of C is $|S|$. The *certificate complexity* $C_x(f)$ of f on input x is the minimal size of an $f(x)$-certificate that is consistent with x. We define the 1-certificate complexity of f as $C^{(1)}(f) = \max_{x:f(x)=1} C_x(f)$. We define $C^{(0)}(f)$ similarly. For example, $C^{(1)}(\text{OR}) = 1$ and $C^{(0)}(\text{OR}) = n$, but $C^{(1)}(\overline{\text{OR}}) = n$ and $C^{(0)}(\overline{\text{OR}}) = 1$.

An n-variate *multilinear polynomial* is a function $p : \mathbb{C}^n \to \mathbb{C}$ that can be written

$$p(x) = \sum_{S \subseteq [n]} a_S X_S.$$

Here S ranges over all sets of indices of variables, a_S is a complex number, and the monomial X_S is the product $\Pi_{i \in S} x_i$ of all variables in S. The *degree* $deg(p)$

of p is the degree of a largest monomial with nonzero coefficient. It is well known that every total Boolean f has a unique polynomial p such that $p(x) = f(x)$ for all $x \in \{0,1\}^n$. Let $deg(f)$ be the degree of this polynomial, which is at most n. For example, $\mathrm{OR}(x_1, x_2) = x_1 + x_2 - x_1x_2$, which has degree 2. Every multilinear polynomial $p = \sum_S a_S X_S$ can also be written out uniquely in the so-called *Fourier basis*:

$$p(x) = \sum_S c_S(-1)^{x \cdot S}.$$

Again S ranges over all sets of indices of variables (we often identify a set S with its characteristic n-bit vector), c_S is a complex number, and $x \cdot S$ denotes the inner product of the n-bit strings x and S, or, equivalently, $x \cdot S = |x \wedge S| = \sum_{i \in S} x_i$. It is easy to see that $deg(p) = \max\{|S| \mid c_S \neq 0\}$. For example, $\mathrm{OR}(x_1, x_2) = \frac{3}{4} - \frac{1}{4}(-1)^{x_1} - \frac{1}{4}(-1)^{x_2} - \frac{1}{4}(-1)^{x_1+x_2}$ in the Fourier basis. We refer to [8, 42, 20] for more details about polynomial representations of Boolean functions.

We introduce the notion of a *nondeterministic polynomial* for f. This is a polynomial p such that $p(x) \neq 0$ iff $f(x) = 1$. Let the *nondeterministic degree* of f, denoted $ndeg(f)$, be the minimum degree among all nondeterministic polynomials p for f. For example, $p(x) = \sum_{i=1}^n x_i$ is a nondeterministic polynomial for OR; hence $ndeg(\mathrm{OR}) = 1$.

We mention some upper and lower bounds for $ndeg(f)$. Let f be a nonconstant symmetric function (i.e., $f(x)$ depends only on $|x|$). Suppose f achieves value 0 on the z Hamming weights, $k_1, \ldots, k_z$. Since $|x| = \sum_i x_i$, it is easy to see that $(|x|-k_1)(|x|-k_2)\cdots(|x|-k_z)$ is a nondeterministic polynomial for f; hence $ndeg(f) \leq z$. This upper bound is tight for AND (see below) but not for PARITY. For example, $p(x_1, x_2) = x_1 - x_2$ is a degree-1 nondeterministic polynomial for PARITY on two variables: it assumes value 0 on x-weights 0 and 2 and ± 1 on weight 1. By squaring $p(x)$ and then using standard symmetrization techniques (as used, for instance, in [39, 42, 6]), we can also show the general lower bound $ndeg(f) \geq z/2$ for symmetric f. Furthermore, it is easy to show that $ndeg(f) \leq C^{(1)}(f)$ for every f. (Take a polynomial that is the "sum" over all 1-certificates for f.)

Finally, we mention a general lower bound on $ndeg(f)$. Let $\Pr[p \neq 0] = |\{x \in \{0,1\}^n \mid p(x) \neq 0\}|/2^n$ denote the probability that a random Boolean input x makes a function p nonzero. A lemma of Schwartz [46] (see also [42, section 2.2]) states that if p is a nonconstant multilinear polynomial of degree d, then $\Pr[p \neq 0] \geq 2^{-d}$, and hence $d \geq \log(1/\Pr[p \neq 0])$. Since a nondeterministic polynomial p for f is nonzero iff $f(x) = 1$, it follows that

$$ndeg(f) \geq \log(1/\Pr[f \neq 0]) = \log(1/\Pr[f = 1]).$$

Accordingly, functions with a very small fraction of 1-inputs will have high nondeterministic degree. For instance, $\Pr[\mathrm{AND} = 1] = 2^{-n}$, so $ndeg(\mathrm{AND}) = n$.

2.2. Quantum computing. We assume familiarity with classical computation and briefly sketch the setting of quantum computation (see, e.g., [40] for more details). An *m-qubit state* is a linear combination of all classical m-bit states

$$|\phi\rangle = \sum_{i \in \{0,1\}^m} \alpha_i |i\rangle,$$

where $|i\rangle$ denotes the basis state i (a classical m-bit string) and α_i is a complex number that is called the *amplitude* of $|i\rangle$. We require $\sum_i |\alpha_i|^2 = 1$. Viewing $|\phi\rangle$ as

a 2^m-dimensional column vector, we use $\langle\phi|$ for the row vector that is the conjugate transpose of $|\phi\rangle$. Note that the inner product $\langle i||j\rangle = \langle i|j\rangle$ is 1 if $i = j$ and 0 if $i \neq j$. When we observe $|\phi\rangle$, we will see $|i\rangle$ with probability $|\langle i|\phi\rangle|^2 = |\alpha_i|^2$, and the state will collapse to the observed $|i\rangle$. A quantum operation which is not an observation corresponds to a unitary (i.e., norm-preserving) transformation U on the 2^m-dimensional vector of amplitudes.

2.3. Query complexity. Suppose we want to compute some function $f : \{0,1\}^n \to \{0,1\}$. For input $x \in \{0,1\}^n$, a *query* corresponds to the unitary transformation O that maps $|i,b,z\rangle \to |i,b \oplus x_i,z\rangle$. Here $i \in [n]$ and $b \in \{0,1\}$; the z-part corresponds to the workspace, which is not affected by the query. We assume that the input can be accessed only via such queries. A T-query quantum algorithm has the form $A = U_T O U_{T-1} \cdots O U_1 O U_0$, where the U_k are fixed unitary transformations, independent of the input x. This A depends on x via the T applications of O. We sometimes write A_x to emphasize this. The algorithm starts in initial state $|\vec{0}\rangle$, and its *output* is the bit obtained from observing the leftmost qubit of the final superposition $A|\vec{0}\rangle$. The *acceptance probability* of A (on input x) is its probability of outputting 1 (on x).

We will consider classical and quantum algorithms and will count only the number of queries these algorithms make on a worst-case input. Let $D(f)$ and $Q_E(f)$ be the query complexities of optimal deterministic classical and exact quantum algorithms for computing f, respectively. $D(f)$ is also known as the decision tree complexity of f. Similarly we can define $R_2(f)$ and $Q_2(f)$ to be the query complexity of f for bounded-error classical and quantum algorithms, respectively. Quantum query complexity and its relation to classical complexity has been well studied in recent years; see, for example, [6, 4, 20].

We define a *nondeterministic algorithm* for f to be an algorithm that has positive acceptance probability on input x iff $f(x) = 1$. Let $N(f)$ and $NQ(f)$ be the query complexities of optimal nondeterministic classical and quantum algorithms for f, respectively. It is easy to show that the 1-certificate complexity fully characterizes the classical nondeterministic complexity of f.

PROPOSITION 2.1. $N(f) = C^{(1)}(f)$.

Proof. A classical algorithm that guesses a 1-certificate, queries its variables, and outputs 1 iff the certificate holds is a nondeterministic algorithm for f. Hence $N(f) \leq C^{(1)}(f)$.

A nondeterministic algorithm for f can only output 1 if the outcomes of the queries that it has made force the function to 1. Hence, if x is an input where all 1-certificates have size at least $C^{(1)}(f)$, then the algorithm will have to query at least $C^{(1)}(f)$ variables before it can output 1 (which it must do on some runs). Hence $N(f) \geq C^{(1)}(f)$. □

2.4. Algebraic characterization. Here we show that $NQ(f)$ is equal to $ndeg(f)$, using the following result from [6].

LEMMA 2.2 (see [6]). *The amplitudes of the basis states in the final superposition of a T-query quantum algorithm can be written as multilinear complex-valued polynomials of degree $\leq T$ in the n x_i-variables. Therefore, the acceptance probability of the algorithm (which is the sum of squares of some of those amplitudes) can be written as an n-variate multilinear polynomial $P(x)$ of degree $\leq 2T$.*

Note that the acceptance probability of a nondeterministic quantum algorithm is actually a nondeterministic polynomial for f, since it is positive iff $f(x) = 1$. By Lemma 2.2, this polynomial will have degree at most twice the number of queries

of the algorithm, which immediately implies $ndeg(f)/2 \leq NQ(f)$. Below we will show how we can get rid of the factor 1/2 in this lower bound, improving it to $ndeg(f) \leq NQcc(f)$. We show that this lower bound is in fact optimal by deriving a nondeterministic algorithm from a nondeterministic polynomial. This derivation uses a trick similar to the one used in [24] to show that co-C$_=$P $\subseteq$ quantum-NP.

THEOREM 2.3. $NQ(f) = ndeg(f)$.

Proof. Upper bound. Let $p(x)$ be a nondeterministic polynomial for f of degree $d = ndeg(f)$. Recall that $x \cdot S$ denotes $|x \wedge S|$, identifying $S \subseteq [n]$ with its characteristic n-bit vector. We write p in the Fourier basis:

$$p(x) = \sum_S c_S (-1)^{x \cdot S}.$$

Since $deg(p) = \max\{|S| \mid c_S \neq 0\}$, we have that $c_S \neq 0$ only if $|S| \leq d$.

We can construct a unitary transformation F that uses d queries to x and maps $|S\rangle \rightarrow (-1)^{x \cdot S}|S\rangle$ whenever $|S| \leq d$. Informally, this transformation does a controlled parity-computation: it computes $|x \cdot S| \pmod 2$ using $|S|/2$ queries [6, 23], then adds a phase "-1" if that answer is 1, and then reverses the computation to clean up the workspace and the answer at the cost of another $|S|/2$ queries. (If $|S|$ is odd, then one variable is treated separately, still using $|S|$ queries in total.)

Now consider the following quantum algorithm:

1. Start with $c \sum_S c_S |S\rangle$ (an n-qubit state, where $c = 1/\sqrt{\sum_S |c_S|^2}$ is a normalizing constant).
2. Apply F to the state.
3. Apply a Hadamard transform H to each qubit.
4. Measure the final state, and output 1 if the outcome is the all-zero state $|\vec{0}\rangle$, and output 0 otherwise.

The state after step 2 is $c \sum_S c_S (-1)^{x \cdot S} |S\rangle$. Note that the sum of the amplitudes in this state is $c \cdot p(x)$, which is nonzero iff $f(x) = 1$. The Hadamard transform in step 3 gives us this sum as amplitude of the $|\vec{0}\rangle$-state, with a normalizing factor of $1/\sqrt{2^n}$. Accordingly, the probability of observing $|\vec{0}\rangle$ at the end is

$$\begin{aligned} P(x) &= \left| \langle \vec{0} | H^{\otimes n} F c \sum_S c_S |S\rangle \right|^2 \\ &= \frac{c^2}{2^n} \left| \sum_{S'} \langle S' | \sum_S c_S (-1)^{x \cdot S} |S\rangle \right|^2 \\ &= \frac{c^2}{2^n} \left| \sum_S c_S (-1)^{x \cdot S} \right|^2 \\ &= \frac{c^2 p(x)^2}{2^n}. \end{aligned}$$

Since $p(x)$ is nonzero iff $f(x) = 1$, $P(x)$ will be positive iff $f(x) = 1$. Hence we have a nondeterministic quantum algorithm for f with $d = ndeg(f)$ queries.

Lower bound. Let $T = NQ(f)$, and consider a T-query nondeterministic quantum algorithm for f. By Lemma 2.2, the amplitudes α_i in the final state,

$$|\phi^x\rangle = \sum_i \alpha_i(x) |i\rangle,$$

on input x are n-variate polynomials of x of degree $\leq T$. We use the probabilistic method [3] to show that some linear combination of these polynomials is a nondeterministic polynomial for f, thus avoiding losing the factor $1/2$ mentioned after Lemma 2.2.

Let S be the set of basis states having a 1 as leftmost bit (observing such a state will lead the algorithm to output 1). Since the algorithm is nondeterministic, we have the following properties:

If $f(x) = 0$, then $\alpha_i(x) = 0$ for all $i \in S$.

If $f(x) = 1$, then $\alpha_i(x) \neq 0$ for at least one $i \in S$.

Let I be an arbitrary set of more than 2^n numbers. For each $i \in S$, pick a coefficient c_i uniformly at random from I, and define $p(x) = \sum_{i \in S} c_i \alpha_i(x)$. By the first property, we have $p(x) = 0$ whenever $f(x) = 0$. Now consider an x for which $f(x) = 1$, and let $k \in S$ satisfy $a = \alpha_k(x) \neq 0$. Such a k must exist by the second property. We want to show that the event $p(x) = 0$ happens only with very small probability (probability taken over the random choices of the c_i). In order to do this, we fix the random choices c_i for all $i \neq k$ and view $p(x) = ac_k + b$ as a linear function in the only not-yet-chosen coefficient c_k. Since $a \neq 0$, at most one out of $|I| > 2^n$ many possible choices of c_k can make $p(x) = 0$, so

$$\Pr[p(x) = 0] < 2^{-n}.$$

However, then, by the union bound we have

$$\Pr\left[\text{there is an } x \in f^{-1}(1) \text{ for which } p(x) = 0\right] \leq \sum_{x \in f^{-1}(1)} \Pr[p(x) = 0] < 2^n \cdot 2^{-n} = 1.$$

This probability is strictly less than 1, which shows that there exists a way of setting the coefficients c_i that satisfies $p(x) \neq 0$ for all $x \in f^{-1}(1)$, thus making p a nondeterministic polynomial for f. Since p is a sum of polynomials of degree $\leq T$, it follows that $ndeg(f) \leq deg(p) \leq T = NQ(f)$. □

2.5. Quantum-classical separation. What is the biggest possible gap between quantum and classical nondeterministic query complexity? Consider the total Boolean function $f : \{0,1\}^n \to \{0,1\}$ defined by

$$f(x) = 1 \text{ iff } |x| \neq 1.$$

It is easy to see that $N(f) = C^{(1)}(f) = C^{(0)}(f) = n$. On the other hand, the following is a degree-1 nondeterministic polynomial for f:

$$p(x) = \left(\sum_{i=1}^{n} x_i\right) - 1 = \frac{n}{2} - 1 - \frac{1}{2}\sum_{i=1}^{n}(-1)^{x_i}. \tag{2.1}$$

Thus we have that $NQ(f) = ndeg(f) = 1$. Explicitly, the 1-query algorithm that we get from the proof is as follows:

1. Start with $c\,((n/2-1)|\vec{0}\rangle - (1/2)\sum_i |e_i\rangle)$, where $c = 1/\sqrt{n^2/4 - 3n/4 + 1}$ and $|e_i\rangle$ has a 1 only at the ith bit.
2. Using one query, we can map $|e_i\rangle \to (-1)^{x_i}|e_i\rangle$.
3. Applying a Hadamard transform turns the amplitude of $|\vec{0}\rangle$ into $\alpha_{\vec{0}} = \frac{c}{\sqrt{2^n}}\,((n/2-1) - \sum_i (-1)^{x_i}/2) = cp(x)/\sqrt{2^n}$.

4. Hence the probability of observing $|\vec{0}\rangle$ at the end is $\alpha_{\vec{0}}^2 = c^2 p(x)^2/2^n$.

For the complement of f, we can easily show $NQ(\overline{f}) = ndeg(\overline{f}) \geq n-1$ (the "-1" is tight for $n=2$; witness $p(x) = x_1 - x_2$). In sum, we have the following theorem.

THEOREM 2.4. *For the above f, we have $NQ(f) = 1$, $NQ(\overline{f}) \geq n-1$, and $N(f) = N(\overline{f}) = n$.*

2.6. Relation to some other complexity measures. Many relations are known between all sorts of complexity measures of Boolean functions, such as polynomial degree, certificate complexity, various classical and quantum decision tree complexities, etc. A survey may be found in [20]. In this subsection, we will similarly embed $ndeg(f)$ $(= NQ(f))$ in this web of relations and give upper bounds on $D(f)$ in terms of $ndeg(f)$, $C(f)$, and the *block sensitivity* $bs(f)$, which is defined as follows. A set of (indices of) variables $B \subseteq [n]$ is called a *sensitive block* for f on input x if $f(x) \neq f(x^B)$; B is *minimal* if no $B' \subset B$ is sensitive. The block sensitivity $bs_x(f)$ is the maximal number of disjoint minimal sensitive blocks in x, and $bs^{(b)}(f) = \max_{x \in f^{-1}(b)} bs_x(f)$.

LEMMA 2.5. *If $f(x) = 0$ and B is a minimal sensitive block for f on x, then $|B| \leq ndeg(f)$.*

Proof. Assume without loss of generality that $x = \vec{0}$. Because B is minimal, for every proper subset B' of B, we have $f(x) = f(x^{B'}) = 0$, but on the other hand $f(x^B) = 1$. Accordingly, if we fix all variables outside of B to zero, then we obtain the AND-function of $|B|$ variables, which requires nondeterministic degree $|B|$. Hence $|B| \leq ndeg(f)$. □

LEMMA 2.6. $C^{(0)}(f) \leq bs^{(0)}(f) ndeg(f)$.

Proof. Consider any input x. As Nisan [41] proved, the union of a maximal set of sensitive blocks forms a certificate for that input (for otherwise there would be one more sensitive block). If $f(x) = 0$, then there can be at most $bs^{(0)}(f)$ disjoint sensitive blocks, and by the previous lemma each block contains at most $ndeg(f)$ variables. Hence each 0-input contains a certificate of at most $bs^{(0)}(f) ndeg(f)$ variables. □

The following theorem improves upon an argument of Nisan and Smolensky, described in [20].

THEOREM 2.7. $D(f) \leq C^{(0)}(f) ndeg(f)$.

Proof. Let p be a nondeterministic polynomial for f of degree $d = ndeg(f)$. Note that if we take a 0-certificate $C : S \to \{0,1\}$ and fix the S-variables accordingly, then p must reduce to the constant-0 polynomial. This implies that S intersects all degree-d monomials of p, because a nonintersected degree-d monomial would still be present in the reduced polynomial, which would then not be constant-0. Thus taking a minimal 0-certificate and querying its variables reduces the degree of p by at least 1. Repeating this at most $ndeg(f)$ times, we reduce p to a constant polynomial and know $f(x)$. This algorithm takes at most $C^{(0)}(f) ndeg(f)$ queries. □

Combining this with the fact that $bs^{(0)}(f) \leq 6Q_2(f)^2$ [6], we obtain the following.

COROLLARY 2.8. $D(f) \leq bs^{(0)}(f) ndeg(f)^2 \leq 6\ Q_2(f)^2 NQ(f)^2$.

This corollary has the somewhat paradoxical consequence that if the nondeterministic complexity $NQ(f)$ is small, then the bounded-error complexity $Q_2(f)$ must be large (i.e., close to $D(f)$). For instance, if $NQ(f) = O(1)$, then $Q_2(f) = \Omega(\sqrt{D(f)})$. We hope that this result will help tighten the relation $D(f) = O(Q_2(f)^6)$ that was proved in [6].

3. Nondeterministic quantum communication complexity.

3.1. Communication complexity. In the standard version of communication complexity, two parties (Alice and Bob) want to compute some function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. For example, $\mathrm{EQ}(x,y) = 1$ iff $x = y$, $\mathrm{NE}(x,y) = 1$ iff $x \neq y$, and $\mathrm{DISJ}(x,y) = 1$ iff $|x \wedge y| = 0$. A *rectangle* is a subset $R = S \times T$ of the domain of f. R is a 1-*rectangle* (for f) if $f(x,y) = 1$ for all $(x,y) \in R$. A 1-*cover* for f is a set of 1-rectangles whose union contains all 1-inputs of f. $Cov^1(f)$ denotes the minimal size (i.e., minimal number of rectangles) of a 1-cover for f. Similarly, we define 0-rectangles, 0-covers, and $Cov^0(f)$.

The *communication matrix* M_f of f is the $2^n \times 2^n$ Boolean matrix whose (x,y)-entry is $f(x,y)$, and *rank*(f) denotes the rank of M_f over the field of complex numbers. A $2^n \times 2^n$ matrix M is called a *nondeterministic communication matrix* for f if it has the property that $M(x,y) \neq 0$ iff $f(x,y) = 1$. Thus M is any matrix obtainable by replacing 1-entries in M_f by nonzero complex numbers. Let the *nondeterministic rank* of f, denoted *nrank*(f), be the minimum rank (over the complex field) among all nondeterministic matrices M for f.[2]

We consider classical and quantum communication protocols and count only the amount of communication (bits or qubits) that these protocols make on a worst-case input. For classical communication protocols, we refer to [36]. Here we briefly define quantum communication protocols, referring to the surveys [49, 15, 33, 11, 51] for more details. The space in which the quantum protocol works consists of three parts: Alice's part, the communication channel, and Bob's part. (We do not write the dimensions of these spaces explicitly.) Initially these three parts contain only 0-qubits,

$$|0\rangle|0\rangle|0\rangle.$$

We assume Alice starts the protocol. She applies a unitary transformation $U_1^A(x)$ to her private space and part of the channel. This corresponds to her initial computation and her first message. The length of this message is the number of channel qubits on which $U_1^A(x)$ acts. The total state is now

$$(U_1^A(x) \otimes I^B)|0\rangle|0\rangle|0\rangle,$$

where $\otimes$ denotes tensor product, and I^B denotes the identity transformation on Bob's part. Then Bob applies a unitary transformation $U_2^B(y) = V_2^B(y)S_2^B$ to his part and the channel. First, the operation S_2^B "reads" Alice's message by swapping the contents of the channel with some fresh $|0\rangle$-qubits in Bob's private space. After this, the unitary $V_2^B(y)$ is applied to Bob's private space and part of the channel. This corresponds to Bob's private computation and his putting a message to Alice on the channel. The length of this new message is the number of channel-qubits on which

[2]This definition looks somewhat similar to the definition of the *Colin de Verdiére parameter* $\mu(G)$ of an undirected graph G [27]. For $G = (V,E)$ with $|V| = n$, $\mu(G)$ is defined to be the maximal corank ($= n - rank$) among all real symmetric $n \times n$ matrices M having the following three properties: (1) $M_{ij} < 0$ if $(i,j) \in E$ and $M_{ij} = 0$ if $i \neq j$ and $(i,j) \notin E$; (2) M has exactly one negative eigenvalue of multiplicity 1; (3) there is no real symmetric matrix $X \neq 0$ such that $MX = 0$ and $X_{ij} = 0$ whenever $i = j$ or $M_{ij} \neq 0$. Such a matrix M is a nondeterministic matrix for the communication complexity problem $f : [n] \times [n] \to \{0,1\}$ defined by $f(i,j) = 1$ iff $(i,j) \in E$, with the promise that the inputs i and j are distinct. However, the Colin de Verdière requirement appears to be more stringent, since it constrains the nondeterministic matrix further by properties (2) and (3).

$V_2^B(y)$ acts. This process goes back and forth for some k messages, so the final state of the protocol on input (x, y) will be (in case Alice goes last as well)

$$(U_k^A(x) \otimes I^B)(I^A \otimes U_{k-1}^B(y)) \cdots (I^A \otimes U_2^B(y))(U_1^A(x) \otimes I^B)|0\rangle|0\rangle|0\rangle.$$

The total *cost* of the protocol is the total length of all messages sent, on a worst-case input (x, y). For technical convenience, we assume that at the end of the protocol the output bit is the first qubit on the channel. Thus the acceptance probability $P(x, y)$ of the protocol is the probability that a measurement of the final state gives a "1" in the first channel-qubit. Note that we do not allow intermediate measurements during the protocol. This is without loss of generality; it is well known that such measurements can be postponed until the end of the protocol at no extra communication cost.

Let $Dcc(f)$ and $Qcc_E(f)$ be the communication complexities of optimal deterministic classical and quantum protocols for computing f, respectively. A *nondeterministic protocol* for f is a protocol that has positive acceptance probability on input (x, y) iff $f(x, y) = 1$. Let $Ncc(f)$ and $NQcc(f)$ be the communication complexities of optimal nondeterministic classical and quantum protocols for f, respectively. Our $Ncc(f)$ is called $N^1(f)$ in [36].

It is not hard to show that $Ncc(f) = \lceil \log Cov^1(f) \rceil + 1$, where the "+1" is due to the fact that we want Alice and Bob both to know the output at the end of the protocol.

3.2. Algebraic characterization. Here we characterize $NQcc(f)$ in terms of $nrank(f)$. We use the following lemma. It was stated without proof by Yao [54] and in more detail by Kremer [35] and is key to many of the earlier lower bounds on quantum communication complexity as well as to ours. It is easily proven by induction on ℓ.

LEMMA 3.1 (see Yao [54] and Kremer [35]). *The final state of an ℓ-qubit protocol on input (x, y) can be written as*

$$\sum_{i \in \{0,1\}^\ell} |A_i(x)\rangle|i_\ell\rangle|B_i(y)\rangle,$$

where the $A_i(x), B_i(y)$ are vectors (of norm ≤ 1), and i_ℓ denotes the last bit of the ℓ-bit string i (the output bit).

The acceptance probability $P(x, y)$ of the protocol is the squared norm of the part of the final state that has $i_\ell = 1$. Letting a_{ij} be the 2^n-dimensional complex column vector with the inner products $\langle A_i(x)|A_j(x)\rangle$ as entries and b_{ij} the 2^n-dimensional column vector with entries $\langle B_i(y)|B_j(y)\rangle$, we can write P (viewed as a $2^n \times 2^n$ matrix) as the sum $\sum_{i,j:i_\ell=j_\ell=1} a_{ij} b_{ij}^T$ of $2^{2\ell-2}$ matrices, each of rank at most 1, so the rank of P is at most $2^{2\ell-2}$. For example, for exact protocols this gives immediately that $\ell \geq \frac{1}{2} \log rank(f) + 1$, and for nondeterministic protocols $\ell \geq \frac{1}{2} \log nrank(f) + 1$.

Below we show how we can get rid of the factor $\frac{1}{2}$ in the nondeterministic case and show that the lower bound of $\log nrank(f) + 1$ is actually optimal. The lower bound part of the proof relies on the following technical lemma.

LEMMA 3.2. *If there exist two families of vectors $\{A_1(x), \ldots, A_m(x)\} \subseteq \mathbb{C}^d$ and $\{B_1(y), \ldots, B_m(y)\} \subseteq \mathbb{C}^d$ such that, for all $x \in \{0,1\}^n$ and $y \in \{0,1\}^n$, we have*

$$\sum_{i=1}^m A_i(x) \otimes B_i(y) = 0 \text{ iff } f(x, y) = 0,$$

then $nrank(f) \leq m$.

Proof. Assume there exist two such families of vectors. Let $A_i(x)_j$ denote the jth entry of vector $A_i(x)$, and similarly let $B_i(y)_k$ denote the kth entry of vector $B_i(y)$. We use pairs $(j,k) \in \{1,\ldots,d\}^2$ to index entries of vectors in the d^2-dimensional tensor space. Note that

if $f(x,y)=0$, then $\sum_{i=1}^m A_i(x)_j B_i(y)_k = 0$ for all (j,k), and
if $f(x,y)=1$, then $\sum_{i=1}^m A_i(x)_j B_i(y)_k \neq 0$ for some (j,k).

As a first step, we want to replace the vectors $A_i(x)$ and $B_i(y)$ by numbers $a_i(x)$ and $b_i(y)$ that have similar properties. We use the probabilistic method to show that this can be done.

Let I be an arbitrary set of 2^{2n+1} numbers. Choose coefficients $\alpha_1,\ldots,\alpha_d$ and $\beta_1,\ldots,\beta_d$, each coefficient picked uniformly at random from I. For every x define $a_i(x) = \sum_{j=1}^d \alpha_j A_i(x)_j$, and for every y define $b_i(y) = \sum_{k=1}^d \beta_k B_i(y)_k$. Consider the number

$$v(x,y) = \sum_{i=1}^m a_i(x)b_i(y) = \sum_{j,k=1}^d \alpha_j\beta_k \left(\sum_{i=1}^m A_i(x)_j B_i(y)_k\right).$$

If $f(x,y)=0$, then $v(x,y)=0$ for all choices of the α_j,β_k.

Now consider some (x,y) with $f(x,y)=1$. There is a pair (j',k') for which $\sum_{i=1}^m A_i(x)_{j'} B_i(y)_{k'} \neq 0$. We want to prove that $v(x,y)=0$ happens only with very small probability. In order to do this, fix the random choices of all α_j, $j \neq j'$, and β_k, $k \neq k'$, and view $v(x,y)$ as a function of the two remaining not-yet-chosen coefficients $\alpha = \alpha_{j'}$ and $\beta = \beta_{k'}$,

$$v(x,y) = c_0\alpha\beta + c_1\alpha + c_2\beta + c_3.$$

Here we know that $c_0 = \sum_{i=1}^m A_i(x)_{j'} B_i(y)_{k'} \neq 0$. There is at most one value of α for which $c_0\alpha + c_2 = 0$. All other values of α turn $v(x,y)$ into a linear equation in β, so for those α there is at most one choice of β that gives $v(x,y)=0$. Hence out of the $(2^{2n+1})^2$ different ways of choosing (α,β), at most $2^{2n+1} + (2^{2n+1}-1)\cdot 1 < 2^{2n+2}$ choices give $v(x,y)=0$. Therefore,

$$\Pr[v(x,y)=0] < \frac{2^{2n+2}}{(2^{2n+1})^2} = 2^{-2n}.$$

Using the union bound, we now have

$$\Pr\left[\text{there is an } (x,y)\in f^{-1}(1) \text{ for which } v(x,y)=0\right] \leq \sum_{(x,y)\in f^{-1}(1)} \Pr[v(x,y)=0] < 2^{2n}\cdot 2^{-2n} = 1.$$

This probability is strictly less than 1, so there exist sets $\{a_1(x),\ldots,a_m(x)\}$ and $\{b_1(y),\ldots,b_m(y)\}$ that make $v(x,y)\neq 0$ for every $(x,y)\in f^{-1}(1)$. We thus have that

$$\sum_{i=1}^m a_i(x)b_i(y) = 0 \text{ iff } f(x,y)=0.$$

View the a_i and b_i as 2^n-dimensional vectors, let A be the $2^n \times m$ matrix having the a_i as columns, and let B be the $m \times 2^n$ matrix having the b_i as rows. Then $(AB)_{xy} = \sum_{i=1}^m a_i(x)b_i(y)$, which is 0 iff $f(x,y)=0$. Thus AB is a nondeterministic matrix for f, and $nrank(f) \leq rank(AB) \leq rank(A) \leq m$. ∎

Lemma 3.2 allows us to prove the following tight characterization.

THEOREM 3.3. $NQcc(f) = \lceil \log nrank(f) \rceil + 1$.

Proof. Upper bound. Let $r = nrank(f)$, and let M be a rank-r nondeterministic matrix for f. Let $M^T = U\Sigma V$ be the singular value decomposition of the transpose of M [28], so U and V are unitary, and Σ is a diagonal matrix whose first r diagonal entries are positive real numbers and whose other diagonal entries are 0. Below we describe a one-round nondeterministic protocol for f, using $\lceil \log r \rceil + 1$ qubits.

First, Alice prepares the state $|\phi_x\rangle = c_x \Sigma V |x\rangle$, where $c_x > 0$ is a normalizing real number that depends on x. Because only the first r diagonal entries of Σ are nonzero, only the first r amplitudes of $|\phi_x\rangle$ are nonzero, so $|\phi_x\rangle$ can be compressed into $\lceil \log r \rceil$ qubits. Alice sends these qubits to Bob. Bob then applies U to $|\phi_x\rangle$ and measures the resulting state. If he observes $|y\rangle$, then he puts 1 on the channel, and otherwise he puts 0 there. The acceptance probability of this protocol is

$$P(x,y) = |\langle y|U|\phi_x\rangle|^2 = c_x^2 |\langle y|U\Sigma V|x\rangle|^2 = c_x^2 |M_{yx}^T|^2 = c_x^2 |M_{xy}|^2.$$

Since M_{xy} is nonzero iff $f(x,y) = 1$, $P(x,y)$ will be positive iff $f(x,y) = 1$. Thus we have a nondeterministic quantum protocol for f with $\lceil \log r \rceil + 1$ qubits of communication.

Lower bound. Consider a nondeterministic ℓ-qubit protocol for f. By Lemma 3.1, its final state on input (x,y) can be written as

$$\sum_{i \in \{0,1\}^\ell} |A_i(x)\rangle |i_\ell\rangle |B_i(y)\rangle.$$

Without loss of generality, we assume the vectors $A_i(x)$ and $B_i(y)$ all have the same dimension d. Let $S = \{i \in \{0,1\}^\ell \mid i_\ell = 1\}$, and consider the part of the state that corresponds to output 1 (we drop the $i_\ell = 1$ and the $|\cdot\rangle$-notation here),

$$\phi(x,y) = \sum_{i \in S} A_i(x) \otimes B_i(y).$$

Because the protocol has acceptance probability 0 iff $f(x,y) = 0$, this vector $\phi(x,y)$ will be the zero vector iff $f(x,y) = 0$. The previous lemma gives $nrank(f) \leq |S| = 2^{\ell-1}$; hence $\log(nrank(f)) + 1 \leq NQcc(f)$. □

Note that any nondeterministic matrix for the equality function has nonzeros on its diagonal and zeros off-diagonal and hence has full rank. Thus we obtain $NQcc(\mathrm{EQ}) = n + 1$. Similarly, a nondeterministic matrix for disjointness has full rank, because reversing the ordering of the columns in M_f gives an upper triangular matrix with nonzero elements on the diagonal. This gives tight bounds for the nondeterministic as well as for the exact setting, neither of which was known prior to this work.

COROLLARY 3.4. $Qcc_E(\mathrm{EQ}) = NQcc(\mathrm{EQ}) = n+1$ *and* $Qcc_E(\mathrm{DISJ}) = NQcc(\mathrm{DISJ}) = n + 1$.

3.3. Quantum-classical separation. To repeat, classically we have $Ncc(f) = \lceil \log Cov^1(f) \rceil + 1$, and quantumly we have $NQcc(f) = \lceil \log nrank(f) \rceil + 1$. We now give a total function f with an exponential gap between $Ncc(f)$ and $NQcc(f)$. For $n > 1$, define f by

$$f(x,y) = 1 \text{ iff } |x \wedge y| \neq 1.$$

We first show that the quantum complexity $NQcc(f)$ is low.

THEOREM 3.5. *For the above f, we have $NQcc(f) \leq \lceil \log(n+1) \rceil + 1$.*

Proof. By Theorem 3.3, it suffices to prove $nrank(f) \leq n+1$. We will derive a low-rank nondeterministic matrix from the polynomial p of (2.1), using a technique from [43]. Let M_i be the matrix defined by $M_i(x,y) = 1$ if $x_i = y_i = 1$ and by $M_i(x,y) = 0$ otherwise. Notice that M_i has rank 1. Define a $2^n \times 2^n$ matrix M by

$$M(x,y) = \left(\sum_{i=1}^{n} M_i(x,y) \right) - 1.$$

Note that $M(x,y) = p(x \wedge y)$. Since p is a nondeterministic polynomial for the function which is 1 iff its input does not have weight 1, it can be seen that M is a nondeterministic matrix for f. Because M is the sum of $n+1$ rank-1 matrices, M itself has rank at most $n+1$. □

Now we show that the classical $Ncc(f)$ is high (both for f and its complement).

THEOREM 3.6. *For the above f, we have $Ncc(f) \in \Omega(n)$ and $Ncc(\overline{f}) \geq n-1$.*

Proof. Let $R_1, \ldots, R_k$ be a minimal 1-cover for f. We use the following result from [36, Example 3.22 and section 4.6], which is essentially due to Razborov [45].

> There exist sets $A, B \subseteq \{0,1\}^n \times \{0,1\}^n$ and a probability distribution $\mu : \{0,1\}^n \times \{0,1\}^n \to [0,1]$ such that all $(x,y) \in A$ have $|x \wedge y| = 0$, all $(x,y) \in B$ have $|x \wedge y| = 1$, $\mu(A) = 3/4$, and there are $\alpha, \delta > 0$ (independent of n) such that for all rectangles R, $\mu(R \cap B) \geq \alpha \cdot \mu(R \cap A) - 2^{-\delta n}$.

Since the R_i are 1-rectangles, they cannot contain elements from B. Hence $\mu(R_i \cap B) = 0$ and $\mu(R_i \cap A) \leq 2^{-\delta n}/\alpha$. However, since all elements of A are covered by the R_i, we have

$$\frac{3}{4} = \mu(A) = \mu\left(\bigcup_{i=1}^{k} (R_i \cap A) \right) \leq \sum_{i=1}^{k} \mu(R_i \cap A) \leq k \cdot \frac{2^{-\delta n}}{\alpha}.$$

Therefore, $Ncc(f) = \lceil \log k \rceil + 1 \geq \delta n + \log(3\alpha/4)$.

For the lower bound on $Ncc(\overline{f})$, consider the set $S = \{(x,y) \mid x_1 = y_1 = 1, x_i = \overline{y_i}$ for $i > 1\}$. This S contains 2^{n-1} elements, all of which are 1-inputs for $\overline{f}$. Note that if (x,y) and (x',y') are two elements from S, then $|x \wedge y'| > 1$ or $|x' \wedge y| > 1$, so a 1-rectangle for $\overline{f}$ can contain at most one element of S. This shows that a minimal 1-cover for $\overline{f}$ requires at least 2^{n-1} rectangles and $Ncc(\overline{f}) \geq n-1$. □

Another quantum-classical separation was obtained earlier by Massar et al. [37]. We include it for the sake of completeness. It shows that the nondeterministic complexity of the *non*equality problem is extremely low, in sharp contrast to the equality problem itself.

THEOREM 3.7 (see [37]). *For the nonequality problem on n bits, $NQcc(\mathrm{NE}) = 2$ versus $Ncc(\mathrm{NE}) = \log n + 1$.*

Proof. $Ncc(\mathrm{NE}) = \log n + 1$ is well known (see [36, Example 2.5]). Below we give the protocol for NE from [37].

Viewing her input x as a number $\in [0, 2^n - 1]$, Alice rotates a $|0\rangle$-qubit over an angle $x\pi/2^n$, obtaining a qubit $\cos(x\pi/2^n)|0\rangle + \sin(x\pi/2^n)|1\rangle$ which she sends to Bob. Bob rotates the qubit back over an angle $y\pi/2^n$, obtaining $\cos((x-y)\pi/2^n)|0\rangle + \sin((x-y)\pi/2^n)|1\rangle$. Bob now measures the qubit and sends back the observed bit. If $x = y$, then $\sin((x-y)\pi/2^n) = 0$, so Bob will always send 0. If $x \neq y$, then $\sin((x-y)\pi/2^n) \neq 0$, so Bob will send 1 with positive probability. □

In another direction, Klauck [34] showed that $NQcc(f)$ is in general incomparable to bounded-error quantum communication complexity: the latter may be exponentially larger or smaller, depending on f.

4. Future work. One of the main reasons for the usefulness of nondeterministic query and communication complexities in the classical case is the tight relation of these complexities with deterministic complexity.

In the query complexity (decision tree) setting, we have the well-known bound

$$\max\{N(f), N(\overline{f})\} \leq D(f) \leq N(f)N(\overline{f}).$$

We conjecture that something similar holds in the quantum case:

$$\max\left\{NQ(f), NQ(\overline{f})\right\} \leq Q_E(f) \leq D(f) \overset{?}{\leq} O(NQ(f)NQ(\overline{f})).$$

The ?-part is open and ties in with tightly embedding $NQ(f)$ and $ndeg(f)$ into the web of known relations between various complexity measures (section 2.6). This conjecture implies, for instance, $D(f) \in O(deg(f)^2)$, which would be close to optimal [42]. Similarly, it would imply $D(f) \in O(Q_0(f)^2)$, which would be close to optimal as well [18]. In both cases, the currently best relation has a fourth power instead of a square.

Similarly, for communication complexity, the following is known [36, section 2.11]:

$$\max\{Ncc(f), Ncc(\overline{f})\} \leq Dcc(f) \leq O(Ncc(f)Ncc(\overline{f})).$$

An analogous result might be true in the quantum setting, but we have been unable to prove it. So far, the best result in this direction is Klauck's observation that $Dcc(f) = O(Ncc(f)NQcc(\overline{f}))$ [33, Theorem 1].

Appendix. Comparison with alternative definitions. As mentioned in the introduction, three different definitions of nondeterministic quantum complexity are possible. We may consider the complexity of quantum algorithms that

1. output 1 iff given an appropriate *classical* certificate (and such certificates must exist iff $f(x) = 1$),
2. output 1 iff given an appropriate *quantum* certificate (and such certificates must exist iff $f(x) = 1$), or
3. output 1 with positive probability iff $f(x) = 1$.

The third definition is the one we adopted for this paper. Clearly definition 2 is at least as strong as definition 1 in the sense that the complexity of a function according to definition 2 will be less than or equal to the complexity according to definition 1. In fact, in the setting of query complexity, these two definitions are equivalent, because without loss of generality the certificate can be taken to be the purported input. See Aaronson [1] for some recent results about "quantum certificate (query) complexity."

Here we show that definition 3 is at least as strong as definition 2. We give the proof for the query complexity setting, but the same proof can be modified to work for communication complexity and other nonuniform settings as well. We then give an example in which the query complexity according to definition 3 is much less than according to definition 2. This shows that our $NQ(f)$ is in fact the most powerful definition of nondeterministic quantum query complexity.

We formalize definition 2 as follows. A *T-query quantum verifier* for f is a T-query quantum algorithm V together with a set $\mathcal{C}$ of m-qubit states, such that for all $x \in \{0,1\}^n$ we have (1) if $f(x) = 1$, then there is a $|\phi_x\rangle \in \mathcal{C}$ such that $V_x|\phi_x\rangle$ has

acceptance probability 1; and (2) if $f(x) = 0$, then $V_x|\phi\rangle$ has acceptance probability 0 for every $|\phi\rangle \in \mathcal{C}$. Informally, the set $\mathcal{C}$ contains all possible certificates: (1) for every 1-input, there is a verifiable 1-certificate in $\mathcal{C}$; and (2) for 0-inputs, there are not any. We do not put any constraints on $\mathcal{C}$. However, note that the definition implies that if $f(x) = 0$ for some x, then $\mathcal{C}$ cannot contain *all* m-qubit states; otherwise, $|\phi_x\rangle = V_x^{-1}|1\vec{0}\rangle$ would be a 1-certificate in $\mathcal{C}$ even for x with $f(x) = 0$.

We now prove that a T-query quantum verifier can be turned into a T-query nondeterministic quantum algorithm according to our third definition. This shows that the third definition is at least as powerful as the second. In fact, this even holds if we replace the acceptance probability 1 in clause (1) of the definition of a quantum verifier by just positive acceptance probability—in this case, both definitions are equivalent.

THEOREM A.1. *If there is a T-query quantum verifier V for f, then $NQ(f) \leq T$.*

Proof. The verifier V and the associated set $\mathcal{C}$ satisfy the following:

1. If $f(x) = 1$, then there is a $|\phi_x\rangle \in \mathcal{C}$ such that $V_x|\phi_x\rangle$ has acceptance probability 1.
2. If $f(x) = 0$, then $V_x|\phi\rangle$ has acceptance probability 0 for all $|\phi\rangle \in \mathcal{C}$.

Let $X_1 = \{z \mid f(z) = 1\}$. For each $z \in X_1$, choose one specific 1-certificate $|\phi_z\rangle \in \mathcal{C}$. Now let us consider some input x and see what happens if we run $V_x \otimes I$ (where I is the $2^n \times 2^n$ identity operation) on the $m + n$-qubit state

$$|\phi\rangle = \frac{1}{\sqrt{|X_1|}} \sum_{z \in X_1} |\phi_z\rangle|z\rangle.$$

V_x acts on only the first m qubits of $|\phi\rangle$; the $|z\rangle$-part remains unaffected. Therefore, running $V_x \otimes I$ on $|\phi\rangle$ gives the same acceptance probabilities as when we first randomly choose some $z \in X_1$ and then apply V_x to $|\phi_z\rangle$. In the case when $f(x) = 0$, this $V_x|\phi_z\rangle$ will have acceptance probability 0, so $(V_x \otimes I)|\phi\rangle$ will have acceptance probability 0 as well. In the case when the input x is such that $f(x) = 1$, the specific certificate $|\phi_z\rangle$ that we chose for this x will satisfy that $V_x|\phi_x\rangle$ has acceptance probability 1. However, then $(V_x \otimes I)|\phi\rangle$ has acceptance probability at least $1/|X_1| > 0$. Accordingly, $(V_x \otimes I)|\phi\rangle$ has positive acceptance probability iff $f(x) = 1$. By prefixing $V_x \otimes I$ with a unitary transformation that maps $|\vec{0}\rangle$ (of $m + n$ qubits) to $|\phi\rangle$, we have constructed a nondeterministic quantum algorithm for f with T queries. □

The above proof shows that our definition of $NQ(f)$ is at least as strong as the certificate-verifier definition. Could it be that both definitions are in fact equivalent (i.e., yield the same complexity)? The function we used in section 2.5 shows that this is not the case. Consider again

$$f(x) = 1 \text{ iff } |x| \neq 1.$$

It satisfies $NQ(f) = 1$. On the other hand, if we take a T-query verifier for f and fix the certificate for the all-0 input, we obtain a T-query algorithm that always outputs 1 on the all-0 input and that outputs 0 on all inputs of Hamming weight 1. The quantum search lower bounds [9, 6] immediately imply $T = \Omega(\sqrt{n})$. This shows that our definition of $NQ(f)$ is strictly more powerful than the certificate-verifying one.

Acknowledgments. Many thanks to Peter Høyer for stimulating discussions and for his permission to include the nondeterministic parts of our joint paper [29] in this paper. I also thank Scott Aaronson, Harry Buhrman, Richard Cleve, Wim

van Dam, Lance Fortnow, Hartmut Klauck, Peter Shor, and John Watrous for various helpful discussions, comments, and pointers to the literature. Thanks to the anonymous referees for suggesting some improvements.

REFERENCES

[1] S. Aaronson, *Quantum Certificate Complexity*, http://arxiv.org/abs/quant-ph/0210020 (2 Oct 2002); to appear in Proceedings of the 18th IEEE Conference on Computational Complexity, IEEE Computer Society, Los Alamitos, CA, 2003.

[2] L. M. Adleman, J. Demarrais, and M. A. Huang, *Quantum computability*, SIAM J. Comput., 26 (1997), pp. 1524–1540.

[3] N. Alon and J. H. Spencer, *The Probabilistic Method*, Wiley-Interscience, New York, 1992.

[4] A. Ambainis, *Quantum lower bounds by quantum arguments*, in Proceedings of the 32nd ACM Symposium on Theory of Computing, ACM, New York, 2000, pp. 636–643.

[5] A. Ambainis, L. Schulman, A. Ta-Shma, U. Vazirani, and A. Wigderson, *The quantum communication complexity of sampling*, in Proceedings of the 39th IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, Los Alamitos, CA, 1998, pp. 342–351.

[6] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf, *Quantum lower bounds by polynomials*, in Proceedings of the 39th IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, Los Alamitos, CA, 1998, pp. 352–361.

[7] N. de Beaudrap, R. Cleve, and J. Watrous, *Sharp quantum vs. classical query complexity separations*, Algorithmica, 34 (2002), pp. 449–461.

[8] R. Beigel, *The polynomial method in circuit complexity*, in Proceedings of the 8th IEEE Structure in Complexity Theory Conference, IEEE, New York, 1993, pp. 82–95.

[9] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, *Strengths and weaknesses of quantum computing*, SIAM J. Comput., 26 (1997), pp. 1510–1523.

[10] G. Brassard, *Personal communication via email*, Université de Montréal, Canada, 2000.

[11] G. Brassard, *Quantum Communication Complexity (A Survey)*, http://arxiv.org/abs/quant-ph/0101005 (1 Jan 2001).

[12] G. Brassard and P. Høyer, *An exact quantum polynomial-time algorithm for Simon's problem*, in Proceedings of the 5th Israeli Symposium on Theory of Computing and Systems, Ramat-Gan, Israel, 1997, pp. 12–23.

[13] G. Brassard, P. Høyer, and A. Tapp, *Quantum algorithm for the collision problem*, ACM SIGACT News (Cryptology Column), 28 (1997), pp. 14–19.

[14] G. Brassard, P. Høyer, and A. Tapp, *Quantum counting*, in Proceedings of the 25th ICALP, Lecture Notes in Comput. Sci. 1443, Springer-Verlag, New York, 1998, pp. 820–831.

[15] H. Buhrman, *Quantum computing and communication complexity*, Bull. Eur. Assoc. Theor. Comput. Sci. EATCS, 70 (2000), pp. 131–141.

[16] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, *Quantum fingerprinting*, Phys. Rev. Lett., 87 (2001), article 167902.

[17] H. Buhrman, R. Cleve, and A. Wigderson, *Quantum vs. classical communication and computation*, in Proceedings of the 30th ACM Symposium on Theory of Computing, ACM, New York, 1998, pp. 63–68.

[18] H. Buhrman, R. Cleve, R. de Wolf, and Ch. Zalka, *Bounds for small-error and zero-error quantum algorithms*, in Proceedings of the 40th IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, Los Alamitos, CA, 1999, pp. 358–368.

[19] H. Buhrman and R. de Wolf, *Communication complexity lower bounds by polynomials*, in Proceedings of the 16th IEEE Conference on Computational Complexity, IEEE Computer Society, Los Alamitos, CA, 2001, pp. 120–130.

[20] H. Buhrman and R. de Wolf, *Complexity measures and decision tree complexity: A survey*, Theoret. Comput. Sci., 288 (2002), pp. 21–43.

[21] R. Cleve and H. Buhrman, *Substituting quantum entanglement for communication*, Phys. Rev. A (3), 56 (1997), pp. 1201–1204.

[22] D. Deutsch and R. Jozsa, *Rapid solution of problems by quantum computation*, Proc. Roy. Soc. London Ser. A, 439 (1992), pp. 553–558.

[23] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, *A limit on the speed of quantum computation in determining parity*, Phys. Rev. Lett., 81 (1998), pp. 5442–5444.

[24] S. Fenner, F. Green, S. Homer, and R. Pruim, *Determining acceptance possibility for a quantum computation is hard for the polynomial hierarchy*, in Proceedings of the 6th Italian Conference on Theoretical Computer Science, Prato, Italy, 1998, pp. 241–252.

[25] L. Fortnow and J. Rogers, *Complexity limitations on quantum computation*, J. Comput. System Sci., 59 (1999), pp. 240–252.
[26] L. K. Grover, *A fast quantum mechanical algorithm for database search*, in Proceedings of the 28th ACM Symposium on Theory of Computing, ACM, New York, 1996, pp. 212–219.
[27] H. van der Holst, L. Lovász, and A. Schrijver, *The Colin de Verdière graph parameter*, in Graph Theory and Combinatorial Biology, János Bolyai Mathematical Society, Budapest, 1999, pp. 29–85.
[28] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
[29] P. Høyer and R. de Wolf, *Improved quantum communication complexity bounds for disjointness and equality*, in Proceedings of the 19th Annual Symposium on Theoretical Aspects of Computer Science, Lecture Notes in Comput. Sci. 2285, Springer-Verlag, New York, 2002, pp. 299–310.
[30] A. Kitaev, A. Shen, and M. Vyalyi, *Classical and Quantum Computation*, AMS, Providence, RI, 2002.
[31] A. Kitaev and J. Watrous, *Parallelization, amplification, and exponential time simulation of quantum interactive proof systems*, in Proceedings of the 32nd ACM Symposium on Theory of Computing, ACM, New York, 2000, pp. 608–617.
[32] A. Y. Kitaev, *Quantum NP*, talk given at the 2nd Workshop on Algorithms in Quantum Information Processing, DePaul University, Chicago, 1999.
[33] H. Klauck, *Quantum communication complexity*, in Proceedings of the Workshop on Boolean Functions and Applications at the 27th International Colloquium on Automata, Languages and Programming, Geneva, Switzerland, 2000, pp. 241–252.
[34] H. Klauck, *Lower bounds for quantum communication complexity*, in Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, Los Alamitos, CA, 2001, pp. 288–297.
[35] I. Kremer, *Quantum Communication*, Master's thesis, Computer Science Department, Hebrew University, Jerusalem, Israel, 1995.
[36] E. Kushilevitz and N. Nisan, *Communication Complexity*, Cambridge University Press, Cambridge, UK, 1997.
[37] S. Massar, D. Bacon, N. Cerf, and R. Cleve, *Classical simulation of quantum entanglement without local hidden variables*, Phys. Rev. A (3), 63 (2001), article 052305.
[38] C. Meinel and S. Waack, *The "log rank" conjecture for modular communication complexity*, in Proceedings of the 13th Annual Symposium on Theoretical Aspects of Computer Science, Lecture Notes in Comput. Sci. 1046, Springer-Verlag, New York, 1996, pp. 619–630.
[39] M. Minsky and S. Papert, *Perceptrons*, MIT Press, Cambridge, MA, 1968, 1988.
[40] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, UK, 2000.
[41] N. Nisan, *CREW PRAMs and decision trees*, SIAM J. Comput., 20 (1991), pp. 999–1007.
[42] N. Nisan and M. Szegedy, *On the degree of Boolean functions as real polynomials*, Comput. Complexity, 4 (1994), pp. 301–313.
[43] N. Nisan and A. Wigderson, *On rank vs. communication complexity*, Combinatorica, 15 (1995), pp. 557–565.
[44] R. Raz, *Exponential separation of quantum and classical communication complexity*, in Proceedings of the 31st ACM Symposium on Theory of Computing, ACM, New York, 1999, pp. 358–367.
[45] A. Razborov, *On the distributional complexity of disjointness*, Theoret. Comput. Sci., 106 (1992), pp. 385–390.
[46] J. T. Schwartz, *Fast probabilistic algorithms for verification of polynomial identities*, J. ACM, 27 (1980), pp. 701–717.
[47] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comput., 26 (1997), pp. 1484–1509.
[48] D. R. Simon, *On the power of quantum computation*, SIAM J. Comput., 26 (1997), pp. 1474–1483.
[49] A. Ta-Shma, *Classical versus quantum communication complexity*, ACM SIGACT News (Complexity Column 23), 30 (1999), pp. 25–34.
[50] R. de Wolf, *Characterization of non-deterministic quantum query and quantum communication complexity*, in Proceedings of the 15th IEEE Conference on Computational Complexity, IEEE Computer Society, Los Alamitos, CA, 2000, pp. 271–278.
[51] R. de Wolf, *Quantum communication and complexity*, Theoret. Comput. Sci., 287 (2002), pp. 337–353.

[52] T. Yamakami and A. C.-C. Yao, $NQP_C = co\text{-}C_{=}P$, Inform. Process. Lett., 71 (1999), pp. 63–69.

[53] A. C.-C. Yao, *Some complexity questions related to distributive computing*, in Proceedings of the 11th ACM Symposium on Theory of Computing, ACM, New York, 1979, pp. 209–213.

[54] A. C.-C. Yao, *Quantum circuit complexity*, in Proceedings of the 34th IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, Los Alamitos, CA, 1993, pp. 352–360.