Communication Complexity Lower Bounds by Polynomials*

Harry Buhrman Ronald de Wolf
CWI
P.O. Box 94709
Amsterdam, The Netherlands
{buhrman,rdewolf}@cwi.nl

Abstract

The quantum version of communication complexity allows the two communicating parties to exchange qubits and/or to make use of prior entanglement (shared EPR-pairs). Some lower bound techniques are available for qubit communication complexity, but except for the inner product function, no bounds are known for the model with unlimited prior entanglement. We show that the "log rank" lower bound extends to the strongest variant of quantum communication complexity (qubit communication + unlimited prior entanglement). By relating the rank of the communication matrix to properties of polynomials, we are able to derive some strong bounds for exact protocols. In particular, we prove both the "log rank conjecture" and the polynomial equivalence of quantum and classical communication complexity for various classes of functions. We also derive some weaker bounds for bounded-error quantum protocols.

1 Introduction

Communication complexity deals with the following kind of problem. There are two separated parties, usually called Alice and Bob. Alice receives some input $x \in X$, Bob receives some $y \in Y$, and together they want to compute some function $f(x, y)$ that depends on both x and y. Alice and Bob are allowed infinite computational power, but communication between them is expensive and has to be minimized. How many bits do Alice and Bob have to exchange in the worst-case in order to be able to compute $f(x, y)$? This model was introduced by Yao [35] and has been studied extensively, both for its applications (like lower bounds on VLSI and circuits) and for its own sake. We refer to [20, 15] for definitions and results.

*Partially supported by the EU fifth framework project QAIP, IST–1999–11234. Both authors are also affiliated with the University of Amsterdam.

An interesting variant of the above is *quantum* communication complexity: suppose that Alice and Bob each have a quantum computer at their disposal and are allowed to exchange quantum bits (qubits) and/or can make use of the quantum correlations given by pre-shared EPR-pairs (these are entangled 2-qubit states $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ of which Alice has the first qubit and Bob the second) — can they do with fewer communication than in the classical case? The answer is yes. Quantum communication complexity was first considered by Yao [36] and the first example where quantum beats classical communication complexity was given in [10]. Bigger (even exponential) gaps have been shown since [8, 2, 32, 7].

The question arises how big the gaps between quantum and classical can be for various (classes of) functions. In order to answer this, we need to exhibit limits on the power of quantum communication complexity, i.e., establish lower bounds — few of which are known currently. The main purpose of this paper is to develop tools for proving lower bounds on quantum communication protocols. We present some new lower bounds for the case where f is a total Boolean function. Most of our bounds apply only to exact quantum protocols, which always output the correct answer. However, we also have some extensions of our techniques to the case of bounded-error quantum protocols.

1.1 Lower bounds for exact protocols

Let $D(f)$ denote the classical deterministic communication complexity of f, $Q(f)$ the qubit communication complexity, and $Q^*(f)$ the qubit communication required if Alice and Bob can also make use of an unlimited supply of pre-shared EPR-pairs. Clearly $Q^*(f) \leq Q(f) \leq D(f)$. Ultimately, we would like to show that $Q^*(f)$ and $D(f)$ are polynomially related for all total functions f (as are their query complexity counterparts [4]). This requires stronger lower bound tools than we have at present. Some lower bound methods are available for $Q(f)$ [36, 19, 11, 2], but the only lower bound known for $Q^*(f)$ is for the inner prod-

uct function [11]. A strong and well known lower bound for the *classical* complexity $D(f)$ is given by the logarithm of the rank of the communication matrix for f [23]. As first noted in [8], techniques of [36, 19] imply that an $\Omega(\log rank(f))$-bound also holds for $Q(f)$. Our first result is to extend this bound to $Q^*(f)$ and to derive the optimal constant:[1]

$$Q^*(f) \geq \frac{\log rank(f)}{2}. \qquad (1)$$

This implies $n/2$ lower bounds for the Q^*-complexity of the equality and disjointness problems, for which no good bounds were known before. This $n/2$ is tight up to 1 bit, since Alice can send her n-bit input to Bob with $n/2$ qubits and $n/2$ EPR-pairs using superdense coding [6]. Our corresponding lower bound also provides a new proof of *optimality* of superdense coding. In fact, the same $n/2$ bound holds for almost all functions. Furthermore, proof of the well-known "log rank conjecture" ($D(f) \leq (\log rank(f))^k$ for some k) would now imply our desired polynomial equivalence between $D(f)$ and $Q^*(f)$ (as already noted for $D(f)$ and $Q(f)$ in [2]). However, this conjecture is a long standing open question that is probably hard to solve in full generality.

Secondly, in order to get an algebraic handle on $rank(f)$, we relate it to a property of polynomials. It is well known that every total Boolean function $g : \{0,1\}^n \to \{0,1\}$ has a unique representation as a multilinear polynomial in its n variables. For the case where Alice and Bob's function has the form $f(x,y) = g(x \wedge y)$, we show that $rank(f)$ equals the number of monomials $mon(g)$ of the polynomial that represents g ($rank(f) \leq mon(g)$ was shown in [31]). This number of monomials is often easy to count and allows to determine $rank(f)$. The functions $f(x,y) = g(x \wedge y)$ form an important class that includes inner product, disjointness, and the functions that give the biggest gaps known between $D(f)$ and $\log rank(f)$ [31] (similar techniques work for the class of functions where $f(x,y) = g(x \vee y)$ or $g(x \oplus y)$).

We use this to show that $Q^*(f) \in \Theta(D(f))$ if g is symmetric. In this case we also show that $D(f)$ is close to the classical randomized complexity. Furthermore, $Q^*(f) \leq D(f) \in O(Q^*(f)^2)$ if g is monotone. For the latter result we re-derive a result of Lovász and Saks [22] using our tools.

1.2 Lower bounds for bounded-error protocols

For the case of bounded-error quantum communication protocols, very few lower bounds are currently known (exceptions are inner product [11] and the general discrepancy bound [19]). In particular, no good lower bounds are known for the disjointness problem. The best known upper bound for this is $O(\sqrt{n}\log n)$ qubits [8], contrasting with linear classical randomized complexity [16, 33]. Since disjointness is a co-NP-complete communication problem [3], a good lower bound for this problem would imply lower bounds for all NP-hard communication problems.

In order to attack this problem, we make an effort to extend the above polynomial-based approach to bounded-error protocols. We consider the approximate rank $\widetilde{rank}(f)$, and show the bound $Q_2(f) \geq (\log \widetilde{rank}(f))/2$ for 2-sided bounded-error qubit protocols (again using techniques from [36, 19]). Unfortunately, lower bounds on $\widetilde{rank}(f)$ are much harder to obtain than for $rank(f)$. If we could prove for the case $f(x,y) = g(x \wedge y)$ that $\widetilde{rank}(f)$ roughly equals the number of monomials $\widetilde{mon}(g)$ of an approximating polynomial for g, then a $\sqrt{n}$ lower bound would follow for disjointness, because we show that disjointness requires at least $2^{\sqrt{n}}$ monomials to approximate. Since we prove that the quantities $rank(f)$ and $mon(g)$ are in fact equal in the exact case, this gives some hope for a similar result $\widetilde{rank}(f) \approx \widetilde{mon}(g)$ in the approximating case, and hence for resolving the complexity of disjointness.

The specific bounds that we actually were able to *prove* for disjointness are more limited at this point: $Q_2^*(\text{DISJ}_n) \in \Omega(\log n)$ for the general case (by an extension of techniques of [11]; the $\log n$ bound without entanglement was already known [2]), $Q_2^*(\text{DISJ}_n) \in \Omega(n)$ for 1-round protocols (using a result of [25]), and $Q_2(\text{DISJ}_n) \in \Omega(\log(n/\varepsilon))$ if the error probability has to be $< \varepsilon$.

Below we sum up the main results, contrasting the exact and bounded-error case.

- We show that $Q^*(f) \geq \log rank(f)/2$ for exact protocols with unlimited prior EPR-pairs and $Q_2(f) \geq \log \widetilde{rank}(f)/2$ for bounded-error qubit protocols without prior EPR-pairs.
- If $f(x,y) = g(x \wedge y)$ for some Boolean function g, then $rank(f) = mon(g)$. An analogous result $\widetilde{rank}(f) \approx \widetilde{mon}(g)$ for the approximate case is open.
- A polynomial for disjointness, $\text{DISJ}_n(x,y) = \text{NOR}_n(x \wedge y)$, requires 2^n monomials in the exact case (implying $Q^*(\text{DISJ}_n) \geq n/2$), and roughly $2^{\sqrt{n}}$ monomials in the approximate case.

2 Preliminaries

We use $|x|$ to denote the Hamming weight (number of 1s) of $x \in \{0,1\}^n$, x_i for the ith bit of x ($x_0 = 0$), and e_i for the string whose only 1 occurs at position i. If

[1] During discussions we had with Michael Nielsen in Cambridge (UK) in the summer of 1999 after having obtained this result, it appeared that an equivalent theorem can be derived from results about *Schmidt numbers* in [27, Section 6.4.2].

$x, y \in \{0,1\}^n$, we use $x \wedge y \in \{0,1\}^n$ for the string obtained by bitwise ANDing x and y, and similarly $x \vee y$. Let $g : \{0,1\}^n \to \{0,1\}$ be a Boolean function. We call g *symmetric* if $g(x)$ only depends on $|x|$, and *monotone* if g cannot decrease if we set more variables to 1. It is well known that each $g : \{0,1\}^n \to R$ has a unique representation as a multilinear polynomial $g(x) = \sum_{S \subseteq \{1,\ldots,n\}} a_S X_S$, where X_S is the product of the variables in S and a_S is a real number. The term $a_S X_S$ is called a *monomial* of g and $mon(g)$ denotes the number of non-zero monomials of g. A polynomial p *approximates* g if $|g(x) - p(x)| \leq 1/3$ for all $x \in \{0,1\}^n$. We use $\widetilde{mon}(g)$ for the minimal number of monomials among all polynomials that approximate g. The *degree* of a monomial is the number of its variables, and the degree of a polynomial is the largest degree of its monomials.

Let X and Y be finite sets (usually $X = Y = \{0,1\}^n$) and $f : X \times Y \to \{0,1\}$ be a Boolean function. For example, *equality* has $\mathrm{EQ}_n(x,y) = 1$ iff $x = y$, *disjointness* has $\mathrm{DISJ}_n(x,y) = 1$ iff $|x \wedge y| = 0$ (equivalently, $\mathrm{DISJ}_n(x,y) = \mathrm{NOR}_n(x \wedge y)$), and *inner product* has $\mathrm{IP}_n(x,y) = 1$ iff $|x \wedge y|$ is odd. M_f denotes the $|X| \times |Y|$ Boolean matrix whose x,y entry is $f(x,y)$, and $rank(f)$ denotes the rank of M_f over the reals. A *rectangle* is a subset $R = S \times T \subseteq X \times Y$ of the domain of f. A *1-cover* for f is a set of (possibly overlapping) rectangles that covers all and only 1s in M_f. $C^1(f)$ denotes the minimal size of a 1-cover for f. For $m \geq 1$, we use $f^{\wedge m}$ to denote the Boolean function that is the AND of m independent instances of f. That is, $f^{\wedge m} : X^m \times Y^m \to \{0,1\}$ and $f^{\wedge m}(x_1,\ldots,x_m,y_1,\ldots,y_m) = f(x_1,y_1) \wedge f(x_2,y_2) \wedge \ldots \wedge f(x_m,y_m)$. Note that $M_{f^{\wedge 2}}$ is the Kronecker product $M_f \otimes M_f$ and hence $rank(f^{\wedge m}) = rank(f)^m$.

Alice and Bob want to compute some $f : X \times Y \to \{0,1\}$. After the protocol they should both know $f(x,y)$. Their system has three parts: Alice's part, the 1-qubit channel, and Bob's part. For definitions of quantum states and operations, we refer to [28]. In the initial state, Alice and Bob share k EPR-pairs and all other qubits are zero. For simplicity we assume Alice and Bob send 1 qubit in turn, and at the end the output-bit of the protocol is put on the channel. The assumption that 1 qubit is sent per round can be replaced by a fixed number of qubits q_i for the ith round. However, in order to be able to run a quantum protocol on a superposition of inputs, it is important that the number of qubits sent in the ith round is independent of the input (x,y). An ℓ-qubit protocol is described by unitary transformations $U_1(x), U_2(y), U_3(x), U_4(y), \ldots, U_\ell(x/y)$. First Alice applies $U_1(x)$ to her part and the channel, then Bob applies $U_2(y)$ to his part and the channel, etc.

$Q(f)$ denotes the (worst-case) cost of an optimal qubit protocol that computes f exactly without prior entanglement, $C^*(f)$ denotes the cost of a protocol that communicates classical bits but can make use of an unlimited (but finite) number of shared EPR-pairs, and $Q^*(f)$ is the cost of a qubit protocol that can use shared EPR-pairs. A *clean* quantum protocol is a protocol without prior entanglement that starts with $|0\rangle|0\rangle|0\rangle$ and ends with $|0\rangle|f(x,y)\rangle|0\rangle$. We use $Q_c(f)$ to denote the minimal cost of such protocols for f. We add the superscript "1 round" for 1-round protocols, where Alice sends a message to Bob and Bob then sends the output bit. Some simple relations that hold between these measures are $Q^*(f) \leq Q(f) \leq D(f) \leq D^{1round}(f)$, and $Q(f) \leq Q_c(f) \leq 2Q(f)$ because a clean protocol can be obtained by running an unclean exact protocol, copying the answer, and reversing the unclean protocol to reset the workspace. We also have $Q^*(f) \leq C^*(f) \leq 2Q^*(f)$ because teleportation allows to send a qubit using 1 EPR-pair and 2 classical bits of communication [5], so the C^*-model can simulate the Q^*-model. For bounded-error protocols we analogously define $Q_2(f)$, $Q_2^*(f)$, $C_2^*(f)$ for quantum protocols that give the correct answer with probability at least $2/3$ on every input. We use $R_2^{pub}(f)$ for the classical bounded-error complexity in the public-coin model [20].

3 Log rank lower bound

As first noted in [8, 2], techniques of Kremer and Yao [36, 19] imply $Q(f) \in \Omega(\log rank(f))$. We first state and prove a lemma from [36, 19], then show how this gives a lower bound $Q_c(f) \geq \log rank(f) + 1$ for clean protocols without prior entanglement, and then extend this to the new result $Q^*(f) \geq (\log rank(f))/2$.

Lemma 1 (Kremer/Yao) *The final state of an ℓ-qubit protocol (without prior entanglement) on input (x,y) can be written as*

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

where the $\alpha_i(x), \beta_i(y)$ are complex numbers and the $A_i(x), B_i(y)$ are unit vectors.

Proof The proof is by induction on ℓ:

Base step. For $\ell = 0$ the lemma is obvious.

Induction step. Suppose after ℓ qubits of communication the state can be written as

$$\sum_{i \in \{0,1\}^\ell} \alpha_i(x)\beta_i(y)|A_i(x)\rangle|i_\ell\rangle|B_i(y)\rangle. \qquad (2)$$

We assume without loss of generality that it is Alice's turn: she applies $U_{\ell+1}(x)$ to her part and the channel. Note that there exist complex numbers $\alpha_{i0}(x), \alpha_{i1}(x)$ and unit vectors $A_{i0}(x), A_{i1}(x)$ such that

$$(U_{\ell+1}(x) \otimes I)|A_i(x)\rangle|i_\ell\rangle|B_i(y)\rangle =$$

$$\alpha_{i0}(x)|A_{i0}(x)\rangle|0\rangle|B_i(y)\rangle + \alpha_{i1}(x)|A_{i1}(x)\rangle|1\rangle|B_i(y)\rangle.$$

Thus every element of the superposition (2) "splits in two" when we apply $U_{\ell+1}$. Accordingly, we can write the state after $U_{\ell+1}$ in the form required by the lemma. □

Theorem 1 $Q_c(f) \geq \log rank(f) + 1$.

Proof Consider a clean ℓ-qubit protocol for f. By Lemma 1, we can write its final state as

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

The protocol is clean, so the final state is $|0\rangle|f(x,y)\rangle|0\rangle$. Hence all parts of $|A_i(x)\rangle$ and $|B_i(y)\rangle$ other than $|0\rangle$ will cancel out, and we can assume without loss of generality that $|A_i(x)\rangle = |B_i(y)\rangle = |0\rangle$ for all i. Now the amplitude of the $|0\rangle|1\rangle|0\rangle$-state is simply the sum of the amplitudes $\alpha_i(x)\beta_i(y)$ of the i for which $i_\ell = 1$. This sum is either 0 or 1, and equals the acceptance probability $P(x,y)$ of the protocol. Letting $\alpha(x)$ (resp. $\beta(y)$) be the dimension-$2^{\ell-1}$ vector whose entries are $\alpha_i(x)$ (resp. $\beta_i(y)$) for the i with $i_\ell = 1$:

$$P(x,y) = \sum_{i:i_\ell=1} \alpha_i(x)\beta_i(y) = \alpha(x)^T \cdot \beta(y).$$

Since the protocol is exact, we must have $P(x,y) = f(x,y)$. Hence if we define A as the $|X| \times d$ matrix having the $\alpha(x)$ as rows and B as the $d \times |Y|$ matrix having the $\beta(y)$ as columns, then $M_f = AB$. But now $rank(M_f) = rank(AB) \leq rank(A) \leq d \leq 2^{l-1}$, and the theorem follows. □

The previous lower bound on clean protocols suffices to prove a log rank lower bound also for the strongest model of quantum communication complexity:

Theorem 2 $Q^*(f) \geq \dfrac{\log rank(f)}{2}$.

Proof Suppose we have some exact protocol for f that uses ℓ qubits of communication and k prior EPR-pairs. We will build a clean qubit protocol without prior entanglement for $f^{\wedge m}$. First Alice makes k EPR-pairs and sends one half of each pair to Bob (at a cost of k qubits of communication). Now they run the protocol to compute the first instance of f (ℓ qubits of communication). Alice and Bob each copy the answer to a safe place, which we will call their respective 'answer bits', and they reverse the protocol (again ℓ qubits of communication). This gives them back the k EPR-pairs (and an otherwise clean workspace), which they can reuse. Now they compute the second instance of f, they each AND the answer into their answer bits (which can be done cleanly), and they reverse the protocol, etc. After all m instances of f have been computed, Alice and Bob both have the answer $f^{\wedge m}(x,y)$ left and the k EPR-pairs. Bob now sends his halves of the k pairs to Alice who sets each of the k pairs back to $|00\rangle$. The protocol thus ends up with the answer and a clean workspace, so we have a clean protocol for $f^{\wedge m}$ that uses $2m\ell + 2k$ qubits and no prior entanglement. By Theorem 1:

$$\begin{aligned} 2m\ell + 2k \geq Q_c(f^{\wedge m}) &\geq \log rank(f^{\wedge m}) + 1 \\ &= m \log rank(f) + 1, \end{aligned}$$

hence

$$\ell \geq \frac{\log rank(f)}{2} - \frac{2k-1}{2m}.$$

Since this holds for every $m > 0$, the theorem follows. □

We can derive a stronger bound for $C^*(f)$:

Theorem 3 $C^*(f) \geq \log rank(f)$.

Proof Since a qubit and an EPR-pair can be used to send 2 classical bits [6], we can devise a qubit protocol for $f \wedge f$ using $C^*(f)$ qubits (compute the two copies of f in parallel using the classical bit protocol). Hence by the previous theorem $C^*(f) \geq Q^*(f \wedge f) \geq (\log rank(f \wedge f))/2 = \log rank(f)$. □

Below we draw some consequences from these log rank lower bounds. Firstly, $M_{\mathrm{EQ_n}}$ is the identity matrix, so $rank(\mathrm{EQ}_n) = 2^n$. This gives the bounds $Q^*(\mathrm{EQ}_n) \geq n/2$, $C^*(\mathrm{EQ}_n) \geq n$ (in contrast, $Q_2(\mathrm{EQ}_n) \in \Theta(\log n)$ and $C_2^*(\mathrm{EQ}_n) \in O(1)$). The disjointness function on n bits is the AND of n disjointnesses on 1 bit (which have rank 2 each), so $rank(\mathrm{DISJ}_n) = 2^n$. The complement of the inner product function has $rank(f) = 2^n$. Thus we have the following strong lower bounds, all tight up to 1 bit:[2]

Corollary 1 $Q^*(\mathrm{EQ}_n), Q^*(\mathrm{DISJ}_n), Q^*(\mathrm{IP}_n) \geq n/2$ *and* $C^*(\mathrm{EQ}_n), C^*(\mathrm{DISJ}_n), C^*(\mathrm{IP}_n) \geq n$.

Komlós [18] has shown that the fraction of $m \times m$ Boolean matrices that have determinant 0 goes to 0 as $m \to \infty$. Hence almost all $2^n \times 2^n$ Boolean matrices have full rank 2^n, which implies that almost all functions have maximal quantum communication complexity:

Corollary 2 *Almost all* $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ *have* $Q^*(f) \geq n/2$ *and* $C^*(f) \geq n$.

[2] These bounds for IP_n are also given in [11]. The bounds for EQ_n and DISJ_n are new, and can also be shown to hold for *zero-error* protocols.

We say f satisfies the *quantum direct sum property* if computing m independent copies of f (without prior entanglement) takes $mQ(f)$ qubits of communication in the worst case. (We have no example of an f without this property.) Using the same technique as before, we can prove an equivalence between the qubit models with and without prior entanglement for such f:

Corollary 3 *If f satisfies the quantum direct sum property, then $Q^*(f) \leq Q(f) \leq 2Q^*(f)$.*

Proof $Q^*(f) \leq Q(f)$ is obvious. Using the techniques of Theorem 2 we have $mQ(f) \leq 2mQ^*(f) + k$, for all m and some fixed k, hence $Q(f) \leq 2Q^*(f)$. □

Finally, because of Theorem 2, the well-known "log rank conjecture" now implies the polynomial equivalence of deterministic classical communication complexity and exact quantum communication complexity (with or without prior entanglement) for all total f:

Corollary 4 *If $D(f) \in O((\log rank(f))^k)$, then $Q^*(f) \leq Q(f) \leq D(f) \in O(Q^*(f)^k)$ for all f.*

4 A lower bound technique via polynomials

4.1 Decompositions and polynomials

The previous section showed that lower bounds on $rank(f)$ imply lower bounds on $Q^*(f)$. In this section we relate $rank(f)$ to the number of monomials of a polynomial for f and use this to prove lower bounds for some classes of functions.

We define the *decomposition number* $m(f)$ of some function $f : \{0,1\}^n \times \{0,1\}^n \to R$ as the minimum m such that there exist functions $a_1(x), \ldots, a_m(x)$ and $b_1(y), \ldots, b_m(y)$ (from R^n to R) for which $f(x,y) = \sum_{i=1}^{m} a_i(x)b_i(y)$ for all x, y. We say that f can be *decomposed* into the m functions $a_i b_i$. Without loss of generality, the functions a_i, b_i may be assumed to be multilinear polynomials. It turns out that the decomposition number equals the rank:[3]

Lemma 2 $rank(f) = m(f)$.

Proof

$\mathbf{rank(f)} \leq \mathbf{m(f)}$: Let $f(x,y) = \sum_{i=1}^{m(f)} a_i(x)b_i(y)$, M_i be the matrix defined by $M_i(x,y) = a_i(x)b_i(y)$, r_i be the row vector whose yth entry is $b_i(y)$. Note that the xth row of M_i is $a_i(x)$ times r_i. Thus *all* rows of M_i are scalar multiples of each other, hence M_i has rank 1. Since $rank(A+B) \leq rank(A) + rank(B)$ and $M_f = \sum_{i=1}^{m(f)} M_i$, we have $rank(f) = rank(M_f) \leq \sum_{i=1}^{m(f)} rank(M_i) = m(f)$.

$\mathbf{m(f)} \leq \mathbf{rank(f)}$: Suppose $rank(f) = r$. Then there are r columns $c_1, \ldots, c_r$ in M_f that span the column space of M_f. Let A be the $2^n \times r$ matrix that has these c_i as columns. Let B be the $r \times 2^n$ matrix whose ith column is formed by the r coefficients of the ith column of M_f when written out as a linear combination of $c_1, \ldots, c_r$. Then $M_f = AB$, hence $f(x,y) = M_f(x,y) = \sum_{i=1}^{r} A_{xi}B_{iy}$. Defining functions a_i, b_i by $a_i(x) = A_{xi}$ and $b_i(y) = B_{iy}$, we have $m(f) \leq rank(f)$. □

[3] The first part of the proof employs a technique of Nisan and Wigderson [31]. They used this to prove $\log rank(f) \in O(n^{\log_3 2})$ for a specific f. Our Corollary 6, together with an easy lower bound on the number of monomials in the polynomial for their function, implies that this is tight: $\log rank(f) \in \Theta(n^{\log_3 2})$ for their f.

Combined with Theorems 2 and 3 we obtain

Corollary 5 $Q^*(f) \geq \frac{\log m(f)}{2}$ *and* $C^*(f) \geq \log m(f)$.

Accordingly, for lower bounds on quantum communication complexity it is important to be able to determine the decomposition number $m(f)$. Often this is hard. It is much easier to determine the number of monomials $mon(f)$ of f (which upper bounds $m(f)$). Below we show that in the special case where $f(x,y) = g(x \wedge y)$, these two numbers are the same.[4]

Below, a monomial is called *even* if it contains x_i iff it contains y_i, for example $2x_1x_3y_1y_3$ is even and $x_1x_3y_1$ is not. A polynomial is *even* if each of its monomials is even.

Lemma 3 *If $p : \{0,1\}^n \times \{0,1\}^n \to R$ is an even polynomial with k monomials, then $m(p) = k$.*

Proof Clearly $m(p) \leq k$. To prove the converse, consider $\mathrm{DISJ}_n(x,y) = \Pi_{i=1}^{n}(1 - x_iy_i)$, the unique polynomial for the disjointness function. Note that this polynomial contains all and only even monomials (with coefficients ± 1). Since DISJ_n has rank 2^n, it follows from Lemma 2 that DISJ_n cannot be decomposed in fewer then 2^n terms. We will show how a decomposition of p with $m(p) < k$ would give rise to a decomposition of DISJ_n with fewer than 2^n terms. Suppose we can write

$$p(x,y) = \sum_{i=1}^{m(p)} a_i(x)b_i(y).$$

Let aX_SY_S be some even monomial in p and suppose the monomial X_SY_S in DISJ_n has coefficient $c = \pm 1$. Now whenever bX_S occurs in some a_i, replace that bX_S by $(cb/a)X_S$. Using the fact that p contains only even monomials, it is not hard to see that the new polynomial obtained in this way is the same as p, except that the monomial aX_SY_S is replaced by cX_SY_S.

[4] After learning about this result, Mario Szegedy (personal communication) came up with an alternative proof of this, using Fourier transforms.

Doing this sequentially for all monomials in p, we end up with a polynomial p' (with k monomials and $m(p') \leq m(p)$) that is a subpolynomial of DISJ$_n$, in the sense that each monomial in p' also occurs with the same coefficient in DISJ$_n$. Notice that by adding all $2^n - k$ missing DISJ$_n$-monomials to p', we obtain a decomposition of DISJ$_n$ with $m(p') + 2^n - k$ terms. But any such decomposition needs at least 2^n terms, hence $m(p') + 2^n - k \geq 2^n$, which implies $k \leq m(p') \leq m(p)$. □

If $f(x,y) = g(x \wedge y)$ for some Boolean function g, then the polynomial that represents f is just the polynomial of g with the ith variable replaced by $x_i y_i$. Hence such a polynomial is even, and we obtain:

Corollary 6 *If $g : \{0,1\}^n \to \{0,1\}$ and $f(x,y) = g(x \wedge y)$, then $mon(g) = mon(f) = m(f) = rank(f)$.*

This gives a tool for lower bounding (quantum and classical) communication complexity whenever f is of the form $f(x,y) = g(x \wedge y)$: $\log mon(g) \leq C^*(f) \leq D(f)$. Below we give some applications.

4.2 Symmetric functions

As a first application we show that $D(f)$ and $Q^*(f)$ are linearly related if $f(x,y) = g(x \wedge y)$ and g is symmetric (this follows from Corollary 8 below). Furthermore, we show that the classical randomized public-coin complexity $R_2^{pub}(f)$ can be at most a $\log n$-factor less than $D(f)$ for such f (Theorem 4). We will assume without loss of generality that $g(\vec{0}) = 0$, so the polynomial representing g does not have the constant-1 monomial.

Lemma 4 *If g is a symmetric function whose lowest-weight 1-input has Hamming weight $t > 0$ and $f(x,y) = g(x \wedge y)$, then $D^{1round}(f) = \log\left(\sum_{i=t}^n \binom{n}{i} + 1\right) + 1$.*

Proof It is known (and easy to see) that $D^{1round}(f) = \log r + 1$, where r is the number of different rows of M_f (this equals the number of different columns in our case, because $f(x,y) = f(y,x)$). We count r. Firstly, if $|x| < t$ then the x-row contains only zeroes. Secondly, if $x \neq x'$ and both $|x| \geq t$ and $|x'| \geq t$ then it is easy to see that there exists a y such that $|x \wedge y| = t$ and $|x' \wedge y| < t$ (or vice versa), hence $f(x,y) \neq f(x',y)$ so the x-row and x'-row are different. Accordingly, r equals the number of different x with $|x| \geq t$, $+1$ for the 0-row, which gives the lemma. □

Lemma 5 *If g is a symmetric function whose lowest-weight 1-input has weight $t > 0$, then $(1 - o(1)) \log\left(\sum_{i=t}^n \binom{n}{i}\right) \leq \log mon(g) \leq \log\left(\sum_{i=t}^n \binom{n}{i}\right)$.*

Proof The upper bound follows from the fact that g cannot have monomials of degree $< t$. For the lower bound we distinguish two cases.

Case 1: $\mathbf{t \leq n/2}$. It is known that every non-constant symmetric function f on m variables has degree $deg(f) = m - O(m^{0.548})$ [13]. This implies that g must contain a monomial of degree d for some $d \in [n/2, n/2 + b]$ with $b \in O(n^{0.548})$, for otherwise we could set $n/2 - b$ variables to zero and obtain a non-constant symmetric function on $m = n/2 + b$ variables with degree $< n/2 \leq m - O(m^{0.548})$. But because g is symmetric, it must then contain *all* $\binom{n}{d}$ monomials of degree d. Hence by Stirling's approximation $mon(g) \geq \binom{n}{d} \geq 2^{n - O(n^{0.548})}$, which implies the lemma.

Case 2: $\mathbf{t > n/2}$. It is easy to see that g must contain all $\binom{n}{t}$ monomials of degree t. Now

$$(n - t + 1) mon(g) \geq (n - t + 1)\binom{n}{t} \geq \sum_{i=t}^n \binom{n}{i}.$$

Hence $\log mon(g) \geq \log\left(\sum_{i=t}^n \binom{n}{i}\right) - \log(n - t + 1) = (1 - o(1)) \log\left(\sum_{i=t}^n \binom{n}{i}\right)$. □

The number $mon(g)$ may be less then $\sum_{i=t}^n \binom{n}{i}$. Consider the function $g(x_1, x_2, x_3) = x_1 + x_2 + x_3 - x_1x_2 - x_1x_3 - x_2x_3$ [30]. Here $mon(g) = 6$ but $\sum_{i=1}^3 \binom{3}{i} = 7$. Hence the $1 - o(1)$ of Lemma 5 cannot be improved to 1 in general (it can if g is a threshold function).

Combining the previous results:

Corollary 7 *If g is a symmetric function whose lowest-weight 1-input has weight $t > 0$ and $f(x,y) = g(x \wedge y)$, then $(1 - o(1)) \log\left(\sum_{i=t}^n \binom{n}{i}\right) \leq C^*(f) \leq D(f) \leq D^{1round}(f) = \log\left(\sum_{i=t}^n \binom{n}{i} + 1\right) + 1$.*

Accordingly, for symmetric g the communication complexity (quantum and classical, with or without prior entanglement, 1-round and multi-round) equals $\log rank(f)$ up to small constant factors. In particular:

Corollary 8 *If g is symmetric and $f(x,y) = g(x \wedge y)$, then $(1 - o(1))D(f) \leq C^*(f) \leq D(f)$.*

We have shown that $Q^*(f)$ and $D(f)$ are equal up to constant factors whenever $f(x,y) = g(x \wedge y)$ and g is symmetric. For such f, $D(f)$ is also nearly equal to the classical bounded-error communication complexity $R_2^{pub}(f)$, where we allow Alice and Bob to share public coin flips. In order to prove this, we introduce the notion of *0-block sensitivity* in analogy to the notion of block sensitivity of Nisan [29]. For input $x \in \{0,1\}^n$, let $\mathrm{bs0}_x(g)$ be the maximal number of disjoint sets $S_1, \ldots, S_b$ of indices of variables, such that for every i we have (1) all S_i-variables have value 0 in x and (2) $g(x) \neq g(x^{S_i})$, where x^{S_i} is the string obtained from x by setting all S_i-variables to 1. Let $\mathrm{bs0}(g) = \max_x \mathrm{bs0}_x(g)$. We now have:

Lemma 6 *If g is symmetric, then $mon(g) \leq n^{2\mathrm{bs}0(g)}$.*

Proof Let t be the smallest number such that $g_t \neq g_{t+1}$, then $\mathrm{bs}0(g) \geq n - t$. If $t \leq n/2$ then $\mathrm{bs}0(g) \geq n/2$, so $mon(g) \leq 2^n \leq n^{2\mathrm{bs}0(g)}$. If $t > n/2$ then g has no monomials of degree $\leq t$, hence $mon(g) \leq \sum_{i=t+1}^{n} \binom{n}{i} \leq n^{2\mathrm{bs}0(g)}$. □

Theorem 4 *If g is a symmetric function and $f(x,y) = g(x \wedge y)$, then $D(f) \in O(R_2^{pub}(f) \log n)$.*

Proof By Corollary 7 we have $D(f) \leq (1 + o(1)) \log mon(g)$. Lemma 6 implies $D(f) \in O(\mathrm{bs}0(g) \log n)$. Moreover, $R_2^{pub}(f) \in \Omega(\mathrm{bs}0(g))$ immediately follows from Razborov's lower bound for disjointness [33] (see also [20, Section 4.6]). This implies the theorem. □

This theorem is tight for the function defined by $g(x) = 1$ iff $|x| \geq n - 1$. We have $mon(g) = n + 1$, so $\log n \leq D(f) \leq (1 + o(1)) \log n$. On the other hand, an $O(1)$ bounded-error public coin protocol can easily be derived from the well-known $O(1)$-protocol for equality: Alice tests if $|x| < n - 1$, sends a 0 if so and a 1 if not. In the first case Alice and Bob know that $f(x,y) = 0$. In the second case, we have $f(x,y) = 1$ iff $x = y$ or $y = \vec{1}$, which can be tested with 2 applications of the equality-protocol. Hence $R_2^{pub}(f) \in O(1)$.

4.3 Monotone functions

A second application concerns monotone problems. Lovász and Saks [22] prove the log rank conjecture for (among others) the following problem, which they call the *union problem for* $\mathbf{C}$. Here $\mathbf{C}$ is a monotone set system (i.e., $(A \in \mathbf{C} \wedge A \subseteq B) \Rightarrow B \in \mathbf{C}$) over some size-$n$ universe. Alice and Bob receive sets x and y (respectively) from this universe, and their task is to determine whether $x \cup y \in \mathbf{C}$. Identifying sets with their representation as n-bit strings, this problem can equivalently be viewed as a function $f(x,y) = g(x \vee y)$, where g is a monotone increasing Boolean function. Note that it doesn't really matter whether we take g increasing or decreasing, nor whether we use $x \vee y$ or $x \wedge y$, as these problems can all be converted into each other via De Morgan's laws. Our translation of rank to number of monomials now allows us to re-derive the Lovász-Saks result without making use of their combinatorial lattice theoretical machinery. We just need the following, slightly modified, result from their paper. For the sake of completeness, we have included a proof in Appendix A. A somewhat more general result may be found in [21, Section 3].

Theorem 5 (Lovász and Saks) $D(f) \in O(\log(C^1(f)) \log rank(f))$.

Theorem 6 (Lovász and Saks) *If g is monotone and $f(x,y) = g(x \wedge y)$, then $D(f) \in O((\log rank(f))^2)$.*

Proof Let $M_1, \ldots, M_k$ be all the minimal monomials in g. Each M_i induces a rectangle $R_i = S_i \times T_i$, where $S_i = \{x \mid M_i \subseteq x\}$ and $T_i = \{y \mid M_i \subseteq y\}$. Because g is monotone increasing, $g(z) = 1$ iff z makes at least one M_i true. Hence $f(x,y) = 1$ iff there is an i such that $(x,y) \in R_i$. Accordingly, the set of R_i is a 1-cover for f and $C^1(f) \leq k \leq mon(g) = rank(f)$ by Corollary 6. Plugging into Theorem 5 gives the theorem. □

Corollary 9 *If g is monotone and $f(x,y) = g(x \wedge y)$, then $D(f) \in O(Q^*(f)^2)$.*

This result can be tightened for the special case of d-level AND-OR-trees. For example, let g be a 2-level AND-of-ORs on n variables with fan-out $\sqrt{n}$ and $f(x,y) = g(x \wedge y)$. Then g has $(2\sqrt{n} - 1)^{\sqrt{n}}$ monomials and hence $Q^*(f) \geq n/2$. In contrast, the zero-error quantum complexity of f is $O(n^{3/4} \log n)$ [9].

5 Bounded-error protocols

Here we generalize the above approach to bounded-error quantum protocols. Define the *approximate rank* of f, $\widetilde{rank}(f)$, as the minimum rank among all matrices M that approximate M_f entry-wise up to $1/3$. Let the *approximate decomposition number* $\widetilde{m}(f)$ be the minimum m such that there exist functions $a_1(x), \ldots, a_m(x)$ and $b_1(y), \ldots, b_m(y)$ for which $|f(x,y) - \sum_{i=1}^{m} a_i(x) b_i(y)| \leq 1/3$ for all x, y. By the same proof as Lemma 2 we get:

Lemma 7 $\widetilde{rank}(f) = \widetilde{m}(f)$.

By a proof similar to Theorem 1 we show

Theorem 7 $Q_2(f) \geq \dfrac{\log \widetilde{m}(f)}{2}$.

Proof By Lemma 1 we can write the final state of an ℓ-qubit bounded-error protocol for f as

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

Let $\phi(x,y) = \sum_{i \in \{0,1\}^{\ell-1}} \alpha_{i1}(x) \beta_{i1}(y) |A_{i1}(x)\rangle |1\rangle |B_{i1}(y)\rangle$ be the part of the final state that corresponds to a 1-output of the protocol. For $i, j \in \{0,1\}^{\ell-1}$, define functions a_{ij}, b_{ij} by

$$a_{ij}(x) = \overline{\alpha_{i1}(x)} \alpha_{j1}(x) \langle A_{i1}(x) | A_{j1}(x) \rangle$$

$$b_{ij}(y) = \overline{\beta_{i1}(y)}\beta_{j1}(y)\langle B_{i1}(y)|B_{j1}(y)\rangle$$

Note that the acceptance probability is

$$P(x,y) = \langle\phi(x,y)|\phi(x,y)\rangle = \sum_{i,j\in\{0,1\}^{\ell-1}} a_{ij}(x)b_{ij}(y).$$

We have now decomposed $P(x,y)$ into $2^{2\ell-2}$ functions. However, we must have $|P(x,y) - f(x,y)| \leq 1/3$ for all x, y, hence $2^{2\ell-2} \geq \widetilde{m}(f)$. It follows that $\ell \geq (\log \widetilde{m}(f))/2 + 1$. □

Unfortunately, it is much harder to prove bounds on $\widetilde{m}(f)$ than on $m(f)$.[5] In the exact case we have $m(f) = mon(g)$ whenever $f(x,y) = g(x \wedge y)$, and $mon(g)$ is often easy to determine. If something similar is true in the approximate case, then we obtain strong lower bounds on $Q_2(f)$, because our next theorem gives a bound on $\widetilde{mon}(g)$ in terms of the 0-block sensitivity defined in the previous section (the proof is deferred to Appendix B).

Theorem 8 *If g is a Boolean function, then $\widetilde{mon}(g) \geq 2^{\sqrt{\mathrm{bs0}(g)/12}}$.*

In particular, for $\mathrm{DISJ}_n(x,y) = \mathrm{NOR}_n(x \wedge y)$ it is easy to see that $\mathrm{bs0}(\mathrm{NOR}_n) = n$, so $\log \widetilde{mon}(\mathrm{NOR}_n) \geq \sqrt{n/12}$ (the upper bound $\log \widetilde{mon}(\mathrm{NOR}_n) \in O(\sqrt{n}\log n)$ follows from the construction of a degree-$\sqrt{n}$ polynomial for OR_n in [30]). Consequently, a proof that the approximate decomposition number $\widetilde{m}(f)$ roughly equals $\widetilde{mon}(g)$ would give $Q_2(\mathrm{DISJ}_n) \in \Omega(\sqrt{n})$, nearly matching the $O(\sqrt{n}\log n)$ upper bound of [8]. Since $m(f) = mon(g)$ in the exact case, a result like $\widetilde{m}(f) \approx \widetilde{mon}(g)$ might be doable.

We end this section by proving some weaker lower bounds for disjointness. Firstly, disjointness has a bounded-error protocol with $O(\sqrt{n}\log n)$ qubits and $O(\sqrt{n})$ rounds [8], but if we restrict to 1-round protocols then a linear lower bound follows from a result of Nayak [25].

Proposition 1 $Q_2^{1round}(\mathrm{DISJ}_n) \in \Omega(n)$.

Proof Suppose there exists a 1-round qubit protocol with m qubits: Alice sends a message $M(x)$ of m qubits to Bob, and Bob then has sufficient information to establish whether Alice's x and Bob's y are disjoint. Note that $M(x)$ is independent of y. If Bob's input is $y = e_i$, then $\mathrm{DISJ}_n(x,y)$ is the negation of Alice's ith bit. But then the message is an $(n, m, 2/3)$ quantum random access code [1]: by choosing input $y = e_i$ and continuing the protocol, Bob can extract from $M(x)$ the ith bit of Alice (with probability $\geq 2/3$), for any $1 \leq i \leq n$ of his choice. For this the lower bound $m \geq (1 - H(2/3))n > 0.08\ n$ is known [25], where $H(\cdot)$ is the binary entropy function. □

[5] It is interesting to note that $\overline{\mathrm{IP}_n}$ (the negation of IP_n) has less than maximal approximate decomposition number. For example for $n = 2$, $m(f) = 4$ but $\widetilde{m}(f) = 3$.

Independently from our work, Klauck [17] recently noted the stronger result that k-round protocols ($k \in O(1)$) for disjointness require $\Omega(n^{1/k})$ qubits of communication (see also [26]).

For unlimited-rounds bounded-error quantum protocols for disjointness we can only prove a logarithmic lower bound, using a technique from [11] (for the model without entanglement, the bound $Q_2(\mathrm{DISJ}_n) \in \Omega(\log n)$ was already shown in [2]).

Proposition 2 $Q_2^*(\mathrm{DISJ}_n) \in \Omega(\log n)$.

Proof We sketch the proof for a protocol which maps $|x\rangle|y\rangle \rightarrow (-1)^{\mathrm{DISJ}_n(x,y)}|x\rangle|y\rangle$. Alice chooses some $i \in \{1,\ldots,n\}$ and starts with $|e_i\rangle$, Bob starts with $(1/\sqrt{2^n})\sum_y |y\rangle$. After running the protocol, Bob has state

$$|\phi_i\rangle = \sum_y \frac{(-1)^{\mathrm{DISJ}_n(e_i,y)}}{\sqrt{2^n}}|y\rangle = \sum_y \frac{(-1)^{1-y_i}}{\sqrt{2^n}}|y\rangle.$$

Note that

$$\langle\phi_i|\phi_j\rangle = \frac{1}{2^n}\sum_y (-1)^{y_i+y_j} = \delta_{ij}.$$

Hence the $|\phi_i\rangle$ form an orthogonal set, and Bob can determine exactly which $|\phi_i\rangle$ he has and thus learn i. Alice now has transmitted $\log n$ bits to Bob and the extension of Holevo's theorem that is given in [11] implies that at least $(\log n)/2$ qubits must have been communicated to achieve this, no matter how much entanglement Alice and Bob share initially. A similar analysis works for bounded-error (as in [11]). □

Finally, for the case where we want to compute disjointness with very small error probability, we can prove an $\Omega(\log(n/\varepsilon))$ bound. Here we use the subscript "ε" to indicate qubit protocols without prior entanglement whose error probability is $< \varepsilon$. We first give a tight bound for equality:

Proposition 3 *If $\varepsilon \geq 2^{-n}$, then $Q_\varepsilon(\mathrm{EQ}_n) \in \Omega(\log(\frac{n}{\varepsilon}))$.*

Proof For simplicity we assume $1/\varepsilon$ is an integer. Suppose that matrix M approximates $M_{\mathrm{EQ}_n} = I$ entry-wise up to ε. Consider the $1/\varepsilon \times 1/\varepsilon$ matrix M' that is the upper left block of M. This M' is strictly diagonally dominant: $|M'_{ii}| > 1 - \varepsilon = (\frac{1}{\varepsilon} - 1)\varepsilon > \sum_{j\neq i} |M'_{ij}|$. A strictly diagonally dominant matrix has full rank [14, Theorem 6.1.10.a], hence M itself has rank at least $1/\varepsilon$. Using Lemma 7 and Theorem 7, we now have $Q_\varepsilon(\mathrm{EQ}_n) \in \Omega(\log(1/\varepsilon))$.

Since it is also known that $Q_\varepsilon(\mathrm{EQ}_n) \in \Omega(\log n)$ for all fixed $\varepsilon < 1/2$ (this follows for instance from the result that $Q_2(f) \in \Omega(\log D(f))$ [19]), we have

$$Q_\varepsilon(\mathrm{EQ}_n) \in \Omega(\max(\log(1/\varepsilon), \log n)) = \Omega(\log(n/\varepsilon)).$$

□

We now reduce equality to disjointness. Let $x, y \in \{0,1\}^{n/2}$. Define $x' \in \{0,1\}^n$ by replacing x_i by $x_i\overline{x_i}$ in x, and $y' \in \{0,1\}^n$ by replacing y_i by $\overline{y_i}y_i$ in y. It is easy to see that $\mathrm{EQ}_{n/2}(x,y) = \mathrm{DISJ}_n(x',y')$ so from the previous proposition we obtain:

Proposition 4 *If $\varepsilon \geq 2^{-\frac{n}{2}}$, then $Q_\varepsilon(\mathrm{DISJ}_n) \in \Omega(\log(\frac{n}{\varepsilon}))$.*

In particular, both equality and disjointness require $\Omega(n)$ qubits of communication if we want the error probability ε to be exponentially small.

6 Open problems

To end this paper, we identify three important open questions in quantum communication complexity. First, are $Q^*(f)$ and $D(f)$ polynomially related for *all* total f, or at least for all f of the form $f(x,y) = g(x \wedge y)$? We have proven this for some special cases here (g symmetric or monotone), but the general question remains open. There is a close analogy between the quantum communication complexity lower bounds presented here, and the quantum query complexity bounds obtained in [4]. Let $deg(g)$ and $mon(g)$ be, respectively, the degree and the number of monomials of the polynomial that represents $g : \{0,1\}^n \to \{0,1\}$. In [4] it was shown that a quantum computer needs at least $deg(g)/2$ queries to the n variables to compute g, and that $O(deg(g)^4)$ queries suffice (see also [30]). This implies that classical and quantum query complexity are polynomially related for all total f. Similarly, we have shown here that $(\log mon(g))/2$ qubits need to be communicated to compute $f(x,y) = g(x \wedge y)$. An analogous upper bound like $Q^*(f) \in O((\log mon(g))^k)$ might be true. A similar resemblance holds in the bounded-error case. Let $\widetilde{deg}(g)$ be the minimum degree of polynomials that approximate g. In [4] it was shown that a bounded-error quantum computer needs at least $\widetilde{deg}(g)/2$ queries to compute g and that $O(\widetilde{deg}(g)^6)$ queries suffice. Here we showed that $(\log \widetilde{m}(f))/2$ qubits of communication are necessary to compute f. A similar upper bound like $Q_2(f) \in O((\log \widetilde{m}(f))^k)$ may hold.

A second open question: how do we prove good lower bounds on *bounded-error* quantum protocols? Theorems 7 and 8 of the previous section show that $Q_2(f)$ is lower bounded by $\log \widetilde{m}(f)/2$ and $\log \widetilde{mon}(g)$ is lower bounded by $\sqrt{\mathrm{bs0}(g)}$. If we could show $\widetilde{m}(f) \approx \widetilde{mon}(g)$ whenever $f(x,y) = g(x \wedge y)$, we would have $Q_2(f) \in \Omega(\sqrt{\mathrm{bs0}(g)})$. Since $m(f) = mon(g)$ in the exact case, this may well be true. As mentioned above, this is particularly interesting because it would give a near-optimal lower bound $Q_2(\mathrm{DISJ}_n) \in \Omega(\sqrt{n})$.

Third and last, does prior entanglement add much power to qubit communication, or are $Q(f)$ and $Q^*(f)$ roughly equal up to small additive or multiplicative factors? Similarly, are $Q_2(f)$ and $Q_2^*(f)$ roughly equal? The biggest gap that we know is $Q_2(\mathrm{EQ}_n) \in \Theta(\log n)$ versus $Q_2^*(\mathrm{EQ}_n) \in O(1)$.

Acknowledgments. We acknowledge helpful discussions with Alain Tapp, who first came up with the idea of reusing entanglement used in Section 3. We also thank Michael Nielsen, Mario Szegedy, Barbara Terhal for discussions, and John Tromp for help with the proof of Lemma 9 in Appendix B.

References

[1] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani. Quantum dense coding and a lower bound for 1-way quantum finite automata. In *Proceedings of 31st ACM STOC*, pages 376–383, 1999. quant-ph/9804043.

[2] A. Ambainis, L. Schulman, A. Ta-Shma, U. Vazirani, and A. Wigderson. The quantum communication complexity of sampling. In *Proceedings of 39th IEEE FOCS*, pages 342–351, 1998.

[3] L. Babai, P. Frankl, and J. Simon. Complexity classes in communication complexity theory. In *Proceedings of 27th IEEE FOCS*, pages 337–347, 1986.

[4] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. In *Proceedings of 39th IEEE FOCS*, pages 352–361, 1998. quant-ph/9802049.

[5] C. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70:1895–1899, 1993.

[6] C. Bennett and S. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69:2881–2884, 1992.

[7] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum fingerprinting. quant-ph/0102001, 1 Feb 2001.

[8] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and computation. In *Proceedings of 30th ACM STOC*, pages 63–68, 1998. quant-ph/9802040.

[9] H. Buhrman, R. Cleve, R. de Wolf, and Ch. Zalka. Bounds for small-error and zero-error quantum algorithms. In *Proceedings of 40th IEEE FOCS*, pages 358–368, 1999. cs.CC/9904019.

[10] R. Cleve and H. Buhrman. Substituting quantum entanglement for communication. *Physical Review A*, 56(2):1201–1204, 1997. quant-ph/9704026.

[11] R. Cleve, W. van Dam, M. Nielsen, and A. Tapp. Quantum entanglement and the communication complexity of the inner product function. In *Proceedings of 1st NASA QCQC conference*, volume 1509 of *Lecture Notes in Computer Science*, pages 61–74. Springer, 1998. quant-ph/9708019.

[12] H. Ehlich and K. Zeller. Schwankung von Polynomen zwischen Gitterpunkten. *Mathematische Zeitschrift*, 86:41–44, 1964.

[13] J. von zur Gathen and J. R. Roche. Polynomials with two values. *Combinatorica*, 17(3):345–362, 1997.

[14] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[15] J. Hromkovič. *Communication Complexity and Parallel Computing*. Springer, 1997.

[16] B. Kalyanasundaram and G. Schnitger. The probabilistic communication complexity of set intersection. *SIAM Journal on Computing*, 5(4):545–557, 1992.

[17] H. Klauck. On rounds in quantum communication. quant-ph/0004100, 26 Apr 2000. To appear in STOC'01, combined with [26].

[18] J. Komlós. On the determinant of $(0,1)$-matrices. *Studia scientiarum mathematicarum Hungarica*, 2:7–21, 1967.

[19] I. Kremer. Quantum communication. Master's thesis, Hebrew University, Computer Science Department, 1995.

[20] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[21] L. Lovász. Communication complexity: A survey. In *Path, Flows, and VLSI-Layout*, pages 235–265. Springer, 1990.

[22] L. Lovász and M. Saks. Communication complexity and combinatorial lattice theory. *Journal of Computer and Systems Sciences*, 47:322–349, 1993. Earlier version in FOCS'88.

[23] K. Mehlhorn and E. Schmidt. Las Vegas is better than determinism in VLSI and distributed computing. In *Proceedings of 14th ACM STOC*, pages 330–337, 1982.

[24] M. Minsky and S. Papert. *Perceptrons*. MIT Press, Cambridge, MA, 1968. Second, expanded edition 1988.

[25] A. Nayak. Optimal lower bounds for quantum automata and random access codes. In *Proceedings of 40th IEEE FOCS*, pages 369–376, 1999. quant-ph/9904093.

[26] A. Nayak, A. Ta-Shma, and D. Zuckerman. Interaction in quantum communication complexity. quant-ph/0005106, 25 May 2000. To appear in STOC'01, combined with [17].

[27] M. A. Nielsen. *Quantum Information Theory*. PhD thesis, University of New Mexico, Albuquerque, 1998. quant-ph/0011036.

[28] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[29] N. Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, 1991. Earlier version in STOC'89.

[30] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4(4):301–313, 1994. Earlier version in STOC'92.

[31] N. Nisan and A. Wigderson. On rank vs. communication complexity. *Combinatorica*, 15(4):557–565, 1995. Earlier version in FOCS'94.

[32] R. Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of 31st ACM STOC*, pages 358–367, 1999.

[33] A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

[34] T. J. Rivlin and E. W. Cheney. A comparison of uniform approximations on an interval and a finite subset thereof. *SIAM Journal on Numerical Analysis*, 3(2):311–320, 1966.

[35] A. C.-C. Yao. Some complexity questions related to distributive computing. In *Proceedings of 11th ACM STOC*, pages 209–213, 1979.

[36] A. C.-C. Yao. Quantum circuit complexity. In *Proceedings of 34th IEEE FOCS*, pages 352–360, 1993.

A Proof of Theorem 5

Theorem 5 (Lovász and Saks) $D(f) \leq (1+\log(C^1(f)+1))(2+\log rank(f))$.

Proof We will first give a protocol based on a 0-cover. Let $c = C^0(f)$ and $R_1, \ldots, R_c$ be an optimal 0-cover. Let $R_i = S_i \times T_i$. We will also use S_i to denote the $|S_i| \times 2^n$ matrix of S_i-rows and T_i for the $2^n \times |T_i|$ matrix of T_i-columns. Call R_i *type 1* if $rank(S_i) \leq rank(M_f)/2$, and *type 2* otherwise. Note that $rank(S_i) + rank(T_i) \leq rank(M_f)$, hence at least one of $rank(S_i)$ and $rank(T_i)$ is $\leq rank(M_f)/2$.

The protocol is specified recursively as follows. Alice checks if her x occurs in some type 1 R_i. If no, then she sends a 0 to Bob; if yes, then she sends the index i and they continue with the reduced function g (obtained by shrinking Alice's domain to S_i), which has $rank(g) = rank(S_i) \leq rank(M_f)/2$. If Bob receives a 0, he checks if his y occurs in some type 2 R_j. If no, then he knows that (x,y) does not occur in any R_i, so $f(x,y) = 1$ and he sends a 0 to Alice to tell her; if yes, then he sends j and they continue with the reduced function g, which has $rank(g) = rank(T_i) \leq rank(M_f)/2$ because R_j is type 2. Thus Alice and Bob either learn $f(x,y)$ or reduce to a function g with $rank(g) \leq rank(f)/2$, at a cost of at most $1+\log(c+1)$ bits. It now follows by induction on the rank that $D(f) \leq (1+\log(C^0(f)+1))(1+\log rank(f))$. Since $C^1(f) = C^0(\overline{f})$ and $|rank(f) - rank(\overline{f})| \leq 1$, we have $D(f) = D(\overline{f}) \leq (1+\log(C^0(\overline{f})+1))(1+\log rank(\overline{f})) \leq (1+\log(C^1(f)+1))(2+\log rank(f))$. □

B Proof of Theorem 8

Here we prove Theorem 8. The proof uses some tools from the degree-lower bound proofs of Nisan and Szegedy [30, Section 3], including the following result from [12, 34]:

Theorem 9 (Ehlich, Zeller; Rivlin, Cheney) *Let p be a single-variate polynomial of degree $deg(p)$ such that $b_1 \leq$*

$p(i) \leq b_2$ for every integer $0 \leq i \leq n$, and the derivative satisfies $|p'(x)| \geq c$ for some real $0 \leq x \leq n$. Then $deg(p) \geq \sqrt{cn/(c + b_2 - b_1)}$.

A *hypergraph* is a set system $H \subseteq \mathcal{P}ow\{1, \ldots, n\}$. The sets $E \in H$ are called the *edges* of H. We call H an *s-hypergraph* if all $E \in H$ satisfy $|E| \geq s$. A set $S \subseteq \{1, \ldots, n\}$ is a *blocking set* for H if it "hits" every edge: $S \cap E \neq \emptyset$ for all $E \in H$.

Lemma 8 *Let $g : \{0,1\}^n \to \{0,1\}$ be a Boolean function for which $g(\vec{0}) = 0$ and $g(e_i) = 1$, p be a multilinear polynomial that approximates g (i.e., $|g(x) - p(x)| \leq 1/3$ for all $x \in \{0,1\}^n$), and H be the $\sqrt{n/12}$-hypergraph formed by the set of all monomials of p that have degree $\geq \sqrt{n/12}$. Then H has no blocking set of size $\leq n/2$.*

Proof Assume, by way of contradiction, that there exists a blocking set S of H with $|S| \leq n/2$. Obtain restrictions h and q of g and p, respectively, on $n - |S| \geq n/2$ variables by fixing all S-variables to 0. Then q approximates h and all monomials of q have degree $< \sqrt{n/12}$ (all p-monomials of higher degree have been set to 0 because S is a blocking set for H). Since q approximates h we have $q(\vec{0}) \in [-1/3, 1/3]$, $q(e_i) \in [2/3, 4/3]$, and $q(x) \in [-1/3, 4/3]$ for all other $x \in \{0,1\}^n$. By standard symmetrization techniques [24, 30], we can turn q into a single-variate polynomial r of degree $< \sqrt{n/12}$, such that $r(0) \in [-1/3, 1/3]$, $r(1) \in [2/3, 4/3]$, and $r(i) \in [-1/3, 4/3]$ for $i \in \{2, \ldots, n/2\}$. Since $r(0) \leq 1/3$ and $r(1) \geq 2/3$, we must have $p'(x) \geq 1/3$ for some real $x \in [0,1]$. But then $deg(r) \geq \sqrt{(1/3)(n/2)/(1/3 + 4/3 + 1/3)} = \sqrt{n/12}$ by Theorem 9, contradiction. Hence there is no blocking set S with $|S| \leq n/2$. □

The next lemma shows that H must be large if it has no blocking set of size $\leq n/2$:

Lemma 9 *If H is an s-hypergraph of size $m < 2^s$, then H has a blocking set of size $\leq n/2$.*

Proof We use the probabilistic method to show the existence of a blocking set S. Randomly choose a set S of $n/2$ elements. The probability that S does not hit some specific $E \in H$ is

$$\frac{\binom{n-|E|}{n/2}}{\binom{n}{n/2}} = \frac{\frac{n}{2}(\frac{n}{2}-1)\ldots(\frac{n}{2}-|E|+1)}{n(n-1)\ldots(n-|E|+1)} \leq 2^{-|E|}.$$

Then the probability that there is some edge $E \in H$ that is not hit by S is

$$\Pr[\bigvee_{E \in H} \text{S does not hit E}] \leq \sum_{E \in H} \Pr[\text{S does not hit E}] \leq$$

$$\sum_{E \in H} 2^{-|E|} \leq m \cdot 2^{-s} < 1.$$

Thus with positive probability, S hits all $E \in H$, which proves the existence of a blocking set. □

The above lemmas allow us to prove:

Theorem 8 *If g is a Boolean function, then $\widetilde{mon}(g) \geq 2^{\sqrt{\mathrm{bs0}(g)/12}}$.*

Proof Let p be a polynomial that approximates g with $\widetilde{mon}(g)$ monomials. Let $b = \mathrm{bs0}(g)$, and z and $S_1, \ldots, S_b$ be the input and sets that achieve the 0-block sensitivity of g. We assume without loss of generality that $g(z) = 0$.

We derive a b-variable Boolean function $h(y_1, \ldots, y_b)$ from $g(x_1, \ldots, x_n)$ as follows: if $j \in S_i$ then we replace x_j in g by y_i, and if $j \notin S_i$ for any i, then we fix x_j in g to the value z_j. Note that h satisfies

1. $h(\vec{0}) = g(z) = 0$
2. $h(e_i) = g(z^{S_i}) = 1$ for all unit $e_i \in \{0,1\}^b$
3. $\widetilde{mon}(h) \leq \widetilde{mon}(g)$, because we can easily derive an approximating polynomial for h from p, without increasing the number of monomials in p.

It follows easily from combining the previous lemmas that any approximating polynomial for h requires at least $2^{\sqrt{b/12}}$ monomials, which concludes the proof. □