Quantum Information and Computation, Vol. 10, No. 5&6 (2010) 0435–0455

ON QUANTUM-CLASSICAL EQUIVALENCE FOR COMPOSED COMMUNICATION PROBLEMS

ALEXANDER A. SHERSTOV
Department of Computer Sciences, University of Texas at Austin
Austin, Texas 78757 USA

Received October 28, 2009
Revised January 6, 2010

An open problem in communication complexity proposed by several authors is to prove that for every Boolean function f, the task of computing $f(x \wedge y)$ has polynomially related classical and quantum bounded-error complexities. We solve a variant of this question. For every f, we prove that the task of computing, on input x and y, *both* of the quantities $f(x \wedge y)$ and $f(x \vee y)$ has polynomially related classical and quantum bounded-error complexities. We further show that the quantum bounded-error complexity is polynomially related to the classical deterministic complexity and the block sensitivity of f. This result holds regardless of prior entanglement.

Keywords: Quantum communication complexity, lower bounds, quantum-classical equivalence, pattern matrix method, block sensitivity
Communicated by: R Cleve &M Mosca

1 Introduction

Quantum communication complexity, introduced by Yao [34], studies the amount of quantum communication necessary to compute a Boolean function F whose arguments are distributed among several parties. In the canonical setting, one considers a function $F\colon X \times Y \to \{0,1\}$, where X and Y are some finite sets. One of the parties, Alice, receives an input $x \in X$, and the other party, Bob, receives an input $y \in Y$. Their objective is to evaluate $F(x,y)$. To this end, Alice and Bob can exchange messages through a shared quantum communication channel. They can additionally take advantage of arbitrary *prior entanglement*. The cost of a communication protocol is the total number of qubits exchanged in the worst case on any input (x,y). The bounded-error quantum communication complexity of F with prior entanglement, denoted $Q^*_{1/3}(F)$, is the least cost of a protocol that computes F correctly with probability at least $2/3$ on every input. Quantum communication has an obvious classical counterpart, the *randomized model*, in which the parties exchange classical bits (0 and 1) and additionally share an unlimited supply of unbiased random bits. The bounded-error classical communication complexity of F, denoted $R_{1/3}(F)$, is the least cost of a randomized protocol that computes F correctly with probability at least $2/3$ on every input.

A central goal of the field is to determine whether quantum communication can be significantly more powerful than classical communication, i.e., whether a superpolynomial gap exists between the quantities $Q^*_{1/3}(F)$ and $R_{1/3}(F)$ for some function $F\colon X \times Y \to \{0,1\}$.

Exponential separations between quantum and classical complexity are well known in several alternate models of communication [2, 24, 6, 3, 12, 13, 11, 10, 14], such as one-way communication, simultaneous message passing, sampling, and computing a partial function or relation. However, these results do not apply to the original question about $Q^*_{1/3}(F)$ and $R_{1/3}(F)$, and the largest known separation between the two quantities is the quadratic gap for the disjointness function [25, 1].

It is conjectured that $Q^*_{1/3}(F)$ and $R_{1/3}(F)$ are polynomially related for all $F\colon X \times Y \to \{0,1\}$. Despite consistent research efforts, this conjecture appears to be beyond the reach of the current techniques. An intermediate goal, proposed by several authors [8, 16, 32, 31] and still unattained, is to prove the conjecture for the class of communication problems $F\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ of the form

$$F(x,y) = f(x \wedge y)$$

for an arbitrary function $f\colon \{0,1\}^n \to \{0,1\}$. There has been encouraging progress on this problem. In a breakthrough result, Razborov [25] solved it for the special case of symmetric f. Using unrelated techniques, a polynomial relationship between quantum and classical complexity was proved in [30] for the broader class of problems $F\colon \{0,1\}^{4n} \times \{0,1\}^{4n} \to \{0,1\}$ given by

$$F(x,y) = f(\dots,\ (x_{i,1}y_{i,1} \vee \dots \vee x_{i,4}y_{i,4}),\ \dots)$$

for an arbitrary function $f\colon \{0,1\}^n \to \{0,1\}$. Independently, Shi and Zhu [32] used a different approach to prove a polynomial relationship between quantum and classical communication complexity for the family of functions $F\colon \{0,1\}^{kn} \times \{0,1\}^{kn} \to \{0,1\}$ given by

$$F(x,y) = f(\dots, g(x_{i,1}, y_{i,1}, \dots, x_{i,k}, y_{i,k}), \dots),$$

where $f\colon \{0,1\}^n \to \{0,1\}$ is arbitrary and g is any gadget on $2k \geqslant \Omega(\log n)$ variables that has certain pseudorandom analytic properties. More recently, Montanaro and Osborne [22] studied quantum-classical equivalence for functions of the form $f(x \oplus y)$, where the combining function f obeys certain constraints such as monotonicity or suitable Fourier structure.

1.1 Our Results

While the above results give further evidence that quantum and classical communication complexities are polynomially related, it remains open to prove this conjecture for all functions of the form $F(x,y) = f(x \wedge y)$. In this paper, we solve a variant of this question. Specifically, we consider the communication problem of computing, on input $x, y \in \{0,1\}^n$, *both* of the quantities $f(x \wedge y)$ and $f(x \vee y)$. Our main result is a polynomial relationship between the quantum and classical complexity of any such problem, regardless of f. We further show that the quantum complexity of any such problem is polynomially related to its *deterministic* classical complexity $D(F)$ and to the *block sensitivity* $\mathrm{bs}(f)$ of f. A formal definition of block sensitivity, a well-studied combinatorial complexity measure, will be given later in Section 2.3.

Theorem 1.1 (On quantum-classical equivalence) *Let $f\colon \{0,1\}^n \to \{0,1\}$ be arbitrary. Let F denote the communication problem of computing, on input $x, y \in \{0,1\}^n$, both of the quantities $f(x \wedge y)$ and $f(x \vee y)$. Then*

$$D(F) \geqslant R_{1/3}(F) \geqslant Q^*_{1/3}(F) \geqslant \Omega(\mathrm{bs}(f)^{1/4}) \geqslant \Omega(D(F)^{1/12}).$$

A corollary of Theorem 1.1 is that given any f, a polynomial relationship between the classical and quantum complexities is assured for at least one of the communication problems $f(x \wedge y)$, $f(x \vee y)$. More precisely, we have:

Corollary 1.2 *Let $f\colon \{0,1\}^n \to \{0,1\}$ be arbitrary. Let F_1 and F_2 denote the communication problems of computing $f(x \wedge y)$ and $f(x \vee y)$, respectively. Then either*

$$D(F_1) \geqslant R_{1/3}(F_1) \geqslant Q^*_{1/3}(F_1) \geqslant \Omega(\mathrm{bs}(f)^{1/4}) \geqslant \Omega(D(F_1)^{1/12}) \tag{1.1}$$

or

$$D(F_2) \geqslant R_{1/3}(F_2) \geqslant Q^*_{1/3}(F_2) \geqslant \Omega(\mathrm{bs}(f)^{1/4}) \geqslant \Omega(D(F_2)^{1/12}) \tag{1.2}$$

or both.

Proof. Theorem 1.1 implies (1.1) if $Q^*_{1/3}(F_1) \geqslant Q^*_{1/3}(F_2)$ and implies (1.2) otherwise. □

Remark 1.3 As a matter of formalism, the communication problem in Theorem 1.1 can be expressed in standard form $F\colon X \times Y \to \{0,1\}$ by introducing an additional bit $b \in \{0,1\}$ to indicate the desired output, i.e., $f(x \wedge y)$ or $f(x \vee y)$.

Apart from giving a polynomial relationship between the quantum and classical complexity of our functions, Theorem 1.1 shows that prior entanglement does not affect their quantum complexity by more than a polynomial. It is an open problem [8] to prove a polynomial relationship for quantum communication complexity with and without prior entanglement, up to an additive logarithmic term. Known separations here are quite modest: entanglement allows for a factor of 2 savings via superdense coding, as well as an additive $\Theta(\log n)$ savings for the equality function. Finally, we prove in Section 6 that the communication problems in Theorem 1.1 satisfy another well-known conjecture, the *log-rank conjecture* of Lovász and Saks [18].

Up to this point, we have focused on the communication problem of computing $f(x \wedge y)$ and $f(x \vee y)$. In Section 7, we consider quantum-classical equivalence and the log-rank conjecture in a broader context. Specifically, we consider general compositions of the form $f(\dots, g_i(x^{(i)}, y^{(i)}), \dots)$, where one has a combining function $f\colon \{0,1\}^n \to \{0,1\}$ that receives input from intermediate functions $g_i\colon X_i \times Y_i \to \{0,1\}$, $i = 1, 2, \dots, n$. We show that under natural assumptions on $g_1, \dots, g_n$, the composed function will have polynomially related quantum and classical bounded-error complexities and will satisfy the log-rank conjecture.

1.2 Our Techniques

We obtain our main result by bringing together *analytic* and *combinatorial* views of the uniform approximation of Boolean functions. The analytic approach and combinatorial approach have each found important applications in isolation, e.g., [23, 4, 8, 25, 30, 32]. The key to our work is to find a way to combine them.

On the analytic side, a key ingredient in our solution is the *pattern matrix method*, developed in [29, 30]. Let $f\colon \{0,1\}^n \to \{0,1\}$ be a given function. The pattern matrix method

centers around a communication game in which Alice is given a string $x \in \{0,1\}^N$, where $N \geqslant 4n$; Bob is given a subset $S \subset \{1,2,\ldots,N\}$, where $|S| = n$; and their objective is to compute $f(x|_S)$, where $x|_S = (x_{i_1},\ldots,x_{i_n}) \in \{0,1\}^n$ and $i_1 < \cdots < i_n$ are the elements of S. The pattern matrix method gives a lower bound on the communication complexity of this problem in a given model (e.g., randomized, bounded-error quantum with prior entanglement, unbounded-error, weakly-unbounded error) in terms of the corresponding analytic property of f (e.g., its approximate degree or threshold degree).

Essential to the pattern matrix method, as applied in this paper, is a closed-form expression for the singular values of every matrix of the form

$$\Psi = \Big[\psi(x|_S \oplus w)\Big]_{x,(S,w)} \tag{1.3}$$

in terms of the Fourier spectrum of the function $\psi\colon \{0,1\}^n \to \mathbb{R}$, where x and S are as described in the previous paragraph and w ranges over $\{0,1\}^n$. The method critically exploits the fact that the rows of Ψ are applications of the *same* function ψ to various subsets of the variables or their negations. In the communication problems of this paper, this assumption is violated: as Bob's input y ranges over $\{0,1\}^n$, the induced functions $f_y(x) = f(x \wedge y)$ may have nothing to do with each other. This obstacle is fundamental: allowing a distinct function ψ in each row of (1.3) disrupts the spectral structure of Ψ and makes it impossible to force the desired spectral bounds.

We overcome this obstacle by exploiting the additional *combinatorial* structure of the base function $f\colon \{0,1\}^n \to \{0,1\}$, which did not figure in previous work [29, 30]. Specifically, we consider the sensitivity of f, the block sensitivity of f, and their polynomial equivalence in our restricted setting, as proved by Kenyon and Kutin [15]. We use this combinatorial structure to identify a large submatrix inside $[f(x \wedge y)]_{x,y}$ or $[f(x \vee y)]_{x,y}$ which, albeit not directly representable in the form (1.3), has a certain *dual* matrix that can be represented precisely in this way. Since the pattern matrix method relies only on the spectral structure of this dual matrix, we are able to achieve our goal and place a strong lower bound on the quantum communication complexity. The corresponding upper bound for classical protocols has a short proof using a well-known argument in the literature [7, 4, 25, 30, 32].

The above program can be equivalently described in terms of polynomials rather than functions. Let $\mathcal{F}$ be a subset of Boolean functions $\{0,1\}^n \to \{0,1\}$ none of which can be approximated within ϵ in the ℓ_∞ norm by a polynomial of degree less than d. For each $f \in \mathcal{F}$, linear programming duality implies the existence of a function $\psi\colon \{0,1\}^n \to \mathbb{R}$ such that $\sum_{x\in\{0,1\}^n} \psi(x)f(x) > \epsilon \sum_{x\in\{0,1\}^n} |\psi(x)|$ and ψ has zero Fourier mass on the characters of order less than d. This dual object ψ witnesses the fact that f has no low-degree approximant. Now, there is no reason to believe that a *single* witness ψ can be found that works for every function in $\mathcal{F}$. A key technical challenge in this work is to show that, under suitable combinatorial constraints that hold in our setting, the family $\mathcal{F}$ will indeed have a common witness ψ. In conjunction with the pattern matrix method, we are then able to solve the original problem. To clarify the relevance of this discussion to the study of functions of the form $f(x \wedge y)$, the family $\mathcal{F}$ in question is the family of the induced functions $f_y(x) = f(x \wedge y)$ as the input y ranges over $\{0,1\}^n$.

2 Preliminaries

For convenience of notation, we will view Boolean functions in the remainder of the paper as mappings $f\colon X \to \{-1,+1\}$ for some finite set X, where -1 corresponds to "true." Note that this is a departure from the introduction, where we used the more traditional range $\{0,1\}$. For $x \in \{0,1\}^n$, we define $|x| = x_1 + x_2 + \cdots + x_n$. The symbol P_d stands for the set of all univariate real polynomials of degree at most d. For a given function $f\colon \{0,1\}^n \to \mathbb{R}$ and a string $z \in \{0,1\}^n$, we let f_z stand for the function $f_z\colon \{0,1\}^n \to \mathbb{R}$ given by $f_z(x) \equiv f(x \oplus z)$. For $b \in \{0,1\}$, we use the notation $\overline{b} = 1 - b = 1 \oplus b$. The *characteristic vector* of a set $S \subseteq \{1,\ldots,n\}$ is the string $\mathbf{1}_S \in \{0,1\}^n$ such that $(\mathbf{1}_S)_i = 1$ for $i \in S$, and $(\mathbf{1}_S)_i = 0$ otherwise. For a string $x \in \{0,1\}^n$ and a set $S \subseteq \{1,\ldots,n\}$, we define $x|_S = (x_{i_1}, x_{i_2}, \ldots, x_{i_{|S|}}) \in \{0,1\}^{|S|}$, where $i_1 < i_2 < \cdots < i_{|S|}$ are the elements of S.

2.1 Matrices

The symbol $\mathbb{R}^{m\times n}$ refers to the family of all $m \times n$ matrices with real entries. We specify a matrix by its generic entry, e.g., the notation $A = [F(i,j)]_{i,j}$ means that the (i,j)th entry of A is given by the expression $F(i,j)$. In most matrices that arise in this work, the exact ordering of the columns (and rows) is irrelevant. In such cases we describe a matrix by the notation $[F(i,j)]_{i\in I,\, j\in J}$, where I and J are some index sets.

Let $A = [A_{ij}] \in \mathbb{R}^{m\times n}$ be given. We adopt the shorthands $\|A\|_\infty = \max|A_{ij}|$ and $\|A\|_1 = \sum|A_{ij}|$. We denote the singular values of A by $\sigma_1(A) \geqslant \sigma_2(A) \geqslant \cdots \geqslant \sigma_{\min\{m,n\}}(A) \geqslant 0$. Recall that the spectral norm of A is given by

$$\|A\| = \max_{x\in\mathbb{R}^n,\ \|x\|_2=1} \|Ax\|_2 = \sigma_1(A),$$

where $\|\cdot\|_2$ is the Euclidean vector norm. For $A, B \in \mathbb{R}^{m\times n}$, we write $\langle A, B\rangle = \sum A_{ij}B_{ij}$. We denote the rank of A over the reals by $\operatorname{rk} A$.

We will need the following formulation of linear programming duality in matrix notation.

Theorem 2.1 (Duality) *For $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, the system $Ax \geqslant b$ has no solution in $x \in \mathbb{R}^n$ if and only if there is a vector $y \in [0,\infty)^m$ such that $y^{\mathsf{T}}A = 0$ but $y^{\mathsf{T}}b > 0$.*

The monograph by Schrijver [27, Chap. 7] provides detailed background on Theorem 2.1 and various other formulations of linear programming duality, along with historical notes.

2.2 Fourier Transform

Consider the vector space of real functions on $\{0,1\}^n$, equipped with the inner product

$$\langle f, g\rangle = 2^{-n} \sum_{x\in\{0,1\}^n} f(x)g(x)$$

and normed by

$$\|f\|_\infty = \max_{x\in\{0,1\}^n} |f(x)|.$$

For $S \subseteq \{1,\ldots,n\}$, define $\chi_S\colon \{0,1\}^n \to \{-1,+1\}$ by $\chi_S(x) = (-1)^{\sum_{i\in S} x_i}$. Then the functions χ_S, $S \subseteq \{1,\ldots,n\}$, form an orthonormal basis for the inner product space in question.

As a result, every function $f\colon \{0,1\}^n \to \mathbb{R}$ has a unique representation of the form

$$f = \sum_{S \subseteq \{1,\ldots,n\}} \hat{f}(S)\, \chi_S,$$

where $\hat{f}(S) = \langle f, \chi_S \rangle$ is the *Fourier coefficient* of f that corresponds to the character χ_S. The following bound is immediate from the definition of Fourier coefficients:

$$\max_{S \subseteq \{1,\ldots,n\}} |\hat{f}(S)| \leqslant 2^{-n} \sum_{x \in \{0,1\}^n} |f(x)|. \tag{2.1}$$

2.3 Monomial Count, Sensitivity, and Decision Trees

Every function $f\colon \{0,1\}^n \to \mathbb{R}$ has a unique representation of the form

$$f(x) = \sum_{S \subseteq \{1,\ldots,n\}} \alpha_S \prod_{i \in S} x_i$$

for some reals α_S. We define the *degree* of f by $\deg(f) = \max\{|S| : \alpha_S \neq 0\}$ and the *number of monomials* in f by $\mathrm{mon}(f) = |\{S : \alpha_S \neq 0\}|$.

For $i = 1, 2, \ldots, n$, we let $e_i \in \{0,1\}^n$ stand for the vector with 1 in the ith component and zeroes everywhere else. For a set $S \subseteq \{1,\ldots,n\}$, we define $e_S \in \{0,1\}^n$ by $e_S = \sum_{i \in S} e_i$. In particular, $e_\varnothing = 0$. Fix a Boolean function $f\colon \{0,1\}^n \to \{-1,+1\}$. For $\ell = 1, 2, \ldots, n$, the *ℓ-block sensitivity* of f, denoted $\mathrm{bs}_\ell(f)$, is defined as the largest k for which there exist nonempty disjoint sets $S_1, \ldots, S_k \subseteq \{1,\ldots,n\}$, each containing no more than ℓ elements, such that

$$f(z \oplus e_{S_1}) = f(z \oplus e_{S_2}) = \cdots = f(z \oplus e_{S_k}) \neq f(z)$$

for some $z \in \{0,1\}^n$. One distinguishes two extremal cases. The *sensitivity* of f, denoted $\mathrm{s}(f)$, is defined by $\mathrm{s}(f) = \mathrm{bs}_1(f)$. The *block sensitivity* of f, denoted $\mathrm{bs}(f)$, is defined by $\mathrm{bs}(f) = \mathrm{bs}_n(f)$. In this context, the term *block* simply refers to a subset $S \subseteq \{1,\ldots,n\}$. We say that a block $S \subseteq \{1,\ldots,n\}$ is *sensitive* for f on input z if $f(z) \neq f(z \oplus e_S)$.

Following Buhrman and de Wolf [8], we define one additional variant of sensitivity. The *zero block sensitivity* of f, denoted $\mathrm{zbs}(f)$, is the largest k for which there exist nonempty disjoint sets $S_1, \ldots, S_k \subseteq \{1,\ldots,n\}$ such that

$$f(z \oplus e_{S_1}) = f(z \oplus e_{S_2}) = \cdots = f(z \oplus e_{S_k}) \neq f(z)$$

for some $z \in \{0,1\}^n$ with $z|_{S_1 \cup \cdots \cup S_k} = (0, 0, \ldots, 0)$.

For a function $f\colon \{0,1\}^n \to \{-1,+1\}$, we let $\mathrm{dt}(f)$ stand for the least depth of a decision tree for f. The following inequalities are known.

Theorem 2.2 (Beals et al. [4, §5]) *Every function $f\colon \{0,1\}^n \to \{-1,+1\}$ satisfies*

$$\mathrm{dt}(f) \leqslant \mathrm{bs}(f)^3.$$

Theorem 2.3 (Midrijanis [20]) *Every function $f\colon \{0,1\}^n \to \{-1,+1\}$ satisfies*

$$\mathrm{dt}(f) \leqslant O(\deg(f)^3).$$

For further background on these combinatorial complexity measures, we refer the reader to the excellent survey by Buhrman and de Wolf [9].

2.4 Symmetric Functions

Let S_n denote the symmetric group on n elements. For $\sigma \in S_n$ and $x \in \{0,1\}^n$, we denote by σx the string $(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \in \{0,1\}^n$. A function $\phi\colon \{0,1\}^n \to \mathbb{R}$ is called *symmetric* if $\phi(x) = \phi(\sigma x)$ for every $x \in \{0,1\}^n$ and every $\sigma \in S_n$. Equivalently, ϕ is symmetric if $\phi(x)$ is uniquely determined by $|x|$. Observe that for every $\phi\colon \{0,1\}^n \to \mathbb{R}$ (symmetric or not), the derived function

$$\phi_{\mathrm{sym}}(x) = \mathop{\mathbf{E}}_{\sigma \in S_n}[\phi(\sigma x)]$$

is symmetric. Symmetric functions on $\{0,1\}^n$ are intimately related to univariate polynomials, as demonstrated by Minsky and Papert's *symmetrization argument* [21]:

Proposition 2.4 (Minsky and Papert) *Let $\phi\colon \{0,1\}^n \to \mathbb{R}$ be given such that $\hat{\phi}(S) = 0$ for $|S| > r$. Then there is a polynomial $p \in P_r$ with*

$$\mathop{\mathbf{E}}_{\sigma \in S_n}[\phi(\sigma x)] = p(|x|), \qquad x \in \{0,1\}^n.$$

2.5 Pattern Matrices

Pattern matrices, introduced in [29, 30], play an important role in this paper. Relevant definitions and results from [30] follow.

Let n and N be positive integers with $n \mid N$. Split $[N]$ into n contiguous blocks, with N/n elements each:

$$[N] = \left\{1, 2, \ldots, \frac{N}{n}\right\} \cup \left\{\frac{N}{n} + 1, \ldots, \frac{2N}{n}\right\} \cup \cdots \cup \left\{\frac{(n-1)N}{n} + 1, \ldots, N\right\}.$$

Let $\mathcal{V}(N,n)$ denote the family of subsets $V \subseteq \{1, \ldots, N\}$ that have exactly one element from each of these blocks (in particular, $|V| = n$). Clearly, $|\mathcal{V}(N,n)| = (N/n)^n$.

Definition 2.5 (Pattern matrix) For $\phi\colon \{0,1\}^n \to \mathbb{R}$, the (N,n,ϕ)-*pattern matrix* is the real matrix A given by

$$A = \Big[\phi(x|_V \oplus w)\Big]_{x \in \{0,1\}^N,\, (V,w) \in \mathcal{V}(N,n) \times \{0,1\}^n}.$$

In words, A is the matrix of size 2^N by $(N/n)^n 2^n$ whose rows are indexed by strings $x \in \{0,1\}^N$, whose columns are indexed by pairs $(V,w) \in \mathcal{V}(N,n) \times \{0,1\}^n$, and whose entries are given by $A_{x,(V,w)} = \phi(x|_V \oplus w)$.

The logic behind the term "pattern matrix" is as follows: a mosaic arises from repetitions of a pattern in the same way that A arises from applications of ϕ to various subsets of the variables. We are going to need the following expression for the spectral norm of a pattern matrix [30, Thm. 4.3].

Theorem 2.6 (Sherstov) *Let $\phi\colon \{0,1\}^n \to \mathbb{R}$ be given. Let A be the (N,n,ϕ)-pattern matrix. Then*

$$\|A\| \;=\; \sqrt{2^{N+n}\binom{N}{n}^n} \max_{S \subseteq \{1,\ldots,n\}} \left\{|\hat{\phi}(S)| \left(\frac{n}{N}\right)^{|S|/2}\right\}.$$

By identifying a set $S \subseteq \{1, 2, \ldots, N\}$ with its characteristic vector $\mathbf{1}_S \in \{0,1\}^N$, we may alternately regard $\mathcal{V}(N,n)$ as a family of strings in $\{0,1\}^N$ rather than as a family of sets. This view will be useful in the proof of Theorem 4.2 below. Detailed background on the pattern matrix method is available in the survey article [28].

2.6 Communication Complexity

This section reviews the quantum model of communication complexity. We include this review mainly for completeness; our proofs rely solely on a standard matrix-analytic property of quantum protocols and on no other aspect of quantum communication.

There are several equivalent ways to describe a quantum communication protocol, e.g., [5, 33, 25]. Our description closely follows Razborov [25]. Let $\mathcal{A}$ and $\mathcal{B}$ be complex finite-dimensional Hilbert spaces. Let $\mathcal{C}$ be a Hilbert space of dimension 2, whose orthonormal basis we denote by $|0\rangle$, $|1\rangle$. Consider the tensor product $\mathcal{A} \otimes \mathcal{C} \otimes \mathcal{B}$, which is itself a Hilbert space with an inner product inherited from $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$. The *state* of a quantum system is a unit vector in $\mathcal{A} \otimes \mathcal{C} \otimes \mathcal{B}$, and conversely any such unit vector corresponds to a distinct quantum state. The quantum system starts in a given state and traverses a sequence of states, each obtained from the previous one via a unitary transformation chosen according to the protocol. Formally, a *quantum communication protocol* is a finite sequence of unitary transformations

$$U_1 \otimes I_{\mathcal{B}}, \quad I_{\mathcal{A}} \otimes U_2, \quad U_3 \otimes I_{\mathcal{B}}, \quad I_{\mathcal{A}} \otimes U_4, \quad \ldots, \quad U_{2k-1} \otimes I_{\mathcal{B}}, \quad I_{\mathcal{A}} \otimes U_{2k},$$

where: $I_{\mathcal{A}}$ and $I_{\mathcal{B}}$ are the identity transformations in $\mathcal{A}$ and $\mathcal{B}$, respectively; $U_1, U_3, \ldots, U_{2k-1}$ are unitary transformations in $\mathcal{A} \otimes \mathcal{C}$; and $U_2, U_4, \ldots, U_{2k}$ are unitary transformations in $\mathcal{C} \otimes \mathcal{B}$. The *cost* of the protocol is the length of this sequence, namely, $2k$. On Alice's input $x \in X$ and Bob's input $y \in Y$ (where X, Y are given finite sets), the computation proceeds as follows.

1. The quantum system starts out in an initial state $\mathsf{Initial}(x, y)$.
2. Through successive applications of the above unitary transformations, the system reaches the state
$$\mathsf{Final}(x,y) = (I_{\mathcal{A}} \otimes U_{2k})(U_{2k-1} \otimes I_{\mathcal{B}}) \cdots (I_{\mathcal{A}} \otimes U_2)(U_1 \otimes I_{\mathcal{B}})\, \mathsf{Initial}(x,y).$$
3. Let v denote the projection of $\mathsf{Final}(x, y)$ onto $\mathcal{A} \otimes \mathrm{span}(|1\rangle) \otimes \mathcal{B}$. The output of the protocol is -1 with probability $\langle v, v \rangle$, and $+1$ with the complementary probability $1 - \langle v, v \rangle$.

All that remains is to specify how the initial state $\mathsf{Initial}(x, y) \in \mathcal{A} \otimes \mathcal{C} \otimes \mathcal{B}$ is constructed from x, y. It is here that the model with prior entanglement differs from the model without prior entanglement. In the model without prior entanglement, $\mathcal{A}$ and $\mathcal{B}$ have orthonormal bases $\{|x, w\rangle : x \in X,\ w \in W\}$ and $\{|y, w\rangle : y \in Y,\ w \in W\}$, respectively, where W is a finite set corresponding to the private workspace of each of the parties. The initial state is the pure state

$$\mathsf{Initial}(x,y) = |x, 0\rangle\, |0\rangle\, |y, 0\rangle,$$

where $0 \in W$ is a certain fixed element. In the model with prior entanglement, the spaces $\mathcal{A}$ and $\mathcal{B}$ have orthonormal bases $\{|x, w, e\rangle : x \in X,\ w \in W,\ e \in E\}$ and $\{|y, w, e\rangle : y \in Y,\ w \in$

$W,\ e \in E\}$, respectively, where W is as before and E is a finite set corresponding to the prior entanglement. The initial state is now the entangled state

$$\mathsf{Initial}(x,y) = \frac{1}{\sqrt{|E|}} \sum_{e \in E} |x,0,e\rangle\,|0\rangle\,|y,0,e\rangle.$$

Apart from finite size, no assumptions are made about W or E. In particular, the model with prior entanglement allows for an unlimited supply of entangled qubits. This mirrors the unlimited supply of shared random bits in the classical public-coin randomized model.

Let $f\colon X \times Y \to \{-1,+1\}$ be a given function. A quantum protocol P is said to compute f with error ϵ if

$$\mathbf{P}[P(x,y) \neq f(x,y)] \leqslant \epsilon$$

for all x, y, where the random variable $P(x,y) \in \{-1,+1\}$ is the output of the protocol on input (x,y). Let $Q_\epsilon(f)$ denote the least cost of a quantum protocol without prior entanglement that computes f with error ϵ. Define $Q^*_\epsilon(f)$ analogously for protocols with prior entanglement. The precise choice of a constant $\epsilon \in (0,1/2)$ affects $Q_\epsilon(f)$ and $Q^*_\epsilon(f)$ by at most a constant factor, and thus the setting $\epsilon = 1/3$ entails no loss of generality. By the communication complexity of a Boolean matrix $F = [F_{ij}]_{i \in I,\, j \in J}$ will be meant the communication complexity of the associated function $f\colon I \times J \to \{-1,+1\}$ given by $f(i,j) = F_{ij}$.

A useful technique for proving lower bounds on quantum communication complexity, regardless of prior entanglement, is the *generalized discrepancy method,* originally applied by Klauck [16] and reformulated more broadly by Razborov [25]. The following is an adaptation by the author [30, Sec. 2.4].

Theorem 2.7 (Generalized discrepancy method) *Fix finite sets X, Y and a given function $f\colon X \times Y \to \{-1,+1\}$. Let $\Psi = [\Psi_{xy}]_{x \in X,\, y \in Y}$ be any real matrix with $\|\Psi\|_1 = 1$. Then for each $\epsilon > 0$,*

$$4^{Q_\epsilon(f)} \geqslant 4^{Q^*_\epsilon(f)} \geqslant \frac{\langle \Psi, F\rangle - 2\epsilon}{3\,\|\Psi\|\sqrt{|X|\,|Y|}},$$

where $F = [f(x,y)]_{x \in X,\, y \in Y}$.

Apart from quantum communication, we will consider two classical models. For a function $f\colon X \times Y \to \{-1,+1\}$, we let $D(f)$ stand for the *deterministic* communication complexity of f. We let $R_{1/3}(f)$ stand for the public-coin randomized communication complexity of f, with error probability at most $1/3$. The following result of Mehlhorn and Schmidt [19] gives a powerful technique for proving lower bounds on deterministic communication.

Theorem 2.8 (Mehlhorn and Schmidt) *Let $f\colon X \times Y \to \{-1,+1\}$ be a given function, where X, Y are finite sets. Put $F = [f(x,y)]_{x \in X,\, y \in Y}$. Then*

$$D(f) \geqslant \log_2 \operatorname{rk} F.$$

An excellent reference on classical communication complexity is the monograph by Kushilevitz and Nisan [17].

3 Combinatorial Ingredients

In this section, we develop the combinatorial component of our solution. We start by recalling an elegant result, due to Kenyon and Kutin [15, Cor. 3.1], that the sensitivity and ℓ-block sensitivity of a Boolean function are polynomially related for all constant ℓ. For the purposes of this paper, the case $\ell = 2$ is all that is needed.

Theorem 3.1 (Kenyon and Kutin) *Let $f : \{0,1\}^n \to \{-1,+1\}$ be given. Then*

$$\mathrm{s}(f) \geqslant \alpha\sqrt{\mathrm{bs}_2(f)}$$

for some absolute constant $\alpha > 0$.

Remark 3.2 The lower bound in Theorem 3.1 is asymptotically tight, by a construction due to Rubinstein [26].

For our purposes, the key consequence of Kenyon and Kutin's result is the following lemma.

Lemma 3.3 *Let $f\colon \{0,1\}^n \to \{-1,+1\}$ be a given function. Then there exists $g\colon \{0,1\}^n \to \{-1,+1\}$ such that*

$$\mathrm{s}(g) \geqslant \alpha\sqrt{\mathrm{bs}(f)} \tag{3.1}$$

for some absolute constant $\alpha > 0$ and

$$g(x) \equiv f(x_{i_1}, x_{i_2}, \ldots, x_{i_n}) \tag{3.2}$$

for some $i_1, i_2, \ldots, i_n \in \{1, 2, \ldots, n\}$.

Proof. Put $k = \mathrm{bs}(f)$ and fix disjoint sets $S_1, \ldots, S_k \subseteq \{1, \ldots, n\}$ such that one has $f(z \oplus e_{S_1}) = f(z \oplus e_{S_2}) = \cdots = f(z \oplus e_{S_k}) \neq f(z)$ for some $z \in \{0,1\}^n$. Let I be the set of all indices i such the string $z|_{S_i}$ features both zeroes and ones. Put $|I| = r$. For convenience of notation, we will assume that $I = \{1, 2, \ldots, r\}$. For $i = 1, 2, \ldots, r$, form the partition $S_i = A_i \cup B_i$, where

$$A_i = \{j \in S_i : z_j = 0\}, \qquad B_i = \{j \in S_i : z_j = 1\}.$$

Now let

$$g(x) = f\left(\bigoplus_{i=1}^{r} x_{\min A_i} e_{A_i} \oplus \bigoplus_{i=1}^{r} x_{\min B_i} e_{B_i} \oplus \bigoplus_{i=r+1}^{k} x_{\min S_i} e_{S_i} \oplus \bigoplus_{i \notin S_1 \cup \cdots \cup S_k} x_i e_i\right).$$

Then (3.2) is immediate. By the properties of f, we have $\mathrm{bs}_2(g) \geqslant k$, with the blocks $\{\min A_1, \min B_1\}, \ldots, \{\min A_r, \min B_r\}$ and $\{\min S_{r+1}\}, \ldots, \{\min S_k\}$ being sensitive for g on input $x = z$. As a result, Theorem 3.1 implies (3.1). $\square$

4 Analytic Ingredients

We now turn to the analytic component of our solution. The main results of this section can all be derived by modifying Razborov's proof of the quantum lower bound for the disjointness

function [25]. The alternate derivation presented here has some advantages, as we discuss in Remark 4.3. We start by exhibiting a large family of Boolean functions whose inapproximability by low-degree polynomials in the uniform norm can be witnessed by a single, common dual object.

Theorem 4.1 *Let $\mathcal{F}$ denote the set of all functions $f\colon \{0,1\}^n \to \{-1,+1\}$ such that $f(e_1) = f(e_2) = \cdots = f(e_n) \neq f(0) = 1$. Let $\delta > 0$ be a sufficiently small absolute constant. Then there exists a function $\psi\colon \{0,1\}^n \to \mathbb{R}$ such that:*

$$\hat{\psi}(S) = 0, \qquad |S| < \delta\sqrt{n}, \tag{4.1}$$

$$\sum_{x\in\{0,1\}^n} |\psi(x)| = 1, \tag{4.2}$$

$$\sum_{x\in\{0,1\}^n} \psi(x)f(x) > \frac{1}{3}, \qquad f \in \mathcal{F}. \tag{4.3}$$

Proof. Let p be a univariate real polynomial that satisfies

$$\begin{aligned} &p(0) \in [2/3, 4/3], \\ &p(1) \in [-4/3, -2/3], \\ &p(i) \in [-4/3, 4/3], \qquad\qquad i = 2, 3, \ldots, n. \end{aligned}$$

It follows from basic approximation theory (viz., the inequalities due to A. A. Markov and S. N. Bernstein) that any such polynomial p has degree at least $\delta\sqrt{n}$ for an absolute constant $\delta > 0$. See Nisan and Szegedy [23], pp. 308–309, for a short derivation.

By the symmetrization argument (Proposition 2.4), there does not exist a multivariate polynomial $\phi(x_1, \ldots, x_n)$ of degree less than $\delta\sqrt{n}$ such that

$$\begin{aligned} &\phi(0) \in [2/3, 4/3], \\ &\phi(e_i) \in [-4/3, -2/3], \qquad\qquad i = 1, 2, \ldots, n, \\ &\phi(x) \in [-4/3, 4/3], \qquad\qquad x \in \{0,1\}^n \setminus \{0, e_1, e_2, \ldots, e_n\}. \end{aligned}$$

Equivalently, the following system of linear constraints has no solution in the reals α_S:

$$\begin{aligned} &\sum_{|S|<\delta\sqrt{n}} \alpha_S \chi_S(0) \in [2/3, 4/3], \\ &\sum_{|S|<\delta\sqrt{n}} \alpha_S \chi_S(e_i) \in [-4/3, -2/3], \qquad\qquad i = 1, 2, \ldots, n, \\ &\sum_{|S|<\delta\sqrt{n}} \alpha_S \chi_S(x) \in [-4/3, 4/3], \qquad\qquad x \in \{0,1\}^n \setminus \{0, e_1, e_2, \ldots, e_n\}. \end{aligned}$$

The duality of linear programming (Theorem 2.1) now implies the existence of ψ that obeys (4.1), (4.2), and additionally satisfies

$$\psi(0) - \sum_{i=1}^{n} \psi(e_i) - \sum_{\substack{x\in\{0,1\}^n \\ |x|\geqslant 2}} |\psi(x)| > \frac{1}{3},$$

which forces (4.3). $\square$

We are now in a position to prove our main technical criterion for high quantum communication complexity. Our proof is based on the pattern matrix method [29, 30]. The novelty of the development below resides in allowing the rows of the given Boolean matrix to derive from distinct Boolean functions, which considerably disrupts the spectral structure. We are able to force the same quantitative conclusion by using the fact that these Boolean functions, albeit distinct, share the relevant dual object.

Theorem 4.2 *Let $g\colon \{0,1\}^n \to \{-1,+1\}$ be a function such that $g(z \oplus e_1) = g(z \oplus e_2) = \cdots = g(z \oplus e_k) \neq g(z)$ for some $z \in \{0,1\}^n$ with $z_1 = \cdots = z_k = 0$. Then the matrix $G = [g(x \wedge y)]_{x,y \in \{0,1\}^n}$ satisfies*

$$Q^*_{1/3}(G) \geqslant \Omega(\sqrt{k}).$$

Remark 4.3 As formulated above, Theorem 4.2 can be derived by modifying Razborov's proof of the $\Omega(\sqrt{n})$ quantum lower bound for the disjointness function [25, §5.3]. The derivation that we are about to give offers some advantages. First, it is simpler and in particular does not require tools such as Hahn matrices in [25]. Second, it generalizes to any family $\mathcal{F}$ of functions with a common dual polynomial, whereas the method in [25] is restricted to symmetrizable families.

Proof (of Theorem 4.2) Without loss of generality, we may assume that k is divisible by 4. Let $\mathcal{F}$ denote the system of all functions $f\colon \{0,1\}^{k/4} \to \{-1,+1\}$ such that $f(e_1) = f(e_2) = \cdots = f(e_{k/4}) \neq f(0) = 1$. By Theorem 4.1, there exists $\psi\colon \{0,1\}^{k/4} \to \mathbb{R}$ such that

$$\hat{\psi}(S) = 0, \qquad |S| < \delta\sqrt{k}, \tag{4.4}$$

$$\sum_{x \in \{0,1\}^{k/4}} |\psi(x)| = 1, \tag{4.5}$$

$$\sum_{x \in \{0,1\}^{k/4}} \psi(x) f(x) > \frac{1}{3}, \qquad f \in \mathcal{F}, \tag{4.6}$$

where $\delta > 0$ is an absolute constant. Now, let Ψ be the $(k/2, k/4, 2^{-3k/4}\psi)$-pattern matrix. It follows from (4.5) that

$$\|\Psi\|_1 = 1. \tag{4.7}$$

By (2.1) and (4.5),

$$\max_S |\hat{\psi}(S)| \leqslant 2^{-k/4}. \tag{4.8}$$

In view of (4.4) and (4.8), Theorem 2.6 yields

$$\|\Psi\| \leqslant 2^{-\delta\sqrt{k}/2}\, 2^{-k/2}. \tag{4.9}$$

Now, put

$$M = g(z) \left[g\left(z \oplus \bigoplus_{i=1}^{k/2} \{x_i y_{2i-1} e_{2i-1} \oplus \overline{x_i} y_{2i} e_{2i}\} \right) \right]_{x \in \{0,1\}^{k/2},\, y \in \mathcal{V}(k,k/4)},$$

where we identify each $y \in \mathcal{V}(k, k/4)$ in the natural way with a string in $\{0,1\}^k$. Observe that

$$M = \Big[f_{V,w}(x|_V \oplus w) \Big]_{x \in \{0,1\}^{k/2}, (V,w) \in \mathcal{V}(k/2,k/4) \times \{0,1\}^{k/4}}$$

for some functions $f_{V,w} \in \mathcal{F}$. This representation makes it clear, in view of (4.6), that

$$\langle \Psi, M \rangle > \frac{1}{3}. \tag{4.10}$$

By (4.7), (4.9), (4.10) and the generalized discrepancy method (Theorem 2.7), we have $Q^*_{1/10}(M) \geqslant \Omega(\sqrt{k})$. It remains to note that M is a submatrix of $g(z)G$, so that $Q^*_{1/10}(G) \geqslant Q^*_{1/10}(M)$. $\square$

We will also need the following equivalent formulation of Theorem 4.2, for disjunctions instead of conjunctions.

Corollary 4.4 *Let $g\colon \{0,1\}^n \to \{-1,+1\}$ be a function such that $g(z \oplus e_1) = g(z \oplus e_2) = \cdots = g(z \oplus e_k) \neq g(z)$ for some $z \in \{0,1\}^n$ with $z_1 = \cdots = z_k = 1$. Then the matrix $G = [g(x \vee y)]_{x,y \in \{0,1\}^n}$ satisfies*

$$Q^*_{1/3}(G) \geqslant \Omega(\sqrt{k}).$$

Proof. Put $\tilde{g} = g_{(1,\ldots,1)}$ and $\tilde{z} = (1,\ldots,1) \oplus z$. Then $\tilde{z}_1 = \cdots = \tilde{z}_k = 0$ and $\tilde{g}(\tilde{z} \oplus e_1) = \tilde{g}(\tilde{z} \oplus e_2) = \cdots = \tilde{g}(\tilde{z} \oplus e_k) \neq \tilde{g}(\tilde{z})$. By Theorem 4.2, the matrix $\tilde{G} = [\tilde{g}(x \wedge y)]_{x,y \in \{0,1\}^n}$ satisfies $Q^*_{1/3}(\tilde{G}) \geqslant \Omega(\sqrt{k})$. It remains to note that G and $\tilde{G}$ are identical, up to permutations of rows and columns. $\square$

We point out another simple corollary to Theorem 4.2.

Corollary 4.5 *Let $f\colon \{0,1\}^n \to \{-1,+1\}$ be given. Then for some $z \in \{0,1\}^n$, the matrix $F = [f_z(x \wedge y)]_{x,y} = [f(\ldots, (x_i \wedge y_i) \oplus z_i, \ldots)]_{x,y}$ obeys*

$$Q^*_{1/3}(F) = \Omega(\sqrt{\mathrm{bs}(f)}).$$

Proof. Put $k = \mathrm{bs}(f)$ and fix $z \in \{0,1\}^n$ such that $\mathrm{zbs}(f_z) = k$. By an argument analogous to Lemma 3.3, one obtains a function $g\colon \{0,1\}^n \to \{-1,+1\}$ such that $g(e_1) = g(e_2) = \cdots = g(e_k) \neq g(0)$ and $g(x) \equiv f_z(\xi_1, \xi_2, \ldots, \xi_n)$ for some symbols $\xi_1, \xi_2, \ldots, \xi_n \in \{x_1, x_2, \ldots, x_n, 0, 1\}$. Then Theorem 4.2 implies that the matrix $G = [g(x \wedge y)]_{x,y \in \{0,1\}^n}$ satisfies $Q^*_{1/3}(G) \geqslant \Omega(\sqrt{k})$. On the other hand, $Q^*_{1/3}(F) \geqslant Q^*_{1/3}(G)$ by construction. $\square$

5 Quantum-Classical Equivalence

We now combine the combinatorial and analytic development of the previous sections to obtain our main results. We start by proving relevant lower bounds against quantum protocols.

Theorem 5.1 *Let $f\colon \{0,1\}^n \to \{-1,+1\}$ be given. Put $F_1 = [f(x \wedge y)]_{x,y}$ and $F_2 = [f(x \vee y)]_{x,y}$, where the row and column indices range over $\{0,1\}^n$. Then*

$$\max\{Q^*_{1/3}(F_1), Q^*_{1/3}(F_2)\} = \Omega(\mathrm{bs}(f)^{1/4}).$$

Proof. By Lemma 3.3, there exists a function $g\colon \{0,1\}^n \to \{-1,+1\}$ such that

$$\mathrm{s}(g) \geqslant \Omega(\sqrt{\mathrm{bs}(f)}) \tag{5.1}$$

and

$$g(x) \equiv f(x_{i_1}, x_{i_2}, \ldots, x_{i_n}) \tag{5.2}$$

for some $i_1, i_2, \ldots, i_n \in \{1, 2, \ldots, n\}$. By renumbering the variables if necessary, we see that at least one of the following statements must hold:

(1) $g(z \oplus e_1) = g(z \oplus e_2) = \cdots = g(z \oplus e_{\lceil \mathrm{s}(g)/2 \rceil}) \neq g(z)$ for some $z \in \{0,1\}^n$ with $z_1 = z_2 = \cdots = z_{\lceil \mathrm{s}(g)/2 \rceil} = 0$;

(2) $g(z \oplus e_1) = g(z \oplus e_2) = \cdots = g(z \oplus e_{\lceil \mathrm{s}(g)/2 \rceil}) \neq g(z)$ for some $z \in \{0,1\}^n$ with $z_1 = z_2 = \cdots = z_{\lceil \mathrm{s}(g)/2 \rceil} = 1$.

In the former case, Theorem 4.2 implies that the matrix $G_1 = [g(x \wedge y)]_{x,y \in \{0,1\}^n}$ satisfies $Q^*_{1/3}(G_1) \geqslant \Omega(\sqrt{\mathrm{s}(g)})$, whence $Q^*_{1/3}(F_1) \geqslant Q^*_{1/3}(G_1) \geqslant \Omega(\mathrm{bs}(f)^{1/4})$ in view of (5.1) and (5.2).

In the latter case, Corollary 4.4 implies that $G_2 = [g(x \vee y)]_{x,y \in \{0,1\}^n}$ satisfies $Q^*_{1/3}(G_2) \geqslant \Omega(\sqrt{\mathrm{s}(g)})$, whence $Q^*_{1/3}(F_2) \geqslant Q^*_{1/3}(G_2) \geqslant \Omega(\mathrm{bs}(f)^{1/4})$ in view of (5.1) and (5.2). □

Having obtained the desired lower bounds on quantum communication, we now turn to classical protocols. The bound that we seek here follows easily from the work of Buhrman et al. [7] and Beals et al. [4]. Related observations have been used in a number of recent papers in the area [25, 30, 32].

Theorem 5.2 (Classical upper bound; cf. [7, 4]) *Let $f\colon \{0,1\}^n \to \{-1,+1\}$ be given. Put $F_1 = [f(x \wedge y)]_{x,y}$ and $F_2 = [f(x \vee y)]_{x,y}$, where the row and column indices range over $\{0,1\}^n$. Then*

$$\max\{D(F_1), D(F_2)\} \leqslant 2\,\mathrm{dt}(f) \leqslant 2\,\mathrm{bs}(f)^3.$$

Proof (adapted from [7, 4]). The second inequality follows immediately by Theorem 2.2, so we will focus on the first. Fix an optimal-depth decision tree for f. The protocol for F_1 is as follows. On input x and y, Alice and Bob start at the top node of the tree, read its label i, and exchange the two bits x_i and y_i. This allows them to compute $x_i \wedge y_i$ and to determine which branch to take next. The process repeats at the new node and so on, until the parties have reached a leaf node. Since the longest root-to-leaf path has length $\mathrm{dt}(f)$, the claim follows. The proof for F_2 is entirely analogous. □

Theorems 5.1 and 5.2 immediately imply our main result on quantum-classical equivalence, stated above as Theorem 1.1.

6 Masked Problems and the Log-Rank Conjecture

As we showed in the previous section, the communication problem of computing $f(x \wedge y)$ and $f(x \vee y)$ has polynomially related quantum and classical complexities. Here, we will see that this communication problem additionally satisfies the *log-rank conjecture* of Lovász and Saks [18].

The log-rank conjecture states that the deterministic communication complexity of every Boolean matrix F satisfies $D(F) \leqslant (\log_2 \operatorname{rk} F)^c + c$ for some absolute constant $c > 0$. By Theorem 2.8, this is equivalent to saying that $D(F)$ is polynomially related to $\log_2 \operatorname{rk} F$. The development in this section is based on the following result of Buhrman and de Wolf [8], who studied the special case of symmetric functions f in the same context.

Theorem 6.1 (Buhrman and de Wolf) *Let $f\colon \{0,1\}^n \to \mathbb{R}$ be a given function. Put $M = [f(x \wedge y)]_{x,y}$, where the row and column indices range over $\{0,1\}^n$. Then*

$$\operatorname{rk} M = \operatorname{mon}(f).$$

Our first observation is as follows.

Lemma 6.2 *Let $f\colon \{0,1\}^n \to \mathbb{R}$ be a given function, where $f \not\equiv 0$ and $d = \deg(f)$. Then for some $z \in \{0,1\}^n$,*

$$\operatorname{mon}(f_z) \geqslant \left(\frac{3}{2}\right)^d.$$

Proof. The proof is by induction on d. The base case $d = 0$ holds since $f \not\equiv 0$. Assume that the claim holds for all f of degree $d-1$. By renumbering the variables if necessary, we have $f(x) = x_1 p(x_2, \ldots, x_n) + q(x_2, \ldots, x_n)$ for some polynomial p of degree $d-1$. The inductive assumption guarantees the existence of $u \in \{0,1\}^{n-1}$ such that $\operatorname{mon}(p_u) \geqslant (3/2)^{d-1}$. Note that $\operatorname{mon}(f_{(0,u)}) = \operatorname{mon}(p_u) + \operatorname{mon}(q_u)$ and $\operatorname{mon}(f_{(1,u)}) \geqslant \operatorname{mon}(p_u) + |\operatorname{mon}(q_u) - \operatorname{mon}(p_u)|$. Therefore,

$$\max\{\operatorname{mon}(f_{(0,u)}), \operatorname{mon}(f_{(1,u)})\} \geqslant \frac{3}{2} \operatorname{mon}(p_u) \geqslant \left(\frac{3}{2}\right)^d,$$

as desired. □

We will also need the following technical lemma.

Lemma 6.3 *Let $f\colon \{0,1\}^n \to \mathbb{R}$ be given. Fix an index $i = 1, 2, \ldots, n$. Define*

$$\tilde{f}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n).$$

Then

$$\max\{\operatorname{mon}(\tilde{f}),\ \operatorname{mon}(f_{e_i})\} \geqslant \frac{1}{2} \operatorname{mon}(f).$$

Proof. Write

$$f(x) = x_i p(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) + \tilde{f}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

It is clear by inspection that $\operatorname{mon}(f_{e_i}) \geqslant \operatorname{mon}(p)$. Thus, we have $\operatorname{mon}(\tilde{f}) + \operatorname{mon}(f_{e_i}) \geqslant \operatorname{mon}(\tilde{f}) + \operatorname{mon}(p) = \operatorname{mon}(f)$, as desired. □

At last, we arrive at the main result of this section.

Theorem 6.4 *Let $f\colon \{0,1\}^n \to \{-1,+1\}$ be given, $d = \deg(f)$. Put $F_1 = [f(x \wedge y)]_{x,y}$ and $F_2 = [f(x \vee y)]_{x,y}$, where the row and column indices range over $\{0,1\}^n$. Then*

$$\max\{\operatorname{rk} F_1,\, \operatorname{rk} F_2\} \geqslant \left(\frac{3}{2\sqrt{2}}\right)^d \geqslant 1.06^d. \tag{6.1}$$

In particular, the communication problem of computing, on input $x, y \in \{0,1\}^n$, both of the quantities $f(x \wedge y)$ and $f(x \vee y)$, satisfies the log-rank conjecture.

Proof. To see how the last statement follows from the lower bound (6.1), note that $\max\{D(F_1), D(F_2)\} \leqslant 2\,\mathrm{dt}(f)$ by Theorem 5.2 and $\mathrm{dt}(f) \leqslant O(\deg(f)^3)$ by Theorem 2.3. In the remainder of the proof, we focus on (6.1) alone.

We assume that $d \geqslant 1$, the claim being trivial otherwise. By renumbering the variables if necessary, we may write

$$f(x) = \alpha x_1 x_2 \cdots x_d + \sum_{S \neq \{1,\ldots,d\}} \alpha_S \prod_{i \in S} x_i,$$

where $\alpha \neq 0$. Define $g(x_1, \ldots, x_d) = f(x_1, \ldots, x_d, 0, \ldots, 0)$. Then g is a nonzero polynomial of degree d, and Lemma 6.2 yields a vector $z \in \{0,1\}^d$ such that

$$\mathrm{mon}(g_z) \geqslant \left(\frac{3}{2}\right)^d.$$

By renumbering the variables if necessary, we may assume that $z = 0^t 1^{d-t}$. We complete the proof by analyzing the cases $t \leqslant d/2$ and $t > d/2$.

Suppose first that $t \leqslant d/2$. Let $\mathcal{F}$ be the set whose elements are the identity function on $\{0,1\}$ and the constant-one function on $\{0,1\}$. Lemma 6.3 provides functions $\phi_1, \ldots, \phi_t \in \mathcal{F}$ such that the polynomial $h(x_1, \ldots, x_d) = g_{1^d}(\phi_1(x_1), \ldots, \phi_t(x_t), x_{t+1}, \ldots, x_d)$ features at least $2^{-t}\,\mathrm{mon}(g_z) \geqslant (3/\{2\sqrt{2}\})^d$ monomials. By Theorem 6.1, the matrix $H = [h(x \wedge y)]_{x,y \in \{0,1\}^d}$ has rank at least $(3/\{2\sqrt{2}\})^d$. Since H is a submatrix of F_2, the theorem holds in this case.

The case $t > d/2$ is entirely symmetric, with F_1 playing the role of F_2. □

Remark 6.5 By the results of Buhrman and de Wolf [8], Theorem 6.4 alone would suffice to obtain a polynomial relationship between classical and quantum communication complexity in the *exact* model. However, for our main result we need a polynomial relationship in the *bounded-error* model, which requires the full development of Sections 3–5.

7 Results for Composed Functions

Up to this point, we have focused on the communication problem of computing $f(x \wedge y)$ and $f(x \vee y)$. Here we point out that our results on quantum-classical equivalence and the log-rank conjecture immediately apply to a broader class of communication problems. Specifically, we will consider compositions of the form $f(g_1(x^{(1)}, y^{(1)}), \ldots, g_n(x^{(n)}, y^{(n)}))$, where one has a combining function $f\colon \{0,1\}^n \to \{-1,+1\}$ that receives input from intermediate functions $g_i\colon X_i \times Y_i \to \{0,1\}$, $i = 1, 2, \ldots, n$. We will show that under natural assumptions on $g_1, \ldots, g_n$, this composed function will have polynomially related quantum and classical bounded-error complexities and will satisfy the log-rank conjecture. To simplify notation, we will henceforth abbreviate $f(g_1(x^{(1)}, y^{(1)}), \ldots, g_n(x^{(n)}, y^{(n)}))$ to $f(\ldots, g_i(x^{(i)}, y^{(i)}), \ldots)$.

Theorem 7.1 *Let $f\colon \{0,1\}^n \to \{-1,+1\}$ be a given function. Fix functions $g_i\colon X_i \times Y_i \to \{0,1\}$, for $i = 1,2,\ldots,n$. Assume that for each i, the matrix $[g_i(x^{(i)}, y^{(i)})]_{x^{(i)}\in X_i, y^{(i)}\in Y_i}$ contains the following submatrices*

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \tag{7.1}$$

up to permutations of rows and columns. Put $F = [f(\ldots, g_i(x^{(i)}, y^{(i)}), \ldots)]$. Assume that for some constant $\alpha > 0$,

$$Q^*_{1/3}(g_i) \geqslant R_{1/3}(g_i)^\alpha, \qquad i = 1,2,\ldots,n. \tag{7.2}$$

Then for some constant $\beta = \beta(\alpha) > 0$,

$$R_{1/3}(F) \geqslant Q^*_{1/3}(F) \geqslant R_{1/3}(F)^\beta.$$

Proof. Without loss of generality, we may assume that f depends on all of its n inputs (otherwise, disregard any irrelevant inputs from among $g_1,\ldots,g_n$ in the analysis below). In particular, we have

$$Q^*_{1/3}(F) \geqslant Q^*_{1/3}(g_i), \qquad i = 1,2,\ldots,n. \tag{7.3}$$

Since each g_i contains the two-variable functions AND and OR as subfunctions, Corollary 4.5 shows that

$$Q^*_{1/3}(F) \geqslant \Omega(\sqrt{\operatorname{bs}(f)}). \tag{7.4}$$

Letting $d = \operatorname{dt}(f)$, we claim that

$$R_{1/3}(F) \leqslant O(d \log d) \max_{i=1,\ldots,n} \{R_{1/3}(g_i)\}. \tag{7.5}$$

The proof of this bound is closely analogous to that of Theorem 5.2. Namely, Alice and Bob evaluate a depth-d decision tree for f. When a tree node calls for the ith variable, the parties run an optimal randomized protocol for g_i with error probability $\frac{1}{3d}$, which requires at most $O(R_{1/3}(g_i) \log d)$ bits of communication. Since all root-to-leaf paths have length at most d, the final answer will be correct with probability at least $2/3$.

In view of Theorem 2.2, the sought polynomial relationship between $R_{1/3}(F)$ and $Q^*_{1/3}(F)$ follows from (7.2)–(7.5). $\square$

We now record an analogous result for the log-rank conjecture.

Theorem 7.2 *Let $f\colon \{0,1\}^n \to \{-1,+1\}$ be a given function. Fix functions $g_i\colon X_i \times Y_i \to \{0,1\}$, for $i = 1,2,\ldots,n$. Assume that for each i, the matrix $[g_i(x^{(i)}, y^{(i)})]_{x^{(i)}\in X_i, y^{(i)}\in Y_i}$ contains* (7.1) *as submatrices, up to permutations of rows and columns. Assume that for some constant $c > 0$,*

$$D(g_i) \leqslant (\log_2 \operatorname{rk} G_i)^c + c, \qquad i = 1,2,\ldots,n, \tag{7.6}$$

where $G_i = [(-1)^{g_i(x^{(i)},y^{(i)})}]_{x^{(i)}\in X_i, y^{(i)}\in Y_i}$. *Then the matrix* $F = [f(\dots, g_i(x^{(i)}, y^{(i)}), \dots)]$ *obeys*

$$D(F) \leqslant (\log_2 \operatorname{rk} F)^C + C$$

for some constant $C = C(c) > 0$. *In particular,* F *satisfies the log-rank conjecture.*

Proof. Without loss of generality, we may assume that f depends on all of its n inputs (otherwise, disregard any irrelevant inputs from among $g_1, \dots, g_n$ in the analysis below). In particular, we have

$$\operatorname{rk} F \geqslant \operatorname{rk} G_i, \qquad i = 1, 2, \dots, n. \tag{7.7}$$

Since each g_i contains the two-variable functions AND and OR as subfunctions, Theorem 6.4 shows that

$$\operatorname{rk} F \geqslant \left(\frac{3}{2\sqrt{2}}\right)^{\deg(f)}. \tag{7.8}$$

Finally, we claim that

$$D(F) \leqslant 2 \operatorname{dt}(f) \max_{i=1,\dots,n} \{D(g_i)\}. \tag{7.9}$$

The proof of this bound is closely analogous to that of Theorem 5.2. Namely, Alice and Bob evaluate an optimal-depth decision tree for f. When a tree node calls for the ith variable, the parties run an optimal deterministic protocol for g_i.

In view of (7.6)–(7.9) and Theorem 2.3, the proof is complete. □

The key property of $g_1, \dots, g_n$ that we have used in this section is that their communication matrices contain (7.1) as submatrices. We close this section by observing that this property almost always holds. More precisely, we show that matrices that do not contain the submatrices (7.1) have a very restricted structure.

Theorem 7.3 *A matrix* $G \in \{0,1\}^{N\times M}$ *does not contain*

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

as a submatrix if and only if $G = 0$, $G = J$, *or*

$$G' \sim \begin{bmatrix} J_1 & & & & \\ & J_2 & & 0 & \\ & & J_3 & & \\ & & & \ddots & \\ & 0 & & & J_k \end{bmatrix}, \tag{7.10}$$

where: G' *is the result of deleting any columns and rows in* G *that consist entirely of zeroes;* $J, J_1, J_2, \dots, J_k$ *are all-*1 *matrices of appropriate dimensions; and* $\sim$ *denotes equality up to permutations of rows and columns.*

Proof. The "if" part is clear. We will prove the other direction by induction on the number of columns, M. The base case is trivial. For the inductive step, let $G \neq 0$ be a given matrix. Let J_1 be a maximal submatrix of G with all entries equal to 1. Then

$$G \sim \begin{bmatrix} J_1 & Z_1 \\ Z_2 & H \end{bmatrix}$$

for suitable matrices Z_1, Z_2, and H, possibly empty. By the maximality of J_1 and the fact that G does not contain A as a submatrix, it follows that either Z_1 is empty or $Z_1 = 0$. Likewise for Z_2. By the inductive hypothesis for H, the proof is complete. $\square$

By reversing the roles of 0 and 1, one obtains from Theorem 7.3 an analogous characterization of all matrices $G = \{0,1\}^{N \times M}$ that do not contain

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

as a submatrix.

Remark 7.4 The communication complexity of a Boolean matrix remains unaffected if one modifies it to retain only one copy of each column, removing any duplicates. An analogous statement holds for the rows. In light of Theorem 7.3, this means that there are only four types of intermediate functions g for which our composition results (Theorem 7.1 and 7.2) fail. These are the functions g with matrix representations

$$I, \qquad \begin{bmatrix} I & \\ & 0 \end{bmatrix}, \qquad \begin{bmatrix} I \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} I & 0 \end{bmatrix}, \tag{7.11}$$

and their negations, where I is the identity matrix. The reason that Theorems 7.1 and 7.2 fail for such g is that the underlying quantum lower bound in terms of block sensitivity of the combining function f is no longer valid. For example, the first matrix type, I, corresponds to letting g be the equality function. Now, the conjunction of n equality functions is still an equality function, and its communication complexity is $O(1)$ both in the randomized and quantum models [17], which is much less than a hypothetical lower bound of $\Omega(\sqrt{n})$ that one would expect from the block sensitivity of $f = \mathrm{AND}_n$. The same $O(1)$ upper bound holds for a conjunction of arbitrarily many functions g of the second, third, and fourth type.

Acknowledgments

The author would like to thank Dima Gavinsky, Adam Klivans, Sasha Razborov, and the anonymous reviewers for their useful comments on a preliminary version of this paper.

References

1. S. Aaronson and A. Ambainis. Quantum search of spatial regions. *Theory of Computing*, 1(1):47–79, 2005.
2. A. Ambainis, L. J. Schulman, A. Ta-Shma, U. V. Vazirani, and A. Wigderson. The quantum communication complexity of sampling. *SIAM J. Comput.*, 32(6):1570–1585, 2003.
3. Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM J. Comput.*, 38(1):366–384, 2008.

4. R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001.
5. H. Buhrman. Quantum computing and communication complexity. *Bulletin of the EATCS*, 70:131–141, 2000.
6. H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum fingerprinting. *Phys. Rev. Lett.*, 87(16), 2001. Article no. 167902.
7. H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and computation. In *Proc. of the 13th Symposium on Theory of Computing (STOC)*, pages 63–68, 1998.
8. H. Buhrman and R. de Wolf. Communication complexity lower bounds by polynomials. In *Proc. of the 16th Conf. on Computational Complexity (CCC)*, pages 120–130, 2001.
9. H. Buhrman and R. de Wolf. Complexity measures and decision tree complexity: A survey. *Theor. Comput. Sci.*, 288(1):21–43, 2002.
10. D. Gavinsky. Classical interaction cannot replace a quantum message. In *Proc. of the 40th Symposium on Theory of Computing (STOC)*, pages 95–102, 2008.
11. D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proc. of the 39th Symposium on Theory of Computing (STOC)*, pages 516–525, 2007.
12. D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf. Bounded-error quantum state identification and exponential separations in communication complexity. In *Proc. of the 38th Symposium on Theory of Computing (STOC)*, pages 594–603, 2006.
13. D. Gavinsky, J. Kempe, and R. de Wolf. Strengths and weaknesses of quantum fingerprinting. In *Proc. of the 21st Conf. on Computational Complexity (CCC)*, pages 288–298, 2006.
14. D. Gavinsky and P. Pudlák. Exponential separation of quantum and classical non-interactive multi-party communication complexity. In *Proc. of the 23rd Conf. on Computational Complexity (CCC)*, pages 332–339, 2008.
15. C. Kenyon and S. Kutin. Sensitivity, block sensitivity, and ℓ-block sensitivity of Boolean functions. *Information and Computation*, 189(1):43–53, 2004.
16. H. Klauck. Lower bounds for quantum communication complexity. *SIAM J. Comput.*, 37(1):20–46, 2007.
17. E. Kushilevitz and N. Nisan. *Communication complexity*. Cambridge University Press, New York, 1997.
18. L. Lovász and M. E. Saks. Lattices, Möbius functions and communication complexity. In *Proc. of the 29th Symposium on Foundations of Computer Science (FOCS)*, pages 81–90, 1988.
19. K. Mehlhorn and E. M. Schmidt. Las Vegas is better than determinism in VLSI and distributed computing. In *Proc. of the 14th Symposium on Theory of Computing (STOC)*, pages 330–337, 1982.
20. G. Midrijanis. Exact quantum query complexity for total Boolean functions. Available at `http://arxiv.org/abs/quant-ph/0403168`, 2004.
21. M. L. Minsky and S. A. Papert. *Perceptrons: An Introduction to Computational Geometry*. MIT Press, Cambridge, Mass., 1969.
22. A. Montanaro and T. Osborne. On the communication complexity of XOR functions. Available at `http://arxiv.org/abs/0909.3392`, 2009.
23. N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4:301–313, 1994.
24. R. Raz. Exponential separation of quantum and classical communication complexity. In *Proc. of the 31st Symposium on Theory of Computing (STOC)*, pages 358–367, 1999.
25. A. A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145–159, 2003.
26. D. Rubinstein. Sensitivity vs. block sensitivity of Boolean functions. *Combinatorica*, 15(2):297–299, 1995.
27. A. Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, Inc., New York, 1998.

28. A. A. Sherstov. Communication lower bounds using dual polynomials. *Bulletin of the EATCS*, 95:59–93, 2008.
29. A. A. Sherstov. Separating AC^0 from depth-2 majority circuits. *SIAM J. Comput.*, 38(6):2113–2129, 2009. Preliminary version in 39th STOC, 2007.
30. A. A. Sherstov. The pattern matrix method. *SIAM J. Comput.*, 2010. To appear. Preliminary version in 40th STOC, 2008.
31. Y. Shi. Approximate polynomial degree of Boolean functions and its applications. In *Proc. of the 4th International Congress of Chinese Mathematicians*, 2007. Available online at `http://www.eecs.umich.edu/~shiyy`.
32. Y. Shi and Y. Zhu. Quantum communication complexity of block-composed functions. *Quantum Information & Computation*, 9(5–6):444–460, 2009.
33. R. de Wolf. *Quantum Computing and Communication Complexity*. PhD thesis, University of Amsterdam, 2001.
34. A. C.-C. Yao. Quantum circuit complexity. In *Proc. of the 34th Symposium on Foundations of Computer Science (FOCS)*, pages 352–361, 1993.