Electronic Colloquium on Computational Complexity, Report No. 110 (2013)

Exponential Quantum-Classical Gaps in Multiparty Nondeterministic Communication Complexity

Xiaoming Sun* Marcos Villagra†

August 12, 2013

Abstract

There are three different types of nondeterminism in quantum communication: i) **NQP**-communication, ii) **QMA**-communication, and iii) **QCMA**-communication. In this paper we show that multiparty **NQP**-communication can be exponentially stronger than **QCMA**-communication. This also implies an exponential separation with respect to classical multiparty nondeterministic communication complexity. We argue that there exists a total function that is hard for **QCMA**-communication and easy for **NQP**-communication. The proof of it involves an application of the pattern tensor method and a new lower bound for polynomial threshold degree. Another important consequence of this result is that nondeterministic rank can be exponentially lower than the discrepancy bound.

Keywords: nondeterministic communication complexity, tensor-rank, norm-bound, pattern-tensor, threshold degree

1 Introduction

1.1 Background

Nondeterministic computation plays a fundamental role in complexity theory. For instance, the **P** vs **NP** problem asks if nondeterministic polynomial-time Turing machines are strictly more powerful than deterministic polynomial-time Turing machines. A nondeterministic Turing machine can be defined as a proof-verifying machine, or, as a probabilistic machine with a possibly large 1-sided error probability. In the former definition, a Yes-instance is accepted if and only if there exists a proof (witness or certificate) that makes the machine to accept, and for every No-instance there is no such proof. In the latter definition, a Yes-instance is accepted with positive probability, and every No-instance is rejected with probability 1.

In classical computation (i.e., models of computation based on classical Turing machines), the two definitions of nondeterminism are equivalent. However, in the quantum world this is not the case. In fact, we have three different definitions: 1) quantum nondeterministic computation which takes on the probabilistic definition of nondeterminism; and quantum nondeterministic computation where the proof is either 2) quantum or 3) classical. When the underlying model of computation is communication complexity, these three notions of nondeterminism yield three different types of communication called **QMA**, **QCMA** and **NQP** communication.

The study of nondeterministic quantum communication complexity started with de Wolf [dW03]. In that work, it was proved that **NQP**-communication can be exponentially stronger than classical nondeterministic communication. Le Gall [LG06] studied a different type of **QCMA**-communication where the length of the proof is not considered in the communication cost, and he showed a quadratic quantum-classical gap. Along this line of work, Klauck [Kla11] gave general lower bound techniques for **QMA**-communication and Raz and Shpilka [RS04] showed an exponential separation between **QMA**-communication and **MA**-communication complexities. All these previous works were in the 2-player setting.

*sunxiaoming@ict.ac.cn, Institute of Computing Technology, Chinese Academy of Sciences, China. The author is supported by the National Natural Science Foundation of China Grant 61170062 and 61222202.

†marcos.villagra@acm.org, Department of Information Science, University of Fuki, Japan. The author is supported by a JSPS Research Fellowship and KAKENHI grant No. 25.6353. Part of this work was done while the author was a graduate student at Nara Institute of Science and Technology and visiting the Institute of Computing Technology, Chinese Academy of Sciences.

ISSN 1433-8092

A very important lower bound technique for quantum communication is the norm-bound discovered by Linial and Shraibman [LS09c]. It essentially relates the 2-sided bounded-error quantum and classical communication complexities with the γ_2^α and μ^α norms of their corresponding communication matrices, where $1 \leq \alpha < \infty$ is a measure of approximation related to the error of the protocol. The norm-bounds were further extended to multiparty communication in the works of Lee and Shraibman [LS09a] and Lee, Schechtman and Schraibman [LSS09].

1.2 Our Results

In this paper we show exponential gaps between different modes of classical and quantum multiparty nondeterministic communication complexity.

Let $\mathbf{C}_k^{cc}$ be a k-party communication complexity class [BFS86]. We say that a boolean function f has a k-party $\mathbf{C}$-communication protocol if f can be computed by a k-party communication protocol whose "mode of communication" corresponds to the class $\mathbf{C}$. For example, a **BPP**-communication protocol for f is a protocol computing f with 2-sided bounded-error communication, and, an **NQP**-communication protocol for f outputs 1 with positive probability if and only if $f(x) = 1$. See Raz and Schpilka [RS04] and Klauck [Kla11] for the definition of **QMA** and **QCMA** nondeterministic communication modes. A boolean function f is in $\mathbf{C}_k^{cc}$ if and only if there exists a k-party $\mathbf{C}$-communication protocol for f with $polylog(n)$ cost where n is the size of the input.

Let $w_n(x_1, \ldots, x_k) = 1$ if $|x_1 \wedge \cdots \wedge x_k| \neq 1$ and -1 otherwise, where each $x_i \in \{-1, 1\}^n$, $|x|$ denotes the Hamming weight of x, and $\wedge$ is the bit-wise AND operator. We refer to this function as *de Wolf's function* [dW03]. The main result is the following theorem.

Theorem 1. *For any $k \geq 2$, $\log \gamma_{2,k}^\infty(w_n) = \Omega(\frac{n}{k2^{2^k}} - k)$.*

For any $k \geq 2$, Theorem 1 immediately implies the same lower bound for de Wolf's function in **NP**-communication, **BPP**-communication, **BQP**-communication, and **QCMA**-communication[1] complexities in the Number-On-Forehead and Number-In-Hand models [LS09c, LS09a, LSS09]. Furthermore, by previous work of de Wolf [dW03] we know that for any $k \geq 2$ there is a Number-On-Forehead **NQP**-protocol for de Wolf's function with cost $\mathcal{O}(\log n)$. This gives a gap between all modes of communication mentioned above and **NQP**-communication complexity which is upper-bounded by the nondeterministic tensor-rank of the communication tensor in the Number-On-Forehead model. The separation is exponential whenever $k = \mathcal{O}(1)$ and super-polynomial when $k = o(\log \log n)$. In complexity-theoretic terms $\mathbf{NQP}_k^{cc} \not\subseteq \mathbf{QCMA}_k^{cc}$ and hence $\mathbf{NQP}_k^{cc} \not\subseteq \mathbf{NP}_k^{cc}$ whenever the number of players is $o(\log \log n)$. Theorem 1 also partly solves an open problem of Klauck [Kla11] who conjectured the existence of a (partial) function with hard **QCMA**-communication complexity.

The main reason of these separations lays in another important consequence of Theorem 1: an exponential separation between nondeterministic rank [dW03, VNYN13] and the discrepancy bound. This is in contrast of the well known result by Nisan and Widgerson [NW95] that small rank implies large discrepancy for boolean matrices.

The proof of Theorem 1 follows from an application of the pattern tensor method and a new lower bound for polynomial threshold degree.

1.3 Open Problems

The modes of nondeterministic communication studied in this work might seem esoteric with no real implication to computation. However, previous work of Aaronson and Widgerson [AW09] showed that separations of complexity classes in communication complexity imply that *non-algebrazing* techniques will be required to show the same separations for Turing machines. Therefore, here we give a list of open problems left by this and previous work that we believe might be of interest for our understanding of quantum nondeterministic communication and computation.

1. **Lower bound method for QMA-communication.** Klauck [Kla11] gave two different ways to lower-bound **QMA**-communication, one based on Razborov's method and another the author called 1-sided smooth discrepancy. It is open if the norm-bound can also yield a lower bound for **QMA**-communication.

[1] To see the lower bound on **QCMA**-communication in terms of the $\gamma_{2,k}^\infty$-norm refer to [BBLV09].

2. **Separations for protocols with more players.** We believe that the denominator in the lower bound of Theorem 1 can be improved by using the techniques of [BDPW10]. The authors give a randomized reduction, different from [LS09a], and then derandomized it to obtain a 2^k factor in the denominator.

3. **The power of quantum vs classical proofs.** One important open problem in quantum complexity theory is about how much computational power is obtained with a quantum proof compared to a classical proof. This question was previously explored by Aaronson and Kuperberg [AK07]. To show a separation in communication, it is enough to show the existence of a total function with high γ_2^∞-norm and low **QMA**-communication complexity.

1.4 Outline

The rest of the paper is organized as follows. In Section 2 we introduce notations and a brief introduction to the norm-bound and the pattern tensor method. In Section 3 we show the upper-bound on de Wolf's function and Section 4 presents the proof of Theorem 1.

2 Preliminaries

In this paper we will deal without loss of generality with the sign versions of boolean functions. Let $f : (\{-1,1\}^n)^k \to \{-1,1\}$ be a sign-function. We will sometimes identify f with its communication tensor T_f where $T_f[x] = f(x)$ and is of order k. The Hadamard or entry-wise product of two tensors T and S is denoted by $T \circ S$. The inner product of T and S is $\langle T, S\rangle = \sum_{x_1,\dots,x_k} T[x_1,\dots,x_k]S[x_1,\dots,x_k]$. We also denote $[n] = \{0,\dots,n-1\}$.

2.1 Nondeterministic Quantum Communication Complexity

In this section we will define the different modes of nondeterministic quantum communication. For reference on classical nondeterministic communication we refer the reader to [KN97].

In a quantum communication protocol, $k \geq 2$ players can interchange qubits. The Hilbert space is defined as $\mathcal{H} = \mathcal{P}_1 \otimes \cdots \otimes \mathcal{P}_k \otimes \mathcal{C}$, where each $\mathcal{P}_i$ is the register of player i, and $\mathcal{C}$ is the channel. Each register $\mathcal{P}_i$ should have enough space to contain the inputs plus some extra workspace for the computations. To communicate, player i applies a unitary U_i to its register and the channel. This will correspond to the act of performing some private computation and sending a message. The length of this message will be the number of channel qubits affected by U_i. At the end of the protocol, one player will make a measurement to determine the output.

When there is no entanglement, the initial state of the protocol on input $x = (x_1,\dots,x_k)$ is

$$|\Psi^0\rangle = |x_1, 0\rangle \otimes \cdots \otimes |x_k, 0\rangle|0\rangle. \tag{1}$$

In the model with shared entanglement, the initial state is

$$|\Psi^0\rangle = \sum_z |x_1, z\rangle \otimes \cdots \otimes |x_k, z\rangle|0\rangle. \tag{2}$$

Before the protocol starts, there is a predefined order for the actions of the players. After kt rounds of communication the state is

$$|\Psi^{kt}\rangle = \overbrace{(U_k \cdots U_1)\cdots(U_k \cdots U_1)}^{t \text{ times}}|\Psi^0\rangle. \tag{3}$$

After t-rounds of communication we project the state $|\Psi^t\rangle$ onto the $|1\rangle$ state of the channel using an operator Π_1. The probability of measuring a 1 on the channel is thus

$$p = \langle\Psi^t|\Pi_1|\Psi^t\rangle. \tag{4}$$

The different modes of computation stem from the way we define the accepting probabilities. For instance, for bounded-error protocols, a Yes-instance is accepted if $p \geq 1 - \epsilon$ for some $\epsilon > 0$, and a No-instance is a accepted if $p \leq \epsilon$. Also, any protocol naturally defines a *communication tensor* T_f where $T[x_1,\dots,x_k] = f(x_1,\dots,x_k)$.

In this paper, we will be interested in *quantum nondeterministic protocols*. There are three different types of nondeterminism in quantum communication: i) **NQP**-communication, ii) **QMA**-communication, and iii) **QCMA**-communication. An **NQP**-communication protocol for a boolean function f outputs 1 with positive probability if and only if $f(x) = 1$. On the other hand, to define the other two modes of nondeterministic communication we need to introduce the notion of a proof. A **QMA**-communication (**QCMA**-communication) protocol outputs 1 if $f(x) = 1$ and there exists a quantum (classical) proof (known to all players) that makes the protocol accept with probability bounded away from $1/2$; if $f(x) = -1$ then for all quantum (classical) proofs the protocol will reject with probabiilty bounded away from $1/2$. Note that for **QMA** and **QCMA** protocols the communication cost is defined as the sum of the length of all messages plus the length of the proof. This way we can define the k-party $(\mathbf{NQP}, \mathbf{QMA}, \mathbf{QCMA})$-*communication complexity* of a function $f : (\{-1, 1\}^n)^k \to \{-1, 1\}$ as the minimum cost of a k-party $(\mathbf{NQP}, \mathbf{QMA}, \mathbf{QCMA})$ protocol for f respectively.

Furthermore, there are two common ways of communication: The Number-On-Forehead (NOF) model where the i-th player knows all inputs except x_i; and the Number-In-Hand (NIH) model where the i-th player only knows x_i.

2.2 The γ_2-norm

Linial and Shraibman [LS09c] introduced the use of factorization norms as tools for proving lower bounds in randomized and quantum communication complexities in the 2-player setting. In particular, they showed that a variation of this kind of norms yield the lower bounds. Given any real matrix M, its γ_2 norm is defined as

$$\gamma_2(M) = \min_{M=AB^T} \sigma(A)\sigma(B), \tag{5}$$

where $\sigma(A)$ is the largest ℓ_2 norm of a row of A (the number 2 in γ_2 stems from the fact that we take the ℓ_2-norm in $\sigma(A)$). Then, the approximate norm γ_2^α with approximation factor $\alpha \geq 1$ is given by

$$\gamma_2^\alpha(M) = \min_{1 \leq M' \circ M \leq \alpha} \gamma_2(M'), \tag{6}$$

where $1 \leq M' \circ M \leq \alpha$ indicates that each entry in $M' \circ M$ is bounded between 1 and α. In particular, when $\alpha \to \infty$,

$$\gamma_2^\infty(M) = \min_{1 \leq M' \circ M} \gamma_2(M'). \tag{7}$$

We define the *dual* norm of γ_2 as

$$\gamma_2^*(M) = \max_{M' : \gamma_2(M') \leq 1} \langle M, M' \rangle. \tag{8}$$

When the number of players is three or more, Lee, Schechtman and Shraibman [LSS09] extended the definition of the γ_2-norm to the multiplayer setting. First the authors identified the set of simple objects into which a successful quantum protocol decomposes the communication tensor T_f. This is defined as

$$\mathcal{C}_k = \left\{ C \;\middle|\; \begin{array}{l} \exists \text{ set of vectors } \{|\phi_i\rangle\} \text{ s.t. } C[x_1, \ldots, x_k] = \langle \phi_1(x_1), \ldots, \phi_i(x_i), \ldots, \phi_k(x_k) \rangle \\ \text{where } \||\phi_i\rangle\| \leq 1 \text{ for all } i, x_1, \ldots, x_k \end{array} \right\}, \tag{9}$$

where $\langle \phi_i, \ldots, \phi_k \rangle$ is a k-multilinear product. $\gamma_{2,k}$ is defined as

$$\gamma_{2,k}(T_f) = \min \left\{ \sum_i |\sigma_i| : T_f = \sum_i \sigma_i C_i, \text{ where } C_i \in \mathcal{C}_k \right\}. \tag{10}$$

The approximate norm is defined in the same way as in equations (6) (7). A characterization in terms of a SDP was also given by Lee and Shraibman [LS09b].

Lemma 1. *For any order-k sign tensor T and $\alpha \geq 1$*

$$\begin{array}{rcl} \gamma_{2,k}^\alpha(T) & = & \max_A \frac{(1+\alpha)}{2} \langle T, A \rangle + \frac{(1-\alpha)}{2} \|A\|_1, \\ & & s.t. \quad \gamma_2^*(A) \leq 1, \end{array}$$

where we maximize over all real matrices A with γ_2^*-norm at most 1. In particular,

$$\begin{array}{rcl} \gamma_{2,k}^\infty(T) & = & \max_A \langle T, A \rangle, \\ & & s.t. \quad \gamma_2^*(A) \leq 1. \end{array}$$

Let $R_{\epsilon,k}$, $Q_{\epsilon,k}$ and N_k denote the k-party randomized, quantum and classical nondeterministic communication complexities respectively.

Lemma 2 ([LS09a, LSS09]). *For any function $f : (\{-1,1\}^n)^k \to \{-1,1\}$ and for any $0 < \epsilon < 1/2$, $R_{\epsilon,k}(f) \geq Q_{\epsilon,k}(f) = \Omega(\log \gamma_{2,k}^{\alpha_\epsilon})$ and $N_k(f) = \Omega(\log \gamma_{2,k}^{\infty})$, where $\alpha_\epsilon = 1/(1-2\epsilon)$.*

2.3 Approximating Polynomials and The Pattern Tensor Method

In this section we give a brief overview of the pattern tensor method which relates communication complexity to the degree of an approximating polynomial [She08]. An alternative technique relating polynomial degree and communication was given in [SZ09] (see also [LZ10]).

We start by defining the notion of *approximating polynomials* as presented in [LS09a]. Let $f : \{-1,1\}^n \to \{-1,1\}$. For any $\alpha \geq 1$, a multilinear polynomial $p(\cdot)$ gives an α-approximation of f if $1 \leq f(x)p(x) \leq \alpha$ for all $x \in X$. Similarly, $p(\cdot)$ gives an ∞-approximation of f if $1 \leq f(x)p(x)$ for all $x \in X$. The α-approximate degree of f, denoted by $deg_\alpha(f)$, is the smallest degree of a polynomial p that α-approximates f (similarly for deg_∞).

As noted in [LS09a], deg_α is equivalent to the more typical approximate degree $\widetilde{deg}_\epsilon$ defined as $\widetilde{deg} = \epsilon(f) = \min\{deg(g) : \|f-g\|_\infty \leq \epsilon\}$, where f is a 0/1 valued function. Indeed, if you let $0 < \epsilon < 1/2$ and $\alpha_\epsilon = (1+2\epsilon)/(1-2\epsilon)$ we have $\widetilde{deg}_\epsilon(f_{0/1}) = deg_{\alpha_\epsilon}(f_\pm)$, where $f_{0/1}$ and $f_\pm$ are the boolean and sign versions of the same function f.

The following lemma was proved by Lee and Shraibman [LS09a] based on a generalization of the pattern matrix method developed by Sherstov [She08]. An order-k pattern tensor is defined by natural numbers t, m and a function $\phi : \{-1,1\}^t \to \mathbb{R}$. Let $x = (x_1, \ldots, x_t)$ where each x_i is an order-$(k-1)$ tensor with side length m, i.e., x_i is an element of the tensor product of $k-1$ vector spaces on $\{-1,1\}$ each of dimension m. Let $S_i \in [m]^t$ for $i = 1, \ldots, k-1$ be ordered sets. Let $S_i[r] \in [m]$ refer to the r-th element of S_i, which can be thought of as a pointer into the i-th dimension of x_r. The set $S = (S_1, \ldots, S_{k-1})$ selects a t-bit string from x as

$$x|_S = x_1\Big[S_1[1], \ldots, S_{k-1}[1]\Big] \cdots x_t\Big[S_1[t], \ldots, S_{k-1}[t]\Big]. \tag{11}$$

The (k, m, t, ϕ)-pattern tensor F is given by

$$F[x, S_1, \ldots, S_{k-1}] = \phi(x|_S). \tag{12}$$

Lemma 3 (The Pattern Tensor Method [LS09a]). *For nonnegative integers k, t and a boolean function ϕ on m variables, let F be the (k, m, t, ϕ)-patter tensor, then*

$$\log \mu^\alpha(F) = \Omega(\widetilde{deg}_\epsilon(\phi)/2^{k-1}),$$

provided $m \geq 2e(k-1)2^{2^{k-1}}/\widetilde{deg}_\epsilon(\phi)$. Furthermore,

$$\log \mu^\infty(F) \geq deg_\infty(\phi)/2^{k-1},$$

provided $m \geq 2e(k-1)2^{2^{k-1}}/deg_\infty(\phi)$.

The ∞-approximation degree is equivalent to the older notion of *polynomial threshold degree.* If the sign of a polynomial $p(x)$ equals $f(x)$ for all $x \in X$ we say that p *sign-represents* f. We denote by $thr(f)$ the minimum degree over all polynomials that sign-represent f[2].

Lemma 4. *For any boolean function $f : X \to \{-1,1\}$, $deg_\infty(f) = thr(f)$.*

Proof. Let p be a multilinear polynomial of degree d that ∞-approximates f with $deg_\infty(f) = d$. Hence, $p(x)f(x) \geq 1$ and p also sign-represents f. Thus, $thr(f) \leq d$.

Now consider the case when p sign-represents f and $thr(f) = d$. Then, no matter how small $p(x)$ is, we can always construct a polynomial $\hat{q}$ that ∞-approximates f with degree at most d for which $|\hat{q}(x)| \geq 1$ for all inputs x. For instance, if we let $\beta = \min_x |p(x)|$, we can make $\hat{q}(x) = p(x)/\beta$. Thus, $\hat{q}(x)p(x) \geq 1$ and hence $deg_\infty(f) \leq d$. □

[2]There is also the notion of *weak sign-representing polynomials* where $p(x)$ could be 0 for some $x \in \{0,1\}^n$. In this paper, we only deal with *strong sign-representing polynomials* as defined above.

By Lemma 4 and the pattern tensor method we can obtain a different lower bound on γ_2^∞ in terms of the threshold degree by applying the *multi-dimensional Grothendieck's inequality* as given in [LSS09, Theorem 6].

Lemma 5. *Let F be a (k, n, t, ϕ)-pattern tensor, then* $\log \gamma_2^\infty(F) = \Omega(thr(\phi)/2^{k-1} - k)$.

3 Upper Bound on de Wolf's Function

In previous work, de Wolf [dW03] studied the following function

$$w_n(x_1, \ldots, x_k) = \begin{cases} 1 & \text{if } |x_1 \wedge \cdots \wedge x_k| \neq 1 \\ -1 & \text{otherwise} \end{cases}, \tag{13}$$

where each $x_i \in \{-1, 1\}^n$ (it is the complement of the Unique-Intersection function). In [dW03] this function, which we refer to as *de Wolf's function*, was used to show an exponential separation between classical nondeterministic and **NQP**-comunication complexity in the 2-player setting.

Let $NQP_k^{NOF}(f)$ denote the **NQP**-communication complexity of f for k players in the Number-On-Forehead model. By previous work of de Wolf [dW03] and Villagra et al. [VNYN13] we have the following upper bound whose proof is included for the sake of completeness.

Lemma 6. $NQP_k^{NOF}(w_n) = \mathcal{O}(\log n)$.

Proof. For each i let $x_i = x_{i,j_1} \ldots x_{i,j_n}$ and let T_j be an order-k tensor where $T_j[x_1, \ldots, x_k] = 1$ if $x_{1,j} \wedge \cdots \wedge x_{k,j} = 1$ and $T_j[x_1, \ldots, x_k] = 0$ otherwise. Note that for each j the tensor T_j has rank 1. Define the order-k tensor T by

$$T[x_1, \ldots, x_k] = \sum_{j=1}^{n} T_j[x_1, \ldots, x_k] - 1.$$

This tensor has rank n. Also T is a nondeterministic communication tensor[3] for f since $T[x_1, \ldots, x_k] = 0$ if and only if $|x_1 \wedge \cdots \wedge x_k| = 1$. Hence, by previous results of [dW03] and [VNYN13], the strong nondeterministic communication complexity in the Number-On-Forehead model is upper-bounded by the logarithm of the tensor rank of T. □

4 Proof of Theorem 1

4.1 Preparation for the Proof

To prove the theorem we make use of Lemma 5. For the lower bound on threshold degree, we rely on a powerful technique by O'Donnell and Servedio [OS10] which restates the lower bound problem as a feasibility question of a linear program.

Let $\Delta : X \to \mathbb{R}^{\geq 0}$ be a distribution over some set X. The *support* of Δ is the set $\{x : \Delta(x) > 0\}$. If the support is the whole set of X we say that Δ is a *total distribution*. If $\sum_x \Delta(x) = 1$ then Δ is a *probability distribution*. Given a monomial x_S, $S \subseteq [n]$, the *correlation* of x_S with a boolean function f under a distribution Δ is

$$\mathbf{E}_\Delta[f(x)x_S] = \sum_{x \in \{-1,1\}} f(x) x_S \Delta(x). \tag{14}$$

Theorem 2 (Theorem of the Alternative [OS10])**.** *Let $f : X \to \{-1, 1\}$ be a boolean function, and let $S \subseteq 2^{[n]}$ be any set of monomials. Then exactly one of the following holds:*

1. *f can be sign-represented by a polynomial whose non-zero coefficients correspond to monomials in S; or,*

2. *there is a distribution on X under which f has zero correlation to every monomial in S.*

The technique by O'Donnell and Servedio [OS10] relies on the theorem of the alternative. Construct a probability distribution for a function f with zero correlation with a set of low-degree monomials S. Immediately, by Theorem 2, there is no polynomial that sign-represents f with non-zero coefficients corresponding to monomials in S. Hence, the polynomial threshold degree must be high.

[3] T is a *nondeterministic communication tensor* if $T[x_1, \ldots, x_k] \neq 0$ if and only if $f(x_1, \ldots, x_k) = 1$.

4.2 Main Proof

To prove the lower bound we rely heavily on the pattern tensor method (Lemma 3). Let $h_n : [2^n] \to \{-1, 1\}$ be defined by

$$h_n(z) = \begin{cases} -1 & \text{if } z \in [2^n] \text{ is a power of } 2 \\ 1 & \text{otherwise} \end{cases}. \tag{15}$$

Note that h_n is the complement of the Unique-OR function. Define the function $\phi_t : \{-1,1\}^t \to \{-1,1\}$ as $\phi_t(x) = h_t\left(\frac{(x_1+1)}{2}2^{t-1} + \cdots + \frac{(x_t+1)}{2}2^0\right)$ and let $c_k = 2e(k-1)2^{2^{k-1}}$. Let F be the (k, m, t, ϕ_t)-pattern tensor with $m = c_k t / thr(\phi_t)$ and $t = \left\lfloor \frac{n}{c_k^{k-1}} \right\rfloor$. Lemma 5 implies that

$$\log \gamma_2^\infty(F) = \Omega(thr(h_t)/2^{k-1} - k). \tag{16}$$

Let M_{w_n} be the communication tensor for de Wolf's function

$$M_{w_n} = [w_n(x_1, \ldots, x_k)]_{x_1,\ldots,x_k \in \{-1,1\}^n}. \tag{17}$$

If F is a sub-tensor of M_w then

$$\log \gamma_2^\infty(M_{w_n}) \geq \log \gamma_2^\infty(F) = \Omega(thr(h_t)/2^{k-1} - k). \tag{18}$$

Thus, Theorem 1 will follow from the following two lemmas.

Lemma 7. *F is a sub-tensor of M_{w_n}.*

Lemma 8. *$thr(h_n) = \Omega(n)$.*

The proof of Lemma 7 goes exactly as the proof given by Lee and Shraibman [LS09a] for the disjointness function. For the sake of completeness we give the proof in Appendix A. The proof of Lemma 8 makes use of the technique by O'Donnell and Servedio [OS10] and is presented next.

4.3 Proof of Lemma 8

As was previously done in [OS10], it is sufficient to find a support $\mathcal{Z} \subseteq [2^n]$ and a probability distribution Δ over $\mathcal{Z}$ such that

$$\forall 0 \leq i \leq d, \quad \mathbf{E}_\Delta[h_n(y)y^i] = \sum_{y \in \mathcal{Z}} \Delta(y) h_n(y) y^i = 0 \tag{19}$$

for some fixed d and y^i is the i-th power of y. By looking each $\Delta(y)$ as a variable we can restate Equation (19) as a system of linear equations. Let $y_i \in \mathcal{Z}$ and let $z = size(\mathcal{Z}) = \max\{y_i \in \mathcal{Z}\}$. Intuitively, $size(\mathcal{Z})$ is the greatest element of $\mathcal{Z}$. Denote $\Delta_i = \Delta(y_i)$, then

$$\begin{bmatrix} h_z(y_1)y_1^0 & h_z(y_2)y_2^0 & h_z(y_3)y_3^0 & \ldots & h_z(y_{|\mathcal{Z}|})y_{|\mathcal{Z}|}^0 \\ h_z(y_1)y_1^1 & h_z(y_2)y_2^1 & h_z(y_3)y_3^1 & & h_z(y_{|\mathcal{Z}|})y_{|\mathcal{Z}|}^1 \\ \vdots & & & \ddots & \vdots \\ h_z(y_1)y_1^d & h_z(y_2)y_2^d & h_z(y_3)y_3^d & & h_z(y_{|\mathcal{Z}|})y_{|\mathcal{Z}|}^d \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_{|\mathcal{Z}|-1} \\ \Delta_{|\mathcal{Z}|} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{20}$$

where each $\Delta_i \geq 0$. The last line in the coefficient matrix indicates that we want Δ to be a probability distribution. If the system of equations has a feasible solution, then by the theorem of the alternative we immediately obtain a $d+1$ lower bound on the polynomial threshold degree of h_z.

With the help of an LP-solver we are able to come out with three different support sets for the cases $d = 1$, $d = 2$ and $d \geq 3$. Denote these sets by $\mathcal{Z}_{d=1}$, $\mathcal{Z}_{d=2}$, and $\mathcal{Z}_{d\geq 3}$ respectively. Below we show that there are support sets $\mathcal{Z}_{d=1}$, $\mathcal{Z}_{d=2}$, $\mathcal{Z}_{d\geq 3}$ that yield feasibility of (20) when $size(\mathcal{Z}_{d=1}) = 4$, $size(\mathcal{Z}_{d=2}) = 5$, and $size(\mathcal{Z}_{d\geq 3}) = 2^d$. Given that $size(\mathcal{Z})$ for any support $\mathcal{Z}$ can be as large as $\Theta(2^n)$ we have $thr(h_n) = \Omega(n)$.

In the following we analyze each support set separately. First we use an LP-solver to find a support for the cases $d = 1$ and $d = 2$. Then we use induction for $d \geq 3$ with base case $d = 3$.

4.3.1 Case $d = 1$

The support set $\mathcal{Z}_{d=1} = \{1, 3, 4\}$ gives the following system of equations

$$\begin{bmatrix} -1 & 1 & -1 \\ -1 & 3 & -4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_3 \\ \Delta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{21}$$

A feasible solution is $\Delta_1 = 1/6, \Delta_3 = 1/2, \Delta_4 = 1/3$.

4.3.2 Case $d = 2$

The support set $\mathcal{Z}_{d=2} = \{1, 3, 4, 5\}$ gives the following system of equations

$$\begin{bmatrix} -1 & 1 & -1 & 1 \\ -1 & 3 & -4 & 5 \\ -1 & 9 & -16 & 25 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_3 \\ \Delta_4 \\ \Delta_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{22}$$

A feasible solution is $\Delta_1 = 1/18, \Delta_3 = 1/3, \Delta_4 = 4/9, \Delta_5 = 1/6$.

4.3.3 Case $d \geq 3$

For this case we select $\mathcal{Z}_{d\geq 3} = \{1, \ldots, 2^d\}$ as support set and prove by induction on d that there are feasible solutions for all $d \geq 3$.

The base case of the induction is $d = 3$ which has the following system of linear equations

$$\begin{bmatrix} -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 \\ -1 & -2 & 3 & -4 & 5 & 6 & 7 & -8 \\ -1 & -4 & 9 & -16 & 25 & 36 & 49 & -64 \\ -1 & -8 & 27 & -64 & 125 & 216 & 343 & -512 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \\ \Delta_4 \\ \Delta_5 \\ \Delta_6 \\ \Delta_7 \\ \Delta_8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{23}$$

A feasible solution is $\Delta_1 = 5/98, \Delta_3 = 9/28, \Delta_4 = 5/14, \Delta_7 = 5/28, \Delta_9/98$ and $\Delta_2 = \Delta_5 = \Delta_6 = 0$. The size is 9.

Now assume that for $d-1$ there is a feasible solution $(\Delta'_1, \ldots, \Delta'_{2^{d-1}})$ with support of size 2^{d-1}. The system of linear equations for $d-1$ is

$$\begin{bmatrix} h_{d-1}(1)\cdot 1^0 & h_{d-1}(2)\cdot 2^0 & \ldots & h_{d-1}(2^{d-1})\cdot (2^{d-1})^0 \\ h_{d-1}(1)\cdot 1^1 & h_{d-1}(2)\cdot 2^1 & & h_{d-1}(2^{d-1})\cdot (2^{d-1})^1 \\ \vdots & & \ddots & \vdots \\ h_{d-1}(1)\cdot 1^{d-1} & h_{d-1}(2)\cdot 2^{d-1} & & h_{d-1}(2^{d-1})\cdot (2^{d-1})^{d-1} \\ 1 & 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} \Delta'_1 \\ \Delta'_2 \\ \vdots \\ \Delta'_{2^{d-1}-1} \\ \Delta'_{2^{d-1}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{24}$$

The system of equations for d is

$$\begin{bmatrix} h_d(1)\cdot 1^0 & h_d(2)\cdot 2^0 & \ldots & h_d(2^d)\cdot (2^d)^0 \\ h_d(1)\cdot 1^1 & h_d(2)\cdot 2^1 & & h_d(2^d)\cdot (2^d)^1 \\ \vdots & & \ddots & \vdots \\ h_d(1)\cdot 1^d & h_d(2)\cdot 2^d & & h_d(2^d)\cdot (2^d)^d \\ 1 & 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_{2^d-1} \\ \Delta_{2^d} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{25}$$

We will use the feasible solutions from (24) to construct the new solution for (25). First let $(\Delta_1, \ldots, \Delta_{2^{d-1}-1}) = (\Delta'_1, \ldots, \Delta'_{2^{d-1}-1})$. Also set $(\Delta_{2^{d-1}}, \ldots, \Delta_{2^d-2}) = (0, \ldots, 0)$. With these assignments, we will solve (25) only for the variables Δ_{2^d-1} and Δ_{2^d}. From now on, denotes these two variables by σ and ξ respectively.

After the assignation of values to variables made above, we have that the coefficient matrix of (25) looks like

$$\begin{bmatrix} h_d(1)\cdot 1^0 & \dots & h_d(2^{d-1}-1)\cdot(2^{d-1}-1)^0 & h_d(2^d-1)\cdot(2^d-1)^0 & h_d(2^d)\cdot(2^d)^0 \\ h_d(1)\cdot 1^1 & \dots & h_d(2^{d-1}-1)\cdot(2^{d-1}-1)^1 & h_d(2^d-1)\cdot(2^d-1)^1 & h_d(2^d)\cdot(2^d)^1 \\ \vdots & \ddots & & \vdots & \\ h_d(1)\cdot 1^d & \dots & h_d(2^{d-1}-1)\cdot(2^{d-1}-1)^d & h_d(2^d-1)\cdot(2^d-1)^d & h_d(2^d)\cdot(2^d)^d \\ 1 & \dots & 1 & 1 & 1 \end{bmatrix}, \tag{26}$$

where the variable vector is $(\Delta'_1, \dots, \Delta'_{2^{d-1}-1}, \sigma, \xi)$ having only σ and ξ as free-variables. This system can be rewritten as a system with two constraints by adding all rows together except the last in the following way

$$\begin{cases} A + C\sigma + D\xi & = 0 \\ B + \sigma + \xi & = 1 \end{cases}, \tag{27}$$

where

$$\begin{aligned} A &= \left(h_d(1)\cdot 1^0 + \dots + h_d(1)\cdot 1^d\right)\Delta'_1 + \cdots \\ &\quad + \left(h_d(2^{d-1}-1)\cdot(2^{d-1}-1)^0 + \dots + h_d(2^{d-1}-1)\cdot(2^{d-1}-1)^d\right)\Delta'_{2^{d-1}-1}, \\ B &= \Delta'_1 + \dots + \Delta'_{2^{d-1}-1}, \\ C &= h_d(2^d-1)\cdot(2^d-1)^0 + \dots + h_d(2^d-1)\cdot(2^d-1)^d, \\ D &= h_d(2^d)\cdot(2^d)^0 + \dots + h_d(2^d)\cdot(2^d)^d. \end{aligned}$$

A solution for this new system of equations is

$$\sigma = 1 - B - \xi \quad \text{and} \quad \xi = \frac{-A + C(B-1)}{D-C}.$$

To finish the proof, we just need to show that σ and ξ are positive. By taking a closer look at the values B, C, D we note that

1. $0 < B < 1$ because the values $(\Delta'_1, \dots, \Delta'_{2^{d-1}-1})$ are all positive values,
2. $C > 0$ because $h_{2^d}(2^d-1)$ is positive, and
3. $D < 0$ because $h_{2^d}(2^d)$ is negative.

Thus, if $A >$ then $\xi > 0$ and $\sigma > 0$ and the support is of size 2^d.

Claim 1. $A > 0$.

Proof. To show that $A > 0$ write

$$A = A' + A'' \tag{28}$$

where

$$\begin{aligned} A' &= \left(h_d(1)\cdot 1^0 + \dots + h_d(1)\cdot 1^{d-1}\right)\Delta'_1 + \cdots \\ &\quad + \left(h_d(2^{d-1}-1)\cdot(2^{d-1}-1)^0 + \dots + h_d(2^{d-1}-1)\cdot(2^{d-1}-1)^{d-1}\right)\Delta'_{2^{d-1}-1} \\ &= \sum_{i=1}^{2^{d-1}-1} h_d(i)\cdot i^0 \Delta'_i + \dots + \sum_{i=1}^{2^{d-1}-1} h_d(i)\cdot i^{d-1}\Delta'_i, \end{aligned} \tag{29}$$

and

$$\begin{aligned} A'' &= h_d(1)\cdot 1^d \cdot \Delta'_1 + \dots + h_d(2^{d-1}-1)\cdot(2^{d-1}-1)^d\cdot\Delta'_{2^{d-1}-1} \\ &= \sum_{i=1}^{2^{d-1}-1} h_d(i)\cdot i^d \Delta'_i. \end{aligned} \tag{30}$$

Let A_t be each summation term in A' and A'', i.e., $A' = A_1 + \cdots + A_{d-1}$ and $A'' = A_d$ where

$$A_t = \sum_{i=1}^{2^{d-1}-1} h_d(i) \cdot i^t \Delta'_i \tag{31}$$

$$= \sum_{i=1}^{2^{d-1}-1} i^t \Delta'_i - 2\sum_{j=0}^{d-2} (2^j)^t \Delta'_{2^j}. \tag{32}$$

Note that for each $t \in [d]$, A_t corresponds to the sum of one row in (24) with the exception of the last element in that row. Also note that the last column only contains negative numbers because $h_d(2^{d-1})$ is negative. This necessarily makes each $A_t > 0$ for $t \in [d]$ in order to cancel out with the last element of each row of (24). Hence, $A' > 0$.

A closer look at (32) also reveals that A_t is a monotone increasing function in t, hence, $A_t < A_{t+1}$ for all t. This way, given that $A_t > 0$ for $t \in [d]$ we have that $0 < A_{d-1} < A_d = A''$. Thus $A > 0$. □

References

[AK07] Scott Aaronson and Greg Kuperberg. Quantum versus classcial proofs and advice. *Theory of Computing*, 3:129–157, 2007.

[AW09] Scott Aaronson and Avi Widgerson. Algebrization: A new barrier in complexity theory. *ACM Transactions on Computation Theory*, 1(1), 2009.

[BBLV09] Jop Briet, Harry Buhrman, Troy Lee, and Thomas Vidick. Multiplayer XOR games and quantum communication complexity with clique-wise entanglement. *arXiv:0911.4007*, 2009.

[BDPW10] Paul Beame, Matei David, Toniann Pitassi, and Philipp Woelfel. Separating deterministic from randomized multiparty communication complexity. *Theory of Computing*, 6:201–225, 2010.

[BFS86] László Babai, Péter Frankl, and János Simon. Complexity classes in communication complexity theory. In *Proceedings of the 27th Annual Symposium on Foundations of Computer Science*, pages 337–347, 1986.

[dW03] Ronald de Wolf. Nondeterministic quantum query and quantum communication complexities. *SIAM Journal on Computing*, 32(3):681–699, 2003.

[Kla11] Hartmut Klauck. On arthur merlin games in communication complexity. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity*, pages 189–199, 2011.

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[LG06] François Le Gall. Quantum weakly nondeterministic communication complexity. In *Proceedings of the 31st International Symposium on Mathematical Foundations of Computer Science*, volume 4162 of *Lecture Notes in Computer Science*, pages 658–669. Springer, 2006.

[LS09a] Troy Lee and Adi Shraibman. Disjointness is hard in the multi-party Number-On-The-Forehead model. *Computational Complexity*, 18(2):309–336, 2009.

[LS09b] Troy Lee and Adi Shraibman. Lower Bounds in Communication Complexity. *Foundations and Trends in Theoretical Computer Science*, 3(4):263–399, 2009.

[LS09c] Nati Linial and Adi Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures and Algorithms*, 34(3):368–394, 2009.

[LSS09] Troy Lee, Gideon Schechtman, and Adi Shraibman. Lower bounds on quantum multiparty communication complexity. In *Proceedings of the 24th IEEE Conference on Computational Complexity*, 2009.

[LZ10] Troy Lee and Shengyu Zhang. Composition theorems in communication complexity. In *Proceedings of the 37th International Colloqium on Automata, Languages and Programming*, volume 6198 of *Lecture Notes in Computer Science*, pages 475–489. Springer Berlin Heidelberg, 2010.

[NW95] Noam Nisan and Avi Widgerson. On rank vs. communication complexity. *Combinatorica*, 15(4):557–565, 1995.

[OS10] Ryan O'Donnell and Rocco Servedio. New degree bounds for polynomial threshold functions. *Combinatorica*, 30(3):327–358, 2010.

[RS04] Ran Raz and Amir Shpilka. On the power of quantum proofs. In *Proceedings of the 19th IEEE Annual Conference on Computational Complexity*, pages 260–274, 2004.

[She08] Alexander A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pages 85–94, 2008.

[SZ09] Yaoyun Shi and Yufan Zhu. The quantum communication complexity of block-composed functions. *Quantum Information and Computation*, 9(5):444–460, 2009.

[VNYN13] Marcos Villagra, Masaki Nakanishi, Shigeru Yamashita, and Yasuhiko Nakashima. Tensor rank and strong quantum nondeterminism in multiparty communication. *IEICE Transactions on Information and Systems*, E96-D(1):1–8, 2013.

A Proof of Lemma 7

First we review the definitions. The communication tensor for de Wolf's function, denoted M_{w_n}, is given by

$$M_{w_n} = \left[w_n(x_1, \ldots, x_k)\right]_{x_1,\ldots,x_k \in \{-1,1\}^n}. \tag{33}$$

Define the function $\phi_t : \{-1,1\}^t \to \{-1,1\}$ as $\phi_t(z) = h_t\left(\frac{(z_1+1)}{2}2^{t-1} + \cdots + \frac{(z_t+1)}{2}2^0\right)$ and let $c_k = 2e(k-1)2^{2^{k-1}}$. Let F be the (k, m, t, ϕ_t)-pattern tensor. In particular,

$$F[y, S_1, \ldots, S_{k-1}] = \phi_t\left(y_1\Big[S_1[1], \ldots, S_{k-1}[1]\Big] \ldots y_t\Big[S_1[t], \ldots, S_{k-1}[t]\Big]\right) = \phi_t(y|_S), \tag{34}$$

where each $S_j[i] \in [m]$ and $y = (y_1, \ldots, y_t)$ is a vector of t tensors each of order $k-1$. We want to prove that F is a sub-tensor of M_{w_n}, i.e., there is a reduction from the problem of computing F to M_{w_n}.

Let $n = tm^{k-1}$. To each S_i we associate a vector of order-$(k-1)$ tensors $z_i = (z_i^1, \ldots, z_i^t)$ with side length m. We set $z_i^j[u_1, \ldots, u_{k-1}] = 1$ if and only if $u_i = S_i[j]$ and 0 otherwise.

Consider the vector $z_1 \wedge z_2 = (z_1^1 \wedge z_2^1 \cdots z_1^t \wedge z_2^t)$. In this example, $z_1^1 \wedge z_2^1$ is 1 in coordinate $(u_1, \ldots, u_{k-1})$ if and only if $u_1 = S_1[1] \wedge u_2 = S_2[1]$. Extrapolating this reasoning to the vector $z_1 \wedge \cdots \wedge z_{k-1}$ we have that the coordinates that are taken in y when restricting to the set S are exactly the same coordinates where the vector $z_1 \wedge \cdots \wedge z_{k-1}$ is equal to 1. Hence,

$$\begin{aligned}
\phi_t(y|_S) &= \phi_t'(|y|_S|) \\
&= \phi_n'(|y \wedge (z_1 \wedge \cdots \wedge z_{k-1})|) \\
&= \phi_n(y \wedge (z_1 \wedge \cdots \wedge z_{k-1})) \\
&= \phi_n(x_1 \wedge \cdots \wedge x_k) \\
&= w_n(x_1, \ldots, x_k),
\end{aligned}$$

where the first and fourth equalities follow from the fact that ϕ_t is a symmetric function for any t.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de