Lower bounds on quantum multiparty communication complexity

Troy Lee
Department of Computer Science
Columbia University
New York, USA
troyjlee@gmail.com

Gideon Schechtman & Adi Shraibman
Department of Mathematics
Weizmann Institute of Science
Rehovot, Israel
{*gideon.schechtman, adi.shraibman*}*@weizmann.ac.il*

***Abstract*—A major open question in communication complexity is if randomized and quantum communication are polynomially related for all total functions. So far, no gap larger than a power of two is known, despite significant efforts.**

We examine this question in the number-on-the-forehead model of multiparty communication complexity. We show that essentially all lower bounds known on randomized complexity in this model also hold for quantum communication. This includes bounds of size $\Omega(n/2^k)$ for the k-party complexity of explicit functions, bounds for the generalized inner product function, and recent work on the multiparty complexity of disjointness. To the best of our knowledge, these are the first lower bounds of any kind on quantum communication in the general number-on-the-forehead model.

We show this result in the following way. In the two-party case, there is a lower bound on quantum communication complexity in terms of a norm γ_2, which is known to subsume nearly all other techniques in the literature. For randomized complexity there is another natural bound in terms of a different norm μ which is also one of the strongest techniques available. A deep theorem in functional analysis, Grothendieck's inequality, implies that γ_2 and μ are equivalent up to a constant factor. This connection is one of the major obstacles to showing a larger gap between randomized and quantum communication complexity in the two-party case.

The lower bound technique in terms of the norm μ was recently extended to the multiparty number-on-the-forehead model. Here we show how the γ_2 norm can be also extended to lower bound quantum multiparty complexity. Surprisingly, even in this general setting the two lower bounds, on quantum and classical communication, are still very closely related. This implies that separating quantum and classical communication in this setting will require the development of new techniques. The relation between these extensions of μ and γ_2 is proved by a multi-dimensional version of Grothendieck's inequality.

***Keywords*-Communication complexity, quantum computing, number-on-the-forehead model**

I. Introduction

Since its introduction thirty years ago [Abe78], [Yao79], communication complexity has become a key concept in complexity theory and theoretical computer science in general. Part of its appeal is that it has applications to many different computational models, for example to formula size and circuit depth, proof complexity, branching programs, VLSI design, and time-space trade-offs for Turing machines (see [KN97] for more details).

A major open question in communication complexity is if randomized and quantum communication complexity are polynomially related for all total functions. While an exponential separation between these models has been exhibited for a promise problem [Raz99], for total functions currently the largest known gap is a power of two, realized by the disjointness function which has bounded-error randomized complexity $\Theta(n)$ [KS87] and quantum complexity $\Theta(\sqrt{n})$ [Raz03], [AA05].

Part of the difficulty of showing larger gaps between these models is that there are very few techniques known that lower bound randomized complexity and do not also work for quantum complexity. In the two-party model we currently have a relatively good understanding of how the various lower bound techniques are related. For the classical case, a powerful lower bound technique is in terms of a norm μ. For quantum communication there is another bound in terms of a different norm γ_2.

The μ norm is the norm induced by the absolute convex hull of combinatorial rectangles. The norm γ_2 is a factorization norm—it seeks the best factorization of an operator from ℓ_1 to ℓ_∞ via ℓ_2. Formally, for every real matrix B

$$\gamma_2(B) = \min_{XY=B} \|X\|_{2\to\infty}\|Y\|_{1\to 2}.$$

Where $\|X\|_{2\to\infty} = \max_{v:\|v\|_2\leq 1} \|Xv\|_\infty$ and the operator norm $\|Y\|_{1\to 2} = \max_{v:\|v\|_1\leq 1} \|Yv\|_2$. For more details see Section IV.

To deal with bounded-error models, the appropriate quantity is an approximated version of the underlying norm. For example, for quantum communication complexity the lower bound is in terms of γ_2^α, defined next. For a real number $\alpha \geq 1$ and a sign matrix A, $\gamma_2^\alpha(A)$ is defined by

$$\gamma_2^\alpha(A) = \min_{B:1\leq b_{ij}a_{ij}\leq\alpha} \gamma_2(B).$$

The approximation variant μ^α of μ is defined analogously. In both cases the parameter α is related to the maximum allowed error probability of the algorithm.

All lower bounds on randomized and quantum communication complexity that use the structure of Euclidean space in any way can be shown to follow from the μ^α bound for randomized communication complexity, and from the

γ_2^α bound for quantum communication complexity [LS07], [LS08a]. This includes the discrepancy method [KN97], bounds using Fourier analysis [Raz95], [Kla01], bounds in terms of singular values [Kla01], [Raz03], approximation rank [BW01], and more.

Grothendieck's inequality, a deep theorem from functional analysis, shows, however, that μ and γ_2 are related by a constant factor. Thus, one cannot use any of the aforementioned lower bound techniques to separate quantum and classical communication complexity. Notable exceptions of lower bound methods which can prove separations between randomized and quantum communication complexity include the corruption bound [Yao83] and information theory methods [CSWY01]. Both of these can show an $\Omega(n)$ lower bound on the communication complexity of disjointness [Raz92], [BYJKS04], whereas the quantum communication complexity is $\Theta(\sqrt{n})$ for this problem [Raz03], [AA05].

It is an open question whether $\log \mu^\alpha$ is polynomially related to randomized communication complexity [LS07]. If there is such a polynomial relation, then by Grothendieck's inequality, quantum and classical communication complexity are also polynomially related. It is also known that μ^α is polynomially related to approximation rank [LS08a], thus this question is also nicely linked to the famous log rank conjecture for deterministic communication complexity [LS88].

In this paper we generalize this theory to multiparty communication complexity. The major application we discuss is for the multiparty number-on-the-forehead (NOF) model of communication complexity originally introduced by Chandra, Furst and Lipton [CFL83]. In this model there are k-players trying to evaluate a function $f(x_1, \ldots, x_k)$, but now player i knows the entire input except for x_i. This large overlap in information makes showing lower bounds very difficult in this model. This difficulty, however, is rewarded by implications for circuit complexity [HG91] and proof complexity [BPS06].

To generalize the results from the two-player model we need first to extend the μ^α and γ_2^α bounds. The μ^α bound was extended to randomized multiparty communication complexity in [LS08b], and independently in [CA08] where it is called the generalized discrepancy method. In analogy with this extension of the μ norm, we similarly show that a natural extension of the γ_2 norm provides a lower bound on multiparty quantum communication complexity. Interestingly, while the γ_2 norm provides a lower bound on the model of quantum communication complexity even with entanglement in the two-party case, we are only able to show that the extension of this norm to the multiparty case is a lower bound in the model without entanglement.

Having generalized the norm-based bounds for multiparty communication complexity, the natural question is whether the corresponding version of Grothendieck's inequality holds. There are many results regarding high dimensional extensions of Grothendieck's inequality in the literature. A large portion of the results are negative, implying that for certain type of extensions, a corresponding Grothendieck type inequality does not hold [Ble01], [Smi88]. It is therefore somewhat surprising that in our case a strong version of Grothendieck's type inequality does hold, and the two generalized norms are closely related.

This result allows us to immediately transfer essentially all known lower bounds on randomized multiparty communication complexity to the quantum case. We now list some examples. Babai, Nisan, Szegedy [BNS89] adapted the discrepancy method, one of the earliest and most general techniques for showing lower bounds on randomized two-party complexity, to the multi-party case to obtain among other things a bound of $\Omega(n/2^{2k})$ on the k-party complexity of the generalized inner product function. The discrepancy method has seen many more applications [CT93], [Raz00], [FG05], [Cha07], in particular to show bounds of size $\Omega(n/2^k)$ on k-party complexity of explicit functions, the largest bounds currently known. The discrepancy method is a special case of the μ^α bound—in fact, it is exactly the limiting case μ^∞ [LS08b]— thus we are able to obtain that these bounds also hold in the quantum case.

More recently, a series of works have used the extension of the μ norm to the multiparty case, together with a generalization of the pattern matrix framework of Sherstov [She07], [She08], to show lower bounds that the discrepancy method cannot [LS08b], [CA08], [DPV08], [BHN08], including a bound of $\Omega(n^{1/(k+1)}/2^{2^k})$ on the k-party complexity of the disjointness function. These bounds also transfer to the quantum case.

To the best of our knowledge, these are the first lower bounds of any kind on quantum communication complexity in the number-on-the-forehead model. We should mention that there are bounds known on quantum NOF complexity in more restricted models: for example, [BARW08] show a bound of $\tilde{\Omega}(\sqrt{n})$ on the complexity of disjointness in the three-party one-way model, and a bound of $\tilde{\Omega}(n^{1/3})$ on disjointness in the case of three parties where the first player sends a message and then players two and three interact arbitrarily.

On the other hand, our results also mean that quantum and classical communication complexity cannot be separated with current techniques unless the number of players is either two or very large—two players as in this case we have techniques for showing lower bounds on randomized communication complexity, like the corruption bound and information theory methods, that can be larger than quantum communication complexity; very large as our lower bound on multiparty quantum communication complexity loses a multiplicative factor of $1/k$, which the classical bound does not.

Although we focus on the number-on-the-forehead model of communication complexity, all our results hold in a more

general setting. In particular, the corresponding results for the number-in-the-hand (NIH) model of multiparty communication complexity follow by a simple adjustment of the definitions and proofs.

II. Preliminaries

We let $[n] = \{1, \ldots, n\}$. For multiparty communication complexity it is convenient to work with tensors, the generalization of matrices to higher dimensions. If an element of a tensor A is specified by k indices, we say that A is a k-tensor. A tensor where all entries are in $\{-1, +1\}$ we call a sign tensor. For a function $f : X_1 \times \ldots \times X_k \to \{-1, +1\}$, we define the communication tensor corresponding to f to be a k-tensor A_f where $A_f[x_1, \ldots, x_k] = f(x_1, \ldots, x_k)$. We identify f with its communication tensor.

We use the shorthand $A \geq c$ to indicate that all of the entries of A are at least c. The Hadamard or entrywise product of two tensors A and B is denoted by $A \circ B$. Their inner product is denoted $\langle A, B\rangle = \sum_{x_1,\ldots,x_k} A[x_1, \ldots, x_k]B[x_1, \ldots, x_k]$. For vectors $u_1, \ldots, u_k \in \mathbb{R}^d$ we define a k-linear form $\langle u_1, \ldots, u_k\rangle = \sum_{j=1}^{d} \prod_{i=1}^{k} u_i(j)$. We write $\|u\|$ for the ℓ_2 norm of a vector u and use $\mathcal{S}^{d-1}$ to denote the set $\{u \in \mathbb{R}^d : \|u\| = 1\}$. For a norm φ, we denote its dual norm $\varphi^*(A) = \max_{B:\varphi(B)\leq 1}\langle A, B\rangle$.

III. Quantum multiparty communication complexity

We now define the NOF model of quantum multiparty communication complexity (see also [Ker07]). Let $f : (\{0,1\}^n)^k \to \{-1, +1\}$ be a function of k strings $x_1, \ldots, x_k$ where each $x_i \in \{0,1\}^n$. In the classical NOF model, player i receives as input all the strings except x_i. In the quantum setting, we can represent the NOF model as follows. If there are k players then we work in a Hilbert space $H_1 \otimes \cdots H_k \otimes C$, composed of $k+1$ many registers, one for each player in addition to a one qubit channel C. On the turn of player i, an arbitrary unitary independent of x_i is applied on $H_i \otimes C$ and acts as the identity everywhere else. The players take turns in an arbitrary order fixed at the beginning of the protocol.

We will only discuss the model without shared entanglement. Such a protocol begins in a pure state $|v^1\rangle \cdots |v^k\rangle|0\rangle$ independent of the input. The protocol outputs 1 with probability the norm squared of the projection of the final state onto the $|1\rangle$ state of the channel qubit. As we use a 1-qubit channel, the cost of a protocol is simply the number of rounds. For a sign tensor A, we define $Q_k^\epsilon(A)$ as the minimum cost of a k-player NOF protocol which computes A with error probability at most ϵ.

The next lemma extracts the property of quantum protocols which we use in our lower bound. This statement follows similar statements in the two-party case [Yao93], [Kre95], [LS07], and we defer the proof to the appendix.

Lemma 1: After c qubits of communication on input $(x_1, \ldots, x_k)$, the state of a quantum NOF protocol without shared entanglement can be written as

$$\sum_{r\in R} |v_r^1\rangle|v_r^2\rangle \cdots |v_r^k\rangle|0\rangle + \sum_{s\in S} |v_s^1\rangle|v_s^2\rangle \cdots |v_s^k\rangle|1\rangle,$$

where the set of vectors $\{v_r^t\}_{r\in R}$ is a function of $(x_1, \ldots, x_{t-1}, x_{t+1}, \ldots, x_k)$ and c.

Moreover, $\sum_{r\in R} \|v_r^t\|^2 + \sum_{s\in S} \|v_s^t\|^2 \leq 2^c$ for every $1 \leq t \leq k$. Note that vectors indexed by t belong to the space H_t.

IV. The multiparty norm

We describe the μ_k and γ_k norms and how they are applied to obtain lower bounds for classical and quantum bounded error communication complexity, respectively. Only the part about quantum communication is new. We discuss the classical case as well, for completeness.

A. Classical communication complexity and μ_k

In a series of recent works [LMSS07], [LS07], [LS08b], a general framework has been developed for showing lower bounds on communication complexity in terms of norms. The basic idea of this approach is that a successful communication protocol allows one to express the communication matrix, or tensor, as a linear sum of simpler objects. This set of simple objects depends upon the model under consideration. For example: in the deterministic two-party case, it is the set of combinatorial rectangles; in the deterministic multiparty NOF case, it is the set of cylinder intersections. The technique then actually bounds how efficiently the communication tensor can be expressed in terms of these simpler objects.

Let us see an instantiation of this framework. Consider the 2-player deterministic communication model. Let A be an $m \times n$ sign matrix and let $\mathcal{C}$ be the set of combinatorial rectangles on $[m] \times [n]$. Define the norm μ_2 by

$$\mu_2(A) = \min\{\sum_j |\alpha_j| : A = \sum_j \alpha_j \chi(C_j)\}$$

where each C_j belongs to $\mathcal{C}$, and $\chi(X)$ stands for the characteristic matrix of the subset $X \subseteq [m] \times [n]$.

The fact that a deterministic protocol for A that uses at most c bits of communication partitions A into at most 2^c combinatorial rectangles clearly gives that $\log \mu(A)$ is a lower bound on the deterministic communication complexity of A.

The randomized communication complexity of A can similarly be bounded by the following approximation variant of μ_2

$$\mu_2^\alpha(A) = \min_{A':1\leq A\circ A'\leq \alpha} \mu_2(A').$$

Denoting by ϵ the maximum allowed error, then the following bound holds for randomized communication complexity

$$R_\epsilon(A) \geq \log \mu_2^\alpha(A) - \log \alpha,$$

for $\alpha = \alpha(\epsilon) = \frac{1}{1-2\epsilon}$.

A similar norm (and its approximation variant) can be defined for other models of classical communication, and the corresponding lower bounds for deterministic and randomized communication complexity will hold accordingly. What changes between different models is the set of simple objects $\mathcal{C}$ which reflects structural properties of the underlying model.

In particular, one can apply the above principles with the basic sets being cylinder intersections. The corresponding norm is the norm induced by k cylinder intersections, denoted by μ_k. As shown by Lee and Shraibman [LS08b] using the framework described above, the μ_k norm and its approximation variant yield lower bounds on classical NOF communication complexity (see also [CA08] where a formulation of μ_k dual to the one described here is used).

Very closely related to the μ_k norm, and sometimes more convenient to analyze [Raz00], [FG05], is the ν_k norm where one considers $\{-1,+1\}$ valued functions rather than $\{0,1\}$ valued functions.

Definition 2: Let A be a k-tensor

$$\nu_k(A) = \min\{\sum_j |\alpha_j| : A = \sum_j \alpha_j C_j\}$$

where each C_j can be written as $C_j[x_1,\ldots,x_k] = \prod_{t=1}^k \phi^t(x_1,\ldots,x_k)$ for $\{-1,+1\}$ valued functions ϕ^t which are independent of x_t.

It is not too difficult to show that the μ_k and ν_k are closely related: $\nu_k(A) \leq \mu_k(A) \leq 2^k \nu_k(A)$.

B. Quantum communication complexity and γ_k

To lower bound quantum NOF communication complexity, we first want to identify the set of simple objects into which a successful protocol decomposes the communication tensor. This is indicated by Lemma 1. Formally, we define the set of simple objects as

$$\begin{aligned}\mathcal{C}_k = \{&C : C[x_1,\ldots,x_k] = \\ &\langle \phi^1(x_2,\ldots,x_k),\ldots,\phi^k(x_1,\ldots,x_{k-1})\rangle \\ &\text{ and } \|\phi^t(x_1,\ldots,x_k)\| \leq 1 \text{ for all } t,x_1,\ldots,x_k\}\end{aligned}$$

where each $\phi^t(x_1,\ldots,x_k)$ is a vector independent of x_t. The $\gamma_{2,k}$ norm is then defined as follows

Definition 3:

$$\gamma_{2,k}(A) = \min\left\{\sum_j |\alpha_j| : A = \sum_j \alpha_j C_j, \text{ where } C_j \in \mathcal{C}_k\right\}$$

When $k=2$ and A is a matrix, this agrees with the γ_2 norm of [LMSS07]. Note that the intention of the 2 in γ_2 is not to indicate 2 players, but rather that the normalization is taken with respect to the ℓ_2 norm. For this reason, we use the notation $\gamma_{2,k}$ to indicate that we normalize with respect to the ℓ_2 norm but consider the k-fold inner product. One could alternatively consider this definition with respect to any ℓ_p norm. For the rest of the paper, however, we stick to the ℓ_2 norm and drop the subscript of 2 to simply write γ_k.

To work with protocols with some probability of error, we will also use an approximate version of the γ_k norm.

Definition 4 (Approximate quantum norm): Let $\alpha \geq 1$, and A be a sign k-tensor.

$$\gamma_k^\alpha(A) = \min_{A':1\leq A\circ A'\leq \alpha} \gamma_k(A').$$

Observe that $\gamma_k^\alpha(A)$ is a decreasing function of α.

The γ_k norm can be used to lower bound communication complexity in the quantum number-on-the-forehead model as follows.

Theorem 5: Let A be a sign k-tensor. Then

$$Q_k^\epsilon(A) \geq \frac{\log \gamma_k^{\alpha_\epsilon}(A) - \log \alpha_\epsilon - 2}{k},$$

where $\alpha_\epsilon = 1/(1-2\epsilon)$.

Proof: Let P be the k-tensor whose entry $(x_1,\ldots,x_k)$ is the probability that the protocol outputs 1 on input $(x_1,\ldots,x_k)$. Let

$$\left(\sum_{s\in S} |v_s^1\rangle \cdots |v_s^k\rangle\right) \otimes |1\rangle,$$

where $|v_s^t\rangle$ is independent of x_t, be the projection of the final state of the algorithm on input $(x_1,\ldots,x_k)$ onto the $|1\rangle$ state of the channel. As the probability that the algorithm outputs 1 on $(x_1,\ldots,x_k)$ is given by the norm squared of this vector, we have

$$P[x_1,\ldots,x_k] = \sum_{s_1,s_2\in S} \langle v_{s_1}^1, v_{s_2}^1\rangle\langle v_{s_1}^2, v_{s_2}^2\rangle \cdots \langle v_{s_1}^k, v_{s_2}^k\rangle. \tag{1}$$

Let us now upper bound $\gamma_k(P)$. We have

$$\begin{aligned}\sum_{s_1,s_2\in S} \langle v_{s_1}^t, v_{s_2}^t\rangle^2 &\leq \left(\sum_{s\in S} \|v_s^t\|^2\right)^2 \\ &\leq 2^{2c},\end{aligned}$$

as Lemma 1 implies in particular that $\sum_{s\in S}\langle v_s^t, v_s^t\rangle \leq 2^c$ for each t.

This means that

$$\frac{1}{2^{ck}} \sum_{s_1,s_2\in S} \langle v_{s_1}^1, v_{s_2}^1\rangle\langle v_{s_1}^2, v_{s_2}^2\rangle \cdots \langle v_{s_1}^k, v_{s_2}^k\rangle \in \mathcal{C}_k,$$

and so $\gamma_k(P) \leq 2^{ck}$.

Now as the protocol has error probability at most ϵ, we have that if $A[x_1,\ldots,x_k] = 1$ then

$P[x_1,\ldots,x_k] \geq 1-\epsilon$ and if $A[x_1,\ldots,x_k] = -1$ then $P[x_1,\ldots,x_k] \leq \epsilon$. Thus the matrix $P' = \alpha_\epsilon(2P - J)$, where J is the all one tensor, satisfies $1 \leq A \circ P' \leq \alpha_\epsilon$. Therefore we conclude

$$\gamma_k^{\alpha_\epsilon}(A) \leq \gamma_k(P') \leq \alpha_\epsilon(2^{ck+1}+1)$$

which gives the theorem. ∎

V. A Grothendieck type inequality in high dimension

As mentioned earlier, all existing lower bounds in the randomized NOF model for more than two players can be shown using the μ_k norm. In this section, we show that the γ_k and μ_k norms are equivalent up to a factor of C^k for some universal constant C. Thus we can immediately transfer all randomized NOF lower bounds to the quantum case, up to the loss of an additive $O(k)$ factor, and the multiplicative factor of $\frac{1}{k}$ in Theorem 5 which does not appear in the randomized case. We do this by presenting a Grothendieck type inequality for γ_k and μ_k. Our inequality holds in a more general framework, as described next.

Fix a family of partitions $\mathcal{P} = \{P_j\}_{j=1}^k$ of $\mathbb{N}^k$, and let $N = [n_1] \times [n_2] \times \cdots \times [n_k]$. For $j = 1 \ldots k$, and $d \in \mathbb{N}$, let $\mathcal{F}_j$ be the family of all functions $f : N \to \mathcal{S}^{d-1}$, which are constant on each set in the partition P_j. We define a semi-norm Φ_d as follows

$$\Phi_d(A) = \sup_{\substack{f_j \in \mathcal{F}_j \\ j=1\ldots k}} \sum_{I \in N} A[I]\langle f_1(I), f_2(I), \ldots, f_k(I)\rangle.$$

Where A is a $n_1 \times n_2 \times \cdots \times n_k$ real tensor. We prove that for any fixed tensor A, $\Phi_d(A)$ depends very weakly on d.

Theorem 6: For every k-tensor A, let $\Phi(A) = \sup_d \Phi_d(A)$, then

$$\Phi_1(A) \leq \Phi(A) \leq C(k)\Phi_1(A).$$

We first prove the theorem with $C(k) = (C \log k)^{k/2}$. We then provide a slightly more involved proof in Appendix B that gives the statement of the theorem with $C(k) = C^k$ for an absolute constant C.

Note that in the above theorem Φ_1 and Φ_d are defined with respect to the same family of partitions. Observe that when the underlying family of partitions is $\mathcal{P} = \{P_j\}_{j=1}^k$ where P_j partitions $\mathbb{N}^k$ according to all coordinates except the jth coordinate, then $\Phi_1 = \nu_k^*$ and $\Phi = \gamma_k^*$. Hence as $\mu_k^* \leq \nu_k^* \leq 2^k \mu_k^*$ we obtain the following corollary.

Corollary 7: For every k-tensor A,

$$\gamma_k(A) \leq \mu_k(A) \leq C^k \gamma_k(A),$$

for some absolute constant C.

Another interesting special case of Theorem 6 is when P_j partition $\mathbb{N}^k$ according to the j-coordinate. In this case $\Phi_d(A)$ takes the form

$$\sup_{f_j:[n_j]\to\mathcal{S}^{d-1}} \sum_{i_1=1,\ldots,i_k=1}^{n_1,\ldots,n_k} A[i_1,\ldots,i_k]\langle f_1(i_1),\ldots,f_k(i_k)\rangle.$$

This instance of Theorem 6 is related to the NIH model of multiparty communication complexity. It was first proved by Blei [Ble79], and with constant C^k by Tonge [Ton78]. In fact, it is possible to reduce the instance of Theorem 6 related to the NOF model to the NIH case. We feel, however, that Theorem 6 is superior to statements in [Ble79], [Ton78] as it works in the broadest generality and moreover the proof we give provides the clearest insight to the machinery behind the scenes.

The proof is a generalization of the proofs of the 2-dimensional Grothendieck's inequality in [DJT95] and [JL01]. We think that this is currently the most elegant and accessible proof for the 2-dimensional Grothendieck's inequality. We note that the generalization is not straightforward, and requires some additional ideas. The main difficulty stems from the fact that the problem is no longer unitarily invariant [1]; it is important, for example, that we use $\mathbb{R}^d$ as the underlying Hilbert space and not L_2 over a probability space, say.

A. Auxiliary lemmas

To prove Theorem 6 we require a few facts from probability theory. We describe these next.

Let $\{g_{i,j}\}$ for $j = 1 \ldots d$ and $i = 1 \ldots k-1$ be independent Bernoulli random variables, and let $g_{k,j} = \prod_{i=1}^{k-1} g_{i,j}$. Notice that

1) $\mathbb{E}(g_{i,j}) = 0$ for every $i = 1 \ldots k$ and $j = 1 \ldots d$.
2) $\mathbb{E}(g_{i,j}^2) = 1$ for $i = 1 \ldots k-1$ and $j = 1 \ldots d$.
3) Let π be a function $\pi : [k] \to [d]$. Then $\mathbb{E}(\prod_i g_{i,\pi(i)}) = 0$ if the image of π contains at least two elements, and otherwise $\mathbb{E}(\prod_i g_{i,\pi(i)}) = 1$.

Lemma 8: Let $g_1, \ldots, g_d$ be independent Bernoulli random variables. For a vector $u \in \mathcal{S}^{d-1}$ consider the random variable $G(u) = \sum u_i g_i$.

Furthermore, for some constant T, denote by $\bar{G}(u)$ the random variable which is equal to $G(u)$ whenever $|G(u)|$ is greater than T and zero otherwise. Then

1) $\mathbb{E}(|G(u)|^2) = 1$.
2) If $T \geq 2$
$$\mathbb{E}(|\bar{G}(u)|^2) \leq 3T^2 e^{-T^2/2}.$$

Proof: The first part of the lemma follows from the following simple calculation

$$\mathbb{E}(|G(u)|^2) \quad = \quad \mathbb{E}((\sum u_i g_i)^2) = \sum u_i^2 = 1.$$

For the second part of the lemma, recall that for every random variable X

$$\mathbb{E}(|X|^k) = k \int_0^\infty t^{k-1} \Pr(|X| > t) dt.$$

[1] Conventional (bivariate) inner product is unitarily invariant, i.e. it holds that $\langle x, y\rangle = \langle Ux, Uy\rangle$ for every pair of vectors x and y and any unitary transformation U. The k-dimensional Grothendieck type inequality involves the k-linear form $\langle x_1, \ldots, x_k\rangle$. For $k \geq 3$ this form is no longer unitarily invariant.

(See, for example, page 42 of [Dur05].) Also by Hoeffding's inequality [Hoe63], the random variable $G(u)$ is sub-Gaussian, with constant $1/2$. i.e., $\Pr(|G(u)| > t) \leq 2e^{-t^2/2}$.

Using these two facts we get that the 2nd moment of $|\bar{G}(u)|$ is bounded by

$$\begin{aligned}
\mathbb{E}(|\bar{G}(u)|^2) &= 2\int_0^\infty t\Pr(|\bar{G}(u)| > t)dt \\
&= 2\int_0^T t\Pr(|\bar{G}(u)| > t)dt \\
&+ 2\int_T^\infty t\Pr(|\bar{G}(u)| > t)dt \\
&\leq 4\int_0^T te^{-T^2/2}dt + 4\int_T^\infty te^{-t^2/2}dt \\
&\leq 2T^2e^{-T^2/2} + 4e^{-T^2/2} \\
&\leq 3T^2e^{-T^2/2},
\end{aligned}$$

when $T \geq 2$. ∎

B. Proof of Theorem 6

Let $g_{i,j}$ be random variables as defined in Section V-A. For a vector $u \in \mathcal{S}^{d-1}$ and $i \in [k]$ consider the random variable $G_i(u) = \sum u_j g_{i,j}$. Observe that for any vectors $u_1, \ldots, u_k \in \mathbb{R}^d$,

$$\begin{aligned}
\mathbb{E}(\prod_i G_i(u_i)) &= \mathbb{E}(\prod_i(\sum_j u_{i,j}g_{i,j})) \\
&= \sum_{j_1,\ldots,j_k}\prod_t u_{t,j_t}\mathbb{E}(\prod_t g_{t,j_t}) \\
&= \langle u_1, \ldots, u_k\rangle.
\end{aligned}$$

Let $N = [n_1] \times \cdots \times [n_k]$. Then $\Phi_d(A)$ can be equivalently written as

$$\begin{aligned}
\Phi_d(A) &= \sup_{\substack{f_i\in\mathcal{F}_i\\ i=1\ldots k}}\sum_{I\in N} A[I]\langle f_1(I), f_2(I), \ldots, f_k(I)\rangle \\
&= \sup_{f_i\in\mathcal{F}_i : i=1\ldots k}\sum_{I\in N} A[I]\mathbb{E}(\prod_{i=1}^k G_i(f_i(I))). \quad (2)
\end{aligned}$$

Fix a constant T. For any vector $u \in \mathbb{R}^m$ and $i \in [k]$, write the random variable $G_i(u)$ as a sum of two random variables $G_i(u) = G_i^1(u) + G_i^2(u)$ such that $G_i^1(u)$ is equal to $G_i(u)$ if $|G_i(u)| \leq T$ and is zero otherwise, and $G_i^2(u) = G_i(u) - G_i^1(u)$. Then, the right hand side of (2) is bounded by

$$\sum_{b\in\{1,2\}^k}\sup_{f_i\in\mathcal{F}_i : i=1\ldots k}\sum A[I]\mathbb{E}(\prod_{i=1}^k G_i^{b_i}(f_i(I))).$$

When $b = (1, 1, \ldots, 1)$ we can bound the corresponding expression as follows

$$\sup_{\substack{f_i\in\mathcal{F}_i\\ i=1\ldots k}}\sum A[I]\mathbb{E}(\prod_{i=1}^k G_i^1(f_i(I))) =$$

$$\sup_{\substack{f_i\in\mathcal{F}_i\\ i=1\ldots k}}\mathbb{E}(\sum A[I]\prod_{i=1}^k G_i^1(f_i(I))) \leq$$

$$\mathbb{E}(\sup_{\substack{f_i\in\mathcal{F}_i\\ i=1\ldots k}}\sum A[I]\prod_{i=1}^k G_i^1(f_i(I))) \leq$$

$$T^k \cdot \Phi_1(A).$$

To bound the rest of the terms we use the Fourier representation of the random variables $G_i^{b_i}(f_i(I))$. For a subset $S \subset [d]$ we denote by W_S the corresponding Walsh function (or character). Fix $b \in \{1, 2\}^k$, the Fourier representation of $G_i^{b_i}(f_i(I))$ is

$$G_i^{b_i}(f_i(I)) = \sum_S \hat{G}_{i,S}(f_i(I))W_S.$$

Here we think of the random variable as a function from $\{\pm1\}^d$ to $\mathbb{R}$. For convenience we identify $\{\pm1\}^d$ with $\mathbb{Z}_2^d$ with addition modulo 2 and for $\mathbf{x} \in \mathbb{Z}_2^d$ write $\mathbf{x} = (x_1, \ldots, x_k)$. Denote by $\triangle$ set theoretic symmetric difference, we see that

$$\mathbb{E}(\prod_{i=1}^k G_i^{b_i}(f_i(I))) =$$

$$\sum_{S_1,\ldots,S_k}\prod_i \hat{G}_{i,S_i}(f_i(I))\mathbb{E}_{\mathbf{x}}(W_{S_1}(x_1)\cdots W_{S_k}(\sum_{j=1}^k x_j)) =$$

$$\sum_{S_1,\ldots,S_k}\prod_i \hat{G}_{i,S_i}(f_i(I))\mathbb{E}_{\mathbf{x}}(W_{S_1}(x_1)\cdots \prod_{j=1}^{k-1} W_{S_k}(x_j)) =$$

$$\sum_{S_1,\ldots,S_k}\prod_i \hat{G}_{i,S_i}(f_i(I))\mathbb{E}_{\mathbf{x}}(\prod_{j=1}^{k-1} W_{S_j\triangle S_k}(x_j)) =$$

$$\sum_S\prod_i \hat{G}_{i,S}(f(I))$$

Therefore

$$\begin{aligned}
&\sup_{\substack{f_i\in\mathcal{F}_i\\ i=1\ldots k}}\sum A[I]\mathbb{E}(\prod_{i=1}^k G_i^{b_i}(f_i(I))) = \\
&\sup_{\substack{f_i\in\mathcal{F}_i\\ i=1\ldots k}}\sum A[I]\langle \hat{G}_1(f_1(I)), \ldots, \hat{G}_k(f_k(I))\rangle,
\end{aligned} \quad (3)$$

where $\hat{G}_i(f_i(I))$ is a vector in $\mathbb{R}^{2^d}$, whose coordinate corresponding to $S \subset [d]$ is equal to $\hat{G}_{i,S}(f_i(I))$.

Note that the right hand side of (3) is bounded from above by Φ_{2^d} times the product of the lengths of the vectors $\hat{G}_i(f_i(I))$ for $i = 1 \ldots k$. But by Parseval's identity $\|\hat{G}_i(f_i(I))\|_2 = \mathbb{E}(G_i^{b_i}f_i(I)^2)^{1/2}$ and therefore Lemma 8

provides a bound on the length of these vectors. Assume that exactly L of the entries of b are equal to 2, this give us

$$\sup_{\substack{f_i \in \mathcal{F}_i \\ i=1\ldots k}} \sum A[I]\mathbb{E}(\prod_{i=1}^{k} G_i^{b_i}(f_i(I))) \leq (\sqrt{3}Te^{-T^2/4})^L \Phi_{2^d}(A). \tag{4}$$

Finally, for large enough T

$$\begin{aligned}\Phi_d(A) &\leq T^k \Phi_1(A) + \Phi_{2^d}(A) \sum_{L>0} \binom{k}{L} (\sqrt{3}Te^{-T^2/4})^L \\ &= T^k \Phi_1(A) + \left[(1+\sqrt{3}Te^{-T^2/4})^k - 1\right] \Phi_{2^d}(A) \\ &\leq T^k \Phi_1(A) + \sqrt{12} \cdot kTe^{-T^2/4} \Phi_{2^d}(A).\end{aligned}$$

The last inequality is because $(1+x)^k \leq 1 + 2kx$ for $0 \leq x \leq \frac{1}{2(k-1)}$.

Taking large enough d (so that both $\Phi_d(A)$ and $\Phi_{2^d}(A)$ are basically $\Phi(A)$) and $T \sim \sqrt{\log k}$ we get the statement of the theorem.

ACKNOWLEDGMENTS

GS is supported by the Israel Science Foundation. This work conducted while TL was at Rutgers University, supported by a NSF Mathematical Sciences Postdoctoral Fellowship.

REFERENCES

[AA05] S. Aaronson and A. Ambainis. Quantum search of spatial regions. *Theory of Computing*, 1:47–79, 2005.

[Abe78] H. Abelson. Lower bounds on information transfer in distributed computations. In *Proceedings of the 19th IEEE Symposium on Foundations of Computer Science*, pages 151–158. IEEE, 1978.

[BARW08] A. Ben-Aroya, O. Regev, and R. de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing. In *Proceedings of the 49th IEEE Symposium on Foundations of Computer Science*. IEEE, 2008. arXiv:0705.3806 [quant-ph].

[BHN08] P. Beame and D. Huynh-Ngoc. Multiparty communication complexity of AC^0. Technical Report TR-08-082, ECCC, 2008.

[Ble79] R. C. Blei. Multidimensional extensions of the Grothendieck inequality and applications. *Ark. Mat.*, 17(1):51–68, 1979.

[Ble01] R. Blei. *Analysis in integer and fractional dimensions*, volume 71 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2001.

[BNS89] L. Babai, N. Nisan, and M. Szegedy. Multiparty protocols and Logspace-hard pseudorandom sequences. In *Proceedings of the 21st ACM Symposium on the Theory of Computing*, pages 1–11. ACM, 1989.

[BPS06] P. Beame, T. Pitassi, and N. Segerlind. Lower bounds for Lovász-Schrijver systems and beyond follow from multiparty communication complexity. *SIAM Journal on Computing*, 37(3):845–869, 2006.

[BW01] H. Buhrman and R. de Wolf. Communication complexity lower bounds by polynomials. In *Proceedings of the 16th IEEE Conference on Computational Complexity*, pages 120–130, 2001.

[BYJKS04] Z. Bar-Yossef, T. Jayram, R. Kumar, and D. Sivakumar. Information statistics approach to data stream and communication complexity. *Journal of Computer and System Sciences*, 68(4):702–732, 2004.

[CA08] A. Chattopadhyay and A. Ada. Multiparty communication complexity of disjointness. Technical Report TR-08-002, ECCC, 2008.

[CFL83] A. Chandra, M. Furst, and R. Lipton. Multi-party protocols. In *Proceedings of the 15th ACM Symposium on the Theory of Computing*, pages 94–99. ACM, 1983.

[Cha07] A. Chattopadhyay. Discrepancy and the power of bottom fan-in depth-three circuits. In *Proceedings of the 48th IEEE Symposium on Foundations of Computer Science*, pages 449–458. IEEE, 2007.

[CSWY01] A. Chakrabarti, Y. Shi, A. Wirth, and A. Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science*, pages 270–278. IEEE, 2001.

[CT93] F. Chung and P. Tetali. Communication complexity and quasi-randomness. *SIAM Journal Discrete Math*, 6:110–123, 1993.

[DJT95] J. Diestel, H. Jarchow, and A. Tonge. *Absolutely summing operators*, volume 43 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.

[DPV08] M. David, T. Pitassi, and E. Viola. Improved separations between nondeterministic and randomized multiparty communication. In *APPROX-RANDOM*, volume 5171 of *Lecture Notes in Computer Science*, pages 371–384. Springer, 2008.

[Dur05] R. Durrett. *Probability: Theory and Examples*. Duxbury Press, 2005.

[FG05] J. Ford and A. Gál. Hadamard tensors and lower bounds on multiparty communication complexity. In *Proceedings of the 32th International Colloquium On Automata, Languages and Programming*, pages 1163–1175, 2005.

[HG91] J. Håstad and M. Goldmann. On the power of small-depth threshold circuits. *Computational Complexity*, 1:113–129, 1991.

[Hoe63] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[JL01] W. Johnson and J. Lindenstrauss. Basic concepts in the geometry of Banach spaces. In *Handbook of the geometry of Banach spaces, Vol. I*, pages 1–84. North-Holland, Amsterdam, 2001.

[Ker07] I. Kerenidis. Quantum multiparty communication complexity and circuit lower bounds. *Mathematical structures in computer sciences*, 2007.

[Kla01] H. Klauck. Lower bounds for quantum communication complexity. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science*. IEEE, 2001.

[KN97] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[Kre95] I. Kremer. Quantum communication. Technical report, Hebrew University of Jerusalem, 1995.

[KS87] B. Kalyanasundaram and G. Schnitger. The probabilistic communication complexity of set intersection. In *Proceedings of the 2nd Annual Conference on Structure in Complexity Theory*, pages 41–49, 1987.

[LMSS07] N. Linial, S. Mendelson, G. Schechtman, and A. Shraibman. Complexity measures of sign matrices. *Combinatorica*, 27(4):439–463, 2007.

[LS88] L. Lovász and M. Saks. Möbius functions and communication complexity. In *Proceedings of the 29th IEEE Symposium on Foundations of Computer Science*, pages 81–90. IEEE, 1988.

[LS07] N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms. In *Proceedings of the 39th ACM Symposium on the Theory of Computing*, pages 699–708. ACM, 2007.

[LS08a] T. Lee and A. Shraibman. An approximation algorithm for approximation rank. Technical Report arXiv:0809.2093 [cs.CC], arXiv, 2008.

[LS08b] T. Lee and A. Shraibman. Disjointness is hard in the multiparty number-on-the-forehead model. In *Proceedings of the 23rd IEEE Conference on Computational Complexity*, pages 81–91. IEEE, 2008.

[Raz92] A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106:385–390, 1992.

[Raz95] R. Raz. Fourier analysis for probabilistic communication complexity. *Computational Complexity*, 5(3/4):205–221, 1995.

[Raz99] R. Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the 31st ACM Symposium on the Theory of Computing*, pages 358–367. ACM, 1999.

[Raz00] R. Raz. The BNS-Chung criterion for multi-party communication complexity. *Computational Complexity*, 9(2):113–122, 2000.

[Raz03] A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145–159, 2003.

[She07] A. Sherstov. Separating AC^0 from depth-2 majority circuits. In *Proceedings of the 39th ACM Symposium on the Theory of Computing*, pages 294–301. ACM, 2007.

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

[Smi88] R. R. Smith. Completely bounded multilinear maps and grothendieck's inequality. *Bull. London Math. Soc.*, 20(6):606–612, 1988.

[Ton78] A. Tonge. The von Neumann inequality for polynomials in several Hilbert-Schmidt operators. *J. London Math. Soc. (2)*, 18(3):519–526, 1978.

[Yao79] A. Yao. Some complexity questions related to distributive computing. In *Proceedings of the 11th ACM Symposium on the Theory of Computing*, pages 209–213. ACM, 1979.

[Yao83] A. Yao. Lower bounds by probabilistic arguments. In *Proceedings of the 24th IEEE Symposium on Foundations of Computer Science*, pages 420–428, 1983.

[Yao93] A. Yao. Quantum circuit complexity. In *Proceedings of the 34th IEEE Symposium on Foundations of Computer Science*, pages 352–360. IEEE, 1993.

APPENDIX A.
PROOF OF LEMMA 1

For convenience, we restate the lemma here.

Lemma 1: After c qubits of communication on input $(x_1, \ldots, x_k)$, the state of a quantum NOF protocol can be written as

$$\sum_{r \in R} |v_r^1\rangle |v_r^2\rangle \cdots |v_r^k\rangle |0\rangle + \sum_{s \in S} |v_s^1\rangle |v_s^2\rangle \cdots |v_s^k\rangle |1\rangle,$$

where the set of vectors $\{v_r^t\}_{r \in R}$ is a function of $(x_1, \ldots, x_{t-1}, x_{t+1}, \ldots, x_k)$ and c.

Moreover, $\sum_{r \in R} \|v_r^t\|^2 + \sum_{s \in S} \|v_s^t\|^2 \leq 2^c$ for every $1 \leq t \leq k$. Note that vectors indexed by t belong to the space H_t.

Proof: We prove by induction. The statement clearly holds after 0 qubits of communication. Assume $c > 0$ qubits were transmitted, then by the induction hypothesis we have some state

$$\sum_{r \in R} |v_r^1\rangle |v_r^2\rangle \cdots |v_r^k\rangle |0\rangle + \sum_{s \in S} |v_s^1\rangle |v_s^2\rangle \cdots |v_s^k\rangle |1\rangle.$$

with $\sum_{r \in R} \|v_r^t\|^2 + \sum_{s \in S} \|v_s^t\|^2 \leq 2^c$ for every t.

For simplicity, suppose it is the turn of player 1, who applies a unitary which does not depend on x_1 and acts

as identity everywhere except for the first register and the channel. We can then write the new state as

$$\sum_{r\in R}|v^1_{r0}\rangle|v^2_r\rangle\cdots|v^k_r\rangle|0\rangle+|v^1_{r1}\rangle|v^2_r\rangle\cdots|v^k_r\rangle|1\rangle+$$
$$\sum_{s\in S}|v^1_{s0}\rangle|v^2_s\rangle\cdots|v^k_s\rangle|0\rangle+|v^1_{s1}\rangle|v^2_s\rangle\cdots|v^k_s\rangle|1\rangle=$$
$$\sum_{i\in R\cup S}|v^1_{i0}\rangle|v^2_i\rangle\cdots|v^k_i\rangle|0\rangle+\sum_{i\in R\cup S}|v^1_{i1}\rangle|v^2_i\rangle\cdots|v^k_i\rangle|1\rangle$$

For every $t=2,\ldots,k$

$$\sum_{i\in R\cup S}\|v^t_i\|^2\le\sum_{r\in R}\|v^t_r\|^2+\sum_{s\in S}\|v^t_s\|^2\le 2^c$$

And for $t=1$ we get

$$\sum_{i\in R\cup S}\|v^1_{i0}\|^2+\sum_{i\in R\cup S}\|v^1_{i1}\|^2=\sum_{i\in R\cup S}\|v^1_i\|^2\le 2^c.$$

■

Appendix B. Improving the constant

We start again from equation (2), using the same notation as we had before. We first claim that for $s=0,\ldots,k-1$

$$\Phi_d(A)\le$$
$$C^s\sup_{\substack{f_i\in\mathcal{F}_i\\ i=1\ldots k}}\sum_I A[I]\mathbb{E}(\prod_{i=1}^{s}G^1_i(f_i(I))\prod_{j=s+1}^{k}G_j(f_j(I))).$$

for some absolute constant C. The proof is by induction on s. The case $s=0$ is trivial. For simplicity we show the induction step for $s=1$. By linearity of expectation

$$\mathbb{E}(\prod_{i=1}^{k}G_i(f_i(I)))=\sum_{b=1,2}\mathbb{E}(G^b_1(f_1(I))\prod_{i=2}^{k}G_i(f_i(I))). \quad (5)$$

Consider the Fourier representation of the random variables $G_i(f_i(I))$ and recall that they are defined as the linear sum of Bernoulli (=Rademacher) random variables. This means that its Fourier representation is the linear sum of Rademacher functions (with the same coefficients) and the coefficients of all the other characters are zero. A Rademacher function is a function of the form $f(\epsilon_1,\ldots,\epsilon_k)=\epsilon_j$ for some $1\le j\le k$. This is not necessarily the case with $G^2_1(f_1(I))$; its Fourier expansion may involve the other characters. But the orthogonality properties of the Fourier characters, as used in the proof of Theorem 6, implies that we can ignore all Fourier coefficients of $G^2_1(f_1(I))$ for non-Rademacher functions to achieve another random variable $G_1(r_1(I))$ without changing the expectation. That is,

$$\mathbb{E}(G^2_1(f_1(I))\prod_{i=2}^{k}G_i(f_i(I)))=\mathbb{E}(G_1(r_1(I))\prod_{i=2}^{k}G_i(f_i(I))). \quad (6)$$

Now $G_1(r_1(I))$ is the sum of Bernoulli random variables, and by Lemma 8

$$\|r_1(I)\|_2=\|\hat{r}_1(I)\|_2\le\|\hat{f}_1(I)\|_2\le\mathbb{E}(G^2_1(f_1(I)))^{1/2}\le\sqrt{3}Te^{-T^2/4}.$$

Therefore we can move the second term in the right hand side of (5) to the left hand side, and since we consider the supremum of the linear sum of such expectations, we get the desired result. Hence,

$$\Phi_d(A)\le$$
$$C^{k-1}\sup_{f_i\in\mathcal{F}_i:i=1\ldots k}\sum_I A[I]\mathbb{E}(\left(\prod_{i=1}^{k-1}G^1_i(f_i(I))\right)G_k(f_k(I))).$$

For the last step we cannot use the analogue of (6) since in the right hand side we do not have any more a variable whose Fourier expansion involves only the Rademacher functions. However, we can use the same argument as in the end of the proof of Theorem 6. Since we are using it to eliminate only one term (rather than 2^k-1, as in the proof of Theorem 6) we only pay at most another factor of C^{k-1}. This concludes the proof.