The NOF Multiparty Communication Complexity of Composed Functions*

Anil Ada[1], Arkadev Chattopadhyay[2], Omar Fawzi[1], and Phuong Nguyen[2]

[1] Department of Computer Science, McGill University.
[2] Department of Computer Science, University of Toronto.

Abstract. We study the k-party 'number on the forehead' communication complexity of composed functions $f \circ g$, where $f : \{0,1\}^n \to \{\pm 1\}$, $g : \{0,1\}^k \to \{0,1\}$ and for $(x_1,\ldots,x_k) \in (\{0,1\}^n)^k$, $f \circ g(x_1,\ldots,x_k) = f(\ldots,g(x_{1,i},\ldots,x_{k,i}),\ldots)$. We show that there is an $O(\log^3 n)$ cost simultaneous protocol for $\mathrm{SYM} \circ g$ when $k > 1+\log n$, SYM is any symmetric function and g is *any function*. Previously, an efficient protocol was only known for $\mathrm{SYM} \circ g$ when g is symmetric and "compressible". We also get a non-simultaneous protocol for $\mathrm{SYM} \circ g$ of cost $O(n/2^k \cdot \log n + k \log n)$ for any $k \geq 2$.
In the setting of $k \leq 1+\log n$, we study more closely functions of the form $\mathrm{MAJORITY} \circ g$, $\mathrm{MOD}_m \circ g$, and $\mathrm{NOR} \circ g$, where the latter two are generalizations of the well-known and studied functions Generalized Inner Product and Disjointness respectively. We characterize the communication complexity of these functions with respect to the choice of g. In doing so, we answer a question posed by Babai et al. (*SIAM Journal on Computing, 33:137–166, 2004*) and determine the communication complexity of $\mathrm{MAJORITY} \circ \mathrm{QCSB}_k$, where QCSB_k is the "quadratic character of the sum of the bits" function.

1 Introduction

The 'number on the forehead' (NOF) model of communication complexity was introduced by Chandra, Furst and Lipton [9] who used it to obtain branching program lower bounds. In this model, k players wish to evaluate a function $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{\pm 1\}$ on a given input $(x_1,\ldots,x_k)$. The input is distributed among the players in a way that Player i sees every x_j for $j \neq i$. This scenario is visualized as x_i being written on the forehead of Player i. In order to compute $F(x_1,\ldots,x_k)$, the players communicate by means of broadcasting, according to a protocol which they have agreed upon beforehand. The goal is to compute $F(x_1,\ldots,x_k)$ by communicating as few bits as possible. Note that for $k = 2$, this model is equivalent to the standard two player model introduced by Yao [39]. We are mainly interested in the case $\mathcal{X}_i = \{0,1\}^n$ for all i. Here, every function can be trivially computed using $n+1$ bits of communication, and protocols of cost at most polylogarithmic in n are considered to be efficient. Deterministic, non-deterministic, randomized and quantum communication complexity models naturally manifest themselves in this setting. The overlap of information among the players makes the NOF model interesting, powerful and fruitful in terms of applications. Apart

* Full version of the paper is given in the Appendix.

from the aforementioned application in branching programs, this model also has important applications in circuit complexity, proof complexity and pseudorandom generators.

The class ACC^0 represents functions computable by polynomial-size, constant-depth circuits with unbounded fan-in AND, OR, NOT and MOD_m gates. Showing NP is not in ACC^0 is one of the frontiers in complexity theory. It is well known that a function in ACC^0 has a $\text{polylog}(n)$ k-party deterministic communication complexity, where k is $\text{polylog}(n)$ [17, 7]. In fact the protocol is *simultaneous* where all the players, without interacting, speak once to an external referee who determines the output based only on the messages she receives. Proving that a function in NP requires super-polylogarithmic communication in the simultaneous model for polylogarithmic number of players would result in a major breakthrough. Currently no non-trivial lower bound is known for an explicit function for $k = \log n$ and this has proven to be a formidable barrier. Despite intense effort, even the 3 player model is far from being well understood and many important problems that are solved in the 2 player setting remain open in the 3 player setting. For example, in the 3 player setting, there is no known explicit function that is hard in the deterministic model but easy in the randomized model. On the other hand, the *equality* function is a canonical example of such a function in the 2 player setting. More relevant to our work, no characterization results are known for 3 player *composed* functions, which we discuss further below.

Most of the well known and studied functions in the standard two party as well as the multiparty model have the following 'composed' structure. Let $f : \{0,1\}^n \to \{\pm 1\}$ be a function and $\overrightarrow{g} = (g_1, \dots, g_n)$ be a vector of functions $g_i : \{0,1\}^k \to \{0,1\}$. Define $f \circ \overrightarrow{g}(x_1, \dots, x_k) = f(\dots, g_i(x_{1,i}, x_{2,i}, \dots, x_{k,i}), \dots)$, where $x_{j,i}$ denotes the ith coordinate of the n-bit string x_j. When all the g_i are the same, say g, we denote $f \circ \overrightarrow{g}$ by $f \circ g$. In this notation, the famous communication functions *generalized inner product*, *disjointness* and *hamming distance* can be written as $\text{GIP} = \text{MOD}_2 \circ \text{AND}$, $\text{DISJ} = \text{NOR} \circ \text{AND}$, and $\text{HD} = \text{THR}_t \circ \text{XOR}$ respectively. In an important paper [31], Razborov characterizes the 2 party communication complexity of $\text{SYM} \circ \text{AND}$ functions, where SYM denotes a symmetric function. Shi and Zhang [34] obtain a similar characterization for $\text{SYM} \circ \text{XOR}$ functions. Note that when $k = 2$, AND and XOR are the only interesting "inside functions" as other functions are either trivial or reduce to the case of AND or XOR.

In this paper, we study the multiparty communication complexity of composed functions with two goals in mind. The first goal is to better understand the power of $\log n$ and more players. The second and more general goal is to understand which combinations of the "inside" function g and the "outside" function f lead to hard communication problems and which combinations lead to easy communication problems. The focus of previous research has been on proving lower bounds for composed functions by selecting a "hard" outside function and a convenient inside function (see e.g. [32, 35, 21, 10, 6, **?**]). Our approach is to study composed functions without putting any restriction on g and obtain characterizations for the communication complexity of composed functions with respect to the choice of g. This *dual* approach is particularly interesting in the multiparty setting where the choice for g increases double exponentially in k.

First, we consider $\text{SYM} \circ g$ functions in the setting of $k > \log n$. This rich class contains many interesting functions and it is tempting to conjecture that some of these functions are candidates to break the $\log n$ barrier mentioned earlier. In particular, since the *majority* function $\text{MAJ} = \text{THR}_{n/2}$ is conjectured to be outside of ACC^0 [37], it is of interest to try to determine the communication complexity of $\text{MAJ} \circ g$ for all g. For instance, Babai, Kimmel and Lokam [4] identified $\text{MAJ} \circ \text{MAJ}$ as a candidate function to be hard for more than $\log n$ many players. Later, in a significantly expanded version of [4], Babai et al. [3] show that $\text{MAJ} \circ \text{MAJ}$ has an efficient simultaneous protocol when $k > 1 + \log n$. Their upper bound in fact applies to $\text{SYM} \circ g$ where SYM is any symmetric function and g is any symmetric "compressible" function. Although the class of symmetric compressible functions contains natural functions like THR_t and MOD_m, this class is only a small portion of all symmetric functions as a random symmetric function is not compressible with high probability. Babai et al. [3] in fact identify QCSB, the *quadractic character of the sum of bits* function, as a symmetric inside function g for which their method fails. In this paper, we remove the symmetry and compressibility conditions on g and show that functions of the form $\text{SYM} \circ g$ are easy in the simultaneous model when $k > 1 + \log n$, for *any* choice of the inside function g.

In the setting of $k \leq \log n$, we study more closely functions of the form $\text{MAJ} \circ g$, $\text{MOD}_m \circ g$ and $\text{NOR} \circ g$, where the latter two are generalizations of arguably the most well known and studied functions GIP and DISJ respectively. We are able to obtain dichotomies, with respect to the choice of g, that characterize the communication complexity of $\text{MAJ} \circ g$, $\text{MOD}_m \circ g$ and $\text{NOR} \circ g$ for every g. Furthermore, our results show that these functions have polynomially related quantum and classical communication complexities[3]. It is worth noting that these characterizations are tightly connected to our upper bound result mentioned above. The upper bounds for these functions in the setting of $k \leq \log n$ use crucially the ideas developed for the upper bound for $\text{SYM} \circ g$ in the setting of $k > \log n$. Perhaps surprisingly, even our lower bounds for $\text{MOD}_m \circ g$ functions use these ideas as well.

Grolmusz [14] presented an efficient non-simultaneous protocol for the function $\text{SYM} \circ \text{AND}$ and $k \geq \log n$ players. Using Grolmusz's ideas, Pudlák [28] obtained the same result with a slightly different protocol. The insight for our protocols is from the work of Grolmusz and Pudlák. We also discover a simple, yet powerful lemma (Lemma 3) which is used in all our protocols presented here. Additionally, we obtain simultaneous protocols when k is sufficiently large by employing a beautiful lemma of Babai et al [3, Lemma 6.10].

The first strong lower bounds in the NOF model were obtained by Babai, Nisan and Szegedy [5] for the $\text{GIP} = \text{MOD}_2 \circ \text{AND}$ function. Grolmusz [15] extended the technique of [5] to show a lower bound for $\text{MOD}_m \circ \text{AND}$. The method of [5] has been analyzed in [11, 30]. Here we obtain our main lower bound result (Theorem 3 **b**) by extending the analysis of [11, 30].

Our Results:

Symmetric of $\overrightarrow{g}$. We show that, for any g, there is a simultaneous deterministic k-party protocol for $\text{SYM} \circ g$ of cost $O(\log^3 n)$ when $k > 1 + \log n$. This improves a result of Babai et. al. [3] which exhibits an efficient simultaneous protocol for $\text{SYM} \circ g$ only when g is both symmetric and "compressible." When $k > 1 + 2\log n$, our simultaneous protocol applies to $\text{SYM} \circ \overrightarrow{g}$ for any vector of functions $\overrightarrow{g}$. Furthermore, we obtain a

[3] Note that by the work of [20], all our lower bounds hold in the quantum model, but we confine ourselves to the classical setting for simplicity.

deterministic protocol (non-simultaneous) for $\mathrm{SYM} \circ \overrightarrow{g}$ of cost $O(n/2^k \cdot \log n + k \log n)$ for any k (Theorem 2). Our result rules out functions of the form $\mathrm{SYM} \circ g$ as candidates to break the $\log n$ barrier. Moreover, by the well known connections of the multiparty model with Ramsey theory [9], our $k+1$ party protocol for $\mathrm{NOR} \circ \mathrm{XOR}$ gives the first non-trivial upper bound on the number of colors needed to color $(\mathbb{F}_2^n)^k$ so that no k dimensional *corner* is monochromatic. Although communication complexity bounds have been proven using Ramsey theory, no bounds on Ramsey numbers have been proven via communication complexity bounds before.

Mod m of g. Let $S_0 = \{y \in g^{-1}(1) \ : \ y \text{ has even weight}\}$ and $S_1 = \{y \in g^{-1}(1) \ : \ y \text{ has odd weight}\}$. First we show that if m divides $|S_0| - |S_1|$, $\mathrm{MOD}_m \circ g$ has a simultaneous deterministic protocol of cost $O(k \log m)$. On the other hand, if m does not divide $|S_0| - |S_1|$, $\mathrm{MOD}_m \circ g$ is a very hard function[4] in the randomized model, up to $\approx \frac{1}{2} \log n$ many players and m up to $n^{\frac{1}{2}-\delta}$ for a constant $\delta > 0$ (Theorem 3). For other m for which $\mathrm{MOD}_m \circ g$ is hard (i.e., m and $|S_0| - |S_1|$ are not coprime but m does not divide $|S_0| - |S_1|$), the previous analysis does not apply. In this case, we obtain the lower bound through a reduction to the previous case. This reduction vitally uses ideas from our upper bound for $\mathrm{SYM} \circ g$.

Majority of g. First, we show that if $|S_0| = |S_1|$, $\mathrm{MAJ} \circ g$ has a k-party simultaneous deterministic protocol of cost $O(k \log n)$. On the other hand, if $|S_0| \neq |S_1|$, then $\mathrm{MAJ} \circ g$ is hard in the randomized bounded error model for k up to $\approx \frac{1}{2} \log n$ (Theorem 4). This is in fact obtained by a (standard) reduction to the lower bound for $\mathrm{MOD}_m \circ g$ mentioned above. As immediate applications, we show for instance that $\mathrm{MAJ} \circ \mathrm{MAJ}$ and $\mathrm{MAJ} \circ \mathrm{XOR}$ are hard in the randomized model for k up to $\approx \frac{1}{2} \log n$. Moreover, from this answers an open question posed by Babai et al. [3], see Corollary 1.

Nor of g. Observe that if g's support size is 1, then it follows from [33] that $\mathrm{NOR} \circ g$ is hard in the randomized bounded error model for k up to $\approx \frac{1}{2} \log n$. On the other hand, we show that if g's support size is not 1, we show that $\mathrm{NOR} \circ g$ has a randomized protocol of cost $O(k)$ (Theorem 5). In other words, the hardness of DISJ crucially relies on the fact that g has singleton support. An important ingredient in our upper bound is the use of our characterization for $\mathrm{MOD}_m \circ g$.

2 Preliminaries

We refer the reader to [19] for details about the communication complexity models discussed in this paper. For $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{\pm 1\}$, we denote by $\mathbf{D}_k(F)$, $\mathbf{D}_k^{||}(F)$ and $\mathbf{R}_k^{\varepsilon}(F)$ the k-party deterministic, simultaneous deterministic and randomized ε-error communication complexities of F respectively. A stronger model allowing quantum communication between the players can similarly be defined, and in fact, all the lower bounds in the randomized model that we prove here carry over to the quantum model using the results of [20].

[4] Here 'very hard' means that even if the error probability of the protocol is allowed to be exponentially close to 1/2, the function does not have an efficient protocol. Note that achieving error probability 1/2 is trivial for any function.

A subset C_i of $\mathcal{X}_1 \times \cdots \times \mathcal{X}_k$ is a cylinder in the ith direction if membership in C_i does not depend on the ith coordinate, i.e., if $(x_1, \ldots, x_i, \ldots, x_k) \in C_i$, then $(x_1, \ldots, x_i', \ldots, x_k) \in C_i$ for every $x_i' \in \mathcal{X}_i$. A cylinder intersection C is an intersection of k cylinders, one in each direction. It is well known that a k-party deterministic protocol for F of cost c partitions the input space into at most 2^c monochromatic (with respect to F's output) cylinder intersections. We identify a cylinder intersection $C \subseteq \mathcal{X}_1 \times \cdots \times \mathcal{X}_k$ with its characteristic function $C : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{0,1\}$.

We define the discrepancy of $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \mathbb{C}$ under μ and with respect to a cylinder intersection C as $\mathrm{disc}_\mu(F,C) = \left|\mathbf{E}_{x\sim\mu}[F(x)C(x)]\right|$. The discrepancy of F under μ is $\mathrm{disc}_\mu(F) = \max_C \mathrm{disc}_\mu(F,C)$, where the maximum is over all possible cylinder intersections C. By the well-known discrepancy method:

$$\mathbf{R}_k^\varepsilon(F) \geq \log\left(\frac{1-2\varepsilon}{\mathrm{disc}_\mu(F)}\right). \tag{1}$$

In order to upper bound the discrepancy we will use the *cube measure*. Let μ be a product distribution over $\mathcal{X}_1 \times \cdots \times \mathcal{X}_k$, i.e., $\mu(x_1, \ldots, x_k) = \mu_1(x_1)\cdots\mu_k(x_k)$, where μ_i is a distribution over $\mathcal{X}_i$. We define the cube measure of a complex valued function F under μ as

$$\mathcal{E}_\mu(F) = \mathbf{E}_{\substack{x_1^0, x_2^0, \ldots, x_k^0 \\ x_1^1, x_2^1, \ldots, x_k^1}}\left[\prod_{u\in\{0,1\}^k} \mathcal{C}^{u_1+\cdots+u_k}(F(x_1^{u_1}, \ldots, x_k^{u_k}))\right],$$

where in the expectation, x_i^0 and x_i^1 are distributed according to μ_i, and $\mathcal{C}$ denotes the complex conjugation operator: $\mathcal{C}^b(z) = z$ if b is even, and $\mathcal{C}^b(z) = \bar{z}$ otherwise. It is not difficult to verify that the cube measure is always a non-negative real number. In fact, the quantity $(\mathcal{E}_\mathcal{U}(F))^{1/2^k}$, where $\mathcal{U}$ is the uniform distribution, is known as the *hypergraph uniformity norm* and is a measure of "quasirandomness" of F. When $F(x_1, \ldots, x_k) = f(x_1 \oplus \cdots \oplus x_k)$, the hypergraph uniformity norm of F corresponds to Gowers uniformity norm of f over $\mathbb{F}_2^n$.

Lemma 1 ([11, 30, 38]). *Let $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \mathbb{C}$ be a complex valued function and μ_i a distribution over $\mathcal{X}_i$. Define the distribution μ as the product of the μ_i. Then,* $\mathrm{disc}_\mu(F) \leq (\mathcal{E}_\mu(F))^{1/2^k}$.

In this paper $\mathcal{X}_i = \{0,1\}^n$ for all i. We let $x = (x_1, \ldots, x_k)$ denote an input in $(\{0,1\}^n)^k$. Often we will view the input as a $k \times n$ dimensional matrix X, where the ith row of X is x_i. We reserve the variables x_i to denote an n-bit string whose j-th bit is denoted by $x_{i,j}$, and reserve the variables y_i to denote a single bit. Let $\mathcal{H}_k$ denote the k dimensional hypercube where the vertex set is $\{0,1\}^k$ and there is an edge between two vertices iff their Hamming distance is 1. Given an input in the $k \times n$ dimensional matrix form X, we associate each column of X with the corresponding vertex of $\mathcal{H}_k$. For each $v \in \{0,1\}^k$, define n_v as the number of times v occurs as a column of X.

3 Communication complexity of composed functions

3.1 SYM $\circ g$

A boolean function $f : \{0,1\}^n \to \{\pm 1\}$ is called *symmetric* if the output depends only on the Hamming weight of the input. In this section we present a deterministic protocol for SYM $\circ \overrightarrow{g}$ where $\overrightarrow{g}$ is any vector of functions. This protocol becomes efficient (i.e. poly-logarithmic in n) for $k \geq \log n - O(\log\log n)$ players. Our protocol is perhaps an easy extension of Grolmusz's protocol [14, 28] that is nevertheless not observed before.

Moreover, for $k > 1 + \log n$ our protocol can be made simultaneous, and this improves an earlier result by Babai et al. [3], who gave an efficient simultaneous protocol for SYM $\circ g$ only for functions g which are symmetric and *compressible*. We observe further that for $k > 1 + 2\log n$ players we can allow an arbitrary vector of functions $\overrightarrow{g}$, as oppose to just a single function g. Our simultaneous protocols are obtained using the following lemma of Babai et al. [3, Lemma 6.10]:

Lemma 2 ([3]). *Suppose $k > 1 + \log n$ and let X be a $k \times n$ boolean matrix given as an input for a k-party communication problem. Let n_i be the number of columns of X with Hamming weight i. Then by communicating $O(k^2 \log n)$ bits, the players can compute n_i for all i in the simultaneous deterministic model.*

Theorem 2. *Let $f : \{0,1\}^n \to \{\pm 1\}$ be a symmetric function, $g : \{0,1\}^n \to \{0,1\}$ an arbitrary function, and $\overrightarrow{g} = (g_1, \ldots, g_n)$ a vector of n functions where $g_i : \{0,1\}^k \to \{0,1\}$ are arbitrary functions. Then,*

(a) $\mathbf{D}_k(f \circ \overrightarrow{g}) \leq O(n/2^k \cdot \log n + k \log n)$,
(b) *for* $k > 1 + \log n$: $\mathbf{D}_k^{||}(f \circ g) \leq O(\log^3 n)$,
(c) *for* $k > 1 + 2\log n$: $\mathbf{D}_k^{||}(f \circ \overrightarrow{g}) \leq O(\log^3 n)$.

Proof. We outline the proof here, for details see [1, Theorem 3.2]. We first prove part **(a)**. Fix an input for $f \circ \overrightarrow{g}$ given in $k \times n$ matrix form X. The protocol proceeds in two steps. In the first step, the players determine a specific $u \in \mathcal{H}_k$ and the set C of precisely all columns that contain u. In the second step, they use this information to compute the output of $f \circ \overrightarrow{g}$.

The first step is roughly the same as in Grolmusz's original protocol [14, 28]. This is done by two specific players (e.g. Player 1 and Player 2) and the cost is $O(k + n/2^k \cdot \log n)$ bits. The second step can be done simultaneously as follows. Let S_j denote the support of g_j: $S_j = g_j^{-1}(1)$. For $v \in \{0,1\}^k$, let $\mathbf{1}_j(v) = 1$ if v is in column j, and $\mathbf{1}_j(v) = 0$ otherwise. Now, to compute the output of $f \circ \overrightarrow{g}$, it suffices to compute

$$\sum_{j \notin C} \sum_{v \in S_j} \mathbf{1}_j(v), \tag{2}$$

For a given v, consider a shortest path from v to u: $v = w_1, w_2, \ldots, w_t = u$. Then, since $\mathbf{1}_j(u) = 0$,

$$\mathbf{1}_j(v) = \sum_{i=1}^{t-1} (-1)^{i+1} (\mathbf{1}_j(w_i) + \mathbf{1}_j(w_{i+1})). \tag{3}$$

Each term $(\mathbf{1}_j(w_i)+\mathbf{1}_j(w_{i+1}))$ above is known by some player because w_i and w_{i+1} differ only in one coordinate. As a result, (2) can be written as a sum of n terms, one for each player. So to compute (2), each player announces her part of the sum. In addition, since $\sum_j \sum_{v\in S_j} \mathbf{1}_j(v) \le n$, it suffices for players to send their part of the sum modulo $n+1$. Therefore this step of the protocol has cost at most $k\cdot\lceil\log(n+1)\rceil$.

To obtain simultaneous protocols for parts **(b)** and **(c)** we show, essentially, that the first step above can be bypassed, because there are many players. Consider for example part **(c)**. Let $\ell = 2+2\log n$. Only the first ℓ players will participate in the protocol. Thus, for each column j, the rows $\ell+1$ to k naturally induce a function $g'_j : \{0,1\}^\ell \to \{0,1\}$ as follows: $g'_j(u) = g_j(u\cdot v)$ where $v \in \{0,1\}^{k-\ell}$ appears in column j from row $\ell+1$ to k. Our task then reduces to finding a protocol for $f\circ \overrightarrow{g'}$ with ℓ players. Step 1 is bypassed by taking u to be the column 0^ℓ and apply Lemma 2 above. See the full version of this paper [1, Theorem 3.2] for details.

3.2 $\text{MOD}_m \circ g$

For $(y_1,y_2,\ldots,y_n)\in\{0,1\}^n$, let $\text{MOD}_m(y_1,y_2,\ldots,y_n) = -1$ iff $\sum_{j=1}^n y_j = 0 \mod m$. In this section we show that the complexity of $\text{MOD}_m \circ g$ is determined by the quantity $\big||S_0|-|S_1|\big|$, where S_i is the subset of the support of g that consists of all inputs whose Hamming weight has parity i.

Theorem 3. *Let $m \ge 2$ be an integer. The function $\text{MOD}_m \circ g$ satisfies:*

(a) *If m divides $|S_0|-|S_1|$, then $\mathbf{D}_k^{||}(\text{MOD}_m\circ g) \le k\lceil\log m\rceil$.*
(b) *Otherwise, $\mathbf{R}_k^{\varepsilon}(\text{MOD}_m\circ g) \ge \frac{5n}{m^2 4^k} + \log(1-2\varepsilon) - (k+1)\lceil\log m\rceil - 1$.*

Before sketching the proof, we first state a fact which we will use.

Fact 3 *Let $S_0 = \{u_1,\ldots,u_r\}$ and $S_1 = \{v_1,\ldots,v_r\}$ be two subsets of the vertices of $\mathcal{H}_k$ such that for each i, the distance between u_i and v_i is odd. The sum $\sum_{i=1}^r n_{u_i} + \sum_{i=1}^r n_{v_i} \mod m$ can be computed by the players in the simultaneous model using at most $k\cdot\lceil\log m\rceil$ bits. Similarly, if for each i, the distance between u_i and v_i is even, $\sum_{i=1}^r n_{u_i} - \sum_{i=1}^r n_{v_i} \mod m$ can be computed in the simultaneous model using at most $k\cdot\lceil\log m\rceil$ bits.*

Proof. Note that we are interested in computing $\sum_{i=1}^r (n_{u_i}+n_{v_i}) \mod m$. Each term $(n_{u_i}+n_{v_i})$ can be written as a telescoping sum as in (3). Each term in the telescoping sum is known by a player. Since we can do arithmetic modulo m, the desired value can be computed with each player sending their part of the sum modulo m. So the total cost is $k\cdot\lceil\log m\rceil$. The second part holds similarly.

Proof (Proof of Theorem 3). **Part (a):** Suppose m divides $|S_0|-|S_1|$ and assume without loss of generality that $|S_0|\ge|S_1|$. We choose (arbitrarily) a subset $S'_0 \subseteq S_0$ of size $|S_1|$. As the distance between an element of S'_0 and an element of S_1 is odd, we can compute $\sum_{v\in S'_0} n_v + \sum_{v\in S_1} n_v \mod m$ using Fact 3. For the remaining elements in $S_0 - S'_0$, we pair them with $\overrightarrow{0}$. Hence, using Fact 3 once again, we can compute $(|S_0|-|S_1|)n_{\overrightarrow{0}} +$

$\sum_{v\in S_0-S'_0} n_v \equiv \sum_{v\in S_0-S'_0} n_v \mod m$. Thus, we have computed $\sum_{v\in S_0\cup S_1} n_v \mod m$, from which the output of $\text{MOD}_m \circ g$ is determined. Observe that the sums $\sum_{v\in S'_0} n_v + \sum_{v\in S_1} n_v \mod m$ and $\sum_{v\in S_0-S'_0} n_v \mod m$ need not be computed separately and that we can compute $\sum_{v\in S_0\cup S_1} n_v \mod m$ in one shot using $k\lceil \log m\rceil$ bits.

Part (b), Case 1: We consider two cases, depending on whether m and $|S_0|-|S_1|$ are coprime or not. The first case is when m and $|S_0|-|S_1|$ are coprime. The proof makes use of the characterization of the MOD_m function in terms of exponential sums. Fix $2 \le m \in \mathbb{N}$ and $0 \le a,b \le m-1$. Let $\omega = e^{2\pi i/m}$ be an m-th root of unity. The function $\text{EXP}_m^{a,b}$ is defined as $\text{EXP}_m^{a,b}(y_1,y_2,\ldots,y_n) = \omega^{a((\sum_{j=1}^n y_j)-b)}$.

The strategy is as follows. Define $f_m(y_1,\ldots,y_n) = \sum_j y_j \mod m$. First we show that for any cylinder intersection, the fraction of points x in the cylinder intersection that satisfy $f_m \circ g(x) = b$ is roughly (with exponentially small error) $1/m$ for all $b \in \{0,1,\ldots,m-1\}$. This step uses an estimate of the cube measure of $\text{EXP}_m^{a,b} \circ g$ under the uniform distribution. Define the distribution μ that puts equal weight to all x with $f_m \circ g(x) = 0$ and $f_m \circ g(x) = 1$. All other points get 0 weight. It will easily follow that $\text{disc}_\mu(\text{MOD}_m \circ g)$ is exponentially small and thus the desired lower bound is achieved using the discrepancy method (Inequality (1)). The details of the proof can be found in the appendix.

Part (b), Case 2: To handle the case where m and $|S_0|-|S_1|$ are not coprime, we construct a reduction to Case 1 using ideas from the protocol of Theorem 2. The proof is provided in the Appendix.

3.3 $\text{MAJ} \circ g$

For each $n \ge 1$, the *majority* function $\text{MAJ}^n : \{0,1\}^n \to \{-1,1\}$ is defined as $\text{MAJ}^n(y_1,\ldots,y_n) = -1$ iff $\sum_i y_i \ge n/2$. When no confusion arises we drop the supercript n from MAJ^n. It is not difficult to show that $\text{MAJ} \circ g$ cannot be much easier than $\text{SYM} \circ g$:

Proposition 1. *Let $g : \{0,1\}^k \to \{0,1\}$ be a boolean function and $f : \{0,1\}^n \to \{-1,1\}$ be a symmetric function on n variables. For any $\varepsilon \ge 0$, $\mathbf{R}_k^{\varepsilon'}(f \circ g) \le \mathbf{R}_k^{\varepsilon}(\text{MAJ}^{2n} \circ g) \cdot \lceil \log(n+1)\rceil$, where $\varepsilon' = \varepsilon \lceil \log(n+1)\rceil$.*

We can combine Proposition 1 with our lower bounds for $\text{MOD}_m \circ g$ functions (Theorem 3) to obtain a characterization for the communication complexity of $\text{MAJ} \circ g$ for every g.

Theorem 4. *Let $g : \{0,1\}^k \to \{0,1\}$ be a boolean function and S be its support. The function $\text{MAJ} \circ g$ satisfies:*

- *If $|S_0| = |S_1|$, then $\mathbf{D}_k^{||}(\text{MAJ} \circ g) \le k \cdot \lceil \log(n+1)\rceil$.*
- *Otherwise, $\mathbf{R}_k^{1/3}(\text{MAJ} \circ g) \ge \Omega\left(\frac{n}{(k\log k)^2 \cdot 4^k \log n \log\log n}\right)$.*

Theorem 4 can be used to determine the communication complexity of a class of functions considered by Babai et al. [3]. For an odd prime k, define the function $\text{QCSB}_k : \{0,1\}^k \to \{0,1\}$ by $\text{QCSB}_k(y_1,\ldots,y_k) = 1$ if and only if $y_1 + \cdots + y_k$ is a quadratic residue modulo k. Recall that $z \in \mathbb{F}_k$ is a quadratic residue if there exists

$a \in \mathbb{F}_k$ such that $z = a^2$. The authors of [3] prove that QCSB_k is not 'compressible', so their protocol for $k > 1 + \log n$ does not apply for $\text{SYM} \circ \text{QCSB}_k$. They leave as an open question the problem of finding good bounds for the communication complexity of the function $\text{MAJ} \circ \text{QCSB}_k$. The following corollary completely determines the hardness of this function for any number of players k, except in the range between $\approx 1/2 \log n$ and $\log n$.

Corollary 1. *Let k be an odd prime.*

- *If $k \equiv 1 \mod 4$, then $\mathbf{D}_k^{||}(\text{MAJ} \circ \text{QCSB}_k) \leq O(k \log n)$.*
- *If $k \equiv 3 \mod 4$, then $\mathbf{R}_k^{1/3}(\text{MAJ} \circ \text{QCSB}_k) \geq \Omega\left(\frac{n}{(k \log k)^2 4^k \log n \log\log n}\right)$.*
- *If $k > 1 + \log n$, then $\mathbf{D}_k^{||}(\text{MAJ} \circ \text{QCSB}_k) \leq O(\log^3 n)$.*

Proof. Let S be the support of QCSB_k and define S_0 and S_1 as in Theorem 4. It is known that when $k \equiv 1 \mod 4$, $z \in \{0, \ldots, k-1\}$ is a quadratic residue modulo k if and only if $-z \equiv k - z \mod k$ is a quadratic residue modulo k; see e.g., [36, Theorem 2.21]. As k is odd, z is even if and only if $k - z$ is odd. In other words, the function $(y_1, \ldots, y_k) \mapsto (1 - y_1, \ldots, 1 - y_k)$ defines a bijection between S_0 and S_1. Thus, $|S_0| = |S_1|$ whenever $k \equiv 1 \mod 4$. Otherwise, if $k \equiv 3 \mod 4$, then the number $|S|$ of quadratic residues modulo k is odd; see e.g., [36, Theorem 2.20]. This implies that $|S_0| \neq |S_1|$. For $k > 1 + \log n$, we can use Theorem 2.

3.4 $\text{NOR} \circ g$

In this section, we obtain a simple and perhaps surprising characterization for the k-player randomized communication complexity of $\text{NOR} \circ g$, where $\text{NOR}(y_1, \ldots, y_n) = -1$ iff $(y_1, \ldots, y_n) = (0, \ldots, 0)$. In a very recent paper, Sherstov [33] significantly improves on the bounds of [21],[10] and [6] on the multiparty bounded error communication complexity of disjointness: $\mathbf{R}_k^{1/3}(\text{DISJ}) \geq \Omega\left(\frac{n}{4^k}\right)^{1/4}$. First we observe that this lower bound applies - via a simple reduction - to $\text{NOR} \circ g$ when g's support size is 1. We complement this with an efficient randomized protocol for $\text{NOR} \circ g$ when g's support size is more than one.

Theorem 5. *Let $g : \{0,1\}^k \to \{0,1\}$ be a boolean function and $S = \{y \in \{0,1\}^k : g(y) = 1\}$ be its support.*

- *If $|S| = 1$, $\mathbf{R}_k^{1/3}(\text{NOR} \circ g) \geq \Omega\left(\frac{n}{4^k}\right)^{1/4}$,*
- *Otherwise, $\mathbf{R}_k^{\varepsilon}(\text{NOR} \circ g) \leq O(k)$ for a constant ε.*

Proof. The lower bound follows from the lower bound on the disjointness function [33] via a simple reduction.

For the upper bound, first assume that $|S|$ is even. In this case, by Theorem 3, we have a deterministic protocol Π for $\text{MOD}_2 \circ g$ of cost k. We will use this protocol Π as a subroutine to compute $\text{NOR} \circ g$. As before, denote by X the $k \times n$ dimensional matrix representing the input. Denote by X_r a random matrix obtained from X by deleting every column independently with probability 1/2. The players can agree on X_r without any communication using their public random bits. We output -1 if $\Pi(X_r) = -1$ and output 1 otherwise.

Observe that if $\text{NOR} \circ g(X) = -1$, then $\text{NOR} \circ g(X_r) = -1$, and so $\text{MOD}_2 \circ g(X_r) = -1$. In this case our protocol does not make an error. In this case, the error probability is $1/2$. Repeating this protocol t times would reduce the error probability to $1/2^t$.

Now assume $|S|$ is odd and $|S| > 1$. Divide S into two non-disjoint parts S_1 and S_2 of even size each. Let g_1 be the boolean function with support S_1 and g_2 be the boolean function with support S_2. Observe that $\text{NOR} \circ g(X) = -1$ iff both $\text{NOR} \circ g_1(X) = -1$ and $\text{NOR} \circ g_2(X) = -1$. Since we covered the case of even support size, we are done.

References

1. Anil Ada, Arkadev Chattopadhyay, Omar Fawzi, and Phuong Nguyen. The NOF Multiparty Communication Complexity of Composed Functions. Technical report, In Electronic Colloquium on Computational Complexity (ECCC) TR11–155, 2011.
2. Noga Alon, Yossi Matias, and Mario Szegedy. The space complexity of approximating the frequency moments. *Journal of Computer and System Sciences*, 58:137–147, 1999.
3. László Babai, Anna Gál, Peter G. Kimmel, and Satyanarayana V. Lokam. Communication complexity of simultaneous messages. *SIAM Journal on Computing*, 33:137–166, 2004.
4. László Babai, Peter G. Kimmel, and Satyanarayana V. Lokam. Simultaneous messages vs. communication. In *In 12th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 361–372. Springer, 1995.
5. László Babai, Noam Nisan, and Mario Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. Syst. Sci.*, 45(2):204–232, 1992.
6. Paul Beame and Dang-Trinh Huynh-Ngoc. Multiparty communication complexity and threshold circuit size of AC^0. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '09, pages 53–62, Washington, DC, USA, 2009. IEEE Computer Society.
7. Richard Beigel and Jun Tarui. On ACC. *Computational Complexity*, 4:350–366, 1994.
8. Eric Blais, Joshua Brody, and Kevin Matulef. Property testing lower bounds via communication complexity. In *In Proceedings of the 2011 IEEE 26th Annual Conference on Computational Complexity (CCC)*, 2011.
9. Ashok K. Chandra, Merrick L. Furst, and Richard J. Lipton. Multi-party protocols. In *Proceedings of the fifteenth annual ACM symposium on Theory of computing*, STOC '83, pages 94–99, New York, NY, USA, 1983. ACM.
10. Arkadev Chattopadhyay and Anil Ada. Multiparty communication complexity of disjointness. Technical report, In Electronic Colloquium on Computational Complexity (ECCC) TR08–002, 2008.
11. Fan R.K. Chung and Prasad Tetali. Communication complexity and quasi randomness. *SIAM Journal on Discrete Mathematics*, 6(1):110–123, 1993.
12. Vincent Conitzer and Tuomas Sandholm. Communication complexity as a lower bound for learning in games. In *International Conference on Machine Learning*, 2004.
13. Jürgen Forster, Matthias Krause, Satyanarayana V. Lokam, Rustam Mubarakzjanov, Niels Schmitt, and Hans ulrich Simon. Relations between communication complexity, linear arrangements, and computational complexity. In *Foundations of Software Technology and Theoretical Computer Science*, pages 171–182, 2001.
14. Vince Grolmusz. The BNS lower bound for multi-party protocols is nearly optimal. *Information and Computation*, 112:51–54, 1994.

15. Vince Grolmusz. Separating the communication complexities of MOD m and MOD p circuits. In *Proceedings of the 33rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 278–287, 1995.
16. Vince Grolmusz. Circuits and multi-party protocols. *Computational Complexity*, 7:1–18, 1998.
17. Johan Håstad and Mikael Goldmann. On the power of small-depth threshold circuits. *Computational Complexity*, 1:610–618, 1991.
18. Hartmut Klauck. Lower bounds for quantum communication complexity. *Siam Journal on Computing*, 37:20–46, 2007.
19. Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge university press, 1997.
20. Troy Lee, Gideon Schechtman, and Adi Shraibman. Lower bounds on quantum multiparty communication complexity. In *In Proceedings of the 24th Annual IEEE Conference on Computational Complexity (CCC)*, pages 254–262, 2009.
21. Troy Lee and Adi Shraibman. Disjointness is hard in the multiparty number-on-the-forehead model. *Computational Complexity*, 18:309–336, 2009.
22. Peter Bro Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. On data structures and asymmetric communication complexity. *Journal of Computer and System Sciences*, 57(1):37 – 49, 1998.
23. Ashley Montanaro and Tobias Osborne. On the communication complexity of XOR functions. *Arxiv preprint arXiv:0909.3392*, 2009.
24. N. Nisan. The communication complexity of threshold gates. *Combinatorica*, 1993.
25. Noam Nisan and Ilya Segal. The communication requirements of efficient allocations and supporting prices. *Journal of Economic Theory*, 129:192–224, 2006.
26. Noam Nisan and Avi Wigderson. Rounds in communication complexity revisited. *Siam Journal on Computing*, 22:211–219, 1993.
27. Beame Paul, Toniann Pitassi, and Nathan Segerlind. Lower bounds for Lovász-Schrijver systems and beyond follow from multiparty communication complexity. *SIAM J. Comput.*, 37:845–869, June 2007.
28. Pavel Pudlák, 2006. Personal communication.
29. Ran Raz. Fourier analysis for probabilistic communication complexity. *Computational Complexity*, 5:205–221, 1995.
30. Ran Raz. The BNS-Chung criterion for multi-party communication complexity. *Computational Complexity*, 9(2):113–122, 2000.
31. Alexander Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145–159, 2003.
32. Alexander A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In *In Proceedings of the 40th Symposium on Theory of Computing (STOC)*, pages 85–94, 2007.
33. Alexander A. Sherstov. The multiparty communication complexity of set disjointness. Technical report, In Electronic Colloquium on Computational Complexity (ECCC) TR11–145, 2011.
34. Yaoyun Shi and Zhiqiang Zhang. Communication complexities of symmetric XOR functions. *Quantum Information and Computation*, 9:255–263, 2009.
35. Yaoyun Shi and Yufan Zhu. Quantum communication complexity of block-composed functions. *Quantum Information and Computation*, 9:444–460, May 2009.
36. Victor Shoup. *A computational introduction to number theory and algebra*. Cambridge University Press, 2009.
37. Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, STOC '87, pages 77–82, New York, NY, USA, 1987. ACM.

38. Emanuele Viola and Avi Wigderson. Norms, xor lemmas, and lower bounds for polynomials and protocols. *Theory of Computing*, 4(1):137–168, 2008.
39. Andrew C. Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 209–213, New York, NY, USA, 1979. ACM Press.