Electronic Colloquium on Computational Complexity, Report No. 155 (2011)

The NOF Multiparty Communication Complexity of Composed Functions

Anil Ada* Arkadev Chattopadhyay† Omar Fawzi‡ Phuong Nguyen§

November 21, 2011

Abstract

We study the k-party 'number on the forehead' communication complexity of composed functions $f \circ \vec{g}$, where $f : \{0,1\}^n \to \{\pm 1\}$, $\vec{g} = (g_1, \ldots, g_n)$, $g_i : \{0,1\}^k \to \{0,1\}$ and for $(x_1, \ldots, x_k) \in (\{0,1\}^n)^k$, $f \circ \vec{g}(x_1, \ldots, x_k) = f(\ldots, g_i(x_{1,i}, \ldots, x_{k,i}), \ldots)$. When $\vec{g} = (g, g, \ldots, g)$ we denote $f \circ \vec{g}$ by $f \circ g$. We show that there is an $O(\log^3 n)$ cost simultaneous protocol for SYM $\circ\, g$ when $k > 1 + \log n$, SYM is any symmetric function and g is any function. When $k > 1 + 2\log n$, our simultaneous protocol applies to SYM $\circ\, \vec{g}$ with $\vec{g}$ being a vector of any n functions. We also get a non-simultaneous protocol for SYM $\circ\, \vec{g}$ 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 MAJORITY $\circ\, g$, MOD$_m \circ\, g$, and 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, 2003*) and determine the communication complexity of MAJORITY $\circ$ QCSB$_k$, where QCSB$_k$ is the "quadratic character of the sum of the bits" function.

In the second part of our paper we utilize the connection between the 'number on the forehead' model and Ramsey theory to construct a large set without a k-dimensional corner (k-dimensional generalization of a k-term arithmetic progression) in $(\mathbb{F}_2^n)^k$, thereby obtaining the first non-trivial bound on the corresponding Ramsey number. Furthermore, we give an explicit coloring of $[N] \times [N]$ without a monochromatic 2-dimensional corner and use this to obtain an explicit 3-party protocol of cost $O(\sqrt{n})$ for the EXACT$_N$ function. For x_1, x_2, x_3 n-bit integers, EXACT$_N(x_1, x_2, x_3) = -1$ iff $x_1 + x_2 + x_3 = N$.

*Department of Computer Science, McGill University. aada@cs.mcgill.ca. Supported by Natural Sciences and Engineering Research Council (NSERC) of Canada.

†Department of Computer Science, University of Toronto. arkadev@cs.toronto.edu. Supported by a postdoctoral fellowship of Natural Sciences and Engineering Research Council (NSERC) of Canada and a postdoctoral fellowship of the Ontario Ministry of Research and Innovation.

‡Department of Computer Science, McGill University. ofawzi@cs.mcgill.ca. Supported by CIFAR, NSERC and ONR grant No. N000140811249.

§Dép. d'informatique et de recherche opérationnelle, Université de Montréal. pnguyen@cs.toronto.edu. Supported by a postdoctoral fellowship of Natural Sciences and Engineering Research Council (NSERC) of Canada.

ISSN 1433-8092

1 Introduction

The 'number on the forehead' (NOF) model of communication complexity was introduced by Chandra, Furst and Lipton [CFL83] 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 [Yao79]. We will be mainly interested in the case of $\mathcal{X}_i = \{0,1\}^n$ 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 is what makes NOF model interesting, powerful and fruitful in terms of applications. Apart from the aforementioned application in branching programs, this model also has very important applications in boolean 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. Separating ACC^0 from NP is one of the frontiers in complexity theory. It is well known that a function in ACC^0 has a polylog(n) k-party deterministic communication complexity, where k is polylog(n) [HG91, BT94]. 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 have been solved in the 2 player setting remain open for 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.

Arguably the most well known and studied functions in the standard two party as well as the multiparty models are the *generalized inner product* function GIP and the *disjointness* function DISJ. The GIP function is a hard function (or conjectured to be hard) in almost every model of communication complexity. As such, strong lower bounds can be proven for many different kinds of boolean circuits using GIP [HG91, Nis93, FKL$^+$01, Gro98]. It is also used in obtaining decision tree lower bounds [Nis93], in the construction of pseudorandom generators, time/space tradeoffs for Turing Machines and branching program lower bounds [BNS92].

The DISJ function, unlike GIP, is easy in the non-deterministic model. Proving lower bounds for DISJ in the randomized model even for 2 players was a major challenge. A strong lower bound for 3 players has been proven only very recently [LS09, CA08]. Both the 2 player and multiplayer lower bounds on DISJ lead to the development of interesting techniques and a deeper understanding of communication complexity in general. Apart from this, the interest in studying DISJ also stems from the fact that it is very suitable for reductions: communication complexity lower bounds for DISJ (and slight variations) have been successfully used to give lower bounds in the context of data streaming [AMS99], proof complexity [PPS07], data structures [MNSW98],

game theory [CS04, NS06], boolean circuits [NW93], and property testing [BBM11].

The functions GIP and DISJ have the following 'composed' structure. Let $f : \{0,1\}^n \to \{\pm 1\}$ be a function and $\vec{g} = (g_1, \ldots, g_n)$ be a vector of functions $g_i : \{0,1\}^k \to \{0,1\}$. Define $f \circ \vec{g}(x_1, \ldots, x_k) = f(\ldots, g_i(x_{1,i}, x_{2,i}, \ldots, x_{k,i}), \ldots)$, where $x_{j,i}$ denotes the ith coordinate of the n-bit string x_j. When all the g_i are the same function, say g, we denote $f \circ \vec{g}$ by $f \circ g$. In this notation, GIP $=$ $\text{MOD}_2 \circ \text{AND}$ and DISJ $=$ NOR $\circ$ AND, where NOR is the negation of OR. In both the two party and the multiparty models, functions of the form $f \circ$ AND have been studied extensively [Raz95, Raz03, Kla07, She07, SZ09b, LS09, CA08, BHN09], with an emphasis on SYM $\circ$ AND, where SYM represents a symmetric function. For instance, in the important paper [Raz03], Razborov shows that the 2 party quantum and classical communication complexities of SYM $\circ$ AND are polynomially related. Functions of the form $f \circ$ XOR have also received a lot of attention in the 2 player setting [Raz95, Kla07, SZ09a, MO09], especially the *hamming distance* problem $\text{THR}_t \circ$ XOR where THR_t is a threshold function[1]. Notably, Shi and Zhang [SZ09a] obtain 2 party classical and quantum equivalence of functions of the form SYM $\circ$ XOR.

Observe that the focus so far in the literature has been to fix an inside function g and vary the outside function f. In this paper, we propose a new dual approach. We study the multiparty communication complexity of composed functions by fixing the outside function to some natural function and vary the inside function. This dual approach is particularly interesting in the multiparty setting where the choice for inside function increases as k increases. Note that when $k = 2$, AND and XOR are really the only interesting inside functions as other functions are either trivial or reduce to the case of AND or XOR.

First, we consider functions of the form SYM $\circ\, g$ in the setting of $k > \log n$. This rich class contains many interesting functions and possible candidates to break the $\log n$ barrier mentioned earlier. Since the *majority* function $\text{MAJ} = \text{THR}_{n/2}$ is conjectured to be outside of ACC^0, it is of interest to try to determine the communication complexity of MAJ $\circ\, g$ for all g. For instance, Babai, Kimmel and Lokam [BKL95] identify MAJ $\circ$ MAJ as a candidate function to be hard for more than $\log n$ many players. Later, in a significantly expanded version of [BKL95], Babai et al. [BGKL03] show that MAJ $\circ$ MAJ has an efficient simultaneous protocol when $k > 1 + \log n$. Their upper bound in fact applies to SYM $\circ\, g$ where SYM is any symmetric function and g is any symmetric "compressible" function, a small subset of all symmetric functions[2]. In the same paper, the authors ask about the communication complexity of MAJ$\circ g$ for a specific symmetric g called "the quadratic character of the sum of the bits", which they show is not compressible. We show that functions of the form 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 MAJ $\circ\, g$, $\text{MOD}_m \circ g$ and NOR $\circ\, g$, where the latter two are generalizations of GIP and DISJ respectively. We are able to obtain dichotomies, with respect to the choice of g, that characterize the communication complexity of MAJ $\circ\, g$, $\text{MOD}_m \circ g$ and NOR $\circ\, g$ for every g. Furthermore, our results show that these functions have polynomially related quantum and classical communication complexities. Below we summarize our results. Note that by the work of [LSS09], all our lower bounds hold in the quantum model, but we confine ourselves to the classical setting for simplicity.

[1] Even though XOR and MOD_2 both represent the *parity* function, we use the notation XOR for an inside function g and MOD_2 for an outside function f.

[2] A random symmetric function is not compressible with high probability.

Our Results:

Symmetric of $\vec{g}$. 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 $k > 1 + \log n$. When $k > 1 + 2\log n$, our simultaneous protocol applies to $\text{SYM} \circ \vec{g}$ for any vector of functions $\vec{g}$. Furthermore, we obtain a deterministic protocol (non-simultaneous) for $\text{SYM} \circ \vec{g}$ of cost $O(n/2^k \cdot \log n + k \log n)$ for any k (Theorem 3.2). Our result rules out functions of the form $\text{SYM} \circ g$ as a candidate to break the $k = \log n$ barrier. Furthermore, this result has an application in Ramsey theory which we will discuss in the next subsection.

Previously, an efficient non-simultaneous protocol for $k \geq \log n$ players was known for $\text{SYM} \circ \text{AND}$ due to Grolmusz [Gro94]. Using Grolmusz's ideas, Pudlák [Pud06] obtained the same result with a slightly different protocol. Babai et al. [BGKL03] gave an efficient simultaneous protocol, which works only when $k > 1 + \log n$, for $\text{SYM} \circ g$ when g is symmetric and compressible. We obtain our protocols by extending the ideas of Grolmusz and Pudlák, and employing a beautiful lemma of Babai et al [BGKL03, Lemma 6.10]. We note that our upper bound results for the functions listed below use the key insights of the protocol for $\text{SYM} \circ \vec{g}$, in particular Fact 3.3.

Majority of g. Let $S_0 = \{y \in g^{-1}(1) \ : \ y \text{ has even weight}\}$ and $S_1 = \{y \in g^{-1}(1) \ : \ y \text{ has odd weight}\}$. We show that if $|S_0| = |S_1|$, $\text{MAJ} \circ g$ has a k-party simultaneous deterministic protocol of cost $O(k \log n)$, and if $|S_0| \neq |S_1|$, then $\text{MAJ} \circ g$ is hard in the randomized bounded error model for up to $\approx \frac{1}{2} \log n$ many players (Theorem 3.8). As immediate applications, we can show for instance that $\text{MAJ} \circ \text{MAJ}$ and $\text{MAJ} \circ \text{XOR}$ are hard in the randomized model for up to $\approx \frac{1}{2} \log n$ many players. Our upper bound for $\text{MAJ} \circ g$ when $|S_0| = |S_1|$ follows directly from Fact 3.3. For the lower bound, we observe that $\text{MAJ} \circ g$ is hard if there is a symmetric f such that $f \circ g$ is hard. This follows from a reduction that uses a binary search strategy. We get the desired lower bound for $\text{MAJ} \circ g$ using our lower bound for $\text{MOD}_m \circ g$, which is described below.

As a corollary to our result, we answer an open question posed by Babai et al. [BGKL03]. For any odd prime k, let $\text{QCSB}_k : \{0,1\}^k \to \{0,1\}$ be defined as $\text{QCSB}_k(y_1, \ldots, y_k) = 1$ iff $y_1 + \cdots + y_k$ is a quadratic residue modulo k. Babai et al. show that QCSB_k is not compressible and they ask the question of determining the communication complexity of $\text{MAJ} \circ \text{QCSB}_k$. Our result implies that if $k \equiv 1 \mod 4$, $\text{MAJ} \circ \text{QCSB}_k$ has cost $O(k \log n)$ in the simultaneous deterministic model, and if $k \equiv 3 \mod 4$, the function is hard in the randomized model for up to $c \log n$ many players with $c < 1/2$ (Corollary 3.9). For $k > 1 + \log n$, our efficient simultaneous protocol for $\text{SYM} \circ g$ implies an $O(\log^3 n)$ upper bound on the simultaneous communication complexity for every g.

Mod m of g. We show that if m divides $|S_0| - |S_1|$, then $\text{MOD}_m \circ g$ has a simultaneous deterministic protocol of cost $O(k \log m)$, and if m does not divide $|S_0| - |S_1|$, $\text{MOD}_m \circ g$ is a very hard function[3] 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 3.4). These types of functions generalize the GIP $= \text{MOD}_2 \circ \text{AND}$ function. The first strong lower bounds in the NOF model were obtained by Babai, Nisan and Szegedy [BNS92], who showed a very strong lower bound for the GIP function. Grolmusz [Gro95] extended the technique of [BNS92] to show a lower bound for $\text{MOD}_m \circ \text{AND}$. We obtain our lower bound for $\text{MOD}_m \circ g$, where m is coprime to $|S_0| - |S_1|$, by extending the analysis of [CT93, Raz00]. For other m for which $\text{MOD}_m \circ g$ is hard (i.e., m and $|S_0| - |S_1|$ are not coprime but m does not divide $|S_0| - |S_1|$), we obtain the lower bound through a reduction to the previous case, employing ideas from our protocol for $\text{SYM} \circ g$. Our upper bound result follows from Fact 3.3 with a minor addition.

[3] 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.

Nor of g. Functions of the form NOR $\circ\, g$ generalize the DISJ function. We observe that if g's support size is 1, then NOR $\circ\, g$ is hard in the randomized bounded error model for k up to $\frac{1}{2}\log n$. This follows trivially from the best known lower bound result on the DISJ function [She11]. If g's support size is not 1, we show that NOR $\circ\, g$ has a classical randomized protocol of cost $O(k)$ (Theorem 3.11). Thus, the hardness of DISJ crucially relies on the fact that g has singleton support. Our upper bound is obtained by combining our deterministic upper bound for $\text{MOD}_2 \circ g$ with a random sampling strategy.

1.1 Connections with Ramsey theory

For an Abelian group G, define $c_k(G)$ to be the minimum number of colors we can use to color G so that no k-term arithmetic progression is monochromatic. Also let $r_k(G)$ be the cardinality of the largest subset of G that contains no length k arithmetic progression. Let $N = |G|$. We use the notation $c_k(N)$ and $r_k(N)$ when working over $[N]$ rather than an Abelian group G. The famous Van der Waerden's Theorem and Szemerédi's Theorem are equivalent to showing $c_k(N) = \omega(1)$ and $N/r_k(N) = \omega(1)$ respectively. Obviously we have $N/r_k(N) \leq c_k(N)$. Obtaining good quantitative bounds on $c_k(N)$ and $r_k(N)$ is one of the major challenges in combinatorial mathematics.

The best known bounds for $r_k(N)$ are as follows (we write the bounds in terms of $N/r_k(N)$ as the interest is in this fraction). Sanders [San11] recently showed that

$$\frac{N}{r_3(N)} \geq \Omega\left(\frac{\log N}{(\log\log N)^5}\right),$$

and the best upper bound comes from Behrend's construction of a set without a 3-term progression [Beh46] (in [Elk10], Elkin obtains a minor improvement):

$$\frac{N}{r_3(N)} \leq O\left(2^{\sqrt{8\log N}}(\log N)^{1/4}\right).$$

For general k, the best bounds are

$$\frac{N}{r_k(N)} \geq \Omega\left((\log\log N)^{t_k}\right),$$

(t_k is a positive constant that depends only on k) due to Gowers [Gow01], and

$$\frac{N}{r_k(N)} \leq C \cdot 2^{O((\log N)^{1/\log k} + \log\log N)},$$

for a constant C, due to O'Bryant [O'B11].

It has been observed several times that the above *lower bound* results are in fact easier and cleaner to handle when working over $\mathbb{F}_p^n$ as one can exploit linear algebraic tools. As Green notes [Gre05], another motivation to work in the finite field setting is inspired by Bourgain's work [Bou99], which can be interpreted to show how to convert results obtained in the finite fields setting to arbitrary groups.

Very recently Bateman and Katz [BK11], in a breakthrough work, show that

$$\frac{N}{r_3(\mathbb{F}_3^n)} \geq \Omega\left((\log N)^{1+\epsilon}\right).$$

Non-trivial upper bounds are harder to come by in the finite field setting. Behrend's construction does not work over $\mathbb{F}_p^n$. The best upper bound we have for $N/r_3(\mathbb{F}_3^n)$ is much weaker and is about $N^{0.28}$, which comes from design theory; see e.g., [Gre05, Section 4]. It is reasonable to expect, both in the setting of $[N]$ and $\mathbb{F}_p^n$, that the lower bounds are far from being tight. For instance, Green [Gre05] conjectures that

$$\frac{N}{r_3(\mathbb{F}_3^n)} \geq N^{\delta},$$

for an absolute constant δ.

A well known generalization of Van der Waerden's Theorem and Szemerédi's Theorem is called the *multidimensional version* or the *corners* problem. In the k dimensional setting, our space is G^k rather than G, and the structure we are looking for is a *corner* rather than an arithmetic progression. A k dimensional *corner* is a set of $k+1$ points in G^k of the form

$$(x_1, x_2, \ldots, x_k), (x_1 + \lambda, x_2, \ldots, x_k), (x_1, x_2 + \lambda, \ldots, x_k), \ldots, (x_1, x_2, \ldots, x_k + \lambda),$$

for some non-zero $\lambda \in G$.

Let $c_k^{\measuredangle}(G)$ be the minimum number of colors we can use to color G^k so that no k-dimensional corner is monochromatic. Also let $r_k^{\measuredangle}(G)$ be the cardinality of the largest subset of G^k that contains no k-dimensional corner. As before, we use the notation $c_k^{\measuredangle}(N)$ and $r_k^{\measuredangle}(N)$ when working over $[N]^k$.

In a far reaching extension of Szemerédi's Theorem, Gowers [Gow07] obtains an explicit lower bound on $N^k/r_k^{\measuredangle}(N)$, but the bound is of Ackerman type and we do not state it here[4]. This bound remains best known for arbitrary fixed k. In the two dimensional case (which can be thought of as the generalization of Roth's Theorem [Rot53], i.e., Szemerédi's Theorem for $k = 3$.), Shkredov [Shk06b, Shk06a] obtains the bound

$$\frac{N^2}{r_2^{\measuredangle}(N)} \geq (\log\log N)^{\epsilon}.$$

The best upper bound comes from Behrend's construction via a reduction. In the finite field setting, a better lower bound is obtained by Lacey and McClain [LM07]:

$$\frac{N^2}{r_2^{\measuredangle}(\mathbb{F}_p^n)} \geq \frac{\log\log N}{\log\log\log N}.$$

To the best of our knowledge, no non-trivial upper bound on $N^k/r_k^{\measuredangle}(\mathbb{F}_p^n)$ is mentioned in the literature.

There is an interesting connection between the coloring number for corners and multiparty communication complexity. Define the EXACT_N function to be equal to -1 if and only if $x_1 + \cdots + x_k = N$, where x_i are the inputs, each an n-bit integer in $[N]$. Chandra Furst and Lipton [CFL83] show that the $k+1$ party deterministic communication complexity of EXACT_N is essentially equal to $\log c_k^{\measuredangle}(N)$. The known lower bounds on $N^k/r_k^{\measuredangle}(N)$ [FK78, Gow07] imply superconstant lower bounds on $c_k^{\measuredangle}(N)$ and using this, they conclude that the deterministic k-party communication complexity of EXACT_N is superconstant for all constant k. Furthermore, they convert the known upper bound on $N/r_3(N)$ due to Behrend into an upper bound on $c_2^{\measuredangle}(N)$ and obtain a surprising

[4]The bound Gowers obtains is similar to what Szemerédi obtains in the setting of progressions. This is because Gowers generalizes Szemerédi's Regularity Lemma to hypergraphs and this step is responsible for the horrendous bound.

non-explicit protocol of cost $O(\sqrt{n})$ for the EXACT_N function for 3 players. Although this and other kinds of communication complexity bounds have been proven using Ramsey theory (e.g. [CFL83, Pud03, Tes03, CKK+07, BGG06]), no bounds on Ramsey numbers have been proven via communication complexity bounds before.

For an Abelian group G, define $\text{EVAL}_G : G^k \to \{\pm 1\}$ to be equal to -1 if and only if $x_1 + \cdots + x_k = 0$, where the $x_i \in G$ are the inputs, and 0 denotes the identity element of G. As observed in [BGG06], the proof of [CFL83] also shows that the $k+1$ party communication complexity of EVAL_G is essentially equal to $\log c_k^{\swarrow}(G)$. In this paper, we are interested in upper bounds on $c_k^{\swarrow}(N)$ and $c_k^{\swarrow}(\mathbb{F}_2^n)$, which in return give upper bounds for $N^k/r_k^{\swarrow}(N)$ and $N^k/r_k^{\swarrow}(\mathbb{F}_2^n)$.

Our Results:

- We observe that $\text{EVAL}_{\mathbb{F}_2^n}$ is the same function as NOR $\circ$ XOR. Using our protocol for SYM $\circ$ g discussed in the first part of the introduction, we get the upper bound $N^k/r_k^{\swarrow}(\mathbb{F}_2^n) \leq c_k^{\swarrow}(\mathbb{F}_2^n) \leq O(N^{1/2^{k-2}} \log^{k+1} N)$ (Corollary 4.3). As far as we are aware, this result gives the first non-trivial upper bound and we suspect that it is essentially tight. For $k \geq \log\log N$, our bounds imply the following strong bounds: $N^k/r_k^{\swarrow}(\mathbb{F}_2^n) \leq c_k^{\swarrow}(\mathbb{F}_2^n) \leq O((\log N)^{5+\log\log N})$. The coloring induced by the protocol does not give an explicit large set without a corner. We provide such an explicit set with a simple description (Theorem 4.4). Our results can be considered as the first application of communication complexity to Ramsey theory.

- Recall that Behrend [Beh46] showed $N/r_3(N) \leq O(2^{\sqrt{8\log N}}(\log N)^{1/4})$. This result does not imply any bounds for $c_3(N)$. We observe that Behrend's idea can be used to give an explicit coloring of $[N]$ and obtain the bound $c_3(N) \leq 2^{\sqrt{8\log N}}(2\log N)^{1/2}$. This upper bound, via a standard reduction, also gives an upper bound for $c_2^{\swarrow}(N)$. Using this, we present an *explicit* protocol of cost $O(\sqrt{n})$ for the EXACT_N function for 3 players. As mentioned before, [CFL83] gets the same upper bound with a non-explicit protocol using a probabilistic argument.

Organization of the paper: In Section 2, we set the notation and give the necessary background on communication complexity. In Section 3, we present our results on the communication complexity of composed functions. Section 4 is devoted to the connections with Ramsey theory. Finally in Section 5, we conclude with some open problems.

2 Preliminaries

In the k-party 'number on the forehead' model of communication complexity, there are k players who are given k inputs $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2, \ldots, x_k \in \mathcal{X}_k$ respectively. Player i sees every input that is not assigned to her, i.e., player i sees x_j for all $j \neq i$. Given a function $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{\pm 1\}$, the players communicate, by broadcasting bits, according to a predetermined protocol. Their goal is to compute the value $F(x_1, \ldots, x_k)$. The protocol determines, in every round:

- whose turn it is to communicate, as a function of the communicated bits thus far,

- what the player communicates, as a function of the inputs the player sees and the communicated bits thus far.

Once the protocol determines that communication between players is over, it determines the output bit as a function of all the communicated bits. The cost of the protocol is the maximum over all possible inputs of the number of bits communicated. The k-party deterministic communication complexity of F, denoted $\mathbf{D}_k(F)$, is the cost of the most efficient protocol that computes F correctly on every input.

In the *simultaneous* model, the players, without interacting with each other, communicate bits to a referee who does not see the input. The referee then determines the output based on the messages she receives. The simultaneous communication complexity of F, denoted by $\mathbf{D}_k^{||}(F)$, is the cost of the most efficient simultaneous protocol that computes F.

In the randomized model, the players have access to an unbounded length random string which they all see. Furthermore, we allow the randomized protocol to make an error with probability at most ϵ on every input. The ϵ-error randomized communication complexity of F, denoted $\mathbf{R}_k^{\epsilon}(F)$, is the cost of the most efficient randomized protocol for F (the number of random bits used does not count towards the cost). A stronger model allowing quantum communication between the players can similarly be defined; see e.g., [LSS09]. As mentioned earlier, we point out that all the lower bounds in the randomized model that we prove here carry over to the quantum model using the results of [LSS09].

We say that a subset C_i of the input space $\mathcal{X}_1 \times \cdots \times \mathcal{X}_k$ 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 $C : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{0, 1\}$.

An equivalent way of defining a cylinder intersection is through the notion of a *star*. A set of k points

$$(x_1', x_2, \ldots, x_k), (x_1, x_2', \ldots, x_k), \ldots, (x_1, x_2, \ldots, x_k')$$

in $\mathcal{X}_1 \times \cdots \times \mathcal{X}_k$ is called a star if $x_i' \neq x_i$ for all $i \in [k]$. The point $(x_1, x_2, \ldots, x_k)$ is called the *center* of the star. It is not difficult to show that C is a cylinder intersection if and only if the center of every star in C also belongs to C as well.

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

$$\mathrm{disc}_\mu(F, C) = \left| \mathbf{E}_{x \sim \mu} \left[F(x) C(x) \right] \right|.$$

Note that when F is ± 1 valued, the discrepancy measures how balanced F is under μ in the cylinder intersection C. 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. If the discrepancy of a function is small (e.g. exponentially small), then all large cylinder intersections are balanced with respect to F's output. The well-known discrepancy method gives a lower bound for $\mathbf{R}_k^{\epsilon}(F)$ in terms of $\mathrm{disc}_\mu(F)$. Informally it states that to lower bound $\mathbf{R}_k^{\epsilon}(F)$, it suffices to find a distribution μ and upper bound the discrepancy $\mathrm{disc}_\mu(F)$.

Lemma 2.1 (Discrepancy Method). *Let $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{\pm 1\}$ be a function and μ a distribution over the input space. Then,*

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

(All logarithms in this paper are to base 2.)

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

$$\mathcal{E}_{\mu}(F) = \mathbf{E}_{\substack{x^0_1, x^0_2, \ldots, x^0_k \\ x^1_1, x^1_2, \ldots, x^1_k}} \left[\prod_{u \in \{0,1\}^k} \mathcal{C}^{u_1 + \cdots + u_k}(F(x^{u_1}_1, \ldots, x^{u_k}_k)) \right],$$

where in the expectation, x^0_i and x^1_i are distributed according to μ_i, and $\mathcal{C}$ denotes the complex conjugation operator: $\mathcal{C}^b(z) = z$ if b is even, and $\mathcal{C}^b(z) = \overline{z}$ 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 $\mathbb{F}^n_2$ (here $\oplus$ denotes bit-wise xor of the strings).

Lemma 2.2 ([CT93, Raz00, VW08]). *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 we will be mainly interested in the case where $\mathcal{X}_i = \{0,1\}^n$ 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. Most of the communication functions that we are interested in will be composed functions $f \circ \vec{g}$, where $f : \{0,1\}^n \to \{\pm 1\}$, $\vec{g} = (g_1, \ldots, g_k)$ with $g_i : \{0,1\}^k \to \{0,1\}$, and $f \circ \vec{g}(x_1, \ldots, x_k) = f(\ldots, g_i(x_{1,i}, x_{2,i}, \ldots, x_{k,i}), \ldots)$. That is, we apply g_i to the ith column of X, and then apply f to the resulting n-bit string to obtain the output. When all the g_i are the same function g, we denote the composed function by $f \circ g$. Whenever the domain of f is $\{0,1\}^{n'}$ instead of the usual $\{0,1\}^n$, the function is denoted $f^{n'}$.

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 vertex $v \in \{0,1\}^k$, we define n_v to be the number of occurrences of v 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 number of input variables set to 1. In other words, for any permutation σ on $[n]$, $f(y_1, \ldots, y_n) = f(y_{\sigma(1)}, \ldots, y_{\sigma(n)})$ holds for every $(y_1, \ldots, y_n) \in \{0,1\}^n$. In this section we present a deterministic

protocol for $\text{SYM} \circ \vec{g}$ where SYM denotes an arbitrary symmetric function and $\vec{g} = (g_1, \ldots, g_n)$ is a vector of n arbitrary boolean functions g_i. The protocol becomes simultaneous and efficient when $k > 1 + 2\log n$. For $k > 1 + \log n$ we obtain an efficient simultaneous protocol for $\text{SYM} \circ g$ for an arbitrary function g.

A multiparty non-simultaneous protocol for such a function, $\text{GIP} = \text{MOD}_2 \circ \text{AND}$, was first discovered by Grolmusz [Gro94]. This protocol is non-trivial for all k but only efficient when k reaches $\log n$. It is not difficult to see that the protocol also works for $\text{SYM} \circ \text{AND}$. Later Pudlák [Pud06] gave a non-simultaneous protocol for $\text{SYM} \circ \text{AND}$, which can be considered as a very elegant reinterpretation of Grolmusz's protocol (Pudlák's protocol is described in detail in [Cha08]). Babai et al. [BGKL03], using a new idea, obtained a simultaneous protocol for $\text{SYM} \circ g$ where g is a symmetric and compressible function, when $k > 1 + \log n$ (see [BGKL03, Section 6] for the definition of a 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. [BGKL03] in fact identify the *quadractic character of the sum of bits* function as a symmetric inside function g for which their method fails.

We improve upon the result of [BGKL03] in two ways. First, we remove the symmetry and compressibility conditions on g and allow inside function(s) to be selected arbitrarily, and second, we provide a non-trivial protocol even when $k \leq 1 + \log n$. We obtain our protocols in the non-simultaneous model by extending the ideas of Grolmusz and Pudlák. We combine this with a beautiful lemma of Babai et al. [BGKL03, Lemma 6.10] in order to make our protocols simultaneous.

Lemma 3.1 ([BGKL03])**.** *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 there is a simultaneous deterministic protocol in which each player sends at most $O(k \log n)$ bits to a referee, who then can compute n_i for all $i \in \{0, \ldots, n\}$.*

Theorem 3.2. *Let $f : \{0,1\}^n \to \{\pm 1\}$ be a symmetric function, $g : \{0,1\}^n \to \{0,1\}$ an arbitrary function, and $\vec{g} = (g_1, \ldots, g_n)$ a vector of n functions where $g_i : \{0,1\}^k \to \{0,1\}$ are arbitrary functions. Then,*

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

Proof. We first prove part **(a)**. Fix an input for $f \circ \vec{g}$ given in $k \times n$ matrix form X. The protocol proceeds in two steps. In the first step, the players determine the column positions of some $u \in \mathcal{H}_k$. At the second step, they use this information to compute the output of $f \circ \vec{g}$.

We now describe the first step. Let $X^{\geq 3}$ denote the $(k-2) \times n$ dimensional submatrix of X where the first two rows are deleted. Since $X^{\geq 3}$ has n columns and there are 2^{k-2} possible strings of length $k-2$, the string $s \in \{0,1\}^{k-2}$ that appears the least number of times as a column of $X^{\geq 3}$ appears at most $n/2^{k-2}$ times. Without any communication, Player 1 and Player 2 agree on this string (breaking ties in say lexicographical order). Player 2, using at most $n/2^{k-2}$ bits of communication, sends Player 1 the bits on Player 1's forehead corresponding to the positions where string s appears. With this information, Player 1 knows the positions of four vertices $00s$,

$01s$, $10s$ and $11s$ in $\mathcal{H}_k$. Now Player 1 announces one of these vertices (call it u) and the column indices corresponding to u. The total cost is at most $O(k + n/2^k \cdot \log n)$.

We proceed to step 2. To convey the main idea of the protocol, we start by considering functions of the form $f \circ g$, i.e., the inner function is g for all the columns. Let $S = g^{-1}(1)$ be the support of g. Recall that n_v denotes the number of occurrences of v as a column of X. As f is symmetric, $f \circ g(X)$ can be computed without communication from the quantity $\sum_{v \in S} n_v$. Note that after the first step, n_u is known to all the players. Now fix some $v \in S$. Consider a shortest path between v and u in $\mathcal{H}_k$: $v = w_1, w_2, \ldots, w_t = u$, where $t \leq k + 1$. It is not difficult to see that we can write a telescoping sum to compute n_v:

$$n_v = \left(\sum_{i=1}^{t-1} (-1)^{i+1} (n_{w_i} + n_{w_{i+1}}) \right) + (-1)^{t+1} n_u. \tag{1}$$

A crucial observation is that each term $(-1)^{i+1}(n_{w_i} + n_{w_{i+1}})$ is known by some player. In fact, as w_i and w_{i+1} are adjacent vertices of the hypercube, they differ in exactly one coordinate $p \in \{1, \ldots, k\}$. Thus, Player p can compute $(-1)^{i+1}(n_{w_i} + n_{w_{i+1}})$ simply by looking at the portion of the input she sees; we say that Player p is responsible for the term $(-1)^{i+1}(n_{w_i} + n_{w_{i+1}})$. If each player communicates the term she is responsible for, it is simple to compute n_v. We should mention that these observations are already present in [Gro94, Pud06]. But recall that the quantity we want to compute is $\sum_{v \in S} n_v$. For this, we write

$$\sum_{v \in S} n_v = \sum_{v \in S} \left(\left(\sum_{i=1}^{t-1} (-1)^{i+1} (n_{w_i} + n_{w_{i+1}}) \right) + (-1)^{t+1} n_u \right), \tag{2}$$

where we omitted the dependence on v for w_i and t to keep the notation simple. For every v, each player can compute the term she is responsible for as described above. But observe that the players need not announce their terms for different vertices v individually. In fact, it is sufficient for each player to announce *the sum* of the terms she is responsible for. Moreover, since $\sum_{v \in S} n_v \leq n$, it actually suffices to send this value evaluated modulo $n + 1$. This requires at most $\lceil \log(n + 1) \rceil$ bits of communication so that in total we only need $k \cdot \lceil \log(n + 1) \rceil$ bits of communication for step 2 of the protocol.

Extending step 2 of the protocol for general functions $f \circ \vec{g}$ with possibly different functions for different columns is straightforward. One can associate a hypercube to each column j of the input and the objective is to compute a sum over all columns of a term analogous to (2). We provide the details of the proof for completeness and to introduce notation that will be used for the proof of part **(c)**. Denote by S_j the support of g_j, that is, $S_j = g_j^{-1}(1)$. 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. To compute the output of $f \circ \vec{g}$, it suffices to compute

$$\sum_{j=1}^{n} \sum_{v \in S_j} \mathbf{1}_j(v). \tag{3}$$

Observe that for the columns j of X that u appears in (which are known to all players using the first step), we can easily compute $\sum_{v \in S_j} \mathbf{1}_j(v)$. Indeed, we have in this case $\sum_{v \in S_j} \mathbf{1}_j(v) = 1$ if $u \in S_j$ and 0 otherwise. It now remains to compute the sum over the remaining columns. Let j be a column in which u does not appear, and consider a shortest path from v to u in $\mathcal{H}_k$: $v = w_1, w_2, \ldots, w_t = u$. Since $\mathbf{1}_j(u) = 0$, we have

$$\mathbf{1}_j(v) = \sum_{i=1}^{t-1} (-1)^{i+1} (\mathbf{1}_j(w_i) + \mathbf{1}_j(w_{i+1})). \tag{4}$$

As described earlier, 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. Thus, to compute (3), each player announces the part of the sum she is responsible for, modulo $n+1$. Therefore this step of the protocol has cost at most $k \cdot \lceil \log(n+1) \rceil$. This completes the proof of part **(a)**. Observe that the second step of the protocol is simultaneous while the first step is not. When k is sufficiently large, we will be able to bypass the first step using Lemma 3.1.

We now prove part **(c)**. Let $\ell = \lfloor 2 + 2\log n \rfloor$. Only the first ℓ players will speak. For each column j, the rows $\ell+1$ to k naturally induce a function $g'_j : \{0,1\}^\ell \to \{0,1\}$; $g'_j(u) = g_j(u \cdot v)$ where $v \in \{0,1\}^{k-\ell}$ appears in column j from row $\ell+1$ to k. Thus our task reduces to finding a protocol for $f \circ \vec{g'}$ with ℓ players. From now on we drop the superscript in g'_j and denote the inner function by g_j.

As before we are interested in computing

$$\sum_{j=1}^{n} \sum_{v \in S_j} \mathbf{1}_j(v). \tag{5}$$

Let $\vec{0}$ be the all 0 vertex. Let $v \in S_j$ and let $v = w_1, \ldots, w_t = \vec{0}$ be a shortest path between v and $\vec{0}$. Then we have

$$\mathbf{1}_j(v) = \left(\sum_{i=1}^{t-1} (-1)^{i+1} (\mathbf{1}_j(w_i) + \mathbf{1}_j(w_{i+1})) \right) + (-1)^{|v|} \mathbf{1}_j(\vec{0}). \tag{6}$$

Substitute (6) into (5). Since the quantity in (5) is at most n, we can do arithmetic modulo $n+1$. As before, each term $(\mathbf{1}_j(w_i) + \mathbf{1}_j(w_{i+1}))$ in the sum is known to a player so the part of the sum involving these terms can be computed by the players using at most $\ell \cdot \lceil \log(n+1) \rceil$ bits. For each $j \in \{1, \ldots, n\}$, we group the terms involving $\mathbf{1}_j(\vec{0})$ when substituting (6) into (5) and let c_j be the coefficient of $\mathbf{1}_j(\vec{0})$ modulo $n+1$. We need to compute $\sum_j c_j \mathbf{1}_j(\vec{0})$, which can be done as follows. From the original $\ell \times n$ input matrix X, we create a new matrix X' by duplicating the jth column c_j many times. Observe that the coefficient c_j only depends on the function g_j and not on the input X and is thus known to all the players. As X' has at most n^2 columns and $\ell > 1 + \log(n^2)$ rows, we can apply Lemma 3.1 on X' to compute the number of all 0 columns in X', which is exactly what we want. In this step, each player sends a message of size $O(\log^2 n)$ leading to a total cost of $O(\log^3 n)$. So putting things together, we can compute (5) with at most $O(\log^3 n)$ bits of communication. The whole protocol is easily seen to be simultaneous. This completes the proof of part **(c)**.

We conclude with the proof of part **(b)**. The strategy is exactly the same as above. We need to calculate $\sum_j c_j \mathbf{1}_j(\vec{0})$. Since all the g_j are the same, $c_j = c$ for all j for some c. So we want to compute $c \sum_j \mathbf{1}_j(\vec{0})$, which is precisely $c n_{\vec{0}}$. We can compute $n_{\vec{0}}$ using Lemma 3.1 when $k > 1 + \log n$. So putting things together, we can compute (5) using at most $O(k^2 \log n)$ bits of communication. Using part **(c)** whenever $k > 1 + 2\log n$, this concludes the proof of part **(b)**. □

Remark. For functions of the form SYM$\circ g$, we can make a small improvement to part **(a)** and show $\mathbf{D}_k(\text{SYM} \circ g) \leq O(n/2^{k-2} + (k+1)\log n)$ as follows. In light of the proof of part **(b)** above, in step 1 of the protocol, all Alice needs to communicate is a vertex u and the value n_u. The column indices corresponding to u are not needed. Thus the cost of step 1 is at most $n/2^{k-2} + k + \lceil \log(n/2^{k-2}) \rceil = n/2^{k-2} + \lceil \log n \rceil + 2$. Combined with step 2, the total cost is at most $n/2^{k-2} + (k+1) \cdot \lceil \log(n+1) \rceil + 2$. In addition, we can also improve part **(c)** when we allow ourselves to be non-simultaneous and show $\mathbf{D}_k(\text{SYM} \circ \vec{g}) \leq O(\log^2 n)$. To see this, set $\ell = \lceil \log(n+1) \rceil$ in the proof of part **(c)**. Observe

that there is a vertex $u \in \{0,1\}^\ell$ that does not appear as a column in the first ℓ rows of the input matrix. Player k announces this vertex using ℓ bits. We replace $\vec{0}$ with u in the proof and note that $\mathbf{1}_j(u) = 0$ for all j. Therefore the desired output can be computed using $\ell + \ell \cdot \lceil \log(n+1) \rceil = \lceil \log(n+1) \rceil^2 + \lceil \log(n+1) \rceil$. These slightly improved upper bounds will be used in Section 4.1.

In what follows, we will determine the communication complexities of $\text{MOD}_m \circ g$, $\text{MAJ} \circ g$ and $\text{NOR} \circ g$, for any boolean function g. All these functions are of the form $\text{SYM} \circ g$ and so for $k > 1 + \log n$, the $O(\log^3 n)$ simultaneous communication complexity upper bound just presented applies to these functions. We note that we will not mention this $O(\log^3 n)$ upper bound explicitly and consider ourselves in the setting of $k \leq 1 + \log n$.

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

For $(y_1, y_2, \ldots, y_n) \in \{0,1\}^n$, the function $\text{MOD}_m(y_1, y_2, \ldots, y_n)$ takes values in $\{-1, 1\}$, and

$$\text{MOD}_m(y_1, y_2, \ldots, y_n) = -1 \text{ if and only if } \sum_{j=1}^{n} y_j = 0 \mod m.$$

In this section, we determine the k-party communication complexity of $\text{MOD}_m \circ g$, for every function g. Babai, Nisan and Szegedy [BNS92] show a lower bound of $\Omega(n/4^k)$ for the k-party randomized communication complexity of generalized inner product $\text{GIP} = \text{MOD}_2 \circ \text{AND}$. Their proof is later refined by [CT93, Raz00], where the technique of upper bounding the discrepancy via the cube measure (Lemma 2.2) is introduced. Grolmusz [Gro95] extends the analysis of [BNS92] to get an $\Omega(n/4^k)$ lower bound for $\text{MOD}_m \circ \text{AND}$, for constant m. Viola and Wigderson [VW08] obtain the same result by extending the analysis of [CT93, Raz00].

In this section we show that in general, the communication 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. (For the case where $g = \text{AND}$, considered in the mentioned papers, the support of g is $(1, 1, \ldots, 1)$, so $\big||S_0| - |S_1|\big| = 1$.) We prove a dichotomy for the communication complexity of $\text{MOD}_m \circ g$. When m divides $|S_0| - |S_1|$, we exhibit an efficient protocol by using ideas from the protocol for $\text{SYM} \circ g$ presented in the previous section. On the other hand, when m does not divide $|S_0| - |S_1|$, we show an $\Omega(n/m^2 4^k)$ lower bound (ignoring some additive logarithmic factors). The case of m not dividing $|S_0| - |S_1|$ is analyzed in two parts. When m and $|S_0| - |S_1|$ are coprime, we use the Discrepancy Method (Lemma 2.1) in conjunction with a careful analysis of the cube measure to obtain the desired lower bound. We prove that there is also a strong lower bound for randomized protocols in the remaining case (where m and $|S_0| - |S_1|$ are not coprime but m does not divide $|S_0| - |S_1|$) by giving a reduction to the previous case.

In the analysis of discrepancy, we will make use of the characterization of the MOD_m function in terms of exponential sums. Fix $2 \leq m \in \mathbb{N}$ and $0 \leq a, b \leq m-1$. Let $\omega = e^{2\pi i/m}$ be an m-th root of unity. For $(y_1, y_2, \ldots, y_n) \in \{0,1\}^n$ the function $\text{EXP}_m^{a,b}(y_1, y_2, \ldots, y_n)$ is defined to be

$$\text{EXP}_m^{a,b}(y_1, y_2, \ldots, y_n) = \omega^{a((\sum_{j=1}^n y_j) - b)}.$$

It is straightforward to check that for any b,

$$\frac{1}{m} \sum_{a=0}^{m-1} \text{EXP}_m^{a,b}(y_1, y_2, \ldots, y_n) \in \{0,1\}$$

and

$$\frac{1}{m}\sum_{a=0}^{m-1} \mathrm{EXP}_m^{a,0}(y_1, y_2, \ldots, y_n) = 1 \text{ if and only if } \mathrm{MOD}_m(y_1, y_2, \ldots, y_n) = -1. \tag{7}$$

Before presenting the main result of this section, we first state a fact which we need for our upper bound (when m divides $|S_0| - |S_1|$) and our reduction (when m and $|S_0| - |S_1|$ are not coprime but m does not divide $|S_0| - |S_1|$). This fact essentially follows from the argument presented in the proof of Theorem 3.2.

Fact 3.3. *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 (1). 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. □

Theorem 3.4. *Let $m \geq 2$ be an integer, $g : \{0,1\}^k \to \{0,1\}$ be a boolean function and $S = \{y \in \{0,1\}^k : g(y) = 1\}$ be its support. Define $S_0 = \{y \in S : y \text{ has even weight}\}$ and $S_1 = \{y \in S : y \text{ has odd weight}\}$. Then the function $\mathrm{MOD}_m \circ g$ satisfies the following:*

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

Proof. **Part (a):** Suppose that m divides $|S_0| - |S_1|$; we will give an efficient protocol for $\mathrm{MOD}_m \circ g$. Assume without loss of generality that $|S_0| \geq |S_1|$. 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 3.3. For the remaining elements in $S_0 - S_0'$, we simply pair them with $\vec{0}$. Therefore, using Fact 3.3 once again, we can compute $(|S_0| - |S_1|)n_{\vec{0}} + \sum_{v \in S_0 - S_0'} n_v \equiv \sum_{v \in S_0 - S_0'} n_v \mod m$. Thus, we have computed $\sum_{v \in S_0 \cup S_1} n_v \mod m$, from which the output of $\mathrm{MOD}_m \circ g$ 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 of communication.

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.

For $(y_1, y_2, \cdots, y_n) \in \{0,1\}^n$, define $f_m(y_1, \ldots, y_n) = \sum_j y_j \mod m$. Also for $b \in \{0, 1, \ldots, m-1\}$, let $f_m^b(y_1, \ldots, y_n) = 1$ if $\sum_j y_j \equiv b \mod m$, and 0 otherwise. Note that f_m^b are 0/1 valued functions rather than ± 1 valued like MOD_m. We define $F_m = f_m \circ g$ and $F_m^b = f_m^b \circ g$.

The strategy is as follows. Assume g is not constant. First note that by an elementary argument, one can show that the fraction of points x with $F_m(x) = b$ is roughly (with an exponentially small error) $1/m$ for all $b \in \{0, 1, \ldots, m-1\}$. It is possible to show that the same holds within any cylinder intersection that is not very small by analyzing the cube measure of the functions $\mathrm{EXP}_m^{a,b} \circ g$

with respect to the uniform distribution. This step uses the assumption that $|S_0| - |S_1|$ and m are coprime. It follows that in any sufficiently large cylinder intersection, the number of points x with $F_m^0(x) = 1$ is roughly the same as the number of points x with $F_m^1(x) = 1$. Define the distribution μ that puts equal weight to all x with $F_m^0(x) = 1$ and $F_m^1(x) = 1$. All other points get 0 weight. The discrepancy $\mathrm{disc}_\mu(\mathrm{MOD}_m \circ g)$ can now be easily upper bounded and this yields the desired lower bound via the Discrepancy Method (Lemma 2.1).

Let C be a cylinder intersection and $\mathcal{U}$ denote the uniform distribution over $(\{0,1\}^n)^k$. Recall that we use C to denote both a set $C \subseteq (\{0,1\}^n)^k$ and this set's characteristic function. For any b, we have

$$\mathbf{E}_x\left[F_m^b(x)C(x)\right] = \mathbf{E}_x\left[\frac{1}{m}\sum_{a=0}^{m-1}\mathrm{EXP}_m^{a,b}\circ g(x)C(x)\right] = \frac{1}{m}\sum_{a=0}^{m-1}\mathbf{E}_x\left[\mathrm{EXP}_m^{a,b}\circ g(x)C(x)\right],$$

where all the expectations are with respect to the uniform distribution. The term corresponding to $a = 0$ contributes $\frac{1}{m}\frac{|C|}{2^{nk}}$ to the sum, and thus we can write

$$\frac{1}{m}\frac{|C|}{2^{nk}} - error \le \mathbf{E}_x\left[F_m^b(x)C(x)\right] \le \frac{1}{m}\frac{|C|}{2^{nk}} + error, \tag{8}$$

where $error = \frac{1}{m}\sum_{a=1}^{m-1}\left|\mathbf{E}_x\left[\mathrm{EXP}_m^{a,b}\circ g(x)C(x)\right]\right|$. Note that the terms of this sum are exactly $\mathrm{disc}_{\mathcal{U}}(\mathrm{EXP}_m^{a,b}\circ g, C)$, which can be upper bounded using the cube measure (Lemma 2.2). The following lemma gives an upper bound on the cube measure of $\mathrm{EXP}_m^{a,b}\circ g$.

Lemma 3.5. *Assume m and $|S_0|-|S_1|$ are coprime. For any $a \in \{1,2,\ldots,m-1\}$ and $b \in \{0,1,\ldots,m-1\}$,*

$$\mathcal{E}_{\mathcal{U}}(\mathrm{EXP}_m^{a,b}\circ g) \le \frac{1}{e^{8n/(m^2 2^k)}}.$$

We defer the proof of this lemma to the end of the section to not break the flow. We can now upper bound the error:

$$error < \frac{1}{e^{8n/(m^2 4^k)}}.$$

From this, it easily follows that the number of points with $F_m^0(x) = 1$ is very close to the number of points with $F_m^1(x) = 1$, with exponentially small error:

$$\left|\mathbf{E}_x\left[(F_m^0(x) - F_m^1(x))C(x)\right]\right| = \left|\mathbf{E}_x\left[F_m^0(x)C(x)\right] - \mathbf{E}_x\left[F_m^1(x)C(x)\right]\right| \le 2\cdot error.$$

Recall the definition of μ, and let $\alpha > 0$ be the non-zero weight that μ assigns to a point in the support of F_m^0 and F_m^1. Then,

$$\begin{aligned}
\mathrm{disc}_\mu(\mathrm{MOD}_m\circ g, C) &= \left|\sum_x \mathrm{MOD}_m\circ g(x)C(x)\mu(x)\right| \\
&= \alpha\cdot\left|\sum_{\substack{x:\\ F_m^0(x)=1 \text{ or } F_m^1(x)=1}} \mathrm{MOD}_m\circ g(x)C(x)\right| \\
&= \alpha\cdot\left|\sum_x (F_m^0(x) - F_m^1(x))C(x)\right| \\
&= \alpha\cdot 2^{nk}\cdot\left|\mathbf{E}_x\left[(F_m^0(x) - F_m^1(x))C(x)\right]\right| \\
&\le \alpha\cdot 2^{nk}\cdot 2\cdot error.
\end{aligned}$$

It is not hard to check that α will be roughly $m/2 \cdot 1/2^{nk}$ and so the error term above will dominate. To see this, note that the whole input space $(\{0,1\}^n)^k$ is a cylinder intersection and so we can use (8) to obtain $\mathbf{E}_x\left[F_m^0(x)\right] = 1/2^{nk} \cdot |\text{support}(F_m^0)| = 1/m \pm error$. Similarly $1/2^{nk} \cdot |\text{support}(F_m^1)| = 1/m \pm error$. Since $\alpha \cdot 2^{nk} = 2^{nk}/(|\text{support}(F_m^0)| + |\text{support}(F_m^1)|)$, we have

$$\alpha \cdot 2^{nk} \leq \frac{1}{2/m - 2 \cdot error}.$$

Putting things together we get

$$\frac{1}{\text{disc}_\mu(\text{MOD}_m \circ g, C)} \geq \frac{1/m - error}{error} \geq \frac{e^{8n/(m^2 4^k)}}{m} - 1 \geq \frac{2^{11n/(m^2 4^k)}}{m} - 1.$$

Finally, we can apply the discrepancy method to conclude

$$\mathbf{R}_k^\epsilon(\text{MOD}_m \circ g) \geq \frac{11n}{m^2 4^k} + \log(1-2\epsilon) - \log m - 1. \tag{9}$$

Part (b), Case 2: We now consider the case where m does not divide $|S_0| - |S_1|$, but $\gcd(m, |S_0| - |S_1|) > 1$. The lower bound here is obtained via a reduction to the previous case. We assume for the remainder of the proof that $|S_0| - |S_1| > 0$. The case $|S_0| - |S_1| < 0$ can be handled in the same way. Let $1 < d = \gcd(m, |S_0| - |S_1|)$, and let $m = dq$ and $|S_0| - |S_1| = dr$, where q and r are coprime integers. Because m does not divide $|S_0| - |S_1|$, $q \geq 2$. Our strategy is to use a protocol for $\text{MOD}_m \circ g$ in order to construct a protocol for $\text{MOD}_q \circ g'$ for some function g' for which we can apply the lower bound on the randomized communication complexity given in (9).

We start by partitioning the set S_0 into sets $S_0', T_1, \ldots, T_d$ with $|S_0'| = |S_1|$ and $|T_1| = \cdots = |T_d| = r$. Let g' be the function whose support is T_1. Note that the support of g' has size r and consists only of inputs of even Hamming weight. So we can apply the lower bound (9) to $\text{MOD}_q \circ g'$.

Using a protocol for $\text{MOD}_m \circ g$, we will construct a protocol for $\text{MOD}_q \circ g'$ as follows. Fix an input $X \in \{0,1\}^{k \times n'}$ in matrix form. Recall that for each $v \in \{0,1\}^k$, n_v denotes the number of occurrences of v as a column of X. First, using Fact 3.3 we can compute $\sum_{v \in S_0' \cup S_1} n_v \mod m$ using $k\lceil \log m \rceil$ bits of communication. Again using Fact 3.3, for any $\ell \in \{2, \ldots, d\}$, the difference $\sum_{v \in T_\ell} n_v - \sum_{v \in T_1} n_v \mod m$ can also be computed at a cost of $k\lceil \log m \rceil$ bits. As a result, we can compute

$$\sum_{v \in S_0' \cup S_1} n_v + \sum_{\ell=2}^{d} \Big(\sum_{v \in T_\ell} n_v - \sum_{v \in T_1} n_v \Big) \equiv \sum_{v \in S} n_v - d \sum_{v \in T_1} n_v \mod m.$$

Let $s = s(X)$ denote this number. Observe that $\sum_{v \in T_1} n_v \equiv 0 \mod q$ if and only if $d \sum_{v \in T_1} n_v \equiv 0 \mod m$. So $\sum_{v \in T_1} n_v \equiv 0 \mod q$ if and only if $\sum_{v \in S} n_v \equiv s \mod m$. The latter can be determined by running the protocol for $\text{MOD}_m \circ g$ on the input which is obtained from X (viewed as an $k \times n'$ array) by appending $m - s$ columns all of which belong to S.

In short, the protocol for $\text{MOD}_q \circ g'$ on inputs from $(\{0,1\}^{n'})^k$ consists of two steps: First, the players compute s. Then they simulate the protocol for $\text{MOD}_m \circ g$ on the input of size $(\{0,1\}^n)^k$ specified above, where $n = n' + (m - s)$.

Suppose that we can use c bits to compute $\text{MOD}_m \circ g(X)$ when X is of size $k \times n$. Then the cost of the above protocol is $c + k\lceil \log m \rceil$. Using the fact that $n' = n - (m - s) > n/2$, and (9), we conclude

$$c + k\lceil \log m \rceil \geq \frac{5n}{m^2 4^k} + \log(1-2\epsilon) - \log m - 1.$$

That is,

$$\mathbf{R}_k^{\epsilon}(\mathrm{MOD}_m \circ g) \geq \frac{5n}{m^2 4^k} + \log(1-2\epsilon) - (k+1)\lceil \log m \rceil - 1.$$

□

Corollary 3.6. *If $g : \{0,1\}^k \to \{0,1\}$ has even support size, then $\mathbf{D}_k^{\|}(\mathrm{MOD}_2 \circ g) \leq k$. Otherwise, $\mathbf{R}_k^{\epsilon}(\mathrm{MOD}_2 \circ g) \geq \frac{n}{4^k} + \log(1-2\epsilon) - k - 2$.*

Proof of Lemma 3.5

By definition of the cube measure, we have

$$\begin{aligned}
\mathcal{E}_{\mathcal{U}}(\mathrm{EXP}_m^{a,b} \circ g) &= \mathbf{E}_{\substack{x_1^0, x_2^0, \ldots, x_k^0 \\ x_1^1, x_2^1, \ldots, x_k^1}} \left[\prod_{u \in \{0,1\}^k} \mathcal{C}^{u_1 + \cdots + u_k}(\mathrm{EXP}_m^{a,b} \circ g(x_1^{u_1}, \ldots, x_k^{u_k})) \right] \\
&= \mathbf{E}_{\substack{x_1^0, x_2^0, \ldots, x_k^0 \\ x_1^1, x_2^1, \ldots, x_k^1}} \left[\prod_{u \in \{0,1\}^k} \mathcal{C}^{u_1 + \cdots + u_k} \left(\omega^{a \sum_{j=1}^n g(x_{1,j}^{u_1}, x_{2,j}^{u_2}, \ldots, x_{k,j}^{u_k}) - ab} \right) \right].
\end{aligned}$$

In the exponent of ω, we can safely ignore ab since exactly half of the terms in the product are conjugated. So without loss of generality we assume $b = 0$.

The first standard step is to write the cube measure $\mathcal{E}_{\mathcal{U}}(\mathrm{EXP}_m^{a,0} \circ g)$ as the n-fold product of the cube measure of $\omega^{a \cdot g(y_1, \ldots, y_k)}$. That is,

$$\mathcal{E}_{\mathcal{U}}(\mathrm{EXP}_m^{a,0} \circ g) = \left(\mathbf{E}_{\substack{y_1^0, \ldots y_k^0 \\ y_1^1, \ldots, y_k^1}} \left[\omega^{a \sum_{(u_1, \ldots, u_k) \in \{0,1\}^k} (-1)^{u_1 + \cdots + u_k} \cdot g(y_1^{u_1}, \ldots, y_k^{u_k})} \right] \right)^n,$$

where in the expectation, y_j^0 and y_j^1 are independent and uniformly distributed over $\{0,1\}$.

Observe that for every setting of y_j^0, y_j^1 (for $1 \leq j \leq k$) where $y_j^0 = y_j^1$ for some j, the sum in the exponent is 0, and thus the expression inside the expectation evaluates to 1. This happens with probability $(1 - \frac{1}{2^k})$. Now consider a setting of y_j^0, y_j^1 (for $1 \leq j \leq k$) where $y_j^0 \neq y_j^1$ for all $1 \leq j \leq k$. Simply write y_j for y_j^0. Then we can nicely write y_j^u as $y_j \oplus u$, for $u \in \{0,1\}$. Consequently,

$$\sum_{(u_1, \ldots, u_k) \in \{0,1\}^k} (-1)^{u_1 + \cdots + u_k} \cdot g(y_1^{u_1}, \ldots, y_k^{u_k}) = \sum_{(u_1, \ldots, u_k) \in \{0,1\}^k} (-1)^{u_1 + \cdots + u_k} g(y_1 \oplus u_1, \ldots, y_k \oplus u_k).$$

By letting $v_i = y_i \oplus u_i$, the last sum becomes

$$\begin{aligned}
(-1)^{y_1 + \cdots + y_k} \sum_{(v_1, \ldots, v_k) \in \{0,1\}^k} (-1)^{v_1 + \cdots + v_k} g(v_1, \ldots, v_k) &= (-1)^{y_1 + \cdots + y_k} \sum_{(v_1, \ldots, v_k) \in S} (-1)^{v_1 + \cdots + v_k} \\
&= (-1)^{y_1 + \cdots + y_k} (|S_0| - |S_1|).
\end{aligned}$$

This is either $|S_0| - |S_1|$ or $|S_1| - |S_0|$, depending on the parity of $y_1 + y_2 + \cdots + y_k$. Among all

tuples $(y_1, y_2, \ldots, y_k)$, exactly half of them have even parity. As a result,

$$\begin{aligned}
\mathbf{E}_{\substack{y_1^0,\ldots y_k^0 \\ y_1^1,\ldots,y_k^1}} \left[\omega^{a\sum_{u\in\{0,1\}^k}(-1)^{u_1+\cdots+u_k}\cdot g(y_1^{u_1},\ldots,y_k^{u_k})}\right] &= (1-\frac{1}{2^k}) + \frac{\omega^{a(|S_0|-|S_1|)}+\omega^{a(|S_1|-|S_0|)}}{2^{k+1}} \\
&= (1-\frac{1}{2^k}) + \frac{\mathrm{Re}(\omega^{a(|S_0|-|S_1|)})}{2^k} \\
&= 1 - \frac{1-\cos\left(\frac{2\pi}{m}\cdot a(|S_0|-|S_1|)\right)}{2^k} \\
&= 1 - \frac{2\sin^2(a(|S_0|-|S_1|)\pi/m)}{2^k}.
\end{aligned}$$

Because m and $|S_0|-|S_1|$ are coprime, $a(|S_0|-|S_1|)\pi/m$ is not a multiple of π, for $1 \le a \le m-1$. So $\sin^2(a(|S_0|-|S_1|)\pi/m) \ge \sin^2(\pi/m) \ge 4/m^2$. (Here we use the fact that $\sin(x) \ge 2x/\pi$ for $0 \le x \le \pi/2$.) Thus,

$$\mathcal{E}_{\mathcal{U}}(\mathrm{EXP}_m^{a,0} \circ g) \le \left(1 - \frac{8}{m^2 2^k}\right)^n \le \frac{1}{e^{8n/(m^2 2^k)}}.$$

3.3 MAJ $\circ$ g

For each $n \ge 1$, the *majority* function $\mathrm{MAJ}^n : \{0,1\}^n \to \{-1,1\}$ is defined as $\mathrm{MAJ}^n(y_1,\ldots,y_n) = -1$ if and only if $\sum_i y_i \ge n/2$. When no confusion arises we will drop the supercript n from MAJ^n. It is not difficult to show that the MAJ $\circ$ g functions are the hardest among the functions of the form SYM $\circ$ g:

Proposition 3.7. *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 $\epsilon \ge 0$,*

$$\mathbf{R}_k^{\epsilon'}(f \circ g) \le \mathbf{R}_k^{\epsilon}(\mathrm{MAJ}^{2n} \circ g) \cdot \lceil \log(n+1) \rceil,$$

where $\epsilon' = \epsilon \lceil \log(n+1) \rceil$.

Proof. If g is constant, the statement clearly holds. We assume g is not constant in the following. By a binary search strategy we will show how to use a communication protocol for $\mathrm{MAJ}^{2n} \circ g$ to compute a function $f \circ g$. Let Π_{2n} be a randomized protocol with cost c computing $\mathrm{MAJ}^{2n} \circ g$ with error ϵ. We are going to use this protocol to build a protocol that determines the number, w, of ones in $\{g(x_{1,1},\ldots,x_{k,1}),\ldots,g(x_{1,n},\ldots,g_{k,n})\}$. Since f is symmetric, $f \circ g(x_1,\ldots,x_k)$ can then be computed from w without communication.

The binary search algorithm for computing w proceeds in stages. During the search we maintain the condition that $w \in [\ell, u]$ for some interval $[\ell, u]$ whose length is halved after each stage. Initially, $\ell = 0$ and $u = n$. Suppose that at some stage we have $\ell \le w \le u$. In order to determine the values ℓ', u' for the next stage, we will determine whether $w \le \left\lfloor \frac{\ell+u}{2} \right\rfloor$ or not. Then, if $w \le \left\lfloor \frac{\ell+u}{2} \right\rfloor$, we set $\ell' = \ell$ and $u' = \left\lfloor \frac{\ell+u}{2} \right\rfloor$, otherwise we set $\ell' = \left\lfloor \frac{\ell+u}{2} \right\rfloor + 1$ and $u' = u$. Clearly, it takes at most $\lceil \log(n+1) \rceil$ stages to arrive at the exact value of w.

Players use a protocol for $\mathrm{MAJ}^{2n} \circ g$ to compare w and $\left\lfloor \frac{\ell+u}{2} \right\rfloor$ as follows. As g is not constant, we can define auxiliary input variables $x'_1,\ldots,x'_k$, all of which are bit strings of length n, such that the number of ones in $g(x'_{1,1},\ldots,x'_{k,1}),\ldots,g(x'_{1,n},\ldots,x'_{k,n})$ is exactly $n - \left\lfloor \frac{\ell+u}{2} \right\rfloor$. Now run the

protocol for $\text{MAJ}^{2n} \circ g$ on the input $x_1x'_1, \ldots, x_kx'_k$, where each $x_ix'_i$ is a $2n$-bit string obtained by concatenating x_i and x'_i. Clearly the output of this protocol tells us whehter $w \leq \lfloor \frac{\ell+u}{2} \rfloor$ or not.

We now analyze the error and communication cost of this protocol. Since there are $\lceil \log(n+1) \rceil$ stages, the total cost is at most $\lceil \log(n+1) \rceil$ times the cost for the majority protocol. Also, by a union bound, the protocol makes an error with probability at most $\lceil \log(n+1) \rceil \cdot \epsilon$. □

We can combine Proposition 3.7 with our lower bounds for $\text{MOD}_m \circ g$ functions (Theorem 3.4) to obtain a characterization for the communication complexity of $\text{MAJ} \circ g$ for every g.

Theorem 3.8. *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. Define $S_0 = \{y \in S : y \text{ has even weight}\}$ and $S_1 = \{y \in S : y \text{ has odd weight}\}$. Then the function $\text{MAJ} \circ g$ satisfies the following:*

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

Proof. The case where $|S_0| = |S_1|$ follows from Fact 3.3 by setting $m = n+1$.

Now consider the case where $|S_0| \neq |S_1|$. We use Proposition 3.7 to prove a lower bound on the randomized communication complexity of $\text{MAJ}^{2n} \circ g$. Observe that for some large enough constant c, $\prod_{p \leq ck \log k : p \text{ prime}} p > 2^k \geq ||S_0| - |S_1||$ because there are at least k primes in the set $\{2, 3, \ldots, ck \log k\}$. Thus, there exists a prime $m \leq ck \log k$ that does not divide $|S_0| - |S_1|$. Now applying Proposition 3.7 with $\epsilon = \frac{1}{3\lceil \log(n+1) \rceil}$, together with Theorem 3.4, and also using $m \leq k \log k$ and $k \leq \log n$, we get

$$\mathbf{R}^{\epsilon}_k(\text{MAJ}^{2n} \circ g) \geq \mathbf{R}^{1/3}_k(\text{MOD}_m \circ g) / \lceil \log(n+1) \rceil \geq \Omega\left(\frac{n}{(k \log k)^2 \cdot 4^k \log n} - \log\log n\right).$$

By a standard boosting argument we have

$$\mathbf{R}^{1/3}_k(\text{MAJ}^{2n} \circ g) \geq \Omega\left(\frac{n}{(k \log k)^2 \cdot 4^k \log n \ \log\log n}\right).$$

Finally, since $\text{MAJ}^{2n+1} \circ g$ is at least as hard as $\text{MAJ}^{2n} \circ g$, we obtain the desired result. □

To illustrate the above theorem, we apply it to some natural choices of inner functions g. Let $\text{THR}_t(y_1, \ldots, y_k) = 1$ if $\sum y_i \geq t$ and $\text{THR}_t(y_1, \ldots, y_k) = 0$ otherwise. If g is the threshold function THR_t for some $0 < t < n$, then it is simple to show that $\text{MAJ} \circ \text{THR}_t$ is always a hard function as long as the number of players is at most $\approx 1/2 \log n$. The functions $\text{MAJ} \circ \text{MOD}_m$ exhibit an interesting behaviour: For even m, the function $\text{MAJ} \circ \text{MOD}_m$ is always hard as long as the number of players is at most $\approx 1/2 \log n$. By contrast, for odd m, it has an efficient protocol for some values of k, namely when k is an odd multiple of m.

Theorem 3.8 can also be used to determine the communication complexity of a class of functions considered by Babai et al. [BGKL03]. 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 [BGKL03] 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 upper or lower

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 3.9. *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 $\mathbf{R}^{1/3}_k(\text{MAJ} \circ \text{QCSB}_k) \geq \Omega\left(\frac{n}{(k \log k)^2 4^k \log n \log\log n}\right)$.*
- *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 3.8. 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., [Sho09, 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., [Sho09, Theorem 2.20]. This implies that $|S_0| \neq |S_1|$. For $k > 1 + \log n$, we can use Theorem 3.2. □

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 significantly improves on the bounds of [LS09],[CA08] and [BHN09] on the multiparty bounded error communication complexity of disjointness:

Theorem 3.10 ([She11])**.**

$$\mathbf{R}^{1/3}_k(\text{DISJ}) \geq \Omega\left(\frac{n}{4^k}\right)^{1/4}.$$

First we observe that the above lower bound for disjointness applies - via a simple reduction - to $\text{NOR} \circ g$ where 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 3.11. *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. For some constant $\epsilon < 1/2$,*

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

Proof. For the first part, let $S = \{v\}$ with $v \in \{0,1\}^k$. Then, we can solve $\text{NOR} \circ \text{AND}$ on input X by first flipping all the input bits of the rows i for which $v_i = 0$ and then run a protocol for $\text{NOR} \circ g$. The lower bound then follows from Theorem 3.10.

For the upper bound, first assume that $|S|$ is even. In this case, by Corollary 3.6, 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. If $\text{NOR} \circ g(X) = 1$, then the bit string $g(X)$ is not the all-zero string and thus the parity of a random subset is uniformly distributed on $\{0, 1\}$, i.e., $\text{MOD}_2 \circ g(X_r) = 1$ with probability $1/2$. So 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 greater than 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 $\text{NOR} \circ g(X) = -1$ if and only if 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. □

4 Upper bounds on coloring numbers for corners

Recall from Section 1.1 the definitions of the Ramsey numbers $c_k^{\swarrow}(N)$, $c_k^{\swarrow}(G)$, $r_k^{\swarrow}(N)$, $r_k^{\swarrow}(G)$, and the functions EXACT_N and EVAL_G, where G is an Abelian group.

First, we state a result by Chandra, Furst and Lipton that connects multiparty communication complexity with coloring numbers for corners:

Lemma 4.1 ([CFL83])**.**

$$\log\left(c_k^{\swarrow}\left(\left\lceil\frac{N-1}{k}\right\rceil\right)\right) \leq \mathbf{D}_{k+1}(\text{EXACT}_N) \leq k + \log(c_k^{\swarrow}(N)).$$

As observed in [BGG06], such a connection, with essentially the same proof, also holds for the EVAL_G function. We provide a proof in the Appendix.

Lemma 4.2.

$$\log(c_k^{\swarrow}(G)) \leq \mathbf{D}_{k+1}(\text{EVAL}_G) \leq k + \log(c_k^{\swarrow}(G)).$$

4.1 Finite field setting

In Section 3.1, we presented a protocol for functions of the form $\text{SYM} \circ g$. Observe that $\text{EVAL}_{\mathbb{F}_2^n}$ can be written as $\text{NOR} \circ \text{XOR}$ and therefore the protocol described in Theorem 3.2 works for $\text{EVAL}_{\mathbb{F}_2^n}$. Using Lemma 4.2, we get an upper bound on $c_k^{\swarrow}(\mathbb{F}_2^n)$, and this in return gives a lower bound on $r_k^{\swarrow}(\mathbb{F}_2^n)$. The bounds below are obtained using the remark made right after the proof of Theorem 3.2.

Corollary 4.3. *Let $N = 2^n$. For any k,*

$$c_k^{\swarrow}(\mathbb{F}_2^n) \leq 16N^{1/2^{k-1}} \log^{k+2} N.$$

In particular, when $k > \log n$,

$$c_k^{\swarrow}(\mathbb{F}_2^n) \leq 32(\log N)^{3+\log\log N}.$$

The coloring above does not give an explicit large set that does not contain a corner. Below, we provide such an explicit set with a simple description that is inspired from the protocol of Theorem 3.2.

For $X = (x_1, \ldots, x_k) \in (\mathbb{F}_2^n)^k$, we denote by $n_i(X)$ the number of columns of X with Hamming weight i, i.e. $n_i(X) = |\{j \in \{1, \ldots, n\} : \sum_{\ell=1}^k X_{\ell,j} = i\}|$, where the sum $\sum_{\ell=1}^k X_{\ell,j}$ should be understood as an operation over the integers. Let $N_i = \left\lfloor \frac{\binom{k}{i}}{2^k - 1} n \right\rfloor$ for $i \in \{1, \ldots, k-1\}$ and $N_k = n - \sum_{i=1}^{k-1} N_i$ and

$$S^k = \left\{ X \in (\mathbb{F}_2^n)^k : \forall i \in \{1, \ldots, k\}, n_i(X) = N_i \right\}. \tag{10}$$

Observe that this implies that for all $X \in S^k$, $n_0(X) = 0$ and $n_i(X) \geq 1$ for $i \in \{1, \ldots, k\}$.

Theorem 4.4. *Let $n \geq 2$ and $2 \leq k \leq \lceil \log n \rceil$, and let $N = 2^n$. The set S^k defined above does not contain a corner, and*

$$|S^k| \geq C_k \frac{N^k}{N^{-\log(1-2^{-k})} \log^{k/2} N}$$

where C_k only depends on k. For $k = \lceil \log n \rceil$,

$$|S^k| \geq \frac{N^k}{(\log N)^{C \log \log N}}.$$

for some constant $C > 0$.

Proof. We first prove that S^k does not contain a corner. This part of the proof does not make use of the particular values chosen for N_i, in fact we prove that S^k as defined in (10) does not contain a corner for any choice of $N_1, \ldots, N_k$ satisfying $\sum_i N_i = n$. Recall that $n_0(X) = 0$ for all $X \in S^k$, and this is crucial for the argument. Assume that there exists $X \in S^k$ and non-zero $\lambda \in \mathbb{F}_2^n$ such that for all $\ell \in \{1, \ldots, k\}$, $X + \lambda^\ell \in S^k$ where $\lambda^\ell \in (\mathbb{F}_2^n)^k$ is zero except for the ℓ-th row where it is equal to λ. Consider the columns of X corresponding to indices j such that $\lambda_j = 1$. Let t denote the minimum Hamming weight among these columns. Note that $t > 0$. By the minimality of t, the columns of X with Hamming weight $t - 1$ remain intact in $X + \lambda^\ell$ for all $\ell \in \{1, \ldots, k\}$. So $n_{t-1}(X + \lambda^\ell) \geq n_{t-1}(X) = N_{t-1}$ for every ℓ. On the other hand, observe that by the choice of t, there is some ℓ' such that $n_{t-1}(X + \lambda^{\ell'}) > n_{t-1}(X)$. This is a contradiction.

We now move on to estimate the size of S^k. The values of N_i were picked so that S^k is as large as possible while keeping the size estimation simple. We will prove generally that for any $k \geq 2$,

$$|S^k| \geq (2^k - 1)^n \cdot \frac{1}{n^{k/2}} \cdot \frac{1}{2} \frac{e^{-2k-2k^2}}{\sqrt{2\pi}^{k-1}(1+k)}.$$

Then, to obtain the advertised bound, we write

$$(2^k - 1)^n \geq (2^n)^k (1 - 2^{-k})^n = \frac{N^k}{N^{-\log(1-2^{-k})}},$$

and we define $C_k = \frac{1}{2} \frac{e^{-2k-2k^2}}{\sqrt{2\pi}^{k-1}(1+k)}$. To obtain the bound for $k = \lceil \log n \rceil$, we observe that $N^{-\log(1-2^{-k})} \leq (1 - 1/n)^{-n} \leq 4$.

We use Stirling's approximation: for all $n \geq 1$

$$\left(\frac{n}{e}\right)^n \sqrt{2\pi n} \leq n! \leq e \cdot \left(\frac{n}{e}\right)^n \sqrt{2\pi n}.$$

We define the reals α_i such that $n_i = \frac{\alpha_i}{2^k-1}n$. Note that $\alpha_i \leq \binom{k}{i}$ for all $i \in \{1, \ldots, k-1\}$ and $n_k \leq \frac{1}{2^k-1}n + k$ so that $\alpha_k \leq 1 + \frac{2^k-1}{n}k$.

$$\begin{aligned}
|S^k| &= \binom{n}{n_1 \quad n_2 \quad \cdots \quad n_k} \cdot \binom{k}{1}^{n_1} \cdots \binom{k}{k}^{n_k} \\
&\geq \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}}{e^k \left(n_1 e^{-1}\right)^{n_1} \cdots \left(n_k e^{-1}\right)^{n_k} \sqrt{(2\pi)^k n_1 \cdots n_k}} \cdot \binom{k}{1}^{n_1} \cdots \binom{k}{k}^{n_k} \\
&\geq \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}}{e^k \left(\frac{\alpha_1}{2^k-1} n e^{-1}\right)^{n_1} \cdots \left(\frac{\alpha_k}{2^k-1} n e^{-1}\right)^{n_k} \sqrt{(2\pi)^k n_1 \ldots n_k}} \cdot \binom{k}{1}^{n_1} \cdots \binom{k}{k}^{n_k} \\
&\geq \frac{(2^k-1)^n \sqrt{2\pi n}}{e^k \alpha_k^{n_k} \sqrt{(2\pi)^k n_1 \ldots n_k}}.
\end{aligned}$$

Observe that

$$\alpha_k^{n_k} \leq \left(1 + \frac{(2^k-1)k}{n}\right)^{\frac{n}{2^k-1}+k} \leq e^{k+2k^2},$$

where we used the fact that $(2^k-1)/n \leq 2$ as $k \leq \lceil \log n \rceil$. Moreover,

$$n_1 \cdots n_k \leq \frac{\binom{k}{1} \cdots \binom{k-1}{k}(1+k)}{(2^k-1)^k} n^k \leq \frac{2^{k^2} \cdot (1+k) n^k}{(2^k-1)^k},$$

which gives the desired bound. □

4.2 An explicit $O(\sqrt{n})$-protocol for EXACT_N for 3 players

Using an elegant argument, Behrend [Beh46] shows that for any N sufficiently large, there is a subset of $[N] = \{1, 2, \ldots, N\}$ of size at least

$$\Omega\left(\frac{N}{2^{\sqrt{8 \log N}} (\log N)^{1/4}}\right) \tag{11}$$

that does not contain any (nontrivial) 3-term arithmetic progressions. We observe that Behrend's argument can actually be made to give an *explicit* coloring of $[N]$ using at most

$$2^{\sqrt{8 \log N}} (2 \log N)^{1/2} \tag{12}$$

colors such that there is no monochromatic 3-term arithmetic progression. Furthermore, Behrend's argument also shows that in our coloring there is a color class of the size stated in (11). This coloring will be used to obtain an explicit protocol for EXACT_N for three players.

Note that Behrend's result has been used in [CFL83] to show the existence of a $O(\sqrt{n})$-cost protocol for the EXACT_N function for three players. In fact, the large set which exists by Behrend's argument gives a large subset of $[N] \times [N]$ that does not contain a corner. Then a probabilistic argument shows that with high probability, a sufficiently large number of translations of this subset will cover the whole space $[N] \times [N]$. Each of these translations is colored by a distinct color, and this shows the existence of a protocol of cost $O(\sqrt{n})$.

Our observation shows that we can bypass the probabilistic step above. Moreover, the explicit protocol we obtain might give insight into how three players can cooperatively offer a much more efficient protocol than two players. Note also that the recent improvements on Behrend's result [Elk10, GW08, O'B11] only give a large set without a corner; to show the existence of a coloring, and hence of a protocol, one may have to use the probabilistic argument as in [CFL83]. So these do not give a protocol more efficient than the protocol presented below.

Let m and d be some parameters to be determined later. The first step is as in [Beh46]: for each $x \in [N]$ we write $x-1$ in base $(2m)$ as

$$x - 1 = x_0 + x_1(2m) + x_2(2m)^2 + \ldots + x_{d-1}(2m)^{d-1}, \tag{13}$$

where $0 \le x_i < 2m$ for $0 \le i < d$. Then, our coloring for $[N]$ is as follows. Define $S(x)$ to be the subset of indices i such that $x_i < m$, and define

$$t(x) = \big(\sum_{j \in S(x)} x_j^2\big) + \big(\sum_{j \notin S(x)} (x_j - m)^2\big).$$

Now we color x with the pair $(S(x), t(x))$.

Lemma 4.5. *In the above coloring of $[N]$ there is no monochromatic 3-term arithmetic progression. Moreover, for $d = \sqrt{2\log N}$ and $m = 2^{\sqrt{\frac{1}{2}\log N - 1}}$, the total number of colors needed is at most $2^{\sqrt{8\log N}}(2\log N)^{1/2}$.*

Proof. The fact that there is no monochromatic 3-term arithmetic progression can be seen as follows. Suppose that x, y, z have the same color, and that $x + y = 2z$. First, since x, y, z have the same color, we have

$$S(x) = S(y) = S(z).$$

From this and the hypothesis that $x + y = 2z$, we can prove by reverse induction on i that $x_i + y_i = 2z_i$, for all $0 \le i < d$. From this it follows that $x_i^2 + y_i^2 \ge 2z_i^2$ and $(x_i - m)^2 + (y_i - m)^2 \ge 2(z_i - m)^2$, and equality holds if and only if $x_i = y_i = z_i$. As a result, $t(x) + t(y) \ge 2t(z)$, and equality holds if and only if $x = y = z$. Now because x, y, z have the same color,

$$t(x) = t(y) = t(z),$$

so equality does indeed hold. It follows that $x = y = z$. This shows that there is no monochromatic (nontrivial) 3-term arithmetic progression.

Now the total number of colors is at most $2^d(d(m-1)^2 + 1)$, because there are 2^d possible sets $S(x)$, and $t(x) \le d(m-1)^2$. So for the values of d and m given in the lemma, it is easy to verify that the total number of colors needed is as stated in (12). Also, with this setting of the parameters, the analysis in [Beh46] shows that there is a color class of size given in (11). □

Note that the above setting of the parameters is optimal for the total number of colors needed in our coloring. This is because we need $(2m)^d \ge N$ in order to write $N - 1$ as in (13). This setting is also optimal for Behrend's analysis.

By a standard argument, i.e., mapping each $(x, y) \in [N]^2$ to $x + 2y$, we can exhibit an explicit coloring of $[N]^2$ without a monochromatic corner. Here we will use this to describe an explicit $O(\sqrt{n})$-protocol for the 3 player communication problem EXACT_{2^n}. Recall that in this problem there are three players: Alice, Bob, and Charlie, with inputs x, y, z ($0 \le x, y, z \le 2^n$) respectively on their foreheads. The players want to determine whether $x + y + z = 2^n$.

Our protocol is obtained by combining the above explicit coloring and the argument from [CFL83] (that shows how to obtain a protocol from a coloring, as in the proof of Lemma 4.2). It is based on the following observation. Let

$$x' = 2^n + y - z, \;\; y' = 2^{n+1} - x - 2z, \;\; z' = x + 2y.$$

Then $y' + z' = 2x'$, i.e. y', x', z' form a 3-term arithmetic progression. Moreover, $x + y + z = 2^n$ if and only if $x' = y' = z'$. In addition, x', y', and z' can be computed by Alice, Bob, and Charlie, respectively, without any communication.

The Protocol:

- Alice sends the color of x';
- Bob and Charlie send one bit each indicating whether y' and z' have the same color as x';
- The players conclude that $x + y + z = 2^n$ if and only if x', y', z' have the same color.

Here the colors of x', y', z' are determined as above, but note that $x', y', z' \in [2N]$. So we set $d = \sqrt{2(n+1)}$ and $m = 2^{\sqrt{(n+1)/2}}$. The cost of the protocol is at most

$$2 + \left\lceil \log(2^d d m^2) \right\rceil \leq 2\sqrt{2(n+1)} + \frac{1}{2}\log(n+1) + 4.$$

5 Open problems

There are many interesting open problems related to the topics studied in this paper. We state here a few of them.

The most natural direction to pursue is determining the communication complexity of $f \circ g$ for other natural functions f. For example, characterizing the communication complexity of $\text{THR}_t \circ g$ for all t and g would be quite interesting.

Getting good bounds on $c_k^{\swarrow}(G)$ and $r_k^{\swarrow}(G)$ is a major challenge. Can one make progress on this using the connection with communication complexity? Observe that EVAL_G has $O(1)$ complexity in the randomized model as it reduces to the 2 player EQUALITY function, which is the canonical example of a function with a very efficient randomized protocol. Hence, to show a good lower bound on $\mathbf{D}_k(\text{EVAL}_G)$, one needs to use a lower bound technique that does not apply to randomized protocols. So far, the only strong lower bound technique we have in the NOF model is the discrepancy method (and its extension called the generalized discrepancy method) which proves lower bounds for the randomized model. It is a major open problem in the NOF model to exhibit an explicit function which is easy in the randomized model but hard in the deterministic model, even for 3 players[5]. The EVAL_G function is of course a natural candidate. Can we develop new lower bound techniques that work only for the deterministic model?

Chandra, Furst and Lipton [CFL83] showed that the EXACT_N function for 3 players has an $O(\sqrt{n})$-cost protocol. Our protocol for $\text{EVAL}_{\mathbb{F}_2^n}$ has cost $\Theta(n)$ when k is a constant, but it has cost

[5]It is proved in [BDPW10], by a clever counting argument, that such functions exist.

$O(\log^2 n)$ when $k \geq \log n$. Does EXACT_N have an efficient protocol for $\log n$ many players? Is it possible to get a $o(n)$ cost protocol for $\text{EVAL}_{\mathbb{F}_2^n}$ for 3 players? We suspect that the answer to the latter question is no.

Our protocol for $\text{EVAL}_{\mathbb{F}_2^n}$ does not work for $\text{EVAL}_{\mathbb{F}_3^n}$. Can one get a similar bound for $\text{EVAL}_{\mathbb{F}_3^n}$, and for $\text{EVAL}_{\mathbb{F}_p^n}$ in general? The complexity of EVAL_G for other G is also interesting to study.

References

[AMS99] 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.

[BBM11] 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.

[BDPW10] Paul Beame, Matei David, Toniann Pitassi, and Philipp Woelfel. Separating deterministic from randomized multiparty communication complexity. *Theory of Computing*, 6(1):201–225, 2010.

[Beh46] Felix A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci*, 32:331–332, 1946.

[BGG06] Richard Beigel, William Gasarch, and James Glenn. The multiparty communication complexity of exact t: Improved bounds and new problems. In Rastislav Královic and Pawel Urzyczyn, editors, *Mathematical Foundations of Computer Science 2006*, volume 4162 of *Lecture Notes in Computer Science*, pages 146–156. Springer Berlin / Heidelberg, 2006.

[BGKL03] László Babai, Anna Gál, Peter G. Kimmel, and Satyanarayana V. Lokam. Communication complexity of simultaneous messages. *SIAM Journal on Computing*, 33:137–166, 2003.

[BHN09] 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.

[BK11] Michael Bateman and Nets H. Katz. New bounds on cap sets. http://arxiv.org/abs/1101.5851v2, 2011.

[BKL95] 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.

[BNS92] 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.

[Bou99] Jean Bourgain. On triples in arithmetic progression. *Geometric and Functional Analysis*, 1999.

[BT94] Richard Beigel and Jun Tarui. On ACC. *Computational Complexity*, 4:350–366, 1994.

[CA08] Arkadev Chattopadhyay and Anil Ada. Multiparty communication complexity of disjointness. Technical report, In Electronic Colloquium on Computational Complexity (ECCC) TR08–002, 2008.

[CFL83] 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.

[Cha08] Arkadev Chattopadhyay. *Circuits, Communication and Polynomials*. PhD thesis, McGill University, 2008.

[CKK+07] Arkadev Chattopadhyay, Andreas Krebs, Michal Koucky, Mario Szegedy, Pascal Tesson, and Denis Thérien. Languages with bounded multiparty communication complexity. In *Proceedings of the 24th annual conference on Theoretical aspects of computer science*, STACS'07, pages 500–511, Berlin, Heidelberg, 2007. Springer-Verlag.

[CS04] Vincent Conitzer and Tuomas Sandholm. Communication complexity as a lower bound for learning in games. In *International Conference on Machine Learning*, 2004.

[CT93] Fan R.K. Chung and Prasad Tetali. Communication complexity and quasi randomness. *SIAM Journal on Discrete Mathematics*, 6(1):110–123, 1993.

[Elk10] Michael Elkin. An Improved Construction of Progression-Free Sets. Accepted to Israeli J. Math, 2010.

[FK78] Harry Furstenberg and Yitzhak Katznelson. An ergodic Szemerédi theorem for commuting transformations. *Journal d'Analyse Mathematique*, 34:275–291, 1978.

[FKL+01] 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.

[Gow01] W. Timothy Gowers. A new proof of Szemerédi's theorem. *Geometric and Functional Analysis*, 11:465–588, 2001.

[Gow07] W. Timothy Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Annals of Mathematics*, 166:897–946, 2007.

[Gre05] Ben Green. *Surveys in Combinatorics 2005*, chapter Finite field models in additive combinatorics, pages 1–27. London Math. Soc. Lecture Notes 327. Cambridge Univ Press, 2005.

[Gro94] Vince Grolmusz. The BNS lower bound for multi-party protocols is nearly optimal. *Information and Computation*, 112:51–54, 1994.

[Gro95] 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.

[Gro98] Vince Grolmusz. Circuits and multi-party protocols. *Computational Complexity*, 7:1–18, 1998.

[GW08] Ben Green and Julia Wolf. A note on Elkin's improvement of Behrend's construction. http://arxiv.org/abs/0810.0732, 2008.

[HG91] Johan Håstad and Mikael Goldmann. On the power of small-depth threshold circuits. *Computational Complexity*, 1:610–618, 1991.

[Kla07] Hartmut Klauck. Lower bounds for quantum communication complexity. *Siam Journal on Computing*, 37:20–46, 2007.

[LM07] Michael T. Lacey and William McClain. On an argument of Shkredov on two-dimensional corners. *Online Journal of Analytic Combinatorics*, 2007.

[LS09] Troy Lee and Adi Shraibman. Disjointness is hard in the multiparty number-on-the-forehead model. *Computational Complexity*, 18:309–336, 2009.

[LSS09] 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.

[MNSW98] 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.

[MO09] Ashley Montanaro and Tobias Osborne. On the communication complexity of XOR functions. *Arxiv preprint arXiv:0909.3392*, 2009.

[Nis93] N. Nisan. The communication complexity of threshold gates. *Combinatorica*, 1993.

[NS06] Noam Nisan and Ilya Segal. The communication requirements of efficient allocations and supporting prices. *Journal of Economic Theory*, 129:192–224, 2006.

[NW93] Noam Nisan and Avi Wigderson. Rounds in communication complexity revisited. *Siam Journal on Computing*, 22:211–219, 1993.

[O'B11] Kevin O'Bryant. Sets of integers that do not contain long arithmetic progressions. *The Electronic Journal of Combinatorics*, 18(1):59, 2011.

[PPS07] 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.

[Pud03] Pavel Pudlák. An application of hindman's theorem to a problem on communication complexity. *Comb. Probab. Comput.*, 12:661–670, November 2003.

[Pud06] Pavel Pudlák, 2006. Personal communication.

[Raz95] Ran Raz. Fourier analysis for probabilistic communication complexity. *Computational Complexity*, 5:205–221, 1995.

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

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

[Rot53] Klaus F. Roth. On certain sets of integers. *Journal of The London Mathematical Society-second Series*, s1-28:104–109, 1953.

[San11] Tom Sanders. On Roths theorem on progressions. *Annals of Mathematics*, 174:619–636, 2011.

[She07] 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.

[She11] Alexander A. Sherstov. The multiparty communication complexity of set disjointness. Technical report, In Electronic Colloquium on Computational Complexity (ECCC) TR11–145, 2011.

[Shk06a] Ilya D. Shkredov. On a generalization of Szemerédi's theorem. *Proc. London Math. Soc.*, 93:723–760, 2006.

[Shk06b] Ilya D. Shkredov. On a problem of Gowers. *Izvestiya: Mathematics*, 70:385, 2006.

[Sho09] Victor Shoup. *A computational introduction to number theory and algebra*. Cambridge University Press, 2009.

[SZ09a] Yaoyun Shi and Zhiqiang Zhang. Communication complexities of symmetric XOR functions. *Quantum Information and Computation*, 9:255–263, 2009.

[SZ09b] Yaoyun Shi and Yufan Zhu. Quantum communication complexity of block-composed functions. *Quantum Information and Computation*, 9:444–460, May 2009.

[Tes03] Pascal Tesson. *Computational complexity questions related to finite semigroups and monoids*. PhD thesis, McGill University, 2003.

[VW08] Emanuele Viola and Avi Wigderson. Norms, xor lemmas, and lower bounds for polynomials and protocols. *Theory of Computing*, 4(1):137–168, 2008.

[Yao79] 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.

Appendix

Proof of Lemma 4.2

Upper bound: Fix a coloring of G^k with $c_k^{\not<}(G)$ colors so that there is no monochromatic corner. Denote the players' input by $x_1, \ldots, x_{k+1}$. For $i = 1, \ldots, k$, define $x_i' = -\sum_{j \neq i} x_j$, where the addition represents the operation of the group. Observe that $\mathrm{EVAL}_G(x_1, \ldots, x_{k+1}) = 1$ if and only if $x_i = x_i'$ for all $i = 1, \ldots, k$. Now, for $i = 1, \ldots, k$, Player i computes the color of $(x_1, \ldots, x_i', \ldots, x_k)$. Player $k+1$ computes the color of $(x_1, \ldots, x_k)$. One player announces her color and the rest compare it with their color. If the colors are the same, they accept. Otherwise they reject. If $\mathrm{EVAL}_G(x_1, \ldots, x_{k+1}) = 1$ then obviously all the colors are the same. If $\mathrm{EVAL}_G(x_1, \ldots, x_{k+1}) = 0$, then setting $z = -\sum_{j=1}^{n} x_i$, we have $x_i' = x_i + z$ for all $i \in \{1, \ldots, n\}$. Thus the $k+1$ points that

the players compute the colors for form a corner. By assumption, this corner is not monochromatic and the correctness of the protocols follows. The number of bits communicated is clearly as advertised.

Lower bound: Let c be the cost of an optimal $(k+1)$-party protocol for EVAL_G. We will color G^k with 2^c colors so that no corner is monochromatic. The coloring is as follows. We know the protocol partitions the input space G^{k+1} into at most 2^c cylinder intersections $C_1, C_2, \ldots, C_{2^c}$, each of which has the same value with respect to EVAL_G's output. We color a point $(x_1, \ldots, x_k)$ in G^k with the label of the cylinder intersection that contains $(x_1, \ldots, x_k, -(x_1 + \cdots + x_k))$. To show that this is indeed a legal coloring, suppose there is a corner which is monochromatic:

$$
\begin{gathered}
(x_1, x_2, \ldots, x_k), \\
(x_1 + \lambda, x_2, \ldots, x_k), \\
(x_1, x_2 + \lambda, \ldots, x_k), \\
\vdots \\
(x_1, x_2, \ldots, x_k + \lambda).
\end{gathered}
$$

These are colored respectively with the colors of

$$
\begin{gathered}
(x_1, x_2, \ldots, x_k, -(x_1 + \cdots + x_k) - \lambda + \lambda), \\
(x_1 + \lambda, x_2, \ldots, x_k, -(x_1 + \cdots + x_k) - \lambda), \\
(x_1, x_2 + \lambda, \ldots, x_k, -(x_1 + \cdots + x_k) - \lambda), \\
\vdots \\
(x_1, x_2, \ldots, x_k + \lambda, -(x_1 + \cdots + x_k) - \lambda).
\end{gathered}
$$

This is a star, contained in a cylinder intersection with value 1, and its center is $(x_1, x_2, \ldots, x_k, -(x_1 + \cdots + x_k) - \lambda)$. Hence the center must also be in the same cylinder intersection and must have the value 1. But this is not true since the sum of the coordinates is λ which is non-zero by definition.

ECCC ISSN 1433-8092
http://eccc.hpi-web.de