THEORY OF COMPUTING, Volume 14 (21), 2018, pp. 1–23
www.theoryofcomputing.org

Separation of Unbounded-Error Models in Multi-Party Communication Complexity

Arkadev Chattopadhyay* Nikhil S. Mande†

Received September 7, 2016; Revised March 24, 2017; Published December 28, 2018

Abstract: We construct a simple function that has small unbounded-error communication complexity in the k-party number-on-forehead (NOF) model but every probabilistic protocol that solves it with subexponential advantage over random guessing has cost essentially $\Omega(\sqrt{n}/4^k)$ bits. This separates these classes up to $k \leq \delta \log n$ players for any constant $\delta < 1/4$, and gives the largest known separation by an explicit function in this regime of k. Our analysis is elementary and self-contained, inspired by the methods of Goldmann, Håstad, and Razborov (Computational Complexity, 1992). After initial publication of our work as a preprint (ECCC, 2016), Sherstov pointed out to us that an alternative proof of an $\Omega((n/4^k)^{1/7})$ separation is implicit in his prior work (SICOMP, 2016). Furthermore, based on his prior work (SICOMP, 2013 and SICOMP, 2016), Sherstov gave an alternative proof of our constructive $\Omega(\sqrt{n}/4^k)$ separation and also produced a stronger non-constructive $\Omega(n/4^k)$ separation. These results are explained in Sherstov's preprint (ECCC, 2016) and in article 14/22 in *Theory of Computing*.

A preliminary version of this paper appeared as ECCC technical report TR 16-095.

*This work was done while the author was partially supported by a Ramanujan fellowship of the Department of Science and Technology, India.

†This work was done while the author was supported by a fellowship of the Department of Atomic Energy, India.

ACM Classification: F.1.3, F.2.3

AMS Classification: 68Q05, 68Q10, 68Q15, 68Q17

Key words and phrases: complexity theory, communication complexity, weakly unbounded error, unbounded error, NOF model

 DOI: 10.4086/toc.2018.v014a021

Our result has the following consequence for Boolean threshold circuits. Let THR and MAJ denote the classes of linear threshold functions that have unbounded weights and polynomially bounded weights, respectively. Further, let PAR_k (ANY_k) denote the class of functions that are parities of k bits (any k-junta, respectively). Denote by $\mathrm{THR} \circ \mathrm{PAR}_k$ the class of depth-2 circuits where the top gate computes a linear threshold function, and the bottom gates compute functions in PAR_k. For every $2 \leq k \leq \delta \log n$, we show that there exists a function computable by a linear-size $\mathrm{THR} \circ \mathrm{PAR}_k$ circuit, but requires $\exp\left(n^{\Omega(1)}\right)$-size circuits in the class $\mathrm{MAJ} \circ \mathrm{SYM} \circ \mathrm{ANY}_{k-1}$, where the gates in the middle layer compute symmetric functions. Applying a result of Goldmann et al. (loc. cit.) to the above, similar lower bounds on the size of circuits of the form $\mathrm{MAJ} \circ \mathrm{THR} \circ \mathrm{ANY}_{k-1}$ for computing the function follow.

The main technical ingredient of our proof is to exhibit a composed function of the form $f \circ \mathrm{XOR}$ which has exponentially small discrepancy while f has sign-degree just 1. An interesting aspect of our composed function is that the block size of the inner XOR function is fixed to 1, independent of k, the number of players.

1 Introduction

Over thirty years ago Chandra, Furst and Lipton [10] introduced the "number-on-forehead" (NOF) model of multi-party communication to obtain lower bounds on the size of branching programs. In the NOF model, there are k players each having an input that is metaphorically held on their foreheads. Every forehead is visible to a player except her own. The two features that make this model much more subtle than its classical two-party counterpart, are the mutual overlap of information and the fact that as k grows, each player misses less information. Indeed, starting with the surprising result of Grolmusz [18], several sets of authors (see, for example, [3, 1, 15]) constructed counter-intuitive protocols especially when k is greater than $\log n$. This makes proving multi-party lower bounds on the cost of protocols quite challenging. However, researchers have been well motivated to take on this challenge due to many well-known applications of such lower bounds in diverse areas like circuit complexity, proof complexity, and pseudo-random generators. More recently new applications have emerged in areas like data structures [25] and distributed computing [16].

In a seminal paper, Babai, Frankl and Simon [2] introduced communication complexity classes for the 2-party model. Babai et al. [2] argue that protocols with poly-log (of input length) communication cost is a natural notion of efficient protocols, just as polynomial time is a notion of efficient computation on Turing machines. Armed with this notion, most computational complexity classes have their analogues in communication complexity. This also extends easily to the NOF model and gives rise to complexity classes $\mathsf{P}_k^{\mathsf{cc}}, \mathsf{BPP}_k^{\mathsf{cc}}, \mathsf{NP}_k^{\mathsf{cc}}, \mathsf{PP}_k^{\mathsf{cc}}$ etc. While it is very hard to separate Turing machine complexity classes, many separations in the communication world are known when $k = 2$. For instance, the Equality function easily separates $\mathsf{P}_2^{\mathsf{cc}}$ from $\mathsf{BPP}_2^{\mathsf{cc}}$. Set-Disjointness famously separates $\mathsf{BPP}_2^{\mathsf{cc}}$ from $\mathsf{PP}_2^{\mathsf{cc}}$ [2] (cf. [22, 28]). However, for $k \geq 3$, things become more delicate. While for $k \geq 3$ Beame et al. [5] separated $\mathsf{P}_k^{\mathsf{cc}}$ from $\mathsf{BPP}_k^{\mathsf{cc}}$ not too long ago, it is still outstanding to find an explicit function witnessing this separation even for $k = 3$. A recent line of work [24, 13, 12, 33, 31, 26] showed that Set-Disjointness also separates $\mathsf{BPP}_k^{\mathsf{cc}}$ and $\mathsf{PP}_k^{\mathsf{cc}}$ for $k \leq \delta \cdot \log n$ for some constant $\delta < 1$.

In this paper, we consider the class $\mathsf{PP}_k^{\mathsf{cc}}$. Babai et al. [2] realized that the Turing machine complexity class PP has two different natural versions in the communication world. Let ε be the advantage of a probabilistic protocol over random guessing. Then, one way to measure the cost of the protocol is to add a $\log(1/\varepsilon)$ term to the total number of bits communicated in the worst case. Functions that admit k-party probabilistic protocols of poly-logarithmic cost in this model form the class $\mathsf{PP}_k^{\mathsf{cc}}$. The other model is unrestricted: it does not penalize by adding the $\log(1/\varepsilon)$ term to the cost, i. e., the cost is just the total number of bits communicated in the worst case. Protocols in this model are allowed to use only *private* random coins (see Section 2.1) and must, on each input, have non-zero advantage over random guessing. Functions that have efficient k-party protocols in this model form the class $\mathsf{UPP}_k^{\mathsf{cc}}$. It is not difficult to see that $\mathsf{PP}_k^{\mathsf{cc}} \subseteq \mathsf{UPP}_k^{\mathsf{cc}}$. The fact that this inclusion is strict for $k = 2$ was shown independently by Buhrman, Vereshchagin and de Wolf [8] and by Sherstov [29]. The two papers use two different functions. However the corresponding separation question for $k \geq 3$ players remained unaddressed in the literature.

We consider a simple and natural extension of the function defined by Goldmann, Håstad and Razborov [17], which we define as follows.

Definition 1.1. Let

$$P(x,y_1,\dots,y_k) := \sum_{i=0}^{n-1}\sum_{j=0}^{n4^k-1} 2^i y_{1j}\dots y_{kj}(x_{i,2j}+x_{i,2j+1})$$

where $x \in \{\pm 1\}^{2n^2 4^k}, y_i \in \{\pm 1\}^{n4^k}$ for each i.
We set

$$\mathrm{GHR}_k^N(x,y_1,\dots,y_k) := \mathrm{sgn}(2P(x,y_1,\dots,y_k)+1)\,,$$

where $N = 2n^2 4^k$.

Our main theorem in this article uses GHR_k^N to separate $\mathsf{PP}_k^{\mathsf{cc}}$ from $\mathsf{UPP}_k^{\mathsf{cc}}$ for $k \leq \delta \log N$, for any constant $\delta < 1/4$. Note that there is a natural way to assign the input variables to GHR_k^N to $k+1$ players as follows: x is Player 1's input, and y_i is Player $(i+1)$'s input (for $i = 1,\dots,k$). Our main theorem is given below.

Theorem 1.2 (Main Theorem)**.** *Let Π be any $(k+1)$-party probabilistic public-coin protocol computing the GHR_k^N function with advantage $\varepsilon > 0$. Then,*

$$\mathrm{Cost}(\Pi) + \log(1/\varepsilon) \geq \Omega\left(\frac{\sqrt{N}}{4^k} - \log N - k\right).$$

Observe that Theorem 1.2 gives a lower bound precisely for the cost of a $\mathsf{PP}_{k+1}^{\mathsf{cc}}$ protocol computing GHR_k^N. On the other hand, note that GHR_k^N is a composition of a linear threshold function with N parities of arity $k+1$. A well-known simple fact (refer to Section 3 for a proof) says that every such function has a $\mathsf{UPP}_{k+1}^{\mathsf{cc}}$ protocol of cost $O(\log N)$. This immediately yields the following separation result.

Corollary 1.3. *For all $1 \leq k \leq \delta \log n$, the GHR_k^N function is not in $\mathsf{PP}_{k+1}^{\mathsf{cc}}$ but is in the class $\mathsf{UPP}_{k+1}^{\mathsf{cc}}$, for any constant $0 < \delta < 1/4$.*

An additional motivation for our work comes from the study of constant-depth circuits with Threshold gates. There are two types of Threshold gates that have been considered in the literature. The first one is with unbounded weights and the class of such gates is denoted by THR. Formally, define a gate G to be a Threshold gate if there exist integer weights $a_0, a_1, \ldots, a_n$ such that

$$G(x_1, \ldots, x_n) = \operatorname{sgn}\left(a_0 + \sum_{i=1}^{n} a_i x_i\right).$$

The second class of gates is those Threshold gates with polynomially bounded weights, called Majority gates. We denote the class of such gates by MAJ.

Goldmann et al. [17] showed that although linear threshold functions with unbounded weights can be simulated by polynomial-size $\mathsf{MAJ} \circ \mathsf{MAJ}$ circuits, a simple function computable by linear-size circuits of the form $\mathsf{THR} \circ \mathsf{PAR}_2$ requires exponential size to be computed by $\mathsf{MAJ} \circ \mathsf{SYM}$ circuits, where SYM denotes gates computing arbitrary symmetric functions. We strengthen their result to depth-3 circuits as follows.

Theorem 1.4. *For each $k \geq 2$, the function GHR_k^N can be computed by linear-size $\mathsf{THR} \circ \mathsf{PAR}_{k+1}$ circuits, but requires size*

$$\exp\left(\Omega\left(\frac{\sqrt{N}}{k4^k} - \frac{\log N}{k}\right)\right)$$

to be computed by depth-3 circuits of the form $\mathsf{MAJ} \circ \mathsf{SYM} \circ \mathsf{ANY}_k$.

Let us remark that Theorem 1.4 continues to yield non-trivial bounds as long as $k < \delta \log n$ for any constant $0 < \delta < 1/4$. It is also worth noting that a result of [17] immediately yields, from the above theorem, the following interesting result.

Corollary 1.5. *The function GHR_k^N can be computed by linear-size $\mathsf{THR} \circ \mathsf{PAR}_{k+1}$ circuits but requires size*

$$\exp\left(\Omega\left(\frac{\sqrt{N}}{k4^k} - \frac{\log N}{k}\right)\right)$$

to be computed by depth-3 circuits of the form $\mathsf{MAJ} \circ \mathsf{THR} \circ \mathsf{ANY}_k$.

1.1 Related work

An anonymous reviewer, and subsequently Sherstov [32], pointed out that a comparatively off-the-shelf $\Omega((n/4^k)^{1/7})$ separation between $\mathsf{PP}_k^{\mathsf{cc}}$ and $\mathsf{UPP}_k^{\mathsf{cc}}$ is implicit in prior work by combining known results of Sherstov [33] and Beigel [7]. The best $\mathsf{PP}_k^{\mathsf{cc}}$ lower bound that one would get using functions obtained in this way is $\Omega((n/4^k)^{2/5})$, using a later result of Sherstov [31] and a more recent result of Thaler [35]. In contrast, our Theorem 1.2 obtains a bound of $\Omega(\sqrt{n}/4^k)$ for a function that is a natural multi-party adaptation of the function used by Goldmann et al. [17]. Our bound is stronger than the above bounds for $k \leq (1/4 - \varepsilon)\log n$ players. In particular, for any constant k, we get an $\Omega(n^{1/2})$ bound for our function whereas the best previous implicit bounds are $\Omega(n^{2/5})$. After our result was published in a technical report [14], Sherstov [34] showed that by carefully piecing together approximation-theoretic ideas from his earlier work [30] and the result in [33], one can obtain an $\Omega(n/4^k)$ lower bound for a non-explicit function. This implies, invoking a standard technique, our lower bound, for an explicit function that is similar to ours. We note that while our result separates $\mathsf{PP}_k^{\mathsf{cc}}$ from $\mathsf{UPP}_k^{\mathsf{cc}}$ for up to $k \le (1/4-\varepsilon)\log n$ players, Sherstov's separation [34] extends to $k \le (1/2-\varepsilon)\log n$ players. On the other hand, our method is elementary and self-contained. Using first principles, we prove a strong $\mathsf{PP}_k^{\mathsf{cc}}$ lower bound for a function which remained unanalyzed until this result.

The route of combining earlier work of Sherstov [33] uses unique-disjointness as the inner function. With such an inner function, the previous techniques work with any outer function, like ODD-MAX-BIT, that has large approximation error for any polynomial of degree sufficiently smaller than n. This is in contrast to our use of XOR as the inner function. It is not very difficult to see that ODD-MAX-BIT $\circ$ XOR has very efficient $\mathsf{PP}_k^{\mathsf{cc}}$ protocols for all $k \ge 2$. Thus, our argument has to exploit some feature of the outer function that is not possessed by functions like ODD-MAX-BIT. We find this an independently interesting aspect of the technique used in this article. Indeed, there has been considerable recent interest in studying the communication complexity of XOR functions (see, for example, [36, 21]).

In summary, progress on separating communication complexity classes in the NOF model has been slow. This article is the first one to explicitly address the question of separating $\mathsf{PP}_k^{\mathsf{cc}}$ and $\mathsf{UPP}_k^{\mathsf{cc}}$ for $k > 2$.

1.2 Our proof technique and organization

We work with the GHR function which is easily seen to be the composition of the *universal* threshold function [20] and Parity. The universal threshold function derives its name from the fact that by setting some of its bits appropriately one can induce any arbitrary threshold function. In that sense, it is the hardest function of sign-degree 1. To estimate the discrepancy of GHR_k^N, we extend ideas from [17] where this was estimated in the setting of two players. The basic intuition can be seen after observing that for a given setting of $y_1, \dots, y_k$ the GHR_k^N function essentially depends on the sign of a plus-minus combination of the A_j for $0 \le j \le n4^k - 1$, where

$$A_j := \frac{1}{2} \sum_{i=0}^{n-1} 2^i \left(x_{i,2j} + x_{i,2j+1}\right).$$

The relevant sign of each A_j depends on the parity of the bits $y_{1,j}, \dots, y_{k,j}$. Further, the set of bits in x that each A_j depends on is disjoint from the set of bits that $A_{j'}$ depends on, whenever $j \neq j'$. We sample x such that each A_j is an i.i.d. binomial distribution centered at 0 with range $[-2^n+1, 2^n-1]$. Let this distribution be μ_X. We sample each y_i uniformly at random. We want to ensure that the GHR_k^N function, under this distribution, behaves in a way that leaves the players with little clue about the outcome unless the relevant sign to be associated with each A_j is determined. The distribution defined above is a product distribution. Sherstov [29] showed that the GHR function has large discrepancy under product distributions. Thus, as done in [17], one is forced to sample in a slightly more involved way. First sample the y uniformly at random. Then sample x according to μ_X, conditioned on the fact that $P = \sum_{j=0}^{n4^k-1} A_j y_{1,j} \cdots y_{k,j}$ is very close to its mean compared to its standard deviation (which is as high as $\exp(\Omega(n))$). Note that the mean of each A_j is 0, which gives us plenty of room to exploit. This turns out to be the hard distribution but to establish this requires technical work. This is mainly because analyzing the discrepancy under non-product distributions is difficult. As a first step to overcome this difficulty, we follow the ideas of Goldmann et al. [17], and show that it is sufficient to prove an upper bound on the discrepancy of a function related to the GHR function under a particular product distribution. Analyzing the discrepancy of this related function on the obtained product distribution is still non-trivial, and this is the main technical contribution of our result.

Organization: Section 2 develops the basic notions and lemmas. Section 3 establishes our main technical result, Theorem 3.1, an upper bound on the k-wise discrepancy of the GHR function. Using this, we prove Theorem 1.2 and Corollary 1.3. Section 4 derives the circuit consequences of Theorem 1.4 and Corollary 1.5. Finally, Section 5 concludes with some open problems.

2 Preliminaries

2.1 The NOF model

In the k-party model of Chandra et al. [10], k players with unlimited computational power wish to compute a function $f : X_1 \times \cdots \times X_k \to \{-1,1\}$ on some input $x = (x_1,\ldots,x_k)$. For the purpose of this paper, we consider inputs of the form $X_i = \{-1,1\}^{n_i}$. On input x, player i is given $(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k)$, which is why it is figuratively said that x_i is on the i-th player's forehead. Players communicate by writing on a blackboard, so every player sees every message. We denote by $D_k(f)$ the deterministic k-party communication complexity of f, namely the number of bits exchanged in the best deterministic protocol for f on the worst-case input.

A probabilistic protocol Π with access to public (private) randomness computes f with advantage ε if the probability that Π and f agree is at least $1/2+\varepsilon$ for all inputs. The cost of Π is the maximum number of bits it communicates over its internal random choices in the worst case. Let us define $R_\varepsilon^{\mathrm{pub}}(f)$ ($R_\varepsilon^{\mathrm{priv}}(f)$) to be the cost of the best such protocol. Note that for convenience, we deviate from the notation defined in [23]. For the purpose of this paper, all logarithms are taken in base 2. We now define two other notions.

$$\mathrm{PP}_k(f) := \min_\varepsilon \left[R_\varepsilon^{\mathrm{pub}}(f) + \log\left(\frac{1}{\varepsilon}\right)\right], \qquad \mathrm{UPP}_k(f) := \min_\varepsilon \left[R_\varepsilon^{\mathrm{priv}}(f)\right]. \tag{2.1}$$

Note that privateness of the random coins is essential in the definition of UPP_k. It is a simple exercise to show that every function can be computed by unbounded-error protocols using 2 bits if allowed public coins. Define $\mathsf{PP}_k^{\mathsf{cc}} = \{f : \mathrm{PP}_k(f) = \mathrm{polylog}(n)\}$ and $\mathsf{UPP}_k^{\mathsf{cc}} = \{f : \mathrm{UPP}_k(f) = \mathrm{polylog}(n)\}$, where n is the maximum length of an input to a player. Each element in either of these classes refers to a family of functions, f, one for each input length.

2.2 Cylinder intersections, discrepancy and the cube norm

Let $f : X_1 \times \cdots \times X_k \to \{-1,1\}$. A subset $S_i \subseteq X_1 \times \cdots \times X_k$ is a cylinder in the i-th direction if membership in S_i does not depend on the i-th coordinate. A subset S is called a cylinder intersection if it can be represented as the intersection of k cylinders, $S = \bigcap_{i=1}^k S_i$, where S_i is a cylinder in the i-th direction.

Definition 2.1. Let μ be a distribution on $X_1 \times \cdots \times X_k$. The discrepancy of f according to μ, $\mathrm{Disc}^k_\mu(f)$ is

$$\max_S \left| \Pr_\mu[f(x_1,\dots,x_k) = 1 \wedge (x_1,\dots,x_k) \in S] - \Pr_\mu[f(x_1,\dots,x_k) = -1 \wedge (x_1,\dots,x_k) \in S] \right|$$

where the maximum is taken over all cylinder intersections S.

The k in Disc^k_μ denotes the dimension of the underlying cylinder intersections. We will drop this superscript when it is clear from the context what k is. Let $\mathrm{Disc}(f) = \min_\mu \mathrm{Disc}^k_\mu(f)$.

The discrepancy method is a powerful tool that gives a lower bound on the randomized communication complexity in terms of the discrepancy. The following lemma, due to Babai et al. [4], can be found, for example, in [23, Sec. 3.5].

Lemma 2.2. $R^{\mathrm{pub}}_\varepsilon(f) \geq \log(2\varepsilon/\mathrm{Disc}(f))$.

We now recall a useful technique that shows upper bounds on the discrepancy of a function under a product distribution. It is a standard lemma (see, for example, [12] and [27]).

Lemma 2.3. *Let $f : X \times Y_1 \times \cdots \times Y_k \to \mathbb{R}$, $\mu = \mu_X \times \mu_1 \times \cdots \times \mu_k$ be any product distribution, and let $\phi : X \times Y_1 \times \cdots \times Y_k \to \{0,1\}$ be the characteristic function of a cylinder intersection. Then,*

$$\left| \mathop{\mathbb{E}}_\mu [f(x,y_1,\dots,y_k)\phi(x,y_1,\dots,y_k)] \right| \leq \left(\mathop{\mathbb{E}}_{y_1^0,y_1^1,\dots,y_k^0,y_k^1} \left[\left| \mathop{\mathbb{E}}_{x\sim\mu_X} \Pi_{a\in\{0,1\}^k} f(x, y_1^{a_1}, \dots, y_k^{a_k}) \right| \right] \right)^{1/2^k}$$

where $y_i^0 \sim \mu_i$, $y_i^1 \sim \mu_i$ are sampled independently for each $i \in [k]$.

Remark 2.4. When f is $\{-1,1\}$ valued, the left hand side represents the discrepancy of f over the cylinder intersection ϕ with respect to the distribution μ. However, for our purposes, we are required to use the inequality when f is $\{-1,1,0\}$ valued.

2.3 The binomial distribution

Definition 2.5. Let $B(N)$ denote the distribution obtained as the sum of $2N$ independent Bernoulli variables, each of which take values $1/2, -1/2$ with probability $1/2$ each.

A few important things to observe are that $B(N)$ takes only integral values, it is centered and symmetric around 0, so $B(N)$ is identically distributed to $-B(N)$. Its range is $[-N,N]$.
Let us denote $\Pr[B(N)=0]$ by p_0. It is a well-known fact that

$$p_0 = \binom{2N}{N} / 4^N = \Theta\left(\frac{1}{N^{1/2}}\right).$$

The following lemma tells us that the probability of a binomial distribution taking any value close to its mean is significantly high.

Lemma 2.6. *Let W be a binomial random variable distributed according to $B(N)$. Let p_0 denote* $\Pr[W=0]$. *Then for all $j \in [-N,N]$,*

$$p_0 - O\left(\frac{j^2}{N^{3/2}}\right) \le \Pr[W=j] \le p_0 \,.$$

Proof. Note that for $|j| \ge N/2$, the bound to be proved is trivial. Thus we assume $|j| < N/2$.

$$\begin{aligned}
\Pr[W=j-1]-\Pr[W=j] &= \left[\binom{2N}{N+j-1}-\binom{2N}{N+j}\right]\cdot\frac{1}{2^{2N}} \\
&= \left[\frac{(2N)!}{(N+j-1)!(N-j+1)!}-\frac{(2N)!}{(N+j)!(N-j)!}\right]\cdot\frac{1}{2^{2N}} \\
&= \frac{(2N)!}{(N+j-1)!(N-j)!}\cdot\frac{2j-1}{(N-j+1)(N+j)}\cdot\frac{1}{2^{2N}} \\
&= \binom{2N}{N+j}\cdot\frac{2j-1}{N-j+1}\cdot\frac{1}{2^{2N}} \\
&\le \binom{2N}{N}\cdot\frac{1}{2^{2N}}\cdot\frac{2j}{N-j}
\end{aligned}$$

since the middle binomial coefficient is the maximum. Thus, we have $\forall i, |i| \le j$,

$$\Pr[W=i-1]-\Pr[W=i] \le \binom{2N}{N}\frac{2j}{N-j}\cdot\frac{1}{2^{2N}} \,.$$

Since $\binom{2N}{N}/4^N = \Theta\left(\frac{1}{N^{1/2}}\right)$,

$$\begin{aligned}
\Pr[W=0]-\Pr[W=j] \le \sum_{i=1}^{j}|\Pr[W=i-1]-\Pr[W=i]| &\le \frac{2j^2}{N-j}\cdot O\left(\frac{1}{N^{1/2}}\right) \\
&\le \frac{2\cdot 2j^2}{N}\cdot O\left(\frac{1}{N^{1/2}}\right) && \text{(since } |j| \le N/2\text{)} \\
&= O\left(\frac{j^2}{N^{3/2}}\right) \,.
\end{aligned}$$

□

3 A discrepancy upper bound for the multiparty GHR function

In this section, we prove essentially an $\exp\left(-\sqrt{N}/4^k\right)$ upper bound on the discrepancy of the GHR_k^N function where the first player gets N input bits. This gives us an $\exp\left(-n^{\Omega(1)}\right)$ upper bound on the discrepancy if $k \le \varepsilon \log(N)$ for any constant $0 < \varepsilon < 1/4$.

Goldmann et al. [17] showed that when $k=2$, if there is a low cost one-way protocol for GHR_2^N, then it must have low advantage. Sherstov [29] noted that the same proof technique can be adapted to prove an upper bound on the discrepancy on GHR_2^N. We generalize this for higher k. In particular, we show

Theorem 3.1. *For any $k \geq 1$,*

$$\mathrm{Disc}(\mathrm{GHR}_k^N) = O\left(\frac{(8\mathrm{e})^k N^{1/4}}{2^{\sqrt{N}/4^k} \cdot 2^{k/2}}\right),$$

where GHR_k^N is defined as in Definition 1.1, and N is the maximum number of bits a player gets (in this case the first player).

Proof of Theorem 1.2. It follows directly from Theorem 3.1 and Lemma 2.2. □

Proof of Corollary 1.3. From Theorem 1.2, it follows that for all $1 \leq k \leq \delta \cdot \log n$, the GHR_k^N function is not in $\mathsf{PP}_{k+1}^{\mathsf{cc}}$ for any constant $0 < \delta < 1/4$. Let us see an easy unbounded error protocol for GHR_k^N. Note that all the weights of the top threshold are positive. One player chooses and announces a Parity gate at the bottom layer with probability proportional to its corresponding weight. The cost of announcing this is $O(\log(N))$. The probability of success equals $\sum w_i^+/w$, where the w_i^+ are the weights of the gates which agree with the output. Since $\sum w_i^+ > \sum w_i^-$ (the weights of the gates which disagree with the output), the probability of success is strictly greater than $1/2$. □

Recall that $N = 2n^2 4^k$. The proof technique of Theorem 3.1 is inspired from that of Goldmann et al. [17].

Proof of Theorem. 3.1. Let

$$A_j = \frac{1}{2}\sum_{i=0}^{n-1} 2^i (x_{i,2j} + x_{i,2j+1}).$$

It is easy to see that A_j can take any integer value in $[-2^n+1, 2^n-1]$. Let μ_X be a distribution on the variables x that make the variables A_j independent and binomially distributed according to $B(2^n-1)$ as defined in Definition 2.5. Such a distribution exists because each A_j depends on a disjoint set of variables. Let $\mathcal{U}$ be the uniform distribution on $\{-1,1\}^{n4^k}$. We choose a tuple $(x, y_1, \ldots, y_k)$ by first picking $y_i \sim \mathcal{U}$ independently for each i, and then picking $x \sim \mu_X$ under the condition that $|P(x, y_1, \ldots, y_k)| = 2^k$. Let us define μ to be the distribution obtained by this sampling procedure.

We will now show an upper bound on the discrepancy of GHR_k^N under the distribution μ. Let S denote the characteristic function (0-1 valued) of a cylinder intersection. By Definition 2.1, the discrepancy of GHR_k^N according to μ is

$$\mathrm{Disc}_\mu(\mathrm{GHR}_k^N) = \max_S \left| \mathop{\mathbb{E}}_{\mu} \left[\mathrm{GHR}_k^N(x, y_1, \ldots, y_k) S(x, y_1, \ldots, y_k)\right]\right|. \tag{3.1}$$

The following lemma will enable us to switch to working with a product distribution on the inputs, for which we have convenient techniques for proving discrepancy upper bounds via Lemma 2.3.

Lemma 3.2. *For $\mu_X, \mathcal{U}$ as defined above,*

$$\Pr_{\mu_X \times \mathcal{U}^k}[|P(x, y_1, \ldots, y_k)| = 2^k] \geq \Omega\left(\frac{1}{\sqrt{n}2^{(n+2k)/2}}\right).$$

Proof. We will show that for any fixed $y_1, \dots, y_k$, if we sample x according to μ_X, then

$$P(x, y_1, \dots, y_k)/2 = \sum_{j=0}^{n4^k-1} A_j y_{1j} \cdots y_{kj}$$

is distributed according to $B(n4^k(2^n-1))$. Note that $A_j y_{1j} \cdots y_{kj}$ is always distributed according to $B(2^n-1)$, no matter what the values of $y_1, \dots, y_k$ are. Next, observe that the sum of binomial distributions is a binomial distribution. This shows that

$$\sum_{j=0}^{n4^k-1} A_j y_{1j} \cdots y_{kj}$$

is distributed according to $B(n4^k(2^n-1))$.

Hence, by plugging in $N = n4^k(2^n-1)$ and $j = 2^k$ in Lemma 2.6,

$$\begin{aligned}\Pr_{\mu_X \times \mathcal{U}^k}[|P(x, y_1, \dots, y_k)| = 2^k] &\geq \Omega\left(\frac{1}{(n4^k(2^n-1))^{1/2}}\right) - O\left(\frac{4^k}{(n4^k(2^n-1))^{3/2}}\right) \\ &= \Omega\left(\frac{1}{\sqrt{n}2^{(n+2k)/2}}\right).\end{aligned}$$

We can discard the second term since it equals

$$O\left(\left(4^{k/2} \cdot n(2^n-1)^{3/2}\right)^{-1}\right),$$

and is dominated by the first term. $\square$

Let us now recall the law of total expectation.

Fact 3.3 (Law of total expectation)**.** *For any probability space $(\Omega, \mathcal{F}, \nu)$, any event $E \in \mathcal{F}$, and any random variable Z, the following equality holds.*

$$\mathop{\mathbb{E}}_{\nu}[Z] = \mathop{\mathbb{E}}_{\nu}[Z \mid E] \cdot \Pr_{\nu}[E] + \mathop{\mathbb{E}}_{\nu}[Z \mid \bar{E}] \cdot (1 - \Pr_{\nu}[E]).$$

Define a function q by

$$q(x, y_1, \dots, y_k) = \begin{cases} P(x, y_1, \dots, y_k)/2^k & \text{if } |P(x, y_1, \dots y_k)| = 2^k, \\ 0 & \text{otherwise.} \end{cases}$$

This means that if $(x, y_1, \dots, y_k)$ is chosen according to the distribution $\mu_X \times \mathcal{U}^k$, then $q(x, y_1, \dots, y_k) = \mathrm{GHR}_k^N(x, y_1, \dots, y_k)$ on the support of μ, and 0 otherwise. For any cylinder intersection S, let Z denote the random variable $q(x, y_1, \dots, y_k) \cdot S(x, y_1, \dots, y_k)$, let E denote the event $|P(x, y_1, \dots y_k)| = 2^k$. Using Fact 3.3 and the fact that $\mathbb{E}_{\mu_X \times \mathcal{U}^k}[Z \mid \bar{E}] = 0$, we obtain

$$\mathop{\mathbb{E}}_{\mu}[\mathrm{GHR}(x, y_1, \dots, y_k) S(x, y_1, \dots, y_k)] = \frac{\mathbb{E}_{\mu_X \times \mathcal{U}^k}[q(x, y_1, \dots, y_k) \cdot S(x, y_1, \dots, y_k)]}{\Pr_{\mu_X \times \mathcal{U}^k}[|P(x, y_1, \dots y_k)| = 2^k]}. \tag{3.2}$$

Using Equation (3.1), Lemma 3.2 and Equation (3.2), we obtain the following.

$$\mathrm{Disc}_\mu(\mathrm{GHR}_k^N) \leq \max_S \left| \mathop{\mathbb{E}}_{\mu_X \times \mathcal{U}^k} [q(x, y_1, \ldots, y_k) S(x, y_1, \ldots, y_k)] \right| \cdot O\left(\sqrt{n} 2^{\frac{n+2k}{2}}\right) \tag{3.3}$$

where S denotes a cylinder intersection. It therefore suffices to show that for all cylinder intersections S,

$$\left| \mathop{\mathbb{E}}_{\mu_X \times \mathcal{U}^k} [q(x, y_1, \ldots, y_k) S(x, y_1, \ldots, y_k)] \right| \leq O\left(2^{-\frac{n+2k}{2} - \varepsilon n}\right) \tag{3.4}$$

for some constant $\varepsilon > 0$ to give us a discrepancy upper bound of $\exp\left(-n^{\Omega(1)}\right)$. For notational convenience, we may switch between the notations $\mathbb{E}_x$ and $\mathbb{E}_{x \sim \mu_X}$ from now on. Now that we have a product distribution, we can use Lemma 2.3.

$$\left| \mathop{\mathbb{E}}_{\mu_X \times \mathcal{U}^k} [q(x, y_1, \ldots, y_k) S(x, y_1, \ldots, y_k)] \right| \leq \left(\mathop{\mathbb{E}}_{y_1^0, y_1^1, \ldots, y_k^0, y_k^1} \left| \mathop{\mathbb{E}}_x \left[\prod_{a_1, \ldots, a_k \in \{0,1\}} q(x, y_1^{a_1}, \ldots, y_k^{a_k}) \right] \right| \right)^{1/2^k} . \tag{3.5}$$

We will now show an upper bound on the RHS of the above equation by splitting the outer expectation into two terms, the first of which has low probability. We will require certain properties of Hadamard matrices to prove an upper bound on the second term. Let $\beta \in \{0,1\}^k$. Define 2^k subsets of indices as

$$I_\beta = \{ j \in [n4^k] : \forall i \in [k],\ (y_i^0)_j = (-1)^{\beta_i} \cdot (y_i^1)_j \} .$$

Note that $\{I_\beta : \beta \in \{0,1\}^k\}$ forms a partition of the indices. Since our distributions on the pairs y_i^0, y_i^1 are uniform and independent, each I_β is empty with equal probability. An easy counting argument tells us that the probability of I_β being empty is

$$\left((2^k - 1)/2^k \right)^{n4^k} .$$

By a union bound, the probability that any one of them is empty is at most

$$2^k \cdot \left((2^k - 1)/2^k \right)^{n4^k} .$$

We have the following.

$$\left(\mathop{\mathbb{E}}_{y_1^0, y_1^1, \ldots, y_k^0, y_k^1} \left| \mathop{\mathbb{E}}_x \left[\prod_{a_1, \ldots, a_k \in \{0,1\}} q(x, y_1^{a_1}, \ldots, y_k^{a_k}) \right] \right| \right)^{1/2^k} \leq \left(2^k \left(1 - \frac{1}{2^k}\right)^{n4^k} + Z \right)^{1/2^k}$$

where

$$Z = \mathop{\mathbb{E}}_{y_1^0, y_1^1, \ldots, y_k^0, y_k^1 : \forall \beta, I_\beta \neq \emptyset} \left| \mathop{\mathbb{E}}_x \prod_{a_1, \ldots, a_k \in \{0,1\}} q(x, y_1^{a_1}, \ldots, y_k^{a_k}) \right| .$$

Claim 3.4. *For all $y_1^0, \dots, y_k^0, y_1^1, \dots, y_k^1$ such that I_β is non-empty for each $\beta \in \{0,1\}^k$, we have*

$$\left| \mathop{\mathbb{E}}_x \left[\prod_{a_1,\dots,a_k \in \{0,1\}} q(x, y_1^{a_1}, \dots, y_k^{a_k}) \right] \right| \le O\left(2^{k \log(\mathrm{e}) 2^k} \cdot 2^{2^k} \frac{1}{(2^{n/2})^{2^k - 1}} \cdot \frac{2^{(k+1)2^{k+1}}}{2^{3n/2}} \right).$$

Let us assume the claim to be true for now. We have from Equation (3.3) that

$$\begin{aligned}
\mathrm{Disc}_\mu(\mathrm{GHR}_k^N) &\le \left| \mathop{\mathbb{E}}_{\mu_X \times \mathcal{U}^k} [q(x, y_1, \dots, y_k) S(x, y_1, \dots, y_k)] \right| O\left(\sqrt{n} 2^{\frac{n+2k}{2}}\right) \\
&\le \left(2^k \left(1 - \frac{1}{2^k}\right)^{n4^k} + O\left(2^{k \log(\mathrm{e}) 2^k} \cdot 2^{2^k} \frac{1}{(2^{n/2})^{2^k - 1}} \cdot \frac{2^{(k+1)2^{k+1}}}{2^{3n/2}} \right) \right)^{1/2^k} \\
&\quad \cdot O\left(\sqrt{n} 2^{\frac{n+2k}{2}}\right) \\
&\le \left[2^{k/2^k} \left(1 - \frac{1}{2^k}\right)^{n2^k} + O\left(\frac{(4\mathrm{e})^k}{(2^{\frac{n}{2}})^{1 - \frac{1}{2^k}} \cdot 2^{\frac{3n}{2} \cdot \frac{1}{2^k}}} \right) \right] O\left(\sqrt{n} 2^{\frac{n+2k}{2}}\right) \\
&\le O\left(\left(\mathrm{e}^{-1/2^k}\right)^{n2^k} \cdot 2^{n/2 + k + k/2^k} \cdot \sqrt{n} + \frac{(8\mathrm{e})^k \sqrt{n}}{2^{(\frac{3n}{2} - \frac{n}{2}) \cdot \frac{1}{2^k}}} \right) \\
&\qquad\qquad\qquad\qquad \text{Using the fact that } \left(1 - \tfrac{1}{\gamma}\right) < \mathrm{e}^{-1/\gamma} \\
&= O\left(\mathrm{e}^{-n} \cdot 2^{n/2 + k + k/2^k} \cdot \sqrt{n} + \frac{(8\mathrm{e})^k \sqrt{n}}{2^{n/2^k}} \right) = O\left(\frac{(8\mathrm{e})^k \sqrt{n}}{2^{n/2^k}} \right) \qquad \text{Assuming } k < n/3 \\
&= O\left(\frac{(8\mathrm{e})^k N^{1/4}}{2^{\sqrt{N}/4^k} \cdot 2^{k/2}} \right) \qquad\qquad \text{Recall that } N = 2n^2 4^k
\end{aligned}$$

which proves Theorem 3.1. □

Now it only remains to prove Claim 3.4.

3.1 Proof of Claim 3.4

Recall that we need to show the following. For all $y_1^0, \dots, y_k^0, y_1^1, \dots, y_k^1$ such that I_β is non-empty for each β, we want

$$\left| \mathop{\mathbb{E}}_x \left[\prod_{a_1,\dots,a_k \in \{0,1\}} q(x, y_1^{a_1}, \dots, y_k^{a_k}) \right] \right| \le O\left(2^{k \log(\mathrm{e}) 2^k} \cdot 2^{2^k} \frac{1}{(2^{n/2})^{2^k - 1}} \cdot \frac{2^{(k+1)2^{k+1}}}{2^{3n/2}} \right).$$

Fix any such $y_1^0, \dots, y_k^0, y_1^1, \dots, y_k^1$. Note that the LHS of the above equation is

$$\left| \Pr_x \left[\prod_{a_1,\dots,a_k \in \{0,1\}} q(x, y_1^{a_1}, \dots, y_k^{a_k}) = 1 \right] - \Pr_x \left[\prod_{a_1,\dots,a_k \in \{0,1\}} q(x, y_1^{a_1}, \dots, y_k^{a_k}) = -1 \right] \right|.$$

For convenience, for all $a \in \{0,1\}^k$ let us denote $P(x, y_1^{a_1}, \ldots, y_k^{a_k})$ by P_a and let S_a denote $P_a/2$. By the definition of q, we have

$$\left|\mathbb{E}_x\left[\prod_{a\in\{0,1\}^k} q(x, y_1^{a_1}, \ldots, y_k^{a_k})\right]\right| = \left|\Pr_x\left[\prod_{a\in\{0,1\}^k} \frac{P_a}{2^k} = 1\right] - \Pr_x\left[\prod_{a\in\{0,1\}^k} \frac{P_a}{2^k} = -1\right]\right|$$
$$= \left|\Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = 2^{(k-1)2^k}\right] - \Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}\right]\right|. \quad (3.6)$$

Let

$$W_\beta = \sum_{j\in I_\beta} A_j (y_1^0)_j \ldots (y_k^0)_j .$$

It will be useful to note here that W_β only takes integral values. We will use this fact crucially later. Let $\mathbf{p_k}$ denote the $2^k \times 1$ column vector whose elements are indexed by $a = (a_1, \ldots, a_k) \in \{0,1\}^k$, and the a-th element of $\mathbf{p_k}$ is $P(x, y_1^{a_1}, \ldots, y_k^{a_k})$. Similarly define column vectors $\mathbf{s_k}$ ($\mathbf{w_k}$, respectively) whose a-th entries are S_a (W_a, respectively) for all $a \in \{0,1\}^k$. Although $\mathbf{p_k}, \mathbf{s_k}$ and $\mathbf{w_k}$ depend on $x, y_1^0, \ldots, y_k^0, y_1^1, \ldots, y_k^1$, we do not make this dependence explicit in the following discussion in order to avoid clutter.

Claim 3.5. *The following holds true for all k, and all $x, y_1^0, \ldots, y_k^0, y_1^1, \ldots, y_k^1$.*

$$\mathbf{p_k} = 2\mathbf{s_k} = 2\mathbf{H_k} \cdot \mathbf{w_k} \quad (3.7)$$

where $\mathbf{H_k}$ is a $2^k \times 2^k$ Hadamard matrix defined[1] *as* $\mathbf{H_k} = \begin{bmatrix} \mathbf{H_{k-1}} & \mathbf{H_{k-1}} \\ \mathbf{H_{k-1}} & -\mathbf{H_{k-1}} \end{bmatrix}$ *and* $\mathbf{H_0} = [1]$.

Let us first state a well-known property of $\mathbf{H_k}$.

Fact 3.6. *Let $\mathbf{H_k}$ be as defined above. Then, $(\mathbf{H_k})_{ij} = (-1)^{\langle i,j \rangle}$ for all $i, j \in \{0,1\}^k$.*

In other words, $\mathbf{H_k}$ is the communication matrix of the inner product (modulo 2) function. Let us now prove Claim 3.5.

Proof of Claim 3.5. Let

$$a \in \{0,1\}^k, \qquad P_a = 2\sum_{j=1}^{n4^k} A_j (y_1^{a_1})_j \cdots (y_k^{a_k})_j, \qquad \text{and} \qquad W_\beta = \sum_{j\in I_\beta} A_j (y_1^0)_j \cdots (y_k^0)_j .$$

Say $j \in I_\beta$ where $\beta \in \{0,1\}^k$. Note that $(y_i^{a_i})_j = -1 \cdot (y_i^0)_j$ iff $a_i = 1, \beta_i = 1$. Hence, we have

$$(y_1^{a_1})_j \cdots (y_k^{a_k})_j = (-1)^{(\sum_i a_i \cdot \beta_i)} (y_1^0)_j \cdots (y_k^0)_j = (-1)^{\langle a,\beta \rangle} (y_1^0)_j \cdots (y_k^0)_j .$$

We conclude that

$$P_a = 2\sum_{j=1}^{n4^k} A_j (y_1^{a_1})_j \cdots (y_k^{a_k})_j = 2\left(\sum_{\beta\in\{0,1\}^k} \sum_{j\in I_\beta} (-1)^{\langle a,\beta \rangle} A_j (y_1^0)_j \cdots (y_k^0)_j\right)$$

[1] This is the Sylvester construction of Hadamard matrices.

$$= 2\left(\sum_{\beta\in\{0,1\}^k}(-1)^{\langle a,\beta\rangle}W_\beta\right)$$
$$= 2(\mathbf{H_k})_a\cdot\mathbf{w_k}$$

where $(\mathbf{H_k})_a$ denotes the a-th row of $\mathbf{H_k}$. Thus, $\mathbf{p_k} = 2\mathbf{s_k} = 2\mathbf{H_k}\cdot\mathbf{w_k}$. □

3.1.1 On integral solutions to Hadamard constraints

In the remainder of this section, we shall refer to an integral assignment to $\mathbf{w_k}$ as a *valid integral assignment* if it satisfies Equation (3.7) for some setting of $x, y_1^0, \ldots, y_k^0, y_1^1, \ldots, y_k^1$. The conditions on $\mathbf{s_k}$ will be explicitly stated in each usage.

First, we prove that the number of valid integral assignments to $\mathbf{w_k}$ satisfying

$$\prod_{a\in\{0,1\}^k} S_a = 2^{(k-1)2^k}$$

is equal to the number of valid integral assignments to $\mathbf{w_k}$ satisfying

$$\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}.$$

Moreover, we show that the total number of such valid integral assignments is small, and the values of $|W_a|$ are not too large in any such valid assignment. Recall from Equation (3.6) that for all $y_1^0, \ldots, y_k^0, y_1^1, \ldots, y_k^1$ such that I_β is non-empty for each β, we have

$$\left|\mathop{\mathbb{E}}_x\left[\prod_{a\in\{0,1\}^k} q(x, y_1^{a_1}, \ldots, y_k^{a_k})\right]\right| = \left|\Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = 2^{(k-1)2^k}\right] - \Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}\right]\right|.$$

Thus, we can pair the valid "positive" and "negative" assignments. Higher-order terms in the difference of probabilities

$$\left|\Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = 2^{(k-1)2^k}\right] - \Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}\right]\right|$$

cancel out. We require the following well-known property of Hadamard matrices.

Fact 3.7. *Let* $\mathbf{H}$ *be an* $N\times N$ *Hadamard matrix. Then,* $\mathbf{H}$ *is invertible, and* $\mathbf{H}^{-1} = \frac{1}{N}\mathbf{H}$.

Claim 3.8. *The number of valid integral assignments to* $\mathbf{w_k}$ *such that*

$$\prod_{a\in\{0,1\}^k} S_a = +2^{(k-1)2^k}$$

equals the number of valid integral assignments such that

$$\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}.$$

Proof. The constraints we have are $\mathbf{H_k} \cdot \mathbf{w_k} = \mathbf{s_k}$. Since W_a is integral for all a, and $\mathbf{H_k}$ is a ± 1 matrix, this implies that the S_a are integral as well. Thus, using Fact 3.7 we get $(1/2^k)\mathbf{H_k} \cdot \mathbf{s_k} = \mathbf{w_k}$, or $\mathbf{H_k} \cdot (\mathbf{s_k}/2^k) = \mathbf{w_k}$. Let us consider two cases, one where $\forall a \in \{0,1\}^k, \left|S_a/2^k\right| = 1/2$, and another where there exists an a such that $\left|S_a/2^k\right| \neq 1/2$.

- Let us assume $\forall a, \left|S_a/2^k\right| = 1/2$. We show something slightly stronger, namely that every setting of each $S_a/2^k$ to $\pm 1/2$ gives us a valid assignment to the W_a. Since $\mathbf{H_k}$ is a ± 1 matrix of even dimension, the parity of the number of appearances of $+1/2$ equals the parity of number of appearances of $-1/2$ in the sum $(\mathbf{H_k})_R \cdot (\mathbf{s_k}/2^k)$, where $(\mathbf{H_k})_R$ is the R-th row of $\mathbf{H_k}$. This holds for every row R. Thus, W_R is always an integer. This means the number of valid positive assignments equals the number of valid negative assignments in this case.

- The absolute value of S_a must equal a power of 2 for each a since the product of them is a power of 2. If there exists an S_a whose value is not $\pm 2^{k-1}$, then there must exist an S_b (consider the last such one) which is a multiple of 2^k since
$$\prod_{a \in \{0,1\}^k} S_a = \pm 2^{(k-1)2^k}.$$
Since $S_b/2^k$ is an integer, and we had a valid integral assignment to $\mathbf{w_k}$, flipping the sign of S_b can change the value of any W_c to $W_c \pm 2 \cdot S_b/2^k$, which remains an integer. This is a bijection between valid positive and negative assignments. $\square$

Lemma 3.9. *The number of valid integral assignments to* $\mathbf{w_k}$ *satisfying*
$$\prod_{a \in \{0,1\}^k} S_a = \pm 2^{(k-1)2^k}$$
is at most $2^{k \log(\mathrm{e}) 2^k}$.

We use the following standard fact about binomial coefficients.

Fact 3.10. *For all* $n \in \mathbb{N}$ *and for all* $k \in [n]$, $\binom{n}{k} \leq (n \cdot \mathrm{e}/k)^k$.

Proof of Lemma 3.9. Suppose $\prod_{a \in \{0,1\}^k} S_a = \pm 2^{(k-1)2^k}$. This means we have to distribute $(k-1)2^k$ powers of 2 among among the integers $2^k S_a$. The total number of ways to do this equals the number of non-negative integer solutions to $m_1 + \cdots + m_{2^k} = (k-1)2^k$, which equals
$$\binom{k2^k - 1}{(k-1)2^k}.$$
Note that
$$\binom{k2^k - 1}{(k-1)2^k} = \binom{k2^k}{(k-1)2^k} \leq \left(k2^k \cdot \mathrm{e}/((k-1)2^k)\right)^{(k-1)2^k},$$
where the last inequality follows by Fact 3.10. Now we use the fact that $1 + x \leq \mathrm{e}^x$ and conclude that
$$\left(k2^k \cdot \mathrm{e}/((k-1)2^k)\right)^{(k-1)2^k}$$

is bounded above by e^{k2^k}, which equals $2^{k\log(\mathrm{e})2^k}$. Each of these can give at most one integral assignment to $\mathbf{w_k}$ because the system of constraints is linearly independent. ☐

We now state an upper bound on the value of $|W_a|$ in every integral assignment.

Lemma 3.11. *For all* $a \in \{0,1\}^k, |W_a| \leq 2^{(k+1)2^k}$ *for any valid integral assignment to* $\mathbf{w_k}$ *satisfying* $\prod_{a\in\{0,1\}^k} S_a = \pm 2^{(k-1)2^k}$.

Proof. First note that for each a, $|W_a| \leq \sum_{a\in\{0,1\}^k} |S_a|/2^k$ since $\mathbf{H_k} \cdot \mathbf{s_k} = \mathbf{w_k}$. We show that $\sum_{a\in\{0,1\}^k} |S_a|$ is at most 2^{k2^k}. Suppose not. By a simple averaging argument, there must be a b such that

$$|S_b| > 2^{k2^k}/2^k,$$

which is $2^{k(2^k-1)}$, which is at least $2^{(k-1)2^k}$ if $k \geq 1$. But this is not possible since

$$\prod_{a\in\{0,1\}^k} S_a = \pm 2^{(k-1)2^k}$$

and each S_a is an integer. ☐

3.1.2 Using properties of the binomial distribution

Recall from Equation (3.6) that for all $y_1^0,\dots,y_k^0,y_1^1,\dots,y_k^1$ such that I_β is non-empty for each β, we want to show an upper bound on

$$\left|\Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = 2^{(k-1)2^k}\right] - \Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}\right]\right|.$$

Recall that we defined

$$W_\beta = \sum_{j\in I_\beta} A_j (y_1^0)_j \dots (y_k^0)_j\,.$$

For any $\beta \in \{0,1\}^k$, note that W_β is always distributed according to $B(c_\beta(2^n-1))$, where $c_\beta = \left|I_\beta\right| \geq 1$. We can prove this in a manner similar to that in the proof of Lemma 3.2. In Claim 3.8, we showed that the number of valid integral assignments to $\mathbf{w_k}$ such that

$$\prod_{a\in\{0,1\}^k} S_a = 2^{(k-1)2^k}$$

equals the number of integral assignments such that

$$\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}\,.$$

Note that if the assignment to $\mathbf{w_k}$ is not integral, then it has probability 0, since for each a, W_a takes only integral values. Let us call an assignment to $\mathbf{w_k}$ *positive* if the corresponding value of

$$\prod_{a\in\{0,1\}^k} S_a = 2^{(k-1)2^k}\,,$$

and *negative* if the value of

$$\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}\,.$$

Arbitrarily form a matching, denoted by $\mathcal{M}$, between the positive and negative assignments. We will bound the difference of probabilities of each match.

$$\begin{aligned}&\left|\Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = 2^{(k-1)2^k}\right] - \Pr_x\left[\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}\right]\right|\\ &\leq \sum_{(w,w')\in\mathcal{M}} \left|\Pr_x[\mathbf{w_k}=w] - \Pr_x[\mathbf{w_k}=w']\right|\end{aligned}$$

where $w=(w_a)_{a\in\{0,1\}^k}$ is a valid positive assignment and $w'=(w'_a)_{a\in\{0,1\}^k}$ is the valid negative assignment that is the unique match of w according to $\mathcal{M}$. The term $\Pr_x[\mathbf{w_k}=w]$ is equal to

$$\Pr_x\left[\bigwedge_{a\in\{0,1\}^k} W_a = w_a\right]\,.$$

In Lemma 3.11 we showed that for each β, the absolute value of W_β in any integral assignment can be at most $2^{(k+1)2^k}$. Each W_β is distributed according to $B(c_\beta(2^n-1)), c_\beta>0$, since $\forall\beta\in\{0,1\}^k, |I_\beta|>0$. For a particular positive assignment w, negative assignment w' and any $y_1^0,\dots,y_k^0,y_1^1,\dots,y_k^1$ such that I_β is non-empty for each β,

$$\left|\Pr_x[\mathbf{w_k}=w] - \Pr_x\left[\mathbf{w_k}=w'\right]\right| = \left|\Pr\left[\bigwedge_{a\in\{0,1\}^k} W_a = w_a\right] - \Pr\left[\bigwedge_{a\in\{0,1\}^k} W_a = w'_a\right]\right|$$

By plugging in $N=c_\beta(2^n-1)$ and $j=2^{(k+1)2^k}$ in Lemma 2.6, we obtain

$$p_0 \geq \Pr_x[W_\beta = w_\beta] \geq p_0 - O\left(2^{(k+1)2^{k+1}}/2^{3n/2}\right)\,,$$

where $p_0=\Pr[W_\beta=0]=O\left(1/2^{n/2}\right)$. For convenience in calculations, let us say

$$\Pr_x[W_\beta = w_\beta] \in \left(p_0 \pm O\left(2^{(k+1)2^{k+1}}/2^{3n/2}\right)\right)\,.$$

Recall that the variables W_β are independent since they depend on disjoint sets of variables. Thus,

$$\begin{aligned}&\left|\Pr[\textstyle\bigwedge_{a\in\{0,1\}^k} W_a = w_a] - \Pr[\bigwedge_{a\in\{0,1\}^k} W_a = w'_a]\right|\\ &\leq \left|\left(p_0 \pm O\left(\frac{2^{(k+1)2^{k+1}}}{2^{3n/2}}\right)\right)^{2^k} - \left(p_0 \pm O\left(\frac{2^{(k+1)2^{k+1}}}{2^{3n/2}}\right)\right)^{2^k}\right| \leq \frac{2\cdot 2^{2^k}}{(2^{n/2})^{2^k-1}}\cdot\frac{2^{(k+1)2^{k+1}}}{2^{3n/2}}\,.\end{aligned}$$

The last inequality holds because the highest-order term after binomially expanding both terms is $(p_0)^{2^k}$, which cancel each other. Note that the sum of the binomial coefficients is 2^{2^k}, and each term after the first is at most

$$\left(1/(2^{n/2})^{2^k-1}\right) \cdot \left(2^{(k+1)2^{k+1}}/2^{3n/2}\right).$$

Thus, the sum of the remaining terms can be bounded above by

$$2^{2^k} \cdot \left(1/(2^{n/2})^{2^k-1}\right) \cdot \left(2^{(k+1)2^{k+1}}/2^{3n/2}\right).$$

By Lemma 3.9, the number of assignments is at most $2^{k\log(\mathrm{e})2^k}$. Thus,

$$\begin{aligned}&\left|\Pr_x\left[\textstyle\prod_{a\in\{0,1\}^k} S_a = 2^{(k-1)2^k}\right] - \Pr_x\left[\textstyle\prod_{a\in\{0,1\}^k} S_a = -2^{(k-1)2^k}\right]\right| \\ &\le \sum_{(w,w')\in\mathcal{M}} \left|\Pr_x[\mathbf{w_k}=w] - \Pr_x[\mathbf{w_k}=w']\right| \le 2^{k\log(\mathrm{e})2^k} \cdot 2^{2^k} \frac{1}{(2^{n/2})^{2^k-1}} \cdot \frac{2^{(k+1)2^{k+1}}}{2^{3n/2}}\end{aligned}$$

which proves Claim 3.4. Using Equation (3.3), this proves Theorem 3.1.

4 Circuit lower bounds

In this section, we will show how we obtain lower bounds on the size of depth-3 circuits of the type $\mathsf{MAJ} \circ \mathsf{THR} \circ \mathsf{ANY}_k$ computing the GHR_k^N function. Recall that GHR_k^N can be computed by linear-size $\mathsf{THR} \circ \mathsf{PAR}_{k+1}$ circuits. First let us state the results that were known prior to this work.

Lemma 4.1 (Folklore)**.** *Any function f computable by size s circuits of the type $\mathsf{SYM} \circ \mathsf{ANY}_k$ has a deterministic simultaneous $(k+1)$-player protocol of cost $O(k\log(s))$ for any partitioning of the input bits.*

Proof. Since each of the bottom-layer gates has fan-in at most k, there must exist a player who sees all the inputs to it. The protocol decides beforehand which gate "belongs" to which player. All players simultaneously broadcast their contribution to the top SYM gate using at most $\log(s)$ bits each. □

A consequence of this is an upper bound for randomized protocols for depth-3 circuits, which may be found in [11], for example, and is stated below without proof.

Lemma 4.2 (Folklore)**.** *Given any function f computable by size s circuits of the type $\mathsf{MAJ} \circ \mathsf{SYM} \circ \mathsf{ANY}_k$, and any partition of the input bits, there exists a public coin $(k+1)$-player randomized protocol computing f with advantage $\Omega(1/s)$ and cost $O(k\log(s))$.*

Let us now prove Theorem 1.4.

Proof. Suppose GHR_k^N could be computed by $\mathsf{MAJ} \circ \mathsf{SYM} \circ \mathsf{ANY}_k$ circuits of size s. Using the protocol mentioned in Lemma 4.2, the cost of the protocol is $O(k\log(s))$ and advantage $\Omega(1/s)$. Using Theorem 1.2,

$$O(k\log(s) + \log(s)) \ge \Omega\left(\frac{\sqrt{N}}{4^k} - \log(N) - k\right),$$

which gives

$$\log(s) \geq \Omega\left(\frac{\sqrt{N}}{k4^k} - \frac{\log(N)}{k} - 1\right).$$

Thus,

$$s \geq \exp\left(\Omega\left(\frac{\sqrt{N}}{k4^k} - \frac{\log(N)}{k} - 1\right)\right) \geq \exp\left(\Omega\left(\frac{\sqrt{N}}{k4^k} - \frac{\log(N)}{k}\right)\right). \qquad \square$$

By definition, polynomial-size $\mathsf{MAJ} \circ \mathsf{MAJ}$ circuits can be simulated by polynomial-size $\mathsf{MAJ} \circ \mathsf{SYM}$ circuits. Also, Goldmann et al. [17] (Theorem 26) showed that $\mathsf{MAJ} \circ \mathsf{THR}$ circuits can be simulated by $\mathsf{MAJ} \circ \mathsf{MAJ}$ circuits with a polynomial blowup. More precisely, a $\mathsf{MAJ} \circ \mathsf{THR}$ circuit of size s can be simulated by a $\mathsf{MAJ} \circ \mathsf{MAJ}$ circuit of size $s^\alpha \cdot n^\beta$ for some constants α, β. Hence, Corollary 1.5 follows by a similar proof as that of Lemma 4.2.

5 Conclusion

We have shown that GHR_k^N requires essentially $\Omega(\sqrt{N}/4^k)$ cost to be solved in the $\mathsf{PP}_{k+1}^{\mathsf{cc}}$ model. Since it follows almost from the definition of GHR_k^N that it has $O(\log N)$ cost $\mathsf{UPP}_{k+1}^{\mathsf{cc}}$ protocols, this provides a separation of $\mathsf{PP}_k^{\mathsf{cc}}$ from $\mathsf{UPP}_k^{\mathsf{cc}}$ for the NOF model when $k \leq \delta \cdot \log N$ for some constant $\delta > 0$. In general, current techniques do not allow us to go beyond $\log N$ players to prove lower bounds for the cost of even deterministic protocols. This remains one of the most interesting problems in NOF complexity. However, let us remark that for many of the functions used in the literature (see, for example, [18, 3, 1, 15]), there are surprisingly efficient protocols when $k > \log N$. Moreover these protocols are typically deterministic and either simultaneous or barely interactive. On the other hand, we do not immediately see an efficient randomized interactive protocol for GHR_k^N at $k > \log N$. This raises the question of whether GHR_k^N is a hard function even for $k > \log N$.

Our result shows that the $\mathsf{PP}_k^{\mathsf{cc}}$ complexity of GHR_k^N is $\Omega(\sqrt{N})$ for any constant k. As mentioned in Section 1.1, Sherstov [34] shows existence of functions with $\Omega(N)$ cost in PP_k but that have efficient UPP_k protocols. A question that may be with reach is whether one can come up with an explicit function in $\mathsf{UPP}_k^{\mathsf{cc}}$ that requires $\Omega(N)$ PP_k cost.

Finally, proving super-logarithmic lower bounds for $\mathsf{UPP}_k^{\mathsf{cc}}$ protocols for any explicit function remains a very interesting challenge even for $k = 3$. Hansen and Podolskii [19] have shown that meeting this challenge is enough to yield super-polynomial lower bounds for $\mathsf{THR} \circ \mathsf{THR}$ circuits.

Acknowledgements

We would like to thank Kristoffer Hansen for directing our attention to the result of Goldmann, Håstad and Razborov [17] during the Summer School on Lower Bounds, held in Prague in the summer of 2015. We thank Michal Koucký for inviting us there, where we started this project. We are grateful to an anonymous reviewer for pointing out to us that the results of Sherstov [31] and Beigel [7] can be combined to get a separation between $\mathsf{PP}_k^{\mathsf{cc}}$ and $\mathsf{UPP}_k^{\mathsf{cc}}$ for k at most $O(\log\log n)$. We would like to thank Alexander Sherstov [32] for his various helpful comments and pointers. Finally, we are thankful to the anonymous ToC referees and Laci Babai, who provided detailed comments and valuable advice that helped in significantly polishing this paper.

References

[1] ANIL ADA, ARKADEV CHATTOPADHYAY, OMAR FAWZI, AND PHUONG NGUYEN: The NOF multiparty communication complexity of composed functions. *Comput. Complexity*, 24(3):645–694, 2015. Preliminary version in ICALP'12. [doi:10.1007/s00037-013-0078-4] 2, 19

[2] LÁSZLÓ BABAI, PETER FRANKL, AND JANOS SIMON: Complexity classes in communication complexity theory. In *Proc. 27th FOCS*, pp. 337–347. IEEE Comp. Soc. Press, 1986. [doi:10.1109/SFCS.1986.15] 2, 3

[3] LÁSZLÓ BABAI, ANNA GÁL, PETER G. KIMMEL, AND SATYANARAYANA V. LOKAM: Communication complexity of simultaneous messages. *SIAM J. Comput.*, 33(1):137–166, 2003. Preliminary version in STACS'95. [doi:10.1137/S0097539700375944] 2, 19

[4] LÁSZLÓ BABAI, NOAM NISAN, AND MARIO SZEGEDY: Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. System Sci.*, 45(2):204–232, 1992. Preliminary version in STOC'89. [doi:10.1016/0022-0000(92)90047-M] 7

[5] PAUL BEAME, MATEI DAVID, TONIANN PITASSI, AND PHILIPP WOELFEL: Separating deterministic from randomized multiparty communication complexity. *Theory of Computing*, 6(9):201–225, 2010. Preliminary version in ICALP'07. [doi:10.4086/toc.2010.v006a009] 2

[6] PAUL BEAME AND DANG-TRINH HUYNH-NGOC: Multiparty communication complexity and threshold circuit size of AC^0. *SIAM J. Comput.*, 41(3):484–518, 2012. Preliminary version in FOCS'09. [doi:10.1137/100792779]

[7] RICHARD BEIGEL: Perceptrons, PP, and the Polynomial Hierarchy. *Comput. Complexity*, 4(4):339–349, 1994. Preliminary version in SCT'92. [doi:10.1007/BF01263422] 4, 19

[8] HARRY BUHRMAN, NIKOLAI K. VERESHCHAGIN, AND RONALD DE WOLF: On computation and communication with small bias. In *Proc. 22nd IEEE Conf. on Computational Complexity (CCC'07)*, pp. 24–32. IEEE Comp. Soc. Press, 2007. [doi:10.1109/CCC.2007.18] 3

[9] MARK BUN AND JUSTIN THALER: Approximate degree and the complexity of depth three circuits. In *Proc. 22nd Internat. Workshop on Randomization and Computation (RANDOM'18)*, pp. 35:1–35:18. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018. [doi:10.4230/LIPIcs.APPROX-RANDOM.2018.35]

[10] ASHOK K. CHANDRA, MERRICK L. FURST, AND RICHARD J. LIPTON: Multi-party protocols. In *Proc. 15th STOC*, pp. 94–99. ACM Press, 1983. [doi:10.1145/800061.808737] 2, 6

[11] ARKADEV CHATTOPADHYAY: Discrepancy and the power of bottom fan-in in depth-three circuits. In *Proc. 48th FOCS*, pp. 449–458. IEEE Comp. Soc. Press, 2007. [doi:10.1109/FOCS.2007.30] 18

[12] ARKADEV CHATTOPADHYAY: *Circuits, Communication and Polynomials*. Ph. D. thesis, McGill University, 2009. 2, 7

[13] ARKADEV CHATTOPADHYAY AND ANIL ADA: Multiparty communication complexity of disjointness. *Electron. Colloq. on Comput. Complexity (ECCC)*, January 2008. [ECCC:TR08-002] 2

[14] ARKADEV CHATTOPADHYAY AND NIKHIL MANDE: Small error versus unbounded error protocols in the NOF model. *Electron. Colloq. on Comput. Complexity (ECCC)*, September 2016. [ECCC:TR16-095] 4

[15] ARKADEV CHATTOPADHYAY AND MICHAEL E. SAKS: The power of super-logarithmic number of players. In *Proc. 18th Internat. Workshop on Randomization and Computation (RANDOM'14)*, pp. 596–603. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2014. [doi:10.4230/LIPIcs.APPROX-RANDOM.2014.596] 2, 19

[16] ANDREW DRUCKER, FABIAN KUHN, AND ROTEM OSHMAN: On the power of the congested clique model. In *Proc. 33th Symp. on Principles of Distributed Comput. (PODC'14)*, pp. 367–376. ACM Press, 2014. [doi:10.1145/2611462.2611493] 2

[17] MIKAEL GOLDMANN, JOHAN HÅSTAD, AND ALEXANDER A. RAZBOROV: Majority gates vs. general weighted threshold gates. *Comput. Complexity*, 2(4):277–300, 1992. Preliminary version in SCT'92. [doi:10.1007/BF01200426] 3, 4, 5, 6, 8, 9, 19

[18] VINCE GROLMUSZ: The BNS lower bound for multi-party protocols in nearly optimal. *Inform. and Comput.*, 112(1):51–54, 1994. [doi:10.1006/inco.1994.1051] 2, 19

[19] KRISTOFFER ARNSFELT HANSEN AND VLADIMIR V. PODOLSKII: Polynomial threshold functions and boolean threshold circuits. *Inform. and Comput.*, 240:56–73, 2015. Preliminary version in MFCS'13. [doi:10.1016/j.ic.2014.09.008] 19

[20] JOHAN HÅSTAD: On the size of weights for threshold gates. *SIAM J. Discrete Math.*, 7(3):484–492, 1994. [doi:10.1137/S0895480192235878] 5

[21] HAMED HATAMI, KAAVE HOSSEINI, AND SHACHAR LOVETT: Structure of protocols for XOR functions. In *Proc. 57th FOCS*, pp. 282–288. IEEE Comp. Soc. Press, 2016. [doi:10.1109/FOCS.2016.38] 5

[22] BALA KALYANASUNDARAM AND GEORG SCHNITGER: The probabilistic communication complexity of set intersection. *SIAM J. Discrete Math.*, 5(4):545–557, 1992. Preliminary version in SCT'87. [doi:10.1137/0405044] 2

[23] EYAL KUSHILEVITZ AND NOAM NISAN: *Communication complexity*. Cambridge University Press, 1997. 6, 7

[24] TROY LEE AND ADI SHRAIBMAN: Disjointness is hard in the multiparty number-on-the-forehead model. *Comput. Complexity*, 18(2):309–336, 2009. Preliminary version in CCC'08. [doi:10.1007/s00037-009-0276-2] 2

[25] MIHAI PĂTRAŞCU: Towards polynomial lower bounds for dynamic problems. In *Proc. 42nd STOC*, pp. 603–610. ACM Press, 2010. [doi:10.1145/1806689.1806772] 2

[26] ANUP RAO AND AMIR YEHUDAYOFF: Simplified lower bounds on the multiparty communication complexity of disjointness. In *Proc. 30th IEEE Conf. on Computational Complexity (CCC'15)*, pp. 88–101. DROPS, 2015. [doi:10.4230/LIPIcs.CCC.2015.88] 2

[27] RAN RAZ: The BNS-Chung criterion for multi-party communication complexity. *Comput. Complexity*, 9(2):113–122, 2000. [doi:10.1007/PL00001602] 7

[28] ALEXANDER A. RAZBOROV: On the distributional complexity of disjointness. *Theoret. Comput. Sci.*, 106(2):385–390, 1992. Preliminary version in ICALP'90. [doi:10.1016/0304-3975(92)90260-M] 2

[29] ALEXANDER A. SHERSTOV: Halfspace matrices. *Comput. Complexity*, 17(2):149–178, 2008. Preliminary version in CCC'07. [doi:10.1007/s00037-008-0242-4] 3, 5, 8

[30] ALEXANDER A. SHERSTOV: The intersection of two halfspaces has high threshold degree. *SIAM J. Comput.*, 42(6):2329–2374, 2013. Preliminary version in FOCS'09. [doi:10.1137/100785260] 5

[31] ALEXANDER A. SHERSTOV: Communication lower bounds using directional derivatives. *J. ACM*, 61(6):34:1–34:71, 2014. Preliminary version in STOC'13. [doi:10.1145/2629334] 2, 4, 19

[32] ALEXANDER A. SHERSTOV: Private Communication, 2016. 4, 19

[33] ALEXANDER A. SHERSTOV: The multiparty communication complexity of set disjointness. *SIAM J. Comput.*, 45(4):1450–1489, 2016. Preliminary version in STOC'12. [doi:10.1137/120891587] 2, 4, 5

[34] ALEXANDER A. SHERSTOV: On multiparty communication with large versus unbounded error. *Theory of Computing*, 14(22), 2018. Preliminary version ECCC TR16-138. [doi:10.4086/toc.2018.v014a022] 4, 5, 19

[35] JUSTIN THALER: Lower bounds for the approximate degree of block-composed functions. In *Proc. 43rd Internat. Colloq. on Automata, Languages and Programming (ICALP'16)*, pp. 17:1–17:15. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016. [doi:10.4230/LIPIcs.ICALP.2016.17, ECCC:TR14-150] 4

[36] SHENGYU ZHANG: Efficient quantum protocols for XOR functions. In *Proc. 25th SODA*, pp. 1878–1885. ACM Press, 2014. [doi:10.1137/1.9781611973402.136] 5

AUTHORS

Arkadev Chattopadhyay
Associate professor
Tata Institute of Fundamental Research
Mumbai, India
arkadev.c@tifr.res.in
http://www.tcs.tifr.res.in/~arkadev/

Nikhil S. Mande
Postdoctoral fellow
Georgetown University
nm.936@georgetown.edu
http://www.tcs.tifr.res.in/~nikhil/

ABOUT THE AUTHORS

ARKADEV CHATTOPADHYAY is currently an Associate Professor at the School of Technology and Computer Science at TIFR, Mumbai, which he joined in September 2012. He obtained his Ph. D. from McGill University in 2008 under the supervision of Denis Thérien. He was a member of the School of Mathematics at the Institute for Advanced Study, Princeton for the year 2008-09. He was a postdoctoral member in the Department of Computer Science at the University of Toronto from 2009 to 2012. His research interests are in computational and communication complexity. Outside of this, he is also interested in literature, films and politics.

NIKHIL S. MANDE is a postdoctoral fellow at the Department of Computer Science at Georgetown University. This work was done while he was a Ph. D. student at the School of Technology and Computer Science at TIFR, Mumbai, where he was advised by Arkadev Chattopadhyay and graduated in 2018. His research interests lie in communication complexity and circuit complexity. In his spare time, he enjoys speedsolving Rubik's cube and related puzzles; his official records can be found here.