Electronic Colloquium on Computational Complexity, Report No. 61 (2008)

Multiparty Communication Complexity of AC^0

Paul Beame*
Computer Science and Engineering
University of Washington
Seattle, WA 98195-2350
beame@cs.washington.edu

Dang-Trinh Huynh-Ngoc †
Computer Science and Engineering
University of Washington
Seattle, WA 98195-2350
trinh@cs.washington.edu

June 10, 2008

Abstract

We prove $n^{\Omega(1)}$ lower bounds on the multiparty communication complexity of AC^0 functions in the number-on-forehead (NOF) model for up to $\Theta(\log n)$ players. These are the first lower bounds for any AC^0 function for $\omega(\log\log n)$ players. In particular we show that there are families of depth 3 read-once AC^0 formulas having k-player randomized multiparty NOF communication complexity $n^{\Omega(1)}/2^{O(k)}$. We show similar lower bounds for depth 4 read-once AC^0 formulas that have nondeterministic communication complexity $O(\log^2 n)$, yielding exponential separations between k-party nondeterministic and randomized communication complexity for AC^0 functions.

As a consequence of the latter bound, we obtain an $n^{\Omega(1/k)}/2^{O(k)}$ lower bound on the k-party NOF communication complexity of set disjointness. This is non-trivial for up to $\Theta(\sqrt{\log n})$ players which is significantly larger than the up to $\Theta(\log\log n)$ players allowed in the best previous lower bounds for multiparty set disjointness given by Lee and Shraibman [LS08] and Chattopadhyay and Ada [CA08] (though our complexity bounds themselves are not as strong as those in [LS08, CA08] for $o(\log\log n)$ players).

We derive these results by extending the k-party generalization in [CA08, LS08] of the pattern matrix method of Sherstov [She07, She08]. Using this technique, we derive a new sufficient criterion for strong communication complexity lower bounds based on functions having many diverse subfunctions that do not have good low-degree polynomial approximations. This criterion guarantees that such functions have orthogonalizing distributions that are "max-smooth" as opposed to the "min-smooth" orthogonalizing distributions used by Razborov and Sherstov [RS08] to analyze the sign-rank of AC^0.

*Research supported by NSF grant CCF-0514870
†Research supported by a Vietnam Education Foundation Fellowship

ISSN 1433-8092

1 Introduction

Recently, Sherstov introduced the so-called pattern matrix method to derive discrepancy bounds [She07, She08] yielding a new strong method for obtaining lower bounds for 2-party quantum communication complexity. His method was then generalized for $k \geq 2$ players [Cha07, CA08, LS08] to yield the first lower bounds for the general multiparty number-on-forehead communication complexity of set disjointness for more than 2 players, improving a long line of research on the problem. The communication lower bound for k players is $\Omega(n^{\frac{1}{2k}}/2^{2^{\Theta(k)}})$ which yields a non-trivial separation between randomized and nondeterministic k-party models for $k \leq \epsilon \log \log n$ for some constant $\epsilon > 0$. This separation between randomized and nondeterministic communication complexity was extended by David and Pitassi and David, Pitassi, and Viola to $\Omega(\log n)$ players for significantly more complex functions than disjointness that based on pseudorandom generators [DPV]. Their construction uses a more complex criterion than the simple masking version of the pattern matrix method used in [CA08]. Set disjointness is an AC^0 function and David, Pitassi, and Viola asked the question of whether one could prove a separation for $\Omega(\log n)$ players using an AC^0 function or even whether one could prove any non-trivial lower bound for $\omega(\log \log n)$ players for any AC^0 function since their functions are also only in AC^0 for $k = O(\log \log n)$.

We resolve this question positively by showing there is a read-once function in AC^0_3 that has $n^{\Omega(1)}$ communication complexity for $k = \Omega(\log n)$ players. Moreover there is a read-once function in AC^0_4 that for $\Omega(\log n)$ players has nondeterministic communication complexity $O(\log^2 n)$ and randomized communication complexity $n^{\Omega(1)}$; i.e., $(\mathsf{NP}^{\mathsf{cc}}_{\mathsf{k(n)}} - \mathsf{BPP}^{\mathsf{cc}}_{\mathsf{k(n)}}) \cap \mathsf{AC}^0_4 \neq \emptyset$ for $k(n) \leq \delta \log n$ for some explicit constant $\delta > 0$. Our method significantly improves the power of the pattern matrix method for proving strong communication complexity lower bounds.

As a consequence of the lower bound for the function we use to separate $\mathsf{NP}^{\mathsf{cc}}_{\mathsf{k(n)}}$ from $\mathsf{BPP}^{\mathsf{cc}}_{\mathsf{k(n)}}$, we obtain $n^{\Omega(1/k)}/2^{O(k)}$ lower bounds on the k-party NOF communication complexity of set disjointness which is non-trivial for up to $\Theta(\sqrt{\log n})$ players. The best previous lower bounds of Lee and Shraibman [LS08] and Chattopadhyay and Ada [CA08] for set disjointness describe above do not apply for $\omega(\log \log n)$ players.

The high-level idea of the k-party version of the pattern matrix method as described in [CA08] is as follows. Suppose that we want to prove k-party lower bounds for a function $\mathcal{F}$. The general idea is to show that $\mathcal{F}$ can express some $\mathcal{F}^f_k$ (specified below) which is a function that under many projection patterns is the same as a function f of large approximate degree. If f has large approximate degree, then Sherstov showed that there exists another function g and a distribution μ on inputs such that with respect to μ, g is both highly correlated with f and orthogonal to all low-degree polynomials. It follows that $\mathcal{F}^f_k$ is also highly correlated with $\mathcal{F}^g_k$ and, using the generalized discrepancy method for communication complexity lower bounds it suffices to prove a discrepancy lower bound for the latter function. Thanks to the orthogonality of g to all low degree polynomials this is possible using an iterated application of the Cauchy-Schwartz inequality as in Babai, Nisan, and Szegedy [BNS92]. For example, the bound for set disjointness $\textsc{Disj}_{k,n}(x) = \vee_{i=1}^{n} \wedge_{j=1}^{k} x_{ij}$, which more properly should be called set intersection, corresponds to the case that $f = \textsc{Or}$ which has approximate degree $\Omega(\sqrt{n})$.

In the two party case, Razborov and Sherstov [RS08] extended Sherstov's method to yield sign-rank lower bounds for the AC^0_3 function $\mathcal{F}^{MP}_2$ where $MP(x) = \wedge_{i=1}^{m} \vee_{j=1}^{4m^2} x_{ij}$ is the so-called Minsky-Papert function which has approximate degree $\Omega(m)$. The key to their argument is to show that there is an orthogonalizing distribution μ for MP that is "min-smooth" in that it assigns probability at least $8^{-m}2^{-n-1}$ to any input vector on which MP is true.

We prove our results by showing that any function f for which there is a diverse collection of partial assignments ρ such that each of the subfunctions $f|_\rho$ of f requires large approximate degree, there is an orthogonalizing distribution μ for f that is "max-smooth" in that the probability of subsets defined by partial assignments cannot be too much larger than under the uniform distribution. The diversity of the partial assignments is determined by a parameter α so we call the degree bound the (ϵ, α)-approximate degree. This property is somewhat delicate but applies directly to PARITY and, much more importantly, to $\text{TRIBES}_{p,q}(x) = \vee_{i=1}^{q} \wedge_{j=1}^{p} x_{ij}$ for certain choices of p and q. Since $\text{TRIBES}_{p,q}$ is a subfunction of many other functions we can use it to obtain lower bounds for many functions in AC^0. (The property unfortunately does not apply to OR but we are able to derive our lower bounds for $\text{DISJ}_{k,n}$ via reduction.) Our lower bound method also shows that the simple masking version of the pattern matrix method is sufficient to obtain strong lower bounds.

Results Let T be the set of all Boolean functions that map the all 0's input to false and each input with precisely one 1 to true. For any integers $m, s, k > 0$, any Boolean function f on m bits, and any s-bit function $t \in T$, we define the following function on msk bits:

$$\mathcal{H}_k^{f,t}(x_1, \dots, x_k) := f(t(\wedge_{i=1}^k x_{11i}, \dots, \wedge_{i=1}^k x_{1si}), \dots, t(\wedge_{i=1}^k x_{m1i}, \dots, \wedge_{i=1}^k x_{msi})),$$

for any $x_1, \dots, x_k \in \{0,1\}^{ms}$. Let $n = ms$. We associate each such $\mathcal{H}_K^{f,t}$ with the k-party NOF communication problem in which player i can see all x_j except for x_i and they want to compute $\mathcal{H}_k^{f,t}$.

For instance, setting f and t to be OR makes $\mathcal{H}_k^{f,t}$ the set disjointness function $\text{DISJ}_{k,n}$ and setting both f and t to be PARITY makes $\mathcal{H}_k^{f,t}$ the Generalized Inner Product (GIP) function.

Given f as above and $n = ms$, we also define a function on nk bits by:

$$\mathcal{F}_k^f(x, y_1, \dots, y_{k-1}) := f(x|\phi_{AND}(\wedge_{j=1}^{k-1} y_j)),$$

where $\phi_{AND}(z)$ returns the set of non-zero indices in z and $x|S$ is the bit vector obtained by restricting x to indices in S. We also associate with each $\mathcal{F}_k^f$ the k-party NOF communication problem on $x, y_1, \dots, y_{k-1} \in \{0,1\}^n$ in which the player 0 holds x and for $1 \le i \le k-1$, player i holds y_i, and they want to compute $\mathcal{F}_k^f$.

If we partition the above n-bit input string x into m blocks of size s and we restrict the inputs $y_1, \dots, y_{k-1}$ such that the set $S = \phi_{AND}(\wedge_{j=1}^{k-1} y_j)$ as above selects exactly one bit in each of the m blocks, then it is easy to see that in this case $\mathcal{F}_k^f$ is a subfunction of $\mathcal{H}_k^{f,t}$. From now on, unless stated otherwise, we will assume that the inputs always satisfy this restriction.

We show that (ϵ, α)-approximate degree lower bounds for a function f allows one to derive lower bounds for $\mathcal{F}_k^f$.

Theorem 1.1. *Let $n = ms$ where $s = \lceil 4e \rceil^{k-1}$. For $0 \le \alpha < 1$ and any Boolean function f on m bits with $(5/6, \alpha)$-approximate degree d, the function $\mathcal{F}_k^f$ defined on $n \cdot k$ bits requires $R_{1/3}^k(\mathcal{F}_k^f)$ that is $\Omega(d/2^k)$ for $k \le (1-\alpha)\log_2 d$.*

Corollary 1.2. *Under the same conditions as above, if t is any s-bit function in T then $\mathcal{H}_k^{f,t}$ has k-party randomized NOF communication complexity $\Omega(d/2^k)$.*

By analyzing the approximation properties of the $m = ps$ bit function $\textsc{Tribes}_{p,q}$ for suitable choices of p and q, we obtain the first AC^0 function separating $\mathsf{NP}_\mathsf{k}^{\mathsf{cc}}$ from $\mathsf{BPP}_\mathsf{k}^{\mathsf{cc}}$ for $k = \omega(\log \log n)$. The separation is non-trivial for k up to $\Theta(\log n)$.

Theorem 1.3. *There is a constant $a > 0$ such for any n and integer $k = k(n) \le a \log_2 n$ such that $n \ge ms$ for $s = \lceil 4e \rceil^{k-1}$ and m is of the form pq such that $\lceil q^{0.2} \rceil < 2^p \le \frac{1}{6} q^{0.3} \ln 2$ and the following holds. The randomized k-party NOF communication complexity of $\mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}_s}$ is $\Omega(n^{0.3-\epsilon}/c^k)$ for $c = 2(4e)^{0.3}$ and its nondeterministic k-party NOF communication complexity is $O(\log^2 n)$.*

Since the function $\mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}}$ is given by a read-once depth 4 formula we have the following theorem.

Corollary 1.4. *There is a constant $\delta > 0$ such that for any $k(n) \le \delta \log_2 n$ there is a function in $\mathsf{AC}_4^0 \cap (\mathsf{NP}_\mathsf{k}^{\mathsf{cc}} - \mathsf{BPP}_\mathsf{k}^{\mathsf{cc}})$.*

By a reduction from $\mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}_s}$ to $\textsc{Disj}_{k,n}$ we obtain the following lower bound.

Theorem 1.5. *The randomized k-party NOF communication complexity of $\textsc{Disj}_{k,n}$ is $n^{\Omega(1/k)}/2^{O(k)}$.*

Write $\textsc{Tribes}'_{p,q}$ for the dual function to $\textsc{Tribes}_{p,q}$, $\textsc{Tribes}'_{p,q}(x) = \wedge_{i=1}^{q} \vee_{j=1}^{p} x_{ij}$. Observe that $\mathcal{H}_k^{\textsc{Tribes}'_{p,q},\textsc{Or}}$ is a read-once depth 3 AC^0 function since the two layers of $\vee$ gates can be combined. Since $\textsc{Tribes}'_{p,q}$ has the same degree approximation properties as $\textsc{Tribes}_{p,q}$, we obtain a similar lower bound for read-once AC_3^0 functions.

Theorem 1.6. *There is a constant $a > 0$ such for any n and integer $k = k(n) \le a \log_2 n$ such that following holds. There is a function $\mathcal{F}$ in read-once AC_3^0, namely $\mathcal{H}_k^{\textsc{Tribes}'_{p,q},\textsc{Or}_s}$ for $s = \lceil 4e \rceil^{k-1}$ and $\lceil q^{0.2} \rceil < 2^p \le \frac{1}{6} q^{0.3} \ln 2$, whose randomized k-party NOF communication complexity is $\Omega(n^{0.3-\epsilon}/c^k)$ for $c = 2(4e)^{0.3}$.*

Our technique yields a new sufficient criterion for functions to have high randomized communication complexity. It can be used to provide strong lower bounds for $k = O(\log n)$ for many other functions including the extension of the Minsky-Papert function MP considered by Razborov and Sherstov, and the Generalized Inner Product.

Our paper is organized as follows. In Section 2 we give an overview of the method of [She08, CA08] based on orthogonalizing distributions for functions of large ϵ-approximate degree and briefly discuss its limitations. In Sections 3 and 4 we define a new notion which we call the (ϵ, α)-approximate degree of a function and show how we can use it to prove Theorem 1.1. In Section 5 we prove that the function $\textsc{Tribes}_{p,q}$ has large (ϵ, α)-approximate degree. Then we prove Theorems 1.3, 1.5, and 1.6 in Section 6 and conclude by considering a variety of other functions.

2 Preliminaries

2.1 Notations and Terminology

We follow the notation used in [DPV]. We will assume that a Boolean function on m bits is a map $f : \{0,1\}^m \to \{-1,1\}$.

Correlation Let $f, g : \{0,1\}^m \mapsto \mathbb{R}$ be two functions, and let μ be a distribution on $\{0,1\}^m$. We define the *correlation* between f and g under μ to be $\mathrm{Cor}_\mu(f,g) := \mathbf{E}_{x\sim\mu}[f(x)g(x)]$. If $\mathcal{G}$ is a class of functions $g : \{0,1\}^m \mapsto \mathbb{R}$, we define the correlation between f and $\mathcal{G}$ under μ to be $\mathrm{Cor}_\mu(f,G) := \max_{g\in\mathcal{G}} \mathrm{Cor}_\mu(f,g)$.

Communication complexity We denote by $R^k_\epsilon(f)$ the cost of the best k-party randomized NOF communication protocol for f with two-sided error at most ϵ, and $N^k(f)$ the cost of the best k-party nondeterministic communication protocol for f. We denote by Π^c_k the class of all deterministic k-party communication protocols of cost at most c.

Fact 2.1. *[KN97] If there exists a distribution μ such that $\mathrm{Cor}_\mu(f,\Pi^c_k) \le 1/3$ then $R^k_{1/3}(f) \ge c$.*

Lemma 2.2 ([BNS92])**.** *Let $f : \{0,1\}^{m\times k} \mapsto \mathbb{R}$ and U_m be the uniform distribution on $\{0,1\}^m$. Then,*

$$\mathrm{Cor}_{U_m}(f,\Pi^c_k)^{2^{k-1}} \le 2^{c\cdot 2^{k-1}}\cdot \mathbf{E}_{y^0_1,\ldots,y^0_{k-1},y^1_1,\ldots,y^1_{k-1}\in\{0,1\}^m}\left[\left|\mathbf{E}_{x\in\{0,1\}^m}\left[\Pi_{u\in\{0,1\}^{k-1}} f(x,y^{u_1}_1,\ldots,y^{u_{k-1}}_{k-1})\right]\right|\right].$$

Approximate degree The ϵ-approximate degree of f, $deg_\epsilon(f)$, is the smallest d for which there exists a multivariate real-valued polynomial p of degree d such that $||f-p||_\infty = \max_x |f(x)-p(x)| \le \epsilon$. Following [NS94] we have the following property of approximate degree of OR.

Proposition 2.3. *Let $\mathrm{OR}_m : \{0,1\}^m \to \{1,-1\}$. For $0 \le \epsilon < 1$, $deg_\epsilon(\mathrm{OR}_m) \ge \sqrt{(1-\epsilon)m/2}$.*

Define an inner product $\langle,\rangle$ on the set of functions $f : \{0,1\}^m \to \mathbb{R}$ by $\langle f,g\rangle = \mathbf{E}[f\cdot g]$. For $S \subseteq [m]$, let $\chi_S : \{0,1\}^m \to \{-1,1\}$ be the function $\chi_S = \prod_{i\in S}(-1)^{x_i}$. The χ_S for $S \subseteq [m]$ form an orthonormal basis of this space.

Lemma 2.4 ([She08])**.** *If $f : \{0,1\}^m \mapsto \{-1,1\}$ is a Boolean function with $deg_\epsilon(f) \ge d$ then there exists a function $g : \{0,1\}^m \mapsto \{-1,1\}$ and a distribution μ on $\{0,1\}^m$ such that:*

1. *$\mathrm{Cor}_\mu(g,f) > \epsilon$; and*
2. *for every $S \subseteq [m]$ with $|S| < d$ and every function $h : \{0,1\}^{|S|} \mapsto \mathbb{R}$, $\mathbf{E}_{x\sim\mu}[g(x)\cdot h(x|S)] = 0$.*

Proof. Let Φ_d be the space of polynomials of degree less than d. By definition, $deg_\epsilon(f) \ge d$ if and only if $\min_{q\in\Phi_d} ||f-q||_\infty > \epsilon$. By duality of norms we have $\min_{q\in\Phi_d} ||f-q||_\infty = \max_{p\in\Phi^\perp_d,\ ||p||_1=1}\langle f,p\rangle$. Writing $\mu(x) = |p(x)|$ the condition $||p||_1 = 1$ implies that μ is a probability distribution and letting $g(x) = p(x)/\mu(x)$ for $\mu(x) \ne 0$ and $g(x) = 1$ if $\mu(x) = 0$. Then $p(x) = \mu(x)g(x)$. Therefore

$$\epsilon < \langle f,p\rangle = \mathbf{E}[f\cdot p] = \mathbf{E}[f\cdot g\cdot\mu] = \mathbf{E}_{x\sim\mu}[f(x)g(x)] = \mathrm{Cor}_\mu(f,g).$$

Moreover since $p \in \Phi^\perp_d$, we have $0 = \langle\chi_S,p\rangle = \mathbf{E}_{x\sim\mu}[\chi_S(x)g(x)]$. Now for $h : \{0,1\}^{|S|} \to \mathbb{R}$ for $|S| \le d$, $h(x|S)$ can be expressed as a degree $|S|$ polynomial and by linearity $\mathbf{E}_{x\sim\mu}[g(x)\cdot h(x|S)] = 0$. □

We will extend this lemma in Section 3 using more general LP duality.

2.2 The correlation method

We give an overview of the method as described in [CA08], which extends ideas of [She07, She08] from 2-party to k-party communication complexity, with specific details at those points that we are extending in this paper.

Given a Boolean function f on m bits, where f has large 5/6-approximate degree d (i.e, d is polynomial in m), we want to lower bound $R^k_{1/3}(\mathcal{F}^f_k)$, where $\mathcal{F}^f_k(x, y_1, \ldots, y_{k-1})$ is on $n \cdot k$ bits for $n = m \cdot s$.

From Lemma 2.4, we obtain another Boolean function g and a distribution μ such that:

1. $\mathrm{Cor}_\mu(g, f) \geq 5/6$; and
2. for every $S \subseteq [m]$ with $|S| < d$ and every function $h : \{0,1\}^{|S|} \mapsto \mathbb{R}$, $\mathbf{E}_{x \sim \mu}[g(x) \cdot h(x|S)] = 0$.

Divide each player's n-bit input into m blocks of size s. Let ℓ be that $n/m = s = \ell^{k-1}$. Hence we can imagine that x consists of m arrays, each having $k-1$ dimensions. For $1 \leq i \leq k-1$, each of the m blocks in y_i is (a bit vector representing) an index in $[\ell]$. Therefore we can view each y_i as in $[\ell]^m$. Thus $\phi_{AND}(y_1, \ldots, y_{k-1})$ selects exactly one bit of x in each of m blocks.

Based on μ, we define a distribution λ on $n \cdot k$ bits in a straightforward way as follows:

$$\lambda(x, y_1, \ldots, y_{k-1}) := \frac{\mu(x|\phi_{AND}(y_1, \ldots, y_{k-1}))}{\ell^{km} 2^{n-m}}$$

for eligible $y_1, \ldots, y_{k-1}$ and 0 otherwise. Here "eligible" means that $y_1, \ldots, y_{k-1}$ satisfy the above requirements. Then it can be verified that $\mathrm{Cor}_\gamma(\mathcal{F}^f_k, \mathcal{F}^g_k) = \mathrm{Cor}_\mu(f, g) \geq 5/6$. Consequently,

$$\mathrm{Cor}_\lambda(\mathcal{F}^f_k, \Pi^c_k) \leq \mathrm{Cor}_\lambda(\mathcal{F}^g_k, \Pi^c_k) + 1/6.$$

Therefore we only need to bound $\mathrm{Cor}_\lambda(\mathcal{F}^g_k, \Pi^c_k)$. Then by Lemma 2.2,

$$\begin{aligned}\mathrm{Cor}_\lambda(\mathcal{F}^g_k, \Pi^c_k))^{2^{k-1}} &= 2^{m2^{k-1}} \mathrm{Cor}_{U_m}(\mu(x|\phi_{AND}(y_1, \ldots, y_{k-1}))g(x|\phi_{AND}(y_1, \ldots, y_{k-1}), \Pi^c_k)^{2^{k-1}} \\ &\leq 2^{(c+m)\cdot 2^{k-1}} \cdot \mathbf{E}_{y^0_1,\ldots,y^0_{k-1},y^1_1,\ldots,y^1_{k-1}} H(y^0_1, \ldots, y^0_{k-1}, y^1_1, \ldots, y^1_{k-1}),\end{aligned}$$

where

$$H(y^0_1, \ldots, y^0_{k-1}, y^1_1, \ldots, y^1_{k-1}) := \left| \mathbf{E}_x \left[\Pi_{u \in \{0,1\}^{k-1}} \mu(x|\phi_{AND}(y^{u_1}_1, \ldots, y^{u_{k-1}}_{k-1}))g(x|\phi_{AND}(y^{u_1}_1, \ldots, y^{u_{k-1}}_{k-1})) \right] \right|.$$

For $1 \leq i \leq k-1$, let $r_i \in \{0, \ldots, m\}$ be the number of blocks for which y^0_i and y^1_i give the same index. Let $r = \sum r_i$. We rely on the following three propositions to continue the proof. Proposition 2.5 and Proposition 2.7 are the same as in [CA08], so we do not give their proofs. We will prove an extension of Proposition 2.6 in Section 3.

Proposition 2.5. *If $r < d$, then $H(y^0_1, \ldots, y^0_{k-1}, y^1_1, \ldots, y^1_{k-1}) = 0$.*

Proposition 2.6. $H(y^0_1, \ldots, y^0_{k-1}, y^1_1, \ldots, y^1_{k-1}) \leq \frac{2^{(2^{k-1}-1)r}}{2^{2^{k-1}m}}$.

Proposition 2.7. *For $d \leq j \leq (k-1)m$,* $\Pr[r = j] \leq (\frac{e(k-1)m}{j(\ell-1)})^j (1 - \frac{1}{\ell})^{(k-1)m}$.

In [CA08, LS08], to prove the lower bound for $\textsc{Disj}_{k,n}$, the function f is set to $\textsc{Or}_m$ and t is set to $\textsc{Or}_s$. By Proposition 2.3, $d = deg_{5/6}(\textsc{Or}_m) \geq \sqrt{m/12}$. Plugging the bound in Proposition 2.7 together with the bounds from Proposition 2.5 for $r < d$ and from Proposition 2.6 when $r \geq d$ into the above correlation inequality it is not hard to show that

$$\mathrm{Cor}_\lambda(\mathcal{F}_k^g, \Pi_k^c)) \leq \frac{2^c}{2^{d/2^k}},$$

for $\ell > \frac{2^{2^k} kem}{d}$. Hence for $k = O(\log\log n)$ and c a small enough polynomial in n, we have a polynomial lower bound for $R_{1/3}^k(\textsc{Disj}_{k,n}) \geq c$.

The key limitation of the above technique is the required lower bound on ℓ which follows from the weakness of the upper bound in Proposition 2.6. That weakness is implied by how little can be assumed about the orthogonalizing distribution μ given by Lemma 2.4. In particular, the arguments in [She08, CA08, LS08] all allow that μ may assign all of its probability mass to small subsets of points defined by partial assignments. Indeed, when the function f is $\textsc{Or}_m$, this is the case. However, we will show that for other very simple functions f one can choose the orthogonalizing distribution μ so that it does not assign too much weight on such small sets of points; that is, μ is "max-smooth". To guarantee this property of μ we need to strengthen Lemma 2.4 by assuming more of f than just large approximate degree.

3 Beyond approximate degree: a new sufficient criterion for strong communication complexity bounds

A $\rho \in \{0, 1, *\}^m$ is called a *restriction*. For any restriction ρ, let $\mathrm{unset}(\rho) \subseteq [m]$ be the set of star positions in ρ, let $|\rho| = m - |\mathrm{unset}(\rho)|$, and let C_ρ be the set of all $x \in \{0, 1\}^m$ such that for any $1 \leq i \leq m$, either $\rho_i = *$ or $\rho_i = x_i$. Hence $|C_\rho| = 2^{m-|\rho|}$. Given a restriction $\rho \in \{0, 1, *\}^m$ and a function f on $\{0, 1\}^m$, we define $f|_\rho$ on $\{0, 1\}^{m-|\rho|}$ in the natural way.

The approximate degree of a function f says how hard it is to approximate f. In this paper, we need a stronger notion which requires that many widely distributed restrictions of f also require large approximate degree.

Definition Given $0 < \epsilon, \alpha \leq 1$ and $d > 0$, let $\Pi = \Pi_{d,\epsilon}(f) \subseteq \{0, 1, *\}^m$ be a set of restrictions such that for any $\pi \in \Pi$, $deg_\epsilon(f|_\pi) \geq d$. We say that f has (ϵ, α)-approximate degree at least d, denoted as $deg_{\epsilon,\alpha}(f) \geq d$, if restrictions in Π are spread out "evenly". Formally, there is a distribution ν on Π such that for any $\rho \in \{0, 1, *\}^m$ with $|\rho| \geq d^\alpha$, then

$$\Pr_{\pi \sim \nu}[C_\rho \cap C_\pi \neq \emptyset] \leq 2^{|\rho|^\alpha - |\rho|}.$$

The set Π and the distribution ν are the *witnesses* for the (ϵ, α)-approximate degree of f. Note that $deg_\epsilon(f) = deg_{\epsilon,1}(f)$.

We will use this definition to prove the following theorem.

Theorem 3.1 (restatement of Theorem 1.1)**.** *Let $n = ms$ where $s = \lceil 4e \rceil^{k-1}$. For $0 \leq \alpha < 1$ and any Boolean function f on m bits with $(5/6, \alpha)$-approximate degree d, the function $\mathcal{F}_k^f$ defined on $n \cdot k$ bits requires $R_{1/3}^k(\mathcal{F}_k^f)$ that is $\Omega(d/2^k)$ for $k \leq (1 - \alpha)\log_2 d$.*

To prove the theorem, we first need the following consequence of large (ϵ, α)-approximate degree. We postpone its proof to Section 4.

Lemma 3.2 (extension of Lemma 2.4)**.** *Given $0 < \epsilon, \alpha \leq 1$. If $f : \{0,1\}^m \mapsto \{-1,1\}$ is a Boolean function with (ϵ, α)-approximate degree d, there exist a function $g : \{0,1\}^m \mapsto \{-1,1\}$ and a distribution μ on $\{0,1\}^m$ such that:*

1. $\mathrm{Cor}_\mu(g, f) \geq \epsilon$;
2. *for every $T \subseteq [m]$ with $|T| < d$ and every function $h : \{0,1\}^{|T|} \mapsto \mathbb{R}$, $\mathbf{E}_{x\sim\mu}[g(x) \cdot h(x|T)] = 0$; and*
3. *for any restriction ρ with $|\rho| \geq d^\alpha$, $\mu(C_\rho) \leq 2^{|\rho|^\alpha - |\rho|}/\epsilon$.*

Note that, although the upper bound on $\mu(C_\rho)$ may seem quite weak, it will be sufficient to obtain an exponential improvement in the dependence of communication complexity lower bounds on k. Moreover, we note in Section 4 that for any function f computed by an AC^0 circuit the assumption and the upper bound are essentially the best possible for $d = \log^{\omega(1)} n$.

We now use Lemma 3.2 to prove an improvement of Proposition 2.6. This is the key to our improved bounds.

Lemma 3.3. *If $f : \{0,1\}^m \to \{1,-1\}$ has (ϵ, α)-approximate degree d, if g and μ are given by the application of Lemma 3.2 to f, and if $r \geq d$, then*

$$H(y_1^0, \dots, y_{k-1}^0, y_1^1, \dots, y_{k-1}^1) \leq \frac{2^{(2^{k-1}-1)r^\alpha}}{2^{2^{k-1}m} \epsilon^{2^{k-1}-1}}.$$

Proof. The proof of this lemma is similar to that of [CA08] except that we apply the upper bound from the third condition of Lemma 3.2. Let $Y_{0^{k-1}}$ represent the set of m variables indexed jointly by $y_1^0, \dots, y_{k-1}^0$. There is precisely one variable chosen from each of the m blocks. Then in increasing order for each nonzero $u \in \{0,1\}^{k-1}$, we let Y_u represent the set of variables indexed jointly by $y_1^{u_1}, \dots, y_{k-1}^{u_{k-1}}$ that are not in $Y_{0^{k-1}} \cup \bigcup_{u'<u} Y_{u'}$. By definition we then have for each nonzero u, $|Y_u| \geq m - r$. Let $Z = \bigcup Y_{u \in \{0,1\}^{k-1}}$.

Since g is 1/-1 valued,

$$\begin{aligned}
H(y_1^0, \dots, y_{k-1}^0, y_1^1, \dots, y_{k-1}^1) &= \left|\mathbf{E}_x\left[\Pi_{u\in\{0,1\}^{k-1}} \mu(x|\phi_{AND}(y_1^{u_1}, \dots, y_{k-1}^{u_{k-1}})) g(x|\phi_{AND}(y_1^{u_1}, \dots, y_{k-1}^{u_{k-1}})\right]\right| \\
&\leq \mathbf{E}_Z \Pi_{u\in\{0,1\}^{k-1}} \mu(x|\phi_{AND}(y_1^{u_1}, \dots, y_{k-1}^{u_{k-1}})) \\
&= \mathbf{E}_{Y_0^{k-1}} \mu(x|\phi_{AND}(y_1^0, \dots, y_{k-1}^0)) \qquad (1) \\
&\times \max_{Y_{0^{k-1}}} \mathbf{E}_{Y_{0\dots01}} \mu(x|\phi_{AND}(y_1^0, \dots, y_{k-1}^1)) \\
&\times \max_{Y_{0\dots0} \cup Y_{0\dots01}} \mathbf{E}_{Y_{0\dots10}} \mu(x|\phi_{AND}(y_1^0, \dots, y_{k-1}^0)) \\
&\times \dots
\end{aligned}$$

and so on repeatedly for all 2^{k-1} of the Y_u. The term at line (1) equals 2^{-m} because μ is a distribution. Now we bound each of the remaining terms. For each non-zero $u \in \{0,1\}^{k-1}$, the corresponding term with u is

$$T_u = \max_{\cup_{u'<u} Y_{u'}} \mathbf{E}_{Y_u} \mu(x|\phi_{AND}(y_1^{u_1}, \dots, y_{k-1}^{u_{k-1}})).$$

Let $Y_u = m - i \geq m - r$. If $i < r^\alpha$, then we can upper bound T_u as

$$T_u \leq \frac{1}{2^{m-i}} < 2^{r^\alpha - m}.$$

Otherwise, $i \geq r^\alpha \geq d^\alpha$. Since μ is as defined, we can then bound T_u by

$$T_u \leq \frac{2^{i^\alpha - i}/\epsilon}{2^{m-i}} \leq \frac{2^{r^\alpha - m}}{\epsilon}.$$

Thus in both cases, $T_u \leq \frac{2^{r^\alpha - m}}{\epsilon}$. Hence the lemma follows. □

Now we are ready to prove the main theorem of this section.

Proof of Theorem 3.1. Apply Lemma 3.2 with $\epsilon = 5/6$ to obtain g and μ. Then follow the approach as outlined in Section 2. What remains is to show that $\mathrm{Cor}_\lambda(\mathcal{F}_k^g, \Pi_k^c) \leq 1/6$. Now we have, by Proposition 2.5, Lemma 3.3, and Proposition 2.7,

$$\begin{aligned}
\mathrm{Cor}_\lambda(\mathcal{F}_k^g, \Pi_k^c))^{2^{k-1}} &\leq 2^{(c+m)\cdot 2^{k-1}} \cdot \mathbf{E}_{y_1^0,\ldots,y_{k-1}^0,y_1^1,\ldots,y_{k-1}^1} H(y_1^0,\ldots,y_{k-1}^0,y_1^1,\ldots,y_{k-1}^1) \\
&\leq 2^{c2^{k-1}} \sum_{j=d}^{(k-1)m} 2^{(2^{k-1}-1)j^\alpha} (\frac{6}{5})^{2^{k-1}-1} \Big(\frac{e(k-1)m}{j(\ell-1)}\Big)^j (1-\frac{1}{\ell})^{(k-1)m}. \qquad (2)
\end{aligned}$$

Since $k \leq (1-\alpha)\log_2 d$, we have $(2^{k-1}-1)j^\alpha < d^{1-\alpha} j^\alpha \leq j$ for $j \geq d$ so (2) is

$$\begin{aligned}
&\leq (\frac{6}{5}2^c)^{2^{k-1}} \sum_{j=d}^{(k-1)m} \Big(\frac{2e(k-1)m}{j(\ell-1)}\Big)^j (1-\frac{1}{\ell})^{(k-1)m} \\
&\leq \frac{(\frac{6}{5}2^c)^{2^{k-1}}}{2^d},
\end{aligned} \qquad (3)$$

for $\ell = \lceil 4e \rceil$. Hence

$$\mathrm{Cor}_\lambda(\mathcal{F}_k^g, \Pi_k^c)) \leq \frac{\frac{6}{5}2^c}{2^{d/2^{k-1}}} \leq 1/6,$$

as long as $c \leq \log_2(\frac{5}{36}2^{d/2^{k-1}})$. Hence $R_{1/3}^k$ is $\Omega(d/2^k)$ for $k \leq (1-\alpha)\log_2 d$. □

4 Proof of Lemma 3.2

Proof. As in the proof for Lemma 2.4, we write the requirements down as a linear program and study its dual. The lemma is implied by proving that the following linear program $\mathcal{P}$ has optimal value 1:

Minimize η subject to

$$y_S : \quad \sum_{x\in\{0,1\}^m} h(x)\chi_S(x) = 0 \qquad |S| < d \tag{4}$$

$$\beta : \quad \sum_{x\in\{0,1\}^m} h(x)f(x) \geq \epsilon \tag{5}$$

$$v_x : \quad \mu(x) - h(x) \geq 0 \qquad x \in \{0,1\}^m \tag{6}$$

$$w_x : \quad \mu(x) + h(x) \geq 0 \qquad x \in \{0,1\}^m \tag{7}$$

$$a_\rho : \quad \eta - 2^{|\rho|-|\rho|^\alpha} \sum_{x\in C_\rho} \mu(x) \geq 0 \qquad \rho \in \{0,1,*\}^m, |\rho| \geq d^\alpha \tag{8}$$

$$\gamma : \quad \sum_{x\in\{0,1\}^m} \mu(x) = 1 \tag{9}$$

Suppose that we have optimum $\eta = 1$. In this LP formulation, inequality γ ensures that the function μ is a probability distribution, and inequalities v_x and w_x ensure that $\mu(x) \geq |h(x)|$ so $||h||_1 \leq 1$. If $||h||_1 = 1$, then we must have $\mu(x) = |h(x)|$ and we can write $h(x) = \mu(x)g(x)$ as in the proof of Lemma 2.4 and then the inequalities y_S will ensure that $\mathrm{Cor}_\mu(g, \chi_S) = 0$ for $|S| < d$ and inequality β will ensure that $\mathrm{Cor}_\mu(f, g) \geq \epsilon$ as required. Finally, each inequality a_ρ ensures that $\mu(C_\rho) \leq 2^{-|\rho|+|\rho|^\alpha} = 2^{-|\rho|+|\rho|^\alpha}$ which is actually a little stronger than our claim.

The only issue is that an optimal solution might have $||h||_1 < 1$. However in this case inequality β ensures that $||h||_1 \geq \epsilon$. Therefore, for any solution of the above LP with function h, we can define another function $h'(x) = h(x)/||h||_1$ with $||h'||_1 = 1$ and a new probability distribution μ' by $\mu'(x) = |h'(x)| \leq \mu(x)/||h||_1 \leq \mu(x)/\epsilon$. This new h' and μ' still satisfy all the inequalities as before except possibly inequality a_ρ but in this case if we increase η by a $1/||h||_1$ factor it will also be satisfied. Therefore, the $\mu'(C_\rho) \leq 2^{-|\rho|+|\rho|^\alpha}/\epsilon$.

Here is the dual LP:

Maximize $\beta \cdot \epsilon + \gamma$ subject to

$$\eta : \quad \sum_{\rho\in\{0,1,*\}^m, |\rho|\geq d^\alpha} a_\rho = 1 \tag{10}$$

$$\mu(x) : \quad v_x + w_x + \gamma - \sum_{C_\rho \ni x, |\rho|\geq d^\alpha} 2^{|\rho|-|\rho|^\alpha} a_\rho = 0 \qquad x \in \{0,1\}^m \tag{11}$$

$$g(x) : \quad \beta f(x) + \sum_{|S|<d} y_S\chi_S(x) + w_x - v_x = 0 \qquad x \in \{0,1\}^m \tag{12}$$

$$\beta, v_x, w_x, a_\rho \geq 0 \qquad x \in \{0,1\}^m \tag{13}$$

Since y_S are arbitrary we can replace $\sum_{|S|<d} y_S\chi_S(x)$ by $p_d(x)$ where p_d is an arbitrary polynomial of degree $< d$ to obtain the modified dual:

Maximize $\beta \cdot \epsilon + \gamma$ subject to

$$\eta : \qquad \sum_{\rho \in \{0,1,*\}^m, |\rho| \geq d^\alpha} a_\rho = 1 \tag{14}$$

$$\mu(x) : \qquad v_x + w_x + \gamma - \sum_{C_\rho \ni x, |\rho| \geq d^\alpha} 2^{|\rho| - |\rho|^\alpha} a_\rho = 0 \qquad x \in \{0,1\}^m \tag{15}$$

$$g(x) : \qquad \beta f(x) + p_d(x) + w_x - v_x = 0 \qquad x \in \{0,1\}^m \tag{16}$$

$$\beta, v_x, w_x, a_\rho \geq 0 \qquad x \in \{0,1\}^m \tag{17}$$

Equations (15) and (16) for $x \in \{0,1\}^m$ together are equivalent to:

$$2w_x + \beta f(x) + p_d(x) + \gamma - \sum_{C_\rho \ni x, |\rho| \geq d^\alpha} 2^{|\rho| - |\rho|^\alpha} a_\rho = 0$$

and

$$2v_x - \beta f(x) - p_d(x) + \gamma - \sum_{C_\rho \ni x, |\rho| \geq d^\alpha} 2^{|\rho| - |\rho|^\alpha} a_\rho = 0.$$

Since these are the only constraints on v_x and w_x respectively other than negativity these can be satisfied by any solution to

$$\beta f(x) + p_d(x) + \gamma \leq \sum_{C_\rho \ni x, |\rho| \geq d^\alpha} 2^{|\rho| - |\rho|^\alpha} a_\rho$$

and

$$-\beta f(x) - p_d(x) + \gamma \leq \sum_{C_\rho \ni x, |\rho| \geq d^\alpha} 2^{|\rho| - |\rho|^\alpha} a_\rho,$$

which together are equivalent to

$$|\beta f(x) + p_d(x)| + \gamma \leq \sum_{C_\rho \ni x, |\rho| \geq d^\alpha} 2^{|\rho| - |\rho|^\alpha} a_\rho.$$

Since $p_d(x)$ is an arbitrary polynomial function of degree less than d we can write $p_d = -\beta p'_d$ where p'_d is another arbitrary polynomial function of degree less than d and we can replace the terms $|\beta f(x) + p_d(x)|$ by $\beta|f(x) - p'_d(x)|$.

Therefore the dual program $\mathcal{D}$ is equivalent to maximizing $\beta \cdot \epsilon + \gamma$ subject to

$$\beta|f(x) - p'_d(x)| + \gamma \leq \sum_{C_\rho \ni x, |\rho| \geq d^\alpha} 2^{|\rho| - |\rho|^\alpha} a_\rho$$

for all $x \in \{0,1\}^m$, a_ρ is probability distribution on the set of all restrictions of size at least d^α, and p'_d is a real-valued function of degree $< d$.

Now, let B be the set of points at which $|f(x) - p'_d(x)| \geq \epsilon$. For any $x \in B$, the value of the objective function of $\mathcal{D}$, which is $\beta \cdot \epsilon + \gamma$, is not more than

$$\beta|f(x) - p'_d(x)| + \gamma \leq \sum_{C_\rho \ni x, |\rho| \geq d^\alpha} 2^{|\rho| - |\rho|^\alpha} a_\rho. \tag{18}$$

Let $R(x)$ denote the right-hand side of inequality (18). It suffices to prove that $R(x) \leq 1$ for some $x \in B$. This is, in turn, equivalent to proving that

$$\min_{x \in B} R(x) \leq 1,$$

for any distribution a_ρ. Suppose, by contradiction, that there exists a distribution a_ρ such that $R(x) > 1$ for any $x \in B$. Let Π, the set of restrictions, and ν, a distribution on Π, be the witnesses for the (ϵ, α)-approximate degree of f. Picking $\pi \in \Pi$ randomly according to ν, we define the random variable

$$I_\pi := \sum_{\rho:|\rho| \geq d^\alpha,\ C_\rho \cap C_\pi \neq \emptyset} 2^{|\rho|-|\rho|^\alpha} a_\rho.$$

Then,

$$\mathbf{E}_{\pi \sim \nu}(I_\pi) = \sum_{\rho:|\rho|\geq d^\alpha} \Pr[C_\rho \cap C_\pi \neq \emptyset] \cdot 2^{|\rho|-|\rho|^\alpha} a_\rho \leq \sum_{\rho:|\rho|\geq d^\alpha} 2^{|\rho|^\alpha - |\rho|} \cdot 2^{|\rho|-|\rho|^\alpha} a_\rho \leq 1.$$

Therefore there exists $\pi \in \Pi$ for which $I_\pi \leq 1$. If there exists $x \in B$ such that $x \in C_\pi$, then since

$$R(x) = \sum_{C_\rho \ni x, |\rho| \geq d^\alpha} 2^{|\rho|-|\rho|^\alpha} a_\rho > 1,$$

we would have $I_\pi > 1$. Thus $C_\pi \cap B = \emptyset$. So for any $x \in C_\pi$, we have $|f(x) - p'_d(x)| \leq \epsilon$. But since the degree of p'_d is less than d this contradicts the fact that $deg_\epsilon(f|_\pi) \geq d$. Thus the lemma follows. □

We note that the bounds in Lemma 3.2 are essentially the best possible for AC^0 functions: By results of Linial, Mansour, and Nisan [LMN89], for any AC^0 function f there is a function p_d of $\log^{O(1)} n$ degree, such that $||f - p_d||_2^2 \leq 2^{n-n^\delta}$ for some constant $\delta > 0$. Let B_n be the set of x such that $|f(x) - p_d(x)| \geq \epsilon$. Then $|B_n|\epsilon^2 \leq \sum_{x \in B_n} |f(x) - p_d(x)|^2 \leq ||f - p_d(x)||_2^2 \leq 2^{n-n^\delta}$ so $|B_n| \leq 2^{n-n^\delta}/\epsilon^2$. Also, if we tried to replace the upper bound on $\mu(C_\rho)$ by some $c(|\rho|)$ where $c(n)$ is $\omega(1/|B_n|)$ then we could choose $a_x = 1/|B_n|$ for $x \in B_n$ and $a_\rho = 0$ for all other ρ and for these values β would be unbounded.

5 Tribes has large (ϵ, α)-approximate degree

It is not obvious that any function, let alone a function in AC^0, has large (ϵ, α)-approximate degree for $\alpha < 1$. Recall that the function $\textsc{Tribes}_{p,q}$ on $m = pq$ bits is defined by

$$\textsc{Tribes}_{p,q}(x) = \vee_{i=1}^q \wedge_{j=1}^p x_{i,j}.$$

Usually the function Tribes is defined so that 2^p is linear or nearly-linear in q. We will show that, with a different relationship in which $q \gg 2^p$ but p is still $\Theta(\log q)$, the (ϵ, α)-approximate degree of $\textsc{Tribes}_{p,q}$ is large.

Lemma 5.1. *Let r, q, p be positive integers with $q > r > p \geq 2$ and let $1 > \alpha > \beta > 0$ be such that $q^\beta \geq rp$, $2^p - 1 \geq q^{1-\beta}$, $q^\alpha \geq \frac{6}{\ln 2} 2^p r$, and $r^{\alpha(\alpha-\beta)} \geq 12(3p/\ln 2)^2$. Then $\textsc{Tribes}_{p,q}$ has $(5/6, \alpha)$-approximate degree at least $\sqrt{r/12}$.*

Proof. We define a distribution ν on restrictions R_m^{pr} that leave pr out of the m variables unset as follows: pick uniformly at random a subset of $q-r$ of the q terms of $\textsc{Tribes}_{p,q}$; then for each of these terms, assign values to the variables in the term uniformly at random from $\{\{0,1\}^p - \mathbf{1}^p\}$. It is clear that for any π with $\nu(\pi) > 0$, $\textsc{Or}_r$ is a subfunction of $\textsc{Tribes}_{p,q}|_\pi$ so $deg_{5/6}(\textsc{Tribes}_{p,q}|_\pi) \geq deg_{5/6}(\textsc{Or}_r) \geq \sqrt{r/12}$.

Let ρ be any restriction of size $i = |\rho| \geq (r/12)^{\alpha/2}$. By definition, we need to prove that

$$\Pr_{\pi\sim\nu}[C_\rho \cap C_\pi \neq \emptyset] \leq 2^{i^\alpha - i}.$$

Now

$$\Pr_{\pi\sim\nu}[C_\rho \cap C_\pi \neq \emptyset] = \frac{1}{\binom{q}{q-r}} \sum_{S\subset[q],|S|=q-r} \Pi_{j\in S}\, p_j,$$

where p_j is the probability that π and ρ agree on the variables in the j-th term in $\textsc{Tribes}_{p,q}$. Write $i = i_1 + \ldots + i_q$, where i_j is the number of assignments ρ makes to variables in the j-th term of $\textsc{Tribes}_{p,q}$. Then

$$p_j \leq \frac{2^{p-i_j}}{2^p - 1} = 2^{-i_j}(1 + \frac{1}{2^p - 1}).$$

Let $i_S = \sum_{j\in S} i_j$ be the number of assignments ρ makes to variables in terms in S and $k_S = |\{j \in S : i_j > 0\}|$ be the number of terms in S in which ρ assigns least one value. Hence,

$$\Pr_{\pi\sim\nu}[C_\rho \cap C_\pi \neq \emptyset] \quad < \quad \frac{1}{\binom{q}{q-r}} \sum_{S\subset[q],|S|=q-r} 2^{-i_S}(1 + \frac{1}{2^p - 1})^{k_S}. \tag{19}$$

Let $k = |\{j : i_j > 0\}|$ be the total number of terms in which ρ assigns at least one value. There are 2 cases: (I) $k \geq q/2$, and (II) $k < q/2$.

Now consider case (I). Thus $i \geq q/2$. In Equation 19, we have $k_S \leq q$ for every S. Thus,

$$\Pr_{\pi\sim\nu}[C_\rho \cap C_\pi \neq \emptyset] \leq \frac{1}{\binom{q}{q-r}} \sum_{S\subset[m],|S|=q-r} 2^{-i_S}(1 + \frac{1}{2^p - 1})^q.$$

It is easy to see that $i_S \geq i - pr$ for every such S. Hence we get

$$\frac{1}{\binom{q}{q-r}} \sum_{S\subset[q],|S|=q-r} 2^{-i_S} \leq 2^{pr-i} \leq 2^{(2i)^\beta - i},$$

since $pr \leq q^\beta \leq (2i)^\beta$ in this case. Thus,

$$\Pr_{\pi\sim\nu}[C_\rho \cap C_\pi \neq \emptyset] \leq 2^{(2i)^\beta - i}(1 + \frac{1}{2^p - 1})^q \leq 2^{(2i)^\beta - i} e^{q^\beta} \leq 2^{2^\beta(1+1/\ln 2) i^\beta - i},$$

since $q^{1-\beta} \leq 2^p - 1$ and $i \geq q/2$. We upper bound the term $2^\beta(1+1/\ln 2)\, i^\beta$ by i^α as follows: Since $i \geq (r/12)^{\alpha/2}$,

$$i^{\alpha-\beta} \geq (r/12)^{\alpha(\alpha-\beta)/2} \geq (r^{\alpha(\alpha-\beta)}/12)^{1/2} \geq 3p/\ln 2 \tag{20}$$

by our assumption in the statement of the lemma. Since $p \geq 2$, we have $i^{\alpha-\beta} > 6 > 2^\beta(1 + 1/\ln 2)$ which is all that we need to derive that $\Pr_{\pi\sim\nu}[C_\rho \cap C_\pi \neq \emptyset] < 2^{i^\alpha - i}$ in case I.

Next, we consider case (II). We must have $k \le p^{1-\beta}(2^p-1)\, i^\beta$, because otherwise

$$i \ge k > p^{1-\beta}(2^p-1) i^\beta \ge p^{1-\beta} q^{1-\beta} i^\beta,$$

which implies $i^{1-\beta} > (pq)^{1-\beta}$ and hence $i > pq = m$ which is impossible. Therefore

$$(1+\frac{1}{2^p-1})^{k_S} \le e^{\frac{k_S}{2^p-1}} \le e^{\frac{k}{2^p-1}} \le e^{p^{1-\beta} i^\beta}.$$

So,

$$\Pr_{\pi\sim\nu}[C_\rho \cap C_\pi \ne \emptyset] < e^{p^{1-\beta} i^\beta} \mathcal{S} \quad \text{where} \quad \mathcal{S} = \frac{1}{\binom{q}{q-r}} \sum_{S\subset[q],|S|=q-r} 2^{-i_S} = E_{S\sim U}[2^{-i_S}].$$

and U is the uniform distribution on subsets of $[q]$ of size $q-r$.

Now we continue by upper bounding $\mathcal{S}$. For the moment let us assume that i is divisible by p. If we view the terms as the bins, and the assigned positions by ρ as balls placed in corresponding bins, then we observe that $\mathcal{S}$ can only increase if we move one ball from a bin A of $x > 0$ balls to another bin B of $y \ge x$ balls. This is because only those i_S with S containing exactly one of these two bins are affected by this move. Then, we can write the contribution of these S's in $\mathcal{S}$ before the move as

$$\mathcal{S}' = \sum_{S\subset[q],\ |S|=q-r,\ S\cap\{A,B\}=1} 2^{-i_S} = \sum_{S'\subset[q]-\{A,B\},\ |S'|=q-r-1} 2^{-i_{S'}}(2^{-x}+2^{-y}),$$

and after the move as

$$\mathcal{S}'' = \sum_{S'\subset[q]-\{A,B\},\ |S'|=q-r-1} 2^{-i_{S'}}(2^{-x+1}+2^{-y-1}).$$

Since $y \ge x$, $\mathcal{S}'' > \mathcal{S}'$.

Hence w.l.o.g. and with the assumption that p divides i, we can assume that the balls are distributed such that every bin is either full, i.e containing p balls, or empty. Hence $k = i/p$ and for any $1 \le j \le q$, either $i_j = 0$ or $i_j = p$.

Claim 5.2. *If i is divisible by p then $\mathcal{S} \le 2^{-i}\ e^{2^{p+1}rk/q}$.*

We first see how the claim suffices to prove the lemma. If i is not divisible by p then we note that $\mathcal{S}$ is a decreasing function of i and apply the claim for the first $i' = p\lfloor i/p\rfloor > i-p$ positions set by ρ to obtain an upper bound of $\mathcal{S} < 2^{p-i} e^{2^{p+1}ri/(pq)}$ that applies for all choices of i. The overall bound we obtain in this case is then

$$\begin{aligned}\Pr_{\pi\sim\nu}[C_\rho \cap C_\pi \ne \emptyset] &< e^{p^{1-\beta} i^\beta} 2^p e^{2^{p+1}ri/(pq)} 2^{-i} \\ &= 2^{i^\beta p^{1-\beta}/\ln 2 + p + 2^{p+1}ri/(pq\ln 2)} 2^{-i}.\end{aligned}$$

We now consider the exponent $i^\beta p^{1-\beta}/\ln 2 + p + 2^{p+1}ri/(pq\ln 2)$ and show that it is at most i^α. For the first term observe that by (20), $i^{\alpha-\beta} \ge 3p/\ln 2$ so $i^\beta p^{1-\beta}/\ln 2 \le i^\alpha/3$. For the second term again by (20) we have $p \le i^{\alpha-\beta}/3 \le i^\alpha/3$. For the last term, since $q^\alpha \ge \frac{6}{\ln 2}2^p r$, we have

$$\frac{2^{p+1}ri}{pq\ln 2} \le \frac{q^\alpha i}{3pq} \le i(pq)^{\alpha-1}/3 \le i^\alpha/3,$$

since $i \leq pq$. Therefore in case II we have $\Pr_{\pi \sim \nu}[C_\rho \cap C_\pi \neq \emptyset] < 2^{i^\alpha - i}$ as required. It only remains to prove the claim.

Proof of Claim: Let $T = \{i_j \mid i_j = p\}$ be the subset of k terms assigned by ρ. Therefore $i_S = |S \cap T| p$ where S is a random set of size $q - r$ and T is a fixed set of size k and both are in $[q]$. We have two subcases: (IIa) when $k \leq r$ and (IIb) when $q/2 \geq k > r$.

If $k \leq r$ then we analyze $\mathcal{S}$ based on the number j of elements of S contained in T. There are $\binom{k}{j}$ choices of elements of T to choose from and $q - r - j$ elements to select from the $q - k$ elements of $\overline{T}$. Therefore

$$\mathcal{S} = \frac{\sum_{j=0}^{k} \binom{r}{j} \binom{q-k}{q-r-j} 2^{-jp}}{\binom{q}{q-r}}.$$

Now since

$$\frac{\binom{q-k}{q-r-j}}{\binom{q}{q-r}} = \frac{(q-k)!(q-r)!r!}{q!(q-r-j)!(r-(k-j))!} < \frac{(q-r)^j r^{k-j}}{(q-k)^k} = \left(\frac{r}{q-k}\right)^k \left(\frac{q-r}{r}\right)^j,$$

we can upper bound $\mathcal{S}$ by

$$\begin{aligned}
\left(\frac{r}{q-k}\right)^k \sum_{j=0}^{k} \binom{k}{j} 2^{-pj} \left(\frac{q-r}{r}\right)^j &= \left(\frac{r}{q-k}\right)^k \left(1 + \frac{q-r}{2^p r}\right)^k \\
&= 2^{-pk} \left(\frac{r}{q-k}\right)^k \left(\frac{2^p r + (q-r)}{r}\right)^k \\
&= 2^{-i} \left(\frac{q + (2^p - 1) r}{q-k}\right)^k \\
&= 2^{-i} \left(1 + \frac{(2^p - 1) r + k}{q-k}\right)^k \\
&\leq 2^{-i} \left(1 + \frac{2^p r}{q-k}\right)^k \\
&\leq 2^{-i} \, e^{2^p rk/(q-k)} \\
&\leq 2^{-i} \, e^{2^{p+1} rk/q}.
\end{aligned}$$

since $k \leq q/2$.

In the case that $r \leq k \leq q/2$ we observe that by symmetry we can equivalently view the expectation $\mathcal{S}$ as the result of an experiment in which the set S of size $q - r$ is chosen first and the set T of size k is chosen uniformly at random. We analyze this case based on the number j of elements of $\overline{S}$ contained in T. There are $\binom{r}{j}$ choices of elements of $\overline{S}$ to choose from and $k - j$ elements to select from the $q - r \geq q/2 \geq k$ elements of S. Therefore

$$\mathcal{S} = \frac{\sum_{j=0}^{r} \binom{r}{j} \binom{q-r}{k-j} 2^{-(k-j)p}}{\binom{q}{k}}.$$

Using the fact that

$$\frac{\binom{q-r}{k-j}}{\binom{q}{k}} = \frac{(q-r)!(q-k)!k!}{q!(k-j)!(q-r-k+j)!} < \frac{(q-k)^{r-j} k^j}{(q-r)^r} = \left(\frac{q-k}{q-r}\right)^r \left(\frac{k}{q-k}\right)^j,$$

we upper bound $\mathcal{S}$ by

$$\begin{aligned}
2^{-pk}\Big(\frac{q-k}{q-r}\Big)^r \sum_{j=0}^{r} \binom{r}{j}\Big(\frac{2^p k}{q-k}\Big)^j &= 2^{-pk}\Big(\frac{q-k}{q-r}\Big)^r \Big(1+\frac{2^p k}{(q-k)}\Big)^r \\
&= 2^{-i}\Big(\frac{q-k}{q-r}\Big)^r \Big(\frac{q+(2^p-1)k}{q-k}\Big)^r \\
&= 2^{-i}\Big(\frac{q+(2^p-1)k}{q-r}\Big)^r \\
&= 2^{-i}\Big(1+\frac{(2^p-1)k+r}{q-r}\Big)^r \\
&\leq 2^{-i}\Big(1+\frac{2^p k}{q-r}\Big)^r \\
&\leq 2^{-i}e^{2^p rk/(q-r)} \\
&\leq 2^{-i}e^{2^{p+1} rk/q}
\end{aligned}$$

since $r \leq q/2$. □

Corollary 5.3. *Given any $1 > \epsilon > 0$. Let q, p be positive integers with $q > p \geq 2$ such that $\lceil q^{1-\beta} \rceil < 2^p \leq \frac{1}{6} q^{\alpha+\epsilon-1} \ln 2$ for some fixed constants $1 > \alpha > \beta > 1-\epsilon$. Then for large enough q, $\text{TRIBES}_{p,q}$ has $(5/6, \alpha)$-approximate degree at least $\sqrt{q^{1-\epsilon}/12}$.*

Proof. We apply Lemma 5.1 with $r := \lfloor q^{1-\epsilon} \rfloor$. All conditions in the statement of the lemma would then be satisfied for q large enough. In particular, for q large enough,

$$q^\beta / r \geq q^{\beta+\epsilon-1} > \log q > p,$$

and

$$r^{\alpha(\alpha-\beta)} = q^{(1-\epsilon)\alpha(\alpha-\beta)} > 12(3\log q/\ln 2)^2 > 12(3p/\ln 2)^2.$$

□

Corollary 5.4. *Let q, p be positive integers with $q > p \geq 2$ such that $\lceil q^{0.2} \rceil < 2^p \leq \frac{1}{6} q^{0.3} \ln 2$. Then for large enough q, $\text{TRIBES}_{p,q}$ has $(5/6, 0.9)$-approximate degree at least $\sqrt{q^{0.6}/12}$.*

Proof. Follows from the last corollary with $\epsilon = 0.4$, $\alpha = 0.9$, and $\beta = 0.8$. □

6 Multiparty communication complexity of AC^0

6.1 A separating function for $\mathsf{NP}^{\mathsf{cc}}_{\mathsf{k}}$ and $\mathsf{BPP}^{\mathsf{cc}}_{\mathsf{k}}$ for $k = O(\log n)$

In this subsection we will show that $\mathcal{F}_k^{\text{TRIBES}_{p,q}}$ separates $\mathsf{NP}^{\mathsf{cc}}_{\mathsf{k}}$ and $\mathsf{BPP}^{\mathsf{cc}}_{\mathsf{k}}$ for $k = O(\log n)$ for some appropriately chosen values of p and q.

Lemma 6.1. *$N^k(\mathcal{F}_k^{\text{TRIBES}_{p,q}})$ is $O(\log q + p \log n)$ for any $k \geq 2$.*

Proof. The lemma is easy to see as follows. The 0-th player (who holds x), guesses one of the q branches and sends this guess to all other players. Then he also broadcasts the positions of all the p bits in that branch. Finally any other player, who can see x and is given the p positions, can compute the output of $\mathcal{F}_k^{\textsc{Tribes}_{p,q}}$. The communication cost is then $O(\log q + p \log n)$ bits. □

Lemma 6.2. *Let $0 < \epsilon < 1$. Let q, p be sufficiently large positive integers with $q > p \geq 2$ such that $\lceil q^{1-\beta} \rceil < 2^p \leq \frac{1}{6} q^{\alpha+\epsilon-1} \ln 2$ for some fixed constants $1 > \alpha > \beta > 1 - \epsilon$. Let $m = pq$ and $n \geq pqs$ for $s = \lceil 4e \rceil^{k-1}$ and $k \leq a \log_2 n$ for some constant $a > 0$ depending only on α. Then for any $\delta > 0$, $R^k_{1/3}(\mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}_s}) \geq R^k_{1/3}(\mathcal{F}_k^{\textsc{Tribes}_{p,q}})$ is $\Omega(q^{(1-\epsilon)/2}/2^k)$, which is $\Omega(n^{(1-\epsilon)/2-\delta}/c^k)$ for $c = 2(4e)^{(1-\epsilon)/2}$.*

Proof. Applying Corollary 5.4, we get that for large enough q, $\textsc{Tribes}_{p,q}$ has $(5/6, \alpha)$-approximate degree of $\Omega(q^{(1-\epsilon)/2})$. Then it follows from Theorem 3.1 that $R^k_{1/3}(\mathcal{F}_k^{\textsc{Tribes}_{p,q}})$ is $\Omega(q^{(1-\epsilon)/2}/2^k)$, when k is $O(\log n)$.

Since $q = m/p$ and p is $O(\log q)$, we get that q is $\Omega(m/(\log m))$. Moreover, $m \geq n/(4e)^k$. Hence $R^k_{1/3}(\mathcal{F}_k^{\textsc{Tribes}_{p,q}})$ is $\Omega(n^{(1-\epsilon)/2-\delta}/c^k)$, for $c = 2(4e)^{(1-\epsilon)/2}$ and any $\delta > 0$. □

In particular, we get the following corollary.

Corollary 6.3. *Let q, p be positive integers with $q > p \geq 2$ such that $\lceil q^{0.2} \rceil < 2^p \leq \frac{1}{6} q^{0.3} \ln 2$. Then for $k \leq a \log_2 n$ for some constant $a > 0$ the following holds. For q large enough and any $\delta > 0$, $R^k_{1/3}(\mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}_s}) \geq R^k_{1/3}(\mathcal{F}_k^{\textsc{Tribes}_{p,q}})$ is $\Omega(q^{0.3}/2^k)$ which is $\Omega(n^{0.3-\delta}/c^k)$ for $c = 2(4e)^{0.3}$.*

Proof. Follows from Lemma 6.2 with $\epsilon = 0.4$, $\alpha = 0.9$, and $\beta = 0.8$. □

Combining Lemma 6.1 and Corollary 6.3 we obtain our desired separation.

Theorem 6.4. *Let q, p be large enough positive integers with $q > p \geq 2$ such that $\lceil q^{0.2} \rceil < 2^p \leq \frac{1}{6} q^{0.3} \ln 2$. Then $\mathcal{F}_k^{\textsc{Tribes}_{p,q}} \in \mathsf{NP}^{\mathsf{cc}}_{\mathsf{k}} - \mathsf{BPP}^{\mathsf{cc}}_{\mathsf{k}}$ for $k \leq a \log n$ for some constant $a > 0$.*

6.2 Lower bound for $\textsc{Disj}_{k,n}$

In this subsection we reduce $\mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}_s}$ to $\textsc{Disj}_{k,n}$ for a suitable value of n to obtain a NOF communication complexity lower bound on $\textsc{Disj}_{k,n}$ for k up to $\Theta(\sqrt{\log n})$ players.

Theorem 6.5. *Given constants $1 > \epsilon, \alpha > 0$ such that $1 > \alpha > 1 - \epsilon$. Then for $k \leq a\sqrt{\log_2 n}$ for some constant $a > 0$ depending only on α, ϵ, $R^k_{1/3}(\textsc{Disj}_{n,k})$ is $\Omega(n^{c/k}/2^k)$, where $c = \frac{1-\epsilon}{7\alpha+7\epsilon-6}$. In particular, $R^k_{1/3}(\textsc{Disj}_{n,k})$ is $\Omega(n^{1/(6k)}/2^k)$.*

Proof. Recall that

$$\textsc{Disj}_{k,n}(x) = \vee_{i=1}^{n} \wedge_{j=1}^{k} x_{i,j}.$$

For any $x \in \{0,1\}^{Nk}$, where $N = pqs$ for integers p, q, and s we rewrite $\mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}_s}$ as

$$\begin{aligned} \mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}_s}(x) &= \vee_{i=1}^{q} \wedge_{j=1}^{p} \vee_{u=1}^{s} \wedge_{v=1}^{k} x_{i,j,u,v} \\ &= \vee_{i=1}^{q} \vee_{I \in [s]^p} \wedge_{j=1}^{p} \wedge_{v=1}^{k} x_{i,j,I(j),v} \end{aligned}$$

by expanding the second "$\wedge$", where $I(j)$ is the j-th index of I. This in turn equals

$$
\begin{aligned}
&= \bigvee_{i=1}^{q} \bigvee_{I\in[s]^p} \bigwedge_{v=1}^{k} \bigwedge_{j=1}^{p} x_{i,j,I(j),v} \\
&= \bigvee_{i=1}^{q} \bigvee_{I\in[s]^p} \bigwedge_{v=1}^{k} y_{i,I,v} \\
&= \bigvee_{i\in[q],\ I\in[s]^p} \bigwedge_{v=1}^{k} y_{i,I,v} \\
&= \textsc{Disj}_{n,k}(y),
\end{aligned}
$$

where the bits of vector $y \in \{0,1\}^{nk}$ for $n = qs^p$, indexed by $i \in [q]$, $I \in [s]^p$, and $v \in [k]$, are given by

$$y_{i,I,v} = \bigwedge_{j=1}^{p} x_{i,j,I(j),v}.$$

Observe that for any two players $v \neq v'$, player v' can compute any value $y_{i,I,v}$. Thus the k players can compute $\mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}_s}$ by executing a NOF randomized communication protocol for $\textsc{Disj}_{n,k}$ on y of length nk, where $n = qs^p$.

Let β be a constant such that $\alpha > \beta > 1 - \epsilon$. Let $q > p \geq 2$ be sufficiently large and satisfy $\lceil q^{1-\beta} \rceil < 2^p \leq \frac{1}{6} q^{\alpha+\epsilon-1} \ln 2$. Let $s = \lceil 4e \rceil^{k-1}$. From Lemma 6.2, we know that for $k \leq a \log_2 q$ for some constant $a > 0$ depending only on α, $R_{1/3}^k(\mathcal{H}_k^{\textsc{Tribes}_{p,q},\textsc{Or}_s})$ is $\Omega(q^{(1-\epsilon)/2}/2^k)$. Observe that since $2^p \leq \frac{1}{6} q^{\alpha+\epsilon-1} \ln 2 \leq q^{\alpha+\epsilon-1}$ we have

$$n \;=\; qs^p < q\lceil 4e \rceil^{kp} < q^{1+(\alpha+\epsilon-1)k \log_2 \lceil 4e \rceil}.$$

Since $\log_2 \lceil 4e \rceil < 3.5$, letting $b = 7(\alpha + \epsilon - 1) + 1$, we get $q > n^{2/(bk)}$. Therefore $R_{1/3}^k(\textsc{Disj}_{n,k})$ is

$$\Omega(n^{\frac{1-\epsilon}{bk}}/2^k).$$

The rest of the theorem follows by taking $\epsilon = 0.4$ and $\alpha = 0.9$. $\square$

6.3 The randomized communication complexity of depth-3 AC^0

In the last two subsections, we showed that there is a depth-4 read-once AC^0 function separating NP_k^{cc} and BPP_k^{cc} for k up to $\Theta(\log n)$, and there is a depth-2 read-once AC^0 function separating NP_k^{cc} and BPP_k^{cc} for k up to $\Theta(\sqrt{\log n})$. In this subsection we show that there is a depth-3 AC^0 function that is hard for randomized NOF communication complexity for k up to $\Theta(\log n)$.

Corollary 6.6. *Let q, p be positive integers with $q > p \geq 2$ such that $\lceil q^{0.2} \rceil < 2^p \leq \frac{1}{6} q^{0.3} \ln 2$, then for q large enough and any $\epsilon > 0$, $R_{1/3}^k(\mathcal{H}_k^{f,t}) \geq R_{1/3}^k(\mathcal{F}_k^f)$ is $\Omega(q^{0.3}/2^k)$ which is $\Omega(n^{0.3-\epsilon}/c^k)$ when k is $O(\log n)$, and $c = 2(4e)^{0.3}$, where f is $\textsc{Tribes}'_{p,q}$, the dual of the $\textsc{Tribes}_{p,q}$ function on $m = pq$ bits, and t is the $\textsc{Or}$ function on s bits and $n = m \cdot s$ for $s = \lceil 4e \rceil^{k-1}$. Moreover,*

$$\mathcal{H}_k^{f,t}(x) = \bigwedge_{i\in[q]} \bigvee_{j\in[p],\ u\in[s]} \bigwedge_{v\in[k]} x_{i,j,u,v}$$

is a depth 3 read-once formula.

Proof. The first part follows directly from the proof of Lemma 6.2. Since the second layer of f can be combined with t into one layer, the second part follows. $\square$

Although we have shown non-trivial lower bounds for $\textsc{Disj}_{k,n}$ for k up to $\Theta(\sqrt{\log n})$ it is open whether one can prove similar lower bounds to Corollary 6.6 for $k = \omega(\sqrt{\log n})$ players for $\textsc{Disj}_{k,n}$

or any other depth-2 AC^0 function. The difficulty of extending our lower bound methods is our inability to apply Lemma 3.2 to OR since the constant function 1 approximates OR on all but one point.

The above function is far from unique. In particular, since it contains the above function as a subfunction, a similar $m^{\Omega(1)}/2^{O(k)}$ lower bound applies to the k-party NOF communication complexity of the extension of the Minsky-Papert function that is the k-party analogue of the function considered by Razborov and Sherstov [RS08], namely $\wedge_{i=1}^{m} \vee_{j=1}^{m^2} \wedge_{v=1}^{k} x_{i,j,v}$.

7 Outside of AC^0

7.1 Generalized Inner Product (GIP)

We can use Theorem 3.1 to prove a $\Omega(n^{1-\epsilon}/c^k)$-lower bound on k-party NOF randomized communication complexity for the GIP function defined on $n \cdot k$ bits as follows:

$$\mathrm{GIP}_{n,k}(x_1, \ldots, x_k) := \oplus_{i=1}^{n}\Big(\wedge_{j=1}^{k} x_{j,i}\Big),$$

for any constant $\epsilon > 0$, where $c > 4$ is some constant. With the lower bound of $\Omega(n/4^k)$ proved in [BNS92], our technique shows that we do not lose too much by applying the pattern matrix method with a simple masking scheme.

Lemma 7.1. *For any* $0 < \gamma < \alpha \leq 1$, $\textsc{Parity}_m$ *has* $(5/6, \alpha)$*-approximation degree of* $\Omega(m^\gamma)$*.*

Proof. We define the witnesses for the $(5/6, \alpha)$-approximate degree of $\textsc{Parity}_m$ to be the set $R_m^{m^\gamma}$ and ν taken to be the uniform distribution on the set. It follows that for any $\pi \sim \nu$, $deg_{5/6}(\textsc{Parity}_m|_\pi) = \Omega(m^\gamma)$. Thus what remains is to show that for any integers $m \geq i \geq \Omega(m^{\gamma\alpha})$, if we pick uniformly at random ρ such that $|\rho| = i$ and $\pi \sim \nu$, then

$$\Pr[C_\rho \cap C_\pi \neq \emptyset] < 2^{i^\alpha - i},$$

for sufficiently large m.

We consider two cases: (1) $i \leq m^\gamma$ and (2) $i > m^\gamma$. These are exactly like the cases in the proof of the claim in Lemma 5.1 with $q = m$, $r = m^\gamma$, $p = 1$ and $k = i$. The bound follows since $2 \cdot 2^{-i} e^{8m^{\gamma-1}i} < 2^{i^\alpha - i}$. □

Corollary 7.2. $R_{1/3}^k(GIP_{n,k})$ *is* $\Omega(n^{1-\epsilon}/c^k)$ *for any constant* $\epsilon > 0$ *and* $c = 8e$*.*

Proof. It is easy to see that $\mathrm{GIP}_{n,k} = \mathcal{H}_k^{f,t}$, where $f = \textsc{Parity}_m$ and $t = \textsc{Parity}_s$ and $n = m \cdot s$. From Lemma 7.1, the $(5/6, \alpha)$-approximate degree of f is $d = \Omega(m^\gamma)$ for any $0 < \gamma < \alpha < 1$. Then by applying Theorem 3.1, since $m = n/(4e)^k$ and for $\alpha > \gamma = 1 - \epsilon$, we have $R_{1/3}^k(\mathcal{H}_k^{f,t}) \geq R_{1/3}^k(F_k^f)$ and the latter is $\Omega(d/2^k)$ which is $\Omega(n^{1-\epsilon}/c^k)$ for $c = 8e$. □

References

[BNS92] L. Babai, N. Nisan, and M. Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *Journal of Computer and System Sciences*, 45(2):204–232, October 1992.

[CA08] Arkadev Chattopadhyay and Anil Ada. Multiparty communication complexity of disjointness. Technical Report TR08-002, Electronic Colloquium in Computation Complexity, http://www.eccc.uni-trier.de/eccc/, 2008.

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

[DPV] Matei David, Toniann Pitassi, and Emanuele Viola. Improved separations between nondeterministic and randomized multiparty communication. Submission to RANDOM 2008.

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

[LMN89] M. Linial, Y. Mansour, and N. Nisan. Constant depth circuits, Fourier transform, and learnability. In *30th Annual Symposium on Foundations of Computer Science*, Research Triangle Park,NC, October 1989.

[LS08] Troy Lee and Adi Shraibman. Disjointness is hard in the multi-party number-on-the-forehead model. Technical Report TR08-003, Electronic Colloquium in Computation Complexity, http://www.eccc.uni-trier.de/eccc/, 2008.

[NS94] N. Nisan and M. Szegedy. On the degree of boolean functions as real polynomials. *Computational Complexity*, 4:301–314, 1994.

[RS08] A. A. Razborov and A. A. Sherstov. The sign-rank of AC^0. Technical Report TR08-016, Electronic Colloquium in Computation Complexity, http://www.eccc.uni-trier.de/eccc/, 2008.

[She07] A. A. Sherstov. Separating AC^0 from depth-2 majority circuits. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, pages 294–301, San Diego, CA, June 2007.

[She08] A. A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, Victoria, BC, May 2008. To appear.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/