Electronic Colloquium on Computational Complexity, Report No. 2 (2008)

# Multiparty Communication Complexity of Disjointness

Arkadev Chattopadhyay and Anil Ada*

School of Computer Science
McGill University, Montreal, Canada
achatt3,aada@cs.mcgill.ca

December 19, 2007

**Abstract**

We extend the 'Generalized Discrepancy' technique suggested by Sherstov [8] to the 'Number on the Forehead' model of multiparty communication. This allows us to prove strong lower bounds of $n^{\Omega(1)}$ on the communication needed by $k$ players to compute the Disjointness function, provided $k$ is a constant. In general, our method yields strong bounds for functions induced by a symmetric predicate if the approximation degree of the predicate is $n^{\Omega(1)}$. Similar bounds have been independently obtained recently by Lee and Shraibman.

In this note, we obtain a lower bound of $n^{\Omega(1)}$ on the $k$-party randomized communication complexity of the Disjointness function in the 'Number on the Forehead' model of multiparty communication for constant $k$. The previous best lower bound for three players until recently was $\Omega(\log n)$. We are told that this has been recently pushed to $n^{\Omega(1)}$ by Lee and Shraibman independent of our work. Our strong bounds follow surprisingly in a simple way from the Approximation/Orthogonality principle in the beautiful recent work of Sherstov [8] and the technique used to prove the Multiparty Degree-Discrepancy Lemma by Chattopadhyay [4].

We obtain our bounds developing a technique that extends the classical discrepancy method to what is called the Generalized Discrepancy Method. This idea also originates in the work of [8] and the earlier work of Razborov

*authors are supported by research grants of Prof. D. Thérien.



ISSN 1433-8092

[7]. It was known that the classical discrepancy method fails to give stronger than $\log n$ lower bounds for Disjointness.

Our strong bounds on Disjointness has consequences for proof complexity (ex. see [10, 3, 2]).

## 1 Extension of the classical discrepancy method

Babai, Nisan and Szegedy introduced the following notion of discrepancy on boolean functions: The key combinatorial object that arises in the study of multiparty communication is a *cylinder-intersection*. A $k$-cylinder in the $i$th dimension is a subset $S$ of $Y_1 \times \cdots \times Y_k$ with the property that membership in $S$ is independent of the $i$th co-ordinate. A set $S$ is called a cylinder-intersection if $S = \cap_{i=1}^{k} S_i$, where $S_i$ is a cylinder in the $i$th dimension. Equivalently, a $k$-cylinder in the $i$th dimension can be viewed as a function $\phi^i : (\{0,1\}^n)^k \to \{0,1\}$ such that $\phi^i(y_1, ..., y_k)$ does not depend on $y_i$. A cylinder intersection is viewed as the product

$$\phi(y_1, ..., y_k) = \phi^1(y_1, ..., y_k)...\phi^k(y_1, ..., y_k).$$

An important measure, defined for a function $f : Y_1 \times ... \times Y_k \to \{-1, 1\}$, is its *discrepancy*. With respect to any probability distribution $\mu$ over $Y_1 \times \cdots \times Y_k$ and cylinder intersection $\phi$, define

$$\mathrm{disc}^{\phi}_{k,\mu}(f) = \left| \Pr_{\mu}\left[f(y_1, \ldots, y_k) = 1 \wedge \phi(y_1, \ldots, y_k) = 1\right] - \Pr_{\mu}\left[f(y_1, \ldots, y_k) = -1 \wedge \phi(y_1, \ldots, y_k) = 1\right.\right. \tag{1}$$

Since $f$ is -1/1 valued, it is not hard to verify that equivalently:

$$\mathrm{disc}^{\phi}_{k,\mu}(f) = \left|\mathbb{E}_{y_1,\ldots,y_k \sim \mu} f(y_1, \ldots, y_k)\phi(y_1, \ldots, y_k)\right| \tag{2}$$

The discrepancy of $f$ w.r.t. $\mu$, denoted by $\mathrm{disc}_{k,\mu}(f)$ is $\max_{\phi}\mathrm{disc}^{\phi}_{k,\mu}(f)$. For removing notational clutter, we will often drop $\mu$ from the subscript when the distribution is clear from the context. Let $R^{\epsilon}_k(f)$ denote the $k$-party randomized communication complexity of $f$ with (two-sided) error probability $\epsilon$. We now state the well-known connection between discrepancy and the randomized communication complexity of a function:

**Theorem 1 (see [1, 5])** *Let $0 < \epsilon < 1/2$ be any real and $k \geq 2$ be any integer. For every function $f : Y_1 \times ... \times Y_k \to \{1, -1\}$ and distribution $\mu$ on inputs from $Y_1 \times \cdots \times Y_k$,*

$$R_k^{1/2-\epsilon}(f) \geq \log\left(\frac{2\epsilon}{disc_{k,\mu}(f)}\right) \tag{3}$$

BNS estimated the discrepancy of functions like $\text{GIP}_k$ w.r.t $k$-wise cylinder intersections and the uniform distribution. These estimates resulted in the first strong lower bounds in the k-party model via Theorem 1. Unfortunately, the applicability of Theorem 1 is limited to those functions that have small discrepancy. Disjointess is a classical example of a function that does not have small discrepancy. For dealing with such functions we will need to generalize the discrepancy method. First, extend the definition of discrepancy provided in (2) to real valued functions. For two functions $f, g : X \to \mathbb{R}$ and a distribution $\mu$ on $\{0, 1\}^n$, define

$$\text{Corr}_\mu(f, g) = \left|\mathbb{E}_{x\sim\mu} f(x)g(x)\right| \tag{4}$$

**Lemma 2 (Generalized Discrepancy Method)** *Denote $X = Y_1 \times ... \times Y_k$. Let $f : X \to \{-1, 1\}$ and $g : X \to \{-1, 1\}$ be such that under some distribution $\mu$ we have $Corr_\mu(f, g) > \epsilon$. Then*

$$R_k^\delta(f) \geq \log\left(\frac{\epsilon - 2\delta}{disc_{k,\mu}(g)}\right) \tag{5}$$

*Proof:* Let $P$ be a $k$-party randomized protocol that computes $f$ with error probability of at most $\delta$ and cost $c$. Then for every distribution $\mu$ over the inputs, we can derive a deterministic $k$-player protocol $P'$ for $f$ that errs only on at most $\delta$ fraction of the inputs (w.r.t $\mu$) and has cost $c$. Take $\mu$ to be a distribution satisfying the correlation inequality. We know $P'$ partitions the input space into at most $2^c$ monochromatic cylinder intersections. Let $\mathcal{C}$ denote this set of cylinder intersections. Then,

$$\begin{aligned}
\epsilon &< \left|\mathbb{E}_{x\sim\mu} f(x)g(x)\right| \\
&= \left|\sum_x f(x)g(x)\mu(x)\right| \\
&\leq \left|\sum_x P'(x)g(x)\mu(x)\right| + \left|\sum_x (f(x) - P'(x))g(x)\mu(x)\right|
\end{aligned}$$

Note that $P'$ is a constant over every cylinder intersection $S$ in $\mathcal{C}$. Therefore,

$$\begin{aligned}\epsilon &< \sum_{S\in\mathcal{C}} \Big|\sum_{x\in S} P'(x)g(x)\mu(x)\Big| + \sum_x \big|g(x)\big|\big|f(x)-P'(x)\big|\mu(x)\\ &\leq \sum_{S\in\mathcal{C}} \Big|\sum_{x\in S} g(x)\mu(x)\Big| + \sum_x \big|f(x)-P'(x)\big|\mu(x)\\ &\leq 2^c \mathrm{disc}_{k,\mu}(g) + 2\delta.\end{aligned}$$

This gives us immediately (5). ∎

## 2 Approximation Degree of boolean functions

We consider the vector space of functions from $\{0,1\}^n \to \mathbb{R}$. Equip this space with the standard inner product $< f, g >$

$$< f, g >= \mathbb{E}_x f(x)g(x) \tag{6}$$

For each $S \subseteq [n]$, define $\chi_S(x) = (-1)^{\sum_{i\in S} x_i}$. Then it is easy to verify that the set of functions $\{\chi_S | S \subseteq [n]\}$ forms an orthonormal basis for this inner product space, and so every $f$ can be expanded in terms of its *Fourier coefficients*

$$f(x) = \sum_{S\subseteq[n]} \hat{f}(S)\chi_S(x) \tag{7}$$

where $\hat{f}(S)$ is defined as $< f, \chi_S >$. This expansion is unique and the *exact degree* of $f$ is defined to be largest $d$ such that there exists $S \subseteq [n]$ with $|S| = d$ and $\hat{f}(S) \neq 0$.

A natural question is how large degree is needed if we want to simply approximate $f$ well. Define the *$\epsilon$-approximate degree of $f$*, denoted by $\deg_\epsilon(f)$ to be the smallest integer $d$ for which there exists a function $\phi$ of exact degree $d$ such that

$$\max_{\{0,1\}^n} \Big| f(x) - \phi(x) \Big| \leq \epsilon$$

Sherstov [8] proves a beautiful result using linear programming duality that we restate in an equivalent form:

**Lemma 3** *Let $f : \{0,1\}^m \to \mathbb{R}$ be given with $deg_\epsilon(f) = d \geq 1$. Then there exists $g : \{0,1\}^m \to \{-1,1\}$ and a distribution $\mu$ on $\{0,1\}^m$ such that:*

$$\mathbb{E}_{x\sim\mu} g(x)\chi_S(x) = 0 \text{ for } |S| < d,$$

$$Corr_\mu(f,g) > \epsilon.$$

The following restatement suits our needs here.

## 3 Using degree to generate a hard problem

Let $f : \{0,1\}^m \to \mathbb{R}$ be any function (called the *base* function) on inputs of length $m$. Let $k \geq 2$ be any integer. We will create a function $F_k$ that takes as input a string $x$ of length $\ell^{k-1}m$ for some suitable $\ell$, and a set of bits that *mask* every bit of $x$ except some $m$ bits that are left unmasked. $F_k$ essentially computes $f$ on the unmasked bits. Here it is convenient to view $x$ as a $k$-dimensional $m \times \ell \times \ell \times \cdots \times \ell$ array. More precisely, define $F_k : X \times ([\ell]^m)^{k-1} \to \mathbb{R}$, where $X = \{0,1\}^{\ell^{k-1}m}$ and $F_k(x, S^1, \ldots, S^{k-1}) = f(x[1, S^1[1], S^2[1], \ldots, S^{k-1}[1]], \ldots, x[m, S^1[m], S^2[m], \ldots, S^{k-1}[m]])$ where the $S^i$ are the masking inputs and we view each as a one dimensional array in $[\ell]^m$. Note that our construction is inspired by the construction of 'pattern matrices' in [8].

It can be easily verified that the proof of the Multiparty Degree-Discrepancy Lemma in [4], that extended the work of [9] yields the following :

**Lemma 4** *Let $f : \{-1,1\}^m \to \{-1,1\}$ have the property that $\mathbb{E}_{x\sim\mu} f(x)\chi_S(x) = 0$, for some probability distribution $\mu$ and all $|S| < d$, where $d \geq 2$ is an integer. Define a probability distribution $\lambda$ on $\{0,1\}^{m\ell^{k-1}} \times \left([\ell]^m\right)^{k-1}$ as $\lambda(x, S^1, \ldots, S^{k-1}) = \frac{\mu_k(x,S^1,\ldots,S^{k-1})}{\ell^{m(k-1)}2^{m\ell^{k-1}-m}}$ where $\mu_k$ is the $k$-wise masked function induced by the base distribution $\mu$.*

$$\left(disc_{k,\lambda}(F_k)\right)^{2^{k-1}} \leq \sum_{j=d}^{m} \binom{(k-1)m}{j} \left(\frac{2^{2^{k-1}-1}}{\ell}\right)^j \tag{8}$$

*Hence, for $\ell \geq 2^{2^k}(k-1)em$ and $d > 2$,*

$$disc_{k,\lambda}(F_k) \leq \frac{1}{2^{d/2^{k-1}}} \tag{9}$$

*Proof:* The starting point is to write the expression for discrepancy w.r.t an arbitrary cylinder intersection $\phi$,

$$\begin{aligned}&\mathrm{disc}_k^{\phi}(F_k) = \\ &\Big| \sum_{x,S^1,\ldots,S^{k-1}} F_k(x,S^1,\ldots,S^{k-1})\phi(x,S^1,\ldots,S^k) \\ &\qquad\qquad \cdot\lambda(x,S^1,\ldots,S^{k-1})\Big| \end{aligned} \tag{10}$$

where $\phi$ is the intersection of $k$ cylinders $\phi_1,\ldots,\phi_k$, and can be expressed as below:

$$\begin{aligned}&\phi(x,S^1,\ldots,S^k) = \\ &\Big(\prod_{i=1}^{k-1} \phi^i(x,S^1,\ldots,S^{k-i-1},S^{k-i+1},\ldots,S^{k-1})\Big) \\ &\quad \times \phi^k(S^1,\ldots,S^{k-1})\end{aligned}$$

This changes to the more convenient expected value notation as follows:

$$\begin{aligned}\mathrm{disc}_k^{\phi}(F_k) = \quad & 2^m\Big|\mathbf{E}_{x,S^1,\ldots,S^{k-1}} F_k(x,S^1,\ldots,S^{k-1}) \\ &\times\phi(x,S^1,\ldots,S^{k-1})\mu_{S^1,\ldots,S^{k-1}}\big(x\big)\Big|\end{aligned} \tag{11}$$

where, as before, $(x,S^1,\ldots,S^{k-1})$ is now uniformly distributed over $\{0,1\}^{m\ell^{k-1}} \times \big([\ell]^m\big)^{k-1}$. Then, we use the trick of repeatedly combining triangle inequality with Cauchy-Schwarz exactly as done in Chattopadhyay[4] to obtain the following:

$$\begin{aligned}&\big(\mathrm{disc}_k^{\phi}(F_k)\big)^{2^{k-1}} \leq \\ &2^{2^{k-1}m}\mathbf{E}_{S_0^1,S_1^1,\ldots,S_0^{k-1},S_1^{k-1}} G_k\big(S_0^1,S_1^1,\ldots,S_0^{k-1},S_1^{k-1}\big)\end{aligned} \tag{12}$$

where,

$$\begin{aligned}&G_k\big(S_0^1,S_1^1,\ldots,S_0^{k-1},S_1^{k-1}\big) \\ &= \Big|\mathbf{E}_{x\in\{0,1\}^{m\ell^{k-1}}} \prod_{u\in\{0,1\}^{k-1}} \Big(F_k(x,S_{u_1}^1,\ldots,S_{u_{k-1}}^{k-1}) \\ &\qquad\qquad \times \mu_{S_{u_1}^1,\ldots,S_{u_{k-1}}^{k-1}}(x)\Big)\Big|\end{aligned} \tag{13}$$

where $\mu_{S^1,\ldots,S^{k-1}}(x) = \mu_k(x, S^1, \ldots, S^{k-1})$. As before we look at a fixed $S_0^i, S_1^i$, for $i = 1, \ldots, k-1$. Let $r = \max\{|S_0^1 \cap S_1^1|, \ldots, |S_0^{k-1} \cap S_1^{k-1}|\}$. We now make two claims.

**Claim 5**

$$G_k\big(S_0^1, S_1^1, \ldots, S_0^{k-1}, S_1^{k-1}\big) \leq \frac{2^{(2^{k-1}-1)r}}{2^{2^{k-1}m}} \tag{14}$$

**Claim 6** *Let $r < d$. Then,*

$$G_k\big(S_0^1, S_1^1, \ldots, S_0^{k-1}, S_1^{k-1}\big) = 0 \tag{15}$$

We leave it to the reader to verify that Claim 5 and Claim 6 made above can be deduced quite easily respectively from Claim 15 and Claim 16 given in the Appendix section of [4]. Applying these two claims we obtain the following:

$$\begin{aligned}
&(\mathrm{disc}_k^{\phi}(F_k))^{2^{k-1}} \\
&\leq \sum_{j=d}^{m} 2^{(2^{k-1}-1)j} \\
&\quad \times \sum_{j_1+\cdots+j_{k-1}=j} \Pr\left[N^1 = j_1 \wedge \cdots \wedge N^{k-1} = j_{k-1}\right]
\end{aligned} \tag{16}$$

where, $N^i = |S_0^i \cap S_1^i|$ for $1 \leq i \leq k-1$.

Substituting the value of the probability, we further obatin:

$$\begin{aligned}
&(\mathrm{disc}_k^{\phi}(F_k))^{2^{k-1}} \\
&\leq \sum_{j=d}^{m} 2^{(2^{k-1}-1)j} \\
&\times \sum_{j_1+\cdots+j_{k-1}=j} \binom{m}{j_1} \cdots \binom{m}{j_{k-1}} \frac{(\ell-1)^{m-j_1} \cdots (\ell-1)^{m-j_{k-1}}}{\ell^{(k-1)m}}
\end{aligned} \tag{17}$$

Applying simple combinatorial identities as in [4], (17) leads to (8), proving the Lemma. ∎

We will now apply the above Lemma, in conjunction with the Generalized Discrepancy Method i.e. Lemma 2, to conclude the following:

**Theorem 7** *Let $f : \{0,1\}^m \to \{1,-1\}$ be such that $deg_\epsilon(f) = d \geq 2$. Consider the k-wise masked function $F_k$. Then, for any $\delta > 0$ and $\ell \geq 2^{2^k}(k-1)em$,*

$$R_k^\delta(F_k) \geq \frac{d}{2^{k-1}} + \log(\epsilon - 2\delta) \tag{18}$$

*Proof:* Applying Lemma 3 we obtain a function $g$ and a distribution $\mu$ such that $\mathrm{Corr}_\mu(f,g) > \epsilon$ and $\mathbb{E}_{x\sim\mu} g(x)\chi_S(x) = 0$ for $|S| < d$. These $g$ and $\mu$ satisfy the conditions of Lemma 4, therefore we have

$$\mathrm{disc}_{k,\lambda}\big(G_k\big) \leq \frac{1}{2^{d/2^{k-1}}}$$

where $\lambda$ is obtained from $\mu$ as stated in Lemma 4.

It can be easily verified that $\mathrm{Corr}_\lambda(F_k, G_k) = \mathrm{Corr}_\mu(f,g) > \epsilon$. Thus, by plugging the value of $\mathrm{disc}_{k,\lambda}\big(G_k\big)$ in (5) of the generalized discrepancy method we get the desired result. ∎

## 4 Approximability of symmetric functions

For any $D : \{0,1,\ldots,n\} \to \{1,-1\}$, define

$$\ell_0(D) \in \{0,1,\ldots,\lfloor n/2 \rfloor\}$$

$$\ell_1(D) \in \{0,1,\ldots,\lceil n/2 \rceil\}$$

such that $D$ is constant over the interval $[\ell_0(D), n - \ell_1(D)]$ and $\ell_0(D)$ and $\ell_1(D)$ are the smallest possible values for which this happens.

Paturi's theorem provides bounds on the approximate degree of symmetric functions.

**Theorem 8 (Paturi[6])** *Let $f : \{0,1\}^n \to \{1,-1\}$ be any symmetric function induced from the predicate $D : \{0,\ldots,n\} \to \{1,-1\}$. Then,*

$$deg_{1/3}(f) = \Theta\big(\sqrt{n(\ell_0(D) + \ell_1(D))}\big) \tag{19}$$

## 5 Disjointness and Other Symmetric Functions

Let $f : \{0,1\}^n \to \{1,-1\}$. Define $F_k^{Com} : (\{0,1\}^n)^k \to \{1,-1\}$ such that $F_k^{Com}(x_1, x_2, \ldots, x_k) = f(z)$ where $z$ is the $n$-bit string obtained from $x_1, ...x_k$ in the following way: $z_i = x_{1i} \wedge x_{2i} \wedge \cdots \wedge x_{ki}$. Here $z_i$ denotes the $i$th bit of $z$ and $x_{ji}$ denotes the $i$th bit of the string $x_j$.

Given $\psi : X_1 \times \cdots \times X_k \to \mathbb{R}$, we associate the $|X_1| \times \cdots \times |X_k|$ $k$-dimensional input matrix $A_\psi$ where $A_\psi[x_1, \ldots, x_k] = \psi(x_1, \ldots, x_k)$.

**Observation 9** *Let $g : \{0,1\}^n \to \{-1,1\}$ be a symmetric function induced from predicate $D : \{0,1,\ldots,n\} \to \{-1,1\}$. Let $f : \{0,1\}^m \to \{-1,1\}$ be another symmetric function such that $f(z) = D(|z|)$. Then, $A_{F_k}$ is a submatrix of $A_{G_k^{Com}}$ when $n = m\ell^{k-1}$.*

Note that if $g$ is the NOR (the complemant of the OR function), then $G_k^{Com}$ is precisely the famous $k$-wise DISJOINTNESS function denoted by $\mathrm{DISJ}_k$.

**Theorem 10**

$$R_k^\delta(DISJ_k) = \Omega\Big(\frac{n^{\frac{1}{2k}}}{2^{2^k} e(k-1)2^{k-1}}\Big)$$

*as long as $\delta < 1/6$.*

*Proof:* We let $g : \{0,1\}^n \to \{-1,1\}$ be the NOR function. As in Observation 9, define $f : \{0,1\}^m \to \{-1,1\}$ using $g$. Thus $f$ is the NOR function on $m$ bits. Then $A_{F_k}$ is a submatrix of $A_{G_k^{Com}}$ for $n = m\ell^{k-1}$. Thus for this setting of parameters, $R_k^\delta(G_k^{Com}) \geq R_k^\delta(F_k)$. By Theorem 7, we know that

$$R_k^\delta(F_k) \geq \frac{\deg_{1/3}(f)}{2^{k-1}} + \log(1/3 - 2\delta) \geq \frac{d}{2^{k-1}}$$

for $\delta < 1/6$ and $\ell \geq 2^{2^k}(k-1)em^2$. We know $\deg_{1/3}(f) = \Omega(\sqrt{m})$ by Theorem 8. Plugging in the values gives the result. ∎

## 6 Conclusion

It is easy to verify that our method allows us to prove strong lower bounds on the $k$-party communication complexity of every function induced by a symmetric predicate whose approximation degree is $n^{\Omega(1)}$.

## References


[1] L. Babai, N. Nisan, and M. Szegedy. Multiparty protocols, pseudorandom generators for Logspace. and time-space trade-offs. *J.Comput.Syst.Sci.*, 45(2):204–232, 1992.

[2] P. Beame, T. Pitassi, and N. Segerlind. Lower bounds for lovsz-schrijver systems and beyond follow from multiparty communication complexity. In *ICALP*, pages 1176–1188, 2005.

[3] P. Beame, T. Pitassi, N. Segerlind, and A. Wigderson. A strong direct product theorem for corruption and the multiparty communication complexity of Disjointness. *Computational Complexity*, 15(4):391–432, 2006.

[4] A. Chattopadhyay. Discrepancy and the power of bottom fan-in in depth-three circuits. In *FOCS*, 2007.

[5] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1996.

[6] R. Paturi. On the degree of polynomials that approximate symmetric boolean functions. In *STOC*, pages 468–474, 1992.

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

[8] A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In *Electronic Colloquium on Computational Complexity*, number TR07-100. 2007.

[9] A. Sherstov. Separating $AC^0$ from depth-2 Majority circuits. In *STOC*, pages 294–301, 2007.

[10] E. Viola and A. Wigderson. One-way multi-party communication lower bound for pointer jumping with applications. In *FOCS*, pages 427–437, 2007.



ECCC ISSN 1433-8092

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