Electronic Colloquium on Computational Complexity, Report No. 60 (2014)

Simplified Lower Bounds on the Multiparty Communication Complexity of Disjointness

Anup Rao* Amir Yehudayoff†

April 20, 2014

Abstract

We show that the deterministic multiparty communication complexity of set disjointness for k parties on a universe of size n is $\Omega(n/4^k)$. We also simplify Sherstov's proof showing an $\Omega(\sqrt{n}/(k2^k))$ lower bound for the randomized communication complexity of set disjointness.

1 Introduction

Given a family of k sets $\mathcal{F} = (X_1, \ldots, X_k)$ over the universe $[n]$, the *disjointness* function is defined as

$$\mathsf{Disjoint}(\mathcal{F}) = \begin{cases} 1 & \text{if } \bigcap_{i=1}^{k} X_i = \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

We study the communication complexity of computing disjointness in the number-on-forehead model [CFL83]. We consider k parties that attempt to compute $\mathsf{Disjoint}(\mathcal{F})$ by exchanging messages about $X_1, \ldots, X_k$, until one of the parties announces the value of $\mathsf{Disjoint}(\mathcal{F})$. The i'th party can see all of the inputs except for X_i, and can send messages that depend on the inputs she sees and all previous messages. All messages are visible to all parties. The communication complexity is the minimum number of bits that should be transmitted to compute $\mathsf{Disjoint}(\mathcal{F})$. In a randomized communication protocol, the parties use shared randomness to pick a deterministic communication protocol, and then run the chosen deterministic protocol. The protocol computes $\mathsf{Disjoint}(\mathcal{F})$ correctly if it outputs $\mathsf{Disjoint}(\mathcal{F})$ with probability at least $2/3$, for every family $\mathcal{F}$. For formal definitions of multiparty communication complexity and its significance, we refer the interested reader to [KN97].

Grolmusz [Gro94] gave a beautiful deterministic protocol showing that $\mathsf{Disjoint}(\mathcal{F})$ can be computed deterministically with communication $O(\log^2(n) + k^2 n/2^k)$. This paper is about proving lower bounds on the communication complexity.

*Computer Science and Engineering, University of Washington, anuprao@cs.washington.edu. Supported by an Alfred P. Sloan Fellowship, the National Science Foundation under agreement CCF-1016565, an NSF Career award, and by the Binational Science Foundation under agreement 2010089.

†Department of Mathematics, Technion-IIT, Haifa, Israel, amir.yehudayoff@gmail.com. Horev Fellow – supported by the Taub Foundation. Supported by the Israel Science Foundation and by the Binational Science Foundation under agreement 2010089.

ISSN 1433-8092

1.1 Motivation and related work

Lower bounds on multiparty communication complexity are important because several computational models such as circuits, branching programs, and propositional proofs can be used to obtain efficient communication protocols. Strong enough communication complexity lower bounds for the computation of any explicit function can therefore be used to prove lower bounds on these models [BNS92, CT93, BHK01, Raz00, VW08]. In particular, lower bounds on the complexity of computing disjointness imply lower bounds on proof systems [BPS07], circuit lower bounds [HG91, RW93, NW93, Vio07], lower bounds on communication for problems related to combinatorial auctions [CS04, NS06, Nis02, DN11, HM07, PSS08], and give oracle separations for complexity classes [AW09].

Attempts to prove lower bounds on the communication complexity of disjointness have led to many interesting ideas. When the number of parties is $k = 2$, Kalyanasundaram and Schnitger [KS92] proved that $\Omega(n)$ communication is required in the randomized setting. Alternate proofs and tight bounds have since been obtained [Raz92, BYJKS04, BM13] using methods involving information theory. These methods have found many other applications that we do not discuss here.

When k is large, Tesson [Tes03] and Beame, Pitassi, Segerlind and Wigderson [BPSW06] proved that the deterministic communication complexity is $\Omega(\log(n)/k)$. Then Sherstov [She09, She11] introduced the *pattern matrix method* for proving lower bounds in the case $k = 2$. The method was used to separate certain circuit classes by relating their complexity to analytic properties of boolean functions, like their approximate degree. This technique was generalized to $k > 2$ by Chattopadhyay [Cha07], Lee and Shraibman [LS09], and Chattopadhyay and Ada [CA08]. These last two papers proved lower bounds of the type $\Omega\left(n^{1/(k+1)}/2^{2^{O(k)}}\right)$ on the randomized communication complexity. Beame and Huynh-Ngoc [BHN09] extended these methods further to prove that the randomized communication complexity is at least $2^{\Omega\left(\sqrt{\log(n)/k}\right)}2^{-k}$. Finally, Sherstov [She12, She13] proved the best known lower bounds prior to our work, showing that the randomized communication complexity is at least $\Omega(\sqrt{n}/(k2^k))$. In fact, Sherstov proved lower bounds for a much broader class of functions, as we discuss below.

These results use powerful techniques such as Fourier analysis, Gowers norms, directional derivatives, and bounds on the approximate degree. The last two works of Sherstov are the main inspiration for our work.

1.2 Results

In what follows, k is the number of players in the number-on-forehead model, and n is the size of the universe. Our work follows the ideas in the recent papers of Sherstov [She12, She13]. We prove a linear lower bound on the multiparty communication complexity of disjointness:

Theorem 1. *The deterministic communication complexity of disjointness is* $\Omega\left(\frac{n}{4^k}\right)$.

Given our interpretation of Sherstov's work in [She12], the proof of Theorem 1 is short. We also simplify the proof of the randomized lower bound from [She13]:

Theorem 2 ([She13])**.** *The randomized communication complexity of disjointness is* $\Omega\left(\frac{\sqrt{n}}{k2^k}\right)$.

Sherstov proved lower bounds for functions of the type $f(\mathsf{Disjoint}(\mathcal{F}_1), \ldots, \mathsf{Disjoint}(\mathcal{F}_m))$, where f is an arbitrary multivariate function, and $\mathcal{F}_1, \ldots, \mathcal{F}_m$ are families on disjoint parts of the universe. In our proof, we focus on the case where f is the AND function, which corresponds to computing disjointness on the whole universe. Our proof is then obtained by taking his proof and *symmetrizing* it, by viewing f as a univariate function $f : \{0, 1, \ldots, m\} \to \{0, 1\}$, rather than as a multivariate function.

A key part of Sherstov's proof of Theorem 2 is a method to control the error in an approximation of f. This corresponds to a bound on the error in an approximation of the Kronecker delta function (i.e. the univariate function f that is the indicator of m) in the symmetrized proof. We bound the error via the following theorem, which shows that if a polynomial is not correlated with any parity, then it has a low-degree approximation:

Theorem 3. *Let m be a power of 2. For $j \in [m]$, let J denote the smallest power of 2 such that $J \geq j$. Let $Y_1, \ldots, Y_m \in \{0, 1\}$ be distributed uniformly and independently. Suppose f is a real univariate polynomial of degree at most m such that for every $j \geq d > 0$,*

$$\left|\mathbb{E}\left[f((Y_1 + \ldots + Y_j)m/J) \cdot (-1)^{Y_1+\ldots+Y_j}\right]\right| \leq 2^{-15J}, \tag{1}$$

then there exists a polynomial g of degree at most $d - 1$ such that $|g(x) - f(x)| \leq 2^{-3d}$ for all $x \in [0, m]$.

The proof of Theorem 3 relies on the choice of a useful basis $b_0(x), \ldots, b_m(x)$ for the space of polynomials, where each b_i is of degree i. Given this basis, the polynomial g is just the projection of f to the space spanned by $b_0, \ldots, b_{d-1}$. This basis may be of independent interest. For the analogous part of the proof, Sherstov finds a low-degree approximation of f using a different basis. We prove Theorem 3 in Section 3.

2 The lower bounds

Without loss of generality, we assume that $n = m\ell$, where m is a power of 2 and ℓ is a function of k to be determined. Any family $\mathcal{F} = (X_1, \ldots, X_k)$ can be described using the m families $\mathcal{F}_1, \ldots, \mathcal{F}_m$, each over a universe of size ℓ, defined as

$$\mathcal{F}_i = (X_1 \cap [(i-1)\ell + 1, i\ell], \ldots, X_k \cap [(i-1)\ell + 1, i\ell]).$$

Moreover,

$$\mathsf{Disjoint}(\mathcal{F}) = \begin{cases} 1 & \text{if } \sum_{i=1}^{m} \mathsf{Disjoint}(\mathcal{F}_i) = m, \\ 0 & \text{otherwise.} \end{cases}$$

In order to prove Theorems 1 and 2, we consider distributions on families $\mathcal{F}$, where each $\mathcal{F}_i$ is independent and identically distributed. Sherstov shows that there are distributions of this type under which every protocol with small communication complexity must have low correlation with $(-1)^{\sum_{i=1}^{m} \mathsf{Disjoint}(\mathcal{F}_i)}$. Consider the distribution μ given in Figure 1 as a way to sample each $\mathcal{F}_i$. The following theorem is an easy consequence of Theorem 4.2 in [She12], and the fact that every communication protocol can be expressed as a sum of cylinder intersections:

Distribution μ on $\mathcal{F}_1 \subseteq 2^{[\ell]}$
Let $S_1, \ldots, S_{k-1} \subseteq [\ell]$ be uniformly random sets conditioned on $\lvert S_1 \cap S_2 \cap \ldots \cap S_{k-1} \rvert = 1$. Let $S_k \subseteq [\ell]$ be uniform and independent. Set $$\mathcal{F}_1 = (S_1, \ldots, S_{k-1}, S_k).$$

Figure 1: The distribution μ

Theorem 4 ([She13])**.** *If each family $\mathcal{F}_i$ is sampled independently according to μ, and π is a k party protocol with communication complexity C, then*

$$\left| \mathbb{E}\left[\pi(\mathcal{F}) \cdot (-1)^{\sum_{i=1}^{m} \mathsf{Disjoint}(\mathcal{F}_i)} \right] \right| \leq 2^C \cdot \left(\frac{2^k - 1}{\sqrt{\ell}} \right)^m .$$

The proof of Theorem 4 involves ideas analogous to [BNS92] and some subtle reasoning about the distribution μ. Given Theorem 4, Theorem 1 easily follows:

Proof of Theorem 1. Let π be a deterministic protocol that computes $\mathsf{Disjoint}(\mathcal{F})$ with communication complexity C. When $\mathsf{Disjoint}(\mathcal{F}) = 1$, we have $(-1)^{\sum_{i=1}^{m} \mathsf{Disjoint}(\mathcal{F}_i)} = (-1)^m$ and $\pi(\mathcal{F}) = 1$. On the other hand, when $\mathsf{Disjoint}(\mathcal{F}) = 0$, we have $\pi(\mathcal{F}) = 0$. In addition, $\Pr[\mathsf{Disjoint}(\mathcal{F}_i) = 1] = 1/2$ for all $i \in [m]$, which implies $\Pr[\mathsf{Disjoint}(\mathcal{F}) = 1] = 2^{-m}$. Thus,

$$\left| \mathbb{E}\left[\pi(\mathcal{F}) \cdot (-1)^{\sum_{i=1}^{m} \mathsf{Disjoint}(\mathcal{F}_i)} \right] \right| = 2^{-m}. \tag{2}$$

Now set $\ell = 16(2^k - 1)^2$. Theorem 4 and (2) imply that

$$2^C \cdot ((2^k - 1)/\sqrt{\ell})^m \geq 2^{-m} \;\Rightarrow\; C \geq m = \Omega\left(\frac{n}{4^k} \right).$$

□

The proof of Theorem 1 does not give anything meaningful in the randomized setting, since it may be the case that the protocol is not correlated with $(-1)^{\sum_{i=1}^{m} \mathsf{Disjoint}(\mathcal{F}_i)}$. To prove lower bounds on the randomized communication, following Sherstov, we use a more complicated distribution on inputs, as well as approximation theory.

For the rest of this section, we work with the distribution γ described in Figure 2. A crucial feature of this distribution is that if $\rho : [\ell] \to [\ell]$ is a uniformly random permutation independent of $\mathcal{F}_1$, then the families $(\mathcal{F}_1, \rho(\mathcal{F}_1))$ have the same joint distribution as two independent samples $(\mathcal{F}_1, \mathcal{F}_1')$ from γ conditioned on $\mathsf{Disjoint}(\mathcal{F}_1) = \mathsf{Disjoint}(\mathcal{F}_1')$. Here by $\rho(\mathcal{F}_1)$, we mean the family obtained by permuting the underlying universe. In analogy with Theorem 4, Sherstov shows (Corollary 4.19 in [She13]) that no protocol can be significantly correlated with $(-1)^{\sum_{i=1}^{m} \mathsf{Disjoint}(\mathcal{F}_i)}$ under the distribution γ:

Distribution γ on $\mathcal{F}_1 \subseteq 2^{[\ell]}$

Let t be such that $\ell = 2^{k-1} + t(2^k - 1)$. Let $b \in \{0,1\}$ be uniformly random. Define

$$A = A(b) = \left\{(0, v) : v \in \{0,1\}^k, v_1 + v_2 + \ldots + v_k = b \mod 2\right\},$$

and

$$W = \{(r, v) \in [t] \times \{0,1\}^k : v \neq 1^k\}.$$

Define the k sets $S_1, S_2, \ldots, S_k$ as

$$S_j = \{(r, v) \in A \cup W : v_j = 1\}.$$

Sample a bijection[a] $\sigma : A \cup W \to [\ell]$ uniformly at random and independently of b, and set

$$\mathcal{F}_1 = (\sigma(S_1), \sigma(S_2), \sigma(S_3), \ldots, \sigma(S_k)).$$

[a] Observe $|A \cup W| = \ell$.

Figure 2: The distribution γ

Theorem 5 ([She13])**.** *If each family $\mathcal{F}_i$ is sampled independently according to γ, and π is a protocol with communication complexity C, then*

$$\left|\mathbb{E}\left[\pi(\mathcal{F}) \cdot (-1)^{\sum_{i=1}^m \mathsf{Disjoint}(\mathcal{F}_i)}\right]\right| \leq 2^C \cdot \left(\frac{c_0 k^2 4^k}{\ell}\right)^{m/4},$$

where $c_0 > 0$ is a universal constant.

The proof of Theorem 5 is delicate, especially if one wishes to optimize the dependence on k. The symmetric structure of the distribution γ is very useful, and we shall exploit it next. Given any protocol π computing $\mathsf{Disjoint}(\mathcal{F})$, define f_π as the degree m polynomial so that for all $j \in \{0, 1, \ldots, m\}$,

$$f_\pi(j) = \Pr\left[\pi(\mathcal{F}) = 1 \middle| \sum_{i=1}^m \mathsf{Disjoint}(\mathcal{F}_i) = j\right]. \tag{3}$$

Since the protocol computes $\mathsf{Disjoint}(\mathcal{F})$ with probability at least $2/3$, we have that $|f_\pi(j)| \leq 1/3$, for $j = 0, 1, \ldots, m-1$, and $|1 - f_\pi(m)| \leq 1/3$. The following well known theorem [EZ64, RC66, NS94] shows that any such function must have degree $\sqrt{m}/3$:

Theorem 6 ([EZ64, RC66, NS94])**.** *Let $\epsilon \in (0, 1/2)$. If $f : \{0, 1, \ldots, m\} \to \mathbb{R}$ is a polynomial such that $|f(j)| \leq \epsilon$ for $j = 0, 1, \ldots, m-1$, and $|1 - f(m)| \leq \epsilon$, then the degree of f is at least $\sqrt{m(1-2\epsilon)}/3$.*

Theorem 6 is proved via a clever reduction to Markov's bound on the magnitude of derivatives in bounded polynomials. We give the short proof in Appendix A. We shall prove that if the

Protocol $\tau_{\pi,j}(\mathcal{F}_1, \ldots, \mathcal{F}_j)$

1. Let J denote the smallest power of 2 such that $J \geq j$. Note that m/J is an integer.
2. Using public randomness, sample $J - j$ families $\mathcal{F}_{j+1}, \mathcal{F}_{j+2}, \ldots, \mathcal{F}_J$ according to γ, conditioned on the event that $\mathsf{Disjoint}(\mathcal{F}_{j+1}) = \mathsf{Disjoint}(\mathcal{F}_{j+2}) = \mathsf{Disjoint}(\mathcal{F}_J) = 0$.
3. Let $\mathcal{G} = (\mathcal{F}_1, \ldots, \mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_2, \ldots, \mathcal{F}_J, \ldots, \mathcal{F}_J)$ be the m families obtained by repeating each family $\mathcal{F}_i$ exactly m/J times.
4. Let $\rho_1, \rho_2, \ldots, \rho_m : [\ell] \to [\ell], \eta : [m] \to [m]$ be independent uniformly random permutations chosen using public randomness.
5. Output $\pi(\rho_1(\mathcal{G}_{\eta(1)}), \rho_2(\mathcal{G}_{\eta(2)}), \ldots, \rho_m(\mathcal{G}_{\eta(m)}))$.

Figure 3: The protocol $\tau_{\pi,j}$

communication of π is of order $\sqrt{n}/(k2^k)$, then Theorem 5 implies that f_π can be approximated by a polynomial whose degree is of order $\sqrt{m}$.

Now we show how to use Theorem 3 to complete the proof of the lower bound. For each $j \in [m]$, define the protocol $\tau_{\pi,j}$ as in Figure 3. The symmetric structure of γ implies the following lemma:

Lemma 7. *For every fixed input $\mathcal{F}_1, \ldots, \mathcal{F}_j$,*

$$\Pr[\tau_{\pi,j}(\mathcal{F}_1, \ldots, \mathcal{F}_j) = 1] = f_\pi\left(\left(\sum_{i=1}^{j} \mathsf{Disjoint}(\mathcal{F}_i)\right) m/J\right).$$

Lemma 7 and Theorem 5 place many restrictions on the function f_π. The restrictions are captured by the following lemma, which shows that the correlation of f_π with parity is small.

Lemma 8. *Let J be the smallest power of 2 so that $J \geq j$. If the communication complexity of π is C, and $Y_1, \ldots, Y_j \in \{0,1\}$ are uniformly random and independent, then*

$$\mathbb{E}\left[f_\pi((Y_1 + \ldots + Y_j)m/J) \cdot (-1)^{Y_1 + \ldots + Y_j}\right] \leq 2^C \cdot \left(\frac{c_0 k^2 4^k}{\ell}\right)^{j/4}.$$

Proof. If $\mathcal{F}_1$ is distributed according to γ, then $\mathsf{Disjoint}(\mathcal{F}_1)$ is a uniformly random bit. Thus,

$$\begin{aligned}
&\mathbb{E}\left[f_\pi((Y_1 + \ldots + Y_j)m/J) \cdot (-1)^{Y_1 + \ldots + Y_j}\right] \\
&= \mathbb{E}\left[f_\pi\left(\left(\sum_{i=1}^{j} \mathsf{Disjoint}(\mathcal{F}_i)\right) m/J\right) \cdot (-1)^{\sum_{i=1}^{j} \mathsf{Disjoint}(\mathcal{F}_i)}\right]
\end{aligned}$$

Using Lemma 7,

$$= \mathbb{E}\left[\tau_{\pi,j}(\mathcal{F}_1, \ldots, \mathcal{F}_j) \cdot (-1)^{\sum_{i=1}^{j} \mathsf{Disjoint}(\mathcal{F}_i)}\right] \leq 2^C \cdot \left(\frac{c_0 k^2 4^k}{\ell}\right)^{j/4},$$

where the last inequality is by Theorem 5, since the communication complexity of $\tau_{\pi,j}$ is equal to that of π. □

Given Lemma 8, the proof is completed as follows:

Proof of Theorem 2. We set $\ell = 2^{16 \cdot 4} c_0 k^2 4^k$, so that the right hand side of Lemma 8 is 2^{C-16j}. Fix any randomized protocol π that computes $\mathsf{Disjoint}(\mathcal{F})$ on the distribution induced by γ with C bits of communication. Let f_π be as defined in (3).

We see that f_π satisfies the hypothesis of Theorem 3, with $d = C$. Thus we conclude that there is a degree $C-1$ polynomial g that agrees with f_π up to an error of 2^{-3C}. Theorem 6 implies that

$$C \geq \sqrt{m(1 - 2(1/3 + 2^{-3C}))/3} = \Omega(\sqrt{n}/(k2^k)).$$

□

3 Approximating functions that are not correlated with parity

Here we prove Theorem 3, which shows that if a polynomial $f(j)$ has low correlation with parity, then it can be approximated by a low degree polynomial. In what follows, let $Y_1, \ldots, Y_m \in \{0,1\}$ be independent and uniformly random bits, and let I, J be the smallest powers of 2 such that $I \geq i$ and $J \geq j$.

To prove the theorem, we define a useful basis for the space of polynomials. Let $b_0(x) = 1$. For $i > 0$, let

$$b_i(x) = 2^i \binom{xI/m}{i} = \frac{2^i x(x - m/I)(x - 2m/I) \ldots (x - (i-1)m/I)}{i! \cdot (m/I)^i}.$$

Since b_i is of degree i, the polynomials $b_0, \ldots, b_m$ form a basis for the space of polynomials of degree at most m. To prove Theorem 3, we express f in this basis and then argue that all coefficients corresponding to high degree terms are negligible. The polynomials in our basis can be bounded by the following lemma:

Lemma 9. *For every $i \in \{0, 1, \ldots, m\}$, $\max_{x \in [0,m]} |b_i(x)| \leq 8^i$.*

Proof. We show that the maximum of b_i is attained when $x = m$, and so

$$\max_{x \in [0,m]} |b_i(x)| = |b_i(m)| = 2^i \binom{mI/m}{i} \leq 2^i \cdot 2^I \leq 8^i.$$

Note that the magnitude of b_i is symmetric around the point $(i-1)m/(2I)$,

$$|b_i(x + (i-1)m/(2I))| = |b_i(-x + (i-1)m/(2I))|.$$

So the maximum is attained with $x \in [(i-1)m/(2I), m]$. For any such x that is not a root of b_i,

$$\left| \frac{b_i(x + m/I)}{b_i(x)} \right| = \left| \frac{x + m/I}{x - (i-1)m/I} \right| \geq \left| \frac{x + m/I}{x} \right| > 1,$$

proving that the maximum is attained with $x \in [m - m/I, m]$. For such x, every term $(x - jm/I)$ with $j \in \{0, 1, \ldots, i-1\}$ in $b_i(x)$ is non-negative, and so the maximum is attained when $x = m$. □

The basis polynomials behave nicely under the random experiments from (1):

Lemma 10. *For all $i \in \{0, 1, 2, \ldots, m\}$ and $j \in [m]$,*

$$\left|\mathbb{E}\left[b_i((Y_1 + \ldots + Y_j)m/J) \cdot (-1)^{Y_1+\ldots+Y_j}\right]\right| \begin{cases} = 0 & \text{if } i < j, \\ = 1 & \text{if } i = j, \\ = 0 & \text{if } j < i \leq J, \\ \leq 8^i & \text{if } J < i. \end{cases}$$

Proof. When $i < j$, the polynomial $b_i(y_1 + \ldots + y_j)$ has degree i in the variables $y_1, \ldots, y_j$. Since every monomial must exclude one of the j variables, the contribution of each of the monomials to the expectation is 0. When $i = j$, $b_i((y_1 + \ldots + y_i)m/I) = 2^i\binom{y_1+\ldots+y_i}{i}$ is non-zero only when $y_1 = y_2 = \ldots = y_i = 1$. Thus the expectation is $2^{-i} \cdot 2^i\binom{i}{i} = 1$ in this case. When $j < i \leq J$, we have $I = J$. Since for $r \in [i-1]$, $b_i(rm/I) = 2^i\binom{r}{i} = 0$, the expectation is 0. When $i > J$, by Lemma 9, the expectation is at most 8^i. □

Proof of Theorem 3. Write $f(x) = \sum_{j=0}^{m} a_j b_j(x)$, and let $g(x)$ be the degree $d-1$ polynomial $g(x) = \sum_{j=0}^{d-1} a_j b_j(x)$. To prove the theorem, we show that $|g(x) - f(x)| \leq 8^{-d}$ for all $x \in [0, m]$. Lemma 10 and (1) imply that for $j = d, \ldots, m$,

$$|a_j| - \sum_{i=J+1}^{m} 8^i |a_i| \leq |\,\mathbb{E}\left[f((Y_1 + \ldots + Y_j)m/J) \cdot (-1)^{Y_1+\ldots+Y_j}\right]\,| \leq 8^{-5J}$$

$$\Rightarrow |a_j| \leq 8^{-5J} + \sum_{i=J+1}^{m} 8^i |a_i|. \tag{4}$$

We now prove by induction that for $j = m, m-1, \ldots, d$,

$$\sum_{t=j}^{J} |a_t| \leq 8^{-3J}. \tag{5}$$

When $m/2 < j \leq m$, (4) implies

$$\sum_{t=j}^{m} |a_t| \leq (m/2)8^{-5m} \leq 8^{-3m}.$$

In the general case (4) implies

$$(2/J)\sum_{t=j}^{J} |a_t| \leq 8^{-5J} + \sum_{t=J+1}^{m} 8^t |a_t| \leq 8^{-5J} + \sum_{r=\log(J)+1}^{\log(m)} 8^{2^r} \sum_{t=1+2^{r-1}}^{2^r} |a_t|$$

Applying the induction hypothesis, we get

$$\leq 8^{-5J} + \sum_{r=\log(J)+1}^{\log(m)} 8^{2^r} 8^{-3\cdot 2^r} \leq 8^{-5J} + 8^{-4J} \sum_{q=0}^{\infty} 8^{-q} \leq 8^{-3J}(2/J),$$

which proves the general case of (5).

Finally, for every $x \in [0, m]$, Lemma 9 and (5) imply

$$|g(x) - f(x)| \leq \sum_{j=d}^{m} |a_j||b_j(x)| \leq \sum_{r=\lceil \log d \rceil}^{\log m} 8^{-3 \cdot 2^r} \cdot 8^{2^r} \leq 8^{-2d} \sum_{q=0}^{\infty} 8^{-2q} \leq 8^{-d}.$$

□

Acknowledgements

We thank Paul Beame, Pavel Hrubeš, and Alexander Sherstov for useful discussions.

A Approximation theory

The proof relies on a fundamental theorem of Markov, relating the degree of a bounded polynomial to the maximum value of its derivative.

Theorem 11 (Markov's Theorem [Che66])**.** *Let $g : [-1, 1] \to [-1, 1]$ be computed by a polynomial of degree d. Then $|g'(y)| \leq d^2$ for every $y \in [-1, 1]$.*

Markov's theorem allows us to prove the statement about approximation that we need:

Proof of Theorem 6. Let d be the degree of f and let $D = \max_{x \in [0,m]} |f'(y)|$. We can bound f using D as follows. The value $|f(j)|$ is at most $1 + \epsilon$ for $j \in \{0, 1, \ldots, m\}$, and so $|f(x)| \leq 1 + \epsilon + D/2$ for $x \in [0, m]$. On the other hand, $D \geq \frac{f(m)-f(m-1)}{1} \geq 1 - 2\epsilon$.

Now consider the degree d polynomial $g : [-1, 1] \to [-1, 1]$ given by $g(y) = \frac{f(my/2+m/2)}{1+\epsilon+D/2}$. Since $g'(y) = \frac{(m/2)f'(my/2+m/2)}{1+\epsilon+D/2}$, there is a $y \in [-1, 1]$ such that $|g'(y)| = \frac{Dm/2}{1+\epsilon+D/2}$. By Theorem 11,

$$d^2 \geq \frac{Dm/2}{1 + \epsilon + D/2} \geq \frac{m(1/2 - \epsilon)}{1 + \epsilon + 1/2 - \epsilon} = \frac{2m(1/2 - \epsilon)}{3} \;\Rightarrow\; d \geq \sqrt{m(1 - 2\epsilon)/3}.$$

□

References

[AW09] Scott Aaronson and Avi Wigderson. Algebrization: A new barrier in complexity theory. *TOCT*, 1(1), 2009.

[BHK01] László Babai, Thomas P. Hayes, and Peter G. Kimmel. The cost of the missing bit: Communication complexity with help. *Combinatorica*, 21(4):455–488, 2001.

[BHN09] Paul Beame and Dang-Trinh Huynh-Ngoc. Multiparty communication complexity and threshold circuit size of AC^0. In *FOCS*, pages 53–62. IEEE Computer Society, 2009.

[BM13] Mark Braverman and Ankur Moitra. An information complexity approach to extended formulations. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *STOC*, pages 161–170. ACM, 2013.

[BNS92] László Babai, Noam Nisan, and Mario Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. Syst. Sci.*, 45(2):204–232, 1992.

[BPS07] Paul Beame, Toniann Pitassi, and Nathan Segerlind. Lower bounds for Lovász–Schrijver systems and beyond follow from multiparty communication complexity. *SIAM Journal on Computing*, 37(3):845–869, 2007.

[BPSW06] Paul Beame, Toniann Pitassi, Nathan Segerlind, and Avi Wigderson. A strong direct product theorem for corruption and the multiparty communication complexity of disjointness. *Computational Complexity*, 15(4):391–432, 2006.

[BYJKS04] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *Journal of Computer and System Sciences*, 68(4):702–732, 2004.

[CA08] Arkadev Chattopadhyay and Anil Ada. Multiparty communication complexity of disjointness. *CoRR*, abs/0801.3624, 2008.

[CFL83] Ashok K. Chandra, Merrick L. Furst, and Richard J. Lipton. Multi-party protocols. In *STOC*, pages 94–99, 1983.

[Cha07] A. Chattopadhyay. Discrepancy and the power of bottom fan-in in depth-three circuits. In IEEE, editor, *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science: [FOCS 2007]: October 20–23, 2007, Providence, Rhode Island*, pages 449–458, pub-IEEE:adr, 2007. IEEE Computer Society Press.

[Che66] Elliot W. Cheney. *Introduction to Approximation Theory.* McGraw-Hill Book Co., New York, 1966.

[CS04] Vincent Conitzer and Tuomas Sandholm. Communication complexity as a lower bound for learning in games. In Carla E. Brodley, editor, *ICML*, volume 69 of *ACM International Conference Proceeding Series.* ACM, 2004.

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

[DN11] Shahar Dobzinski and Noam Nisan. Limitations of VCG-based mechanisms. *Combinatorica*, 31(4):379–396, 2011.

[EZ64] H. Ehlich and K. Zeller. Schwankung von polynomen zwischen gitterpunkten. *Mathematische Zeitschrift*, 86:41–44, 1964.

[Gro94] Vince Grolmusz. The bns lower bound for multi-party protocols in nearly optimal. *Inf. Comput.*, 112(1):51–54, 1994.

[HG91] Johan Håstad and Mikael Goldmann. On the power of small-depth threshold circuits. *Computational Complexity*, 1:113–129, 1991.

[HM07] Sergiu Hart and Yishay Mansour. The communication complexity of uncoupled nash equilibrium procedures. In David S. Johnson and Uriel Feige, editors, *STOC*, pages 345–353. ACM, 2007.

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, 1997.

[KS92] Bala Kalyanasundaram and Georg Schnitger. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics*, 5(4):545–557, November 1992.

[LS09] Troy Lee and Adi Shraibman. Disjointness is hard in the multiparty number-on-the-forehead model. *Computational Complexity*, 18(2):309–336, 2009.

[Nis02] Noam Nisan. The communication complexity of approximate set packing and covering. *Lecture Notes in Computer Science*, 2380:868–??, 2002.

[NS94] Noam Nisan and Mario Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4(4):301–313, 1994.

[NS06] Noam Nisan and Ilya Segal. The communication requirements of efficient allocations and supporting prices. *J. Economic Theory*, 129(1):192–224, 2006.

[NW93] Noam Nisan and Avi Wigderson. Rounds in communication complexity revisited. *SIAM Journal on Computing*, 22(1):211–219, February 1993.

[PSS08] Christos H. Papadimitriou, Michael Schapira, and Yaron Singer. On the hardness of being truthful. In *FOCS*, pages 250–259. IEEE Computer Society, 2008.

[Raz92] Razborov. On the distributed complexity of disjointness. *TCS: Theoretical Computer Science*, 106, 1992.

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

[RC66] Theodore J. Rivlin and Elliott W. Cheney. A comparison of uniform approximations on an interval and a finite subset thereof. *SIAM Journal on Numerical Analysis*, 3(2):311–320, June 1966.

[RW93] Alexander A. Razborov and Avi Wigderson. $n^{\omega}(logn)$ lower bounds on the size of depth-3 threshold circuits with AND gates at the bottom. *Information Processing Letters*, 45(6):303–307, 1993.

[She09] Alexander A. Sherstov. Separating AC^0 from depth-2 majority circuits. *SIAM Journal of Computing*, 38(6):2113–2129, 2009.

[She11] Alexander A. Sherstov. The pattern matrix method. *SIAM Journal of Computing*, 40(6):1969–2000, 2011.

[She12] Alexander A. Sherstov. The multiparty communication complexity of set disjointness. In *STOC*, pages 525–548, 2012.

[She13] Alexander A. Sherstov. Communication lower bounds using directional derivatives. In *STOC*, pages 921–930, 2013.

[Tes03] Pascal Tesson. Computational complexity questions related to finite monoids and semigroups, 2003.

[Vio07] Emanuele Viola. Pseudorandom bits for constant-depth circuits with few arbitrary symmetric gates. *SIAM Journal on Computing*, 36(5):1387–1403, ???? 2007.

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

ECCC ISSN 1433-8092

http://eccc.hpi-web.de