MAX-PLANCK-INSTITUT FÜR INFORMATIK

Harmonic Analysis, Real Approximation, and the Communication Complexity of Boolean Functions

Technical Report No. MPII–1993–161

Vince Grolmusz
Max Planck Institute for Computer Science
and Eötvös University

November 22, 1993

INFORMATIK

Im Stadtwald
66123 Saarbrücken
Germany

Harmonic Analysis, Real Approximation, and the Communication Complexity of Boolean Functions

Technical Report No. MPII–1993–161

Vince Grolmusz
Max Planck Institute for Computer Science
and Eötvös University

November 22, 1993

Harmonic Analysis, Real Approximation, and the Communication Complexity of Boolean Functions

Technical Report No. MPII-1993-161

Vince Grolmusz

Max Planck Institute and Eötvös University

ABSTRACT:

In this paper we prove several fundamental theorems, concerning the multi–party communication complexity of Boolean functions.

Let g be a real function which approximates Boolean function f of n variables with error less than 1/5. Then — from our Theorem 1 — there exists a $k = O(\log(n\mathrm{L}_1(g)))$–party protocol which computes f with a communication of $O(\log^3(n\mathrm{L}_1(g)))$ bits, where $\mathrm{L}_1(g)$ denotes the L_1 spectral norm of g.

We show an upper bound to the symmetric k–party communication complexity of Boolean functions in terms of their L_1 norms in our Theorem 3. For $k = 2$ it was known that the communication complexity of Boolean functions are closely related with the *rank* of their communication matrix [Ya1]. No analogous upper bound was known for the k–party communication complexity of *arbitrary* Boolean functions, where $k > 2$.

For a Boolean function of exponential L_1 norm our protocols need $n^{\Omega(1)}$ bits of communication. However, if the *Fourier–coefficients* of a Boolean function f are *unevenly* distributed, more exactly, if they can be divided into two groups: one with small L_1 norm (say, L), and the other with small enough L_2 norm (say, ε), then there exists a $O(\log(nL))$–party protocol which computes f with $O(\log^3(Ln))$ communication on the $(1-\varepsilon^2)$ fraction of all inputs.

In contrast, we prove that almost all Boolean functions of n variables has a k–party communication complexity of at least $n/k - 4\log n$. This result, along with our upper bounds, shows that for almost all Boolean function no real approximating function of small L_1 norm can be found, or: almost all Boolean function has exponential L_1 norm, or: for almost all Boolean function the distribution of the Fourier–coefficients is "even": they cannot be divided into two classes: one with small L_1, the other with small L_2 norms.

Our results suggest that in the multi–party communication theory, instead of the well–studied *degree* of a polynomial representation of a Boolean function, its L_1 norm can be an important measure of complexity.

Address: Max Planck Institute for Computer Science, Im Stadtwald, D-66123 Saarbruecken, GERMANY; email: grolmusz@mpi–sb.mpg.de

1. INTRODUCTION

1.1 Multi–party games

The *multi–party communication game*, defined by *Chandra, Furst* and *Lipton* [CFL], is an interesting generalization of the 2–party communication game. In this game, k players: $P_1, P_2 ..., P_k$ intend to compute a Boolean function $f(x_1, x_2, ..., x_n) : \{0,1\}^n \to \{0,1\}$. On set $S = \{x_1, x_2, ..., x_n\}$ of variables there is a fixed partition A of k classes $A_1, A_2, ..., A_k$, and player P_i knows every variable, *except* those in A_i, for $i = 1, 2, ..., k$. The players have unlimited computational power, and they communicate with the help of a blackboard, viewed by all players. Only one player may write on the blackboard at a time. The goal is to compute $f(x_1, x_2, ..., x_n)$, such that at the end of the computation, every player knows this value. The cost of the computation is the number of bits written on the blackboard for the given $x = (x_1, x_2, ..., x_n)$ and $A = (A_1, A_2, ..., A_k)$. The cost of a multi–party protocol is the maximum number of bits communicated for any x from $\{0,1\}^n$ and the given A. The k-party communication complexity, $C_A^{(k)}(f)$, of a function f, with respect to partition A, is the minimum of costs of those k-party protocols which compute f. The k-party symmetric communication complexity of f is defined as

$$C^{(k)}(f) = \max_A C_A^{(k)}(f),$$

where the maximum is taken over all k–partitions of set $\{x_1, x_2, ..., x_n\}$.

The theory of the k–party communication games for $k = 2$ is well developed (see [BFS] or [L] for a survey), but much less is known about the $k > 2$ case. As a general upper bound both for two and more players, let us suppose that A_1 is one of the smallest classes of $A_1, A_2, ..., A_k$. Then P_1 can compute any Boolean function of S with $|A_1| + 1$ bits of communication: P_2 writes down the $|A_1|$ bits of A_1 on the blackboard, P_1 reads it, and computes and announces the value $g(x_1, x_2, ..., x_n) \in \{0,1\}$. So

$$C^{(k)}(f) \leq \left\lfloor \frac{n}{k} \right\rfloor + 1.$$

We show in Theorem 7 that this upper bound is nearly optimal for almost all Boolean function.

For two players, the communication complexity of a function f is known to be between the rank and the logarithm of the rank of the *communication matrix* of f [Ya1], [L]. Better upper bounds were given for special classes of functions by *Lovász* and *Saks* [LS], using extensively lattice–theory and Moebius functions. For more than two players, no analogue results were known.

Chandra, Furst and *Lipton* [CFL] proved non–trivial upper and lower bounds for the k–communication complexity of a specific function, using intricate Ramsey-theoretic arguments.

An important progress was made by *Babai, Nisan* and *Szegedy*, [BNS], proving an $\Omega(\frac{n}{4^k})$ lower bound for the k–party communication complexity of the GIP function. It is proved in [G] that their lower bound is close to the optimal.

We proved in [G3] that any function, computed by a depth–2 MOD p circuit of size N can be computed with p players and $O(p)$ bits of communication, and the number of communicated bits do not depend on N.

In this paper we give several fundamental upper bounds to the symmetric multi–party communication complexity of *arbitrary* Boolean functions. Our bounds depend on the L_1 *spectral norm* of functions.

1.2 Spectral Norms

There is a vast literature on representing the Boolean functions by polynomials above some field or ring (see, e.g. [ABFR], [BBR], [Be], [BRS], [BS], [LMN], [NS], [Sm]). One reason for this may be that the polynomials offer a more developed machinery than the "pure" Boolean functions. One tool in this machinery is the Fourier–expansion of Boolean functions [LMN], [BS], [KKL], [NS]:

Let us represent Boolean function f as a function $f : \{-1,1\}^n \to \{-1,1\}$ where -1 stays for "true". The set of all real valued functions over $\{-1,1\}^n$ forms a 2^n dimensional vector–space over the reals with an inner product:

$$< g,h >= 2^{-n} \sum_{x\in\{-1,1\}^n} g(x)h(x).$$

Let us define for $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n) \in \{0,1\}^n$

$$X^\alpha = \prod_{i=1}^{n} x_i^{\alpha_i}.$$

The monomials X^α for $\alpha \in \{0,1\}^n$ form an *orthonormal basis* in this 2^n–dimensional vector space; consequently, any function $h : \{-1,1\}^n \to \Re$ can be uniquely expressed as

$$(1) \qquad h(x_1, x_2, ..., x_n) = \sum_{\alpha\in\{0,1\}^n} a_\alpha X^\alpha$$

The right-hand-side of (1) is called the *Fourier–expansion* of h, and numbers a_α for $\alpha \in \{0,1\}^n$ are called *the spectral (or Fourier–) coefficients* of h.

The L_1 norm of h is:

$$L_1(h) = \sum_{\alpha\in\{0,1\}^n} |a_\alpha|$$

The L_2 norm:

$$L_2(h) = \left(\sum_{\alpha\in\{0,1\}^n} a_\alpha^2 \right)^{\frac{1}{2}} =< h,h >^{\frac{1}{2}} .$$

Example. *The PARITY function in this setting is $x_1x_2...x_n$, its* L_1 *and* L_2 *norms are 1, while its degree is n.*

Linial, Mansour and *Nisan* [LMN] proved that if f is a Boolean function computed by a bounded–depth, polynomial–size Boolean circuit, then the L_2 norm of the end–segments of the Fourier–expansion of f are decreasing exponentially fast.

Bruck and *Smolensky* [BS] established a relation between the L_1 norm and the computability of f by polynomial threshold functions. A generalization of one of their results plays a main role in the present work (Lemma 9).

1.3 Our results

Our Theorem 1 shows, that if a Boolean function can be approximated by a *real* function with small error, then there exists a k–party protocol which computes the Boolean function, and the number of communicated bits in this protocol depends only on the L_1 norm of the *approximating real function.*

Theorem 1. *Let f be a Boolean function: $f:\{-1,1\}^n \to \{-1,1\}$, and g be a real function $g:\{-1,1\}^n \to \Re$. Suppose that for all $x \in \{-1,1\}^n$,*

$$|g(x) - f(x)| < \frac{1}{5}.$$

Then the k–party symmetric communication complexity of f is

$$O\left(k^2 \log(nL_1(g)) \left\lceil \frac{nL_1^2(g)}{2^k} \right\rceil\right).$$

Specially:

Corollary 2. *Suppose that the conditions of Theorem 1 are satisfied, and let $k = \Omega(\log(nL_1(g)))$. Then*

$$C^{(k)}(f) = O(\log^3(nL_1(g))).$$

∎

In other words, if the L_1 spectral norm of g is bounded by a polynomial in n, then the *symmetric* k–party communication complexity of f is at most $O(\log^3 n)$, with $k = \Omega(\log n)$. Choosing $f = g$ in Theorem 1, we shall get:

Theorem 3. *[G2] Let f be an arbitrary Boolean function of n variables. Then the k–party symmetric communication complexity of f,*

$$C^{(k)}(f) = O\left(k^2 \log\left(nL_1(f)\right) \left\lceil \frac{nL_1^2(f)}{2^k} \right\rceil\right).$$

∎

Or, in another setting:

Corollary 4. *Suppose that $L_1(f) > n^\varepsilon$ for some $\varepsilon > 0$. Then there exists a multi–party protocol with $\Omega(\log L_1(f))$ players and of $O(\log^3 L_1(f))$ communication which computes f.* ∎

Another corollary of Theorem 1:

Corollary 5. *Let*

$$\gamma = \inf\left\{ \mathrm{L}_1(g) | g : \{-1,1\}^n \to \Re, \text{ and } \forall x \in \{-1,1\}^n : \; |g(x) - f(x)| < \frac{1}{5} \right\}.$$

Then

$$C^{(k)}(f) = O\left(k^2 \log(n\gamma) \left\lceil \frac{n\gamma^2}{2^k} \right\rceil \right).$$

∎

Suppose that f is a Boolean function of large (say, exponential in n) L_1 norm. Our Theorem 3 can guarantee only a communication protocol with too many communicated bits: the trivial $\lfloor \frac{n}{k} \rfloor$ protocol is usually better. Suppose now, that the set of Fourier–coefficients of f can be divided into two parts: one with small L_1, the other with small L_2 norms.

Example. *Let* $|a_1| = |a_2| = \frac{1}{2} - \delta$, *and*

$$|a_3| = |a_4| = \ldots = |a_{2^n}| = 2^{-\frac{2}{3}n},$$

where $\delta = (2^{n-1} - 1)/2^{(4/3)n} = O(2^{-\frac{n}{3}})$. *Then the* L_1 *norm*

$$\sum_{i=1}^{2^n} |a_i| \geq 2^{\frac{n}{3}}$$

is exponentially large, while

$$\sum_{i=3}^{2^n} a_i^2 \leq \frac{1}{2^{\frac{n}{3}}},$$

is exponentially small, and

$$|a_1| + |a_2| < 1.$$

When the Fourier coefficients are so unevenly distributed, then we can give a much better protocol to compute f. The price: the computation will not be correct on a small fraction of the inputs.

Theorem 6. *Let*

$$f(x) = \sum_{\alpha \in \{0,1\}^n} a_\alpha X^\alpha,$$

and let $S \subset \{0,1\}^n$ *such that*

$$\sum_{\alpha \in S} a_\alpha^2 \leq \varepsilon,$$

for some $\varepsilon < \frac{1}{2500}$. *Let*

$$g(x) = \sum_{\alpha \in \{0,1\}^n - S} a_\alpha X^\alpha.$$

Then for all $k \geq 2$ and for all k–partition of the inputs, there exists a k–party protocol with

$$O\left(k^2 \log(n\mathrm{L}_1(g)) \left\lceil \frac{n\mathrm{L}_1^2(g)}{2^k} \right\rceil\right)$$

bits of communication, and this protocol computes f correctly on at least on the $(1-25\varepsilon) > \frac{99}{100}$ fraction of the inputs.

The following results of [G4] show the power of our upper bounds in Theorems 1, 3 and 6, proving that almost all Boolean function has very high communication complexity:

Theorem 7. *[G4] Let f be a uniformly chosen random member of set*

$$\{f | f : \{-1,1\}^n \to \{-1,1\}\}.$$

Then the probability, that for some A k–equipartition of $X = \{x_1, x_2, ..., x_n\}$, there exists a k–party protocol, which computes f with communication of at most $\lfloor \frac{n}{k} \rfloor - 4\log n$ bits, is less than

$$2^{-2^{\Omega(n)}}.$$

The communication complexity remains high even if we compute f on *most* of the inputs:

Theorem 8. *[G4] Let f be a uniformly chosen random member of set*

$$\{f | f : \{-1,1\}^n \to \{-1,1\}\}.$$

Then the probability, that for some A k–equipartition of $X = \{x_1, x_2, ..., x_n\}$, there exists a k–party protocol, which correctly computes f on a fraction of at least $\frac{1}{2} + \varepsilon$ of inputs, with communication of at most $\lfloor \frac{n}{k} \rfloor - 4\log \frac{n}{\varepsilon}$ bits, is less than

$$2^{-2^{\Omega(n)}}.$$

The proofs of Theorems 7 and 8 need a thoughtful analysis of the underlying structure of *cylinder intersections*, and have been appeared in [G4]. ∎

Comparing Theorems 1, 3 with Theorem 7, and Theorem 6 with Theorem 8, we have got that for almost all Boolean function f:

– f has exponential L_1–norm,

– If f is approximated by a real function g with error less than $1/5$, then the L_1 norm of g is exponential in n,

– the Fourier–coefficients of f are "evenly distributed": they cannot be divided into two sets, one with subexponential L_1 norm, the other with a small L_2 norm.

In some fields of complexity theory, the *degree* of the polynomial, which approximates, or represents a Boolean function f, has been proved to be a good characterization of the hardness of f (e.g. [NS], [Sm]). In the multi–party communication theory, as we show in this work, instead of the degree, the L_1 norm can be an important measure of complexity.

3. THE PROOF OF THEOREM 1.

The following lemma is a generalization of a lemma of *Bruck* and *Smolensky* [BS].

Lemma 9. *Let $U \subset \{-1,1\}^n$ such that $|U| \geq (1-\frac{1}{100})2^n$. Let $g : \{-1,1\}^n \to \Re$. Suppose that for all $x \in U$, $\frac{4}{5} < |g(x)| < \frac{6}{5}$ is satisfied. Then there exists polynomial $G_0(x)$ with integer coefficients and with L_1 norm*

$$L_1(G_0) \leq 400nL_1^2(g)$$

such that

$$\mathrm{sgn}(G_0(x)) = \mathrm{sgn}(g(x))$$

for all $x \in U$.

Proof. The Fourier–expansion of g:

$$g(x) = \sum_{\alpha\in\{0,1\}^n} a_\alpha X^\alpha$$

where a_α for $\alpha \in \{0,1\}^n$ are the Fourier–coefficients of g. Then by definition

$$L_1(g) = \sum_{\alpha\in\{0,1\}^n} |a_\alpha|.$$

and

$$L_2(g) = < g, g > = 2^{-n} \sum_{x\in\{-1,1\}^n} g^2(x) = \sum_{\alpha\in\{0,1\}^n} a_\alpha^2,$$

using the *Parseval*-identity.

Since $|g(x)| \geq \frac{4}{5}$ for $x \in U$, and $|U| \geq (1 - \frac{1}{100})2^n$,

$$L_2(g) \geq \left(1 - \frac{1}{100}\right)\frac{16}{25}.$$

Our next step is giving a lower bound to the L_1 norm of g.

Case I. Suppose that there exists an α: $|a_\alpha| > \frac{1}{2}$. If $\mathrm{sgn}(X^\alpha) = \mathrm{sgn}(g(x))$ for all $x \in U$, then we are done, $G_0(x) = X^\alpha$ suffices. Otherwise, for some $x \in U$, $\mathrm{sgn}(X^\alpha) \neq \mathrm{sgn}(g(x))$. Then the other terms of g must compensate X^α, so the sum of the absolute values of their coefficients should be greater than $\frac{4}{5}$. So

$$L_1(g) \geq \frac{4}{5} + |a_\alpha| \geq \frac{13}{10}.$$

Case II. If all $|a_\alpha| \leq \frac{1}{2}$, then

$$\left(1 - \frac{1}{100}\right)\frac{16}{25} \leq \sum_{\alpha\in\{0,1\}^n} a_\alpha^2 \leq \frac{1}{2} \sum_{\alpha\in\{0,1\}^n} |a_\alpha|,$$

so

$$\left(1-\frac{1}{100}\right)\frac{32}{25} \leq \sum_{\alpha\in\{0,1\}^n} |a_\alpha| = \mathrm{L}_1(g).$$

Consequently, either we have found a suitable $G_0(x)$, or we have concluded that

$$\mathrm{L}_1(g) \geq \left(1-\frac{1}{100}\right)\frac{32}{25} \geq \frac{127}{100}. \tag{3}$$

Let us define random monomials Z_i as follows:

$$Z_i = \mathrm{sgn}(a_\alpha)X^\alpha \quad \text{with probability} \quad \frac{|a_\alpha|}{\mathrm{L}_1(g)}.$$

Let $G(x)$ random polynomial be the sum of $N = \lfloor 400nL_1^2(g)\rfloor$ monomials Z_i:

$$G(x) = \sum_{i=1}^{N} Z_i.$$

Computing the expectation of Z_i:

$$\mathrm{E}(Z_i(x)) = \sum_{\alpha\in\{0,1\}^n} \frac{|a_\alpha|}{\mathrm{L}_1(g)}\mathrm{sgn}(a_\alpha)X^\alpha = \frac{g(x)}{\mathrm{L}_1(g)},$$

where we used the fact that $\mathrm{sgn}(v)|v| = v$.
The expectation of $G(x)$

$$\mathrm{E}(G(x)) = \frac{Ng(x)}{\mathrm{L}_1(g)}. \tag{4}$$

The variance of Z_i:

$$\mathrm{Var}(Z_i(x)) = \mathrm{E}(Z_i^2) - \mathrm{E}^2(Z_i) = 1 - \frac{g^2(x)}{\mathrm{L}_1^2(g)}.$$

The variance of $G(x)$:

$$\mathrm{Var}(G(x)) = N\left(1 - \frac{g^2(x)}{\mathrm{L}_1^2(g)}\right).$$

Since $|g(x)| \leq \frac{6}{5}$, and because of (3):

$$\frac{g^2(x)}{\mathrm{L}_1^2(g)} \leq \left(\frac{120}{127}\right)^2 \leq \frac{9}{10},$$

so

$$\frac{N}{10} \leq Var(G(x)) \leq N$$

or

$$\sqrt{\frac{N}{10}} \leq D(G(x)) \leq \sqrt{N}, \tag{5}$$

where $D(G(x)) = \sqrt{Var(G(x))}$, the standard deviation of $G(x)$.

From (4), the sign of $E(G(x))$ is the same as the sign of $g(x)$. Consequently,

$$\Pr\Big(\text{sgn}(G(x)) \neq \text{sgn}(g(x)\Big) = \Pr\Big(\text{sgn}(G(x)) \neq \text{sgn}(\text{E}(G(x))\Big) \leq$$

$$\leq \Pr\Big(|G(x) - \text{E}(G(x))| \geq \frac{N|g(x)|}{\text{L}_1(g)}\Big)\Big) \leq \Pr\Big(|G(x) - \text{E}(G(x))| \geq \frac{4N}{5\text{L}_1(g)}\Big)\Big).$$

From the *Bernstein–inequality* (see [Re1] or [Re2]), (or from the Central Limit Theorem), with $D = D(G(x))$, we have got:

$$\Pr(|G(x) - \text{E}(G(x))| \geq \mu D) \leq 2\exp\left(-\frac{\mu^2}{2\left(1+\frac{\mu}{D}\right)^2}\right), \tag{6}$$

where $0 < \mu < \frac{D}{2}$.

For $\mu = 3\sqrt{n}$, $N = \lfloor 400nL_1^2(g) \rfloor$ we got that the probability in (6) is less than e^{-n}. On the other hand,

$$\mu D \leq \frac{4N}{5\text{L}_1(g)},$$

so

$$\Pr\Big(\text{sgn}(G(x)) \neq \text{sgn}(g(x)\Big) < e^{-n}.$$

Consequently,

$$\Pr(\exists x \in U : \text{sgn}(G(x)) \neq \text{sgn}(g(x)) \leq$$

$$\leq \sum_{x \in U} \Pr(\text{sgn}(G(x)) \neq \text{sgn}(g(x))) \leq |U|e^{-n} \leq 2^n e^{-n} < 1,$$

and since this probability is less than one, there exists a polynomial $G_0(x)$ for which $\text{sgn}(G_0(x)) = \text{sgn}(g(x))$ for all $x \in U$. The coefficients of this G_0 are integers, and its L_1–norm is at most N. ∎

Proof of Theorem 1. Function g satisfies the requirements of Lemma 9, for $U = \{-1,1\}^n$. Then there exists a polynomial $G_0(x)$ with integer coefficients and an L_1 norm of at most $400n\text{L}_1^2$, such that

$$\text{sgn}(g(x)) = \text{sgn}(G_0(x))$$

for all $x \in \{-1,1\}^n$. Since $\mathrm{sgn}(g(x)) = f(x)$, we have got that $\mathrm{sgn}(G_0(x)) = f(x)$, for all $x \in \{-1,1\}^n$. And, by the following Theorem 10, $G_0(x)$ has the needed symmetric k–party communication complexity. ∎

Theorem 10. *Let*

$$G(x) = \sum_{i=1}^{N} Z_i,$$

where $Z_i = X^\alpha$ or $Z_i = -X^\alpha$, for some $\alpha \in \{0,1\}^n$, and for $x \in \{-1,1\}^n$. Then the symmetric k–party communication complexity of G is

$$O\left(k^2 \log(nN) \left\lceil \frac{nN^2}{2^k} \right\rceil\right).$$

Proof. Let $G_1(x)$ be the sum of Z_i's with positive sign, and let $G_2(x)$ be the sum of $(-Z_i)$'s, where Z_i has a negative sign. So:

$$G(x) = G_1(x) - G_2(x),$$

and G_1 has N_1 terms, G_2 has N_2 terms, $N_1 + N_2 = N$.

Let us observe that $G_j(x)$ is the sum of N_j terms of form

$$X^\alpha = \prod_{i=1}^{n} x_i^{\alpha_i} = \prod_{i:\alpha_i=1} x_i$$

for $j = 1,2$.
Clearly,

$$X^\alpha = \begin{cases} -1, \text{ if } |\{i : x_i = -1, \alpha_i = 1\}| \text{ is odd} \\ 1 \text{ otherwise} \end{cases}$$

For $j = 1,2$ let b_j the number (counting the possible multiplicity) of those terms X^α in $G_j(x)$ for which $|\{i : x_i = -1, \alpha_i = 1\}|$ is odd . Then $G_j(x) = (N_j - b_j) - b_j = N_j - 2b_j$, so:

$$G(x) = G_1(x) - G_2(x) = N_1 - N_2 + 2b_2 - 2b_1. \tag{2}$$

Let us denote

$$y_i = \begin{cases} 1, \text{ if } x_i = -1 \\ 0, \text{ if } x_i = 1 \end{cases}$$

then

$$X^\alpha = -1 \Longleftrightarrow \sum_{i=1}^{n} y_i \alpha_i = 1 \bmod 2.$$

Let us form a matrix $M^{(j)}$ with N_j rows and n columns, for $j = 1,2$. Each row is corresponded to a term X^α in $G_j(x)$, and the i^{th} entry of that row is $y_i \alpha_i$.

Obviously, the number of those rows of $M^{(j)}$ which have odd sum is equal to b_j.

Suppose now that we are given polynomial $G(x)$, players $P_1, P_2, ..., P_k$ and a k-partition $A = (A_1, A_2, ..., A_k)$ of the set $\{x_1, x_2, ..., x_n\}$. We assume that player P_ℓ knows function $G(x)$, partition A, functions $G_1(x)$, $G_2(x)$, and the values of all variables, except those in A_ℓ, for $\ell = 1, 2, ..., k$. Then the players, without any communication can compute privately matrices $M^{(1)}$ and $M^{(2)}$, and exactly those entries of these matrices will be not known for player P_ℓ which were corresponded to variables in class A_ℓ. The set of these entries will be called B_ℓ, for $\ell = 1, 2, ..., k$. The following lemma shows a protocol by which the players can first compute b_1 and then b_2, and consequently, $G(x)$, by equation (2).

Lemma 11. *Let $M \in \{0,1\}^{m\times n}$, $M = \{m_{ij}\}$, and let $B = \{B_1, B_2, ..., B_k\}$ a partition of the set $\{m_{ij} : 1 \le i \le m, 1 \le j \le n\}$, such that player P_ℓ knows every m_{ij} except those in B_ℓ, for $\ell = 1, 2, ..., k$. Then there exists a k-party protocol which computes the number of the rows with odd sum in M with communicating*

$$O\left(k^2 \log m \left\lceil \frac{m}{2^k} \right\rceil\right)$$

bits.

Proof. First, the players compute a matrix $Q \in \{0,1\}^{m\times k}$ from M, with no communication: for each row of M a row of Q is corresponded; the first element of row j of Q is the mod 2 sum of those entries of the j^{th} row of M which are the elements of B_1 at the same time. Analogously, the i^{th} element of row j of Q is the mod 2 sum of those entries of the j^{th} row of M which are the elements of B_i at the same time.
Clearly, the number of rows with odd sum in M and in Q is the same. Moreover, player P_ℓ knows every column of matrix Q, except column ℓ, for $\ell = 1, 2, ..., k$.

With an additional assumption Lemma 12 gives a protocol with $O(k^2 \log m)$ communication:

Lemma 12. *Let $\beta \in \{0,1\}^k$. Suppose it is known to each player that β does not occur as a row of Q. Then there exists a k–party protocol which computes the number of the odd rows with a communication of $O(k^2 \log m)$ bits.*

Proof. Without restricting the generality we may suppose that β is the all–1 vector of length k.

Let $ODD(\gamma_1\gamma_2...\gamma_\ell)$ and $EVEN(\gamma_1\gamma_2...\gamma_\ell)$ denote the number of those rows of Q which have odd (respectively, even) sums, and they begin with $\gamma_1\gamma_2...\gamma_\ell$, $\ell \le k$, $\gamma_i \in \{0,1\}$. For example, P_1 do not know the first column of Q, but he can communicate $ODD(0) + EVEN(1)$ if P_1 counts those rows which has odd sum in its second through kth position. Similarly P_2 can communicate $ODD(10)+EVEN(11)$ if he counts those rows which begins with 1, and the sum of their first, 3rd, 4th,...,kth elements is odd.

This observation motivates the following protocol:

PROTOCOL ODDCOUNT

The goal: to compute b, the number of rows with odd sum in Q. Number b will be the sum of values u_i announced by player P_i, $i = 1, 2, ..., k$.

P_1 announces $u_1 = ODD(0) + EVEN(1)$.

remark: $b = u_1 + ODD(1) - EVEN(1)$.

P_2 announces $u_2 = ODD(10) + EVEN(11) - EVEN(10) - ODD(11)$.

remark: $b = u_1 + u_2 - 2EVEN(11) + 2ODD(11)$

P_3 announces $u_3 = 2ODD(110) + 2EVEN(111) - 2EVEN(110) - 2ODD(111)$.

remark: $b = u_1 + u_2 + u_3 - 4EVEN(111) + 3ODD(111)$

P_i announces $u_i = 2^{i-2}ODD(11...10) + 2^{i-2}EVEN(11...11) - 2^{i-2}EVEN(11...10) - 2^{i-2}ODD(11...11)$

remark: $b = \sum_{j=1}^{i} u_j - 2^{i-1}EVEN(\overbrace{11...1}^{i \text{ times}}) + (2^{i-1} - 1)ODD(\overbrace{11...1}^{i \text{ times}})$.

After P_k announces u_k, the players privately add up the u_i's from $i = 1$ through k. Let us remark that

$$b = \sum_{j=1}^{k} u_j - 2^{k-1}EVEN(\overbrace{11...1}^{k \text{ times}}) + (2^{k-1} - 1)ODD(\overbrace{11...1}^{k \text{ times}}).$$

However, as we assumed at the beginning, there are no all–1 rows in Q, so

$$b = \sum_{j=1}^{k} u_j$$

and we are done. Each u_i can be communicated using $O(k \log m)$ bits, so the total communication is $O(k^2 \log m)$. ∎

Now we return to the proof of Lemma 11. Let us divide the rows of matrix Q into blocks of $2^{k-1} - 1$ contiguous rows plus a leftover of at most $2^{k-1} - 1$ rows. The players cooperatively determine the number of the odd rows in each block, and then privately add up the results.

Next we show how to obtain the number of the odd rows for a single block at the cost of $O(k^2 \log m)$ bits of communication. P_1 knows all the columns, except the first, so he knows at most $2^{k-1} - 1$ rows of length $k - 1$ in a block, so he can find an $\beta' \in \{0,1\}^{k-1}$, $\beta' = (\beta_2, \beta_3, \ldots, \beta_k)$ which is not a row of the $k - 1$ column wide part of the block seen by P_1. Let $\beta = (1, \beta_2, \beta_3, \ldots, \beta_k)$. Then β does not occur as a row in this block. So if P_0 communicates β, and they play protocol ODDCOUNT of Lemma 12 for a given block.

They use $k^2 \log m$ bits for a block, and, since there are at most $\left\lceil \frac{m}{2^{k-1}-1} \right\rceil$ blocks, the total communication is

$$O\left(k^2 \log m \left\lceil \frac{m}{2^k} \right\rceil\right).$$

■

4. PROOF OF THEOREM 6.

Lemma 13. *Let f be a Boolean function and let $h : \{-1,1\}^n \to \Re$ such that*

$$\mathrm{L}_2^2(f-h) = < f-h, f-h > \leq \varepsilon.$$

Then

$$\mathrm{Pr}_x(|f(x)-h(x)| > \frac{1}{5}) \leq 25\varepsilon,$$

where Pr_x is the probability measure associated with the uniform distribution over $\{-1,1\}^n$.

Proof.

$$\varepsilon \geq < f(x)-h(x), f(x)-h(x) >= \mathrm{E}_x(f(x)-h(x))^2 \geq \frac{1}{25}\mathrm{Pr}_x\Big(|f(x)-h(x)| > \frac{1}{5}\Big).$$

■

Now we prove Theorem 6. Let U be defined as

$$U = \Big\{x \in \{-1,1\}^n : |f(x)-g(x)| \leq \frac{1}{5}\Big\}.$$

From Lemma 13, $|U| \geq (1-25\varepsilon)2^n$. If $\varepsilon \leq \frac{1}{2500}$ then we can apply Lemma 9 for g. The proof proceeds exactly as in the proof of Theorem 1. ■

REFERENCES

[ABFR] J. Aspnes, R. Beigel, M. Furst, S. Rudich: The expressive power of voting polynomials, Proc. 23rd ACM STOC, 1991, pp. 402-409

[BBR] D. A. Barrington, R. Beigel, S. Rudich: Representing Boolean functions as polynomials modulo composite numbers, Proc. 24th ACM STOC, 1992, pp. 455-461

[Be] R. Beigel: When do extra MAJORITY gates help?, Proc. 24th ACM STOC, 1992, pp. 450-454

[BFS] L. Babai, P. Frankl, J. Simon: Complexity classes in communication complexity theory, Proc. 27th IEEE FOCS, 1986, pp. 337-347.

[BNS] L. Babai, N. Nisan, M. Szegedy: Multiparty Protocols and Pseudorandom Sequences, Proc. 21st ACM STOC, 1989, pp. 1-11.

[BRS] R. Beigel, N. Reingold, D. Spielman: The perceptron strikes back, Proc. 6th Annual Conference on Structure in Complexity Theory, IEEE Comp. Soc. Press, 1991

[BS] J. Bruck, R. Smolensky: Polynomial threshold functions, AC^0 functions and spectral norms, Proc. 32nd IEEE FOCS, 1991, pp. 632–641

[CFL] A. K. Chandra, M. L. Furst, R. J. Lipton: Multi-party Protocols, Proc. 15th ACM STOC, 1983, pp. 94–99.

[G] V. Grolmusz: The BNS Lower Bound for Multi–Party Protocols is Nearly Optimal, to appear in "Information and Computation".

[G2] V. Grolmusz: Multi–Party Protocols and Spectral Norms, Technical Report MPII-1993-132, Max Planck Institute for Computer Science, Saarbruecken, Germany, 1993.

[G3] V. Grolmusz: Separating the communication complexities of MOD m and MOD p circuits, Proc. 33rd IEEE FOCS, 1992, pp. 278-287.

[G4] V. Grolmusz: On Multi–Party Communication Complexity of Random Functions, Technical Report MPII-1993-162, Max Planck Institute for Computer Science, Saarbruecken, Germany, 1993.

[KKL] J. Kahn, G. Kalai, N. Linial: The influence of variables on Boolean functions, Proc. 29th IEEE FOCS, 1988, pp. 68–80.

[L] L. Lovász: Communication Complexity: A Survey, Technical Report, CS–TR–204–89, Princeton University, 1989.

[LMN] N. Linial, Y. Mansour, N. Nisan: Constant depth circuits, Fourier transform and learnability, Proc. Proc. 30th IEEE FOCS, 1989, pp. 574–579

[NS] N. Nisan, M. Szegedy: On the degree of Boolean functions as real polynomials, Proc. 24th ACM STOC, 1992, pp. 462–467

[Re1] A. Rényi: Wahrscheinlichkeitsrechnung, VEB Deutscher Verlag der Wissenschaften, Berlin, 1962

[Re2] A. Rényi: Valószínűségszámítás, Tankönyvkiadó, Budapest, 1973

[Sm] R. Smolensky, Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity, Proc. 19th ACM STOC, pp. 77-82, (1987).

[Ya1] A.C. Yao: Some Complexity Questions Related to Distributive Computing, Proc. 11th ACM STOC, 1979, pp. 209–213.