MAX-PLANCK-INSTITUT FÜR INFORMATIK

On Multi–Party Communication Complexity of Random Functions

Technical Report No. MPII–1993–162

Vince Grolmusz
Max Planck Institute for Computer Science
and Eötvös University

December 1, 1993

INFORMATIK

Im Stadtwald
66123 Saarbrücken
Germany

On Multi–Party Communication Complexity of Random Functions

Technical Report No. MPII–1993–162

Vince Grolmusz
Max Planck Institute for Computer Science
and Eötvös University

December 1, 1993

On Multi–Party Communication Complexity of Random Functions

Technical Report No. MPII–1993–162

Vince Grolmusz

Max Planck Institute and Eötvös University

Keywords: communication complexity, random functions, communication protocols

ABSTRACT:

We prove that almost all Boolean function has a high k–party communication complexity. The 2–party case was settled by *Papadimitriou* and *Sipser* [PS]. Proving the k–party case needs a deeper investigation of the underlying structure of the k–cylinder–intersections; (the 2–cylinder–intersections are the rectangles).
First we examine the basic properties of k–cylinder–intersections, then an upper estimation is given for their number, which facilitates to prove the lower–bound theorem for the k–party communication complexity of random Boolean functions. In the last section we extend our results to communication protocols, which are correct only on *most* of the inputs.

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 $\mathcal{A}$ 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 $f(x_1, x_2, ..., x_n)$, and write it down to the blackboard. 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 $\mathcal{A} = (A_1, A_2, ..., A_k)$. The cost of a multi–party protocol is the maximum number of bits communicated for any x from $\{0,1\}^n$ and the given $\mathcal{A}$. The k-party communication complexity, $C_{\mathcal{A}}^{(k)}(f)$, of a function f, with respect to partition $\mathcal{A}$, is the minimum of costs of those k-party protocols which compute f. The k-party symmetric communication complexity of f is defined as

$$C^{(k)}(f) = \max_{\mathcal{A}} C_{\mathcal{A}}^{(k)}(f),$$

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

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

In this paper we consider only the "hard" case, when all the classes are of the same size, in other words:

$$n = mk, \quad |A_1| = |A_2| = ... = |A_k| = m.$$

Then

$$C^{(k)}(f) \leq m + 1.$$

For $k = 2$, *Papadimitriou* and *Sipser* [PS] proved that for almost all Boolean functions f

$$C^{(2)}(f) = m + 1.$$

For $k > 2$ analogous results were not known. Our main result is the following theorem:

Theorem 1 . *Let f be a uniformly chosen random member of set*

$$\{f | f : \{0,1\}^{mk} \to \{0,1\}\}.$$

Then the probability, that for some $\mathcal{A}$ k–equipartition of $X = \{x_1, x_2, ..., x_{mk}\}$ there exists a k–party protocol, which computes f with communication of at most $m - (\log m + 2 \log k + 1)$ bits is less than

$$2^{-2^{m(k-1)}}.$$

1.2 Correct Computation on Most of the Inputs

Babai, Nisan and *Szegedy* [BNS] investigated the k–party communication complexity of computing a specific function correctly on *most* of the possible inputs. Our next theorem shows that for almost all Boolean functions, this task also needs almost m bits to communicate:

Theorem 2 . *Let f be a uniformly chosen random member of set*

$$\{f | f : \{0,1\}^{mk} \to \{0,1\}\}.$$

Then the probability, that for some $\mathcal{A}$ 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 $m - (\log m + 2\log k + 3\log\frac{1}{\varepsilon} + 2)$ bits, is less than

$$2^{-2^{m(k-1)}}.$$

2. PRELIMINARIES

2.1 Protocols and Cylinder-Intersections

The notion of *cylinder-intersections* plays an important role in the theory of multi–party communication games.

Let $X = \{x_1, x_2, ..., x_n\}$ be a set of symbols, and let $\mathcal{A} = \{A_1, A_2, ..., A_k\}$ be a k-partition on X, i.e., for $1 \le i \le k$ $A_i \subset X$, and

$$\bigcup_{i=1}^{k} A_i = X.$$

For a set S let $\{0,1\}^S$ denote the set of all functions of form $h : S \to \{0,1\}$. Clearly, $\{0,1\}^n$ is isomorphic with $\{0,1\}^{A_1 \cup A_2 \cup ... \cup A_k}$. In this paper we do not make any distinction between them.

Let

$$p_i : \{0,1\}^n \to \{0,1\}^{A_1 \cup A_2 \cup ... \cup A_{i-1} \cup A_{i+1} \cup ... \cup A_k}$$

be a projection for $i = 1, 2, ..., k$. In other words, p_i simply cuts out those coordinates of an n-bit sequence, which corresponds to the elements of A_i.

Definition 3 . *Let $1 \le i \le k$ and let*

$$Q_i \subset p_i(\{0,1\}^n).$$

Then

$$p_i^{-1}(Q_i) \subset \{0,1\}^n$$

is called an $(i, \mathcal{A})$–cylinder on Q_i. Set $Q \in \{0,1\}^n$ is called a $(k, \mathcal{A})$ cylinder–intersection (or, k-cylinder–intersection, if $\mathcal{A}$ is fixed), if there exist $Q_1, Q_2, ..., Q_k$ such that $Q_i \subset p_i(\{0,1\}^n)$, $i = 1, 2, ..., k$, and

$$Q = \bigcap_{i=1}^{k} p_i^{-1}(Q_i). \tag{1}$$

Babai, Nisan and *Szegedy* [BNS] proved the following lemma:

Lemma 4 . *Let s be a string, written by the k players to the blackboard, in case of a fixed input. Then the set of all inputs, which imply that the string s is written onto the blackboard, is a k–cylinder–intersection.* ∎

2.2 Basic properties

By Definition 3, if $q \in \{0,1\}^n$, and the projection of q to $Y^{(i)}$ is in Q_i, for $i = 1, 2, ..., k$, then q itself is also in Q. This observation is formalized in Theorem 5 :

Theorem 5 .

$Q \subset \{0,1\}^n$ is a $(k, \mathcal{A})$ cylinder–intersection

if and only if

$$(q_1', q_2, q_3, ..., q_k) \in Q$$
$$(q_1, q_2', q_3, ..., q_k) \in Q$$
$$(q_1, q_2, q_3', ..., q_k) \in Q$$
$$(q_1, q_2, q_3, ..., q_k') \in Q$$

implies

$$(q_1, q_2, q_3, ..., q_k) \in Q,$$

where $q_i, q_i' \in \{0,1\}^{A_i}$, and $q_i \neq q_i'$, for $i = 1, 2, ..., k$.

Proof. The proof of the "only if" part appears just before the statement of the theorem. The proof of the "if" part:
Let $Q_i = p_i(Q)$, for $i = 1, 2, ..., k$. Let

$$Q' = \bigcap_{i=1}^{k} p_i^{-1}(Q_i).$$

Obviously, $Q \subset Q'$. Suppose that $q \in Q'$. Since $q \in p_i^{-1}(Q_i)$, there exists a q_i' such that

$$(q_1, q_2, ..., q_{i-1}, q_i', q_{i+1}, ..., q_k) \in Q,$$

for $i = 1, 2, ..., k$. But this implies $q \in Q$, consequently, $Q' \subset Q$. ∎

Definition 6 . *Let $q, q^{(j)} \in \{0,1\}^n$ for $j = 1, 2, ..., k$. Set $\{q^{(1)}, q^{(2)}, ..., q^{(k)}, q\}$ is called a $(k, \mathcal{A})$–pyramid, if there exist $q_i, q_i' \in \{0,1\}^{A_i}$ for $i = 1, 2, ..., k$, such that*

$$(q_1', q_2, q_3, ..., q_k) = q^{(1)}$$
$$(q_1, q_2', q_3, ..., q_k) = q^{(2)}$$
$$(q_1, q_2, q_3', ..., q_k) = q^{(3)}$$
$$(q_1, q_2, q_3, ..., q_k') = q^{(k)}$$
$$(q_1, q_2, q_3, ..., q_k) = q$$

q is the top of the pyramid, while $\{q^{(1)}, q^{(2)}, ..., q^{(k)}\}$ is the fundament of the pyramid.

Using the definition of the pyramids, Theorem 2 can also be stated as follows: Q is a k–cylinder–intersection $\Longleftrightarrow$ If Q contains the fundament of a pyramid then it also contains its top.

Definition 7. *$I \subset \{0,1\}^n$ is called a c–independent set (where "c" stays for "cylinder") if I does not contain any k–pyramids. Let Q be a k-cylinder–intersection. Set $G \subset Q$ is called c–generator for Q if every $q \in Q$ can be written as the top of a pyramid, with all the fundamental points in G.*

Definition 8. *Let Q be a k–cylinder–intersection. $B \subset Q$ is called a c-basis of Q, if B is a c-independent c-generator for Q.*

Theorem 9 . *Every k–cylinder–intersection contains a c–basis.*

Proof. Let G be a minimal c–generator for k–cylinder–intersection Q. We show that G is c–independent. Suppose that G contains a pyramid:

$$\begin{aligned}
(q_1', q_2, q_3, ..., q_k) &= q^{(1)} \\
(q_1, q_2', q_3, ..., q_k) &= q^{(2)} \\
(q_1, q_2, q_3', ..., q_k) &= q^{(3)} \; . \\
(q_1, q_2, q_3, ..., q_k') &= q^{(k)} \\
(q_1, q_2, q_3, ..., q_k) &= q
\end{aligned}$$

Then $G - \{q\}$ is also a c–generator for Q: suppose that q is the i^{th} element in a pyramid which generates element $\bar{q}$. Then q can be substituted in the same pyramid by $q^{(i)}$. Consequently, G cannot contain a pyramid. ∎

2.3 Theorems for the equipartition

Let $n = mk$, and let $\mathcal{A}$ be an equipartition of X: $|A_1| = |A_2| = \ldots = |A_k| = m$. In this section partition $\mathcal{A}$ remains fixed.

The following theorem gives an upper bound to the size of any c–independent set in $\{0,1\}^{mk}$:

Theorem 10. *Let $I \subset \{0,1\}^{mk}$ be a c–independent set. Then*

$$|I| \leq k2^{(k-1)m}.$$

The next structural lemma is needed in the proof of Theorem 10 :

Lemma 11 . *Let $I \subset \{0,1\}^{A_1 \cup A_2 \cup \ldots \cup A_k}$ be a c–independent set, and let $q \in I$. Then there exists an i: $1 \leq i \leq k$:*

$$p_i^{-1}(r^{(i)}) \cap I = \{q\},$$

where $r^{(i)} = p_i(q) \in Y^{(i)}$.

Proof. Suppose that all intersection has at least two elements: for $i = 1, 2, ..., k$, $\exists q^{(i)} \in \{0,1\}^n$: $p_i^{-1}(r^{(i)}) \cap I = \{q, q^{(i)}\}$, where $q^{(i)} \neq q$. Let us observe that $q^{(i)} \neq q^{(j)}$ for $1 \leq i < j \leq k$. Then

$$q^{(1)}, q^{(2)}, ..., q^{(k)}, q$$

form a pyramid, entirely in I, which is a contradiction. ∎

Proof of Theorem 10 . From Lemma 11 :

$$|I| \leq \sum_{i=1}^{k} |p_i(I)|.$$

On the other hand,

$$p_i(I) \subset Y^{(i)}, \quad |Y^{(i)}| = 2^{(k-1)m}$$

so

$$|I| \leq k2^{k-1}m.$$

∎

Theorem 12 . *The number of k–cylinder–intersections in $\{0,1\}^{A_1 \cup A_2 \cup \ldots \cup A_k}$ is at most*

$$\binom{2^{mk}}{k2^{m(k-1)}}.$$

Proof. By Theorem 9, every k–cylinder–intersection has a c–basis. Every k–cylinder–intersection can be corresponded to one of its c–basises. Since the c–basis generates the cylinder–intersection, different cylinder–intersections are corresponded to different c–basises. Every c–basis is c–independent, so, from Theorem 10 , its size is less than or equal to $k2^{(k-1)m}$. The statement follows. ∎

3. THE PROOF OF THEOREM 1

Now we are ready to prove Theorem 1 . Let $\mathcal{A}$ be fixed. Suppose that function $f : \{0,1\}^{mk} \to \{0,1\}$ is computed by a k–party protocol, where player P_i knows the value of every variable, except those in A_i. We also suppose, that the players use at most $m - \alpha$ bits for the communication. Every possible communication-sequence corresponds to a cylinder–intersection by Lemma 4 , on which f is constant: either 0 or 1.
Since there are at most $2^{m-\alpha}$ possible communication-sequence, there exists a cylinder–intersection Q such that

$$|Q| \geq \frac{2^{mk}}{2^{m-\alpha}} = 2^{m(k-1)+\alpha},$$

and $f(Q) = \{1\}$ or $f(Q) = \{0\}$.
Obviously, there are

$$2^{2^{mk}}$$

different functions $f : \{0,1\}^{mk} \to \{0,1\}$. By the uniform distribution, let us choose randomly an f among these functions.
The probability, that f is constant on Q is at most

$$2^{1-2^{m(k-1)+\alpha}}.$$

By Theorem 12 , there are at most

$$\binom{2^{mk}}{k2^{m(k-1)}} \leq 2^{mk^2 2^{m(k-1)}}$$

cylinder–intersections in $\{0,1\}^{A_1 \cup A_2 \cup \ldots \cup A_k}$.
So, the probability, that a random f is constant on *at least one* cylinder–intersection of size at least $2^{m(k-1)+\alpha}$) is at most

$$2^{mk^2 2^{m(k-1)}} 2^{1-2^{m(k-1)+\alpha}} = 2^{2^{m(k-1)}(mk^2 - 2^\alpha)+1},$$

for a fixed $\mathcal{A}$. There are at most

$$\frac{(mk)!}{(m!)^k} \leq (ke)^{mk}$$

equipartitions of $X = \{x_1, x_2, ..., x_{mk}\}$ into k classes $A_1, A_2, ..., A_k$.
We have got that the probability, that for some equipartition $\mathcal{A}$ there exists a k–party protocol which computes f with communication $m - \alpha$ is at most

$$2^{2^{m(k-1)}(2mk^2 - 2^\alpha)}.$$

Consequently, for $\alpha = \log m + 2\log k + 1$, the probability that a randomly chosen f can be computed by a k-party protocol with $m - \alpha$ communication is double–exponentially small. ∎

4. PROOF OF THEOREM 2

Let $\mathcal{A} = \{A_1, A_2, ..., A_k\}$ be a fixed equipartition of X. Suppose that there exists a k–party protocol which correctly computes function f on a fraction of at least

$$\frac{1}{2} + \varepsilon$$

of all inputs, communicating $m - \alpha$ bits.
First, we need a combinatorial lemma:

Lemma 13 . *Let $u = \frac{\varepsilon}{2} 2^{m(k-1)+\alpha}$. Then there exists a cylinder–intersection Q of size at least u, such that the protocol correctly computes f on a fraction of at least $\frac{1}{2} + \frac{\varepsilon}{2}$ of Q.*

Proof. Suppose that the statement does not hold. Then the fraction of the inputs, for which f is correctly computed is less than

$$\frac{1}{2} + \frac{\varepsilon}{2}$$

in all cylinder–intersections, which have size at least u. Then the missing

$$\frac{\varepsilon}{2}$$

fraction of all inputs:

$$\frac{\varepsilon}{2} 2^{mk}$$

should be computed correctly in cylinder intersections of size less than u.
However, there are at most $2^{m-\alpha}$ cylinder–intersections, so if even all of them has size $u - 1$, and f is correctly computed on them, then less than

$$2^{m-\alpha} u = \frac{\varepsilon}{2} 2^{mk}$$

inputs are computed correctly on these small cylinder–intersections, contradiction. ∎
So, there exists a Q of size at least u on which at least

$$\frac{1}{2} + \frac{\varepsilon}{2}$$

fraction are computed correctly. This means, that on a

$$\frac{1}{2} + \frac{\varepsilon}{2}$$

part of Q f is constant 0 or constant 1.
By the *Chernoff–bound* [ES], the probability that a random f is constant on at least a

$$\frac{1}{2} + \frac{\varepsilon}{2}$$

fraction of Q is at most

$$2e^{-\frac{\varepsilon^3}{4}2^{(k-1)m+\alpha}}.$$

By Theorem 12 , there are at most

$$\binom{2^{mk}}{k2^{m(k-1)}} \leq 2^{mk^2 2^{m(k-1)}}$$

cylinder–intersections in $\{0,1\}^{A_1 \cup A_2 \cup \ldots \cup A_k}$.
So, the probability, that a random f is constant on the

$$\frac{1}{2} + \frac{\varepsilon}{2}$$

fraction of *at least one* cylinder–intersection of size at least u is at most

$$2^{mk^2 2^{m(k-1)}} 2e^{-\frac{\varepsilon^3}{4}2^{(k-1)m+\alpha}} \leq 2^{2^{m(k-1)}(mk^2 - \frac{\varepsilon^3}{4}2^{\alpha})},$$

for a fixed $\mathcal{A}$.
There are at most

$$\frac{(mk)!}{(m!)^k} \leq (ke)^{mk}$$

equipartitions of $X = \{x_1, x_2, ..., x_{mk}\}$ into k classes $A_1, A_2, ..., A_k$.
We have got that the probability, that for some equipartition $\mathcal{A}$ there exists a k–party protocol which computes f with communication $m - \alpha$, correctly on the fraction of

$$\frac{1}{2} + \varepsilon$$

of all inputs, is at most

$$2^{2^{m(k-1)}(2mk^2 - \frac{\varepsilon^3}{4}2^{\alpha})}.$$

Consequently, for $\alpha = \log m + 2\log k + 3\log\frac{1}{\varepsilon}+2$, the probability, that a randomly chosen f can be computed by a k-party protocol with $m - \alpha$ communication, is double–exponentially small. ∎

REFERENCES

[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.

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

[ES] P. Erdős, J. Spencer: Probabilistic Methods in Combinatorics, Academic Press, New York and London, 1974.

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

[PS] C.H. Papadimitriou, M. Sipser: Communication Complexity, Proc. 14th ACM STOC, 1982, pp. 196–200