GEOMETRICAL REALIZATION OF SET SYSTEMS AND PROBABILISTIC COMMUNICATION COMPLEXITY

N. Alon* - P. Frankl** - V. Rödl**

*Department of Mathematics, Tel Aviv University, Tel Aviv and
Bell Communications Research, Morristown, New Jersey 07960 USA
**AT&T Bell Labs, Murray Hill, New Jersey, 07974 USA

ABSTRACT

Let $d = d(n)$ be the minimum d such that for every sequence of n subsets $F_1, F_2, \ldots, F_n$ of $\{1, 2, \ldots, n\}$ there exist n points $P_1, P_2, \ldots, P_n$ and n hyperplanes $H_1, H_2, \ldots, H_n$ in R^d such that P_j lies in the positive side of H_i *iff* $j \in F_i$. Then

$$n/32 \leq d(n) \leq (\frac{1}{2} + o(1)) \cdot n \ . \qquad (1)$$

This implies that the probabilistic unbounded-error 2-way complexity of almost all the Boolean functions of $2p$ variables is between p-5 and p, thus solving a problem of Yao and another problem of Paturi and Simon.

The proof of (1) combines some known geometric facts with certain probabilistic arguments and a theorem of Milnor from real algebraic geometry.

1. INTRODUCTION

Let $N = \{1, 2, \ldots, n\}$ and let $\mathrm{F} = \{F_1, F_2, \ldots, F_m\}$ be a family of m subsets of N. We say that F is *realizable* in the d-dimensional Euclidean space R^d if there exist n points $P_1, P_2, \ldots, P_n$ and m hyperplanes $H_1, H_2, \ldots, H_m$ such that P_j lies in the positive side of H_i if and only if $j \in F_i$. Define $d(\mathrm{F})$, the *dimension of* F, to be the minimal dimension d such that F is realizable in R^d. Also put $d(n, m) = \max\{d(\mathrm{F}) : \mathrm{F}$ is a family of m subsets of $N\}$. Clearly $d(n, m) \leq n-1$ (simply take n points in general position in R^n).

It is also easy to see that

$$d(n\ ,m) \geq \left\lfloor \log_2 m \right\rfloor .$$

Indeed, if $X \subseteq N$ is separated in all $2^{|X|-1}$ possibilities by the subsets of F (i.e., for every partition of $X = X_1 \cup X_2$ there is an $F \in \mathrm{F}$ such that $F \cap X = X_1$ or $F \cap X = X_2$), then, by Radon's theorem (cf. e.g. [Gr, p. 16] $d(\mathrm{F}) \geq |X|-1$.

In this paper we prove:

Theorem 1.1.

If $n, m \to \infty$ and $\log_2 m = o(n)$ *then* $d(n,m) \leq (\frac{1}{2} + o(1))n$ (i)

Put $d = d(n, m)$ *then* (ii)

$$2^{n^3 + O(n^2) + H(\frac{d}{n}) \cdot n \cdot m + m} \geq 2^{n \cdot m} \ ,$$

where $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function.

Put $d = d(n, m)$, then for every integer $1 \leq h \leq nm$ (iii)

$$(8\left\lceil \frac{nm}{h} \right\rceil)^{(n+m)d+h+m} \geq 2^{nm} \ .$$

We specify two special cases separately.

Corollary 1.2.

$$n/32 \leq d(n, n) \leq (\frac{1}{2} + o(1)) \cdot n \ .$$

(The constant 1/32 can be somewhat improved).

Corollary 1.3.

If $m/n^2 \to \infty$ and $(log_2 m)/n \to 0$ then ,

$$d(n, m) = (\frac{1}{2} + o(1)) \cdot n \ .$$

Corollary 1.2 solves a problem of Paturi and Simon [PS], who showed that $\left\lfloor \log_2 n \right\rfloor \leq d(n, n) \leq n-1$ and asked if one can prove a superlogarithmic lower bound for $d(n, n)$. As shown in Section 5, Corollary 1.2 also enables us to improve the lower bound of [PS] for the maximal possible unbounded-error probabilistic communication complexity of a Boolean function of $2p$ bits from $\Omega(\log p)$ to $p-5$. (This complexity is always at most p).

Yao [Ya] introduced a model of bounded-error probabilistic communication complexity and showed that there are functions of $2p$ Boolean variables whose complexity is $\Omega(log\ p)$. This was improved by Vazirani [V] to $\Omega(p/\log p)$. Our $p-5$ lower bound applies to this model, as well, and improves the bound to $\Omega(p)$. Moreover, our proof shows that the (bounded or unbounded) probabilistic communication complexity of almost every Boolean function on $2p$ variables is between $p-5$ and p. This answers a question raised in [Ya]. A slightly weaker result for the bounded error case has been recently obtained also by Chor and Goldreich [CG], who showed that almost every Boolean function of $2p$ variables has an $\Omega(p)$ bounded error probabilistic communication complexity.

Our paper is organized as follows. In Sections 2 and 3 we show how to use the moment curve together with some simple probabilistic arguments to prove the upper bound part of Theorem 1.1. In Section 4 we combine some recent results of Goodman and Pollack [GP2] with some of the results of [Al] and simple counting arguments to prove the lower bounds. The results of [GP2] and [Al] both follow from a theorem of Milnor from real algebraic geometry. In Section 5 we discuss the application to probabilistic communication complexity. The final Section 6 contains some concluding remarks and related results.

0272-5428/85/0000/0277$01.00 © 1985 IEEE

2. TWO PROBABILISTIC LEMMAS

Let $\vec{a} = (a_1, a_2, \ldots, a_n)$ be a sequence, where $a_i \in \{1,-1\}$. Let $\sigma(\vec{a})$ denote the number of sign changes in $\vec{a}$, i.e. $\sigma(\vec{a}) = |\{i : 1 \le i \le n-1 \quad \text{and} \quad a_i \ne a_{i+1}\}|$. By a *random permutation* we mean a random variable π such that $Pr(\pi = \rho) = 1/n!$ for all $\rho \in S_n$, the symmetric group on $N = \{1, 2, \ldots, n\}$. For $\rho \in S_n$ put $\rho(\vec{a}) = (a_{\rho(1)}, a_{\rho(2)}, \ldots, a_{\rho(n)})$.

Lemma 2.1.

For every $\epsilon > 0$ there exists a $\delta = \delta(\epsilon) > 0$ such that if π is a random permutation then

$$P_r(\sigma(\pi(\vec{a})) > (\frac{1}{2} + \epsilon)n) < (1 + \delta)^{-n} .$$

Proof.

Suppose that $\vec{a}$ has k plus 1's and $n-k$ minus 1's and fix t, $t > (\frac{1}{2} + \epsilon)n$. Let us compute the number of permutations $\rho \in S_n$ with $\sigma(\rho(\vec{a})) = t$. To define such a ρ we have to choose the order among the k plus 1's and among the -1's (altogether $k!(n-k)!$ choices), then to decide if the first position is a plus or a minus, and finally to determine the $\lceil t/2 \rceil$, $\lfloor t/2 \rfloor$ cutting points (after which the sign changes take place). For definiteness we consider the case $t = 2s$, (the case $t = 2s + 1$ can be bounded similarly). In this case the cutting points can be chosen in $\binom{k-1}{s}\binom{n-k-1}{s}$ ways, and hence

$$Pr(\sigma(\pi(\vec{a})) = t) = \frac{k!\,(n-k)!2\binom{k-1}{s}\binom{n-k-1}{s}}{n!} = \frac{2\binom{k-1}{s}\binom{n-k-1}{s}}{\binom{n}{k}}$$

One can easily check that for fixed n and s the right hand side is maximized for $k = \lfloor n/2 \rfloor$ and then it is still bounded by $\simeq 2^{(H(\frac{1}{2}+\epsilon)-1)\cdot n}$, where $H(x) = -x\log x-(1-x)\log(1-x)$ is the binary entropy function. Since $Pr(\sigma(\pi(\vec{a})) > (\frac{1}{2} + \epsilon)n) = \sum\{Pr(\sigma(\pi(\vec{a})) = t : t > (\frac{1}{2} + \epsilon)n)\}$ the desired result follows.

□

Suppose now that $\mathrm{F} = \{F_1, F_2, \ldots, F_m\}$ is a family of subsets of $N = \{1, 2, \ldots, n\}$. Let $M = (m_{ij})$ be the *incidence matrix* of F defined as follows; it is an m by n matrix where $m_{ij} = +1$ if $j \in F_i$ and $m_{ij} = -1$ if $j \notin F_i$.

Lemma 2.2.

Let ϵ and δ be as in Lemma 2.1 and suppose $m < (1 + \delta)^n$. Then there is a permutation of the columns of M such that the number of sign changes in each row is $\le (\frac{1}{2} + \epsilon) \cdot n$.

Proof:

Consider a random permutation π of the columns of M. Call a permuted row *bad* if the number of sign changes in it is $> (\frac{1}{2} + \epsilon) \cdot n$. By Lemma 2.1 the probability that each fixed row is bad is $< (1 + \delta)^{-n}$. Hence, the expected number of bad rows is smaller than $m(1 + \delta)^{-n} < 1$, thus the desired permutation exists.

□

Remark 2.3.

Note that the bound $(\frac{1}{2} + \epsilon) \cdot n$ cannot be substantially improved even for $m = n$. Indeed, if M is an n by n Hadamard matrix then the total number of sign changes in its permuted rows is precisely $\frac{n}{2}(n-1)$. Hence under any permutation at least one row has at least $(n-1)/2$ sign changes.

3. THE UPPER BOUND

Let $\epsilon > 0$ be arbitrarily small and let $\delta = \delta(\epsilon)$ be defined by Lemma 2.1. We will show that if $m < (1 + \delta)^n$ then $d(n, m) \le (\frac{1}{2} + \epsilon)n$. Note that this proves part (i) of Theorem 1.1. Suppose $m < (1 + \delta)^n$ and let $\mathrm{F} = \{F_1, F_2, \ldots, F_m\}$ be a family of subsets of N. By Lemma 2.2 we can assume that each row of the incidence matrix of F has at most $(\frac{1}{2} + \epsilon) \cdot n$ sign changes. Let $d = \left\lfloor (\frac{1}{2} + \epsilon) \cdot n \right\rfloor$. We will show, using the well-known moment curve (cf. e.g. [Gr, pp. 61-63], that F is realizable in R^d. Let $t_1 < t_2 < \ldots < t_n$ be real numbers and put $P_j = (t_j, t_j^2, \ldots, t_j^d)$. These will be the points of our realization. Consider the i^{th} row of the incidence matrix of F. Suppose that the sign changes in this row appear after positions $j_1 < j_2 < \ldots < j_r$. Then $r \le d$. Choose real numbers y_{j_k} satisfying $t_{j_k} < y_{j_k} < t_{j_{k+1}}$ and consider the polynomial $p_i(t) = \prod_{k=1}^{r}(t-y_{j_k})$. Since $p_i(t)$ has degree at most d, we can write $p_i(t) = \sum_{k=0}^{d} a_k t^k$. Let H_i be the hyperplane $H_i = \{(x_1, x_2, \ldots, x_d) : \sum_{k=1}^{d} a_k \cdot x_k = -a_0\}$. Clearly $y_{j_1}, \ldots, y_{j_r}$ are the only sign changes of the polynomial $p_i(t)$. Since $p_i(t) = \sum_{k=0}^{d} a_k t^k$, the point $(t, t^2, \ldots, t^d)$ lies on the positive side of H_i if and only if $p_i(t) > 0$. Hence the hyperplane H_i separates the points $P_1, P_2, \ldots, P_n$ according to the sign pattern of $(m_{ij})_{j=1}^{n}$ and after adjusting the sign of the H_i-s, if necessary, we conclude that the H_i-s and the P_j-s form a realization of F, as desired.

□

4. THE LOWER BOUNDS

We first note that if a family F is realizable in R^d by the points $P_1, \ldots, P_n$, and the hyperplanes $H_1, \ldots, H_m$ then it is also realizable by $P_1, \ldots, P_n$ and the same H_i-s, whenever P_j is sufficiently close to P_j. Hence we can always assume that the points P_j of a realization are in general position in R^d. Let us call two ordered sets $P_1, P_2, \ldots, P_n$ and $Q_1, Q_2, \ldots, Q_n$ of points in general position in R^d *equivalent* if they can be partitioned by hyperplanes in precisely the same way, i.e., there exists a hyperplane H separating $P_{j_1}, \ldots, P_{j_r}$ from the rest of the P_j-s if and only if there exists a hyperplane H' separating $Q_{j_1}, \ldots, Q_{j_r}$ from the rest of the Q_j-s. For a sequence $(P_0, \ldots, P_d)$ of points in R^d with $P_i = (x_{i1}, \ldots, x_{id})$, we say that it has a *positive orientation*, written $P_0 \ldots P_d > 0$, if

$$det\,(x_{ij}) > 0$$

where $x_{io} = 1$ for each i. $P_o \ldots P_d < 0$ is defined similarly. The *order type* of an ordered set of points $P_1, P_2, \ldots, P_n$ (in general position) in R^d is the set of all d+1-tuples $j_1 < j_2 < \ldots < j_{d+1}$ such that $P_{j_1} \ldots P_{j_{d+1}} > 0$. It is easy and well known (see e.g. [GP1]) that if $P_1 \ldots P_n$ and $Q_1 \ldots Q_n$ have the same order type then they are equivalent. Very recently, Goodman and Pollack

[GP2] have found a clever (and simple) way to apply a result of Milnor [Mi] from real algebraic geometry in order to obtain an asymptotically best possible upper bound for the number of order types (and hence for the number of equivalence classes) of n labeled points in R^d. Here we need an easy modification of their result, proved in [Al].

Lemma 4.1 ([Al])

For every d, n the number of equivalence classes of n labeled points in R^d is at most $2^{n^3+O(n^2)}$.

Another known result we will need is the following (see, e.g. [Ha] or [Za]).

Lemma 4.2 ([Ha, [Za])

The number of ways to partition n points in R^d into two disjoint subsets separated by a hyperplane is at most

$$\sum_{i=0}^{d} \binom{n-1}{i} \left(\leq 2^{H(\frac{d}{n})\cdot n} \right) . \qquad \square$$

We can now prove part (ii) of Theorem 1.1. By Lemma 4.2, every given point set of n points in R^d realizes at most $(2 \cdot 2^{H(\frac{d}{n})\cdot n})^m$ ordered sequences of m subsets of $N = \{1, 2, \ldots, n\}$. Thus, by Lemma 4.1, the total number of sequences of m subsets of N that can be realized by n points in R^d is at most $2^{n^3+O(n^2)} \cdot 2^{m+H(\frac{d}{n})nm}$.

Consequently, for $d = d(n, m)$, we have $2^{n^3+O(n^2)+m+H\left(\frac{d}{n}\right)\cdot nm} \geq 2^{nm}$, since every ordered family of m subsets of N is realizable in R^d. This proves part (ii) of Theorem 1.1.

$\square$

To prove part (iii) of Theorem 1.1 we need another result from [Al]. Let $P_1 = P_1(x_1, x_2, \ldots, x_t)$, $P_2 = P_2(x_1, x_2, \ldots, x_t)$, $\ldots$, $P_m = P_m(x_1, x_2, \ldots, x_t)$ be real polynomials. For $c = (c_1, c_2, \ldots, c_n) \in R^n$ and $1 \leq j \leq m$, let $P_j(c)$ denote $P_j(c_1, c_2, \ldots, c_n)$. Assume $P_j(c) \neq 0$ for all $1 \leq j \leq m$. The *sign-pattern* of the P_j-s at c is the m-tuple $(\epsilon_1, \epsilon_2, \ldots, \epsilon_m) \in \{-1, 1\}^m$, where $\epsilon_j = \text{sign } P_j(c)$. The total number of sign patterns as c ranges over all points of R^n for which $P_j(c) \neq 0$ for all $1 \leq j \leq m$, denoted by $s(P_1, P_2, \ldots, P_m)$, is clearly at most 2^m. The following result is an easy modification of Theorem 2.2 of [Al], and can be derived, as in [Al], from the theorems of Milnor [Mi] and Thom [Th].

Lemma 4.3

Let $P_1, P_2, \ldots, P_m$ be as above and let $d_j = deg\ P_j\ (\geq 1)$ be the degree of P_j, $1 \leq j \leq m$. Put $J = \{1, 2, \ldots, m\}$ and let $J = J_1 \cup J_2 \cup \ldots \cup J_h$ be a partition of J into h pairwise disjoint parts. Define $k = 2 \max_{1 \leq i \leq h} \left(\sum_{j \in J_i} d_j \right)$. Then

$$s(P_1, P_2, \ldots, P_m) \leq k \cdot (2k-1)^{t+h-1} . \qquad \square$$

We can now prove part (iii) of Theorem 1.1. Suppose $\mathrm{F} = (F_1, F_2, \ldots, F_m)$ is a sequence of m subsets of $N = \{1, 2, \ldots, n\}$ that is realizable in R^d. Then we can associate d real variables $x_{j1}, x_{j2}, \ldots, x_{jd}$ to each $1 \leq j \leq n$, and $d+1$ real variables $y_{j0}, y_{i1}, y_{i2}, \ldots, y_{id}$ for each $1 \leq i \leq m$ such that $y_{i0} + \sum_{r=1}^{d} x_{jr}\, y_{ir} > 0$ if $j \in F_i$ and $y_{i0} + \sum_{r=1}^{d} x_{jr}\, y_{ir} < 0$ if $j \notin F_i$. Thus the total number of sequences of m subsets of N that are realizable in R^d is bounded by the number of sign patterns of $n \cdot m$ quadratic polynomials with $dn + (d+1)\, m$ variables. The assertion of part (iii) of Theorem 1.1 now follows from Lemma 4.3

$\square$

Corollary 1.2 follows from part (i) of Theorem 1.1 and part (iii) with $h = 2d \cdot n$, $m = n$. Corollary 1.3 follows from parts (i) and (ii) of Theorem 1.1.

5. PROBABILISTIC COMMUNICATION COMPLEXITY

The model of [PS] is similar to that of [Ya], and considers the following problem. Two processors P_0 and P_1 wish to compute a Boolean function $f : \{0, 1\}^p \times \{0, 1\}^p \rightarrow \{0, 1\}$ of two arguments, each consisting of p bits. The first argument, x_0, is know only to P_0 and the second, x_1, only to P_1. In order to compute f, P_0 and P_1 communicate by sending each other in turns sequences of bits according to some (probabilistic) protocol ψ. Both processors have unlimited local computing power and can realize an arbitrary probability distribution over the set of messages they transmit. The last message is always sent by P_1 and is the output produced. We say that the protocol ψ *outputs* bit b if the probability that their last produced bit is b is greater than 1/2. The protocol *computes* f if for every x_0, x_1 it outputs b if and only if $f(x_0, x_1) = b$. The *communication complexity* C_ψ of ψ is the maximum number of bits transmitted by P_0 and P_1 during the protocol. The *unbounded-error probabilistic communication complexity* C_f of f is $\min\{C_\psi : \psi \text{ computes } f\}$, i.e., the complexity of the most efficient protocol to compute f. It is shown in [PS] that the power of this unrestricted probabilistic model is considerable. E.g., the unbounded error probabilistic communication complexities of the functions $I(x, y) = (x = y)$, $I(x, y) = (x \neq y)$ and $G(x, y) = (x \geq y)$ are all shown to be ≤ 2. On the other hand, it is shown that for some f-s, $C_f = \Omega(\log p)$. Our results imply that for some f-s $p-5 \leq C_f\ (\leq p)$. This follows immediately from Corollary 1.2 and the following result of [PS].

Theorem 5.1. [PS]

Let $f : \{0, 1\}^p \times \{0, 1\}^p \rightarrow \{0, 1\}$ be a Boolean function and let $M = (m_{X_0 X_1})_{X_0, X_1 \in \{0,1\}^p}$ be its matrix. Put $n = 2^p$ and let N be the set of all binary vectors of length p. For every $X_0 \in N$ put $F_{X_0} = \{X_1 : f(X_0, X_1) = 1\}$, and put $\mathrm{F} = \mathrm{F}(f) = \{F_{X_0} : X_0 \in N\}$. Then

$$\left\lceil \log d(\mathrm{F}) \right\rceil \leq C_f \leq \left\lceil \log(d(\mathrm{F})) \right\rceil + 1 .$$

Note that F is a family of n subsets of N and that, conversely, for each such family G there is a Boolean f such that $\mathrm{F}(f) = G$. Hence, combining this theorem with Corollary 1.2 we obtain:

Theorem 5.2

There are functions $f : \{0, 1\}^p \times \{0, 1\}^p \rightarrow \{0, 1\}$ such that

$$C_f \geq p-5 .$$

This settles the problem raised in [PS]. Notice that the proof actually gives that for almost all functions $f : \{0, 1\}^p \times \{0, 1\}^p \rightarrow \{0, 1\}$ $C_f \geq p-5$. (Obviously, for each such f, $C_f \leq p$). This answers a problem raised in [Ya].

6. CONCLUDING REMARKS

1. The *sign-pattern* of an m by n matrix A with real entries $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ is an m by n matrix $Z(A) = (z_{ij})$ of 1-s and -1-s where $z_{ij} = \text{sign } a_{ij}$. For an m by n matrix Z of 1,-1 entries, let $r(Z)$ be the minimum possible rank of a matrix A such that $Z(A) = Z$. Define $r(n, m) = Max\{r(Z) : Z \text{ is an } m \text{ by } n \text{ matrix over } \{1,-1\}\}$. One can easily check that

$$d(n, m) \leq r(n, m) \leq d(n, m) + 1 .$$

Thus our results imply, e.g., that

$$n/32 \leq r(n, n) \leq (1+o(1))\ \frac{n}{2}\ .$$

2. The method of proving the lower bound can be easily extended to the case of realizing the family of sets by points and hypersurfaces defined by polynomials of given degrees. We omit the details.

3. There are certain applications of our results to problems of embedding bipartite graphs on the unit sphere. These will appear in the full version of the paper.

4. It would be nice to determine more precisely the asymptotic behavior of $d(n, n)$. It seems reasonable that

$$d(n, n) = (1+o(1)) \cdot \frac{n}{2}\ .$$

Acknowledgement.

We would like to thank N. Pippenger for telling us about the problem discussed in this paper. We would also like to thank M. Ben-Or and N. Linial for improving and simplifying the proof of Theorem 1.1 part (iii).

REFERENCES

[AL] N. Alon, "The number of polytopes, configurations and real matroids", to appear.

[CG] B. Chor and O. Goldreich, "Unbiased bits from weak sources of randomness", Proc. 26th Focs (1985), to appear.

[GP1] J. E. Goodman and R. Pollack, "Multidimensional sorting", SIAM J. Comput. 12 (1983), 484-507.

[GP2] J. E. Goodman and R. Pollack, "Upper bounds for configurations and polytopes in R^d", to appear.

[Gr] B. Grünbaum, Convex Polytopes, Interscience, Wiley, London 1967.

[Ha] E. F. Harding, "The number of partitions of a set of n points in k dimensions induced by hyperplanes", Proc. Edinburg Math. Soc. 15 (1967), 285-289.

[Mi] J. Milnor, "On the betti numbers of real algebraic varieties", Proc. AMS 15 (1964), 275-280.

[PS] R. Paturi and J. Simon, "Probabilistic communication complexity", Proc. 25th FOCS, Florida (1984), 118-126.

[Th] R. Thom, "Sur l'homologie des variétés algébriques réelles," Differential and Combinatorial Topology, Ed. S.S. Cairns, Princeton Univ. Press, 1965.

[V] U. V. Vazirani, "Towards a strong communication complexity theory or generating quasi-random sequences from two communicating slightly random sources", Proc. 17th STOC, Providence, RI (1985), 366-378.

[Ya] A. C. C. Yao, "Some complexity questions related to distributed computations", Proc. 11th ACM STOC (1979), 209-213.

[Za] T. Zaslavsky, Facing up to Arrangements : Face–Count Formulas for Partitions of Space by Hyperplanes, Memoir 154, American Mathematical Society, Providence, RI (1975).