Probabilistic communication complexity over the reals

Dima Grigoriev

Probabilistic communication complexity over the reals

Dima Grigoriev
CNRS, IRMAR, Université de Rennes
Beaulieu, 35042, Rennes, France
dmitry.grigoryev@univ-rennes1.fr
http://perso.univ-rennes1.fr/dmitry.grigoryev

Abstract

Deterministic and probabilistic communication protocols are introduced in which parties can exchange the values of polynomials (rather than bits in the usual setting). It is established a sharp lower bound $2n$ on the communication complexity of recognizing the $2n$-dimensional orthant, on the other hand the probabilistic communication complexity of its recognizing does not exceed 4. A polyhedron and a union of hyperplanes are constructed in $\mathbb{R}^{2n}$ for which a lower bound $n/2$ on the probabilistic communication complexity of recognizing each is proved. As a consequence this bound holds also for the EMPTINESS and the KNAPSACK problems.

Introduction

Communication complexity (see [15], a survey one can find in [12], [13]) in the usual (bit) setting counts the number of bit exchanges between two (or more) parties who altogether compute a certain function (one of the goals of the communication complexity was to provide a framework to analyze distributed computations and to obtain lower bounds on other complexity ressources). In [2] one can find the relations of the communication complexity with the question of representing a function as a composition of functions of a special form (this question stems from the Hilbert's 13th problem). In [5] the communication complexity of quantum computations was studied.

In the present paper we introduce the model of communication protocols over real (or complex) numbers when the parties exchange the values of polynomials. The variables of polynomials are supposed to be partitioned in two groups: $X = \{X_1, \ldots, X_{n_1}\}, Y = \{Y_1, \ldots, Y_{n_2}\}$, the first party is able to calculate polynomials in X, the second party in Y. It is worthwhile to mention that in [11] a different (less restrictive) concept of a communication protocol was introduced in which the parties can exchange arbitrary real numbers (rather than just values of a given family of polynomials as in the present paper). After the present paper had been submitted the paper [3] has appeared in which a similar algebraic communication protocol was introduced and several lower bounds on the algebraic communication complexity for computing rational functions and recognizing algebraic varieties were established. Unlike [3] we obtain lower bounds on *probabilistic* communication complexity and in addition, for recognizing real *semi-algebraic* sets.

We note that parallel to the numerous customary (boolean or discrete) complexity classes one develops also their continuous (algebraic or semi-algebraic) counterparts (see e. g. [4], [6]). This paper presents an attempt to introduce and study the probabilistic continuous communication complexity.

For illustration of the results obtained in the present paper we consider the KNAPSACK problem: whether for given sets $\{x_1,\dots,x_n\}$ and $\{y_1,\dots,y_n\}$ there exist subsets $I_1, I_2 \subseteq \{1,\dots,n\}$ such that $\sum_{i_1\in I_1} x_{i_1} + \sum_{i_2\in I_2} y_{i_2} = 0$? There is an evident *deterministic* communication protocol for the KNAPSACK problem with the communication complexity $2n$ when two parties just yield $\{x_1,\dots,x_n\}$ and $\{y_1,\dots,y_n\}$, respectively. In Section 4 we show a lower bound $n/4$ in the complex case and $n/2$ in the real case on the *probabilistic* communication complexity for the KNAPSACK problem.

In Section 1 we define the *communication complexity of computing a function* (polynomial for simplicity) and show a lower bound on it being the rank of the matrix of its second derivatives, earlier this matrix in the frames of communication complexity was employed in [1]. This slightly resembles the lower bound on the bit communication complexity being the logarithm of the rank of the communication matrix [15].

In Section 2 we describe the (deterministic) *communication protocols* (respectively, *probabilistic communication protocols*) and relying on this we define the (deterministic) *communication complexity of recognizing a set* (respectively, *probabilistic communication complexity*). As an application of the matrix of the second derivatives we establish a lower bound $n-3$ on a probabilistic communication complexity of recognizing a constructible set in $\mathbb{C}^{2n}$ whose Zariski closure contains the hypersurface $\{f = X_1Y_1+\dots+X_nY_n = 0\}$. As a real counterpart we establish the same bound $n-3$ for a semialgebraic set in $\mathbb{R}^{2n}$ whose euclidean closure has (full) $(2n-1)$-dimensional intersection with the hypersurface $\{f=0\}$.

In Section 3 we demonstrate a possible exponential gap between the deterministic and probabilistic communication complexities. Namely, we prove a (sharp) lower bound n_1+n_2 on the deterministic communication complexity of recognizing the orthant

$$\{(x_1,\dots,x_{n_1},y_1,\dots,y_{n_2})\in\mathbb{R}^{n_1+n_2} : x_i>0, y_j>0, 1\le i\le n_1, 1\le j\le n_2\}.$$

On the other hand, we show that the probabilistic communication complexity of recognizing the orthant does not exceed 4.

In Section 2 the lower bound was established for a set which involves a polynomial f with a big communication complexity of its computation. In Section 4 we construct sets defined by linear contraints which nevertheless have big probabilistic communication complexity (clearly, any linear function has the communication complexity of its computation at most 2). Namely, we consider the polyhedron $\{X_i+Y_i>0, 1\le i\le n\}\subset\mathbb{R}^{2n}$ and the arrangement $\cup_{1\le i,j\le n}\{X_i+Y_j=0\}\subset\mathbb{R}^{2n}$ and for each of both prove a lower bound $n/2$ on the probabilistic communication complexity of its recognizing. For the complex arrangement $\cup_{1\le i,j\le n}\{X_i+Y_j=0\}\subset\mathbb{C}^{2n}$ we establish a lower bound $n/4$. As applications the obtained lower bounds imply the same bounds for the EMPTINESS problem, i. e whether $\{x_1,\dots,x_n\}\cap\{y_1,\dots,y_n\}=\emptyset$, and for the KNAPSACK problem.

1 Lower bound on the communication complexity of computing a function

First we describe computational models for the communication complexity over complex or real numbers. Let two families of variables $X = \{X_1,\dots,X_{n_1}\}$ and $Y = \{Y_1,\dots,Y_{n_2}\}$ be given. As usually in communication complexity studies, there are two parties. We assume that one party is able to calculate polynomials $a_1(X),\dots,a_{r_1}(X)$ in X and the second party is able to calculate polynomials $b_1(Y),\dots,b_{r_2}(Y)$

in Y. Then the result is obtained by means of calculating suitable polynomials $P_1(a_1(X),\dots,a_{r_1}(X),b_1(Y),\dots,b_{r_2}(Y)),\dots,P_N(a_1(X),\dots,a_{r_1}(X),b_1(Y),\dots,b_{r_2}(Y))$. The goal is to minimize r_1+r_2 viewed as a measure of communication complexity.

We study the communication complexity of two problems: computing a polynomial $g(X,Y)$ and recognizing a subset S in (n_1+n_2)-dimensional complex or real space.

Definition 1.1 *A polynomial $g(X,Y)$ has a communication complexity $c(g)$ less or equal to r_1+r_2 if $g = P(a_1(X),\dots,a_{r_1}(X),b_1(Y),\dots,b_{r_2}(Y))$ for appropriate polynomials $P,a_1,\dots,a_{r_1},b_1,\dots,b_{r_2}$.*

Obviously, the communication complexity of g does not exceed n_1+n_2.

By $H(g)$ denote $n_1\times n_2$ matrix of the second derivatives $(\frac{\partial^2 g}{\partial X_i\partial Y_j})$, by $H(P)$ denote $r_1\times r_2$ matrix $(\frac{\partial^2 P}{\partial a_{i_1}\partial b_{j_1}})$, by the Jacobian $J(a_1,\dots,a_{r_1})$ denote $n_1\times r_1$ matrix of the first derivatives $(\frac{\partial a_{i_1}}{\partial X_i})$, similar $J(b_1,\dots,b_{r_2})=(\frac{\partial b_{j_1}}{\partial Y_j})$. Then we have

$$H(g)=J(a_1,\dots,a_{r_1})H(P)(J(b_1,\dots,b_{r_2}))^T.$$

Lemma 1.2 *(cf. [1]) In the notations of Definition 1.1 we have*

$$c(g)\geq min\{r_1,r_2\}\geq rk(H(g)).$$

Corollary 1.3 $c(f=X_1Y_1+\dots+X_nY_n)\geq n$

To deal in the sequel with communication protocols we need the following statement generalizing the latter corollary.

Lemma 1.4 *Let a polynomial g be a multiple of f. Then $rk(H(g))\geq n-3$.*

Proof. We write $g=f^mh$ where f does not divide h (evidently, f is absolutely irredicible when $n\geq 2$, we assume here that $n_1=n_2=n$). We have

$$H(g)=mf^{m-1}h\left(\frac{\partial^2 f}{\partial X_i\partial Y_j}\right)+f^m\left(\frac{\partial^2 h}{\partial X_i\partial Y_j}\right)+$$

$$m(m-1)f^{m-2}h\left(\frac{\partial f}{\partial X_i}\right)\left(\frac{\partial f}{\partial Y_j}\right)+mf^{m-1}\left(\frac{\partial f}{\partial X_i}\right)\left(\frac{\partial h}{\partial Y_j}\right)+mf^{m-1}\left(\frac{\partial h}{\partial X_i}\right)\left(\frac{\partial f}{\partial Y_j}\right).$$

Each of the latter three matrices has rank at most 1, so it suffices to verify that the sum of the former two matrices divided by f^{m-1} is non-singular, it equals

$$M=mh\left(\frac{\partial^2 f}{\partial X_i\partial Y_j}\right)+f\left(\frac{\partial^2 h}{\partial X_i\partial Y_j}\right)$$

We have $\det(M)=(mh)^n+ff_1$ for a certain polynomial f_1, hence $\det(M)\neq 0$. ∎

It would be interesting to clarify, whether one can majorate $c(g)$ via an appropriate function in $rk(H(g))$?

2 Probabilistic communication protocols

Now we define a *communication protocol* for recognizing a set S. We consider two cases: $S \subset \mathbb{C}^{n_1+n_2}$ is a constructible set or $S \subset \mathbb{R}^{n_1+n_2}$ is a semialgebraic set. A protocol is a rooted tree, and to its root an input $(x,y) = (x_1,\dots,x_{n_1},y_1,\dots,y_{n_2})$ is attached. To every vertex v of the tree (including the root, but excluding the leaves) either a certain polynomial $a_v(X)$ or a polynomial $b_v(Y)$ is attached (so, it is calculated either by the first party or by the second party, respectively). To every vertex v (of a depth r) leads a unique path from the root, denote by $q_1(X,Y),\dots,q_r(X,Y)$ the polynomials attached to the vertices $v_1,\dots,v_r = v$ along this path, thus for every $1 \leq k \leq r$ either $q_k(X,Y) = a_{v_k}(X)$ or $q_k(X,Y) = b_{v_k}(Y)$, respectively. In addition, to the vertex v a family of *testing polynomials* $P_{v,1}(Q_1,\dots,Q_r),\dots,P_{v,N_v}(Q_1,\dots,Q_r)$ is assigned. Similar to the usual decision trees (see e.g. [14], [9], [10]) the protocol ramifies at v according to the set of the signs $sgn(P_{v,1}(q_1(x,y),\dots,q_r(x,y))),\dots,sgn(P_{v,N_v}(q_1(x,y),\dots,q_r(x,y)))$. Similar to decision trees in the complex case the sign can attain two values: $=,\neq$, in the real case three values: $=,<,>$. To every leaf a label either "accept" or "reject" is assigned which provides an output of the protocol. To the protocol naturally corresponds a decision tree (without restrictions on the degrees of testing polynomials). To any input (x,y) corresponds a unique leaf of the protocol and a path leading to this leaf, according to the signs of testing polynomials: the output assigned to the leaf is "accept" if and only if $(x,y) \in S$.

The *communication complexity of the recognizing protocol* is defined as its depth. We note that the communication complexity counts just the number of the polynomials $a_{v_i}(X)$ or $b_{v_i}(Y)$, respectively, calculated (separately) by each of both parties in several rounds along a path of the protocol and ignores the (jointly) calculated polynomials $P_{v,1},\dots,P_{v,N_v}$.

Now we introduce *probabilistic communication protocols*. One can define it similar to probabilistic decision trees (cf. [14], [9], [8], [10]) as a finite family $C = \{C_i\}_i$ of communication protocols C_i, chosen with a certain probability $p_i \geq 0$, where $\sum_i p_i = 1$. As for decision trees we require that a probabilistic communication protocol for any input returns a correct output with the probability greater than $2/3$ (we suppose that a certain continuous probabilistic measure is fixed in the ambient space, e.g. one can take the Gaussian measure). The maximal depth of communication protocols which constitute a probabilistic communication protocol is called the *probabilistic communication complexity*.

First we consider probabilistic communication protocols over complex numbers.

Proposition 2.1 *The probabilistic communication complexity of an $(2n-1)$-dimensional constructible set $W \subset \mathbb{C}^{2n}$ such that its Zariski closure $\overline{W}$ contains the hypersurface $U = \{f = X_1Y_1 + \dots + X_nY_n = 0\}$ is greater or equal to $n-3$.*

Proof. Let a probabilistic communication protocol C recognize W. Among communication protocols which constitute C there exists C_0 such that it gives the correct outputs for at least of $1/3$ of the points from U and for at least of $1/3$ of the points outside of U (in fact, for the arguments below, instead of $1/3$ any positive constant would suffice).

Distinguish in the decision tree corresponding to C_0 a (unique) path along which all the signs in the ramifications are $\neq$. Denote by $\{P_j(q_1(x,y),\dots,q_r(x,y))\}_{1\leq j\leq N}$ the collection of all the testing polynomials along this path, clearly r does not exceed the communication complexity of C_0. Denote $P = \prod_{1\leq j\leq N} P_j(q_1,\dots,q_r)$. Then the inputs from the Zariski-open set $V = \{(x,y) : P(x,y) \neq 0\} \subset \mathbb{C}^{2n}$ follow this path in C_0.

Due to the choice of C_0 we conclude that f divides P. Indeed, C_0 rejects all the points from a suitable (constructive) subset of $\mathbb{C}^{2n}$ of the dimension $2n$ because C_0 rejects a subset of a positive (namely, at least $1/3$) measure, whence if f did not divide P then C_0 would reject all the points of U except for its certain (constructive) subset of the dimension at most $2n-2$, but on the other hand, C_0 should accept a subset of a positive measure (at least $1/3$) from U. Therefore, Lemma 1.4 and Lemma 1.2 imply that $r \geq c(P) \geq rk(H(P)) \geq n-3$. ∎

For a semialgebraic set $S \subset \mathbb{R}^{n_1+n_2}$ denote by $\partial(S) \subset \mathbb{R}^{n_1+n_2}$ its boundary, being a semialgebraic set as well. The following proposition is a real counterpart of Proposition 2.1.

Corollary 2.2 *The probabilistic communication complexity of a semialgebraic set S such that $dim(\partial(S) \cap U) = 2n-1$ is greater or equal to $n-3$.*

Proof. For any communication protocol C_i from C consider the product $P^{(C_i)} = \prod_{1 \leq j \leq N} P_j$ of all the testing polynomials from C_i (cf. the proof of Proposition 2.1 where a similar product of the polynomials along a particular path was taken). For any point $u \in \partial(S) \cap U$ there exists C_0 such that $P^{(C_0)}(u) = 0$, otherwise all the points from an appropriate ball centered at u would get the same output for all communication protocols C_i from C which would contradict the definition of the boundary. Hence there exists C_0 for which f divides $P^{(C_0)}$. Therefore, we complete the proof as at the end of Proposition 2.1. ∎

3 Communication complexity of recognizing the orthant

Now we proceed to estimating the communication complexity of the orthant $T = \{(x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}) \in \mathbb{R}^{n_1+n_2} : x_i > 0, y_j > 0, 1 \leq i \leq n_1, 1 \leq j \leq n_2\}$. For this goal we use infinitesimals $\epsilon_1 > \cdots > \epsilon_{n_1+n_2} > 0$ (see e.g.[7], [9], [8], [10]). Namely, denote by $\mathbb{R}_i = \widetilde{\mathbb{R}(\epsilon_1, \ldots, \epsilon_i)}$ by recursion on i the real closure of the field $\mathbb{R}(\epsilon_1, \ldots, \epsilon_i)$, for the base of recursion we put $\mathbb{R}_0 = \mathbb{R}$. Then ϵ_{i+1} is transcendental over $\mathbb{R}_i$ and for any positive element $0 < d \in \mathbb{R}_i$ we have $0 < \epsilon_{i+1} < d$.

For a polynomial $g \in \mathbb{R}[X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}]$ denote by $lt(g)$ its *least term* with respect to the following (lexicographical) ordering: take the terms with a minimal degree in Y_{n_2}, among them with a minimal degree in Y_{n_2-1} and so on. If $lt(g) = g_0 X_1^{i_1} \cdots X_{n_1}^{i_{n_1}} Y_1^{j_1} \cdots Y_{n_2}^{j_{n_2}}$ for a certain $g_0 \in \mathbb{R}$, we call $(i_1, \ldots, i_{n_1}, j_1, \ldots, j_{n_2})$ the exponent vector of $lt(g)$. Take $e_1, \ldots, e_{n_1+n_2} \in \{-1, 1\}$, then we have (cf. [9], [8], [10])

$$sgn(g(e_1\epsilon_1, \ldots, e_{n_1+n_2}\epsilon_{n_1+n_2})) = sgn(lt(g)(e_1\epsilon_1, \ldots, e_{n_1+n_2}\epsilon_{n_1+n_2})) \quad (1)$$

Lemma 3.1 *Let $g_1, \ldots, g_s \in \mathbb{R}[X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}]$ and $P_1, \ldots, P_N \in \mathbb{R}[G_1, \ldots, G_s]$. Then among the exponent vectors of the least terms of $P_1(g_1, \ldots, g_s), \ldots, P_N(g_1, \ldots, g_s)$ there are at most s linearly independent.*

Proof. We claim that if exponent vectors of any family of polynomials $h_1, \ldots, h_t \in \mathbb{R}[X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}]$ are linearly independent then $h_1, \ldots, h_t$ are algebraically independent over $\mathbb{R}$. Indeed, denote the exponent vectors of $lt(h_1), \ldots, lt(h_t)$ by $l_1, \ldots, l_t$, respectively, and denote by L the $t \times (n_1+n_2)$ matrix with the rows $l_1, \ldots, l_t$, then for any polynomial $P = \sum_K p_K G^K \in \mathbb{R}[G_1, \ldots, G_t]$ the exponent of the least term of $P(h_1, \ldots, h_t)$ coincides with the least vector among the pairwise distinct vectors KL for all $K \in \mathbb{Z}^t$ such that $p_K \neq 0$. The proved claim entails the lemma immediately. ∎

Theorem 3.2 *The communication complexity of recognizing the orthant T (as well as its closure $\overline{T}$ in the euclidean topology) is greater or equal to $n_1 + n_2$.*

Proof. Let a communication protocol C_0 recognize T (the arguing for $\overline{T}$ is similar). Using the Tarski's transfer principle (see e. g. [7], [9], [8], [10]) one can extend the inputs of C_0 over the field $\mathbb{R}_{n_1+n_2}$, then C_0 recognizes the set $T^{(\mathbb{R}_{n_1+n_2})} = \{(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \in \mathbb{R}^{n_1+n_2}_{n_1+n_2} : x_i > 0, y_j > 0, 1 \leq i \leq n_1, 1 \leq j \leq n_2\}$. Take in C_0 the path which follows the input $(\epsilon_1, \dots, \epsilon_{n_1+n_2}) \in T^{(\mathbb{R}_{n_1+n_2})}$. Let r be the length of this path and denote by $q_1(X,Y), \dots, q_r(X,Y)$ the polynomials attached to the vertices along this path (we use the notations introduced in Section 2 and recall that every q_i depends either on X or on Y, although the latter is not used in the proof of the Theorem, cf. Remark 3.3 below). Let $P_1(q_1, \dots, q_r), \dots, P_N(q_1, \dots, q_r)$ be all the testing polynomials along this path.

Lemma 3.1 implies that among the exponent vectors of $lt(P_1(q_1, \dots, q_r)), \dots, lt(P_N(q_1, \dots, q_r))$ there are at most r linearly independent $K_1, \dots, K_{r_0}, r_0 \leq r$. Suppose that the theorem is wrong and $r < n_1 + n_2$. Pick a boolean vector $0 \neq (m_1, \dots, m_{n_1+n_2}) \in (\mathbb{Z}/2\mathbb{Z})^{n_1+n_2}$ orthogonal to all $K_i (mod\, 2), 1 \leq i \leq r_0$. Then

$$sgn(P_j(q_1, \dots, q_r)(\epsilon_1, \dots, \epsilon_{n_1+n_2})) = sgn(P_j(q_1, \dots, q_r)((-1)^{m_1}\epsilon_1, \dots, (-1)^{m_{n_1+n_2}}\epsilon_{n_1+n_2}))$$

for $1 \leq j \leq N$ (cf. the proof of lemma 1 [9]). This means that the output of C_0 is the same for the inputs $(\epsilon_1, \dots, \epsilon_{n_1+n_2})$ and $((-1)^{m_1}\epsilon_1, \dots, (-1)^{m_{n_1+n_2}}\epsilon_{n_1+n_2})$. The obtained contradiction with the supposition completes the proof of the theorem. ∎

Remark 3.3 *The bound in Theorem 3.2 still holds if instead of communication protocols one considers more general decision trees omitting the condition that each of the polynomials $q_1(X,Y), \dots, q_r(X,Y)$ depends either on X or on Y. This strengthens slightly lemma 1 [9] since here we consider decision trees without a priori bound on fan-out of branching, unlike [9] where the fan-out did not exceed 3.*

Remark 3.4 *Clearly, the communication complexity in the theorem equals $n_1 + n_2$.*

Remark 3.5 *The probabilistic communication complexity of recognizing the closure $\overline{T}$ does not exceed $\log^{O(1)}(n_1 + n_2)$. Indeed, the first party tests whether for an input $(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2})$ the inequalities $x_1 \geq 0, \dots, x_{n_1} \geq 0$ hold by means of a probabilistic decision tree of the depth $\log^{O(1)} n_1$ due to Theorem 1 [9]. The second party tests the inequalities $y_1 \geq 0, \dots, y_{n_2} \geq 0$ by the same token.*

The latter remark demonstrates an exponential gap between the probabilistic and deterministic communication complexities for recognizing the closure $\overline{T}$. The next proposition provides even a bigger gap for T.

Proposition 3.6 *The probabilistic communication complexity of recognizing T is at most 4.*

Proof. For an input $(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2})$ consider the partition of the indices $\{1, \dots, n_1 + n_2\} = I_0 \cup I_+ \cup I_-$ into the subsets for which the corresponding coordinates of the input are zero, positive or negative, respectively. If $I_0 \cup I_- \neq \emptyset$ then for a randomly

chosen subset $I \subseteq \{1, \ldots, n_1+n_2\}$ the probability of the event that $I \cap I_0 = \emptyset$ and that $|I \cap I_-|$ is even is less or equal to $1/2$. The latter statement is obvious when $I_0 \neq \emptyset$, and when $I_0 = \emptyset$ this probability equals to $1/2$.

Therefore, when $(x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}) \notin T$ and if one chooses randomly a product $\prod_{i_1 \in I_1} x_{i_1} \prod_{i_2 \in I_2} y_{i_2}$ then this product is positive with the probability less or equal to $1/2$. Thus, the first party chooses randomly independently two subsets $I^{(1)}, I^{(2)} \subseteq \{1, \ldots, n_1\}$ and calculates the products $\prod_{i \in I^{(1)}} x_i$ and $\prod_{i \in I^{(2)}} x_i$ (in a similar way the second party). If all 4 calculated products are positive then the output is "accept", otherwise "reject". ∎

4 Lower bound on probabilistic communication complexity

Corollary 2.2 together with Lemma 1.4 show that if the (n_1+n_2-1)-dimensional boundary of a semialgebraic set contains a "facet" with a great communication complexity of computing the polynomial which determines this facet, then the probabilistic communication complexity of recognizing this set is great as well. Now we construct a set (being a polyhedron) with a great probabilistic communication complexity (note that any facet of the polyhedron being determined by a linear function, has a communication complexity at most 2).

Consider the polyhedron $S = \{(x_1, \ldots, x_n, y_1, \ldots, y_n) \in \mathbb{R}^{2n} : x_i + y_i > 0, 1 \leq i \leq n\}$ and an *arrangement* R either real (i. e. $\subset \mathbb{R}^{2n}$) or complex (i. e. $\subset \mathbb{C}^{2n}$) being a union of hyperplanes among which there appear n hyperplanes $\{X_i + Y_i = 0\}, 1 \leq i \leq n$.

Theorem 4.1 *The probabilistic communication complexity of recognizing over the reals the set S or the set R is greater than $n/2$.*

Proof. Denote $Z_i = X_i + Y_i$, $1 \leq i \leq n$. We consider the new coordinates $(X_1, \ldots, X_n, Z_1, \ldots, Z_n)$ in $\mathbb{R}^{2n}$ and the point $u = (\epsilon_1, \ldots, \epsilon_{2n})$. Let a probabilistic communication protocol C recognize S (respectively, R). Introduce n points $u_i = (\epsilon_1, \ldots, \epsilon_{n+i-1}, -\epsilon_{n+i}, \epsilon_{n+i+1}, \ldots, \epsilon_{2n})$ (respectively, $u_i^{(0)} = (\epsilon_1, \ldots, \epsilon_{n+i-1}, 0, \epsilon_{n+i+1}, \ldots, \epsilon_{2n})$), $1 \leq i \leq n$. Clearly, $u \in S, u_i \notin S$ (respectively, $u \notin R, u_i^{(0)} \in R$).

There exists a communication protocol C_0 from the family constituting C which gives correct outputs for the input u and for at least of $n/2$ inputs among u_i (respectively, $u_i^{(0)}$). Without loss of generality one can assume that the outputs are correct for all $u_i, 1 \leq i \leq \lceil n/2 \rceil$ (respectively, for $u_i^{(0)}$).

Take the path in C_0 which follows the input u and consider the testing polynomials $P_1(q_1, \ldots, q_r), \ldots, P_N(q_1, \ldots, q_r)$ along this path (cf. Section 2). Denote $P = \prod_{1 \leq j \leq N} P_j(q_1, \ldots, q_r)$. We claim that the least term $lt(P) = \prod_{1 \leq j \leq N} lt(P_j(q_1, \ldots, q_r))$ divides on each $Z_i, 1 \leq i \leq \lceil n/2 \rceil$ (recall that the least term is defined with respect to the coordinates $(X_1, \ldots, X_n, Z_1, \ldots, Z_n)$). Otherwise, if $lt(P)$ does not divide on Z_i then we have

$$sgn(P_j(q_1, \ldots, q_r)(u)) = sgn(P_j(q_1, \ldots, q_r)(u_i)), 1 \leq j \leq N$$

(respectively,

$$sgn(P_j(q_1, \ldots, q_r)(u)) = sgn(P_j(q_1, \ldots, q_r)(u_i^{(0)}))).$$

Hence C_0 gives the same output for both inputs u and u_i (respectively, $u_i^{(0)}$). The obtained contradiction proves the claim.

Thus, the theorem would follow from the next lemma taking into account Lemma 1.2. ∎

Lemma 4.2 *If for a certain $k > 1$ the product $Z_1 \cdots Z_k$ divides $lt(P)$ then for the rank of $n \times n$ matrix we have*

$$rk\left(\frac{\partial^2 P}{\partial X_i \partial Y_j}\right) \geq k$$

Proof. Let $lt(P) = p_0 X_1^{m_1} \cdots X_n^{m_n} Z_1^{l_1} \cdots Z_n^{l_n}$ where $p_0 \in \mathbb{R}$. Then the highest term (cf. (1)) of a non-diagonal entry $\frac{\partial^2 P}{\partial X_i \partial Y_j}(u)$ when $i \neq j, 1 \leq i, j \leq k$ equals

$$\frac{l_i l_j \epsilon_1^{m_1} \cdots \epsilon_n^{m_n} \epsilon_{n+1}^{l_1} \cdots \epsilon_{2n}^{l_n}}{\epsilon_{n+i}\epsilon_{n+j}}$$

The highest term of a diagonal entry $\frac{\partial^2 P}{\partial X_i \partial Y_i}(u)$ either equals

$$\frac{l_i(l_i-1)\epsilon_1^{m_1} \cdots \epsilon_n^{m_n} \epsilon_{n+1}^{l_1} \cdots \epsilon_{2n}^{l_n}}{\epsilon_{n+i}^2}$$

when $l_i > 1$ or is less than

$$\frac{\epsilon_1^{m_1} \cdots \epsilon_n^{m_n} \epsilon_{n+1}^{l_1} \cdots \epsilon_{2n}^{l_n}}{\epsilon_{n+i}^2}.$$

Denote by M $k \times k$ matrix with the diagonal (i,i)-entries $l_i(l_i - 1)$ and the non-diagonal (i,j)-entries $l_i l_j, 1 \leq i, j \leq k$. Then $det(M) = (-1)^{k+1} l_1 \cdots l_k (l_1 + \cdots + l_k - 1) \neq 0$ when $k > 1$. Therefore, the coefficient of the $k \times k$ minor

$$det\left(\frac{\partial^2 P}{\partial X_i \partial Y_j}\right)(u)$$

where $1 \leq i, j \leq k$ at its highest term

$$(\epsilon_1^{m_1} \cdots \epsilon_n^{m_n})^k \epsilon_{n+1}^{kl_1-2} \cdots \epsilon_{2n}^{kl_n-2}$$

equals to $det(M)$ and thereby, it does not vanish, which proves the lemma. ∎

Remark 4.3 *The same bound as in the theorem holds as well for the (euclidean) closure $\overline{S}$.*

Corollary 4.4 *The probabilistic communication complexity over complex numbers of R is greater than $n/4$.*

Proof. Having a probabilistic communication protocol C over $\mathbb{C}$ which recognizes R, one can convert it into a probabilistic communication protocol $C^{(\mathbb{R})}$ over reals which recognizes R at the cost of increasing the complexity at most twice. For this purpose the first party replaces every polynomial $a(X)$ in C which the first party calculates by a pair of polynomials $Re(a), Im(a) \in \mathbb{R}[X]$ in $C^{(\mathbb{R})}$ where $a = Re(a) + \sqrt{-1}Im(a)$. The same for the second party. Then for each testing polynomial $P_j(q_1, \ldots, q_r)$ its real and imaginary parts $Re(P_j(q_1, \ldots, q_r)), Im(P_j(q_1, \ldots, q_r))$ can be expressed as polynomials over $\mathbb{R}$ in $Re(q_l), Im(q_l), 1 \leq l \leq r$. Any ramification condition $P_j(q_1, \ldots, q_r) = 0$ in C we replace in

$C^{(\mathbb{R})}$ by $Re(P_j(q_1,\ldots,q_r)) = Im(P_j(q_1,\ldots,q_r)) = 0$. To complete the proof of the corollary we apply Theorem 4.1 to $C^{(\mathbb{R})}$. ∎

As particular cases consider the problem EMPTINESS: whether the intersection of two finite sets $\{x_1,\ldots,x_n\} \cap \{y_1,\ldots,y_n\} = \emptyset$ is empty? It corresponds to the arrangement $\cup_{i,j}\{x_i = y_j\}$ (in $\mathbb{C}^{2n}$ or $\mathbb{R}^{2n}$). Another example is the KNAPSACK problem: whether there exist subsets $I_1, I_2 \subseteq \{1,\ldots,n\}$ such that $\sum_{i_1\in I_1} x_{i_1} + \sum_{i_2\in I_2} y_{i_2} = 0$? It can be also represented as an arrangement (cf. [10]).

Corollary 4.5 *The probabilistic communication complexity of both EMPTINESS and KNAPSACK problems is greater than $n/4$ over $\mathbb{C}$ and greater than $n/2$ over $\mathbb{R}$.*

Acknowledgements. The author is grateful to the Max-Planck Institut fuer Mathematik, Bonn where the paper was written, to Farid Ablayev and to Harry Buhrman for interesting discussions and to anonymous referees for very detailed comments, which helped to improve the presentation of the paper.

References

[1] H. Abelson, *Lower bounds on information transfer in distributed computations*, J. Assoc. Comput. Mach., **27** (1980), 384–392.

[2] F. Ablayev, S. Ablayeva, *A discrete approximation and communication complexity approach to the superposition problem*, in Proc. Intern. Symp. Fundamentals of Computation Theory, Lect. Notes Comput. Sci., **2138**, (2001), Springer, 47–58.

[3] M. Bläser, E. Vicari, *Algebraic communication complexity*, Preprint (2007).

[4] L. Blum, F. Cucker, M. Shub, S. Smale, *Complexity and real computations*, Springer (1998).

[5] H. Buhrman, R. de Wolf, *Communication complexity lower bounds by polynomials*, Proc. IEEE Conf. Computational Complexity (2001), 120–130.

[6] P. Bürgisser, *Completeness and reduction in algebraic complexity theory*, Springer (2000).

[7] D. Grigoriev, N. Vorobjov, *Solving systems of polynomial inequalities in subexponential time*, J. Symb. Comput., **5** (1988), 37–64.

[8] D. Grigoriev, M. Karpinski, F. Meyer auf der Heide, R. Smolensky, *A lower bound for randomized algebraic decision trees*, Computational Complexity, **6** (1996/1997), 357–375.

[9] D. Grigoriev, M. Karpinski, R. Smolensky, *Randomization and the computational power of analytic and algebraic decision trees*, Computational Complexity, **6** (1996/1997), 376–388.

[10] D. Grigoriev, *Randomized complexity lower bounds for arrangements and polyhedra*, Discrete Computational Geometry, **21** (1999), 329–344.

[11] J. Krajiĉek. *Interpolation by a game*, Math. Logic Quat., **44** (1998), 450–458.

[12] E. Kushilevitz, N. Nisan, *Communication complexity*, Cambridge (1997).

[13] L. Lovasz, *Communication complexity: a survey*, in "Paths, flows and VLSI layout", Korte, Lovasz, Proemel, Schrijver Eds. (1990), Springer, 235–266.

[14] F. Meyer auf der Heide, *Simulating probabilistic by deterministic algebraic computation trees*, Theor. Comp. Sci., **41** (1984), 325–330.

[15] A. Yao, *Some complexity questions related to distributive computing*, in Proc. ACM Symp. Theory on Computing (1979), 209–213.