On the Communication Complexity of Distributed Algebraic Computation

ZHI-QUAN LUO AND JOHN N. TSITSIKLIS

Massachusetts Institute of Technology, Cambridge, Massachusetts

Abstract. We consider a situation where two processors P_1 and P_2 are to evaluate a collection of functions $f_1, \ldots, f_s$ of two-vector variables x, y, under the assumption that processor P_1 (respectively, P_2) has access only to the value of the variable x (respectively, y) and the functional form of $f_1, \ldots, f_s$. We provide some new bounds on the communication complexity (the amount of information that has to be exchanged between the processors) for this problem. An almost optimal bound is derived for the case of one-way communication when the functions $f_1, \ldots, f_s$ are polynomials. We also derive some new lower bounds for the case of two-way communication that improve on earlier bounds by Abelson [2]. As an application, we consider the case where x and y are $n \times n$ matrices and $f(x, y)$ is a particular entry of the inverse of $x + y$. Under a certain restriction on the class of allowed communication protocols, we obtain an $\Omega(n^2)$ lower bound, in contrast to the $\Omega(n)$ lower bound obtained by applying Abelson's results. Our results are based on certain tools from classical algebraic geometry and field extension theory.

Categories and Subject Descriptors: F.1.2 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Algebraic computation, communication complexity, lower bound

1. *Introduction*

In several situations of practical interest, there is a set of processors who wish to perform some computational task and who must communicate because none of them possesses all of the problem data. Communication resources are often limited and we are led to the study of the minimal amount of required information transfer, that is, the "communication complexity" of the problem under consideration. For example, in parallel computation [5], communication is often much slower than computation and excessive communication may create bottlenecks to the speed of an algorithm. A similar argument applies to computations using special purpose VLSI chips [20] in which communications

This research was supported by the Office of Naval Research (ONR) under contract N00014-84-K-0519 (NR 649-003) and the Army Research Office (ARO) under contract DAAL03-86-K-0171.

Authors' current addresses: Z.-Q. Luo, Room 225/CRL, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, L8S 4L7, Canada; J. N. Tsitsiklis, Operations Research Center and Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

capabilities are constrained by physical and topological considerations. Finally, there are several applications in signal processing: for example, in decentralized estimation and detection, or in distributed sensor networks [18], data are collected at geographically distant sites. Then, summaries of the data are communicated so as to enable a particular processor or sensor to make certain statistical inferences (see, e.g., [22]). Communication resources are often costly in such contexts, and it is again natural to minimize the amount of information exchange.

1.1. COMMUNICATION PROTOCOLS. In this subsection, we introduce the class of protocols that will be considered and we formulate the general problem to be studied.

Let there be two processors P_1 and P_2. Processor P_1 (respectively, P_2) has access to the value of a vector $x \in \Re^m$ (respectively, $y \in \Re^n$). Let there be a given a finite collection $\vec{f}$ of functions $f_1, f_2, \dots, f_s$: $D_{\vec{f}} \mapsto \Re$, where $D_{\vec{f}}$ is some subset of $\Re^m \times \Re^n$ on which these functions are defined. (For example, if each f_i is a rational function expressed as a ratio of two relatively prime polynomials, it is natural to let $D_{\vec{f}}$ be the set of all vectors at which none of the denominators of these functions vanishes.)

The objective of the processors is to exchange messages and compute the values $f_1(x, y), \dots, f_s(x, y)$. It is assumed that both processors know the formulas defining these functions. (For instance, if each f_i is a polynomial, then each processor knows the coefficients of these polynomials.) Ideally, a protocol should work for all possible values $(x, y) \in D_{\vec{f}}$ of the "inputs." We occasionally consider, however, protocols that are defined only when (x, y) belongs to some possibly smaller set $D \subset D_{\vec{f}}$.

In a *two-way communication protocol* π, messages can be exchanged in both directions. We use $r(\pi)$ to denote the number of exchanged messages and we let $T_{1\to 2}$ (respectively, $T_{2\to 1}$) denote the set of is for which the ith message is transmitted from P_1 to P_2 (respectively, from P_2 to P_1). The protocol is defined in terms of a collection of functions $m_1, \dots, m_{r(\pi)}$ mapping a set $D \subseteq D_{\vec{f}}$ into $\Re$. (In particular, $m_i(x, y)$ is the value of the ith message and the set D is called the *domain* of the protocol.) Since a message by a processor can only be a function of the information available to that processor, we impose the requirement that for for each i, there exists some real-valued function $\hat{m}_i$ such that

$$m_i(x, y) = \hat{m}_i(x, m_1(x, y), \dots, m_{i-1}(x, y)), \qquad \forall (x, y) \in D, \quad \text{if} \quad i \in T_{i\to 2}, \tag{1.1}$$

and

$$m_i(x, y) = \hat{m}_i(y, m_1(x, y), \dots, m_{i-1}(x, y)), \qquad \forall (x, y) \in D \quad \text{if} \quad i \in T_{2\to 1}. \tag{1.2}$$

We say that the protocol is *legitimate* if either of the following conditions is true:

(a) There exist functions $h_1, \dots, h_s$ such that

$$f_i(x, y) = h_i\big(x, m_1(x, y), \dots, m_{r(\pi)}(x, y)\big), \qquad \forall (x, y) \in D, \quad i = 1, \dots, s. \tag{1.3}$$

(This corresponds to the case where processor P_1 evaluates the final result.)

(b) There exist functions $h_1, \ldots, h_s$ such that

$$f_i(x, y) = h_i\big(y, m_1(x, y), \ldots, m_{r(\pi)}(x, y)\big), \qquad \forall (x, y) \in D, \quad i = 1, \ldots, s. \quad (1.4)$$

Let $\Pi(\vec{f}; D, \leftrightarrow)$ denote the class of all legitimate two-way protocols, with domain D, for computing the functions $f_1, \ldots, f_s$, subject to some additional restrictions to be introduced later. We define the two-way communication complexity $C(\vec{f}; D, \leftrightarrow)$ for computing $\vec{f}$ on the domain D to be

$$C\big(\vec{f}; D, \leftrightarrow\big) = \inf_{\pi \in \Pi(\vec{f}; D, \leftrightarrow)} r(\pi).$$

The definition of an *one-way communication protocol* π is identical, except that messages can only be transmitted by processor P_1. That is, the set $T_{2 \to 1}$ is assumed empty. Let $\Pi(\vec{f}; D, \to)$ denote the set of all legitimate one-way communication protocols with domain D. We define the one-way (from P_1 to P_2) communication complexity $C(\vec{f}; D, \to)$ on the domain D to be

$$C\big(\vec{f}; D, \to\big) = \inf_{\pi \in \Pi(\vec{f}; D, \to)} r(\pi).$$

Notice that in the above models the protocols are "continuous" in the sense that the messages to be sent are real numbers. Given that real numbers can only be encoded with an infinite number of bits, such protocols might seem impossible to implement in practice. However, parallel and distributed numerical algorithms are almost always described and analyzed as if real numbers can be communicated, with the understanding that in practice these numbers will be encoded with a finite number of bits that is sufficient to obtain a desired accuracy. Furthermore, if the messages being transmitted are rational functions of the data and if the data consist of rational numbers, then an implementation using a finite number of bits is clearly possible. Finally, in practice, it is usually the case that a field of a fixed length is used for transmitting an encoded version of a real number. For this reason, it is reasonable to count the number of real-valued messages being transmitted, as opposed to counting individual bits. Our model is therefore a fairly realistic way of capturing the communication resources needed in a number of practical applications. Let us also note that the formal model of real-number computation introduced by [6] has been motivated by similar considerations.

Typically, some smoothness constraints have to be imposed on the message functions $m_1, \ldots, m_{r(\pi)}$. This is because there exist one-to-one functions from $\Re^m$ into $\Re$, and processor P_1 could transmit the value of its vector x by using a single message. In particular, P_1 can simply interleave the binary expansions of the components of x and use the resulting number as a message. This is not a useful protocol, for the purposes of numerical computation, and is unlike any protocol that is used in practice. In contrast to the above-described interleaving, a good protocol should compress the information in x or y intelligently, and then transmit only the compressed information. For this reason, we shall impose some smoothness requirements on the message functions m_i. From a

TABLE I. VARIOUS RESTRICTIONS ON ALLOWED PROTOCOLS

Notations	Restrictions on the message functions $\hat{m}_1, \ldots, \hat{m}_r$ (cf. Eqs. (1 1)–(1.2)).	Restrictions on the final evaluation functions $h_1, \ldots, h_s$ (cf Eqs (1.3)–(1.4)).
$\Pi_1(\vec{f}, D)$	continuously differentiable	continuously differentiable
$\Pi_2(\vec{f}, D)$	twice continuously differentiable	twice continuously differentiable
$\Pi_\infty(\vec{f}, D)$	infinitely diffcrentiable	infinitely differentiable
$\Pi_{rat}(\vec{f}, D)$	rational	rational
$\Pi_{poly}(\vec{f}, D)$	polynomial	rational
$\Pi_{linear}(\vec{f}, D)$	linear	polynomial

technical point of view, smoothness assumptions prohibit the use of one-to-one functions from $\Re^m$ into $\Re$, if $m > 1$. From a practical point of view, such smoothness is present in the vast majority of practical numerical methods for algebraic problems. Furthermore, in this paper, we concentrate on the case where each one of the functions in $f_1, \ldots, f_s$ is rational. It is then natural to restrict attention even further to protocols involving only rational functions of the data. This is equivalent to an assumption that each processor can only perform the elementary arithmetic operations. Such an assumption is common in complexity studies for algebraic problems [7].

In the sequel, we use the shorter notations $\Pi(\vec{f}; D)$ and $C(\vec{f}; D)$ whenever it is clear from the context whether we are dealing with one-way or two-way protocols. Furthermore, we use the notation $\Pi(f; D)$ and $C(f; D)$ whenever $s = 1$ and the collection $\vec{f}$ of functions consists of the single function f.

In this paper, we consider various restrictions on the set of allowed protocols. We indicate these restrictions in our notation, as shown in Table I.

We use notation like $C_1(\vec{f}; D)$, $C_2(\vec{f}; D)$, etc., to denote the communication complexity under the restrictions on the protocols introduced in Table I. Notice that, as we go down the table, additional restrictions are introduced and, therefore, the corresponding communication complexity can only increase. Finally, assuming that D is a nonempty open set, we see that the set $\Pi_{rat}(\vec{f}; D)$ (respectively, $\Pi_{linear}(\vec{f}; D)$) is empty unless $\vec{f}$ is a rational (respectively, polynomial) function.

All of our definitions can be extended, in the obvious way, to the case where the real number field $\Re$ is replaced by the complex field $\mathscr{C}$. Here, all the functions f_i are defined on a subset D_{f_i} of $\mathscr{C}^m \times \mathscr{C}^n$ and take values in $\mathscr{C}$. Furthermore, a protocol has a domain $D \subset \mathscr{C}^m \times \mathscr{C}^n$ and the message functions m_i and $\hat{m}_i$ [cf. Eqs. (1.1)–(1.2)] are defined on D.

1.2. RELATED RESEARCH. The problem formulation we are using is due to Abelson [1, 2] who established lower bounds on one-way and two-way communication complexity, assuming that the message functions are once (respectively, twice) continuously differentiable. (These results are stated and discussed in Sections 3 and 5, respectively.)

Communication complexity has also been studied under discrete models of communication. In these models, the messages exchanged are binary and the functions evaluated are such that a finite number of binary messages are actually sufficient. For example, Yao [27] and Papadimitriou and Sipser [16] consider the computation of Boolean functions using binary messages. The approach in these references is combinatorial in nature and very different from ours. A fair amount of research has dealt with extensions of the results of [27] and with the evaluation of the communication complexity of selected combinatorial problems [3, 14–17, 20]. A different framework is considered in [19] for the problem of approximately minimizing (within a desired accuracy) the sum of two convex functions, with each function known by a different processor. Here, the objective is to minimize the number of binary messages, as a function of the desired accuracy of the solution.

1.3. OUTLINE OF THE PAPER. The rest of this paper is organized as follows: In Section 2, we present some background results from field extension theory that will be used in our study of one-way communication complexity.

In Section 3, we study the one-way communication complexity of computing a set $f_1, \dots, f_s$ of polynomials. The results of [1] (stated in Section 3.1) provide a complete solution for the case of a single function f, smooth message functions, and polynomials whose domain is a (possibly very small) open set. We extend these results to the case of $s > 1$. We also show that we can restrict to the class of polynomial protocols while increasing the communication complexity by at most one. Furthermore, the polynomial protocols we construct have a domain that is almost all of $\Re^m \times \Re^n$ (except for a set of measure zero). We also consider the special case where $m = n$ and each one of the polynomials $f_i : \Re^n \times \Re^n$ is of the form $f_i(x, y) = \hat{f}_i(x + y)$, where $\hat{f}_i$ is a polynomial in n variables. For this case, we obtain a complete characterization of the communication complexity, a proof that linear protocols are optimal, and a constructive procedure for designing such protocols.

In Section 4, we present some background from algebraic geometry (e.g., Hilbert's Nullstellensatz) that will be needed later.

In Section 5, we derive several general lower bounds on two-way communication complexity of computing a rational function f when the messages are constrained to be rational functions of the data. Our results are obtained by combining an earlier result of Abelson [2] with the tools of Section 4. We also identify certain instances where the lower bounds of [2] are tight.

In Section 6, we apply the results of Section 5 to the problem of computing a particular entry of the inverse of $x + y$, where x and y are $n \times n$ complex matrices. We derive an $n^2 - 1$ lower bound (which agrees with the obvious upper bound, within one message), while the results of [2] could only provide an $\Omega(n)$ lower bound.

2. *Preliminaries*

In this section, we introduce some algebraic results (see, e.g., [25, pp. 95–125] or [21]) that will be needed in Section 3.

Notations. Let $\{a_i : i \in I\}$ be a collection of vectors in $\Re^n$, where I is a finite index set. We use $[a_i : i \in I]$ to denote the matrix with columns a_i, $i \in I$. Whenever the range of the index variable i (i.e., the index set I) is evident from the context, we use the simpler notation $[a_i : i]$. For any function $f: \Re^n \mapsto \Re$, we use ∇f to denote the vector-valued function whose components are the partial derivatives of f. We also use $\nabla f(p)$ to denote the value of ∇f evaluated at some $p \in \Re^n$.

Let F_1 be a field and let F_2 be an extension field of F_1. An element $\lambda \in F_2$ is called a *primitive element* of the extension F_2/F_1 if $F_2 = F_1(\lambda)$, that is, if F_2 is generated by λ over the field F_1. The following result (see, e.g., [25, page 84]) is called the theorem of primitive element and will be used in Section 3.

THEOREM 2.1. *Every finite separable algebraic extension F_2/F_1 has a primitive element. Furthermore, if $F_2 = F_1(\lambda_1, \dots, \lambda_k)$, then there exists a primitive element of the form $\lambda = \sum_{j=1}^{k} \gamma_j \lambda_j$ where $\gamma_j \in F_1$ for each j.*

Remark. In fact, the proof of Theorem 2.1 given in [25, page 84] shows that a primitive element λ is obtained for an arbitrary choice of the coefficients $\gamma_1, \dots, \gamma_k$, as long as they do not lie in the zero set of a certain polynomial.

We now turn our attention to the case of transcendental extensions. Let F_2/F_1 be a field extension. The transcendental degree of F_2/F_1, denoted by tr.d.F_2/F_1, is defined as the smallest number t such that there exist elements $\lambda_1, \lambda_2, \dots, \lambda_t$ in F_2 with the property that F_2 is an algebraic extension of $F_1(\lambda_1, \lambda_2, \dots, \lambda_t)$. The following theorem summarizes some important properties of the transcendental degree of a field extension.

THEOREM 2.2. *Let F_2 be a finitely generated extension field of F_1 and let F_3 be a finitely generated extension field of F_2. (In particular, F_3 is also a finitely generated extension field of F_1.) Suppose that $F_3 = F_1(\lambda_1, \lambda_2, \dots, \lambda_n)$ and that tr.d.$F_3/F_1 = t$. Then, $t =$ tr.d.$F_3/F_1 =$ tr.d.$F_3/F_2 +$ tr.d.F_2/F_1.*

The following is the definition of a derivation over a field, which is a generalized notion of differentiation.

Definition 2.1. *Let F_2 be a finitely generated extension field of F_1 and let F_3 be an extension field of F_2. A mapping D of F_2 into F_3 is said to be an F_1-derivation of F_2 (with values in F_3) if for every λ in F_1 and every x, y in F_2 the mapping D has the following three properties*: (1) $D(\lambda) = 0$; (2) $D(x+y) = D(x) + D(y)$; (3) $D(xy) = xD(y) + yD(x)$.

The well-known chain rules remain true for derivations. We now let $\mathscr{D}_{F_2/F_1}(F_3)$ stand for the space of all F_1-derivations of F_2 with values in F_3. Then $\mathscr{D}_{F_2/F_1}(F_3)$ can be viewed as a vector space over F_3. It can be shown (see [25, pp. 120–127]) that the dimension of the vector space $\mathscr{D}_{F_2/F_1}(F_3)$ does not depend on the particular choice of F_3. It is for this reason that we usually drop F_3 from the notation $\mathscr{D}_{F_2/F_1}(F_3)$ and use simply $\mathscr{D}_{F_2/F_1}$ to denote the space of F_1-derivations of F_2 with values in any extension field of F_2.

Example. We now consider in some detail the space of derivations for an important special case and derive a result that will be needed in Section 3. Let $F_1 = \Re$ and let $F_3 = \Re(x_1, x_2, \dots, x_m)$, the field of rational functions over $\Re$ with indeterminates $x_1, x_2, \dots, x_m$. Furthermore, we let F_2 be the subfield of F_3 that is generated by polynomials $f_1, f_2, \dots, f_n \in F_3$. In other words, F_2 is the set of all rational functions that can be expressed as rational functions of the f_j's. As is well known, we have, for any $D \in \mathscr{D}_{F_3/F_1}(F_3)$,

$$D = \sum_{k=1}^{m} D(x_k) \frac{\partial}{\partial x_k}.$$

Hence, D is completely determined by the choice of $D(x_k) \in F_3$, $k = 1, 2, \dots, m$, and $\{(\partial/\partial x_1), \dots, (\partial/\partial x_m)\}$ is a basis for $\mathscr{D}_{F_3/F_1}(F_3)$. Now suppose that $D \in \mathscr{D}_{F_2/F_1}(F_3)$. Since F_2 has characteristic 0, it follows that D can be extended to a derivation $\bar{D}$ in $\mathscr{D}_{F_3/F_1}(F_3)$ (see [25, pp. 125–127]). From the above discussion, we see that

$$\bar{D} = \sum_{k=1}^{m} \bar{D}(x_k) \frac{\partial}{\partial x_k}. \tag{2.1}$$

Therefore, the map D, which is equal to the restriction of $\bar{D}$ on F_2, can be written as a linear combination of the $(\partial/\partial x_k)$'s (cf. Eq. (2.1)). Conversely, for each choice of $\bar{D}(x_k) \in F_3$, Eq. (2.1) defines a derivation in $\mathscr{D}_{F_2/F_1}(F_3)$. However, two different choices of $\bar{D}(x_k)$ may give rise to the same derivation in $\mathscr{D}_{F_2/F_1}(F_3)$. As a matter of fact, any $f \in F_2$ can be expressed in the form of $f = g(f_1, f_2, \dots, f_n)$, where $g(z_1, z_2, \dots, z_n)$ is a rational function. By the chain rule, we have

$$D(f) = \frac{\partial g}{\partial z_1} D(f_1) + \frac{\partial g}{\partial z_2} D(f_2) + \cdots + \frac{\partial g}{\partial z_n} D(f_n),$$

where $\partial g/\partial z_j$ is the partial derivative of g with respect to z_j defined in the usual sense. Since the $\partial g/\partial z_j$'s are independent of D, we see that D is completely determined by its operation on f_j, $j = 1, 2, \dots, n$. Moreover, since the f_j's belong to F_2 we see that different choices of the $D(f_j)$'s will result in different derivations in $\mathscr{D}_{F_2/F_1}$.

We now develop an explicit formula for the dimensions of $\mathscr{D}_{F_2/F_1}$ (eq. (2.4) below), in the context of the particular example we have been considering. This formula will be crucial for the results of Section 3.

Notice that for every j and any $D \in \mathscr{D}_{F_2/F_1}$, one has

$$\begin{aligned} D(f_j) &= \left(\sum_{k=1}^{m} \bar{D}(x_k) \frac{\partial}{\partial x_k} \right)(f_j) = \sum_{k=1}^{m} \bar{D}(x_k) \frac{\partial f_j}{\partial x_k} \\ &= \big(\bar{D}(x_1), \bar{D}(x_2), \dots, \bar{D}(x_m)\big) \nabla f_j. \end{aligned} \tag{2.2}$$

We now rewrite Eq. (2.2) in the matrix form

$$(D(f_1), D(f_2), \dots, D(f_n)) = \big(\bar{D}(x_1), \bar{D}(x_2), \dots, \bar{D}(x_m)\big)[\nabla f_j \colon j \in J],$$

where $J = \{1, 2, \dots, n\}$. Since $\bar{D}(x_k)$ can be taken arbitrarily, we see that the vector space $\mathscr{D}_{F_2/F_1}(F_3)$ is isomorphic to the space spanned by the rows of the matrix $[\nabla f_j \colon j]$. Hence

$$\dim \mathscr{D}_{F_2/F_1} = \operatorname{rank}[\nabla f_j \colon j], \tag{2.3}$$

where the entries of $[\nabla f_j \colon j]$ are polynomials in variables $x_1, x_2, \dots, x_m$ and the rank is evaluated in the field F_3. We can now assign real values to $x_1, x_2, \dots, x_m$ and calculate the rank in $\Re$. Let $[\nabla f_j(p) \colon j]$ denote the matrix $[\nabla f_j \colon j]$ evaluated at the point $p \in \Re^m$. Notice that if $\Sigma a_i(p) \nabla f_i(p) = 0$, $\forall p$, then, by solving the linear system formally with Gaussian elimination, we see that each $a_i(p)$ can be chosen as a rational function of p. This implies

$$\max_{p \in \Re^m} \operatorname{rank}\big([\nabla f_j(p) \colon j]\big) = \operatorname{rank}\big([\nabla f_j \colon j]\big).$$

Combining this with Eq. (2.3), we obtain the following basic result:

$$\dim \mathscr{D}_{F_2/F_1} = \max_{p \in \Re^m} \operatorname{rank}\left(\left[\nabla f_j(p) : j\right]\right). \tag{2.4}$$

We close this section with a result that relates the transcendental extension degree and the dimension of the associated space of derivations (see [25, pp. 125–127]).

THEOREM 2.3. *Let F_1 be a field and let F_2 be a finitely generated extension field of F_1 such that tr.d.$F_2/F_1 = d$ and dim $\mathscr{D}_{F_2/F_1} = t$. Then t is equal to the smallest number r such that there exist elements $\lambda_1, \lambda_2, \ldots, \lambda_r$ with the property that F_2 is separable algebraic over $F_1(\lambda_1, \lambda_2, \ldots, \lambda_r)$. In particular, $t \geq d$. Furthermore, if F_1 has characteristic* 0, *then the equality $t = d$ holds.*

3. *One-Way Communication Complexity*

In this section, we study the one-way communication complexity of evaluating a set $f_1, \ldots, f_s$ of polynomials, when the messages transmitted are restricted to be polynomial functions of the data. We apply the tools of field extension theory (presented in Section 2) to obtain a bound for the communication complexity that is almost optimal (within one message). It will be seen that our results strengthen earlier results in a number of directions. We also show that the restriction to polynomial protocols can increase the communication complexity of the problem by at most one message. We then specialize to the case where the polynomials f_j to be evaluated are of the form $f_j(x, y) = \hat{f}_j(x + y)$, for some functions $\hat{f}_j$, and we show that there exist optimal protocols with a very simple structure: they consist of messages that are linear functions of the data.

3.1. General Results.

The main available result on one-way protocols is due to Abelson [1]:[1]

THEOREM 3.1. *Let $f: \Re^m \times \Re^n \mapsto \Re$ be an infinitely differentiable function.*

(*a*) *Let D be a subset of $\Re^m \times \Re^n$. There holds $C_\infty(f; D) \leq r$ if and only if there exist infinitely differentiable functions $m_1, m_2, \ldots, m_r$: $\Re^n \mapsto \Re$ and h: $\Re^{r+n} \mapsto \Re$ such that*

$$f(x, y) = h(y, m_1(x), m_2(x), \ldots, m_r(x)), \qquad \forall (x, y) \in D. \tag{3.1}$$

(*b*) *Let (x^*, y^*) be some element of $\Re^m \times \Re^n$. There exists some open set $D \subset \Re^m \times \Re^n$ containing (x^*, y^*) for which $C_\infty(f; D) \leq r$ if and only if*

$$dim(span\{g_{1,x^*}, g_{2,x^*}, \ldots, g_{m,x^*}\}) \leq r, \tag{3.2}$$

where $g_{i,x^*}(y) = (\partial f/\partial x_i)(x^*, y)$ and where the span is taken in the vector space of functions of y defined on an open set containing (x^*, y^*).

[1] We state this result for the class $\Pi_\infty(f; D)$ of protocols that use infinitely differentiable functions. The result was actually proved in [1] for the class $\Pi_1(f; D)$ but the proof remains valid for $\Pi_\infty(f; D)$.

Let us consider protocols whose domain D is all of $\Re^m \times \Re^n$. By varying (x^*, y^*) over all possible elements of $\Re^m \times \Re^n$ and applying part (b) of the theorem to each one of these points we obtain

$$C_\infty(f; \Re^m \times \Re^n) \geq \max_{x^* \in \Re^m} \dim(\operatorname{span}\{g_{1,x^*}, \dots, g_{m,x^*}\}). \tag{3.3}$$

Part (b) of the theorem states that this lower bound is also tight in a local sense: there exist protocols whose number of messages equals the lower bound and that evaluate f correctly when (x, y) is restricted to a suitably small domain D. However, nothing can be inferred on the tightness of this bound when one considers protocols whose domain is all of $\Re^m \times \Re^n$. Furthermore, the message functions m_i in Eq. (3.1) are not guaranteed to be polynomials, even if the function f is a polynomial. Both of these deficiencies will be remedied in the sequel.

Throughout this section, we assume that we are dealing with a given set $\vec{f} = \{f_1, \dots, f_s\}$ of polynomial functions mapping $\Re^m \times \Re^n$ into $\Re$ and that only one-way protocols are considered. We start by proving a lower bound similar to Theorem 3.1(b), but more general, because Theorem 3.1 dealt only with the case $s = 1$.

Notation. For $i = 1, \dots, s$, and for any sequence $\alpha = (\alpha_1, \dots, \alpha_n)$ of nonnegative integer indices, we define a function $g_i^\alpha \colon \Re^{m+n} \mapsto \Re$ by letting

$$g_i^\alpha(x, y) = \frac{\partial^\alpha f_i}{\partial y_1^{\alpha_1} \partial y_2^{\alpha_2} \cdots \partial y_n^{\alpha_n}}(x, y). \tag{3.4}$$

(We use the convention $g_i^0 = f_i$.) Let $\mathscr{A}$ be the set of all α such that g_i^α is not identically zero for some i. (Clearly, $\mathscr{A}$ is a finite set, since each f_i is a polynomial.) For any function $g(x, y) \colon \Re^m \times \Re^n \mapsto \Re$, we use $\nabla_x g$ to denote the vector-valued function of dimension m whose components are the partial derivatives of g with respect to the first m coordinates.

THEOREM 3.2. *Let D be some open subset of $\Re^m \times \Re^n$.*

(*a*) *If $C_\infty(\vec{f}; D) \leq r$, then there exist infinitely differentiable functions $m_1, \dots, m_r \colon \Re^m \mapsto \Re$ and $h_i^\alpha \colon \Re^{r+n} \mapsto \Re$, $i = 1, \dots, s$, $\alpha \in \mathscr{A}$, such that*

$$g_i^\alpha(x, y) = h_i^\alpha(y, m_1(x), \dots, m_2(x)), \qquad \forall (x, y) \in D, \qquad i = 1, \dots, s. \tag{3.5}$$

(*b*) *There holds*

$$C_\infty\left(\vec{f}; D\right) \geq \max_{(x, y) \in D} \mathit{rank}[\nabla_x g_i^\alpha(x, y) \colon i = 1, 2, \dots, s;\ \alpha \in \mathscr{A}]. \tag{3.6}$$

PROOF

(a) Since $C_\infty(\vec{f}; D) \leq r$, there exist infinitely differentiable functions $m_1, \dots, m_r$ and $h_1, \dots, h_s$ such that

$$f_i(x, y) = h_i(y, m_1(x), \dots, m_r(x)), \qquad \forall (x, y) \in D, \qquad i = 1, \dots, s.$$

We differentiate both sides of this equation, with respect to y. The left-hand side yields $g_i^\alpha(x, y)$. The right-hand side remains an infinitely differentiable function of $m_j(x)$, $j = 1, \dots, r$ and y, and h_i^α can be taken equal to that function.

(b) Suppose that $C_\infty(\vec{f}; D) = r$. Then Eq. (3.5) holds for some suitable functions h_i^α and for all $(x, y) \in D$. By differentiating both sides with respect to x, we obtain

$$\nabla_x g_i^\alpha(x, y) = \sum_{k=1}^{r} \frac{\partial}{\partial m_k} h_i^\alpha(y, m_1(x), \dots, m_r(x)) \cdot \nabla_x m_k(x),$$
$$\forall (x, y) \in D, \quad \forall i. \quad (3.7)$$

Thus, each column of the matrix $[\nabla_x g_i^\alpha(x, y): i = 1, 2, \dots, s;\ \alpha \in \mathscr{A}]$ is a linear combination of the vectors $\nabla_x m_1(x), \dots, \nabla_x m_r(x)$. It follows that the rank of that matrix is at most r for every $(x, y) \in D$. Q.E.D.

We now notice that any polynomial f_i can be written in the form

$$f_i(x, y) = \sum_{(\alpha_1, \dots, \alpha_n) \in \mathscr{A}} f_{i\alpha}(x) y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_n^{\alpha_n}, \quad (3.8)$$

where each $f_{i\alpha}$ is a suitable polynomial. By differentiating both sides of (3.8), setting $y = 0$, and comparing with Eq. (3.4), we see that for each i, α, there exists a positive constant $c_{i\alpha}$ such that

$$f_{i\alpha}(x) = c_{i\alpha} g_i^\alpha(x, 0), \qquad \forall x \in \Re^m. \quad (3.9)$$

Let us define

$$t = \max_{x \in \Re^m} \operatorname{rank}[\nabla f_{i\alpha}(x): i = 1, \dots, s;\ \alpha \in \mathscr{A}]. \quad (3.10)$$

Using Eq. (3.9), we see that

$$\begin{aligned} t &= \max_{x \in \Re^m} \operatorname{rank}[\nabla_x g_i^\alpha(x, 0): i = 1, 2, \dots, s;\ \alpha \in \mathscr{A}] \\ &\le \max_{(x, y) \in \Re^m \times \Re^n} \operatorname{rank}[\nabla_x g_i^\alpha(x, y): i = 1, 2, \dots, s;\ \alpha \in \mathscr{A}]. \quad (3.11) \end{aligned}$$

COROLLARY 3.1. $C_{poly}(\vec{f}; \Re^m \times \Re^n) \ge C_\infty(\vec{f}; \Re^m \times \Re^n) \ge t$.

PROOF. The first inequality is trivial since we are considering a restricted class of protocols. The second follows from (3.6) and (3.11). Q.E.D.

We make a short digression to verify that the bound t of Corollary 3.1 is a generalization Theorem 3.1.

THEOREM 3.3. *For the case $s = 1$, that is, for the problem of computing a single polynomial $f(x, y) = \sum_{\alpha \in \mathscr{A}} f_\alpha(x) y_1^{\alpha_1} \cdots y_n^{\alpha_n}$, the value of t is equal to the right-hand side of Eq.* (3.3).

PROOF. Let us fix some $x^* \in \Re^m$. Let $r(x^*)$ be the dimension of the span of $\{\partial f / \partial x_j(x^*, y), j = 1, \dots, m.\}$, where the span is formed in the vector space of functions of the variable y. We only need to show that $\max_{x^* \in \Re^n} r(x^*) = t$. Notice that

$$\nabla_x f(x^*, y) = \sum_{\alpha \in \mathscr{A}} \nabla_x f_\alpha(x^*) y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_n^{\alpha_n}.$$

Using the definition of $r(x^*)$, we see that there exist $m - r(x^*)$ linearly independent vectors $\mu_1, \mu_2, \ldots, \mu_{m-r(x^*)}$ in $\Re^m$ with

$$\sum_{j=1}^{m} \mu_{ij} \frac{\partial f}{\partial x_j}(x^*, y) \equiv 0, \qquad i = 1, \ldots, m - r(x^*), \qquad \forall y,$$

where μ_{ij} denotes the jth component of μ_i. The above relation implies that $\mu_1, \mu_2, \ldots, \mu_{m-r}(x^*)$ are orthogonal to $\nabla_x f(x^*, y)$ for all y. This is clearly equivalent to

$$\mu_i^T \nabla_x f_\alpha(x^*) = 0, \qquad \forall \alpha, \qquad i,$$

and implies that $\operatorname{rank}[\nabla_x f_\alpha \colon \alpha \in \mathscr{A}] \leq r(x^*)$. Taking the maximum over all x^*, we have $t \leq r(x^*)$. The proof of the reverse inequality is just the reverse of the preceding argument. Q.E.D.

We now come to the main result of this section, which shows that the lower bound of Corollary 3.1 is quite tight.

THEOREM 3.4. *There exists an open set $D_0 \subset \Re^m$ whose complement has Lebesgue measure zero and such that $C_{poly}(\vec{f}; D_0 \times \Re^n) \leq t + 1$.*

PROOF. We show the existence of an open set D_0 and of a set of polynomial message functions $m_1, m_2, \ldots, m_{t+1}$, such that each $f_{i\alpha}$ can be expressed in the form

$$f_{i\alpha}(x) = h_{i\alpha}(m_1(x), \ldots, m_{t+1}(x)), \qquad \forall x \in D_0, \tag{3.12}$$

where $h_{i\alpha}$ is a suitable rational function. In light of Eq. (3.8), processor P_2 is able, upon receipt of the messages $m_1(x), m_2(x), \ldots, m_{t+1}(x)$, to evaluate $f_i(x, y)$ for each i, and this will prove that $C_{poly}(\vec{f}; D_0 \times \Re^n) \leq t + 1$, as desired.

Let $F_1 = \Re$ (the field of real numbers). Let $F_3 = F_1(\{f_{i\alpha}\})$ be the field generated by the polynomials $\{f_{i\alpha} \colon i = 1, \ldots, s;\ \alpha \in \mathscr{A}\}$ over F_1. Since F_1 has characteristic 0 and F_3/F_1 is finitely generated, Theorem 2.3 applies and shows that

$$\text{tr.d.} F_3/F_1 = \dim \mathscr{D}_{F_3/F_1}. \tag{3.13}$$

Notice that we are dealing with the situation considered in the example of Section 2. In particular, Eq. (2.4) shows that

$$\dim \mathscr{D}_{F_3/F_1} = \max_{x \in \Re^m} \operatorname{rank}[\nabla f_{i\alpha}(x) \colon i = 1, \ldots, s;\ \alpha \in \mathscr{A}]. \tag{3.14}$$

By comparing with Eq. (3.10), we see that $t = \dim \mathscr{D}_{F_3/F_1}$ and using Eq. (3.13), we obtain

$$t = \text{tr.d.} F_3/F_1.$$

Let us choose a set of indices such that

$$t = \max_{x \in \Re^n} \operatorname{rank}\left[\nabla f_{i_1\alpha_1}(x), \ldots, \nabla f_{i_t\alpha_t}(x)\right],$$

and let F_2 stand for the field generated by $f_{i_1\alpha_1}, \ldots, f_{i_t\alpha_t}$ over F_1. By repeating the argument in the preceding paragraph, we obtain $t = \dim \mathscr{D}_{F_2/F_1} = \text{tr.d.}F_2/F_1$. We then invoke Theorem 2.2 to obtain

$$t = \text{tr.d.}F_3/F_1 = \text{tr.d.}F_2/F_1 + \text{tr.d.}F_3/F_2 = t + \text{tr.d.}F_3/F_2,$$

which shows that $\text{tr.d.}F_3/F_2 = 0$.

We notice that F_3 is a finitely generated extension of F_2, and F_2 clearly has characteristic zero. Therefore, we are in a position to apply Theorem 2.3 to F_3/F_2, to conclude that F_3/F_2 is a separable algebraic field extension. Since every finitely generated algebraic extension is finite (see [25, pp. 60–61]), we see that F_3/F_2 is also a finite algebraic extension. We can therefore apply the theorem of primitive element (Theorem 2.1) to F_3/F_2. This leads to the conclusion that $F_3 = F_2(f^*)$ where f^* is some linear combination (over the field F_2) of the polynomials $\{f_{i\alpha} : (i, \alpha) \neq (i_k, \alpha_k), \forall k\}$. More precisely,

$$f^* = \sum_{\alpha \in \mathscr{A}} \sum_{i=1}^{s} \epsilon_{i\alpha} f_{i\alpha}, \tag{3.15}$$

where each $\epsilon_{i\alpha}$ is an element of F_2 and where $\epsilon_{i_k\alpha_k} = 0$ for $k = 1, \ldots, t$. In particular, using the definition of F_2, each $\epsilon_{i\alpha}$ can be expressed as a rational function of $f_{i_1\alpha_1}, \ldots, f_{i_t\alpha_t}$.

Since $F_3 = F_2(f^*) = F_1(f_{i_1\alpha_1}, \ldots, f_{i_t\alpha_t}, f^*)$, it follows that each $f_{i\alpha}$ can be expressed as a rational function of the functions $f_{i_1\alpha_1}, \ldots, f_{i_t\alpha_t}, f^*$. Thus, there exist rational functions $\bar{h}_{i\alpha}$ such that

$$f_{i\alpha} = \bar{h}_{i\alpha}(f_{i_1\alpha_1}, \ldots, f_{i_t\alpha_t}, f^*). \tag{3.16}$$

Note that (3.16) is similar to (3.12) except that it refers to the equality of two elements in F_3 and that f^* need not be a polynomial. Let S be the set in $\Re^m$ on which the denominator of some of the rational functions under consideration vanishes. The set S has measure zero. Let us denote the complement of S by D_0. Clearly, D_0 is an open set. By evaluating both sides of Eq. (3.16) at an arbitrary vector $x \in D_0$, Eq. (3.12) is obtained, provided that we can replace f^* by a polynomial.

To see that f^* can be replaced by a polynomial, we recall the representation (3.15) of f^*. Since each $\epsilon_{i\alpha}$ is a rational function of $f_{i_1\alpha_1}, \ldots, f_{i_t\alpha_t}$, the function f^* can be expressed as the ratio of two polynomials, $f^* = p/q$, where q is a common multiple of the denominators of each one of the rational functions $\epsilon_{i\alpha}$. It follows that q is a polynomial function of $f_{i_1\alpha_1}, \ldots, f_{i_t\alpha_t}$. Let us consider the one-way protocol defined by $m_k = f_{i_k\alpha_k}$, $k = 1, \ldots, t$ and $\bar{m}_{t+1} = f^*$. Then, q is known to a processor who has already received the values of $f_{i_1\alpha_1}, \ldots, f_{i_t\alpha_t}$. Consequently, transmitting the value $p(x)$ (as the last message) carries the same information as transmitting the value $f^*(x)$. We have therefore constructed a one-way protocol (with $m_k = f_{i_k\alpha_k}$, $k = 1, \ldots, t$, and $m_{t+1} = p$) that uses $t + 1$ messages, and all messages are polynomial functions of the input x. Furthermore, by Eq. (3.16) and the fact that q is a polynomial function of $m_1, \ldots, m_k$, we see that Eq. (3.12) holds for some suitable rational functions $h_{i\alpha}$. Q.E.D.

In order to turn Theorem 3.4 into a useful result, one needs a computationally effective method for evaluating t^* and for constructing a protocol that uses

$t + 1$ messages. The solution to this problem is not apparent and depends on the structure of the field F_3. However, our proof does suggest a randomized procedure, which we now outline. Assuming that the number of functions $f_{i\alpha}$ is not excessive, we can evaluate the rank of the matrix consisting of the gradients $\nabla_x f_{i\alpha}$ at a random point. Obviously, except for a closed set of zero measure (an algebraic set) we find the maximum rank t, as well as polynomials $f_{i_1\alpha_1}, \ldots, f_{i_t\alpha_t}$ with the desired properties. Moreover, according to the remark following Theorem 2.2, we know that the overwhelming majority of choices of the coefficients $\epsilon_{i\alpha}$ in Eq. (3.15) are acceptable.

To summarize the results in this subsection, we have shown that (as long as we are willing to disregard a set of points of measure zero) the restriction to polynomial messages can increase the communication complexity by at most one. This is in contrast to the earlier results (Theorem 3.1) that asserted the existence of protocols that are not necessarily polynomials and whose domain is only some (possibly very small) open set.

3.2. COMPUTING POLYNOMIALS OF THE FORM $f(x+y)$. In this section, we consider the special case where all of the polynomials $f_i: \mathfrak{R}^n \times \mathfrak{R}^n \mapsto \mathfrak{R}$ to be computed are of the form

$$f_i(x,y) = \hat{f}_i(x+y), \qquad i = 1,2,\ldots,s,$$

where each $\hat{f}_i: \mathfrak{R}^n \mapsto \mathfrak{R}$ is a polynomial. We exploit this special structure and show that linear protocols (i.e., the messages are linear functions of the input) are optimal within the class of protocols that use infinitely differentiable message functions.

Let, as in the preceding subsection,

$$g_i^\alpha(x,y) = \frac{\partial^\alpha f_i}{\partial y_1^{\alpha_1} \cdots \partial y_n^{\alpha_n}}(x,y).$$

We view $\hat{f}_i$ as a function of a variable $x \in \mathfrak{R}^n$ and we define

$$\hat{g}_i^\alpha(z) = \frac{\partial^\alpha \hat{f}_i}{\partial z_1^{\alpha_1} \cdots \partial z_n^{\alpha_n}}(z).$$

Let

$$t = \max_{z\in\mathfrak{R}^n} \operatorname{rank}[\nabla_z \hat{g}_i^\alpha(z): i = 1,\ldots,s;\ \alpha \in \mathscr{A}]. \tag{3.17}$$

THEOREM 3.5. $C_\infty(\vec{f}; \mathfrak{R}^n \times \mathfrak{R}^n) = C_{linear}(\vec{f}; \mathfrak{R}^n \times \mathfrak{R}^n) = t$.

PROOF. We first prove a lower bound. Using Theorem 3.2(b), we have

$$C_\infty\left(\vec{f}; \mathfrak{R}^n \times \mathfrak{R}^n\right) \geq \max_{(x,y)\in\mathfrak{R}^n\times\mathfrak{R}^n} \operatorname{rank}[\nabla_x g_i^\alpha(x,y); i, \alpha].$$

We notice that $\hat{g}_i^\alpha(z) = g_i^\alpha(x,y)$ and $\nabla_z \hat{g}_i^\alpha(z) = \nabla_x g_i^\alpha(x,y)$, where $z = x + y$. We thus obtain

$$\begin{aligned} C_\infty\left(\vec{f}; \mathfrak{R}^n \times \mathfrak{R}^n\right) &\geq \max_{(x,y)\in\mathfrak{R}^n\times\mathfrak{R}^n} \operatorname{rank}[\nabla_x \hat{g}_i^\alpha(x+y); i, \alpha] \\ &= \max_{z\in\mathfrak{R}^n} \operatorname{rank}[\nabla_z \hat{g}_i^\alpha(z); i, \alpha] \\ &= t, \end{aligned}$$

which proves the lower bound. Given that $C_{\infty}(\vec{f}; \Re^n \times \Re^n) \leq C_{linear}(\vec{f}; \Re^n \times \Re_n)$, the proof of the theorem will be completed once we establish that $C_{linear}(\vec{f}; \Re^n \times \Re^n) \leq t$.

We first consider the case where $t = n$. In this case, we can use the protocol defined by $m_k(x) = x_k$, $k = 1, \ldots, n$. (That is, processor P_1 transmits its entire vector to processor P_2.) This is clearly a linear protocol with t messages and establishes the desired result for the case $t = n$. Notice also that the case $t > n$ cannot occur since t is the rank of a matrix with n rows.

The proof of the upper bound for the general case ($t \leq n$) proceeds by induction on n. For the basis of the induction, we consider the case where $n = 1$. If $t = n = 1$, then the result is true, by the argument of the preceding paragraph. If on the other hand $t = 0$, then $\nabla_z \hat{g}_i^{\alpha}(z) = 0$ for all $z \in \Re$ and all i, α. By letting $\alpha = (0, 0, \ldots, 0)$, we see that $\nabla_z \hat{f}_i(z) = 0$ for all z and i. Therefore, each $\hat{f}_i$ is a constant function. In this case, processor P_2 can compute $f_i(x, y)$ for each i, without receiving any messages, and $C_{linear}(\vec{f}; \Re^n \times \Re^n) = 0 = t$, as desired.

We now assume that the result has been proved for $n - 1$ ($n \geq 2$) and we prove it for n as well. The case $t = n$ has already been dealt with and we assume that $t < n$.

LEMMA 3.1. *If $t < n$, then there exists a nonzero vector $c = (c_1, c_2, \ldots, c_n) \in \Re^n$ such that*

$$\sum_{j=1}^{n} c_j \frac{\partial \hat{f}_i}{\partial z_j}(z) = 0, \qquad \forall i, z. \tag{3.18}$$

PROOF. The left-hand side of Eq. (3.18) is a polynomial, therefore, it suffices to show that the coefficient corresponding to each term $z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n}$ is identically zero. Let us denote the coefficient corresponding to the term $z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n}$ of $\partial \hat{f}_i / \partial z_j$ by $d_{\alpha}(ij)$. Then Eq. (3.18) becomes equivalent to $\sum_{j=1}^{n} c_j d_{\alpha}(ij) = 0$ for all i and α.

Let $H(z) = [\nabla_z \hat{g}_i(z); i = 1, \ldots, s; \alpha \in \mathscr{A}]$, and consider the matrix $H(0)$. Note that the column of $H(0)$ corresponding to indices i, α, is equal to

$$\alpha!(d_{\alpha}(i1), d_{\alpha}(i2), \ldots, d_{\alpha}(in)),$$

where $\alpha! \overset{\text{def}}{=} \alpha_1! \alpha_2! \cdots \alpha_n!$. (This is because the terms corresponding to $\alpha' \neq \alpha$ are either washed out by the differentiations or are set to zero when we let $z = (0, 0, \ldots, 0)$.) We have $\operatorname{rank} H(0) \leq \max_{z \in \Re^n} \operatorname{rank} H(z) = t < n$. Therefore, there exists a nonzero vector $c = (c_1, \ldots, c_n) \in \Re^n$ that is orthogonal to each one of the columns of $H(0)$. This implies that $\sum_{j=1}^{n} c_j d_{\alpha}(ij) = 0$ and concludes the proof of the lemma. Q.E.D.

Without loss of generality, we assume that $c_n \neq 0$, where c_n is the last coordinate of the nonzero vector c given by Lemma 3.1. We define an invertible linear transformation $T: \Re^n \mapsto \Re^n$ by means of the formula

$$Tz = (z_1 + c_1 z_n, z_2 + c_2 z_n, \ldots, z_{n-1} + c_{n-1} z_n, c_n z_n).$$

We show that this coordinate transformation leads to polynomials that are independent of the last coordinate of their argument, which will then allow us to use the induction hypothesis.

Consider the polynomials $\hat{f}_1', \dots, \hat{f}_s'$ and $f_1', \dots, f_s'$ defined by

$$\hat{f}_i'(z) = \hat{f}_i(Tz) = \hat{f}_i(z_1 + c_1 z_n, \dots, z_{n-1} + c_{n-1} z_n, c_n z_n), \tag{3.19}$$

$$f_i'(x, y) = \hat{f}_i'(x + y). \tag{3.20}$$

Using the chain rule and Eq. (3.18), we see that

$$\frac{\partial \hat{f}_i'}{\partial z_n} = \sum_{j=1}^{n} c_j \frac{\partial \hat{f}_i}{\partial z_j} \equiv 0.$$

Therefore, the polynomials $\hat{f}_i'$ are independent of the last coordinate of their argument and can be viewed as mappings defined on $\Re^{n-1}$ (instead of $\Re^n$).

Given that T is an invertible linear transformation, it is easily seen that the rank of the matrix considered in (3.17) does not change if each $\hat{f}_i$ is replaced by $\hat{f}_i'$. We now apply the induction hypothesis on the functions $\vec{f}' = \{f_1', \dots, f_s'\}$ to conclude that

$$C_{linear}\left(\vec{f}'; \Re^n \times \Re^n\right) \leq t.$$

Let the linear functions $m_1'(x), \dots, m_t'(x)$ correspond to a linear protocol for the problem of evaluating the functions in $\vec{f}'$. It follows that there exist polynomials $h_1', \dots, h_s'$ such that

$$f_i'(x, y) = h_i'(y, m_1'(x), \dots, m_t'(x)), \qquad \forall i, x, y.$$

Therefore,

$$\begin{aligned} f_i(x, y) &= \hat{f}_i(x + y) = \hat{f}_i'(T^{-1}(x + y)) = f_i'(T^{-1}x, T^{-1}y) \\ &= h_i'\big(T^{-1}y, m_1'(T^{-1}x), \dots, m_t'(T^{-1}x)\big), \qquad \forall i, x, y, \end{aligned}$$

where we have use of the definitions (3.19) and (3.20). Thus, the functions $m_1, \dots, m_t$ defined by $m_i(x) = m_i'(T^{-1}x)$, $i = 1, \dots, t$, define a one-way protocol for the problem of evaluating $f_1, f_2, \dots, f_s$. Furthermore, each m_i is linear, since it is the composition of linear functions. Therefore, $C_{linear}(\vec{f}; \Re^n \times \Re^n) \leq t$. This completes the induction and the proof of the theorem. Q.E.D.

We remark that the proof of Theorem 3.5 actually provides a procedure for constructing a linear and optimal protocol. Furthermore, the proof shows that we do not need to evaluate $\max_{z \in \Re^n} \operatorname{rank} H(z)$ but only the rank of $H(0)$. If the latter rank is equal to n, the problem is trivial, and if it is less than n, Lemma 3.1 applies and the problem can be reduced to one with a smaller dimension. Another point worth mentioning is that our proof actually suggests a deterministic procedure for constructing the optimal linear protocol. In fact, one can first compute the rank of $H(0)$. If $\operatorname{rank} H(0) = n$, then $m_k(x) = x_k$ is an optimal protocol. If $H(0)$ has rank less than n, then one can use, for example, Gaussian elimination method to find a nonzero vector c such that $c^T H(0) = 0$. As shown in the proof, the problem is reduced to one with a smaller dimension by a suitable change of variables. By repeating this process at most finitely many times, one will find an optimal linear protocol for computing functions f_i, $i = 1, \dots, s$.

4. *Preliminaries Continued*

In this section, we review some results (e.g., Hilbert's Nullstellensatz) from algebraic geometry (see, e.g., [4] and [10]) that will be needed in Section 5.

Let $\mathscr{C}[x_1, x_2, \dots, x_n]$ denote the ring of polynomials of variables $x_1, \dots, x_n$ over $\mathscr{C}$, the field of complex numbers. Let $f, g \in \mathscr{C}[x_1, x_2, \dots, x_n]$. We use the notation $f|g$ and say that f divides g if there exists some $h \in \mathscr{C}[x_1, x_2, \dots, x_n]$ such that $g = f \cdot h$. We say that a polynomial $g \in \mathscr{C}[x_1, x_2, \dots, x_n]$ is *irreducible* if $g = f \cdot h$ implies that either f or h is an element of $\mathscr{C}$. As is well known, $\mathscr{C}[x_1, x_2, \dots, x_n]$ is a *unique factorization ring*, that is, each one of its elements can be expressed as a product of irreducible polynomials. Furthermore, this factorization is unique up to reordering of the factors and up to multiplication of each factor by an element of $\mathscr{C}$.

Let $f_1, \dots, f_r$ be some polynomials in $\mathscr{C}[x_1, x_2, \dots, x_n]$. We define the zero set of $f_1, \dots, f_r$ by

$$V(f_1, \dots, f_r) = \{(x_1, x_2, \dots, x_n) \in \mathscr{C}^n | f_k(x_1, x_2, \dots, x_n) = 0,$$
$$\forall k \in \{1, \dots, r\}\}.$$

We now state a simple version of Hilbert's Nullstellensatz [4, p. 85] that will be used in Section 6.

THEOREM 4.1 (HILBERT'S NULLSTELLENSATZ). *Let $f_1, \dots, f_r$ be some polynomials in $\mathscr{C}[x_1, \dots, x_n]$. If $g \in \mathscr{C}[x_1, \dots, x_n]$ and $V(f_1, \dots, f_r) \subset V(g)$, then there exist some polynomials $g_1, \dots, g_r \in \mathscr{C}[x_1, \dots, x_n]$ and some positive integer k such that*

$$g^k = g_1 f_1 + g_2 f_2 + \cdots + g_r f_r. \tag{4.1}$$

Notice that if Eq. (4.1) holds, then $V(f_1, \dots, f_r) \subset V(g^k) = V(g)$. The fact that the converse is also true is exactly the content of Hilbert's theorem.

COROLLARY 4.1. *If $f, g \in \mathscr{C}[x_1, \dots, x_n]$, and if $f(x) = 0$ implies that $g(x) = 0$, that is, if $V(f) \subset V(g)$, then there is an integer k and some $h \in \mathscr{C}[x_1, \dots, x_n]$ such that $g^k = fh$. (In other words, $f | g^k$.)*

One can assign a topology to the field $\mathscr{C}^n$ by taking the family $\{V(S) | S$ is an ideal$\}$ as the closed sets. (It is a simple exercise to check that these sets satisfy the usual requirements for the closed sets of a topology.) Traditionally, this topology is called the *Zariski topology* on $\mathscr{C}$. An important property of Zariski topology is the following (see [10]).

THEOREM 4.2. *Every two nonempty Zariski open sets of $\mathscr{C}^n$ have nonempty intersection and every closed set has zero Lebesgue measure.*

5. *Two-Way Communication Complexity*

In this section we study the two-way communication complexity of evaluating a function $f: D_f \mapsto \mathscr{C}$, where D_f, the domain of f is an open subset of $\mathscr{C}^m \times \mathscr{C}^n$. Throughout, we assume that f is twice continuously differentiable on D_f.

5.1. ABELSON'S LOWER BOUND

Definition 5.1. We let $H_{xy}(f)$ be the matrix (of size $m \times n$) whose (i, j)th entry is given by $\partial^2 f / \partial x_i \partial y_j$. We use the alternative notations $(H_{xy}(f))(p)$ and

$H_{xy}(f)|_p$ to denote the value of $H_{xy}(f)$ at some vector $p \in D_f$. Also, we let $\nabla_x f$ and $\nabla_y f$ stand for the vectors of dimensions m and n (respectively) with the partial derivatives of f with respect to the components of x and y, respectively.

The following basic result has been established by Abelson [2]:[2]

THEOREM 5.1. *For any open set $D \subset D_f$ and any $p \in D$, we have*

$$C_2(f; D) \geq rank\big(H_{xy}(f)\big)(p). \tag{5.1}$$

Theorem 5.1 has an obvious corollary:

COROLLARY 5.1. *For any open set $D \subset D_f$, we have*

$$C_2(f; D) \geq \max_{p \in D} rank\big(H_{xy}(f)\big)(p). \tag{5.2}$$

The matrix $H_{xy}(f)$ is defined in terms of the cross derivatives of f and in some sense provides information on how x and y are interrelated in the formula for $f(x, y)$. On the other hand, Eq. (5.1) only takes into account the second order derivatives of f and ignores the higher-order derivatives or the first-order derivatives of f. Thus, this bound should not be expected to be tight, in general. As an example, let f be a linear function, that is, $f(x, y) = a^T x + b^T y$ ($a \in \mathscr{C}^m$, $b \in \mathscr{C}^n$, $a \neq 0$, $b \neq 0$). It is clear that $C_2(f; D) = 1$, for any nonempty open set D, while Eq. (5.1) gives a vacuous lower bound of zero. The following corollary strengthens Eq. (5.1) somewhat, by incorporating the first order derivatives of f as well. It is only a minor improvement because it can increase the lower bound by at most 1.

COROLLARY 5.2. *For any open set $D \subset D_f$, we have*

$$C_2(f; D) \geq \max_{c \in \mathscr{C}} \max_{p \in D} rank\Big[\big(H_{xy}(f)\big)(p) + c\nabla_x f(p) \cdot \nabla_y f(p)^T\Big].$$

PROOF. We notice that $C_2(f; D) \geq C_2(g \circ f; D)$, for any twice continuously differentiable function $g\colon \mathscr{C} \mapsto \mathscr{C}$, where $g \circ f$ denotes the composition of f and g. For any $p \in D$, and $c \in \mathscr{C}$, consider a function g such that $g'(f(p)) \neq 0$ and $c = g''(f(p))/g'(f(p))$. The result then follows by applying Theorem 5.1 to the function $g \circ f$. Q.E.D.

In the remainder of this section, as well as in the next section, we investigate the extent to which Abelson's bound is tight and we derive some tighter bounds. We mostly restrict attention to the case where f is a rational function and we require the messages to be rational functions of the input. In the next section, we identify two instances where Abelson's lower bound (Theorem 5.1) is tight. Then, in Section 5.3, we establish some new general lower bounds by making use of Hilbert's Nullstellensatz.

5.2. SOME CASES WHERE ABELSON'S BOUND IS TIGHT. We consider here two particular cases in which Abelson's bound (Theorem 5.1) can be shown to be tight. This is in contrast to the results in Section 6 in which it will be shown to be far from tight.

[2] This result was actually proved in [2] for real-valued functions defined on $\Re^{m+n}$ but the proof remains valid when $\Re$ is replaced by $\mathscr{C}$.

THEOREM 5.2. *Suppose that $f(x, y) = x^TQy$, where Q is a matrix of size $m \times n$ and $x \in \Re^m$, $y \in \Re^n$. Then $C_2(f; \Re^{n+m}) = rankH_{xy}(f) = rank(Q)$. In fact, the lower bound can be attained by an one-way protocol with linear messages.*

PROOF. Let $\text{rank}(Q) = r$. By Theorem 5.1, we see that $C_2(f; \Re^{n+m}) \geq \text{rank}(H_{xy}(f) = \text{rank}(Q) = r$. To prove the other direction of the inequality, we present an one-way linear protocol that uses exactly r messages. Using the singular value decomposition of Q, there exist vectors $u_1, \ldots, u_r \in \Re^m$ and $v_1, \ldots, v_r \in \Re^n$ such that

$$Q = u_1v_1^T + u_2v_2^T + \cdots + u_rv_r^T,$$

from which we obtain

$$x^TQy = x^Tu_1v_1^Ty + x^Tu_2v_2^Ty + \cdots + x^Tu_rv_r^Ty. \tag{5.3}$$

Notice that in Eq. (5.3) each one of the expressions x^Tu_i and v_i^Ty is a scalar. Thus, the one-way protocol with r linear messages, defined by $m_i(x) = x^Tu_i$, $i = 1, \ldots, r$, is adequate for computing f. Q.E.D.

Theorem 5.2 states that Abelson's bound is tight for homogeneous quadratic polynomials. What happens for polynomials of degree greater than 2? In what follows, we show the tightness of Abelson's bound for computing functions of the form: $f(x, y) = g(x + y)$, where g is a nonlinear homogeneous polynomial in no more than four variables. Although this result determines a case for which $C_2(f; \Re^{2n})$ can be determined completely, it is of little use in practice. This is because we have $n \leq 4$ and the naive protocol $m_i(x) = x_i$, $i = 1, \ldots, n$, uses at most four messages and cannot be too far from being optimal. Our result makes use of the following theorem proved by Gordan and Nöether in 1876 [9].

THEOREM 5.3. *Let $f: \Re^n \mapsto \Re$ be a nonlinear homogeneous polynomial in $n \leq 4$ variables and let $H(f)$ be its Hessian matrix, that is, the matrix with entries $(\partial^2 f/\partial x_i \partial x_j)$. If $detH(f) \equiv 0$, then there exists a linear mapping T from $\Re^n$ onto $\Re^{n-1}$ and a homogeneous polynomial $g: \Re^{n-1} \to \Re$ such that $f(x) = g(Tx)$.*

Our result is the following:

THEOREM 5.4. *Let $g: \Re^n \mapsto \Re$ be a nonlinear homogeneous polynomial and let the polynomial $f: \Re^{2n} \mapsto \Re^n$ be defined by $f(x, y) = g(x + y)$. If $n \leq 4$, then*

$$C_2(f; \Re^{2n}) = \max_{(x,y) \in \Re^{2n}} rankH_{xy}(f)|_{(x,y)} = C_{linear}(f; \Re^{2n}).$$

PROOF. Let $z = x + y$. We regard g as a nonlinear polynomial in the variable $z \in \Re^n$. Let k be the smallest integer such that there exists some linear mapping T from $\Re^n$ onto $\Re^k$ and some homogeneous polynomial $\hat{g}: \Re^k \mapsto \Re$ such that $g(z) = \hat{g}(Tz)$, $\forall z$. Since g is nonlinear and T is linear, we see that $\hat{g}$ is also nonlinear. We claim that there exists some vector $\hat{z} = (\hat{z}_1, \ldots, \hat{z}_k) \in \Re^k$ at which $H(\hat{g})$ is nonsingular. Indeed, if this is not so, then by Theorem 5.3, there exists another linear mapping $\bar{T}$ from $\Re^k$ onto $\Re^{k-1}$ and some homogeneous polynomial $\bar{g}: \Re^{k-1} \mapsto \Re$ such that $\hat{g}(\hat{z}) = \bar{g}(\bar{T}\hat{z})$, for all $\hat{z}$. But this implies that $g(z) = \bar{g}(\bar{T}Tz)$. Since the composition of T and $\bar{T}$ maps $\Re^n$ onto $\Re^{k-1}$, this contradicts the definition of k.

A simple calculation shows that $H(g)|_z = T^T H(\hat{g})|_{Tz} T$. Since T maps $\mathfrak{R}^n$ onto $\mathfrak{R}^k$, the matrix T has full column rank and we obtain rank $H(g)|_z =$ rank $H(\hat{g})|_{Tz}$. Since the range of T is all of $\mathfrak{R}^k$, we have

$$\max_{z \in \mathfrak{R}^n} \operatorname{rank} H(g)|_z = \max_{\hat{z} \in \mathfrak{R}^k} \operatorname{rank} H(\hat{g})|_{\hat{z}} = k.$$

Since $H_{xy}(f)|_{(x,y)} = H(g)|_{z=x+y}$, we obtain that $\max_{x,y}\text{rank } H_{xy}(f) = k$. It then follows from Theorem 5.1 that $C_2(f; \mathfrak{R}^{2n}) \geq k$. To establish the reverse inequality, we present a protocol for computing f that uses exactly k messages. Let $m_i(x) = T_i x$, $i = 1, \dots, k$, where T_i is the ith row of the matrix T. Then,

$$\begin{aligned} f(x,y) &= g(x+y) = \hat{g}(T(x+y)) = \hat{g}(T_1(x+y), \dots, T_k(x+y)) \\ &= \hat{g}(m_1(x) + T_1 y, \dots, m_k(x) + T_k y). \end{aligned}$$

This last formula shows that f can be computed using the one-way protocol with messages $m_i(x) = T_i x$, $i = 1, \dots, k$. In particular, $C_2(f; \mathfrak{R}^{2n}) \leq C_{linear}(f; \mathfrak{R}^{2n}) \leq k$, which completes the proof. Q.E.D.

Unfortunately, Theorem 5.4 is not true for the case $n > 4$, for the simple reason that Theorem 5.3 fails to hold. Historically, Hess had published a paper in which he gave an erroneous proof of Theorem 5.3 for all n. It was later discovered by Gordan and Nöether that Hess' proof was incorrect and proved that the largest value of n for which Theorem 5.3 holds is 4.

5.3. Some New Lower Bounds. Throughout this subsection we assume that $f\colon D_f \mapsto \mathscr{C}$ is a complex rational function, where $D_f \subset \mathscr{C}^m \times \mathscr{C}^n$ is the set of vectors (x, y) at which f is finite. In this context, it is natural to consider "rational" protocols, in which the messages transmitted are rational functions of the input data (x, y).

We present two new methods for establishing lower bounds on the two-way communication complexity in this setting. The first method provides lower bounds on $C_{rat}(f; D)$ for any open subset $D \in D_f$. The second method requires that $D = D_f$ but usually gives sharper lower bounds.

Our first method (Theorem 5.5–5.7) exploits the fact that any rational protocol can be converted into a protocol in which the messages are polynomial function of (x, y) and that uses at most twice as many messages:

Theorem 5.5. *Let f be a rational function and let D be an open subset of D_f. Then there holds*

$$C_{rat}(f; D_f) \leq C_{poly}(f; D_f) \leq 2C_{rat}(f; D_f).$$

The idea behind the proof of Theorem 5.5 is that each rational message of a rational protocol can be replaced by two polynomial messages consisting of the numerator and denominator polynomials (respectively) of the original message. The proof can be found in [12] and is omitted because it is relatively straightforward and also because Theorem 5.5 will not be invoked in subsequent proofs.

Suppose that $f(x,y) = p(x,y)/q(x,y)$ where p and q are two relatively prime polynomials. Let $D_f = \{(x,y) | q(x,y) \neq 0\}$ and let $D \subset D_f$ be an open subset. Consider some rational protocol $\pi \in \Pi_{rat}(f; D)$ with r messages, where $r = C_{rat}(f; D)$ (cf. Section 1). Then, by Theorem 5.5, there exists a polynomial protocol $\pi' \in \Pi_{poly}(f; D)$ that uses $2r$ messages. Let $m_1, \ldots, m_{2r}$ be the message functions of the protocol π'. Assuming that processor P_1 performs the final evaluation of $f(x, y)$, we must have $f(x, y) = h(x, m_1(x, y), \ldots, m_{2r}(x, y))$ for all $(x, y) \in D$, where h is a rational function. Since h is rational, we must have $f(x, y) = p'(x, y)/q'(x, y)$, where p' and q' are some polynomials whose values (on the set D) are completely determined by the values of message functions $m_1, \ldots, m_{2r}$ and x. This implies that $C_{poly}(p'; D) \leq 2r$ and $C_{poly}(q'; D) \leq 2r$. Notice that $p/q = p'/q'$. Using the unique factorization property of rational functions over $\mathscr{C}$ (cf. Section 4), we see that $p' = pg$ and $q' = qg$ for some nonzero polynomial g. We conclude that there exists some nonzero polynomial g such that

$$C_{rat}(f; D) \geq \frac{1}{2} C_{poly}(pg; D)$$

and

$$C_{rat}(f; D) \geq \frac{1}{2} C_{poly}(qg; D).$$

This shows that we can bound from below the communication complexity of f by bounding from below the communication complexity of pg or qg. The difficulty, however, is that the polynomial g is not known and we are forced to develop a bound which is valid for an arbitrary choice of g. Ideally, we would like to be able to say that if p has high communication complexity then the same is true for pg. Although this does not seem to be true in general, the following result makes a step in that direction:

THEOREM 5.6. *Let $f, g \in \mathscr{C}[x_1, \ldots, x_m, y_1, \ldots, y_n]$ be two nonzero polynomials that are relatively prime. Then,*

$$C_2(fg; \mathscr{C}^{m+n}) \geq \max_{(x,y) \in V(f)} rankH_{xy}(f)|_{(x,y)} - 2,$$

where $V(f) = \{(x, y) | f(x, y) = 0\}$ is the zero set of polynomial f.

PROOF. By Theorem 5.1, we have

$$\begin{aligned} C_2(fg; \mathscr{C}^{m+n}) &\geq \max_{(x,y) \in \mathscr{C}^{m+n}} \operatorname{rank}\big(H_{xy}(fg)\big)|_{(x,y)} \\ &= \max_{(x,y) \in \mathscr{C}^{m+n}} \operatorname{rank}\Big(f(x,y) H_{xy}(g)|_{(x,y)} + g(x,y) H_{xy}(f)|_{(x,y)} \\ &\quad + (\nabla_x f(x,y))\big(\nabla_y g(x,y)\big)^T + (\nabla_x g(x,y))\big(\nabla_y f(x,y)\big)^T\Big) \\ &\geq \max_{(x,y) \mathscr{C}^{n+m}} \operatorname{rank}\big(f(x,y) H_{xy}(g)|_{(x,y)} + g(x,y) H_{xy}(f)|_{(x,y)}\big) - 2 \\ &\geq \max_{(x,y) \in V(f)} \operatorname{rank}\big(g(x,y) H_{xy}(f)|_{(x,y)}\big) - 2, \end{aligned} \tag{5.4}$$

where the third step follows from the fact $(\nabla_x f(x, y))(\nabla_y g(x, y))^T$ and $(\nabla_x g(x, y))(\nabla_y f(x, y))^T$ have rank at most 1. Choose some $(x_0, y_0) \in V(f)$ such that

$$\operatorname{rank}\big(H_{xy}(f)|_{(x_0, y_0)}\big) = \max_{(x,y) \in V(f)} \operatorname{rank} H_{xy}(f)|_{(x,y)} = r.$$

Then, there exists a submatrix M of size $r \times r$ embedded in $H_{xy}(f)$, which is nonsingular at (x_0, y_0). We view this submatrix as a function of (x, y) and we consider its determinant $\det(M)$, which is a polynomial in (x, y). We have just shown that $V(f)$ is not contained in $V(\det(M))$. In other words, if we write f as a product of irreducible polynomials, then at least one of the irreducible factors of f, call it f_1, does not divide $\det(M)$. But since f and g are relatively prime, it follows that f_1 does not divide g either. We conclude that f_1 does not divide $g \cdot \det(M)$. We now claim that $V(f) \not\subset V(g \cdot \det(M))$. If the claim is not true, then $V(f) \subset V(g \cdot \det(M))$. Hilbert's Nullstellensatz applies and shows that $(g \cdot \det(M))^k = fh$ for some positive integer k and some polynomial h. By the unique factorization property, we see that the irreducible polynomial f_1 would have to be a factor of either g or $\det(M)$, which is a contradiction and establishes our claim.

Since $V(f) \not\subset V(g \cdot \det(M))$, there exists some $(x^*, y^*) \in V(f)$ such that $g(x^*, y^*)\det(M)|_{(x^*, y^*)} \neq 0$. Consequently,

$$\begin{aligned}\max_{(x,y)\in V(f)} \operatorname{rank}\left(g(x,y)H_{xy}(f)|_{(x,y)}\right) &\geq \operatorname{rank}\left(g(x^*,y^*)H_{xy}(f)|_{(x^*,y^*)}\right)\\ &= r\\ &= \max_{(x,y)\in V(f)} \operatorname{rank}\left(H_{xy}(f)|_{(x,y)}\right).\end{aligned}$$

which when combined with (5.4) completes the proof of the theorem. Q.E.D.

The above theorem states that if rank $H_{xy}(f)$ is large for some $(x, y) \in V(f)$, then fg also has large communication complexity for any polynomial g which is relatively prime to f. Unfortunately, Theorem 5.6 is not always sufficient for proving tight lower bounds for fg because there exist functions f for which rank $H_{xy}(f)$ is small for every $(x, y) \in V(f)$ even though $H_{xy}(f)$ has high rank when the restriction $(x, y) \in V(f)$ is removed. A specific example will be seen in the next section.

The following is a result from algebraic geometry that gives a sufficient condition on f under which $H_{xy}(f)$ has high rank at some point belonging to $V(f)$.

THEOREM 5.7. *Let $\hat{f}$ be a nonlinear homogeneous polynomial in n variables such that $\nabla\hat{f}(x) \neq 0$ for every $x \in V(\hat{f})$. Let the polynomial $f: \mathscr{C}^{2n} \mapsto \mathscr{C}$ be defined by $f(x, y) = \hat{f}(x + y)$. Then,*

$$\max_{(x,y)\in V(f)} \mathit{rank}\left(H_{xy}(f)|_{(x,y)}\right) \geq n - 1.$$

The proof of the above theorem can be found in [24] and [11]. As an immediate consequence of Theorems 5.6 and 5.7, we have the following:

COROLLARY 5.3. *Let f and $\hat{f}$ be as in Theorem 5.7. Then,*

$$C_2(f \cdot g; \mathscr{C}^{2n}) \geq n - 3$$

for any polynomial g that is relatively prime to f.

Unfortunately, the above corollary is not easy to apply, because the set $V(\hat{f})$ is usually hard to determine. Accordingly, the condition $\nabla\hat{f}(x) \neq 0$ on the set $V(\hat{f})$ cannot be easily tested. In fact, it seems a lot easier to just compute the rank of $H_{xy}(f)$ at a random point of $V(f)$ because $\max_{(x,y)\in V(f)}$ rank $H_{xy}(f)|_{(x,y)}$ is attained at the majority of points on $V(f)$ (a Zariski open set of $V(f)$).

We have so far shown that lower bounds on the communication complexity of a rational function $f = p/q$ (p and q are relatively prime) can be obtained by developing lower bounds on the communication complexity of pg or qg, where g is an arbitrary nonzero polynomial. We now develop our second method for establishing lower bounds by exploiting the fact that if the domain of a protocol is the set D_f on which f is finite, then the polynomial g is not entirely arbitrary. We have shown earlier that if f can be evaluated by a rational protocol with domain D_f, then there exist polynomials p' and q' such that $f(x,y) = p'(x,y)/q'(x,y)$ for all $(x,y) \in D_f$ and $C_{poly}(f; D_f) \geq C_{poly}(q'; D_f)/2$. The polynomial q' must certainly satisfy $q' = qg$, for some g, but it must also be nonzero at every point in the domain D_f of f because otherwise the expression $p'(x,y)/q'(x,y)$ will be meaningless for some $(x,y) \in D_f$. This additional constraint is used in an essential way in the following result.

THEOREM 5.8. *Suppose that f is a rational function and that $f = p/q$, where p, $q \in \mathscr{C}[x_1,\dots,x_m,y_1,\dots,y_n]$ are relatively prime polynomials. If q is irreducible, then*

(*a*)

$$C_{rat}(f; D_f) \geq \max_{(x,y)\in\mathscr{C}^m\times\mathscr{C}^n} rankH_{xy}(q)|_{(x,y)} - 1. \tag{5.5}$$

(*b*)

$$C_{rat}(f; D_f) \geq \frac{1}{2} \max_{(x,y)\in\mathscr{C}^m\times\mathscr{C}^n} rankH_{xy}(p)|_{(x,y)} - \frac{3}{2}. \tag{5.6}$$

PROOF

(a) Consider a rational protocol for computing f on D_f that uses $r = C_{rat}(f; D_f)$ messages and let $m_1,\dots,m_r\colon D_f \mapsto \mathscr{C}$ be the corresponding message functions. We first consider the special case where each one of the message functions is a polynomial. Without loss of generality, we assume that the final evaluation of the function f is performed by processor P_1. By the definition of a rational protocol (cf. Section 1), there exists a rational function h such that $f(x,y) = h(x, m_1(x,y),\dots,m_r(x,y))$ for all $(x,y) \in D_f$. Note that h can be expressed in the form

$$h(x, m_1(x,y),\dots,m_r(x,y)) = \frac{h_1(x, m_1(x,y),\dots,m_r(x,y))}{h_2(x, m_1(x,y),\dots,m_r(x,y))},$$

where h_1 and h_2 are relatively prime polynomials. Let $h'_2(x,y) = h_2(x, m_1(x,y),\dots,m_r(x,y))$. The functions $m_1,\dots,m_r$ were originally defined on D_f. On the other hand, since they are polynomials they can be uniquely extended to polynomial functions on the entire of $\mathscr{C}^{m+n}$. Furthermore, the representation $h'_2(x,y) = h_2(x, m_1(x,y),\dots,m_r(x,y))$ must be also valid over $\mathscr{C}^{m+n}$ and this implies that $C_{poly}(h'_2; \mathscr{C}^{m+n}) \leq r$. We now notice that we must have $h'_2(x,y) \neq 0$ for all $(x,y) \in D_f$, because the function h must be defined for all $(x,y) \in D_f$. Equivalently, $V(h'_2) \subset V(q)$, where q is the denominator polynomial of f, assumed irreducible. Hilbert's Nullstellensatz shows that $q^k = h'_2 g$, for some polynomial g and some positive integer k. We factor the polynomial $h'_2 g$ as a product of irreducible factors. Since q is irreducible, it follows that each one of these factors must be equal to q. We conclude that $h'_2 = cq^K$ for some nonzero constant $c \in \mathscr{C}$ and some positive integer K, and therefore $r \geq C_{poly}(h'_2; \mathscr{C}^{m+n}) = C_{poly}(q^K; \mathscr{C}^{m+n})$.

We now consider the case where there exists some i such that the message function m_i is not a polynomial and let us choose, in particular, the first such index i. Suppose, without loss of generality, that $i \in T_{1 \to 2}$. We have $m_i(x, y) = \hat{m}_i(x, m_1(x, y), \dots, m_{i-1}(x))$ for some rational function $\hat{m}_i$ and each one of the functions $m_1, \dots, m_{i-1}$ is a polynomial. We write $\hat{m}_i$ in the form $\hat{m}_i = h_1/h_2$, where h_1 and h_2 are relatively prime polynomials, so that

$$\hat{m}_i(x, m_1(x, y), \dots, m_{i-1}(x, y)) = \frac{h_1(x, m_1(x, y), \dots, m_{i-1}(x, y))}{h_2(x, m_1(x, y), \dots, m_{i-1}(x, y))}.$$

Let $h'_2(x, y) = h_2(x, m_1(x, y), \dots, m_{i-1}(x, y))$. We now repeat the argument of the preceding paragraph. Since the domain of the protocol is all of D_f, it follows that $V(h'_2) \subset V(q)$ and $h'_2 = cq^K$ for some nonzero constant $c \in \mathscr{C}$ and some positive integer K. Notice that h'_2 can be expressed as a polynomial function of x, $m_1(x, y), \dots, m_{i-1}(x, y)$, and this implies that $r > i - 1 \geq C_{poly}(h'_2; \mathscr{C}^{m+n}) = C_{poly}(q^K; \mathscr{C}^{m+n})$.

To summarize, we have shown that in both cases that there exists a positive integer K for which $C_{rat}(f; D_f) = r \geq C_{poly}(q^K; \mathscr{C}^{m+n})$. It now remains to derive a lower bound on $C_{poly}(q^K; \mathscr{C}^{m+n})$. To this effect, we apply Theorem 5.1. We have

$$\begin{aligned}
C_{poly}(q^K; \mathscr{C}^{m+n}) &\geq C_2(q^K; \mathscr{C}^{m+n}) \\
&\geq \max_{(x,y) \in \mathscr{C}^{m+n}} \operatorname{rank} H_{xy}(q^K)|_{(x,y)} \\
&= \max_{(x,y) \in \mathscr{C}^{m+n}} \operatorname{rank}\Big(Kq^{K-1}(x, y) H_{xy}(q)|_{(x,y)} \\
&\qquad + K(K-1) q^{K-2}(x, y) (\nabla_x q(x, y)) \big(\nabla_y q(x, y)\big)^T \Big) \qquad (5.7) \\
&\geq \max_{(x,y) \in \mathscr{C}^{m+n}} \operatorname{rank}\Big(Kq^{K-1}(x, y) H_{xy}(q)|_{(x,y)} \Big) - 1 \qquad (5.8) \\
&\geq \max_{(x,y) \in \mathscr{C}^{m+n}} \operatorname{rank} H_{xy}(q)|_{(x,y)} - 1, \qquad (5.9)
\end{aligned}$$

Here, the first equality (5.7) is a simple calculation and the next step (5.8) is due to the fact $(\nabla_x q)(\nabla_y q)^T$ has rank at most 1. The last step is obtained as follows: The set $\{(x, y) | q(x, y) \neq 0\}$ is a Zariski open set. Furthermore, the maximum rank of $H_{xy}(q)$ is attained at the set of points where the determinant of a suitable submatrix of $H_{xy}(q)$ does not vanish and is also a Zariski open set. Since every two nonempty Zariski open sets have nonempty intersection (Theorem 4.2), it suffices to consider a vector (x, y) in the intersection of these two sets.

(b) Let (x, y) be an arbitrary vector of D_f. Note that

$$H_{xy}\left(\frac{p}{q}\right) = \frac{1}{q}H_{xy}(p) - \frac{p}{q^2}H_{xy}(q) - \frac{2}{q^2}(\nabla_x p)(\nabla_y q)^T + \frac{2p}{q^3}(\nabla_x q)(\nabla_y q)^T.$$

By evaluating the rank of both sides at (x, y) and noticing that both $(\nabla_x p)(\nabla_y q)^T|_{(x,y)}$ and $(\nabla_x q)(\nabla_y q)^T|_{(x,y)}$ have rank at most 1, we see that

$$\operatorname{rank} H_{xy}\left(\frac{p}{q}\right)\bigg|_{(x,y)} \geq \operatorname{rank} H_{xy}(p)|_{(x,y)} - \operatorname{rank} H_{xy}(q)|_{(x,y)} - 2.$$

Therefore,

$$\begin{aligned} C_{rat}(f; D_f) &\geq C_2(f; D_f) \\ &\geq \operatorname{rank} H_{xy}\left(\frac{p}{q}\right)\bigg|_{(x,y)} \\ &\geq \operatorname{rank} H_{xy}(p)|_{(x,y)} - \operatorname{rank} H_{xy}(q)|_{(x,y)} - 2 \\ &\geq \operatorname{rank} H_{xy}(p)|_{(x,y)} - C_{rat}(f; D_f) - 3, \qquad \forall (x,y) \in D_f, \end{aligned}$$

where the last step follows from Eq. (5.5). After rearranging the above inequality, we see that

$$C_{rat}(f; D_f) \geq \frac{1}{2}\operatorname{rank} H_{xy}(p)|_{(x,y)} - \frac{3}{2}, \qquad \forall (x,y) \in D_f. \tag{5.10}$$

Since $\max_{(x,y)\in \mathcal{C}^{m+n}} \operatorname{rank} H_{xy}(p)$ is attained at a Zariski open set and D_f is also a Zariski open set, by Theorem 4.2, there exists some vector $(x^*, y^*) \in D_f$ such that

$$\max_{(x,y)\in \mathcal{C}^{m+n}} \operatorname{rank} H_{xy}(p) = \operatorname{rank} H_{xy}(p)|_{(x^*,y^*)}.$$

Now Eq. (5.6) becomes evident when one considers (5.10) at (x^*, y^*). Q.E.D.

6. *An* $\Omega(n^2)$ *Lower Bound for Computing* $[(x+y)^{-1}]_{11}$

Let x and y be $n \times n$ matrices. As an application of the results of Section 5, we consider the communication complexity of the function $f(x, y) = [(x+y)^{-1}]_{11}$ (the (1,1)th entry of $(x+y)^{-1}$) within the class of rational protocols. Although Abelson's lower bound is only $\Omega(n)$, we derive a lower bound of $n^2 - 1$, which is almost equal to the obvious upper bound of n^2. In particular, this example will show that Abelson's bound can be far from tight.

We motivate our choice of the problem. The value of $[(x+y)^{-1}]_{11}$ can be thought of as the first entry of the solution of the system of linear equations: $(x+y)u = b$, where $b = (1, 0, \ldots, 0)$ and u is the unknown. Thus, the problem under consideration captures the essential difficulties of a distributed solution of a system of the form $(x+y)u = b$, when x and y are possessed by different processors. Since the solution of linear systems of equations is the most basic problem in numerical computation, the problem we are studying is an interesting paradigm.

It is easy to see that n^2 messages would be needed if we had required that a particular processor, say P_1, should eventually evaluate all entries of the inverse matrix $(x+y)^{-1}$. (This is because P_1 could then invert $(x+y)^{-1}$ to obtain $x+y$ and use its knowledge of x to infer the value of y, and this is possible only if at least n^2 messages have been exchanged.) However, the fact that the evaluation of the whole inverse matrix $(x+y)^{-1}$ is hard does not imply that the computation of a particular entry is also difficult. In fact, we shall see that the derivation of tight bounds on the communication complexity of $[(x+y)^{-1}]_{11}$ is surprisingly hard. As a first indication, we show that Abelson's result (Theorem 5.1) gives only an $\Omega(n)$ lower bound.

THEOREM 6.1. *Let* $f(x,y) = [(x+y)^{-1}]_{11}$. *Then*

$$\mathit{max}_{(x,y)\in D_f} \mathit{rankH}_{xy}(f)|_{(x,y)} \le 3n. \tag{6.1}$$

PROOF. Let us fix a pair $p = (x_0, y_0) \in D_f$ of $n \times n$ matrices. We show that the rank of $H_{xy}(f)|_p$ is at most $3n$. Let Δ_1, Δ_2 be two $n \times n$ perturbation matrices. We consider the Taylor series expansion of f at the point p:

$$\begin{aligned} f(x_0+\Delta_1, y_0+\Delta_2) &= \Big[((x_0+y_0)+(\Delta_1+\Delta_2))^{-1}\Big]_{11} \\ &= \Big[(x_0+y_0)^{-1}\Big]_{11} - \Big[(x_0+y_0)^{-1}(\Delta_1+\Delta_2)(x_0+y_0)\Big]_{11} \\ &\quad + \Big[(x_0+y_0)^{-1}((\Delta_1+\Delta_2)(x_0+y_0))^2\Big]_{11} + \cdots \end{aligned}$$

Notice that the value of $H_{xy}(f)|_p$ is completely determined by the second-order terms of this expansion. Thus, if we let

$$g(\Delta_1, \Delta_2) = \Big[(x_0+y_0)^{-1}((\Delta_1+\Delta_2)(x_0+y_0))^2\Big]_{11},$$

then $H_{xy}(f)|_{(x_0,y_0)} = H_{\Delta_1\Delta_2}(g)|_{(0,0)}$. Therefore, we only need to show that rank $H_{\Delta_1\Delta_2}(g)|_{(0,0)} \le 3n$. We present a two-way polynomial protocol for computing g that only uses $3n$ messages.

Notice that as far as the computation of g is concerned, the matrices x_0, y_0 are constant and the matrices Δ_i $(i = 1, 2)$ are the inputs. Let $e = (1, 0, \ldots, 0)^T$. The protocol proceeds as follows:

(1) Processor P_1 sends the vector $\Delta_1(x_0+y_0)e$ to processor P_2 (n messages).

(2) Processor P_2 computes $(\Delta_1+\Delta_2)(x_0+y_0)e$ and sends the following two vectors ($2n$ messages) to P_1:

$$(\Delta_1+\Delta_2)(x_0+y_0)e$$

and

$$\Delta_2(x_0+y_0)(\Delta_1+\Delta_2)(x_0+y_0)e.$$

(3) Once processor P_1 receives these messages, it can use its knowledge of Δ_1 to evaluate $((\Delta_1+\Delta_2)(x_0+y_0))^2e$. It follows that $g(\Delta_1, \Delta_2) = [(x_0+y_0)^{-1}((\Delta_1+\Delta_2)(x_0+y_0))^2]_{11}$ can also be evaluated by P_1.

By Abelson's result (Theorem 5.1), we see that for any open set D containing $(0,0)$, we have

$$\text{rank } H_{\Delta_1\Delta_2}(g)|_{(0,0)} \le C_{poly}(g; D) \le 3n,$$

which completes the proof. Q.E.D.

Let D_f be the set of all $(x, y) \in \mathscr{C}^{n^2} \times \mathscr{C}^{n^2}$ at which the rational function $f(x, y) = [(x + y)^{-1}]_{11}$ is well defined. Clearly, D_f is the same as the set of all (x, y) such that $\det(x + y) \neq 0$. Our main result is the following:

THEOREM 6.2

$$C_{rat}(f; D_f) \geq n^2 - 1. \tag{6.2}$$

The proof is based on two lemmas:

LEMMA 6.1. *The polynomial $g(x, y) = det(x + y)$ is irreducible.*

LEMMA 6.2. *Suppose that $n > 1$ and let $g(x, y) = det(x + y)$. Then the rank of $H_{xy}(g)$ evaluated at $(I, 0)$ (I is the identity matrix) is n^2.*

Once these two lemmas are proved, the desired result is obtained as follows: If $n = 1$, then Eq. (6.2) holds trivially. For $n > 1$, we have $f(x, y) = \det_{11}(x + y)/\det(x + y) = \det_{11}(x + y)/g(x, y)$, where $\det_{11}(x + y)$ is the cofactor of the $(1, 1)$th entry of $x + y$. It is seen that $g(x + y)$ does not divide $\det_{11}(x + y)$, because otherwise $[(x + y)^{-1}]_{11}$ would be a polynomial in the entries of x and y, which is easily shown not to be the case. Since g is irreducible (Lemma 6.1), we conclude that the polynomials $\det_{11}(x + y)$ and $g(x, y)$ are relatively prime. Then, Theorem 5.8 applies and shows that

$$C_{rat}(f; D_f) \geq \max_{(x, y) \in \mathscr{C}^{2n^2}} H_{xy}(g)|_{(x, y)} - 1 \geq n^2 - 1,$$

where the last inequality has made use of Lemma 6.2. Thus, it only remains to prove the two lemmas.

PROOF OF LEMMA 6.1. In $n = 1$, then $g(x, y) = x + y$, which is obviously an irreducible polynomial. For $n > 1$, we assume, in order to derive a contradiction, that $g(x, y) = A(x, y)B(x, y)$ where A, B are nonconstant polynomial functions of the entries of x, y. Let x_{ij} (respectively, y_{ij}) denote the (i, j)th entry of x (respectively, y). Let us restrict x and y by letting $x_{1i} = -y_{1i} = 1$, $i = 2, \ldots, n$. With such a restriction, g, A, and B can be expressed as polynomials $\hat{g}$, $\hat{A}$, and $\hat{B}$, respectively, of the unrestricted variables. Note that

$$\hat{g}(x, y) = (x_{11} + y_{11})\det_{11}(x + y) = \hat{A}(x, y)\hat{B}(x, y).$$

By the unique factorization property of polynomials, we see that $(x_{11} + y_{11})$ must be a factor of either $\hat{A}(x, y)$ or $\hat{B}(x, y)$. Since $\det(x + y)$ is a linear function of $x_{11} + y_{11}$, we conclude that x_{11}, y_{11} appear together in either $\hat{A}$ or $\hat{B}$, but not in both. It then follows that x_{11}, y_{11} appear together in either $A(x, y)$ or $B(x, y)$, but not in both. Repeating our argument for all (i, j) $(1 \leq i, j \leq n)$, we see that either x_{ij} and y_{ij} both appear only in $A(x, y)$ or they both appear only in $B(x, y)$. Therefore, the set $\{(i, j),\ i, j = 1, 2, \ldots, n\}$ can be partitioned into two subsets R_1, R_2 (with R_1 being nonempty) such that $A(x, y)$ depends only on the entries x_{ij}, y_{ij} with $(i, j) \in R_1$ and $B(x, y)$ depends only on the entries x_{ij}, y_{ij} with $(i, j) \in R_2$. Let us express each one of the polynomials A and B as a sum of products and then carry out the cross-multiplications to expand $A(x, y)B(x, y)$ as a sum of products. Since A and B depend on different entries, it is seen that this expansion leads to no cancellations. Hence, if (i, j) is in R_1, then (i, k) and (k, j), $k = 1, \ldots, n$, also belong R_1, since otherwise there would be a term in the expansion of

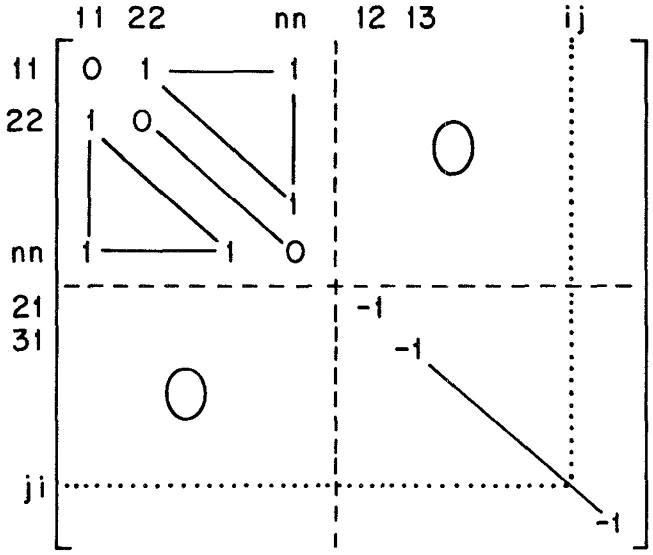

FIGURE 1

$A(x, y)B(x, y) = \det(x + y)$ with two entries from the same row or column. This implies that all of the entries must be in R_1, and R_2 is empty. Consequently, $B(x, y)$ is a constant polynomial, which contradicts our original assumption. Q.E.D.

PROOF OF LEMMA 6.2. An easy calculation yields

$$\left.\frac{\partial^2 g}{\partial x_{ij}\partial y_{lm}}\right|_{(I,0)} = \begin{cases} 1 & \text{if } i = j,\ l = m \text{ and } i \neq l \\ -1 & \text{if } i = m,\ j = l \text{ and } i \neq l \\ 0 & \text{otherwise.} \end{cases}$$

Thus, if the rows and columns of $H_{xy}(g)|_{(I,0)}$ are suitably rearranged, the matrix $H_{xy}(g)|_{(I,0)}$ has the structure shown in Fig. 1. It is not hard to see that this matrix is nonsingular and therefore has rank n^2. Q.E.D.

We would like to be able to strengthen Theorem 6.2 in a number of directions. First, Theorem 6.2 refers to the computation of $[(x + y)^{-1}]_{11}$, where x, y are complex matrices. This does not lead to a lower bound when we restrict x and y to be real, even though this is the case of main practical interest. A related deficiency is that the lower bound applies only to protocols whose domain is equal to all of D_f. It would be interesting to know whether the communication complexity of the problem can be reduced by an order of magnitude when we restrict to real matrices, or if we only consider the evaluation of f in an open set of real matrices. We conjecture that this is not the case, but we are not aware of any proof technique that could lead to such a result.

One possible approach for proving a stronger lower bound is based on Theorem 5.6 of Section 5. This result shows that an $\Omega(n^2)$ lower bound will be established if we manage to find a pair (x, y) of matrices such that $g(x, y) = 0$ and rank $H_{xy}(g)|_{(x,y)} = \Omega(n^2)$, where $g(x, y) = \det(x + y)$. Unfortunately, the determinant function is particularly nasty in that respect. It can be shown [12] that the rank of $H_{xy}(g)$ is n^2 at each point (x, y) such that $x + y$ is invertible but it is no more than $3n + 3$ at each point (x, y) at which $g(x, y) = 0$.

Finally, let us mention that an $\Omega(n^2)$ bound can also be obtained for the special case where x and y are restricted to be symmetric matrices. The proof is similar to the proof of Theorem 6.2.

7. *Conclusions and Extensions*

We have presented a variety of new results on the one-way and two-way communication complexity for algebraic problems. We have used, in several occasions, the results of [2], but our results are often stronger because they exploit the algebraic structure of the problem.

There are several directions for further research on the subject. One direction concerns the derivation of lower bounds on two-way communication complexity that involve information other than the second order derivatives. (One such result can be found in [13].) Another direction concerns two-way protocols for computing a collection $\{f_1, \dots, f_s\}$ of functions, with $s > 1$. Here, even if one assumes that the functions f_i are quadratic, the evaluation of the communication complexity is surprisingly hard and leads to problems with a combinatorial flavor. (Some partial results can be found in [12].) A final direction concerns "multi-party" protocols in which more than two processors are involved. There is very little literature on this subject [8] and it is not completely clear what are the interesting problems in this area.

ACKNOWLEDGMENTS. We are indebted to Professors Steve Kleiman and Michael Artin and Mr. Siye Wu of MIT for several stimulating discussions. We also wish to thank Professor Peter Olver of University of Minnesota for suggesting the reference [9]. Thanks are also due to an anonymous referee for a careful reading of the manuscript.

REFERENCES

1. ABELSON, H. Towards a theory of local and global computation. *Theoret. Comput. Sci. 6* (1978), 41–67.
2. ABELSON, H. Lower bounds on information transfer in distributed computations. *J. ACM 27*, 2 (1980), 384–392.
3. AHO, A. V., ULLMAN, J. D., AND YANNAKAKIS, M. On notions of information transfer in VLSI circuits. In *Proceedings of the 15th Annual Symposium on Theory of Computing* (Boston, Mass., Apr. 25–27). ACM, New York, 1983, pp. 133–139.
4. ATIYAH, M., AND MACDONALD, I. *Introduction to Commutative Algebra*. Addison-Wesley, Reading, Pa., 1969.
5. BERTSEKAS, D. P., AND TSITSIKLIS, J. N. *Parallel and Distributed Computation. Numerical Methods*. Prentice-Hall, Englewood Cliffs, N.J., 1989.
6. BLUM, L., SHUB, M., AND SMALE, S. On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *Bull. AMS 21*, 1 (1989), 1–47.
7. BORODIN, A., AND MUNRO, I. *The Computational Complexity of Algebraic and Numeric Problems*. American Elsevier, New York, 1975.
8. CHANDRA, A. K., FURST, M. L., AND LIPTON, R. J. Multi-party protocols. In *Proceedings of the 15th Annual Symposium on Theory of Computing* (Boston, Mass., Apr. 25–27). ACM, New York, 1983, pp. 94–99.
9. GORDAN, P., AND NÖETHER, M. Ueber die Algebraischen Formen, deren Hesse'sche Determinante Identisch Verschwindet. *Math. Ann. 10* (1876), 547–568.
10. HARTSHORNE, R. *Algebraic Geometry*. Springer-Verlag, New York, 1977.

11. KLEIMAN, S. L. Tangency and dualıty. In *Proceedings of the CMS Summer Instıtute in Algebraic Geometry* (Vancouver, B.C., Canada), 1984.
12. LUO, Z. Q. Communication complexity of some problems in distributed computation. Ph.D. dissertation. Operations Research Center, Tech. Rep. LIDS-TH-1909. Laboratory for Information and Decision Systems. MIT, Cambridge, Mass., 1989.
13. LUO, Z. Q., AND TSITSIKLIS, J. N. On the communication complexity of solving a polynomial equation. *SIAM J. Comput. 20*, 5 (1991), 936–950.
14. MEHLHORN, K., AND SCHMIDT, E. M. Las Vegas is better than determinism in VLSI and distributed computing. In *Proceedings of the 14th Symposium on Theory of Computing* (San Francisco, Calif., May 5–7). ACM, New York, 1982, pp. 330–337.
15. PANG, K. F., AND EL GAMAL, A. Communication complexity of computing the hamming distance. *SIAM J. Comput. 15*, 4 (1986), 932–947.
16. PAPADIMITRIOU, C. H., AND SIPSER, M. Communication complexity. In *Proceedings of the 14th Symposium on Theory of Computıng* (San Francisco, Calif., May 5–7). ACM, New York, 1982, pp. 196–200.
17. PAPADIMITRIOU, C. H., AND TSITSIKLIS, J. N. On the complexity of designing distributed protocols. *Inf. Cont. 53*, 3 (1982), pp. 211–218.
18. TENNEY, R. R., AND SANDELL, N. R., JR. Detection with distributed sensors. *IEEE Trans. Aerospace Electronic Syst. AES-17*, 4 (1981), pp. 501–510.
19. TSITSIKLIS, J. N., AND LUO, Z. Q. Communication complexity of convex optimization. *J. Complex. 3* (1987), 231–243.
20. ULLMAN, J. D. *Computational Aspects of VLSI.* Computer Science Press, Rockville, Md., 1984.
21. VAN DER WAERDEN, B. L. *Modern Algebra*, Vol. 1 & 2. Ungar, New York, 1953.
22. WILLSKY, A. S., BELLO, M. G., CASTANON, D. A., LEVY, B. C., AND VERGHESE, G. C. Combining and updating of local estimates and regional maps along sets of one-dimensional tracks. *IEEE Trans. Autom. Cont. AC-27*, 4 (1982), 799–812.
23. YAO, A. C. Some complexıty questions related to distributed computing. In *Proceedings of the 11th Symposium on Theory of Computing* (Atlanta, Ga., Apr. 30–May 2). ACM, New York, 1979, pp. 209–213.
24. ZAK, F. Projection of algebraic varieties. *Math. U.S.S.R. Sbornık 44* (1983), 535–554.
25. ZARISKI, O., AND SAMUEL, P. *Commutative Algebra*, vol. 1. Van Nostrand, New Jersey, 1965.

RECEIVED MARCH 1989; REVISED AUGUST 1992; ACCEPTED FEBRUARY 1992