Continuous-Model Communication Complexity with Application in Distributed Resource Allocation in Wireless Ad hoc Networks

Husheng Li\textsuperscript{1} and Huaiyu Dai\textsuperscript{2}

\textsuperscript{1}Department of Electrical Engineering and Computer Science, the University of Tennessee (husheng@eecs.utk.edu)
\textsuperscript{2}Department of Electrical and Computer Engineering, North Carolina State University (hdai@ece.ncsu.edu)

Abstract

Distributed resource allocation is an important problem in wireless ad hoc networks, in which there is no centralized scheduler and the resource allocation is carried out in a distributed way. Information exchange in the distributed resource allocation incurs overhead since it does not convey data information. The communication complexity, defined as the minimum number of exchanged messages needed for computing a common function with distributed inputs, is studied and the resource allocation is considered to be the procedure of computing a common function whose inputs are the parameters of multiple communication links. A lower bound for the communication complexity is provided based on the first order differentiation of the output function of resource allocation by extending the two-input-single-output case in [9] to multi-input-multi-output case. The conclusion is then applied to a concrete example of distributed resource allocation.

I. INTRODUCTION

In recent years, distributed resource allocation is receiving intensive study in the community of wireless ad hoc networks [4] [5] [8] [13] [15]. Due to the nature of broadcast in wireless communications, substantial interference may be incurred across different communication links. Therefore, resource allocation is needed to arrange the communications. Quite often, there is no centralized scheduler, e.g. a base station, in wireless ad hoc networks. Therefore, the resource allocation has to be carried out in a distributed way, i.e. the parameters for resource allocation are located at different communication links and have to be exchanged for accomplishing the resource allocation. However, the information exchange for distributed resource allocation incurs overhead since it does not convey data information. Therefore, a complicated resource allocation algorithm may not be beneficial if it requires much information exchange, although it may yield good performance for the data communication stage. Researchers have paid considerable attention to the complexity of distributed resource allocation and proposed plenty of low-complexity distributed resource allocation algorithms [4] [5] [8]. However, research on analyzing the complexity of distributed resource allocation is still lacking.

For analyzing the complexity of distributed computing, a powerful theoretical tool called \textit{communication complexity}, defined as the minimum amount of exchanged information needed for computing a common function of multiple nodes in a distributed way. The concept of communication complexity was proposed by A. C. Yao [16] for discrete mapping, and H. Abelson in [1] for computing continuous functions. For discrete case, multiple tools, e.g. monochromatic rectangle and discrepancy, have been proposed for analyzing the lower bound of communication complexity [6]. For continuous case, a lower bound for the communication complexity is proposed based on the second order differentiation in [1]. Relaxed to the family of first order continuously differentiable functions, a novel lower bound for the communication complexity is provided in [9]. The communication complexity for distributed convex optimization is analyzed in [14]. Similar communication issues have been analyzed for resource allocation in economics in [2] [11].

In this paper, we consider the distributed resource allocation as computing a multi-input-multi-output function. For example, in a distributed bandwidth allocation in orthogonal frequency division multiplexing (OFDM) networks, the inputs are the parameters of different links (e.g. deadlines of data packets and
channel conditions) and the outputs are the bandwidth allocation of all active links. For simplicity, we assume that the inputs and outputs in the resource allocation are both real numbers and do not consider quantization errors. Although it requires infinite channel capacity to transmit real numbers, we can use high-precision quantization to approximate the real number. Therefore, our analysis can be applied to situations where the channel quality is sufficiently good and cost is dominated by the rounds of communications. For other contexts, our analysis also provides insights for compressing the procedure of resource allocation. We follow the framework in [9] and extend the lower bound of communication complexity to multi-input-multi-output case (only two nodes and one function output are considered in [9]). Note that some study for the case of discrete inputs and outputs has been provided in [7].

The remainder of this paper is organized as follows: the system model and main result, Prop. 2.1, are introduced in Section II; the proof of Prop. 2.1 is given in Section III and then Prop. 2.1 is applied in an example of distributed resource allocation in Section IV; conclusions and some open problems are provided in Section V.

The following mathematical notation is used in the remainder of this paper:

- for a set \( A = \{a_1, ..., a_n\} \), then we define \( A_{/i} = \{a_1, ..., a_{i-1}, a_{i+1}, ..., a_n\} \); \(|A|\) denotes the cardinality of set \( A \);
- for a differentiable function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \), \( \nabla_x f(x) = \left( \frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n} \right)^T \);
- for a vector-valued function \( f \), the \( m \)-th element of \( f \) is denoted by \( f_m \);
- for a family of vectors \( x_{j \in J} \), \( \text{span} \left( x_{j \in J} \right) \) means the subspace spanned by these vectors.

II. SYSTEM MODEL AND MAIN RESULT

A. System Model

We consider the general case of distributed resource allocation, which can be abstracted into a distributed computing problem. We assume that there exist \( N \) nodes in a computing network (in this section, network means the abstract network for computing a common function, instead of the concrete wireless communication network), each of which has a real vector, denoted by \( x_n \in D_n \subset \mathbb{R}^{d_n} \) for node \( n \), where \( d_n \) is the dimension of \( x_n \) and \( D_n \) is an open subset. We also assume that each node knows only its own vector. The \( N \) nodes collaborate to compute a continuously differentiable \( M \)-dimensional function:

\[
f(x) : \prod_{n=1}^N D_n \rightarrow \mathbb{R}^M, \quad (1)
\]

where \( x = \{x_1, ..., x_N\} \).

Mapping the abstract model to a concrete distributed resource allocation scheme in wireless networks, we can consider the variable \( x_n \) to be the parameters of each transmitter (e.g. the urgency of packet, current channel condition or history of resource allocation), which are unknown to other transmitters, and the function output to be resource allocation for each transmitter (e.g. power and bandwidth).

B. Communication Complexity

We assume that the \( N \) nodes broadcast serially, one broadcast per time slot and one real number per broadcast (we allow a node to broadcast in multiple consecutive time slots). The order of broadcast is predetermined and independent of \( x \). Each broadcast can be decoded by all other nodes, i.e. the network topology is a complete graph (it is very useful and difficult to study a general topology). We denote the \( i \)-th broadcast by \( m_i(x) \), which is a function of \( x \). Each broadcast is a function of broadcast history. Therefore, we can rewrite \( m_i(x) \) as (suppose \( i \in B_n \))

\[
m_i(x) = \hat{m}_i(x_n, m_1(x), ..., m_{i-1}(x)). \quad (2)
\]

We place the following assumptions on the broadcast protocol:
• there are totally \( r \) broadcasts; we denote by \( B_n \) the set of time slots used by node \( n \), i.e. node \( n \) broadcasts at time slot \( i \) if and only if \( i \in B_n \);
• every node broadcasts at least once, i.e. its private vector is always needed for computing the function.

Then, there exist a family of functions \( \{ h_n \} \) such that

\[
f(x) = h_n (x_n, m_1(x), ..., m_r(x)),
\]

since every node knows \( f(x) \) after \( r \) broadcasts. We define the corresponding communication complexity, denoted by \( C \left( f; \prod_{n=1}^{N} D_n \right) \), as the minimum \( r \) over all possible protocols computing \( f(x) \) and satisfying the above assumptions. For the scenario of wireless networks, the communication complexity represents the overhead for resource allocation transmissions. Note that we do not consider quantization error, which facilitates the analysis and can be justified by using sufficient precision for quantization in each broadcast.

The discussion of optimal quantization and the impact of quantization error on data communications are beyond the scope of this paper.

C. Main Result

Similar to [9], the following assumptions are placed for the distributed computing problem:
• for any \( 1 \leq n \leq N \), we define

\[
S_n(x_n) \triangleq \left\{ S \subset \prod_{k=1, k \neq n}^{N} D_k \left| f(S, x_n) \text{contains an open set} \right. \right\};
\]

• for any \( 1 \leq n \leq N \) and any nonempty open set \( S \) in \( \prod_{k=1, k \neq n}^{N} D_k \), we have \( S \subset S_n(x_n) \);
• for any \( 1 \leq n \leq N \) and some \( k_n \in \mathbb{N} \), we have

\[
\dim \left[ \text{span} \left( \nabla_{x_n} f_l(x) | 1 \leq l \leq M, x_n \in S \right) \right] \geq k_n, \quad \forall x_n \in \prod_{k=1, k \neq n}^{N} D_k, S \in S_n(x_n). \quad (3)
\]

Then, the main result of this paper is given in the following proposition, which provides a lower bound for the communication complexity.

**Proposition 2.1:** The following inequality holds:

\[
C \left( f; \prod_{n=1}^{N} D_n \right) > \min_{1 \leq n \leq N} k_n. \quad (4)
\]

Note that, when \( N = 2 \), Prop. 2.1 is equivalent to the conclusion (Theorem 2.1) in [9].

It is easy to verify the following Corollary:

**Corollary 2.2:** If \( 1 \leq n \leq N, d_n = 1 \) and \( k_n = N - 1 \), then we have

\[
C \left( f; \prod_{n=1}^{N} D_n \right) = N, \quad (5)
\]

and the optimal algorithm is to let all nodes broadcast all of their parameters and compute the function according to the broadcasts.
III. PROOF OF PROP. 2.1

The proof of Prop. 2.1 follows the framework of that in [9]. Throughout the paper, we assume that

$$C \left( f; \prod_{n=1}^{N} D_n \right) = \min_{\{D_n\}} C \left( f; \prod_{n=1}^{N} \tilde{D}_n \right),$$  \hspace{1cm} (6)

where \{\tilde{D}_n\} are open subsets and satisfy \( \tilde{D}_n \subset D_n \), \( \forall 1 \leq n \leq N \). The justification of the assumption is the same as that in [9].

Now, we fix a protocol having \( r + M \) broadcasts and denote by \( \nabla m(x) \) the \( \sum_{n=1}^{N} d_n \times r \) matrix (recall that \( d_n \) denotes the dimension of \( x_n \)), in which the \( i \)-th column vector is the gradient vector of scalar function \( m_i \) with respect to the elements in \( x \). Then, we can prove the following lemma in the same way as that in [9]:

**Lemma 3.1:** The following equation holds:

$$\max_{x \in \prod_{n=1}^{N} D_n} \text{rank} \left[ \nabla m(x) \right] = r.$$  \hspace{1cm} (7)

Then, we define \{\tilde{D}_n\} as a family of open sets such that \( \tilde{D}_n \subset D_n \), \( \forall 1 \leq n \leq N \), and \( \nabla m(x) \) is of full rank in \( \tilde{D} \triangleq \prod_{n=1}^{N} \tilde{D}_n \). We also define \( r_n = \| B_n \| \) (recall that \( B_n \) denotes the set of time slots for node \( n \)), i.e. the number of time slots allocated to node \( n \). For a vector \( c = (c_1, ..., c_r) \in \mathbb{R}^r \), we denote by \( c^i = (c_1, ..., c_i) \), \( \forall 1 \leq i \leq r \). Then, we can prove the following lemma:

**Lemma 3.2:** For \( x \in \prod_{n=1}^{N} D_n \), we have

$$\text{rank} \left[ \text{span} \left( \nabla_{x/n} \hat{m}_i (x_n, c^i) | i \in B_n \right) \right] = r - r_n,$$  \hspace{1cm} (8)

where \( c = (m_1(x), ..., m_r(x)) \).

**Proof:** We can rewrite \( \nabla m(x) \) as

$$\nabla m(x) = \begin{pmatrix}
A_{11} & A_{12} & \cdots & A_{1N} \\
A_{21} & A_{22} & \cdots & A_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
A_{N1} & A_{N2} & \cdots & A_{NN}
\end{pmatrix},$$  \hspace{1cm} (9)

where the submatrices are defined as

$$A_{ij} = (\nabla_{x} m_k(x)| k \in B_j), \hspace{1cm} 1 \leq i, j \leq N.$$  \hspace{1cm} (10)

For any \( 1 \leq n \leq N \) and \( i \in B_n \), the message \( m_i \) is a function of \( x_n \) and \{\( m_l \)\}_{1 \leq l < i, l \notin B_n} \ and therefore can be written as

$$m_i(x) = M_i \left( x_n, \{m_l(x)\}_{1 \leq l < i, l \notin B_n} \right),$$  \hspace{1cm} (11)

where \( M_i \) is a function mapping from \( \mathbb{R}^{l-1+d_i} \) to \( \mathbb{R} \).

Considering the corresponding gradient vector, we have

$$\nabla_{x/n} m_i(x) = \sum_{l \notin B_n} \frac{\partial M_i}{\partial m_l} \bigg|_{x} \nabla_{x/n} m_l(x),$$  \hspace{1cm} (12)

which implies that

$$\nabla_{x/n} m_i(x) \in \text{span} \left( \nabla_{x/n} m_l(x) | l \notin B_n \right).$$  \hspace{1cm} (13)
Therefore, for every $1 \leq n, m \leq N$ and $m \neq n$, all column vectors in submatrix $A_{/n,m}$, which is defined as

$$A_{/n,m} = \begin{pmatrix} A_{1,m} \\ \vdots \\ A_{n-1,m} \\ A_{n,m} \end{pmatrix},$$

(14)

can be linearly represented by the column vectors in $\{ A_{/n,k} \}_{k \neq m}$ since $\nabla_{x/m} m_i(x)$ is a column vector in $A_{/n,m}$. Then, fixing $m = n + 1$ when $n < N$ and $m = 1$ when $n = N$, we have

$$\text{rank } (A_{/n,1}, \ldots, A_{/n,n}, A_{/n,n+2}, \ldots, A_{/n,N})$$
$$= \text{rank } (A_{/n,1}, \ldots, A_{/n,n}, A_{/n,n+1}, \ldots, A_{/n,N}),$$

(15)

where the first matrix ignores $A_{/n,n+1}$ while the second matrix includes $A_{/n,n+1}$ and the last inequality is because the first matrix has only $r - r_{n+1}$ columns.

Then, we have

$$(N - 1)r = \sum_{n=1}^{N} \left( r - r_n \right)$$
$$\geq \sum_{n=1}^{N} \text{rank } (A_{/n,1}, \ldots, A_{/n,n}, A_{/n,n+2}, \ldots, A_{/n,N})$$
$$= \sum_{n=1}^{N} \text{rank } (A_{/n,1}, \ldots, A_{/n,n}, A_{/n,n+1}, \ldots, A_{/n,N}),$$

(16)

where the last inequality is proved as follows: due to Lemma 3.1, we have

$$\text{rank } \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} \end{pmatrix} = r.$$ 

(17)

Then, we can always find $r$ row vectors to form a linearly independent set. Each of these row vectors is also within a maximum linear independent set in matrix

$$(A_{/n,1}, \ldots, A_{/n,n}, A_{/n,n+1}, \ldots, A_{/n,N})$$

if it appears in the matrix. Then the summation of ranks in (16) is larger than or equal to the number of these row vectors appearing in the corresponding matrices. It is easy to check that each row vector is counted for $N - 1$ times since it disappears only once. Therefore, the sum is larger than or equal to $(N - 1)r$. This concludes the proof.

For a vector $c \in \mathbb{R}^r$, we define the set of all $x$ leading to the broadcast sequence $c$:

$$S(c) = \left\{ x \in \prod_{n=1}^{N} D_n \left| m_i(x) = c_i, i = 1, \ldots, r \right. \right\},$$

(18)
and the set of \( x_n \) leading to \( c \):

\[
S_n(c) = \left\{ x_n \in \prod_{n=1}^{N} D_n \, \mid \, \hat{m}_i(x_n, c^{i-1}) = c_i, \forall i \in B_n \right\},
\]

(19)

and the set of all possible broadcast sequences (we consider only the first \( r \) broadcasts):

\[
R = \left\{ (m_1(x), ..., m_r(x)) \, \mid \, \forall x \in \prod_{n=1}^{N} D_n \right\}.
\]

(20)

Then, we have the following lemma, whose proof is the same as Lemma 2.3 in [9].

**Lemma 3.3:** For any vector \( c \in R \), we have

\[
S(c) = \prod_{n=1}^{N} S_n(c).
\]

(21)

Next, we fix an \( x^* \in D \), let \( c^* = m(x^*) \) and, for \( i \in B_n \), define

\[
F_i(x_n, c) = m_i(x_n, c^{i-1}) - c_i, \quad c \in \mathbb{R}, \ x_n \in D_n.
\]

(22)

Obviously, we have

\[
F_i(x^*_n, c^*) = 0,
\]

(23)

when \( i \in B_n \).

Following the same argument as that in [9], we conclude that there exists an open set \( U \in \mathbb{R}^r \) such that \( c^* \in U \) and an open subset \( \hat{D}_n \subset \mathbb{R}^{d_n} \) containing \( x^*_n \), for \( 1 \leq n \leq N \), such that \( \hat{S}_n(c) = S_n(c) \cap \hat{D}_n \) is nonempty and connected, \( \forall c \in U \).

In the remainder of the proof, suppose that \( r \in B_n \), i.e. the \( r \)-th broadcast is sent out by node \( n \). Since all other nodes know the result \( f(x) \), there exist \( \{\hat{h}_l^j\}_{j \neq n} \) such that

\[
f_l(x) = h_l^j(x_j, c^r), \quad \forall j \neq n, m = 1, ..., M.
\]

(24)

Following the same argument as that in [9], we obtain that \( \{\hat{h}_l^j\}_{j \neq n} \) are all continuous functions and there exists some \( \hat{c} \in U \) such that not all \( h_j(x_i^l, \hat{c}) \) are constant functions of \( x_j \) on \( \hat{S}_j(\hat{c}) \) (different from the two-node case in [9], not all \( h_l^j(x_j, \hat{c}) \) \( (j \neq n) \) are necessarily nonconstant functions of \( x_j \) in \( \hat{S}_j(\hat{c}) \) since some nodes may know the result before the \( r \)-th broadcast).

Now, we fix \( x_j \in \hat{S}_j(\hat{c}), j = 1, ..., n-1, n+1, ..., N \). According to (3), we can find \( x^1_n, ..., x^{k_n}_n \) (all in \( \hat{S}_n(\hat{c}) \)) and \( l^1, ..., l^{k_n} \) such that \( (x^i_l, m^l) \neq (x^j_l, m^j) \) if \( i \neq j \), and vectors in set

\[
I = \{ \nabla_{x/n} f_l(x_n, \{x_j\}_{j \neq n}), ..., \nabla_{x/n} f_{l^{k_n}}(x_{k_n}, \{x_j\}_{j \neq n}) \}
\]

are linearly independent. We also have that, for the fixed \( x_n \in \hat{S}_n(\hat{c}), f_l(x) = h_l^j(x_n, \hat{c}), l = 1, ..., M \), are constant functions of \( \{x_j\}_{j \neq n} \) on \( \prod_{j=1,j \neq n}^{N} \hat{S}_j(\hat{c}) \). According to (26), it is easy to check that, \( \forall i \in B_n \),

\[
\left\{ x_n \, \mid \, F_i(\{x_n, x_n\}, \hat{c}) = 0, x_n \in \prod_{j=1,j \neq n}^{N} \hat{D}_j \right\} = \prod_{j=1,j \neq n}^{N} \hat{S}_j(\hat{c}).
\]

(25)

Then applying Theorem A.4 in [9], we have

\[
\nabla_{x/n} f_l(x) \in \text{span} \left( \nabla_{x/n} \hat{m}_i(x_n, c^{i-1}) \mid i \in B_n \right), \quad l = 1, ..., M.
\]

(26)

Then, all vectors in set \( I \) in (25) are within \( \text{span} \left( \nabla_{x/n} \hat{m}_i(x_n, c^{i-1}) \mid i \in B_n \right) \). From Lemma 3.2, we have

\[
\text{rank} \left[ \text{span} \left( \nabla_{x/n} \hat{m}_i(x_n, c^{i-1}) \mid i \in B_n \right) \right] = r - r_n.
\]

(27)
Then, we have (recall that \( r_n > 0 \), since every node broadcasts at least once)
\[
r \geq r - r_n > k_n \geq \min_{1 \leq j \leq N} k_j.
\]
(28)

This concludes the proof.

IV. APPLICATION IN DISTRIBUTED BANDWIDTH RESOURCE ALLOCATION

As a simple application of Prop. 2.1, we consider the following problem of bandwidth resource allocation: suppose that there are \( N \) communication links sharing a total bandwidth of \( W \). We denote by \( W_n, P_n \) and \( g_n \) the bandwidth, transmit power and channel gain of link \( n \) (suppose that frequency bands allocated to different links are orthogonal, thus inducing no inter-user interference). We assume that the resource allocation targets at maximizing the sum of utilities, as functions of the channel capacities of different links. Then, the distributed resource allocation is formulated into an optimization problem:

\[
\max \left\{ U_n(R_n) \right\}_{n=1,...,N} \quad \text{s.t.} \quad R_n = W_n \log \left( 1 + \frac{P_n g_n}{W_n N_0} \right),
\]
\[
\sum_{n=1}^N W_n = W,
\]
(29)

where \( g_n \) is channel power gain for link \( n \), \( U_n \) is the utility function of link \( n \) and \( N_0 \) is the noise power spectral density (PSD). Since the channel gain \( g_n \) of link \( n \) is unknown to other links, the optimization has to be carried out in a distributed way. For applying Prop. 2.1, we consider the optimal bandwidth allocation as an \( N \)-dimensional function of distributed parameters \( \gamma \triangleq \left( \frac{g_1 P_1}{N_0}, \ldots, \frac{g_N P_N}{N_0} \right) \), \( n = 1, \ldots, N \), denoted by \( f(\gamma) \). Since there are only \( N - 1 \) degrees of freedom in selecting \( \{W_n\}_{n=1,\ldots,N} \), we consider only \( \{W_n\}_{n=1,\ldots,N-1} \) as inputs to function \( f \). We assume that the utility functions \( \{U_n\}_{n=1,\ldots,N} \) are known to all users. It is easy to extend our results to the case in which each link has finite private parameters of its utility function. Note that the case in which even the forms of utility functions are private is beyond the discussion of this paper.

We place the following assumptions for the bandwidth allocation problem:

- \( U_n \) is a strictly convex and strictly increasing function of \( R_n \) and is second order continuously differentiable;
- \( U_n(0) = -\infty \).

Based on the above assumptions, it is easy to verify the following lemma:

Lemma 4.1: The optimization problem in (29) is a convex optimization; the optimal solution is an interior point within the feasible region, which is uniquely determined by the following \( N - 1 \) equations (note that \( U_N \) is also a function of \( W_n, n < N \), since \( W_N = W - W_1 - \ldots - W_{N-1} \)):

\[
F_n \triangleq \frac{\partial U_n}{\partial W_n} + \frac{\partial U_N}{\partial W_n} = 0, \quad n = 1, \ldots, N - 1.
\]
(30)

We assume that the following matrix

\[
\begin{pmatrix}
\frac{\partial F_1}{\partial W_1} & \frac{\partial F_1}{\partial W_2} & \cdots & \frac{\partial F_1}{\partial W_{N-1}} \\
\frac{\partial F_2}{\partial W_1} & \frac{\partial F_2}{\partial W_2} & \cdots & \frac{\partial F_2}{\partial W_{N-1}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial F_{N-1}}{\partial W_1} & \frac{\partial F_{N-1}}{\partial W_2} & \cdots & \frac{\partial F_{N-1}}{\partial W_{N-1}}
\end{pmatrix}
\]
(31)
which is of rank $N$ elements. Then, we need the following simple lemma:

$$
\left( \frac{\partial F_i}{\partial (\gamma_n)_k} \right)_{k} = - \left( \frac{\partial F_i}{\partial W_1} \right)_{n=1,...,N-1} \cdot \left( \frac{\partial F_i}{\partial W_N} \right)_{n=1,...,N-1} \cdot \left( \frac{\partial F_i}{\partial (\gamma_n)_k} \right)_{n=1,...,N-1}, \quad \forall n = 1,...,N,
$$

from which we can obtain the partial derivatives of $\{W_n\}_{n=1,...,N-1}$ with respect to $\{\gamma_n\}_{n=1,...,N}$. Fixing an arbitrary $1 \leq n^* < N$, we have

$$
\left( \frac{\partial F_i}{\partial \gamma_n} \right)_{n \neq n^*} = - \left( \frac{\partial F_i}{\partial W_n} \right)_{n=1,...,N-1} \cdot \left( \frac{\partial F_i}{\partial (\gamma_n)_k} \right)_{n=1,...,N-1}, \quad \forall n = 1,...,N-1.
$$

It is easy to check

$$
\left( \frac{\partial F_i}{\partial \gamma_n} \right)_{n \neq n^*} = \left( \begin{array}{cccc}
\frac{\partial F_1}{\partial \gamma_1} & 0 & \cdots & 0 \\
0 & \frac{\partial F_2}{\partial \gamma_2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{\partial F_{N-1}}{\partial \gamma_{N-1}}
\end{array} \right),
$$

which is of rank $N-1$ (recall that $\frac{\partial F_i}{\partial \gamma_{n^*}}$ is not included). Then, we obtain the following lemma:

**Lemma 4.2:** If the matrix $\left( \frac{\partial F_i}{\partial W_n} \right)_{i=1,...,N-1, n=1,...,N-1}$ is nonsingular at any point, the communication complexity of the distributed resource allocation in (29) is $N$.

Now, the communication complexity is determined by the singularity of matrix $\left( \frac{\partial F_i}{\partial W_n} \right)_{i=1,...,N-1, n=1,...,N-1}$. We notice that $U_n$ is explicitly dependent on only $W_n$ when $n \neq N$, which implies

$$
\frac{\partial^2 U_i}{\partial W_n^2} = 0, \quad \text{if } i \neq n.
$$

Then, the matrix $\left( \frac{\partial F_i}{\partial W_n} \right)_{i=1,...,N-1, n=1,...,N-1}$ can be rewritten as

$$
\left( \frac{\partial F_i}{\partial W_n} \right)_{i=1,...,N-1, n=1,...,N-1} = \left( \begin{array}{cccc}
\frac{\partial^2 U_1}{\partial W_1^2} & 0 & \cdots & 0 \\
0 & \frac{\partial^2 U_2}{\partial W_2^2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{\partial^2 U_{N-1}}{\partial W_{N-1}^2}
\end{array} \right) + \left( \begin{array}{cccc}
\frac{\partial^2 U_1}{\partial W_1 \partial W_N} & \frac{\partial^2 U_2}{\partial W_2 \partial W_N} & \cdots & \frac{\partial^2 U_{N-1}}{\partial W_{N-1} \partial W_N}
\end{array} \right),
$$

Noticing $W_N = W - W_1 - ... - W_{N-1}$, we have

$$
\frac{\partial U_N}{\partial W_n} = - \frac{\partial U_N}{\partial W_N}, \quad n \neq N,
$$

and

$$
\frac{\partial^2 U_N}{\partial W_n^2} = \frac{\partial^2 U_N}{\partial W_N^2}, \quad n \neq N.
$$

Therefore, all terms in the second matrix in (36) are identical and positive.

According to (36), the matrix $\left( \frac{\partial F_i}{\partial W_n} \right)_{i=1,...,N-1, n=1,...,N-1}$ is the sum of a diagonal matrix (all the diagonal elements are positive due to the strict convexity of utility functions) and a matrix having identical positive elements. Then, we need the following simple lemma:
Lemma 4.3: Suppose that a matrix has the following expression:

\[
\begin{pmatrix}
 a_1 + b & b & \cdots & b \\
 b & a_2 + b & \cdots & b \\
 \vdots & \vdots & \ddots & \vdots \\
 b & b & \cdots & a_N + b
\end{pmatrix},
\]

(39)

where \(a_n > 0\) and \(b > 0\), \(\forall 1 \leq n \leq N\). Then, the matrix is of full rank.

Proof: Subtracting the \(n\)-th column from the \(n-1\)-th column \((n = 2, \ldots, N)\), we obtain that the rank of matrix (39) is the same as that of the following matrix.

\[
M_N \triangleq \begin{pmatrix}
 a_1 & 0 & \cdots & 0 & b \\
 -a_2 & a_2 & \cdots & 0 & b \\
 0 & -a_3 & \cdots & 0 & b \\
 \vdots & \vdots & \ddots & \vdots & \vdots \\
 0 & 0 & \cdots & a_{N-1} & b \\
 0 & 0 & \cdots & -a_N & a_N + b
\end{pmatrix}.
\]

(40)

We do induction on \(N\) for the determinant of the above matrix. When \(N = 1\), \(\det(M_1) = a_1 > 0\). We assume that \(\det(M_N) > 0\), when \(N = k\). Now, we consider the case \(N = k + 1\). For proving \(\det(M_{k+1}) > 0\), we expand the determinant through the first row, which yields

\[
\det(M_{k+1}) = a_1 \times \det
\begin{pmatrix}
 a_2 & \cdots & 0 & b \\
 -a_3 & \cdots & 0 & b \\
 \vdots & \vdots & \ddots & \vdots \\
 0 & \cdots & a_k & b \\
 0 & \cdots & -a_{k+1} & a_{k+1} + b
\end{pmatrix}
\] + \((-1)^{N+1}b \times \det
\begin{pmatrix}
 -a_2 & a_2 & 0 & \cdots & 0 \\
 0 & -a_3 & a_3 & \cdots & 0 \\
 0 & 0 & -a_4 & \cdots & 0 \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & 0 & \cdots & a_k \\
 0 & 0 & 0 & \cdots & -a_{k+1}
\end{pmatrix}.
\]

(41)

Due to the induction assumption, the first term in the above equation is positive (recall that \(a_1 > 0\)). The matrix in the second term has the same determinant as the following matrix

\[
\begin{pmatrix}
 -a_2 & 0 & 0 & \cdots & 0 \\
 0 & -a_3 & 0 & \cdots & 0 \\
 0 & 0 & -a_4 & \cdots & 0 \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & 0 & \cdots & -a_{k+1}
\end{pmatrix},
\]

(42)

whose determinant is \((-1)^{N-1} \prod_{n=2}^{N} a_n\). Therefore, the sign of the second term in (41) is \(\prod_{n=2}^{N} a_n\), thus also being positive. Then, the determinants of all \(M_N\) are positive. This concludes the proof.

Summarizing the above lemmas, we obtain the following proposition:

Proposition 4.4: The communication complexity of the distributed resource allocation problem in (29) is \(N\).

Remark 4.5: We notice that the expression of channel capacity is not used in the proof. Therefore, the conclusion also applies to any expression of transmission rate that is a strictly convex function of allocated bandwidth.
V. Conclusions and Open Problems

In this paper, we have considered the communication complexity of distributed resource allocation in wireless ad hoc networks. We have proposed a lower bound for the communication complexity of computing a common function by extending the argument for two-input-single-output case in [9] to multi-input-multi-output case. We have also shown how to apply the main result in Prop. 2.1 to an example of utility based distributed resource allocation and obtained that the corresponding communication complexity is the same as the number of communication links.

The following problems are still open:

- the impact of turn-around time, i.e. the time needed for converting from transmitting to receiving (or the reverse);
- communication complexity measured by bits when quantization is considered, as well as the tradeoff between performance of data communication and information exchange overhead;
- the communication complexity of resource allocation with continuous inputs and binary outputs, e.g. the exponential rule in [12].

REFERENCES