Distributed Approach for Joint Symbol and Channel Estimation in Heterogeneous Cellular Networks

Igor F. S. de Sousa, André L. F. de Almeida and Tarcisio F. Maciel

***Abstract*— This work considers the uplink of a multiuser heterogeneous cellular network based on direct-sequence code division multiple access (DS-CDMA). Two cases are considered: i) a *centralized case*, where a base-station (BS) gathers the information of all users; and ii) a *distributed case*, where micro-BS nodes perform the signal processing in a distributed fashion. A distributed approach for data estimation and detection is studied in this paper for the uplink of this heterogeneous network which employs spreading codes at the transmitters and consensus-based averaging at the micro-BS nodes. We compare the distributed and centralized approaches considering, in both cases, a receiver structure based on a joint trilinear tensor model for channel gains, spreading codes and transmitted symbols of all users. Numerical results show that the performance of the distributed approach is superior to that of the centralized one.**

***Keywords*— Distributed estimation, PARAFAC, DS-CDMA, multiuser communication, consensus averaging.**

I. Introduction

The interest in autonomous systems has grown in several areas such as military communications and surveillance. In the last years, the deployment of Wireless Sensors Networks (WSNs) for collaborative monitoring, information processing, and control has drawn considerable attention. Specifically, WSNs can operate autonomously, i.e. without a data-fusion center collecting and processing all measurements, thus exhibiting desirable properties such as robustness against node failure [1]. As there is no central controller, the cooperation among nodes is necessary for estimating/detecting a common system parameter or for taking reliable decisions and, in order to maintain the coordinated action between the different nodes, local information exchange is also needed.. The concept of consensus averaging (CA) is used to achieve this cooperation.

In the downlink of a WSN, a distributed consensus problem may be useful when some network nodes are interested in a common message sent by a remote transmitter, but each node has limited hardware-capabilities thus making it unable to decode the message individually [2]. A few years ago, some researchers have studied the distributed consensus problem applied to communications systems [2]–[4]. Several theoretical studies on the problem of iterative decoding of common messages sent over broadcast channels to a pair of users have been developed, for instance, in [4]. The optimal link weights for fast convergence of CAs was designed by Xiao and Boyd in [5]. Some other works developed algorithms which take advantage of collaboration among more than two nodes [6]. The CA problem in networks with random link failure has been analyzed by some authors assuming noiseless consensus [7], [8]. Differently, Zhu *et al* [3] have considered the demodulation, detection and estimation problem using CA-like single iteration consensus averaging (CA-SI) and consensus averaging variant based in method of multipliers (CA-MoM), where they consider algorithms that are robust to both random link failure and noisy consensus.

In recent years, a few works based on cooperative wireless networks have been developed using tensor-based estimation algorithms [9], [10]. In [10] a supervised technique was proposed for two-way relaying cooperative systems, where the relay node is assumed to operate with multiple antennas. A distributed tensor-based receiver is proposed in [9] for joint channel estimation and detection in a DS-CDMA based WSN.

In this work, we present a distributed data estimation and detection approach for the uplink of a multiuser heterogeneous cellular network which employs DS-CDMA at the transmitters (users) and consensus-based averaging at the receivers (micro-BS nodes). The micro-BS nodes are supposed to cooperate to jointly recover users' transmitted signals in a distributed way, i.e., without the help of a central base-station (BS). This cooperation allows the use of a distributed Alternating Least Squares (D-ALS) algorithm that exploits the Parallel Factor (PARAFAC) model [11], [12] of the overall data gathered by the network for a joint in-network channel estimation and symbol detection. We then compare the distributed estimation approach with a centralized one, where a central base-station (BS) gathers the information of all users.

Notation and Properties: the following notation is used throughout the paper: Scalar are denoted by lower-case letters $(a, b, \dots)$, vectors and matrices are written as boldface lower-case $(\mathbf{a}, \mathbf{b}, \dots)$ and upper-case letters $(\mathbf{A}, \mathbf{B}, \dots)$ respectively. Tensors are denoted by calligraphic upper-case letter $(\mathcal{A}, \mathcal{B}, \dots)$. $\mathbf{A}_{i.}$ and $\mathbf{A}_{.j}$ denote, respectively, the i^{th} row and the j^{th} column of the $I \times J$ matrix $\mathbf{A}$. The transpose of a matrix is denote by $\mathbf{A}^T$ whereas $\mathbf{A}^H$ stands for its conjugate transpose. $\mathbf{1}$ is a column vector with all the elements equal to 1. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. The Kronecker and Hadamard product are denoted by $\otimes$ and $\odot$, respectively. Frobenius norm is denoted by $\|.\|_F$. The pseudo inverse operator is $(\cdot)^\dagger$, while $\text{diag}(\cdot)$ is the operator that forms a diagonal matrix from vector argument. $\text{vec}(\cdot)$ forms a vector by stacking the columns of its matrix argument. $\mathbf{X}_{k..}$ is the k^{th} slice of tensor $\mathcal{X}$ in K dimension, $\mathbf{X}_{.p.}$ is the p^{th} slice of tensor $\mathcal{X}$ in dimension P and $\mathbf{X}_{..n}$ is the n^{th} slice of tensor $\mathcal{X}$ in

Igor F. S. de Sousa and André L. F. de Almeida are with the Department of Teleinformatics Engineering, Federal University of Ceará, Fortaleza, Brazil (e-mails: igorflavio@gtel.ufc.br, andre@gtel.ufc.br). Tarcisio F. Maciel is with Center of Technology, Federal University of Ceará, Fortaleza, Brasil (e-mail: maciel@gtel.ufc.br). This work was partially supported by the Innovation Center, Ericsson Telecomunicações S.A., Brazil, and the CAPES."

dimension N for $\mathcal{X} \in \mathbb{C}^{K\times P\times N}$. The Katri-Rao (columnwise Kronecker) product between two matrices $\mathbf{A} \in \mathbb{C}^{K\times R}$ and $\mathbf{B} \in \mathbb{C}^{J\times R}$ is symbolized by $\mathbf{A} \diamond \mathbf{B}$ can be defined as

$$\mathbf{A} \diamond \mathbf{B} = \begin{bmatrix} \mathbf{B}\mathrm{diag}(\mathbf{A}_{1.}) \\ \vdots \\ \mathbf{B}\mathrm{diag}(\mathbf{A}_{K.}) \end{bmatrix} \in \mathbb{C}^{KJ\times R}. \tag{1}$$

For any matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{X}$, we can be rewrite the $\mathrm{vec}(\mathbf{AXB})$ as [13]

$$\mathrm{vec}(\mathbf{AXB}) = (\mathbf{B}^T \diamond \mathbf{A})\mathrm{vecd}(\mathbf{X}), \tag{2}$$

where $\mathrm{vecd}(\mathbf{X})$ indicates the vectorization which selects only the diagonal elements of matrix $\mathbf{X}$.

II. Tensor Prerequisites

We can decompose a third-order tensor $\mathcal{X} \in \mathbb{C}^{I_1\times I_2\times I_3}$ using the Parallel Factor (PARAFAC) decomposition [12], also known as Canonical Decomposition (CANDECOMP) [11], can represent $\mathcal{X}$ in scalar form as

$$x_{i_1,i_2,i_3} = \sum_{q=1}^{Q} u^{(1)}_{i_1,q} u^{(2)}_{i_2,q} u^{(3)}_{i_3,q} \tag{3}$$

where $\mathbf{U}^{(l)} = [u^{(l)}_{i_l,q}] \in \mathbb{C}^{I_l\times Q}|_{l=1,2,3}$.

It is possible to see that (3) admits a Q-component PARAFAC decomposition [12]. Therefore, the data tensor is completely characterized by three factor matrices $\mathbf{U}^{(l)}|_{l=1,2,3}$. Due to the uniqueness property of the PARAFAC model, if the decomposition is essentially unique then, the sets of matrices $\{\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}\}$ and $\{\tilde{\mathbf{U}}^{(1)}, \tilde{\mathbf{U}}^{(2)}, \tilde{\mathbf{U}}^{(3)}\}$that give rise to same tensor $\mathcal{X}$ are linked by

$$\tilde{\mathbf{U}}^{(1)} = \mathbf{U}^{(1)}\mathbf{\Pi}\mathbf{\Delta}_1, \tilde{\mathbf{U}}^{(2)} = \mathbf{U}^{(2)}\mathbf{\Pi}\mathbf{\Delta}_2, \tilde{\mathbf{U}}^{(3)} = \mathbf{U}^{(3)}\mathbf{\Pi}\mathbf{\Delta}_3, \tag{4}$$

where $\mathbf{\Delta}_1$, $\mathbf{\Delta}_2$ and $\mathbf{\Delta}_3$ are $Q \times Q$ diagonal matrices that scale/counter-scale the columns of $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$ and $\mathbf{U}^{3}$, respectively, with $\mathbf{\Delta}_1\mathbf{\Delta}_2\mathbf{\Delta}_3 = \mathbf{I}_Q$, and $\mathbf{\Pi}$ is a $Q \times Q$ permutation matrix.

A sufficient condition for a PARAFAC decomposition to provide unique parameter estimates (up to column scaling and permutation) was earlier established in [14] for the real numbers case, and later, extended by Sidiropoulos and Bro [15] for the complex numbers case. This condition, known as Kruskal condition, is given $\kappa_{\mathbf{U}^{(1)}} + \kappa_{\mathbf{U}^{(2)}} + \kappa_{\mathbf{U}^{(3)}} \geq 2(Q+1)$, where $\kappa_{\mathbf{U}^{(1)}}$, $\kappa_{\mathbf{U}^{(2)}}$ and $\kappa_{\mathbf{U}^{(3)}}$ are the Kruskal-rank of the factor matrices $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$ and $\mathbf{U}^{(3)}$. In order to ensure Kruskal condition it is necessary that the dimension of the system matrices follow $\min(I_1,Q) + \min(I_2,Q) + \min(I_3,Q) \geq 2Q+2$.

For a tensor $\mathcal{X} \in \mathbb{C}^{I_1\times I_2\times I_3}$, we can be define $\mathbf{X}_{i_1..}$ as i_1-th matrix slice in first mode of $\mathcal{X}$ given [16]

$$\mathbf{X}_{i_1..} = \mathbf{U}^{(2)}\mathrm{diag}(\mathbf{U}^{(1)}_{i_1.})\mathbf{U}^{(3)T}. \tag{5}$$

When the matrix slices (5) are stacking in the same dimension we obtain the matrix unfolding, $\mathbf{X}_1 = \left[\mathbf{X}^T_{1..}, \ldots, \mathbf{X}^T_{I_1..}\right]^T \in \mathbb{C}^{I_1 I_2\times I_3}$, in first-mode and matrix unfolding, $\mathbf{X}_2 = \left[\mathrm{vec}(\mathbf{X}^T_{1..}), \ldots, \mathrm{vec}(\mathbf{X}^T_{I_1..})\right] \in \mathbb{C}^{I_2 I_3\times I_1}$, in second-mode. Using properties (1) and (2), respectively, we have following equalities

$$\mathbf{X}_1 = \left(\mathbf{U}^{(1)} \diamond \mathbf{U}^{(2)}\right)\mathbf{U}^{(3)T}, \tag{6a}$$

$$\mathbf{X}_2 = \left(\mathbf{U}^{(2)} \diamond \mathbf{U}^{(3)}\right)\mathbf{U}^{(1)T}. \tag{6b}$$

III. System Overview and Signal Model

Consider a cell which has a single BS, Q co-channel users and B clusters of micro-BS nodes. Each cluster has K micro-BS nodes. Both user equipment (UE) and micro-BS nodes are single-antenna devices. The micro-BS nodes may serve UEs in order to improve performance, e.g., in terms of cell coverage and system capacity. Figure 1 illustrates a possible configuration of the described cell.

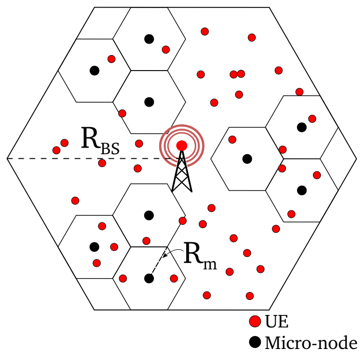


Fig. 1. Cell configuration, with R_{BS} is radius of BS cell and R_m is radius of the micro-BS node cell.

We assume that the UEs may communicate with BS in a direct manner or by using a cluster of micro-BS node. The UEs transmit their information to the micro-BS node or to the BS antenna using direct-sequence code division multiple access (DS-CDMA) through a flat Rayleigh fading channel with zero-mean additive white Gaussian noise (AWGN). Data exchanges among micro-BS nodes are considered to be error-free. The baseband signal received at each antenna (from the BS or a micro-BS node) is sampled at the chip rate and decomposed into its polyphase components. Using (3), the signal received at the k^{th} antenna, n^{th} symbol and p^{th} is given by

$$x_{k,p,n} = \sum_{q=1}^{Q} a_{k,q} c_{p,q} s_{n,q}, \tag{7}$$

where $a_{k,q}$ is the channel coefficient between UEq and antenna k, $c_{p,q}$ and $s_{n,q}$ are the code of length P and the n^{th}symbol of user q, respectively. In turn, each UE encodes its information sequence $[s_{n,q}]_{n=1,\ldots,N}$ using the code $c_{p,q}$ before transmission. By comparing (3) and (7), with following correspondences:

$$(\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}) \longleftrightarrow (\mathbf{A}, \mathbf{C}, \mathbf{S}), \tag{8a}$$

$$(I_1, I_2, I_3) \longleftrightarrow (K, P, N). \tag{8b}$$

Hence, by analogy with (6a) and (6b), $\mathbf{X}_{i=1,2}$ can be written as

$$\mathbf{X}_1 = (\mathbf{A} \diamond \mathbf{C})\,\mathbf{S}^T, \quad \mathbf{X}_2 = (\mathbf{C} \diamond \mathbf{S})\,\mathbf{A}^T \tag{9}$$

There are two different manners that the tensor-based processing can be performed: i) the *centralized case*, where the BS gathers the information of all users, and ii) the *distributed case*, where micro-BS nodes perform the signal processing in a distributed fashion. In the former, each BS antenna receives a copy of $\mathbf{X}_{k..}$ and the BS can then build $\mathcal{X}$ for further processing. In the latter, the micro-BS nodes form a network, following some network topology, such that they estimate $\mathcal{X}$ in a distributed fashion. Such a technique is now discussed.

IV. CONSENSUS AVERAGING ALGORITHM

The consensus averaging (CA) algorithm is based on an iterative exchange procedure for solving a distributed consensus problem. In [17], the author has introduced a CA algorithm for describe a group might reach agreement for the parameter by pooling their initial parameter in time-invariant scenarios. However there are works where the time-varying cases are dealt with, such as [18], [19].

The network analyzed in this work can be represented as an undirected and connected[1] graph $\mathtt{G} := \{\mathtt{E}, \mathtt{I}\}$, where $\mathtt{I} := \{1, \ldots, I\}$ denotes the set of nodes, and $\mathtt{E} \subset \{\mathtt{I} \times \mathtt{I}\}$ is the set of edges. Particularly, nodes in $\mathtt{I}$ represent the micro-BS nodes and edges in $\mathtt{E}$ are related to the connectivity among nodes, thus defining a list of neighborhoods. In other words, each node has a list of neighbor nodes so that $\mathtt{N}_i \subset \mathtt{I}$ is the set of neighbors of node i.

A. Convergence conditions

Given each node i has a scalar value $x_i(t)|_{t=0} = x_i(0)$ as its initial state. Let us group the initial states of the nodes in the vector at time t denoted by $\mathbf{x}(t) = \begin{bmatrix} x_1(t) & x_2(t) & \ldots & x_I(t) \end{bmatrix}^T$. The consensus update equation for $\mathbf{x}(t)$ is given by

$$x_i(t+1) = \sum_{j=1}^{I} w_{ij} x_j(t), \tag{10}$$

where w_{ij} belongs to a non-negative stochastic matrix $\mathbf{W}$ whose rows sum up to 1 each.. Thus, $x_i(t+1)$ is a weighted average of $x_j(t)$ owned to nodes at time t. As an example of matrix $\mathbf{W}$, we have

$$w_{ij} = \begin{cases} \dfrac{1}{\delta_i}, & \text{if } j \in \mathtt{N}_i(t) \\ 0, & \text{if } j \notin \mathtt{N}_i \end{cases}, \tag{11}$$

where $\mathtt{N}_i = \{j \mid (j,i) \in \mathtt{E}\}$ is the set of nodes j whose values are taken into account by i at time t, while δ_i is the degree[2] of node i. We can rewrite (10) in matrix form as $\mathbf{x}(t+1) = \mathbf{W}\mathbf{x}(t)$ and we can obtain $\mathbf{x}(t) = (\mathbf{W})^t \mathbf{x}(0)$ for all t. The weight matrix $\mathbf{W}$ must be chosen to ensure that with the initial value $\mathbf{x}(0)$, $\mathbf{x}(t)$ will converge to the average vector $\bar{\mathbf{x}} = m\mathbf{1}$, where m is the consensus value. Xiao and Boyd have shown in [5] that necessary and sufficient conditions for the convergence in the case of a fixed network are

$$\mathbf{W}\mathbf{1} = \mathbf{1}, \tag{12a}$$

$$\mathbf{1}^T\mathbf{W} = \mathbf{1}^T, \tag{12b}$$

$$\rho\left(\mathbf{W} - \frac{1}{K}\mathbf{1}\mathbf{1}^T\right) < 1, \tag{12c}$$

where $\rho(\mathbf{W})$ is the spectral radius of matrix $\mathbf{W}$.

According to (12b), $\mathbf{1}$ is the left eigenvector of $\mathbf{W}$, which has a single nonzero eigenvalue equal to 1. This implies that the average of all the estimates is conserved for all t, proving that the final consensus value corresponds to the average of $\mathbf{x}(0)$. In (12a), the consensus stability is guaranteed. This implies that if $\mathbf{x}(t_f) = \bar{m}\mathbf{1} = \bar{\mathbf{m}}$, then $\mathbf{x}(t) = \bar{\mathbf{m}}$, $\forall\ t > t_f$, $\bar{\mathbf{m}}$ is the result of consensus average. By satisfying all the conditions given in (12), we guarantee that $\mathbf{W}$ has 1 as eigenvalue, while all the other eigenvalues are significantly smaller than one.

B. Local degree weights

There may be several matrices $\mathbf{W}$ that satisfy the conditions ensuring the convergence of the estimates to $\bar{m}$, but each choice may have a different convergence rate. With the purpose of speeding up the convergence, a judicious choice of the matrix $\mathbf{W}$ is important. There are different algorithms to accomplish this task.

Herein, we consider that each node can calculate its weight locally from the knowledge of the connectivity degree of its neighbors. This approach is known as "local degree weighting". We have:

$$w_{ij} = \begin{cases} \dfrac{1}{\max\{\delta_i, \delta_j\}}, & \text{if } (i,j) \in \mathtt{E} \text{ and } i \neq j \\ 0, & \text{if } (i,j) \notin \mathtt{E} \text{ and } i \neq j \end{cases}, \tag{13}$$

For satisfying the conditions in (12), we calculate w_{ii} as

$$w_{ii} = 1 - \sum_{j \in \mathtt{N}_i} w_{ij}. \tag{14}$$

V. ALGORITHMS

A. Bilinear alternating least squares

Various algorithms can be used for estimating the factor matrices of a tensor $\mathcal{X}$. Since the spreading code matrix $\mathbf{C}$ assumed to known, we are concerned with the estimation of $\mathbf{S}$ and $\mathbf{A}$ by exploiting the two matrix unfolding in (9), respectively, this can be done using the Bilinear Alternating Least Squares (BALS) algorithm. It combines by alternately minimizing the following least squares cost functions in the least square (LS) sense:

[1] A graph $\mathtt{G}$ is connected, in the sense of a topological space, when any two nodes of the graph are connected through a path.

[2] Degree of a node i is the number of edges incident with i

$$\hat{\mathbf{S}} = \arg\min_{\mathbf{S}} \|\mathbf{X}_1 - \mathbf{Y}\mathbf{S}^T\|_F^2 = \sum_{k=1}^{K} \|\mathbf{X}_{k..} - \mathbf{Y}_k\mathbf{S}^T\|_F^2, \tag{15a}$$

$$\hat{\mathbf{A}} = \arg\min_{\{\mathbf{A}_{1.}\ldots\mathbf{A}_{k.}\}} \|\mathbf{X}_2 - \mathbf{Z}\mathbf{A}^T\|_F^2 = \sum_{k=1}^{K} \|\text{vec}(\mathbf{X}_{k..}^T) - \mathbf{Z}\mathbf{A}_{k.}^T\|_F^2. \tag{15b}$$

where $\mathbf{Y} = (\mathbf{A} \diamond \mathbf{C})$, using $\mathbf{Y}_k = \mathbf{C}\text{diag}(\mathbf{A}_{k.})$ and $\mathbf{Z} = (\mathbf{C} \diamond \mathbf{S})$.

In Algorithm 1 we can see the summary of the BALS.

Algorithm 1 BALS

1) Set $i = 0$;
2) Initialize $\hat{\mathbf{A}}(i)$;
3) Use $\mathbf{X}_1$ to find an LS estimate of $\hat{\mathbf{S}}(i)$:
   $\mathbf{Y} = (\hat{\mathbf{A}}(i) \diamond \mathbf{C})$;
   $\hat{\mathbf{S}}(i) = [(\mathbf{Y}^H\mathbf{Y})^{-1}\mathbf{Y}^H\mathbf{X}_1]^T$;
4) $i = i + 1$;
5) Use $\mathbf{X}_2$ to find an LS estimate of $\hat{\mathbf{A}}(i)$:
   $\mathbf{Z} = (\mathbf{C} \diamond \hat{\mathbf{S}}(i))$;
   $\hat{\mathbf{A}}(i) = [(\mathbf{Z}^H\mathbf{Z})^{-1}\mathbf{Z}^H\mathbf{X}_2]^T$;
6) Repeat steps 3-4 until convergence.

The convergence is achieved when the error described in (15) at the i^{th} iteration does not substantially change between iteration i and $i + 1$. Moreover, in this work we use random initializations for matrices $\mathbf{A}$ and $\mathbf{S}$.

B. Distributed Bilinear Alternating Least Squares

In this work, we developed a Distributed Bilinear Alternating Least Squares (D-BALS) algorithm relying on the D-ALS algorithm derived by Kibangou and de Almeida in [9]. We have one consensus estimation case, which corresponds to estimate the symbol matrix $\mathbf{S}$.

An estimate of $\mathbf{S}$ can be obtained by (15a). Thus, we can obtain $\mathbf{S}$ as

$$\hat{\mathbf{S}}^T = \left(\frac{1}{K}\sum_{k=1}^{K} \mathbf{Y}_k^H\mathbf{Y}_k\right)^{-1} \left(\frac{1}{K}\sum_{k=1}^{K} \mathbf{Y}_k^H\mathbf{X}_{k..}\right), \tag{16}$$

where, using property (1), we have $\mathbf{Y}_k = \mathbf{C}\text{diag}(\mathbf{A}_{k.})$. Then, we have to consider two average consensus problems, $\mathbf{\Gamma}_k(0) = \mathbf{Y}_k^H\mathbf{Y}_k$ and $\mathbf{\Theta}_k(0) = \mathbf{Y}_k^H\mathbf{X}_{k..}$. Using (2) in (15b) we can locally estimate $\mathbf{A}$ as $\hat{\mathbf{A}}_k^T = \left(\mathbf{C} \diamond \hat{\mathbf{S}}_k\right)^{\dagger} \text{vec}(\mathbf{X}_{k..}^T)$. Algorithm 2 summarizes the D-BALS.

VI. Simulation Results

In this section, computer simulation results are provided for the performance evaluation of the proposed receive algorithm in some selected system configurations. We consider a scenario similar to that described in Figure 1, where we have a single BS with K antennas, K micro-BS nodes , and Q users are drawn in the area covered by the micro-BS nodes' cluster. The BS use the BALS algorithm (*centralized case*) and the micro-BS node's cluster use D-BALS (*distributed case*). In our simulations, we consider the parameters described in Table I. The results are obtained by an average over a large number of independent Monte Carlo runs, where each run corresponds to a redrawn a users' position, transmitted symbols and channel matrix given $\mathbf{A} = \mathbf{A}^{(R)} \odot \mathbf{A}^{(G)}$, where $\mathbf{A}^{(R)} = [a_{k,q}^{(R)}]$ is small-scale channel matrix, $\mathbf{A}^{(G)} = [a_{k,q}^{(G)}]$ is the large-scale channel matrix, and additive noise. Our goal is to evaluate the impact of path gain in both centralized and distributed cases. As previously mentioned, we consider that data exchanges among micro-BS nodes are error-free.

Algorithm 2 D-BALS

1) For $k = 1, \ldots, K$:
   Initialize $\hat{\mathbf{A}}_k \in \mathbb{C}^{1\times Q}$;
   Compute $\mathbf{Y}_k = \mathbf{C}_k\text{diag}(\hat{\mathbf{A}}_k)$,
   $\mathbf{\Gamma}_k(0) = \mathbf{Y}_k^H\mathbf{Y}_k$ and $\mathbf{\Theta}_k(0) = \mathbf{Y}_k^H\mathbf{X}_{k..}$;
2) Run the consensus algorithm for $\mathbf{\Gamma}$ and $\mathbf{\Theta}$:
   For $t = 0, 1, \ldots, T - 1$,
   $\mathbf{\Gamma}_k(t+1) = \mathbf{\Gamma}_k(t) + \sum_{j\in\mathcal{N}_k} w_{k,j}(\mathbf{\Gamma}_j(t) - \mathbf{\Gamma}_k(t))$.
   $\mathbf{\Theta}_k(t+1) = \mathbf{\Theta}_k(t) + \sum_{j\in\mathcal{N}_k} w_{k,j}(\mathbf{\Theta}_j(t) - \mathbf{\Theta}_k(t))$.
3) Set $\mathbf{\Gamma}_k(0) = \mathbf{\Gamma}_k(T)$ and $\mathbf{\Theta}_k(0) = \mathbf{\Theta}_k(T)$.
4) Compute the local estimates of $\mathbf{S}$:
   $\hat{\mathbf{S}}_k = \mathbf{\Gamma}_k^{-1}(0)\mathbf{\Theta}_k(0)$;
5) Compute local estimates of the channel:
   $\hat{\mathbf{A}}_k^T = \left(\mathbf{C} \diamond \hat{\mathbf{S}}_k\right)^{\dagger} \text{vec}(\mathbf{X}_{k..}^T)$;
6) Return for step 2 until convergence.

TABLE I
SIMULATION PARAMETERS

# of BS antennas / micro-BS nodes K	3,4,5
Code's length P	12
Number of symbols per UE N	100
Transmitter antennas (users) Q	12
Modulation	4QAM
Monte Carlo runs	2000
SNR_0	-6 dB to 30 dB
Consensus iterations	4
Connectivity topology	Ring (degree equal to 3)
BS cell radius	1000-2000m
Micro-node cell radius	200m

The path gain coefficients are obtained following the simplified path-loss model [20] described in (17).

$$PL(d) = 20\log_{10}\left(\frac{\lambda}{4\pi d_0}\right) - 10\gamma\log_{10}\left(\frac{d}{d_0}\right) \text{ in dB, and}$$
$$G(d) = -PL(d) \text{ in dB.} \tag{17}$$

where $PL(d)$ represents the path-loss for distance d(in meters) between receiver and transmitter, d_0 is reference distance, λ is the wavelength, γ is path-loss exponent and $G(d)$ is the path-gain. In this work is considered $d_0 = 20m$, $f_c = 2$Ghz.

The SNR_0, in Table I, is given by $\text{SNR}_0 = \bar{\text{SNR}}_{node}$, where $\bar{\text{SNR}}_{node}$ is the average received signal-to-noise ratio (SNR) for all nodes. Then, the $\bar{\text{SNR}}_{BS}$ is given by $\bar{\text{SNR}}_{BS} = \text{SNR}_0 - \Delta_{PL}$, where $\Delta_{PL} = \bar{PL}_{BS} - \bar{PL}_{node}$, where PL_a represents the path-loss for $a = BS, node$.

Figure 2 shows the average of bit error rate (BER) versus the SNR_0 in dB scale, when the BS radius assumes

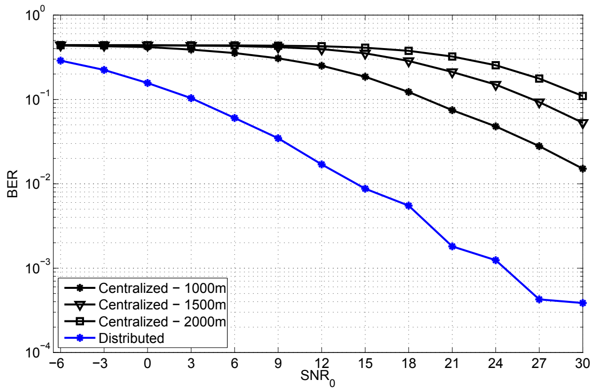
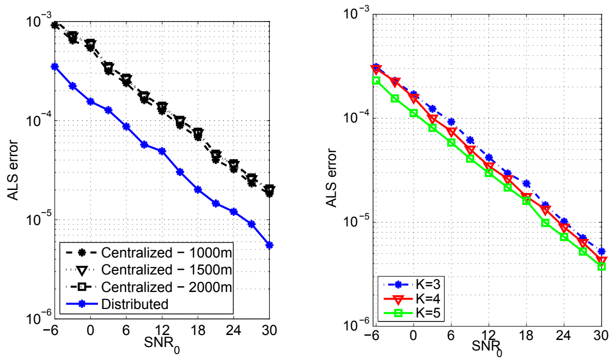


Fig. 2. BER vs SNR, where $K = 3$.



(a) Distributed and centralized with $K = 3$

(b) Distributed with $K = 3, 4, 5$

Fig. 3. Average ALS error.

different values. It can be seen that the distributed processing improves significantly, the BER performance when compared to centralized one. This occurs because the distance between UEs and BS is greater than the distance between UEs and micro-BS nodes. Thus the channel in centralized case link has gain smaller than in the distributed case. Therefore the BS has more difficulty to demodulate the signal.

In Figure 3(a), we can evaluate the convergence of Alternating Least Squares (ALS) algorithms in centralized and distributed cases with same parameter of simulation result in Figure 2. As the previous result, we can see that distributed case has better values than the three distinct configurations of centralized case. It occurs because the difference in received signal power for both cases. Figure 3(b) shows the comparative of the convergence of ALS to distinct number of micro-BS nodes. We have better values when we increase the number of nodes. With the increased number of nodes, we increase the spatial diversity and it contributes to increase the performance of the distributed algorithm.

VII. Conclusions and Future Work

We compared the centralized and distributed cases in uplink of a DS-CDMA system considering centralized and distributed processing cases. We use a trilinear PARAFAC modeling for composite received tensor in both cases, where our receiver exploits space (first dimension), code (second dimension) and time (third dimension) diversity for estimating the user's transmitted symbols and channels. This work has several extensions, e.g.:

- The impact of the cell size on the convergence of the centralized and distributed estimation approaches.
- The impact of different spatial distribuitions of the UEs and micro-BS nodes.
- Computational complexity study of the proposes algorithm or different nodes' connection topologies.

References

[1] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[2] H. Zhu, A. Cano, and G. Giannakis, "Distributed consensus-based demodulation: algorithms and error analysis," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 2044–2054, Jun. 2010.

[3] H. Zhu, A. Cano, and G. B. Giannakis, "Distributed demodulation using consensus averaging in wireless sensor networks," in *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, 2008, pp. 1170–1174.

[4] S. C. Draper, B. J. Frey, and F. R. Kschischang, "Interactive decoding of a broadcast message," in *PROCEEDINGS OF THE ANNUAL ALLERTON CONFERENCE ON COMMUNICATION CONTROL AND COMPUTING*, vol. 41, 2003, pp. 170–180.

[5] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[6] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *Journal of Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33–46, Jan. 2007.

[7] Y. Hatano and M. Mesbahi, "Agreement over random networks," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 2, 2004, pp. 2010–2015.

[8] M. Porfiri and D. J. Stilwell, "Stochastic consensus over weighted directed networks," in *American Control Conference, 2007. ACC'07*, 2007, pp. 1425–1430.

[9] A. Y. Kibangou and A. L. F. De Almeida, "Distributed PARAFAC based DS-CDMA blind receiver for wireless sensor networks," in *Proceedings of the 11th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2010)*, 2010.

[10] F. Roemer and M. Haardt, "Tensor-based channel estimation and iterative refinements for two-way relaying with multiple antennas and spatial reuse," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5720–5735, Nov. 2010.

[11] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.

[12] R. A. Harshman, "Foundations of the parafac procedure: Model and conditions for an 'explanatory' multi-mode factor analysis," *UCLA Working Papers Phonetics*, vol. 16, pp. 1–84, Dec. 1970.

[13] S. Liu and G. Trenkler, "Hadamard, khatri-rao, kronecker and other matrix products," *Int. J. Inform. Syst. Sci*, vol. 4, no. 1, pp. 160–177, 2008.

[14] J. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear algebra and its applications*, vol. 18, no. 2, pp. 95–138, 1977.

[15] N. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of n-way arrays," *Journal of chemometrics*, vol. 14, no. 3, pp. 229–239, 2000.

[16] C. C. Cavalcante, R. F. Colares, and P. C. Barbosa, *Telecommunications: Advances and Trends in Transmission, Networking and Applications*. University of Fortaleza Press, 2006.

[17] M. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[18] J. Tsitsiklis, "Problems in decentralized decision making and computation." DTIC Document, Tech. Rep., 1984.

[19] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *Automatic Control, IEEE Transactions on*, vol. 31, no. 9, pp. 803–812, 1986.

[20] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.