Blind PARAFAC Receivers for Multiple Access-Multiple Antenna Systems

Alexandre de Baynast*, Lieven De Lathauwer† and Behnaam Aazhang*
*Department of Electrical and Computer Engineering
Rice University
MS-366 - 6100 Main Street
Houston, Texas 77005
Telephone: (713) 348-2359
Fax: (713) 334-5686
Email: debaynas@ece.rice.edu
†Equipe Traitement des Images et du Signal
Université de Cergy-Pontoise/ENSEA
6 avenue du Ponceau
Cergy-Pontoise Cedex, France 95014

***Abstract*— In this paper, we present a new blind receiver for multiple access channel with multiple transmit antennas per user and multiple receive antennas (MIMO channel). After being multiplied by a spreading sequence, each user's data is split into N_t streams that are simultaneously transmitted using N_t transmit antennas. The received signal at each receive antenna is a linear superposition of the N_t transmitted signals of the N_u users perturbed by noise.**

We propose a new blind detection/identification algorithm under the assumption that the fading is slow and frequency non-selective. This algorithm relies on a generalization of *parallel factor* analysis (PARAFAC analysis, [Kruskal, Lin. Alg. Appl.'77, Sidiropoulos, Tr. on Sig. Proc.'00]): we show that a generalized canonical decomposition (CANDECOMP) of the 3D data tensor is unique under mild assumptions without noise. Neither algebraic orthogonality nor independence between sources is needed for uniqueness of the decomposition. By performing this decomposition, in rank $(N_t, N_t, 1)$ terms, we are able to retrieve the three sets of parameters: the symbols, the channel fading coefficients (including the antenna gains) and the spreading sequences. In a noisy context, we propose a simple algorithm of the alternating least squares (ALS) type, which yields a performance close to the linear minimum mean square error (LMMSE) receiver which requires knowledge of the channel and spreading sequences.

I. Introduction

In this paper, we address the problem of blind detection/identification for multiple access with spectral and spatial diversities at the transmitter and the receiver.

A common way to ensure good performance at the receiver is the use of spreading sequences to separate the contribution of all users first and then, the use of "space-time" coding to separate the transmitted data streams from each antenna of a single user. Various techniques to exploit the capabilities of multiple transmit antenna - multiple receive antenna (MIMO) channels have been proposed in the literature. When the channel fading is slow (stationary over a few symbols), some authors have proposed an algebraic approach (see for instance [1], [2]). In order to separate the contribution of all users, most researchers rely on "orthogonality" of the sources. All the proposed methods require knowledge of the channel taps and the antenna gains at least at the receiver. This assumption is hard to be satisfied in practical situations.

In this paper, we propose a new simple algorithm for blind source detection and channel taps/antenna gains estimation. Nor knowledge of channel taps/antenna gains neither orthogonality between sources is required in our case. This algorithm relies on a generalization of *parallel factor* analysis (PARAFAC analysis, [3], [4]). Furthermore, our scheme can have a transmission rate up to $N_u N_t (N-1)/(N S_F)$ where N is the frame length over which the channel is supposed to be stationary, N_t the number of transmit antennas per user, S_F the spreading factor and N_u the number of users. The (induced) spectral diversity is realized by a spreading step of rate $1/S_F$.

First, we describe the data model. In the second section, we present the theorem of identifiability in the noiseless case. Then, in a noisy context, we propose a simple practical algorithm of the alternating least squares (ALS) type. Finally, we provide simulation results indicating that this algorithm yields a performance close to the linear minimum mean square error (LMMSE) receiver which requires knowledge of the channel and spreading sequences.

Some notation conventions that will be used in this paper follow.

Lower case bold (as in $\mathbf{v}$) denotes a column vector. $\mathbf{v}^T$ is the transpose, $\otimes$ is the Kronecker product. Upper case (as in A) denotes a matrix; $A^\dagger$ denotes the Moore-Penrose pseudo-inverse. Operator vec(.) for matrix $A = [\mathbf{v}_1 \ldots \mathbf{v}_N] \in \mathbb{C}^{M \times N}$ is defined by $\text{vec}(A) = \left[\mathbf{v}_1^T \cdots \mathbf{v}_N^T\right]^T$.

II. Data model

We consider a mobile communication system where the base-station is equipped with N_r antennas and the N_u mobiles are equipped with N_t antennas. The information frame of each user goes through a serial-to-parallel converter, and is divided

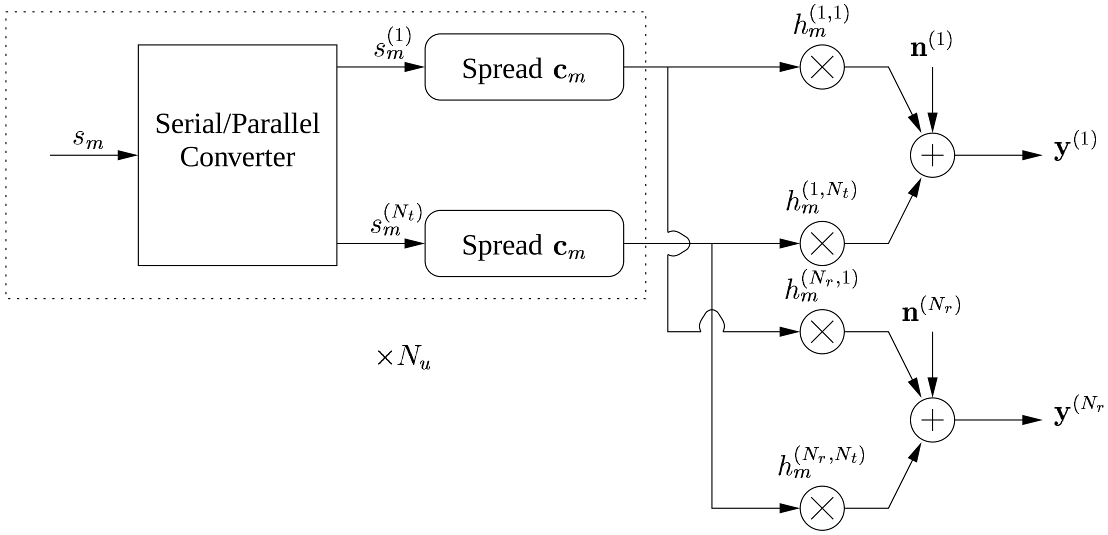

Fig. 1. MIMO transmission model for multiple access.

into N_t streams of data $\mathbf{s}_m = [s_m^{(l)}[1]\ s_m^{(l)}[2]\ \ldots\ s_m^{(l)}[N]],\ 1 \leq l \leq N_t,\ 1 \leq m \leq N_u$. Each transmitted frame $\mathbf{s'}_m^{(l)},\ 1 \leq l \leq N_t,\ 1 \leq m \leq N_u$, results from spreading the information frame $\mathbf{s}_m^{(l)}$ by means of the spreading sequence $\mathbf{c}_m$. This transmission scheme requires only N_u different spreading sequences. We emphasize that the N_t of the N_u users are transmitted simultaneously each from a different transmit antenna and that all these signals have the same transmission period. We assume a narrow-band channel so that the channel is frequency non-selective. Let $\mathbf{h}_m^{(l)} = [h_m^{(1,l)}\ h_m^{(2,l)}\ \ldots\ h_m^{(N_r,l)}]^T$ be the $N_r \times 1$ channel vector from the l-th transmitter of the m-th user to the receiver. The signal at each receive antenna is a noisy superposition of the $N_u \times N_t$ transmitted signals corrupted by Rayleigh fading. This multiple access MIMO transmission scheme is shown on Fig. 1. The data received by the i-th sensor at time $k + jS_F$ are now given by:

$$y^{(i,k)}[j] = \sum_{m=1}^{N_u} \sum_{l=1}^{N_t} h_m^{(i,l)} c_m^{(k)} s_m^{(l)}[j] + n^{(i,k)}[j] \qquad (1)$$

with

i	:	$1 \to N_r$	receive antenna (spatial diversity)
j	:	$1 \to N$	time (temporal diversity)
k	:	$1 \to S_F$	spreading (spectral diversity)
l	:	$1 \to N_t$	transmit antenna
m	:	$1 \to N_u$	users

This decomposition is a generalized CANDECOMP since the CANDECOMP involves only triads (rank-one elements) and here we have a sum of rank-$(N_t, N_t, 1)$ elements. As shown in Fig. II, by stacking the samples along the three diversity dimensions — spectral, spatial (at the receiver) and temporal — we obtain a three-way tensor $\mathcal{Y}$ of dimension $N_r \times N \times S_F$. Symbols matrices $S_m = [\mathbf{s}_m[1]\ \ldots\ \mathbf{s}_m[N]],\ 1 \leq m \leq N_u$, are of dimension $N_t \times N$. $\mathbf{s}_m[j] = [s_m^{(1)}[j]\ \ldots\ s_m^{(N_t)}[j]]^T,\ m = 1, \ldots, N_u,\ j = 1, \ldots, N$. Matrix $\mathbf{S}$ of dimension $N \times N_t N_u$ results of the concatenation of all matrices S_m. The matrices $H_m,\ m = 1, \ldots, N_u$ contain the channel fading coefficients and the antenna gains and are defined as follows: $H_m = [\mathbf{h}_m^{(1)}\ \mathbf{h}_m^{(2)}\ \ldots\ \mathbf{h}_m^{(N_t)}]$, $1 \leq m \leq N_u$, with $\mathbf{h}_m^{(l)} = [h_m^{(1,l)}\ \ldots\ h_m^{(N_r,l)}]^T,\ 1 \leq l \leq N_t$. Matrix $\mathbf{H}$ of dimension $N_r \times N_t N_u$ results of the concatenation of all matrices H_m. Finally, $\mathbf{c}^{(k)},\ 1 \leq k \leq N_r$ contains the k-th spreading coefficient for each user, *i.e.* , $\mathbf{c}^{(k)} = [c_1^{(k)}\ \ldots\ c_{N_u}^{(k)}]^T,\ 1 \leq k \leq S_F$. Matrix C of dimension $S_F \times N_u$ is equal to $C = [\mathbf{c}^{(1)}\ \mathbf{c}^{(2)}\ \ldots\ \mathbf{c}^{(S_F)}]$.

III. Identifiability

In the noiseless case, we are looking for a decomposition of a three-way array in a sum of terms consisting of outer product between a vector and a matrix which itself results from the (inner) product of two matrices, as shown on Figure II. Then, we highlight some algebraic definitions related to the independence of a set of column vectors extracted from a matrix M.

Definition 1 (k-rank of matrix M, [4] page 115): If $M \in \mathbb{C}^{J \times K}$, k_M is the maximum number of columns of matrix M such that any set of k_M columns of M results is linearly independent. By definition, we have: $k_M \leq$ Rank$\,(M)$.

Definition 2 (k'-rank): Given a block matrix $\mathbf{M} = (M_1\ M_2\ \ldots\ M_L)$, with $M_n \in \mathbb{C}^{J \times K}, 1 \leq n \leq L$, the k'-rank is the maximum number of submatrices such that any matrix $\left(M_{i_1}, \ldots, M_{i_{k'}}\right)$ consisting of k' submatrices is full column rank.

We have now the following fundamental theorem of identifiability, applying to the case of MIMO systems.

Theorem 1: Consider the set of data $y^{(i,k)}[j],\ i = 1, \ldots, N_r,\ j = 1, \ldots, N,\ k = 1, \ldots, S_F$, defined by Eq. (1) without noise. Suppose that the matrix C is full k-rank and the matrices $\mathbf{H}$ and $\mathbf{S}$ are full k'-rank (which is always satisfied in practical situations). Suppose also that each matrix $S_m, 1 \leq m \leq N_u$ contains no more than $N_t(N - N_t)$ distinct

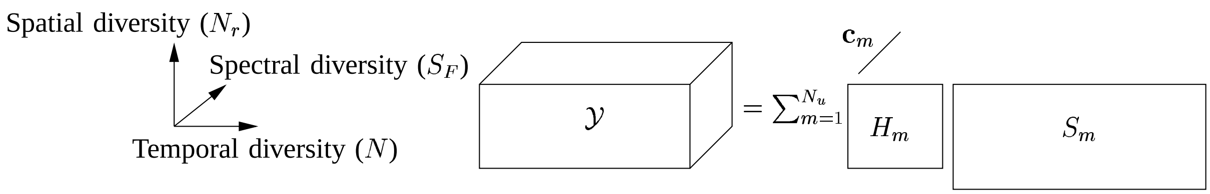

Fig. 2. Observed data tensor of which the dimensions correspond to spectral, spatial (at the receiver) and temporal diversity (noiseless case).

entries and $N-1$ distinct entries per row at the maximum. If

$$k_C + k'_{\mathbf{H}} + k'_{\mathbf{S}^T} \geq 2(N_u+1), \quad (2)$$

then the matrix $C \in \mathbb{C}^{S_F \times N_u}$ is unique up to right multiplication with a permutation matrix and a diagonal matrix and the matrices $\mathbf{H} \in \mathbb{C}^{N_r \times N_t N_u}$ $\mathbf{S} \in \mathbb{C}^{N_t N_u \times N}$ are unique up to right multiplication with a block diagonal matrix of which all the diagonal blocks are equal to the same product of a permutation matrix and a diagonal matrix. More precisely, we have that any other triple $\overline{C}$, $\overline{\mathbf{H}}$, $\overline{\mathbf{S}}$ that leads to an equivalent decomposition of $\{Y_k\}$, is related to C, $\mathbf{H}$ and $\mathbf{S}$ via:

$$\overline{C} = C\Pi\Delta_1,\ \overline{\mathbf{H}} = \mathbf{H}\left(\Pi\Delta_2 \otimes I\right),\ \overline{\mathbf{S}} = \mathbf{S}\Pi\Delta_3 \quad (3)$$

where Π is a permutation matrix and $\{\Delta_i\}_{i=1,2,3}$ are diagonal matrices satisfying $\Delta_1\Delta_2\Delta_3 = I$. ■

Given the constraint that each matrix $S_m, 1 \leq m \leq N_u$ contains no more than $N_t(N-N_t)$ distinct entries, the resulting transmission rate per user and per antenna becomes $\frac{N-N_t}{N}$ which is close to 1 for N large.

To prove this theorem, we use the following Theorem. The proof is given in [5]. First we look what happens when no constraint on the number of distinct entries is imposed on the factors. Then the constraint is incorporated.

Theorem 2 (Generalized identification theorem): Assume a $(N_a \times N_b \times N_c)$-tensor $\mathcal{Y}$. Call $\{Y_i\}_{i=1,\ldots,N_a}$ the matrices that are obtained by slicing the tensor along the first direction. Let these slices have the following structure:

$$Y_i = \sum_{m=1}^{L} a_m^{(i)} B_m C_m^T,\ \forall i, 1 \leq i \leq N_a, \quad (4)$$

with complex scalars $a_m^{(i)},\ 1 \leq m \leq L,\ \ \forall i, 1 \leq i \leq N_a$, and complex matrices $B_m,\ 1 \leq m \leq L$ and $C_m,\ 1 \leq m \leq L$, of dimensions $N_b \times M$ and $N_c \times M$, respectively. We define the matrix A of dimension $N_a \times L$ and the block matrices $\mathbf{B}$ of dimension $N_b \times LM$ and $\mathbf{C}$ of dimension $N_c \times LM$ as follows:

$$A = \begin{pmatrix} a_1^{(1)} & \ldots & a_L^{(1)} \\ a_1^{(2)} & \ldots & a_L^{(2)} \\ \vdots & \vdots & \vdots \\ a_1^{(N_a)} & \ldots & a_L^{(N_a)} \end{pmatrix}$$

$$\mathbf{B} = (B_1\ B_2\ \ldots\ B_{N_b})$$

$$\mathbf{C} = (C_1\ C_2\ \ldots\ C_{N_c})$$

We suppose that matrix A is full k-rank and matrices $\mathbf{B}$ and $\mathbf{C}$ are full k'-rank.

We then have that, if

$$k_A + k'_{\mathbf{B}} + k'_{\mathbf{C}^T} \geq 2(L+1), \quad (5)$$

the matrix A is unique up to a permutation matrix and a diagonal matrix, and the matrices $\mathbf{B}$ and $\mathbf{C}$ are unique up to a block permutation matrix and a non-singular block diagonal matrix. More precisely, any other triple $\overline{A}$, $\overline{\mathbf{B}}$, $\overline{\mathbf{C}}$ that gives rise to Eq. 4, is related to A, $\mathbf{B}$ and $\mathbf{C}$ via:

$$\overline{A} = A\Pi\Delta,\ \overline{\mathbf{B}} = \mathbf{B}(\Pi \otimes I)T_{\text{block}},\ \overline{\mathbf{C}} = \mathbf{C}(\Pi \otimes I)\tilde{T}_{\text{block}}, \quad (6)$$

where Π is a permutation matrix, Δ a diagonal matrix and T_{block} and $\tilde{T}_{\text{block}}$ non-singular block diagonal matrices. Moreover, the m_{th} coefficient $\delta_m,\ 1 \leq m \leq L$ of the diagonal matrix Δ and the m_{th} blocks T_m and $\tilde{T}_m$ of the matrices T_{block} and $\tilde{T}_{\text{block}}$ satisfy: $\delta_m T_m \tilde{T}_m = I,\ 1 \leq m \leq L$. So the matrices $T_m,\ 1 \leq m \leq L$ are the inverses of the matrices $\tilde{T}_m,\ 1 \leq m \leq L$, up to a scaling factor. ■

This theorem allows for a certain indeterminacy of the components. However, if each matrix C_m contains no more than $N_c(L-N_c)$ distinct entries and $L-1$ distinct entries per row at the maximum, and the pattern, according to which entries are repeated, is the same for both matrices, then the matrix T_m is generically unique up to a scaling of its rows. We call a property "generic" when it holds everywhere except for a set of Lebesgue measure zero. Since $\delta_m T_m \tilde{T}_m = I,\ 1 \leq m \leq L$, the matrices $\tilde{T}_m$ are also fixed. The matrix A is then defined up to a permutation matrix and a diagonal matrix, and the matrices $\mathbf{B}$ and $\mathbf{C}$ are unique up to a block permutation matrix and a diagonal matrix.

IV. Alternating least squares algorithm

In the previous section, we saw that, in the absence of noise, the unknown parameters correspond to the factors of a generalized canonical decomposition of the data tensor; by the fundamental identification theorem, this decomposition is essentially unique. In this section, following an ALS approach, we present a simple algorithm that can be used for the estimation of parameters, possibly from noisy observations. The basic idea behind ALS is simple: each time update *one* matrix using least squares conditioned on previously obtained estimates for the remaining matrices; proceed to update the other matrices; repeat until convergence of the least squares cost function.

We will now derive explicit expressions for the conditional updates. Consider Eq. (1):

$$y^{(i,k)}[j] = \sum_{m=1}^{N_u} \sum_{l=1}^{N_t} h_m^{(i,l)} c_m^{(k)} s_m^{(l)}[j] + n^{(i,k)}[j] \qquad (7)$$

First let us consider the conditional update of $\mathbf{H}$, given C and $\mathbf{S}$. By slicing $\mathcal{Y}$ along its first dimension, we obtain:

$$\text{vec}(Y_i) = \begin{pmatrix} c_1^{(1)} S_1^T & \cdots & c_{N_u}^{(1)} S_{N_u}^T \\ & \vdots & \\ c_1^{(S_F)} S_1^T & \cdots & c_{N_u}^{(S_F)} S_{N_u}^T \end{pmatrix} \text{vec}(H^{(i)}) + \text{vec}(N_i),$$

where $Y_i, i = 1, \ldots, N_r$ is of size $S_F \times N$ and

$$H^{(i)} = \begin{pmatrix} h_1^{(i,1)} & \cdots & h_{N_u}^{(i,1)} \\ \vdots & & \vdots \\ h_1^{(i,N_t)} & \cdots & h_{N_u}^{(i,N_t)} \end{pmatrix}$$

This equation will be written as

$$\mathbf{y}_{\underline{1},i} = M(C, \mathbf{S}) \text{vec}\left(H^{(i)}\right) + \mathbf{n}_{\underline{1},i}, \ 1 \leq i \leq N_r. \qquad (8)$$

This equation can be used for a conditional update of matrix $\mathbf{H}$.

Next let us consider the conditional update of $\mathbf{S}$, given C and $\mathbf{H}$. By slicing $\mathcal{Y}$ along its second dimension, we obtain:

$$\text{vec}(Y_j) = \begin{pmatrix} c_1^{(1)} H_1 & \cdots & c_{N_u}^{(1)} H_{N_u} \\ & \vdots & \\ c_1^{(S_F)} H_1 & \cdots & c_{N_u}^{(S_F)} H_{N_u} \end{pmatrix} \text{vec}(S[j]) + \text{vec}(N_i),$$

where $Y_j, j = 1, \ldots, N$ is of size $S_F \times N_r$ and $S[j] = [\mathbf{s}_1[j] \ \ldots \ \mathbf{s}_{N_u}[j]]$ This equation will be written as

$$\mathbf{y}_{\underline{2},j} = M(C, \mathbf{H}) \text{vec}\left(S[j]\right) + \mathbf{n}_{\underline{2},j}, \ 1 \leq j \leq N. \qquad (9)$$

Finally, let us consider the conditional update of C, given $\mathbf{H}$ and $\mathbf{S}$. By slicing the tensor $\mathcal{Y}$ along its first dimension, we obtain:

$$Y_k = \sum_{m=1}^{N_u} c_m^{(k)} H_m S_m^T, \ \forall k, 1 \leq k \leq S_F, \qquad (10)$$

Eq. (10) is equivalent to

$$\text{vec}(\mathbf{Y}_k) = \left[\text{vec}(H_1 \cdot S_1^T) \ldots \text{vec}(H_{N_u} \cdot S_{N_u}^T)\right] \mathbf{c}^{(k)} + \text{vec}(\mathbf{N}_k).$$

This will be written as

$$\mathbf{y}_{\underline{3},k} = M(\mathbf{H}, \mathbf{S}) \mathbf{c}^{(k)}) + \mathbf{n}_{\underline{3},k}, \ 1 \leq k \leq S_F. \qquad (11)$$

The overall algorithm consists of the following steps:

Initialization: randomly initialize two of the three estimated matrices, *e.g.*, $\hat{C}^{(0)}$ and $\hat{\mathbf{S}}^{(0)}$.

p-th step:

1) Update the estimates of the channel fading coefficients $\hat{\mathbf{H}}^{(p)}$:

$$\text{vec}\left(\hat{H}^{(i,p)}\right) = M^\dagger(\hat{C}^{(p-1)}, \hat{\mathbf{S}}^{(p-1)}) \mathbf{y}_{\underline{1},i}, \ 1 \leq i \leq N_r. \qquad (12)$$

2) Update the symbol estimates:

$$\text{vec}\left(\hat{S}^{(p)}[j]\right) = M^\dagger(\hat{C}^{(p-1)}, \hat{\mathbf{H}}^{(p)}) \mathbf{y}_{\underline{2},j}, \ 1 \leq j \leq N. \qquad (13)$$

Normalize the estimates of the sequences emitted by the different users w.r.t. the scaling ambiguity.

3) Update the estimate of $\hat{C}^{(p)}$:

$$\hat{\mathbf{c}}^{(k,p)} = M^\dagger(\hat{\mathbf{H}}^{(p)}, \hat{\mathbf{S}}^{(p)}) \mathbf{y}_{\underline{3},k}, \ 1 \leq k \leq S_F. \qquad (14)$$

End: The iteration is terminated when $\|\hat{S}^{(p)} - \hat{S}^{(p-1)}\| < \epsilon$.

The normalization of the symbol sequences may involve a scaling to unit-norm, to avoid under- and overflow.

V. Simulation results

We have shown in Section III that each matrix $S_m, 1 \leq m \leq N_u$ should not contain more than $N_t(N - N_t)$ distinct entries and should have up to $N - 1$ distinct entries per row in order to guarantee the uniqueness of the decomposition of the 3D data tensor. We propose to assign N_t^2 entries of each matrix $S_m, 1 \leq m \leq N_u$ to zero as shown on Fig. 3. Although any scheme suitable with the previous constraint works, assignment of N_t entries to zero per row improves convergence speed of the ALS algorithm.

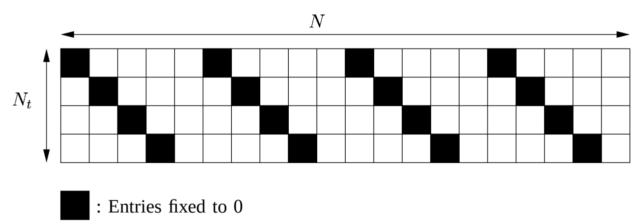

Fig. 3. Zero padding scheme for the symbol matrices S_m, $m = 1, \ldots, N_u$ to guarantee uniqueness of the decomposition of the 3D data tensor.

We compare the performance of the ALS algorithm with the Block-LMMSE receiver which requires knowledge of channel fading coefficients and the spreading sequences. The system load is equal to $N_u/S_F = 2/2 = 1$ in presence of 2 transmit antennas and 2 receive antennas. The frame length is fixed to 10 symbols and modulation is QPSK. On Figure 4, we show that the performance of our algorithm is close to the Block-LMMSE algorithm (about 4 decibels).

VI. Conclusion

In this paper, we have proposed a deterministic way to solve the problem of detection/identification for MIMO systems in a multiuser case in presence of transmit and receive spatial diversities and spectral diversity based on the PARAFAC analysis. This method requires neither channel estimation nor antenna gains. Moreover, the spreading sequences of the different users need not to be orthogonal and the different emitted signals need not to be independent.

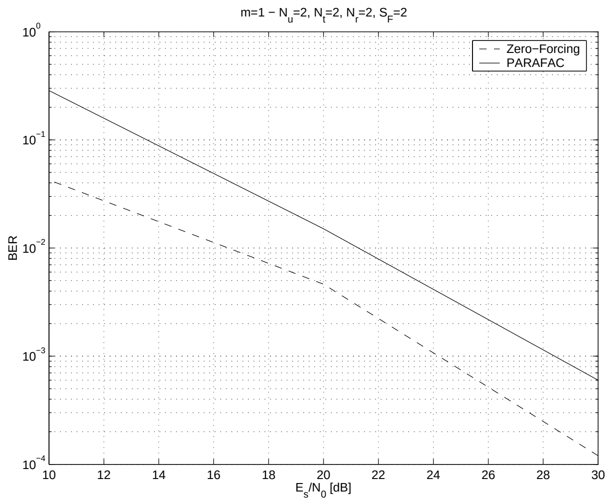

Fig. 4. Symbol Error Rate vs. Signal to noise ratio ($N_u = 2$, $N_t = 2$, $N_r = 2$, $S_F = 2$, $N = 10$).

ACKNOWLEDGMENT

Alexandre de Baynast is currently visiting the Rice University, Houston, TX, supported by a DGA/EGIDE (French Defense) post-doctoral grant.

REFERENCES

[1] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE JSAC*, vol. 16, no. 10, pp. 1451–1458, 1998.

[2] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction,," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 744–765, 1998.

[3] N. Sidiropoulos, G. Giannakis, and R. Bro, "Blind PARAFAC Receivers for DS-CDMA Systems," *IEEE Trans. on Signal Processing*, vol. 48, no. 3, pp. 810–823, March 2000.

[4] J. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Applications*, vol. 18, pp. 95–138, 1977.

[5] L. De Lathauwer and A. de Baynast, "A PARAFAC Generalization and Applications in DS-CDMA," 2002, to be submitted.