TENSOR-BASED BLIND IDENTIFICATION OF MIMO VOLTERRA CHANNELS IN A MULTIUSER CDMA ENVIRONMENT

Carlos Alexandre R. FERNANDES , Gérard FAVIER, and João C. M. MOTA

I3S Laboratory
University of Nice-Sophia Antipolis/CNRS
2000 route des Lucioles, BP 121,06903
Sophia-Antipolis Cedex, France.
phone: + (33) 492942736, fax: + (33) 492942896
email: {acarlos,favier}@i3s.unice.fr

Dep. de Engenharia de Teleinformática
Universidade Federal do Ceará
Campus do Pici, 60.755-640,6007 Fortaleza, Brazil.
phone: +(55 85) 33669467 Fax:+(55 85) 33669468
email: carlosalexandre@deti.ufc.br, mota@gtel.ufc.br

ABSTRACT

This paper is concerned with the blind identification of Multiple-Input-Multiple-Output (MIMO) Volterra channels in a multiuser Code Division Multiple Access (CDMA) environment. The channel is modeled using the most generic representation of complex-valued Volterra systems. A Parallel Factor (PARAFAC) decomposition of a third-order tensor composed of channel output covariances is used, the transmitted signals being assumed to be Phase Shift Keying (PSK) modulated. The channel estimation is carried out by two algorithms: the Alternating Least Squares (ALS) algorithm and a non-iterative least squares algorithm that exploits the redundancy provided by the Khatri-Rao product. The performance of the proposed estimation methods is illustrated by means of computer simulations.

1. INTRODUCTION

This paper proposes two blind identification methods for identifying Multiple-Input-Multiple-Output (MIMO) Volterra channels in the context of a multiuser Code Division Multiple Access (CDMA) communication system. The channel is modeled as the most generic representation of complex-valued Volterra systems. This kind of nonlinear models has important applications in the field of telecommunications, e.g. to model uplink channels in Radio Over Fiber (ROF) multiuser communication systems [1, 2, 3], the nonlinearity of which is introduced by the electrical-optical conversion. These links can be modeled as a MIMO Wiener filter, that constitutes a particular case of MIMO Volterra filters. Another application of such models can be found in CDMA systems with power amplifiers driven at or near saturation to achieve the power consumption requirements [4].

The proposed identification methods use second-order statistics of the signals received by an antenna array, assuming that the transmitted signals are Phase Shift Keying (PSK) modulated. These methods are based on a Parallel Factor (PARAFAC) decomposition [5] of a third-order tensor (three way array) composed of channel output covariances. One of the great advantages of these methods is that they allow working with weak uniqueness conditions compared with previous works [4, 6, 7, 8], that require a number of channel outputs greater than the number of Volterra filter parameters. Indeed, the proposed tensor-based algorithms provide a great flexibility on the number of antennas, which is particularly important when identifying Volterra systems. Moreover, PARAFAC decomposition avoids the use of a pre-whitening step, an operation that increases the computational complexity and may degrade the channel estimation.

The PARAFAC decomposition is first carried out by a two-steps Alternating Least Squares (ALS) algorithm [5]. As the ALS algorithm may need many iterations to converge [9], a non-iterative estimation method exploiting the redundancy of the Khatri-Rao product is also proposed. Few works deal with blind channel estimation or source separation in the context of multiuser nonlinear communication channels. For instance, [4] proposes a blind zero forcing receiver for multiuser CDMA systems with nonlinear channels and [10] develops blind and semi-blind source separation algorithms for memoryless Volterra channels in ultra-wide-band systems.

The paper is organized as follows. Section 2 introduces the nonlinear CDMA channel model used in this work. In Section 3, the chip-level channel output covariance matrices are characterized. In Section 4, a tensor composed of these covariances is defined and a sufficient condition for its uniqueness is given. In Section 5, we present two new blind channel estimation algorithms. In Section 6, we evaluate the performance of these algorithms by means of simulations and some conclusions are drawn in Section 7.

2. THE CDMA SYSTEM WITH NONLINEAR CHANNEL

The discrete time signal transmitted by the t^{th} user ($1 \leq t \leq T$) at time instant n and associated with chip p, is given by

$$u_t(\bar{n}) = c_t(p)s_t(n), \tag{1}$$

where $\bar{n} = (n-1)P+p$, T is the number of users, $c_t(p)$ ($p = 1,..,P$) is the p^{th} element of the spreading code of the t^{th} user, P is the length of the spreading code and $s_t(n)$ is the information signal of the t^{th} user at the n^{th} symbol period. The signal $s_t(n)$ ($1 \leq t \leq T$) is assumed to be PSK modulated, stationary and independent from $s_{t'}(n)$, for $t \neq t'$.

The sampled baseband equivalent channel is modeled as a complex-valued MIMO Volterra filter:

$$y_r(\bar{n}) = \sum_{k=0}^{K} \sum_{t_1=1}^{T} \sum_{t_3=1}^{T} \cdots \sum_{t_{2k+1}=0}^{T} \sum_{m_1=0}^{M} \sum_{m_3=0}^{M} \cdots \sum_{m_{2k+1}=0}^{M} h_{2k+1}^{(r)}(t_1,t_3,\ldots,t_{2k+1},m_1,m_3,\ldots,m_{2k+1}) \prod_{i=1}^{k+1} u_{t_i}(\bar{n}-m_i) \prod_{i=k+2}^{2k+1} u_{t_i}^{*}(\bar{n}-m_i) + \upsilon_r(\bar{n}), \tag{2}$$

where $y_{r,p}(n) = y_r(\bar{n}) = y_r((n-1)P+p)$ ($1 \leq r \leq R$) is the chip rate sampled signal received by antenna r at time instant n and associated with chip p, R is the number of receive antennas, $(2K+1)$ is the nonlinearity order of the model, M is the channel memory, $h_{2k+1}^{(r)}(t_1,\ldots,t_{2k+1},m_1,\ldots,m_{2k+1})$ are the kernel coefficients of the r^{th} sub-channel and $\upsilon_r(\bar{n})$ is the Additive White Gaussian Noise (AWGN) at antenna r. It is assumed that the noise components $\upsilon_r(\bar{n})$ are zero mean, independent from each other and from the transmitted signals $u_t(\bar{n})$.

The model (2) contains only odd power terms composed of products of $k+1$ non-conjugated and k conjugated delayed inputs.

C. A. R. FERNANDES is scholarship supported by CAPES/Brazil agency.

The other nonlinear combinations of delayed input signals generate distortions producing spectral components lying outside of the channel bandwidth, which implies their elimination by bandpass filtering [11, 12, 13, 14, 15].

Assuming that the channel memory is in the order of a few chips, i.e. $M < P$, nonlinear Inter-Symbol Interference (ISI) can be avoided by considering spreading codes that contain "guard-chips" [16]. In this case, the M last elements $c_t(p)$ of spreading codes are equal to zero, i.e. $c_t(p) = 0$, for $P-M+1 \leq p \leq P$. From (1), we may write:

$$\begin{aligned} u_t(\bar{n}-m) &= u_t((n-1)P+p-m) \\ &= \begin{cases} c_t(p-m)\,s_t(n), & \text{if} \quad 1 \leq p-m \leq P, \\ c_t(P+p-m)\,s_t(n-1), & \text{if} \quad p-m \leq 0. \end{cases} \end{aligned} \tag{3}$$

For $p-m \leq 0$, we have $P-M+1 \leq P+p-m \leq P$ and the spreading codes $c_t(P+p-m)$ are null, leading to $c_t(P+p-m) = c_t(p-m) = 0$. We can therefore replace $u_t(\bar{n}-m)$ by $c_t(p-m)s_t(n)$, and equation (2) can be rewritten as:

$$y_{r,p}(n) = \sum_{k=0}^{K} \sum_{t_1=1}^{T} \cdots \sum_{t_{2k+1}=1}^{T} \bar{g}_{2k+1}^{(r,p)}(t_1,\ldots,t_{2k+1}) \prod_{i=1}^{k+1} s_{t_i}(n) \prod_{i=k+2}^{2k+1} s_{t_i}^*(n) + \upsilon_r(\bar{n}), \tag{4}$$

where

$$\bar{g}_{2k+1}^{(r,p)}(t_1,\ldots,t_{2k+1}) = \sum_{m_1=0}^{M} \cdots \sum_{m_{2k+1}=0}^{M} h_{2k+1}^{(r)}(t_1,\ldots,t_{2k+1},m_1,\ldots,m_{2k+1}) \prod_{i=1}^{k+1} c_{t_i}(p-m_i) \prod_{i=k+2}^{2k+1} c_{t_i}^*(p-m_i). \tag{5}$$

The use of "guard-chips" leads to an equivalent memoryless Volterra representation of the channel, the kernel coefficients $\bar{g}_{2k+1}^{(r,p)}(t_1,\ldots,t_{2k+1})$, given by (5), depending on the spreading codes $c_t(p)$ and the original kernel coefficients $h_{2k+1}^{(r)}(t_1,\ldots,t_{2k+1},m_1,\ldots,m_{2k+1})$. Note that the linear kernel $\bar{g}_1^{(r,p)}(t)$ is given by the linear convolution of the linear kernel $h_1^{(r)}(t,m)$ with the spreading code $c_t(p)$. It should be also highlighted that each value of p generates a new output for the equivalent channel $\bar{g}_{2k+1}^{(r,p)}(t_1,\ldots,t_{2k+1})$, leading to a Volterra system with RP outputs at each symbol period n.

Some input terms of (4) are redundant. They can be eliminated by rewriting the Volterra filter in a triangular form. Moreover, as the information signals $s_t(n)$ are PSK modulated, the nonlinear terms corresponding to $t_i = t_j$, for all $i \in \{1,...,k+1\}$ and $j \in \{k+2,...,2k+1\}$, can be eliminated from (4) due to the fact that the term $|s_{t_i}(n)|^2$ reduces to a multiplicative constant that can be absorbed by the associated kernel coefficient, leading to the following equivalent model writing:

$$y_{r,p}(n) = \sum_{k=0}^{K} \sum_{t_1=1}^{T} \cdots \sum_{t_{k+1}=t_k}^{T} \underbrace{\sum_{t_{k+2}=1}^{T} \cdots \sum_{t_{2k+1}=t_{2k}}^{T}}_{t_{k+2},\ldots,t_{2k+1} \neq t_1,\ldots,t_{k+1}} g_{2k+1}^{(r,p)}(t_1,\ldots,t_{2k+1}) \prod_{i=1}^{k+1} s_{t_i}(n) \prod_{i=k+2}^{2k+1} s_{t_i}^*(n) + \upsilon_{r,p}(n), \tag{6}$$

where $\upsilon_{r,p}(n) = \upsilon_r(\bar{n}) = \upsilon_r((n-1)P+p)$.

The RP output signals can be expressed in a matrix form:

$$\mathbf{y}(n) = \mathbf{G}\tilde{\mathbf{s}}(n) + \mathbf{v}(n), \tag{7}$$

where $\mathbf{y}(n) = [y_{1,1}(n)\ldots y_{R,1}(n)\cdots y_{1,P}(n)\ldots y_{R,P}(n)]^T \in \mathbb{C}^{RP\times 1}$ is the vector composed of the signals received by the R antennas and associated with the P chips, $\mathbf{G} = [\mathbf{g}_{1,1}\cdots\mathbf{g}_{R,1}\cdots\mathbf{g}_{1,P}\ldots\mathbf{g}_{R,P}]^T \in \mathbb{C}^{RP\times Q}$ is the channel matrix, with $\mathbf{g}_{r,p} = [g_{r,p,1}\, g_{r,p,2}\ldots g_{r,p,Q}]^T \in \mathbb{C}^{Q\times 1}$ containing the Volterra kernel coefficients $g_{2k+1}^{(r,p)}(t_1,\ldots,t_{2k+1})$ of the $((p-1)R+r)^{th}$ sub-channel, Q being the dimension of each vector $\mathbf{g}_{r,p}$, $\mathbf{v}(n) = [v_{1,1}(n)\ldots v_{R,1}(n)\cdots v_{1,P}(n)\ldots v_{R,P}(n)]^T \in \mathbb{C}^{RP\times 1}$ is the noise vector and $\tilde{\mathbf{s}}(n) = [\tilde{s}_1(n)\ldots\tilde{s}_Q(n)]^T \in \mathbb{C}^{Q\times 1}$ is the nonlinear input vector formed from the spread signals. This vector contains all the linear and nonlinear terms in $s_t(n)$ and $s_t^*(n)$ of (6).

3. CHANNEL OUTPUT COVARIANCE MATRICES

The proposed identification methods rely on the use of covariances of the chip-rate sampled received signals. The covariance matrix of the output signal vector $\mathbf{y}(n)$ is given by:

$$\mathbf{R}_y(d) = \mathbb{E}\left[\mathbf{y}(n+d)\mathbf{y}^H(n)\right] = \mathbf{G}\mathbf{R}_{\tilde{s}}(d)\mathbf{G}^H \in \mathbb{C}^{RP\times RP}, \tag{8}$$

with

$$\mathbf{R}_{\tilde{s}}(d) = \mathbb{E}\left[\tilde{\mathbf{s}}(n+d)\tilde{\mathbf{s}}^H(n)\right] \in \mathbb{C}^{Q\times Q} \tag{9}$$

and $0 \leq d \leq D-1$, where D is the number of delays (covariance matrices) taken into account. The noise covariance matrix is not considered in (8) since it can be estimated and then subtracted from $\mathbf{R}_y(d)$ [17].

In telecommunication systems, the transmitted signals are often assumed to be white. That means that some precoding must be used to introduce time correlation in the signals, otherwise, the covariance matrices $\mathbf{R}_{\tilde{s}}(d)$ are null for $d \neq 0$. In [6, 18], by exploiting some properties of PSK signals, we developed a precoding scheme that introduces a modulation memory in such a way that the matrices $\mathbf{R}_{\tilde{s}}(d)$ are non-null and diagonal. In other words, time correlation is added to the transmitted signals while keeping the orthogonality between products of the transmitted signals. The following theorem states sufficient conditions for assuring the diagonality of the matrices $\mathbf{R}_{\tilde{s}}(d)$.

Theorem 1: Assuming that the information signals $s_t(n)$ are PSK modulated with cardinality $C > 2K+1$, the matrices $\mathbf{R}_{\tilde{s}}(d)$, $d = 0,..,D-1$ are diagonal if the following conditions are satisfied for $(T-1)$ users:

- $\mu_t^{(i,j)}(d) = 0$, for all $0 \leq i,j \leq K+1$ with $i \neq j$;
- $\rho_t^{(i,j)}(d) = 0$, for all $0 \leq i,j \leq K+1$ with i or/and $j \neq K+1$;

where

$$\mu_t^{(i,j)}(d) \equiv \mathbb{E}\left[s_t^i(n+d)\left[s_t^j(n)\right]^*\right] \tag{10}$$

and

$$\rho_t^{(i,j)}(d) \equiv \mathbb{E}\left[s_t^i(n+d)s_t^j(n)\right]. \tag{11}$$

See [18] for the proof. The precoding scheme is designed to ensure that the transmitted signals satisfy the constraints of Theorem 1. The precoding consists in modeling the transmitted signals as Discrete Time Markov Chains (DTMC) that introduce redundancy by expanding the order of the PSK constellation, which means that the number of constellation points is higher than the number needed to modulate the useful data. The states of the DTMC are given by the PSK symbols $a_c = \{A_t \cdot e^{j2\pi(c-1)/C}\}$, for $c = 1,2,...,C$, where

A_t is the amplitude of the signal of the t^{th} user and C is the number of points of the PSK constellation. The state transitions are defined by a set of L_B bits, denoted by $B_n = \{b_n^{(1)}, b_n^{(2)}, ..., b_n^{(L_B)}\}$, that are uniformly distributed over the set $\{0,1\}$ with $2^{L_B} < C$. In addition, it is assumed that the $b_n^{(l)}$ (l=1,...,L_B) are mutually independent. For each of the C states, the set B_n of bits defines 2^{L_B} equiprobable possible transitions. Therefore, this coding imposes some restrictions on the symbol transitions. For each state, there is $\left(C - 2^{L_B}\right)$ not assigned transitions. For further details about this coding scheme, see [18].

The diagonality of $\mathbf{R}_{\tilde{s}}(d)$ implies that (8) can be rewritten as:

$$\mathbf{R}_y(d) = \mathbf{G}\mathsf{diag}_{d+1}[\mathbf{Z}]\mathbf{G}^H, \tag{12}$$

where $\mathrm{diag}_i[\cdot]$ denotes the diagonal matrix formed from the i^{th} row of the matrix argument and the rows of the matrix $\mathbf{Z} \in \mathbb{C}^{D\times Q}$ contain the diagonal elements of $\mathbf{R}_{\tilde{s}}(d)$ for $0 \le d \le D-1$, i.e. $z_{d+1,q} = [\mathbf{Z}]_{d+1,q} = [\mathbf{R}_{\tilde{s}}(d)]_{q,q}$.

4. TENSOR OF CHANNEL OUTPUT COVARIANCES

4.1 Description of the Tensor

A third-order tensor $\underline{\mathbf{R}} \in \mathbb{C}^{D\times RP\times RP}$ composed of channel output covariances can be defined in such a way that $r_{d+1,r_1,r_2} \equiv [\underline{\mathbf{R}}]_{d+1,r_1,r_2} = [\mathbf{R}_y(d)]_{r_1,r_2}$, for $0 \le d \le D-1$ and $1 \le r_1, r_2 \le RP$. This means that the element (r_1, r_2) of the matrix $\mathbf{R}_y(d)$ is placed at position $(d+1, r_1, r_2)$ of $\underline{\mathbf{R}}$. 2D-slices or *matrix slices* of the tensor $\underline{\mathbf{R}}$ are obtained by fixing one index of the tensor and varying the two other ones. For instance, the first-mode matrix slice $\mathbf{R}_{d+1,\cdot\cdot} \in \mathbb{C}^{RP\times RP}$, obtained by fixing the first index of $\underline{\mathbf{R}}$ to $(d+1)$, and by varying the second and third indices, is given by (12). That corresponds to a matrix slice writing of the PARAFAC decomposition of a third-order tensor with rank Q and matrix factors (components) $\mathbf{G}$, $\mathbf{G}^*$ and $\mathbf{Z}$. The corresponding scalar writing of the tensor $\underline{\mathbf{R}}$ is:

$$r_{(d+1),r_1,r_2} = \sum_{q=1}^{Q} g_{r_1,q} g^*_{r_2,q} z_{d,q}, \tag{13}$$

where $g_{r_1,q} = [\mathbf{G}]_{r_1,q}$. The other matrix slices of $\underline{\mathbf{R}}$ are given by:

$$\mathbf{R}_{\cdot r_1\cdot} = \mathbf{G}^* \,\mathrm{diag}_{r_1}[\mathbf{G}]\mathbf{Z}^T \in \mathbb{C}^{RP\times D} \tag{14}$$

and

$$\mathbf{R}_{\cdot,\cdot,r_2} = \mathbf{Z}\,\mathsf{diag}_{r_2}[\mathbf{G}^*]\;\mathbf{G}^T \in \mathbb{C}^{D\times RP}. \tag{15}$$

All the elements of a tensor can be organized in *unfolded matrices* by stacking all the matrix slices of a given type. The estimation algorithms presented in the next section are based on the following unfolded matrices of the tensor:

$$\mathbf{R}_{[1]} \equiv \begin{bmatrix} \mathbf{R}_{1\cdot\cdot} \\ \vdots \\ \mathbf{R}_{D\cdot\cdot} \end{bmatrix}, \qquad \mathbf{R}_{[2]} \equiv \begin{bmatrix} \mathbf{R}_{\cdot 1\cdot} \\ \vdots \\ \mathbf{R}_{\cdot RP\cdot} \end{bmatrix} \tag{16}$$

and

$$\mathbf{R}_{[3]} \equiv \begin{bmatrix} \mathbf{R}_{\cdot\cdot 1} \\ \vdots \\ \mathbf{R}_{\cdot\cdot RP} \end{bmatrix}, \tag{17}$$

where $\mathbf{R}_{[1]} \in \mathbb{C}^{RPD\times RP}$, $\mathbf{R}_{[2]} \in \mathbb{C}^{R^2P^2\times D}$ and $\mathbf{R}_{[3]} \in \mathbb{C}^{RPD\times RP}$ denote respectively the first, second and third-mode unfolded matrices of the tensor $\underline{\mathbf{R}}$. These unfolded matrices are given by:

$$\mathbf{R}_{[1]} = (\mathbf{Z} \diamond \mathbf{G})\,\mathbf{G}^H, \tag{18}$$

$$\mathbf{R}_{[2]} = (\mathbf{G} \diamond \mathbf{G}^*)\,\mathbf{Z}^T \tag{19}$$

and

$$\mathbf{R}_{[3]} = (\mathbf{G}^* \diamond \mathbf{Z})\,\mathbf{G}^T, \tag{20}$$

where $\diamond$ denotes the Khatri-Rao (column-wise Kronecker) product.

It is important to note that, in the case of a memoryless channel ($M = 0$), equation (5) becomes:

$$\bar{g}^{(r,p)}_{2k+1}(t_1,\ldots,t_{2k+1}) = h^{(r)}_{2k+1}(t_1,\ldots,t_{2k+1},0,\ldots,0) \prod_{i=1}^{k+1} c_{t_i}(p) \prod_{i=k+2}^{2k+1} c^*_{t_i}(p). \tag{21}$$

This means that the contributions of the channel coefficients $h^{(r)}_{2k+1}(\cdot)$ and of the codes $c_t(p)$ in the Volterra kernel coefficients $\bar{g}^{(r,p)}_{2k+1}(\cdot)$ can be decoupled. In this case, a fifth-order PARAFAC tensor can be formed from the channel output covariances, with two dimensions corresponding to the receive antennas, two dimensions corresponding to the chips and one corresponding to the covariance delay [19].

4.2 Uniqueness Condition

The main property of the PARAFAC decomposition is its essential uniqueness, demonstrated in [20]. Let us denote by $k_{\mathbf{A}}$ the k-rank of matrix $\mathbf{A}$, i.e. the greatest integer $k_{\mathbf{A}}$ such that every set of $k_{\mathbf{A}}$ columns of $\mathbf{A}$ is linearly independent. Considering a F^{th}-order tensor $\underline{\mathbf{A}}$ of rank Q, with matrix factors $\mathbf{A}_f \in \mathbb{C}^{L_f\times Q}$, $f = 1,...,F$, it is proved in [20] that if:

$$\sum_{f=1}^{F} k_{\mathbf{A}_f} \ge 2Q + F - 1, \tag{22}$$

then the matrix factors $\mathbf{A}_f$ are unique up to column scaling and permutation ambiguities. In the case of the tensor $\underline{\mathbf{R}}$, essential uniqueness means that any other set of matrices $\mathbf{G}'$, $\mathbf{G}''$ and $\mathbf{Z}'$ satisfying (13) is such that $\mathbf{G}' = \mathbf{G}\Pi\Lambda_a$, $\mathbf{G}'' = \mathbf{G}^*\Pi\Lambda_b$ and $\mathbf{Z}' = \mathbf{Z}\Pi\Lambda_c$, where Λ_a, Λ_b and Λ_c are diagonal matrices such that $\Lambda_a\Lambda_b\Lambda_c = \mathbf{I}_Q$ and Π is a permutation matrix.

The matrix $\mathbf{Z}$ containing the information about the time correlation introduced by the precoding scheme, can be assumed to be known, as shown in [18]. So, if condition (22) is verified, we have $\mathbf{Z}' = \mathbf{Z}$ and, hence, $\Pi = \Lambda_c = \mathbf{I}_Q$ and $\Lambda_b = \Lambda_a^{-1}$. Thus, $\mathbf{G}' = \mathbf{G}\Lambda_a$ and $\mathbf{G}'' = \mathbf{G}^*\Lambda_a^{-1}$. This means that the permutation ambiguity is eliminated. Moreover, the scaling ambiguity does not represent an effective problem, as it can be removed by a gain control at the receiver or using a few pilot symbols.

Assuming that the matrix factors $\mathbf{G}$ and $\mathbf{Z}$ are full k-rank, the condition (22) for the tensor $\underline{\mathbf{R}}$ becomes:

$$2\min(RP, Q) + \min(D, Q) \ge 2Q + 2, \tag{23}$$

This uniqueness condition is weaker than that associated with other estimation methods [4, 6, 7, 8, 18]. The flexibility on the choice of R provided by condition (23) is one of the main advantages of using a tensor-based approach, which is particularly interesting when identifying nonlinear systems that are characterized by a high number of parameters. In particular, it is possible to choose $R \ll Q$ (underdetermined case).

We have to note that the sufficient condition (22) is not necessary for the uniqueness of the tensor decomposition. In the next section, we state an alternative sufficient uniqueness condition when the matrix $\mathbf{Z}$ is known.

5. CHANNEL ESTIMATION

5.1 Two-Steps ALS algorithm

The PARAFAC factors are classically estimated by means of the ALS algorithm [16]. If all the matrix factors are unknown, the ALS algorithm allows to estimate these matrix factors in an alternating way. A two-steps version of the classical ALS algorithm can then be used by exploiting the fact that the matrix $\mathbf{Z}$ is assumed to be known. The algorithm provides two channel estimates, denoted by $\hat{\mathbf{G}}_a$ and $\hat{\mathbf{G}}_b$, corresponding respectively to the matrices $\mathbf{G}$ and $\mathbf{G}^*$. The channel estimation problem is solved by minimizing the following cost functions in an alternate way:

$$J_{[3]} = \left\| \hat{\mathbf{R}}_{[3]} - (\mathbf{G}_b \diamond \mathbf{Z})\, \mathbf{G}_a^T \right\|_F^2, \tag{24}$$

$$J_{[1]} = \left\| \hat{\mathbf{R}}_{[1]} - (\mathbf{Z} \diamond \hat{\mathbf{G}}_a)\, \mathbf{G}_b^T \right\|_F^2, \tag{25}$$

where $\hat{\mathbf{R}}_{[1]}$ and $\hat{\mathbf{R}}_{[2]}$ are respectively the sample estimate of $\mathbf{R}_{[1]}$ and $\mathbf{R}_{[2]}$, and $\|\cdot\|_F$ denotes the *Frobenius norm*. The it^{th} iteration of the ALS algorithm is given by:

$$\hat{\mathbf{G}}_a^{(it)} = \left[\left(\hat{\mathbf{G}}_b^{(it-1)} \diamond \mathbf{Z} \right)^{\dagger} \hat{\mathbf{R}}_{[3]} \right]^T, \tag{26}$$

$$\hat{\mathbf{G}}_b^{(it)} = \left[\left(\mathbf{Z} \diamond \hat{\mathbf{G}}_a^{(it)} \right)^{\dagger} \hat{\mathbf{R}}_{[1]} \right]^T, \tag{27}$$

where the initial value $\hat{\mathbf{G}}_b^{(0)}$ is chosen as an $RP \times Q$ Gaussian random matrix and $(\cdot)^{\dagger}$ denotes the matrix pseudo-inverse. The algorithm iterates until the convergence of the estimated parameters is achieved. The existence of the left inverse of the matrices $(\mathbf{Z} \diamond \mathbf{G})$ and $(\mathbf{G}^* \diamond \mathbf{Z})$ is assured if condition (23) is satisfied [21].

5.2 Single-LS algorithm

The ALS algorithm is monotonically convergent but it may require a large number of iterations to converge and/or it can converge towards a local minimum. In order to avoid these convergence problems, we propose a non-iterative channel estimation method. This method exploits the redundancy of the Khatri-Rao product in the following cost function:

$$J_{[2]} = \left\| \hat{\mathbf{R}}_{[2]} - (\mathbf{G} \diamond \mathbf{G}^*)\, \mathbf{Z}^T \right\|_F^2 \tag{28}$$

Defining $\mathbf{W} \equiv (\mathbf{G} \diamond \mathbf{G}^*) \in \mathbb{C}^{R^2P^2 \times Q}$, the LS estimate of $\mathbf{W}$ is given by:

$$\hat{\mathbf{W}} = \hat{\mathbf{R}}_{[2]} \left(\mathbf{Z}^T \right)^{\dagger}. \tag{29}$$

The channel matrix $\mathbf{G}$ can be estimated from $\hat{\mathbf{W}}$ by using the fact that:

$$\mathbf{G} \diamond \mathbf{G}^* = \begin{pmatrix} \mathbf{G}^* \Lambda_1 \\ \vdots \\ \mathbf{G}^* \Lambda_{RP} \end{pmatrix}, \tag{30}$$

where Λ_r, $r = 1, \ldots, RP$, is a diagonal matrix formed with the elements of the r^{th} row of $\mathbf{G}$. Let us define $\hat{\mathbf{W}}^{(r)}$, for $r = 1, \ldots, RP$, as the matrix formed from the lines $[(r-1)RP+1]$ up to (rRP) of $\hat{\mathbf{W}}$. Thus, the channel matrix $\hat{\mathbf{G}}$ can be estimated up to a diagonal matrix as the mean:

$$\hat{\mathbf{G}} = \frac{1}{RP} \sum_{r=1}^{RP} \left[\hat{\mathbf{W}}^{(r)} \right]^*. \tag{31}$$

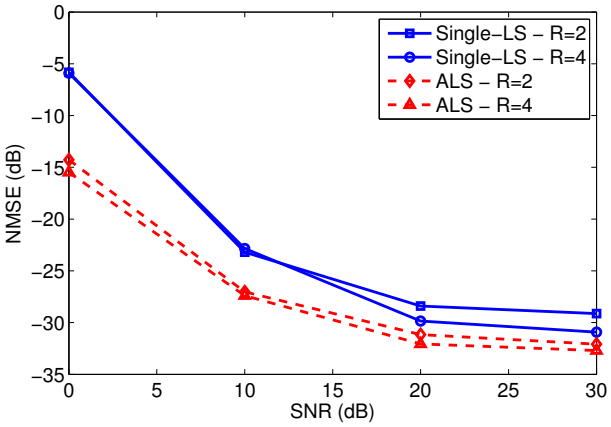

Figure 1: NMSE versus SNR for $R = 2$ and $R = 4$.

The Single-LS method requires that the following identifiability condition be satisfied: $r_{\mathbf{Z}} = Q$, i.e. $\mathbf{Z}^T$ is right invertible or equivalently $\mathbf{Z}$ is full column rank. So, another advantage of this approach is that it does not impose constraints on the number of antennas, contrarily to the ALS algorithm and to other methods [4, 6, 7, 8, 18].

6. SIMULATION RESULTS

In this section, the proposed channel estimation algorithms are evaluated by means of simulations. A MIMO Wiener model of an uplink channel of a ROF multiuser communication system [1, 6] is considered for the simulations. The multipath wireless link is characterized by a $R \times 2$ convolutive linear mixer, with $T = 2$ users ($Q = 4$) and R half-wavelength spaced receive antennas. The E/O conversion is modeled by the following polynomial $c_1 x + c_3 |x|^2 x$, with $c_1 = -0.35$ and $c_3 = 1$ [1, 22]. All the Monte Carlo simulation results were obtained using $N_R = 100$ independent data realizations and the modulation of the transmitted signals is 8-PSK. The spreading codes are complex exponentials with an unitary modulus and a phase uniformly distributed over the set $[-\pi, \pi]$.

The performance is evaluated in terms of the Normalized Mean Squared Error (NMSE) of the estimated channel parameters, defined as: $NMSE = \frac{1}{N_R} \sum_{l=1}^{N_R} \frac{\|\mathbf{G} - \hat{\mathbf{G}}_l\|_F^2}{\|\mathbf{G}\|_F^2}$, where $\hat{\mathbf{G}}_l$ represents the channel matrix estimated at the l^{th} Monte Carlo simulation.

Figure 1 shows the NMSE versus Signal-to-Noise-Ratio (SNR) provided by the ALS and Single-LS algorithms for $R = 2$ and 4, with $P = 3$ and $D = 5$. From these simulation results, we can conclude that the performance of the proposed estimation methods is not deteriorated in the underdetermined case ($R = 2$), with respect to the overdetermined case ($R = 4$). Moreover, it can be viewed that the ALS algorithm provides better performances than the Single-LS algorithm. However, the Single-LS algorithm has a computational cost significantly smaller. For instance, in Figure 1, for $SNR = 0dB$, the ALS algorithm needs approximatively 6 iterations to converge, with two LS estimate computations per step, while the Single-LS algorithm computes only one LS estimate. For small SNR values and $D = 3$ or 4, the ALS algorithm can take more than 50 iterations to converge.

Figure 2 shows the NMSE versus the number of used covariance delays D with the ALS and Single-LS algorithms for $SNR = 10dB$ and $SNR = 30dB$, with $P = 3$ and $R = 2$. In order to provide a performance reference, we also show the NMSE obtained with the supervised Wiener solution, given by:

$$\hat{\mathbf{H}} = \hat{\mathbf{R}}_{\mathbf{y}\tilde{\mathbf{s}}} \hat{\mathbf{R}}_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}^{-1}, \tag{32}$$

where $\hat{\mathbf{R}}_{\mathbf{y}\tilde{\mathbf{s}}}$ and $\hat{\mathbf{R}}_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}}$ are respectively the sample estimates of $\mathbf{R}_{\mathbf{y}\tilde{\mathbf{s}}} =$

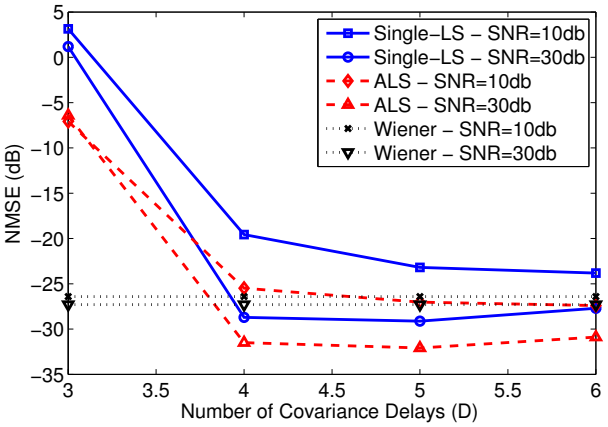

Figure 2: NMSE versus the number of covariance delays D for $SNR = 10dB$ and $SNR = 30dB$.

$\mathrm{E}\left[\mathbf{y}(n)\tilde{\mathbf{s}}^H(n)\right]$ and $\mathbf{R}_{\tilde{\mathbf{s}}\tilde{\mathbf{s}}} = \mathrm{E}\left[\tilde{\mathbf{s}}(n)\tilde{\mathbf{s}}^H(n)\right] = \mathbf{I}_Q$, $\mathbf{I}_Q$ being the identity matrix of order Q. From these simulation results, we can conclude that the accuracy of the ALS estimate does not change significantly for $D \geq 4$. Besides, the performance of the two proposed estimation methods is better than that of the supervised Wiener solution for $SNR = 30dB$. This is due to the fact that the Wiener solution does not exploit the time correlation of the transmitted signals, while the proposed methods do.

7. CONCLUSION

In this paper, we have proposed two new methods for identifying MIMO Volterra communication channels using the PARAFAC decomposition of a tensor composed of channel output covariances, with PSK input signals. This tensor-based approach provides weak uniqueness conditions, leading to weaker constraints on the number of antennas than those imposed by other existing estimation methods. The proposed channel estimation algorithms have been applied for identifying an uplink channel in a ROF multiuser communication system. Some simulation results have illustrated the performance of these algorithms, the ALS providing better performance, at the price of a higher computational cost with respect to the Single-LS algorithm. Both algorithms outperform the Wiener solution for high SNR values.

REFERENCES

[1] S. Z. Pinter and X. N. Fernando, "Estimation of radio-over-fiber uplink in a multiuser CDMA environment using PN spreading codes," in *Canadian Conf. on Elect. and Comp. Eng.*, May 1-4, 2005, pp. 1–4.

[2] X. N. Fernando and A. B. Sesay, "Higher order adaptive filter based predistortion for nonlinear distortion compensation of radio over fiber links," in *Intern. Conf. on Comm. (ICC)*, New-Orleans, LA, USA, June 2000, vol. 1/3, pp. 367–371.

[3] X. N. Fernando and A. B. Sesay, "A Hammerstein-type equalizer for concatenated fiber-wireless uplink," *IEEE Trans. on Vehicular Tech.*, vol. 54, no. 6, pp. 1980–1991, 2005.

[4] A. J. Redfern and G. T. Zhou, "Blind zero forcing equalization of multichannel nonlinear CDMA systems," *IEEE Trans. on Sig. Proc.*, vol. 49, no. 10, pp. 2363–2371, Oct. 2001.

[5] R. A. Harshman, *Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis*, UCLA Working Papers in Phonetics, pp. 1-84, 16*th* edition, Dec. 1970.

[6] C. A. R. Fernandes, G. Favier, and J. C. M. Mota, "Blind source separation and identification of nonlinear multiuser channels using second order statistics and modulation codes," in *IEEE Signal Processing Advances in Wireless Communications (SPAWC) workshop*, Helsinki, Finland, 17-20 Jun. 2007.

[7] G. B. Giannakis and E. Serpedin, "Linear multichannel blind equalizers of nonlinear FIR Volterra channels," *IEEE Trans. on Sig. Proc.*, vol. 45, no. 1, pp. 67–81, Jan. 1997.

[8] R. Lopez-Valcarce and S. Dasgupta, "Blind equalization of nonlinear channels from second-order statistics," *IEEE Trans. on Sig. Proc.*, vol. 49, no. 12, pp. 3084–3097, Dec. 2001.

[9] R. Bro, *Multi-way analysis in the food industry: Models, algorithms and applications*, Ph.D. thesis, University of Amsterdam, Amsterdam, 1998.

[10] N. Petrochilos and K. Witrisal, "Semi-blind source separation for memoryless Volterra channels in UWB and its uniqueness," in *IEEE Workshop on Sensor Array and Multichannel Proc.*, Waltham, MA, USA, 12-14 Jul. 2006, pp. 566–570.

[11] S. Benedetto, E. Biglieri, and R. Daffara, "Modeling and performance evaluation of nonlinear satellite links - A Volterra series approach," *IEEE Trans. on Aerospace Electronic Systems*, vol. 15, pp. 494–507, Jul. 1979.

[12] S. Benedetto and E. Biglieri, "Nonlinear equalization of digital satellite channels," *IEEE J. on Select. Areas in Comm.*, vol. 1, no. 1, pp. 57–62, Jan. 1983.

[13] E. Biglieri, A. Gersho, R. Gitlin, and T. Lim, "Adaptive cancellation of nonlinear intersymbol interference for voiceband data transmission," *IEEE J. on Select. in Areas Commun.*, vol. 2, no. 5, pp. 765–777, 1984.

[14] S. Serfaty, J.L. LoCicero, and G.E. Atkin, "Cancellation of nonlinearities in bandpass QAM systems," *IEEE Trans. on Comm.*, vol. 38, no. 10, pp. 1835–1843, 1990.

[15] C.-H. Cheng and E.J. Powers, "Optimal Volterra kernel estimation algorithms for a nonlinear communication system for PSK and QAM inputs," *IEEE Trans. on Sig. Proc.*, vol. 49, no. 1, pp. 147–163, 2001.

[16] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. on Sig. Proc.*, vol. 48, no. 3, pp. 810–823, Mar. 2000.

[17] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. on Sig. Proc.*, vol. 45, no. 2, pp. 434–444, Feb. 1997.

[18] C. A. R. Fernandes, G. Favier, and J. C. M. Mota, "A modulation code-based blind receiver for memoryless multiuser Volterra channels," in *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 4-7 2007.

[19] C. A. R. Fernandes, G. Favier, and J. C. M. Mota, "Blind estimation of nonlinear instantaneous channels in multiuser CDMA systems with PSK inputs," in *IEEE Signal Processing Advances in Wireless Communications (SPAWC) workshop*, Recife, Brazil, Jul. 6-9 2008.

[20] N.D. 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.

[21] X. Liu and N.D. Sidiropoulos, "Cramer-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Trans. on Sig. Proc.*, vol. 49, no. 9, pp. 2074–2086, Sept. 2001.

[22] S. Z. Pinter and X. N. Fernando, "Concatenated fiber-wireless channel identification in a multiuser CDMA environment," *IET Communications*, vol. 1, no. 5, pp. 937–944, Oct. 2007.