BLIND SOURCE SEPARATION AND IDENTIFICATION OF NONLINEAR MULTIUSER CHANNELS USING SECOND ORDER STATISTICS AND MODULATION CODES

Carlos Alexandre R. FERNANDES, Gérard FAVIER*

I3S Laboratory
CNRS - University of Nice-Sophia Antipolis
2000 route des Lucioles, BP 121,
06903, Sophia-Antipolis Cedex, France.
acarlos@i3s.unice.fr and favier@i3s.unice.fr

João Cesar M. MOTA

GTEL Laboratory
Federal University of Ceará
Campus do Pici, 60.755-640,
6007 Fortaleza, Brazil.
mota@gtel.ufc.br

ABSTRACT

In this paper, a Blind Source Separation (BSS) and channel identification method using second order statistics is proposed for nonlinear multiuser communication channels. It is based on the joint diagonalization of a set of covariance matrices. Modulation codes (constrained codes) are used to ensure the orthogonality of nonlinear combinations of the transmitted signals, allowing the application of a joint diagonalization based estimation algorithm. This constitutes a new application of modulation codes, used to introduce temporal redundancy and to ensure some statistical constraints. Identifiability conditions for the problem under consideration are addressed and some simulation results illustrate the performance of the proposed method.

1. INTRODUCTION

This work deals with the Blind Source Separation (BSS) and system identification problem for nonlinear multiuser communication channels. The considered channel is modelled as a complex-valued linear-cubic Multiple-Input-Multiple-Output (MIMO) Volterra filter, which consists of a generical representation of instantaneous linear-cubic polynomial mixtures. This kind of nonlinear models has important applications in the field of telecommunications to model wireless communication links with nonlinear power amplifiers [1] and uplink channels in Radio Over Fiber (ROF) multiuser communication systems [2]. The ROF links have found a new important application with their introduction in microcellular wireless networks [3, 4]. This kind of network architecture provides to the system a better capacity, coverage and power consumption. Thus, it can also improve the system reliability and Quality of Service. The uplink transmission of such systems is done from a mobile station towards a Radio Access Point, where the transmitted signals are converted in optical frequencies by a laser diode and then retransmitted through optical fibers. Important nonlinear distortions are introduced by the electrical-optical (E/O) conversion [3, 4]. When the length of the optical fiber is short (few kilometers) and the radio frequency has an order of GHz, the dispersion of the fiber is negligible [5]. In this case, the nonlinear distortion arising from the E/O conversion process becomes preponderant [3, 4, 5]. Up to several Mbps, the ROF channel can be considered as a memoryless link [2, 3]. Thus, in a multiuser system, the wireless link can be viewed as a linear mixture and the overall uplink channel as a memoryless MIMO Wiener system [2], which is a particular case of the model considered in this work. Moreover, in this application, the channel nonlinearity is modelled as a third order polynomial [2, 3].

In what concerns the BSS problem for nonlinear systems, there are many works dealing with Post Nonlinear (PNL) mixtures [1, 6, 7]. Some nonlinear BSS techniques are performed in several steps (multi-stage processing) [1, 7] and some works use a joint diagonalization of spatio-temporal covariance matrices to perform some of these stages [1, 8]. However, in the context of communication systems, the transmitted signals are often assumed to be white, which requires the use of some coding, like those developed in this paper, to introduce time correlation in the signals. There are few works dealing with the problem of BSS or/and identification of nonlinear systems in the context of multiuser communication channels. Among them, we cite [9] that proposes a blind zero forcing based equalization technique for Code Division Multiple Access (CDMA) systems.

The technique proposed in this paper exploits the use of Second Order Statistics (SOS) of the received signals. Modulation codes (constrained codes) [10] are used to ensure the orthogonality of nonlinear combinations of the transmitted signals for several time delays, allowing the application of a joint diagonalization algorithm [11, 12] to a set of estimated spatio-temporal covariance matrices. The proposed modulation codes introduce redundancy by expanding the signal constellation, generating multilevel modulations. Modulation expanding is often used in bandwidth-constrained channels, where a performance gain can be achieved without expanding the channel bandwidth or the transmission power [10]. Modulation codes have applications in magnetic record, optical recording and in digital communications over cable systems, with the goal of achieving spectral shaping and minimizing the DC content in the baseband signal [10]. This kind of coding was also used in [13] to reduce intrachannel nonlinear effects in high-speed optical transmissions.

In this work, the modulation codes are explored with a different purpose: the nonlinear channel identification. The redundancy provided by the codes introduces temporal correlation in a controlled way, in order that the transmitted signals verify some statistical constraints associated with the channel nonlinearities.

The method developed in this paper can be viewed as an extension of the Second Order Blind Identification (SOBI) algorithm [12] to nonlinear channels. The SOBI algorithm is a blind source separation technique for linear instantaneous channels. It uses a joint diagonalization based estimator that exploits the temporal correlation of the sources. Joint diagonalization has been addressed by some other authors in the context of communication systems in the case of

*Thanks to CAPES/Brazil agency for funding.

1-4244-0955-1/07/$25.00 ©2007 IEEE.

linear channels, like in [14] that proposes a time-varying user power loading to enable the application of the PARAFAC analysis, with the goal of performing blind estimation of spatial signatures.

2. SYSTEM MODEL

The sampled baseband equivalent model of the communication channel under consideration is assumed to be expressed as complex linear-cubic polynomials of the form:

$$x^{(i)}(n) = \sum_{m_1=1}^{M} h_1^{(i)}(m_1)s_{m_1}(n) + \sum_{m_1=1}^{M}\sum_{m_2=m_1}^{M}\sum_{\substack{m_3=1\\ m_3\neq m_1\\ m_3\neq m_2}}^{M} h_3^{(i)}(m_1,m_2,m_3)s_{m_1}(n)s_{m_2}(n)s_{m_3}^*(n) + v^{(i)}(n), \quad (1)$$

where $x^{(i)}(n)$ is the signal received by the antenna i $(i = 1, 2, ..., I)$ at the time instant n, I is the number of antennae, M is the number of users, $h_{2k+1}^{(i)}(m_1, \ldots, m_{2k+1})$, for $k = 0, 1$, are the channel coefficients, $s_m(n)$, for $1 \leq m \leq M$, are the unknown stationary and statistically independent transmitted signals and $v^{(i)}(n)$ is the Additive White Gaussian Noise (AWGN). The noise components $v^{(i)}(n)$, $1 \leq i \leq I$, are assumed to be zero mean, independent from each other and from the transmitted signals $s_m(n)$.

The cubic terms corresponding to $m_3 = m_1$ and $m_3 = m_2$ are absent in (1) due to the fact that, for constant modulus signals, like PSK modulated signals, they have the form: $s_{m_1}(n)\ |s_{m_2}(n)|^2$, where $|s_{m_2}(n)|^2$ is a multiplicative constant absorbed by the associated channel coefficient. As a consequence, these cubic terms degenerate in linear terms. In addition, the quadratic terms are absent in (1) due to the fact that distortions generated by even-power terms produce spectral components lying outside the channel bandwidth, which can be eliminated by bandpass filters at the receiver.

The channel model (1) represents a complex-valued truncated triangular MIMO Volterra filter, the inputs of which are user indexed signals, instead of a single time indexed input as in traditional Volterra filters. It represents a generical representation of instantaneous linear-cubic polynomial mixtures.

The signals received on the I antennae, at the time instant n, can also be expressed in a compact way:

$$\mathbf{x}(n) \quad = \quad \mathbf{H}\mathbf{s}(n) + \mathbf{v}(n), \quad (2)$$

where $\mathbf{x}(n) = [x^{(1)}(n) \ldots x^{(I)}(n)]^T \in \mathbb{C}^{I\times 1}$, $\mathbf{v}(n) = [v^{(1)}(n) \ldots v^{(I)}(n)]^T \in \mathbb{C}^{I\times 1}$ and $\mathbf{H} = [\mathbf{h}^{(1)} \ldots \mathbf{h}^{(I)}]^T \in \mathbb{C}^{I\times M_V}$, the vector $\mathbf{h}^{(i)}$ $(1 \leq i \leq I)$ containing the parameters $h_{2k+1}^{(i)}(m_1, \ldots, m_{2k+1})$, $k = 0, 1$, and M_V being the number of channel coefficients of each filter $\mathbf{h}^{(i)}$ in (1). Moreover, $\mathbf{s}(n) \in \mathbb{C}^{M_V\times 1}$ is the input vector containing the linear $\{s_{m_1}(n)\}$ $(1 \leq m_1 \leq M)$ and cubic terms $\{s_{m_1}(n)s_{m_2}(n)\ s_{m_3}^*(n)\}$ $(1 \leq m_1, m_2, m_3 \leq M$, $m_1 \neq m_3$, $m_2 \neq m_3$, $m_2 \geq m_1)$. Note that $M_V = \frac{M}{2}(M^2 - M + 2)$.

3. IDENTIFIABILITY CONDITIONS FROM SOS

The proposed nonlinear BSS and channel identification method relies on the joint diagonalization of a set of spatio-temporal covariance matrices of the received signals, given by:

$$\mathbf{R}(\tau) = \mathbb{E}\left[\mathbf{x}(n+\tau)\mathbf{x}^H(n)\right] = \mathbf{H}\mathbf{C}(\tau)\mathbf{H}^H + \sigma^2\mathbf{I_I}\delta(\tau), \quad (3)$$

with

$$\mathbf{C}(\tau) = \mathbb{E}\left[\mathbf{s}(n+\tau)\mathbf{s}^H(n)\right], \quad (4)$$

where $\tau \in \Upsilon = \{\tau_1, \tau_2, ..., \tau_T\}$, the superscript H denotes the complex conjugate transpose of a matrix, $\delta(\tau)$ is the Kronecker symbol, σ^2 is the AWGN variance and $\mathbf{I_I}$ is the $I \times I$ identity matrix. If $I \geq M_V$, the noise variance σ^2 can be estimated as the mean of the $(I - M_V)$ smallest eigenvalues of $\mathbf{R}(0)$ [12], allowing the subtraction of the noise term in (3). Thus, this noise term will be omitted in the sequel.

In order to enable the application of a joint diagonalization algorithm, the matrices $\mathbf{C}(\tau)$ must be diagonal for $\tau \in \Upsilon$. The following theorem states sufficient conditions to ensure this constraint.

Theorem 1: Suppose that all the signals transmitted by the users are mutually independent and have constant moduli. The following conditions are sufficient to ensure the diagonality of the covariance matrices $\mathbf{C}(\tau)$, $\tau \in \Upsilon$:

(i). $\mathbb{E}[s_m(n)] = 0$, for all the users;

(ii). $\mathbb{E}\left[s_m^2(n)\right] = 0$, for $(M-1)$ users;

(iii). $\mathbb{E}\left[s_m^2(n+\tau)s_m(n)\right] = 0$ and $\mathbb{E}\left[s_m^2(n)s_m(n+\tau)\right] = 0$, for $(M-1)$ users, $\forall\,\tau \in \Upsilon$;

(iv). $\mathbb{E}[s_m(n+\tau)s_m(n)] = 0$, for $(M-1)$ users, $\forall\,\tau \in \Upsilon$.

The proof is omitted due to a lack of space.

The following theorem proves that some conditions of Theorem 1 are verified if all the users transmit uniformly distributed PSK signals with more than 2 symbols in the constellation.

Theorem 2: Suppose that all the users transmit uniformly distributed PSK signals with $R_m > 2$, $\forall\, m \in \{1, 2, ..., M\}$, where R_m is the number of constellation symbols of the m^{th} user. Then, conditions (i) and (ii) of Theorem 1, and conditions (iii) and (iv), for $\tau = 0$, are verified.

Proof: If $s_m(n)$, $m = 1, ..., M$, takes an equiprobable value from the set $\left\{A_m.e^{j2\pi(r-1)/R_m}; r = 1, 2, ..., \quad R_m; R_m > 2\right\}$, then we have

$$\mathbb{E}[s_m^p(n)] = \frac{A_m^p}{R_m}\sum_{r=1}^{R_m} e^{j2\pi(r-1)p/R_m} = \frac{A_m^p\left(e^{j2\pi p}-1\right)}{R_m\left(e^{j2\pi p/R_m}-1\right)}, \quad (5)$$

which is equal to zero for $p = 1, 2, 3$ and $R_m > 2$. □

4. DESIGN OF CODING SCHEMES

In this section, some modulation codes are designed to ensure that the transmitted signals satisfy the constraints listed in Theorem 1. In these modulation code schemes, the modulation makes part of the encoding process and it introduces redundancy by expanding the signal constellation. This means that a modulation memory is introduced in a controlled way with the purpose of keeping the orthogonality between nonlinear combinations of the transmitted signals.

This constitutes a new application of modulation codes, since they are used to ensure some statistical properties associated with the channel nonlinearities. Moreover, the code redundancy could also be explored in the symbol recovery process to provide Bit Error Rate (BER) improvements, by exploiting the fact that introduced redundancy imposes some constraints on the symbol transitions. This subject will be investigated in future works.

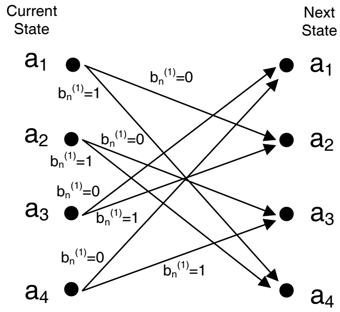

Fig. 1. Miller Code State diagram.

The modulated signals are characterized by Discrete Time Markov Chains (DTMC) with R_m states, given by the PSK symbols $a_r = \{A_m \cdot e^{j2\pi(r-1)/R_m}\}$, for $r = 1, 2, ..., R_m$, where A_m is the amplitude of the signal of the m^{th} user. The state transitions are defined by a block of k_m bits, denoted by $B_n = \{b_n^{(1)}, b_n^{(2)}, ..., b_n^{(k_m)}\}$, where $b_n^{(k)}$, for k=1, ..., k_m, is uniformly distributed over the set $\{0, 1\}$ and $2^{k_m} < R_m$. In addition, it is assumed that $b_n^{(k)}$ (k=1, ..., k_m) are mutually independent. For each of the R_m states, the block of bits B_n defines 2^{k_m} equiprobable possible transitions. Therefore, the coding imposes some restrictions on the symbol transitions. For each state, there is $\left(R_m - 2^{k_m}\right)$ not assigned transitions. The code rate of the m^{th} user is then given by (k_m/l_m), where $l_m = \log_2 R_m$.

Let us denote by $\mathbf{T} = \{T_{r_1,r_2}\}$, with $r_1,\, r_2 \in \{1, 2, ..., R_m\}$ the Transition Probability Matrix, T_{r_1,r_2} being the probability of a transition from the state r_1 to the state r_2. Note that $\sum_{r_2=1}^{R_m} T_{r_1,r_2} = 1$ and $T_{r_1,r_2} \in \{0, 1/2^{k_m}\}$. So, the matrix $\mathbf{T}$ defines which are the possible state transitions for each state.

An example of mapping from the bits B_n to the corresponding PSK symbols is illustrated in Fig. 1 for a 4-PSK signal, where $\{a_1,\, a_2,\, a_3,\, a_4\}$ are the constellation symbols (states) and $k_m = 1$. This state diagram corresponds to the run-length-limited code known as Miller Code, associated with the transition probability matrix $\mathbf{T}_{2,B}$ given in (13). The Miller Code widely used in digital magnetic recording and in Binary-PSK carrier modulation systems [10]. Similar state diagrams can be obtained for the other transition probability matrices given in the Appendix.

According to Theorem 2, if all the users transmit uniformly distributed PSK signals, then conditions of Theorem 1 are verified for $\tau = 0$. So, the following theorem proposes some constraints in the transition probability matrix $\mathbf{T}$ associated with the users in such a way that all the users transmit uniformly distributed PSK signals.

Theorem 3: Let us assume that the DTMC associated with the coding is irreducible and aperiodic. If $\sum_{r_1=1}^{R_m} T_{r_1,r_2} = 1$, for $1 \leq r_2 \leq R_m$, then, for a large number of time steps, the average fraction of time steps during which the DTMC is in the state a_{r_1} converges to $1/R_m$, for $1 \leq r_1 \leq R_m$.

Proof: The aperiodicity and irreducibility properties assure that [15]: (i) all the limiting probabilities of a DTMC exist and are positive, (ii) the stationary distribution exists and is unique, and (iii) the distribution of limiting probabilities is equal to the stationary distribution. So, the limiting probabilities $\mathbf{P} = [p_1\ p_2\ ...\ p_{R_m}]$ can be obtained by the following system of equations

$$\begin{cases} \mathbf{P}\,\mathbf{T} = \mathbf{P}, \\ \sum_{r=1}^{R_m} p_r = 1. \end{cases} \tag{6}$$

It can be easily verified that if $\sum_{r_1=1}^{R_m} T_{r_1,r_2} = 1$, then $\mathbf{P} = [1/R_m\ ...\ 1/R_m]$ is a solution of the system (6). And finally, it can be proved (the proof is omitted due to a lack of space) that if the limiting probability of a state a_{r_1} exists, then it is equal to the long-run time average spent in the state a_{r_1}, i.e. for a large number of time steps, the average fraction of time steps that the DTMC spends in the state a_{r_1} converges to the limiting probability of the state a_{r_1}. □

In the sequel, some restrictions to the transition probability matrix are developed in order that the conditions of Theorem 1 are verified for $\tau \neq 0$. Let T_{r_1,r_2}^n be the $(r_1, r_2)^{th}$ entry of $\mathbf{T}^n$. By definition, T_{r_1,r_2}^n represents the probability of being in the state a_{r_2} after n transitions, supposing that the current state is a_{r_1}. So, we may write:

$$\mathbb{E}\left[s_m^k(n+\tau)s_m^l(n)\right] = \frac{1}{R_m}\mathbf{a}_l^T\mathbf{T}^\tau\mathbf{a}_k, \tag{7}$$

where $\mathbf{a} = [a_1,\, a_2,\, ...\, a_{R_m}]^T$ and $\mathbf{a}_k = \left[a_1^k,\, a_2^k,\, ...\, a_{R_m}^k\right]^T$. Thus, the conditions (iii) and (iv) of Theorem 1 can be rewritten as:

$$\mathbf{a}^T\mathbf{T}^\tau\mathbf{a}_2 = 0, \quad \mathbf{a}_2^T\mathbf{T}^\tau\mathbf{a} = 0 \quad \text{and} \quad \mathbf{a}^T\mathbf{T}^\tau\mathbf{a} = 0. \tag{8}$$

The results found in this section may be summarized in the following corollary.

Corollary 1: If the following conditions hold for all the users:

(i). the transition probability matrix corresponds to an irreducible and aperiodic DTMC;

(ii). $\sum_{r_1=1}^{R_m} T_{r_1,r_2} = 1, \forall\, r_2\,,\, 1 \leq r_2 \leq R_m$;

and, in addition, equations (8) hold for $(M-1)$ users $\forall\,\tau \in \Upsilon$, then all the conditions of Theorem 1 are satisfied and, therefore, the covariance matrix $\mathbf{C}(\tau)$ is diagonal $\forall\,\tau \in \Upsilon$.

It should be highlighted that equations (8) only depend on the matrix $\mathbf{T}$ and the constellation order. That means that transition probability matrices can be a priori designed to verify these equations. In the Appendix, some examples of such matrices verifying these constraints are listed, with the corresponding admissible delays.

5. CHANNEL ESTIMATION ALGORITHM

Provided that the conditions of Corollary 1 hold, the channel $\mathbf{H}$ can be estimated from the set of covariance matrices $\mathbf{R}(\tau)$ by using a joint diagonalization algorithm. The proposed method can then be viewed as an extension of the SOBI algorithm [12] to nonlinear channels. The uniqueness of the joint diagonalizer based estimator is given by the following theorem [12]. The covariation matrix $\mathbf{C}(0)$ is assumed to be normalized, i.e. $\mathbf{C}(0) = \mathbf{I}_{M_V}$.

Theorem 4: Let $\mathcal{B} = \{\mathbf{B}_1, ..., \mathbf{B}_T\}$ be a set of T matrices $M_V \times M_V$ such that $\mathbf{B}_t = \mathbf{M}\mathbf{C}_t\mathbf{M}^H$, for $t = 1, ..., T$, where $\mathbf{M} \in \mathbb{C}^{M_V \times M_V}$ is a unitary matrix and $\mathbf{C}_t \in \mathbb{C}^{M_V \times M_V}$, for $t = 1, ..., T$, are diagonal matrices, the elements of which are denoted by $c_t(r) = [\mathbf{C}_t]_{r,r}$. If

$$\exists\ t \in \{1, ..., T\} \text{ such that } c_t(r_1) \neq c_t(r_2), \quad \forall\, r_1, r_2 \in \{1, ..., M_V\}, \text{with } r_1 \neq r_2, \tag{9}$$

then any joint diagonalizer of $\mathcal{B}$ is equal to $\Pi\Lambda\mathbf{M}$, where Λ is a diagonal matrix and Π a permutation matrix.

In the Appendix, some examples of configurations of transition probability matrices for 2 users are given, verifying the uniqueness condition (9). The estimation algorithm can be summarized as follows:

(i). Calculate the whitening matrix $\mathbf{U}$ from:

$$\mathbf{U} = \left[\lambda_1^{-\frac{1}{2}}\mathbf{w}_1 \cdots \lambda_{M_V}^{-\frac{1}{2}}\mathbf{w}_{M_V}\right]^H, \tag{10}$$

where $\{\lambda_k\}_{k=1}^{M_V}$ are the M_V largest eigenvalues of $\hat{\mathbf{R}}(0)$ and $\{\mathbf{w}_k\}_{k=1}^{M_V}$, the corresponding eigenvectors, $\hat{\mathbf{R}}(0)$ being the sampled estimate of $\mathbf{R}(0)$. We have considered that the estimated noise variance $\hat{\sigma}^2$ was already subtracted from $\hat{\mathbf{R}}(0)$, as mentioned earlier.

(ii). Calculate the following set of prewhitened matrices: $\hat{\mathbf{R}}_W(\tau) = \mathbf{U}\hat{\mathbf{R}}(\tau)\mathbf{U}^H$, for $\tau \in \Upsilon$, where $\hat{\mathbf{R}}(\tau)$ are the sampled covariance matrices.

(iii). Obtain a unitary matrix $\hat{\mathbf{M}}$ as the joint diagonalizer of the matrices $\hat{\mathbf{R}}_W(\tau)$, for $\tau \in \Upsilon$.

(iv). Estimate the channel matrix as $\hat{\mathbf{H}} = \mathbf{U}^\dagger\hat{\mathbf{M}}$ and the transmitted signals as $\hat{\mathbf{s}}(n) = \hat{\mathbf{M}}^H\mathbf{U}\mathbf{x}(n)$, where $(\cdot)^\dagger$ denotes the matrix pseudo-inverse.

Step (iii) of the method is carried out by using the joint diagonalization algorithm of [11]. Note that the joint diagonalization estimator does not assume the knowledge of the covariance matrices of the sources $\mathbf{C}(\tau)$ and that it requires $I \geq M_V$.

6. SIMULATION RESULTS

In this section, the proposed nonlinear BSS and channel identification method is evaluated by means of computer simulations with an uplink channel of a Radio Over Fiber (ROF) multiuser communication system. The linear wireless interface is modeled as a memoryless multiuser channel. The I antennae are half-wavelength spaced and the transmitted signals are narrowband with respect to the array aperture. Moreover, the propagation scenario is characterized by two users, the angles of arrival of which are $30°$ and $70°$, respectively. The E/O conversion in each antenna is modelled by the linear-cubic polynomial $c_1x + c_3|x|^2x$, with $c_1 = -0.291$, $c_3 = 1.078$ (see [4]). The used modulation is 4-PSK and all the results were obtained via Monte Carlo simulations using $N_R = 200$ independent data realizations.

Fig. 2 shows the Normalized Mean Squared Error (NMSE) of the estimated transmitted symbols versus SNR for the configurations of transition probability matrices given in Table 1 of the Appendix, for $M = 2$, $T = 4$, $N_s = 3000$ and $I = 4$, where N_s is the length of the data block used for the moment estimation. The NMSE of the transmitted signals is defined as:

$$e_S(n) = \frac{1}{N_R}\sum_{j=1}^{N_R}\frac{\sum_{n=1}^{N_s} \| \mathbf{s}_j(n) - \hat{\mathbf{s}}_j(n) \|_2^2}{\sum_{n=1}^{N_s} \| \mathbf{s}_j(n) \|_2^2}, \tag{11}$$

where $\mathbf{s}_j(n)$ and $\hat{\mathbf{s}}_j(n)$ represents respectively the transmitted signals and the estimated transmitted signals at the j^{th} Monte Carlo simulation and $\| \cdot \|_2$ denotes the l^2 norm. The performance of our technique is compared with that of the Minimum Mean Square Error

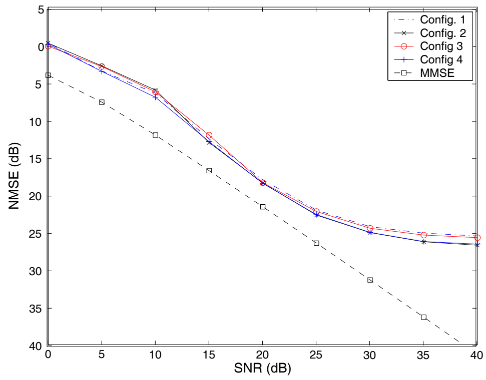

Fig. 2. NMSE versus SNR for the configurations given in Table 1 - $M = 2$, $T = 4$, $N_s = 3000$ and $I = 4$.

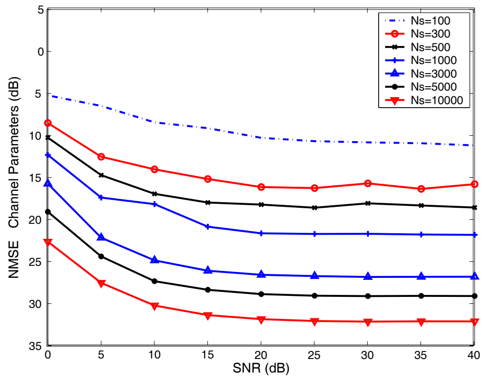

Fig. 3. NMSE of the channel parameters versus SNR for various values of N_s using Config. 2, $M = 2$, $T = 4$ and $I = 10$.

(MMSE) receiver [10] assuming perfect knowledge of the channel. Note that the tested configurations provide quite similar NMSE performances, not far from that of the MMSE receiver when the SNR is smaller than 25 dB.

Fig. 3 shows the NMSE of the channel parameters versus SNR for various values of N_s, where the NMSE of the channel parameters is defined as: $e_H = 1/N_R\left(\sum_{l=1}^{N_R} \| \mathbf{H} - \hat{\mathbf{H}}_l \|_F^2\right) / (\| \mathbf{H} \|_F^2)$, where $\| \cdot \|_F$ denotes the *Frobenius norm*. In this case, we have used Config. 2 of Table 1, $M = 2$, $T = 4$ and $I = 10$. It can be seen that the quality of the channel estimation can be considerably improved by increasing the length of the data block. This result indicates that the errors in the estimation of the covariance matrices constitute one of the main sources of performance degradation. In fact, if the theoretical values of the covariance matrices $\mathbf{R}(\tau)$ are used, the estimation algorithm provides a very low NMSE for the estimated channel parameters, limited by the machine precision.

Fig. 4 shows the Symbol Error Rate (SER) versus SNR provided by the proposed technique and the MMSE receiver assuming perfect knowledge of the channel, using Config. 2 of Table 1, $M = 2$, $T = 4$ and $N_s = 3000$, for $I = 4$ and $I = 8$. We remark that the performance of the joint diagonalization algorithm is close to that of the MMSE solution if the number of antennae is increased.

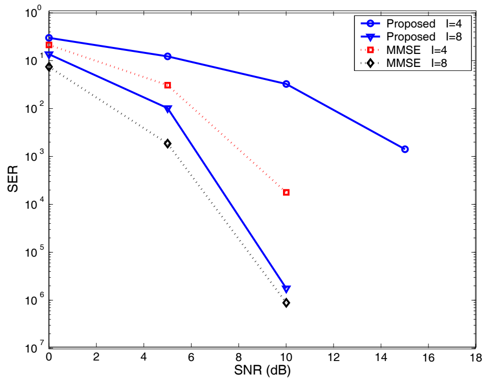

Fig. 4. SER versus SNR using Config. 2, $M = 2$, $T = 4$ and $N_s = 3000$, for $I = 4$ and $I = 8$.

7. CONCLUSION

In this paper, the problem of BSS and channel identification for nonlinear multiuser communication channels has been solved in assuming that the channel is modeled as a multiuser MIMO Volterra filter. The proposed method is based on the joint diagonalization of a set of spatio-temporal covariance matrices. We have made use of modulation codes to ensure the orthogonality of nonlinear interfering terms for different time delays, which constitutes a new application of modulation codes. The proposed technique was tested by means of computer simulations with an uplink channel of a multiuser ROF communication system. In future works, other estimation algorithms will be tested and the impact of the modulation codes on the bit recovery will be investigated.

A. APPENDIX - CONFIGURATIONS OF TRANSITION PROBABILITY MATRICES

As pointed out, the transition probability matrices can be a priori designed to verify the conditions of Corollary 1. In the following, we present some examples of such matrices corresponding to $1/2$-rate codes for 4-PSK signals.

It can be proved by mathematical induction that the following matrices:

$$\mathbf{T}_{1,A} = 0.5 \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \mathbf{T}_{1,B} = 0.5 \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}, \quad (12)$$

verify all the conditions of Corollary 1 $\forall \tau \in \mathbb{I}$. In this case $\mathbf{a} = [1 \; j \; -1 \; -j]^T$. In addition,

$$\mathbf{T}_{2,A} = 0.5 \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \mathbf{T}_{2,B} = 0.5 \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}. \quad (13)$$

also verify conditions (i) and (ii) of Corollary 1.

The identifiability test indicated in Theorem 4 only depends on the covariance matrices $\mathbf{C}(\tau)$, for $\tau \in \Upsilon$, which can be calculated from the transition probability matrices by using (7) and:

$$\mathbb{E}\left[s_m^k(n+\tau){s^{l}}_m^*(n)\right] = \frac{1}{R_m}\mathbf{a}_l^H \mathbf{T}^\tau \mathbf{a}_k. \quad (14)$$

This means that, if the matrices $\mathbf{T}$ of the users are known, the identifiability test can be carried out. Thus, it can be verified that the configurations of transition probability matrices for 2 users given in Table 1 verify the condition of Theorem 4. The corresponding admissible delays are $\Upsilon = \{0, 1, ..., T-1\}$, with $T \geq 2$.

Table 1. Configurations of transition probability matrices for M=2.

Config.	User 1	User 2
1	$\mathbf{T}_{1,A}$	$\mathbf{T}_{2,B}$
2	$\mathbf{T}_{1,A}$	$\mathbf{T}_{2,A}$
3	$\mathbf{T}_{1,B}$	$\mathbf{T}_{2,B}$
4	$\mathbf{T}_{1,B}$	$\mathbf{T}_{2,A}$

B. REFERENCES

[1] A. Ziehe, M. Kawanabe, S. Harmeling, and K.-R. Muller, "Blind separation of post-nonlinear mixtures using linearizing transformations and temporal decorrelation," *Journal of Machine Learning Research*, vol. 4, no. 7-8, pp. 1319–1338, 2003.

[2] 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 Computer Eng.*, May 1-4, 2005, pp. 1–4.

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

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

[5] W.I. Way, "Optical fiber based microcellular systems. An overview," *IEICE Trans. Commun.*, vol. E76-B, no. 9, pp. 1091–1102, Sept. 1993.

[6] C. Jutten, M. Babaie-Zadeh, and S. Hosseini, "Three easy ways for separating nonlinear mixtures?," *Signal Processing*, vol. 84, no. 2, pp. 217–229, Feb. 2004.

[7] A. Taleb and C. Jutten, "Source separation in post-nonlinear mixtures," *IEEE Trans. on Sig. Proc.*, vol. 47, no. 10, pp. 2807–2820, Sept. 1999.

[8] S. Harmeling, A. Ziehe, M. Kawanabe, and K.-R. Muller, "Kernel-based nonlinear blind source separation," *Neural Comp.*, vol. 15, no. 5, pp. 1089–1124, May 2003.

[9] 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.

[10] J. G. Proakis, *Digital Communications*, McGraw-Hill, 4*rd* edition, 2001.

[11] J.-F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM Journal on Matrix Anal. and Applicat.*, vol. 17, no. 1, pp. 161–164, Jan. 1996.

[12] 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.

[13] I. B. Djordjevic, B. Vasic, and V. S. Rao, "Rate 2/3 modulation code for suppression of intrachannel nonlinear effects in high-speed optical transmission," *IEE Proc.-Optoelectron.*, vol. 153, no. 2, pp. 87–92, Apr 2006.

[14] Y. Rong, S. A. Vorobyov, A. B. Gershman, and N. D. Sidiropoulos, "Blind spatial signature estimation via time-varying user power loading and parallel factor analysis," *IEEE Trans. on Signal Proc.*, vol. 53, no. 5, pp. 1697–1709, May 2005.

[15] O. Haggstrom, *Finite Markov Chains and Algorithmic Applications*, Cambridge University Press, 2002.