# Closed-Form Blind Channel Identification and Source Separation in SDMA Systems Through Correlative Coding

João M. F. Xavier, *Student Member, IEEE,* Victor A. N. Barroso, *Member, IEEE,* and José M. F. Moura, *Fellow, IEEE*

***Abstract*— We address the problem of blind identification of multiuser multiple-input multiple-output (MIMO) finite-impulse response (FIR) digital systems. This problem arises in spatial division multiple access (SDMA) architectures for wireless communications. We present a *closed-form,* i.e., noniterative, consistent estimator for the MIMO channel based only on *second-order* statistics. To obtain this closed form we introduce spectral/correlation asymmetry between the sources by filtering each source output with adequate correlative filters. Our algorithm uses the closed form MIMO channel estimate to cancel the intersymbol interference (ISI) due to multipath propagation and to discriminate between the sources at the wireless base station receiver. Simulation results show that, for single-user channels, this technique yields better channel estimates in terms of mean-square error (MSE) and better probability of error than a well-known alternative method. Finally, we illustrate its performance for MIMO channels in the context of the global system for mobile communications (GSM) system.**

***Index Terms*— Blind channel identification, global system for mobile communications (GSM), intersymbol interference, MIMO, multipath propagation, SIMO, wireless digital communications.**

## I. Introduction

We study blind channel identification in the context of digital multiple-input/multiple-output (MIMO) systems. This problem arises naturally in spatial division multiple access (SDMA) architectures for wireless communications. In SDMA, multiple users transmit simultaneously in time using the same frequency narrowband channel, thus increasing the cellular capacity without the need for additional RF spectrum [1]. In code division multiple access (CDMA) systems, the users also transmit simultaneously in time but each user's signal is spread over a larger frequency region.

Tong *et al.* [2] presented a major breakthrough for the blind identification of digital single-input/multiple-output (SIMO) systems, i.e., multichannel filters driven by a single digital source. They assume that the array at the base station cell site receives an *oversampled* weighted linear superposition of the emitted signal. Their algorithm exploits the cyclostationary property associated with virtual channels created by temporal and/or spatial oversampling of digital communications signals. These authors derived a *closed-form* (noniterative) asymptotically exact estimator, i.e., a consistent estimator, for the SIMO channel which relies only on *second-order* statistics of the received signals. When compared to other well-known blind equalization methods, their approach exhibits two main advantages, though at the expense of added computational complexity.

- Because it is a closed-form algorithm, it does not suffer from the irregular convergence properties [5] of most adaptive iterative methods, such as the constant modulus (CM) algorithms [3], [4]. This drawback is due to the existence of many local minima attractors in the cost functions. Also, the overall performance of gradient-based optimization techniques strongly depends on the learning rate parameter, usually chosen by a trial-and-error procedure.
- Since only second-order statistics are involved, the approach in [2] is feasible with short data packets. This contrasts with high-order cumulants-based methods [6]–[8] which require a significant amount of data in order to attain equivalent results in terms of mean-square error (MSE). In wireless radio communications, this is an important issue since only a few data samples are available for processing during the time interval along which the channel can be assumed time invariant.

After [2], many other closed-form approaches have been proposed for the blind identification of digital SIMO-finite impulse response (FIR) systems. Xu *et al.* [9] model the input process as a deterministic signal and exploit special algebraic properties of oversampled systems. Since no statistical model is assumed, the algorithm works with a very small number of data samples. Moulines *et al.* [10] take advantage of the orthogonality between the signal and noise subspaces spanned by the data covariance matrix. They exploit the structure of the filtering matrix connecting the transmitted digital sequence to the oversampled array outputs in order to significantly reduce the number of estimated coefficients. The linear prediction approach, introduced by Slock [11] and generalized by Abed-Meraim *et al.* [12] and Gesbert *et al.* [13], exhibits robustness

Manuscript received August 26, 1997; revised April 13, 1998. This work was supported in part by the NATO Collaborative Research Grants Programme, SA. 5-02-05 (CRG.971184) 1202/97/JARC-501. The work of the J. M. F. Xavier and V. A. N. Barroso was also supported by FEDER and PRAXIS XXI, under contract PRAXIS/3/3.1/TPR/23/94.

J. M. F. Xavier and V. A. N. Barroso are with the Instituto Superior Técnico–Instituto de Sistemas e Robótica, 1096 Lisboa Codex, Portugal (e-mail: jxavier@isr.ist.utl.pt; vab.@isr.ist.utl.pt).

J. M. F. Moura is with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: moura@ece.cmu.edu).

Publisher Item Identifier S 0733-8716(98)07898-6.

to channel order overestimation and is also second-order statistics-based.

In addition to the identification of the channel, the task of the SDMA receiver includes:

1) suppressing the intersymbol interference (ISI) induced by the multipath propagation;
2) discriminating the signals transmitted by the sources present in the cell.

These two goals may be accomplished once the time/space channel is identified, although some approaches estimate directly the transmitted digital sequences, avoiding the channel estimation. The iterative least squares with projection (ILSP) and iterative least squares with enumeration (ILSE) algorithms introduced by Talwar *et al.* [14], as well as the algorithms in [15]–[19], exploit the finite alphabet property of the binary shift keying (BSK), phase shift keying (PSK), and quadrature amplitude modulation (QAM) digital modulation formats. Being iterative algorithms, global convergence is not guaranteed *a priori* and several runnings starting from different initializations may be required. For some of these algorithms, this situation can occur even at high values of the signal-to-noise ratio (SNR). The approach in [20] uses the geometrical properties induced by the binary signal constellation but tends to result in an identification scheme more sensitive to the noise level. Papadias *et al.* [21] extended the CM concept to the multiuser situation. Being a gradient-based algorithm, it shares the convergence and implementation problems already discussed for the single-source case.

Liu *et al.* [9] and Abed-Meraim *et al.* [22] considered the extension of closed-form SIMO system identification strategies to the MIMO situation. Their algorithms solve only the problem up to a unknown $P \times P$ residual instantaneous mixing matrix, where $P$ is the number of users. The general consensus is that SIMO methods can perform blind equalization in a MIMO environment eliminating the ISI [point 1) above], but are unable to solve the second issue, i.e., the sources' mixture problem. With no additional information, an iterative sources' separation algorithm such as the ILSP or algorithms based on high-order statistics [23] must be used. Ding [24] has also proposed a closed-form procedure which relies on previous knowledge of the pulse-shaping filters. Ding's method is a semiblind procedure since prior knowledge of the shaping filter used by each source is equivalent to knowing part of the MIMO channel transfer function.

In this paper, we present a *closed-form second-order* statistics-based blind channel identification algorithm for MIMO channels that yields asymptotically exact estimates of the FIR communication MIMO channels. The method extends that of Tong *et al.*, which applies only to SIMO channels. Our algorithm cancels the ISI and separates the sources, solving problems 1) and 2) above.

We obtain a closed-form blind estimator for the MIMO channel, and separate the sources, by introducing *spectral asymmetry* between the sources through correlative coding. Correlative coding or partial-response signaling schemes are used in communication systems to attain the Nyquist capacity upper bound for ideal bandlimited channels [25]. Basically, at the transmitter, a correlative filter enables the overall system impulse response to exhibit smoother rolloffs than Nyquist pulse shaping. This results in communication systems less sensitive to symbol clock timing errors. Since fractionally spaced equalization (FSE), i.e., equalization based on signals sampled faster than the data rate, exhibits robustness to symbol clock jitter by itself, we use partial-response signaling with a different objective, namely, to assigning distinct spectral signatures to each user. As we show in the paper, the spectral asymmetry thus introduced leads to a closed-form second-order-based solution to the blind identification problem.

Our algorithm exhibits several relevant properties. As in [2], the identification is exact if the second-order statistical description of the output process is known exactly. In practice, these statistics are consistently estimated from the available set of observations. It is a full blind identification algorithm that assumes no previous knowledge about the pulse-shaping filters nor any other channel component. Hence, it is robust with respect to any eventual mismatches. The penalty incurred by our approach is not in terms of additional power or bandwidth consumption nor synchronization between the sources, but increased computational effort and an extra decoding delay due to the controlled ISI introduced at the transmitter. This ISI is removed at the receiver using maximum-likelihood sequence estimation (MLSE) or Viterbi decoding [26]. In this paper, we do not address the impact of our results in CDMA systems.

The paper is organized as follows. Section II introduces the signal model and states the blind identification problem. Section III establishes the framework of correlative coding for the separation of the sources. It presents a theorem proved in Appendix A which establishes that, under this framework, the MIMO channel is uniquely determined (up to a phase offset for each user) from the *second-order* statistics of its outputs. Section IV describes the blind channel identification algorithm. We discuss a computationally simpler variation of this algorithm which is adequate in high SNR environments. We present the technique that separates the sources and recovers the transmitted sequences. Section V considers the performance of our method. For single-users, the algorithm outperforms that in [2]. For multiusers, since there is no alternative closed-form, second-order-based blind scheme, we illustrate with simulation studies its performance in the context of the GSM system. Section VI concludes the paper.

## II. Problem Formulation

*Notation:* Throughout the paper, we adopt the following notation. $\mathbb{Z}, \mathbb{R}$, and $\mathbb{C}$ denote, respectively, the set of integer, real, and complex numbers. Matrices (capital letter) and vectors are in boldface type. The set of all $m \times n$ matrices with complex entries is $\mathbb{C}^{n \times m}$, and $\mathbb{C}^n$ denotes the set of $n$-dimensional complex column vectors. For an arbitrary matrix $\boldsymbol{A}$, its $(i,j)$ entry is denoted by $A(i,j)$, its $i$th row by $\boldsymbol{A}(i,\cdot)$, and its $j$h column by $\boldsymbol{A}(\cdot,j)$. $\mathcal{R}(\boldsymbol{A})$ and $\mathcal{N}(\boldsymbol{A})$ denote, respectively, the range space and the null space (kernel) of $\boldsymbol{A}$. The notations $(\cdot)^T, (\cdot)^*, (\cdot)^H, (\cdot)^\#$, and $\mathrm{tr}(\cdot)$ stand for transpose, complex conjugate, Hermitian, the Moore-Penrose pseudoinverse, and the trace operator, respectively.

The symbols $\boldsymbol{I}_n, \boldsymbol{B}_n$, and $\boldsymbol{K}_n$ stand for the $n \times n$ identity, backward identity (ones in the antidiagonal), and forward-shift (ones in the first lower diagonal) matrices, respectively. The symbol $\boldsymbol{e}_{i;n}$ represents the $i$th column of the $n \times n$ identity matrix $\boldsymbol{I}_n$. Diagonal matrices with entries $\lambda_1, \cdots, \lambda_N$ are represented by $\mathrm{diag}[\lambda_1, \cdots, \lambda_N]$. For a zero-mean wide sense stationary scalar process $x(k)$, the autocorrelation function is given by $r_x(l) = \mathrm{E}\{x(k)x(k-l)^*\}$, and for a vector process $\boldsymbol{x}(k)$, the autocorrelation matrices of lag $l$ are defined as $\boldsymbol{R}_{\boldsymbol{x}}(l) = \mathrm{E}\{\boldsymbol{x}(k)\boldsymbol{x}(k-l)^H\}$. The $L_2$-norm of a finite energy continuous-time signal $g(t)$ supported in a region $I \subset \mathbb{R}$ is

$$\|g\|_2 = \left( \int_I |g(t)|^2 \, dt \right)^{1/2}. \tag{1}$$

For a discrete-time signal $g(l)$ supported in $I \subset \mathbb{Z}$, we use the $l_2$-norm

$$\|g\|_2 = \left( \sum_{l \in I} |g(l)|^2 \right)^{1/2}. \tag{2}$$

Additional notation is introduced as needed.

*MIMO System: Problem Statement:* Consider a generic $P$-input/$N$-output causal discrete-time noisy linear time-invariant (LTI) digital MIMO system described by the convolution equation

$$\boldsymbol{x}(k) = \sum_{p=1}^{P} \sum_{l=0}^{L_p - 1} \boldsymbol{h}_p(l) s_p(k-l) + \boldsymbol{w}(k) \tag{3}$$

where $\boldsymbol{x}(k) \in \mathbb{C}^N$ is a column vector of system outputs, $\{\boldsymbol{h}_p(l)\colon\ l = 0, 1, \cdots, L_p - 1\}$ is the FIR associated to the $p$th user ($\boldsymbol{h}_p(l) \in \mathbb{C}^N$), $L_p$ is the length of the impulse response corresponding to the $p$th user, $s_p(k) \in \mathbb{C}$ is the signal transmitted by the $p$th user, and $\boldsymbol{w}(k) \in \mathbb{C}^N$ denotes additive noise. Equation (3) can be compactly rewritten as

$$\boldsymbol{x}(k) = \sum_{p=1}^{P} \boldsymbol{H}_p \boldsymbol{s}_p(k) + \boldsymbol{w}(k) = \boldsymbol{H}\boldsymbol{s}(k) + \boldsymbol{w}(k). \tag{4}$$

Here, the $N \times L_p$ matrix $\boldsymbol{H}_p \equiv [\boldsymbol{h}_p(0) \cdots \boldsymbol{h}_p(L_p - 1)]$, and $\boldsymbol{H} \equiv [\boldsymbol{H}_1 \cdots \boldsymbol{H}_P]$ is the $N \times L$ *channel convolution matrix*. $L \equiv L_1 + \cdots + L_P$ is the overall order of the system. The vector $\boldsymbol{s}(k)$ is obtained by stacking the $P$ vectors $\boldsymbol{s}_p(k) \equiv [s_p(k) \cdots s_p(k - L_p + 1)]^T$. The input signals $s_p(k)$ are taken from a finite modulation-dependent alphabet $\mathcal{A}_p \subset \mathbb{C}$. In SDMA communications, (4) models the complex baseband transfer function between the information symbols $s_p(k)$ emitted by the sources and the oversampled array outputs (see [2], [10], [19], and [24] for details); the matrix $\boldsymbol{H}$ is fully parameterized by the samples of the equivalent *composite* transmitter/propagation/receiver FIR channels employed by the users; and $\boldsymbol{w}[k]$ is sensor noise.

We study the blind identification of $\boldsymbol{H}$ under the following assumptions:

A1) The number $P$ of users is known and the $N \times L$ channel matrix $\boldsymbol{H}$ is full-rank with $N \geq L$. In practice, this assumption is satisfied by properly dimensioning certain physical parameters of the communication system such as the number of antennas and/or the oversampling factor (see [2], [19], and [22] for discussion of these identifiability/implementation issues).

A2) The sources are uncorrelated and the noise process $\boldsymbol{w}(k)$ is zero-mean, wide sense stationary, and statistically independent of $\boldsymbol{s}(k)$, with known autocorrelation matrices $\boldsymbol{R}_{\boldsymbol{w}}(l)$.

## III. Spectral Diversity: Correlative Filters

Usually, the data sequences generated by the sources are assumed to be zero-mean white random signals. For instance, in most digital communications systems, those sequences consist of independent and identically distributed (i.i.d.) symbols extracted from a given constellation set, and thus their respective power spectral densities exhibit a similar flat pattern. That is, even if the users exhibit diversity in the time and spatial domains (in the sense of transmitting different time sequences at distinct cell sites), they tend to remain indistinguishable from that *statistical* domain point of view. The key idea in our approach is to break this symmetry condition by judiciously coloring the information sequences prior to transmission, thus assigning distinct spectral patterns to the random processes transmitted by the sources. This will permit the finding of a closed-form solution for the identification problem considered here.

*Spectral Diversity:* To introduce this additional degree of diversity, we pass the white sequence generated by the $p$th user, say $a_p(k) \in \mathcal{A}_p$, through a correlative filter with impulse response

$$c_p(k) = c_p(0)\delta(k) + c_p(L_c)\delta(k - L_c) \tag{5}$$

and transmit the colored sequence $s_p(k) \equiv (c_p \odot a_p)(k)$. Here, $\delta(k)$ is the discrete-time impulse signal, the symbol $\odot$ denotes the linear convolution operator, and the filter coefficients $\{c_p(0), c_p(L_c)\} \subset \mathbb{C}$ are variables to be designed. For mathematical convenience, the memory of the filters, i.e., $L_c$ in (5), is the same for all users, and is greater than or equal to the memory of any FIR channel in the MIMO system, i.e., $L_c \geq \max\{L_1, L_2, \cdots, L_P\}$. The delay spreads $L_p$ are unknown *a priori* by the receiver. In many application scenarios, it is possible to adequately overestimate these parameters on the basis of previous field experiments. For example, in the GSM mobile system, a typical multipath channel profile is available for several environments (urban, hilly terrain, etc.) [27] from which the maximum expected degree of ISI is easily inferred. With this correlative preprocessing, the autocorrelation function of the filtered process is

$$r_{s_p}(l) = \eta_p^* \delta(l + L_c) + \sigma_p^2 \delta(l) + \eta_p \delta(l - L_c) \tag{6}$$

where $\sigma_p^2 \equiv |c_p(0)|^2 + |c_p(L_c)|^2$ and $\eta_p \equiv c_p(0)^* c_p(L_c)$ denote, respectively, the power and the new correlation peak appearing at indexes $\pm L_c$. Notice that we have assumed that the input sequences of the correlative filters are white with unit power. This entails no loss of generality since $\boldsymbol{H}$ absorbs any multiplicative factor. Also, for further reference, the autocorrelation matrices of the process $\boldsymbol{s}(k)$ in (4) are, for

$l \geq L_c$, given by

$$\boldsymbol{R_s}(0) = \mathrm{diag}\left[\sigma_1^2 \boldsymbol{I}_{L_1}, \sigma_2^2 \boldsymbol{I}_{L_2}, \cdots, \sigma_P^2 \boldsymbol{I}_{L_P}\right]$$
$$\boldsymbol{R_s}(l) = \mathrm{diag}\left[\eta_1 \boldsymbol{K}_{L_1}^{l-L_c}, \eta_2 \boldsymbol{K}_{L_2}^{l-L_c}, \cdots \eta_P \boldsymbol{K}_{L_P}^{l-L_c}\right]. \quad (7)$$

*Correlative Filters:* We now consider the design of the correlative filters. The goal is to induce distinct spectral characteristics in the filtered processes without increasing the transmitted power. Define the power spectral density (psd) associated to the $p$th user as

$$\Gamma_{s_p}(f) = \mathcal{F}\{r_{s_p}(l)\} = \sum_{l=-\infty}^{+\infty} r_{s_p}(l) e^{-j2\pi f l}, \qquad f \in I = (-1/2, 1/2)$$

where $\mathcal{F}\{\cdot\}$ denotes the Fourier operator. We design the correlative filters by the following optimization problem: find

$$[\hat{c}_1(0), \hat{c}_1(L_c), \cdots, \hat{c}_P(0), \hat{c}_P(L_c)]^T = \underset{[c_1(0), c_1(L_c), \cdots, c_P(0), c_P(L_c)]^T \in \mathbb{C}^{2P}}{\mathrm{argmax}} \min_{p \neq q} \|\Gamma_{s_p} - \Gamma_{s_q}\|_2^2 \quad (8)$$

subject to the power constraint $\mathrm{E}\{|s_p(k)|^2\} = 1$. The filters in (8) maximize the distance between the users closest psd's. Since the Fourier operator $\mathcal{F}\{\cdot\}$ is unitary, $\|\Gamma_{s_p} - \Gamma_{s_q}\|_2 = \|r_{s_p} - r_{s_q}\|_2$; then, using (6) in (8), we can restate the optimization problem as: subject to $\sigma_p^2 = 1$, find

$$[\hat{c}_1(0), \hat{c}_1(L_c), \cdots, \hat{c}_P(0), \hat{c}_P(L_c)]^T = \underset{[c_1(0), c_1(L_c), \cdots, c_P(0), c_P(L_c)]^T \in \mathbb{C}^{2P}}{\mathrm{argmax}} \cdot \min_{p \neq q} |\sigma_p^2 - \sigma_q^2|^2 + |\eta_p - \eta_q|^2. \quad (9)$$

By symmetry, it is readily seen that a possible solution to this optimization is given by

$$c_p(0) = 1/\sqrt{2}$$
$$c_p(L_c) = \frac{1}{\sqrt{2}} e^{j2\pi(p-1)/P}.$$

This solution distributes the correlation coefficients $\eta_p = c_p(0)^* c_p(L_c)$ in (6) uniformly around the circle of radius $r = 1/2$ in the complex plane.

*No Mixing:* We state our last assumption.

A3) The $p$th user correlates its zero-mean unit-power symbols $a_p(k) \in \mathcal{A}_p$ so that (6) and (7) hold. The finite alphabets and the correlative filters are known to the receiver.

We now formulate an identifiability theorem based on the framework established by A1)–A3). It asserts that, in a certain sense, MIMO systems are uniquely defined by this set of conditions.

*Theorem 1:* Consider the signal model in (4) and suppose that A1)–A3) are satisfied. Then, each users' convolution matrix $\boldsymbol{H}_p$ is uniquely determined up to a phase factor by the output autocorrelation matrices $\boldsymbol{R_x}(0), \boldsymbol{R_x}(L_c)$, and $\boldsymbol{R_x}(L_c+1)$. In other words, if $\tilde{\boldsymbol{H}} = [\tilde{\boldsymbol{H}}_1\ \tilde{\boldsymbol{H}}_2\ \cdots\ \tilde{\boldsymbol{H}}_P]$ is any channel matrix inducing these same statistics, then

$$\tilde{\boldsymbol{H}} = \boldsymbol{H\Theta}, \boldsymbol{\Theta} = \mathrm{diag}[e^{j\theta_1} \boldsymbol{I}_{L-1}, e^{j\theta_2} \boldsymbol{I}_{L_2}, \cdots e^{j\theta_P} \boldsymbol{I}_{L_P}]$$

or, equivalently, $\tilde{\boldsymbol{H}}_p = \boldsymbol{H}_p e^{j\theta_p}$.

*Proof:* See the Appendix.

The key point to emphasize is that the theorem establishes the absence of a residual mixing matrix coupling the sources together. Here, the residual ambiguity matrix $\boldsymbol{\Theta}$ is *diagonal*. Therefore, not only are the sources resolved, but also, for each user, the corresponding FIR coefficients in $\boldsymbol{H}_p(l)$ are obtained. As it is well known, the remaining phase factors cannot be resolved by just using second-order statistics, but they are easily compensated at the receiver.

We could attempt applying Theorem 1 as the basis of a moments matching identification strategy [28] by computing a full-rank $N \times L$ matrix $\tilde{\boldsymbol{H}}$ such that the matrix equalities

$$\boldsymbol{H R_s}(0) \boldsymbol{H}^H = \tilde{\boldsymbol{H}} \boldsymbol{R_s}(0) \tilde{\boldsymbol{H}}^H \quad (10)$$
$$\boldsymbol{H R_s}(L_c) \boldsymbol{H}^H = \tilde{\boldsymbol{H}} \boldsymbol{R_s}(L_c) \tilde{\boldsymbol{H}}^H \quad (11)$$
$$\boldsymbol{H R_s}(L_c+1) \boldsymbol{H}^H = \tilde{\boldsymbol{H}} \boldsymbol{R_s}(L_c+1) \tilde{\boldsymbol{H}}^H \quad (12)$$

are satisfied. The quantities on the left-hand side of (10)–(12) can be estimated at the receiver by subtracting the noise correlation matrix $\boldsymbol{R_w}(l)$ in

$$\boldsymbol{R_x}(l) = \boldsymbol{H R_s}(l) \boldsymbol{H}^H + \boldsymbol{R_w}(l), \qquad l = 0, L_c, L_{c+1}$$

which can be estimated from the array output data using the sample mean operator. Two main reasons prevent this from being straightforward.

- The users' delay spreads $L_p$ are not known *a priori* and thus the receiver is unable to construct the necessary sources' autocorrelation matrices $\boldsymbol{R_s}(l)$ [(7)]. One exception would be the case of digital communication scenarios where the bandwidth of the information-bearing signals is small when compared with the coherence bandwidth of the propagation channels [25]. That is, the maximum multipath delay is a negligible fraction of the sources' signaling period, and thus the ISI effect is virtually null. In these situations, $L$ always coincides with the number of users $P$ since $L_p = 1$ for all $p$. Examples of such coherent multipath environments include microcell wireless systems or the Advanced Mobile Phone Systems (AMPS) [29].
- Even if the delay spreads $L_p$ are known, it remains to extract a least-squares solution of the optimization problem in (10). The equalities (10)–(12) could be translated into an appropriate cost function, in turn minimized using an iterative gradient-based technique. This approach is outside of the main goal of this paper, which is to achieve the blind identification of the MIMO channel based on noniterative methods, and so it is not pursued further.

## IV. Blind Channel Identification Algorithm

We develop the closed-form algorithm that estimates the channel matrix $\boldsymbol{H}$. We refer to it as the XBM algorithm

as opposed to the TXK algorithm [2]. The algorithm determines each of the users' channel matrices $\boldsymbol{H}_p$ up to a phase factor, so meeting the identifiability property of Theorem 1. However, the channel identification may require a larger set of output autocorrelation matrices, not to exceed the set $\{\boldsymbol{R_x}(0), \boldsymbol{R_x}(L_c), \boldsymbol{R_x}(L_c+1), \cdots, \boldsymbol{R_x}(L_c+L)\}$. The exact number of matrices used depends on the actual system's delay spreads. The nice structure of the sources' autocorrelation functions induced by the correlative filters with equal memory $L_c$ plays a major role in the derivation of the algorithm. When compared with TXK, XBM trades computational complexity for a smaller mean-square error (MSE) of the estimates. XBM is specially suited for implementation in parallel machines which alleviates this concern. We will also describe a simpler variation of XBM which decreases its computational cost and is efficient at high SNR scenarios. Finally, we present an equalization technique that uses the estimate of the channel matrix $\boldsymbol{H}$. This technique separates the sources' signals from the observations and recovers the transmitted sequences.

### A. The XBM Algorithm

The algorithm consists of four main steps.

1) Determination of the order $L$ of the system and estimation of $\boldsymbol{H}$ in the form $\boldsymbol{H}_0 = \boldsymbol{H}\boldsymbol{Q}^H$, where $\boldsymbol{Q}$ is an $L \times L$ unitary mixing matrix.
2) Determination of each user's delay spread, that is, the parameters $L_1, L_2, \cdots, L_P$. From this, the matrix $\boldsymbol{Q}$ will be partitioned into $P$ submatrices, $\boldsymbol{Q} = [\boldsymbol{Q}_1 \ \boldsymbol{Q}_2 \ \cdots \ \boldsymbol{Q}_P]$, where the $p$th submatrix $\boldsymbol{Q}_p$ has dimensions $L \times L_p$.
3) Determination of each of the $P$ submatrices $\boldsymbol{Q}_p$. This step can be optionally broken into $P$ parallel substeps, each one leading to the computation of $\hat{\boldsymbol{Q}}_p = \boldsymbol{Q}_p e^{j\theta_p}$.
4) Determination of the MIMO channel matrix $\hat{\boldsymbol{H}} = \boldsymbol{H}_0\hat{\boldsymbol{Q}}$, where $\hat{\boldsymbol{Q}} = [\hat{\boldsymbol{Q}}_1 \ \hat{\boldsymbol{Q}}_2 \ \cdots \ \hat{\boldsymbol{Q}}_P]$.

These four steps are detailed in the sequel. We derive the algorithm based on the true second-order statistics of the MIMO channel's output. In practice, the corresponding sample correlation matrices are employed.

*Step 1:* Determination of $L$ and $\boldsymbol{H}_0$.

The starting point is the estimation of the system's order $L$. Given (4) and the assumptions A1)–A3), we have $\boldsymbol{R_x}(0) = \boldsymbol{H}\boldsymbol{H}^H + \boldsymbol{R_w}(0)$. The order $L$ is determined from the spectrum profile of $\boldsymbol{R_x}(0)$ after subtracting the noise autocorrelation matrix $\boldsymbol{R_w}(0)$. As $\operatorname{rank}(\boldsymbol{R_x}(0) - \boldsymbol{R_w}(0)) = \operatorname{rank}(\boldsymbol{H}\boldsymbol{H}^H) = L$, $L$ is the number of nonzero eigenvalues of $\boldsymbol{R_x}(0) - \boldsymbol{R_w}(0)$. In practice, a threshold test is applied. When the noise process is modeled as spatio-temporal white Gaussian noise, i.e., $\boldsymbol{R_w}(k,l) = \mathrm{E}\{\boldsymbol{w}(k)\boldsymbol{w}(l)^H\} = \sigma_{\boldsymbol{w}}^2 \boldsymbol{I}_N \delta(k-l)$, statistical procedures based in information theory concepts, such as the Aikaike information criterion (AIC) or the minimum description length (MDL) test [30], are preferable.

Regardless of the approach chosen, a truncated eigenvalue decomposition (EVD) leads to

$$\boldsymbol{R_x}(0) - \boldsymbol{R_w}(0) = \boldsymbol{V}\boldsymbol{\Lambda}^2\boldsymbol{V}^H \tag{13}$$

where $\boldsymbol{V}$ is an $N \times L$ matrix with orthonormal columns and $\boldsymbol{\Lambda}$ is a diagonal matrix with nonzero entries. Since both terms in (13) equal $\boldsymbol{H}\boldsymbol{H}^H$, an estimate $\boldsymbol{H}_0 = \boldsymbol{H}\boldsymbol{Q}^H$ can be obtained from that factorization by setting $\boldsymbol{H}_0 \equiv \boldsymbol{V}\boldsymbol{\Lambda}$. The unknown

$$\boldsymbol{Q} = [\boldsymbol{Q}_1 \quad \boldsymbol{Q}_2 \quad \cdots \quad \boldsymbol{Q}_P]$$

is an $L \times L$ unitary mixing matrix. For further reference, we define the matrices

$$\boldsymbol{M}_l \equiv \boldsymbol{H}_0^{\#}(\boldsymbol{R_x}(l) - \boldsymbol{R_w}(l))\boldsymbol{H}_0^{\#\,H} = \boldsymbol{Q}\boldsymbol{R_s}(l)\boldsymbol{Q}^H$$

where the pseudo-inverse of $\boldsymbol{H}_0$ is given by $\boldsymbol{H}_0^{\#} = \boldsymbol{\Lambda}^{-1}\boldsymbol{V}^H$.

*Step 2:* Determination of $L_1, L_2, \cdots, L_P$.

This step determines the delay spreads of the users, i.e., the coefficients $L_p$ which add up to $L$. This is accomplished by exploiting the structure of $\boldsymbol{R_s}(L_c)$ in (7). Recall that

$$\boldsymbol{R_s}(L_c) = \operatorname{diag}[\eta_1\boldsymbol{I}_{L_1}, \eta_2\boldsymbol{I}_{L_2}, \cdots, \eta_P\boldsymbol{I}_{L_P}].$$

Hence, in the matrix $\boldsymbol{M}_{L_c} = \boldsymbol{Q}\boldsymbol{R_s}(L_c)\boldsymbol{Q}^H$, the delay spread $L_p$ appears as the algebraic multiplicity of the eigenvalue $\eta_p$. To access the eigenvalues of $\boldsymbol{M}_{L_c}$, we perform a Schur decomposition

$$\boldsymbol{M}_{L_c} = \boldsymbol{U}\boldsymbol{T}\boldsymbol{U}^H \tag{14}$$

where $\boldsymbol{U} = [\boldsymbol{U}_1 \ \boldsymbol{U}_2 \ \cdots \ \boldsymbol{U}_P]$ is unitary and $\boldsymbol{T}$ is upper-triangular (both $L \times L$). Although $\boldsymbol{M}_{L_c}$ is not Hermitian, it is a normal matrix (it commutes with its conjugate transpose) and thus it turns out that $\boldsymbol{T}$ in the Schur decomposition (14) is diagonal [31], [32] (in practical applications, the strictly upper diagonal entries are negligible). The spectrum of $\boldsymbol{M}_{L_c}$ is revealed in the diagonal of $\boldsymbol{T}$, and the delay spreads $L_1, L_2, \cdots, L_P$ are obtained by direct inspection by counting the multiplicity of each eigenvalue. In practice, the diagonal entries of $\boldsymbol{T}$ will cluster around the theoretical values $\{\eta_1, \eta_2, \cdots, \eta_P\}$ and a decision rule must be employed. We notice that for many sources in high noise scenarios, this method may fail due to the possible overlap of the eigenvalues' clusters. This topic deserves further investigation, a possible approach being the redesign of the correlative filters.

*Step 3:* Determination of $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_P$.

Additional information may be retrieved from the Schur factorization (14) due to the diagonal nature of $\boldsymbol{T}$. Without loss of generality, let the diagonal of $\boldsymbol{T}$ be ordered as $\boldsymbol{R_s}(L_c)$, see (7). Then we get $\boldsymbol{Q}_p\boldsymbol{Q}_p^H = \boldsymbol{U}_p\boldsymbol{U}_p^H$. That is, the projector associated to the $p$th user, i.e., the orthogonal projector onto $\mathcal{R}(\boldsymbol{Q}_p)$, given by $\boldsymbol{\Pi}_p = \boldsymbol{Q}_p\boldsymbol{Q}_p^H$, is also available from the Schur decomposition in (14). We compute each $\boldsymbol{Q}_p = [\boldsymbol{q}_p(0) \ \cdots \ \boldsymbol{q}_p(L_p-1)]$ in the residual unitary mixing matrix $\boldsymbol{Q}$. Focus on the $p$th user, and consider the matrices

$$\begin{aligned}\boldsymbol{P}_l &\equiv \boldsymbol{\Pi}_p\boldsymbol{M}_{L_c+l} + (\boldsymbol{I}_L - \boldsymbol{\Pi}_p) \\ &= \eta_p\boldsymbol{Q}_p\boldsymbol{R_s}(l)\boldsymbol{Q}_p^H + (\boldsymbol{I}_L - \boldsymbol{\Pi}_p).\end{aligned}$$

It is straightforward to verify that

$$\begin{cases}\mathcal{N}(\boldsymbol{P}_{l+1}^H) = \operatorname{span}\{\boldsymbol{q}_p(0), \boldsymbol{q}_p(1), \cdots, \boldsymbol{q}_p(l)\} \\ \mathcal{N}(\boldsymbol{P}_{L_p-l}) = \operatorname{span}\{\boldsymbol{q}_p(l), \boldsymbol{q}_p(l+1), \cdots, \boldsymbol{q}_p(L_p-1)\}\end{cases}$$

for $l = 0, 1, \cdots, L_p - 1$. Therefore, each column $\boldsymbol{q}_p(l)$ of $\boldsymbol{Q}_p$ can be determined by intersecting the kernels of $\boldsymbol{P}_{l+1}^H$ and

$\boldsymbol{P}_{L_p-l}$. This can be accomplished [31] by calculating the right singular vector corresponding to the minimum singular value of the $2L \times L$ matrix

$$\boldsymbol{T}_l \equiv \begin{bmatrix} \boldsymbol{P}_{l+1}^H \\ \boldsymbol{P}_{L_p-l} \end{bmatrix}$$

or equivalently, by extracting the eigenvector associated with the minimum eigenvalue of the $L \times L$ semipositive definite matrix $\boldsymbol{T}_l^H \boldsymbol{T}_l$. In either case, an estimate $\boldsymbol{q}_p(l)e^{j\omega_l}$ is generated, i.e., a matrix $\boldsymbol{W} = \boldsymbol{Q}_p\boldsymbol{\Omega}$ is determined, where $\boldsymbol{\Omega} = \mathrm{diag}[e^{j\omega_0}, e^{j\omega_1}, \cdots, e^{j\omega_{L_p-1}}]$ contains (possibly distinct) phase factors.

We now equalize the remaining phases in $\boldsymbol{\Omega}$, i.e., we condense the phase ambiguity in just one parameter, $\boldsymbol{\Omega} = e^{j\omega}\boldsymbol{I}_{L_p}$. This will show that the algorithm verifies the identifiability result expressed in Theorem 1 when using the exact process correlation matrices. We exploit the shift property structure of

$$\begin{aligned} \boldsymbol{N} &\equiv \boldsymbol{\Pi}_p \boldsymbol{M}_{L_c+1} \\ &= \eta_p \boldsymbol{Q}_p \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \boldsymbol{Q}_p^H \end{aligned}$$

by defining the matrix

$$\boldsymbol{\Delta} \equiv (\boldsymbol{N}\boldsymbol{W}/\eta_p)^H \boldsymbol{W} = \boldsymbol{\Omega}^* \boldsymbol{K}_{L_p}^T \boldsymbol{\Omega}.$$

The first upper diagonal of $\boldsymbol{\Delta}$ contains information about two consecutive phase ambiguity factors

$$\Delta(l, l+1) = e^{j(\omega_l - \omega_{l-1})}$$

for $l = 1, 2, \cdots, L_p - 1$. This can be used to correct all phase drifts from the last factor, by setting

$$\hat{\boldsymbol{Q}}_p \equiv \boldsymbol{W}\ \mathrm{diag}\left[\beta_0/|\beta_0|, \beta_1/|\beta_1|, \cdots, \beta_{L_p-1}/|\beta_{L_p-1}|\right] \quad (15)$$

where $\beta_l \equiv \Pi_{j=l}^{L_p-2}\ \Delta(j+1, j+2)$. It is straightforward to verify that this results in an estimate $\hat{\boldsymbol{Q}}_p = \boldsymbol{Q}_p e^{j\theta_p}$, where $\theta_p = \omega_{L_p-1}\ (\mathrm{mod}\ 2\pi)$.

*Step 4:* Determination of $\hat{\boldsymbol{H}}$.

Defining $\hat{\boldsymbol{H}}_p \equiv \boldsymbol{H}_0 \hat{\boldsymbol{Q}}_p$, the matrix $\hat{\boldsymbol{H}} = [\hat{\boldsymbol{H}}_1\ \hat{\boldsymbol{H}}_2\ \cdots\ \hat{\boldsymbol{H}}_P]$ meets the theoretical identifiability result of Theorem 1. The $P$ submatrices $\hat{\boldsymbol{H}}_p$ can be computed in parallel at the receiver since no cross-related data is ever needed.

*MIMO Blind Channel Identification:* Table I summarizes the MIMO blind channel identification algorithm. We assume that there are $K$ data samples available, i.e., $\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots, \boldsymbol{x}(K)$. The output correlation matrices $\boldsymbol{R}_{\boldsymbol{x}}(l)$ are replaced by their respective estimates

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}}(l) \equiv \frac{1}{K-l} \sum_{k=l+1}^{K} \boldsymbol{x}(k)\boldsymbol{x}(k-l)^H.$$

For simplicity, we assume that the noise process $\boldsymbol{w}(k)$ is a spatio-temporal white Gaussian process, $\boldsymbol{R}_{\boldsymbol{W}}(k,l) = \sigma_{\boldsymbol{w}}^2 \boldsymbol{I}_N \delta(k-l)$. Notice that, for this case, the MDL criteria also yields an estimate of $\sigma_{\boldsymbol{w}}^2$, i.e., previous knowledge of this system parameter is not required.

*Variation:* We discuss briefly a possible alternative to avoid the Schur decomposition, step 2 in the algorithm of Table I, necessary to estimate $L_p$ and the projector $\boldsymbol{\Pi}_p = \boldsymbol{Q}_p\boldsymbol{Q}_p^H$. It consists in exploiting the concept of interpolating Lagrange polynomials. Namely, if

$$P_p(t) = \prod_{q \neq p} \frac{t - \eta_q}{\eta_p - \eta_q} \quad (16)$$

is the Lagrange polynomial corresponding to the $p$th user, then, given the unitary structure of $\boldsymbol{Q}$ in

$$\begin{aligned} \boldsymbol{M}_{L_c} &= \boldsymbol{Q}\boldsymbol{R}_{\boldsymbol{s}}(L_c)\boldsymbol{Q}^H \\ &= \boldsymbol{Q}\ \mathrm{diag}\left[\eta_1 \boldsymbol{I}_{L_1}, \eta_2 \boldsymbol{I}_{L_2}, \cdots, \eta_P \boldsymbol{I}_{L_P}\right]\boldsymbol{Q}^H \end{aligned}$$

we conclude that the matrix polynomial $P_p(\boldsymbol{M}_{L_c}) = \boldsymbol{Q}_p\boldsymbol{Q}_p^H = \boldsymbol{\Pi}_p$. This method is more sensitive to errors in the (sample) estimate of $\boldsymbol{M}_{L_c}$ than the Schur decomposition, since it depends on the degree of *exact* matching between the (sample) eigenvalues of $\boldsymbol{M}_{L_c}$ and the expected $\eta_p$'s. This alternative is best suited for higher SNR scenarios with large data sets. To obtain each user delay spread, we use the consistent estimators

$$\hat{L}_p = \lfloor \mathrm{tr}(\boldsymbol{\Pi}_p) \rceil \quad (17)$$

where $\lfloor x \rceil$ denotes the integer nearest to $x \in \mathbb{R}$, in the Euclidean metric. Set $\boldsymbol{l} \equiv [L_1\ L_2\ \cdots\ L_P]^T$ and $\hat{\boldsymbol{l}} \equiv [\hat{L}_1\ \hat{L}_2\ \cdots\ \hat{L}_P]^T$. Also, let $\mathbb{Z}_+^n$ denote the set of $n$-dimensional positive integer column vectors and, for $\boldsymbol{x}_0 \in \mathbb{Z}_+^n$, let $\mathcal{D}(\boldsymbol{x}_0) \equiv \{\boldsymbol{x} \in \mathbb{Z}_+^n \colon \boldsymbol{x} = \boldsymbol{x}_0 \pm \boldsymbol{e}_{i;n}, i = 1, 2, \cdots, n\}$; i.e., $\mathcal{D}(\boldsymbol{x}_0)$ contains the $2n$ vectors in $\mathbb{Z}_+^n$ which differ from $\boldsymbol{x}_0$ by $\pm 1$ in one of its $n$ entries. For finite (large) data sets, simulations studies showed that the most likely error event in the estimates (17) corresponds to $\hat{\boldsymbol{l}} \in \mathcal{D}(\boldsymbol{l})$. We now describe a method that overcomes such errors. It is based on the following straightforward fact, for which we omit the proof.

*Fact:* Let $\hat{\boldsymbol{l}} \in \mathcal{D}(\boldsymbol{l})$. Let $\tilde{\boldsymbol{l}} = [\tilde{L}_1\ \tilde{L}_2\ \cdots\ \tilde{L}_P]^T$ be a global minimizer of the function $\chi \colon \mathbb{Z}_+^P \to \mathbb{R}$

$$\begin{aligned} \boldsymbol{x} &= [x_1\ \ x_2\ \ \cdots\ \ x_P]^T \mapsto \chi(\boldsymbol{x}) \\ &= \left| \left( \sum_{p=1}^{P} \eta_p x_p \right) - \mathrm{tr}(\boldsymbol{M}_{L_c}) \right|^2 \end{aligned}$$

over the finite set $\mathcal{S} \equiv \mathcal{D}(\hat{\boldsymbol{l}}) \cap \{\boldsymbol{x} \in \mathbb{Z}_+^P \colon x_1 + x_2 + \cdots + x_P = L\}$. Then, $\tilde{\boldsymbol{l}} = \boldsymbol{l}$.

From (14), we see that $\mathrm{tr}(\boldsymbol{M}_{L_c}) = \Sigma_{p=1}^{P}\ \eta_p L_p$; thus, $\chi(\boldsymbol{l}) = 0$, where $\boldsymbol{l} = [L_1\ L_2\ \cdots\ L_P]^T$. What the fact emphasizes is that $\boldsymbol{l}$ is the *only* zero of $\chi$ in the neighborhood $\mathcal{S}$ of $\hat{\boldsymbol{l}}$. We exploit this result in the following manner. We compute the vector of estimates $\hat{\boldsymbol{l}}$ as in (17). We check if its elements sum up to $L$ (estimated in Step 1). If not, we assume that $\hat{\boldsymbol{l}} \in \mathcal{D}(\boldsymbol{l})$ (most likely error event) and minimize $\chi$ over the set $\mathcal{S}$ by an exhaustive search. The cardinality of $\mathcal{S}$ is upper-bounded by $2P$, so this is not computationally expensive. This yields the new vector of estimates. It will be seen in Section V that this method allows for the substantial improvement of the initial estimates in (17).

TABLE I
MIMO XBM BLIND CHANNEL IDENTIFICATION ALGORITHM AS PRESENTED IN THE PAPER

**Step 1: Determination of $L$ and $\mathbf{H}_0$**

Perform an EVD, $\mathbf{R}_{\mathbf{x}}(0) = \mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^H$ ($\mathbf{\Lambda}$ sorted in decreasing order)
Estimate both $L$ and $\sigma_{\mathbf{w}}^2$ by the MDL criteria [30]
Partition $\mathbf{V} = [\ \mathbf{V}_1 \quad \mathbf{V}_2\ ]$ and $\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \\ & \mathbf{\Lambda}_2 \end{bmatrix}$ ($\mathbf{V}_1 : N \times L$, $\mathbf{\Lambda}_1 : L \times L$)
Compute $\mathbf{H}_0 = \mathbf{V}_1\tilde{\mathbf{\Lambda}}_1$ and $\mathbf{H}_0^{\#} = \tilde{\mathbf{\Lambda}}_1^{-1}\mathbf{V}_1^H$, where $\tilde{\mathbf{\Lambda}}_1 \equiv (\mathbf{\Lambda}_1^2 - \sigma_{\mathbf{w}}^2\mathbf{I}_L)^{1/2}$
Compute the matrices $\mathbf{M}_l = \mathbf{H}_0^{\#}\mathbf{R}_{\mathbf{x}}(l)\,\mathbf{H}_0^{\#\,H}$, for $l = L_c, L_c+1, \ldots, L_c+L$

**Step 2: Determination of $L_1, L_2, \ldots, L_P$**

Perform a Schur decomposition, $\mathbf{M}_{L_c} = \mathbf{U}\mathbf{T}\mathbf{U}^H$
Estimate each $L_p$ by counting the diagonal entries of $\mathbf{T}$ closest to $\eta_p$
Compute the projectors $\mathbf{\Pi}_p = \mathbf{U}_p\mathbf{U}_p^H$, where $\mathbf{U}_p : L \times L_p$ contain the columns of $\mathbf{U}$ associated to $\eta_p$

**Step 3: Determination of $\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_P$**

for $p = 1, 2, \ldots, P$
  Compute the matrices $\mathbf{P}_l = \mathbf{\Pi}_p\mathbf{M}_{L_c+l} + (\mathbf{I}_L - \mathbf{\Pi}_p)$, for $l = 0, 1, \ldots, L_p - 1$
  Compute $\mathbf{W} = [\ \mathbf{w}(0) \quad \mathbf{w}(1) \quad \cdots \quad \mathbf{w}(L_p-1)\ ]$, where $\mathbf{w}(l)$ is the minimum right singular vector of $\mathbf{T}_l = \left[\mathbf{P}_{l+1}\,\mathbf{P}_{L_p-l}^H\right]^H$
  Compute $\mathbf{N} = \mathbf{\Pi}_p\mathbf{M}_{L_c+1}$ and $\mathbf{\Delta} = (\mathbf{N}\mathbf{W}/\eta_p)^H\mathbf{W}$
  Compute $\hat{\mathbf{Q}}_p = [\ \hat{\mathbf{q}}_p(0) \quad \hat{\mathbf{q}}_p(1) \quad \cdots \quad \hat{\mathbf{q}}_p(L_p-1)\ ]$, where $\hat{\mathbf{q}}_p(l) = \mathbf{w}(l)\,\beta_l/|\beta_l|$ and $\beta_l = \prod_{j=l}^{L_p-2}\Delta(j+1, j+2)$
end

**Step 4: Determination of $\hat{\mathbf{H}}$**

Let $\hat{\mathbf{Q}} = [\ \hat{\mathbf{Q}}_1 \quad \hat{\mathbf{Q}}_2 \quad \cdots \quad \hat{\mathbf{Q}}_P\ ]$
Compute $\hat{\mathbf{H}} = \mathbf{H}_0\hat{\mathbf{Q}}$.

### *B. Sources Separation and Equalization*

Let $\boldsymbol{H} = [\boldsymbol{H}_1 \ \boldsymbol{H}_2 \ \cdots \ \boldsymbol{H}_P]$ be the MIMO channel matrix. We estimate each source's emitted sequence $s_p(k)$. The extraction of the information symbols $a_p(k)$ from $s_p(k)$ is straightforward and is efficiently carried out by an MLSE approach [25], [26]. Notice that, due to the controlled ISI introduced at the transmitter, the complexity of the detection step increases since we need to use a Viterbi algorithm whose trellis contains now at least $(\#\mathcal{A}_p)^{L_c}$ states, where $\#\mathcal{A}_p$ is the cardinality of the modulation alphabet corresponding to the user $p$. We focus on the $p$th user.

*Sources' Separation:* The first step consists in isolating its signal from the observations. This can be achieved by applying an oblique projector to the data [33]. More specifically, let us define

$$\mathit{\Xi}_p \equiv \boldsymbol{H}_p\left(\boldsymbol{H}_p^H\boldsymbol{\Upsilon}_p\boldsymbol{H}_p\right)^{-1}\boldsymbol{H}_p^H\boldsymbol{\Upsilon}_p$$

where $\boldsymbol{\Upsilon}_p = \boldsymbol{I}_N - \tilde{\boldsymbol{H}}_p(\tilde{\boldsymbol{H}}_p^H\tilde{\boldsymbol{H}}_p)^{-1}\tilde{\boldsymbol{H}}_p^H$, and $\tilde{\boldsymbol{H}}_p$ is obtained from $\boldsymbol{H}$ by deleting the $p$th submatrix, i.e., $\tilde{\boldsymbol{H}}_p = [\boldsymbol{H}_1 \ \cdots \ \boldsymbol{H}_{p-1}\ \boldsymbol{H}_{p+1} \ \cdots \ \boldsymbol{H}_P]$. As described in [33], $\mathit{\Xi}_p$ nulls every vector lying in $\mathcal{R}(\tilde{\boldsymbol{H}}_p)$ and preserves those contained in $\mathcal{R}(\boldsymbol{H}_p)$. Thus

$$\boldsymbol{y}(k) \equiv \mathit{\Xi}_p\boldsymbol{x}(k) = \boldsymbol{H}_p\boldsymbol{s}_p(k) + \boldsymbol{n}(k)$$

where the new noise samples are defined by $\boldsymbol{n}(k) = \mathit{\Xi}_p\boldsymbol{w}(k)$, having autocorrelation matrices given by $\boldsymbol{R}_{\boldsymbol{n}}(l) = \mathit{\Xi}_p\boldsymbol{R}_{\boldsymbol{w}}(l)\mathit{\Xi}_p^H$. In practice, an estimate of $\boldsymbol{H}_p$ is used so we expect a residual interuser interference. In the performance studies in Section V, it will be seen that the algorithm effectively combats this residual crosstalk interference.

*Sources' Equalization:* We now address the problem of estimating $s_p(k)$ under the Gaussian noise assumption. For estimating $s_p(k)$, we exploit the fact that multiple delayed replicas of this signal are available in the vector $\boldsymbol{s}_p(k) = [s_p(k) \ \ s_p(k-1) \ \cdots \ \ s_p(k-L_p+1)]^T$. Thus, instead of picking up just one element of $\hat{\boldsymbol{s}}_p(k) = \boldsymbol{H}_p^{\#}\boldsymbol{y}(k)$ as in [2], we start by synchronizing the time-delayed signals in $\hat{\boldsymbol{s}}_p(k)$. This produces data samples given by

$$\check{\boldsymbol{s}}_p(k) = \mathbf{1}s_p(k) + \check{\boldsymbol{n}}(k) \tag{18}$$

where $\mathbf{1} \equiv [1\ 1\ \cdots\ 1]^T$ ($L_p$ times), and the $l$th element of $\check{\boldsymbol{n}}(k)$ is $\boldsymbol{e}_{l;L_p}^H\boldsymbol{H}_p^{\#}\boldsymbol{n}(k+l-1)$. Thus, the $(l,m)$ entry of $\boldsymbol{R}_{\check{\boldsymbol{n}}}(0)$ equals the $(l,m)$ entry of $\boldsymbol{H}_p^{\#}\boldsymbol{R}_{\boldsymbol{n}}(l-m)\boldsymbol{H}_p^{\#\,H}$. Estimation of $s_p(k)$ given (18) is a well-known studied problem in the area of statistical signal processing [28]. It is equivalent to the estimation of a dc-level in colored Gaussian noise, and the minimum variance unbiased (MVU) estimator is given by

$$\hat{s}_p(k) = \frac{\mathbf{1}^T\boldsymbol{R}_{\check{\boldsymbol{n}}}(0)^{-1}}{\mathbf{1}^T\boldsymbol{R}_{\check{\boldsymbol{n}}}(0)^{-1}\mathbf{1}}\check{\boldsymbol{s}}_p(k). \tag{19}$$

If the noise process $\boldsymbol{w}(k)$ is modeled as a spatio-temporal white Gaussian process, then we have $\boldsymbol{R}_{\boldsymbol{w}}(k,l) = \sigma_{\boldsymbol{w}}^2\boldsymbol{I}_N\delta(k-l)$ and (19) simplifies to

$$\hat{s}_p(k) = \frac{1}{\mathrm{tr}(\boldsymbol{R}_{\check{\boldsymbol{n}}}(0)^{-1})}\mathbf{1}^T\boldsymbol{R}_{\check{\boldsymbol{n}}}(0)^{-1}\check{\boldsymbol{s}}_p(k)$$

TABLE II
SIMO XBM BLIND CHANNEL IDENTIFICATION ALGORITHM AS DESCRIBED IN THE PAPER

**Determination of $L$ and $\mathbf{H}_0$**

Perform an EVD, $\mathbf{R_x}(0) = \mathbf{V}\boldsymbol{\Lambda}^2\mathbf{V}^H$ ($\boldsymbol{\Lambda}$ sorted in decreasing order)
Estimate both $L$ and $\sigma_{\mathbf{w}}^2$ by the MDL criteria [30]
Partition $\mathbf{V} = [\ \mathbf{V}_1 \quad \mathbf{V}_2\ ]$ and $\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_1 & \\ & \boldsymbol{\Lambda}_2 \end{bmatrix}$ ($\mathbf{V}_1 : N \times L,\ \boldsymbol{\Lambda}_1 : L \times L$)
Compute $\mathbf{H}_0 = \mathbf{V}_1\tilde{\boldsymbol{\Lambda}}_1$ and $\mathbf{H}_0^{\#} = \tilde{\boldsymbol{\Lambda}}_1^{-1}\mathbf{V}_1^H$, where $\tilde{\boldsymbol{\Lambda}}_1 \equiv (\boldsymbol{\Lambda}_1^2 - \sigma_{\mathbf{w}}^2\mathbf{I}_L)^{1/2}$
Compute the matrices $\mathbf{M}_l = \mathbf{H}_0^{\#}\mathbf{R_x}(l)\,\mathbf{H}_0^{\#\,H}$, for $l = 0, 1, \ldots, L-1$ ($\mathbf{M}_L \equiv \mathbf{0}$)

**Determination of $\mathbf{Q}$**

Compute $\mathbf{W} = [\ \mathbf{w}(0) \quad \mathbf{w}(1) \quad \cdots \quad \mathbf{w}(L-1)\ ]$, where $\mathbf{w}(l)$ is the minimum right singular vector of $\mathbf{T}_l = \left[\mathbf{M}_{l+1}\,\mathbf{M}_{L-l}^H\right]^H$
Compute $\boldsymbol{\Delta} = (\mathbf{M}_1\mathbf{W})^H\mathbf{W}$
Compute $\hat{\mathbf{Q}} = [\ \hat{\mathbf{q}}(0) \quad \hat{\mathbf{q}}(1) \quad \cdots \quad \hat{\mathbf{q}}(L-1)\ ]$, where $\hat{\mathbf{q}}(l) = \mathbf{w}(l)\,\beta_l/|\beta_l|$ and $\beta_l = \prod_{j=l}^{L_p-2}\Delta(j+1, j+2)$

**Determination of $\hat{\mathbf{H}}$**

Compute $\hat{\mathbf{H}} = \mathbf{H}_0\hat{\mathbf{Q}}$.

where $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}(0)$ is a diagonal matrix having the same diagonal entries as $\boldsymbol{R}_{\boldsymbol{n}}(0) = \sigma_{\boldsymbol{w}}^2\boldsymbol{H}_p^{\#}\boldsymbol{\Xi}_p\boldsymbol{\Xi}_p^H\boldsymbol{H}_p^{\#\,H}$.

## V. COMPUTER SIMULATIONS

We present two sets of simulations. In the first set, we consider a SIMO channel and compare our XBM algorithm with TXK [2]. This comparison is based on the mean-square error of the channel matrix estimates and on the probability of error at the receiver output. The second set of experiments is performed in the context of the GSM system. We run a Monte Carlo simulation over several randomly generated MIMO channel matrices. We evaluate the resulting mean-square error of the channel matrix estimates and the ability of our method to discriminate the sources.

*Single-User Context:* In a single-source environment there is no need for coloring the information sequence before transmission since there are no multiple users to be discriminated. Thus, the user transmits directly a sequence of i.i.d. unit-power symbols randomly chosen in its modulation alphabet. Adaptation of the XBM algorithm to the SIMO situation is straightforward and consists basically in eliminating the Schur decomposition step; see Table II.

We start by comparing the normalized error (NE)

$$\mathrm{NE} \equiv \frac{||\hat{\boldsymbol{H}} - \boldsymbol{H}||}{||\boldsymbol{H}||}$$

achieved by the TXK and the XBM algorithms. Here, $||\cdot||$ denotes the Frobenius norm, which is an MSE-like norm. We generated randomly $S = 500$ SIMO channels, each parameterized by an $N \times L$ convolution matrix $\boldsymbol{H}$, where $N = 20$ and $L = 15$. The complex entries of each $\boldsymbol{H}$ are independent samples of a complex zero-mean normal random variable with variance two. For each SIMO channel, a QAM16 input digital source is considered, and the channel output is corrupted by spatio-temporal white Gaussian noise, i.e., $\boldsymbol{R}_{\boldsymbol{w}}(k,l) = \sigma_{\boldsymbol{w}}^2\boldsymbol{I}_N\delta(k,l)$. The noise power is adjusted for each SIMO channel in order to guarantee an SNR $= 15$ dB, where the SNR is defined by

$$\mathrm{SNR} \equiv \frac{\mathrm{E}\{||\boldsymbol{H}\boldsymbol{s}(k)||^2\}}{\boldsymbol{E}\{||\boldsymbol{w}(k)||^2\}}.$$

We used the MDL criteria [30] to estimate both the system's order $L$ and the noise variance $\sigma_{\boldsymbol{w}}^2$. The identification of each channel matrix $\boldsymbol{H}$ is based on $K = 200$ data samples. Fig. 1 exhibits the results obtained. Observing this figure, we conclude that XBM (lower curve) is clearly more robust than the TXK algorithm (upper curve). In fact, the average of the normalized error taken over the ensemble of the generated SIMO channels, mean$(\mathrm{NE}_{\mathrm{XBM}}) = 0.1387$, is smaller than that achieved by the TXK algorithm, mean$(\mathrm{NE}_{\mathrm{TXK}}) = 0.2294$. Moreover, the same conclusion applies with respect to its variance, since we have var$(\mathrm{NE}_{\mathrm{XBM}}) = 0.0004$ and var$(\mathrm{NE}_{\mathrm{TXK}}) = 0.0010$.

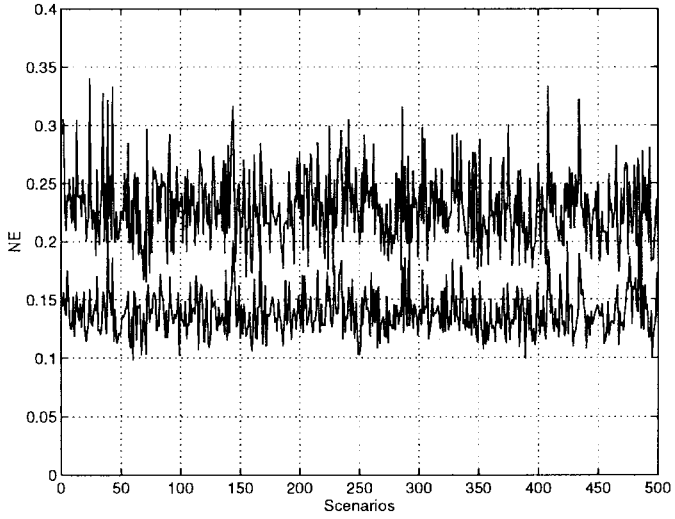



Fig. 1. Normalized error of the TXK algorithm of Tong, Xu, and Kailath (upper curve) and of the XBM algorithm of this paper (lower curve).

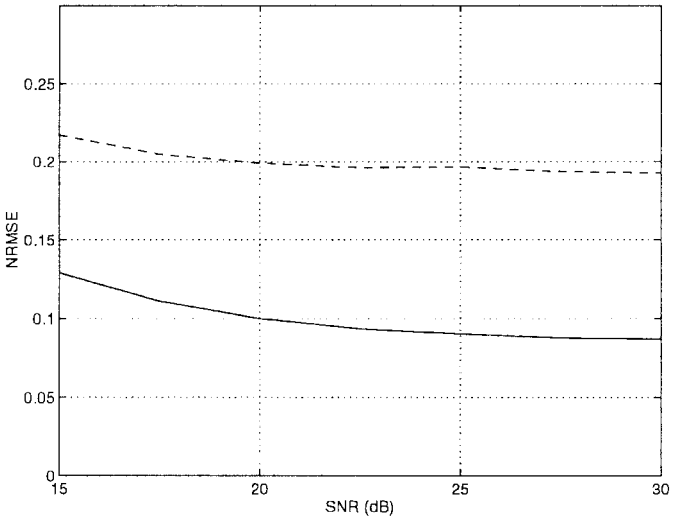



Fig. 2. Normalized root mean-square error of the TXK algorithm of Tong, Xu, and Kailath (dashed line) and of the XBM (solid) algorithm of this paper (solid line).

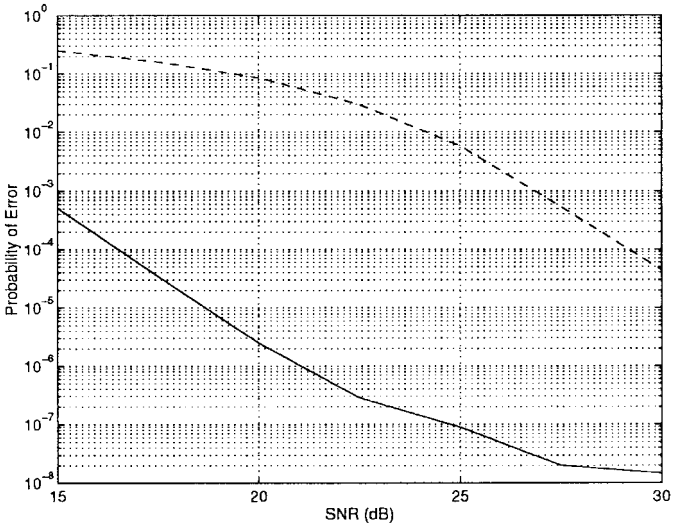



Fig. 3. Probability of error of the TXK algorithm of Tong, Xu, and Kailath (dashed line) and of the XBM algorithm of this paper (solid line).

To obtain a more accurate performance analysis, we fixed one of the generated scenarios and made the SNR vary between $\mathrm{SNR}_{\min} = 15$ dB and $\mathrm{SNR}_{\max} = 30$ dB, in steps of $\mathrm{SNR}_{\mathrm{step}} = 2.5$ dB. For each SNR, $M = 500$ statistically independent trials were considered. In each trial, the channel identification is based on $K = 200$ data samples. Fig. 2 shows the normalized root mean-square error (NRMSE)

$$\mathrm{NRMSE} \equiv \frac{1}{\|\boldsymbol{H}\|}\sqrt{\frac{1}{M}\sum_{m=1}^{M}\|\hat{\boldsymbol{H}}^{(m)} - \boldsymbol{H}\|^2}$$

for the two algorithms, where $\hat{\boldsymbol{H}}^{(m)}$ is the channel estimate at the $m$th run. The dashed and solid lines refer to the TXK and XBM algorithms, respectively. As we see, smaller NRMSE's are produced by XBM for all the SNR's considered. To compare the combined identification and equalization capability of the complete algorithm described in Section IV-B (also referred for short as XBM), with the one in [2] (also referred to as TXK), we determined the probability of error associated with the reconstructed input signal sequence. In Fig. 3 the dashed line refers to TXK and the solid line to XBM. For each value of the SNR, the probability of error was determined in the following manner. In each of the $M = 500$ trials, the channel's estimate (based on the first $K = 200$ samples) was used to equalize the next $\tilde{K} = 2000$ new data samples and the variance at the equalizer's output was computed. The probability of error was then evaluated on the basis of the mean value of these $M = 500$ variances. That is, for each SNR, we have $P_{\mathrm{error}} = P_e(\hat{\sigma}^2)$, where

$$\hat{\sigma}^2 = \frac{1}{M}\sum_{m=1}^{M}\sigma_m^2$$

is the mean variance over all trials, and

$$P_e(\sigma^2) = 1 - \frac{1}{4}\left(2 - \frac{3}{2}\,\mathrm{erfc}\left(\frac{1}{\sigma}\right)\right)^2$$

is the probability of error associated to the ML estimation of a QAM16 symbol in zero-mean Gaussian noise with variance $\sigma^2$ [34]. As we see in Fig. 3, significant gains are obtained through the entire range of SNR's considered. Figs. 4 and 5 confirm these probability of error computations. They show the spread of $\tilde{K} = 2000$ samples after equalization by the TXK and the XBM algorithms, respectively. The figures are typical of any one of the 500 simulated trials at SNR $= 15$ dB. As we see, the channel is well equalized by the XBM algorithm, whereas several decision errors are likely to occur at the TXK equalizer's output. The better performance of the XBM over the TXK algorithm is perhaps best explained by the fact that, in the former, the columns of the residual unitary matrix $\boldsymbol{Q}$ are identified independently of each other, whereas, in the latter, the estimation relies on a sequential unfolding of this matrix starting from the last column. This process introduces propagation errors which accumulate from column to column.

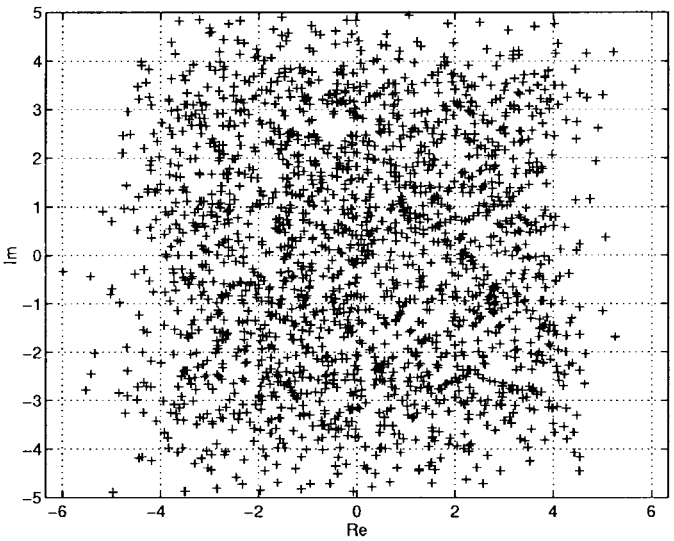



Fig. 4. The output of the TXK equalizer of Tong, Xu, and Kailath.

*Multiple-User Context:* To assess the performance of the XBM algorithm in a multiuser environment, we considered $P = 3$ binary sources. For physical communication param-

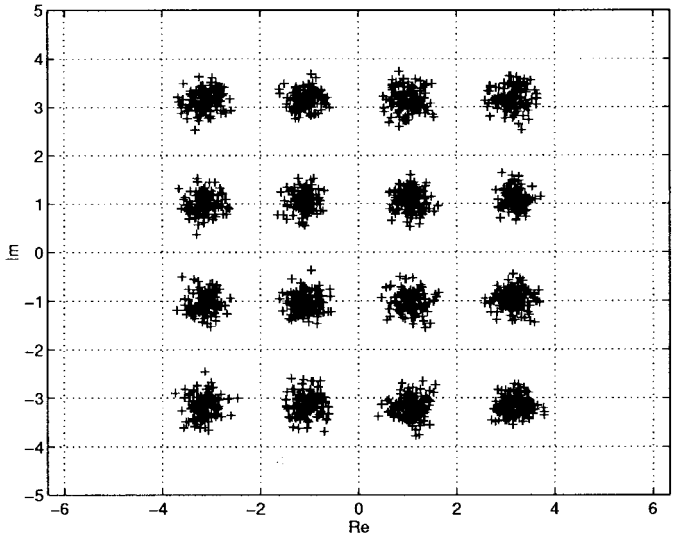



Fig. 5. The output of the XBM equalizer of this paper.

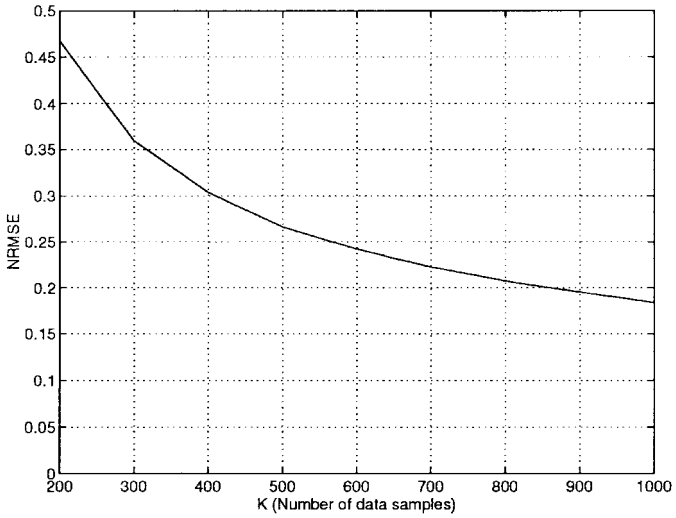



Fig. 6. MIMO: normalized root mean-square error of the XBM algorithm of this paper.

eters, we used typical values of the mobile GSM system. Namely, the standardized impulse response corresponding to the more hostile propagation environment [27] was chosen to model the channel between each source and one antenna element. This impulse response is specified by six delta impulses modeling six equal-power independent Rayleigh-fading paths. The delays associated to these paths range from $\tau_{\min} = 0$ $\mu$s to $\tau_{\max} = 16$ $\mu$s in equal steps of $\tau_{\text{step}} = 3.2$ $\mu$s. In addition to the amplitude random modulation, each path is also randomly modulated in phase. The sources use the symbol period $T = 3.7$ $\mu$s and transmit raised-cosine pulses $r_{T/2}(t - T/2)$ with 100% rolloff. It is readily seen that, with these parameters, each system's impulse response spans over $L_p = 5$ symbol's intervals. At the base station, we assumed an antenna array with $D = 6$ sensors and an oversampling factor of $J = 4$. This results in MIMO channel matrices $\boldsymbol{H}$ of dimension $(DJ) \times (L_1 + L_2 + L_3) = 24 \times 15$. Notice also that each matrix $\boldsymbol{H}$ contains samples of $M \times P = 18$ GSM impulse responses. The correlative filters were designed as in Section III, with $L_c = 5$. The system's output is corrupted by spatio-temporal white Gaussian noise.

We considered $S = 10$ distinct scenarios $\boldsymbol{H}$. For each one, the noise variance was chosen to fix the SNR $= 20$ dB. The number of data samples used to estimate each $\boldsymbol{H}$ ranges from $K_{\min} = 200$ to $K_{\max} = 1000$ in steps of $K_{\text{step}} = 100$. For each $K$, we ran a Monte Carlo simulation consisting of $M = 200$ independent trials. We computed the corresponding NRMSE. Fig. 6 displays the average results over the $S = 10$ scenarios considered, which are monotonically decreasing.

We also evaluate the ability of XBM to discriminate among the users. Recall that the $p$th user's signal is separated from the observations by applying an oblique projector to the data samples

$$\begin{aligned}\boldsymbol{y}(k) &\equiv \boldsymbol{\Xi}_p \boldsymbol{x}(k) \\ &= \boldsymbol{\Xi}_p \boldsymbol{H}_p \boldsymbol{s}_p(k) + \sum_{q \neq p} \boldsymbol{\Xi}_p \boldsymbol{H}_q \boldsymbol{s}_q(k) + \boldsymbol{n}(k). \end{aligned} \quad (20)$$

The second term on the right-hand side of (20) measures the residual user's interference crosstalk. Its relative power can be measured by the signal-to-interference ratio (SIR)

$$\text{SIR}_p \equiv \frac{\text{E}\{\|\dot{\boldsymbol{H}}_p \boldsymbol{s}_p(k)\|^2\}}{\text{E}\left\{\left\|\sum_{q \neq p} \dot{\boldsymbol{H}}_q \boldsymbol{s}_q(k)\right\|^2\right\}} = \frac{\text{tr}(\dot{\boldsymbol{H}}_p \dot{\boldsymbol{H}}_p^H}{\sum_{qn \neq p} \text{tr}\,(\dot{\boldsymbol{H}}_q \dot{\boldsymbol{H}}_q^H)}$$

where we have defined $\dot{\boldsymbol{H}}_p \equiv \boldsymbol{\Xi}_p \boldsymbol{H}_p$. Fig. 7 shows the average results obtained in terms of $\text{SIR}_p$ over the $S = 10$ scenarios simulated. The solid, dashed, and dotted lines refer to the first, second, and third user, respectively. As seen, XBM satisfactorily rejects interuser interference even for small data blocks, e.g., $K = 300$ observations.

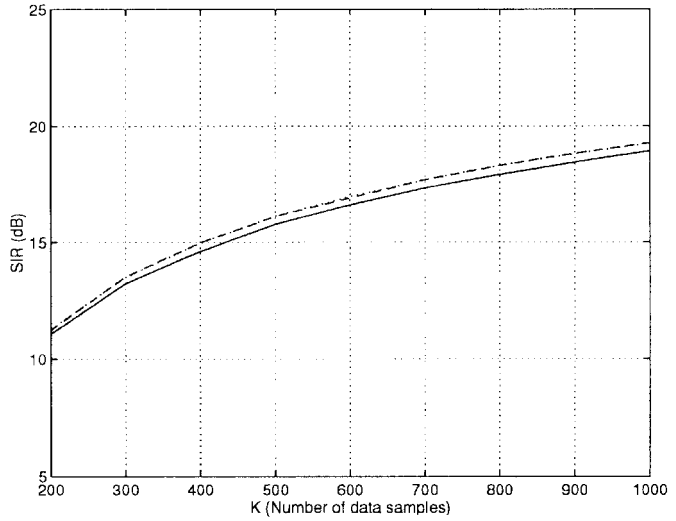



Fig. 7. MIMO: signal-to-interference ratios of User 1 (solid line), User 2 (dashed line), and User 3 (dashdot line).

The last set of simulations compares the performance of the XBM algorithm with and without the variation discussed for Step 2. Recall that the objective of the alternative algorithm is

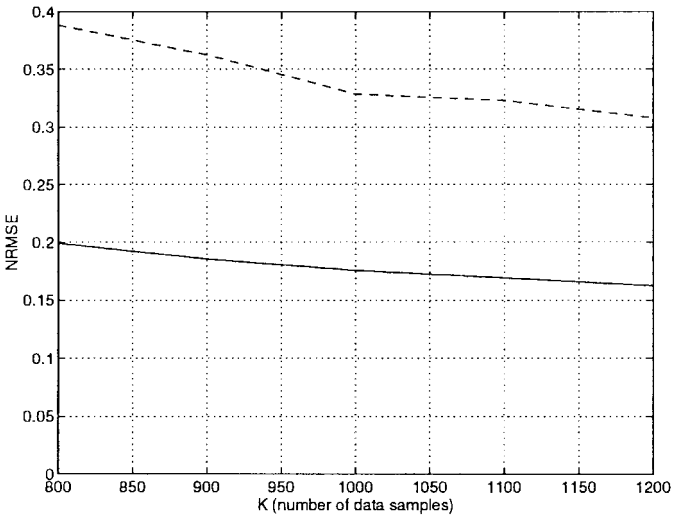



Fig. 8. MIMO: XBM original (solid line) and alternative (dashed line) algorithms as presented in this paper.

to lower the computational burden in high SNR environments. We have chosen one of the ten generated $\boldsymbol{H}$ GSM channel matrices. The lengths of the data packets range from $K_{\min} = 800$ to $K_{\max} = 1200$ in steps of $K_{\text{step}} = 100$. For each $K$, we ran $M = 100$ independent trials at SNR $= 40$ dB. We computed the NRMSE for both algorithms. Fig. 8 displays the results obtained. The dashed and solid lines refer to the XBM algorithm with and without the variation, respectively. As expected, the alternative algorithm presents a higher error in the estimate. Also, we have discarded the runs at which the alternative method fails to correctly estimate the right delay spreads $L_p$ of the users. This situation occurred for $K = 800, 900, 1000, 1100$, and 1200 in 3.0, 1.5, 1.5, 1.0, and 1.0% of the runs, respectively (as compared to 11.0, 10.0, 7.5, 7.0, and 5.0% if the estimates in (17) were not subject to the improvement discussed in Section IV-A).

## VI. Conclusions

We derived XBM, a closed-form solution for the blind identification of MIMO-FIR systems driven by digital sources. It uses spectral diversity obtained by coloring the output of each source prior to transmission. The receiver exploits the spectral diversity to achieve the blind identification of the unknown MIMO channel matrix. The algorithm is in closed form, avoiding the convergence problems of iterative approaches, and uses only second-order statistics, making it suitable for wireless communications. The penalty is the added computational complexity, which is mitigated by its parallelizable structure. The algorithm separates the MIMO sources and equalizes the corresponding channels. The approach is based on synchronization of the replicas delayed by the multipath and on their coherent recombination to attain higher SNR.

Computer simulations illustrated the good performance of XBM. In the context of a single-user, when compared to TXK, XBM produces smaller biases and variances in the estimated channels, which results in a significant decrease of the probability of error at the equalizer's output. For multiusers, XBM is asymptotically exact both in terms of the mean-square error of the channels estimates and in its ability in suppressing interuser interference crosstalk.

## Appendix
## Proof of Theorem 1

Let $\boldsymbol{R_x}(l;\boldsymbol{H}) = \boldsymbol{H}\boldsymbol{R_s}(l)\boldsymbol{H}^H + \boldsymbol{R_w}(l)$ be the output autocorrelation matrices of a MIMO system parameterized by the channel matrix $\boldsymbol{H} = [\boldsymbol{H}_1\ \boldsymbol{H}_2\ \cdots\ \boldsymbol{H}_P]$. It is understood that (4) models the system and conditions A1)–A3) hold. Let $\tilde{\boldsymbol{H}}$ be a channel matrix satisfying $\boldsymbol{R_x}(l;\boldsymbol{H}) = \boldsymbol{R_x}(l;\tilde{\boldsymbol{H}})$ for $l = 0, L_c$, and $L_c + 1$, i.e., that (10)–(12) hold.

Since $\boldsymbol{R_s}(0) = \boldsymbol{I}_L$, the equality in (10) is just $\boldsymbol{H}\boldsymbol{H}^H = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^H$ and thus $\tilde{\boldsymbol{H}} = \boldsymbol{H}\boldsymbol{Q}^H$, where $\boldsymbol{Q}$ is an $L \times L$ unitary matrix. By using this representation in (11) and (12), and summing both sides of these two equalities, we have $\boldsymbol{Q}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{Q}$, where $\boldsymbol{A} \equiv \boldsymbol{R_s}(L_c) + \boldsymbol{R_s}(L_c + 1) = \operatorname{diag}[\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_P]$ is a block diagonal matrix, and $\boldsymbol{A}_p = \eta_p(\boldsymbol{I}_{L_p} + \boldsymbol{K}_{L_p})$. Due to its structure, $\boldsymbol{A}$ is also a nonderogatory matrix [32] (the linear dimension of every eigenspace is one). Since $\boldsymbol{Q}$ commutes with $\boldsymbol{A}$, there exists a polynomial $P(t)$ such that $\boldsymbol{Q} = P(\boldsymbol{A})$ [32]. Thus, $\boldsymbol{Q} = \operatorname{diag}[\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_P]$ is also a block diagonal matrix, where the $p$th block on the diagonal is given by $\boldsymbol{Q}_p = P(\boldsymbol{A}_p)$. Notice that each $\boldsymbol{Q}_p$ is by itself an $L_p \times L_p$ unitary matrix. It is straightforward to verify that every power of $\boldsymbol{A}_p$ is a lower-triangular Toeplitz matrix, i.e., it can be written as

$$\begin{bmatrix} a_1 & & & & \\ a_2 & a_1 & & & \\ a_3 & a_2 & a_1 & & \\ \vdots & \ddots & \ddots & \ddots & \\ & \cdots & a_3 & a_2 & a_1 \end{bmatrix}.$$

Since this structure is preserved under scalar multiplications and matrix additions, we conclude that each $\boldsymbol{Q}_p$ has that same structure. Using the fact that $\boldsymbol{Q}_p$ is also a unitary matrix, it follows that $\boldsymbol{Q}_p = e^{j\theta_p}\boldsymbol{I}_{L_p}$. Therefore, $\tilde{\boldsymbol{H}}_p = \boldsymbol{H}_p e^{j\theta_p}$, and Theorem 1 is proved.

## References

[1] V. A. N. Barroso and J. M. F. Moura, "Beamforming with correlated arrivals in mobile communications," in V. Madisetti and D. Williams, Eds., *Digital Signal Processing Handbook*, invited paper, chapter on Array Signal Processing, M. Kaveh, Ed. Boca Raton, Fl: CRC Press, 1997.

[2] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.

[3] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communications systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867–1875, Nov. 1980.

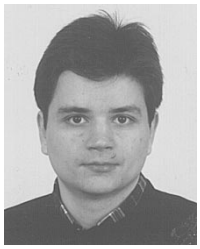
[4] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, pp. 871–883, Aug. 1984.

[5] Z. Ding, R. Kennedy, B. Anderson, and C. Johnson, "Ill-convergence of Godard blind equalizers in data communication systems," *IEEE Trans. Commun.*, vol. 39, pp. 1313–1327, Sept. 1991.

[6] D. Hatzinakos and C. Nikias, "Estimation of multipath channel response in frequency selective channels," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 12–19, Jan. 1989.

[7] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–320, Mar. 1990.

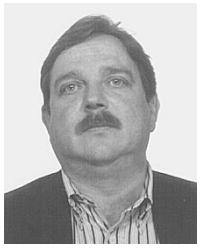
[8] J. Gomes and V. A. N. Barroso, "A super-exponential algorithm for blind fractionally spaced equalization," *IEEE Signal Processing Lett.*, vol. 3, pp. 283–285, Oct. 1996.
[9] H. Liu and G. Xu, "Closed-form blind symbol estimation in digital communications," *IEEE Trans. Signal Processing*, vol. 43, pp. 2714–2723, Nov. 1995.
[10] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.
[11] D. Slock, "Blind fractionally-spaced equalization, perfect- reconstruction filter-banks and multichannel linear prediction," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'94)*, vol. 4, pp. 585–588.
[12] K. Abed-Meraim, E. Moulines, and P. Loubaton, "Prediction error method for second-order blind identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 694–705, Mar. 1997.
[13] D. Gesbert and P. Duhamel, "Robust blind channel identification and equalization based on multi-step predictors," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'97)*, vol. 5, pp. 3621–3624.
[14] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *IEEE Signal Processing Lett.*, vol. 1, pp. 29–31, Feb. 1994.
[15] A. van der Veen, S. Talwar, and A. Paulraj, "Blind estimation of multiple digital signals transmitted over FIR channels," *IEEE Signal Processing Lett.*, vol. 2, pp. 99–102, May 1995.
[16] B. Halder, B. Ng, A. Paulraj, and T. Kailath, "Unconditional maximum likelihood approach for blind estimation of digital signals," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'96)*, vol. 2, pp. 1081–1084.
[17] V. A. N. Barroso, J. M. F. Moura, and J. M. F. Xavier, "Blind array channel division multiple access (AChDMA) for mobile communications," *IEEE Trans. Signal Processing*, vol. 46, pp. 737–752, Mar. 1998.
[18] J. M. F. Xavier and V. A. N. Barroso, "Blind SDMA with a signal space geometry matched algorithm," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'97)*, vol. 5, pp. 3437–3440.
[19] ———, "Blind source separation, ISI cancellation and carrier phase recovery in SDMA systems for mobile communications," *Wireless Personal Commun.*, special issue on Wireless Broadband Communications, to be published.
[20] K. Anand, G. Mathew, and V. Reddy, "Blind separation of multiple co-channel BPSK signals arriving at an antenna array," *IEEE Signal Processing Lett.*, vol. 2, pp. 176–178, Sept. 1995.
[21] C. Papadias and A. Paulraj, "A space-time constant modulus algorithm for SDMA systems," in *Proc. IEEE Vehicular Technology Conf.*, Atlanta, GA, 1996, pp. 86–90.
[22] K. Abed-Meraim and Y. Hua, "Blind identification of multi-input multi-output system using minimum noise subspace," *IEEE Trans. Signal Processing*, vol. 45, pp. 254–258, Jan. 1997.
[23] J. Cardoso, "Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'91)*, vol. 5, pp. 3109–3112.
[24] Z. Ding, "Blind wiener filter estimation for multi-channel systems based on partial information," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'97)*, vol. 5, pp. 3609–3612.
[25] J. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
[26] G. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-8, pp. 363–378, May 1972.
[27] R. Steele, *Mobile Radio Communications*. Pentech Press, 1992.
[28] B. Porat, *Digital Processing of Random Signals*. Prentice-Hall International, 1994.
[29] D. Gerlach and A. Paulraj, "Base station transmitting antenna arrays for multipath environments," *Signal Processing*, no. 54, pp. 59–73, 1996.
[30] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-33, pp. 387–392, Apr. 1985.
[31] G. Golub and C. van Loan, *Matrix Computations*. Baltimore, MD: John Hopkins, 1983.
[32] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge University Press, 1990.
[33] R. Behrens and L. Scharf, "Signal processing applications of oblique projections operators," *IEEE Trans. Signal Processing*, vol. 42, pp. 1413–1424, June 1994.
[34] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.



**João M. F. Xavier** (S'97) was born in Bumba, Zaire, on February 2, 1972. He received the E.C.E. degree in 1995 and the M.Sc. degree in electrical and computer engineering in 1997, both from the Instituto Superior Técnico (IST), Lisbon, Portugal. He is currently pursuing the Ph.D. degree at IST.

In 1997, he joined the faculty of the Department of Electrical and Computer Engineering at IST. He is with the Signal and Image Processing Group at the Institute of Systems and Robotics (ISR), Lisbon, Portugal. His research interests are in statistical signal array processing, performance analysis, digital communications, and blind equalization.

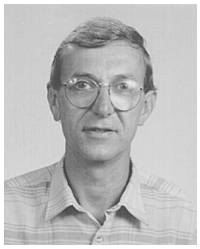
Mr. Xavier is affiliated with the IEEE Signal Processing Society.



**Victor A. N. Barroso** (M'89) was born in Lisbon, Portugal, on December 22, 1952. He received the E.E. degree in 1976 and the Ph.D. degree in electrical and computer engineering in 1990, both from the Instituto Superior Técnico (IST), Lisbon, Portugal.

In 1976, he joined the faculty of the Department of Electrical and Computer Engineering at IST. He was promoted to Assistant Professor in 1990 and to Associate Professor in 1993. He has taught courses in systems and signal theory, control systems, signal and array processing, computer networks, and communications. He is also a Researcher of the Signal and Image Processing Group at the Institute of Systems and Robotics (ISR), Lisbon, Portugal. His research interests are in statistical signal/image and array processing, communications, and underwater acoustics.

Dr. Barroso has been an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (Underwater Acoustics) since 1996. He is affiliated with several IEEE societies.



**José M. F. Moura** (S'71–M'75–SM'90–F'94) received the engenheiro electrotécnico degree in 1969 from Instituto Superior Técnico (IST), Lisbon, Portugal, and the M.Sc., E.E., and the D.Sc. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 1973 and 1975, respectively.

He is presently a Professor of Electrical and Computer Engineering at Carnegie Mellon University (CMU), Pittsburgh, PA, which he joined in 1986. Prior to this, he was on the faculty of IST where he was an Assistant Professor (1975), Professor Agregado (1978), and Professor Catedrático (1979). He has had visiting appointments at several institutions, including MIT (Genrad Associate Professor of Electrical Engineering and Computer Science, 1984–1986) and the University of Southern California (research Scholar, Department of Aerospace Engineering, Summers 1978–1981). His research interests include statistical signal processing (one- and two-dimensional), digital communications, image and video processing, radar and sonar, and multiresolution techniques. He has organized and codirected two international scientific meetings on signal processing theory and applications. He has over 190 published technical contributions and is co-editor of two books: *Nonlinear Stochastic Problems* (Reidal, 1983) and *Accoustic Signal Processing for Ocean Exploration* (Kluwer, 1993).

Dr. Moura is currently the Editor-in-Chief for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and an elected member of the Board of Governors of the IEEE Signal Processing Society. He is a member of the IEEE Signal Processing Society Publications Board, and of the IEEE Signal Processing Society Technical Committees on Underwater Accoustics and on MultiMedia Signal Processing. He was a member of the IEEE Press Board (1991–1995), a Technical Associate Editor for the IEEE SIGNAL PROCESSING LETTERS (1993–1995), and an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1988–1992). He was a program committee member for the IEEE International Conference on Image Processing (ICIP'95) and for the IEEE International Symposium on Information Theory (ISIT'93). He is a corresponding member of the Academy of Sciences of Portugal (Section of Sciences). He is affiliated with several IEEE societies, Sigma Xi, AMS, IMS, and SIAM.