Hybrid Higher-Order Statistics Learning in Multiuser Detection

Antonio J. Caamaño, *Member, IEEE*, Rafael Boloix-Tortosa, Javier Ramos, *Member, IEEE*, and Juan José Murillo-Fuentes, *Member, IEEE*

***Abstract*—In this paper, we explore the significance of second- and higher-order statistics learning in communication systems. The final goal in spread-spectrum communication systems is to receive a signal of interest completely free from interference caused by other concurrent signals. To achieve this end, we exploit the structure of the interference by designing second-order statistics detectors, such as the minimum square error, in conjunction with higher-order statistics (HOS) techniques, such as the blind source separation (BSS). This hybrid higher-order statistics (HyHOS) approach enables us to alleviate BSS algorithms of one of their main problems, that is, their sensitiveness to high levels of noise. In addition, we benefit from remarkable properties of BSS in learning such as fast learning (superefficiency) and independence of the initial settings of the problem (equivariance). We successfully applied the results of this approach to the design of multiuser detectors in code-division multiple access channels.**

***Index Terms*—Array signal processing, blind source separation, code-division multiple access (CDMA), higher-order statistics (HOS), independent component analysis, unsupervised learning.**

I. Introduction

INTERFERENCE limitation due to the simultaneous access of multiple users has been the stimulus to the development of a powerful family of signal processing techniques, namely multiuser detection (MUD). These techniques have been extensively applied to direct-sequence code-division multiple access (DS-CDMA) systems. Thus, most of last generation digital communication systems such as global positioning system (GPS), wireless 802.11 b, Universal Mobile Telecommunication System (UMTS), etc., may take advantage of any improvement on this topic. In the blind case, the algorithms should cope not only with noise and the near-far problem but with no training sequence available. In this sense, the minimum mean square error (MMSE) criteria [1] provides a good blind linear solution to the problem. Due to its computational complexity, some other alternative algorithms have been proposed [2]–[4].

On the other hand, blind source separation (BSS) techniques allow us to reconstruct a set of nonobservable signals, regarded as sources, from mixtures of them. BSS has lately been a main field of research due to the great number of applications in different areas. BSS have been successfully applied (see [5] and references therein) to communications signals (mainly in array processing), biomedical signals such as electrocardiogram (ECG) or electroencephalogram (EEG), as an alternative to principal component analysis in image or financial data processing, in monitoring, etc… In addition, they have been lately applied to other problems such as spread-spectrum-based communications [6], [7], digital watermarking of images [8], audio spectrum basis functions computation [9], image classification, encoding, or compression [10]. Several methods have been proposed as solutions to the BSS problem. We bring out here two main approaches that rest on the assumption that sources are statistical independent. On the one hand, we have techniques based on the cancellation of estimation equations by using the natural-gradient (NG) steepest descent algorithm [11]–[13]. These methods usually have the maximum-likelihood (ML) approach as a starting point and are not usually robust as they need some knowledge on the probability density functions (PDF) [12], [14]. Here, robustness stands for the methods to be available for sources of any statistical distributions. In this paper, we will use the M-EASI algorithm [7] as it exhibits good properties in the separation of digital communication signals. On the other hand, we have contrast functions. Contrasts are cost functions whose minimization yields the solution to the BSS. Assuming zero-mean signals, these algorithms use higher-order statistics (HOS) in the computation of a unitary transformation to diagonalize higher-order cumulant tensors of the whitened (decorrelated) outputs. This tensor may be diagonalized by either canceling the whole set of cumulants out of the diagonal [minimization of the mutual information (MI)] or by the maximization of the absolute value of the diagonal entries (minimization of the marginal entropy (ME) [15]). The SICA [16] method used in this paper is based on this last approach.

Some of these BSS algorithms have been proposed as *fully* blind MUDs, as spreading codes are unknown [7], [17]–[19]. However, these blind detectors usually are quite sensitive to noise and computationally expensive. In this paper, we propose to introduce BSS [5], [16] into the structure of the MMSE MUD to exploit its proven noise robustness, inheriting BSS remarkable properties as equivariance [11], superefficiency [20], and, in some approaches, low computational cost. We will focus on the performance of BSS-MUD detectors based on the offline SICA algorithm and the online M-EASI [17] methods.

The paper is organized as follows. Section II summarizes the matrix model, main assumptions, and definitions for a DS-CDMA communication system. Linear MUDs will be

Manuscript received February 5, 2003; revised June 20, 2003. This work was supported by the Spanish Government under Contract MCYT-TIC2003-03781. This paper was recommended by Guest Editors J. L. Fernandez and A. Artés.

A. J. Caamaño and J. Ramos are with the DTSC, EPS-Telecomunicaciones, Universidad Carlos III de Madrid, Leganés 28911, Madrid, Spain (e-mail: ton@tsc.uc3m.es).

R. Boloix-Tortosa and J. J. Murillo-Fuentes are with ATSC, Escuela Superior de Ingenieros, Universidad de Sevilla, Sevilla 41092, Spain (e-mail: murillo@esi.us.es).

Digital Object Identifier 10.1109/TSMCC.2004.833299

also addressed. Then, BSS techniques are introduced in Section III, where we will focus on the SICA algorithm and natural gradient-based methods such as the M-EASI. Section IV is devoted to relate both of these approaches BSS and the MMSE to propose new HOS-based detectors. These novel MUDs are referred to throughout this paper as hybrid high-order statistics (HyHOS) detectors. In Section V, we focus on the characterization of these detectors. Section VI includes experiments to compare MMSE with the HyHOS developed in this work and to show their noise resistance and superefficient learning. Section VII is devoted to conclusions.

II. Communication System Model: Linear Multiuser Detection

Consider a synchronous code-division multiple access (CDMA) transmitter [1]. The baseband discrete-time model at the chip rate for the received signal at symbol tth and chip lth, assuming a noisy channel, yields

$$x_t(l) = \sum_{i=1}^{N} \chi_i s_t(i) c_i(l) + n_t(l) \quad 1 \le l \le L \tag{1}$$

where L is the length of the spreading codes, N is the number of active users, $s_t(i)$ is the ith user's tth symbol, and χ_i is a scaling factor specified by the power control loop. Besides, the normalized spreading code of user i is denoted by $\boldsymbol{c}_i = [c_i(1), \ldots, c_i(L)]^T$ with each $c_i(\cdot) \in \pm 1$. Finally, $n_t(l)$ is white Gaussian noise added to the tth symbol at the lth chip. By using a discrete model, we assume that the transmitted pulse-shape filters are Nyquist and that the receiver uses a chip-rate-sampled matched filter. If this filter is followed by a serial-to-parallel converter, we have the following matrix model of the system:

$$\boldsymbol{x}_t = \boldsymbol{A}\boldsymbol{s}_t + \boldsymbol{n}_t \tag{2}$$

where $\boldsymbol{x}_t = [x_t(1), \ldots, x_t(L)]^{\mathrm{T}}$, $\boldsymbol{s}_t = [s_t(1), \ldots, s_t(N)]^{\mathrm{T}}$, $\boldsymbol{n}_t =$ is a $L \times 1$ vector with white Gaussian random entries, and $\boldsymbol{A}$ is a $L \times N$ matrix including the spreading codes and scaling factors. Hence

$$\boldsymbol{A} = \boldsymbol{H}\boldsymbol{\chi} \tag{3}$$

where the ith column of $\boldsymbol{H}$ is vector $\boldsymbol{c}_i$ and $\boldsymbol{\chi} = \mathrm{diag}([\chi_1, \ldots, \chi_N])$. The optimum receiver for this system when $N = 1$ is the matched filter (MF). However, if $N > 1$, the difference between powers of users is high enough and user's codes are not orthogonal, we have the near-far problem. Under these circumstances, other well-known second-order statistics (SOS)-based linear detectors, such as the MMSE, have much better performance.

A. Linear Multiuser Detection

A linear multiuser detector $\boldsymbol{C}$ gives an estimation of the original transmitted signals as

$$\boldsymbol{y}_t = \boldsymbol{C}\boldsymbol{x}_t = \boldsymbol{C}\boldsymbol{A}\boldsymbol{s}_t + \boldsymbol{C}\boldsymbol{n}_t. \tag{4}$$

If we know the number of users N and their spreading codes, we may use this information in the design of the detector. This is the so-called *centralized* (i.e., base station) detector. On the contrary, if we have access just to the spreading code of the desired user, we have the *noncentralized* (i.e., user equipment) detector. The centralized detector minimizing the MSE for each user i, $\mathrm{MSE}_i = \mathrm{E}[|y_t(i) - s_t(i)|^2]$, yields

$$\boldsymbol{C}_{\mathrm{MMSE}}^{(c)} = \boldsymbol{R}_v^{-1}\boldsymbol{H}^* = \boldsymbol{H}^*\boldsymbol{H}\boldsymbol{W}^*\boldsymbol{W}\boldsymbol{H}^* \tag{5}$$

where $\boldsymbol{v}_t = \boldsymbol{H}^*\boldsymbol{x}_t$ is the output of the MF and $\boldsymbol{W}$ is the whitening matrix for vector $\boldsymbol{v}$, $\mathrm{E}[\boldsymbol{W}\boldsymbol{v}(\boldsymbol{W}\boldsymbol{v})^*] = \boldsymbol{I}$, and $\boldsymbol{R}_v$ its covariance matrix. Here, as in the following, * denotes transpose-conjugate or simply conjugate for scalar values. Besides, the noncentralized MMSE detector in matrix form is defined as follows:

$$\boldsymbol{C}_{\mathrm{MMSE}}^{(nc)} = \boldsymbol{H}^*\boldsymbol{W}^*\boldsymbol{W} \tag{6}$$

where $\boldsymbol{W}$ is the whitening matrix for vector $\boldsymbol{x}_t$. Notice that although we have matrix $\boldsymbol{H}$ in (6) with the whole set of spreading codes, each of its columns leads us to a different user's detector.

In the context of the BSS, we propose to compute the detector that minimizes statistical dependence at the output by using HOS. The application of BSS approaches to fully blind MUDs is almost immediate [7], [18], [19]. However, they usually involve a high computational complexity as they operate on the chip space (L) rather than on the user subspace (N). Besides, they do not cope with noise. In this paper, we assume we have the active user's spreading codes available and we face the use of this knowledge in the development of blind HOS-BSS-based detectors [21]. We first introduce these HOS-BSS techniques.

III. Blind Separation of Sources

A. Main Assumptions and Model

BSS involve the task of obtaining a nonobservable set of signals, the so-called sources, from another set of observable signals regarded as mixtures or observations. Here, the adjective "blind" stands for the fact that neither the original sources nor the mixture itself are known. Usually, and in the context of this paper, the assumption of spatial statistical independence is the key to achieve separation. If we deal with temporally white signals, the use of HOS is mandatory [15]. In statistical terms, we aim to compute the projection of a set of components (mixtures in BSS) that minimizes the statistical dependence of the outputs [15], [22]. This is the independent component analysis (ICA) of the observations.

In its simplest form, the BSS reduces to the following matrix form. The $m = n$ mixtures sampled at time t, $\boldsymbol{x}_t$, are instantaneous linear combinations of the n sources $\boldsymbol{s}_t$, that is

$$\boldsymbol{x}_t = \boldsymbol{A}\boldsymbol{s}_t \tag{7}$$

where $\boldsymbol{A}$ is the mixing or transfer matrix. If $\boldsymbol{x}_t$ is a stationary ergodic random sequence and the mixing matrix $\boldsymbol{A}$ is nonsingular, it is possible to estimate a separation matrix $\boldsymbol{B}$ to obtain the sources as

$$\boldsymbol{y}_t = \boldsymbol{B}\boldsymbol{x}_t = \boldsymbol{B}\boldsymbol{A}\boldsymbol{s}_t = \boldsymbol{C}\boldsymbol{s}_t. \tag{8}$$

In BSS, we compute matrix $\boldsymbol{B}$ so that $\boldsymbol{C}$ is ideally the identity matrix. However, the original scaling and arrangement cannot be estimated from the independence assumption. In this sense, $\boldsymbol{C}$ is a nonmixing matrix if it has one and only one nonzero entry in each column and each row. Matrix $\boldsymbol{B}$ can be decomposed into the product of a whitening $\boldsymbol{W}$ and a unitary $\boldsymbol{V}$ matrix. The whitening stage gives us signals $z_t(i)$

$$\boldsymbol{y}_t = \boldsymbol{B}\boldsymbol{x}_t = \boldsymbol{V}\boldsymbol{W}\boldsymbol{A}\boldsymbol{s}_t = \boldsymbol{V}\boldsymbol{z}_t. \tag{9}$$

In addition to the source independence hypothesis, there are other minor considerations. First, although the instantaneous mixture model considered here applies to many of the applications referred at the Introduction, it may be extended to the convolutive case [23], [24] when needed. Besides, we may consider here the case where the number of mixtures m is greater or equal to the number of sources n (i.e., $m \geq n$). In the case $m > n$, a subspace approach, such as a singular value decomposition (SVD), may be used to project the mixture space onto the signal subspace, reducing the effect of weak noises. For noisy mixtures, see [25]–[27]. In addition, a necessary and sufficient condition for the waveform-preserving source estimation to be feasible is that no more than one Gaussian distributed source be present in the mixture [15], [25], [28].

The model in (7) for instantaneous BSS may be easily identified with narrowband m-sensor linear-array applications in communications and particularly with the synchronous-CDMA case with spreading factor $m = L$ and $n = N$ users. In synchronous CDMA communications, matrix $\boldsymbol{A}$ may be decomposed as in (3). Thus, the mixing matrix includes the spreading codes and different powers of the signals to estimate the sources. Notice that BSS directly applied to CDMA would lead to a detector where neither a training sequence nor the spreading codes themselves are needed. This detector is referred to as *fully* blind MUD. Note that the noise vector in (2) is not included in the BSS model in (7). If noise is high enough, the performance of the BSS based on this model deteriorates.

Adaptive solutions to this problem are usually based on maximum likelihood (ML) and the natural (or relative) gradient [11], [29], [30]. On the contrary, most of the offline solutions to ICA [31] minimize one criterion, contrast function, or cancellation of multiple criteria. In this sense, we consider here the contrast function based on the minimization of the marginal entropies [15].

B. Minimum Marginal Entropy-Based Contrast

A contrast function $\phi(\cdot)$ maps $\mathcal{Y}$, the set of random vectors $\boldsymbol{y}$ (multivariate PDFs), on $\mathbb{R}$. It has a minimum when the entries of $\boldsymbol{y}$ are statistically independent (i.e., if $\boldsymbol{y}$ has independent component $\phi(\boldsymbol{y}) \leq \phi(\boldsymbol{A}\boldsymbol{y})\, \forall \boldsymbol{A}$ nonsingular). Thus, the minimization of a contrast function yields the solution to the BSS problem. If the number of sources $n > 2$, these contrasts are functions that are difficult to minimize. The "Jacobi optimization" (JO) [15] consists of solving the n-dimensional problem by decomposing it in a set of 2-dimensional (2-D) optimizations. Hence, the rotation matrix $\boldsymbol{V}$ in (9) is decomposed into $g = m(m-1)/2$ Givens rotations. Minimizing the contrast for each angle $\theta_h, h = 1, \ldots, g$ drives the system to the solution. We will first focus on these 2-D contrasts.

Under the whiteness constraint (i.e., $\mathrm{E}[\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}] = I$), it yields the following orthogonal (denoted by ϕ_2) contrast [22]:

$$\phi_2^{ME}(\boldsymbol{y}) \stackrel{c}{=} \sum_{i=1}^{n} \mathrm{H}[y_i] \tag{10}$$

where the criteria to minimize is the marginal entropy $\mathrm{H}[\boldsymbol{y}]$. Here, as in the following, $\stackrel{c}{=}$ means equality up to an additive constant c. Besides, we will denote by $\mu_{ijkl}^{y} = \mathrm{E}[y_i y_j^* y_k y_l^*]$ and $\mu_{ij}^{y} = \mathrm{E}[y_i y_j^*]$ the fourth-order and second-order moments. If we approximate the possible distributions for $\boldsymbol{s}$ by and Edgeworth expansion [32], rewrite the result in terms of second-order and fourth-order cumulants of the zero-mean outputs $C_{ijkl}^{\boldsymbol{y}} = \mu_{ijkl}^{y} - \mu_{ij}^{y}\mu_{kl}^{y} - \mu_{ik}^{y}\mu_{jl}^{y} - \mu_{il}^{y}\mu_{jk}^{y}$, and then minimize it for all possible distributions [15], it follows that:

$$\phi_{24}^{ME}(\boldsymbol{y}) \approx \frac{1}{48}\phi_{24}^{ME}(\boldsymbol{y}) = -\frac{1}{48}\sum_i \left(C_{iiii}^{y}\right)^2 \tag{11}$$

where $C_{iiii}^{y} = \mu_{iiii}^{y} - 3\mu_{ii}^2$ is the autocumulant of output y_i. Thus, the minimization of the fourth-order cumulant-based ME contrast yields the diagonalization of the fourth-order cumulant tensor when at most one marginal fourth-order cumulant is null. In the following, we will face the real case. The extension to complex-valued sources is somehow immediate [33].

By using complex (polar) notation $\check{\boldsymbol{y}}_t = y_t(p) + jy_t(q) = r_t e^{j\rho_t}$ and under the whitening constraint, the independent components $y_t(p)$ and $y_t(q)$ in the 2-D approach yield

$$\check{\boldsymbol{y}}_t = r_t e^{j\rho_t} = r_t e^{j(\theta+\beta_t)} = e^{j\theta}\check{\boldsymbol{z}}_t \tag{12}$$

where $j = \sqrt{-1}$. The whitened mixture vector $\check{\boldsymbol{z}}_t = z_t(p) + jz_t(q) = r_t e^{j\beta_t}$ is a simple rotation of the normalized (unit variance) sources $\check{\underline{\boldsymbol{s}}}_t = \underline{s}_t(p) + j\underline{s}_t(q) = r_t e^{j\alpha_t}$. Notice that at the solution $\rho_t = \theta + \beta_t = \alpha_t + k\pi/2,\ k = 0, 1, 2, \ldots$. The estimate of the rotation angle θ minimizing 2-D contrasts may be easily expressed as a closed function of the following complex-valued linear combinations (*centroids*) of the statistics of the outputs [34]:

$$\xi_\gamma = \mathrm{E}\left[r_t^4 e^{j4\beta_t}\right] \tag{13}$$

$$\xi_\eta = \mathrm{E}^2\left[r_t^4 e^{j2\beta_t}\right] \tag{14}$$

$$\gamma = \mathrm{E}\left[r_t^4\right] - 8. \tag{15}$$

Based on the so-called "weighted estimators" (WEs) or WAML [35], [36] a general 2-D estimator yields

$$\hat{\theta}_{GWE}(\omega_\gamma, \omega_\xi) = \frac{1}{4}\angle\left(\omega_\xi\omega_\gamma\xi_\gamma + (1-\omega_\xi)\xi_\eta\right)$$
$$0 < \omega_\xi < 1, \quad \omega_\gamma = \{\pm 1, \gamma\}. \tag{16}$$

This estimates yields a wide variety of estimators such as the EML in [34] $\hat{\theta}_{\text{EML}} = \hat{\theta}_{\text{GWE}}(\text{sign}(\gamma), 1)$ or the [37], MK [31], [38], SKSE and SKDE [39], ML [40] estimates $\hat{\theta}_{\text{GWE}}(\pm 1, 1)$, $\hat{\theta}_{\text{MaSSFOC}} = \hat{\theta}_{\text{GWE}}(\gamma, 1/2)$ in [39], or the $\hat{\theta}_{\text{AML}} = \hat{\theta}_{\text{GWE}}(\gamma, 1/3)$ [35]. The discussion on the optimum w_ξ [36] is open as it is difficult to compute the statistics of these estimates and analyze them for all possible PDFs. In this paper, we propose to use

$$\theta_{\text{SICA}} = \hat{\theta}_{\text{GWE}}\left(\gamma, \frac{3}{7}\right) \tag{17}$$

as it may be proved [16] that the SICA contrast function is equivalent to the ME approach $\phi_{24}^{\text{ME}}(\boldsymbol{y})$ in (11). Thus, it provides a solution to the ICA problem with the only condition of no more than one source with null fourth-order statistics in the mixture. The minimization of $\phi_{\text{SICA}}(\theta)$ is immediate as the solution is the phase of the resulting function, a sinusoid. This may be easily implemented by a lookup table so that we minimize the number of operations needed.

The 2-D case above may be easily extended to n dimensions by using the JO [15]. We have rewritten the algorithm using $\phi_{SICA}(\theta)$. Such an algorithm can be summarized as follows:

Algorithm 1: n-dimensional SICA using Jacobi optimization

1) *Whitenning*. Compute a whitening matrix $\boldsymbol{W}$ and the output vector $\boldsymbol{z}_t = \boldsymbol{W}\boldsymbol{x}_t$. Set $c = 1$ and $\boldsymbol{y}_t = \boldsymbol{z}_t$.

2) *Sweep* c. For all $g = n(n-1)/2$ pairs, i.e., for $1 \le p < q \le n$, do

   (a) Compute the Givens angle $\theta = \theta_{SICA}(\boldsymbol{y})$ in (17) with $[z_t(p), z_t(q)]^{\text{T}} = [y_t(p), y_t(q)]^{\text{T}}$.

   (b) If $\theta > \theta_{min}$, do rotate the pair $(y_t(p), y_t(q))$ by θ.

3) *End?* If the number of sweeps c satisfies $c \ge K = 1 + \sqrt{n}$ or no Givens angle has been updated, stop. Otherwise, go to step 2 for another sweep with $c = c + 1$.

C. Natural Gradient in BSS

The steepest descent method updates $\boldsymbol{C}$ according to the direction of the gradient ∇L of a loss function $L(\boldsymbol{C}) = \text{E}[l(\boldsymbol{C})]$. The natural [24] or relative gradient [5] proposes to use $\widetilde{\nabla} L(\boldsymbol{C}) = \nabla L(\boldsymbol{C})\boldsymbol{C}^{\text{T}}\boldsymbol{C}$. The stochastic version uses the instantaneous value $\nabla l(\boldsymbol{C})$. The learning law yields

$$\boldsymbol{C} \longleftarrow \boldsymbol{C} - \lambda \nabla l(\boldsymbol{C})\boldsymbol{C}^{\text{T}}\boldsymbol{C}. \tag{18}$$

In this paper, we assume statistical independence at the outputs $\boldsymbol{y}_t$. ML is an extended technique to derive a loss function for this criteria [5], [24]. It can be shown that we achieve independence at the output by canceling the estimating function

$$\boldsymbol{K}(\boldsymbol{y}) = \nabla l(C) \cdot \boldsymbol{C}^T = \boldsymbol{\varphi}(\boldsymbol{y})\boldsymbol{y}^{\text{T}} - \boldsymbol{I} \tag{19}$$

where in the ML approach, the ith entry of vector $\boldsymbol{\varphi}$ yield $\varphi_i(y(i)) = -q_i'(y(i))/q_i(y(i))$, being $q_i(\cdot)$ the probability density function of the ith source and φ_i its corresponding score function. However, as source distributions are supposed unknown, each author introduces his own activation function $\varphi_i(y(i))$. A family of them may be found, for sources with negative or positive kurtoses, in [41]. Here, as in the following, the temporal reference $_t$ has been removed for the sake of simplicity. The learning law to cancel (19) yields

$$\boldsymbol{C} \longleftarrow \boldsymbol{C} - \lambda \boldsymbol{K}(\boldsymbol{y})\boldsymbol{C}. \tag{20}$$

By normalizing (20), the learning law may be written as follows

$$\boldsymbol{C} \longleftarrow \boldsymbol{C} - \lambda \frac{\boldsymbol{\varphi}(\boldsymbol{y})\boldsymbol{y}^{\text{T}} - \boldsymbol{I}}{1 + \lambda\left|\boldsymbol{\varphi}^{\text{T}}(\boldsymbol{y})\boldsymbol{y}\right|}\boldsymbol{C}. \tag{21}$$

In [17], the authors proposed a NG-based algorithm to separate signals in digital communication, the M-EASI (Median-Equivariant Adaptive Separation via Independence). This method assumes zero-mean, symmetric, "circularly distributed" complex signals and introduces the sign function to reduce the bias introduced by noise. Its estimating function yields

$$\boldsymbol{K}(\boldsymbol{y}) = \frac{\boldsymbol{y}\text{sgn}(\boldsymbol{y})^* - I}{1 + \lambda\text{sgn}(\boldsymbol{y})^*\boldsymbol{y}} + \frac{1}{\alpha}\frac{\boldsymbol{\varphi}\text{sgn}(\boldsymbol{y})^* - \text{sgn}(\boldsymbol{y})\boldsymbol{\varphi}^*}{1 + \lambda|\boldsymbol{y}^*\boldsymbol{\varphi}|} \tag{22}$$

where $\text{sgn}(\boldsymbol{y}) = \text{sgn}(\Re(\boldsymbol{y})) + j\text{sgn}(\Im(\boldsymbol{y}))$, $\boldsymbol{y} \in \mathbb{C}$. With this method [30], we improve the stability of the algorithm, provide the method with phase recovering properties, and make the method more robust against noise.

IV. Hybrid High-Order Statistics Multiuser Detectors

We could use the natural gradient to compute a detector $\boldsymbol{C}$ by imposing some criteria such as the MSE. But we may directly exploit the MMSE structure in (5) for the centralized case. In this sense, we propose first a BSS-based solution by using the SICA algorithm presented in Section III-B. Then, we will introduce the natural gradient-based detector.

A. SICA-Based MUD: HyHOS I

We propose to substitute the matrix product $\boldsymbol{H}^*\boldsymbol{H}\boldsymbol{W}^*\boldsymbol{W}$ in (5) by a matrix $\boldsymbol{B}$ that makes the outputs $y(i)$ as statistically independent as possible. Attending to the matrix product, this is the BH detector, referred to as hybrid higher-order statistics (HyHOS) detector

$$\boldsymbol{C}_{\text{HyHOS}} = \boldsymbol{C}_{\text{BH}} = \boldsymbol{B}\boldsymbol{H}^* = \boldsymbol{V}\boldsymbol{W}\boldsymbol{H}^* \tag{23}$$

where $\boldsymbol{B}$ is a separating matrix. Notice that in the centralized case, the dimensional reduction is carried out by the matched filter $\boldsymbol{H}^*$. As the spreading codes are usually available (at least at the base station), it is straightforward to introduce them as a subspace algorithm at a null complexity cost [3]. In fact, we have no information loss in this subspace projection [1] and it allows us to cope with noise.

By computing the separating matrix $\boldsymbol{B}$ in (23) with the SICA algorithm in (17) [16], we have the HyHOS I. When applying SICA to the projected signal $\boldsymbol{v} = \boldsymbol{H}^*\boldsymbol{x}$, we first use SOS in the whitening stage. This gives us $\boldsymbol{z} = \boldsymbol{Wv}$. This matrix may be computed as in the MMSE algorithm (e.g., by using the singular value decomposition of the covariance matrix). Then, this algorithm uses JO to compute the unitary matrix $\boldsymbol{V}$ so that the contrast function in (11) with $\boldsymbol{y} = \boldsymbol{VWv}$ is minimized. Thus, this last stage may be seen as a fourth-order decorrelation. We go further than the second-order-based MMSE to find a new HOS-based solution to the problem. Notice that the JO-SICA Algorithm 1 is an offline method that computes matrix $\boldsymbol{B}$ from a whole set of observations, similar to the MMSE method.

On the other hand, it is not possible to define a BSS-based architecture for the noncentralized MUD similar to that of the MMSE in (6). Suppose the candidate now to be the detector

$$\boldsymbol{C}_{HB} = \boldsymbol{H}^*\boldsymbol{B} = \boldsymbol{H}^*\boldsymbol{VW}. \tag{24}$$

Here, matrix $\boldsymbol{C}_B = \boldsymbol{B}$ is already a solution, as independence is imposed at the outputs. Thus, $\boldsymbol{C}_{HB}$ is not a detector as $\boldsymbol{H} \neq \boldsymbol{I}$. Computing matrix $\boldsymbol{C}_B$ is basically a blind separation problem and it is out of the scope of this paper. Notice that this detector is fully blind as it does not use the spreading codes [7], [17], [18], but rather estimates them.

B. Natural Gradient-Based MUD: HyHOS II

In this subsection, we use the NG in the design of BSS-based MUDs, the NG-MUD. Notice that the natural gradient may be used in the whitening (SOS) process by setting $\varphi(y) = y$ in (19). The resulting detectors [21] are centralized and noncentralized adaptive MMSE MUDs. The description of these detectors is out of the scope of this paper as we focus on NG and HOS applied to MUD.

Proceeding as in the previous subsection, we have the following centralized BSS-based detector algorithm. It is possible to substitute the matrix product $\boldsymbol{H}^*\boldsymbol{HW}^*\boldsymbol{W}$ in (5) by a separating matrix $\boldsymbol{B}$. Thus, we have $\boldsymbol{C}_{BH} = \boldsymbol{BH}^*$ as in (23) where $\boldsymbol{B}$ is computed as a BSS. In this section, the separating matrix $\boldsymbol{B}$ is computed by using the M-EASI algorithm

$$\boldsymbol{B} \longleftarrow \boldsymbol{B} - \lambda \boldsymbol{K}_B(\boldsymbol{y}) \cdot \boldsymbol{B} \tag{25}$$

where $\boldsymbol{K}_B(\boldsymbol{y})$ was given in (22), $\boldsymbol{y} = \boldsymbol{BH}^*\boldsymbol{x} = \boldsymbol{Bv}$, and $\boldsymbol{\varphi}(\boldsymbol{y})$ may be chosen as described in [41]. Thus, the MMSE structure is further enhanced by the properties of the M-EASI algorithm discussed in the next section. This detector will be referred to as the HyHOS II detector. Again, the noncentralized detector leads us to a fully blind MUD.

C. User of Interest

BSS based on statistical independence allow us to recover the sources (i.e., the users' symbols). However, as described in Section III, the original arrangement cannot be estimated from the independence assumption. Thus, each row of the C_{HyHOS} (either I or II) is a detector for one of the active users, but we do not know which one corresponds to the user of interest (UOI). As we face a blind detector where no training sequence is transmitted, the only information available is that of the spreading codes of each user. Thus, we propose to correlate the rows of $\boldsymbol{C}_{\text{HyHOS}}$ with the user of interest's spreading code. This leads us to the UOI's HyHOS detector.

V. Discussion

We include here a discussion on some theoretical aspects of the methods above. We focus on the near-far problem, noise, convergence, and complexity.

The main point in using the natural gradient in MUD is that it is *equivariant* [5] (i.e., the convergence has a uniform performance). Thus, the convergence of the HyHOS II does not depend on the mixing matrix $\boldsymbol{A}$. This matrix in the CDMA problem yields $\boldsymbol{A} = \boldsymbol{H\chi}$ as in (3). Similar conclusions may be drawn for the SICA-based HyHOS I method [36]. It can be concluded that the detectors proposed in this paper are near-far resistance, as convergence is independent of the user's amplitudes (i.e, they are *scale invariant*).

Algorithms in BSS usually do not cope with the noisy case whenever the number of sources and mixtures are the same $m = n$. On the other hand, if $m > n$ a signal subspace projection allows noise reduction. Previous BSS approaches to fully blind MUDs use a singular value decomposition (SVD) decomposition, a MPLL [18], ... However, this involves a higher computational complexity. As spreading codes are usually available (at least at the base station), it is straightforward to introduce them as a subspace algorithm at a null complexity cost [1], [4]. In this sense, the structure of the centralized MMSE detector has been exploited. In addition, by using the M-EASI algorithm, we combat the effect of noise in the separation process for digital communication signals [7], reducing the variance of the estimation.

Another important characteristic of natural gradient BSS techniques is that of superefficiency [20]. In this sense, provided $E[\varphi(y)] = 0$ (an usual case), the covariance between two outputs decreases of the order of $1/t^2$ in batch estimation and of the order of λ^2 in online learning. Furthermore, $\lambda = 1/t$ gives, asymptotically, the best performance, which is the same as the optimal batch estimator. Thus, with $\lambda = 1/t$ we achieve an online algorithm with batch features at every t. On the other hand, if adaptive features are needed, we may use $\lambda < 1$ to achieve an output covariance decreasing as λ^2. The HyHOS II inherits these properties. In the SICA algorithm used in the HyHOS I, the variance of the estimator in (17) decreases of the order of $1/t$ [36].

On the question of complexity, the MMSE solution at every time t involves computing the SVD of the covariance matrix of the inputs in the whitening stage. This leads us to a number of multiplications and accumulations (MACs) of the order $O(N^3)$, where N is the number of active users. Thus, the computational resources needed are significant. The SICA algorithm used in the HyHOS I aggravates this problem as after $\boldsymbol{W}$ we need to compute the unitary matrix $\boldsymbol{V}$ by using fourth-order moments. However, if we compute (20)–(22) in right-to-left matrix multiplication order, the number of MACs yields of the order $O(N^2)$. Hence, being the stochastic NG superefficient, the HyHOS II included in this paper may be designed as an online algorithm with solutions close to those of the offline methods at a lower computational cost.

VI. Experimental Results

The test environment used here is a synchronous CDMA system where the BPSK symbols were spread using GOLD sequences with spreading factor $L = 31$ chips. We consider here two scenarios. On the one hand, we face the typical evaluation of the performance in a digital communications system (i.e., bit error rates for $\mathrm{SNR} \geq 8\ \mathrm{dB}$). This is a low-noise environment (LNE). On the other hand, as the BSS problem is posed in the *absence* of noise, one of the main issues is to test the performance of the resulting algorithms in noisy environments. Therefore, we include an evaluation in a high noise environment (HNE) where we test the *noise resistance* capabilities of the HyHOS algorithms with bit error rates as low as $2 \times 10^{-1} \leq \mathrm{BER} < 5 \times 10^{-1}$ and equivalent signal-to-noise ratio (SNR) for the user of interest (UOI) of $-15 \leq \mathrm{SNR} \leq 0\ \mathrm{dB}$. Besides, we will use the signal-to-interference-plus-noise ratio (SINR) as performance index in these environments.

We compare here the SVD-based MMSE–the HyHOS I and the HyHOS II. This comparison must be done with equal convergence speed (i.e., learning speed) on one hand and equal final variance on the other to faithfully characterize their performance. The useful methods described in [42] and [43] are to be used to analyze the final variance and the convergence speed of each algorithm. Those methods are straightforward in the case of linear algorithms such as the MMSE but, in the case of the nonlinear functionals of the HyHOS algorithms, this is a much difficult task and, as such, merits a separate analysis which is currently under way. In the present work, and as a means to compare the algorithms on an equal footing, the convergence speed of the NG-based MUDs was fixed so that if we have a faster average convergence than the MMSE we do not have a worse SINR (or BER) at any sample size. In this sense, the learning rate of the NG was set to $\lambda = 1/t$ and the activation function used in the HyHOS II was the cubic function $\varphi(y) = y^3$. The number of Monte Carlo simulations averaged is 100 in the HNE and 500 in the LNE, which amounts to a confidence margin of at least 95% for any BER estimation presented in this work. The number of samples used in the learning of the detectors used in the BER estimation was 2000.

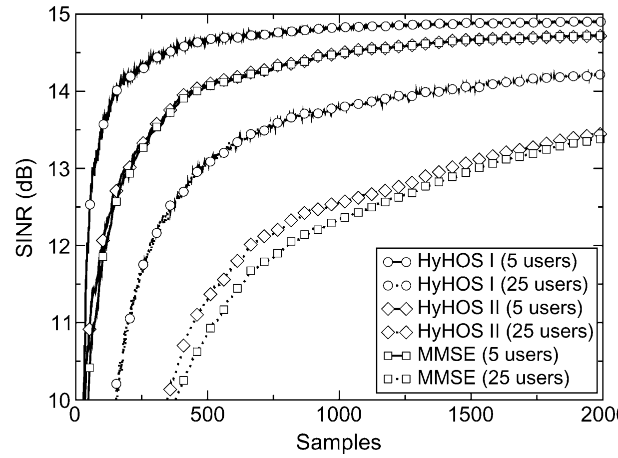

Fig. 1. Signal-to-interference plus noise rate for the HyHOS detectors and the MMSE detector, both with 15 and 20 users and a SNR of the User Of Interest $(\mathrm{UOI}) = 15\ \mathrm{dB}$. The powers of the interfering users are distributed uniformly between 0 and 30 dB above that of the UOI.

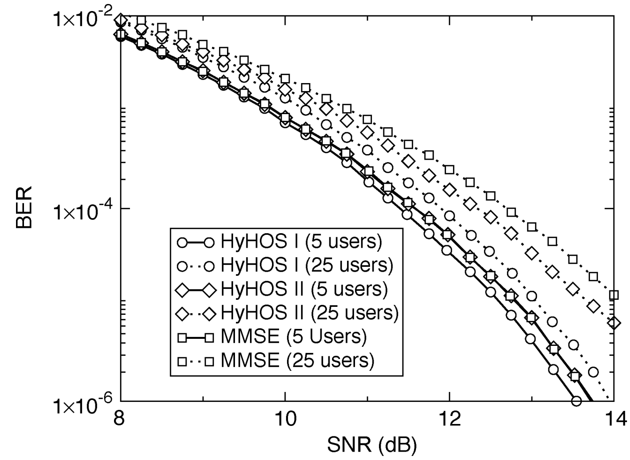

Fig. 2. Bit error rate in LNE (i.e., SNR of the $\mathrm{UOI} \geq 8\ \mathrm{dB}$) for the HyHOS detectors and the MMSE detector, both with five and 25 users. The powers of the interfering users are distributed uniformly between 0 and 30 dB above that of the UOI.

A. LNE

In the LNE we present, for the sake of clarity, the convergence performance of the HyHOS and the MMSE algorithms both with five and 25 users. In Fig. 1, we see a clear difference in the convergence behavior of the HyHOS I and the MMSE algorithm. That of the HyHOS II online algorithm closely resembles that of the offline MMSE method. It can be observed that 2000 samples is more than enough training length for the HyHOS I algorithm while for the MMSE, convergence cannot be achieved in a highly occupied channel.

In the evaluation of the bit error rate (see Fig. 2), this speed of convergence of the different algorithms is reflected in the performance of all the algorithms. In this LNE, the sought features in a multiuser detector are that of a controlled degradation of bit error rate with an increase in the number of users. Both the SICA-based HyHOS algorithm (HyHOS I) and the natural-gradient-based HyHOS (HyHOS II) show a performance close to that of the MMSE (if not better in the HyHOS I case) with a low occupation of the channel (five users). But the difference in behavior arises in a highly occupied channel (25 users), where the BER of the MMSE degrades from 10^{-6} at approximately 14 dB of SNR to 10^{-5}, while the HyHOS I algorithm keeps the BER to approximately 10^{-6} in the same conditions. The HyHOS II is halfway between both methods at a lower computational cost.

B. HNE

In the case of HNE, the convergence speed (see Fig. 3) of both HyHOS algorithms is completely similar but for higher variance in the HyHOS I algorithm due to its sensitiveness to noise. Both for a low or high number of users, the convergence is fairly similar. The MMSE shows a lack of convergence in a high-noise, low-interference channel. This was expected for the MMSE because the interference noise structure is completely masked by Gaussian noise, avoiding this algorithm to "lock" on to the signal of the user of interest. Quite the contrary is the case of the HyHOS algorithms.

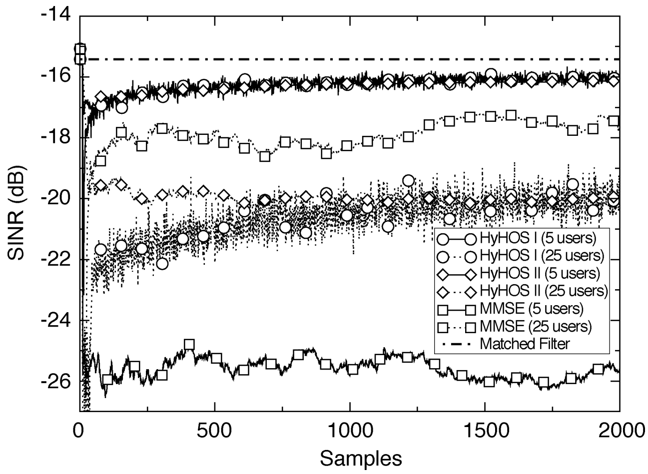

Fig. 3. Signal to interference plus noise rate, with SNR of the $\mathrm{UOI} = -15\ \mathrm{dB}$, for the HyHOS detectors and the MMSE detector, with five and 25 users. The powers of the interfering users are distributed uniformly between 0 and 30 dB above that of the UOI.

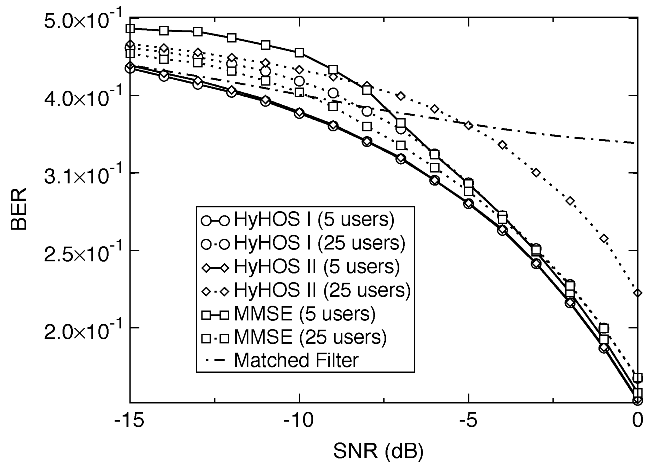

Fig. 4. Bit error rate in high noise environment (HNE) [i.e., SNR of the $\mathrm{UOI} \leq 0\ \mathrm{dB}$], for the HyHOS detectors and the MMSE detector, with five and 25 users. The powers of the interfering users are distributed uniformly between 0 and 30 dB above that of the UOI.

The bit error rates of the HyHOS algorithms show the noise resistance proper to the MMSE without showing its demeanors. The convergence to a matched filter solution in the asymptotic limit of $\mathrm{SNR(dB)} \to -\infty$ is shown experimentally in Fig. 4.

VII. Conclusion

In this paper, the authors propose HOS-based BSS as valid techniques to exploit the matrix structure of the parameters involved in MUD, an application of the more general narrowband m-sensor linear-array digital communications problem. We exploit BSS approaches based on contrast functions and estimation equations such as the SICA and the M-EASI. The Jacobi optimization and the natural gradient may be used in the optimization of these HOS-based cost functions. We developed a novel hybrid MUD by introducing these BSS techniques into the structure of the blind centralized MMSE MUD for synchronous CDMA systems. This way, we palliate one of the main problems of the BSS/ICA algorithms–their sensitiveness to noise. In addition, by using HOS, we achieve a better solution than with SOS-based approaches as the MMSE. The results included here show a good near-far resistant performance in synchronous CDMA as these BSS methods are equivariant. Besides these BSS methods being superefficient, these new detectors present good convergence. Finally, natural gradient-based MUD exhibits a low computational cost for adaptive versions. Some experiments have been included to demonstrate these properties.

Acknowledgment

The authors would like to acknowledge Dr. Zarzoso and Dr. Martínez-Ramón for their valuable comments.

References

[1] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[2] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Processing*, vol. 43, pp. 95–107, Jan. 1995.

[3] X. Wang and H. V. Poor, "Blind multiuser detection: a subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–690, Mar. 1998.

[4] D. S.-V. A. J. Caamaño and J. Ramos, "Blind adaptive krylov subspace multiuser detection," in *Proc. IEEE 54th Vehicular Technology Conf.*, Oct. 2001.

[5] J. F. Cardoso, "Blind signal separation: statistical principles," *Proc. IEEE*, vol. 86, pp. 2009–2025, Oct. 1998.

[6] A. J. Caamaño, J. J. Murillo-Fuentes, F. J. González-Serrano, and J. Ramos, "Natural gradient based multiuser detection," in *Proc. Personal Indoor Mobile Radio Communications*, Lisbon, Portugal, Sept. 2002.

[7] J. Murillo-Fuentes and F. González-Serrano, "Median equivariant adaptive separation via independence: Application to communications," *Neurocomput.*, vol. 49, no. 1, pp. 389–409, 2002.

[8] J. Murillo-Fuentes, H. Molina-Bulla, and F. González-Serrano, "Independent component analysis applied to digital image watermarking," in *Proc. Int. Conf. Audio, Speech Signal Processing*, vol. III, Salt Lake City, UT, May 2001, pp. 1997–2000.

[9] "Working Draft Information Technology Multimedia Content Description Interface,", Tech. Rep. ISO/IEC CD 15938-4, 2000.

[10] T. Lee, M. Lewicki, and T. Sejnowski, "Unsupervised classification with nongaussian mixture models using ICA," in *Advances in Neural Information Processing Systems*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds. Cambridge, MA: MIT Press, 1999, vol. 11, pp. 58–64.

[11] J. F. Cardoso and B. H. Laheld, "Equivariant adaptive source separation," *IEEE Trans. Signal Processing*, vol. 44, pp. 3017–3030, Dec. 1996.

[12] S. I. Amari and J. Cardoso, "Blind source separation-semiparametric statistical approach," *IEEE Trans. Signal Processing*, vol. 45, pp. 2692–2700, Nov. 1997.

[13] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Networks*, vol. 10, pp. 626–634, May 1999.

[14] D. T. Pahm, P. Garrat, and C. Jutten, "Separation of a mixture of independent sources through a ML approach," in *Proc. European Signal Processing Conf.*, 1992, p. 771, 774.

[15] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, pp. 287–314, Apr. 1994.

[16] J. Murillo-Fuentes and F. González-Serrano, "Independent component analysis with sinusoidal fourth-order contrast," in *Proc. Int. Conf. Audio, Speech Signal Processing*, vol. V, Salt Lake City, UT, May 2001, pp. 2785–2788.

[17] J. J. Murillo-Fuentes, M. Sánchez-Fernández, A. Caamaño-Fernández, and F. J. González-Serrano, "Adaptive blind joint source-phase separation in digital communications," in *Proc. IEEE Int. Conf. Communications*, Helsinki, Finland, June 2001.

[18] R. T. Causey and J. R. Barry, "Blind multiuser detection using linear prediction," *IEEE J. Selected Areas Commun.*, vol. 16, pp. 1702–1710, Dec. 1998.

[19] J. Joutsensalo and T. Ristaniemi, "Learning algorithms for blind multiuser detection in CDMA downlink," in *Proc. Personal Indoor Mobile Radio Commun.*, 1998.

[20] S. Amari, "Superefficiency in blind source separation," *IEEE Trans. Signal Processing*, vol. 47, pp. 936–944, Apr. 1999.

[21] F. G.-S. A. J. Caamaño, J. J. Murillo-Fuentes, and J. Ramos, "Natural gradient based multiuser detection," in *Proc. IEEE PIRMC*, Lisbon, Portugal, Sept. 2002.

[22] D. Pham, “Blind separation of instantaneous mixture of sources via an independent component analysis,” *IEEE Trans. Signal Processing*, vol. 44, pp. 2668–2779, Nov. 1996.

[23] P. Comon, “Contrast for multichannel blind deconvolution,” *IEEE Signal Processing Lett.*, vol. 3, pp. 209–211, July 1996.

[24] S. I. Amari and A. Cichocki, “Adaptive blind signal processing—Neural network approaches,” *Proc. IEEE*, vol. 86, pp. 2026–2048, Oct. 1998.

[25] L. Tong, Y. Inouye, and R. Liu, “Waveform preserving blind estimation of multiple independent sources,” *IEEE Trans. Signal Processing*, vol. 41, pp. 2461–2470, July 1993.

[26] A. Cichocki, S. Douglas, and S. Amari, “Robust techniques for independent component analysis (ICA) with noisy data,” *Neurocomput.*, no. 22, pp. 113–129, 1998.

[27] S. Douglas, A. Cichocki, and S. Amari, “Self-whitening algorithms for adaptive equalization and deconvolution,” *IEEE Trans. Signal Processing*, vol. 47, pp. 1161–1165, Apr. 1999.

[28] X. R. Cao and R. W. Liu, “General approach to blind source separation,” *IEEE Trans. Signal Processing*, vol. 44, pp. 562–571, Mar. 1996.

[29] S. I. Amari, “Neural learning in structured parameter spaces—Natural riemannian gradient,” in *Advances in Neural Information Processing Systems*, M. C. Mozer, M. I. Jordan, and T. Petsche, Eds. Cambridge, MA: MIT Press, 1997, vol. 9, p. 127.

[30] J. Murillo-Fuentes and F. González-Serrano, “Improving stability in blind source separation with the stochastic median gradient,” *Electron. Lett.*, vol. 36, pp. 1662–1663, Sept. 2000.

[31] J. F. Cardoso, “High-order contrasts for independent component analysis,” *Neural Comput.*, vol. 11, pp. 157–192, Jan. 1999.

[32] A. Stuart and J. K. Ord, *Kendall’s Advanced Theory of Statistics*, 6th ed. London, U.K.: Edward Arnold, 1994, vol. I.

[33] V. Zarzoso and K. Nandi, “Unified formulation of closed-form estimators for blind source separation incomplex instantaneous mixtures,” in *Proc. European Signal Processing Conf.*, vol. I, Tampere, Finland, Sept. 2000.

[34] V. Zarzoso and A. K. Nandi, “Blind separation of independent sources for virtually any source probability density function,” *IEEE Trans. Signal Processing*, vol. 47, pp. 2419–2432, Sept. 1999.

[35] M. Ghogho, A. Swami, and T. Durrani, “Approximate maximum likelihood blind source separation with arbitrary source pdfs,” in *Proc. IEEE Workshop Statistical Signal Array Processing*, Pocono Manor, PA, Aug. 2000.

[36] V. Zarzoso, F. Herrmann, and K. Nandi, “Weighted closed-form estimators for blind source separation,” in *Proc. 11th Int.l Workshop Statistical Signal Processing*, Singapore, Aug. 2001.

[37] P. Comon and E. Moreau, “Improved contrast dedicated to blind separation in communications,” in *Proc. IEEE Int. Confe. Acoustics, Speech Signal Processing*, vol. V, Munich, Germany, 1997, pp. 3453–3456.

[38] E. Moreau and O. Macchi, “Higher order contrast for self-adaptive source separation,” *Proc. Int. J. Adaptive Control Signal Processing*, vol. 10, pp. 19–46, Jan. 1996.

[39] F. Herrmann and A. Nandi, “Blind separation of linear instantaneous mixture using close forms estimators,” *Signal Processing*, vol. 81, pp. 1537–1556, 2001.

[40] F. Harroy and J.-L. Lacoume, “Maximum likelihood estimators and cramer-rao bounds in source separation,” *Signal Process.*, vol. 55, no. 2, pp. 167–177, 1996.

[41] S. I. Amari, “Natural gradient works efficiently in learning,” *Neural Comput.*, vol. 10, pp. 251–276, Jan. 1998.

[42] N. Bershad, P. Celka, and J. Vesin, “Stochastic analysis of gradient adaptive identification of nonlinear systems with memory for gaussian data and noisy input and output measurements,” *IEEE Trans. Signal Process.*, vol. 47, pp. 675–689, Mar. 1999.

[43] J. Bucklew, T. Kurtz, and W. Sethares, “Weak convergence and local stability properties of fixed step size recursive algorithms,” *IEEE Trans. Inform. Theory*, vol. 39, pp. 966–978, May 1993.

Antonio J. Caamaño (M’99) received the B.Sc and M.Sc. degrees in theoretical physics from the Universidad Autónoma of Madrid, Madrid, Spain in 1995, and the Ph.D. degree from the Carlos III University of Madrid in 2003, with a dissertation in the area of multiuser detection algorithms based in blind source separation and the analysis of differential geometric nature of lie groups.

Currently, he is Associate Professor in the Rey Juan Carlos University of Madrid, Madrid, Spain. He published four papers in the area of theoretical condensed matter physics. In 1998, he changed the subject of his investigations to the area of signal theory and communications. His main subjects of investigation are channel estimation in OFDM communications through wavelet analysis and the theoretical analysis of the self-organizing nature of different MAC layers in wireless sensor networks. He is also working in the implementation of the aforementioned algorithms in a Software Defined Radio platform.

Rafael Boloix-Tortosa received the telecommunication engineering degree in 2000 from the University of Seville, Seville, Spain, where he is currently pursuing the Ph.D. degree.

Currently, he is an Assistant Professor at the University of Seville, where he was a Research Assistant from 1999 to 2000. His research interests include higher-order statistics and algorithm for independent component analysis applied to multiuser detection.

Javier Ramos (M’93) received the B.Sc and M.Sc. degrees from the Polytechnic University of Madrid, Madrid, Spain. He received the Ph.D. degree in 1995.

Currently, he is Associate Professor at Carlos III University of Madrid. From 1992 to 1995, he cooperated in several research projects at Purdue University, West Lafayete, IN, working in signal processing for communications. In 1996, he was Post-Doctoral Research Associate at Purdue University. Since 2003, he has been the Dean of the Telecommunications Engineering Department at the Rey Juan Carlos University of Madrid. During the last 12 years, he has worked on the development of new technologies for communication equipment, dedicating special attention to advanced techniques for radio transmission.

Dr. Ramos received the Ericsson Award for the best Ph.D. dissertation on mobile communications in 1996.

Juan José Murillo-Fuentes (M’99) received the Telecommunication Engineering degree from the University of Seville, Seville, Spain, in 1996, and the Ph.D. degree from the Carlos III University of Madrid in 2001.

Currently, he is Assistant Professor at the University of Seville. His main research interests include machine learning, image processing, and array processing with emphasis on algorithms for independent component analysis applied to multiuser detection in mobile communications.