Performance limits of alphabet diversities for FIR SISO channel identification

Jean-Pierre Delmas, Yann Meurisse, Pierre Comon

Performance limits of alphabet diversities for FIR SISO channel identification

Jean Pierre Delmas, *Senior Member, IEEE*, Yann Meurisse and Pierre Comon, *Fellow Member, IEEE*

***Abstract*—Finite Impulse Responses (FIR) of Single-Input Single-Output (SISO) channels can be blindly identified from second order statistics of transformed data, for instance when the channel is excited by Binary Phase Shift Keying (BPSK), Minimum Shift Keying (MSK) or Quadrature Phase Shift Keying (QPSK) inputs. Identifiability conditions are derived by considering that noncircularity induces diversity. Theoretical performance issues are addressed to evaluate the robustness of standard subspace-based estimators with respect to these identifiability conditions. Then benchmarks such as asymptotically minimum variance (AMV) bounds based on various statistics are presented. Some illustrative examples are eventually given where Monte Carlo experiments are compared to theoretical performances. These comparisons allow to quantify limits to the use of the alphabet diversities for the identification of FIR SISO channels, and to demonstrate the robustness of algorithms based on High-Order Statistics.**

***Index Terms*—Blind estimation, SISO channel, Noncircularity, Performance analysis**

I. Introduction

SISO blind identification has been long considered to need High-Order Statistics (HOS) [1], [2]. Actually, it is now well known that the use of an additional diversity at the receiver permits to build a SIMO channel that can be identified with the sole help of second order statistics, e.g., via subspace techniques [3]; if spatial diversity is not available at the receiver, oversampling allows to increase diversity only in the presence of sufficient excess bandwidth, which is however rarely encountered. This is one of the reasons why HOS-based techniques are still often preferred. Other more recent techniques incorporate the knowledge of the symbol constellation, which eventually amounts to using noncircularity of the symbol sequence [4], [5], [6], [7], [8]. It is thus legitimate to ask oneself the question whether the latter are more attractive than HOS-based approaches. We address this question in rather favorable conditions, but our answer is still negative, as will be subsequently shown.

Second or higher order noncircularity can be utilized in order to restore identifiability of FIR SISO channels in the absence of space or bandwidth diversities. This results in simple SIMO-type blind identification algorithms based on second order statistics of *transformed data*; these transformations include complex conjugation, but also monomial functions. This is one of the differences between our analysis and that of [5], where the performance analysis only takes into account second-order statistics of the inputs. Identifiability conditions are derived for BPSK, MSK and QPSK inputs. Theoretical performance issues are addressed to evaluate the robustness of standard subspace-based estimators when the impulse response approaches an unidentifiable channel. In this contribution, it is demonstrated that HOS-based blind identification algorithms exhibit a much better robustness than alphabet-based.

For this purpose, benchmarks such as AMV bounds based on second order or extended second order statistics, or based on orthogonal projectors, are presented. In comparison to previous works dedicated to performance analysis of identification of impulse responses, which suppose the stacked samples of the received signal to be independent and Gaussian distributed [9], [10], the theoretical performance analysis we propose is based on the actual distribution of the observed signals, including the possibility of *correlated* signals.

This paper is organized as follows. Section II introduces FIR SISO data models. Identifiability results are given in section III, and performance issues are addressed in section IV. Some illustrative examples are reported in section V. Finally, conclusions are drawn in Section VI.

II. Data model

Limiting our discussion to linear modulations, the complex envelope of a transmitted signal $s(t)$ takes the baseband expression $s(t) = \sum_k g(t - kT)\, x_k$, where x_k denotes the discrete sequence of transmitted symbols, T the symbol period, and $g(\cdot)$ the transmit filter. After propagation through a time-invariant channel, the signal received on the antenna is of the form $s(t) = \sum_k h(t - kT)\, x_k$, for some complex linear filter $h(\cdot)$ representing the global channel, combining transmit and receive filters with the channel. It is subsequently assumed that the global channel can be approximated by a FIR filter. Thus, if sampled exactly at the rate $1/T$, the received discrete-time signal may be modelled as

$$y_t = \sum_{k=0}^{M} h_k\, x_{t-k} + n_t$$

where $h_k = h(kT)$, $0 \leq k \leq M$, and n_t denotes an additive noise, which is assumed to be second-order circular

Jean Pierre Delmas and Yann Meurisse are with Institut TELECOM, TELECOM et Management SudParis, Département CITI, CNRS UMR 5157, 91011 Evry Cedex, France, e-mail:jean-pierre.delmas@it-sudparis.eu, yann.meurisse@it-sudparis.eu, phone: +(33).1.60.76.46.32, fax: +(33).1.60.76.44.33.

Pierre Comon is with the I3S laboratory, University of Nice, CNRS, Sophia-Antipolis France, email: p.comon@ieee.org, phone: +(33).4.92.94.27.26, fax: +(33).4.92.94.28.96/98.

(i.e. $\mathrm{E}\{n_t^2\} = 0$), white, zero-mean and with finite variance $\sigma_n^2 = \mathrm{E}\{|n_t|^2\}$. The information symbol sequence, x_t is a stationary process; in addition, x_t and noise n_t are assumed to be statistically independent. It is convenient to use a multivariate representation by stacking $M+1$ samples[1] of the received signal:

$$\mathbf{y}_t = (y_t, y_{t-1}, ..., y_{t-M})^T = \mathcal{H}(\mathbf{h})\mathbf{x}_t + \mathbf{n}_t \qquad (1)$$

with $\mathbf{x}_t = (x_t, x_{t-1}, ..., x_{t-2M})^T$ and $\mathbf{n}_t = (n_t, n_{t-1}, ..., n_{t-M})^T$, and where $\mathcal{H}(\mathbf{h})$ is the following $(M+1) \times (2M+1)$ Toeplitz matrix:

$$\mathcal{H}(\mathbf{h}) = \begin{pmatrix} h_0 & \cdots & \cdots & h_M & & \\ & \ddots & & & \ddots & \\ & & h_0 & \cdots & \cdots & h_M \end{pmatrix}$$

with $\mathbf{h} \stackrel{\text{def}}{=} (h_0, h_1, ..., h_M)^T$. In the following, the cases of BPSK, MSK and QPSK modulations will be considered, as working examples.

III. Identifiability

A. BPSK modulation

In this section, x_t is a stationary process, possibly colored, taking its values in the set $\{-1, +1\}$ with equal probabilities. It is assumed that the so-called *noncircular* covariance $\mathbf{R}_x \stackrel{\text{def}}{=} \mathrm{E}\{\mathbf{x}_t\mathbf{x}_t^T\}$ is nonsingular. The set of nonzero circular and noncircular second-order statistics of y_t can be gathered in the covariance matrix of the extended vector $\tilde{\mathbf{y}}_t = (\mathbf{y}_t^T, \mathbf{y}_t^H)^T$, so that from (1):

$$\begin{aligned} \mathbf{R}_{\tilde{y}} &\stackrel{\text{def}}{=} \mathrm{E}\{\tilde{\mathbf{y}}_t\tilde{\mathbf{y}}_t^H\} \\ &= \begin{bmatrix} \mathcal{H}(\mathbf{h}) \\ \mathcal{H}(\mathbf{h}^*) \end{bmatrix} \mathbf{R}_x \begin{bmatrix} \mathcal{H}(\mathbf{h})^H & \mathcal{H}(\mathbf{h}^*)^H \end{bmatrix} + \sigma_n^2 \mathbf{I}_{2M+2} \end{aligned} \qquad (2)$$

Consequently, we obtain this way a structured covariance matrix similar to that obtained in the SIMO case; here the two channels have impulse responses $\mathbf{h}$ and $\mathbf{h}^*$. Therefore the results (see e.g., [11]) concerning the *global* identifiability of SIMO channels can be applied. Because $2(M+1) > 2M+1$ and $\mathbf{R}_x$ is nonsingular, the range space of the filtering matrix $\tilde{\mathcal{H}}(\mathbf{h}) \stackrel{\text{def}}{=} \begin{bmatrix} \mathcal{H}(\mathbf{h}) \\ \mathcal{H}(\mathbf{h}^*) \end{bmatrix}$ is identifiable from $\mathbf{R}_{\tilde{y}}$, and this range space determines the channel coefficients up to a multiplicative constant if channels $\mathbf{h}$ and $\mathbf{h}^*$ do not share any common zeros. This ambiguity can be fixed by using the knowledge of the alphabet; we have proved the following.

Result 1: With a BPSK modulation and additive noise satisfying the above assumptions, the impulse response $\mathbf{h}$ of a SISO channel is globally identifiable from the circular and noncircular second-order statistics of its output if the polynomial $h(z) = \sum_{k=0}^{M} h_k z^k$ has neither real zeros nor conjugated zeros.

Using the notion of *local* identifiability for which it has been shown in [10] that a SIMO channel is identifiable if and only if the different channels do not share common conjugate reciprocal zeros, or equivalently here $h(z)$, $h^*(z)$, $h(1/z^*)$ and $h^*(1/z^*)$ do not share common zeros, we have proved the following

Result 2: Under the conditions of Result 1, the impulse response $\mathbf{h}$ of a SISO channel is locally identifiable from the circular and noncircular second-order statistics of its output if and only if $h(\pm 1) \neq 0$.

Remark: If this condition is satisfied, there is no other channel in the neighborhood of a channel $\mathbf{h}$, but possibly other channels outside this neighborhood satisfying the constraints imposed by the second-order statistics of its output. Consequently in practice, one must know the whereabouts of the true channel in order to identify it under this condition only. Such knowledge may come from some prior information. A locally consistent estimator that gives accurate estimate of a channel satisfying $h(\pm 1) \neq 0$ but with real zeros or conjugated zeros will be exhibited in Section IV. This is in contrast to the global identifiability, for which we know that if a channel is identifiable, there is no other channel $\mathbf{h}$ in the space $\mathbb{C}^{M+1}$ satisfying the constraints imposed by the second-order statistics of its output.

B. MSK modulation

Now, we suppose x_t is a MSK modulated signal defined by $x_{t+1} \stackrel{\text{def}}{=} ix_tc_t$ where c_t is a sequence of independent BPSK symbols $\{-1, +1\}$ with equal probabilities where the original value x_0 remains unspecified in the set $\{+1, +i, -1, -i\}$. This process may be equivalently modelled as $x_t = i^t b_t x_0$ where b_t is another sequence of independent BPSK symbols $\{-1, +1\}$ with equal probabilities.

We note that x_t (and thus y_t) is not stationary. But by demodulating[2] each y_t, we get

$$y'_t \stackrel{\text{def}}{=} y_t i^{-t} = \sum_{k=0}^{M} h'_k b_{t-k} + n'_t, \qquad (3)$$

with $h'_k \stackrel{\text{def}}{=} x_0 h_k i^{-k}, k = 0, ..., M$ and where $n'_t \stackrel{\text{def}}{=} i^{-t} n_t$ is still second order stationary. Gathering again the set of nonzero circular and noncircular second-order statistics of y'_t in the covariance matrix of the extended vector $\tilde{\mathbf{y}}'_t = (\mathbf{y}'^T_t, \mathbf{y}'^H_t)^T$ we obtain because $\mathbf{R}_b = \mathbf{I}_{2M+1}$ in contrast to $\mathbf{R}_x$ of (2) which is not necessarily equal to $\mathbf{I}_{2M+1}$ from the assumptions

$$\begin{aligned} \mathbf{R}_{\tilde{y}'} &\stackrel{\text{def}}{=} \mathrm{E}\{\tilde{\mathbf{y}}'_t\tilde{\mathbf{y}}'^H_t\} \\ &= \begin{bmatrix} \mathcal{H}(\mathbf{h}') \\ \mathcal{H}(\mathbf{h}'^*) \end{bmatrix} \begin{bmatrix} \mathcal{H}(\mathbf{h}')^H & \mathcal{H}(\mathbf{h}'^*)^H \end{bmatrix} + \sigma_n^2 \mathbf{I}_{2M+2} \end{aligned} \qquad (4)$$

[1] The length $M+1$ of the observed output samples y_t has been selected for two reasons. First, it is the minimal length to assure the SIMO channel identifiability with the so-called zero condition. Second, this length allows one to simplify the subspace algorithm because in this case, the dimension of the noise subspace reduces to one.

[2] In [4], this demodulation is performed on vector $\mathbf{y}_t$ and consequently $\mathbf{y}'_t = \mathbf{y}_t i^{-t}$ becomes stationary at the second order but not y'_t.

Consequently, we can use the same approach as for the BPSK modulation. Because $\sum_{k=0}^{M} h'_k z^k = x_0 \sum_{k=0}^{M} h_k (i^{-1}z)^k$, we obtain the result below.

Result 3: With a MSK modulation and additive noise satisfying the above assumptions, the impulse response $\mathbf{h}$ of a SISO channel is globally [resp. locally] identifiable from the circular and noncircular second-order statistics of its demodulated output if the polynomial $h(z)$ has neither purely imaginary zeros nor paired zeros of the form (iz_0, iz_0^*), $z_0 \in \mathbb{C}$ [resp. if $h(\pm i) \neq 0$].

C. QPSK modulation

Now, we suppose x_t is a QPSK modulated signal defined as a sequence of i.i.d. r.v. taking their values in the set $\{+1, +i, -1, -i\}$ with equal probabilities and now n_t is Gaussian distributed and circular. Consequently y_t is now second-order circular. However, noting that x_t^2 is a BPSK modulation, a similar approach can still be used by squaring the outputs y_t. Indeed define $y'_t \stackrel{\text{def}}{=} y_t^2$ and $h'_k \stackrel{\text{def}}{=} h_k^2, k = 0, ..., M$. Then, by using (i) the multi-linearity of moments and cumulants [12] [13], (ii) relations between moments and cumulants, (iii) properties specific to the QPSK alphabet, namely $x_t^4 = 1$, $\mathrm{E}\{x_t^2\} = 0$ and $\mathrm{E}\{|x_t^2|\} = 1$, and (iv) the whiteness of x_t at order 4, we obtain the following second-order statistics of the modified input, y'_t:

$$\begin{aligned} -r_{y'}(\ell) + 2r_y^2(\ell) &\stackrel{\text{def}}{=} -\mathrm{E}\{y'_t y'^*_{t-\ell}\} + 2\left[\mathrm{E}\{y_t y^*_{t-\ell}\}\right]^2 \\ &= \sum_{k=\ell}^{M} h'_k h'^*_{k-\ell} \\ r'_{y'}(\ell) &\stackrel{\text{def}}{=} \mathrm{E}\{y'_t y'_{t-\ell}\} = \sum_{k=\ell}^{M} h'_k h'_{k-\ell} \end{aligned}$$

for $\ell = 0, ..., M$. Gathering the extended second-order statistics $r''_y(l) \stackrel{\text{def}}{=} -r_{y'}(\ell) + 2r_y^2(\ell)$ and $r'_{y'}(l)$ in an Hermitian 2×2 block Toeplitz matrix, we obtain the extended covariance matrix

$$\mathbf{R}_{\tilde{y}'} = \begin{bmatrix} \mathcal{H}(\mathbf{h}') \\ \mathcal{H}(\mathbf{h}'^*) \end{bmatrix} \begin{bmatrix} \mathcal{H}(\mathbf{h}')^H & \mathcal{H}(\mathbf{h}'^*)^H \end{bmatrix}. \quad (5)$$

As a consequence, a structured covariance matrix has been obtained, which is similar to that obtained in the SIMO case, with two channels of impulse responses $\mathbf{h}'$ and $\mathbf{h}'^*$. Therefore the results (see e.g., [11]) concerning the identifiability of SIMO channels can be applied as well. The range space of $\mathbf{R}_{\tilde{y}'}$ determines the channel coefficients of $\mathbf{h}'$ and $\mathbf{h}'^*$ up to a multiplicative constant if these channels do not share any common zeros. This ambiguity can be reduced by using the knowledge of the symbol alphabet. Returning to $\mathbf{h}$, each coefficient of $\mathbf{h}$ is determined up to a sign ambiguity, which may be cleared up using the successive second-order statistics $\mathrm{E}\{y_t y^*_{t-l}\}, l = M, M-1, ..., \lfloor M/2 \rfloor$. Consequently, we obtain the following result:

Result 4: With a QPSK modulation and additive noise satisfying the above assumptions, the impulse response $\mathbf{h}$ of a channel is globally [resp. locally] identifiable from the second-order statistics of the squared output if the polynomial $h'(z) = \sum_{k=0}^{M} h_k^2 z^k$ has neither real zeros nor conjugated zeros [resp. if $h'(\pm 1) \neq 0$].

IV. Asymptotic performance

The above identifiability results naturally raise the important issue of performance analysis of algorithms based on the second-order covariance $\mathbf{R}_{\tilde{y}}$, and on the modified second-order statistics $\mathbf{R}_{\tilde{y}'}$. In particular, it is essential to evaluate the performance of algorithms identifying channels whose impulse response approaches the conditions of non-identifiability previously given.

Usually, the deterministic Cramer-Rao bound (CRB) serves as a useful benchmark for unbiased estimators yielded by identification algorithms. Because this CRB is neither attainable for increasing SNR, nor for increasing the number of symbols, this CRB is only a loose lower bound for the variance of unbiased estimator, in contrast to the stochastic CRB, which is attainable in the previous conditions. In the present context, the distribution of $(y_t)_{t=1,..,T}$ is a mixture of a large number of Gaussian distributions and consequently the stochastic CRB appears to be computationally prohibitive. In these conditions, the notion of asymptotically (in the number of measurements) minimum variance (AMV) bound introduced by Porat and Friedlander [14], and Stoica *et al* with their asymptotically best consistent (ABC) estimator [15], are considered.

A. AMV bounds based on extended covariance matrices

We first note that $\mathbf{h}$ is identifiable from $\mathbf{R}_{\tilde{y}}$ or $\mathbf{R}_{\tilde{y}'}$ except an intrinsic ambiguity, viz: a sign ambiguity for BPSK, or a rotation of $\pi/2$ ambiguity for MSK and QPSK modulations. Furthermore, the subspace-based algorithms that are proposed estimate $\mathbf{h}$ up to a multiplicative constant. To compare the asymptotic performance of these algorithms to given AMV bounds based on various statistics, $\mathbf{h}$ must be strictly identifiable from these statistics. Consequently, one parameter of $\mathbf{h}$ is fixed to a predefined value, say, $h_0 = 1$ throughout Sections IV and V.

Let $\boldsymbol{\alpha} = (\boldsymbol{\theta}^T, \boldsymbol{\rho}^T)^T$ denote the real-valued unknown parameters (containing the real and imaginary parts of the complex parameters) of the extended covariance matrices $\mathbf{R}_{\tilde{y}}$ or $\mathbf{R}_{\tilde{y}'}$, where $\boldsymbol{\theta} \stackrel{\text{def}}{=} [\Re(h_1), ..., \Re(h_M), Im(h_1), ..., Im(h_M)]^T$, and where $\boldsymbol{\rho}$ collects the nuisance[3] parameters for the BPSK and MSK modulations. Depending on the a priori knowledge of the inputs, $\boldsymbol{\rho} = \sigma_n^2$ if the BPSK sequence x_t or the driving sequence b_t for the MSK modulation are white, and $\boldsymbol{\rho} = [r_x(1).., r_x(2M), \sigma_n^2]^T$ with $r_x(k) \stackrel{\text{def}}{=} \mathrm{E}\{x_t x_{t-k}\}$ for correlated BPSK symbols. We note that if the conditions of Section III are satisfied $\boldsymbol{\alpha}$ is identifiable from $\mathbf{R}_{\tilde{y}}$ or $\mathbf{R}_{\tilde{y}'}$.

These block Toeplitz matrices $\mathbf{R}_{\tilde{y}}$ and $\mathbf{R}_{\tilde{y}'}$ are traditionally consistently estimated from T successive received

[3]Note that all the previous performance analyses (e.g., [10]) consider that the nuisance parameters are known, and consequently give optimistic AMV bounds.

signals $\mathbf{y}_t$ by replacing the various expectations by the associated sample correlations: $r_{y,T}(\ell) \stackrel{\text{def}}{=} \frac{1}{T}\sum_{t=1}^{T} y_t y_{t-\ell}^*$, $r'_{y,T}(\ell) \stackrel{\text{def}}{=} \frac{1}{T}\sum_{t=1}^{T} y_t y_{t-\ell}$, $r_{y',T}(\ell) \stackrel{\text{def}}{=} \frac{1}{T}\sum_{t=1}^{T} y'_t y_{t-\ell}^{'*}$ and $r'_{y',T}(\ell) \stackrel{\text{def}}{=} \frac{1}{T}\sum_{t=1}^{T} y'_t y'_{t-\ell}$. In order to apply the AMV bound [14], [15] to an arbitrary consistent extended second-order algorithm based on the sample estimates $\mathbf{R}_{\tilde{y},T}$ or $\mathbf{R}_{\tilde{y}',T}$ of $\mathbf{R}_{\tilde{y}}$ or $\mathbf{R}_{\tilde{y}'}$ respectively, the involved statistics $\mathbf{s}_T$ must collect real-valued sample correlations and complex-valued sample correlations and their conjugate [16]. For example for BPSK modulations, we have:

$$\begin{aligned}\mathbf{s}_T &= (r_{y,T}(0), r_{y,T}(1), ..., r_{y,T}(M), r^*_{y,T}(1), ..., r^*_{y,T}(M),\\ &\quad r'_{y,T}(0), ..., r'_{y,T}(M), r'^*_{y,T}(0), ..., r'^*_{y,T}(M))^T. \qquad (6)\end{aligned}$$

Under these conditions, the asymptotic covariance $\mathbf{Cov}_{\boldsymbol{\alpha}}^{\text{Alg}}$ of an estimator of $\boldsymbol{\alpha}$ given by an arbitrary consistent second-order algorithm based on these statistics $\mathbf{s}_T$ is bounded below by the real symmetric positive definite matrix $[\mathbf{S}^H(\boldsymbol{\alpha})\mathbf{C}_{\mathbf{s}}^{-1}(\boldsymbol{\alpha})\mathbf{S}(\boldsymbol{\alpha})]^{-1}$ where[4] $\mathbf{S}(\boldsymbol{\alpha}) \stackrel{\text{def}}{=} \frac{d\mathbf{s}(\boldsymbol{\alpha})}{d\boldsymbol{\alpha}}$ with $\mathbf{s}(\boldsymbol{\alpha}) = \lim_{T\to\infty} \mathbf{s}_T$ and where $\mathbf{C}_s(\boldsymbol{\alpha})$ is the circular covariance[5] of the asymptotic distribution of $\mathbf{s}_T$ whose expression is given in Appendix A. Furthermore, there exists a nonlinear least square algorithm (dubbed the AMV algorithm [14]), for which the covariance of the asymptotic distribution of the estimate of $\boldsymbol{\alpha}$ attains this lower bound.

For white BPSK sequences and white driving sequences of MSK modulations, σ_n^2 is the unique nuisance parameter and $r_y(k) = \sum_{i=k}^{M} h_i h_{i-k}^* + \delta_k \sigma_n^2$ and $r'_y(k) = \sum_{i=k}^{M} h_i h_{i-k}$ for BPSK sequences. Consequently $\mathbf{s}(\boldsymbol{\alpha})$ is structured as $\mathbf{s}(\boldsymbol{\alpha}) = \boldsymbol{\Phi}(\boldsymbol{\theta}) + \sigma_n^2 \mathbf{e}_1$, where $\mathbf{e}_1$ is the first unit vector in $\mathbb{R}^{4M+3}$. This implies $\mathbf{S} = [\mathbf{S}_1, \mathbf{e}_1]$, and the matrix inversion lemma yields:

$$\mathbf{Cov}_{\boldsymbol{\theta}}^{\text{Alg}} \geq \mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{R}_{\tilde{y}})^{(1)}}, \qquad (7)$$

with

$$\begin{aligned}\mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{R}_{\tilde{y}})^{(1)}} &\stackrel{\text{def}}{=} \left[\mathbf{S}^H(\boldsymbol{\alpha})\mathbf{C}_{\mathbf{s}}^{-1}(\boldsymbol{\alpha})\mathbf{S}(\boldsymbol{\alpha})\right]^{-1}_{(1:2M,1:2M)}\\ &= \left[\begin{bmatrix}\mathbf{S}_1^H\\ \mathbf{e}_1^H\end{bmatrix}\mathbf{C}_{\mathbf{s}}^{-1}(\boldsymbol{\alpha})[\mathbf{S}_1, \mathbf{e}_1]\right]^{-1}_{(1:2M,1:2M)}\\ &= \left(\mathbf{S}_1^H \mathbf{C}_{\mathbf{s}}^{-1/2} \boldsymbol{\Pi}^{\perp}_{\mathbf{C}_{\mathbf{s}}^{-1/2}\mathbf{e}_1} \mathbf{C}_{\mathbf{s}}^{-1/2} \mathbf{S}_1\right)^{-1},\end{aligned}$$

where $\boldsymbol{\Pi}_{\mathbf{A}}^{\perp}$ denotes the projector onto the orthogonal complement of the columns of $\mathbf{A}$.

For the BPSK modulation with correlated symbols, $r_y(k) = \sum_{i=0}^{M}\sum_{j=0}^{M} h_i h_j^* r_x(k+j-i) + \delta_k\sigma_n^2$ and $r'_y(k) = \sum_{i=0}^{M}\sum_{j=0}^{M} h_i h_j r_x(k+j-i)$ where $r_x(0) = 1$. Consequently $\mathbf{s}(\boldsymbol{\alpha}) = (r_y(0), r_y(1), ..., r_y(M), r_y^*(1), ..., r_y^*(M), r'_y(0), ..., r'_y(M), r'^*_y(0), ..., r'^*_y(M))^T$ is structured as $\mathbf{s}(\boldsymbol{\alpha}) = \boldsymbol{\Psi}(\boldsymbol{\theta})\boldsymbol{\rho} + \boldsymbol{\psi}(\boldsymbol{\theta})$, which implies $\mathbf{S} = [\mathbf{S}_1, \boldsymbol{\Psi}]$ with $\mathbf{S}_1 \stackrel{\text{def}}{=} \frac{\partial \mathbf{s}(\boldsymbol{\alpha})}{\partial \boldsymbol{\theta}}$ and the matrix inversion lemma gives this time:

$$\mathbf{Cov}_{\boldsymbol{\theta}}^{\text{Alg}} \geq \mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{R}_{\tilde{y}})^{(1)}}, \qquad (8)$$

with

$$\begin{aligned}\mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{R}_{\tilde{y}})^{(2)}} &\stackrel{\text{def}}{=} \left[\mathbf{S}^H(\boldsymbol{\alpha})\mathbf{C}_{\mathbf{s}}^{-1}(\boldsymbol{\alpha})\mathbf{S}(\boldsymbol{\alpha})\right]^{-1}_{(1:2M,1:2M)}\\ &= \left(\mathbf{S}_1^H \mathbf{C}_{\mathbf{s}}^{-1/2} \boldsymbol{\Pi}^{\perp}_{\mathbf{C}_{\mathbf{s}}^{-1/2}\boldsymbol{\Psi}} \mathbf{C}_{\mathbf{s}}^{-1/2} \mathbf{S}_1\right)^{-1}.\end{aligned}$$

We note that when the BPSK symbols are uncorrelated,

$$\mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{R}_{\tilde{y}})^{(1)}} \leq \mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{R}_{\tilde{y}})^{(2)}},$$

where here $\mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{R}_{\tilde{y}})^{(2)}}$ is derived under the assumption of uncorrelated symbols.

Lastly for the QPSK modulation, where there is no nuisance parameter in $\mathbf{R}_{\tilde{y}'}$, we obtain the lower bound

$$\mathbf{Cov}_{\boldsymbol{\theta}}^{\text{Alg}} \geq \mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{R}_{\tilde{y}'})}, \qquad (9)$$

with

$$\mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{R}_{\tilde{y}'})} = [\mathbf{S}^H(\boldsymbol{\theta})\mathbf{C}_{\mathbf{s}}^{-1}(\boldsymbol{\theta})\mathbf{S}(\boldsymbol{\theta})]^{-1},$$

where $\mathbf{C}_{\mathbf{s}}$ is here associated with the statistics $\mathbf{s}_T = [r''_{y,T}(0), ..., r''_{y,T}(M), r''^*_{y,T}(1), ..., r''^*_{y,T}(M), r'_{y',T}(0), ..., r'_{y',T}(M), r'^*_{y',T}(0), ..., r'^*_{y',T}(M)]^T$ where $r''_{y,T}(k) \stackrel{\text{def}}{=} -\frac{1}{T}\sum_{t=1}^{T} y'_t y'^*_{t-k} + 2\left[\frac{1}{T}\sum_{t=1}^{T} y_t y^*_{t-k}\right]^2$ and $r'_{y',T}(k) \stackrel{\text{def}}{=} \frac{1}{T}\sum_{t=1}^{T} y'_t y'_{t-k}$ (see Section III-C), which is a consistant estimate of $\mathbf{s}(\boldsymbol{\theta}) = [r''_y(0), ...r''_y(M), r''^*_y(1), ...r''^*_y(M), r'_{y'}(0), ..., r'_{y'}(M), r'^*_{y'}(0), ..., r'^*_{y'}(M)]^T$, whose expression is derived in Appendix B.

B. AMV bounds based on orthogonal projectors

In order to assess the performance of subspace-based algorithms built from $\mathbf{R}_{\tilde{y},T}$ or $\mathbf{R}_{\tilde{y}',T}$, it is relevant to replace the previous statistics $\mathbf{s}_T$ by the orthogonal projectors $\boldsymbol{\Pi}_{\tilde{y},T}$ or $\boldsymbol{\Pi}_{\tilde{y}',T}$ onto the noise subspace of $\mathbf{R}_{\tilde{y},T}$ or $\mathbf{R}_{\tilde{y}',T}$ respectively. In this case, $\mathbf{s}_T = \text{vec}(\boldsymbol{\Pi}_{\tilde{y},T})$ or $\text{vec}(\boldsymbol{\Pi}_{\tilde{y}',T})$, consistent estimates of $\boldsymbol{\Pi}_{\tilde{y}}$ and $\boldsymbol{\Pi}_{\tilde{y}'}$ respectively which here depends only on $\boldsymbol{\theta}$ and are therefore denoted by $\mathbf{s}(\boldsymbol{\theta})$.

The circular covariances of the asymptotic distribution of $\text{vec}(\boldsymbol{\Pi}_{\tilde{y},T})$ and $\text{vec}(\boldsymbol{\Pi}_{\tilde{y}',T})$ are singular, but it is proved in [17] that the AMV bound definition can be extended and is given by

$$\mathbf{Cov}_{\boldsymbol{\theta}}^{\text{Alg}} \geq \mathbf{C}_{\boldsymbol{\theta}}^{\text{AMV}(\mathbf{s})} = [\mathbf{S}^H(\boldsymbol{\theta})\mathbf{C}_{\mathbf{s}}^{\dagger}(\boldsymbol{\theta})\mathbf{S}(\boldsymbol{\theta})]^{-1} \qquad (10)$$

where $\dagger$ denotes the Moore-Penrose inverse of the covariance $\mathbf{C}_{\mathbf{s}}$ of the asymptotic distribution of $\text{vec}(\boldsymbol{\Pi}_{\tilde{y},T})$ or $\text{vec}(\boldsymbol{\Pi}'_{\tilde{y},T})$ which are derived in Appendix C and where $\mathbf{S}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \frac{d\mathbf{s}(\boldsymbol{\theta})}{d\boldsymbol{\theta}}$.

Remark 1 Noting that the block Toeplitz matrix $\mathbf{R}_{\tilde{y},T}$ and the sample covariance $\frac{1}{T}\sum_{t=1}^{T} \tilde{\mathbf{y}}_t \tilde{\mathbf{y}}_t^H$ (which is not block Toeplitz structured) have the same asymptotic Gaussian

[4]Note that the local identifiability used in Result 2 is equivalent to that the Jacobian $\mathbf{S}$ is full column rank [10]. When $h_0 = 1$, the number of columns of $\mathbf{S}$ is reduced by 2 w.r.t. the case of an arbitrary $\mathbf{h}$ and $\mathbf{h}$ becomes locally identifiable for arbitrary channels.

[5]Note that the circular and the noncircular covariance matrices contain the same terms [16].

distribution [18], the circular covariance matrix $\mathbf{C}_{\mathbf{R}_{\tilde{y}}}$ of the asymptotic distribution of $\mathbf{R}_{\tilde{y},T}$ can be derived from that of $\frac{1}{T}\sum_{t=1}^{T}\tilde{\mathbf{y}}_t\tilde{\mathbf{y}}_t^H$. Based on the linear structure $\tilde{\mathcal{H}}(\mathbf{h})x_t+\tilde{\mathbf{n}}_t$ of the correlated data $\tilde{\mathbf{y}}_t$, the robustness [19, Result 3] property that states that the circular covariance of the asymptotic distribution of the associated projection matrix does not depend on the distribution of x_t but on the temporal correlation of $(x_t)_{t=\mathbb{N}}$ applies. Using $\tilde{\mathbf{y}}'_t=\tilde{\mathcal{H}}(\mathbf{h}')x_t+\tilde{\mathbf{n}}'_t$, this property extends to the MSK case. This is in contrast with $\mathbf{C}_{\mathbf{s}}(\boldsymbol{\theta})$, associated with the statistics $\mathbf{R}_{\tilde{y},T}$ and $\mathbf{R}_{\tilde{y}',T}$, which depend on the second and fourth-order terms of x_t. Naturally for QPSK symbols, $\mathbf{C}_{\mathbf{s}}(\boldsymbol{\theta})$ depends on the second, fourth, sixth and eighth-order terms.

Remark 2 Note that lower bounds (8), (7) and (9) associated with arbitrary consistent algorithms based on $\mathbf{R}_{\tilde{y}}$ or $\mathbf{R}_{\tilde{y}'}$ satisfy, for the same a priori knowledge:

$$\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\mathbf{R}_{\tilde{y}})}\leq\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\boldsymbol{\Pi}_{\tilde{y}})}\quad\text{and}\quad\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\mathbf{R}_{\tilde{y}'})}\leq\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\boldsymbol{\Pi}_{\tilde{y}'})}$$

where $\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\boldsymbol{\Pi}_{\tilde{y}})}$ and $\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\boldsymbol{\Pi}_{\tilde{y}'})}$ denote lower bound (10) associated with $\boldsymbol{\Pi}_{\tilde{y},T}$ and $\boldsymbol{\Pi}_{\tilde{y}',T}$ respectively, because these statistics are functions of $\mathbf{R}_{\tilde{y},T}$ and $\mathbf{R}_{\tilde{y}',T}$ respectively.

C. Subspace-based algorithms

Because the structure of the covariance matrices $\mathbf{R}_{\tilde{y}}$ and $\mathbf{R}_{\tilde{y}'}$ of (2) (4) and (5) are similar to those obtained in the SIMO case, all the algorithms devised in this case can be used in the present context. In the sequel by lack of place, only the so-called least square (LS) and subspace (SS) will be considered [20]. We note that (i) the LS and SS estimates coincide in the two-channel case [20], and (ii) the noise subspace is of dimension one.

For the BPSK modulation, any eigenvector $\tilde{\mathbf{v}}$ associated with the smallest eigenvalue of the block structured matrix $\mathbf{R}_{\tilde{y}}=\left[\begin{array}{cc}\times & \square\\ \square^* & \times^*\end{array}\right]$ satisfies $\tilde{\mathbf{v}}=(\mathbf{v}_1^T,\mathbf{v}_1^He^{i\phi})^T$ and from [21]:

$$\left[\begin{array}{c}\mathbf{h}\\ \mathbf{h}^*\end{array}\right]=c\left[\begin{array}{c}\mathbf{v}_1\\ -\mathbf{v}_1^*e^{i\phi}\end{array}\right].$$

From now, $\mathbf{v}_1$ is constrained to have its first component to be unity: $\mathbf{h}=\mathbf{v}_1$ ($c=1$, $\phi=\pi$). Consequently, the LS and SS estimates are given by $\mathbf{h}_T=\mathbf{v}_{1,T}$ where $\tilde{\mathbf{v}}_T=(\mathbf{v}_{1,T}^T,\mathbf{v}_{1,T}^He^{i\phi_T})^T$ denotes the eigenvector associated with the smallest eigenvalue of the block structured matrices $\mathbf{R}_{\tilde{y},T}=\left[\begin{array}{cc}\times & \square\\ \square^* & \times^*\end{array}\right]$ whose first component is unity.

The asymptotic performance of this algorithm can be deduced from the asymptotic distribution of $\tilde{\mathbf{v}}_T$ whose circular covariance $\mathbf{C}_{\tilde{\mathbf{v}}}$ is derived from the mapping[6] $\mathbf{R}_{\tilde{y},T}\longmapsto\tilde{\mathbf{u}}_T\longmapsto\tilde{\mathbf{v}}_T$ where $\tilde{\mathbf{u}}_T$ denotes the eigenvector associated with the smallest eigenvalue of $\mathbf{R}_{\tilde{y},T}$ satisfying the constraint $\tilde{\mathbf{u}}_T^H\tilde{\mathbf{u}}=1$ where $\tilde{\mathbf{u}}$ is an arbitrary unit norm eigenvector of $\mathbf{R}_{\tilde{y}}$ associated with its smallest eigenvalue.

[6]In practice, the unique eigenvector $\tilde{\mathbf{v}}_T$ is obtained from $\mathbf{R}_{\tilde{y},T}$ through an eigenvector $\tilde{\mathbf{u}}_T$ satisfying other constraints. For example, the SVD function of MATLAB forces all eigenvectors to have a real first element and a unit norm.

Using the standard perturbation theory [22, p. 162]

$$\delta\tilde{\mathbf{u}}=-\mathcal{S}^{\dagger}\boldsymbol{\Pi}_{\tilde{y}}\delta(\mathbf{R}_{\tilde{y}})\tilde{\mathbf{u}}+o(\delta\mathbf{R}_{\tilde{y}})$$

with $\mathcal{S}\stackrel{\mathrm{def}}{=}\tilde{\mathcal{H}}(\mathbf{h})\tilde{\mathcal{H}}(\mathbf{h})^H$, we obtain by the chain rule and the standard theorem of continuity (see e.g., [28, p. 122]):

$$\mathbf{C}_{\mathbf{h}}=\mathbf{C}_{\mathbf{v}_1}=\mathbf{D}\mathcal{S}^{\dagger}(\tilde{\mathbf{u}}^T\otimes\boldsymbol{\Pi}_{\tilde{y}})\mathbf{C}_{\mathbf{R}_{\tilde{y}}}(\tilde{\mathbf{u}}^*\otimes\boldsymbol{\Pi}_{\tilde{y}})\mathcal{S}^{\dagger}\mathbf{D}^H,\quad(11)$$

where $\mathbf{D}=\frac{1}{[\tilde{\mathbf{u}}]_1}[\mathbf{I}_{M+1},\mathbf{O}_{M+1}][\mathbf{I}_{2(M+1)}-[\tilde{\mathbf{v}},\mathbf{O}_{2(M+1),2M+1}]]$ is the Jacobian matrix associated with the mapping $\tilde{\mathbf{u}}_T\longmapsto\tilde{\mathbf{v}}_T\longmapsto\mathbf{v}_{1,T}$.

For the MSK and QPSK modulations, this analysis extends by replacing $\mathbf{R}_{\tilde{y},T}$ and $\mathbf{h}_T$ by $\mathbf{R}_{\tilde{y}',T}$ and $\mathbf{h}'_T$ respectively, and appending the mapping $\mathbf{h}'_T\longmapsto\mathbf{h}_T$.

V. Illustrative examples

Four experiments are considered to illustrate the performance of the impulse response estimates and particularly the robustness of estimators when the impulse response is close to unidentifiable.

In Figs. 1 and 2, the channel is given by $h(z)=(1-z_1^{-1}z)(1-z_2^{-1}z)$ with $z_1=0.8e^{i\alpha}$ and $z_2=1.25e^{i\pi/4}$ where α varies from 0 to 2π for $T=1000$, $SNR=\frac{\|\mathbf{h}\|^2}{\sigma_n^2}=15$dB and 100 Monte Carlo runs. These figures exhibit for independent BPSK and MSK inputs driven by independent symbols:

- the theoretical AMV bound $\mathrm{MSE}(\mathbf{h})=\frac{1}{T}\mathrm{Tr}(\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\mathbf{R}_{\tilde{y}})})$ with the prior information of independent symbols (denoted AMV1 bound) and without any a priori information (denoted AMV2 bound);
- the theoretical AMV bound $\mathrm{MSE}(\mathbf{h})=\frac{1}{T}\mathrm{Tr}(\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\boldsymbol{\Pi}_{\tilde{y}})})$ (denoted AMV3 bound);
- the theoretical asymptotic $\mathrm{MSE}(\mathbf{h})=\frac{1}{T}\mathrm{Tr}(\mathbf{Cov}_{\boldsymbol{\theta}}^{\mathrm{LS/SS}})$ given by the LS/SS algorithm;
- the actual (Monte Carlo) MSE given by the LS/SS algorithm.

The following relation is satisfied for arbitrary α as noted in Section IV

$$\begin{aligned}\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\mathbf{R}_{\tilde{y}})^{(1)}}&\leq\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\mathbf{R}_{\tilde{y}})^{(2)}}|_{\boldsymbol{\rho}=(0,\ldots,0,\sigma_n^2)^T}\\&\leq\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\boldsymbol{\Pi}_{\tilde{y}})}\leq\mathbf{Cov}_{\boldsymbol{\theta}}^{\mathrm{LS/SS}}.\end{aligned}$$

We note that the theoretical AMV bound $\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\mathbf{R}_{\tilde{y}})^{(2)}}$ is bounded when the zero z_1 approaches the real [resp. imaginary] axis for the BPSK [resp. MSK] modulation for which $\mathbf{h}$ becomes globally nonidentifiable but remains locally identifiable. This is in contrast with $\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\mathbf{R}_{\tilde{y}})^{(1)}}$, $\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\boldsymbol{\Pi}_{\tilde{y}})}$ and $\mathbf{Cov}_{\boldsymbol{\theta}}^{\mathrm{LS/SS}}$ which increase dramatically in this case. The behavior of the two bounds $\mathrm{AMV}(\mathbf{R}_{\tilde{y}})$ is explained by $\mathbf{S}=[\mathbf{S}_1,\mathbf{e}_1]$ in (7) which remains full rank contrary to $\mathbf{S}=[\mathbf{S}_1,\boldsymbol{\Psi}]$ in (8), which becomes rank deficient. The behavior of $\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\boldsymbol{\Pi}_{\tilde{y}})}$ and $\mathbf{Cov}_{\boldsymbol{\theta}}^{\mathrm{LS/SS}}$ is explained by the pseudoinverses $\mathbf{C}_{\mathbf{s}}^{\#}(\boldsymbol{\theta})$ and $\mathcal{S}^{\#}$ in (10) and (11) respectively, for which $\tilde{\mathcal{H}}(\mathbf{h})$ becomes rank deficient. This behavior of the LS/SS algorithm is interpreted by the "noise eigenvector" that is mistaken for a "signal eigenvector" when the channel is close to the non identifiability

conditions. If the length $L \geq M+1$ of the stacked observed output samples y_t increases, the dimension $2L-(2M+1)$ of the noise subspace increases and it is well known that the performance improves (see e.g., [23] where Fig. 1 shows the two channel case with order two), but at an expense of additive complexity. Furthermore, we note that for arbitrary processing window lengths $L \geq M+1$, only the correlations $r_y(0),...,r_y(M)$ and $r'_y(0),...,r'_y(M)$ are involved in the LS/SS algorithm for independent inputs x_t. Consequently, the AMV2 bound given if Fig. 1 keeps on being a lower bound for the MSE of the impulse response estimated by arbitrary second order algorithms derived for $L \geq M+1$ samples of stacked observed outputs. And we see from this figure that this AMV2 bound is also very sensitive to the non identifiable conditions.

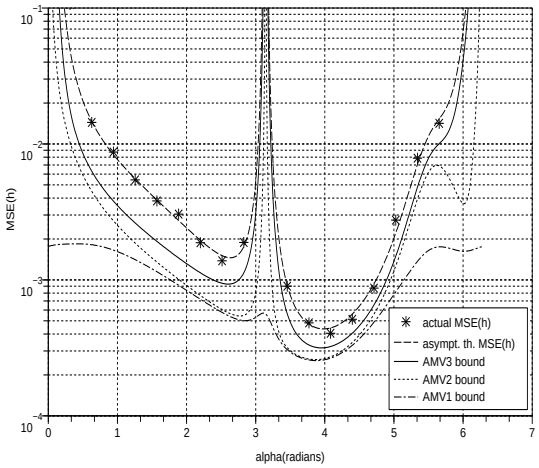

Fig.1 For white BPSK inputs three AMV bounds, asymptotic theoretical MSE(**h**) and ($\star$) actual MSE(**h**), given by the LS/SS algorithm as a function of the phase α of a zero of **h**.

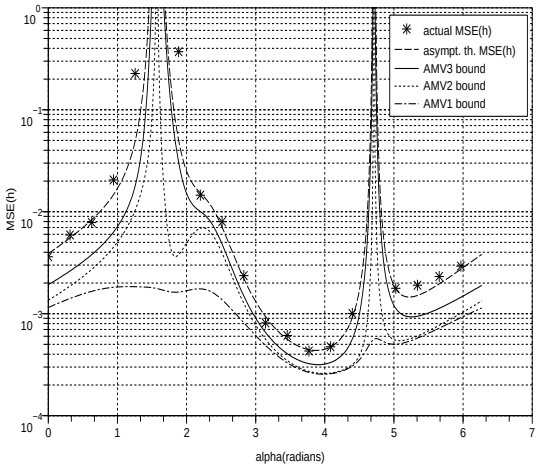

Fig. 2 For MSK inputs driven by white symbols 3 AMV bounds, asymptotic theoretical MSE(**h**) and ($\star$) actual MSE(**h**), given by the LS/SS algorithm as a function of the phase α of a zero of **h**.

We now consider the influence of the correlation of BPSK symbols on the performance of the identification of **h** by using a differential coding for which the symbols x_t are generated from $x_t = x_{t-1}b_t$ with b_t is an i.i.d. sequence where $P(b_t = 1) = p$ and $P(b_t = -1) = 1-p$. In this case x_t is a zero mean correlated sequence for which $r_x(k) = \mathrm{E}\{x_t x_{t-k}\} = (2p-1)^k$ and $\mathrm{E}\{x_t x_{t-k} x_{t-l} x_{t-m}\} = (2p-1)^{k+m-l}$ for $0 \leq k \leq l \leq m$ implying $\mathrm{Cum}(x_t, x_{t-k}, x_{t-l}, x_{t-m}) = -2(2p-1)^{l+m-k}$ for $0 \leq k \leq l \leq m$. Incorporating these values in (14) (15), Appendix A allows you to study the performance. Fig. 3 exhibits for the channel used in the previous figure with $\alpha = 4$ radians, as functions of the probability p for $T = 1000$ and $SNR = 5$dB:

- the theoretical AMV bound $\mathrm{MSE}(\mathbf{h}) = \frac{1}{T}\mathrm{Tr}(\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\mathbf{R}_{\bar{y}})})$ with the prior information of differential coding (denoted AMV4 bound) and without any a priori information (denoted AMV2 bound);
- the theoretical asymptotic MSE $(\mathbf{h}) = \frac{1}{T}\mathrm{Tr}(\mathbf{Cov}_{\boldsymbol{\theta}}^{\mathrm{LS/SS}})$ given by the LS/SS algorithm;
- the actual (Monte Carlo) MSE given by the LS/SS algorithm.

We see from this figure[7] that the different MSE are sensitive to the probability p except for $p \approx 0.5$ and degrade rapidly for a strong correlation i.e. $p \approx 0$, $p \approx 1$.

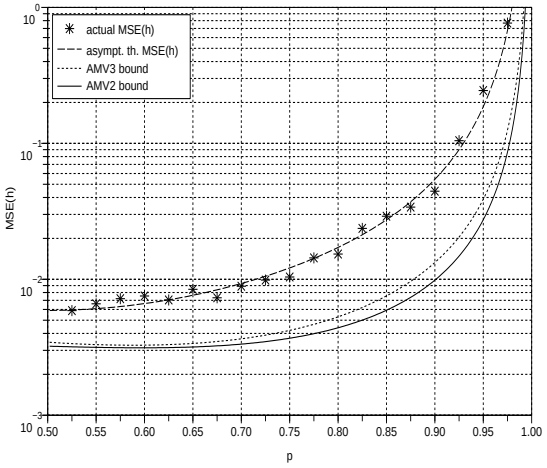

Fig. 3 For correlated BPSK symbols asymptotic, 2 AMV bounds, asymptotic theoretical MSE(**h**) and ($\star$) actual MSE(**h**), given by the LS/SS algorithm as a function of the probability p.

To test the occurrence of an unidentifiable channel, we consider now a random second-order FIR channel where $\Re(h_1)$, $\Im(h_1)$, $\Re(h_2)$ and $\Re(h_2)$ are i.i.d. uniformly distributed in $[-1,+1]$. For each realization of such a channel, the asymptotic theoretical MSE $(\mathbf{h}) = \frac{1}{T}\mathrm{Tr}(\mathbf{Cov}_{\boldsymbol{\theta}}^{\mathrm{LS/SS}})$ is computed with the associated matched SNR. Fig.4 exhibits the empirical distribution function of the random variable MSE(**h**) from 1000 realizations for $T = 1000$. From this figure, we see for example that $P\{\mathrm{MSE}(\mathbf{h}) < 3\ 10^{-2}\} = 0.5$, for $SNR = 15$dB. Comparing this median with the different MSE given by the LS/SS algorithm for the specific channel of Fig. 1 for this same SNR, we see that this specific channel is representative enough. Consequently, channels close to unidentifiable are not uncommon.

[7]We note that the different MSE are symmetric with respect to $p = 0.5$.

Unfortunately for more realistic random channels, the empirical distribution function of the random variable MSE($\mathbf{h}$) is much worse. For example, for Clarke filters in the typical urban mode, where each generated channel is specular and contains six paths, where the transmit filter is a raised cosine with a rolloff of 0.1, and where delays and attenuation standard deviations are given in [24] according to the ETSI norm, the probability of a high asymptotic theoretical MSE($\mathbf{h}$) is much larger than for the channels of Fig. 4.

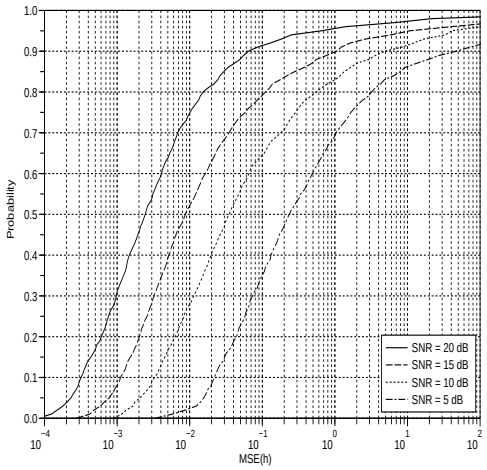

Fig. 4 For white BPSK symbols, empirical distribution function of the random variable MSE($\mathbf{h}$) for four values of the SNR.

In the last experiment, we consider the channel given by $h(z) = 1 - z\sqrt{z_1^{-1} + z_2^{-1}} + z^2\sqrt{z_1^{-1}z_2^{-1}}$ (where the real part of the square roots are positive) associated with $h'(z) = (1 - z_1^{-1}z)(1 - z_2^{-1}z)$ with $z_1 = 0.8e^{i\alpha}$ and $z_2 = 1.25e^{i\pi/4}$ where α varies from 0 to 2π for $T = 100000$, $SNR = 20$dB and 100 Monte Carlo runs. This figure exhibits for independent QPSK symbols:

- the theoretical AMV bound MSE($\mathbf{h}$) $= \frac{1}{T}\mathrm{Tr}(\mathbf{C}_{\boldsymbol{\theta}}^{\mathrm{AMV}(\mathbf{R}_{\tilde{y}'})})$ based on the statistics used in the LS/SS algorithm;
- the theoretical asymptotic MSE ($\mathbf{h}$) $= \frac{1}{T}\mathrm{Tr}(\mathbf{Cov}_{\boldsymbol{\theta}}^{\mathrm{LS/SS}})$ given by the LS/SS algorithm;
- the theoretical asymptotic MSE ($\mathbf{h}$) $= \frac{1}{T}\mathrm{Tr}(\mathbf{Cov}_{\boldsymbol{\theta}}^{\mathrm{Ckq}})$ given by the so-called $C(k,q)$ formula of Giannakis [27], [2] ($h(k,T) = c_{y,T}(0,k,2)/c_{y,T}(0,0,2)$ where $c_{y,T}(t_1,t_2,t_3)$ denotes the estimated cumulant associated with the sample moment $\frac{1}{T}\sum_{t=1}^{T} y_t y_{t+t_1} y_{t+t_2} y_{t+t_3}$;
- the actual (Monte Carlo) MSE given by the LS/SS algorithm;
- the actual (Monte Carlo) MSE given by the $C(k,q)$ formula of Giannakis [27], [2].

We see from this figure that the AMV bound outperforms by far the LS/SS algorithm, and the latter is outperformed in a large domain of α by the $C(k,q)$ formula of Giannakis[8]. Next, we note that these algorithms require very large data samples for an effective estimation (here $T = 100000$). Furthermore when $\mathbf{h}$ becomes nonidentifiable, i.e. α approaches 0 or π, the performance of the LS/SS algorithms degrades in contrast to those of the $C(k,q)$ formula of Giannakis. Finally, we note that the theoretical and actual MSE's of the LS/SS present some discontinuities in the neighborhood of $\alpha = \frac{3\pi}{4}$ for which the real part of h_2 is zero, implying some artifacts in the derivation of the theoretical asymptotic MSE given by the LS/SS algorithm and large errors in the estimates.

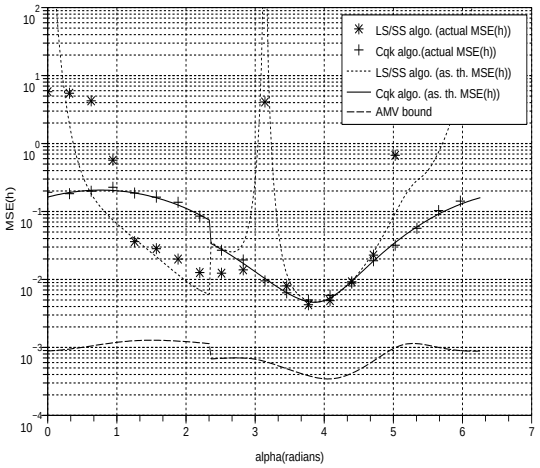

Fig. 5 For white QPSK symbols, AMV bound, asymptotic theoretical MSE($\mathbf{h}$) and actual MSE($\mathbf{h}$) given by the LS/SS and the $C(k,q)$ algorithms as a function of the phase α of a zero of $\mathbf{h}'$.

VI. Conclusion

In this paper, the problem of blindly estimating FIR SISO channels has been considered, when the channel is excited by discrete inputs. Identifiability has been stated from second order statistics of transformed data. One may say that the knowledge of the alphabet has induced a form of diversity. Identifiability conditions have been derived and limitations have been pointed out: some particular channels are not identifiable. Theoretical performance issues have been addressed in order to evaluate the robustness of standard subspace-based estimators with respect to these identifiability conditions, as well as the performance of the so-called $C(k,q)$-algorithm [27], [2]. More precisely, it has been shown that the performance of alphabet-based algorithms seriously degrades in the neighborhood of nonidentifiable channels, implying that these approaches are unapplicable in practice. On the other hand, the HOS-based algorithms not based on subspace approaches, as the $C(k,q)$-algorithm which does not utilize noncircularity, do not present difficult identifiability conditions and consequently are much more robust.

[8] Of course, this algorithm may be outperformed by other HOS-algorithms that use these statistics more efficiently. This $C(k,q)$ algorithm has nevertheless been selected simply to reveal the lack of robustness of the standard subspace alphabet-based estimators.

APPENDIX

Appendix A: derivation of $\mathbf{C_s}(\boldsymbol{\alpha})$ associated with the BPSK and MSK modulation

To derive the expression of the circular covariance of the asymptotic distribution of $\mathbf{s}_T$, we number the channels $\mathbf{h}$ and $\mathbf{h}^*$ by $\mathbf{h}^{(1)}$ and $\mathbf{h}^{(2)}$ respectively and the associated output by $y_t^{(1)}$ and $y_t^{(2)}$. Consequently the terms of $\mathbf{C_s}(\boldsymbol{\alpha}) = \lim_{T\to\infty} T\mathrm{E}\{(\mathbf{s}_T - \mathbf{s}(\boldsymbol{\alpha}))(\mathbf{s}_T - \mathbf{s}(\boldsymbol{\alpha}))^H\}$ are constituted by the terms

$$\lim_{T\to\infty} T\mathrm{E}\{(r_{i,j,T}(k) - r_{i,j}(k))((r_{i',j',T}(l) - r_{i',j'}(l))\} \quad (12)$$

for $k = 0,...,M$, $i \le j$, $l = 0,...,M$, $i' \le j'$ and $k = 1,...,M$, $i > j$, $l = 1,...,M$, $i' > j'$, $i,j,i',j' = 1,2$ to eliminate redundancies, in which $r_{i,j,T}(k) \stackrel{\mathrm{def}}{=} \frac{1}{T}\sum_{t=1}^{T} y_t^{(i)} y_{t-k}^{(j)}$ and $r_{i,j}(k) \stackrel{\mathrm{def}}{=} \mathrm{E}\{y_t^{(i)} y_{t-k}^{(j)}\}$ for the BPSK modulation. Limit (12) deduced from the Bartlett formula (see e.g., [25]) is given by

$$\sum_{n=-M}^{M} r_{i,i'}(n) r_{j,j'}(n+l-k) + r_{i,j'}(n+l) r_{j,i'}(n-k) + \mathrm{Cum}(y_n^{(i)}, y_{n-k}^{(j)}, y_0^{(i')}, y_{-l}^{(j')}) \quad (13)$$

with

$$r_{i,j}(k) = \sum_{l=k}^{M} h_l^{(i)} h_{l-k}^{(j)} + \sigma_n^2 \delta_k \delta_{i-1} \delta_{j-2}$$

and

$$\mathrm{Cum}(y_n^{(i)}, y_{n-k}^{(j)}, y_0^{(i')}, y_{-l}^{(j')}) = \kappa_4(x) \sum_{t=0}^{M} h_{t+n}^{(i)} h_{t+n-k}^{(j)} h_t^{(i')} h_{t-l}^{(j')}$$

where $\kappa_4(x) \stackrel{\mathrm{def}}{=} \mathrm{E}\{x_t^4\} - 3\mathrm{E}\{x_t^2\} = -2$ in the white case and with

$$r_{i,j}(k) = \sum_{l=0}^{M}\sum_{l'=0}^{M} h_l^{(i)} h_{l'}^{(j)} r_x(k+l'-l) + \sigma_n^2 \delta_k \delta_{i-1} \delta_{j-2} \quad (14)$$

and

$$\mathrm{Cum}(y_n^{(i)}, y_{n-k}^{(j)}, y_0^{(i')}, y_{-l}^{(j')}) = \sum_{t_1=0}^{M}\sum_{t_2=0}^{M}\sum_{t_3=0}^{M}\sum_{t_4=0}^{M} \mathrm{Cum}(x_{n-t_1}, x_{n-k-t_2}, x_{-t_3}, x_{-l-t_4}) h_{t_1}^{(i)} h_{t_2}^{(j)} h_{t_3}^{(i')} h_{t_4}^{(j')} \quad (15)$$

in the general correlated case.

The expression of the circular covariance of the asymptotic distribution of $\mathbf{s}_T$ associated with the MSK modulation driven by white symbols follows the same lines as the white BPSK case replacing y_t and h_t by $y'_t =$ and h'_t respectively according to (3).

Appendix B: derivation of $\mathbf{C_s}(\boldsymbol{\theta})$ associated with the QPSK modulation

First, we note that the statistics $\mathbf{s}_T$ is a simple function of the statistics $\mathbf{r}_T \stackrel{\mathrm{def}}{=} [r_{y,T}(0),...r_{y,T}(M), r_{y,T}^*(1),...r_{y,T}^*(M), r_{y,T}(0),...r_{y,T}(M), r_{y,T}^*(1),...r_{y,T}^*(M), r'_{y',T}(0),..., r'_{y',T}(M), r_{y',T}^{'*}(0), ..., r_{y',T}^{'*}(M)]^T$. Consequently, using the standard theorem of continuity (see e.g., [28]) on regular functions of asymptotically Gaussian statistics, $\mathbf{C_s}(\boldsymbol{\alpha})$ is derived from $\mathbf{C_r}(\boldsymbol{\alpha})$ by

$$\mathbf{C_s}(\boldsymbol{\theta}) = \frac{\partial \mathbf{s}}{\partial \mathbf{r}^T} \mathbf{C_r}(\boldsymbol{\alpha}) \frac{\partial \mathbf{s}^H}{\partial \mathbf{r}^T}.$$

Then the different terms of $\mathbf{C_r}(\boldsymbol{\alpha}) = \lim_{T\to\infty} T\mathrm{E}\{(\mathbf{r}_T - \mathbf{r}(\boldsymbol{\alpha}))(\mathbf{r}_T - \mathbf{r}(\boldsymbol{\alpha}))^H\}$ similar to terms (12) are given (see e.g., [26, rel.10.5.2]) by the following finite limits:

$$\sum_t \Big(\mathrm{E}\{y_0^{(i_1)} y_0^{(i_2)} y_k^{(i_3)} y_k^{(i_4)} y_t^{(i_5)} y_t^{(i_6)} y_{t+l}^{(i_7)} y_{t+l}^{(i_8)}\} - \mathrm{E}\{y_0^{(i_1)} y_0^{(i_2)} y_k^{(i_3)} y_k^{(i_4)}\}\mathrm{E}\{y_0^{(i_5)} y_0^{(i_6)} y_l^{(i_7)} y_l^{(i_8)}\}\Big),$$

$$\sum_t \Big(\mathrm{E}\{y_0^{(i_1)} y_k^{(i_2)} y_l^{(i_3)} y_t^{(i_4)} y_{t+m}^{(i_5)} y_{t+n}^{(i_6)}\} - \mathrm{E}\{y_0^{(i_1)} y_k^{(i_2)} y_l^{(i_3)}\}\mathrm{E}\{y_0^{(i_4)} y_m^{(i_5)} y_n^{(i_6)}\}\Big),$$

$$\sum_t \Big(\mathrm{E}\{y_0^{(i_1)} y_k^{(i_2)} y_t^{(i_3)} y_{t+l}^{(i_4)}\} - \mathrm{E}\{y_0^{(i_1)} y_k^{(i_2)}\}\mathrm{E}\{y_0^{(i_3)} y_l^{(i_4)}\}\Big)$$

where the exponents i_k take values 1 or 2. The last sum reduces to (13) using the Bartlett formula but the first two sums are much more intricate. We could use the extended Bartlett formulas derived in [25] in the case of infinite sums (ARMA models), but here because all the sums are finite, we use a specific symbolic calculus akin to a high-level language to automatically generate the expressions of these sums. This calculus uses $y_t^{(i)} = \sum_{k=0}^{M} h_k^{(i)} x_{t-k} + n_t^{(i)}$, the circularity of n_t with $\mathrm{E}\{|n_t^{2p}|\} = p!\,\sigma_n^2$ and the following specific properties of the QPSK alphabet, namely $\mathrm{E}\{(x_t)^p (x_t^*)^q\} = \begin{cases} 1 & \text{for} |p-q| = 0 \equiv 4 \\ 0 & \text{elsewhere} \end{cases}$.

Appendix C: derivation of the covariance of the asymptotic distribution of $\mathrm{vec}(\boldsymbol{\Pi}_{\tilde{y},T})$ or $\mathrm{vec}(\boldsymbol{\Pi}_{\tilde{y}',T})$

First, noting that the Hermitian block Toeplitz matrix $\mathbf{R}_{\tilde{y},T}$ is built from its first column and row constituted by the statistics $\mathbf{s}_T$ given by (6), the circular covariance $\mathbf{C}_{\mathbf{R}_{\tilde{y}}}$ of the asymptotic distribution of $\mathbf{R}_{\tilde{y},T}$ is deduced from $\mathbf{C_s}(\boldsymbol{\theta})$ given in Appendix A by

$$\mathbf{C}_{\mathbf{R}_{\tilde{y}}} = \frac{d\mathrm{vec}(\mathbf{R}_{\tilde{y}})}{d\mathbf{s}^T} \mathbf{C_s}(\boldsymbol{\theta}) \frac{d\mathrm{vec}(\mathbf{R}\tilde{y})^T}{d\mathbf{s}^T}.$$

Using the standard perturbation result for orthogonal projectors [30] (see also [29]) applied to $\boldsymbol{\Pi}_{\tilde{y}}$ associated with the noise subspace of $\mathbf{R}_{\tilde{y}}$

$$\delta(\boldsymbol{\Pi}_{\tilde{y}}) = -\boldsymbol{\Pi}_{\tilde{y}} \delta(\mathbf{R}_{\tilde{y}}) \mathcal{S}^\dagger - \mathcal{S}^\dagger \delta(\mathbf{R}_{\tilde{y}}) \boldsymbol{\Pi}_{\tilde{y}} + o\left(\delta(\mathbf{R}_{\tilde{y}})\right) \quad (16)$$

with $\mathcal{S} \stackrel{\mathrm{def}}{=} \tilde{\mathcal{H}}(\mathbf{h}) \mathbf{R}_x \tilde{\mathcal{H}}(\mathbf{h})^H$, the asymptotic behaviors of $\mathrm{vec}(\boldsymbol{\Pi}_{\tilde{y},T})$ and $\mathbf{R}_{\tilde{y},T}$ are directly related. The standard theorem (see e.g., [28, p. 122]) on regular functions of asymptotically Gaussian statistics applies and the circular covariance matrix of the asymptotically Gaussian distribu-

tion of $\mathrm{vec}(\mathbf{\Pi}_{\tilde{y},T})$ can be written as

$$\mathbf{C}_{\mathbf{\Pi}_{\tilde{y}}} = \left((\mathbf{\Pi}_{\tilde{y}}^{*} \otimes \mathcal{S}^{\dagger}) + (\mathcal{S}^{\dagger^{*}} \otimes \mathbf{\Pi}_{\tilde{y}})\right) \mathbf{C}_{\mathbf{R}_{\tilde{y}}} \left((\mathbf{\Pi}_{\tilde{y}}^{*} \otimes \mathcal{S}^{\dagger}) + (\mathcal{S}^{\dagger^{*}} \otimes \mathbf{\Pi}_{\tilde{y}})\right).$$

For the case of MSK and QPSK symbols, $\mathbf{C}_{\mathbf{\Pi}_{\tilde{y}'}}$ is deduced in the same way from the circular covariance matrix $\mathbf{C}_{\mathbf{s}}(\boldsymbol{\theta})$ of the asymptotic distribution of the associated second-order statistics of transformed data where $\mathcal{S} = \tilde{\mathcal{H}}(\mathbf{h}')\tilde{\mathcal{H}}(\mathbf{h}')^{H}$.

REFERENCES

[1] J. K. Tugnait, L. Tong, and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 17–28, May 2000.

[2] G. Giannakis, Y. Inouye, and J. Mendel, "Cumulant-based identification of multichannel moving-average models," *IEEE Trans. Automatic Control*, vol. 34, no. 7, pp. 783–787, Jul. 1989.

[3] E. Moulines et al., "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Sig. Proc.*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[4] M. Kristensson, B. Ottersten, and D. Slock, "Blind subspace identification of a BPSK communication channel," in *Asilomar*, Pacific Grove, CA, , pp. 828–832, Nov. 1997.

[5] D. Darsena, G. Gelli, L. Paura, F. Verde, "Subspace-based blind channel identification of SISO-FIR systems with improper random inputs," *Signal Processing*, vol. 84, pp. 2021-2039, 2004.

[6] Z. Ding, G. Li, "Single-channel blind equalization for GSM cellular systems," *IEEE J. on Sel. A. in Comm.*, vol. 16, no. 8, Oct. 1998.

[7] P. Comon, "Contrasts, Independent Component Analysis, and Blind Deconvolution," *Int. Journal Adapt. Control Sig. Proc.*, Wiley, vol. 18, no. 3, pp. 225–243, April 2004.

[8] P. Chevalier and F. Pipon, "New Insights into optimal widely linear array receivers for the demodulation of BPSK, MSK and GMSK signals corrupted by noncircular interferences - Application to SAIC," *IEEE Trans. Sig. Proc.*, Vol 54, no. 3, pp. 870-883, March 2006.

[9] G.B. Giannakis and S.D. Halford, "Performance analysis of blind equalizers based on cyclostationarity statistics," *in Proc. 26th Conf. Inform. Sci. Syst., Princeton*, pp. 711–716, March 1994.

[10] H.H. Zeng and L. Tong, "Blind channel estimation using the second order statistics: asymptotic performance and limitations," *IEEE Trans. Sig. Proc.*, vol. 45, no. 8, pp. 2060–20701, Aug. 1997.

[11] K. Abed-Meraim et al., "On subspace methods for blind identification of SIMO FIR systems," *IEEE Trans. Sig. Proc.*, vol. 45, no. 1, pp. 42–55, Jan. 1997, Special issue on communications.

[12] M. Kendall and A. Stuart, *The Advanced Theory of Statistics, Distribution Theory*, vol. 1, C. Griffin, 1977.

[13] P. McCullagh, *Tensor Methods in Statistics*, Monographs on Statistics and Applied Probability. Chapman and Hall, 1987.

[14] B. Porat and B. Friedlander, "Asymptotic accuracy of ARMA parameter estimation methods based on sample covariances," *Proc.7th IFAC/IFORS Symposium on Identification and System Parameter Estimation, York*, 1985.

[15] P. Stoica, B. Friedlander and T. Söderström, "An approximate maximum approach to ARMA spectral estimation," *in Proc. Decision and control, Fort Lauderdale*, 1985.

[16] J.P. Delmas, "Asymptotically minimum variance second-order estimation for non-circular signals with application to DOA estimation," *IEEE Trans. Sig. Proc.*, vol. 52, no. 5, pp. 1235–1241, May 2004.

[17] H. Abeida, J.P. Delmas, "Efficiency of subspace-based DOA estimators," *Signal Processing*, vol. 87, no. 9, pp. 2075-2084, Sept. 2007.

[18] J.P. Delmas, "Asymptotic normality of sample covariance matrix for mixed spectra time series: Application to sinusoidal frequencies estimation," *IEEE Transactions on Information theory*, vol. 47, no. 4, pp. 1681-1687, May 2001.

[19] J.P. Delmas, "Asymptotic performance of second-order algorithms," *IEEE Transactions on Signal Processing*, vol. 50, no. 1, pp. 49-57, January 2002.

[20] H. Zeng and L. Tong, "Connections between the least squares and subspace approaches to blind channel estimation," *IEEE Trans. Sig. Proc.*, vol. 44, pp. 1593-1596, June 1996.

[21] G. Xu, H. Liu, L. Tong and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, pp. 2982-2993, Dec. 1995.

[22] J.R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*, Wiley series in probability and statistics, revised edition, 1999.

[23] W. Qui, Y. Hua, "Performance analysis of the subspace method for blind channel identification," *Signal Processing*, vol. 50, pp. 71-81, 1996.

[24] O. Grellier, P. Comon, B. Mourrain, and P. Trebuchet, "Analytical blind channel identification," *IEEE Trans. Signal Processing*, vol. 50, no. 9, pp. 2196–2207, Sept. 2002.

[25] J.P. Delmas, Y. Meurisse, "Extension of the matrix Bartlett's formula to the third and fourth order and to noisy linear models with application to parameter estimation," *IEEE Trans. Sig. Proc.*, vol. 53, no. 8, pp. 2765–2776, August 2005.

[26] B. Porat, *Digital processing of random signals, Theory and Methods*, Prentice Hall, 1994.

[27] G.B. Giannakis, "Cumulants: A powerful tool in signal processing," *Proc. of the IEEE*, vol. 75, no. 9, pp. 1333-1334, Sept. 1987.

[28] R.J. Serfling, *Approximation Theorems of Mathematical Statistics*, John Wiley and Sons, 1980.

[29] H. Krim, P. Forster and G. Proakis, "Operator approach to performance analysis of root-MUSIC and root-min-norm," *IEEE Trans. Signal Processing*, vol. 40, no. 7, pp. 1687-1696, July 1992.

[30] T. Kato, *Perturbation Theory for Linear Operators*, Springer Berlin, 1995.

PLACE PHOTO HERE

Jean Pierre Delmas (M'00-SM'06) was born in France in 1950. He received the Engineering degree from Ecole Centrale de Lyon, France in 1973, the Certificat d'études supérieures from the Ecole Nationale Supérieure des Télécommunications, Paris, France in 1982 and the Habilitation à diriger des recherches (HDR) degree from the University of Paris, Orsay, France in 2001. Since 1980, he has been with the Institut TELECOM SudParis (ex INT), where he is presently a Professor in the CITI department and Deputy director of UMR CNRS 5157 (SAMOVAR) laboratory. His teaching and research interests are in the areas of statistical signal processing with application to communications and antenna array. He has served as Associate Editor with the IEEE Transactions on Signal Processing (2002-2006) and is an IEEE Senior Member.

PLACE PHOTO HERE

Yann Meurisse received the Engineer degree in Telecommunications from the Institut National des Télécommunications (INT), Evry, France in 1983, the Diplôme d'Etudes Approfondies in Automatic Control and Signal Processing and the Ph.D degree from University of Paris Sud, Orsay in 1987 and 2002 respectively. Since 1985, he has been with the Institut TELECOM SudParis (ex INT), where he is currently an Assistant Professor in the CITI Department. His research interests are in antenna array processing.

PLACE PHOTO HERE

Pierre Comon (M'87 - SM'95 - F'07) graduated in 1982, and received the Doctorate degree in 1985, both from the University of Grenoble, France. He later received the Habilitation to Lead Researches in 1995, from the University of Nice, France. He has been for nearly 13 years in industry, first with Crouzet-Sextant, Valence, France, between 1982 and 1985, and then with Thomson Marconi, Sophia-Antipolis, France, between 1988 and 1997. He spent 1987 with the ISL laboratory, Stanford University, CA. He joined in 1997 the Eurecom Institute, Sophia-Antipolis, France, and left during the fall of 1998. He is now research director with CNRS since 1998 at laboratory I3S, Sophia-Antipolis, France, where he heads both the STIC Doctoral School and the SIS department. His research interests include High-Order Statistics (HOS), Blind Deconvolution and Equalization, tensor-based Factor Analysis and Statistical Signal and Array Processing.

Dr. Comon was Associate Editor of the *IEEE Transactions on Signal Processing* from 1995 to 1998, and a member of the French National Committee of Scientific Research from 1995 to 2000. He was the coordinator of the European Basic Research Working Group on HOS, from 1992 to 1995. Between 1992 and 1998, he was a member of the Technical and Scientific Council of the Thomson Group. Between 2001 and 2004 he acted as launching Associate Editor with the *IEEE Transactions on Circuits and Systems I*, in the area of Blind Techniques. He is currently Associate Editor of the Elsevier journal *Signal Processing*, and a member of several IEEE Technical Committees.