Performance limits of alphabet diversities for FIR SISO channel identification

Jean-Pierre Delmas, Yann Meurisse, Pierre Comon

To cite this version:
Jean-Pierre Delmas, Yann Meurisse, Pierre Comon. Performance limits of alphabet diversities for FIR SISO channel identification. IEEE Transactions on Signal Processing, Institute of Electrical and Electronics Engineers (IEEE), 2009, 57 (1), pp.73-82. <10.1109/TSP.2008.2007926>. <hal-00354686>

HAL Id: hal-00354686
https://hal.archives-ouvertes.fr/hal-00354686
Submitted on 20 Jan 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
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) \cite{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 \cite{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 \cite{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 \cite{9}, 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 \cite{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.
(i.e. $E\{n_i^2\} = 0$), white, zero-mean and with finite variance $\sigma_n^2 = E\{|n_i|^2\}$. 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:

$$ y_t = (y_t, y_{t-1}, ..., y_{t-M})^T = H(h)x_t + n_t \quad (1) $$

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

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

with $h \overset{\text{def}}{=} (h_0, h_1, ..., h_M)^T$. 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 $R_x = E\{x_t x_t^H\}$ 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 $\tilde{y}_t = (y_t^T, y_{t-1}^T)^T$, so that from (1):

$$ R_{\tilde{y}} \overset{\text{def}}{=} E\{\tilde{y}_t \tilde{y}_t^H\} = \begin{bmatrix} H(h) & \mathcal{H}(h^*)^H \end{bmatrix} R_x \begin{bmatrix} H(h) & \mathcal{H}(h^*)^H \end{bmatrix} + \sigma_n^2 I_{2M}. $$

Consequently, we obtain this way a structured covariance matrix similar to that obtained in the SIMO case; here the two channels have impulse responses $h$ and $h^*$. Therefore the results (see e.g., \cite{1}) concerning the global identifiability of SIMO channels can be applied. Because $2(M + 1) > 2M + 1$ and $R_x$ is nonsingular, the range space of the filtering matrix $H(h)$ is identifiable from $R_{\tilde{y}}$, and this range space determines the channel coefficients up to a multiplicative constant if channels $h$ and $h^*$ 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 $h$ 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 \cite{2} 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 \cite{2}, the impulse response $h$ of a SISO channel is locally identifiable from the circular and noncircular second-order statistics of its output if and only if $h(1) \neq 0$.

**Remark**: If this condition is satisfied, there is no other channel in the neighborhood of a channel $h$, 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 $h(1) \neq 0$ but with real zeros or conjugated zeros will be exhibited in Section \cite{3}. This is in contrast to the global identifiability, for which we know that if a channel is identifiable, there is no other channel in the space $C^{M+1}$ 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 $x_{t+1} \overset{\text{def}}{=} i x_t c_t$ 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 $x_t = i b_t x_0$ 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

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

with $h'_k = x_0 h_k i^{-k}, k = 0, ..., M$ and where $n'_t = i^{-t} n_t$ 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 $y'_t = (y'_t^T, y'_{t-1}^T)^T$ we obtain $R_{y'} = I_{2M+1}$ in contrast to $R_x$ of (2) which is not necessarily equal to $I_{2M+1}$ from the assumptions

$$ R_{y'} \overset{\text{def}}{=} E\{y'_t y'_t^H\} = \begin{bmatrix} \mathcal{H}(h') & \mathcal{H}(h')^H \end{bmatrix} + \sigma_n^2 I_{2M}. $$

\(^1\) 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 \cite{2}, this demodulation is performed on vector $y_t$ and consequently $y'_t = y_t i^{-t}$ becomes stationary at the second order but not $y'_t$.\[\text{IEEE TRANSACTIONS ON SIGNAL PROCESSING}\]
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 $h$ 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 $h(\pm i) \neq 0$].

IV. ASYMPOTIC PERFORMANCE

The above identifiability results naturally raise the important issue of performance analysis of algorithms based on the second-order covariance $R_{\hat{g}}$, and on the modified second-order statistics $R_{\hat{g}'}$. 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 $(y_t)_{t=1,\ldots,T}$ 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 $h$ is identifiable from $R_{\hat{g}}$ or $R_{\hat{g}'}$ 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 $h$ up to a multiplicative constant.

To compare the asymptotic performance of these algorithms to given AMV bounds based on various statistics, $h$ must be strictly identifiable from these statistics. Consequently, one parameter of $h$ is fixed to a predefined value, say, $h_0 = 1$ throughout Sections 4 and 5.

Let $\alpha = (\theta^T, \rho^T)^T$ denote the real-valued unknown parameters (containing the real and imaginary parts of the complex parameters) of the extended covariance matrices $R_{\hat{g}}$ or $R_{\hat{g}'}$, where $\theta \defeq [\theta(h_1), \ldots, \theta(h_M), M(h_1), \ldots, M(h_M)]^T$, and where $\rho$ collects the nuisance parameters for the BPSK and MSK modulations. Depending on the a priori knowledge of the inputs, $\rho = \sigma^2_n$ if the BPSK sequence $x_t$ or the driving sequence $b_t$ for the MSK modulation are white, and $\rho = [r_x(1), \ldots, r_x(2M), \sigma^2_n]^T$ with $r_x(k) \defeq E\{x_t x_{t-k}\}$ for correlated BPSK symbols. We note that if the conditions of Section 4 are satisfied $\alpha$ is identifiable from $R_{\hat{g}}$ or $R_{\hat{g}'}$.

These block Toeplitz matrices $R_{\hat{g}}$ and $R_{\hat{g}'}$ are traditionally consistently estimated from $T$ successive received

---

3Note that all the previous performance analyses (e.g., [10]) consider that the nuisance parameters are known, and consequently give optimistic AMV bounds.
signals $y_t$ by replacing the various expectations by the associated sample correlations: $r_{y,T}(\ell) \equiv \frac{1}{T} \sum_{t=1}^{T} y_t y_{t-\ell}$, $r'_{y,T}(\ell) \equiv \frac{1}{T} \sum_{t=1}^{T} y_t y_{t+\ell}$ and $r''_{y,T}(\ell) \equiv \frac{1}{T} \sum_{t=1}^{T} y_t^2$. In order to apply the AMV bound to an arbitrary consistent extended second-order algorithm based on the sample estimates $R_{y,T}$ or $R_{y',T}$ respectively, the involved statistics $s_T$ must collect real-valued sample correlations and complex-valued sample correlations and their conjugate. For example for BPSK modulations, we have:

$$s_T = (r_{y,T}(0), r_{y,T}(1), ..., r_{y,T}(M), r'_{y,T}(1), ..., r'_{y,T}(M), r''_{y,T}(0), ..., r''_{y,T}(M))^T.$$  

(6)

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

For white BPSK sequences and white driving sequences of MSK modulations, $\sigma_n^2$ is the unique nuisance parameter and $r_y(k) = \sum_{i=k}^{M} h_i h_{i-k} + \delta_i \sigma_n^2$ and $r'_y(k) = \sum_{i=k}^{M} h_i h_{i-k}$ for BPSK sequences. Consequently $s(\alpha)$ is structured as $s(\alpha) = \Phi(\theta) + \sigma_n^2 e_1$, where $e_1$ is the first unit vector in $\mathbb{R}^{1M+3}$. This implies $S = [S_1, e_1]$, and the matrix inversion lemma yields:

$$\text{Cov}_\vartheta^{\text{Alg}} \geq C_\vartheta^{\text{AMV}(R_y)}(\phi)(1),$$  

(7)

with

$$C_\vartheta^{\text{AMV}(R_y)}(\phi)(1) \equiv [S^H(\alpha)C_s^{\text{Alg}}(\alpha)S(\alpha)]^{-1}$$

$$= \left[ \begin{array}{c} S_1 \ 0 \end{array} \right]^{-1} \left[ \begin{array}{c} S_1^H \ e_1^H \end{array} \right]^{-1} \left[ \begin{array}{c} S_1^H \ e_1 \end{array} \right]^{-1}$$

$$= \left[ S^H_1 S_1^{-1} \Pi_{C_s^{\text{Alg}}(\alpha)} S_1^{-1} \right]^{-1},$$

where $\Pi_{\Lambda}$ denotes the projector onto the orthogonal complement of the columns of $\Lambda$.

For the BPSK modulation with correlated symbols, $r_y(k) = \sum_{i=0}^{M} \sum_{j=0}^{M} h_i h_{j-k} r_x(k+j-i) + \delta_i \sigma_n^2$ and $r'_y(k) = \sum_{i=0}^{M} \sum_{j=0}^{M} h_i h_{j-k} r_x(k+j-i)$ where $r_x(k) = 1$. Consequently $s(\alpha)$ is

$$s(\alpha) = (r_y(0), r_y(1), ..., r_y(M), r'_y(1), ..., r'_y(M), r''_y(0), ..., r''_y(M))^T.$$  

(4)

Note that the local identifiability used in Result 3 is equivalent to that the Jacobian $S$ is full column rank [14]. When $h_0 = 1$, the number of columns of $S$ is reduced by 2 w.r.t. the case of an arbitrary $h$ and $h$ becomes locally identifiable for arbitrary channels.

3Note that the circular and the noncircular covariance matrices contain the same terms [4].

B. AMV bounds based on orthogonal projectors

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

The circular covariances of the asymptotic distribution of $\text{vec}(\Pi_{y,T})$ and $\text{vec}(\Pi_{y',T})$ are singular, but it is proved in [7] that the AMV bound definition can be extended and is given by

$$\text{Cov}_\vartheta^{\text{Alg}} \geq C_\vartheta^{\text{AMV}(s)} = [S^H(\theta)C_s^{\text{Alg}}(\theta)S(\theta)]^{-1}$$

(10)

where $\Pi$ denotes the Moore-Penrose inverse of the covariance $C_s^{\text{Alg}}$ of the asymptotic distribution of $\text{vec}(\Pi_{y,T})$ or $\text{vec}(\Pi_{y',T})$ which are derived in Appendix C and where

$$S(\theta) \equiv \frac{ds(\theta)}{d\theta}.$$  

Remark 1: Noting that the block Toeplitz matrix $R_{y,T}$ and the sample covariance $\frac{1}{T} \sum_{t=1}^{T} y_t y_H^T$ (which is not block Toeplitz structured) have the same asymptotic Gaussian
distribution \([3]\), the circular covariance matrix \(C_{R_g}\) of the asymptotic distribution of \(R_{g,T}\) can be derived from that of \(\frac{1}{T} \sum_{t=1}^{T} \tilde{y}_t \tilde{y}_t^H\). Based on the linear structure \(\tilde{y}(h) x_t + \tilde{u}_t\) of the correlated data \(\tilde{y}_t\), the robustness \([12]\). 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 \((x_t)_t \in \mathbb{N}\) applies. Using \(\tilde{y}^* = \tilde{\chi}(h) x_t + \tilde{u}_t\), this property extends to the MSK case. This is in contrast with \(C_{\mathcal{S}}(\theta)\), associated with the statistics \(R_{\tilde{g},T}\) and \(R_{g',T}\), which depend on the second and fourth-order terms of \(x_t\). Naturally for QPSK symbols, \(C_{\mathcal{S}}(\theta)\) depends on the second, fourth, sixth and eighth-order terms.

Remark 2 Note that lower bounds \([3, 1]\) and \([3, 2]\) associated with arbitrary consistent algorithms based on \(R_{\tilde{g}}\) or \(R_{g'}\) satisfy, for the same a priori knowledge:

\[
C_{\mathcal{S}}^{AMV}(R_{\tilde{g}}) \leq C_{\mathcal{S}}^{AMV}(\Pi_{\tilde{g}}) \leq C_{\mathcal{S}}^{AMV}(R_{g'}) \leq C_{\mathcal{S}}^{AMV}(\Pi_{g'})
\]

where \(C_{\mathcal{S}}^{AMV}(\Pi_{\tilde{g}})\) and \(C_{\mathcal{S}}^{AMV}(\Pi_{g'})\) denote lower bound \([3, 2]\) associated with \(\Pi_{\tilde{g},T}\) and \(\Pi_{g',T}\) respectively, because these statistics are functions of \(R_{\tilde{g},T}\) and \(R_{g',T}\) respectively.

C. Subspace-based algorithms

Because the structure of the covariance matrices \(R_\phi\) and \(R_{\phi'}\) of \([3, 1]\) and \([3, 2]\) 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 \(\tilde{v}\) associated with the smallest eigenvalue of the block structured matrix \(R_{\tilde{g}} = \begin{bmatrix} \times & \tilde{v} \times \end{bmatrix} \) satisfies \(\tilde{v} = (v_1^T, v_2^H e^{i\phi})^T\) and from \([21]\):

\[
\begin{bmatrix} h \\ h^* \end{bmatrix} = c \begin{bmatrix} v_1 \\ -v_1^* e^{i\phi} \end{bmatrix}.
\]

From now, \(v_1\) is constrained to have its first component to be unity: \(h = v_1 (c = 1, \phi = \pi)\). Consequently, the LS and SS estimates are given by \(h_{T} = v_{1,T}\) where \(v_{1,T} = (v_1^T, v_2^H e^{i\phi_T})^T\) denotes the eigenvector associated with the smallest eigenvalue of the block structured matrices \(R_{\tilde{g},T} = \begin{bmatrix} \times & \times \\ \times^* & \times^* \end{bmatrix}\) whose first component is unity.

The asymptotic performance of this algorithm can be deduced from the asymptotic distribution of \(v_{1,T}\) whose circular covariance \(C_{\mathcal{V}}\) is derived from the mapping \(R_{\tilde{g},T} \mapsto \tilde{u}_T \mapsto v_{1,T}\) where \(\tilde{u}_T\) denotes the eigenvector associated with the smallest eigenvalue of \(R_{\tilde{g},T}\) satisfying the constraint \(\tilde{u}_T^H \tilde{u} = 1\) where \(\tilde{u}\) is an arbitrary unit norm eigenvector of \(R_{\tilde{g}}\) associated with its smallest eigenvalue.

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

\[
\delta \tilde{u} = -S^T \Pi_{\tilde{g}} \delta (R_{\tilde{g}}) \tilde{u} + o(\delta R_{\tilde{g}})
\]

with \(S \overset{\text{def}}{=} \tilde{y}(h) \tilde{y}(h)^H\), we obtain by the chain rule and the standard theorem of continuity (see e.g., \([20, p. 122]\)):

\[
C_{h} = C_{v_1} = D S^T (\tilde{u}_T^\dagger \Pi_{\tilde{g}}) C_{R_{\tilde{g}}} (\tilde{u}_T^* \Pi_{\tilde{g}}) S^T D^H,
\]

where \(D = \frac{1}{\|v_T\|} [I_{M+1}, O_{M+1}] [I_{2(M+1)} - [v, O_{2(M+1)}]2(M+1)]\) is the Jacobian matrix associated with the mapping \(\tilde{u}_T \mapsto v_T \mapsto v_{1,T}\).

For the MSK and QPSK modulations, this analysis extends by replacing \(R_{\tilde{g},T}\) and \(h_T\) by \(R_{g',T}\) and \(h_{T}^r\) respectively, and appending the mapping \(h_T^r \mapsto h_T\).

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 \(z_1 = 0.8 e^{i\alpha}\) and \(z_2 = 1.25 e^{i\pi/4}\) where \(\alpha\) varies from 0 to \(2\pi\) for \(T = 1000\), \(SNR = \frac{|\|h\|^2|}{\sigma^2} = 15dB\) and 100 Monte Carlo runs. These figures exhibit for independent BPSK and MSK inputs driven by independent symbols:

- the theoretical AMV bound MSE(h) = \(\frac{1}{T} \text{Tr}(C_{\mathcal{S}}^{AMV}(R_{\tilde{g}}))\) with the prior information of independent symbols (denoted AMV1 bound) and without any a priori information (denoted AMV2 bound);
- the theoretical AMV bound MSE(h) = \(\frac{1}{T} \text{Tr}(C_{\mathcal{S}}^{AMV}(\Pi_{\tilde{g}}))\) (denoted AMV3 bound);
- the theoretical asymptotic MSE (h) = \(\frac{1}{T} \text{Tr}(C_{\mathcal{S}}^{LS/SS})\) given by the LS/SS algorithm;
- the actual (Monte Carlo) MSE given by the LS/SS algorithm.

The following relation is satisfied for arbitrary \(\alpha\) as noted in Section \(V\):

\[
C_{\mathcal{S}}^{AMV}(R_{\tilde{g}}) \overset{(1)}{\leq} C_{\mathcal{S}}^{AMV}(R_{\tilde{g}}) \overset{(2)}{\leq} C_{\mathcal{S}}^{AMV}(\Pi_{\tilde{g}}) \overset{(3)}{\leq} C_{\mathcal{S}}^{AMV}(\Pi_{g'}) \overset{(4)}{\leq} C_{\mathcal{S}}^{LS/SS}.
\]

We note that the theoretical AMV bound \(C_{\mathcal{S}}^{AMV}(R_{\tilde{g}})\) is bounded when the zero \(z_1\) approaches the real [resp. imaginary] axis for the BPSK [resp. MSK] modulation for which \(h\) becomes globally nonidentifiable but remains locally identifiable. This is in contrast with \(C_{\mathcal{S}}^{AMV}(R_{\tilde{g}})\) and \(C_{\mathcal{S}}^{AMV}(\Pi_{\tilde{g}})\) which increase dramatically in this case. The behavior of the two bounds AMV(R_{\tilde{g}}) is explained by \(S = [S_1, S_2]\) in \([3, 2]\) which remains full rank contrary to \(S = [S_1, \Psi]\) in \([3, 1]\), which becomes rank deficient. The behavior of \(C_{\mathcal{S}}^{AMV}(\Pi_{\tilde{g}})\) and \(C_{\mathcal{S}}^{LS/SS}\) is explained by the pseudoinverses \(C_{\mathcal{S}}^{\dagger}(\theta)\) and \(S^\#\) in \([3, 2]\) respectively, for which \(\tilde{y}(h)\) 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 \( 2L - 2(M + 1) \) 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 \( r_p(0), \ldots, r_p(M) \) and \( r_p'(0), \ldots, r_p'(M) \) 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.

1. In this case \( x_t \) is a zero mean correlated sequence for which \( r_z(k) = \mathbb{E}(x_t x_{t-k}) = (2p - 1)^k \) and \( \mathbb{E}(x_t x_{t-k} x_{t-l} x_{t-m}) = (2p - 1)^{k+m-l} \) for \( 0 \leq k \leq l \leq m \) implying \( \text{Cum}(x_t, x_t-k, x_t-l, x_t-m) = -2(2p - 1)^{l+m-k} \) for \( 0 \leq k \leq l \leq m \). Incorporating these values in (14) (5), Appendix A allows you to study the performance.

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

Fig. 2 For MSK inputs driven by white symbols 3 AMV bounds, asymptotic theoretical MSE(h) and \((*)\) actual MSE(h), given by the LS/SS algorithm as a function of the phase \( \alpha \) 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 \( b_t \) an i.i.d. sequence where \( P(b_t = 1) = p \) and \( P(b_t = -1) = 1 - p \). We note that the different MSE are symmetric with respect to \( p = 0.5 \).

Fig. 3 For correlated BPSK symbols asymptotic, 2 AMV bounds, asymptotic theoretical MSE(h) and \((*)\) 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 \( \Im(h_2) \) are i.i.d. uniformly distributed in \([-1, +1]\). For each realization of such a channel, the asymptotic theoretical MSE \((h) = \frac{1}{T} \text{Tr}(\text{Cov}_{\theta}^{\text{LS/SS}})\) 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 \( P\{\text{MSE}(h) < 3 \times 10^{-2}\} = 0.5 \), for \( SNR = 15 \text{dB} \). 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.

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 $\text{MSE}(h)$ 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 $\text{MSE}(h)$ is much larger than for the channels of Fig. 4.

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 $h'(z) = (1 - z_1^{-1} z_2^{-1})(1 - z_2^{-1})$ with $z_1 = 0.8 e^{i\alpha}$ and $z_2 = 1.25 e^{i\pi/4}$ where $\alpha$ varies from 0 to 2$\pi$ for $T = 100000$, $SNR = 20$dB and 100 Monte Carlo runs. This figure exhibits for independent QPSK symbols:

- the theoretical AMV bound $\text{MSE}(h) = \frac{1}{T} \text{Tr}(C^\text{AMV}(R_y))$ based on the statistics used in the LS/SS algorithm;
- the theoretical asymptotic $\text{MSE}(h) = \frac{1}{T} \text{Tr}(C^\text{LS/SS})$ given by the LS/SS algorithm;
- the theoretical asymptotic $\text{MSE}(h) = \frac{1}{T} \text{Tr}(C^\text{Cov}(\theta))$ 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 $\frac{1}{T} \sum_{t=1}^T y_1 y_{t_1} y_{t_2} y_{t_3}$;
- the actual (Monte Carlo) $\text{MSE}$ given by the LS/SS algorithm;
- the actual (Monte Carlo) $\text{MSE}$ 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 $\alpha$ by the $C(k, q)$ formula.

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 non-identifiable 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 $C_n(\alpha)$ associated with the BPSK and MSK modulation

To derive the expression of the circular covariance of the asymptotic distribution of $s_T$, we number the channels $h$ and $h^*$ by $h^{(1)}$ and $h^{(2)}$ respectively and the associated output by $y^{(1)}$ and $y^{(2)}$. Consequently the terms of $C_n(\alpha) = \lim_{T \to \infty} TE\{ (s_T - s(\alpha))(s_T - s(\alpha))^H \}$ are constituted by the terms

$$\lim_{T \to \infty} TE\{ (r_{i,j,T}(k) - r_{i,j}(k))(r_{i',j',T}(l) - r_{i',j'}(l)) \}$$

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

$$\sum_{n=-M}^{M} r_{i,j}(n) r_{i',j'}(n + l - k) + \text{Cum}(y^{(i)}_{n}, y^{(j)}_{n-k}, y^{(i')}_{l}, y^{(j')}_{l-k})$$

with

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

and

$$\text{Cum}(y^{(i)}_{n}, y^{(j)}_{n-k}, y^{(i')}_{l}, y^{(j')}_{l-k}) = \kappa_4(x) \sum_{l=0}^{M} h^{(i)}_{l+n} h^{(j)}_{l}$$

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

$$r_{i,j}(k) = \sum_{l=0}^{M} \sum_{l'=0}^{M} h^{(i)}_{l} h^{(j)}_{l'} x_0(k + l' - l) + \sigma_n^2 \delta_{l-1} \delta_{j-2}$$

and

$$\text{Cum}(y^{(i)}_{n}, y^{(j)}_{n-k}, y^{(i')}_{l}, y^{(j')}_{l-k}) = \kappa_4(x) \sum_{t_1=0}^{M} \sum_{t_2=0}^{M} \sum_{t_3=0}^{M} \sum_{t_4=0}^{M} \text{Cum}(x_{n-t_1}, x_{n-k-t_2}, x_{-t_3}, x_{-t_4}) \kappa_4(x)$$

in the general correlated case.

The expression of the circular covariance of the asymptotic distribution of $s_T$ 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 $C_n(\theta)$ associated with the QPSK modulation

First, we note that the statistics $s_T$ is a simple function of the statistics $r_T \equiv [r_{y,T}(0), ..., r_{y,T}(M), r_{y,T}^*(1), ..., r_{y,T}^*(M), r_{y,T}(0), ..., r_{y,T}(M)].$ Consequently, using the standard theorem of continuity (see e.g., [28]) on regular functions of asymptotically Gaussian statistics, $C_n(\alpha)$ is derived from $C_n(\theta)$ by

$$C_n(\theta) = \frac{\partial s}{\partial \theta} C_n(\theta) \frac{\partial s^H}{\partial \theta}$$

Then the different terms of $C_n(\theta) = \lim_{T \to \infty} TE\{ (r_T - r(\theta))(r_T - r(\theta))^H \}$ similar to terms (12) are given (see e.g., [24] ref.10.5.2) by the following finite limits:

$$\sum_{t} \left( E\{ y_0^{(i_1)} y_0^{(i_2)} y_k^{(i_3)} y_k^{(i_4)} y_t^{(i_5)} y_t^{(i_6)} \} - E\{ y_0^{(i_1)} y_0^{(i_2)} y_k^{(i_3)} y_k^{(i_4)} y_t^{(i_5)} y_t^{(i_6)} \} \right)$$

$$\sum_{t} \left( E\{ y_0^{(i_1)} y_0^{(i_2)} y_k^{(i_3)} y_k^{(i_4)} y_t^{(i_5)} y_t^{(i_6)} \} - E\{ y_0^{(i_1)} y_0^{(i_2)} y_k^{(i_3)} y_k^{(i_4)} y_t^{(i_5)} y_t^{(i_6)} \} \right)$$

$$\sum_{t} \left( E\{ y_0^{(i_1)} y_0^{(i_2)} y_k^{(i_3)} y_k^{(i_4)} y_t^{(i_5)} y_t^{(i_6)} \} - E\{ y_0^{(i_1)} y_0^{(i_2)} y_k^{(i_3)} y_k^{(i_4)} y_t^{(i_5)} y_t^{(i_6)} \} \right)$$

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 $y_t^{(i)} = \sum_{k=0}^{M} h^{(i)}_{k} x_{t-k} + n_t^{(i)}$, the circularity of $n_t$ with $E\{ n_t^{(p)} n_t^{(q)} \} = p^l \sigma_n^2$ and the following specific properties of the QPSK alphabet, namely

$$E\{ x_t^{|p-q|} x_t^{|q-p|} \} = \begin{cases} 1 & \text{for } |p-q| = 0 \equiv 4 \\ 0 & \text{elsewhere} \end{cases}$$

Appendix C: derivation of the covariance of the asymptotic distribution of vec($\Pi_{\hat{g}} T$) or vec($\Pi_{\hat{g}} T$)

First, noting that the Hermitian block Toeplitz matrix $R_{\hat{g}} T$ is built from its first column and row constituted by the statistics $s_T$ given by (6), the circular covariance $C_{R_{\hat{g}}}$ of the asymptotic distribution of $R_{\hat{g}}$ is deduced from $C_n(\theta)$ given in Appendix A by

$$C_{R_{\hat{g}}} = \frac{dvec(R_{\hat{g}})}{ds T} C_n(\theta) \frac{dvec(R_{\hat{g}})^T}{ds T}$$

Using the standard perturbation result for orthogonal projectors [20] (see also [25]) applied to $\Pi_{\hat{g}}$ associated with the noise subspace of $R_{\hat{g}}$

$$\delta(\Pi_{\hat{g}}) = -\Pi_{\hat{g}} \delta(R_{\hat{g}}) S^T - S^H \delta(R_{\hat{g}}) \Pi_{\hat{g}} + o(\delta(R_{\hat{g}}))$$

with $S \equiv \tilde{H}(h) R_{\hat{g}} \tilde{H}(h)^H$, the asymptotic behaviors of vec($\Pi_{\hat{g}} T$) and $R_{\hat{g}} T$ 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 $\text{vec}(\Pi_y^T)$ can be written as
\[
C_{\Pi_y} = \left( \left( \Pi_y^T \otimes S^T \right) + \left( S^T \otimes \Pi_y^T \right) \right) C_R y + \left( \left( \Pi_y^T \otimes S^T \right) + \left( S^T \otimes \Pi_y^T \right) \right) C_{\Pi_y^T}.
\]

For the case of MSK and QPSK symbols, $C_{\Pi_y}$ is deduced in the same way from the circular covariance matrix $C_R(\theta)$ of the asymptotic distribution of the associated second-order statistics of transformed data where $S = \hat{h}(\hat{h})^H$.

**REFERENCES**


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.

Yann Meurisse received the Engineer degree in Telecommunications from the Institut National des Télécommunications (INT), Evry, France in 1985, the Diplôme d’Études 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.
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.