Analytical Blind Channel Identification

Olivier Grellier, Pierre Comon, *Senior Member, IEEE*, Bernard Mourrain, and Philippe Trébuchet

***Abstract*—In this paper, a novel analytical blind single-input single-output (SISO) identification algorithm is presented, based on the noncircular second-order statistics of the output. It is shown that statistics of order higher than two are not mandatory to restore identifiability. Our approach is valid, for instance, when the channel is excited by phase shift keying (PSK) inputs. It is shown that the channel taps need to satisfy a polynomial system of degree 2 and that identification amounts to solving the system. We describe the algorithm that is able to solve this particular system entirely analytically, thus avoiding local minima. Computer results eventually show the robustness with respect to noise and to channel length overdetermination. Identifiability issues are also addressed.**

***Index Terms*—Blind channel estimation, minimum shift keying, multipath channels, noncircularity, second-order statistics, time-varying channels.**

I. Introduction

BLIND identification methods depend on the characteristics of the input sources. For example, it is known that a system can only be identified up to an allpass filter when its input is Gaussian circular. Consequently, particular attention has been paid to the non-Gaussian inputs during the last two decades. In those situations, the phase information can be accessed using high-order statistics of the observations, and in the single-input single-output (SISO) case, the system is identified up to a scalar factor only. This has been studied in numerous papers, including the works of Shalvi–Weinstein [24] or Tugnait [28]. Here, we focus our attention on the noncircular character of inputs.

An interesting class of noncircular signals is the discrete, which appears in wireless communications. In the SISO case, the discrete character has been used by few authors; Li [15], and Yellin and Porat [31] proposed deterministic approaches. The former is valid for binary inputs and is iterative. To our knowledge, the latter, which is quite complicated, is the only work available in the open literature addressing analytical blind identification of SISO channels with discrete inputs; it includes a clustering stage that is rather sensitive to noise. On the other hand, the discrete character has been broadly used for equalization [15] but often in an iterative manner [4]; key references are not cited here since equalization is out of the scope of the present paper. The constant modulus (CM) property, which is widely used in blind equalization, can hardly be used in blind identification.

The studied signals also have nonzero cyclo-stationary statistics, which allows identification using second-order statistics only [11], [16]. However, for those signals, it is more interesting to use the cyclo-stationarity as a time diversity, which leads to the study of SIMO systems.

Slock in [25] and Tong *et al.* in [26] have first taken advantage of oversampling of cyclo-stationary sources. With single-input multiple-output (SIMO) systems, second-order statistics only can be used, provided that the channels do not share a common root [1], [23], [30]. In this sense, the SIMO problem can be considered to be easier than the SISO, in which one conventionally resorts either to cyclostationarity (which induces diversity) or to high-order moments, e.g., constant modulus algorithms (CMAs). SIMO second-order methods can be divided into three families:

1) subchannel response matching (SRM) approach introduced by Xu [30],
2) subspace methods [18];
3) linear prediction techniques [2], [25].

In this paper, the oversampling method (inducing a diversity) is not used, i.e., only SISO systems are studied. The novelty of our contribution is twofold. First, only second-order moments are used; they are shown to be sufficient to restore identifiability without resorting to higher order statistics. Second, an algebraic solution to a class of polynomial systems, constructed from a block of data, is introduced. Our approach is described mainly in the case of minumum shift keying (MSK) modulations, effectively approximating the digital modulation used in the GSM standard, but it holds valid for differential binary PSK (DBPSK) or quadrature PSK (QPSK) modulations. In addition, block methods are well matched to burst-mode communication systems (TDMA).

For instance, at 900 MHz and 190 km/h, the coherence time is of order 2 ms; in the GSM system, this corresponds to only two bursts, or about 300 symbol periods. This example shows that block algorithms become necessary in a blind context and for reduced coherence times.

The paper is organized as follows. Section II introduces the assumptions made on the input and the related second-order properties. Section III describes the principles of the novel procedure used to solve polynomial systems; this procedure is detailed in Appendix C, whereas the standard technique of resultants is recalled in Appendix A but is not used in the paper. The selection of the best solution is described in Section IV. Some identifiability results are proved in Section V, and computer experiments are eventually presented in Section VI.

II. Model and Basic Properties

Assume that a finite sequence of input samples $x(m)$ is fed into a finite impulse response (FIR) linear system of length M,

Manuscript received June 30, 2000; revised January 17, 2002. The associate editor coordinating the review of this paper and approving it for publication was Prof. Hideaki Sakai.

O. Grellier is with Amadeus Development, Sophia-Antipolis, France.

P. Comon is with the the I3S Laboratory, Algorithmes/Euclide-B, Sophia Antipolis, France.

B. Mourrain and P. Trébuchet are with Galuud, Inrai Institute, Sophia Antipolis, France.

Publisher Item Identifier 10.1109/TSP.2002.801887.

with (*a priori* complex) taps $h(m)$, $0 \leq m \leq M-1$. Denote as $y(n)$ the corresponding output sequence of length N, satisfying

$$y(n) = \sum_{m=0}^{M-1} h(m)x(n-m) + w(n) \overset{\text{def}}{=} \boldsymbol{x}(n;M)^{\mathrm{T}}\boldsymbol{h} + w(n)$$

where $w(n)$ stands for a noise with unknown distribution, and $(^{\mathrm{T}})$ denotes transposition. In a standard manner, multidimensional variables are stored in column vectors and denoted by boldface letters; for instance, $\boldsymbol{x}(n;M) = [x(n), \ldots x(n-M+1)]^{\mathrm{T}}$ by construction.

The input sequence $x(m)$ is i.i.d. and assumed to follow a discrete distribution, stemming from BPSK, MSK, or QPSK digital modulations [5], [22], and the channel $\boldsymbol{h}$ is supposed time-invariant during the observation record, which can be very short. The noise is introduced to take into account modeling errors, and computer experiments are run in Section VI for various noise levels. However, noise is ignored in the theoretical developments so that it is considered only to be a nuisance for its distribution is assumed to be unknown.

Complex Gaussian random variables are nothing but a pair of real random variables. What allows simpler expressions of its distribution, moments, and related statistical objects is its circularity, which induces a correlation between real and imaginary parts [12], [29]. For a scalar random variable Z, the circularity property at order 2 is characterized by the equation $\mathrm{E}\{Z^2\} = 0$. A random variable is referred to as *noncircular* at order 2 if the latter moment is nonzero.

For non-Gaussian random variables, strict-sense circularity means invariance of the distribution by multiplication of a unit modulus complex number (that is, a rotation in the complex plane), hence, the terminology. The concept has been introduced independently in [6] and [21]. Various properties are investigated in depth in [21]. Some statistical aspects have been addressed in [3]. Random variables whose distribution is not circularly invariant are referred to as *noncircular.*

The key statistical property used in this paper is that discrete signals are noncircular at given orders (at order k for a PSK-k random variable [3], [14]). However, only second-order statistics are used, so that only *noncircularity at order 2* will be exploited. More precisely, for DBPSK modulated signals, noncircular and circular second-order correlations are given by

$$\begin{aligned} \mathrm{E}\{x(n)x(n-\ell)|x(0)\} &= x(0)^2\delta(\ell) \\ \mathrm{E}\{x(n)x(n-\ell)^*\} &= \delta(\ell) \end{aligned} \tag{1}$$

respectively. Next, we have, for MSK signals

$$\mathrm{E}\{x(n)x(n-\ell)|x(0)\} = (-1)^n x(0)^2\delta(\ell) \tag{2}$$

$$\mathrm{E}\{x(n)x(n-\ell)^*|x(0)\} = \delta(\ell) \tag{3}$$

and last, for DQPSK modulated signals

$$\begin{aligned} \mathrm{E}\{\mathrm{Re}[x(n)]\mathrm{Re}[x(n-\ell)]|x(0)\} &= \mathrm{Re}\,[x(0)]^2\,\delta(\ell) \\ \mathrm{E}\{\mathrm{Im}[x(n)]\mathrm{Im}[x(n-\ell)]|x(0)\} &= \mathrm{Im}[x(0)]^2\delta(\ell) \\ \mathrm{E}\{x(n)x(n-\ell)^*\} &= \delta(\ell) \end{aligned}$$

where $\delta(\ell) \overset{\text{def}}{=} 1$ if $\ell = 0$ and $\delta(\ell) = 0$ elsewhere. Note the conditional expectation, which is necessary under the assumption that the initial value $x(0)$ is uniformly distributed, exhibiting cyclostationarity in the noncircular moment of MSK inputs.

Based on these properties, it is possible to derive a set of polynomial equations that the channel must satisfy. In the MSK case, we obtain

$$\mathrm{E}[y(n)y(n-\ell)|x(0)] = x(0)^2 \sum_{m=0}^{M-1} (-1)^m h(m)h(m+\ell) \tag{4}$$

and in the BPSK case

$$\mathrm{E}[y(n)y(n-\ell)|x(0)] = x(0)^2 \sum_{m=0}^{M-1} h(m)h(m+\ell). \tag{5}$$

In the QPSK case, we consider the equivalent problem (up to a rotation of $\pi/4$) of a QAM4 distributed source, where $x(n)$ is the sum of purely real and purely imaginary binary white and processes. It is necessary to consider real and imaginary parts separately at the receiver because $x(0)^2$ is not deterministic, whereas the real and imaginary parts of $x(n)$ have a deterministic square. This yields four families of equations. For simplicity, setting $\mathrm{Re}[x(0)]^2 = \mathrm{Im}[x(0)]^2 = 1$ without restricting the generality, one gets

$$\begin{aligned} E[y^r(n)y^r(n-\ell)|x(0)] &= \sum_q a(q)a(q+\ell) - b(q)b(q+\ell) \\ E[y^r(n)y^i(n-\ell)|x(0)] &= \sum_q a(q+\ell)b(q) - a(q)b(q+\ell) \\ E[y^i(n)y^r(n-\ell)|x(0)] &= \sum_q a(q)b(q+\ell) - a(q+\ell)b(q) \\ E[y^i(n)y^i(n-\ell)|x(0)] &= \sum_q b(q)b(q+\ell) + a(q)a(q+\ell) \end{aligned} \tag{6}$$

where $a(m)$ and $b(m)$ denote the real and imaginary parts of $h(m)$, and $y^r(n)$ and $y^i(n)$ those of $y(n)$, respectively.

Another obvious possibly would be to use fourth-order moments, which would yield the family of equations

$$\begin{aligned} &E[y(n)y(n-\ell_1)y(n-\ell_2)y(n-\ell_3)|x(0)] \\ &\qquad = x(0)^4 \sum_{m=0}^{M-1} h(m)h(m+\ell_1)h(m+\ell_2)h(m+\ell_3). \end{aligned}$$

In this paper, only polynomial systems of degree 2 will be considered; therefore, the latter property will not be utilized.

III. Solving the Polynomial System

In order to concentrate on principles, we will explain in detail the algorithm in the case of an MSK input, which seems to be a good compromise between simplicity of developments and generality. The algorithm described in Section III-D is, nevertheless, valid for other cases. Without restricting the generality, assume a channel of length $M = 3$. Then, from (4), the polynomial system given above based on noncircular statistics can be explicitly written as

$$\begin{cases} f_1(\boldsymbol{h}) = h(0)^2 - h(1)^2 + h(2)^2 - \alpha_0 = 0 \\ f_2(\boldsymbol{h}) = h(0)h(1) - h(1)h(2) - \alpha_1 = 0 \\ f_3(\boldsymbol{h}) = h(0)h(2) - \alpha_2 = 0 \end{cases} \tag{7}$$

where $\boldsymbol{h} = \{h(0), h(1), h(2)\}$ denotes the taps vector. The goal of this paper is to solve this polynomial system for taps $h(m)$. In a polynomial system having generally several solutions in a finite number, equations provided by standard circular statistics (covariance matching) are used to pick up the best solution in a final stage.

A. Example in the Case of a Real Channel

As a simple particular case, consider a real channel, but outside this section, the channel is *always assumed to be complex with no real roots,* in accordance with identifiability results proved in Section V. Then, circular statistics yield

$$\begin{cases} g_1(\boldsymbol{h}) = h(0)^2 + h(1)^2 + h(2)^2 - \beta_0 = 0 \\ g_2(\boldsymbol{h}) = h(0)h(1) + h(1)h(2) - \beta_1 = 0 \\ g_3(\boldsymbol{h}) = h(0)h(2) - \alpha_2 = 0 \end{cases}$$

where $\alpha_i = \mathrm{E}\{y(n)y(n-i)|x(0)\}$ and $\beta_i = \mathrm{E}\{y(n)y(n-i)^*|x(0)\}$ are given (they depend on statistics of observations y). Grouping of those equations results in

$$\begin{cases} h(0)^2 + h(2)^2 = (\alpha_0 + \beta_0)/2 \\ h(0)h(1) = (\alpha_1 + \beta_1)/2 \\ h(0)h(2) = \alpha_2. \end{cases}$$

Using the first and third equations, one obtains

$$\begin{aligned} (h(0) - \jmath h(2))^2 &= h(0)^2 + h(2)^2 - 2\jmath h(0)h(2) \\ &= (\alpha_0 + \beta_0)/2 - 2\jmath\alpha_2. \end{aligned}$$

This equation eventually allows the calculation of $h(0)$ and $h(2)$, up to a sign, and then $h(1)$.

This particular example shows that it is possible to identify a real channel by using the *noncircular second-order* statistics together with *circular second-order* ones, using a simple elimination procedure. Of course, this was valid only for real channels. For general complex FIR channels, which is the case in which we are actually interested, the elimination procedure is more complicated and is described in Section III-D.

B. Preliminaries

Consider the ring $\mathcal{R} = \mathbb{C}[\boldsymbol{\xi}]$ of polynomials in variables $\boldsymbol{\xi} \overset{\text{def}}{=} [\xi(0), \xi(1), \ldots \xi(M-1)]$ with coefficients in the complex field $\mathbb{C}$; the dual space of $\mathcal{R}$ is the set of linear forms from $\mathcal{R}$ to $\mathbb{C}$, which is denoted as $\hat{\mathcal{R}}$. The evaluation of a polynomial p at a point $\zeta \in \mathbb{C}^M$, which is denoted as $\mathbf{1}_\zeta\colon p \mapsto p(\zeta)$, is the linear form that most interests us.

Let $\{f_1, \ldots, f_M\}$ be polynomials of degree D belonging to $\mathcal{R}$.

Definition III.1: The ideal $\mathcal{I}$ spanned by polynomials $\{f_1, \ldots, f_M\}$ is the set of polynomials $p \in \mathcal{R}$ of the form

$$p = \sum_{i=1}^{M} f_i q_i, \qquad \text{with } q_i \in \mathcal{R}.$$

The quotient ring $\mathcal{A} = \mathcal{R}/\mathcal{I}$ is then defined as follows.

Definition III.2: For any ideal $\mathcal{I}$ included in $\mathcal{R}$, the quotient algebra $\mathcal{A} = \mathcal{R}/\mathcal{I}$ is the set of polynomial classes $p \in \mathcal{R}$ modulo ideal $\mathcal{I}$, *viz*

$$p \equiv q \quad \text{iff} \quad p - q \in \mathcal{I}.$$

The dual space $\hat{\mathcal{A}}$ of $\mathcal{A}$ is the subspace of $\hat{\mathcal{R}}$ of linear forms vanishing on the ideal $\mathcal{I}$. In particular, the evaluation $\mathbf{1}_\zeta$ is in $\hat{\mathcal{A}}$ if and only if ζ is a root of all polynomials belonging to $\mathcal{I}$. This is the fundamental property on which our approach is based.

Given a polynomial $a \in \mathcal{A}$, define the multiplication operator by a as the mapping $\mathcal{M}_a$ that associates q with qa

$$\begin{aligned} \mathcal{M}_a\colon \mathcal{A} &\to \mathcal{A} \\ q &\mapsto qa. \end{aligned} \tag{8}$$

The transposed operator $\mathcal{M}_a^{\mathrm{T}}$ is by definition the mapping from $\hat{\mathcal{A}}$ onto itself such that $(\mathcal{M}_a^{\mathrm{T}}\Lambda)(q) = \Lambda(\mathcal{M}_a q)$, $\forall\ \Lambda \in \hat{\mathcal{A}}$, $\forall\ q \in \mathcal{A}$ or, equivalently, $(\mathcal{M}_a^{\mathrm{T}}\Lambda)(q) = \Lambda(qa)$.

C. Lemmas

Let $\mathcal{P}$ be the subset of polynomials $\{f_1, \ldots, f_M\}$ of degree D and belonging to $\mathcal{R}$. Bézout's theorem [13, p. 227] states that such a system

$$\mathcal{P}\colon \{f_m(\boldsymbol{\xi}) = 0,\ 1 \le m \le M\} \tag{9}$$

where $\boldsymbol{\xi} \overset{\text{def}}{=} [\xi(0), \xi(1), \ldots \xi(M-1)]$, has either an infinity of solutions or a number of solutions smaller than or equal to D^M. This extends more well-known results for a single polynomial ($M = 1$) or for linear systems ($D = 1$). In what remains, we consider *generic systems* having exactly D^M solutions.

When the system has a finite number of solutions, the quotient $\mathcal{A}$ is of finite dimension (in fact, the variety of solutions is zero-dimensional, which implies that $\mathcal{A}$ is of finite dimension because Hilbert's polynomial [8] is of degree zero). Therefore, one conventional way to compute the solutions is to reduce the problem to an eigenvector computation, as shown by the following lemma.

Lemma III.3: Let a be any given polynomial in $\mathcal{R}$. Then, the eigenvalues of the multiplication map $\mathcal{M}_a$ in $\mathcal{A}$ are the values of a at the roots of the polynomial system $\mathcal{P}$.

Proof: Consider the polynomial $b(\boldsymbol{\xi}) = \prod_{\boldsymbol{z}\in\mathcal{P}}(a(\boldsymbol{\xi}) - a(\boldsymbol{z}))$. It vanishes at all the roots $\boldsymbol{z} \in \mathcal{P}$. Thus, by Hilbert's zero theorem (Nullstellensatz) [8], there exists a positive integer N, such that $b^N \in \mathcal{I}$ or, equivalently, such that $\prod_{\boldsymbol{z}\in\mathcal{P}}(a(\boldsymbol{\xi}) - a(\boldsymbol{z}))^N \equiv 0$ in $\mathcal{A}$. In terms of operators, this means that

$$\prod_{\boldsymbol{z}\in\mathcal{P}} (\mathcal{M}_a - a(\boldsymbol{z})I)^N \equiv 0$$

where I denotes the identity operator of $\mathcal{A}$. Thus, for any eigenpair $(\lambda, \boldsymbol{v})$ of $\mathcal{M}_a$, we have $\prod_{\boldsymbol{z}\in\mathcal{P}}(\mathcal{M}_a - a(\boldsymbol{z})I)^N \cdot \boldsymbol{v} = 0$ and thus

$$\prod_{\boldsymbol{z}\in\mathcal{P}} (\lambda - a(\boldsymbol{z}))^N \cdot \boldsymbol{v} = 0.$$

As $\boldsymbol{v} \neq 0$, λ must be one the values $a(\boldsymbol{z})$, $\boldsymbol{z} \in \mathcal{P}$, which completes the proof. ■

The reverse inclusion will be proved by Lemma III.4, among others.

Besides, if $a = \xi(0)$, the eigenvalues of matrix $\boldsymbol{M}_a$ of operator $\mathcal{M}_a$ give the values of coordinate $\xi(0)$ of the D^M solutions. If we repeat this operation for each tap $\xi(m)$, we have the D^M solutions. However, a somewhat simpler solution is introduced by the following lemma and avoids the computation of all eigenvalues of every multiplication operator $\mathcal{M}_{\xi(m)}$. In fact, *all* eigenvectors of a *single* operator $\mathcal{M}_a$ are actually required.

Lemma III.4: Linear forms $\mathbf{1}_{\boldsymbol{z}}$: $\mathbf{p} \mapsto \mathbf{p}(\boldsymbol{z})$, where $\boldsymbol{z}$ is any solution of $\mathcal{P}$, are the eigenvectors of all matrices $(\boldsymbol{M}_a^{\mathrm{T}})_{a\in\mathcal{A}}$ associated with the eigenvalues $a(\boldsymbol{z})$, $a \in \mathcal{A}$.

Proof: Using the definition of matrix $\boldsymbol{M}_a^{\mathrm{T}}$ and applying it to the linear form $\mathbf{1}_{\boldsymbol{z}}$, we get

$$\boldsymbol{M}_a^{\mathrm{T}}(\mathbf{1}_{\boldsymbol{z}})(q) = \mathbf{1}_{\boldsymbol{z}}(\mathbf{aq}) = \mathbf{a}(\boldsymbol{z})\mathbf{1}_{\boldsymbol{z}}(\mathbf{q}). \qquad \forall\, \mathbf{q} \in \mathcal{A}.$$

In other words, we have $\boldsymbol{M}_a^{\mathrm{T}}(\mathbf{1}_{\boldsymbol{z}}) = \mathbf{a}(\boldsymbol{z})\mathbf{1}_{\boldsymbol{z}}$. Therefore, $\mathbf{1}_{\boldsymbol{z}}$ and $a(\boldsymbol{z})$ are, respectively, the eigenvectors and eigenvalues of matrix $\boldsymbol{M}_a^{\mathrm{T}}$. This proves the lemma. ∎

If eigenvalues are not distinct, some eigenvectors are not uniquely determined, which makes the previous lemma less useful. Yet, it can be proved that the common eigenvectors of all operators $\mathcal{M}_a^{\mathrm{T}}$ are exactly the evaluation forms $\mathbf{1}_{\boldsymbol{z}}$ [19]. Therefore, a solution consists of taking several forms, say, $a(\xi)$ and $b(\xi)$, instead of a single one $a(\xi)$. $\boldsymbol{M}_a$ and $\boldsymbol{M}_b$ commute and have the same eigenspaces. The indeterminacy can be handled in this way, but it is not reported here in detail.

Hence, the computation of the multiplication matrix $\boldsymbol{M}_a$ appears as a key step in the proposed algorithm since the eigenvectors of $\boldsymbol{M}_a^{\mathrm{T}}$ allow the finding of all the solutions of $\mathcal{P}$. Indeed, if we take for a basis of $\mathcal{A}$ the set $\mathcal{B} = \{1, \xi(0), \xi(1), \ldots, \xi(0)\xi(1), \ldots\}$ of monomials of global degree at most D, the entries of the eigenvector $\mathbf{1}_z$ are equal to $\{1, z(0), z(1), \ldots, z(0)z(1), \ldots\}$ in the dual basis of $\mathcal{B}$, where $\boldsymbol{z}$ stands for any possible solution of $\mathcal{P}$. More precisely, entries 2 to $M+1$ of any eigenvector of $\boldsymbol{M}_a^{\mathrm{T}}$, whose first entry is normalized to 1, yield a solution $\boldsymbol{h}$ to $\mathcal{P}$. This property is used in the numerical algorithm of Section III-D by merely choosing $a = \xi(0)$. Any other polynomial could have made it.

D. Computing $\boldsymbol{M}_a$ Directly From the Polynomial System

As already mentioned, we prefer a more direct approach than that (more standard) described in Appendix A for computational reasons. In order to simplify the discussion, we will use the following example: Suppose that a channel of length $M = 3$ is excited by a MSK input. System $\mathcal{P}$ is then equal to that in (7). In that case, the computation of the multiplication matrix can be split into five steps.

First Step—Change in Variables: Suppose we use the following change in variables: $\boldsymbol{z} = \boldsymbol{T}^{-1}\boldsymbol{h}$. The system in $\boldsymbol{z}$ is implicitly defined by the system $\mathcal{P}$ in $\boldsymbol{h}$

$$\boldsymbol{P}\mathcal{H} = \mathbf{0} \tag{10}$$

where $\mathcal{H} = [1, h(0), h(1), h(2), h(0)h(1), h(0)h(2), h(1)h(2), h(0)^2, h(1)^2, h(2)^2]$. The matrix $\boldsymbol{T}$ is chosen so that a simple basis can always be found for $\mathcal{A}$ (that basis will contain monomials of degree at most 1 in every variable). Any choice of $\boldsymbol{T}$ (by drawing it randomly) would lead to a system that can be solved by $\mathcal{B}_3$ with probability one. Moreover, if $\boldsymbol{T}$ is not well chosen, the algorithm detects it because one of the matrices involved in the subsequent steps is rank deficient.

The next steps consist of expressing second-, third-, and fourth-degree monomials in the basis. We give, in Appendix C, the general procedure, but for more clarity, these steps are explained in detail for the particular case of system (7), which consists of three equations of degree 2 in three variables.

Second Step—Choosing a Basis: A basis that can generically solve our kind of polynomial systems is composed of the neutral element 1, the M unknowns $z(m)$, and all the cross monomials of degree less than or equal to M, where each unknown appears with a power equal to 1 or 0. In fact, it has been proved by Macaulay that such a basis is sufficient when there are no zeros at infinity (generic case) [27]. Here are two examples: $M = 2$ and $M = 3$.

- Channel length $M = 2$:

$$\mathcal{B}_2 = \{1, z(0), z(1), z(0)z(1)\}$$

- Channel length $M = 3$:

$$\mathcal{B}_3 = \{1, z(0), z(1), z(2), z(0)z(1), z(0)z(2) \\ z(1)z(2), z(0)z(1)z(2)\}.$$

In the following, the column vector containing the elements of the basis $\mathcal{B}_i$ will be denoted as $\boldsymbol{b}_i(\boldsymbol{z})$. In this example, $\boldsymbol{b}_3(\boldsymbol{z})$ is of size 8. System (10) can then be rewritten as $\boldsymbol{A}\boldsymbol{b}_3(\boldsymbol{z}) = \mathbf{0}$ for some matrix $\boldsymbol{A}$ depending on $\boldsymbol{P}$ and $\boldsymbol{T}$ only.

This basis can always be used in our problem [27] because the system is a complete intersection (there are a finite number of solutions), and we first applied a generic change in the variables (there are no longer zeros at infinity). The reason is that in (7), some equations linked monomials $h(i)h(j)$ and 1, which would have not been linearly independent in $\mathcal{A}$.

Third Step—Expression of the Second-Degree Monomials: Suppose we want to find in $\mathcal{B}_3$ the matrix associated with multiplication by $z(0)$. The monomials to be expressed are as follows: If $M = 3$

Monomials of the basis		Monomials to be expressed
1		$z(0)$
$z(0)$		$z(0)^2$
$z(1)$		$z(0)z(1)$
$z(2)$		$z(0)z(2)$
$z(0)z(1)$	$\overset{\times z(0)}{\longmapsto}$	$z(0)^2z(1)$
$z(0)z(2)$		$z(0)^2z(2)$
$z(1)z(2)$		$z(0)z(1)z(2)$
$z(0)z(1)z(2)$		$z(0)^2z(1)z(2)$.

Some of these monomials are already in the basis, such as $z(0)$, $z(0)z(1)$, $z(0)z(2)$, and $z(0)z(1)z(2)$. The other monomials $z(0)^2$, $z(0)^2z(1)$, $z(0)^2z(2)$, and $z(0)^2z(1)z(2)$ have to be expressed using the polynomial system.

According to (10), monomials $z(0)^2$, $z(1)^2$, and $z(2)^2$ can be expressed directly as a function of 1, $z(0)$, $z(1)$, $z(2)$, $z(0)z(1)$, $z(0)z(2)$, and $z(1)z(2)$, provided that $\boldsymbol{T}$ is chosen correctly.

In other words, monomials $z(0)^2$, $z(1)^2$, and $z(2)^2$ can be expressed directly as a function of the basis using (10)

$$\begin{bmatrix} z(0)^2 \\ z(1)^2 \\ z(2)^2 \end{bmatrix} = \boldsymbol{B} \begin{bmatrix} 1 \\ z(0) \\ z(1) \\ z(2) \\ z(0)z(1) \\ z(0)z(2) \\ z(1)z(2) \end{bmatrix} \tag{11}$$

for some matrix $\boldsymbol{B}$ depending on $\boldsymbol{A}$ only. Therefore, the monomial $z(0)^2$ is now expressed in the basis $\mathcal{B}_3$; in fact, $z(0)z(1)z(2)$ has here a null coefficient.

Fourth Step—Expression of the Third-Degree Monomials: Monomials $z(0)^2z(1)$ and $z(0)^2z(2)$ are now of interest. These monomials can be written using the expression of the monomial $z(0)^2$ in the first row of (11). If we multiply this equation by $z(1)$, monomial $z(0)^2z(1)$ appears in the left-hand side, and monomials $z(1)$, $z(0)z(1)$, $z(1)^2$, $z(2)z(1)$, $z(0)z(1)^2$, $z(0)z(1)z(2)$, and $z(1)^2z(2)$ appear on the right-hand side. Among these monomials, one can distinguish those that are in the basis, like $z(1)$, $z(0)z(1)$, $z(2)z(1)$, and $z(0)z(1)z(2)$, those that have already been expressed in the basis, like $z(1)^2$, and those that are unknown, like $z(1)^2z(2)$ and $z(0)z(1)^2$. However, these unknown monomials are of the same type as monomials $z(0)^2z(1)$ and $z(0)^2z(2)$, and one can show that expressing monomials $z(0)^2z(1)$, $z(0)^2z(2)$, $z(1)^2z(0)$, $z(1)^2z(2)$, $z(2)^2z(0)$, and $z(2)^2z(1)$ all together using (11) leads to

$$\begin{bmatrix} z(0)^2z(1) \\ z(0)^2z(2) \\ z(1)^2z(0) \\ z(1)^2z(2) \\ z(2)^2z(0) \\ z(2)^2z(1) \end{bmatrix} = \boldsymbol{C} \begin{bmatrix} 1 \\ z(0) \\ z(1) \\ z(2) \\ z(0)z(1) \\ z(0)z(2) \\ z(1)z(2) \\ z(0)z(1)z(2) \end{bmatrix} \tag{12}$$

for some matrix $\boldsymbol{C}$ depending on $\boldsymbol{B}$ only.

Fifth Step—Expression of the Fourth-Degree Monomials: Using the same method as before, one can express monomials such as $z(0)^2z(1)z(2)$ using (12). See Appendix C for more details.

Having expressed all the monomials in the chosen basis, the multiplication matrix can be constructed. Once the multiplication matrix $\boldsymbol{M}_a$ is found for the arbitrarily chosen $a = z(0)$, one computes the 2^M eigenvectors of $\boldsymbol{M}_a^{\mathrm{T}}$, $\boldsymbol{v}^{(m)}$. Next, all the possible solutions $\boldsymbol{h}^{(m)}$ to the polynomial system are obtained as $\boldsymbol{z}^{(m)} = \boldsymbol{v}^{(m)}(2\colon M+1)/\boldsymbol{v}^{(m)}(1)$. Then, the solutions in the original coordinate system are given by $\boldsymbol{h}^{(m)} = \boldsymbol{T}\boldsymbol{z}^{(m)}$. See Appendix C for details.

For the sake of clarity, we have described the elimination algorithm for a channel length of $M = 3$, but the principles hold exactly the same for larger values of M (with a larger number of unknowns), such as $M = 5$, as in some subsequent computer simulations. As explicited in [27], the general algorithm goes along the lines described in Table I.

TABLE I
STEPS OF THE ELIMINATION PROCEDURE

1. Make a change of variables $\boldsymbol{z} = \boldsymbol{T}^{-1}\boldsymbol{h}$
2. Choose the basis $\mathcal{B}_M$ containing all 2^M monomials of global degree at most M, and partial degree at most 1, in variables $z_0, z_1, \ldots, z_{M-1}$
3. For $d = 2$ to M, express monomials of degree d, that are at Hamming distance 1 of basis $\mathcal{B}_M$, as a function of monomials of $\mathcal{B}_M$

IV. Estimation of the Channel

Selection of a Solution: In a final step, one chooses the solution $\boldsymbol{h}$ among $\{\boldsymbol{h}^{(m)}, 1 \le m \le 2^M\}$ that best matches the actual channel by a moment matching method, as we explain now.

A polynomial system rarely admits a unique solution, regardless of the number of unknowns. Therefore, it is very likely that we obtain in practice 2^M distinct solutions that we can denote as $\boldsymbol{h}^{(m)} = [h^{(m)}(0), \ldots h^{(m)}(M-1)]$, $1 \le m \le 2^M$. However, circular statistics (3) have yet to be utilized. With this goal, denote

$$R_y(\ell) \overset{\mathrm{def}}{=} \mathrm{E}\{y(n)y(n-\ell)^* | x(0)\}. \tag{13}$$

Then, because of (3), we have the well-known phase-blind relation

$$R_y(\ell) = \sum_{i=\ell}^{M-1} h(i)h(i-\ell)^* + R_w(\ell), \qquad 0 \le \ell \le M-1. \tag{14}$$

The procedure proposed in this paper consists of choosing the solution $\boldsymbol{h}$ minimizing the distance:

$$\boldsymbol{h} = \operatorname*{Arg\,Min}_{\boldsymbol{h}^{(m)}} \sum_{\ell=0}^{M-1} \left| R_y(\ell) - \sum_{i=\ell}^{M-1} h^{(m)}(i)h^{(m)}(i-\ell)^* \right|^2 \tag{15}$$

hence, the name of *moment matching* method. The whole algorithm is summarized in Table II.

Computational Complexity: It is worth noting that the largest part of the computational load consists of building the multiplication matrix $\boldsymbol{M}_a$, which depends on the modulation (discrete alphabet and trellis). Yet, it can be shown [9] that this matrix $\boldsymbol{M}_a$ is itself a polynomial function of the data moments when the polynomial system $\mathcal{P}$ has no infinite solution. Thus, in an operational context, for a given modulation, one can only store the polynomial coefficients of $\boldsymbol{M}_a$ in a ROM; as a consequence, the computation of $\boldsymbol{M}_a$ becomes negligible, and the overall data-dependent computations are dominated by the calculation of the eigenvectors of $\boldsymbol{M}_a^{\mathrm{T}}$.

TABLE II
SUMMARY OF THE ALGORITHM

- Choose a presumed channel length, $\widehat{M} = L$
- Compute the sample correlation

$$R_y(\ell) = \frac{1}{N-\ell} \sum_{n=\ell+1}^{N} y(n)y(n-\ell)^*,\ 0 \le \ell \le L-1$$

- Compute the non-circular sample correlation

$$\alpha(\ell) = \frac{(-1)^\ell}{N-\ell} \sum_{n=\ell+1}^{N} (-1)^n y(n)y(n-\ell),\ 0 \le \ell \le L-1$$

- Choose a transform matrix, $\boldsymbol{T}$.

From $\alpha(\ell)$, calculate the multiplication matrix, $\boldsymbol{M}^{\mathrm{T}}_{z(0)}$ according to the procedure described in sections III-D and VIII-C, as well as table I.

- Compute all its 2^L eigenvectors, $\boldsymbol{v}^{(m)}$
- Compute the 2^L possible solutions, $\boldsymbol{z}^{(m)}$, by normalizing the first entry of $\boldsymbol{v}^{(m)}$ to 1, and retaining entries 2 to $L+1$
- Transform back every candidate into the original coordinate system: $\boldsymbol{h}^{(m)} = \boldsymbol{T}\boldsymbol{z}^{(m)}$
- For each of these candidates, compute the corresponding theoretical correlation (14)
- Select the solution $\boldsymbol{h}$ minimizing (15)
- Equalize the channel and compute the BER of the estimated input.

V. IDENTIFIABILITY

It is well known that blind identifiability can be carried out only up to a complex multiplicative factor. Therefore, there are infinitely many solutions. However, if we arbitrarily fix the source variance to 1 and the scalar phase inherent indeterminacy, which contains a fixed phase and a delay z^γ, then identifiability results can be stated. Thus, we may assume from now on that the channel is causal of degree $M-1$, and we want to prove that the joint use of circular and noncircular second-order output correlation functions yields an *essentially* unique solution (i.e., up to a unit modulus multiplicative factor and up to a time delay).

Lemma V.1: Suppose we look for a causal FIR channel H of length M from given second-order circular statistics; then, the number of solutions is *essentially* finite and bounded by 2^{M-1}.

Proof: The z-transform of the circular covariance $c(n)$ of the output $y(n)$ is equal to $C(z) = H(z)H^*(1/z^*)$ since the input is white and of unit variance. This shows that if $H(z)$ is causal, it can be determined from $C(z)$ up to two kinds of indeterminacies. First, $H(z)$ can only be determined up to a multiplicative constant phase factor. Second, if $H(z)$ is transformed into $H(z)\Phi(z)$, where $\Phi(z)$ verifies $\Phi(z)\Phi^*(1/z^*) = 1$, $C(z)$ remains the same. It is well known that $\Phi(z)$ is then an allpass filter, i.e., of the following form, up to a delay

$$\Phi(z) = \prod_{i=1}^{Q} \frac{1 - a_i^* z^{-1}}{a_i - z^{-1}}.$$

Since $H(z)$ must be FIR and since $\Phi(z)$ is not FIR, $H(z)\Phi(z)$ is FIR only if each pole of $\Phi(z)$ is associated with one of the $M-1$ roots of $H(z)$. As a consequence, there is a finite number of allpass filters such that $H(z)\Phi(z)$ is FIR. Therefore, if the phase indeterminacy is fixed, there are at most 2^{M-1} possible FIR filters that correspond to $C(z)$. If the roots of $H(z)$ are not all distinct or of unit modulus, there are indeed fewer possibilities than 2^{M-1}. This proves Lemma V.1, which has been known for many years. ■

Theorem V.2: Suppose we look for an FIR channel H of length M from given second-order circular and noncircular statistics of the output when the input is stationary white. Then, up to complex phase and time delay indeterminacies, we have the following:

- a unique solution if H has no real root,
- 2^Q solutions if H has Q distinct real roots.

Proof: Suppose now that we also use the noncircular covariance $\overline{c}(\ell) \overset{\mathrm{def}}{=} \mathrm{E}\{y(n)y(n-\ell)|x(0)\}$. Its z-transform is equal to $\overline{C}(z) = x(0)^2 H(z)H(1/z)$. Yet, one can easily show that rational filters satisfying $\Psi(z)\Psi(1/z) = 1$ are of the following form, up to a delay:

$$\Psi(z) = \prod_j \frac{1 - b_j z^{-1}}{b_j - z^{-1}}.$$

Now, using statements made in the proof of Lemma V.1, allpass rational filters $\Phi(z)$ that also satisfy $\Phi(z)\Phi(1/z) = 1$ must have real poles (and zeros). As a consequence, if $H(z)$ has no real roots, $\Phi(z)$ cannot have real poles since $H(z)\Phi(z)$ is FIR, and the allpass filter $\Phi(z)$ must be equal to ±1. In this case, there is an *essentially* unique solution for $H(z)$, up to a sign. The fact that $x(0)$ is known or not is of no importance since it is of unit modulus and can be pulled into the inherent indeterminacies. If $H(z)$ has Q real roots, one must use the result of Lemma V.1. The number of solutions is then *essentially* equal to 2^Q. ■

When the input source is MSK, it is white but not stationary. However, identifiability still holds.

Corollary V.3: When the input source is white MSK, the joint use of second-order circular and noncircular statistics of the output yield a unique solution for $H(z)$, up to a sign.

The proof is given in Appendix B. These identifiability results justify the algorithm we have proposed and are summarized in Section IV.

VI. COMPUTER RESULTS

Tests are run on a random FIR channel ($M = 5$). At each run, the channel is a realization of a Clarke filter in the typical urban (TU) mode. Every channel generated is specular and contains six paths, whose delays and attenuations are given in the following table, according to the ETSI norm (GSM 05.05, December 1995, six taps TU setting):

τ_i (μs)	0	0.2	0.5	1.6	2.3	5.0
$\mathrm{E}\{a_i^2\}$ (dB)	-3	0	-2	-6	-8	-10

In other words, each coefficient is a random complex circular Gaussian variable whose standard deviation is given in the

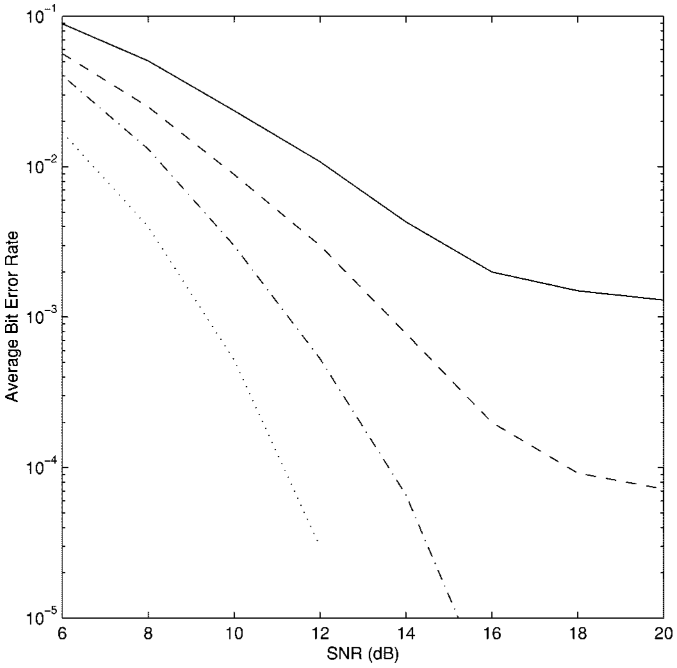

Fig. 1. Average BER at the output of a Viterbi algorithm using our channel estimate as a function of the SNR for various block lengths.

table. Note that the first path is not the strongest. The symbol rate is 271 ksymbol/s, which corresponds to a symbol period of 3.68 μs, whereas the delays above are also given in μs (these cases are those of GSM). The transmit filter is a raised cosine with rolloff $\beta = 0.1$, and four samples have been generated per symbol period. The channel is excited by a MSK input. The performance is presented as a function of the SNR and of the length N of the observation block and averaged over 500 independent channel trials.

Fig. 1 shows the average bit error rate (BER) obtained at the output of a Viterbi equalizer that uses our channel estimate. The solid, dashed, and dash-dotted lines correspond to block lengths $N = 200$, $N = 500$, and $N = 1000$, respectively. These performances are compared with the average BER obtained with the actual channel (dotted line).

For high SNRs and $N = 200$ or $N = 500$, the results show the effects of moment estimation errors. These effects disappear for $N = 1000$, where the performances exhibit a loss of 2 dB compared with the actual channel results.

Fig. 2 shows the average mean squares distance between the actual and the estimated channel. The solid, dashed, and dash-dotted lines correspond again to block lengths $N = 200$, $N = 500$, and $N = 1000$, respectively.

It appeared relevant to test the robustness of the algorithm with respect to channel overestimation. Keeping $M = 5$ in the algorithm, tests have been run on a random FIR channel of actual length $M = 3$. Fig. 3 shows the average BER obtained at the output of a Viterbi equalizer that uses our channel estimate when $\hat{M} = 3$ (solid line) and $\hat{M} = 5$ (dashed line). Fig. 4 shows the average mean square estimation error of the channel as a function of the SNR and in the presence of overdetermination. These two performances are compared with the one obtained when the actual channel is used (dash-dotted line). Hence, this test illustrates the loss in performance encountered in the presence of channel order overdetermination.

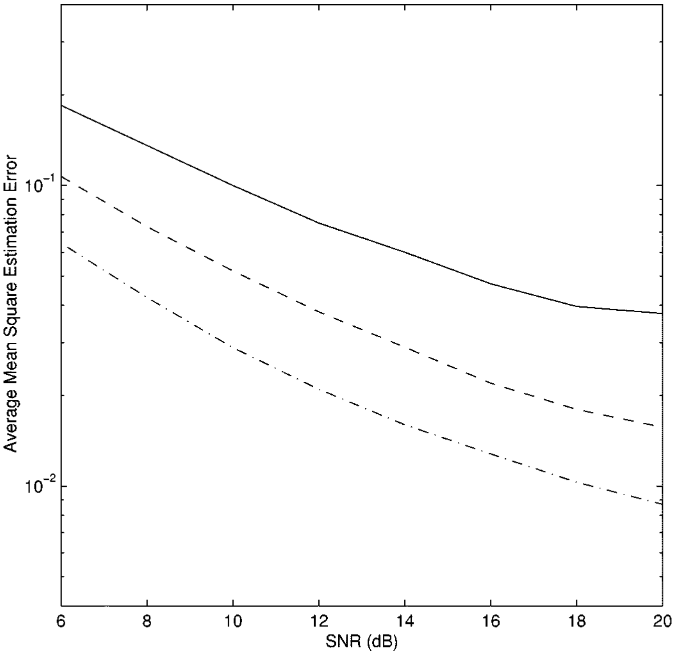

Fig. 2. Average mean square estimation error of the channel as a function of the SNR for BER at the output of the SNR for various block lengths.

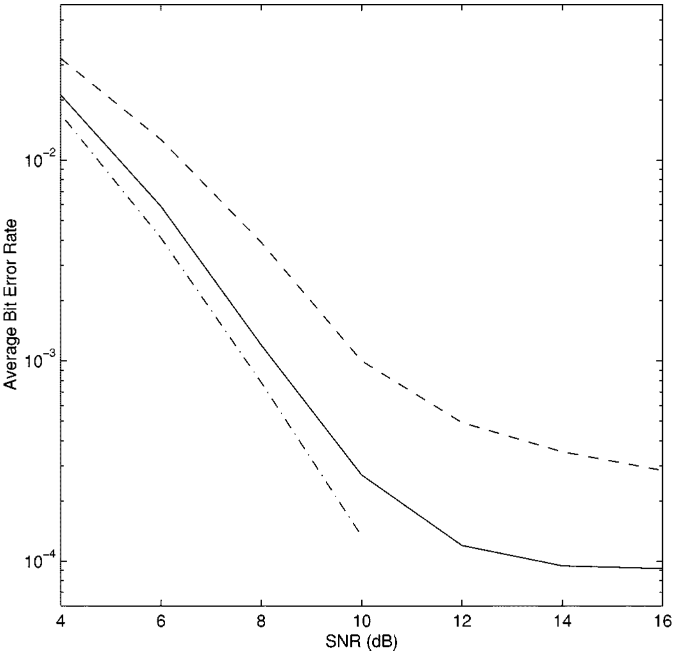

Fig. 3. Average BER at the output of a Viterbi algorithm using our channel estimate as a function of the SNR and in presence of overdetermination.

Experiments for QPSK modulated inputs have not been run, but one can either treat real and imaginary parts separately, as explained in Section II, or use fourth-order moments. Of course, the former is preferred (e.g., for IS95 standard), but the latter seems unavoidable in the case of π/4-QPSK modulations, which are found, for example, in the IS54 standard.

VII. Concluding Remarks

The blind identification method described in the paper is based only on second-order statistics of the observation. We

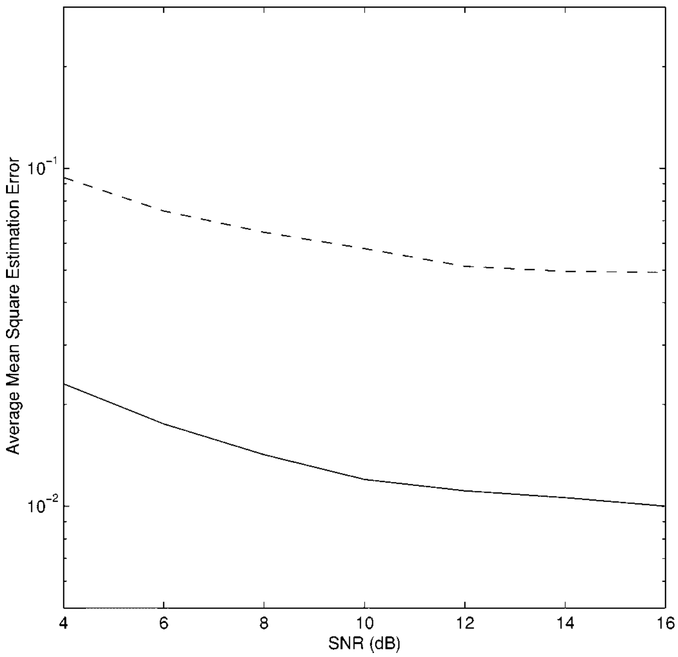

Fig. 4. Average mean square estimation error of the channel as a function of the SNR and in presence of overdetermination.

proved that it is not necessary to resort to higher order statistics, explicitly or implicitly (like in CMA), to identify a SISO channel; next, we gave an analytical solution that was free of local extrema problems. In addition, identifiability results were derived when an FIR channel is searched for, from both circular and noncircular moments. Finally, computer results showed that the behavior of our algorithm depends on the estimation quality of the moments and on the accuracy of the determination of the channel length.

The extension to SIMO channels is rather straightforward and consists essentially of code writing and computer experiments. In fact, there is no fundamental change in the algorithm. However, there exist specific algorithms that work in the SIMO case and not in the SISO case. For this reason, we believe our algorithm is more attractive in the SISO case.

APPENDIX

A. Computation of Matrix $\boldsymbol{M}_a$ by Resultants

Many methods can be used to compute the multiplication matrix. Among them, the Gröbner bases are the most well known. They can be used to build a base of the quotient algebra $\mathcal{A}$ and then to compute the multiplication matrix with normal forms computation. However, this method must be run in exact arithmetic and thus implies the use of big numbers. Another use of the Gröbner bases consists of the elimination of $M-1$ variables and leads to the computation of the roots of a mono-variate polynomial of degree D^M. Since these methods are too expensive in terms of computation load, we prefer to use a modification of the old method by Macaulay [17], which has been used to build resultants.

This approach, which is discussed in this section for the sake of completeness, can be considered to be an extension of the Sylvester theorem to multivariate polynomials and, thus, could seem natural. However, another simpler (but less standard) approach is discussed in Section III D and Appendix C and has been used throughout the paper. We now present Sylvester and Macaulay matrices and their properties.

1) Sylvester Matrices: Sylvester matrices are matrices S that represent maps of the form

$$\mathcal{S}\colon V_0 \times V_1 \to V$$
$$(q_0, q_1) \mapsto q_0 f_0 + q_1 f_1$$

where V_0, V_1, and V are the spaces generated by monomials $\{1, \ldots, x^{d_1-1}\}$, $\{1, \ldots, x^{d_0-1}\}$, and $\{1, \ldots, x^{d_0+d_1-1}\}$, respectively. Matrix S, of size $d_0 + d_1 \times d_0 + d_1$, thus has the following structure:

$$S = \begin{bmatrix} \boldsymbol{f}_0 & & & \boldsymbol{f}_1 & & \\ & \ddots & & & \ddots & \\ & & \boldsymbol{f}_0 & & & \boldsymbol{f}_1 \end{bmatrix}$$

where the entries of $\boldsymbol{f}_i$ are the coefficients of the polynomials f_i so that $f_0 = [1, \ldots, x^{d_0}]\boldsymbol{f}_0$ and $f_1 = [1, \ldots, x^{d_1}]\boldsymbol{f}_1$. The determinant of S is also called the resultant of polynomials f_0 and f_1. Of course, the construction procedure also applies for more than two polynomials.

2) Macaulay Matrices: The approach described above can be generalized to multivariate polynomials. Let $f_0, f_1, \ldots, f_M \in \mathcal{R}$ be $M + 1$ polynomials of degree $d_0, \ldots, d_M$ in variables $\boldsymbol{\xi} = \{\xi(1), \cdots, \xi(M)\}$ (Note that in the algorithm of Section III-D, all degrees d_i are equal to D, except $d_0 = 1$.) Macaulay matrices [17] used to build the resultant of these polynomials are matrices associated with maps of the form

$$\mathcal{S}\colon V_0 \times \cdots \times V_M \to V$$
$$(q_0, \ldots, q_M) \mapsto \sum_{i=0}^{M} q_i f_i$$

where the V_m are subspaces spanned by a finite number of monomials in variables $\{\xi(1), \ldots, \xi(M)\}$, which are denoted as $V_m = \langle \boldsymbol{\xi}^{E_m} \rangle$, where $\boldsymbol{\xi}^{E_m} = \{\xi^\beta, \beta \in E_m\}$ and $\xi^\beta = \xi_1^{\beta_1} \cdots \xi_M^{\beta_M}$. In this notation, E_m is the set of respective powers of the variables of the monomials that span the subspace V_m, $E_m = \{\beta_{m,1}, \ldots, \beta_{m,N_m}\}$, each $\beta_{m,n}$ containing M integers. In a similar manner, define F such that $V = \langle \boldsymbol{\xi}^F \rangle$.

The matrix S associated with this multivariate map $\mathcal{S}$ is such that its columns are the image, in the canonical basis, of the monomials $\xi^{\beta_{m,i}}$ in the polynomials f_m. If d_m denote the degree of polynomial f_m, it can be shown that $\mathcal{S}$ is of dimension $\binom{\nu+M}{M}$, where $\nu = \sum_i d_i - M$. In other words, its size is that of the basis $\mathcal{B}$.

Matrix S can be split in $M + 1$ blocks: $S = [S_0, S_1, \ldots, S_M]$, where block S_m is the canonical representation of the image subspace of the monomials $\boldsymbol{\xi}^{E_m}$ by the polynomial f_m. Let us now explain the construction procedure on an example.

Example: Suppose that $M = 2$, $D = 2$, $f_0(x, y) = a_0 x$, $f_1(x, y) = a_1 x^2 + b_1 y$, and $f_2(x, y) = a_2 y + b_2$.

The critical degree is $\nu = 1 + 2 + 1 - 2 = 2$ in this example, and we arbitrarily choose the basis $\xi^F = \{1, x, y, x^2, xy, y^2\}$.

At each step $M-m+1$ of the procedure $M \geq m \geq 1$, one looks for monomials left in ξ^F that can be divided by $\xi_m^{d_m}$; in order to construct a basis of space V_m, one divides these monomials by $\xi_m^{d_m}$. In *Step 1*, monomials of ξ^F divisible by $\xi_2^{d_2} = y$ are $\{y, xy, y^2\}$. Ones sets $V_2 = \{1, x, y\}$ removes $\{y, xy, y^2\}$ from ξ^F, and go to the next step. In *Step 2*, monomials of $\{1, x, x^2\}$ divisible by $\xi_1^{d_1} = x^2$ reduce to $\{x^2\}$. Thus, one sets $V_1 = \{1\}$ and removes $\{x^2\}$ from what remained in ξ^F. It eventually remains that $\{1, x\}$ in ξ^F, which consists of the basis of V_0.

Every row of S is associated with a monomial of ξ^F, in a preassigned order; here, we choose arbitrarily $\xi^F = \{1, x, y, x^2, xy, y^2\}$. The first two columns span the space $f_0(V_0)$, the third one spans $f_1(V_1)$, and the last three columns span $f_2(V_2)$

$$S = \left|\begin{array}{cc|c|ccc} 0 & 0 & 0 & b_2 & 0 & 0 \\ a_0 & 0 & 0 & 0 & b_2 & 0 \\ 0 & 0 & b_1 & a_2 & 0 & b_2 \\ 0 & a_0 & a_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_2 \end{array}\right|.$$

The submatrices comprised within vertical bars can be denoted as S_0, S_1, and S_2, respectively.

The sets of exponents E_m associated with V_m are the following: $E_0 = \{[0, 0], [1, 0]\}$, $E_1 = \{[0, 0]\}$, $E_2 = \{[0, 0], [1, 0], [0, 1]\}$, and the set F associated with the whole space V is $F = \{[0, 0], [1, 0], [0, 1], [2, 0], [1, 1], [0, 2]\}$ and actually describes $\mathcal{B}$.

3) Properties of Macaulay's Matrices: The proof of the following theorems are not given but can be found in [10], [17], and [20].

Theorem VIII.1: Consider the generic polynomial system $f_1, \ldots, f_M$, choose a polynomial f_0, and build the Macaulay matrix S associated with these polynomials. The set of monomials $\boldsymbol{\xi}^{E_0}$ is a basis of the quotient algebra $\mathcal{A} = \mathcal{R}/(f_1, \ldots, f_M)$ and where E_0 is defined in Appendix A2.

Hence, the construction of the Macaulay matrix yields a basis of the quotient algebra. The following theorem gives a means to compute the matrix associated with the multiplication by f_0 thanks to the Macaulay matrices.

Theorem VIII.2: Suppose E_0 is a subset of F, which is the set of exponents indexing the rows of S; this is the case if f_0 includes a constant term. Then, matrix S rewrites

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where the rows and the columns of A are indexed by the monomials $\boldsymbol{\xi}^{E_0}$, and the columns of B and D are indexed by the monomials $\{\boldsymbol{\xi}^{E_m}\}$, where m is different from zero. Then, for all polynomial systems $f_1, \ldots, f_M$, the matrix of the multiplication by f_0 in the quotient algebra $\mathcal{A} = \mathcal{R}/(f_1, \ldots, f_M)$ is the Schur complement of the block D in the matrix S

$$\boldsymbol{M}_{f_0} = A - BD^{-1}C.$$

Thus, matrix $\boldsymbol{M}_a$ can be directly computed from the Macaulay matrix associated with polynomials $\{f_1, \ldots, f_M\}$ [7]. However, if we take into account the relationships between the monomials introduced in the polynomial system $\mathcal{P}$, there exists a much simpler procedure to compute $\boldsymbol{M}_a$, as explained in Section III-D and Appendix C.

B. Proof of Corollary V.3

Proof: When the input source is MSK, let us first show that the noncircular covariance $\overline{c}(\ell; n) \overset{\text{def}}{=} \mathrm{E}\{y(n)y(n-\ell)|x(0)\}$ has a z-transform given by

$$\overline{C}(z; n) = (-1)^n x(0)^2 H(-z)H(1/z). \tag{16}$$

In fact, $\overline{c}(\ell; n) = \sum_p \sum_q h(p)h(q)\mathrm{E}\{x(n-p)x(n-q-\ell)|x(0)\}$, which, from (2), yields

$$\overline{c}(\ell;\, n) = (-1)^n x(0)^2 \sum_p \sum_q (-1)^p h(p)h(q)\delta(p-\ell-q)$$

or $\overline{c}(\ell; n) = (-1)^n x(0)^2 \sum_p (-1)^p h(p)h(p-\ell))$. Now, take the z-transform over time index ℓ, and get $\overline{C}(z; n) = (-1)^n x(0)^2 \sum_\ell \sum_p (-1)^p h(p)h(p-\ell)z(-\ell)$. Making the change of variables $k = p - \ell$ eventually gives (16).

Now, suppose we have one solution $H_o(z)$ satisfying (16). From the result above, we see that the whole set of rational filters satisfying (16) is then generated by $H_o(z)\Upsilon(z)$, where $\Upsilon(-z)\Upsilon(1/z) = 1$, and where poles of $\Upsilon(z)$ coincide with zeros of $H_o(z)$.

Yet, it can be easily shown that such filters take the following form, up to an even delay:

$$\Upsilon(z) = \prod_k \jmath \frac{1 + b_k z^{-1}}{b_k - z^{-1}}.$$

Now, it can be seen that the only allpass filter satisfying $\Phi(z)\Phi^*(1/z^*)$ that also satisfies $\Phi(-z)\Phi(1/z) = 1$ is $\Phi(z) = \pm 1$. Therefore, the whole set of solution reduces to $\pm H_o(z)$, up to a delay. ■

C. Computational Details

In this section, we detail the steps described in Section III-D in the particular case of system (7), which contains three equations of degree 2 in three variables. Computer codes can be downloaded from the web page of the second author. From the Bézout lemma quoted in Section III-C, we already know that this system will have generally $2^3 = 8$ solutions.

1) First Step: The identity matrix is not suitable for $\boldsymbol{T}$ because some steps in the procedure will involve singular matrices. This comes from the fact that pure squares all appear in the first equation only. Thus, the following choice is made:

$$\boldsymbol{h} = \boldsymbol{T}\boldsymbol{z}, \qquad \boldsymbol{T} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}. \tag{17}$$

2) Second Step: One chooses the basis $\mathcal{B}_3$, which contains the eight monomials in the following, that we can arrange in a vector:

$$\boldsymbol{b}_3 = [1,\, z(0),\, z(1),\, z(2),\, z(0)z(1) \\ z(0)z(2),\, z(1)z(2),\, z(0)z(1)z(2)]^{\mathrm{T}}.$$

The system (7) can now be expressed in terms of the new variables $z(i)$ as

$$\boldsymbol{A}\begin{bmatrix} 1 \\ z(0)z(1) \\ z(0)z(2) \\ z(1)z(2) \\ z(0)^2 \\ z(1)^2 \\ z(2)^2 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \qquad (18)$$

where

$$\boldsymbol{A} = \begin{bmatrix} \alpha_0 & -4 & 2 & 2 & -3 & -2 & 0 \\ \alpha_1 & -2 & -1 & -2 & -2 & 0 & 1 \\ \alpha_2 & 2 & -2 & -1 & 0 & 1 & 0 \end{bmatrix}.$$

The goal is now to express the matrix $\boldsymbol{M}_{z(0)}$ of operator $\mathcal{M}_{z(0)}$ in the basis $\mathcal{B}_3$. For doing this, we need to have the expression of all monomials of $z(0)\boldsymbol{b}_3$, in particular, monomials such as $z(0)^2$, $z(0)^2z(1)$, or $z(0)^2z(1)z(2)$, as a function of monomials of $\boldsymbol{b}_3$ itself.

3) Third Step: The pure squares $z(i)^2$ are not in $\mathcal{B}_3$, but we can obtain them from (18) by isolating the squares in the left-hand side

$$\begin{bmatrix} z(0)^2 \\ z(1)^2 \\ z(2)^2 \end{bmatrix} = \boldsymbol{A}_1^{-1}\boldsymbol{B}_1\boldsymbol{b}_3(\boldsymbol{z}) \overset{\mathrm{def}}{=} \boldsymbol{C}_1\boldsymbol{b}_3(\boldsymbol{z}) \qquad (19)$$

where matrices $\boldsymbol{A}_1$ and $\boldsymbol{B}_1$ are functions of $\boldsymbol{A}$:

$$\begin{aligned} \boldsymbol{A}_1 &= [A(:,5) \quad A(:,6) \quad A(:,7)] \\ \boldsymbol{B}_1 &= -[A(:,1) \quad \mathbf{0} \quad \mathbf{0} \quad \mathbf{0} \quad A(:,2) \quad A(:,3) \quad A(:,4) \quad \mathbf{0}] \end{aligned}$$

and where $A(:,i)$ denotes the ith column of $\boldsymbol{A}$ (as in the Matlab notation).

4) Fourth Step: Let us turn now to monomials of the form $z(i)^2z(j)$, $i \neq j$. Take the first row of (19), and multiply it by $z(1)$. A careful inspection then shows that $z(0)^2z(1)$ depends only on $\{z(1), z(0)z(1)^2, z(0)z(1)z(2), z(1)^2z(2)\}$ because of the presence of zeros in matrix $\boldsymbol{C}_1$. Similarly, $z(0)^2z(2)$ depends only on $\{z(2), z(0)z(1)z(2), z(0)z(2)^2, z(1)z(2)^2\}$. In order to solve this problem, it suffices to write jointly the other equations, which are obtained on one hand by multiplying the second row of (19) by $z(0)$ and $z(2)$ and on the other hand by multiplying the third row of (19) by $z(0)$ and $z(1)$. This eventually yields the following system, after some manipulations:

$$\begin{bmatrix} z(0)^2z(1) \\ z(0)^2z(2) \\ z(1)^2z(0) \\ z(1)^2z(2) \\ z(2)^2z(0) \\ z(2)^2z(1) \end{bmatrix} = -\boldsymbol{A}_2^{-1}\boldsymbol{B}_2\boldsymbol{b}_3(\boldsymbol{z}) \overset{\mathrm{def}}{=} \boldsymbol{C}_2\boldsymbol{b}_3(\boldsymbol{z}) \qquad (20)$$

where

$$\boldsymbol{A}_2 = \begin{bmatrix} -1 & 0 & C_1(1,5) & C_1(1,7) & 0 & 0 \\ 0 & -1 & 0 & 0 & C_1(1,6) & C_1(1,7) \\ C_1(2,5) & C_1(2,6) & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & C_1(2,6) & C_1(2,7) \\ C_1(3,5) & C_1(3,6) & 0 & 0 & -1 & 0 \\ 0 & 0 & C_1(3,5) & C_1(3,7) & 0 & -1 \end{bmatrix}$$

$$\boldsymbol{B}_2 = \begin{bmatrix} 0 & 0 & C_1(1,1) & 0 & 0 & 0 & 0 & C_1(1,6) \\ 0 & 0 & 0 & C_1(1,1) & 0 & 0 & 0 & C_1(1,5) \\ 0 & C_1(2,1) & 0 & 0 & 0 & 0 & 0 & C_1(2,7) \\ 0 & 0 & 0 & C_1(2,1) & 0 & 0 & 0 & C_1(2,5) \\ 0 & C_1(3,1) & 0 & 0 & 0 & 0 & 0 & C_1(3,7) \\ 0 & 0 & C_1(3,1) & 0 & 0 & 0 & 0 & C_1(3,6) \end{bmatrix}.$$

Consequently, we have indeed expressed monomials $z(0)^2z(1)$ and $z(0)^2z(2)$ in the basis $\mathcal{B}_3$.

5) Fifth Step: Last, let us turn to monomial $z(0)^2z(1)z(2)$. From (20), it can be seen that

$$z(0)^2z(1) = [C_2(1, 2:4), C_2(1,8)]\begin{bmatrix} z(0) \\ z(1) \\ z(2) \\ z(0)z(1)z(2) \end{bmatrix}.$$

Multiplying this relation by $z(2)$ shows that $z(0)^2z(1)z(2)$ can be expressed as a function of monomials $z(2)^2$, $z(0)z(1)$, $z(0)z(2)$, and $z(0)z(1)z(2)^2$; the first monomial is expressed in $\mathcal{B}_3$ thanks to (19), the second and the third thanks to (20), and the fourth is of the same nature as $z(0)^2z(1)z(2)$. In order to obtain its expression, we will proceed along the same lines as in the previous paragraph. We will express jointly all three monomials of the same type. This leads eventually to the linear system

$$\begin{bmatrix} z(0)^2z(1)z(2) \\ z(1)^2z(0)z(2) \\ z(2)^2z(0)z(1) \end{bmatrix} = \boldsymbol{A}_3^{-1}\boldsymbol{B}_{31}\boldsymbol{B}_{32}\boldsymbol{b}_3(\boldsymbol{z}) \overset{\mathrm{def}}{=} \boldsymbol{C}_3\boldsymbol{b}_3(\boldsymbol{z}) \qquad (21)$$

where

$$\boldsymbol{A}_3 = \begin{bmatrix} 1 & 0 & -C_2(1,8) \\ 0 & 1 & -C_2(3,8) \\ -C_2(4,8) & 1 & 0 \end{bmatrix}$$

$$\boldsymbol{B}_{32} = \begin{bmatrix} \boldsymbol{I}_7 & 0 \\ \multicolumn{2}{c}{\boldsymbol{C}_1} \\ 0 & 1 \end{bmatrix}$$

$$\boldsymbol{B}_{31} = \begin{bmatrix} 0 & 0 & 0 & C_2(1,1) & 0 & C_2(1,2) & C_2(1,3) & 0 & 0 & C_2(1,4) & C_2(1,5) \\ 0 & 0 & 0 & C_2(3,1) & 0 & C_2(3,2) & C_2(3,3) & 0 & 0 & C_2(3,4) & C_2(3,5) \\ 0 & C_2(4,1) & 0 & 0 & C_2(4,3) & C_2(4,4) & 0 & C_2(4,2) & 0 & 0 & C_2(4,7) \end{bmatrix}.$$

6) Solutions to the Polynomial System: The matrix of the operator $\mathcal{M}_{z(0)}^{\mathrm{T}}$ can now be obtained as

$$M_{z(0)}^{\mathrm{T}} = \begin{bmatrix} 0 & & 0 & 0 & & & 0 & \\ 1 & & 0 & 0 & & & 0 & \\ 0 & & 0 & 0 & & & 0 & \\ 0 & C_1(1,:)^{\mathrm{T}} & 0 & 0 & C_2(1,:)^{\mathrm{T}} & C_2(2,:)^{\mathrm{T}} & 0 & C_3(1,:)^{\mathrm{T}} \\ 0 & & 1 & 0 & & & 0 & \\ 0 & & 0 & 1 & & & 0 & \\ 0 & & 0 & 0 & & & 0 & \\ 0 & & 0 & 0 & & & 1 & \end{bmatrix}^{\mathrm{T}}.$$

It admits generally eight eigenvectors $\boldsymbol{v}^{(m)}$, $1 \leq m \leq 8$. Each of these eigenvectors gives a solution $\boldsymbol{z}^{(m)} = \boldsymbol{v}^{(m)}(2 : 4)/v^{(m)}(1)$ because the first four entries of $\boldsymbol{b}_3$ are $[1, z(0), z(1), z(2)]$. The corresponding solutions $\boldsymbol{h}^{(m)}$ are obtained by transforming back to the original coordinate system $\boldsymbol{h}^{(m)} = \boldsymbol{T}\boldsymbol{z}^{(m)}$.

References

[1] K. Abed-Meraim *et al.*, "On subspace methods for blind identification of SIMO FIR systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 42–55, Jan. 1997.

[2] K. Abed-Meraim, E. Moulines, and P. Loubaton, "Prediction error method for second-order blind identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 694–705, Mar. 1997.

[3] P. O. Amblard, M. Gaeta, and J. L. Lacoume, "Statistics for complex variables and signals," *Signal Process.*, vol. 53, pp. 1–25, 1996.

[4] S. Barbarossa and A. Scaglione, "Blind equalization using cost function matched to the signal constellation," in *Proc. Asilomar Conf, Signals, Syst., Comput.*, Pacific Grove, CA, 1997.

[5] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[6] P. Comon, "Circularité et signaux aléatoires à temps discret," *Traitement du Signal*, vol. 11, no. 5, pp. 417–420, Dec. 1994.

[7] P. Comon, O. Grellier, and B. Mourrain, "Closed-form blind channel identification with MSK inputs," in *Proc. Asilomar Conf, Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1–4, 1998, pp. 1569–1573.

[8] D. Cox, J. Little, and D. O'Shea, "Ideals, varieties, and algorithms: An introduction to computational algebraic geometry and commutative algebra," in *Undergraduate Texts in Mathematics*. New York: Springer-Verlag, 1992.

[9] M. Elkadi and B. Mourrain, "Algorithms for residues and Lojasiewicz exponents," *J. Pure Appl. Algebra*, vol. 153, pp. 27–44, 2000.

[10] I. Emiris and A. Rege, "Monomial bases and polynomial system solving," in *Proc. ACM Int. Symp. Symbolic Algebraic Comput.*, Oxford, U.K., 1994, pp. 114–122.

[11] W. A. Gardner, "A new method of channel identification," *IEEE Trans. Commun.*, vol. 39, pp. 813–817, June 1991.

[12] N. R. Goodman, "Statistical analysis based on certain multivariate complex normal distributions," *Ann. Math. Stat.*, vol. 34, pp. 152–177, 1963.

[13] J. Harris, "Algebraic geometry, a first course," in *Graduate Texts in Math*. New York: Springer, 1992, vol. 133.

[14] J. L. Lacoume, P. O. Amblard, and P. Comon, "Statistiques d'ordre supérieur pour le traitement du signal," in *Coll. Sci. l'Ingénieur*. Paris, France: Masson, 1997.

[15] T. H. Li, "Blind identification and deconvolution of linear systems driven by binary random sequences," *Trans. Inform. Theory*, vol. 38, pp. 26–38, Jan. 1992.

[16] Y. Li and Z. Ding, "ARMA system identification based on second-order cyclostationarity," *IEEE Trans. Signal Processing*, vol. 42, pp. 3483–3494, Dec. 1994.

[17] F. S. Macaulay, "Some formulae in elimination," *Proc. London Math. Soc.*, vol. 1, no. 33, pp. 3–27, 1902.

[18] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrague, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.

[19] B. Mourrain, "Computing isolated polynomial roots by matrix methods," *J. Symbolic Comput., Special Issue on Symbolic–Numeric Algebra for Polynomials*, vol. 26, no. 6, pp. 715–738, Dec. 1998.

[20] B. Mourrain and V. Y. Pan, "Multivariate polynomials, duality and structured matrices," *J. Complexity*, vol. 16, pp. 110–180, 2000.

[21] B. Picinbono, "On circularity," *IEEE Trans. Signal Processing*, vol. 42, pp. 3473–3482, Dec. 1994.

[22] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[23] E. Serpedin and G. B. Giannakis, "A simple proof of a known blind channel identifiability result," *IEEE Trans. Signal Processing*, vol. 47, pp. 591–593, Feb. 1999.

[24] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.

[25] D. T. M. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," in *Proc. ICASSP Conf.*, Adelaide, Australia, Apr. 1994.

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

[27] P. Trebuchet and B. Mourrain, "Solving projective complete intersection faster," in *Proc. Int. Symp. Symbolic Algebraic Comput.*, C. Traverso, Ed., New York, 2000, pp. 231–238.

[28] J. Tugnait, "Comments on 'New criteria for blind deconvolution of nonminimum phase systems'," *IEEE Trans. Inform. Theory*, vol. 38, pp. 210–213, Jan. 1992.

[29] R. A. Wooding, "The multivariate distribution of complex normal variables," *Biometrika*, vol. 43, pp. 212–215, 1956.

[30] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, pp. 813–817, Dec. 1995.

[31] D. Yellin and B. Porat, "Blind identification of FIR systems excited by discrete-alphabet inputs," *IEEE Trans. Signal Processing*, vol. 41, pp. 1331–1339, Mar. 1993.

Olivier Grellier was born in 1972. He graduated from Supelec, Paris, France, in 1995 and received the D.E.A. degree in 1995 from the University of Rennes I, Rennes, France. He received the Ph.D. degree in 2000 under the direction of P. Comon.

His Ph.D. dissertation was on blind deconvolution and separation of discrete sources. His works were focused, in particular, on analytical techniques. He is now with Amadeus Development, Sophia-Antipolis, France.

Pierre Comon (M'87–SM'95) graduated in 1982 and received the Ph.D. degree in 1985, both from the University of Grenoble, Grenoble, France. He later received the Habilitation to lead Researches degree in 1995 from the University of Nice, Nice, France.

He has been in industry for nearly 13 years, first with Crouzet-Sextant between 1982 and 1985 and then with Thomson Marconi between 1988 and 1997. He spent 1987 in the ISL Laboratory, Stanford University, Stanford, CA. He joined Eurecom in 1997 and left there in the fall of 1998. He was an Associate Research Director with CNRS from 1994 to 1998. He is now Research Director with Laboratory I3S, CNRS, Sophia-Antipolis, France. His research interests include high-order statistics, blind deconvolution and equalization, digital communications, 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 ATHOS from 1992 to 1995. Between 1992 and 1998, he was a member of the Technical and Scientific Council of the Thomson Group. Since July 2001, he has been an Associate Editor with the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I, in the area of blind techniques; he is currently an IEEE Distinguished Lecturer for 2002 to 2003.

Bernard Mourrain was born in 1964. He received the Agrégation de Mathématiques degree from École Normale Supérieure de Cachan in 1988 and defended his thesis on invariant theory and effective algebra and geometry in 1991, under the direction of M. Demazure.

Since 1991, he has been a Researcher at INRIA, where he leads the new group Galaad, which focusses on geometry, algebra, and applications. His main interests are polynomial systems solving and symbolic computation. He is currently working on symbolic and numerical methods in effective algebraic geometry based on linear algebra tools and resultant formulations. Applications of these methods to computer vision, robotics, signal processing and three-dimensional curves and surface manipulations are a significant part of his work.

Philippe Trébuchet was born in 1975. He received the Agrégation de Mathématiques degree from École Normale Supérieure de Cachan in 2000.

He is currently working on algebraic methods for polynomial system solving with the Galaad team, INRIA, France, under the direction of D. Lazard and B. Mourrain. His main interests are polynomial systems solving and symbolic computation. He is currently working on symbolic and numerical methods in effective algebraic geometry based on linear algebra tools and rewriting techniques. In addition, he is concerned with practical applications, particularly in the fields of signal processing, computer vision, computational geometry, and robotics.