

Combining the Frisch scheme and Yule–Walker equations for identifying multivariable errors–in–variables models

Roberto Diversi and Roberto Guidorzi

Abstract—Errors–in–Variables (EIV) models, i.e. models whose stochastic environment considers measurement errors on both inputs and outputs are intrinsically more realistic than representations assuming an exact knowledge of the input but are also more difficult to estimate. The difficulties increase in a non trivial way passing from the SISO and MISO cases to the MIMO one. This paper proposes a procedure for EIV identification of MIMO processes based on the Frisch scheme that assumes additional white noises on all inputs and outputs and shows its effectiveness by means of Monte Carlo simulations.

I. INTRODUCTION

The identification of errors–in–variables (EIV) models is a challenging problem that, despite the advantages that it can offer in many contexts, has received an increasing attention only in the last three decades [1]–[16]. The EIV environment enjoys some peculiarities not shared by other classes of models; among them, the symmetrical presence of additive noise on all variables (inputs and outputs if a model orientation is considered). EIV models are particularly suitable when the interest is mainly focused on the inner laws that describe the considered process and/or on filtering applications rather than on the prediction of its future behavior. This class of models has thus found many applications in system engineering, signal processing and econometrics [17], [18].

The identification of errors–in–variables models is more complex than that of traditional equation error models like, for instance, ARX and ARMAX where the input is considered as exactly known. An interesting EIV identification approach is based on the Frisch scheme, i.e. on the stochastic context originally introduced by Ragnar Frish in the static case for evaluating the family of linear relations linking sets of noisy observations compatible with the data covariance matrix [2], [19]. The extension of the original algebraic context to the dynamic one leads to the use of EIV models where the measures of the inputs and outputs are affected by additive and mutually uncorrelated white noises.

The dynamic case shares many interesting algebraic and geometric properties with the static one like the existence of convex singularity hypersurfaces defining the family of solutions that can be associated with a set of observations. Moreover, the dynamic properties of the system allow the asymptotic determination of the exact model of the process that has generated the data, not just the definition of a class

of compatible models [4]. This nice property is based on the existence of a common point between the singularity hypersurfaces associated with all models with orders greater than that of the process that has generated the data. Unfortunately this property cannot be observed on finite sequences so that it is necessary to introduce suitable cost functions in order to extract a single model from the data [8], [9], [10].

The extension of the Frisch scheme to the identification of multi–input multi–output (MIMO) errors–in–variables models introduces further problems not present in the SISO and MISO cases [6], [20]. In fact, the whole systems can be viewed, in this case, as a set of m (m = number of outputs) interconnected subsystems, that can be associated with m singularity hypersurfaces so that these subsystems could be separately identified by means of the same procedures used for the MISO case. An approach of this kind requires m separate searches and, even worse, leads to m different estimates for every input and output noise variance with a consequent loss of congruence for the global model. It would be thus preferable to select cost functions treating the model as a whole and not as a set of separate models. A geometric approach to the solution of this problem has been proposed in [6], [15] where multivariable EIV models have been no longer associated to single points on singularity hypersurfaces but to directions in the noise space. This approach, that relies, from a computational standpoint, on the easy computation of the intersection between a straight line from the origin and a singularity hypersurface [21], offers the additional advantage of an easy introduction of congruent cost functions [16].

The identification procedure introduced in this paper takes advantage of the properties of both the dynamic Frisch scheme and high order Yule–Walker equations and is based on the aforementioned geometric approach so that congruent estimations of the system parameters and noise variances can be obtained. The method is characterized by a good estimation accuracy and gives better results than the approach introduced in [16], as shown by the performed Monte Carlo simulations.

The organization of the paper is as follows. Section II defines the EIV MIMO identification problem. Section III describes the properties of the dynamic Frisch scheme in the multivariable case while Section IV introduces the new identification procedure. The results obtained in some Monte Carlo simulations are reported in Section V. Short concluding remarks are finally given in Section VI.

Roberto Diversi and Roberto Guidorzi are with the Department of Electronics, Computer Science and Systems, University of Bologna, Viale del Risorgimento 2, 40136 Bologna, Italy roberto.diversi@unibo.it, roberto.guidorzi@unibo.it

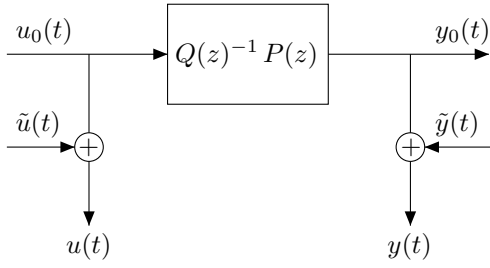


Fig. 1. MIMO errors-in-variables model

II. PROBLEM STATEMENT

Consider the MIMO EIV model represented in Fig. 1, described by the equations

$$Q(z)y_0(t) = P(z)u_0(t) \quad (1)$$

$$u(t) = u_0(t) + \tilde{u}(t) \quad (2)$$

$$y(t) = y_0(t) + \tilde{y}(t), \quad (3)$$

where $u_0(t) \in \mathcal{R}^r$, $y_0(t) \in \mathcal{R}^m$ are the noise-free input and output

$$u_0(t) = [u_{01}(t) \ u_{02}(t) \ \dots \ u_{0r}(t)]^T \quad (4)$$

$$y_0(t) = [y_{01}(t) \ y_{02}(t) \ \dots \ y_{0m}(t)]^T, \quad (5)$$

$u(t), y(t)$ are the available observations affected by the additive noises $\tilde{u}(t), \tilde{y}(t)$ and $Q(z), P(z)$ are $(m \times m)$ and $(m \times r)$ polynomial matrices in the unitary advance operator z ($z y(t) = y(t+1)$).

By selecting a suitable minimal parametrization [22], [23], model (1) can be partitioned into the set of m relations

$$y_{0i}(t + \nu_i) = \sum_{j=1}^m \sum_{k=1}^{\nu_{ij}} \alpha_{ijk} y_{0j}(t + k - 1) + \sum_{j=1}^r \sum_{k=1}^{\nu_i+1} \beta_{ijk} u_{0j}(t + k - 1) \quad i = 1, \dots, m \quad (6)$$

where the integers ν_i ($i = 1, \dots, m$) are the Kronecker observability invariants of the system while the integers ν_{ij} are completely defined by these invariants through the relations

$$\begin{aligned} \nu_{ij} &= \nu_i & \text{for } i = j \\ \nu_{ij} &= \min(\nu_i + 1, \nu_j) & \text{for } i > j \\ \nu_{ij} &= \min(\nu_i, \nu_j) & \text{for } i < j. \end{aligned}$$

The maximal observability invariant will be denoted as ν_M i.e.

$$\nu_M = \max_i \{\nu_i\} \quad i = 1, \dots, m;$$

the system order n , i.e. the McMillan degree of $Q(z)^{-1}P(z)$, is given by

$$n = \sum_{i=1}^m \nu_i.$$

Assumptions

- A1. System (1) is asymptotically stable.
- A2. $Q(z)$ and $P(z)$ are left coprime.
- A3. The noiseless input $u_0(t)$ can be either a zero-mean ergodic process or a quasi-stationary bounded deterministic signal, i.e. such that the limit

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N u_0(t) u_0^T(t - \tau)$$

exists $\forall \tau$ [24]. Moreover, $u_0(t)$ is assumed as persistently exciting of sufficiently high order.

- A4. $\tilde{u}(t)$ and $\tilde{y}(t)$ are zero-mean ergodic white processes with unknown diagonal covariance matrices

$$\tilde{\Sigma}_u^* = \text{diag} [\tilde{\sigma}_{u_1}^{2*}, \tilde{\sigma}_{u_2}^{2*}, \dots, \tilde{\sigma}_{u_r}^{2*}]$$

$$\tilde{\Sigma}_y^* = \text{diag} [\tilde{\sigma}_{y_1}^{2*}, \tilde{\sigma}_{y_2}^{2*}, \dots, \tilde{\sigma}_{y_m}^{2*}].$$

- A5. $\tilde{u}(t)$ and $\tilde{y}(t)$ are mutually uncorrelated and uncorrelated with the noise-free input $u_0(t)$.

- A6. The structural indexes $\nu_1, \nu_2, \dots, \nu_m$ will be considered *a priori* known.

Under these assumptions, the EIV MIMO identification problem can be stated as follows.

Problem 1: Given the set of noisy input-output observations $u(1), \dots, u(N)$, $y(1), \dots, y(N)$, estimate the coefficients $\alpha_{ijk}, \beta_{ijk}$ and the noise covariance matrices $\tilde{\Sigma}_u^*, \tilde{\Sigma}_y^*$.

Remark 1: The considered class of models is quite general and can describe any MIMO system; by contrast the commonly used “full polynomial form” representations, despite their practical usefulness, meet many limits (e.g. the order of the system must always be a multiple of the number of outputs). Another advantage of the considered models is associated with their minimal parametrization that assures not only a lower covariance of the estimate but the much more important identifiability of every parameter (no parameter can be expressed as a function of remaining ones); this property can be extremely relevant in fault diagnosis applications.

Remark 2: The integers ν_1, \dots, ν_m that describe the structure of the model have been assumed as known. Their determination (structural identification) will be described in another paper.

III. PROPERTIES OF THE DYNAMIC FRISCH SCHEME IN THE MIMO CASE

Let us define the vectors

$$\varphi_{0i}(t) = [-y_{01}(t) \ \dots \ -y_{01}(t + \nu_{i1} - 1) \ \dots \ -y_{0i}(t) \ \dots \ -y_{0i}(t + \nu_i) \ \dots \ -y_{0m}(t) \ \dots \ -y_{0m}(t + \nu_{im} - 1) \ \dots \ u_{01}(t) \ \dots \ u_{01}(t + \nu_i) \ \dots \ u_{0r}(t) \ \dots \ u_{0r}(t + \nu_i)]^T \quad (7)$$

$$\varphi_i(t) = [-y_1(t) \ \dots \ -y_1(t + \nu_{i1} - 1) \ \dots \ -y_i(t) \ \dots \ -y_i(t + \nu_i) \ \dots \ u_r(t) \ \dots \ u_r(t + \nu_i)]^T \quad (8)$$

$$\tilde{\varphi}_i(t) = [-\tilde{y}_1(t) \ \dots \ -\tilde{y}_1(t + \nu_{i1} - 1) \ \dots \ -\tilde{y}_i(t) \ \dots \ -\tilde{y}_i(t + \nu_i) \ \dots \ \tilde{u}_r(t) \ \dots \ \tilde{u}_r(t + \nu_i)]^T, \quad (9)$$

for $i = 1, \dots, m$, whose entries are noise-free, noisy and noise samples respectively. By defining also the parameter vectors

$$\theta_i^* = \begin{bmatrix} \alpha_{i11} \dots \alpha_{i1\nu_{i1}} \alpha_{i21} \dots \alpha_{i2\nu_{i2}} \dots \\ \alpha_{ii1} \dots \alpha_{iiv_i} 1 \dots \alpha_{im1} \dots \alpha_{im\nu_{im}} \\ \beta_{i11} \dots \beta_{i1(\nu_i+1)} \dots \beta_{ir1} \dots \beta_{ir(\nu_i+1)} \end{bmatrix}^T \quad (10)$$

for $i = 1, \dots, m$, and taking into account (6), (2), (3) it is possible to write

$$\varphi_{0i}^T(t) \theta_i^* = 0 \quad (11)$$

$$\varphi_i(t) = \varphi_{0i}(t) + \tilde{\varphi}_i(t), \quad (12)$$

for $i = 1, \dots, m$.

With reference to the i -th subsystem, define the covariance matrices of noise-free data, noisy data and noise

$$\Sigma_{0i} = E[\varphi_{0i}(t) \varphi_{0i}^T(t)] \quad (13)$$

$$\Sigma_i = E[\varphi_i(t) \varphi_i^T(t)] \quad (14)$$

$$\tilde{\Sigma}_i = E[\tilde{\varphi}_i(t) \tilde{\varphi}_i^T(t)]. \quad (15)$$

From (11), (12) and assumption A5 it follows that

$$\Sigma_{0i} \theta_i^* = 0 \quad (16)$$

$$\Sigma_i = \Sigma_{0i} + \tilde{\Sigma}_i \quad (17)$$

where, because of assumptions A4 and A5

$$\tilde{\Sigma}_i^* = \text{diag} [\tilde{\sigma}_{y_1}^{2*} I_{\nu_{i1}}, \dots, \tilde{\sigma}_{y_i}^{2*} I_{\nu_i+1}, \dots, \tilde{\sigma}_{y_m}^{2*} I_{\nu_{im}}, \\ \tilde{\sigma}_{u_1}^{2*} I_{\nu_i+1}, \dots, \tilde{\sigma}_{u_i}^{2*} I_{\nu_i+1}, \dots, \tilde{\sigma}_{u_r}^{2*} I_{\nu_i+1}]. \quad (18)$$

Note that Σ_{0i} is a positive semidefinite matrix whose null space has dimension one and defines the parameter vector θ_i^* . This matrix can be obtained from the noisy data covariance matrix Σ_i if the noise variances $\tilde{\sigma}_{y_1}^{2*}, \dots, \tilde{\sigma}_{y_m}^{2*}, \tilde{\sigma}_{u_1}^{2*}, \dots, \tilde{\sigma}_{u_r}^{2*}$ are known.

Consider now the set of all diagonal matrices $\tilde{\Sigma}_i(P)$ of type (18) that exhibit the same properties of $\tilde{\Sigma}_i^*$, i.e. the set of points $P = (\tilde{\sigma}_1^2, \dots, \tilde{\sigma}_m^2, \tilde{\sigma}_{m+1}^2, \dots, \tilde{\sigma}_{m+r}^2) \in \mathcal{R}^{m+r}$ such that

$$\Sigma_i - \tilde{\Sigma}_i(P) \geq 0, \quad \min \text{eig}(\Sigma_i - \tilde{\Sigma}_i(P)) = 0, \quad (19)$$

where

$$\tilde{\Sigma}_i(P) = \text{diag} [\tilde{\sigma}_1^2 I_{\nu_{i1}}, \dots, \tilde{\sigma}_i^2 I_{\nu_i+1}, \dots, \tilde{\sigma}_m^2 I_{\nu_{im}}, \\ \tilde{\sigma}_{m+1}^2 I_{\nu_i+1}, \dots, \tilde{\sigma}_{m+r}^2 I_{\nu_i+1}]. \quad (20)$$

It is possible to prove the following result [4].

Theorem 1: The set of points P satisfying (19) defines an hypersurface $\mathcal{S}(\Sigma_i)$ belonging to the first orthant of \mathcal{R}^{m+r} . Every point P of $\mathcal{S}(\Sigma_i)$ is an admissible solution in the noise space and can thus be associated with a possible vector of parameters $\theta_i(P)$ given by

$$(\Sigma_i - \tilde{\Sigma}_i(P)) \theta_i(P) = 0. \quad (21)$$

As an example, Fig. 2 shows a typical hypersurface for a system with one input and two outputs.

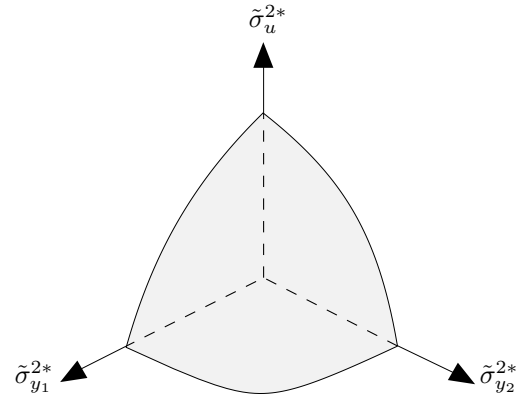


Fig. 2. Typical shape of $\mathcal{S}(\Sigma_i)$ for a one-input two-output EIV model.

The following corollary is a direct consequence of (16)–(17).

Corollary 1: The point $P^* = (\tilde{\sigma}_{y_1}^{2*}, \tilde{\sigma}_{y_2}^{2*}, \dots, \tilde{\sigma}_{y_m}^{2*}, \tilde{\sigma}_{u_1}^{2*}, \dots, \tilde{\sigma}_{u_r}^{2*})$, defined by the true noise variances, belongs to $\mathcal{S}(\Sigma_i)$ and is associated with the true parameter vector $\theta_i(P^*) = \theta_i^*$.

The above considerations are valid for each subsystem so that it is possible to define m hypersurfaces $\mathcal{S}(\Sigma_1), \mathcal{S}(\Sigma_2), \dots, \mathcal{S}(\Sigma_m)$ that exhibit the following important property.

Theorem 2: All hypersurfaces $\mathcal{S}(\Sigma_1), \mathcal{S}(\Sigma_2), \dots, \mathcal{S}(\Sigma_m)$ have the common point P^* corresponding to the actual input and output noise variances.

Define now the vectors

$$\varphi_0(t) = [-y_{01}(t) \dots -y_{01}(t + \nu_1) \quad (22)$$

$$-y_{02}(t) \dots -y_{02}(t + \nu_2) \dots -y_{0m}(t) \dots -y_{0m}(t + \nu_m)$$

$$u_{01}(t) \dots u_{01}(t + \nu_M) \dots u_{0r}(t) \dots u_{0r}(t + \nu_M)]^T$$

$$\varphi(t) = [-y_1(t) \dots -y_1(t + \nu_1) \quad (23)$$

$$-y_2(t) \dots -y_2(t + \nu_2) \dots -y_m(t) \dots -y_m(t + \nu_m)$$

$$u_1(t) \dots u_1(t + \nu_M) \dots u_r(t) \dots u_r(t + \nu_M)]^T$$

$$\tilde{\varphi}(t) = [-\tilde{y}_1(t) \dots -\tilde{y}_1(t + \nu_1) \quad (24)$$

$$-\tilde{y}_2(t) \dots -\tilde{y}_2(t + \nu_2) \dots -\tilde{y}_m(t) \dots -\tilde{y}_m(t + \nu_m)$$

$$\tilde{u}_1(t) \dots \tilde{u}_1(t + \nu_M) \dots \tilde{u}_r(t) \dots \tilde{u}_r(t + \nu_M)]^T$$

and the matrix of parameters

$$\Theta^* = [\bar{\theta}_1^* \bar{\theta}_2^* \dots \bar{\theta}_m^*], \quad (25)$$

where

$$\bar{\theta}_i^* = [\alpha_{i11} \dots \alpha_{i1\nu_{i1}} \underbrace{0 \dots 0}_{(\nu_i+1-\nu_{i1})} \dots \quad (26)$$

$$\alpha_{i11} \dots \alpha_{iiv_i} 1 \dots \alpha_{im1} \dots \alpha_{im\nu_{im}} \underbrace{0 \dots 0}_{(\nu_m+1-\nu_{im})}$$

$$\beta_{i11} \dots \beta_{i1(\nu_i+1)} \underbrace{0 \dots 0}_{(\nu_M-\nu_i)} \dots \beta_{ir1} \dots \beta_{ir(\nu_i+1)} \underbrace{0 \dots 0}_{(\nu_M-\nu_i)}]^T,$$

for $i = 1, \dots, m$. Note that $\bar{\theta}_i^*$ differs from θ_i^* only for the presence of some structural zeros. Because of (6) and (2)–

(3) it is easy to show that

$$\varphi_0^T(t) \Theta^* = 0 \quad (27)$$

$$\varphi(t) = \varphi_0(t) + \tilde{\varphi}(t). \quad (28)$$

Introduce also the covariance matrices

$$\Sigma_0 = E[\varphi_0(t) \varphi_0^T(t)] \quad (29)$$

$$\Sigma = E[\varphi(t) \varphi^T(t)] \quad (30)$$

$$\tilde{\Sigma}^* = E[\tilde{\varphi}(t) \tilde{\varphi}^T(t)]. \quad (31)$$

From (27) and (28) it follows immediately that

$$\Sigma_0 \Theta^* = 0 \quad (32)$$

$$\Sigma = \Sigma_0 + \tilde{\Sigma}^* \quad (33)$$

where

$$\tilde{\Sigma}^* = \text{diag} [\tilde{\sigma}_{y_1}^{2*} I_{\nu_{1+1}}, \dots, \tilde{\sigma}_{y_m}^{2*} I_{\nu_{m+1}}, \tilde{\sigma}_{u_1}^{2*} I_{\nu_{M+1}}, \dots, \tilde{\sigma}_{u_r}^{2*} I_{\nu_{M+1}}]. \quad (34)$$

Matrix Σ_0 is thus positive semidefinite with a null space of dimension m .

Consider now the family of all diagonal matrices $\tilde{\Sigma}(P)$ of type (34) which exhibit the same properties of $\tilde{\Sigma}^*$, i.e. the set of points $P \in \mathcal{R}^{m+r}$ such that

$$\Sigma - \tilde{\Sigma}(P) \geq 0, \quad \min \text{eig}(\Sigma - \tilde{\Sigma}(P)) = 0, \quad (35)$$

where $\tilde{\Sigma}(P)$ has the structure (34). This set is described by the following theorem that constitutes a simple extension of the result reported in [4].

Theorem 3: The set of points satisfying (35) defines an hypersurface $\mathcal{S}(\Sigma)$ belonging to the first orthant of \mathcal{R}^{m+r} and lying under $\mathcal{S}(\Sigma_1), \dots, \mathcal{S}(\Sigma_m)$. Because of (32) and (33), the point P^* , characterized by the true noise variances, belongs to $\mathcal{S}(\Sigma)$ and $\Sigma - \tilde{\Sigma}(P^*) = \Sigma_0$.

The properties of $\mathcal{S}(\Sigma)$ play an important role in the solution of Problem 1, as it will be shown in the next section.

IV. IDENTIFICATION OF MIMO EIV MODELS

Although of great theoretical interest, the properties described by Theorems 2 and 3 cannot be directly used in practice. In fact, in presence of a finite number N of noisy data, the covariance matrices Σ_i , $i = 1, \dots, m$ and Σ must be replaced by their sample estimates so that the associated hypersurfaces do no longer share any common point.

It is of course possible to identify the m subsystems separately by using one of the selection criteria proposed for the SISO or MISO cases [8], [9], [10]. The generic subsystem i would thus be treated as a MISO model where all outputs except the i -th one play the role of inputs. Nevertheless, an approach of this kind is characterized by two main drawbacks. Firstly, it is necessary to solve m optimization problems of complexity $m+r$. This increases the computational burden but it is much more important to observe that this procedure leads to m different estimates for every input and output noise variance with a consequent loss of congruence. To overcome these problems, the search

procedure and the selection criterion should treat the m subsystems as a whole MIMO system.

For this reason, the solution of Problem 1 will be based on the geometric approach proposed in [6], [15], where multivariable EIV models have been no longer associated to single points on singularity hypersurfaces but to directions in the noise space. This approach relies on the computation of the intersection between a straight line from the origin and a singularity hypersurface described by the following theorem [21].

Theorem 4: Let $\xi = (\xi_1, \xi_2, \dots, \xi_{m+r})$ be a generic point of the first orthant of \mathcal{R}^{m+r} and ρ the straight line from the origin through ξ . Its intersection with $\mathcal{S}(\Sigma_i)$ is the point P_i given by

$$P_i = \frac{\xi}{\lambda_{M_i}}, \quad \lambda_{M_i} = \max \text{eig}(\Sigma_i^{-1} \tilde{\Sigma}_i^\xi) \quad (36)$$

where

$$\tilde{\Sigma}_i^\xi = \text{diag} [\xi_1 I_{\nu_{i1}}, \dots, \xi_i I_{\nu_{i+1}}, \dots, \xi_m I_{\nu_{im}}, \xi_{m+1} I_{\nu_{i+1}}, \dots, \xi_{m+r} I_{\nu_{i+1}}]. \quad (37)$$

Remark 3: To compute the intersection \bar{P} between $\mathcal{S}(\Sigma)$ and ρ , (36) will be replaced with

$$\bar{P} = \frac{\xi}{\lambda_M}, \quad \lambda_M = \max \text{eig}(\Sigma^{-1} \tilde{\Sigma}^\xi), \quad (38)$$

where $\tilde{\Sigma}^\xi$ has the structure (34).

It is thus possible to consider the m intersections P_1, P_2, \dots, P_m of a line ρ from the origin with the m singularity surfaces $\mathcal{S}(\Sigma_1), \dots, \mathcal{S}(\Sigma_m)$ that define univocally the parameters $\theta_1(P_1), \theta_2(P_2), \dots, \theta_m(P_m)$ i.e. a whole MIMO model. The use of this approach, which is valid also for finite numbers of data, still leads to m different values of each noise variance. However, since P_1, P_2, \dots, P_m belong to the same straight line from the origin, the ratios between the noise variances are univocally defined.

It is now necessary to introduce a selection criterion to identify the “true” direction in the noise space, i.e. that associated with the true model. The criterion introduced in the next subsection takes advantage of the properties of both the high-order Yule-Walker equations and the hypersurface $\mathcal{S}(\Sigma)$.

A. The selection criterion

Let us define the $r q \times 1$ (being q a user-chosen parameter) vector of delayed inputs

$$\varphi_u^d(t) = [u_1(t-q) \dots u_1(t-1) \dots u_2(t-q) \dots u_2(t-1) \dots u_r(t-q) \dots u_r(t-1)]^T, \quad (39)$$

that, because of (2), satisfies the condition

$$\varphi_u^d(t) = \varphi_{u0}^d(t) + \tilde{\varphi}_u^d(t), \quad (40)$$

where

$$\varphi_{u_0}^d(t) = [u_{01}(t-q) \dots u_{01}(t-1) \quad (41)$$

$$u_{02}(t-q) \dots u_{02}(t-1) \dots u_{0r}(t-q) \dots u_{0r}(t-1)]^T,$$

$$\tilde{\varphi}_u^d(t) = [\tilde{u}_1(t-q) \dots \tilde{u}_1(t-1) \quad (42)$$

$$\tilde{u}_2(t-q) \dots \tilde{u}_2(t-1) \dots \tilde{u}_r(t-q) \dots \tilde{u}_r(t-1)]^T.$$

Define also the $r q \times [n + m + (r + 1) \nu_M]$ matrix

$$\Sigma^d = E [\varphi_u^d(t) \varphi^T(t)] \quad (43)$$

$$= E [(\varphi_{u_0}^d(t) + \tilde{\varphi}_u^d(t)) (\varphi_0(t) + \tilde{\varphi}(t))^T], \quad (44)$$

that can be estimated from the noisy data. Because of assumptions A4–A5 we have

$$\Sigma^d = E [\varphi_{u_0}^d(t) \varphi_0^T(t)], \quad (45)$$

so that, from (27)

$$\Sigma^d \Theta^* = 0. \quad (46)$$

Relation (46) represents a set of high order Yule–Walker equations that can be used to define a selection criterion.

For this purpose, consider a direction ρ in the noise space, its intersections P_1, \dots, P_m with $\mathcal{S}(\Sigma_1), \dots, \mathcal{S}(\Sigma_m)$ and the associated MIMO model $\theta_1(P_1), \theta_2(P_2), \dots, \theta_m(P_m)$. Relation (46) allows the introduction of the cost function

$$J(P_1, \dots, P_m) = J(\rho) \quad (47)$$

$$= \text{trace} (\Theta^T(P_1, \dots, P_m) (\Sigma^d)^T \Sigma^d \Theta(P_1, \dots, P_m)),$$

where $\Theta(P_1, \dots, P_m) = \Theta(\rho)$ has the same structure of Θ^* and has been constructed with the entries of $\theta_1(P_1), \dots, \theta_m(P_m)$. This function satisfies the conditions

- 1) $J(\rho) \geq 0$
- 2) $J(\rho) = 0$ if $P_1 = \dots = P_m = P^*$,

where the second condition follows directly from (46). Problem 1 can thus be solved by minimizing $J(\rho)$ in the noise space. Of course, when the covariance matrices are replaced by their sample estimates relation (46) does no longer hold exactly and the minimum of the cost function (47) will be greater than zero.

B. The identification algorithm

The whole identification procedure is summarized by the following algorithm.

Algorithm 1.

- 1) Compute, on the basis of the noisy observations $u(1), \dots, u(N), y(1), \dots, y(N)$, the sample estimates of the covariance matrices $\Sigma_1, \dots, \Sigma_m$ and Σ :

$$\hat{\Sigma}_i = \frac{1}{N - \nu_M} \sum_{t=1}^{N - \nu_M} \varphi_i(t) \varphi_i^T(t), \quad i = 1, \dots, m$$

$$\hat{\Sigma} = \frac{1}{N - \nu_M} \sum_{t=1}^{N - \nu_M} \varphi(t) \varphi^T(t).$$

- 2) Choose a value of the delay q in (39) and compute the sample estimate of the covariance matrix Σ^d :

$$\hat{\Sigma}^d = \frac{1}{N - \nu_M - q} \sum_{t=q+1}^{N - \nu_M} \varphi_u^d(t) \varphi^T(t).$$

- 3) Start from a generic direction ρ in the first orthant of \mathcal{R}^{m+r} .
- 4) Compute the intersections P_1, \dots, P_m between ρ and $\mathcal{S}(\hat{\Sigma}_1), \dots, \mathcal{S}(\hat{\Sigma}_m)$.
- 5) Construct the diagonal matrices $\tilde{\Sigma}(P_1), \dots, \tilde{\Sigma}(P_m)$ and compute the parameter vectors $\theta_1(P_1), \theta_2(P_2), \dots, \theta_m(P_m)$ by means of the relations

$$(\hat{\Sigma}_i - \tilde{\Sigma}_i(P_i)) \theta_i(P_i) = 0 \quad i = 1, \dots, m.$$
- 6) Construct, with the entries of $\theta_1(P_1), \dots, \theta_m(P_m)$ the matrix $\Theta(\rho)$ of type (25).
- 7) Compute the cost function $J(\rho)$ (47).
- 8) Move to a new direction $\rho \pm \Delta\rho$ corresponding to a decrease of $J(\rho)$.
- 9) Repeat steps 3–8 until the direction ρ° corresponding to the minimum of $J(\rho)$ is found.

Algorithm 1 involves the computation of the points P_1, \dots, P_m and \bar{P} in the noise space so that it gives $m + 1$ different estimations for each noise variance $\hat{\sigma}_{y_1}^{2*}, \dots, \hat{\sigma}_{y_m}^{2*}, \hat{\sigma}_{u_1}^{2*}, \dots, \hat{\sigma}_{u_r}^{2*}$. Even if other choices are possible we will consider as noise variance estimates the values given by the point \bar{P} on $\mathcal{S}(\hat{\Sigma})$. From Theorem 3 it is easy to show that these values lead to matrices $\hat{\Sigma}_i - \tilde{\Sigma}_i(\bar{P})$, $i = 1, \dots, m$, that are all positive definite. The algorithm is thus concluded as follows

- 10) Compute the intersection \bar{P} between ρ° and $\mathcal{S}(\hat{\Sigma})$ and consider the entries of \bar{P} as estimates of the true noise variances.

Remark 4: It is important to observe that the whole procedure requires only the computation of a very limited number of points of the considered singularity hypersurfaces, i.e. only their intersections with the directions tested by the adopted optimization algorithm. The whole procedure is thus very efficient from a numerical point of view.

V. NUMERICAL RESULTS

The effectiveness of the proposed identification method has been tested by means of Monte Carlo simulations and compared to that of the identification procedure described in [16]. The numerical example concerns the two–input two–output model considered in [16], characterized by the polynomial matrices

$$Q(z) = \begin{bmatrix} z^2 - 0.4z + 0.3 & z + 0.2 \\ -0.2z + 0.1 & z^2 - 0.4 \end{bmatrix}$$

$$P(z) = \begin{bmatrix} 0.6z^2 - 0.1 & 0.2z^2 + 0.8z + 1 \\ 0.2z^2 + 0.5z + 0.5 & 0.4z^2 - z + 0.5 \end{bmatrix}.$$

The true inputs $u_{01}(t), u_{02}(t)$ are pseudo random binary sequences of unit variance and length $N = 2000$. The true

TABLE I
TRUE AND ESTIMATED VALUES OF THE COEFFICIENTS OF $Q(z)$.

	α_{111}	α_{112}	α_{121}	α_{122}
true	0.3	-0.4	0.2	1
YW (Exp.1)	0.3018 ± 0.0148	-0.4013 ± 0.0279	0.1980 ± 0.0478	0.9973 ± 0.0325
NS (Exp.1)	0.3035 ± 0.0217	-0.4060 ± 0.0508	0.1933 ± 0.0727	1.0091 ± 0.0722
YW (Exp.2)	0.3130 ± 0.0238	-0.4131 ± 0.0379	0.1656 ± 0.0696	0.9352 ± 0.0933
NS (Exp.2)	0.3535 ± 0.0358	-0.4957 ± 0.0757	0.0078 ± 0.1145	0.8171 ± 0.0570
	α_{211}	α_{212}	α_{221}	α_{222}
true	0.1	-0.2	-0.4	0
YW (Exp.1)	0.1017 ± 0.0103	-0.2011 ± 0.0216	-0.4014 ± 0.0316	0.0013 ± 0.0211
NS (Exp.1)	0.1047 ± 0.0166	-0.2078 ± 0.0358	-0.4131 ± 0.0572	0.0023 ± 0.0615
YW (Exp.2)	0.1050 ± 0.0122	-0.1985 ± 0.0222	-0.3872 ± 0.0366	-0.0059 ± 0.0318
NS (Exp.2)	0.0543 ± 0.0130	-0.0855 ± 0.0279	-0.1693 ± 0.0506	0.2137 ± 0.0927

TABLE II
TRUE AND ESTIMATED VALUES OF THE COEFFICIENTS OF $P(z)$.

	β_{111}	β_{112}	β_{113}	β_{121}	β_{122}	β_{123}
true	-0.1	0	0.6	1	0.8	0.2
YW (Exp.1)	-0.1038 ± 0.0340	0.0008 ± 0.0288	0.6029 ± 0.0182	0.9991 ± 0.0491	0.7993 ± 0.0281	0.2023 ± 0.0169
NS (Exp.1)	-0.0956 ± 0.0501	-0.0004 ± 0.0331	0.5980 ± 0.0388	0.9828 ± 0.1334	0.7993 ± 0.0300	0.2010 ± 0.0172
YW (Exp.2)	-0.1425 ± 0.0598	-0.0182 ± 0.0408	0.6078 ± 0.0202	1.0431 ± 0.0892	0.7721 ± 0.0428	0.2028 ± 0.0166
NS (Exp.2)	-0.2982 ± 0.0623	-0.1203 ± 0.0677	0.8260 ± 0.0740	0.9978 ± 0.1485	0.6684 ± 0.0479	0.1959 ± 0.0187
	β_{211}	β_{212}	β_{213}	β_{221}	β_{222}	β_{223}
true	0.5	0.5	0.2	0.5	-1	0.4
YW (Exp.1)	0.5030 ± 0.0220	0.4985 ± 0.0182	0.1998 ± 0.0131	0.4994 ± 0.0398	-1.0009 ± 0.0178	0.4001 ± 0.0123
NS (Exp.1)	0.4998 ± 0.0633	0.4909 ± 0.0482	0.1982 ± 0.0186	0.4906 ± 0.0914	-0.9981 ± 0.0582	0.3983 ± 0.0185
YW (Exp.2)	0.5201 ± 0.0383	0.5138 ± 0.0265	0.2068 ± 0.0185	0.5272 ± 0.0632	-1.0284 ± 0.0495	0.4096 ± 0.0200
NS (Exp.2)	0.9062 ± 0.1496	0.8402 ± 0.1057	0.2794 ± 0.0314	0.4216 ± 0.1003	-0.8475 ± 0.0751	0.3681 ± 0.0222

TABLE III
TRUE AND ESTIMATED VALUES OF THE NOISE VARIANCES.

	$\tilde{\sigma}_{u_1}^{2*}$	$\tilde{\sigma}_{u_2}^{2*}$	$\tilde{\sigma}_{y_1}^{2*}$	$\tilde{\sigma}_{y_2}^{2*}$
true	0.1	0.1	0.75	0.22
YW (Exp.1)	0.0982 ± 0.0133	0.0971 ± 0.0077	0.7222 ± 0.0407	0.2129 ± 0.0108
NS (Exp.1)	0.0886 ± 0.0605	0.0946 ± 0.0311	0.7402 ± 0.0289	0.2248 ± 0.0294
YW (Exp.2)	0.1058 ± 0.0192	0.1000 ± 0.0156	0.7620 ± 0.0628	0.1561 ± 0.0828
NS (Exp.2)	0.3606 ± 0.0671	0.0380 ± 0.0458	0.8040 ± 0.0445	0.0459 ± 0.0465

values of the noise variances are

$$\tilde{\sigma}_{u_1}^{2*} = 0.1 \quad \tilde{\sigma}_{u_2}^{2*} = 0.1 \quad \tilde{\sigma}_{y_1}^{2*} = 0.75 \quad \tilde{\sigma}_{y_2}^{2*} = 0.22,$$

corresponding to signal-to-noise ratios of about 10 dB on both inputs and outputs.

Two experiments have been performed. In the first experiment the tested algorithms are initialized by choosing a direction ρ near the true one while in the second experiment the initial direction is far from the true one. For both cases a Monte Carlo simulation of 100 independent runs has been carried out.

The value of q in step 2 has been set to 5 since it has been observed that greater values do not lead to significant improvements. The obtained results are shown in Tables 1–3 where the true value of parameters and noise variances, the means of their estimates and the corresponding standard deviations are reported. The proposed Yule–Walker based algorithm is denoted as YW while the approach [16], based on the null space of Σ_0 is denoted as NS. It can be observed that the YW method is characterized by better accuracy and higher robustness with respect to the initial estimate of the noise space direction.

VI. CONCLUSIONS

This paper has proposed a MIMO EIV identification procedure based on the use of a general class of minimally parameterized models particularly suitable for diagnosis and filtering applications. This procedure is based on a geometric approach that associates EIV models with directions in the noise space and leads to the computation of both model parameters and noise variances. The performed Monte Carlo simulations show the effectiveness of this approach.

REFERENCES

- [1] T. Söderström, Identification of stochastic linear systems in presence of input noise, *Automatica*, vol. 17, 1981, pp. 713–725.
- [2] R. E. Kalman, System identification from noisy data, in A. R. Bednarek and L. Cesari (Eds.), *Dynamical Systems II*, Academic Press, 1982, pp. 135–164.
- [3] B. D. O. Anderson and M. Deistler, Identifiability of dynamic errors-in-variables models, *Journal of Time Series Analysis*, vol. 5, 1984, pp. 1–13.
- [4] S. Beghelli, R. Guidorzi and U. Soverini, The Frisch scheme in dynamic system identification, *Automatica* vol. 26, 1990, pp. 171–176.
- [5] P. Stoica, M. Cedervall and A. Eriksson, Combined instrumental variable and subspace fitting approach to parameter estimation of noisy input–output systems, *IEEE Transactions on Signal Processing*, vol. 43, 1995, pp. 2386–2397.
- [6] R. Guidorzi, Identification of multivariable processes in the Frisch scheme context, *MTNS '96*, St. Louis, June 1996.
- [7] W. X. Zheng, Transfer function estimation from noisy input and output data, *International Journal of Adaptive Control and Signal Processing*, vol. 12, 1998, pp. 365–380.
- [8] S. Beghelli, P. Castaldi, R.P. Guidorzi and U. Soverini, A robust criterion for model selection in identification from noisy data, *Proc. of Ninth International Conference on Systems Engineering*, Las Vegas, Nevada, 1993, pp. 480–484.
- [9] R. Diversi, R. Guidorzi and U. Soverini, Frisch scheme-based algorithms for EIV identification, *Proc. of the 12th IEEE Mediterranean Conference on Control and Automation*, Kusadasi, Turkey, 2004.
- [10] R. Diversi, R. Guidorzi and U. Soverini, Yule–Walker equations in the Frisch scheme solution of errors-in-variables identification problems, *Proc. of the 17th International Symposium on Mathematical Theory of Networks and Systems*, Kyoto, Japan, 2006, pp. 391–395.
- [11] R. Diversi, R. Guidorzi and U. Soverini, Maximum likelihood identification of noisy input–output models, *Automatica*, vol. 43, 2007, pp. 464–472.
- [12] T. Söderström, Errors-in-Variables methods in system identification, *Automatica*, vol. 43, 2007, pp. 939–958.
- [13] K. Mahata, An improved bias-compensation approach for errors-in-variables model identification, *Automatica*, vol. 43, 2007, pp. 1339–1354.
- [14] J. C. Aguero and G. C. Goodwin, Identifiability of errors in variables dynamic systems, *Automatica*, vol. 44, 2008, pp. 371–382.
- [15] R. Guidorzi and R. Diversi, A geometric approach to multivariable errors-in-variables identification, *Prep. of the 15th IFAC Symposium on System Identification*, St. Malo, France, 2009, pp. 1615–1619.
- [16] R. Diversi and R. Guidorzi, Frisch scheme-based identification of multivariable errors-in-variables models, *Prep. of the 15th IFAC Symposium on System Identification*, St. Malo, France, 2009, pp. 1563–1567.
- [17] S. Van Huffel (Ed.), *Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modelling*, SIAM, Philadelphia, PA, 1997.
- [18] S. Van Huffel and P. Lemmerling (Eds.), *Total Least Squares Techniques and Errors-in-Variables Modelling: Analysis, Algorithms and Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
- [19] R. Frisch, Statistical confluence analysis by means of complete regression systems, *Pub. No. 5, Economic Institute*, Oslo University, Norway, 1934.
- [20] R. Guidorzi, R. Diversi and U. Soverini, The Frisch scheme in algebraic and dynamic identification problems, *Kybernetika*, vol. 44, 2008, pp. 585–616.
- [21] R. Guidorzi and M. Pierantoni, A new parametrization of Frisch scheme solutions, *Proc. of the XII Int. Conf. on Systems Science*, Wroclaw, 1995, pp. 114–120.
- [22] R. Guidorzi, Invariants and canonical forms for system structural and parametric identification, *Automatica*, vol. 17, 1981, pp. 117–133.
- [23] R. Guidorzi, Equivalence, invariance and dynamical system canonical modelling – Part I, *Kybernetika*, vol. 25, 1989, pp. 233–257.
- [24] L. Ljung, *System Identification – Theory for the User*, Prentice Hall, Englewood Cliffs, NJ, 1999.