Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Frequency domain EIV identification: a Frisch Scheme approach

Umberto Soverini * **Torsten Söderström** **

* *Department of Electrical, Electronic and Information Engineering, University of Bologna, Italy (e-mail: umberto.soverini@unibo.it)*
** *Department of Information Technology, Uppsala University, Sweden (e-mail: ts@it.uu.se)*

Abstract: This paper deals with the identification in the frequency domain of errors–in–variables models corrupted by additive and uncorrelated white noises. The noise–free input is an arbitrary signal, not necessarily periodic. The frequency domain algorithm proposed here resembles in many aspects the so–called Frisch Scheme approach, originally developed in the time domain.

Keywords: Linear dynamic systems; System identification; Errors–in–variables models; Discrete Fourier Transform; Frisch Scheme.

1. INTRODUCTION

Representations where errors or measurement noises are present on both inputs and outputs are usually called errors–in–variables (EIV) models and play an important role in several engineering applications. The identification of EIV models has been deeply investigated in the literature and many solutions have been proposed with different approaches (Söderström, 2007, 2012).

Among them, frequency domain identification techniques have received particular attention (Pintelon and Schoukens, 2012). The frequency domain approach has some special features, not present in time domain methods. In particular, filtering can be reduced to the selection of appropriate frequencies in a limited band of the signal spectrum; moreover, continuous–time and discrete–time models can be handled with equal difficulties (McKelvey, 2002). From a theoretic point of view, there is a full equivalence between time and frequency domain identification methods, also for finite data records (Agüero et al., 2010).

In this work the EIV identification problem for SISO systems is addressed by using a frequency domain approach, when the noise-free input is an arbitrary sequence. The system input and output are assumed to be affected by additive white noises with unknown variances. The proposed solution can be viewed as a frequency domain version of the so–called Frisch Scheme approach, originally developed for time domain techniques (Guidorzi et al., 2008). In particular, the proposed identification criterion can be considered as the frequential counterpart of the time domain covariance–matching criterion, originally presented in (Diversi et al., 2003).

The organization of the paper is as follows. Section 2 defines the EIV identification problem in the frequency domain. In Section 3 the problem is reformulated as a Frisch Scheme problem and the search for the solution is analyzed within this context. Section 4 describes a possible identification criterion, based on covariance–matching properties that can be directly computed in the frequency domain. In Section 5 the effectiveness of the proposed approach is verified by means of Monte Carlo simulations. Finally some concluding remarks are reported in Section 6.

2. STATEMENT OF THE PROBLEM

Consider the linear time–invariant SISO system described in Figure 1. The noise–free input and output $\hat{u}(t)$, $\hat{y}(t)$ are linked by the linear difference equation

$$A(z^{-1})\,\hat{y}(t) = B(z^{-1})\,\hat{u}(t), \tag{1}$$

where $A(z^{-1})$ and $B(z^{-1})$ are polynomials in the backward shift operator z^{-1}

$$A(z^{-1}) = 1 + \alpha_1\, z^{-1} + \cdots + \alpha_n\, z^{-n} \tag{2}$$
$$B(z^{-1}) = \beta_0 + \beta_1\, z^{-1} + \cdots + \beta_n\, z^{-n}. \tag{3}$$

In the EIV environment the input and output measurements are assumed to be corrupted by additive noise so that the available observations are

$$u(t) = \hat{u}(t) + \tilde{u}(t) \tag{4}$$
$$y(t) = \hat{y}(t) + \tilde{y}(t). \tag{5}$$

The following assumptions are satisfied.

- A1. System (1) is asymptotically stable.
- A2. $A(z^{-1})$ and $B(z^{-1})$ do not share any common factor.
- A3. The order n of the system is assumed as *a priori* known.
- A4. The noise–free input $\hat{u}(t)$ is a quasi–stationary bounded deterministic signal (Ljung, 1999) and is persistently exciting of sufficiently high order.
- A5. $\tilde{u}(t)$ and $\tilde{y}(t)$ are zero–mean, mutually uncorrelated Gaussian white processes with *unknown* variances λ_u^* and λ_y^*, respectively.

Let $\{u(t)\}_{t=0}^{N-1}$ and $\{y(t)\}_{t=0}^{N-1}$ be a set of input and output observations at N equidistant time instants. The corresponding Discrete Fourier Transforms (DFTs) are defined as

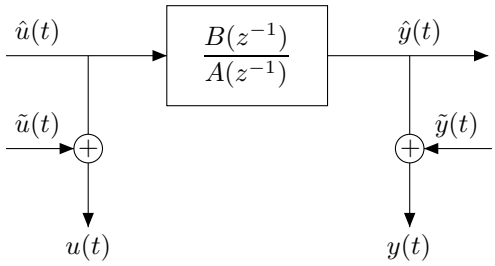

Fig. 1. Errors–in–variables model

$$U(\omega_k) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} u(t)\, \mathrm{e}^{-j\omega_k t} \tag{6}$$

$$Y(\omega_k) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} y(t)\, \mathrm{e}^{-j\omega_k t}\ , \tag{7}$$

where $\omega_k = 2\pi k/N$ and $k = 0, \ldots, N-1$. The system transfer function is represented as

$$G(\mathrm{e}^{-j\omega_k}) = \frac{B(\mathrm{e}^{-j\omega_k})}{A(\mathrm{e}^{-j\omega_k})}. \tag{8}$$

The DFTs defined in (6) and (7) can be expressed in matrix form by introducing the $N \times N$ Fourier matrix F_N (Agüero et al., 2010) whose entries are defined as follows

$$F_N = [f_{ik}] \tag{9}$$

$$f_{ik} = \frac{1}{\sqrt{N}} \mathrm{e}^{-j\frac{2\pi}{N}(i-1)(k-1)} \quad i,k = 1, \ldots, N\ . \tag{10}$$

In can be proved that matrix F_N is unitary, i.e. $F_N F_N^H = I$, where $(\cdot)^H$ denotes the transpose and conjugate operation.

Defining the following vectors in time and frequency domain

$$v_u = [u(0) \ldots u(N-1)]^T \tag{11}$$

$$v_y = [y(0) \ldots y(N-1)]^T \tag{12}$$

$$V_U = [U(\omega_0) \ldots U(\omega_{N-1})]^T \tag{13}$$

$$V_Y = [Y(\omega_0) \ldots Y(\omega_{N-1})]^T \tag{14}$$

the relations (6) and (7) can be represented by the linear transformations

$$V_U = F_N\, v_u \tag{15}$$

$$V_Y = F_N\, v_y\ . \tag{16}$$

In the frequency domain, the problem under investigation can be stated as follows.

Problem 1. Let $U(\omega_k)$, $Y(\omega_k)$ be a set of noisy measurements generated by an EIV system of type (1)–(5), under Assumptions A1–A5, where $\omega_k = 2\pi k/N$ and $k = 0, \ldots, N-1$. Estimate the system parameters α_i $(i = 1, \ldots, n)$, β_i $(i = 0, \ldots, n)$ and the noise variances λ_u^*, λ_y^*.

Under the additional assumption that the noise variances λ_u^*, λ_y^* are unknown but their ratio $\rho = \lambda_y^*/\lambda_u^*$ is known, a maximum likelihood solution of Problem 1. can be obtained, as described in (Soverini and Söderström, 2014).

3. FREQUENCY DOMAIN FRISCH SCHEME

With reference to the noise–free signals $\hat{u}(t)$ and $\hat{y}(t)$, definitions similar to (11)–(14) and (15)–(16) hold, i.e.

$$\hat{v}_u = [\hat{u}(0) \ldots \hat{u}(N-1)]^T \tag{17}$$

$$\hat{v}_y = [\hat{y}(0) \ldots \hat{y}(N-1)]^T \tag{18}$$

$$\hat{V}_U = [\hat{U}(\omega_0) \ldots \hat{U}(\omega_{N-1})]^T \tag{19}$$

$$\hat{V}_Y = [\hat{Y}(\omega_0) \ldots \hat{Y}(\omega_{N-1})]^T\ , \tag{20}$$

where

$$\hat{V}_U = F_N\, \hat{v}_u \tag{21}$$

$$\hat{V}_Y = F_N\, \hat{v}_y\ . \tag{22}$$

It is a well–known fact (Pintelon et al., 1997) that for finite N, even in absence of noise, the ratio of the DFTs $\hat{Y}(\omega_k)$ and $\hat{U}(\omega_k)$ $(\omega_k = 2\pi k/N)$ is not equal to the true transfer function

$$G(\mathrm{e}^{-j\omega_k}) \neq \frac{\hat{Y}(\omega_k)}{\hat{U}(\omega_k)}\ . \tag{23}$$

In fact, it can be proved that the DFTs $\hat{Y}(\omega_k)$ and $\hat{U}(\omega_k)$ exactly satisfy an extended model that includes also a transient term, i.e.

$$A(\mathrm{e}^{-j\omega_k})\, \hat{Y}(\omega_k) = B(\mathrm{e}^{-j\omega_k})\, \hat{U}(\omega_k) + T(\mathrm{e}^{-j\omega_k})\ , \tag{24}$$

where $T(z^{-1})$ is a polynomial of order $n-1$

$$T(z^{-1}) = \tau_0 + \tau_1\, z^{-1} + \cdots + \tau_{n-1}\, z^{-n+1} \tag{25}$$

that takes into account the effects of the initial and final conditions of the experiment.

By considering the whole number of frequencies, eq. (24) can be rewritten in a matrix form. For this purpose, introduce the parameter vectors

$$\theta_\alpha = [\alpha_n \ldots \alpha_1 1\ ]^T \tag{26}$$

$$\theta_\beta = [\beta_n \ldots \beta_1 \beta_0]^T \tag{27}$$

$$\theta_\tau = [\tau_{n-1} \ldots \tau_0]^T\ . \tag{28}$$

and define the following vector Θ, with dimension $p = 3n+2$, containing the whole number of parameters

$$\Theta = [-\theta_\beta^T\ -\theta_\tau^T\ \ \theta_\alpha^T\ ]^T\ . \tag{29}$$

In absence of noise, the system parameters can be recovered by means of the following procedure. Define the row vectors

$$Z_{n+1}(\omega_k) = [\mathrm{e}^{-jn\omega_k}\, \mathrm{e}^{-j(n-1)\omega_k}\ \ldots \mathrm{e}^{-j\omega_k}\, 1] \tag{30}$$

$$Z_n(\omega_k) = [\mathrm{e}^{-j(n-1)\omega_k}\ \ldots \mathrm{e}^{-j\omega_k}\, 1]\ , \tag{31}$$

whose entries are constructed with multiple frequencies of ω_k, and construct the following matrices

$$\Pi = \begin{bmatrix} Z_{n+1}(\omega_0) \\ \vdots \\ Z_{n+1}(\omega_{N-1}) \end{bmatrix} \qquad \Psi = \begin{bmatrix} Z_n(\omega_0) \\ \vdots \\ Z_n(\omega_{N-1}) \end{bmatrix}\ . \tag{32}$$

of dimension $N \times (n+1)$ and $N \times n$, respectively.

With the noise–free input–output DFTs (19) and (20) construct the following $N \times N$ diagonal matrices

$$\hat{V}_U^{diag} = \begin{bmatrix} \hat{U}(\omega_0) & 0 & \ldots & 0 \\ 0 & \hat{U}(\omega_1) & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \hat{U}(\omega_{N-1}) \end{bmatrix} \tag{33}$$

$$\hat{V}_Y^{diag} = \begin{bmatrix} \hat{Y}(\omega_0) & 0 & \ldots & 0 \\ 0 & \hat{Y}(\omega_1) & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \hat{Y}(\omega_{N-1}) \end{bmatrix}\ . \tag{34}$$

Compute the $N \times (n+1)$ matrices

$$\hat{\Phi}_B = \hat{V}_U^{diag}\,\Pi \quad \hat{\Phi}_A = \hat{V}_Y^{diag}\,\Pi \tag{35}$$

and set

$$\hat{\Phi}_T = \Psi\,. \tag{36}$$

Construct the $N \times p$ matrix

$$\hat{\Phi} = [\hat{\Phi}_B \,|\, \hat{\Phi}_T \,|\, \hat{\Phi}_A]\,. \tag{37}$$

Thus, eq. (24) can be rewritten as

$$\hat{\Phi}\,\Theta = 0\,. \tag{38}$$

It then holds

$$\hat{\Sigma}\,\Theta = 0\,, \tag{39}$$

where $\hat{\Sigma}$ is the $p \times p$ matrix

$$\hat{\Sigma} = \frac{1}{N}(\hat{\Phi}^H \hat{\Phi})\,. \tag{40}$$

In presence of noise, the previous procedure can be modified as follows. With the noisy input–output DFTs (13) and (14) construct the $N \times N$ diagonal matrices

$$V_U^{diag} = \begin{bmatrix} U(\omega_0) & 0 & \dots & 0 \\ 0 & U(\omega_1) & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & U(\omega_{N-1}) \end{bmatrix} \tag{41}$$

$$V_Y^{diag} = \begin{bmatrix} Y(\omega_0) & 0 & \dots & 0 \\ 0 & Y(\omega_1) & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & Y(\omega_{N-1}) \end{bmatrix}, \tag{42}$$

compute the matrices

$$\Phi_B = V_U^{diag}\,\Pi \quad \Phi_A = V_Y^{diag}\,\Pi \quad \Phi_T = \Psi \tag{43}$$

and construct the $N \times p$ matrix

$$\Phi = [\Phi_B \,|\, \Phi_T \,|\, \Phi_A]\,. \tag{44}$$

Because of Assumptions A5, when $N \to \infty$, we obtain the following $p \times p$ positive definite matrix

$$\Sigma = \frac{1}{N}(\Phi^H \Phi) = \hat{\Sigma} + \tilde{\Sigma}^*\,, \tag{45}$$

where

$$\tilde{\Sigma}^* = \begin{bmatrix} \lambda_u^*\, I_{n+1} & 0 & 0 \\ 0 & 0_n & 0 \\ 0 & 0 & \lambda_y^*\, I_{n+1} \end{bmatrix}. \tag{46}$$

From (39) and (45), the parameter vector Θ, defined in (29), can be obtained as the kernel of

$$\left(\Sigma - \tilde{\Sigma}^*\right)\Theta = 0\,. \tag{47}$$

Starting from knowledge of the noisy matrix Σ, the determination of the system parameter vector Θ and of the noise variances λ_u^*, λ_y^* can be seen as a Frisch Scheme problem (Beghelli et al., 1990), (Guidorzi et al., 2008).

Consider the set of non–negative definite diagonal matrices of type

$$\tilde{\Sigma} = \begin{bmatrix} \lambda_u\, I_{n+1} & 0 & 0 \\ 0 & 0_n & 0 \\ 0 & 0 & \lambda_y\, I_{n+1} \end{bmatrix} \tag{48}$$

such that

$$\Sigma - \tilde{\Sigma} \geq 0 \qquad \det\left(\Sigma - \tilde{\Sigma}\right) = 0\,. \tag{49}$$

With the same reasoning of (Beghelli et al., 1990), the following statements can be proved.

Theorem 1. The set of all matrices $\tilde{\Sigma}$ satisfying conditions (49) defines the points $P = (\lambda_u, \lambda_y)$ of a convex curve $\mathcal{S}(\Sigma)$ belonging to the first quadrant of the noise space $\mathcal{R}^2$ whose concavity faces the origin. At every point $P = (\lambda_u, \lambda_y)$ can be associated the noise matrix $\tilde{\Sigma}(P)$ and the coefficient vector $\Theta(P)$ satisfying the relation

$$\left(\Sigma - \tilde{\Sigma}(P)\right)\Theta(P) = 0\,. \diamond \tag{50}$$

In Figure 2 an example of $\mathcal{S}(\Sigma)$ is reported.

Theorem 2. Because of the relations (46) and (47), the point $P^* = (\lambda_u^*, \lambda_y^*)$, associated with the true variances of $\tilde{u}(t)$ and $\tilde{y}(t)$, belongs to $\mathcal{S}(\Sigma)$ and the corresponding coefficient vector $\Theta(P^*)$ is characterized (after a normalization of its last entry to 1) by the true system parameter vector, i.e. $\Theta(P^*) = \Theta$. $\diamond$

Theorem 3. Partition the matrix Σ as follows

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \tag{51}$$

where Σ_{11} is the square matrix of dimension $(n+1)$. The intersection of $\mathcal{S}(\Sigma)$ with the λ_u axis is the point $P_u = (\lambda_u^{max}, 0)$ where

$$\lambda_u^{max} = \min \operatorname{eig}\left(\Sigma_{11} - \Sigma_{12}\,\Sigma_{22}^{-1}\,\Sigma_{21}\right). \tag{52}$$

Partition the matrix Σ instead as follows

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \tag{53}$$

where Σ_{22} is the square matrix of dimension $(n+1)$. The intersection of $\mathcal{S}(\Sigma)$ with the λ_y axis is the point $P_y = (0, \lambda_y^{max})$ where

$$\lambda_y^{max} = \min \operatorname{eig}\left(\Sigma_{22} - \Sigma_{21}\,\Sigma_{11}^{-1}\,\Sigma_{12}\right). \diamond \tag{54}$$

The next theorem describes a parametrization of the curve $\mathcal{S}(\Sigma)$ that allows to associate a solution of (49) with every straight line departing from the origin and lying in the first quadrant (Guidorzi et al., 2008). This parametrization plays an important role in the practical implementation of the identification algorithm.

Theorem 4. Let $\xi = (\xi_1, \xi_2)$ be a generic point of the first quadrant of $\mathcal{R}^2$ and r the straight line from the origin through ξ. Its intersection with $\mathcal{S}(\Sigma)$ is the point $P = (\lambda_u, \lambda_y)$ given by

$$\lambda_u = \frac{\xi_1}{\lambda_M} \qquad \lambda_y = \frac{\xi_2}{\lambda_M} \tag{55}$$

where

$$\lambda_M = \max \operatorname{eig}\left(\Sigma^{-1}\tilde{\Sigma}_\xi\right) \tag{56}$$

$$\tilde{\Sigma}_\xi = \begin{bmatrix} \xi_1\, I_{n+1} & 0 & 0 \\ 0 & 0_n & 0 \\ 0 & 0 & \xi_2\, I_{n+1} \end{bmatrix}. \diamond \tag{57}$$

Remark 1. The described procedure allows to construct the curve $\mathcal{S}(\Sigma)$ in the noise space (λ_u, λ_y) also when only a subset of the whole frequency range is used, i.e. $\omega_k \in W = [\omega_i,\, \omega_f]$, with $i \geq 0$ and $f \leq N-1$, on condition that the number of frequencies $L = f - i + 1$ is large enough. The subset $W = [\omega_i,\, \omega_f]$ must be chosen by the user on the basis of *a priori* knowledge of the frequency properties of the transfer function $G(\mathrm{e}^{-j\omega_k})$ and of the noise–free input $\hat{U}(\omega_k)$.

As asserted in Theorem 2, the determination of the point P^* on $\mathcal{S}(\Sigma)$ leads to the solution of Problem 1. For this

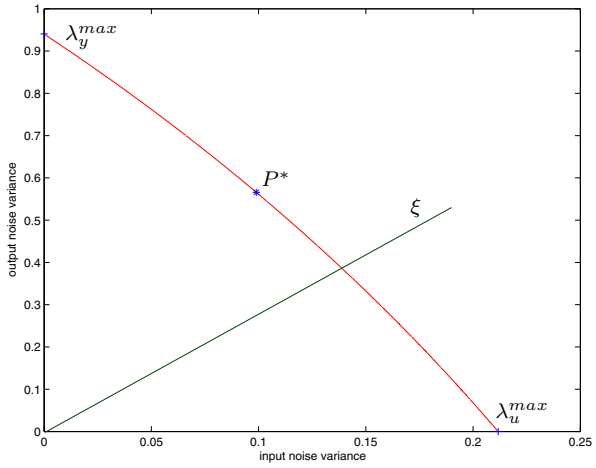

Fig. 2. Typical shape of $\mathcal{S}(\Sigma)$.

purpose a search criterion must be introduced. Unfortunately, the theoretic properties of $\mathcal{S}(\Sigma)$ described so far do not allow to distinguish point P^* from the other points of the curve.

4. A COVARIANCE–MATCHING CRITERION

In this section we will describe a possible search criterion in order to find the point P^* on $\mathcal{S}(\Sigma)$. The proposed criterion is based on considerations analogue to those reported in (Diversi et al., 2003) with reference to time domain EIV identification techniques.

Define the process

$$\begin{aligned}\Gamma(\omega_k) &= A(\mathrm{e}^{-j\omega_k})\,Y(\omega_k) - B(\mathrm{e}^{-j\omega_k})\,U(\omega_k) - T(\mathrm{e}^{-j\omega_k}) \\ &= A(\mathrm{e}^{-j\omega_k})\,\tilde{Y}(\omega_k) - B(\mathrm{e}^{-j\omega_k})\,\tilde{U}(\omega_k)\,. \end{aligned} \tag{58}$$

By introducing the diagonal matrices

$$\tilde{V}_Y^{diag} = \begin{bmatrix} \tilde{Y}(\omega_0) & 0 & \dots & 0 \\ 0 & \tilde{Y}(\omega_1) & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \tilde{Y}(\omega_{N-1}) \end{bmatrix} \tag{59}$$

$$\tilde{V}_U^{diag} = \begin{bmatrix} \tilde{U}(\omega_0) & 0 & \dots & 0 \\ 0 & \tilde{U}(\omega_1) & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \tilde{U}(\omega_{N-1}) \end{bmatrix} \tag{60}$$

it is possible to define the $N \times (n+1)$ matrices

$$\tilde{\Phi}_A = \tilde{V}_Y^{diag}\,\Pi \qquad \tilde{\Phi}_B = \tilde{V}_U^{diag}\,\Pi \tag{61}$$

so that from (58) we can write the following expressions in vector form

$$V_\Gamma = \Phi_A\,\theta_\alpha^T - \Phi_B\,\theta_\beta^T - \Phi_T\,\theta_\tau^T \tag{62}$$

$$= \tilde{\Phi}_A\,\theta_\alpha^T - \tilde{\Phi}_B\,\theta_\beta^T \tag{63}$$

where

$$V_\Gamma = [\Gamma(\omega_0)\dots\Gamma(\omega_{N-1})]^T\,. \tag{64}$$

Equation (63) is the DFT vector form of the time domain relation

$$\gamma(t) = A(z^{-1})\,\tilde{y}(t) - B(z^{-1})\,\tilde{u}(t)\,. \tag{65}$$

It results

$$V_\Gamma = F_N v_\gamma\,, \tag{66}$$

where

$$v_\gamma = [\gamma(0)\dots\gamma(N-1)]^T\,. \tag{67}$$

and F_N has been defined in (9)–(10).

From (65), the process $\gamma(t)$ is the sum of two MA processes driven by the white noises $\tilde{y}(t)$ and $\tilde{u}(t)$. Because of the assumptions on $\tilde{y}(t)$ and $\tilde{u}(t)$ the autocorrelations of $\gamma(t)$, $r_\gamma(\tau) = \mathrm{E}\,[\gamma(t)\,\gamma(t-\tau)] = r_\gamma(-\tau)$, are given by

$$r_\gamma(0) = \tilde{\lambda}_y^* \sum_{i=0}^{n} \alpha_i^2 + \tilde{\lambda}_u^* \sum_{i=0}^{n} \beta_i^2 \tag{68}$$

$$r_\gamma(\tau) = \tilde{\lambda}_y^* \sum_{i=0}^{n-\tau} \alpha_i\,\alpha_{i+\tau} + \tilde{\lambda}_u^* \sum_{i=0}^{n-\tau} \beta_i\,\beta_{i+\tau} \tag{69}$$

$$\text{for } \tau = 1,\dots,n$$

$$r_\gamma(\tau) = 0 \qquad \text{for } \tau > n\,, \tag{70}$$

where $\alpha_0 = 1$. For every point $P = (\lambda_u, \lambda_y)$ of $\mathcal{S}(\Sigma)$, define now the vector

$$r_\tau(P) = [r_\gamma(0,P)\;\; r_\gamma(1,P)\dots\;\; r_\gamma(\tau,P)]^T \tag{71}$$

whose entries are computed by means of (68)–(70) using the variances λ_u, λ_y and the parameters $\Theta(P)$.

It is also possible to compute the sample vector

$$\bar{r}_\tau(P) = [\bar{r}_\gamma(0,P)\;\; \bar{r}_\gamma(1,P)\dots\;\; \bar{r}_\gamma(\tau,P)]^T \tag{72}$$

by means of the time domain sequence $\gamma(t)$. For this purpose we can proceed as follows. Consider the Hankel matrix

$$H_\tau(\gamma) = \begin{bmatrix} \gamma(n) & \dots & \gamma(n+\tau) \\ \gamma(n+1) & \dots & \gamma(n+\tau+1) \\ \vdots & & \vdots \\ \gamma(N-\tau-1) & \dots & \gamma(N-1) \end{bmatrix}\,. \tag{73}$$

with $\tau \geq n$.

Remark 2. The value of τ is a user choice that can influence the performance of the identification algorithm. This aspect will be illustrated by means of a numerical example in Section 5.

Let us define with v_γ^i $(i = 0,\dots,\tau)$ the column vectors of $H_\tau(\gamma)$, i.e.

$$H_\tau(\gamma) = \begin{bmatrix} v_\gamma^\tau & \dots & v_\gamma^1 & v_\gamma^0 \end{bmatrix}\,. \tag{74}$$

Each vector v_γ^i has dimension $M = N - \tau - n$ and can be obtained from the N–dimensional vector v_γ (67) by a proper selection of its entries

$$v_\gamma^i = J^i\,v_\gamma \tag{75}$$

where

$$J^i = [0_{M\times(N-M-i)} \,|\, I_M \,|\, 0_{M\times i}]\,. \tag{76}$$

The sample autocorrelations $\bar{r}_\gamma(i,P)$ can thus be obtained as

$$\bar{r}_\gamma(i,P) = \frac{1}{M}(v_\gamma^0)^T\,(v_\gamma^i) \qquad i = 0,\dots,\tau\,. \tag{77}$$

In the following we will show that the autocorrelations $\bar{r}_\gamma(i,P)$ can also be computed starting from the frequency domain vector V_Γ obtained from the data by means of (62).

Let us denote with V_Γ^i the DFT vector corresponding to v_γ^i, i.e.

$$V_\Gamma^i = F_M\,v_\gamma^i \qquad i = 0,\dots,\tau \tag{78}$$

where F_M is the $M \times M$ Fourier matrix defined according to equations (9)–(10).

Since

$$V_\Gamma^i = F_M\,v_\gamma^i = F_M\,J^i\,v_\gamma = F_M\,J^i\,F_N^{-1}\,V_\Gamma\,, \tag{79}$$

it results

$$V_\Gamma^i = Z^i\,V_\Gamma \tag{80}$$

where

$$Z^i = F_M\,J^i\,F_N^{-1}\,. \tag{81}$$

Thus, recalling that

$$F_M^{-H}\, F_M^{-1} = (F_M\, F_M^H)^{-1} = I_M\ , \tag{82}$$

we have

$$\begin{aligned}\bar{r}_\gamma(i,P) &= \frac{1}{M}(v_\gamma^0)^T\,(v_\gamma^i) = \frac{1}{M}(V_\Gamma^0)^H\, F_M^{-H}\, F_M^{-1}\, V_\Gamma^i \\ &= \frac{1}{M}(V_\Gamma^0)^H\, V_\Gamma^i\ .\end{aligned} \tag{83}$$

Finally, we obtain

$$\begin{aligned}\bar{r}_\gamma(i,P) &= \frac{1}{M}(V_\Gamma^0)^H\, V_\Gamma^i = \frac{1}{M} V_\Gamma^H\,(Z^0)^H\, Z^i\, V_\Gamma \\ &= \frac{1}{M} V_\Gamma^H\, R^i\, V_\Gamma\end{aligned} \tag{84}$$

where

$$R^i = (Z^0)^H\, Z^i = F_N^{-H}\,(J^0)^T\, J^i\, F_N^{-1}\ . \tag{85}$$

Note that the matrices R^i $(i = 0, \ldots, \tau)$ can be evaluated in advance, before computing V_Γ through (62).

Since, for $N \to \infty$

$$r_\tau(P^*) = \bar{r}_\tau(P^*) = [r_\gamma(0)\ r_\gamma(1) \ldots\ r_\gamma(\tau)]^T\ , \tag{86}$$

the following covariance–matching cost function can be considered

$$J(P) = \|r_\tau(P) - \bar{r}_\tau(P)\|_2 \tag{87}$$

where the theoretical statistical properties of $\gamma(t)$ are compared with those computed from the data.

Problem 1 can thus be solved by minimizing $J(P)$ along $\mathcal{S}(\Sigma)$. In fact, the point P° associated with the minimum of $J(P)$ can be considered as an estimate of the point $P^* = (\lambda_u^*, \lambda_y^*)$. Once P^* is estimated, an estimate of the system parameter vector Θ can be obtained through (50).

It is thus possible to develop the following algorithm.

Algorithm 1.

(1) Compute, on the basis of the available data $U(\omega_k)$, $Y(\omega_k)$ with $\omega_k = 2\pi k/N$ $(k = 0, \ldots, N-1)$, the matrices Φ_A, Φ_B and Φ_T as in (43) and construct the matrix Φ as in (44).
(2) Compute, as in (45), the sample estimate of matrix

$$\Sigma = \frac{1}{N}(\Phi^H \Phi)\ . \tag{88}$$

(3) Select $\tau \geq n$ and compute matrices R^i $(i = 0, \ldots, \tau)$ with (85).
(4) Start from a generic point ξ (a generic direction) in the first quadrant of $\mathcal{R}^2$ and compute, by means of (55)–(57) the corresponding point $P = (\lambda_u, \lambda_y)$ on $\mathcal{S}(\Sigma)$.
(5) Compute the estimates of $\hat{\Sigma}(P)$ and $\Theta(P)$ by means of the relations

$$\hat{\Sigma}(P) = \Sigma - \mathrm{diag}\,[\lambda_u\, I_{n+1} \mid 0_n \mid \lambda_y\, I_{n+1}]\ , \tag{89}$$

$$\hat{\Sigma}(P)\Theta(P) = 0\ . \tag{90}$$

(6) Construct the vector $r_\tau(P)$ as in (71) through (68)–(70).
(7) Compute the frequency vector V_Γ by means of (62).
(8) Compute through (84) the entries of vector $\bar{r}_\tau(P)$, defined in (72).
(9) Compute the value of the cost function $J(P)$ (87).
(10) Search on the curve $\mathcal{S}(\Sigma)$ for the point P° associated with the minimum of $J(P)$.

Table 1. True and estimated parameters obtained with Alg.1 and the Frisch-CM algorithm

	true	Alg.1	Frisch − CM
α_1	-0.5	-0.4965 ± 0.0599	-0.4969 ± 0.0600
α_2	0.3	0.2986 ± 0.0648	0.2996 ± 0.0664
β_0	2	1.9973 ± 0.0908	1.9958 ± 0.0903
β_1	-1.2	-1.1970 ± 0.1371	-1.1970 ± 0.1372
β_2	-0.6	-0.6034 ± 0.1754	-0.6007 ± 0.1798
λ_u^*	0.1	0.0997 ± 0.0360	0.1001 ± 0.0363
λ_y^*	0.6	0.5643 ± 0.1755	0.5743 ± 0.1763

5. NUMERICAL EXAMPLES

Example 1. The proposed algorithm has been tested on sequences generated by a second–order model of type (1), already proposed in (Diversi et al., 2007)

$$A(z^{-1}) = 1 - 0.5\, z^{-1} + 0.3\, z^{-2} \tag{91}$$

$$B(z^{-1}) = 2 - 1.2\, z^{-1} - 0.6\, z^{-2}\ . \tag{92}$$

The input is a pseudo random binary sequence with unit variance and length $N = 250$. A Monte Carlo simulation of 100 independent runs have been performed by adding to the noise–free sequences $\hat{u}(\cdot)$, $\hat{y}(\cdot)$ different white noise realizations with variances $\lambda_u^* = 0.1$, $\lambda_y^* = 0.6$, corresponding to a signal to noise ratio (SNR) of about 10 dB on both input and output. Algorithm 1 has been implemented by choosing $\tau = 4$ for the number of autocorrelations.

Table 1 reports the empirical means of the system parameter estimates and of the noise variance estimates, together with the corresponding standard deviations, obtained with the Algorithm 1 and with the covariance–matching algorithm proposed in (Diversi et al., 2003), denoted with Frisch-CM.

The table shows that the described identification method yields good results, comparable with those obtained by means of the corresponding time domain version.

Example 2. As a second example, the features of the new method have been illustrated by means of the following second–order model

$$A(z^{-1}) = 1 - 1.5\, z^{-1} + 0.7\, z^{-2} \tag{93}$$

$$B(z^{-1}) = 0.54 - 0.36\, z^{-1} - 0.18\, z^{-2}\ , \tag{94}$$

characterized by a pair of complex poles much closer to the unit circle. The input is a white noise process with unit variance and length N. The performance of the proposed method has been evaluated by varying the number of the available frequencies, with $N = 65$, $N = 125$, $N = 250$ and $N = 500$. For every value of N, a Monte Carlo simulation of 100 independent runs has been performed by adding to the noise–free sequences $\hat{u}(\cdot)$, $\hat{y}(\cdot)$ different white noise realizations with variances $\lambda_u^* = 0.1$, $\lambda_y^* = 0.1$, corresponding to a SNR of about 10 dB on both input and output.

Defining the system parameters vector as

$$\theta = [\theta_\beta^T\ \ \theta_\alpha^T\,]^T\ , \tag{95}$$

the performance of the estimation algorithm has been evaluated by means of the normalized root mean square error

$$\mathrm{NRMSE} = \frac{1}{\|\theta\|}\sqrt{\frac{1}{100}\sum_{i=1}^{100}\|\hat{\theta}^i - \theta\|^2}, \tag{96}$$

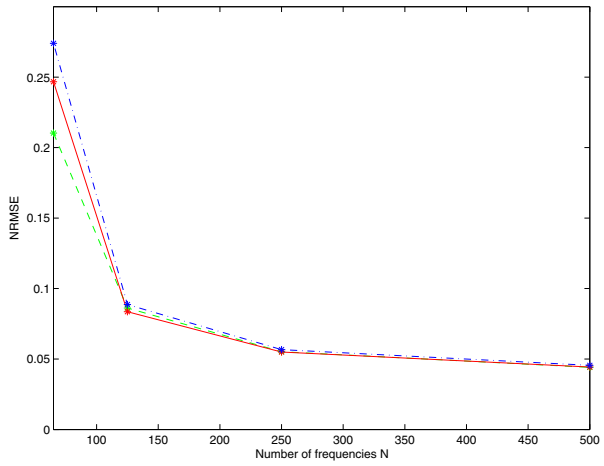

Fig. 3. NRMSE versus N for different values of τ: $\tau = 2$ green (dashed); $\tau = 4$ red (solid); $\tau = 6$ blue (dash–dotted).

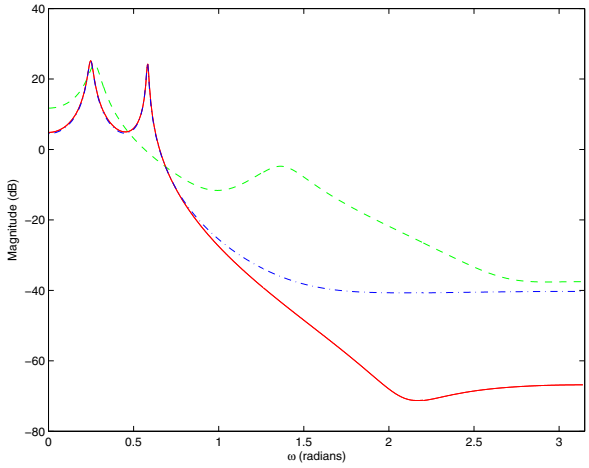

Fig. 4. True TF: red, solid; Estimated TF, $N = 200$: green, dashed; Estimated TF, $N = 2000$, $L = 200$: blue, dash–dotted.

where $\hat{\theta}^i$ denotes the estimate of θ obtained in the i–th trial of the Monte Carlo simulation.

Figure 3 reports the NRMSE versus the number of frequencies N, for different values of τ: $\tau = n$, $\tau = 2n$ and $\tau = 3n$. It can be observed that increasing the number of autocorrelations τ does not necessarily improve the quality of the parameter estimates. Moreover, the improvement is often marginal in comparison with the increase of the computational burden. From the simulation experiences, a good choice is $\tau = 2n$.

Example 3. In order to verify the selective properties described in Remark 1, the algorithm has been tested on the following fourth–order model, also considered in (Zhang et al., 2013)

$$A(z^{-1}) = 1 - 3.57\,z^{-1} + 5.13\,z^{-2} - 3.5\,z^{-3} + 0.96\,z^{-4}$$
$$B(z^{-1}) = 10^{-2} \times p(z^{-1})$$
$$p(z^{-1}) = 1 + 1.15\,z^{-1} + 1.02\,z^{-2} + 0.27\,z^{-3} + 0.05\,z^{-4}\,.$$

The noise–free input $\hat{u}(t)$ is a white noise process with unit variance and length N. Two Monte Carlo simulations of 100 independent runs have been performed by adding different white noise realizations with variances $\lambda_u^* = 0.03$, $\lambda_y^* = 0.2$, corresponding to a signal to noise ratio (SNR) of about 15 dB on both input and output.

In both cases the system has been identified under conditions of equal computational burden for Algorithm 1, by choosing $\tau = 2n$ and by using only 200 frequencies. The first Monte Carlo simulation has been performed with $N = 200$. The system transfer function $G(\mathrm{e}^{-j\omega_k})$, defined in (8), has been identified by using all the frequency data. The second Monte Carlo simulation has been performed with $N = 2000$ frequencies. However, the system transfer function has been identified by using the $L = 200$ frequencies in the window $W = [\omega_0,\, \omega_{199}]$.

Figure 4 reports the true value of $|G(\mathrm{e}^{-j\omega_k})|_{dB}$, together with the means of the two estimated transfer functions (TF). In the frequency region around the two peaks of $|G(\mathrm{e}^{-j\omega_k})|$ the advantageous effects of filtering are evident.

6. CONCLUSIONS

In this paper a new frequency domain identification method has been proposed for EIV models affected by additive white noises with unknown variances. The method does not require any assumption of periodicity for the input process. The effectiveness of the proposed method has been verified by means of Monte Carlo simulations.

REFERENCES

Agüero, J.C., Yuz, J.I., Goodwin, G.C. and Delgado, R.A. (2010). On the equivalence of time and frequency domain maximum likelihood estimation. *Automatica*, 46, 260–270.

Beghelli, S., Guidorzi, R., and Soverini, U. (1990). The Frisch scheme in dynamic system identification. *Automatica*, 26(1), 171–176.

Diversi, R., Guidorzi, R., and Soverini, U. (2003). A new criterion in EIV identification and filtering applications. *Preprints of the 13–th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, 1993–1998.

Diversi, R., Guidorzi, R., and Soverini, U. (2007). Maximum likelihood identification of noisy input–output models. *Automatica*, 43(3), 464–472.

Guidorzi, R., Diversi, R. and Soverini, U. (2008). The Frisch Scheme in algebraic and dynamic identification problems. *Kybernetika*, 44(5), 585–616.

Ljung, L. (1999). *System identification–Theory for the user*. (2nd ed.), Upper Saddle River, NJ, USA: Prentice–Hall.

McKelvey, T. (2002). Frequency domain identification methods. *Circuits Systems Signal Processing*, 21(1), 39–55.

Pintelon, R., Schoukens, J. (2012). *System identification: a frequency domain approach* (2nd ed.). NY: IEEE Press.

Pintelon, R., Schoukens, J., and Vandersteen G. (1997). Frequency domain system identification using arbitrary signals. *IEEE Transactions on Automatic Control*, 42, 1717–1720.

Söderström, T. (2007). Errors–in–variables methods in system identification. *Automatica*, 43(6), 939–958.

Söderström, T. (2012). System identification for the errors–in–variables problem *Transaction of the Institute of Measurement and Control*, 34(7), 780–792.

Soverini, U., Söderström, T. (2014). Frequency domain maximum likelihood identification of noisy input–output models. *Preprints of the 19–th IFAC World Congress*, Cape Town, South Africa.

Zhang E., Pintelon, R. and Schoukens, J. (2013). Errors–in–variables identification of dynamic systems excited by arbtrary non–white input. *Automatica*, 49(10), 3032–3041.