OPTIMAL ERRORS–IN–VARIABLES FILTERING IN THE MIMO CASE

Roberto Diversi * **Roberto Guidorzi** * **Umberto Soverini** *

* *Dipartimento di Elettronica, Informatica e Sistemistica*
Università di Bologna
Viale del Risorgimento 2, 40136, Bologna, Italy
e–mail: {rdiversi, rguidorzi, usoverini}@deis.unibo.it

Abstract: The Errors–in–Variables (EIV) stochastic environment constitutes a superset of most common stochastic environments considered, for instance, in Kalman filtering or in equation–error identification where the process input is assumed as noise–free. Errors–in–variables models assume, on the contrary, the presence of unknown additive noise also on the inputs; the associated filtering procedures concern thus the optimal (minimal variance) estimation not only of the system state and output but also of the input. Optimal EIV filtering has been formulated and solved only recently (Guidorzi *et al.*, 2003) making reference to SISO models; this paper extends the efficient algorithm proposed in (Diversi *et al.*, 2003a), based on the Cholesky factorization, to the more general multivariable case.

Keywords: Optimal filtering, linear filtering, dynamic errors–in–variables models, recursive filtering.

1. INTRODUCTION

Most stochastic environments considered in identification and filtering assume the presence of disturbances on the process output and/or state but consider the input as exactly known. This assumption can be considered as realistic in some cases, for instance in all control applications where the process input is generated by known laws, but is restrictive in many others, where measurement errors are associated to all variables.

Errors–in–Variables models are free from this restriction but require remarkably more complex identification and filtering procedures. The EIV filtering problem, i.e. the optimal (minimal variance) estimation of the true system input and output on the basis of their noisy observations and of the knowledge of the process and noise models, has been, in fact, formulated and solved only recently (Guidorzi *et al.*, 2002; Guidorzi *et al.*, 2003) making reference to polynomial and to state–space models. The computational aspects of EIV filtering have then been analyzed in (Diversi *et al.*, 2003a) with the goal of deriving a fast and robust formulation suitable for real–time implementations.

In these works, the analysis has been limited to the SISO case in order to focus the presentation only on the peculiarities of the new algorithms and avoid the introduction of the inherently more complex notation required by multivariable models. The extension to the MIMO case has been subsequently described both in stochastic contexts (Diversi *et al.*, 2003b) and in deterministic ones (Markovsky and De Moor, 2003), making reference to state–space models.

This paper considers the extension to the multivariable case of the efficient algorithm described in (Diversi *et al.*, 2003a). This algorithm, differently from previous ones, works on input–output models and does not require, at every update, the solution of a Riccati equation.

The contents are organized as follows. Section 2 defines the optimal EIV interpolation and filtering problems. Sections 3 briefly recalls the solution of optimal EIV interpolation while Section 4 proposes a solution of optimal EIV filtering on the basis of the Cholesky factorization. The expression of the expected performance of the filter is derived in Section 5 while Section 6 reports the results of a Monte Carlo simulation. Some concluding remarks are finally reported in Section 7.

2. PROBLEM STATEMENT

Consider the multi–input multi–output, linear, time–invariant system described by the difference equation

$$
\begin{aligned}
&A_{n+1}\,\hat{y}(t) + A_n\,\hat{y}(t-1) + \cdots + A_1\,\hat{y}(t-n) = \\
&\quad B_{n+1}\,\hat{u}(t) + B_n\,\hat{u}(t-1) + \cdots + B_1\,\hat{u}(t-n), \quad (1)
\end{aligned}
$$

where $\hat{y}(t) \in R^m$, $\hat{u}(t) \in R^r$ are the system output and input respectively and A_i, $B_i (i = 1, \ldots, n+1)$ are $m \times m$ and $m \times r$ coefficient matrices. Models of this type can always be used for completely observable systems (Guidorzi, 1989) and can also be written in the compact form

$$
Q(z^{-1})\,\hat{y}(t) = P(z^{-1})\,\hat{u}(t),
$$

where z^{-1} denotes the backward shift operator and

$$
\begin{aligned}
Q(z^{-1}) &= A_{n+1} + A_n\,z^{-1} + \cdots + A_1\,z^{-n}, \\
P(z^{-1}) &= B_{n+1} + B_n\,z^{-1} + \cdots + B_1\,z^{-n}.
\end{aligned}
$$

In errors–in–variables contexts the input and output observations are affected by additive noise:

$$
\begin{aligned}
u(t) &= \hat{u}(t) + \tilde{u}(t), \quad (2) \\
y(t) &= \hat{y}(t) + \tilde{y}(t), \quad (3)
\end{aligned}
$$

where $\tilde{u}(t)$ and $\tilde{y}(t)$ are assumed as zero–mean white noises, uncorrelated with $\hat{u}(t)$ and with covariances

$$
\begin{aligned}
E\left[\tilde{u}(t)\,\tilde{u}^T(t-\tau)\right] &= \tilde{\Sigma}_u\,\delta(\tau), \quad (4) \\
E\left[\tilde{y}(t)\,\tilde{y}^T(t-\tau)\right] &= \tilde{\Sigma}_y\,\delta(\tau), \quad (5) \\
E\left[\tilde{y}(t)\,\tilde{u}^T(t-\tau)\right] &= \tilde{\Sigma}_{yu}\,\delta(\tau), \quad (6)
\end{aligned}
$$

where $\delta(\tau)$ denotes the Kronecker delta function.

When a sequence of N input–output samples is available, it is possible to write equation (1) for $t = n+1, \ldots, N$ and to express this set of relations in the compact form

$$
G\,\hat{v} = 0, \quad (7)
$$

where G is the $(N-n)m \times N(m+r)$ matrix

$$
G = \left[\begin{array}{ccccccccc}
A_1 & -B_1 & A_2 & -B_2 & \cdots & A_n & -B_n & A_{n+1} & -B_{n+1} \\
0 & 0 & A_1 & -B_1 & A_2 & -B_2 & \cdots & A_n & -B_n \\
\vdots & & & & & & & & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & A_1 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots
\end{array}\right.
$$

$$
\left.\begin{array}{ccccccccc}
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & 0 \\
A_{n+1} & -B_{n+1} & 0 & \cdots & \cdots & \cdots & \cdots & & 0 \\
\vdots & & & & & & & & \vdots \\
-B_1 & A_2 & -B_2 & \cdots & A_{n+1} & -B_{n+1} & 0 & & 0 \\
0 & A_1 & -B_1 & \cdots & A_n & -B_n & A_{n+1} & & -B_{n+1}
\end{array}\right]
$$

and $\hat{v}$ the $N(m+r) \times 1$ vector

$$
\hat{v} = \left[\,\hat{y}^T(1)\ \hat{u}^T(1)\ \cdots\ \hat{y}^T(N)\ \hat{u}^T(N)\,\right]^T. \quad (8)
$$

From (2), (3) and (7) it follows that

$$
G\,v = G\,(\hat{v} + \tilde{v}) = G\,\tilde{v} = \Gamma, \quad (9)
$$

where

$$
\begin{aligned}
v &= \left[\,y^T(1)\ u^T(1)\ \cdots\ y^T(N)\ u^T(N)\,\right]^T, \quad (10) \\
\tilde{v} &= \left[\,\tilde{y}^T(1)\ \tilde{u}^T(1)\ \cdots\ \tilde{y}^T(N)\ \tilde{u}^T(N)\,\right]^T. \quad (11)
\end{aligned}
$$

The paper will treat the following problems.

Problem 1 (Optimal interpolation). Determine the minimal variance estimate of the noiseless signals (8) on the basis of the knowledge of model (1), covariance matrices (4)–(6) and the noisy observations (10).

Problem 2 (Optimal filtering). Determine, at every t, the minimal variance estimate of the current values of $\hat{u}(t), \hat{y}(t)$ on the basis of the knowledge of model (1), covariance matrices (4)–(6) and the input–ouput observations $\{u(1), y(1), \ldots, u(t), y(t)\}$.

3. OPTIMAL INTERPOLATION

If $\tilde{u}(t)$ and $\tilde{y}(t)$ are jointly gaussian processes, Problem 1 can be solved by considering the maximum likelihood estimation of $\hat{v}$ under constraint (7), as done in (Guidorzi *et al.*, 2003). An alternative approach can be based on the computation of the minimal variance estimate of $\tilde{v}$ conditioned by Γ:

$$
\tilde{v}^* = E\,[\,\tilde{v}\,|\,\Gamma\,] = E\,\left[\,\tilde{v}\,\Gamma^T\,\right]\ E\,\left[\,\Gamma\,\Gamma^T\,\right]^{-1}\,\Gamma\,.
$$

In fact, since $v = \hat{v} + \tilde{v}$ and $E\,[\tilde{v}\,\hat{v}^T] = 0$, we have

$$
\tilde{v}^* = \tilde{\Sigma}\,G^T\big(G\,\tilde{\Sigma}\,G^T\big)^{-1}\,\Gamma = \tilde{\Sigma}\,G^T\big(G\,\tilde{\Sigma}\,G^T\big)^{-1}G\,v, \quad (12)
$$

where

$$
\tilde{\Sigma} = E\,\left[\,\tilde{v}\,\tilde{v}^T\right] = \begin{bmatrix}
\tilde{\Sigma}_y & \tilde{\Sigma}_{yu} & \ldots & 0 & 0 \\
\tilde{\Sigma}_{yu} & \tilde{\Sigma}_u & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & \tilde{\Sigma}_y & \tilde{\Sigma}_{yu} \\
0 & 0 & \ldots & \tilde{\Sigma}_{yu} & \tilde{\Sigma}_u
\end{bmatrix},
$$

so that

$$\hat{v}^* = v - \tilde{v}^* = \Big[I - \tilde{\Sigma}\, G^T \big(G\,\tilde{\Sigma}\,G^T\big)^{-1} G \Big]\, v. \quad (13)$$

Remark 1. When $\tilde{u}(t)$ and $\tilde{y}(t)$ are not gaussian, expression (13) constitutes the best (minimal variance) linear unbiased estimate of $\hat{v}$ that can be obtained from Γ, i.e. from the noisy observations v under condition (7).

Remark 2. The covariance matrix of the estimation error $e = \hat{v} - \hat{v}^*$ is given, as shown in (Guidorzi *et al.*, 2003), by

$$E\,[e\,e^T] = \tilde{\Sigma}\left(I - G^T (G\tilde{\Sigma}G^T)^{-1} G\tilde{\Sigma} \right).$$

4. OPTIMAL FILTERING

The solution of Problem 1 does not require the introduction of the time t because interpolation is a batch procedure based on all available samples. Since, on the contrary, filtering is an on–line procedure working with increasing–time sequences $\{u(1), y(1), \ldots, u(t), y(t)\}$, it is convenient to replace N with t in all definitions and to rewrite (12) and (13) as

$$\begin{aligned} \tilde{v}^*(t) &= E\,\left[\,\tilde{v}(t)\,|\,\Gamma(t)\,\right] & (14)\\ &= \tilde{\Sigma}(t)\,G^T(t)\big[G(t)\,\tilde{\Sigma}(t)\,G^T(t)\big]^{-1}\,\Gamma(t), \\ \hat{v}^*(t) &= v(t) - \tilde{v}^*(t). & (15) \end{aligned}$$

It is now possible to observe that in the whole interpolated sequence (13) the last two terms $\hat{u}^*(t), \hat{y}^*(t)$ are, actually, filtered:

$$\hat{u}^*(t) = u(t) - \tilde{u}^*(t) = u(t) - E\,\left[\,\tilde{u}(t)\,|\,\Gamma(t)\,\right], \quad (16)$$
$$\hat{y}^*(t) = y(t) - \tilde{y}^*(t) = y(t) - E\,\left[\,\tilde{y}(t)\,|\,\Gamma(t)\,\right]. \quad (17)$$

A possible way to solve Problem 2 could thus consist in computing or updating (15) as t increases. A procedure of this kind would however exhibit poor efficiency because it does not rely on previous computations and requires a computational load that increases at every step with the dimensions of $G(t)$. These problems can be overcome by developing a finite–memory recursive filtering algorithm or a state–space solution based on the realization of $\Gamma(t)$ (Guidorzi *et al.*, 2003). A more efficient solution can, however, be based on Cholesky factorization and Bauer's algorithm as described in (Diversi *et al.*, 2003a) for the SISO case. To extend this approach to the MIMO case, note that the entries of $\Gamma(t)$ are the samples of a m–dimensional stochastic process $\gamma(t)$:

$$\begin{aligned} \Gamma(t) &= G(t)\,v(t) & (18)\\ &= \left[\,\gamma^T(n+1)\;\gamma^T(n+2)\;\cdots\;\gamma^T(t)\,\right]^T, \end{aligned}$$

where $\gamma(t)$ is the sum of two moving average processes driven by the white noises $\tilde{y}(t), \tilde{u}(t)$. In fact

$$\begin{aligned} \gamma(t) &= Q(z^{-1})\,y(t) - P(z^{-1})\,u(t) \\ &= Q(z^{-1})\,\tilde{y}(t) - P(z^{-1})\,\tilde{u}(t) & (19)\\ &= A_{n+1}\,\tilde{y}(t) + A_n\,\tilde{y}(t-1) + \cdots + A_1\,\tilde{y}(t-n) \\ &\quad - B_{n+1}\,\tilde{u}(t) - B_n\,\tilde{u}(t-1) - \cdots - B_1\,\tilde{u}(t-n). \end{aligned}$$

The autocorrelations $R_\gamma(k) = E\left[\gamma(t)\,\gamma^T(t-k)\right]$ are given by

$$\begin{aligned} R_\gamma(0) &= \sum_{i=1}^{n+1} \Big(A_i\,\tilde{\Sigma}_y\,A_i^T + B_i\,\tilde{\Sigma}_u\,B_i^T \\ &\qquad - A_i\,\tilde{\Sigma}_{yu}\,B_i^T - B_i\,\tilde{\Sigma}_{yu}^T\,A_i^T \Big), & (20)\\ R_\gamma(k) &= \sum_{i=1}^{n-k+1} \Big(A_i\,\tilde{\Sigma}_y\,A_{i+k}^T + B_i\,\tilde{\Sigma}_u\,B_{i+k}^T \\ &\qquad - A_i\,\tilde{\Sigma}_{yu}\,B_{i+k}^T - B_i\,\tilde{\Sigma}_{yu}^T\,A_{i+k}^T \Big),\ 1 \le k \le n & (21)\\ R_\gamma(k) &= 0, \qquad \text{for } k > n. & (22) \end{aligned}$$

The optimal estimate $\tilde{u}^*(t)$ can be rewritten as

$$\tilde{u}^*(t) = E\left[\tilde{u}(t)\,|\,\gamma(t), \gamma(t-1), \ldots, \gamma(n+1)\right]$$

or, equivalently,

$$\begin{aligned} \tilde{u}^*(t) &= E\left[\tilde{u}(t)\,|\,\varepsilon(t), \varepsilon(t-1), \ldots, \varepsilon(n+1)\right] \\ &= E\left[\tilde{u}(t)\,|\,\varepsilon(t)\right] \\ &= E\left[\tilde{u}(t)\,\varepsilon^T(t)\right]\,E\left[\varepsilon(t)\,\varepsilon^T(t)\right]^{-1}\varepsilon(t), & (23) \end{aligned}$$

where $\varepsilon(t)$ is the innovation process of $\gamma(t)$ (Anderson and Moore, 1979; Caines, 1988). A similar result holds for the optimal estimate $\tilde{y}^*(t)$

$$\begin{aligned} \tilde{y}^*(t) &= E\left[\tilde{y}(t)\,|\,\varepsilon(t), \varepsilon(t-1), \ldots, \varepsilon(n+1)\right] \\ &= E\left[\tilde{y}(t)\,|\,\varepsilon(t)\right] \\ &= E\left[\tilde{y}(t)\,\varepsilon^T(t)\right]\,E\left[\varepsilon(t)\,\varepsilon^T(t)\right]^{-1}\varepsilon(t). & (24) \end{aligned}$$

The evaluation of (23) and (24) can rely on an approach similar to that proposed by Rissanen and Barbosa (Rissanen and Barbosa, 1969; Caines, 1988). In fact, the covariance matrix of $\Gamma(t)$ exhibits a band Toeplitz structure

$$E\,[\,\Gamma(t)\,\Gamma(t)^T] = \Sigma_\Gamma(t) = G(t)\tilde{\Sigma}(t)\,G^T(t) =$$

$$\begin{bmatrix} R_\gamma(0) & R_\gamma^T(1) & \cdots & R_\gamma^T(n) & 0 & \cdots & \cdots & & \cdots & 0 \\ R_\gamma(1) & R_\gamma(0) & \cdots & R_\gamma^T(n-1) & R_\gamma^T(n) & \cdots & \cdots & & \cdots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \ddots & & & \vdots \\ R_\gamma(n) & R_\gamma(n-1) & \cdots & R_\gamma(0) & R_\gamma^T(1) & \cdots & \cdots & R_\gamma^T(n) & 0 \; \cdots & 0 \\ 0 & & \ddots & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & & \vdots & \cdots & \cdots & R_\gamma(n) & \cdots \; \cdots & R_\gamma(0) \end{bmatrix}$$

and admits, thus, the Cholesky factorization

$$\Sigma_\Gamma(t) = L_\Gamma(t) L_\Gamma^T(t),$$

$$L_\Gamma(t) = \left[\begin{array}{cccccccc} L(1,1) & 0 & \dots & 0 & 0 & \dots & 0 \\ L(2,1) & L(2,2) & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ L(n+1,1) & L(n+1,2) & \dots & L(n+1,n+1) & 0 & \dots & 0 \\ 0 & L(n+2,2) & \dots & L(n+2,n+1) & L(n+2,n+2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & L(t,t-n) & L(t,t-n+1) & \dots & L(t,t) \end{array}\right]. \tag{25}$$

The elements of $L_\Gamma(t)$ can then be computed by means of Bauer's algorithm

$$L(t,t)\, L^T(t,t) = R_\gamma(0) - \sum_{\tau=t-n}^{\tau=t-1} L(t,\tau)\, L^T(t,\tau), \tag{26}$$

$$L(t,t-k) = \left(R_\gamma(k) - \sum_{\tau=t-n}^{\tau=t-k-1} L(t,\tau)\, L^T(t-k,\tau) \right) \times L^{-T}(t-k,t-k), \tag{27}$$

where $k = 1,\dots,n$, $L(1,1)\, L^T(1,1) = R_\gamma(0)$ and L^{-T} is a short notation for $(L^{-1})^T$. The innovation representation of $\gamma(t)$ is given by the time–varying model

$$\gamma(t) = L(t,t)\,\varepsilon(t) + L(t,t-1)\,\varepsilon(t-1) + \dots + L(t,t-n)\,\varepsilon(t-n). \tag{28}$$

By comparing the expressions of $E\,[\tilde{u}(t)\,\gamma^T(t)]$ computed by means of (19) and (28) it is possible to write the relation

$$E\,[\tilde{u}(t)\,\varepsilon^T(t)] = \left(\tilde{\Sigma}_{yu}^T\, A_{n+1}^T - \tilde{\Sigma}_u\, B_{n+1}^T\right) L^{-T}(t,t).$$

Finally, by recalling that $E\,[\varepsilon(t)\,\varepsilon^T(t)] = I$, relation (23) leads to

$$\tilde{u}^*(t) = \left(\tilde{\Sigma}_{yu}^T\, A_{n+1}^T - \tilde{\Sigma}_u\, B_{n+1}^T\right) L^{-T}(t,t)\,\varepsilon(t). \tag{29}$$

The same considerations can be repeated for $\tilde{y}(t)$ in order to derive the expression

$$\tilde{y}^*(t) = \left(\tilde{\Sigma}_{y}\, A_{n+1}^T - \tilde{\Sigma}_{yu}\, B_{n+1}^T\right) L^{-T}(t,t)\,\varepsilon(t). \tag{30}$$

The minimal variance estimates of $\hat{u}(t)$ and $\hat{y}(t)$ can then be computed by means of (16), (17)

$$\hat{u}^*(t) = u(t) - \left(\tilde{\Sigma}_{yu}^T\, A_{n+1}^T - \tilde{\Sigma}_u\, B_{n+1}^T\right) \times L^{-T}(t,t)\,\varepsilon(t), \tag{31}$$

$$\hat{y}^*(t) = y(t) - \left(\tilde{\Sigma}_{y}\, A_{n+1}^T - \tilde{\Sigma}_{yu}\, B_{n+1}^T\right) \times L^{-T}(t,t)\,\varepsilon(t). \tag{32}$$

The following optimal filtering algorithm can thus be formulated.

Algorithm 1.

(1) Start at time $t = n+1$ by computing $R_\gamma(0),\dots,R_\gamma(n)$ (20)–(22).
(2) Compute the terms $L(t,t-n),\dots,L(t,t-1), L(t,t)$ by means of (26)–(27).
(3) Compute $\gamma(t)$:

$$\gamma(t) = A_{n+1}\, y(t) + \dots + A_1\, y(t-n) - B_{n+1}\, u(t) - \dots - B_1\, u(t-n).$$

(4) Compute the innovation $\varepsilon(t)$:

$$\varepsilon(t) = L^{-1}(t,t) \Big(\gamma(t) - L(t,t-1)\,\varepsilon(t-1) - \dots - L(t,t-n)\,\varepsilon(t-n) \Big).$$

(5) Compute the optimal estimates $\hat{u}^*(t), \hat{y}^*(t)$ by means of (31), (32).
(6) Set $t \leftarrow t+1$ and return to step 2.

5. COVARIANCE OF THE ESTIMATION ERRORS

The purpose of this section is to derive the expected performance of the filter, i.e to determine the expression of the covariance matrices of the estimation errors

$$e_u(t) = \hat{u}(t) - \hat{u}^*(t) = \left(\tilde{\Sigma}_{yu}^T\, A_{n+1}^T - \tilde{\Sigma}_u\, B_{n+1}^T\right)\, L^{-T}(t,t)\,\varepsilon(t) - \tilde{u}(t),$$

$$e_y(t) = \hat{y}(t) - \hat{y}^*(t) = \left(\tilde{\Sigma}_{y}\, A_{n+1}^T - \tilde{\Sigma}_{yu}\, B_{n+1}^T\right)\, L^{-T}(t,t)\,\varepsilon(t) - \tilde{y}(t).$$

Since

$$E\left[\tilde{u}(t)\,\varepsilon^T(t)\right] = \left(\tilde{\Sigma}_{yu}^T\, A_{n+1}^T - \tilde{\Sigma}_u\, B_{n+1}^T\right) L^{-T}(t,t),$$

$$E\left[\tilde{y}(t)\,\varepsilon^T(t)\right] = \left(\tilde{\Sigma}_{y}\, A_{n+1}^T - \tilde{\Sigma}_{yu}\, B_{n+1}^T\right) L^{-T}(t,t),$$

it is easy to obtain the relations

$$P_u(t) = E\left[e_u(t)\, e_u^T(t)\right] = \tilde{\Sigma}_u - \left(\tilde{\Sigma}_{yu}^T\, A_{n+1}^T - \tilde{\Sigma}_u\, B_{n+1}^T\right)\, L^{-T}(t,t) \times L^{-1}(t,t)\, \left(A_{n+1}\,\tilde{\Sigma}_{yu} - B_{n+1}\,\tilde{\Sigma}_u\right),$$

$$P_y(t) = E\left[e_y(t)\, e_y^T(t)\right]$$

$$= \tilde{\Sigma}_y - \left(\tilde{\Sigma}_y \, A_{n+1}^T - \tilde{\Sigma}_{yu} \, B_{n+1}^T \right) \, L^{-T}(t,t)$$
$$\times \, L^{-1}(t,t) \, \left(A_{n+1} \, \tilde{\Sigma}_y - B_{n+1} \, \tilde{\Sigma}_{yu}^T \right).$$

From (20)–(22) it follows that $\gamma(t)$ has the polynomial spectrum

$$\Pi(z) = \sum_{k=-n}^{k=n} R_\gamma(k) \, z^{-k},$$

where z denotes a complex variable. In (Rissanen and Barbosa, 1969) it has been proved that the elements of the t-th row of matrix (25) converge for $t \to \infty$ to stationary terms

$$\lim_{t \to \infty} L(t, t-k) = L_k$$

if and only if $\Sigma_\Gamma(t) > 0$ for every t; under this condition the associated polynomial

$$L(z) = L_0 + L_1 \, z^{-1} + \cdots + L_n \, z^{-n}$$

is the minimum–phase spectral factor of $\Pi(z)$, i.e.

$$\Pi(z) = L(z) \, L^T(z^{-1}),$$

where $L(z)$ is asymptotically stable. It follows that

$$\begin{aligned} P_u &= \lim_{t \to \infty} P_u(t) \\ &= \tilde{\Sigma}_u - \left(\tilde{\Sigma}_{yu}^T \, A_{n+1}^T - \tilde{\Sigma}_u \, B_{n+1}^T \right) \, L_0^{-T} \\ &\times L_0^{-1} \, \left(A_{n+1} \, \tilde{\Sigma}_{yu} - B_{n+1} \, \tilde{\Sigma}_u \right), \qquad (33) \\ P_y &= \lim_{t \to \infty} P_y(t) \\ &= \tilde{\Sigma}_y - \left(\tilde{\Sigma}_y \, A_{n+1}^T - \tilde{\Sigma}_{yu} \, B_{n+1}^T \right) \, L_0^{-T} \\ &\times L_0^{-1} \, \left(A_{n+1} \, \tilde{\Sigma}_y - B_{n+1} \, \tilde{\Sigma}_{yu}^T \right). \qquad (34) \end{aligned}$$

6. SIMULATION RESULTS

A 200 runs Monte Carlo simulation has been performed on a two–input two–output system with coefficient matrices

$$A_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -0.4 & 1 \\ -0.2 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & -0.4 \end{bmatrix},$$
$$B_3 = \begin{bmatrix} 1.2 & 0 \\ 0 & 0.8 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0.8 \\ 0.5 & -1 \end{bmatrix}, \quad B_1 = \begin{bmatrix} -0.1 & 1 \\ 0.5 & 0.5 \end{bmatrix}.$$

The inputs are pseudo random binary sequences with unit variance and length $N = 2000$. In every run the noiseless signals $\hat{u}(\cdot)$, $\hat{y}(\cdot)$ have been corrupted by adding gaussian white noise sequences $\tilde{u}(\cdot)$, $\tilde{y}(\cdot)$ generated with the function *randn* in MATLAB and characterized by the following covariance and cross–covariance matrices

$$\tilde{\Sigma}_u = \begin{bmatrix} 0.12 & 0.15 \\ 0.15 & 0.25 \end{bmatrix}, \; \tilde{\Sigma}_y = \begin{bmatrix} 1.3 & 1.5 \\ 1.5 & 2.1 \end{bmatrix}, \; \tilde{\Sigma}_{yu} = \begin{bmatrix} 0.3 & 0.5 \\ 0.4 & 0.7 \end{bmatrix}.$$

Figures 1 and 2 report, as an example, the last 100 samples of the second noiseless input and output (continuous line) and their observations (dotted line) in a typical run of the Monte Carlo simulation.

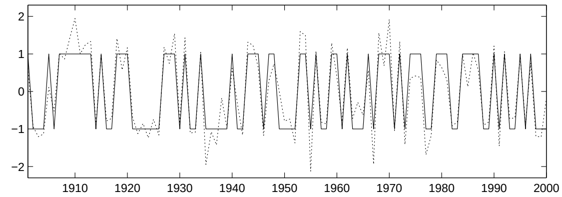

Fig. 1. Comparison between the second noiseless input (continuous line) and its observation (dotted line).

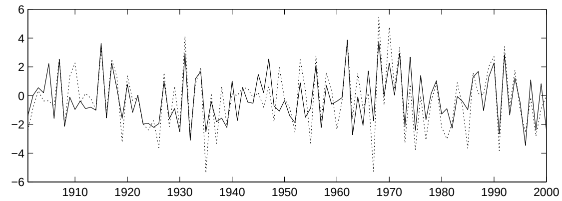

Fig. 2. Comparison between the second noiseless output (continuous line) and its observation (dotted line).

The effectiveness of the filter can be observed, in the same typical case, in Figures 3 and 4, where the second noiseless input and output (continuous line) are compared with the corresponding sequences filtered by means of Algorithm1 (dotted line).

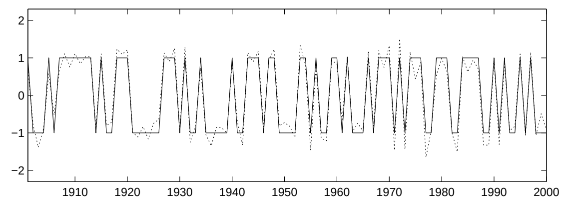

Fig. 3. Comparison between the second noiseless input (continuous line) and its optimal estimation (dotted line).

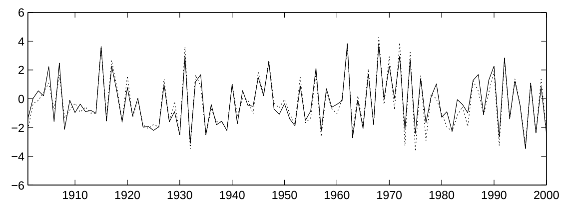

Fig. 4. Comparison between the second noiseless output (continuous line) and its optimal estimation (dotted line).

The asymptotic covariance matrices of the estimation errors, computed by means of (33) and (34) are

$$P_u = \begin{bmatrix} 0.0592 & 0.0440 \\ 0.0440 & 0.0602 \end{bmatrix},$$
$$P_y = \begin{bmatrix} 0.2012 & 0.1798 \\ 0.1798 & 0.2934 \end{bmatrix},$$

while the mean values obtained in the 200 runs and the associated standard deviations are

$$\bar{P}_u = \begin{bmatrix} 0.0591 \pm 0.0014 & 0.0440 \pm 0.0020 \\ 0.0440 \pm 0.0020 & 0.0603 \pm 0.0029 \end{bmatrix},$$
$$\bar{P}_y = \begin{bmatrix} 0.2015 \pm 0.0090 & 0.1800 \pm 0.0129 \\ 0.1800 \pm 0.0129 & 0.2939 \pm 0.0195 \end{bmatrix}.$$

The theoretical results are thus in complete agreement with the numerical simulation.

7. CONCLUSIONS

This paper has described a solution of the errors–in–variables filtering problem for multi–input multi–output processes. The efficient algorithm proposed in (Diversi *et al.*, 2003a) for SISO models, has been extended to the more general multivariable case. The Monte Carlo simulation that has been performed shows an excellent agreement between the expected performance of the filter and the observed one.

REFERENCES

Anderson, B.D.O. and J.B. Moore (1979). *Optimal Filtering*. Prentice–Hall, Englewood Cliffs, New Jersey.

Caines, P.E. (1988). *Linear Stochastic Systems*. Wiley.

Diversi, R., R. Guidorzi and U. Soverini (2003a). Algorithms for optimal errors–in–variables filtering. *Systems & Control Letters*, 48, 1–13.

Diversi, R., R. Guidorzi and U. Soverini (2003b). Kalman filtering in symmetrical noise environments. *Proceedings of the 11th IEEE Mediterranean Conference on Control and Automation*, Rhodes, Greece.

Guidorzi, R. (1989). Equivalence, invariance and dynamical system canonical modelling–Part I. *Kybernetika*, 25, 233–257.

Guidorzi R., R. Diversi and U. Soverini, (2002). Errors–in–variables filtering in behavioural and state–space contexts. In: *Total Least Squares and Errors–in–Variables Modelling: Analysis, Algorithms and Applications* (S. Van Huffel and P. Lemmerling (Eds.)), pp. 281–291. Kluwer Academic Publishers, Dordrecht.

Guidorzi, R., R. Diversi and U. Soverini (2003). Optimal errors–in–variables filtering. *Automatica*, 39, 281–289.

Markovsky, I. and B. De Moor (2003). Linear dynamic filtering with noisy input and output. *Preprints of the 13th IFAC Symposium on System Identification*, pp. 1749–1754, Rotterdam, The Netherlands.

Rissanen J. and L. Barbosa (1969). Properties of infinite covariance matrices and stability of optimum predictors. *Information Sciences*, 1, 221–236.