[5] C. I. Byrnes and A. Isidori, "Limit sets, zero dynamics, and internal models in the problem of nonlinear output regulation," *IEEE Trans. Autom. Control*, vol. 48, no. 10, pp. 1712–1723, Oct. 2003.
[6] R. De Santis and A. Isidori, "On the output regulation for linear systems in the presence of input saturation," *IEEE Trans. Autom. Control*, vol. 46, no. 1, pp. 156–160, Jan. 2001.
[7] F. D. Priscoli, "Robust tracking for polynomial plants," in *Proc. 2nd Eur. Control Conf.*, Groningen, The Netherlands, 1993, pp. 369–373.
[8] B. A. Francis and W. M. Wonham, "The internal model principle of control theory," *Automatica*, vol. 12, pp. 457–465, 1976.
[9] J. Huang, "K-fold exosystem and the robust servomechanism problem," *ASME J. Dyna. Syst., Meas., Control*, pp. 149–153, 1998.
[10] ——, "Remarks on the robust output regulation problem for nonlinear systems," *IEEE Trans. Autom. Control*, vol. 46, no. 12, pp. 2028–2031, Dec. 2001.
[11] ——, *Nonlinear Output Regulation: Theory and Applications*. Philadelphia, PA: SIAM, 2004.
[12] J. Huang and Z. Chen, "A general framework for tackling the output regulation problem," *IEEE Trans. Autom. Control*, vol. 49, no. 12, pp. 2203–2218, Dec. 2004.
[13] J. Huang and C.-F. Lin, "On a robust servomechanism problem," *IEEE Trans. Autom. Control*, vol. 39, no. 7, pp. 1510–1513, Jul. 1994.
[14] A. Isidori, *Nonlinear Control Systems*, 3rd ed. New York: Springer-Verlag, 1995.
[15] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 35, no. 2, pp. 131–140, Feb. 1990.
[16] Z.-P. Jiang and I. Kanellakopoulos, "Global output-feedback tracking for a benchmark nonlinear system," *IEEE Trans. Autom. Control*, vol. 45, no. 5, pp. 1023–1027, May 2000.
[17] G. Kaliora and A. Astolfi, "Nonlinear control of feedforward systems with bounded signals," *IEEE Trans. Autom. Control*, vol. 49, no. 11, pp. 1975–1990, Nov. 2004.
[18] H. K. Khalil, "Robust servomechanism output feedback controllers for feedback linearizable systems," *Automatica*, vol. 30, pp. 1587–1599, 1994.
[19] L. Marconi and A. Isidori, "Robust global stabilization of a class of uncertain feedforward nonlinear systems," *Syst. Control Lett.*, vol. 41, no. 4, pp. 281–290, 2000.
[20] L. Marconi, A. Isidori, and A. Serrani, "Input disturbance suppression for a class of feedforward uncertain nonlinear systems," *Syst. Control Lett.*, vol. 45, pp. 227–236, 2002.
[21] F. Mazenc and L. Praly, "Adding integrations, saturated controls, and stabilization for feedforward systems," *IEEE Trans. Autom. Control*, vol. 41, no. 11, pp. 1559–1578, Nov. 1996.
[22] F. Mazenc and J. C. Vivalda, "Global asymptotic output feedback stabilization of feedforward systems," in *Proc. Eur. Control Conf.*, 2001, pp. 2564–2569.
[23] R. Ortega, A. Loria, P. J. Nicklasson, and H. Sira-Ramirez, *Passivity-Based Control of Euler-Lagrange Systems*. New York: Springer-Verlag, 1998.
[24] A. Rapaport and A. Astolfi, "A remark on the stability of interconnected nonlinear systems," *IEEE Trans. Autom. Control*, vol. 49, no. 1, pp. 120–124, Jan. 2004.
[25] R. Sepulchre, M. Jankovic, and P. V. Kokotović, *Constructive Nonlinear Control*. New York: Springer-Verlag, 1996.
[26] A. Serrani, A. Isidori, and L. Marconi, "Semiglobal nonlinear output regulation with adaptive internal model," *IEEE Trans. Autom. Control*, vol. 46, no. 8, pp. 1178–1194, Aug. 2001.
[27] A. R. Teel, "A nonlinear small gain theorem for the analysis of control systems with saturation," *IEEE Trans. Autom. Control*, vol. 11, no. 9, pp. 1256–1270, Sep. 1996.
[28] P. Tsiotras, M. Corless, and M. A. Rotea, "An L_2 disturbance attenuation approach to the nonlinear benchmark problem," in *Proc. Amer. Control Conf.*, vol. 6, 1995, pp. 4352–4367.
[29] G. Weiss and M. Haefele, "Repetitive control of MIMO systems using H_∞ design," *Automatica*, pp. 1185–1199, 1999.

Kalman Filtering in Extended Noise Environments

Roberto Diversi, Roberto Guidorzi, and Umberto Soverini

***Abstract*—This note introduces an extended environment for Kalman filtering that considers also the presence of additive noise on input observations in order to solve the problem of optimal (minimal variance) estimation of noise-corrupted input and output sequences. This environment includes as subcases both errors-in-variables filtering (optimal estimate of inputs and outputs from noisy observations) and traditional Kalman filtering (optimal estimate of state and output in presence of state and output noise). A Monte Carlo simulation shows that the performance of this extended filtering technique leads to the expected minimal variance estimates.**

***Index Terms*—Errors-in-variables filtering, Kalman filtering, optimal filtering, recursive filtering.**

I. Introduction

Very few algorithms, if any, have seen so many applications in different areas as Kalman filtering that constitutes the *de facto* standard for information retrieval from noisy data generated by known processes and affected by noise with known statistical properties. This wide success could be considered somehow surprising in front of the relatively simple stochastic environment considered by this approach and of the requirement of information, usually unknown, on noise statistics. These drawbacks are, in fact, overcome by the robustness of the filter versus modeling errors and by the possibility of monitoring its performance and of deducing from the innovation of nonoptimal filters some information about the unknown noise properties [1]–[4].

A limit of the stochastic environment of Kalman filtering concerns its asymmetrical description of the uncertainties on the observations; in fact, while the output is considered as affected by additive noise, the input is assumed as exactly known. This condition is met in all control applications where the process input is generated by known laws but can be restrictive in other contexts.

Symmetrical environments are, instead, at the basis of errors-in-variables (EIV) models that consider all system attributes as affected by unknown additive disturbances. This symmetry allows, in many cases, to avoid system orientation, i.e., the necessity of partitioning observations into inputs and outputs [5].

Kalman filtering cannot be directly applied, as discussed in [6], to EIV filtering problems, i.e., to the optimal reconstruction of inputs and outputs of EIV models on the basis of their noise-corrupted observations. The problem of EIV filtering has been recently solved in [6] and [7] by making reference to both behavioral and state-space contexts, starting from the solution of optimal EIV interpolation. The numerical aspects of this problem have then been investigated in [8] where an high-efficiency algorithm, based on the properties of Cholesky factorization, has been described.

Optimal EIV filtering can be approached also in a deterministic context, as an optimization problem along the line followed by Roorda and Heij [9], as described in [6]. An approach of this kind has been recently used also in [10].

Manuscript received May 6, 2003; revised April 27, 2005. Recommended by Associate Editor R. Boel.

The authors are with the Department of Electronics, Computer Science and Systems, the University of Bologna, 40136 Bologna, Italy (e-mail: rdiversi@deis.unibo.it; rguidorzi@deis.unibo.it; usoverini@deis.unibo.it).

Digital Object Identifier 10.1109/TAC.2005.854627

A unified context for both Kalman and EIV filtering has subsequently been developed in [11] by extending Kalman filtering to the more general case of symmetrical noise environments in order to include, as particular cases, both Kalman and EIV filtering. This note completes the results described in [11] and considers also time-varying processes and the possibility of mutual noise correlations.

The note is organized as follows. Section II is dedicated to a statement of the extended filtering problem while Section III considers its solution. The evaluation of the expected performance of the filter is then considered in Section IV. The results of a Monte Carlo simulation are reported in Section V and some concluding remarks are finally reported in Section VI.

II. Statement of the Problem

The models considered in this note are described by the state–space equations

$$x(t+1) = A(t)x(t) + B(t)\hat{u}(t) + G(t)w(t) \quad x(0) = x_0 \tag{1}$$

$$\hat{y}(t) = C(t)x(t) + D(t)\hat{u}(t) \tag{2}$$

where $x(t) \in \mathcal{R}^n$, $\hat{u}(t) \in \mathcal{R}^r$, and $\hat{y}(t) \in \mathcal{R}^m$ denote the state, input, and output processes while $w(t) \in \mathcal{R}^p$ is the noise acting on the state. The true input and output are unknown and only the noisy observations

$$u(t) = \hat{u}(t) + \tilde{u}(t) \tag{3}$$

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

are available, where $\tilde{u}(t)$ and $\tilde{y}(t)$ denote the additive noise on $\hat{u}(t)$ and $\hat{y}(t)$. We will assume, in the sequel, that $w(t), \tilde{u}(t)$ and $\tilde{y}(t)$ are zero mean white processes, uncorrelated with $\hat{u}(t)$ and with covariances

$$E[w(t)w^T(t-\tau)] = \Sigma_w(t)\delta(\tau) \tag{5}$$

$$E[\tilde{u}(t)\tilde{u}^T(t-\tau)] = \tilde{\Sigma}_u(t)\delta(\tau) \tag{6}$$

$$E[\tilde{y}(t)\tilde{y}^T(t-\tau)] = \tilde{\Sigma}_y(t)\delta(\tau) \tag{7}$$

$$E[\tilde{u}(t)\tilde{y}^T(t-\tau)] = \tilde{\Sigma}_{uy}(t)\delta(\tau) \tag{8}$$

$$E[w(t)\tilde{u}^T(t-\tau)] = 0 \qquad \forall\tau \tag{9}$$

$$E[w(t)\tilde{y}^T(t-\tau)] = 0 \qquad \forall\tau \tag{10}$$

where $\delta(\tau)$ denotes the Kronecker delta function. The initial state x_0 is a random vector with mean $\bar{x}_0$ and covariance matrix P_0, uncorrelated with $w(t), \tilde{u}(t)$, and $\tilde{y}(t), \forall t$.

Remark 1: Assumptions (9) and (10), concerning the uncorrelation between the state noise $w(t)$ and the measurement noise $\tilde{u}(t), \tilde{y}(t)$, have been introduced only for the sake of simplicity and can be easily removed.

The optimal filtering problem can be defined as follows.

Problem 1: Given model (1)–(4), covariance matrices (5)–(10) and the input-ouput observations $\{u(0), y(0), \ldots, u(t), y(t)\}$, determine, at every t, the optimal (minimal variance) linear estimates of $\hat{u}(t)$ and $\hat{y}(t)$.

III. Optimal Filtering

The solution of Problem 1 is based on the following theorem, whose proof is reported in the Appendix.

Theorem 1: The optimal estimates $\hat{u}^*(t), \hat{y}^*(t)$ of $\hat{u}(t), \hat{y}(t)$ that can be obtained from $\{u(0), y(0), \ldots, u(t), y(t)\}$, under constraints (1)–(4) are given by

$$\hat{u}^*(t) = u(t) - \tilde{u}^*(t) = u(t) - E[\tilde{u}(t)|z(t)] \tag{11}$$

$$\hat{y}^*(t) = y(t) - \tilde{y}^*(t) = y(t) - E[\tilde{y}(t)|z(t)] \tag{12}$$

with

$$z(t) = y(t) - D(t)u(t) \tag{13}$$

where $E[\cdot]$ denotes mathematical expectation and $E[x|y]$ is the linear minimal variance estimator, i.e., $E[x|y] = E[x] + E[xy^T]E[yy^T]^{-1}(y - E[y])$ that coincides with the conditional expectation in the gaussian case [1], [2].

Expressions (11) and (12) can be computed recursively by taking advantage of the standard Kalman filter equations. Note that relations (3) and (4) allow to write model (1)–(2) in the form

$$x(t+1) = A(t)x(t) + B(t)u(t) - B(t)\tilde{u}(t) + G(t)w(t) \tag{14}$$

$$y(t) = C(t)x(t) + D(t)u(t) - D(t)\tilde{u}(t) + \tilde{y}(t). \tag{15}$$

By introducing the auxiliary white processes

$$n_x(t) = G(t)w(t) - B(t)\tilde{u}(t) \tag{16}$$

$$n_y(t) = \tilde{y}(t) - D(t)\tilde{u}(t) \tag{17}$$

with covariances

$$Q(t) = E\left[n_x(t)n_x^T(t)\right] = G(t)\Sigma_w(t)G(t)^T + B(t)\tilde{\Sigma}_u(t)B^T(t) \tag{18}$$

$$R(t) = E\left[n_y(t)n_y^T(t)\right] = \tilde{\Sigma}_y(t) - \tilde{\Sigma}_{uy}^T(t)\,D^T(t) - D(t)\tilde{\Sigma}_{uy}(t) + D(t)\tilde{\Sigma}_u(t)D^T(t) \tag{19}$$

$$S(t) = E\left[n_x(t)n_y^T(t)\right] = B(t)\left[\tilde{\Sigma}_u(t)D^T(t) - \tilde{\Sigma}_{uy}(t)\right] \tag{20}$$

it is possible to rewrite relations (14)–(15) as

$$x(t+1) = A(t)x(t) + B(t)u(t) + n_x(t) \tag{21}$$

$$z(t) = C(t)x(t) + n_y(t). \tag{22}$$

In the previous formulation, $z(t)$ plays the role of measured output so that the corresponding Kalman filter equations are

$$x(t+1|t) = A(t)x(t|t-1) + B(t)u(t) + K(t)\varepsilon(t) \tag{23}$$

$$x(0|-1) = \bar{x}_0 \tag{24}$$

$$K(t) = [A(t)P(t|t-1)C^T(t) + S(t)]\Sigma_\varepsilon(t)^{-1} \tag{25}$$

$$P(t+1|t) = A(t)P(t|t-1)A^T(t) + Q(t) - [A(t)P(t|t-1)C^T(t) + S(t)]\Sigma_\varepsilon(t)^{-1} \times [A(t)P(t|t-1)C^T(t) + S(t)]^T \tag{26}$$

$$P(0|-1) = P_0 \tag{27}$$

where

$$x(t+1|t) = E[x(t+1)|z(t), z(t-1), \ldots, z(0)] \tag{28}$$

while $\varepsilon(t)$ and $\Sigma_\varepsilon(t)$ denote the innovation of $z(t)$ and its covariance matrix, given by

$$\begin{aligned}\varepsilon(t) &= z(t) - C(t)x(t|t-1)\\ &= C(t)(x(t) - x(t|t-1)) + n_y(t) \qquad (29)\end{aligned}$$

$$\Sigma_\varepsilon(t) = E[\varepsilon(t)\varepsilon^T(t)] = C(t)P(t|t-1)C^T(t) + R(t). \qquad (30)$$

To compute $\tilde{u}^*(t), \tilde{y}^*(t)$ in (11)–(12), we can replace $z(t)$ with its innovation

$$\tilde{u}^*(t) = E[\tilde{u}(t)|\varepsilon(t)] = E[\tilde{u}(t)\varepsilon^T(t)]E[\varepsilon(t)\varepsilon^T(t)]^{-1}\varepsilon(t) \quad (31)$$

$$\tilde{y}^*(t) = E[\tilde{y}(t)|\varepsilon(t)] = E[\tilde{y}(t)\varepsilon^T(t)]E[\varepsilon(t)\varepsilon^T(t)]^{-1}\varepsilon(t) \quad (32)$$

then, since $\tilde{u}(t)$ and $\tilde{y}(t)$ are uncorrelated with $x(t) - x(t|t-1)$ it is easy to show that

$$E\left[\tilde{u}(t)\varepsilon^T(t)\right] = E\left[\tilde{u}(t)n_y^T(t)\right] = \tilde{\Sigma}_{uy}(t) - \tilde{\Sigma}_u(t)D^T(t) \quad (33)$$

$$E\left[\tilde{y}(t)\varepsilon^T(t)\right] = E\left[\tilde{y}(t)n_y^T(t)\right] = \tilde{\Sigma}_y(t) - \tilde{\Sigma}_{uy}^T(t)D^T(t) \quad (34)$$

hence

$$\tilde{u}^*(t) = \left[\tilde{\Sigma}_{uy}(t) - \tilde{\Sigma}_u(t)D^T(t)\right]\Sigma_\varepsilon(t)^{-1}\varepsilon(t) \qquad (35)$$

$$\tilde{y}^*(t) = \left[\tilde{\Sigma}_y(t) - \tilde{\Sigma}_{uy}^T(t)D^T(t)\right]\Sigma_\varepsilon(t)^{-1}\varepsilon(t). \qquad (36)$$

Finally, by using (11) and (12), the minimal variance estimates of $\hat{y}(t)$ and $\hat{u}(t)$ can be written in the form

$$\hat{u}^*(t) = u(t) - \left[\tilde{\Sigma}_{uy}(t) - \tilde{\Sigma}_u(t)D^T(t)\right]\Sigma_\varepsilon(t)^{-1}\varepsilon(t) \qquad (37)$$

$$\hat{y}^*(t) = y(t) - \left[\tilde{\Sigma}_y(t) - \tilde{\Sigma}_{uy}^T(t)D^T(t)\right]\Sigma_\varepsilon(t)^{-1}\varepsilon(t). \qquad (38)$$

These equations constitute the solution of Problem 1.

Remark 2: An alternative expression for the filtered output $\hat{y}^*(t)$ is given by

$$\hat{y}^*(t) = C(t)x(t|t) + D(t)\hat{u}^*(t) \qquad (39)$$

where $x(t|t)$ is the filtered state given by

$$x(t|t) = x(t|t-1) + P(t|t-1)C^T(t)\Sigma_\varepsilon(t)^{-1}\varepsilon(t). \qquad (40)$$

Proof: To verify the equivalence between (38) and (39) first replace $x(t|t)$ and $\hat{u}^*(t)$ in (39) with (40) and (37) in order to obtain

$$\begin{aligned}\hat{y}^*(t) = {}& C(t)x(t|t-1) + D(t)u(t) + [C(t)P(t|t-1)C^T(t)\\ & - D(t)\tilde{\Sigma}_{uy}(t) + D(t)\tilde{\Sigma}_u(t)D^T(t)]\Sigma_\varepsilon(t)^{-1}\varepsilon(t). \qquad (41)\end{aligned}$$

Now, by using (19) and (30), it can be shown that

$$\begin{aligned}\hat{y}^*(t) = {}& C(t)x(t|t-1) + D(t)u(t) + \varepsilon(t)\\ & - \left[\tilde{\Sigma}_y(t) - \tilde{\Sigma}_{uy}^T(t)D^T(t)\right]\Sigma_\varepsilon(t)^{-1}\varepsilon(t). \qquad (42)\end{aligned}$$

Finally, by replacing $\varepsilon(t)$ with $z(t) - C(t)x(t|t-1)$ and $z(t)$ with $y(t) - D(t)u(t)$, (38) can be obtained in a straightforward way. ■

Remark 3: When the system is purely dynamic, i.e., $D(t) = 0, \forall t$, the optimal estimate of $\hat{y}(t)$ is given by $C(t)x(t|t)$, as in standard Kalman filtering. Note also that if $\tilde{\Sigma}_y(t) = \tilde{\Sigma}_{uy}^T(t)D^T(t), \forall t$, the optimal estimate of $\hat{y}(t)$ concides with its observation $y(t)$. Finally, when $D(t) = 0, \forall t$ and $\tilde{\Sigma}_{uy}(t) = 0, \forall t$, the optimal estimate of the noiseless input $\hat{u}(t)$ concides with its observation $u(t)$.

IV. Evaluation of the Expected Performance

The purpose of this section is to develop an expression for the expected performance of the filter (37)–(38), i.e., for the covariance matrices of the estimation errors

$$\begin{aligned}e_u(t) &= \hat{u}(t) - \hat{u}^*(t)\\ &= -\tilde{u}(t) + \left[\tilde{\Sigma}_{uy}(t) - \tilde{\Sigma}_u(t)D^T(t)\right]\Sigma_\varepsilon(t)^{-1}\varepsilon(t) \qquad (43)\\ e_y(t) &= \hat{y}(t) - \hat{y}^*(t)\\ &= -\tilde{y}(t) + \left[\tilde{\Sigma}_y(t) - \tilde{\Sigma}_{uy}^T(t)D^T(t)\right]\Sigma_\varepsilon(t)^{-1}\varepsilon(t). \qquad (44)\end{aligned}$$

For simplicity, rewrite (43) and (44) as

$$e_u(t) = -\tilde{u}(t) + H_u(t)\Sigma_\varepsilon(t)^{-1}\varepsilon(t) \qquad (45)$$

$$e_y(t) = -\tilde{y}(t) + H_y(t)\Sigma_\varepsilon(t)^{-1}\varepsilon(t) \qquad (46)$$

where

$$H_u(t) = \tilde{\Sigma}_{uy}(t) - \tilde{\Sigma}_u(t)D^T(t) \qquad (47)$$

$$H_y(t) = \tilde{\Sigma}_y(t) - \tilde{\Sigma}_{uy}^T(t)D^T(t). \qquad (48)$$

By taking into account (30), (33), and (34) it is easy to show that

$$P_u(t) = E\left[e_u(t)e_u^T(t)\right] = \tilde{\Sigma}_u(t) - H_u(t)\Sigma_\varepsilon(t)^{-1}H_u^T(t) \quad (49)$$

$$P_y(t) = E\left[e_y(t)e_y^T(t)\right] = \tilde{\Sigma}_y(t) - H_y(t)\Sigma_\varepsilon(t)^{-1}H_y^T(t). \quad (50)$$

Remark 4: Consider now a time-invariant system described by the matrices A, B, G, C, D, whose state, input, and output noise are stationary stochastic processes characterized by covariance and cross-covariance matrices $\Sigma_w, \tilde{\Sigma}_u, \tilde{\Sigma}_y$, and $\tilde{\Sigma}_{uy}$. In this case, when the pair $(A - SR^{-1}C, C)$ is detectable and the pair $(A - SR^{-1}C, \bar{Q})$ is stabilizable ($\bar{Q}\bar{Q}^T = Q - SR^{-1}S^T$), $P(t+1|t)$ converges, for $t \to \infty$, to the unique solution P of the algebraic Riccati equation

$$\begin{aligned}P = {}& APA^T + Q - [APC^T + S] \qquad (51)\\ & \times [CPC^T + R]^{-1}[APC^T + S]^T \qquad (52)\end{aligned}$$

so that $P_u(t)$ and $P_y(t)$ converge to the the matrices P_u, P_y given by

$$P_u = \lim_{t\to\infty} P_u(t) = \tilde{\Sigma}_u - H_u[CPC^T + R]^{-1}H_u^T \qquad (53)$$

$$P_y = \lim_{t\to\infty} P_y(t) = \tilde{\Sigma}_y - H_y[CPC^T + R]^{-1}H_y^T \qquad (54)$$

with

$$H_u = \tilde{\Sigma}_{uy} - \tilde{\Sigma}_u\, D^T \qquad (55)$$

$$H_y = \tilde{\Sigma}_y - \tilde{\Sigma}_{uy}^T\, D^T. \qquad (56)$$

Moreover, the filter (23)–(27) is asymptotically stable for $t \to \infty$.

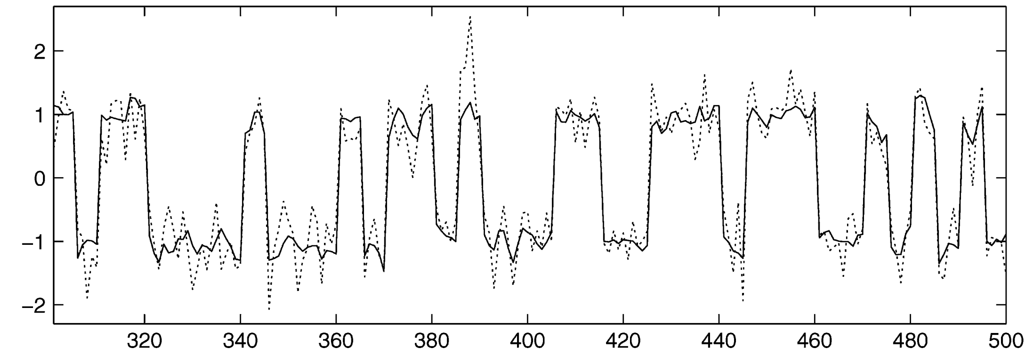

Fig. 1. Comparison between the first noiseless input (continuous line) and its observation (dotted line).

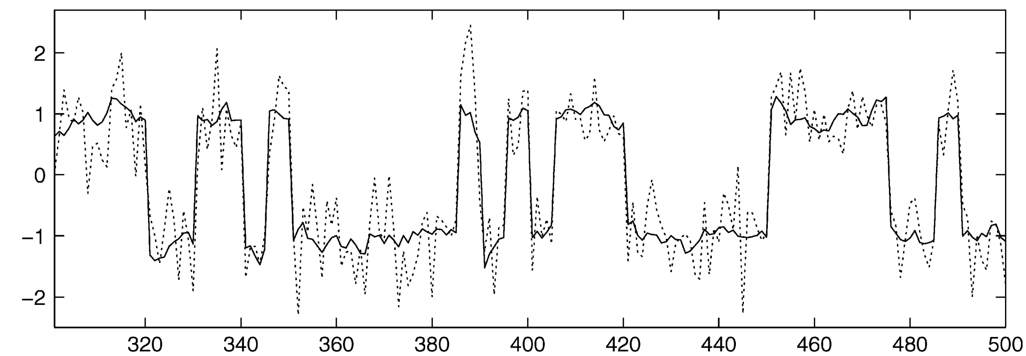

Fig. 2. Comparison between the second noiseless input (continuous line) and its observation (dotted line).

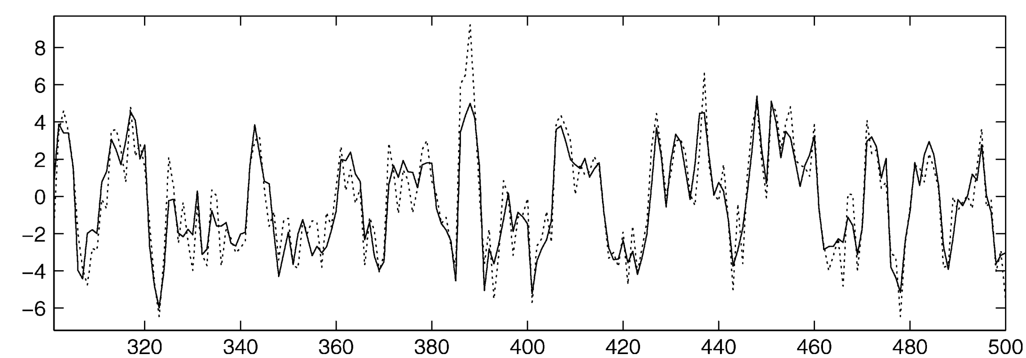

Fig. 3. Comparison between the first noiseless output (continuous line) and its observation (dotted line).

V. Numerical Results

The results obtained in previous sections have been numerically verified by means of a 100 runs Monte Carlo simulation performed on the two inputs-two outputs time-invariant system described by the following matrices:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -0.3 & 0.4 & -0.2 \\ -0.1 & 0.2 & 0.4 \end{bmatrix} \quad B = \begin{bmatrix} 0.8 & -0.8 \\ 0.17 & -0.37 \\ 1.09 & 1.1 \end{bmatrix}$$

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1.7 & 1.5 \\ 0.51 & -1 \end{bmatrix}.$$

The number of samples is 500. The input sequences $\hat{u}_1(\cdot), \hat{u}_2(\cdot)$ have unit variance and are shown in Figs. 1 and 2 (last 200 samples). In every run, the noise sequences $w(\cdot), \tilde{u}(\cdot), \tilde{y}(\cdot)$ are characterized by the following covariance and cross-covariance matrices:

$$\Sigma_w = \begin{bmatrix} 0.56 & 0.26 & 0.45 \\ 0.26 & 0.17 & 0.23 \\ 0.45 & 0.23 & 0.39 \end{bmatrix} \quad \tilde{\Sigma}_u = \begin{bmatrix} 0.12 & 0.15 \\ 0.15 & 0.25 \end{bmatrix}$$

$$\tilde{\Sigma}_y = \begin{bmatrix} 1.3 & 1.7 \\ 1.7 & 2.4 \end{bmatrix} \quad \tilde{\Sigma}_{uy} = \begin{bmatrix} 0.38 & 0.51 \\ 0.46 & 0.7 \end{bmatrix}.$$

The initial state x_0 is a random vector and (23) and (26) have been initialized as $x(0|-1) = 0$ and $P(0|-1) = I_n$.

Figs. 1–4 report the noiseless inputs and outputs (continuous line) and the associated noisy observations (dotted line) in a typical case of the Monte Carlo simulation (last 200 samples).

The effectiveness of the filter can be observed, in the same typical case, in Figs. 5–8, where the noiseless inputs and outputs (continuous

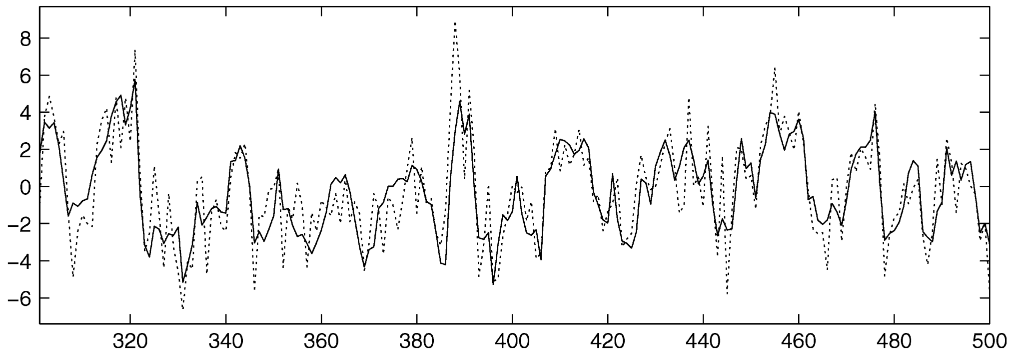

Fig. 4. Comparison between the second noiseless output (continuous line) and its observation (dotted line).

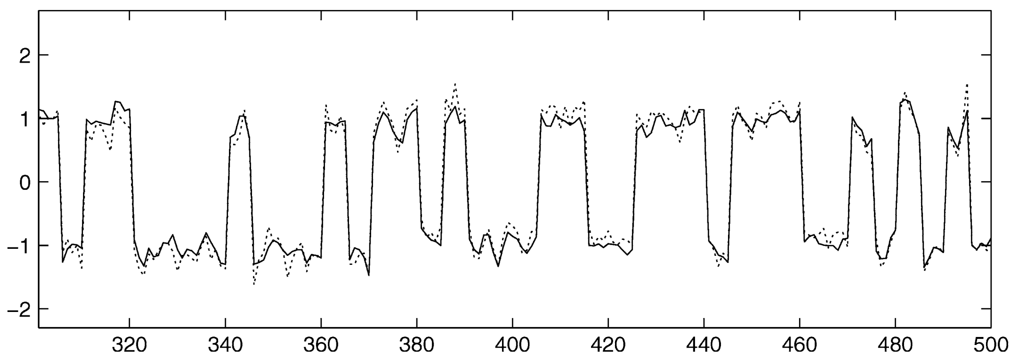

Fig. 5. Comparison between the first noiseless input (continuous line) and its optimal estimate (dotted line).

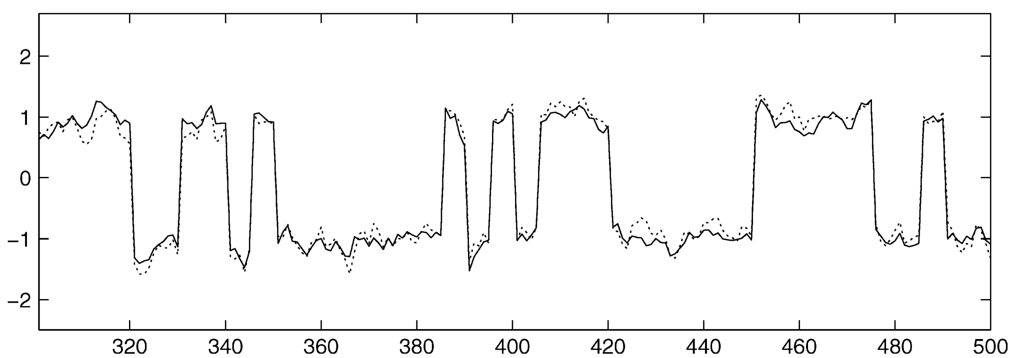

Fig. 6. Comparison between the second noiseless input (continuous line) and its optimal estimate (dotted line).

line) are compared with the corresponding filtered quantities (dotted line).

The covariance matrices of the estimation errors, obtained by means of the asymptotic relations (53) and (54) are

$$P_u = \begin{bmatrix} 0.0271 & 0.0083 \\ 0.0083 & 0.0251 \end{bmatrix}$$
$$P_y = \begin{bmatrix} 0.3343 & 0.2912 \\ 0.2912 & 0.3189 \end{bmatrix}$$

while the means of the actual values obtained in the 100 runs of the Monte Carlo simulation are

$$\bar{P}_u = \begin{bmatrix} 0.0272 \pm 0.0018 & 0.0085 \pm 0.0027 \\ 0.0085 \pm 0.0027 & 0.0253 \pm 0.0043 \end{bmatrix}$$

$$\bar{P}_y = \begin{bmatrix} 0.3343 \pm 0.0206 & 0.2912 \pm 0.0254 \\ 0.2912 \pm 0.0254 & 0.3194 \pm 0.0334 \end{bmatrix}.$$

The theoretical results are thus confirmed in a very accurate way by the numerical simulation.

VI. Conclusion

In this note, the extension of the stochastic context of Kalman filtering to the presence of additive noise on input observations has been considered. This extended filter has then been used to solve the problem of optimal (minimal variance) estimation of noise-corrupted input and output sequences. A unified context for both Kalman filtering and errors-in-variables filtering has thus been established. A Monte Carlo simulation has shown the effectiveness of this extended filter and the excellent agreement between expected and observed performances.

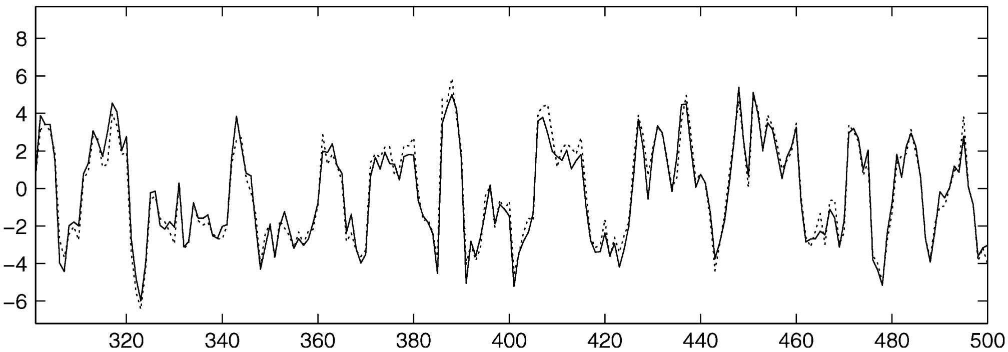

Fig. 7. Comparison between the first noiseless output (continuous line) and its optimal estimate (dotted line).

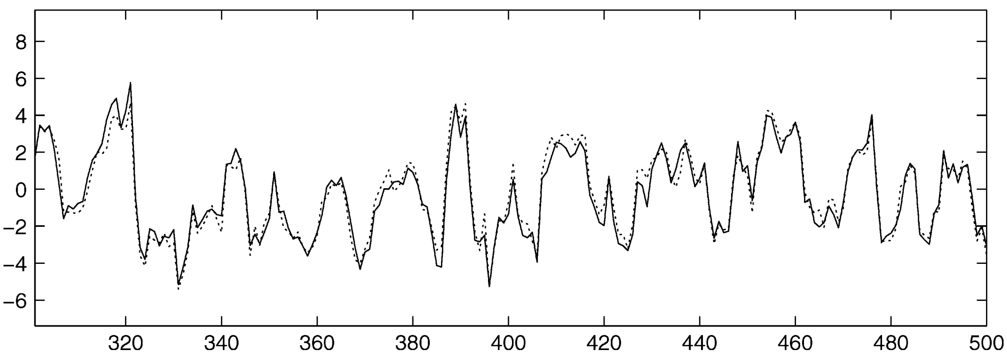

Fig. 8. Comparison between the second noiseless output (continuous line) and its optimal estimate (dotted line).

APPENDIX

Relations (11) and (12) in Theorem 1 can be obtained by considering the following interpolation problem.

Problem 2: Given model (1)–(4), covariance matrices (5)–(10) and the input–ouput observations $\{u(0), y(0), \ldots, u(t), y(t)\}$, determine the optimal (minimal variance) linear estimate of the noise-free input–output sequence $\{\hat{u}(0), \hat{y}(0), \ldots, \hat{u}(t), \hat{y}(t)\}$.

Solution of Problem 2: For the sake of simplicity, we will make reference to a time-invariant system described by the matrices A, B, G, C, D and to stationary white processes described by the covariances $\Sigma_w, \tilde{\Sigma}_u, \tilde{\Sigma}_u$, and $\tilde{\Sigma}_{uy}$. As it will be clear later, it is not restrictive to assume $x_0 = 0$.

Let us define the following vectors:

$$\hat{u} = [\hat{u}(0)\hat{u}(1)\ldots\hat{u}(t)]^T \tag{57}$$
$$\hat{y} = [\hat{y}(0)\hat{y}(1)\ldots\hat{y}(t)]^T \tag{58}$$
$$w = [w(0)w(1)\ldots w(t)]^T. \tag{59}$$

It is easy to verify that, because of (1)–(2), they are linked by the relation

$$\hat{y} = M_u\hat{u} + M_w w \tag{60}$$

where

$$M_u = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ CA^{t-1}B & CA^{t-2}B & \cdots & D \end{bmatrix} \tag{61}$$

$$M_w = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ CG & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ CA^{t-1}G & \cdots & CG & 0 \end{bmatrix}. \tag{62}$$

By denoting with I_k the $k \times k$ identitiy matrix and with $\otimes$ the Kronecker product, relation (60) can also be expressed in the form

$$M\hat{v} = 0 \tag{63}$$

where

$$\hat{v} = [\hat{y}^T \, \hat{u}^T \, w^T]^T \tag{64}$$

and

$$M = [I_{t+1} \otimes I_m \;\; -M_u \;\; -M_w]. \tag{65}$$

Define now the vectors

$$v = [y^T \, u^T \, \underbrace{0\ldots0}_{(t+1)\times p}]^T \tag{66}$$
$$\tilde{v} = [\tilde{y}^T \, \tilde{u}^T \, -w^T]^T \tag{67}$$

where

$$u = [u(0)u(1)\ldots u(t)]^T \tag{68}$$
$$\tilde{u} = [\tilde{u}(0)\tilde{u}(1)\ldots\tilde{u}(t)]^T \tag{69}$$
$$y = [y(0)y(1)\ldots y(t)]^T \tag{70}$$
$$\tilde{y} = [\tilde{y}(0)\tilde{y}(1)\ldots\tilde{y}(t)]^T. \tag{71}$$

From (3)–(4), it follows immediately:

$$v = \hat{v} + \tilde{v} \tag{72}$$

so that (63) leads to

$$Mv = M\tilde{v} = \Psi. \tag{73}$$

The solution of Problem 2 consists thus in finding the optimal estimate $\hat{v}^*$ of $\hat{v}$ satisfying condition (63), starting from the data v, the model M and the covariances (5)–(10). The solution can rely on two different approaches whether $\tilde{u}(t), \tilde{y}(t)$, and $w(t)$ are Gaussian processes or not. In the former case, the problem can be solved by maximizing the likelihood function

$$J(\hat{v}^*) = -\frac{1}{2}(v - \hat{v}^*)^T \tilde{\Sigma}^{-1}(v - \hat{v}^*) \tag{74}$$

under the constraint (63), where

$$\begin{aligned}\tilde{\Sigma} &= E[\tilde{v}\,\tilde{v}^T] \\ &= \begin{bmatrix} I_{t+1} \otimes \tilde{\Sigma}_y & I_{t+1} \otimes \tilde{\Sigma}_{uy}^T & 0 \\ I_{t+1} \otimes \tilde{\Sigma}_{uy} & I_{t+1} \otimes \tilde{\Sigma}_u & 0 \\ 0 & 0 & I_{t+1} \otimes \Sigma_w \end{bmatrix}.\end{aligned} \tag{75}$$

By introducing the Lagrange multipliers vector λ the solution can be obtained by minimizing the loss function

$$J(\hat{v}^*, \lambda) = (v - \hat{v}^*)^T \tilde{\Sigma}^{-1}(v - \hat{v}^*) + \lambda^T M \hat{v}^*. \tag{76}$$

By equating to zero the gradient vectors of (76) with respect to $\hat{v}^*$ and λ we obtain

$$\lambda = 2(M\tilde{\Sigma}M^T)^{-1}Mv \tag{77}$$

$$\hat{v}^* = v - \frac{\tilde{\Sigma}M^T\lambda}{2}. \tag{78}$$

The maximum likelihood (minimum variance) solution is thus given by

$$\hat{v}^* = v - \tilde{\Sigma}M^T(M\tilde{\Sigma}M^T)^{-1}Mv. \tag{79}$$

Note that the same result can be obtained also by considering the minimal variance estimate of the noise vector $\tilde{v}$, given by the conditional expectation

$$\begin{aligned}\tilde{v}^* &= \mathrm{E}[\tilde{v}|\Psi] = \mathrm{E}[\tilde{v}\Psi^T]\mathrm{E}[\Psi\Psi^T]^{-1}\Psi \\ &= \tilde{\Sigma}M^T(M\tilde{\Sigma}M^T)^{-1}\Psi\end{aligned} \tag{80}$$

where Ψ is defined by (73); in fact

$$\hat{v}^* = v - \tilde{v}^*. \tag{81}$$

In the non-Gaussian case, relation (80) can be obtained by solving a weighted least-squares problem and constitutes the best linear estimate of $\tilde{v}$ (in the least-squares sense) that can be obtained from v under condition (73).

Proof of Theorem 1: The generic samples $\tilde{u}^*(\tau), \tilde{y}^*(\tau), (0 \leq \tau \leq t)$ obtained from (80) depend on both past and future data; the only exception concerns $\tilde{u}^*(0), \tilde{y}^*(0)$ that do not depend on past samples and $\tilde{u}^*(t), \tilde{y}^*(t)$ that do not depend on future samples. $\tilde{u}^*(t)$ and $\tilde{y}^*(t)$ are thus the filtered quantities required for the solution of Problem 1. Because of (80), they are given by

$$\tilde{u}^*(t) = E[\tilde{u}(t)|\Psi] = E[\tilde{u}(t)|\psi(t), \psi(t-1), \ldots, \psi(0)] \tag{82}$$

$$\tilde{y}^*(t) = E[\tilde{y}(t)|\Psi] = E[\tilde{y}(t)|\psi(t), \psi(t-1), \ldots, \psi(0)] \tag{83}$$

where, from (73)

$$\psi(\tau) = y(\tau) - \sum_{k=0}^{\tau-1} CA^{\tau-1-k}Bu(k) - D\,u(\tau). \tag{84}$$

Since $\tilde{u}(t), \tilde{y}(t)$ are white and uncorrelated with $\hat{u}(t)$ they do not depend on the past values of $u(t), y(t)$ so that

$$\tilde{u}^*(t) = E[\tilde{u}(t)|\Psi] = E[\tilde{u}(t)|\psi(t)] = E[\tilde{u}(t)|y(t) - Du(t)] \tag{85}$$

$$\tilde{y}^*(t) = E[\tilde{y}(t)|\Psi] = E[\tilde{y}(t)|\psi(t)] = E[\tilde{y}(t)|y(t) - Du(t)]. \tag{86}$$

By recalling (81), the filtered samples $\hat{u}^*(t), \hat{y}^*(t)$ can finally be obtained by means of the relations

$$\hat{u}^*(t) = u(t) - \tilde{u}^*(t) = u(t) - E[\tilde{u}(t)|y(t) - Du(t)] \tag{87}$$

$$\hat{y}^*(t) = y(t) - \tilde{y}^*(t) = y(t) - E[\tilde{y}(t)|y(t) - Du(t)]. \tag{88}$$

Remark 5: If the initial state x_0 is not zero the term $M_x x_0$, where $M_x = [C^T A^T C^T \ldots (A^t)^T C^T]^T$, will appear in (60). This quantity must be included in the constraint (63); it does not affect, however, the obtained results since x_0 has been assumed as uncorrelated with $\tilde{u}(t)$ and $\tilde{y}(t)$.

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Upper Saddle River, NJ: Prentice-Hall, 1979.

[2] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*. Upper Saddle River, NJ: Prentice-Hall, 1995.

[3] R. K. Mehra, "On the identification of variances and adaptive Kalman filtering," *IEEE Trans. Autom. Control*, vol. AC-15, no. 2, pp. 175–184, Apr. 1970.

[4] S. S. Godbole, "Kalman filtering with no a priori information about noise—White noise case: Identification of covariances," *IEEE Trans. Autom. Control*, vol. AC-19, no. 5, pp. 561–563, Oct. 1974.

[5] M. Deistler, "Linear errors-in-variables models," in *Time Series and Linear Systems*, S. Bittanti, Ed. Berlin, Germany: Springer-Verlag, 1986, pp. 37–68.

[6] R. Guidorzi, R. Diversi, and U. Soverini, "Optimal errors-in-variables filtering," *Automatica*, vol. 39, pp. 281–289, Feb. 2003.

[7] ——, "Errors-in-variables filtering in behavioral and state-space contexts," in *Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications*, S. Van Huffel and P. Lemmerling, Eds, Dordrecht, The Netherlands: Kluwer, 2002, pp. 281–291.

[8] R. Diversi, R. Guidorzi, and U. Soverini, "Algorithms for optimal errors-in-variables filtering," *Syst. Control Lett.*, vol. 48, pp. 1–13, Jan. 2003.

[9] B. Roorda and C. Heij, "Global total least squares modeling of multivariable time series," *IEEE Trans. Autom. Control*, vol. 40, no. 1, pp. 50–63, Jan. 1995.

[10] I. Markovsky and B. De Moor, "Linear dynamic filtering with noisy input and output," in *Preprints 13th IFAC Symp. System Identification*, Rotterdam, The Netherlands, Aug. 2003, pp. 1749–1754.

[11] R. Diversi, R. Guidorzi, and U. Soverini, "Kalman filtering in symmetrical noise environments," in *Proc. 11th IEEE Mediterranean Conf. Control and Automation*, Rhodes, Greece, Jun. 2003.