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

Quadratic Filtering of non-Gaussian Linear Systems with Random Observation Matrices

F. Cacace * **A. Fasano** * **A. Germani** **

* *Università Campus Bio-Medico di Roma, Rome, Italy (e-mail: {a.fasano,f.cacace}@unicampus.it)*
** *Università dell'Aquila, Center of Excellence DEWS, L'Aquila, Italy (e-mail: a.germani@univaq.it)*

Abstract: In this paper we consider the problem of state estimation for linear discrete-time non-Gaussian systems with random observation matrices. This is the model for systems with observation losses due to propagation through unreliable communication channels. Losses may result from intermittent failures that cause packet dropouts, as in the case of networks, or fading phenomena in propagation channel, as in the case of wireless networks. These are common problems in wireless sensor network, or networked control systems. In this paper, we do not make any assumption about the distribution of the observation matrix, thus encompassing a great variety of possible scenarios. We derive the quadratic estimate of the state by means of a recursive algorithm. The solution is obtained by applying the Kalman filter to a suitably augmented system, which is fully observable. The augmented system is constructed as the aggregate of the actual system and the observable part of a system having as state the second Kronecker power of the original state, namely the quadratic system. To extract the observable part of the quadratic system we exploit the knowledge of the rank of the corresponding observability matrix. This approach guarantees the internal stability of the estimation filter.

Keywords: Uncertain linear systems, Kalman filters, Nonlinear filters, Quadratic filtering, Non-Gaussian processes.

1. INTRODUCTION

The problem of state estimation for linear systems has been widely studied due to its crucial role in many scientific applications, ranging from engineering to econometrics. For linear Gaussian systems the well-known Kalman filter (Kalman, 1960) provides the optimal state estimate in the minimum mean square error sense. However, in many situations the hypothesis of Gaussianity does not represent a realistic statistical description of the system, and should be removed (Spall and Wall, 1984; Spall, 1985; Wu and Chen, 1993; Spall, 2003). In the non-Gaussian case, the minimum mean square error estimation, is in general an infinite dimensional problem (Zakai, 1969), whose solution cannot be easily computed numerically. In this case the Kalman filter provides the suboptimal affine estimate, offering a good compromise between computational simplicity and goodness of the estimate. Unfortunately, in many other cases this procedure is unsatisfactory, and more accurate approaches, such as polynomial algorithms (De Santis et al., 1995; Carravetta et al., 1996), are needed.

Another quite common case where the Kalman filter loses its optimality is when the system is characterized by uncertain observations or when system model parameters are not known exactly. These conditions translate into observation and state transition matrices that are random. In particular, in this paper we focus on the problem of state estimation for system with random observation matrices. This is the model, for example, of communication systems where the channel undergoes random interruptions or is affected by fading phenomena. More in general, uncertain observations are quite common in networked control systems, where the state of a remote plant is estimated from measurements carried through a lossy network (Hespanha et al., Jan.). Such systems can be described as linear dynamic systems with deterministic parameter matrices and state-dependent measurement noise (De Koning, 1984). Thus, the assumptions of the Kalman filter are violated and the Kalman filter recursive estimate cannot be derived directly.

In De Koning (1984) and Luo and Zhu (2008) the linear minimum variance state estimation is derived under some mild conditions, for linear discrete system with random state transition and measurement matrices. It is in the form of a modified Kalman filter.

In Nahi (1969) the problem of observation losses is investigated and the optimal linear state estimator is derived under the assumption that the sequence of packet dropouts, modeled as a Bernoulli process, is independent. The estimator has a recursive structure very similar to the Kalman filter. In Hadidi and Schwartz (1979) the result of Nahi (1969) is generalized investigating the least mean-square error recursive estimator over the class of linear filters, when the sequence of packet dropouts is not necessarily independent. The authors show that, in general, the optimal linear estimator is not recursive. However, provided some conditions are met, a necessary and sufficient condition for the existence of a linear recursive estimator is given.

In Sinopoli et al. (2004) the statistical performance of the Kalman filter are investigated assuming that the estimator knows exactly whether the observation contains the signal to be estimated or contains noise alone, i.e., whether there is a packet dropout or not. Uncertain observations are thus modeled as a Bernoulli process. The statistical convergence properties of the estimation error covariance are studied and the authors show that there exists a critical value of the dropout rate, above which the state estimation error covariance is unbounded.

In Nakamori et al. (2003a,b) no assumptions on the state-space model for the signal are made and the problem of uncertain observation is tackled using covariance information. In particular in Nakamori et al. (2003a) the linear estimator is derived using the auto-covariance function of the signal in a semi-degenerate kernel form, whereas a quadratic filter is considered in Nakamori et al. (2003b).

In this paper, we consider the case of random observation matrices, assuming to know only the moments of the entries up to the fourth order. No assumptions are made about the distribution of the observation matrices. Note that, even in the case of linear systems with Gaussian state and output noise, the presence of random observation matrices makes the overall system non-Gaussian. Nonetheless, we assume non-Gaussian state and output noise. Under these hypotheses, the Kalman filter or the linear filter in Nahi (1969) are no longer optimal and it is necessary consider better suboptimal estimates. In this regard, polynomial algorithms (De Santis et al., 1995; Carravetta et al., 1996; Fasano et al., 2013) are more accurate than the linear one, and preserve the nice features of easy computability and recursivity. To cope with the randomness of the observation matrix, we replace it with its mean and move the state-dependent randomness to the noise term. Following an approach similar to the one in Fasano et al. (2013), we construct an augmented system as the aggregate of the original system and a suitably devised quadratic subsystem, which is fully observable. We derive the quadratic estimate of the state by means of a recursive algorithm obtained by applying the Kalman filter to the augmented system. Since this last is fully observable, the internal stability of the estimation filter is guaranteed (Fasano et al., 2013).

The paper is organized as follows. In Section 2, we present the system model with non-Gaussian state and output noise and random observation matrix. Moreover, we construct the augmented system comprising the original system and a suitably devised quadratic subsystem exploiting the results of Fasano et al. (2013). In Section 3, we derive the quadratic estimate. In Section 4, we evaluate the performance of the proposed approach by a numerical example. Finally, Section 5 follows with concluding remarks. To improve readability of the paper, some proofs are given in the Appendix.

2. SYSTEM MODEL

Consider the following class of linear discrete-time systems

$$x_{k+1} = Ax_k + f_k\ , \qquad k \geq 0 \tag{1}$$

$$y_k = C_k x_k + g_k \tag{2}$$

where $x_k, f_k \in \mathbb{R}^n$, $y_k, g_k \in \mathbb{R}$, $A \in \mathbb{R}^{n\times n}$ and $C_k \in \mathbb{R}^{1\times n}$. The initial state x_0 and the random sequences $\{f_k\}$, $\{g_k\}$, $\{C_k\}$ satisfy the following conditions for $k \geq 0$:

(1) $\mathbb{E}\{x_0\} = 0$, $\mathbb{E}\{f_k\} = 0$, $\mathbb{E}\{g_k\} = 0$,
(2) $\{f_k\}$ is a sequence of independent random vectors,
(3) $\{g_k\}$ is a sequence of independent random variables,
(4) $\{C_k\}$ is a sequence of i.i.d. [1] random vectors,
(5) $\{f_k\}$, $\{g_k\}$, $\{C_k\}$, x_0 are statistically independent,
(6) the components of x_0, f_k, C_k, and g_k have finite fourth moments,
(7) $\mathbb{E}\{x_0^{[i]}\} = m_{x_0}^{(i)}$, $\mathbb{E}\{f_k^{[i]}\} = m_{f_k}^{(i)}$, $\mathbb{E}\{g_k^i\} = m_{g_k}^{(i)}$, for $i = 2, 3, 4$, where $m_{x_0}^{(i)}$, $m_{f_k}^{(i)}$ are known vectors and $m_{g_k}^{(i)}$ is a known scalar,
(8) $\mathbb{E}\left\{C_k^{[i]}\right\} = \overline{C_i}$, for $i = 1, \ldots, 4$, where $\overline{C_i}$ are known constant vectors,
(9) the pair $\left(\overline{C_1}, A\right)$ is observable, and rank $\mathcal{O}\left(\overline{C_2}, A^{[2]}\right) =$ rank $\mathcal{O}\left(\overline{C_1}^{[2]}, A^{[2]}\right)$.

In condition 7 we make use of the *Kronecker power* $z^{[i]}$ (Bellman, 1997), defined as $z^{[1]} = z$, $z^{[i+1]} = z^{[i]} \otimes z$, where $\otimes$ is the Kronecker product. Note that since g_k is scalar, $g_k^{[i]} = g_k^i$. Condition 6 guarantees the existence of moments up to the fourth order (Feller, 1968). In condition 9, $\mathcal{O}(\,\cdot\,,\cdot\,)$ denotes the observability matrix of the pair within parentheses.

Hereinafter, we use the following notation to denote centered quantities

$$\widehat{Z} = Z - \mathbb{E}\{Z\}$$

where Z is a random matrix or vector. This notation improves readability and makes formulas more compact. Moreover, assuming z a (column) random vector, we denote by

$$\Sigma_z^{(i,j)} = \mathbb{E}\left\{z^{[i]} z^{[j]T}\right\} \tag{3}$$

and

$$\widehat{\Sigma}_z^{(i,j)} = \mathbb{E}\left\{\widehat{z^{[i]}}\,\widehat{z^{[j]T}}\right\} = \Sigma_z^{(i,j)} - \mathbb{E}\left\{z^{[i]}\right\}\mathbb{E}\left\{z^{[j]T}\right\} \tag{4}$$

with $i, j \geq 1$, the cross-correlation and cross-covariance matrix, respectively, of the Kronecker powers of z. If z is a row vector, it is replaced by its transpose within expectation in formulas (3) and (4). When $i = j = 1$, we use

$$\Sigma_z = \Sigma_z^{(1,1)} \tag{5}$$

$$\widehat{\Sigma}_z = \widehat{\Sigma}_z^{(1,1)} \tag{6}$$

to keep the notation simple.

In order to derive the quadratic estimate of the state of the system (1)–(2), we construct an augmented system that has as observations both y_k and y_k^2. Towards this end, we follow Fasano et al. (2013) and derive the augmented system considering the quadratic Kronecker power of (1)

$$x_{k+1}^{[2]} = (Ax_k + f_k)^{[2]} = A^{[2]}x_k^{[2]} + m_{f_k}^{(2)} + f_k^{(2)} \tag{7}$$

[1] Independent identically distributed.

where $f_k^{(2)} = Ax_k \otimes f_k + f_k \otimes Ax_k + \widehat{f_k^{[2]}}$, with $\mathbb{E}\{f_k^{(2)}\} = 0$. Introducing the $n^2 \times n^2$ permutation matrix Π that guarantees $f_k \otimes Ax_k = \Pi(Ax_k \otimes f_k)$ (Horn and Johnson, 1994), we can write $f_k^{(2)} = (I + \Pi)(Ax_k \otimes f_k) + \widehat{f_k^{[2]}}$. This expression will be useful later in Section 3.

Proceeding in the same way with (2) and taking into account that $y_k^2 = y_k^{[2]}$ as y_k is scalar, we get

$$\begin{aligned} y_k^2 &= (C_k x_k + g_k)^{[2]} = C_k^{[2]} x_k^{[2]} + g_k^2 + 2C_k x_k g_k \\ &= C_k^{[2]} x_k^{[2]} + g_k^2 + 2C_k x_k g_k \\ &\quad + \mathbb{E}\left\{C_k^{[2]}\right\} x_k^{[2]} - \mathbb{E}\left\{C_k^{[2]}\right\} x_k^{[2]} + m_{g_k}^{(2)} - m_{g_k}^{(2)} \\ &= \overline{C_2} x_k^{[2]} + m_{g_k}^{(2)} + 2C_k x_k g_k + \widehat{C_k^{[2]}} x_k^{[2]} + \widehat{g_k^2} \\ &= \overline{C_2} x_k^{[2]} + m_{g_k}^{(2)} + g_k^{(2)} \end{aligned} \tag{8}$$

where $g_k^{(2)} = 2C_k x_k g_k + \widehat{C_k^{[2]}} x_k^{[2]} + \widehat{g_k^2}$, with $\mathbb{E}\{g_k^{(2)}\} = 0$.

Rearranging (2) in a more convenient form, so that C_k is moved to the noise term we get

$$\begin{aligned} y_k &= \mathbb{E}\{C_k\} x_k + (C_k - \mathbb{E}\{C_k\}) x_k + g_k \\ &= \overline{C_1} x_k + \widehat{C_k} x_k + g_k \\ &= \overline{C_1} x_k + g_k^{(1)} \end{aligned} \tag{9}$$

where $g_k^{(1)} = \widehat{C_k} x_k + g_k$, with $\mathbb{E}\{g_k^{(1)}\} = 0$. Finally, combining (1), (2), (8), and (9) we get the full-order quadratic system derived from system (1)–(2)

$$x_{k+1} = Ax_k + f_k \qquad k \geq 0 \tag{10}$$
$$x_{k+1}^{[2]} = A^{[2]} x_k^{[2]} + m_{f_k}^{(2)} + f_k^{(2)} \tag{11}$$
$$y_k = \overline{C_1} x_k + g_k^{(1)} \tag{12}$$
$$y_k^2 = \overline{C_2} x_k^{[2]} + m_{g_k}^{(2)} + g_k^{(2)} \tag{13}$$

where $m_{f_k}^{(2)}$ and $m_{g_k}^{(2)}$ are deterministic inputs, and the noise terms, namely f_k, $f_k^{(2)}$, $g_k^{(1)}$, $g_k^{(2)}$, are zero mean.

By hypothesis the pair $\left(\overline{C_1}, A\right)$ is observable, that is $\operatorname{rank} \mathcal{O}\left(\overline{C_1}, A\right) = n$. However, the augmented system (10)–(13) is not fully observable, as proved in Fasano et al. (2013). This is a consequence of the fact that the quadratic part of the system, comprising (11) and (13), is *never* fully observable when $n > 1$, since $\operatorname{rank} \mathcal{O}\left(\overline{C_2}, A^{[2]}\right) = \operatorname{rank} \mathcal{O}\left(\overline{C_1}^{[2]}, A^{[2]}\right) < n^2$.

Nevertheless, exploiting the results of Fasano et al. (2013), where a closed form expression of $\operatorname{rank} \mathcal{O}(\overline{C_1}^{[2]}, A^{[2]})$ is provided, it is easy to find a coordinate transformation for extracting the observable part of the quadratic subsystem in (10)–(13).

Indeed, let $r = \operatorname{rank} \mathcal{O}\left(\overline{C_2}, A^{[2]}\right)$, which is equal to $\operatorname{rank} \mathcal{O}\left(\overline{C_1}^{[2]}, A^{[2]}\right)$ by hypothesis, and partition the observability matrix of the pair $\left(\overline{C_2}, A^{[2]}\right)$ as follows

$$\mathcal{O}\left(\overline{C_2}, A^{[2]}\right) = \begin{bmatrix} R \\ Q \end{bmatrix}$$

where $R \in \mathbb{R}^{r \times n^2}$, $Q \in \mathbb{R}^{(n^2 - r) \times n^2}$ and $\operatorname{rank} R = r$. Now, consider the singular value decomposition (SVD) of R

$$R = U \Sigma V^T \tag{14}$$

where $U \in \mathbb{R}^{r \times r}$ and $V \in \mathbb{R}^{n^2 \times n^2}$ are orthogonal matrices. Using the coordinate transformation

$$T = V^{\mathrm{T}} \tag{15}$$

it is easy to verify that the matrices $A^{[2]}$ and $\overline{C_2}$ become

$$T A^{[2]} T^{-1} = V^{\mathrm{T}} A^{[2]} V = \begin{bmatrix} A_o^{(2)} & 0_{r \times (n^2 - r)} \\ A_1^{(2)} & A_2^{(2)} \end{bmatrix} \tag{16}$$

and

$$\overline{C_2} T^{-1} = \overline{C_2} V = \begin{bmatrix} C_o^{(2)} & 0_{1 \times (n^2 - r)} \end{bmatrix} \tag{17}$$

where $A_o^{(2)} \in \mathbb{R}^{r \times r}$, $C_o^{(2)} \in \mathbb{R}^{1 \times r}$ and the pair $(C_o^{(2)}, A_o^{(2)})$ is fully observable.

Denoting by $T_r \in \mathbb{R}^{r \times n^2}$ the rectangular matrix collecting the first $r = \operatorname{rank} \mathcal{O}(\overline{C_2}, A^{[2]})$ rows of T, namely

$$T_r = T(1\!:\!r, :) \tag{18}$$

using Matlab® notation, matrices $A_o^{(2)}$ and $C_o^{(2)}$ in (16) and (17) become

$$A_o^{(2)} = T_r A^{[2]} T_r^T \tag{19}$$
$$C_o^{(2)} = \overline{C_2}\, T_r^T. \tag{20}$$

The key point in the above analysis is the *knowledge* of $\operatorname{rank} \mathcal{O}\left(\overline{C_1}^{[2]}, A^{[2]}\right)$, whose closed form expression is provided in Fasano et al. (2013) as a function of the spectrum of A.

It is worth stressing that the knowledge of the (exact) rank of the observability matrix of the quadratic subsystem is crucial to extract the observable subspace. Indeed, numerical computation of the rank of an observability matrix is an ill-conditioned problem (Paige, 1981). Moreover, iterative procedures for computing the observable subspace, like the ones in Rosenbrock (1970), fail to recognize the right dimension.

3. THE OBSERVABLE QUADRATIC SYSTEM

Using the $r \times n^2$ matrix T_r defined in (18), and denoting by $\tilde{x}_k^{(2)} = T_r x_k^{[2]}$, the state vector collecting the *observable* states of the quadratic subsystem, (10)–(13) can be rewritten as

$$x_{k+1} = Ax_k + f_k \qquad k \geq 0 \tag{21}$$
$$\tilde{x}_{k+1}^{(2)} = A_o^{(2)} \tilde{x}_k^{(2)} + \tilde{m}_{f_k}^{(2)} + \tilde{f}_k^{(2)} \tag{22}$$
$$y_k = \overline{C_1} x_k + g_k^{(1)} \tag{23}$$
$$y_k^2 = C_o^{(2)} \tilde{x}_k^{(2)} + m_{g_k}^{(2)} + g_k^{(2)} \tag{24}$$

where $\tilde{m}_{f_k}^{(2)} = T_r m_{f_k}^{(2)}$, $\tilde{f}_k^{(2)} = T_r f_k^{(2)}$, and $C_o^{(2)} = \overline{C_2}\, T_r^T$ as in (20).

Defining the extended state $X_k = \begin{bmatrix} x_k^T & \tilde{x}_k^{(2)T} \end{bmatrix}^T$, the extended input $U_k = \begin{bmatrix} \tilde{m}_{f_k}^{(2)T} & m_{g_k}^{(2)} \end{bmatrix}^T$, and the extended output $Y_k = \begin{bmatrix} y_k & y_k^2 \end{bmatrix}^T$, we get the following dynamical system

$$X_{k+1} = \mathcal{A} X_k + \mathcal{B} U_k + V_k \tag{25}$$
$$Y_k = \mathcal{C} X_k + \mathcal{D} U_k + W_k \tag{26}$$

where

$$\mathcal{A} = \begin{bmatrix} A & 0_{n\times r} \\ 0_{r\times n} & A_o^{(2)} \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} 0_{n\times r} & 0_{n\times 1} \\ I_{r\times r} & 0_{r\times 1} \end{bmatrix} \tag{27}$$

$$\mathcal{C} = \begin{bmatrix} \overline{C_1} & 0_{1\times r} \\ 0_{1\times n} & C_o^{(2)} \end{bmatrix}, \quad \mathcal{D} = \begin{bmatrix} 0_{1\times r} & 0 \\ 0_{1\times r} & 1 \end{bmatrix} \tag{28}$$

$$V_k = \begin{bmatrix} f_k \\ \tilde{f}_k^{(2)} \end{bmatrix}, \quad W_k = \begin{bmatrix} g_k^{(1)} \\ g_k^{(2)} \end{bmatrix}. \tag{29}$$

3.1 State estimate

Equation (25) represents a linear system whose state evolves in the space $\mathbb{R}^{n+r}$. Output equation (26) has the structure we need to calculate the optimal quadratic estimate for the state vector of system (1)–(2). The noise sequences $\{V_k\}$, $\{W_k\}$ are zero mean and white. Moreover, they are mutually orthogonal and orthogonal to the initial state $X_0 = \begin{bmatrix} x_0^T & \tilde{x}_0^{(2)T} \end{bmatrix}^T$. This is a consequence of the following propositions.

Proposition 1. The random process $\{V_k\}$ is wide-sense stationary with zero mean and autocorrelation function

$$\mathbb{E}\{V_h V_k^T\} = Q_k \delta_{h,k} \tag{30}$$

where Q_k is defined in (32).

Proof. See Appendix.

Proposition 2. The random process $\{W_k\}$ is wide-sense stationary with zero mean and autocorrelation function

$$\mathbb{E}\{W_h W_k^T\} = R_k \delta_{h,k} \tag{31}$$

where R_k is defined in (36).

Proof. See Appendix.

Proposition 3. The random processes $\{V_k\}$, $\{W_k\}$ are mutually orthogonal, i.e.,

$$\mathbb{E}\{V_h W_k^T\} = 0_{(n+r)\times 2}$$

and orthogonal to $X_0 = \begin{bmatrix} x_0^T & \tilde{x}_0^{(2)T} \end{bmatrix}^T$.

Proof. The proof is omitted due to space limitations.

The covariance matrix Q_k in (30) can be computed by exploiting conditions 1–7 of Section 2 and using some Kronecker algebra. We get

$$Q_k = \begin{bmatrix} \Sigma_{f_k} & \Sigma_{f_k}^{(1,2)} T_r^T \\ T_r \Sigma_{f_k}^{(2,1)} & Q_k(2,2) \end{bmatrix} \tag{32}$$

with

$$Q_k(2,2) = $$
$$= T_r \left\{ (I + \Pi) \left[A \Sigma_{x_k} A^T \otimes \Sigma_{f_k} \right] (I + \Pi) + \widehat{\Sigma}_{f_k}^{(2,2)} \right\} T_r^T$$

where Π is the permutation matrix that guarantees $f_k \otimes x_k = \Pi(x_k \otimes f_k)$ (Horn and Johnson, 1994), whereas

$$\Sigma_{x_k} = \mathbb{E}\left\{ x_k x_k^T \right\}, \qquad \Sigma_{f_k} = \mathbb{E}\left\{ f_k f_k^T \right\}, \tag{33}$$

$$\Sigma_{f_k}^{(1,2)} = \Sigma_{f_k}^{(2,1)T} = \mathbb{E}\{ f_k f_k^{[2]T} \}, \tag{34}$$

$$\widehat{\Sigma}_{f_k}^{(2,2)} = \mathbb{E}\left\{ \widehat{f_k^{[2]}} \widehat{f_k^{[2]T}} \right\} = \Sigma_{f_k}^{(2,2)} - m_{f_k}^{(2)} m_{f_k}^{(2)T}, \tag{35}$$

consistently with (3)–(4), and $m_{f_k}^{(2)} = \mathbb{E}\{f_k^{[2]}\}$, as defined in Section 2. Note that the entries of Σ_{f_k} in (33), $\Sigma_{f_k}^{(1,2)}$ in (34), and $\Sigma_{f_k}^{(2,2)}$ in (35) are known by hypothesis since they are the elements of $m_{f_k}^{(2)}$, $m_{f_k}^{(4)}$, $m_{f_k}^{(3)}$, respectively, which are defined in Section 2.

Working in a similar way for R_k in (31), we get

$$R_k = \begin{bmatrix} \mathbb{E}\{g_k^{(1)} g_k^{(1)}\} & \mathbb{E}\{g_k^{(1)} g_k^{(2)}\} \\ \mathbb{E}\{g_k^{(1)} g_k^{(2)}\} & \mathbb{E}\{g_k^{(2)} g_k^{(2)}\} \end{bmatrix} \tag{36}$$

with

$$\mathbb{E}\left\{ g_k^{(1)} g_k^{(1)} \right\} = \left(\overline{C_2} - \overline{C_1} \otimes \overline{C_1} \right) \mathbb{E}\left\{ x_k^{[2]} \right\} + m_{g_k}^{(2)} \tag{37}$$

$$\mathbb{E}\left\{ g_k^{(1)} g_k^{(2)} \right\} = \left(\overline{C_3} - \overline{C_1} \otimes \overline{C_2} \right) \mathbb{E}\left\{ x_k^{[3]} \right\} + m_{g_k}^{(3)}$$

$$\mathbb{E}\left\{ g_k^{(2)} g_k^{(2)} \right\} = 4 m_{g_k}^{(2)} \overline{C_2}\, \mathbb{E}\left\{ x_k^{[2]} \right\} + \left(\overline{C_4} - \overline{C_2} \otimes \overline{C_2} \right)$$
$$\times\, \mathbb{E}\left\{ x_k^{[4]} \right\} + m_{g_k}^{(4)} - m_{g_k}^{(2)} m_{g_k}^{(2)} \tag{38}$$

where $m_{g_k}^{(i)} = \mathbb{E}\{g_k^{(i)}\}$, for $i = 2, 3, 4$, and $\overline{C_i} = \mathbb{E}\left\{ C_k^{[i]} \right\}$, for $i = 1, \ldots, 4$, as defined in Section 2, whereas $\mathbb{E}\left\{ x_k^{[i]} \right\}$, for $i = 2, 3, 4$, can be computed recursively.

Now, the optimal linear estimate $\hat{X}_k$ of X_k can be computed using the Kalman filter, as follows

$$\hat{X}_0 = \mathbb{E}\{X_0\}, \quad P_{0|0} = \mathbb{E}\{\mathring{X}_0 \mathring{X}_0^T\} \tag{39}$$

$$\hat{X}_{k|k-1} = \mathcal{A}\hat{X}_{k-1} + \mathcal{B}U_{k-1}, \qquad k > 0 \tag{40}$$

$$\hat{X}_k = \hat{X}_{k|k-1} + K_k \left(Y_k - \mathcal{C}\hat{X}_{k|k-1} - \mathcal{D}U_k \right) \tag{41}$$

$$K_k = P_{k|k-1}\mathcal{C}^T \left(\mathcal{C}P_{k|k-1}\mathcal{C}^T + R_k \right)^{-1} \tag{42}$$

$$P_{k|k-1} = \mathcal{A}P_{k-1|k-1}\mathcal{A}^T + Q_{k-1} \tag{43}$$

$$P_{k|k} = (I - K_k\mathcal{C})\, P_{k|k-1} \tag{44}$$

where $\mathring{X}_0 = X_0 - \mathbb{E}\{X_0\}$. It is worth stressing that (39)–(44) provide the optimal linear estimate $\hat{X}_k$ as projection of X_k on the linear space of $(n+r)$-dimensional linear functions of the vector

$$\begin{bmatrix} 1 & Y_1^T & \cdots & Y_k^T \end{bmatrix}^T = \begin{bmatrix} 1 & y_1 & y_1^2 & \cdots & y_k & y_k^2 \end{bmatrix}^T. \tag{45}$$

This amounts to the optimal quadratic [2] estimate of X_k as a function of the observations $y_1, \ldots, y_k$. The vector constituted by the first n entries of $\hat{X}_k$, namely $\hat{x}_k$, gives the required optimal quadratic estimate of the state x_k of the system (1)–(2). The $n \times n$ leading principal submatrix of $P_{k|k}$ is the corresponding error covariance matrix of the quadratic estimate.

4. SIMULATION RESULTS

In this section we evaluate the performance of the proposed approach by a numerical example. Consider the system (1)–(2), with

$$A = \begin{bmatrix} 0.5 & 1 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.7 \end{bmatrix}, \tag{46}$$

$$C_k = \begin{bmatrix} -0.85\eta_k^{(1)} & \eta_k^{(2)} & -\eta_k^{(3)} \end{bmatrix} \tag{47}$$

[2] With no cross terms.

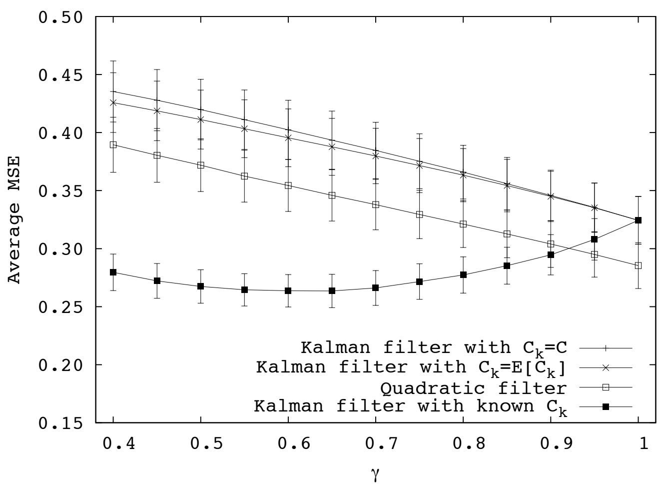

Fig. 1. Comparison of the average MSE of four state estimation algorithms, with confidence intervals.

where $\{\eta_k^{(i)}, k \geq 0\}$, for $i = 1, 2, 3$, are independent Bernoulli processes with parameter γ, i.e., $\mathbb{E}\{\eta_k^{(i)}\} = \gamma$. The state and measurement noises are non-Gaussian, as they are discrete random processes. In particular, $f_k = [\, f_k^{(1)} \;\; f_k^{(2)} \;\; f_k^{(3)} \,]^T$, with i.i.d. components $f_k^{(i)} \in \{-0.1,\, 0.3,\, 0.9\}$ with probabilities 15/18, 2/18, 1/18, respectively. The output noise $g_k \in \{0.05,\, -0.15,\, -0.45\}$ with probabilities 15/18, 2/18, 1/18, respectively. It is easily verified that $\mathbb{E}\{f_k\} = 0$ and $\mathbb{E}\{g_k\} = 0$. Fig. 2 compares the performance of four different algorithms in terms of average mean square error (MSE) as a function of γ. The algorithms considered are: i) the Kalman filter that uses the knowledge of C_k; ii) the Kalman filter that uses $C_k = C \doteq [-0.85,\, 1,\, -1]$; iii) the Kalman filter that uses as observation matrix the average C_k, namely $C_k = \mathbb{E}\{C_k\} = [-0.85\gamma,\, \gamma,\, -\gamma]$; iv) the proposed quadratic filter. Note that the first algorithm cannot be implemented when C_k is stochastic and not know, as is the case considered in this paper. Thus, it serves as an ideal reference for comparison with the other algorithms.

The performance has been evaluated averaging the MSE of the four methods over 100 realizations of $N = 2 \times 10^3$ points each, for values of γ ranging in the interval $[0.4\,1]$. The average MSE is computed as

$$\overline{\mathrm{MSE}} = \frac{1}{100}\sum_{i=1}^{100} \varepsilon_i, \quad \varepsilon_i \;=\; \frac{1}{N}\sum_{k=50}^{N} \left\| \hat{x}_k^{(i)} - x_k^{(i)} \right\|^2$$

where ε_i is the MSE of the ith realization.

Fig. 1 shows that for $\gamma = 1$ all the linear filters, namely the algorithms i)–iii), have the same performance, since $C_k = \mathbb{E}\{C_k\} = C$. In this case, the quadratic filter has better performance than the Kalman filter because both have complete knowledge of C_k but, being the noise non-Gaussian, the Kalman filter is no longer optimal. As γ decreases, the MSE of the algorithms ii)–iv) increases at roughly the same rate, and the quadratic filter exhibits uniformly better performance. It is worthwhile noting that above a certain threshold for γ, the quadratic filter outperforms the (ideal) Kalman filter of algorithm i), which relies on the perfect knowledge of C_k that is not available when C_k is random.

5. CONCLUSION

In this paper we consider the problem of state estimation for linear discrete-time non-Gaussian systems with random observation matrices. This is the model for uncertain observations resulting from losses in the propagation channel due to fading phenomena or packet dropouts. This is common in wireless sensor networks, networked control systems, or remote sensing applications. In this paper, we do not make any assumption about the distribution of the observation matrix, thus encompassing a great variety of possible scenarios. We derive the quadratic estimate of the state by means of a recursive algorithm. The solution is obtained by applying the Kalman filter to a suitably augmented system, which is fully observable. To extract the observable part of the quadratic system we exploit the knowledge of the rank of the corresponding observability matrix. This approach guarantees the internal stability of the estimation filter.

Appendix A. PROOFS OF PROPOSITIONS

A.1 Proof of Proposition 1

Proof. From the definition of V_k in (29) it follows

$$\mathbb{E}\{V_k\} = \begin{bmatrix} \mathbb{E}\{f_k\} \\ T_r\, \mathbb{E}\{f_k^{(2)}\} \end{bmatrix} = 0_{(n+r)}.$$

With respect the the autocorrelation function, consider that

$$V_k = \begin{bmatrix} I_n & 0_{n\times r} \\ 0_{r\times n} & T_r \end{bmatrix} \begin{bmatrix} f_k \\ f_k^{(2)} \end{bmatrix} = \mathrm{diag}\,(I_n, T_r) \begin{bmatrix} f_k \\ f_k^{(2)} \end{bmatrix}$$

and in De Santis et al. (1995) it is proved that the sequence $\left\{ \left[\, f_k^T \;\; f_k^{(2)T} \,\right]^T \right\}$ is uncorrelated thus implying that $\{V_k\}$ is also an uncorrelated sequence. The correlation (covariance) matrix Q_k is computed in (32).

The autocorrelation function is

$$\mathbb{E}\left\{V_h V_k^T\right\} = \begin{bmatrix} \mathbb{E}\left\{f_h f_k^T\right\} & \mathbb{E}\left\{f_h f_k^{(2)T}\right\} T_r^T \\ T_r\, \mathbb{E}\left\{f_h^{(2)} f_h^T\right\} & T_r\, \mathbb{E}\{f_h^{(2)} f_k^{(2)T}\} T_r^T \end{bmatrix}$$

with

$$\mathbb{E}\left\{f_h f_k^T\right\} = \Sigma_{f_k} \delta_{h,k} \tag{A.1}$$

$$\mathbb{E}\left\{f_h f_k^{(2)T}\right\} = \mathbb{E}\left\{ f_h \left[(I+\Pi)(A x_k \otimes f_k) + \widehat{f_k^{[2]}} \right]^T \right\}$$

$$= \mathbb{E}\left\{ f_h (x_k^T \otimes f_k^T) \right\} \left(A^T \otimes I\right)(I+\Pi) + \mathbb{E}\left\{ f_h \widehat{f_k^{[2]T}} \right\}$$

taking into account that f_h and $(x_k \otimes f_k)$ are orthogonal, we get

$$= \mathbb{E}\left\{ f_h f_k^{[2]T} \right\} = \Sigma_{f_k}^{(1,2)} \delta_{h,k} \tag{A.2}$$

$$\mathbb{E}\left\{ f_h^{(2)} f_k^T \right\} = \mathbb{E}\left\{ f_h f_k^{(2)T} \right\}^T \tag{A.3}$$

$$\mathbb{E}\left\{ f_h^{(2)} f_k^{(2)T} \right\} =$$

exploiting that $(x_h \otimes f_h)$ and $\widehat{f_k^{[2]}}$ are orthogonal, we get

$$=\left[(I+\Pi)\left(A\Sigma_{x_k}A^T \otimes \Sigma_{f_k}\right)(I+\Pi)+\widehat{\Sigma}_{f_k}^{(2,2)}\right]\delta_{h,k} \quad \text{(A.4)}$$

A.2 Proof of Proposition 2

Proof. From the definition of W_k in (29) it is easily verified that

$$\mathbb{E}\{W_k\} = 0_2.$$

The autocorrelation function is

$$\mathbb{E}\left\{W_h W_k^T\right\} = \begin{bmatrix} \mathbb{E}\{g_h^{(1)}g_k^{(1)}\} & \mathbb{E}\{g_h^{(1)}g_k^{(2)}\} \\ \mathbb{E}\{g_h^{(2)}g_k^{(1)}\} & \mathbb{E}\{g_h^{(2)}g_k^{(2)}\} \end{bmatrix}$$

with

$$\mathbb{E}\left\{g_h^{(1)}g_k^{(1)}\right\} = \mathbb{E}\left\{\left[\widehat{C_h}x_h + g_h\right]\left[\widehat{C_k}x_k + g_k\right]\right\}$$

exploiting $\widehat{C_h}x_h\widehat{C_k}x_k = \left(\widehat{C_h}\otimes\widehat{C_k}\right)(x_h \otimes x_k)$, we get

$$\begin{aligned} &= \mathbb{E}\left\{\widehat{C_h}\otimes\widehat{C_k}\right\}\mathbb{E}\{x_h\otimes x_k\} + \mathbb{E}\{g_h g_k\} \\ &= \left[\operatorname{tr}\left(\widehat{\Sigma}_{C_k}^{(1,1)}\Sigma_{x_k}\right) + m_{g_k}^{(2)}\right]\delta_{h,k} \qquad \text{(A.5)} \end{aligned}$$

$$\begin{aligned} \mathbb{E}\left\{g_h^{(1)}g_k^{(2)}\right\} &= \mathbb{E}\left\{\left[\widehat{C_h}x_h + g_h\right]\left[2C_k x_k g_k + \widehat{C_k^{[2]}}x_k^{[2]} + \widehat{g_k^2}\right]\right\} \\ &= \left[\mathbb{E}\left\{\widehat{C_k}x_k\widehat{C_k^{[2]}}x_k^{[2]}\right\} + \mathbb{E}\left\{g_k\widehat{g_k^2}\right\}\right]\delta_{h,k} \end{aligned}$$

exploiting $\widehat{C_k}x_k\widehat{C_k^{[2]}}x_k^{[2]} = \left(\widehat{C_k}\otimes\widehat{C_k^{[2]}}\right)x_k^{[3]}$, we get

$$= \left[\operatorname{tr}\left(\widehat{\Sigma}_{C_k}^{(1,2)}\Sigma_{x_k}^{(1,2)}\right) + m_{g_k}^{(3)}\right]\delta_{h,k} \qquad \text{(A.6)}$$

$$\mathbb{E}\left\{g_h^{(2)}g_k^{(1)}\right\} = \mathbb{E}\left\{g_h^{(1)}g_k^{(2)}\right\} \qquad \text{(A.7)}$$

$$\begin{aligned} \mathbb{E}\left\{g_h^{(2)}g_k^{(2)}\right\} &= 4\,\mathbb{E}\{C_h x_h C_k x_k\}\,\mathbb{E}\{g_h g_k\} \\ &+ \mathbb{E}\left\{\widehat{C_h^{[2]}}x_h^{[2]}\widehat{C_k^{[2]}}x_k^{[2]}\right\} + \mathbb{E}\left\{\widehat{g_h^2}\widehat{g_k^2}\right\} \end{aligned}$$

exploiting $C_h x_h C_k x_k = (C_h \otimes C_k)(x_h \otimes x_k)$, we get

$$\begin{aligned} &= 4\,\mathbb{E}\{C_h\otimes C_k\}\,\mathbb{E}\{x_h\otimes x_k\}\,\mathbb{E}\{g_h g_k\} \\ &+ \mathbb{E}\left\{\widehat{C_h^{[2]}}\otimes\widehat{C_k^{[2]}}\right\}\mathbb{E}\left\{x_h^{[2]}\otimes x_k^{[2]}\right\} + \mathbb{E}\left\{\widehat{g_h^2}\widehat{g_k^2}\right\} \\ &= \left[4m_{g_k}^{(2)}\operatorname{tr}\left(\Sigma_{C_k}^{(1,1)}\Sigma_{x_k}\right) + \operatorname{tr}\left(\widehat{\Sigma}_{C_k}^{(2,2)}\Sigma_{x_k}^{(2,2)}\right)\right. \\ &\left. + m_{g_k}^{(4)} - m_{g_k}^{(2)}m_{g_k}^{(2)}\right]\delta_{h,k} \qquad \text{(A.8)} \end{aligned}$$

REFERENCES

Bellman, R. (1997). *Introduction to Matrix Analysis.* Society for Industrial and Applied Mathematics, 2nd edition.

Carravetta, F., Germani, A., and Raimondi, M. (1996). Polynomial filtering for linear discrete time non-Gaussian systems. *SIAM J. Control Optim.*, 34(5), 1666–1690.

De Koning, W.L. (1984). Optimal estimation of linear discrete-time systems with stochastic parameters. *Automatica*, 20(1), 113–115.

De Santis, A., Germani, A., and Raimondi, M. (1995). Optimal quadratic filtering of linear discrete-time non-Gaussian systems. *IEEE Trans. Autom. Control*, 40(7), 1274–1278.

Fasano, A., Germani, A., and Monteriù, A. (2013). Reduced-order quadratic Kalman-like filtering of non-Gaussian systems. *IEEE Trans. Autom. Control*, 58(7), 1744–1757.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications, Vol. 1.* Wiley, 3rd edition.

Hadidi, M.T. and Schwartz, S.C. (1979). Linear recursive state estimators under uncertain observations. *IEEE Trans. Autom. Control*, AC-24(6), 944–948.

Hespanha, J., Naghshtabrizi, P., and Xu, Y. (Jan.). A survey of recent results in networked control systems. *Proc. IEEE*, 95(1), 138–162.

Horn, R.A. and Johnson, C.R. (1994). *Topics in Matrix Analysis.* Cambridge University Press.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(Series D), 35–45.

Luo, D. and Zhu, Y. (2008). Applications of random parameter matrices Kalman filtering in uncertain observation and multi-model systems. *Proc. 17th World Congr. Int. Fed. Automat. Control (IFAC)*, 12516–12521.

Nahi, N.E. (1969). Optimal recursive estimation with uncertain observations. *IEEE Trans. Inf. Theory*, IT-15(7), 457–462.

Nakamori, S., Caballero-Águila, R., Hermoso-Carazo, A., and Linares-Pérez, J. (2003a). New design of estimators using covariance information with uncertain observations in linear discrete-time systems. *Applied Mathematics and Computation*, 135, 429 – 441.

Nakamori, S., Caballero-Águila, R., Hermoso-Carazo, A., and Linares-Pérez, J. (2003b). Second-order polynomial estimators from uncertain observations using covariance information. *Applied Mathematics and Computation*, 143, 319 – 338.

Paige, C. (1981). Properties of numerical algorithms related to computing controllability. *IEEE Trans. Autom. Control*, 26(1), 130–138.

Rosenbrock, H. (1970). *State-Space and Multivariable Theory.* Thomas Nelson.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M.I., and Sastry, S.S. (2004). Kalman filtering with intermittent observations. *IEEE Trans. Autom. Control*, 49(9), 1453–1464.

Spall, J.C. (1985). Validation of state-space models from a single realization of non-Gaussian measurements. *IEEE Trans. Autom. Control*, 30(12), 1212–1214.

Spall, J.C. (2003). Estimation via Markov chain Monte Carlo. *IEEE Control Syst. Mag.*, 23(2), 34–45.

Spall, J.C. and Wall, K. (1984). Asymptotic distribution theory for the Kalman filter state estimator. *Communications in Statistics - Theory and Methods*, 13(16), 1981–2003.

Wu, H. and Chen, G. (1993). Suboptimal Kalman filtering for linear systems with non-Gaussian noise. *Approximate Kalman Filtering*, 2, 113–136.

Zakai, M. (1969). On the optimal filtering of diffusion processes. *Probability Theory and Related Fields*, 11(3), 230–243.