To appear in the *International Journal of Systems Science*
Vol. 00, No. 00, Month 20XX, 1–20

Optimal linear filter design for systems with correlation in the measurement matrices and noises: recursive algorithm and applications

J. Linares-Pérez[a], * R. Caballero-Águila[b] and I. García-Garrido[a]

[a] *Dpto. de Estadística e I.O., Universidad de Granada, 18071 Granada, Spain;*
[b] Dpto. de Estadística e I.O., Universidad de Jaén, 23071 Jaén, Spain

(*v4.0 released May 2013*)

This paper addresses the optimal least-squares linear estimation problem for a class of discrete-time stochastic systems with random parameter matrices and correlated additive noises. The system presents the following main features: (1) one-step correlated and cross-correlated random parameter matrices in the observation equation are assumed; (2) the process and measurement noises are one-step autocorrelated and two-step cross-correlated. Using an innovation approach and these correlation assumptions, a recursive algorithm with a simple computational procedure is derived for the optimal linear filter. As a significant application of the proposed results, the optimal recursive filtering problem in multi-sensor systems with missing measurements and random delays can be addressed. Numerical simulation examples are used to demonstrate the feasibility of the proposed filtering algorithm, which is also compared with other filters that have been proposed.

Keywords: random parameter matrices; correlated noises; optimal least-squares estimation; innovation approach

1. Introduction

The least-squares (LS) state estimation problem in discrete-time linear systems from noise measurements has been widely considered, due to its applicability in many practical situations. The Kalman filter provides a recursive algorithm for the optimal LS estimator when the model parameter matrices are deterministic and the additive white noises and the initial state are Gaussian and mutually independent. However, many real systems do not meet these requirements and new filtering algorithms have been reported for models representing the relationship between the unknown state and the observable variables and under different assumptions for the noise processes.

In recent decades, the filtering problem in multi-sensor systems, where sensor networks are used to obtain all available information on the system state, has become an issue of great interest for researchers. In data transmission, unreliable network characteristics can produce random sensor delays, multiple packet dropouts and uncertain observations (missing measurements). Due to these random uncertainties, standard observation models are not appropriate and estimation algorithms cannot be derived directly from Kalman filter theory. Accordingly, new algorithms are needed, and many research papers have been presented concerning the state estimation problem in multi-sensor systems with some of the aforementioned uncertainties (see Moayedi et al. (2010); Ma and Sun (2011); Sun and Xiao (2013), among others).

In systems with uncertain observations, besides the usual additive noise, the observation equation includes a multiplicative noise; hence, these systems are a special case of random measure-

*Corresponding author. Email: jlinares@ugr.es

ment matrices. Moreover, systems with random sensor delays or multiple packet dropouts are transformed into systems with random measurement matrices in Sahebsara et al. (2007). Systems with random state transition matrices can be used, for example, to describe randomly variant dynamic systems with multiple models (Luo et al. 2008) or linear systems with state-dependent multiplicative noise (Feng et al. 2013). Discrete-time systems with random state transition and measurement parameter matrices also arise in areas such as the digital control of chemical processes, systems with human operators, economic systems and stochastically sampled digital control systems (De Koning 1984).

In De Koning (1984) and Luo et al. (2008), the optimal linear filtering problem in systems with independent random state transition and measurement matrices is addressed by transforming the original system into one with deterministic parameter matrices and state-dependent process and measurement noises, to which the Kalman filter is applied. Although De Koning (1984) applies the Kalman filter without providing any theoretical justification, Luo et al. (2008) shows that under mild conditions, the transformed system satisfies the Kalman filter requirements and, hence, optimal linear estimators are derived for systems with independent random parameter matrices.

However, in many practical situations the random parameter matrices are not independent but correlated; for example, when random sensor delays and/or multiple packet dropouts are converted into observation models with random measurement matrices (Sahebsara et al. 2007), or when a nonlinear system is linearized around the random state estimate to apply the extended Kalman filter (for other realistic systems and backgrounds where the model parameter matrices are random and correlated, see Luo et al. (2012) and Shen et al. (2012)). In Shen et al. (2011), systems with deterministic transition matrices and one-step correlated measurement matrices are considered, and the optimal recursive state estimation is derived by converting the observation equation into one with deterministic measurement matrices and applying the optimal Kalman filter for the case of one-step correlated measurement noise. In addition, a specific class of systems, where both the state transition and the measurement matrices are one-step moving average matrix sequences driven by a common independent zero-mean parameter sequence, is considered in the latter paper.

In the above-mentioned papers, although the noises of the transformed system with deterministic matrices depend on the system state and therefore can be correlated, the original system noises are assumed to be independent white processes. This assumption can be restrictive in many real-world problems in which correlation and cross-correlation of the noises may be present. In systems with deterministic (state transition and measurement) matrices and with correlated and cross-correlated noises, the estimation problem has aroused significant research interest recently (see Feng et al. (2013); Caballero-Águila et al. (2013a) and references therein).

In view of the above considerations, this study focuses on the optimal LS linear filtering problem in systems with random parameter matrices and autocorrelated and cross-correlated noises, assuming independent random state transition matrices and one-step correlated and cross-correlated random parameter matrices in the observation equation. The proposed optimal LS linear recursive filtering algorithm can be applied to two significant classes of systems: *(i)* multi-sensor systems with missing measurements when the missing measurement phenomenon in each sensor is described by different sequences of correlated scalar random variables with arbitrary discrete probability distribution over the interval [0,1] (see Section 4); and *(ii)* multi-sensor systems with correlated random delays in the observations (see Section 5). In both cases, correlated and cross-correlated noises are considered. *This paper makes a substantial and novel contribution in two respects: (1) Unlike most existing results with random parameter matrices, in which a system transformation is carried out, the proposed optimal LS linear recursive filtering algorithm is obtained by using an innovation approach, without requiring any system transformation and, moreover, in which noise correlation is considered; (2) multi-sensor systems with missing and randomly delayed measurements can be considered as particular cases of the current random measurement matrices model and, hence, the proposed filter can be applied to these* *kind of multi-sensor systems. Besides these advantages, the filtering algorithm described is very simple computationally.*

The rest of this paper is organized as follows. Section 2 describes the system model with random state transition and measurement matrices, and autocorrelated and cross-correlated noises. In addition, some properties of the state and noise processes derived from the correlation assumptions are specified. In Section 3, by using an innovation approach, a recursive algorithm for the optimal LS linear filter is obtained. In Sections 4 and 5, applications to multi-sensor systems with missing and randomly delayed measurements, respectively, are considered. In both sections, a numerical simulation example is presented to show the effectiveness of the proposed recursive filtering algorithm. Finally, some conclusions are drawn in Section 6.

Notation: The notation used throughout the paper is standard. For any matrix A, the notation symbols A^T and A^{-1} represent its transpose and inverse, respectively; $\mathbb{R}^n$ denotes the n-dimensional Euclidean space and $\mathbb{R}^{m\times n}$ is the set of $m \times n$ real matrices. The shorthand $Diag(A_1, \ldots, A_m)$ denotes a block diagonal matrix with matrices $A_1, \ldots, A_m$, and $[A_1 \mid \cdots \mid A_m]$ denotes a partitioned matrix into sub-matrices $A_1, \ldots, A_m$. If a matrix dimension is not explicitly stated, it is assumed to be compatible for algebraic operations. I and 0 represent the identity and zero matrices of appropriate dimensions. $\delta_{k,s}$ is the Kronecker delta function, which is equal to one, if $k = s$, and zero otherwise. $\circ$ denotes the Hadamard product. Moreover, for arbitrary random vectors X and Y, we denote $Cov[X, Y] = E[(X - E[X]) (Y - E[Y])^T]$ and $Cov[X] = Cov[X, X]$, where $E[\cdot]$ stands for the mathematical expectation operator.

2. Discrete-time system model with random parameter matrices

Our aim in this paper is to address the optimal LS linear filtering problem in a class of discrete-time stochastic systems with random parameter matrices (independent random transition matrices and one-step correlated and cross-correlated matrices in the observation equation) and autocorrelated and cross-correlated noises. In this section, the system model is described and the statistical properties of the initial state, the random parameter matrices and the noise processes are identified.

Consider a class of discrete-time linear stochastic systems whose $n-$dimensional state process, $\{x_k\}_{k\geq 0}$, is perturbed by $n \times n$ random parameter matrices $\{F_k\}_{k\geq 0}$ and by an additive process noise $\{w_k\}_{k\geq 0}$; specifically, the state evolution is given by:

$$x_{k+1} = F_k x_k + w_k, \quad k \geq 0. \tag{1}$$

The measurements of the state are described by the following observation equation:

$$y_k = H_k x_k + B_k v_k, \quad k \geq 1, \tag{2}$$

where $\{y_k\}_{k\geq 1}$ is the $r-$dimensional observation process; the measurement matrices, $\{H_k\}_{k\geq 1}$, are $r \times n$ random parameter matrices; $\{B_k\}_{k\geq 1}$ are $r \times m$ random parameter matrices and the additive measurement noise, $\{v_k\}_{k\geq 1}$, is an $m-$dimensional process.

It is known that, if the state x_k and the observations $y_1, \ldots, y_k$ have finite second-order moments, then the optimal LS linear filter of x_k is the orthogonal projection of the vector x_k onto $\mathcal{L}(y_1, \ldots, y_k)$, i.e., the space of n-dimensional random variables obtained as linear transformations of $y_1, \ldots, y_k$. The hypotheses about the processes in (1) and (2) that guarantee the existence of the second-order moments of the vectors $y_1, \ldots, y_k$, as well as the correlation assumptions of the noise processes and the random parameter matrices in the observation equation are as follows:

(a) The initial state x_0 is a random vector with $E[x_0] = \overline{x}_0$ and $Cov[x_0] = P_0$, and it is independent of the random parameter matrices and noise processes.

(b) The random parameter matrices $\{F_k\}_{k\geq 0}$, $\{H_k\}_{k\geq 1}$ and $\{B_k\}_{k\geq 1}$ satisfy:

$$\begin{array}{l} E[F_k]=\overline{F}_k, \quad E[H_k]=\overline{H}_k, \quad E[B_k]=\overline{B}_k, \\ Cov[f_{ij}^k, f_{pq}^s]=C_{f_{ij}^k f_{pq}^k}\delta_{k,s}, \\ Cov[h_{ij}^k, h_{pq}^s]=C_{h_{ij}^k h_{pq}^k}\delta_{k,s}+C_{h_{ij}^k h_{pq}^s}\delta_{k,s-1}+C_{h_{ij}^k h_{pq}^s}\delta_{k,s+1}, \\ Cov[b_{ij}^k, b_{pq}^s]=C_{b_{ij}^k b_{pq}^k}\delta_{k,s}+C_{b_{ij}^k b_{pq}^s}\delta_{k,s-1}+C_{b_{ij}^k b_{pq}^s}\delta_{k,s+1}, \\ Cov[h_{ij}^k, b_{pq}^s]=C_{h_{ij}^k b_{pq}^k}\delta_{k,s}+C_{h_{ij}^k b_{pq}^s}\delta_{k,s-1}+C_{h_{ij}^k b_{pq}^s}\delta_{k,s+1}, \end{array}$$

where f_{ij}^k, h_{ij}^k and b_{ij}^k denote the $(i,j)-$th entries of matrices F_k, H_k and B_k, respectively. The mean matrices $\overline{F}_k$, $\overline{H}_k$ and $\overline{B}_k$ are known matrices $\forall k$, and $C_{f_{ij}^k f_{pq}^k}$, $C_{h_{ij}^k h_{pq}^s}$, $C_{b_{ij}^k b_{pq}^s}$ and $C_{h_{ij}^k b_{pq}^s}$, the covariances of the entries of the system random parameter matrices, are also assumed to be known $\forall k$ and $\forall s=k-1,k,k+1$.

(c) The process noise, $\{w_k\}_{k\geq 0}$, and the measurement noise, $\{v_k\}_{k\geq 1}$, are zero-mean sequences with the following covariances and cross-covariances:

$$\begin{array}{l} Cov[w_k, w_s]=Q_{k,k}\delta_{k,s}+Q_{k,s}\delta_{k,s-1}+Q_{k,s}\delta_{k,s+1}, \\ Cov[v_k, v_s]=R_{k,k}\delta_{k,s}+R_{k,s}\delta_{k,s-1}+R_{k,s}\delta_{k,s+1}, \\ Cov[w_k, v_s]=S_{k,k}\delta_{k,s}+S_{k,s}\delta_{k,s-1}+S_{k,s}\delta_{k,s-2}. \end{array}$$

(d) Independence assumptions:
- $\{F_k\}_{k\geq 0}$ is independent of $(\{H_k\}_{k\geq 1}, \{B_k\}_{k\geq 1}, \{w_k\}_{k\geq 0}, \{v_k\}_{k\geq 1})$.
- $(\{H_k\}_{k\geq 1}, \{B_k\}_{k\geq 1})$ is independent of $(\{F_k\}_{k\geq 0}, \{w_k\}_{k\geq 0}, \{v_k\}_{k\geq 1})$.

Remark 1: correlation of the noise processes. The correlation hypothesis *(c)* of the process noise and the measurement noise is the same as those given in Feng et al. (2013) and Caballero-Águila et al. (2013a). Specifically, both noise processes are correlated at consecutive sampling times and independent otherwise, and the measurement noise vector v_k is correlated with the noise vectors w_s, for $s=k,k-1,k-2$, and independent otherwise. Systems with only finite-step correlated process noise or multi-step correlated process and measurement noise are considered in Song et al. (2008), Fu et al. (2008) and Feng et al. (2011), among others.

As a consequence of the noise correlation assumptions, it is easy to see that:

- The vectors w_k and y_k are correlated, with

$$\mathcal{W}_k := E\left[w_k y_k^T\right]=Q_{k,k-1}\overline{H}_k^T+S_{k,k}\overline{B}_k^T, \quad k\geq 1. \tag{3}$$

- The state vector x_k is correlated with the observation noise vectors v_k and v_{k-1}, with

$$\begin{array}{c} \mathcal{E}_{k,k} := E[x_k v_k^T]=\overline{F}_{k-1}S_{k-2,k}+S_{k-1,k}, \quad k\geq 2; \qquad \mathcal{E}_{1,1}=S_{0,1}, \\ \mathcal{E}_{k,k-1} := E[x_k v_{k-1}^T]=\overline{F}_{k-1}\mathcal{E}_{k-1,k-1}+S_{k-1,k-1}, \quad k\geq 2. \end{array} \tag{4}$$

Remark 2: correlation of the random parameter matrices in the observation equation. Besides considering autocorrelated and cross-correlated noises (assumption (c)), the correlation assumption (b) of the random parameter matrices $\{H_k\}_{k\geq 1}$ and $\{B_k\}_{k\geq 1}$ in the observation equation is the main difference between the current model and the model in Luo et al. (2008), where $\{H_k\}_{k\geq 1}$ is assumed to be a sequence of independent random parameter matrices, and the observation noise is not multiplied by random parameter matrices. The correlation of the measurement matrices $\{H_k\}_{k\geq 1}$ at consecutive sampling times allows us to apply the results proposed in this paper to multi-sensor systems with correlated missing measurements (see Section 4). In addition, the correlation and cross-correlation of the random parameter matrices $\{H_k\}_{k\geq 1}$ and $\{B_k\}_{k\geq 1}$ allow us to consider multi-sensor systems with correlated randomly delayed measurements as a particular case of the current study (see Section 5).

As a consequence of the correlation assumptions of the noises and random parameter matrices, using the conditional expectation properties, and denoting $\widetilde{H}_k = H_k - \overline{H}_k$ and $\widetilde{B}_k = B_k - \overline{B}_k$, it can be seen that the vector $B_k v_k$ is correlated with the observation vector y_{k-1}, and

$$\begin{aligned}\mathcal{V}_{k,k-1} := E\left[B_k v_k y_{k-1}^T\right] = \overline{B}_k\left(\overline{H}_{k-1}S_{k-2,k} + \overline{B}_{k-1}R_{k-1,k}\right)^T \\ + E\left[\widetilde{B}_k S_{k-2,k}^T \widetilde{H}_{k-1}^T\right] + E\left[\widetilde{B}_k R_{k,k-1}\widetilde{B}_{k-1}^T\right], \quad k \geq 2.\end{aligned} \tag{5}$$

The matrix $E\left[\widetilde{B}_k S_{k-2,k}^T \widetilde{H}_{k-1}^T\right]$ is yielded by both the cross-correlation of the noise processes and that of the random parameter matrices $\{H_k\}_{k\geq1}$ and $\{B_k\}_{k\geq1}$, while the matrix $E\left[\widetilde{B}_k R_{k,k-1}\widetilde{B}_{k-1}^T\right]$ arises because of the one-step correlation of the measurement noise and also that of the matrices $\{B_k\}_{k\geq1}$. From assumption *(b)*, the $(p,q)-$th entries of these matrices are obtained by:

$$\begin{aligned}\left(E[\widetilde{B}_k S_{k-2,k}^T \widetilde{H}_{k-1}^T]\right)_{pq} &= \sum_{j=1}^{m}\sum_{i=1}^{n} C_{b_{qj}^k h_{pi}^{k-1}}\left(S_{k-2,k}^T\right)_{ji} \quad (p,q=1,2,\ldots,r),\\ \left(E[\widetilde{B}_k R_{k,k-1}\widetilde{B}_{k-1}^T]\right)_{pq} &= \sum_{j=1}^{m}\sum_{i=1}^{m} C_{b_{qj}^k b_{pi}^{k-1}}\left(R_{k,k-1}\right)_{ji} \quad (p,q=1,2,\ldots,r).\end{aligned}$$

Remark 3: state transition equation. Linear discrete-time systems with random state transition matrices, $\{F_k\}_{k\geq0}$, have important applications; for example, they can be used to describe randomly variant dynamic systems with multiple models (Luo et al. 2008) or linear systems with state-dependent multiplicative noise (Feng et al. 2013). Furthermore, bilinear stochastic systems (Carravetta et al. 1997) can be reduced to models with random transition matrices.

After denoting $\widetilde{F}_k = F_k - \overline{F}_k$, from the state equation (1) and the conditional expectation properties, it is easy to deduce that $\mathcal{D}_{k+1} = E[x_{k+1}x_{k+1}^T]$ is recursively calculated by:

$$\begin{aligned}&\mathcal{D}_{k+1} = \overline{F}_k\mathcal{D}_k\overline{F}_k^T + E[\widetilde{F}_k\mathcal{D}_k\widetilde{F}_k^T] + Q_{k,k} + \overline{F}_kQ_{k-1,k} + Q_{k,k-1}\overline{F}_k^T, \quad k\geq1;\\ &\mathcal{D}_1 = \overline{F}_0\mathcal{D}_0\overline{F}_0^T + E[\widetilde{F}_0\mathcal{D}_0\widetilde{F}_0^T] + Q_{0,0}, \qquad \mathcal{D}_0 = P_0 + \overline{x}_0\overline{x}_0^T,\end{aligned} \tag{6}$$

where, from assumption *(b)*, the $(p,q)-$th entry of the matrix $E[\widetilde{F}_k\mathcal{D}_k\widetilde{F}_k^T]$ is obtained by

$$\left(E[\widetilde{F}_k\mathcal{D}_k\widetilde{F}_k^T]\right)_{pq} = \sum_{j=1}^{n}\sum_{i=1}^{n} C_{f_{qj}^k f_{pi}^k}\left(\mathcal{D}_k\right)_{ji} \quad (p,q=1,2,\ldots,n).$$

Also, from the state equation (1), it is immediately clear that $\mathcal{G}_{k+1,k} = E[x_{k+1}x_k^T]$ satisfies

$$\mathcal{G}_{k+1,k} = \overline{F}_k\mathcal{D}_k + Q_{k,k-1}, \quad k\geq1; \qquad \mathcal{G}_{1,0} = \overline{F}_0\mathcal{D}_0. \tag{7}$$

3. Optimal LS linear estimation problem

Given the observations up to time k, $\{y_1,\ldots,y_k\}$, our aim is to derive a recursive algorithm for the optimal LS linear filter, $\widehat{x}_{k/k}$, of the state x_k. Since $\widehat{x}_{k/k}$ is the orthogonal projection of x_k onto the space $\mathcal{L}(y_1,\ldots,y_k)$ of linear transformations of $y_1,\ldots,y_k$, and these observations are generally non-orthogonal vectors, an innovation approach will be used. This approach considerably simplifies the algorithm derivation, because the innovation process is a white noise. The innovation approach is based on the Gram-Schmidt orthogonalization procedure by means of which the observation process $\{y_k\}_{k\geq 1}$ is transformed into an equivalent process (*innovation process*) $\{\mu_k\}_{k\geq 1}$, equivalent in the sense that $\mathcal{L}(\mu_1,\dots,\mu_k) = \mathcal{L}(y_1,\dots,y_k)$; that is, each set $\{\mu_1,\dots,\mu_k\}$ spans the same linear subspace as $\{y_1,\dots,y_k\}$.

The innovation at time k is defined as $\mu_k = y_k - \widehat{y}_{k/k-1}$, where $\widehat{y}_{k/k-1}$, the one-stage LS linear predictor of y_k, is the projection of y_k onto $\mathcal{L}(\mu_1,\dots,\mu_{k-1})$. The orthogonality property allows us to find the projection by projecting onto each of the previous orthogonal vectors separately; that is,

$$\widehat{y}_{k/k-1} = \sum_{i=1}^{k-1} E[y_k\mu_i^T](E[\mu_i\mu_i^T])^{-1}\mu_i, \quad k\geq 2; \qquad \widehat{y}_{1/0} = \overline{H}_1\widehat{x}_{1/0}. \tag{8}$$

Similarly, by denoting $\mathcal{X}_{k,i} = E[x_k\mu_i^T]$ and $\Pi_i = E[\mu_i\mu_i^T]$, a general expression for the optimal LS linear filter, $\widehat{x}_{k/k}$, as a linear combination of the innovations is obtained; namely,

$$\widehat{x}_{k/k} = \sum_{i=1}^{k} \mathcal{X}_{k,i}\Pi_i^{-1}\mu_i, \quad k\geq 1; \quad \widehat{x}_{0/0} = \overline{x}_0,$$

and, the following expression for the filter, $\widehat{x}_{k/k}$, in terms of the predictor, $\widehat{x}_{k/k-1}$, is obvious:

$$\widehat{x}_{k/k} = \widehat{x}_{k/k-1} + \mathcal{X}_{k,k}\Pi_k^{-1}\mu_k, \quad k\geq 1; \quad \widehat{x}_{0/0} = \overline{x}_0. \tag{9}$$

Next, we obtain the state predictor $\widehat{x}_{k/k-1}$, the innovation μ_k and its covariance matrix Π_k, and the matrix $\mathcal{X}_{k,k}$, which together with (9) will constitute the proposed recursive linear filtering algorithm.

3.1 *State predictor* $\widehat{x}_{k/k-1}$

In systems with random parameter matrices and uncorrelated additive white noises (Luo et al. 2008), the one-stage state predictor is calculated as $\widehat{x}_{k/k-1} = \overline{F}_{k-1}\widehat{x}_{k-1/k-1}$; this is because the uncorrelation assumption of the noises guarantees that $\widehat{w}_{k-1/k-1} = 0$. However, this is not true for the problem at hand, where the noise estimator $\widehat{w}_{k-1/k-1}$ must be taken into account in order to derive the one-stage state predictor. From the Orthogonal Projection Lemma (OPL), we have

$$\widehat{x}_{k/k-1} = \overline{F}_{k-1}\widehat{x}_{k-1/k-1} + \widehat{w}_{k-1/k-1}, \quad k\geq 1,$$

and hence, an expression for the noise filter $\widehat{w}_{k/k}$ is necessary. Taking into account that w_k is independent of $\mu_1,\dots,\mu_{k-1}$ and $\widehat{y}_{k/k-1}$, we have,

$$\widehat{w}_{k/k} = \sum_{i=1}^{k} E[w_k\mu_i^T]\Pi_i^{-1}\mu_i = E[w_k\mu_k^T]\Pi_k^{-1}\mu_k = E[w_ky_k^T]\Pi_k^{-1}\mu_k, \quad k\geq 1; \quad \widehat{w}_{0/0} = 0.$$

Therefore, the state predictor, $\widehat{x}_{k/k-1}$, satisfies

$$\widehat{x}_{k/k-1} = \overline{F}_{k-1}\widehat{x}_{k-1/k-1} + \mathcal{W}_{k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k\geq 2; \qquad \widehat{x}_{1/0} = \overline{F}_0\widehat{x}_{0/0}, \tag{10}$$

where $\mathcal{W}_k$ is given by (3).

3.2 *Prediction $P_{k/k-1}$ and filtering $P_{k/k}$ error covariance matrices*

From (1) and (10), it is easy to see that the prediction error covariance matrix, $P_{k/k-1}$, satisfies

$$\begin{aligned} P_{k/k-1} &= \overline{F}_{k-1}P_{k-1/k-1}\overline{F}_{k-1}^T + E[\widetilde{F}_{k-1}\mathcal{D}_{k-1}\widetilde{F}_{k-1}^T] + Q_{k-1,k-1} + \overline{F}_{k-1}\mathcal{J}_{k-1} + \mathcal{J}_{k-1}^T\overline{F}_{k-1}^T \\ &\quad - \mathcal{W}_{k-1}\Pi_{k-1}^{-1}\mathcal{W}_{k-1}^T, \quad k \geq 2; \\ P_{1/0} &= \overline{F}_0 P_{0/0}\overline{F}_0^T + E[\widetilde{F}_0\mathcal{D}_0\widetilde{F}_0^T] + Q_{0,0}, \end{aligned} \tag{11}$$

where, using (9), it is clear that $\mathcal{J}_k = E\left[\left(x_k - \widehat{x}_{k/k}\right)w_k^T\right]$ is calculated by

$$\mathcal{J}_k = Q_{k-1,k} - \mathcal{X}_{k,k}\Pi_k^{-1}\mathcal{W}_k^T, \quad k \geq 1. \tag{12}$$

Again, from (9), the filtering error covariance matrix, $P_{k/k}$, is given by

$$P_{k/k} = P_{k/k-1} - \mathcal{X}_{k,k}\Pi_k^{-1}\mathcal{X}_{k,k}^T, \quad k \geq 1; \qquad P_{0/0} = P_0. \tag{13}$$

3.3 *Innovation $\mu_k = y_k - \widehat{y}_{k/k-1}$*

In Luo et al. (2008), the one-stage observation predictor is calculated as $\widehat{y}_{k/k-1} = \overline{H}_k\widehat{x}_{k/k-1}$; this is because the uncorrelation assumption of the noises guarantees that $\widehat{v}_{k/k-1} = 0$. However, due to the correlation assumptions of the measurement matrices, *(b)*, and the noise processes, *(c)*, this is not true for the problem at hand and both the correlation of H_{k-1} and H_k and the noise estimator $\widehat{v}_{k/k-1}$, must be taken into account in deriving the predictor $\widehat{y}_{k/k-1}$.

Therefore, to obtain the innovation $\mu_k = y_k - \widehat{y}_{k/k-1}$, it is necessary to find a new expression for $\widehat{y}_{k/k-1}$. For this purpose, taking into account (8), we first calculate

$$E[y_k\mu_i^T] = E[H_k x_k\mu_i^T] + E[B_k v_k\mu_i^T] = \begin{cases} \overline{H}_k\mathcal{X}_{k,i}, & i \leq k-2, \\ E[H_k x_k\mu_{k-1}^T] + \mathcal{V}_{k,k-1}, & i = k-1, \end{cases} \tag{14}$$

and, substituting the expectations (14) in (8), we obtain

$$\widehat{y}_{k/k-1} = \overline{H}_k\sum_{i=1}^{k-1}\mathcal{X}_{k,i}\Pi_i^{-1}\mu_i + \left(E[H_k x_k\mu_{k-1}^T] + \mathcal{V}_{k,k-1} - \overline{H}_k\mathcal{X}_{k,k-1}\right)\Pi_{k-1}^{-1}\mu_{k-1}.$$

Now, from the conditional expectation properties, we obtain that

$$\begin{aligned} E[H_k x_k\mu_{k-1}^T] - \overline{H}_k\mathcal{X}_{k,k-1} &= E[\widetilde{H}_k x_k\mu_{k-1}^T] = E[\widetilde{H}_k x_k y_{k-1}^T] \\ &= E[\widetilde{H}_k\mathcal{G}_{k,k-1}\widetilde{H}_{k-1}^T] + E[\widetilde{H}_k\mathcal{E}_{k,k-1}\widetilde{B}_{k-1}^T], \end{aligned}$$

where $\mathcal{G}_{k,k-1}$ and $\mathcal{E}_{k,k-1}$ are given in (7) and (4), respectively.

Hence, it is concluded that the one-stage observation predictor satisfies

$$\widehat{y}_{k/k-1} = \overline{H}_k\widehat{x}_{k/k-1} + \Psi_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k \geq 1, \tag{15}$$

where

$$\Psi_{k,k-1} = E[\widetilde{H}_k\mathcal{G}_{k,k-1}\widetilde{H}_{k-1}^T] + E[\widetilde{H}_k\mathcal{E}_{k,k-1}\widetilde{B}_{k-1}^T] + \mathcal{V}_{k,k-1}, \quad k \geq 2; \qquad \Psi_{1,0} = 0. \tag{16}$$

It can be observed that the matrices $E[\widetilde{H}_k \mathcal{G}_{k,k-1} \widetilde{H}_{k-1}^T]$ and $E[\widetilde{H}_k \mathcal{E}_{k,k-1}^T \widetilde{B}_{k-1}^T]$ are yielded by the correlation of the random parameter matrices of the observation equation at consecutive sampling times. From *(b)*, the $(p,q)-$th entries of these matrices are obtained by

$$\begin{aligned}\left(E[\widetilde{H}_k \mathcal{G}_{k,k-1} \widetilde{H}_{k-1}^T]\right)_{pq} &= \sum_{j=1}^{n}\sum_{i=1}^{n} C_{h_{qj}^k h_{pi}^{k-1}} \left(\mathcal{G}_{k,k-1}\right)_{ji} \quad (p,q=1,2,\ldots,r),\\ \left(E[\widetilde{H}_k \mathcal{E}_{k,k-1}^T \widetilde{B}_{k-1}^T]\right)_{pq} &= \sum_{j=1}^{n}\sum_{i=1}^{m} C_{h_{qj}^k b_{pi}^{k-1}} \left(\mathcal{E}_{k,k-1}^T\right)_{ji} \quad (p,q=1,2,\ldots,r).\end{aligned}$$

Hence, the innovation μ_k is obtained as a linear combination of the new observation, the state predictor and the previous innovation:

$$\mu_k = y_k - \overline{H}_k \widehat{x}_{k/k-1} - \Psi_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k \geq 1. \tag{17}$$

3.4 *Matrix* $\mathcal{X}_{k,k} = E[x_k \mu_k^T]$

Next, an expression for the matrix $\mathcal{X}_{k,k} = E[x_k \mu_k^T] = E[x_k y_k^T] - E[x_k \widehat{y}_{k/k-1}^T]$ is derived. From (2) and (4), it is clear that

$$E[x_k y_k^T] = \mathcal{D}_k \overline{H}_k^T + \mathcal{E}_{k,k}\overline{B}_k^T, \quad k \geq 1.$$

From (15) and since, from the OPL, $E[x_k \widehat{x}_{k/k-1}^T] = \mathcal{D}_k - P_{k/k-1}$, we obtain:

$$E[x_k \widehat{y}_{k/k-1}^T] = \left(\mathcal{D}_k - P_{k/k-1}\right)\overline{H}_k^T + \mathcal{X}_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T, \quad k \geq 1,$$

where $\mathcal{X}_{k,k-1} = E\left[x_k \mu_{k-1}^T\right]$ satisfies

$$\mathcal{X}_{k,k-1} = \overline{F}_{k-1}\mathcal{X}_{k-1,k-1} + \mathcal{W}_{k-1}, \quad k \geq 2. \tag{18}$$

By subtracting the above expectations, the following expression for $\mathcal{X}_{k,k}$ is derived

$$\mathcal{X}_{k,k} = P_{k/k-1}\overline{H}_k^T + \mathcal{E}_{k,k}\overline{B}_k^T - \mathcal{X}_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T, \quad k \geq 1. \tag{19}$$

3.5 *Innovation covariance matrix* $\Pi_k = E[\mu_k \mu_k^T]$

Finally, we obtain an expression for $\Pi_k = E[\mu_k \mu_k^T] = E[y_k y_k^T] - E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T]$. From (2) and again using the conditional expectation properties, we have

$$\begin{aligned}E[y_k y_k^T] =&\overline{H}_k \mathcal{D}_k \overline{H}_k^T + E[\widetilde{H}_k \mathcal{D}_k \widetilde{H}_k^T] + \overline{B}_k R_{k,k}\overline{B}_k^T + E[\widetilde{B}_k R_{k,k}\widetilde{B}_k^T]\\ &+ \overline{H}_k \mathcal{E}_{k,k}\overline{B}_k^T + E[\widetilde{H}_k \mathcal{E}_{k,k}\widetilde{B}_k^T] + \overline{B}_k \mathcal{E}_{k,k}^T \overline{H}_k^T + E[\widetilde{B}_k \mathcal{E}_{k,k}^T \widetilde{H}_k^T],\end{aligned}$$

where $\mathcal{D}_k$ and $\mathcal{E}_{k,k}$ are given in (6) and (4), respectively.

Using (15) and since $E[\widehat{x}_{k/k-1}\mu_{k-1}^T] = E[x_k \mu_{k-1}^T] = \mathcal{X}_{k,k-1}$, the following identity holds:

$$\begin{aligned}E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T] =& \overline{H}_k(\mathcal{D}_k - P_{k/k-1})\overline{H}_k^T + \Psi_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T\\ &+ \overline{H}_k \mathcal{X}_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T + \Psi_{k,k-1}\Pi_{k-1}^{-1}\mathcal{X}_{k,k-1}^T\overline{H}_k^T.\end{aligned}$$

From the above expectations, using (19) and after some manipulations, the following expression for the innovation covariance matrix Π_k is obtained:

$$\begin{aligned}\Pi_k =& E[\widetilde{H}_k\mathcal{D}_k\widetilde{H}_k^T] + E[\widetilde{B}_kR_{k,k}\widetilde{B}_k^T] + E[\widetilde{B}_k\mathcal{E}_{k,k}^T\widetilde{H}_k^T] + E[\widetilde{H}_k\mathcal{E}_{k,k}\widetilde{B}_k^T] + \overline{B}_kR_{k,k}\overline{B}_k^T \\ &+ \overline{H}_k\mathcal{X}_{k,k} + \mathcal{X}_{k,k}^T\overline{H}_k^T - \overline{H}_kP_{k/k-1}\overline{H}_k^T - \Psi_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T, \quad k \geq 1.\end{aligned} \tag{20}$$

It can be observed that the matrices $E[\widetilde{H}_k\mathcal{D}_k\widetilde{H}_k^T]$, $E[\widetilde{B}_kR_{k,k}\widetilde{B}_k^T]$ and $E[\widetilde{B}_k\mathcal{E}_{k,k}^T\widetilde{H}_k^T]$ are yielded by the correlation of the random matrices of the observation equation at the same time instant. From *(b)*, the $(p,q)-$th entries of these matrices are obtained by

$$\begin{aligned}\left(E[\widetilde{H}_k\mathcal{D}_k\widetilde{H}_k^T]\right)_{pq} &= \sum_{j=1}^{n}\sum_{i=1}^{n} C_{h_{qj}^k h_{pi}^k}\left(\mathcal{D}_k\right)_{ji} \quad (p,q=1,2,\ldots,r),\\ \left(E[\widetilde{B}_kR_{k,k}\widetilde{B}_k^T]\right)_{pq} &= \sum_{j=1}^{m}\sum_{i=1}^{m} C_{b_{qj}^k b_{pi}^k}\left(R_{k,k}\right)_{ji} \quad (p,q=1,2,\ldots,r),\\ \left(E[\widetilde{B}_k\mathcal{E}_{k,k}^T\widetilde{H}_k^T]\right)_{pq} &= \sum_{j=1}^{m}\sum_{i=1}^{n} C_{b_{qj}^k h_{pi}^k}\left(\mathcal{E}_{k,k}^T\right)_{ji} \quad (p,q=1,2,\ldots,r).\end{aligned}$$

3.6 *Filtering algorithm: computational procedure and advantages*

The optimal LS linear filtering algorithm is constituted by equations (9)-(13) and (17)-(20), and the computational procedure can be summarized as follows:

i) The matrices $\mathcal{W}_k$, $\mathcal{E}_{k,k}$, $\mathcal{E}_{k,k-1}$ and $\mathcal{V}_{k,k-1}$ are computed by expressions (3)-(5). We then recursively compute $\mathcal{D}_k$ by (6) and thus $\mathcal{G}_{k,k-1}$ by (7); with the matrices $\mathcal{E}_{k,k-1}$, $\mathcal{V}_{k,k-1}$ and $\mathcal{G}_{k,k-1}$, we can compute $\Psi_{k,k-1}$ by (16). The matrices $E[\widetilde{B}_k\mathcal{E}_{k,k}^T\widetilde{H}_k^T]$, $E[\widetilde{H}_k\mathcal{D}_k\widetilde{H}_k^T]$ and $E[\widetilde{B}_kR_{k,k}\widetilde{B}_k^T]$ are also computed in order to obtain the innovation covariance matrix Π_k. Note that all these matrices depend only on the system model information and can be obtained before the observations are available.

ii) At the sampling time k, when the $(k-1)$th iteration is finished and the new observation y_k is available, the proposed filtering algorithm operates as follows (Figure 1):

1) By (18), we compute $\mathcal{X}_{k,k-1} = \overline{F}_{k-1}\mathcal{X}_{k-1,k-1} + \mathcal{W}_{k-1}$ and, from this, $\mathcal{X}_{k,k}$ by (19):

$$\mathcal{X}_{k,k} = P_{k/k-1}\overline{H}_k^T + \mathcal{E}_{k,k}\overline{B}_k^T - \mathcal{X}_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T.$$

2) The innovation μ_k and its covariance matrix Π_k are computed by (17) and (20), respectively:

$$\begin{aligned}\mu_k =& y_k - \overline{H}_k\widehat{x}_{k/k-1} - \Psi_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1},\\ \Pi_k =& E[\widetilde{H}_k\mathcal{D}_k\widetilde{H}_k^T] + E[\widetilde{B}_kR_{k,k}\widetilde{B}_k^T] + E[\widetilde{B}_k\mathcal{E}_{k,k}^T\widetilde{H}_k^T] + E[\widetilde{H}_k\mathcal{E}_{k,k}\widetilde{B}_k^T] + \overline{B}_kR_{k,k}\overline{B}_k^T\\ &+ \overline{H}_k\mathcal{X}_{k,k} + \mathcal{X}_{k,k}^T\overline{H}_k^T - \overline{H}_kP_{k/k-1}\overline{H}_k^T - \Psi_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T.\end{aligned}$$

3) The filter $\widehat{x}_{k/k}$ and the filtering error covariance matrix $P_{k/k}$ are computed by (9) and (13), respectively:

$$\begin{aligned}\widehat{x}_{k/k} &= \widehat{x}_{k/k-1} + \mathcal{X}_{k,k}\Pi_k^{-1}\mu_k,\\ P_{k/k} &= P_{k/k-1} - \mathcal{X}_{k,k}\Pi_k^{-1}\mathcal{X}_{k,k}^T.\end{aligned}$$

4) To implement the above steps at time $k+1$, we must:
- Compute the state predictor by (10): $\widehat{x}_{k+1/k} = \overline{F}_k \widehat{x}_{k/k} + \mathcal{W}_k \Pi_k^{-1} \mu_k$.
- Compute $\mathcal{J}_k = Q_{k-1,k} - \mathcal{X}_{k,k} \Pi_k^{-1} \mathcal{W}_k^T$ by (12), and from this, the prediction error covariance matrix $P_{k+1/k}$ by (11):

$$P_{k+1/k} = \overline{F}_k P_{k/k} \overline{F}_k^T + E[\widetilde{F}_k \mathcal{D}_k \widetilde{F}_k^T] + Q_{k,k} + \overline{F}_k \mathcal{J}_k + \mathcal{J}_k^T \overline{F}_k^T - \mathcal{W}_k \Pi_k^{-1} \mathcal{W}_k^T.$$

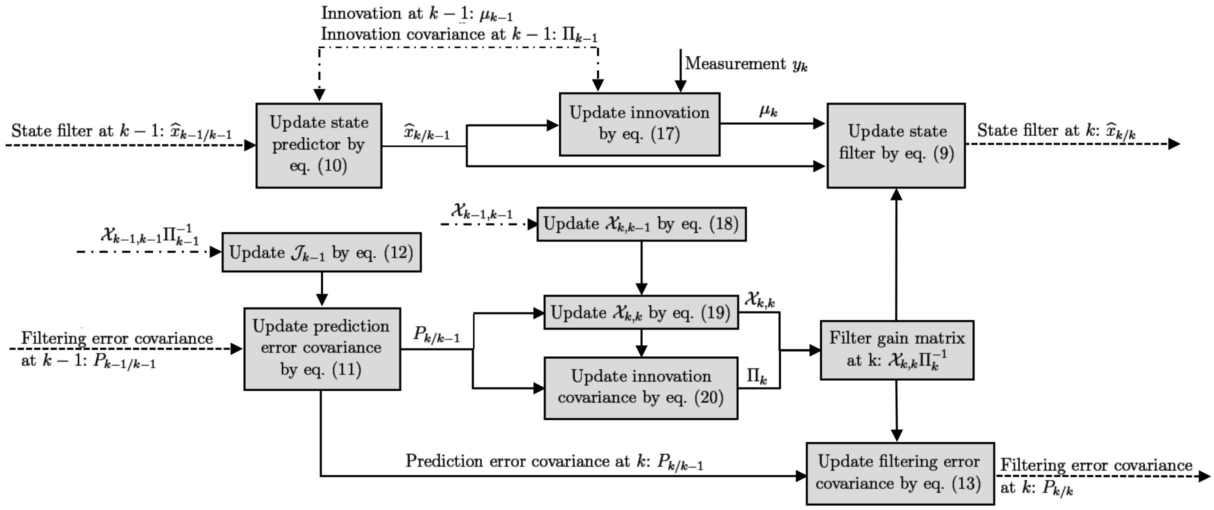


Figure 1. Optimal LS linear filtering algorithm.

The proposed algorithm has the following advantages: 1) the filter is globally optimal in the linear LS sense; 2) the filter structure is recursive, very simple computationally and suitable for online applications; 3) the algorithm takes into account both the influence of the correlation of the random parameter matrices and that of the noises; 4) the algorithm, obtained by an innovation approach, does not require any transformation of the original system into one with deterministic parameter matrices; 5) the proposed filter can be applied to multi-sensor systems with correlated missing measurements considering at each sensor the possibility of observations containing only partial information about the state (Section 4); 6) the proposed filter can be applied to multi-sensor systems with randomly delayed measurements correlated at consecutive sampling times (Section 5).

4. Application to multi-sensor systems with missing measurements

Over the past few decades, considerable research has been carried out into multi-sensor systems with missing measurements, due to the importance of this question and its applicability to modelling a broad class of real-world problems. Most papers concerning systems with missing measurements transmitted by multiple sensors assume that the missing probabilities in all the sensors are identical (see Wang et al. (2006); Nakamori et al. (2005)). In recent years, however, this situation has been generalized to address missing measurements whose statistical properties are not assumed to be the same for all the sensors (see Hounkpevi and Yaz (2007a); Qu and Zhou (2013))). Different missing probabilities have also been considered for some classes of nonlinear systems in Wang et al. (2012) and Hu et al. (2013), where quantized H_∞ control and filtering problems are addressed, respectively. This is a realistic assumption in several application fields, for instance, in networked communication systems involving heterogeneous measurement devices. In Hounkpevi and Yaz (2007a), different sequences of independent Bernoulli random variables are used to describe the missing measurement phenomenon at each sensor. Caballero-Águila et al. (2011) subsequently generalized these results, by weakening the independence restriction and considering sequences of Bernoulli random variables correlated at consecutive sampling times. This form of correlation covers practical situations where the state cannot be missing in two successive observations (for example, transmission models with stand-by sensors in which any failure in the transmission is detected immediately and the old sensor is then replaced). In all of the above papers, it is assumed that the state measurement is either completely lost or successfully transferred, and Bernoulli random variables are used to model the missing measurement phenomenon. In a more recent study, this missing measurement model was generalized to consider an arbitrary discrete distribution in the interval $[0,1]$, thus covering some practical applications where only partial information is missing (see Hu et al. (2012); Caballero-Águila et al. (2013a) and references therein).

Our aim in this section is to show that the observation model in multi-sensor systems with missing measurements can be considered a special case of the observation model with random measurement matrices (2). Hence, the proposed optimal LS linear filtering algorithm can be applied to multiple missing measurement systems with correlated and cross-correlated noises, when the missing measurement phenomenon at each sensor is described by different sequences of correlated (at consecutive sampling times) scalar random variables with arbitrary discrete probability distribution over the interval [0,1]. In particular, the proposed optimal LS linear filtering algorithm extends the results in Luo et al. (2008), Caballero-Águila et al. (2011) and Caballero-Águila et al. (2013a), among others.

Accordingly, consider the state equation given by (1), with $\{F_k\}_{k\geq 0}$ and $\{w_k\}_{k\geq 0}$ verifying the hypotheses (b) and (c), and r sensors which, at any time k, provide scalar measurements of the state, perturbed by additive and multiplicative noises according to the following observation model:

$$y_k^i = \theta_k^i C_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2, \ldots, r, \tag{21}$$

where $\{y_k^i\}_{k\geq 1}$ are the measured data from the $i-$th sensor, $\{C_k^i\}_{k\geq 1}$, are known time-varying matrices with compatible dimensions, $\{v_k^i\}_{k\geq 1}$ are zero-mean measurement noises, and $\{\theta_k^i\}_{k\geq 1}$ are different sequences of scalar discrete-time random variables over the interval $[0,1]$, with $E[\theta_k^i] = \overline{\theta}_k^i$. For $i, j = 1, \ldots, r$, the following noise correlation assumptions are made:

$$\begin{aligned}
Cov[\theta_k^i, \theta_s^j] &= K_{k,k}^{\theta^{ij}} \delta_{k,s} + K_{k,s}^{\theta^{ij}} \delta_{k,s-1} + K_{k,s}^{\theta^{ij}} \delta_{k,s+1},\\
Cov[v_k^i, v_s^j] &= R_{k,k}^{ij} \delta_{k,s} + R_{k,s}^{ij} \delta_{k,s-1} + R_{k,s}^{ij} \delta_{k,s+1},\\
Cov[w_k, v_s^i] &= S_{k,k}^{i} \delta_{k,s} + S_{k,s}^{i} \delta_{k,s-1} + S_{k,s}^{i} \delta_{k,s-2}.
\end{aligned}$$

The observation model (21) can be rewritten as follows:

$$y_k = H_k x_k + v_k, \quad k \geq 1,$$

where $y_k = (y_k^1, \ldots, y_k^r)^T$ is the $r-$dimensional observation vector, $H_k = \Theta_k C_k$, with $\Theta_k = Diag(\theta_k^1, \ldots, \theta_k^r)$ and $C_k = \left[C_k^{1T} \mid \cdots \mid C_k^{rT}\right]^T$, are $r \times n$ random parameter matrices, and $v_k = (v_k^1, \ldots, v_k^r)^T$ is the $r-$dimensional noise vector. Hence, the observation model (21) is a particular case of (2) with $B_k = I$, and clearly verifies that:

- the additive noise $\{v_k\}_{k\geq 1}$ is autocorrelated and cross-correlated with $\{w_k\}_{k\geq 1}$, with

$$Cov[v_k, v_s] = R_{k,k}\delta_{k,s} + R_{k,s}\delta_{k,s-1} + R_{k,s}\delta_{k,s+1},$$
$$Cov[w_k, v_s] = S_{k,k}\delta_{k,s} + S_{k,s}\delta_{k,s-1} + S_{k,s}\delta_{k,s-2},$$

  where $R_{k,s} = \left(R^{ij}_{k,s}\right)_{i,j=1,\cdots,r}$ and $S_{k,s} = \left[S^1_{k,s} \mid \cdots \mid S^r_{k,s}\right]$.
- the random parameter matrices $\{H_k\}_{k\geq 1}$ satisfy:
  - $E[H_k] = \overline{H}_k = \overline{\Theta}_k C_k$ where $\overline{\Theta}_k = E[\Theta_k] = Diag\left(\overline{\theta}^1_k, \ldots, \overline{\theta}^r_k\right)$.
  - Denoting $\theta_k = (\theta^1_k, \ldots, \theta^r_k)^T$, we have $Cov[\theta_k, \theta_s] = K^\theta_{k,k}\delta_{k,s} + K^\theta_{k,s}\delta_{k,s-1} + K^\theta_{k,s}\delta_{k,s+1}$, where $K^\theta_{k,s} = \left(K^{\theta^{ij}}_{k,s}\right)_{i,j=1,\cdots,r}$.
  - For any matrix $\mathcal{A} \in \mathbb{R}^{n\times n}$, we have

$$E[\widetilde{H}_k \mathcal{A} \widetilde{H}^T_s] = E[(\Theta_k - \overline{\Theta}_k)C_k\mathcal{A}C^T_s(\Theta_s - \overline{\Theta}_s)] = K^\theta_{k,s} \circ \left(C_k\mathcal{A}C^T_s\right), \quad s = k,\ k-1.$$

- since $B_k = I$, we have $\widetilde{B}_k = 0$, $\forall k$, and consequently all the expectations in which $\widetilde{B}_k$ or $\widetilde{B}_{k-1}$ appears, are zero.

Thus, the proposed optimal filtering algorithm in this case of multi-sensor systems with missing measurements is the following:

$$\begin{aligned}
&\widehat{x}_{k/k} = \widehat{x}_{k/k-1} + \mathcal{X}_{k,k}\Pi^{-1}_k\mu_k, \quad k \geq 1; \quad \widehat{x}_{0/0} = \overline{x}_0,\\
&\widehat{x}_{k/k-1} = \overline{F}_{k-1}\widehat{x}_{k-1/k-1} + \mathcal{W}_{k-1}\Pi^{-1}_{k-1}\mu_{k-1}, \quad k \geq 2; \qquad \widehat{x}_{1/0} = \overline{F}_0\widehat{x}_{0/0},\\
&\mathcal{X}_{k,k} = P_{k/k-1}C^T_k\overline{\Theta}_k + \mathcal{E}_{k,k} - \left(\overline{F}_{k-1}\mathcal{X}_{k-1,k-1} + \mathcal{W}_{k-1}\right)\Pi^{-1}_{k-1}\Psi^T_{k,k-1}, \quad k \geq 1,\\
&\mu_k = y_k - \overline{\Theta}_k C_k \widehat{x}_{k/k-1} - \Psi_{k,k-1}\Pi^{-1}_{k-1}\mu_{k-1}, \quad k \geq 1,\\
&\Pi_k = K^\theta_{k,k} \circ \left(C_k\mathcal{D}_kC^T_k\right) + R_{k,k} + \overline{\Theta}_kC_k\mathcal{X}_{k,k} + \mathcal{X}^T_{k,k}C^T_k\overline{\Theta}_k\\
&\qquad - \overline{\Theta}_kC_kP_{k/k-1}C^T_k\overline{\Theta}_k - \Psi_{k,k-1}\Pi^{-1}_{k-1}\Psi^T_{k,k-1}, \quad k \geq 1,
\end{aligned}$$

where $\mathcal{E}_{k,k}$, $\mathcal{D}_k$ and $P_{k/k-1}$ are given by (4), (6) and (11), respectively, and

$$\begin{aligned}
&\mathcal{W}_k = Q_{k,k-1}C^T_k\overline{\Theta}_k + S_{k,k}, \quad k \geq 1,\\
&\Psi_{k,k-1} = K^\theta_{k,k-1} \circ \left(C_k\left(\overline{F}_{k-1}\mathcal{D}_{k-1} + Q_{k-1,k-2}\right)C^T_{k-1}\right) + \mathcal{V}_{k,k-1}, \quad k \geq 2; \quad \Psi_{1,0} = 0,\\
&\mathcal{V}_{k,k-1} = S^T_{k-2,k}C^T_{k-1}\overline{\Theta}_{k-1} + R_{k,k-1}, \quad k \geq 2.
\end{aligned}$$

4.1 *Numerical simulation example*

Consider the following system with state-dependent multiplicative noise, and missing measurements from two sensors, with different missing characteristics and noise correlation:

$$\begin{aligned}
&x_k = (0.95 + 0.2\epsilon_{k-1})x_{k-1} + w_{k-1}, \quad k \geq 1,\\
&y^i_k = \theta^i_k x_k + v^i_k, \quad k \geq 1, \quad i = 1, 2.
\end{aligned}$$

The initial state x_0 is a zero-mean Gaussian variable with $P_0 = 1$. The multiplicative state noise $\{\epsilon_k\}_{k\geq 0}$ is a zero-mean Gaussian white process with unit variance. The additive noise processes $\{w_k\}_{k\geq 0}$ and $\{v^i_k\}_{k\geq 1}$, $i = 1, 2$, are the same as those in Caballero-Águila et al. (2013a), i.e.,

$w_k = 0.6(\eta_k + \eta_{k+1})$ and $v_k^i = c_k^i(\eta_{k-1} + \eta_k),\ i = 1, 2$, where $c_k^1 = 1$, $c_k^2 = 0.5$, and $\{\eta_k\}_{k\geq 0}$ is a zero-mean Gaussian white process with variance 0.5.

Two different independent sequences of random variables with a probability distribution over the interval $[0, 1]$ are used to model the missing phenomenon:

- In the first sensor, the missing phenomenon is modelled by a sequence $\{\theta_k^1\}_{k\geq 1}$ of Bernoulli random variables correlated at consecutive sampling times; specifically, $\theta_k^1 = 1 - \beta_{k-1}(1 - \beta_k)$, where $\{\beta_k\}_{k\geq 0}$ is a sequence of independent Bernoulli random variables with $P[\beta_k = 1] = \beta$. Since the variables β_k and β_s are independent, it is clear that θ_k^1 and θ_s^1 are also independent for $|k - s| \geq 2$. Moreover, if $\theta_k^1 = 0$, then $\beta_{k-1} = 1$ and $\beta_k = 0$, and consequently $\theta_{k+1}^1 = 1$; hence, in the first sensor the state cannot be missing in two successive observations.
- In the second sensor, the missing phenomenon is modelled by a sequence $\{\theta_k^2\}_{k\geq 1}$ of independent and identically distributed random variables with the following probability distribution: $P[\theta_k^2 = 0] = 0.1,\quad P[\theta_k^2 = 0.5] = 0.5,\quad P[\theta_k^2 = 1] = 0.4$.

Under these assumptions, for all k, the mean $\overline{\Theta}_k$ and the covariances $K_{k,s}^{\theta},\ s = k,\ k-1$, are given by

$$\overline{\Theta}_k = \begin{bmatrix} \overline{\theta}^1 & 0 \\ 0 & \overline{\theta}^2 \end{bmatrix} = \begin{bmatrix} 1-\beta(1-\beta) & 0 \\ 0 & 0.65 \end{bmatrix},$$

$$K_{k,k}^{\theta} = \begin{bmatrix} \overline{\theta}^1(1-\overline{\theta}^1) & 0 \\ 0 & 0.1025 \end{bmatrix} \quad \text{and} \quad K_{k,k-1}^{\theta} = \begin{bmatrix} -(1-\overline{\theta}^1)^2 & 0 \\ 0 & 0 \end{bmatrix}.$$

To analyze the effectiveness of the proposed estimator, one hundred iterations of the proposed filtering algorithm were performed and the filtering error variances were calculated for different values of the probability β, which provide different values of the probability $\overline{\theta}^1$ that the state is not missing from the observations of the first sensor. Since $\overline{\theta}^1$ is the same if the value $1 - \beta$ is considered instead of β, it is sufficient to consider $\beta \leq 0.5$ (note that, in this case, $\overline{\theta}^1$ is a decreasing function of β). Specifically, the values $\beta = 0.1, 0.2, 0.3, 0.4$ and 0.5 (leading to $\overline{\theta}^1 = 0.91, 0.84, 0.78, 0.76$ and 0.75, respectively) are examined here.

Figure 2 shows that the filtering error variances become greater as β increases or, equivalently, as $\overline{\theta}^1$ decreases. This means that, as the probability of only noise measurements (false alarm probability) increases in the first sensor, worse estimations are obtained; note that for $\beta = 0.3, 0.4, 0.5$ the difference is smaller since the corresponding values of $\overline{\theta}^1$ are very close to each other.

Finally, we present a comparative analysis of four filters: the Kalman filter in systems with independent random parameter matrices and uncorrelated white noises (Luo et al. 2008); the linear filter in systems with uncertain observations with correlated uncertainty and uncorrelated white noises (Caballero-Águila et al. 2011); the centralized filter in systems with missing measurements and correlated and cross-correlated noises (Caballero-Águila et al. 2013a); and the filter proposed here. Using one thousand independent simulations of the mentioned algorithms, the different filtering estimates were compared using the mean square error (MSE) criterion. The filtering MSE at time k is calculated by $\text{MSE}_k = \dfrac{1}{1000}\sum_{s=1}^{1000}(x_k^{(s)} - \widehat{x}_{k/k}^{(s)})^2$, where $\{x_k^{(s)}\}_{1\leq k\leq 100}$ denote the s-th set of artificially simulated data and $\widehat{x}_{k/k}^{(s)}$ is the filter at the sampling time k in the s-th simulation run. These values are shown in Figure 3, where it can be seen that: 1) the performance of the filters with correlated uncertainty or with correlated and cross-correlated noises is better than that of the Kalman filter with independent random parameter matrices and uncorrelated white noises, since this filter ignores any correlation assumption; 2) the proposed filtering algorithm provides better estimations than other filtering algorithms reported previously.

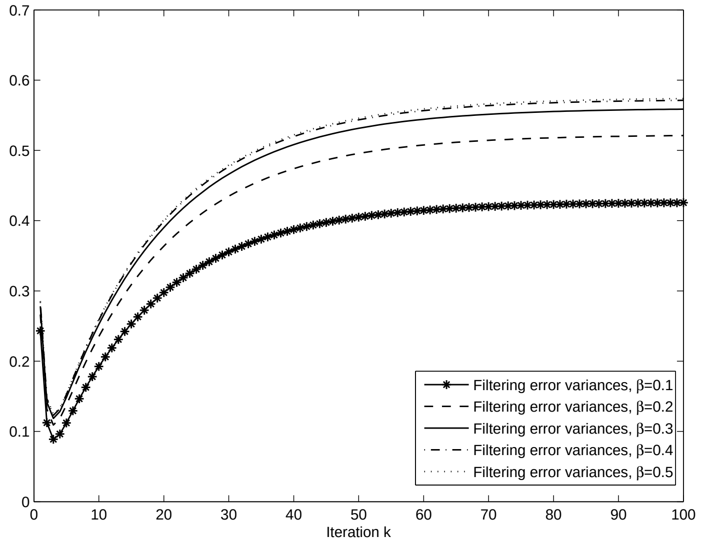


Figure 2. Filtering error variances for $\beta = 0.1$, 0.2, 0.3, 0.4 and 0.5.

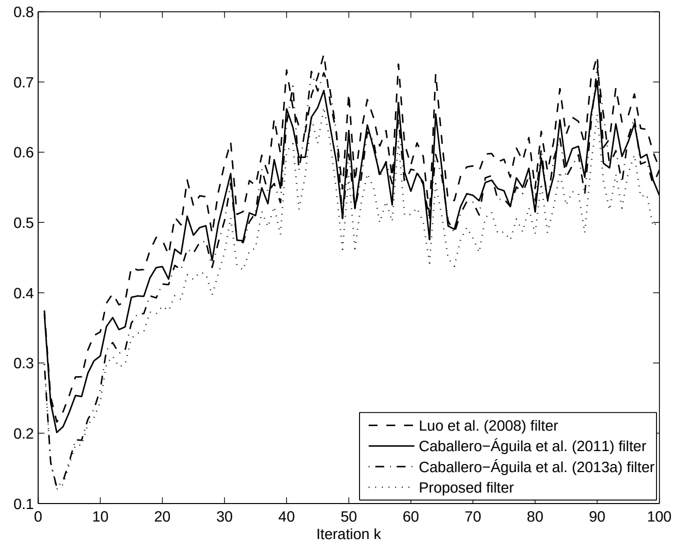


Figure 3. Comparison of MSE_k for different filters.

5. Application to multi-sensor systems with randomly delayed measurements

The estimation problem in multi-sensor systems with randomly delayed measurements is arousing increasing interest due to its broad scope of application. In networked systems, time delays are usually unavoidable, due to numerous causes including network congestion, random failures in the transmission mechanism or data inaccessibility at certain times. Since delays in measurement arrivals often occur randomly, the standard estimation algorithms are not applicable and several modifications have been proposed to incorporate the effects of randomly delayed measurements (see Matveev and Savkin (2003); Sinopoli et al. (2004); Wang et al. (2004)).

Most papers on estimation in multi-sensor systems with randomly delayed observations assume that all the sensors have the same delay characteristics. Nevertheless, such an assumption is not realistic in many practical situations, where the information is gathered by an array of heterogeneous sensors, and the delay probability at each individual sensor can be different from the others. In recent years, this approach has been generalized in Hounkpevi and Yaz (2007b) and Caballero-Águila et al. (2010) considering multiple delayed sensors with different delay characteristics and assuming that the delays are mutually independent. Caballero-Águila et al. (2013b), recently weakened this assumption of independence by considering different sequences of Bernoulli variables correlated at consecutive sampling times to model the delays at each sensor. Similarly to the case of missing measurements, this correlation model avoids the possibility of two successive delayed observations, and so it can be applied to networked systems with stand-by sensors for the immediate replacement of a failed unit.

In this section we show that the current observation model with random measurement matrices (2) includes the observation model in multi-sensor systems with correlated random delays as a particular case; thus the current study generalizes the above results (Hounkpevi and Yaz (2007b); Caballero-Águila et al. (2010, 2013b)).

Assume that the state equation is given by (1), with $\{F_k\}_{k\geq 0}$ and $\{w_k\}_{k\geq 0}$ verifying hypotheses (b) and (c), and consider that z_k^i, $i = 1, \ldots, r$, are scalar sensor outputs perturbed by zero-mean additive noises v_k^i; namely,

$$z_k^i = c_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, \ldots, r. \tag{22}$$

For $i, j = 1, \ldots, r$, it is assumed that $Cov[v_k^i, v_s^j] = R_k^{ij}\delta_{k,s}$ and $Cov[w_k, v_s^i] = S_{k,k}^i\delta_{k,s} + S_{k,s}^i\delta_{k,s-1}$.

Consider that, at the initial time $k = 1$, the ith sensor outputs, z_1^i, are always available for the estimation but, at time $k \geq 2$, the ith sensor measurement, y_k^i, may be randomly delayed by one sampling time according to different delay characteristics, due to possible failures in data transmission. Therefore, the measurement model is described by

$$y_k^i = (1 - \gamma_k^i)z_k^i + \gamma_k^i z_{k-1}^i, \quad k \geq 2; \quad y_1^i = z_1^i, \quad i = 1, \ldots, r, \tag{23}$$

where $\left\{\gamma_k^i\right\}_{k\geq 2}$, $i = 1, \ldots, r$, denote sequences of Bernoulli variables with $P\left[\gamma_k^i = 1\right] = p_k^i$ and $Cov[\gamma_k^i, \gamma_s^j] = K_{k,k}^{\gamma^{ij}}\delta_{k,s} + K_{k,s}^{\gamma^{ij}}\delta_{k,s-1} + K_{k,s}^{\gamma^{ij}}\delta_{k,s+1}$.

By denoting $z_k = \left(z_k^1, \ldots, z_k^r\right)^T$, $C_k = \left[\begin{array}{c|c|c} c_k^{1T} & \cdots & c_k^{rT}\end{array}\right]^T$, $v_k = \left(v_k^1, \ldots, v_k^r\right)^T$, and $\Gamma_k = Diag\left(\gamma_k^1, \ldots, \gamma_k^r\right)$, (22) and (23) can be rewritten as:

$$\begin{aligned} z_k &= C_k x_k + v_k, \quad k \geq 1, \\ y_k &= (I - \Gamma_k)z_k + \Gamma_k z_{k-1}, \quad k \geq 2; \quad y_1 = z_1, \end{aligned} \tag{24}$$

and, from the correlation assumptions of the noises, it is clear that $Cov[v_k, v_s] = R_k\delta_{k,s}$ and $Cov[w_k, v_s] = S_{k,k}\delta_{k,s} + S_{k,s}\delta_{k,s-1}$, where $R_k = \left(R_k^{ij}\right)_{i,j=1,\cdots,r}$ and $S_{k,s} = \left[S_{k,s}^1 \mid \cdots \mid S_{k,s}^r\right]$. Moreover, by denoting $\gamma_k = \left(\gamma_k^1, \ldots, \gamma_k^r\right)^T$, we have $Cov[\gamma_k, \gamma_s] = K_{k,k}^{\gamma}\delta_{k,s} + K_{k,s}^{\gamma}\delta_{k,s-1} + K_{k,s}^{\gamma}\delta_{k,s+1}$, where $K_{k,s}^{\gamma} = \left(K_{k,s}^{\gamma^{ij}}\right)_{i,j=1,\cdots,r}$.

Now, as in Hounkpevi and Yaz (2007b), equations (1) and (24) are rewritten as follows, with random parameter matrices:

$$\begin{array}{ll} X_{k+1} = \mathcal{F}_k X_k + W_k, & k \geq 1, \\ y_k = H_k X_k + B_k V_k, & k \geq 2, \end{array} \tag{25}$$

where

$$X_k = \left[\begin{array}{c} x_k \\ \hline x_{k-1} \end{array}\right], \quad W_k = \left[\begin{array}{c} w_k \\ \hline 0 \end{array}\right], \quad V_k = \left[\begin{array}{c} v_k \\ \hline v_{k-1} \end{array}\right], \quad \mathcal{F}_k = \left[\begin{array}{c|c} F_k & 0 \\ \hline I & 0 \end{array}\right],$$

$$H_k = \left[\begin{array}{c|c} (I-\Gamma_k)C_k & \Gamma_k C_{k-1} \end{array}\right], \quad B_k = \left[\begin{array}{c|c} I-\Gamma_k & \Gamma_k \end{array}\right].$$

It is clear that the random parameter matrices and noise processes of system (25) verify the hypotheses to apply the algorithm proposed in this paper. Specifically, we have:

- $E[\mathcal{F}_k] = \overline{\mathcal{F}}_k = \left[\begin{array}{c|c} \overline{F}_k & 0 \\ \hline I & 0 \end{array}\right]$, $\overline{H}_k = \left[\begin{array}{c|c} (I-\overline{\Gamma}_k)C_k & \overline{\Gamma}_k C_{k-1} \end{array}\right]$ and $\overline{B}_k = \left[\begin{array}{c|c} I-\overline{\Gamma}_k & \overline{\Gamma}_k \end{array}\right]$, where $\overline{\Gamma}_k = Diag\left(p_k^1, \ldots, p_k^r\right)$.
- The process noise, $\{W_k\}_{k\geq 1}$, and the measurement noise, $\{V_k\}_{k\geq 2}$, are zero-mean sequences with covariances and cross-covariances:

$$\begin{array}{l} Cov[W_k, W_s] = \mathbb{Q}_{k,k}\delta_{k,s} + \mathbb{Q}_{k,s}\delta_{k,s-1} + \mathbb{Q}_{k,s}\delta_{k,s+1}, \\ Cov[V_k, V_s] = \mathbb{R}_{k,k}\delta_{k,s} + \mathbb{R}_{k,s}\delta_{k,s-1} + \mathbb{R}_{k,s}\delta_{k,s+1}, \\ Cov[W_k, V_s] = \mathbb{S}_{k,k}\delta_{k,s} + \mathbb{S}_{k,s}\delta_{k,s-1} + \mathbb{S}_{k,s}\delta_{k,s-2}, \end{array}$$

where

$$\mathbb{Q}_{k,k} = \left[\begin{array}{c|c} Q_{k,k} & 0 \\ \hline 0 & 0 \end{array}\right], \quad \mathbb{Q}_{k,k-1} = \left[\begin{array}{c|c} Q_{k,k-1} & 0 \\ \hline 0 & 0 \end{array}\right], \quad \mathbb{R}_{k,k} = \left[\begin{array}{c|c} R_k & 0 \\ \hline 0 & R_{k-1} \end{array}\right], \quad \mathbb{R}_{k,k-1} = \left[\begin{array}{c|c} 0 & 0 \\ \hline R_{k-1} & 0 \end{array}\right],$$

$$\mathbb{S}_{k,k} = \left[\begin{array}{c|c} S_{k,k} & 0 \\ \hline 0 & 0 \end{array}\right], \quad \mathbb{S}_{k-1,k} = \left[\begin{array}{c|c} S_{k-1,k} & S_{k-1,k-1} \\ \hline 0 & 0 \end{array}\right], \quad \mathbb{S}_{k-2,k} = \left[\begin{array}{c|c} 0 & S_{k-2,k-1} \\ \hline 0 & 0 \end{array}\right].$$

Then, it is clear that

- $\mathbb{D}_k = E[X_k X_k^T] = \left[\begin{array}{c|c} \mathcal{D}_k & \mathcal{G}_{k,k-1} \\ \hline \mathcal{G}_{k,k-1}^T & \mathcal{D}_{k-1} \end{array}\right]$, $E\left[\widetilde{\mathcal{F}}_k \mathbb{D}_k \widetilde{\mathcal{F}}_k^T\right] = \left[\begin{array}{c|c} E[\widetilde{F}_k \mathcal{D}_k \widetilde{F}_k^T] & 0 \\ \hline 0 & 0 \end{array}\right]$ and $\mathbb{G}_{k+1,k} = E[X_{k+1}X_k^T] = \left[\begin{array}{c|c} \mathcal{G}_{k+1,k}^T & \overline{F}_k \mathcal{G}_{k,k-1} \\ \hline \mathcal{D}_k & \mathcal{G}_{k,k-1} \end{array}\right]$, where $\mathcal{D}_k$ and $\mathcal{G}_{k,k-1}$ are given by (6) and (7), respectively.
- Analogously to (3) and (4), we have

$$\begin{array}{l} \mathbb{W}_k = E[W_k y_k^T] = \mathbb{Q}_{k,k-1}\overline{H}_k^T + \mathbb{S}_{k,k}\overline{B}_k^T, \quad k \geq 1, \\ \mathbb{E}_{k,k} = E[X_k v_k^T] = \overline{\mathcal{F}}_{k-1}\mathbb{S}_{k-2,k} + \mathbb{S}_{k-1,k}, \quad k \geq 2; \qquad \mathbb{E}_{1,1} = \mathbb{S}_{0,1}, \\ \mathbb{E}_{k,k-1} = E[X_k v_{k-1}^T] = \overline{\mathcal{F}}_{k-1}\mathbb{E}_{k-1,k-1} + \mathbb{S}_{k-1,k-1}, \quad k \geq 2. \end{array}$$

– For arbitrary matrices $\mathcal{A}_1 \in \mathbb{R}^{2n\times 2n}$, $\mathcal{A}_2 \in \mathbb{R}^{2r\times 2r}$ and $\mathcal{A}_3 \in \mathbb{R}^{2r\times 2n}$, we have

$$E[\widetilde{H}_k \mathcal{A}_1 \widetilde{H}_s^T] = K_{k,s}^{\gamma} \circ \left([-C_k \mid C_{k-1}]\, \mathcal{A}_1 \, [-C_s \mid C_{s-1}]^T \right), \quad s = k,\ k-1,$$
$$E[\widetilde{B}_k \mathcal{A}_2 \widetilde{B}_s^T] = K_{k,s}^{\gamma} \circ \left([-I \mid I\,]\, \mathcal{A}_2 \, [-I \mid I\,]^T \right), \quad s = k,\ k-1,$$
$$E[\widetilde{B}_k \mathcal{A}_3 \widetilde{H}_s^T] = K_{k,s}^{\gamma} \circ \left([-I \mid I\,]\, \mathcal{A}_3 \, [-C_s \mid C_{s-1}]^T \right), \quad s = k,\ k-1.$$

By applying the above expressions, we obtain $E[\widetilde{H}_k \mathbb{D}_k \widetilde{H}_k^T]$, $E[\widetilde{B}_k \mathbb{R}_{k,k} \widetilde{B}_k^T]$ and $E[\widetilde{B}_k \mathbb{E}_{k,k}^T \widetilde{H}_k^T]$, which are necessary to calculate the innovation covariance matrices. We also have

$$\mathcal{V}_{k,k-1} = \overline{\Gamma}_k \left(C_{k-1} S_{k-2,k-1} + R_{k-1} \right)^T \left(I - \overline{\Gamma}_{k-1} \right) - K_{k,k-1}^{\gamma} \circ \left(C_{k-1} S_{k-2,k-1} + R_{k-1} \right)^T,$$
$$\Psi_{k,k-1} = K_{k,k-1}^{\gamma} \circ \left([-C_k \mid C_{k-1}] \left(\mathbb{G}_{k,k-1} [-C_{k-1} \mid C_{k-2}]^T + \mathbb{E}_{k,k-1} [-I \mid I\,]^T \right) \right) + \mathcal{V}_{k,k-1}.$$

Hence, the proposed optimal filtering algorithm for multi-sensor systems with randomly delayed measurements is:

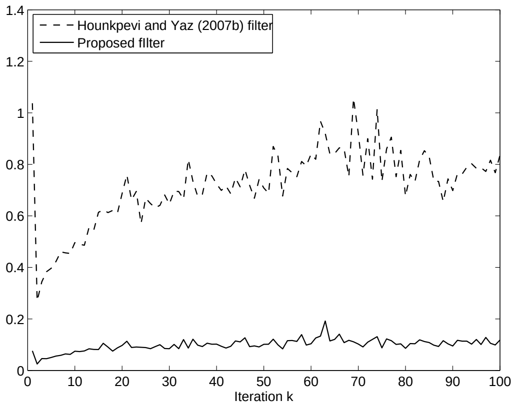
$$\widehat{X}_{k/k} = \widehat{X}_{k/k-1} + \mathcal{X}_{k,k} \Pi_k^{-1} \mu_k, \quad k \geq 1,$$
$$\widehat{X}_{k/k-1} = \overline{\mathcal{F}}_{k-1} \widehat{X}_{k-1/k-1} + \mathbb{W}_{k-1} \Pi_{k-1}^{-1} \mu_{k-1}, \quad k \geq 2; \qquad \widehat{X}_{1/0} = \left[\frac{\overline{F}_0 \overline{x}_0}{\overline{x}_0} \right],$$
$$\mathcal{X}_{k,k} = P_{k/k-1} \overline{H}_k^T + \mathbb{E}_{k,k} \overline{B}_k^T - \left(\overline{\mathcal{F}}_{k-1} \mathcal{X}_{k-1,k-1} + \mathbb{W}_{k-1} \right) \Pi_{k-1}^{-1} \Psi_{k,k-1}^T, \quad k \geq 1,$$
$$\mu_k = y_k - \overline{H}_k \widehat{X}_{k/k-1} - \Psi_{k,k-1} \Pi_{k-1}^{-1} \mu_{k-1}, \quad k \geq 1,$$
$$\Pi_k = E[\widetilde{H}_k \mathbb{D}_k \widetilde{H}_k^T] + E[\widetilde{B}_k \mathbb{R}_{k,k} \widetilde{B}_k^T] + E[\widetilde{B}_k \mathbb{E}_{k,k}^T \widetilde{H}_k^T] + E[\widetilde{H}_k \mathbb{E}_{k,k} \widetilde{B}_k^T] + \overline{B}_k \mathbb{R}_{k,k} \overline{B}_k^T$$
$$+ \overline{H}_k \mathcal{X}_{k,k} + \mathcal{X}_{k,k}^T \overline{H}_k^T - \overline{H}_k P_{k/k-1} \overline{H}_k^T - \Psi_{k,k-1} \Pi_{k-1}^{-1} \Psi_{k,k-1}^T,$$
$$P_{k/k} = P_{k/k-1} - \mathcal{X}_{k,k} \Pi_k^{-1} \mathcal{X}_{k,k}^T, \quad k \geq 1,$$
$$P_{k/k-1} = \overline{\mathcal{F}}_{k-1} P_{k-1/k-1} \overline{\mathcal{F}}_{k-1}^T + E[\widetilde{\mathcal{F}}_{k-1} \mathbb{D}_{k-1} \widetilde{\mathcal{F}}_{k-1}^T] + \mathbb{Q}_{k-1,k-1} + \overline{\mathcal{F}}_{k-1} \mathbb{J}_{k-1} + \mathbb{J}_{k-1}^T \overline{\mathcal{F}}_{k-1}^T$$
$$- \mathbb{W}_{k-1} \Pi_{k-1}^{-1} \mathbb{W}_{k-1}^T, \quad k \geq 2;$$
$$P_{1/0} = \left[\begin{array}{c|c} \overline{F}_0 P_0 \overline{F}_0^T + E[\widetilde{F}_0 \mathcal{D}_0 \widetilde{F}_0^T] + Q_{0,0} & \overline{F}_0 P_0 \\ \hline P_0 \overline{F}_0^T & P_0 \end{array} \right],$$
$$\mathbb{J}_k = \mathbb{Q}_{k-1,k} - \mathcal{X}_{k,k} \Pi_k^{-1} \mathbb{W}_k^T, \quad k \geq 1.$$

5.1 *Numerical simulation example*

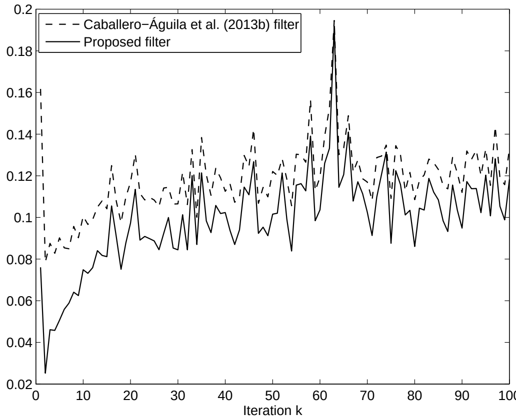
In this example, it is assumed that the state $\{x_k\}_{k\geq 0}$ is generated by the same model as that in Section 4.1, and we consider measured outputs coming from two sensors, $z_k^i = x_k + v_k^i,\ k \geq 1,\ i = 1, 2$, where the additive noises are defined by $v_k^1 = \eta_k$ and $v_k^2 = 0.5\eta_k$.

According to the proposed observation model, it is assumed that, at any sampling time $k \geq 2$, the measured output from the ith sensor, z_k^i, can be randomly delayed by one sampling period during network transmission; thus, the measurement model is described by

$$y_k^i = (1 - \gamma_k^i) z_k^i + \gamma_k^i z_{k-1}^i, \quad k \geq 2; \quad y_1^i = z_1^i, \quad i = 1, 2.$$



(a) Proposed filter vs. Hounkpevi and Yaz (2007b) filter.



(b) Proposed filter vs. Caballero-Águila et al. (2013b) filter.

Figure 4. Comparison of MSE_k for different filters.

As in Caballero-Águila et al. (2013b), it is assumed that the delays are correlated at consecutive sampling times, which guarantees that two successive observations cannot be delayed; specifically, the variables γ_k^i are defined by $\gamma_k^i = \alpha_{k+1}^i(1-\alpha_k^i)$, where $\{\alpha_k^i\}_{k\geq 1}$, $i = 1, 2$, are two independent sequences of independent Bernoulli variables with probabilities $P[\alpha_k^1 = 1] = 0.5$ and $P[\alpha_k^2 = 1] = 0.1$, respectively.

To illustrate the accuracy of the proposed algorithm in comparison with other estimation methods that have been proposed, one thousand independent simulations were considered and one hundred iterations of each algorithm performed to compute the filtering MSE at each time instant k. A comparative analysis was carried out between the suboptimal Kalman-type filter for systems with independent random delays (Hounkpevi and Yaz 2007b), the optimal linear filter using covariance information for systems with one-step correlated random delays (Caballero-Águila et al. 2013b), and the current filter for multi-sensor systems with randomly delayed measurements. The results of this comparison are shown in Figure 4, where it can be seen that the proposed filter performs better than the other two. The difference with respect to Hounkpevi and Yaz (2007b) is greater since the correlation assumption on the delays and the noises is not taken into account and moreover the estimator in Hounkpevi and Yaz (2007b) is suboptimal.

6. Conclusions

This paper reports a study of the optimal LS linear filtering problem for discrete-time linear systems with random parameter matrices and correlated additive noise. The main contributions of this approach are:

(1) The current system model includes independent random state transition matrices and one-step correlated and cross-correlated random parameter matrices in the observation equation. The process and measurement noises are assumed to be one-step autocorrelated and two-step cross-correlated.
(2) An optimal LS linear recursive filtering algorithm with a simple computational procedure is derived by an innovation approach.
(3) The proposed optimal LS linear filtering algorithm was applied to systems with multiple missing measurements with correlated and cross-correlated noises, when the missing measurement phenomenon in each sensor is described by different sequences of scalar random variables with arbitrary discrete probability distribution over the interval [0,1] correlated at consecutive sampling times. This kind of multi-sensor system is found in various real-

world problems, such as transmission models with stand-by sensors or situations involving the partial loss of measurements.

(4) Multi-sensor systems with randomly delayed measurements, correlated at consecutive sampling times, with correlated and cross-correlated noises are also treated as a particular case of the model described in this paper. These models cover the situations in which two successive observations cannot be delayed. This kind of delay frequently occurs, in situations such as network congestion, random failures in the transmission mechanism or data inaccessibility at certain times.
(5) For both particular cases, the feasibility of the proposed filtering algorithm is analyzed by two numerical simulation examples, which show that the proposed filter performs better than others that have been reported.
(6) A similar study to that performed in this paper would allow us to generalize the current results by considering correlation between random state transition matrices and the random matrices in the observation equation. This extension would cover systems with multiple packet dropouts as a particular case, and would constitute an interesting research topic.
(7) Another interesting future direction would be to complement the current study with a detailed analysis of the convergence and computational complexity of the proposed filtering algorithm.
(8) The filtering methodology proposed in this paper can be applied to other, related problems, such as fault detection or control systems, which constitute interesting and challenging topics for future research (Dong et al. (2012a), Dong et al. (2012b), Wang et al. (2012)).

Acknowledgements

This research is supported by Ministerio de Ciencia e Innovación (FPU programme and Grant No. MTM2011-24718).

References

Caballero-Águila, R., Hermoso-Carazo, A., Jiménez-López, J.D., Linares-Pérez, J., and Nakamori, S. (2010), "Signal estimation with multiple delayed sensors using covariance information," *Digital Signal Processing*, 20, 528–540.

Caballero-Águila, R., Hermoso-Carazo, A. and Linares-Pérez, J. (2011), "Linear and quadratic estimation using uncertain observations from multiple sensors with correlated uncertainty", *Signal Processing,* 91 (2), 330-337.

Caballero-Águila, R., García-Garrido, I. and Linares-Pérez, J. (2013a), "Optimal Fusion Filtering in Multisensor Stochastic Systems with Missing Measurements and Correlated Noises", *Mathematical Problems in Engineering,* Volume 2013, Article ID 418678, 14 pages.

Caballero-Águila, R., Hermoso-Carazo, A. and Linares-Pérez, J. (2013b), "Linear estimation based on covariances for networked systems featuring sensor correlated random delays," *International Journal of Systems Science,* 44(7), 1233-1244.

Carravetta, F., Germani, A. and Raimondi, M. (1997), "Polynomial Filtering of Discrete-Time Stochastic Linear Systems with Multiplicative State Noise", *IEEE Transactions on Automatic Control*, 42 (8), 1106-1126.

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

Dong, H., Wang, Z. and Gao, H. (2012a), "Fault detection for Markovian jump systems with sensor saturations and randomly varying nonlinearities", *IEEE Transactions on Circuits and Systems - Part I,* 59(10), 2354-2362.

Dong, H., Wang, Z., Lam, J. and Gao, H. (2012b), "Fuzzy-model-based robust fault detection with stochastic mixed time-delays and successive packet dropouts", *IEEE Transactions on Systems, Man, and Cybernetics - Part B,* 42(2), 365-376.

Feng, J., Wang, Z. and Zeng, M. (2011), “Optimal robust non-fragile Kalman-type recursive filtering with finite-step autocorrelated noises and multiple packet dropouts”, *Aerospace Science and Technology,* 15 (6), 486-494.

Feng, J., Wang, Z. and Zeng, M. (2013),“Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises”, *Information Fusion,* 14, 78-86.

Fu, A., Zhu, Y. and Song, E. (2008), “The optimal Kalman type state estimator with multi-step correlated process and measurement noises”, in *The 2008 International Conference on Embedded Software and Systems,* 215-220.

Hounkpevi, F.O., and Yaz, E.E. (2007a), “Robust minimum variance linear state estimators for multiple sensors with different failure rates”, *Automatica,* 43, 1274-1280.

Hounkpevi, F.O., and Yaz, E.E. (2007b), “ Minimum variance generalized state estimators for multiple sensors with different delay rates,” *Signal Processing* , 87, 602–613.

Hu, J., Wang, Z., Gao H., and Stergiouslas, L.K. (2012), “Extended Kalman filtering with stochastic nonlinearities and multiple missing measurements”, *Automatica,* 48, 2007-2015.

Hu, J., Wang, Z., Shen, B., and Gao H. (2013), “Quantized recursive filtering for a class of nonlinear systems with multiplicative noises and missing measurements”, *International Journal of Control,* 86 (4), 650-663.

Luo, Y., Zhu, Y., Luo, D., Zhou, J., Song, E. and Wang, D. (2008), “Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications”, *Sensors,* 8 (12), 8086-8103.

Luo, Y.T., Zhu, Y.M., Shen X.J. and Song E.B. (2012), “Novel Data Association Algorithm Based on Integrated Random Coefficient Matrices Kalman Filtering,” *IEEE Transactions on Aerospace and Electronic Systems,* 48(1), 144-158.

Ma, J. and Sun, S.L. (2011), “Optimal linear estimators for systems with random sensor delays, multiple packet dropouts and uncertain observations,” *IEEE Trans. Signal Process.,* 59(11), 5181–5192.

Matveev, A.S., and Savkin, A.V. (2003), “The problem of state estimation via asynchronous communication channels with irregular transmission times,” *IEEE Transactions on Automatic Control*, 48(4), 670–676.

Moayedi, M., Foo, Y.K. and Soh, Y.C. (2010), “Adaptive Kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements,” *IEEE Trans. Signal Process.,* 58(3) 1577-1588.

Nakamori, S., Caballero-Águila, R., Hermoso-Carazo, A., and Linares- Pérez, J. (2005) “New recursive estimators from correlated interrupted observations using covariance information”, *International Journal of Systems Science,* 36 (10) 617-629.

Qu, X., and Zhou, J. (2013), “The optimal robust finite-horizon Kalman filtering for multiple sensors with different stochastic failure rates”, *Applied Mathematics Letters,* 26, 80-86.

Sahebsara, M., Chen, T. and Shah, S.L. (2007) “Optimal H_2 filtering with random sensor delay, multiple packet dropout and uncertain observations,” *Int. J. Control*, 80, 292-301.

Shen X.J., Luo, Y.T., Zhu, Y.M. and Song, E.B. (2012), “Globally optimal distributed Kalman filtering fusion,” *Science China. Information Sciences,* 55(3), 512-529.

Shen X.J., Zhu, Y.M. and Luo, Y.T. (2011), “Optimal State Estimation of Linear Discrete-time Systems with Correlated Random Parameter Matrices,” in *Proceedings of the 30th Chinese Control Conference,* 1488-1493.

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

Song, E., Zhu, Y. and You, Z. (2008), “The Kalman type recursive state estimator with a finite-step correlated process noises”, in *Proceedings of the IEEE International Conference on Automation and Logistics,* 196-200.

Sun, S.L. and Xiao, W. (2013) “Optimal linear estimators for systems with multiple random measurement delays and packet dropouts,” *International Journal of Systems Science,* 44(2), 358-370.

Wang, Z. , Ho, D.W.C., and Liu, X. (2004), “Robust filtering under randomly varying sensor delay with variance constraints,” *IEEE Transactions on Circuits and Systems, II Express Briefs*, 51(6), 320–326.

Wang, Z., Yang, F., Ho, D.W.C., and Liu, X. (2006),“Robust H_∞ filtering for stochastic time-delay systems with missing measurements”, *IEEE Transaction on Signal Processing,* 54 (7), 2579-2587.

Wang, Z., Shen, B., Shu, H., and Wei, G. (2012), “Quantized H_∞ control for nonlinear stochastic time-delay systems with missing measurements”, *IEEE Transactions on Automatic Control,* 57 (6), 1431-1444.