Distributed fusion filtering for uncertain systems with coupled noises, random delays and packet loss prediction compensation

R. Caballero-Águila[a] and J. Linares-Pérez[b]

[a]Dpto. de Estadística. Universidad de Jaén. Paraje Las Lagunillas. 23071. Jaén. Spain;
[b]Dpto. de Estadística. Universidad de Granada. Avda. Fuentenueva. 18071. Granada. Spain

ARTICLE HISTORY
Compiled January 30, 2025

ABSTRACT
The design of recursive estimation algorithms in networked systems is an important research challenge from both theoretical and practical perspectives. The growing number of application fields are demanding the development of new mathematical models and algorithms that accommodate the effect of the unavoidable network-induced uncertainties. Special relevance have transmission delays and packet dropouts, which may yield a significant degradation in the performance of conventional estimators. This paper discusses the distributed fusion estimation problem in a class of linear stochastic uncertain systems whose measurement noises are cross-correlated and coupled with the process noise. The uncertainty of the system is not only described by additive noises, but also by multiplicative noise in the state equation and random parameter matrices in the measurement model. Both one-step delays and packet dropouts can randomly occur during the transmission of the sensor measurements to the local processors and a compensation strategy based on measurement prediction is used. Under the least-squares criterion and using an innovation approach, a recursive algorithm for the local filtering estimators is designed. These local estimators are then fused at a processing centre, where the distributed fusion filter is generated as the least-squares matrix-weighted linear combination of the local ones.

KEYWORDS
Networked systems; distributed fusion estimation; noise correlation; random delays and packet dropouts; loss compensation

1. Introduction

The great development of computer and communication technologies over the past decades has yielded a growing amount of significant new applications in many fields concerning networked systems, such as intelligent transportation, telemedicine, industry and home automation or space and terrestrial exploration, among others. It is, therefore, not surprising that several challenges have arisen in relation to the efficient processing of the information provided by sensor networks and, particularly, that the study of the estimation problem in such systems has become a highly relevant research topic. Despite the unquestionable advantages of sensor networks, the presence of external environment uncertainties and internal network vulnerabilities often cause certain phenomena, which are generally random in nature, such as lack of information about

CONTACT R. Caballero-Águila. Email: raguila@ujaen.es

the signal –commonly referred to as uncertain or missing observations–, fading measurements or sensor gain degradation, transmission delays and/or packet dropouts, just to mention a few ones related to this work (Hu et al., 2016). Due to these random uncertainties, which usually lead to errors in the data available for estimation, the quality of the estimators proposed in conventional systems can be considerably affected, thus leading to new challenges in the study of the estimation problem in networked systems influenced by random imperfections.

The use of random matrices in the mathematical model of the sensor measurements provides a unified framework to describe different random phenomena that arise in numerous application fields, such as digital control of chemical processes, radar control, navigation systems and economic systems (see e.g. Hu et al., 2013a,b; Linares-Pérez et al., 2014). On the one hand, random measurement matrices include systems with multiplicative noises, which are of great interest due to their applications in various areas of communication, image processing, etc. and, on the other hand, systems with uncertain observations or sensor gain degradation are also special cases of systems with random measurement matrices. These considerations, among other reasons, justify the fact that, over the last decade, the study of the estimation problem in this type of systems has gained considerable interest; see, for example, Hu et al. (2013b), Linares-Pérez et al. (2014), Caballero-Águila et al. (2015), Yang et al. (2016), Sun et al. (2017), Wang & Zhou (2017), Han et al. (2018), Caballero-Águila et al. (2019c), Sun (2020), Liu et al. (2020), and references therein, for some representative findings.

Another source of uncertainty that commonly affects networked systems is the presence of multiplicative random disturbances in the system model, that account for different stochastic phenomena, such as random failures of the components, noisy communication channels, state-dependent sensor errors and environmental changes. The design of estimation algorithms under the effects of multiplicative noises in the system state and measurement equations has become a hot research topic. The minimum-variance linear estimation problem in discrete-time linear systems with white multiplicative noises is investigated in Wang et al. (2016). Centralized and distributed fusion filtering algorithms are designed in Tian et al. (2016) –for networked systems with correlated white multiplicative noises in the state and measurement equations– and, also, in Caballero-Águila et al. (2017b) –for sensor networks whose measured outputs include stochastic uncertainties caused by multiplicative noises and randomly missing measurements–. The state estimation problem for uncertain networked systems in the presence of multiplicative noises and fading channels has also been addressed, for example, in Li et al. (2018) and Wang & Sun (2019), under different assumptions about the control input and the additive noises involved in the system model. More recently, linear and non-linear uncertain systems with correlated white multiplicative noises have been considered in Ma & Sun (2020b) and Cheng et al. (2021), respectively, and the Tobit Kalman fusion estimation problem is investigated in Geng et al. (2021) for a class of multi-sensor systems subject to parametric uncertainties (characterized by multiplicative noises in both state and observation equations) and measurement censoring.

Concerning the system additive noises, there are many engineering applications in which the measurement noises are state dependent and, consequently, the different sensor noises are cross-correlated to each other and, also, correlated with the process noise. Moreover, noise correlation is found, for example, when the state process is observed or transmitted in a common noisy environment, as well as in the augmented systems frequently used to deal with random transmission delays and packet dropouts.

Regarding the correlation between the additive noises of the state and measurement equations, even though several papers consider that the sensor noises are correlated with the system noise at the same time step (see e.g. Ding et al., 2019; Lin & Sun, 2018; Ma & Sun, 2019; Zhao et al., 2019, and references therein), the case of correlation at the previous time step is of great interest, since this kind of correlation arises in many practical situations and, also, when a continuous-time system is discretised (see e.g. Caballero-Águila et al., 2019a; Lin & Sun, 2019; Zhao et al., 2020, among others).

In networked systems, random delays and/or packet dropouts during data transmissions are usually unavoidable, due to unreliable communication medium or space and bandwidth limitations that may cause nodes to fail or collapse. Therefore, it is not surprising that, over the last decades, numerous authors have addressed the fusion estimation problem from measurements subject to these random phenomena. Using independent Bernoulli random variables to model the possibility of random delays, the fusion estimation problem has been addressed in several papers (see e.g. Feng, 2014; Hu et al., 2013c). Also, many research efforts have been focused on removing the independence hypothesis about the delays and, simultaneously, considering some kind of uncertainties in the sensor measurements (see e.g. Caballero-Águila et al., 2016b; García-Ligero et al., 2020, among others).

When dealing with random packet dropouts, a key issue is how to compensate the lost measurements. A popular and simple compensation approach consists of using the last successfully transmitted measurement if the real one is not available –hold-input compensation–. Based on this compensation strategy, many estimation results have been reported (see e.g. Caballero-Águila et al., 2017a; Chen et al., 2017; Guo, 2017, and references therein).

This strategy has the handicap of using only the latest measurement data, ignoring the historical data previously received. To overcome this issue, a new methodology consisting of compensating each lost measurement packet by its predictor –prediction compensation– is being widely used (see e.g. Caballero-Águila et al., 2017c; Cheng et al., 2021; Ding et al., 2019; Ma & Sun, 2020b; Sun et al., 2016; Tan et al., 2020; Zhao et al., 2019, among others). These papers consider systems subject to packet dropouts, but do not cover the possibility of random delays. In many practical situations, when a packet is not received after a certain time delay, it is usually considered as definitely lost, so both phenomena commonly appear in the same context. Using the hold-input compensation strategy, the fusion estimation problem in systems with random parameter matrices, whose measurements are subject to random delays and packet dropouts, is addressed in Caballero-Águila et al. (2016a) and Caballero-Águila et al. (2019b). Based on the system augmentation technique, an optimal linear filtering algorithm is proposed in Ma & Sun (2020a) for time-invariant systems with both random delays and packet dropouts, using the prediction compensation technique to replace lost data. Under the same compensation strategy, in Uribe et al. (2021), unbiased finite impulse response, Kalman-type and game theory H_∞ filtering algorithms are developed for systems with randomly delayed data and packet dropouts.

The above considerations give rise to the conclusion that, in spite of the vast existing literature on the estimation problem over sensor networks, there are still interesting challenges which stimulate the main motivation of the current research. Actually, in this paper, a comprehensive system model incorporating some common random phenomena appearing in real networked environments –namely, state-dependent multiplicative noise, random parameter matrices, random transmission delays and packet dropouts– is considered under the assumption that the sensor noises are cross-correlated and, in turn, they are correlated with the state noise at the previous time step. Using the prediction compensation strategy, a distributed fusion filtering algorithm is designed for this class of systems.

Related work. Some of the most closely related papers in the literature are Caballero-Águila et al. (2016a) and Ma & Sun (2020a). In Caballero-Águila et al. (2016a), random one-step delays and packet dropouts with different rates are supposed to exist in data transmissions but, unlike the current paper, the measurement noises and the process noise are assumed to be independent and packet dropouts are compensated with the last measurement successfully received (hold-input compensation). In Ma & Sun (2020a), the optimal linear filtering problem is addressed for a class of time-invariant systems with one-step random delays and packet dropouts, but no uncertainties – neither in the state equation nor in the outputs before transmission– are considered in the state-space model. Moreover, unlike the current paper, the additive noises are assumed to be mutually independent and the augmentation technique is used to deal with the delays and packet dropouts.

Main contributions. The key advantages and contributions of the proposed system model and distributed fusion filtering estimators are summarized as follows:

- *System model.*

  (a) In order to cover different stochastic uncertainties, the current system model considers multiplicative noise and random parameter matrices in the state and measurement equations, respectively, which allows us to deal with a wide variety of real situations, where the state-space model involves uncertainties that cannot be described only by additive disturbances; moreover, the measurement additive noises are assumed to be cross-correlated and correlated with the process noise at the previous time step, a realistic assumption in many networked environments. Thus, a unified framework is provided to manage different simultaneous network-induced phenomena and a wide class of uncertain networked systems can be considered.

  (b) Besides the uncertainties in the measured outputs, random one-step delays and packet dropouts are assumed to occur during the transmission from the sensor side to the local processors; both phenomena are usually unavoidable during data transmissions and commonly appear in the same context.

  (c) To replace the lost data packets, the prediction compensation technique is used (that is, the estimator of the current sensor measurement based on the information received previously is employed to compensate the packet loss), which, in comparison with the hold-input compensation technique, provides better estimators, since all the historical data received previously are considered.

- *Distributed fusion filtering estimators.*

  (d) The innovation approach is used to obtain local filtering estimators globally optimal in the linear LS sense. The filter structure is recursive, very simple computationally and suitable for online applications.

  (e) A distributed fusion filter is designed as the minimum-variance matrix-weighted linear combination of the LS local linear filtering estimators provided by the different sensor nodes. The algorithm takes into account both the influence of the uncertainties described by the random parameter matrices and that of the transmission packet dropout and delay phenomena.

  (f) Unlike the derivation of most conventional Kalman-type estimation algorithms, the mathematical design of the proposed algorithm is based on a mild factorisation property of the state covariance function (covariance information) which is deduced

from the system state model, but such equation itself is not directly used, thus providing a general frame to deal with different kinds of signals when their evolution equation is not known.

(g) The estimators are obtained without the need of augmenting the state; so, the dimension of the designed estimators is the same as that of the original state, thus reducing the computational cost compared with the augmentation method.

Paper structure. The rest of the paper is organised as follows. In Section 2, the system model is described and the estimation problem is formulated. A recursive algorithm for the local LS linear filters is presented in Section 3. In Section 4, the proposed distributed fusion filter is designed as the optimal matrix-weighted linear combination of the local estimators. A simulation study is discussed in Section 5 and some conclusions are drawn in Section 6. Finally, the proofs of the main results are detailed in three appendixes.

Notation. As far as possible, standard mathematical notation will be used. $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ denote the n-dimensional Euclidean space and the set of all $m\times n$ real matrices, respectively. The superscript T represents the transpose operation of a vector or matrix and M^{-1} denotes the inverse of the matrix M. If not explicitly stated, the vectors and matrices are assumed to be of suitable dimensions, compatible with algebraic operations. The shorthand $(M_1|\cdots|M_k)$ will represent a partitioned matrix whose blocks are the submatrices $M_1,\ldots,M_k$. $\delta_{k,s}$ denotes the Kronecker delta function. $E[\cdot]$ stands for the mathematical expectation operator. $\Sigma_{k,s}^{ab^{(ij)}}$ represents the covariance function of the random vectors $a_k^{(i)}$ and $b_s^{(j)}$, $Cov[a_k^{(i)},b_s^{(j)}]=E\big[\big(a_k^{(i)}-\overline{a}_k^{(i)}\big)\big(b_s^{(j)}-\overline{b}_s^{(j)}\big)^T\big]$ (if $a=b$, we will just write $\Sigma_{k,s}^{a^{(ij)}}$). The LS linear estimator of a random vector $a_k^{(i)}$ based on the observations $y_1^{(i)},\ldots,y_L^{(i)}$ will be denoted by $\widehat{a}_{k/L}^{(i)}$. For any function $G_{k,s}$, depending on the time instants k and s, we will write $G_k=G_{k,k}$ for simplicity; analogously, $F^{(i)}=F^{(ii)}$ will be written for any function $F^{(ij)}$, depending on the sensors i and j.

2. System model and problem statement

We are concerned with the distributed fusion estimation problem in networked systems whose measurements –linearly related to the system state– are susceptible to suffering random delays and packet losses during data transmission. The additive noises at the different sensors are assumed to be cross-correlated and coupled with the process noise. Also, in order to cover some usual network-induced phenomena, random uncertainties, including multiplicative noises in the state equation and random parameter matrices in the sensor output measurement equations, are incorporated in the system model.

State-space model with random measurement matrices. Consider a discrete-time dynamic process, $\{x_k\}_{k\geq 0}$, observed over a networked system with m sensors, and described by the following state-space model:

$$\begin{aligned} x_{k+1}&=(\Phi_k+\xi_k\Psi_k)\,x_k+w_k, \quad k\geq 0,\\ z_k^{(i)}&=H_k^{(i)}x_k+v_k^{(i)}, \quad k\geq 1, \quad i=1,\ldots,m, \end{aligned} \tag{1}$$

where $x_k\in\mathbb{R}^{n_x}$ is the state vector at time k, $\{w_k\}_{k\geq 0}$ is the process noise, $\{\xi_k\}_{k\geq 0}$

is a scalar multiplicative noise and Φ_k, Ψ_k are known time-dependent matrices. For $i = 1, \ldots, m$, $z_k^{(i)} \in \mathbb{R}^{n_z}$ is the measured output from the ith sensor at time k, $\{v_k^{(i)}\}_{k\geq 1}$ are the additive measurement noises and $\{H_k^{(i)}\}_{k\geq 1}$ are sequences of random parameter matrices.

Observation transmission model with random delays and packet losses. As already indicated, in the networked system under consideration, random one-step delays and packet dropouts occur during the sensor measurement transmissions to the local processors. In order to reduce the effect of packet losses without unduly overloading the network traffic, at each time instant, the measured output is re-transmitted twice and, hence, the local processor can receive the same packet in two consecutive time instants. Moreover, it is assumed that at most one packet is received by the local processor at each sampling time. More specifically, at every sampling time k, the local processor of the ith sensor node, where the local filter is obtained, may receive the current sensor measurement $z_k^{(i)}$, the delayed one-step sensor measurement $z_{k-1}^{(i)}$ or nothing, and at most one data packet (the one that arrives first) is used to update the filter. When nothing is received, the predictor $\widehat{z}_{k/k-1}^{(i)}$ of the current sensor measurement $z_k^{(i)}$ is taken to compensate the packet loss. Hence, for $i = 1, \ldots, m$, the following mathematical model for the measurements used to update the filter is considered:

$$\begin{aligned} y_k^{(i)} &= \alpha_k^{(i)} z_k^{(i)} + \lambda_k^{(i)} z_{k-1}^{(i)} + \rho_k^{(i)} \widehat{z}_{k/k-1}^{(i)}, \quad k \geq 2; \\ y_1^{(i)} &= z_1^{(i)}, \end{aligned} \tag{2}$$

where $\alpha_k^{(i)}$, $\lambda_k^{(i)}$ and $\rho_k^{(i)}$ are Bernoulli random variables, such that $\alpha_k^{(i)} + \lambda_k^{(i)} + \rho_k^{(i)} = 1$.

Model assumptions. To address the optimal linear filtering problem under the LS optimality criterion, the following assumptions are made:

(i) The initial state x_0 is a zero-mean random vector with known covariance Σ_0^x.

(ii) $\{\xi_k\}_{k\geq 0}$ is a zero-mean scalar white process with known variance, $Var(\xi_k) = V_k^{\xi}$.

(iii) $\{H_k^{(i)}\}_{k\geq 1}$, $i = 1, \ldots, m$, are independent sequences of independent random parameter matrices with known first-order and second-order moments, $E\big[h_{pq}^{(i)}(k)\big]$ and $E\big[h_{pq}^{(i)}(k) h_{p'q'}^{(j)}(k)\big]$, for $p, p' = 1, \ldots, n_z$ and $q, q' = 1, \ldots, n_x$, where $h_{pq}^{(i)}(k)$ is the (p, q)-th entry of $H_k^{(i)}$. We will denote $\overline{H}_k^{(i)} \equiv E\big[H_k^{(i)}\big]$.

(iv) The noises $\{w_k\}_{k\geq 0}$ and $\{v_k^{(i)}\}_{k\geq 1}$, $i = 1, \ldots, m$, are zero-mean second-order white processes and the following expectations are known:

$$\begin{aligned} &E[w_k w_s^T] = Q_k \delta_{k,s}, \qquad E[v_k^{(i)} v_s^{(j)T}] = R_k^{(ij)} \delta_{k,s}, \\ &E[w_{k-1} v_s^{(i)T}] = S_k^{(i)} \delta_{k,s}, \qquad i, j = 1, \ldots, m. \end{aligned}$$

(v) For $i = 1, \ldots, m$, the processes $\{\alpha_k^{(i)}\}_{k\geq 2}$, $\{\lambda_k^{(i)}\}_{k\geq 2}$ and $\{\rho_k^{(i)}\}_{k\geq 2}$ are sequences of independent Bernoulli random variables, satisfying $\alpha_k^{(i)} + \lambda_k^{(i)} + \rho_k^{(i)} = 1$, with known probabilities

$$P(\alpha_k^{(i)} = 1) = \overline{\alpha}_k^{(i)}, \quad P(\lambda_k^{(i)} = 1) = \overline{\lambda}_k^{(i)}, \quad P(\rho_k^{(i)} = 1) = \overline{\rho}_k^{(i)}, \quad \forall k \geq 2.$$

Also, for $i \neq j$, we assume that $\left(\{\alpha_k^{(i)}\}_{k\geq 2}, \{\lambda_k^{(i)}\}_{k\geq 2}, \{\rho_k^{(i)}\}_{k\geq 2}\right)$ is independent of $\left(\{\alpha_k^{(j)}\}_{k\geq 2}, \{\lambda_k^{(j)}\}_{k\geq 2}, \{\rho_k^{(j)}\}_{k\geq 2}\right)$.

(vi) For $i = 1, \dots, m$, the initial state x_0 and the stochastic processes $\{\xi_k\}_{k\geq 0}$, $\{H_k^{(i)}\}_{k\geq 1}$ and $\left(\{\alpha_k^{(i)}\}_{k\geq 2}, \{\lambda_k^{(i)}\}_{k\geq 2}, \{\rho_k^{(i)}\}_{k\geq 2}\right)$ are mutually independent and they are independent of the additive noises $\left(\{w_k\}_{k\geq 0}, \{v_k^{(i)}\}_{k\geq 1}\right)$.

Remark 1. From assumption *(v)*, it is clear that, for each $i = 1, \dots, m$ and $k \geq 2$, only one of the Bernoulli random variables $\alpha_k^{(i)}, \lambda_k^{(i)}$ and $\rho_k^{(i)}$ will be equal to one. For every sensor node i, these variables model the three mutually exclusive events –receiving the current sensor measurement, the delayed one-step sensor measurement or nothing, respectively– that can occur at the sampling time k with the following probabilities:

- If $\alpha_k^{(i)} = 1$, the current measurement arrives first and is processed. Consequently, $y_k^{(i)} = z_k^{(i)}$. Packet on time reception probability: $\overline{\alpha}_k^{(i)} = P(\alpha_k^{(i)} = 1)$.
- If $\lambda_k^{(i)} = 1$, the measurement at time $k-1$ is first received at time k, so the measurement processed at that time is $y_k^{(i)} = z_{k-1}^{(i)}$. One-step delay probability: $\overline{\lambda}_k^{(i)} = P(\lambda_k^{(i)} = 1)$.
- If $\rho_k^{(i)} = 1$, nothing arrives at time k and this loss is compensated by the predictor of the current measurement, $y_k^{(i)} = \widehat{z}_{k/k-1}^{(i)}$. Packet loss probability: $\overline{\rho}_k^{(i)} = P(\rho_k^{(i)} = 1)$.

Remark 2. State covariance matrices: From assumptions *(i)-(iv)* and *(vi)*, the covariance function of the system state can be written as $\Sigma_{k,s}^x = \Phi_{k,s}\Sigma_s^x,\ s \leq k$, where $\Phi_{k,s} = \Phi_{k-1}\cdots\Phi_s$, for $s < k$, $\Phi_{k,k} = I$, and the matrices $\Sigma_s^x,\ s \geq 1$, satisfy the following recursive relation:

$$\Sigma_s^x = \Phi_{s-1}\Sigma_{s-1}^x\Phi_{s-1}^T + V_{s-1}^\xi\Psi_{s-1}\Sigma_{s-1}^x\Psi_{s-1}^T + Q_{s-1}, \qquad s \geq 1.$$

Consequently, denoting $A_k = \Phi_{k,0}$ and $B_s^T = \Phi_{s,0}^{-1}\Sigma_s^x$, the covariance function of the system state can be factorised as follows:

$$\Sigma_{k,s}^x = A_k B_s^T, \qquad s \leq k.$$

Moreover, using that by assumption *(iv)*, the noise w_{s-1} is correlated with the sensor noises at time s, for $i = 1, \dots, m$ we have that $E[x_k v_s^{(i)T}] = \Phi_{k,s}E[x_s v_s^{(i)T}] = \Phi_{k,s}E[w_{s-1}v_s^{(i)T}] = \Phi_{k,s}S_s^{(i)}$, for $s \leq k$, and $E[x_k v_s^{(i)T}] = 0$, for $s > k$. This means that the sensor noises are correlated with the system state at the same and subsequent time steps and, denoting $C_s^{(i)T} = \Phi_{s,0}^{-1}S_s^{(i)}$, these covariance functions can also be factorised as follows:

$$E[x_k v_s^{(i)T}] = \begin{cases} A_k C_s^{(i)T}, & s \leq k \\ 0, & s > k \end{cases}, \qquad i = 1, \dots, m.$$

Remark 3. Measured output covariance matrices: Using (1) for $z_k^{(i)}$, it is clear that the covariance matrices $\Sigma_{k,s}^{z^{(ij)}} = E\big[z_k^{(i)} z_s^{(j)T}\big]$, $i, j = 1, \dots, m$, satisfy

$$\Sigma_{k,s}^{z^{(ij)}} = E\big[H_k^{(i)} x_k x_s^T H_s^{(j)T}\big] + R_k^{(ij)} \delta_{k,s} + \overline{H}_k^{(i)} E[x_k v_s^{(j)T}] + E[v_k^{(i)} x_s^T] \overline{H}_s^{(j)T}, \quad s \le k.$$

From the model assumptions and the factorisations stated in Remark 2, we have:

- For arbitrary $i, j = 1, \dots, m$,

$$E\big[H_k^{(i)} x_k x_s^T H_s^{(j)T}\big] = E\big[H_k^{(i)} E[x_k x_s^T] H_s^{(j)T}\big] = E\big[H_k^{(i)} A_k B_s^T H_s^{(j)T}\big], \quad s \le k,$$

  where $E\big[H_k^{(i)} A_k B_s^T H_s^{(j)T}\big] = \overline{H}_k^{(i)} A_k B_s^T \overline{H}_s^{(j)T}$ for $j \ne i$ or $s \ne k$, and the entries of $E\big[H_k^{(i)} A_k B_k^T H_k^{(i)T}\big]$ are computed by:

$$\Big(E\big[H_k^{(i)} A_k B_k^T H_k^{(i)T}\big]\Big)_{pq} = \sum_{a=1}^{n_x} \sum_{b=1}^{n_x} E[h_{pa}^{(i)}(k) h_{qb}^{(i)}(k)] (A_k B_k^T)_{ab}, \quad p, q = 1, \dots, n_z.$$

- For $i, j = 1, \dots, m$, and $s \le k$, we have that

$$E[x_k v_s^{(j)T}] = A_k C_s^{(j)T} \quad \textit{and} \quad E[v_k^{(i)} x_s^T] = C_k^{(i)} A_k^T \delta_{k,s}.$$

Hence, the following expression for $\Sigma_{k,s}^{z^{(ij)}}$ is deduced:

$$\Sigma_{k,s}^{z^{(ij)}} = E\left[H_k^{(i)} A_k B_s^T H_s^{(j)T}\right] + \overline{H}_k^{(i)} A_k C_s^{(j)T} + \Big(C_k^{(i)} A_k^T \overline{H}_k^{(j)T} + R_k^{(ij)}\Big) \delta_{k,s}, \quad s \le k. \quad (3)$$

Estimation problem. Our aim is to design a recursive filtering algorithm using the distributed fusion architecture, under which the local processor of each individual node produces a local LS linear filter of the system state, x_k, based on its own sensor measurements; afterwards, these local filters are fused in the processing centre, where the proposed distributed filter is generated as the LS matrix-weighted linear combination of the local ones. First, in *Section 3*, a recursive algorithm for the local LS linear filters, $\widehat{x}_{k/k}^{(i)}$, $i = 1, \dots, m$, will be obtained. Then, in *Section 4*, the cross-correlation matrices between any two local filters, $\Sigma_{k/k}^{\widehat{x}^{(ij)}} = E\big[\widehat{x}_{k/k}^{(i)} \widehat{x}_{k/k}^{(j)T}\big]$, $i, j = 1, \dots, m$, will be derived and the proposed distributed fusion filter weighted by matrices, $\widehat{x}_{k/k}^D$, will be designed under the LS optimality criterion.

Traditionally, when the state evolution equation is known, recursive estimation algorithms are usually based on the state-space model. However, the design of the filtering algorithms proposed in this paper will be based on the factorisation property of the state covariance function deduced from the state transition equation (Remark 2), but such equation itself will not be directly used. This fact makes the structure of the proposed filtering algorithms be different from that of most standard algorithms based on the state-space model, but it also makes such algorithms useful to estimate a wide class of stationary or non-stationary state processes that admit this kind of covariance factorisation, which is undoubtedly a significant advantage.

3. Local LS linear filters

With the aim of designing a recursive algorithm for the local LS linear filters, $\widehat{x}^{(i)}_{k/k}$, $i = 1, \ldots, m$, we will use an innovation approach.

3.1. Innovation process

For each sensor node, $i = 1, \ldots, m$, let us consider the innovation at time k,

$$\mu^{(i)}_k = y^{(i)}_k - \widehat{y}^{(i)}_{k/k-1}, \quad k \geq 2; \qquad \mu^{(i)}_1 = y^{(i)}_1, \tag{4}$$

where $\widehat{y}^{(i)}_{k/k-1}$ denotes the LS linear estimator of $y^{(i)}_k$ based on the observations, $y^{(i)}_1, \ldots, y^{(i)}_{k-1}$. Its relevance lies in the fact that the observation and innovation processes are uniquely determined one by the other, so the observation process can be replaced by the innovation one. Therefore, the LS linear estimator, $\widehat{a}^{(i)}_{k/L}$, of a random vector $a^{(i)}_k$ based on the observations up to an arbitrary time L, $y^{(i)}_1, \ldots, y^{(i)}_L$, can be expressed as linear combination of the innovations $\mu^{(i)}_1, \ldots, \mu^{(i)}_L$ as follows:

$$\widehat{a}^{(i)}_{k/L} = \sum_{h=1}^{L} E\big[a^{(i)}_k \mu^{(i)T}_h\big] \Pi^{(i)-1}_h \mu^{(i)}_h, \quad k \geq 1, \tag{5}$$

where $\Pi^{(i)}_h \equiv E\big[\mu^{(i)}_h \mu^{(i)T}_h\big]$. The main advantage of using this expression for the LS linear estimator arises from the fact that the innovations constitute a zero-mean white process, which simplifies substantially the mathematical derivations. Expression (5) for the LS linear estimators, together with the *Orthogonal Projection Lemma* (OPL), are the key results to deduce the proposed recursive local filtering algorithm.

Starting from (5) and denoting $\mathcal{X}^{(i)}_{k,h} = E\big[x_k \mu^{(i)T}_h\big]$, the local LS linear estimators of the system state, x_k, based on the observations $y^{(i)}_1, \ldots, y^{(i)}_L$, are expressed as

$$\widehat{x}^{(i)}_{k/L} = \sum_{h=1}^{L} \mathcal{X}^{(i)}_{k,h} \Pi^{(i)-1}_h \mu^{(i)}_h, \quad i = 1, \ldots, m, \tag{6}$$

and the main difficulty in obtaining these state estimators is to find an expression for the innovation $\mu^{(i)}_h = y^{(i)}_h - \widehat{y}^{(i)}_{h/h-1}$ (or, equivalently, an expression for the observation predictor $\widehat{y}^{(i)}_{h/h-1}$) that allows us to determine –in a simple way– the coefficients $\mathcal{X}^{(i)}_{k,h} = E\big[x_k \mu^{(i)T}_h\big]$ and the matrices $\Pi^{(i)}_h = E\big[\mu^{(i)}_h \mu^{(i)T}_h\big]$ involved in (6).

From the OPL, we have that $\widehat{z}^{(i)}_{h/h-1} = \overline{H}^{(i)}_h \widehat{x}^{(i)}_{h/h-1}$, then, by substitution in (2),

$$y^{(i)}_h = \alpha^{(i)}_h z^{(i)}_h + \lambda^{(i)}_h z^{(i)}_{h-1} + \rho^{(i)}_h \overline{H}^{(i)}_h \widehat{x}^{(i)}_{h/h-1}, \quad h \geq 2.$$

Now, using (1) for $z^{(i)}_h$ and $z^{(i)}_{h-1}$, we have

$$y^{(i)}_h = \alpha^{(i)}_h H^{(i)}_h x_h + \alpha^{(i)}_h v^{(i)}_h + \lambda^{(i)}_h H^{(i)}_{h-1} x_{h-1} + \lambda^{(i)}_h v^{(i)}_{h-1} + \rho^{(i)}_h \overline{H}^{(i)}_h \widehat{x}^{(i)}_{h/h-1}, \quad h \geq 2.$$

Since $H_{h-1}^{(i)}$ is correlated with the observations up to time $h-1$, in order to simplify the derivation of the observation predictor, we add and subtract $\lambda_h^{(i)}\overline{H}_{h-1}^{(i)}x_{h-1}$, getting the following expression for $y_h^{(i)}$:

$$y_h^{(i)} = X_h^{(i)} + V_h^{(i)}, \quad h \geq 2, \tag{7}$$

where

$$X_h^{(i)} = \alpha_h^{(i)}H_h^{(i)}x_h + \lambda_h^{(i)}\overline{H}_{h-1}^{(i)}x_{h-1} + \rho_h^{(i)}\overline{H}_h^{(i)}\widehat{x}_{h/h-1}^{(i)}, \quad h \geq 2, \tag{8}$$

$$V_h^{(i)} = \alpha_h^{(i)}v_h^{(i)} + \lambda_h^{(i)}\left(H_{h-1}^{(i)} - \overline{H}_{h-1}^{(i)}\right)x_{h-1} + \lambda_h^{(i)}v_{h-1}^{(i)}, \quad h \geq 2. \tag{9}$$

Consequently, $\widehat{y}_{h/h-1}^{(i)} = \widehat{X}_{h/h-1}^{(i)} + \widehat{V}_{h/h-1}^{(i)}$, $h \geq 2$, and $\widehat{X}_{h/h-1}^{(i)}$, $\widehat{V}_{h/h-1}^{(i)}$ must be calculated.

- From the independence assumptions of the model, using (8), we have that

$$\widehat{X}_{h/h-1}^{(i)} = \left(\overline{\alpha}_h^{(i)} + \overline{\rho}_h^{(i)}\right)\overline{H}_h^{(i)}\widehat{x}_{h/h-1}^{(i)} + \overline{\lambda}_h^{(i)}\overline{H}_{h-1}^{(i)}\widehat{x}_{h-1/h-1}^{(i)}, \quad h \geq 2.$$

- Using (5) for $\widehat{V}_{h/h-1}^{(i)}$ and taking into account that $V_h^{(i)}$, defined by (9), is uncorrelated with $\mu_1^{(i)}, \ldots, \mu_{h-2}^{(i)}$, denoting $\mathcal{V}_{h,h-1}^{(i)} = E\left[V_h^{(i)}\mu_{h-1}^{(i)T}\right]$, it is clear that $\mathcal{V}_{h,h-1}^{(i)} = E\left[V_h^{(i)}y_{h-1}^{(i)T}\right]$ and

$$\widehat{V}_{h/h-1}^{(i)} = \mathcal{V}_{h,h-1}^{(i)}\Pi_{h-1}^{(i)-1}\mu_{h-1}^{(i)}, \quad h \geq 2.$$

So, taking into account that $\overline{\alpha}_k^{(i)} + \overline{\rho}_k^{(i)} + \overline{\lambda}_k^{(i)} = 1$, the observation predictor is expressed as

$$\widehat{y}_{h/h-1}^{(i)} = (1-\overline{\lambda}_h^{(i)})\overline{H}_h^{(i)}\widehat{x}_{h/h-1}^{(i)} + \overline{\lambda}_h^{(i)}\overline{H}_{h-1}^{(i)}\widehat{x}_{h-1/h-1}^{(i)} + \mathcal{V}_{h,h-1}^{(i)}\Pi_{h-1}^{(i)-1}\mu_{h-1}^{(i)}, \quad h \geq 2. \tag{10}$$

Note that, in order to obtain the observation predictor –and, from it, the innovation– at each time instant, we need not only the innovation at the previous instant, but also the state prediction and filtering estimators, which will be simultaneously obtained in the next section.

For simplicity, from this section onwards, for $i = 1, \ldots, m$, the following notation will be used:

$$\begin{array}{l}
\mathbb{A}_k \equiv \big(A_k|A_k\big). \\
\mathbf{H}_{\Psi_k}^{(i)} \equiv \overline{\alpha}_k^{(i)}\overline{H}_k^{(i)}\Psi_k + \overline{\lambda}_k^{(i)}\overline{H}_{k-1}^{(i)}\Psi_{k-1}, \ k \geq 2; \ \ \mathbf{H}_{\Psi_1}^{(i)} \equiv \overline{H}_1^{(i)}\Psi_1 \ \text{ for } (\Psi = \mathbb{A}, \ B). \\
\mathbf{H}_{\mathbb{A}_k}^{*(i)} \equiv \big(1-\overline{\lambda}_k^{(i)}\big)\overline{H}_k^{(i)}\mathbb{A}_k + \overline{\lambda}_k^{(i)}\overline{H}_{k-1}^{(i)}\mathbb{A}_{k-1}, \ k \geq 2. \\
\mathbf{C}_k^{(i)} \equiv \overline{\alpha}_k^{(i)}C_k^{(i)} + \overline{\lambda}_k^{(i)}C_{k-1}^{(i)}, \quad k \geq 2; \quad \mathbf{C}_1^{(i)} \equiv C_1^{(i)}.
\end{array} \tag{11}$$

3.2. LS linear prediction and filtering estimators of the state

We start from expression (6) for the prediction and filtering state estimators, $\widehat{x}^{(i)}_{k/s}$, $s \le k$, and we calculate the matrix coefficients $\mathcal{X}^{(i)}_{k,h} = E\big[x_k \mu^{(i)T}_h\big] = E\big[x_k y^{(i)T}_h\big] - E\big[x_k \widehat{y}^{(i)T}_{h/h-1}\big]$, for $1 \le h \le k$,. Using expression (10) of the observation predictor, these coefficients can be written as follows:

$$\begin{aligned} \mathcal{X}^{(i)}_{k,h} = &\ E\big[x_k y^{(i)T}_h\big] - \big(1-\overline{\lambda}^{(i)}_h\big) E\big[x_k \widehat{x}^{(i)T}_{h/h-1}\big]\overline{H}^{(i)T}_h - \overline{\lambda}^{(i)}_h E\big[x_k \widehat{x}^{(i)T}_{h-1/h-1}\big]\overline{H}^{(i)T}_{h-1} \\ & -\mathcal{X}^{(i)}_{k,h-1}\Pi^{(i)-1}_{h-1}\mathcal{V}^{(i)T}_{h,h-1}, \quad h \ge 2; \\ \mathcal{X}^{(i)}_{k,1} = &\ E\big[x_k y^{(i)T}_1\big]. \end{aligned}$$

To obtain $E\big[x_k y^{(i)T}_h\big]$, we write $y^{(i)}_h = X^{(i)}_h + V^{(i)}_h$, $\quad h \ge 2$, with $X^{(i)}_h$ and $V^{(i)}_h$ given in (8) and (9). Then, taking into account the independence assumption *(iv)*, we have

$$\begin{aligned} E[x_k y^{(i)T}_h] = &\ \overline{\alpha}^{(i)}_h E[x_k x^T_h]\overline{H}^{(i)T}_h + \overline{\lambda}^{(i)}_h E[x_k x^T_{h-1}]\overline{H}^{(i)T}_{h-1} + \overline{\rho}^{(i)}_h E[x_k \widehat{x}^{(i)T}_{h/h-1}]\overline{H}^{(i)T}_h \\ & +\overline{\alpha}^{(i)}_h E[x_k v^{(i)T}_h] + \overline{\lambda}^{(i)}_h E[x_k v^{(i)T}_{h-1}], \quad h \ge 2, \end{aligned}$$

and, using the covariance factorisation established in *Remark 2*, we get

$$\begin{aligned} E[x_k y^{(i)T}_h] = &\ \overline{\alpha}^{(i)}_h A_k B^T_h \overline{H}^{(i)T}_h + \overline{\lambda}^{(i)}_h A_k B^T_{h-1}\overline{H}^{(i)T}_{h-1} + \overline{\rho}^{(i)}_h E[x_k \widehat{x}^{(i)T}_{h/h-1}]\overline{H}^{(i)T}_h \\ & +\overline{\alpha}^{(i)}_h A_k C^{(i)T}_h + \overline{\lambda}^{(i)}_h A_k C^{(i)T}_{h-1}, \quad h \ge 2. \end{aligned}$$

Clearly, $E[x_k y^{(i)T}_1] = E[x_k z^{(i)T}_1] = A_k B^T_1 \overline{H}^{(i)T}_1 + A_k C^{(i)T}_1$. Hence, using the notation (11), the following expression is obtained:

$$\begin{aligned} & E\big[x_k y^{(i)T}_h\big] = \mathbb{A}_k\big(\mathbf{H}^{(i)}_{B_h}|\mathbf{C}^{(i)}_h\big)^T + \overline{\rho}^{(i)}_h E\big[x_k \widehat{x}^{(i)T}_{h/h-1}\big]\overline{H}^{(i)T}_h, \quad h \ge 2; \\ & E\big[x_k y^{(i)T}_1\big] = \mathbb{A}_k\big(\mathbf{H}^{(i)}_{B_1}|\mathbf{C}^{(i)}_1\big)^T. \end{aligned}$$

Now, using (6) for $\widehat{x}^{(i)T}_{h/h-1}$ and $\widehat{x}^{(i)T}_{h-1/h-1}$, we have that

$$\begin{aligned} \mathcal{X}^{(i)}_{k,h} = &\ \mathbb{A}_k\big(\mathbf{H}^{(i)}_{B_h}|\mathbf{C}^{(i)}_h\big)^T - \sum_{l=1}^{h-1}\mathcal{X}^{(i)}_{k,l}\Pi^{(i)-1}_l\Big[\overline{\alpha}^{(i)}_h\mathcal{X}^{(i)T}_{h,l}\overline{H}^{(i)T}_h + \overline{\lambda}^{(i)}_h\mathcal{X}^{(i)T}_{h-1,l}\overline{H}^{(i)T}_{h-1}\Big] \\ & -\mathcal{X}^{(i)T}_{k,h-1}\Pi^{(i)-1}_{h-1}\mathcal{V}^{(i)T}_{h,h-1}, \quad h \ge 2; \\ \mathcal{X}^{(i)}_{k,1} = &\ \mathbb{A}_k\big(\mathbf{H}^{(i)}_{B_1}|\mathbf{C}^{(i)}_1\big)^T. \end{aligned}$$

This expression for the matrix coefficients guarantees that

$$\mathcal{X}^{(i)}_{k,h} = \mathbb{A}_k\mathcal{E}^{(i)}_h, \quad 1 \le h \le k, \tag{12}$$

with $\mathcal{E}_h^{(i)}$ given by

$$\begin{aligned}\mathcal{E}_h^{(i)} &= \big(\mathbf{H}_{B_h}^{(i)}|\mathbf{C}_h^{(i)}\big)^T - \sum_{l=1}^{h-1}\mathcal{E}_l^{(i)}\Pi_l^{(i)-1}\mathcal{E}_l^{(i)T}\mathbf{H}_{\mathbb{A}_h}^{(i)T} - \mathcal{E}_{h-1}^{(i)}\Pi_{h-1}^{(i)-1}\mathcal{V}_{h,h-1}^{(i)T},\ h\geq 2;\\ \mathcal{E}_1^{(i)} &= \big(\mathbf{H}_{B_1}^{(i)}|\mathbf{C}_1^{(i)}\big)^T.\end{aligned} \tag{13}$$

Then, by defining the vectors $e_k^{(i)} = \sum_{h=1}^{k}\mathcal{E}_h^{(i)}\Pi_h^{(i)-1}\mu_h^{(i)},\ \ k\geq 1$, it is clear from (6) and (12) that the LS prediction and filtering state estimators are given by

$$\widehat{x}_{k/s}^{(i)} = \mathbb{A}_k e_s^{(i)},\ \ s\leq k. \tag{14}$$

Remark 4. From the definition of the vectors $e_k^{(i)}$, the following recursive relation is clear

$$e_k^{(i)} = e_{k-1}^{(i)} + \mathcal{E}_k^{(i)}\Pi_k^{(i)-1}\mu_k^{(i)},\ \ k\geq 1;\qquad e_0^{(i)} = 0,$$

and, using the orthogonality of the innovations, it is easily deduced that their covariance matrices $\Sigma_k^{e^{(i)}} = E\big[e_k^{(i)}e_k^{(i)T}\big]$ are also recursively obtained by

$$\Sigma_k^{e^{(i)}} = \Sigma_{k-1}^{e^{(i)}} + \mathcal{E}_k^{(i)}\Pi_k^{(i)-1}\mathcal{E}_k^{(i)T},\ \ k\geq 1;\qquad \Sigma_0^{e^{(i)}} = 0.$$

Remark 5. The consideration of measurement noises that are correlated with the system noise at the previous time step adds some complexity to the mathematical derivation of the local filtering estimators and their cross-covariance matrices. The current approach deal with this correlation in a natural way, without resorting to either decorrelation or augmentation methods. Nevertheless, the correlation assumption of the additive noises, $E[w_{k-1}v_h^{(i)T}] = S_k^{(i)}\delta_{k,h}$, means –as stated in Remark 2– that the sensor noises are correlated with the system state at the same and subsequent time steps, $E[x_k v_h^{(i)T}] = A_k C_h^{(i)T},\ h\leq k$. Consequently, the system state x_k is correlated with the observations $y_h^{(i)},\ h\leq k$. The main difficulty associated with this correlation assumption of the additive noises has been to establish the factorisation $\mathbb{A}_k\big(\mathbf{H}_{B_h}^{(i)}|\mathbf{C}_h^{(i)}\big)^T$ during the calculation of the correlations $E\big[x_k y_h^{(i)T}\big]$. In turn, this factorisation has been the key to demonstrate that the matrix coefficients $\mathcal{X}_{k,h}^{(i)}$ can be expressed in the separable form $\mathcal{X}_{k,h}^{(i)} = \mathbb{A}_k\mathcal{E}_h^{(i)},\ h\leq k$, a crucial property for the derivation of the covariance-based algorithms.

3.3. Local LS linear filtering algorithm

Substituting (14) into (10) and using (11), the following alternative expression for the observation predictor is obtained:

$$\widehat{y}_{k/k-1}^{(i)} = \mathbf{H}_{\mathbb{A}_k}^{*(i)}e_{k-1}^{(i)} + \mathcal{V}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mu_{k-1}^{(i)},\ \ k\geq 2. \tag{15}$$

In order to get the design of the filtering algorithm completed, the matrices $\mathcal{V}_{k,k-1}^{(i)} = E\big[V_k^{(i)} y_{k-1}^{(i)T}\big]$, $k \geq 2$, and the innovation covariances, $\Pi_k^{(i)} = E\big[\mu_k^{(i)} \mu_k^{(i)T}\big]$, $k \geq 1$, must be obtained. On the one hand, from (7), $V_k^{(i)} = y_k^{(i)} - X_k^{(i)}$, from which it is clear that $\mathcal{V}_{k,k-1}^{(i)} = \Sigma_{k,k-1}^{y^{(i)}} - E\big[X_k^{(i)} y_{k-1}^{(i)T}\big]$; on the other, from the OPL, $\Pi_k^{(i)} = \Sigma_k^{y^{(i)}} - E\big[\widehat{y}_{k/k-1}^{(i)} \widehat{y}_{k/k-1}^{(i)T}\big]$. Hence, expressions for the covariance matrices $\Sigma_k^{y^{(i)}}$, $k \geq 1$, and $\Sigma_{k,k-1}^{y^{(i)}}$, $k \geq 2$, are needed; the following lemma sets out how these covariance matrices are to be calculated.

Lemma 3.1. *For $i = 1, \dots, m$, the covariance functions $\Sigma_k^{y^{(i)}}$, $k \geq 1$, and $\Sigma_{k,k-1}^{y^{(i)}}$, $k \geq 2$, are given by*

$$\begin{aligned} \Sigma_k^{y^{(i)}} &= \overline{\alpha}_k^{(i)} \Sigma_k^{z^{(i)}} + \overline{\lambda}_k^{(i)} \Sigma_{k-1}^{z^{(i)}} + \overline{\rho}_k^{(i)} \overline{H}_k^{(i)} \mathbb{A}_k \Sigma_{k-1}^{e^{(i)}} \mathbb{A}_k^T \overline{H}_k^{(i)T}, \quad k \geq 2; \\ \Sigma_1^{y^{(i)}} &= \Sigma_1^{z^{(i)}}. \end{aligned} \tag{16}$$

$$\Sigma_{k,k-1}^{y^{(i)}} = \overline{\alpha}_k^{(i)} \Sigma_{k,k-1}^{zy^{(i)}} + \overline{\lambda}_k^{(i)} \Sigma_{k-1}^{zy^{(i)}} + \overline{\rho}_k^{(i)} \overline{H}_k^{(i)} \mathbb{A}_k \Sigma_{k-1}^{ey^{(i)}}, \quad k \geq 2, \tag{17}$$

where

$$\begin{aligned} \Sigma_{k,s}^{zy^{(i)}} &= \overline{\alpha}_s^{(i)} \Sigma_{k,s}^{z^{(i)}} + \overline{\lambda}_s^{(i)} \Sigma_{k,s-1}^{z^{(i)}} + \overline{\rho}_s^{(i)} \overline{H}_k^{(i)} \mathbb{A}_k \Sigma_{s-1}^{e^{(i)}} \mathbb{A}_s^T \overline{H}_s^{(i)T}, \quad 2 \leq s \leq k; \\ \Sigma_{k,1}^{zy^{(i)}} &= \Sigma_{k,1}^{z^{(i)}}, \end{aligned} \tag{18}$$

in which the matrices $\Sigma_{k,s}^{z^{(i)}}$ are obtained by (3).

Proof. See Appendix A. □

Taking into account the above results and the notation introduced in (11), the following local filtering algorithm is deduced.

Theorem 3.2. *For $i = 1, \dots, m$, the local LS linear filtering estimators, $\widehat{x}_{k/k}^{(i)}$, and the error covariances, $\Sigma_{k/k}^{\widetilde{x}^{(i)}} = E\big[(x_k - \widehat{x}_{k/k}^{(i)})(x_k - \widehat{x}_{k/k}^{(i)})^T\big]$, are obtained by*

$$\widehat{x}_{k/k}^{(i)} = \mathbb{A}_k e_k^{(i)}, \quad k \geq 1, \tag{19}$$

$$\Sigma_{k/k}^{\widetilde{x}^{(i)}} = A_k B_k^T - \mathbb{A}_k \Sigma_k^{e^{(i)}} \mathbb{A}_k^T, \quad k \geq 1, \tag{20}$$

where the vectors $e_k^{(i)}$ and their covariances $\Sigma_k^{e^{(i)}}$ are recursively calculated by

$$e_k^{(i)} = e_{k-1}^{(i)} + \mathcal{E}_k^{(i)} \Pi_k^{(i)-1} \mu_k^{(i)}, \quad k \geq 1; \qquad e_0^{(i)} = 0, \tag{21}$$

$$\Sigma_k^{e^{(i)}} = \Sigma_{k-1}^{e^{(i)}} + \mathcal{E}_k^{(i)} \Pi_k^{(i)-1} \mathcal{E}_k^{(i)T}, \quad k \geq 1; \qquad \Sigma_0^{e^{(i)}} = 0, \tag{22}$$

in which the matrices $\mathcal{E}_k^{(i)} = E\big[e_k^{(i)}\mu_k^{(i)T}\big]$ *are given by*

$$\begin{array}{l}\mathcal{E}_k^{(i)} = \big(\mathbf{H}_{B_k}^{(i)}|\mathbf{C}_k^{(i)}\big)^T - \Sigma_{k-1}^{e^{(i)}}\mathbf{H}_{\mathbb{A}_k}^{(i)T} - \mathcal{E}_{k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mathcal{V}_{k,k-1}^{(i)T}, \quad k \geq 2;\\ \mathcal{E}_1^{(i)} = \big(\mathbf{H}_{B_1}^{(i)}|\mathbf{C}_1^{(i)}\big)^T.\end{array} \tag{23}$$

The innovation, $\mu_k^{(i)}$*, is calculated by*

$$\begin{array}{l}\mu_k^{(i)} = y_k^{(i)} - \mathbf{H}_{\mathbb{A}_k}^{*(i)}e_{k-1}^{(i)} - \mathcal{V}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mu_{k-1}^{(i)}, \quad k \geq 2;\\ \mu_1^{(i)} = y_1^{(i)},\end{array} \tag{24}$$

and its covariance matrix, $\Pi_k^{(i)}$*, satisfies*

$$\begin{array}{l}\Pi_k^{(i)} = \Sigma_k^{y^{(i)}} - \mathbf{H}_{\mathbb{A}_k}^{*(i)}\Delta_{k-1,k}^{(i)} - \mathcal{V}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mathcal{Y}_{k,k-1}^{(i)T}, \quad k \geq 2;\\ \Pi_1^{(i)} = \Sigma_1^{y^{(i)}},\end{array} \tag{25}$$

in which the matrices $\mathcal{V}_{k,k-1}^{(i)} = E\big[V_k^{(i)}\mu_{k-1}^{(i)T}\big]$, $\mathcal{Y}_{k,k-1}^{(i)} = E\big[y_k^{(i)}\mu_{k-1}^{(i)T}\big]$ *and* $\Delta_{k-1,k}^{(i)} = E\big[e_{k-1}^{(i)}y_k^{(i)T}\big]$ *are calculated as follows:*

$$\mathcal{V}_{k,k-1}^{(i)} = \overline{\alpha}_k^{(i)}\Sigma_{k,k-1}^{zy^{(i)}} + \overline{\lambda}_k^{(i)}\Sigma_{k-1}^{zy^{(i)}} - \mathbf{H}_{\mathbb{A}_k}^{(i)T}(\Delta_{k-2,k-1}^{(i)} + \mathcal{E}_{k-1}^{(i)}), \quad k \geq 2, \tag{26}$$

$$\mathcal{Y}_{k,k-1}^{(i)} = \mathbf{H}_{\mathbb{A}_k}^{*(i)}\mathcal{E}_{k-1}^{(i)} + \mathcal{V}_{k,k-1}^{(i)}, \quad k \geq 2, \tag{27}$$

$$\Delta_{k-1,k}^{(i)} = \Sigma_{k-1}^{e^{(i)}}\mathbf{H}_{\mathbb{A}_k}^{*(i)T} + \mathcal{E}_{k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mathcal{V}_{k,k-1}^{(i)T}, \quad k \geq 2. \tag{28}$$

The covariance matrices $\Sigma_k^{y^{(i)}}$, $\Sigma_{k,k-1}^{zy^{(i)}}$ *and* $\Sigma_{k-1}^{zy^{(i)}}$ *are given in Lemma 3.1.*

Proof. See Appendix B. □

4. Distributed fusion filtering algorithm

As we have previously indicated, our final goal is to obtain a distributed fusion filter, $\widehat{x}_{k/k}^D$, as a matrix-weighted linear combination of the local estimators $\widehat{x}_{k/k}^{(i)}$, $i = 1, \ldots, m$, in which the matrix-weights will be computed by minimizing the mean squared estimation error. For this purpose, let us consider the following vector obtained by stacking the local filtering estimators: $\widehat{X}_{k/k} = \Big(\widehat{x}_{k/k}^{(1)T}, \ \ldots \ , \widehat{x}_{k/k}^{(m)T}\Big)^T$. Under the LS criterion (see, e.g., (Caballero-Águila et al., 2016a)), the distributed filtering estimator we are looking for is given by

$$\widehat{x}_{k/k}^D = E\left[x_k\widehat{X}_{k/k}^T\right]\left(E\left[\widehat{X}_{k/k}\widehat{X}_{k/k}^T\right]\right)^{-1}\widehat{X}_{k/k}, \quad k \geq 1. \tag{29}$$

Since $E\big[\widehat{X}_{k/k}\widehat{X}^T_{k/k}\big] = \Big(E\big[\widehat{x}^{(i)}_{k/k}\widehat{x}^{(j)T}_{k/k}\big]\Big)_{i,j=1,\dots,m}$ and, from the OPL,

$$E\big[x_k\widehat{X}^T_{k/k}\big] = \Big(E\big[\widehat{x}^{(1)}_{k/k}\widehat{x}^{(1)T}_{k/k}\big]\Big|\cdots\Big|E\big[\widehat{x}^{(m)}_{k/k}\widehat{x}^{(m)T}_{k/k}\big]\Big),$$

the derivation of the distributed filter (29) requires the calculation of the local filtering cross-covariance matrices $\Sigma^{\widehat{x}^{(ij)}}_{k/k} = E\big[\widehat{x}^{(i)}_{k/k}\widehat{x}^{(j)T}_{k/k}\big],\ i,j=1,\dots,m$.

4.1. Cross-covariance matrices of local filters

From (19), the cross-covariance between any two local filtering estimators is clearly given by

$$\Sigma^{\widehat{x}^{(ij)}}_{k/k} = \mathbb{A}_k\Sigma^{e^{(ij)}}_k\mathbb{A}^T_k, \quad k\geq 1, \quad i,j=1,\dots,m.$$

Consequently, we must calculate the cross-covariance matrices $\Sigma^{e^{(ij)}}_k = E\big[e^{(i)}_k e^{(j)T}_k\big]$, for which, taking into account expression (21), the innovation cross-covariances $\Pi^{(ij)}_k = E\big[\mu^{(i)}_k\mu^{(j)}_k\big]$ will also be needed. The following lemma describes how these cross-covariance matrices $\Sigma^{e^{(ij)}}_k$ and $\Pi^{(ij)}_k$ are to be calculated.

Lemma 4.1. *For $i,j=1,\dots,m$, the matrices $\Sigma^{e^{(ij)}}_k = E\big[e^{(i)}_k e^{(j)T}_k\big]$ are obtained by*

$$\begin{array}{l}\Sigma^{e^{(ij)}}_k = \Sigma^{e^{(ij)}}_{k-1} + \mathcal{E}^{(ij)}_{k-1,k}\Pi^{(j)-1}_k\mathcal{E}^{(j)T}_k + \mathcal{E}^{(i)}_k\Pi^{(i)-1}_k\mathcal{E}^{(ji)T}_k, \quad k\geq 1;\\ \Sigma^{e^{(ij)}}_0 = 0,\end{array} \tag{30}$$

in which the matrices $\mathcal{E}^{(ij)}_k = E\big[e^{(i)}_k\mu^{(j)T}_k\big]$ and $\mathcal{E}^{(ij)}_{k-1,k} = E\big[e^{(i)}_{k-1}\mu^{(j)T}_k\big]$ are given by

$$\mathcal{E}^{(ij)}_k = \mathcal{E}^{(ij)}_{k-1,k} + \mathcal{E}^{(i)}_k\Pi^{(i)-1}_k\Pi^{(ij)}_k, \quad k\geq 1, \tag{31}$$

$$\begin{array}{l}\mathcal{E}^{(ij)}_{k-1,k} = \big(\Sigma^{e^{(i)}}_{k-1} - \Sigma^{e^{(ij)}}_{k-1}\big)\mathbf{H}^{(j)T}_{\mathbb{A}_k} + \mathcal{E}^{(i)}_{k-1}\Pi^{(i)-1}_{k-1}\mathcal{V}^{(ji)T}_{k,k-1} - \mathcal{E}^{(ij)}_{k-1}\Pi^{(j)-1}_{k-1}\mathcal{V}^{(j)T}_{k,k-1}, \quad k\geq 2;\\ \mathcal{E}^{(ij)}_{0,1} = 0.\end{array} \tag{32}$$

The matrices $\mathcal{V}^{(ij)}_{k,k-1} = E\big[V^{(i)}_k\mu^{(j)T}_{k-1}\big]$ are calculated by

$$\mathcal{V}^{(ij)}_{k,k-1} = \overline{\alpha}^{(i)}_k\Sigma^{zy^{(ij)}}_{k,k-1} + \overline{\lambda}^{(i)}_k\Sigma^{zy^{(ij)}}_{k-1} - \mathbf{H}^{(i)}_{\mathbb{A}_k}(\Delta^{(j)}_{k-2,k-1} + \mathcal{E}^{(j)}_{k-1}), \quad k\geq 2, \tag{33}$$

with

$$\Sigma^{zy^{(ij)}}_{k,s} = \overline{\alpha}^{(j)}_s\Sigma^{z^{(ij)}}_{k,s} + \overline{\lambda}^{(j)}_s\Sigma^{z^{(ij)}}_{k,s-1} + \overline{\rho}^{(j)}_s\overline{H}^{(i)}_k\mathbb{A}_k\Sigma^{e^{(j)}}_{s-1}\mathbb{A}^T_s\overline{H}^{(j)T}_s, \quad 2\leq s\leq k. \tag{34}$$

The innovation cross-covariance matrices, $\Pi^{(ij)}_k = E\big[\mu^{(i)}_k\mu^{(j)}_k\big]$, satisfy

$$\begin{array}{rl}\Pi^{(ij)}_k = & \overline{\alpha}^{(i)}_k\Sigma^{zy^{(ij)}}_k + \overline{\lambda}^{(i)}_k\Sigma^{zy^{(ij)}}_{k,k-1} - \mathbf{H}^{(j)}_{\mathbb{A}_k}\Big(\Delta^{(j)}_{k-1,k} + \mathcal{E}^{(ij)}_{k-1,k}\Big)\\ & -\mathcal{V}^{(ij)}_{k,k-1}\Pi^{(j)-1}_{k-1}\mathcal{Y}^{(j)T}_{k,k-1} - \mathcal{V}^{(i)}_{k,k-1}\Pi^{(i)-1}_{k-1}\Pi^{(ij)}_{k-1,k}, \quad k\geq 2;\\ \Pi^{(ij)}_1 = & \Sigma^{z^{(ij)}}_1,\end{array} \tag{35}$$

in which

$$\Pi_{k-1,k}^{(ij)} = \left(\mathcal{E}_{k-1}^{(i)} - \mathcal{E}_{k-1}^{(ji)}\right)^T \mathbf{H}_{\mathbb{A}_k}^{(j)T} + \mathcal{V}_{k,k-1}^{(ji)T} - \Pi_{k-1}^{(ij)}\Pi_{k-1}^{(j)-1}\mathcal{V}_{k,k-1}^{(j)T}, \quad k \geq 2. \tag{36}$$

Proof. See Appendix C. □

4.2. Distributed fusion filtering algorithm

Using (19) and the matrices obtained in the previous lemma, an expression for the cross-covariance matrices between any two local estimators is easily obtained and the distributed filter, $\widehat{x}_{k/k}^D$, is calculated from (29). Moreover, the factorisation of the system state covariance (Remark 2) together with (29) allow us to derive a formula for the filtering error covariance matrices $\Sigma_{k/k}^{\widetilde{x}^D} = E\big[(x_k - \widehat{x}_{k/k}^D)(x_k - \widehat{x}_{k/k}^D)^T\big]$. These results are presented in the following theorem.

Theorem 4.2. *Let* $\widehat{X}_{k/k} = \left(\widehat{x}_{k/k}^{(1)T}, \ \ldots \ , \widehat{x}_{k/k}^{(m)T}\right)^T$ *be the vector constituted by the local filtering estimators calculated from the algorithm in Theorem 3.2. The distributed filtering estimator is given by*

$$\widehat{x}_{k/k}^D = \boldsymbol{\Xi}_{k/k}\left(\boldsymbol{\Sigma}_{k/k}\right)^{-1}\widehat{X}_{k/k}, \quad k \geq 1,$$

with $\boldsymbol{\Sigma}_{k/k} = \left(\Sigma_{k/k}^{\widehat{x}^{(ij)}}\right)_{i,j=1,\ldots,m}$ *and* $\boldsymbol{\Xi}_{k/k} = \left(\Sigma_{k/k}^{\widehat{x}^{(11)}}\big|\cdots\big|\Sigma_{k/k}^{\widehat{x}^{(mm)}}\right)$, *where the matrices* $\Sigma_{k/k}^{\widehat{x}^{(ij)}} = E\big[\widehat{x}_{k/k}^{(i)}\widehat{x}_{k/k}^{(j)T}\big]$ *are obtained by*

$$\Sigma_{k/k}^{\widehat{x}^{(ij)}} = \mathbb{A}_k\Sigma_k^{e^{(ij)}}\mathbb{A}_k^T, \quad k \geq 1, \quad i,j = 1,\ldots,m,$$

in which $\Sigma_k^{e^{(ij)}}$ *are calculated from the formulas given in Lemma 4.1.*

The error covariance matrices of the distributed filtering estimators are computed by

$$\Sigma_{k/k}^{\widetilde{x}^D} = A_k B_k^T - \boldsymbol{\Xi}_{k/k}\boldsymbol{\Sigma}_{k/k}^{-1}\boldsymbol{\Xi}_{k/k}^T, \quad k \geq 1.$$

5. Simulation study

In this section, a numerical simulation example is presented with a dual purpose: *(a)* illustrating some of the different sensor uncertainties covered by the current state-space model with random measurement matrices; *(b)* showing the applicability and effectiveness of the proposed distributed filter design scheme and how the estimation accuracy is influenced by the sensor uncertainties and the random transmission delays and packet losses.

System state model. Consider a target tracking system whose two-dimensional state process $\{x_k\}_{k\geq 0}$ is described by the following model with state-dependent multiplicative noise (the first and second component of x_k represent the position and velocity of the target at time k, respectively):

$$x_{k+1} = \left(F + \xi_k \breve{F}\right)x_k + Gw_k, \qquad k \geq 0,$$

where

$$F = \begin{pmatrix} 0.95 & 0.01 \\ 0 & 0.95 \end{pmatrix}, \quad \breve{F} = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}, \quad G = \begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix}.$$

The initial state, x_0, is a two-dimensional zero-mean gaussian vector with $\Sigma_0^x = I$, and $\{\xi_k\}_{k\geq 0}$, $\{w_k\}_{k\geq 0}$ are zero-mean gaussian white processes with unit variance. These noise sequences and the initial state are assumed to be mutually independent; then, it is clear that, as indicated in Remark 2, the state covariance function can be expressed in a separable form as $\Sigma_{k,s}^x = A_k B_s^T$, with $A_k = F^k$ and $B_s^T = F^{-s}\Sigma_s^x$, where $\Sigma_s^x = E[x_s x_s^T]$ is recursively obtained by

$$\Sigma_s^x = F\Sigma_{s-1}^x F^T + \breve{F}\Sigma_{s-1}^x \breve{F}^T + GG^T, \quad s \geq 1.$$

Sensor measured outputs. Let us consider three sensors that provide scalar measurements of the state described by the following model:

$$z_k^{(i)} = H_k^{(i)} x_k + v_k^{(i)}, \qquad k \geq 1, \quad i = 1, 2, 3,$$

and, according to the theoretical study, let us suppose that these measurements are affected by different uncertainties; namely:

- At each sensor, $i = 1, 2, 3$, the random parameter sequences $\{H_k^{(i)}\}_{k\geq 1}$, are chosen to account for different types of network-induced uncertainties:
  - $H_k^{(1)} = (0.8,\ 0.9)\gamma_k^{(1)}$, where $\{\gamma_k^{(1)}\}_{k\geq 1}$ is a sequence of independent random variables uniformly distributed over $[0.3, 0.7]$ (*continuous fading measurements in sensor 1*).
  - $H_k^{(2)} = (0.6,\ 0.7)\gamma_k^{(2)}$, where $\{\gamma_k^{(2)}\}_{k\geq 1}$ is a sequence of independent random variables with $P(\gamma_k^{(2)} = 0) = 0.1$, $P(\gamma_k^{(2)} = 0.5) = 0.5$, $P(\gamma_k^{(2)} = 1) = 0.4$ (*discrete fading measurements in sensor 2*).
  - $H_k^{(3)} = \gamma_k^{(3)}\left((0.9,\ 0.5) + (0,\ 0.95)\varphi_k\right)$, where $\{\gamma_k^{(3)}\}_{k\geq 1}$ is a sequence of independent Bernoulli random variables with $P(\gamma_k^{(3)} = 1) = \overline{\gamma}^{(3)}$, $\forall k \geq 1$, and $\{\varphi_k\}_{k\geq 1}$ is a zero-mean Gaussian white process with unit variance (*missing measurements and multiplicative noise in sensor 3*).

  The sequences $\left\{\gamma_k^{(i)}\right\}_{k\geq 1}$, $i = 1, 2, 3$, and $\left\{\varphi_k\right\}_{k\geq 1}$ have time-invariant probability distributions and they are mutually independent white sequences.

- The sensor additive noises $\{v_k^{(i)}\}_{k\geq 1}$, $i = 1, 2, 3$, are defined by $v_k^{(i)} = v_i w_{k-1}$, with $v_1 = v_2 = 50$ and $v_3 = 25$. Clearly, these noises are correlated to each other, with $R_k^{(ij)} = v_i v_j$, and also correlated with the process noise, with $S_k^{(i)} = Gv_i$. Hence, as indicated in *Remark 2*, the state process and the sensor noises are correlated, with $C_k^{(i)T} = F^{-k}Gv_i$.

Observations with random one-step delays and packet dropouts. In accordance with the theoretical study, let us suppose that random one-step delays and packet dropouts with different rates exist in the data transmissions from the individual sensors to the local processors. Specifically, we assume that the available measurements used for the estimation, $y_k^{(i)}$, $i = 1, 2, 3$, are modelled as follows:

$$
\begin{aligned}
y_k^{(1)} &= \theta_k^{(1)} \beta_k^{(1)} z_k^{(1)} + \theta_k^{(1)} (1 - \beta_k^{(1)}) z_{k-1}^{(1)} + (1 - \theta_k^{(1)}) \widehat{z}_{k/k-1}^{(1)}, \quad k \geq 2; \quad y_1^{(1)} = z_1^{(1)}, \\
y_k^{(2)} &= \theta_k^{(2)} z_k^{(2)} + (1 - \theta_k^{(2)}) \widehat{z}_{k/k-1}^{(2)}, \quad k \geq 2; \quad y_1^{(2)} = z_1^{(2)}, \\
y_k^{(3)} &= \beta_k^{(3)} z_k^{(3)} + (1 - \beta_k^{(3)}) z_{k-1}^{(3)}, \quad k \geq 2; \quad y_1^{(3)} = z_1^{(3)},
\end{aligned}
$$

where $\{\theta_k^{(i)}\}_{k \geq 2}, \ i = 1, 2$, and $\{\beta_k^{(i)}\}_{k \geq 2}, \ i = 1, 3$, are independent white sequences of Bernoulli random variables with time-invariant probabilities $P(\theta_k^{(i)} = 1) = \overline{\theta}^{(i)}$ and $P(\beta_k^{(i)} = 1) = \overline{\beta}^{(i)}$.

Let us observe that this model considers the possibility of delays and packet dropouts in transmissions from sensor 1, while the measurements received by the local processors of sensors 2 and 3 are only subject to random dropouts and delays, respectively. At sensor 1, the packet on time receiving rate is $\overline{\alpha}^{(1)} = \overline{\theta}^{(1)} \overline{\beta}^{(1)}$, the one-step delay receiving rate is $\overline{\lambda}^{(1)} = \overline{\theta}^{(1)} (1 - \overline{\beta}^{(1)})$ and the packet dropout rate (compensation rate) is $\overline{\rho}^{(1)} = 1 - \overline{\theta}^{(1)}$. At sensor 2, the packet dropout rate is $\overline{\rho}^{(2)} = 1 - \overline{\theta}^{(2)}$ and, at sensor 3, the one-step delay receiving rate is $\overline{\lambda}^{(3)} = 1 - \overline{\beta}^{(3)}$.

If $\theta_k^{(i)} = 1, \ i = 1, 2$, all the transmissions at time k are successful; if also $\beta_k^{(i)} = 1, \ i = 1, 3$, the current measured outputs are all received on time, i.e. $y_k^{(i)} = z_k^{(i)}, \ i = 1, 2, 3$, while if $\beta_k^{(i)} = 0$ for some i, the corresponding observation is one-step delayed, $y_k^{(i)} = z_{k-1}^{(i)}$. At the sensors 1 and 2, $\theta_k^{(i)} = 0$ means that the kth transmission fails, no packet is then received at time k, and the one-stage observation predictor, $\widehat{z}_{k/k-1}^{(i)}$, is used instead.

Finally, in order to apply the proposed distributed fusion filtering algorithm, it is assumed that all the processes involved in the observation equations satisfy the independence hypotheses imposed on the theoretical model.

In order to illustrate the feasibility and effectiveness of the proposed algorithm, a MATLAB program has been created to obtain the local and distributed fusion estimators, as well as the corresponding error variances, and fifty iterations of the proposed local and distributed filtering algorithms have been run. The estimation accuracy has been assessed by analyzing the error variances for different probabilities $\overline{\gamma}^{(3)}$ of the Bernoulli random variables modelling the missing measurement phenomenon of the third sensor, and several values of the probabilities $\overline{\theta}^{(i)}, \ i = 1, 2$, and $\overline{\beta}^{(i)}, \ i = 1, 3$, that determine the probabilities of one-step delays and packet dropouts in the transmissions.

Performance of local and distributed fusion filtering estimators. Let us assume the same value, 0.5, for the probabilities $\overline{\gamma}^{(3)}, \ \overline{\theta}^{(i)}, \ i = 1, 2$, and $\overline{\beta}^{(i)}, \ i = 1, 3$. First, the error variance analysis is used to compare the proposed local and distributed fusion filtering estimators with the local and distributed fusion filters in (Caballero-Águila et al., 2016a), where the process noise is uncorrelated with the sensor noises and the

latest measurement previously received is used as loss compensation.
Since the measurements received by the local processor 3 are not influenced by the packet loss phenomenon, only the local filters of processors 1 and 2 are considered. The results of this comparison are displayed in Figure 1. On the one hand, this figure shows that the error variances of the proposed distributed fusion filtering estimators are significantly smaller than those of every local estimator; hence, as expected, the distributed fusion filtering estimators outperform all the local ones. On the other hand, it is observed that the proposed filters outperform the filters in (Caballero-Águila et al., 2016a), due to the fact that they take into account the cross-correlation of additive noises and, moreover, they use all the previous measurements successfully received (and not only the last one) as compensation, as it has been already pointed out. Analogous results are obtained for other values of the probabilities $\overline{\gamma}^{(3)}$, $\overline{\theta}^{(i)}$, $i = 1, 2$, and $\overline{\beta}^{(i)}$, $i = 1, 3$.

Influence of the missing measurement phenomenon on the estimation accuracy (sensor 3). Assuming, as in Figure 1, that the probabilities $\overline{\theta}^{(i)} = 0.5$, $i = 1, 2$, and $\overline{\beta}^{(i)} = 0.5$, $i = 1, 3$, the effect of the missing measurement phenomenon in sensor 3 is studied by analyzing how the distributed filtering error variances are influenced by the probability $\overline{\gamma}^{(3)}$, that the state is present in the measured outputs of sensor 3. For the first state component, Figure 2 (a) shows the distributed filtering error variances for the values $\overline{\gamma}^{(3)} = 0.1$ to 0.9; from this figure, it is observed that the performance of the distributed fusion filtering estimators is indeed influenced by this probability and, as expected, the distributed error variances become smaller as $\overline{\gamma}^{(3)}$ increases, which means that the performance of the distributed filter improves when the probability of missing measurements, $1 - \overline{\gamma}^{(3)}$, decreases. Taking into account that, from $k = 60$ onwards, the behaviour of the error variances is analogous in all the iterations, for a better visualization of this decreasing trend, Figure 2 (b) displays the distributed error variances only at iteration $k = 100$, versus $\overline{\gamma}^{(3)}$. Similar results are obtained for the second component of the state vector, as shown in Figure 3.

Influence of the transmission packet dropout and delay phenomena on the estimation accuracy (sensors 2 and 3). For $\overline{\gamma}^{(3)} = 0.5$, the distributed filtering accuracy is analyzed versus the probabilities $\overline{\theta}^{(2)}$ and $\overline{\beta}^{(3)}$ of the Bernoulli variables modelling the packet loss and delay phenomena in the transmissions from sensors 2 and 3, respectively. Specifically, the influence on the distributed fusion filter performance is analyzed for $\overline{\theta}^{(2)}$ varying from 0.1 to 0.9 and for different fixed values of $\overline{\beta}^{(3)}$. Since, as already indicated, from $k = 60$ onwards, the behavior of the error variances is analogous in all the iterations, only the results of a specific iteration ($k = 100$) are shown here. Figure 4 displays, for the first and second state components, the distributed filtering error variances versus $\overline{\theta}^{(2)}$, for $\overline{\beta}^{(3)} = 0.5$ to 0.9 (for values less than 0.5, the error variance differences are negligible). From this figure, it is concluded that, as $\overline{\theta}^{(2)}$ or $\overline{\beta}^{(3)}$ increase, the distributed filtering error variances become smaller, which means that, as expected, the smaller the packet dropout probabilities, $1 - \overline{\theta}^{(2)}$, and/or the delay probabilities, $1 - \overline{\beta}^{(3)}$, are, the better the estimations are. The error variances have also been compared assuming other values for the probability $\overline{\gamma}^{(3)}$ and similar results have been obtained in all cases.

6. Conclusions

A distributed fusion filtering algorithm has been designed over discrete-time linear systems with stochastic disturbances, including multiplicative state noise and random measurement matrices, under the assumption that the additive noises of the measurement equations at the different sensors are cross-correlated at the same time step and correlated with the system noise at the previous time step. Due to the unreliable network characteristics, the information transmission through the network communication channels is assumed to be subject to random delays and packet dropouts. When a sensor measurement is lost, its prediction estimate is used as a compensator to update the filter. A simulation example has shown that the proposed distributed filter outperforms the local ones and, also, the filter with hold-input compensation strategy, a fact that was expected, since the current filter uses all historical data for compensation and not only the last measurement received.

An interesting future direction would be to complement the current study with a detailed analysis of the steady-state property of the proposed filtering algorithm. In addition, related topics for further research work include the extension of our main results to more complex systems, such as those involving stochastic nonlinearities or sequentially time-correlated additive noises. Another considerable concern at the forefront of research in networked systems subject to random delays and packet dropouts is the number of packets that are processed to update the estimator at each moment. In the future, we will try to investigate the state estimation problem by considering the possibility that, at each sampling time k, the current measurement z_k and the delayed measurement z_{k-1} can arrive simultaneously and both data packets can be processed. Finally, we highlight also the interest in extending the proposed framework to deal with more complicated problems, such as simultaneous state and parameter uncertainty estimation (Geng et al., 2020).

Disclosure statement

The authors report there are no competing interests to declare.

Funding

Grant PID2021-124486NB-I00 funded by MICIU/AEI/ 10.13039/501100011033 and ERDF/EU.

Notes on contributors

R. Caballero-Águila received her MSc degree in Mathematics and her PhD degree in Polynomial Filtering in Systems with Uncertain Observations, both from the University of Granada (Spain), in 1997 and 1999, respectively. In 1997, she joined the University of Jaén (Spain), where she has been a Professor with the Department of Statistics and Operations Research since 2021. Her current research interests include stochastic systems, filtering, prediction and smoothing.

J. Linares-Pérez received her MSc degree in Mathematics and her PhD in Stochastic Differential Equations, both from the University of Granada (Spain) in 1980 and 1982, respectively. She is currently a Professor with the Department of Statistics and Operations Research, University of Granada (Spain). Her research interests are related with stochastic calculus and estimation in stochastic systems.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

Caballero-Águila, R., Hermoso-Carazo, A., & Linares-Pérez, J. (2015). Optimal state estimation for networked systems with random parameter matrices, correlated noises and delayed measurements. *International Journal of General Systems*, *44 (2)*, 142–154.

Caballero-Águila, R., Hermoso-Carazo, A., & Linares-Pérez, J. (2016a). Fusion estimation using measured outputs with random parameter matrices subject to random delays and packet dropouts. *Signal Processing*, *127*, 12–23.

Caballero-Águila, R., Hermoso-Carazo, A., & Linares-Pérez, J. (2016b) Networked fusion filtering from outputs with stochastic uncertainties and correlated random transmission delays. *Sensors*, *16 (6)*, Article ID 847.

Caballero-Águila, R., Hermoso-Carazo, A., & Linares-Pérez, J. (2017a). Distributed fusion filters from uncertain measured outputs in sensor networks with random packet losses. *Information Fusion*, *34*, 70–79.

Caballero-Águila, R., Hermoso-Carazo, A., & Linares-Pérez J. (2017b). Fusion Estimation from Multisensor Observations with Multiplicative Noises and Correlated Random Delays in Transmission. *Mathematics*, *5*, Article ID 45.

Caballero-Águila, R., Hermoso-Carazo, A., & Linares-Pérez, J. (2017c). New distributed fusion filtering algorithm based on covariances over sensor networks with random packet dropouts. *International Journal of Systems Science*, *48 (9)*, 1805–1817.

Caballero-Águila, R., Hermoso-Carazo, A., & Linares-Pérez, J. (2019a). Centralized, distributed and sequential fusion estimation from uncertain outputs with correlation between sensor noises and signal. *International Journal of General Systems*, *48*, 713–737.

Caballero-Águila, R.; Hermoso-Carazo, A., & Linares-Pérez, J. (2019b). Centralized filtering and smoothing algorithms from outputs with random parameter matrices transmitted through uncertain communication channels. *Digital Signal Processing*, *85*, 77–85.

Caballero-Águila, R., Hermoso-Carazo, A., Linares-Pérez, J., & Wang, Z. (2019c). A new approach to distributed fusion filtering for networked systems with random parameter matrices and correlated noises. *Information Fusion*, *45*, 324–332.

Chen, Y., Wang, Z., Qian, W., & Alsaadi, F.E. (2017). Finite-horizon H_∞ filtering for switched time-varying stochastic systems with random sensor nonlinearities and packet dropouts. *Signal Processing*, *138*, 138–145.

Cheng, G., Ma, M., Tan, L., & Song, S. (2021). Gaussian estimation for non-linear stochastic uncertain systems with time-correlated additive noises and packet dropout compensations. *IET Control Theory & Applications*, *16*, 600–614.

Ding, J., Sun, S., Ma, J., & Li, N. (2019). Fusion estimation for multi-sensor networked systems with packet loss compensation. *Information Fusion* **2019**, *45*, 138–149.

Feng, J. (2014) Recursive Estimation for Dynamical Systems with Different Delay Rates Sensor Network and Autocorrelated Process Noises. *Mathematical Problems in Engineering*, Article ID 404907.

García-Ligero, M. J., Hermoso-Carazo, A., & Linares-Pérez, J. (2020). Least-squares estimators for systems with stochastic sensor gain degradation, correlated measurement noises and delays in transmission modelled by Markov chains *International Journal of Systems Science*, *51 (4)*, 731–745.

Geng, H., Haile, M. A., & Fang, H. (2020). SSUE: Simultaneous state and uncertainty estimation for dynamical systems. *International Journal of Robust and Nonlinear Control*, *31*, 1068–1083.

Geng, H., Wang, Z., Chen, Y., Alsaadi, F. E., & Cheng, Y. (2021). Multi-Sensor Filtering Fusion With Parametric Uncertainties and Measurement Censoring: Monotonicity and Boundedness. *IEEE Transactions on Signal Processing*, *69*, 5875–5890.

Guo, Y. (2017). Switched filtering for networked systems with multiple packet dropouts. *Journal of the Franklin Institute*, *354*, 3134–3151.

Han, F., Dong, H., Wang, Z., Li, G., & Alsaadi, F.E. (2018). Improved Tobit Kalman filtering for systems with random parameters via conditional expectation. *Signal Processing*, *147*, 35–45.

Hu, J., Wang, Z., Chen, D., & Alsaadi, F. E. (2016). Estimation, filtering and fusion for networked systems with network-induced phenomena: New progress and prospects. *Information Fusion*, *31*, 65–75.

Hu, J., Wang, Z., Dong, H., & Gao, H. (2013a). Recent Advances on Recursive Filtering and Sliding Mode Design for Networked Nonlinear Stochastic Systems: A Survey. *Mathematical Problems in Engineering*, Article ID 646059.

Hu, J., Wang, Z., & Gao, H. (2013b). Recursive filtering with random parameter matrices, multiple fading measurements and correlated noises, *Automatica*, *49*, 3440–3448.

Hu, J., Wang, Z., Shen, B., & Gao, H. (2013c). Gain-Constrained Recursive Filtering With Stochastic Nonlinearities and Probabilistic Sensor Delays. *IEEE Transactions on Signal Processing*, *61 (5)*, 1230–1238.

Li, Y., Liu, S., Zhong, M., & Ding, S. (2018). State estimation for stochastic discrete-time systems with multiplicative noises and unknown inputs over fading channels. *Applied Mathematics and Computation*, *320*, 116–130.

Lin, H., & Sun, S. (2018). Optimal Sequential Fusion Estimation With Stochastic Parameter Perturbations, Fading Measurements, and Correlated Noises. *IEEE Transactions on Signal Processing*, *66*, 3571–3583.

Lin, H., & Sun, S.(2019). Globally optimal sequential and distributed fusion state estimation for multi-sensor systems with cross-correlated noises. *Automatica*, *101*, 128–137.

Linares-Pérez, J., Caballero-Águila, R., & García-Garrido, I. (2014). Optimal linear filter design for systems with correlation in the measurement matrices and noises: recursive algorithm and applications. *International Journal of Systems Science*, *45 (7)*, 1548–1562.

Liu, W., Xie, X., Qian, W., Xu, X., & Shi, Y. (2020). Optimal Linear Filtering for Networked Control Systems With Random Matrices, Correlated Noises, and Packet Dropouts. *IEEE Access*, *8*, 59987–59997.

Ma, J., & Sun, S. (2019). A general packet dropout compensation framework for optimal prior filter of networked multi-sensor systems. *Information Fusion*, *45*, 128–137.

Ma, J., & Sun, S. (2020a). Optimal Linear Filter for Systems With Random Delay and Packet Dropout Compensations. *IEEE Access*, *8*, 145268–145277.

Ma, J., & Sun, S. (2020b). Optimal linear recursive estimators for stochastic uncertain systems with time-correlated additive noises and packet dropout compensations. *Signal Processing*, *176*, 107704.

Sun, S. (2020). Distributed Optimal Linear Fusion Predictors and Filters for Systems With Random Parameter Matrices and Correlated Noises. *IEEE Transactions on Signal Processing*, *68*, 1064–1074.

Sun, S., Tian, T., & Honglei, L. (2017). State estimators for systems with random parameter matrices, stochastic nonlinearities, fading measurements and correlated noises. *Information Sciences*, *397*, 118–136.

Sun, S., Tian, T., & Lin, H. (2016). Optimal Linear Estimators for Systems With Finite-Step Correlated Noises and Packet Dropout Compensations. *IEEE Transactions on Signal Processing*, *64 (21)*, 5672–5681.

Tan, L., Xu, C., Wang, Y., Wei, H., Zhao, K., & Song, S. (2020). Gaussian Recursive Filter for Nonlinear Systems with Finite-step Correlated Noises and Packet Dropout Compensations. *Measurement Science Review*, *20 (2)*, 80–92.

Tian, T., Sun, S., & Li, N. (2016). Multi-sensor information fusion estimators for stochastic uncertain systems with correlated noises. *Information Fusion*, *27*, 126–137.

Uribe-Murcia, K., Shmaliy, Y. S., & Andrade-Lucio, J. A. (2021). Unbiased FIR, Kalman, and game theory H_∞ filtering under Bernoulli distributed random delays and packet dropouts. *Neurocomputing*, *442*, 89–97.

Wang, S., Fang, H., & Tian, X. (2016). Minimum variance estimation for linear uncertain systems with one-step correlated noises and incomplete measurements. *Digital Signal Processing*, *49*, 126–136.

Wang, X., & Sun, S. (2019). Optimal recursive estimation for networked stochastic uncertain systems with fading measurements and time-correlated channel noises. *Journal of Computational and Applied Mathematics*, *346*, 549–565.

Wang, W., & Zhou, J. (2017). Optimal linear filtering design for discrete-time systems with cross-correlated stochastic parameter matrices and noises. *IET Control Theory* & *Applications*, *11 (18)*, 3353–3362.

Yang, Y., Liang, Y., Pan, Q., Qin, Y., & Yang, F. (2016). Distributed fusion estimation with square-root array implementation for Markovian jump linear systems with random parameter matrices and cross-correlated noises. *Information Sciences*, *370*, 446–462.

Zhao, X., She, S., Li, Y., Zhao, K., & Min, S. (2020). State estimation for the system with unknown sensor inputs, asynchronous correlated noises and packet loss compensation. *Proceedings of the 39th Chinese Control Conference*, 2946–2950.

Zhao, K., Tan, L., & Song, S. (2019). Fusion estimation for nonlinear multi-sensor networked systems with packet loss compensation and correlated noises. *Sensor Review* **2019**, *39 (5)*, 682–696.

Appendix A. Proof of Lemma 3.1

For $i = 1, \dots, m$, substituting $\widehat{z}^{(i)}_{k/k-1} = \overline{H}^{(i)}_k \mathbb{A}_k e^{(i)}_{k-1}$ in (2), we have

$$\begin{aligned} y^{(i)}_k &= \alpha^{(i)}_k z^{(i)}_k + \lambda^{(i)}_k z^{(i)}_{k-1} + \rho^{(i)}_k \overline{H}^{(i)}_k \mathbb{A}_k e^{(i)}_{k-1}, \quad k \geq 2; \\ y^{(i)}_1 &= z^{(i)}_1. \end{aligned} \tag{A1}$$

On the one hand, since $\alpha^{(i)}_k$, $\lambda^{(i)}_k$ and $\rho^{(i)}_k$, $k \geq 2$, are Bernoulli random variables such that $\alpha^{(i)}_k + \lambda^{(i)}_k + \rho^{(i)}_k = 1$, it this clear that $\alpha^{(i)}_k \lambda^{(i)}_k = \alpha^{(i)}_k \rho^{(i)}_k = \lambda^{(i)}_k \rho^{(i)}_k = 0$, and also that $E\big[(\alpha^{(i)}_k)^2\big] = \overline{\alpha}^{(i)}_k$, $E\big[(\lambda^{(i)}_k)^2\big] = \overline{\lambda}^{(i)}_k$ and $E\big[(\rho^{(i)}_k)^2\big] = \overline{\rho}^{(i)}_k$.

On the other hand, according to the model assumptions, it is easy to derive that:

– The variable $\alpha^{(i)}_k$ is independent of the vectors $z^{(i)}_k$ and $z^{(i)}_k y^{(i)}_{k-1}$.
– The variable $\lambda^{(i)}_k$ is independent of the vectors $z^{(i)}_{k-1}$ and $z^{(i)}_{k-1} y^{(i)}_{k-1}$.
– The variable $\rho^{(i)}_k$ is independent of the vectors $e^{(i)}_k$ and $e^{(i)}_k y^{(i)}_{k-1}$.

Using these properties together with expression (A1) for $y^{(i)}_k$ in (16) and (17), and for $y^{(i)}_s$ in (18), Lemma 3.1 can be easily deduced.

Appendix B. Proof of Theorem 3.2

- *Derivation of expressions (19)-(23):*
  - From (14), expression (19) for the filter $\widehat{x}^{(i)}_{k/k}$ is clear.
  - To obtain (20), we use the OPL to write the filtering error covariances as $\Sigma^{\widetilde{x}^{(i)}}_{k/k} = \Sigma^x_k - E\big[\widehat{x}^{(i)}_{k/k}\widehat{x}^{(i)T}_{k/k}\big]$; so, taking into account that, according to the factorisation of the state covariance function set out in Remark 2, $\Sigma^x_k = A_k B^T_k$ and, from (19), $E\big[\widehat{x}^{(i)}_{k/k}\widehat{x}^{(i)T}_{k/k}\big] = \mathbb{A}_k\Sigma^{e^{(i)}}_k\mathbb{A}^T_k$, expression (20) is proven.
  - Expressions (21) and (22) for the vectors $e^{(i)}_k$ and their covariances $\Sigma^{e^{(i)}}_k$, respectively, have been established in Remark 4.
  - Expression (23) for $\mathcal{E}^{(i)}_k$ is deduced from (13) taking into account that, from (22) $\Sigma^{e^{(i)}}_{k-1} = \displaystyle\sum_{h=1}^{k-1}\mathcal{E}^{(i)}_h\Pi^{(i)-1}_h\mathcal{E}^{(i)T}_h$.
- *Derivation of expressions (24) and (25):*
  - From the alternative expression (15) for the observation predictor, expression (24) for the innovation is immediately deduced.
  - Applying the OPL, the innovation covariance matrix is expressed as $\Pi^{(i)}_k = \Sigma^{y^{(i)}}_k - E\big[\widehat{y}^{(i)}_{k/k-1}y^{(i)T}_k\big]$, from which, using again (15) for $\widehat{y}^{(i)}_{k/k-1}$, we have

$$\Pi^{(i)}_k = \Sigma^{y^{(i)}}_k - \mathbf{H}^{*(i)}_{\mathbb{A}_k}E\big[e^{(i)}_{k-1}y^{(i)T}_k\big] + \mathcal{V}^{(i)}_{k,k-1}\Pi^{(i)-1}_{k-1}E\big[\mu^{(i)}_{k-1}y^{(i)T}_k\big], \quad k \geq 2,$$

    and, since $E\big[e^{(i)}_{k-1}y^{(i)T}_k\big] = \Delta^{(i)}_{k-1,k}$ and $E\big[\mu^{(i)}_{k-1}y^{(i)T}_k\big] = \mathcal{Y}^{(i)T}_{k,k-1}$, expression (25) is proven.
- *Derivation of expressions (26)-(28):*
  - As already indicated, $V^{(i)}_k$, defined by (9), is uncorrelated with $\mu^{(i)}_1, \dots, \mu^{(i)}_{k-2}$ and, hence $\mathcal{V}^{(i)}_{k,k-1} = E\Big[V^{(i)}_k\mu^{(i)T}_{k-1}\Big] = E\Big[V^{(i)}_k y^{(i)T}_{k-1}\Big]$. From (7), we write $V^{(i)}_k = y^{(i)}_k - X^{(i)}_k$ with $X^{(i)}_k$ given in (8), then the model assumptions guarantee that

$$\begin{aligned}\mathcal{V}^{(i)}_{k,k-1} = \; & \Sigma^{y^{(i)}}_{k,k-1} - \overline{\alpha}^{(i)}_k\overline{H}^{(i)}_k E\big[x_k y^{(i)T}_{k-1}\big] - \overline{\lambda}^{(i)}_k\overline{H}^{(i)}_{k-1}E\big[x_{k-1}y^{(i)T}_{k-1}\big] \\ & -\overline{\rho}^{(i)}_k\overline{H}^{(i)}_k E\big[\widehat{x}^{(i)}_{k/k-1}y^{(i)T}_{k-1}\big].\end{aligned}$$

    Next, using the OPL and (19), we have that

$$E\big[x_s y^{(i)T}_{k-1}\big] = E\big[\widehat{x}^{(i)}_{s/k-1}y^{(i)T}_{k-1}\big] = \mathbb{A}_s E\big[e^{(i)}_{k-1}y^{(i)T}_{k-1}\big] = \mathbb{A}_s\Sigma^{ey^{(i)}}_{k-1}, \quad s = k, k-1,$$

    and, hence, we obtain

$$\begin{aligned}\mathcal{V}^{(i)}_{k,k-1} & = \Sigma^{y^{(i)}}_{k,k-1} - \overline{\alpha}^{(i)}_k\overline{H}^{(i)}_k\mathbb{A}_k\Sigma^{ey^{(i)}}_{k-1} - \overline{\lambda}^{(i)}_k\overline{H}^{(i)}_{k-1}\mathbb{A}_{k-1}\Sigma^{ey^{(i)}}_{k-1} - \overline{\rho}^{(i)}_k\overline{H}^{(i)}_k\mathbb{A}_k\Sigma^{ey^{(i)}}_{k-1} \\ & = \Sigma^{y^{(i)}}_{k,k-1} - \mathbf{H}^{(j)}_{\mathbb{A}_k}\Sigma^{ey^{(i)}}_{k-1} - \overline{\rho}^{(i)}_k\overline{H}^{(i)}_k\mathbb{A}_k\Sigma^{ey^{(i)}}_{k-1}, \quad k \geq 2,\end{aligned}$$

which, using expression (17) for $\Sigma^{y^{(i)}}_{k,k-1}$, leads us to

$$\mathcal{V}^{(i)}_{k,k-1} = \overline{\alpha}^{(i)}_k \Sigma^{zy^{(i)}}_{k,k-1} + \overline{\lambda}^{(i)}_k \Sigma^{zy^{(i)}}_{k-1,k-1} - \mathbf{H}^{(i)T}_{\mathbb{A}_k} \Sigma^{ey^{(i)}}_{k-1}, \quad k \geq 2.$$

Finally, to calculate $\Sigma^{ey^{(i)}}_{k-1} = E\big[e^{(i)}_{k-1} y^{(i)T}_{k-1}\big]$, we write $y^{(i)}_{k-1} = \mu^{(i)}_{k-1} + \widehat{y}^{(i)}_{k-1/k-2}$, then

$$\Sigma^{ey^{(i)}}_{k-1} = E\big[e^{(i)}_{k-1}\mu^{(i)T}_{k-1}\big] + E\big[e^{(i)}_{k-1}\widehat{y}^{(i)T}_{k-1/k-2}\big] = \mathcal{E}^{(i)}_{k-1} + E\big[e^{(i)}_{k-1}\widehat{y}^{(i)T}_{k-1/k-2}\big], \quad k \geq 2,$$

and, using (21) for $e^{(i)}_{k-1}$, the orthogonality of $\mu^{(i)}_{k-1}$ with $\widehat{y}^{(i)}_{k-1/k-2}$, and the OPL, we have that

$$E\big[e^{(i)}_{k-1}\widehat{y}^{(i)T}_{k-1/k-2}\big] = E\big[e^{(i)}_{k-2}\widehat{y}^{(i)T}_{k-1/k-2}\big] = E\big[e^{(i)}_{k-2} y^{(i)T}_{k-1}\big] = \Delta^{(i)}_{k-2,k-1},$$

and expression (26) for $\mathcal{V}^{(i)}_{k,k-1}$ is hereby proved just by substitution.

- From the OPL, $\mathcal{Y}^{(i)}_{k,k-1} = E\big[y^{(i)}_k \mu^{(i)T}_{k-1}\big] = E\big[\widehat{y}^{(i)}_{k/k-1}\mu^{(i)T}_{k-1}\big]$; then, using the alternative expression (15) for the observation predictor $\widehat{y}^{(i)}_{k/k-1}$, we have

$$\mathcal{Y}^{(i)}_{k,k-1} = \mathbf{H}^{*(i)}_{\mathbb{A}_k} E\big[e^{(i)}_{k-1}\mu^{(i)T}_{k-1}\big] + \mathcal{V}^{(i)}_{k,k-1}\Pi^{(i)-1}_{k-1} E\big[\mu^{(i)}_{k-1}\mu^{(i)T}_{k-1}\big], \quad k \geq 2,$$

and, since $E\big[e^{(i)}_{k-1}\mu^{(i)T}_{k-1}\big] = \mathcal{E}^{(i)}_{k-1}$ and $E\big[\mu^{(i)}_{k-1}\mu^{(i)T}_{k-1}\big] = \Pi^{(i)}_{k-1}$, expression (27) is proven.

- Using the OPL again, $\Delta^{(i)}_{k-1,k} = E\big[e^{(i)}_{k-1} y^{(i)T}_k\big] = E\big[e^{(i)}_{k-1}\widehat{y}^{(i)T}_{k/k-1}\big]$ and, from (15) for the observation predictor $\widehat{y}^{(i)}_{k/k-1}$, expression (28) is easily obtained.

Appendix C. Proof of Lemma 4.1

- *Derivation of expressions (30)-(32):*
  - Using (21) for $e^{(i)}_k$ and taking into account that $\mathcal{E}^{(ij)}_{s,k} = E\big[e^{(i)}_s \mu^{(j)T}_k\big]$, for $s = k, k-1$, we get (30) for $\Sigma^{e^{(ij)}}_k$.
  - Using again (21) for $e^{(i)}_k$, expression (31) for $\mathcal{E}^{(ij)}_k = E\big[e^{(i)}_k \mu^{(j)T}_k\big]$ is directly obtained.
  - To derive (32) for $\mathcal{E}^{(ij)}_{k-1,k} = E\big[e^{(i)}_{k-1}\mu^{(j)T}_k\big] = E\big[e^{(i)}_{k-1} y^{(j)T}_k\big] - E\big[e^{(i)}_{k-1}\widehat{y}^{(j)T}_{k/k-1}\big]$, we proceed to calculate both expectations.

    First expectation $E\big[e^{(i)}_{k-1} y^{(j)T}_k\big]$. From (7), $y^{(j)}_k = X^{(j)}_k + V^{(j)}_k$ with $X^{(j)}_k$ and $V^{(j)}_k$ given in (8) and (9), respectively, then the model assumptions ensure that:

    On the one hand,

$$\begin{aligned} E\big[e^{(i)}_{k-1} X^{(j)T}_k\big] = \; & \overline{\alpha}^{(j)}_k E\big[e^{(i)}_{k-1} x^T_k\big]\overline{H}^{(j)T}_k + \overline{\lambda}^{(j)}_k E\big[e^{(i)}_{k-1} x^T_{k-1}\big]\overline{H}^{(j)T}_{k-1} \\ & + \overline{\rho}^{(j)}_k E\big[e^{(i)}_{k-1}\widehat{x}^{(j)T}_{k/k-1}\big]\overline{H}^{(j)T}_k, \end{aligned}$$

where, applying the OPL and (19),

$$E\big[e^{(i)}_{k-1}x^{T}_{s}\big]=E\big[e^{(i)}_{k-1}\widehat{x}^{(i)T}_{s/k-1}\big]=E\big[e^{(i)}_{k-1}e^{(i)T}_{k-1}\big]\mathbb{A}^{T}_{s}=\Sigma^{e^{(i)}}_{k-1}\mathbb{A}^{T}_{s},\quad s=k,k-1,$$

and, also from (19), $E\big[e^{(i)}_{k-1}\widehat{x}^{(j)T}_{k/k-1}\big]=\Sigma^{e^{(ij)}}_{k-1}\mathbb{A}^{T}_{k}$; then, by substitution, we get

$$E\big[e^{(i)}_{k-1}X^{(j)T}_{k}\big]=\Sigma^{e^{(i)}}_{k-1}\mathbf{H}^{(j)T}_{\mathbb{A}_k}+\overline{\rho}^{(j)}_{k}\Sigma^{e^{(ij)}}_{k-1}\mathbb{A}^{T}_{k}\overline{H}^{(j)T}_{k},\quad k\geq 2. \tag{C1}$$

On the other hand, since $V^{(i)}_{k}$, defined by (9), is uncorrelated with $\mu^{(j)}_{1},\dots,\mu^{(j)}_{k-2}$, it is also uncorrelated with $e^{(j)}_{k-2}$; hence, using (21) for $e^{(i)}_{k-1}$, we have

$$E\big[e^{(i)}_{k-1}V^{(j)T}_{k}\big]=\mathcal{E}^{(i)}_{k-1}\Pi^{(i)-1}_{k-1}\mathcal{V}^{(ji)T}_{k,k-1},\quad k\geq 2. \tag{C2}$$

From (C1) and (C2), it is clear that the first expectation satisfies

$$E\big[e^{(i)}_{k-1}y^{(j)T}_{k}\big]=\Sigma^{e^{(i)}}_{k-1}\mathbf{H}^{(j)T}_{\mathbb{A}_k}+\overline{\rho}^{(j)}_{k}\Sigma^{e^{(ij)}}_{k-1}\mathbb{A}^{T}_{k}\overline{H}^{(j)T}_{k}+\mathcal{E}^{(i)}_{k-1}\Pi^{(i)-1}_{k-1}\mathcal{V}^{(ji)T}_{k,k-1},\quad k\geq 2. \tag{C3}$$

Second expectation $E\big[e^{(i)}_{k-1}\widehat{y}^{(j)T}_{k/k-1}\big]$. From (15) for the observation predictor $\widehat{y}^{(j)T}_{k/k-1}$, the following expression is directly obtained

$$E\big[e^{(i)}_{k-1}\widehat{y}^{(j)T}_{k/k-1}\big]=\Sigma^{e^{(ij)}}_{k-1}\mathbf{H}^{*(j)T}_{\mathbb{A}_k}+\mathcal{E}^{(ij)}_{k-1}\Pi^{(j)-1}_{k-1}\mathcal{V}^{(j)T}_{k,k-1},\quad k\geq 2. \tag{C4}$$

From (C3) and (C4), expression (32) is clear.

- *Derivation of expressions (33) and (34):* By following an analogous reasoning to that used in (26) and (18), expressions (33) and (34), respectively, can be easily deduced and hence the details are omitted for brevity.
- *Derivation of expressions (35) and (36):*
  - To obtain (35), starting from $\Pi^{(ij)}_{k}=E\big[y^{(i)}_{k}\mu^{(j)T}_{k}\big]-E\big[\widehat{y}^{(i)}_{k/k-1}\mu^{(j)T}_{k}\big]$, $k\geq 2$, these two expectations should be calculated.

    (I) Since $E\big[y^{(i)}_{k}\mu^{(j)T}_{k}\big]=\Sigma^{y^{(ij)}}_{k}-E\big[y^{(i)}_{k}\widehat{y}^{(j)T}_{k/k-1}\big]$, $k\geq 2$, firstly, we obtain an expression for $\Sigma^{y^{(ij)}}_{k}$; as in (17), expression (A1) and the model assumptions lead us to

$$\Sigma^{y^{(ij)}}_{k}=\overline{\alpha}^{(i)}_{k}\Sigma^{zy^{(ij)}}_{k}+\overline{\lambda}^{(i)}_{k}\Sigma^{zy^{(ij)}}_{k,k-1}+\overline{\rho}^{(i)}_{k}\overline{H}^{(i)}_{k}\mathbb{A}_{k}E\big[e^{(i)}_{k-1}y^{(j)T}_{k}\big]. \tag{C5}$$

Secondly, we write $y^{(j)}_{k}=X^{(j)}_{k}+V^{(j)}_{k}$, with $X^{(j)}_{k}$ and $V^{(j)}_{k}$ given in (8) and (9), and we reason as in the derivation of (C3) to obtain

$$E\big[y^{(i)}_{k}\widehat{y}^{(j)T}_{k/k-1}\big]=\mathbf{H}^{(i)}_{\mathbb{A}_k}\Delta^{(j)}_{k}+\overline{\rho}^{(i)}_{k}\overline{H}^{(i)}_{k}\mathbb{A}_{k}E\big[e^{(i)}_{k-1}\widehat{y}^{(j)T}_{k/k-1}\big]+\mathcal{V}^{(ij)}_{k,k-1}\Pi^{(j)-1}_{k-1}\mathcal{Y}^{(j)T}_{k,k-1}. \tag{C6}$$

Next, taking into account that $\mathcal{E}^{(ij)}_{k-1,k} = E\big[e^{(i)}_{k-1} y^{(j)T}_{k}\big] - E\big[e^{(i)}_{k-1}\widehat{y}^{(j)T}_{k/k-1}\big]$, from (C5) and (C6), we get

$$\begin{aligned} E\big[y^{(i)}_k \mu^{(j)T}_k\big] = \ & \overline{\alpha}^{(i)}_k \Sigma^{zy^{(ij)}}_k + \overline{\lambda}^{(i)}_k \Sigma^{zy^{(ij)}}_{k,k-1} + \overline{\rho}^{(i)}_k \overline{H}^{(i)}_k \mathbb{A}_k \mathcal{E}^{(ij)}_{k-1,k} - \mathbf{H}^{(i)}_{\mathbb{A}_k} \Delta^{(j)}_k \\ & - \mathcal{V}^{(ij)}_{k,k-1} \Pi^{(j)-1}_{k-1} \mathcal{Y}^{(j)T}_{k,k-1}, \quad k \geq 2. \end{aligned} \tag{C7}$$

(II) From (15) for $\widehat{y}^{(j)T}_{k/k-1}$ with $\mathbf{H}^{*(i)}_{\mathbb{A}_k} = \mathbf{H}^{(i)}_{\mathbb{A}_k} + \overline{\rho}^{(i)}_k \overline{H}^{(i)}_k \mathbb{A}_k$, we have

$$E\big[\widehat{y}^{(i)}_{k/k-1} \mu^{(j)T}_k\big] = \Big(\mathbf{H}^{(i)}_{\mathbb{A}_k} + \overline{\rho}^{(i)}_k \overline{H}^{(i)}_k \mathbb{A}_k\Big) \mathcal{E}^{(ij)}_{k-1,k} + \mathcal{V}^{(i)}_{k,k-1} \Pi^{(i)-1}_{k-1} \Pi^{(ij)}_{k-1,k}, \quad k \geq 2. \tag{C8}$$

Expression (35) for $\Pi^{(ij)}_k$ is immediately clear from (C7) and (C8).

- Finally, (36) is proven just expressing $\Pi^{(ij)}_{k-1,k} = E\big[\mu^{(i)}_{k-1} y^{(j)T}_k\big] - E\big[\mu^{(i)}_{k-1} \widehat{y}^{(j)T}_{k/k-1}\big]$ and adopting a deductive reasoning similar to that used to obtain (C3) and (C4), we derive

$$\begin{aligned} E\big[\mu^{(i)}_{k-1} y^{(j)T}_k\big] &= \mathcal{E}^{(i)T}_{k-1} \mathbf{H}^{(j)T}_{\mathbb{A}_k} + \overline{\rho}^{(j)}_k \mathcal{E}^{(ji)T}_{k-1} \mathbb{A}^T_k \overline{H}^{(j)T}_k + \mathcal{V}^{(ji)T}_{k,k-1}, \\ E\big[\mu^{(i)}_{k-1} \widehat{y}^{(j)T}_{k/k-1}\big] &= \mathcal{E}^{(ji)T}_{k-1} \mathbf{H}^{*(j)T}_{\mathbb{A}_k} + \Pi^{(ij)}_{k-1} \Pi^{(j)-1}_{k-1} \mathcal{V}^{(j)T}_{k,k-1}. \end{aligned}$$

From these expectations, taking into account that $\overline{\rho}^{(i)}_k \overline{H}^{(i)}_k \mathbb{A}_k - \mathbf{H}^{*(i)}_{\mathbb{A}_k} = \mathbf{H}^{(i)}_{\mathbb{A}_k}$, expression (36) is straightforward.

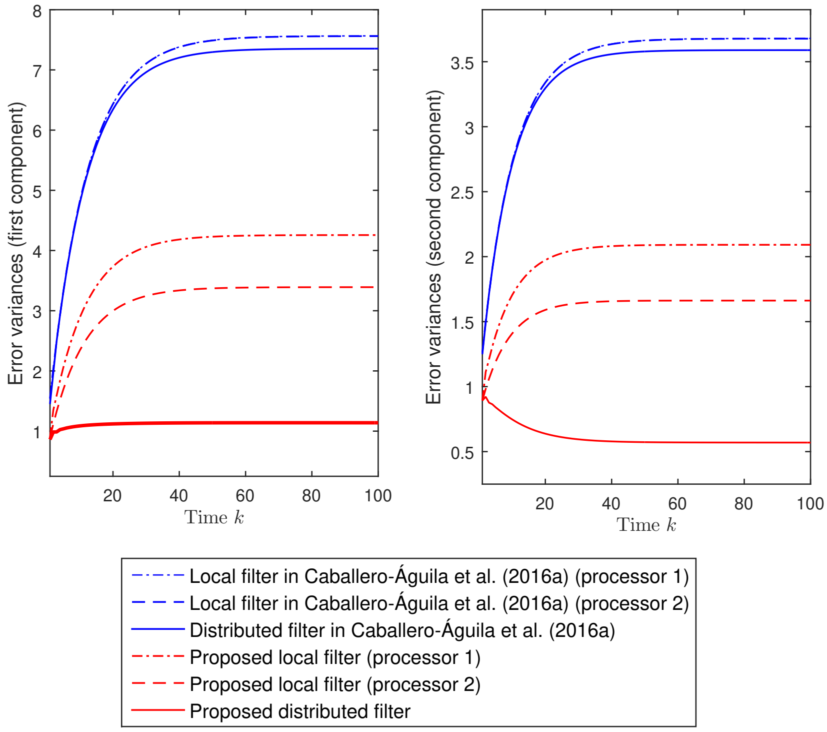

Figure 1.

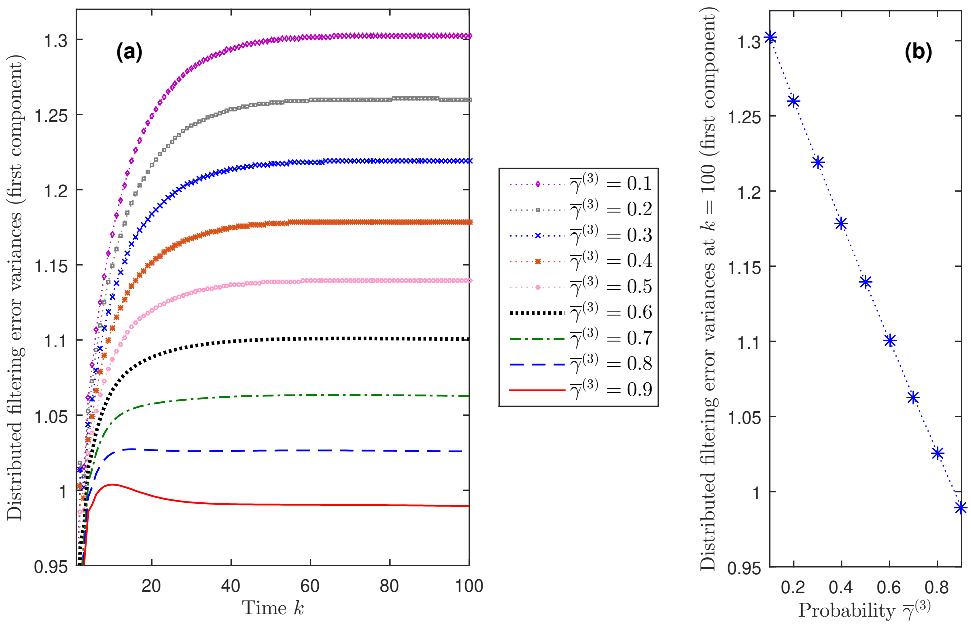

Figure 2.

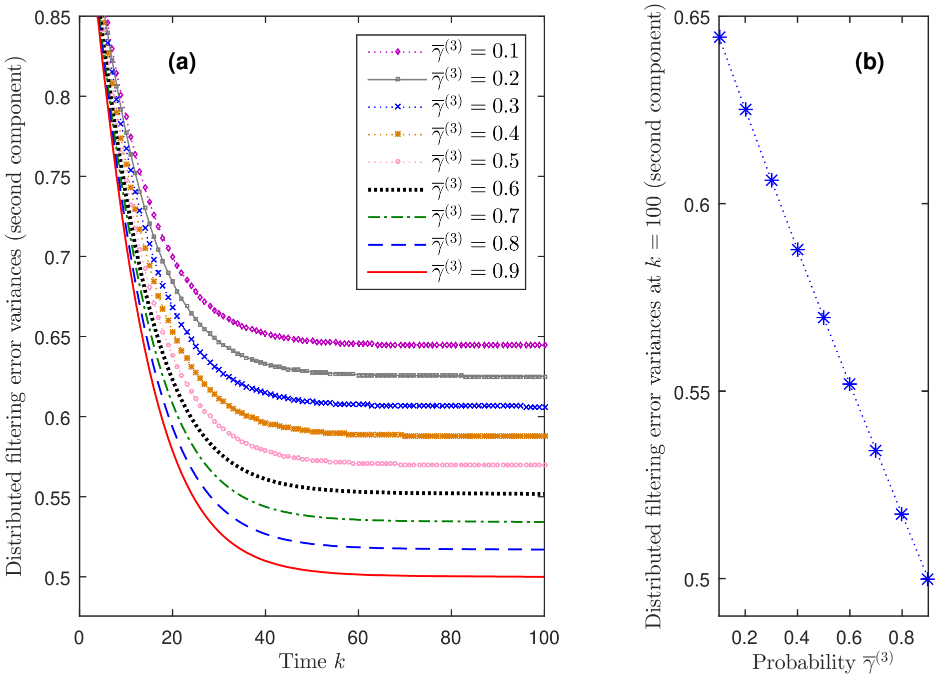

Figure 3.

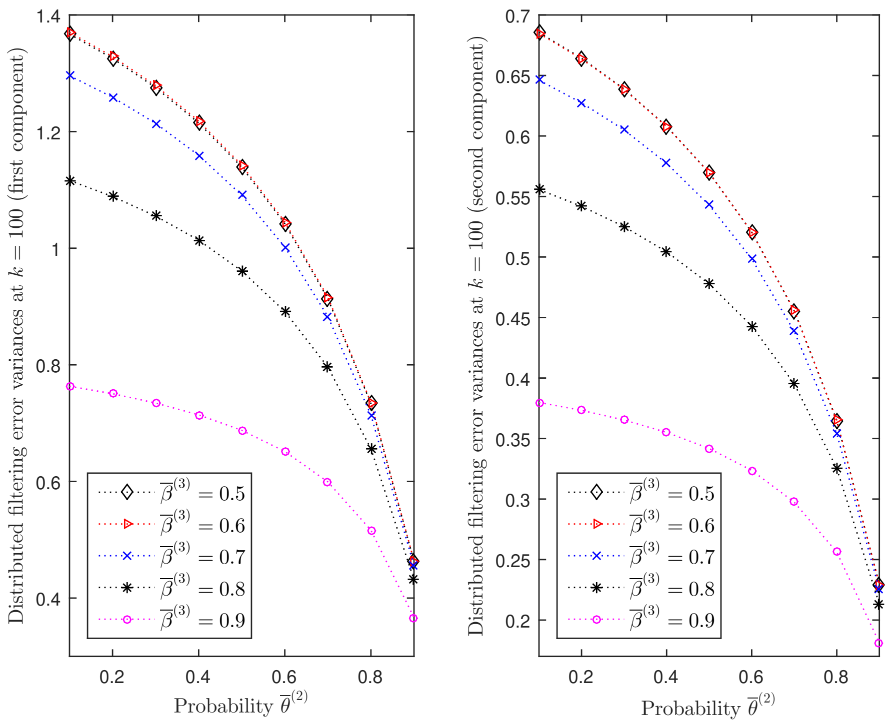

Figure 4.

Figure captions

Figure 1. Error variance comparison of the proposed filtering estimators and the filtering estimators in Caballero-Águila et al. (2016a), when $\overline{\gamma}^{(3)} = 0.5$, $\overline{\theta}^{(i)} = 0.5$, $i = 1, 2$, and $\overline{\beta}^{(i)} = 0.5$, $i = 1, 3$.

Figure 2. Distributed filtering error variances (first component): **(a)** for $\overline{\gamma}^{(3)} = 0.1$ to 0.9; **(b)** at $k = 100$, versus $\overline{\gamma}^{(3)}$.

Figure 3. Distributed filtering error variances (second component): **(a)** for $\overline{\gamma}^{(3)} = 0.1$ to 0.9; **(b)** at $k = 100$, versus $\overline{\gamma}^{(3)}$.

Figure 4. Distributed filtering error variances at $k = 100$, versus $\overline{\theta}^{(2)}$, when $\overline{\beta}^{(3)} = 0.5$ to 0.9.