Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

A new approach to distributed fusion filtering for networked systems with random parameter matrices and correlated noises

R. Caballero-Águila[*,a], A. Hermoso-Carazo[b], J. Linares-Pérez[b], Z. Wang[c]

[a] Departamento de Estadística e I.O., Universidad de Jaén, Paraje Las Lagunillas, s/n, Jaén 23071, Spain
[b] Departamento de Estadística e I.O., Universidad de Granada, Campus Fuentenueva, s/n, Granada 18071, Spain
[c] Department of Computer Science, Brunel University London, Middlesex, UB8 3PH Uxbridge, United Kingdom

Preprint version. Please cite original version:

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

https://doi.org/10.1016/j.inffus.2018.02.006

Abstract

This paper is concerned with the distributed filtering problem for a class of discrete-time stochastic systems over a sensor network with a given topology. The system presents the following main features: *(i)* random parameter matrices in both the state and observation equations are considered; and *(ii)* the process and measurement noises are one-step autocorrelated and two-step cross-correlated. The state estimation is performed in two stages. At the first stage, through an innovation approach, intermediate distributed least-squares linear filtering estimators are obtained at each sensor node by processing available output measurements not only from the sensor itself but also from its neighboring sensors according to the network topology. At the second stage, noting that at each sampling time not only the measurement but also an intermediate estimator is available at each sensor, attention is focused on the design of distributed filtering estimators as the least-squares matrix-weighted linear combination of the intermediate estimators within its neighborhood. The accuracy of both intermediate and distributed estimators, which is measured by the error covariance matrices, is examined by a numerical simulation example where a four-sensor network is considered. The example illustrates the applicability of the proposed results to a linear networked system with state-dependent multiplicative noise and different network-induced stochastic uncertainties in the measurements; more specifically, sensor gain degradation, missing measurements and multiplicative observation noises are considered as particular cases of the proposed observation model.

1. Introduction

Estimation over sensor networks systems. In the last decades, sensor networks have shown to be a persistent focus of research due to their successful applications in a wide variety of areas (e.g., target tracking, habitat monitoring, animal tracking, communications, etc.). Accordingly, considerable research attention has been devoted to state estimation techniques over sensor networks, not only due to the large number of potential applications but also because they provide more information than traditional communication systems with a single sensor. Using different approaches, a large number of research results on the design of fusion estimation algorithms in multi-sensor systems have been reported (see e.g., [1], [2], [3], [4], [5], [6]).

Distributed estimation problem. Usually, the information available at each individual node of the sensor network comes not only from its own measurements but also from those of its neighboring sensors according to a given topology and, instead of sending their information to the fusion center, each sensor node itself can perform an estimation by incorporating all the information from its neighbors. Hence, for distributed estimation problems, it is of fundamental importance to establish a strategy to describe how each node communicates with its neighboring nodes according to the information provided by the network topology. Recently, the distributed filtering problem through sensor networks has gained an ever-increasing interest and, using different filter structures, a great number of distributed algorithms have been proposed (see e.g., [7], [8], [9], [10], [11]). A survey of recent advances on distributed filtering for stochastic systems over sensor networks has been given in [12] and [13] where a comprehensive overview on this field was provided.

Incomplete information. In stochastic systems within a networked environment, certain network-induced phenomena can occur randomly due to many reasons such as network congestion, intermittent sensor failures or accidental loss of measured data, among others. These random phenomena (e.g., missing measurements, random communication delays or packet dropouts, to mention a few), which are referred to in [14] as randomly occurring incomplete information, have a great impact on the performance of the estimators and make it necessary to develop new distributed estimation algorithms that take them into account. In the design of these new distributed estimation algorithms, the difficulties caused by the coupling between the sensors according to the given topology must be overcome in addition to those arising from the random phenomena induced by the network. In recent years, the distributed estimation problem with incomplete information has become a research topic of growing interest (see e.g., [15], [16], [17], [18]).

Random parameter matrices. Usually, some of the systems describing the aforementioned network-induced random phenomena include stochastic parameters, so they can be transformed into systems with random parameter matrices. Some examples are networked systems with random observation losses [19], stochastic sensor gain degradation [20], multiplicative noises in the observation equations [21], missing [22] and fading measurements [23], or measurement multiplicative noises and missing measurements [24]. Also, the original system with random delays and packet dropouts in [25] and [26] can be transformed into an equivalent stochastic parameterized system. . Similarly, systems with two-step random delays have been transformed into systems with random parameter matrices in several papers, e.g., [27] and [28]. Moreover, it is noted that systems with random state transition matrices can be used, for example, to describe linear systems with state-dependent multiplicative noise [29] or randomly variant dynamic systems with multiple models [30]. Consequently, random state transition and measurement parameter matrices can model a great variety of real situations and communication processes, as they provide a unified framework to address different simultaneous network-induced phenomena. Discrete-time systems with random parameter matrices arise in areas such as digital control of chemical processes, systems with human operators, mobile robot localization, navigation systems, economic systems and stochastically sampled digital control systems ([31], [32]). This wide applicability has encouraged an increasing interest on the estimation problem for systems with random parameter matrices (see e.g., [30], [31], [32], [33], [34], [35]).

Noise correlation. In the study of estimation problems, a general assumption about the system noises is that they are uncorrelated or correlated only at the same time instant. However, this assumption is not always true and it can be restrictive in many real-world problems where both correlation and cross-correlation of the noises may be present. For example, when the sensors operate in the same noisy environment, the sensor noises are usually correlated. Also, when the noises are state dependent, there is cross-correlation between the process noise and the sensor noises, as well as between the different sensor noises. Furthermore, the augmented systems used to describe random delays and packet dropouts have correlated noises, and discretized continuous-time systems also have inherently correlated noises. Hence, both in systems with deterministic matrices and systems with random parameter matrices, the estimation problem with correlated and cross-correlated noises has become a challenging research topic. In the first case, under different correlation assumptions of the noises, centralized and distributed fusion algorithms are obtained in [29], for systems with multiplicative noise in the state equation; in [27], when multiplicative noises exist in both the state and observation equations; in [36] for systems with fading measurements; and in [5], for systems with finite-step correlated noises and multiple packet dropouts. For systems with random parameter matrices and autocorrelated and cross-correlated noises, many research efforts have been devoted to the fusion estimation problems (see e.g., [31], [32], [33], [34], [35]).

Addressed problem. Motivated by the above considerations, this paper is concerned with the study of the distributed state estimation problem for systems perturbed by random parameter matrices and correlated additive noises over a sensor network with a given topology. The design of the proposed distributed filtering estimators is carried out in two stages. At the first stage, using an innovation approach, every sensor node collects measurements from neighboring sensors according to the network topology in order to generate intermediate least-squares linear estimators. After that, at the second stage, the intermediate estimators from neighboring sensors are further collected to form the proposed distributed estimators as the least-squares matrix-weighted linear combination of them. Since more measurements from different sensors are used to generate distributed estimators in the second stage compared with the first one, the proposed distributed method steers each distributed estimator closer to the global optimal linear one (based on the measurements of all the network sensors), thus improving the intermediate estimation performance and also reducing disagreements of intermediate estimators among different sensors.

Paper contributions. The main contributions of the current study can be highlighted as follows: (1) the considered system model includes random parameter matrices in both state and measurement equations which provides a unified framework comprehending, for example, multiplicative noise in the state equation and some network-induced phenomena such as missing measurements or sensor gain degradations; hence, the proposed algorithm can be applied to these kinds of network systems with incomplete information; (2) the random parameters are time-varying, thus allowing to cover general situations involving network-induced phenomena that depend explicitly upon time and, moreover, different random phenomena at the different sensor nodes can be considered; (3) one-step autocorrelation of the noises and also two-step cross-correlation between the process noise and different sensor noises are considered; (4) unlike most existing papers on distributed estimation, where optimal linear estimators with a given structure are obtained, in this paper an optimal linear distributed filter is designed, without requiring a particular structure on the estimators, but just using the mean squared error criterion; (5) the innovation technique is used to simplify substantially the derivation of the proposed algorithm which is recursive and computationally simple, thus being suitable for online implementation.

Paper structure. The rest of the paper is organized as follows. In Section 2, we present the system model to be considered and the assumptions under which the distributed estimation problem is addressed. In Section 3, using an innovation approach, a recursive algorithm for the intermediate least-squares linear filter is derived. In Section 4, the proposed distributed filter is generated as a matrix-weighted linear combination of the intermediate filtering estimators within its communication neighborhood, using the mean squared error as optimality criterion. An illustrative example is provided in Section 5 to show the performance of the proposed estimators. Finally, some conclusions are drawn in Section 6.

Notation: The notation used throughout the paper is standard. $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. A^T and A^{-1} denote the transpose and the inverse of a matrix A, respectively. The shorthand $Diag(A_i)_{i=1,\dots,m}$ denotes a block diagonal matrix with matrices $A_1, \dots, A_m$, and $(A_1, \dots, A_m)$ denotes a partitioned matrix into sub-matrices $A_1, \dots, A_m$. I_n is the $n \times n$ identity matrix. If the dimensions of matrices are not explicitly stated, they are assumed to be compatible with algebraic operations. The notation symbol $\otimes$ represents the Kronecker product. $\delta_{k,s}$ denotes the Kronecker delta function, which is equal to one if $s = k$ and zero otherwise. Finally, for any function $G_{k,s}$, dependent of the time instants k and s, we will write $G_k \equiv G_{k,k}$ for simplicity; analogously, $G^{(i)} \equiv G^{(ii)}$ will be written for any function $G^{(ij)}$, dependent of the sensor nodes i and j.

2. System formulation and problem statement

Our purpose is to study the distributed filtering problem in systems with random parameter matrices and correlated noises, over a sensor network with a given topology; at each sensor node, the state filtering estimators are based not only on its own information but also on the information from its neighboring nodes.

Consider a sensor network with a fixed topology represented by a directed graph of order m, $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$. Here, $\mathcal{V} = \{1, \dots, m\}$ is the set of sensor nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges connecting some pairs of nodes. Since the graph is directed, the edges have a specific direction; namely, $(i, j) \in \mathcal{E}$ means that sensor i can obtain information from sensor j. $\mathcal{A} = (a_{ij})_{m \times m}$ is the weighted adjacency matrix, whose elements (the edge weights) are nonnegative finite real numbers indicating whether pairs of vertices are connected or not in the graph, since $a_{ij} > 0 \;\Leftrightarrow\; (i, j) \in \mathcal{E}$. We assume that $a_{ii} = 1, \; \forall i \in \mathcal{V}$, and therefore (i, i) can be regarded as an additional edge. The set of neighbors of node i, plus the node itself, is denoted by $\mathcal{N}_i = \{j \in \mathcal{V} : \; a_{ij} > 0\}, \; \forall i \in \mathcal{V}$, and it is assumed that each node i knows all the relevant information from its adjacent nodes, $j \in \mathcal{N}_i$. A communication graph $\mathcal{G}$ is said to be completely connected if there is an edge between every pair of nodes; that is, $(i, j) \in \mathcal{E}$ for all $i, j \in \mathcal{V}$ or, equivalently, $\mathcal{N}_i = \mathcal{V}, \; \forall i \in \mathcal{V}$.

Consider the following discrete-time linear stochastic system:

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

where $x_k \in \mathbb{R}^{n_x}$ is the state vector at time k, $\{F_k;\; k \geq 0\}$ is a sequence of random parameter matrices and $\{w_k;\; k \geq 0\}$ is the process noise.

The state measured outputs from the different sensor nodes are described by:

$$y_k^{(i)} = H_k^{(i)} x_k + v_k^{(i)}, \quad k \geq 1, \quad i = 1, \dots, m, \tag{2}$$

where $y_k^{(i)} \in \mathbb{R}^{n_y}$ is the output from sensor node i at time k. For $i = 1, \dots, m$, $\{H_k^{(i)};\; k \geq 1\}$ is a sequence of random parameter matrices and $\{v_k^{(i)};\; k \geq 1\}$ is the measurement noise of the sensor node i.

Model assumptions. The distributed filtering problem is addressed under the following assumptions about the initial state, the random parameter matrices and the noises involved in the system model (1)-(2):

(i) The initial state x_0 is a random vector with $E[x_0] = \overline{x}_0$ and $Cov[x_0] = \Sigma_0$.

(ii) $\{F_k;\ k \geq 0\}$ and $\{H_k^{(i)};\ k \geq 1\}$, $i = 1, \dots, m$, are independent sequences of independent random parameter matrices with known means, $E[F_k] = \overline{F}_k$, $E[H_k^{(i)}] = \overline{H}_k^{(i)}$, and the covariances of their entries, $Cov[f_{pq}(k), f_{p'q'}(k)]$, $Cov[h_{pq}^{(i)}(k), h_{p'q'}^{(i)}(k)]$, are also assumed to be known. $f_{pq}(k)$ denotes the (p,q)-th entry of matrix F_k, for $p, q = 1, \dots, n_x$, and $h_{pq}^{(i)}(k)$ denotes the (p,q)-th entry of $H_k^{(i)}$, for $p = 1, \dots, n_y$ and $q = 1, \dots, n_x$.

(iii) The noises $\{w_k;\ k \geq 0\}$ and $\{v_k^{(i)};\ k \geq 1\}$, $i = 1, \dots, m$, are zero-mean sequences with known covariances and cross-covariances:

$$\begin{aligned} &Cov[w_k, w_s] = Q_k\delta_{k,s} + Q_{k,k-1}\delta_{k-1,s}, \quad s \leq k, \\ &Cov[v_k^{(i)}, v_s^{(j)}] = R_k^{(ij)}\delta_{k,s} + R_{k,k-1}^{(ij)}\delta_{k-1,s}, \quad s \leq k, \\ &Cov[w_k, v_s^{(i)}] = S_k^{(i)}\delta_{k,s} + S_{k,k+1}^{(i)}\delta_{k+1,s} + S_{k,k+2}^{(i)}\delta_{k+2,s}. \end{aligned}$$

(iv) For $i = 1, \dots, m$, the initial state x_0 and the processes $\{F_k;\ k \geq 0\}$ and $\{H_k^{(i)};\ k \geq 1\}$ are mutually independent and they are independent of the additive noises $\{w_k;\ k \geq 0\}$ and $\{v_k^{(i)};\ k \geq 1\}$.

Remark 1. By denoting $\widetilde{F}_k = F_k - \overline{F}_k$ and $\widetilde{H}_k^{(i)} = H_k^{(i)} - \overline{H}_k^{(i)}$, $i = 1, \dots, m$, the following identities hold for the (p,q)-th entries of the matrices $E[\widetilde{F}_k G \widetilde{F}_k^T]$ and $E[\widetilde{H}_k^{(i)} G \widetilde{H}_k^{(i)T}]$, being G an arbitrary deterministic matrix:

$$\begin{aligned} \left(E[\widetilde{F}_k G \widetilde{F}_k^T]\right)_{pq} &= \sum_{a=1}^{n_x}\sum_{b=1}^{n_x} Cov[f_{pa}(k), f_{qb}(k)]G_{ab}, \quad p, q = 1, \dots, n_x, \\ \left(E[\widetilde{H}_k^{(i)} G \widetilde{H}_k^{(i)T}]\right)_{pq} &= \sum_{a=1}^{n_x}\sum_{b=1}^{n_x} Cov[h_{pa}^{(i)}(k), h_{qb}^{(i)}(k)]G_{ab}, \quad p, q = 1, \dots, n_y. \end{aligned}$$

Remark 2. Assumptions *(i)-(iv)* lead to the following recursive formula for $\mathcal{D}_k \equiv E[x_k x_k^T]$, the correlation matrix of the state vector x_k (see, e.g., [31]):

$$\begin{aligned} \mathcal{D}_{k+1} &= \overline{F}_k \mathcal{D}_k \overline{F}_k^T + E[\widetilde{F}_k \mathcal{D}_k \widetilde{F}_k^T] + Q_k \\ &\quad + \overline{F}_k Q_{k-1,k} + Q_{k,k-1}\overline{F}_k^T, \quad k \geq 1; \\ \mathcal{D}_1 &= \overline{F}_0 \mathcal{D}_0 \overline{F}_0^T + E[\widetilde{F}_0 \mathcal{D}_0 \widetilde{F}_0^T] + Q_0, \\ \mathcal{D}_0 &= \Sigma_0 + \overline{x}_0 \overline{x}_0^T. \end{aligned} \tag{3}$$

Under the previous assumption that each node i has access to the information from its neighbors, we consider that the communication between adjacent nodes is coordinated and conducted in two stages: in the first one, every node sends out only its local measurements and, in the second stage, every node sends out only the estimators obtained using the local measurements. Our aim is to find the distributed filtering estimator, $\widehat{x}_{k/k}^{(i)}$, of the state x_k based on its own information and that from its neighboring nodes, $j \in \mathcal{N}_i$. Taking into account the communication between adjacent nodes, the proposed estimators are performed in two steps: *Step 1)* An intermediate distributed optimal least-squares (LS) linear filter of the signal x_k, denoted by $\widehat{x}_{k/k}^{d(i)}$, is obtained using the measurements $y_1^{(j)}, \dots, y_k^{(j)}$, for all $j \in \mathcal{N}_i$. *Step 2)* Motivated by the fact that the neighbors of the node i have also their own estimators for the same signal x_k, the proposed distributed estimator, $\widehat{x}_{k/k}^{(i)}$, is generated by a matrix-weighted linear combination of the intermediate estimators within its communication neighborhood, $\widehat{x}_{k/k}^{d(j)}$, $j \in \mathcal{N}_i$, using the mean squared error as optimality criterion.

2.1. Stacked observation model

For notational simplicity in the mathematical derivations, the observation model (2) is rewritten in a stacked form as follows:

$$Y_k = \mathcal{H}_k x_k + V_k, \qquad k \geq 1, \tag{4}$$

where

$$Y_k = \begin{pmatrix} y_k^{(1)} \\ \vdots \\ y_k^{(m)} \end{pmatrix}, \quad \mathcal{H}_k = \begin{pmatrix} H_k^{(1)} \\ \vdots \\ H_k^{(m)} \end{pmatrix}, \quad V_k = \begin{pmatrix} v_k^{(1)T} \\ \vdots \\ v_k^{(m)} \end{pmatrix}.$$

The following properties of the processes in (4) are easily inferred from the model assumptions previously stated:

- $\{\mathcal{H}_k;\ k \geq 1\}$ are independent random parameter matrices with known means, $E[\mathcal{H}_k] = \overline{\mathcal{H}}_k = \left(\overline{H}_k^{(1)T}, \dots, \overline{H}_k^{(m)T}\right)^T$, and for any deterministic matrix G, we have:

  $$E[\widetilde{\mathcal{H}}_k G \widetilde{\mathcal{H}}_k^T] = Diag\left(E[\widetilde{H}_k^{(i)} G \widetilde{H}_k^{(i)T}]\right)_{i=1,\dots,m}$$

  where $\widetilde{\mathcal{H}}_k = \mathcal{H}_k - \overline{\mathcal{H}}_k$ and $E[\widetilde{H}_k^{(i)} G \widetilde{H}_k^{(i)T}]$ is obtained as indicated in *Remark 1.*

- $\{V_k;\ k \geq 1\}$ is a zero-mean process with

$$E[V_k V_s^T] = R_k \delta_{k,s} + R_{k,k-1}\delta_{k-1,s}, \quad s \leq k,$$
$$E[w_k V_s^T] = S_k \delta_{k,s} + S_{k,k+1}\delta_{k+1,s} + S_{k,k+2}\delta_{k+2,s},$$

  being

$$R_{k,s} = \left(R_{k,s}^{(ij)}\right)_{i,j=1,\ldots,m}, \quad S_{k,s} = \left(S_{k,s}^{(1)}, \ldots, S_{k,s}^{(m)}\right).$$

- The initial state x_0 and the random matrix sequences $\{F_k;\ k \geq 0\}$ y $\{\mathcal{H}_k;\ k \geq 1\}$ are mutually independent and independent of the additive noises $\{w_k;\ k \geq 0\}$ and $\{V_k;\ k \geq 1\}$.

From these properties, the following correlation properties of the vector noises w_k and V_k are clear:

- The process noise vector w_k is uncorrelated with $Y_1, \ldots, Y_{k-1}$ and correlated with Y_k; the correlation matrix $\mathcal{W}_k \equiv E[w_k Y_k^T]$ is obtained by

$$\mathcal{W}_k = Q_{k,k-1}\overline{\mathcal{H}}_k^T + S_k, \quad k \geq 1. \tag{5}$$

- The observation noise vector V_k is uncorrelated with $Y_1, \ldots, Y_{k-2}$ and correlated with Y_{k-1}; noting $\mathcal{V}_{k,k-1} \equiv E[V_k Y_{k-1}^T]$, we have

$$\mathcal{V}_{k,k-1} = S_{k-2,k}^T \overline{\mathcal{H}}_{k-1}^T + R_{k,k-1}, \quad k \geq 2. \tag{6}$$

- The state vector x_k is correlated with the noise vector V_k and $\mathcal{B}_k \equiv E[x_k V_k^T]$ satisfy

$$\mathcal{B}_k = \overline{F}_{k-1} S_{k-2,k} + S_{k-1,k}, \quad k \geq 2; \quad \mathcal{B}_1 = S_{0,1}. \tag{7}$$

3. Intermediate distributed LS linear filter

In this section, our aim is to derive the LS linear filter that results when, at each sensor node, not only its own measurements but also all the available measurements from its neighboring nodes are used. Therefore, for every sensor node i, our challenge is to obtain the filter $\widehat{x}_{k/k}^{d(i)}$ based on all the measurements from the nodes in its neighborhood set, $\mathcal{N}_i = \{j \in \mathcal{V} : a_{ij} > 0\}$, up to time k or, equivalently, based on those observations $y_s^{(j)},\ s \leq k$ for which $a_{ij} > 0$.

To give a unified expression of the observations used at each node, let us define $c_{ij} = 1$, if $a_{ij} > 0$, and $c_{ij} = 0$, otherwise; i.e., $c_{ij} = 1$ means that the nodes i and j are connected and a_{ij} is the edge weight. Let us denote $C_{i,y}$ the matrix obtained by removing the all-zero rows of $Diag\,(c_{ij})_{j=1,\dots,m} \otimes I_{n_y}$ and $Z_s^{(i)} = C_{i,y} Y_s$. As Y_s is the stacked observation vector (4), $Z_s^{(i)}$ is the vector constituted only by those observations $y_s^{(j)}$ such that $c_{ij} = 1$ or, equivalently, those observations coming from the neighboring nodes of sensor i. Then, the aim is to derive a recursive algorithm to obtain the LS linear filter of the state x_k based on $\{Z_s^{(i)},\ s \leq k\}$.

Innovation approach. For each $i = 1, \dots, m$, the recursive algorithm for the LS linear filter, $\widehat{x}_{k/k}^{d(i)}$, is derived by using an innovation approach. The innovation at time k is defined as $\mu_k^{(i)} = Z_k^{(i)} - \widehat{Z}_{k/k-1}^{d(i)} = C_{i,y}\left(Y_k - \widehat{Y}_{k/k-1}^{d(i)}\right)$, where $\widehat{Y}_{k/k-1}^{d(i)}$ is the LS linear estimator of Y_k based on $Z_s^{(i)},\ s \leq k-1$.

Replacing the observation process $\{Z_k^{(i)};\ k \geq 1\}$ by the innovation one, $\{\mu_k^{(i)};\ k \geq 1\}$, and considering an arbitrary number of observations, L, the LS linear estimator, $\widehat{\xi}_{k/L}^{d(i)}$, of a random vector ξ_k based on the observations $Z_1^{(i)}, \dots, Z_L^{(i)}$, can be calculated as linear combination of the innovations $\mu_1^{(i)}, \dots, \mu_L^{(i)}$; namely, $\widehat{\xi}_{k/L}^{d(i)} = \sum_{s=1}^{L} h_{k,s,L}^{(i)} \mu_s^{(i)}$. Using now the orthogonality conditions, $E[(\xi_k - \widehat{\xi}_{k/L}^{d(i)})\mu_s^{(i)T}] = 0,\quad s = 1, \dots, L$, and the whiteness of the innovation process, it is deduced that $h_{k,s,L}^{(i)} = E[\xi_k \mu_s^{(i)T}](E[\mu_s^{(i)} \mu_s^{(i)T}])^{-1}$; hence, $h_{k,s,L}^{(i)}$ is independent of L and, noting $\Pi_s^{(i)} = E[\mu_s^{(i)} \mu_s^{(i)T}]$, the following identity holds

$$\widehat{\xi}_{k/L}^{d(i)} = \sum_{s=1}^{L} E[\xi_k \mu_s^{(i)T}] \Pi_s^{(i)-1} \mu_s^{(i)}. \tag{8}$$

This general expression for the LS linear estimators as linear combination of the innovations is the starting point to derive the following recursive filtering algorithm.

Theorem 1. *The LS linear filter, $\widehat{x}_{k/k}^{d(i)}$, is given by*

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

where the one-stage state predictor, $\widehat{x}^{d(i)}_{k/k-1}$, satisfies

$$\begin{array}{l}\widehat{x}^{d(i)}_{k/k-1} = \overline{F}_{k-1}\widehat{x}^{d(i)}_{k-1/k-1} + \mathcal{W}_{k-1}C^T_{i,y}\Pi^{(i)-1}_{k-1}\mu^{(i)}_{k-1}, \ \ k \geq 2; \\ \widehat{x}^{d(i)}_{1/0} = \overline{F}_0\overline{x}_0. \end{array} \tag{10}$$

The filtering error covariance matrix, $\Sigma^{d(i)}_{k/k}$, is given by

$$\Sigma^{d(i)}_{k/k} = \Sigma^{d(i)}_{k/k-1} - \mathcal{X}^{(i)}_k \Pi^{(i)-1}_k \mathcal{X}^{(i)T}_k, \quad k \geq 1; \quad \Sigma^{d(i)}_{0/0} = \Sigma_0, \tag{11}$$

where the prediction error covariance matrix, $\Sigma^{d(i)}_{k/k-1}$, is calculated by

$$\begin{array}{l}\Sigma^{d(i)}_{k/k-1} = \mathcal{D}_k + \overline{F}_{k-1}\left(\Sigma^{d(i)}_{k-1/k-1} - \mathcal{D}_{k-1}\right)\overline{F}^T_{k-1} - \mathcal{X}^{(i)}_{k,k-1}\Pi^{(i)-1}_{k-1}C_{i,y}\mathcal{W}^T_{k-1} \\ \qquad\qquad -\mathcal{W}_{k-1}C^T_{i,y}\Pi^{(i)-1}_{k-1}\mathcal{X}^{(i)T}_{k-1}\overline{F}^T_{k-1}, \quad k \geq 2; \\ \Sigma^{d(i)}_{1/0} = \mathcal{D}_1 - \overline{F}_0\overline{x}_0\overline{x}^T_0\overline{F}^T_0. \end{array} \tag{12}$$

The matrix $\mathcal{X}^{(i)}_k \equiv E[x_k\mu^{(i)T}_k]$ is obtained by

$$\mathcal{X}^{(i)}_k = \left(\Sigma^{d(i)}_{k/k-1}\overline{\mathcal{H}}^T_k + \mathcal{M}^{(i)}_k\right)C^T_{i,y}, \quad k \geq 1, \tag{13}$$

where $\mathcal{M}^{(i)}_k \equiv E[(x_k - \widehat{x}^{d(i)}_{k/k-1})V^T_k]$ is given by

$$\begin{array}{l}\mathcal{M}^{(i)}_k = \mathcal{B}_k - \mathcal{X}^{(i)}_{k,k-1}\Pi^{(i)-1}_{k-1}C_{i,y}\mathcal{V}^T_{k,k-1}, \quad k \geq 2; \\ \mathcal{M}^{(i)}_1 = \mathcal{B}_1, \end{array} \tag{14}$$

with $\mathcal{X}^{(i)}_{k,k-1} \equiv E[x_k\mu^{(i)T}_{k-1}]$ satisfying

$$\mathcal{X}^{(i)}_{k,k-1} = \overline{F}_{k-1}\mathcal{X}^{(i)}_{k-1} + \mathcal{W}_{k-1}C^T_{i,y}, \quad k \geq 2. \tag{15}$$

The innovation, $\mu^{(i)}_k$, is given by

$$\begin{array}{l}\mu^{(i)}_k = C_{i,y}\left(Y_k - \overline{\mathcal{H}}_k\widehat{x}^{d(i)}_{k/k-1} - \mathcal{V}_{k,k-1}C^T_{i,y}\Pi^{(i)-1}_{k-1}\mu^{(i)}_{k-1}\right), \quad k \geq 2; \\ \mu^{(i)}_1 = C_{i,y}\left(Y_1 - \overline{\mathcal{H}}_1\widehat{x}^{d(i)}_{1/0}\right) \end{array} \tag{16}$$

and the innovation covariance matrix, $\Pi^{(i)}_k$, satisfies

$$\Pi^{(i)}_k = C_{i,y}\left(\overline{\mathcal{H}}_k\mathcal{X}^{(i)}_k + E[\widetilde{\mathcal{H}}_k\mathcal{D}_k\widetilde{\mathcal{H}}^T_k]C^T_{i,y} + \mathcal{V}^{(i)}_k\right), \quad k \geq 1, \tag{17}$$

where $\mathcal{V}_k^{(i)} \equiv E[V_k \mu_k^{(i)T}]$ *is obtained by*

$$
\begin{aligned}
\mathcal{V}_k^{(i)} &= \left(\mathcal{M}_k^{(i)T}\overline{\mathcal{H}}_k^T + R_k - \mathcal{V}_{k,k-1}C_{i,y}^T\Pi_{k-1}^{(i)-1}C_{i,y}\mathcal{V}_{k,k-1}^T\right)C_{i,y}^T, \quad k \geq 2;\\
\mathcal{V}_1^{(i)} &= \left(\mathcal{M}_1^{(i)T}\overline{\mathcal{H}}_1^T + R_1\right)C_{i,y}^T.
\end{aligned}
\tag{18}
$$

The matrices $\mathcal{D}_k$*,* $\mathcal{W}_k$*,* $\mathcal{V}_{k,k-1}$ *and* $\mathcal{B}_k$ *are given in (3), (5), (6) and (7), respectively, and* $C_{i,y}$ *is the matrix obtained by removing the all-zero rows of* $Diag\,(c_{ij})_{j=1,\dots,m} \otimes I_{n_y}$.

Proof. From the equations (1) and (4), and the orthogonal projection Lemma (OPL), the state predictor, $\widehat{x}_{k/k-1}^{d(i)}$, and the observation predictor, $\widehat{Y}_{k/k-1}^{d(i)}$, satisfy:

$$\widehat{x}_{k/k-1}^{d(i)} = \overline{F}_{k-1}\widehat{x}_{k-1/k-1}^{d(i)} + \widehat{w}_{k-1/k-1}^{d(i)}, \quad k \geq 1, \tag{19}$$

$$\widehat{Y}_{k/k-1}^{d(i)} = \overline{\mathcal{H}}_k\widehat{x}_{k/k-1}^{d(i)} + \widehat{V}_{k/k-1}^{d(i)}, \quad k \geq 1, \tag{20}$$

where $\widehat{w}_{k-1/k-1}^{d(i)}$ and $\widehat{V}_{k/k-1}^{d(i)}$ are the LS linear estimators of w_{k-1} and V_k, respectively, based on $Z_s^{(i)},\ s \leq k-1$. Note that, due to the noise correlation assumptions, the vectors w_{k-1} and V_k are correlated with the observation $Z_{k-1}^{(i)}$ and, hence, their estimators are not equal to zero; next, their expressions are obtained.

Taking into account that w_k is uncorrelated with $Z_s^{(i)}$, for $s \leq k-1$, and V_k is uncorrelated with $Z_s^{(i)}$, for $s \leq k-2$, we have

$$
\begin{aligned}
E[w_k\mu_k^{(i)T}] &= E[w_k Z_k^{(i)T}] = \mathcal{W}_k C_{i,y}^T,\\
E[V_k\mu_{k-1}^{(i)T}] &= E[V_k Z_{k-1}^{(i)T}] = \mathcal{V}_{k,k-1}C_{i,y}^T,
\end{aligned}
$$

where $\mathcal{W}_k$ and $\mathcal{V}_{k,k-1}$ are given in (5) and (6), respectively. So, the general expression (8) for the LS linear estimators leads to:

$$\widehat{w}_{k/k}^{d(i)} = \mathcal{W}_k C_{i,y}^T \Pi_k^{(i)-1}\mu_k^{(i)}, \quad k \geq 1; \quad \widehat{w}_{0/0}^{d(i)} = 0. \tag{21}$$

$$\widehat{V}_{k/k-1}^{d(i)} = \mathcal{V}_{k,k-1}C_{i,y}^T\Pi_{k-1}^{(i)-1}\mu_{k-1}^{(i)}, \quad k \geq 2; \quad \widehat{V}_{1/0}^{d(i)} = 0. \tag{22}$$

• *Derivation of expressions (9)- (12).* Denoting $\mathcal{X}_k^{(i)} = E[x_k\mu_k^{(i)T}]$, expressions (9) and (11) for the filter and the filtering error covariance matrix are obvious from (8) and the OPL, respectively. Also, expression (10) for the state predictor, with $\mathcal{W}_k$ given in (5), is immediately obtained from (19) and (21). Finally, from the OPL, $\Sigma^{d(i)}_{k/k-1} = \mathcal{D}_k - E[\widehat{x}^{d(i)}_{k/k-1}\widehat{x}^{d(i)T}_{k/k-1}]$, where $\mathcal{D}_k$ is given by (3), and using expression (10) we obtain (12) for the prediction error covariance matrix.

• *Derivation of expressions (13)-(15).* From the OPL, we have that $\mathcal{X}^{(i)}_k = E[(x_k-\widehat{x}^{d(i)}_{k/k-1})\mu^{(i)T}_k] = E[(x_k-\widehat{x}^{d(i)}_{k/k-1})Y^T_k]C^T_{i,y}$ and, from (4) for Y_k, expression (13) for $\mathcal{X}^{(i)}_k$ is obtained. Expression (14) for $\mathcal{M}^{(i)}_k$ is easily deduced from the following recursive formula for the state predictor, which is obtained from (9) and (10), together with (15) for $\mathcal{X}^{(i)}_{k,k-1}$:

$$\widehat{x}^{d(i)}_{k/k-1} = \overline{F}_{k-1}\widehat{x}^{d(i)}_{k-1/k-2} + \mathcal{X}^{(i)}_{k,k-1}\Pi^{(i)-1}_{k-1}\mu^{(i)}_{k-1}, \quad k \geq 2. \tag{23}$$

Finally, expression (15) for $\mathcal{X}^{(i)}_{k,k-1}$ is immediately obtained from (1).

• *Derivation of expressions (16)- (18).* From (20) and (22), the innovation is clearly given by (16), with $\mathcal{V}_{k,k-1}$ satisfying (6). Next, expression (17) for $\Pi^{(i)}_k = E\left[\mu^{(i)}_k\mu^{(i)T}_k\right]$ is derived. From the OPL and (4), we have that

$$\begin{aligned}\Pi^{(i)}_k &= E[Z^{(i)}_k\mu^{(i)T}_k] = C_{i,y}E[Y_k\mu^{(i)T}_k]\\ &= C_{i,y}\left(\overline{\mathcal{H}}_k\mathcal{X}^{(i)}_k + E[\widetilde{\mathcal{H}}_k x_k\mu^{(i)T}_k] + \mathcal{V}^{(i)}_k\right).\end{aligned}$$

Again, from the OPL and (4), together with the conditional expectation properties, the following identities hold:

$$\begin{aligned}E[\widetilde{\mathcal{H}}_k x_k\mu^{(i)T}_k] &= E[\widetilde{\mathcal{H}}_k x_k Z^{(i)T}_k] = E[\widetilde{\mathcal{H}}_k x_k Y^T_k]C^T_{i,y}\\ &= E[\widetilde{\mathcal{H}}_k x_k x^T_k\widetilde{\mathcal{H}}^T_k]C^T_{i,y} = E[\widetilde{\mathcal{H}}_k\mathcal{D}_k\widetilde{\mathcal{H}}^T_k]C^T_{i,y},\end{aligned}$$

and substituting this expectation into the above expression for $\Pi^{(i)}_k$, we obtain (17).

Finally, using (16) for $\mu^{(i)}_k$, with (4) for Y_k, we easily obtain that $\mathcal{V}^{(i)}_k = E[V_k\mu^{(i)T}_k]$ satisfies (18), and the proof is completed. □

4. Distributed filtering estimators

At each sensor node, the optimal LS linear intermediate distributed filtering estimators obtained in the previous section use the information of the measurements from the node itself and its neighboring nodes. Now, this information can be complemented, since each node can access its neighbors intermediate filtering estimators. On this basis, our goal now is to design a new type of distributed filter for every node, by using its own intermediate filtering estimators and those of its neighbors; specifically, at each sensor node i, a distributed filter, $\widehat{x}^{(i)}_{k/k}$, will be generated as a matrix-weighted sum of the intermediate filters, $\widehat{x}^{d(j)}_{k/k}$, for $j \in \mathcal{N}_i$, in which the matrix weights are computed to minimize the mean squared estimation error. So, since $c_{ij} = 1$ for $j \in \mathcal{N}_i$, by denoting $\widehat{X}^{(i)}_{k/k} = C_{i,x}\widehat{X}_{k/k}$, with $\widehat{X}_{k/k} = \left(\widehat{x}^{d(1)T}_{k/k}, \dots, \widehat{x}^{d(m)T}_{k/k}\right)^T$ and $C_{i,x}$ the matrix obtained by removing the all-zero rows of $Diag\,(c_{ij})_{j=1,\dots,m} \otimes I_{n_x}$, the aim is to find $\boldsymbol{\mathcal{A}}^{(i)}_k$ such that the estimator $\boldsymbol{\mathcal{A}}^{(i)}_k \widehat{X}^{(i)}_{k/k}$ minimizes

$$E\left[\left(x_k - \boldsymbol{\mathcal{A}}^{(i)}_k \widehat{X}^{(i)}_{k/k}\right)\left(x_k - \boldsymbol{\mathcal{A}}^{(i)}_k \widehat{X}^{(i)}_{k/k}\right)^T\right].$$

As it is known, the solution of this problem is given by

$$\boldsymbol{\mathcal{A}}^{(i)}_k = E[x_k \widehat{X}^{(i)T}_{k/k}]\left(E[\widehat{X}^{(i)}_{k/k}\widehat{X}^{(i)T}_{k/k}]\right)^{-1}, \quad k \geq 0. \tag{24}$$

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

$$E[x_k \widehat{X}^T_{k/k}] = \left(E[\widehat{x}^{d(1)}_{k/k}\widehat{x}^{d(1)T}_{k/k}], \dots, E[\widehat{x}^{d(m)}_{k/k}\widehat{x}^{d(m)T}_{k/k}]\right),$$

to obtain the optimal matrix $\boldsymbol{\mathcal{A}}^{(i)}_k$ in (24), the cross-covariance matrices between the intermediate estimators of every pair of nodes, $K^{(lj)}_{k/k} \equiv E[\widehat{x}^{d(l)}_{k/k}\widehat{x}^{d(j)T}_{k/k}]$, must be calculated.

From expression (9) for the intermediate filters, it is clear that $K^{(lj)}_{k/k}$ can be obtained from the cross-covariance matrices between the corresponding predictors, $K^{(lj)}_{k/k-1} \equiv E[\widehat{x}^{d(l)}_{k/k-1}\widehat{x}^{d(j)T}_{k/k-1}]$, if the expectations $E[\widehat{x}^{d(l)}_{k/k-1}\mu^{(j)T}_k]$ and $E[\mu^{(l)}_k \mu^{(j)T}_k]$ are known. Expressions for these expectations and for the prediction and filtering cross-covariance matrices are presented as preliminaries. The notations throughout this section are those of Theorem 1.

4.1. Preliminary results

The following lemmas 1 and 2 present expressions for the matrices $L^{(lj)}_k \equiv E[\widehat{x}^{d(l)}_{k/k-1}\mu^{(j)T}_k]$ and $\Pi^{(lj)}_k \equiv E[\mu^{(l)}_k \mu^{(j)T}_k]$, for arbitrary $l, j = 1, \dots, m$.

Lemma 1. *For $l, j = 1, \ldots, m$, the expectation $L_k^{(lj)} = E[\widehat{x}_{k/k-1}^{d(l)} \mu_k^{(j)T}]$ satisfies*

$$\begin{aligned} L_k^{(lj)} &= \Big[\Big(K_{k/k-1}^{(l)} - K_{k/k-1}^{(lj)} \Big) \overline{\mathcal{H}}_k^T + \Big(\mathcal{X}_{k,k-1}^{(l)} \Pi_{k-1}^{(l)-1} C_{l,y} \\ & \qquad - L_{k,k-1}^{(lj)} \Pi_{k-1}^{(j)-1} C_{j,y} \Big) \mathcal{V}_{k,k-1}^T \Big] C_{j,y}^T, \quad k \geq 2; \\ L_1^{(lj)} &= 0, \end{aligned}$$

where $L_{k,k-1}^{(lj)} = E[\widehat{x}_{k/k-1}^{d(l)} \mu_{k-1}^{(j)T}]$ is given by

$$L_{k,k-1}^{(lj)} = \overline{F}_{k-1} L_{k-1}^{(lj)} + \mathcal{X}_{k,k-1}^{(l)} \Pi_{k-1}^{(l)-1} \Pi_{k-1}^{(lj)}, \quad k \geq 2.$$

Proof. Taking into account expression (16) for $\mu_k^{(j)}$, we have that

$$L_k^{(lj)} = \Big(E[\widehat{x}_{k/k-1}^{d(l)} Y_k^T] - K_{k/k-1}^{(lj)} \overline{\mathcal{H}}_k^T - L_{k,k-1}^{(lj)} \Pi_{k-1}^{(j)-1} C_{j,y} \mathcal{V}_{k,k-1}^T \Big) C_{j,y}^T.$$

Then, using (4) for Y_k and (23) for $\widehat{x}_{k/k-1}^{d(l)}$, we obtain

$$E[\widehat{x}_{k/k-1}^{d(l)} Y_k^T] = K_{k/k-1}^{(l)} \overline{\mathcal{H}}_k^T + \mathcal{X}_{k,k-1}^{(l)} \Pi_{k-1}^{(l)-1} C_{l,y} \mathcal{V}_{k,k-1}^T,$$

and the expression of $L_k^{(lj)}$ is immediately derived. The proof is completed with the expression of $L_{k,k-1}^{(lj)}$, which is immediately obtained using again (23) for $\widehat{x}_{k/k-1}^{d(l)}$. □

Lemma 2. *For $l, j = 1, \ldots, m$, the innovation cross-covariance matrix, $\Pi_k^{(lj)} = E[\mu_k^{(l)} \mu_k^{(j)T}]$, satisfies*

$$\begin{aligned} \Pi_k^{(lj)} &= C_{l,y} \Big[\overline{\mathcal{H}}_k \Big(\mathcal{X}_k^{(j)} - L_k^{(lj)} \Big) + E[\widetilde{\mathcal{H}}_k \mathcal{D}_k \widetilde{\mathcal{H}}_k^T] C_{j,y}^T \\ & \qquad + \mathcal{V}_k^{(j)} - \mathcal{V}_{k,k-1} C_{l,y}^T \Pi_{k-1}^{(l)-1} \Pi_{k-1,k}^{(lj)} \Big], \quad k \geq 2; \\ \Pi_1^{(lj)} &= C_{l,y} \Big(\overline{\mathcal{H}}_1 \mathcal{X}_1^{(j)} + E[\widetilde{\mathcal{H}}_1 \mathcal{D}_1 \widetilde{\mathcal{H}}_1^T] C_{j,y}^T + \mathcal{V}_1^{(j)} \Big), \end{aligned}$$

where $\Pi_{k-1,k}^{(lj)} = E[\mu_{k-1}^{(l)} \mu_k^{(j)T}]$, $k \geq 2$, is obtained by

$$\Pi_{k-1,k}^{(lj)} = \Big[\overline{\mathcal{H}}_k \Big(\mathcal{X}_{k,k-1}^{(l)} - L_{k-1}^{(jl)} \Big) + \mathcal{V}_{k,k-1} \Big(C_{l,y}^T - C_{j,y}^T \Pi_{k-1}^{(j)-1} \Pi_{k-1}^{(jl)} \Big) \Big]^T C_{j,y}^T.$$

Proof. Using (16) for the innovation $\mu_k^{(l)}$, it is clear that

$$\Pi_k^{(lj)} = C_{l,y}\left(E[Y_k\mu_k^{(j)T}] - \overline{\mathcal{H}}_k L_k^{(lj)} - \mathcal{V}_{k,k-1}C_{l,y}^T\Pi_{k-1}^{(l)-1}\Pi_{k-1,k}^{(lj)}\right), \quad k \geq 2;$$
$$\Pi_1^{(lj)} = C_{l,y}E[Y_1\mu_1^{(j)T}].$$

Then, the expression for $\Pi_k^{(lj)}$ is deduced since (see derivation of (17))

$$E[Y_k\mu_k^{(j)T}] = \overline{\mathcal{H}}_k\mathcal{X}_k^{(j)} + E[\widetilde{\mathcal{H}}_k\mathcal{D}_k\widetilde{\mathcal{H}}_k^T]C_{j,y}^T + \mathcal{V}_k^{(j)}, \; k \geq 1;$$

expression for $\Pi_{k-1,k}^{(lj)}$ is derived by an analogous reasoning. □

Remark 3. Note that the expressions of Lemma 1 lead to $L_k^{(ll)} = 0$ and $L_{k,k-1}^{(ll)} = \mathcal{X}_{k,k-1}^{(l)}$, which is also immediate if we apply the OPL in the definition of these matrices. Then, it is clear that the expression of $\Pi_k^{(lj)}$ in Lemma 2 for $j = l$ reduces to that in (17), since $\Pi_{k-1,k}^{(ll)} = 0$.

Lemma 3. *The cross-covariance matrices between the intermediate filters,* $K_{k/k}^{(lj)} = E[\widehat{x}_{k/k}^{d(l)}\widehat{x}_{k/k}^{d(j)T}]$, $l, j = 1, \dots, m$, *are computed by*

$$\begin{aligned} K_{k/k}^{(lj)} &= K_{k/k-1}^{(lj)} + L_k^{(lj)}\Pi_k^{(j)-1}\mathcal{X}_k^{(j)T} + \mathcal{X}_k^{(l)}\Pi_k^{(l)-1}L_k^{(jl)T} \\ &\quad + \mathcal{X}_k^{(l)}\Pi_k^{(l)-1}\Pi_k^{(lj)}\Pi_k^{(j)-1}\mathcal{X}_k^{(j)T}, \quad k \geq 1; \\ K_{0/0}^{(lj)} &= \overline{x}_0\overline{x}_0^T, \end{aligned} \tag{25}$$

and those between the intermediate predictors, $K_{k/k-1}^{(lj)} = E\left[\widehat{x}_{k/k-1}^{d(l)}\widehat{x}_{k/k-1}^{d(j)T}\right]$, *are given by*

$$\begin{aligned} K_{k/k-1}^{(lj)} &= \overline{F}_{k-1}K_{k-1/k-1}^{(lj)}\overline{F}_{k-1}^T + \overline{F}_{k-1}\Big(L_{k-1}^{(lj)} \\ &\quad + \mathcal{X}_{k-1}^{(l)}\Pi_{k-1}^{(l)-1}\Pi_{k-1}^{(lj)}\Big)\Pi_{k-1}^{(j)-1}C_{j,y}\mathcal{W}_{k-1}^T \\ &\quad + \mathcal{W}_{k-1}C_{l,y}^T\Pi_{k-1}^{(l)-1}L_{k,k-1}^{(jl)T}, \quad k \geq 2; \\ K_{1/0}^{(lj)} &= \overline{F}_0\overline{x}_0\overline{x}_0^T\overline{F}_0^T. \end{aligned} \tag{26}$$

Proof. Expression (25) follows easily using (9) for the intermediate filters. Now, using (10) for $\widehat{x}_{k/k-1}^{d(l)}$, we obtain

$$K_{k/k-1}^{(lj)} = E[\widehat{x}_{k-1/k-1}^{d(l)}\widehat{x}_{k/k-1}^{d(j)T}] + \mathcal{W}_{k-1}C_{l,y}^T\Pi_{k-1}^{(l)-1}L_{k,k-1}^{(jl)T};$$

then, using again (10) for $\widehat{x}_{k/k-1}^{d(j)}$, and (9) to write $E[\widehat{x}_{k-1/k-1}^{d(l)}\mu_{k-1}^{(j)T}] = L_{k-1}^{(lj)} + \mathcal{X}_{k-1}^{(l)}\Pi_{k-1}^{(l)-1}\Pi_{k-1}^{(lj)}$, expression (26) is immediately obtained. □

4.2. Distributed filter

The following theorem provides the proposed distributed filtering estimators, $\widehat{x}^{(i)}_{k/k}$, and the corresponding error covariance matrices, $\Sigma^{(i)}_{k/k}$.

Theorem 2. *Let* $\widehat{X}_{k/k} = \left(\widehat{x}^{(1)T}_{k/k}, \dots, \widehat{x}^{(m)T}_{k/k}\right)^T$ *be the vector constituted by the intermediate filtering estimators calculated from the recursive algorithm in Theorem 1. Then, the distributed filter from sensor i is given by*

$$\begin{array}{l} \widehat{x}^{(i)}_{k/k} = \boldsymbol{\Xi}_{k/k} C^T_{i,x} \left(C_{i,x} \boldsymbol{K}_{k/k} C^T_{i,x}\right)^{-1} C_{i,x} \widehat{X}_{k/k}, \quad k \geq 1; \\ \widehat{x}^{(i)}_{0/0} = \overline{x}_0, \end{array}$$

where

$$\boldsymbol{\Xi}_{k/k} = \left(K^{(11)}_{k/k}, \dots, K^{(mm)}_{k/k}\right), \quad \boldsymbol{K}_{k/k} = \left(K^{(lj)}_{k/k}\right)_{l,j=1,\dots,m}$$

and $C_{i,x}$ *is the matrix obtained by removing the all-zero rows of*

$$Diag\left(c_{ij}\right)_{j=1,\dots,m} \otimes I_{n_x}.$$

The error covariance matrix of the distributed filter is computed by

$$\begin{array}{l} \Sigma^{(i)}_{k/k} = \mathcal{D}_k - \boldsymbol{\Xi}_{k/k} C^T_{i,x} \left(C_{i,x} \boldsymbol{K}_{k/k} C^T_{i,x}\right)^{-1} C_{i,x} \boldsymbol{\Xi}^T_{k/k}, \quad k \geq 1; \\ \Sigma^{(i)}_{0/0} = \Sigma_0. \end{array}$$

Proof. The proof is immediately derived from (24). □

Remark 4. Until now, we have proposed a new approach to designing the distributed fusion filter for networked systems with random parameter matrices and correlated noises. It is worth mentioning that the random parameter matrix $H^{(i)}_k$ in the measurement equation can model the measurement missing phenomenon induced by the intermittent sensor failures, which is actually an intermittent omission fault. Therefore, our approach is capable of handling the estimation problem when the intermittent omission faults occur in the sensors.

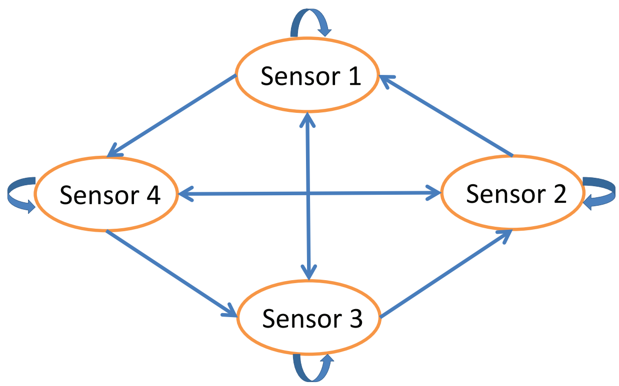

Figure 1: Topological structure of the sensor network.

5. Numerical Simulation Example

Consider the following discrete-time linear networked system with state-dependent multiplicative noise, and scalar measurements coming from four sensor nodes:

$$x_k = (0.95 + 0.2\epsilon_{k-1})x_{k-1} + w_{k-1}, \quad k \geq 1,$$
$$y_k^{(i)} = H_k^{(i)} x_k + v_k^{(i)}, \quad k \geq 1, \quad i = 1, 2, 3, 4,$$

where the initial state, x_0, is a standard Gaussian variable, and $\{\epsilon_k;\ k \geq 0\}$ is a zero-mean Gaussian white process with unit variance. The additive noises are defined as $w_k = 0.6(\eta_k + \eta_{k+1})$ and $v_k^{(i)} = c_k^{(i)}(\eta_{k-1} + \eta_k)$, $i = 1, 2, 3, 4$, where $c_k^{(1)} = 1$, $c_k^{(2)} = 0.5$, $c_k^{(3)} = 0.75$, $c_k^{(4)} = 0.85$, and $\{\eta_k;\ k \geq 0\}$ is a zero-mean Gaussian white process with variance 0.5.

Consider the sensor network displayed in Figure 1, represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ with set of nodes $\mathcal{V} = \{1, 2, 3, 4\}$, set of edges $\mathcal{E} = \{\ (1,1), (1,2), (1,3), (2,2), (2,3), (2,4), (3,1), (3,3), (3,4), (4,1), (4,2), (4,4)\ \}$ and adjacency matrix

$$\mathcal{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

At each sensor node, $i = 1, 2, 3, 4$, the random parameter matrices $H_k^{(i)}$ are defined to model different types of network-induced uncertainties: *continuous and discrete gain degradation* in sensors 1 and 2, respectively, *missing* *measurements* in sensor 3, and both *missing measurements and multiplicative noise* in sensor 4; specifically these matrices are defined as follows:

- $H_k^{(1)} = 0.82\lambda_k^{(1)}$, where $\{\lambda_k^{(1)}; k \geq 1\}$ is a sequence of independent random variables uniformly distributed over $[0.3, 0.7]$.

- $H_k^{(2)} = 0.75\lambda_k^{(2)}$, where $\{\lambda_k^{(2)}; k \geq 1\}$ is a sequence of independent discrete random variables with $P[\lambda_k^{(2)} = 0] = 0.1, \quad P[\lambda_k^{(2)} = 0.5] = 0.5, \quad P[\lambda_k^{(2)} = 1] = 0.4$.

- $H_k^{(3)} = 0.74\lambda_k^{(3)}$, where $\{\lambda_k^{(3)}; k \geq 1\}$ are independent Bernoulli variables with $P[\lambda_k^{(3)} = 1] = 0.7$, $\forall k \geq 1$.

- $H_k^{(4)} = \lambda_k^{(4)}(0.75 + 0.95\zeta_k)$, where $\{\lambda_k^{(4)}; k \geq 1\}$ are independent Bernoulli variables with $P[\lambda_k^{(4)} = 1] = 0.7$, $\forall k \geq 1$, and $\{\zeta_k; \ k \geq 1\}$ is a zero-mean Gaussian white process with unit variance.

Finally, according to the model hypotheses, the sequences $\{\epsilon_k; k \geq 0\}$, $\{\eta_k; k \geq 0\}$, $\{\lambda_k^{(i)}; k \geq 1\}$, $i = 1, 2, 3, 4$, and $\{\zeta_k; k \geq 1\}$ are assumed to be mutually independent.

To illustrate the feasibility and effectiveness of the proposed algorithms, they were implemented in MATLAB, and one hundred iterations of the algorithms were run. For $i = 1, 2, 3, 4$, Figure 2 shows the error variances of the following estimators:

- The local LS linear filter at sensor node i (obtained by using only measurements from the node i itself).

- The proposed intermediate filters at sensor nodes j within the communication neighborhood of node i ($j \in \mathcal{N}_i$).

- The proposed distributed filter at sensor node i.

From Figure 2 it is observed that, at each sensor node i, the error variances of the distributed filter are smaller than those of the intermediate filter, and the error variances of the intermediate filter are significantly less than those of the local filter. Hence, each sensor improves its local performance when the information from its neighbors is used and this is further improved by fusing intermediate filters from its neighborhood. Also it can be seen from Figure

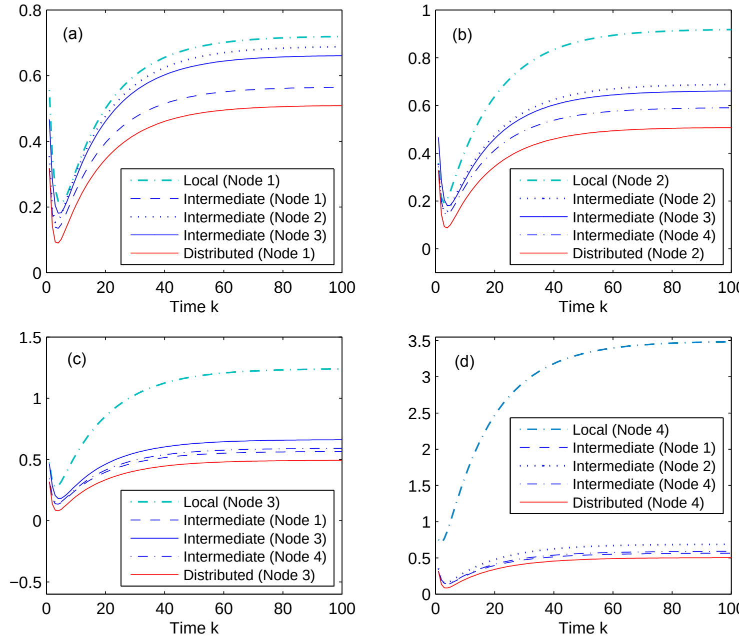

Figure 2: Error variance comparison of the local, intermediate and distributed filtering estimators in: (a) node 1, (b) node 2, (c) node 3 and (d) node 4.

2 that the proposed distributed estimator in a sensor node i outperforms all the intermediate filters in its neighborhood $\mathcal{N}_i$. In summary, it is shown that the proposed distributed estimation method has a satisfactory performance in connected sensor network systems where the measured outputs may be affected by different network-induced uncertainties.

Next, we analyze the disagreements of the proposed estimators among the different sensor nodes. It is obvious that, even when the same type (local, intermediate or distributed) of estimator is used, different measurements are processed at each sensor node and, consequently, the estimators at any two nodes may be different from each other. Clearly, a highly desirable property of an estimator is to reduce such disagreements among different sensor nodes. Figure 3 displays the error variances of the intermediate and distributed filters in the four nodes, as well as the centralized global optimal linear filter based on the measurements from all the network nodes. This figure shows that, in comparison to the intermediate filtering estimators, the proposed distributed filtering estimators reduce significantly the disagreements among

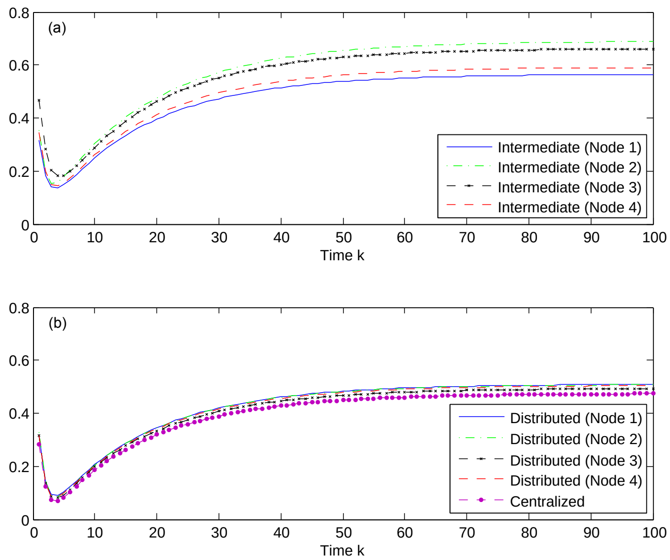

Figure 3: Error variance comparison of (a) intermediate filtering estimators and (b) distributed and centralized global optimal filtering estimators.

different nodes. Moreover, the closeness between the error variances of the global optimal filter and those of the proposed distributed filters show a highly accurate performance of the latest.

Finally, it is noted that analogous results are obtained when other probability distributions are assumed to model the network-induced uncertainties at the different sensor nodes. Moreover, concerning the missing measurement phenomena in nodes 3 and 4, it can be shown that, as expected, the filtering error variances become smaller and, hence, better estimations are obtained as the probability of missing observations decreases.

6. Conclusions

In this paper, the distributed filtering problem has been investigated in networked multi-sensor systems with random parameter matrices and correlated noises. The main outcomes and results can be summarized as follows:

- Using an innovation approach, a recursive LS linear estimation algorithm, very simple computationally and suitable for online applications, has been designed to obtain intermediate filtering estimators, which are calculated at each sensor node, using not only its own measurements, but also those from its neighboring sensors according to the network topology.

- Once the intermediate filtering estimators have been obtained, a new distributed estimator is designed at each sensor node as the LS matrix-weighted linear combination of its own intermediate estimator and those from its neighbors. The error covariance matrices of the distributed filters have been also derived.

- A numerical simulation example has shown the applicability of both intermediate and distributed estimators. This example has also highlighted the usefulness of the proposed algorithms for a great variety of sensor networked systems featuring different random network-induced uncertainties at the different sensor nodes, such as sensor gain degradation, missing measurements or multiplicative observation noises, which are covered by the observation model with random measurement matrices considered in this paper.

- A different approach to the estimation problem in this kind of systems with a given network topology, suggested by the anonymous reviewers, would be to consider the filter $\widehat{x}^{(i)}_{k-1/k-1}$ instead of $\widehat{x}^{d(i)}_{k-1/k-1}$ in equation (10), so the intermediate estimators would be based not only on the observations coming from the adjacent nodes, but also on the local estimators of these adjacent nodes. This new approach, as well as other different filter structures proposed in the existing literature (see [14] and references therein), might be interesting issues for further research.

Acknowledgments

This research is supported by *Ministerio de Economa y Competitividad* and *Fondo Europeo de Desarrollo Regional FEDER* (grant no. MTM2014-52291-P, MTM2017-84199-P).

References

[1] L. Yan, X. Li, R. Xia, M. Fu, Optimal sequential and distributed fusion for state estimation in cross-correlated noise, Automatica 49 (2013) 3607–3612.

[2] B. Chen, W. Zhang, L. Yu, Networked fusion Kalman filtering with multiple uncertainties, IEEE Trans. Aerosp. Electron. Syst. 51(3) (2015) 2332–2349.

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

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

[5] S. Sun, T. Tian, H. Lin, Optimal linear estimators for systems with finite-step correlated noises and packet dropout compensations, IEEE Trans. Signal Process. 64(21) (2016) 5672–5681.

[6] R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez, New distributed fusion filtering algorithm based on covariances over sensor networks with random packet dropouts, Int. J. Syst. Sci. 48(9) (2017) 1805–1817.

[7] J. Fernandez-Bes, L.A. Azpicueta-Ruiza, J. Arenas-García, M.T.M. Silva, Distributed estimation in diffusion networks using affine least-squares combiners, Digit. Signal Process. 36 (2015) 1-14.

[8] D.E. Marelli, M. Fu, Distributed weighted least-squares estimation with fast convergence for large-scale systems, Automatica 51 (2015) 27-39.

[9] J. Szurley, A. Bertrand, M. Moonen, Distributed adaptive node-specific signal estimation in heterogeneous and mixed-topology wireless sensor networks, Signal Process. 117 (2015) 44-60.

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

[11] G. Wang, N. Li, Y. Zhang, Diffusion distributed Kalman filter over sensor networks without exchanging raw measurements, Signal Process. 132 (2017) 1–7.

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

[13] S. Sun,H. Lin, J. Ma, X. Li, Multi-sensor distributed fusion estimation with applications in networked systems: A review paper, Inform. Fusion 38 (2017) 122–134.

[14] D. Ding, Z. Wang, B. Shen, Recent advances on distributed filtering for stochastic systems over sensor networks, Int. J. Gen. Syst. 43(3-4) (2014) 372–386.

[15] H. Dong, S.X. Ding, W. Ren, Distributed filtering with randomly occurring uncertainties over sensor networks: the channel fading case, Int. J. Gen. Syst. 43(3-4) (2014) 254–266.

[16] P. Millán L. Orihuela, C. Vivas, F.R. Rubio, Distributed consensus-based estimation considering network induced delays and dropouts, Automatica 48 (2012) 2726-2729.

[17] D. Ding, Z. Wang, D. W.C. Ho, G. Wei, Distributed recursive filtering for stochastic systems under uniform quantizations and deception attacks through sensor networks, Automatica 78 (2017) 231–240.

[18] Q. Liu, Z. Wang, X. He, G. Ghinea, F. E. Alsaadi, A resilient approach to distributed filter design for time-varying systems under stochastic nonlinearities and sensor degradation, IEEE Trans. Signal Process. 65(5) (2017), 1300–1309.

[19] S. Gao, P. Chen, Suboptimal filtering of networked discrete-time systems with random observation losses, Math. Probl. Eng. 2014 (2014) Article ID 151836.

[20] Y. Liu, Z. Wang, X. He, D. Zhou, Minimum-variance recursive filtering over sensor networks with stochastic sensor gain degradation: Algorithms and performance analysis, IEEE Trans. Control Netw. Syst. 3(3) (2016) 265–274.

[21] F. Peng, S. Sun, Distributed fusion estimation for multisensor multirate systems with stochastic observation multiplicative noises, Math. Probl. Eng. 2014 (2014) Article ID 373270.

[22] H. Lin, S. Sun, State estimation for a class of non-uniform sampling systems with missing measurements, Sensors (2016) 16, 1155.

[23] X. Liu, L. Li, Z. Li, H.C. Iu, T. Fernando, Stochastic stability of modifie d extende d Kalman filter over fading channels with transmission failure and signal fluctuation, Signal Processing 138 (2017) 220–232.

[24] J. Ma, S. Sun, Centralized fusion estimators for multi-sensor systems with multiplicative noises and missing measurements, Journal of Networks 7(10) (2012) 1538–1545.

[25] S. Sun, J. Ma, Linear estimation for networked control systems with random transmission delays and packet dropouts, Inf. Sci. 269 (2014) 349–365.

[26] J. Ma, S. Sun, Distributed fusion filter for networked stochastic uncertain systems with transmission delays and packet dropouts, Signal Process. 130 (2017), 268–278.

[27] D. Chen, Y. Yu, L. Xu, X. Liu, Kalman filtering for discrete stochastic systems with multiplicative noises and random two-step sensor delays, Discrete Dyn. Nat. Soc. 2015 (2015) Article ID 809734, 11 pages

[28] S. Wang, H. Fang, X. Tian, Recursive estimation for nonlinear stochastic systems with multi-step transmission delays, multiple packet dropouts and correlated noises, Signal Process. 115 (2015) 164–175.

[29] J. Feng, Z. Wang, M. Zeng, Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises, Inform. Fusion 14 (2013) 78-86.

[30] Y. Luo, Y. Zhu, D. Luo, J. Zhou, E. Song, D. Wang, Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications, Sensors 8 (12) (2008) 8086–8103.

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

[32] J. Linares-Pérez, R. Caballero-Águila, I. García-Garrido, Optimal linear filter design for systems with correlation in the measurement matrices

and noises: recursive algorithm and applications, Int. J. Syst. Sci. 45(7) (2014) 1548–1562.

[33] R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez, Optimal state estimation for networked systems with random parameter matrices, correlated noises and delayed measurements, Int. J. Gen. Syst. 44(2) (2015) 142–154.

[34] Y. Yang, Y. Liang, Q. Pan, Y. Qin, F. Yang, Distributed fusion estimation with square-root array implementation for Markovian jump linear systems with random parameter matrices and cross-correlated noises, Inf. Sci. 370–371 (2016) 446–462.

[35] S. Sun, T. Tian, H. Lin, State estimators for systems with random parameter matrices, stochastic nonlinearities, fading measurements and correlated noises, Inf. Sci. 397–398 (2017) 118–136.

[36] W. Li, Y. Jia, J. Du, Distributed filtering for discrete-time linear systems with fading measurements and time-correlated noise, Digit. Signal Process. 60 (2017) 211–219.