



---

*Research article*

## Unreliable networks with random parameter matrices and time-correlated noises: distributed estimation under deception attacks

Raquel Caballero-Águila<sup>1,\*</sup>, María J. García-Ligero<sup>2</sup>, Aurora Hermoso-Carazo<sup>2</sup> and Josefa Linares-Pérez<sup>2</sup>

<sup>1</sup> Departamento de Estadística e I.O., Universidad de Jaén, Campus Las Lagunillas, 23071 Jaén, Spain

<sup>2</sup> Departamento de Estadística e I.O., Universidad de Granada, Campus Fuentenueva, 18071 Granada, Spain

\* **Correspondence:** Email: raguila@ujaen.es; Tel: +34-953-212-926.

**Abstract:** This paper examines the distributed filtering and fixed-point smoothing problems for networked systems, considering random parameter matrices, time-correlated additive noises and random deception attacks. The proposed distributed estimation algorithms consist of two stages: the first stage creates intermediate estimators based on local and adjacent node measurements, while the second stage combines the intermediate estimators from neighboring sensors using least-squares matrix-weighted linear combinations. The major contributions and challenges lie in simultaneously considering various network-induced phenomena and providing a unified framework for systems with incomplete information. The algorithms are designed without specific structure assumptions and use a covariance-based estimation technique, which does not require knowledge of the evolution model of the signal being estimated. A numerical experiment demonstrates the applicability and effectiveness of the proposed algorithms, highlighting the impact of observation uncertainties and deception attacks on estimation accuracy.

**Keywords:** Networked systems; random parameter matrices; time-correlated additive noise; random deception attacks; distributed estimation

---

### 1. Introduction

Sensor networks have generated considerable research interest due to their wide range of practical applications and the rapid development of communication and information technologies (see, e.g., [1] and [2]). A widely-explored research topic in this domain is the *distributed estimation problem*, which arises in situations where the sensor nodes are spatially distributed according to a predetermined net-

work topology and each sensor node can use data information from itself and its neighboring nodes to estimate the signal of interest. Each sensor node then acts as a local fusion centre, combining its own information with that obtained from adjacent nodes to improve the performance of the local estimators, based solely on its own measurement information. This collaborative signal estimation strategy offers many advantages, including ease of implementation, robustness, scalability and high reliability. In [3], the distributed estimation problem is addressed in the presence of random packet dropouts during data transmission. A distributed event-based filtering structure is proposed in [4] for a class of sensor networks with sensor saturations and cyber-attacks. Nonlinear systems over sensor networks, whose topologies are changeable subject to Round-Robin protocol in the finite horizon case, are considered in [5]. A recursive distributed filtering algorithm for networked systems with random parameter matrices and correlated noises is designed in [6]. In recent years, a large variety of distributed estimation algorithms have been designed for sensor networks exposed to various challenges and vulnerabilities, under a range of filter schemes. For example, a new distributed filtering strategy has been presented in [7] by fully taking both the prediction estimations and its own measurement innovation into consideration. In [8], a novel distributed filtering compensation algorithm is presented in terms of the available transmitted data. A new design of an innovation-based stealthy attack strategy against distributed state estimation over a sensor network is proposed in [9], while the state estimation problem in linear time-invariant systems by using a network of distributed observers with switching communication topology is studied in [10]. Some significant achievements in the field of distributed estimation for stochastic systems over sensor networks are reviewed in [11] and [12].

In the fields of physics, electronics and engineering, many applications can involve *infinite-step colored measurement noises*, particularly when the sampling frequency is high enough to make the noises significantly correlated over two or more consecutive sampling periods. In recent years, researchers have addressed the estimation problem under the assumption that the measurements are affected by infinite-step time-correlated channel noise, modeled as the output of a linear system driven by white noise. Two popular methods for dealing with this type of noise correlation are state augmentation, which is simple and direct but computationally expensive, and measurement differencing, which avoids the problem of increasing dimensions but requires two consecutive measurements to compute the difference. By referencing to the measurement differencing method, in [13] the time-correlation of the measurement noises is transformed into the cross-correlation between the equivalent measurement noise and the process noise. Convergence conditions of the optimal linear estimator are obtained in [14], by using a new measurement obtained from measurement differencing. In [15], by using the time-differencing approach, the available measurements are transformed into an equivalent set of observations that do not depend on the time-correlated noise. Alternative non-augmentation and non-differencing methods to address the state estimation problem, based on the direct estimation of the time-correlated additive noise, are described in [16] and [17].

Communication networks are typically subject to resource limitations, which can provoke network-related issues during signal measurement or transmission [18]. Some such issues, such as the presence of multiplicative noise, missing observations or fading measurements, can be described by introducing *stochastic parameter matrices* into the measurement equations. In recent years, considerable research has been conducted into estimation problems arising in systems with random parameter matrices. For example, in [19] the Tobit Kalman filtering problem is studied for a class of linear discrete-time system with random parameters, where the elements of both the system matrix and the measurement matrix

are allowed to be random variables in order to reflect the reality. The distributed fusion estimation problem for networked systems whose multisensor measured outputs involve uncertainties modeled by random parameter matrices is investigated in [20]. In [21], the optimal linear filtering problem for linear discrete-time stochastic systems with random matrices in both the state and measurement equations is addressed. A Kalman-like recursive distributed optimal linear fusion predictor without feedback, for discrete-time linear stochastic systems with correlated random parameter matrices, is presented in [22]. In [23], the distributed fusion estimation problem is discussed in the presence of coupled noises, random delays and packet dropouts, for a class of uncertain systems, where the uncertainty in the measurement model is described by random parameter matrices.

In addition to the problem of random uncertainties in measurements and transmissions, a critical issue that cannot be ignored in any study of the estimation problem in networked systems is the possibility of suffering cyber-attacks. Security vulnerability is a common weakness that has been widely discussed in the literature; a bibliographical review of recent advances and challenges in this regard can be found in [24]. In particular, *deception attacks* have attracted significant research attention. This form of attack seeks to compromise data integrity by maliciously and randomly falsifying information. In this respect, [25] examined the centralized security-guaranteed filtering problem for linear time-invariant stochastic systems with multirate-sensor fusion subjected to deception attacks. In other approaches, the  $H_\infty$ -consensus filtering problem for discrete-time systems with multiplicative noises and deception attacks has been investigated by [26], and the distributed estimation problem in sensor networks with a specific topology structure under deception attacks has been addressed by [27, 28, 30]. More specifically, the distributed filtering problem in networked systems with fading measurements and multiplicative noises in both the signal and measurement equations is discussed in [27]. In [28], a positive system over a sensor network with simultaneous deception attacks and various network-induced constraints on sensor measurements is considered. The distributed secure state estimation problem is addressed in [29] for a class of general nonlinear systems over sensor networks under unknown deception attacks on innovations, while networked uncertain systems, containing uncertainties due to multiplicative and additive noises in the state and measurement equations, are considered in [30].

In view of the above considerations, the present study focuses on a class of networked systems, whose sensor nodes are distributed in space according to a fixed network topology. The main *study goal* is to address the least-squares linear distributed estimation problem from measurements affected by random parameter matrices and time-correlated additive noises, and simultaneously exposed to random deception attacks. For this purpose, a recursive algorithm for the distributed estimators is generated in two stages. First, each sensor node collects measurements from its neighbors to create intermediate least-squares linear estimators by an innovation approach. In the second stage, the intermediate estimators from neighboring sensors are combined to form distributed estimators through least-squares matrix-weighted linear combinations. A greater volume of information from different sensors is used in the second stage than in the first. This enhances the intermediate estimation performance and reduces disagreements among intermediate estimators from different sensors, by steering each distributed estimator closer to the global optimal linear estimator (hypothetically based on measurements from all network sensors).

This study makes the following *main contributions*. First, the consideration of random parameter matrices in the measurement equations provides a unified framework for handling common network-induced phenomena, such as multiplicative noise, missing observations and missing or fading mea-

surements; thus, the proposed algorithm can be applied to a wide range of network systems with incomplete information. The model also integrates stochastic deception attacks, to which many networked systems are vulnerable. Second, both the random parameter matrices and the random variables modeling the deception attacks are time-varying, making it possible to consider general situations involving time-dependent network-induced phenomena and different random phenomena at different sensor nodes. Third, the study considers the infinite-step time correlation of the measurement noises and facilitates the direct estimation of time-correlated additive noise, without relying on the differencing method. Fourth, the covariance-based estimation technique employed does not require knowledge of the evolution model for the signal being estimated. Furthermore, and unlike most previous studies of distributed estimation, which usually obtain optimal linear estimators based on a given structure, we present optimal linear-distributed estimation algorithms under the mean squared error criterion, without requiring a particular structure for the estimators. Another advantage of the proposed distributed estimation scheme is that, while most other approaches incorporate an upper bound for the estimation error covariance, this paper derives an exact expression for the error covariance, which can be calculated offline, regardless of the specific measurement set to be processed. Finally, in comparison with the most closely related authors' previous studies on distributed estimation, the current system model, in contrast to the one in [6], incorporates deception attacks and infinite-step time-correlated noise. The system model in [27], as the system considered in this work, includes stochastic deception attacks, but, in contrast to the present study, white additive noises are considered. Furthermore, it is worth noting that the derivation of the distributed filtering algorithm in both [6] and [27] relies on the state-space model equations, whereas the algorithms proposed in this paper do not require explicit information about the state transition equation. Instead, they rely solely on the factorization of the state covariance matrix in a separable form. The superiority of the proposed distributed filter over the one in [27] in the presence of infinite-step time-correlated additive noises will be experimentally tested in a numerical simulation example.

The rest of this paper is structured as follows. The networked system model with random parameter matrices, time-correlated noises and deception attacks is presented in Section 2, together with the assumptions required of the stochastic processes involved. The distributed estimation problem is then formulated in Section 3. After this, the distributed estimators are derived in two steps, described in Section 4 and Section 5, respectively. Finally, Section 6 provides an illustrative example highlighting the effectiveness of the proposed distributed estimation algorithms and the main conclusions drawn are summarized in Section 7.

### *Notation and abbreviations*

The mathematical notation and abbreviations used in this paper are detailed in the following table. If not explicitly stated, all vector and matrix dimensions are assumed to be compatible with algebraic operations.

$\mathbb{R}^n$	Set of $n$ -dimensional real vectors
$0$	Zero scalar or matrix of compatible dimension
$\mathbf{1}_{n \times n}$	$n \times n$ all-ones matrix
$I_n$	$n \times n$ identity matrix
$M^T, M^{-1}$ and $M^{-T}$	Transpose, inverse and transpose of the inverse of matrix $M$
$Diag(d_1, \dots, d_m)$	Diagonal matrix with entries $d_1, \dots, d_m$
$(M_1   \dots   M_k)$	Partitioned matrix whose blocks are the submatrices $M_1, \dots, M_k$
$G_k = G_{k,k}$	Function $G_{k,h}$ , depending on time instants $k$ and $h$ , when $h = k$
$M^{(i)} = M^{(ii)}$	Function $M^{(ij)}$ , depending on sensors $i$ and $j$ , when $j = i$
$\otimes$	Kronecker product of matrices
$\circ$	Hadamard product of matrices
$\delta_{k,h}$	Kronecker delta function
LS	Least-squares
OPL	Orthogonal projection lemma

## 2. Model description and assumptions

Consider a second-order  $n_x$ -dimensional discrete-time signal process  $\{x_k\}_{k \geq 1}$ , measured by different sensors which are spatially distributed according to a fixed network topology. Specifically, the sensor network is represented by a digraph  $\mathcal{G} = (\mathcal{N}, \mathcal{E}, \mathcal{D})$  of order  $m$ , where  $\mathcal{N} = \{1, \dots, m\}$  denotes the set of sensor nodes,  $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$  is the set of edges connecting different nodes and  $(j, i) \in \mathcal{E}$  means that the sensor node  $i$  receives the information from node  $j$ . These link relations among sensors are specified by the adjacency matrix,  $\mathcal{D} = (d_{ij})_{m \times m}$ , with  $d_{ij} = 1$  for  $(i, j) \in \mathcal{E}$  and  $d_{ij} = 0$  otherwise; since any sensor receives its own information, it is clear that  $d_{ii} = 1$ . For each node  $i \in \mathcal{N}$ , the set of adjacent nodes plus the node itself is denoted by  $\mathcal{N}_i = \{j \in \mathcal{N} : d_{ji} = 1\}$ ; therefore  $\mathcal{N}_i$ , which we term the neighborhood of node  $i$ , is the set of sensor nodes that transmit their information to node  $i$ .

The network sensors provide noisy measurements of the signal with multiplicative perturbations described by random parameter matrices, and the measurement noises in each sensor are assumed to be sequentially correlated. In our study context, deception attacks may be launched by potential adversaries to replace these measurements by deception noises before they are processed.

We now address the distributed filtering and smoothing problems at each node  $i \in \mathcal{N}$ , based on the available signal information in that node, i.e. the measurements derived from all its neighbor nodes  $j \in \mathcal{N}_i$ . The estimation is addressed under the *LS approach*, using only covariance information on the processes involved in the model. Therefore, assumptions (A1)–(A6), listed below, are set to guarantee the existence and knowledge of the first and second-order moments of the signal to be estimated and those of the observations on which the estimation is based. Regarding the signal process, the following assumption—which is key to the recursivity of the estimation algorithms—is required.

**(A1) (On the signal).** The signal process  $\{x_k\}_{k \geq 1}$  is a zero mean second-order process whose covariance function is expressed in a separable form as follows:

$$E[x_k x_l^T] = \mathcal{A}_k \mathcal{B}_l^T, \quad 1 \leq l \leq k,$$

where the factors  $\mathcal{A}_k, \mathcal{B}_k$  are  $n_x \times N$ -dimensional known matrices.

In the following, we describe the measurement model and specify the assumptions made for the random measurement matrices and additive noises.

### 2.1. Noisy measurements with random parameter matrices

As previously indicated, at any time  $k \geq 1$ , the signal  $x_k$  is measured by all the network sensors, which provide local outputs perturbed by random parameter matrices and time-correlated additive measurement noises. To describe this situation, consider the following measurement model:

$$\check{y}_k^{(i)} = C_k^{(i)} x_k + v_k^{(i)}, \quad k \geq 1; \quad i \in \mathcal{N}, \quad (2.1)$$

where  $\check{y}_k^{(i)} \in \mathbb{R}^{n_y}$  is the measurement of the signal provided by the  $i$ -th sensor at time  $k$ ,  $C_k^{(i)}$  is a random parameter matrix and  $v_k^{(i)}$  is the time-correlated measurement noise, which is assumed to be generated by a white process,  $\{\xi_k^{(i)}\}_{k \geq 0}$ , from an initial noise  $v_0^{(i)}$ :

$$v_k^{(i)} = H_{k-1}^{(i)} v_{k-1}^{(i)} + \xi_{k-1}^{(i)}, \quad k \geq 1; \quad i \in \mathcal{N}, \quad (2.2)$$

where  $H_k^{(i)}$  are non-singular known deterministic matrices.

To address the estimation problem using covariance information, the following assumptions about the first and second-order moments of the random parameter matrices and the measurement noises are required.

**(A2)** (On the random parameter matrices).  $\{C_k^{(i)}\}_{k \geq 1}$ ,  $i \in \mathcal{N}$ , with  $C_k^{(i)} = (c_{pq}^{(i)}(k))_{n_y \times n_x}$ , are independent sequences of independent random parameter matrices whose entries have known first and second-order moments. The mean matrices are denoted by  $\bar{C}_k^{(i)} = E[C_k^{(i)}]$ , and their  $(p, q)$ -entries are given by  $E[c_{pq}^{(i)}(k)]$ .

**Remark 1.** The existence of second-order moments guarantees that of  $E[C_k^{(i)} G C_k^{(j)T}]$ , for any random matrix  $G$  with mean  $\bar{G}$ . Moreover, if  $G$  is independent of the matrices  $C_k^{(i)}$  and  $C_k^{(j)}$ , the  $(p, q)$ -entry of this expectation is given by  $\sum_{a=1}^{n_x} \sum_{b=1}^{n_x} E[c_{pa}^{(i)}(k) c_{qb}^{(j)}(k)] \bar{G}_{ab}$ .

**(A3)** (On the measurement noises). The measurement noises  $\{v_k^{(i)}\}_{k \geq 1}$ ,  $i \in \mathcal{N}$ , are time-correlated sequences as described in (2.2), where:

- The initial vectors  $v_0^{(i)}$ ,  $i \in \mathcal{N}$ , have zero-mean and known cross-covariance matrices,  $\Sigma_0^{v^{(ij)}} = E[v_0^{(i)} v_0^{(j)T}]$ ,  $i, j \in \mathcal{N}$ .
- The white processes  $\{\xi_k^{(i)}\}_{k \geq 0}$ ,  $i \in \mathcal{N}$ , are assumed to be independent of each other at different times (consequently,  $E[\xi_k^{(i)} \xi_l^{(j)T}] = 0$ ,  $l \neq k$ ;  $i, j \in \mathcal{N}$ ). Their covariance and cross-covariance functions are denoted by  $\Sigma_k^{\xi^{(ij)}} = E[\xi_k^{(i)} \xi_k^{(j)T}]$ ,  $k \geq 0$ ;  $i, j \in \mathcal{N}$ .
- $\{\xi_k^{(i)}\}_{k \geq 0}$ ,  $i \in \mathcal{N}$ , are independent of the initial vectors  $v_0^{(i)}$ ,  $i \in \mathcal{N}$ , and, consequently,  $E[v_0^{(i)} \xi_k^{(j)T}] = 0$ ,  $k \geq 0$ ;  $i, j \in \mathcal{N}$ .

**Remark 2.** As a consequence of (A3), the measurement noises,  $\{v_k^{(i)}\}_{k \geq 1}$ ,  $i \in \mathcal{N}$ , are zero-mean second-order processes with cross-covariance matrices  $\Sigma_{k,l}^{v^{(i)}} = E[v_k^{(i)} v_l^{(j)T}] = H_{k-1}^{(i)} \cdots H_l^{(i)} \Sigma_l^{v^{(i)}}$ ,  $1 \leq l < k$ , and  $\Sigma_k^{v^{(i)}}$  is recursively obtained from the relation  $\Sigma_k^{v^{(i)}} = H_{k-1}^{(i)} \Sigma_{k-1}^{v^{(i)}} H_{k-1}^{(j)T} + \Sigma_{k-1}^{e^{(i)}}$ ,  $k \geq 1$ , with initial condition  $\Sigma_0^{v^{(i)}}$ .

## 2.2. Random deception attack model

In practice, sensor networks are often exposed to attacks from potential adversaries, seeking to modify or deteriorate the real measurements. In this situation, the observations to be processed for estimation may differ from the actual measurements, and so the mathematical model for the measurement outputs after the attacks must be specified.

In this paper, we assume that the measurements are subject to deception attacks that, if successful, will neutralize them and insert deceptive information. Such attacks may or may not be successful, and this uncertainty is incorporated into the model via Bernoulli random variables. Therefore, the mathematical model that describes the potentially attacked measurements to be processed in the signal estimation, which will be denoted by  $y_k^{(i)}$ , is formulated as:

$$y_k^{(i)} = \check{y}_k^{(i)} + \lambda_k^{(i)} \dot{y}_k^{(i)}, \quad k \geq 1; \quad i \in \mathcal{N}, \quad (2.3)$$

where  $\lambda_k^{(i)}$  are Bernoulli random variables modeling the randomness on the success ( $\lambda_k^{(i)} = 1$ ) or failure ( $\lambda_k^{(i)} = 0$ ) of the attacks, and  $\dot{y}_k^{(i)} = -\check{y}_k^{(i)} + w_k^{(i)}$  is the signal inserted by the attacker, neutralizing the actual measurement,  $\check{y}_k^{(i)}$ , and replacing it with a deceptive noise represented by  $w_k^{(i)}$ .

Therefore,  $\lambda_k^{(i)} = 0$  means that the attack against the  $i$ -th sensor at time  $k$  has failed and the measurement to be processed is the actual one ( $y_k^{(i)} = \check{y}_k^{(i)}$ ), while  $\lambda_k^{(i)} = 1$  means that the attack was successful and the processed measurement is the deceptive one ( $y_k^{(i)} = w_k^{(i)}$ ). An equivalent expression for the attacked measurement outputs (2.3) is given by:

$$y_k^{(i)} = (1 - \lambda_k^{(i)}) \check{y}_k^{(i)} + \lambda_k^{(i)} w_k^{(i)}, \quad k \geq 1; \quad i \in \mathcal{N}. \quad (2.4)$$

The following assumptions are made regarding the processes involved in these observation equations.

**(A4)** (*On the success of attacks*). The processes  $\{\lambda_k^{(i)}\}_{k \geq 1}$ ,  $i \in \mathcal{N}$ , are independent sequences of independent Bernoulli random variables with known success probabilities,  $P(\lambda_k^{(i)} = 1) = \bar{\lambda}_k^{(i)}$ . As a consequence, the first and second-order moments of these variables are:

$$E[\lambda_k^{(i)}] = E[(\lambda_k^{(i)})^2] = \bar{\lambda}_k^{(i)}, \quad k \geq 1; \quad E[\lambda_k^{(i)} \lambda_l^{(j)}] = \bar{\lambda}_k^{(i)} \bar{\lambda}_l^{(j)}, \quad l \neq k \text{ or } j \neq i; \quad l, k \geq 1; \quad i, j \in \mathcal{N}.$$

**(A5)** (*On the deception noises*). The noises inserted by successful attacks,  $\{w_k^{(i)}\}_{k \geq 1}$ ,  $i \in \mathcal{N}$ , consist of independent white processes with known covariance and cross-covariance matrices,  $\Sigma_k^{w^{(ij)}} = E[w_k^{(i)} w_k^{(j)T}]$ ,  $k \geq 1$ ;  $i, j \in \mathcal{N}$ .

**Remark 3.** In this paper, the attack probabilities of success and the covariances and cross-covariances of the noises of the attacks are assumed to be known. If they were unknown, they should

be identified before applying the proposed algorithms. To deal with this issue, a distributed self-tuning filtering algorithm is proposed in [30], based on the identification of unknown characteristics through the sample zero-order and first-order correlation functions of the observations.

Finally, the following assumption on the processes involved in the described model is required to address the estimation problem.

**(A6) (Mutual independence).** For each  $i \in \mathcal{N}$ , the signal process,  $\{x_k\}_{k \geq 1}$ , the random parameter matrices,  $\{C_k^{(i)}\}_{k \geq 1}$ , the measurement noise,  $\{v_k^{(i)}\}_{k \geq 1}$ , and the processes involved in the attacks,  $\{\lambda_k^{(i)}\}_{k \geq 1}$  and  $\{w_k^{(i)}\}_{k \geq 1}$ , are mutually independent.

### 3. Estimation problem formulation

In formulating the estimation problem, our aim is to use the distributed fusion method to obtain estimators of the signal at each sensor node,  $i \in \mathcal{N}$ , based on the information available at that node. In accordance with the network topology described in Section 2, this information comes not only from the sensor itself but also from all the neighboring ones that transmit their information to it. Therefore, in the  $i$ -th sensor, the distributed fusion estimator of the signal  $x_k$  based on the information up to time  $h$ , denoted by  $\widehat{x}_{k/h}^{D(i)}$ , is obtained by using the local information from the sensor together with that from the nodes in its neighborhood,  $\mathcal{N}_i = \{j \in \mathcal{N} : d_{ji} = 1\}$ .

The proposed distributed estimators are derived in two steps for each  $i \in \mathcal{N}$ : a) *derive the intermediate LS linear estimators*; b) *fuse the neighboring intermediate estimators*. In the first step, the LS linear estimator,  $\widehat{x}_{k/h}^{(i)}$ , is obtained from the potentially attacked observations (2.4) coming from all the sensor nodes  $j \in \mathcal{N}_i$ . Then, in the second step, the distributed estimator  $\widehat{x}_{k/h}^{D(i)}$  is determined as the LS matrix-weighted linear combination of the neighboring intermediate estimators,  $\widehat{x}_{k/h}^{(j)}$ ,  $j \in \mathcal{N}_i$ .

In order to unify the derivation of the intermediate estimators in all sensors, we jointly consider all the network output information at each sampling time  $k \geq 1$ , which is described by the gathered vector  $y_k = (y_k^{(1)T}, \dots, y_k^{(m)T})^T$ ; the outputs corresponding to the sensors in  $\mathcal{N}_i$  are then extracted from this vector to be processed in each sensor  $i \in \mathcal{N}$ . Therefore, the intermediate estimator  $\widehat{x}_{k/h}^{(i)}$ , for each  $i \in \mathcal{N}$ , is based on the measurements described by  $Y_l^{(i)} = D_y^{(i)} y_l$ ,  $l = 1, \dots, h$ , where  $D_y^{(i)}$  is the matrix obtained by removing the all-zero rows of  $\text{Diag}(d_{1i}, \dots, d_{mi}) \otimes I_{n_y}$ .

Next, we specify the gathered measurement model and the statistical properties of the processes involved—derived from our prior assumptions about the local measurements—.

#### 3.1. Gathered measurement model and properties

To describe the gathered observations to be processed in the estimation, the components of the local observation models in Section 2 are stacked as follows:

$$y_k = \begin{pmatrix} y_k^{(1)} \\ \vdots \\ y_k^{(m)} \end{pmatrix}, \quad \check{y}_k = \begin{pmatrix} \check{y}_k^{(1)} \\ \vdots \\ \check{y}_k^{(m)} \end{pmatrix}, \quad v_k = \begin{pmatrix} v_k^{(1)} \\ \vdots \\ v_k^{(m)} \end{pmatrix}, \quad w_k = \begin{pmatrix} w_k^{(1)} \\ \vdots \\ w_k^{(m)} \end{pmatrix},$$



$$C_k = \begin{pmatrix} C_k^{(1)} \\ \vdots \\ C_k^{(m)} \end{pmatrix}, \quad \Lambda_k = \begin{pmatrix} \lambda_k^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_k^{(m)} \end{pmatrix} \otimes I_{n_y}, \quad H_k = \begin{pmatrix} H_k^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & H_k^{(m)} \end{pmatrix}.$$

From this, the  $mn_y$ -dimensional observation vector  $y_k$ , which gathers all the sensor measurements at time  $k$ , is expressed by:

$$y_k = (I_{mn_y} - \Lambda_k)\check{y}_k + \Lambda_k w_k, \quad k \geq 1, \quad (3.1)$$

where  $\check{y}_k$  is the actual measurement vector gathered before the attacks, given by:

$$\check{y}_k = C_k x_k + v_k, \quad k \geq 1. \quad (3.2)$$

The following properties of the processes involved in (3.1)-(3.2) are directly obtained from the corresponding assumptions stated in Section 2.

- $\{C_k\}_{k \geq 1}$  is a sequence of independent second-order random parameter matrices with means  $\bar{C}_k = (\bar{C}_k^{(1)T} | \cdots | \bar{C}_k^{(m)T})^T$ . Moreover, for any first-order random matrix,  $G$ ,  $\exists E[C_k G C_k^T] = (E[C_k^{(i)} G C_k^{(j)T}])_{i,j \in \mathcal{N}}$ , with  $(i, j)$ -components as defined in Remark 1.
- The measurement noise,  $\{v_k\}_{k \geq 1}$ , is a zero-mean second-order time-correlated sequence. From Remark 2, the corresponding covariance matrices,  $\Sigma_{k,l}^v = E[v_k v_l^T]$ , are given by  $\Sigma_{k,l}^v = H_{k-1} \cdots H_l \Sigma_l^v$ ,  $1 \leq l < k$ , and  $\Sigma_k^v$  is recursively obtained from the relation  $\Sigma_k^v = H_{k-1} \Sigma_{k-1}^v H_{k-1}^T + \Sigma_{k-1}^\xi$ ,  $k \geq 1$ , with initial condition  $\Sigma_0^v = (\Sigma_0^{v(ij)})_{i,j \in \mathcal{N}}$  and  $\Sigma_k^\xi = (\Sigma_k^{\xi(ij)})_{i,j \in \mathcal{N}}$ .

**Remark 4.** The non-singularity of the matrices  $H_k$  allows us to factorize the noise covariance matrices in a similar way to those of the signal in (A1); namely:

$$\Sigma_{k,l}^v = \mathcal{H}_k \mathcal{F}_l^T, \quad 1 \leq l \leq k, \quad (3.3)$$

where  $\mathcal{H}_k = H_{k-1} \cdots H_0$  and  $\mathcal{F}_k = \Sigma_k^v \mathcal{H}_k^{-T}$ ,  $k \geq 1$ .

- $\{\Lambda_k\}_{k \geq 1}$  are diagonal independent matrices with means  $\bar{\Lambda}_k = \text{Diag}(\bar{\lambda}_k^{(1)}, \dots, \bar{\lambda}_k^{(m)}) \otimes I_{n_y}$ ,  $k \geq 1$ . For the purpose of further developments, if  $G$  is a random matrix independent of  $\{\Lambda_k\}_{k \geq 1}$ , these matrices operate as follows:

$$E[\Lambda_k G \Lambda_k] = K_k^\lambda \circ E[G], \quad E[(I_{mn_y} - \Lambda_k) G (I_{mn_y} - \Lambda_k)] = K_k^{1-\lambda} \circ E[G], \quad k \geq 1,$$

where  $K_k^\lambda = (E[\lambda_k^{(i)} \lambda_k^{(j)}])_{i,j \in \mathcal{N}} \otimes \mathbf{1}_{n_y \times n_y}$ ,  $K_k^{1-\lambda} = (E[(1 - \lambda_k^{(i)})(1 - \lambda_k^{(j)})])_{i,j \in \mathcal{N}} \otimes \mathbf{1}_{n_y \times n_y}$ , and the entries  $E[\lambda_k^{(i)} \lambda_k^{(j)}]$  are given in (A4).

- The deception noise,  $\{w_k\}_{k \geq 1}$ , is a white process whose covariance matrices,  $\Sigma_k^w = E[w_k w_k^T]$ , are obtained from (A5),  $\Sigma_k^w = (\Sigma_k^{w(ij)})_{i,j \in \mathcal{N}}$ ,  $k \geq 1$ .

Finally, assumption (A6) implies the independence of the processes involved in the gathered model (3.1)-(3.2):

- The signal process  $\{x_k\}_{k \geq 1}$ , the random parameter matrices,  $\{C_k\}_{k \geq 1}$ , the measurement noise,  $\{v_k\}_{k \geq 1}$ , and the processes  $\{\Lambda_k\}_{k \geq 1}$  and  $\{w_k\}_{k \geq 1}$ , are all mutually independent.

The above properties guarantee that the observations  $\{y_k\}_{k \geq 1}$  constitute a zero-mean second-order process, whose covariance matrices,  $\Sigma_k^y = E[y_k y_k^T]$ , are given by:

$$\Sigma_k^y = K_k^{1-\lambda} \circ \left( E \left[ C_k \mathcal{A}_k \mathcal{B}_k^T C_k^T \right] + \mathcal{H}_k \mathcal{F}_k^T \right) + K_k^\lambda \circ \Sigma_k^w, \quad k \geq 1. \quad (3.4)$$

### 3.2. Innovation approach to the LS estimation

In this paper, the intermediate LS linear estimation problem is addressed by an innovation approach, according to which the observation process at the  $i$ -th sensor,  $\{Y_k^{(i)}\}_{k \geq 1}$ , with  $Y_k^{(i)} = D_y^{(i)} y_k$ , is replaced by a white process  $\{\eta_k^{(i)}\}_{k \geq 1}$ —*innovation process*—, such that the LS linear estimator of an arbitrary zero-mean second-order vector  $a_k$  based on the observations  $\{Y_1^{(i)}, \dots, Y_h^{(i)}\}$  can be obtained by the following linear combination of the corresponding innovations,  $\{\eta_1^{(i)}, \dots, \eta_h^{(i)}\}$ :

$$\widehat{a}_{k/h}^{(i)} = \sum_{l=1}^h E \left[ a_k \eta_l^{(i)T} \right] \left( \Sigma_l^{\eta^{(i)}} \right)^{-1} \eta_l^{(i)}, \quad h \geq 1; \quad \widehat{a}_{k/0}^{(i)} = 0, \quad (3.5)$$

where  $\Sigma_k^{\eta^{(i)}} = E \left[ \eta_k^{(i)} \eta_k^{(i)T} \right]$  denotes the innovation covariance.

The innovation at time  $k$  is defined as  $\eta_k^{(i)} = Y_k^{(i)} - \widehat{Y}_{k/k-1}^{(i)}$ , where  $\widehat{Y}_{k/k-1}^{(i)}$  is the LS linear one-stage predictor of the observation  $Y_k^{(i)} = D_y^{(i)} y_k$ , which is clearly given by  $\widehat{Y}_{k/k-1}^{(i)} = D_y^{(i)} \widehat{y}_{k/k-1}^{(i)}$ . Hence, the innovations are written as:

$$\eta_k^{(i)} = D_y^{(i)} \left( y_k - \widehat{y}_{k/k-1}^{(i)} \right), \quad k \geq 1, \quad (3.6)$$

where, according to (3.5), expressions (3.1)-(3.2) and the independence properties on the model, the predictor of  $y_k$  is expressed in terms of those for the signal and noise as follows:

$$\widehat{y}_{k/k-1}^{(i)} = \left( I_{mny} - \overline{\Lambda}_k \right) \left( \overline{C}_k \widehat{x}_{k/k-1}^{(i)} + \widehat{v}_{k/k-1}^{(i)} \right), \quad k \geq 1. \quad (3.7)$$

Expressions (3.6) and (3.7) provide the basis for obtaining the intermediate estimation algorithms presented in the following section.

## 4. Intermediate LS linear estimation

In this section, the LS linear filtering and fixed-point smoothing estimators of the signal  $x_k$  are obtained in each sensor according to the observations available at that sensor. More precisely, for each  $i \in \mathcal{N}$ , we first derive a recursive algorithm for the intermediate LS linear filter,  $\widehat{x}_{k/k}^{(i)}$ , based on the observations  $Y_l^{(i)} = D_y^{(i)} y_l$ ,  $l = 1, \dots, k$ . The filtering estimator of  $x_k$  is then recursively updated with successive observations,  $Y_l^{(i)} = D_y^{(i)} y_l$ ,  $l = k + 1, \dots, h$ , in order to obtain the intermediate LS linear smoothing estimators,  $\widehat{x}_{k/h}^{(i)}$ ,  $h > k$ .

The intermediate filtering and fixed-point smoothing algorithms are presented in Theorems 1 and 2, respectively. Both theorems also include recursive formulas for the error covariance matrices  $\Sigma_{k/h}^{\bar{x}^{(i)}} = E \left[ \bar{x}_{k/h}^{(i)} \bar{x}_{k/h}^{(i)T} \right]$ , with  $\bar{x}_{k/h}^{(i)} = x_k - \hat{x}_{k/h}^{(i)}$ , thus providing the natural measure of the estimation accuracy under the LS criterion.

**Theorem 1.** For any  $i \in \mathcal{N}$ , the intermediate LS linear filtering estimators,  $\hat{x}_{k/k}^{(i)}$ , and the error covariance matrices,  $\Sigma_{k/k}^{\bar{x}^{(i)}}$ , are calculated as:

$$\hat{x}_{k/k}^{(i)} = (\mathcal{A}_k | 0) \mathbf{z}_k^{(i)}, \quad k \geq 1, \quad (4.1)$$

$$\Sigma_{k/k}^{\bar{x}^{(i)}} = \mathcal{A}_k \mathcal{B}_k^T - (\mathcal{A}_k | 0) \Sigma_k^{\mathbf{z}^{(i)}} (\mathcal{A}_k | 0)^T, \quad k \geq 1. \quad (4.2)$$

The vectors  $\mathbf{z}_k^{(i)}$  and their covariance matrices,  $\Sigma_k^{\mathbf{z}^{(i)}}$ , satisfy the following recursive formulas:

$$\mathbf{z}_k^{(i)} = \mathbf{z}_{k-1}^{(i)} + \mathcal{Z}_k^{(i)} (\Sigma_k^{\eta^{(i)}})^{-1} \eta_k^{(i)}, \quad k \geq 1; \quad \mathbf{z}_0^{(i)} = 0, \quad (4.3)$$

$$\Sigma_k^{\mathbf{z}^{(i)}} = \Sigma_{k-1}^{\mathbf{z}^{(i)}} + \mathcal{Z}_k^{(i)} (\Sigma_k^{\eta^{(i)}})^{-1} \mathcal{Z}_k^{(i)T}, \quad k \geq 1; \quad \Sigma_0^{\mathbf{z}^{(i)}} = 0, \quad (4.4)$$

where the gain coefficients,  $\mathcal{Z}_k^{(i)}$ , are:

$$\mathcal{Z}_k^{(i)} = \left( (\bar{\mathcal{C}}_k \mathcal{B}_k | \mathcal{F}_k) - (\bar{\mathcal{C}}_k \mathcal{A}_k | \mathcal{H}_k) \Sigma_{k-1}^{\mathbf{z}^{(i)}} \right)^T (I_{m_y} - \bar{\Lambda}_k) D_y^{(i)T}, \quad k \geq 1, \quad (4.5)$$

and the innovations,  $\eta_k^{(i)}$ , and their covariance matrices,  $\Sigma_k^{\eta^{(i)}}$ , are given by:

$$\eta_k^{(i)} = D_y^{(i)} \left( y_k - (I_{m_y} - \bar{\Lambda}_k) (\bar{\mathcal{C}}_k \mathcal{A}_k | \mathcal{H}_k) \mathbf{z}_{k-1}^{(i)} \right), \quad k \geq 1, \quad (4.6)$$

$$\Sigma_k^{\eta^{(i)}} = D_y^{(i)} \left( \Sigma_k^y - (I_{m_y} - \bar{\Lambda}_k) (\bar{\mathcal{C}}_k \mathcal{A}_k | \mathcal{H}_k) \Sigma_{k-1}^{\mathbf{z}^{(i)}} (\bar{\mathcal{C}}_k \mathcal{A}_k | \mathcal{H}_k)^T (I_{m_y} - \bar{\Lambda}_k) \right) D_y^{(i)T}, \quad k \geq 1. \quad (4.7)$$

The matrices  $\mathcal{H}_k$  and  $\mathcal{F}_k$  are the noise covariance factors defined in (3.3) and the observation covariance,  $\Sigma_k^y$ , is given in (3.4).

**Proof.** See Appendix A.

The key of the recursive algorithm in Theorem 1 is the intermediate filter factorization (4.1), in which the vectors  $\mathbf{z}_k^{(i)}$  can be recursively obtained as indicated in (4.3). Equation (4.2) provides the filtering error covariance matrices, which are needed to evaluate the performance of the filters; since these covariances do not depend of the observations, the filter performance can be measured before the observation set is available. In view of (4.1) and (4.2), the problem is focused on obtaining the vectors  $\mathbf{z}_k^{(i)}$  and their covariances,  $\Sigma_k^{\mathbf{z}^{(i)}}$ , which are derived from the innovations and their covariance matrices by the recursive equations (4.3) and (4.4), respectively. Specifically, by starting from  $\mathbf{z}_{k-1}^{(i)}$ ,  $\Sigma_{k-1}^{\mathbf{z}^{(i)}}$  and the new observation  $y_k^{(i)}$ , the gain, innovation and innovation covariance are calculated by equations (4.5), (4.6) and (4.7), respectively; then, equations (4.3) and (4.4) are used for obtaining  $\mathbf{z}_k^{(i)}$  and  $\Sigma_k^{\mathbf{z}^{(i)}}$ .

**Theorem 2.** For any  $i \in \mathcal{N}$ , the intermediate LS linear smoothing estimators,  $\hat{x}_{k/h}^{(i)}$ ,  $h > k$ , and the error covariance matrices,  $\Sigma_{k/h}^{\bar{x}^{(i)}}$ , are recursively calculated from the filter,  $\hat{x}_{k/k}^{(i)}$ , and the error covariance  $\Sigma_{k/k}^{\bar{x}^{(i)}}$ , respectively, given in Theorem 1:

$$\hat{x}_{k/h}^{(i)} = \hat{x}_{k/k}^{(i)} + \mathcal{X}_{k,h}^{(i)} (\Sigma_h^{\eta^{(i)}})^{-1} \eta_h^{(i)}, \quad h > k \geq 1, \quad (4.8)$$

$$\Sigma_{k/h}^{\bar{x}^{(i)}} = \Sigma_{k/h-1}^{\bar{x}^{(i)}} - \mathcal{X}_{k,h}^{(i)} \left( \Sigma_h^{\eta^{(i)}} \right)^{-1} \mathcal{X}_{k,h}^{(i)T}, \quad h > k \geq 1, \quad (4.9)$$

where:

$$\mathcal{X}_{k,h}^{(i)} = \left( (\mathcal{B}_k | 0) - \Sigma_{k,h-1}^{xz^{(i)}} \right) \left( \bar{\mathcal{C}}_h \mathcal{A}_h | \mathcal{H}_h \right)^T \left( I_{m_{ny}} - \bar{\Lambda}_h \right) D_y^{(i)T}, \quad h > k \geq 1, \quad (4.10)$$

and

$$\Sigma_{k,h}^{xz^{(i)}} = \Sigma_{k,h-1}^{xz^{(i)}} + \mathcal{X}_{k,h}^{(i)} \left( \Sigma_h^{\eta^{(i)}} \right)^{-1} \mathcal{Z}_h^{(i)T}, \quad h > k \geq 1; \quad \Sigma_k^{xz^{(i)}} = (\mathcal{A}_k | 0) \Sigma_k^z, \quad k \geq 1. \quad (4.11)$$

**Proof.** See Appendix B.

By starting from the filter, equation (4.8) explains how the estimators of  $x_k$  are updated when successive observations are available. Besides the innovations –which are given in Theorem 1–, for the update it is necessary to know the gain coefficients  $\mathcal{X}_{k,h}^{(i)}$ , which are obtained from equation (4.10) and require, in turn, the cross-covariance matrices  $\Sigma_{k,h}^{xz^{(i)}}$ , that are given by equation (4.11). Finally, equation (4.9) provides the smoothing error covariances in a recursive way.

**Remark 5.** By replacing  $D_y^{(i)}$  by the matrix obtained taking  $d_{ji} = 0$ ,  $j \neq i$ , the LS local linear filter and smoothers, based only on the measurements from the  $i$ -th sensor, can be obtained from the formulas in Theorems 1 and 2, respectively. Furthermore, the hypothetical global optimal linear estimators –based on measurements from all sensors– can be calculated by replacing  $D_y^{(i)}$  by the identity matrix.

## 5. Distributed fusion estimation

The second step of the distributed fusion method is performed in order to derive the distributed filter and fixed-point smoother of the signal,  $x_k$ . As indicated above, for each  $i \in \mathcal{N}$ , the distributed estimators,  $\widehat{x}_{k/h}^{D(i)}$ ,  $h \geq k$ , are derived as the matrix-weighted linear combination of the neighboring intermediate estimators,  $\widehat{x}_{k/h}^{(j)}$ ,  $j \in \mathcal{N}_i$ , using the mean squared error as an optimality criterion.

Specifically, the distributed estimators are expressed as  $\widehat{x}_{k/h}^{D(i)} = F_{k/h}^{(i)} \widehat{X}_{k/h}^{(i)}$ ,  $h \geq k$ , for each  $i \in \mathcal{N}$ , where  $\widehat{X}_{k/h}^{(i)}$  is the vector obtained by assembling all the intermediate estimators of the sensors connected to the  $i$ -th one, and where  $F_{k/h}^{(i)}$  is a matrix to be determined in order to minimize the mean squared error,  $E \left[ \left( x_k - F_{k/h}^{(i)} \widehat{X}_{k/h}^{(i)} \right)^T \left( x_k - F_{k/h}^{(i)} \widehat{X}_{k/h}^{(i)} \right) \right]$ .

As it is known, the solution to this problem is given by:

$$F_{k/h}^{(i)opt} = E \left[ x_k \widehat{X}_{k/h}^{(i)T} \right] \left( E \left[ \widehat{X}_{k/h}^{(i)} \widehat{X}_{k/h}^{(i)T} \right] \right)^{-1}$$

and, consequently, both matrices,  $E \left[ x_k \widehat{X}_{k/h}^{(i)T} \right]$  and  $E \left[ \widehat{X}_{k/h}^{(i)} \widehat{X}_{k/h}^{(i)T} \right]$ , must be determined for each  $i \in \mathcal{N}$ .

The entries of the above matrices are extracted from  $E \left[ x_k \widehat{X}_{k/h}^T \right]$  and  $E \left[ \widehat{X}_{k/h} \widehat{X}_{k/h}^T \right]$ , respectively, where  $\widehat{X}_{k/h} = \left( \widehat{x}_{k/h}^{(1)T}, \dots, \widehat{x}_{k/h}^{(m)T} \right)^T$  denotes the vector stacking all the intermediate estimators in the network. For this purpose, we write  $\widehat{X}_{k/h}^{(i)} = D_x^{(i)} \widehat{X}_{k/h}$ , where  $D_x^{(i)}$  is the matrix obtained by removing the all-zero rows

of  $\text{Diag}(d_{1i}, \dots, d_{mi}) \otimes I_{n_x}$ . The optimal distributed filter and the smoothers of the signal  $x_k$  are then given by:

$$\widehat{x}_{k/h}^{D(i)} = E \left[ x_k \widehat{X}_{k/h}^T \right] D_x^{(i)T} \left( D_x^{(i)} E \left[ \widehat{X}_{k/h} \widehat{X}_{k/h}^T \right] D_x^{(i)T} \right)^{-1} D_x^{(i)} \widehat{X}_{k/h}, \quad h \geq k \geq 1. \quad (5.1)$$

Hence, both matrices,  $E \left[ x_k \widehat{X}_{k/h}^T \right]$  and  $E \left[ \widehat{X}_{k/h} \widehat{X}_{k/h}^T \right]$  must be determined. Taking into account the OPL, each intermediate estimator  $\widehat{x}_{k/h}^{(r)}$  is uncorrelated with the estimation error; therefore, the entries of  $E \left[ x_k \widehat{X}_{k/h}^T \right]$  can be rewritten as  $E \left[ x_k \widehat{x}_{k/h}^{(r)T} \right] = E \left[ \widehat{x}_{k/h}^{(r)T} \widehat{x}_{k/h}^{(r)T} \right]$ ,  $r \in \mathcal{N}$  and then, denoting the cross-covariance between intermediate estimators as  $\Sigma_{k/h}^{\widehat{x}^{(rs)}} = E \left[ \widehat{x}_{k/h}^{(r)} \widehat{x}_{k/h}^{(s)T} \right]$ ,  $r, s \in \mathcal{N}$ , the matrices involved in (5.1) are given by  $E \left[ \widehat{X}_{k/h} \widehat{X}_{k/h}^T \right] = \left( \Sigma_{k/h}^{\widehat{x}^{(rs)}} \right)_{r,s \in \mathcal{N}}$  and  $E \left[ x_k \widehat{X}_{k/h}^T \right] = \left( \Sigma_{k/h}^{\widehat{x}^{(1)}} \mid \dots \mid \Sigma_{k/h}^{\widehat{x}^{(m)}} \right)$ .

The algorithms needed to obtain the cross-covariances between the intermediate filters and smoothers,  $\Sigma_{k/h}^{\widehat{x}^{(rs)}}$ ,  $h \geq k$ ,  $r, s \in \mathcal{N}$ , are stated in the following lemmas.

Lemma 1 presents the expression of the cross-covariances between the filters,  $\Sigma_{k/k}^{\widehat{x}^{(rs)}}$ ,  $r, s \in \mathcal{N}$ , which is directly obtained from (4.1). Thus, it depends on the cross-covariances  $\Sigma_k^{\mathbf{z}^{(rs)}} = E \left[ \mathbf{z}_k^{(r)} \mathbf{z}_k^{(s)T} \right]$ ; a recursive algorithm is then stated for  $\Sigma_k^{\mathbf{z}^{(rs)}}$ , which in turn depends on the innovation cross-covariance matrices,  $\Sigma_k^{\eta^{(rs)}} = E \left[ \eta_k^{(r)} \eta_k^{(s)T} \right]$ .

**Lemma 1.** For  $r, s \in \mathcal{N}$ , the cross-covariance matrices between the intermediate filters,  $\Sigma_{k/k}^{\widehat{x}^{(rs)}} = E \left[ \widehat{x}_{k/k}^{(r)} \widehat{x}_{k/k}^{(s)T} \right]$ , are computed by:

$$\Sigma_{k/k}^{\widehat{x}^{(rs)}} = (\mathcal{A}_k \mid 0) \Sigma_k^{\mathbf{z}^{(rs)}} (\mathcal{A}_k \mid 0)^T, \quad k \geq 1, \quad (5.2)$$

where  $\Sigma_k^{\mathbf{z}^{(rs)}}$  satisfies the following recursive relation:

$$\Sigma_k^{\mathbf{z}^{(rs)}} = \Sigma_{k-1}^{\mathbf{z}^{(rs)}} + \mathcal{Z}_{k-1,k}^{(rs)} \left( \Sigma_k^{\eta^{(s)}} \right)^{-1} \mathcal{Z}_k^{(s)T} + \mathcal{Z}_k^{(r)} \left( \Sigma_k^{\eta^{(r)}} \right)^{-1} \left( \mathcal{Z}_{k-1,k}^{(sr)T} + \Sigma_k^{\eta^{(rs)}} \left( \Sigma_k^{\eta^{(s)}} \right)^{-1} \mathcal{Z}_k^{(s)T} \right), \quad k \geq 1; \quad \Sigma_0^{\mathbf{z}^{(rs)}} = 0. \quad (5.3)$$

The matrices  $\mathcal{Z}_{k-1,k}^{(rs)} = E \left[ \mathbf{z}_{k-1}^{(r)} \eta_k^{(s)T} \right]$  are given by:

$$\mathcal{Z}_{k-1,k}^{(rs)} = \left( \Sigma_{k-1}^{\mathbf{z}^{(r)}} - \Sigma_{k-1}^{\mathbf{z}^{(rs)}} \right) \left( \overline{\mathcal{C}}_k \mathcal{A}_k \mid \mathcal{H}_k \right)^T \left( I_{m_y} - \overline{\Lambda}_k \right) D_y^{(s)T}, \quad k \geq 1, \quad (5.4)$$

and the innovation cross-covariance matrices,  $\Sigma_k^{\eta^{(rs)}}$ , satisfy:

$$\Sigma_k^{\eta^{(rs)}} = D_y^{(r)} \left( \Sigma_k^{\mathbf{y}} - \left( I_{m_y} - \overline{\Lambda}_k \right) \left( \overline{\mathcal{C}}_k \mathcal{A}_k \mid \mathcal{H}_k \right) \left( \Sigma_{k-1}^{\mathbf{z}^{(r)}} + \Sigma_{k-1}^{\mathbf{z}^{(s)}} - \Sigma_{k-1}^{\mathbf{z}^{(rs)}} \right) \left( \overline{\mathcal{C}}_k \mathcal{A}_k \mid \mathcal{H}_k \right) \left( I_{m_y} - \overline{\Lambda}_k \right) \right) D_y^{(s)T}, \quad k \geq 1. \quad (5.5)$$

**Proof.** See Appendix C.

In the second lemma, for any  $r, s \in \mathcal{N}$ , the cross-covariances between the intermediate smoothers,  $\Sigma_{k/h}^{\widehat{x}^{(rs)}}$ ,  $h > k$ , are obtained from that of the filter,  $\Sigma_{k/k}^{\widehat{x}^{(rs)}}$ , by a recursive algorithm.

**Lemma 2.** For  $r, s \in \mathcal{N}$ , the cross-covariance matrices between the intermediate fixed-point smoothers,  $\Sigma_{k/h}^{\widehat{x}^{(rs)}} = E \left[ \widehat{x}_{k/h}^{(r)} \widehat{x}_{k/h}^{(s)T} \right]$ ,  $h > k$ , are recursively obtained from:

$$\Sigma_{k/h}^{\widehat{x}^{(rs)}} = \Sigma_{k/h-1}^{\widehat{x}^{(rs)}} + \widehat{\mathcal{X}}_{k,h}^{(rs)} \left( \Sigma_h^{\eta^{(s)}} \right)^{-1} \mathcal{X}_{k,h}^{(s)T} + \mathcal{X}_{k,h}^{(r)} \left( \Sigma_h^{\eta^{(r)}} \right)^{-1} \left( \widehat{\mathcal{X}}_{k,h}^{(sr)T} + \Sigma_h^{\eta^{(rs)}} \left( \Sigma_h^{\eta^{(s)}} \right)^{-1} \mathcal{X}_{k,h}^{(s)T} \right), \quad h > k \geq 1, \quad (5.6)$$

with initial condition  $\widehat{\Sigma}_{k/k}^{(rs)}$ , given in Lemma 1.

The matrices  $\widehat{\mathcal{X}}_{k,h}^{(rs)} = E \left[ \widehat{x}_{k/h-1}^{(r)} \eta_h^{(s)T} \right]$  are derived as:

$$\widehat{\mathcal{X}}_{k,h}^{(rs)} = \left( \Sigma_{k,h-1}^{xz^{(r)}} - \Sigma_{k,h-1}^{xz^{(rs)}} \right) \left( \overline{C}_h \mathcal{A}_h \mid \mathcal{H}_h \right)^T \left( I_{m_y} - \overline{\Lambda}_h \right) D_y^{(s)T}, \quad h > k \geq 1, \quad (5.7)$$

where  $\Sigma_{k,h}^{xz^{(r)}}$  is given in Theorem 2 and  $\widehat{\Sigma}_{k,h}^{(rs)}$  is recursively computed by:

$$\widehat{\Sigma}_{k,h}^{(rs)} = \widehat{\Sigma}_{k,h-1}^{(rs)} + \widehat{\mathcal{X}}_{k,h}^{(rs)} \left( \Sigma_h^{\eta^{(s)}} \right)^{-1} \mathcal{Z}_h^{(s)T} + \mathcal{X}_{k,h}^{(r)} \left( \Sigma_h^{\eta^{(r)}} \right)^{-1} \left( \mathcal{Z}_{h-1,h}^{(sr)T} + \Sigma_h^{\eta^{(rs)}} \left( \Sigma_h^{\eta^{(s)}} \right)^{-1} \mathcal{Z}_h^{(s)T} \right), \quad h > k \geq 1, \quad (5.8)$$

with initial condition  $\widehat{\Sigma}_k^{(rs)} = (\mathcal{A}_k \mid 0) \Sigma_k^{z^{(rs)}}$ ,  $k \geq 1$ .

**Proof.** See Appendix D.

The above results for the distributed filters and fixed-point smoothers are presented in the following theorem, which also includes the formula to calculate their estimation error covariance matrices,  $\Sigma_{k/h}^{D(i)} = E \left[ \left( x_k - \widehat{x}_{k/h}^{D(i)} \right) \left( x_k - \widehat{x}_{k/h}^{D(i)} \right)^T \right]$ .

**Theorem 3.** For each  $i \in \mathcal{N}$ , the distributed filter and fixed-point smoothers,  $\widehat{x}_{k/h}^{D(i)}$ ,  $h \geq k$ , are given by:

$$\widehat{x}_{k/h}^{D(i)} = \left( \widehat{\Sigma}_{k/h}^{x^{(1)}} \mid \cdots \mid \widehat{\Sigma}_{k/h}^{x^{(m)}} \right) D_x^{(i)T} \left( D_x^{(i)} \left( \widehat{\Sigma}_{k/h}^{x^{(rs)}} \right)_{r,s \in \mathcal{N}} D_x^{(i)T} \right)^{-1} D_x^{(i)} \widehat{X}_{k/h}, \quad h \geq k \geq 1, \quad (5.9)$$

where  $\widehat{X}_{k/h}$  is the stacked vector of intermediate estimators and  $\widehat{\Sigma}_{k/h}^{x^{(rs)}}$  are the cross-covariances between them, given in Lemma 1 and Lemma 2 for the filter and smoothers, respectively.

The estimation error covariance matrices,  $\Sigma_{k/h}^{D(i)}$ , are given by:

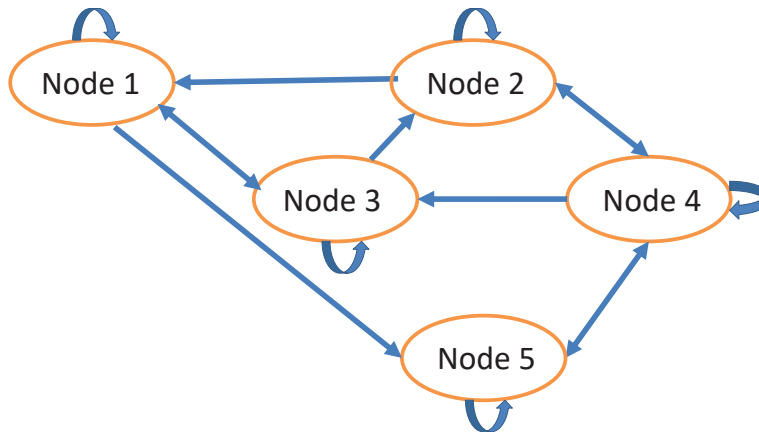
$$\Sigma_{k/h}^{D(i)} = \mathcal{A}_k \mathcal{B}_k^T - \left( \widehat{\Sigma}_{k/h}^{x^{(1)}} \mid \cdots \mid \widehat{\Sigma}_{k/h}^{x^{(m)}} \right) D_x^{(i)T} \left( D_x^{(i)} \left( \widehat{\Sigma}_{k/h}^{x^{(rs)}} \right)_{r,s \in \mathcal{N}} D_x^{(i)T} \right)^{-1} D_x^{(i)} \left( \widehat{\Sigma}_{k/h}^{x^{(1)}} \mid \cdots \mid \widehat{\Sigma}_{k/h}^{x^{(m)}} \right)^T, \quad h \geq k \geq 1. \quad (5.10)$$

**Proof.** Expression (5.9) is merely a rewrite of (5.1). To derive (5.10), the OPL must be used to express  $\Sigma_{k/h}^{D(i)} = E \left[ x_k x_k^T \right] - E \left[ \widehat{x}_{k/h}^{D(i)} \widehat{x}_{k/h}^{D(i)T} \right]$ , and the formula is straightforwardly achieved by using the signal covariance factorization, (A1), together with (5.9).  $\square$

## 6. Numerical simulation example

The applicability and effectiveness of the proposed distributed estimators are illustrated by the sensor network shown in Figure 1, where the topology is represented by a digraph  $\mathcal{G} = (\mathcal{N}, \mathcal{E}, \mathcal{D})$ , with the set of nodes  $\mathcal{N} = \{1, 2, 3, 4, 5\}$ . The elements in the adjacency matrix,  $\mathcal{D}$ , are  $d_{ji} = 1$  when sensor  $i$  can obtain information from sensor  $j$ ; otherwise,  $d_{ji} = 0$ :

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



**Figure 1.** Topological structure of the sensor network with five nodes.

## 6.1. Signal process and measurement model

### 6.1.1. Signal process

Consider a two-dimensional stochastic signal,  $x_k = \begin{pmatrix} x_{1k} \\ x_{2k} \end{pmatrix}$ , whose first and second component,  $x_{1k}$  and  $x_{2k}$ , denote the position and velocity of a target, respectively. Moreover, assume that the signal evolution is the same as in [30]; namely:

$$x_{k+1} = \left( A + \sum_{n=1}^2 \alpha_k^n A_n \right) x_k + \left( B + \sum_{n=1}^2 \beta_k^n B_n \right) \mu_k, \quad k \geq 0,$$

where

$$A = \begin{pmatrix} 0.95 & 0.01 \\ 0 & 0.95 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.01 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.02 \end{pmatrix}; \quad B = \begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The initial signal,  $x_0$ , is a two-dimensional standard Gaussian vector. The multiplicative noises  $\{\alpha_k^n\}_{k \geq 0}$ ,  $\{\beta_k^n\}_{k \geq 0}$ ,  $n = 1, 2$ , are white Gaussian scalar processes, with variances  $\text{Var}(\alpha_k^n) = 0.16$ ,  $\text{Var}(\beta_k^n) = 0.11$ ,  $n = 1, 2$ . The additive noise  $\{\mu_k\}_{k \geq 0}$  is a zero-mean Gaussian scalar process with variance 0.5. All of these noise sequences and the initial signal are assumed to be mutually independent. The signal covariance function can then be expressed in a separable form as  $E[x_k x_l^T] = \mathcal{A}_k \mathcal{B}_l^T$ ,  $l \leq k$ , with  $\mathcal{A}_k = A^k$  and  $\mathcal{B}_l^T = A^{-l} \Sigma_l^x$ , where  $\Sigma_l^x = E[x_l x_l^T]$  is recursively obtained by:

$$\Sigma_l^x = A \Sigma_{l-1}^x A^T + 0.16 \sum_{n=1}^2 A_n \Sigma_{l-1}^x A_n^T + 0.5 (B B^T + 0.11 \sum_{n=1}^2 B_n B_n^T), \quad l \geq 1; \quad \Sigma_0^x = I_2.$$

### 6.1.2. Sensor measurements with multiplicative perturbations and time-correlated noises

Scalar measurements of the signal, simultaneously disturbed by fading effects and both additive and multiplicative noises, are provided by the five sensor nodes, according to model (2.1), where the

processes involved have the following characteristics:

- The multiplicative perturbations are described by  $C_k^{(i)} = \theta_k^{(i)}(\dot{C}^{(i)} + \rho_k^{(i)}\ddot{C}^{(i)})$ ,  $i \in \mathcal{N}$ , where:
  - $\dot{C}^{(1)} = (0.8, 0.9)$ ,  $\dot{C}^{(2)} = (0.9, 0.7)$ ,  $\dot{C}^{(3)} = (0.6, 0.7)$ ,  $\dot{C}^{(4)} = (0.7, 0.8)$ ,  $\dot{C}^{(5)} = (0.9, 0.5)$ ,  
 $\ddot{C}^{(i)} = (1, 0)$ ,  $i = 1, 2$ , and  $\ddot{C}^{(i)} = (0, 1)$ ,  $i = 3, 4, 5$ .
  - $\{\theta_k^{(i)}\}_{k \geq 1}$  and  $\{\rho_k^{(i)}\}_{k \geq 1}$ ,  $i \in \mathcal{N}$ , are independent sequences of independent and identically distributed random variables; namely:
    - \*  $\theta_k^{(1)}$  and  $\theta_k^{(2)}$  are random variables uniformly distributed over  $[0.3, 0.7]$  and  $[0.2, 0.8]$ , respectively, representing a continuous random fading effect on the measurements from sensors 1 and 2.
    - \*  $\theta_k^{(3)}$  and  $\theta_k^{(4)}$  are discrete random variables, representing a discrete random fading effect on the measurements from sensors 3 and 4, according to the following probability mass functions:

$$P(\theta_k^{(3)} = 0) = 0.1, \quad P(\theta_k^{(3)} = 0.5) = 0.5, \quad P(\theta_k^{(3)} = 1) = 0.4.$$

$$P(\theta_k^{(4)} = 0.2) = 0.2, \quad P(\theta_k^{(4)} = 0.5) = 0.5, \quad P(\theta_k^{(4)} = 0.8) = 0.3.$$

- \*  $\theta_k^{(5)}$  are Bernoulli random variables, where  $P(\theta_k^{(5)} = 1) = \bar{\theta}$ , represents the randomly missing measurement phenomenon in sensor 5.
  - \* The multiplicative components  $\rho_k^{(i)}$ ,  $i \in \mathcal{N}$ , are standard Gaussian variables.
- The time-correlated noises  $\{v_k^{(i)}\}_{k \geq 1}$ ,  $i \in \mathcal{N}$ , are defined by (2.2), with  $H^{(1)} = H^{(3)} = H^{(5)} = 0.8$  and  $H^{(2)} = H^{(4)} = 0.7$ . These noises are generated by the processes  $\{\xi_k^{(i)}\}_{k \geq 0}$ , defined by  $\xi_k^{(i)} = a^{(i)}\xi_k$ ,  $i \in \mathcal{N}$ , where  $\{\xi_k\}_{k \geq 0}$  is a standard Gaussian white process,  $a^{(1)} = a^{(3)} = a^{(5)} = 0.5$ ,  $a^{(2)} = a^{(4)} = 0.25$ , and the initial conditions are  $v_0^{(i)} = v_0$ , for  $i \in \mathcal{N}$ , and where  $v_0$  is a standard Gaussian variable. Under these assumptions, the covariance matrices of the stacked noise  $v_k = (v_k^{(1)}, \dots, v_k^{(5)})^T$  can always be factorized as  $\Sigma_{k,l}^v = \mathcal{H}_k \mathcal{F}_l^T$ ,  $l \leq k$ , where  $\mathcal{H}_k = H^k$  and  $\mathcal{F}_l^T = H^{-l} \Sigma_l^v$ , being  $H = \text{Diag}(0.8, 0.7, 0.8, 0.7, 0.8)$  and where  $\Sigma_l^v$  is recursively obtained by  $\Sigma_l^v = H \Sigma_{l-1}^v H^T + \Sigma_{l-1}^\xi$ ,  $l \geq 1$ , with  $\Sigma_l^\xi = (a^{(i)} a^{(j)})_{i,j \in \mathcal{N}}$  and initial condition  $\Sigma_0^v = \mathbf{1}_{5 \times 5}$ .

### 6.1.3. Sensor measurements subject to deception attacks

In accordance with the theoretical model, let us suppose that the measurements at each sensor are subject to deception attacks and that the attacked measurement outputs are given by (2.4), where:

- The noises of the false data injection attacks are defined as  $w_k^{(i)} = b^{(i)} w_k$ , for  $i \in \mathcal{N}$ , where  $b^{(1)} = b^{(2)} = b^{(5)} = 0.5$ ,  $b^{(3)} = b^{(4)} = 0.75$ , and  $\{w_k\}_{k \geq 1}$  is a standard Gaussian white process. Clearly, these attack noises are correlated and  $\Sigma_k^{w^{(ij)}} = b^{(i)} b^{(j)}$ ,  $i, j \in \mathcal{N}$ .
- The status of the attacks is described by mutually independent sequences of independent and identically distributed Bernoulli random variables,  $\{\lambda_k^{(i)}\}_{k \geq 1}$ ,  $i \in \mathcal{N}$ , with known probabilities  $P(\lambda_k^{(i)} = 1) = \bar{\lambda}$ .

## 6.2. Applicability of the proposed distributed algorithms

To illustrate the effectiveness of the proposed distributed filtering and fixed-point smoothing algorithms and to quantify the estimation accuracy obtained, the estimation error variances of the first and



second signal components (position and velocity) were calculated at every sensor node  $i \in \mathcal{N}$ . First, the local, intermediate, distributed and global estimators were compared, for fixed values of the probabilities  $\bar{\theta}$  (probability that the signal is present in the measured outputs of sensor 5) and  $\bar{\lambda}$  (probability of a successful attack). Different values of the probabilities  $\bar{\theta}$  and  $\bar{\lambda}$  were then considered to highlight the effects of the missing measurement and attack phenomena, respectively, on the performance of the proposed distributed estimators, analyzing how these probabilities influence the estimation error variances for both signal components.

### 6.2.1. Comparison of local, intermediate and distributed estimators

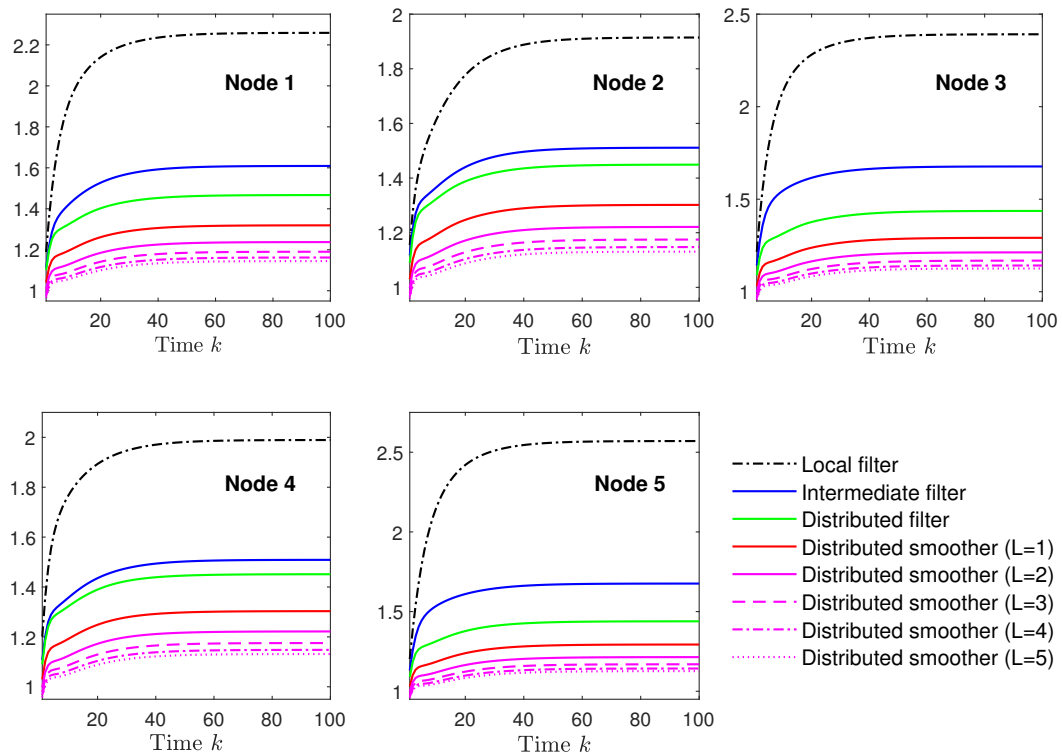
Considering the same value of 0.5 for the probabilities  $\bar{\theta}$  and  $\bar{\lambda}$ , Figure 2 depicts –for the first signal component– the error variances of the local filters (obtained using only the measurements from the sensor itself) and those of the proposed intermediate filters,  $\hat{x}_{k/k}^{(i)}$ , and distributed filters,  $\hat{x}_{k/k}^{D(i)}$ , and smoothers,  $\hat{x}_{k/k+L}^{D(i)}$  (with lag  $L = 1, 2, 3, 4, 5$ ), at every sensor node  $i \in \mathcal{N}$ . On the one hand, this figure shows that the error variances corresponding to the intermediate filters are significantly less than those of the local filters and that the distributed filters outperform the intermediate ones. In addition, it is apparent that the distributed smoothing error variances are less than the filtering ones and, also, that at each fixed-point  $k$ , the fixed-point smoothers become more accurate as the number of available observations,  $k + L$ , increases. As expected, the improvement is smaller as  $L$  increases; indeed, in this example, the improvement is practically imperceptible for  $L \geq 5$ . Similar results are obtained for the second signal component.

### 6.2.2. Comparison of intermediate, distributed and global optimal linear estimators

Close inspection of Figure 2 reveals little difference in the values of the distributed estimation error variances over the five nodes. For a better evaluation of these differences, Figure 3 shows, for the first and second signal components, the error variances of the intermediate filters together with those of the distributed filters and smoothers (for  $L = 1, 2$ ) at the different nodes, as well as those of the global optimal linear filtering and smoothing estimators, based on the set of measurements obtained from the five nodes of the network. A desirable property for distributed estimators over sensor networks is that the discrepancies between different nodes should be as small as possible; indeed, as we can see in Figure 3, the proposed distributed estimators considerably reduce the disagreements among intermediate estimators from different sensors. Thus, the distance between the error variances of the distributed estimators at the different nodes is fairly small and close to the global optimal error variances. In consequence, not only do the proposed distributed estimators present only slight discrepancies among the sensors, but they also provide a very similar level of performance to that of the global optimal estimators. Moreover, the proximity between the error variances of the global optimal filtering and smoothing estimators and those of the corresponding proposed distributed estimators shows that the latter estimators perform well.

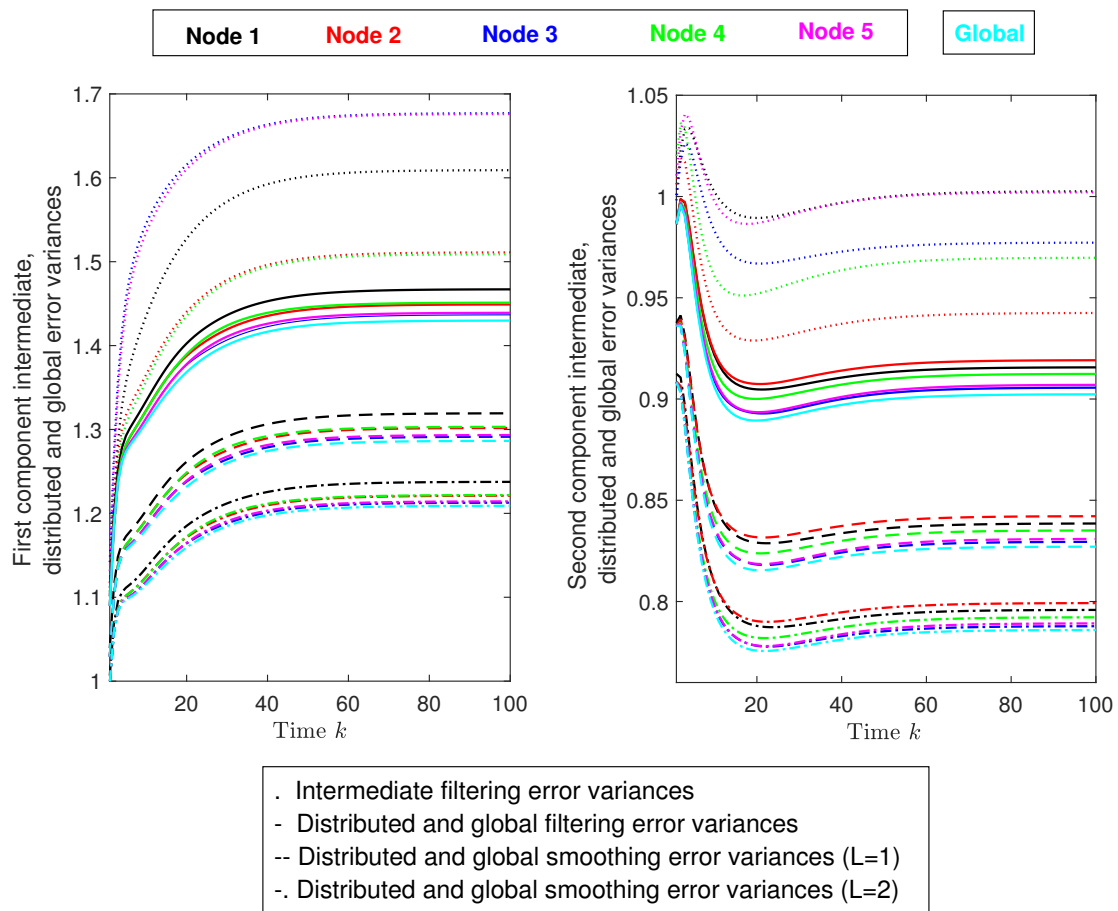
### 6.2.3. Effect of the probability of missing measurements on estimation accuracy

Assuming, as above, that the attack probability is  $\bar{\lambda} = 0.5$ , we now evaluate the performance of the proposed distributed estimators with respect to the missing measurements phenomenon in sensor node 5. To do so, the distributed filtering and smoothing ( $L = 2$ ) error variances for the first signal



**Figure 2.** Error variance comparison of the local and intermediate filters, distributed filters and smoothers, for the first component of the signal vector.

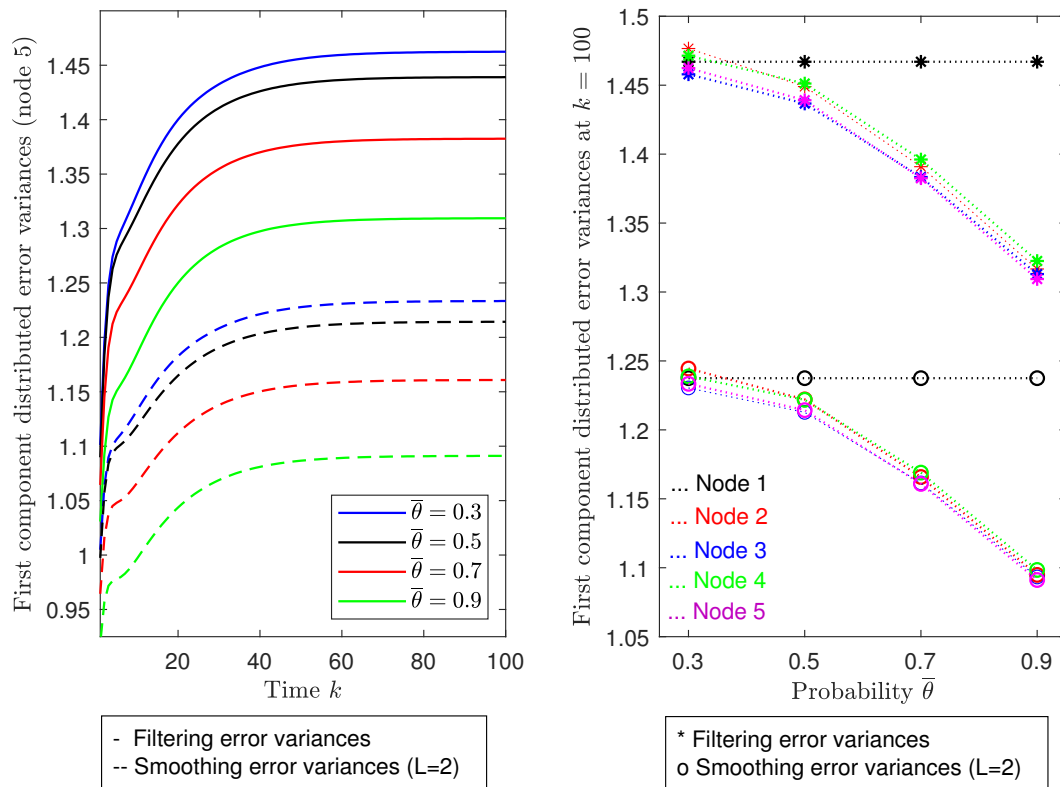
component are plotted in Figure 4 for different values of the probability  $\bar{\theta}$  (namely,  $\bar{\theta} = 0.3, 0.5, 0.7$  and  $0.9$ ). For these values, the distributed estimation error variances at sensor node 5 are plotted in the left-side panel of Figure 4. This figure shows that the performance of the distributed estimators is indeed influenced by these probabilities and, as expected, that the distributed estimation error variances decrease as the probability  $\bar{\theta}$  increases. Hence, both the filtering and the smoothing distributed estimators achieve better estimation accuracy when  $1 - \bar{\theta}$ , the probability of missing measurements, decreases, as this means that further information about the signal is available. Analogous results were obtained for sensor nodes 2, 3 and 4, since these nodes all utilize information from node 5; in fact, node 4 uses the measurements from node 5 to obtain its intermediate estimators, which are subsequently sent to nodes 2 and 3 to construct the distributed estimators at these nodes. In node 1, however, the distributed error variances do not change since the estimators in this node do not use the measurements from node 5. Since the behavior of the distributed error variances is analogous in all the iterations, for a better visualization of the decreasing trend of the error variances as the probability  $\bar{\theta}$  increases at all sensor nodes –except sensor 1, in which it remains constant–, the right-side panel of Figure 4 displays the distributed error variance at the iteration  $k = 100$ . As in the previous figures, Figure 4 also shows that the error variances corresponding to the smoothers are less than those of the filters. Similar results are obtained for the second signal component and therefore the same conclusions are drawn.



**Figure 3.** Error variance comparison of intermediate, distributed and global estimators for the first and second signal components.

#### 6.2.4. Effect of the probability of a successful attack on estimation accuracy

Next, our aim is to examine the association between the probability of a successful attack and the performance of the estimators. For this purpose, we compare the distributed filtering error variances for different values of this probability (namely,  $\bar{\lambda} = 0.1$  to  $0.9$ ). The distributed filtering error variances at sensor node 1, corresponding to the second signal component, are plotted in the left-side panel of Figure 5. Here, as expected, these error variances rise in line with  $\bar{\lambda}$ . Furthermore, this increase is more pronounced for higher values of  $\bar{\lambda}$ . Similar results are obtained in all nodes, as shown in the right-side panel of Figure 5, which displays the distributed filtering and smoothing error variances at  $k = 100$  versus  $\bar{\lambda}$  in the five sensor nodes. The discrepancies between the different nodes are negligible and, as shown in all the other figures, the smoother with lag  $L = 2$  outperforms the one with  $L = 1$  which, in turn, outperforms the filter. Similar results are inferred for the first signal component, and therefore the same conclusions are drawn.



**Figure 4.** The left-side (resp. right-side) panel depicts the distributed estimation error variances for  $\bar{\theta} = 0.3, 0.5, 0.7, 0.9$  in node 5 (resp. in all nodes at  $k = 100$ ).

### 6.2.5. Comparative analysis with the distributed filter [27]

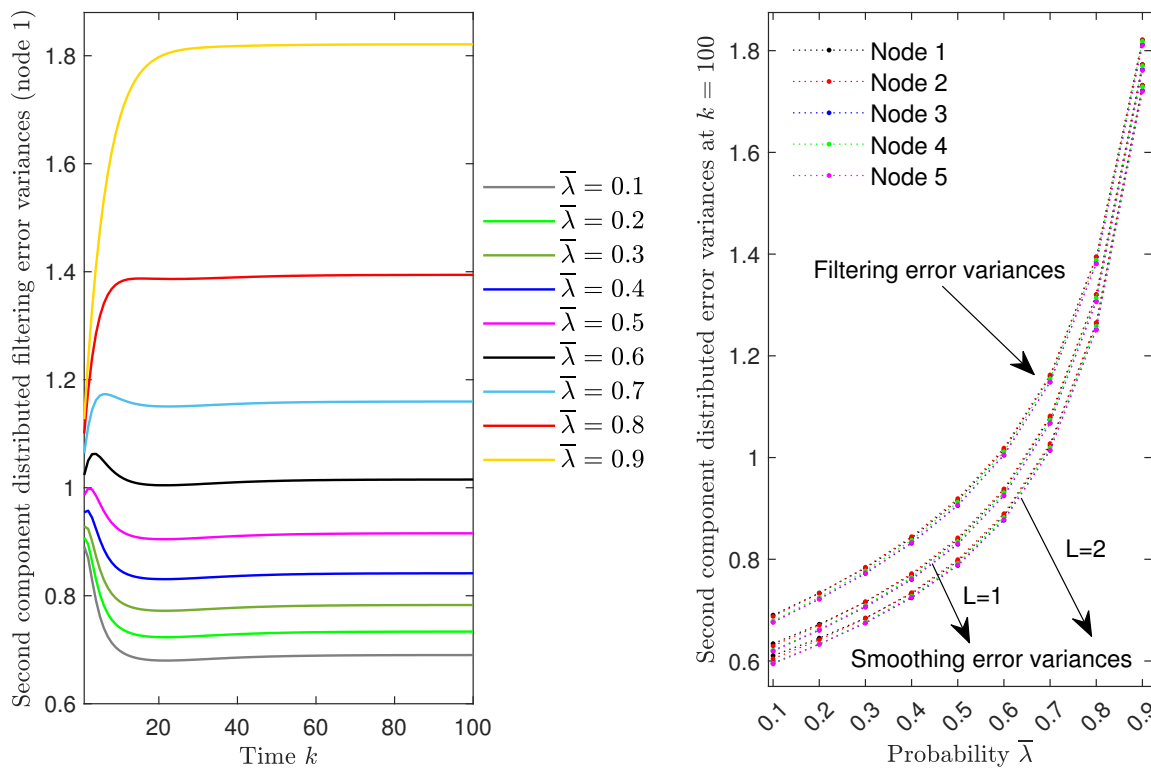
Our final aim in this section is to show the superior performance of the proposed distributed filter in the presence of infinite-step time-correlated additive noises. For this purpose, we conduct a comparative analysis between the distributed filter proposed in this paper and the one proposed in [27] for networked systems with fading measurements, multiplicative noises in both the signal and measurements equations and stochastic deception attacks, but without infinite-step time correlation of the measurement noises.

At every sensor node and for each of the two components of the distributed filtering estimates, the comparison is made on the basis of the empirical values of the mean-squared error at each time instant, which are calculated from two thousand independent simulations by

$$\text{MSE}_{a,k}^{(i)} = \frac{1}{2000} \sum_{s=1}^{2000} (x_{a,k}^{(s)} - \hat{x}_{a,k/k}^{D(i,s)})^2, \quad 1 \leq k \leq 100, i \in \mathcal{N}, a = 1, 2,$$

where, for each sampling time  $k$  and for the  $s$ -th simulation run,  $x_{a,k}^{(s)}$  denotes the  $a$ -th component of the simulated signal, and  $\hat{x}_{a,k/k}^{D(i,s)}$  is the  $a$ -th component of the distributed filter calculated in the  $i$ -th sensor node.

Assuming again the same fixed value 0.5 for the probabilities  $\bar{\theta}$  and  $\bar{\lambda}$ , the results are displayed

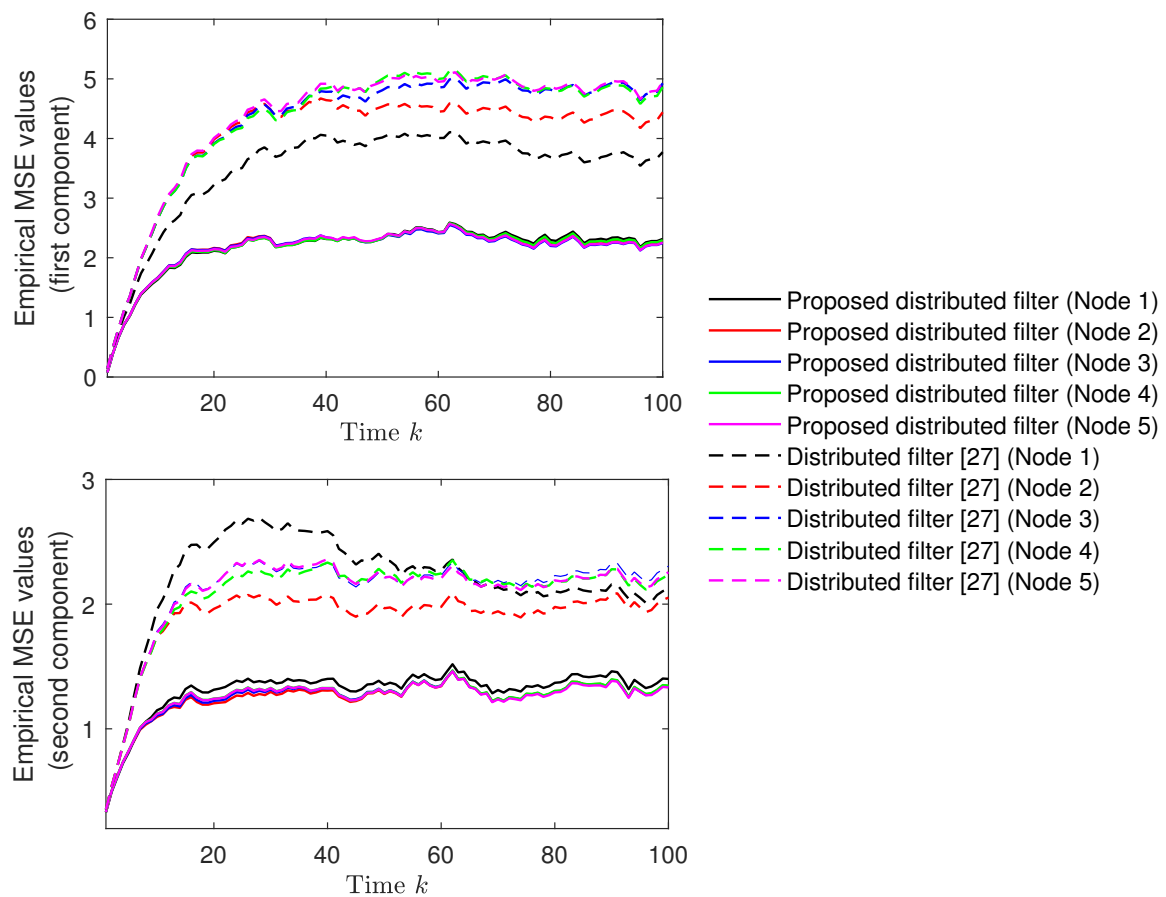


**Figure 5.** The left-side (resp. right-side) panel depicts the distributed filtering (resp. filtering and smoothing) error variances for  $\bar{\lambda} = 0.1$  to  $0.9$  in node 1 (resp. in all nodes at  $k = 100$ ).

in Figure 6, which shows that for all the sensor nodes and for both the first and second components, the empirical mean-squared error values of the proposed distributed filtering estimates are lower than those of the distributed filtering estimates in [27]. Note that the proposed filter was indeed expected to outperform the filter in [27], since the latter does not take into account the infinite-step time correlation of the sensor measurement noises.

## 7. Conclusions

In this paper, we investigate the distributed estimation problem –including filtering and fixed-point smoothing– in networked systems whose sensor nodes are spatially distributed according to a predetermined network topology, represented by a directed graph. Random parameter matrices and stochastic deception attacks are incorporated into the measurement model. Thus, a broad theoretical framework is provided with which to address general stochastic multi-sensor systems with different network-induced uncertainties. The presence of time-correlated additive noise in the observation model is handled by a non-augmentation method, based on the direct estimation of the noise. For every sensor node, the proposed distributed estimation algorithm runs in two phases. The first yields an intermediate least-squares linear estimator using its own local measurements and those received from its neighboring nodes. In the second phase, the own-sensor intermediate estimator is combined with those calculated



**Figure 6.** Empirical mean-squared error comparison of the proposed distributed filter and the distributed filter in [27].

by its neighboring nodes to obtain the desired distributed estimator as the minimum mean squared error matrix-weighted linear combination of the intermediate estimators. The proposed estimation strategy does not rely on explicit information provided by the signal evolution equation, but rather on the factorization of the signal and time-correlated noise covariance matrices in a separable form. As a result, whether or not the signal evolution model is completely known, the proposed distributed estimation technique can be used to estimate a wide class of stochastic signals, including those whose evolution is affected by multiplicative noises. The simulation experiment performed shows that the theoretical system model we present covers common random imperfections, such as the presence of multiplicative noise, missing observations and fading effects. The numerical results obtained were used to examine the influence of two degrading effects on estimation performance: a) the probability of missing measurements; b) the probability of successful attacks. Comparative analysis of the estimation error variances shows that the proposed distributed estimators outperform the intermediate ones and reduce the disagreements between different sensors by bringing each distributed estimator closer to the global optimal linear estimator, based on the full set of measurements of the entire network. Finally, in the

presence of infinite-step colored measurement noises, the proposed estimators are shown to outperform the distributed estimators in the authors' previous work [27].

A challenging topic for future studies is the derivation of self-tuning estimation algorithms for the case when the attack probabilities and/or the covariances and cross-covariances of the attack noises are unknown. It would also be interesting to consider the possibility that attack noise is not stochastic but a constant or time-varying deterministic sequence, as well as the scenario of random packet dropouts in the transmissions among sensor nodes.

### Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

### Acknowledgments

This research was supported by the “Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación” of Spain and the European Regional Development Fund [grant number PID2021-124486NB-I00].

### Conflict of interest

The authors declare there is no conflict of interest.

### References

1. U. Singh, A. Abraham, A. Kaklauskas, T. Hong, *Smart Sensor Networks. Analytics, Sharing and Control*, Springer, Switzerland, 2022. <https://doi.org/10.1007/978-3-030-77214-7>
2. Z. Zhou, H. Xu, H. Feng, W. Li, A Non-Equal Time Interval Incremental Motion Prediction Method for Maritime Autonomous Surface Ships, *Sensors*, **23** (2023), 2852. <https://doi.org/10.3390/s23052852>
3. 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. <https://doi.org/10.1016/j.inffus.2016.06.008>
4. J. Liu, Y. Gu, J. Cao, S. Fei, Distributed event-triggered  $H_\infty$  filtering over sensor networks with sensor saturations and cyber-attacks, *ISA Trans.*, **81** (2018), 63–75. <https://doi.org/10.1016/j.isatra.2018.07.018>
5. X. Bu, H. Dong, F. Han, N. Hou, G. Li, Distributed filtering for time-varying systems over sensor networks with randomly switching topologies under the round-robin protocol, *Neurocomputing*, **346** (2019), 58–64. <https://doi.org/10.1016/j.neucom.2018.07.087>
6. R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez, Z. Wang, A new approach to distributed fusion filtering for networked systems with random parameter matrices and correlated noises, *Inform. Fusion*, **45** (2019), 324–332. <https://doi.org/10.1016/j.inffus.2018.02.006>

7. J. Hu, Z. Wang, G.-P. Liu, H. Zhang, R. Navaratne, A prediction-based approach to distributed filtering with missing measurements and communication delays through sensor networks, *IEEE Trans. Syst. Man Cybern. -Syst.*, **51** (2021), 7063–7074. <https://doi.org/10.1109/TSMC.2020.2966977>
8. J. Li, J. Hu, J. Cheng, Y. Wei, H. Yu, Distributed filtering for time-varying state-saturated systems with packet disorders: An event-triggered case, *Appl. Math. Comput.*, **434** (2022), 127411. <https://doi.org/10.1016/j.amc.2022.127411>
9. M. Niu, G. Wen, Y. Lv, G. Chen, Innovation-based stealthy attack against distributed state estimation over sensor networks, *Automatica*, **152** (2023), 110962. <https://doi.org/10.1016/j.automatica.2023.110962>
10. G. Yang, H. Rezaee, A. Alessandri, T. Parisini, State estimation using a network of distributed observers with switching communication topology, *Automatica*, **147** (2023), 110690. <https://doi.org/10.1016/j.automatica.2022.110690>
11. 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. <https://doi.org/10.1016/j.inffus.2016.01.001>
12. 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. <https://doi.org/10.1016/j.inffus.2017.03.006>
13. H. Geng, Z. Wang, Y. Cheng, F. Alsaadi, A. M. Dobaie, State estimation under non-Gaussian Lévy and time-correlated additive sensor noises: A modified Tobit Kalman filtering approach, *Signal Process.*, **154** (2019), 120–128. <https://doi.org/10.1016/j.sigpro.2018.08.005>
14. W. Liu, P. Shi, Convergence of optimal linear estimator with multiplicative and time-correlated additive measurement noises, *IEEE Trans. Autom. Control*, **64** (2019), 2190–2197. <https://doi.org/10.1109/TAC.2018.2869467>
15. R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez, Networked fusion estimation with multiple uncertainties and time-correlated channel noise, *Inform. Fusion*, **54** (2020), 161–171. <https://doi.org/10.1016/j.inffus.2019.07.008>
16. J. Ma, S. Sun, Optimal linear recursive estimators for stochastic uncertain systems with time-correlated additive noises and packet dropout compensations, *Signal Process.*, **176** (2020), 107704. <https://doi.org/10.1016/j.sigpro.2020.107704>
17. R. Caballero-Águila, J. Hu, J. Linares-Pérez, Two Compensation Strategies for Optimal Estimation in Sensor Networks with Random Matrices, Time-Correlated Noises, Deception Attacks and Packet Losses, *Sensors*, **22** (2022), 8505. <https://doi.org/10.3390/s22218505>
18. Q. Liu, Z. Wang, X. He, *Stochastic Control and Filtering over Constrained Communication Networks*, Springer, Switzerland, 2019. <https://doi.org/10.1007/978-3-030-00157-5>
19. F. Han, H. Dong, Z. Wang, G. Li, F. E. Alsaadi, Improved Tobit Kalman filtering for systems with random parameters via conditional expectation, *Signal Process.*, **147** (2018), 35–45. <http://dx.doi.org/10.1016/j.sigpro.2018.01.015>



20. R. Caballero-Águila, A. Hermoso-Caraz, J. oLinares-Pérez, Centralized filtering and smoothing algorithms from outputs with random parameter matrices transmitted through uncertain communication channels, *Digit. Signal Process.*, **85** (2019), 77–85. <https://doi.org/10.1016/j.dsp.2018.11.010>
21. W. Liu, X. Xie, W. Qian, X. Xu, Y. Shi, Optimal linear filtering for networked control systems with random matrices, correlated noises, and packet dropouts, *IEEE Access*, **8** (2020), 59987–59997. <http://dx.doi.org/10.1109/ACCESS.2020.2983122>
22. S. Sun, Distributed optimal linear fusion predictors and filters for systems with random parameter matrices and correlated noises, *IEEE Trans. Signal Process.*, **68** (2020), 1064–1074. <https://doi.org/10.1109/TSP.2020.2967180>
23. R. Caballero-Águila, J. Linares-Pérez, Distributed fusion filtering for uncertain systems with coupled noises, random delays and packet loss prediction compensation, *Int. J. Syst. Sci.*, **54** (2023), 371–390. <https://doi.org/10.1080/00207721.2022.2122905>
24. M. S. Mahmoud, M. M. Hamdan, U. A. Baroudi, Modeling and control of Cyber-Physical Systems subject to cyber attacks: A survey of recent advances and challenges, *Neurocomputing*, **338** (2019), 101–115. <https://doi.org/10.1016/j.neucom.2019.01.099>
25. Z. Wang, D. Wang, B. Shen, F. E. Alsaadi, Centralized security-guaranteed filtering in multirate-sensor fusion under deception attacks, *J. Frankl. Inst.*, **355** (2018), 406–420. <https://doi.org/10.1016/j.jfranklin.2017.11.010>
26. F. Han, H. Dong, Z. Wang, G. Li, Local design of distributed  $H_\infty$ -consensus filtering over sensor networks under multiplicative noises and deception attacks, *Int. J. Robust Nonlinear Control*, **29** (2019), 2296–2314. <https://doi.org/10.1002/rnc.4493>
27. R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez, A two-phase distributed filtering algorithm for networked uncertain systems with fading measurements under deception attacks, *Sensors*, **20** (2020), 6445. <https://doi.org/10.3390/s20226445>
28. S. Xiao, Q. Han, X. Ge, Y. Zhang, Secure distributed finite-time filtering for positive systems over sensor networks under deception attacks, *IEEE Trans. Cybern.*, **50** (2020), 1200–1228. <https://doi.org/10.1109/tcyb.2019.2900478>
29. L. Ma, Z. Wang, Y. Chen, X. Yi, Probability-guaranteed distributed secure estimation for nonlinear systems over sensor networks under deception attacks on innovations, *IEEE Trans. Signal Inf. Proc. Netw.*, **7** (2021), 465–477. <https://doi.org/10.1109/TSIPN.2021.3097217>
30. Y. Ma, S. Sun, Distributed Optimal and Self-Tuning Filters Based on Compressed Data for Networked Stochastic Uncertain Systems with Deception Attacks, *Sensors*, **23** (2023), 335. <https://doi.org/10.3390/s23010335>

## Appendix A. Proof of Theorem 1

According to expression (3.5) for the LS linear estimators, the coefficients  $\mathcal{X}_{k,l}^{(i)} = E[x_k \eta_l^{(i)T}]$ ,  $1 \leq l \leq k$ , with  $\eta_l^{(i)}$  given from (3.6) and (3.7), must be calculated for the signal filter  $\hat{x}_{k/k}^{(i)}$ . For this purpose, we use expressions (3.1) and (3.2) for  $y_l$  which, together with (A1) on the signal covariance and the independence properties, easily lead to:

$$\mathcal{X}_{k,l}^{(i)} = \left( \mathcal{A}_k \mathcal{B}_l^T \bar{\mathcal{C}}_l^T - E \left[ x_k \widehat{x}_{l/l-1}^{(i)T} \right] \bar{\mathcal{C}}_l^T - E \left[ x_k \widehat{v}_{l/l-1}^{(i)T} \right] \right) (I_{m_{n_y}} - \bar{\Lambda}_l) D_y^{(i)T}, \quad 1 \leq l \leq k.$$

Now, by denoting  $\mathcal{V}_{k,l}^{(i)} = E \left[ v_k \eta_l^{(i)T} \right]$  and using (3.5) for the one-stage predictors  $\widehat{x}_{l/l-1}^{(i)T}$  and  $\widehat{v}_{l/l-1}^{(i)T}$ , we obtain:

$$E \left[ x_k \widehat{x}_{l/l-1}^{(i)T} \right] = (1 - \delta_{l,1}) \sum_{l'=1}^{l-1} \mathcal{X}_{k,l'}^{(i)} (\Sigma_{l'}^{\eta^{(i)}})^{-1} \mathcal{X}_{l,l'}^{(i)T}; \quad E \left[ x_k \widehat{v}_{l/l-1}^{(i)T} \right] = (1 - \delta_{l,1}) \sum_{l'=1}^{l-1} \mathcal{X}_{k,l'}^{(i)} (\Sigma_{l'}^{\eta^{(i)}})^{-1} \mathcal{V}_{l,l'}^{(i)T}, \quad l \geq 1,$$

and  $\mathcal{X}_{k,l}^{(i)}$  is then written as:

$$\mathcal{X}_{k,l}^{(i)} = \left( \mathcal{A}_k \mathcal{B}_l^T \bar{\mathcal{C}}_l^T - (1 - \delta_{l,1}) \sum_{l'=1}^{l-1} \mathcal{X}_{k,l'}^{(i)} (\Sigma_{l'}^{\eta^{(i)}})^{-1} (\bar{\mathcal{C}}_l \mathcal{X}_{l,l'}^{(i)} + \mathcal{V}_{l,l'}^{(i)})^T \right) (I_{m_{n_y}} - \bar{\Lambda}_l) D_y^{(i)T}, \quad 1 \leq l \leq k.$$

In a similar way, using now the factorization (3.3) for the noise covariance, the coefficients  $\mathcal{V}_{k,l}^{(i)}$  are also written as:

$$\mathcal{V}_{k,l}^{(i)} = \left( \mathcal{H}_k \mathcal{F}_l^T - (1 - \delta_{l,1}) \sum_{l'=1}^{l-1} \mathcal{V}_{k,l'}^{(i)} (\Sigma_{l'}^{\eta^{(i)}})^{-1} (\bar{\mathcal{C}}_l \mathcal{X}_{l,l'}^{(i)} + \mathcal{V}_{l,l'}^{(i)})^T \right) (I_{m_{n_y}} - \bar{\Lambda}_l) D_y^{(i)T}, \quad 1 \leq l \leq k.$$

These expressions guarantee the following factorizations:

$$\mathcal{X}_{k,l}^{(i)} = \mathcal{A}_k \mathcal{Z}_l^{x(i)}; \quad \mathcal{V}_{k,l}^{(i)} = \mathcal{H}_k \mathcal{Z}_l^{v(i)}, \quad 1 \leq l \leq k, \quad (7.1)$$

where

$$\mathcal{Z}_l^{x(i)} = \left( \mathcal{B}_l^T \bar{\mathcal{C}}_l^T - (1 - \delta_{l,1}) \sum_{l'=1}^{l-1} \mathcal{Z}_{l'}^{x(i)} (\Sigma_{l'}^{\eta^{(i)}})^{-1} (\bar{\mathcal{C}}_l \mathcal{A}_l \mathcal{Z}_{l'}^{x(i)} + \mathcal{H}_l \mathcal{Z}_{l'}^{v(i)})^T \right) (I_{m_{n_y}} - \bar{\Lambda}_l) D_y^{(i)T}, \quad l \geq 1,$$

$$\mathcal{Z}_l^{v(i)} = \left( \mathcal{F}_l^T - (1 - \delta_{l,1}) \sum_{l'=1}^{l-1} \mathcal{Z}_{l'}^{v(i)} (\Sigma_{l'}^{\eta^{(i)}})^{-1} (\bar{\mathcal{C}}_l \mathcal{A}_l \mathcal{Z}_{l'}^{x(i)} + \mathcal{H}_l \mathcal{Z}_{l'}^{v(i)})^T \right) (I_{m_{n_y}} - \bar{\Lambda}_l) D_y^{(i)T}, \quad l \geq 1,$$

and, hence, the augmented matrices  $\mathcal{Z}_k^{(i)} = \left( \mathcal{Z}_k^{x(i)T} \mid \mathcal{Z}_k^{v(i)T} \right)^T$  satisfy:

$$\mathcal{Z}_k^{(i)} = \left( (\bar{\mathcal{C}}_k \mathcal{B}_k \mid \mathcal{F}_k)^T - (1 - \delta_{k,1}) \sum_{l=1}^{k-1} \mathcal{Z}_l^{(i)} (\Sigma_l^{\eta^{(i)}})^{-1} \mathcal{Z}_l^{(i)T} (\bar{\mathcal{C}}_k \mathcal{A}_k \mid \mathcal{H}_k)^T \right) (I_{m_{n_y}} - \bar{\Lambda}_k) D_y^{(i)T}, \quad k \geq 1. \quad (7.2)$$

From these preliminaries, we can now derive the expressions of the theorem at hand. Expression (4.1) for the signal filter comes directly from (3.5) using the factorization (7.1) for  $\mathcal{X}_{k,l}^{(i)}$ , and defining

$\mathbf{z}_k^{(i)} \equiv \sum_{l=1}^k \mathcal{Z}_l^{(i)} (\Sigma_l^{\eta^{(i)}})^{-1} \eta_l^{(i)}$  which clearly satisfies (4.3). Now, using the OPL—which guarantees the uncorrelation between the estimation error and the observations—, we express  $\Sigma_{k/k}^{\bar{x}^{(i)}} = E \left[ x_k x_k^T \right] - E \left[ \widehat{x}_{k/k}^{(i)} \widehat{x}_{k/k}^{(i)T} \right]$ , and (4.2) is easily obtained from (4.1), by denoting  $\Sigma_k^{\mathbf{z}^{(i)}} = E \left[ \mathbf{z}_k^{(i)} \mathbf{z}_k^{(i)T} \right]$ . Recursive formula (4.4)

for these matrices is derived by from the fact that the innovation process is white, which leads to  $\Sigma_k^{z^{(i)}} = \sum_{l=1}^k \mathcal{Z}_l^{(i)} (\Sigma_l^{\eta^{(i)}})^{-1} \mathcal{Z}_l^{(i)T}$ ; now, (4.5) is merely expression (7.2) for  $\mathcal{Z}_k^{(i)}$  rewritten in terms of  $\Sigma_{k-1}^{z^{(i)}}$ .

Again, taking again into account the factorizations (7.1), the one-stage predictors of the signal and noise are given by  $\widehat{x}_{k/k-1}^{(i)} = (\mathcal{A}_k | 0) \mathbf{z}_{k-1}^{(i)}$  and  $\widehat{v}_{k/k-1}^{(i)} = (0 | \mathcal{H}_k) \mathbf{z}_{k-1}^{(i)}$ , respectively, and so (3.7) can be rewritten as:

$$\widehat{y}_{k/k-1}^{(i)} = (I_{m_{n_y}} - \bar{\Lambda}_k) (\bar{C}_k \mathcal{A}_k | \mathcal{H}_k) \mathbf{z}_{k-1}^{(i)}. \quad (7.3)$$

Expression (4.6) for the innovation is obtained immediately from (3.6), and (4.7) is deduced by again using the OPL to express the innovation covariance matrix as  $\Sigma_k^{\eta^{(i)}} = D_y^{(i)} (\Sigma_k^y - E [\widehat{y}_{k/k-1}^{(i)} \widehat{y}_{k/k-1}^{(i)T}]) D_y^{(i)T}$ .  $\square$

## Appendix B. Proof of Theorem 2

Recursive formula (4.8) is an immediate consequence of the general expression (3.5). From this, the estimation error is expressed as  $\widehat{x}_{k/h}^{(i)} = \widehat{x}_{k/h-1}^{(i)} - \mathcal{X}_{k,h}^{(i)} (\Sigma_h^{\eta^{(i)}})^{-1} \eta_h^{(i)}$ , and (4.9) is derived straightforwardly taking into account that, from the OPL,  $E [\widehat{x}_{k/h-1}^{(i)} \eta_h^{(i)T}] = E [x_k \eta_h^{(i)T}] = \mathcal{X}_{k,h}^{(i)}$ .

Since, for  $h > k$ ,  $E [x_k y_h^T] = \mathcal{B}_k \mathcal{A}_h^T \bar{C}_h^T (I_{m_{n_y}} - \bar{\Lambda}_h)$ , expression (4.10) for the gain coefficients,  $\mathcal{X}_{k,h}^{(i)} = E [x_k \eta_h^{(i)T}]$ , is obtained using (4.6) for  $\eta_h^{(i)}$  and denoting  $\Sigma_{k,h}^{xz^{(i)}} = E [x_k \mathbf{z}_h^{(i)T}]$ ,  $h \geq k$ .

Finally, recursive formula (4.11) is deduced from (4.3) for  $\mathbf{z}_h^{(i)}$ ; its initial condition is another consequence of the OPL, from which  $E [x_k \mathbf{z}_k^{(i)T}] = E [\widehat{x}_{k/k}^{(i)} \mathbf{z}_k^{(i)T}]$ , and it is obtained simply by using (4.1) for  $\widehat{x}_{k/k}^{(i)}$ .  $\square$

## Appendix C. Proof of Lemma 1

Clearly, expression (5.2) for the cross-covariance matrices is obtained from (4.1) for  $\widehat{x}_{k/k}^{(r)}$  and  $\widehat{x}_{k/k}^{(s)}$ , being  $\Sigma_k^{z^{(rs)}} = E [\mathbf{z}_k^{(r)} \mathbf{z}_k^{(s)T}]$ . Expression (5.3) for  $\Sigma_k^{z^{(rs)}}$  is derived by using (4.3) for  $\mathbf{z}_k^{(r)}$  and  $\mathbf{z}_k^{(s)}$ , and denoting  $\mathcal{Z}_{k-1,k}^{(rs)} = E [\mathbf{z}_{k-1}^{(r)} \eta_k^{(s)T}]$ .

From (4.6) for  $\eta_k^{(s)}$ , we have  $\mathcal{Z}_{k-1,k}^{(rs)} = (E [\mathbf{z}_{k-1}^{(r)} y_k^T] - \Sigma_{k-1}^{z^{(rs)}} (\bar{C}_k \mathcal{A}_k | \mathcal{H}_k)^T (I_{m_{n_y}} - \bar{\Lambda}_k)) D_y^{(s)T}$ . Now, using the OPL, we express  $E [\mathbf{z}_{k-1}^{(r)} y_k^T] = E [\mathbf{z}_{k-1}^{(r)} \widehat{y}_{k/k-1}^{(r)T}]$ , and from (7.3) for  $\widehat{y}_{k/k-1}^{(r)}$ , the following identity holds:

$$E [\mathbf{z}_{k-1}^{(r)} y_k^T] = \Sigma_{k-1}^{z^{(r)}} (\bar{C}_k \mathcal{A}_k | \mathcal{H}_k)^T (I_{m_{n_y}} - \bar{\Lambda}_k), \quad k \geq 1, \quad (7.4)$$

and (5.4) is obtained.

To end the proof, we again use (4.6) for  $\eta_k^{(r)}$  together with the definition  $\mathcal{Z}_{k-1,k}^{(rs)} = E [\mathbf{z}_{k-1}^{(r)} \eta_k^{(s)T}]$  to express the innovation cross-covariance as  $\Sigma_k^{\eta^{(rs)}} = D_y^{(rs)} (E [y_k \eta_k^{(s)T}] - (I_{m_{n_y}} - \bar{\Lambda}_k) (\bar{C}_k \mathcal{A}_k | \mathcal{H}_k) \mathcal{Z}_{k-1,k}^{(rs)})$ . Then, from (4.6) for  $\eta_k^{(s)}$ , and taking into account (7.4) for  $E [y_k \mathbf{z}_{k-1}^{(s)T}]$ , it is clear that:

$$E [y_k \eta_k^{(s)T}] = (\Sigma_k^y - (I_{m_{n_y}} - \bar{\Lambda}_k) (\bar{C}_k \mathcal{A}_k | \mathcal{H}_k) \Sigma_{k-1}^{z^{(s)}} (\bar{C}_k \mathcal{A}_k | \mathcal{H}_k)^T (I_{m_{n_y}} - \bar{\Lambda}_k)) D_y^{(s)T}, \quad k \geq 1,$$

and expression (5.5) is concluded simply by using (5.4) for  $\mathcal{Z}_{k-1,k}^{(rs)}$ .  $\square$

## Appendix D. Proof of Lemma 2

Expression (5.6) for  $\Sigma_{k/h}^{(rs)}$ ,  $h > k$ , is directly obtained from (4.8), by denoting  $\widehat{\mathcal{X}}_{k,h}^{(rs)} = E \left[ \widehat{x}_{k/h-1}^{(r)} \eta_h^{(s)T} \right]$ . In order to derive (5.7), we use (4.6) for  $\eta_h^{(s)}$ , and denote  $\Sigma_{k,h-1}^{(rs)} = E \left[ \widehat{x}_{k/h-1}^{(r)} \mathbf{z}_{h-1}^{(s)T} \right]$  to obtain  $\widehat{\mathcal{X}}_{k,h}^{(rs)} = \left( E \left[ \widehat{x}_{k/h-1}^{(r)} y_h^T \right] - \Sigma_{k,h-1}^{(rs)} \left( \overline{C}_h \mathcal{A}_h \mid \mathcal{H}_h \right)^T \left( I_{m_y} - \overline{\Lambda}_h \right) \right) D_y^{(s)T}$ . Then, taking into account that –from the OPL–  $E \left[ \widehat{x}_{k/h-1}^{(r)} y_h^T \right] = E \left[ x_k \widehat{y}_{h/h-1}^{(r)T} \right]$ , we need only use (7.3) for  $\widehat{y}_{h/h-1}^{(r)}$ .

Finally, expression (5.8) of  $\Sigma_{k,h}^{(rs)}$ ,  $h > k$ , is obtained straightforwardly simply by using (4.8) for  $\widehat{x}_{k/h}^{(r)}$  and (4.3) for  $\mathbf{z}_h^{(s)}$ . Its initial condition is derived directly from (4.1) and  $\Sigma_k^{(rs)} = E \left[ \mathbf{z}_k^{(r)} \mathbf{z}_k^{(s)T} \right]$ .  $\square$



AIMS Press

© 2023 the author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)