Article

Distributed Fusion Estimation in Network Systems Subject to Random Delays and Deception Attacks

María Jesús García-Ligero *,† , Aurora Hermoso-Carazo † and Josefa Linares-Pérez †

Departamento de Estadística e I. O., Universidad de Granada, Avda Fuentenueva s/n, 18071 Granada, Spain; ahermoso@ugr.es (A.H.-C.); jlinares@ugr.es (J.L.-P.)
* Correspondence: mjgarcia@ugr.es
† These authors contributed equally to this work.

Abstract: This paper focuses on the distributed fusion estimation problem in which a signal transmitted over wireless sensor networks is subject to deception attacks and random delays. We assume that each sensor can suffer attacks that may corrupt and/or modify the output measurements. In addition, communication failures between sensors and their local processors can delay the receipt of processed measurements. The randomness of attacks and transmission delays is modelled by different Bernoulli random variables with known probabilities of success. According to these characteristics of the sensor networks and assuming that the measurement noises are cross-correlated at the same time step between sensors and are also correlated with the signal at the same and subsequent time steps, we derive a fusion estimation algorithm, including prediction and filtering, using the distributed fusion method. First, for each sensor, the local least-squares linear prediction and filtering algorithm are derived, using a covariance-based approach. Then, the distributed fusion predictor and the corresponding filter are obtained as the matrix-weighted linear combination of corresponding local estimators, checking that the mean squared error is minimised. A simulation example is then given to illustrate the effectiveness of the proposed algorithms.

Keywords: distributed fusion estimation; sensor networks; deception attacks; random delays; correlated noises

MSC: 60G35; 62M20; 93E10; 93E11

1. Introduction

Wireless sensor networks (WSNs) are a growing focus of research, with applications in a wide variety of fields including industrial processes, signal processing, communications and navigation systems. In general, each WSN is composed of a large number of sensors that can be interconnected by wireless channels and also linked to the processing centre. While they are widely used, the structure of WSNs makes them vulnerable to external attacks and/or unreliability in data transmission. In this context, a topic that has attracted considerable attention is that of signal fusion estimation. This problem can be addressed by centralised or distributed fusion methods. Both are commonly used to tackle the question of fusion estimation; however, the distributed method is usually more suitable, since the structure of the sensors means that less online communication is required, because each local estimator is obtained using only its corresponding local data, and therefore manufacturing costs are reduced. Consequently, this method usually has stronger fault tolerance and lower calculation burdens than the centralised approach. Accordingly, the distributed fusion estimation problem has been of particular interest to researchers, and a large number of distributed fusion estimation algorithms have been developed, using different techniques, in the context of WSNs (see [1–7] and references therein).

As commented above, the characteristics of WSNs make them vulnerable to external attacks and hence less reliable. In general, three kinds of attack have been considered

(classified according to the attack mechanisms and mathematical models involved): Denial-of-service (DoS) attacks, replay attacks and deception attacks. A DoS attack deteriorates system performance by temporarily or indefinitely interrupting the service, preventing information from reaching the destination. In a replay attack, the information is intercepted, captured and forwarded elsewhere, thus altering the system's behaviour. Deception attacks tamper with the original signal measures by accessing the sensors and introducing false information. A problem of significant interest in WSNs subject to cyber-attacks is that of the distributed fusion estimation of the signal. This problem can be addressed in various ways, depending on the type of attack considered. Chen et al. [8] designed a distributed resilient filter for power systems exposed to DoS attacks and gain perturbations. Similarly, the distributed robust fusion estimation problem for systems under DoS attacks and uncertain covariances was studied by [9]. For replay attacks, [10] designed a recursive distributed Kalman fusion estimator focused on the linear minimum variance, for bandwidth-constrained cyber-physical systems. Finally, deception attacks are generally considered the most dangerous and complex types of aggression; for this reason, the scientific community is highly active in researching estimation problems for WSNs subject to this kind of attack. For instance, [11] developed a novel filter for a class of discrete time-delayed stochastic systems subject to the effects of both uniform quantisation and deception attacks. In this field, too, [12] considered a class of time-varying systems subject to multiplicative noises and deception attacks, with unknown but bounded disturbances, and designed an appropriate variance-constrained distributed filter. [13] derived a distributed H_∞-consensus filter for a class of discrete time-varying systems subject to both multiplicative noises and deception attacks. In a related approach, assuming clustering sensor networks subject to random deception attacks, [14] considered distributed fusion filtering and fixed-point smoothing problems.

Another significant problem in WSNs is the limitation of network communication bandwidth, which can produce network congestion and hence delay reception of the measures at the processing centre, thus impairing the performance of the estimator. Accordingly, new algorithms must be developed to take these effects into account in distributed fusion estimation. In studies of one-step random delays modelled by Bernoulli random variables, various types of distributed fusion estimation algorithms have been proposed, assuming either that these variables are independent (see for example, [15–18]) or that they are correlated at consecutive sampling times [19,20]. In addition, various distributed fusion estimation algorithms describing random delays by Markov chains have been derived (see, for example, [21–23]).

Many early studies of the estimation problem in multi-sensor systems assumed that the system noises were white and uncorrelated. However, in practice, this assumption is not always realistic; for example, when all sensors observe a common target within a noisy environment, there may be cross-correlations between the different sensor noises. Cross-correlation between process and measurement noises can also arise, for example, when in a target tracking system both noises are dependent of the state of the system. In this respect, Li et al. [24] showed that if a discrete-time linear system is obtained from the discretisation of a continuous-time system, then the measurement noise is correlated with the system noise of the previous time step. These common situations in multi-sensor systems have led researchers to address the fusion estimation problem under the assumption of correlated noises. Thus, assuming that the measurement noises from different sensors are both cross-correlated and correlated with the system noise at the previous time step, Yan et al. [25] derived the optimal state estimator, in a sequential form and by means of distributed fusion. Similarly, ref. [26] developed globally optimal sequential and distributed fusion estimation algorithms assuming linear minimum variance. The distributed estimation problem has also been addressed when the noises between sensors are not correlated at the same instant, but the process and measurement noises are two-step cross-correlated; under this assumption, distributed filtering algorithms have been obtained assuming that

the system is subject to stochastic uncertainties or multiplicative noises [27] and in systems with random matrices over a sensor network with a given topology [28].

Motivated by the above analysis, our aim in this paper is to address the distributed fusion estimation problem in WSNs subjected to the three situations described above, as often encountered in real life: Deception attacks, one-step random transmission delays and measurement noises cross-correlated between sensors and correlated with the signal. To the best of our knowledge, this problem has not been considered previously, although some papers have discussed deception attacks and delays or correlated noises. Specifically, for a single sensor, the estimation problem is studied assuming deception attacks and random communication delays in [29,30], and in a multi-sensor environment, ref. [31] investigated the distributed filtering problem in WSNs receiving deception attacks and when sensor additive noises are both cross-correlated and correlated with the process noise.

In this paper, we assume that each sensor of the WSN can suffer random attacks that modify the output measurements. The randomness of these attacks is modelled by independent random Bernoulli variables reflecting whether or not the attack succeeded in modifying the measure. Furthermore, the measurement noises from different sensors are cross-correlated at the same time step and are also correlated with the signal process at the same and subsequent time steps. Due to possible imperfections in communication channels, the processed measurements can be one-step delayed with different delay rates. Transmission delays, which occur randomly, are described by different random Bernoulli variables describing whether the processed measurements arrive on time, or are delayed by one sampling time. In this context, assuming that no signal evolution equation is available, a distributed fusion prediction and filtering algorithm is derived using the first and second-order moments of the signal and the processes involved in the observations. The design of the proposed distributed prediction and filtering estimators is structured in two stages. In the first, an innovation approach is used to obtain the local least-squares (LS) linear prediction and the filtering estimators. In the second stage, the information provided by the local estimators is combined to derive the fusion estimators; specifically, the distributed fusion predictor and filter are obtained as a matrix-weighted linear combination of the local LS linear estimators, using the mean squared error as the optimality criterion.

The rest of this paper is organised as follows. In the next section, we describe the model proposed and specify the assumptions under which the distributed fusion estimation problem is addressed. The distributed fusion prediction and filtering algorithm and the associated error covariance matrices, which provide a measure of the estimator performance, are then derived in Section 3. In order to illustrate the feasibility of the proposed distributed fusion estimators, a simulation example is given in Section 4. Finally, the main conclusions drawn are summarised in Section 5.

Notation. The notations used in this paper are standard. $\mathbb{R}^n$ and $\mathbb{R}^{n \times m}$ denote the n-dimensional Euclidean space and the set of all $n \times m$ matrices, respectively. If the dimensions of vectors or matrices are not explicitly stated, they are assumed to be compatible with algebraic operations. The shorthand $(A_1 | \ldots | A_m)$ stands for a matrix partitioned into submatrices A_i, $i = 1, \ldots, m$. The Kronecker delta function is denoted as $\delta_{k,s}$. Finally, $L^{(i)} = L^{(ii)}$ is written for any function $L^{(ij)}$ depending on sensors i and j; similarly, for any function $G_{k,s}$, depending on the time instants k and s, we simply write G_k when $k = s$.

2. Problem Formulation

This paper considers the linear estimation problem of a discrete-time signal over WSNs subject to deception attacks, random delays and correlated noises. More specifically, the problem is to estimate a signal from the sensor measurements $z_k^{(i)}$, $i = 1, \ldots, m$, which are described by Equation (1). We assume that deception attacks are launched by potential adversaries, inserting false information that may perturb the real measurements. Assuming that the attacks may randomly succeed or not, the sensor-measured outputs subject to these attacks, $\breve{z}_k^{(i)}$, are modeled by (2). It is also assumed that each sensor transmits its outputs to

a local processor over a packet-erasure channel, where one-step random delays can occur during transmission; the measures received by the local processors, $y_k^{(i)}$, are given by (5). The linear estimation problem is addressed using the distributed fusion method, by which each local processor produces LS linear predictors and filters, $\widehat{x}_{k/s}^{(i)}$, $s = k - 1, k$, of the signal, x_k, based on the measures received from the corresponding sensor, $y_1^{(i)}, \ldots, y_s^{(i)}$; afterwards, these estimators are transmitted, over perfect connections, to the fusion center, where new prediction and filtering estimators, $\widehat{x}_{k/s}$, $s = k - 1, k$, are obtained as matrix-weighted linear combination of the local estimators using the mean squared error as the optimality criterion (see Figure 1).
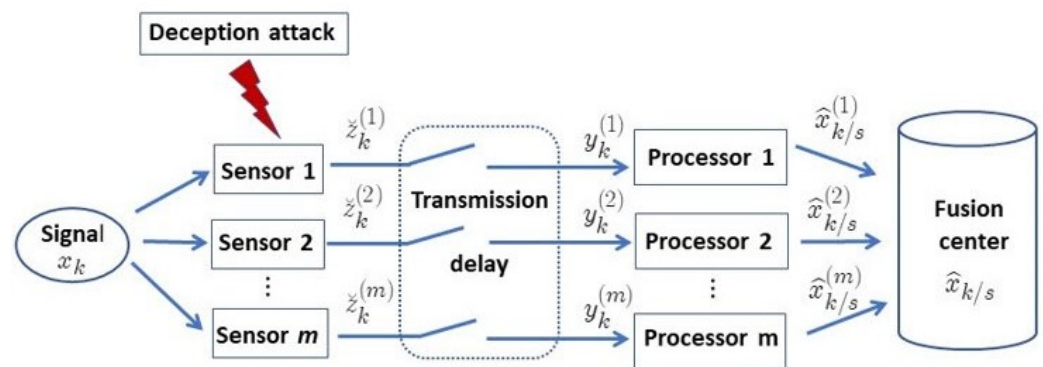

Figure 1. Distributed fusion estimation scheme in WNS with attacks and delays.

2.1. Multisensor Measurement

Consider a networked system consisting of m sensors which measure a n_x-dimensional discrete-time signal x_k according the following model:

$$z_k^{(i)} = H_k^{(i)} x_k + v_k^{(i)}, \quad k \geq 1, \quad i = 1, \ldots m,$$ (1)

where $z_k^{(i)} \in \mathbb{R}^{n_z}$, $H_k^{(i)} \in \mathbb{R}^{n_z \times n_x}$ are known matrices and $v_k^{(i)}$ are the measurement noises.
The following assumptions on the measurement model (1) are required:

Assumption 1. *The n_x-dimensional signal $\{x_k, \ k \geq 1\}$ is a second-order process with zero mean and autocovariance function expressed in a separable form as follows:*

$$E\left[x_k x_h^T\right] = A_k B_h^T, \quad h \leq k,$$

where $A_k, B_h \in \mathbb{R}^{n_x \times M}$ are known matrices.

Assumption 2. *The measurement noises $\left\{v_k^{(i)}, \ k \geq 1\right\}$, $i = 1, \ldots, m$, are white second-order processes with the following statistical properties:*

(i) $E[v_k^{(i)}] = 0$ *and their correlation functions are known and given by*

$$E[v_k^{(i)} v_h^{(j)T}] = R_k^{(ij)} \delta_{k,h}, \quad i, j = 1, \ldots, m.$$

(ii) *For $i = 1, \ldots, m$, $\left\{v_k^{(i)}, \ k \geq 1\right\}$, is correlated with the signal process $\{x_k, \ k \geq 1\}$ and its correlation function is expressed in separable form as*

$$E[x_k v_h^{(i)T}] = \begin{cases} C_k D_h^{(i)T}, & h \leq k, \\ 0, & h > k, \end{cases}$$

where $C_k \in \mathbb{R}^{n_x \times M'}$ and $D_h^{(i)} \in \mathbb{R}^{n_z \times M'}$ are known matrices.

Remark 1. *The general assumption of uncorrelation for the system noises 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. Assumption 2 weakens this condition of uncorrelated noises. Concretely:*

- *The sensor noise cross-correlation specified in (i) can be found, for example, when the signal process is observed by sensors that operate in a common noisy environment, or when the noises are state dependent; in this case, the cross-correlation between the process noise and the sensor noises leads to correlation between the different sensor noises. Moreover, after augmenting a system subject to uncertainties such as random delays or packet-dropouts, the transformed system presents correlated noises.*

- *Condition (ii) allows us to consider models in which the sensor noises are correlated with the process noise at the previous time (see Section 4). One of the most common situations in which this type of correlation occurs is in the discretisation of continuous-time systems; actually, if a discrete-time linear system is obtained from the discretisation of a continuous-time system, then the measurement noise is correlated with the process noise at the previous time.*

2.2. Deception Attack Model

As discussed above, the measurement process can be subject to attacks launched by malicious attackers. Specifically, at each sensor, attackers could inject false information that would degrade or deteriorate the real measurement. If deception attacks are launched randomly, the output measure may or may not be modified; therefore, the mathematical model for current output after a randomly occurring deception attack is modelled as:

$$\check{z}_k^{(i)} = z_k^{(i)} + \beta_k^{(i)} \xi_k^{(i)}, \quad k \geq 1, \ i = 1, \ldots, m, \tag{2}$$

where $\left\{ \beta_k^{(i)}, \ k \geq 1 \right\}$, $i = 1, \ldots, m$, are sequences of Bernoulli random variables, which model the success, $\beta_k^{(i)} = 1$, or failure, $\beta_k^{(i)} = 0$, of the attack on the *ith*-sensor and $\xi_k^{(i)}$ denotes the deception attack signal sent by adversaries to the *ith*-sensor, which can be described as follows:

$$\xi_k^{(i)} = -z_k^{(i)} + \varepsilon_k^{(i)}, \quad k \geq 1, \ i = 1, \ldots, m, \tag{3}$$

where $\varepsilon_k^{(i)}$ represents a random deception signal, which is inaccessible to the defenders.

According to the deception attacks (3), the attacked measurement outputs (2) can be rewritten as:

$$\check{z}_k^{(i)} = (1 - \beta_k^{(i)}) z_k^{(i)} + \beta_k^{(i)} \varepsilon_k^{(i)}, \quad k \geq 1, \ i = 1, \ldots, m. \tag{4}$$

The following assumptions are set on the sequences of random variables that model the occurrence or otherwise of the attack and the deception signal:

Assumption 3. $\left\{ \beta_k^{(i)}, \ k \geq 1 \right\}$, $i = 1, \ldots, m$, *are independent sequences of independent Bernoulli random variables with known success probabilities* $P(\beta_k^{(i)} = 1) = \overline{\beta}_k^{(i)}$, $k \geq 1$.

Assumption 4. $\left\{ \varepsilon_k^{(i)}, \ k \geq 1 \right\}$, $i = 1, \ldots, m$, *are independent white second-order processes with zero-mean and known covariance matrices* $E[\varepsilon_k^{(i)} \varepsilon_h^{(j)T}] = T_k^{(ij)} \delta_{k,h}$.

Remark 2. *For various reasons, such as the possible presence of safety protection devices, the attacks launched by adversaries are not always successful. From the defender's perspective, the uncertainty about the success or failure of a deception attack should be considered random and included in the model with the certain-success rates, different for every sensor and every time instant. The nature of the uncertainty (successful or failed attacks) suggests that the randomness should be modelled using different sequences of random binary variables,* $\left\{ \beta_k^{(i)}, \ k \geq 1 \right\}$, *each taking the value 0 ($\beta_k^{(i)} = 0$, meaning that the attack against the ith-sensor at time k has failed) or the value*

1 ($\beta_k^{(i)} = 1$ representing the success of the ith-sensor attack at time k). Clearly, therefore, the measurement output of the ith-sensor at time k, $\check{z}_k^{(i)}$, is modelled by (3); that is, it is only noise, $\check{z}_k^{(i)} = \varepsilon_k^{(i)}$, if $\beta_k^{(i)} = 1$ (successful attack) or, alternatively, it is the real measure, $\check{z}_k^{(i)} = z_k^{(i)}$, if $\beta_k^{(i)} = 0$ (failed attack). Assumption 3 on $\{\beta_k^{(i)}, k \geq 1\}$ specifies the independence of the success or failure of attacks against different sensors or at different times. Consequently, under this assumption, each measurement output is either the real measure or consists only of noise, regardless of the other.

2.3. Randomly Delayed Observations

It is assumed that the measures arriving from each sensor to its processing centre may have suffered delays during the transmission process, due to possible imperfections in the communication channels. For each sensor, the absence or presence of delays in the transmission is modelled by different sequences of Bernoulli variables describing whether the measures arrive on time or are delayed by one sampling time; thus, the measure to be processed at time k is $\check{z}_k^{(i)}$ or $\check{z}_{k-1}^{(i)}$. In addition, it is assumed that at the initial instant the actual outputs are always available. Therefore, our mathematical model of randomly one-step delayed received measures is described by

$$y_k^{(i)} = (1 - \gamma_k^{(i)})\check{z}_k^{(i)} + \gamma_k^{(i)}\check{z}_{k-1}^{(i)}, \quad k \geq 2; \ y_1^{(i)} = \check{z}_1^{(i)}, \ i = 1, \ldots, m, \tag{5}$$

and the following assumption is made for the random variables $\gamma_k^{(i)}$ modelling the delays:

Assumption 5. $\left\{\gamma_k^{(i)}, k \geq 2\right\}$, $i = 1, \ldots, m$, are independent sequences of independent Bernoulli random variables with known success probabilities $P(\gamma_k^{(i)} = 1) = \overline{\gamma}_k^{(i)}$, $k \geq 2$.

Finally, the following assumption is made about the signal and processes involved in the measurement model used to derive the LS linear estimators.

Assumption 6. For $i = 1, \ldots, m$, the processes $\left(\{x_k, k \geq 1\}, \left\{v_k^{(i)}, k \geq 1\right\}\right)$, $\left\{\varepsilon_k^{(i)}, k \geq 1\right\}$, $\left\{\beta_k^{(i)}, k \geq 1\right\}$ and $\left\{\gamma_k^{(i)}, k \geq 2\right\}$ are mutually independent.

From the previous assumptions about the measurement, attacked measurement and observation processes, the expressions of their second order moments are deduced in the following lemma.

Lemma 1. Under the model assumptions, the following statistical properties are satisfied:

(i) The measurement processes $\left\{z_k^{(i)}, k \geq 1\right\}$, $i = 1, \ldots, m$, have zero-mean and their covariance functions, $\Sigma_{k,h}^{z^{(ij)}} \equiv E\left[z_k^{(i)} z_h^{(j)T}\right]$, are given by

$$\Sigma_{k,h}^{z^{(ij)}} = H_k^{(i)} A_k B_h^T H_h^{(j)T} + R_k^{(ij)} \delta_{k,h} + H_k^{(i)} C_k D_h^{(j)T} + D_k^{(i)} C_k^T H_h^{(j)T} \delta_{k,h}, \ 1 \leq h \leq k.$$

(ii) The measurements $\left\{\check{z}_k^{(i)}, k \geq 1\right\}$, $i = 1, \ldots, m$, are zero-mean processes with

$$\Sigma_k^{\check{z}^{(i)}} \equiv E\left[\check{z}_k^{(i)} \check{z}_k^{(i)T}\right] = (1 - \overline{\beta}_k^{(i)})\Sigma_k^{z^{(i)}} + \overline{\beta}_k^{(i)} T_k^{(i)}, \ k \geq 1.$$
$$\Sigma_{k,h}^{\check{z}^{(i)}} \equiv E\left[\check{z}_k^{(i)} \check{z}_h^{(i)T}\right] = (1 - \overline{\beta}_k^{(i)})(1 - \overline{\beta}_h^{(i)})\Sigma_{k,h}^{z^{(i)}}, \ k,h \geq 1; \ h \neq k.$$
$$\Sigma_{k,h}^{\check{z}^{(ij)}} \equiv E\left[\check{z}_k^{(i)} \check{z}_h^{(j)T}\right] = (1 - \overline{\beta}_k^{(i)})(1 - \overline{\beta}_h^{(j)})\Sigma_{k,h}^{z^{(ij)}} + \overline{\beta}_k^{(i)}\overline{\beta}_h^{(j)} T_k^{(ij)} \delta_{k,h}, \ k,h \geq 1; \ j \neq i.$$

(iii) The observation processes $\left\{y_k^{(i)}, k \geq 1\right\}$, $i = 1, \ldots, m$, have zero-mean and their covariance functions, $\Sigma_{k,h}^{y^{(ij)}} \equiv E\left[y_k^{(i)} y_h^{(j)T}\right]$, satisfy:

- For $i = j$

$$\Sigma_k^{y(i)} = (1 - \overline{\gamma}_k^{(i)})\Sigma_k^{\breve{z}(i)} + \overline{\gamma}_k^{(i)}\Sigma_{k-1}^{\breve{z}(i)}, \ k \geq 2; \ \Sigma_1^{y(i)} = \Sigma_1^{\breve{z}(i)}.$$

$$\Sigma_{k,h}^{y(i)} = (1 - \overline{\gamma}_k^{(i)})(1 - \overline{\gamma}_h^{(i)})\Sigma_{k,h}^{\breve{z}(i)} + (1 - \overline{\gamma}_k^{(i)})\overline{\gamma}_h^{(i)}\Sigma_{k,h-1}^{\breve{z}(i)} + \overline{\gamma}_k^{(i)}(1 - \overline{\gamma}_h^{(i)})\Sigma_{k-1,h}^{\breve{z}(i)}$$

$$+ \overline{\gamma}_k^{(i)}\overline{\gamma}_h^{(i)}\Sigma_{k-1,h-1}^{\breve{z}(i)}, \ 1 < h < k; \ \Sigma_{k,1}^{y(i)} = (1 - \overline{\gamma}_k^{(i)})\Sigma_{k,1}^{\breve{z}(i)} + \overline{\gamma}_k^{(i)}\Sigma_{k-1,1}^{\breve{z}(i)} \ k \geq 2.$$

- For $i \neq j$

$$\Sigma_{k,h}^{y(ij)} = (1 - \overline{\gamma}_k^{(i)})(1 - \overline{\gamma}_h^{(j)})\Sigma_{k,h}^{\breve{z}(ij)} + (1 - \overline{\gamma}_k^{(i)})\overline{\gamma}_h^{(j)}\Sigma_{k,h-1}^{\breve{z}(ij)} + \overline{\gamma}_k^{(i)}(1 - \overline{\gamma}_h^{(j)})\Sigma_{k-1,h}^{\breve{z}(ij)}$$

$$+ \overline{\gamma}_k^{(i)}\overline{\gamma}_h^{(j)}\Sigma_{k-1,h-1}^{\breve{z}(ij)}, \ 1 < h \leq k;$$

$$\Sigma_{k,1}^{y(ij)} = (1 - \overline{\gamma}_k^{(i)})\Sigma_{k,1}^{\breve{z}(ij)} + \overline{\gamma}_k^{(i)}\Sigma_{k-1,1}^{\breve{z}(ij)}, \ k \geq 2; \ \Sigma_1^{y(ij)} = \Sigma_1^{\breve{z}(ij)}.$$

Proof. These statistical properties are easily obtained taking into account Equations (1), (4) and (5), and model Assumptions (1)–(6). □

3. Least-Squares Linear Distributed Fusion Estimation Problem

The LS linear estimation problem of the signal from the observations coming from the m sensors is addressed by the distributed fusion method. This method is structured in two stages: The first consists of obtaining, for each sensor, LS linear estimators; in the second one, the distributed fusion estimators are calculated as matrix-weight linear combination of the local ones, using the LS optimisation criterion. Following this methodology, algorithms for the prediction and filtering problems are developed below.

3.1. Stage One: LS Local Prediction and Filtering Algorithm

In this stage, our aim is to obtain a recursive algorithm for the local estimators of the signal x_k based on the observations $\left\{ y_1^{(i)}, \ldots, y_s^{(i)} \right\}$, for $s \leq k$, applying an innovation approach. Under this approach, the observation process is transformed, by an orthogonalisation procedure, into an equivalent one that provides the same information; this is termed the innovation process. The innovation at time k is defined as $\mu_k^{(i)} = y_k^{(i)} - \widehat{y}_{k/k-1}^{(i)}$, where $\widehat{y}_{k/k-1}^{(i)}$ is the one-stage observation predictor. Due to this equivalence, the local estimator of a vector u_k based on the observations $\left\{ y_1^{(i)}, \ldots, y_s^{(i)} \right\}$ can be expressed as a linear combination of the innovations $\left\{ \mu_1^{(i)}, \ldots, \mu_s^{(i)} \right\}$, as follows:

$$\widehat{u}_{k/s}^{(i)} = \sum_{h=1}^{s} E[u_k \mu_h^{(i)T}]\Pi_h^{(i)-1}\mu_h^{(i)}, \tag{6}$$

where $\Pi_h^{(i)} \equiv E[\mu_h^{(i)}\mu_h^{(i)T}]$ denotes the innovation covariance matrix.

Then, our first objective is to find an adequate expression for the innovations $\mu_h^{(i)} = y_h^{(i)} - \widehat{y}_{h/h-1}^{(i)}$, $h \leq s$, or, equivalently, for the one-stage observation predictors $\widehat{y}_{h/h-1}^{(i)}$, which simplifies the general expression of the local estimators of the signal, $\widehat{x}_{k/s}^{(i)} = \sum_{h=1}^{s} E[x_k \mu_h^{(i)T}]\Pi_h^{(i)-1}\mu_h^{(i)}$, $s \leq k$. This expression of the innovations will enable us to obtain the coefficients $E[x_k \mu_h^{(i)T}]$ and the innovation covariance matrices $\Pi_h^{(i)}$ that appear in the above expression of $\widehat{x}_{k/s}^{(i)}$.

One-stage observation predictor.

To facilitate the derivation of the one-stage observation predictor, we rewrite the observation expression (5) as

$$y_k^{(i)} = (1 - \gamma_k^{(i)})(1 - \beta_k^{(i)})(H_k^{(i)}x_k + v_k^{(i)}) + (1 - \gamma_k^{(i)})\beta_k^{(i)}\varepsilon_k^{(i)} + \gamma_k^{(i)}(1 - \overline{\beta}_{k-1}^{(i)})H_{k-1}^{(i)}x_{k-1} + w_{k-1}^{(i)}, \ k \geq 2;$$

$$y_1^{(i)} = \breve{z}_1^{(i)}, \tag{7}$$

where

$$w_{k-1}^{(i)} = \gamma_k^{(i)}(\overline{\beta}_{k-1}^{(i)} - \beta_{k-1}^{(i)})H_{k-1}^{(i)}x_{k-1} + \gamma_k^{(i)}(1 - \beta_{k-1}^{(i)})v_{k-1}^{(i)} + \gamma_k^{(i)}\beta_{k-1}^{(i)}\varepsilon_{k-1}^{(i)}, \quad k \geq 2.$$

Then, taking into account the independence assumption and the Orthogonal Projection Lemma (OPL), it is clear that the one-stage observation predictor is given by

$$\widehat{y}_{k/k-1}^{(i)} = (1 - \overline{\gamma}_k^{(i)})(1 - \overline{\beta}_k^{(i)})H_k^{(i)}\widehat{x}_{k/k-1}^{(i)} + \overline{\gamma}_k^{(i)}(1 - \overline{\beta}_{k-1}^{(i)})H_{k-1}^{(i)}\widehat{x}_{k-1/k-1}^{(i)} + \widehat{w}_{k-1/k-1}^{(i)}, \quad k \geq 2,$$

where $\widehat{x}_{k/k-1}^{(i)}$ and $\widehat{x}_{k-1/k-1}^{(i)}$ are the predictor and filter of the signal, respectively, and $\widehat{w}_{k-1/k-1}^{(i)}$ is the filter of $w_{k-1}^{(i)}$.

From (6) for $\widehat{w}_{k-1/k-1}^{(i)}$, and taking into account that $w_{k-1}^{(i)}$ is independent of $y_1^{(i)}, \ldots, y_{k-2}^{(i)}$, it is easy to obtain that

$$\widehat{w}_{k-1/k-1}^{(i)} = \mathcal{W}_{k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mu_{k-1}^{(i)}, \quad k \geq 2,$$

where $\mathcal{W}_{k-1}^{(i)} = E[w_{k-1}^{(i)}\mu_{k-1}^{(i)T}]$.

Therefore, the one-stage observation predictor is given by

$$\widehat{y}_{k/k-1}^{(i)} = (1 - \overline{\gamma}_k^{(i)})(1 - \overline{\beta}_k^{(i)})H_k^{(i)}\widehat{x}_{k/k-1}^{(i)} + \overline{\gamma}_k^{(i)}(1 - \overline{\beta}_{k-1}^{(i)})H_{k-1}^{(i)}\widehat{x}_{k-1/k-1}^{(i)} + \mathcal{W}_{k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mu_{k-1}^{(i)}, \quad k \geq 2; \tag{8}$$
$$\widehat{y}_{1/0}^{(i)} = 0.$$

To simplify the expressions of the local estimation algorithm derived in Theorem 1, the following notations are used for $i = 1, \ldots, m$:

$$\mathcal{H}_{G_k}^{(i)} = (1 - \overline{\gamma}_k^{(i)})(1 - \overline{\beta}_k^{(i)})H_k^{(i)}G_k + \overline{\gamma}_k^{(i)}(1 - \overline{\beta}_{k-1}^{(i)})H_{k-1}^{(i)}G_{k-1}, \quad k \geq 2; \tag{9}$$
$$\mathcal{H}_{G_1}^{(i)} = (1 - \overline{\beta}_1^{(i)})H_1^{(i)}G_1.$$

$$\mathcal{D}_k^{(i)} = (1 - \overline{\gamma}_k^{(i)})(1 - \overline{\beta}_k^{(i)})D_k^{(i)} + \overline{\gamma}_k^{(i)}(1 - \overline{\beta}_{k-1}^{(i)})D_{k-1}^{(i)}, \quad k \geq 2; \tag{10}$$
$$\mathcal{D}_1^{(i)} = (1 - \overline{\beta}_1^{(i)})D_1^{(i)},$$

where $G_k = A_k, B_k$ and C_k.

Theorem 1. *Under the model assumptions set out in Section 2, for $i = 1, \ldots, m$, the local predictor and filter, $\widehat{x}_{k/s}^{(i)}$, $s \leq k$, and the associated error covariances matrices, $P_{k/s}^{(i)} \equiv E[(x_k - \widehat{x}_{k/s}^{(i)})(x_k - \widehat{x}_{k/s}^{(i)})^T]$, are given by*

$$\widehat{x}_{k/s}^{(i)} = (A_k|C_k)O_s^{(i)}, \quad s \leq k, \tag{11}$$

$$P_{k/s}^{(i)} = A_k B_s^T - (A_k|C_k)r_s^{(i)}(A_k|C_k)^T, \quad s \leq k, \tag{12}$$

where the vectors $O_s^{(i)}$ and the matrices $r_s^{(i)} \equiv E[O_s^{(i)}O_s^{(i)T}]$ are recursively obtained from

$$O_s^{(i)} = O_{s-1}^{(i)} + J_s^{(i)}\Pi_s^{(i)-1}\mu_s^{(i)}, \quad s \geq 1; \quad O_0^{(i)} = 0, \tag{13}$$

$$r_s^{(i)} = r_{s-1}^{(i)} + J_s^{(i)}\Pi_s^{(i)-1}J_s^{(i)T}, \quad s \geq 1; \quad r_0^{(i)} = 0, \tag{14}$$

and $J_s^{(i)} \equiv E[O_s^{(i)}\mu_s^{(i)T}]$ satisfies

$$J_s^{(i)} = \left(\mathcal{H}_{B_s}^{(i)}|\mathcal{D}_s^{(i)}\right)^T - r_{s-1}^{(i)}\left(\mathcal{H}_{A_s}^{(i)}|\mathcal{H}_{C_s}^{(i)}\right)^T - J_{s-1}^{(i)}\Pi_{s-1}^{(i)-1}\mathcal{W}_{s-1}^{(i)T}, \quad s \geq 2; \tag{15}$$
$$J_1^{(i)} = \left(\mathcal{H}_{B_1}^{(i)}|\mathcal{D}_1^{(i)}\right)^T.$$

The innovations, $\mu_s^{(i)} = y_s^{(i)} - \hat{y}_{s/s-1}^{(i)}$, and their covariance matrices, $\Pi_s^{(i)} \equiv E[\mu_s^{(i)} \mu_s^{(i)T}]$ are calculated as

$$\begin{aligned} \mu_s^{(i)} &= y_s^{(i)} - \left(\mathcal{H}_{A_s}^{(i)} | \mathcal{H}_{C_s}^{(i)}\right) O_{s-1}^{(i)} - \mathcal{W}_{s-1}^{(i)} \Pi_{s-1}^{(i)-1} \mu_{s-1}^{(i)}, \quad s \geq 2; \\ \mu_1^{(i)} &= y_1^{(i)}, \end{aligned} \tag{16}$$

$$\Pi_s^{(i)} = \Sigma_s^{y^{(i)}} - \left(\mathcal{H}_{A_s}^{(i)} | \mathcal{H}_{C_s}^{(i)}\right) \left(\left(\mathcal{H}_{B_s}^{(i)} | \mathcal{D}_s^{(i)}\right)^T - J_s^{(i)}\right) - \mathcal{W}_{s-1}^{(i)} \Pi_{s-1}^{(i)-1} \left(\left(\mathcal{H}_{A_s}^{(i)} | \mathcal{H}_{C_s}^{(i)}\right) J_{s-1}^{(i)} + \mathcal{W}_{s-1}^{(i)}\right)^T, \quad s \geq 2; \tag{17}$$
$$\Pi_1^{(i)} = \Sigma_1^{y^{(i)}}.$$

Finally, the coefficients $\mathcal{W}_{s-1}^{(i)} \equiv E[w_{s-1}^{(i)} \mu_{s-1}^{(i)T}]$, $s \geq 2$, are obtained by

$$\mathcal{W}_{s-1}^{(i)} = \Sigma_{s,s-1}^{y^{(i)}} - \left(\mathcal{H}_{A_s}^{(i)} | \mathcal{H}_{C_s}^{(i)}\right) \left(\mathcal{H}_{B_{s-1}}^{(i)} | \mathcal{D}_{s-1}^{(i)}\right)^T, \quad s \geq 2, \tag{18}$$

and the matrices $\Sigma_s^{y^{(i)}}$ and $\Sigma_{s,s-1}^{y^{(i)}}$ are given in Lemma 1.

Proof. See Appendix A. □

3.2. Stage Two: Distributed LS Fusion Predictor and Filter

In this stage, the distributed fusion predictor and filter, $\hat{x}_{k/s}$, $s \leq k$, are obtained as the matrix-weight linear combination of the m corresponding local estimators, $\hat{x}_{k/s}^{(i)}$, $s \leq k$, $i = 1, \ldots, m$, by using the mean squared error as the optimality criterion.

Let $\hat{X}_{k/s} = \left(\hat{x}_{k/s}^{(1)T}, \ldots, \hat{x}_{k/s}^{(m)T}\right)^T$ be the vector constituted by the local estimators calculated from the algorithm given in Theorem 1; then, the distributed estimators and the error covariance matrices are obtained by the following algorithm.

Distributed fusion prediction and filtering algorithm. The distributed fusion predictor and filter are given by

$$\hat{x}_{k/s} = \left(\hat{\Sigma}_{k/s}^{(1)} | \ldots | \hat{\Sigma}_{k/s}^{(m)}\right) \hat{\Sigma}_{k/s}^{-1} \hat{X}_{k/s}, \quad s \leq k, \tag{19}$$

and the error covariance matrices $P_{k/s} = E[(x_k - \hat{x}_{k/s})(x_k - \hat{x}_{k/s})^T]$ are obtained as

$$P_{k/s} = A_k B_s^T - \left(\hat{\Sigma}_{k/s}^{(1)} | \ldots | \hat{\Sigma}_{k/s}^{(m)}\right) \hat{\Sigma}_{k/s}^{-1} \left(\hat{\Sigma}_{k/s}^{(1)} | \ldots | \hat{\Sigma}_{k/s}^{(m)}\right)^T, \quad s \leq k, \tag{20}$$

where $\hat{\Sigma}_{k/s} \equiv \left(\hat{\Sigma}_{k/s}^{(ij)}\right)_{i,j=1,\ldots,m}$ and $\hat{\Sigma}_{k/s}^{(ij)} \equiv E\left[\hat{x}_{k/s}^{(i)} \hat{x}_{k/s}^{(j)T}\right]$, the cross-correlation matrices between any two local estimators, are given in the following theorem.

Expressions (19) and (20) for the estimators and the error covariance matrices can be easily obtained by applying the LS criterion (see e.g., García-Ligero et al. [21]).

Theorem 2. *Under the model assumptions, the cross-correlation matrices $\hat{\Sigma}_{k/s}^{(ij)} \equiv E[\hat{x}_{k/s}^{(i)} \hat{x}_{k/s}^{(j)T}]$, $i, j = 1, \ldots, m$, are calculated by*

$$\hat{\Sigma}_{k/s}^{(ij)} = (A_k | C_k) r_s^{(ij)} (A_k | C_k)^T, \quad s \leq k. \tag{21}$$

The matrices $r_s^{(ij)} \equiv E[O_s^{(i)} O_s^{(j)T}]$ verify the following recursive relation

$$r_s^{(ij)} = r_{s-1}^{(ij)} + J_{s-1,s}^{(ij)} \Pi_s^{(j)-1} J_s^{(j)T} + J_s^{(i)} \Pi_s^{(i)-1} J_s^{(ji)T}, \quad s \geq 1; \quad r_0^{(ij)} = 0, \tag{22}$$

where $J_{s-1,s}^{(ij)} \equiv E[O_{s-1}^{(i)} \mu_s^{(j)T}]$ *and* $J_s^{(ij)} \equiv E[O_s^{(i)} \mu_s^{(j)T}]$ *satisfy*

$$
\begin{aligned}
J_{s-1,s}^{(ij)} &= \left(r_{s-1}^{(i)} - r_{s-1}^{(ij)} \right) \left(\mathcal{H}_{A_s}^{(j)} | \mathcal{H}_{C_s}^{(j)} \right)^T + J_{s-1}^{(i)} \Pi_{s-1}^{(i)-1} \mathcal{W}_{s-1}^{(ji)T} - J_{s-1}^{(ij)} \Pi_{s-1}^{(j)-1} \mathcal{W}_{s-1}^{(j)T}, \quad s \geq 2; \\
J_{0,1}^{(ij)} &= 0,
\end{aligned}
\tag{23}
$$

and

$$
J_s^{(ij)} = J_{s-1,s}^{(ij)} + J_s^{(i)} \Pi_s^{(i)-1} \Pi_s^{(ij)}, \quad s \geq 1.
\tag{24}
$$

The innovations cross-covariance matrices $\Pi_s^{(ij)} \equiv E[\mu_s^{(i)} \mu_s^{(i)T}]$ *are obtained as*

$$
\begin{aligned}
\Pi_s^{(ij)} &= \Sigma_s^{y^{(ij)}} - \left(\mathcal{H}_{A_s}^{(i)} | \mathcal{H}_{C_s}^{(i)} \right) \left(\left(\mathcal{H}_{B_s}^{(j)} | \mathcal{D}_s^{(j)} \right)^T - J_s^{(j)} - J_{s-1,s}^{(ij)} \right) - \mathcal{W}_{s-1}^{(ij)} \Pi_{s-1}^{(j)-1} \left(\left(\mathcal{H}_{A_s}^{(j)} | \mathcal{H}_{C_s}^{(j)} \right) J_{s-1}^{(j)} + \mathcal{W}_{s-1}^{(j)} \right)^T \\
&\quad - \mathcal{W}_{s-1}^{(i)} \Pi_{s-1}^{(i)-1} \Pi_{s-1,s}^{(ij)}, \quad s \geq 2; \\
\Pi_1^{(ij)} &= \Sigma_1^{y^{(ij)}},
\end{aligned}
\tag{25}
$$

where $\Pi_{s,s-1}^{(ij)} \equiv E[\mu_s^{(i)} \mu_{s-1}^{(i)T}]$ *verifies*

$$
\Pi_{s,s-1}^{(ij)} = \left(\mathcal{H}_{A_s}^{(i)} | \mathcal{H}_{C_s}^{(i)} \right) \left(J_{s-1}^{(j)} - J_{s-1}^{(ij)} \right) + \mathcal{W}_{s-1}^{(ij)} - \mathcal{W}_{s-1}^{(i)} \Pi_{s-1}^{(i)-1} \Pi_{s-1}^{(ij)}, \quad s \geq 2.
\tag{26}
$$

Finally, the coefficients $\mathcal{W}_{s-1}^{(ij)} \equiv E[w_{s-1}^{(i)} \mu_{s-1}^{(j)T}]$ *are given by*

$$
\mathcal{W}_{s-1}^{(ij)} = \Sigma_{s,s-1}^{y^{(ij)}} - \left(\mathcal{H}_{A_s}^{(i)} | \mathcal{H}_{C_s}^{(i)} \right) \left(\mathcal{H}_{B_{s-1}}^{(j)} | \mathcal{D}_{s-1}^{(j)} \right)^T, \quad s \geq 2,
\tag{27}
$$

and the matrices $\Sigma_s^{y^{(ij)}}$ *and* $\Sigma_{s,s-1}^{y^{(ij)}}$ *are given in Lemma 1.*

Proof. See Appendix B.　□

4. Illustrative Example

In this section, a numerical simulation example is used to illustrate the effectiveness of the proposed prediction and filtering algorithm in WSNs subject to deception attacks, correlated noises and random delays. The quality of the estimators is analysed considering different probability distributions for the random variables modelling the attacks and delays.

Let us consider a two-dimensional signal with the following evolution equation:

$$
x_{k+1} = F x_k + G \tau_k, \quad k \geq 0,
$$

where $F = \begin{pmatrix} 0.95 & 0.03 \\ 0.01 & 0.95 \end{pmatrix}$ and $G = \begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix}$. The additive noise $\{\tau_k, \ k \geq 0\}$ is a standard white Gaussian scalar noise and the initial signal, x_0, is a zero-mean vector with covariance matrix $E[x_0 x_0^T] = 0.1I$. Assuming that the initial signal vector and additive noise are independent, the autocovariance function of the signal is expressed as:

$$
E[x_k x_h^T] = F^{k-h} E[x_h x_h^T], \quad h \leq k,
$$

which is clearly factorised, according to Assumption 1, taking, for example, $A_k = F^k$ and $B_h^T = F^{-h} E[x_h x_h^T]$, where

$$
E[x_h x_h^T] = F E[x_{h-1} x_{h-1}^T] F^T + G G^T, \quad h \geq 1.
$$

Consider a WSN with two sensors, in which the measured signal is described by model (1). As in the theoretical model, we assume that the measurements at each sensor can experience deception attacks such as those described by (3) and that during the transmission the measurements can be delayed by one sample period; then, the received measurements in

the local processing centers are modelled by (5). To analyze the performance of the proposed estimation algorithm, the following assumptions are made on the involved processes.

Simulation assumptions:

- $\left\{ v_k^{(i)}, \ k \geq 1 \right\}$, $i = 1, 2$, are defined as $v_k^{(i)} = c^{(i)} \tau_{k-1}$, $\forall k \geq 1$.

- $\left\{ \varepsilon_k^{(i)}, \ k \geq 1 \right\}$, $i = 1, 2$, are given by $\varepsilon_k^{(i)} = e^{(i)} \varepsilon_k$, where $\{ \varepsilon_k, \ k \geq 1 \}$ is a standard Gaussian white process.

- $\left\{ \beta_k^{(i)}, \ k \geq 1 \right\}$, $i = 1, 2$, are independent sequences of independent Bernoulli random variables with $P(\beta_k^{(i)} = 1) = \overline{\beta}^{(i)}$, $\forall k \geq 1$.

- $\left\{ \gamma_k^{(i)}, \ k \geq 1 \right\}$, $i = 1, 2$, are independent sequences of independent Bernoulli random variables with $P(\gamma_k^{(i)} = 1) = \overline{\gamma}^{(i)}$, $\forall k \geq 1$.

The parameter values involved in this observation model are given in Table 1.

Table 1. Model parameter values.

	$H_k^{(i)}$	$c^{(i)}$	$e^{(i)}$	$\overline{\beta}_k^{(i)}$	$\overline{\gamma}_k^{(i)}$
Sensor 1	(0.8, 0.9)	5	0.25	0.7	0.3
Sensor 2	(0.6, 0.7)	10	0.5	0.2	0.6

From the above assumptions it is easy to find out the following properties:

(*i*) The measurement noises $\left\{ v_k^{(i)}, \ k \geq 1 \right\}$, $i = 1, 2$, have zero-mean and are correlated, with $R_k^{(ij)} = c^{(i)} c^{(j)}$.

(*ii*) The signal process and measurement noises are correlated, with the correlation function given by $E[x_k v_h^{(i)}] = 0$, $h > k$, and $E[x_k v_h^{(i)}] = c^{(i)} F^{k-h} G$, $h \leq k$, which can be expressed in separable form, as indicated in Assumption 2, taking, for example, $C_k = F^k$ and $D_h^{(i)T} = c^{(i)} F^{-h} G$, $h \leq k$.

(*iii*) The noises $\left\{ \varepsilon_k^{(i)}, \ k \geq 1 \right\}$, $i = 1, 2$, have zero-mean and variances $T_k^{(ij)} = e^{(i)} e^{(j)}$.

Performance of the local and distributed estimators. Figure 2 illustrates the local and distributed prediction and filtering error variances for $s = k - 3, k - 1, k$, of the first and second signal components. For each signal component, this figure shows, on the one hand, that the error variances of the distributed estimators are lower than those of local ones; hence, the distributed fusion estimators provide better estimations of the signal. On the other hand, this figure reveals that the performance of the estimators improves as more observations are used; that is, the error variances of the local and distributed filters are smaller than those of the predictors, and these predictors became more accurate as s increases.

Influence of deception attacks. To analyse the effect of deception attacks, let us assume different values of $\overline{\beta}^{(1)} = 0, 0.1, \ldots, 0.9$. Figure 3 shows that the distributed filtering error variances for both signal components decrease in line with the probability of success $\overline{\beta}^{(1)}$. Hence, as expected, better estimations of the signal are obtained for lower values of successful attack probability and the best ones are obtained when $\overline{\beta}^{(1)} = 0$, since in this case Sensor 1 has not been attacked. A similar analysis for different values of $\overline{\beta}^{(2)}$ as well as for the distributed prediction error variances leads us to the same conclusions.
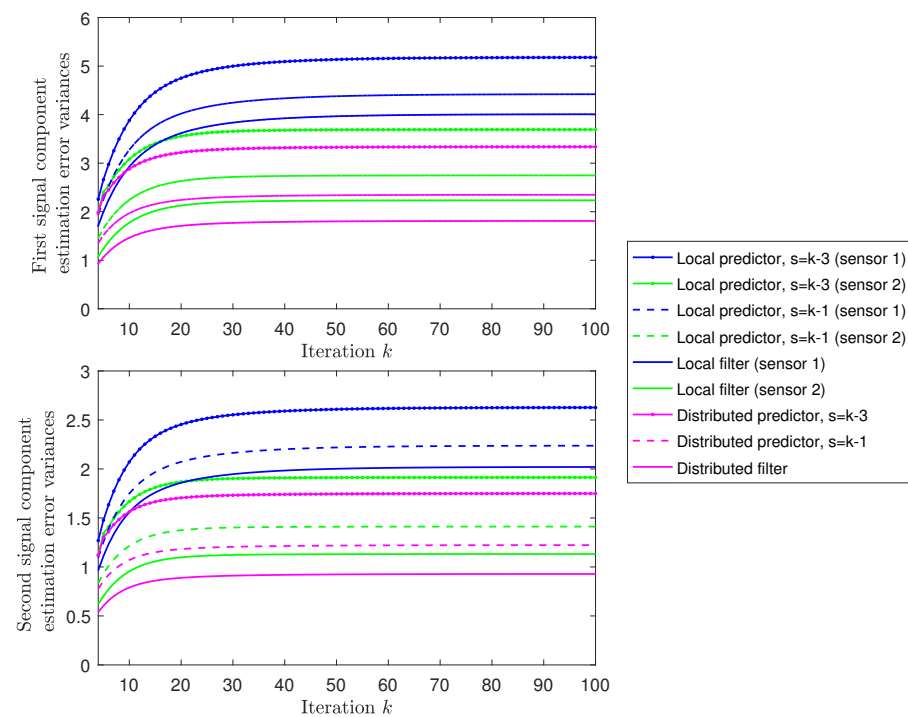
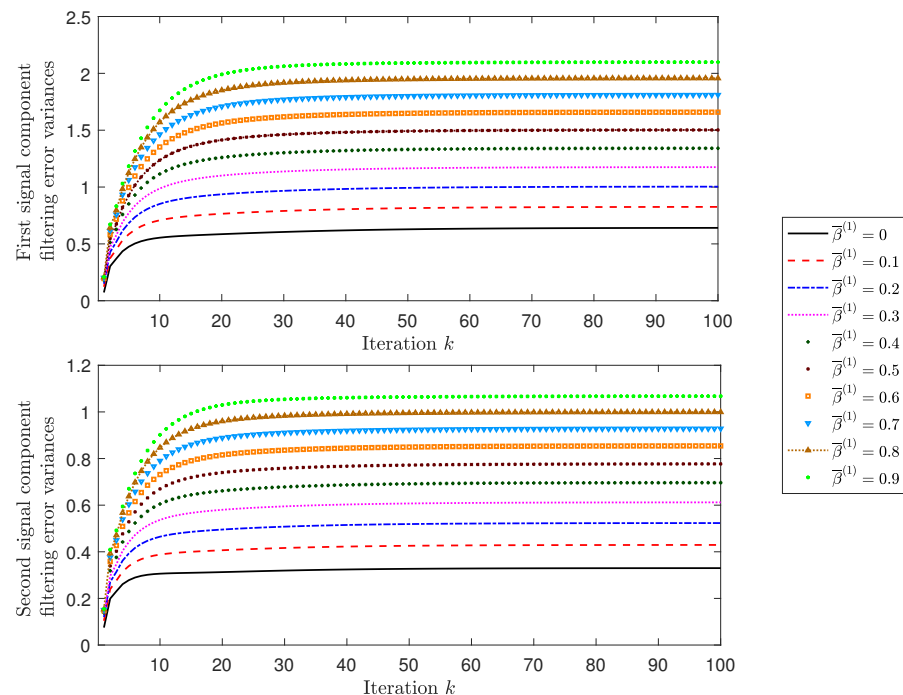**Figure 2.** Local and distributed prediction and filtering error variances.

Figure 3. Distributed filtering error variances for different values of $\overline{\beta}^{(1)}$.

Influence of transmission delays. As in our to analysis of the impact of the deception attack, here we consider that only the delay probabilities for the first sensor change, while the other probabilities remain fixed. Figure 4 shows the distributed filtering error variances when $\overline{\gamma}^{(1)} = 0, 0.1, 0.3, 0.5, 0.7$; values greater than 0.7 have not been taken into account since in this case the difference in the results is negligible. From this figure it can be concluded that, as the values of $\overline{\gamma}^{(1)}$ decrease, the distributed filtering error variances become smaller, which means that smaller probabilities of transmission delays provide better results.
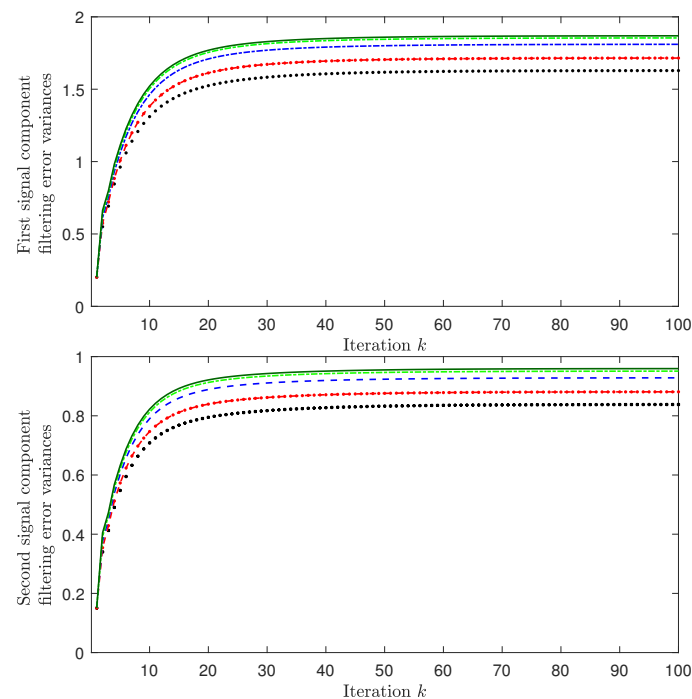
Figure 4. Distributed filtering error variances for different values of $\overline{\gamma}^{(1)}$.

Finally, Figure 5 shows the distributed filtering error variances at $k = 55$, varying jointly the different values of $\overline{\beta}^{(1)}$ and $\overline{\gamma}^{(1)}$ previously considered (actually, the results from $k = 30$ are very similar, with only negligible variation). In this figure it can be clearly seen that the variances of the filtering error increase with the probability of success of the attacks on the first sensor ($\overline{\beta}^{(1)}$) and also with the probability of delay in the measurements from the first sensor ($\overline{\gamma}^{(1)}$). Similar results are obtained by varying these probabilities in the second sensor.
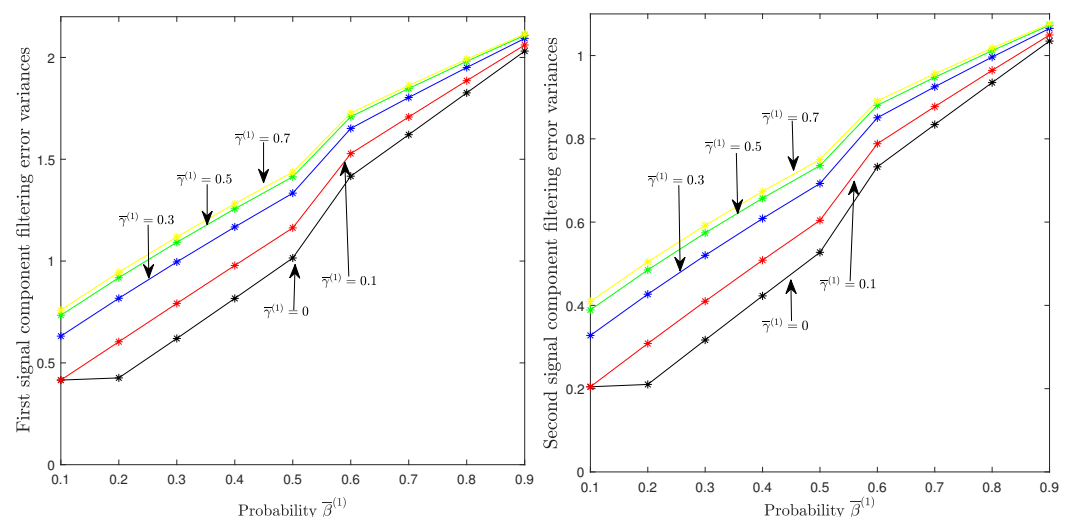

Figure 5. Distributed filtering error variances for different values at $k = 55$, of $\overline{\gamma}^{(1)}$ and $\overline{\beta}^{(1)}$.

5. Conclusions

In this paper, we investigate the distributed fusion linear estimation problem in WSNs that are affected by three types of phenomena that are very common in practical situations, namely deception attacks, correlated noises and delays during transmission. At each sensor, the adversaries may launch random attacks, falsifying the output measurements; these deception attacks are modelled by independent sequences of independent Bernoulli

random variables with different and known success probabilities. The measurement noises from different sensors are cross-correlated at the same time step and are also correlated with the signal at the same and subsequent time steps, an assumption that is fulfilled in many practical situations. During the transmission of the sensor outputs through the communication channels to their respective processing centres, these signals can suffer delays, which are modelled by Bernoulli variables describing whether the measures arrive on time or are delayed by one sampling time. For this model, we have derived a distributed fusion prediction and filtering algorithm, using only the information provided by the first and second-order moments of the processes involved in the model; that is, they do not require knowledge of the signal evolution model, although they are also applicable to the classical formulation using this model. In addition, we use a simulation example to show that the estimator accuracy is influenced by the success probabilities of deception attacks and by transmission delays.

Among further research topics of interest in this field, we investigate the stability property in the framework of estimation algorithms using covariance information, and address the problem of the vulnerability or destabilisation of systems subject to sensor deception attacks, as recently discussed in [32,33] for stealthy attacks.

Finally, it would also be of interest to design distributed fusion estimation algorithms in energy-efficient wireless sensor networks whilst considering some of the alternatives that have generated great interest in recent years; for example, event-triggered strategies [34] or other energy-efficient sensor transmission schemes, such as those described in [35].

Author Contributions: M.J.G.-L., A.H.-C. and J.L.-P. contributed equally in deriving and implementing the proposed estimation algorithms and in writing this manuscript. M.J.G.-L., A.H.-C. and J.L.-P. have read and agreed to the published version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare they have no conflict of interest.

Appendix A. Proof of Theorem 1

The expression for the LS estimators $\widehat{x}_{k/s}^{(i)}$, $s \leq k$, of the signal x_k, is obtained from general expression (6). To do so, the coefficients $\mathcal{X}_{k,h}^{(i)} = E[x_k \mu_h^{(i)T}] = E[x_k y_h^{(i)T}] - E[x_k \widehat{y}_{h/h-1}^{(i)T}]$, $1 \leq h \leq k$, must be calculated.

Using (5) for the observations and taking into account the independence assumption and that the covariance function of the signal and the correlation function of the signal and measurement noises are separable (Assumptions 2 and 3, respectively), it can easily be seen that the first expectation is given by

$$E[x_k y_h^{(i)T}] = (A_k | C_k)(\mathcal{H}_{B_h}^{(i)} | \mathcal{D}_h^{(i)})^T, \ \ 1 \leq h \leq k,$$

with $\mathcal{H}_{B_h}^{(i)}$ and $\mathcal{D}_h^{(i)}$ given by (9) and (10), respectively.

Now, we calculate $E[x_k \widehat{y}_{h/h-1}^{(i)T}]$, $h \geq 2$; using (8) for the one-stage observation predictor and general expression (6) for $\widehat{x}_{h/h-1}^{(i)}$ and $\widehat{x}_{h-1/h-1}^{(i)}$, we obtain that

$$E[x_k \widehat{y}_{h/h-1}^{(i)T}] = \sum_{l=1}^{h-1} \mathcal{X}_{k,l}^{(i)} \Pi_l^{(i)-1} \big[(1-\overline{\gamma}_h^{(i)})(1-\overline{\beta}_h^{(i)}) \mathcal{X}_{h,l}^{(i)T} H_h^{(i)T} + \overline{\gamma}_h^{(i)}(1-\overline{\beta}_{h-1}^{(i)}) \mathcal{X}_{h-1,l}^{(i)T} H_{h-1}^{(i)T} \big]$$
$$- \mathcal{X}_{k,h-1}^{(i)} \Pi_{h-1}^{(i)-1} \mathcal{W}_{h-1}^{(i)T}, \ \ 2 \leq h \leq k.$$

Hence, the coefficients are expressed as

$$\mathcal{X}_{k,h}^{(i)} = (A_k|C_k)(\mathcal{H}_{B_h}^{(i)}|\mathcal{D}_h^{(i)})^T - \sum_{l=1}^{h-1} \mathcal{X}_{k,l}^{(i)}\Pi_l^{(i)-1}[(1-\overline{\gamma}_h^{(i)})(1-\overline{\beta}_h^{(i)})\mathcal{X}_{h,l}^{(i)T}H_h^{(i)T} + \overline{\gamma}_h^{(i)}(1-\overline{\beta}_{h-1}^{(i)})\mathcal{X}_{h-1,l}^{(i)T}H_{h-1}^{(i)T}]$$
$$- \mathcal{X}_{k,h-1}^{(i)}\Pi_{h-1}^{(i)-1}\mathcal{W}_{h-1}^{(i)T}, \quad 2 \le h \le k;$$
$$\mathcal{X}_{k,1}^{(i)} = (A_k|C_k)(\mathcal{H}_{B_1}^{(i)}|\mathcal{D}_1^{(i)})^T,$$

or, equivalently,

$$\mathcal{X}_{k,h}^{(i)} = (A_k|C_k)J_h^{(i)}, \quad 1 \le h \le k,$$

where $J_h^{(i)}$ is a function satisfying

$$J_h^{(i)} = \left(\mathcal{H}_{B_h}^{(i)}|\mathcal{D}_h^{(i)}\right)^T - \sum_{l=1}^{h-1} J_l^{(i)}\Pi_l^{(i)-1}J_l^{(i)T}\left(\mathcal{H}_{A_h}^{(i)}|\mathcal{H}_{C_h}^{(i)}\right)^T - J_{h-1}^{(i)}\Pi_{h-1}^{(i)-1}\mathcal{W}_{h-1}^{(i)T}, \quad h \ge 2;$$
$$J_1^{(i)} = \left(\mathcal{H}_{B_1}^{(i)}|\mathcal{D}_1^{(i)}\right)^T. \tag{A1}$$

Therefore, (11) is obtained immediately by defining $O_s^{(i)} = \sum_{h=1}^{s} J_h^{(i)}\Pi_h^{(i)-1}\mu_h^{(i)}$, $s \ge 1$, and $O_0^{(i)} = 0$.

Again using the OPL, the error covariance matrices are expressed as $P_{k/s}^{(i)} = E[x_k x_s^T] - E[\hat{x}_{k/s}^{(i)}\hat{x}_{k/s}^{(i)T}]$, $s \le k$, and from (11), expression (12) is obtained by denoting $r_s^{(i)} \equiv E[O_s^{(i)}O_s^{(i)T}] = \sum_{h=1}^{s} J_h^{(i)}\Pi_h^{(i)-1}J_h^{(i)T}$, $s \ge 1$, and $r_0^{(i)} = 0$.

The recursive relations (13) and (14) are directly deduced from their respective definitions and expression (15) for $J_s^{(i)}$ is immediate from (A1), taking into account the definition of $r_s^{(i)}$.

The following equivalent expression for the one-stage observation predictor (8) is obtained using (11) for the predictor, $\hat{x}_{s/s-1}$, and filter, $\hat{x}_{s-1/s-1}$:

$$\hat{y}_{s/s-1}^{(i)} = (\mathcal{H}_{A_s}^{(i)}|\mathcal{H}_{C_s}^{(i)})O_{s-1}^{(i)} + \mathcal{W}_{s-1}^{(i)}\Pi_{s-1}^{(i)-1}\mu_{s-1}^{(i)}, \quad s \ge 2, \tag{A2}$$

and then, expression (16) for innovation is immediate.

Now, applying the OPL, the innovation covariance matrices are expressed as $\Pi_s^{(i)} = E[y_s^{(i)}y_s^{(i)T}] - E[\hat{y}_{s/s-1}^{(i)}y_s^{(i)T}]$ and, using (A2), we have

$$\Pi_s^{(i)} = \Sigma_s^{y^{(i)}} - (\mathcal{H}_{A_s}^{(i)}|\mathcal{H}_{C_s}^{(i)})E[O_{s-1}^{(i)}y_s^{(i)T}] - \mathcal{W}_{s-1}^{(i)}\Pi_{s-1}^{(i)-1}E[\mu_{s-1}^{(i)}y_s^{(i)T}], \quad s \ge 2;$$
$$\Pi_1^{(i)} = \Sigma_1^{y^{(i)}}.$$

Next, we calculate the two expectations that appear in the above expression. Taking into account (13) and that $\Pi_s^{(i)} = E[\mu_s^{(i)}y_s^{(i)T}]$, we have $E[O_{s-1}^{(i)}y_s^{(i)T}] = E[O_s^{(i)}y_s^{(i)T}] - J_s^{(i)}$. Now, substituting $y_s^{(i)} = \mu_s^{(i)} + \hat{y}_{s/s-1}^{(i)}$, using that $J_s^{(i)} \equiv E[O_s^{(i)}\mu_s^{(i)T}]$ and expression (15) for $J_s^{(i)}$, together with (A2) for the one-stage observation predictor, it is deduced that $E[O_s^{(i)}y_s^{(i)T}] = (\mathcal{H}_{B_s}^{(i)}|\mathcal{D}_s^{(i)})^T$ and, therefore,

$$E[O_{s-1}^{(i)}y_s^{(i)T}] = (\mathcal{H}_{B_s}^{(i)}|\mathcal{D}_s^{(i)})^T - J_s^{(i)}. \tag{A3}$$

Again, using the OPL and (A2) it is clear that

$$E[\mu_{s-1}^{(i)}y_s^{(i)T}] = E[\mu_{s-1}^{(i)}\hat{y}_{s/s-1}^{(i)T}] = ((\mathcal{H}_{A_s}^{(i)}|\mathcal{H}_{C_s}^{(i)})J_{s-1}^{(i)} + \mathcal{W}_{s-1}^{(i)})^T, \quad s \ge 2. \tag{A4}$$

Then, from (A3) and (A4), expression (17) is obtained.

Finally, we obtain (18) for the coefficients $\mathcal{W}_{s-1}^{(i)}$. Since $w_{s-1}^{(i)}$ is uncorrelated with $y_h^{(i)T}$, $h \leq s-2$, the coefficients $\mathcal{W}_{s-1}^{(i)}$ are expressed as $\mathcal{W}_{s-1}^{(i)} = E[w_{s-1}^{(i)} y_{s-1}^{(i)T}]$ and the expression of $w_{s-1}^{(i)}$ obtained from (7) leads to

$$\mathcal{W}_{s-1}^{(i)} = \Sigma_{s,s-1}^{y^{(i)}} - (1 - \overline{\gamma}_s^{(i)})(1 - \overline{\beta}_s^{(i)}) H_s^{(i)} E[x_s y_{s-1}^{(i)T}] - \overline{\gamma}_s^{(i)}(1 - \overline{\beta}_{s-1}^{(i)}) H_{s-1}^{(i)} E[x_{s-1} y_{s-1}^{(i)T}].$$

Now, using (5) for $y_{s-1}^{(i)}$ and the model assumptions, expression (18) is easily obtained.

Appendix B. Proof of Theorem 2

From (11), using that $r_s^{(ij)} \equiv E[O_s^{(i)} O_s^{(j)T}]$, expression (21) for the cross-correlation matrices is directly obtained. The recursive relation (22) is immediate using (13) together with $J_{s-1,s}^{(ij)} \equiv E[O_{s-1}^{(i)} \mu_s^{(j)T}]$ and $J_s^{(ij)} \equiv E[O_s^{(i)} \mu_s^{(j)T}]$.

Next, to facilitate obtaining Equations (23)–(27), we determine an expression for the one-stage observation predictor of $y_s^{(j)}$ based on the observations of sensor i, which will be denoted by $\hat{y}_{s/s-1}^{(j/i)}$. This expression is easily obtained from (7), taking into account the model assumptions and (11) for $\hat{x}_{s/s-1}^{(j)}$ and $\hat{x}_{s-1/s-1}^{(j)}$,

$$\hat{y}_{s/s-1}^{(j/i)} = (\mathcal{H}_{A_s}^{(j)} | \mathcal{H}_{C_s}^{(j)}) O_{s-1}^{(j)} + \mathcal{W}_{s-1}^{(ji)} \Pi_{s-1}^{(j)-1} \mu_{s-1}^{(j)}, \quad s \geq 2, \tag{A5}$$

where $\mathcal{W}_{s-1}^{(ji)} \equiv E[w_{s-1}^{(j)} \mu_{s-1}^{(i)T}]$.

Now, we obtain (23) and (24). Clearly, by the OPL, $E[O_{s-1}^{(i)} y_s^{(j)T}] = E[O_{s-1}^{(i)} \hat{y}_{s/s-1}^{(j/i)T}]$ and then $J_{s-1,s}^{(ij)} = E[O_{s-1}^{(i)} \hat{y}_{s/s-1}^{(j/i)T}] - E[O_{s-1}^{(i)} \hat{y}_{s/s-1}^{(j)T}]$. Taking into account (A2) and (A5) for $\hat{y}_{s/s-1}^{(j)}$ and $\hat{y}_{s/s-1}^{(j/i)}$, respectively, it is seen that $J_{s-1,s}^{(ij)}$ satisfies (23). Expression (24) is straightforwardly derived using (13) and $\Pi_s^{(ij)} \equiv E[\mu_s^{(i)} \mu_s^{(j)T}]$.

In order to calculate the innovation covariance matrices, $\Pi_s^{(ij)}$ and $\Pi_{s,s-1}^{(ij)}$, $i \neq j$, we apply the OPL to express these matrices as:

$$\Pi_s^{(ij)} = \Sigma_s^{y^{(ij)}} - E[\hat{y}_{s/s-1}^{(i/j)} y_s^{(j)T}] - E[\hat{y}_{s/s-1}^{(i)} \mu_s^{(j)T}], \quad s \geq 2; \quad \Pi_1^{(ij)} = \Sigma_1^{y^{(ij)}},$$
$$\Pi_{s,s-1}^{(ij)} = E[\hat{y}_{s/s-1}^{(i/j)} \mu_{s-1}^{(j)T}] - E[\hat{y}_{s/s-1}^{(i)} \mu_{s-1}^{(j)T}], \quad s \geq 2.$$

Again, using (A2) and (A5) for the one-stage observation predictors, $\hat{y}_{s/s-1}^{(i)}$ and $\hat{y}_{s/s-1}^{(i/j)}$, and taking into account (A3) and (A4) as well as the definitions of $J_{s-1,s}^{(ij)}$ and $\Pi_{s,s-1}^{(ij)}$, the expressions (25) and (26) are confirmed.

By reasoning similar to that used to obtain (18) for the coefficients $\mathcal{W}_{s-1}^{(i)}$, the expression (27) is derived for the coefficients $\mathcal{W}_{s-1}^{(ij)}$.

References

1. Ding, D.; Wang, Z.; Shen, B. Recent advances on distributed filtering for stochastic systems over sensor networks. *Int. J. Gen. Syst.* **2014**, *43*, 372–386. [CrossRef]
2. Dong, H.; Wang, Z.; Ding, S.X.; Gao, H. A Survey on Distributed Diltering and Fault Detection for Sensor Networks. *Math. Probl. Eng.* **2014**, *2014*, 858624. [CrossRef]
3. Li, W.; Wang, Z.; Wei, G.; Ma, L.; Hu, J.; Ding, D. A Survey on Multisensor Fusion and Consensus Filtering for Sensor Networks. *Discrete Dyn. Nat. Soc.* **2015**, *2015*, 683701. [CrossRef]
4. Sun, S.L.; Lin, H.; Ma, J.; Li, X. Multi-sensor distributed fusion estimation with applications in networked systems: A review paper. *Inf. Fusion.* **2017**, *38*, 122–134. [CrossRef]
5. Yang, C.; Yang, Z.; Deng, Z. Robust weighted state fusion Kalman estimators for networked systems with mixed uncertainties. *Inf. Fus.* **2019**, *45*, 246–265. [CrossRef]
6. Xia, J.; Gao, S.; Qi, X.; Zhang, J.; Li, G. Distributed cubature H-infinity information filtering for target tracking against uncertain noise statistics. *Signal Proccess.* **2020**, *177*, 107725. [CrossRef]

7. Hu, Z.; Hu, J.; Yang, G. A survey on distributed filtering, estimation and fusion for nonlinear systems with communication constraints: New advances and prospects. *Syst. Sci. Control Eng.* **2020**, *8*, 189–205. [CrossRef]

8. Chen, W.; Ding, D.; Dong, H.; Wei, G. Distributed Resilient Filtering for Power Systems Subject to Denial-of-Service Attacks. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *49*, 1688–1697. [CrossRef]

9. Xu, M.; Zhang, Y.; Zhang, D.; Chen, B. Distributed Robust Dimensionality Reduction Fusion Estimation under DoS Attacks and Uncertain Covariances. *IEEE Access.* **2021**, *9*, 10328–10337. [CrossRef]

10. Chen, B.; Ho, D.; Hu, G.; Yu, L. Secure Fusion Estimation for Bandwidth Constrained Cyber-Physical Systems under Replay Attacks. *IEEE Trans. Cybern.* **2018**, *48*, 1862–1876. [CrossRef]

11. Ding, D.; Wang, Z.; Ho, D.; Wei, G. Distributed recursive filtering for stochastic systems under uniform quantizations and deception attacks through sensor networks. *Automatica* **2017**, *78*, 231–240. [CrossRef]

12. Ma, L.; Wang, Z.; Han, Q.-L.; Lam, H.-K. Variance Constrained Distributed Filtering for Time-Varying Systems with Multiplicative Noises and Deception Attacks over Sensor Networks. *IEEE Sens. J.* **2017**, *17*, 2279–2288. [CrossRef]

13. Han, F.; Dong, H.; Wang, Z.; Li, G. Local design of distributed H_∞-consensus filtering over sensor networks under multiplicative noises and deception attacks. *Int. J. Robust Nonlinear Control.* **2019**, *29*, 2296–2314. [CrossRef]

14. Caballero-Águila, R.; Hermoso-Carazo, A.; Linares-Pérez, J. Covariance-Based Estimation for Clustered Sensor Networks subject to Random Deception Attacks. *Sensors* **2019**, *19*, 3112. [CrossRef]

15. Feng, J.; Zeng, M. Descriptor recursive estimation for multiple sensors with different delay rates. *Int. J. Control.* **2011**, *84*, 584–596. [CrossRef]

16. Li, N.; Sun, S.; Ma, J. Multi-sensor distributed fusion filtering for networked systems with different delay and loss rates. *Digit. Signal Process.* **2014**, *34*, 29–38. [CrossRef]

17. Chen, B.; Zhang, W.; Yu, L. Distributed Fusion Estimation with Missing Measurements, Random Transmission Delays and Packet Dropouts. *IEEE Trans. Automat. Control.* **2014**, *59*, 1961–1967. [CrossRef]

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

19. Caballero-Águila, R.; Hermoso-Carazo, A.; Linares-Pérez, J. Networked Fusion Filtering from Outputs with Stochastic Uncertainties and Correlated Random Transmission Delays. *Sensors* **2016**, *16*, 847. [CrossRef]

20. Caballero-Águila, R.; Hermoso-Carazo, A.; Linares-Pérez, J. Fusion Estimation from Multisensor Observations with Multiplicative Noises and Correlated Random Delays in Transmission. *Mathematics* **2017**, *5*, 45. [CrossRef]

21. García-Ligero, M.J.; Hermoso-Carazo, A.; Linares-Pérez, J. Distributed and centralized fusion estimation from multiple sensors with Markovian delays. *Appl. Math. Comput.* **2012**, *219*, 2932–2948. [CrossRef]

22. Shang, Y. Consensus seeking over Markovian switching networks with time-varying delays and uncertain topologies. *Appl. Math. Comput.* **2016**, *273*, 1234–1245. [CrossRef]

23. García-Ligero, M.J.; Hermoso-Carazo, A.; Linares-Pérez, J. Distributed Fusion Estimation with Sensor Gain Degradation and Markovian Delays. *Mathematics* **2020**, *8*, 1948. [CrossRef]

24. Li, X.; Zhu, Y.; Wang, J.; Han, C. Optimal Linear Estimation Fusion—Part I: Unified Fusion Rules. *IEEE Trans. Inf. Theory* **2003**, *49*, 2192–2208. [CrossRef]

25. Yan, L.; Li, X.; Xia, Y.; Fu, M. Optimal sequential and distributed fusion for state estimation in cross-correlated noise. *Automatica* **2013**, *49*, 3607–3612. [CrossRef]

26. Lin, H.; Sun, S. Globally optimal sequential and distributed fusion state estimation for multi-sensor systems with cross-correlated noises. *Automatica* **2019**, *101*, 128–137. [CrossRef]

27. Feng, J.; Wang, Z.; Zeng, M. Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises. *Inf. Fus.* **2013**, *14*, 78–86. [CrossRef]

28. Caballero-Águila, R.; Hermoso-Carazo, A.; Linares-Pérez, J.; Wang, Z. A new approach to distributed fusion filtering for networked systems with random parameter matrices and correlated noises. *Inf. Fus.* **2019**, *45*, 324–332. [CrossRef]

29. Wang, D.; Wang, Z.; Shen, B.; Alsaadi, F. Security-guaranteed filtering for discrete-time stochastic delayed systems with randomly occurring sensor saturations and deception attacks. *Int. J. Robust Nonlinear Control* **2017**, *27*, 1194–1208. [CrossRef]

30. Zhang, D.; Xie, J.; Ning, B. Network-based filtering for positive systems with random communication delays and deception attacks. *Neurocomputing* **2020**, *400*, 450–457. [CrossRef]

31. Caballero-Águila, R.; Hermoso-Carazo, A.; Linares-Pérez, J. A Two-Phase Distributed Filtering Algorithm for Networked Uncertain with Fading Measuremente under Deception Attacks. *Sensors* **2020**, *20*, 6445. [CrossRef] [PubMed]

32. Sui, T.; Mo, Y.; Marelli, D.; Sun, X.; Fu, M. The Vulnerability of Cyber-Physical System Under Stealthy Attacks. *IEEE Trans. Automat. Control* **2021**, *66*, 637–650. [CrossRef] [PubMed]

33. Sui, T.; Sun, X. The vulnerability of distributed state estimator under stealthy attacks. *Autoatica* **2021**, *133*, 1–12. [CrossRef]

34. Jiang, L.; Yan, L.; Xia, Y.; Guo, Q.; Fu, M.; Li, L. Distributed fusion in wireless sensor networks based on a novel event-triggered strategy. *J. Franklin Inst.* **2019**, *356*, 10315–10334. [CrossRef]

35. Sriranga, N.; Nagananda, K.G.; Blum, R.S.; Saucan, A.; Varshney, P.K. Energy-efficient Decision Fusion for Distributed Detection in Wireless Sensor Networks. In Proceedings of the 21st International Conference on Information Fusion (FUSION), Cambridge, UK, 10–13 July 2018; pp. 1541–1547. [CrossRef]