

Article

Distributed Fusion Filtering with Prediction Compensation for Multi-Sensor Systems Subject to DoS-Attack-Induced Packet Dropouts

Fengtao Hu ¹ and Jing Ma ^{1,2,*} ¹ School of Mathematics and Statistics, Heilongjiang University, Harbin 150080, China; stephan_hu@163.com² Heilongjiang Provincial Key Laboratory of the Theory and Computation of Complex Systems, Harbin 150080, China

* Correspondence: majing@hlju.edu.cn

Abstract

This paper investigates the distributed fusion estimation problem for multi-sensor cyber-physical systems (CPSs), where the communication channels from local estimators to the fusion center are subject to random packet dropouts. Packet dropouts induced by either network congestion or intermittent denial-of-service (DoS) attacks are modeled as Bernoulli random variables. When a local estimate is lost, a prediction compensation strategy is adopted at the fusion center, where the missing data are replaced by their one-step predictors. By constructing an augmented state consisting of the original state, local prediction errors, and virtual measurements, the multi-sensor system is transformed into a stochastic system with random parameter matrices and one-step autocorrelated noises. Based on the transformed system, a distributed state fusion (DSF) filter is proposed via the innovation analysis method, whose filter gain depends on the successful-reception probabilities. The stability of the proposed DSF filter is analyzed, and a sufficient condition for the existence of a steady-state filter is obtained. The steady-state gain can be pre-computed offline, thereby reducing the online computational burden. Simulation results validate the effectiveness of the proposed algorithm.

Keywords: distributed fusion; denial-of-service attacks; prediction compensation; innovation analysis approach

1. Introduction

With the rapid development of CPSs, multi-sensor information fusion is critical in applications such as autonomous driving, Unmanned Aerial Vehicle (UAV) formation, smart grid monitoring, and industrial automation [1,2]. By combining complementary sensors like radar, LiDAR, cameras, IMU, and GPS, multi-sensor systems achieve better reliability and accuracy than single-sensor systems. However, the growing reliance on open communication networks exposes these systems to a key threat: DoS attacks [3]. In practice, such attacks degrade data availability by jamming channels or causing bursty packet delays and losses [4]. In a multi-sensor fusion scenario, even an intermittent DoS attack can block sensor data streams, resulting in incomplete, outdated, or irregularly sampled measurements at the fusion center. This can seriously degrade estimation performance and may even cause the filter to diverge. Therefore, designing a fusion filter that handles DoS attacks is a key issue in state estimation.



Academic Editor: Patrick Seeling

Received: 30 July 2026

Revised: 1 September 2026

Accepted: 2 September 2026

Published: 4 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Various approaches have been proposed to model DoS attacks. For constant-rate attacks, ref. [5] uses a stochastic differential equation with Lévy jump noise to quantify their impact. In UAV confrontation scenarios, ref. [6] formulates the attack-defense interaction as a zero-sum game. For nonlinear CPSs with unknown dynamics, ref. [7] treats frequency-duration-constrained DoS attacks as intermittent channel blockages. In [8], energy-constrained attacks are modeled as multi-mode switching based on attack frequency and duration bounds. Unlike the models in [5–8], which account for adversarial intelligence and temporal correlations, another common approach in the literature relies on Bernoulli variables [9–16]. This model is mathematically equivalent to random packet losses caused by network congestion or poor channel conditions, and it does not capture the time-correlated or energy-limited features emphasized in [5–8]. In this paper, we adopt the Bernoulli-type dropout model for the distributed fusion filtering problem under DoS-induced packet dropouts.

Within this Bernoulli packet-dropout modeling framework, considerable studies have been made on filter design and state estimation. For systems suffering from random packet dropouts in both sensor-to-controller and controller-to-actuator channels, an optimal H_∞ filter is proposed in [17] where two independent sets of Bernoulli random variables are introduced to describe packet loss phenomena separately. Similarly, focusing on the packet dropout problem with two sides from a sensor to an estimator and from a controller to an actuator, ref. [18] further considers bounded random transmission delays and obtains the optimal linear estimators including a filter, predictor and smoother. However, refs. [17,18] are only applicable to single-rate sampling systems. With the widespread adoption of multi-rate sampling systems, a least-squares linear estimator for multi-rate uncertain systems under DoS attacks is developed in [10]. The compensation strategy used for packet loss in [17] is the hold-input (HI) strategy, where the input is held at the most recent data that are received. In contrast, ref. [10] adopts both the HI and prediction compensation (PC) strategies. In the PC strategy, the lost measurement packet is compensated by its predictor. Compared with the HI strategy, the PC strategy exhibits stronger robustness against successive packet dropouts and achieves higher estimation accuracy, since the PC strategy utilizes all historical received data for compensation.

In contrast to the single-sensor systems in [10,17,18], multi-sensor systems are more popular [11,12,19–22]. Centralized and distributed fusion estimators are designed for systems under different sensors with different dropout rates in [19]. This framework is further extended to networked stochastic uncertain systems subject to transmission delays and packet dropouts [20]. Avoiding computing the second-order moment of the state which is required in [19,20], a matrix-weighted distributed fusion filter is designed in [21]. More recently, several additional practical engineering problems are also considered. Outlier-resistant sequential filtering fusion is addressed under DoS attacks with quantized measurements in [11]. The protocol-based distributed secure fusion estimation problem for networked systems with unknown bounded noise under quantization is considered in [12]. In the nonlinear domain, a fusion filter is developed where the local filtering gain matrix and the fusion weighting matrix are designed by solving convex optimization problems [22]. When the noise covariances are unknown but bounded, a robust dimensionality reduction fusion estimation is studied in [23]. Under the random parameter matrices, time-correlated additive noises, deception attacks and random packet dropouts, a centralized fusion linear estimator is developed by designing two strategies in [24]. To address the packet loss problem, the HI strategy is adopted in [12,19,20,22,23], the PC strategy is employed in [21] to compensate for lost data packets, and the zero-input (ZI) strategy, which sets the measurement to zero, is utilized in [11]. Among these three compensation strategies, the PC strategy achieves the best estimation performance.

In addition to single DoS attacks, hybrid attacks in which DoS attacks coexist with other attack types such as false data injection (FDI) or deception constitute another common scenario that has been investigated in some existing studies. From an application perspective, an adaptive distributed Kalman-like filter for power systems in [25] and event-triggered distributed fusion in [13] are studied. In addition, researchers also consider other system complexities: nonlinear hybrid attacks [14]; systems with quantized measurements [15]; systems involving random parameter matrices, correlated noises and quantized measurements [16] and multi-rate sampling [26]. The DoS attacks considered in the aforementioned literature are the losses of raw measurements. However, the transmission channels from estimators to the fusion center, as well as those between estimators, are also frequent targets of attackers, and the data lost over these channels are the state estimates [27–30]. Hybrid DoS and FDI attacks along with missing measurements are considered in [27], and a consensus estimator is proposed. For linear discrete systems, ref. [28] designs a distributed fusion filter with DoS compensation, but the process noise is uncorrelated with the measurement noise, and the measurement noises are mutually uncorrelated. Multi-sensor systems with missing data are considered in [29] and a consensus filter is proposed, in which not only the loss of raw measurements from local sensors is considered, but also the reception of local estimates from neighboring nodes is taken into account, and these local estimates are treated as measurement data. More recently, ref. [30] considers a large-scale clustered networks with a hierarchical event-triggered estimator. As shown in Table 1, none of the related works [13–16,21,25–30] simultaneously consider PC compensation for lost state estimates, correlated noises, and steady-state existence. This motivates the work presented in this paper.

Table 1. Related work [13–16,21,25–30]: comparison with this paper.

Ref.	Type of Lost Data	Compensation Strategy	Noise-Correlation Assumptions	Steady-State Existence
[21]	measurements	PC	Correlated	Not discussed
[25]	measurements	PC	Uncorrelated	No
[13]	measurements	PC	Partially correlated	Not discussed
[14]	measurements	PC	Unknown but bounded	Not discussed
[15]	measurements	PC	Temporally correlated	Not discussed
[16]	measurements	PC	Correlated	Yes
[26]	measurements	PC	Uncorrelated	Yes
[27]	State estimates	PC	Uncorrelated	Boundedness
[28]	State estimates	PC	Uncorrelated	No
[29]	State estimates	ZI	Uncorrelated	Yes
[30]	State estimates	ZI	Uncorrelated	Boundedness
This paper	State estimates	PC	Correlated	Yes

“Not discussed” means that the corresponding property is not addressed in the reference; “Partially correlated” means that the process noise is uncorrelated with the measurement noises, but the measurement noises of different sensors are correlated with each other; “temporally correlated” means that the measurement noise of each sensor is temporally correlated (modeled as an autoregressive process) but is uncorrelated with all other random variables.

Motivated by the above research gap, in this paper, we investigate a distributed fusion filtering problem for multi-sensor systems, where packet losses over the transmission channels from local estimators to the fusion center are modeled as Bernoulli random variables. Such packet dropouts can be induced by network congestion or DoS attacks. Consequently, the lost data packets are the local estimates rather than the raw measurements considered in [13–16,21,25,26]. When a local estimate is missing, the PC strategy is adopted; i.e., the missing data are replaced by their one-step predictors. Furthermore, the steady-state behavior of the proposed distributed state fusion (DSF) filter is rigorously analyzed, and a sufficient condition for the existence of the steady-state filter is established.

The main contributions of this paper are summarized as follows.

- (1) This paper considers the multi-sensor system, where communication channels from local estimators to the fusion center are subject to Bernoulli packet dropouts. The system is transformed into an augmented system with random parameter matrices. The augmented state includes the original state, local prediction errors, and virtual measurements.
- (2) For the augmented systems, a re-filtering-based DSF filter adopting the PC strategy is developed via the innovation analysis approach. Although the distributed weighted fusion (DWF) filter in [28] also considers DoS attacks on the channels from estimator to the fusion center and employs the PC strategy, the proposed DSF filter can handle correlated process and measurement noises. Furthermore, its steady-state filter gain can be pre-computed offline, which reduces the online computational cost.
- (3) The proposed DSF filter is calculated using local one-step predictors that include historical raw measurements, whereas the fusion filter in [21] is based on the raw measurements that have missing data. Thus, the proposed DSF filter achieves superior fusion accuracy compared with that in [21] under the same packet reception probability. The steady-state existence is not discussed in [21].

It should be noted that the proposed DSF filter depends on the reception probabilities rather than random variables, which guarantees the existence of the steady-state filter. This may lead to a moderate accuracy loss compared with the DWF filter in [28] under low reception probabilities and independent noise, where the DWF filter achieves higher accuracy by using random variables. Under correlated noise conditions, DWF is not applicable, whereas the proposed DSF filter remains valid. This is further discussed in Section 6.

2. Problem Formulation

Consider the multi-sensor discrete linear stochastic system with l sensors:

$$x(t+1) = \Phi x(t) + \Gamma w(t), \quad (1)$$

$$y_i(t) = H_i x(t) + v_i(t), i = 1, 2, \dots, l, \quad (2)$$

where $x(t) \in R^n$ is the state, $y_i(t) \in R^{m_i}$ is the raw measurement of the i -th sensor, l denotes the number of sensors, $w(t) \in R^r$ and $v_i(t) \in R^{m_i}, i = 1, 2, \dots, l$ are the process and measurement noises of the i -th sensor, Φ, Γ and $H_i, i = 1, 2, \dots, l$ are known constant matrices with appropriate dimensions.

Assumption 1. $w(t)$ and $v_i(t), i = 1, 2, \dots, l$ are correlated white noises with the following statistical information:

$$E \left\{ \begin{bmatrix} w(t) \\ v_i(t) \end{bmatrix} \begin{bmatrix} w^T(k) & v_j^T(k) \end{bmatrix} \right\} = \begin{bmatrix} Q^w & S_j^w \\ S_i^T & Q_{ij}^v \end{bmatrix} \delta_{t,k},$$

where $Q_{ii}^v = Q_i^v$.

Assumption 2. The initial state $x(0)$ is uncorrelated with $w(t)$ and $v_i(t), i = 1, 2, \dots, l$ and it satisfies

$$E\{x(0)\} = x_0, \quad E\{[x(0) - x_0][x(0) - x_0]^T\} = P_0.$$

First, we review the following Lemma 1, which will be used in the subsequent analysis.

Lemma 1 ([31]). Given Assumptions 1 and 2, for the systems (1) and (2), the local one-step predictor $\hat{x}_i(t+1|t)$ derived from the raw measurements of the i -th sensor is expressed as

$$\hat{x}_i(t+1|t) = \Psi_i(t)\hat{x}_i(t|t-1) + L_i(t)y_i(t), \quad (3)$$

$$L_i(t) = [\Phi P_i(t|t-1)H_i^T + \Gamma S_i][H_i P_i(t|t-1)H_i^T + Q_i^v]^{-1}, \quad (4)$$

$$P_i(t+1|t) = \Psi_i(t)P_i(t|t-1)\Psi_i^T(t) + \Gamma Q^w \Gamma^T - \Gamma S_i L_i^T(t) - L_i(t)S_i^T \Gamma^T + L_i(t)Q_i^v L_i^T(t), \quad (5)$$

where $\Psi_i(t) = \Phi - L_i(t)H_i$, the initial values are $\hat{x}_i(0|-1) = x_0$ and $P_i(0|-1) = P_0$.

The local one-step predictor $\hat{x}_i(t+1|t)$ is designed based on the local measurements of the i -th sensor, $i = 1, 2, \dots, l$. The local one-step prediction is then transmitted to the fusion center. DoS attacks occur on the channels from the local estimators to the fusion center, and the packet loss phenomenon induced by DoS attacks is described by Bernoulli random variables $\gamma_i(t)$. $\gamma_i(t) = 1$ means that the predictor $\hat{x}_i(t+1|t)$ and the corresponding gain matrices are successfully received by the fusion center, while $\gamma_i(t) = 0$ indicates that $\hat{x}_i(t+1|t)$ is lost. Such packet dropouts can be caused by network congestion or, in particular, by DoS attacks modeled as Bernoulli variables. The successful data reception probabilities are $\text{Prob}\{\gamma_i(t) = 1\} = \alpha_i$ and the packet-dropout probabilities are $\text{Prob}\{\gamma_i(t) = 0\} = 1 - \alpha_i$, where $0 \leq \alpha_i \leq 1$, $i = 1, 2, \dots, l$. The following stochastic independence properties of $\gamma_i(t)$, $i = 1, 2, \dots, l$, are assumed throughout this paper: (a) for each sensor, $\gamma_i(t)$ at different time instants are mutually independent; (b) $\gamma_i(t)$ and $\gamma_j(t)$ are mutually independent at the same time instant for $i \neq j$; (c) the random variables $\gamma_i(t)$, $i = 1, 2, \dots, l$, are independent of other random variables in the considered system.

It should be pointed out that the Bernoulli model treats DoS-induced packet losses as independent random dropouts. Accordingly, it is not applicable to modeling intelligent and time-correlated attack scenarios.

Remark 1. The reception probabilities α_i , $i = 1, 2, \dots, l$, can be obtained in practice via system identification. For example, we may analyze historical packet-loss records or employ the recursive extended least-squares algorithm for the online identification of packet reception rates using real-time received observation data (see [32] for the complete identification algorithm). Thus, it is assumed that the reception probabilities α_i are pre-identified and available. If the assumed α_i deviate from the true values, the covariance matrices obtained from Theorem 2 will no longer coincide with the actual filtering error covariances, and the performance degradation becomes more significant as the deviation increases. The sensitivity of the DSF filter to reception probability mismatches is examined via simulations in Section 6.

For ease of understanding, the overall communication and estimation structure is illustrated in Figure 1. The aim of this paper is to design the DSF filter $\hat{x}_o(t|t)$ based on the received local one-step predictors $\hat{x}_i(t+1|t)$ with the gain matrices $L_i(t)$ in the fusion center.

In the next section, we will transform the system in the fusion center.

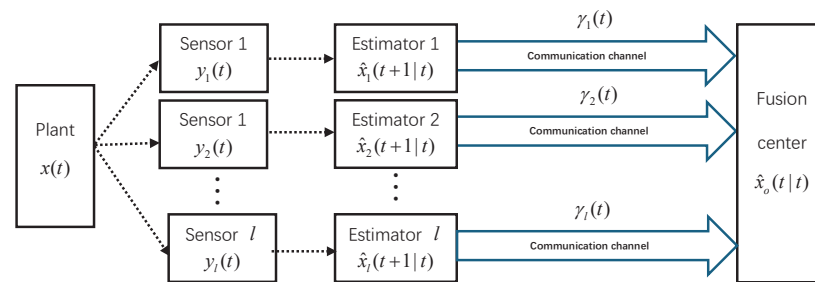


Figure 1. Communication and estimation setup of the multi-sensor system under DoS attacks.

3. System Transformation

Let $\bar{y}_i(t) = \hat{x}_i(t+1|t) = x(t+1) - \tilde{x}_i(t+1|t)$; then, the locally computed predictor $\hat{x}_i(t+1|t)$ can be regarded as a virtual measurement for the state $x(t)$ (similar to [31]). The prediction error $\tilde{x}_i(t+1|t)$ is zero-mean under Assumptions 1 and 2, and its covariance is determined by the local Kalman filter. Therefore, the above relation satisfies the standard form of a measurement equation, namely, a linear function of the state together with zero-mean noise. In this formulation, the local prediction error $-\tilde{x}_i(t+1|t)$ acts as an equivalent ‘measurement noise’. Thus, a new virtual measurement is constructed as

$$z_i(t) = \gamma_i(t)\bar{y}_i(t) + [1 - \gamma_i(t)]\hat{y}_i(t|t-1), i = 1, 2, \dots, l, \quad (6)$$

where $\hat{y}_i(t|t-1) = \text{proj}[\bar{y}_i(t) | z_i(1), z_i(2), \dots, z_i(t-1)]$, and $\text{proj}[\bullet | \circ]$ denotes the projection of $\bullet$ onto the linear space $L(\circ)$.

Remark 2. According to the projection theory [33], we have $\hat{x}_i(t+1|t) \in L(y_i(1), y_i(2), \dots, y_i(t))$, i.e., $\bar{y}_i(t) \in L(y_i(1), y_i(2), \dots, y_i(t))$, $i = 1, 2, \dots, l$. From (6), we have $z_i(t) \in L(\bar{y}_i(1), \bar{y}_i(2), \dots, \bar{y}_i(t))$. Consequently, there is $L(z_i(1), z_i(2), \dots, z_i(t)) \subseteq L(\bar{y}_i(1), \bar{y}_i(2), \dots, \bar{y}_i(t)) \subseteq L(y_i(1), y_i(2), \dots, y_i(t))$. From systems (1) and (2) and Assumption 1, we can obtain $w(k) \perp L(y_i(1), y_i(2), \dots, y_i(t))$, $v_j(k) \perp L(y_i(1), y_i(2), \dots, y_i(t))$, $k > t$, $i, j = 1, 2, \dots, l$. Hence, $w(k) \perp L(z_i(1), z_i(2), \dots, z_i(t))$ and $v_j(k) \perp L(z_i(1), z_i(2), \dots, z_i(t))$, $k > t$, $i, j = 1, 2, \dots, l$.

From (2) and (3) in Lemma 1, measurement input $\bar{y}_i(t)$ can be rewritten as

$$\bar{y}_i(t) = \Phi\bar{y}_i(t-1) + L_i(t)H_i\tilde{x}_i(t|t-1) + L_i(t)v_i(t). \quad (7)$$

When the DoS attacks occurs at time instant t and the $t-1$ time instant is free of DoS attacks, according to the projection theory [33], from (7), we have the one-step predictor of the measurement input $\bar{y}_i(t)$, $i = 1, 2, \dots, l$

$$\begin{aligned} \hat{\bar{y}}_i(t|t-1) &= \text{proj}[\bar{y}_i(t) | z_i(0), z_i(1), \dots, z_i(t-1)] \\ &= \Phi z_i(t-1) + \text{proj}[L_i(t)H_i\tilde{x}_i(t|t-1) + L_i(t)v_i(t) | \\ &\quad z_i(0), z_i(1), \dots, z_i(t-1)]. \end{aligned} \quad (8)$$

From the orthogonality of projection $\tilde{x}_i(t|t-1) \perp L(y_i(0), y_i(1), \dots, y_i(t-1))$ and Remark 2, it can be derived that $\tilde{x}_i(t|t-1) \perp L(z_i(0), z_i(1), \dots, z_i(t-1))$ and $v_i(t) \perp L(z_i(0), z_i(1), \dots, z_i(t-1))$. Therefore, from (8), we can have

$$\hat{\bar{y}}_i(t|t-1) = \Phi z_i(t-1). \quad (9)$$

In fact, a compensator using all received local estimates can yield higher accuracy. However, it would introduce cross-correlations among local estimation errors, which would complicate the filter derivation. The compensator in (9) has the simple form $\Phi z_i(t-1)$ and

relies only on the history of each local estimator. We adopt it as a design choice to keep the structure simple and the derivation feasible.

(9) is also applicable to consecutive packet dropouts. Without loss of generality, it is supposed that DoS attacks induce two successive packet loss events; i.e., packet loss occurs at both time instants $t - 1$ and t but not at time instant $t - 2$. From (6) and (9), we have

$$z_i(t-1) = \hat{y}_i(t-1|t-2) = \Phi z_i(t-2), \quad (10)$$

thus, from (7), it can be obtained that

$$\hat{y}_i(t|t-1) = \Phi^2 z_i(t-2) = \Phi z_i(t-1). \quad (11)$$

By (1), (9) and the estimation error definition $\tilde{x}_i(t+1|t) = x(t+1) - \hat{x}_i(t+1|t)$, the virtual measurement $z_i(t)$ can be rewritten as

$$z_i(t) = \gamma_i(t)\Phi x(t) - \gamma_i(t)\tilde{x}_i(t+1|t) + [1 - \gamma_i(t)]\Phi z_i(t-1) + \gamma_i(t)\Gamma w(t). \quad (12)$$

It follows from (1)–(3) that

$$\tilde{x}_i(t+1|t) = x(t+1) - \hat{x}_i(t+1|t) = \Psi_i(t)\tilde{x}_i(t|t-1) + \Gamma w(t) - L_i(t)v_i(t). \quad (13)$$

From (1), (12) and (13), it can be obtained that $x(t)$, $z_i(t-1)$ and $\tilde{x}_i(t+1|t)$ are temporally coupled. Therefore, these three components are augmented into a new state vector $X(t)$ to avoid complicated derivations resulting from temporal coupling. Based on their respective recursive equations, the augmented state Equation (14) can then be constructed at the fusion center.

Let $X(t) = [x^T(t) \ \tilde{x}_1^T(t+1|t) \cdots \tilde{x}_l^T(t+1|t) \ z_1^T(t-1) \cdots z_l^T(t-1)]^T$ be an augmented state and $z(t) = [z_1^T(t) \ z_2^T(t) \cdots z_l^T(t)]^T$ be an augmented measurement. In the augmented state vector, $x(t)$ is the original system state, $\tilde{x}_i(t+1|t)$ is the one-step prediction error of the i -th local estimator, and $z_i(t-1)$ is the virtual measurement. Using (1), (12) and (13), the systems (1) and (2) can be transformed at the fusion center into the following new system

$$X(t+1) = \bar{\Phi}(t)X(t) + \bar{w}(t), \quad (14)$$

$$z(t) = H(t)X(t) + D(t)w(t), \quad (15)$$

$$\bar{w}(t) = \Gamma_0(t)w(t) + \Gamma_1 w(t+1) + \Gamma_2(t+1)v(t+1), \quad (16)$$

where

$$\bar{\Phi}(t) = \begin{bmatrix} \Phi & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Psi(t+1) & \mathbf{0} \\ \gamma(t)\Phi^{(1)} & -\gamma(t) & [\mathbf{I} - \gamma(t)]\Phi^{(2)} \end{bmatrix},$$

$$H(t) = [\gamma(t)\Phi^{(1)} \quad -\gamma(t) \quad [\mathbf{I} - \gamma(t)]\Phi^{(2)}], \quad D(t) = \gamma(t)\Gamma^{(1)},$$

$$\Gamma_0(t) = \begin{bmatrix} \Gamma \\ \mathbf{0} \\ \gamma(t)\Gamma^{(1)} \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} \mathbf{0} \\ \Gamma^{(1)} \\ \mathbf{0} \end{bmatrix}, \quad \Gamma_2(t) = \begin{bmatrix} \mathbf{0} \\ -L^a(t) \\ \mathbf{0} \end{bmatrix}, \quad v(t) = \begin{bmatrix} v_1(t) \\ \vdots \\ v_l(t) \end{bmatrix},$$

$$\Psi(t) = \text{diag}(\Psi_1(t), \Psi_2(t), \dots, \Psi_l(t)), \quad \gamma(t) = \text{diag}(\gamma_1(t)\mathbf{I}, \gamma_2(t)\mathbf{I}, \dots, \gamma_l(t)\mathbf{I}),$$

$$\Phi^{(1)} = [\Phi^T \quad \dots \quad \Phi^T]^T, \quad \Phi^{(2)} = \text{diag}(\Phi, \Phi, \dots, \Phi),$$

$$\Gamma^{(1)} = [\Gamma^T \quad \dots \quad \Gamma^T]^T, \quad L^a(t) = \text{diag}(L_1(t), L_2(t), \dots, L_l(t)),$$

$\mathbf{0}$ is a zero matrix of appropriate dimensions, and $\mathbf{I}$ denotes the identity matrix of appropriate dimension. For clarity, the dimensions of the main quantities in (14)–(16) are summarized in Table 2.

Table 2. Dimensions of the variables and matrices in the transformed system (14)–(16).

Symbol	Definition	Dimension
$X(t)$	Augmented state in (14)	$(2l+1)n \times 1$
$z(t)$	Augmented measurement in (15)	$ln \times 1$
$\bar{\Phi}(t)$	Transition matrix of the augmented system	$(2l+1)n \times (2l+1)n$
$H(t)$	Measurement matrix of the augmented system	$ln \times (2l+1)n$
$D(t)$	Coefficient matrix of $w(t)$ in (15)	$ln \times r$
$\Gamma_0(t)$	Coefficient matrix of $w(t)$ in $\bar{w}(t)$	$(2l+1)n \times r$
Γ_1	Coefficient matrix of $w(t+1)$ in $\bar{w}(t)$	$(2l+1)n \times r$
$\Gamma_2(t+1)$	Coefficient matrix of $v(t+1)$ in $\bar{w}(t)$	$(2l+1)n \times m$

Here, $m = \sum_{i=1}^l m_i$ is the total dimension of the stacked measurement noise $v(t)$.

To facilitate understanding of the system transformation process, a simple schematic diagram of the derivation procedure is provided in Figure 2.

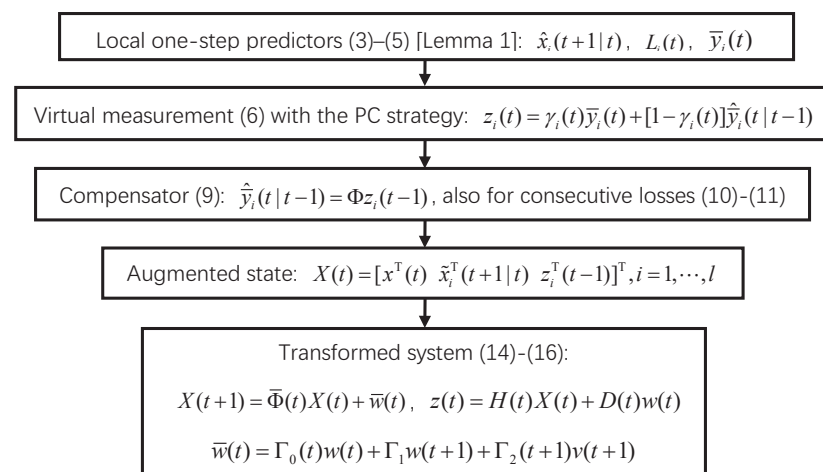


Figure 2. Schematic diagram illustrating the derivation flow of the system transformation procedure.

In the next section, a DSF filter is designed based on the transformed new system (14)–(16).

4. Distributed State Fusion Filter

In this section, the transformed new systems (14)–(16) are re-filtered via the innovation analysis approach in the fusion center, and the DSF filter is obtained. First of all, for ease of presentation, the following Remarks are given.

Remark 3.

$$\begin{aligned}
 & \begin{bmatrix} a_1 \mathbf{I} & & \\ & \ddots & \\ & & a_k \mathbf{I} \end{bmatrix}_{r \times r} (b_{ij})_{r \times s} \begin{bmatrix} c_1 \mathbf{I} & & \\ & \ddots & \\ & & c_l \mathbf{I} \end{bmatrix}_{s \times s} \\
 &= \begin{bmatrix} a_1 c_1 \mathbf{1}_{m_1 n_1} & a_1 c_2 \mathbf{1}_{m_1 n_2} & \cdots & a_1 c_l \mathbf{1}_{m_1 n_l} \\ a_2 c_1 \mathbf{1}_{m_2 n_1} & a_2 c_2 \mathbf{1}_{m_2 n_2} & \cdots & a_2 c_l \mathbf{1}_{m_2 n_l} \\ \vdots & \vdots & \ddots & \vdots \\ a_k c_1 \mathbf{1}_{m_k n_1} & a_k c_2 \mathbf{1}_{m_k n_2} & \cdots & a_k c_l \mathbf{1}_{m_k n_l} \end{bmatrix}_{r \times s} \odot (b_{ij})_{r \times s}, \quad (17)
 \end{aligned}$$

where $\mathbf{1}_{m_k n_l}$ denotes an $m_k \times n_l$ all-ones matrix. $\odot$ is the Hadamard product with the definition $B \odot C = (b_{ij} \times c_{ij})_{r \times s}$, $B = (b_{ij})_{r \times s}$ and $C = (c_{ij})_{r \times s}$ are $r \times s$ matrices.

Remark 4. Based on the probability distribution of the Bernoulli random variable $\gamma_i(t)$, we have that $E[\gamma_i(t)] = \alpha_i$, $E[\gamma_i^2(t)] = \alpha_i$, $E[\gamma_i(t)\gamma_j(t)] = \alpha_i\alpha_j$, $i \neq j$, $i, j = 1, 2, \dots, l$.

Remark 5. Regarding systems (14)–(16), the following results hold.

$$\hat{\Phi}(t) = E[\Phi(t)] = \begin{bmatrix} \Phi & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Psi(t+1) & \mathbf{0} \\ \alpha\Phi^{(1)} & -\alpha & [\mathbf{I} - \alpha]\Phi^{(2)} \end{bmatrix}, \alpha = \text{diag}(\alpha_1\mathbf{I}, \alpha_2\mathbf{I}, \dots, \alpha_l\mathbf{I}),$$

$$\hat{H} = E[H(t)] = \begin{bmatrix} \alpha\Phi^{(1)} & -\alpha & [\mathbf{I} - \alpha]\Phi^{(2)} \end{bmatrix}, \hat{D} = E[D(t)] = \alpha\Gamma^{(1)}, \hat{\Gamma}_0 = E[\Gamma_0(t)] = \begin{bmatrix} \Gamma \\ \mathbf{0} \\ \alpha\Gamma^{(1)} \end{bmatrix}.$$

And

$$\tilde{\Phi}(t) = \Phi(t) - \hat{\Phi}(t) = \Lambda(t)\hat{\Phi}^{(1)}, \quad \tilde{H}(t) = H(t) - \hat{H} = [\gamma(t) - \alpha]\hat{H}^{(1)}, \\ \tilde{D}(t) = D(t) - \hat{D} = [\gamma(t) - \alpha]\Gamma^{(1)}, \quad \tilde{\Gamma}_0(t) = \Gamma_0(t) - \hat{\Gamma}_0 = \Lambda(t)\hat{\Gamma}_0^{(1)},$$

where

$$\Lambda(t) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \gamma(t) - \alpha \end{bmatrix}, \quad \hat{\Phi}^{(1)} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \Phi^{(1)} & -\mathbf{I} & -\Phi^{(2)} \end{bmatrix}, \\ \hat{H}^{(1)} = \begin{bmatrix} \Phi^{(1)} & -\mathbf{I} & -\Phi^{(2)} \end{bmatrix}, \quad \hat{\Gamma}_0^{(1)} = \begin{bmatrix} \mathbf{0} \\ \Gamma^{(1)} \end{bmatrix}.$$

It can be observed that $E[\tilde{\Phi}(t)] = \mathbf{0}$, $E[\tilde{H}(t)] = \mathbf{0}$, $E[\tilde{D}(t)] = \mathbf{0}$, $E[\tilde{\Gamma}_0(t)] = \mathbf{0}$.

Under the Assumption 1, the noise statistics for systems (14)–(16) satisfy as follows.

Lemma 2. According to the Assumption 1, the system noises satisfy the following statistical characteristics.

$$E[\bar{w}(t)] = \mathbf{0}, \quad E\left\{\begin{bmatrix} \bar{w}(t) \\ w(t) \end{bmatrix} \begin{bmatrix} \bar{w}^T(t) & w^T(t) \end{bmatrix}\right\} = \begin{bmatrix} Q^{\bar{w}}(t) & Q^{\bar{w}w} \\ Q^{w\bar{w}} & Q^w \end{bmatrix},$$

$$Q^{\bar{w}}(t) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{\hat{\Gamma}_0^{(1)} Q^w [\hat{\Gamma}_0^{(1)}]^T\} + \hat{\Gamma}_0 Q^w \hat{\Gamma}_0^T + \Gamma_1 Q^w \Gamma_1^T + \Gamma_2(t+1) Q^v \Gamma_2^T(t+1) \\ + \Gamma_1 S \Gamma_2^T(t+1) + \Gamma_2(t+1) S^T \Gamma_1^T, \quad (18)$$

$$Q^{\bar{w}w} = \hat{\Gamma}_0 Q^w, \quad (19)$$

where $Q^{w\bar{w}} = (Q^{\bar{w}w})^T$, $S = [S_1 \ \dots \ S_l]$, $Q^v = (Q_{ij}^v)_{m \times m}$, $i, j = 1, 2, \dots, l$ is a block matrix with $m = \sum_{i=1}^l m_i$, $A = \text{diag}((\alpha_1 - \alpha_1^2)\mathbf{1}_{m_1}, \dots, (\alpha_l - \alpha_l^2)\mathbf{1}_{m_l})$.

Proof. See the Appendix A. $\square$

Remark 6. From (12), it can be seen that $X(t) \in L(w(t), v(t), w(t-1), v(t-1), \dots, w(1), v(1), w(0))$; under Assumption 1, we have $w(j) \perp X(t)$ and $v(j) \perp X(t)$, $j > t$. Since variable $\gamma(t)$ is uncorrelated with all other variables, substituting (14) into $Q^{X\bar{w}}(t) = E[X(t)\bar{w}^T(t)]$ under Assumption 1, there is

$$Q^{X\bar{w}}(t) = \Gamma_1 Q^w \hat{\Gamma}_0^T + \Gamma_2(t) S^T \hat{\Gamma}_0^T.$$

Similarly, substituting (15) into $Q^{Xw}(t) = E[X(t)w^T(t)]$, there is

$$Q^{Xw}(t) = \Gamma_1 Q^w + \Gamma_2(t) S^T,$$

$$\text{and } Q^{X\bar{w}}(t) = Q^{Xw}(t) \hat{\Gamma}_0^T.$$

Before deriving the DSF filter, the following Theorem 1 is first presented.

Theorem 1. Under Assumptions 1 and 2 and Lemma 2, we can obtain the following second-order moment $q(t) = E[X(t)X^T(t)]$ of the state $X(t)$ in system (14).

$$\begin{aligned} q(t+1) &= \hat{\Phi}(t)q(t)\hat{\Phi}^T(t) + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{\hat{\Phi}^{(1)} q(t) [\hat{\Phi}^{(1)}]^T\} + \hat{\Phi}(t)Q^{X\bar{w}}(t) + Q^{\bar{w}X}(t)\hat{\Phi}^T(t) \\ &\quad + Q^{\bar{w}}(t) + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{\hat{\Phi}^{(1)} Q^{Xw}(t) [\hat{\Gamma}_0^{(1)}]^T + \hat{\Gamma}_0^{(1)} Q^{wX}(t) [\hat{\Phi}^{(1)}]^T\}, \end{aligned} \quad (20)$$

the initial value is $q(0) = \begin{bmatrix} x_0 x_0^T + P_0 & q^{x\bar{x}}(0) & q^{x\hat{x}}(0) \\ q^{\bar{x}x}(0) & q^{\bar{x}\bar{x}}(0) & q^{\bar{x}\hat{x}}(0) \\ q^{\hat{x}x}(0) & q^{\hat{x}\bar{x}}(0) & q^{\hat{x}\hat{x}}(0) \end{bmatrix}$, where $q^{x\bar{x}}(0) = (q^{x\bar{x}_i})_{n \times nl}$, $q^{x\hat{x}}(0) = (q^{x\hat{x}_i})_{n \times nl}$, $q^{\bar{x}\bar{x}}(0) = (q^{\bar{x}_i\bar{x}_j})_{nl \times nl}$, $q^{\bar{x}\hat{x}}(0) = (q^{\bar{x}_i\hat{x}_j})_{nl \times nl}$, $q^{\hat{x}\hat{x}}(0) = (q^{\hat{x}_i\hat{x}_j})_{nl \times nl}$, $q^{\bar{x}\bar{x}}(0) = \mathbf{0}$, $q^{x\bar{x}_i} = P_0 \Phi^T$, $q^{x\hat{x}_i} = x_0 x_0^T$, $q^{\bar{x}_i\bar{x}_j} = \Phi P_0 \Phi^T + \Gamma Q^w \Gamma^T$, $q^{\hat{x}_i\hat{x}_j} = x_0 x_0^T$, $i, j = 1, 2, \dots, l$.

Proof. See Appendix B. $\square$

In the following, a fusion filter is derived based on the noise statistical properties in Lemma 2, and the second-order moment $q(t)$ is given in Theorem 1.

Theorem 2. Under Assumptions 1 and 2 for the system (14)–(16), the fusion filter $\hat{X}(t|t)$ and the one-step predictor are

$$\hat{X}(t|t) = \hat{X}(t|t-1) + K(t)\varepsilon(t), \quad (21)$$

$$\hat{X}(t+1|t) = \hat{\Phi}(t)\hat{X}(t|t-1) + L(t)\varepsilon(t). \quad (22)$$

The innovation process $\varepsilon(t)$ and the innovation covariance matrix $Q^\varepsilon(t)$ are

$$\varepsilon(t) = z(t) - \hat{H}\hat{X}(t|t-1), \quad (23)$$

$$\begin{aligned} Q^\varepsilon(t) &= \hat{H}P(t|t-1)\hat{H}^T + \hat{D}Q^w\hat{D}^T + \hat{H}Q^{Xw}(t)\hat{D}^T + \hat{D}Q^{wX}(t)\hat{H}^T + A \\ &\quad \odot \left\{ \hat{H}^{(1)}q(t)[\hat{H}^{(1)}]^T + \Gamma^{(1)}Q^w[\Gamma^{(1)}]^T + \hat{H}^{(1)}Q^{Xw}(t)[\Gamma^{(1)}]^T + \Gamma^{(1)}Q^{wX}(t)[\hat{H}^{(1)}]^T \right\}, \end{aligned} \quad (24)$$

where the innovation process $\varepsilon(t)$ denotes the one-step prediction error of the measurement $z(t)$, i.e., $\varepsilon(t) = z(t) - \text{proj}[z(t) | z(1), z(2), \dots, z(t-1)]$. The filter gain matrix $K(t)$ and the predictor gain matrix $L(t)$ are

$$K(t) = [P(t|t-1)\hat{H}^T + Q^{Xw}(t)\hat{D}^T][Q^\varepsilon(t)]^{-1}, \quad (25)$$

$$\begin{aligned} L(t) &= \left\{ \begin{bmatrix} \mathbf{0} \\ A \end{bmatrix} \odot \left\{ \hat{\Phi}^{(1)}q(t)[\hat{H}^{(1)}]^T + \hat{\Phi}^{(1)}Q^{Xw}(t)[\Gamma^{(1)}]^T + \hat{\Gamma}_0^{(1)}Q^{wX}(t)[\hat{H}^{(1)}]^T + \hat{\Gamma}_0^{(1)}Q^w \right. \right. \\ &\quad \times [\Gamma^{(1)}]^T \left. \right\} + \hat{\Phi}(t)P(t|t-1)\hat{H}^T + \hat{\Phi}(t)Q^{Xw}(t)\hat{D}^T + \hat{\Gamma}_0 Q^{wX}(t)\hat{H}^T + \hat{\Gamma}_0 Q^w\hat{D}^T \left. \right\} \\ &\quad \times [Q^\varepsilon(t)]^{-1}. \end{aligned} \quad (26)$$

The filtering error covariance matrix $P(t|t)$ and the prediction error covariance matrix $P(t+1|t)$ are

$$P(t|t) = P(t|t-1) - K(t)Q^e(t)K^T(t), \quad (27)$$

$$\begin{aligned} P(t+1|t) &= \hat{\Phi}(t)P(t|t-1)\hat{\Phi}^T(t) + \hat{\Phi}(t)Q^{X\bar{w}}(t) + Q^{\bar{w}X}(t)\hat{\Phi}^T(t) + Q^{\bar{w}}(t) - L(t)Q^e(t)L^T(t) \\ &\quad + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{ \hat{\Phi}^{(1)} q(t) [\hat{\Phi}^{(1)}]^T + \hat{\Phi}^{(1)} Q^{Xw}(t) [\hat{\Gamma}_0^{(1)}]^T + \hat{\Gamma}_0^{(1)} Q^{wX}(t) [\hat{\Phi}^{(1)}]^T \}. \end{aligned} \quad (28)$$

The initial values are $\hat{X}(0|-1) = [x_0^T \ \mathbf{0} \ \cdots \ \mathbf{0} \ x_0^T \ \cdots \ x_0^T]^T$ and $P(0|-1) = \begin{bmatrix} P_0 & P^{x\bar{x}}(0|-1) & \mathbf{0} \\ P^{\bar{x}x}(0|-1) & P^{\bar{x}\bar{x}}(0|-1) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$, where $P^{x\bar{x}}(0|-1) = (P^{x\bar{x}_i})_{n \times nl}$, $P^{\bar{x}\bar{x}}(0|-1) = (P^{\bar{x}_i\bar{x}_j})_{nl \times nl}$, $P^{x\bar{x}_i} = P_0\Phi^T$, $P^{\bar{x}_i\bar{x}_j} = \Phi P_0\Phi^T + \Gamma Q^w\Gamma^T$, $P^{\bar{x}x}(0|-1) = [P^{x\bar{x}}(0|-1)]^T$, $i, j = 1, 2, \dots, l$.

Proof. See Appendix C. $\square$

Remark 7. For systems (1) and (2) under Assumptions 1 and 2, the distributed fusion filter $\hat{x}_o(t|t) = [\mathbf{I}_n \ \mathbf{0}] \hat{X}(t|t)$ and the filtering error variance matrix $P_o(t|t) = [\mathbf{I}_n \ \mathbf{0}] P(t|t) [\mathbf{I}_n \ \mathbf{0}]^T$ can be obtained.

Remark 8. It should be pointed out that the augmented state vector $X(t)$ is formed by augmenting $x(t)$, $\bar{x}_i(t+1|t)$ and $z_i(t-1)$, $i = 1, 2, \dots, l$. When the number of sensors l is large, the aggregate computational load across all local estimators grows, which may pose a non-negligible burden for real-time implementation. Moreover, transmitting both local predictor $\hat{x}_i(t+1|t)$ and gain matrices $L_i(t)$ additionally increases the communication burden. However, in this paper, the steady-state case is considered; i.e., $L_i(t)$ in Lemma 1, $P(t+1|t)$, $K(t)$ and $L(t)$ in Theorem 2 converge to constant matrices. Consequently, the local gain matrices and the fusion gain matrices can be pre-computed offline, which significantly reduces the computational burden and avoids the need to transmit time-varying gain matrices $L_i(t)$ from the local estimators to the fusion center, thus further reducing the communication burden.

To better demonstrate the implementing procedure of the proposed DSF filter algorithm, the computing steps are listed as follows:

Step 1 (Set the initial values). $\hat{x}_i(0|-1) = x_0$, $P_i(0|-1) = P_0$, $i = 1, 2, \dots, l$, $\hat{X}(0|-1)$, $P(0|-1)$ and $q(0)$.

Step 2 (Local predictor update). At each time t , for each sensor i , compute the gain $L_i(t)$ by (4), the local predictor $\hat{x}_i(t+1|t)$ by (3), and the local prediction error covariance $P_i(t+1|t)$ by (5). Transmit $\hat{x}_i(t+1|t)$ and $L_i(t)$ to the fusion center after calculation.

Step 3 (Packet reception/DoS realization). Observe the Bernoulli variables $\gamma_i(t)$: $\gamma_i(t) = 1$ means that $\hat{x}_i(t+1|t)$ is received by the fusion center, while $\gamma_i(t) = 0$ means that a DoS attack occurs and the packet is lost.

Step 4 (Prediction compensation). Compute the virtual measurement $z_i(t)$ by (6) and (9), and form $z(t) = [z_1^T(t) \ \cdots \ z_l^T(t)]^T$.

Step 5 (Innovation computation). Compute the innovation $\varepsilon(t)$ by (23).

Step 6 (Gain and covariance computation). Compute $q(t+1)$ by (20), $Q^e(t)$ by (24), and the gains $K(t)$ by (25) and $L(t)$ by (26).

Step 7 (Fusion update). Compute $\hat{X}(t|t)$ by (21), $\hat{X}(t+1|t)$ by (22), $P(t|t)$ by (27), and $P(t+1|t)$ by (28); output the fused estimate $\hat{x}_o(t|t) = [\mathbf{I}_n \ \mathbf{0}] \hat{X}(t|t)$.

Step 8 (Let $t = t + 1$). Proceed to Step 2.

Each step above is implemented at each time t .

In the next section, the stability and steady-state behavior of the proposed DSF filter will be analyzed.

5. Stability Analysis and Steady-State Filter

Now, we will study the stability of the DSF filter obtained in Section 4.

First of all, to ensure the existence of a steady state for the local predictor, we make the following Assumption 3.

Assumption 3. (Φ, H_i) is detectable, and $(\Phi - \Gamma S_i(Q_i^v)^{-1} H_i, \bar{R}_i)$ is stabilizable, where $\bar{R}_i \bar{R}_i^T = \Gamma Q^w \Gamma^T - \Gamma S_i(Q_i^v)^{-1} S_i^T \Gamma^T$.

From [33], we know that for systems (1) and (2) under Assumption 3, all local one-step predictors have steady states. Under the premise that the local predictor possesses a steady state, we present a sufficient condition for the convergence of the proposed DSF filter.

Assumption 4. $\rho(\hat{\Phi} \otimes \hat{\Phi} + \sum_{i=1}^l \alpha_i (1 - \alpha_i) I_i \hat{\Phi}^{(1)} \otimes \hat{\Phi}^{(1)} I_i) < 1$, where $\otimes$ is the Kronecker product, $I_i = \text{diag}(\mathbf{0}, \mathbf{I}, \mathbf{0})$ is a diagonal block matrix, the $(l + 1 + i)$ -th diagonal block is an identity matrix of appropriate dimension, and $\rho(\bullet)$ is the spectrum radius of the matrix $\bullet$.

Theorem 3. For systems (14)–(16) under Assumptions 3 and 4, the solution $q(t)$ to (20) with any initial value $q(0) \geq 0$ will converge to the unique positive semi-definite solution q to the following algebraic Lyapunov equation:

$$\begin{aligned} q = & \hat{\Phi} q \hat{\Phi}^T + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{ \hat{\Phi}^{(1)} q [\hat{\Phi}^{(1)}]^T \} + \hat{\Phi} Q^{Xw} + Q^{wX} \hat{\Phi}^T + Q^{w\bar{w}} \\ & + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{ \hat{\Phi}^{(1)} Q^{Xw} [\hat{\Gamma}_0^{(1)}]^T + \hat{\Gamma}_0^{(1)} Q^{wX} [\hat{\Phi}^{(1)}]^T \}. \end{aligned} \quad (29)$$

Proof. See Appendix D. $\square$

Theorem 4. For systems (14)–(16) under Theorem 3, if $(\hat{\Phi}, \hat{H})$ is detectable, $(\hat{\Phi} - \check{S} \check{R}^{-1} \hat{H}, \bar{Q})$ is stabilizable, where $\bar{Q} \bar{Q}^T = \check{Q} - \check{S} \check{R}^{-1} \check{S}^T$ and

$$\begin{aligned} \check{Q} = & \hat{\Phi} Q^{Xw} \hat{\Gamma}_0^T + \hat{\Gamma}_0 Q^{wX} \hat{\Phi}^T + \Gamma_1 Q^w \Gamma_1^T + \Gamma_1 S \Gamma_2^T + \Gamma_2 S^T \Gamma_1^T + \Gamma_2 Q^v \Gamma_2^T \\ & + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \left\{ \hat{\Phi}^{(1)} q [\hat{\Phi}^{(1)}]^T + \hat{\Phi}^{(1)} Q^{Xw} [\hat{\Gamma}_0^{(1)}]^T + \hat{\Gamma}_0^{(1)} Q^{wX} [\hat{\Phi}^{(1)}]^T + \hat{\Gamma}_0^{(1)} Q^{wX} [\hat{\Phi}^{(1)}]^T \right\}, \end{aligned}$$

$$\begin{aligned} \check{S} = & \begin{bmatrix} \mathbf{0} \\ A \end{bmatrix} \odot \left\{ \hat{\Phi}^{(1)} q [\hat{H}^{(1)}]^T + \hat{\Phi}^{(1)} Q^{Xw} [\Gamma^{(1)}]^T + \hat{\Gamma}_0^{(1)} Q^{wX} [\hat{H}^{(1)}]^T + \hat{\Gamma}_0^{(1)} Q^{wX} [\Gamma^{(1)}]^T \right\} \\ & + \hat{\Phi} Q^{Xw} \hat{D}^T + \hat{\Gamma}_0 Q^{wX} \hat{H}^T + \hat{\Gamma}_0 Q^{wX} \hat{D}^T, \end{aligned}$$

$$\begin{aligned} \check{R} = & \hat{H} Q^{Xw} \hat{D}^T + \hat{D} Q^{wX} \hat{H}^T + \hat{D} Q^{wX} \hat{D}^T \\ & + A \odot \left\{ \hat{H}^{(1)} q [\hat{H}^{(1)}]^T + \hat{H}^{(1)} Q^{Xw} [\Gamma^{(1)}]^T + \Gamma^{(1)} Q^{wX} [\hat{H}^{(1)}]^T + \Gamma^{(1)} Q^{wX} [\Gamma^{(1)}]^T \right\}, \end{aligned}$$

the solution $P(t+1|t)$ to (28) with initial $P(0|-1) \geq 0$ will converge to the unique positive semi-definite solution Σ to the following algebraic Riccati equation:

$$\Sigma = (\hat{\Phi} - L \hat{H}) \Sigma (\hat{\Phi} - L \hat{H})^T + \check{Q} - L \check{S}^T - \check{S} L^T + L \check{R} L^T, \quad (30)$$

Further, we have $L = \lim_{t \rightarrow \infty} L(t)$, $K = \lim_{t \rightarrow \infty} K(t)$ and $\Sigma = \lim_{t \rightarrow \infty} P(t|t-1)$.
Furthermore, the steady-state fusion predictor

$$\hat{X}^s(t+1|t) = (\hat{\Phi} - L\hat{H})\hat{X}^s(t|t-1) + Lz(t) \quad (31)$$

is asymptotically stable.

Proof. See Appendix E. $\square$

Remark 9. Compared with the real-time DSF filter computed by Steps 1–8 in the previous section, the steady-state DSF filter only requires Steps 1–5, 7 with the following replacements, where Steps 3, 4 and 5 remain unchanged.

- (1) In Step 1, only $\hat{X}^s(0|-1)$ needs to be initialized and its value can be chosen arbitrarily, since the steady-state fusion predictor (31) is asymptotically stable.
- (2) In Step 2, the time-varying gain $L_i(t)$ is replaced by its constant steady-state value L_i ; i.e., only the predictor $\hat{x}_i(t+1|t)$ needs to be transmitted to the fusion center.
- (3) Step 6 is omitted, because $q(t)$, $Q^\varepsilon(t)$, $K(t)$ and $L(t)$ converge to the constant matrices q , Q^ε , K and L , which can be pre-computed offline according to Theorems 3 and 4.
- (4) In Step 7, the fusion filter and predictor are updated by the steady-state filter $\hat{X}^s(t|t) = \hat{X}^s(t|t-1) + K\varepsilon(t)$ and the steady-state predictor (31), where the innovation $\varepsilon(t)$ is computed by (23) with $\hat{X}(t|t-1)$ replaced by $\hat{X}^s(t|t-1)$.

6. Simulation Examples

6.1. Example 1

In this example, two cases are considered to compare the estimation accuracy: Case 1 with correlated noise and Case 2 with independent noise.

Consider a tracking system with three sensors, which is described as follows:

$$x(t+1) = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} T^2/2 \\ T \end{bmatrix} w(t), \quad (32)$$

$$y_i(t) = H_i x(t) + v_i(t), \quad i = 1, 2, 3. \quad (33)$$

The sampling interval is $T = 0.25$. The zero-mean white noise $w(t)$ with variance $Q^w = 1$ correlates with the measurement noises $v_i(t)$ via $v_i(t) = c_i w(t) + \eta_i(t)$, where c_i represents the noise correlation coefficients. $\eta_i(t)$, $i = 1, 2, 3$ are uncorrelated white noises with zero mean with the covariance matrices $Q^{\eta_1} = 1$, $Q^{\eta_2} = 2$, $Q^{\eta_3} = 3$. The measurement matrices are $H_1 = \begin{bmatrix} 0.5 & 1 \end{bmatrix}$, $H_2 = \begin{bmatrix} 1 & 1 \end{bmatrix}$, $H_3 = \begin{bmatrix} 1 & 0.5 \end{bmatrix}$. The packet reception probabilities of three sensors are set to $\alpha_1 = 0.9$, $\alpha_2 = 0.85$ and $\alpha_3 = 0.8$, respectively. The initial values are $x_0 = [0 \ 0]^T$ and $P_0 = 0.1I_2$, where I_2 denotes the 2×2 identity matrix.

Case 1 (correlated noise case): The noise correlation coefficients are set to $c_1 = 0.2$, $c_2 = 0.4$, and $c_3 = 0.6$.

Figure 3 shows the tracking performance of the proposed fusion filter: (a) is the first state component and (b) is the second state component. It can be seen that the proposed DSF filter has the effective performance.

Figure 4 shows the Root Mean Square Error based on 2000 Monte Carlo trials (RMSE). The RMSE for the i -th state component is defined as $\text{RMSE}_i(t) = \sqrt{\frac{1}{M} \sum_{k=1}^M (x_i^k(t) - \hat{x}_i^k(t|t))^2}$, where M is the number of Monte Carlo runs, $x_i^k(t)$ is the true value of the i -th state component in the k -th run, and $\hat{x}_i^k(t|t)$ is the corresponding filtered estimate. It can be observed that the RMSE values of DSF filter are lower than those of CF filter, which verifies the superior estimation accuracy of the proposed DSF filter.

The averaged steady-state RMSE ($ARMSE = \frac{1}{N-t_0+1} \sum_{t=t_0}^N RMSE(t)$) computes from the 20th step to the 100th step, i.e., $N = 100$, $t_0 = 20$. The ARMSE in all simulations is calculated using the definition given above. And the improvement is calculated as $Improvement = \frac{ARMSE_{ref} - ARMSE_{DSF}}{ARMSE_{ref}} \times 100\%$, $ARMSE_{DSF}$ denotes the ARMSE of the DSF filter, and $ARMSE_{ref}$ denotes the ARMSE of the compared filter. As shown in Table 3, the ARMSEs of the proposed DSF filter are lower than those of the CF filter in [21] for both state components, with improvements of 7.02% and 2.81%, respectively, which further shows that the proposed filter achieves superior estimation accuracy.

We evaluate the sensitivity of the proposed filter to reception probabilities by setting the true probabilities to $\alpha_1 = 0.5$, $\alpha_2 = 0.6$, $\alpha_3 = 0.7$, and considering two mismatched cases with $(\alpha'_i = \alpha_i \pm 0.15)$. The variation rates of the reception probability and the ARMSE are computed as $Variation_{\alpha} = \frac{|\alpha_i - \alpha'_i|}{\alpha_i} \times 100\%$ and $Variation_{ARMSE} = \frac{|ARMSE - ARMSE'|}{ARMSE} \times 100\%$, where ARMSE and ARMSE' correspond to the correct and mismatched cases, respectively. As shown in Table 4, when the assumed reception probabilities deviate from the true values by 0.15 (a relative change of 21.43% to 30%), the ARMSEs of the two state components change by no more than 1.05%. Figure 5 shows that the RMSE curves under these mismatches remain close to the correct-probability case. Figure 5 and Table 4 show that the DSF filter performance degrades only slightly under the considered mismatched cases.

Table 3. ARMSEs and percentage improvements of the proposed DSF filter versus CF [21] in the correlated noise case.

Fusion Filter	DSF Filter	CF Filter in [21]	Improvement
First component	0.2980	0.3205	7.02%
Second component	0.3362	0.3460	2.81%

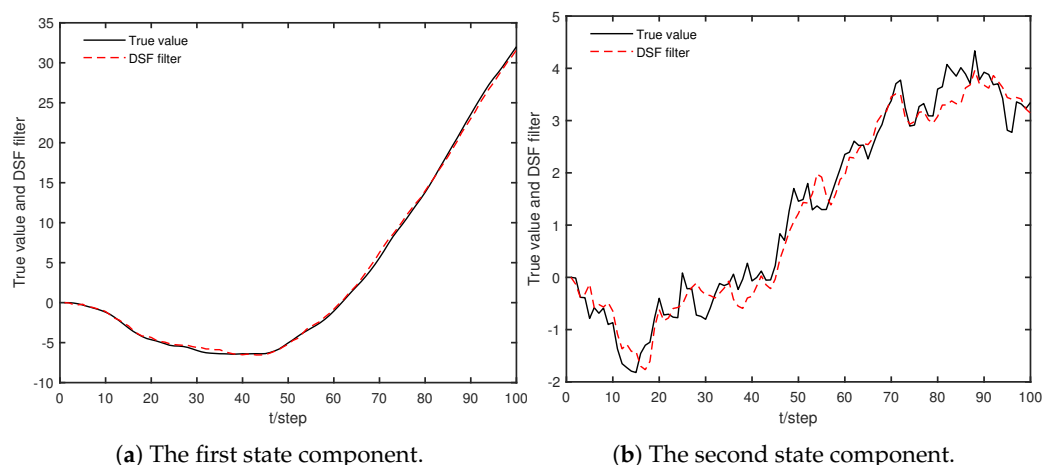


Figure 3. The tracking performance of DSF filter.

Case 2 (uncorrelated noise case): In this case, the process noise $w(t)$ is uncorrelated with the measurement noises $v_i(t)$, $i = 1, 2, 3$, i.e., $c_1 = c_2 = c_3 = 0$. This setting is used to ensure a fair comparison with the DWF filter in [28], since the DWF filter is derived under the assumption of independent noises.

Under reception probabilities $\alpha_1 = 0.9$, $\alpha_2 = 0.85$ and $\alpha_3 = 0.8$, Figure 6 shows the RMSE curves of the three fusion filters (DSF filter, CF filter, DWF filter). The RMSE of the proposed DSF filter is lower than that of the other two compared filters. Figure 7 shows the estimation variance for the DWF filter in [28]. As illustrated in Figure 7, the DWF filter cannot achieve steady-state behavior, as it depends on the random variables $\gamma_i(t)$, $i = 1, 2, 3$.

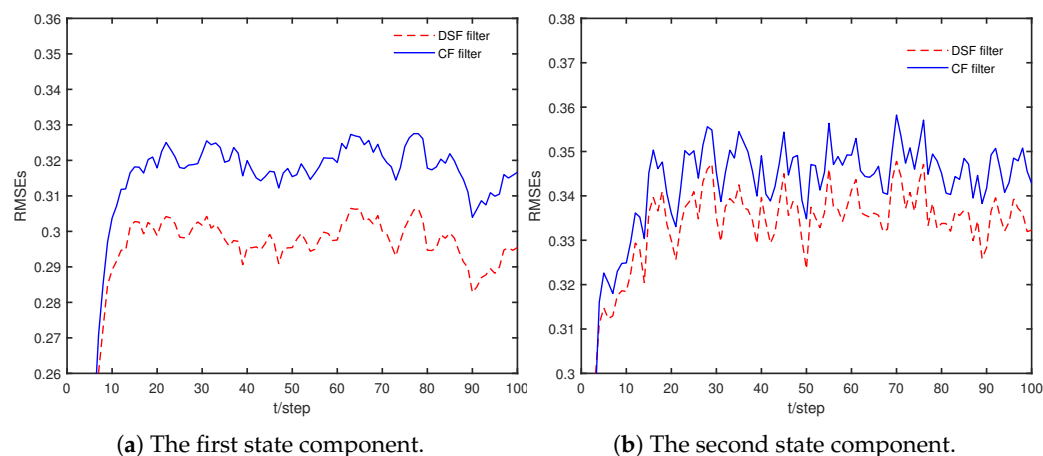


Figure 4. RMSEs of DSF filter and CF filter in [21].

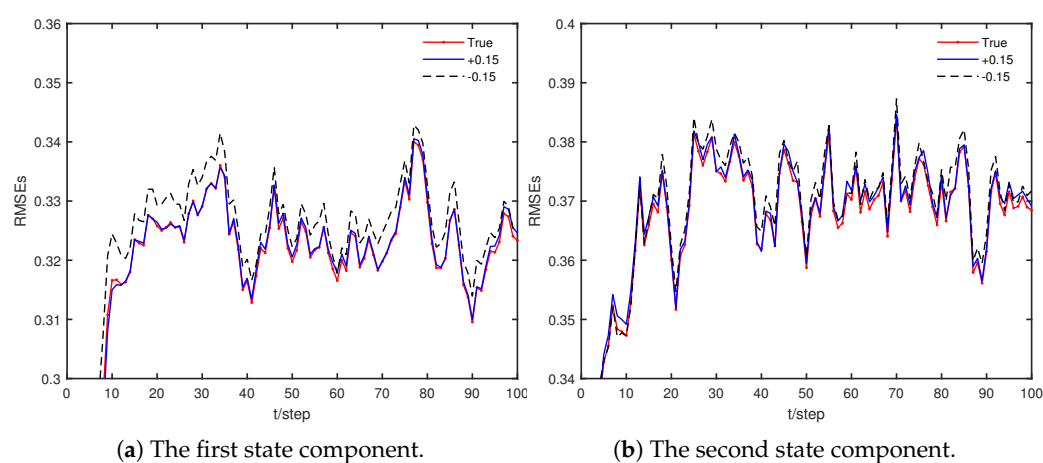


Figure 5. Sensitivity of the DSF filter to the reception probability mismatch.

Table 4. Percentage ARMSE increase of the DSF filter under reception probability mismatch.

Reception Probability Mismatch	First Component	Second Component
+30.00% ($\alpha_1 + 0.15$)	0.18%	0.29%
+25.00% ($\alpha_2 + 0.15$)		
+21.42% ($\alpha_3 + 0.15$)		
−30.00% ($\alpha_1 - 0.15$)	1.05%	0.55%
−25.00% ($\alpha_2 - 0.15$)		
−21.42% ($\alpha_3 - 0.15$)		

For further comparison, we set $\alpha_1 = 0.3$, $\alpha_2 = 0.5$ and $\alpha_3 = 0.4$, i.e., under low reception probabilities. Figure 8 shows the RMSE for the DSF filter, the CF filter in [21] and the DWF filter in [28]. Specifically, the CF filter exhibits the highest RMSE; the DSF filter achieves a medium level with improved accuracy over the CF filter; and the DWF filter maintains the lowest RMSE throughout, since the DWF filter depends on random variables. However, the DWF filter is only applicable to the case of independent noises, and it does not have steady-state behavior.

Table 5 shows the ARMSE values and percentage improvements of the proposed DSF filter over the CF [21] and DWF filter [28] under both high and low reception probabilities with uncorrelated noise. As shown in Table 5, under high reception probabilities, the DSF filter outperforms both the CF filter and the DWF filter, whereas under low reception

probabilities, the DWF filter achieves lower ARMSEs than the DSF filter (indicated by the negative improvements), which is consistent with Figure 8.

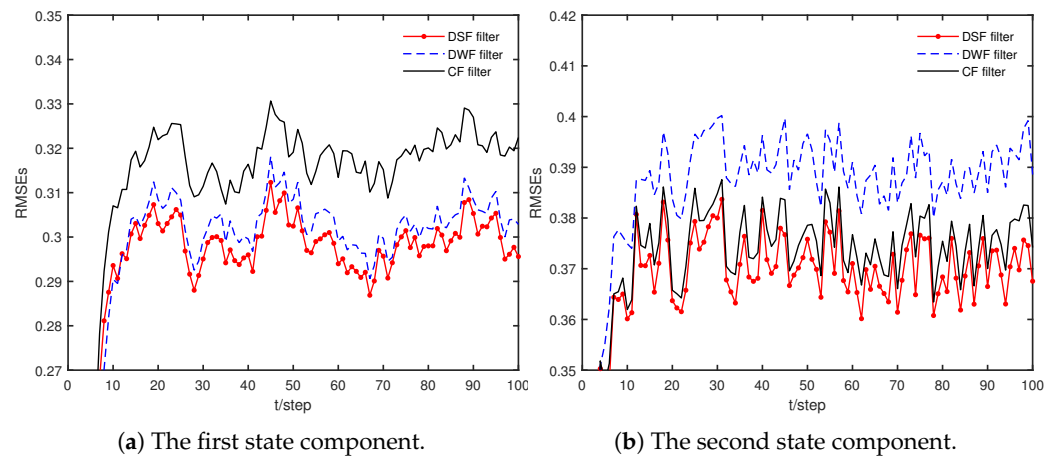


Figure 6. RMSEs of DSF filter, DWF filter in [28], and CF filter in [21] under reception probabilities 0.9, 0.85 and 0.8, respectively.

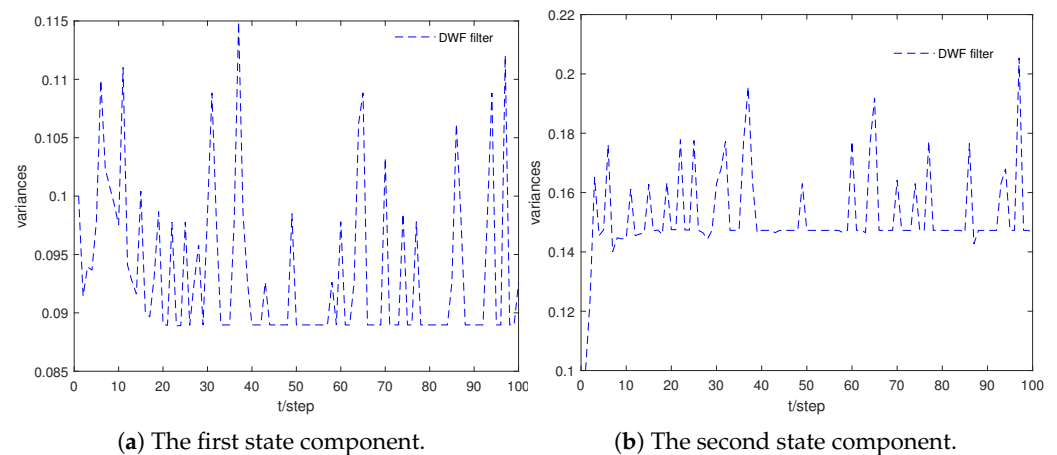


Figure 7. Estimation error variances of DWF filter in [28].

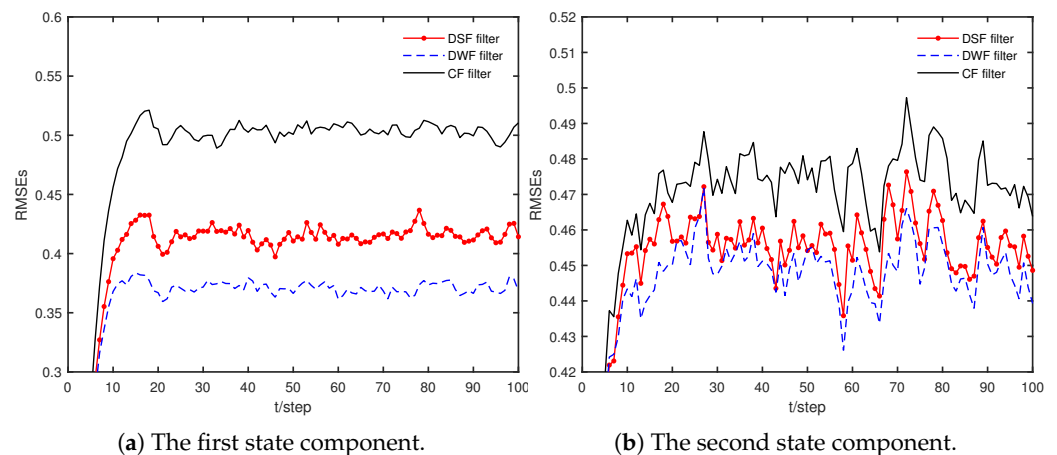


Figure 8. RMSEs of DSF filter, DWF filter in [28], and CF filter in [21] under reception probabilities 0.3, 0.5 and 0.4.

Table 5. ARMSE comparison and percentage improvement of the proposed DSF filter over CF [21] and DWF [28] under high and low reception probabilities.

Reception Probability	State Component	DSF (Proposed)	CF [21]	Improvement
High	First	0.2991	0.3196	+6.41%
High	Second	0.3718	0.3764	+1.21%
Low	First	0.4197	0.5063	+17.11%
Low	Second	0.4597	0.4790	+4.02%
Reception Probability	State Component	DSF (Proposed)	DWF [28]	Improvement
High	First	0.2991	0.3042	+1.68%
High	Second	0.3718	0.3905	+4.76%
Low	First	0.4197	0.3742	−12.13%
Low	Second	0.4597	0.4518	−1.74%

6.2. Example 2

In this example, the steady-state case is considered. It shows the existence of the steady-state filter when Assumptions 3 and 4 are satisfied. Unlike Cases 1 and 2, where the system parameters do not satisfy these assumptions, the system matrices in this example are modified as follows. All other variables remain the same as in Case 1.

$$x(t+1) = \begin{bmatrix} 0.9 & T \\ 0 & 0.9 \end{bmatrix} x(t) + \begin{bmatrix} T^2/2 \\ T \end{bmatrix} w(t).$$

In this case, $\text{rank} \begin{bmatrix} H_1 \\ H_1\Phi \end{bmatrix} = 2$, $\text{rank} \begin{bmatrix} H_2 \\ H_2\Phi \end{bmatrix} = 2$, $\text{rank} \begin{bmatrix} H_3 \\ H_3\Phi \end{bmatrix} = 2$ and $\text{rank} [\Gamma, \Phi\Gamma] = 2$; i.e., the system is completely controllable and completely observable. It means that (Φ, H_i) is detectable, and $(\Phi - \Gamma S_i(Q_i^v)^{-1} H_i, \bar{R}_i)$ is stabilizable; according to Assumption 3, the local estimators have steady-state properties. In this example, $\rho(\hat{\Phi} \otimes \hat{\Phi} + \sum_{i=1}^l \alpha_i (1 - \alpha_i) I_i \hat{\Phi}^{(1)} \otimes \hat{\Phi}^{(1)} I_i) = 0.81 < 1$; i.e., Assumption 4 is also satisfied. From Theorem 3, the steady-state filter exists.

To illustrate the steady-state behavior of the proposed DSF filter, Figure 9 compares its tracking performance under three different initial values: initial value 1: $\hat{X}^s(0|-1) = [14 \ 13 \ 12 \ 11 \ 10 \ 9 \ 8 \ 7 \ 6 \ 5 \ 4 \ 3 \ 2 \ 1]^T$ and $P(0|-1) = 0.1\mathbf{I}$, initial value 2: $\hat{X}^s(0|-1) = [1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10 \ 11 \ 12 \ 13 \ 14]^T$ and $P(0|-1) = \mathbf{I}$, initial value 3: $\hat{X}^s(0|-1) = [1 \ 5 \ 0.5 \ 3 \ 10 \ 0 \ 7 \ -1 \ -2 \ -3 \ -4 \ -5 \ -6 \ -7]^T$ and $P(0|-1) = 5\mathbf{I}$. The results show that the steady-state DSF filter is asymptotically stable and that the initial values can be chosen arbitrarily.

As shown in Table 6, the estimation errors remain unchanged from the 80th step to the 100th step, which indicates that the DSF filter has reached its steady state. This is consistent with the theoretical result in Theorem 4 and further verifies the steady-state property of the proposed filter.

All simulations were carried out on a Xiaomi laptop equipped with an Intel Core 5 210H CPU (2.20 GHz). The total computational time of the real-time fusion algorithm over the 100-step simulation is 0.031207500 s, whereas that of the steady-state fusion algorithm is only 0.002913800 s. This demonstrates that the steady-state gains pre-computed offline reduce the total computational time, which effectively alleviate the online computational burden of the proposed DSF filter.

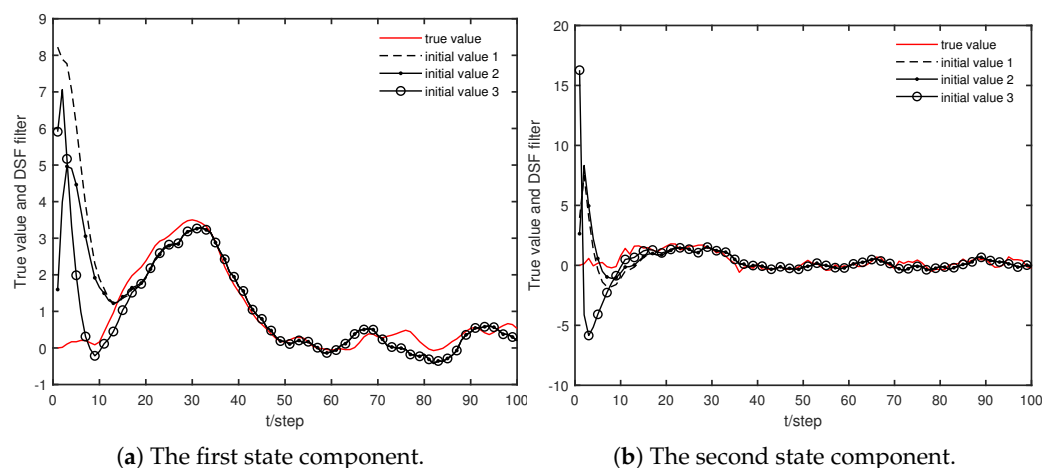


Figure 9. Tracking performance of DSF filter under different initial values.

Table 6. Steady-state filter error variance of the DSF filter at different time steps.

Step	80	90	100
First component	0.04687722996758	0.04687722996758	0.04687722996758
Second component	0.08508771245807	0.08508771245807	0.08508771245807

7. Conclusions

This paper focuses on a distributed fusion filtering problem with prediction-based compensation for DoS-induced packet dropouts modeled as Bernoulli random variables. A distributed fusion filtering algorithm has been developed for multi-sensor systems whose transmission channels from local estimators to the fusion center are subject to random packet dropouts. Such packet dropouts are caused by various random events, in particular, DoS attacks. When a local estimate is lost, a prediction compensation strategy is employed at the fusion center, where the missing data are replaced by their one-step predictors. The DSF filter is derived via the innovation analysis method based on an augmented new system. The proposed framework accommodates correlated process and measurement noises as well as sensors with heterogeneous dropout rates. The steady-state behavior of the proposed DSF filter is rigorously analyzed, and a sufficient condition for the existence of the steady-state filter is established. Under the steady-state condition, the filter gains and error covariances converge to constant matrices. This will reduce the computational and communication burden. And the initial values for the steady-state filter can be chosen arbitrarily. Compared with the existing centralized fusion filter with packet dropouts in [21], the proposed DSF filter achieves superior estimation accuracy under the same reception probabilities. The proposed DSF filter is intermediate in RMSE compared with DWF in [28] under low reception probabilities and independent noise. Under correlated noise conditions, however, DSF is applicable whereas DWF is not.

The current model cannot distinguish DoS attacks from ordinary channel losses. Future work will consider intelligent and time-correlated attack scenarios, design corresponding filter algorithms for time-varying attack probabilities, and explore corresponding attack-detection mechanisms.

Author Contributions: Conceptualization, F.H.; Methodology, F.H.; Software, F.H.; Writing—original draft, F.H.; Supervision, J.M.; Validation, J.M.; Writing—review and editing, J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Heilongjiang Provincial Natural Science Foundation of China (PL2024F032) and the Program for Young Talents of Basic Research in Universities of Heilongjiang Province (YQJH2024184).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

To facilitate understanding of the subsequent proofs, cross-references are provided for the main symbols and notations that appear repeatedly in the appendices.

Table A1. Summary of the key symbols and notations used in the appendices.

Symbol and Notations	Meaning	Index
$z_i(t)$	Virtual measurement	(6)
$X(t)$	Augmented state composed of $x(t)$, $\tilde{x}_i(t+1 t)$ and $z_i(t-1)$	(14)
$z(t)$	Augmented measurement composed of $z_i(t)$	(15)
$\bar{w}(t)$	Augmented noise of system (14), one-step autocorrelated	(16)
$\Phi(t)$	Transition matrix of the augmented system	(14)
$H(t)$	Measurement matrix	(15)
$D(t)$	Matrix multiplying $w(t)$ in (15)	(15)
$\Gamma_0(t), \Gamma_1, \Gamma_2(t)$	Coefficient matrices of $w(t)$, $w(t+1)$ and $v(t+1)$ in $\bar{w}(t)$	(16)
$\Lambda(t)$	Matrix such that $\tilde{\Phi}(t) = \Lambda(t)\hat{\Phi}^{(1)}$	Remark 5
α	$\text{diag}(\alpha_1 \mathbf{I}, \dots, \alpha_l \mathbf{I})$	Remark 5
$\hat{\Phi}^{(1)}, \hat{H}^{(1)}, \hat{\Gamma}_0^{(1)}$	Deterministic matrices in the fluctuation decompositions	Remark 5
A	$\text{diag}((\alpha_1 - \alpha_1^2) \mathbf{1}_{nn}, \dots, (\alpha_l - \alpha_l^2) \mathbf{1}_{nn})$	Lemma 2

The explicit entries of the matrices are provided in the main text, which can be directly retrieved through the index.

Proof of Lemma 2. Under Assumption 1, substituting (16) into $Q^{\bar{w}}(t) = E[\bar{w}(t)\bar{w}^T(t)]$ yields

$$Q^{\bar{w}}(t) = E[\Gamma_0(t)w(t)w^T(t)\Gamma_0^T(t)] + \Gamma_1 Q^w \Gamma_1^T + \Gamma_2(t+1)Q^v \Gamma_2^T(t+1) + \Gamma_1 S \Gamma_2^T(t+1) + \Gamma_2(t+1)S^T \Gamma_1^T, \quad (\text{A1})$$

where the property that $\gamma_i(t), i = 1, 2, \dots, l$ are uncorrelated with other variables is utilized. Substituting $\tilde{\Gamma}_0(t) = \Gamma_0(t) - \hat{\Gamma}_0 = \Lambda(t)\hat{\Gamma}_0^{(1)}$ into $E[\Gamma_0(t)w(t)w^T(t)\Gamma_0^T(t)]$ and using $E[\tilde{\Gamma}_0(t)] = \mathbf{0}$, we obtain

$$E[\Gamma_0(t)w(t)w^T(t)\Gamma_0^T(t)] = \hat{\Gamma}_0 Q^w \hat{\Gamma}_0^T + E\{\Lambda(t)\hat{\Gamma}_0^{(1)}w(t)w^T(t)[\hat{\Gamma}_0^{(1)}]^T \Lambda(t)\}. \quad (\text{A2})$$

Using (17), and then from (A2), we obtain

$$E[\Gamma_0(t)w(t)w^T(t)\Gamma_0^T(t)] = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{\hat{\Gamma}_0^{(1)} Q^w [\hat{\Gamma}_0^{(1)}]^T\} + \hat{\Gamma}_0 Q^w \hat{\Gamma}_0^T. \quad (\text{A3})$$

Substituting (A3) into (A1) yields (18).

Under Assumption 1, substituting (16) into $Q^{\bar{w}w}(t) = E[\bar{w}(t)w^T(t)]$ yields (19). The proof is completed. $\square$

Appendix B

Proof of Theorem 1. Using $\bar{\Phi}(t) = \hat{\Phi}(t) + \tilde{\Phi}(t)$, system (14) can be rewritten as

$$X(t+1) = \hat{\Phi}(t)X(t) + \tilde{\Phi}(t)X(t) + \bar{w}(t), \quad (\text{A4})$$

substituting (A4) into $q(t+1) = E[X(t+1)X^T(t+1)]$, using $E[\tilde{\Phi}(t)] = \mathbf{0}$ and $\gamma(t)$ being uncorrelated with all other variables, we have

$$\begin{aligned} q(t+1) &= \hat{\Phi}(t)q(t)\hat{\Phi}^T(t) + \hat{\Phi}(t)Q^{X\bar{w}}(t) + E[\tilde{\Phi}(t)X(t)X^T(t)\tilde{\Phi}^T(t)] \\ &\quad + E[\tilde{\Phi}(t)X(t)\bar{w}^T(t)] + Q^{\bar{w}X}(t)\hat{\Phi}^T(t) + E[\bar{w}(t)X^T(t)\tilde{\Phi}^T(t)] + Q^{\bar{w}}(t). \end{aligned} \quad (\text{A5})$$

In the following, the derivation of $E[\tilde{\Phi}(t)X(t)X^T(t)\tilde{\Phi}^T(t)]$ and $E[\tilde{\Phi}(t)X(t)\bar{w}^T(t)]$ are presented. Substituting $\tilde{\Phi}(t) = \Lambda(t)\hat{\Phi}^{(1)}$ into $E[\tilde{\Phi}(t)X(t)X^T(t)\tilde{\Phi}^T(t)]$, and using (17), we have

$$E[\tilde{\Phi}(t)X(t)X^T(t)\tilde{\Phi}^T(t)] = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{\hat{\Phi}^{(1)} q(t) [\hat{\Phi}^{(1)}]^T\}. \quad (\text{A6})$$

Substituting $\tilde{\Phi}(t) = \Lambda(t)\hat{\Phi}^{(1)}$, $\tilde{\Gamma}_0(t) = \Gamma_0(t) - \hat{\Gamma}_0 = \Lambda(t)\hat{\Gamma}_0^{(1)}$ into $E[\tilde{\Phi}(t)X(t)\bar{w}^T(t)]$ and using $E[\tilde{\Phi}(t)] = \mathbf{0}$, $E[\tilde{\Gamma}_0(t)] = \mathbf{0}$, the property that variable $\gamma(t)$ is uncorrelated with all other variables, it can be obtained from (17)

$$E[\tilde{\Phi}(t)X(t)\bar{w}^T(t)] = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{\hat{\Phi}^{(1)} Q^{X\bar{w}}(t) [\hat{\Gamma}_0^{(1)}]^T\}. \quad (\text{A7})$$

Substituting (A6) and (A7) into (A5), after rearrangement, we can obtain (20).

Finally, we present the proof of the initial values $q(0)$. Substituting $z_i(-1) = \hat{x}_i(0|-1)$, $i = 1, 2, \dots, l$ into $X(0)$ yields $X(0) = [x^T(0) \quad \hat{x}_1^T(1|0) \quad \dots \quad \hat{x}_l^T(1|0) \quad \hat{x}_1^T(0|-1) \quad \dots \quad \hat{x}_l^T(0|-1)]^T$, substituting it into $q(0) = E[X(0)X^T(0)]$ yields

$$q(0) = \begin{bmatrix} E\{x(0)x^T(0)\} & q^{x\tilde{x}}(0) & q^{x\hat{x}}(0) \\ q^{\tilde{x}x}(0) & q^{\tilde{x}\tilde{x}}(0) & q^{\tilde{x}\hat{x}}(0) \\ q^{\hat{x}x}(0) & q^{\hat{x}\tilde{x}}(0) & q^{\hat{x}\hat{x}}(0) \end{bmatrix}, \quad (\text{A8})$$

where $q^{x\tilde{x}}(0) = (E\{x(0)\hat{x}_i^T(1|0)\})$ and $q^{x\hat{x}}(0) = (E\{x(0)\hat{x}_i^T(0|-1)\})$, $i = 1, 2, \dots, l$, are $n \times nl$ matrices, $q^{\tilde{x}\tilde{x}}(0) = (E\{\hat{x}_i(1|0)\hat{x}_j^T(1|0)\})$, $q^{\tilde{x}\hat{x}}(0) = (E\{\hat{x}_i(1|0)\hat{x}_j^T(0|-1)\})$ and $q^{\hat{x}\hat{x}}(0) = E\{\hat{x}_i(0|-1)\hat{x}_j^T(0|-1)\}$, $i, j = 1, 2, \dots, l$ are $nl \times nl$ matrices, and $q^{\tilde{x}x}(0) = [q^{x\tilde{x}}(0)]^T$, $q^{\hat{x}x}(0) = [q^{x\hat{x}}(0)]^T$, $q^{\hat{x}\tilde{x}}(0) = [q^{\tilde{x}\hat{x}}(0)]^T$.

From $\hat{x}_i(0|-1) \perp \tilde{x}_i(0|-1)$, substituting $x(0) = \hat{x}_i(0|-1) + \tilde{x}_i(0|-1)$ into $E\{x(0)x^T(0)\}$, it follows from Lemma 1 that $E\{x(0)x^T(0)\} = x_0x_0^T + P_0$. Similarly, based on $\hat{x}_i(0|-1) \perp \tilde{x}_i(0|-1)$, $L_i(0) = \mathbf{0}$, (13), Assumption 1, Assumption 2 and Lemma 1, we have $E\{x(0)\hat{x}_i^T(1|0)\} = P_0\Phi^T$, $E\{x(0)\hat{x}_i^T(0|-1)\} = x_0x_0^T$, $E\{\hat{x}_i(1|0)\hat{x}_j^T(1|0)\} = \Phi P_0\Phi^T + \Gamma Q^w \Gamma^T$, $E\{\hat{x}_i(1|0)\hat{x}_j^T(0|-1)\} = \mathbf{0}$, $E\{\hat{x}_i(0|-1)\hat{x}_j^T(0|-1)\} = x_0x_0^T$, where the unbiasedness of the projection $E[\tilde{x}_i(1|0)] = \mathbf{0}$ is used. The proof is completed. $\square$

Appendix C

Proof of Theorem 2. From the recursive projection formula [33], we directly obtain (21), where the filter gain matrix $K(t) = E[X(t)\varepsilon^T(t)]\{E[\varepsilon(t)\varepsilon^T(t)]\}^{-1}$. According to Remark 2,

we have $w(t+1) \perp L(z(1), z(2), \dots, z(t))$, $v(t+1) \perp L(z(1), z(2), \dots, z(t))$, $w(t) \perp L(z(1), z(2), \dots, z(t-1))$; thus, based on (14), (22) is obtained similarly with (21), where the predictor gain matrix $L(t) = E\{[\bar{\Phi}(t)X(t) + \Gamma_0(t)w(t)]\varepsilon^T(t)\} \{E[\varepsilon(t)\varepsilon^T(t)]\}^{-1}$.

Substituting (15) into $\varepsilon(t) = z(t) - \text{proj}[z(t) | z(1), z(2), \dots, z(t-1)]$, (23) is obtained, where $w(t) \perp L(z(1), z(2), \dots, z(t-1))$ is used. Substituting $H(t) = \hat{H} + \tilde{H}(t)$, $D(t) = \hat{D} + \tilde{D}(t)$ into (15), (23) can be rewritten as

$$\varepsilon(t) = \hat{H}\tilde{X}(t|t-1) + \tilde{H}(t)X(t) + \hat{D}w(t) + \tilde{D}(t)w(t), \quad (\text{A9})$$

where the one-step prediction error is defined as $\tilde{X}(t|t-1) = X(t) - \hat{X}(t|t-1)$. Based on $E[\tilde{H}(t)] = \mathbf{0}$ and $E[\tilde{D}(t)] = \mathbf{0}$, substituting (A9) into innovation covariance matrix $Q^\varepsilon(t) = E[\varepsilon(t)\varepsilon^T(t)]$, we have

$$\begin{aligned} Q^\varepsilon(t) &= \hat{H}P(t|t-1)\hat{H}^T + \hat{H}E[\tilde{X}(t|t-1)w^T(t)]\hat{D}^T + E[\tilde{H}(t)X(t)X^T(t)\tilde{H}^T(t)] \\ &\quad + E[\tilde{H}(t)X(t)w^T(t)\tilde{D}^T(t)] + \hat{D}E[w(t)\tilde{X}^T(t|t-1)]\hat{H}^T + \hat{D}Q^w\hat{D}^T \\ &\quad + E[\tilde{D}(t)w(t)X^T(t)\tilde{H}^T(t)] + E[\tilde{D}(t)w(t)w^T(t)\tilde{D}^T(t)], \end{aligned} \quad (\text{A10})$$

where the prediction error covariance matrix $P(t|t-1) = E[\tilde{X}(t|t-1)\tilde{X}^T(t|t-1)]$. From the projection theory, we have $\hat{X}(t|t-1) \in L(z(1), z(2), \dots, z(t-1))$, but $w(t) \perp L(z(1), z(2), \dots, z(t-1))$; thus, we have $w(t) \perp \hat{X}(t|t-1)$, and based on $\tilde{X}(t|t-1) = X(t) - \hat{X}(t|t-1)$, we have

$$E[\tilde{X}(t|t-1)w^T(t)] = E[X(t)w^T(t)] = Q^{Xw}(t). \quad (\text{A11})$$

Using $\tilde{H}(t) = [\gamma(t) - \alpha]\hat{H}^{(1)}$ and $\tilde{D}(t) = [\gamma(t) - \alpha]\Gamma^{(1)}$, under Remark 4 and (17), we have, respectively,

$$E[\tilde{H}(t)X(t)X^T(t)\tilde{H}^T(t)] = A \odot \{\hat{H}^{(1)}q(t)[\hat{H}^{(1)}]^T\}, \quad (\text{A12})$$

$$E[\tilde{H}(t)X(t)w^T(t)\tilde{D}^T(t)] = A \odot \{\hat{H}^{(1)}Q^{Xw}(t)[\Gamma^{(1)}]^T\}, \quad (\text{A13})$$

$$E[\tilde{D}(t)w(t)w^T(t)\tilde{D}^T(t)] = A \odot \{\Gamma^{(1)}Q^w[\Gamma^{(1)}]^T\}, \quad (\text{A14})$$

substituting (A11)–(A14) into (A10), after rearrangement, (24) is obtained.

In the following, we derive the filter gain matrix $K(t)$ and the prediction gain matrix $L(t)$. Substituting (A9) into $E[X(t)\varepsilon^T(t)]$, using $E[\tilde{H}(t)] = \mathbf{0}$ and $E[\tilde{D}(t)] = \mathbf{0}$, we have

$$E[X(t)\varepsilon^T(t)] = E[X(t)\tilde{X}^T(t|t-1)]\hat{H}^T + Q^{Xw}(t)\hat{D}^T. \quad (\text{A15})$$

Based on $\tilde{X}(t|t-1) = X(t) - \hat{X}(t|t-1)$ and $\hat{X}(t|t-1) \perp \tilde{X}(t|t-1)$, (A15) can be written as $E[X(t)\varepsilon^T(t)] = P(t|t-1)\hat{H}^T + Q^{Xw}(t)\hat{D}^T$; thus, (25) can be obtained. Substituting (A9), $\bar{\Phi}(t) = \tilde{\Phi}(t) + \hat{\Phi}(t)$ and $\Gamma_0(t) = \tilde{\Gamma}_0(t) + \hat{\Gamma}_0$ into $E\{[\bar{\Phi}(t)X(t) + \Gamma_0(t)w(t)]\varepsilon^T(t)\}$, based on $E[\tilde{\Phi}(t)] = \mathbf{0}$ and $E[\tilde{\Gamma}_0(t)] = \mathbf{0}$, we have

$$\begin{aligned} E\{[\bar{\Phi}(t)X(t) + \Gamma_0(t)w(t)]\varepsilon^T(t)\} &= E[\tilde{\Phi}(t)X(t)X^T(t)\tilde{H}^T(t)] + E[\tilde{\Phi}(t)X(t)w^T(t)\tilde{D}^T(t)] \\ &\quad + \hat{\Phi}(t)E[X(t)\tilde{X}^T(t|t-1)]\hat{H}^T + \hat{\Phi}(t)Q^{Xw}(t)\hat{D}^T + E[\tilde{\Gamma}_0(t)w(t)X^T(t)\tilde{H}^T(t)] \\ &\quad + E[\tilde{\Gamma}_0(t)w(t)w^T(t)\tilde{D}^T(t)] + \hat{\Gamma}_0E[w(t)\tilde{X}^T(t|t-1)]\hat{H}^T + \hat{\Gamma}_0Q^w\hat{D}^T. \end{aligned} \quad (\text{A16})$$

Based on $\tilde{\Phi}(t) = \Lambda(t)\hat{\Phi}^{(1)}$, $\tilde{\Gamma}_0(t) = \Lambda(t)\hat{\Gamma}_0^{(1)}$, $\tilde{D}(t) = [\gamma(t) - \alpha]\Gamma^{(1)}$, and $\tilde{H}(t) = [\gamma(t) - \alpha]\hat{H}^{(1)}$, from (17), we have, respectively,

$$E[\tilde{\Phi}(t)X(t)X^T(t)\hat{H}^T(t)] = \begin{bmatrix} \mathbf{0} \\ A \end{bmatrix} \odot \{\hat{\Phi}^{(1)}q(t)[\hat{H}^{(1)}]^T\}, \quad (\text{A17})$$

$$E[\tilde{\Phi}(t)X(t)w^T(t)\hat{D}^T(t)] = \begin{bmatrix} \mathbf{0} \\ A \end{bmatrix} \odot \{\hat{\Phi}^{(1)}Q^{Xw}(t)[\Gamma^{(1)}]^T\}, \quad (\text{A18})$$

$$E[\tilde{\Gamma}_0(t)w(t)X^T(t)\hat{H}^T(t)] = \begin{bmatrix} \mathbf{0} \\ A \end{bmatrix} \odot \{\hat{\Gamma}_0^{(1)}Q^{wX}(t)[\hat{H}^{(1)}]^T\}, \quad (\text{A19})$$

$$E[\tilde{\Gamma}_0(t)w(t)w^T(t)\hat{D}^T(t)] = \begin{bmatrix} \mathbf{0} \\ A \end{bmatrix} \odot \{\hat{\Gamma}_0^{(1)}Q^w[\Gamma^{(1)}]^T\}. \quad (\text{A20})$$

In the derivation of (A11), $w(t) \perp \hat{X}(t|t-1)$ is used, and based on $\tilde{X}(t|t-1) = X(t) - \hat{X}(t|t-1)$, we have, respectively,

$$\hat{\Phi}(t)E[X(t)\tilde{X}^T(t|t-1)]\hat{H}^T = \hat{\Phi}(t)P(t|t-1)\hat{H}^T, \quad (\text{A21})$$

$$\hat{\Gamma}_0E[w(t)\tilde{X}^T(t|t-1)]\hat{H}^T = \hat{\Gamma}_0Q^{wX}(t)\hat{H}^T. \quad (\text{A22})$$

Substituting (A16)–(A22) into $L(t) = E\{[\tilde{\Phi}(t)X(t) + \Gamma_0(t)w(t)]\varepsilon^T(t)\}\{E[\varepsilon(t)\varepsilon^T(t)]\}^{-1}$, (26) can be obtained.

Subtracting (21) from state $X(t)$ yields the filtering error $\tilde{X}(t|t) = \tilde{X}(t|t-1) - K(t)\varepsilon(t)$; by rearranging the terms, we have

$$\tilde{X}(t|t) + K(t)\varepsilon(t) = \tilde{X}(t|t-1). \quad (\text{A23})$$

It follows from $\tilde{X}(t|t) \perp \varepsilon(t)$, taking the variance on both sides of (A23), we have $P(t|t) + K(t)Q^\varepsilon(t)K^T(t) = P(t|t-1)$ where $P(t|t) = E[\tilde{X}(t|t)\tilde{X}^T(t|t)]$, $P(t|t-1) = E[\tilde{X}(t|t-1)\tilde{X}^T(t|t-1)]$; thus, (27) is obtained. Similarly, subtracting (22) from $X(t+1)$, based on (14), we have $\tilde{X}(t+1|t) + L(t)\varepsilon(t) = \hat{\Phi}(t)\tilde{X}(t|t-1) + \tilde{\Phi}(t)X(t) + \bar{w}(t)$, and taking the variance on both sides, it can be obtained that

$$P(t+1|t) + L(t)Q^\varepsilon(t)L^T(t) = E\{[\hat{\Phi}(t)\tilde{X}(t|t-1) + \tilde{\Phi}(t)X(t) + \bar{w}(t)] \times [\hat{\Phi}(t)\tilde{X}(t|t-1) + \tilde{\Phi}(t)X(t) + \bar{w}(t)]^T\}, \quad (\text{A24})$$

using $E[\tilde{\Phi}(t)] = \mathbf{0}$, the right-hand side of (A24) can be expanded as

$$\begin{aligned} & \hat{\Phi}(t)P(t|t-1)\hat{\Phi}^T(t) + \hat{\Phi}(t)E[\tilde{X}(t|t-1)\bar{w}^T(t)] + E[\tilde{\Phi}(t)X(t)X^T(t)]\tilde{\Phi}^T(t) \\ & + E[\tilde{\Phi}(t)X(t)\bar{w}^T(t)] + E[\bar{w}(t)\tilde{X}(t|t-1)]\hat{\Phi}^T(t) + E[\bar{w}(t)X(t)X^T(t)]\tilde{\Phi}^T(t) + Q^{\bar{w}}(t). \end{aligned} \quad (\text{A25})$$

It is known that $\hat{X}(t|t-1) \in L(z(t-1), z(t-2), \dots, z(1))$, $w(t) \perp L(z(t-1), z(t-2), \dots, z(1))$, $w(t+1) \perp L(z(t-1), z(t-2), \dots, z(1))$, $v(t+1) \perp L(z(t-1), z(t-2), \dots, z(1))$, hence, we have $\bar{w}(t) \perp \hat{X}(t|t-1)$. Based on $\tilde{X}(t|t-1) = X(t) - \hat{X}(t|t-1)$ and $\bar{w}(t) \perp \hat{X}(t|t-1)$, we have

$$E[\tilde{X}(t|t-1)\bar{w}^T(t)] = E[X(t)\bar{w}^T(t)] = Q^{X\bar{w}}(t). \quad (\text{A26})$$

Based on $E[\tilde{\Phi}(t)] = \mathbf{0}$, $E[\tilde{\Gamma}_0(t)] = \mathbf{0}$, $\tilde{\Phi}(t) = \Lambda(t)\hat{\Phi}^{(1)}$, $\tilde{\Gamma}_0(t) = \Lambda(t)\hat{\Gamma}_0^{(1)}$, from (17), we have

$$E[\tilde{\Phi}(t)X(t)\bar{w}^T(t)] = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{\hat{\Phi}^{(1)}Q^{X\bar{w}}(t)[\hat{\Gamma}_0^{(1)}]^T\}. \quad (\text{A27})$$

Based on $E[\tilde{\Phi}(t)] = \mathbf{0}$, $\tilde{\Phi}(t) = \Lambda(t)\hat{\Phi}^{(1)}$, from (17), we have

$$E[\tilde{\Phi}(t)X(t)X^T(t)\tilde{\Phi}^T(t)] = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{\hat{\Phi}^{(1)}q(t)[\hat{\Phi}^{(1)}]^T\}. \quad (\text{A28})$$

After substituting (A26)–(A28) into (A25), rearranging (A24) yields (28).

In the end, we give the proof of the initial values. Substituting $z_i(-1) = \hat{x}_i(0|-1)$, $i = 1, 2, \dots, l$ into $X(0)$, we have

$$X(0) = \begin{bmatrix} x^T(0) & \hat{x}_1^T(1|0) & \cdots & \hat{x}_l^T(1|0) & \hat{x}_1^T(0|-1) & \cdots & \hat{x}_l^T(0|-1) \end{bmatrix}^T.$$

Under Assumption 2, there is $E[x(0)] = x_0$; from the unbiasedness of $\hat{x}_i(1|0)$, we have $E[\hat{x}_i(1|0)] = \mathbf{0}$ and under Lemma 1, there is $\hat{x}_i(0|-1) = x_0$, $i = 1, 2, \dots, l$. From $E[\tilde{X}(0|-1)] = E[X(0) - \hat{X}(0|-1)] = \mathbf{0}$, we have $\hat{X}(0|-1) = E[X(0)] = \begin{bmatrix} x_0^T & \mathbf{0} & \cdots & \mathbf{0} & x_0^T & \cdots & x_0^T \end{bmatrix}^T$. Substituting

$$\tilde{X}(0|-1) = \begin{bmatrix} (x(0) - x_0)^T & \hat{x}_1^T(1|0) & \cdots & \hat{x}_l^T(1|0) & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}^T$$

into $P(0|-1) = E[\tilde{X}(0|-1)\tilde{X}^T(0|-1)]$, we have

$$P(0|-1) = \begin{bmatrix} E\{(x(0) - x_0)(x(0) - x_0)^T\} & P^{\tilde{x}\tilde{x}}(0|-1) & \mathbf{0} \\ P^{\tilde{x}\tilde{x}}(0|-1) & P^{\tilde{x}\tilde{x}}(0|-1) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $P^{\tilde{x}\tilde{x}}(0|-1) = (E\{(x(0) - x_0)\hat{x}_i^T(1|0)\})$ is an $n \times nl$ matrix, $P^{\tilde{x}\tilde{x}}(0|-1) = (E\{\hat{x}_i(1|0)\hat{x}_j^T(1|0)\})$ is a $nl \times nl$ matrix, and $P^{\tilde{x}\tilde{x}}(0|-1) = [P^{\tilde{x}\tilde{x}}(0|-1)]^T$. The derivation of the initial value can be referred to the proof of Theorem 1. The proof is completed. $\square$

Appendix D

Proof of Theorem 3. Under Assumption 3, the local one-step predictor will have the stability [33]. When $t \rightarrow \infty$, we have $L_i(t) = L_i$, $\Psi_i(t) = \Psi_i$; furthermore, there will be $L^a(t) = L^a$, $\Psi(t) = \Psi$, $\Gamma_2(t) = \Gamma_2$, $\hat{\Phi}(t) = \hat{\Phi}$, $Q^{\bar{w}}(t) = Q^{\bar{w}}$, $Q^{X\bar{w}}(t) = Q^{X\bar{w}}$, $Q^{Xw}(t) = Q^{Xw}$.

(20) can be rewritten as

$$q(t+1) = \hat{\Phi}q(t)\hat{\Phi}^T + E\{\Lambda(t)\hat{\Phi}^{(1)}X(t)X^T(t)[\hat{\Phi}^{(1)}]^T\Lambda^T(t)\} + \Sigma^q, \quad (\text{A29})$$

where $\Sigma^q = \hat{\Phi}Q^{X\bar{w}} + Q^{\bar{w}X}\hat{\Phi}^T + Q^{\bar{w}} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \{\hat{\Phi}^{(1)}Q^{Xw}[\hat{\Gamma}_0^{(1)}]^T + \hat{\Gamma}_0^{(1)}Q^{wX}[\hat{\Phi}^{(1)}]^T\}$.

Based on $\Lambda(t) = \sum_{i=1}^l [\gamma_i(t) - \alpha_i]\mathbf{I}_i$ and $E[(\gamma_i(t) - \alpha_i)(\gamma_j(t) - \alpha_j)] = 0$, we can have

$$E\{\Lambda(t)\hat{\Phi}^{(1)}X(t)X^T(t)[\hat{\Phi}^{(1)}]^T\Lambda^T(t)\} = \sum_{i=1}^l \alpha_i(1 - \alpha_i)\mathbf{I}_i\hat{\Phi}^{(1)}q(t)[\hat{\Phi}^{(1)}]^T\mathbf{I}_i, \quad (\text{A30})$$

Substituting (A30) into (A29) yields

$$q(t+1) = \hat{\Phi}q(t)\hat{\Phi}^T + \sum_{i=1}^l \alpha_i(1 - \alpha_i)\mathbf{I}_i\hat{\Phi}^{(1)}q(t)[\hat{\Phi}^{(1)}]^T\mathbf{I}_i + \Sigma^q, \quad (\text{A31})$$

similar to [18], when $\rho(\bar{A}) < 1$, $\bar{A} = \hat{\Phi} \otimes \hat{\Phi} + \sum_{i=1}^l \alpha_i(1 - \alpha_i)\mathbf{I}_i\hat{\Phi}^{(1)} \otimes \hat{\Phi}^{(1)}\mathbf{I}_i$, the solution $q(t)$ to (A31) with any initial value $q(0) \geq 0$ converges to the positive semi-definite solution q to the algebraic Lyapunov (29), i.e., $q = \lim_{t \rightarrow \infty} q(t)$ finite. The proof is completed. $\square$

Appendix E

Proof of Theorem 4. When $t \rightarrow \infty$, substituting (A9) into $\tilde{X}(t+1|t) + L(t)\varepsilon(t) = \hat{\tilde{\Phi}}(t)\tilde{X}(t|t-1) + \tilde{\tilde{\Phi}}(t)X(t) + \tilde{w}(t)$ and from Theorem 3 yields

$$\begin{aligned}\tilde{X}(t+1|t) = & [\hat{\tilde{\Phi}} - L(t)\hat{H}]\tilde{X}(t|t-1) + [\tilde{\tilde{\Phi}}(t) - L(t)\tilde{H}(t)]X(t) + \Gamma_0(t)w(t) \\ & + \Gamma_1w(t+1) + \Gamma_2v(t+1) - L(t)\hat{D}w(t) - L(t)\tilde{D}(t)w(t).\end{aligned}\quad (\text{A32})$$

Substituting (A32) into $P(t+1|t) = E[\tilde{X}(t+1|t)\tilde{X}^T(t+1|t)]$ and using $E[\tilde{\tilde{\Phi}}(t)] = \mathbf{0}$, $E[\tilde{H}(t)] = \mathbf{0}$, $E[\tilde{\Gamma}_0(t)] = \mathbf{0}$ yield

$$\begin{aligned}P(t+1|t) = & E\{\{[\hat{\tilde{\Phi}} - L(t)\hat{H}]\tilde{X}(t|t-1) + [\tilde{\tilde{\Phi}}(t) - L(t)\tilde{H}(t)]X(t) + \Gamma_0(t)w(t) + \Gamma_1w(t+1) \\ & + \Gamma_2v(t+1) - L(t)\hat{D}w(t) - L(t)\tilde{D}(t)w(t)\}\{*\}^T\} \\ = & [\hat{\tilde{\Phi}} - L(t)\hat{H}]P(t|t-1)[\hat{\tilde{\Phi}} - L(t)\hat{H}]^T + E\{[\hat{\tilde{\Phi}} - L(t)\hat{H}]\tilde{X}(t|t-1)w^T(t)\Gamma_0^T(t)\} \\ & + E\{*\}^T - E\{[\hat{\tilde{\Phi}} - L(t)\hat{H}]\tilde{X}(t|t-1)w^T(t)\hat{D}^TL^T(t)\} - E\{*\}^T \\ & + E\{[\tilde{\tilde{\Phi}}(t) - L(t)\tilde{H}(t)]X(t)X^T(t)[\tilde{\tilde{\Phi}}(t) - L(t)\tilde{H}(t)]^T\} \\ & + E\{[\tilde{\tilde{\Phi}}(t) - L(t)\tilde{H}(t)]X(t)w^T(t)\Gamma_0^T(t)\} + E\{*\}^T \\ & - E\{[\tilde{\tilde{\Phi}}(t) - L(t)\tilde{H}(t)]X(t)w^T(t)\hat{D}^TL^T(t)\} - E\{*\}^T \\ & - E\{[\tilde{\tilde{\Phi}}(t) - L(t)\tilde{H}(t)]X(t)w^T(t)\tilde{D}^T(t)L^T(t)\} - E\{*\}^T \\ & + E\{\Gamma_0(t)w(t)w^T(t)\Gamma_0^T(t)\} + E\{\Gamma_1w(t+1)w^T(t+1)\Gamma_1^T\} \\ & + E\{\Gamma_2v(t+1)v^T(t+1)\Gamma_2^T\} + E\{\Gamma_1w(t+1)v^T(t+1)\Gamma_2^T\} + E\{*\}^T \\ & - E\{\Gamma_0(t)w(t)w^T(t)\hat{D}^TL^T(t)\} - E\{*\}^T \\ & - E\{\Gamma_0(t)w(t)w^T(t)\tilde{D}^T(t)L^T(t)\} - E\{*\}^T \\ & + E\{L(t)\hat{D}w(t)w^T(t)\hat{D}^TL^T(t)\} + E\{L(t)\hat{D}w(t)w^T(t)\tilde{D}^T(t)L^T(t)\} + E\{*\}^T \\ & + E\{L(t)\tilde{D}(t)w(t)w^T(t)\tilde{D}^T(t)L^T(t)\} \\ = & [\hat{\tilde{\Phi}} - L(t)\hat{H}]P(t|t-1)[\hat{\tilde{\Phi}} - L(t)\hat{H}]^T + \check{Q} - L(t)\check{S}^T - \check{S}L^T(t) + L(t)\check{R}L^T(t),\end{aligned}\quad (\text{A33})$$

where $\{*\}$ denotes the term inside the preceding curly braces, $w(t)$ is uncorrelated with $w(t+1)$ and $v(t+1)$, $X(t)$ is uncorrelated with $w(t+1)$ and $v(t+1)$, $E[\tilde{X}(t|t-1)w^T(t)] = Q^{Xw}$. $\check{Q}$, $\check{S}$, and $\check{R}$ are, respectively, defined as follows:

$$\begin{aligned}\check{Q} = & E\{\hat{\tilde{\Phi}}\tilde{X}(t|t-1)w^T(t)\Gamma_0^T(t)\} + E\{*\}^T + E\{\tilde{\tilde{\Phi}}(t)X(t)X^T(t)\tilde{\tilde{\Phi}}^T(t)\} + E\{\tilde{\tilde{\Phi}}(t)X(t)w^T(t) \times \\ & \Gamma_0^T(t)\} + E\{*\}^T + E\{\Gamma_0(t)w(t)w^T(t)\Gamma_0^T(t)\} + E\{\Gamma_1w(t+1)w^T(t+1)\Gamma_1^T\} \\ & + E\{\Gamma_2v(t+1)v^T(t+1)\Gamma_2^T\} + E\{\Gamma_1w(t+1)v^T(t+1)\Gamma_2^T\} + E\{*\}^T \\ = & \hat{\tilde{\Phi}}Q^{Xw}\hat{\Gamma}_0^T + \hat{\Gamma}_0Q^{wX}\hat{\tilde{\Phi}}^T + \Gamma_1Q^w\Gamma_1^T + \Gamma_1S\Gamma_2^T + \Gamma_2S^T\Gamma_1^T + \Gamma_2Q^v\Gamma_2^T \\ & + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \odot \left\{ \hat{\tilde{\Phi}}^{(1)}q[\hat{\tilde{\Phi}}^{(1)}]^T + \hat{\tilde{\Phi}}^{(1)}Q^{Xw}[\hat{\Gamma}_0^{(1)}]^T + \hat{\Gamma}_0^{(1)}Q^{wX}[\hat{\tilde{\Phi}}^{(1)}]^T + \hat{\Gamma}_0^{(1)}Q^{wX}[\hat{\tilde{\Phi}}^{(1)}]^T \right\},\end{aligned}$$

$$\begin{aligned}\check{S} = & E\{\Gamma_0(t)w(t)\tilde{X}^T(t|t-1)\hat{H}^T\} + E\{\hat{\tilde{\Phi}}\tilde{X}(t|t-1)w^T(t)\hat{D}^T\} + E\{\tilde{\tilde{\Phi}}(t)X(t)X^T(t)\tilde{H}^T(t)\} \\ & + E\{\Gamma_0(t)w(t)X^T(t)\tilde{H}^T(t)\} + E\{\tilde{\tilde{\Phi}}(t)X(t)w^T(t)\tilde{D}^T(t)\} + E\{\Gamma_0(t)w(t)w^T(t)\hat{D}^T\} \\ & + E\{\Gamma_0(t)w(t)w^T(t)\tilde{D}^T(t)\} \\ = & \begin{bmatrix} \mathbf{0} \\ A \end{bmatrix} \odot \left\{ \hat{\tilde{\Phi}}^{(1)}q[\hat{H}^{(1)}]^T + \hat{\tilde{\Phi}}^{(1)}Q^{Xw}[\Gamma^{(1)}]^T + \hat{\Gamma}_0^{(1)}Q^{wX}[\hat{H}^{(1)}]^T + \hat{\Gamma}_0^{(1)}Q^w[\Gamma^{(1)}]^T \right\} \\ & + \hat{\tilde{\Phi}}Q^{Xw}\hat{D}^T + \hat{\Gamma}_0Q^{wX}\hat{H}^T + \hat{\Gamma}_0Q^w\hat{D}^T,\end{aligned}$$

$$\begin{aligned}
\tilde{R} &= E\{\hat{H}\tilde{X}(t|t-1)w^T(t)\hat{D}^T\} + E\{*\}^T + E\{\tilde{H}(t)X(t)X^T(t)\tilde{H}^T(t)\} + E\{\tilde{H}(t)X(t)w^T(t) \times \\
&\quad \tilde{D}^T(t)\} + E\{*\}^T + E\{\hat{D}w(t)w^T(t)\hat{D}^T\} + E\{\tilde{D}(t)w(t)w^T(t)\tilde{D}^T(t)\} \\
&= \hat{H}Q^{Xw}\hat{D}^T + \hat{D}Q^{wX}\hat{H}^T + \hat{D}Q^w\hat{D}^T \\
&\quad + A \odot \left\{ \hat{H}^{(1)}q[\hat{H}^{(1)}]^T + \hat{H}^{(1)}Q^{Xw}[\Gamma^{(1)}]^T + \Gamma^{(1)}Q^{wX}[\hat{H}^{(1)}]^T + \Gamma^{(1)}Q^w[\Gamma^{(1)}]^T \right\}.
\end{aligned}$$

Similar to [18], the solution $P(t+1|t)$ to (28) with any initial condition $P(0|-1) \geq 0$ converges to the unique positive semi-definite solution Σ to (30), i.e., $\lim_{t \rightarrow \infty} P(t|t-1) = \Sigma$. Furthermore, $\hat{\Phi} - L\hat{H}$ is a stable matrix, which implies the asymptotical stability of the steady-state fusion predictor. The proof is completed. $\square$

References

1. Sun, S.L.; Lin, H.L.; Ma, J.; Li, X.Y. Multi-sensor distributed fusion estimation with applications in networked systems: A review paper. *Inf. Fusion* **2017**, *38*, 122–134. [\[CrossRef\]](#)
2. Hu, J.; Wang, Z.D.; Chen, D.Y.; Alsaadi, F.E. Estimation, filtering and fusion for networked systems with network-induced phenomena: New progress and prospects. *Inf. Fusion* **2016**, *31*, 65–75. [\[CrossRef\]](#)
3. Xing, W.; Shen, J. Security control of cyber-physical systems under cyber attacks: A survey. *Sensors* **2024**, *24*, 3815. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Geng, H.; Wang, Z.D.; Chen, Y.; Yi, X.J.; Cheng, Y.H. Variance-constrained filtering fusion for nonlinear cyber-physical systems with the denial-of-service attacks and stochastic communication protocol. *IEEE-CAA J. Autom. Sin.* **2022**, *9*, 978–989. [\[CrossRef\]](#)
5. Huang, K.J.; Tan, L.S.; Peng, G. Stability of SDE-LJN system in the internet to mitigate constant-rate DDoS attacks. *Secur. Commun. Netw.* **2021**, *2021*, 4733190. [\[CrossRef\]](#)
6. Miao, S.T.; Li, Y.; Pan, Q. Honeypot game theory against DoS attack in UAV Cyber. *CMC-Comput. Mater. Contin.* **2023**, *76*, 2745–2762. [\[CrossRef\]](#)
7. Bu, X.H.; Yu, W.; Yin, Y.L.; Hou, Z.S. Event-triggered data-driven control for nonlinear systems under frequency-duration-constrained DoS attacks. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 1449–1460. [\[CrossRef\]](#)
8. Wang, Y.Z.; Wang, Z.D.; Zou, L.; Chen, Y.; Yue, D. Distributed Proportional-Integral fuzzy state estimation over sensor networks under energy-constrained denial-of-service attacks. *IEEE Trans. Cybern.* **2024**, *54*, 641–654. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Song, H.F.; Ding, D.R.; Dong, H.L. Distributed entropy filtering subject to DoS attacks in non-Gauss environments. *Int. J. Robust Nonlinear Control* **2019**, *30*, 1240–1257. [\[CrossRef\]](#)
10. Caballero-Aguila, R.; Frías-Bustamante, M.P.; Oya-Lechuga, A. Least-squares linear estimation for multirate uncertain systems subject to DoS attacks. *Int. J. Netw. Dyn. Intell.* **2025**, *4*, 100014. [\[CrossRef\]](#)
11. Geng, H.; Wang, Z.D.; Hu, J.; Alsaadi, F.E.; Cheng, Y.H. Outlier-resistant sequential filtering fusion for cyber-physical systems with quantized measurements under denial-of-service attacks. *Inf. Sci.* **2023**, *628*, 488–503. [\[CrossRef\]](#)
12. Zha, L.J.; Guo, Y.P.; Liu, J.L.; Xie, X.P.; Tian, E.G.; Cao, J.D. Protocol-based distributed security fusion estimation for networked systems with unknown bounded noise under quantization. *IEEE Trans. Control Netw. Syst.* **2024**, *11*, 2139–2150. [\[CrossRef\]](#)
13. Bi, X.R.; Sun, R.Z.; Li, J.X.; Hu, J. Event-triggered distributed fusion estimation for multi-sensor systems against mixed network attacks. *Int. J. Adapt. Control Signal Process.* **2026**, *40*, 50–62. [\[CrossRef\]](#)
14. Chen, K.Z.; Song, H.Y.; Shi, P.; Zhang, W.A.; Yu, L. Probability-guaranteed distributed set-membership secure fusion estimation against nonlinear hybrid attacks. *IEEE Trans. Autom. Sci. Eng.* **2025**, *22*, 5539–5550. [\[CrossRef\]](#)
15. Caballero-Águila, R.; Hu, J.; Linares-Pérez, J. Distributed estimation for uncertain systems subject to measurement quantization and adversarial attacks. *Inf. Fusion* **2025**, *120*, 103044. [\[CrossRef\]](#)
16. Tian, T.; Sun, S.L. Fusion estimation against mixed network attacks for systems with random parameter matrices, correlated noises, and quantized measurements. *Digit. Signal Process.* **2024**, *150*, 104523. [\[CrossRef\]](#)
17. Sahebsara, M.; Chen, T.; Shah, S.L. Optimal H_∞ filtering in networked control systems with multiple packet dropouts. *Syst. Control Lett.* **2008**, *57*, 696–702. [\[CrossRef\]](#)
18. Sun, S.L.; Ma, J. Linear estimation for networked control systems with random transmission delays and packet dropouts. *Inf. Sci.* **2014**, *269*, 349–365. [\[CrossRef\]](#)
19. Ma, J.; Sun, S.L. Information fusion estimators for systems with multiple sensors of different packet dropout rates. *Inf. Fusion* **2011**, *12*, 213–222. [\[CrossRef\]](#)
20. Ma, J.; Sun, S.L. Distributed fusion filter for networked stochastic uncertain systems with transmission delays and packet dropouts. *Signal Process.* **2017**, *130*, 268–278. [\[CrossRef\]](#)
21. Ding, J.; Sun, S.L.; Ma, J.; Li, N. Fusion estimation for multi-sensor networked systems with packet loss compensation. *Inf. Fusion* **2019**, *45*, 138–149. [\[CrossRef\]](#)

22. Zhu, D.P.; Chen, B.; Hong, Z.; Yu, L. Networked nonlinear fusion estimation under DoS attacks. *IEEE Sens. J.* **2021**, *21*, 7058–7066. [[CrossRef](#)]
23. Xu, M.; Zhang, Y.C.; Zhang, D.M.; Chen, B. Distributed robust dimensionality reduction fusion estimation under DoS attacks and uncertain covariances. *IEEE Access* **2021**, *9*, 10328–10337. [[CrossRef](#)]
24. Caballero-Aguila, R.; Hu, J.; Linares-Pérez, J. Two compensation strategies for optimal estimation in sensor networks with random matrices, time-correlated noises, deception attacks and packet losses. *Sensors* **2022**, *22*, 8505. [[CrossRef](#)] [[PubMed](#)]
25. Yang, J.; Zhang, W.A.; Guo, F.H. Adaptive distributed Kalman-like filter for power system with cyber attacks. *Automatica* **2022**, *137*, 110091. [[CrossRef](#)]
26. Song, H.Y.; Ye, S.Q.; Shi, P.; Zhang, W.A.; Yu, L. Multi-rate sampled-data secure fusion estimation against malicious hybrid attacks. *IEEE Trans. Signal Inf. Process. Netw.* **2025**, *11*, 401–412. [[CrossRef](#)]
27. Cheng, Z.J.; Yang, L.; Yuan, Q.Y.; Long, Y.R.; Ren, H.R. Distributed Consensus Estimation for Networked Multi-Sensor Systems under Hybrid Attacks and Missing Measurements. *Sensors* **2024**, *24*, 4071. [[CrossRef](#)] [[PubMed](#)]
28. He, G.Y.; Sun, S.L. Distributed fusion filter for linear discrete-time systems with DoS attack compensation. *J. Nat. Sci. Heilongjiang Univ.* **2024**, *41*, 347–354.
29. Jin, H.; Sun, S.L. Distributed filtering for multi-sensor systems with missing data. *Information Fusion* **2022**, *86–87*, 116–135. [[CrossRef](#)]
30. Niu, M.F.; Wen, G.H.; Zheng, W.X.; Ge, Q.B. Event-triggered distributed hierarchical secure estimation under DoS attacks. *IEEE Trans. Autom. Control* **2026**, *71*, 902–914. [[CrossRef](#)]
31. Ma, J.; Sun, S.L. Globally optimal distributed and sequential state fusion filters for multi-sensor systems with correlated noises. *Inf. Fusion* **2023**, *99*, 101885. [[CrossRef](#)]
32. Wang, M.H.; Sun, S.L. Self-tuning distributed fusion filter for multi-sensor networked systems with unknown packet receiving rates, noise variances, and model parameters. *Sensors* **2019**, *19*, 4436. [[CrossRef](#)] [[PubMed](#)]
33. Anderson, B.D.O.; Moore, J.B. *Optim. Filtering*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1979.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.