Distributed Filtering for Nonlinear Multi-Agent Systems with Biased Observations

Xingkang He\textsuperscript{a,b}, Wenchao Xue \textsuperscript{a,∗}, Xiaocheng Zhang \textsuperscript{a}, Haitao Fang \textsuperscript{a}

\textsuperscript{a} LSC, NCMIS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
\textsuperscript{b}Department of Automatic Control, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

Abstract

This paper considers the distributed filtering problem for a class of discrete-time stochastic nonlinear multi-agent systems with biased observations over switching communication topologies. We first build a general model for the systems by considering the distributed output feedback control and state-correlated observation bias. Then, we propose a three-staged distributed Kalman filter with guaranteed consistency, which means that the upper bounds of estimation error covariances can be online calculated by each agent. To alleviate the effect of biased observations, the event-triggered update scheme is proposed and proven to have a tighter bound of error covariance than the typical time-driven update scheme. Also, the proposed scheme can perform better in energy-constrained situations via abandoning redundant observations. Moreover, we rigorously prove the stability of the estimation error covariances for the proposed two distributed filters, respectively, under the mild conditions of collective observability of multi-agent system and joint connectivity of switching topologies. Finally, we carry out several numerical simulations to validate the theoretical results developed in this paper.

Key words: Multi-agent system; Distributed filtering; Kalman filter; Nonlinear system; Biased observation

1 Introduction

In recent years, multi-agent systems are broadly applied to sensor network \cite{1}, environment sensing \cite{2}, target tracking \cite{3,4}, smart grid \cite{5}, etc. As is well known, the state estimation problems of multi-agent systems are usually modeled as the distributed filtering problems, thus more and more researchers and engineers around the world are paying their attention to the research on the methods and theories of distributed filtering.

In the existing literatures on distributed filtering problems of linear multi-agent systems, many effective approaches and analysis tools have been provided. As is known, for linear stochastic systems, the optimal centralized Kalman filter can provide the estimation error covariances in a recursive manner. However, in distributed Kalman filters \cite{6-12} for multi-agent systems, the covariances cannot be obtained by each agent due to the unknown correlation between estimates of agents. To evaluate the estimation precision in terms of error covariance, we will investigate the consistency of a filter, which means an upper bound of estimation error covariance can be online calculated. Next, the main existing results on distributed filters will be investigated from two aspects, i.e., consistency and stability.

The estimation consistency plays an important role in real-time precision evaluation and covariance-based information fusion, which are essential to distributed state estimation of multi-agent systems. For linear time-invariant systems, without considering the consistency, \cite{13,14} proposed distributed filters with constant filtering gains, which yet confined the instability of the system dynamics. \cite{7,8,15} studied the distributed Kalman filters (DKFs) based on consensus or diffusion strategies, where the state estimates of agents were simply fused by scalar weights. As thus, the information of neighboring agents were not effectively utilized. There were still many other filters focusing on specific problems without maintaining the consistency. The authors of \cite{16} studied the stochastic link activation problem of distributed filtering under a scenario of sensor power constraint. The robust distributed filtering problems were well studied in \cite{17,18}. Recently, for linear multi-agent systems, the algorithms in \cite{9,11} possessed consistency, which enabled a sequence of upper bounds of estimation error covariances to be available by each agent.
Stability is one of fundamental properties for filtering algorithms. For unstable linear multi-agent systems with very limited local observable information, finding mild conditions to guarantee the stability of distributed filtering algorithms is a challenging problem. In the results on stability analysis, \cite{1,8,19} assumed local observability conditions of linear multi-agent systems, which confined the application scope of distributed filters. On the other hand, the observation errors including both observation bias and stochastic noise, can directly influence the consistency as well as stability of filters. Compared with stochastic observation noise, the observation bias may lead to larger loss in estimation performance of algorithms if the case is not well handled. This is attributed to that it is difficult for each agent to fuse the biased estimates from neighbors. Hence, it is necessary to provide information metric for the quality of observations, so as to judge whether the observations corrupted by state-correlated bias can be utilized at the update stage or simply abandoned. Moreover, existing results usually required the independence between the system state and the random bias \cite{20,21} which is difficult to be satisfied for feedback control systems with colored random bias processes. Therefore, the research on consistency and stability of distributed Kalman filters for the nonlinear systems seem still open, especially under collective observability, switching topologies, output feedback control input and biased observations, which are much more reasonable for engineering systems.

On the other hand, even based on the traditional centralized frameworks, designing filters with stability and consistency for nonlinear systems is a challenging problem \cite{6,22,23}. For known-model nonlinear systems, \cite{6,24} studied the linearized Kalman filter based algorithms. However, to guarantee the stability of nonlinear filters, they required the initial estimation error and noise should be sufficiently small, which is difficult to be met in practical applications. Moreover, due to the existence of outer disturbance or unmodeled dynamics, many practical systems contain uncertain nonlinearities besides nominally known nonlinearities. To deal with the unknown nonlinearities, some robust estimation methods, such as $H_{\infty}$ filters and set valued filters, were studied by researchers \cite{26,28}. But the estimation performance seem to be quite conservative in practical applications. Considering the instability of the linearization methods and the conservativeness of the robust filters, \cite{29} proposed an novel extended state based Kalman filter (ESKF) for a class of nonlinear uncertain systems. By employing the scheme of \cite{29} to handle the nonlinear uncertainty, this paper will construct the distributed filtering algorithm for nonlinear uncertain multi-agent systems.

In this paper, we consider the distributed filtering problem for a class of discrete-time stochastic nonlinear multi-agent systems with biased observations over switching communication topologies. The main contributions of this paper are threefold.

1. A three-staged time-driven distributed Kalman filter with guaranteed consistency is proposed for a class of nonlinear uncertain multi-agent systems. It is shown that the proposed filter enables a sequence of upper bounds of error covariances to be online calculated by each agent.

2. Based on an information metric for local observation statistics, we present an event-triggered observation update scheme. Moreover, based on this scheme, we propose an event-triggered distributed Kalman filter, which is shown to have a tighter bound of error covariance than that based on the typical time-driven update scheme.

3. We rigorously prove the stability of the estimation error covariances for the proposed two distributed filters. More importantly, our results do not require the assumptions of the noise independence among agents and the uniform nonsingularity of transition matrices, which are usually in the existing results but hardly to be satisfied in practice. Besides, the results suit a class of distributed output feedback control systems, such as the coupled tanks system \cite{30}.

The remainder of the paper is organized as follows: Section 2 is on the graph preliminaries and some useful definitions. Section 3 is on the problem formulation. Section 4 analyzes the distributed filter with time-driven update scheme. Section 5 studies the distributed filter with event-triggered update scheme. Section 6 shows the numerical simulations. The conclusion of this paper is given in Section 7.

### 1.1 Notations

The notations on mathematics and graphs used in this paper are standard and expressed in Table 1.

### 2 Graph Preliminaries and Useful Definitions

In this section, we will provide some graph preliminaries and useful definitions serving for the subsequents of this paper. The main notations are provided in Table 1.

We model the communication topologies of a multi-agent system as switching weighted digraphs $\mathcal{G}_i^{K}$, where $\mathcal{G}_i = (\mathcal{V}, \mathcal{E}_i, \mathcal{A}_i)$. In a weighted adjacency matrix $\mathcal{A}_i$, all the elements are nonnegative, row stochastic and the diagonal elements are all positive, i.e., $a_{i,i}^G > 0$, $a_{i,j}^G \geq 0, \sum_j a_{i,j}^G = 1$. If $a_{i,j}^G > 0$, $j \neq i$, there is a link $(i,j) \in \mathcal{E}_i$, which means node $i$ can directly receive the information of node $j$ through the communication channel. In this situation, node $j$ is called the neighbor or one of the neighbors of node $i$ and all the neighbors of node $i$ including itself can be represented by the set $\mathcal{N}_i^{G} = \{ j \in \mathcal{V} | (i,j) \in \mathcal{E}_i \} \bigcup \{ i \}$.

For a given positive integer $K$, the union of the $K$ digraphs $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}_1, \mathcal{A}_1), \cdots, \mathcal{G}_K = (\mathcal{V}, \mathcal{E}_K, \mathcal{A}_K)$ is denoted as $\sum_{k=1}^{K} \mathcal{G}_i = (\mathcal{V}, \sum_{k=1}^{K} \mathcal{E}_i, \sum_{k=1}^{K} \mathcal{A}_i)$. $\mathcal{G}$ is called strongly connected if for any pair nodes $(i_1, i_2)$, there exists a direct path from $i_1$ to $i_2$ consisting of
edges \((i_1, i_2), (i_2, i_3), \ldots, (i_{l-1}, i_l)\). We call \(\mathcal{G}_1, \ldots, \mathcal{G}_K\) jointly strongly connected if \(\sum_{l=1}^K \mathcal{G}_l\) is strongly connected.

In the Kalman filter for the linear time-varying systems with known exact noise statistics, the estimation error covariances can be recursively calculated. However, for the distributed Kalman filters \([9, 10]\), due to the unknown correlation between estimates of agents, the error covariances are usually unaccessible. In order to evaluate the estimation performance, the following definition of consistency is introduced.

**Definition 2.1** (\([37]\)) Suppose \(x_k\) is a random vector and \(\hat{x}_k\) is the estimate of \(x_k\). Then the pair \((\hat{x}_k, \Pi_k)\) is said to be **consistent** if \(E\{(\hat{x}_k - x_k)(\hat{x}_k - x_k)^T\} \leq \Pi_k\).

To study the estimation stability of filtering algorithms, the following definition is provided.

**Definition 2.2** Let \(e_{k,i}\) be the state estimation error of agent \(i\) at time \(k\), then the sequence of estimation error covariances \(E\{e_{k,i}e_{k,i}^T\}, k \in \mathbb{N}\), are said to be **stable** if \(\sup_{k \in \mathbb{N}} E\{e_{k,i}e_{k,i}^T\} < \infty\).

Denote \((\Omega, \mathcal{F}, P)\) as the basic probability space. \(\mathcal{F}_k\) stands for a filtration of \(\sigma\)-algebra \(\mathcal{F}\). A discrete-time sequence \(\{\xi_k, \mathcal{F}_k\}\) is said to be adapted if \(\xi_k\) is measurable to \(\mathcal{F}_k\). The definitions of ‘filtration’, ‘\(\sigma\)-algebra’ and ‘measurable’ are given in \([32]\).

**Definition 2.3** A discrete-time adapted sequence \(\{\xi_k, \mathcal{F}_k\}\) is called a **martingale difference sequence**, if \(E\{|\xi_k|\} < \infty\) and \(E\{\xi_k|\mathcal{F}_{k-1}\} = 0\), almost surely.

Since this paper studies a class of time-varying multi-agent systems, an useful definition is provided. Given a positive integer \(L\), a matrix sequence \(\{M_k, k \in \mathbb{N}\}\) and a positive scalar \(\beta\), define the time sequence \(\{T_l, l \in \mathbb{N}\}\) as

\[
T_{l+1} = \inf \{k \geq T_l + l | \lambda_{\min}(M_{T_l+s}M_{T_l+s}^T) \geq \beta, s \in [0 : L]\},
\]

**Definition 2.4** Given \(L\), if there exists an integer \(T > 0\) and a scalar \(\beta > 0\), such that for the defined time sequence \(\{T_l, l \in \mathbb{N}\}\) in (1),

\[
T_{l+1} - T_l \leq T,
\]

then \(\{T_l, l \in \mathbb{N}\}\) is called a **L-step supporting sequence (L-SS)** of \(\{M_k, k \in \mathbb{N}\}\).

**Remark 2.1** The definition of L-SS is introduced to study the nonsingularity of the time-varying transition matrices \(\{A_{k,i}\}\) given in next section. In many existing results, \(A_{k,i}\) is usually assumed to be nonsingular for \(\forall k \in \mathbb{N}\), which is removed in our paper.

### 3 Model Description and Problem setup

Consider the following model for a class of stochastic multi-agent systems with nonlinear uncertain dynamics and biased observations

\[
\begin{align*}
x_{k+1} &= \bar{A}_k x_k + \bar{G}_k f(x_k, k) + \bar{\omega}_k, \\
y_{k,i} &= H_{k,i} x_k + b_{k,i} + v_{k,i}, i \in \mathcal{V},
\end{align*}
\]

where \(x_k\) is the unknown \(n\)-dimensional system state, \(\bar{A}_k\) is the known state transition matrix and \(\bar{\omega}_k\) is the unknown zero-mean white process noise. \(f(x_k, k)\) is the \(p\)-dimensional nonlinear uncertain dynamics consisting of the known nominal model \(f(x_k, k)\) and some unknown disturbance. \(\bar{G}_k\) is the known matrix subject to \(\sup_{k \in \mathbb{N}} \{\bar{G}_k \bar{G}_k^T\} < \infty\). \(y_{k,i}\) is the \(m_i\)-dimensional observation vector obtained via agent \(i\), \(H_{k,i}\) is the known observation matrix, \(b_{k,i}\) is the unknown state-correlated stochastic observation bias of agent \(i\), and \(v_{k,i}\) is the stochastic zero-mean observation noise. \(\mathcal{V}\) is the number of agents over the system. Note that \(\bar{A}_k, \bar{G}_k, H_{k,i}\) and \(\bar{\omega}_k\) are simply known to agent \(i\). The above matrices and vectors have compatible dimensions.

Let \(\mathcal{F}_k = \sigma\{x_0, b_{0,i}, \bar{\omega}_j, v_{j,i}, i \in \mathcal{V}, 0 \leq j \leq k\}, f_k = f(x_k, k)\) for simplicity. In the following, we will provide several assumptions on the system structure and network topology.

**Assumption 3.1** On the multi-agent system (2), the following conditions hold.
1) The process noise \( \{w_k\}_{k=0}^{\infty} \) is independent of \( x_0 \) and \( \{v_{k,i}\}_{k=0}^{\infty}, i \in \mathcal{V} \), subject to \( E\{w_k \sigma_1^2\} \leq Q_k \), where \( \inf_{k \in \mathbb{N}} Q_k \geq Q > 0 \) and \( \sup_{k \in \mathbb{N}} Q_k \leq Q < \infty \).

2) The stochastic biases \( \{b_{k,i}\}_{i \in \mathcal{V}} \) are measurable to \( F_{k-1}, k \geq 1 \), and \( E\{b_{k,i}b_{k,i}^T\} \leq B_{k,i} \).

3) \( \{v_{k,i}, F_{k}\}_{i \in \mathcal{V}} \) are \( N \) martingale difference sequences such that \( E(v_{k,i}v_{k,i}^T) \leq R_{k,i}, \) where \( R_{k,i} \) are positive definite matrices such that \( \sup_{k \in \mathbb{N}} \left( H_{k,i}^T R_{k,i} H_{k,i}\right) < \infty, \forall i \in \mathcal{V} \).

4) \( E(\{X_0 - \bar{X}_{0,i}\}(X_0 - \bar{X}_{0,i})^T) \leq P_{0,i}, \) where \( \bar{X}_{0,i} \) is the estimate of \( X_0 \) for the \( i \)-th agent with \( X_0 \triangleq [x_0^T, f_0^T]^T \), and \( P_{0,i} > 0, i \in \mathcal{V} \).

Under the condition 2) of Assumption 3.1, the bias sequences \( \{b_{k,i}, F_{k-1}\}_{i=1}^{N} \) are \( N \) adapted sequences, thus the bias model is built in a general framework, which includes both deterministic phenomena \([33]\) and random noise. Different from \([11]\) where the observation noises of agents are independent, \( 3) \) of Assumption 3.1 allows the noise dependence between agents. In addition, compared with \([6, 11]\), which required the initial estimation error is sufficiently small, \( 4) \) of Assumption 3.1 is quite general and satisfied by setting a large \( P_{0,i} \).

**Assumption 3.2** For the system (2), there exists a positive integer \( L \), such that the matrix sequence \( \{A_k, k \in \mathbb{N}\} \) has an L-SS and \( \sup_{k \in \mathbb{N}} \{A_k A_k^T\} < \infty \).

Assumption 3.2 poses no requirement on the stability of the original system (2) which is necessary in many existing studies \([34, 35]\). Besides, within the scope of distributed filtering for time-varying systems, Assumption 3.2 is milder than that in \([6, 11]\), where the non-singularity of the system state transition matrix is needed for each time.

**Assumption 3.3** On the nonlinear uncertain dynamics \( f_k \), the following conditions hold.

1) The nonlinear dynamics \( f_k \) is measurable to \( F_k = \sigma\{x_0, b_{0,i}, \omega_{j-1}, v_{j,i}, i \in \mathcal{V}, 1 \leq j \leq k\} \) and \( F_0 = \sigma\{b_{0,i}, v_{0,i}, i \in \mathcal{V}\} \).

2) Denote \( u_k = f_{k+1} - f_k \) and \( u_k(i) \) is the \( i \)-th element of \( u_k \), then there exists a vector \( q_k \in \mathbb{R}^p \) such that
   \[
   E\{u_k^2(j)\} \leq q_k(j), j \in [1 : p],
   \]
   with \( \inf_{k \in \mathbb{N}} q_k(j) > \bar{q} > 0 \) and \( \sup_{k \in \mathbb{N}} \{q_k(j)\} < \infty, j \in [1 : p]\).

The first registration of Assumption 3.3 permits \( f_k \) to be implicitly related with \( \{x_j, y_j\}_{j=0}^{\infty}, i \in \mathcal{V} \). Under this setting, the model built in \((2)\) also considers the distributed output feedback control systems, such as the coupled tanks system \([30]\). Different from the existing result that treats the uncertain dynamics as a bounded total disturbance \([37]\), the requirement for the increment of the nonlinear dynamics in \((2)\) of Assumption 3.3 has no restriction on the boundedness of uncertain dynamics.

For the system (2), a new state vector, consisting of the original state \( x_k \) and the nonlinear uncertain dynamics \( f_k \), can be constructed. Then a modified system model with respect to the new state vector is given in the following.

\[
\begin{aligned}
\begin{bmatrix}
 x_{k+1} \\
 f_{k+1}
\end{bmatrix}
&= \begin{bmatrix}
 \bar{A}_k & \bar{G}_k \\
 0 & I_p
\end{bmatrix}
\begin{bmatrix}
 x_k \\
 f_k
\end{bmatrix}
+ \begin{bmatrix}
 \bar{w}_k \\
 0
\end{bmatrix} + \begin{bmatrix}
 0 \\
 f_{k+1} - f_k
\end{bmatrix} \\
\end{aligned}
\]

Under this setting, \( \{x_k, f_k\} \) are measurable to \( F_k \) and \( F_k \) is positive definite for time-varying systems, Assumption 3.2 is milder than that in \([6, 11]\), where the non-singularity of the system state transition matrix is needed for each time.

**Definition 3.1** On the relationship between \( \bar{A}_k \) in \((2)\) and \( A_k \) in \((3)\), the following conclusions hold.

1) Suppose \( \sup_{k \in \mathbb{N}} \{A_k A_k^T\} < \infty \) if and only if \( \sup_{k \in \mathbb{N}} \{\bar{A}_k \bar{A}_k^T\} < \infty \).

2) \( \{A_k | A_k \in \mathbb{R}^{n+p}, k \in \mathbb{N}\} \) has an L-SS if and only if \( \{\bar{A}_k | \bar{A}_k \in \mathbb{R}^{n}, k \in \mathbb{N}\} \) has an L-SS.

**Proof.** Please see the proof in Appendix A.

**Assumption 3.4** (Collective observability) There exist two positive integers \( N, M \), and a constant \( \alpha > 0 \) such that for any \( k \geq M \), there is

\[
\sum_{i=1}^{N} \sum_{j=k}^{k+N} \Phi_{j,i}^T H_{j,i} (R_{j,i} + B_{j,i})^{-1} H_{j,i} \Phi_{j,i} \geq \alpha I_n, \quad (3)
\]

where

\[
\Phi_{k,k} = I_n, \Phi_{k+1,k} = A_k, \Phi_{j,k} = \Phi_{j,j-1} \cdots \Phi_{k+1,k}(j > k).
\]
Assumption 3.4 is a standard collective observability condition for time-varying stochastic systems. If the system is time-invariant, then Assumption 3.4 degenerates to \((A, H)\) being observable \([9, 38]\). Besides, if the local observability conditions are satisfied \([11, 8, 19]\), then Assumption 3.4 holds, but not vice versa.

In this paper, the topologies of the networks are assumed to be switching digraphs \(\{G_{\sigma_k}, k \in \mathbb{N}\}\). \(\sigma_k\) is the graph switching signal defined \(\sigma_k : \mathbb{N} \rightarrow \Omega\), where \(\Omega\) is the set of the underlying network topology numbers. For convenience, the weighted adjacency matrix of the digraph \(G_{\sigma_k}\) is denoted as \(A_{\sigma_k} = [a_{i,j}(k)] \in \mathbb{R}^{N \times N}\). To analyze the switching topologies, we consider the infinity interval sequence of bounded, non-overlapping and contiguous time intervals \([k_l, k_{l+1})\), \(l = 0, 1, \cdots\), with \(k_0 = 0\) and \(0 \leq k_{l+1} - k_l \leq k^0\) for some integer \(k^0\). On the switching topologies of the multi-agent system, the following assumption is needed.

**Assumption 3.5** The digraph set \(\{G_{\sigma_k}, k \in [k_l, k_{l+1})\}\) is jointly strongly connected across the time interval \([k_l, k_{l+1})\) and the elements \(a_{i,j}(k)\) belong to \(\Psi, k \in \mathbb{N}^+, \) where \(\Psi\) is a finite set of arbitrary nonnegative real numbers.

Assumption 3.5 is on the conditions of the network topologies. Since the joint connectivity of the switching digraphs admits the network is unconnected at each moment, it is quite general for the networks confronting link failures. If the network remains connected at each moment or fixed \([9, 10]\), then Assumption 3.5 holds.

Due to the existence of stochastic biases with unknown correlation to the system state, the observations of the system (2) become more unreliable than that simply with random noises. In other words, employing the observations with typical time-driven methods many lead to the degradation of the estimation performance. Thus, for the time-varying system (3), different observation update protocols should be studied. In this paper, we consider two observation update schemes, namely time-driven update and event-triggered update, whose difference lies in whether the biased observation \(y_{k,i}\) is utilized at the update stage. Obeying a peer-to-peer communication strategy, we propose the following three-staged distributed filter structure of the system (3) for agent \(i\), \(\forall i \in \mathcal{V}\),

\[
\begin{aligned}
\dot{\tilde{x}}_{k,i} &= A_{k-1} \tilde{x}_{k-1,i} + D \hat{u}_{k-1} \\
\text{if } y_{k,i} \text{ is utilized:} & \quad \tilde{x}_{k,i} = \tilde{x}_{k,i} + K_{k,i}(y_{k,i} - H_{k,i}\tilde{x}_{k,i}) \\
\text{Otherwise:} & \quad \\
\dot{\tilde{x}}_{k,i} &= \tilde{x}_{k,i} \\
\tilde{x}_{k,i} &= \sum_{j \in N_i(k)} W_{k,i,j} \tilde{x}_{k,j},
\end{aligned}
\]

where \(\tilde{x}_{k,i}, \tilde{x}_{k,i}\) and \(\tilde{x}_{k,i}\) are the extended state’s prediction, update and estimate for agent \(i\) at the \(k\)th moment, respectively. \(K_{k,i}\) and \(W_{k,i,j}, j \in N_i(k)\), are the filtering gain matrix and the local fusion matrices, respectively. They are remain to be designed. Additionally,

\[
\begin{aligned}
\hat{u}_{k-1} &= (\hat{u}_{k-1}(1), \cdots, \hat{u}_{k-1}(p))^T \\
\hat{u}_{k-1}(j) &= \text{sat}(\hat{u}_{k-1}(j), \sqrt{\nu_{k-1}(j)}) \\
\tilde{u}_{k-1}(j) &= \tilde{f}_k(j) - \tilde{f}_{k-1}(j), j = 1, \cdots, p
\end{aligned}
\]

where \(\tilde{f}_k\) stands for the estimate of nominal model \(f_k\) by employing the former state estimates \(\{\hat{x}_{j,i}\}_{j=0}^k\). It is noted that the saturation function \(\text{sat}(\cdot, \cdot)\) is utilized to guarantee the boundedness of \(\hat{u}_{k-1}\).

The objectives of this paper are twofold:

- **a)** Under the time-driven and event-triggered update schemes, construct two distributed recursive filters in a distributed manner, respectively, such that the filters are consistent to estimate the extended state \(\hat{x}_k\).

- **b)** Based on the provided conditions on the system structure and network topology, prove the stability of the estimation error covariances for the proposed filters.

### 4 Distributed filter: time-driven update

In this section, for the filtering structure (5) with \(y_{k,i}\) employed each time, we will study the design methods of \(K_{k,i}\) and \(W_{k,i,j}\). Then we will find the conditions to guarantee the stability of the estimation error covariances for the proposed filter with the designed \(K_{k,i}\) and \(W_{k,i,j}\).

#### 4.1 Filter design

In the next, Lemma 4.1 provides a design method of fusion matrices \(\{W_{k,i,j}\}_{j \in N_i(k)}\), which can lead to the consistent estimate of each agent.

**Lemma 4.1** Consider the multi-agent system (3) with the filtering structure (5). Under Assumptions 3.1 and 3.3, for \(i \in \mathcal{V}, \forall \theta_{k,i} > 0, \forall \mu_{k,i} > 0\), the pairs \((\tilde{x}_{k,i}, \tilde{P}_{k,i})\) and \((\hat{x}_{k,i}, \hat{P}_{k,i})\) are all consistent, if

\[
W_{k,i,j} = a_{i,j}(k)P_{k,i}^{-1} \tilde{P}_{k,j}^{-1},
\]

where \(\hat{P}_{k,i}, \tilde{P}_{k,i}\) and \(P_{k,i}\) are recursively calculated through

\[
\begin{aligned}
\hat{P}_{k,i} &= (1 + \theta_{k,i} A_{k-1} K_{k-1,i} A_{k-1}^T + \frac{1 + \theta_{k,i}}{\theta_{k,i}} \hat{Q}_{k-1} + \tilde{Q}_{k-1})K_{k-1,i} \\
\tilde{P}_{k,i} &= (1 + \mu_{k,i} (I - K_{k,i} H_{k,i}) P_{k,i} (I - K_{k,i} H_{k,i})^T + K_{k,i} (R_{k,i} + \frac{1 + \mu_{k,i}}{\mu_{k,i}} B_{k,i}) K_{k,i}^T)K_{k-1,i} \\
P_{k,i} &= \left(\sum_{j \in N_i(k)} a_{i,j}(k) \hat{P}_{k,j}^{-1}\right)^{-1}
\end{aligned}
\]

with \(\hat{Q}_{k-1}\) and \(\tilde{Q}_{k-1}\) are defined in (B.2).
For the filtering gain matrix $K_{k,i}$, its design can be casted into an optimization problem with closed-form solution given in the following lemma.

**Lemma 4.2** Solving the optimization problem

$$K_{k,i}^* = \arg \min_{K_{k,i}} tr(\hat{P}_{k,i})$$

yields

$$K_{k,i}^* = \hat{P}_{k,i} H_{k,i}^T \left( H_{k,i} \hat{P}_{k,i} H_{k,i}^T + \frac{R_{k,i}}{1 + \mu_{k,i}} + \frac{B_{k,i}}{\mu_{k,i}} \right)^{-1}.$$

**Remark 4.1** In Lemma 4.1, it is shown the upper bounds of estimation error covariances at three typical stages can be obtained by each agent. The bounds can not only contribute to the design of fusion weights as well as filtering gain, but also evaluate the estimation accuracy in real time.

Summing up the results of Lemmas 4.1 and 4.2, the extended state based distributed Kalman filter (ESDKF) is provided in Algorithm 1.

**Algorithm 1** Extended State Based Distributed Kalman Filter (ESDKF):

**Prediction:** Each agent carries out a prediction operation

$$\begin{align*}
\hat{X}_{k,i} &= A_{k-1} \hat{X}_{k-1,i} + D \tilde{u}_{k-1}, \\
\hat{P}_{k,i} &= (1 + \theta_{k,i}) A_{k-1} \hat{P}_{k-1,i} A_{k-1}^T + \frac{1 + \theta_{k,i}}{\theta_{k,i}} \tilde{Q}_{k-1} + \tilde{Q}_{k-1}, \forall \theta_{k,i} > 0,
\end{align*}$$

where $\tilde{u}_{k-1}$ is provided in (6), $\tilde{Q}_{k-1}$ and $\tilde{\hat{Q}}_{k-1}$ are given in (B.2).

**Update:** Each agent uses its own observations to update the estimation

$$\begin{align*}
\hat{X}_{k,i} &= \hat{X}_{k,i} + \Delta \theta_{k,i} (\hat{Y}_{k,i} - H_{k,i} \hat{X}_{k,i}) \\
K_{k,i} &= \hat{P}_{k,i} H_{k,i}^T \left( H_{k,i} \hat{P}_{k,i} H_{k,i}^T + \frac{R_{k,i}}{1 + \mu_{k,i}} + \frac{B_{k,i}}{\mu_{k,i}} \right)^{-1} \\
\hat{P}_{k,i} &= (1 + \mu_{k,i}) (I - K_{k,i} H_{k,i}) \hat{P}_{k,i},
\end{align*}$$

**Local Fusion:** Each agent fuses ($\hat{X}_{k,i}$, $\hat{P}_{k,i}$) received from its neighbors

$$\begin{align*}
\hat{X}_{k,i} &= \hat{P}_{k,i} \sum_{j \in N_i(k)} a_{i,j} (k) \hat{P}_{k,j}^{-1} \hat{x}_{k,j}^{-1} \\
\hat{P}_{k,i} &= \left( \sum_{j \in N_i(k)} a_{i,j} (k) \hat{P}_{k,j}^{-1} \right)^{-1}.
\end{align*}$$

**Remark 4.2** Algorithm 1 is a fully distributed filter, i.e., its implementation simply requires the local information and the messages received from neighboring agents. Note the observations and observation statistics (i.e., $H_{k,i}, y_{k,i}, R_{k,i},$ and $D_{k,i}$) are not shared between agents, which contributes to the privacy of distributed estimation for multi-agent systems [39].

### 4.2 Stability

In this subsection, we will find the conditions to guarantee the stability of the estimation error covariances for ESDKF in Algorithm 1. Before that, we provide the following lemma for proof convenience.

**Lemma 4.3** Under Assumptions 3.1-3.3, if there are positive constants $\{\theta_{1,2}, \mu_{1,2}\}$ such that $\theta_{k,i} \in (\theta_{1,2})$ and $\mu_{k,i} \in (\mu_{1,2})$, the following two conclusions hold.

1) At the observation update stage, it holds that

$$\hat{P}_{k,i}^{-1} = \frac{P_{k,i}^{-1}}{1 + \mu_{k,i}} + H_{k,i}^T \Delta R_{k,i}^{-1} H_{k,i}, k \in \mathbb{N}, i \in \mathcal{V},$$

where $\Delta R_{k,i} = R_{k,i} + \frac{1 + \mu_{k,i}}{\mu_{k,i}} B_{k,i}$.

2) At the prediction stage, there exists a positive scalar $\eta$ such that

$$\hat{P}_{k,i}^{-1} \geq \eta \hat{A}_{k,i}^{-T} \hat{A}_{k,i}^{-1},$$

where $\{T_l, l \in \mathbb{N}\}$ is an L-SS of $\{A_k, k \in \mathbb{N}\}$.

**PROOF.** Please see the proof in Appendix D.

**Theorem 4.1** Consider the multi-agent system (3) with Algorithm 1. Under Assumptions 3.1-3.5, if $L > \max\{k^0, N\} + N$ and there are positive constants $\{\theta_{1,2}, \mu_{1,2}\}$ such that $\theta_{k,i} \in (\theta_{1,2})$ and $\mu_{k,i} \in (\mu_{1,2})$, then the estimation error covariances of each agent are stable, i.e.,

$$\sup_{k \in \mathbb{N}} \left\{ E\left\{ \langle \hat{X}_{k,i} - X_k \rangle \langle \hat{X}_{k,i} - X_k \rangle^T \right\} \right\} < +\infty, \forall i \in \mathcal{V}.$$

**PROOF.** Due to the consistency in Lemma 4.1, we turn to prove $\sup_{k \in \mathbb{N}} \hat{P}_{k,i} < \infty$. Under Assumption 3.2, $\{A_k \mid A_k \in \mathbb{R}^{n+p}, k \in \mathbb{N}\}$ has an L-SS, which is supposed to be $\{T_l, l \in \mathbb{N}\}$ subject to $L \leq T_{l+1} - T_l < T < \infty, \forall l \geq 0$, where $L > \max\{k^0, N\} + N$. Without loss of generality, we assume $T_0 \geq M$, where $M$ is given in Assumption 3.4. Otherwise, a subsequence of $\{T_l, l \in \mathbb{N}\}$ can always be obtained to satisfy the requirement. To prove the boundedness of $\hat{P}_{k,i}$, we divide the sequence set $\{T_l, l \geq 0\}$ into two non-overlapping time set: $\{T_l + L, l \geq 0\}$ and $\bigcup_{l \geq 0} [T_l + L + 1 : T_{l+1} + L - 1]$.

1) First, we consider the case of $k = T_l + L$, $l \geq 1$. For convenience, let $\bar{k} = T_l + L$. According to Lemma 4.3, we
obtain
\[
P^{-1}_{k,i} \geq \sum_{j \in \mathcal{N}_i(k)} a_{i,j}(k) \tilde{P}^{-1}_{k,j}
\]
(8)
\[
\geq \sum_{j \in \mathcal{N}_i(k)} a_{i,j}(k) \frac{\tilde{P}^{-1}_{k,j}}{1 + \mu_{k,j}} + \sum_{j \in \mathcal{N}_i(k)} a_{i,j}(k) H^T_{k,j} \Delta R^{-1}_{k,j} H_{k,j},
\]
(9)
\[
\geq \eta \left( \frac{\eta}{1 + \mu_{k,i}} \right) \sum_{j \in \mathcal{N}_i(k)} a_{i,j}(k) L^T_{k,j} \Delta R^{-1}_{k,j} H_{k,j}.
\]

Denote \( \delta = \frac{\eta}{(1 + \mu_{k,i})} \). By recursively applying (8) for \( L \) times, one has
\[
P^{-1}_{k,i} = \sum_{s=0}^{L-1} \delta^s \left[ \Phi^{-1}_{k,k-L} \sum_{j \in \mathcal{V}} a_{i,j} a_{j,k-s} P^{-1}_{k-L,j} \Phi^{-1}_{k,k-L} + \tilde{P}^{-1}_{k,i} \right],
\]
(10)
where
\[
\tilde{P}^{-1}_{k,i} = \sum_{s=0}^{L-1} \delta^s \left[ \Phi^{-1}_{k,k-s} \sum_{j \in \mathcal{V}} a_{i,j} a_{j,k-s} H^T_{k,s,j} \Delta R^{-1}_{k,s,j} H_{k,s,j} \right] \Phi^{-1}_{k,k-s},
\]
and \( a_{i,j} a_{j,k-s} \) is the \((i,j)\)th element of \( \Pi_{l=k-s}^L A_{\sigma_l} \). Since the first term on the right hand side of (9) is positive definite, we consider the second term \( \tilde{P}^{-1}_{k,i} \). Under Assumption 3.5, the jointly strongly connected network can lead to \( a_{i,j} a_{j,k-s} > 0, s \geq \max\{k^0, N\} \). From (10), one can obtain
\[
\tilde{P}^{-1}_{k,i} \geq a_{\min} \delta^L \sum_{s=0}^{L-1} \Phi^{-1}_{k,k-s} \left( \sum_{j \in \mathcal{V}} H^T_{k,s,j} \Delta R^{-1}_{k,s,j} H_{k,s,j} \right) \Phi^{-1}_{k,k-s}.
\]
It can be seen from Theorem 4.1 that, under mild conditions including collective observability of a multi-agent system and jointly strong connectedness of switching topologies, the proposed filter can effectively estimate the extended state, which consists of the original state and the nonlinear dynamics.

4.3 Design of parameters

Although the design principles on \( \theta_{k,i} \) and \( \mu_{k,i} \) have been provided in Theorem 4.1 to guarantee the stability of Algorithm 1, to improve the estimation performance, in this subsection, we give some optimization based design methods for the parameters \( \theta_{k,i} \) and \( \mu_{k,i} \). First of all, the objective functions ought to be given. Due to the unknown correlation between estimates of agents, the estimation covariance of distributed Kalman filters is usually not attainable. Thanks to the consistency of Algorithm 1, we can use \( \hat{P}_{k,i} \) and \( \tilde{P}_{k,i} \) to take the roles.

a) Design of \( \theta_{k,i} \)

At the prediction stage, the design of the parameter \( \theta_{k,i} \) is aimed to minimize the trace of \( \hat{P}_{k,i} \). Mathematically, the optimization problem on \( \theta_{k,i} \) is given as
\[
\theta^*_{k,i} = \arg \min_{\theta_{k,i}} tr(\hat{P}_{k,i})
\]
where \( \hat{P}_{k,i} = (1 + \theta_{k,i}) A_{k-1} \hat{P}_{k-1,i} A^T_{k-1} + \frac{1 + \theta_{k,i}}{\theta_{k,i}} \bar{Q}_{k-1} + \bar{Q}_{k-1} \).

2) Second, we consider the time set \( \bigcup_{i=0}^{T} [T_i + L - 1] \). Considering (9), we have \( P_{k+i} \leq \tilde{P}_k \). Since the length of the interval is bounded by \( T < \infty \), we can just consider the prediction stage to study the boundedness of \( P_{k,i} \), for \( k \in [T_i + L - 1 : T_{i+1} + L - 1] \). Under Assumption 3.2 and Lemma 3.1, there is a scalar \( \beta_1 > 0 \) such that \( A_k A_k^T \leq \beta_1 I_n \). Due to \( sup_k Q_k < \infty \), it is safe to conclude that there exists an constant matrix \( P^\text{mid} \), such that
\[
P_{k,i} \leq P^\text{mid}, k \in [T_i + L - 1 : T_{i+1} + L - 1]
\]
(12)
3) Finally, for the time interval \([0 : T_0 + L]\), there exists a constant matrix \( \tilde{P} \), such that
\[
P_{k,i} \leq \tilde{P}, k \in [0 : T_0 + L].
\]
(13)

According to (11), (12) and (13), we have \( sup_k \subseteq P_{k,i} < \infty \). Q.E.D.
Since Problem 1 is a convex optimization problem, which can be numerically solved by existing convex optimization methods. In Proposition 4.1, we provide the closed-form solution of Problem 1.

**Proposition 4.1** Solving the Problem 1 yields the closed-form solution

\[ \theta_{k,i}^* = \sqrt{\frac{\text{tr}(\hat{Q}_{k-1})}{\text{tr}(A_{k-1}P_{k-1,i}A_{k-1}^T)}} , \quad i \in \mathcal{V}, \]

subject to \( \theta_{k,i}^* > 0. \)

**PROOF.** Consider \( \text{tr}(\hat{P}_{k,i}) \), then we have \( \text{tr}(\hat{P}_{k,i}) = (1 + \theta_{k,i}^*)\text{tr}(A_{k-1}P_{k-1,i}A_{k-1}^T) + \frac{\text{tr}(\hat{Q}_{k-1})}{\theta_{k,i}^*} + \text{tr}(\hat{Q}_{k-1}) \).

Then \( \theta_{k,i}^* = \arg \min_{\theta_{k,i}^*} \text{tr}(\hat{P}_{k,i}) = \arg \min_{\theta_{k,i}^*} f_k(\theta_{k,i}^*), \) where \( f_k(\theta_{k,i}^*) = \theta_{k,i}^*\text{tr}(A_{k-1}P_{k-1,i}A_{k-1}^T) + \frac{\text{tr}(\hat{Q}_{k-1})}{\theta_{k,i}^*} \), which is minimized if \( \theta_{k,i}^*\text{tr}(A_{k-1}P_{k-1,i}A_{k-1}^T) = \frac{\text{tr}(\hat{Q}_{k-1})}{\theta_{k,i}^*} \). As a result, \( \theta_{k,i}^* = \sqrt{\frac{\text{tr}(\hat{Q}_{k-1})}{\text{tr}(A_{k-1}P_{k-1,i}A_{k-1}^T)}} \). Due to \( P_{k-1,i} > 0, \hat{Q}_{k-1} > 0 \) and \( A_{k-1} \neq 0 \), we have \( \theta_{k,i}^* > 0. \) Q.E.D.

From Proposition 4.1, it can be seen that the time-varying \( \theta_{k,i}^* \) can be calculated in a step-wise manner based on the current information of agent \( i \).

**b) Design of \( \mu_{k,i} \)**

On the parameter \( \mu_{k,i} \) in the observation update, considering the consistency, we aim to design \( \mu_{k,i} \) to minimize \( \text{tr}(\hat{P}_{k,i}) \), which is cast into the following optimization problem

\[ \mu_{k,i}^* = \arg \min_{\mu_{k,i}} \text{tr}(\hat{P}_{k,i}), \quad \text{s.t.} \quad \mu_{k,i} > 0, \]

where

\[
\hat{P}_{k,i} = (1 + \mu_{k,i})(I - K_{k,i}H_{k,i})\hat{P}_{k,i}, \\
K_{k,i} = \hat{P}_{k,i}H_{k,i}^T\left(H_{k,i}\hat{P}_{k,i}H_{k,i}^T + \frac{R_{k,i}}{1 + \mu_{k,i}} + \frac{B_{k,i}}{\mu_{k,i}}\right)^{-1}.
\]

**Remark 4.3** Problem 2 is not a convex optimization problem, which can be solved by many existing non-convex optimization methods, such as the quasi-Newton methods or the nonlinear least square [40].

5 Distributed filter: event-triggered update

In this section, we will study an even-triggered observation update scheme to deal with the corrupted observations by random noise and state-correlated bias. Based on the update scheme, a new distributed Kalman filter is proposed and analyzed in terms of the covariance stability.

5.1 Event-triggered update scheme

Due to the influence of random noise and observation bias over the multi-agent system (2), some corrupted observations may lead to the performance degradation of individual agent. Thus, we aim to provide a scheme to decide when the observation is utilized or abandoned. First, to measure the information quantity of local observations with random noise and state-correlated bias for agent \( i \), we introduce the information metric \( S_{k,i} \), defined as

\[ S_{k,i} \triangleq H_{k,i}^T \left( R_{k,i} + \frac{(1 + \mu_{k,i})}{\mu_{k,i}} B_{k,i} \right)^{-1} H_{k,i}. \]  

(14)

Note that \( S_k \) is defined based on the known local rough statistics, which admits individual agent can online calculate the metric to statistically evaluate the quantity of current observation information. If the eigenvalues of \( R_{k,i} \) and \( B_{k,i} \) are pretty larger than the observation matrix \( H_{k,i} \), then the eigenvalues of \( S_{k,i} \) will be quite small, which means the quantity of the observation information is not sufficient. In the following, we define the update event and the event-triggered scheme.

**Update Event \( \mathcal{E} \):**

At time \( k \), agent \( i \) utilizes the observation \( y_{k,i} \) to update the estimate of the system state \( X_{k,i} \) using Algorithm 1.

**Event-triggered Scheme:**

The event \( \mathcal{E} \) is triggered if

\[ \lambda_{\max} \left( S_{k,i} - \frac{\mu_{k,i}}{1 + \mu_{k,i}} \bar{P}_{k,i}^{-1} \right) > \tau, \]  

(15)

where \( \tau > 0 \) is the preset triggering threshold of the observation update. Otherwise, agent \( i \) skips the current observation update to the fusion stage, which means \( y_{k,i} \) will be abandoned.

**Remark 5.1** The triggering scheme in (15) shows if the current information is more sufficient in one channel at least than the prediction information, the observation is worth utilizing at the update stage. The triggering threshold \( \tau \) is used as a measure of information increment.

The following lemma provides an equivalent form of the triggering scheme.

**Lemma 5.1** The event-triggered scheme (15) is satisfied if and only if

\[ \lambda_{\max} \left( \tilde{P}_{k,i}^{-1} - \bar{P}_{k,i}^{-1} \right) > \tau, \]  

(16)
where $\hat{P}_{k,i} = (1 + \mu_{k,i}) (I - K_{k,i} H_{k,i}) \tilde{P}_{k,i}$ and $\tau \geq 0$.

**PROOF.** Employing the matrix inverse formula on $\hat{P}_{k,i}$ yields $\hat{P}_{k,i}^{-1} = \frac{\hat{P}_{k,i}^{-1}}{1 + \mu_{k,i}} + H_{k,i}^T \Delta R_{k,i}^{-1} H_{k,i},$ where $\Delta R_{k,i} = R_{k,i} + \frac{1}{\mu_{k,i}} B_{k,i}$. Substituting $\hat{P}_{k,i}^{-1}$ into (16) and considering (14), the conclusion of this lemma holds. Q.E.D.

Since the estimates of Algorithm 1 are consistent, $\hat{P}_{k,i}^{-1}$ and $\tilde{P}_{k,i}^{-1}$ stand for the lower bounds of information matrices at the update stage and the prediction stage, respectively. $(\hat{P}_{k,i}^{-1} - \tilde{P}_{k,i}^{-1})$ reflects the variation of statistical information resulted from a new observation. In light of Lemma 5.1, the event $E$ is triggered if sufficiently new information is accumulated at the update stage.

### 5.2 ESDKF with event-triggered update

Based on the event-triggered scheme (15) and Algorithm 1, we can obtain the extended state based distributed Kalman filter with event-triggered update scheme in Algorithm 2.

**Algorithm 2** ESDKF based on event-triggered update:

**Prediction:** Each agent carries out a prediction operation

\[
\tilde{x}_{k,i} = A_{k,i-1} \tilde{x}_{k-1,i} + D \hat{u}_{k-1},
\]

\[
\tilde{P}_{k,i} = (1 + \theta_{k,i}) A_{k-1} \tilde{P}_{k-1,i} A_{k,i}^T + \frac{1 + \theta_{k,i}}{\theta_{k,i}} \tilde{Q}_{k-1} + \tilde{Q}_{k-1}, \forall \theta_{k,i} > 0,
\]

where $\hat{u}_{k-1}$ is provided in (6), $\tilde{Q}_{k-1}$ and $\tilde{Q}_{k-1}$ are given in (B.2).

**Event-triggered update:**

\[
K_{k,i} = \hat{P}_{k,i} H_{k,i}^T \left( H_{k,i} \hat{P}_{k,i} H_{k,i}^T + \frac{R_{k,i}}{1 + \mu_{k,i}} + \frac{B_{k,i}}{\mu_{k,i}} \right)^{-1} H_{k,i}^T \tilde{P}_{k,i},
\]

\[
\hat{P}_{k,i} = (1 + \mu_{k,i}) (I - K_{k,i} H_{k,i}) \hat{P}_{k,i},
\]

If (15) is satisfied, then

\[
\tilde{x}_{k,i} = \tilde{x}_{k,i} + K_{k,i} (y_{k,i} - H_{k,i} \tilde{x}_{k,i})
\]

\[
\hat{P}_{k,i} = \hat{P}_{k,i},
\]

Otherwise,

\[
\tilde{x}_{k,i} = \tilde{x}_{k,i}, \quad \hat{P}_{k,i} = \hat{P}_{k,i}.
\]

**Local Fusion:** Each agent fuses ($\tilde{x}_{k,j}, \hat{P}_{k,j}$) received from its neighbors

\[
\tilde{x}_{k,i} = \hat{P}_{k,i} \sum_{j \in N_i} a_{i,j} (k) \tilde{P}_{k,j}^{-1} \tilde{x}_{k,j},
\]

\[
\hat{P}_{k,i} = \left( \sum_{j \in N_i} a_{i,j} (k) \tilde{P}_{k,j}^{-1} \right)^{-1}.
\]

**Lemma 5.2** Consider the multi-agent system (3) with Algorithm 2. Under Assumptions 3.1 and 3.3, the pairs ($\tilde{x}_{k,i}, \hat{P}_{k,i}$), ($\tilde{x}_{k,i}, \tilde{P}_{k,i}$) and ($\tilde{x}_{k,i}, \hat{P}_{k,i}$) are all consistent.

**PROOF.** Since the difference between Algorithm 1 and Algorithm 2 lies in the observation update stage, here for convenience, we simply consider the observation update of Algorithm 2. If (15) is satisfied, then the estimate $\tilde{x}_{k,i}$ will be updated with the current observation $y_{k,i}$. In this case, the proof is the same as that of Lemma 4.2. If (15) is not satisfied, then $\tilde{x}_{k,i} = \tilde{x}_{k,i}, \hat{P}_{k,i} = \hat{P}_{k,i}$, which guarantees the consistency of ($\tilde{x}_{k,i}, \hat{P}_{k,i}$). Q.E.D.

Lemma 5.2 shows that, although the observation update scheme has been modified, Algorithm 2 inherits the consistency of Algorithm 1. On Algorithm 2, we have the following lemma.

**Lemma 5.3** Under the event-triggered update scheme, for Algorithm 2 it holds that

\[
\hat{P}_{k,i}^{-1} \geq \frac{\hat{P}_{k,i}^{-1}}{1 + \mu_{k,i}} + H_{k,i}^T \Delta R_{k,i}^{-1} H_{k,i} - \tau I_{n+p},
\]

where $\Delta R_{k,i} = R_{k,i} + \frac{1 + \mu_{k,i}}{\mu_{k,i}} B_{k,i}$.

**PROOF.** If the update event in (16) is triggered, then $\hat{P}_{k,i} = \hat{P}_{k,i}$. According to 2) of Lemma 4.3, we have $\hat{P}_{k,i}^{-1} = \frac{\hat{P}_{k,i}^{-1}}{1 + \mu_{k,i}} + H_{k,i}^T \Delta R_{k,i}^{-1} H_{k,i}$, Thus, conclusion of Lemma 5.3 holds in this case. If the update event in (16) is not triggered, then $\hat{P}_{k,i} = \hat{P}_{k,i}$. Besides, according to the scheme, it follows that $\hat{P}_{k,i} \geq \hat{P}_{k,i} - \tau I_{n}$. Q.E.D.

The relationship between Algorithm 1 and Algorithm 2 is studied in the following proposition, which shows the event-triggered observation update scheme can lead to a tighter bound of error covariance than the typical time-driven observation update.

**Proposition 5.1** Let $P_{A_{k,i}}^{s}$, $s = 1, 2$, be the $P_{k,i}$ matrix of Algorithm 1 and Algorithm 2, respectively. If the two algorithms share the same initial setting and $\tau = 0$, then $P_{A_{k,i}}^{s} \leq P_{A_{k,i}}^{s}$. **PROOF.** In light of Lemma 5.3, for $\tau = 0$, $\hat{P}_{k,i}^{-1} \geq \hat{P}_{k,i}^{-1} + H_{k,i}^T \Delta R_{k,i}^{-1} H_{k,i} = \hat{P}_{k,i}^{-1},$ which means $\hat{P}_{k,i} \leq \hat{P}_{k,i}$, where $\hat{P}_{k,i}$ corresponds to the observation update of Algorithm 1. By using the mathematical induction method, the proof of this proposition can be finished. Q.E.D.

**Remark 5.2** Compared with the ESDKF in Algorithm 1, Algorithm 2 is able to obtain better estimation performance since it discards corrupted observations that may deteriorate the estimation performance. Meanwhile, for some applications [27, 28] that estimator and sensor are located at different geographical locations with energy-constrained


communication channels, it is suggested to judge which observations contain novel information and to decide when the observations are transmitted from the sensor to the remote estimator. These tasks can be achieved by the proposed event-triggered update scheme of this paper.

5.3 Stability

In this subsection, considering the system (3) with Algorithm 2, we will provide the stability analysis with respect to the sequence of estimation error covariances.

Theorem 5.1 Consider the multi-agent system (3) with Algorithm 2. Under Assumptions 3.1-3.5, if $L > \max\{k, N\} + N$ and there are positive constants $\{\theta_1, \theta_2, \mu_1, \mu_2\}$ such that $\theta_{k,i} \in (\theta_1, \theta_2)$ and $\mu_{k,i} \in (\mu_1, \mu_2)$, then there exists a scalar $\vartheta > 0$, such that for $0 \leq \tau < \vartheta$, the estimation error covariances of each agent are stable, i.e.,

$$\sup_{k \in \mathbb{N}} \left\{ E\{(\tilde{X}_{k,i} - X_k)(\tilde{X}_{k,i} - X_k)^T\} \right\} < +\infty, \forall i \in V.$$

**PROOF.** According to Lemma 5.2, we turn to prove $\sup_{k \in \mathbb{N}} P_{k,i} < \infty$. Similar to the proof of Theorem 4.1, we also consider the time set $\{T_l + L, l \geq 0\}$ and $\bigcup_{l \geq 0}[T_l + L + 1 : T_{l+1} + L - 1]$. To prove the upper boundedness of $P_{k,i}$ in Algorithm 2 over $\mathbb{N}$, we focus on proving its boundedness over the set $\{T_l + L, l \geq 0\}$. The rest part can be covered by taking the same method as the proof of Theorem 4.1. Let $\bar{k} \triangleq T_l + L$. According to Lemmas 4.3 and 5.3, we obtain

$$P_{k,i}^{-1} \geq \sum_{j \in N_i(\bar{k})} a_{i,j}(\bar{k}) \bar{P}_{k,j}^{-1} \geq \sum_{j \in N_i(\bar{k})} a_{i,j}(\bar{k}) \frac{P_{k,j}^{-1}}{1 + \mu_{k,j}} + \sum_{j \in N_i(\bar{k})} a_{i,j}(\bar{k}) H_{k,j}^T R_{k,j}^{-1} H_{k,j} - \tau I_{n+p},$$

$$\geq \frac{\eta}{1 + \mu_2} \sum_{j \in N_i(\bar{k})} a_{i,j}(\bar{k}) A_{k-1,j}^T A_{k-1,j}^{-1} + \sum_{j \in N_i(\bar{k})} a_{i,j}(\bar{k}) \left( H_{k,j}^T R_{k,j}^{-1} H_{k,j} - \tau I_{n+p} \right).$$

Denote $\delta = \frac{\eta}{1 + \mu_2}$. By recursively applying (17) for $L$ times, one has

$$P_{k,i}^{-1} \geq \delta^L \Phi_{k,i}^{-1} \left[ \sum_{j \in V} a_{i,j}^T k-P_{k-1,j} \right] \Phi_{k,i}^{-1} + \bar{P}_{k,i}^{-1},$$

where

$$\bar{P}_{k,i}^{-1} = \sum_{s=0}^{L-1} \delta^s \left[ \sum_{j \in V} a_{i,j}^T \Phi_{k,s}^{-1} + \bar{P}_{k,i}^{-1} \right].$$

Since the first term on the right hand side of (18) is positive definite, we have $P_{k,i} \geq \bar{P}_{k,i}^{-1}$. To prove the conclusion, we turn to prove there is a constant matrix $S > 0$, such that $P_{k,i} \leq S$. Under the conditions of this theorem, it follows from the proof of Theorem 4.1 that there exists a constant $\pi > 0$, such that

$$\sum_{s=0}^{L-1} \delta^s \Phi_{k,s}^{-T} \left[ \sum_{j \in V} a_{i,j}^T \Phi_{k,s}^{-1} \right] \Phi_{k,s}^{-1} \geq \pi I_{n+p},$$

Denote $\Xi_k = \sum_{s=0}^{L-1} \delta^s \Phi_{k,s}^{-T} \Phi_{k,s}^{-1}$. To guarantee $\inf_{k} \bar{P}_{k,i} > 0$, it is sufficient to prove that there exists a constant matrix $S$, such that

$$\pi I_{n+p} - \tau \Xi_k \geq S > 0.$$

Under Assumption 3.2, there exists a constant scalar $\vartheta > 0$, such that $\sup_{k \in \mathbb{N}} P_{k,i} \leq \vartheta I_{n+p}$. Let $S = (\pi - \tau \vartheta)I_{n+p}$, it can be seen that a sufficient condition to guarantee (20) is $0 \leq \tau < \frac{\vartheta}{\vartheta}$. Choose $\vartheta = \frac{\vartheta}{\vartheta} > 0$, then we can obtain the conclusion of this theorem. Q.E.D.

Remark 5.3 According to Theorem 5.1, the estimation error covariances of Algorithm 2 remain stable by designing proper event-triggered update parameter $\tau$. However, if we set a quite large $\tau$ such that the triggering condition (15) of the update scheme is hardly satisfied, then most of observation information will be lost. As a result, the collective observability condition in Assumption 3.4 may not hold, which means the stability of estimation error covariances is not guaranteed.

6 Numerical Simulations

In this section, numerical simulations are carried out to demonstrate the aforementioned theoretical results and show the effectiveness of the proposed algorithms. To this end, let us consider an object whose motion is described by the
kinematic model [43] with uncertain dynamics:
\[
x_{k+1} = \begin{pmatrix} 1 & 0 & T \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} x_k + \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & T/3 \end{pmatrix} \left( \sin (x_k (3)) + k \right) + \omega_k, k = 1, 2, \ldots ,
\]
where \( T = 0.1 \) is the sampling step and \( x_k \) is the unknown state vector consisting of four-dimensional components along the coordinate axes. The covariance of process noise \( \omega_k \) is \( Q_k = \text{diag} ([4, 4, 1, 1]) \).

The kinematic state of the object is observed by means of four agents modeled as
\[
y_{k,i} = \tilde{H}_{k,i} x_k + b_{k,i} + v_{k,i}, i = 1, 2, 3, 4.
\]
The observation matrices are supposed to be switching with time, given by
\[
\begin{align*}
\tilde{H}_{k,i} &= \tilde{H}_{\text{mod}(i+[\frac{n-i}{4}],4)+1} \\
\tilde{H}_1 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tilde{H}_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\
\tilde{H}_3 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tilde{H}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.
\end{align*}
\]
The state-correlated bias \( b_{k,i} \) satisfies
\[
b_{k,i} = \text{sat} (\sin (x_k^i (1) + x_k^i (2)) + b_{0,i}, 1),
\]
where the initial bias \( b_{0,i} \) is generated uniformly within \([-1,1]\). The observation noise \( [v_{k,1}, \ldots , v_{k,N}]^{T} \) is i.i.d. Gaussian with covariance \( R_k = 10 \cdot I_4 \). Additionally, the agent network’s communication topology is assumed to be directed and switching, whose adjacency matrix is selected from
\[
\begin{align*}
A_1 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}, A_2 = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0.3 & 0.4 & 0.3 \\ 0 & 0.5 & 0.5 & 0 \end{pmatrix}, \\
A_3 &= \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 1 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}.
\end{align*}
\]
And the topology switching signal \( \sigma_k \) is
\[
\begin{align*}
\sigma_k &= 1, k = [1 : 5], [16 : 20], \cdots \\
\sigma_k &= 2, k = [6 : 10], [21 : 25], \cdots \\
\sigma_k &= 3, k = [11 : 15], [26 : 30], \cdots
\end{align*}
\]

Next, we conduct the numerical simulations through Monte Carlo experiment, in which 500 runs for the considered algorithms are implemented, respectively. The Root Mean Square Error (RMSE) averaged over all the agents is defined as
\[
\text{RMSE}_k = \sqrt{\frac{1}{3 \cdot n} \sum_{i=1}^{n} \text{MSE}_{k,i}},
\]
where
\[
\text{MSE}_{k,i} = \frac{1}{500} \sum_{j=1}^{500} \left( x_{k,i}^j - \hat{x}_{k,i}^j \right)^T \left( x_{k,i}^j - \hat{x}_{k,i}^j \right),
\]
and \( \hat{x}_{k,i}^j \) is the state estimate of the \( j \)th run of agent \( i \) at the \( k \)th time instant. Besides, we denote \( P_k = \frac{1}{4} \sum_{i=1}^{4} P_{k,i} \).

It is assumed that the initial state is a zero-mean random vector with covariance \( P_0 = \text{diag} ([10, 10, 1, 1]) \). The parameters of Algorithm 1 and Algorithm 2 are set in the following
\[
\begin{align*}
X_{i,0} &= 0_{6 \times 1}, P_{i,0} = \begin{pmatrix} P_0 & 0_{4 \times 2} \\ 0_{2 \times 4} & I_2 \times 2 \end{pmatrix}, \\
q_{k,i} &= 1, \mu_{k,i} = 0.05, \tau_i = 0.01, \forall i = 1, 2, 3, 4.
\end{align*}
\]

6.1 Performance Evaluation

First, we carry out numerical simulations for Algorithm 1 (i.e., ESDKF) and Algorithm 2 based on the above parameter setting. The results are given in Fig. 1 to Fig. 3. Fig. 1 shows the behavior of the RMSE and \( \sqrt{\text{tr} (P_k)} \) of the proposed ESDKF, from which one can see the estimation error covariances of the proposed ESDKF keep stable in the given period and the consistency of each agent remains. Meanwhile, because of the switching of observation matrices and communication topologies, the RMSE and \( \sqrt{\text{tr} (P_k)} \) show periodic fluctuations of rise and fall. Fig. 2 shows the difference value between the RMSE of Algorithm 1 and the RMSE of Algorithm 2, and the difference value between the
Fig. 2. RMSE difference between Algorithm 1 and Algorithm 2

From Fig. 2, one can see the event-triggered update scheme of update in Algorithm 2 can, to some extent, reduce the $P$ matrix and the estimation error covariance. Meanwhile, Fig. 3 shows the triggering time instants of observation update for each agent. Because of the periodic switching of observation matrices, the triggering time instants of all agents are also periodic. Thus, compared with Algorithm 1, Algorithm 2 can reduce the frequency of observation update with competitive estimation performance.

6.2 Comparisons with other algorithms

Second, in this subsection, we carry out numerical simulations to compare the proposed ESDKF with other three algorithms, namely, centralized Kalman filter (CKF), Distributed State Estimation with Consensus on the Posteriors (DSEA-CP) [9] and Collaborative Scalar-gain Estimator (CSGF) [13]. The performance comparison result of the above algorithms is shown in Fig. 4. From this figure, one can see that in the given period both CSGF, DSEA-CP and CKF become unstable, but the proposed ESDKF still keep stable. The stability of ESDKF lies in its capability in handling with uncertain nonlinear dynamics and state-correlated observation bias. For the CSGF, since it is a distributed filter based on a fixed filtering gain with limited ability in tracking an unstable dynamics, the estimation error covariance of CSGF is divergent at a very fast speed. As for the DSEA-CP, due to the switching topologies, the observability condition of this filter may not be satisfied at some periods. As a result, the estimation error covariances of DSEA-CP are divergent with a periodic fluctuations of rise and fall. While, for the CKF, since it is a centralized filter without being affected by the topology changing, its error covariances are divergent at a stable speed.

Based on the above results, we can see the proposed ESDKF in Algorithm 1 and its event-triggered version in Algorithm 2 are effective distributed filtering algorithms.

Fig. 3. Update triggering time of Algorithm 2 for agents

Fig. 4. Estimation performance of different algorithms

7 Conclusion

In this paper, we first built a model for a class of general nonlinear stochastic multi-agent systems with biased observations by considering the distributed output feedback control systems. Then, we proposed a three-staged distributed Kalman filter with guaranteed consistency. To alleviate the effect of biased observations, we proposed a novel event-triggered update scheme and proved this scheme can lead to a tighter bound of error covariance than the typical observation update. Moreover, we rigorously proved the stability of the estimation error covariances for the proposed two distributed filters under some quite mild conditions, including the collective observability of a multi-agent system and the joint connectivity of switching topologies. Finally, we carried out several numerical simulations to validate the effec-
tiveness of the distributed filters and the theoretical results of our paper.

A Proof of Lemma 3.1

Recall \( A_k = \begin{pmatrix} \bar{A}_k & \bar{G}_k \\ 0 & I_p \end{pmatrix} \), and \( \sup_{k \in \mathbb{N}} \{ \bar{G}_k \bar{G}_k^T \} < \infty \), then we have

\[
A_k A_k^T = \begin{pmatrix} \bar{A}_k \bar{A}_k^T + \bar{G}_k \bar{G}_k^T & \bar{G}_k^T I_p \\ I_p & 0 \end{pmatrix}. \tag{A.1}
\]

First, we consider the proof of (1). We consider the sufficiency of (1). According to (A.1), we have \( A_k A_k^T \leq 2 \begin{pmatrix} \bar{A}_k \bar{A}_k^T + \bar{G}_k \bar{G}_k^T & \bar{G}_k^T I_p \\ I_p & 0 \end{pmatrix} \). Thus, \( \sup_{k \in \mathbb{N}} \{ \bar{A}_k \bar{A}_k^T \} < \infty \) and \( \sup_{k \in \mathbb{N}} \{ \bar{G}_k \bar{G}_k^T \} < \infty \) lead to \( \sup_{k \in \mathbb{N}} \{ A_k A_k^T \} < \infty \). We consider the necessity of (1). If \( \sup_{k \in \mathbb{N}} \{ A_k A_k^T \} < \infty \), it follows from (A.1) that \( \sup_{k \in \mathbb{N}} \{ \bar{A}_k \bar{A}_k^T + \bar{G}_k \bar{G}_k^T \} < \infty \). Due to \( \bar{G}_k \bar{G}_k^T \geq 0 \), then \( \sup_{k \in \mathbb{N}} \{ A_k \bar{A}_k^T \} < \infty \). Next, we consider the proof of (2). According to (A.1), it holds that

\[
\lambda_{\min}(A_k A_k^T) = \left( \begin{array}{cc} \frac{1}{\bar{A}_k} & \frac{\bar{G}_k \bar{G}_k^T}{\bar{A}_k} \\ \bar{G}_k^T I_p & \end{array} \right). \tag{B.1}
\]

Considering that the estimation error \( e_{k-1,i} \) is measurable to \( \sigma(F_{k-2} \cup v_{k-1,i}, i \in V) \), and \( \omega_{k-1} \) is independent from \( F_{k-2} \) and \( v_{k-1,i}, i \in V \), we have \( E(e_{k-1,i} \omega_{k-1}^T) = 0 \). Similarly, it holds that \( E(\hat{\mu}_k, x_{k} x_{k}^T + \frac{1}{\theta_{k,i}} y y^T) \), \( \forall \theta_{k,i} > 0 \), we have

\[
E(e_{k,i} e_{k,i}^T) = A_k A_k^T \tag{B.2}
\]

According to (B.1) and the definition of \( \tilde{P}_{k,i} \), we have \( E(e_{k,i} e_{k,i}^T) \preceq \tilde{P}_{k,i} \). At the update stage, there is \( \tilde{e}_{k,i} = X_{k,i} - X_k = (I - K_{k,i} H_{k,i}) e_{k,i} + K_{k,i} b_{k,i} + K_{k,i} v_{k,i} \). Under Assumption 3.1 and the fact that \( \tilde{e}_{k,i} \) is measurable to \( F_{k-1} \), it follows that

\[
E(\tilde{e}_{k,i} v_{k,i}^T) = E\left( E(\tilde{e}_{k,i} v_{k,i}^T | F_{k-1}) \right) = E\left( (A_{k-1} e_{k-1,i} + w_{k-1}) v_{k,i}^T | F_{k-1} \right) \tag{B.3}
\]

Similarly, it holds that \( E(\tilde{e}_{k,i} v_{k,i}^T) = 0 \) by noting that \( b_{k,i} \) is measurable to \( F_{k-1} \). Then, given the inequality \( E(x y^T + y x^T) \leq E(\mu_k x x^T + \frac{1}{\mu_k} y y^T), \forall \mu_k > 0 \), we have

\[
E(\tilde{e}_{k,i} \hat{e}_{k,i}^T) = (I - K_{k,i} H_{k,i}) E(e_{k,i} e_{k,i}^T (I - K_{k,i} H_{k,i})^T + K_{k,i} E(v_{k,i} v_{k,i}^T) K_{k,i}^T + 1 + \frac{1}{\mu_k} K_{k,i} E(b_{k,i} b_{k,i}^T) K_{k,i}^T \leq (1 + \mu_k) (I - K_{k,i} H_{k,i}) E(e_{k,i} e_{k,i}^T) (I - K_{k,i} H_{k,i})^T + K_{k,i} E(v_{k,i} v_{k,i}^T) K_{k,i}^T \leq (1 + \mu_k) (I - K_{k,i} H_{k,i}) \hat{P}_{k,i} (I - K_{k,i} H_{k,i})^T + K_{k,i} \left( R_{k,i} + \frac{1}{\mu_k} b_{k,i} B_{k,i} \right) K_{k,i} \hat{P}_{k,i}.
\]
Notice that $e_{k,i} = \hat{X}_{k,i} - X_k = P_{k,i} \sum_{j \in N_i(k)} a_{i,j}(k) \hat{P}_{k,j}^{-1} \hat{e}_{k,j}$.

According to the consistent estimate of Covariance Intersection [44], there is $E\{e_{k,i} e_{k,i}^T\} \leq P_{k,i}$. Therefore, the proof is finished.

C Proof of Lemma 4.2

Consider (B.3), then we have

$$\hat{P}_{k,i} = (1 + \mu_{k,i}) \left( (I - K_{k,i} H_{k,i}) \hat{P}_{k,i} (I - K_{k,i} H_{k,i})^T + K_{k,i} \left( \frac{R_{k,i}}{1 + \mu_{k,i}} + \frac{1}{\mu_{k,i}} B_{k,i} K_{k,i}^T \right) \right).$$  \tag{C.1}

Denote $\tilde{R}_{k,i} = \frac{R_{k,i}}{1 + \mu_{k,i}} + \frac{1}{\mu_{k,i}} B_{k,i} K_{k,i}^T$, then we have

$$\Delta \tilde{P}_{k,i} = (I - K_{k,i} H_{k,i}) \hat{P}_{k,i} (I - K_{k,i} H_{k,i})^T + K_{k,i} \tilde{R}_{k,i} K_{k,i}^T,$$

$$= \hat{P}_{k,i} - K_{k,i} H_{k,i} \hat{P}_{k,i} + \hat{P}_{k,i} H_{k,i}^T K_{k,i}^T + K_{k,i} \tilde{R}_{k,i} K_{k,i}^T,$$

$$= \hat{P}_{k,i} - K_{k,i} H_{k,i} \hat{P}_{k,i} + \hat{P}_{k,i} H_{k,i}^T K_{k,i}^T + K_{k,i} \tilde{R}_{k,i} K_{k,i}^T,$$

$$= (K_{k,i} - K_{k,i}^*) (H_{k,i} \hat{P}_{k,i} H_{k,i}^T + \tilde{R}_{k,i}) (K_{k,i} - K_{k,i}^*)^T + (I - K_{k,i} H_{k,i}) \hat{P}_{k,i}.$$

Therefore, the $K_{k,i}^*$ is minimized in the sense of positive definiteness (i.e., $tr(\hat{P}_{k,i})$ is minimized) when $K_{k,i} = K_{k,i}^*$.

D Proof of Lemma 4.3

For the proof of 1), exploiting the matrix inverse formula on $\hat{P}_{k,i}$ directly yields the conclusion. Next, we consider the proof of 2). First, we prove there exists a constant positive definite matrix $\bar{P}$, such that $P_{k,i} \geq \bar{P}$. Consider

$$\hat{P}_{k,i} = (1 + \theta_{k,i}) A_{k-1} \hat{P}_{k-1,i} A_{k-1}^T + \frac{1}{\theta_{k,i}} \tilde{Q}_{k-1} + \tilde{Q}_{k-1} \geq \frac{1 + \theta_{k,i}}{\theta_{k,i}} \tilde{Q}_{k-1} + \tilde{Q}_{k-1} \geq 1 + \frac{1}{\theta_{k,i}} \tilde{Q}_{k-1} + \tilde{Q}_{k-1} \geq Q^* > 0,$$

where $Q^*$ can be obtained by noting $\inf_i Q_k \geq Q > 0$ and Assumption 3.3. According to 1) of Lemma 4.3, then we have

$$\hat{P}_{k,i} \geq \frac{Q^*}{1 + \mu_{k,i}} + H_{k,i}^T \Delta R_{k,i}^{-1} H_{k,i},$$

$$\leq \frac{Q^*}{1 + \mu_{k,i}} + H_{k,i}^T R_{k,i}^{-1} H_{k,i} \leq Q_*, \quad \tag{D.1}$$

where $Q_*$ is obtained by employing the condition 3) of Assumption 3.1. Recall $P_{k,i} = \left( \sum_{j \in N_i(k)} a_{i,j}(k) \hat{P}_{k,j}^{-1} \right)^{-1}$, then employing (D.1) leads to $P_{k,i} \leq Q_*$. Thus, $P_{k,i} \geq Q^{-1} > 0$. Consider the time sequence $\{T_i, i \in N\}$, which is the L-SS of $\{A_k, k \in N\}$. Under Assumption 3.2, there exists a scalar $\bar{\beta}$, such that $A_{T_i+s} P_{T_i+s,i} A_{T_i+s}^T \geq \beta Q^{-1} > 0$. Due to $\sup_i Q_k \leq \bar{Q} < \infty$, there is a scalar $\varpi > 0$, such that $Q_{T_i+s} \leq \varpi A_{T_i+s} P_{T_i+s,i} A_{T_i+s}^T$. Then

$$\hat{P}_{T_i+s,1,i} = A_{T_i+s} P_{T_i+s,i} A_{T_i+s}^T + Q_{T_i+s} \leq (1 + \varpi) A_{T_i+s} P_{T_i+s,i} A_{T_i+s}^T.$$

Let $\eta = \frac{1}{1 + \varpi}$, then the conclusion 2) of this lemma holds.

References

[1] W. Yang, G. Chen, X. Wang, and L. Shi, “Stochastic sensor activation for distributed state estimation over a sensor network,” Automatika, vol. 50, no. 8, pp. 2070–2076, 2014.

[2] X. Cao, J. Chen, Z. Yan, and Y. Sun, “Development of an integrated wireless sensor network micro-environmental monitoring system,” ISA Transactions, vol. 47, no. 3, pp. 247–255, 2008.

[3] X. Wang, M. Fu, and H. Zhang, “Target tracking in wireless sensor networks based on the combination of KF and MLE using distance measurements,” IEEE Transactions on Mobile Computing, vol. 11, no. 4, pp. 567–576, 2012.

[4] J. Y. Yu, M. J. Coates, M. G. Rabbat, and S. Blouin, “A distributed particle filter for bearings-only tracking on spherical surfaces,” IEEE Signal Processing Letters, vol. 23, no. 3, pp. 326–330, 2016.

[5] C. Dou and B. Liu, “Multi-agent based hierarchical hybrid control for smart microgrid,” IEEE Transactions on Smart Grid, vol. 4, no. 2, pp. 771–778, 2013.

[6] G. Battistelli and L. Chisci, “Stability of consensus extended Kalman filter for distributed state estimation,” Automatica, vol. 68, pp. 169–178, 2016.

[7] R. Olaf-Saber, “Distributed Kalman filtering for sensor networks,” in Proceedings of the IEEE Conference on Decision and Control, pp. 5492–5498, 2007.

[8] F. S. Cattivelli and A. H. Sayed, “Diffusion strategies for distributed Kalman filtering and smoothing,” IEEE Transactions on Automatic Control, vol. 55, no. 9, pp. 2069–2084, 2010.

[9] G. Battistelli and L. Chisci, “Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability,” Automatica, vol. 50, no. 3, pp. 707–718, 2014.

[10] X. He, W. Xue, and H. Fang, “Consistent distributed state estimation with global observability over sensor network,” Automatica, vol. 92, pp. 162 – 172, 2018.
G. Calafiore, “Reliable localization using set-valued nonlinear filters,” IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, vol. 35, no. 2, pp. 189–197, 2005.

W. Bai, W. Xue, Y. Huang, and H. Fang, “On extended state based Kalman filter design for a class of nonlinear time-varying uncertain systems,” Science China Information Sciences, vol. 61, no. 4, p. 042201, 2018.

L. Orehuela, P. Millán, C. Vivas, and F. R. Rubio, “Suboptimal distributed control and estimation: application to a four coupled tanks system,” International Journal of Systems Science, vol. 47, no. 8, pp. 1755–1771, 2016.

S. J. Julier and J. K. Uhlmann, “A non-divergent estimation algorithm in the presence of unknown correlations,” in American Control Conference, pp. 2369–2373, 1997.

Y. S. Chow and H. Teicher, Probability theory: independence, interchangeability, martingales. Springer Science & Business Media, 2012.

S. Aranovskii, A. A. Bobtsov, and A. Pyrkın, “Adaptive observer of an unknown sinusoidal output disturbance for linear plants,” Automation and Remote Control, vol. 70, no. 11, pp. 1862–1870, 2009.

Y. Liu, Z. Wang, X. He, and D. Zhou, “Minimum-variance recursive filtering over sensor networks with stochastic sensor gain degradation: Algorithms and performance analysis,” IEEE Transactions on Control of Network Systems, vol. 3, no. 3, pp. 265–274, 2016.

W. Li, Y. Jia, and J. Du, “Distributed filtering for discrete-time linear systems with fading measurements and time-correlated noise,” Digital Signal Processing, vol. 60, pp. 211–219, 2017.

C. Wen, Z. Wang, Q. Liu, and F. E. Alsaadi, “Recursive distributed filtering for a class of state-saturated systems with fading measurements and quantization effects,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2016.

Z. Cai, M. S. D. Queiroz, and D. M. Dawson, “Robust adaptive asymptotic tracking of nonlinear systems with additive disturbance,” IEEE Transactions on Automatic Control, vol. 51, no. 3, pp. 524–529, 2006.

X. He, C. Hu, W. Xue, and H. Fang, “On event-based distributed Kalman filter with information matrix triggers,” in IFAC World Congress, pp. 14873–14878, 2017.

J. Le Ny and G. J. Pappas, “Differentially private filtering,” IEEE Transactions on Automatic Control, vol. 59, no. 2, pp. 341–354, 2014.

M. Bartholomew-Biggs, Nonlinear optimization with engineering applications. Springer Science & Business Media, 2008.

D. Han, K. You, L. Xie, J. Wu, and L. Shi, “Optimal parameter estimation under controlled communication over sensor networks,” IEEE Transaction on Signal Processing, vol. 63, no. 24, pp. 6473–6485, 2015.

S. Weerakkody, Y. Mo, B. Sinopoli, D. Han, and L. Shi, “Multi-sensor scheduling for state estimation with event-based, stochastic triggers,” IEEE Transactions on Automatic Control, vol. 61, no. 9, pp. 2695–2701, 2016.

D. Simon, “Kalman filtering with state constraints: a survey of linear and nonlinear algorithms,” IET Control Theory & Applications, vol. 4, no. 8, pp. 1303–1318, 2010.

W. Niehsen, “Information fusion based on fast covariance intersection filtering,” in International Conference on Information Fusion, pp. 901–904, 2002.