SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

July 2021, Vol. 64 172204:1–172204:14
https://doi.org/10.1007/s11432-020-2962-9

Stability of the distributed Kalman filter using general random coefficients

Die GAN[1,2], Siyu XIE[3] & Zhixin LIU[1,2*]

[1] *Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing* 100190*, China;*
[2] *School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing* 100190*, China;*
[3] *The Department of Electrical and Computer Engineering, Wayne State University, Detroit* 48202*, USA*

Received 11 January 2020/Revised 4 April 2020/Accepted 29 May 2020/Published online 18 May 2021

Abstract In this paper, we propose a distributed Kalman filter (DKF) for the dynamical system with general random coefficients. In the proposed method, each estimator shares local innovation pairs with its neighbors to collectively complete the estimation task. Further, we introduce a collective random observability condition by which the L_p-stability of the covariance matrix and the L_p-exponential stability of the homogeneous part of the estimation error equation can be established. In contrast, the stringent conditions on the coefficient matrices, such as independency and stationarity are not required. Besides, the stability of the DKF, i.e., the boundedness of the filtering errors, can be established. Finally, from the simulation result, we demonstrate the cooperative effect of the sensors.

Keywords distributed Kalman filter, collective random observability, L_p-stable, L_p-exponentially stable, state estimation

Citation Gan D, Xie S Y, Liu Z X. Stability of the distributed Kalman filter using general random coefficients. Sci China Inf Sci, 2021, 64(7): 172204, https://doi.org/10.1007/s11432-020-2962-9

1 Introduction

The state estimation plays an important role in the areas of signal processing and control engineering [1]. The Kalman filter (KF) is one of the most widely used recursive estimation algorithms. It is well-known that for the discrete-time linear dynamical systems, the KF is an optimal state estimate in the sense of the minimum mean square error when the noises have Gaussian distribution [2, 3]. The KF and its variants have many applications in practical systems, such as guidance and navigation of vehicles.

Over the last decade, the sensor networks have an increasing development, meaning that we can acquire more data. However, using these data to improve the observability brings great challenges both in the theoretical and practical aspects. The centralized and distributed frameworks are two main approaches involved in processing the data collected from the sensors. For the centralized method, there is a fusion center, which can receive and merge the data from all the sensors at each time instant, such as measurements, observation matrices, and state estimates [4, 5]. Although the centralized method can realize an optimum estimate, it is vulnerable to a connection failure, delay, and packet loss. The centralized method lacks robustness in addition to bringing amounts of computational overhead and communication load among the nodes and the processor. Therefore, the approach to estimating the state in a distributed way arises [6–15], where each sensor only utilizes the local information and the data from its neighboring sensors to estimate the unknown states. The distributed state estimation has applications in some practical systems, such as surveillance and tracking systems [16, 17].

The distributed Kalman filter (DKF) is one of the important algorithms to collectively estimate the states of the dynamical systems. Much effort has been devoted to the investigation of the DKF where the coefficient matrices of the systems are deterministic. For example, Battistelli and Chisci [12] studied

* Corresponding author (email: lzx@amss.ac.cn)

the upper boundedness of the error covariance matrix of the DKF for the systems with time-invariant coefficients. He et al. [18] established the boundedness of the error covariance matrix and the exponentially asymptotic unbiasedness of the state estimate for time-varying coefficient matrices. Further, Wang and Ren [19] investigated the convergence of a distributed hybrid information fusion algorithm based on optimized weights of time-varying topology and coefficient matrices. The stability analysis of continuous Kalman-consensus filtering algorithm on a mobile sensor network with a flocking-based mobility control model was studied in [20], where the coefficient matrices of the dynamical system are considered time-varying. Some results of the special case with intermittent observations are also obtained. Yang et al. [21] proposed a random sensor activation scheme for a consensus-based distributed estimation algorithm. Also, a distributed Kalman filtering algorithm of a linear time-invariant discrete-time system using data packet drops was proposed in [13]. From the current literature, very few results are obtained relating to the DKF for the dynamical systems with general random coefficients.

In this paper, the stability of the DKF for the linear dynamical systems with general random coefficient matrices is considered. We note that the main challenge arising in the theoretical investigation of the DKF lies in analyzing the properties of the product of the random matrices. Most of the existing studies on the theoretical analysis of the adaptive filtering algorithms require that the signals satisfy some stringent conditions, such as independency and statistical stationarity [22, 23]. This makes it hard or even impossible to apply these theoretical results to practical feedback control systems. Further, we introduce a collective random observation condition, which can be regarded as the extension of the random observation condition proposed in [24] using a distributed case. Hence, some good properties of the product on random matrices can be obtained with the stability of the DKF established. The main contributions of this paper are summarized as follows.

- First, the DKF algorithm for the dynamical systems with general random coefficients was proposed, where each sensor diffuses the local innovation pairs $(\boldsymbol{H}_{k,i}^{\mathrm{T}}\boldsymbol{R}_{k,i}^{-1}\boldsymbol{H}_{k,i}, \boldsymbol{H}_{k,i}^{\mathrm{T}}\boldsymbol{R}_{k,i}^{-1}\boldsymbol{y}_{k,i})$ with its neighbors. We know that most engineering systems are nonlinear and disturbed by noise, and the extended Kalman filter (EKF) is often used to estimate the states of the system. The algorithm and the results in this paper may give some clues for the effectiveness of the EKF.
- We introduce a collective random observability condition, in which the L_p-stability of the covariance matrix can be established by relying on an auxiliary system. Then, the L_p-exponential stability of the homogeneous part of the estimation error equation can be obtained, which helps to establish the stability of the DKF further.
- Compared with the current literature, the stability of DKF can be obtained without relying on the assumptions of the independency and statistical stationarity of the regression signals, which makes it possible for applications to the stochastic feedback systems. We observe that by the collective random observability condition, the estimation task can still be fulfilled using the cooperation of multiple sensors even if any of them cannot do it individually.

The remainder of this paper is organized as follows. We first introduce some preliminaries and propose the DKF with random coefficients in Section 2. In Section 3, we present the collective random observability condition. The main results, including L_p-stability of the covariance matrix, L_p-exponential stability of the estimation error equation, and the stability of DKF are also presented. We illustrate the cooperative effect of the sensors by a simulation example in Section 4. Section 5 presents the conclusion of the paper.

2 Problem formulation

2.1 Some preliminaries

In this paper, we use $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ to denote an $m\times n$-dimensional real matrix. For a matrix $\boldsymbol{A}$, $\|\boldsymbol{A}\|$ denotes the Euclidean norm, i.e., $\|\boldsymbol{A}\| = (\lambda_{\max}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}))^{\frac{1}{2}}$, where the notation T denotes the transpose operator and $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of the matrix. Correspondingly, we use $\lambda_{\min}(\cdot)$ to denote the smallest eigenvalue of the matrix. Moreover, $\mathrm{E}[\cdot]$, $P[\cdot|\cdot]$ and $\mathrm{E}[\cdot|\cdot]$ denote the expectation, the conditional probability and the conditional expectation operator, respectively. We define $\|\boldsymbol{A}\|_{L_p} \triangleq (\mathrm{E}[\|\boldsymbol{A}\|^p])^{\frac{1}{p}}$ as the L_p-norm of the random matrix $\boldsymbol{A}$. Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n\times n}$ be two symmetric matrices, and then $\boldsymbol{A} \geqslant \boldsymbol{B}$ means that $\boldsymbol{A}-\boldsymbol{B}$ is a positive semi-definite matrix. If all elements of a matrix $\boldsymbol{A} = \{a_{ij}\} \in \mathbb{R}^{n\times n}$ are nonnegative, then it is a nonnegative matrix, and furthermore if $\sum_{j=1}^{n} a_{ij} = 1$ holds for all $i = 1$, then it is called a stochastic matrix. The Kronecker product of two matrices $\boldsymbol{A} = \{a_{ij}\} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{p\times q}$ is defined as

$$\boldsymbol{A}\otimes\boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{pmatrix} \in \mathbb{R}^{mp\times nq}.$$

In this paper, our purpose is to propose a distributed Kalman filtering algorithm based on the information from neighboring sensors and investigate the stability of the proposed algorithm. In order to describe the relationship between sensors, an undirected weighted graph $\mathscr{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ is introduced here, where $\mathcal{V} = \{1, 2, 3, \ldots, n\}$ is the set of sensors, $\mathcal{E} \in \mathcal{V}\times\mathcal{V}$ is the edge set which is used to describe the communication between sensors, and $\mathcal{A} = \{a_{ij}\} \in \mathbb{R}^{n\times n}$ is the weighted adjacency matrix. The elements of the matrix $\mathcal{A}$ satisfy $a_{ij} > 0$ if and only if $(i, j) \in \mathcal{E}$. The set of the neighbors of sensor i is denoted by $N^i = \{j \in \mathcal{V}|(i, j) \in \mathcal{E}\}$, and we assume that agent i also belongs to N^i, and each sensor can only exchange information with its neighbors. A path of length ℓ is a sequence of nodes $\{i_1, \ldots, i_\ell\}$ satisfying $(i_j, i_{j+1}) \in \mathcal{E}$ for all $1 \leqslant j \leqslant \ell - 1$. The graph $\mathscr{G}$ is called connected if for any two sensors i and j, there is a path connecting them. The diameter $D(\mathscr{G})$ of the graph $\mathscr{G}$ is defined as the maximum length of the path between any two sensors. For simplicity of analysis, the stability of the distributed algorithm is considered under the condition that the weighted adjacency matrix $\mathcal{A}$ is symmetric and stochastic. Thus, it is obvious that $\mathcal{A}$ is doubly stochastic.

In order to proceed with our discussion, we need to introduce some definitions similar to those used in [25].

Definition 1. A random matrix sequence $\{\boldsymbol{A}_k, k \geqslant 0\}$ defined on the basic probability space $(\Omega, \mathscr{F}, P)$ is called L_p-stable $(p > 0)$ if

$$\sup_{k\geqslant 0} \mathrm{E}[\|\boldsymbol{A}_k\|^p] < \infty.$$

Definition 2. A sequence of $n\times n$ random matrices $\boldsymbol{A} = \{\boldsymbol{A}_k, k \geqslant 0\}$ is called L_p-exponentially stable $(p \geqslant 0)$ with parameter $\lambda \in [0, 1)$, if it belongs to the following set:

$$S_p(\lambda) = \left\{ \boldsymbol{A} : \left\| \prod_{i=j+1}^{k} \boldsymbol{A}_i \right\|_{L_p} \leqslant M\lambda^{k-j}, \forall k \geqslant j, \forall j \geqslant 0, \text{for some } M > 0 \right\}.$$

For convenience of discussion, we introduce the following subclass of $S_1(\lambda)$ for a scalar sequence $a = \{a_k, k \geqslant 0\}$:

$$S^0(\lambda) = \left\{ a : a_k \in [0, 1), \mathrm{E}\left[\prod_{i=j+1}^{k} a_i\right] \leqslant M\lambda^{k-j}, \forall k \geqslant j, \forall j \geqslant 0, \text{for some } M > 0 \right\}.$$

Remark 1. It is clear that if there exists a constant $a_0 \in [0, 1)$ such that $a_k \leqslant a_0$, then $\{a_k\} \in S^0(a_0)$. More properties about the set $S^0(\lambda)$ can be found in [26].

Definition 3. Let $\{\boldsymbol{A}_k\}$ be a matrix sequence and $\{b_k\}$ be a positive scalar sequence. Then by $\boldsymbol{A}_k = O(b_k)$ we mean that there exists a constant $M > 0$ such that

$$\|\boldsymbol{A}_k\| \leqslant Mb_k, \quad \forall k \geqslant 0.$$

2.2 Distributed Kalman filter

In this paper, we consider the following discrete-time linear dynamical system:

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{F}_k\boldsymbol{x}_k + \boldsymbol{w}_{k+1}, \\ \boldsymbol{y}_{k,i} = \boldsymbol{H}_{k,i}\boldsymbol{x}_k + \boldsymbol{v}_{k,i}, \end{cases} \tag{1}$$

where $\boldsymbol{x}_k$ is the s-dimensional state vector to be estimated and $\boldsymbol{F}_k \in \mathbb{R}^{s\times s}$ is the random state evolution matrix, $\boldsymbol{y}_{k,i} \in \mathbb{R}^d$ and $\boldsymbol{H}_{k,i} \in \mathbb{R}^{d\times s}$ represent the measurement and the random measurement matrices for agent i at time instant k respectively, $\{\boldsymbol{w}_k \in \mathbb{R}^s, \boldsymbol{v}_{k,i} \in \mathbb{R}^d, k \geqslant 0\}$ is an independent noise process which satisfies the following conditions:

$$\begin{aligned} &\mathrm{E}[\boldsymbol{w}_k] = 0, \quad \mathrm{E}[\boldsymbol{w}_k \boldsymbol{w}_k^{\mathrm{T}}] = \boldsymbol{Q}_k \geqslant \boldsymbol{Q} > 0; \\ &\mathrm{E}[\boldsymbol{v}_{k,i}] = 0, \quad \mathrm{E}[\boldsymbol{v}_{k,i} \boldsymbol{v}_{k,i}^{\mathrm{T}}] = \boldsymbol{R}_{k,i} \geqslant a_i \boldsymbol{I} > 0, \ \mathrm{E}[\boldsymbol{v}_{k,i} \boldsymbol{w}_k^{\mathrm{T}}] = 0. \end{aligned}$$

Remark 2. The investigation of the linear system with random coefficients has practical significance. We know that most engineering systems are nonlinear which can be written as the following general expression:

$$\begin{cases} x_{k+1} = f(x_k) + w_{k+1}, \\ y_k = g(x_k) + v_k, \end{cases}$$

where $f(\cdot)$ and $g(\cdot)$ are nonlinear functions, and $\{w_k\}$ and $\{v_k\}$ are two noise sequences. In order to estimate the state of the above system, the EKF is widely used based on the following linearization:

$$\begin{cases} x_{k+1} = f(\hat{x}_k) + \frac{\partial f(x)}{\partial x}|_{x_k=\hat{x}_k}(x_k - \hat{x}_k) + w_{k+1}, \\ y_k = g(\hat{x}_k) + \frac{\partial g(x)}{\partial x}|_{x_k=\hat{x}_k}(x_k - \hat{x}_k) + v_k, \end{cases}$$

where $\hat{x}_k$ denotes the estimate of the state at time k. It is clear that $\frac{\partial f(x)}{\partial x}|_{x_k=\hat{x}_k}$ and $\frac{\partial g(x)}{\partial x}|_{x_k=\hat{x}_k}$ are two random matrices.

We assume that the initial state $\boldsymbol{x}_0$ is a random vector with mean $\hat{\boldsymbol{x}}_0$ and covariance matrix $\boldsymbol{P}_0' > 0$, and is independent of the sequence $\{\boldsymbol{w}_k, \boldsymbol{v}_{k,i}, k \geqslant 0\}$. We first present the standard non-cooperative KF [27] in Algorithm 1, where $\hat{\boldsymbol{x}}_{k,i}'$ and $\hat{\boldsymbol{x}}_{k,i}$ denote a prior and a posteriori estimates of $\boldsymbol{x}_k$ for node i at time instant k.

Algorithm 1 Standard non-cooperative Kalman filter

For any given sensor $i \in \{1, \ldots, n\}$, start with an initial estimate $\hat{\boldsymbol{x}}_{0,i}' \in \mathbb{R}^s$ and an initial covariance matrix $\boldsymbol{P}_{0,i}' > 0 \in \mathbb{R}^{s\times s}$. The standard KF is given as follows.

(Measurement update process)

$$\boldsymbol{P}_{k,i}^{-1} = \boldsymbol{P}_{k,i}'^{-1} + \boldsymbol{H}_{k,i}^{\mathrm{T}} \boldsymbol{R}_{k,i}^{-1} \boldsymbol{H}_{k,i}, \tag{2}$$

$$\hat{\boldsymbol{x}}_{k,i} = \hat{\boldsymbol{x}}_{k,i}' + \boldsymbol{K}_{k,i}(\boldsymbol{y}_{k,i} - \boldsymbol{H}_{k,i}\hat{\boldsymbol{x}}_{k,i}'), \tag{3}$$

$$\boldsymbol{K}_{k,i} = \boldsymbol{P}_{k,i}' \boldsymbol{H}_{k,i}^{\mathrm{T}} (\boldsymbol{H}_{k,i} \boldsymbol{P}_{k,i}' \boldsymbol{H}_{k,i}^{\mathrm{T}} + \boldsymbol{R}_{k,i})^{-1}, \tag{4}$$

(State prediction process)

$$\hat{\boldsymbol{x}}_{k+1,i}' = \boldsymbol{F}_k \hat{\boldsymbol{x}}_{k,i}, \tag{5}$$

$$\boldsymbol{P}_{k+1,i}' = \boldsymbol{F}_k \boldsymbol{P}_{k,i} \boldsymbol{F}_k^{\mathrm{T}} + \boldsymbol{Q}_k. \tag{6}$$

The matrix inversion formula is used in our analysis, and we list it here.

Lemma 1 ([26])**.** For any matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ with suitable dimensions, the following formula

$$(\boldsymbol{A} + \boldsymbol{BDC})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{D}^{-1} + \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}\boldsymbol{A}^{-1}$$

holds, provided that the relevant matrices are invertible.

In order to introduce the DKF, a preliminary result on the standard non-cooperative KF given in Algorithm 1 is presented here.

Lemma 2. Eq. (3) can be rewritten as the following form:

$$\boldsymbol{P}_{k,i}^{-1} \hat{\boldsymbol{x}}_{k,i} = \boldsymbol{P}_{k,i}'^{-1} \hat{\boldsymbol{x}}_{k,i}' + \boldsymbol{H}_{k,i}^{\mathrm{T}} \boldsymbol{R}_{k,i}^{-1} \boldsymbol{y}_{k,i}. \tag{7}$$

Proof. By (3) and (4), we have

$$\begin{aligned} \hat{\boldsymbol{x}}_{k,i} &= \hat{\boldsymbol{x}}_{k,i}' + \boldsymbol{P}_{k,i}' \boldsymbol{H}_{k,i}^{\mathrm{T}} (\boldsymbol{H}_{k,i} \boldsymbol{P}_{k,i}' \boldsymbol{H}_{k,i}^{\mathrm{T}} + \boldsymbol{R}_{k,i})^{-1} (\boldsymbol{y}_{k,i} - \boldsymbol{H}_{k,i} \hat{\boldsymbol{x}}_{k,i}') \\ &= \hat{\boldsymbol{x}}_{k,i}' - \boldsymbol{P}_{k,i}' \boldsymbol{H}_{k,i}^{\mathrm{T}} (\boldsymbol{H}_{k,i} \boldsymbol{P}_{k,i}' \boldsymbol{H}_{k,i}^{\mathrm{T}} + \boldsymbol{R}_{k,i})^{-1} \boldsymbol{H}_{k,i} \hat{\boldsymbol{x}}_{k,i}' \end{aligned}$$

$$\begin{aligned} &+\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}(\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{R}_{k,i})^{-1}\boldsymbol{y}_{k,i}\\ =&\left[\boldsymbol{P}'_{k,i}-\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}(\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{R}_{k,i})^{-1}\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\right]\boldsymbol{P}'^{-1}_{k,i}\hat{\boldsymbol{x}}'_{k,i}\\ &+\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}(\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{R}_{k,i})^{-1}\boldsymbol{y}_{k,i}. \end{aligned}$$

Then by the matrix inversion formula and (2), we have

$$\begin{aligned} \hat{\boldsymbol{x}}_{k,i}=&(\boldsymbol{P}'^{-1}_{k,i}+\boldsymbol{H}^{\mathrm{T}}_{k,i}\boldsymbol{R}^{-1}_{k,i}\boldsymbol{H}_{k,i})^{-1}\boldsymbol{P}'^{-1}_{k,i}\hat{\boldsymbol{x}}'_{k,i}\\ &+\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}(\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{R}_{k,i})^{-1}\boldsymbol{y}_{k,i}\\ =&\boldsymbol{P}_{k,i}\boldsymbol{P}'^{-1}_{k,i}\hat{\boldsymbol{x}}'_{k,i}+\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}(\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{R}_{k,i})^{-1}\boldsymbol{y}_{k,i}. \end{aligned} \tag{8}$$

Notice that

$$\begin{aligned} &\boldsymbol{P}^{-1}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}(\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{R}_{k,i})^{-1}\\ &=(\boldsymbol{P}'^{-1}_{k,i}+\boldsymbol{H}^{\mathrm{T}}_{k,i}\boldsymbol{R}^{-1}_{k,i}\boldsymbol{H}_{k,i})\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}(\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{R}_{k,i})^{-1}\\ &=(\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{H}^{\mathrm{T}}_{k,i}\boldsymbol{R}^{-1}_{k,i}\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i})(\boldsymbol{H}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{R}_{k,i})^{-1}\\ &=\boldsymbol{H}^{\mathrm{T}}_{k,i}\boldsymbol{R}^{-1}_{k,i}. \end{aligned}$$

Then by (8) we have

$$\begin{aligned} \hat{\boldsymbol{x}}_{k,i}&=\boldsymbol{P}_{k,i}\boldsymbol{P}'^{-1}_{k,i}\hat{\boldsymbol{x}}'_{k,i}+\boldsymbol{P}_{k,i}\boldsymbol{P}^{-1}_{k,i}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}(\boldsymbol{H}_{k}\boldsymbol{P}'_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}+\boldsymbol{R}_{k,i})^{-1}\boldsymbol{y}_{k,i}\\ &=\boldsymbol{P}_{k,i}\boldsymbol{P}'^{-1}_{k,i}\hat{\boldsymbol{x}}'_{k,i}+\boldsymbol{P}_{k,i}\boldsymbol{H}^{\mathrm{T}}_{k,i}\boldsymbol{R}^{-1}_{k,i}\boldsymbol{y}_{k,i}, \end{aligned}$$

which completes the proof of the lemma.

Inspired by the expressions of (2) and (7), we propose a DKF for the dynamical system (1) with random coefficients in Algorithm 2. In order to avoid confusion, the notations $\hat{\boldsymbol{x}}_{k,i}$ and $\hat{\boldsymbol{x}}'_{k,i}$ represent a prior and a posteriori estimates of $\boldsymbol{x}_k$ for node i at time instant k with the corresponding error covariances $\boldsymbol{P}_{k,i}$ and $\boldsymbol{P}'_{k,i}$ of the DKF.

Algorithm 2 Distributed Kalman filter

For any given sensor $i\in\{1,\ldots,n\}$, start with an initial estimate $\hat{\boldsymbol{x}}'_{0,i}\in\mathbb{R}^s$ and an initial covariance matrix $\boldsymbol{P}'_{0,i}>0\in\mathbb{R}^{s\times s}$, the DKF is given as follows.

Step 1. Set the initial value at each time instant k as

$$\boldsymbol{\xi}_{k,i}(0)=\boldsymbol{H}^{\mathrm{T}}_{k,i}\boldsymbol{R}^{-1}_{k,i}\boldsymbol{H}_{k,i},\quad \boldsymbol{\eta}_{k,i}(0)=\boldsymbol{H}^{\mathrm{T}}_{k,i}\boldsymbol{R}^{-1}_{k,i}\boldsymbol{y}_{k,i}; \tag{9}$$

Step 2. Perform the following diffusion process for $l=0,1,2,\ldots,L$ with $L\geqslant D(\mathscr{G})$,

$$\boldsymbol{\xi}_{k,i}(l+1)=\sum_{j\in N^i}a_{ij}\boldsymbol{\xi}_{k,j}(l),\quad \boldsymbol{\eta}_{k,i}(l+1)=\sum_{j\in N^i}a_{ij}\boldsymbol{\eta}_{k,j}(l); \tag{10}$$

Step 3. Measurement update process

$$\boldsymbol{P}^{-1}_{k,i}=\boldsymbol{P}'^{-1}_{k,i}+\boldsymbol{\xi}_{k,i}(L),\quad \boldsymbol{P}^{-1}_{k,i}\hat{\boldsymbol{x}}_{k,i}=\boldsymbol{P}'^{-1}_{k,i}\hat{\boldsymbol{x}}'_{k,i}+\boldsymbol{\eta}_{k,i}(L); \tag{11}$$

Step 4. State prediction process

$$\hat{\boldsymbol{x}}'_{k+1,i}=\boldsymbol{F}_k\hat{\boldsymbol{x}}_{k,i}, \tag{12}$$

$$\boldsymbol{P}'_{k+1,i}=\boldsymbol{F}_k\boldsymbol{P}_{k,i}\boldsymbol{F}^{\mathrm{T}}_k+\boldsymbol{Q}_k. \tag{13}$$

Indeed, the DKF (Algorithm 2) proposed in this paper is obtained by employing the structure of the standard KF, saving that the innovation pairs $(\boldsymbol{H}^{\mathrm{T}}_{k,i}\boldsymbol{R}^{-1}_{k,i}\boldsymbol{H}_{k,i},\boldsymbol{H}^{\mathrm{T}}_{k,i}\boldsymbol{R}^{-1}_{k,i}\boldsymbol{y}_{k,i})$ are diffused with its neighbors. Such a diffusion strategy was used to design the distributed algorithm to estimate the state in [6, 7]. One can see that when the diffusion step L is large enough, the performance of DKF in Algorithm 2 is close to the centralized KF which is optimal in the sense of mean square error [28]. We remark that how to design the coefficients of $\xi_{k,i}(l)$ to improve the performance is an important topic in the investigation of DKF, and some investigations focus on the design of the coefficients for the systems where F_k and $H_{k,i}$ are deterministic (e.g., [29–31]). However, for the systems with random parameters,

Table 1 Some notations

Notation	Definition	Dimension
$\boldsymbol{P}_k$	$\mathrm{diag}\{\boldsymbol{P}_{k,1},\ldots,\boldsymbol{P}_{k,n}\}$	$sn\times sn$
$\boldsymbol{P}'_k$	$\mathrm{diag}\{\boldsymbol{P}'_{k,1},\ldots,\boldsymbol{P}'_{k,n}\}$	$sn\times sn$
$\bar{\boldsymbol{F}}_{\boldsymbol{k}}$	$\mathrm{diag}\{\boldsymbol{F}_k,\ldots,\boldsymbol{F}_k\}$	$sn\times sn$
$\bar{\boldsymbol{Q}}_{\boldsymbol{k}}$	$\mathrm{diag}\{\boldsymbol{Q}_k,\ldots,\boldsymbol{Q}_k\}$	$sn\times sn$
$\widetilde{\boldsymbol{X}}_k$	$\mathrm{col}\{\widetilde{\boldsymbol{x}}'_{k,1},\ldots,\widetilde{\boldsymbol{x}}'_{k,n}\}$	$sn\times 1$
$\mathscr{A}$	$\mathcal{A}\otimes\boldsymbol{I}_s$	$sn\times sn$
$\boldsymbol{H}_{\boldsymbol{k}}$	$\mathrm{diag}\{\boldsymbol{H}_{k,1},\ldots,\boldsymbol{H}_{k,n}\}$	$sn\times sn$
$\boldsymbol{R}_{\boldsymbol{k}}$	$\mathrm{diag}\{\boldsymbol{R}_{k,1},\ldots,\boldsymbol{R}_{k,n}\}$	$sn\times sn$
$\boldsymbol{W}_k$	$\mathrm{col}\{\boldsymbol{w}_k,\ldots,\boldsymbol{w}_k\}$	$sn\times 1$
$\boldsymbol{V}_k$	$\mathrm{col}\{\boldsymbol{v}_{k,1},\ldots,\boldsymbol{v}_{k,n}\}$	$dn\times 1$

it is hard to obtain an explicit solution of the coefficients of $\xi_{k,i}(l)$ since the objective function in the optimization problem is often taken as the mean square error, and thus the methods used in [29–31] may not be suitable to deal with our case. For simplicity of analysis, we choose the constant coefficients of $\xi_{k,i}(l)$ in this paper. The investigation of seeking the optimal filter gain falls into our future research.

From Steps 1–3 in Algorithm 2, we obtain

$$\boldsymbol{P}_{k,i}^{-1} = \boldsymbol{P}_{k,i}'^{-1} + \sum_{j=1}^{n} a_{ij}^{(L)} \boldsymbol{H}_{k,j}^{\mathrm{T}} \boldsymbol{R}_{k,j}^{-1} \boldsymbol{H}_{k,j}, \tag{14}$$

$$\boldsymbol{P}_{k,i}^{-1}\hat{\boldsymbol{x}}_{k,i} = \boldsymbol{P}_{k,i}'^{-1}\hat{\boldsymbol{x}}'_{k,i} + \sum_{j=1}^{n} a_{ij}^{(L)} \boldsymbol{H}_{k,j}^{\mathrm{T}} \boldsymbol{R}_{k,j}^{-1} \boldsymbol{y}_{k,j}, \tag{15}$$

where $a_{ij}^{(L)}$ is the i-th row, j-th column entry of $\mathcal{A}^L$. Hence by (14), we have

$$\boldsymbol{x}_k = \boldsymbol{P}_{k,i}\boldsymbol{P}_{k,i}^{-1}\boldsymbol{x}_k = \left(\boldsymbol{P}_{k,i}\boldsymbol{P}_{k,i}'^{-1} + \boldsymbol{P}_{k,i}\sum_{j=1}^{n} a_{ij}^{(L)} \boldsymbol{H}_{k,j}^{\mathrm{T}} \boldsymbol{R}_{k,j}^{-1} \boldsymbol{H}_{k,j}\right)\boldsymbol{x}_k. \tag{16}$$

Let $\widetilde{\boldsymbol{x}}'_{k,i} = \boldsymbol{x}_k - \hat{\boldsymbol{x}}'_{k,i}$. By (1) and (12), we have $\widetilde{\boldsymbol{x}}'_{k+1,i} = \boldsymbol{F}_k(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k,i}) + \boldsymbol{w}_{k+1}$. Combining this with (15) and (16), we can obtain the following error equation for the estimate:

$$\widetilde{\boldsymbol{x}}'_{k+1,i} = \boldsymbol{F}_k\boldsymbol{P}_{k,i}\boldsymbol{P}_{k,i}'^{-1}\widetilde{\boldsymbol{x}}'_{k,i} - \boldsymbol{F}_k\boldsymbol{P}_{k,i}\sum_{j=1}^{n} a_{ij}^{(L)} \boldsymbol{H}_{k,j}^{\mathrm{T}} \boldsymbol{R}_{k,j}^{-1} \boldsymbol{v}_{k,j} + \boldsymbol{w}_{k+1}. \tag{17}$$

For convenience of analysis, we introduce the following notations (see Table 1). In Table 1, $\mathrm{col}(\cdot,\ldots,\cdot)$ denotes a vector stacked by the specified vectors, and $\mathrm{diag}(\cdot,\ldots,\cdot)$ denotes a block matrix formed in a diagonal manner of the corresponding vectors or matrices.

By (17) and the notations in Table 1, we obtain the following matrix form of the estimation error:

$$\widetilde{\boldsymbol{X}}_{k+1} = \bar{\boldsymbol{F}}_k\boldsymbol{P}_k\boldsymbol{P}_k'^{-1}\widetilde{\boldsymbol{X}}_k - \bar{\boldsymbol{F}}_k\boldsymbol{P}_k\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-1}\boldsymbol{V}_k + \boldsymbol{W}_{k+1}, \tag{18}$$

and by (13), it is clear that

$$\boldsymbol{P}'_{k+1} = \bar{\boldsymbol{F}}_{\boldsymbol{k}}\boldsymbol{P}_k\bar{\boldsymbol{F}}_k^{\mathrm{T}} + \bar{\boldsymbol{Q}}_{\boldsymbol{k}}. \tag{19}$$

3 Stability of the distributed Kalman filter

3.1 Some assumptions

In order to proceed with our analysis, we introduce some assumptions concerning the graph, the state evolution matrix, and the measurement matrix.

Assumption 1. The graph $\mathcal{G}$ is connected.

Define the state transition matrix $\boldsymbol{\Phi}(k,j)$ as follows:

$$\boldsymbol{\Phi}(k,j)=\boldsymbol{F}_{k-1}\cdots\boldsymbol{F}_j,\quad \forall k\geqslant j+1;\quad \boldsymbol{\Phi}(j,j)=\boldsymbol{I}. \tag{20}$$

Assumption 2 (Collective random observability condition)**.** For any $\varepsilon>0$, there exists $\delta>0$ such that for any k,

$$P\{\lambda_{\min}(\boldsymbol{G}(k+h,k))>\delta|\mathscr{G}_{k-1}\}>1-\varepsilon,$$

where $h>0$ is an integer, and $\mathscr{G}_k=\sigma\{\boldsymbol{F}_j,\boldsymbol{H}_{j,i},j\leqslant k,1\leqslant i\leqslant n\}$, $\boldsymbol{G}(k+h,k)$ is the collective observability matrix, i.e.,

$$\boldsymbol{G}(k+h,k)=\sum_{i=1}^{n}\sum_{j=k+1}^{k+h}\boldsymbol{\Phi}^{\mathrm{T}}(j,k)\boldsymbol{H}_{j,i}^{\mathrm{T}}\boldsymbol{H}_{j,i}\boldsymbol{\Phi}(j,k).$$

Remark 3. Wang and Guo [24] proved the stability of the random Riccati equation under the following random observability condition.

For any $\varepsilon>0$, there exists $\delta>0$ such that

$$P\left\{\lambda_{\min}\left(\sum_{j=k+1}^{k+h_0}\boldsymbol{\Phi}^{\mathrm{T}}(j,k)\boldsymbol{H}_{j,i}^{\mathrm{T}}\boldsymbol{H}_{j,i}\boldsymbol{\Phi}(j,k)\right)>\delta\,\middle|\,\mathscr{F}_{k-1}\right\}>1-\varepsilon, \tag{21}$$

where h_0 is a positive integer. Assumption 2 is actually an extension of the condition (21) to the distributed case.

Remark 4. It is clear that Assumption 2 can be satisfied by the cooperation of all sensors, even if none of them satisfy the observation condition (21), which reveals that the estimation task can still be fulfilled by the cooperation of multiple sensors even if any of them cannot do it individually.

Assumption 3. For some $r\geqslant 1$, there exist positive constants M_1,M_2,M_3 and α such that

$$\begin{aligned}
&\text{(i)}\ \sup_k \mathrm{E}[\|\boldsymbol{H}_{k,i}\|^{32r}]<M_1<\infty,\ i\in\{1,\ldots,n\};\\
&\text{(ii)}\ \sup_{k\leqslant j\leqslant m\leqslant k+h}\mathrm{E}[\|\boldsymbol{\Phi}(m,j)\|^{32r+\alpha}]<M_2<\infty,\quad \forall k\geqslant 0;\\
&\text{(iii)}\ \sup_{k\geqslant 0}\mathrm{E}[\|\boldsymbol{\Phi}(k+h,k)\|^{16r}|\mathscr{G}_{k-1}]<M_3<\infty,
\end{aligned}$$

where $h>0$ is defined in Assumption 2.

Assumption 4. There exist constants N_1 and N_2 such that

$$\sup_{k\leqslant j\leqslant k+h}\mathrm{E}[\|\boldsymbol{H}_{j,i}\|^{8}|\mathscr{G}_{k-1}]<N_1<\infty,\ i\in\{1,\ldots,n\},\ \forall k,$$

and

$$\sup_{k\leqslant j\leqslant m\leqslant k+h}\mathrm{E}[\|\boldsymbol{\Phi}(m,j)\|^{8+\alpha}|\mathscr{G}_{k-1}]<N_2<\infty,\quad \forall k,$$

where $h>0$ is defined in Assumption 2 and α is defined in Assumption 3.

Remark 5. It is easy to see that Assumptions 3 and 4 are automatically satisfied if $\{\boldsymbol{F}_k,\boldsymbol{H}_{k,i}\}$ is a bounded sequence which is usually used in [12, 32].

To study the stability of the DKF, the assumptions on the initial value $\boldsymbol{x}_0$ and the noises $\boldsymbol{w}_{k+1}$ and $\boldsymbol{v}_{k,i}$ are needed.

Assumption 5. The initial value and the noises satisfy the following conditions:

$$\mathrm{E}[\|\boldsymbol{x}_0\|^{2r}]<\infty,\quad \sup_k \mathrm{E}[\|\boldsymbol{w}_{k+1}\|^{2r}+\|\boldsymbol{v}_{k,i}\|^{4r}]<\infty,\quad i\in\{1,\ldots,n\}.$$

3.2 The main results

We will first establish the stability of the covariance matrix $\boldsymbol{P}'_k$.

Theorem 1. Under Assumptions 1–3, the matrix $\boldsymbol{P}'_k$ defined in Table 1 is L_{2r}-stable.

Proof. By Assumption 1, we have $0 < a_{ij}^{(L)} < 1$ for $L \geqslant D(\mathscr{G})$ (see [33]). Let $A_\delta(k) = \{\lambda_{\min}(\boldsymbol{G}(k+h,k)) > \delta\}$, $\delta > 0$. We use $\{\boldsymbol{F}_k, \boldsymbol{H}_{k,i}, k \geqslant 0, i = 1,\ldots,n\}$ to construct an auxiliary time-varying linear system:

$$\begin{cases}\boldsymbol{\theta}_{k+1} = \boldsymbol{F}_k\boldsymbol{\theta}_k + \boldsymbol{\delta}_{k+1},\\ \boldsymbol{y}_{k,i}^0 = \sqrt{a}\boldsymbol{H}_{k,i}\boldsymbol{\theta}_k + \boldsymbol{\varepsilon}_{k,i},\end{cases} \tag{22}$$

where $a \triangleq \min_{i,j\in\{1,\ldots,n\}} a_{ij}^{(L)}$ is a positive constant. The initial condition $\boldsymbol{\theta}_0$ has a Gaussian distribution with mean $\hat{\boldsymbol{\theta}}'_0$ and covariance matrix $\boldsymbol{P}'_0$, and $\{\boldsymbol{\delta}_k, \boldsymbol{\varepsilon}_{k,i}, k \geqslant 0\}$ is a sequence of independent Gaussian random vectors, independent of $\{\boldsymbol{F}_k, \boldsymbol{H}_{k,i}, k \geqslant 0\}$ with the following properties:

$$\begin{aligned}&\mathrm{E}[\boldsymbol{\delta}_k] = 0, \quad \mathrm{E}[\boldsymbol{\delta}_k\boldsymbol{\delta}_k^{\mathrm{T}}] = \boldsymbol{Q}_k, \quad \mathrm{E}[\boldsymbol{\varepsilon}_{k,i}] = 0, \quad \mathrm{E}[\boldsymbol{\varepsilon}_{k,i}\boldsymbol{\varepsilon}_{k,i}^{\mathrm{T}}] = \boldsymbol{R}_{k,i},\\ &\mathrm{E}[\boldsymbol{\delta}_k\boldsymbol{\varepsilon}_{k,i}^{\mathrm{T}}] = 0, \quad \mathrm{E}[\boldsymbol{\varepsilon}_{k,i}\boldsymbol{\varepsilon}_{k,j}^{\mathrm{T}}] = 0, \quad i \neq j.\end{aligned}$$

Let

$$\begin{aligned}\boldsymbol{Y}_k &= \mathrm{col}\{\boldsymbol{y}_{k,1}^0,\ldots,\boldsymbol{y}_{k,n}^0\},\\ \boldsymbol{\Psi}_k &= \sqrt{a}\mathrm{col}\{\boldsymbol{H}_{k,1},\ldots,\boldsymbol{H}_{k,n}\},\\ \boldsymbol{\Xi}_k &= \mathrm{col}\{\boldsymbol{\varepsilon}_k^1,\ldots,\boldsymbol{\varepsilon}_k^n\}.\end{aligned}$$

Hence, Eq. (22) can be rewritten as

$$\begin{cases}\boldsymbol{\theta}_{k+1} = \boldsymbol{F}_k\boldsymbol{\theta}_k + \boldsymbol{\delta}_{k+1},\\ \boldsymbol{Y}_k = \boldsymbol{\Psi}_k\boldsymbol{\theta}_k + \boldsymbol{\Xi}_k.\end{cases}$$

Then, the minimum-variance linear estimate of $\boldsymbol{\theta}_k$ in (22) can be written as

$$\begin{aligned}&\hat{\boldsymbol{\theta}}'_{k+1} = \boldsymbol{F}_k\hat{\boldsymbol{\theta}}'_k + \boldsymbol{F}_k\hat{\boldsymbol{P}}'_k\boldsymbol{\Psi}_k^{\mathrm{T}}(\boldsymbol{\Psi}_k\hat{\boldsymbol{P}}'_k\boldsymbol{\Psi}_k^{\mathrm{T}} + \boldsymbol{R}_k)^{-1}(\boldsymbol{Y}_k - \boldsymbol{\Psi}_k\hat{\boldsymbol{\theta}}'_k),\\ &\hat{\boldsymbol{\theta}}'_k = \mathrm{E}[\boldsymbol{\theta}_k|\mathscr{F}_{k-1}],\\ &\hat{\boldsymbol{P}}'_k = \mathrm{E}[\widetilde{\boldsymbol{\theta}}_k\widetilde{\boldsymbol{\theta}}_k^{\mathrm{T}}|\mathscr{F}_{k-1}],\\ &\widetilde{\boldsymbol{\theta}}_k = \boldsymbol{\theta}_k - \hat{\boldsymbol{\theta}}'_k,\end{aligned}$$

where $\mathscr{F}_k = \sigma\{\mathscr{G}_\infty, \boldsymbol{Y}_0,\ldots,\boldsymbol{Y}_k\}$. Note that $\{\hat{\boldsymbol{P}}'_k\}$ can be recursively generated as follows:

$$\hat{\boldsymbol{P}}_k^{-1} = \hat{\boldsymbol{P}}_k'^{-1} + \sum_{j=1}^n a\boldsymbol{H}_{k,j}^{\mathrm{T}}\boldsymbol{R}_{k,j}^{-1}\boldsymbol{H}_{k,j}, \tag{23}$$

$$\hat{\boldsymbol{P}}'_{k+1} = \boldsymbol{F}_k\hat{\boldsymbol{P}}_k\boldsymbol{F}_k^{\mathrm{T}} + \boldsymbol{Q}_k. \tag{24}$$

For $h > 0$ defined in Assumption 2, we introduce another estimate of $\boldsymbol{\theta}_{k+h}$, denoted by $\boldsymbol{\theta}^*_{k+h}$, which is recursively defined by

$$\boldsymbol{\theta}^*_{k+h} = \frac{1}{a}\boldsymbol{\Phi}(k+h,k)\boldsymbol{G}^{-1}(k+h,k)\sum_{j=k+1}^{k+h}\boldsymbol{\Phi}^{\mathrm{T}}(j,k)\boldsymbol{\Psi}_j^{\mathrm{T}}\boldsymbol{Y}_jI_{A_\delta(k)} + \boldsymbol{\Phi}(k+h,k)\boldsymbol{\theta}^*_kI_{A^c_\delta(k)},$$

where the initial values $\boldsymbol{\theta}^*_m$ $(m = 1,2,\ldots,h-1)$ are defined as

$$\begin{cases}\boldsymbol{\theta}^*_0 = \hat{\boldsymbol{\theta}}'_0,\\ \boldsymbol{\theta}^*_{m+1} = \boldsymbol{F}_m\boldsymbol{\theta}^*_m.\end{cases} \tag{25}$$

Similar to the proof of Theorem 1.1 in [24], we can obtain that

$$\mathrm{E}[\|\hat{\boldsymbol{P}}'_{jh}\|^{2r}] \leqslant \mathrm{E}[\|\boldsymbol{\theta}_{jh} - \boldsymbol{\theta}^*_{jh}\|^{2r}] = O(1), \quad j = 1, 2 \ldots.$$

Then we have

$$\sup_k \mathrm{E}[\|\hat{\boldsymbol{P}}'_k\|^{2r}] < \infty. \tag{26}$$

In the following, we will prove that $\boldsymbol{P}'_{k,i}$ defined in Algorithm 2 is L_{2r}-stable for all $i \in \{1, \ldots, n\}$.

We use the induction method to prove that for $k \geqslant 0$, $\boldsymbol{P}'_{k,i} \leqslant \hat{\boldsymbol{P}}'_k, i = 1, 2, \ldots, n$ holds. For $k = 0$, we can choose the initial $\boldsymbol{P}'_{0,i}$ and $\hat{\boldsymbol{P}}'_0$ such that $\boldsymbol{P}'_{0,i} \leqslant \hat{\boldsymbol{P}}'_0$ holds for $i = 1, 2, \ldots, n$. Next, we assume that $\boldsymbol{P}'_{k,i} \leqslant \hat{\boldsymbol{P}}'_k$ holds for any $i \in \{1, \ldots, n\}$. Then, $\boldsymbol{P}'^{-1}_{k,i} \geqslant \hat{\boldsymbol{P}}'^{-1}_k$. Hence we have

$$\boldsymbol{P}'^{-1}_{k,i} + \sum_{j=1}^{n} a^L_{ij} \boldsymbol{H}^{\mathrm{T}}_{k,j} \boldsymbol{R}^{-1}_{k,j} \boldsymbol{H}_{k,j} \geqslant \hat{\boldsymbol{P}}'^{-1}_k + \sum_{j=1}^{n} a \boldsymbol{H}^{\mathrm{T}}_{k,j} \boldsymbol{R}^{-1}_{k,j} \boldsymbol{H}_{k,j}.$$

By (14) and (23), we have $\boldsymbol{P}^{-1}_{k,i} \geqslant \hat{\boldsymbol{P}}^{-1}_k$. Then

$$\boldsymbol{F}_k \boldsymbol{P}_{k,i} \boldsymbol{F}^{\mathrm{T}}_k + \boldsymbol{Q}_k \leqslant \boldsymbol{F}_k \hat{\boldsymbol{P}}_k \boldsymbol{F}^{\mathrm{T}}_k + \boldsymbol{Q}_k$$

holds, i.e.,

$$\boldsymbol{P}'_{k+1,i} \leqslant \hat{\boldsymbol{P}}'_{k+1}. \tag{27}$$

Hence, $\boldsymbol{P}'_{k,i} \leqslant \hat{\boldsymbol{P}}'_k, i = 1, 2, \ldots, n$ holds for all $k \geqslant 0$ by induction. Thus, by (26), we obtain that $\boldsymbol{P}'_{k,i}$ is L_{2r}-stable. This completes the proof of the theorem.

The following inequality is often used in the proof of the main results:

$$\lambda_{\max}(\boldsymbol{B}\boldsymbol{A}) = \lambda_{\max}(\boldsymbol{A}\boldsymbol{B}) = \lambda_{\max}(\boldsymbol{A}^{\frac{1}{2}} \boldsymbol{B} \boldsymbol{A}^{\frac{1}{2}}) \leqslant \lambda_{\max}(\boldsymbol{A}) \lambda_{\max}(\boldsymbol{B}), \tag{28}$$

where $\boldsymbol{A}$ and $\boldsymbol{B} \in \mathbb{R}^{n \times n}$ are two nonnegative definite matrices.

Lemma 3. For all $k > m \geqslant 0$, we have the following inequality:

$$\left\| \prod_{j=m}^{k-1} \bar{\boldsymbol{F}}_j \boldsymbol{P}_j \boldsymbol{P}'^{-1}_j \right\|^2 \leqslant \prod_{j=m}^{k-1} \left(1 - \frac{1}{1 + \|\bar{\boldsymbol{Q}}^{-1}_j\| \|\boldsymbol{P}'_{j+1}\|} \right) \|\boldsymbol{P}'_k\| \|\boldsymbol{P}'^{-1}_m\|.$$

Proof. Let us consider the following equation for $k > m$,

$$\boldsymbol{z}_{j+1} = \bar{\boldsymbol{F}}_j \boldsymbol{P}_j \boldsymbol{P}'^{-1}_j \boldsymbol{z}_j, \qquad j \in [m, k-1],$$

where $\boldsymbol{z}_m$ is taken to be deterministic and $\|\boldsymbol{z}_m\| = 1$. Then we have

$$\boldsymbol{z}_k = \prod_{j=m}^{k-1} \bar{\boldsymbol{F}}_j \boldsymbol{P}_j \boldsymbol{P}'^{-1}_j \boldsymbol{z}_m. \tag{29}$$

Taking the following Lyapunov function $V_j = \boldsymbol{z}^{\mathrm{T}}_j \boldsymbol{P}'^{-1}_j \boldsymbol{z}_j$, we have by Lemma 1

$$\begin{aligned} V_{j+1} &= \boldsymbol{z}^{\mathrm{T}}_{j+1} \boldsymbol{P}'^{-1}_{j+1} \boldsymbol{z}_{j+1} = \boldsymbol{z}^{\mathrm{T}}_j \boldsymbol{P}'^{-1}_j \boldsymbol{P}_j \bar{\boldsymbol{F}}^{\mathrm{T}}_j \boldsymbol{P}'^{-1}_{j+1} \bar{\boldsymbol{F}}_j \boldsymbol{P}_j \boldsymbol{P}'^{-1}_j \boldsymbol{z}_j \\ &= \boldsymbol{z}^{\mathrm{T}}_j \boldsymbol{P}'^{-1}_j \boldsymbol{P}_j \bar{\boldsymbol{F}}^{\mathrm{T}}_j (\bar{\boldsymbol{F}}_j \boldsymbol{P}_j \bar{\boldsymbol{F}}^{\mathrm{T}}_j + \bar{\boldsymbol{Q}}_j)^{-1} \bar{\boldsymbol{F}}_j \boldsymbol{P}_j \boldsymbol{P}'^{-1}_j \boldsymbol{z}_j \\ &= \boldsymbol{z}^{\mathrm{T}}_j \boldsymbol{P}'^{-1}_j [\boldsymbol{P}_j - (\boldsymbol{P}^{-1}_j + \bar{\boldsymbol{F}}^{\mathrm{T}}_j \bar{\boldsymbol{Q}}^{-1}_j \bar{\boldsymbol{F}}_j)^{-1}] \boldsymbol{P}'^{-1}_j \boldsymbol{z}_j \\ &= \boldsymbol{z}^{\mathrm{T}}_j \boldsymbol{P}'^{-\frac{1}{2}}_j \boldsymbol{P}'^{-\frac{1}{2}}_j [\boldsymbol{P}_j - (\boldsymbol{P}^{-1}_j + \bar{\boldsymbol{F}}^{\mathrm{T}}_j \bar{\boldsymbol{Q}}^{-1}_j \bar{\boldsymbol{F}}_j)^{-1}] \boldsymbol{P}'^{-\frac{1}{2}}_j \boldsymbol{P}'^{-\frac{1}{2}}_j \boldsymbol{z}_j \\ &\leqslant \lambda_{\max} \left(\boldsymbol{P}'^{-\frac{1}{2}}_j [\boldsymbol{P}_j - (\boldsymbol{P}^{-1}_j + \bar{\boldsymbol{F}}^{\mathrm{T}}_j \bar{\boldsymbol{Q}}^{-1}_j \bar{\boldsymbol{F}}_j)^{-1}] \boldsymbol{P}'^{-\frac{1}{2}}_j \right) \boldsymbol{z}^{\mathrm{T}}_j \boldsymbol{P}'^{-1}_j \boldsymbol{z}_j. \end{aligned}$$

By (14), it is clear that $\boldsymbol{P}_k'^{-1} \leqslant \boldsymbol{P}_k^{-1}$. Then by (28) we can obtain

$$\begin{aligned}V_{j+1} &\leqslant \lambda_{\max}([\boldsymbol{P}_j-(\boldsymbol{P}_j^{-1}+\bar{\boldsymbol{F}}_j^{\mathrm{T}}\bar{\boldsymbol{Q}}_j^{-1}\bar{\boldsymbol{F}}_j)^{-1}]\boldsymbol{P}_j'^{-1})\boldsymbol{z}_j^{\mathrm{T}}\boldsymbol{P}_j'^{-1}\boldsymbol{z}_j\\ &=\lambda_{\max}\Big([\boldsymbol{P}_j-(\boldsymbol{P}_j^{-1}+\bar{\boldsymbol{F}}_j^{\mathrm{T}}\bar{\boldsymbol{Q}}_j^{-1}\bar{\boldsymbol{F}}_j)^{-1}]^{\frac{1}{2}}\boldsymbol{P}_j'^{-1}\\ &\quad\cdot[\boldsymbol{P}_j-(\boldsymbol{P}_j^{-1}+\bar{\boldsymbol{F}}_j^{\mathrm{T}}\bar{\boldsymbol{Q}}_j^{-1}\bar{\boldsymbol{F}}_j)^{-1}]^{\frac{1}{2}}\Big)\boldsymbol{z}_j^{\mathrm{T}}\boldsymbol{P}_j'^{-1}\boldsymbol{z}_j\\ &\leqslant\lambda_{\max}\Big([\boldsymbol{P}_j-(\boldsymbol{P}_j^{-1}+\bar{\boldsymbol{F}}_j^{\mathrm{T}}\bar{\boldsymbol{Q}}_j^{-1}\bar{\boldsymbol{F}}_j)^{-1}]^{\frac{1}{2}}\boldsymbol{P}_j^{-1}\\ &\quad\cdot[\boldsymbol{P}_j-(\boldsymbol{P}_j^{-1}+\bar{\boldsymbol{F}}_j^{\mathrm{T}}\bar{\boldsymbol{Q}}_j^{-1}\bar{\boldsymbol{F}}_j)^{-1}]^{\frac{1}{2}}\Big)\boldsymbol{z}_j^{\mathrm{T}}\boldsymbol{P}_j'^{-1}\boldsymbol{z}_j\\ &=\lambda_{\max}\Big([\boldsymbol{P}_j-(\boldsymbol{P}_j^{-1}+\bar{\boldsymbol{F}}_j^{\mathrm{T}}\bar{\boldsymbol{Q}}_j^{-1}\bar{\boldsymbol{F}}_j)^{-1}]\boldsymbol{P}_j^{-1}\Big)\boldsymbol{z}_j^{\mathrm{T}}\boldsymbol{P}_j'^{-1}\boldsymbol{z}_j\\ &\leqslant\left\{1-\frac{1}{1+\lambda_{\max}(\bar{\boldsymbol{Q}}_j^{-1}\bar{\boldsymbol{F}}_j\boldsymbol{P}_j\bar{\boldsymbol{F}}_j^{\mathrm{T}})}\right\}\boldsymbol{z}_j^{\mathrm{T}}\boldsymbol{P}_j'^{-1}\boldsymbol{z}_j\\ &\leqslant\left\{1-\frac{1}{1+\|\bar{\boldsymbol{Q}}_j^{-1}\|\|\boldsymbol{P}_{j+1}'\|}\right\}V_j,\end{aligned}$$

where Eq. (19) is used in the last inequality. Hence we have

$$V_k \leqslant \prod_{j=m}^{k-1}\left(1-\frac{1}{1+\|\bar{\boldsymbol{Q}}_j^{-1}\|\|\boldsymbol{P}_{j+1}'\|}\right)V_m.$$

Combining this with (29), we have

$$\begin{aligned}\left\|\prod_{j=m}^{k-1}\bar{\boldsymbol{F}}_j\boldsymbol{P}_j\boldsymbol{P}_j'^{-1}\right\|^2 &= \max_{\|\boldsymbol{z}_m\|=1}\|\boldsymbol{z}_k\|^2=\max_{\|\boldsymbol{z}_m\|=1}\left\|\boldsymbol{z}_k\boldsymbol{P}_k'^{-\frac{1}{2}}\boldsymbol{P}_k'^{\frac{1}{2}}\right\|^2\\ &\leqslant \max_{\|\boldsymbol{z}_m\|=1}\left\|\boldsymbol{z}_k\boldsymbol{P}_k'^{-\frac{1}{2}}\right\|^2\left\|\boldsymbol{P}_k'^{\frac{1}{2}}\right\|^2=\max_{\|\boldsymbol{z}_m\|=1}(V_k\|\boldsymbol{P}_k'\|)\\ &\leqslant\left\{\prod_{j=m}^{k-1}\left(1-\frac{1}{1+\|\bar{\boldsymbol{Q}}_j^{-1}\|\|\boldsymbol{P}_{j+1}'\|}\right)\right\}\{\|\boldsymbol{P}_k'\|\max_{\|\boldsymbol{z}_m\|=1}V_m\}\\ &\leqslant\left\{\prod_{j=m}^{k-1}\left(1-\frac{1}{1+\|\bar{\boldsymbol{Q}}_j^{-1}\|\|\boldsymbol{P}_{j+1}'\|}\right)\right\}\{\|\boldsymbol{P}_k'\|\|\boldsymbol{P}_m'^{-1}\|\},\end{aligned}$$

which completes the proof.

In the following, we establish the L_p-exponential stability of the coefficient matrix $\{\bar{\boldsymbol{F}}_k\boldsymbol{P}_k\boldsymbol{P}_k'^{-1}, k\geqslant 0\}$ of (18).

Lemma 4. Under Assumptions 2 and 4, the scalar sequence in Lemma 3 belongs to $S^0(\lambda)$, i.e.,

$$\left\{1-\frac{1}{1+\|\bar{\boldsymbol{Q}}_{\boldsymbol{k}}^{\boldsymbol{-1}}\|\|\boldsymbol{P}_{k+1}'\|},\ k\geqslant 0\right\}\in S^0(\lambda).$$

Proof. By (27) and Lemma 2.4 in [24], we know that the following inequality:

$$\begin{aligned}&\mathrm{E}\left[\prod_{j=m}^{k-1}\left(1-\frac{1}{1+\|\boldsymbol{Q}_{\boldsymbol{j}}^{\boldsymbol{-1}}\|\|\boldsymbol{P}_{j+1,i}'\|}\right)\right]\\ &\leqslant\mathrm{E}\left[\prod_{j=m}^{k-1}\left(1-\frac{1}{1+\|\boldsymbol{Q}_{\boldsymbol{j}}^{\boldsymbol{-1}}\|\|\hat{\boldsymbol{P}}_{j+1}'\|}\right)\right]\leqslant C\lambda^{k-m}\end{aligned}$$

holds for all $i\in\{1,\ldots,n\}$. By the fact that $\boldsymbol{P}_k'$ is a partitioned diagonal matrix, we have

$$\mathrm{E}\left[\prod_{j=m}^{k-1}\left(1-\frac{1}{1+\|\bar{\boldsymbol{Q}}_{\boldsymbol{j}}^{\boldsymbol{-1}}\|\|\boldsymbol{P}_{j+1}'\|}\right)\right]$$

$$= \mathrm{E}\left[\prod_{j=m}^{k-1}\left(1-\frac{1}{1+\max_i\{\|\boldsymbol{P}'_{j+1,i}\|\}\|\boldsymbol{Q}_{\boldsymbol{j}}^{-\mathbf{1}}\|}\right)\right] \leqslant C\lambda^{k-m},$$

which completes the proof.

Theorem 2. Under Assumptions 1–4, the matrix sequence $\{\bar{\boldsymbol{F}}_k\boldsymbol{P}_k\boldsymbol{P}_k'^{-1}, k \geqslant 0\}$ is L_{2r}-exponentially stable, where r is defined in Assumption 3.

Proof. By Lemma 3 and the fact $\boldsymbol{P}'^{-1}_{k+1,i} \leqslant \boldsymbol{Q}_k^{-1} \leqslant \boldsymbol{Q}^{-1}$, we have

$$\begin{aligned}
&\left\|\prod_{j=m}^{k-1}\bar{\boldsymbol{F}}_k\boldsymbol{P}_k\boldsymbol{P}_k'^{-1}\right\|_{L_{2r}}^{2r}\\
&\leqslant \mathrm{E}\left[\left\{\prod_{j=m}^{k-1}\left(1-\frac{1}{1+\|\bar{\boldsymbol{Q}}_j^{-1}\boldsymbol{P}'_{j+1}\|}\right)\right\}^r\|\boldsymbol{P}'_k\|^r\|\boldsymbol{P}'^{-1}_m\|^r\right]\\
&\leqslant \|\boldsymbol{Q}^{-1}\|^r\sqrt{\mathrm{E}\left[\prod_{j=m}^{k-1}\left(1-\frac{1}{1+\|\bar{\boldsymbol{Q}}_j^{-1}\boldsymbol{P}'_{j+1}\|}\right)\right]^{2r}}\sqrt{\mathrm{E}[\|\boldsymbol{P}'_k\|^{2r}]}\\
&\leqslant \|\boldsymbol{Q}^{-1}\|^r\sqrt{\mathrm{E}\left[\prod_{j=m}^{k-1}\left(1-\frac{1}{1+\|\bar{\boldsymbol{Q}}_j^{-1}\boldsymbol{P}'_{j+1}\|}\right)\right]}\sqrt{\mathrm{E}[\|\boldsymbol{P}'_k\|^{2r}]}.
\end{aligned}$$

Therefore, by Theorem 1 and Lemma 4, the matrix sequence $\{\bar{\boldsymbol{F}}_k\boldsymbol{P}_k\boldsymbol{P}_k'^{-1}, k \geqslant 0\}$ is L_{2r}-exponentially stable, which completes the proof.

In the following, we consider the stability of the DKF.

Lemma 5 ([34])**.** Let $\mathscr{A} = \mathcal{A}\otimes\boldsymbol{I}_s$ with $\mathcal{A} = \{a_{ij}\} \in \mathbb{R}^{n\times n}$ being a doubly stochastic matrix, and

$$\boldsymbol{Q} = \mathrm{diag}\{\boldsymbol{Q}_1,\ldots,\boldsymbol{Q}_n\}, \quad \boldsymbol{Q}' = \mathrm{diag}\{\boldsymbol{Q}'_1,\ldots,\boldsymbol{Q}'_n\},$$

where $\boldsymbol{Q}_i \in \mathbb{R}^{s\times s}, i = 1,\ldots,n$ are nonnegative definite matrices satisfying $\boldsymbol{Q}'_i = \sum_{j=1}^n a_{ij}\boldsymbol{Q}_j$. Then the following inequality holds:

$$\mathscr{A}^{\mathrm{T}}\boldsymbol{Q}\mathscr{A} \leqslant \boldsymbol{Q}'.$$

Theorem 3. Under Assumptions 1–5, the estimation error $\widetilde{\boldsymbol{X}}_k$ defined in (18) is bounded, i.e.,

$$\sup_k \|\widetilde{\boldsymbol{X}}_k\|_{L_r} < \infty.$$

Proof. Let $\boldsymbol{\Delta}_{k+1} = \boldsymbol{W}_{k+1} - \bar{\boldsymbol{F}}_k\boldsymbol{P}_k\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-1}\boldsymbol{V}_k$. Then by (18), we have

$$\widetilde{\boldsymbol{X}}_{k+1} = \prod_{j=0}^{k}(\bar{\boldsymbol{F}}_j\boldsymbol{P}_j\boldsymbol{P}_j'^{-1})\widetilde{\boldsymbol{X}}_0 + \sum_{j=1}^{k+1}\prod_{m=j}^{k}(\bar{\boldsymbol{F}}_m\boldsymbol{P}_m\boldsymbol{P}_m'^{-1})\boldsymbol{\Delta}_j.$$

By Theorem 2, Assumption 5, and the Hölder inequality, we have

$$\begin{aligned}
\|\widetilde{\boldsymbol{X}}_{k+1}\|_r &\leqslant \left\|\prod_{j=0}^{k}(\bar{\boldsymbol{F}}_j\boldsymbol{P}_j\boldsymbol{P}_j'^{-1})\right\|_{L_{2r}}\|\widetilde{\boldsymbol{X}}_0\|_{L_{2r}} + \sum_{j=1}^{k+1}\left\|\prod_{m=j}^{k}(\bar{\boldsymbol{F}}_m\boldsymbol{P}_m\boldsymbol{P}_m'^{-1})\right\|_{L_{2r}}\|\boldsymbol{\Delta}_j\|_{L_{2r}}\\
&\leqslant O(1)C\lambda^{k+1} + C\sum_{j=1}^{k+1}\lambda^{k-j+1}\|\boldsymbol{\Delta}_j\|_{L_{2r}}\\
&= O(1) + C\sum_{j=0}^{k}\lambda^{j}\|\boldsymbol{\Delta}_{k-j+1}\|_{L_{2r}}. \qquad (30)
\end{aligned}$$

By (19) and (28), we have

$$\begin{aligned}
&\|\bar{\boldsymbol{F}}_k\boldsymbol{P}_k\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-1}\|^2\\
&=\lambda_{\max}(\bar{\boldsymbol{F}}_k\boldsymbol{P}_k\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-2}\boldsymbol{H}_k\mathscr{A}^L\boldsymbol{P}_k\bar{\boldsymbol{F}}_k^{\mathrm{T}})\\
&=\lambda_{\max}\left(\bar{\boldsymbol{F}}_k\boldsymbol{P}_k^{\frac{1}{2}}\boldsymbol{P}_k^{\frac{1}{2}}\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-2}\boldsymbol{H}_k\mathscr{A}^L\boldsymbol{P}_k^{\frac{1}{2}}\boldsymbol{P}_k^{\frac{1}{2}}\bar{\boldsymbol{F}}_k^{\mathrm{T}}\right)\\
&=\lambda_{\max}\left(\boldsymbol{P}_k^{\frac{1}{2}}\bar{\boldsymbol{F}}_k^{\mathrm{T}}\bar{\boldsymbol{F}}_k\boldsymbol{P}_k^{\frac{1}{2}}\boldsymbol{P}_k^{\frac{1}{2}}\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-2}\boldsymbol{H}_k\mathscr{A}^L\boldsymbol{P}_k^{\frac{1}{2}}\right)\\
&\leqslant\lambda_{\max}\left(\boldsymbol{P}_k^{\frac{1}{2}}\bar{\boldsymbol{F}}_k^{\mathrm{T}}\bar{\boldsymbol{F}}_k\boldsymbol{P}_k^{\frac{1}{2}}\right)\lambda_{\max}\left(\boldsymbol{P}_k^{\frac{1}{2}}\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-\frac{1}{2}}\boldsymbol{R}_k^{-1}\boldsymbol{R}_k^{-\frac{1}{2}}\boldsymbol{H}_k\mathscr{A}^L\boldsymbol{P}_k^{\frac{1}{2}}\right)\\
&\leqslant\lambda_{\max}(\boldsymbol{P}_k')\lambda_{\max}\left(\boldsymbol{R}_k^{-\frac{1}{2}}\boldsymbol{H}_k\mathscr{A}^L\boldsymbol{P}_k^{\frac{1}{2}}\boldsymbol{P}_k^{\frac{1}{2}}\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-\frac{1}{2}}\boldsymbol{R}_k^{-1}\right)\\
&\leqslant\lambda_{\max}(\boldsymbol{P}_k')\lambda_{\max}\left(\boldsymbol{P}_k^{\frac{1}{2}}\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\mathscr{A}^L\boldsymbol{P}_k^{\frac{1}{2}}\right)\lambda_{\max}(\boldsymbol{R}_k^{-1}). \qquad (31)
\end{aligned}$$

Notice that $\mathcal{A}^L$ is still a doubly stochastic matrix. By Lemma 5 and (14), we have

$$\begin{aligned}
\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\mathscr{A}^L &\leqslant \operatorname{diag}\left\{\sum_{j=1}^n a_{1j}^{(L)}\boldsymbol{H}_{k,j}^{\mathrm{T}}\boldsymbol{R}_{k,j}^{-1}\boldsymbol{H}_{k,j},\ldots,\sum_{j=1}^n a_{nj}^{(L)}\boldsymbol{H}_{k,j}^{\mathrm{T}}\boldsymbol{R}_{k,j}^{-1}\boldsymbol{H}_{k,j}\right\}\\
&\leqslant \operatorname{diag}\{\boldsymbol{P}_{k,1}^{-1},\ldots,\boldsymbol{P}_{k,n}^{-1}\}\\
&=\boldsymbol{P}_k^{-1}.
\end{aligned}$$

Combining this with (31), we have

$$\|\bar{\boldsymbol{F}}_k\boldsymbol{P}_k\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-1}\|^2\leqslant\|\boldsymbol{P}_k'\|\operatorname{tr}(\boldsymbol{R}_k^{-1})\leqslant d\|\boldsymbol{P}_k'\|\sum_{i=1}^n\alpha_i^{-1},$$

where d is the dimension of $\boldsymbol{v}_{k,i}$.

Hence, by Theorem 1, it is easy to see that $\sup_k\|\bar{\boldsymbol{F}}_k\boldsymbol{P}_k\mathscr{A}^L\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-1}\|_{L_{4r}}=O(1)$. Then by Assumption 5, we have for all $k\geqslant 1$,

$$\begin{aligned}
\|\boldsymbol{\Delta}_k\|_{2r}&=\|\boldsymbol{W}_k-\bar{\boldsymbol{F}}_{k-1}\boldsymbol{P}_{k-1}\mathscr{A}^L\boldsymbol{H}_{k-1}^{\mathrm{T}}\boldsymbol{R}_{k-1}^{-1}\boldsymbol{V}_{k-1}\|_{L_{2r}}\\
&\leqslant\|\boldsymbol{W}_k\|_{L_{2r}}+\|\bar{\boldsymbol{F}}_{k-1}\boldsymbol{P}_{k-1}\mathscr{A}^L\boldsymbol{H}_{k-1}^{\mathrm{T}}\boldsymbol{R}_{k-1}^{-1}\|_{L_{4r}}\|\boldsymbol{V}_{k-1}\|_{L_{4r}}\\
&=O(1).
\end{aligned}$$

By (30), we have

$$\|\widetilde{\boldsymbol{X}}_{k+1}\|_{L_r}=O(1)+O(1)\sum_{j=0}^k\lambda^j=O(1).$$

This completes the proof of the theorem.

4 A simulation example

In this section, we illustrate the cooperative effect of the sensors by Example 1.

Example 1. Consider the one-dimensional case where the network is composed of three sensors. The dynamics of each sensor obeys (1) with $\boldsymbol{F}_k=2$. The noises $\{\boldsymbol{w}_k\}$ and $\{\boldsymbol{v}_{k,i}\}$ are identically independently distributed (i.i.d.) sequences and satisfy $\boldsymbol{w}_k\sim N(0,1)$, $\boldsymbol{v}_{k,1}\sim N(0,0.1)$, $\boldsymbol{v}_{k,2}\sim N(0,0.2)$, $\boldsymbol{v}_{k,3}\sim N(0,0.3)$. Suppose that $\{\boldsymbol{H}_{k,i}\}$ is an i.i.d. sequence and obeys the following distribution:

$$\begin{aligned}
&P(\boldsymbol{H}_{k,1}=0,\boldsymbol{H}_{k,2}=0,\boldsymbol{H}_{k,3}=1)=0.1,\quad P(\boldsymbol{H}_{k,1}=0,\boldsymbol{H}_{k,2}=2,\boldsymbol{H}_{k,3}=0)=0.2,\\
&P(\boldsymbol{H}_{k,1}=0,\boldsymbol{H}_{k,2}=2,\boldsymbol{H}_{k,3}=1)=0.15,\quad P(\boldsymbol{H}_{k,1}=1,\boldsymbol{H}_{k,2}=0,\boldsymbol{H}_{k,3}=0)=0.15,\\
&P(\boldsymbol{H}_{k,1}=1,\boldsymbol{H}_{k,2}=0,\boldsymbol{H}_{k,3}=1)=0.1,\quad P(\boldsymbol{H}_{k,1}=1,\boldsymbol{H}_{k,2}=2,\boldsymbol{H}_{k,3}=0)=0.1,\\
&P(\boldsymbol{H}_{k,1}=1,\boldsymbol{H}_{k,2}=2,\boldsymbol{H}_{k,3}=1)=0.2.
\end{aligned}$$

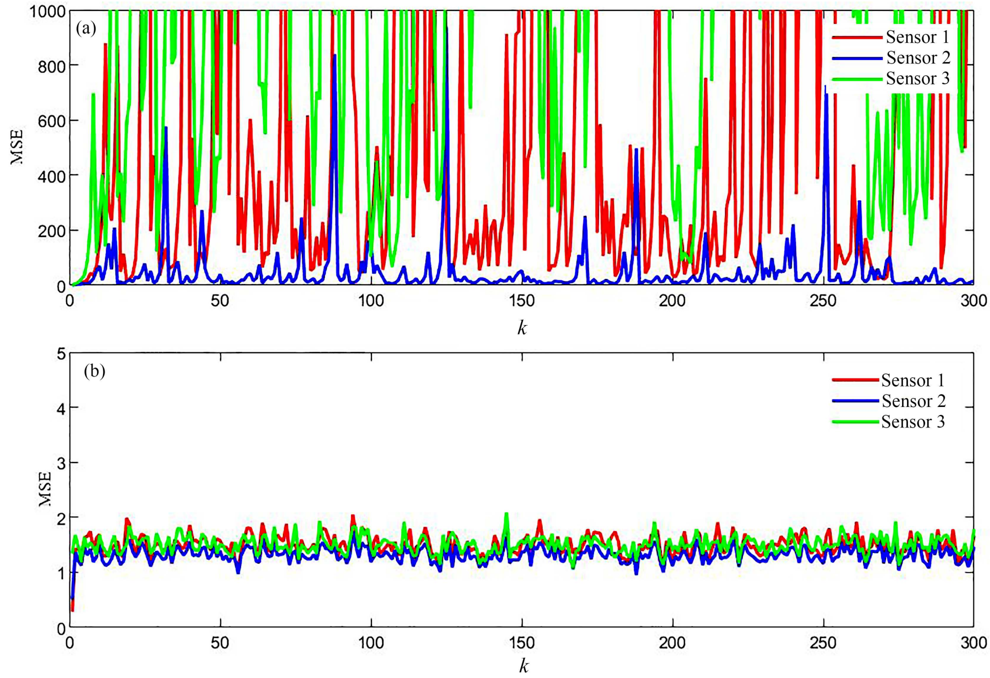

Figure 1 (Color online) MSE of three sensors in Example 1. (a) Non-cooperative KF; (b) DKF.

The adjacency matrix is taken as

$$\mathcal{A} = \begin{pmatrix} 2/3 & 1/3 & 0 \\ 1/3 & 1/2 & 1/6 \\ 0 & 1/6 & 5/6 \end{pmatrix}.$$

By Example 1.1 in [24], one can verify that none of the regression signals $\boldsymbol{H}_{k,i}$ $(i = 1, 2, 3)$ of the three individual sensors satisfy the excitation condition (21), but they can cooperate to satisfy Assumption 2. In order to estimate the unknown state $\boldsymbol{x}_k$, The mean square error (MSE) of the sensors (averaged over 200 runs) is shown in Figure 1. From this figure, we see that the MSE of each sensor using the non-cooperative KF is unbounded while the MSE of the sensors using the DKF proposed in this paper is bounded, which means that the estimation task can be still fulfilled through exchanging the information between sensors even though any individual sensor cannot.

5 Concluding remarks

In this paper, we proposed a DKF to estimate the state of the dynamical system with random coefficients by diffusing the local innovation pairs over the network. Further, we introduced the collective random observability condition. The theoretical analysis of the stability of the proposed DKF was established, and the boundedness of the state estimation error was presented without the independency and statistical stationarity assumptions on the measurement signals. However, the collective random observability condition may still be conservative, and how to relax the excitation condition to guarantee the stability of the DKF with random dropouts suggests suitable future investigations.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 11688101) and National Key R&D Program of China (Grant No. 2018YFA0703800).

References

1 Tang J G, Li B, Chang J, et al. Unscented Kalman-filter-based sliding mode control for an underwater gliding snake-like robot. Sci China Inf Sci, 2020, 63: 112207
2 Kalman R E. A new approach to linear filtering and prediction problems. J Basic Eng, 1960, 82: 35–45
3 Kailath T, Sayed A H, Hassibi B. Linear Estimation. Upper Saddle River: Prentice-Hall, 2000
4 Deng Z L, Gao Y, Mao L, et al. New approach to information fusion steady-state Kalman filtering. Automatica, 2005, 41: 1695–1707
5 Battistelli G, Chisci L, Mugnai G, et al. Consensus-based linear and nonlinear filtering. IEEE Trans Automat Contr, 2015, 60: 1410–1415

6 Olfati-Saber R. Distributed Kalman filtering for sensor networks. In: Proceedings of the 46th IEEE Conference on Decision and Control, New Orleans, 2007. 5492–5498

7 Li W L, Jia Y M. Distributed consensus filtering for discrete-time nonlinear systems with non-Gaussian noise. Signal Process, 2012, 92: 2464–2470

8 Zhang Z, Li J X, Liu L. Distributed state estimation and data fusion in wireless sensor networks using multi-level quantized innovation. Sci China Inf Sci, 2016, 59: 022316

9 Zhao Y, He X, Zhou D H. Distributed filtering for time-varying networked systems with sensor gain degradation and energy constraint: a centralized finite-time communication protocol scheme. Sci China Inf Sci, 2018, 61: 092208

10 Chen C, Liu Z X, Guo L. Performance bounds of distributed adaptive filters with cooperative correlated signals. Sci China Inf Sci, 2016, 59: 112202

11 Xie S Y, Guo L. Analysis of distributed adaptive filters based on diffusion strategies over sensor networks. IEEE Trans Automat Contr, 2018, 63: 3643–3658

12 Battistelli G, Chisci L. Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability. Automatica, 2014, 50: 707–718

13 Zhou J M, Gu G X, Chen X. Distributed Kalman filtering over wireless sensor networks in the presence of data packet drops. IEEE Trans Automat Contr, 2019, 64: 1603–1610

14 Liu Q Y, Wang Z D, He X, et al. On Kalman-consensus filtering with random link failures over sensor networks. IEEE Trans Automat Contr, 2018, 63: 2701–2708

15 Carli R, Chiuso A, Schenato L, et al. Distributed Kalman filtering based on consensus strategies. IEEE J Sel Areas Commun, 2008, 26: 622–633

16 Leng M, Tay W P, Quek T Q S, et al. Distributed local linear parameter estimation using Gaussian SPAWN. IEEE Trans Signal Process, 2015, 63: 244–257

17 Sayed A H, Tu S Y, Chen J S, et al. Diffusion strategies for adaptation and learning over networks: an examination of distributed strategies and network behavior. IEEE Signal Process Mag, 2013, 30: 155–171

18 He X K, Hu C, Hong Y H, et al. Distributed Kalman filters with state equality constraints: time-based and event-triggered communications. IEEE Trans Automat Contr, 2020, 65: 28–43

19 Wang S, Ren W. On the convergence conditions of distributed dynamic state estimation using sensor networks: a unified framework. IEEE Trans Contr Syst Technol, 2018, 26: 1300–1316

20 Olfati-Saber R, Jalalkamali P. Coupled distributed estimation and control for mobile sensor networks. IEEE Trans Automat Contr, 2012, 57: 2609–2614

21 Yang W, Chen G, Wang X, et al. Stochastic sensor activation for distributed state estimation over a sensor network. Automatica, 2014, 50: 2070–2076

22 Schizas I D, Mateos G, Giannakis G B. Distributed LMS for consensus-based in-network adaptive processing. IEEE Trans Signal Process, 2009, 57: 2365–2382

23 Takahashi N, Yamada I, Sayed A H. Diffusion least-mean squares with adaptive combiners: formulation and performance analysis. IEEE Trans Signal Process, 2010, 58: 4795–4810

24 Wang Y, Guo L. On stability of random Riccati equations. Sci China Ser E-Technol Sci, 1999, 42: 136–148

25 Guo L. Stability of recursive stochastic tracking algorithms. SIAM J Control Opt, 1994, 32: 1195–1225

26 Guo L. Time-varying Stochastic Systems. Jilin: Jilin Scientific and Technological Press, 1993

27 Welch G, Bishop G. An Introduction to the Kalman Filter. Department of Computer Science, University of North Carolina at Chapel Hill, Technical Report, 2000

28 Olfati-Saber R. Distributed Kalman filter with embedded consensus filters. In: Proceedings of the 44th IEEE Conference on Decision and Control, Seville, 2005. 8179–8184

29 Speranzon A, Fischione C, Johansson K H, et al. A distributed minimum variance estimator for sensor networks. IEEE J Sel Areas Commun, 2008, 26: 609–621

30 Boem F, Xu Y, Fischione C, et al. A distributed pareto-optimal dynamic estimation method. In: Proceedings of European Control Conference, 2015. 3673–3680

31 He X K, Xue W C, Fang H T. Consistent distributed state estimation with global observability over sensor network. Automatica, 2018, 92: 162–172

32 Talebi S P, Werner S. Distributed Kalman filtering and control through embedded average consensus information fusion. IEEE Trans Automat Contr, 2019, 64: 4396–4403

33 Godsil C, Royle G. Algebraic Graph Theory. Berlin: Springer, 2014

34 Xie S Y, Zhang Y Q, Guo L. Learning and prediction theory of distributed least squares. 2019. ArXiv: 1912.11686