Available online at www.sciencedirect.com

ScienceDirect

Journal of the Franklin Institute 354 (2017) 6200–6216

www.elsevier.com/locate/jfranklin

Distributed dynamic state estimation with parameter identification for large-scale systems☆

Yibing Sun[a,b], Minyue Fu[c,d,*], Bingchang Wang[a], Huanshui Zhang[a]

[a] School of Control Science and Engineering, Shandong University, Jinan 250061, China
[b] School of Mathematical Sciences, University of Jinan, Jinan 250022, China
[c] School of Electrical Engineering and Computer Science, University of Newcastle, NSW 2308, Australia
[d] Guangdong Key Laboratory of IoT Information Processing, School of Automation, Guangdong University of Technology, Guangzhou 510006, China

Received 14 June 2016; received in revised form 27 March 2017; accepted 2 July 2017
Available online 15 July 2017

Abstract

In this paper, we consider a distributed dynamic state estimation problem for time-varying systems. Based on the distributed maximum *a posteriori* (MAP) estimation algorithm proposed in our previous study, which studies the linear measurement models of each subsystem, and by weakening the constraint condition as that each time-varying subsystem is observable, this paper proves that the error covariances of state estimation and prediction obtained from the improved algorithm are respectively positive definite and have upper bounds, which verifies the feasibility of this algorithm. We also use new weighting functions and time-varying exponential smoothing method to ensure the robustness and improve the forecast accuracy of the distributed state estimation method. At last, an example is used to demonstrate the effectiveness of the proposed algorithm together with the parameter identification.

1. Introduction

In recent years, large-scale complex systems, such as power systems, traffic networks and multi-agent systems, have received significant attention from researchers in different

☆ This paper was not presented at any IFAC meeting. This work is supported by the National Science Foundation of China under Grants 61120106011, 61573221, 61403233.

* Corresponding author at: School of Electrical Engineering and Computer Science, University of Newcastle, NSW 2308, Australia.

E-mail address: Minyue.fu@newcastle.edu.au (M. Fu).

http://dx.doi.org/10.1016/j.jfranklin.2017.07.019

practical domains. The emergence of sensor networks calls for the development of distributed algorithms to replace the centralized methods [1]. This problem is motivated by many applications [2,3], which involve large-scale networked systems. We take the state estimation problem in power systems as an example, where local state of each subsystem is estimated by using measurements provided by Supervisory Control and Data Acquisition (SCADA) systems [4] and Phasor Measurement Units (PMUs) [5]. For large power systems, it is unrealistic and unnecessary for each subsystem to estimate the global state of systems [6]. On the other hand, considering the location problem for sensor networks, the locations of all sensors are estimated by utilizing relative position measurements between neighboring sensors [7]. For a small sensor network with a few sensors, it is possible to aggregate all measurements at the fusion center to calculate a whole location estimate for all sensors. However, this requires a great deal of computing resources with the increase of network size, and it is needless for each sensor to localize other sensors [8]. All this naturally promotes the development of distributed estimation algorithms.

State estimation methods can be mainly divided into two categories: centralized and distributed algorithms. In the centralized state estimation [9–12], all measured data are communicated to the fusion or control center for processing. It can provide the optimal state estimate for the entire network, due to the smallest information loss. But the computational burden at the fusion or control center increases exponentially when the system becomes large [13]. The distributed estimation methods, which are attractively alternative, can be classified into the static and dynamic estimation. At present, there have been much research activity focused on the distributed methods to static state estimation for many practical domains [8,13–16]. In particular, these methods are widely used in power systems to implement security monitoring of the current operating status [8,13,14]. In distributed static estimation, local state or a set of parameters of each subsystem is estimated by using the measurements at a fixed time. Marelli and Fu [8] studied a distributed weighted least square (WLS) estimation problem for a large-scale system. For the network with loops, this method was used for each subsystem to asymptotically compute the globally optimal estimate of parameters using its own measurement and information transmitted from its neighbors, and a preconditioning method was also used to speed up the convergence rate. Distributed dynamic estimation [17–21], which is mainly related to the Kalman filter, uses current measurements and the predicted values obtained from the previous time to estimate system state. Furthermore, the ability of predicting the future state, which is not captured by static estimation, plays an important role in real-time control and security analysis. Sharma et al. [20] proposed a multi-agent-based multi-area dynamic state estimator for power systems utilizing the cubature Kalman filter, which was run for all the extended subsystems independently and in parallel to reduce the overall execution time. But the partitioned power system was overlapping, and [20] needed the constraint condition that the voltage magnitudes and phase angles of the same boundary buses of the nearby subsystems must be equal.

Based on the distributed MAP estimation algorithm proposed in [21], this paper extends and generalizes this algorithm to nonlinear systems, and we will improve it in the following ways: (1) the work [21] needs the assumption that the local measurement matrix C_i has full column rank, which is relaxed in this paper as that each subsystem is observable; (2) taking into account the bad data, we use new weighting functions for all subsystems to ensure the robustness of the distributed algorithm, and we also use a modified exponential smoothing method to compute model parameters. The main contributions of this paper are as below:

- Under the relaxed assumption, we prove that the error covariance matrices of state estimation and prediction obtained from the proposed algorithm are positive definite by using the mathematical induction, and further prove that error covariance matrices have upper bounds by the rank criterion of the observability for time-varying system;
- In order to online compute the state-space parameters, we extend the method proposed in [22] to a time-varying model, which improves the precision of state prediction;
- In the simulations, the centralized and distributed state estimators are tested through the IEEE 118-bus system under different scenarios, which include normal operating condition, sudden load change and bad measurements.

Test results verify the validity of the distributed state estimation algorithm together with parameter identification, demonstrate the main results of this paper, and support the feasibility of the proposed algorithm for state estimation applications in large-scale systems.

The rest of the paper is organized as follows. Section 2 presents some related graph notations, and describes the system model. Section 3 discusses the centralized state estimator and proves that its error covariance matrices are bounded. Section 4 introduces the distributed state estimation algorithm together with the parameter identification, and gives the main results of this paper. Section 5 shows the simulation results. The conclusions are drawn in Section 6.

The following notations will be used throughout this paper. $\mathbb{N}_0$ denotes the set of non-negative integers, while $\mathbb{N}$ is the set of positive integers. $\mathbb{R}^l$ denotes the set of $l-$dimensional real column vectors and $\mathbb{R}^{l\times q}$ denotes the set of $l\times q$ real matrices. I denotes the identity matrix with appropriate dimension. For a given vector or matrix M, M^T denotes its transpose. For square symmetric matrices X and Y, $X<Y$ means that the matrix $Y-X$ is positive definite. The shorthand $\mathrm{diag}\{A_1,\ A_2,\ldots,\ A_n\}$ denotes a block diagonal matrix with diagonal blocks being matrices $A_1,\ldots,\ A_n$.

2. System model and problem description

In this section, we first present some preliminary notations on graph theory, which will be used in this paper. An undirected graph $\mathcal{G}$ contains a node set $\mathcal{V}=\{1,\ldots,\ N\}$ and an edge set of unordered pairs $\mathcal{E}=\{(i,j),\ \ i,j\in\mathcal{V}\}$. If there exists an edge between nodes i and j, then nodes i and j are called adjacent. A path from node i_1 to i_k is a sequence of edges $(i_1,i_2),\ (i_2,i_3),\ldots,\ (i_{k-1},i_k)$. If there exists a path between any two nodes of $\mathcal{G}$, then graph $\mathcal{G}$ is connected. The radium Γ_i of node i denotes the maximum length of a path between node i and any other node in $\mathcal{G}$, and the diameter Γ of graph $\mathcal{G}$ is the maximum radium of all nodes, i.e., $\Gamma=\max\{\Gamma_i,\ \ i\in\mathcal{V}\}$. We assume that the considered graph $\mathcal{G}$ in this paper is undirected and connected, which is void of self-loops and multiple edges. $\mathcal{N}_i=\{j\in\mathcal{V}:\ \ (i,\ \ j)\in\mathcal{E}\}$ denotes the neighbor set of node i, and $\mathcal{N}_i/\{j\}$ denotes that node j is removed from $\mathcal{N}_i$.

Consider a network formed by N nonoverlapping subsystems (called nodes). From the local viewpoint of node i, we consider the linear dynamic system as follows

$$x_i(k+1)=A_i(k)x_i(k)+G_i(k)+\omega_i(k), \tag{1}$$

where $k\in\mathbb{N}_0$ is the time sample, $x_i(k)\in\mathbb{R}^{s_i}$ is the local state of node i, $A_i(k)\in\mathbb{R}^{s_i\times s_i}$ is the system matrix which is usually assumed diagonal (see [9] and [11]), $G_i(k)\in\mathbb{R}^{s_i}$ describes the trend behavior of the state trajectory, and $\omega_i(k)\sim\mathcal{N}(0,R_i(k))$ is the model noise with $R_i(k)\geq 0$. It is also assumed that the initial state $x_i(0)\sim\mathcal{N}(\bar{x}_i(0),P_i(0))$ is irrelevant to the model and measurement noises, and $P_i(0)$ is a positive definite matrix.

The measurements of each node can be classified into two types, i.e., *the local measurement* which is only functions of the state of every node, and *the edge measurement* representing the tie-line measurements related to the neighboring nodes. Therefore, the two types of measurement equations of node i can be represented respectively as follows

$$z_{i,i}(k) = f_i(x_i(k)) + v_{i,i}(k), \tag{2}$$

$$z_{i,j}(k) = h_{i,j}(x_i(k), x_j(k)) + v_{i,j}(k), \tag{3}$$

where $z_{i,i}(k) \in \mathbb{R}^{q_{i,i}}$ is the local measurement vector of node i, $z_{i,j}(k) \in \mathbb{R}^{q_{i,j}}$ describes the interaction between nodes i and j,$v_{i,i}(k)$ and $v_{i,j}(k)$ are the associated measurement noises, assumed to be independent white Gaussian with zero mean and covariances $S_i(k) > 0$ and $T_{i,j}(k) > 0$. $f_i(\cdot)$ and $h_{i,j}(\cdot)$ are vectors consisting of nonlinear functions. Furthermore, we assume that the model and measurement noises are mutually independent.

Linearizing around the operating points $x_i^0(k)$ and $x_j^0(k)$, the Jacobian matrices of $f_i(\cdot)$ and $h_{i,j}(\cdot)$ are derived as

$$C_i(x_i^0) = \left.\frac{\partial f_i(x_i)}{\partial x_i}\right|_{x_i = x_i^0},$$
$$B_{i,j}(x_i^0) = \left.\frac{\partial h_{i,j}(x_i, x_j)}{\partial x_i}\right|_{x_i = x_i^0},$$
$$B_{j,i}(x_j^0) = \left.\frac{\partial h_{i,j}(x_i, x_j)}{\partial x_j}\right|_{x_j = x_j^0}.$$

In the rest of the paper, we choose the state prediction $\tilde{x}_i(k)$ of each node i as $x_i^0(k)$ at each time k, and we express the above mentioned matrices as $C_i(k)$, $B_{i,j}(k)$ and $B_{j,i}(k)$ for simplicity.

3. Centralized state estimation process

In order to propose the centralized state estimator, we firstly describe the state and measurement models for the whole system. Aggregating the states and measurements in Eqs. (1)–(3), the stacked state and measurement equations can be written as

$$x(k+1) = A(k)x(k) + G(k) + \omega(k), \tag{4}$$

$$z(k) = f(x(k)) + v(k), \tag{5}$$

where

$$x(k) = \left(x_1^T(k), \ldots, x_N^T(k)\right)^T \in \mathbb{R}^p,$$
$$z(k) = \left(\cdots, z_{i,i}^T(k), \ldots, z_{i,j}^T(k), \ldots\right)^T,$$
$$A(k) = \mathrm{diag}\left\{A_1(k), \ldots, A_N(k)\right\},$$

$$G(k) = \left(G_1^T(k), \ldots, \ G_N^T(k) \right)^T,$$

$$f(x(k)) = \left(\cdots, \ f_i^T(x_i(k)), \ldots, \ h_{i,j}^T(x_i(k), x_j(k)), \ldots \right)^T,$$

and $\Sigma_{i=1}^N s_i = p$. The noises are $\omega(k) \sim \mathcal{N}(0, R(k))$ and $\nu(k) \sim \mathcal{N}(0, R_*(k))$ with

$$R(k) = \text{diag}\left\{ R_1(k), \ldots, \ R_N(k) \right\},$$

$$R_*(k) = \text{diag}\left\{ \cdots, \ S_i(k), \ldots, \ T_{i,j}(k), \ldots \right\}.$$

With respect to the operating point $x^0(k)$, the Jacobian matrix of $f(x(k))$ is $H(k)$, which is an aggregation of Jacobian matrices of local and edge functions, and can be written as

$$H(k) = \begin{bmatrix} & & & \cdots & & & \\ \cdots & 0 & C_i(k) & 0 & \cdots & & \\ & & & \cdots & & & \\ \cdots & 0 & B_{i,j}(k) & 0 & B_{j,i}(k) & 0 & \cdots \\ & & & \cdots & & & \end{bmatrix}. \tag{6}$$

Also, the initial state $x(0)$ has mean $\bar{x}(0) = (\bar{x}_1^T(0), \ldots, \ \bar{x}_N^T(0))^T$ and covariance $P(0) = \text{diag}\{P_1(0), \ldots, \ P_N(0)\}$.

3.1. Centralized state estimator

The centralized state estimator used in this paper is the extended Kalman filter, where the state estimation $\hat{x}(k)$ along with its covariance matrix $\Sigma(k)$ can be expressed as follows:

$$\begin{aligned} \hat{x}(k) &= Q^{-1}(k)\alpha(k), \\ \Sigma(k) &= Q^{-1}(k), \end{aligned} \tag{7}$$

where

$$\begin{aligned} \alpha(k) &= H^T(k)R_*^{-1}(k)\Big(z(k) + H(k)\tilde{x}(k) - f(\tilde{x}(k)) \Big) + M^{-1}(k)\tilde{x}(k), \\ Q(k) &= H^T(k)R_*^{-1}(k)H(k) + M^{-1}(k), \end{aligned} \tag{8}$$

are initialized by $\tilde{x}(0) = \bar{x}(0)$ and $M(0) = P(0)$.

Executing the conditional expectation on Eq. (4), we obtain

$$\begin{aligned} \tilde{x}(k+1) &= A(k)\hat{x}(k) + G(k), \\ M(k+1) &= A(k)\Sigma(k)A^T(k) + R(k), \end{aligned} \tag{9}$$

where $\tilde{x}(k+1)$ is the predicted state vector based on the measurements from time 0 to k, and $M(k+1)$ denoted its error covariance.

3.2. Boundedness of the error covariances

From the derivation of $\Sigma(k)$ and $M(k+1)$, we can easily get that $\Sigma(k)$ and $M(k+1)$ are positive definite by sequential. In what follows, we will present upper bounds for these

covariance matrices. Firstly, we introduce the state transition matrix of the centralized system model, which can be expressed as

$$\Phi(k+1,k_0)=A(k)\Phi(k,k_0),\ \ k\geq k_0,\ \ \Phi(k_0,k_0)=I. \tag{10}$$

The observability matrix $O(k_0, d)$ is defined dually as

$$O(k_0,d)=\begin{bmatrix} H(k_0)\\ H(k_0+1)\Phi(k_0+1,k_0)\\ \vdots\\ H(d-1)\Phi(d-1,k_0)\end{bmatrix}, \tag{11}$$

where $d>k_0$, $k_0\in N_0$. From Eq. (10), we get that $\Phi(k+1,k_0)$ is invertible. Multiplying $\Phi^{-1}(d-1,k_0)$ on the right hand of Eq. (11), we obtain

$$\tilde{O}(k_0,d)=\begin{bmatrix} H(k_0)\Phi^{-1}(d-1,k_0)\\ H(k_0+1)\Phi^{-1}(d-1,k_0+1)\\ \vdots\\ H(d-1)\end{bmatrix},$$

which is called the modified observability matrix.

Secondly, we will use the following lemma to prove Theorem 1, which can be found in [23].

Lemma 1. *The pair $(A(k), H(k))$ is observable over $[k_0, d)$, if and only if the observability matrix $O(k_0, d)$ as defined in Eq. (11) has the full column rank.*

Theorem 1. *If $(A(k), H(k))$ is observable over $[k_0, d)$, and $A(k)$ is invertible, then there exists a constant $\gamma>0$, such that the covariance matrices $\Sigma(k)$ and $M(k+1)$, $k\geq d-1$ are bounded by*

$$\Sigma(k)<(1+\gamma)^{k-(d-1)}\Omega^{-1}(k_0,d), \tag{12}$$

$$M(k+1)<(1+\gamma)^{k-(d-1)}A(k)\Omega^{-1}(k_0,d)A^T(k)+R(k), \tag{13}$$

where

$$\Omega(k_0,d)=\sum_{l=k_0}^{d-1}(1+\gamma)^{-(d-1-l)}\Big(H(l)\Phi^{-1}(d-1,l)\Big)^T R_*^{-1}(l)H(l)\Phi^{-1}(d-1,l).$$

Proof. At each time instant k, since $\Sigma(k)>0$, $R(k)\geq 0$ and $A(k)$ is invertible, $A(k)\Sigma(k)A^T(k)$ is positive definite, and there exists a constant $\gamma>0$, such that

$$R(k)<\gamma A(k)\Sigma(k)A^T(k). \tag{14}$$

From Eqs. (9) and (14), we have

$$M(k+1)<(1+\gamma)A(k)\Sigma(k)A^T(k), \tag{15}$$

for each k. Based on Eqs. (7), (8) and (15), we get

$$\Sigma^{-1}(k)>\Xi(k)+\Big((1+\gamma)A(k-1)\Sigma(k-1)A^T(k-1)\Big)^{-1}, \tag{16}$$

where

$$\Xi(k) = H^T(k)R_*^{-1}(k)H(k).$$

Multiplying $A^T(k-1)$ and $A(k-1)$ on the left and right hands of Eq. (16), respectively, we obtain

$$A^T(k-1)\Sigma^{-1}(k)A(k-1) > A^T(k-1)\Xi(k)A(k-1) + ((1+\gamma)\Sigma(k-1))^{-1}. \tag{17}$$

Similarly, for the time instant $k-1$, we get

$$\begin{aligned} A^T(k-2)\Sigma^{-1}(k-1)A(k-2) > {} & A^T(k-2)\Xi(k-1)A(k-2) \\ & +((1+\gamma)\Sigma(k-2))^{-1}. \end{aligned} \tag{18}$$

It follows from Eqs. (17) and (18) that

$$\begin{aligned} &\Phi^T(k,k-2)\Sigma^{-1}(k)\Phi(k,k-2) > \Phi^T(k,k-2)\Xi(k)\Phi(k,k-2) \\ &+(1+\gamma)^{-1}A^T(k-2)\Xi(k-1)A(k-2) + (1+\gamma)^{-2}\Sigma^{-1}(k-2). \end{aligned}$$

Sequentially,

$$\begin{aligned} \Phi^T(k,k_0)\Sigma^{-1}(k)\Phi(k,k_0) > {} & \Phi^T(k,k_0)\Xi(k)\Phi(k,k_0) \\ & +(1+\gamma)^{-1}\Phi^T(k-1,k_0)\Xi(k-1)\Phi(k-1,k_0) \\ & +\cdots+(1+\gamma)^{-(k-k_0-1)}\Phi^T(k_0+1,k_0)\Xi(k_0+1) \\ & \times\Phi(k_0+1,k_0) + (1+\gamma)^{-(k-k_0)}\Sigma^{-1}(k_0). \end{aligned}$$

Since

$$\Sigma^{-1}(k_0) = H^T(k_0)R_*^{-1}(k_0)H(k_0) + M^{-1}(k_0),$$

and $M(k_0) > 0$, $k_0 \in N_0$, we obtain

$$\Phi^T(k,k_0)\Sigma^{-1}(k)\Phi(k,k_0) > \sum_{l=k_0}^{k}(1+\gamma)^{-(k-l)}\Big(H(l)\Phi(l,k_0)\Big)^T R_*^{-1}(l)H(l)\Phi(l,k_0).$$

Multiplying $(\Phi^{-1}(k,k_0))^T$ and $\Phi^{-1}(k,k_0)$ on the left and right hands of the above inequality, respectively, we have

$$\Sigma^{-1}(k) > \sum_{l=k_0}^{k}(1+\gamma)^{-(k-l)}\Big(H(l)\Phi^{-1}(k,l)\Big)^T R_*^{-1}(l)H(l)\Phi^{-1}(k,l).$$

Since $k \geq d-1$ and $R_*(k)$ is positive definite, we get

$$\begin{aligned} \Sigma^{-1}(k) &> \sum_{l=k_0}^{d-1}(1+\gamma)^{-(k-l)}\Big(H(l)\Phi^{-1}(d-1,l)\Big)^T R_*^{-1}(l)H(l)\Phi^{-1}(d-1,l) \\ &= (1+\gamma)^{-(k-(d-1))}\Omega(k_0,d). \end{aligned}$$

Based on Lemma 1 and $A(k)$ is invertible, we can see that the modified observability matrix $\tilde{O}(k_0,d)$ also has the full column rank. As a result, $\Omega(k_0,d) > 0$. Then we obtain Eq. (12). From Eqs. (9) and (12), one has

$$M(k+1) < A(k)\Big((1+\gamma)^{-(k-(d-1))}\Omega(k_0,d)\Big)^{-1}A^T(k) + R(k),$$

which leads to Eq. (13). This completes the proof. □

In the next section, we will use the demonstration method of Theorem 1 to prove the boundedness of the distributed state estimator.

4. Distributed state estimation process

In this section, we will investigate the distributed state estimator, which generalizes and extends the algorithm proposed in [21].

4.1. Distributed state estimator

Notations $\tilde{x}_i^*(k)$ and $M_i^*(k)$ are the state prediction and its error covariance matrix of node i for time instant k. Define

$$\bar{\alpha}_i(k) = C_i^T(k)S_i^{-1}(k)Z_{i,i}(k) + (M_i^*(k))^{-1}\tilde{x}_i^*(k),$$
$$\bar{Q}_i(k) = C_i^T(k)S_i^{-1}(k)C_i(k) + (M_i^*(k))^{-1},$$
$$Z_{i,i}(k) = z_{i,i}(k) + C_i(k)\tilde{x}_i^*(k) - f_i(\tilde{x}_i^*(k)).$$

Algorithm 1. Distributed state estimation algorithm

Initialization: (1) At time instant $k \in \mathbb{N}_0$, each node i computes the local state estimation and its error covariance:

$$\breve{x}_i^j(k,0) = \bar{Q}_i^{-1}(k)\bar{\alpha}_i(k),$$
$$\breve{\Sigma}_i^j(k,0) = \bar{Q}_i^{-1}(k).$$

If $k = 0$, then $\tilde{x}_i^*(k)$ and $M_i^*(k)$ are replaced by $\bar{x}_i(0)$ and $P_i(0)$, respectively.

(2) Node i transmits the following information to node $j \in \mathcal{N}_i$:

$$\eta_i^j(k,0) = B_{i,j}(k)\breve{x}_i^j(k,0),$$
$$\Upsilon_i^j(k,0) = B_{i,j}(k)\breve{\Sigma}_i^j(k,0)B_{i,j}^T(k). \tag{19}$$

Main loop: At iteration $h \in \mathbb{N}$, and for each node i:

(1) Using the information $\eta_j^i(k,h-1)$ and $\Upsilon_j^i(k,h-1)$ received from node $j \in \mathcal{N}_i$, node i updates the edge information as follows:

$$y_j^i(k,h) = Z_{i,j}(k) - \eta_j^i(k,h-1),$$
$$D_j^i(k,h) = T_{i,j}(k) + \Upsilon_j^i(k,h-1), \tag{20}$$

where

$$Z_{i,j}(k) = z_{i,j}(k) + B_{i,j}(k)\tilde{x}_i^*(k) + B_{j,i}(k)\tilde{x}_j^*(k) - h_{i,j}(\tilde{x}_i^*(k), \tilde{x}_j^*(k)).$$

(2) Node i calculates the current state estimation and the associated covariance:

$$\hat{x}_i(k,h) = Q_i^{-1}(k,h)\alpha_i(k,h),$$
$$\Sigma_i(k,h) = Q_i^{-1}(k,h),$$

where

$$\alpha_i(k,h) = \bar{\alpha}_i(k) + \sum_{j\in\mathcal{N}_i} B_{i,j}^T(k)\left(D_j^i(k,h)\right)^{-1} y_j^i(k,h), \tag{21}$$

$$Q_i(k,h) = \bar{Q}_i(k) + \sum_{j\in\mathcal{N}_i} B_{i,j}^T(k)\left(D_j^i(k,h)\right)^{-1} B_{i,j}(k). \tag{22}$$

(3) Meanwhile, node i computes

$$\check{x}_i^j(k,h) = \left[Q_i^j(k,h)\right]^{-1}\alpha_i^j(k,h),$$
$$\check{\Sigma}_i^j(k,h) = \left[Q_i^j(k,h)\right]^{-1},$$

where

$$\alpha_i^j(k,h) = \bar{\alpha}_i(k) + \sum_{m\in\mathcal{N}_i/\{j\}} B_{i,m}^T(k)(D_m^i(k,h))^{-1} y_m^i(k,h),$$
$$Q_i^j(k,h) = \bar{Q}_i(k) + \sum_{m\in\mathcal{N}_i/\{j\}} B_{i,m}^T(k)(D_m^i(k,h))^{-1} B_{i,m}(k), \tag{23}$$

and transmits $\eta_i^j(k,h)$ and $\Upsilon_i^j(k,h)$ to node $j\in\mathcal{N}_i$.

Theorem 1 in [21] has proven that, if the partitioned network is acyclic, then the state estimate of node i at each time instant k converges after a finite number of iterations, i.e.,

$$\hat{x}_i(k,h) = \hat{x}_i^*(k),$$
$$\Sigma_i(k,h) = \Sigma_i^*(k), \tag{24}$$

for all $h\geq\Gamma_i$.

The prediction equations of node i can be written as

$$\tilde{x}_i^*(k+1) = A_i(k)\hat{x}_i^*(k) + G_i(k),$$
$$M_i^*(k+1) = A_i(k)\Sigma_i^*(k)A_i^T(k) + R_i(k). \tag{25}$$

4.2. Parameter identification

1. Identification of state-space parameters

It is shown in [22] that the Holt's 2-parameter linear exponential smoothing method is converted into Eq. (1) to compute $A_i(k)$ and $G_i(k)$, where the smoothing parameters are constants. [10] has pointed out that when a significant load or power change happens, the method proposed in [22] will lead to enormous prediction errors. So we employ the time-varying parameter model as follows:

$$\tilde{x}_{q_i}^*(k+1) = a_{q_i}(k) + b_{q_i}(k), \tag{26}$$
$$a_{q_i}(k) = \alpha_{q_i}(k)\hat{x}_{q_i}^*(k) + (1-\alpha_{q_i}(k))\tilde{x}_{q_i}^*(k),$$
$$b_{q_i}(k) = \beta_{q_i}(k)(a_{q_i}(k) - a_{q_i}(k-1)) + (1-\beta_{q_i}(k))b_{q_i}(k-1),$$

where $\tilde{x}_{q_i}^*(k+1)$ and $\hat{x}_{q_i}^*(k)$ are the q_ith components of the state prediction $\tilde{x}_i^*(k+1)$ and state estimation $\hat{x}_i^*(k)$ with $q_i = 1,\ldots,s_i$. $A_i(k)$ is a diagonal matrix and $G_i(k)$ is a vector, the elements of which can be written as

$$A_{q_i}(k) = \alpha_{q_i}(k)(1+\beta_{q_i}(k)),$$

$$G_{q_i}(k) = (1+\beta_{q_i}(k))(1-\alpha_{q_i}(k))\tilde{x}^*_{q_i}(k) - \beta_{q_i}(k)a_{q_i}(k-1) + (1-\beta_{q_i}(k))b_{q_i}(k-1).$$

Parameters $\alpha_{q_i}(k)$ and $\beta_{q_i}(k)$ are time-varying, which are designed by the following method. Firstly, at each time instant k, we choose three groups of values: central values $(\alpha^0_{q_i}(k), \beta^0_{q_i}(k))$, low values $(\alpha^L_{q_i}(k), \beta^L_{q_i}(k))$ and high values $(\alpha^H_{q_i}(k), \beta^H_{q_i}(k))$. The relationship among these values can be described as

$$\begin{cases} \alpha^L_{q_i}(k) = \alpha^0_{q_i}(k) - \Delta\alpha_{q_i}(k), & \alpha^H_{q_i}(k) = \alpha^0_{q_i}(k) + \Delta\alpha_{q_i}(k), \\ \beta^L_{q_i}(k) = \beta^0_{q_i}(k) - \Delta\beta_{q_i}(k), & \beta^H_{q_i}(k) = \beta^0_{q_i}(k) + \Delta\beta_{q_i}(k), \end{cases}$$

where $\Delta\alpha_{q_i}(k)$ and $\Delta\beta_{q_i}(k)$ are the given steps. Using these three groups of parameters, we obtain five combinations, which are

$$d^1_{q_i}(k) = \left(\alpha^0_{q_i}(k), \beta^0_{q_i}(k)\right),\ d^2_{q_i}(k) = \left(\alpha^L_{q_i}(k), \beta^L_{q_i}(k)\right),\ d^3_{q_i}(k) = \left(\alpha^H_{q_i}(k), \beta^H_{q_i}(k)\right)$$

$$d^4_{q_i}(k) = \left(\alpha^H_{q_i}(k), \beta^L_{q_i}(k)\right),\ d^5_{q_i}(k) = \left(\alpha^L_{q_i}(k), \beta^H_{q_i}(k)\right).$$

Secondly, based on each designed parameter $d^c_{q_i}(k)$, $c = 1, \ldots, 5$, Eq. (26) implies state prediction $\tilde{x}^c_{q_i}(k+1)$, respectively. Meanwhile, we can compute the prediction error and mean square error as follows:

$$e^c_{q_i}(k+1) = \tilde{x}^c_{q_i}(k+1) - \hat{x}^*_{q_i}(k), \tag{27}$$

$$\Phi^c(k+1) = \frac{1}{s_i}\sum_{q_i=1}^{s_i}(e^c_{q_i}(k+1))^2,$$

where $q_i = 1, \ldots, s_i$ and $c = 1, \ldots, 5$. Let

$$\Phi^{\min}(k+1) = \min_c\left\{\Phi^c(k+1)\right\}. \tag{28}$$

We choose $\tilde{x}^c_{q_i}(k+1)$ which satisfies Eq. (28) as $\tilde{x}^*_{q_i}(k+1)$, and the corresponding $d^c_{q_i}(k)$ is used to compute $A_{q_i}(k)$ and $G_{q_i}(k)$. In order to verify the stability of state prediction, parameters $\alpha_{q_i}(k)$ and $\beta_{q_i}(k)$ are strictly in the range of 0.01 and 0.99.

Based on the prediction error Eq. (27), this method adjusts the time-varying parameters of the system model on-line, which improves the precision of state prediction at each time instant.

2. Identification of measurement error covariances

Bad data conditions affect the innovation vector, and then worsen the performance of algorithm. [11] proposed a robust algorithm for dynamic state estimation, by incorporating a new weighting function to inhibit the effect of anomaly conditions and ensure the robustness of computation. In this paper, we also use this method on each control center in order to strengthen the robustness of the distributed algorithm. By using the local and edge measurements of each node i at each time instant k,

$$W^*_{i,i}(k) = W_{i,i}(k)\exp\left(-|z_{i,i}(k) - f_i(\tilde{x}^*_i(k))|\right),$$

$$W^*_{i,j}(k) = W_{i,j}(k)\exp\left(-|z_{i,j}(k) - h_{i,j}(\tilde{x}^*_i(k), \tilde{x}^*_j(k))|\right)$$

are constantly updated weighting functions. Furthermore, the measurement error covariances can be defined as

$$S_i(k) = \left(W^*_{i,i}(k)\right)^{-1}, \ T_{i,j}(k) = \left(W^*_{i,j}(k)\right)^{-1}. \tag{29}$$

When some measurements are considerably distorted at time k, $z_{i,i}(k)$ or $z_{i,j}(k)$ will significantly change. Eq. (29) can reduce the weighting functions and mitigate the measurement errors well, and the weighting functions will not diverge.

4.3. Main results of boundedness

We will prove $\Sigma^*_i(k)$ and $M^*_i(k+1)$ are bounded in the following two theorems. Firstly, Theorem 2 shows that $\Sigma^*_i(k)$ and $M^*_i(k+1)$ are positive definite matrices, which also verifies the feasibility of Algorithm 1.

Theorem 2. *At each time instant k, the error covariance matrices of state estimate and prediction of node i are positive definite, i.e.,*

$$\Sigma^*_i(k) > 0, \ M^*_i(k+1) > 0.$$

Proof. We will use the mathematical induction to prove this theorem. Based on the initial conditions, we get that $\bar{Q}_i(0) > 0$ and $\breve{\Sigma}^j_i(0,0) > 0$ for each node i. From Eqs. (19) and (20), we have $D^i_j(0,1) > 0$ and then we get that $Q_i(0,1)$ and $Q^j_i(0,1)$ are positive definite based on Eqs. (22) and (23), which also lead to $\Sigma_i(0,1) > 0$ and $\breve{\Sigma}^j_i(0,1) > 0$. Assume that this result is valid for $h > 1$, i.e., for each node i, we have $\Sigma_i(0,h) > 0$ and $\breve{\Sigma}^j_i(0,h) > 0$, $h > 1$. At the iteration $h+1$, Eqs. (19) and (20) imply $D^i_j(0,h+1) > 0$. Following from Eqs. (22) and (23) again, one has that $\Sigma_i(0,h+1)$ and $\breve{\Sigma}^j_i(0,h+1)$ are positive definite. When $h \geq \varepsilon_i$, Eq. (24) implies $\Sigma^*_i(0) > 0$, and then $M^*_i(1) > 0$ obtained from Eq. (25), since $A_i(k)$ is invertible and $R_i(k) \geq 0$.

Assume that this result is valid for the time instant $k > 0$, i.e., for each node i, we obtain that $\Sigma^*_i(k)$ and $M^*_i(k+1)$ are positive definite matrices for $k > 0$. At the time instant $k+1$, we can easily get $\bar{Q}_i(k+1) > 0$, which implies $\breve{\Sigma}^j_i(k+1,0) > 0$ for each node i. By Eqs. (19) and (20), we obtain $D^i_j(k+1,1)$ is a positive definite matrix. Then from Eqs. (22) and (23), one has $Q_i(k+1,1) > 0$ and $Q^j_i(k+1,1) > 0$, which imply that $\Sigma_i(k+1,1)$ and $\breve{\Sigma}^j_i(k+1,1)$ are positive definite. Following from the derivation in the above paragraph, we can also get $\Sigma^*_i(k+1) > 0$ and $M^*_i(k+2) > 0$. This completes the proof. □

Theorem 2 also proves that for the iteration h, $D^i_j(k,h)$ is positive definite at each time instant k.

Secondly, we will give upper bounds on the error covariance matrices. Eqs. (21) and (22) can be written as

$$\begin{aligned} \alpha_i(k,h) &= \bar{H}^T_i(k)\bar{R}^{-1}_i(k,h)Z_i(k,h) + (M^*_i(k))^{-1}\tilde{x}_i(k), \\ Q_i(k,h) &= \bar{H}^T_i(k)\bar{R}^{-1}_i(k,h)\bar{H}_i(k) + (M^*_i(k))^{-1}, \end{aligned} \tag{30}$$

which means that $\bar{H}_i(k)$ is composed by $C_i(k)$ and $B_{i,j}(k)$. The centralized matrix $H(k)$ admits a block partition of the form $H(k) = \left[H_1(k) \ldots \ H_N(k)\right]$. Each node $i \in \mathcal{V}$ knows $H_i(k)$, which can also be partitioned, and there exist some zero blocks in $H_i(k)$, which can be seen in Eq.

(6). Deleting zero blocks from $H_i(k)$, the obtained matrix is the above mentioned $\bar{H}_i(k)$. Specially, if $\mathcal{N}_i = \{1, \ldots, m_i\}$, then one has

$$\bar{H}_i(k) = \begin{pmatrix} C_i(k) \\ B_{i,1}(k) \\ \vdots \\ B_{i,m_i}(k) \end{pmatrix}, \quad Z_i(k,h) = \begin{pmatrix} Z_{i,i}(k) \\ y_1^i(k,h) \\ \vdots \\ y_{m_i}^i(k,h) \end{pmatrix},$$

$$\bar{R}_i(k,h) = \begin{bmatrix} S_i(k) & 0 & \cdots & 0 \\ 0 & D_1^i(k,h) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_{m_i}^i(k,h) \end{bmatrix}.$$

Actually, when the estimation error covariance $\Sigma_i(k,\ h)$ converges to $\Sigma_i^*(k)$, $\breve{\Sigma}_i^j(k,h)$ also converges, since there is no more information transmitted after $h \geq \Gamma_i$ steps at each time instant. This means that $D_i^j(k,h)$ is convergence, based on Eqs. (19) and (20). Define $\tilde{R}_i(k) = \bar{R}_i(k,h)$, for $h \geq \Gamma_i$. The state transition matrix of each node i can be described as

$$\Phi_i(k+1,k_0) = A_i(k)\Phi_i(k,k_0), \ \ k \geq k_0, \ \ \Phi_i(k_0,k_0) = I.$$

Theorem 3. *If $(A_i(k), H_i(k))$ is observable over $[k_0, d_i)$, and $A_i(k)$ is invertible, then there exists a constant $\gamma_i > 0$, such that the error covariance matrices $\Sigma_i^*(k)$ and $M_i^*(k+1)$, $k \geq d_i - 1$ are bounded by*

$$\Sigma_i^*(k) < (1+\gamma_i)^{k-(d_i-1)}\Omega_i^{-1}(k_0,d_i), \tag{31}$$

$$M_i^*(k+1) < (1+\gamma_i)^{k-(d_i-1)}A_i(k)\Omega_i^{-1}(k_0,d_i)A_i^T(k) + R_i(k), \tag{32}$$

where

$$\Omega_i(k_0,d_i) = \sum_{l=k_0}^{d_i-1}(1+\gamma_i)^{-(d_i-1-l)}\left(\bar{H}_i(l)\Phi_i^{-1}(d_i-1,l)\right)^T\tilde{R}_i^{-1}(l)\bar{H}_i(l)\Phi_i^{-1}(d_i-1,l).$$

Proof. At each time instant k, when the iteration $h \geq \Gamma_i$, one has

$$(\Sigma_i^*(k))^{-1} = \bar{H}_i^T(k)\tilde{R}_i^{-1}(k)\bar{H}_i(k) + (M_i^*(k))^{-1},$$

for each node i, based on Algorithm 1. Since $\Sigma_i^*(k) > 0$, $R_i(k) \geq 0$ and $A_i(k)$ is invertible, there exists a constant $\gamma_i > 0$, such that

$$M_i^*(k+1) \leq (1+\gamma_i)A_i(k)\Sigma_i^*(k)A_i^T(k).$$

The rest of the proof is similar to that of Theorem 1, and we can prove that Eqs. (31) and (32) are valid. So we omit it here. □

5. Simulation results

In this section, we will show the effectiveness of Algorithm 1 together with the parameter identification. In Section 5.1, we will describe the test systems and performance indices. The detailed simulation results are investigated in Section 5.2.

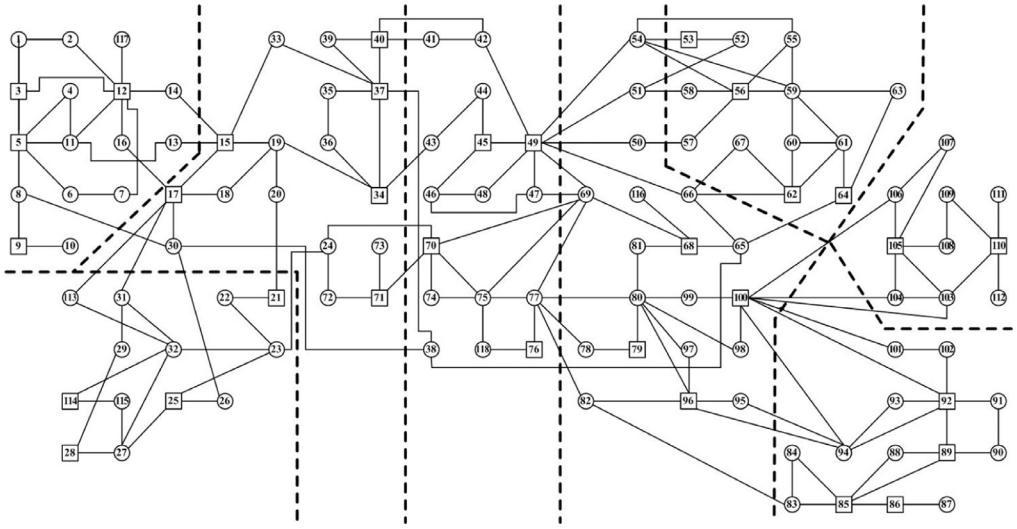

Fig. 1. Topological structure of the IEEE 118-bus system.

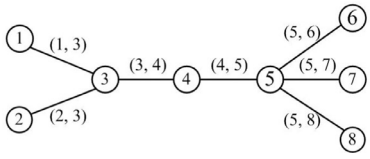

Fig. 2. The graph $\mathcal{G}$ depicting the partition of the 118-bus system.

5.1. Test system and description

The proposed algorithm of this paper is scalable to several large-scale networks in practice, such as power systems, traffic networks and multi-agent systems. In this section, the IEEE 118-bus system, whose specifications are given in [24], is used to test the performance of the distributed state estimator and the boundedness of the state and prediction error covariances. As shown in Fig. 1, the IEEE 118-bus system can be divided into 8 nonoverlapping subsystems, and all buses within a subsystem are treated as a whole to measure its local state and exchange information with its neighbors. Each node in Fig. 2 corresponds to a subsystem in Fig. 1, and we can see that this graph is acyclic. The placement of PMUs is designed by using the method presented in [6], and the buses installed with PMU measurements are marked as blocks in Fig. 1.

In the simulations, the considered power system is treated as a quasi-static system, which means that the state vector is varied slowly and smoothly. A linear trend between 1 and 3% along with a randomly distributed fluctuation of $\pm 4\%$ are added to the load curve, where we use different values for different buses. With regard to the local measurements of each node i, the actual measurement $z_{i,\,i}(k)$ is obtained by adding the true value with random normally distributed noise of 1% for SCADA measurements, or 0.2% for PMU measurements, while the edge measurements are obtained by adding true values with normally distributed noise of 2% standard deviation for power injections.

In order to respectively show the boundedness of state estimation and prediction error covariances, and the performance of the distributed state estimator compared with the centralized state estimator, the simulation results are divided into the following two parts:

(1) Boundedness: for the distributed state estimator, the sums of the traces of estimation and prediction error covariances are described as $\sum_{i=1}^{N} \mathrm{Tr}\{\Sigma_i^*(k)\}$ (TEC) and $\sum_{i=1}^{N} \mathrm{Tr}\{M_i^*(k+1)\}$ (TPC). Furthermore, the upper bounds of TEC and TPC are respectively denoted by BEC and BPC.

(2) Performance: the filtering index is evaluated by using the relative error between the state estimation $\hat{x}_i(k)$ and the true state vector $x_i^+(k)$, which is the following index

$$\varepsilon^f(k) = \frac{\Sigma_{i=1}^{N} |\hat{x}_i(k) - x_i^+(k)|}{\Sigma_{i=1}^{N} |x_i^+(k)|} \times 100\%,$$

where $\hat{x}_i(k)$ is obtained by the centralized and distributed state estimators.

The time step of tests is chosen as one second. We can see that the diameter Γ of the graph in Fig. 2 is 4, and Theorem 1 in [21] has proven that the local state estimate of each node converges after 4 steps at each time instant. Therefore, we let each distributed state estimator iterating 4 steps per time k. Here, we use Monte Carlo simulations to compute the error covariances of state estimation and prediction as well as the filtering index, and 1000 Monte Carlo runs are taken.

5.2. Test results

In order to validate the effectiveness of the distributed state estimation algorithm together with the parameter identification (IDSE) of this paper, we investigate it through the IEEE 118-bus system under the following test scenarios: normal operating condition, sudden load change and presence of bad data. Furthermore, the distributed state estimator (DSE), whose parameters are not identified on-line as in Section 4.2, is used here as a reference. Simulation results are presented and discussed as below.

Case 1. Normal operation condition

For the test system operated in this case, Fig. 3 describes the time evolutions of TEC, TPC and their upper bounds BEC and BPC for IDSE. Fig. 4 shows the results of filtering indices of the centralized state estimator (CSE), DSE and IDSE. We can see that Figs. 3 and 4 contain three parts corresponding to the three test conditions and simulations of each part are carried out through 20 time samples. The first parts of Figs. 3 and 4 are tested on the normal operating condition. One can see from Fig. 3 that TEC and TPC are less strictly than BEC and BPC, respectively, which verifies the feasibility and stability of IDSE. Although the $\varepsilon^f(k)$ values of IDSE in Fig. 4 is larger than the ones of CSE at each time instant, the gap between them is not quite significant, and we compute that the average values of CSE and IDSE are 0.403% and 0.612%, respectively, which mean that the estimated precision of IDSE is marginally less than that of CSE. Furthermore, Fig. 4 also shows that the performance of IDSE is more accurate than DSE at all time samples, which supports the efficiency of the parameter identification.

Case 2. Sudden load change condition

In this case, the estimation performance of CSE, DSE and IDSE is investigated under sudden load change conditions. For the IEEE-118 bus system, the following scenarios are simulated:

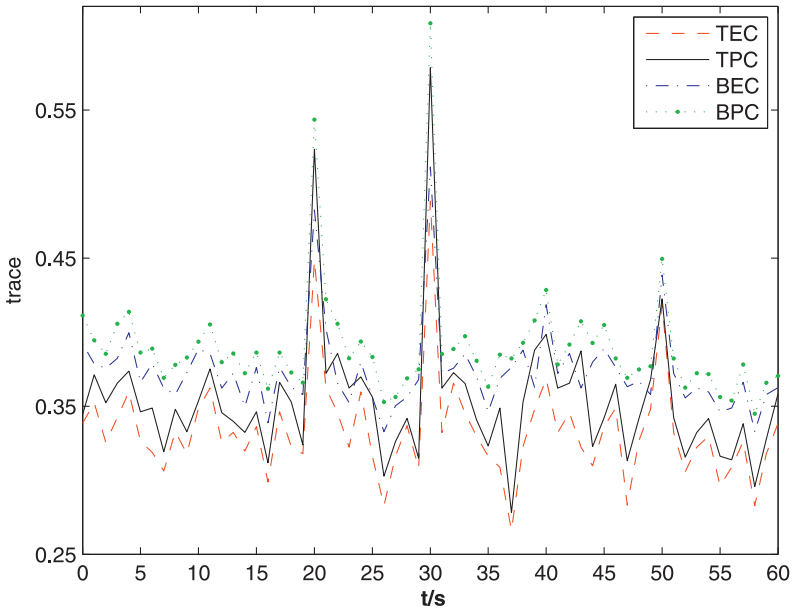

Fig. 3. Comparison of the trace of error covariance and upper bound.

1. 30% loads are cut on buses 8, 11, 12, 16, 19, 21, 23, 24, 30, 31, 37, 40, 49, 70, 77 at the 20th time sample;
2. 30% loads are increased on buses 51, 54, 56, 59, 65, 69, 89, 92, 94, 95, 100, 103, 106 at the 30th time sample.

We can see from the second part of Fig. 3 that although the state estimation and prediction error covariances of IDSE are affected by the sudden load changes, the degrees of variations are not drastic, and they are still positive definite and less than their upper bounds. The filtering indices of these methods are shown in the second part of Fig. 4. From the simulation results, the performance comparison of IDSE with CSE and DSE is similar to Case 1, and all the methods suffer the sudden load changes at the 20th and 30th time samples, the impact of which on IDSE is smaller than DSE with a few improvement on the maximal value. Besides, after sudden load occurs, the $\varepsilon^f(k)$ values of IDSE reduce quickly to the normal level, which also validate the feasibility of our proposed algorithm.

Case 3. Bad data condition

In this case, these methods are applied to the scenario where bad data is included in measurements. Two different conditions are investigated, which contain

1. one raw measurement error of 20% at the 40th time sample;
2. one raw measurement is mistaken as zero at the 50th time sample.

From the third part of Fig. 4, we can see that DSE is heavily affected by bad data, and the degree of variation of CSE is also drastic at the 40th and 50th time samples. Yet the maximal value of IDSE is much smaller than that of DSE. Accurately the maximal value of $\varepsilon^f(k)$ obtained by DSE decreases from 1.733% down to 0.742% by IDSE at the 50th time sample.

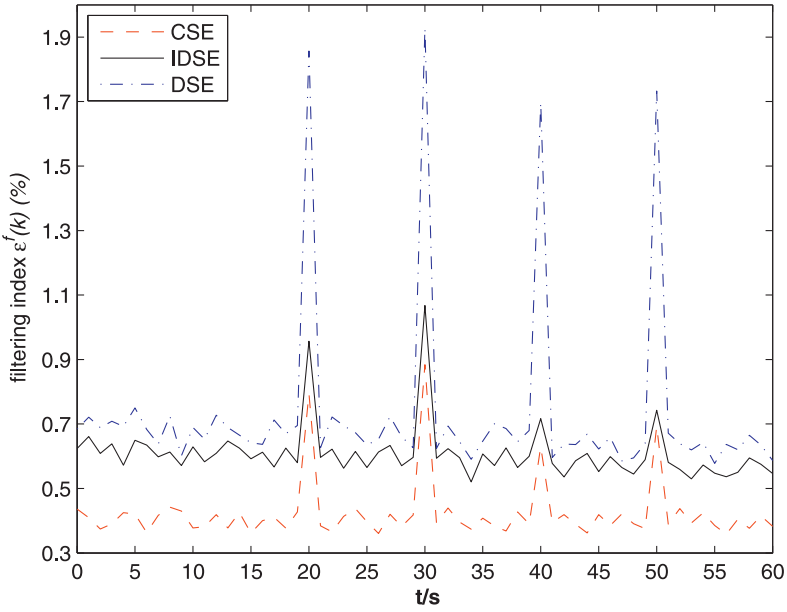

Fig. 4. Variations of filtering indices for different scenarios of different methods.

These mean that the overall influence caused by polluted measurements is well restrained by new weighting functions. Furthermore, the estimation performance of IDSE is better than DSE due to the identification of state-space parameters. Fig. 3 also shows that new weighting functions help to ensure the robustness of IDSE, since at the 40th and 50th time samples, TEC and TPC are little impacted by polluted measurements. All this reveals that IDSE owns excellent performance to the influence of bad data.

6. Conclusions

The distributed state estimation algorithm considered in this paper for large-scale complex systems can be applied to several practical domains, such as power systems, sensor networks and traffic networks. The main contribution of this paper is that the assumption that the measurement matrix must have full column rank in our previous study is relaxed to that each subsystem is observable, which can also verify the feasibility of the distributed algorithm. New weighting functions are used to restrain the influence of polluted measurements and model parameters are identified on-line to enhance the predicted precision. In order to approve the effectiveness of the distributed state estimation algorithm, it is tested through the IEEE 118-bus system under various operating conditions, which support the feasibility and efficiency of the proposed algorithm in applications to large-scale systems.

References

[1] W. Dargie, C. Poellabauer, Fundamentals of Wireless Sensor Networks: Theory and Practice, Wiley, Chichester, 2010.

[2] J. Feng, M. Zeng, Optimal distributed Kalman filtering fusion for a linear dynamic system with cross-correlated noises, Int. J. Syst. Sci. 43 (2) (2012) 385–398.

[3] J. Liang, Z. Wang, X. Liu, Distributed state estimation for discrete-time sensor networks with randomly varying nonlinearities and missing measurements, IEEE Trans. Neural Netw. 22 (3) (2011) 66–86.
[4] A. Abur, A.G. Exposito, Power System State Estimation: Theory and Implementation, Marcel Dekker, New York, 2004.
[5] J.D.L. Ree, V. Centeno, J.S. Thorp, A.G. Phadke, Synchronized phasor measurement applications in power systems, IEEE Trans. Smart Grid 1 (1) (2010) 20–27.
[6] X. Tai, D. Marelli, E. Rohr, M.Y. Fu, Optimal PMU placement for power system state estimation with random component outages, J. Electr. Power Energy Syst. 51 (2013) 35–42.
[7] Y.F. Diao, Z.Y. Lin, M.Y. Fu, H.S. Zhang, Localizability and distributed localization of sensor networks using relative position measurements, in: Proceedings of the IFAC Symposium on Large Scale Complex Systems: Theory and Applications, Shanghai, China, 2013, pp. 1–6.
[8] D. Marelli, M.Y. Fu, Distributed weighted least-squares estimation with fast convergence for large-scale systems, Automatica 51 (2015) 27–39.
[9] A.M.L.d. Silva, M.B.D.C. Filho, J.F. de Queiroz, State forecasting in electric power systems, IEE Proc. Gener. Transm. Distrib. 130 (5) (1983) 237–244.
[10] B.M. Zhang, S.Y. Wang, N.D. Xiang, Estimation and forcasting of real-time power system operation states, Proc. CSEE 11 (S1) (1991) 68–74.
[11] K.R. Shih, S.J. Huang, Application of a robust algorithm for dynamic state estimation of a power system, IEEE Trans. Power Syst. 17 (1) (2002) 141–147.
[12] J.L. Liang, J. Lam, Robust state estimation for stochastic genetic regulatory networks, Int. J. Syst. Sci. 41 (1) (2010) 47–63.
[13] X. Tai, Z.Y. Lin, M.Y. Fu, Y.Z. Sun, A new distributed state estimation technique for power networks, in: Proceedings of the American Control Conference, Washington, DC, 2013, pp. 3338–3343.
[14] L. Xie, D.H. Choi, S. Kar, H.V. Poor, Fully distributed state estimation for wide-area monitoring systems, IEEE Trans. Smart Grid 3 (3) (2012) 1154–1169.
[15] S. Kar, J.M.F. Moura, K. Ramanan, Distributed parameter estimation in sensor networks: nonlinear observation models and imperfect communication, IEEE Trans. Inf. Theory 58 (6) (2012) 3575–3605.
[16] H. Karimipour, V. Dinavahi, Parallel domain-decomposition-based distributed state estimation for large-scale power systems, IEEE Trans. Ind. Appl. 52 (2) (2015) 1265–1269.
[17] P. Alriksson, A. Rantzer, Distributed Kalman filtering using weighted averaging, in: Proceedings of the Seventeenth International Symposium on Mathematical Theory of Networks and Systems, Kyoto, Japan, 2006, pp. 1–6.
[18] J.L. Liang, Z.D. Wang, X.H. Liu, Distributed state estimation for discrete-time sensor networks with randomly varying nonlinearities and missing measurements, IEEE Trans. Neural Netw. 22 (3) (2011) 66–86.
[19] J.X. Feng, M. Zeng, Optimal distributed Kalman filtering fusion for a linear dynamic system with cross-correlated noises, Int. J. Syst. Sci. 43 (2) (2012) 385–398.
[20] A. Sharma, S.C. Srivastava, S. Chakrabarti, Multi-agent-based dynamic state estimator for multi-area power system, IET Gener. Transm. Distrib. 10 (1) (2016) 131–141.
[21] Y.B. Sun, M.Y. Fu, B.C. Wang, H.S. Zhang, A distributed MAP approach to dynamic state estimation with applications in power networks, in: Proceedings of the European Control Conference, Linz, Austria, 2015, pp. 235–240.
[22] S. Makridakis, S.C. Wheelwright, Forecasting Methods and Applications, Wiley, New York, 1978.
[23] G.X. Gu, Discrete-Time Linear Systems: Theory and Design with Applications, Springer-Verlag, New York, 2012.
[24] R. Christie, 118 bus power flow test case, 1993, http://www.ee.washington.edu/research/pstca/pf118/pg_tca118bus.htm.