2013 American Control Conference (ACC)
Washington, DC, USA, June 17-19, 2013

A New Distributed State Estimation Technique for Power Networks

Xin Tai[1], Zhiyun Lin[1,2], Minyue Fu[1,2], Yuanzhang Sun[3]

Abstract— **The paper develops a new distributed state estimation technique for power networks. This distributed state estimator aims to provide the same optimal estimate as the centralized one via a small amount information exchange between neighbors, which are physically in parallel with the partition of power networks. Moreover, we show that the distributed state estimator converges to the optimal estimate in a finite number of iterations depending on the size of the abstract graph. Both rigorous analysis and simulations are provided to show the correctness.**

I. Introduction

Power system state estimation as a core application of the on-line energy management system (EMS) has been well developed since it was introduced in early 1970's [12]. The traditional centralized state estimator is typically installed in a central control center in order to provide the best estimate of the state based on all the measurements collected by SCADA systems as well as a global power system model [5], [11]. However, the significant growth of demand for electricity has even stressed the importance of the development of renewable and distributed power generation along with decentralized energy storage technologies in the last decade. In this process, the so-called "smart grid" is considered to replace the traditional highly centralized power system at low voltage levels, where the architecture of the future power grid is fully distributed to distributed generation and distributed control [3]. Correspondingly, distributed state estimation techniques become more desirable in smart grids.

To this end, this paper investigates the distributed state estimation problem replacing a single central estimator by a set of local estimators distributed over a power network. In recent years, there have been developed quite a few distributed state estimation techniques [2]. However, the algorithms either still require a coordinator to cope with the interaction of local estimators [4], [6], [7], [10], [16], or they are not able to provide the optimal estimate as the centralized one or do not converge in finite time [1], [8], [9].

This paper develops a new distributed state estimation technique for power networks, which does not require a coordinator like those hierarchical estimation schemes ([4], [6], [7], [10], [16]) and is fully distributed. The execution of the algorithm only needs a small amount information exchange between neighbors via a data communication network that is topologically the same and physically in parallel with the partition of a power network, which is unlike [9] where a particular communication structure has to be defined. Among all these distributed estimation techniques, the most distinct feature of our proposed scheme is that each local estimator only estimates the local state in the region it is in charge of, rather than the global state of the entire power network. Thus, the topological information of the whole network and the global power network measurement model are not the requisite. Instead, each local estimator only needs to get access to the local measurements in its own region and also the boundary measurements related to its physical neighbors. By carrying out a local estimation based on the local measurements and updating its local estimation iteratively based on the boundary measurements and the exchanged estimates of its neighbors' states, our proposed scheme ensures a finite time convergence towards the same optimal estimate as obtained by a centralized estimator. By comparison, our algorithm results in the optimal estimate while [9] provides an approximate estimate with the estimate errors upper bounded. Moreover, our algorithm requires to transmit low dimensional estimation information of the boundary states between two neighbors while the algorithms in [9] need to transmit individual copy of the estimation information of the whole states. Our proposed algorithm is general, with applications not only on power systems but also other networked systems such as traffic systems, sensor networks, etc.

II. System Modeling and Problem Statement

Consider an interconnected system that can be represented by an acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ (see for example Fig. 1).

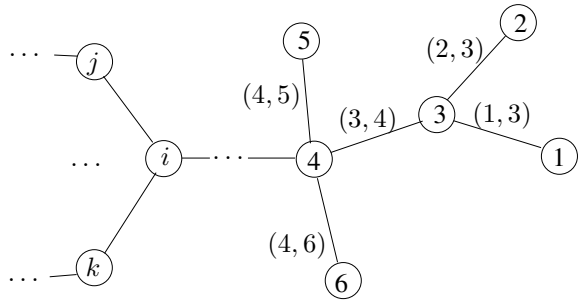


Fig. 1. An interconnected system represented by an acyclic graph.

Each node $i \in \mathcal{V} = \{1, 2, \cdots, n\}$ represents a subsystem of the interconnected system, whose state is $x_i \in \mathbb{R}^{s_i}$. The edges in the graph indicate the links between different

[1]School of Electrical Engineering and Computer Science, The University of Newcastle, NSW 2308, Australia.
[2]State Key Laboratory of Industrial Control Technology and College of Electrical Engineering, Zhejiang University, Hangzhou 310027, P. R. China.
[3]School of Electrical Engineering, Wuhan University, Wuhan 430072, P. R. China. (e-mail: xin.tai@uon.edu.au; linz@zju.edu.cn; minyue.fu@newcastle.edu.au; yzsun@whu.edu.cn)

subsystems. Lots of practical systems can be modeled in this way, including power grids, traffic networks and internet. Taking a power grid as an example, a node i refers to a subnetwork in a local area, while the edges refer to the transmission lines between different subnetworks.

Suppose on each node i in the graph, we have a node measurement, namely,

$$z_i = A_i x_i + w_i,\ i \in \mathcal{V}, \tag{1}$$

where $z_i \in \mathbb{R}^{q_i}$ is called the measurement on node i, $A_i \in \mathbb{R}^{q_i \times s_i}$ is the measurement matrix, and w_i stands for the measurement noise, which is white Gaussian with zero mean and covariance $\mathrm{cov}\,(\omega_i) = R_i$.

Moreover, on each edge $(i,j) \in \mathcal{E}$, we have an edge measurement for which the measurement relates to the states of both nodes associated with this edge. That is,

$$z_{(i,j)} = B_{ij} x_i + B_{ji} x_j + w_{(i,j)}, \tag{2}$$

where $z_{(i,j)} \in \mathbb{R}^{r_{ij}}$ is called the measurement on edge (i,j), B_{ij} and B_{ji} with appropriate dimensions are the measurement matrices, and $\omega_{(i,j)}$ is also white Gaussian with zero mean and covariance $\mathrm{cov}\left(\omega_{(i,j)}\right) = R_{(i,j)}$.

For the example of power grid, the node measurement represents the internal measurement for subnetworks where a measurement equipment is placed, and the edge measurement represents the tie-line measurement between two subnetworks.

In the paper, we have the following assumptions.

Assumption 1: The topological structure of graph $\mathcal{G}$ is acyclic.

Assumption 2: A_i is of full column rank for $i \in \mathcal{V}$.

Denote the aggregated state $x = (x_1, \ldots, x_n)$ and the aggregated measurement $z = (\cdots z_i \cdots z_{ij} \cdots)$. Then the measurement equation for the graph can be written as

$$z = Hx + \omega, \tag{3}$$

where $H = \begin{bmatrix} \cdots & H_i & \cdots & H_{(i,j)} & \cdots \end{bmatrix}^T$, with $H_i = [0\ A_i\ 0]$ and $H_{(i,j)} = [0\ B_{ij}\ 0\ B_{ji} 0]$. In the above 0 indicates a zero matrix with proper dimensions. The covariance of ω is of the block diagonal form, i.e.,

$$R := \mathrm{cov}\,(\omega) = \mathrm{diag}\begin{pmatrix} \cdots & R_i & \cdots & R_{(i,j)} & \cdots \end{pmatrix}.$$

Then based on the typical linear measurement equation (3), the traditional weighted least squares (WLS) estimation method can be utilized to calculate the optimal estimation value as

$$\hat{x}^* = \arg\ \min \left((z - Hx)^T R^{-1} (z - Hx) \right) \tag{4}$$

$$= \left(H^T R^{-1} H \right)^{-1} H^T R^{-1} z. \tag{5}$$

The estimation error covariance is given by

$$P^* = E\left((x - \hat{x}^*)(x - \hat{x}^*)^T \right) = \left(H^T R^{-1} H \right)^{-1}. \tag{6}$$

The above centralized WLS estimation method assumes the complete knowledge of the matrices H and R. In other words, all the measurements z need to be transmitted to a central monitor. Furthermore, Eq. (5) requires the inversion of $H^T R^{-1} H$. Hence a heavy burden of computation and communication is the great challenge for the centralized WLS estimation method applied in very large-scale power systems.

In the paper, we are seeking for light solutions for each node so that it is capable of estimating its own local state, without degrading any estimation performance compared with the centralized estimation scheme. In the proposed distributed estimation algorithm, a number of distributed local monitors are installed to replace the single central monitor. The local monitors need to finish the following two task in estimation: 1) **Local communication**: interchange the boundary estimates with its neighbor nodes; 2) **Local calculation**: run the local estimator based on its own local measurement and the boundary estimates received from its neighbor nodes.

The following notions and notations will be used throughout the paper. The set of neighbor nodes of i is defined as $\mathcal{N}_i$. Denote $\mathcal{N}_i/j$ the set of all the neighbor nodes of node i except j. The set of edges linking i and $j \in \mathcal{N}_i$, is defined as $\mathcal{M}_i$.

Now we are ready to introduce the problem formally.

Distributed WLS Problem. Design a local estimation scheme for each node i based on its own node measurement (1), its directly related edge measurement (2), and communication information from its neighbors, to find the unbiased estimates $\hat{x}_i, i \in \mathcal{V}$, minimizing the global objective function

$$J(\hat{x}) := (z - H\hat{x})^T R^{-1} (z - H\hat{x}) \tag{7}$$

where $\hat{x}$ is the aggregated state of local state estimates $\hat{x}_i$.

The motivation for the problem is clear, as for a distributed network over a large scale regions such as power networks, it is usually impossible or costly to collect all the measurements and then make a centralized estimation.

III. Distributed WLS Algorithm

In this section, we present a distributed WLS estimation scheme and then show that it offers the optimal solution as the centralized one.

We first present our proposed algorithm formally and next give detailed interpretations.

In the algorithm presented below, we use the form of information filter, for which we define $\alpha_i(k) = P_i(k)^{-1} \hat{x}_i(k)$ and $Q_i(k) = P_i(k)^{-1}$ for any $i \in \mathcal{V}$ and $k = 0, 1, \ldots$, called the information vector and information matrix respectively.

Algorithm 3.1: For each $i \in \mathcal{V}$,

1) Initialization ($k = 0$).

   a) Local estimation:

$$\alpha_i(0) = A_i^T R_i^{-1} z_i \tag{8}$$

$$Q_i(0) = A_i^T R_i^{-1} A_i. \tag{9}$$

b) Information transmitting to neighbor $j \in \mathcal{N}_i$:

$$\beta_i^j(0) = B_{ij} Q_i^{-1}(0)\, \alpha_i(0) \tag{10}$$

$$\Phi_i^j(0) = B_{ij} Q_i^{-1}(0)\, B_{ij}^T. \tag{11}$$

2) Iteration ($k = 1, 2, \ldots$).

a) Information receiving and update:
For any edge (i, j) with $j \in \mathcal{N}_i$, edge measurement update using the estimation from its neighbor:

$$y_{ji}(k) = z_{(i,j)} - \beta_j^i(k-1); \tag{12}$$

Corresponding noise covariance update:

$$S_{ji}(k) = R_{(i,j)} + \Phi_j^i(k-1). \tag{13}$$

b) Local estimation update:

$$\begin{aligned} \alpha_i(k) &= \alpha_i(0) + \sum_{j \in \mathcal{N}_i} B_{ij}^T S_{ji}^{-1}(k) y_{ji}(k) \\ Q_i(k) &= Q_i(0) + \sum_{j \in \mathcal{N}_i} B_{ij}^T S_{ji}^{-1}(k) B_{ij}. \end{aligned} \tag{14}$$

c) Information transmitting to neighbors $j \in \mathcal{N}_i$:
Compute

$$\begin{aligned} \alpha_i^j(k) &= \alpha_i(k) - B_{ij}^T S_{ji}^{-1}(k) y_{ji}(k) \\ Q_i^j(k) &= Q_i(k) - B_{ij}^T S_{ji}^{-1}(k) B_{ij}, \end{aligned} \tag{15}$$

and then transmit the following information to node j

$$\beta_i^j(k) = B_{ij} \left[Q_i^j(k) \right]^{-1} \alpha_i^j(k) \tag{16}$$

$$\Phi_i^j(k) = B_{ij} \left[Q_i^j(k) \right]^{-1} B_{ij}^T. \tag{17}$$

First, each node i receives information transmitted from its neighbors, which is calculated via the estimate and corresponding estimation error covariance: (k indicates the iteration step)

$$\beta_j^i(k) = B_{ji} \hat{x}_j(k) \tag{18}$$

$$\Phi_j^i(k) = B_{ji} P_j(k) B_{ji}^T, \tag{19}$$

and then updates the edge measurement using the received information according to the following formula:

$$z_{(i,j)} - \beta_j^i(k) = B_{ij} x_i + B_{ji}(x_j - \hat{x}_j(k)) + w_{(i,j)}. \tag{20}$$

The left-hand side of (20) is treated as a known measurement and the last two terms in the right-hand side of (20) is treated as the measurement noise with covariance $\Phi_j^i(k) + R_{(i,j)}$. Here we would like to point out that the transmitting information (β_j^i and Φ_j^i) relating to the edge measurement is of very small dimension and this indicates a very light communication load.

Second, each node runs a local estimator based on its own measurement (1) and the updated edge measurements (20) to estimate its own state $\hat{x}_i(k+1)$ and also provides the estimation error covariance $P_i(k+1)$.

Third, update the information going to send to the neighbour nodes. We take the information from node i to j, i.e., $\beta_i^j(k+1)$ and $\Phi_i^j(k+1)$, as an example. The local estimator on node i is ran again based on (1) and (20) (without using the information from j) to calculate $\hat{x}_i^{'}(k+1)$ and $P_i^{'}(k+1)$. Then

$$\beta_i^j(k+1) = B_{ij} \hat{x}_i^{'}(k+1) \tag{21}$$

$$\Phi_i^j(k+1) = B_{ij} P_i^{'}(k+1) B_{ij}^T \tag{22}$$

will be sent to node j.

Iterations continue until every estimate attains the optimal one. It will be shown that after a finite number of steps related to the "diameter" of the graph, all the estimates will attain the optimal values and further iterations will result in the same values. As a result, each node can terminate its estimation after a fixed number of steps.

In the proposed distributed estimation scheme, the resource being used only includes local calculation and local communication with neighbors. Local calculation occurs on each node by only estimating its own state, while communication is only required between neighbors in transmitting a very light data. Comparing to the centralized WLS estimation scheme that needs to collect the knowledge of all the nodes via communication and needs a powerful computer to do relatively heavy calculations, the proposed distributed scheme can save considerable computation burden and communication bandwidth and thus is particularly useful for large scale networks.

Next we present our main result, but firstly we introduce several notions from graph theory. The maximum distance of a path between node i and any other node j in the graph $\mathcal{G}$ is defined as the *eccentricity* of node i, denoted as ε_i. The maximum eccentricity is called the graph *diameter*, L. Also denote P_i^* ($i = 1, \ldots, n$) the block diagonal sub-matrix in P^* corresponding to the state x_i on node i, i.e., $P_i^* = E\left((x_i - \hat{x}_i^*)(x_i - \hat{x}_i^*)^T \right)$.

We start to analyze the proposed algorithm from the following preliminary result.

Lemma 3.1: Consider a distributed estimation scheme under Assumption 2. Let $\hat{x}_i$ be the resulting estimate and P_i be the associated estimation error covariance, $i = 1, \ldots, n$. If $P_i = P_i^*$ for all $i = 1, \ldots, n$, then $(\hat{x}_1, \cdots, \hat{x}_n)$ equals to $\hat{x}^*$ and minimizes $J(\hat{x})$.

Proof: Assuming the measurement noise has zero mean and diagonal covariance, it is known that if the weighted matrix is chosen as R^{-1}, the WLS estimation $\hat{x}^*$ is also the best linear unbiased estimation that minimizes the mean squared estimation error [13], i.e.,

$$\begin{aligned} \hat{x}^* &= \arg\min \left(E\left(\|x - \hat{x}\|^2 \right) \right) \\ &= \arg\min \left(\mathrm{Tr}\left(E\left((x - \hat{x})(x - \hat{x})^T \right) \right) \right) \end{aligned}$$

For a distributed estimation scheme generating an estimation $\hat{x}_i$ and the associated estimation error covariance P_i ($i = 1, \ldots, n$), if $P_i = P_i^*$ for all $i = 1, \ldots, n$, then we know that $\mathrm{Tr}(P^*) = \sum_{i=1}^n \mathrm{Tr}(P_i)$, which means

$$E(\|x - \hat{x}\|^2) = E(\|x - \hat{x}^*\|^2).$$

Note that $E(\|x - \hat{x}\|^2)$ is a convex function and that

Assumption 2 ensures the uniqueness of the global minimum for $J(\hat{x})$, so it follows that $(\hat{x}_1, \cdots, \hat{x}_n)$ equals to $\hat{x}^*$ and minimizes $J(\hat{x})$. □

Ensured by Lemma 3.1, we only need to show $P_i = P_i^*,\ i \in \mathcal{V}$ in the remaining of the paper for a distributed solution.

Theorem 3.1: Consider an acyclic graph $\mathcal{G}$ and the distributed WLS estimation algorithm. For each node i ($i = 1, \ldots, n$), if $k = \varepsilon_i$, then

$$P_i\left(k+l\right) = P_i^*,\ for\, all\, l \geq 0. \tag{23}$$

Notice that, from (23), the estimate on all nodes converge after $L = \max\left\{\varepsilon_i, i \in \mathcal{V}\right\}$ steps.

The proof of the theorem requires the following lemmas.

For an acyclic graph $\mathcal{G}$, when we cut off an edge (i, j), the graph $\mathcal{G}$ becomes two separated subgraphs, which do not connect each other. Denote the subgraph containing node i as $\mathcal{G}_{(i,j)}$. An example is given in Fig. 2.

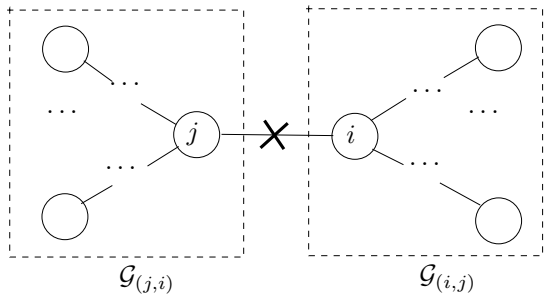


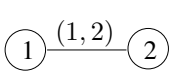
Fig. 2. Cutting off edge (i, j) results in two separated subgraphs.

Lemma 3.2: $(\alpha_i^j(k), Q_i^j(k))$ in the distributed WLS estimator for an acyclic graph $\mathcal{G}$ in step k is the distributed WLS estimate on node i for the subgraph $\mathcal{G}_{(i,j)}$ in step k.

Proof: From (14) and (15), we can get

$$\begin{aligned}\alpha_i^j\left(k\right) &= \alpha_i\left(0\right) + \sum_{l \in \mathcal{N}_i/j} B_{il}^T S_{li}^{-1}(k) y_{li}(k)\\ Q_i^j\left(k\right) &= Q_i\left(0\right) + \sum_{l \in \mathcal{N}_i/j} B_{il}^T S_{li}^{-1}(k) B_{il},\end{aligned}$$

Comparing with the formula (14), it is certain that $(\alpha_i^j(k), Q_i^j)$ is the local estimation update on node i for the induced subgraph $\mathcal{G}_{(i,j)}$ after removing edge (i, j). □



$$\begin{aligned}&1:\ z_1 = A_1 x_1 + w_1\\ &2:\ z_2 = A_2 x_2 + w_2\\ &(1,2):\ z_{(1,2)} = B_{12} x_1 + B_{21} x_2 + w_{(1,2)}\end{aligned}$$

Fig. 3. An interconnected system with two nodes.

Lemma 3.3: Consider an interconnected system of two nodes represented by $\mathcal{G}$ in Fig. 3. Under the distributed WLS scheme,

$$P_i\left(\varepsilon_i + l\right) = P_i^*,\ i = 1, 2,\ for\, all\, l \geq 0$$

where $\varepsilon_i = 1$.

Due to space limitations, the proof is omitted. The readers may refer to [14].

Before next lemma, we give a definition of the topological structure of a particular graph.

Definition 3.1: An acyclic graph is called a radial graph (shown in Fig. 4), if in this graph, one and only one node's eccentricity equals to 1, and the eccentricities of all the other nodes equal to 2. The node with 1 eccentricity is called the central node, which is indicated by c.

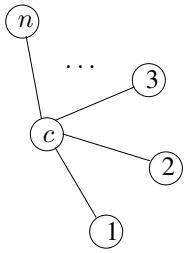


Fig. 4. An interconnected system represented by a radial graph.

Lemma 3.4: Consider an interconnected system represented by a radial graph $\mathcal{G}$ in Fig. 4 with the measurements given by

$$z_i = \quad A_i x_i + w_i,\ i \in \mathcal{V} \tag{24}$$

$$z_{(c,j)} = \quad B_{cj} x_c - B_{jc} x_j + w_{(c,j)},\ j \in \mathcal{N}_c. \tag{25}$$

Under the distributed WLS estimation scheme, we have

$$P_c\left(\varepsilon_c + l\right) = P_c^*,\ for\, all\, l \geq 0$$

where $\varepsilon_c = 1$.

Proof: Note that the local estimation update on node c is

$$\alpha_c\left(k\right) = \alpha_c\left(0\right) + \sum_{j \in \mathcal{N}_c} B_{cj}^T S_{jc}^{-1}(k) y_{jc}(k)$$

and

$$Q_c\left(k\right) = Q_c\left(0\right) + \sum_{j \in \mathcal{N}_c} B_{cj}^T S_{jc}^{-1}(k) B_{cj}.$$

They can be rewritten as

$$\alpha_c\left(k\right) = \alpha_c\left(0\right) + \begin{bmatrix} B_{c1} \\ B_{c2} \\ \vdots \\ B_{cn} \end{bmatrix}^T \widetilde{S} \begin{bmatrix} y_{1c}\left(0\right) \\ y_{2c}\left(0\right) \\ \vdots \\ y_{nc}\left(0\right) \end{bmatrix} \tag{26}$$

and

$$Q_c\left(k\right) = Q_c\left(0\right) + \begin{bmatrix} B_{c1} \\ B_{c2} \\ \vdots \\ B_{cn} \end{bmatrix}^T \widetilde{S} \begin{bmatrix} B_{c1} \\ B_{c2} \\ \vdots \\ B_{cn} \end{bmatrix} \tag{27}$$

where $\widetilde{S} = diag\left(\ S_{1c}^{-1}\left(0\right),\ \ S_{2c}^{-1}\left(0\right),\ \ \cdots,\ \ S_{nc}^{-1}\left(0\right)\ \right)$. Eqs. (26) and (27) imply that the radial graph can be treated as a two-node graph by combining nodes $1, \ldots, n$ into a single node. Thus, applying Lemma 3.3, the conclusion follows immediately. □

Now we come to prove Theorem 3.1.

Proof of Theorem 3.1: Consider an acyclic graph $\mathcal{G}$ and a node i in $\mathcal{G}$. Relabel the nodes if necessary so that the graph looks as in Fig. 5 where nodes 1 and 2 are considered to have the maximal distance away from node i.

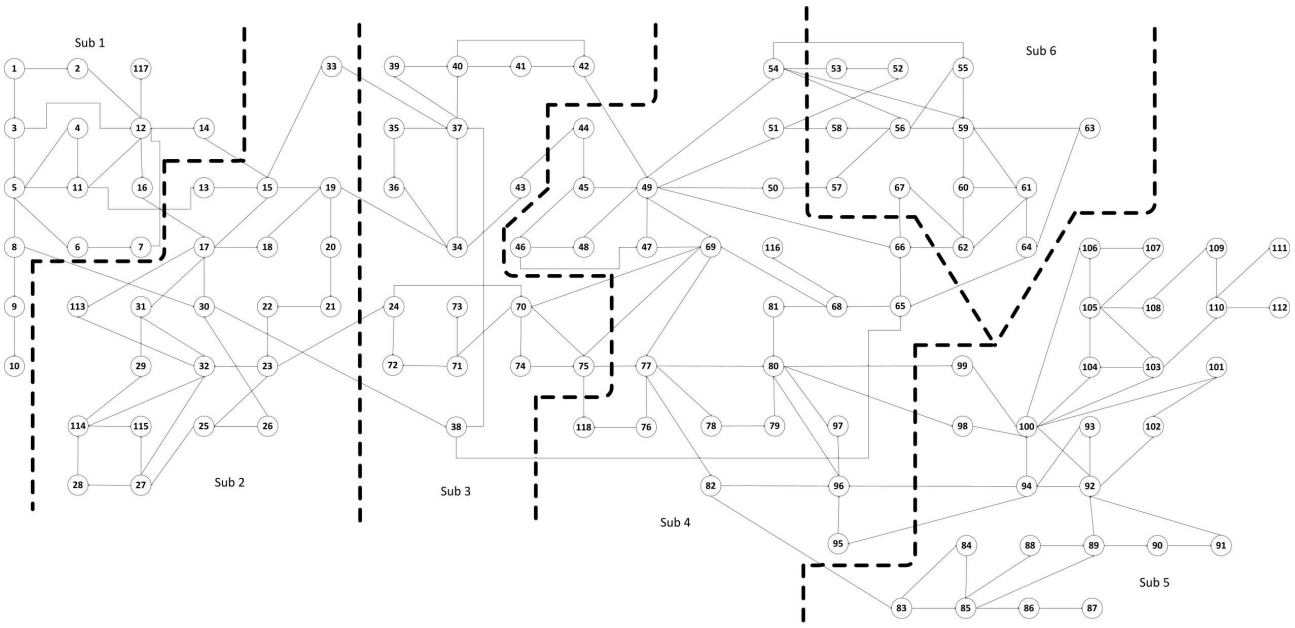


Fig. 6. Topological structure of the IEEE 118-bus system.

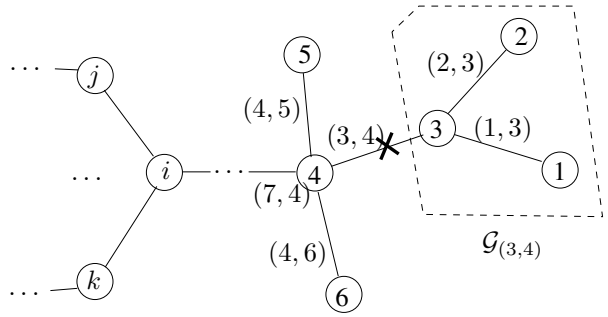


Fig. 5. Illustration for the proof of Theorem 3.1.

In other words, the distance from node 1 and 2 to node i equals to ε_i, while the distance from all other nodes to node i is no more than ε_i.

By cutting off edge $(3,4)$, it is known that $\mathcal{G}_{(3,4)}$ is a radial graph. Denote ε_3' the eccentricity of node 3 in the subgraph $\mathcal{G}_{(3,4)}$, which is 1 here. Then it follows from Lemma 3.4 and Lemma 3.2 that $\alpha_3^4(\varepsilon_3') = \alpha_3^4(\varepsilon_3'+1) = \cdots$ and $Q_3^4(\varepsilon_3') = Q_3^4(\varepsilon_3'+1) = \cdots$, satisfying $Q_3^4(\varepsilon_3') = (P_3^{sub})^{-1}$ and $\alpha_3^4(\varepsilon_3') = (P_3^{sub})^{-1}\hat{x}_3'$, where $\hat{x}_3'$ and P_3^{sub} are the optimal estimation on node 3 and its corresponding estimation error covariance for the graph $\mathcal{G}_{(3,4)}$ (namely, only using the measurements in $\mathcal{G}_{(3,4)}$).

Next, cut off edge $(4,7)$ and look at the subgraph $\mathcal{G}_{(3,4)}$. Note that $\alpha_3^4(1) = \alpha_3^4(2) = \cdots$ and $Q_3^4(1) = Q_3^4(2) = \cdots$, as we have just shown. Moreover, by the same way, it can be shown that $\alpha_5^4(0) = \alpha_5^4(1) = \cdots$ and $Q_5^4(0) = Q_5^4(1) = \cdots$, since the eccentricity of node 5 is 0 in the subgraph $\mathcal{G}_{(5,4)}$. Also, for the same reason we have $\alpha_6^4(0) = \alpha_6^4(1) = \cdots$ and $Q_6^4(0) = Q_6^4(1) = \cdots$. Thus according to the formula (15), $\alpha_4^7(\varepsilon_4') = \alpha_4^7(\varepsilon_4'+1) = \cdots$ and $Q_4^7(\varepsilon_4') = Q_4^7(\varepsilon_4'+1) = \cdots$, for which $\varepsilon_4' = 2$ is the eccentricity of node 4 in the subgraph $\mathcal{G}_{(4,7)}$. Moreover, as $Q_3^4(\varepsilon_3') = (P_3^{sub})^{-1}$ and $\alpha_3^4(\varepsilon_3') = (P_3^{sub})^{-1}\hat{x}_3'$, by checking the formula (15) at step $k=2$, it is equivalent to the first-step local estimate on node 4 based on the subgraph $\mathcal{G}_{(4,7)}$ with $\mathcal{G}_{(3,4)}$ being treated as a single node. Thus, applying Lemma 3.4 and Lemma 3.2, we know that $Q_4^7(\varepsilon_4') = (P_4^{sub})^{-1}$ and $\alpha_4^7(\varepsilon_4') = (P_4^{sub})^{-1}\hat{x}_4'$, where $\hat{x}_4'$ and P_4^{sub} are the optimal estimation on node 4 and its corresponding estimation error covariance for the graph $\mathcal{G}_{(4,7)}$.

Since the eccentricity of any node $l \in \mathcal{N}_i$ must be less than or equal to $\varepsilon_i - 1$, repeating this argument eventually leads to the conclusion that for any $l \in \mathcal{N}_i$,

$$\alpha_l^i(\varepsilon_i - 1) = \alpha_l^i(\varepsilon_i) = \cdots = (P_l^{sub})^{-1}\hat{x}_l'$$
$$Q_l^i(\varepsilon_i - 1) = Q_l^i(\varepsilon_i) = \cdots = (P_l^{sub})^{-1}$$

where $\hat{x}_l'$ and P_l^{sub} are the optimal estimation on node l and its corresponding estimation error covariance for the graph $\mathcal{G}_{(l,i)}$. Thus, from the formula (14), we could get $\alpha_i(\varepsilon_i) = \alpha_i(\varepsilon_i+1) = \cdots$ and $Q_i(\varepsilon_i) = Q_i(\varepsilon_i+1) = \cdots$, which implies $P_i(\varepsilon_i) = P_i(\varepsilon_i+1) = \cdots$.

Furthermore, since

$$\alpha_l^i(\varepsilon_i - 1) = (P_l^{sub})^{-1}\hat{x}_l' \text{ and } Q_l^i(\varepsilon_i - 1) = (P_l^{sub})^{-1}$$

for any $l \in \mathcal{N}_i$, the local estimate on node i at step $k = \varepsilon_i$ is equivalent to the first-step local estimate on node i for a radial graph with $\mathcal{G}_{(l,i)}$ being treated as a single node. Thus, it follows from Lemma 3.4 that $P_i(\varepsilon_i) = P_i^*$. □

IV. Examples and Simulations

The IEEE 118-bus system is utilized to test the proposed algorithm in this paper. The partition of the system is shown in Fig. 6, which can be generally described by the graph in Fig. 7. As mentioned in Section II, each subsystem is seen as a node and the links with tieline measurements

are referred as edges. In order to simplify the simulation, only PMU measurements which offer linear measurements are considered. The placement of PMU is optimized by the method proposed in [15]. We point out that the traditional SCADA measurements also can be utilized in Algorithm 3.1, after linearizing around the operating point [12].

Under Assumptions 1 and 2, we run Algorithm 3.1 on the system described above. The sum of the difference between the distributed and centralized estimates, i.e., $\sum_{i=1}^{n} |\hat{x}_i(k) - \hat{x}_i^*|$, and sum of the trace of the estimation error covariance of each subsystem, i.e., $\sum_{i=1}^{n} \mathrm{Tr}\{P_i(k)\}$, are utilized to compare the performances of distributed and centralized estimation methods. Here, we use a Monte Carlo simulation to compute the estimation error covariance, and 1000 Monte Carlo runs are taken.

Fig. 8 and Fig. 9 show the simulation results, which show that finite-time-convergence is guaranteed, and as indicated in Theorem 3.1, the algorithm converges after $k \geq L = 4$ steps.

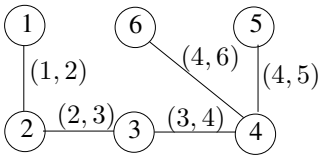


Fig. 7. The graph $\mathcal{G}$ describing the partition of the IEEE 118-bus system.

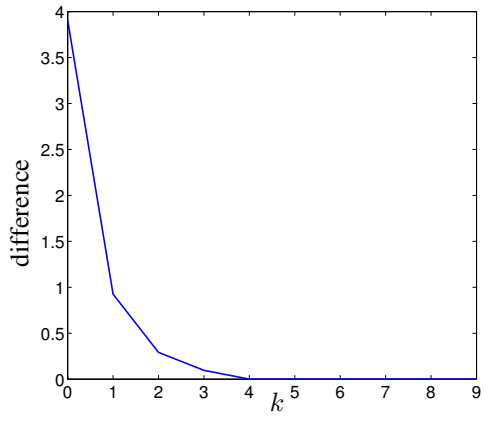


Fig. 8. Difference between the distributed and centralized estimates.

V. Conclusions

A novel distributed state estimation algorithm for decentralized local monitors to compute the WLS estimate of the power system state is proposed. The algorithm only requires local topological structure information, local measurements and low dimensional boundary information from neighbors. After a finite number of iterations, the local estimates can reach the same estimation values obtained via the centralized estimator. And the total iteration step equals to the diameter of the graph.

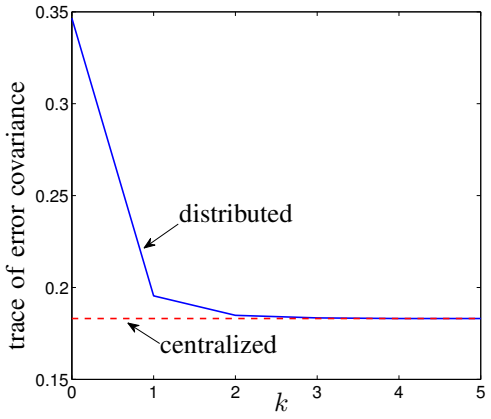


Fig. 9. Trace of the distributed and centralized estimation error covariances.

References

[1] A. J. Conejo, S. de la Torre, and M. Canas, "An optimization approach to multiarea state estimation," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 213–221, 2007.

[2] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, and T. V. Cutsem, "A taxonomy of multi-area state estimation methods," *Electric Power Systems Research*, vol. 81, pp. 1060–1069, 2011.

[3] A. Gomez-Exposito, A. Abur, A. de la Villa Jaen, and C. Gomez-Quiles, "A multilevel state estimation paradigm for smart grids," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 952–976, 2011.

[4] A. Gomez-Exposito and A. de la Villa Jaen, "Two-level state estimation with local measurement pre-processing," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 676–684, 2009.

[5] Z. Huang, R. Guttromson, J. Nieplocha, and R. Pratt, "Transforming power grid operations," *Scientific Computing*, vol. 24, no. 5, 2007.

[6] W. Jiang, V. Vittal, and G. T. Heydt, "A distributed state estimator utilizing synchronized phasor measurements," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 563–571, 2007.

[7] G. N. Korres, "A distributed multiarea state estimation," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 73–84, 2011.

[8] S. Y. Lin, "A distributed state estimator for electric power systems," *IEEE Transactions on Power Systems*, vol. 7, no. 2, pp. 551–557, 1992.

[9] F. Pasqualetti, R. Carli, and F. Bullo, "Distributed estimation via iterative projections with application to power network monitoring," *Automatica*, vol. 48, no. 5, pp. 747–758, 2012.

[10] M. Y. Patel and A. A. Girgis, "Two-level state estimation for multi-area power system," in *Proceedings of the 2007 IEEE Power Engineering Society General Meeting*, 2007, pp. 1–6.

[11] J. Ren and Z. Guo, "Period-oriented state estimation approach for power system operational control," in *Proceedings of the 2005 IEEE/PES Transmission and Distribution Conference and Exhibition*, 2005, pp. 1–5.

[12] F. C. Schweppe and J. Wildes, "Power system static-state estimation, part I, II, III," *IEEE Transactions on Power Apparatus and Systems*, vol. 89, pp. 120–135, 1970.

[13] M. K. Steven, *Fundamentals of Statistical Signal Processing*. Prentice Hall, 1993.

[14] X. Tai, D. Marelli, Z. Lin, and M. Fu, "A distributed approach to globally optimal state estimation in smart electricity networks," *submitted to IEEE Transactions on Power Systems*, 2012.

[15] X. Tai, D. Marelli, E. Rohr, and M. Fu, "Optimal PMU placement for power system state estimation with random communication packet losses," in *Proceedings of the 9th IEEE International Conference on Control and Automation*, 2011, pp. 444–448.

[16] X. Yang, X. Zhang, and S. Zhou, "Coordinated algorithms for distributed state estimation with synchronized phasor measurements," *Applied Energy*, vol. 96, no. 0, pp. 253–260, 2012.