Distributed Multi-Area State Estimation for Power Systems With Switching Communication Graphs

Jiexiang Wang and Tao Li, *Senior Member, IEEE*

***Abstract*—We consider distributed multi-area state estimation algorithms for power systems with switching communication graphs. The power system is partitioned into multiple geographically non-overlapping areas and each area is assigned with an estimator to give a local estimate of the entire power system's state. The inter-area communication networks are assumed to switch among a finite set of digraphs. Each area runs a distributed estimation algorithm based on consensus+innovations strategies. By the binomial expansion of matrix products, time-varying system and algebraic graph theories, we prove that all area's local estimates converge to the global least square estimate with probability 1 if the measurement system is jointly observable and the communication graphs are balanced and jointly strongly connected. Finally, we demonstrate the theoretical results by an IEEE 118-bus system.**

***Index Terms*—Power system, distributed state estimation, switching communication graph, convergence analysis.**

I. Introduction

STATE estimation plays an important role in monitoring and control of power systems. The Supervisory Control and Data Acquisition (SCDA) system collects the active and reactive power flow, bus injection power and voltage amplitude information, measured by remote terminal units and phasor measurement units (PMUs), to obtain the optimal estimate of the power system's state consisting of voltage phase angles and amplitudes at all buses, and then provides accurate information for the monitoring and control of power systems ([1]–[2]).

Centralized state estimation has several drawbacks. Firstly, since the power system is inevitably a large-scale network, the center bears great computation and communication burdens and thus is prone to failure. Secondly, in case that the center encounters a breakdown or attacks from enemies, the estimation task fails. Thirdly, it is usually difficult to collect information from the entire power system due to privacy protection. In distributed multi-area state estimation, the power system is partitioned into multiple geographically non-overlapping areas. Each area is assigned with an estimator to give a local estimate via information exchange with the estimators of neighbouring areas. The central unit is no longer needed. Thus, distributed state estimation is an ideal alternative to centralized state estimation thanks to its security, resilience and privacy protection ([3]–[4]). In addition, power industry deregulation also calls for state estimation in a distributed way to enable the monitoring of the overall power system ([5]).

In the early research on distributed state estimation for power systems, most literature was focused on incomplete distributed estimation ([6]–[9]), where a fusion center was used to coordinate all areas' local estimates. For fully distributed estimation of power systems, no central unit is needed. Conejo *et al.* [10] proposed a robust distributed state estimation algorithm and discussed the performance of the algorithm through several case studies. Xie *et al.* [11] put forward a distributed multi-area state estimation algorithm based on consensus+innovations strategies with a fixed and undirected communication graph. They proved that each area's local estimate converges to the global least square estimate with probability 1 if the graph is connected and the algorithm gain decays with the rate $1/(k+1)^{\tau}$ where $\tau \in (0, 1]$ and k is the iterative step. Wang *et al.* [12] focused on using composite optimization techniques for the distributed state estimation problem. Kekatos and Giannakis [13] developed a distributed state estimation algorithm based on the alternating direction method of multipliers. For the summaries of methods of state estimation for power systems, readers may refer to [14]–[17].

The above literature assumed a fixed communication graph. However, in reality, the inter-area communication networks are usually time-varying and not connected at each time step due to remote communication distances and bad external environments. In addition, to alleviate increasingly serious environment pollution and energy shortage, more and more renewable-based distributed generation units are integrated into power systems. The inherent intermittency of renewable energy will intensify the volatility of the power system's states ([18]–[19]). To monitor the states in real time, measurement devices have to acquire data more frequently ([15]). Then, more frequent inter-area communications are inevitable. This will aggravate intermittent link failures and consequently lead to switching inter-area communication graphs.

In this article, we consider distributed multi-area state estimation algorithms for power systems with switching communication graphs. The inter-area communication networks

Manuscript received November 20, 2019; revised March 18, 2020 and July 15, 2020; accepted August 17, 2020. Date of publication August 21, 2020; date of current version December 21, 2020. This work was supported by the National Natural Science Foundation of China under Grant 61977024 and the Basic Research Project of Shanghai Science and Technology Commission under Grant 20JC1414000. Paper no. TSG-01757-2019. *(Corresponding author: Tao Li.)*

Jiexiang Wang is with the School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China (e-mail: 920331447@qq.com).

Tao Li is with the Key Laboratory of Pure Mathematics and Mathematical Practice, School of Mathematical Sciences, East China Normal University, Shanghai 200241, China (e-mail: tli@math.ecnu.edu.cn).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2020.3018486

are assumed to switch among a finite set of digraphs. Each area runs a consensus+innovations type distributed estimation algorithm based on its previous estimate and the estimates of neighbouring areas and the local measurement. Among the aforementioned literature, Xie *et al.* [11] is most relevant to our work, which required the communication graph to be fixed. Compared with [11], the switching of communication graphs brings essential difficulties to the convergence analysis. By the binomial expansion method of matrix products and algebraic graph theory, we prove that all area's local estimates converge to the global least square estimate with probability 1 if the measurement system is jointly observable and the graphs are balanced and jointly strongly connected. This shows that the communication graphs are not necessarily strongly connected at each time step. Besides, the algorithm gains are more general and not restricted to specific forms. The contributions of this article are summarized below.

- As stated above, the inter-area communication is actually intermittent and not connected at each time step. Motivated by this, we propose a fully distributed state estimation algorithm with switching communication networks and shows that convergence can still be achieved even if the communication networks are not strongly connected at each time step. In [11], the communication graph is fixed, undirected and connected. Compared with [11], we consider a more practical scenario where the communication graph is directed and the communication between the areas of the system is intermittent. By the binomial expansion of matrix products, matrix and algebraic graph theories, we first prove that the product of the system matrices converges if the measurement system is jointly observable and the graphs are balanced and jointly strongly connected over a sequence of fixed-length intervals, then prove that all area's local estimates converge to the global least square estimate with probability 1.
- Both the innovation gain and consensus gain of the proposed algorithm are not restricted to special forms as in [11]. This expands the scope of the algorithm gains such that proper gains can be chosen to accelerate convergence of the algorithm, which is demonstrated in the simulation.

The rest of the paper is arranged as follows. In Section II, we give the related preliminaries and formulate the problem. In Section III, we give the distributed multi-area state estimation algorithm for power systems with switching communication graphs and the convergence conditions for the algorithm. In Section IV, we demonstrate the theoretical results by an IEEE 118-bus system. In Section V, we conclude our work. Notation and symbols: $\otimes$ denotes the Kronecker product; $\|A\|$ denotes the 2-norm of matrix A; A^T denotes the transpose of matrix A; $\mathbb{E}[\xi]$ denotes the mathematical expectation of random variable ξ; $\rho(A)$ denotes the spectral radius of matrix A; $\mathbf{1}_n$ denotes the n-dimensional column vector with all entries being one; denote $J_n = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$; $\mathbf{0}_{n\times m}$ denotes the $n\times m$-dimensional matrix with all entries being zero; I_n denotes the n-dimensional identity matrix; $\mathbb{R}^n$ denotes the n-dimensional real vector space; $\mathbb{Z}^+$ denotes the set of positive integers; for a real symmetric matrix A, $\lambda_2(A)$ denotes the second smallest eigenvalue and $\lambda_{\min}(A)$ denotes the minimum eigenvalue; $\lfloor x\rfloor$ denotes the largest integer less than or equal to x; $\lceil x\rceil$ denotes the smallest integer greater than or equal to x; $b_n = o(r_n)$ means $\lim_{n\to\infty}\frac{b_n}{r_n} = 0$, where $\{b_n, n\geq 0\}$ is a sequence of real numbers and $\{r_n, n\geq 0\}$ is a sequence of real positive numbers; for a sequence of n-dimensional matrices $\{Z(k), k\geq 0\}$, denote $\Phi_Z(j,i) = Z(j)\cdots Z(i)$ if $j\geq i$ and $\Phi_Z(j,i) = I_n$ if $j<i$.

II. Preliminaries and Problem Formulation

A. Preliminaries

Let the triplet $\mathcal{G} = \{\mathcal{V}, \mathcal{E}_\mathcal{G}, \mathcal{A}_\mathcal{G}\}$ be a weighted digraph, where $\mathcal{V} = \{1, 2, \ldots, N\}$ is the node set, $\mathcal{E}_\mathcal{G}$ is the edge set and $(j,i)\in\mathcal{E}_\mathcal{G}$ if and only if node j can send information to node i directly. Denote the neighborhood of node i by $\mathcal{N}_i = \{j\in\mathcal{V}|(j,i)\in\mathcal{E}_\mathcal{G}\}$. The weighted adjacency matrix of $\mathcal{G}$ is $\mathcal{A}_\mathcal{G} = [a_{ij}]\in\mathbb{R}^{N\times N}$, where $a_{ii} = 0$ and $a_{ij}>0 \Leftrightarrow j\in\mathcal{N}_i$, $i\in\mathcal{V}$. The in-degree of node i is $deg_{in}(i) = \sum_{j=1}^N a_{ij}$ and out-degree of node i is $deg_{out}(i) = \sum_{j=1}^N a_{ji}$. If $deg_{in}(i) = deg_{out}(i), i\in\mathcal{V}$, then $\mathcal{G}$ is balanced. An undirected graph, whose weighted adjacency matrix is symmetric, is always balanced. The Laplacian matrix of $\mathcal{G}$ is $\mathcal{L}_\mathcal{G} = \mathcal{D}_\mathcal{G} - \mathcal{A}_\mathcal{G}$, where $\mathcal{D}_\mathcal{G} = diag(deg_{in}(1), \ldots, deg_{in}(N))$ is the degree matrix of $\mathcal{G}$. The union digraph of $\mathcal{G}_1 = \{\mathcal{V}, \mathcal{E}_{\mathcal{G}_1}, \mathcal{A}_{\mathcal{G}_1}\}$ and $\mathcal{G}_2 = \{\mathcal{V}, \mathcal{E}_{\mathcal{G}_2}, \mathcal{A}_{\mathcal{G}_2}\}$ with the common node set is denoted by $\mathcal{G}_1+\mathcal{G}_2 = \{\mathcal{V}, \mathcal{E}_{\mathcal{G}_1}\cup\mathcal{E}_{\mathcal{G}_2}, \mathcal{A}_{\mathcal{G}_1}+\mathcal{A}_{\mathcal{G}_2}\}$. A sequence of edges $(i_1,i_2), (i_2,i_3), \ldots, (i_{k-1},i_k)$ is called a directed path from i_1 to i_k. If for all $i,j\in\mathcal{V}$, there exists a directed path from i to j, then $\mathcal{G}$ is strongly connected.

B. Problem Formulation

In a power system with $n+1$ buses, the power system's state is a vector $x\in\mathbb{R}^{2n+1}$ consisting of the voltage phase angles at n buses (except for the known voltage phase angle at the reference bus) and the voltage amplitudes at $n+1$ buses. In distributed multi-area state estimation, the power system is partitioned into N geographically non-overlapping areas and each area is assigned with an estimator which is only allowed to communicate with the estimators of neighbouring areas. The measurement of area i, denoted by $z_i\in\mathbb{R}^{m_i}$, consists of the active and reactive power flow, bus injection power and voltage amplitude information measured by remote terminal units and phasor measurement units in the i-th area. Generally, z_i is nonlinearly related to the power system's state x, given by

$$z_i = f_i(x) + v_i,\ \ i\in\mathcal{V}, \tag{1}$$

where $f_i(\cdot):\mathbb{R}^{2n+1}\to\mathbb{R}^{m_i}$ is the nonlinear measurement function of area i and v_i is the zero-mean measurement noise of area i with $R_i \triangleq \mathbb{E}(v_iv_i^T)>0$. After a DC power flow approximation ([20]), the state to be estimated reduces to a vector of voltage phase angles at n buses, denoted by $x_0\in\mathbb{R}^n$, and z_i is linearly related to x_0, given by

$$z_i = H_ix_0 + v_i,\ \ i\in\mathcal{V}. \tag{2}$$

Here, $H_i\in\mathbb{R}^{m_i\times n}$ is the observation matrix of area i, determined by the parameters of power lines. Denote $z =$

$[z_1^T, \ldots, z_N^T]^T$, $H = [H_1^T, \ldots, H_N^T]^T$ and $v = [v_1^T, \ldots, v_N^T]^T$. Rewrite (2) as

$$z = Hx_0 + v. \quad (3)$$

Remark 1: DC power flow is a commonly used tool for power system analysis. For high voltage modes in large-scale power grids, the voltage magnitude fluctuates at a much slower rate than the phase angle ([21]–[22]), then, the DC power flow approximation is reasonable. About the usefulness and accuracy of DC power flow approximation, readers may refer to [23].

We make the following assumption for the system (2).

Assumption 1: The measurement system (2) is jointly observable, i.e., $H^T H$ is invertible.

Remark 2: In state estimation, joint observability is a typical assumption. It shows that enough number of measurements is required to ensure the state of the entire power system to be observed and local observability in each area is not necessary. If the power system is not jointly observable, a measurement placement algorithm can be used to add measurements at proper locations to make Assumption 1 hold ([24]).

Denote $R = diag\{R_1, \ldots, R_N\}$. The global least square estimate for x_0 is given by

$$x_c \triangleq \arg\min_{x \in \mathbb{R}^n} [z - Hx]^T R^{-1} [z - Hx].$$

Let $[\partial((z - Hx)^T R^{-1}(z - Hx))]/\partial x = 2H^T R^{-1}(Hx - z) = 0$. If Assumption 1 holds, then

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

The centralized weighted least square estimator computes x_c by collecting the observation matrices, covariances and measurements of all areas, which lacks privacy protection and resilience. In the next section, we give the distributed multi-area state estimation algorithm with switching communication graphs, where no central unit is needed and the global least square estimate x_c can be asymptotically obtained by all areas' estimators via information exchange.

III. Algorithms and Convergence Analysis

We use the switching digraphs $\mathcal{G}(k) = \{\mathcal{V}, \mathcal{A}_{\mathcal{G}(k)}\}, k \geq 0$ to model the inter-area communication networks, where $\mathcal{A}_{\mathcal{G}(k)} = [a_{ij}(k)]_{1 \leq i,j \leq N}$ is the weighted adjacency matrix of $\mathcal{G}(k)$. Assume that the networks switch among a finite set of digraphs, that is, $\mathcal{G}(k) \in \{\mathcal{G}_1, \ldots, \mathcal{G}_q\}$, $k \geq 0$ with q being the total number of digraphs. For example, a network with time-varying link failures always switches among a finite set of topology graphs. Let $x_i(k)$ represent the estimate of x_c by the i-th area at the kth iteration. The initial estimate $x_i(0)$ is assumed to be deterministic, $i \in \mathcal{V}$. Each area collects the measurement z_i and then produce a local estimate by performing a distributed offline iterative procedure. In detail, at each $k \geq 0$, area i takes a weighted sum of its own estimate and the estimates from its neighbouring areas, and then incorporates the local measurement to update the estimate $x_i(k+1)$. Formally, the distributed multi-area state estimation algorithm for power systems with switching communication graphs is given by

$$\begin{aligned} x_i(k+1) = x_i(k) &+ a(k) H_i^T R_i^{-1} [z_i - H_i x_i(k)] \\ &+ b(k) \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \big[x_j(k) - x_i(k)\big], \\ &i \in \mathcal{V}, \ k \geq 0. \end{aligned} \quad (4)$$

Here, $a(k)$ and $b(k)$ are the innovation gain and the consensus gain, respectively, and $\mathcal{N}_i(k)$ denotes the set of neighbouring areas of area i at the kth iteration. For the algorithm gains, we make the following assumption.

Assumption 2: The sequences $\{a(k), k \geq 0\}$ and $\{b(k), k \geq 0\}$ are positive real sequences, which monotonically decrease to zero and satisfy $a(k) = o(b(k))$, $k \to \infty$ and $\sum_{k=0}^{\infty} a(k) = \infty$.

Before giving the following assumption on the communication graphs, we give the concept of joint strong connectivity. Recalling the definition of union graph given in Section II-A, joint strong connectivity of $\mathcal{G}(k)$ over an interval $[m, n]$, $m \leq n$ means that the union digraph of $\mathcal{G}(m), \ldots, \mathcal{G}(n)$ is strongly connected ([25]).

Assumption 3: The sequence $\{\mathcal{G}(k), k \geq 0\}$ is a sequence of balanced digraphs and jointly strongly connected over the sequence of intervals $[mh, (m+1)h - 1]$, $m \geq 0$ for some integer $h > 0$.

Remark 3: Assumption 3 means that the inter-area communication network is jointly strongly connected over a sequence of intervals with fixed length h. The constant h reflects the inter-area communication quality. If the communication network is strongly connected at each time step, then $h = 1$. The larger h is, the worse the communication quality is, and vice versa. Also, note that an undirected graph is always balanced. Hence, Assumption 3 is more realistic and easier to be satisfied than the conditions in [11] on the network graph.

We now give the main result, whose proof is put in Appendix A.

Theorem 1: For the algorithm (4), if Assumptions 1–3 hold, then all areas' local estimates converge to x_c with probability 1, i.e., $\mathbb{P}\{\lim_{k \to \infty} x_i(k) = x_c, \ i \in \mathcal{V}\} = 1$ where $\mathbb{P}\{\cdot\}$ denotes the probability of an event.

Remark 4: Theorem 1 shows that all areas' local estimates converge to the global least square estimate x_c with probability 1 provided the measurement model (2) is jointly observable, the communication graphs are balanced and jointly strongly connected and the algorithm gains satisfy Assumption 2. Xie *et al.* [11] studied the case of fixed communication graph, where the algorithm gains are restricted to the form $a(k) = a/(k+1)^{\tau_1}$ and $b(k) = b/(k+1)^{\tau_2}$ with $0 < \tau_1 \leq 1$, $0 < \tau_2 < \tau_1$ and a, b being positive constants. In Theorem 1, both the innovation gain and consensus gain of the proposed algorithm are not restricted to special forms. This expands the scope of the algorithm gains such that proper gains can be chosen to accelerate convergence of the algorithm, which will be demonstrated in Section IV.

Remark 5: For the case of AC state estimation, the measurement equation is represented by (1). To accommodate the nonlinearity of the measurement equation, the algorithm (4) is

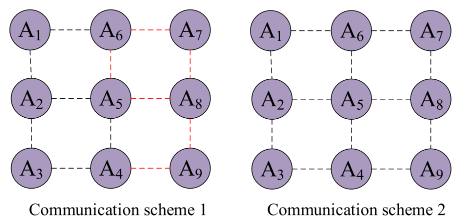


Fig. 1. Two communication schemes. In Communication Scheme 1, only the links represented by the black lines are active at odd time instants and only the links represented by the red lines are active at even time instants, i.e., $h = 2$. In Communication Scheme 2, all these links are active at each time instant.

modified as

$$\bar{x}_i(k+1) = \bar{x}_i(k) + b(k) \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k)\big[\bar{x}_j(k) - \bar{x}_i(k)\big] + a(k)\overline{H}_i^T(k,p)R_i^{-1}\big[z_i - f_i(\bar{x}_i(k))\big], \quad i \in \mathcal{V},\ k \geq 0. \tag{5}$$

Here, $\bar{x}_i(k)$ is the local estimate of the real grid state x and when $k \in [jp, (j+1)p-1], j \geq 0$, $\overline{H}_i(k,p) = \frac{\partial f_i(x)}{\partial x}|_{x=\bar{x}_i(jp)}$ where p is the updating period of the local Jacobian matrix. The algorithm (5) can be used to estimate both voltage phase angles and amplitudes at all buses simultaneously. For the case of switching communication networks, the length h of time interval over which the communication network is jointly strongly connected is only dependent on the communication topology and is not related to the updating period p of the local Jacobian matrix. In this article, as a preliminary research, we focus on the convergence conditions of DC state estimation algorithms, and remain the rigorous mathematical analysis of algorithm (5) as a future research topic.

IV. Numerical Simulations

We use an IEEE 118-bus system for the tests ([9], [26]). The system is partitioned into 9 areas $A_i, i = 1, \ldots, 9$, which contains 13, 13, 12, 13, 14, 13, 13, 14 and 13 buses, respectively. The state to be estimated is $x_0 = [\delta_2, \ldots, \delta_{118}]^T$ where δ_m represents the voltage phase angle at bus $m, m = 2, \ldots, 118$.

We consider undirected graphs with 0-1 weights. Communication Scheme 1 in Fig. 1 is used to model the intermittent communication among the areas, where only the links represented by the black lines are active at odd time instants and only the links represented by the red lines are active at even time instants, i.e., Assumption 3 holds with $h = 2$. Communication Scheme 2 in Fig. 1, where all communication links remain active during the whole process, is used for the comparison purpose in the later. Power injection measurements are placed at all generator buses and power flow measurements are placed at a subset of transmission lines. The total number of measurements is 176. The measurement configuration in each area is shown in Table I where Type 1 represents the power injection measurement and Type 2 represents the power flow measurement. By the physical connections among the buses and the distributions of the measurements, we know that $\lambda_{\min}(H^T H) = 0.00272$, i.e., Assumption 1 holds. The measurements noises are

TABLE I
Number of Measurements in Each Area

Areas	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9
Type 1	5	5	4	6	7	6	8	6	6
Type 2	14	16	13	12	10	15	16	17	10

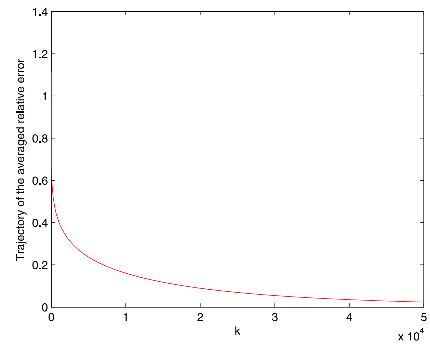


Fig. 2. Trajectory of the averaged relative error, $\frac{1}{9}\sum_{i=1}^{9} \frac{\|x_i(k) - x_c\|}{\|x_c\|}$.

TABLE II
Computation Time Corresponding to Fig. 2

Iterations	10000	20000	30000	50000
Averaged relative error	23%	13%	7%	3%
Elapsed time (s)	1.8826	3.9573	5.7030	9.5703

subject to Guassian distribution $\mathcal{N}(0, 0.01)$. We first choose $a(k) = \frac{0.03}{(k+1)^{0.1}}$ and $b(k) = \frac{0.1}{(k+1)^{0.01}}$. Then, Assumption 2 holds.

The results are numerically tested with an Intel(R) Core(TM) i5-4210U @ 1.7 GHz (4 GB RAM) computer using MATLAB. Fig. 2 is depicted with the trajectory of the averaged relative error $\frac{1}{9}\sum_{i=1}^{9} \frac{\|x_i(k) - x_c\|}{\|x_c\|}$. It shows convergence of all areas' local estimates to x_c even if areas A_7, A_8 and A_9 are isolated at odd time instants and areas A_1, A_2 and A_3 are isolated at even time instants. This simulation result validates the scalability of the proposed algorithm. The computational time of the algorithm for this case is recorded in Table II, showing that after about 9.57 seconds, the averaged relative error is 3%. Since the system is large and the observation matrix H is rather sparse, the number of iterations is high. Note that the communication time at each iteration is on the nanosecond level ([11]) and thus can be ignored. Hence, despite a high number of iterations, it is practical to apply the proposed algorithm into the power system.

To understand the effect of intermittent communication losses on the rate of convergence, we make a comparison between Communication Scheme 1 and Communication Scheme 2. Fig. 3 shows that the distributed estimation algorithm with intermittent communication losses has a slower convergence rate than the case without link failures.

We now give simulation results to illustrate the choices of the algorithm gains. The variance of the measurement noises is set to 1. Note that the role of $a(k)$ is to attenuate the measurement noises and the role of $b(k)$ is to drive all areas' local estimates to reach consensus. They only need to satisfy Assumption 2 and their choices are not system based and are independent of the operating condition of the system. In

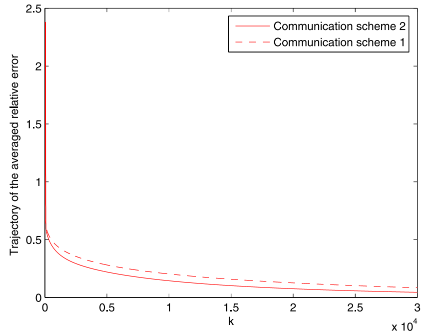


Fig. 3. Trajectories of the averaged relative errors under the communication schemes 1 and 2.

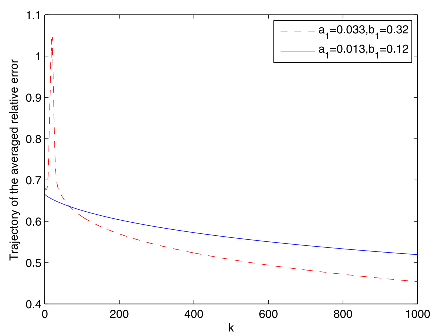


Fig. 4. Trajectory of the averaged relative error when $a(k) = \frac{a_1}{(k+1)^{0.1}}$ and $b(k) = \frac{b_1}{(k+1)^{0.01}}$.

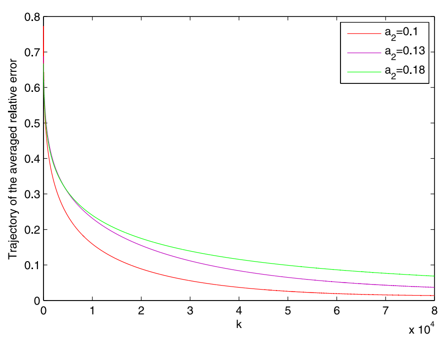


Fig. 5. Trajectory of the averaged relative error when $a(k) = \frac{0.03}{(k+1)^{a_2}}$ and $b(k) = \frac{0.1}{(k+1)^{0.01}}$.

fact, they have a significant effect on the convergence rate. Take the special form as an example: $a(k) = \frac{a_1}{(k+1)^{a_2}}$ and $b(k) = \frac{b_1}{(k+1)^{b_2}}$, where $a_1 > 0$, $b_1 > 0$, $0 < b_2 < a_2 \leq 1$. Fig. 4 shows that larger a_1 and b_1 leads to a faster convergence rate at the expense of more oscillations. Fig. 5 shows that the convergence speeds up as the decrease of a_2. Hence, the convergence rate is sensitive to the parameters a_1, b_1 and a_2. They should be carefully chosen.

We now demonstrate the benefit of generality of the algorithm gains. The innovation gain is changed as $a(k) = \frac{0.0033\ln(k+1)}{(k+1)^{0.1}}$, which satisfies Assumption 2 but is not allowed in [11]. The trajectory of the averaged relative error for this case is depicted in Fig. 6 and the computational time is recorded in Table III, which shows a faster convergence than

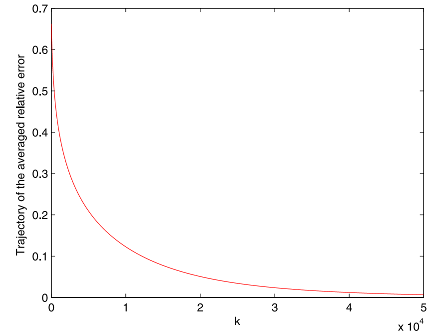


Fig. 6. Trajectory of the averaged relative error when $a(k) = \frac{0.0033\ln(k+1)}{(k+1)^{0.1}}$ and $b(k) = \frac{0.1}{(k+1)^{0.01}}$.

TABLE III
COMPUTATION TIME CORRESPONDING TO FIG. 6

Iterations	10000	20000	30000	50000
Averaged relative error	17%	7%	3%	0.8%
Elapsed time (s)	1.9324	3.9082	5.6701	9.3724

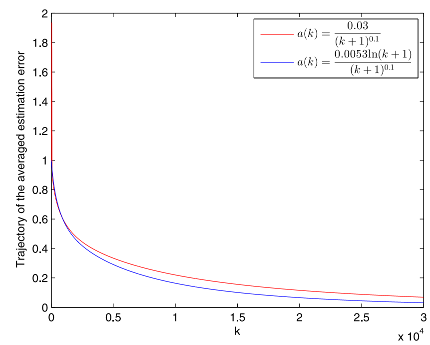


Fig. 7. Trajectory of the averaged relative error when $a(k) = \frac{0.0053\ln(k+1)}{(k+1)^{0.1}}$ and $a(k) = \frac{0.03}{(k+1)^{0.1}}$.

the case in Fig. 2 and Table II. For the case of fixed communication network as considered in [11], the benefit of generality of the algorithm gains is further confirmed in Fig. 7, where Communication Scheme 2 is considered. It shows that the convergence with $a(k) = \frac{0.0053\ln(k+1)}{(k+1)^{0.1}}$ is faster than the case with $a(k) = \frac{0.03}{(k+1)^{0.1}}$. In summary, Fig. 6 and Fig. 7 show that compared with the choices of algorithm gains in [11], even for the case of fixed communication graph, by Assumption 2, the algorithm gains can be chosen without satisfying the conditions in [11] such that a faster convergence of the algorithm is achieved.

We next give the simulation result showing the effect of the number of measurements $\sum_{i=1}^{9} m_i$ on convergence of the algorithm. The settings are the same as in the first paragraph except that the total number of measurements decreases to 137. The system is still jointly observable. Fig. 8 shows that the distributed estimation algorithm with less number of measurements has a much lower convergence rate. The computational time for this case is recorded in Table IV.

Finally, we demonstrate the applicability and performance of the AC state estimation algorithm (5). The real grid state

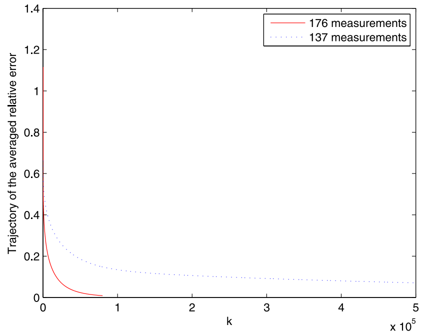


Fig. 8. Trajectory of the averaged relative error under different measurement configurations.

TABLE IV
COMPUTATION TIME OF THE CASE WITH 137 MEASUREMENTS

Iterations	10000	50000	100000	500000
Averaged relative error	50%	28%	20%	10%
Elapsed time (s)	1.5781	7.8519	15.9079	86.0216

to be estimated is given by $x = [\delta_2, \dots, \delta_{118}, V_1, \dots, V_{118}]^T$ where V_l is the voltage amplitude at bus l, $l = 1, \dots, 118$.

Denote the set of buses connected to bus l by $\mathcal{B}_l$. Denote

$$\begin{aligned} \mathcal{S}_i = \{&\text{branch } (l, m) | \text{bus } l \text{ is in area } i, m \in \mathcal{B}_l, m > l, \\ &\text{power flow measurements are placed at } (l, m)\}, \\ \mathcal{S}'_i = \{&\text{bus } l | \text{bus } l \text{ is in area } i, \text{injection measurements} \\ &\text{are placed at bus } l\}, \quad i = 1, \dots, 9. \end{aligned}$$

Let P^i_{lm} and Q^i_{lm} represent the active and reactive power flow measurements of branch $(l, m) \in \mathcal{S}_i$, respectively. Let $P^i_{l'}$ and $Q^i_{l'}$ represent the active and reactive injection measurements of bus $l' \in \mathcal{S}'_i$, respectively. Let $\overline{V}^i_q$ represent the voltage magnitude measurement of some bus q in area i. They are nonlinearly related to x ([8]):

$$\begin{cases} P^i_{lm} &= -V_l V_m (g_{lm} \cos(\delta_l - \delta_m) + b_{lm} \sin(\delta_l - \delta_m)) \\ &\quad + V_l^2 (g_{lm} + g_{slm}) + v^+_{ilm}, \\ Q^i_{lm} &= -V_l V_m (g_{lm} \sin(\delta_l - \delta_m) - b_{lm} \cos(\delta_l - \delta_m)) \\ &\quad - V_l^2 (b_{lm} + b_{slm}) + v^-_{ilm}, \quad (l, m) \in \mathcal{S}_i, \\ P^i_{l'} &= V_{l'}^2 g_{l'} + \sum_{m \in \mathcal{B}_{l'}} P_{l'm} + v^+_{il'}, \\ Q^i_{l'} &= -V_{l'}^2 b_{l'} + \sum_{m \in \mathcal{B}_{l'}} Q_{l'm} + v^-_{il'}, \quad l' \in \mathcal{S}'_i, \\ \overline{V}^i_q &= V_q + \tilde{v}_{iq}, \end{cases} \quad (6)$$

where $g_{lm}+jb_{lm}$ is the series admittance of branch (l, m) with j being the imaginary unit, $g_{slm}+jb_{slm}$ is the shunt admittance of branch (l, m), $g_{l'}+jb_{l'}$ is the shunt admittance connected at bus l' and v^+_{ilm}, v^-_{ilm}, $v^+_{il'}$, $v^-_{il'}$ and $\tilde{v}_{iq}$ are the measurement noises. For area $i = 1, \dots, 9$, take q =2, 15, 27, 34, 70, 51, 100, 89, 63, respectively. The measurement vector of area i, z_i, consists of P^i_{lm}, Q^i_{lm}, $P^i_{l'}$, $Q^i_{l'}$ and $\overline{V}^i_q$ in (6). The measurement configuration is shown in Table V where Type 3 represents the voltage magnitude measurement. The measurement noises all are independent and subject to Guassian distribution $\mathcal{N}(0, 0.01)$. The initial estimates for the voltage magnitudes at all buses are the rated voltage magnitude. The initial estimates for the voltage phases at all buses are zero. Set $p = 60$. Choose $a(k) = \frac{0.0016}{(k+1)^{0.1}}$ and $b(k) = \frac{0.2}{(k+1)^{0.01}}$. Fig. 9 is depicted with

TABLE V
NUMBER OF MEASUREMENTS IN EACH AREA

Areas	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9
Type 1	28	32	34	36	38	42	36	30	32
Type 2	12	14	12	16	14	16	14	14	12
Type 3	1	1	1	1	1	1	1	1	1

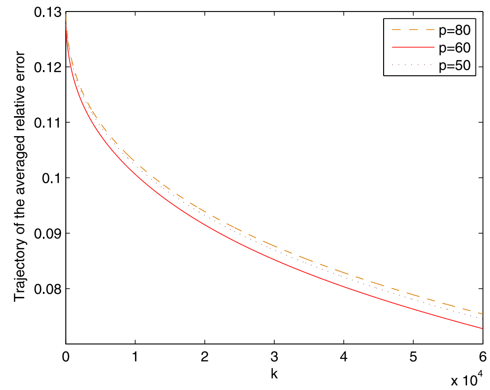


Fig. 9. Trajectory of the averaged relative error for the algorithm (5) under Communication Scheme 1.

the trajectories of the averaged relative error under different values of p, It is shown that the convergence slows down if p is too large or too small.

V. CONCLUSION

We have studied convergence of distributed multi-area state estimation algorithms for power systems with switching communication graphs. The graphs belong to a finite set of digraphs. The power system is partitioned into multiple geographically non-overlapping areas. Each area runs a distributed offline estimation algorithm of the consensus+innovations type. By the binomial expansion of matrix products and algebraic graph theory, we have proved that all area's local estimates converge to the global least square estimate with probability 1 if the measurement system is jointly observable and the communication graphs are balanced and jointly strongly connected. This shows that the communication graphs are not necessarily strongly connected at each time step. Besides, the algorithm gains are more general and not restricted to specific forms.

APPENDIX A
PROOF OF THEOREM 1

Lemma A.1 [27]: Let $\{s_1(k), k \geq 0\}$ and $\{s_2(k), k \geq 0\}$ be real sequences satisfying $0 < s_2(k) \leq 1$ and $\sum_{k=0}^{\infty} s_2(k) = \infty$. Assume that $\lim_{k\to\infty} \frac{s_1(k)}{s_2(k)}$ exists. Then $\lim_{k\to\infty} \sum_{i=1}^{k} s_1(i) \prod_{l=i+1}^{k} (1 - s_2(l)) = \lim_{k\to\infty} \frac{s_1(k)}{s_2(k)}$.

Let $x(k) = [x_1^T(k), \dots, x_N^T(k)]^T$, $\mathcal{H} = diag\{H_1, \dots, H_N\}$, $G(k) = b(k)\mathcal{L}_{\mathcal{G}(k)} \otimes I_n + a(k)\mathcal{H}^T R^{-1}\mathcal{H}$ and $F(k) = I_{Nn} - G(k)$, where $\mathcal{L}_{\mathcal{G}(k)}$ is the Laplacian matrix of $\mathcal{G}(k)$. By (4), we have

$$x(k+1) = F(k)x(k) + a(k)\mathcal{H}^T R^{-1} z. \quad (7)$$

Denote $\widehat{\mathcal{L}}_{\mathcal{G}(k)} = (\mathcal{L}_{\mathcal{G}(k)} + \mathcal{L}^T_{\mathcal{G}(k)})/2$. For any $m, h \in \mathbb{Z}^+$, denote

$$\lambda_m^h = \lambda_{\min}\left(\sum_{k=mh}^{(m+1)h-1}\left[\frac{b(k)}{a(k)}\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n + \mathcal{H}^T R^{-1}\mathcal{H}\right]\right).$$

For clarity, we decompose the proof of Theorem 1 into several intermediate lemmas, i.e., Lemmas A.2–A.8.

Lemma A.2: If Assumptions 2–3 hold, then $\inf_{m\geq 0}\lambda_m^h > 0$ if and only if Assumption 1 holds.

Proof: We first prove the necessity. Since $\inf_{m\geq 0}\lambda_m^h > 0$, we know from the definition of λ_m^h that $\sum_{k=mh}^{(m+1)h-1}(\frac{b(k)}{a(k)}\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n + \mathcal{H}^T R^{-1}\mathcal{H})$ is positive definite, $m \geq 0$. For any given $w \in \mathbb{R}^n$ with $w \neq \mathbf{0}_{n\times 1}$, let $x_w = \mathbf{1}_N \otimes w$. Then, by Assumption 3, we have $w^T(\sum_{i=1}^N H_i^T R_i^{-1} H_i)w = x_w^T \mathcal{H}^T R^{-1}\mathcal{H} x_w = \frac{1}{h}x_w^T \sum_{k=mh}^{(m+1)h-1}(\frac{b(k)}{a(k)}\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n + \mathcal{H}^T R^{-1}\mathcal{H})x_w > 0$. Noting that w is an arbitrary nonzero vector, then $\sum_{i=1}^N H_i^T R_i^{-1} H_i$ is positive definite, implying Assumption 1.

Next, we prove the sufficiency. By Assumption 2, we have $\frac{b(k)}{a(k)} \to \infty$. Hence, there exists a positive constant c_1 such that

$$\inf_{m\geq 0}\lambda_m^h \geq \inf_{m\geq 0}\lambda_{\min}\left(\sum_{k=mh}^{(m+1)h-1}\left(c_1\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n + \mathcal{H}^T R^{-1}\mathcal{H}\right)\right). \tag{8}$$

For any given $x \in \mathbb{R}^{Nn}$, if $x = \mathbf{1}_N \otimes d$, $\exists\, d \in \mathbb{R}^n$ with $d \neq \mathbf{0}_{n\times 1}$, then we have from Assumption 1 that

$$\begin{aligned} x^T &\sum_{k=mh}^{(m+1)h-1}\left(c_1\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n + \mathcal{H}^T R^{-1}\mathcal{H}\right)x \\ &= hd^T\left(\sum_{i=1}^N H_i^T R_i^{-1} H_i\right)d > 0,\ m \geq 0; \end{aligned} \tag{9}$$

otherwise, let $x = [x_1^T, \ldots, x_N^T]^T$, $x_i \in \mathbb{R}^n$. Then, we must have $x_i \neq x_j$, $\exists\, i \neq j$. By Assumption 3, we have

$$\begin{aligned} x^T &\sum_{k=mh}^{(m+1)h-1}\left(c_1\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n + \mathcal{H}^T R^{-1}\mathcal{H}\right)x \\ &\geq c_1 x^T \sum_{k=mh}^{(m+1)h-1}\left(\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n\right)x > 0,\ m \geq 0. \end{aligned} \tag{10}$$

Since the set $\{\mathcal{G}(k), k \geq 0\}$ is finite, from (9)–(10), we have $\inf_{m\geq 0}\lambda_{\min}(\sum_{k=mh}^{(m+1)h-1}(c_1\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n + \mathcal{H}^T R^{-1}\mathcal{H})) = \min\{\lambda_{\min}(\sum_{k=mh}^{(m+1)h-1}(c_1\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n + \mathcal{H}^T R^{-1}\mathcal{H})), m \geq 0\} > 0$. By the above and (8), we get the conclusion. ■

Let $\underline{m}_k = \lfloor \frac{k}{h} \rfloor$ and $\overline{m}_k = \lceil \frac{k}{h} \rceil$ for any given $k, h \in \mathbb{Z}^+$.

Lemma A.3: If Assumptions 1–3 hold, then there exist positive constants c_2 and c_3 such that

$$\|\Phi_F(k_1, k_2)\| \leq c_2 \prod_{s=\overline{m}_{k_2}}^{\underline{m}_{k_1}} \sqrt{1 - a(sh)c_3},\ k_1 \geq k_2 \geq 0. \tag{11}$$

Proof: By the binomial expansion and the definition of $G(k)$,

$$\begin{aligned} &\left\|\Phi_F^T((m+1)h-1, mh)\Phi_F((m+1)h-1, mh)\right\| \\ &= \left\|\left(I_{Nn} - G^T(mh)\right)\cdots\left(I_{Nn} - G^T((m+1)h-1)\right)\right. \\ &\quad \left.\times\ (I_{Nn} - G((m+1)h-1))\cdots(I_{Nn} - G(mh))\right\| \\ &= \left\|I_{Nn} - \sum_{k=mh}^{(m+1)h-1}\left(G(k) + G^T(k)\right)\right. \\ &\quad \left. +\ M_2(m) + \cdots + M_{2h}(m)\right\| \\ &\leq \left\|I_{Nn} - \sum_{k=mh}^{(m+1)h-1}\left(G(k) + G^T(k)\right)\right\| \\ &\quad +\ \|M_2(m) + \cdots + M_{2h}(m)\|, \end{aligned} \tag{12}$$

where $M_i(m)$ represents the i-th order term of $\Phi_F^T((m+1)h - 1, mh)\Phi_F((m+1)h-1, mh)$, $i = 2, \ldots, 2h$. Notice that the 2-norm of a symmetric matrix is equal to its spectral radius. Then, by Assumption 2 and the definition of $G(k)$, we have

$$\begin{aligned} &\left\|I_{Nn} - \sum_{k=mh}^{(m+1)h-1}\left(G(k) + G^T(k)\right)\right\| \\ &\quad = 1 - \lambda_{\min}\left(\sum_{k=mh}^{(m+1)h-1}\left(G(k) + G^T(k)\right)\right), \\ &\quad m \geq m_1, \end{aligned} \tag{13}$$

for some integer $m_1 > 0$.

Since $\inf_{m\geq 0}\lambda_m^h > 0$ by Lemma A.2, from the definition of λ_m^h and Assumption 2, we have

$$\begin{aligned} &1 - \lambda_{\min}\left(\sum_{k=mh}^{(m+1)h-1}\left(G(k) + G^T(k)\right)\right) \\ &= 1 - \lambda_{\min}\left(\sum_{k=mh}^{(m+1)h-1} a(k)\left[2\frac{b(k)}{a(k)}\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n + 2\mathcal{H}^T R^{-1}\mathcal{H}\right]\right) \\ &\leq 1 - 2a((m+1)h)\inf_{m\geq 0}\lambda_m^h < 1,\quad m \geq m_1. \end{aligned} \tag{14}$$

It is direct from Assumption 2 that $\sup_{k\geq 0}\|G(k)\| \leq \sup_{k\geq 0}[b(k)\|\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n\| + a(k)\|\mathcal{H}^T R^{-1}\mathcal{H}\|] \leq c_4$ for some constant $c_4 > 0$. Let $\mathbb{C}_m^p$ represent the combinatorial number of choosing p elements from m elements, $m \geq p$. By termwise multiplication, it follows from the definitions of $M_i(m)$ and Assumption 2 that $\|M_i(m)\| \leq b^2(mh)\mathbb{C}_{2h}^i c_4^i$, $i = 2, \ldots, 2h$, $m \geq 0$. Then, $\|M_2 + \cdots + M_{2h}\| \leq b^2(mh)\sum_{i=2}^{2h}\mathbb{C}_{2h}^i c_4^i = b^2(mh)c_5$ with $c_5 = c_4^{2h} - 1 - 2hc_4$. This together with (12)–(14) and Assumption 2 yields

$$\begin{aligned} &\left\|\Phi_F^T((m+1)h-1, mh)\Phi_F((m+1)h-1, mh)\right\| \\ &\leq 1 - 2a((m+1)h)\inf_{m\geq 0}\lambda_m^h + b^2(mh)c_5 \\ &\leq 1 - a((m+1)h)c_6,\ m \geq m_2, \end{aligned} \tag{15}$$

for some positive integer $m_2 \geq m_1$ and constant $c_6 > 0$. Since $\|\Phi_F((m+1)h - 1, mh)\| =$

$\sqrt{\|\Phi_F^T((m+1)h-1, mh)\Phi_F((m+1)h-1, mh)\|}$, by (15), $\|\Phi_F((m+1)h-1, mh)\| \leq \sqrt{1-a((m+1)h)c_6}$, $m \geq m_2$. Let $\mu = \sup_{k\geq 0}\|F(k)\|$. By Assumption 2, we have

$$\begin{aligned}\|\Phi_F(k_1,k_2)\| &= \left\|F(k_1)\cdots F\big(\underline{m}_{k_1}h\big)\Phi_F\big(\underline{m}_{k_1}h-1, \overline{m}_{k_2}h\big)\right.\\ &\quad \left.\times F\big(\overline{m}_{k_2}h-1\big)\cdots F(k_2)\right\|\\ &\leq \mu^{2h}\left\|\Phi_F\big(\underline{m}_{k_1}h-1, \overline{m}_{k_2}h\big)\right\|\\ &\leq \mu^{2h}\left\|\Phi_F\big(\underline{m}_{k_1}h-1, \big(\underline{m}_{k_1}-1\big)h\big)\right\|\\ &\quad \times \cdots \times \left\|\Phi_F\big(\big(\overline{m}_{k_2}+1\big)h-1, \overline{m}_{k_2}h\big)\right\|\\ &\leq \mu^{2h}\prod_{s=\overline{m}_{k_2}}^{\underline{m}_{k_1}}\sqrt{1-a(sh)c_6},\\ &\quad k_1-k_2 \geq 3h-1, k_2 \geq m_2h. \end{aligned} \tag{16}$$

For the case of $k_1 - k_2 < 3h-1$, $k_2 \geq 0$, by Assumption 2,

$$\begin{aligned}\|\Phi_F(k_1,k_2)\| &\leq \mu^{3h}\frac{\prod_{s=\overline{m}_{k_2}}^{\underline{m}_{k_1}}\sqrt{1-a(sh)c_6}}{\prod_{s=\overline{m}_{k_2}}^{\underline{m}_{k_1}}\sqrt{1-a(sh)c_6}}\\ &\leq \frac{\mu^{3h}\prod_{s=\overline{m}_{k_2}}^{\underline{m}_{k_1}}\sqrt{1-a(sh)c_6}}{(1-a(0)c_6)^{3h/2}}.\end{aligned} \tag{17}$$

Let $c_2 = \max\{[\mu/\sqrt{1-a(0)c_6}]^{3h}, \mu^{2h}\}$. Combining (16) and (17) gives (11). ■

Lemma A.4: For the algorithm (4), if Assumptions 1–3 hold, then $\sup_{k\geq 0}\|x(k)\|<\infty$ a.s.

Proof: By (7), $x(k+1) = \Phi_F(k,0)x(0) + \sum_{i=0}^k a(i)\Phi_F(k,i+1)\mathcal{H}^TR^{-1}z$. Taking the 2-norm on this equation gives

$$\begin{aligned}\|x(k+1)\| &\leq \|\Phi_F(k,0)\|\|x(0)\|\\ &\quad + \sum_{i=0}^k a(i)\|\Phi_F(k,i+1)\|\left\|\mathcal{H}^TR^{-1}z\right\|.\end{aligned} \tag{18}$$

It follows from Assumption 2 that

$$\begin{aligned}\sum_{s=0}^{\infty}a((s+1)h) &\geq \frac{1}{h}\sum_{s=0}^{\infty}\sum_{i=(s+1)h}^{(s+2)h-1}a(i)\\ &= \frac{1}{h}\sum_{t=h}^{\infty}a(t) = \infty.\end{aligned} \tag{19}$$

By Lemma A.3, we have $\|\Phi_F(k,0)\| \leq c_2\prod_{s=0}^{\underline{m}_k}\sqrt{1-a(sh)c_3} \leq c_2\prod_{s=0}^{\underline{m}_k}e^{-\frac{c_3}{2}a(sh)} = c_2e^{-\frac{c_3}{2}\sum_{s=0}^{\underline{m}_k}a(sh)}$. By this relation and (19), we have

$$\lim_{k\to\infty}\|\Phi_F(k,0)\| = 0. \tag{20}$$

Let $\eta = \|\mathcal{H}^TR^{-1}z\|c_2$. Noting that $\sqrt{1-x} \leq 1 - \frac{1}{2}x$, $0 \leq x \leq 1$, by Lemma A.3, it follows that

$$\begin{aligned}&\sum_{i=0}^k a(i)\|\Phi_F(k,i+1)\|\left\|\mathcal{H}^TR^{-1}z\right\|\\ &\leq \eta\sum_{i=0}^k a(i)\prod_{s=\overline{m}_{i+1}}^{\underline{m}_k}\sqrt{1-a(sh)c_3}\\ &= \eta\sum_{q=0}^{\underline{m}_k}\sum_{j=qh}^{(q+1)h-1}a(j)\prod_{s=q+1}^{\underline{m}_k}\sqrt{1-a(sh)c_3}\\ &\quad - \eta\sum_{j=k+1}^{(\underline{m}_k+1)h-1}a(j)\\ &\leq \eta\sum_{q=0}^{\underline{m}_k}\sum_{j=qh}^{(q+1)h-1}a(j)\prod_{s=q+1}^{\underline{m}_k}\left(1-\frac{c_3}{2}a(sh)\right)\\ &\quad - \eta\sum_{j=k+1}^{(\underline{m}_k+1)h-1}a(j).\end{aligned} \tag{21}$$

By (19), Assumption 2 and [28, Lemma 4.2], we have

$$\begin{aligned}&\limsup_{k\to\infty}\sum_{q=0}^{\underline{m}_k}\sum_{j=qh}^{(q+1)h-1}a(j)\prod_{s=q+1}^{\underline{m}_k}\left(1-\frac{c_3}{2}a(sh)\right)\\ &\leq \limsup_{k\to\infty}\frac{2\sum_{j=\underline{m}_kh}^{(\underline{m}_k+1)h-1}a(j)}{c_3a\big(\underline{m}_kh\big)}\\ &\leq \limsup_{k\to\infty}\frac{2ha\big(\underline{m}_kh\big)}{c_3a\big(\underline{m}_kh\big)} = \frac{2h}{c_3}.\end{aligned} \tag{22}$$

Since $h \leq (\underline{m}_k+1)h - k \leq 2h$, we have by Assumption 2 that $\lim_{k\to\infty}\sum_{j=k+1}^{(\underline{m}_k+1)h-1}a(j) = 0$. By the above, (21)–(22) and $\eta < \infty$ a.s., we have $\limsup_{k\to\infty}\sum_{i=0}^k a(i)\|\Phi_F(k,i+1)\|\|\mathcal{H}^TR^{-1}z\| < \infty$ a.s. Then, by (18) and (20), the lemma is proved. ■

Lemma A.5: If Assumptions 2–3 hold, then there exists a positive integer m_3 and a positive constant c_7 such that

$$\begin{aligned}&\inf_{m\geq 0}\lambda_{\min}\left[\left(\sum_{k=mh}^{(m+1)h-1}2b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)}\right)+J_N\right]\\ &\quad \geq c_7b((m+1)h),\ m \geq m_3.\end{aligned} \tag{23}$$

Proof: For any nonzero vector $x \in \mathbb{R}^N$, there exists a constant $q \in \mathbb{R}$ and a vector $\delta_x \in \text{span}\{\mathbf{1}_N\}^{\perp}$ such that $x = q\mathbf{1}_N + \delta_x$, where $\text{span}\{\mathbf{1}_N\}^{\perp}$ is the orthogonal complement space of $\text{span}\{\mathbf{1}_N\}$. Then, we have

$$\begin{aligned}&x^T\left(\sum_{k=mh}^{(m+1)h-1}2b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)}+J_N\right)x\\ &= (q\mathbf{1}_N+\delta_x)^T\left(\sum_{k=mh}^{(m+1)h-1}2b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)}+J_N\right)\\ &\quad \times (q\mathbf{1}_N+\delta_x)\\ &= \left(q\mathbf{1}_N^T+2\delta_x^T\sum_{k=mh}^{(m+1)h-1}b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)}\right)(q\mathbf{1}_N+\delta_x)\\ &= q^2N+2\delta_x^T\left(\sum_{k=mh}^{(m+1)h-1}b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)}\right)\delta_x\\ &\geq q^2N+2\lambda_2\left(\sum_{k=mh}^{(m+1)h-1}b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)}\right)\|\delta_x\|^2\end{aligned}$$

$$\geq \min\left\{1, 2\lambda_2\left(\sum_{k=mh}^{(m+1)h-1} b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)}\right)\right\} \times \left(q^2N + \|\delta_x\|^2\right) = \min\left\{1, 2\lambda_2\left(\sum_{k=mh}^{(m+1)h-1} b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)}\right)\right\}\|x\|^2, \quad (24)$$

where the second equality follows by Assumption 3 and the last equality follows by $\|x\|^2 = \|q\mathbf{1}_N+\delta_x\|^2 = q^2N+\|\delta_x\|^2$. Let $m_3 = \inf\{m'|2\lambda_2(\sum_{k=mh}^{(m+1)h-1} b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)}) < 1, \forall\ m \geq m' \geq 0\}$. Then, by (24), we have

$$x^T\left(\sum_{k=mh}^{(m+1)h-1} b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)} + J_N\right)x \geq 2b((m+1)h)\inf_{m\geq 0}\lambda_2\left(\sum_{k=mh}^{(m+1)h-1}\widehat{\mathcal{L}}_{\mathcal{G}(k)}\right)\|x\|^2, \quad m \geq m_3. \quad (25)$$

Since the set $\{\mathcal{G}(k), k \geq 0\}$ is finite, we have from Assumption 3 that $c_7 \triangleq \inf_{m\geq 0}\lambda_2(\sum_{k=mh}^{(m+1)h-1}\widehat{\mathcal{L}}_{\mathcal{G}(k)}) = \min\{\lambda_2(\sum_{k=mh}^{(m+1)h-1}\widehat{\mathcal{L}}_{\mathcal{G}(k)}), m \geq 0\} > 0$. Then, by (25), we have (23). ■

Let $F_1(k) = I_{Nn} - b(k)\mathcal{L}_{\mathcal{G}(k)} \otimes I_n$ and $W = I_{Nn} - J_N \otimes I_n$.

Lemma A.6: If Assumptions 2–3 hold, then there exist positive constants c_8 and c_9 such that

$$\left\|W\Phi_{F_1}(k_1, k_2)\right\| \leq c_8 \prod_{s=\overline{m}_{k_2}}^{\underline{m}_{k_1}} \sqrt{1 - b(sh)c_9},\ k_1 \geq k_2 \geq 0.$$

Proof: By Assumption 3, we have $J_N\mathcal{L}_{\mathcal{G}(k)} = \mathbf{0}_N,\ k \geq 0$. Then, by the binomial expansion, we have

$$\left\|\Phi_{F_1}^T((m+1)h-1, mh)W\Phi_{F_1}((m+1)h-1, mh)\right\| = \left\|W - \sum_{k=mh}^{(m+1)h-1} b(k)\left(\mathcal{L}_{\mathcal{G}(k)}^T \otimes I_n\right)W - W\sum_{k=mh}^{(m+1)h-1} b(k)\mathcal{L}_{\mathcal{G}(k)} \otimes I_n + \mathbb{M}(m)\right\| \leq \left\|W - 2\sum_{k=mh}^{(m+1)h-1} b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n\right\| + \|\mathbb{M}(m)\|, \quad (26)$$

where $\mathbb{M}(m)$ represents the sum of high-order terms of $\Phi_{F_1}^T((m+1)h-1, mh)W\Phi_{F_1}((m+1)h-1, mh)$. By Lemma A.5 and Kronecker product manipulations, we have

$$\left\|W - \sum_{k=mh}^{(m+1)h-1} 2b(k)\widehat{\mathcal{L}}_{\mathcal{G}(k)} \otimes I_n\right\| \leq 1 - b((m+1)h)c_7, \quad m \geq m_3. \quad (27)$$

From Assumption 2 and the definition of $\mathbb{M}(m)$, we have $\|\mathbb{M}(m)\| \leq b^2(mh)c_{10}$ for some positive constant c_{10}. Thus, we have from (26)–(27) that $\|\Phi_{F_1}^T((m+1)h-1, mh)W\Phi_{F_1}((m+1)h-1, mh)\| \leq 1-b((m+1)h)c_7+b^2(mh)c_{10}, m \geq m_3$. By this and Assumption 2, we have $\|\Phi_{F_1}^T((m+1)h-1, mh)W\Phi_{F_1}((m+1)h-1, mh)\| \leq 1 - b((m+1)h)c_{11},\ m \geq m_4$ for some positive constant c_{11} and integer $m_4 \geq m_3$. Then, it follows that

$$\left\|W\Phi_{F_1}((m+1)h-1, mh)\right\| \leq \sqrt{1 - b((m+1)h)c_{11}}, \quad m \geq m_4. \quad (28)$$

By Assumption 3, we have $WF_1(k) = F_1(k)W,\ k \geq 0$. Notice that $W^p = W, \forall\ p \in \mathbb{Z}^+$. Let $\overline{\mu} = \sup_{k\geq 0}\|F_1(k)\|$. Then, we have from (28) that

$$\begin{aligned}
&\left\|W\Phi_{F_1}(k_1, k_2)\right\| \\
&\leq \left\|W\Phi_{F_1}\left(k_1, \underline{m}_{k_1}h\right) \times \Phi_{F_1}\left(\underline{m}_{k_1}h-1, \overline{m}_{k_2}h\right)\Phi_{F_1}\left(\overline{m}_{k_2h}-1, k_2\right)\right\| \\
&= \left\|\Phi_{F_1}\left(k_1, \underline{m}_{k_1}h\right)W \times \Phi_{F_1}\left(\underline{m}_{k_1}h-1, \overline{m}_{k_2}h\right)\Phi_{F_1}\left(\overline{m}_{k_2h}-1, k_2\right)\right\| \\
&\leq \overline{\mu}^{2h}\left\|W\Phi_{F_1}\left(\underline{m}_{k_1}h-1, \overline{m}_{k_2}h\right)\right\| \\
&= \overline{\mu}^{2h}\left\|W\Phi_{F_1}\left(\underline{m}_{k_1}h-1, (\underline{m}_{k_1}-1)h\right) \times \cdots \times W\Phi_{F_1}\left(\left(\overline{m}_{k_2}+1\right)h-1, \overline{m}_{k_2}h\right)\right\| \\
&\leq \overline{\mu}^{2h}\prod_{s=\overline{m}_{k_2}}^{\underline{m}_{k_1}}\sqrt{1 - b(sh)c_{11}}, \quad k_1 - k_2 \geq 3h-1, k_2 \geq m_4h.
\end{aligned}$$

The case of $k_1 - k_2 < 3h-1, k_2 \geq 0$ is similar to (17). The proof is completed. ■

Denote $x_a(k) = \frac{1}{N}\sum_{i=1}^N x_i(k)$.

Lemma A.7: For the algorithm (4), if Assumptions 1–3 hold, then $\lim_{k\to\infty}\|x_i(k) - x_a(k)\| = 0$ a.s., $i \in \mathcal{V}$.

Proof: By (7) and the definition of $F_1(k)$, $x(k+1) = F(k)x(k) + a(k)\mathcal{H}^TR^{-1}z = F_1(k)x(k) + a(k)[\mathcal{H}^TR^{-1}z - \mathcal{H}^TR^{-1}\mathcal{H}x(k)] = \Phi_{F_1}(k, 0)x(0) + \sum_{i=0}^k a(i)\Phi_{F_1}(k, i+1)[\mathcal{H}^TR^{-1}z - \mathcal{H}^TR^{-1}\mathcal{H}x(i)]$. Premultiplying this equation by W gives

$$\|Wx(k+1)\| \leq \left\|W\Phi_{F_1}(k, 0)\right\|\|x(0)\| + \sup_{k\geq 0}\left\|\mathcal{H}^TR^{-1}z - \mathcal{H}^TR^{-1}\mathcal{H}x(k)\right\| \times \sum_{i=0}^k a(i)\left\|W\Phi_{F_1}(k, i+1)\right\|. \quad (29)$$

From Lemma A.6 and $\sqrt{1-x} \leq 1 - \frac{1}{2}x,\ 0 \leq x \leq 1$, we have $\|W\Phi_{F_1}(k, 0)\| \leq c_8\prod_{s=0}^{\underline{m}_k}\sqrt{1 - b(sh)c_9} \leq c_8e^{-\frac{c_9}{2}\sum_{s=0}^{\underline{m}_k}b(sh)}$. This together with $\sum_{s=0}^{\infty}b(sh) = \infty$ gives

$$\lim_{k\to\infty}\|W\Phi_{F_1}(k, 0)\| = 0. \quad (30)$$

By Lemma A.6 and Assumption 2, we have

$$\begin{aligned}
&\sum_{i=0}^k a(i)\left\|W\Phi_{F_1}(k, i+1)\right\| \\
&\leq c_8\sum_{i=0}^k a(i)\prod_{s=\overline{m}_{i+1}}^{\underline{m}_k}\sqrt{1 - b(sh)c_9} \\
&\leq c_8\sum_{i=0}^k a(i)\prod_{s=\overline{m}_{i+1}}^{\underline{m}_k}\left(1 - \frac{c_9}{2}b(sh)\right)
\end{aligned}$$

$$= c_8 \sum_{q=0}^{m_k} \sum_{j=qh}^{(q+1)h-1} a(j) \prod_{s=q+1}^{m_k} \left(1 - \frac{c_9}{2} b(sh)\right) - c_8 \sum_{j=k+1}^{(m_k+1)h-1} a(j). \quad (31)$$

Since $\lim_{k\to\infty} \frac{a(m_k h)}{b(m_k h)} = 0$ and $\frac{2\sum_{j=m_k h}^{(m_k+1)h-1} a(j)}{c_9 b(m_k h)} \le \frac{2ha(m_k h)}{c_9 b(m_k h)}$ by Assumption 2, we know that $\lim_{k\to\infty} \frac{2\sum_{j=m_k h}^{(m_k+1)h-1} a(j)}{c_9 b(m_k h)} = 0$, which together with Assumption 2 and Lemma A.1 gives

$$\lim_{k\to\infty} \sum_{q=0}^{m_k} \sum_{j=qh}^{(q+1)h-1} a(j) \prod_{s=q+1}^{m_k} \left(1 - \frac{c_9}{2} b(sh)\right) = \lim_{k\to\infty} \frac{2\sum_{j=m_k h}^{(m_k+1)h-1} a(j)}{c_9 b(m_k h)} = 0. \quad (32)$$

By (31)–(32) and $\lim_{k\to\infty} \sum_{j=k+1}^{(m_k+1)h-1} a(j) = 0$, we have

$$\lim_{k\to\infty} \sum_{i=0}^{k} a(i) \| W\Phi_{F_1}(k, i+1)\| = 0. \quad (33)$$

By Lemma A.4, $\sup_{k\ge 0} \|\mathcal{H}^T R^{-1} z - \mathcal{H}^T R^{-1} \mathcal{H} x(k)\| < \infty$ a.s. Then, we have from (29)–(30) and (33) that $\lim_{k\to\infty} \|Wx(k+1)\| = 0$ a.s. The lemma is proved. ■

Lemma A.8: For the algorithm (4), if Assumptions 1–3 hold, then $\lim_{k\to\infty} \|x_c(k) - x_a(k)\| = 0$ a.s., where $x_c(k)$ is given by $x_c(k+1) = (I_n - \frac{1}{N} a(k) H^T R^{-1} H) x_c(k) + \frac{1}{N} a(k) H^T R^{-1} z$, $k \ge 0$ with any given $x_c(0) \in \mathbb{R}^n$.

Proof: By (7) and the definitions of $x_a(k)$ and $F(k)$, we have

$$x_a(k+1) = \left(I_n - \frac{1}{N} a(k) H^T R^{-1} H\right) x_a(k) + \frac{a(k)}{N} \sum_{i=1}^{N} H_i^T R_i^{-1} z_i + \frac{a(k)}{N} \sum_{i=1}^{N} H_i^T R_i^{-1} H_i (x_a(k) - x_i(k)).$$

Let $e(k) = x_a(k) - x_c(k)$. By the above, we have

$$e(k+1) = a(k) \frac{1}{N} \sum_{i=1}^{N} H_i^T R_i^{-1} H_i (x_a(k) - x_i(k)) + \left(I_n - \frac{1}{N} a(k) H^T R^{-1} H\right) e(k). \quad (34)$$

By Lemma A.7, we know that for any given $\epsilon > 0$, there exists a positive integer k_ϵ such that $\|H_i^T R_i^{-1} H_i\| \|x_a(k) - x_i(k)\| \le \epsilon, i \in \mathcal{V}, k \ge k_\epsilon$. By Assumptions 1–2, we have $\|(I_n - \frac{1}{N} a(k) H^T R^{-1} H)\| \le 1 - c_{12} a(k)$, $k \ge k^*$ for some positive constant c_{12} and positive integer k^*. Let $\bar{k}_\epsilon = \max\{k_\epsilon, k^*\}$. Then, taking the 2-norm on (34) gives

$$\begin{aligned} \|e(k+1)\| \le & \prod_{i=\bar{k}_\epsilon}^{k} \left\| I_n - \frac{1}{N} a(i) H^T R^{-1} H \right\| \left\| e(\bar{k}_\epsilon) \right\| \\ & + \sum_{i=\bar{k}_\epsilon}^{k} a(i) \prod_{j=i+1}^{k} \left\| I_n - \frac{1}{N} a(j) H^T R^{-1} H \right\| \\ & \times \left\| \frac{1}{N} \sum_{s=1}^{N} H_i^T R_s^{-1} H_s (x_a(i) - x_s(i)) \right\| \\ \le & \ \epsilon \sum_{i=\bar{k}_\epsilon}^{k} a(i) \prod_{j=i+1}^{k} (1 - c_{12} a(j)) \\ & + \prod_{i=\bar{k}_\epsilon}^{k} (1 - c_{12} a(i)) \left\| e(\bar{k}_\epsilon) \right\|, k \ge \bar{k}_\epsilon. \end{aligned} \quad (35)$$

We have by Assumption 2 that $\lim_{k\to\infty} \prod_{i=\bar{k}(\epsilon)}^{k} (1 - c_{12} a(i)) = 0$. By Lemma A.1 and Assumption 2, we have $\lim_{k\to\infty} \sum_{i=\bar{k}(\epsilon)}^{k} a(i) \prod_{j=i+1}^{k} (1 - c_{12} a(j)) = c_{12}$. Then, by the arbitrariness of ϵ and (35), it follows that $\lim_{k\to\infty} \|e(k)\| = 0$ a.s. ■

Proof of Theorem 1: By the definition of $x_c(k)$ and Assumptions 1–2, we have $\lim_{k\to\infty} x_c(k) = x_c$ ([11]). Then, by Lemmas A.7–A.8, the theorem is proved. ■

REFERENCES

[1] F. F. Wu, K. Moslehi, and A. Bose, "Power system control centers: Past, present, and future," *Proc. IEEE*, vol. 93, no. 11, pp. 1890–1908, Nov. 2005.

[2] F. C. Schweppe, "Power system static-state estimation, part III: Implementation," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 130–135, Jan. 1970.

[3] D. P. Bertsekas and J. N. Tsitsikilis, *Parallel and Distributed Computation*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.

[4] L. Murphy and F. F. Wu, "An open design approach for distributed energy management systems," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1172–1179, Aug. 1993.

[5] V. Terzija *et al.*, "Wide-area monitoring, protection, and control of future electric power networks," *Proc. IEEE*, vol. 99, no. 1, pp. 80–93, Jan. 2011.

[6] D. M. Falcao, F. F. Wu, and L. Murphy, "Parallel and distributed state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 2, pp. 724–730, May 1995.

[7] R. Ebrahimian and R. Baldick, "State estimation distributed processing [for power systems]," *IEEE Trans. Power Syst.*, vol. 15, no. 4, pp. 1240–1246, Nov. 2000.

[8] G. N. Korres, "A distributed multiarea state estimation," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 73–84, Feb. 2011.

[9] L. Zhao and A. Abur, "Multi area state estimation using synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 611–617, May 2005.

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

[11] L. Xie, D-H. Choi, S. Kar, and H. Poor, "Fully distributed state estimation for wide-area monitoring systems," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1154–1169, Sep. 2012.

[12] G. Wang, G. B. Giannakis, and J. Chen, "Robust and scalable power system state estimation via composite optimization," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6137–6147, Nov. 2019, doi: 10.1109/TSG.2019.2897100.

[13] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1617–1626, May 2013.

[14] 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," *Elect. Power Syst. Res.*, vol. 81, pp. 1060–1069, Apr. 2011.

[15] Y-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, "State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 33–43, Sep. 2012.

[16] L. Hu, Z. Wang, X. Liu, A. V. Vasilakos, and F. E. Alsaadi, "Recent advances on state estimation for power grids with unconventional measurements," *IET Control Theory Appl.*, vol. 11, no. 18, pp. 3221–3232, Dec. 2017.

[17] S. Kar, G. Hug, J. Mohammadi, and J. M. F. Moura, "Distributed state estimation and energy management in smart grids: A consensus+ innovations approach," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1022–1038, Dec. 2014.

[18] M. McGranaghan, D. Houseman, L. Schmitt, F. Cleveland, and E. Lambert, "Enabling the integrated grid: Leveraging data to integrate distributed resources and customers," *IEEE Power Energy Mag.*, vol. 14, no. 1, pp. 83–93, Jan./Feb. 2016.

[19] S. J. Plathottam and H. Salehfar, "Unbiased economic dispatch in control areas with conventional and renewable generation sources," *Elect. Power Syst. Res.*, vol. 119, no. 2, pp. 313–321, Feb. 2015.

[20] A. Wood and B. Wollenberg, *Power Generation, Operation, and Control*. New York, NY, USA: Wiley, 2012.

[21] B. Stott, J. Jardim, and O. Alsaç, "DC power flow revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.

[22] M. D. Ilic and J. Zaborszky, *Dynamics and Control of Large Electric Power Systems*. New York, NY, USA: Wiley, 2000.

[23] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans, and W. Kling, "Usefulness of DC power flow for active power flow analysis with flow controlling devices," in *Proc. 8th IEE Int. Conf. AC DC Power Transm.*, London, U.K., Mar. 2006, pp. 58–62.

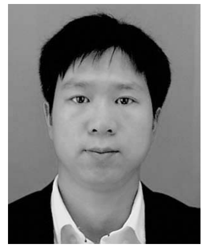
[24] B. Gou and A. Abur, "An improved measurement placement algorithm for network observability," *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 819–824, Nov. 2001.

[25] A. Jadbabaie, J. Lin, and S. M. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, Jun. 2003.

[26] *Power Systems Test Case Archive*, Univ. Washington, Seattle, WA, USA, 1999. [Online]. Available: http://www.ee.washington.edu/research/pstca/

[27] L. Guo, *Time-Varying Stochastic Systems: Stability, Estimation and Control*. Jilin, China: Jilin Sci. Technol. Press, 1990.

[28] T. Li and L. Xie, "Distributed consensus over digital networks with limited bandwidth and time-varying topologies," *Automatica*, vol. 47, no. 9, pp. 2006–2015, Sep. 2011.



Jiexiang Wang received the B.E. degree in automation from East China Jiaotong University, Nanchang, China, in 2014. He is currently pursuing the Ph.D. degree in control theory and engineering with the School of Mechatronic Engineering and Automation, Shanghai University. His research interests are stochastic systems, distributed state estimation and optimization. He received the 28th Zhang Siying Outstanding Youth Paper Award in 2016.

Tao Li (Senior Member, IEEE) received the B.E. degree in automation from Nankai University, Tianjin, China, in 2004, and the Ph.D. degree in systems theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in 2009. Since January 2017, he has been with East China Normal University, Shanghai, China, where he is currently a Professor and the Director of the Department of Mathematical Intelligence Sciences, School of Mathematical Sciences. His current research interests include stochastic systems, cyber-physical multiagent systems, distributed algorithms and game theory. He received the 28th Zhang Siying Outstanding Youth Paper Award in 2016, the Best Paper Award of the 7th Asian Control Conference in 2009, and honourable mentioned as one of five finalists for Young Author Prize of the 17th IFAC Congress in 2008. He received the 2009 Singapore Millennium Foundation Research Fellowship and the 2010 Australian Endeavor Research Fellowship. He was entitled Dongfang Distinguished Professor by Shanghai Municipality in 2012, received the Excellent Young Scholar Fund from NSFC in 2015 and was elected to the Chang Jiang Scholars Program, Ministry of Education, China, Youth Scholar in 2018. He is currently serves as an Associate Editor for several journals, including *Systems and Control Letters*, IEEE CONTROL SYSTEMS LETTERS, and *Science China Information Sciences*. He is a member of IFAC Technical Committee on Networked Systems.