

Distributed Weighted Least-Squares Estimation for Power Networks

Damián Marelli* Brett Ninness* Minyue Fu*,**

* *School of Electrical Engineering and Computer Science, University of Newcastle, University Drive, Callaghan, NSW 2308, Australia.*

** *Department of Control Science and Engineering, Zhejiang University, 38 Zhe Da Road, Hangzhou, 310007, China.*

Abstract: The paper presents a fully distributed scheme to optimal parameter estimation, with application in multi-area interconnected power systems. In this scheme, each area performs its own local parameter estimation based on low-dimensional local and boundary measurements as well as the estimate of boundary parameters from its neighbors. We show, under certain assumptions, that the distributed estimation scheme provides the same accurate unbiased estimate as the centralized weighted least-squares estimate with finite time convergence. The distributed estimation scheme is robust to communication link failures, delays, and asynchronism of control centers in different areas. A simulation using the IEEE 118-bus system is included to demonstrate the performance.

Keywords: Distributed parameters, parameter estimation, least-squares estimation, power systems, smart power applications.

1. INTRODUCTION

Electric power networks are undergoing profound changes recently and receiving increasing attention from researchers in different fields. By incorporating a communication, computing and control overlay, more efficient and intelligent processes are integrated into the electric power networks. Parameter estimation for the so-called *quasi-steady state* parameters is considered to be one of the key integrating components for the real time energy management system (EMS).

Quasi-steady state parameter estimation (or simply called state estimation) in power networks has been introduced in the early 1970's Schweppe and Wildes (1970). The traditional centralized parameter estimator is typically installed in a control center collecting all measurements over the entire network, and providing the optimal estimate of the parameters for the power network. The data is measured by SCADA systems and the estimation usually takes minutes to get a snapshot of a normal sized power network Ren and Guo (2005); Huang et al. (2007). However, due to the deregulation of energy markets, large amounts of power are transferred over high-rate, long-distance lines spanning multiple areas to form a very large scale network Gomez-Exposito et al. (2011). Also, policy and privacy considerations make a centralized estimation inappropriate for a power network spanned over multiple areas, regional transmission organizations (RTOs), and/or countries. Thus, it leads to the emergence of *hierarchical estimation* Gomez-Exposito and de la Villa Jaen (2009); Korres (2011); Zhao and Abur (2005); Jiang et al. (2007);

Patel and Girgis (2007); Jiang et al. (2008); Yang et al. (2012) and *distributed estimation* Lin (1992); Lin and Lin (1994); Falcao et al. (1995); Conejo et al. (2007); Xie et al. (2011); Pasqualetti et al. (2012). A good recent survey on quasi-steady state parameter estimation for power networks refers to Gómez-Expósito et al. (2011).

In principle, a desirable distributed estimator for large-scale power networks should own the following properties: (1) the distributed estimator should be able to deal with the case where local parameters might not be uniquely identifiable due to bad data removal and finer decomposition of control areas; (2) the information exchange between different control centers should be kept as low as possible to reduce the communication load and improve the estimation response time; (3) the resulting estimate should be accurate or close enough to the optimal one obtained from the traditional centralized estimator; (4) the estimation scheme should exhibit a fast convergence rate or even finite time convergence for the purpose of real time monitoring; (5) the convergence to a correct estimate should be robust to link failures and time delays commonly occurred in communication networks, and asynchronism inherent in the distributed setup. However, no estimation scheme in recent development incorporates all the desired properties. For example, most algorithms such as hierarchical estimation of Iwamoto et al. (1989); Zhao and Abur (2005); Korres (2011) and distributed estimation of Lin and Lin (1994); Falcao et al. (1995); Conejo et al. (2007), presume locally topological observability. The locally topological observability assumption is no longer necessary in the distributed estimation schemes of Xie et al. (2011) and Pasqualetti et al. (2012). However, in Xie et al. (2011) and Pasqualetti et al. (2012) each control center has to communicate its own estimate of the entire high-dimensional parameter

¹ *Email addresses:* Damián.Marelli@newcastle.edu.au (Damián Marelli), Brett.Ninness@newcastle.edu.au (Brett Ninness), Minyue.Fu@newcastle.edu.au (Minyue Fu).

vector to its neighboring control centers, which scales unfavorably with the size of the power network. Moreover, in Pasqualetti et al. (2012) only an approximate estimate is ensured though an analytical estimation error bound is provided within a finite number of iterations. On the other hand, Xie et al. (2011) shows almost sure convergence towards the centralized parameter estimation result, which is similar to asymptotic convergence of distributed estimation strategies Cattivelli et al. (2008); Schizas et al. (2009) developed in other fields. The performance is not well suited for applications in power networks as the convergence is only asymptotic. Besides, no analytical study is conducted in Xie et al. (2011); Pasqualetti et al. (2012) for the behavior of the estimation schemes in response to communication link failures, transmission delays and asynchronism of distributed control centers.

The objective of this paper is to propose a fully distributed estimation scheme that incorporates all the desired features. To this end, we propose a new distributed estimator for each control center, called *local estimator*, which requires only local (own area) and boundary information. Each local estimator provides an estimate of the parameters of its own area, and the connection with neighboring areas is done by exchanging only small amounts of boundary estimation data. Notice that the physical linkage (tie-lines) between different areas is usually low-dimensional and most measurements taken in one area are not affected by the parameters in other areas. The proposed scheme is based on the distributed method for weighted least squares (WLS) developed in Marelli and Fu (2015). A key property of this method is that it converges to the globally optimal estimate. Moreover, this convergence is achieved in a finite number of iterations, being equal to the diameter of the graph modeling the interconnection of control areas. Finally, we show that the finite-time convergence is also ensured in the presence of link failures and communication delays, as well as asynchronism between distributed control centers.

2. PROBLEM DESCRIPTION

2.1 Measurements in power networks

Consider a power network, for example, the IEEE 118-bus system shown in Fig. 1. Let $x \in \mathbb{R}^n$ be the parameter vector at a certain time instant consisting of the voltage phasors at all buses. The measurements of the whole power network usually take the following form:

$$z = h(x) + \eta, \quad (1)$$

where $z \in \mathbb{R}^m$ is the measurement vector, $h(\cdot)$ is a measurement function, and η is the Gaussian random measurement error vector satisfying $\eta \sim \mathcal{N}(0, R_\eta)$.

The traditional SCADA measurements typically contain voltage magnitudes, power injections at the measured buses, and power flows along the measured transmission lines. In this case, the parameter vector (i.e., quasi-steady state) x in eq. (1) is usually defined in the polar coordinate form. Adopting the approximated estimation model presented in Schweppe and Wildes (1970) which follows from the linearization around an operating point \tilde{x} of eq. (1), the measurements can be expressed as

$$z = Hx + \nu, \quad (2)$$

where $H \in \mathbb{R}^{m \times n}$ is the Jacobian matrix of $h(\cdot)$ and ν is still assumed to satisfy $\nu \sim \mathcal{N}(0, R_\nu)$. This model is also appropriate if the measurements are made using phasor measurement units (PMU).

2.2 Centralized parameter estimation

For the linear measurement model (2), the centralized WLS estimation (Steven, 1993, Section 8.4) is represented as

$$\hat{x}^{opt} = \arg \min_{\hat{x}} (z - H\hat{x}) R_\nu^{-1} (z - H\hat{x}). \quad (3)$$

If H has full column rank and R_ν is invertible, the WLS estimate \hat{x}^{opt} and the estimation error covariance Σ take the following explicit forms:

$$\hat{x}^{opt} = \Psi^{-1}\alpha; \quad \Sigma = \Psi^{-1}; \quad (4)$$

where

$$\alpha = H^T R_\nu^{-1} z; \quad \Psi = H^T R_\nu^{-1} H. \quad (5)$$

To have the centralized WLS parameter estimation solution, it requires to collect all the measurements distributed in different areas and assume the complete knowledge of the matrix H and R_ν . Thus, both communication and computation burden scales unfavorably with the size of the power networks.

2.3 Partition of power networks

For a practical power network, the measurements are either related to the parameters of one bus (such as power injections and voltage phasor measurements), or indicate the relationship between two adjacent buses (such as power flows and current phasor measurements). These characteristics naturally lead to a sparse measurement matrix H . The measurements can also be classified as *local measurements*, meaning that the measurements are only functions of the parameters of one control area, and *boundary measurements*, consisting of tie-line measurements related to the parameters of more than one control areas. The local measurements are given by

$$y_i = A_i x_i + v_i, \quad (6)$$

where x_i and y_i represent the local parameter vector and the local measurement vector in the control area i , and $v_i \sim \mathcal{N}(0, R_i)$ is the local measurement noise. The boundary measurements linking the parameters in control area i and j are represented as

$$z_{i,j} = D_{i,j} x_i + C_{i,j} x_j + w_{i,j}, \quad (7)$$

where $w_{i,j} \sim \mathcal{N}(0, S_{i,j})$ is the boundary measurement noise. Note that the boundary measurement $z_{i,j}$ is usually of very low dimension and that the boundary measurement (7) is shared by both control center i and j , i.e., $z_{i,j} = z_{j,i}$. All measurement noises are assumed to be uncorrelated.

We use a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ to abstract the partition of a power network, where $\mathcal{V} = \{1, \dots, I\}$ is a set of nodes with each node corresponding to a control area and \mathcal{E} is an edge set with an edge $(i, j) \in \mathcal{E}$ indicating that there is a boundary measurement $z_{i,j}$ relying on both x_i and x_j . In the following, we use \mathcal{N}_i to represent the neighbor set of i , i.e., $\mathcal{N}_i = \{j : (i, j) \in \mathcal{E}\}$. For a partition made in Fig. 1 for the IEEE 118-bus system, the graph \mathcal{G} is depicted in Fig. 2, for which an edge (1,3) represents the boundary measurement related to the parameters in areas 1 and 3.

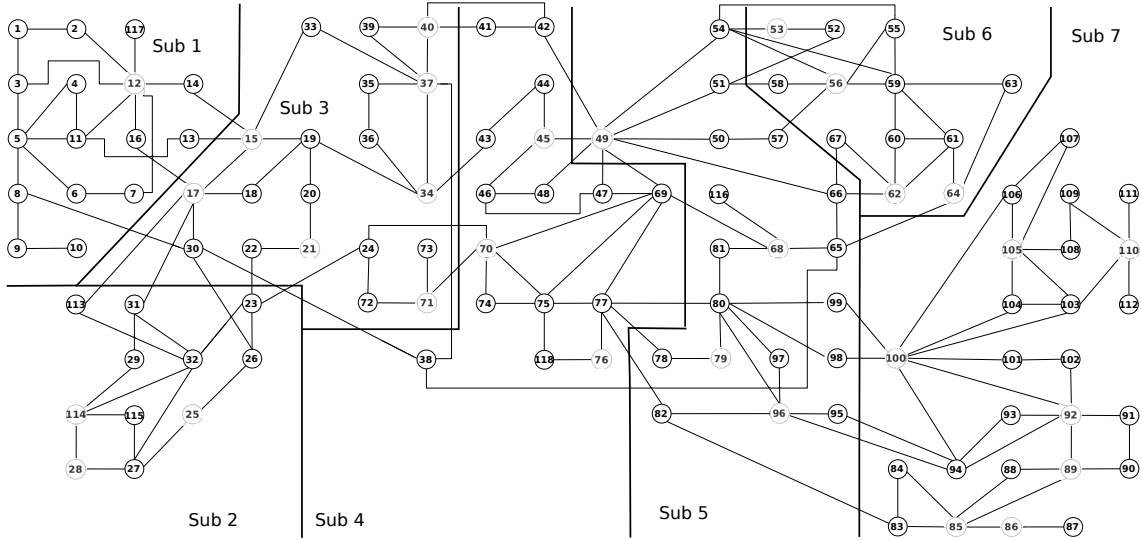


Fig. 1. Topological structure of the IEEE 118-bus system.

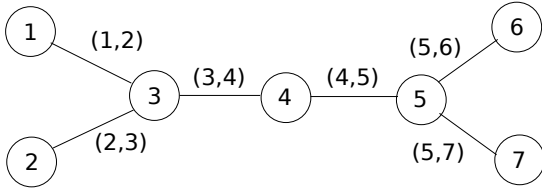


Fig. 2. The graph \mathcal{G} describing the partition of the IEEE 118-bus system.

2.4 Distributed parameter estimation problems

The distributed parameter estimation problem we address in this paper is as follows: For a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ reflecting the abstract structure of a power network, design a local estimator at each control center i ($i \in \mathcal{V}$) to have an unbiased local estimate \hat{x}_i , based on the local measurements (6) and the boundary measurements (7), as well as its neighbors' local estimate \hat{x}_j ($j \in \mathcal{N}_i$), minimizing the global objective function

$$J(\hat{x}) := (z - H\hat{x})^T R_\nu^{-1} (z - H\hat{x}), \quad (8)$$

where \hat{x} is the aggregated parameters of the local estimation \hat{x}_i , i.e., $\hat{x} = [\hat{x}_1^T, \dots, \hat{x}_I^T]^T$.

We make the following assumptions in the paper:

- **Assumption 1:** The graph \mathcal{G} is acyclic;
- **Assumption 2:** The measurement matrix H has full column rank and R_ν is invertible.

Assumption 1 means that the graph corresponding to a partition of a power network does not have a cycle, which in practice can be made by properly partitioning the power network and by properly placing the measurement units. An example is given in Fig. 1 for the partition, and Jiang et al. (2007) gives another example. Assumption 2 is also necessary for centralized estimators.

3. DISTRIBUTED ESTIMATION ALGORITHM

In this section, we propose a distributed estimation algorithm, in which each control center builds an estimate \hat{x}_i

for the local parameters x_i in its region. In Section 3.1, we state the algorithm for a network whose graph has a particular topology, and in Section 3.2 we use this result to state the proposed algorithm.

3.1 Preliminary

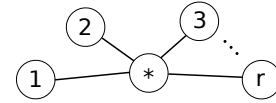


Fig. 3. Topological structure of a star graph.

Fig. 3 shows a *star graph of order $r + 1$* with a *central node* denoted by \star and r neighbor nodes. For this network topology, the measurement equation is given by

$$\begin{bmatrix} y_\star \\ y_1 \\ \vdots \\ y_r \\ z_1 \\ \vdots \\ z_r \end{bmatrix} = \begin{bmatrix} A_\star & 0 & \cdots & 0 \\ 0 & A_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & A_r \\ D_1 & C_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ D_r & 0 & \cdots & C_r \end{bmatrix} \begin{bmatrix} x_\star \\ x_1 \\ \vdots \\ x_r \end{bmatrix} + \begin{bmatrix} v_\star \\ v_1 \\ \vdots \\ v_r \\ w_1 \\ \vdots \\ w_r \end{bmatrix}, \quad (9)$$

where x_\star, x_1, \dots, x_r represent the local parameters to be estimated, y_\star, y_1, \dots, y_r denote the measurements on the nodes, and z_1, \dots, z_r denote the measurements on the edges $(1, \star), \dots, (r, \star)$. Also, $[v_\star, v_1, \dots, v_r, w_1, \dots, w_r]^T \sim \mathcal{N}(0, T)$ is the measurement noise with $T = \text{diag}\{R_\star, R_1, \dots, R_r, S_1, \dots, S_r\}$.

Consider the following subsystem

$$\begin{bmatrix} y_\star \\ z_1 \\ \vdots \\ z_r \end{bmatrix} = \begin{bmatrix} A_\star \\ D_1 \\ \vdots \\ D_r \end{bmatrix} x_\star + \begin{bmatrix} v_\star \\ w_1 \\ \vdots \\ w_r \end{bmatrix} \quad (10)$$

obtained from eq. (9) by eliminating the measurements and terms not related to the parameters x_\star . This can be thought as how node \star understands the measurement from

its local point of view. The information vector $\check{\alpha}_\star$ and information matrix $\check{\Psi}_\star$ for the WLS associated to (10) are

$$\check{\alpha}_\star = A_\star^T R_\star^{-1} y_\star + \sum_{j=1}^r D_j^T S_j^{-1} z_j,$$

$$\check{\Psi}_\star = A_\star^T R_\star^{-1} A_\star + \sum_{i=1}^r D_j^T S_j^{-1} D_j.$$

Correspondingly, the WLS estimate and the estimation error covariance based on the modified measurement equation (10) are respectively

$$\check{x}_\star = \check{\Sigma}_\star \check{\alpha}_\star, \quad \check{\Sigma}_\star = \check{\Psi}_\star^{-1}.$$

Similarly, from node j 's local point of view ($j = 1, \dots, r$), the modified measurement equation becomes

$$\begin{bmatrix} y_j \\ z_j \end{bmatrix} = \begin{bmatrix} A_j \\ C_j \end{bmatrix} x_j + \begin{bmatrix} v_1 \\ w_j \end{bmatrix}.$$

Thus,

$$\check{\alpha}_j = A_j^T R_j^{-1} y_j + C_j^T S_j^{-1} z_j,$$

$$\check{\Psi}_j = A_j^T R_j^{-1} A_j + C_j^T S_j^{-1} C_j,$$

and

$$\check{x}_j = \check{\Sigma}_j \check{\alpha}_j, \quad \check{\Sigma}_j = \check{\Psi}_j^{-1}.$$

In order for node \star to obtain an optimal estimate for the local parameters x_\star , it has to combine its own estimate with its neighbors' estimates. The following lemma, which follows from Marelli and Fu (2015, Lemma 13), makes this statement precise.

Lemma 1. Suppose Assumption 2 holds.

- (1) The optimal WLS estimate \hat{x}_\star of x_\star , and its associated estimation error covariance Σ_\star , are given by

$$\hat{x}_\star = \Sigma_\star \left(\check{\alpha}_\star - \sum_{j=1}^r \beta_{\star,j} \right), \quad (11)$$

$$\Sigma_\star = \left(\check{\Psi}_\star - \sum_{j=1}^r \Phi_{\star,j} \right)^{-1}, \quad (12)$$

where

$$\beta_{\star,j} = D_j^T S_j^{-1} C_j \check{x}_j, \quad (13)$$

$$\Phi_{\star,j} = D_j^T S_j^{-1} C_j \check{\Sigma}_j C_j^T S_j^{-1} D_j \quad (14)$$

- (2) The matrices $\check{\Psi}_j$, $j = \star, 1, \dots, r$ and $\check{\Psi}_\star - \sum_{j=1}^r \Phi_{\star,j}$ are invertible.

3.2 Distributed WLS Algorithm

Based on Lemma 1 we can devise a distributed estimation algorithm, which is summarized in Algorithm 1.

The idea of the distributed parameter estimation algorithm is simple. Initially, each control center i calculates the estimate $\check{x}_i^{(0)}$ and estimation error covariance $\check{\Sigma}_i^{(0)}$, for the local parameters x_i , based on the reduced measurement equation, as for the central node in the star graph case. Then each control center updates its estimate $\hat{x}_i^{(t+1)}$ and estimation error covariance $\Sigma_i^{(t+1)}$, by combining its

own estimate $\check{x}_i^{(0)}$ and estimation error covariance $\check{\Sigma}_i^{(0)}$ with the correction factors $(\gamma_{i,j}^{(t)}, \Upsilon_{i,j}^{(t)})$ received from its neighbors. Meanwhile, each control center calculates the correction factors $(\gamma_{j,i}^{(t+1)}, \Upsilon_{j,i}^{(t+1)})$ related to its own most recently updated local estimate and estimation error covariance, and next transmits them to its neighbors.

Algorithm 1 Distributed WLS estimation: (Algorithm on node i , $i \in \mathcal{V}$)

Initialization:

- (1) Compute the local estimate and its associated estimation error covariance

$$\check{x}_i^{(0)} = \check{\Sigma}_i^{(0)} \check{\alpha}_i, \quad \check{\Sigma}_i^{(0)} = \check{\Psi}_i^{-1},$$

with

$$\check{\alpha}_i^0 = A_i^T R_i^{-1} y_i + \sum_{j \in \mathcal{N}_i} D_{i,j}^T S_{i,j}^{-1} z_{i,j},$$

$$\check{\Psi}_i^0 = A_i^T R_i^{-1} A_i + \sum_{j \in \mathcal{N}_i} D_{i,j}^T S_{i,\mathcal{N}_i(k)}^{-1} D_{i,j}.$$

- (2) Transmit to every node $j \in \mathcal{N}_i$ the correction factor

$$\gamma_{j,i}^{(0)} = D_{i,j} \check{x}_j^{(0)}, \quad \Upsilon_{j,i}^{(0)} = D_{i,j} \check{\Sigma}_j^{(0)} D_{i,j}^T.$$

Main loop: For $t = 0, 1, \dots$, let $\gamma_{i,j}^{(t)}, \Upsilon_{i,j}^{(t)}$ be the correction factor received by node i from node j .

- (1) Update the local estimate and its associated estimation error covariance based on the received correction factor:

$$\hat{x}_i^{(t+1)} = \Sigma_i^{(t+1)} \left(\check{\alpha}_i^0 - \sum_{j \in \mathcal{N}_i} \beta_{i,j}^{(t)} \right), \quad (15)$$

$$\Sigma_i^{(t+1)} = \left(\check{\Psi}_i^0 - \sum_{j \in \mathcal{N}_i} \Phi_{i,j}^{(t)} \right)^{-1}, \quad (16)$$

where

$$\beta_{i,j}^{(t)} = D_{i,j}^T S_{i,j}^{-1} \gamma_{i,j}^{(t)}, \quad (17)$$

$$\Phi_{i,j}^{(t)} = D_{i,j}^T S_{i,j}^{-1} \Upsilon_{i,j}^{(t)} S_{i,j}^{-1} D_{i,j}. \quad (18)$$

- (2) Compute

$$\check{x}_{j,i}^{(t+1)} = \check{\Sigma}_{j,i}^{(t+1)} \left(\check{\alpha}_i^0 - \sum_{k \in \mathcal{N}_i/\{j\}} \beta_{i,k}^{(t)} \right),$$

$$\check{\Sigma}_{j,i}^{(t+1)} = \left(\check{\Psi}_i^0 - \sum_{k \in \mathcal{N}_i/\{j\}} \Phi_{i,k}^{(t)} \right)^{-1}$$

for every $j \in \mathcal{N}_i$, and then transmit to node j the correction factor

$$\gamma_{j,i}^{(t+1)} = C_{j,i} \check{x}_{j,i}^{(t+1)}, \quad \Upsilon_{j,i}^{(t+1)} = C_{j,i} \check{\Sigma}_{j,i}^{(t+1)} C_{j,i}^T.$$

Remark 2. Notice that the information $(\gamma_{j,i}^{(t)}, \Upsilon_{j,i}^{(t)})$ exchanged by nodes is usually of very low dimension.

4. FINITE-TIME-CONVERGENCE AND OPTIMALITY

4.1 Ideal Setup

In this section we assume that inter-node communications do not have communication delays and packet loss, and all local control centers operate synchronously. We then state a result asserting that the local estimate \hat{x}_i , given by Algorithm 1, at every node $i \in \mathcal{V}$, converges after a finite number of steps, to the (block) component \hat{x}_i^{opt} of the globally optimal estimate \hat{x}^{opt} .

We introduce the following definitions. For a given graph, a *path* is a concatenation of adjacent edges (without loops), and its *length* is the number of edges forming it. The *radius* Γ_i of node i is defined as the maximum length of a path from node i to any other node in the graph. Also, the *diameter* of the graph is the maximum radius over all nodes and is denoted by Γ . The convergence result is stated in the following theorem, which follows from Marelli and Fu (2015, Theorem 15).

Theorem 3. Consider the system (2) with Assumption 1 and 2. If Algorithm 1 is used, then for every $i \in \mathcal{V}$,

$$\hat{x}_i^{(t)} = \hat{x}_i^{opt} \quad \text{for all } t \geq \Gamma_i. \quad (19)$$

Moreover, for any $i \in \mathcal{V}$ and $t \in \mathbb{N}_0$ the matrices $\tilde{\Psi}_i^0 - \sum_{j \in \mathcal{N}_i} \Phi_{i,j}^{(t)}$ and $\tilde{\Psi}_i^0 - \sum_{j \in \mathcal{N}_i/\{k\}} \Phi_{i,k}^{(t)}$ for any $k \in \mathcal{N}_i$ are invertible.

Remark 4. From (19) in Theorem 3, it follows that the local estimates on all nodes converge to the optimal one after $\Gamma = \max\{\Gamma_i, i \in \mathcal{V}\}$ steps.

Remark 5. Notice that the invertibility of the matrices $\tilde{\Psi}_i^0 - \sum_{j \in \mathcal{N}_i} \Phi_{i,j}^{(t)}$ and $\tilde{\Psi}_i^0 - \sum_{k \in \mathcal{N}_i/\{j\}} \Phi_{i,k}^{(t)}$, for all $j \in \mathcal{N}_i$ is necessary for the validity of Algorithm 1.

4.2 Communication Packet Loss, Delays and Asynchronism

In practical applications, communication packets among different control centers may be lost or delayed. Also, clocks in different control centers may not be perfectly synchronized, causing mismatches between time-stamps at different nodes. In this section we introduce a modification to our proposed algorithm to deal with these issues.

We introduce the following modification to the notation used Algorithm 1. Recall that $(\gamma_{i,j}^{(t)}, \Upsilon_{i,j}^{(t)})$ denotes the correction factor transmitted by node j to node i at time t . This information can be lost, delayed, or affected by timing mismatch. Consequently, a number of packets, say $(\gamma_{i,j}^{(t-\tau_1)}, \Upsilon_{i,j}^{(t-\tau_1)}), \dots, (\gamma_{i,j}^{(t-\tau_N)}, \Upsilon_{i,j}^{(t-\tau_N)})$, $n = 1, \dots, N$, may be simultaneously received by node i and time t . Let $\tilde{\tau} = \min\{\tau_n : n = 1, \dots, N\}$ denote the smallest time delay among those packets, and

$$(\tilde{\gamma}_{i,j}^{(t)}, \tilde{\Upsilon}_{i,j}^{(t)}) = (\gamma_{i,j}^{(t-\tilde{\tau})}, \Upsilon_{i,j}^{(t-\tilde{\tau})}), \quad (20)$$

be the packet associated to that delay. Then, the modification to the algorithm consists in using $(\tilde{\gamma}_{i,j}^{(t)}, \tilde{\Upsilon}_{i,j}^{(t)})$ in place of $(\gamma_{i,j}^{(t)}, \Upsilon_{i,j}^{(t)})$ in (17)-(18).

We do the following assumption.

Assumption 3: The timing mismatch between every two nodes is upper bounded by a finite integer. Also, the probability distribution of delays and packet losses is such that, for each t , and with probability one, there exists t^* and $s > t$ satisfying

$$(\tilde{\gamma}_{i,j}^{(t^*)}, \tilde{\Upsilon}_{i,j}^{(t^*)}) = (\tilde{\gamma}_{i,j}^{(s)}, \tilde{\Upsilon}_{i,j}^{(s)}). \quad (21)$$

In words (21) guarantees that, for every t , a packet from node j equal or newer than $(\gamma_{i,j}^{(t)}, \Upsilon_{i,j}^{(t)})$ is guaranteed to eventually arrive to node i . Under this condition, the following property follows,

Corollary 6. Consider the system (2) together with Assumptions 1-3. If Algorithm 1 runs with the modification (20), then, with probability one, there exists t^* such that

$$\hat{x}_i^{(t)} = \hat{x}_i^{opt} \quad \text{for all } t \geq t^*. \quad (22)$$

5. SIMULATION RESULTS

We use the IEEE 118-bus system to test the performance of the proposed distributed estimation algorithm. The system composed of 118 buses is partitioned into 7 subsystems connected by tie-lines, as shown in Figs. 1-2. The measurement locations are given in Table 1.

Subsystem	Buses with local measurements	
Sub 1	3 5 9 12	
Sub 2	25 28 114	
Sub 3	15 17 21 34 37 40 71	
Sub 4	45 70 76	
Sub 5	49 68 79 96	
Sub 6	53 56 62 64	
Sub 7	85 86 89 92 100 105 110	

Tie-line	Boundary buses
(1, 3)	12 13 14 15 16
(2, 3)	17 31 113
(3, 4)	34 37 38 40 41 42 43 70 71
(4, 5)	42 45 47 48 49 68 69
(5, 6)	53 54 56 62 64 65 66
(5, 7)	94 96 98 99 100

Table 1. Measurement Locations

We run Algorithm 1 on the system described above and use 2000 Monte Carlo simulations to compute, at iteration t , the expected value of the norm squared of the difference between the distributed and centralized estimates, i.e., $\sum_{i=1}^I \left\| \hat{x}_i^{(t)} - \hat{x}_i^{opt} \right\|^2$, and the expected value of $\sum_{i=1}^I \text{Tr} \left\{ \Sigma_i^{(t)} \right\}$. Random i.i.d. packet loss with rate p of different values is considered. Fig. 4 and Fig. 5 show the simulation results. As expected, Fig. 4 and Fig. 5 show that a higher packet loss rate causes longer convergence steps, but the algorithm still converges in a finite number of iterations. The curves corresponding to $p = 0$ in both of Fig. 4 and Fig. 5 show that, as indicated in Theorem 3, the algorithm converges after $t \geq \Gamma = 4$ steps.

6. CONCLUSIONS

We have proposed a novel distributed parameter estimation algorithm for large-scale power networks. The algo-

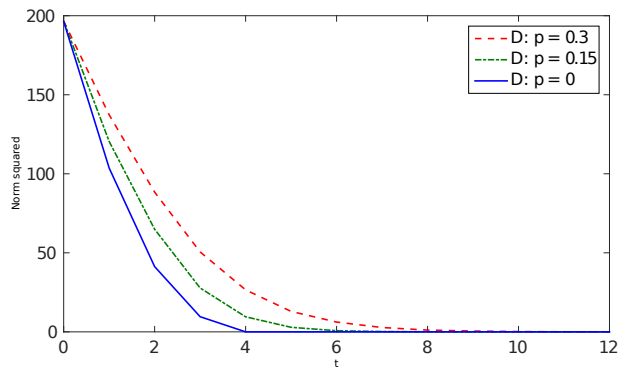


Fig. 4. The norm squared of the difference between the distributed (D) and centralized estimates.

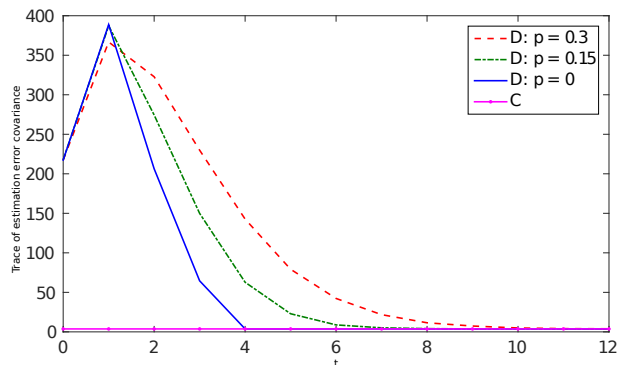


Fig. 5. The trace of the estimation error covariances of the distributed (D) and centralized (C) estimates.

algorithm only requires locally topological structure information, local measurements and low-dimensional boundary information exchange from its neighbors. The local estimates converge in finite time to the centralized WLS estimate and they exhibit good robustness against packet losses, time delays or asynchronous processing.

REFERENCES

Cattivelli, F.S., Lopes, C.G., and Sayed, A.H. (2008). Diffusion recursive least-squares for distributed estimation over adaptive networks. *IEEE Transactions on Signal Processing*, 56(5), 2069–2084.

Conejo, A.J., de la Torre, S., and Canas, M. (2007). An optimization approach to multiarea state estimation. *IEEE Transactions on Power Systems*, 22(1), 213–221.

Falcao, D.M., Wu, F.F., and Murphy, L. (1995). Parallel and distributed state estimation. *IEEE Transactions on Power Systems*, 10(2), 724–730.

Gomez-Exposito, A., Abur, A., de la Villa Jaen, A., and Gomez-Quiles, C. (2011). A multilevel state estimation paradigm for smart grids. *Proceedings of the IEEE*, 99(6), 952–976.

Gomez-Exposito, A. and de la Villa Jaen, A. (2009). Two-level state estimation with local measurement preprocessing. *IEEE Transactions on Power Systems*, 24(2), 676–684.

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

Huang, Z., Guttromson, R., Nieplocha, J., and Pratt, R. (2007). Transforming power grid operations. *Scientific Computing*, 24(5).

Iwamoto, S., Kusano, M., and Quintana, V.H. (1989). Hierarchical state estimation using a fast rectangular-coordinate method. *IEEE Transactions on Power Systems*, 4(3), 870–880.

Jiang, W., Vittal, V., and Heydt, G.T. (2007). A distributed state estimator utilizing synchronized phasor measurements. *IEEE Transactions on Power Systems*, 22(2), 563–571.

Jiang, W., Vittal, V., and Heydt, G.T. (2008). Diakoptic state estimation using phasor measurement units. *IEEE Transactions on Power Systems*, 23(4), 1580–1589.

Korres, G.N. (2011). A distributed multiarea state estimation. *IEEE Transactions on Power Systems*, 26(1), 73–84.

Lin, S.Y. (1992). A distributed state estimator for electric power systems. *IEEE Transactions on Power Systems*, 7(2), 551–557.

Lin, S.Y. and Lin, C.H. (1994). An implementable distributed state estimator and distributed bad data processing schemes for electric power systems. *IEEE Transactions on Power Systems*, 9(3), 1277–1284.

Marelli, D. and Fu, M. (2015). Distributed weighted least-squares estimation with fast convergence for large-scale systems. *Automatica*, 51(1), 27 – 39.

Pasqualetti, F., Carli, R., and Bullo, F. (2012). Distributed estimation via iterative projections with application to power network monitoring. *Automatica*, 48(5), 747–758.

Patel, M.Y. and Girgis, A.A. (2007). Two-level state estimation for multi-area power system. In *Proceedings of the 2007 IEEE Power Engineering Society General Meeting*, 1–6. doi:10.1109/PES.2007.386113.

Ren, J. and Guo, Z. (2005). Period-oriented state estimation approach for power system operational control. In *Proceedings of the 2005 IEEE/PES Transmission and Distribution Conference and Exhibition*, 1–5.

Schizas, I.D., Mateos, G., and Giannakis, G.B. (2009). Distributed lms for consensus-based in-network adaptive processing. *IEEE Transactions on Signal Processing*, 57(6), 2365–2382.

Schweppe, F.C. and Wildes, J. (1970). Power system static-state estimation, part I, II, III. *IEEE Transactions on Power Apparatus and Systems*, 89, 120–135.

Steven, M.K. (1993). *Fundamentals of Statistical Signal Processing*. Prentice Hall.

Xie, L., Choi, D.H., and Kar, S. (2011). Cooperative distributed state estimation: local observability relaxed. In *Proceedings of IEEE Power and Energy Society General Meeting*, 1–11. San Diego, CA, USA.

Yang, X., Zhang, X., and Zhou, S. (2012). Coordinated algorithms for distributed state estimation with synchronized phasor measurements. *Applied Energy*, 96(0), 253–260.

Zhao, L. and Abur, A. (2005). Multi area state estimation using synchronized phasor measurements. *IEEE Transactions on Power Systems*, 20(2), 611–617.