2015 American Control Conference
Palmer House Hilton
July 1-3, 2015. Chicago, IL, USA

A Fully Distributed State Estimation Using Matrix Splitting Methods

Ariana Minot and Na Li

Abstract— **In this paper, we propose a fully distributed state estimation algorithm for the electric power grid using matrix splitting methods. Our method incorporates both traditional SCADA measurements as well as new PMU data. The proposed distributed scheme relies only on local information and a minimal amount of information from neighboring areas which reduces communication burdens and increases the robustness of state estimation. The distributed estimate is guaranteed to exponentially converge to the centralized optimal estimate. The effectiveness of the method is demonstrated in various numerical experiments.**

Index terms–distributed algorithms, state estimation, matrix splitting.

I. Introduction

Accurate state estimation is critical for robust and secure operation of the electric power grid. In order to efficiently operate the power grid, the system operator relies on the output from state estimation to assess real-time operating conditions. Advancing state estimation algorithms will facilitate further improvements in state-of-the-art system control functionalities, such as fault diagnosis, oscillation damping, and online power dispatch. In order to realize an electric grid with a high penetration of renewable energy sources, such advanced control capabilities are increasingly important. At the same time, estimating the system state quickly with high accuracy is challenging in such a setting. For these reasons, recently there has been much interest in using advanced sensor technology to enhance state estimation. One such example is the development of a Wide-Area Measurement System (WAMS) that uses Phasor Measurement Units (PMUs) [1].

Processing these new measurements introduces additional computational demands, motivating the need for state estimation algorithms that execute in a distributed manner. A distributed approach allows control areas, or geographically aggregated buses, to locally estimate their state rather than require a central processer. Furthermore, since each control area utilizes a subset of information from the global system, distributed algorithms increase robustness of state estimation while reducing the computation and memory requirements within a control area. Fully-distributed state estimation methods will be of great use for achieving wide-area control between large interconnected areas of the power grid, as well as successfully monitoring down at the micro-grid level [2].

Many research efforts in developing distributed state estimation methods for electric power systems consider a hierarchical communication scheme. In hierarchical distributed approaches, state estimation is performed locally and then information is communicated to a central processor. The central processor is responsible for coordinating the local estimates to produce a global state estimate. Hierarchical approaches to state estimation in power systems include [3], [4], [5], [6]. A principal disadvantage of relying on a central coordinator is the possibility of communication bottlenecks and reduced robustness.

In contrast to hierachical distributed algorithms are fully-distributed approaches to state estimation, in which information is exchanged only via neighbor-to-neighbor communication. Gossip-based algorithms for fully distributed state estimation have been of recent interest [7], [8], [9]. For such methods, an estimate of the global state is required at each area. This presents a shortcoming, since for large networks, the memory requirements can be prohibitive. Fully distributed methods that require only local rather than global estimates per area have recently been explored. These include algorithms based on matrix decomposition methods[10], [11]; alternating direction method of multipliers (ADMM) [12]; and information filter-based techniques [13]. The matrix decomposition-based methods of [10], [11] lack guarantees for convergence of the distributed state estimates to the estimates obtained by a centralized state estimator. While the ADMM approach in [12] guarantees asymptotic convergence, possible disadvantages include the computation and storage of additional information in the form of Lagrange multipliers and complications in extending the algorithm to an asynchronous setting. The method of [13] limits the structure of the network to be acyclic. Although this algorithm has guarantees to converge in a finite number of iterations, the number of iterations needed for convergence increases linearly with the network size. For large-scale networks, asymptotically convergent methods can be advantageous, especially if convergence speed scales independently of the network size.

In this paper, we explore the use of matrix splitting techniques [14], [15] for developing a new distributed state estimation algorithm that exploits inherent sparse structure in power systems. For ease of exposition, we adopt the DC power flow model [2]. In the DC setting, the optimal state estimation problem is equivalent to solving a system of linear equations. Applying matrix splitting to solve this system of equations leads to an iterative solution to the state estimation problem. The choice of matrix splitting is made to ensure convergence of the distributed state estimates. In our algorithm, each bus calculates its own state estimate, and information is exchanged between neighboring nodes. The contributions of this work include a new fully distributed

A. Minot and N. Li are with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. emailsminot@fas.harvard.edu, nali@seas.harvard.edu

method for power system state estimation with the following features:

1) The method combines the use of both traditional SCADA (supervisory control and data acquisition) system measurements of power injections and flows with PMU measurements of voltages and currents.
2) The convergence of the distributed estimates to the optimal estimates obtained by a centralized algorithm is guaranteed analytically.
3) Each local bus or area only needs to hold an estimate of its own local states rather than an estimate of the entire global state of the system. This frees a large amount of communication and memory resources.

The paper is outlined as follows. In Section II, we provide the mathematical problem statement and introduce the centralized state estimation problem. In Section III, we present our distributed state estimation algorithm with analysis of its communication requirements and convergence properties. In Section IV, numerical simulations demonstrate the effectiveness of our method.

Notations: We use v_k to denote the k^{th} entry of a vector $\mathbf{v}$. The $(i,j)^{th}$ entry of a matrix $\mathbf{M}$ is given by M_{ij}, and the i^{th} row and j^{th} column are given by $[\mathbf{M}]_\mathbf{i}$ and $[\mathbf{M}]^\mathbf{j}$ respectively. The transpose of a vector or matrix $\mathbf{X}$ is denoted $\mathbf{X^T}$.

II. Problem Statement and Formulation

We consider an interconnected power network, denoted by an undirected graph $(\mathcal{N},\mathcal{E})$ with a set $\mathcal{N} \triangleq \{1,2,\ldots,n\}$ of buses and a set $\mathcal{E} \subseteq \mathcal{N}\times\mathcal{N}$ of transmission lines connecting the buses. The goal of state estimation in power systems is to infer the unknown voltages (magnitude and phase angle) at each bus from a set of measurements of the system. Figure 1 provides an example of the IEEE 14-bus test system with different measurements. In this work, we consider the linearized DC state estimation problem, which is a good approximation under the assumptions that 1) the voltage angle differences between neighboring buses are small, 2) the voltage magnitudes are all close to 1 per unit (p.u.), and 3) the transmission lines have negligible resistance [2]. Thus in the setting of this paper, the voltage magnitudes are assumed to be 1 p.u., leaving only the voltage phase angles, $\boldsymbol{\theta}$, as unknown variables. The objective of state estimation is to infer $\boldsymbol{\theta}$ from the measurements.

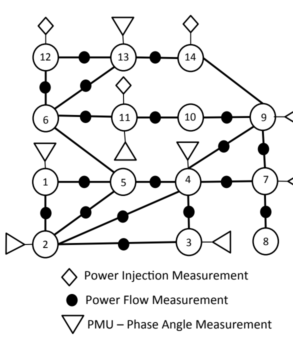


Fig. 1. Schematic illustration of IEEE 14-bus test system with location and types of measurements used.

There are two typical power measurement systems. One is the traditional SCADA measurements including power flows along transmission lines and power injections at buses; the other system uses PMUs to measure the voltages and currents directly. Since we only consider the DC state estimation where the voltage magnitudes are assumed to be 1 p.u., we assume that the power flow measurements and current measurements are interchangeable and thus only consider power flow measurements. Therefore, we consider the following three measurements: 1) a measurement of power flow along the transmission line between buses, i and j, denoted by $\hat{P}_{ij}$, 2) a measurement of power injection at bus i, denoted by $\hat{P}_i$, and 3) a measurement of phase angle at bus i, denoted by $\hat{\theta}_i$. The cost of measurement units is too prohibitive to deploy enough sensors to take all possible measurements of the system. We denote the set of measurements as $\mathcal{Z}$ and the ordered vector of measurements as $\mathbf{z}$. The DC model linearly relates the measurements $\mathbf{z}$ and unknown voltage phase angles $\boldsymbol{\theta}$,

$$\mathbf{z} = \mathbf{H}\,\boldsymbol{\theta} + \mathbf{e},$$

where $\mathbf{e} \sim \mathcal{N}(\mathbf{0},\mathbf{R})$. Let the neighbor set of node i be denoted $\mathcal{N}_i$. Denote the susceptance of the transmission line between nodes i and j as B_{ij}. The entries of the measurement matrix $\mathbf{H}$ are given as follows:

1) If $z_k = \hat{P}_{ij}$, then

$$H_{kl} = \begin{cases} B_{ij}, & l = i \\ -B_{ij}, & l = j \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

2) If $z_k = \hat{P}_i$, then

$$H_{kl} = \begin{cases} \sum_{j\in\mathcal{N}_i} B_{ij}, & l = i \\ -B_{ij}, & l = j, j \in \mathcal{N}_i \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

3) If $z_k = \hat{\theta}_i$, then

$$H_{kl} = \begin{cases} 1, & l = i \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Given an invertible matrix $\mathbf{W}$ that weights the measurements, we seek an estimate of $\boldsymbol{\theta}$ which solves the following optimization problem:

$$\min_{\boldsymbol{\theta}} \, (\mathbf{z} - \mathbf{H}\boldsymbol{\theta})^\mathbf{T}\mathbf{W}(\mathbf{z} - \mathbf{H}\boldsymbol{\theta}), \tag{4}$$

where one possible choice for $\mathbf{W}$ is the inverse covariance matrix $\mathbf{R^{-1}}$. The following assumption is made on the weighting matrix:

Assumption 1. *The weighting matrix* $\mathbf{W}$ *is invertible and diagonal with entries* $W_{kk} \equiv w_k > 0$.

The solution to (4) is given by choosing an estimate $\boldsymbol{\theta}_C^*$ to satisfy the first-order optimality conditions,

$$(\mathbf{H^T W H})\boldsymbol{\theta}_\mathrm{C}^* = \mathbf{H^T W z}.$$

We make the following assumption.

Assumption 2. *(Observability) The measurement matrix* $\mathbf{H}$ *has full column rank.*

With this, we can calculate the unique solution as

$$\boldsymbol{\theta}_C^* = (\mathbf{H^T W H})^{-1}\mathbf{H^T W z}. \quad (5)$$

Calculating the centralized solution $\boldsymbol{\theta}_C^*$ depends on full knowledge of $\mathbf{H}, \mathbf{W}$, and $\mathbf{z}$. This requires all measurements being sent to a central coordinator, as well as calculating the inversion of the matrix $\mathbf{H^T W H}$, which can be too large to be practical for wide-area control. These communication and computational burdens present a difficulty for state estimation in large-scale power systems. To address this challenge, we propose a new distributed algorithm for state estimation that requires only local information internal to the local control area of interest and information communicated from neighbors. For ease of exposition, we assume that each bus is a control area with a local processor. Our method can be extended to cases where a control area includes multiple buses in the network as described in [16].

III. Distributed State Estimation Algorithm

We propose a new state estimation algorithm based on a matrix-splitting technique which allows us to calculate $(\mathbf{H^T W H})^{-1}$ in a distributed way. This method is inspired by the use of matrix-splitting for developing a distributed Newton method for the Network Utility Maximization (NUM) problem of Wei *et al.* in [15].

A. Matrix Splitting for Distributed State Estimation (DSE) Problem

First, we introduce the idea behind matrix splitting for solving linear systems [14]. Given a square linear system of m equations, $\mathbf{Ax} = \mathbf{y}$, we seek to write $\mathbf{A}$ as the sum of an invertible matrix, $\mathbf{M}$, and a matrix $\mathbf{N}$, *i.e.* $\mathbf{A} = \mathbf{M}+\mathbf{N}$. Then for an arbitrary $\mathbf{x^0} \in \mathbb{R}^\mathbf{n}$, consider the following iterative scheme:

$$\mathbf{x^{t+1}} = -\mathbf{M^{-1}Nx^t} + \mathbf{M^{-1}y}. \quad (6)$$

The sequence $\{\mathbf{x^t}\}$ converges to its limit $\mathbf{x}^*$ as $t \to \infty$ if and only if the spectral radius of the matrix $\mathbf{M^{-1}N}$ is strictly less than 1 [14]. In the event that the sequence converges, the limit $\mathbf{x}^*$ is the solution of the system, i.e., $\mathbf{Ax}^* = \mathbf{y}$.

To apply this to the DSE problem, we define

$$\mathbf{A} \equiv \mathbf{H^T W H}. \quad (7)$$

We can decompose $\mathbf{A}$ into the sum of a matrix containing its diagonal entries, $\mathbf{D}$, and a matrix containing its off-diagonal entries $\mathbf{E}$. Specifically, let

$$D_{ij} = \begin{cases} A_{ii}, & j = i \\ 0, & j \neq i \end{cases} \quad (8)$$

and

$$E_{ij} = \begin{cases} A_{ij}, & j \neq i \\ 0, & j = i \end{cases} \quad (9)$$

yielding $\mathbf{A} = \mathbf{D} + \mathbf{E}$. The main challenge is to identify matrices $\mathbf{M}$ and $\mathbf{N}$ such that $\mathbf{A} = \mathbf{M} + \mathbf{N}$ and the spectral radius $\rho(\mathbf{M^{-1}N}) < \mathbf{1}$. To this end, we introduce a diagonal matrix $\bar{\mathbf{E}}$ whose i^{th} entry equals:

$$\bar{E}_{ii} \equiv \alpha \sum_{j=1}^{n} |E_{ij}|, \quad (10)$$

for some positive constant α.

Proposition 1. *Let* $\mathbf{M} = \mathbf{D} + \bar{\mathbf{E}}$ *and* $\mathbf{N} = \mathbf{E} - \bar{\mathbf{E}}$. *Then, for* $\alpha \geq \frac{1}{2}, \rho(\mathbf{M^{-1}N}) < \mathbf{1}$.

Proof. By Theorem 2.5.3 of [17], to prove that $\rho(\mathbf{M^{-1}N}) < \mathbf{1}$, it is sufficient to show that $\mathbf{M}+\mathbf{N}$ and $\mathbf{M}-\mathbf{N}$ are positive definite. First, it is straightforward to verify that $\mathbf{H^T W H} = \mathbf{M} + \mathbf{N}$ is positive definite given Assumptions 1 and 2.

Second, we show that $\mathbf{M} - \mathbf{N}$ is positive definite. As a corollary to the Gershgorin Circle Theorem [18], it is sufficient to show that $\mathbf{M}-\mathbf{N}$ is strictly diagonally dominant with strictly positive diagonal entries.[1] To show this, we insert our choice for $\mathbf{M}$ and $\mathbf{N}$ to obtain $\mathbf{M} - \mathbf{N} = \mathbf{D} + \mathbf{2}\bar{\mathbf{E}} - \mathbf{E}$. By definition, $\mathbf{D}$ is a diagonal matrix and $D_{ii} = (\mathbf{H^T W H})_{\mathrm{ii}} > 0$. Thus, given that $\alpha \geq \frac{1}{2}$,

$$\begin{aligned}(\mathbf{D} + \mathbf{2}\bar{\mathbf{E}} - \mathbf{E})_{\mathrm{ii}} &= D_{ii} + 2\bar{E}_{ii} - E_{ii} \\ &> 2\,\alpha \sum_{j=1}^{n} |E_{ij}| - 0 \\ &\geq \sum_{j \neq i} |E_{ij}| \\ &= \sum_{j \neq i} |(\mathbf{D} + \mathbf{2}\bar{\mathbf{E}} - \mathbf{E})_{\mathrm{ij}}|.\end{aligned}$$

Therefore, $M - N$ is a positive definite matrix. □

B. Proposed Algorithm and Analysis of Information Communication Requirements

We now study the information required to compute an individual phase angle estimate. Applying the iterative scheme of equation (6) to the state estimation problem with our choice of matrix splitting, we obtain

$$\boldsymbol{\theta}^{(t+1)} = -(\mathbf{D} + \bar{\mathbf{E}})^{-1}(\mathbf{E} - \bar{\mathbf{E}})\boldsymbol{\theta}^{(\mathbf{t})} + (\mathbf{D} + \bar{\mathbf{E}})^{-1}\mathbf{H^T W z} \quad (11)$$

Besides our choice of matrix splitting satisfying conditions necessary for convergence, we also note that the only matrix to invert, namely $(\mathbf{D} + \bar{\mathbf{E}})$, is diagonal and therefore easy to invert distributedly. In order to analyze the information communication of this iterative approach, we study the expression for the estimate of the unknown phase angle at a specific bus. From equation (11), we have

$$\theta_i^{(t+1)} = \frac{1}{D_{ii} + \bar{E}_{ii}}(\bar{E}_{ii}\boldsymbol{\theta}_i^{(t)} - [\mathbf{E}]_\mathbf{i}\boldsymbol{\theta}^{(\mathbf{t})} + [\mathbf{H}^\mathrm{T}\mathbf{W}]_\mathbf{i}\mathbf{z}) \quad (12)$$

We then utilize the particular structure of the power grid state estimation problem in order to verify the information from neighboring areas needed to calculate $\theta_i^{(t+1)}$. Each node

[1] A matrix $\mathbf{A}$ is diagonal dominant if $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|\ \forall i$.

is assumed to know measurements of its own phase angle, power injection, and local power flows. We introduce the following quantities:

$$\sigma_{ij} = \begin{cases} w_k, & z_k = \hat{P}_{ij} \\ 0, & \hat{P}_{ij} \notin \mathcal{Z} \end{cases},$$

$$\sigma_{i,P} = \begin{cases} w_k, & z_k = \hat{P}_{i} \\ 0, & \hat{P}_{i} \notin \mathcal{Z} \end{cases},$$

and

$$\sigma_{i,\theta} = \begin{cases} w_k, & z_k = \hat{\theta}_{i} \\ 0, & \hat{\theta}_{i} \notin \mathcal{Z} \end{cases}. \quad (13)$$

In addition, we define the set of "1-hop" neighbor nodes of node i as nodes which are not direct neighbors but share a common neighbor:

$$\mathcal{N}_i^{\dagger} \equiv \{ j \mid j \notin \mathcal{N}_i, \mathcal{N}_i \cap \mathcal{N}_j \neq \emptyset \}.$$

By definition, we have that $\mathcal{N}_i \cap \mathcal{N}_i^{\dagger} = \emptyset$. For notational simplicity in what follows, let

$$\mathcal{N}_i^{*} = \mathcal{N}_i \cup \mathcal{N}_i^{\dagger}.$$

With this, we present the final expression for an estimate at an individual bus in terms of the local information available and only the subset of external information needed.

Lemma 1. *The updates to the distributed state estimates of equation (12) have the form given in equation (14), where the entries of* $A_{ij} =$

$$\begin{cases} \sigma_{i,\theta} + \sum\limits_{j\in\mathcal{N}_i} (\sigma_{ij} + \sigma_{ji} + \sigma_{j,P}) B_{ij}^2 + \\ \sigma_{i,P} \Big(\sum\limits_{j\in\mathcal{N}_i} B_{ij} \Big)^2, & j = i \\ -(\sigma_{ij} + \sigma_{ji}) B_{ij}^2 + \sum\limits_{l\in\mathcal{N}_i\cap\mathcal{N}_j} \sigma_{l,P} B_{il} B_{jl} \\ -B_{ij} \Big(\sigma_{i,P} \sum\limits_{l\in\mathcal{N}_i} B_{il} + \sigma_{j,P} \sum\limits_{l\in\mathcal{N}_j} B_{jl} \Big), & j \in \mathcal{N}_i \\ \sum\limits_{l\in\mathcal{N}_i\cap\mathcal{N}_j} \sigma_{l,P} B_{il} B_{jl}, & j \in \mathcal{N}_i^{\dagger} \\ 0, & \textit{else.} \end{cases} \quad (15)$$

Proof. Expanding equation (12) using the definitions for matrices $\mathbf{D}, \mathbf{E}$, and $\bar{\mathbf{E}}$ from equations (8)-(10), we obtain an expression for $\theta_i^{(t+1)}$ soley in terms of entries of the matrices $\mathbf{H}$ and $\mathbf{A} = \mathbf{H^T W H}$,

$$\theta_i^{(t+1)} = \frac{1}{A_{ii} + \alpha \sum\limits_{j\neq i} |A_{ij}|} \Big\{ (\alpha \sum_{j\neq i} |A_{ij}|) \theta_i^{(t)} - \sum_{j\neq i} A_{ij} \theta_j^{(t)} + \sum_{k=1}^{m} H_{ki} w_k z_k \Big\} \quad (16)$$

Since $\mathbf{W}$ is diagonal, we have that

$$A_{ij} = \sum_{k=1}^{m} w_k H_{ki} H_{kj}. \quad (17)$$

From equations (1)-(3), we see that the topology of the power grid network and the measurement model induces a sparsity on the entries of $\mathbf{A}$. We determine the entries A_{ij} by considering four cases.

1) Case 1: $(i = j)$
According to equations (1)-(2), for measurements of power flow on a line incident with node i or a power injection at a neighboring node, $H_{ki} = \pm B_{ij}$, so $H_{ki}^2 = B_{ij}^2$. For measurement of the power injection at node i, then $H_{ki}^2 = (\sum_{j\in\mathcal{N}_i} B_{ij})^2$. According to equation (3), for a measurement of the voltage phase angle at node i, $H_{ki}^2 = 1$. All other measurements k in the network will not contribute to the sum in equation (17) since $H_{ki} = 0$. Therefore, we see that we need only consider measurements local to node i or power injections at its neighbors. Since in general not all possible power flows, power injections, and voltage phase angles are measured, we use the quantities defined in equation (13) to include only the available measurements.
2) Case 2: $(j \in \mathcal{N}_i)$
Similar logic applies as in Case 1 except in the instance of a power injection at node j, we need the following information about node j's incident edges, $\sum_{l\in\mathcal{N}_j} B_{jl}$. In addition, we must include possible contributions from a power injection at a mutual neighbor of nodes i and j, which does not require any communication of information that is neither local to node i nor to node j.
3) Case 3: $(j \in \mathcal{N}_i^{\dagger})$
In this case, we consider "1-hop" neighbor nodes of i. If nodes i and j are not direct neighbors but share a neighbor l with a power injection measurement k, then $H_{ki} H_{kl} = B_{il} B_{jl} > 0$, yielding a non-zero entry for A_{ij}.
4) Case 4: Otherwise $(j \notin \mathcal{N}_i^{*}, j \neq i)$
If j does not satisfy the conditions for the first three cases, then we must have either H_{ki} or H_{kj} be zero for all possible measurements.

With this, we obtain the entries of A_{ij} in terms of the available measurements, their variances, the network connectivity, and the bus susceptance parameters. Therefore in equation (16), instead of requiring *all* entries $\{A_{ij} | j \neq i\}$, we only require $\{A_{ij} | j \in \mathcal{N}_i^{*}\}$. Furthermore, the only measurements that contribute to the term $\sum_{k=1}^{m} H_{ki} w_k z_k$ are those locally available to node i and power injections at its neighbors. These simplifications provide the final result in equation (14). □

From this lemma, we know that if there are only power flow and phase angle measurements at node i (i.e. no power injection measurement), then node i needs to communicate only its current estimate $\theta_i^{(t)}$ to its neighbors. If there is a

$$\theta_i^{(t+1)} = \frac{1}{A_{ii} + \alpha \sum_{\mathrm{j} \in \mathcal{N}_{\mathrm{i}}^*} |A_{ij}|} \Bigg\{ (\alpha \sum_{\mathrm{j} \in \mathcal{N}_i^*} |A_{ij}|) \theta_{\mathrm{i}}^{(\mathrm{t})} - \sum_{\mathrm{j} \in \mathcal{N}_i^*} A_{ij} \theta_{\mathrm{j}}^{(\mathrm{t})} + \sigma_{i,P} (\sum_{\mathrm{j} \in \mathcal{N}_i} B_{ij}) \hat{P}_i + \sum_{\mathrm{j} \in \mathcal{N}_i} B_{ij} (\sigma_{ij} \hat{P}_{ij} - \sigma_{ji} \hat{P}_{ji} - \sigma_{j,P} \, \hat{P}_j) + \sigma_{i,\theta} \, \hat{\theta}_i \Bigg\} \tag{14}$$

power injection measured at node i, then the following additional information must be communicated to its neighbors:

1) Value of measured power injection, $\hat{P}_i$
2) The set of neighbors' state estimates $\{\theta_j^{(t)}\}_{j \in \mathcal{N}_i}$ must be shared with each neighbor.

Algorithm 1: Matrix-Splitting Based Distributed State Estimation

Initialization : Node i has access to local measurements $\{\hat{\theta}_i, \hat{P}_i, \hat{P}_{ij}, \hat{P}_{ji}\} \in \mathcal{Z}$.
$\forall\ 1 \leq i \leq n, \theta_i^{(0)} = 0$. If $\hat{P}_i \in \mathcal{Z}$, node i sends $\hat{P}_i, w_{k|z_k = \hat{P}_i}$, and $\{B_{ij}\}_{j \in \mathcal{N}_i}$ to nodes $j \in \mathcal{N}_i$.
for $t := 0$ *to* T **do**
  for $i := 1$ *to* n **do**
    $\theta_i^{(t+1)} = f(\theta_i^{(t)}, \{\theta_j^{(t)}\}_{j \in \mathcal{N}_i^*})$ from equation (14)
    Node i sends $\theta_i^{(t+1)}$ to nodes $j \in \mathcal{N}_i$
  end
  if $\hat{P}_i \in \{z\}$ **then**
    Node i sends $\{\theta_j^{(t+1)}\}_{j \in \mathcal{N}_i}$ to nodes $j \in \mathcal{N}_i$
  end
end

It is noted that only the phase angle estimates change with time and need to be communicated at each time step. Power injections are communicated once during initialization. [2]

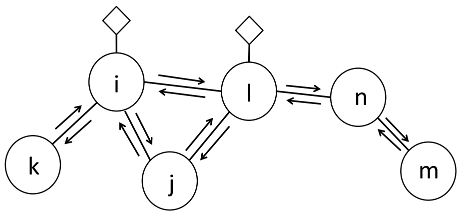


Fig. 2. Example to illustrate information exchange of Algorithm 1. Neighboring nodes are connected via an edge in the graph. There are power injection measurements at nodes i and l, signified by a diamond-shaped sensor. Communication occurs directly between neighbors.

We illustrate the information exchange for nodes with a power injection measurement using the example in Figure 2. Power injections are measured at nodes i and l.
At initialization,

- Node i sends $\hat{P}_i$ to its neighboring nodes j, k, and l.
- Node i receives $\hat{P}_l$ from node l.

[2]The line susceptance parameters and measurement weights are assumed to be known a priori.

At iteration t, after each node has calculated its next estimate using equation (14),

- Node i sends $\theta_i^{(t+1)}, \theta_j^{(t+1)}, \theta_k^{(t+1)}$, and $\theta_l^{(t+1)}$ to its neighboring nodes j, k, and l.
- Node i receives $\theta_j^{(t+1)}, \theta_l^{(t+1)}$, and $\theta_n^{(t+1)}$ from node l. Node i also receives $\theta_k^{(t+1)}$ from node k and $\theta_j^{(t+1)}$ from node j.

C. Convergence Analysis

Under *Assumption 2*, there is a unique solution given by equation (5), which we denote here as $\boldsymbol{\theta}^*$. In Proposition 1, we showed that $\rho(\mathbf{M^{-1}N}) < \mathbf{1}$. Because the iterative scheme in equation (6) forms a discrete linear dynamic system, according to (Thm 6.1) [19], the iterative scheme in (6) exponentially converges to the solution $\boldsymbol{\theta}^*$. The convergence speed is determined by $\rho(\mathbf{M^{-1}N})$. Formally speaking, we have the following theorem about the convergence:

Theorem 1. *The distributed state estimation scheme as described in Algorithm 1 exponentially converges to the optimal solution in (4). The convergence speed is determined by* $\rho(\mathbf{M^{-1}N})$.

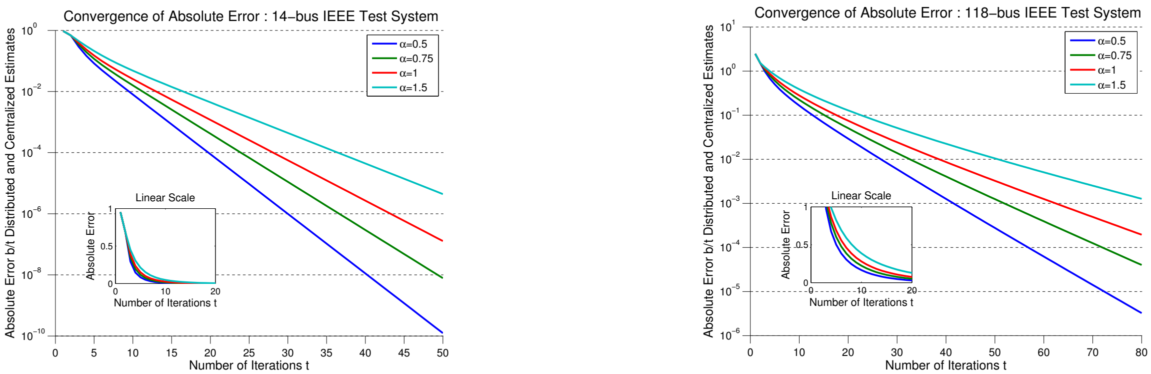
IV. Numerical Results

We study the performance of our algorithm by calculating the error between the distributed estimate and the centralized optimal estimate at each iteration t, $\delta^t \equiv ||\boldsymbol{\theta}_d^{(t)} - \boldsymbol{\theta}_c^*||$. Figure 3 demonstrates the exponential convergence of the distributed estimates for the IEEE 14-bus and 118-bus systems respectively. We study the convergence for different values of the parameter α in equation (10) and empirically find $\alpha = \frac{1}{2}$ to be optimal.

The values of the power flows and power injections in the simulation are on the order of 10^{-1} p.u. The measurements are perturbed by additive zero-mean Gaussian noise $\mathcal{N}(0, \sigma^2)$ with $\sigma = 0.01$. As stated in Theorem 1, the convergence of the distributed algorithm is governed by the spectral radius, $\rho(\mathbf{M^{-1}N})$. The value of $\rho(\mathbf{M^{-1}N})$ is sensitive to the measurement configuration, the choice of weighting matrix $\mathbf{W}$, and the parameter α of the matrix splitting. We used the identity matrix for $\mathbf{W}$ in these tests. From Figure 3, it is demonstrated that the lower value of $\rho(\mathbf{M^{-1}N})$ in the 14-bus system leads to faster convergence than in the 118-bus system.

V. Conclusion

In this paper, we have proposed a new fully distributed state estimation algorithm based on a matrix splitting iterative approach. Attractive features of the algorithm include



(a) 14-bus test system.

(b) 118-bus test system.

Fig. 3. Convergence is shown for different values of the parameter α from equation (10). For the 14-bus system (a), using the measurement configuration in Figure 1 with $\alpha = \frac{1}{2}$, the convergence is exponential with rate determined by $\rho(\mathbf{M^{-1}N}) = 0.64$. For the 118-bus system (b), using the measurement configuration in Table I with $\alpha = \frac{1}{2}$, the convergence is exponential with rate determined by $\rho(\mathbf{M^{-1}N}) = 0.87$.

TABLE I
118-BUS MEASUREMENT CONFIGURATION

Measurement Type	**Locations (Bus Index)**
Missing Power Flows	(2,12), (9,10), (24,72), (93,94)
Power Injections	1, 2, 3, 5, 18, 19, 21, 22, 25, 33, 34, 35, 36, 55, 56, 69, 71, 72, 77, 80, 81, 86, 88, 90, 98, 106,107 110, 114, 118
Missing phase angles	3, 5, 14, 15, 18, 19, 20, 21, 26, 33, 38, 39, 41, 44, 45, 51, 63, 64, 67, 68, 51, 63, 64, 67, 68, 63, 64, 67, 68, 69, 73, 74, 75, 76, 79, 82, 84, 89, 92, 94, 96, 97, 99, 100, 101, 105, 106, 107

the limited amount of information that needs to be shared between neighboring nodes and that each node needs only to store and compute its local estimate. Future work includes: 1) analytically characterizing the spectral radius associated with the matrix splitting in terms of the measurement configuration, network topology, and matrix splitting parameters to increase the exponential rate of convergence, 2) extending the algorithm to detect and identify bad data in a distributed manner, and 3) generalizing the approach to the AC power flow setting.

ACKNOWLEDGMENT

The authors would like to thank Yue Lu in the School of Engineering and Applied Sciences in Harvard University and Ermin Wei in Electrical Engineering in Northwestern University for their helpful and constructive discussion.

REFERENCES

[1] A. Chakrabortty and P. Khargoneka, "Introduction to Wide-Area Control of Power Systems," in *Proceedings of 2013 American Control Conference (ACC)*, 2013, pp. 6773–6785.
[2] G. B. Giannakis, V. Kekatos, N. Gatsis, S.-J. Kim, H. Zhu, and B. F. Wollenberg, "Monitoring and Optimization for Power Grids: A Signal Processing Perspective," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 107–128, 2013.
[3] R. Ebrahimian and R. Baldick, "State Estimation Distributed Processing," *IEEE Trans. Power Systems*, vol. 15, no. 4, pp. 1240–1246, 2000.
[4] L. Zhao and A. Abur, "Multiarea State Estimation Using Synchronized Phasor Measurements," *IEEE Trans. Power Systems*, vol. 20, no. 2, pp. 611–617, 2005.
[5] A. G. Expsito, A. Abur, A. de la Villa Jaen, and C. Gomez-Quiles, "A Multilevel State Estimation Paradigm for Smart Grids," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 952–976, 2011.
[6] G. N. Korres, "A Distributed Multiarea State Estimation," *IEEE Trans. Power Systems*, vol. 26, no. 1, pp. 73–84, 2011.
[7] L. Xie, D.-H. Choi, S. Kar, and H. V. Poor, "Fully Distributed State Estimation for Wide-Area Monitoring Systems," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1154–1169, 2012.
[8] X. Li and A. Scaglione, "Advances in Decentralized State Estimation for Power Systems," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2013 IEEE 5th International Workshop on*, Dec 2013, pp. 428–431.
[9] Y. Hu, A. Kuh, T. Yang, and A. Kavcic, "A Belief Propagation Based Power Distribution System State Estimator," *Computational Intelligence Magazine, IEEE*, vol. 6, no. 3, pp. 36–46, Aug 2011.
[10] A. Conejo, S. De La Torre, and M. Canas, "An Optimization Approach to Multiarea State Estimation," *Power Systems, IEEE Transactions on*, vol. 22, no. 1, pp. 213–221, Feb 2007.
[11] J. Carvalho and F. Barbosa, "Distributed Processing in Power System State Estimation," in *Electrotechnical Conference, 2000. MELECON 2000. 10th Mediterranean*, vol. 3, May 2000, pp. 1128–1131 vol.3.
[12] V. Kekatos and G. Giannakis, "Distributed Robust Power System State Estimation," *Power Systems, IEEE Transactions on*, vol. 28, no. 2, pp. 1617–1626, May 2013.
[13] X. Tai, Z. Lin, M. Fu, and Y. Sun, "A New Distributed State Estimation Technique for Power Networks," in *American Control Conference (ACC), 2013*, June 2013, pp. 3338–3343.
[14] R. Varga, *Matrix Iterative Analysis*. Springer Series in Computational Mathematics, 2000.
[15] E. Wei, A. E. Ozdaglar, and A. Jadbabaie, "A Distributed Newton Method for Network Utility Maximization," in *CDC*. IEEE, 2010, pp. 1816–1821.
[16] A. Gomez-Exposito, A. Abur, A. de la Villa Jaen, C. Gomez-Quiles, P. Rousseaux, and T. Van Cutsem, "A Taxonomy of Multi-Area State Estimation Methods," *Electric Power Systems Research*, vol. 81, pp. 1060–1069, 2011.
[17] R. W. Cottle, J. Pang, and R. Stone, *Linear Complementarity Problem.* Academic Press, 1992.
[18] R. Varga, *Gershgorin and His Circles*. Springer Series in Computational Mathematics, 2004.
[19] J. Demmel, *Applied Numerical Linear Algebra*. SIAM, 1997.