Decentralized PMU-assisted Power System State Estimation with Reduced Inter-Area Communication

Neelabh Kashyap, Stefan Werner, *Senior Member, IEEE,* Yih-Fang Huang, *Fellow, IEEE,*

***Abstract*—This paper presents a decentralized approach to multi-area power system state estimation using a combination of conventional measurement devices and newer phasor measurement units (PMU). We employ a reduced-order approach to state estimation, wherein the PMU observable and unobservable components of the state vector are estimated separately in each control area. The estimation problem is solved by exchanging information through an improved gossip-based protocol which takes into account the weak measurement coupling between control areas. A detailed analysis of the proposed protocol reveals that the amount of information exchange is always less than the naïve gossip based scheme. We show through simulations on the IEEE 118-bus test system that, in practice, the gossip iterations converge in one iteration regardless of the number of control areas. Our simulations also show that the marginal information exchange is also considerably lower when using the proposed method.**

***Index Terms*—Distributed Estimation, Reduced-Order, Phasor Measurement Units, Multi-Area State Estimation**

I. Introduction

As we transition towards a smarter grid, three key factors have emerged which necessitate a rethinking of our energy infrastructure. Firstly, following the trend towards power industry deregulation, portions of the grid are being administered by independent transmission system operators (TSOs) [1]. These TSOs are responsible for the monitoring and maintenance of large parts of the electricity transmission system, sometimes for an entire country. Secondly, the size of the power interconnection is set to increase in order to get access to renewable resources scattered over a large geographical area. The intermittent nature of these renewable sources call for near real-time monitoring of the power grid. Finally, the implementation of a wide-area monitoring system (WAMS) enabled by a network of synchronized phasor measurement units (PMUs) is leading to a significant growth in the number, frequency and accuracy of measurements in the power grid [2], [3].

State estimation plays a key role in the monitoring of an electrical grid. Employing a centralized monitoring scheme in a modernized electrical grid would lead to communication bottlenecks [4] which impacts the timeliness of state estimation. Furthermore, this would also create a single point of failure which goes against the reliability principle of the smart grid. It is evident that decentralized approaches are needed for monitoring the energy infrastructure of the future.

Distributed approaches to power system state estimation, namely, multi-area state estimation (MASE), were first developed to overcome computational issues involved in centralized state estimation of large power systems [5], [6]. A summary of different approaches to MASE can be found in [7]. Some approaches to MASE rely on a hierarchical two-level method wherein the estimates obtained by the local state estimator of each area are sent to a central control center to estimate the system-wide state vector [8]–[11]. On the other hand, decentralized approaches to MASE do not require a central control center; instead, the local state estimators share information amongst themselves through a communication network. This idea was even briefly discussed in Schweppe's seminal paper [12]. Several approaches to decentralized MASE have been proposed since [13]–[18]. Security aspects of such estimation schemes have been studied recently in [19].

A major limitation in both the aforementioned classes of methods is that all the local control areas need to have enough measurement redundancy in order to compute the local estimates [4], i.e., each area needs to be locally observable. In contrast, new methods for decentralized MASE only require that the overall system be observable. Such methods are presented in [20], [21]. These methods, however, do not estimate the system-wide state vector in each area, which is recommended for the implementation of wide-area monitoring and control [2], [22].

Another method has been proposed in [23], which estimates the system-wide state vector in each area by sharing information using network gossiping [24]. The method is shown to converge faster than the method proposed in [4], and is also robust against bad-data. The method is based on the gossip-based Gauss-Newton (GGN) algorithm, and the convergence of this method is analytically proven [25]. The method in [23], however, does not take into account the fact that, in practical power systems, *measurement coupling* between areas is weak [7]. Each area participates in the information exchange related to all state variables regardless of whether the state variable was observed in that area. Hence, the number of gossip iterations performed is larger than necessary, resulting in an undesirable increase in communication between areas. In [26] we outline an information exchange scheme for gossip-based Gauss-Newton which drastically reduces inter-area communication.

Neelabh Kashyap is with the Aalto University School of Electrical Engineering, Finland (e-mail: neelabh.kashyap@aalto.fi). His work was supported, in part, by a grant from the Fortum Foundation. Stefan Werner is with Department of Electronics and Telecommunications at the Norwegian University of Science and Technology (NTNU). His work was supported by the Academy of Finland under Grant 296849. Yih-Fang Huang is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA. His work was supported, in part, by the U.S. National Science Foundation's Cyber-Physical Systems program under grant CNS-1239244.

When performing state estimation with both conventional and PMU measurements, we posit that it is beneficial to separate the state vector into PMU observable and PMU unobservable parts, where the PMU observable states are estimated using PMU measurements alone. In [27], [28], we develop a *reduced-order* formulation of the state estimation problem which allows us to estimate the PMU unobservable states from conventional measurements and estimates of PMU observable states. The following arguments support choosing the reduced-order approach over a combined approach:

- PMU measurement errors have a much lower variance than those of conventional measurements. Employing a *mixed-measurement approach* [29]–[32] or a two-step approach [33], [34] would involve operations on an ill-conditioned matrix when implementing weighted least-squares estimation. This leads to numerical instability [35].
- Since conventional measurements are asynchronous, measurements taken in a window of a few seconds are not consistent with each other [36]. This induces *time-skew errors* in conventional measurements. We show in [27] that time-skew errors could potentially offset the improvement in estimation accuracy gained by the increased redundancy when using all available measurements in a combined state estimator.
- As shown in Fig. 1, the measurement architecture of the PMU network is independent of the conventional supervisory control and data acquisition (SCADA) system. The PMU measurement system is based on the newer standards like IEEE C37.118, while the SCADA system is still largely dependent on a myriad of vendor-dependent, non-standard legacy devices with proprietary software and communication protocols.

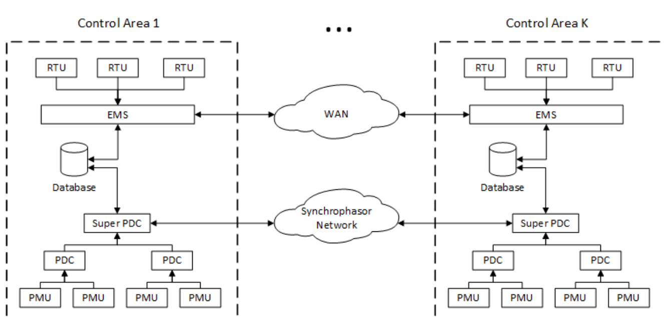


Fig. 1. Monitoring architecture of the smart grid.

The contributions of this paper are twofold. Firstly, we extend the centralized reduced-order formulation of the state estimation problem in [27] to a decentralized multi-area setting. We derive a solution to this problem based on distributed average consensus through network gossiping. Secondly, we present a detailed analytical treatment of the CGS information sharing approach introduced in [26]. We prove that it indeed achieves distributed average consensus. We also attempt to quantify the information exchange in the CGS scheme, and prove that this approach always results in reduced information exchange when compared with a naïve approach to network gossiping. We then go on to present the results of numerical simulations on the IEEE 118-bus test system showing the performance of the proposed reduced-order decentralized multi-area state estimator with CGS-based information exchange. We compare the proposed method with that of the method of [23] and found, in our simulations, that the proposed method renders lower mean-squared estimation error. We also compare the amount of information exchanged in both scheme and show that the proposed method is more scalable as the size of the interconnection grows and more areas are added to the system.

II. System Model

Consider an N-bus, balanced three-phase power system divided into K non-overlapping control areas as discussed in [8]. Let the index set of areas be $\mathcal{K}$. The minimal set of network parameters needed to represent the state of such a power system is the collection of the voltage magnitudes and phase angles at every bus in the power network [37]. These state variables constitute the system-wide state vector, denoted by $\mathbf{v} \in \mathbb{R}^{(2N-1)}$, where $\mathbf{v} = (\Re\{V_1\}, \ldots, \Re\{V_N\}, \Im\{V_2\}, \ldots, \Im\{V_N\})$, and $V_i,\ i = 1, \ldots, N$ are the positive sequence complex voltages at each bus. The phase angles at each bus are measured with respect to the GPS cosine signal when PMU measurements are used. In general, however, the phase angle of the reference bus is assumed to be zero, and therefore, $\Im\{V_1\} = 0$.

Let $N_1 < 2N - 1$ state variables be PMU observable [27]. We denote these components of $\mathbf{v}$ by $\mathbf{v}_p \in \mathbb{R}^{N_1}$. The remaining $N_2 = 2N - N_1 - 1$ PMU unobservable states are denoted by $\mathbf{v}_c \in \mathbb{R}^{N_2}$. Therefore, the state vector can also be written as

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_c \\ \mathbf{v}_p \end{bmatrix} \tag{1}$$

A. WAMS Architecture

PMUs output noisy measurements of the real and imaginary parts of the positive sequence voltage phasor at the bus to which they are connected, as well as the real and imaginary parts of the positive sequence currents flowing through all transmission lines connected that bus. The measurements from all PMUs are forwarded to a phasor data concentrator (PDC) which receives and time-synchronizes phasor data from multiple PMUs to produce a time-aligned output data stream. A PDC can exchange phasor data with PDCs at other locations. The PDCs communicate with each other through the synchrophasor network [38] as shown in Fig. 1. The PDC in Area i collects L_{p_i} measurements from all PMUs in Area i, denoted by $\mathbf{p}_i \in \mathbb{R}^{L_{p_i}}$. The measurement vector $\mathbf{p}_i$ is related to $\mathbf{v}_p$ as

$$\mathbf{p}_i = \mathbf{A}_i \mathbf{v}_p + \boldsymbol{\varepsilon}_i,\ \forall\ i \in \mathcal{K} \tag{2}$$

where $\mathbf{A}_i = \begin{bmatrix} \mathbf{B}_i^{\mathrm{T}} & \mathbf{Y}_i^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$. Each row of the matrix $\mathbf{B}_i$ is a unit vector of appropriate dimension with a 1 placed in the column associated with a particular voltage phasor at a PMU observable bus. The matrix $\mathbf{Y}_i$ is an admittance matrix of appropriate dimension corresponding to the current phasors [34]. The measurement error is denoted by a zero-mean random vector $\boldsymbol{\varepsilon}_i \in \mathbb{R}^{L_{p_i}}$ with covariance matrix $\mathbf{C}_{\varepsilon_i}$.

B. Legacy EMS/SCADA Architecture

The data acquisition part of SCADA system is responsible for acquiring measurements from remote terminal units (RTUs), while state estimation is performed in the EMS, as depicted in Fig. 1. The SCADA measurements consist of asynchronous power flow and injection measurements, as well as voltage magnitude measurements taken by remote terminal units placed throughout the interconnection. Let $\mathbf{c}_i \in \mathbb{R}^{L_i}$ denote a vector consisting of L_i measurements taken in Area i. We have

$$\mathbf{c}_i = \mathbf{g}_i(\mathbf{v}_c, \mathbf{v}_p) + \boldsymbol{\mu}_i \tag{3}$$

where $\mathbf{g}_i(\cdot)$ is a vector valued nonlinear function defined by Kirchhoff's laws, and the zero-mean random vector $\boldsymbol{\mu}_i$ represents measurement uncertainties associated with SCADA measurements. Since SCADA measurements are taken asynchronously, they are affected by time-skew errors and therefore, therefore SCADA measurements have much higher measurement uncertainties than PMU measurements [27].

III. Reduced-Order Decentralized MASE

The state estimation algorithm has two separate stages. In the first stage, the (Super-)PDCs obtain an estimate of the system-wide PMU observable states $\mathbf{v}_p$ by solving the weighted least-squares problem

$$\underset{\mathbf{v}_p}{\text{minimize}} \sum_{i\in\mathcal{K}} \|\mathbf{p}_i - \mathbf{A}_i\mathbf{v}_p\|^2_{\mathbf{C}^{-1}_{\epsilon_i}} \tag{4}$$

The normal equations of the above problem are given by

$$\left(\sum_{i\in\mathcal{K}} \mathbf{A}_i^{\mathrm{T}}\mathbf{C}^{-1}_{\epsilon_i}\mathbf{A}_i\right)\mathbf{v}_p = \sum_{i\in\mathcal{K}} \mathbf{A}_i^{\mathrm{T}}\mathbf{C}^{-1}_{\epsilon_i}\mathbf{p}_i \tag{5}$$

The PDC in each area $i \in \mathcal{K}$ initializes a so-called *gossip vector* [39], [40] $\mathbf{m}_i^{(0)} \triangleq \mathbf{A}_i^{\mathrm{T}}\mathbf{C}^{-1}_{\epsilon_i}\mathbf{p}_i$. Next the gossip vector of Area i is updated by combining it with the gossip vectors of the neighboring areas denoted by $\mathcal{N}_i$, i.e.,

$$\mathbf{m}_i^{(l+1)} = W_{ii}\mathbf{m}_i^{(l)} + \sum_{j\in\mathcal{N}_i} W_{ij}\mathbf{m}^{(l)} \tag{6}$$

for all $i \in \mathcal{K}$ and $l = 0, 1, \ldots$. The weights W_{ij} are elements in the ith row and jth column of a doubly stochastic matrix $\mathbf{W}^{K\times K}$ such that $W_{ij} = W_{ji} = 0$ if Area i is unable to share information with area j. If the spectral radius ρ of $\mathbf{W} - (1/K)\mathbf{1}\mathbf{1}^{\mathrm{T}}$ is less than 1, according to [41], we have

$$\lim_{l\to\infty} \mathbf{m}_i^{(l)} = \frac{1}{K}\sum_{j\in\mathcal{K}} \mathbf{m}_j^{(0)}, \ \forall i \in \mathcal{K} \tag{7}$$

which is the term in the left hand side of (5). These steps are summarized in Algorithm 1.

One advantage of estimating the PMU observable states separately is that they are estimated more frequently than PMU unobservable states. The term on the right hand side of (5) depends only on line parameters, and remains constant throughout several state estimation epochs. Therefore, it can be precomputed and stored in all the state estimators. This reduces the estimation of $\mathbf{v}_p$ to a matrix-vector multiplication. It also reduces the amount of information that needs to be exchanged. These factors allow for a near real-time view of the states of PMU observable buses to be presented to the system operator. Since the second-stage estimator is sensitive to the variance of $\hat{\mathbf{v}}_p$, it is recommended over a buffer several consecutive PMU measurement snapshots to reduce the estimation error as shown in [42], [43].

Algorithm 1 Estimating PMU observable states in Area i

given $\mathbf{p}_i$, $\mathbf{A}_i$, and $\mathbf{C}_{\varepsilon_i}$
set $l = 0$
initialize gossip vector $\mathbf{m}_i^{(l)} = \begin{bmatrix} \text{vec}\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{C}^{-1}_{\epsilon_i}\mathbf{A}_i\right)^{-1} \\ \mathbf{A}_i^{\mathrm{T}}\mathbf{C}^{-1}_{\epsilon_i}\mathbf{p}_i \end{bmatrix}$
repeat
  $\mathbf{m}_i^{(l+1)} = W_{ii}\mathbf{m}_i^{(l)} + \sum_{j\in\mathcal{N}_i} W_{ij}\mathbf{m}_j^{(l)}$
  $l \leftarrow l + 1$
until $\|\mathbf{m}_i^{(l)} - \mathbf{m}_i^{(l-1)}\|_\infty < \delta_p$
set $\mathbf{P} = \text{mat}\left(\mathbf{m}_i^{(l)}[0 : N_2^2 - 1]\right)$
set $\mathbf{q} = \mathbf{m}_i^{(l)}[N_2^2 : N_2(N_2 + 1)]$
compute $\hat{\mathbf{v}}_p = \mathbf{P}^{-1}\mathbf{q}$
Send $\hat{\mathbf{v}}_p$ to Stage II estimator in Area i

The second stage of the state estimation process involves estimating $\mathbf{v}_c$ from (3) by substituting $\hat{\mathbf{v}}_p$ for $\mathbf{v}_p$ in (3). This yields a nonlinear weighted least-squares problem in $\mathbf{v}_c$ given by

$$\underset{\mathbf{v}_c}{\text{minimize}} \sum_{i\in\mathcal{K}} \left\|\mathbf{R}_i^{-1/2}\left\{\mathbf{c}_i - \mathbf{g}_i(\mathbf{v}_c, \hat{\mathbf{v}}_p)\right\}\right\|^2 \tag{8}$$

where the weighting matrix $\mathbf{R}$ must take into account the estimation error of the first stage. The optimal weighting matrix is given by [27]

$$\mathbf{R}_i = \mathbf{C}_{\mu_i} + \boldsymbol{\Gamma}_{p_i}(\mathbf{v}_c, \hat{\mathbf{v}}_p)\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{C}^{-1}_{\epsilon_i}\mathbf{A}_i\right)^{-1}\boldsymbol{\Gamma}^{\mathrm{T}}_{p_i}(\mathbf{v}_c, \hat{\mathbf{v}}_p) \tag{9}$$

for all $i \in \mathcal{K}$, where $\boldsymbol{\Gamma}_{p_i}(\mathbf{v}_c, \mathbf{v}_p) = \partial\mathbf{g}_i(\mathbf{v}_c, \mathbf{v}_p)/\partial\mathbf{v}_p$. The problem is solved by employing Gauss-Newton iterations where the normal equations are written as (10) and (11).

$$\begin{aligned}\left(\sum_{i\in\mathcal{K}} \boldsymbol{\Gamma}^{\mathrm{T}}_{c_i}(\hat{\mathbf{v}}_c^{(j)}, \hat{\mathbf{v}}_p)\mathbf{R}_i^{-1/2}(\mathbf{v}_c^{(j)}, \hat{\mathbf{v}}_p)\boldsymbol{\Gamma}_{c_i}(\hat{\mathbf{v}}_c^{(j)}, \hat{\mathbf{v}}_p)\right)\boldsymbol{\Delta}\hat{\mathbf{v}}_c^{(j)} \\ = \sum_{i\in\mathcal{K}} \boldsymbol{\Gamma}^{\mathrm{T}}_{c_i}(\hat{\mathbf{v}}_c^{(j)}, \hat{\mathbf{v}}_p)\mathbf{R}_i^{-1/2}(\mathbf{v}_c^{(j)}, \hat{\mathbf{v}}_p)\left[\mathbf{c}_i - \mathbf{g}_i(\hat{\mathbf{v}}_c^{(j)}, \hat{\mathbf{v}}_p)\right]\end{aligned} \tag{10}$$

$$\hat{\mathbf{v}}_c^{(j+1)} = \hat{\mathbf{v}}_c^{(j)} + \boldsymbol{\Delta}\hat{\mathbf{v}}_c^{(j)} \tag{11}$$

Here again, the averages on either side of (10) are obtained by iterated network gossiping as described previously. The initial gossip vector for each Area i is given by

$$\mathbf{m}_i^{(0)} = \begin{bmatrix} \text{vec}\left(\boldsymbol{\Gamma}^{\mathrm{T}}_{c_i}(\hat{\mathbf{v}}_c^{(j)}, \hat{\mathbf{v}}_p)\mathbf{R}_i^{-1/2}(\mathbf{v}_c^{(j)}, \hat{\mathbf{v}}_p)\boldsymbol{\Gamma}_{c_i}(\hat{\mathbf{v}}_c^{(j)}, \hat{\mathbf{v}}_p)\right) \\ \boldsymbol{\Gamma}^{\mathrm{T}}_{c_i}(\hat{\mathbf{v}}_c^{(j)}, \hat{\mathbf{v}}_p)\mathbf{R}_i^{-1/2}(\mathbf{v}_c^{(j)}, \hat{\mathbf{v}}_p)\left[\mathbf{c}_i - \mathbf{g}_i(\hat{\mathbf{v}}_c^{(j)}, \hat{\mathbf{v}}_p)\right] \end{bmatrix} \tag{12}$$

where where $\boldsymbol{\Gamma}_{c_i}(\mathbf{v}_c, \mathbf{v}_p) = \partial\mathbf{g}_i(\mathbf{v}_c, \mathbf{v}_p)/\partial\mathbf{v}_c$, and j denotes the iteration index of the Gauss-Newton iterations. After each set of gossip iterations, the estimate $\hat{\mathbf{v}}_c^{(j)}$ is updated by evaluating (10) and (11). Subsequently, another set of gossip

Algorithm 2 Estimating PMU unobservable states in Area i

given $\mathbf{c}_i$, $\mathbf{g}_i(\cdot)$, $\hat{\mathbf{v}}_p$ and $\mathbf{C}_{\mu_i}$
set $j = 0$
set $\hat{\mathbf{v}}_c^{(j)}$ ▷ flat start
repeat
  set $l = 0$
  compute $\mathbf{R}_i$ according to (9)
  initialize gossip vector according to (12)
  repeat
    $\mathbf{m}_i^{(l+1)} = W_{ii}\mathbf{m}_i^{(l)} + \sum_{j \in \mathcal{N}_i} W_{ij}\mathbf{m}_j^{(l)}$
    $l \leftarrow l + 1$
  until $\|\mathbf{m}_i^{(l)} - \mathbf{m}_i^{(l-1)}\|_\infty < \delta_c$
  set $\mathbf{P} = \text{mat}\left(\mathbf{m}_i^{(l)}[0 : N_1^2 - 1]\right)$
  set $\mathbf{q} = \mathbf{m}_i^{(l)}[N_1^2 : N_1(N_1 + 1)]$
  compute $\hat{\mathbf{v}}_c^{(j+1)} = \mathbf{P}^{-1}\mathbf{q}$
  set $j \leftarrow j + 1$
until $\|\hat{\mathbf{v}}_c^{(j)} - \hat{\mathbf{v}}_c^{(j-1)}\|_\infty < \eta$

iterations are carried out. This continues until an appropriate stopping criterion is met, usually given by $\|\boldsymbol{\Delta}\hat{\mathbf{v}}_c^{(j)}\|_\infty < \eta$ where η is a predefined threshold. These steps are summarized in Algorithm 2

Thus, iteratively updating the local gossip vectors of a control area with a weighted sum of the gossip vectors of neighboring areas allows for the system-wide state vector to be obtained in each area. In principle, any gossip scheme would yield an accurate state estimate asymptotically. However, the speed of obtaining this state estimate can be vastly improved compared to the naïve gossiping approach employed in [23].

In particular, the state estimation scheme outlined in [23] uses the *uncoordinated random exchange* (URE) protocol [44] to perform gossiping. However, in practical power systems, most state variables are observed in only one area, and so-called *boundary variables* are observed by a few *topologically adjacent* areas, i.e., areas with at least one transmission line connecting them. Therefore, all areas should not participate in gossip iterations pertaining to state variables observed in only a few areas. In fact, the zeros contributed by these irrelevant areas to the gossip iterations will bias the average, which in turn slows down convergence of the gossip iterations. It would be more efficient to limit the gossip iterations to only those areas that observe the state variable, and transmit the consensus value to the remaining areas. In the following section, we develop such an information exchange protocol, wherein the information exchange is limited to only those areas that actually observe a state variable.

Bad Data Analysis

Bad data analysis is used as a way of identifying and eliminating the effect of gross errors in measurements. Two main approaches are seen in the literature: detection/identification and suppression. The most well known approaches to bad data detection identification are the χ^2-test and the largest normalized residue (LNR) test [37]. Fully distributed approaches to these methods have been presented in [45]. Methods for suppression include robust methods [46] or by manipulating the weights in the weighting matrix of the WLS formulation [47]. Additionally, from the literature on detecting cyber attacks on state estimation systems, several l_1-norm based approached have been proposed (see e.g., [48], [49]), which are also applicable for bad-data analysis.

In the proposed method, PMU measurements alone are used to estimate PMU observable states. This reduces measurement redundancy, causing some measurements to become critical measurements. Therefore gross errors in these measurements will become undetectable. Several steps can be taken to overcome this. Firstly, PMUs must be placed so as to maximize observability and result in the least number of critical measurements [50]. Secondly, since PMU measurements arrive at high frequency we can hold several consecutive measurements in a buffer and apply change detection and anomaly detection methods to identify corrupted measurements. Lastly, the covariance matrices $\mathbf{C}_{\varepsilon_i}$ can be adjusted according to the scheme in [47] to suppress the effects of bad data.

In addition to the above, conventional ways of bad-data detection like LNR can still be applied at the timescale of estimation of PMU unobservable states. This can be applied in a distributed manner since all the quantities required to calculate the normalized residuals are available in each area. The same is true of the normalized Lagrangian test for identifying parameter errors proposed in [51] and the multiple bad data detection scheme in [52]. Also, if the PMU measurements that observe a state variable are found to be corrupted, that state can be designated as PMU unobservable and estimated with conventional measurements.

IV. Clustered Gossip and Share Average Consensus

The information exchange protocol has two steps, namely, the clustered gossip stage, and the sharing stage. In the clustered gossip stage, groups of areas exchange information about the state variables they share. In the subsequent sharing stage, one of the areas in the group transmits the consensus value to the remaining areas. We show analytically that this scheme converges to the average of the initial gossip vectors of all areas.

A. Clustered Gossip Stage

The main idea behind the proposed method is the notion of a *cluster*. Let us denote the initial gossip vector associated with Area k in an N-bus system by $\mathbf{m}_k^{(0)} \in \mathbb{R}^S$ for all $k \in \mathcal{K}$, where $S = 2N(2N-1)$. Let $\mathbf{M} \in \mathbb{R}^{S \times K}$ be a matrix formed by $\mathbf{m}_k^{(0)}$ for all $k \in \mathcal{K}$ such that

$$\mathbf{M} \triangleq \begin{bmatrix} \mathbf{m}_1^{(0)} & \cdots & \mathbf{m}_K^{(0)} \end{bmatrix} \tag{13}$$

A key insight we develop here is that the rows of $\mathbf{M}$ are related to observations of state variables, while its columns represent the areas in which those measurements are taken. Let $\bar{\mathbf{m}}_k \in \mathbb{R}^K$, for all $k = 1, \ldots, S$ be columns of $\mathbf{M}^{\mathrm{T}}$, i.e., rows of $\mathbf{M}$. Furthermore, let the index set of all clusters be $\mathcal{J}$. Now, a cluster is defined as follows:

Definition 1. *For any $j \in \mathcal{J}$, a cluster $\mathcal{C}_j$ is a pair $(\mathcal{K}_j, \mathcal{S}_j)$ such that for a specific $\mathcal{K}_j \in 2^{\mathcal{K}}$,*

$$\mathcal{S}_j = \left\{ \bar{\mathbf{m}}_k \mid \bar{\mathbf{m}}_k = \sum_{i \in \mathcal{K}_j} a_i \mathbf{e}_{K_i}^T,\ k = 1, \ldots, S \right\} \tag{14}$$

where $a_i \in \mathbb{R},\ a_i \neq 0$, $2^{\mathcal{K}}$ is the power set of $\mathcal{K}$, and $\mathbf{e}_{K_i}$ is the ith canonical basis vector in $\mathbb{R}^K$.

Each set $\mathcal{S}_j$ consists of those rows of $\mathbf{M}$ with an identical pattern of non-zero elements, where the positions of the non-zero elements correspond to the areas in $\mathcal{K}_j$.

The gossip vector associated with each cluster is derived as follows: Let $\mathbf{H}_j \in \mathbb{R}^{|\mathcal{S}_j| \times K}$, be a matrix such that each element of $\mathcal{S}_j$ is a column of $\mathbf{H}_j^{\mathrm{T}}$, i.e.,

$$\mathbf{H}_j \triangleq \mathbf{T}_j \mathbf{M}, \quad j \in \mathcal{J} \tag{15}$$

where $\mathbf{T}_j \in \{0,1\}^{|S|_j \times S}$ is a selection matrix that selects those rows of $\mathbf{M}$ that belong to $\mathcal{C}_j$. The initial gossip vector associated with Area i in Cluster $\mathcal{C}_j$, denoted by $\mathbf{h}_{j_i}^{(0)}$, is the ith column of $\mathbf{H}_j$, i.e.,

$$\mathbf{h}_{j_i}^{(0)} = \mathbf{H}_j \mathbf{e}_{K_i} \quad i \in \mathcal{K}_j,\ j \in \mathcal{J} \tag{16}$$

We represent the connectivity model between the areas by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the vertex set $\mathcal{V}$ corresponds to areas, i.e, $\mathcal{V} = \mathcal{K}$, and the edge set $\mathcal{E} \subset \{(i,j) | \forall i, j \in \mathcal{V}\}$ and $(i,j) \in \mathcal{E}$ if and only if Areas i and j can exchange information with each other. Now, the connectivity model associated with each cluster $\mathcal{C}_j$ is denoted by $\mathcal{G}_j = (\mathcal{V}_j, \mathcal{E}_j)$, where $\mathcal{V}_j = \mathcal{K}_j$ and $\mathcal{E}_j \subseteq \{(i,j) | i,j \in \mathcal{K}_j,\ (i,j) \in \mathcal{E}\}$ for all $j \in \mathcal{J}$. Note that $\mathcal{G}_j$ is a subgraph of $\mathcal{G}$. The local state estimator in each area in Cluster $\mathcal{C}_j$ iteratively updates its gossip vector using its neighbors' gossip vectors such that

$$\mathbf{h}_{j_k}^{(l+1)} = W_{j_{kk}} \mathbf{h}_{j_k}^{(l)} + \sum_{i \in \mathcal{N}_{j_i} \setminus k} W_{j_{ki}} \mathbf{h}_{j_i}^{(l)},\ l = 0, 1 \ldots \tag{17}$$

for all $k \in \mathcal{K}_j$ and $j \in \mathcal{J}$, where the neighborhood of Area i in cluster $\mathcal{C}_j$ is defined as $\mathcal{N}_{j_i} \triangleq \{k \in \mathcal{K}_j | (i,k) \in \mathcal{E}_j\}$. The weights, $W_{j_{ki}}$, are elements of a weighting matrix $\mathbf{W}_j \in \mathbb{R}^{|K|_j \times |\mathcal{K}|_j}$. If $\mathbf{W}_j$ is doubly stochastic and $\rho(\mathbf{W}_j - (1/|\mathcal{K}_j|)\mathbf{1}\mathbf{1}^{\mathrm{T}}) < 1$ for all $j \in \mathcal{J}$, then

$$\mathbf{h}_j^* \triangleq \lim_{l \to \infty} \mathbf{h}_{j_i}^{(l)} = \frac{1}{|\mathcal{K}_j|} \sum_{k \in \mathcal{K}_j} \mathbf{h}_{j_k}^{(0)} \tag{18}$$

for all $j \in \mathcal{J}$ and $i \in \mathcal{K}_j$.

Remark: The CGS scheme is subject to more stringent connectivity requirements for convergence than naïve gossiping. The URE information exchange employed in [23] more robust to failure of communication links since it only requires that the system-wide connectivity graph $\mathcal{G}$ is connected. However, the convergence of each set of clustered gossip iterations requires that each local connectivity model $\mathcal{G}_j$ is connected for all $j \in \mathcal{J}$. In practice, however, this condition is easy to satisfy since clusters only contain *topologically adjacent* areas. Therefore, it is sufficient that communication links exist between such neighboring areas to satisfy the above condition.

Following the discussion in Section III, we need the average of the initial gossip vectors in each area to perform state estimation, i.e., $(1/K)\mathbf{M} \cdot \mathbf{1}$. We will now show that this is obtained by means of a permutation on the vector $\mathbf{h} \in \mathbb{R}^S$ defined as

$$\mathbf{h}^{\mathrm{T}} \triangleq \frac{1}{K} \begin{bmatrix} |\mathcal{K}_1| \mathbf{h}_1^{*^{\mathrm{T}}} & \cdots & |\mathcal{K}_j| \mathbf{h}_{|\mathcal{J}|}^{*^{\mathrm{T}}} \end{bmatrix} \tag{19}$$

Theorem 1. *Suppose $\mathbf{\Pi}$ is a permutation matrix with dimension S. If $\mathbf{h}$ and $\mathbf{M}$ are given by* (19) *and* (13)*, respectively, then $\mathbf{\Pi} \cdot \mathbf{h} = (1/K)\mathbf{M} \cdot \mathbf{1}$*

Proof. From (18) and (16) we have

$$\begin{aligned} |\mathcal{K}_j| \mathbf{h}_j^* &= \sum_{k \in \mathcal{K}_j} \mathbf{h}_{j_k}^{(0)} \\ &= \sum_{i \in \mathcal{K}} \mathbf{H}_j \mathbf{e}_{K_i} = \mathbf{H}_j \sum_{i \in \mathcal{K}} \mathbf{e}_{K_i} \end{aligned}$$

Substituting for $\mathbf{H}_j$ using (15), and since $\sum_{i \in \mathcal{K}} \mathbf{e}_{K_i} = \mathbf{1}$, we have

$$|\mathcal{K}_j| \mathbf{h}_j^* = \mathbf{T}_j \mathbf{M} \mathbf{1} \tag{20}$$

From (19) and (20), we have

$$\mathbf{h} = \frac{1}{K} \begin{bmatrix} |\mathcal{K}_1| \mathbf{h}_1^* \\ \vdots \\ |\mathcal{K}_{\mathcal{J}}| \mathbf{h}_{|\mathcal{J}|}^* \end{bmatrix} = \frac{1}{K} \begin{bmatrix} \mathbf{T}_1 \\ \vdots \\ \mathbf{T}_{|\mathcal{J}|} \end{bmatrix} \mathbf{M}\mathbf{1} \tag{21}$$

Definition 1 implies that $\mathcal{S}_j,\ \forall j \in \mathcal{J}$ are disjoint sets. Therefore, $\begin{bmatrix} \mathbf{T}_1^{\mathrm{T}} & \cdots & \mathbf{T}_{|\mathcal{J}|}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ is a permutation matrix. Hence, we have $\mathbf{\Pi}^{\mathrm{T}} \mathbf{h} = (1/K)\mathbf{M}\mathbf{1}$. □

Remark: In power systems, buses at the boundary of a control area are connected to neighboring areas through tie-lines. due to current limits on buses, it is very unlikely that a bus is connected to more than a few tie lines. As a result, $|\mathcal{K}_j|$ is very small compared to the number of areas in the system. This number remains small even as we increase the number of areas in the system. A cluster would typically contain a few neighboring areas which can all exchange information with each other. In this case, $|\mathcal{G}_j|$ becomes a fully connected graph and the clustered gossip iterations converge in one iteration.

B. Sharing Stage

Theorem 1 implies that in order to perform state estimation, the local state estimator in each control area requires $\mathbf{h}$. However, at the end of the clustered gossip stage, only a part of $\mathbf{h}$ is available to each local state estimator. In order for $\mathbf{h}$ to be available in each control area, $\mathbf{h}_j^*$ for each cluster in $\mathcal{J}$ must be relayed to every area outside that cluster. In principle, the local state estimator in any area of $\mathcal{K}_j$ can transmit its information, however, it makes sense to choose the local state estimator that accomplishes this in the most communication-efficient manner. A good choice is one with the smallest average distance (in hops) to all other areas outside the cluster.

Let us define a symmetric matrix $\mathbf{D} \in \mathbb{N}_0^{K \times K}$ such that an element D_{ij} is the smallest number of hops between Vertices i and j in the communication graph $\mathcal{G}$. Furthermore, let $\delta : \mathcal{K} \mapsto \mathbb{R}$ such that $\delta(i) = (1/K) \sum_{j \in \mathcal{K}} D_{ij}$. The area in $\mathcal{C}_j$ that must be chosen to transmit its gossip vector to all other areas, denoted by i_j^* is given by

$$i_j^* = \arg\min_{i \in \mathcal{K}_j} \delta(i), \quad j \in \mathcal{J} \tag{22}$$

Accordingly, Area $i^* \in \mathcal{K}_j$ transmits $\mathbf{h}_j^*$ to each area in $\mathcal{K} \setminus \mathcal{K}_j$ for all $j \in \mathcal{J}$. Now, the local state estimator in each control area has access to $\mathbf{h}$. By applying an appropriate permutation on $\mathbf{h}$, the average consensus value is recovered. The steps for performing clustered gossip and share average consensus at Area $i \in \mathcal{K}$ is shown in Algorithm 3

Algorithm 3 Clustered gossip and share in Area i

```
given m_i^(0), K_j , T_j ∀j ∈ J, J_i ≜ {j ∈ J | i ∈ K_j}
for each j ∈ J_i do
    h_{j_i}^(0) = T_j m_i^(0)
    set l = 0
    repeat
        h_{j_i}^(l+1) = W_{j_ii} h_{j_i}^(l) + Σ_{k∈N_{j_k}\i} W_{j_ik} h_{j_k}^(l)
        l ← l + 1
    until ||h_{j_i}^(l) − h_{j_i}^(l−1)||_∞ < δ_j
    Store h_j^* ← |K_j| h_{j_i}^(l)
    Nominate i_j^* according to (22)
    if i = i_j^* then
        Transmit h_j^* to K_j \ K
    end if
end for
```

C. Communication Cost

We will now quantify the information exchange involved in performing MASE using the clustered gossip and share (CGS) scheme. We assume that each element of a gossip vector can be encoded in one symbol. A measure of information exchange is the total number of requires symbols transmitted over a communication link per iteration.

Consider the best-case scenario for naïve gossiping using the URE protocol. This is when it behaves equivalently to centralized estimation, and convergence occurs in one iteration. This is also when the smallest amount of information that must be exchanged to perform state estimation, and happens when the following conditions hold We make following assumptions:

C1 The communication graph $\mathcal{G}$ is complete, i.e., $\mathcal{A}(\mathcal{G}) = \mathbf{1}\mathbf{1}^{\mathrm{T}} - \mathbf{I}$.

C2 The *maximum-degree* weighting scheme [53] is used to obtain the weighting matrix.

Equation (26) then becomes

$$\mathcal{I}_{URE} = SK(K-1) \tag{23}$$

Let n_j be the number of gossip iterations required to achieve consensus in $\mathcal{C}_j$. The amount of communication (in symbols) during the clustered gossip stage, denoted by $\mathcal{I}_{CG}$, is given by

$$\mathcal{I}_{CG} = \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}_j} n_j |\mathcal{S}_j| (|\mathcal{N}_{j_k}| - 1) \tag{24}$$

During the sharing stage, multiple communication links (hops) may be used to transmit information, as discussed previously. The amount of communication during the sharing stage, denoted by $\mathcal{I}_S$, is given by

$$\mathcal{I}_S = \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K} \setminus \mathcal{K}_j} |\mathcal{S}_j| |\mathcal{K} \setminus \mathcal{K}_j| D_{i_j^* k} \tag{25}$$

We now compare the total amount of information exchange using the CGS scheme with that of the URE scheme employed in [23]. Let the URE gossip iterations converge in n steps. This is denoted by $\mathcal{I}_{URE}$, and

$$\mathcal{I}_{URE} = n \sum_{k \in \mathcal{K}} S(|\mathcal{N}_k| - 1) \tag{26}$$

Theorem 2. *Suppose Conditions C1 and C2 hold. If $\mathcal{I}_{CG}$, $\mathcal{I}_S$ and $\mathcal{I}_{URE}$ are given by* (24), (25) *and* (26)*, respectively, then $\mathcal{I}_{CG} + \mathcal{I}_S \leq \mathcal{I}_{URE}$.*

Proof. Since $\mathcal{G}_j$ is a subgraph of $\mathcal{G}$, $\mathcal{G}_j$ is also complete. Therefore, we have $n_j = 1$, $\forall j \in \mathcal{J}$, and (24) and (25) become

$$\mathcal{I}_{CG} = \sum_{j \in \mathcal{J}} |\mathcal{S}_j| |\mathcal{K}_j| (|\mathcal{K}_j| - 1) \tag{27}$$

$$\mathcal{I}_S = \sum_{j \in \mathcal{J}} |\mathcal{S}_j| (K - |\mathcal{K}_j|) \tag{28}$$

Let $\mathcal{I}_{CGS} \triangleq \mathcal{I}_{CG} + \mathcal{I}_S$. From (27) and (28), we have

$$\begin{aligned} \mathcal{I}_{CGS} &= \sum_{j \in \mathcal{J}} |\mathcal{S}_j| |\mathcal{K}_j| (|\mathcal{K}_j| - 1) + \sum_{j \in \mathcal{J}} |\mathcal{S}_j| (K - \mathcal{K}_j) \\ &= \sum_{j \in \mathcal{J}} |\mathcal{S}_j| \left\{ (|\mathcal{K}_j| - 1)^2 + (K-1) \right\} \end{aligned} \tag{29}$$

Since $S = \sum_{j \in \mathcal{J}} |\mathcal{S}_j|$, we write (23) as

$$\begin{aligned} \mathcal{I}_{URE} &= \sum_{j \in \mathcal{J}} |\mathcal{S}_j| K(K-1) \\ &= \sum_{j \in \mathcal{J}} |\mathcal{S}_j| \left\{ (K-1)^2 + (K-1) \right\} \end{aligned} \tag{30}$$

Since $1 \leq |\mathcal{K}_j| \leq K$, comparing (29) and (30), we have $\mathcal{I}_{CG} + \mathcal{I}_S \leq \mathcal{I}_{URE}$

□

Remark: In practice, the majority of information is contained in singleton clusters, i.e., where the cluster has only one area. These clusters are not involved in the clustered gossip stage, where only information related to boundary buses is exchanged. Most state variables are internal and the elements of the gossip vector related to these variables is only transmitted in the sharing stage. As a result, the communication cost is drastically reduced when compared to the method of [23].

V. Numerical Results

In this section we present the results of simulations carried out on the IEEE 118-bus system. A load-flow calculation is performed using the MatPower toolbox in MATLAB [54] to update voltage magnitudes and phase angles throughout the system. The results of this load-flow calculation, i.e., the voltage magnitudes and phase angles, represent the true state variables of the system. Measurements are generated from these values using (2) and (3).

In the SCADA system, it is assumed that voltage magnitude measurements are taken at every bus in the system and real and reactive power flow measurements are taken on every transmission line. The measurement errors have zero-mean and are normally distributed. The standard deviations of voltage magnitude and power flow measurements are chosen to be 0.1% and 2% of their true value, respectively, following [55]. We assume that PMUs produce both voltage and current phasor measurements, and the standard deviation of PMU measurements is 0.01% of their true value.

We first compare the performance of the DARSE approach in [23] with the proposed reduced-order method. We use an IEEE 118 bus system with 9 areas. We place PMUs on buses 3, 15, 35, 25, 72, 47, 102, 81, and 53. The DARSE method consists of n gossip iterations followed by a Gauss-Netwon descent. Furthermore, note that DARSE follows a mixed measurement approach where PMU and conventional measurements are estimated together. The performance measure chosen is the mean squared error (MSE), given by

$$\text{MSE} = \frac{1}{2N}\left(\sum_{i=1}^{N_2}(\hat{v}_{p_i} - v_{p_i})^2 + \sum_{i=1}^{N_1}(\hat{v}_{c_i} - v_{c_i})^2\right) \quad (31)$$

The information exchange model of the SCADA/EMS systems of control areas is such that only adjacent areas can exchange information with each other. This is because, mainly due to historical reasons, the SCADA/EMS uses proprietary data models and database management systems to handle data, and such data cannot not be easily exchanged [1]. However, it is reasonable to assume that control areas with tie lines between them would have some means of exchanging data. When we employ this model, the clustered gossip stage converges in one iteration, since the cluster graph is fully connected.

Due to standardization efforts for PMUs, it is very likely that PDCs from several vendors will exchange data in a common format [38], [56]. In that case, the information exchange model of the WAMS typically would resemble a fully connected graph (every node is connected to all other nodes). In fact, this is equivalent to performing a centralized state estimation, and the results of Theorem 2 are applicable in this case. However, to demonstrate the efficacy of the CGS approach, we will assume that only adjacent areas can share information with each other.

Fig. 2 shows the MSE of the DARSE approach with 10 and 50 gossip iterations after each Gauss-Newton iteration. We compare this with the MSE of the reduced-order CGS method after each iteration. We see that the proposed method performs better in comparison with the DARSE approach. This is despite the fact that the reduced order approach incurs a performance penalty when there are no time-skew errors in conventional measurements [28].

Fig. 3 shows the convergence of the CGS gossip iterations and the URE gossip iterations for different weighting schemes. Columns of $\mathbf{M}_k$ are gossip vectors for each area at the kth gossip iteration. Let $\mathbf{M}^*$ be a matrix such that each column

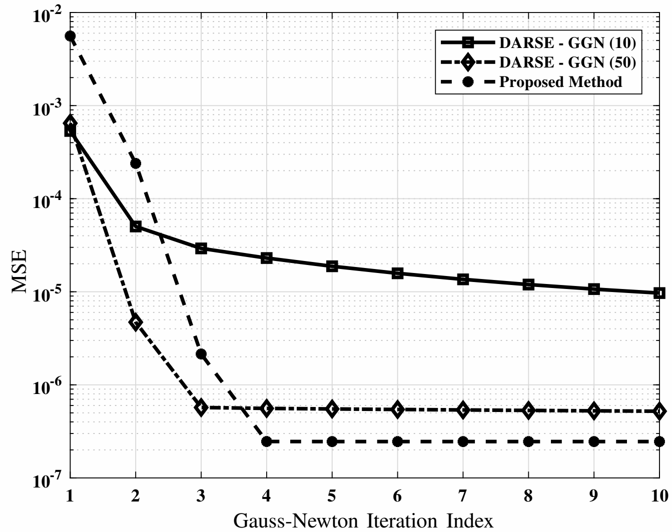


Fig. 2. Performance of the proposed method when compared with the DARSE scheme of [23] with 10 and 50 gossip iterations between Gauss-Newton descent steps.

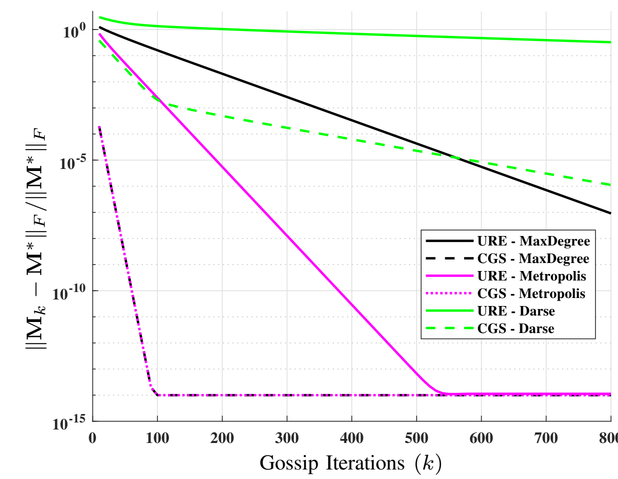


Fig. 3. Convergence of the CGS and URE schemes for different weighting matrices. 'Darse' indicates the weighting scheme employed in [23].

of $\mathbf{M}^*$ is the consensus gossip vector. We set the performance measure to be $\|\mathbf{M}_k - \mathbf{M}^*\|_F / \|\mathbf{M}^*\|_F$. We employ a communication graph that minimally satisfies the conditions for convergence of the CGS method. In this case, the edges of the communication graph are the edges of each cluster subgraph which form a Hamiltonian path for that subgraph. It is seen that in each case, the CGS method outperforms the URE method.

Employing the CGS information exchange scheme results in method is also more scalable than the URE scheme employed in [23]. We see from Figs. 5 and 4 that the convergence speed of the URE scheme used in [23] degrades considerably as the number of control areas in the system increases. This is because weighting matrix becomes sparse and many more gossip iterations are needed to achieve consensus. On the other hand, the CGS scheme does not incur any penalty as the system becomes larger since the local connectively model

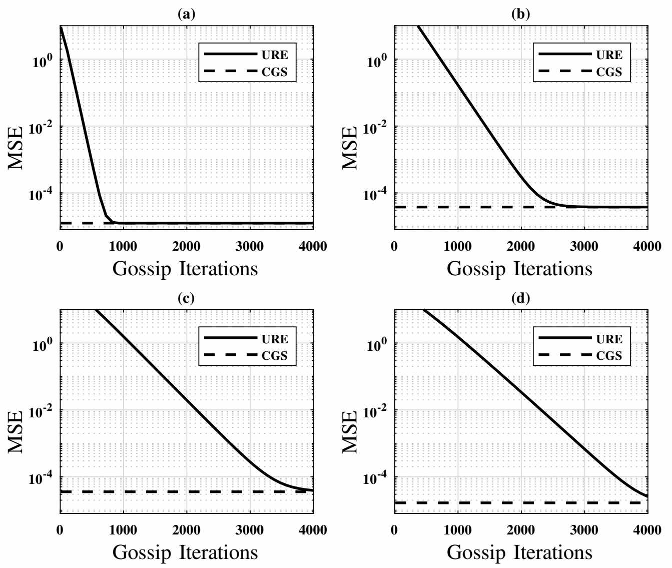


Fig. 4. Comparison of MSE of estimate of the PMU observable states obtained by WAMS for (a)4, (b)7, (c)11, and (d)14 areas when using CGS and URE gossip schemes.

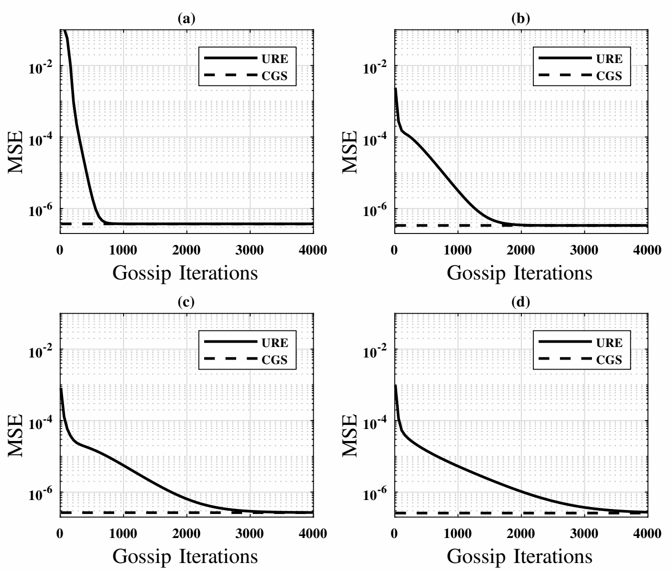


Fig. 5. Comparison MSE of estimate of the PMU unobservable states obtained by reduced order estimator for (a)4, (b)7, (c)11, and (d)14 areas when using CGS and URE gossip schemes.

(and hence the weighting matrix) remains dense.

Assume we employ a coding scheme where each symbol is encoded in 32 bits. With expressions derived in Section IV C, and the convergence rates of the URE scheme in Figs 5 and 4, we plot the total information exchange against the number of areas in Fig. 6. It is evident that employing the CGS scheme leads to a massive reduction in the amount of communication. Furthermore, we see that the amount of communication does not increase as much as traditional gossiping as the number of areas increases. This is because gossip iterations are localized to only a few areas, irrespective of the size of the wider interconnection. The increase in communication is because the number of areas that a cluster needs to share information with increases. In fact, we see that on average, each additional

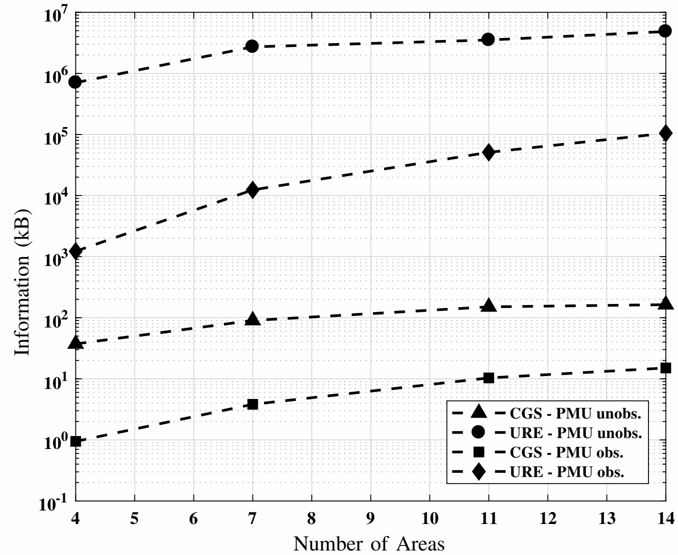


Fig. 6. Total amount of information exchanged in the synchrophasor network of the WAMS to arrive at a state estimate when employing CGS and URE schemes.

control area increases the amount of information by 12kB for the CGS scheme whereas that increase is 41.4 MB for the URE scheme.

VI. Conclusion

In this paper we presented a method to perform decentralized MASE where the PMU observable states are estimated by the PMU network, and the remaining states are estimated using SCADA measurements. We introduce a new protocol for information exchange between local state estimators of control areas. The scheme takes into account the weak measurement coupling between those areas, and estimation is performed such that the amount of communication is reduced. Additionally, by decoupling the SE functionality of the SCADA/EMS and the WAMS, we obtain benefits of improved numerical stability and robustness to time-skew errors. Our simulations show dramatic improvements in the convergence speed and communication load when compared to an existing networked gossip approach to MASE.

References

[1] 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] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. Begovic, and A. Phadke, "Wide-area monitoring, protection, and control of future electric power networks," *Proc. IEEE*, vol. 99, no. 1, pp. 80–93, Jan 2011.

[3] 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.

[4] 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, MayS 2012.

[5] K. Clements, O. Denison, and R. Ringlee, "A multi-area approach to state estimation in power system networks," in *Proc. IEEE PES Summer Meeting*. Paper C 72 465-3, 1972.

[6] H. Karimipour and V. Dinavahi, "Parallel domain decomposition based distributed state estimation for large-scale power systems," *IEEE Trans. Ind. Appl.*, vol. 52, no. 2, pp. 1265–1269, Mar. 2016.

[7] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, and T. Van Cutsem, "A taxonomy of multi-area state estimation methods," *Electr. Power Syst. Res.*, vol. 81, no. 4, pp. 1060–1069, 2011.

[8] T. Van Cutsem, J. Horward, and M. Ribbens-Pavella, "A two-level static state estimator for electric power systems," *IEEE Trans. Power App. Syst.*, vol. 100, no. 8, pp. 3722–3732, Aug. 1981.

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

[10] V. Donde, X. Feng, I. Segerqvist, and M. Callavik, "Distributed state estimation of hybrid ac/hvdc grids by network decomposition," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 974–981, Mar. 2016.

[11] Y. Guo, L. Tong, W. Wu, H. Sun, and B. Zhang, "Hierarchical multi-area state estimation via sensitivity function exchanges," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 442–453, Jan. 2017.

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

[13] A. 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.

[14] E. Caro, A. Conejo, and R. Minguez, "Decentralized state estimation and bad measurement identification: An efficient lagrangian relaxation approach," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2500–2508, Nov. 2011.

[15] P. Chavali and A. Nehorai, "Distributed power system state estimation using factor graphs," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2864–2876, June 2015.

[16] Z. Guo, S. Li, X. Wang, and W. Heng, "Distributed point-based Gaussian approximation filtering for forecasting-aided state estimation in power systems," *Power Systems, IEEE Transactions on*, vol. 31, no. 4, pp. 2597–2608, Jul.

[17] W. Jiang, V. Vittal, and G. Heydt, "Diakoptic state estimation using phasor measurement units," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1580–1589, 2008.

[18] ——, "A distributed state estimator utilizing synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 563 – 571, May 2007.

[19] V. O. and G. Dán, "Security of fully distributed power system state estimation: Detection and mitigation of data integrity attacks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1500–1508, Jul. 2014.

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

[21] H. Zhu and G. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1039–1050, Dec. 2014.

[22] A. Bose, K. Poon, and R. Emami, "Implementation issues for hierarchical state estimators," *PSERC Final Project Report*, Sep. 2010.

[23] X. Li and A. Scaglione, "Robust decentralized state estimation and tracking for power systems via network gossiping," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1184–1194, Jul. 2013.

[24] A. Dimakis, S. Kar, J. Moura, M. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847–1864, Nov. 2010.

[25] X. Li and A. Scaglione, "Convergence and applications of a gossip-based gauss-newton algorithm," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5231–5246, Nov. 2013.

[26] N. Kashyap, S. Werner, and Y.-F. Huang, "Decentralized power system state estimation with reduced inter-area communication," in *Proc. Intl. Conf. on Digital Signal Processing*, Singapore, Jul. 2015, pp. 14–18.

[27] N. Kashyap, S. Werner, Y.-F. Huang, and T. Riihonen, "Power system state estimation under incomplete PMU observability – A reduced-order approach," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1051–1062, Dec. 2014.

[28] N. Kashyap, S. Werner, T. Riihonen, and Y.-F. Huang, "Reduced-order synchrophasor-assisted state estimation for smart grids," in *Proc. 2012 IEEE 3rd Int. Conf. on Smart Grid Commun.*, Tainan City, Taiwan, pp. 605–610.

[29] T. Bi, X. Qin, and Q. Yang, "A novel hybrid state estimator for including synchronized phasor measurements," *Electric Power Systems Research*, vol. 78, no. 8, pp. 1343–1352, Aug. 2008.

[30] S. Chakrabarti, E. Kyriakides, G. Ledwich, and A. Ghosh, "Inclusion of PMU current phasor measurements in a power system state estimator," *IET Generation, Transmission Distribution*, vol. 4, no. 10, pp. 1104–1115, Oct. 2010.

[31] G. Valverde, S. Chakrabarti, E. Kyriakides, and V. Terzija, "A constrained formulation for hybrid state estimation," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1102–1109, Aug. 2011.

[32] G. N. Korres and T. A. Alexopoulos, "A constrained ordering for solving the equality constrained state estimation," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1998–2005, Nov. 2012.

[33] A. S. oes Costa, A. Albuquerque, and D. Bez, "An estimation fusion method for including phasor measurements into power system real-time modeling," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1910–1920, May 2013.

[34] M. Zhou, V. Centeno, J. Thorp, and A. Phadke, "An alternative for including phasor measurements in state estimators," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1930–1937, Nov. 2006.

[35] J. W. Gu, K. Clements, G. R. Krumpholz, and P. Davis, "The solution of ill-conditioned power system state estimation problems via the method of Peters and Wilkinson," *IEEE Power Eng. Rev.*, vol. PER-3, no. 10, pp. 43–44, 1983.

[36] E. Schweitzer and D. Whitehead, "Real-time power system control using synchrophasors," in *Proc. 2008 Annu. Conf. for Protective Relay Engineers*, pp. 78–88.

[37] A. Abur and A. Goméz-Expósito, *Power system state estimation: theory and implementation*. New York: Marcel Dekker, 2004.

[38] M. Patel, S. Aivaliotis, E. Ellen *et al.*, "Real-time application of synchrophasors for improving reliability," *NERC Report*, Oct. 2010.

[39] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. 4th Int. Symp. on Information Processing in Sensor Networks*. IEEE, 2005, pp. 63–70.

[40] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[41] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," in *Proc. 42nd IEEE Conf. on Decision and Control*, vol. 5, Maui, HI, USA, Dec. 2003, pp. 4997–5002.

[42] Q. Zhang, Y. Chakhchoukh, V. Vittal, G. Heydt, N. Logic, and S. Sturgill, "Impact of PMU measurement buffer length on state estimation and its optimization," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1657–1665, May 2013.

[43] N. Kashyap, S. Werner, and J. Lundén, "Signal-dependent preprocessing of buffered PMU measurements for hybrid state estimation," in *Proc. IEEE Intl. Instrumentation and Measurement Technology Conf.*, Taipei,Taiwan, May 2016, pp. 1–6.

[44] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48–61, Jan. 2009.

[45] D. H. Choi and L. Xie, "Fully distributed bad data processing for wide area state estimation," in *Proc. 2012 IEEE 3rd Int. Conf. on Smart Grid Commun.*, Tainan City, Taiwan, Oct 2011, pp. 546–551.

[46] L. Mili, M. G. Cheniae, and P. J. Rousseeuw, "Robust state estimation of electric power systems," *IEEE Trans. Circuits Syst. I*, vol. 41, no. 5, pp. 349–358, May 1994.

[47] H. M. Merrill and F. C. Schweppe, "Bad data suppression in power system static state estimation," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-90, no. 6, pp. 2718–2725, Nov. 1971.

[48] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[49] W. Xu, M. Wang, J. F. Cai, and A. Tang, "Sparse error correction from nonlinear measurements with applications in bad data detection for power networks," *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6175–6187, Dec. 2013.

[50] B. Gou and R. G. Kavasseri, "Unified PMU placement for observability and bad data detection in state estimation," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2573–2580, Nov. 2014.

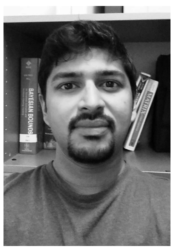
[51] Y. Lin and A. Abur, "A new framework for detection and identification of network parameter errors," *IEEE Trans. Smart Grid*, 2016.

[52] E. Caro, A. J. Conejo, R. Minguez, M. Zima, and G. Andersson, "Multiple bad data identification considering measurement dependencies," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 1953–1961, Nov. 2011.

[53] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Gossip algorithms: design, analysis and applications," in *Proc. IEEE Annu. Joint Conf. of the IEEE Comput. and Commun. Soc.*, vol. 3, March 2005, pp. 1653–1664.

[54] R. Zimmerman, C. Murillo-Sánchez, and R. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems

research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.
[55] A. Leite da Silva, M. Do Coutto Filho, and J. de Queiroz, "State forecasting in electric power systems," *IEE Proc. Gener. Transm. Distrib.*, vol. 130, no. 5, pp. 237–244, Sep. 1983.
[56] "IEEE standard for synchrophasor measurements for power systems," *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1–61, Dec. 2011.



Neelabh Kashyap Neelabh Kashyap (SM'06) received the M.Sc. degree in communications engineering from the Aalto University School of Electrical Engineering in April 2012. He is currently pursuing a D.Sc. degree at the Department of Signal Processing and Acoustics at the Aalto University School of Electrical Engineering.

Stefan Werner Stefan Werner (SM'07) received the M.Sc. degree in electrical engineering from the Royal Institute of Technology, Stockholm, Sweden, in 1998 and the D.Sc. degree (Hons.) in electrical engineering from the Signal Processing Laboratory, Helsinki University of Technology, Espoo, Finland, in 2002. He is currently a Professor in the Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway. He is also an Adjunct Professor with Aalto University in Finland, and an Adjunct Senior Research Fellow with the Institute for Telecommunications Research, University of South Australia. He was holding an Academy Research Fellowship funded by the Academy of Finland from 2009 to 2014. His research interests include adaptive and statistical signal processing, signal processing for communications, and smart electric power grids. He is a member of the editorial board for the EURASIP Journal of Signal Processing.

Yih-Fang Huang Yih-Fang Huang (F95) received his B.S.E.E. degree from National Taiwan University, M.S.E.E. degree from the University of Notre Dame, and Ph.D. degree from Princeton University. He is currently Professor of Electrical Engineering and Senior Associate Dean of the College of Engineering for Education and Undergraduate Programs at the University of Notre Dame. He served as Chair of the Department of Electrical Engineering at the University of Notre Dame from 1998 to 2006. In Spring 1993, he received the Toshiba Fellowship and was a Toshiba Visiting Professor at the Waseda University, Tokyo, Japan. From April to July 2007, he was a Visiting Professor at the Munich University of Technology, Munich, Germany. In Fall 2007, he was awarded the Fulbright-Nokia scholarship for lectures/research at the Helsinki University of Technology, Finland. Prof. Huang received the Golden Jubilee Medal of the IEEE Circuits and Systems Society in 1999, served as the Vice President for Publications in 199798 and was a Distinguished Lecturer for the same society in 20002001. At the University of Notre Dame, he received the Presidential Award in 2003, the Electrical Engineering departments Outstanding Teacher Award in 1994 and 2011, the Rev. Edmund P. Joyce, CSC Award for Excellence in Undergraduate Teaching in 2011, and the College of Engineerings Teacher of the Year Award in 2013. Dr. Huang's research interests focus on theory and applications of statistical signal detection and estimation, adaptive signal processing, and smart electric power grid.