Decentralized Power System State Estimation

Vassilis Kekatos and Georgios B. Giannakis
Dept. of ECE, University of Minnesota
Minneapolis, MN 55455, USA
Emails:{kekatos,georgios}@umn.edu

***Abstract*—Recent advances in metering technology, deregulation of the energy market, rapid penetration of renewables, and the smart grid vision for situational awareness, all call for a system-wide power system state estimation (PSSE). On the other hand, complexity of an interconnected power grid, privacy policies of local authorities, and vulnerability concerns preclude a centralized estimator. In this context, several possible manifestations of PSSE are treated in the present paper under a unified and systematic framework. Building on the alternating direction method of multipliers, a novel decentralized PSSE approach is developed. The obtained algorithm waives local observability concerns, leverages existing software, while requiring local system operators exchange minimal information only across tie lines. Numerical tests on two commonly used benchmarks, namely IEEE 14- and 118-bus power grids, show that the desired statistical accuracy is attained within 5-10 inter-area exchanges.**

***Index Terms*—Alternating direction method of multipliers; phasor measurement units; multi-area state estimation.**

I. Introduction

Power system state estimation (PSSE) has been traditionally performed at regional control centers with limited cross-regional interaction. Due to the recent deregulation of the electric energy market, large amounts of power are transfered over high-rate, long-distance lines spanning several control areas [6]. These so-called tie lines, originally constructed for emergency situations, are now fully operational and must be accurately monitored. Renewable sources further intensify inter-area power transfers, while they necessitate more frequent state acquisition. The latter can be now supported by the advanced metering infrastructure currently being installed, which includes phasor measurement units (PMU) [6].

Since any control area can be strongly affected by events and decisions elsewhere, independent system operators (ISO) can no longer operate in a truly independent fashion. Rather PSSE should be performed at an interconnection level. An interconnection typically consists of several ISOs and may include some thousands of buses, while 2-3 measurements per state are typically needed [6]. Requiring also real-time processing along with resilience to malicious data attacks on measurements make centralized state estimation computationally intractable. Further, a centralized approach is vulnerable and violates policy and privacy considerations. It is further worth mentioning that power grid information processing should be decentralized at several levels [6]: PMU measurements can be processed by phasor data concentrators (PDC) [10]; conventional measurements and PDC results can be aggregated by the ISO; and finally, ISO estimates can be merged at the interconnection level. For all the above reasons, distributed PSSE becomes an essential component of upcoming smart grids.

Distributed solutions were pursued since the statistical formulation of PSSE [11]. In [11, Part III], it was realized that for a chain of interconnected areas, Kalman filter-type updates readily apply. For an arbitrary area dependency graph though, a two-level approach is required [11]: States are locally estimated using solely the measurements related to them. Local estimates of shared variables, their associated covariance matrices, and tie line measurements are then forwarded and processed by a global coordinator. Several approximate variations of this protocol can be found in [14], [6], [7], but they assume local observability, i.e., the local state estimate derived after excluding boundary bus measurements is uniquely identifiable. Such an assumption may not hold due to bad data removal and/or because state estimation is now accomplished at lower grid hierarchies (e.g., substations or PDC). Additionally, the need for a coordinator hinders system's reliability, while such a protocol may be infeasible due to computational, communication, or policy limitations.

Decentralized solutions have been proposed too. Block Jacobi methods are developed in [3]. In [5], a decentralized algorithm is derived upon approximating the optimality conditions involved. But these methods require local observability too, while convergence is not always guaranteed. The auxiliary problem principle is employed in [4], where several parameters must be tuned. In [13] local observability is waived and each area is envisioned to maintain a copy of the interconnection-wide state: a first-order algorithm is proposed, but its linear convergence scales unfavorably with the interconnection size.

The PSSE problem, its unique requirements and challenges are presented in Section II. In Section III, a new distributed PSSE methodology is developed. Based on the alternating direction method of multipliers (AD-MoM) [2], a systematic way of cooperation between local control centers is provided having unique features: several practical PSSE formulations can be handled under the AD-MoM framework; the information exchange between neighboring areas is minimal; convergence to the centralized solution is guaranteed regardless of local observability; the algorithm can be executed by solvers already installed and currently in use by control centers. The

Dr. Kekatos' is supported by a Marie Curie International Outgoing Fellowship within the 7-th European Community Framework Programme (No. 234914). This work was also supported by NSF grants CCF-1016605, and ECCS-1002180. Dr. Kekatos is also with the Computer Engnr. & Informatics Dept., University of Patras, Greece.

numerical tests presented in Sec. IV illustrate that the method can essentially converge within 5-10 iterations. The paper is concluded in Sec. V. Regarding notation, lower- (upper-) case boldface letters denote column vectors (matrices); calligraphic letters stand for sets; and $(\cdot)^T$ denotes transposition.

II. Problem Formulation

Let an interconnected power system consist of K control areas. A control area is defined here as a subset of (geographically adjacent) buses supervised by a control center. The latter is able to (i) collect the electrical measurements recorded at area buses; (ii) reliably communicate with neighboring control centers; and (iii) accomplish specific computational tasks, such as solving a (linearly constrained) least-squares (LS) problem. At the other extreme, a control area here is not restricted to the conventional ISO region, but it can model other power system entities residing at lower grid hierarchies, e.g., a substation or a PDC [10]. A control area may degenerate even to a single bus under a micro-grid setup.

Suppose that the M_k measurements aggregated at the k-th area are stacked in $\tilde{\mathbf{z}}_k \in \mathbb{R}^{M_k}$, obeying

$$\tilde{\mathbf{z}}_k = h_k(\mathbf{x}_k) + \tilde{\mathbf{w}}_k \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^{N_k}$ contains the subset of the interconnected power system states involved in the measurements in $\tilde{\mathbf{z}}_k$; h_k is a vector of M_k functions; and $\tilde{\mathbf{w}}_k$ is an error term capturing measurement error and modeling inaccuracies. Error vectors $\{\tilde{\mathbf{w}}_k\}_{k=1}^K$ are assumed zero mean, having known covariance matrix $\tilde{\mathbf{\Sigma}}_k$, and independent across areas.

Functions $\{h_k(\mathbf{x}_k)\}_{k=1}^K$ depend on the system's admittance matrix and are in general nonlinear – except for some PMU data cases [10]. Performing state estimation with nonlinear h_k's entails solving non-convex optimization problems. Typically, such models are linearized iteratively via the Gauss-Newton method, or by resorting to the so called DC approximation [9], [1]. Hence, the focus here is on the special, yet computationally ubiquitous linear model where (1) becomes

$$\tilde{\mathbf{z}}_k = \tilde{\mathbf{H}}_k \mathbf{x}_k + \tilde{\mathbf{w}}_k \tag{2}$$

with $\tilde{\mathbf{H}}_k$ being a known real $M_k \times N_k$ matrix. Note that when the $\{h_k\}$ are originally nonlinear, (2) is the model assumed at a Gauss-Newton iteration. To simplify the presentation, measurements are premultiplied by the inverse square root of the noise covariance matrix, $\tilde{\mathbf{\Sigma}}_k^{-1/2}$, to yield

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k \tag{3}$$

where $\mathbf{z}_k := \tilde{\mathbf{\Sigma}}_k^{-1/2}\tilde{\mathbf{z}}_k$, $\mathbf{H}_k := \tilde{\mathbf{\Sigma}}_k^{-1/2}\tilde{\mathbf{H}}_k$, and the noise term $\mathbf{w}_k := \tilde{\mathbf{\Sigma}}_k^{-1/2}\tilde{\mathbf{w}}_k$ now is zero-mean but with identity covariance matrix. This prewhitening step is easily implemented and does not destroy the sparsity of $\mathbf{H}_k$ due to the usually (block) diagonal structure of $\tilde{\mathbf{\Sigma}}_k$.

PSSE could be performed locally at each area. Specifically, area k could aim solving

$$\min_{\mathbf{x}_k \in \mathcal{X}_k} f_k(\mathbf{x}_k; \mathbf{z}_k, \mathbf{H}_k) \tag{4}$$

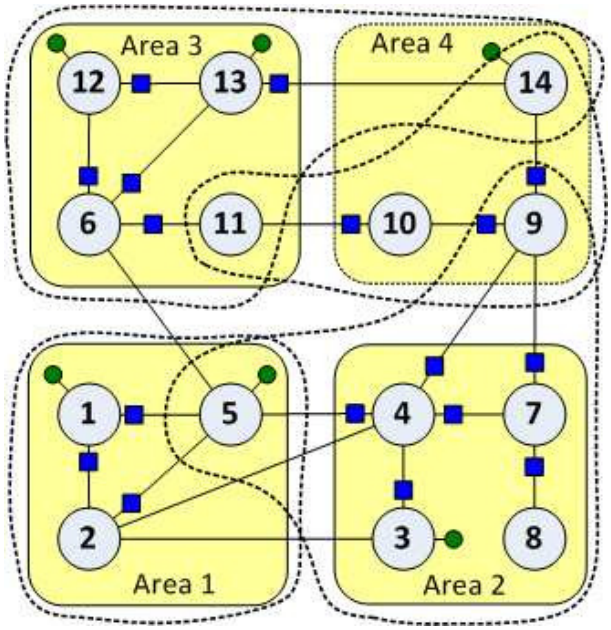

Fig. 1. The IEEE 14-bus system divided into four areas [12], [7]. The states belonging to $\mathbf{x}_k$'s are shown in dotted lassos. PMU bus voltage (line current) measurements shown as green circles (blue squares).

where $f_k(\cdot)$ is a convex function of $\mathbf{x}_k$ for the model in (3); and the convex set $\mathcal{X}_k$ captures possible prior information, such as zero-injection buses, short circuits, or operational limits [9], [1]. Typically, f_k is chosen as $\frac{1}{2}\|\mathbf{z}_k - \mathbf{H_k}\mathbf{x}_k\|_2^2$. For this choice, the minimizer of (4) is the least-squares estimate, that is the maximum-likelihood estimate (MLE) of $\mathbf{x}_k$ if $\mathbf{w}_k$ is Gaussian. To derive MLEs or facilitate bad data analysis, other forms of f_k could be envisioned. For notational simplicity, the dependence of f_k on $\mathbf{z}_k$ and $\mathbf{H}_k$ is henceforth dropped.

One of the characteristics of state estimation in power interconnections is that the local state vectors $\{\mathbf{x}_k\}_{k=1}^K$ overlap partially. To solidify that, consider the toy interconnection depicted in Fig. 1. Supposing that both PMU data (bus voltage and line current measurements) and interconnection states (bus voltages) are expressed in rectangular coordinates, the linear model of (3) is exact. Area 2 supervises buses $\{3, 4, 7, 8\}$. But, since it collects the electric current readings on lines $(7, 9)$ and $(4, 5)$, its state vector $\mathbf{x}_2$ contains not only $\{3, 4, 7, 8\}$, but also the voltages on buses $\{5, 9\}$. Hence, area 2 shares the states of bus 5 (9) with area 1 (4). Similarly, $\mathbf{x}_3$ and $\mathbf{x}_4$ overlap on buses 11 and 14. Notationally, let the $N \times 1$ vector $\mathbf{x}$ collect all the states. For every two areas, say k and l, let $\mathcal{S}_{kl}$ denote the set of their shared states. Let also $\mathbf{x}_k[l]$ ($\mathbf{x}_l[k]$) be the subvector of $\mathbf{x}_k$ ($\mathbf{x}_l$) consisting of their $|\mathcal{S}_{kl}|$ overlapping variables.

Solving the K problems of the form (4) in isolation is apparently suboptimal, let alone that control areas may be locally unobservable. Moreover, disagreement on boundary bus estimates over tie lines is another important limitation of solving (4) independently on a per-area basis. On the other hand, upon defining $\mathcal{X} := \{\mathbf{x} : \ \mathbf{x}_k \in \mathcal{X}_k \ \forall k\}$, jointly

optimizing

$$\min_{\mathbf{x}\in\mathcal{X}} \sum_{k=1}^{K} f_k(\mathbf{x}) \tag{5}$$

waives all these problems and can considerably improve estimation accuracy. Yet this comes at the expense of computational and communication burden, increased vulnerability, and disclosure of the internal system structure. Targeting the sweet spot between these two extremes, a decentralized method is proposed next.

III. Decentralized PSSE Approach

Tying the local tasks of (4) into a single optimization problem equivalent to (5) can be accomplished as

$$\min_{\{\mathbf{x}_k\in\mathcal{X}_k\}} \quad \sum_{k=1}^{K} f_k(\mathbf{x}_k) \tag{6}$$
$$\text{s.t.} \quad \mathbf{x}_k[l] = \mathbf{x}_l[k], \text{ for all } l \in \mathcal{N}_k, \ k = 1, \ldots, K$$

where $\mathcal{N}_k$ is the set of areas sharing states with area k.

The constraints of (6) force neighboring areas to consent on their shared variables, and thus, problems (6) and (5) are equivalent. But, the same constraints couple the optimization across areas. To enable a truly decentralized solution, an auxiliary variable denoted by $\mathbf{x}_{kl} \in \mathbb{R}^{|\mathcal{S}_{kl}|}$ is introduced per pair of interacting areas k, l. To keep the notation uncluttered, symbols $\mathbf{x}_{kl}$ and $\mathbf{x}_{lk}$ are used interchangeably for the same variable. Then, (6) is alternatively expressed as

$$\min_{\{\mathbf{x}_k\in\mathcal{X}_k\},\{\mathbf{x}_{kl}\}} \quad \sum_{k=1}^{K} f_k(\mathbf{x}_k) \tag{7}$$
$$\text{s.t.} \quad \mathbf{x}_k[l] = \mathbf{x}_{kl}, \text{ for all } l \in \mathcal{N}_k, \ k = 1, \ldots, K.$$

The novelty here is solving (7) using AD-MoM, a method with well-appreciated merits that has been successfully applied to several large-scale optimization problems; see [2] for a review. In AD-MoM, Lagrange multipliers $\mathbf{v}_{k,l} \in \mathbb{R}^{|\mathcal{S}_{kl}|}$ are introduced for each constraint of (7). Observe that $\mathbf{v}_{k,l}$ and $\mathbf{v}_{l,k}$ correspond to the distinct constraints $\mathbf{x}_k[l] = \mathbf{x}_{kl}$ and $\mathbf{x}_l[k] = \mathbf{x}_{kl}$, respectively. AD-MoM then exploits the method of multipliers concatenated with an iteration of the Gauss-Seidel algorithm. Specifically for (7), one first defines the augmented Lagrangian function

$$L\left(\{\mathbf{x}_k\},\{\mathbf{x}_{kl}\};\{\mathbf{v}_{k,l}\}\right) := \tag{8}$$
$$\sum_{k=1}^{K}\left[f_k(\mathbf{x}_k) + \sum_{l\in\mathcal{N}_k}\left(\mathbf{v}_{k,l}^T(\mathbf{x}_k[l]-\mathbf{x}_{kl}) + \frac{c}{2}\left\|\mathbf{x}_k[l]-\mathbf{x}_{kl}\right\|_2^2\right)\right]$$

where $c > 0$ is a predefined constant. Letting r denote the iteration index, AD-MoM cycles through three steps:

$$\{\mathbf{x}_k^{r+1}\} := \arg\min_{\{\mathbf{x}_k\in\mathcal{X}_k\}} L\left(\{\mathbf{x}_k\},\{\mathbf{x}_{kl}^r\};\{\mathbf{v}_{k,l}^r\}\right) \tag{9a}$$

$$\{\mathbf{x}_{kl}^{r+1}\} := \arg\min_{\{\mathbf{x}_{kl}\}} L\left(\{\mathbf{x}_k^{r+1}\},\{\mathbf{x}_{kl}\};\{\mathbf{v}_{k,l}^r\}\right) \tag{9b}$$

$$\mathbf{v}_{k,l}^{r+1} := \mathbf{v}_{k,l}^r + c\left(\mathbf{x}_k^{r+1}[l] - \mathbf{x}_{kl}^{r+1}\right) \text{ for all } k, l. \tag{9c}$$

During the first step, $\mathbf{x}_k$ are updated as the minimizers of the augmented Lagrangian function keeping $\mathbf{x}_{kl}$ and $\mathbf{v}_{k,l}$ fixed to their previous iteration values; $\mathbf{x}_{kl}$ and $\mathbf{v}_{k,l}$ can be initialized to zero. Similarly, $\mathbf{x}_{kl}$ are updated in (9b). Finally, (9c) is a simple gradient ascent of $L\left(\{\mathbf{x}_k^{r+1}\},\{\mathbf{x}_{kl}^{r+1}\};\{\mathbf{v}_{k,l}\}\right)$ using the step size c.

Inheriting AD-MoM features, the minimization in (9a) decouples over control areas. Moreover, by exploiting the problem structure, the iterations (9) can be greatly simplified as presented next and detailed in the Appendix.

Proposition 1. *The steps in* (9) *yield the same* $\mathbf{x}_k^r$ *iterates as the following steps*

$$\mathbf{x}_k^{r+1} := \arg\min_{\mathbf{x}_k\in\mathcal{X}_k} f_k(\mathbf{x}_k) + \frac{c}{2}\sum_{\substack{i=1\\ \mathcal{N}_k^i\neq\emptyset}}^{N_k} |\mathcal{N}_k^i|\left(x_k(i) - p_k^r(i)\right)^2, \ \forall k \tag{10a}$$

$$s_k^{r+1}(i) := \frac{1}{|\mathcal{N}_k^i|}\sum_{l\in\mathcal{N}_k^i} x_l^{r+1}[i], \ \forall i \text{ with } \mathcal{N}_k^i \neq \emptyset \tag{10b}$$

$$p_k^{r+1}(i) := p_k^r(i) + s_k^{r+1}(i) - \frac{x_k^r(i) + s_k^r(i)}{2}, \ \forall i, \mathcal{N}_k^i \neq \emptyset \tag{10c}$$

where $x_k(i)$ *is the* i*-th entry of* $\mathbf{x}_k$*; the set* $\mathcal{N}_k^i$ *consists of the areas sharing the variable* $x_k(i)$ *with area* k*; and* $x_l[i]$ *denotes the entry of* $\mathbf{x}_l$ *corresponding to* $x_k(i)$ *defined for all* $l \in \mathcal{N}_k^i$*. Regarding initialization, state variables* $\mathbf{x}_k$ *are set to arbitrary values* $\mathbf{x}_k^0$*;* $p_k^0(i)$ *are initialized to* $(x_k^0(i) + s_k^0(i))/2$*; and* $s_k^0(i)$ *as in* (10c)*.*

In lieu of a prior estimate, state variables can be initialized to the flat profile which assumes unitary real and zero imaginary parts for all bus voltages.

The minimization in (10a) and the simple update of (10c) are performed at the local centers. The averaging step of (10b) is accomplished either by a coordinator, or locally too. Notice though that either in the decentralized, or in the coordinated mode, the information revealed by area k is *minimal*: No measurements or regression matrices, but only the boundary bus states need to be exchanged, and only between the interested neighboring areas.

For notational convenience, define per area k a diagonal matrix $\mathbf{D}_k$ with (i,i)-th entry $|\mathcal{N}_k^i|$. Recall that by definition, for strictly local states $|\mathcal{N}_k^i|$ is zero. Also, define the N_k-dimensional vector $\mathbf{p}_k^r$ with i-th entry the $p_k^r(i)$ in (10c) when $|\mathcal{N}_k^i| > 0$, and 0 otherwise. Hence, the second term in the cost of (10a) is expressed as $\frac{c}{2}\left\|\mathbf{D}_k^{\frac{1}{2}}(\mathbf{x}_k - \mathbf{p}_k^r)\right\|_2^2$. For the typical case of unconstrained LS estimator (LSE), the minimizer of (10a) is clearly given by

$$\hat{\mathbf{x}}_k^{r+1} := \left(\mathbf{H}_k^T\mathbf{H}_k + c\mathbf{D}_k\right)^{-1}\left(\mathbf{H}_k^T\mathbf{z}_k + c\mathbf{D}_k\mathbf{p}_k^r\right) \tag{11}$$

which is an easy yet systematic modification of the local LSE. It can be shown that under mild conditions, the iterates of (10) converge to the optimal value of (7) [2], even when areas are not locally observable. As empirically observed, the algorithm of (10) can be terminated after a few iterations; cf. Sec. IV.

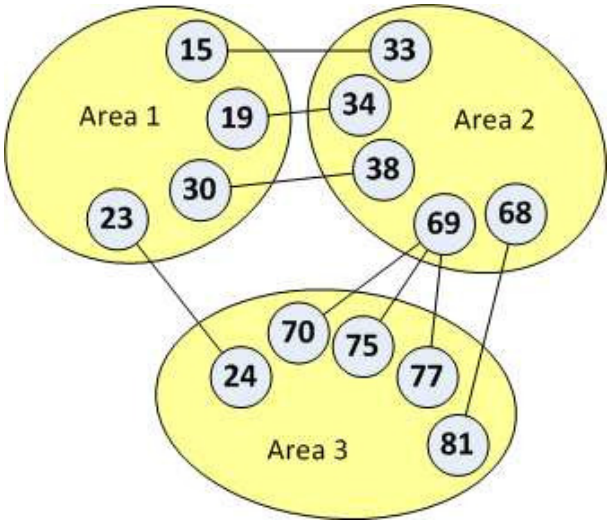

Fig. 2. The IEEE 118-bus power partitioned into three areas [12], [8]; only tie lines are shown here.

TABLE I
EMPIRICAL PER STATE STANDARD DEVIATION

Estimator	IEEE 14-bus grid	IEEE 118-bus grid
Internal LSE	$3.4 \cdot 10^{-3}$	$4.1 \cdot 10^{-4}$
Local LSE	$3.1 \cdot 10^{-3}$	$4.0 \cdot 10^{-4}$
Global LSE	$1.0 \cdot 10^{-3}$	$2.2 \cdot 10^{-4}$

IV. SIMULATED TESTS

In this section, the developed decentralized state estimator is numerically tested. Two power network benchmarks, namely the IEEE 14- and 118-bus systems are considered [12], while their admittance matrices and the underlying power system states are obtained using the MATPOWER software [15].

For both systems, the state vector contains the real and the imaginary parts of all bus voltages. Measurements consist of PMU recordings on bus voltages and line currents, expressed in rectangular coordinates too. Measurement noise is simulated as independent zero Gaussian with standard deviation per real component $\sigma_V = 0.01$ and $\sigma_I = 0.02$, for voltages and currents, respectively [15].

For the IEEE 14-bus network, PMU sites and types are shown in Fig. 1: 6 bus voltage- and 17 line current-meters yield a total of 46 measurements which translates to a redundancy ratio of 1.6 measurements per state [9]. For the IEEE 118-bus network, PMU sites are selected uniformly at random (including tie line current measurements): 77 bus voltage- and 205 line current-meters are assumed, yielding a redundancy ratio of 2.4. Regarding control areas, the IEEE 14-bus grid is partitioned into the 4 areas depicted in Fig. 1, while the IEEE 118-bus interconnection is split into 3 control areas as shown in Fig. 2 [8].

A reasonable question is whether interconnection-wide PSSE offers any improvement over local PSSE. To this end, three estimators are numerically compared: First, an estimator that uses only the measurements related to the states of the area, henceforth called "internal." Second, a "local" estimator which extends its state to boundary buses that can be reached via tie line measurements. Lastly, the interconnection-wide or "global" estimator. The empirical standard deviation per state is computed over 100 Monte Carlo runs. Table I lists the acquired standard deviations for the two power grids. The IEEE 118-bus grid attains better estimation accuracy due to its increased redundancy ratio. More importantly, the improvement of the local over the internal estimator is marginal, whereas the accuracy of the global estimator roughly doubles. This observation speaks for the importance of interconnection-wide PSSE even when local observability is guaranteed.

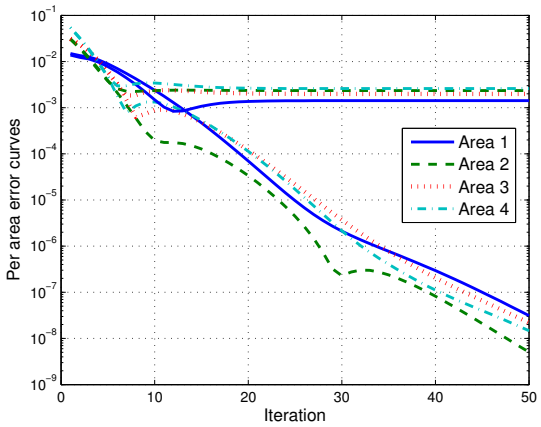

Fig. 3. Per area error curves $e^r_{k,c}$'s (bottom) and $e^r_{k,o}$'s (top) for the decentralized LSE of the IEEE 14-bus system of Fig. 1.

Having recognized the need for interconnection-wide PSSE, the decentralized LSE is evaluated next. State variables are initialized to the flat profile. Even though the iterations of (10) are guaranteed to converge to the centralized solution of (5) for any $c > 0$, the value of c affects the convergence rate. After scaling the problem data to obey the model in (3), c is empirically set to 10^4.

Two performance metrics are adopted: (i) the per area ℓ_2-norm distance to the centralized solution of (5), denoted by $e^r_{k,c} := \|\mathbf{x}^{(c)}_k - \mathbf{x}^r_k\|_2$; and (ii) the per area ℓ_2-norm distance to the true underlying state, defined as $e^r_{k,o} := \|\mathbf{x}_k - \mathbf{x}^r_k\|_2$.

Fig. 3 depicts the $e^r_{k,c}$ and $e^r_{k,o}$ curves obtained for the IEEE 14-bus network. As indicated by the curves, the decentralized minimizer approaches the centralized one to an accuracy of 10^{-3} in 10 iterations. Interestingly, the accuracy dictated by the measurements has been reached within 10-15 iterations. Another important issue is the consent of neighboring areas on the values of shared estimates. Fig. 4 shows the differences on the real part of shared bus voltages across areas. The curves indicate that only 5-10 iterations suffice for adjacent areas to consent.

To evaluate the new algorithm in scenarios where local observability does not hold, the electric current measurement on line $(6, 11)$ is removed from the IEEE 14-bus measurement set (cf. Fig. 1). Since the only measurement directly related

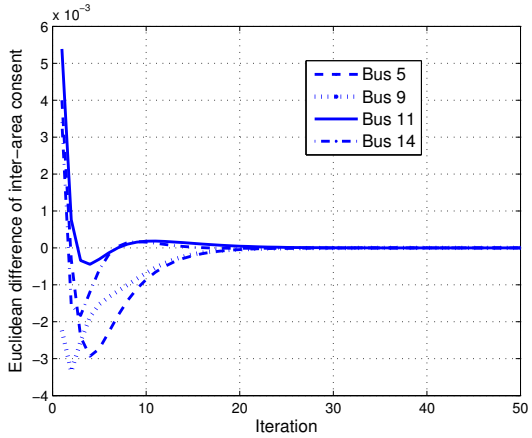

Fig. 4. Voltage difference (real part) between shared buses.

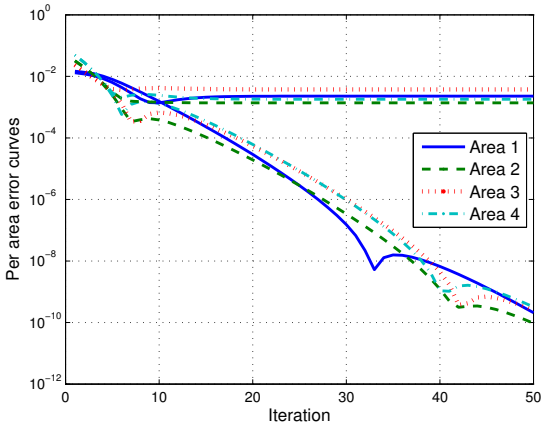

Fig. 5. Per area error curves $e^r_{k,c}$'s (bottom) and $e^r_{k,o}$'s (top) for the decentralized LSE of the IEEE 14-bus system without local observability.

to bus 11 is the current measurement on line $(10, 11)$ and that is collected by control area 4, area 3 is locally unobservable. The error curves obtained and plotted in Fig. 5 verify that the new method does not require local observability.

Switching to the IEEE 118-bus benchmark, similar results are observed. As evidenced by the $e^r_{k,c}$ and $e^r_{k,o}$ depicted in Fig. 6, the decentralized solution attains the desired statistical accuracy within only 5-10 iterations.

V. Conclusions

Decentralized PSSE has been treated here in a unified and systematic manner. The proposed algorithm waives local observability requirements and maintains backward compatibility. Within few iterations, local control centers can acquire highly accurate estimates for the parts of the state which are of interest to them, without violating privacy policies. Building on this approach, several exciting research issues emerge, such as extensions to robust and generalized PSSE, which are of extreme practical interest for realizing the forthcoming smart grid vision.

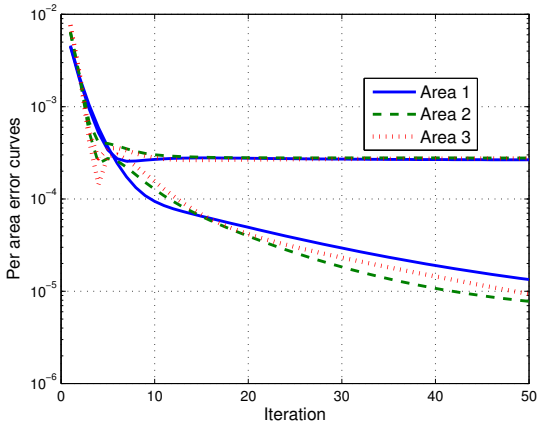

Fig. 6. Per area error curves $e^r_{k,c}$'s (bottom) and $e^r_{k,o}$'s (top) for the decentralized LSE of the IEEE 118-bus system of Fig. 2.

Appendix

A useful lemma is shown first.

Lemma 1. *For every pair of adjacent areas k and l, the Lagrange multipliers updated by* (9c) *satisfy $\mathbf{v}^r_{k,l} + \mathbf{v}^r_{l,k} = \mathbf{0}$ per iteration $r > 0$.*

Proof: Note that step (9b) decouples over the $\mathbf{x}_{kl}$'s as

$$\min_{\mathbf{x}_{kl}} -\mathbf{x}_{kl}^T\left(\mathbf{v}^r_{k,l}+\mathbf{v}^r_{l,k}\right)+\frac{c}{2}\left\|\mathbf{x}_{kl}-\mathbf{x}_k^{r+1}[l]\right\|_2^2 + \frac{c}{2}\left\|\mathbf{x}_{kl}-\mathbf{x}_l^{r+1}[k]\right\|_2^2, \text{ for all } k,l,$$

whose minimizer can be easily shown to be

$$\mathbf{x}_{kl}^{r+1} := \left(\frac{\mathbf{x}_k^{r+1}[l]+\mathbf{x}_l^{r+1}[k]}{2}\right)+\left(\frac{\mathbf{v}^r_{k,l}+\mathbf{v}^r_{l,k}}{2c}\right). \quad (12)$$

Next, consider the updates of $\mathbf{v}_{k,l}$ and $\mathbf{v}_{l,k}$ according to step (9c). Adding the two updates by parts and solving for the common term $\mathbf{x}_{kl}^{r+1}$, yields

$$\mathbf{x}_{kl}^{r+1} = \left(\frac{\mathbf{v}^r_{k,l}+\mathbf{v}^r_{l,k}}{2c}\right)-\left(\frac{\mathbf{v}^{r+1}_{k,l}+\mathbf{v}^{r+1}_{l,k}}{2c}\right) + \left(\frac{\mathbf{x}_k^{r+1}[l]+\mathbf{x}_l^{r+1}[k]}{2}\right). \quad (13)$$

By equating the right-hand sides of (12) and (13), the claim of the lemma follows readily. ■

Proof of Proposition 1: The optimization in (9a) is separable across areas. Upon completing the squares, the optimization task for area k during step (9a) becomes

$$\min_{\mathbf{x}_k} f_k(\mathbf{x}_k)+\frac{c}{2}\sum_{l\in\mathcal{N}_k}\left\|\mathbf{x}_k[l]-\left(\mathbf{x}^r_{kl}-\frac{\mathbf{v}^r_{k,l}}{c}\right)\right\|_2^2. \quad (14)$$

Apparently, the ℓ_2-norms in (14) decouple over the entries of the vectors involved. However, a single entry of $\mathbf{x}_k$, say $x_k(i)$, may be shared not only between areas k and l, but rather among area k and all the areas in $\mathcal{N}^i_k$. If $x_{kl}[i]$ $(v_{k,l}[i])$

denotes the entry of $\mathbf{x}_{kl}$ ($\mathbf{v}_{k,l}$) corresponding to $x_k(i)$, the optimization in (14) can be expressed as

$$\min_{\mathbf{x}_k} f_k(\mathbf{x}_k) + \frac{c}{2} \sum_{\substack{i \in \mathcal{N}_k \\ \mathcal{N}_k^i \neq \emptyset}} |\mathcal{N}_k^i| \left(x_k(i) - p_k^{r+1}(i)\right)^2 \tag{15}$$

where for all k, and $i = 1, \ldots, N_k$ with $\mathcal{N}_k^i \neq \emptyset$,

$$p_k^{r+1}(i) := \frac{1}{|\mathcal{N}_k^i|} \sum_{l \in \mathcal{N}_k^i} \left(x_{kl}^r[i] - \frac{v_{k,l}^r[i]}{c} \right). \tag{16}$$

By Lemma 1, step (9b) simplifies to

$$\mathbf{x}_{kl}^{r+1} = \frac{1}{2} \left(\mathbf{x}_k^{r+1}[l] + \mathbf{x}_l^{r+1}[k] \right). \tag{17}$$

In other words, the auxiliary variable $\mathbf{x}_{kl}$ is the average of the shared state variables across areas k and l per iteration. Based on (17), step (9b) can be dropped after eliminating the auxiliary variables $\mathbf{x}_{kl}$ from the updates of (16) and (9c). Hence, one arrives at the iterates

$$\mathbf{x}_k^{r+1} := \arg\min_{\mathbf{x}_k} f_k(\mathbf{x}_k) + \frac{c}{2} \sum_{\substack{i \in \mathcal{N}_k \\ \mathcal{N}_k^i \neq \emptyset}} |\mathcal{N}_k^i| \left(x_k(i) - p_k^r(i)\right)^2 \tag{18a}$$

$$\mathbf{v}_{k,l}^{r+1} := \mathbf{v}_{k,l}^r + c \left(\frac{\mathbf{x}_k^{r+1}[l] - \mathbf{x}_l^{r+1}[k]}{2} \right) \tag{18b}$$

$$p_k^{r+1}(i) := \frac{1}{2} \left(x_k^{r+1}(i) + \frac{1}{|\mathcal{N}_k^i|} \sum_{l \in \mathcal{N}_k^i} x_l^{r+1}[i] \right) - \frac{1}{|\mathcal{N}_k^i|} \sum_{l \in \mathcal{N}_k^i} \frac{v_{k,l}^{r+1}[i]}{c}. \tag{18c}$$

To further simplify the iterations, define the average of the shared variable $x_k(i)$'s copies over $\mathcal{N}_k^i$ as

$$s_k^r(i) := \frac{1}{|\mathcal{N}_k^i|} \sum_{l \in \mathcal{N}_k^i} x_l^r[i] \tag{19}$$

as well as the average of the weighted Lagrange multipliers $u_k^r(i) := \sum_{l \mathcal{N}_k^i} v_{k,l}^r[i]/(c|\mathcal{N}_k^i|)$. Then, (18c) can be written as

$$p_k^{r+1}(i) := \frac{1}{2} \left(x_k^{r+1}(i) + s_k^{r+1}(i) \right) - u_k^{r+1}(i). \tag{20}$$

With the $\mathbf{v}_{k,l}$'s initialized to zero, the $u_k(i)$'s can be recursively updated as

$$u_k^{r+1}(i) := u_k^r(i) + (x_k^{r+1}(i) - s_k^{r+1}(i))/2.$$

Hence, update (20) can be alternatively performed as

$$p_k^{r+1}(i) := p_k^r(i) + s_k^{r+1}(i) - \frac{x_k^r(i) + s_k^r(i)}{2}. \tag{21}$$

Collecting (18a), the definition in (19), and the recursion of (21), the algorithm of (10) is derived. ■

REFERENCES

[1] A. Abur and A. Gomez-Exposito, *Power System State Estimation: Theory and Implementation*. New York, NY: Marcel Dekker, 2004.

[2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach Learning*, vol. 3, pp. 1–122, 2010.

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

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

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

[6] A. Gómez-Exposito, A. Abur, A. de la Villa Jaén, and C. Gomez-Quiles, "A multilevel state estimation paradigm for smart grids," *Proc. IEEE*, vol. 99, no. 6, pp. 952–976, Jun. 2011.

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

[8] L. Min and A. Abur, "Total transfer capability computation for multi-area power systems," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1141–1147, Aug. 2006.

[9] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.

[10] A. G. Phadke and J. S. Thorp, *Synchronized Phasor Measurements and Their Applications*. New York, NY: Springer, 2008.

[11] F. C. Schweppe, J. Wildes, and D. Rom, "Power system static state estimation: Parts I, II, and III," *IEEE Trans. Power App. Syst.*, vol. 89, pp. 120–135, Jan. 1970.

[12] Power systems test case archive. University of Washington. [Online]. Available: http://www.ee.washington.edu/research/pstca/

[13] L. Xie, D.-H. Choi, and S. Kar, "Cooperative distributed state estimation: Local observability relaxed," in *Proc. IEEE PES General Meeting*, Detroit, MI, Jul. 2011.

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

[15] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. 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.