Power System Nonlinear State Estimation Using Distributed Semidefinite Programming

Hao Zhu, *Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—State estimation (SE) is an important task allowing power networks to monitor accurately the underlying system state, which is useful for security-constrained dispatch and power system control. For nonlinear AC power systems, SE amounts to minimizing a weighted least-squares cost that is inherently nonconvex, thus giving rise to many local optima. As a result, estimators used extensively in practice rely on iterative optimization methods, which are destined to return only *locally* optimal solutions, or even fail to converge. A semidefinite programming (SDP) formulation for SE has been advocated, which relies on the *convex* semidefinite relaxation (SDR) of the original problem and thereby renders it efficiently solvable. Theoretical analysis under simplified conditions is provided to shed light on the near-optimal performance of the SDR-based SE solution at polynomial complexity. The new approach is further pursued toward complementing traditional nonlinear measurements with linear synchrophasor measurements and reducing computational complexity through distributed implementations. Numerical tests on the standard IEEE 30- and 118-bus systems corroborate that the SE algorithms outperform existing alternatives, and approach near-optimal performance.**

***Index Terms*—Distributed state estimation, phasor measurement units, power system state estimation, semidefinite relaxation.**

I. Introduction

THE electric power grid is a complex system consisting of multiple subsystems, each with a transmission infrastructure spanning over a huge geographical area, transporting energy from generations to distribution networks. Monitoring the operational conditions of grid transmission networks is of paramount importance to facilitate system control and optimization tasks, including security analysis and economic dispatch under security constraints; see e.g., [1, Ch. 1] and [10], [25]. For this purpose, various system variables are measured at selected nodes and then transmitted to the control center for estimating the system state variables, namely complex bus voltages at all buses in the grid.

Manuscript received October 01, 2013; revised March 06, 2014; accepted June 11, 2014. Date of publication June 13, 2014; date of current version November 18, 2014. The work of H. Zhu was supported in part by the University of Minnesota Doctoral Dissertation Fellowship. The work of G. B. Giannakis was supported in part by the Institute of Renewable Energy and the Environment under Grant RL-0010-13. The guest editor coordinating the review of this manuscript and approving it for publication was Dr. Shalinee Kishore.

H. Zhu is with the Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL 61801 USA (e-mail: haozhu@illinois.edu).

G. B. Giannakis is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@umn.edu).

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

Digital Object Identifier 10.1109/JSTSP.2014.2331033

Because meter measurements are *nonlinearly* related with state variables, the power system *state estimation* (SE) task is inherently nonconvex giving rise to many local optima. Specifically, SE amounts to solving the associated nonlinear (weighted) least-squares (LS) problem using the Gauss-Newton iterations as the algorithmic foundation; see e.g., [1, Ch. 2], [25]. Since this iterative procedure is in fact related to gradient descent solvers for nonconvex problems, it inevitably faces convergence issues and sensitivity to initialization [4, Sec. 1.5]. Without guaranteed convergence to the global optimum, existing variants have asserted numerical stability or robustness to outliers [1, Chs. 3–6], but they can only improve the linearized error cost per iteration.

At least as important, availability of suitable initializations becomes increasingly difficult with the penetration of renewable energy sources (RES). The reason is two-fold. First, the intermittency of RES leads to growing dynamics of the system state, which challenge usage of temporal information for initializing the SE. Secondly, emerging distributed energy resources (DERs) advocate the importance of SE for distribution networks, in which increased resistance-to-inductance ratios render voltage magnitudes nonflat [9, Ch. 6]. (This is different from standard settings where flat voltage initializations are common for the linear DC flow approximation.)

The latest trend for SE is to incorporate *linear* state measurements offered by synchronized phasor measurement units (PMUs) to develop the so-termed *hybrid* SE; see e.g., [13], [34] and references therein. However, limited PMU deployment currently confines SE to mostly rely on the nonlinear legacy meter measurements, and its companion Gauss-Newton iterative methods. Furthermore, *distributed* SE among multiple control areas are strongly motivated by the deregulation of energy markets, where large amounts of power are transferred among areas over the tie-lines at increasing rates [12]. This is part of the system-level institutional changes aiming at an interconnected network with improved reliability. Since each control area can be strongly affected by events and decisions elsewhere, regional operators can no longer operate in a truly independent fashion. At the same time, central processing of the current energy-related tasks faces several limitations: i) vulnerability to unreliable telemetry; ii) high computational complexity at a single control center; and iii) data security and privacy concerns of regional operators. However, most works on multi-area SE relies on linearized models or iteratively approximating to linear ones; see e.g., [17], [32] and references therein. Hence, with nonlinear (possibly along with linear) measurements the grand challenge remains to develop (distributed) solvers approaching the *global optimum* at *polynomial-time* complexity.

The goal of this paper is to introduce such polynomial-time SE algorithms for AC power systems, having the potential to attain a near-optimal state estimate. Challenged by the nonconvexity in SE, the proposed approaches leverage a well-appreciated optimization technique called *semidefinite relaxation* (SDR) that surrogates nonconvex problems by semidefinite programming (SDP) ones. SDR has been applied successfully to various areas [11], including signal processing and communications; see e.g., [24]. Recently, SDR has been proposed for solving the optimal power flow (OPF) problem [2]; see also [21], [22] for potential guarantee on the global optimality. Compared to our conference precursors [35], [37] where (distributed) SDR was first leveraged for SE, the analysis of the present paper also offers insights on near-optimality of SDR-based SE, as well as useful extensions to include synchrophasor data, and extensive test on larger benchmark systems. Recent works on using SDR for (distributed) SE have been carried out independently [29], [30], which appeared after our works [35], [37]. While [29] only deals with the centralized SDR-based SE as in [35] with no performance analysis or synchrophasor measurements, the iterative SE scheme in [30] requires solving an outer problem of dimension equal to the centralized SDP problem, and thus it is not fully distributed.

The rest of the paper is organized as follows. The SE problem is introduced in Section II. The relaxed SDP formulation is the subject of Section III, which enables computationally efficient solvers regardless of initialization, along with an ideal scenario for achieving the global optimum. Section IV generalizes the SDR-based SE formulation to include linear PMU measurements, while Section V develops a multi-area SDR-SE method using an effective distribute optimization technique. Several numerical tests presented in Section VI corroborate the near-optimal performance of the proposed SDR approaches relative to the Gauss-Newton method. The paper is wrapped up in Section VII.

Notation: Upper (lower) boldface letters will be used for matrices (column vectors); $(\cdot)^{\mathcal{T}}$ denotes transposition; $(\cdot)^{\mathcal{H}}$ complex-conjugate transposition; $\mathcal{R}(\cdot)$ the real part; $\mathcal{I}(\cdot)$ the imaginary part; $\mathrm{Tr}(\cdot)$ the matrix trace; $\mathrm{rank}(\cdot)$ the matrix rank; $\mathbf{0}$ the all-zero matrix; $\|\cdot\|_F$ the matrix Frobenius norm; $\|\cdot\|_p$ the vector p—norm for $p \geq 1$; and $|\cdot|$ the magnitude of a complex number.

II. Modeling and Problem Formulation

Consider a power network with N buses denoted by the set $\mathcal{N} := \{1, \ldots, N\}$, and all lines represented by $\mathcal{E} := \{(n, m)\}$. To estimate the complex voltage V_n per bus $n \in \mathcal{N}$ (in *rectangular* as opposed to *polar* coordinate in most SE literature [1], [25]), a subset of following system variables[1] are measured:

- $P_n(Q_n)$: the real (reactive) power injection at bus n (negative if bus n is connected to a load);
- $P_{mn}(Q_{mn})$: the real (reactive) power flow from bus m to bus n; and
- $|V_n|$: the voltage magnitude at bus n.

Compliant with the well-known AC power flow model [31, Ch. 4], these measurements are nonlinearly related with the system state, namely the vector $\mathbf{v} := [V_1, \ldots, V_N]^{\mathcal{T}} \in \mathbb{C}^N$. To specify this, collect the injected currents of all buses in $\mathbf{i} := [I_1, \ldots, I_N]^{\mathcal{T}} \in \mathbb{C}^N$, and let $\mathbf{Y} \in \mathbb{C}^{N \times N}$ represent the so-termed bus admittance matrix. Kirchoff's law in vector-matrix form dictates $\mathbf{i} = \mathbf{Y}\mathbf{v}$, see e.g., [3, Sec. 9.1], where the (m, n)-th entry of $\mathbf{Y}$ is

$$Y_{mn} := \begin{cases} -y_{mn}, & \text{if } (m,n) \in \mathcal{E} \\ \bar{y}_{nn} + \sum_{\nu \in \mathcal{N}_n} y_{n\nu}, & \text{if } m = n \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

with y_{mn} denoting the line admittance between buses m and n, $\bar{y}_{nn}$ bus n's shunt admittance to the ground, and $\mathcal{N}_n$ the set of all buses linked to n via transmission lines. Letting $\bar{y}_{mn}$ stand for the shunt admittance at bus m associated with the line (m, n), the current flowing from bus m to n is $I_{mn} = \bar{y}_{mn} V_m + y_{mn}(V_m - V_n)$. The AC power flow model further asserts that the complex power injection into bus n is given by $P_n + jQ_n = V_n I_n^{\mathcal{H}}$, while the complex power flow from bus m to bus n by $P_{mn} + jQ_{mn} = V_m I_{mn}^{\mathcal{H}}$. Finally, expressing the squared bus voltage magnitude as $|V_n|^2 = V_n V_n^{\mathcal{H}}$, it is clear that all measurable quantities listed earlier are nonlinearly (in fact quadratically) related to the state $\mathbf{v}$.

Collect these (possibly noisy) measurements in the $L \times 1$ vector $\mathbf{z} := \left[\{\check{P}_n\}_{n \in \mathcal{N}_P}, \{\check{Q}_n\}_{n \in \mathcal{N}_Q}, \{\check{P}_{mn}\}_{(m,n) \in \mathcal{E}_P}, \{\check{Q}_{mn}\}_{(m,n) \in \mathcal{E}_Q}, \{|\check{V}_n|^2\}_{n \in \mathcal{N}_V}\right]^{\mathcal{T}}$, where the check mark differentiates measured values from the error-free variables[2]. The ℓ-th entry of $\mathbf{z}$ can be written as $z_\ell = h_\ell(\mathbf{v}) + \epsilon_\ell$, where $h_\ell(\cdot)$ denotes the quadratic dependence of z_ℓ on $\mathbf{v}$, and ϵ_ℓ accounts for the additive measurement error assumed independent across meters. Without loss of generality (Wlog), $\mathbf{z}$ is pre-whitened so that all error terms have uniform variance. Hence, the maximum-likelihood (ML) criterion for estimating $\mathbf{v}$ boils down to the nonconvex least-squares (LS) one, as given by

$$\hat{\mathbf{v}} := \arg\min_{\mathbf{v}} \sum_{\ell=1}^{L} [z_\ell - h_\ell(\mathbf{v})]^2. \tag{2}$$

The Gauss-Newton iterative solver for nonlinear LS problems [4, Sec. 1.5] has been widely used for SE; see e.g., [1, Ch. 2] and [31, Ch. 12]. Using Taylor's expansion around a given starting point, the pure form of Gauss-Newton methods approximates the cost in (2) with a linear LS one, and relies on its minimizer to initialize the subsequent iteration. This iterative procedure is closely related to gradient descent algorithms for solving nonconvex problems, see e.g., [4, Ch. 1], which are known to encounter two issues: i) sensitivity to the initial guess; and ii) convergence concerns.

Typical Gauss-Newton iterations for SE start with a flat voltage profile, where all bus voltages are initialized with the same real number. Unfortunately, this fails to guarantee convergence to the global optimum, as pointed out in Section I. Existing variants have asserted improved numerical stability and robustness to outliers [1, Chs. 3–6], but they are all limited to improving the approximate error cost per iteration. Recently,

[1]For distribution system SE, line current magnitude measurements may also be available, as detailed in Remark 2; see also [1, Sec 2.6].

[2]For consistency with measurements in Section III, $|V_n|^2$ is considered henceforth. This is possible using $|\check{V}_n| = |V_n| + \epsilon_V$, where ϵ_V is zero-mean Gaussian with small variance σ_V^2, to obtain the approximate model $|\check{V}_n|^2 \approx |V_n|^2 + \epsilon'_V$, where ϵ'_V has variance $4|\check{V}_n|^2\sigma_V^2$.

with the advent of PMU technology, SE has benefited greatly by including synchrophasor data, which adhere to *linear* measurement models with respect to (wrt) the unknown $\mathbf{v}$. Unfortunately, cost and limited penetration of PMUs require linear measurements to be combined with nonlinear ones to ensure observability.

In a nutshell, the grand challenge so far has been to develop a solver attaining or approximating the *global optimum* at *polynomial-time* complexity. The next section addresses this challenge by appropriately reformulating SE to apply the semidefinite relaxation (SDR) technique.

III. Solving SE Using SDR

Consider first expressing each quadratic measurement in $\mathbf{z}$ linearly in terms of the outer-product matrix $\mathbf{V} := \mathbf{v}\mathbf{v}^{\mathcal{H}}$. To this end, let $\{\mathbf{e}_n\}_{n=1}^{N}$ denote the canonical basis of $\mathbb{R}^N$, and define the following admittance-related matrices

$$\mathbf{Y}_n := \mathbf{e}_n \mathbf{e}_n^{\mathcal{T}} \mathbf{Y} \tag{3a}$$
$$\mathbf{Y}_{mn} := (\bar{y}_{mn} + y_{mn})\mathbf{e}_m \mathbf{e}_m^{\mathcal{T}} - y_{mn}\mathbf{e}_m \mathbf{e}_n^{\mathcal{T}} \tag{3b}$$

along with their related Hermitian counterparts

$$\mathbf{H}_{P,n} := \frac{1}{2}\left(\mathbf{Y}_n + \mathbf{Y}_n^{\mathcal{H}}\right), \quad \mathbf{H}_{Q,n} := \frac{j}{2}\left(\mathbf{Y}_n - \mathbf{Y}_n^{\mathcal{H}}\right) \tag{4a}$$
$$\mathbf{H}_{P,mn} := \frac{1}{2}\left(\mathbf{Y}_{mn} + \mathbf{Y}_{mn}^{\mathcal{H}}\right), \mathbf{H}_{Q,mn} := \frac{j}{2}\left(\mathbf{Y}_{mn} - \mathbf{Y}_{mn}^{\mathcal{H}}\right) \tag{4b}$$
$$\mathbf{H}_{V,n} := \mathbf{e}_n \mathbf{e}_n^{\mathcal{T}}. \tag{4c}$$

Using these definitions, the quadratic measurement model easily gives rise to a linear relation wrt $\mathbf{V}$, as

$$P_n = \mathrm{Tr}(\mathbf{H}_{P,n}\mathbf{V}), \qquad Q_n = \mathrm{Tr}(\mathbf{H}_{Q,n}\mathbf{V}) \tag{5a}$$
$$P_{mn} = \mathrm{Tr}(\mathbf{H}_{P,mn}\mathbf{V}), \quad Q_{mn} = \mathrm{Tr}(\mathbf{H}_{Q,mn}\mathbf{V}) \tag{5b}$$
$$|V_n|^2 = \mathrm{Tr}(\mathbf{H}_{V,n}\mathbf{V}). \tag{5c}$$

Thus, each noisy meter measurement z_ℓ can be written as

$$z_\ell = h_\ell(\mathbf{v}) + \epsilon_\ell = \mathrm{Tr}(\mathbf{H}_\ell \mathbf{V}) + \epsilon_\ell \tag{6}$$

where $\mathbf{H}_\ell$ is the Hermitian matrix specified as per (4a)–(4c). Rewriting (2) with $\mathbf{V}$ as the optimization variable yields

$$\hat{\mathbf{V}}_1 := \arg \min_{\mathbf{V}\in\mathbb{C}^{N\times N}} \sum_{\ell=1}^{L} w_\ell \left[z_\ell - \mathrm{Tr}(\mathbf{H}_\ell \mathbf{V})\right]^2 \tag{7a}$$
$$\text{s.t. } \mathbf{V} \succeq \mathbf{0}, \text{ and } \mathrm{rank}(\mathbf{V}) = 1 \tag{7b}$$

where the positive semi-definite (PSD) and the rank-1 constraints jointly ensure that for any $\mathbf{V}$ admissible to (7b), there always exists a $\mathbf{v} \in \mathbb{C}^N$ such that $\mathbf{V} = \mathbf{v}\mathbf{v}^{\mathcal{H}}$.

Although z_ℓ and $\mathbf{V}$ are linearly related as in (7), nonconvexity is still present in two aspects: i) the cost in (7a) has degree 4 wrt the entries of $\mathbf{V}$; and ii) the rank constraint in (7b) is nonconvex. Aiming for an SDP formulation of (7), Schur's complement lemma, see e.g., [6, Appx. 5.5], can be leveraged to convert the summands in (7a) to a linear cost over an auxiliary vector $\boldsymbol{\chi} \in \mathbb{R}^L$. Specifically, with $\mathbf{w} := [w_1, \ldots, w_L]^{\mathcal{T}}$ and likewise for $\boldsymbol{\chi}$, consider a second SE reformulation as

$$\left\{\hat{\mathbf{V}}_2, \hat{\boldsymbol{\chi}}_2\right\} := \arg \min_{\mathbf{V},\boldsymbol{\chi}} \sum_{\ell=1}^{L} w_\ell \chi_\ell = \arg \min_{\mathbf{V},\boldsymbol{\chi}} \mathbf{w}^{\mathcal{T}} \boldsymbol{\chi} \tag{8a}$$
$$\text{s.t. } \mathbf{V} \succeq \mathbf{0}, \text{and } \mathrm{rank}(\mathbf{V}) = 1 \tag{8b}$$
$$\begin{bmatrix} -\chi_\ell & z_\ell - \mathrm{Tr}(\mathbf{H}_\ell \mathbf{V}) \\ z_\ell - \mathrm{Tr}(\mathbf{H}_\ell \mathbf{V}) & -1 \end{bmatrix} \preceq \mathbf{0} \;\; \forall \ell. \tag{8c}$$

The equivalence among all three SE problems (2), (7), and (8) has been shown in [35], where their optimal solutions satisfy

$$\hat{\mathbf{V}}_1 = \hat{\mathbf{V}}_2 = \hat{\mathbf{v}}\hat{\mathbf{v}}^{\mathcal{H}}$$
$$\text{and } \hat{\chi}_{2,\ell} = \left[\hat{z}_\ell - \mathrm{Tr}(\mathbf{H}_\ell \hat{\mathbf{V}}_2)\right]^2 \;\; \forall \ell. \tag{9}$$

The only nonconvexity in the equivalent SE problem (8) lies in the rank-1 constraint. Fortunately, (8) is amenable to the SDR technique, which amounts to dropping the rank constraint and has well-appreciated merits as an optimization tool to handel nonconvex problems; see e.g., the seminal work of [11]. Thanks to its performance guarantees and implementation advantages SDR has recently provided new perspectives for a number of nonconvex problems in various applications, including signal processing and communications [24]. For our SE problem, it can be shown later on that the relaxed problem is able to achieve the global optimum under some simplified conditions. The contribution here consists in permeating the benefits of this powerful optimization tool to estimate the state of AC power systems. In the spirit of SDR, relaxing the rank constraint in (8b) leads to the following SDP

$$\left\{\hat{\mathbf{V}}, \hat{\boldsymbol{\chi}}\right\} := \arg \min_{\mathbf{V},\boldsymbol{\chi}} \mathbf{w}^{\mathcal{T}} \boldsymbol{\chi} \tag{10a}$$
$$\text{s.t. } \mathbf{V} \succeq \mathbf{0}, \tag{10b}$$
$$\begin{bmatrix} -\chi_\ell & z_\ell - \mathrm{Tr}(\mathbf{H}_\ell \mathbf{V}) \\ z_\ell - \mathrm{Tr}(\mathbf{H}_\ell \mathbf{V}) & -1 \end{bmatrix} \preceq \mathbf{0} \; \forall \ell. \tag{10c}$$

To support the near-optimality of the SDR approach, a few assumptions are needed to show its global optimality under a special scenario:

(as1) The graph $(\mathcal{N}, \mathcal{E})$ has a tree topology.
(as2) Every bus is equipped with a voltage magnitude meter.
(as3) All the measurements in $\mathbf{z}$ are noise-free; i.e., $\epsilon_\ell = 0$, $\forall \ell$;

Proposition 1: Under (as1)-(as3), the relaxed problem (10) can attain the global optimum of the SE problem (8); i.e., $rank(\hat{\mathbf{V}}) = 1$.

Proof: The proof sketched here leverages the results on characterizing the geometry of the power flow regions in [22]. The noise-free model in (as3) yields the minimum estimation error cost of (10) to be 0, and thus the inequality constraints in (10c) can be replaced by the strict equality ones

$$\mathrm{Tr}(\mathbf{H}_\ell \mathbf{V}) = z_\ell, \forall \ell. \tag{11}$$

Together with (as2), (11) leads to a fixed voltage magnitude at every bus. As shown in [22], the SDR approach equivalently relaxes the power flow region to its convex hull, where for tree networks with fixed bus voltage magnitude the power flow region in fact consists of the boundary points of its convex hull. The full observability of the measurement model ensures that the solution to (11) falls uniquely on the boundary point of the feasible set. Thus, the optimum of (10) yields a valid power flow solution $\hat{\mathbf{v}}$ such that $\hat{\mathbf{V}} = \hat{\mathbf{v}}\hat{\mathbf{v}}^{\mathcal{H}}$. ■

Although (as1)-(as3) do not hold for most power networks, they may offer a 'close' approximation of the realistic SE scenario. First, it is well known that all power networks are extremely sparse with average node degree around 2 [28], while almost all distribution networks today have tree topology. Therefore, (as1) could closely capture the sparse connectivity of most power networks. Second, bus voltage magnitude meters are deployed in most substations to monitor the voltage rating for ensuing dynamic stability. Otherwise, since the transmission system has to operate at the voltage level within a fixed range, pseudo-measurements are often used to complement metering if no direct measurements are available for some bus voltage magnitudes, and thus fulfill (as2). The ideal metering condition (as3) can be reasonable since most deployed meters are of high accuracy. Even though bad data or outliers could be present in some meters as pointed out later in Remark 3, high measurement redundancy ensured by control centers makes it possible to rely on the highly accurate meters only. Admittedly, (as1)-(as3) are unlikely to be satisfied in realistic SE settings. Hence, Proposition 1 has limitations in assessing the SDR-based SE performance in practice. Nonetheless, it offers insights on the near-optimal performance of the SDR-based SE approach, since a promising SE scheme should at least ensure identifiability in the noiseless or asymptotically high-SNR cases. It further supports analytically the simulation-based evidence that the proposed SDR-based SE attains a solution very close to the global optimum of (8). Near-optimality of the relaxed problem (10) will be more convincingly supported by extensive numerical tests in Section VI.

The solution to the relaxed (10), $\hat{\mathbf{V}}$ is very likely to have rank greater than 1. Hence, it is necessary to recover a feasible estimate $\hat{\mathbf{v}}$ from $\hat{\mathbf{V}}$. This is possible by eigen-decomposing $\hat{\mathbf{V}} = \sum_{i=1}^{r} \lambda_i \mathbf{u}_i \mathbf{u}_i^{\mathcal{H}}$, where $r := \mathrm{rank}(\hat{\mathbf{V}})$, $\lambda_1 \geq \cdots \geq \lambda_r > 0$ denote the positive ordered eigenvalues, and $\{\mathbf{u}_i \in \mathbb{C}^N\}_{i=1}^{r}$ are the corresponding eigenvectors. Since the best (in the minimum-norm sense) rank-one approximation of $\hat{\mathbf{V}}$ is $\lambda_1 \mathbf{u}_1 \mathbf{u}_1^{\mathcal{H}}$, the state estimate can be chosen equal to $\hat{\mathbf{v}}(\mathbf{u}_1) := \sqrt{\lambda_1}\mathbf{u}_1$. Besides this eigenvector approach, randomization offers another way to extract $\hat{\mathbf{v}}$ from $\hat{\mathbf{V}}$, with quantifiable approximation accuracy; see e.g., [24]. The basic idea is to generate multiple Gaussian distributed random vectors $\boldsymbol{\nu} \sim \mathcal{CN}(\mathbf{0}, \hat{\mathbf{V}})$, and pick the one with the minimum WLS cost. Although any vector $\boldsymbol{\nu}$ is feasible for (2), it is still possible to decrease the minimum achievable cost by rescaling to obtain $\hat{\mathbf{v}}(\boldsymbol{\nu}) = \hat{c}\boldsymbol{\nu}$, where the optimal weight can be chosen as the solution to the following convex problem as

$$\hat{c} = \arg\min_{c>0} \sum_{\ell=1}^{L} w_\ell [z_\ell - c^2 \boldsymbol{\nu}^{\mathcal{H}} \mathbf{H}_\ell \boldsymbol{\nu}]^2 = \sqrt{\frac{\sum_{\ell=1}^{L} w_\ell z_\ell \boldsymbol{\nu}^{\mathcal{H}} \mathbf{H}_\ell \boldsymbol{\nu}}{\sum_{\ell=1}^{L} w_\ell (\boldsymbol{\nu}^{\mathcal{H}} \mathbf{H}_\ell \boldsymbol{\nu})^2}}. \quad (12)$$

SDR endows SE with a convex SDP formulation for which efficient schemes are available to obtain the global optimum using, e.g., the interior-point solver SeDuMi [27]. The computational complexity for eigen-decomposition is in the order of matrix multiplication, and thus negligible compared to solving the SDP; see e.g., [24] and references therein. However, the polynomial complexity order of solving the SDP could be a burden for real-time power system monitoring, which motivates us to consider a distributed implementation in Section V.

Remark 1: (Reference bus). The reference bus convention adopted in power systems sets the corresponding bus voltage angle set to 0; see e.g., [31, pg. 76]. Due to the quadratic measurement model in (6), the outer-product $(e^{j\theta}\mathbf{v})(e^{j\theta}\mathbf{v})^{\mathcal{H}} = \mathbf{v}\mathbf{v}^{\mathcal{H}}$ remains invariant to any phase rotation $\theta \in [-\pi, \pi]$. To account for this, once an estimate $\hat{\mathbf{v}}$ is recovered, it can be rotated by multiplying with $\hat{V}_{\mathrm{ref}}^{\mathcal{H}}/|\hat{V}_{\mathrm{ref}}|$, where $\hat{V}_{\mathrm{ref}}$ denotes the estimated reference-bus voltage.

Remark 2: (Additional measurements). For certain distribution-level settings, line current magnitude data $|\check{I}_{mn}|$ are also available; see e.g., [1, Sec. 2.6]. Furthermore, for buses with no generations or loads, it is possible to include "pseudo-measurements," as equality constraints of zero power injections; see e.g., [25]. It is worth stressing that all SDP-based SE reformulations of the present paper can readily handle these types of measurements. Since I_{mn} is linear in $\mathbf{v}$, it follows that $|I_{mn}|^2$ is quadratic in $\mathbf{v}$, and thus linear in $\mathbf{V}$ as in (6). To include pseudo-measurements $[P_n(Q_n) = 0$ or $|V_n| = 1]$, it suffices to add extra equality constraints $\mathrm{Tr}[\mathbf{H}_{P(Q),n}\mathbf{V}] = 0$ or $\mathrm{Tr}(\mathbf{H}_{V,n}\mathbf{V}) = 1$ to (10). Hence, the reformulation equivalence and the ensuing analysis apply even to these additional measurements. As detailed soon in Section IV, it is also possible to include linear synchrophasor measurements offered by the latest PMU technology.

Remark 3: (Robust SE). In addition to being optimal under nominal operation, SE approaches must be resilient to outliers emerging not only due to data contamination, telemetry errors, and asynchronous meter measurements [25] and [1, Chs. 5–6], but also due to those coming from cyber attacks, where attackers maybe able to manipulate remote meters without being noticed [23]. Although robustness issues go beyond the scope of the present work, it is worth mentioning that the proposed SDR-based SE approach can be robustified by introducing an auxiliary outlier (or attack) variable per meter, and regularizing the WLS cost with the number of those variables that are nonzero [36]. With or without regularization, resilience to "bad data" can be also effected by replacing the WLS cost with robust alternatives including the least-absolute error and Huber's costs [25], which are convex and can be accommodated too by the SDP reformulations here. Lastly, prior information such as the probability distribution of state variables has also been used to develop a Bayesian SE framework for more resilience to cyber-attacks [19]. Such extension can be easily incorporated as well by slightly modifying the optimization objective with additional covariance-related terms.

IV. PMU-AIDED SDR-BASED SE

Recent deployment of PMUs suggests complementing with PMU data, the measurements collected by legacy meters to perform SE. This motivates us to expand the SDR-based SE paradigm to include synchrophasor measurements. Compared to legacy measurements, PMUs provide synchronous data that are *linear* functions of the state $\mathbf{v}$. If bus m is equipped with a PMU, then its voltage phasor V_m and related current phasors $\{I_{mn}\}_{n\in\mathcal{N}_m}$ are available to the control center with high accuracy. Hence, with adequate number of PMUs and wisely chosen placement buses, SE using only PMU data boils down to estimating a linear regression coefficient vector for which a batch WLS solution is available in closed form. However, installation and networking costs involved allow only for limited penetration of PMUs in the near future. This means that SE must be performed using *jointly* legacy meters and PMU measurements.

To this end, let $\boldsymbol{\zeta}_m = \boldsymbol{\Phi}_m\mathbf{v} + \boldsymbol{\varepsilon}_m$ collect the noisy PMU data at bus m, where $\boldsymbol{\Phi}_m$ is the measurement matrix, while $\boldsymbol{\varepsilon}_m$ denotes measurement noise, assumed to be complex zero-mean Gaussian with covariance $2\check{\sigma}_m^2\mathbf{I}$, independent across buses and from the legacy meter noise terms $\{\epsilon_\ell\}$. Matrix $\boldsymbol{\Phi}_m$ is constructed in accordance to the bus index m and the line admittances. The voltage phasor measurement $\check{V}_m$ corresponds to an all-zero row except for the m-th entry which is unity, while the row for the current phasor measurement $\check{I}_{mn}$ has line admittance value at the m-th and n-th entries; see e.g., [13], [18].

The SE task now amounts to estimating $\mathbf{v}$ given both $\mathbf{z}$ and $\{\boldsymbol{\zeta}_m\}_{m\in\mathcal{P}}$, where $\mathcal{P}\subseteq\mathcal{N}$ denotes the PMU-instrumented set of buses. Hence, the ML-optimal WLS cost in (2) must be augmented with the log-likelihood induced by PMU data, as

$$\hat{\mathbf{v}} := \arg\min_{\mathbf{v}} \sum_{\ell=1}^{L} w_\ell[z_\ell - h_\ell(\mathbf{v})]^2 + \sum_{m\in\mathcal{P}} \omega_m\|\boldsymbol{\zeta}_m - \boldsymbol{\Phi}_m\mathbf{v}\|_2^2 \tag{13}$$

where $\omega_m := 1/\check{\sigma}_m^2\ \forall m\in\mathcal{P}$. The augmented SE problem (13) is still nonconvex due to the quadratic dependence of legacy measurements in the wanted state $\mathbf{v}$. Existing SE methods that account for PMU measurements can be categorized in two groups. The first one includes the so-termed *hybrid* SE approaches which utilize both PMU and legacy measurements in a WLS solver via iterative linearization; see e.g., [13]. Depending on the number of PMUs, the state can be either expressed using polar coordinates (similar to traditional WLS-based SE), or by rectangular coordinates (as $\mathbf{v}$ is expressed here). The polar representation is preferred when legacy measurements are abundant, because it requires minor adaptations of the existing WLS-based SE. On the other hand, the polar representation is less powerful when it comes to exploiting the linearity of PMU measurements. With increasing penetration of PMUs, the rectangular representation will grow in popularity, especially if full observability can be ensured solely based on PMU data.

An alternative approach to including PMU data is through sequential SE [34], which entails two steps. The WLS-based SE is performed first based only on legacy measurements. Together with PMU data, these estimates serve as linear "pseudo-measurements" for the subsequent step. The post-processing involves linear models only, and is efficiently computable. Clearly, this two-step scheme requires no modifications of existing SE modules, but loses the optimality offered by joint estimation. Even worse, as traditional SE based only on legacy measurements cannot ensure convergence to a global optimum, the post-processing including PMU data is unlikely to improve estimation accuracy.

Since both means of including PMU data suffer from the nonconvexity inherent with legacy measurements, SDR is again well motivated to convexify the augmented SE to

$$\left\{\hat{\mathbf{X}},\hat{\boldsymbol{\chi}}\right\} := \arg\min_{\mathbf{X},\mathbf{V},\mathbf{v},\boldsymbol{\chi}} \mathbf{w}^T\boldsymbol{\chi} + \sum_{m\in\mathcal{P}}\omega_m\Big[\mathrm{Tr}(\boldsymbol{\Phi}_m^{\mathcal{H}}\boldsymbol{\Phi}_m\mathbf{V}) - 2\mathcal{R}(\boldsymbol{\zeta}_m^{\mathcal{H}}\boldsymbol{\Phi}_m\mathbf{v})\Big] \tag{14a}$$

$$\text{s.t. } \mathbf{X} = \begin{bmatrix}\mathbf{V} & \mathbf{v}\\ \mathbf{v}^{\mathcal{H}} & 1\end{bmatrix} \succeq \mathbf{0}, \tag{14b}$$

$$\begin{bmatrix}-\chi_\ell & z_\ell - \mathrm{Tr}(\mathbf{H}_\ell\mathbf{V})\\ z_\ell - \mathrm{Tr}(\mathbf{H}_\ell\mathbf{V}) & -1\end{bmatrix} \preceq \mathbf{0}\forall\ell. \tag{14c}$$

Similar to (10), with an additional constraint $\mathrm{rank}(\mathbf{X}) = 1$ the PSD of $\mathbf{X}$ can ensure $\mathbf{V} = \mathbf{v}\mathbf{v}^{\mathcal{H}}$. Substituting this into (14) leads to the equivalence of rank-constrained (14) with the augmented WLS in (13). The SDP here also offers the advantages of (10), in terms of the near-optimality and distributed implementation as discussed later on. From the solution $\hat{\mathbf{X}}$, either eigenvector approximation or randomization can be employed to generate vectors of length $N+1$. Using the first N entries of any such vector, a feasible $\hat{\boldsymbol{\chi}}$ can be formed by proper rescaling. With linear PMU measurements, the voltage angle ambiguity is no longer present, and the rescaling factor $\hat{c}^\star$ can be found by solving

$$\hat{c}^\star = \arg\min_{c\in\mathbb{C}} \sum_{\ell=1}^{L} w_\ell\left[z_\ell - |c|^2\boldsymbol{\nu}^{\mathcal{H}}\mathbf{H}_\ell\boldsymbol{\nu}\right]^2 + \sum_{m\in\mathcal{P}}\omega_m\|\boldsymbol{\zeta}_m - c\boldsymbol{\Phi}_m\boldsymbol{\nu}\|_2^2. \tag{15}$$

This fourth-order polynomial minimization can be solved numerically. As PMU data are typically more accurate than legacy ones, the weights in (15) must satisfy $\omega_\ell \gg w_\ell$. Hence, it suffices to minimize only the dominant second sum in (15), and efficiently approximate the solution $\hat{c}^\star \approx (\sum_{m\in\mathcal{P}}\omega_m\boldsymbol{\zeta}_m^{\mathcal{H}}\boldsymbol{\Phi}_m\boldsymbol{\nu})/(\sum_{m\in\mathcal{P}}\omega_m\|\boldsymbol{\Phi}_m\boldsymbol{\nu}\|_2^2)$. This approximation will be used next in the numerical tests. As opposed to existing methods, the SDR-SE approach constitutes as a convenient way to include additional synchrophasor measurements.

V. Distributed SDR-Based SE

Although the SDR-SE approach enjoys the polynomial complexity from the convex SDP formulation, its worst-case complexity is still $\mathcal{O}(L^4\sqrt{N}\log(1/\epsilon))$ for a given solution accuracy $\epsilon > 0$ [24]. For typical power networks, L is in the order of N, and thus the worst-case complexity becomes $\mathcal{O}(N^{4.5}\log(1/\epsilon))$. This complexity order could be a computational burden for large-scale power networks with a growing size, which motivates us to consider accelerating the SDR-SE method using distributed parallel implementations. A distributed SE solver is also advocated by the increased interaction among different regional control centers for improving reliability through system interconnection. Due to data confidentiality concerns, it is necessary to develop a distributed

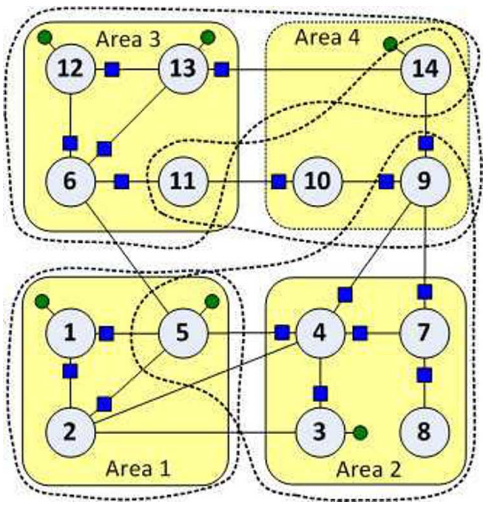


Fig. 1. The IEEE 14-bus system partitioned into four areas [17], [26]. Dotted lassos show the buses in $\mathcal{N}_{(k)}$ that are related to all the measurements in $\mathbf{z}_k$. Squares on the lines mark the power flow meter locations.

SE solver with minimal data exchange among the regional operators while improving the estimation performance.

Suppose there exist K interconnected control areas partitioning the set of all buses as $\mathcal{N} = \bigcup_{k=1}^{K} \mathcal{N}_k$, where $\mathcal{N}_k$ contains the subset of buses supervised by the k-th control center; see e.g., the four-area partitioning in Fig. 1. The same partition also holds for the voltage vector $\mathbf{v}_k$ and the measurement vector $\mathbf{z}_k$ of length M_k. All the measurements in $\mathbf{z}_k$ not necessarily relate to bus voltage phasors only in $\mathbf{v}_k$. Taking for example Area 2 in Fig. 1, the line flow meter placed on the tie-line (4,5) depends on V_5 in Area 1; similarly for the flow meter on the tie-line (7,9). Hence, augment $\mathcal{N}_k$ to $\mathcal{N}_{(k)} \subseteq \mathcal{N}$ to include all the buses that would affect the measurements contained in $\mathbf{z}_k$. The augmented set $\mathcal{N}_{(k)}$ may include buses from neighboring areas that are interconnected through tie lines; e.g., $\mathcal{N}_{(2)} = \mathcal{N}_2 \cup \{5, 9\}$ for the example in Fig. 1. (All the sets $\mathcal{N}_{(k)}$ are indicated by the dotted lassos in Fig. 1.) Based on the overlaps among $\{\mathcal{N}_{(k)}\}$, define the set of neighboring areas for the k-th one as $\mathcal{A}_k := \{j | 1 \leq j \leq K, \mathcal{N}_{(j)} \bigcap \mathcal{N}_{(k)} \neq 0\}$. Area 2 in Fig 1 has neighbors in $\mathcal{A}_2 = \{1, 3\}$. Let also the vector $\mathbf{v}_{(k)}$ collect the state variables in $\mathcal{N}_{(k)}$. Hence, it is possible to re-write the measurement model (6) as

$$z_k^\ell = h_{(k)}^\ell(\mathbf{v}_{(k)}) = \mathrm{Tr}(\mathbf{H}_{(k)}^\ell \mathbf{V}_{(k)}) + \epsilon_k^\ell, \forall k, \ell \tag{16}$$

where $\mathbf{V}_{(k)} := \mathbf{v}_{(k)} \mathbf{v}_{(k)}^{\mathcal{H}}$ denotes a submatrix of $\mathbf{V}$ formed by extracting rows and columns corresponding to buses in $\mathcal{N}_{(k)}$; and similarly for $\mathbf{H}_{(k)}^\ell$. Due to the overlaps among the subsets $\{\mathcal{N}_{(k)}\}$, the outer-product $\mathbf{V}_{(k)}$ of the k-th area overlaps also with $\mathbf{V}_{(j)}$, for each of its neighboring areas $j \in \mathcal{A}_k$.

By reducing the measurement functions at the k-th area to the submatrix $\mathbf{V}_{(k)}$, it is further possible to define the LS error cost per area k as

$$f_k(\mathbf{V}_{(k)}) := \sum_{\ell=1}^{M_k} w_\ell \left[z_k^\ell - \mathrm{Tr}(\mathbf{H}_{(k)}^\ell \mathbf{V}_{(k)}) \right]^2 \tag{17}$$

which only involves the local matrix $\mathbf{V}_{(k)}$. Hence, the centralized SE problem in (10) becomes equivalent to

$$\begin{aligned} \hat{\mathbf{V}} = \arg\min_{\mathbf{V}} & \sum_k f_k(\mathbf{V}_{(k)}) \\ \text{s.t.} & \mathbf{V} \succeq \mathbf{0}. \end{aligned} \tag{18}$$

This equivalent formulation effectively expresses the overall LS cost as the superposition of local costs in (17). Nonetheless, even with such a decomposition of the cost, the main challenge to implement (18) in a distributed manner actually lies in the PSD constraint that couples local matrices $\{\mathbf{V}_{(k)}\}$ which overlap partially. If all submatrices $\{\mathbf{V}_{(k)}\}$ were non-overlapping, the cost would be decomposable as in (18), and the PSD of $\mathbf{V}$ would boil down to a PSD constraint per area k, as given by

$$\begin{aligned} \hat{\mathbf{V}} = \arg\min_{\mathbf{V}} & \sum_k f_k(\mathbf{V}_{(k)}) \\ \text{s.t.} & \mathbf{V}_{(k)} \succeq \mathbf{0}, \forall k. \end{aligned} \tag{19}$$

As detailed soon, the formulation in (19) is amendable to being decomposed into sub-problems, thanks to the separable PSD constraints. It is not always equivalent to the centralized (18) though, because PSD property of all submatrices not necessarily lead to a PSD overall matrix. Nonetheless, the decomposable problem (19) is still a valid SDR-SE reformulation, since with additional constraints $\mathrm{rank}(\mathbf{V}_{(k)}) = 1$ per area k it is actually equivalent to (8). While it is totally legitimate to use (19) as the relaxed SDP formulation for (8), the two relaxed problems are actually equivalent under very mild conditions.

The idea here is to explore valid network topologies to facilitate such PSD constraint decomposition. To this end, it will be instrumental to leverage results on completing partial Hermitian matrices to obtain PSD ones [16]. Upon obtaining the underlying graph formed by the specified entries in the partial Hermitian matrices, these results rely on the so-termed graph "chordal" property to establish the equivalence between the positive semidefiniteness of the overall matrix and that of all submatrices corresponding to the graph's maximal cliques. Interestingly, this technique has been used recently for developing distributed SDP-based optimal power flow (OPF) solvers in [8], [15], [20]. To leverage this, construct first a graph $\mathcal{G}$ over $\mathcal{N}$, with all its edges corresponding to the entries in $\{\mathbf{V}_{(k)}\}$. The graph $\mathcal{G}$ amounts to having all buses within each subset $\mathcal{N}_{(k)}$ to form a clique. Furthermore, the following is assumed:

(as4) The graph with all the control areas as nodes and their edges defined by the neighborhood subset $\{\mathcal{A}_k\}_{k=1}^K$ forms a tree.

(as5) Each control area has at least one bus that does not overlap with any neighboring area.

Condition (as4) is quite reasonable for the control areas in most transmission networks, which in general are loosely connected over large geographical areas by a small number of tie lines. In addition, under the current meter deployment, the tie lines are not monitored everywhere, and thus it is more likely to have tree-connected control areas when requiring neighboring areas to share a tie-line with meter measurements. In Fig. 1 for instance, Areas 1 and 4 are physically interconnected by the tie line (5,6). However, there is no measurement on that line, hence the two areas are not neighbors and the total four areas eventually form a line array. Moreover, condition (as5) easily holds since in practice most of the buses are not connected to any tie line.

Proposition 2: Under (as4) and (as5), the two relaxed problems (18) and (19) are equivalent.

Proposition 2 can be proved by following the arguments in [37], which show that the full PSD matrix $\mathbf{V}$ can be "completable" from all PSD submatrices $\mathbf{V}_{(k)}$. The premise here is that in most practical systems, the relaxed problem (19) has the promise to achieve the same accuracy as the centralized one, where the PSD constraint decomposition of the former is key for developing distributed solvers. To this end, each control area k solves for its own local $\mathbf{V}_{(k)}$, denoted by the complex matrix $\mathbf{W}_k$ of size $|\mathcal{N}_{(k)}| \times |\mathcal{N}_{(k)}|$. For every pair of neighboring areas, say k and j, identify the intersection of their buses as $\mathcal{S}_{kj}$. Let also $\mathbf{W}_k^j$ denote the submatrix extracted from $\mathbf{W}_k$ with both rows and columns corresponding to $\mathcal{S}_{kj}$; and likewise for the submatrix $\mathbf{W}_j^k$ from $\mathbf{W}_j$. To formulate (19) as one involving all local matrices $\{\mathbf{W}_k\}$, it suffices to have additional equality constraints on the overlapping entries, namely

$$\begin{aligned}\{\hat{\mathbf{W}}_k\} = \arg \min_{\{\mathbf{W}_k\}} & \sum_{\kappa} f_\kappa(\mathbf{W}_\kappa) \\ \text{s.t. } & \mathbf{W}_k \succeq \mathbf{0}, \forall k, \\ & \mathbf{W}_k^j = \mathbf{W}_j^k, \ \forall k, \ \forall j \in \mathcal{A}_k. \end{aligned} \tag{20}$$

The equality constraints in (20) enforce neighboring areas to consent on their shared entries, rendering the equivalence between (20) and (19) established as $\hat{\mathbf{W}}_k = \hat{\mathbf{V}}_{(k)}, \forall k$. Interestingly, this allows for powerful distributed implementation modules to realize multi-area SDR-based SE.

A. Alternating Direction Method-Of-Multipliers

As mentioned earlier, the PSD constraint decomposition recently has been used for also developing distributed SDP-based optimal power flow (OPF) solvers, where [20] utilizes either primal or dual iterations to minimize a linear cost function. To extend this to a quadratic cost, the distributed OPF method in [20] has to further require a more complicated outer iteration. As opposed to this approach, it is possible to leverage on the alternating direction method-of-multipliers (ADMM) [4, Sec. 3.4.4], to handle any general error cost f_k, so long as it is convex. The ADMM approach has shown successful in a variety of distributed computation tasks; see e.g., the review in [5] and its application for SDP-based OPF in [8].

To this end, two auxiliary matrices denoted by $\mathbf{R}_{kj}$ and $\mathbf{I}_{kj}$ are introduced per pair of neighboring areas (k,j), to handle the coupling equality constraints in (20). For notational brevity, the symbols $\mathbf{R}_{kj}$ and $\mathbf{R}_{jk}$ are used interchangeably to denote a same matrix; and similarly for $\mathbf{I}_{kj}$ and $\mathbf{I}_{jk}$. With these, (20) can be alternatively expressed as

$$\begin{aligned}\{\hat{\mathbf{W}}_k\} = \arg \min_{\{\mathbf{W}_k \succeq \mathbf{0}\}} & \sum_{\kappa} f_\kappa(\mathbf{W}_\kappa) \\ \text{s.t. } & \mathcal{R}(\mathbf{W}_k^j) = \mathbf{R}_{kj}, \ \forall j \in \mathcal{A}_k, \ \forall k, \\ & \mathcal{I}(\mathbf{W}_k^j) = \mathbf{I}_{kj}, \ \forall j \in \mathcal{A}_k, \ \forall k. \end{aligned} \tag{21}$$

The goal is to solve the penalized dual problem of (21) in a distributed fashion the ADMM solver. Let $\mathbf{\Gamma}_{kj}$ and $\mathbf{\Lambda}_{kj}$ denote the Lagrange multipliers associated with the pair of constraints in (21). With $\mu > 0$ denoting a penalty coefficient, consider the augmented Lagrangian function of (21) as

$$\begin{aligned}\mathcal{L}(\{\mathbf{W}_k\}, &\{\mathbf{R}_{kj}\}, \{\mathbf{I}_{kj}\}, \{\mathbf{\Gamma}_{kj}\}, \{\mathbf{\Lambda}_{kj}\}) \\ := \sum_k \Bigg\{ & f_k(\mathbf{W}_k) + \sum_{j \in \mathcal{A}_k} \mathrm{Tr}[\mathbf{\Gamma}_{kj}(\mathcal{R}(\mathbf{W}_k^j) - \mathbf{R}_{kj})] \\ & + \sum_{j \in \mathcal{A}_k} \frac{\mu}{2} \|\mathcal{R}(\mathbf{W}_k^j) - \mathbf{R}_{kj}\|_F^2 \\ & + \sum_{j \in \mathcal{A}_k} \mathrm{Tr}[\mathbf{\Lambda}_{kj}(\mathcal{I}(\mathbf{W}_k^j) - \mathbf{I}_{kj})] \\ & + \sum_{j \in \mathcal{A}_k} \frac{\mu}{2} \|\mathcal{I}(\mathbf{W}_k^j) - \mathbf{I}_{kj}\|_F^2 \Bigg\}. \end{aligned} \tag{22}$$

The positive coefficient μ is introduced to penalize the mismatch associated with the equality constraints in (21). Letting i denote the iteration index in the superscript, the ADMM operates by cyclically minimizing the augmented Lagrangian $\mathcal{L}$ in (22) wrt one set of variables while fixing the rest. Given all the iterates at the i-th iteration, the ADMM steps proceed to the ensuing iteration as follows.

[S1] Update the primal variables:

$$\{\mathbf{W}_k^{i+1}\} := \arg \min_{\mathbf{W}_k \succeq \mathbf{0}} \mathcal{L}(\{\mathbf{W}_k\}, \{\mathbf{R}_{kj}^i\}, \{\mathbf{I}_{kj}^i\}, \{\mathbf{\Gamma}_{kj}^i\}, \{\mathbf{\Lambda}_{kj}^i\}).$$

[S2] Update the auxiliary variables:

$$\{\mathbf{R}_{kj}^{i+1}, \mathbf{I}_{kj}^{i+1}\} := \arg \min_{\mathbf{R}_{kj}, \mathbf{I}_{kj}} \mathcal{L}(\{\mathbf{W}_k^{i+1}\}, \{\mathbf{R}_{kj}\}, \{\mathbf{I}_{kj}\}, \{\mathbf{\Gamma}_{kj}^i\}, \{\mathbf{\Lambda}_{kj}^i\}).$$

[S3] Update the multipliers:

$$\begin{aligned}\mathbf{\Gamma}_{kj}^{i+1} &:= \mathbf{\Gamma}_{kj}^i + \mu[\mathcal{R}(\mathbf{W}_k^j)^{i+1} - \mathbf{R}_{kj}^{i+1}] \\ \mathbf{\Lambda}_{kj}^{i+1} &= \mathbf{\Lambda}_{kj}^i + \mu[\mathcal{I}(\mathbf{W}_k^j)^{i+1} - \mathbf{I}_{kj}^{i+1}].\end{aligned}$$

All the variables can be easily initialized to $\mathbf{0}$. Clearly, the optimization problem in [S1] is decomposable over all K control areas. Moreover, exploiting the specific problem structure can simplify the three steps as follows.

[S1] Update $\mathbf{W}_k^{i+1}$ per area k:

$$\begin{aligned}\mathbf{W}_k^{i+1} := \arg\min_{\mathbf{W}_k \succeq \mathbf{0}} & f_k(\mathbf{W}_k) + \sum_{j\in\mathcal{A}_k} \mathrm{Tr}[\mathbf{\Gamma}_{kj}^i \mathcal{R}(\mathbf{W}_k^j)] \\ &+ \sum_{j\in\mathcal{A}_k} \mathrm{Tr}[\mathbf{\Lambda}_{kj}^i \mathcal{I}(\mathbf{W}_k^j)] \\ &+ \sum_{j\in\mathcal{A}_k} \frac{\mu}{2}\|\mathcal{R}(\mathbf{W}_k^j) - \mathbf{R}_{kj}^i\|_F^2 \\ &+ \sum_{j\in\mathcal{A}_k} \frac{\mu}{2}\|\mathcal{I}(\mathbf{W}_k^j) - \mathbf{I}_{kj}^i\|_F^2. \end{aligned} \tag{23}$$

[S2] Update the pair of auxiliary variables per area k:

$$\mathbf{R}_{kj}^{i+1} = \frac{1}{2}[\mathcal{R}(\mathbf{W}_k^j)^{i+1} + \mathcal{R}(\mathbf{W}_j^k)^{i+1}] \tag{24}$$

$$\mathbf{I}_{kj}^{i+1} = \frac{1}{2}[\mathcal{I}(\mathbf{W}_k^j)^{i+1} + \mathcal{I}(\mathbf{W}_j^k)^{i+1}]. \tag{25}$$

[S3] Update the pair of multipliers per area k:

$$\mathbf{\Gamma}_{kj}^{i+1} := \mathbf{\Gamma}_{kj}^i + \frac{\mu}{2}[\mathcal{R}(\mathbf{W}_k^j)^{i+1} - \mathcal{R}(\mathbf{W}_j^k)^{i+1}] \tag{26}$$

$$\mathbf{\Lambda}_{kj}^{i+1} = \mathbf{\Lambda}_{kj}^i + \frac{\mu}{2}[\mathcal{I}(\mathbf{W}_k^j)^{i+1} - \mathcal{I}(\mathbf{W}_j^k)^{i+1}]. \tag{27}$$

Since the LS error f_k is quadratic, the cost in (23) is convex in $\mathbf{W}_k$, and can be formulated as an SDP using Schur's complement lemma as in (10). Hence, [S1] amounts to solving local problems that scale with the number of buses controlled by each regional center, greatly reducing the computational burden as compared to the global SDP problem. In addition, both [S2] and [S3] are simplify very efficient linear iterations. This completes the iterative procedure for updating the SE submatrices in a distributed fashion among multiple areas. Upon convergence, each control center obtains the iterate $\mathbf{W}_k^i$ as the estimate of its local matrix $\mathbf{V}_{(k)}$ of the centralized SDR-based SE problem (19). Eigen-decomposition or randomization method can be applied to recover a rank-1 solution from $\mathbf{W}_k^i$ similarly.

Remark 4: (Generalized estimation error costs) Compared to the distributed OPF methods with linear or at most quadratic costs, the proposed distributed SE framework can accommodate more general error cost functions f_k for various estimation purposes pointed out by Remark 3. This includes the aforementioned ℓ_1-norm of estimation error vector, the combination of the $\ell_1 - \ell_2$ norm, as well as log-prior terms. Furthermore, it is possible to develop distributed *hybrid-SE* as well by incorporating linear PMU data. Thanks to the convexity of the error criterion for all cases, it only requires to slightly modify (23) in [S1] to a different local convex SDP problem.

Remark 5: (Data exchange overhead and privacy) At first glance, one may think that in steps [S1]-[S3] neighboring control centers need to exchange the submatrices $\{(\mathbf{W}_k^j)^i\}$ related to the common buses per iteration i, as well as the associated multipliers $\{\mathbf{\Gamma}_{kj}^i\}$ and $\{\mathbf{\Lambda}_{kj}^i\}$. A closer look however, reveals that exchanging common submatrices suffices as the multipliers can be readily updated locally with coordination among neighboring areas, e.g., by initializing all to zero. This suggests considerable reduction in the communication overhead for the ADMM iterations. Furthermore, the proposed scheme neither requires exchanging local measurements nor local network topology. It suffices to only share a small portion of local state matrices. From the data privacy perspective, individual operators enjoy this benefit as a natural bonus of this multi-area SE method.

VI. Numerical Tests

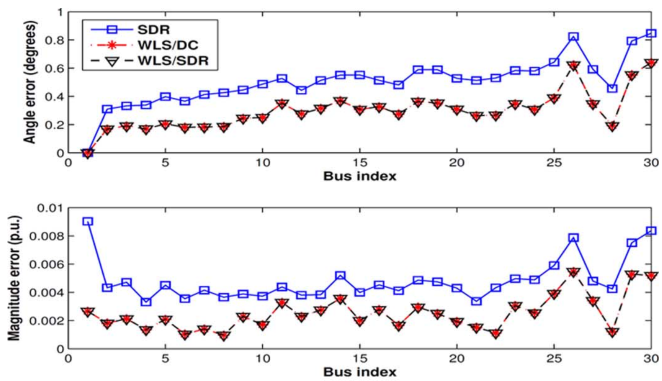
The SDR-based SE algorithms are first tested in a centralized setting using the IEEE 30-bus system with 41 lines from [26], and compared to existing WLS methods that based on Gauss-Newton iterations. Different legacy meter or PMU placements and variable levels of voltage angles are considered. MATPOWER [33] is used to generate the pertinent power flow and meter measurements. In addition, its SE function `doSE` is adapted to realize the WLS Gauss-Newton iterations. The iterations terminate either upon convergence, or, once the condition number of the approximate linearization exceeds 10^8, which flags divergence of the iterates. To solve the (augmented) SDR-based SE problems, the MATLAB-based optimization modeling package `CVX` [14] is used, together with the interior-point solver SeDuMi [27]. Additional tests of PMU-aided SE and distributed implementations are also available at the end of this section.

1) Test Case 1: The real and reactive power flows along all 41 lines are measured, together with voltage magnitudes at 30 buses. Independent Gaussian noise corrupts all measurements, with σ_ℓ equal to 0.02 at power meters, and 0.01 at voltage meters. The empirical estimation errors $\|\mathbf{v} - \hat{\mathbf{v}}\|_2$ are averaged over 500 Monte-Carlo realizations for the SDR approach and the WLS one using various initializations listed in Table I. In each realization, except for the reference bus phasor $V_{\mathrm{ref}} = 1$, each bus has its actual voltage magnitude Gaussian distributed with mean 1 and variance 0.01, and its actual voltage angle uniformly distributed over $[-\theta,\ \theta]$. Three choices of θ are tested, namely $\theta = 0.3\pi, 0.4\pi$, and 0.5π. The percentage of realizations that the iterative WLS method converges is also given in parentheses. The SDR estimator is recovered from the SDR-based SE solution $\hat{\mathbf{V}}$, by picking the minimum-cost vector over the eigenvector solution and 50 randomization samples. The first WLS estimator, termed WLS/FVP, corresponds to the WLS solution initialized by the flat-voltage profile (FVP) point; that is, the one using the all-one vector as initial guess. For a better starting point, the second WLS/DC one is obtained by initializing the voltage angles using the DC model SE [1, Sec. 2.8], and the magnitudes using the corresponding meter measurements. To gauge the SDR approach's near-optimal performance wrt the global solution, the SDR estimator is further used to initialize the WLS iterations, and the abbreviation used for this estimator is WLS/SDR.

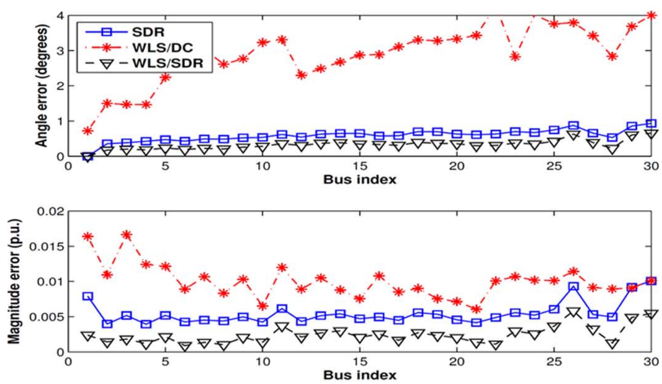
Table I clearly shows that the DC model based SE provides a much better initialization compared to the FVP one, in terms of smaller estimation error and higher probability of convergence. When the actual voltage angles are small ($\theta = 0.3\pi$), the WLS linear approximation is quite accurate with either the FVP or the DC model based initialization, and thus convergence to the global optimum can be guaranteed. Especially for the WLS/DC with $\theta = 0.3\pi$, the empirical error 0.042 can be considered as

TABLE I
ESTIMATION ERROR WITH % OF CONVERGENCE FOR TEST CASE 1

θ	SDR	WLS/FVP	WLS/DC	WLS/SDR
0.3π	0.070	0.097 (98.6%)	0.042 (100%)	0.042 (100%)
0.4π	0.081	0.593(88.6%)	0.255 (97.2%)	0.044 (100%)
0.5π	0.088	2.228 (68.6%)	1.161 (88.0%)	0.047 (100%)

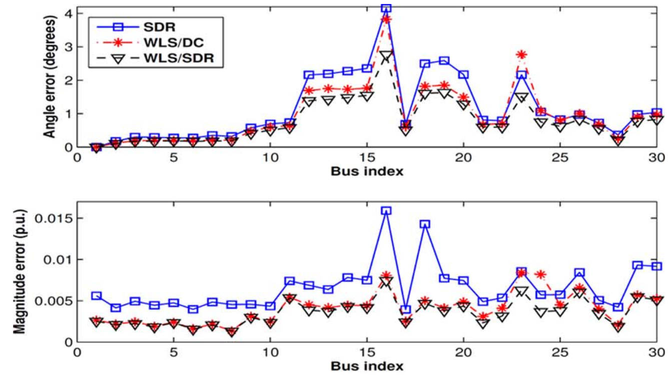


(a)

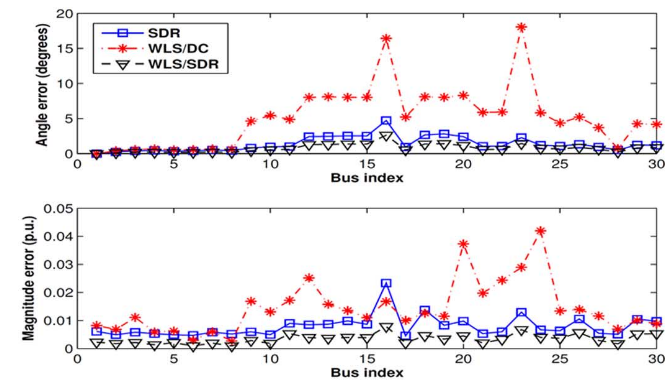


(b)

Fig. 2. Comparing estimation errors in voltage magnitudes and angles between SDR and WLS solvers at different buses for Test Case 1 with (a) $\theta = 0.3\pi$; and (b) $\theta = 0.4\pi$.



(a)



(b)

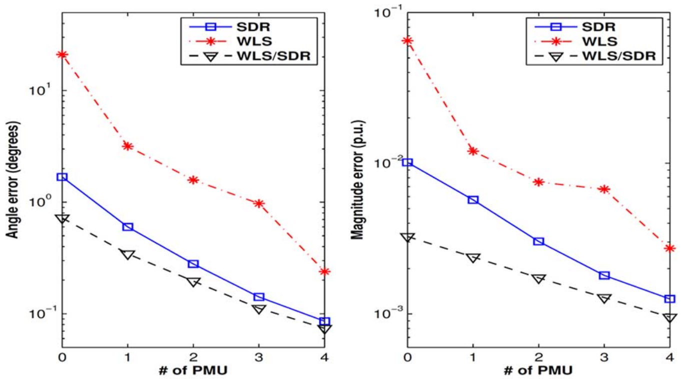
Fig. 3. Comparing estimation errors in voltage magnitudes and angles between SDR and WLS solvers at different buses for Test Case 2 with (a) $\theta = 0.2\pi$; and (b) $\theta = 0.3\pi$.

TABLE II
ESTIMATION ERROR WITH % OF CONVERGENCE FOR TEST CASE 2

θ	SDR	WLS/FVP	WLS/DC	WLS/SDR
0.2π	0.174	0.265 (96.6%)	0.148 (99.4%)	0.115 (100%)
0.3π	0.203	1.759 (68.6%)	0.653 (91.4%)	0.109 (100%)
0.4π	0.247	3.521 (47.0%)	2.141 (66.0%)	0.104 (100%)

the benchmark estimation error achieved for such meter placements and noise levels. As θ increases however, the nonlinearity in the measurement model is responsible for the performance degradation exhibited by the WLS/FVP and WLS/DC estimators. Interestingly, estimation accuracy of the SDR estimator is still competitive to the benchmark and comes close to the global optimum, verifying the analytical insights provided by Proposition 1. With any choice of θ, the WLS/SDR estimator is always convergent and attains the benchmark accuracy 0.042 within numerical accuracy. This suggests that the SDR-based estimator comes with numerically verifiable approximation bounds relative to the global optimum. Further evidence to this effect is provided by the empirical voltage angle and magnitude errors per bus, which are plotted in Fig. 2. With $\theta = 0.3\pi$, Fig. 2(a) demonstrates that the SDR estimator exhibits error variation similar to both WLS/DC and WLS/SDR, which is roughly twice that of these two optimal schemes. However, as θ increases to 0.4π, Fig. 2(b) illustrates that the WLS/DC estimator blows up due to possible divergence especially in the angle estimates, while both the SDR and WLS/SDR show comparable accuracy as well as analogous performance. This test case numerically supports the analytical insights on the near-optimal performance of the proposed SDR-based SE algorithm.

2) Test Case 2: Here 19 line flow meters and 15 bus injection meters are placed according to the setting in Fig. 4 of [7], together with 30 voltage magnitude meters. Although full observability is ensured, a certain number of lines is not directly observed. Thus, quadratic coupling of measurements affects SE in those indirectly observed lines and leads to performance degradation, as confirmed by Table II. The relative performance and convergence probability among different estimators for various choices of θ follow the trends of Test Case 1, but the placement here yields a larger benchmark estimation error around 0.11. As a result, the impact of initialization is more pronounced here, as for $\theta = 0.3\pi$ the WLS/DC iterations diverge in nearly 10% of realizations. A close look at the error plots in Fig. 3(b) reveals that the estimation errors at buses 1 through 8 approach the optimal ones, especially for the angle errors. Hence, divergence of



(a) (b)

Fig. 4. Comparing estimation errors between SDR and WLS solvers versus the number of PMUs in Test Case 3 for (left) angle estimates; and (b) magnitude estimates.

TABLE III
AVERAGE RUNNING TIMES IN SECONDS

# of buses	WLS	SDR
30	0.216	1.62
57	0.558	4.32
118	2.87	21.6

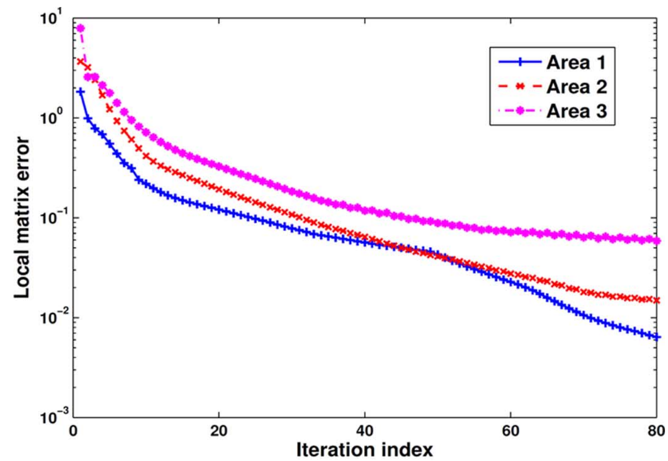
the WLS/DC estimator due to insufficient direct flow measurements affects the estimates at buses 9 through 30. Nonetheless, the SDR-based SE still offers near-optimal performance relative to the benchmark WLS/SDR one for any θ.

The IEEE 57- and 118-bus systems from [26] have also been extensively tested under scenarios similar to those for the 30-bus system. The empirical estimation error performance for larger systems has been observed to be similar to that of the 30-bus system, which again confirms the near-optimality of the proposed SDR-based SE approach. The details are omitted here due to page limitations, but the run time comparison is given using the MATLAB® R2011a software on a typical Windows XP computer with a 2.8GHz CPU. Table III shows that the SDR-based method takes more time (around 20 seconds for the 118-bus system), while the WLS iterations incur increasing computational time mainly due to the higher divergence rate in larger systems.

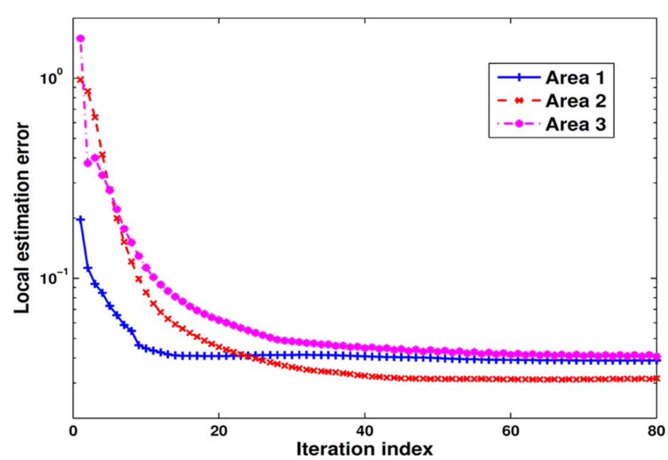
3) Test Case 3: To handle the insufficient direct measurements in Test Case 2, PMUs are deployed to enhance the SE performance offered by legacy measurements with $\theta = 0.4\pi$. The PMU meter noise level is set to $\check{\sigma}_m = 0.002$ at all buses. The convex relaxation approach using the *A-optimal* placement of PMUs [17] selects the four buses from $\{10, 12, 27, 15\}$ to be equipped with PMUs sequentially. Since the WLS iterations with only legacy measurements are not guaranteed to converge, as verified by Table II, the sequential approach of including PMU data in [34] does not lead to improved convergence. Hence, the joint WLS-based SE approach using polar representation of the state is adopted for comparison. The WLS initialization combines the linear estimates at those observable buses based on the PMU measurements, with the DC model angle estimates mentioned earlier. The SDR estimator is obtained from the solution $\hat{\mathbf{X}}$ in (14) with the approximate

TABLE IV
ESTIMATION ERROR WITH % OF CONVERGENCE FOR TEST CASE 3

# of PMU	SDR	WLS	WLS/SDR
0	0.247	2.141 (66.0%)	0.104 (100%)
1	0.122	0.678 (94.4%)	0.063 (100%)
2	0.062	0.335 (96.6%)	0.040 (100%)
3	0.036	0.280 (98.8%)	0.025 (100%)
4	0.019	0.061 (99.6%)	0.015 (100%)



(a)



(b)

Fig. 5. (a) Per area state matrix error and (b) state vector estimation error, versus ADMM iterations.

rescaling factor $\hat{c}^\star$ based only on PMU measurements, which also serves as initial guess to obtain the WLS/SDR estimator. The empirical estimation errors for 0 to 4 PMUs are listed in Table IV, where the PMU absent results are repeated from Table II. As the number of PMUs increases, the estimation accuracy as well as the probability of convergence improve for the WLS estimator. Still, there is a considerable gap relative to the other two estimators based on the SDR solution. It is also worth noticing that using more PMUs the SDR estimator approximates better the optimal WLS/SDR one, with the approximation gap coming very close to 1. This is illustrated in Fig. 4, where empirical angle and magnitude errors in the logarithmic scale are averaged over all 30 buses and plotted versus the number of PMUs deployed. The difference between the SDR and WLS/SDR estimators strictly diminishes as the number of PMUs increases, which suggests that the approximation accuracy of the SDR approach relative to the globally optimum one can be markedly aided by the use of PMU data.

4) Test Case 4: To verify the proposed distributed SE method, the IEEE 118-bus system is tested using the three-area partition in [17]. All three areas measure their local bus voltage magnitudes, as well as real and reactive power flow levels at all lines. The overlaps among the areas form a tree communication graph to construct the equality constraints enforced in (21).

To illustrate convergence of the ADMM iterations to the centralized SE solution $\hat{\mathbf{V}}$, the local matrix error $\|\mathbf{W}_k^i - \hat{\mathbf{V}}_{(k)}\|_F$ is plotted versus the iteration index i in Fig. 5(a) for every control area k. Clearly, all the local iterates converge to (approximately with a linear rate) their counterparts in the centralized solution. In addition, as the estimation task is of interest here, the local estimation error $\|\hat{\mathbf{v}}_k^i - \mathbf{v}_k\|_2$ is also plotted in Fig. 5(b), where $\hat{\mathbf{v}}_k$ is the estimate of bus voltages at $\mathcal{N}_k$ obtained from the iterate $\mathbf{W}_k^i$ using the eigen-decomposition method. Interestingly, the estimation error costs converge within the estimation accuracy of around 10^{-2} after about 20 iterations (less than 10 iterations for Area 1), even though the local matrix has not yet converged. In addition, these error costs decrease much more fast in the first couple of iterations. This demonstrates that even with only a limited number of iterations the estimation accuracy can be greatly boosted in practise, which in turn makes inter-area communication overhead more affordable.

VII. Conclusions and Current Research

New SDR-based SE schemes were developed in this paper for power system monitoring, by tactfully reformulating the nonlinear relationship between legacy meter measurements and complex bus voltages. The nonconvex SE problem was relaxed to a convex SDP one to render it efficiently solvable via existing interior-point methods. In addition, simplified conditions for the SDP-SE approach to attain the global optimum were provided to support the near-optimal performance in practical systems. To account for recent developments in PMU technology, linear state measurements were also incorporated to enhance the proposed SDR-based SE framework. A distributed SDR-SE method has been further developed to reduce the computational burden and facilitate real-time wide-area monitoring, thanks to the interactions among multiple control areas. Extensive numerical tests demonstrated the near-optimal performance of the proposed approaches. Further enhancements to the SDR-based SE framework are currently pursued under the cyber-security context and toward further complexity reduction in terms computations and communications.

References

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

[2] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *Int. J. Elect. Power Energy Syst.*, vol. 30, no. 6–7, pp. 383–392, Jul. 2008.

[3] A. R. Bergen and V. Vittal, *Power System Analysis*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2000.

[4] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 1995.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundat. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[7] J. Chen and A. Abur, "Enhanced topology error processing via optimal measurement design," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 845–852, Aug. 2008.

[8] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sep. 2013.

[9] W. H. Kersting, *Distribution system modeling and analysis*, 2nd ed. Boca Raton, NJ, USA: CRC, 2006.

[10] G. B. Giannakis, V. Kekatos, N. Gatsis, S.-J. Kim, H. Zhu, and B. Wollenberg, "Monitoring and optimization for power grids: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 107–128, Sep. 2013.

[11] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *J. ACM*, vol. 42, no. 6, pp. 1115–1145, 1995.

[12] A. Gomez-Exposito, A. Abur, A. d. l. V. Jaen, and C. Gomez-Quiles, "A multilevel state estimation paradigm for smart grids," *Proc. IEEE*, vol. 99, no. 6, pp. 952–976, Jun. 2011.

[13] A. Gomez-Exposito, A. Abur, P. Rousseaux, A. d. l. V. Jaen, and C. Gomez-Quiles, "On the use of PMUs in power system state estimation," in *Proc. 17th Power Syst. Comput. Conf.*, Stockholm, Sweden, Aug. 22–26, 2011.

[14] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," Apr. 2011, [Online]. Available: http://cvxr.com/cvx/

[15] R. A. Jabr, "Exploiting sparsity in SDP relaxations of the OPF problem," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 1138–1139, May 2012.

[16] R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, "Positive definite completions of partial Hermitian matrices," *Linear Algebra and its Applicat.*, vol. 58, pp. 109–124, Apr. 1984.

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

[18] V. Kekatos, G. B. Giannakis, and B. F. Wollenberg, "Optimal placement of phasor measurement units via convex relaxation," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1521–1530, Aug. 2012.

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

[20] A. Y. S. Lam, B. Zhang, and D. Tse, Distributed algorithms for optimal power flow problem, [Online]. Available: http://arxiv.org/abs/1109.5229

[21] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

[22] J. Lavaei, D. Tse, and B. Zhang, "Geometry of power flows in tree networks," in *Proc. IEEE PES General Meeting*, 2012, pp. 1–8.

[23] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Comm. Security*, Chicago, IL, USA, Nov. 9–13, 2009.

[24] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

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

[26] Power Systems Test Case Archive, Univ. of Washington, [Online]. Available: http://www.ee.washington.edu/research/pstca/

[27] J. F. Sturm, "Using SeDuMi 1.02, a MatLab toolbox for optimization over symmetric cones," *Optimiz. Meth. Software*, vol. 11–12, pp. 625–653, Aug. 1999 [Online]. Available: http://sedumi.mcmaster.ca

[28] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[29] Y. Weng, Q. Li, R. Negi, and M. Ilic, "Semidefinite programming for power system state estimation," in *Proc. PES General Meeting*, Jul. 2012.

[30] Y. Weng, Q. Li, R. Negi, and M. Ilic, "Distributed algorithm for SDP state estimation," in *Proc. Innovative Smart Grid Technol. (ISGT)*, Jan. 2013.

[31] A. J. Wood and B. F. Wollenberg, *Power generation, operation, and control*. New York, NY, USA: Wiley, 1984.
[32] L. Xie, D.-H. Choi, S. Kar, and H. V. Poor, "Fully distributed state estimation for wide-area monitoring systems," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1154–1169, Sep. 2012.
[33] 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.
[34] M. Zhou, V. A. Centeno, J. S. Thorp, and A. G. Phadke, "An alternative for including phasor measurements in state estimators," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1930–1937, Nov. 2006.
[35] H. Zhu and G. B. Giannakis, "Estimating the state of AC power systems using semidefinite programming," in *Proc. 43rd North Amer. Power Symp.*, Boston, MA, USA, Aug. 4–6, 2011.
[36] H. Zhu and G. B. Giannakis, "Robust power system state estimation for the nonlinear AC flow model," in *Proc. 44th North Amer. Power Symp.*, Univ. of Illinois at Urbana-Champaign, Urbana, IL, USA, Sep. 9–11, 2012.
[37] H. Zhu and G. B. Giannakis, "Multi-area state estimation using distributed SDP for nonlinear power systems," in *Proc. 3rd IEEE Smart Grid Commun.*, Nov. 2012, pp. 623–628.

Hao Zhu (M'12) is currently an Assistant Professor of electrical and computer engineering at the University of Illinois, Urbana-Champaign. She received her B.S. from Tsinghua University, Beijing, China, in 2006 and the M.Sc. and Ph.D. degrees from the University of Minnesota, Minneapolis, in 2009 and 2012, respectively. Her current research interests include power system monitoring and operations, dynamics and stability, and energy data analytics. She received the two-year UMN Graduate School Fellowship in 2006, and the UMN Doctoral Dissertation Fellowship in 2011.

Georgios B. Giannakis (F'97) received his Diploma in electrical engineering from the National Technical University of Athens, Greece, 1981. From 1982 to 1986 he was with the University of Southern California (USC), where he received his M.Sc. in electrical engineering in 1983, the M.Sc. in mathematics in 1986, and the Ph.D. in electrical engineering in 1986. Since 1999, he has been a professor with the University of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications in the Electrical and Computer Engineering Department, and serves as director of the Digital Technology Center.

His general interests span the areas of communications, networking and statistical signal processing—subjects on which he has published more than 365 journal papers, 625 conference papers, 21 book chapters, two edited books and two research monographs (h-index 108). Current research focuses on sparsity and big data analytics, wireless cognitive radios, mobile ad hoc networks, renewable energy, power grid, gene-regulatory, and social networks. He is the (co-) inventor of 22 patents issued, and the (co-) recipient of 8 best paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, and the G. W. Taylor Award for Distinguished Research from the University of Minnesota. He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SP Society.