Estimating the State of AC Power Systems using Semidefinite Programming

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

***Abstract*—An important monitoring task for power networks is to estimate accurately the underlying grid state, which is useful for security-constrained dispatch and power system control. For nonlinear AC power systems, the state estimation (SE) problem is inherently nonconvex giving rise to many local optima. As a result, existing estimators used extensively in practice rely on iterative optimization methods, which are destined to return only *locally* optimal solutions. A semidefinite programming (SDP) based approach is introduced in this paper, which relies on *convex* relaxation of the original SE problem and thereby renders it efficiently solvable. A sufficient condition also becomes available to guarantee that the dual SDP problem attains zero duality gap, and thus ensure that the *globally* optimal SE solution is achievable in polynomial time. The novel scheme's ability to markedly outperform existing iterative alternatives is corroborated through numerical tests on the standard IEEE 14-bus benchmark system.**

***Index Terms*—Power system, state estimation, semidefinite programming, polynomial-time algorithm.**

I. INTRODUCTION

Although the power grid in North America has been lauded as the most important engineering achievement of the 20th century by the National Academy of Engineering [15], the modern power grid faces major challenges related to environmental, stability, security, and market diversity issues [13]. The smart grid vision aspires to address the challenges associated with the aging infrastructure by capitalizing on advanced information technologies in sensing, machine learning, control, communications, and optimization. Among all important avenues where such technologies can make significant impact on power networks, agile monitoring techniques that can estimate the status of large-scale grids with precision manifest their significance in facilitating accurate system awareness and prompt operation of the network.

The *state estimation* (SE) task for power systems refers to the process of acquiring estimates of the voltage phasors at all buses in the network. Determination of the system's operating condition through SE is an important prerequisite for system control and economic dispatch with security constraints; see e.g., [1], [10]. In addition to the huge volume of the data requiring real-time processing, the SE problem is inherently nonconvex giving rise to many local optima due to the nonlinear coupling present among measurement variables. Specifically, SE falls under the general class of nonlinear least-squares (LS) problems, for which the Gauss-Newton method is the "workhorse" solver typically employed in practice [3, Sec. 1.5]. This iterative scheme has also been reckoned as the algorithmic foundation of SE; see e.g., [1, Ch. 2] and [14, Ch. 12]. Using the Taylor expansion around a given initialization point, the Gauss-Newton method approximates the nonlinear LS cost by a linear LS one, whose minimum is used to initialize the ensuing iteration. Since this iterative procedure actually relates to a gradient descent solver of nonconvex problems [3, Sec. 1.5], it inevitably faces challenges pertaining to sensitivity of the initial guess and convergence issues. SE iterations typically start with a "flat voltage profile," where all bus voltages are initialized with a common real number. As there are no guarantees ensuring convergence to the global optimum, existing variants have asserted numerical stability [1, Ch. 3], but they are limited to improving the linearized LS cost per iteration. The grand challenge for many years has been to develop a solver attaining the *global optimum* at *polynomial-time* complexity.

The goal of the present paper is to leverage physical properties of AC power systems in order to develop a polynomial-time SE algorithm, which offers the potential to find a globally optimum state estimate. To this end, a novel convex semidefinite programming (SDP) formulation is introduced for the SE problem (Section III). This formulation can afford optimizing the dual SDP problem at reduced complexity, and has the potential to exhibit zero duality gap with the primal SE problem (Section IV). In a different context, namely that of optimum power flow via SDP-based relaxation, a zero-duality gap result was recently reported in [6]. However, its counterpart for the SE problem was not considered. The present SDP formulation enables computationally efficient SDP solvers for SE, which are neither sensitive to initialization nor they require iterative updates. The analytical results and numerical tests presented in Section V corroborate the merits of the proposed SDP approach over the Gauss-Newton method. Conclusions and current research threads are outlined in Section VI.

Notation: Upper (lower) boldface letters will be used for matrices (column vectors); $(\cdot)^{\mathbf{T}}$ denotes transposition; $(\cdot)^*$ complex-conjugate transposition; $\mathrm{Re}(\cdot)$ the real part; $\mathrm{Im}(\cdot)$ the imaginary part; $\mathrm{Tr}(\cdot)$ the matrix trace; $\mathrm{rank}(\cdot)$ the matrix rank; $\mathbf{0}$ the all-zero matrix; $\|\cdot\|_p$ the vector $p-$norm for $p \geq 1$; $[\cdot]_{i:j}$ the sub-vector comprising the i-th through j-th entries; and $|\cdot|$ the magnitude or cardinality of a set.

This work is supported by the NSF grants CCF-0830480, CCF-1016605, ECCS-0824007, and ECCS-1002180.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA. Tel/fax: (612)624-9510/625-2002, emails: {zhuh,georgios}@umn.edu

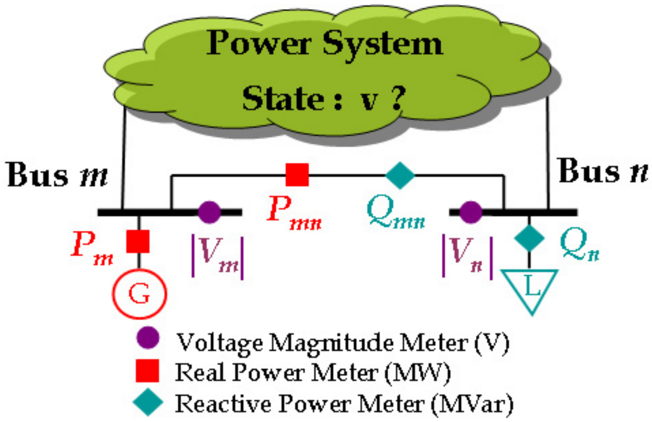


Fig. 1. SE through measuring power system variables. (Here bus m connects to the generation, and bus n to the load.)

II. Modeling and Problem Statement

Consider a power transmission network with N buses denoted by the set of nodes $\mathcal{N} := \{1, \ldots, N\}$, and transmission lines represented by the set of edges $\mathcal{E} := \{(n, m)\} \subseteq \mathcal{N} \times \mathcal{N}$. To estimate the complex voltage V_n at each bus $n \in \mathcal{N}$, measurements are taken for a subset of the following system variables (see also Fig. 1):

- $P_n(Q_n)$: 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 [14, Ch. 4], these measurements are nonlinearly related with the power system state of interest $\mathbf{v} := [V_1, \ldots, V_N]^T \in \mathbb{C}^N$. To specify this relationship, collect the injected currents of all buses in $\mathbf{i} := [I_1, \ \ldots, I_N]^T \in \mathbb{C}^N$, and let $\mathbf{Y} \in \mathbb{C}^{N \times N}$ represent the grid's symmetric bus admittance matrix. Kirchoff's law in vector-matrix form dictates, see e.g., [2, Sec. 9.1]

$$\mathbf{i} = \mathbf{Y}\mathbf{v} \tag{1}$$

where the (m, n)-th entry of $\mathbf{Y}$ is given by

$$Y_{mn} := \begin{cases} -y_{mn}, & \text{if } (m,n) \in \mathcal{E} \\ y_{nn} + \sum_{\nu \in \mathcal{N}_n} y_{n\nu}, & \text{if } m = n \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

with y_{mn} denoting the line admittance between buses m and n, y_{nn} bus n's admittance to the ground, and $\mathcal{N}_n$ the set of all buses linked to bus n through 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

$$I_{mn} = \bar{y}_{mn} V_m + y_{mn}(V_m - V_n). \tag{3}$$

The AC power flow model further asserts that the apparent power injection into bus n is given by

$$P_n + j\ Q_n = V_n I_n^* \tag{4}$$

while the apparent power flow from bus m to bus n by

$$P_{mn} + j\ Q_{mn} = V_m I_{mn}^*. \tag{5}$$

Finally, the squared bus voltage magnitude is given by

$$|V_n|^2 = \text{Re}^2(V_n) + \text{Im}^2(V_n). \tag{6}$$

Clearly, (4)-(6) specify how all measurable variables listed earlier are nonlinearly (quadratically) related to $\mathbf{v}$.

Collect the actual (possibly noisy) measurements of these variables in the $L \times 1$ vector $\mathbf{z} := [\{\check{P}_n\}_{n \in \mathcal{N}_P}, \{\check{Q}_n\}_{n \in \mathcal{N}_Q}, \{\check{P}_{mn}\}_{(m,n) \in \mathcal{E}_P}, \{\check{Q}_{mn}\}_{(m,n) \in \mathcal{E}_Q}, \{|\check{V}_n|^2\}_{n \in \mathcal{N}_V}]^T$, where the check mark ˇ is used to differentiate measured values from the error-free variables[1]. The ℓ-th entry of $\mathbf{z}$ can be written as

$$z_\ell = h_\ell(\mathbf{v}) + \epsilon_\ell \tag{7}$$

where $h_\ell(\cdot)$ denotes the nonlinear relationship specified by (4)-(6), and ϵ_ℓ is the zero-mean Gaussian distributed measurement error at the ℓ-th meter with variance σ_ℓ^2, which is assumed independent across meters.

Given $\mathbf{z}$, the goal of SE is to estimate $\mathbf{v}$. Due to the independence among all measurement errors, the maximum-likelihood (ML) optimality criterion reduces to the weighted least-squares (WLS) one, yielding the state estimator as

$$\boxed{\begin{aligned} &\hat{\mathbf{v}} := \arg\min_{\mathbf{v}} \sum_{\ell=1}^{L} w_\ell [z_\ell - h_\ell(\mathbf{v})]^2 && \text{(8a)} \\ &\text{s.to } \mathbf{i} = \mathbf{Y}\mathbf{v} \text{ and } I_{mn} = \bar{y}_{mn} V_m + y_{mn}(V_m - V_n) && \text{(8b)} \end{aligned}}$$

where the ℓ-th entry of the weight vector $\mathbf{w}$ is $w_\ell := 1/(\sigma_\ell^2)$. The SE problem in (8) is nonconvex; in fact, substituting the constraint (8b) back into (8a) shows that it falls in the general class of nonlinear WLS problems.

The Gauss-Newton iterative solver of nonlinear WLS problems [3, Sec. 1.5] has been widely used for SE; see e.g., [1, Ch. 2] and [14, Ch. 12]. Using the Taylor expansion around a given starting point, the pure form of Gauss-Newton methods approximates the cost in (8a) with a linearized WLS one. Then, the minimizer to the approximate cost initializes the next iteration. This iterative procedure is closely related to gradient descent methods for solving nonconvex problems, see e.g., [3, Ch. 1], which are known to encounter two issues: i) sensitivity to the initial guess; and ii) convergence concerns. Typical WLS-based SE iterations start with a flat voltage profile, where all bus voltages are initialized with the same real number, but this does not guarantee convergence to the global optimum. Existing variants have asserted improved numerical stability [1, Ch. 3], but they are all limited to improving the approximate WLS cost per iteration. The grand challenge for many years has been to develop a solver attaining the *global optimum* at *polynomial-time* complexity. The next

[1]The squared voltage magnitude measurement is considered here instead of the magnitude measurement, for consistency with other measurements included in Section III. However, by adopting the magnitude measurement noise model $|\check{V}_n| = |V_n| + \epsilon_V$, where ϵ_V is zero-mean Gaussian with small variance σ_V^2, an approximate model holds for the squared magnitude too; namely $|\check{V}_n|^2 \approx |V_n|^2 + \epsilon'_V$, where ϵ'_V has variance $4|\check{V}_n|^2\sigma_V^2$.

section addresses this challenge by appropriately reformulating SE to a relaxed semidefinite programming (SDP) problem.

III. Solving SE using SDP

In this section, SE is relaxed to an SDP problem. This SDP formulation will not only lead to efficient solvers by unveiling the problem's hidden convexity structure, but will also turn out to suggest the possibility for achieving global optimality.

To develop an SDP formulation for the SE problem in (8), it is necessary to first show that each system variable measured in $\mathbf{z}$ can be expressed linearly with respect to (w.r.t.) the outer-product matrix

$$\mathbf{X} := \mathbf{x}\mathbf{x}^T, \quad \text{where} \quad \mathbf{x} := [\mathrm{Re}^T(\mathbf{v})\ \mathrm{Im}^T(\mathbf{v})]^T \in \mathbb{R}^{2N}. \tag{9}$$

To this end, let $\mathbf{e}_n$ denote the n-th canonical basis of $\mathbb{R}^N$ for all $n = 1, \ldots, N$, and define the following admittance-related matrices

$$\mathbf{Y}_n := \mathbf{e}_n\mathbf{e}_n^T\mathbf{Y} \tag{10a}$$

$$\mathbf{Y}_{mn} := (\bar{y}_{mn} + y_{mn})\mathbf{e}_m\mathbf{e}_m^T - y_{mn}\mathbf{e}_m\mathbf{e}_n^T\,. \tag{10b}$$

For future use, define also the real matrices

$$\mathbf{H}_{P,n} := \frac{1}{2}\begin{bmatrix} \mathrm{Re}(\mathbf{Y}_n + \mathbf{Y}_n^T) & \mathrm{Im}(\mathbf{Y}_n^T - \mathbf{Y}_n) \\ \mathrm{Im}(\mathbf{Y}_n - \mathbf{Y}_n^T) & \mathrm{Re}(\mathbf{Y}_n + \mathbf{Y}_n^T) \end{bmatrix} \tag{11a}$$

$$\mathbf{H}_{Q,n} := \frac{-1}{2}\begin{bmatrix} \mathrm{Im}(\mathbf{Y}_n + \mathbf{Y}_n^T) & \mathrm{Re}(\mathbf{Y}_n - \mathbf{Y}_n^T) \\ \mathrm{Re}(\mathbf{Y}_n^T - \mathbf{Y}_n) & \mathrm{Im}(\mathbf{Y}_n + \mathbf{Y}_n^T) \end{bmatrix} \tag{11b}$$

$$\mathbf{H}_{P,mn} := \frac{1}{2}\begin{bmatrix} \mathrm{Re}(\mathbf{Y}_{mn} + \mathbf{Y}_{mn}^T) & \mathrm{Im}(\mathbf{Y}_{mn}^T - \mathbf{Y}_{mn}) \\ \mathrm{Im}(\mathbf{Y}_{mn} - \mathbf{Y}_{mn}^T) & \mathrm{Re}(\mathbf{Y}_{mn} + \mathbf{Y}_{mn}^T) \end{bmatrix} \tag{11c}$$

$$\mathbf{H}_{Q,mn} := \frac{-1}{2}\begin{bmatrix} \mathrm{Im}(\mathbf{Y}_{mn} + \mathbf{Y}_{mn}^T) & \mathrm{Re}(\mathbf{Y}_{mn} - \mathbf{Y}_{mn}^T) \\ \mathrm{Re}(\mathbf{Y}_{mn}^T - \mathbf{Y}_{mn}) & \mathrm{Im}(\mathbf{Y}_{mn} + \mathbf{Y}_{mn}^T) \end{bmatrix} \tag{11d}$$

$$\mathbf{H}_{V,n} := \begin{bmatrix} \mathbf{e}_n\mathbf{e}_n^T & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_n\mathbf{e}_n^T \end{bmatrix}. \tag{11e}$$

Using these definitions, the first result is given in the following lemma.

Lemma 1: *All error-free system variables are linearly related with the matrix* $\mathbf{X}$ *in* (9)*, as*

$$P_n = \mathrm{Tr}(\mathbf{H}_{P,n}\mathbf{X}) \tag{12a}$$

$$Q_n = \mathrm{Tr}(\mathbf{H}_{Q,n}\mathbf{X}) \tag{12b}$$

$$P_{mn} = \mathrm{Tr}(\mathbf{H}_{P,mn}\mathbf{X}) \tag{12c}$$

$$Q_{mn} = \mathrm{Tr}(\mathbf{H}_{Q,mn}\mathbf{X}) \tag{12d}$$

$$|V_n|^2 = \mathrm{Tr}(\mathbf{H}_{V,n}\mathbf{X}). \tag{12e}$$

Thus, each meter measurement in (7) *can be written as*

$$z_\ell = h_\ell(\mathbf{v}) + \epsilon_\ell = \mathrm{Tr}(\mathbf{H}_\ell\mathbf{X}) + \epsilon_\ell \tag{13}$$

where $\mathbf{H}_\ell$ *is specified in accordance with* (11a)-(11e).

Proof: To establish (12a), substitute (1) into (4), and use successively (10a), (11a), as well as (9) to obtain

$$\begin{aligned} P_n &= \mathrm{Re}(V_n I_n^*) = \mathrm{Re}(V_n^* I_n) = \mathrm{Re}(\mathbf{v}^*\mathbf{e}_n\mathbf{e}_n^T\mathbf{i}) = \mathrm{Re}(\mathbf{v}^*\mathbf{Y}_n\mathbf{v}) \\ &= \mathbf{x}^T\mathbf{H}_{P,n}\mathbf{x} = \mathrm{Tr}(\mathbf{H}_{P,n}\mathbf{X}). \end{aligned} \tag{14}$$

Likewise, the linear in $\mathbf{X}$ relations (12b)-(12e) for other variables follow the same steps detailed in (14). ■

Lemma 1 implies the following equivalent form for the SE problem in (8) [cf. (13)]

$$\hat{\mathbf{X}}_1 := \arg\min_{\mathbf{X}} \sum_{\ell=1}^{L} w_\ell \left[z_\ell - \mathrm{Tr}(\mathbf{H}_\ell\mathbf{X})\right]^2 \tag{15a}$$

$$\text{s.to } \mathbf{X} \succeq \mathbf{0}, \text{and } \mathrm{rank}(\mathbf{X}) = 1 \tag{15b}$$

where the positive semidefiniteness and rank constraints in (15b) equivalently ensure that there always exists a vector $\mathbf{x} \in \mathbb{R}^{2N}$ such that $\mathbf{X} = \mathbf{x}\mathbf{x}^T$.

Although this new SE formulation in (15) takes advantage of the linear relations between z_ℓ and $\mathbf{X}$ in (13), it is still nonconvex for two reasons: i) its cost (15a) has degree 4 w.r.t. the entries of the positive semidefinite matrix $\mathbf{X}$; and ii) the rank constraint in (15b) is a nonconvex one. To tackle the first source of nonconvexity, Schur's complement lemma, see e.g., [4, Appendix 5.5], can be leveraged to convert the summands in the cost (15a) to a linear function of an auxiliary vector $\boldsymbol{\alpha} \in \mathbb{R}^L$. Specifically, consider the following equivalent SE formulation of (15)

$$\left\{\hat{\mathbf{X}}_2, \hat{\boldsymbol{\alpha}}_2\right\} := \arg\min_{\mathbf{X},\boldsymbol{\alpha}} \left(\sum_{\ell=1}^{L} w_\ell\alpha_\ell\right) = \arg\min_{\mathbf{X},\boldsymbol{\alpha}} \mathbf{w}^T\boldsymbol{\alpha} \tag{16a}$$

$$\text{s.to} \quad \mathbf{X} \succeq \mathbf{0}, \text{and } \mathrm{rank}(\mathbf{X}) = 1 \tag{16b}$$

$$\begin{bmatrix} -\alpha_\ell & z_\ell - \mathrm{Tr}(\mathbf{H}_\ell\mathbf{X}) \\ z_\ell - \mathrm{Tr}(\mathbf{H}_\ell\mathbf{X}) & -1 \end{bmatrix} \preceq \mathbf{0}\ \forall \ell. \tag{16c}$$

The equivalence of all three SE formulations considered so far is asserted in the following proposition.

Proposition 1: *All three nonconvex optimization problems in* (8), (15)*, and* (16) *solve an equivalent SE problem of an AC power system. For the optima of these three problems, it holds that*

$$\hat{\mathbf{X}}_1 = \hat{\mathbf{X}}_2 = [\mathrm{Re}^T(\hat{\mathbf{v}})\ \mathrm{Im}^T(\hat{\mathbf{v}})]^T[\mathrm{Re}^T(\hat{\mathbf{v}})\ \mathrm{Im}^T(\hat{\mathbf{v}})]$$

$$\text{and} \quad \hat{\alpha}_{2,\ell} = \left[z_\ell - \mathrm{Tr}(\mathbf{H}_\ell\hat{\mathbf{X}}_2)\right]^2 \quad \forall \ell. \tag{17}$$

Proof: To establish the equivalence between (8) and (15), notice that the $2N \times 2N$ matrix $[\mathrm{Re}^T(\hat{\mathbf{v}})\ \mathrm{Im}^T(\hat{\mathbf{v}})]^T[\mathrm{Re}^T(\hat{\mathbf{v}})\ \mathrm{Im}^T(\hat{\mathbf{v}})]$ is a feasible solution of the problem (15). Similarly, express the rank-1 solution of (15) as $\hat{\mathbf{X}}_1 = \hat{\mathbf{x}}_1\hat{\mathbf{x}}_1^T$, and consider the $N \times 1$ complex vector $[\hat{\mathbf{x}}_1]_{1:N} + j\ [\hat{\mathbf{x}}_1]_{(N+1):2N}$, which is also feasible for (8). As Lemma 1 establishes the equivalence of the costs in (8a) and (15a), the two optimization problems yield solutions related as in (17).

It remains to show the equivalence of (15) and (16). Using Schur's complement lemma, the semidefinite constraint in (16c) equivalently requires $\alpha_\ell \geq [z_\ell - \mathrm{Tr}(\mathbf{H}_\ell\mathbf{X})]^2$, $\forall \ell = 1, \ldots, L$. Further, since (16a) minimizes a positive weighted sum of $\{\alpha_\ell\}_{\ell=1}^L$, it is ensured that equality holds at the optimum; i.e., $\hat{\alpha}_{2,\ell} = \left[z_\ell - \mathrm{Tr}(\mathbf{H}_\ell\hat{\mathbf{X}}_2)\right]^2$. Substituting this back into (16a) shows that $\hat{\mathbf{X}}_1$ and $\hat{\mathbf{X}}_2$ achieve the same minimum cost, which completes the proof. ■

A. Convexifying SE via Semidefinite Relaxation

Proposition 1 suggests the SDP-based SE formulation in (16), but it is still nonconvex due to the rank constraint of $\mathbf{X}$. However, the semidefinite relaxation technique, which amounts to dropping this rank constraint, has well-appreciated merits in the optimization community; see e.g., the seminal work [5]. Semidefinite relaxation has also recently provided new perspectives for many nonconvex problems in the field of signal processing and communications, thanks to its provable performance guarantees and implementation advantages; see e.g., [9] for a tutorial treatment on this topic. The contribution here consists in permeating the benefits of this powerful tool to estimating the state of AC power systems. Specifically, semidefinite relaxation leads to the following convex SDP formulation for SE:

$$\boxed{\begin{aligned} &\left\{\hat{\mathbf{X}}, \hat{\boldsymbol{\alpha}}\right\} := \arg\min_{\mathbf{X}} \mathbf{w}^T \boldsymbol{\alpha} && \text{(18a)}\\ &\text{s.to} \quad \mathbf{X} \succeq \mathbf{0} && \text{(18b)}\\ &\qquad \begin{bmatrix} -\alpha_\ell & z_\ell - \text{Tr}(\mathbf{H}_\ell \mathbf{X}) \\ z_\ell - \text{Tr}(\mathbf{H}_\ell \mathbf{X}) & -1 \end{bmatrix} \preceq \mathbf{0} \; \forall \ell. && \text{(18c)} \end{aligned}}$$

Several remarks are now in order.

Remark 1 (*Computational Complexity*). The SE problem in (18) is a convex SDP one, for which efficient schemes are available to obtain the global optimum using, e.g., the interior-point solver SeDuMi [12]. In general, the worst-case complexity of the SDP problem is $\mathcal{O}(\max\{L,N\}^4 \sqrt{N} \log(1/\epsilon))$ for a given solution accuracy $\epsilon > 0$ [9]. For typical power networks L is in the order of N, and thus the worst-case complexity becomes $\mathcal{O}(N^{4.5} \log(1/\epsilon))$. However, it is possible to further reduce SDP complexity by taking advantage of the sparsity or other special structures present. Notice that all matrices in (11) are markedly sparse and exhibit very special symmetric structures. For example, for each n, the matrix $\mathbf{H}_{V,n}$ in (11e) has only two non-zero entries, namely at the (n, n)-th and $(N+n, N+n)$-th positions, respectively. It is conceivable that these problem-specific properties can be used to further reduce complexity in solving the relaxed SE (18). Nonetheless, the SDP formulation in (18) yields a computationally efficient solver of the nonconvex SE problem, since its complexity is guaranteed to be polynomial in the problem size N, and the number of constraints L.

Remark 2 (*Finding a feasible solution to* (8)). Since the SDP formulation (18) is an approximation of the SE problem, it is necessary to recover from $\hat{\mathbf{X}}$ a feasible solution of the SE in (16). This is possible by eigen-decomposing $\hat{\mathbf{X}} = \sum_{i=1}^r \lambda_i \mathbf{u}_i \mathbf{u}_i^T$, where $r = \text{rank}(\hat{\mathbf{X}})$, $\lambda_1 \geq \cdots \geq \lambda_r > 0$ are the eigenvalues, and $\{\mathbf{u}_i \in \mathbb{R}^{2N}\}_{i=1}^r$ are the corresponding eigenvectors. Since the best rank-one approximation of $\hat{\mathbf{X}}$ is $\lambda_1 \mathbf{u}_1 \mathbf{u}_1^T$, in the sense of minimizing the Frobenius norm of the error, it follows that $\hat{\mathbf{x}} := \sqrt{\lambda_1}\mathbf{u}_1$, and the solution to the complex voltage state is found as $\hat{\mathbf{v}} = [\hat{\mathbf{x}}]_{1:N} + j\,[\hat{\mathbf{x}}]_{(N+1):2N}$. The randomization technique can also be used to improve this solution, and has the potential to achieve quantifiable approximation accuracy. The latter has been demonstrated when invoking semidefinite relaxation to e.g., communication problems [9].

Remark 3 (*Consensus-based distributed SE*). Existing SE algorithms are implemented at a central controller, namely the energy management system (EMS), to which the measurements are typically telemetered. Such central processing has two serious limitations: i) vulnerability to unreliable telemetry; and ii) high computational complexity at the EMS. Taking advantage of the decomposable convex problem structure, it is possible to devise a *distributed* solver of the proposed convex SDP-based SE in (18) using only local communication in a consensus-based collaboration; see also [17] for a related work in the context of distributed demodulation. Compared to centralized estimators, the distributed SE will not only be robust to failures and disturbances in the long-distance telemetry, but will also greatly reduce the computation burden at the EMS.

These three remarks manifest the benefits that the SDP formulation (18) brings in solving the SE problem efficiently and accurately. Additional insights are offered by the corresponding dual SDP problem, which is the subject of the ensuring section.

IV. Dual SE via SDP

To introduce the dual of the relaxed SE problem (18), start by defining the Lagrange multiplier associated with the ℓ-th matrix inequality in (18c) as

$$\boldsymbol{\mu}_\ell := \begin{bmatrix} \mu_{\ell,0} & \mu_{\ell,1} \\ \mu_{\ell,1} & \mu_{\ell,2} \end{bmatrix} \succeq \mathbf{0}. \tag{19}$$

Using (19), the Lagrangian corresponding to (18) becomes

$$\begin{aligned} &\mathcal{L}(\mathbf{X}, \boldsymbol{\alpha}, \{\boldsymbol{\mu}_\ell\}) \\ &:= \mathbf{w}^T \boldsymbol{\alpha} + \sum_{\ell=1}^L \text{Tr}\left\{ \begin{bmatrix} -\alpha_\ell & z_\ell - \text{Tr}(\mathbf{H}_\ell \mathbf{X}) \\ z_\ell - \text{Tr}(\mathbf{H}_\ell \mathbf{X}) & -1 \end{bmatrix} \boldsymbol{\mu}_\ell \right\} \\ &= \sum_{\ell=1}^L \left\{ (w_\ell - \mu_{\ell,0})\alpha_\ell - \mu_{\ell,2} + 2\left[\mu_{\ell,1} z_\ell - \text{Tr}(\mu_{\ell,1} \mathbf{H}_\ell \mathbf{X})\right] \right\}. \end{aligned} \tag{20}$$

The dual problem amounts to maximizing over $\boldsymbol{\mu}_\ell \succeq \mathbf{0}$, the Lagrangian $\mathcal{L}(\mathbf{X}, \boldsymbol{\alpha}, \{\boldsymbol{\mu}_\ell\})$ minimized over both the unconstrained $\boldsymbol{\alpha}$ and the positive definite $\mathbf{X} \succeq \mathbf{0}$. The Lagrangian is minimized when $w_\ell - \mu_{\ell,0} = 0$ and $-2 \sum_{\ell=1}^L \mu_{\ell,1} \mathbf{H}_\ell \succeq \mathbf{0}$, see e.g., [4, Sec. 4.6], which leads to the following form of the dual problem:

$$\{\hat{\boldsymbol{\mu}}_\ell\}_{\ell=1}^L := \arg\min_{\boldsymbol{\mu}_\ell} \sum_{\ell=1}^L \mu_{\ell,2} - 2\mu_{\ell,1} z_\ell \tag{21a}$$

$$\text{s.to} \quad \boldsymbol{\mu}_\ell = \begin{bmatrix} w_\ell & \mu_{\ell,1} \\ \mu_{\ell,1} & \mu_{\ell,2} \end{bmatrix} \succeq \mathbf{0}, \quad \forall \ell \tag{21b}$$

$$\mathbf{A} := 2 \sum_{\ell=1}^L \mu_{\ell,1} \mathbf{H}_\ell \preceq \mathbf{0}. \tag{21c}$$

Proposition 2: *The SDP problem in* (21) *is the dual of both the SE in* (16) *and the relaxed SE in* (18)*. Strong duality holds between* (21) *and* (18) *while the primal variable* $\mathbf{X} \succeq \mathbf{0}$ *corresponds to the Lagrange multiplier associated with the dual inequality constraint* (21c)*.*

Proof: Given the way $\text{Tr}(\mathbf{AX})$ appears in the Lagrangian $\mathcal{L}$, it follows that matrix $\mathbf{X}$ is the Lagrange multiplier for the dual constraint $\mathbf{A} \preceq \mathbf{0}$ in (21c) . Strong duality between these two convex problems holds because of Slater's condition. Specifically, the dual SE problem admits the trivial feasible point $\boldsymbol{\mu}_\ell = w_\ell \mathbf{I}$, $\forall \ell$.

To show that (21) is also the dual of the original SE problem in (16), recall that the constraint (16b) is equivalent to having $\mathbf{X} = \mathbf{x}\mathbf{x}^T$ for some vector $\mathbf{x}$. Therefore, the corresponding Lagrangian of the SE problem (16) can be obtained by substituting $\mathbf{X}$'s outer-product form into (20). Interestingly, the minimum of this new Lagrangian over $\boldsymbol{\alpha}$ and $\mathbf{x}$ is no different from the one in (20), and thus the same dual problem in (21) follows. ■

The dual SE problem (21) entails $2L$ unknown variables, where for typical power networks the number of measurements L is usually in the order of the number of lines $|\mathcal{E}|$, which is of order N. Compared to the $2N \times 2N$ matrix $\mathbf{X}$ in the primal problem (18), the dual SDP has considerably less variables to optimize over. Thus, additional complexity reduction is envisioned from this dualization; see e.g., [8]. Moreover, the dualization could provide further insights regarding the approximation accuracy of the relaxed SDP problem (18) to the SE problem (16), as summarized in the next remark.

Remark 4 (*Duality Gap*). A zero duality gap result for the SDP relaxation was recently reported for the optimal power flow (OPF) problem [6]. Using the sufficiency claim in [6] it can be shown that for the structured matrix $\mathbf{A}$ in (21c), if there exists a dual optimal solution such that $\mathbf{A}$ has a zero eigenvalue of multiplicity of 2, then there is no duality gap between the primal SE problem (16) and the dual SE one (21). Under this condition, efficiently solving the convex dual problem yields the globally optimal SE to the nonconvex one (16). However, the kind of grid topologies and meter placements ensuring this sufficient condition for zero duality gap remains open, and constitutes an interesting future research direction.

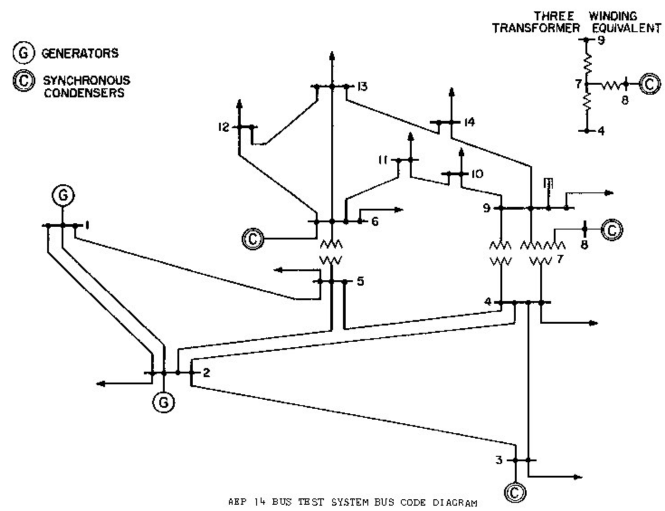


Fig. 2. The circuit of the IEEE 14-bus system taken from [11].

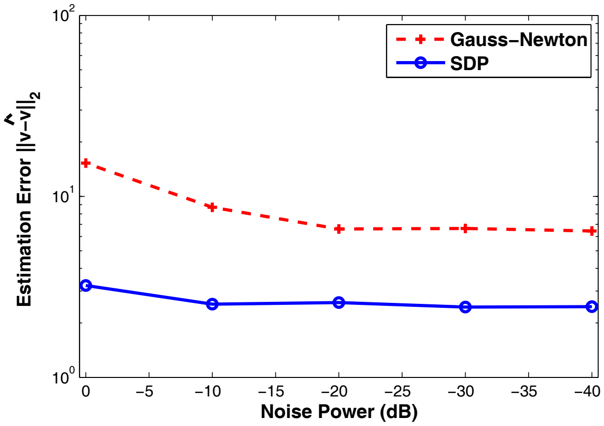
V. Numerical Tests

The proposed SDP-based SE algorithm is tested in this section using the IEEE 14-bus benchmark system, and its performance is compared with that of the traditional iterative Gauss-Newton method.

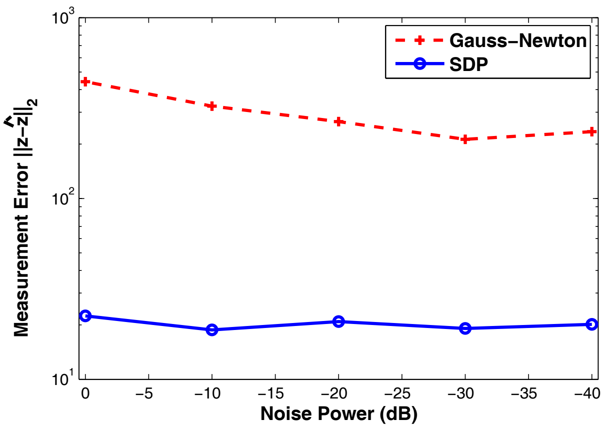
The IEEE 14-bus system depicted in Fig. 2 is a popular test case for power system simulations [11]. The software toolbox MATPOWER [16] is used to generate the measurements corresponding to this 14-bus system, as well as the pertinent power flows. In addition, the SE function `doSE` in MATPOWER is used to realize the WLS iterations based on the Gauss-Newton method for the sake of comparison. The iterations terminate either upon convergence, or, once the condition number of the approximate linearization exceeds 10^4, which flags explosion of the iterates. A total of 39 power meters are installed in the system to ensure observability, including 5 (correspondingly 2) for real (reactive) bus injections, 9 (respectively 12) for real (reactive) line flows, as well as 6 measuring bus voltage magnitudes. Independent Gaussian noise corrupts the measurements, with relative noise power $16:9:4$ for these three types of meters.

To solve the dual SDP problem (21), the MATLAB-based optimization modeling tool YALMIP [8], together with its interior-point method solver SeDuMi [12], is utilized. The primal SDP problem (18) is formulated and autromatically converted to the dual SDP one using the function `dualize` in YALMIP. Once the dual problem (21) is solved, the optimal $\hat{\mathbf{X}}$ is found as the Lagrange multiplier associated with the dual constraint (21c), and the corresponding estimator $\hat{\mathbf{v}}$ is recovered using the eigenvector as detailed in Remark 2. In addition, to fix the phase angle ambiguity, bus 1 is chosen as the reference bus [14, Ch. 12]. Thus, its voltage needs to be set as a real variable by imposing an additional linear equality constraint on $\mathbf{X}$ to have its $(15, 15)$-th entry equal to 0 in the primal problem (18).

Two thousand Monte Carlo realizations are averaged to



(a)



(b)

Fig. 3. Comparisons between the WLS and SDP approaches with various meter noise power for the (a) ℓ_2 estimation error $\|\mathbf{v}-\hat{\mathbf{v}}\|_2$; and (b) ℓ_2 measurement error $\|\mathbf{z}-\hat{\mathbf{z}}\|_2$.

obtain empirical errors of the state estimators $\hat{\mathbf{v}}$, and the corresponding meter measurements $\hat{\mathbf{z}}$ with ℓ-th entry $\hat{z}_\ell = h_\ell(\hat{\mathbf{v}})$. Each bus' voltage magnitude is Gaussian with mean 1 and variance 0.01, while its angle is uniform over $[-0.3\pi,\ 0.3\pi]$. Fig. 3 speaks for the merits of the proposed SDP scheme in both empirical error metrics. Without favorable initialization, the Gauss-Newton iterative method vitally suffers from the potential convergence issue, clearly revealed by its large gaps from the SDP scheme, in both error metrics.

VI. Conclusions and Current Research

A novel SDP-based SE scheme was developed in this paper for power system monitoring, by tactfully reformulating the nonlinear relationship between power meter measurements and complex bus voltages. The nonconvex SE problem was converted to a convex one using semidefinite relaxation, and thus rendered efficiently solvable via existing SDP routines. Interestingly, the convex formulation is also amenable to a consensus-based distributed implementation. In addition, the convex dual SE problem further reduces the number of optimization variables, and is important for investigating sufficient conditions for it to achieve the global optimum of the original nonconvex SE problem. (Results on this subject will be published elsewhere.)

The proposed SE algorithm is currently evaluated on more complicated power testing systems, and further enhancements to its performance are pursued toward recovering the solution of the primal SE problem using the randomization technique. At the same time, the novel SDP formulation for *nonlinear* AC power systems is investigated in a cyber-security context, where approaches to analyze and account for fault data injection have so far been confined to the *linear* DC power flow model [7].

References

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

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

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

[4] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge, UK: Cambridge University Press, 2004.

[5] 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.

[6] J. Lavaei and S. H. Low, "Convexification of optimal power flow problem," in *Proc. of 48th Annual Allerton Conf.*, UIUC, IL., Sep. 29-Oct. 1, 2010.

[7] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against State estimation in electric power grids," in *Proc. of the 16th ACM Conf. on Computer and Comm. Security*, Chicago, IL, Nov. 9–13, 2009.

[8] J. Löfberg, "Dualize it: software for automatic primal and dual conversions of conic programs," *Optimization Methods & Software*, vol. 24, no. 3, pp. 313–325, Jun. 2009.

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

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

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

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

[13] U.S. Department of Energy, *The Smart Grid: An introduction*, 2008. [Online.] Available: http://www.oe.energy.gov/SmartGridIntroduction.htm

[14] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*. New York, NY: John Wiley & Sons, 1984.

[15] W. A. Wulf, "Great achievements and grand challenges," *The Bridge*, vol. 30, no. 3/4, pp. 5–10, Fall 2010. [Online.] Available: http://www.greatachievements.org/

[16] 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. on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[17] H. Zhu, A. Cano and G. B. Giannakis, "Consensus-based distributed MIMO decoding using semidefinite relaxation," in *Proc. of 2nd IEEE Workshop on Computational Advances in Multi-Channel Sensor Array Processing (CAMSAP)*, St. Thomas, Virgin Islands, Dec. 12–14, 2007.

BIOGRAPHIES

Hao Zhu (S'07) received her Bachelor degree from Tsinghua University, Beijing, China in 2006 and the M.Sc. degree from the University of Minnesota, Minneapolis in 2009, both in electrical engineering.

From 2005 to 2006, she was working as a research assistant with the Intelligent Transportation information Systems (ITiS) Laboratory, Tsinghua University. Since September 2006, she has been working towards her Ph.D. degree with the Department of Electrical and Computer Engineering, University of Minnesota. Her current research interests include sparsity-aware statistical signal processing, machine learning, and power system monitoring and optimization.

Georgios B. Giannakis (Fellow'97) received his Diploma in Electrical Engr. from the Ntl. Tech. Univ. of Athens, Greece, 1981. From 1982 to 1986 he was with the Univ. of Southern California (USC), where he received his MSc. in Electrical Engineering, 1983, MSc. in Mathematics, 1986, and Ph.D. in Electrical Engr., 1986. Since 1999 he has been a professor with the Univ. of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications in the ECE 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 300 journal papers, 500 conference papers, two edited books and two research monographs. Current research focuses on compressive sensing, cognitive radios, network coding, cross-layer designs, mobile ad hoc networks, the smart grid, wireless sensor and social networks. He is the (co-) inventor of twenty patents issued, and the (co-) recipient of seven 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.