Power System State Estimation and Bad Data Detection by Means of Conic Relaxation

Ramtin Madani[1], Javad Lavaei[2], Ross Baldick[3], Alper Atamtürk[2]
[1]University of Texas at Arlington [2]University of California, Berkeley [3]University of Texas at Austin
ramtin.madani@uta.edu lavaei@berkeley.edu baldick@ece.utexas.edu atamturk@berkeley.edu

Abstract

This paper is concerned with the power system state estimation problem, which aims to find the unknown operating point of a power network based on a set of available measurements. We design a penalized semidefinite programming (SDP) relaxation whose objective function consists of a surrogate for rank and an ℓ_1-norm penalty accounting for noise. Although the proposed method does not rely on initialization, its performance can be improved in presence of an initial guess for the solution. First, a sufficient condition is derived with respect to the closeness of the initial guess to the true solution to guarantee the success of the penalized SDP relaxation in the noiseless case. Second, we show that a limited number of incorrect measurements with arbitrary values have no effect on the recovery of the true solution. Furthermore, we develop a bound for the accuracy of the estimation in the case where a limited number of measurements are corrupted with arbitrarily large values and the remaining measurements are perturbed with modest noise values. The proposed technique is demonstrated on a large-scale 1354-bus European system.

1. Introduction

Uncertainties and complexities imposed by renewable energy sources and emerging technologies for future grids, such as electric vehicles, necessitate the development of advanced system monitoring mechanisms for the control and operation of power networks. In this regard, state estimation plays a crucial role. Power system state estimation (PSSE) is the problem of determining the operating point of an electrical network based on the given model and the measurements obtained from supervisory control and data acquisition (SCADA) systems [1], [2].

Due to difficulties such as the nonlinearity induced by the laws of physics, PSSE is known to be a non-convex and NP-hard problem in the worst case. Aside from non-convexity, there are further challenges involved in obtaining an accurate estimation for the state of a power network, including the presence of noise and possibly incorrect sensor information. It is essential for power system state estimation algorithms to be robust to a limited number of severely corrupted measurements, known as bad data. Once corrupted measurements are identified, they can be removed from the set of measurements in order to avoid inaccurate estimation. Therefore, several methods have been proposed in the literature to address the problem of bad data detection through different approaches such as statistical techniques, convex relaxation and sensor placement [3]–[7].

In this paper, we employ a semidefinite programming (SDP) relaxation technique in order to tackle the non-convexity of the feasible region described by the AC power flow equations. The papers [8] and [9] have sparked intensive studies of the SDP relaxation technique for solving fundamental problems in power networks. The work [9] develops an SDP relaxation for finding a global minimum of the optimal power flow (OPF) problem. A sufficient and necessary condition is provided to guarantee a zero duality gap, which is satisfied by several benchmark systems. From the perspective of the physics of power systems, the follow-up papers [10] and [11] develop theoretical results to support the success of the SDP relaxation in handling the non-convexity of OPF. The papers [12] and [13] develop a graph-theoretic SDP framework for finding a near-global solution whenever the SDP relaxation fails to find a global minimum. Recently, SDP relaxation has been applied to the PSSE problem and its superiority over traditional approaches has been demonstrated through extensive numerical case studies [14], [15]. The papers [16] and [17] have performed a graph decomposition in order to replace the large-scale SDP matrix variable with smaller sub-matrices, based on which different distributed numerical algorithms have been developed. Moreover, the formulations in [14] and [15] have been extended in [16] and [5] to accommo-

URI: http://hdl.handle.net/10125/41532
ISBN: 978-0-9981331-0-2

date PMU measurements. The work [5] has studied a variety of regularization methods to solve the PSSE problem in the presence of bad data and topology error. These methods include weighted least square (WLS) and weighted least absolute value (WLAV) penalty functions, together with a nuclear norm surrogate for obtaining a low-rank solution.

In this paper, we develop theoretical guarantees for the effectiveness of an SDP-based convex formulation equipped with an ℓ_1-norm estimator for solving the PSSE problem. We design a class of penalized convex problems, where the measurement equations are softly penalized in the objective as opposed to being imposed as equality constraints. The objective function of the convex problem has two terms: (i) one is a linear function that can be designed based on an initial guess for the solution to deal with non-convexity, (ii) and the other term is an ℓ_1-norm penalty that is intended to estimate noise values. We show that the penalized SDP relaxation precisely solves the PSSE problem in the case of noiseless measurements if the available initial guess is relatively close to the solution to be sought. We then consider the scenario where a limited number of measurements are corrupted with possibly large noise values (bad data) and the remaining measurements are noiseless. In this case, we prove that if the number of corrupted measurements is small, they have absolutely no effect on the solution of the penalized convex problem. Finally, we derive a bound on the estimation error in the case where the measurements are subject to perturbations with modest noise values. We demonstrate the efficacy of the proposed mathematical framework through extensive simulations.

1.1. Notations

The symbols $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{C}$ denote the sets of real, nonnegative real and complex numbers, respectively. $\mathbb{S}^n$ denotes the space of $n \times n$ real symmetric matrices and $\mathbb{H}^n$ denotes the space of $n \times n$ complex Hermitian matrices. $\mathrm{Re}\{\cdot\}$, $\mathrm{Im}\{\cdot\}$, $\mathrm{rank}\{\cdot\}$, $\mathrm{trace}\{\cdot\}$, $\det\{\cdot\}$ and $\mathrm{null}\{\cdot\}$ denote the real part, imaginary part, rank, trace, determinant and null space of a given scalar/matrix. $\mathrm{diag}\{\cdot\}$ denotes the vector of diagonal entries of a matrix. For every $p \geq 1$, the symbol $\|\cdot\|_p$ denotes the p-norm of a matrix. Likewise, the notation $\|\cdot\|_F$ denotes the Frobenius norm. Matrices are shown by capital and bold letters. The symbols $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$ and $(\cdot)^{\mathrm{conj}}$ denote transpose, conjugate transpose and conjugate respectively. Furthermore, "$\mathbb{i}$" is reserved to denote the imaginary unit. The notation $\langle \mathbf{A}, \mathbf{B} \rangle$ represents $\mathrm{trace}\{\mathbf{A}^*\mathbf{B}\}$, which is the inner product of the matrices $\mathbf{A}$ and $\mathbf{B}$. The notations $\measuredangle x$ and $|x|$ denote the angle and magnitude of a complex number x. The notation $\mathbf{W} \succeq 0$ means that $\mathbf{W}$ is a Hermitian and positive semidefinite matrix. Likewise, $\mathbf{W} \succ 0$ means that $\mathbf{W}$ is Hermitian and positive definite. Given a matrix $\mathbf{W}$, its Moore Penrose pseudoinverse is denoted as $\mathbf{W}^+$. The (i, j) entry of $\mathbf{W}$ is denoted as W_{ij}. The symbol $\mathbf{0}_n$ and $\mathbf{1}_n$ denote the $n \times 1$ vectors of zeros and ones, respectively. $\mathbf{0}_{m\times n}$ denotes the $m \times n$ zero matrix and $\mathbf{I}_{n\times n}$ is the $n \times n$ identity matrix. The notation $|\mathcal{X}|$ denotes the cardinality of a set $\mathcal{X}$. For an $m \times n$ matrix $\mathbf{W}$, the notation $\mathbf{W}[\mathcal{X}, \mathcal{Y}]$ denotes the submatrix of $\mathbf{W}$ whose rows and columns are chosen form $\mathcal{X}$ and $\mathcal{Y}$, respectively, for given index sets $\mathcal{X} \subseteq \{1, \ldots, m\}$ and $\mathcal{Y} \subseteq \{1, \ldots, n\}$. Similarly, $\mathbf{x}[\mathcal{X}]$ denotes the vector whose entries are chosen from $\mathbf{x}$ based on the index set $\mathcal{X}$.

2. Preliminaries

In this section, we offer some preliminary results on the power system state estimation problem.

2.1. Voltages, Currents, and Admittance Matrices

Let $\mathcal{N}$ and $\mathcal{L}$ denote the sets of buses (nodes) and branches (edges) of the power network under study. Moreover, let n denote the number of buses. Define $\mathbf{v} \triangleq [v_1, v_2, \ldots, v_n]^{\mathrm{T}}$ to be the vector of complex voltages, where $v_k \in \mathbb{C}$ is the complex (phasor) voltage at node $k \in \mathcal{N}$ of the power network. Denote the magnitude and phase of v_k as $|v_k|$ and $\measuredangle v_k$, respectively. Let $i_k \in \mathbb{C}$ denote the net injected complex current at bus $k \in \mathcal{N}$. Given an edge $l \in \mathcal{L}$, there are two current signals entering the transmission line from its both ends. We orient the lines of the network arbitrarily and define $i_{\mathrm{f};l} \in \mathbb{C}$ and $i_{\mathrm{t};l} \in \mathbb{C}$ to be the complex currents entering the branch $l \in \mathcal{L}$ through its *from* and *to* (tail and head) ends, respectively, according to the designated orientation.

Define $\mathbf{Y} \in \mathbb{C}^{n\times n}$ as the admittance matrix of the network, and $\mathbf{Y}_{\mathrm{f}} \in \mathbb{C}^{|\mathcal{L}|\times n}$ and $\mathbf{Y}_{\mathrm{t}} \in \mathbb{C}^{|\mathcal{L}|\times n}$ as the *from* and *to* branch admittance matrices, respectively, such that

$$\mathbf{i} = \mathbf{Y} \times \mathbf{v}, \quad \mathbf{i}_{\mathrm{f}} = \mathbf{Y}_{\mathrm{f}} \times \mathbf{v}, \quad \mathbf{i}_{\mathrm{t}} = \mathbf{Y}_{\mathrm{t}} \times \mathbf{v}, \qquad (1)$$

where $\mathbf{i} \triangleq [i_1, i_2 \ldots, i_n]^{\mathrm{T}}$ is the vector of complex nodal current injections, and $\mathbf{i}_{\mathrm{f}} \triangleq [i_{\mathrm{f},1}, i_{\mathrm{f},2} \ldots, i_{\mathrm{f},|\mathcal{L}|}]^{\mathrm{T}}$and $\mathbf{i}_{\mathrm{t}} \triangleq [i_{\mathrm{t},2}, i_{\mathrm{t},2} \ldots, i_{\mathrm{t},|\mathcal{L}|}]^{\mathrm{T}}$ are the vectors of currents entering the *from* and *to* ends of branches, respectively.

2.2. Power Flow Equations

Let p_k and q_k represent the net active and reactive power injections at every bus $k \in \mathcal{N}$, where $\mathbf{p} \triangleq [p_1\ p_2\ \cdots\ p_n]^{\mathrm{T}} \in \mathbb{R}^n$ and $\mathbf{q} \triangleq [q_1\ q_2\ \cdots\ q_n]^{\mathrm{T}} \in \mathbb{R}^n$ are the vectors containing net injected active and reactive powers, respectively. The power balance equations can be expressed as

$$\mathbf{p} + \mathbb{i}\mathbf{q} = \mathrm{diag}\{\mathbf{v} \times \mathbf{i}^*\}. \tag{2}$$

For every $k \in \mathcal{N}$, define

$$\mathbf{E}_k \triangleq \mathbf{e}_k\mathbf{e}_k^*, \tag{3a}$$

$$\mathbf{Y}_{\mathrm{i};k} \triangleq \mathbf{Y}^*\mathbf{e}_k\mathbf{e}_k^*\mathbf{Y} \tag{3b}$$

$$\mathbf{Y}_{\mathrm{p};k} \triangleq \frac{1}{2}(\mathbf{Y}^*\mathbf{e}_k\mathbf{e}_k^* + \mathbf{e}_k\mathbf{e}_k^*\mathbf{Y}), \tag{3c}$$

$$\mathbf{Y}_{\mathrm{q};k} \triangleq \frac{1}{2\mathbb{i}}(\mathbf{Y}^*\mathbf{e}_k\mathbf{e}_k^* - \mathbf{e}_k\mathbf{e}_k^*\mathbf{Y}), \tag{3d}$$

where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ denote the standard basis vectors in $\mathbb{R}^n$. The nodal parameters $|v_k|^2$, $|i_k|^2$, p_k and q_k can be expressed as the Frobenius inner-product of $\mathbf{v}\mathbf{v}^*$ with the matrices $\mathbf{E}_k$, $\mathbf{Y}_{\mathrm{i};k}$, $\mathbf{Y}_{\mathrm{p};k}$ and $\mathbf{Y}_{\mathrm{q};k}$:

$$|v_k|^2 = \langle \mathbf{v}\mathbf{v}^*, \mathbf{E}_k \rangle, \qquad \forall k \in \mathcal{N} \tag{4a}$$

$$|i_k|^2 = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{i};k} \rangle, \qquad \forall k \in \mathcal{N} \tag{4b}$$

$$p_k = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{p};k} \rangle, \qquad \forall k \in \mathcal{N} \tag{4c}$$

$$q_k = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{q};k} \rangle, \qquad \forall k \in \mathcal{N}. \tag{4d}$$

There are two power flows entering the transmission line from its both ends. Given a line $l \in \mathcal{L}$ from node k to node j, define $s_{\mathrm{f};l} \triangleq p_{\mathrm{f};l} + q_{\mathrm{f};l}\mathbb{i}$ and $s_{\mathrm{t};l} \triangleq p_{\mathrm{t};l} + q_{\mathrm{t};l}\mathbb{i}$ to be the complex power flows entering the branch $l \in \mathcal{L}$ through buses k and j, respectively. One can write:

$$s_{\mathrm{f};l} = v_k \times i_{\mathrm{f};l}^*\ , \qquad s_{\mathrm{t};l} = v_j \times i_{\mathrm{t};l}^*\ . \tag{5}$$

Let $\mathbf{d}_1, \ldots, \mathbf{d}_{|\mathcal{L}|}$ denote the standard basis vectors in $\mathbb{R}^{|\mathcal{L}|}$. Given a line $l \in \mathcal{L}$ from node k to node j, define

$$\mathbf{Y}_{\mathrm{i_f};l} \triangleq \mathbf{Y}_{\mathrm{f}}^*\mathbf{d}_l\mathbf{d}_l^*\mathbf{Y}_{\mathrm{f}}, \tag{6a}$$

$$\mathbf{Y}_{\mathrm{p_f};l} \triangleq \frac{1}{2}(\mathbf{Y}_{\mathrm{f}}^*\mathbf{d}_l\mathbf{e}_k^* + \mathbf{e}_k\mathbf{d}_l^*\mathbf{Y}_{\mathrm{f}}), \tag{6b}$$

$$\mathbf{Y}_{\mathrm{q_f};l} \triangleq \frac{1}{2\mathbb{i}}(\mathbf{Y}_{\mathrm{f}}^*\mathbf{d}_l\mathbf{e}_k^* - \mathbf{e}_k\mathbf{d}_l^*\mathbf{Y}_{\mathrm{f}}), \tag{6c}$$

$$\mathbf{Y}_{\mathrm{i_t};l} \triangleq \mathbf{Y}_{\mathrm{t}}^*\mathbf{d}_l\mathbf{d}_l^*\mathbf{Y}_{\mathrm{t}}, \tag{6d}$$

$$\mathbf{Y}_{\mathrm{p_t};l} \triangleq \frac{1}{2}(\mathbf{Y}_{\mathrm{t}}^*\mathbf{d}_l\mathbf{e}_j^* + \mathbf{e}_j\mathbf{d}_l^*\mathbf{Y}_{\mathrm{t}}), \tag{6e}$$

$$\mathbf{Y}_{\mathrm{q_t};l} \triangleq \frac{1}{2\mathbb{i}}(\mathbf{Y}_{\mathrm{t}}^*\mathbf{d}_l\mathbf{e}_j^* - \mathbf{e}_j\mathbf{d}_l^*\mathbf{Y}_{\mathrm{t}}). \tag{6f}$$

The branch parameters $i_{\mathrm{f};l}$, $p_{\mathrm{f};l}$, $q_{\mathrm{f};l}$, $i_{\mathrm{t};l}$, $p_{\mathrm{t};l}$ and $q_{\mathrm{t};l}$ can be written as the inner product of $\mathbf{v}\mathbf{v}^*$ with the matrices $\mathbf{Y}_{\mathrm{i_f};l}$, $\mathbf{Y}_{\mathrm{p_f};l}$, $\mathbf{Y}_{\mathrm{q_f};l}$, $\mathbf{Y}_{\mathrm{i_t};l}$, $\mathbf{Y}_{\mathrm{p_t};l}$ and $\mathbf{Y}_{\mathrm{q_t};l}$:

$$i_{\mathrm{f};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{i_f};l} \rangle, \quad i_{\mathrm{t};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{i_t};l} \rangle, \quad \forall l \in \mathcal{L} \tag{7a}$$

$$p_{\mathrm{f};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{p_f};l} \rangle, \ p_{\mathrm{t};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{p_t};l} \rangle, \ \forall l \in \mathcal{L} \tag{7b}$$

$$q_{\mathrm{f};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{q_f};l} \rangle, \ q_{\mathrm{t};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{q_t};l} \rangle, \ \forall l \in \mathcal{L} \tag{7c}$$

Equations (4) and (7) offer a compact formulation for common measurements in power networks. In what follows, we will study a general version of the state estimation problem with arbitrary measurements of quadratic forms. Consider the state estimation problem of finding a solution to the quadratic equations

$$x_r = \langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_r \rangle + \omega_r + \eta_r, \qquad r \in \mathcal{M}, \tag{8}$$

where

- $\mathcal{M} = \{1, 2, \ldots, m\}$ is the set of indices associated with the available measurements (or specifications).
- $x_1, \ldots, x_m$ are the known measurement values.
- $\omega_1, \ldots, \omega_m$ are the unknown noise values, for which some *a priori* statistical information may be available (e.g., zero mean Gaussian noise).
- $\eta_1, \ldots, \eta_m$ are unknown and sparsely occurring noise values with arbitrary magnitudes (possibly large), which are expected to impact a small subset of measurements.
- $\mathbf{M}_1, \ldots, \mathbf{M}_m$ are some known $n \times n$ Hermitian matrices (e.g., they could be any subset of the matrices defined in (3) and (6)). With no loss of generality, we assume that

  $$\|\mathbf{M}_r\|_2 \leq 1, \qquad \forall r \in \mathcal{M}. \tag{9}$$

  (this can be achieved by rescaling the measurement equations).

The scalars $\omega_1, \omega_2, \ldots, \omega_m$ are expected to represent modest noise values that may not affect the accuracy of estimation significantly. However, the values of $\eta_1, \eta_2, \ldots, \eta_m$ could purposefully be large (known as bad data), in which case the corresponding equations will not be informative. Therefore, it is of practical interest to separately analyze the effect of measurements that are entirely corrupted (i.e., bad data).

Definition 1: Let $(\mathcal{G}, \mathcal{B})$ denote the partition of $\mathcal{M}$ into the two sets of measurements with zero and nonzero sparse noise values, i.e.,

$$\mathcal{G} = \{r \in \mathcal{M} \,|\, \eta_r = 0\}, \qquad \mathcal{B} = \mathcal{M} \setminus \mathcal{G}. \tag{10}$$

$\mathcal{G}$ and $\mathcal{B}$ are called the sets of *good* and *bad* measurements, respectively.

In the case where the noises $\omega_1, \ldots, \omega_m$ are all equal to zero and $\mathcal{B} = \varnothing$, the problem (8) reduces to the well-known power flow problem. It is straightforward to verify that if $\mathbf{v}$ is a solution to the state estimation problem, then $\alpha\mathbf{v}$ is another solution of this problem for every complex number α with magnitude 1. To resolve the existence of infinitely many solutions due to phase ambiguity, we assume that $\measuredangle v_k$ is equal to zero at a pre-selected bus, named *reference* bus. Hence, in the noiseless case, the state estimation problem with the complex variable $\mathbf{v}$ amounts to $2n-1$ real variables.

2.3. Semidefinite Programming Relaxation

The state estimation problem, as a general case of the power flow problem, is nonconvex due to the quadratic matrix $\mathbf{v}\mathbf{v}^*$. Hence, it is desirable to convexify the problem. By replacing $\mathbf{v}\mathbf{v}^*$ with a new matrix variable $\mathbf{W}$, the quadratic equations in (8) can be formulated linearly in terms of $\mathbf{W}$ as follows:

$$x_r = \langle \mathbf{W}, \mathbf{M}_r \rangle + \omega_r + \eta_r, \qquad \forall r \in \mathcal{M}. \tag{11}$$

Consider the case where the quadratic measurements $x_1, \ldots, x_m$ are noiseless. Solving the non-convex equations in (8) is tantamount to finding a rank-1 and positive semidefinite matrix $\mathbf{W} \in \mathbb{H}^n$ satisfying the above linear equations for $\boldsymbol{\omega} = \boldsymbol{\eta} = 0$ (because such a matrix $\mathbf{W}$ could then be decomposed as $\mathbf{v}\mathbf{v}^*$). The problem of finding a positive semidefinite matrix $\mathbf{W} \in \mathbb{H}^n$ satisfying the linear equations in (11) is regarded as a convex *relaxation* of (8) since it includes no restriction on the rank of $\mathbf{W}$. Although the set of equations (8) normally has a finite number of solutions whenever $m \geq 2n-1$, its SDP relaxation (11) may have infinitely many solutions because the matrix variable $\mathbf{W}$ includes $O(n^2)$ scalar variables as opposed to $2n-1$. The literature of compressed sensing substantiates that minimizing $\mathrm{trace}\{\mathbf{W}\}$ over the feasible set of (11) may yield a rank-1 matrix $\mathbf{W}$ under strong technical assumptions [18]–[20]. However, this approach is not applicable to power system problems since diagonal values of the matrix $\mathbf{W}$ represent voltage magnitudes, which may be fixed or tightly constrained in practice. The main objective of this paper is to develop an alternative approach for power system applications along with theoretical guarantees on the performance of a convexification framework for the PSSE problem in terms of: (i) the closeness of the available initial guess to the true unknown state of the system, (ii) the number of bad measurements, and (iii) the noise level of good measurements.

2.4. Sensitivity Analysis

Let $\mathcal{O}$ denote the set of all buses of the network except the reference bus. The operating point of the power system can be characterized in terms of the real-valued vector

$$\overline{\mathbf{v}} \triangleq \left[\mathrm{Re}\{\mathbf{v}[\mathcal{N}]^{\mathrm{T}}\} \quad \mathrm{Im}\{\mathbf{v}[\mathcal{O}]^{\mathrm{T}}\} \right]^{\mathrm{T}} \in \mathbb{R}^{2n-1}. \tag{12}$$

For every $n \times n$ Hermitian matrix $\mathbf{X}$, let $\overline{\mathbf{X}}$ denote the following $(2n-1)\times(2n-1)$ real-valued and symmetric matrix:

$$\overline{\mathbf{X}} = \begin{bmatrix} \mathrm{Re}\{\mathbf{X}[\mathcal{N},\mathcal{N}]\} & -\mathrm{Im}\{\mathbf{X}[\mathcal{N},\mathcal{O}]\} \\ \mathrm{Im}\{\mathbf{X}[\mathcal{O},\mathcal{N}]\} & \mathrm{Re}\{\mathbf{X}[\mathcal{O},\mathcal{O}]\} \end{bmatrix}. \tag{13}$$

Definition 2: Given an arbitrary subset of measurements $\mathcal{M}' \subseteq \mathcal{M}$ with

$$\alpha_1 < \alpha_2 < \ldots < \alpha_{|\mathcal{M}'|}$$

as the indices of such measurements, define the function $\boldsymbol{f}_{\mathcal{M}'}(\overline{\mathbf{v}}) : \mathbb{R}^{2n-1} \to \mathbb{R}^{|\mathcal{M}'|}$ as the mapping from the real-valued state of the power network (i.e., $\overline{\mathbf{v}}$) to the vector of true (noiseless) measurement values:

$$\boldsymbol{f}_{\mathcal{M}'}(\overline{\mathbf{v}}) \triangleq [\langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_{\alpha_1} \rangle, \ldots, \langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_{\alpha_{|\mathcal{M}|}} \rangle]^{\mathrm{T}}.$$

Define also $\mathbf{J}_{\mathcal{M}'}(\mathbf{z}) \in \mathbb{R}^{(2n-1)\times|\mathcal{M}'|}$ to be the Jacobian of $\boldsymbol{f}_{\mathcal{M}'}$ at the point $\mathbf{z} \in \mathbb{R}^{2n-1}$, i.e.,

$$\mathbf{J}_{\mathcal{M}'}(\mathbf{z}) = 2 \left[\overline{\mathbf{M}}_{\alpha_1} \mathbf{z} \quad \overline{\mathbf{M}}_{\alpha_2} \mathbf{z} \quad \ldots \quad \overline{\mathbf{M}}_{\alpha_{|\mathcal{M}'|}} \mathbf{z} \right]$$

(note that $\overline{\mathbf{M}}_r$ can be found using $\mathbf{M}_r$ via the equation (13) for $r = 1, \ldots, m$).

Notation 1: Given a Hermitian matrix $\mathbf{A}$, denote $\kappa(\mathbf{A})$ as the sum of the two smallest eigenvalues of $\mathbf{A}$.

According to the inverse function theorem, if $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})$ has full row rank, then the inverse of the function $\boldsymbol{f}(\overline{\mathbf{v}})$ exists in a neighborhood of the point $\overline{\mathbf{v}}$. Similarly, it follows from the Kantorovich Theorem that, under the previous assumption, the equation (8) can be solved using Newton's method by starting from any initial point sufficiently close to the point $\mathbf{v}$, provided that the measurements are noiseless.

3. Main Results

Suppose that we are given an initial point $\mathbf{u} \in \mathbb{C}^n$, which is relatively close to the unknown vector of voltages to be sought. We propose a penalized convex optimization problem for solving the system of quadratic equations (8).

Definition 3: The following convex optimization is referred to as the *penalized SDP relaxation* problem with the input $(\mathbf{x}, \mu)$:

$$\underset{\substack{\mathbf{W}\in\mathbb{H}^n \\ \boldsymbol{\nu}\in\mathbb{R}^m}}{\text{minimize}} \quad \langle \mathbf{W}, \mathbf{M} \rangle + \mu \times \|\boldsymbol{\nu}\|_1 \tag{14a}$$

$$\text{subject to} \quad \langle \mathbf{W}, \mathbf{M}_r \rangle + \nu_r = x_r, \quad r \in \mathcal{M}, \tag{14b}$$

$$\mathbf{W} \succeq 0. \tag{14c}$$

where $\mathbf{M}$ is a constant $n \times n$ Hermitian matrix that satisfies the following two properties:

$$\mathbf{M} \times \mathbf{u} = \mathbf{0}_n \tag{15a}$$

$$\kappa(\mathbf{M}) > 0 \tag{15b}$$

Note that the penalized SDP relaxation is not a semidefinite program due to the norm term in its objective, but it can be converted into an SDP by introducing new variables to replace the norm term with linear constraints. The term $\langle \mathbf{W}, \mathbf{M} \rangle$ in the objective of the penalized convex problem (14) is a surrogate for the rank of $\mathbf{W}$ and aims to handle the nonlinearity of measurement equations. In addition, noise values are estimated through the auxiliary variables $\nu_1, \ldots, \nu_m \in \mathbb{R}$ by incorporating the regularization term $\|\boldsymbol{\nu}\|_1$ into the objective function. In this work, we seek to derive certain conditions in order to guarantee that:

1) In the noiseless case, the penalized SDP problem (14) recovers the exact solution if the vector of voltages to be sought is relatively close to the initial guess $\mathbf{u}$.
2) A limited number of wrong or severely corrupted equations (i.e., members of $\mathcal{B}$) have zero effect on the accuracy of estimation, regardless of their corresponding measurement values.
3) The accuracy of estimation is bounded with respect to the level of modest noise values, i.e., $\|\boldsymbol{\omega}\|_1$.

In order to perform an analysis on the effectiveness of the proposed penalized SDP problem in different scenarios, it is necessary to quantify the quality of the initial guess $\mathbf{u}$ and the set of good measurements $\mathcal{G}$ for finding the unknown vector of voltages. To this end, we define functions that measure the closeness of our initial guess to the solution and informativeness of the available measurements.

Definition 4: For an arbitrary subset of measurements $\mathcal{M}' \subseteq \mathcal{M}$ and a real number $\rho \geq 1$, define $\delta_{\mathcal{M}';\rho} : \mathbb{C}^n \to \mathbb{R}$ as follows:

$$\delta_{\mathcal{M}';\rho}(\mathbf{v}) \triangleq \frac{4\|\mathbf{J}^{+}_{\mathcal{M}'}(\overline{\mathbf{v}})\overline{\mathbf{M}}\overline{\mathbf{v}}\|_\rho}{\kappa(\mathbf{M})}. \tag{16}$$

It can be easily observed that if $\mathbf{J}_{\mathcal{M}'}(\mathbf{v})$ is full row rank and $\mathbf{M}$ satisfies the two properties in (15), then $\delta_{\mathcal{M}';\rho}(\mathbf{v})$ serves as a measure for the distance between the two vectors $\|\mathbf{v}\|_2^{-1} \times \mathbf{v}$ and $\|\mathbf{u}\|_2^{-1} \times \mathbf{u}$. The next definition introduces another function that will be later used to quantify the influence of an arbitrary subset of measurements.

Definition 5: For an arbitrary subset of measurements $\mathcal{M}' \subseteq \mathcal{M}$ and a real number $\rho \geq 1$, define $\zeta_{\mathcal{M}';\rho} : \mathbb{C}^n \to \mathbb{R}$ as follows:

$$\zeta_{\mathcal{M}';\rho}(\mathbf{v}) \triangleq \|\mathbf{J}^{+}_{\mathcal{M}'}(\overline{\mathbf{v}})\mathbf{J}_{\mathcal{M}\setminus\mathcal{M}'}(\overline{\mathbf{v}})\|_\rho. \tag{17}$$

3.1. Noiseless Measurements

Our first contribution in this work is for the noiseless case. The following theorem offers conditions under which the penalized SDP problem (14) with a noiseless input is guaranteed to recover the exact vector of voltages.

Theorem 1 (Noiseless case): Suppose that $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})$ has full row rank and

$$\delta_{\mathcal{M};1}(\mathbf{v}) < 1. \tag{18}$$

Then, for every $\mu > \delta_{\mathcal{M};\infty}(\mathbf{v})$, the point $(\mathbf{W}, \boldsymbol{\nu}) = (\mathbf{v}\mathbf{v}^*, \mathbf{0}_m)$ is the unique solution of the penalized SDP relaxation problem (14) with the noiseless input $(\boldsymbol{f}_{\mathcal{M}}(\mathbf{v}), \mu)$.

Proof: The proof is provided in the Appendix. ■

Remark 1: According to Theorem 1, in the noiseless case, if the initial guess $\mathbf{u}$ is sufficiently close to $\mathbf{v}$ and μ is chosen large enough, then the penalized SDP relaxation (14) recovers the exact solution of the problem (8).

3.2. Severely Corrupted Measurements

Theorem 1 offers a sufficient condition to guarantee the exact recovery of $\mathbf{v}$ for cases where noise values are zero and all of the measurements are correct. In this subsection, we extend the result of Theorem 1 to the case where each measurement value is either correct or entirely corrupted. We aim to show that if the unknown vector of complex voltages remains observable through the correct measurements and if the number of corrupted measurements is relatively small, then the unknown state can be recovered precisely.

Theorem 2 (Bad data): Let $\mathcal{G}$ and $\mathcal{B}$ denote the set of good and bad measurements, respectively (see Definition 1). Suppose that $\mathbf{J}_{\mathcal{G}}(\overline{\mathbf{v}})$ has full row rank and the following three conditions are satisfied:

$$\delta_{\mathcal{G};1}(\mathbf{v}) < 1, \tag{19a}$$

$$\zeta_{\mathcal{G};\infty}(\mathbf{v}) < 1, \tag{19b}$$

$$|\mathcal{B}| < \frac{1 - \zeta_{\mathcal{G};\infty}(\mathbf{v})}{1 + \zeta_{\mathcal{G};1}(\mathbf{v})} \times \frac{1 - \delta_{\mathcal{G};1}(\mathbf{v})}{\delta_{\mathcal{G};\infty}(\mathbf{v})}. \tag{19c}$$

Then, for every $\mu \in \mathbb{R}$ such that

$$\frac{\delta_{\mathcal{G};\infty}(\mathbf{v})}{2(1 - \zeta_{\mathcal{G};\infty}(\mathbf{v}))} < \frac{\mu}{\kappa(\mathbf{M})} < \frac{1 - \delta_{\mathcal{G};1}(\mathbf{v})}{2|\mathcal{B}|(1 + \zeta_{\mathcal{G};1}(\mathbf{v}))}, \tag{20}$$

the point $(\mathbf{W}, \boldsymbol{\nu}) = (\mathbf{v}\mathbf{v}^*, \boldsymbol{\eta})$ is the unique solution of the penalized SDP relaxation problem (14) with the input $(\boldsymbol{f}_{\mathcal{M}}(\mathbf{v}) + \boldsymbol{\eta}, \mu)$.

Proof: The proof is provided in the Appendix. ■

3.3. Combination of Modest Noise Values and Bad Data

The guarantee provided by Theorem 2 relies on the number of equations for which the measurement values are corrupted, as opposed to the level of corruption for those equations. It is of particular interest to analyze the case where, in addition to a limited number of arbitrarily corrupted measurements, the other measurements are perturbed by modest noise values $\omega_1, \omega_2, \ldots, \omega_m$. In this subsection, we offer a bound on the estimation error with respect to the level of modest noise values $\|\boldsymbol{\omega}\|_1$.

Theorem 3: Let $(\mathbf{W}^{\text{opt}}, \boldsymbol{\nu}^{\text{opt}})$ be an arbitrary solution of the penalized SDP relaxation problem (14) with the input $(\boldsymbol{f}_{\mathcal{M}}(\mathbf{v}) + \boldsymbol{\omega} + \boldsymbol{\eta}, \mu)$, and $\mathcal{G}$ and $\mathcal{B}$ denote the sets of good and bad measurements, respectively (see Definition 1). Suppose that $\mathbf{J}_{\mathcal{G}}(\overline{\mathbf{v}})$ has full row rank. If the three conditions in (19) hold and μ satisfies (20), then there exists a constant $\alpha > 0$ such that

$$\|\mathbf{W}^{\text{opt}} - \alpha\mathbf{v}\mathbf{v}^*\|_F \leq \sqrt{\tau \times \text{trace}\{\mathbf{W}^{\text{opt}}\} \times \|\boldsymbol{\omega}\|_1} \quad (21)$$

where

$$\tau \triangleq \frac{4\mu \times \kappa(\mathbf{M})^{-1} \times |\mathcal{B}|^{-1}}{(1 - \delta_{\mathcal{G};1}(\mathbf{v})) - 2\mu(\zeta_{\mathcal{G};1}(\mathbf{v}) + 1)}. \quad (22)$$

Proof: The proof is provided in the Appendix. ■

Remark 2: In light of Theorem 3, there exists a constant $\alpha > 0$ for which the estimation error $\|\mathbf{W}^{\text{opt}} - \alpha\mathbf{v}\mathbf{v}^*\|_F$ is upper bounded by

$$\left(\frac{1 - \zeta_{\mathcal{G};\infty}(\mathbf{v})}{1 + \zeta_{\mathcal{G};1}(\mathbf{v})} \times \frac{1 - \delta_{\mathcal{G};1}(\mathbf{v})}{\delta_{\mathcal{G};\infty}(\mathbf{v})} - |\mathcal{B}|\right)^{-1} \times \text{trace}\{\mathbf{W}^{\text{opt}}\} \times \|\boldsymbol{\omega}\|_1$$

if an optimal coefficient μ is used for the penalized SDP problem. This upper bound is non-trivial as long as the inequality (19c) holds.

4. Simulation Results

In this section, numerical results are presented to verify the performance of the proposed penalized convex relaxation technique for the PSSE problem. For all simulations, it is assumed that the measurement set consists of the voltage magnitudes at all buses, the nodal active and reactive power injections at all buses, and the active and reactive power flows in one direction for every line of the network (i.e., entries of $|\mathbf{v}|$, $\mathbf{p}$, $\mathbf{q}$, $\mathbf{p}_\text{f}$ and $\mathbf{q}_\text{f}$).

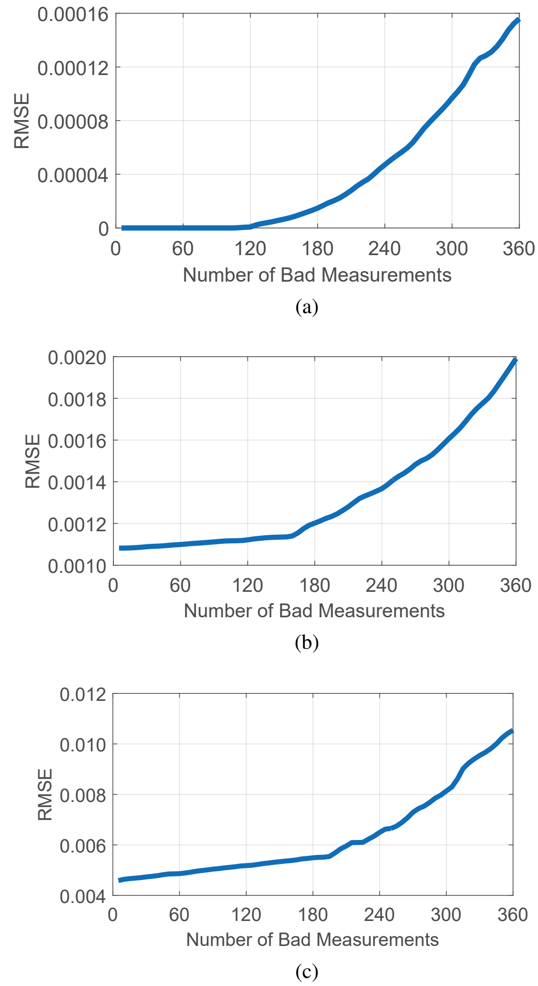

Fig. 1: These figures show the RMSE of the estimated complex voltages for the PEGASE 1354-bus system, using the proposed penalized convex problem, with different numbers of bad measurements. The standard deviation of Gaussian noise values for all measurements is set to (a): 0%, (b): 1%, and (c): 5% of the true values.

4.1. Exploiting Sparsity

The penalized convex problem (14) can be computationally expensive for large-scale systems because of the high-order conic constraints (14c). One method for tackling this issue is to replace the single conic constraint with several lower-order conic constraint as follows:

$$\mathbf{W}[\mathcal{C}_1, \mathcal{C}_1] \succeq 0, \ \ldots, \ \mathbf{W}[\mathcal{C}_d, \mathcal{C}_d] \succeq 0, \quad (23)$$

where $\mathbf{W}[\mathcal{C}_1, \mathcal{C}_1], \mathbf{W}[\mathcal{C}_2, \mathcal{C}_2], \ldots, \mathbf{W}[\mathcal{C}_d, \mathcal{C}_d]$ are principal submatrices of $\mathbf{W}$ with rows and columns chosen from $\mathcal{C}_1, \mathcal{C}_1, \ldots, \mathcal{C}_d \subseteq \mathcal{N}$, respectively. $\mathcal{C}_1, \mathcal{C}_1, \ldots, \mathcal{C}_d$ are some possibly overlapping subsets of $\mathcal{N}$ that can be found through a graph-theoretic analysis of the network graph, named tree decomposition. This procedure breaks down the large-scale conic constraint into several smaller ones. Due to the sparsity and near planarity of power networks, the decomposed penalized convex problem is expected to be significantly lower dimensional. This is due to the fact that all entries of $\mathbf{W}$ that do not appear in any of the above principal submatrices could be removed from the optimization problem. These entries of $\mathbf{W}$, referred to as missing entries, can later be found through a matrix completion algorithm, which enables a rank-1 decomposition of $\mathbf{W}$ for recovering a vector of voltages [13], [20].

4.2. Rank-One Approximation Algorithm

Given an optimal solution $(\mathbf{W}^{\text{opt}}[\mathcal{C}_1, \mathcal{C}_1], \ldots, \mathbf{W}^{\text{opt}}[\mathcal{C}_d, \mathcal{C}_d])$ of the decomposed penalized convex problem, we obtain an approximate solution $\tilde{\mathbf{v}}$ of the set of equations (8) as follows:

1) Set the voltage magnitude $|\tilde{v}_k| := \sqrt{W^{\text{opt}}_{kk}}$ for $k = 1, \ldots, n$.
2) Find the phases of the entries of $\tilde{\mathbf{v}}$ by solving the convex program:

$$\underset{\boldsymbol{\theta} \in [-\pi,\pi]^n}{\text{minimize}} \quad \sum_{(i,j)\in\mathcal{L}} \left| \measuredangle W^{\text{opt}}_{ij} - \theta_i + \theta_j \right| \tag{24a}$$

$$\text{subject to} \quad \theta_o = 0, \tag{24b}$$

where $o \in \mathcal{N}$ is the reference bus.

Note that the above approximation technique is exact in the case where there exists a positive semidefinite filling $\mathbf{W}^{\text{opt}}$ of the known entries such that $\text{rank}\{\mathbf{W}^{\text{opt}}\} = 1$. Under that circumstance, we have $\measuredangle(\mathbf{W}^{\text{opt}})_{ij} - \theta_i + \theta_j = 0$. If there exists a non-rank-one matrix $\mathbf{W}^{\text{opt}}$ with a dominant nonzero eigenvalue, then the above recovery method aims to find a vector $\tilde{\mathbf{v}}$ for which the corresponding line angle differences are as closely as possible to those proposed by $(\mathbf{W}^{\text{opt}}[\mathcal{C}_1, \mathcal{C}_1], \ldots, \mathbf{W}^{\text{opt}}[\mathcal{C}_d, \mathcal{C}_d])$.

4.3. Case Study: Robustness to Bad Data

We have conducted numerical experiments on the Pan European Grid Advanced Simulation and State Estimation (PEGASE) system with 1354 buses to evaluate the robustness of the penalized convex problem to bad measurements and modest noise values. The data is borrowed from [21]. It is assumed that the measurement set consists of the voltage magnitudes at all buses, the nodal active and reactive power injections at all buses, and the active and reactive power flows in both directions for every line of the network.

Let m_0 denote the number of bad measurements. For different values of m_0, up to 3% of the total number of measurements, we have generated 10 random trials by uniformly choosing m_0 measurements and replacing them with large numbers. We have then used the penalized convex problem with normalized measurement coefficients and set $\mu = 10$. The matrix $\mathbf{M}$ in the objective is chosen as $\alpha \times \mathbf{I}_{n\times n} - \mathbf{B}$, where $\mathbf{B}$ is the susceptance matrix and α is the smallest nonnegative number such that $\alpha \times \mathbf{I}_{n\times n} - \mathbf{B} \succeq 0$. This choice is motivated by the fact that $\langle -\mathbf{B}, \mathbf{W} \rangle$ is equal to the total reactive power generation. Since for the power system under study, the matrix $-\mathbf{B}$ has a few negative eigenvalues, we combined it with the identity matrix in order to satisfy the property (15b). Therefore, no initialization is applied for designing the matrix $\mathbf{M}$.

The average root-mean-square error (RMSE) of the estimated voltages for the 10 trials are shown in Figure 1a. As expected from Theorem 2, we empirically observed that for trials with less than 110 bad measurements, the true vector of voltages can be recovered with no estimation error (there would be an extremely small computation error).

As another simulation, in addition to sparsely occurring high intensity noise, we have added zero mean Gaussian noise values to all measurements. The results are reported in Figures 1b and 1c for noise values with standard deviations equal to 1% and 5% of the true values, respectively. In both figures, one realization of Gaussian noise is considered and each point is obtained by averaging the RMSE values of 10 random choices of bad measurements. Then, in each trial, the set of bad measurements is additively expanded for obtaining the next point in diagram. As substantiated in Theorem 3, as long as the number of bad measurements is smaller than a threshold, the estimation error does not grow rapidly.

5. Conclusions

This paper aims to propose a convex relaxation scheme for the power system state estimation (PSSE) problem, which needs to be robust to noise and bad data. We employ an initial guess for the solution in order to build a family of penalized semidefinite programming problems for solving PSSE. The proposed convex relaxation scheme is guaranteed to succeed in the noiseless case as well as the case where a limited number of measurements are corrupted, provided that the initial guess is relatively close to the solution to be sought. We prove that noise values have no effect on the accuracy of estimation in this case. We then consider the case where all measurements are subject to perturbation with modest noise values, in addition to severe corruptions caused by sparse noise. In the later case, we offer an upper bound on the accuracy of estimation with respect to the level of modest noise perturbations. Extensive simulations are performed on a real-world European system with 1354 buses in order to demonstrate the efficacy of the proposed convex relaxation scheme.

7. Acknowledgments

This work was supported by ONR YIP Award, DARPA Young Faculty Award, NSF CAREER Award 1351279, and NSF EECS Awards 1406894 and 1406865. Alper Atamtürk was supported, in part, by grant FA9550-10-1-0168 from the Office of the Assistant Secretary of Defense for Research and Engineering.

6. References

[1] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.
[2] G. 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 Processing Magazine*, vol. 30, no. 5, pp. 107–128, Sept 2013.
[3] W. W. Kotiuga and M. Vidyasagar, "Bad data rejection properties of weighted least absolute value techniques applied to static state estimation," *IEEE Transactions on Power Apparatus and Systems*, no. 4, pp. 844–853, 1982.
[4] D. Deka, R. Baldick, and S. Vishwanath, "Optimal data attacks on power grids: Leveraging detection & measurement jamming," *arXiv preprint arXiv:1506.04541*, 2015.
[5] Y. Weng, M. D. Ilić, Q. Li, and R. Negi, "Convexification of bad data and topology error detection and identification problems in AC electric power systems," *Generation, Transmission & Distribution, IET*, vol. 9, no. 16, pp. 2760–2767, Nov 2015.
[6] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1608–1615, 2006.
[7] J. Zhu and A. Abur, "Bad data identification when using phasor measurements," in *Power Tech, 2007 IEEE Lausanne*. IEEE, 2007, pp. 1676–1681.
[8] J. Lavaei, "Zero duality gap for classical OPF problem convexifies fundamental nonlinear power problems," in *Proceedings of the American Control Conference*, June 2011, pp. 4566–4573.
[9] J. Lavaei and S. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, Feb. 2012.
[10] S. Sojoudi and J. Lavaei, "Physics of power networks makes hard optimization problems easy to solve," July 2012, pp. 1–8.
[11] S. Sojoudi and J. Lavaei, "Exactness of semidefinite relaxations for nonlinear optimization problems with underlying graph structure," *SIAM Journal on Optimization*, vol. 24, no. 4, pp. 1746–1778, 2014.
[12] R. Madani, S. Sojoudi, and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 199–211, Jan. 2015.
[13] R. Madani, M. Ashraphijuo, and J. Lavaei, "Promises of conic relaxation for contingency-constrained optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1297–1307, March 2016.
[14] H. Zhu and G. Giannakis, "Estimating the state of AC power systems using semidefinite programming," in *North American Power Symposium (NAPS)*, Aug 2011, pp. 1–7.
[15] Y. Weng, Q. Li, R. Negi, and M. Ilić, "Semidefinite programming for power system state estimation," in *IEEE Power and Energy Society General Meeting*, July 2012, pp. 1–8.
[16] H. Zhu and G. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1039–1050, 2014.
[17] Y. Weng, Q. Li, R. Negi, and M. Ilić, "Distributed algorithm for SDP state estimation," in *Proceedings of the 2013 IEEE PES Innovative Smart Grid Technologies (ISGT)*, 2013, pp. 1–6.
[18] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, Aug. 2010.
[19] E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, Aug. 2013.
[20] R. Madani, G. Fazelnia, S. Sojoudi, and J. Lavaei, "Low-rank solutions of matrix inequalities with applications to polynomial optimization and matrix completion problems," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, 2014.
[21] C. Josz, S. Fliscounakis, J. Maeght, and P. Panciatici, "AC power flow data in MATPOWER and QCQP format: iTesla, RTE snapshots, and PEGASE," Mar. 2016, [Online]. Available: arxiv.org/abs/1603.01533.

7. Appendix

The dual of the penalized SDP problem (14) can be derived as follows:

$$\underset{\boldsymbol{\lambda}\in\mathbb{R}^m}{\text{maximize}} \quad -\mathbf{x}^{\mathrm{T}}\boldsymbol{\lambda} \tag{25a}$$

$$\text{subject to} \quad \mathbf{M} + \sum_{r\in\mathcal{M}} \lambda_r \mathbf{M}_r \succeq 0 \tag{25b}$$

$$\|\boldsymbol{\lambda}\|_\infty \leq \mu. \tag{25c}$$

Proof of Theorem 1: In order to prove that $(\mathbf{v}\mathbf{v}^*, \mathbf{0}_m)$ is the unique optimal point for the primal problem, it suffices to construct a dual feasible point $\hat{\boldsymbol{\lambda}} \in \mathbb{R}^m$ that certifies the optimality of $(\mathbf{v}\mathbf{v}^*, \mathbf{0}_m)$. To this end, consider the following choice for the dual certificate:

$$\hat{\boldsymbol{\lambda}} := -2\mathbf{J}_{\mathcal{M}}^{+}(\overline{\mathbf{v}})\overline{\mathbf{M}}\overline{\mathbf{v}} \tag{26}$$

and define $\hat{\mathbf{H}} \triangleq \mathbf{M} + \sum_{r\in\mathcal{M}} \hat{\lambda}_r \mathbf{M}_r$. Due to the assumption $\mu > \delta_{\mathcal{M};\infty}(\mathbf{v})$, the dual constraint (25c) is satisfied. In addition, according the definition of $\hat{\boldsymbol{\lambda}}$, it can be easily observed that complementary slackness holds, i.e.,

$$\langle \hat{\mathbf{H}}, \mathbf{v}\mathbf{v}^* \rangle = 0. \tag{27}$$

Moreover, due to the concavity of $\kappa(\cdot)$ and according to the assumption (18), we have

$$\begin{aligned}\kappa(\hat{\mathbf{H}}) &\geq \kappa(\mathbf{M}) + \sum_{r\in\mathcal{M}} \kappa(\hat{\lambda}_r \mathbf{M}_r) \\ &\geq \kappa(\mathbf{M}) - \sum_{r\in\mathcal{M}} 2|\hat{\lambda}_r| \|\mathbf{M}_r\|_2 \\ &\geq \kappa(\mathbf{M}) - 2\|\hat{\boldsymbol{\lambda}}\|_1 = \kappa(\mathbf{M})(1-\delta_{\mathcal{M};1}(\mathbf{v})) > 0. \quad (28)\end{aligned}$$

This implies that the constraint (25b) holds and in addition,

$$\text{rank}\{\hat{\mathbf{H}}\} = n-1. \quad (29)$$

As a result, $\hat{\boldsymbol{\lambda}}$ satisfies the requirements to certify that $(\mathbf{v}\mathbf{v}^*, \mathbf{0}_m)$ is the unique solution for problem (14), provided that strong duality holds.

In order to prove the strong duality, we need to construct a strictly feasible point $\tilde{\boldsymbol{\lambda}} \in \mathbb{R}^m$ for the dual problem. Let o represent the reference bus of the power system. With no loss of generality, we assume that $\text{Im}\{v_o\} = 0$. Since $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})$ has full row rank, we have

$$\overline{\mathbf{v}}^{\text{T}} \mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}}) \neq 0.$$

Therefore, the relation

$$\mathbf{v}^* \mathbf{M}_r \mathbf{v} = \overline{\mathbf{v}}^{\text{T}} \overline{\mathbf{M}}_r \overline{\mathbf{v}} \neq 0 \quad (30)$$

holds for at least one index $r \in \mathcal{M}$. Let $\mathbf{d}_1, \ldots, \mathbf{d}_m$ be the standard basis vectors for $\mathbb{R}^m$. We select $\tilde{\boldsymbol{\lambda}}$ as $\hat{\boldsymbol{\lambda}} + c \times \mathbf{d}_r$, where $c \in \mathbb{R}$ is a nonzero number with an arbitrarily small absolute value such that $c \times \mathbf{v}^* \mathbf{M}_r \mathbf{v} > 0$. Then, one can write:

$$\mathbf{M} + \sum_{r\in\mathcal{M}} \tilde{\lambda}_r \mathbf{M}_r = \hat{\mathbf{H}} + c\mathbf{M}_r \succ 0$$

if c is sufficiently small. Therefore, $\hat{\boldsymbol{\lambda}} + c \times \mathbf{d}_r$ is a strictly feasible point for the dual problem. This proves strong duality. ■

Proof of Theorem 2: The pair $(\mathbf{v}\mathbf{v}^*, \boldsymbol{\eta})$ is feasible for the primal problem (14). With no loss of generality, assume that $\mathcal{G} = \{1, \ldots, |\mathcal{G}|\}$ and $\mathcal{B} = \mathcal{M} \backslash \mathcal{G} = \{|\mathcal{G}| + 1, \ldots, |\mathcal{M}|\}$. Consider the following choice for the dual certificate:

$$\hat{\boldsymbol{\lambda}} := \begin{bmatrix} -\mathbf{J}_{\mathcal{G}}^{+}(\overline{\mathbf{v}}) \left(2\overline{\mathbf{M}}\overline{\mathbf{v}} - \mu \mathbf{J}_{\mathcal{B}}(\overline{\mathbf{v}})\boldsymbol{\gamma}\right) \\ -\mu\boldsymbol{\gamma} \end{bmatrix}, \quad (31)$$

where

$$\boldsymbol{\gamma} := \left[\frac{\eta_{|\mathcal{G}|+1}}{|\eta_{|\mathcal{G}|+1}|}, \frac{\eta_{|\mathcal{G}|+2}}{|\eta_{|\mathcal{G}|+2}|}, \ldots, \frac{\eta_{|\mathcal{M}|}}{|\eta_{|\mathcal{M}|}|}\right]^{\text{T}}. \quad (32)$$

Define $\hat{\mathbf{H}} \triangleq \mathbf{M} + \sum_{r\in\mathcal{M}} \hat{\lambda}_r \mathbf{M}_r$. According to the definition of $\hat{\boldsymbol{\lambda}}$, it is straightforward to verify that complementary slackness holds, i.e.,

$$\langle \hat{\mathbf{H}}, \mathbf{v}\mathbf{v}^* \rangle = 0. \quad (33)$$

In addition, according to the left side of (20), one can write:

$$\begin{aligned}|\hat{\lambda}_r| &\leq \|2\mathbf{J}_{\mathcal{G}}^{+}(\overline{\mathbf{v}})\overline{\mathbf{M}}\overline{\mathbf{v}}\|_\infty + \mu \times \|\mathbf{J}_{\mathcal{G}}^{+}(\overline{\mathbf{v}})\mathbf{J}_{\mathcal{B}}(\overline{\mathbf{v}})\|_\infty \\ &= \frac{\kappa(\mathbf{M})}{2} \times \delta_{\mathcal{G};\infty}(\mathbf{v}) + \mu \times \zeta_{\mathcal{G};\infty}(\mathbf{v}) < \mu, \quad (34)\end{aligned}$$

for every $r \in \mathcal{G}$. This implies that $\hat{\boldsymbol{\lambda}}$ satisfies the dual constraints (25c). Moreover, according to the right side (20), we have

$$\begin{aligned}&\kappa(\hat{\mathbf{H}}) \geq \kappa(\mathbf{M}) + \sum_{r\in\mathcal{M}} \kappa(\hat{\lambda}_r \mathbf{M}_r) \\ &\geq \kappa(\mathbf{M}) - \sum_{r\in\mathcal{M}} 2|\hat{\lambda}_r| \|\mathbf{M}_r\|_2 \\ &\geq \kappa(\mathbf{M}) - 2\|\hat{\boldsymbol{\lambda}}\|_1 \\ &= \kappa(\mathbf{M}) - 2\|2\mathbf{J}_{\mathcal{G}}^{+}\overline{\mathbf{M}}\overline{\mathbf{v}}\|_1 - 2\mu(\|\mathbf{J}_{\mathcal{G}}^{+}\mathbf{J}_{\mathcal{B}}\|_1 + 1)|\mathcal{B}| \\ &= \kappa(\mathbf{M})(1-\delta_{\mathcal{G};1}(\mathbf{v})) - 2\mu(\zeta_{\mathcal{G};1}(\mathbf{v})+1)|\mathcal{B}| > 0 \quad (35)\end{aligned}$$

As a result, the constraint (25b) holds and moreover,

$$\text{rank}\{\hat{\mathbf{H}}\} = n-1. \quad (36)$$

Therefore, $\hat{\boldsymbol{\lambda}}$ is dual feasible and certifies that $(\mathbf{v}\mathbf{v}^*, \boldsymbol{\eta})$ is the unique solution for problem (14) (note that strong duality holds because of the argument made in the proof of Theorem 1). ■

Proof of Theorem 3: With no loss of generality, assume that $\mathcal{G} = \{1, \ldots, |\mathcal{G}|\}$ and $\mathcal{B} = \mathcal{M} \setminus \mathcal{G} = \{|\mathcal{G}| + 1, \ldots, |\mathcal{M}|\}$. We construct the vectors $\hat{\boldsymbol{\lambda}}$ and $\boldsymbol{\gamma}$ as follows:

$$\hat{\boldsymbol{\lambda}} := \begin{bmatrix} -\mathbf{J}_{\mathcal{G}}^{+}(\overline{\mathbf{v}}) \left(2\overline{\mathbf{M}}\overline{\mathbf{v}} - \mu \mathbf{J}_{\mathcal{B}}(\overline{\mathbf{v}})\boldsymbol{\gamma}\right) \\ -\mu\boldsymbol{\gamma} \end{bmatrix}, \quad (37)$$

where

$$\boldsymbol{\gamma} := \left[\frac{\omega_{|\mathcal{G}|+1} + \eta_{|\mathcal{G}|+1}}{|\omega_{|\mathcal{G}|+1} + \eta_{|\mathcal{G}|+1}|}, \ldots, \frac{\omega_{|\mathcal{M}|} + \eta_{|\mathcal{M}|}}{|\omega_{|\mathcal{M}|} + \eta_{|\mathcal{M}|}|}\right]^{\text{T}}. \quad (38)$$

Define $\hat{\mathbf{H}} \triangleq \mathbf{M} + \sum_{r\in\mathcal{M}} \hat{\lambda}_r \mathbf{M}_r$. Similar to the proof of Theorem 2, it can be verified that the following properties hold for $\hat{\boldsymbol{\lambda}}$ and $\hat{\mathbf{H}}$:

$$\|\hat{\boldsymbol{\lambda}}\|_\infty \leq \mu, \quad (39a)$$

$$\langle \hat{\mathbf{H}}, \mathbf{v}\mathbf{v}^* \rangle = 0, \quad (39b)$$

$$\kappa(\hat{\mathbf{H}}) > \kappa(\mathbf{M})(1-\delta_{\mathcal{G};1}(\mathbf{v})) - 2\mu(\zeta_{\mathcal{G};1}(\mathbf{v})+1)|\mathcal{B}|. \quad (39c)$$

Now, primal feasibility of the point $(\mathbf{W}^{\text{opt}}, \boldsymbol{\nu}^{\text{opt}})$, combined with the inequality (39a) implies that:

$$\begin{aligned}\|\boldsymbol{\nu}^{\text{opt}}\|_1 = &\sum_{r\in\mathcal{M}} |\langle \mathbf{M}_r, \mathbf{W}^{\text{opt}} - \mathbf{v}\mathbf{v}^*\rangle - \omega_r - \eta_r| \\ \geq &\sum_{r\in\mathcal{G}} |\langle \mathbf{M}_r, \mathbf{W}^{\text{opt}} - \mathbf{v}\mathbf{v}^*\rangle| - \sum_{r\in\mathcal{G}} |\omega_r| \\ &- \sum_{r\in\mathcal{B}} \gamma_r \langle \mathbf{M}_r, \mathbf{W}^{\text{opt}} - \mathbf{v}\mathbf{v}^*\rangle + \sum_{r\in\mathcal{B}} |\omega_r + \eta_r| \\ \geq &\frac{1}{\mu}\sum_{r\in\mathcal{M}} \hat{\lambda}_r \langle \mathbf{M}_r, \mathbf{W}^{\text{opt}} - \mathbf{v}\mathbf{v}^*\rangle - \sum_{r\in\mathcal{G}} |\omega_r| \\ &+ \sum_{r\in\mathcal{B}} |\omega_r + \eta_r| \\ \geq &\frac{1}{\mu}\langle \hat{\mathbf{H}}, \mathbf{W}^{\text{opt}} - \mathbf{v}\mathbf{v}^*\rangle - \frac{1}{\mu}\langle \mathbf{M}, \mathbf{W}^{\text{opt}} - \mathbf{v}\mathbf{v}^*\rangle \\ &+ \|\boldsymbol{\omega} + \boldsymbol{\eta}\|_1 - 2\|\boldsymbol{\omega}\|_1. \end{aligned} \tag{40}$$

On the other hand, evaluating the objective function of the primal problem at $(\mathbf{v}\mathbf{v}^*, \boldsymbol{\omega} + \boldsymbol{\eta})$ yields that

$$\|\boldsymbol{\nu}^{\text{opt}}\|_1 \leq -\frac{1}{\mu}\langle \mathbf{M}, \mathbf{W}^{\text{opt}} - \mathbf{v}\mathbf{v}^*\rangle + \|\boldsymbol{\omega} + \boldsymbol{\eta}\|_1. \tag{41}$$

Replacing $\|\boldsymbol{\nu}^{\text{opt}}\|_1$ on the left side of (41) with the lower bound offered by (40) leads to

$$\frac{1}{\mu}\langle \hat{\mathbf{H}}, \mathbf{W}^{\text{opt}} - \mathbf{v}\mathbf{v}^*\rangle \leq 2\|\boldsymbol{\omega}\|_1. \tag{42}$$

Hence, according to (39b), we have

$$\langle \hat{\mathbf{H}}, \mathbf{W}^{\text{opt}}\rangle \leq 2 \times \mu \times \|\boldsymbol{\omega}\|_1. \tag{43}$$

Now, consider the eigenvalue decomposition

$$\hat{\mathbf{H}} = \mathbf{U}\,\text{diag}\{\boldsymbol{\tau}\}\mathbf{U}^*, \tag{44}$$

where $\boldsymbol{\tau} = [\tau_n, \ldots, \tau_2, 0]^{\text{T}}$ collects the eigenvalues of $\hat{\mathbf{H}}$ in descending order and $\mathbf{U}$ is a unitary matrix whose last column is equal to $\mathbf{v}/\|\mathbf{v}\|_2$. Define

$$\hat{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{W}} & \tilde{\mathbf{w}} \\ \tilde{\mathbf{w}}^{\text{T}} & \tilde{W}_{nn} \end{bmatrix} = \mathbf{U}^*\mathbf{W}^{\text{opt}}\mathbf{U}, \tag{45}$$

where $\tilde{\mathbf{W}} \in \mathbb{H}^n$, $\tilde{\mathbf{w}} \in \mathbb{C}^n$ and $\tilde{W}_{nn} \in \mathbb{R}$. Therefore,

$$\begin{aligned}\text{trace}\{\tilde{\mathbf{W}}\} &\leq \frac{1}{\tau_2}\langle \text{diag}\{\boldsymbol{\tau}\}, \tilde{\mathbf{W}}\rangle \leq \frac{1}{\tau_2}\langle \hat{\mathbf{H}}, \hat{\mathbf{W}}\rangle \\ &\leq \frac{1}{\tau_2}\langle \hat{\mathbf{H}}, \mathbf{W}^{\text{opt}}\rangle \leq \frac{1}{\tau_2} \times 2 \times \mu \times \|\boldsymbol{\omega}\|_1. \end{aligned} \tag{46}$$

Moreover, due to the positive semidefiniteness of $\hat{\mathbf{W}}$, it can be easily observed that

$$\|\tilde{\mathbf{w}}\|_2^2 \leq \tilde{W}_{nn} \times \text{trace}\{\tilde{\mathbf{W}}\}. \tag{47}$$

Hence, by defining

$$\alpha = \tilde{W}_{nn}/\|\mathbf{v}\|_2^2, \tag{48}$$

one can write:

$$\begin{aligned}\|\mathbf{W}^{\text{opt}} - \alpha\mathbf{v}\mathbf{v}^*\|_F^2 &= \|\hat{\mathbf{W}} - \tilde{W}_{nn}\mathbf{e}_n\mathbf{e}_n^{\text{T}}\|_F^2 \\ &= \|\tilde{\mathbf{W}}\|_F^2 + 2\|\tilde{\mathbf{w}}\|_2^2 \\ &\leq \|\tilde{\mathbf{W}}\|_F^2 + 2\tilde{W}_{nn} \times \text{trace}\{\tilde{\mathbf{W}}\} \\ &\leq \|\tilde{\mathbf{W}}\|_F^2 \\ &\quad + 2\left(\text{trace}\{\mathbf{W}^{\text{opt}}\} - \text{trace}\{\tilde{\mathbf{W}}\}\right)\text{trace}\{\tilde{\mathbf{W}}\} \\ &\leq 2 \times \text{trace}\{\tilde{\mathbf{W}}\} \times \text{trace}\{\mathbf{W}^{\text{opt}}\} \\ &\leq \frac{4}{\tau_2} \times \mu \times \|\boldsymbol{\omega}\|_1 \times \text{trace}\{\mathbf{W}^{\text{opt}}\} \end{aligned} \tag{49}$$

The proof is completed by combining (49) with the lower bound on τ_2 given in (39c). ■