Power System State Estimation with a Limited Number of Measurements

Ramtin Madani, Morteza Ashraphijuo, Javad Lavaei and Ross Baldick

***Abstract*—This paper is concerned with the power system state estimation (PSSE) problem, which aims to find the unknown operating point of a power network based on a given set of measurements. The measurements of the PSSE problem are allowed to take any arbitrary combination of nodal active powers, nodal reactive powers, nodal voltage magnitudes and line flows. This problem is non-convex and NP-hard in the worst case. We develop a set of convex programs with the property that they all solve the non-convex PSSE problem in the case of noiseless measurements as long as the voltage angles are relatively small. This result is then extended to a general PSSE problem with noisy measurements, and an upper bound on the estimation error is derived. The objective function of each convex program developed in this paper has two terms: one accounting for the non-convexity of the power flow equations and another one for estimating the noise levels. The proposed technique is demonstrated on the 1354-bus European network.**

I. Introduction

The power system state estimation (PSSE) is the problem of determining the state of a power network, namely nodal complex voltages, based on certain measurements taken at buses and over branches of the network. This problem plays a crucial role in control and operation of power networks [1]. As a special case, the power flow (PF) problem aims to find the state of the system, given noiseless measurements at buses. This problem has been studied extensively for many years, with the goal of designing an efficient computational method that is able to cope with the non-convexity of the power flow equations. Since 1962, several linearization and local search algorithms have been developed for this classical problem, and the current practice in the power industry relies on linearization and/or Newton's method (depending on the time scale and whether this problem is solved for planning or real-time operation) [2]–[4].

To tackle the non-convexity of the feasible region described by the AC power flow equations, the semidefinite programming (SDP) relaxation technique can be used [5], [6]. Sparked by the papers [7] and [8], the SDP relaxation method has received a significant attention in the power society [9], [10]. The work [8] has developed an SDP relaxation to find a global solution of the optimal power flow (OPF) problem, and showed that the relaxation is exact for IEEE test systems. Recent advances in leveraging the sparsity of power systems have made SDP problems computationally more tractable [11]–[16].

Ramtin Madani, Morteza Ashraphijuo and Javad Lavaei are with the Department of Industrial Engineering and Operations Research, University of California, Berkeley (ramtin.madani@berkeley.edu, ashraphijuo@berkeley.edu and lavaei@berkeley.edu). Ross Baldick is with the Department of Electrical and Computer Engineering, University of Texas at Austin (baldick@ece.utexas.edu). This work was supported by ONR YIP Award, DARPA Young Faculty Award, NSF CAREER Award 1351279, and NSF EECS Awards 1406894 and 1406865.

Recently, the SDP relaxation technique has been applied to the PSSE problem, and gained success in the case where the number of measurements is significantly higher than the dimension of the unknown state of the system (i.e., twice the number of buses minus one) [17], [18]. The papers [19] and [20] 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 using the alternating direction method of multipliers (ADMM) and Lagrange decomposition have been developed. Moreover, the formulations in [17] and [18] have been extended in [19] and [21] to accommodate PMU measurements. The work [21] has studied a variety of regularization methods to solve the PSSE problem in 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 the recent work [22], we have investigated the non-convex PF problem in two steps: (i) PF is transformed into an optimization problem by augmenting PF with a suitable objective function, (ii) the resulting non-convex problem is relaxed to an SDP. The designed objective function is not unique and there are infinitely many choices for this function. It has been proven that if the PF solution belongs to the recovery region of the SDP problem, then the solution can be found precisely using SDP. This recovery region contains voltage vectors with relatively small angles. Note that voltage angles are often small in practice due to practical considerations, which has two practical implications: (i) linearization would be able to find an approximate solution, (ii) Newton's method would converge by initializing all voltage angles at zero. Linearization techniques offer low-complex approximate models that can provide insights into power systems, whereas Newton's method is an attractive numerical algorithm that has been used in the power industry for many years. Some of the advantage of the SDP technique over the aforementioned approaches are as follows:

- A one-time linearization of the power flow equations (known as DC modeling) solves the PF problem approximately by linearizing the laws of physics. However, the SDP problem finds the correct solution (with any arbitrary precision) as long as it belongs to the corresponding recovery region.
- The basin of attraction of Newton's method is chaotic and hard to characterize, but the recovery region of the SDP problem is explicitly characterizable via matrix inequalities.
- The SDP relaxation provides a convex model for the PF problem, which can be solved by many numerical

algorithms (such as Newton's method).

By building upon the results developed in [22], the goal of this paper is to solve the PSSE problem via a penalized convex program (an SDP-type problem), where the measurement equations are softly penalized in the objective function as opposed to being imposed as equality constraints. The objective function of the convex program developed here has two terms: (i) the one previously used for the PF problem in the noiseless case to deal with non-convexity, (ii) another one added to account for the noisy measurements. We prove that the penalized convex program precisely solves the PSSE problem in the case of noiseless measurements as long as the solution belongs to its associated recovery region (the region includes solutions with small voltage angles). In the noisy case, the SDP matrix solution of the convex program may or may not have rank-1 due to corrupted measurements. We design an algorithm to estimate the solution of the PSSE problem from that of the penalized convex program, and derive an upper bound on the estimation error. We demonstrate the efficacy of the proposed technique on a large test system with over 1000 buses.

A. 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. $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. The symbol $\langle\cdot,\cdot\rangle$ represents the Frobenius inner product of matrices. Matrices are shown by capital and bold letters. The notations $(\cdot)^{\mathrm{T}}$ and $(\cdot)^*$ denote transpose and conjugate transpose, respectively. The symbol "$\mathbb{i}$" denotes the imaginary unit. The notation $\langle \mathbf{A}, \mathbf{B}\rangle$ represents $\mathrm{trace}\{\mathbf{A}^*\mathbf{B}\}$, which is the inner product of $\mathbf{A}$ 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. Similarly, $\mathbf{W} \succ 0$ means that $\mathbf{W}$ is Hermitian and positive definite. The (i,j) entry of $\mathbf{W}$ is denoted as W_{ij}. $\mathbf{0}_n$ 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 $\mathcal{X} \subseteq \{1,\dots,m\}$ and $\mathcal{Y} \subseteq \{1,\dots,n\}$. Similarly, $\mathbf{W}[\mathcal{X}]$ denotes the submatrix of $\mathbf{W}$ induced by those rows of $\mathbf{W}$ indexed by $\mathcal{X}$. The interior of a set $\mathcal{D} \in \mathbb{C}^n$ is denoted as $\mathrm{int}\{\mathcal{D}\}$.

II. Preliminaries

Let $\mathcal{N}$ and $\mathcal{L}$ denote the sets of buses (nodes) and branches (edges) of the power network under study. Denote the number of buses as n and let p_k and q_k represent the net active and reactive power injections at every bus $k \in \mathcal{N}$. Define $\mathbf{p} = [p_1\ p_2\ \cdots\ p_n]^{\mathrm{T}}$ and $\mathbf{q} = [q_1\ q_2\ \cdots\ q_n]^{\mathrm{T}}$ as the vectors containing net injected active and reactive powers, respectively. The complex voltage phasor at bus k is denoted by v_k, whose magnitude and phase are shown as $|v_k|$ and $\measuredangle v_k$, respectively. The vector of all nodal voltages is shown as $\mathbf{v}$. We orient the lines of the network arbitrarily and define $s_{f;l} = p_{f;l} + q_{f;l}\mathbb{i}$ and $s_{t;l} = p_{t;l} + q_{t;l}\mathbb{i}$ as the complex power injections at the *from* and *to* ends of each branch $l \in \mathcal{L}$. Note that $p_{f;l}$ and $p_{f;l}$ denote the active powers entered the line from both ends, while $q_{f;l}$ and $q_{f;l}$ denote the reactive powers over the line.

Given an edge $(j,k) \in \mathcal{L}$, let $g_{jk} + \mathbb{i}b_{jk}$ denote the admittance of the transmission line between nodes j and k. Due to the passivity of the line, it is assumed that $g_{jk} \geq 0$ and $b_{jk} \leq 0$. Define $\mathbf{Y} \in \mathbb{C}^{n\times n}$ as the admittance matrix of the network. Likewise, define $\mathbf{Y}_f \in \mathbb{C}^{|\mathcal{L}|\times n}$ and $\mathbf{Y}_t \in \mathbb{C}^{|\mathcal{L}|\times n}$ as the *from* and *to* branch admittance matrices, respectively. These matrices satisfy the equations

$$\mathbf{i} = \mathbf{Y}\mathbf{v}, \quad \mathbf{i}_f = \mathbf{Y}_f\mathbf{v}, \quad \mathbf{i}_t = \mathbf{Y}_t\mathbf{v}, \tag{1}$$

where $\mathbf{i} \in \mathbb{C}^n$ is the complex nodal current injection, and $\mathbf{i}_f \in \mathbb{C}^{|\mathcal{L}|}$ and $\mathbf{i}_t \in \mathbb{C}^{|\mathcal{L}|}$ are the vectors of currents at the *from* and *to* ends of branches, respectively. Although the results to be developed in this paper hold for a general matrix $\mathbf{Y}$, we make a few assumptions to streamline the presentation:

- The network is a connected graph.
- Every line of the network consists of a series impedance with nonnegative resistance and inductance.
- The shunt elements are ignored for simplicity in guaranteeing the observability of the network, which ensures that $\mathbf{Y} \times \mathbf{1}_n = \mathbf{0}_n$.

The power balance equations can be expressed as

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

Let $\mathbf{Y} = \mathbf{G} + \mathbf{B}\mathbb{i}$, where $\mathbf{G}$ and $\mathbf{B}$ are the conductance and susceptance matrices, respectively. For every $k \in \mathcal{N}$, define

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

$$\mathbf{Y}_{p;k} \triangleq (\mathbf{Y}^*\mathbf{e}_k\mathbf{e}_k^* + \mathbf{e}_k\mathbf{e}_k^*\mathbf{Y})/2, \tag{3b}$$

$$\mathbf{Y}_{q;k} \triangleq (\mathbf{Y}^*\mathbf{e}_k\mathbf{e}_k^* - \mathbf{e}_k\mathbf{e}_k^*\mathbf{Y})/2\mathbb{i} \tag{3c}$$

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

$$|v_k|^2 = \langle\mathbf{v}\mathbf{v}^*,\mathbf{E}_k\rangle, \ \ p_k = \langle\mathbf{v}\mathbf{v}^*,\mathbf{Y}_{p;k}\rangle, \ \ q_k = \langle\mathbf{v}\mathbf{v}^*,\mathbf{Y}_{q;k}\rangle, \tag{4}$$

for every $k \in \mathcal{N}$. Moreover, let $\mathbf{d}_1,\dots,\mathbf{d}_{|\mathcal{L}|}$ denote the standard basis vectors in $\mathbb{R}^{|\mathcal{L}|}$. Given a line $l \in \mathcal{L}$ from node i to node j, we define

$$\mathbf{Y}_{p_f;l} \triangleq (\mathbf{Y}_f^*\mathbf{d}_l\mathbf{e}_i^* + \mathbf{e}_i\mathbf{d}_l^*\mathbf{Y}_f)/2, \tag{5a}$$

$$\mathbf{Y}_{q_f;l} \triangleq (\mathbf{Y}_f^*\mathbf{d}_l\mathbf{e}_i^* - \mathbf{e}_i\mathbf{d}_l^*\mathbf{Y}_f)/2\mathbb{i}, \tag{5b}$$

$$\mathbf{Y}_{p_t;l} \triangleq (\mathbf{Y}_t^*\mathbf{d}_l\mathbf{e}_j^* + \mathbf{e}_j\mathbf{d}_l^*\mathbf{Y}_t)/2, \tag{5c}$$

$$\mathbf{Y}_{q_t;l} \triangleq (\mathbf{Y}_t^*\mathbf{d}_l\mathbf{e}_j^* - \mathbf{e}_j\mathbf{d}_l^*\mathbf{Y}_t)/2\mathbb{i} \tag{5d}$$

and then write the branch parameters $p_{f;l}$, $q_{f;l}$, $p_{t;l}$ and $q_{t;l}$ as the inner product of $\mathbf{vv}^*$ with the matrices $\mathbf{Y}_{p_f;l}$, $\mathbf{Y}_{q_f;l}$, $\mathbf{Y}_{p_t;l}$ and $\mathbf{Y}_{q_t;l}$ as follows:

$$p_{f;l} = \langle \mathbf{vv}^*, \mathbf{Y}_{p_f;l} \rangle, \qquad q_{f;l} = \langle \mathbf{vv}^*, \mathbf{Y}_{q_f;l} \rangle,$$
$$p_{t;l} = \langle \mathbf{vv}^*, \mathbf{Y}_{p_t;l} \rangle, \qquad q_{t;l} = \langle \mathbf{vv}^*, \mathbf{Y}_{q_t;l} \rangle,$$

for every $l \in \mathcal{L}$. Equations (4) and (6) offer a compact formulation for common measurements in power networks. A general state estimation problem can be formulated as finding a solution to a system of quadratic equations of the form

$$x_r = \langle \mathbf{vv}^*, \mathbf{M}_r \rangle + \omega_r, \qquad r = 1, \ldots, m, \tag{7}$$

where

- $x_1, \ldots, x_m$ are the known measurements/specifications.
- $\omega_1, \ldots, \omega_m$ are some unknown measurement noises, with known statistical information.
- $\mathbf{M}_1, \ldots, \mathbf{M}_m$ are some known $n \times n$ Hermitian matrices (e.g., they could be any subset of the matrices defined in (3) and (5)).

In the case where all noises $\omega_1, \ldots, \omega_m$ are equal to zero, the above problem reduces to the well-known power flow problem.

A. Semidefinite Relaxation

The state estimation problem, as a general case of the power flow problem, is nonconvex due to the quadratic matrix $\mathbf{vv}^*$. Hence, it is desirable to convexify the problem. By defining $\mathbf{W} \triangleq \mathbf{vv}^*$, the quadratic equations in (7) can be linearly formulated in terms of $\mathbf{W}$:

$$x_r = \langle \mathbf{W}, \mathbf{M}_r \rangle + \omega_r, \qquad r = 1, \ldots, m. \tag{8}$$

Provided that the quadratic measurements $x_1, \ldots, x_m$ are noiseless, solving the non-convex equations (7) is tantamount to finding a rank-1 and positive semidefinite matrix $\mathbf{W} \in \mathbb{H}_+^n$ satisfying the above linear equations (because such a matrix $\mathbf{W}$ could then be decomposed as $\mathbf{vv}^*$). The problem of finding a matrix $\mathbf{W} \in \mathbb{H}_+^n$ satisfying the linear equations in (8) is regarded as a convex *relaxation* of (7) since it includes no restriction on the rank of $\mathbf{W}$. Although (7) normally has a finite number of solutions whenever $m \geq 2n - 1$, its SDP relaxation (8) is expected to have infinitely many solutions because the matrix variable $\mathbf{W}$ includes $O(n^2)$ scalar variables as opposed to $O(n)$. Hence, it is desirable to minimize a convex function of $\mathbf{W}$ subject to the SDP relaxation of the noiseless measurement constraints to make the solution unique.

B. Sensitivity Analysis

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 a simple phase shift, we assume that $\measuredangle v_k$ is equal to zero at a pre-selected bus (named, slack bus).

Notation 1. *Let $\mathcal{O}$ denote the set of all buses of the network except for the slack bus. Then, the operating point of the power system can be characterized in terms of the real-valued vector*

$$\overline{\mathbf{v}} \triangleq \left[\mathrm{Re}\{\mathbf{v}[\mathcal{N}]^{\mathrm{T}}\} \quad \mathrm{Im}\{\mathbf{v}[\mathcal{O}]^{\mathrm{T}}\} \right]^{\mathrm{T}} \in \mathbb{R}^{2n-1}. \tag{9}$$

In addition, for every $n \times n$ Hermitian matrix $\mathbf{X}$, the notation $\overline{\mathbf{X}}$ represents the following $(2n-1) \times (2n-1)$ real-valued and symmetric matrix:

$$\overline{\mathbf{X}} = \begin{bmatrix} \mathrm{Re}\{\mathbf{X}[\mathcal{N}, \mathcal{N}]\} & -\mathrm{Im}\{\mathbf{X}[\mathcal{N}, \mathcal{O}]\} \\ \mathrm{Im}\{\mathbf{X}[\mathcal{O}, \mathcal{N}]\} & \mathrm{Re}\{\mathbf{X}[\mathcal{O}, \mathcal{O}]\} \end{bmatrix}. \tag{10}$$

Definition 1. *Define the function $\mathcal{A}(\overline{\mathbf{v}}) : \mathbb{R}^{2n-1} \to \mathbb{R}^m$ as the mapping from the state of the power network (i.e., $\overline{\mathbf{v}}$) to the vector of noiseless specifications (i.e., $\mathbf{x}$). The r-th component of $\mathcal{A}(\overline{\mathbf{v}})$ can be expressed as*

$$\mathcal{A}_r(\overline{\mathbf{v}}) \triangleq \langle \mathbf{vv}^*, \mathbf{M}_r \rangle, \qquad r = 1, \ldots, m.$$

Define also the sensitivity matrix $\mathbf{J}_{\mathcal{A}}(\overline{\mathbf{v}}) \in \mathbb{R}^{(2n-1) \times m}$ as the Jacobian of $\mathcal{A}(\overline{\mathbf{v}})$ at the point $\overline{\mathbf{v}}$, which is equal to

$$\mathbf{J}_{\mathcal{A}}(\overline{\mathbf{v}}) = 2 \left[\overline{\mathbf{M}}_1 \, \overline{\mathbf{v}} \quad \overline{\mathbf{M}}_2 \, \overline{\mathbf{v}} \quad \ldots \quad \overline{\mathbf{M}}_m \, \overline{\mathbf{v}} \right].$$

According to the inverse function theorem, if $\mathbf{J}_{\mathcal{A}}(\overline{\mathbf{v}})$ has full row rank, then the inverse of the function $\mathcal{A}(\overline{\mathbf{v}})$ exists in a neighborhood of the point $\overline{\mathbf{v}}$. Similarly, it follows from the Kantorovich Theorem that, under the previous assumption, the equation (7) 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. We will show that the invertibility of $\mathbf{J}_{\mathcal{A}}(\overline{\mathbf{v}})$ is beneficial not only for Newton's method but also for the SDP relaxation technique.

Definition 2. *A vector of complex voltages $\mathbf{v}$ is said to be observable through the system of equations (7) if $\mathbf{J}_{\mathcal{A}}(\overline{\mathbf{v}})$ has full row rank. Define $\mathcal{J}_{\mathcal{A}} \in \mathbb{C}^n$ as the set of all such observable voltage vectors.*

The point $\overline{\mathbf{v}} = \overline{\mathbf{1}}_n$ (associated with $\mathbf{v} = \mathbf{1}_n$) is often regarded as a nominal state for: (i) the linearization of the quadratic power flow equations, (ii) the initialization of local search algorithms used for nonlinear power flow equations. Throughout this paper, we assume that $\mathbf{J}_{\mathcal{A}}(\overline{\mathbf{1}}_n)$ has full row rank.

Assumption 1. *The point $\mathbf{1}_n$ is observable through the system of equations (7) (i.e., $\mathbf{1}_n \in \mathcal{J}_{\mathcal{A}}$).*

We have shown in [22] that the above assumption holds for the classical power flow problem to be stated next.

C. Classical Power Flow Problem

The power flow (PF) problem can be regarded as a noiseless state estimation problem, for which $\omega_1, \omega_2, \ldots, \omega_m$ are all equal to zero. As a special case of the PF problem, the classical PF problem is concerned with the case where the number of quadratic constraints (namely m) is equal to $2n - 1$, the measurements are all at buses as opposed to a combination of buses and lines, and there is no measurement noise. To formulate the problem, three basic types of buses are considered based on the parameters known at each bus:

- PQ bus: p_k and q_k are specified.
- PV bus: p_k and $|v_k|$ are specified.
- The slack bus: $|v_k|$ is specified.

Each PQ bus represents a load bus or possibly a generator bus, whereas each PV bus represents a generator bus. Given the specified parameters at every bus of the network, the classical PF problem aims to solve the network equations in order to find an operating point that fits the input values. Note that Assumption 1 holds for the classical power flow problem.

D. Noiseless Case

To be able to proceed with this paper, we present the key results of [22] in this section. Consider the case where $m = 2n-1$ and the measurements in (7) are noiseless:

$$x_r = \langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_r \rangle, \qquad r = 1, \ldots, m. \tag{11}$$

To solve this set of quadratic equations through a convex relaxation, we aim to propose a family of convex optimization problems of the form

$$\begin{aligned} \underset{\mathbf{W}\in\mathbb{H}^n}{\text{minimize}} \quad & \langle \mathbf{W}, \mathbf{M} \rangle & \text{(12a)} \\ \text{subject to} \quad & \langle \mathbf{W}, \mathbf{M}_r \rangle = x_r, \quad r = 1, \ldots, m, & \text{(12b)} \\ & \mathbf{W} \succeq 0, & \text{(12c)} \end{aligned}$$

where the matrix $\mathbf{M} \in \mathbb{H}^n_+$ is to be designed. Unlike the compressing sensing literature that assumes $\mathbf{M} = \mathbf{I}_n$, it is desirable to contrive $\mathbf{M}$ such that the above problem yields a unique rank-1 solution $\mathbf{W}$ from which a feasible solution $\mathbf{v}$ can be recovered for (11). Notice that the existence of such a rank-1 solution depends in part on its input specifications $x_1, x_2, \ldots, x_m$. It is said that the SDP problem (12) solves the set of equations (11) for the input $\mathbf{x} = [x_1, x_2, \ldots, x_m]^{\mathrm{T}}$ if (12) has a unique rank-1 solution.

Definition 3. *Given $\mathbf{M} \in \mathbb{H}^n_+$, a voltage vector $\mathbf{v}$ is said to be recoverable if $\mathbf{W} = \mathbf{v}\mathbf{v}^*$ is the unique solution of the SDP problem* (12) *for some $x_1, x_2 \ldots, x_m \in \mathbb{R}$. Define $\mathcal{R}_{\mathcal{A}}(\mathbf{M})$ as the set of all recoverable vectors of voltages.*

Note that the set $\mathcal{R}_{\mathcal{A}}(\mathbf{M})$ is indeed the collection of all possible operating points $\mathbf{v}$ that can be found through (12) associated with different values of $x_1, x_2, \ldots, x_m$. In order to narrow the search space for the matrix $\mathbf{M}$, we impose some conditions on this matrix below.

Assumption 2. *The matrix $\mathbf{M}$ satisfies the properties:*

- $\mathbf{M} \succeq 0$
- 0 *is a simple eigenvalue of* $\mathbf{M}$
- *The vector $\mathbf{1}_n$ belongs to the null space of* $\mathbf{M}$.

Note that the matrix $-\mathbf{B}$ satisfies Assumption 2. The next lemma reveals an interesting property of (12).

Lemma 1 (see [22])**.** *If Assumptions 1 and 2 hold and $\mathbf{v} \in \mathcal{R}_{\mathcal{A}}(\mathbf{M}) \cap \mathcal{J}_{\mathcal{A}}$, then strong duality holds between the primal SDP* (12) *with the input $\mathbf{x} = \mathcal{A}(\mathbf{v})$ and its dual. Moreover, the vector*

$$\boldsymbol{\lambda}_{\mathcal{A}}(\mathbf{v}, \mathbf{M}) \triangleq -2\mathbf{J}_{\mathcal{A}}(\overline{\mathbf{v}})^{-1}\overline{\mathbf{M}}\,\overline{\mathbf{v}} \tag{13}$$

is the unique vector of Lagrange multipliers associated with the constraints in (12b).

Definition 4. *Define $\mathcal{D}^n$ as the set of all $n \times n$ positive semidefinite Hermitian matrices with the sum of two smallest eigenvalues greater than* 0.

The following theorem offers a nonlinear matrix inequality to characterize the interior of the set of recoverable voltage vectors, except for a subset of measure zero of this interior at which the Jacobian of $\mathcal{A}(\overline{\mathbf{v}})$ loses rank.

Theorem 1 (see [22])**.** *If Assumptions 1 and 2 hold, then the interior of the set $\mathcal{R}_{\mathcal{A}}(\mathbf{M})$ can be characterized as*

$$\mathrm{int}\{\mathcal{R}_{\mathcal{A}}(\mathbf{M})\} \cap \mathcal{J}_{\mathcal{A}} = \{\mathbf{v} \in \mathcal{J}_{\mathcal{A}} |\ \mathbf{F}_{\mathcal{A}}(\mathbf{v}, \mathbf{M}) \in \mathcal{D}^n\},$$

where the matrix function $\mathbf{F}_{\mathcal{A}}: \mathcal{J}_{\mathcal{A}} \times \mathbb{H}^n_+ \to \mathbb{H}^n$ is defined as

$$\mathbf{F}_{\mathcal{A}}(\mathbf{v}, \mathbf{M}) \triangleq \mathbf{M} + \sum_{r=1}^{m} \lambda_r \mathbf{M}_r \tag{14}$$

and λ_r denotes the r^{th} entry of $\boldsymbol{\lambda}_{\mathcal{A}}(\mathbf{v}, \mathbf{M})$ defined in (13).

The following theorem shows that if Assumptions 1 and 2 hold, then the region $\mathcal{R}_{\mathcal{A}}(\mathbf{M})$ contains the nominal point $\mathbf{1}_n$ and a ball around it.

Theorem 2 (see [22])**.** *If Assumptions 1 and 2 hold, then the region $\mathcal{R}_{\mathcal{A}}(\mathbf{M})$ has a non-empty interior containing the point $\mathbf{1}_n$.*

III. Main Results

In the presence of measurement noises, the convex problem (12) may be infeasible (if $m > 2n-1$) or result in a poor approximate solution. To remedy this issue, a standard approach is to estimate the noise values through some auxiliary variables $\nu_1, \ldots, \nu_m \in \mathbb{R}$. This can be achieved by incorporating a regularization term $\phi : \mathbb{R}^m \to \mathbb{R}$ into the objective function that elevates the likelihood of the estimated noise:

$$\begin{aligned} \underset{\substack{\mathbf{W}\in\mathbb{H}^n \\ \boldsymbol{\nu}\in\mathbb{R}^m}}{\text{minimize}} \quad & \langle \mathbf{W}, \mathbf{M} \rangle + \mu \times \phi(\boldsymbol{\nu}) & \text{(15a)} \\ \text{subject to} \quad & \langle \mathbf{W}, \mathbf{M}_r \rangle + \nu_r = x_r, \quad r = 1, \ldots, m, & \text{(15b)} \\ & \mathbf{W} \succeq 0 & \text{(15c)} \end{aligned}$$

where $\mu > 0$ is a fixed parameter. This problem is referred to as the **penalized convex problem**. If the noise parameters admit a zero mean Gaussian distribution with a covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_m^2)$, then $\phi(\boldsymbol{\nu}) = \phi_{\mathrm{WLS}}(\boldsymbol{\nu})$ and $\phi(\boldsymbol{\nu}) = \phi_{\mathrm{WLAV}}(\boldsymbol{\nu})$ lead to the weighted least square (WLS) and weighted least absolute value (WLAV) estimators, where

$$\phi_{\mathrm{WLS}}(\boldsymbol{\nu}) \triangleq \frac{\nu_1^2}{\sigma_1^2} + \cdots + \frac{\nu_m^2}{\sigma_m^2}, \tag{16a}$$

$$\phi_{\mathrm{WLAV}}(\boldsymbol{\nu}) \triangleq \frac{|\nu_1|}{\sigma_1} + \cdots + \frac{|\nu_m|}{\sigma_m} \tag{16b}$$

To solve the state estimation problem under study, we need to address two questions: (i) how to deal with the nonlinearity of the measurement equations, (ii) how to deal with noisy measurements. The terms $\langle \mathbf{W}, \mathbf{M} \rangle$ and $\phi(\boldsymbol{\nu})$ in the objective function of the penalized convex problem (15) aim to handle issues (i) and (ii), respectively. In fact, it can be observed that

- If $\mu = 0$, the objective function (15a) reduces to $\langle \mathbf{W}, \mathbf{M} \rangle$, which may resolve the non-convexity of the quadratic measurement equations by returning a rank-1 solution in the noiseless case, due to Theorem 2.
- If $\mu = +\infty$, the objective function (15a) is equivalent to $\phi(\boldsymbol{\nu})$ (i.e., $\langle \mathbf{W}, \mathbf{M} \rangle$ becomes unimportant). In this case, the resulting objective function aims to estimate the noise values.

A question arises as to whether a finite value for μ could integrate the benefits of the cases $\mu = 0$ and $\mu = +\infty$.

Theorem 3. *Suppose that Assumptions 1 and 2 hold, and that* $m = 2n - 1$. *Consider a function* $\phi(\boldsymbol{\nu}) : \mathbb{R}^m \to \mathbb{R}_+$ *such that*

- $\phi(\mathbf{0}_m) = 0$,
- $\phi(\boldsymbol{\nu}) = \phi(-\boldsymbol{\nu})$,
- $\phi(\boldsymbol{\nu})$ *is continuous, convex, and strictly increasing with respect to all its arguments over the region* $\mathbb{R}^m_+$.

There exists a region $\mathcal{T} \subseteq \mathbb{C}^n$ *containing* $\mathbf{1}_n$ *and its neighborhood such that, for every* $\mathbf{v} \in \mathcal{T}$, *the penalized convex problem* (15) *with the input* $\mathbf{x} = \mathcal{A}(\mathbf{v})$ *has a rank-1 solution, for all finite numbers* $\mu \in \mathbb{R}_+$. *Moreover, this solution is unique if* $\phi(\cdot)$ *is strictly convex.*

Proof. Consider an arbitrary voltage vector $\mathbf{v}$. Let $(\mathbf{W}^{\text{opt}}, \boldsymbol{\nu}^{\text{opt}})$ denote a solution of (15) with the input $\mathbf{x} = \mathcal{A}(\mathbf{v})$. Since $(\mathbf{W}, \boldsymbol{\nu}) = (\mathbf{1}_n \mathbf{1}_n^*, \mathcal{A}(\mathbf{v}) - \mathcal{A}(\mathbf{1}_n))$ is a feasible point, one can write:

$$\langle \mathbf{W}^{\text{opt}}, \mathbf{M} \rangle + \mu \times \phi(\boldsymbol{\nu}^{\text{opt}}) \leq \langle \mathbf{1}_n \mathbf{1}_n^*, \mathbf{M} \rangle + \mu \times \phi(\mathcal{A}(\mathbf{v}) - \mathcal{A}(\mathbf{1}_n)). \quad (17)$$

On the other hand, it follows from the relations $\mathbf{M} \succeq 0$ and $\mathbf{W}^{\text{opt}} \succeq 0$ as well as Assumption 2 that

$$\langle \mathbf{W}^{\text{opt}}, \mathbf{M} \rangle \geq 0 \quad \text{and} \quad \langle \mathbf{1}_n \mathbf{1}_n^*, \mathbf{M} \rangle = 0. \quad (18)$$

Combining (17) and (18) leads to the inequality

$$\phi(\boldsymbol{\nu}^{\text{opt}}) \leq \phi(\mathcal{A}(\mathbf{v}) - \mathcal{A}(\mathbf{1}_n)). \quad (19)$$

On the other hand,

$$\begin{aligned} &\|(\mathcal{A}(\mathbf{v}) - \boldsymbol{\nu}^{\text{opt}}) - \mathcal{A}(\mathbf{1}_n)\| \leq \|\mathcal{A}(\mathbf{v}) - \mathcal{A}(\mathbf{1}_n)\| + \|\boldsymbol{\nu}^{\text{opt}}\| \\ &\leq \|\mathcal{A}(\mathbf{v}) - \mathcal{A}(\mathbf{1}_n)\| + \max\{\|\boldsymbol{\nu}\| \,|\, \phi(\boldsymbol{\nu}) \leq \phi(\boldsymbol{\nu}^{\text{opt}})\} \\ &\leq \|\mathcal{A}(\mathbf{v}) - \mathcal{A}(\mathbf{1}_n)\| + \max\{\|\boldsymbol{\nu}\| \,|\, \phi(\boldsymbol{\nu}) \leq \phi(\mathcal{A}(\mathbf{v}) - \mathcal{A}(\mathbf{1}_n))\}. \end{aligned}$$

Notice that as $\mathbf{v}$ approaches $\mathbf{1}_n$, the right side of the above inequality goes towards zero and hence $\mathcal{A}(\mathbf{v}) - \boldsymbol{\nu}^{\text{opt}}$ becomes arbitrarily close to $\mathcal{A}(\mathbf{1}_n)$. This implies that there exists a region $\mathcal{T} \in \mathbb{C}^n$ containing the point $\mathbf{1}_n$ and its neighborhood such that

$$\mathcal{A}(\mathbf{v}) - \boldsymbol{\nu}^{\text{opt}} \in \text{image}\{\mathcal{R}_{\mathcal{A}}(\mathbf{M})\}, \qquad \forall\ \mathbf{v} \in \mathcal{T} \quad (21)$$

where $\text{image}\{\mathcal{R}_{\mathcal{A}}(\mathbf{M})\}$ denotes the image of the region $\mathcal{R}_{\mathcal{A}}(\mathbf{M})$ under the mapping $\mathcal{A}(\cdot)$. In addition, the penalized convex problem (15) can be written as

$$\begin{aligned} &\underset{\mathbf{W} \in \mathbb{H}^n}{\text{minimize}} && \langle \mathbf{W}, \mathbf{M} \rangle && (22a) \\ &\text{subject to} && \langle \mathbf{W}, \mathbf{M}_r \rangle = \mathcal{A}_r(\mathbf{v}) - \nu_r^{\text{opt}}, \quad r = 1, \ldots, m, && (22b) \\ &&& \mathbf{W} \succeq 0. && (22c) \end{aligned}$$

In other words, $\mathbf{W}^{\text{opt}}$ is a solution of the above problem. Moreover, it follows from (21) and Theorem 2 that $\mathbf{v}(\mu)\mathbf{v}(\mu)^*$ is the only solution of (22) for every $\mathbf{v} \in \mathcal{T}$, where $\mathbf{v}(\mu)$ is a vector satisfying the relation $\mathcal{A}(\mathbf{v}(\mu)) = \mathcal{A}(\mathbf{v}) - \boldsymbol{\nu}^{\text{opt}}$. As a result, the solution of (15) with the input $\mathbf{x} = \mathcal{A}(\mathbf{v})$ is rank-1 for every $\mathbf{v}$ in the region $\mathcal{T}$. Now, it remains to show that $\mathbf{v}(\mu)\mathbf{v}(\mu)^*$ is the only solution of (15). To prove by contradiction, let $(\tilde{\mathbf{W}}^{\text{opt}}, \tilde{\boldsymbol{\nu}}^{\text{opt}})$ denote another solution of (15) with the input $\mathbf{x} = \mathcal{A}(\mathbf{v})$. Due to the strict convexity of $\phi(\cdot)$, the vectors $\boldsymbol{\nu}$ and $\tilde{\boldsymbol{\nu}}$ must be identical. Hence, $\mathbf{W}^{\text{opt}}$ and $\tilde{\mathbf{W}}^{\text{opt}}$ must both be optimal solutions of (22). However, as stated earlier, $\mathbf{v}(\mu)\mathbf{v}(\mu)^*$ is the unique solution of (22) whenever $\mathbf{v} \in \mathcal{T}$. This contradiction completes the proof. □

Theorem 3 considers a large class of $\phi(\cdot)$ functions, including WLS and WLAV. It states that the penalized convex problem (15) associated with the PF problem always returns a rank-1 solution as long as the PF solution $\mathbf{v}$ is sufficiently close to $\mathbf{1}_n$, no matter how small or big the mixing term μ is. A question arises as to whether this rank-1 solution is equal to the matrix $\mathbf{v}\mathbf{v}^*$ being sought. This problem will be addressed below.

Theorem 4. *Suppose that Assumptions 1 and 2 hold. Given an arbitrary vector of voltages* $\mathbf{v} \in \mathcal{R}_{\mathcal{A}}(\mathbf{M}) \backslash \{\mathbf{1}_n\}$, *consider the penalized convex problem* (15) *with the input* $\mathbf{x} = \mathcal{A}(\mathbf{v})$. *The following statements hold:*

i) *If* $\phi(\boldsymbol{\nu}) = \phi_{\text{WLS}}(\boldsymbol{\nu})$ *and* $\mu \in \mathbb{R}_+$, *then* $\mathbf{v}\mathbf{v}^*$ *cannot be a solution of the penalized convex problem.*
ii) *If* $\phi(\boldsymbol{\nu}) = \phi_{\text{WLAV}}(\boldsymbol{\nu})$ *and* μ *is large enough, then* $\mathbf{v}\mathbf{v}^*$ *is a solution of the penalized convex problem.*

Proof. For Part (i), assume that $\phi(\boldsymbol{\nu}) = \phi_{\text{WLS}}(\boldsymbol{\nu})$ and consider the matrix $(1-\varepsilon)\mathbf{v}\mathbf{v}^* + \varepsilon\ \mathbf{1}_n\mathbf{1}_n^*$. Since $\mathbf{v} \neq \mathbf{1}_n$, this matrix is not rank-1. We aim to show that the objective function of the penalized convex problem (15) is smaller at the point $\mathbf{W} = (1-\varepsilon)\mathbf{v}\mathbf{v}^* + \varepsilon\ \mathbf{1}_n\mathbf{1}_n^*$ than the point $\mathbf{W} = \mathbf{v}\mathbf{v}^*$, for a sufficiently small number $\varepsilon \in \mathbb{R}_+$. To this end, notice that the function (15a) evaluated at $\mathbf{W} = \mathbf{v}\mathbf{v}^*$ is equal to

$$\langle \mathbf{W}, \mathbf{M} \rangle + \mu \times \phi(\boldsymbol{\nu}) = \langle \mathbf{v}\mathbf{v}^*, \mathbf{M} \rangle \quad (23)$$

(note that $\boldsymbol{\nu}$ is equal to $\mathbf{0}_m$ in this case). On the other hand, the function (15a) at $\mathbf{W} = (1-\varepsilon)\mathbf{v}\mathbf{v}^* + \varepsilon\ \mathbf{1}_n\mathbf{1}_n^*$ can be calculated as

$$\langle \mathbf{W}, \mathbf{M} \rangle + \mu \times \phi(\boldsymbol{\nu}) = (1-\varepsilon)\langle \mathbf{v}\mathbf{v}^*, \mathbf{M} \rangle + \sum_{r=1}^{m} \frac{\varepsilon^2 \mu}{\sigma_r^2} \left(\langle \mathbf{v}\mathbf{v}^* - \mathbf{1}_n\mathbf{1}_n^*, \mathbf{M}_r \rangle\right)^2 \quad (24)$$

Note that since $\mathbf{v} \neq \mathbf{1}_n$, the term $\langle \mathbf{v}\mathbf{v}^*, \mathbf{M} \rangle$ is strictly positive. Therefore, when ε approaches zero, the first-order term with respect to ε dominates the second-order term and (24) becomes smaller than (23). This completes the proof of Part (i).

Notice that if the constraint $\mathbf{W} \succeq 0$ were missing, Part (ii) would have been an immediate consequence of the exact penalty theorem. We adopt the proof of that theorem given in [23] to prove Part (ii). Assume that $\phi(\boldsymbol{\nu}) = \phi_{\text{WLS}}(\boldsymbol{\nu})$, and let $\rho_1(\mathbf{x})$ and $\rho_2(\mathbf{x})$ denote the optimal objective values of the convex problems (12) and (15) as a function of the input vector $\mathbf{x}$. Due to the convexity of these problems, $\rho_1(\mathbf{x})$ and $\rho_2(\mathbf{x})$ are both convex. Assume for now that $m = 2n-1$. One can write:

$$\rho_2(\mathcal{A}(\mathbf{v})) = \min_{\boldsymbol{\nu}\in\mathbb{R}^m} \left\{ \rho_1(\mathcal{A}(\mathbf{v}) - \boldsymbol{\nu}) + \mu \sum_{r=1}^{m} \frac{|\nu_r|}{\sigma_r} \right\} \tag{25}$$

On the other hand, the Gradient of $\rho_1(\mathcal{A}(\mathbf{v}) - \boldsymbol{\nu})$ at $\boldsymbol{\nu} = \mathbf{0}_m$ is equal to the unique vector $\boldsymbol{\lambda}$ given in (13). For every arbitrary vector $\boldsymbol{\nu}$, it follows from the mean-value theorem that there exists a number $\alpha \in [0,1]$ such that

$$\begin{aligned}\rho_1(\mathcal{A}(\mathbf{v}) - \boldsymbol{\nu}) &= \rho_1(\mathcal{A}(\mathbf{v})) + \boldsymbol{\lambda}^T\boldsymbol{\nu} \\ &+ \frac{1}{2}\boldsymbol{\nu}^T \times \nabla^2 \rho_1(\mathcal{A}(\mathbf{v}) - \alpha\boldsymbol{\nu}) \times \boldsymbol{\nu},\end{aligned} \tag{26}$$

where ∇^2 is the Hessian operator. Therefore, when $\boldsymbol{\nu}$ is sufficiently small, we have

$$\begin{aligned}\rho_1(\mathcal{A}(\mathbf{v}) - \boldsymbol{\nu}) + \mu\sum_{r=1}^{m}\frac{|\nu_r|}{\sigma_r} &= \rho_1(\mathcal{A}(\mathbf{v})) \\ &+ \sum_{r=1}^{m}\left(\frac{\mu}{\sigma_r}|\nu_r| + \lambda_r\nu_r + O(\nu^2)\right)\end{aligned} \tag{27}$$

It can be inferred from the above equation that $\boldsymbol{\nu} = \mathbf{0}_m$ is a local minimum of the function

$$\rho_1(\mathcal{A}(\mathbf{v}) - \boldsymbol{\nu}) + \mu\sum_{r=1}^{m}\frac{|\nu_r|}{\sigma_r} \tag{28}$$

if μ is greater than $\sigma_r|\lambda_r|$ for $r = 1, 2, ..., m$. Note that since $\rho(\cdot)$ is convex, any local minimum is a global solution as well. Now, it follow from (25) that

$$\rho_2(\mathcal{A}(\mathbf{v})) = \rho_1(\mathcal{A}(\mathbf{v}) - \mathbf{0}_m) + \mu\sum_{r=1}^{m}\frac{|0|}{\sigma_r} = \rho_1(\mathcal{A}(\mathbf{v})) \tag{29}$$

This completes the proof for $m = 2n-1$. Now, consider the case $m > 2n-1$. It can be concluded from (25) that

$$\rho_2(\mathcal{A}(\mathbf{v})) \leq \rho_1(\mathcal{A}(\mathbf{v}) - \mathbf{0}_m) + \mu\sum_{r=1}^{m}\frac{|0|}{\sigma_r} = \rho_1(\mathcal{A}(\mathbf{v})) \tag{30}$$

and that

$$\rho_2(\mathcal{A}(\mathbf{v})) \geq \min_{\boldsymbol{\nu}\in\mathbb{R}^m} \left\{ \rho_1(\mathcal{A}(\mathbf{v}) - \boldsymbol{\nu}) + \mu \sum_{r=1}^{2n-1} \frac{|\nu_r|}{\sigma_r} \right\} \tag{31}$$

(note that the sum is taken up to $r = 2n-1$ as opposed to $r = m$). On the other hand, we proved earlier that if μ is large enough, the right side of the above inequality is equal to $\rho_1(\mathcal{A}(\mathbf{v}))$. This implies that $\rho_2(\mathcal{A}(\mathbf{v})) \geq \rho_1(\mathcal{A}(\mathbf{v}))$ in light of (31). Combining this relation with (30) concludes that $\rho_1(\mathcal{A}(\mathbf{v})) = \rho_2(\mathcal{A}(\mathbf{v}))$. □

Corollary 1. *Suppose that Assumptions 1 and 2 hold, and that* $\phi(\boldsymbol{\nu}) = \phi_{\text{WLAV}}(\boldsymbol{\nu})$. *There is a region containing* $\mathbf{1}_n$ *and a neighborhood around this point such that the following statements are satisfied for every* $\mathbf{v}$ *in this region:*

- *The penalized convex problem* (15) *with the input* $\mathbf{x} = \mathcal{A}(\mathbf{v})$ *has a rank-1 solution, for all finite numbers* $\mu \in \mathbb{R}_+$.
- *The penalized convex problem* (15) *with the input* $\mathbf{x} = \mathcal{A}(\mathbf{v})$ *has the unique solution* $\mathbf{v}\mathbf{v}^*$ *and solves the PF problem, for large numbers* $\mu \in \mathbb{R}_+$.

Proof. The proof follows from Theorems 3 and 4. □

Consider the case where the number of measurement (i.e., m) is greater than $2n-1$ and all measurements are noiseless. Let $\mathcal{K} \subseteq \{1, \ldots, m\}$ be a subset of the m measurement equations with only $2n-1$ specifications. According to Theorem 1, the vector $\mathbf{v}$ belongs to the recovery region of the SDP relaxation problem (12) associated with the measurements in $\mathcal{K}$ if the matrix $\mathbf{F}_{\mathbf{B}}(\mathbf{v}, \mathbf{M})$ in (14) is positive semidefinite and its second smallest eigenvalue is strictly positive. In this case, it can be easily verified that the SDP relaxation problem that includes all m measurements (rather than only $2n-1$ specifications) also recovers $\mathbf{v}$. The next theorem generalizes the above result to the noisy case and derives an upper bound on the estimation error in terms of the energy of the noise.

Theorem 5. *Suppose that Assumptions 1 and 2 hold. Consider a vector of voltages* $\mathbf{v} \in \text{int}\{\mathcal{R}_{\mathcal{B}}(\mathbf{M})\} \cap \mathcal{J}_{\mathcal{A}}$, *where*

$$\mathcal{B}(\mathbf{v}) = \left[\langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_{u_1}\rangle, \ldots, \langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_{u_{2n-1}}\rangle\right]^{\mathrm{T}}. \tag{32}$$

and $u_1, \ldots, u_{2n-1} \in \{1, \ldots, m\}$ *correspond to an arbitrary set of* $2n-1$ *linearly independent columns of* $\mathbf{J}_{\mathcal{A}}(\mathbf{v})$. *Let* $(\mathbf{W}^{\text{opt}}, \boldsymbol{\nu}^{\text{opt}})$ *denote an optimal solution of the penalized convex problem* (15) *with the noisy input* $\mathbf{x} = \mathcal{A}(\mathbf{v}) + \boldsymbol{\omega}$ *and* $\phi(\boldsymbol{\nu}) = \phi_{\text{WLAV}}(\boldsymbol{\nu})$. *There exists a scalar* $\alpha > 0$ *such that*

$$\|\mathbf{W}^{\text{opt}} - \alpha\mathbf{v}\mathbf{v}^*\|_F \leq 2\sqrt{\frac{\mu \times \|\boldsymbol{\omega}\|_1 \times \text{trace}\{\mathbf{W}^{\text{opt}}\}}{\eta_{n-1}}}, \tag{33}$$

where η_{n-1} *is the second smallest eigenvalue of* $\mathbf{F}_{\mathcal{B}}(\mathbf{v}, \mathbf{M})$ *defined in* (14).

Proof. The proof developed in [24] can be adopted to prove this theorem. The details are omitted due to space restrictions. □

A. Rank-One Approximation Algorithm

The penalized convex problem (15) could be computationally expensive for large-scale systems because of the high-order conic constraint (15c). One method for tackling this issue is to replace the single conic constraint (15c) with several lower-order conic constraint as follows:

$$\mathbf{W}\{\mathcal{C}_1,\mathcal{C}_1\}\succeq 0,\ \mathbf{W}\{\mathcal{C}_2,\mathcal{C}_2\}\succeq 0,\ \ldots,\ \mathbf{W}\{\mathcal{C}_d,\mathcal{C}_d\}\succeq 0, \tag{34}$$

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 (15c) into several smaller ones. Due to the sparsity and near planarity of power networks, the decomposed penalized convex problem can 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. [25].

In this work, we adopt an alternative approach for recovering the vector of voltages, which does not require calculating the missing entries of $\mathbf{W}$. 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 (7) as follows:

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

$$\underset{\boldsymbol{\theta}\in[-\pi,\pi]^n}{\text{minimize}} \quad \sum_{(i,j)\in\mathcal{L}} \left|\measuredangle W_{ij}^{\text{opt}} - \theta_i + \theta_j\right| \tag{35a}$$

$$\text{subject to} \quad \theta_o = 0, \tag{35b}$$

where $o \in \mathcal{N}$ is the slack 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 $(\mathbf{W}^{\text{opt}}\{\mathcal{C}_1, \mathcal{C}_1\}, \ldots, \mathbf{W}^{\text{opt}}\{\mathcal{C}_d, \mathcal{C}_d\})$.

B. Zero Injection Buses

Real-world power networks have many intermediate buses that do not exchange electrical powers with any external load or generator. In this subsection, we will exploit this feature of power systems to design a number of valid inequalities that can be used to strengthen the convex problems (12) and (15).

Definition 5. *A PQ bus $k \in \mathcal{N}$ is called a zero injection bus if both active and reactive power injections at bus k are equal to zero. Define $\mathcal{Z}$ as the set of all zero injection buses of the network.*

In the PSSE problem, we seek a solution $\mathbf{v}$ whose entries are all nonzero. A zero voltage is regarded as grounding the corresponding bus, which is highly undesirable. This property will be exploited to derive valid inequalities in the next lemma.

Lemma 2. *If $\mathbf{v}$ is a solution to the power flow problem* (11) *with nonzero entries, then the equation*

$$\mathbf{v}\mathbf{v}^*\mathbf{Y}^*\mathbf{e}_k = \mathbf{0}_n \tag{36}$$

holds for every $k \in \mathcal{Z}$.

Proof. Observe that

$$\begin{aligned}
\mathbf{v}\mathbf{v}^*\mathbf{Y}^*\mathbf{e}_k &= (v_k^*)^{-1}(\mathbf{v}\mathbf{v}^*\mathbf{Y}^*\mathbf{e}_k\mathbf{e}_k^*\mathbf{v}) && (37a)\\
&= (v_k^*)^{-1}[\mathbf{v}\mathbf{v}^*(\mathbf{Y}_{p;k} + \mathbf{Y}_{q;k}\mathbb{i})\mathbf{v}] && (37b)\\
&= (v_k^*)^{-1}[\langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{p;k}\rangle + \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{q;k}\rangle\mathbb{i}]\mathbf{v} && (37c)\\
&= (v_k^*)^{-1}(p_k + q_k\mathbb{i})\mathbf{v}, && (37d)
\end{aligned}$$

which concludes (36) since $p_k = q_k = 0$. □

According to Lemma 2, the set of additional constraints

$$\mathbf{W}\mathbf{Y}^*\mathbf{e}_k = \mathbf{0}_n, \qquad k \in \mathcal{Z} \tag{38}$$

can be added to the convex problems (12) and (15) in order to strengthen the relaxations. Notice that each bus $k \in \mathcal{Z}$ has only two power constraints $\langle \mathbf{W}, \mathbf{Y}_{p;k}\rangle = 0$ and $\langle \mathbf{W}, \mathbf{Y}_{q;k}\rangle = 0$. However, equation (38) introduces $2n$ valid scalar constraints for this bus, which would significantly tighten the relaxations. Note that a large number of buses in real-world transmission networks are zero injection buses. As an example, more than one fifth of buses for Polish Grid test systems are zero buses.

IV. Simulation Results

Several papers have shown the superiority of the SDP convex relaxation of the PSSE problem over Newton's method [17]–[22]. That convex relaxation is equivalent to an unpenalized version of (15) by setting $\mathbf{M} = \mathbf{0}$. We have observed in many simulations on IEEE and Polish systems that the penalized convex program with a nonzero matrix M significantly outperforms the SDP convex relaxation of PSSE. Due to space restrictions, we study only the PEGASE 1354-bus system in this paper [26]. Consider a positive number c. Suppose that all measurements are subject to zero mean Gaussian noises, where the standard deviations for squared voltage magnitude, nodal active/reactive power, and branch flow measurements are c, $1.5c$ and $2c$ times the corresponding noiseless values of squared voltage magnitudes, nodal active/reactive powers, and branch flows, respectively.

Let M be equal to $\alpha \times \mathbf{I} - \mathbf{B}$, where the constant α is chosen in such a way that $\alpha \times \mathbf{I} - \mathbf{B}$ satisfies Assumption 2. Consider three scenarios as follows:

- **Scenario 1:** This corresponds to the classical power flow problem, where the measurements are taken at PV and PQ buses. The measurements are then corrupted with Gaussian noise values with $c = 0.01$.
- **Scenario 2:** This is built upon Scenario 1 by taking extra measurements. More precisely, 10% of the line flow parameters (the entries of $\mathbf{p}_f$, $\mathbf{p}_t$, $\mathbf{q}_f$ and $\mathbf{q}_t$) are randomly sampled and added to the measurements used in Scenario 1.
- **Scenario 3:** This is the same as Scenario 2 with the only difference that $c = 0.05$.

We have generated 20 random trials for each scenario and solved the penalized convex program (15) for four objective functions

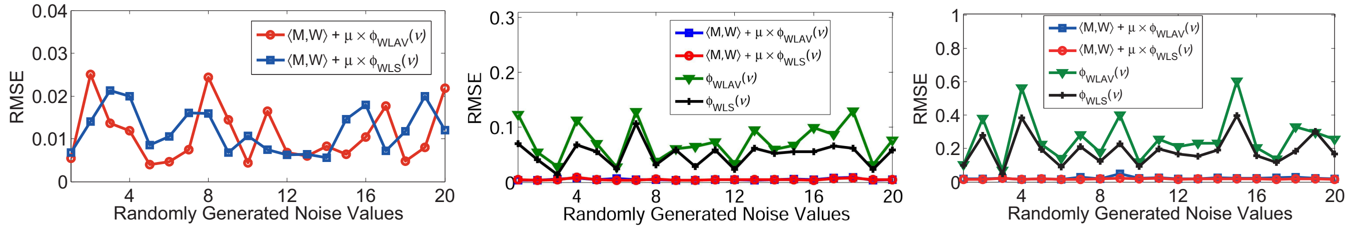
$$f_1(\mathbf{W}, \boldsymbol{\nu}) \triangleq \langle \mathbf{M}, \mathbf{W}\rangle + \mu \times \phi_{\text{WLS}}(\boldsymbol{\nu}), \tag{39a}$$

$$f_2(\mathbf{W}, \boldsymbol{\nu}) \triangleq \langle \mathbf{M}, \mathbf{W}\rangle + \mu \times \phi_{\text{WLAV}}(\boldsymbol{\nu}), \tag{39b}$$

$$f_3(\mathbf{W}, \boldsymbol{\nu}) \triangleq \phi_{\text{WLS}}(\boldsymbol{\nu}), \tag{39c}$$

$$f_4(\mathbf{W}, \boldsymbol{\nu}) \triangleq \phi_{\text{WLAV}}(\boldsymbol{\nu}), \tag{39d}$$

with $\mu = 0.5$. The root mean square errors of the recovered nodal complex voltages are plotted in Figure 1. Note that the curves corresponding to the objective functions f_3 and f_4 are not shown in Figure 1(a) since they are significantly higher than those for the functions f_1 and f_2.



(a) (b) (c)

Fig. 1: These plots compare the accuracy of estimated vector of voltages obtained through minimization of different objective functions for case PEGASE 1354-bus system and 20 randomly generated vector of noise values. In each case power flow measurements are available and (a): $c = 0.01$, (b): additional 10% of line flows are given and $c = 0.01$, and (c): additional 10% of line flows are given and $c = 0.05$.

In order to be able to solve the large-scale problem (15) efficiently, we exploited the sparsity structure of the network. More precisely, through a graph theoretic algorithm from [16], the conic constraint of the penalized convex program was replaced by a set of low-order conic constraints (as discussed in Subsection III-A). In order to preserve the low-complex structure of the problem, only those valid constraints in (38) that did not change the tree decomposition of the underling optimization problem were imposed. The total number of such valid scalar constraints chosen from (38) and incorporated in (15) is equal to 1436.

V. Conclusions

This paper aims to find a convex model for the power system state estimation (PSSE) problem. PSSE is central to the operation of power systems, and has a high computational complexity due to the nonlinearity of power flow equations. In this work, we develop a family of penalized convex problems to solve the PSSE problem. It is shown that each convex program proposed in this paper finds the correct solution of the PSSE problem in the case of noiseless measurements, provided that the voltage angles are relatively small. In presence of noisy measurements, it is proven that the penalized convex problems are all able to find an approximate solution of the PSSE problem, where the estimation error has an explicit upper bound in terms of the energy of the noise. The objective function of each penalized convex problem has two terms: one accounting for the non-convexity of the power flow equations and another one for estimating the noise level. Simulation results elucidate the superiority of the proposed method.

References

[1] 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.

[2] W. F. Tinney and C. E. Hart, "Power flow solution by Newton's method," *IEEE T. Power Ap. Syst.*, vol. 86, no. Nov., pp. 1449–1460, Jun. 1967.

[3] B. Stott and O. Alsaç, "Fast decoupled load flow," *IEEE T. Power Ap. Syst.*, vol. 93, no. 3, pp. 859–869, May. 1974.

[4] R. A. Van Amerongen, "A general-purpose version of the fast decoupled load flow," *IEEE Trans. Power Syst.*, vol. 4, no. 2, pp. 760–770, May. 1989.

[5] 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.

[6] S. Sojoudi and J. Lavaei, "Physics of power networks makes hard optimization problems easy to solve," in *IEEE Power and Energy Society General Meeting*, 2012.

[7] J. Lavaei, "Zero duality gap for classical OPF problem convexifies fundamental nonlinear power problems," in *Proc. Amer. Control Conf.*, June 2011, pp. 4566–4573.

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

[9] S. H. Low, "Convex relaxation of optimal power flow–part I: Formulations and equivalence," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, March 2014.

[10] S. H. Low, "Convex relaxation of optimal power flow–part II: Exactness," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 2, pp. 177–189, June 2014.

[11] A. Lam, B. Zhang, and D. Tse, "Distributed algorithms for optimal power flow problem," in *Proc. IEEE Conf. Decision and Control*, 2012.

[12] B. Zhang, A. Lam, A. Dominguez-Garcia, and D. Tse, "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Trans. Power Syst.*, 2012.

[13] D. K. Molzahn, J. T. Holzer, B. C. Lesieutre, and C. L. DeMarco, "Implementation of a large-scale optimal power flow solver based on semidefinite programming," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3987–3998, Nov. 2013.

[14] M. S. Andersen, A. Hansson, and L. Vandenberghe, "Reduced-complexity semidefinite relaxations of optimal power flow problems," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1855–1863, Jul 2014.

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

[16] R. Madani, M. Ashraphijuo, and J. Lavaei, "Promises of conic relaxation for contingency-constrained optimal power flow problem," to appear in *IEEE T. Power Syst.*, 2015, http://www.ieor.berkeley.edu/∼lavaei/SCOPF_2014.pdf.

[17] 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.

[18] 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.

[19] 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.

[20] 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)*. IEEE, 2013, pp. 1–6.

[21] 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.

[22] R. Madani, J. Lavaei, and R. Baldick, "Convexification of power flow problem over arbitrary networks," in *2015 54th IEEE Conference on Decision and Control (CDC)*, Dec 2015, pp. 1–8.

[23] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[24] Y. Zhang, R. Madani, and J. Lavaei, "Power system state estimation with line measurements," [Online]. Available: http://www.ieor.berkeley.edu/∼lavaei/SE_Line_2016.pdf.

[25] 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 *Conference on Decision and Control*, 2014.

[26] S. Fliscounakis, P. Panciatici, F. Capitanescu, and L. Wehenkel, "Contingency ranking with respect to overloads in very large power systems taking into account uncertainty, preventive, and corrective actions," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4909–4917, 2013.