Convexification of Power Flow Equations in the Presence of Noisy Measurements

Ramtin Madani, Javad Lavaei and Ross Baldick

***Abstract*—This paper is concerned with the power system state estimation (PSSE) problem that aims to find the unknown operating point of a power network based on a given set of measurements. We first study the power flow (PF) problem as an important special case of PSSE. PF is known to be non-convex and NP-hard in the worst case. To this end, we propose a set of semidefinite programs (SDPs) with the property that they all solve the PF problem as long as the voltage angles are relatively small. Associated with each SDP, we explicitly characterize the set of all complex voltages that can be recovered via that convex problem. As a generalization, the design of an SDP problem that recovers multiple nominal points and a neighborhood around each point is also cast as a convex program. The results are then extended to the PSSE problem, where the measurements used in the PF problem are subject to noise. A two-term objective function is employed for each convex program developed for the PSSE problem: (i) one term accounting for the non-convexity of the power flow equations, (ii) another one for estimating the noise levels. An upper bound on the estimation error is derived with respect to the noise level and the proposed techniques are demonstrated on multiple test systems, including a 9241-bus European network. Although the focus of the paper is on power networks, the developed results apply to every arbitrary state estimation problem with quadratic measurement equations.**

I. Introduction

Consider a group of generators (i.e., sources of energy), which are connected to a group of electrical loads (i.e., consumers) via an electrical power network. This network comprises a set of transmission lines connecting various nodes to each other (e.g., a generator to a load). At each node of the network, the associated external devices (loads and generators) exchange a net complex electrical power with the network, where the real and imaginary parts of this complex power are called active and reactive powers. The active power is a measure of the short-term average power delivered to loads that is able to do useful work, whereas the reactive power is a measure of the back-and-forth flow of power in the network between electric and magnetic fields. The nodal complex powers induce active and reactive powers over all lines of the network, which are referred to as line flows. The complex flow entering a line could differ from the flow leaving the line at the other end since transmission lines are lossy in practice. On the other hand, each node of the network (called a bus) is associated with a complex number, named nodal voltage. To describe the relationship among all nodal powers and line flows, it is possible to write each of these parameters as a quadratic function of nodal voltages using *power flow equations*. The set of all nodal complex voltages of a power network is referred to as the state of the system. The power flow (PF) problem involves solving power flow equations in order to find the state of the system, given a set of noiseless measurements. These measurements are usually a subset of voltage magnitudes, nodal active powers, nodal reactive powers, active line flows and reactive line flows.

Ramtin Madani is with the Department of Electrical Engineering, University of Texas at Arlington (ramtin.madani@uta.edu). Javad Lavaei is with the Department of Industrial Engineering and Operations Research, University of California, Berkeley, and also with the Tsinghua-Berkeley Shenzhen Institute (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 DARPA YFA, ONR YIP Award, DARPA Young Faculty Award, AFOSR YIP Award, NSF CAREER Award 1351279, NSF EECS Awards 1406894 and 1406865, and the University of Texas System STARs program. Parts of this work have appeared in the conference papers *"Ramtin Madani, Javad Lavaei and Ross Baldick, Convexification of power flow problem over arbitrary networks, IEEE Conference on Decision and Control, 2015"*, and *"Ramtin Madani, Morteza Ashraphiojuo, Javad Lavaei and Ross Baldick, Power system state estimation with a limited number of measurements, IEEE Conference on Decision and Control, 2016"*. The authors would like to thank the anonymous reviewers for their constructive and invaluable comments.

A. Power Flow Problem

Power Flow equations are central to the analysis and operation of power systems, based upon which several optimization problems are built. These problems include optimal power flow, state estimation, security-constrained optimal power flow, unit commitment, network reconfiguration, and transmission switching [1]–[4]. However, it is well-known that solving these equations, namely the PF problem, is NP-hard for both transmission and distribution networks due to its reduction to the subset sum problem [5], [6]. The recent paper [7] shows the strong NP-hardness of this problem for certain types of systems. The nonlinearity of the power flow equations is imposed by the laws of physics, and is a major impediment to the efficient, optimal and reliable operation of power systems.

Since 1962, several linearization and local search algorithms have been developed for solving power flow equations, and the current practice in the power industry relies on linearization and/or Newton's method (depending on the time scale and whether the problem is solved for planning or real-time operation) [8]–[10]. Traditional methods based on the linearization of power flow equations do not typically capture important quantities such as voltage magnitudes, thermal losses and reactive flows, which make these techniques less appealing for applications such as voltage control. As a result, great effort has been devoted to developing modified linear programming models that incorporate reactive power and voltage magnitudes [11], [12]. Another major focus of the existing literature for solving power flow equations has been on homotopy continuation methods that are widely applied to the PF problem [13]–[15]. As argued in [16], earlier homotopy-based techniques lack performance guarantees, have scalability issues, and may not always find all solutions of the PF problem. Nevertheless, new homotopy algorithms have been recently introduced in [17] and [18] that are provably capable of finding all feasible solutions and attempt to ameliorate the issue of scalability through parallelization. Other approaches for tackling the nonlinearity of power flow equations include Gröebner basis techniques and interval based methods [19]–[21]. The recent papers [22] and [23] propose upper bounds on the number of solutions for the PF problem based on the network topology.

In this paper, we exploit the semidefinite programming relaxation technique to handle the non-convexity of the PF problem. Semidefinite programming (SDP) is a subfield of convex optimization, which has received a considerable amount of attention in the past two decades [24]–[27]. The SDP relaxation technique is a powerful method for tackling quadratic nonlinearities, which has been proven to be effective in the convexification of several hard optimization problems in various areas, including graph theory, approximation theory, quantum mechanics, neural networks, communication networks, and power systems [28]–[33]. SDP relaxation methods have been successfully used for real-world applications such as radar code design, multiple-input and multiple-output (MIMO) beamforming, error-correcting codes, magnetic resonance imaging (MRI), data training, and portfolio selection, among many others [34]–[37]. Several papers have evaluated the performance of SDP relaxations for various problems, by investigating the approximation ratio and the maximum rank of SDP solutions [38]–[42]. Moreover, different global optimization techniques for polynomial optimization have been built upon SDP relaxations [32], [43]–[46].

Our first contribution is related to the PF problem. To handle the non-convexity of this feasibility problem, we propose a class of convex optimization programs in the form of SDPs. We derive an exact recovery region for each convex program in the proposed SDP class. This means that the solution of the PF problem can be found using SDP if and only if it belongs to the associated recovery region. Moreover, we prove that if the voltage angles are small enough, the classical PF problem can be solved precisely using the proposed SDP problems, and this result does not make any assumption on the network topology whatsoever. Note that voltage angles are often small in practice due to practical considerations, which has two 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-complexity approximate models that can provide insights into power systems, whereas Newton's method is an attractive numerical algorithm that has been used in practice for many years [47], [48]. Some of the advantages 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 proposed 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 in general, but the recovery region of the proposed SDP problem is explicitly characterizable via matrix inequalities [49].
- The SDP relaxation provides a convex model for the PF problem, independent of what numerical algorithm will be used to solve it.
- As will be verified on benchmark systems later in this paper, the proposed SDP relaxation has a higher success rate than Newton's method.

Our approach relies on converting the feasibility PF problem into a convex program 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 problem. The designed objective function is not unique and there are infinitely many choices for this function. The question arises as to whether any of these objective functions added to the PF problem would have a physical meaning. To address this problem, we show that one such function is the sum of squares of the nodal current magnitudes, which indirectly accounts for the loss in the network. Note that each objective function produces its own recovery region and therefore it is not always beneficial to use a physically meaningful objective function instead of a synthetically designed function. In this paper, we also study the problem of selecting the best objective function.

B. Power Systems State Estimation

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 [50]. Notice that PSSE is built upon the PF problem by replacing noiseless specifications with noisy measurements, and therefore it is also a nonconvex problem. In order to adapt the proposed approach for the PF problem to PSSE, we adopt a penalized convex relaxation approach similar to [51], 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) the one previously used for the PF problem in the noiseless case to deal with non-convexity, (ii) another term added to account for the noisy measurements.

We show that the penalized convex problem 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 order to assess the accuracy of the proposed estimation framework in the noisy case, we offer a bound on the estimation error with respect to the noise level. In this case, the SDP matrix solution may or may not be rank-1 due to corrupted measurements. We employ the algorithm introduced in [41] to estimate the solution of the PSSE problem from the penalized SDP solution, and demonstrate the efficacy of the proposed technique on multiple test systems, including a network with more than 9000 nodes.

C. Related Work

Started by the papers [52] and [53], the SDP relaxation technique has received a significant attention in the power and optimization societies. The work [53] develops an SDP relaxation for finding a global solution of the optimal power flow (OPF) problem, and shows that the relaxation is exact for IEEE test systems. The follow-up papers [54] and [29] prove that the success of the SDP relaxation in handling the non-convexity of the power flow equations is due in part to the passivity of transmission lines, and moreover the relaxation finds a global solution under two assumptions: (i) load over-satisfaction (by modeling loads as inequality constraints rather than equality constraints), (ii) the presence of a phase shifting transformer in every basic cycle of the network.

The papers [55] and [56] introduce branch-and-bound techniques to obtain feasible solutions for the case where the SDP relaxation is not exact. In order to improve the performance of the SDP relaxation, several valid inequalities and bound tightening techniques have been proposed in [57] and [58]. The work [59] identifies certain classes of mesh power networks for which the SDP relaxation finds a global solution of the OPF problem without using any transformers. The paper [42] shows that the graphs of real-world power networks often have a low treewidth and, as a result, the proposed SDP would possess a low-rank solution. The more recent paper [60] designs a linear program to find an ε-approximation of the solution, where the size of linear program is exponential in the treewidth of the network. In other words, [42] and [60] relate the complexity of the power equations to the treewidth of power networks. In the case where the SDP relaxation fails to work, a graph-theoretic penalized SDP framework has been developed in [59] and [41]. This method identifies the problematic lines of the network (sources of non-convexity) through a graph analysis and then penalizes the loss over those lines in the objective of the SDP relaxation in order to find a near-global solution of the OPF problem. The proposed approach is successful in finding near-global solutions with global optimality guarantees of at least 99% for IEEE and Polish test systems. Inspired by the general technique proposed in [61] and [62], recent advances in leveraging the sparsity of power systems have made SDP problems computationally more tractable [41], [63]–[67]. The paper [68] develops a computationally efficient second-order cone programming (SOCP) relaxation scheme for the OPF problem, whose performance is empirically verified to be close to the SDP relaxation. Further extensions of the above-mentioned SDP relaxation to the AC transmission switching and security-constrained OPF problem have been made in [69] and [41]. The reader is referred to [70] and [71] for a detailed survey of this topic.

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 underlying dimension of the unknown state of the system (i.e., twice the number of buses minus one). The papers [72] and [73] 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 and Lagrange decomposition have been developed. The work [74] 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 this work, our primary focus is mainly on the hard case where the number of measurements is on the same order as the number of unknown parameters. In order to approach the measurement noise, in the present work, we incorporate WLAV estimator into the objective along with a penalty term that promotes rank-1 solutions. We will develop theoretical results and show through simulations that the proposed convexification approach outperforms the WLS and WLAV estimators given in [73]–[75].

D. 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. The symbol $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. 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 Frobenius 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. 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, $\mathbf{W}[\mathcal{X}]$ 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

In this section, we offer some preliminary results on the power flow equations.

A. 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 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 each of its ends, respectively. 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

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

where $\mathbf{i} \triangleq [i_1, i_2 \dots, i_n]^{\mathrm{T}}$ is the vector of complex nodal current injections, and $\mathbf{i}_{\mathrm{f}} \triangleq [i_{\mathrm{f},1}, i_{\mathrm{f},2} \dots, i_{\mathrm{f},|\mathcal{L}|}]^{\mathrm{T}}$and $\mathbf{i}_{\mathrm{t}} \triangleq [i_{\mathrm{t},2}, i_{\mathrm{t},2} \dots, i_{\mathrm{t},|\mathcal{L}|}]^{\mathrm{T}}$ are the vectors of currents entering 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 the following 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 implies that $\mathbf{Y} \times \mathbf{1}_n = \mathbf{0}_n$.

Note that $\mathbf{Y}$ acts as the Laplacian of a weighted graph obtained from the power network where the weight of each edge is equal to the complex admittance (or inverse impedance) of the corresponding branch of the system. Let $\mathbf{Y} = \mathbf{G} + \mathbf{B}\mathbb{i}$, where $\mathbf{G}$ and $\mathbf{B}$ are the conductance and susceptance matrices, respectively. Before proceeding with the main results of this work, we derive a fundamental property of the matrix $\mathbf{B}$ in the next lemma.

Lemma 1. *For every $\mathcal{N}' \subseteq \mathcal{N}$, the relation $\mathbf{B}[\mathcal{N}', \mathcal{N}'] \preceq 0$ holds. Moreover, $\mathbf{B}[\mathcal{N}', \mathcal{N}']$ is singular if and only if $\mathcal{N}' = \mathcal{N}$.*

Proof. Please refer to Section V for the proof. □

B. 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 $\mathbf{p} + \mathbb{i}\mathbf{q} = \mathrm{diag}\{\mathbf{v} \times \mathbf{i}^*\}$. In addition, 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:

$$s_{\mathrm{f};l} = v_k \times i^*_{\mathrm{f};l} \;, \qquad s_{\mathrm{t};l} = v_j \times i^*_{\mathrm{t};l} \;. \tag{2}$$

Define $\mathbf{E}_k \triangleq \mathbf{e}_k \mathbf{e}_k^*$, and

$$\mathbf{Y}_{\mathrm{p};k} \triangleq \frac{\mathbf{Y}^*\mathbf{E}_k + \mathbf{E}_k\mathbf{Y}}{2}, \qquad \mathbf{Y}_{\mathrm{q};k} \triangleq \frac{\mathbf{Y}^*\mathbf{E}_k - \mathbf{E}_k\mathbf{Y}}{2\mathbb{i}}, \tag{3a}$$

for every $k \in \mathcal{N}$, 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 $\mathbf{E}_k$, $\mathbf{Y}_{\mathrm{p};k}$ and $\mathbf{Y}_{\mathrm{q};k}$:

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

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 k to node j, define

$$\mathbf{Y}_{\mathrm{p_f};l} \triangleq \frac{\mathbf{Y}_{\mathrm{f}}^* \mathbf{d}_l \mathbf{e}_k^* + \mathbf{e}_k \mathbf{d}_l^* \mathbf{Y}_{\mathrm{f}}}{2}, \; \mathbf{Y}_{\mathrm{q_f};l} \triangleq \frac{\mathbf{Y}_{\mathrm{f}}^* \mathbf{d}_l \mathbf{e}_k^* - \mathbf{e}_k \mathbf{d}_l^* \mathbf{Y}_{\mathrm{f}}}{2\mathbb{i}}, \tag{5a}$$

$$\mathbf{Y}_{\mathrm{p_t};l} \triangleq \frac{\mathbf{Y}_{\mathrm{t}}^* \mathbf{d}_l \mathbf{e}_j^* + \mathbf{e}_j \mathbf{d}_l^* \mathbf{Y}_{\mathrm{t}}}{2}, \; \mathbf{Y}_{\mathrm{q_t};l} \triangleq \frac{\mathbf{Y}_{\mathrm{t}}^* \mathbf{d}_l \mathbf{e}_j^* - \mathbf{e}_j \mathbf{d}_l^* \mathbf{Y}_{\mathrm{t}}}{2\mathbb{i}}. \tag{5b}$$

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

$$p_{\mathrm{f};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{p_f};l} \rangle, \quad q_{\mathrm{f};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{q_f};l} \rangle, \tag{6a}$$

$$p_{\mathrm{t};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{p_t};l} \rangle, \quad q_{\mathrm{t};l} = \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\mathrm{q_t};l} \rangle, \tag{6b}$$

for every $l \in \mathcal{L}$. Equations (4) and (6) 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

$$x_r = \langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_r \rangle + \omega_r, \qquad \forall r \in \mathcal{M}, \tag{7}$$

where

- $\mathcal{M} = \{1, 2, \dots, m\}$ is a set of indices associated with the available measurements (or specifications).
- $x_1, \dots, x_m$ are the known measurement values.
- $\omega_1, \dots, \omega_m$ are the unknown measurement noises, for which some *a priori* statistical information may be available.
- $\mathbf{M}_1, \dots, \mathbf{M}_m$ are some known $n \times n$ Hermitian matrices (e.g., they could be any subset of the matrices defined in (3) and (5)).

Several algorithms in different contexts, such as signal processing, have been proposed in the literature for solving a system of quadratic equation in the form of (7) [51], [76]–[80]. In this work, we aim to propose a convex relaxation scheme with strong theoretical guarantees, which is tailored to power system applications. In the case where the noises $\omega_1, \dots, \omega_m$ are all equal to zero, the above problem 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 a simple phase shift, we assume that $\measuredangle v_\rho$ is equal to zero at a pre-selected bus ρ, named the reference bus. Hence, the state estimation problem with the complex variable $\mathbf{v}$ amounts to $2n - 1$ real variables.

C. 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 defining $\mathbf{W} \triangleq \mathbf{v}\mathbf{v}^*$, the quadratic equations in (7) can be formulated linearly in terms of $\mathbf{W}$ as follows:

$$x_r = \langle \mathbf{W}, \mathbf{M}_r \rangle + \omega_r, \qquad \forall r \in \mathcal{M}. \tag{8}$$

Consider the case where the quadratic measurements $x_1, \dots, x_{|\mathcal{M}|}$ are noiseless. Solving the non-convex equations in (7) is tantamount to finding a rank-1 and positive semidefinite matrix $\mathbf{W} \in \mathbb{H}^n$ satisfying the above linear equations for $\omega_1 = \cdots = \omega_{|\mathcal{M}|} = 0$ (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 (8) is regarded as a convex *relaxation* of (7) since it includes no restriction on the rank of $\mathbf{W}$.

Although the set of equations (7) normally has a finite number of solutions whenever $|\mathcal{M}| \geq 2n-1$, its SDP relaxation (8) may have infinitely many solutions because the matrix variable $\mathbf{W}$ includes $O(n^2)$ scalar variables as opposed to $2n-1$. However, under some additional assumptions, it is known that the relaxed problem has a unique solution in some applications such as phase retrieval if $|\mathcal{M}| \geq 3n$, and this solution automatically has rank-1 [80]. In the case where the relaxed problem does not have a unique solution, the literature of compressed sensing substantiates that minimizing $\text{trace}\{\mathbf{W}\}$ over the feasible set of (8) may yield a low-rank matrix $\mathbf{W}$ under strong technical assumptions [42], [80]–[82]. One main objective of this paper is to study what objective function should be minimized (instead of $\text{trace}\{\mathbf{W}\}$) to attain a rank-1 solution for the relaxed problem (8) in the noiseless case. Another objective is to generalize the results to noisy measurements.

D. 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

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

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:

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

Definition 1. *Given an index set of measurements $\mathcal{M} = \{1, 2, ..., m\}$, define the function $\boldsymbol{f}_{\mathcal{M}}(\overline{\mathbf{v}}) : \mathbb{R}^{2n-1} \to \mathbb{R}^m$ 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:*

$$\boldsymbol{f}_{\mathcal{M}}(\overline{\mathbf{v}}) \triangleq [\langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_1\rangle, \langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_1\rangle, \ldots, \langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_m\rangle]^{\text{T}}.$$

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

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

(note that $\overline{\mathbf{M}}_i$ can be obtained from $\mathbf{M}_i$ via the equation (10) *for $i = 1, ..., m$).*

According to the inverse function theorem, if $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})$ has full row rank and $|\mathcal{M}| = 2n-1$, 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 assumption that $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})$ has full row rank, 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 later show that the full rank property of $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})$ is beneficial not only for Newton's method but also for the SDP relaxation.

Definition 2. *Given an index set of measurements $\mathcal{M}$, define $\mathcal{J}_{\mathcal{M}} \subseteq \mathbb{C}^n$ as the set of all voltage vectors $\mathbf{v}$ for which $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})$ has full row rank. A vector of complex voltages $\mathbf{v}$ is said to be observable through the set of measurements $\mathcal{M}$ if it belongs to $\mathcal{J}_{\mathcal{M}}$.*

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 $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{1}}_n)$ has full row rank.

Assumption 1. *The vector of voltages $\mathbf{1}_n$ is observable (i.e. $\mathbf{1}_n \in \mathcal{J}_{\mathcal{M}}$).*

We will later show that the above assumption holds for the set of measurements corresponding to the classical power flow problem.

E. Classical Power Flow Problem

The power flow (PF) problem can be defined as the noiseless state estimation problem, i.e., by assuming that $\omega_1 \ldots, \omega_{|\mathcal{M}|}$ 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 $|\mathcal{M}|$) is equal to $2n-1$, the measurements are all made at buses as opposed to lines, and there is no measurement noise. To formulate the problem, three basic types of buses are considered, depending on the parameters that are known at each bus:

- PQ bus: p_k and q_k are specified.
- PV bus: p_k and $|v_k|$ are specified.
- The reference bus: $|v_\rho|$ is specified.

Each PQ bus represents a load bus or possibly a generator bus, whereas each PV bus represents a generator bus. It is also assumed that there is a unique reference 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.

Define $\mathcal{P}$, $\mathcal{Q}$ and $\mathcal{V}$ as the sets of buses for which active powers, reactive powers and voltage magnitudes are known, respectively. Assume that $\mathcal{V} \neq \varnothing$, and let $\mathcal{P}$ and $\mathcal{Q}$ be strict subsets of $\mathcal{N}$. The classical PF problem can be formalized as

$$\begin{aligned} &\text{find} && \mathbf{v} \in \mathbb{C}^n \\ &\text{subject to} && \langle \mathbf{v}\mathbf{v}^*, \mathbf{E}_k\rangle = |v_k|^2, && \forall k \in \mathcal{V} && \text{(11a)} \\ &&& \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\text{q};k}\rangle = q_k, && \forall k \in \mathcal{Q} && \text{(11b)} \\ &&& \langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}_{\text{p};k}\rangle = p_k, && \forall k \in \mathcal{P}. && \text{(11c)} \end{aligned}$$

The problem (11) is in the canonical form (7) after noticing that

- $\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_m$ correspond to the m matrices $\mathbf{E}_k$ ($\forall k \in \mathcal{V}$), $\mathbf{Y}_{\text{q};k}$ ($\forall k \in \mathcal{Q}$), and $\mathbf{Y}_{\text{p};k}$ ($\forall k \in \mathcal{P}$) defined in (3).
- The specifications $x_1, x_2, \ldots, x_m$ correspond to $|v_k|^2$'s, q_k's, and p_k's.
- The measurement noise values $\omega_1, \omega_2, \ldots, \omega_m$ are all equal to zero.

Define $\mathcal{M}_{\text{CPF}}$ as the set of measurements corresponding to the classical power flow problem.

Lemma 2. *If $\mathcal{V} \neq \varnothing$ and $\mathcal{P}$ and $\mathcal{Q}$ are strict subsets of $\mathcal{N}$, then Assumption 1 holds for the classical power flow problem, i.e., $\mathbf{1}_n \in \mathcal{J}_{\mathcal{M}_{\text{CPF}}}$.*

Proof. Please refer to Section V for the proof. □

Remark 1. *It is straightforward to verify that Lemma 2 holds true for every arbitrary power network with shunt elements as long as the matrix* $\mathbf{Y}$ *is generic. In other words, if* $\mathbf{J}_{\mathcal{M}_{\text{CPF}}}(\overline{\mathbf{1}}_n)$ *is singular for a power network possessing shunt elements, an infinitesimal perturbation of the susceptance values of the existing lines makes the resulting matrix* $\mathbf{J}_{\mathcal{M}_{\text{CPF}}}(\overline{\mathbf{1}}_n)$ *non-singular.*

III. Convexification of Power Flow Problem

In this section, assume that the available measurements provided in (7) are all noiseless:

$$x_r = \langle \mathbf{v}\mathbf{v}^*, \mathbf{M}_r \rangle, \qquad \forall r \in \mathcal{M}. \tag{12}$$

To solve this set of quadratic equations through a convex relaxation, consider a family of convex programs of the form

$$\begin{aligned}
&\underset{\mathbf{W}\in\mathbb{H}^n}{\text{minimize}} && \langle \mathbf{W}, \mathbf{M} \rangle && && (13a)\\
&\text{subject to} && \langle \mathbf{W}, \mathbf{M}_r \rangle = x_r, && \forall r \in \mathcal{M}, && (13b)\\
& && \mathbf{W} \succeq 0, && && (13c)
\end{aligned}$$

where the matrix $\mathbf{M} \in \mathbb{H}_+^n$ is to be designed. As an example, the SDP program (13) associated with the classical PF problem can be written as

$$\begin{aligned}
&\underset{\mathbf{W}\in\mathbb{H}^n}{\text{minimize}} && \langle \mathbf{W}, \mathbf{M} \rangle && && (14a)\\
&\text{subject to} && \langle \mathbf{W}, \mathbf{E}_k \rangle = |v_k|^2, && \forall k \in \mathcal{V} && (14b)\\
& && \langle \mathbf{W}, \mathbf{Y}_{\text{q};k} \rangle = q_k, && \forall k \in \mathcal{Q} && (14c)\\
& && \langle \mathbf{W}, \mathbf{Y}_{\text{p};k} \rangle = p_k, && \forall k \in \mathcal{P} && (14d)\\
& && \mathbf{W} \succeq 0. && && (14e)
\end{aligned}$$

We aim for systematically designing $\mathbf{M}$ such that the above problem yields a unique rank-1 solution $\mathbf{W}$, from which a feasible solution $\mathbf{v}$ can be recovered for (12). Notice that the existence of such a rank-1 solution depends in part on its input specifications $\mathbf{x} = [x_1, x_2, \ldots, x_{|\mathcal{M}|}]^{\text{T}}$. It is said that the SDP problem (13) solves the set of equations (12) for the input $\mathbf{x} = [x_1, x_2, \ldots, x_{|\mathcal{M}|}]^{\text{T}}$ if (13) has a unique rank-1 solution.

Definition 3. *Given an index set of measurements* $\mathcal{M}$ *and an objective matrix* $\mathbf{M} \in \mathbb{H}^n$*, 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* (13) *for some input vector* $\mathbf{x} \in \mathbb{R}^{|\mathcal{M}|}$*. Define* $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$ *as the set of all recoverable vectors of voltages.*

Note that the set $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$ is indeed the collection of all possible operating points $\mathbf{v}$ that can be found through (13) given a set of consistent specification values $\mathbf{x} = [x_1, x_2, \ldots, x_{|\mathcal{M}|}]^{\text{T}}$. More precisely, a vector of voltages $\mathbf{v}$ belongs to $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$ if it can be recovered by solving the convex problem (13) with the input $\mathbf{x} = \boldsymbol{f}_{\mathcal{M}}(\bar{\mathbf{v}})$. It is desirable to find out whether there exists a matrix $\mathbf{M}$ for which the recoverable region $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$ covers a large subset of $\mathbb{C}^n$ that contains practical solutions of power flow problems. Addressing this problem is central to this section. 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 following properties:*

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

Remark 2. *Assumption 2 is satisfied for the two sample choices* $\mathbf{M} = -\mathbf{B}$ *and* $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$ *considered in this paper. If there are shunt elements, then the diagonal elements of* $-\mathbf{B}$ *and* $\mathbf{Y}^*\mathbf{Y}$ *may need to be modified to satisfy these conditions. Since shunt elements cannot be ignored for the PEGASE and IEEE systems, we have used the matrix* $\mathbf{M} = \alpha\mathbf{I} - \mathbf{B}$ *in the simulations section, where* α *is the smallest real number such that* $\alpha\mathbf{I} - \mathbf{B} \succeq 0$*. Note that this choice of* $\mathbf{M}$ *does not necessarily satisfy the third condition of Assumption 2.*

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

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

Proof. Please refer to Section V for the proof. □

Theorem 1 states that if $\mathbf{1}_n$ serves as a valid point for the linearization of the power flow equations (i.e. $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{1}}_n)$ has full row rank), then as long as the specifications $x_1, \ldots, x_{|\mathcal{M}|}$ correspond to a vector of voltages with small angles, the exact recovery of the solution of the PF problem is guaranteed through the proposed SDP problem. Note that this result does not require any assumption on the network topology whatsoever. This implies that although the widely-used DC (linearized) model of power flow equations can be used to find an approximate solution around the nominal point, the SDP relaxation is an exact convex model of the PF problem (leading to a solution with an arbitrary precision).

Notice that according to Lemma 2, Assumption 1 is automatically satisfied for the classical PF problem. In addition, if $\mathbf{M}$ is chosen as $\mathbf{Y}^*\mathbf{Y}$, this matrix satisfies Assumption 2 due to the equation $\mathbf{Y} \times \mathbf{1}_n = \mathbf{0}_n$. In this case, the objective of the convex problem (14) corresponds to $|i_1|^2 + |i_2|^2 + \cdots + |i_n|^2$, where i_k denotes the net current at bus k for $k = 1, ..., n$. Therefore, Theorem 1 implies that as long as the voltage angles are small enough, a solution of the feasibility PF problem can be recovered exactly by means of an SDP relaxation whose objective function reflects the minimization of nodal currents. In the case where the PF problem has multiple solutions, the one found using the SDP relaxation is likely the most practical (desirable) solution since it indirectly corresponds to the minimum loss (or voltage drop) in the network.

A. Region of Recoverable Voltages

Given an arbitrary matrix $\mathbf{M} \in \mathbb{H}^n$, the objective is to characterize $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$, i.e., the set of all voltage vectors that can be recovered using the convex problem (13). To this end, it is useful to analyze the vector of Lagrange multipliers $\boldsymbol{\lambda} \in \mathbb{R}^{|\mathcal{M}|}$ associated with the constraints in (13b). Some of the results to be presented in the remainder of the paper do not require Assumption 1 and/or Assumption 2. Whenever any of these assumptions is needed, it will be explicitly mentioned in the statement of the corresponding theorem/proposition/corollary.

Definition 4. *Given an index set of measurements $\mathcal{M}$, a vector $\boldsymbol{\lambda} \in \mathbb{R}^{|\mathcal{M}|}$ is regarded as a dual certificate for the voltage vector $\mathbf{v} \in \mathbb{C}^n$ if it satisfies the two properties:*

$$\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})\boldsymbol{\lambda} = -2\overline{\mathbf{M}}\overline{\mathbf{v}}, \qquad (15a)$$

$$\kappa_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}) > 0, \qquad (15b)$$

where $\kappa_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}) : \mathbb{H}^n \times \mathbb{R}^{|\mathcal{M}|} \to \mathbb{R}$ is called observability factor and defined as the sum of the two smallest eigenvalues of the Hermitian matrix $\mathbf{M} + \sum_{r \in \mathcal{M}} \lambda_r \mathbf{M}_r$. Denote the set of all dual certificates for the voltage vector $\mathbf{v}$ as $\mathcal{D}_{\mathcal{M}}(\mathbf{M}, \mathbf{v})$.

Note that since the sum of the two smallest eigenvalues of a Hermitian matrix variable is a concave function of that matrix, $\kappa_{\mathcal{M}}(\cdot, \cdot)$ is a concave function. Moreover, $\kappa_{\mathcal{M}}(\cdot, \cdot)$ takes both positive and negative values depending on the signs of the eigenvalues of its matrix argument.

Proposition 1. *Consider an arbitrary vector of voltages $\mathbf{v} \in \mathcal{J}_{\mathcal{M}}$. If there exists a dual certificate $\boldsymbol{\lambda} \in \mathcal{D}_{\mathcal{M}}(\mathbf{M}, \mathbf{v})$, then $\mathbf{v}$ belongs to the interior of $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$.*

Proof. Please refer to Section V for the proof. □

The above proposition offers a nonlinear matrix inequality formulation 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 $\boldsymbol{f}_{\mathcal{M}}(\overline{\mathbf{v}})$ loses rank (note that the conditions in (15) can be cast as bilinear matrix inequalities). This proposition is particularly interesting in the special case where the number of equations is equal to the number of unknown parameters. In that case, there exists a unique vector $\boldsymbol{\lambda} = -2\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^{-1}\overline{\mathbf{M}}\,\overline{\mathbf{v}}$ that satisfies (15a).

Proposition 2. *If $|\mathcal{M}| = 2n - 1$, then the interior of the set $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$ can be characterized as*

$$\text{int}\{\mathcal{R}_{\mathcal{M}}(\mathbf{M})\} \cap \mathcal{J}_{\mathcal{M}} = \{\mathbf{v} \in \mathcal{J}_{\mathcal{M}} \,|\, \tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) > 0\},$$

where the function $\tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) : \mathbb{H}^n \times \mathbb{C}^n \to \mathbb{R}$ is defined as sum of the two smallest eigenvalues of the matrix $\mathbf{M} + \sum_{r \in \mathcal{M}} \lambda_r \mathbf{M}_r$, and λ_r denotes the r^{th} entry of the vector $-2\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^{-1}\overline{\mathbf{M}}\,\overline{\mathbf{v}}$.

Proof. Please refer to Section V for the proof. □

If the matrix $\mathbf{M}$ satisfies Assumption 2, then $\tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v})$ serves as a measure for the closeness of the true solution $\mathbf{v}$ to the nominal point $\mathbf{1}_n$. In particular, the vector $\boldsymbol{\lambda} = -2\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^{-1}\overline{\mathbf{M}}\,\overline{\mathbf{v}}$ in the case $|\mathcal{M}| = 2n-1$ can be arbitrarily close to $\mathbf{0}_n$ if $\|\mathbf{v} - \mathbf{1}_n\|_2$ is sufficiently small, which implies that

$$\mathbf{M} + \sum_{r \in \mathcal{M}} \lambda_r \mathbf{M}_r \simeq \mathbf{M} \;\Rightarrow\; \tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) \simeq \tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{1}_n) > 0,$$

where the last inequality is due to Assumption 2.

Multiple illustrative examples are given in [83] to show the recovery region associated with a simple three-bus network.

B. Adjustment of Recoverable Region

Theorem 1 states that if Assumptions 1 and 2 hold, then it is possible to recover voltage vectors that belong to a vicinity of the nominal point $\mathbf{1}_n$. However, there are cases for which it is desirable to find a matrix $\mathbf{M}$ such that the corresponding recoverable set $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$ contains a set of neighborhoods around multiple nominal points rather than the single generic vector $\mathbf{1}_n$. One such case is the dynamic state estimation where the previous operating points and historical data could be used to find the next operating point. The following theorem aims to show that, given a set of pre-specified nominal points $\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_2, \ldots, \widehat{\mathbf{v}}_s$, the problem of designing a matrix $\mathbf{M}$ for which $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$ contains all of nominal voltage vectors and a neighborhood around each of these points can be cast as a convex program.

Theorem 2. *Assume that $|\mathcal{M}| = 2n - 1$. Given an arbitrary natural number s and arbitrary points $\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_2, \ldots, \widehat{\mathbf{v}}_s \in \mathcal{J}_{\mathcal{M}}$, consider the problem*

$$\text{find} \quad \mathbf{M} \in \mathbb{H}^n \qquad (16a)$$

$$\text{subject to} \quad \tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \widehat{\mathbf{v}}_k) \geq \varepsilon, \quad k = 1, 2, \ldots, s, \qquad (16b)$$

where $\varepsilon > 0$ is the desired minimum observability factor. The following statements hold:

i) *The feasibility problem* (16) *is convex.*
ii) *There exists a matrix $\mathbf{M}$ such that the associated recoverable set $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$ contains $\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_2, \ldots, \widehat{\mathbf{v}}_s$ with the observability factor at least ε as well as a ball around each of these points if and only if the convex problem* (16) *has a feasible solution $\mathbf{M}$.*

Proof. Please refer to Section V for the proof. □

Theorem 2 is particularly helpful when the state estimation problem is solved in real-time. In that case, rough predictions of the state are often available, which can be used for the design of the matrix $\mathbf{M}$ via (16). Note that if $|\mathcal{M}| > 2n - 1$, then the above theorem should be used for a subset of the measurement index $\mathcal{M}$ with $2n - 1$ measurements (more details will be provided in the next section).

Remark 3. *If $|\mathcal{M}|$ is large enough to offer a level of restricted isometry [84], then the feasible set of the relaxed problem* (13) *reduces to a single point, and every vector becomes recoverable independent of the choice of the objective function [80]. Hence, the complexity of the problem reduces when the number of measurements is high. However, Theorems 1 and 2 focus on more challenging instances of* (12) *where the number of equations is equal to the number of unknowns (i.e., $|\mathcal{M}| = 2n - 1$). In this case, due to computational complexity boundaries [5], it is useful to have some prior information about the unknown vector of voltages.*

IV. Convexification of State Estimation Problem

In the presence of measurement noises, the convex problem (13) may be infeasible (if $|\mathcal{M}| > 2n - 1$) or result in a poor approximate solution. To remedy this issue, a standard approach is to estimate the noise values through some auxiliary variables $\nu_1, \ldots, \nu_{|\mathcal{M}|} \in \mathbb{R}$. This can be achieved by incorporating a convex regularization term $\phi_{\mathcal{M}} : \mathbb{R}^{|\mathcal{M}|} \to \mathbb{R}$ into the objective function that elevates the likelihood of the estimated noise:

$$\begin{aligned} \underset{\substack{\mathbf{W}\in\mathbb{H}^n \\ \boldsymbol{\nu}\in\mathbb{R}^{|\mathcal{M}|}}}{\text{minimize}} \quad & \langle \mathbf{W}, \mathbf{M}\rangle + \mu \times \phi_{\mathcal{M}}(\boldsymbol{\nu}) && \text{(17a)}\\ \text{subject to} \quad & \langle \mathbf{W}, \mathbf{M}_r\rangle + \nu_r = x_r, \qquad \forall r \in \mathcal{M} && \text{(17b)}\\ & \mathbf{W} \succeq 0, && \text{(17c)} \end{aligned}$$

where $\mu > 0$ is a fixed parameter, [51]. We refer to the above convex program as the *penalized convex problem*. If the noise parameters of the measurement values in $\mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}$ admit a zero mean Gaussian distribution with a covariance matrix $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_{|\mathcal{M}|}^2)$, then $\phi_{\mathcal{M}}(\boldsymbol{\nu}) = \phi_{\mathcal{M};\text{WLS}}(\boldsymbol{\nu})$ and $\phi_{\mathcal{M}}(\boldsymbol{\nu}) = \phi_{\mathcal{M};\text{WLAV}}(\boldsymbol{\nu})$ lead to the weighted least square (WLS) and weighted least absolute value (WLAV) estimators, where

$$\phi_{\mathcal{M};\text{WLS}}(\boldsymbol{\nu}) \triangleq \sum_{r\in\mathcal{M}} \frac{\nu_r^2}{\sigma_r^2}, \quad \phi_{\mathcal{M};\text{WLAV}}(\boldsymbol{\nu}) \triangleq \sum_{r\in\mathcal{M}} \frac{|\nu_r|}{\sigma_r}. \quad (18)$$

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 compensate for the noisy measurements. The terms $\langle \mathbf{W}, \mathbf{M}\rangle$ and $\phi_{\mathcal{M}}(\boldsymbol{\nu})$ in the objective function of the penalized convex problem (17) aim to handle issues (i) and (ii), respectively. A question arises as to whether a properly chosen value for μ could resolve the non-convexity of the quadratic measurement equations and estimate the noise values as well.

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

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

Then, there exists a region $\mathcal{T} \subseteq \mathbb{C}^n$ containing $\mathbf{1}_n$ and a neighborhood around this point such that, for every $\mathbf{v} \in \mathcal{T}$, the penalized convex problem (17) *with the input $\mathbf{x} = \boldsymbol{f}_{\mathcal{M}}(\overline{\mathbf{v}})$ has a rank-1 solution, for every $\mu \in \mathbb{R}_+$. Moreover, this solution is unique if $\phi(\cdot)$ is strictly convex.*

Proof. Please refer to Section V for the proof. □

Proposition 3 considers a large class of $\phi(\cdot)$ functions, including WLS and WLAV. It states that the penalized convex problem (17) 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.

Proposition 4. *Suppose that Assumptions 1 and 2 hold. Given an arbitrary vector of voltages $\mathbf{v} \in \mathcal{R}_{\mathcal{M}}(\mathbf{M})\backslash\{\mathbf{1}_n\}$, consider the penalized convex problem* (17) *with the noiseless input $\mathbf{x} = \boldsymbol{f}_{\mathcal{M}}(\overline{\mathbf{v}})$. The following statements hold:*

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

Proof. Please refer to Section V for the proof. □

Consider the case where the number of measurement (i.e., $|\mathcal{M}|$) is greater than $2n-1$ and all measurements are noiseless. Let $\mathcal{M}' \subseteq \mathcal{M}$ be a subset of the available measurement where $|\mathcal{M}'| = 2n - 1$ specifications. According to Proposition 2, the vector $\mathbf{v}$ belongs to the recovery region of the SDP relaxation problem (13) associated with the measurements in $\mathcal{M}'$ if $\tilde{\kappa}_{\mathcal{M}'}(\mathbf{M}, \mathbf{v})$ is positive. In this case, it can be easily verified that the SDP relaxation problem that includes all measurements in $\mathcal{M}$ (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 noise level (namely $\phi_{\mathcal{M}}(\boldsymbol{\omega})$).

Theorem 3. *Consider an arbitrary index set of measurements $\mathcal{M}$ and a vector of voltages $\mathbf{v} \in \mathcal{J}_{\mathcal{M}}$. Suppose that Assumptions 1 and 2 hold. Let $\boldsymbol{\lambda} \in \mathcal{D}_{\mathcal{M}}(\mathbf{M}, \mathbf{v})$ be an arbitrary dual certificate for $\mathbf{v}$ (see Definition 4), and $(\mathbf{W}^{\text{opt}}, \boldsymbol{\nu}^{\text{opt}})$ be an optimal solution of the penalized convex problem* (17) *with the noisy input $\mathbf{x} = \boldsymbol{f}_{\mathcal{M}}(\mathbf{v}) + \boldsymbol{\omega}$. There exists a scalar $\alpha > 0$ such that*

$$\|\mathbf{W}^{\text{opt}} - \alpha \mathbf{v}\mathbf{v}^*\|_F \leq 2\sqrt{\frac{\mu \times \phi_{\mathcal{M}}(\boldsymbol{\omega}) \times \text{trace}\{\mathbf{W}^{\text{opt}}\}}{\kappa_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})}} \quad (19)$$

if $\phi_{\mathcal{M}}(\boldsymbol{\nu})$ is chosen as the WLAV penalty function and the constant μ is selected appropriately to satisfy the relation

$$\mu \geq \max_{r\in\mathcal{M}} |\sigma_r \lambda_r|. \quad (20)$$

Proof. Please refer to Section V for the proof. □

Theorem 3 relates an arbitrary solution of the penalized convex problem to the unknown state $\mathbf{v}$. It shows that the estimation error depends on the noise level $\phi_{\mathcal{M}}(\boldsymbol{\omega})$ and the observability factor $\kappa_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})$. In particular, the error is zero in the noiseless case. To minimize an upper bound on the estimation error, this theorem suggests selecting μ as small as possible in such a way that the inequality (20) turns into an equality. This corresponds to the worst-case scenario and does not use any detailed information about the statistics of the noise. However, given a particular noise realization, one may use an adaptive technique to solve the penalized convex problem (17) for different values of μ to find a suboptimal value. This can be carried out using cross-validation together with the Akaike information criterion or the Bayesian information criterion, as discussed in [85] for the classic Lasso problem.

According to Definition 4, the vector

$$\boldsymbol{\lambda}^+ \triangleq -2\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^+ \overline{\mathbf{M}}\overline{\mathbf{v}} \quad (21)$$

serves as a candidate for dual certificate, and its norm can be bounded as follows:

$$\|\boldsymbol{\lambda}^+\|_2 \leq 2 \times \|\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^+\|_2 \times \|\overline{\mathbf{M}}\overline{\mathbf{v}}\|_2. \quad (22)$$

Hence, $\|\boldsymbol{\lambda}^+\|_2$ is bounded by two factors: (i) the matrix norm $\|\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^+\|_2$ that is related to the number of measurements, and (ii) the vector norm $\|\overline{\mathbf{M}}\overline{\mathbf{v}}\|_2$ that measures the distance between the true solution $\mathbf{v}$ and the initial guess $\mathbf{1}_n$. Therefore, as the numbers $\|\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^+\|_2$ and $\|\overline{\mathbf{M}}\overline{\mathbf{v}}\|_2$ decrease, the observability factor $\kappa_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}^{+})$ approaches the value $\kappa_{\mathcal{M}}(\mathbf{M}, \mathbf{0}_{|\mathcal{M}|})$, which favors the accuracy of estimation due to the bound (19). Observe that $\|\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^{+}\|_2 = \alpha_{\min}^{-1}$, where $\alpha_{\min}$ is the minimum singular value of $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})$, which can tend to infinity as the number of measurements increases [86]. In light of the error bound (19), we have studied a special case in [87] and proved that if the measurements include voltage magnitudes at all buses as well as line flows at a spanning subgraph of the system, then the tail probability of the estimation error decays exponentially with respect to the number measurements.

Corollary 1. *Suppose that Assumptions 1 and 2 hold, and that $\phi_{\mathcal{M}}(\boldsymbol{\nu}) = \phi_{\mathcal{M};\mathrm{WLAV}}(\boldsymbol{\nu})$. 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* (17) *with the input $\mathbf{x} = \boldsymbol{f}_{\mathcal{M}}(\overline{\mathbf{v}})$ has a rank-1 solution, for every $\mu \in \mathbb{R}_{+}$.*
- *The penalized convex problem* (17) *with the input $\mathbf{x} = \boldsymbol{f}_{\mathcal{M}}(\overline{\mathbf{v}})$ has the unique solution $\mathbf{v}\mathbf{v}^{*}$ and solves the PF problem, for sufficiently large values $\mu \in \mathbb{R}_{+}$.*

Proof. The proof follows from Propositions 3 and 4. The uniqueness of the solution using the WLAV penalty is due to Theorem 3. □

Corollary 1 relies on the WLAV penalty as opposed to WLS. The underlying reason is that the exact penalty method for the conversion of constrained optimization problems to unconstrained optimization problems is known to work with WLAV but not with WLS [88].

V. Proofs

In this section, we will prove the main results of this paper. To this end, it is useful to derive the dual of (13). The Lagrangian function for the primal SDP problem (13) can be expressed as follows:

$$L(\mathbf{W}, \boldsymbol{\lambda}) \triangleq \langle \mathbf{M} + \sum_{r \in \mathcal{M}} \lambda_r \mathbf{M}_r - \mathbf{H}, \mathbf{W} \rangle - \mathbf{x}^{\mathrm{T}} \boldsymbol{\lambda} \tag{23}$$

where $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_m]^{\mathrm{T}}$ denotes the measurement values, and the dual variables $\mathbf{H} \in \mathbb{S}_{+}^{n}$ and $\boldsymbol{\lambda} \in \mathbb{R}^{|\mathcal{M}|}$ are the Lagrange multipliers associated with the constraints (13c) and (13b), respectively. The Lagrangian function $L(\cdot, \boldsymbol{\lambda})$ has a finite minimum if and only if

$$\mathbf{H} = \mathbf{M} + \sum_{r \in \mathcal{M}} \lambda_r \mathbf{M}_r, \tag{24}$$

As a result, the dual problem can be stated as

$$\underset{\boldsymbol{\lambda} \in \mathbb{R}^{|\mathcal{M}|}}{\text{minimize}} \quad \mathbf{x}^{\mathrm{T}} \boldsymbol{\lambda} \tag{25a}$$

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

Definition 5. *Define the matrix function $\mathbf{H}_{\mathcal{M}}(\cdot, \cdot) : \mathbb{H}^n \times \mathbb{R}^{|\mathcal{M}|} \to \mathbb{H}^n$ as*

$$\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}) \triangleq \mathbf{M} + \sum_{r \in \mathcal{M}} \lambda_r \mathbf{M}_r. \tag{26}$$

It can be easily observed that the condition (15a) in Definition 4 is satisfied for $\boldsymbol{\lambda} \in \mathbb{R}^{|\mathcal{M}|}$ if and only if $\mathbf{v}$ belongs to the null space of $\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})$, i.e.,

$$\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}}) \boldsymbol{\lambda} = -2\overline{\mathbf{M}}\overline{\mathbf{v}} \iff \mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}) \mathbf{v} = \mathbf{0}_n. \tag{27}$$

Definition 6. *Given an index set of measurements $\mathcal{M}$ with $|\mathcal{M}| = 2n - 1$, define*

$$\boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) \triangleq -2\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^{-1} \overline{\mathbf{M}}\, \overline{\mathbf{v}}. \tag{28}$$

Note that $\boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v})$ is the unique member of $\mathbb{R}^{|\mathcal{M}|}$ satisfying (15a).

Proof of Lemma 1: Since $\mathbf{B}$ is symmetric, the relation $\mathbf{B} \times \mathbf{1}_n = \mathbf{0}_n$ holds according to the equation $\mathbf{Y} \times \mathbf{1}_n = \mathbf{0}_n$. Moreover, every off-diagonal entry of $\mathbf{B}$ is nonnegative due to the assumption that the inductance of each line is nonnegative. Now, one can write:

$$-B_{kk} = \sum_{j \neq k} B_{kj} \implies -B_{kk} \geq \sum_{j \neq k} |B_{kj}|, \tag{29}$$

for every $k \in \{1, \ldots, n\}$. Therefore $-\mathbf{B}$ is diagonally dominant and positive semidefinite. As a result, the relation $\mathbf{B}[\mathcal{N}', \mathcal{N}'] \preceq 0$ holds for every principal submatrix of $\mathbf{B}$. Since the network is connected by assumption and every entry of $\mathbf{B}$ corresponding to an existing line of the network is positive, it follows from the weighted matrix-tree theorem (see [89]) that if $|\mathcal{N}'| = n - 1$, then $\det\{\mathbf{B}[\mathcal{N}', \mathcal{N}']\} \neq 0$ and subsequently $\mathbf{B}[\mathcal{N}', \mathcal{N}'] \prec 0$ due to the relation $\mathbf{B}[\mathcal{N}', \mathcal{N}'] \preceq 0$. Now, consider the case $|\mathcal{N}'| < n - 1$. There exists a set $\mathcal{N}'' \subset \mathcal{N}$ such that $\mathcal{N}' \subset \mathcal{N}''$ and $|\mathcal{N}''| = n - 1$. Due to the Cauchy interlacing theorem, every eigenvalue of $\mathbf{B}[\mathcal{N}', \mathcal{N}']$ is less than or equal to the largest eigenvalue of $\mathbf{B}[\mathcal{N}'', \mathcal{N}'']$, which implies that $\mathbf{B}[\mathcal{N}', \mathcal{N}']$ is non-singular. □

Proof of Lemma 2: For the classical PF problem, it is straightforward to verify that

$$\mathbf{J}_{\mathcal{M}_{\mathrm{CPF}}}(\overline{\mathbf{1}}_n) = \begin{bmatrix} 2\mathbf{I}_{n \times n}[\mathcal{N}, \mathcal{V}] & \mathbf{B}[\mathcal{N}, \mathcal{Q}] & \mathbf{G}[\mathcal{N}, \mathcal{P}] \\ \mathbf{0}_{(n-1) \times |\mathcal{V}|} & -\mathbf{G}[\mathcal{P}, \mathcal{Q}] & \mathbf{B}[\mathcal{P}, \mathcal{P}] \end{bmatrix}.$$

By Gaussian elimination, $\mathbf{J}_{\mathcal{M}_{\mathrm{CPF}}}(\overline{\mathbf{1}}_n)$ reduces to the matrix

$$\mathbf{S} \triangleq \begin{bmatrix} \mathbf{B}[\mathcal{Q}, \mathcal{Q}] & \mathbf{G}[\mathcal{Q}, \mathcal{P}] \\ -\mathbf{G}[\mathcal{P}, \mathcal{Q}] & \mathbf{B}[\mathcal{P}, \mathcal{P}] \end{bmatrix}.$$

Hence, it suffices to prove that $\mathbf{S}$ is not singular. To this end, one can write, $\det\{\mathbf{S}\} = \det\{\mathbf{S}_1\} \times \det\{\mathbf{S}_2\}$, where $\mathbf{S}_1 \triangleq \mathbf{B}[\mathcal{P}, \mathcal{P}]$ is non-singular and $\mathbf{S}_2$ is the Schur complement of $\mathbf{S}_1$ in $\mathbf{S}$, i.e.,

$$\mathbf{S}_2 \triangleq \mathbf{B}[\mathcal{Q}, \mathcal{Q}] + \mathbf{G}[\mathcal{Q}, \mathcal{P}] \mathbf{B}[\mathcal{P}, \mathcal{P}]^{-1} \mathbf{G}[\mathcal{P}, \mathcal{Q}].$$

On the other hand, $\mathbf{S}_1$ and $\mathbf{S}_2$ are both symmetric, and in addition Lemma 1 yields that $\mathbf{S}_1 \prec 0$ and $\mathbf{B}[\mathcal{Q}, \mathcal{Q}] \prec 0$. This implies that $\mathbf{S}_2 \prec 0$ according to the above equation, which leads to the relation $\det\{\mathbf{S}\} \neq 0$. □

Lemma 3. *Suppose that there exists a vector $\mathbf{u} \in \mathcal{J}_{\mathcal{M}}$ such that $\mathcal{D}_{\mathcal{M}}(\mathbf{M}, \mathbf{u}) \neq \varnothing$. Strong duality holds between the primal SDP* (13) *and the dual SDP* (25)*, for every $\mathbf{x} \in \mathbb{R}^{|\mathcal{M}|}$.*

Proof. Let $\boldsymbol{\lambda} \in \mathcal{D}_{\mathcal{M}}(\mathbf{M}, \mathbf{u})$. The assumption $\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{u}})\boldsymbol{\lambda} = -2\overline{\mathbf{M}}\overline{\mathbf{u}}$ implies that $\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})\mathbf{u} = \mathbf{0}_n$. In addition, 0 is a simple eigenvalue of $\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})$ due to the assumption $\kappa_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}) > 0$. In order to show the strong duality, it suffices to build a strictly feasible point $\tilde{\boldsymbol{\lambda}}$ for the dual problem. Let ρ represent the reference bus of the power system. With no loss of generality, we assume that $\mathrm{Im}\{u_\rho\} = 0$. The assumption $\mathbf{u} \in \mathcal{J}_{\mathcal{M}}$ implies that $\overline{\mathbf{u}}^{\mathrm{T}}[\overline{\mathbf{M}}_1\overline{\mathbf{u}} \quad \overline{\mathbf{M}}_2\overline{\mathbf{u}} \quad \ldots \quad \overline{\mathbf{M}}_{|\mathcal{M}|}\overline{\mathbf{u}}] \neq 0$. Therefore, the relation $\mathbf{u}^*\mathbf{M}_r\mathbf{u} = \overline{\mathbf{u}}^{\mathrm{T}}\overline{\mathbf{M}}_r\overline{\mathbf{u}} \neq 0$ holds for at least one index $r \in \mathcal{M}$. Let $\mathbf{d}_1, \ldots, \mathbf{d}_{|\mathcal{M}|}$ be the standard basis vectors for $\mathbb{R}^{|\mathcal{M}|}$. We select $\tilde{\boldsymbol{\lambda}}$ as $\boldsymbol{\lambda} + c \times \mathbf{d}_r$, where $c \in \mathbb{R}$ is a nonzero number with an arbitrarily small absolute value such that $c \times \mathbf{u}^*\mathbf{M}_r\mathbf{u} > 0$. Then, one can write:

$$\tilde{\mathbf{H}} \triangleq \mathbf{M} + \sum_{r\in\mathcal{M}} \tilde{\lambda}_r\mathbf{M}_r = \mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}) + c\,\mathbf{M}_r. \tag{30}$$

Next, we will argue that if c is sufficiently small, then $\tilde{\mathbf{H}} \succ 0$. Let $\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})$ admit the eigenvalue decomposition $[\mathbf{U}\ \mathbf{u}]\ \mathrm{diag}\{[\boldsymbol{\kappa}\ 0]\}\ [\mathbf{U}\ \mathbf{u}]^*$, where $\boldsymbol{\kappa} \in \mathbb{R}^{|\mathcal{V}|-1}$ is the vector of positive eigenvalues. It yields that

$$\tilde{\mathbf{H}} = \begin{bmatrix}\mathbf{U} & \mathbf{u}\end{bmatrix}\begin{bmatrix}\mathrm{diag}\{\boldsymbol{\kappa}\}+c\ \mathbf{U}^*\mathbf{M}_r\mathbf{U} & c\ \mathbf{U}^*\mathbf{M}_r\mathbf{u}\\ c\ \mathbf{u}^*\mathbf{M}_r\mathbf{U} & c\ \mathbf{u}^*\mathbf{M}_r\mathbf{u}\end{bmatrix}\begin{bmatrix}\mathbf{U} & \mathbf{u}\end{bmatrix}^*.$$

Since $\mathrm{diag}\{\boldsymbol{\kappa}\} \succ 0$, if c is small enough, then $\mathrm{diag}\{\boldsymbol{\kappa}\} + c\ \mathbf{U}^*\mathbf{M}_r\mathbf{U} \succ 0$. Therefore, due to Schur complement, the relation $\tilde{\mathbf{H}} \succ 0$ holds if and only if

$$c\ \mathbf{u}^*\mathbf{M}_r\mathbf{u} > c^2\mathbf{u}^*\mathbf{M}_r\mathbf{U}(\mathrm{diag}\{\boldsymbol{\kappa}\}+c\ \mathbf{U}^*\mathbf{M}_r\mathbf{U})^{-1}\mathbf{U}^*\mathbf{M}_r\mathbf{u},$$

which is satisfied when the absolute value of c is sufficiently small. This completes the proof. □

Lemma 4. *Suppose that strong duality holds between the primal SDP* (13) *and the dual SDP* (25)*. Let* $\mathbf{v} \in \mathcal{J}_{\mathcal{M}}$ *be an optimal solution to the power flow problem* (12) *and* $\boldsymbol{\lambda} \in \mathbb{R}^m$ *be a feasible point for the dual SDP* (25)*. The following two statements are equivalent:*

i) $(\mathbf{v}\mathbf{v}^*, \boldsymbol{\lambda})$ *is a pair of primal and dual optimal solutions for the primal SDP* (13) *and the dual SDP* (25)*,*

ii) $\mathbf{v} \in \mathrm{null}\{\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})\}$.

Proof. $(i) \Rightarrow (ii)$: Due to the complementary slackness, one can write

$$\begin{aligned}0 = \langle \mathbf{v}\mathbf{v}^*, \mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})\rangle &= \mathrm{trace}\{\mathbf{v}\mathbf{v}^*\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})\}\\ &= \mathbf{v}^*\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})\mathbf{v}.\end{aligned} \tag{31}$$

On the other hand, it follows from the dual feasibility that $\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}) \succeq 0$, which together with (31) implies that $\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})\mathbf{v} = 0$.

$(ii) \Rightarrow (i)$: Observe that due to the linearity of the Lagrangian function $L(\cdot, \boldsymbol{\lambda})$ in (23), it suffices to verify complementary slackness in order to prove the optimality of a feasible primal-dual pair. Since $\mathbf{v} \in \mathcal{J}_{\mathcal{M}}$ is a solution of (12), the matrix $\mathbf{v}\mathbf{v}^*$ is a feasible point for (13). On the other hand, since $\mathbf{v} \in \mathrm{null}\{\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})\}$, we have $\langle \mathbf{v}\mathbf{v}^*, \mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda})\rangle = 0$, which certifies the optimality of the pair $(\mathbf{v}\mathbf{v}^*, \boldsymbol{\lambda})$. □

Proof of Proposition 2: First, we show that $\{\mathbf{v} \in \mathcal{J}_{\mathcal{M}} \mid \tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) > 0\}$ is an open set for $m = 2n - 1$. To this end, consider a vector $\mathbf{v} \in \mathcal{J}_{\mathcal{M}}$ such that $\tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) > 0$ and let δ denote the second smallest eigenvalue of $\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}))$. Due to the continuity of the functions $\det\{\mathbf{J}_{\mathcal{M}}(\cdot)\}$, $\boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \cdot)$ and $\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \cdot)$, there exists a neighborhood $\mathcal{B} \in \mathbb{C}^n$ around $\mathbf{v}$ such that $\mathbf{v}' \in \mathcal{J}_{\mathcal{M}}$ and

$$\|\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}')) - \mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}))\|_F < \sqrt{\delta}$$

for every $\mathbf{v}'$ within this neighborhood. Now, through an eigenvalue perturbation argument (see Lemma 5 in [83]), we have $\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}')) \succeq 0$ and $\mathrm{rank}\{\mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}'))\} = n - 1$, which implies that $\tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}') > 0$ for every $\mathbf{v}' \in \mathcal{B}$. This proves that $\{\mathbf{v} \in \mathcal{J}_{\mathcal{M}} \mid \tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) > 0\}$ is an open set.

Now, consider a vector $\mathbf{v} \in \mathcal{J}_{\mathcal{M}}$ such that $\tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) > 0$. The objective is to show that $\mathbf{v} \in \mathrm{int}\{\mathcal{R}_{\mathcal{M}}(\mathbf{M})\}$. Notice that since $\tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) > 0$, we have $\mathbf{H} \succeq 0$, where $\mathbf{H} \triangleq \mathbf{H}_{\mathcal{M}}(\mathbf{M}, \boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}))$. This means that the vector $\boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v})$ is a feasible point for the dual problem (25). Therefore, it follows from Lemmas 3 and 4 that the matrix $\mathbf{v}\mathbf{v}^*$ is an optimal solution for the primal problem (13). In addition, every optimal solution $\mathbf{W}^{\mathrm{opt}}$ must satisfy

$$\langle \mathbf{H}, \mathbf{W}\rangle = 0. \tag{32}$$

According to Lemma 4, $\mathbf{v}$ is an eigenvector of $\mathbf{H}$ corresponding to the eigenvalue 0. Therefore, since $\mathbf{H} \succeq 0$ and $\mathrm{rank}\{\mathbf{H}\} = n - 1$, every positive semidefinite matrix $\mathbf{W}^{\mathrm{opt}}$ satisfying (32) is equal to $c \times \mathbf{v}\mathbf{v}^*$ for a nonnegative constant c. This concludes that $\mathbf{v}\mathbf{v}^*$ is the unique solution to (13), and therefore $\mathbf{v}$ belongs to $\mathcal{R}_{\mathcal{M}}(\mathbf{M})$. Since $\{\mathbf{v} \in \mathcal{J}_{\mathcal{M}} \mid \tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) > 0\}$ is shown to be an open set, the above result can be translated as $\{\mathbf{v} \in \mathcal{J}_{\mathcal{M}} \mid \tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) > 0\} \subseteq \mathrm{int}\{\mathcal{R}_{\mathcal{M}}(\mathbf{M})\} \cap \mathcal{J}_{\mathcal{M}}$.

In order to complete the proof, it is requited to show that $\mathrm{int}\{\mathcal{R}_{\mathcal{M}}(\mathbf{M})\} \cap \mathcal{J}_{\mathcal{M}}$ is a subset of $\{\mathbf{v} \in \mathcal{J}_{\mathcal{M}} \mid \tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) > 0\}$. To this end, consider a vector $\mathbf{v} \in \mathrm{int}\{\mathcal{R}_{\mathcal{M}}(\mathbf{M})\} \cap \mathcal{J}_{\mathcal{M}}$. This means that $\mathbf{v}\mathbf{v}^*$ is a solution to (13), and therefore $\mathbf{H} \succeq 0$ due to Lemma 4, and as a result $\tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) \geq 0$. To prove the aforementioned inclusion by contradiction, suppose that $\tilde{\kappa}_{\mathcal{M}}(\mathbf{M}, \mathbf{v}) = 0$, implying that 0 is an eigenvalue of $\mathbf{H}$ with multiplicity at least 2. Let $\hat{\mathbf{v}}$ denote a second eigenvector corresponding to the eigenvalue 0 such that $\mathbf{v}^*\hat{\mathbf{v}} = 0$. Since $\mathbf{v} \in \mathcal{J}_{\mathcal{M}}$, it results from the inverse function theorem that there exists a constant $\varepsilon_0 > 0$ with the property that for every $\varepsilon \in (0, \varepsilon_0]$, there is a vector $\mathbf{w}_\varepsilon \in \mathbb{C}^n$ satisfying the relation $\boldsymbol{f}_{\mathcal{M}}(\mathbf{w}_\varepsilon) = \boldsymbol{f}_{\mathcal{M}}(\mathbf{v}) + \varepsilon\,\boldsymbol{f}_{\mathcal{M}}(\hat{\mathbf{v}})$. This means that for every $\varepsilon \in (0, \varepsilon_0]$, the rank-2 matrix $\mathbf{W}_\varepsilon = \mathbf{v}\mathbf{v}^* + \varepsilon\hat{\mathbf{v}}\hat{\mathbf{v}}^*$ is a feasible point to the problem (13) with the input $\mathbf{x} = \boldsymbol{f}_{\mathcal{M}}(\mathbf{w}_\varepsilon)$, because $\mathbf{W}_\varepsilon \succeq 0$ and

$$\langle \mathbf{M}_r, \mathbf{W}_\varepsilon\rangle = \langle \mathbf{M}_r, \mathbf{v}\mathbf{v}^*\rangle + \varepsilon\langle \mathbf{M}_r, \hat{\mathbf{v}}\hat{\mathbf{v}}^*\rangle = \langle \mathbf{M}_r, \mathbf{w}_\varepsilon\mathbf{w}_\varepsilon^*\rangle = x_r.$$

Additionally, one can argue that $\mathbf{W}_\varepsilon$ is an optimal solution associated with the dual certificate $\boldsymbol{\lambda}_{\mathcal{M}}(\mathbf{M}, \mathbf{v})$, due to complementary slackness:

$$\langle \mathbf{H}, \mathbf{W}_\varepsilon\rangle = \langle \mathbf{H}, \mathbf{v}\mathbf{v}^* + \varepsilon\hat{\mathbf{v}}\hat{\mathbf{v}}^*\rangle = \langle \mathbf{H}, \mathbf{v}\mathbf{v}^*\rangle + \varepsilon\langle \mathbf{H}, \hat{\mathbf{v}}\hat{\mathbf{v}}^*\rangle = 0.$$

Therefore, for every $\varepsilon \in (0, \varepsilon_0]$, since there is a rank-2 solution $\mathbf{W}_\varepsilon$, the true matrix $\mathbf{w}_\varepsilon\mathbf{w}_\varepsilon^*$ is not a unique solution. As a result,

$\mathbf{w}_\varepsilon \notin \mathcal{R}_\mathcal{M}(\mathbf{M})$. This implies that one can construct a sequence of vectors $\{\mathbf{w}_{\varepsilon_k}\}_{k\in\mathbb{N}}$ not belonging to $\mathcal{R}_\mathcal{M}(\mathbf{M})$ such that $\lim_{k\to\infty}\mathbf{w}_{\varepsilon_k} = \mathbf{v}$. This contradicts the previous assumption that $\mathbf{v}\in \mathrm{int}\{\mathcal{R}_\mathcal{M}(\mathbf{M})\}$. Therefore, we have $\tilde{\kappa}_\mathcal{M}(\mathbf{M},\mathbf{v}) > 0$, which completes the proof. □

Proof of Proposition 1: Given a vector of voltages $\mathbf{v}\in\mathcal{J}_\mathcal{M}$, suppose that there exists $\boldsymbol{\lambda}\in\mathbb{R}^{|\mathcal{M}|}$ such that

$$\mathbf{J}_\mathcal{M}(\overline{\mathbf{v}})\boldsymbol{\lambda} = -2\overline{\mathbf{M}}\overline{\mathbf{v}}, \tag{33a}$$

$$\kappa_\mathcal{M}(\mathbf{M},\boldsymbol{\lambda}) > 0. \tag{33b}$$

It follows from (33a) that $\mathbf{v}\in\mathrm{null}\{\mathbf{H}_\mathcal{M}(\mathbf{M},\boldsymbol{\lambda})\}$. Therefore, according to Lemma 4, the pair $(\mathbf{v}\mathbf{v}^*,\boldsymbol{\lambda})$ is a set of primal and dual optimal solutions for the primal SDP (13) and the dual SDP (25). Let $a_1,\ldots,a_{2n-1}\in\mathcal{M}$ denote the indices for $2n-1$ linearly independent columns of $\mathbf{J}_\mathcal{M}(\mathbf{v})$ and define $\mathcal{M}' = \{a_1,\ldots,a_{2n-1}\}$. Let $\mathbf{H}\triangleq\mathbf{H}_\mathcal{M}(\mathbf{M},\boldsymbol{\lambda})$ and consider the following optimization problem:

$$\underset{\mathbf{W}\in\mathbb{H}^n}{\text{minimize}} \quad \langle\mathbf{W},\mathbf{H}\rangle \tag{34a}$$

$$\text{subject to} \quad \langle\mathbf{W},\mathbf{M}_r\rangle = x_r, \qquad \forall r\in\mathcal{M}', \tag{34b}$$

$$\mathbf{W}\succeq 0. \tag{34c}$$

One can write $\mathbf{v}\in\mathcal{J}_{\mathcal{M}'}$ and $\mathbf{v}\in\mathrm{null}\{\mathbf{H}\}$, which imply that

$$\tilde{\kappa}_{\mathcal{M}'}(\mathbf{H},\mathbf{v}) = \kappa_{\mathcal{M}'}\left(\mathbf{H},-2\mathbf{J}_{\mathcal{M}'}^{-1}(\overline{\mathbf{v}})\overline{\mathbf{H}}\overline{\mathbf{v}}\right) = \kappa_{\mathcal{M}'}\left(\mathbf{H},\mathbf{0}_{2n-1}\right) > 0$$

Hence, according to Proposition 2, every vector of voltages in a vicinity of $\mathbf{v}$ can be recovered via (34). This means that $\mathbf{v}\in\mathrm{int}\left\{\mathcal{R}_{\mathcal{M}'}(\mathbf{H})\right\}$. On the other hand, since additional measurements do not prevent the recovery, every member of $\mathcal{R}_{\mathcal{M}'}(\mathbf{H})$ can be recovered by solving the following optimization problem as well:

$$\underset{\mathbf{W}\in\mathbb{H}^n}{\text{minimize}} \quad \langle\mathbf{W},\mathbf{H}\rangle \tag{35a}$$

$$\text{subject to} \quad \langle\mathbf{W},\mathbf{M}_r\rangle = x_r, \qquad \forall r\in\mathcal{M}, \tag{35b}$$

$$\mathbf{W}\succeq 0, \tag{35c}$$

i.e., $\mathcal{R}_{\mathcal{M}'}(\mathbf{H})\subseteq\mathcal{R}_\mathcal{M}(\mathbf{H})$. Finally, observe that every feasible matrix $\mathbf{W}$ satisfies the relation $\langle\mathbf{W},\mathbf{H}\rangle = \langle\mathbf{W},\mathbf{M}\rangle + \boldsymbol{\lambda}^\mathrm{T}\mathbf{x}$, which means that (34) and (13) are equivalent, and therefore, $\mathcal{R}_{\mathcal{M}'}(\mathbf{H})\subseteq\mathcal{R}_\mathcal{M}(\mathbf{H}) = \mathcal{R}_\mathcal{M}(\mathbf{M})$. Hence, it follows that $\mathbf{v}\in\mathrm{int}\{\mathcal{R}_\mathcal{M}(\mathbf{M})\}$. □

Proof of Theorem 1: It can be inferred from Assumption 1 that $\mathbf{1}_n\in\mathcal{J}_\mathcal{M}$ and therefore the Jacobian $\mathbf{J}_\mathcal{M}(\mathbf{1}_n)$ has full row rank. Let $a_1,\ldots,a_{2n-1}\in\mathcal{M}$ denote the indices for $2n-1$ linearly independent columns of $\mathbf{J}_\mathcal{M}(\mathbf{1}_n)$ and define $\mathcal{M}'\triangleq\{a_1,\ldots,a_{2n-1}\}$. Observe that $\mathcal{R}_{\mathcal{M}'}(\mathbf{M})\subseteq\mathcal{R}_\mathcal{M}(\mathbf{M})$. On the other hand, since $\mathbf{M}\times\mathbf{1}_n = 0$, we have $\boldsymbol{\lambda}_{\mathcal{M}'}(\mathbf{M},\mathbf{1}_n) = \mathbf{0}_{2n-1}$, which implies that $\mathbf{H}_{\mathcal{M}'}(\mathbf{M},\boldsymbol{\lambda}_{\mathcal{M}'}(\mathbf{M},\mathbf{v}')) = \mathbf{M}$. Therefore, it follows from Proposition 1 that $\mathbf{1}_n\in\mathrm{int}\{\mathcal{R}_{\mathcal{M}'}(\mathbf{M})\}$ and therefore $\mathbf{1}_n\in\mathrm{int}\{\mathcal{R}_\mathcal{M}(\mathbf{M})\}$. □

Proof of Theorem 2: Part (i) follows from the facts that the sum of the two smallest eigenvalues of a matrix is a concave function, which implies that the function $\tilde{\kappa}_\mathcal{M}(\mathbf{M},\widehat{\mathbf{v}}_k)$ is concave with respect to $\mathbf{M}$. Part (ii) follows immediately from Proposition 2. □

Proof of Proposition 3: Consider an arbitrary voltage vector $\mathbf{v}$. Let $(\mathbf{W}^\mathrm{opt},\boldsymbol{\nu}^\mathrm{opt})$ denote a solution of (17) with the input $\mathbf{x} = \boldsymbol{f}_\mathcal{M}(\mathbf{v})$. Since $(\mathbf{W},\boldsymbol{\nu}) = (\mathbf{1}_n\mathbf{1}_n^*,\boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{f}_\mathcal{M}(\mathbf{1}_n))$ is a feasible point for this problem, one can write:

$$\langle\mathbf{W}^\mathrm{opt},\mathbf{M}\rangle + \mu\times\phi(\boldsymbol{\nu}^\mathrm{opt}) \leq \langle\mathbf{1}_n\mathbf{1}_n^*,\mathbf{M}\rangle + \mu\times\phi(\boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{f}_\mathcal{M}(\mathbf{1}_n)). \tag{36}$$

On the other hand, since the inner product of every arbitrary pair of Hermitian positive semidefinite matrices is nonnegative, it follows from the relations $\mathbf{M}\succeq 0$ and $\mathbf{W}^\mathrm{opt}\succeq 0$ as well as Assumption 2 that

$$\langle\mathbf{W}^\mathrm{opt},\mathbf{M}\rangle \geq 0, \qquad \langle\mathbf{1}_n\mathbf{1}_n^*,\mathbf{M}\rangle = 0. \tag{37}$$

Combining (36) and (37) leads to the inequality $\phi(\boldsymbol{\nu}^\mathrm{opt})\leq\phi(\boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{f}_\mathcal{M}(\mathbf{1}_n))$. Moreover,

$$\begin{aligned}
&\|\boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{\nu}^\mathrm{opt}-\boldsymbol{f}_\mathcal{M}(\mathbf{1}_n)\|_F\\
&\leq\|\boldsymbol{\nu}^\mathrm{opt}\|_F+\|\boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{f}_\mathcal{M}(\mathbf{1}_n)\|_F\\
&\leq\|\boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{f}_\mathcal{M}(\mathbf{1}_n)\|_F+\max\{\|\boldsymbol{\nu}\|_F\,|\,\phi(\boldsymbol{\nu})\leq\phi(\boldsymbol{\nu}^\mathrm{opt})\}\\
&\leq\|\boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{f}_\mathcal{M}(\mathbf{1}_n)\|_F\\
&+\max\{\|\boldsymbol{\nu}\|_F\,|\,\phi(\boldsymbol{\nu})\leq\phi(\boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{f}_\mathcal{M}(\mathbf{1}_n))\}.
\end{aligned}$$

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

$$\boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{\nu}^\mathrm{opt}\in\mathrm{image}\{\mathcal{R}_\mathcal{M}(\mathbf{M})\}, \qquad \forall\mathbf{v}\in\mathcal{T} \tag{39}$$

where $\mathrm{image}\{\mathcal{R}_\mathcal{M}(\mathbf{M})\}$ denotes the image of the region $\mathcal{R}_\mathcal{M}(\mathbf{M})$ under the mapping $\boldsymbol{f}_\mathcal{M}(\cdot)$. In addition, the penalized convex problem (17) can be written as

$$\underset{\mathbf{W}\in\mathbb{H}^n}{\text{minimize}} \quad \langle\mathbf{W},\mathbf{M}\rangle \tag{40a}$$

$$\text{subject to} \quad \langle\mathbf{W},\mathbf{M}_r\rangle = \boldsymbol{f}_{\mathcal{M},r}(\mathbf{v})-\nu_r^\mathrm{opt}, \quad \forall r\in\mathcal{M} \tag{40b}$$

$$\mathbf{W}\succeq 0, \tag{40c}$$

where $\boldsymbol{f}_{\mathcal{M},r}(\mathbf{v})$ denotes the r^th entry of $\boldsymbol{f}_\mathcal{M}(\mathbf{v})$. In other words, $\mathbf{W}^\mathrm{opt}$ is a solution of the above problem. Moreover, it follows from (39) and Theorem 1 that $\mathbf{v}(\mu)\mathbf{v}(\mu)^*$ is the only solution of (40) for every $\mathbf{v}\in\mathcal{T}$, where $\mathbf{v}(\mu)$ is a vector satisfying the relation $\boldsymbol{f}_\mathcal{M}(\mathbf{v}(\mu)) = \boldsymbol{f}_\mathcal{M}(\mathbf{v})-\boldsymbol{\nu}^\mathrm{opt}$. As a result, the solution of (17) with the input $\mathbf{x} = \boldsymbol{f}_\mathcal{M}(\mathbf{v})$ 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 (17) if $\phi(\cdot)$ is strictly convex. Let $(\tilde{\mathbf{W}}^\mathrm{opt},\tilde{\boldsymbol{\nu}}^\mathrm{opt})$ be a solution of (17) with the input $\mathbf{x} = \boldsymbol{f}_\mathcal{M}(\mathbf{v})$, whose corresponding objective value is denoted as o. Observe that for every $\alpha\in(0,1)$, the linear combination $(\tilde{\alpha}\tilde{\mathbf{W}}^\mathrm{opt}+\alpha\mathbf{W}^\mathrm{opt},\tilde{\alpha}\tilde{\boldsymbol{\nu}}^\mathrm{opt}+\alpha\boldsymbol{\nu}^\mathrm{opt})$ is a feasible point for (17) as well, where $\tilde{\alpha} = 1-\alpha$. If $\tilde{\boldsymbol{\nu}}^\mathrm{opt}\neq\boldsymbol{\nu}^\mathrm{opt}$, then

$$\begin{aligned}
&\langle\mathbf{M},\tilde{\alpha}\tilde{\mathbf{W}}^\mathrm{opt}+\alpha\mathbf{W}^\mathrm{opt}\rangle+\phi(\tilde{\alpha}\tilde{\boldsymbol{\nu}}^\mathrm{opt}+\alpha\boldsymbol{\nu}^\mathrm{opt}) <\\
&\tilde{\alpha}(\langle\mathbf{M},\tilde{\mathbf{W}}^\mathrm{opt}\rangle+\phi(\tilde{\boldsymbol{\nu}}^\mathrm{opt}))+\alpha(\langle\mathbf{M},\mathbf{W}^\mathrm{opt}\rangle+\phi(\boldsymbol{\nu}^\mathrm{opt})) = o,
\end{aligned}$$

which contradicts the optimality of $(\tilde{\mathbf{W}}^\mathrm{opt},\tilde{\boldsymbol{\nu}}^\mathrm{opt})$ and $(\mathbf{W}^\mathrm{opt},\boldsymbol{\nu}^\mathrm{opt})$. Therefore, the vectors $\boldsymbol{\nu}^\mathrm{opt}$ and $\tilde{\boldsymbol{\nu}}^\mathrm{opt}$ must be identical. Hence, $\mathbf{W}^\mathrm{opt}$ and $\tilde{\mathbf{W}}^\mathrm{opt}$ must both be optimal solutions of (40). However, as stated earlier, $\mathbf{v}(\mu)\mathbf{v}(\mu)^*$ is the unique solution of (40) whenever $\mathbf{v} \in \mathcal{T}$. This completes the proof. $\square$

Proof of Proposition 4: For Part (i), assume that $\phi_{\mathcal{M}}(\boldsymbol{\nu}) = \phi_{\mathcal{M};\mathrm{WLS}}(\boldsymbol{\nu})$ 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 (17) 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 (17a) evaluated at $\mathbf{W} = \mathbf{v}\mathbf{v}^*$ is equal to

$$\langle \mathbf{W}, \mathbf{M}\rangle + \mu \times \phi_{\mathcal{M}}(\boldsymbol{\nu}) = \langle \mathbf{v}\mathbf{v}^*, \mathbf{M}\rangle \tag{41}$$

(note that $\boldsymbol{\nu}$ is equal to $\mathbf{0}_m$ in this case). On the other hand, the function (17a) at $\mathbf{W} = (1-\varepsilon)\mathbf{v}\mathbf{v}^* + \varepsilon\ \mathbf{1}_n\mathbf{1}_n^*$ can be calculated as

$$\begin{aligned}\langle \mathbf{W}, \mathbf{M}\rangle + \mu \times \phi_{\mathcal{M}}(\boldsymbol{\nu}) &= (1-\varepsilon)\langle \mathbf{v}\mathbf{v}^*, \mathbf{M}\rangle \\ &+ \sum_{r\in\mathcal{M}} \frac{\varepsilon^2\mu}{\sigma_r^2}\left(\langle \mathbf{v}\mathbf{v}^* - \mathbf{1}_n\mathbf{1}_n^*, \mathbf{M}_r\rangle\right)^2. \end{aligned}\tag{42}$$

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 subsequently (42) becomes smaller than (41). This completes the proof of Part (i).

The proof of Part (ii) is omitted because it is an immediate consequence of Theorem 3. $\square$

Proof of Theorem 3: The proof is an extension of the technique developed in [87] for a special type of the PSSE problem. Observe that the primal feasibility of the point $(\mathbf{W}^{\mathrm{opt}}, \boldsymbol{\nu}^{\mathrm{opt}})$ combined with the inequality (20) implies that

$$\begin{aligned}\phi_{\mathcal{M}}(\boldsymbol{\nu}^{\mathrm{opt}}) &= \sum_{r\in\mathcal{M}} \sigma_r^{-1}|\langle \mathbf{M}_r, \mathbf{W}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*\rangle - \omega_r| \\ &\geq \frac{1}{\mu}\sum_{r\in\mathcal{M}} \lambda_r\langle \mathbf{M}_r, \mathbf{W}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*\rangle - \sum_{r\in\mathcal{G}}\sigma_r^{-1}|\omega_r| \\ &= \frac{1}{\mu}\langle \mathbf{H}_{\mathcal{M}}(\mathbf{M},\boldsymbol{\lambda}) - \mathbf{M}, \mathbf{W}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*\rangle - \phi_{\mathcal{M}}(\boldsymbol{\omega}). \end{aligned}\tag{43}$$

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

$$\phi_{\mathcal{M}}(\boldsymbol{\nu}^{\mathrm{opt}}) \leq -\frac{1}{\mu}\langle \mathbf{M}, \mathbf{W}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*\rangle + \phi_{\mathcal{M}}(\boldsymbol{\omega}). \tag{44}$$

Replacing $\phi_{\mathcal{M}}(\boldsymbol{\nu}^{\mathrm{opt}})$ on the left side of (44) with the lower bound offered by (43) leads to

$$\frac{1}{\mu}\langle \mathbf{H}_{\mathcal{M}}(\mathbf{M},\boldsymbol{\lambda}), \mathbf{W}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*\rangle \leq 2\phi_{\mathcal{M}}(\boldsymbol{\omega}). \tag{45}$$

Due to the assumption $\boldsymbol{\lambda} \in \mathcal{D}_{\mathcal{M}}(\mathbf{M},\mathbf{v})$, we have $\mathbf{v} \in \mathrm{null}\{\mathbf{H}_{\mathcal{M}}(\mathbf{M},\boldsymbol{\lambda})\}$ and therefore

$$\langle \mathbf{H}_{\mathcal{M}}(\mathbf{M},\boldsymbol{\lambda}), \mathbf{W}^{\mathrm{opt}}\rangle \leq 2 \times \mu \times \phi_{\mathcal{M}}(\boldsymbol{\omega}). \tag{46}$$

Now, consider the eigenvalue decomposition $\mathbf{H}_{\mathcal{M}}(\mathbf{M},\boldsymbol{\lambda}) = \mathbf{U}\,\mathrm{diag}\{\boldsymbol{\tau}\}\mathbf{U}^*$, where $\boldsymbol{\tau} = [\tau_n, \ldots, \tau_2, 0]^{\mathrm{T}}$ collects the eigenvalues of $\mathbf{H}_{\mathcal{M}}(\mathbf{M},\boldsymbol{\lambda})$ in descending order and $\mathbf{U}$ is a unitary matrix whose last column is equal to $\mathbf{v}/\|\mathbf{v}\|_2$. Define

$$\hat{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{W}} & \tilde{\mathbf{w}} \\ \tilde{\mathbf{w}}^* & \tilde{W}_{nn}\end{bmatrix} = \mathbf{U}^*\mathbf{W}^{\mathrm{opt}}\mathbf{U} \tag{47}$$

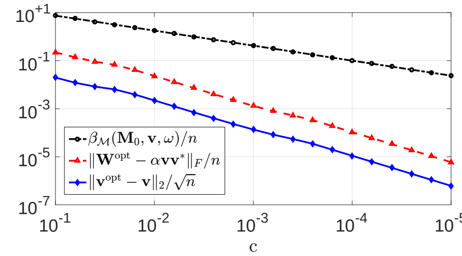


Fig. 1: This figure illustrates the convergence behavior of the error bound (19) relative to the actual estimation error of the problem (17) for the IEEE 118-bus system with the classical PF measurements.

where $\tilde{\mathbf{W}} \in \mathbb{H}^{n-1}$, $\tilde{\mathbf{w}} \in \mathbb{C}^{n-1}$ and $\tilde{W}_{nn} \in \mathbb{R}$. Due to the positive semidefiniteness of $\mathbf{W}^{\mathrm{opt}}$, we have $\hat{\mathbf{W}} \succeq 0$, which implies that the diagonal elements of $\hat{\mathbf{W}}$ are non-negative. Therefore, one can write:

$$\begin{aligned}\mathrm{trace}\{\tilde{\mathbf{W}}\} &\leq \frac{1}{\tau_2}\langle \mathrm{diag}\{\boldsymbol{\tau}\}, \hat{\mathbf{W}}\rangle \leq \frac{1}{\tau_2}\langle \mathbf{U}\mathrm{diag}\{\boldsymbol{\tau}\}\mathbf{U}^*, \mathbf{U}\hat{\mathbf{W}}\mathbf{U}^*\rangle \\ &\leq \frac{1}{\tau_2}\langle \mathbf{H}_{\mathcal{M}}(\mathbf{M},\boldsymbol{\lambda}), \mathbf{W}^{\mathrm{opt}}\rangle \leq \frac{2}{\tau_2} \times \mu \times \phi_{\mathcal{M}}(\boldsymbol{\omega}).\end{aligned}$$

Since $\tilde{\mathbf{W}} \succeq 0$, according to Schur complement, we have $\tilde{W}_{nn} \times \tilde{\mathbf{W}} \succeq \tilde{\mathbf{w}}\tilde{\mathbf{w}}^*$, and therefore,

$$\tilde{W}_{nn}\mathrm{trace}\{\tilde{\mathbf{W}}\} \geq \mathrm{trace}\{\tilde{\mathbf{w}}\tilde{\mathbf{w}}^*\} \geq \mathrm{trace}\{\tilde{\mathbf{w}}^*\tilde{\mathbf{w}}\} = \|\tilde{\mathbf{w}}\|_2^2. \tag{48}$$

Additionally,

$$\|\tilde{\mathbf{W}}\|_F^2 = \sum_{k=1}^{n}\delta_k^2 \leq \left(\sum_{k=1}^{n}\delta_k\right)^2 = \mathrm{trace}\{\tilde{\mathbf{W}}\}^2, \tag{49}$$

where $\delta_1, \ldots, \delta_n$ are the eigenvalues of $\tilde{\mathbf{W}}$. Hence, by defining $\alpha = \tilde{W}_{nn}/\|\mathbf{v}\|_2^2$, it can be concluded that

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

Now, replacing τ_2 with $\kappa_{\mathcal{M}}(\mathbf{M},\boldsymbol{\lambda})$ completes the proof. $\square$

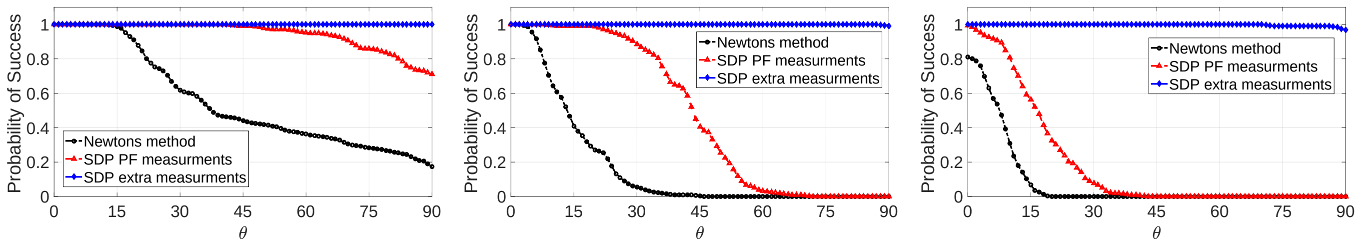
VI. Simulation Results

In what follows, we will offer several simulations on benchmark systems. We will use the OPF Solver for conic optimization (see [90]) and the MATPOWER solver for Newton's method [91] (note that different versions of Newton's method would perform slightly differently for the power flow problem).

Motivated by the promising experiments in [59] and [41] and in order to induce the sparsity of power systems into the semidefinite programs (13) and (17), two choices for the matrix $\mathbf{M}$ are considered in this section: (i) $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$, and (ii) $\mathbf{M} = -\mathbf{B}$. These choices both satisfy Assumption 2 for networks without shunt elements. Observe that:

$$\langle \mathbf{v}\mathbf{v}^*, \mathbf{Y}^*\mathbf{Y}\rangle = \sum_{k\in\mathcal{V}}\|i_k\|_2^2, \qquad \langle \mathbf{v}\mathbf{v}^*, -\mathbf{B}\rangle = \sum_{k\in\mathcal{V}} q_k. \tag{50}$$

Therefore, choices (i) and (ii) minimize the norm of the nodal current injection vector and the overall reactive power injection to the network, respectively.



(a) (b) (c)

Fig. 2: These figures show the probability of success for Newton's method, the SDP relaxation, and the SDP relaxation with extra specifications for (a): IEEE 9-bus system, (b): New England 39-bus system, and (c): IEEE 57-bus system.

A. Case Study: Power Flow Problem

In order to demonstrate the efficacy of the proposed SDP problem (13) in solving the power flow equations, we perform numerical simulations on the IEEE 9-bus, New England 39-bus, and IEEE 57-bus systems [91]. Three recovery methods are considered for each test case:

1) *Newton's method:* We evaluate the probability of convergence for Newton's method (based on default MATPOWER 6.0 settings) in polar coordinates for the classical PF problem with $2n-1$ specifications, where the starting point is $v_k = 1\angle 0^\circ$ for every $k \in \mathcal{N}$.
2) *SDP relaxation:* The probability of obtaining a rank-1 solution for the SDP relaxation (13) with $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$ is evaluated, where the same set of specifications as in Newton's method is used.
3) *SDP relaxation with extra specifications:* The probability of obtaining a rank-1 solution for the SDP relaxation (13) with $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$ is evaluated, under extra specifications compared to the classical PF problem. Active powers at PV and PQ buses, reactive powers at PQ buses, and voltages magnitudes at all buses are measured (as opposed to only PV and reference buses).

For different values of θ, we have generated 500 specification sets $(x_1,\ldots,x_{|\mathcal{M}|})$ by randomly choosing voltage vectors whose magnitudes and phases are uniformly drawn from the intervals $[0.9, 1.1]$ and $[-\theta,\theta]$, respectively. We have then exploited each of the three methods described above to find a feasible voltage vector associated with each specification set. The results are depicted in Figure 2. It can be observed that the SDP relaxations outperform the Newton's method.

B. Case Study: Power System State Estimation

In order to show that the penalized convex program (17) with a nonzero matrix $\mathbf{M}$ significantly outperforms the classical SDP relaxation of PSSE proposed in [72]–[75], we conduct simulations on the PEGASE 1354-bus and 9241-bus systems from [92]. 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 higher than the corresponding noiseless values of squared voltage magnitudes, nodal active/reactive powers, and branch flows, respectively. As pointed out in Remark 2, we choose $\mathbf{M}$ equal to $\alpha\times\mathbf{I}-\mathbf{B}$, where the constant α is the smallest value such that $\alpha\times\mathbf{I}-\mathbf{B}\succeq 0$. This choice of $\mathbf{M}$ makes the function $\langle\mathbf{M},\mathbf{W}\rangle$ account for the reactive loss in the network [41], [59].

We have performed simulations on the PEGASE 9241-bus system with randomly generated noise values corresponding to $c = 0.01$ under different numbers of extra measurements (i.e., 1%, 1.5%, and 2% errors for squared voltage magnitudes, nodal active/reactive powers, and branch flows, respectively). The performance of the penalized convex problem (17) using the two objective functions

$$f_1(\mathbf{W},\boldsymbol{\nu}) \triangleq \langle\mathbf{M},\mathbf{W}\rangle + \mu\times\phi_{\mathcal{M};\mathrm{WLS}}(\boldsymbol{\nu}), \tag{51a}$$
$$f_2(\mathbf{W},\boldsymbol{\nu}) \triangleq \langle\mathbf{M},\mathbf{W}\rangle + \mu\times\phi_{\mathcal{M};\mathrm{WLAV}}(\boldsymbol{\nu}), \tag{51b}$$

is shown in Figure 3. Each histogram shows the distribution of the absolute differences between the actual and estimated values of complex voltages. For these simulations, we have set $\mu = 100$ and $\mathbf{M} = \alpha\times\mathbf{I}-\mathbf{B}$, where α is the smallest number such that $\alpha\times\mathbf{I}-\mathbf{B}\succeq 0$. In each figure, we have assumed that the PV and PQ measurements corresponding to the classical power flow problem are all available, in addition to the specified numbers of additional line flow measurements. It can be observed that the penalized convex problem obtains high-quality estimations with both WLAV and WLS regularization terms, while it typically works better with the WLAV estimator rather than the WLS estimator.

In order to efficiently solve the large-scale semidefinite programming problem (17), we have exploited the sparsity structure of the network. More precisely, the conic constraint of the SDP problems was replaced by a set of low-order conic constraints (as discussed in [41]). For cases where the resulting solution is not rank-1, a recovery algorithm from [41] is deployed to find an approximate rank-1 SDP matrix.

Figure 1 illustrates the convergence behavior of the error bound (19) relative to the actual estimation error of the problem (17) for the classical PF measurements as the parameter c tends to zero. The IEEE 118-bus system is used for this case study with $\mathbf{M}_0 = \mathbf{Y}^*\mathbf{Y}$. The function $\beta_{\mathcal{M}}(\mathbf{M}_0,\mathbf{v},\omega)$ represents the right side of the error bound (19) for $\boldsymbol{\lambda} = -2\mathbf{J}_{\mathcal{M}}(\overline{\mathbf{v}})^{-1}\overline{\mathbf{M}}\,\overline{\mathbf{v}}$ and $\mu = \max_{r\in\mathcal{M}}|\sigma_r\lambda_r|$. Each data point represents the average of 100 experiments with random realizations of noise based on the corresponding value of the parameter c, and voltage vectors with magnitudes and angles uniformly chosen from the intervals $[0.9, 1.1]$ and $[-5^\circ, 5^\circ]$, respectively. Note that the error bound attenuates to zero as expected, and that it differs from the actual error with almost a constant factor. The reason for the discrepancy is that the derived error bound is valid for all possible noise realizations, and is a worst-case analysis.

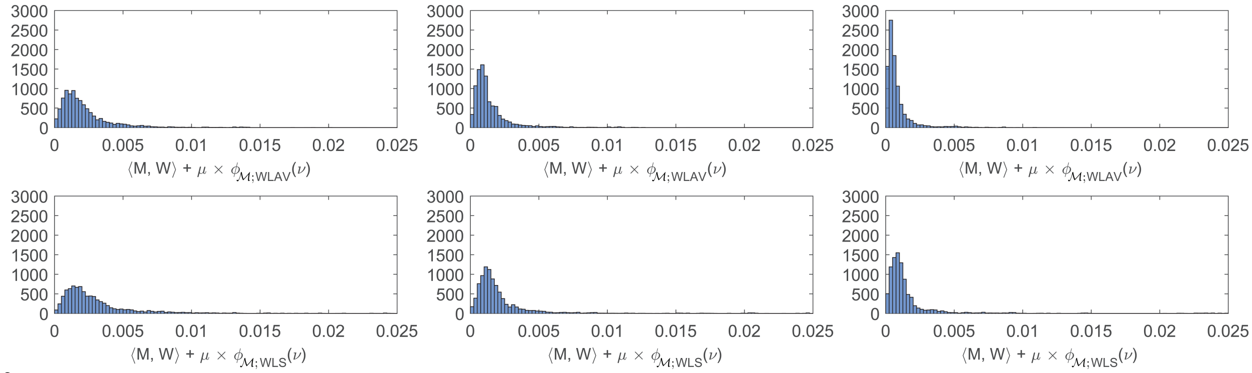


Fig. 3: Histograms of absolute differences between the actual and estimated complex voltages for the PEGASE 9241-bus system based on the penalized convex problem equipped with the WLAV and WLS estimators are given in the first row (top) and second row (bottom), respectively. In addition to the PV and PQ measurements for the classical PF problem, there are 10%, 15% and 20% of line flow measurements (the entries of $\mathbf{p}_\mathrm{f}$, $\mathbf{p}_\mathrm{t}$, $\mathbf{q}_\mathrm{f}$ and $\mathbf{q}_\mathrm{t}$) for the figures in the first column (left), second column (middle) and third column (right), respectively.

C. Case Study: Zero Injection Buses

Real-world power networks have intermediate buses that do not exchange electrical powers with any external load or generator. 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. If $\mathbf{v}$ is a solution to the power flow problem (12) with nonzero entries, then the equation $\mathbf{v}\mathbf{v}^*\mathbf{Y}^*\mathbf{e}_k = \mathbf{0}_n$ holds for every $k \in \mathcal{Z}$. Therefore, for every $k \in \mathcal{Z}$, the set of additional valid constraints $\mathbf{W}\mathbf{Y}^*\mathbf{e}_k = \mathbf{0}_n$ can be added to the SDP problems (13) and (17) in order to strengthen the relaxations.

We have conducted simulations on 20 randomly generated trials on the PEGASE 1354-bus system in the presence of 1436 zero injection valid constraints. Four different penalized SDP problems of the form (17) are tested with the objective functions (51a), (51b), $f_3(\boldsymbol{\nu}) \triangleq \phi_{\mathcal{M};\mathrm{WLS}}(\boldsymbol{\nu})$ and $f_4(\boldsymbol{\nu}) \triangleq \phi_{\mathcal{M};\mathrm{WLAV}}(\boldsymbol{\nu})$, where $\mu = 0.5$. 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 measurements (the entries of $\mathbf{p}_\mathrm{f}$, $\mathbf{p}_\mathrm{t}$, $\mathbf{q}_\mathrm{f}$ and $\mathbf{q}_\mathrm{t}$) 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 $c = 0.05$.

The root mean square errors of the recovered nodal complex voltages are plotted in Figure 4. The curves corresponding to the objective functions $\phi_{\mathcal{M};\mathrm{WLS}}(\boldsymbol{\nu})$ and $\phi_{\mathcal{M};\mathrm{WLS}}(\boldsymbol{\nu})$ are not shown in Figure 4(a) since they are significantly higher than those for the functions f_1 and f_2 in (51).

VII. Conclusion

This paper aims to find a convex model for the power system state estimation (PSSE) problem, which includes the power flow (PF) problem as a special case. 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 power 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 on real-world systems elucidate the superiority of the proposed method in estimating the state of a power system based on non-convex and noisy measurements.

References

[1] A. Pinar, J. Meza, V. Donde, and B. Lesieutre, "Optimization strategies for the vulnerability analysis of the electric power grid," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1786–1810, 2010.

[2] A. Papavasiliou and S. S. Oren, "Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network," *Operations Research*, vol. 61, no. 3, pp. 578–592, 2013.

[3] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-constrained optimal power flow: Risk-aware network control under uncertainty," *SIAM Review*, vol. 56, no. 3, pp. 461–495, 2014.

[4] B. Kocuk, H. Jeon, S. S. Dey, J. Linderoth, J. Luedtke, and X. A. Sun, "A cycle-based formulation and valid inequalities for DC power transmission problems with switching," *Operations Research*, 2016.

[5] K. Lehmann, A. Grastien, and P. V. Hentenryck, "AC-feasibility on tree networks is NP-hard," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 798–801, Jan 2016.

[6] A. Verma, "Power grid security analysis: An optimization approach," Ph.D. dissertation, Columbia University, 2009.

[7] D. Bienstock and A. Verma, "Strong NP-hardness of AC power flows feasibility," *arXiv preprint arXiv:1512.07315*, 2015.

[8] W. F. Tinney and C. E. Hart, "Power flow solution by Newton's method," *IEEE Transactions on Power Systems*, vol. 86, no. Nov., pp. 1449–1460, Jun. 1967.

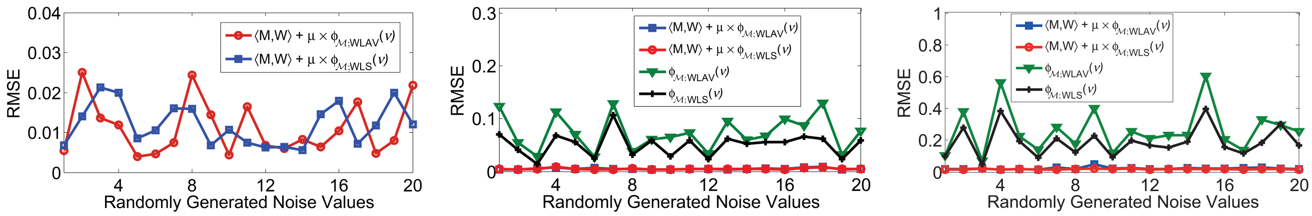
[9] B. Stott and O. Alsaç, "Fast decoupled load flow," *IEEE Transactions on Power Systems*, vol. 93, no. 3, pp. 859–869, May. 1974.

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

[11] D. Bienstock and G. Munoz, "On linear relaxations of OPF problems," *arXiv preprint arXiv:1411.1120*, 2014.

[12] C. Coffrin and P. Van Hentenryck, "A linear-programming approximation of AC power flows," *INFORMS Journal on Computing*, vol. 26, no. 4, pp. 718–734, 2014.

[13] F. Salam, L. Ni, S. Guo, and X. Sun, "Parallel processing for the load flow of power systems: the approach and applications," in *Proceedings of the 28th IEEE Conference on Decision and Control*, 1989, pp. 2173–2178.



(a) (b) (c)

Fig. 4: These figures compare the accuracy of the estimated vector of voltages obtained through the minimization of four different objective functions for the SDP relaxation of the PEGASE 1354-bus system under 20 randomly generated vectors of noise values. In each case, the PV and PQ measurements corresponding to the classical power flow problem are assumed to be available. Figure (a) corresponds to $c = 0.01$, Figure (b) is for the case with 10% additional measurements of line flows with $c = 0.01$. Figure (c) corresponds to 10% additional measurements of line flows with $c = 0.05$. Note that 1436 valid scalar constraints (due to zero injection buses) are imposed on each SDP problem.

[14] W. Ma and J. S. Thorp, "An efficient algorithm to locate all the load flow solutions," *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 1077–1083, 1993.

[15] C.-W. Liu, C.-S. Chang, J.-A. Jiang, and G.-H. Yeh, "Toward a CPFLOW-based algorithm to compute all the type-1 load-flow solutions in electric power systems," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 3, pp. 625–630, 2005.

[16] D. K. Molzahn, B. C. Lesieutre, and H. Chen, "Counterexample to a continuation-based algorithm for finding all power flow solutions," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 564–565, 2013.

[17] D. Mehta, H. Nguyen, and K. Turitsyn, "Numerical polynomial homotopy continuation method to locate all the power flow solutions," *arXiv preprint arXiv:1408.2732*, 2014.

[18] S. Chandra, D. Mehta, and A. Chakrabortty, "Equilibria analysis of power systems using a numerical homotopy method," in *IEEE Power & Energy Society General Meeting*, 2015, pp. 1–5.

[19] A. Montes and J. Castro, "Solving the load flow problem using Gröebner basis," *ACM SIGSAM Bulletin*, vol. 29, no. 1, pp. 1–13, 1995.

[20] J. Ning, W. Gao, G. Radman, and J. Liu, "The application of the Gröebner basis technique in power flow study," in *North American Power Symposium (NAPS)*, 2009, pp. 1–7.

[21] H. Mori and A. Yuihara, "Calculation of multiple power flow solutions with the Krawczyk method," in *Proceedings of the 1999 IEEE International Symposium on Circuits and Systems*, vol. 5, 1999, pp. 94–97.

[22] D. K. Molzahn, D. Mehta, and M. Niemerg, "Toward topologically based upper bounds on the number of power flow solutions," *arXiv preprint arXiv:1509.09227*, 2015.

[23] T. Chen and D. Mehta, "On the network topology dependent solution count of the algebraic load flow equations," *arXiv preprint arXiv:1512.04987*, 2015.

[24] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM review*, vol. 38, no. 1, pp. 49–95, 1996.

[25] F. Alizadeh, "Interior point methods in semidefinite programming with applications to combinatorial optimization," *SIAM Journal on Optimization*, vol. 5, no. 1, pp. 13–51, 1995.

[26] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, "Complementarity and nondegeneracy in semidefinite programming," *Mathematical Programming*, vol. 77, no. 1, pp. 111–128, 1997.

[27] Z. Li, S. He, and S. Zhang, *Approximation methods for polynomial optimization: Models, Algorithms, and Applications*. Springer Science & Business Media, 2012.

[28] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of the ACM (JACM)*, vol. 42, no. 6, pp. 1115–1145, 1995.

[29] 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.

[30] Y. Nesterov, "Semidefinite relaxation and nonconvex quadratic optimization," *Optimization methods and software*, vol. 9, no. 1-3, pp. 141–160, 1998.

[31] S. Zhang and Y. Huang, "Complex quadratic optimization and semidefinite programming," *SIAM Journal on Optimization*, vol. 16, no. 3, pp. 871–890, 2006.

[32] Z.-Q. Luo and S. Zhang, "A semidefinite relaxation scheme for multivariate quartic polynomial optimization with quadratic constraints," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1716–1736, 2010.

[33] A. Beck, Y. Drori, and M. Teboulle, "A new semidefinite programming relaxation scheme for a class of quadratic matrix problems," *Operations Research Letters*, vol. 40, no. 4, pp. 298–302, 2012.

[34] B. Jiang, Z. Li, and S. Zhang, "Approximation methods for complex polynomial optimization," *Computational Optimization and Applications*, vol. 59, no. 1-2, pp. 219–248, 2014.

[35] S. He, Z. Li, and S. Zhang, "Approximation algorithms for discrete polynomial optimization," *Journal of the Operations Research Society of China*, vol. 1, no. 1, pp. 3–36, 2013.

[36] S. He, Z. Li, and S. Zhang, "Approximation algorithms for homogeneous polynomial optimization with quadratic constraints," *Mathematical Programming*, vol. 125, no. 2, pp. 353–383, 2010.

[37] Y. Huang, A. De Maio, and S. Zhang, "Semidefinite programming, matrix decomposition, and radar code design," *Convex Optimization in Signal Processing and Communications*, pp. 192–228, 2010.

[38] Y. Huang and S. Zhang, "Approximation algorithms for indefinite complex quadratic maximization problems," *Science China Mathematics*, vol. 53, no. 10, pp. 2697–2708, 2010.

[39] S. He, Z.-Q. Luo, J. Nie, and S. Zhang, "Semidefinite relaxation bounds for indefinite homogeneous quadratic optimization," *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 503–523, 2008.

[40] Z.-Q. Luo, N. D. Sidiropoulos, P. Tseng, and S. Zhang, "Approximation bounds for quadratic optimization with homogeneous quadratic constraints," *SIAM Journal on optimization*, vol. 18, no. 1, pp. 1–28, 2007.

[41] 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.

[42] R. Madani, S. Sojoudi, G. Fazelnia, and J. Lavaei, "Finding low-rank solutions of sparse linear matrix inequalities using convex optimization," *SIAM Journal on Optimization*, vol. 27, no. 2, pp. 725–758, 2017.

[43] P. A. Parrilo, "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization," Ph.D. dissertation, 2000.

[44] A. Beck and M. Teboulle, "Global optimality conditions for quadratic optimization problems with binary constraints," *SIAM Journal on Optimization*, vol. 11, no. 1, pp. 179–188, 2000.

[45] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo, "New developments in sum of squares optimization and SOSTOOLS," in *Proceedings of the American Control Conference*, 2004, pp. 5606–5611.

[46] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[47] J. F. Bonnans, "Mathematical study of very high voltage power networks I: The optimal DC power flow problem," *SIAM Journal on Optimization*, vol. 7, no. 4, pp. 979–990, 1997.

[48] J. F. Bonnans, "Mathematical study of very high voltage power networks II: The AC power flow problem," *SIAM Journal on Applied Mathematics*, vol. 58, no. 5, pp. 1547–1567, 1998.

[49] J. Thorp and S. Naqavi, "Load flow fractals," in *Proceedings of the 28th IEEE Conference on Decision and Control*, 1989, pp. 1822–1827.

[50] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.

[51] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Review*, vol. 57, no. 2, pp. 225–251, 2015.

[52] 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.

[53] 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.

[54] 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, pp. 1–8.

[55] D. T. Phan, "Lagrangian duality and branch-and-bound algorithms for optimal power flow," *Operations Research*, vol. 60, no. 2, pp. 275–285, 2012.

[56] C. Chen, A. Atamturk, and S. S. Oren, "Bound tightening for the alternating current optimal power flow problem," *IEEE Transactions on Power Systems*, vol. PP, no. 99, pp. 1–8, 2015.

[57] B. Kocuk, S. S. Dey, and X. A. Sun, "Inexactness of SDP relaxation and valid inequalities for optimal power flow," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 642–651, 2016.

[58] C. Coffrin, H. Hijazi, and P. Van Hentenryck, "Strengthening the SDP relaxation of AC power flows with convex envelopes, bound tightening, and lifted nonlinear cuts," *arXiv preprint arXiv:1512.04644*, 2015.

[59] 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.

[60] D. Bienstock and G. Munoz, "LP approximations to mixed-integer polynomial optimization problems," *arXiv preprint arXiv:1501.00288*, 2015.

[61] M. Fukuda, M. Kojima, K. Murota, and K. Nakata, "Exploiting sparsity in semidefinite programming via matrix completion I: General framework," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 647–674, 2001.

[62] K. Nakata, K. Fujisawa, M. Fukuda, M. Kojima, and K. Murota, "Exploiting sparsity in semidefinite programming via matrix completion II: Implementation and numerical results," *Mathematical Programming*, vol. 95, no. 2, pp. 303–327, 2003.

[63] A. Lam, B. Zhang, and D. Tse, "Distributed algorithms for optimal power flow problem," in *Proceedings of the 51st IEEE Conference on Decision and Control*, 2012.

[64] B. Zhang, A. Y. S. Lam, A. D. Domnguez-Garca, and D. Tse, "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1714–1726, July 2015.

[65] 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 Transactions on Power Systems*, vol. 28, no. 4, pp. 3987–3998, Nov. 2013.

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

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

[68] B. Kocuk, S. S. Dey, and X. A. Sun, "Strong SOCP relaxations for the optimal power flow problem," *Operations Research*, 2016.

[69] B. Kocuk, S. S. Dey, and X. A. Sun, "New formulation and strong MISOCP relaxations for AC optimal transmission switching problem," *arXiv preprint arXiv:1510.02064*, 2015.

[70] 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, 2014.

[71] 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, 2014.

[72] 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.

[73] 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.

[74] 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," *IET Generation, Transmission & Distribution*, vol. 9, no. 16, pp. 2760–2767, Nov 2015.

[75] 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.

[76] A. Beck and Y. C. Eldar, "Sparsity constrained nonlinear optimization: Optimality conditions and algorithms," *SIAM Journal on Optimization*, vol. 23, no. 3, pp. 1480–1509, 2013.

[77] A. Beck, Y. C. Eldar, and Y. Shechtman, "Nonlinear compressed sensing with application to phase retrieval," in *Global Conference on Signal and Information Processing*, 2013, pp. 617–617.

[78] Y. Chen and E. Candes, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," in *Advances in Neural Information Processing Systems*, 2015, pp. 739–747.

[79] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.

[80] 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.

[81] 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.

[82] W. Ai, Y. Huang, and S. Zhang, "On the low rank solutions for linear matrix inequalities," *Mathematics of Operations Research*, vol. 33, no. 4, pp. 965–975, 2008.

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

[84] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE transactions on information theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[85] H. Zou, T. Hastie, and R. Tibshirani, "On the "degrees of freedom" of the lasso," *The Annals of Statistics*, vol. 35, no. 5, pp. 2173–2192, 2007.

[86] M. Rudelson and R. Vershynin, "Smallest singular value of a random rectangular matrix," *Communications on Pure and Applied Mathematics*, vol. 62, no. 12, pp. 1707–1739, 2009.

[87] Y. Zhang, R. Madani, and J. Lavaei, "Conic relaxations for power system state estimation with line measurements," *IEEE Transactions on Control of Network Systems*, pp. 1–12, April 2017.

[88] D. G. Luenberger and Y. Ye, *Linear and nonlinear programming*. Springer, 1984.

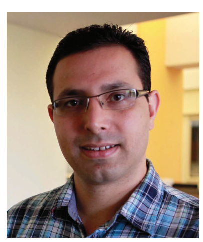
[89] A. Duval, C. Klivans, and J. Martin, "Simplicial matrix-tree theorems," *T. Am. Math. Soc.*, vol. 361, no. 11, pp. 6073–6114, Jun. 2009.

[90] R. Madani, M. Ashraphijuo, and J. Lavaei, "OPF Solver," *http://ieor.berkeley.edu/~lavaei/Software.html*, 2014.

[91] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower's extensible optimal power flow architecture," in *IEEE Power & Energy Society General Meeting*, 2009, pp. 1–7.

[92] 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.

Ramtin Madani is an Assistant Professor at the Electrical Engineering Department of the University of Texas at Arlington. He received the Ph.D. degree in Electrical Engineering from Columbia University in 2015, and was a postdoctoral scholar in the Department of Industrial Engineering and Operations Research at University of California, Berkeley in 2016.



Javad Lavaei is an Associate Professor in the Department of Industrial Engineering and Operations Research at University of California, Berkeley. He obtained his Ph.D. degree in Computing and Mathematical Sciences from California Institute of Technology in 2011. He has won the Donald P. Eckman Award, INFORMS Optimization Society Prize for Young Researchers, INFORMS ENRE Energy Best Publication Award, and SIAM Control and Systems Theory Prize.

Ross Baldick (F07) received the B.Sc. degree in mathematics and physics and the B.E. degree in electrical engineering from the University of Sydney, Sydney, Australia, and the M.S. and Ph.D. degrees in electrical engineering and computer sciences from the University of California, Berkeley, in 1988 and 1990, respectively. From 1991 to 1992, he was a Post-Doctoral Fellow with the Lawrence Berkeley Laboratory, Berkeley, CA. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Texas, Austin.