Title

Conic relaxations for power system state estimation with line measurements

Permalink

https://escholarship.org/uc/item/4vw168fn

Journal

IEEE Transactions on Control of Network Systems, 5(3)

ISSN

2325-5870

Authors

Zhang, Y
Madani, R
Lavaei, J

Publication Date

2018-09-01

DOI

10.1109/TCNS.2017.2691473

Peer reviewed

Conic Relaxations for Power System State Estimation with Line Measurements

Yu Zhang, Ramtin Madani, and Javad Lavaei

arXiv:1704.00133v1 [math.OC] 1 Apr 2017

Abstract—This paper deals with the non-convex power system state estimation (PSSE) problem, which plays a central role in the monitoring and operation of electric power networks. Given a set of noisy measurements, PSSE aims at estimating the vector of complex voltages at all buses of the network. This is a challenging task due to the inherent nonlinearity of power flows, for which existing methods lack guaranteed convergence and theoretical analysis. Motivating by these limitations, we propose a novel convexification framework for the PSSE using semidefinite programming (SDP) and second-order cone programming (SOCP) relaxations. We first study a related power flow (PF) problem as the noiseless counterpart, which is cast as a constrained minimization program by adding a suitably designed objective function. We study the performance of the proposed framework in the case where the set of measurements includes: (i) nodal voltage magnitudes, and (ii) branch active power flows over a spanning tree of the network. It is shown that the SDP and SOCP relaxations both recover the true PF solution as long as the voltage angle difference across each line of the network is not too large (e.g., less than $90°$ for lossless networks). By capitalizing on this result, penalized SDP and SOCP problems are designed to solve the PSSE, where a penalty based on the weighted least absolute value is incorporated for fitting noisy measurements with possible bad data. Strong theoretical results are derived to quantify the optimal solution of the penalized SDP problem, which is shown to possess a dominant rank-one component formed by lifting the true voltage vector. An upper bound on the estimation error is also derived as a function of the noise power, which decreases exponentially fast as the number of measurements increases. Numerical results on benchmark systems, including a 9241-bus European system, are reported to corroborate the merits of the proposed convexification framework.

Index Terms—Power system state estimation, power flow analysis, convex relaxations, semidefinite programming, tree decomposition.

NOMENCLATURE

A. Sets and numbers

$\mathcal{N}$, N	Set and number of buses.
$\mathcal{L}$, L	Set and number of power lines.
$\mathcal{M}$, M	Set and number of measurements.
$\mathcal{N}_p$, $\mathcal{N}_q$	Sets of active and reactive power injection measurements.
$\mathcal{D}$	Set of all dual SDP certificates.

B. Input signals and constants

$\mathbf{v}$, $\mathbf{i}$	N-dimensional complex vectors of nodal voltages (state of the system) and current injections.
$\mathbf{p}$, $\mathbf{q}$	N-dimensional real vectors of net injected active and reactive powers.
$\mathbf{i}_f$, $\mathbf{i}_t$	L-dimensional complex vectors of current injections at the *from* and *to* ends of all branches.
$\mathbf{Y}$, $\mathbf{Y}_f$, $\mathbf{Y}_t$	Matrices of nodal admittance, *from* branch admittance, and *to* branch admittance.
$\mathbf{Y}_{l,p_f}$, $\mathbf{Y}_{l,p_t}$	Coefficient matrices corresponding to active power flow measurements at the *from* and *to* ends over the l-th branch.
$\mathbf{Y}_{l,q_f}$, $\mathbf{Y}_{l,q_t}$	Coefficient matrices corresponding to reactive power flow measurements at the *from* and *to* ends over the l-th branch.
$\mathbf{M}_0$	Designed coefficient matrix in the objective.
$\mathbf{M}_j$	Coefficient matrix corresponding to the j-th measurement.
$\mathbf{z}$	M-dimensional real vector collecting all measurements.
$|v_k|$, $\angle v_k$	Voltage magnitude and angle at the k-th bus.
$\angle y_{st}$	Angle of the branch (s, t) line admittance.
η_j, σ_j	Additive noise and positive weight of the j-th measurement.
ρ	Positive weight trading off the data fitting cost and the designed linear regularizer.
ζ	Defined root-mean-square estimation error of the obtained optimal SDP solution.

C. Variables and functions

$\mathbf{X}$, $\mathbf{H}$	$N \times N$ primal and dual matrix variables.
$\boldsymbol{\mu}$	M-dimensional real vector of Lagrange multipliers.
$\boldsymbol{\nu}$	M-dimensional real vector of slack variables.
$f_{\text{WLAV}}(\cdot)$	Weighted least absolute value cost.
$f_{\text{WLS}}(\cdot)$	Weighted least squares cost.

I. INTRODUCTION

An electrical grid infrastructure is operated for delivering electricity from power generators to consumers via interconnected transmission and distribution networks. Accurately determining the operating point and estimating the underlying state of the system are of paramount importance for the reliable and economic operation of power networks. Power flow analysis and power system state estimation play indispensable

roles in the planning and monitoring of the power grid. The solutions of these two problems are used for many optimal resource allocation problems such as unit commitment, optimal power flow (OPF), security-constrained OPF, and network reconfiguration [1], [2].

A. Power Flow Analysis

The power flow (PF) problem is a numerical analysis of the steady-state electrical power flows, which serves as a necessary prerequisite for future system planning. Specifically, having measured the voltage magnitudes and injected active/reactive powers at certain buses, the PF problem aims to find the unknown voltage magnitude and phase angle at each bus of a power network. Using the obtained voltage phasors and the network admittances, line power flows can then be determined for the entire system. The calculation of power flows is essentially equivalent to solving a set of quadratic equations obeying the laws of physics. Solving a system of nonlinear polynomial equations is NP-hard in general. Beźout's theorem asserts that a well-behaved system can have exponentially many solutions [3]. Upper bounds on the number of PF solutions have been analyzed in the recent work [4] and the references therein. When it comes to the feasibility of AC power flows, it is known that this problem is NP-hard for both transmission and distribution networks [5], [6].

For solving the PF problem, many iterative methods such as the Newton-Raphson method and Gauss-Seidel algorithms have been extensively studied over the last few decades [7]. The Newton-Raphson method features quadratic convergence whenever the initial point is sufficiently close to the solution [8], [9]. Nevertheless, a fundamental drawback of various Newton-based algorithms is that there is no convergence guarantee in general. By leveraging advanced techniques in complex analysis and algebraic geometry, sophisticated tools have been developed for solving PF, including holomorphic embedding load flow and numerical polynomial homotopy continuation [10], [11]. However, these approaches involve costly computations, and are generally not suitable for large-scale power systems. Using the theory of monotone operators and moment relaxations, the papers [12] and [13] identify a "monotonicity domain", within which it is possible to efficiently find the PF solutions or certify their non-existence. A review on recent advances in computational methods for the PF equations can be found in [14].

Facing the inherent challenge of non-convexity, convex relaxation techniques have been recently developed for finding the PF solutions [15]. More specifically, a class of convex programs is proposed to solve the PF problem in the case where the solution belongs to a recovery region that contains voltage vectors with small angles. The proposed convex programs are in the form of semidefinite programming (SDP), where a convex objective is designed as a surrogate of the rank-one constraint to guarantee the exactness of the SDP relaxation.

B. Power System State Estimation

Closely related to the PF problem, the power system state estimation (PSSE) problem plays a key role for grid monitoring. System measurements are acquired through the supervisory control and data acquisition (SCADA) systems, as well as increasingly pervasive phasor measurement units (PMUs). Given these noisy measurements, the PSSE task aims at estimating the complex voltage at each bus, and determining the system's operating conditions. The PSSE is traditionally formulated as a nonlinear least-squares (LS) problem, which is commonly solved by the Gauss-Newton algorithm in practice [16], [17]. The algorithm is based on a sequence of linear approximations of the nonlinear residuals. A descent direction is obtained at each iteration by minimizing the sum of squares of the linearized residuals. However, the Gauss-Newton algorithm has no guaranteed convergence in general. Furthermore, a linear search must be carefully carried out for the damped Gauss-Newton method. The widely-adopted Levenberg-Marquardt algorithm finds only a local optimum of the nonlinear LS problem, and may still be slow for large residual or highly nonlinear problems [18], [19].

For a linear regression model, the classic Gauss-Markov theorem states that if the additive noises are uncorrelated with mean zero and homoscedastic with finite variance, then the ordinary least squares estimator (LSE) of the unknown parameters is the best linear unbiased estimator (BLUE) that yields the least variance estimates. The generalized LSE should be applied when the noise covariance matrix is positive definite [18]. The work [20] shows that even when the noise covariance matrix is singular, the BLUE can be found by utilizing its pseudo-inverse in the generalized normal equations. Analytic solutions of the BLUE and the minimum variances of the estimates are available for the linear model. In addition, minimum variance unbiased estimator (MVUE) and Bayesian-based estimators are studied in [21] and [22]. It is well known that when the linear measurements are normally distributed, the LSE coincides with the maximum-likelihood estimator (MLE). However, LSE for the PSSE problem may not possess these attractive properties due to the inherently nonlinear measurements. There are several issues involved from both optimization and statistical perspectives:

- The problem of nonlinear LS estimation is generally non-convex, which can have multiple local solutions. Hence, finding a globally optimal solution is challenging.
- Newton-based iterative algorithms are sensitive to the initialization and lack a guaranteed convergence. They may converge to a stationary point. It is nevertheless not easy to interpret that point, and quantify its distance relative to the true unknown state of the system.
- Even if a global solution can be obtained, the nonlinear LSE may not correspond to the MVUE. When the noises are not from the exponential family of distributions, the LSE is different from the MLE in general.
- The LSE is vulnerable to the presence of outliers primarily due to its uniformly squaring, which makes data with large residuals have a significant influence on the fitted model.

To deal with bad data, the weighted least absolute value (WLAV) function is proposed as the data fitting cost in [23], [24], for which efficient algorithms are developed in [25],

[26]. The work [27] presents linear transformations to mitigate the deteriorating effect of "leverage points" on the WLAV estimator. Robust or distributed PSSE has also been developed in the papers [28]–[31]. The state estimation problem with line flow measurements using an iterative algorithm is studied in [32] and [33], where complex power flows over all transmission lines and at least one voltage phasor are assumed to be measured to achieve the necessary redundancy for the solution of the problem. The performance of these selected measurements and the proposed algorithm tested on the Ontario hydro power system are reported in [34]. Heuristic optimization techniques are also utilized for PSSE in [35], [36].

Intensive studies of the SDP relaxation technique for solving fundamental problems in power networks have been springing up due to the pioneering papers [37], [38] and [39]. The work [39] develops an SDP relaxation for finding a global minimum of the OPF problem. A sufficient and necessary condition is provided to guarantee a zero duality gap, which is satisfied by several benchmark systems. From the perspective of the physics of power systems, the follow-up papers [40] and [41] develop theoretical results to support the success of the SDP relaxation in handling the non-convexity of OPF. The papers [42] and [43] develop a graph-theoretic SDP framework for finding a near-global solution whenever the SDP relaxation fails to find a global minimum. Recent advances in the convex relaxation of the OPF problem are summarized in the tutorial papers [44] and [45].

The paper [46] initializes the idea of solving the PSSE problem via the SDP relaxation. When the SDP solution is not rank one, its principal eigenvector is used to recover approximate voltage phasors. The work [47] suggests generating a "good" initial point from the SDP optimal solution to improve the performance of Newton's method, while a nuclear norm regularizer is used to promote a low-rank solution in [48]. Distributed or online PSSE using the SDP relaxation can be found in [49]–[51]. However, in the literature there is a lack of theoretical analysis on the quality of the SDP optimal solution for estimating the complex voltages. Hence, to the best of our knowledge, this is still an intriguing open problem.

The aforementioned grand challenges of the PSSE problem motivate us to revisit the design of a high-performance estimator with finite measurements. The novelty and main contributions of the present work are outlined in the ensuing subsection.

C. Contributions

In this paper, we start with a PF problem that can be regarded as the noiseless counterpart of PSSE. In contrast to the standard setup with only nodal measurements at the PV, PQ and slack buses, one objective of this work is to investigate the effect of branch flow measurements on reducing the computational complexity of the PF problem. Motivated by the work [15], we contrive a convex optimization framework for the PF problem using SDP and second-order cone programming (SOCP) relaxations. It is shown that the proposed conic relaxations are both always exact if: (i) the set of

measurements includes the nodal voltage magnitude at each bus and line active power flows over a spanning tree of the power network, and (ii) the line phase voltage differences are not too large (e.g., less than $90°$ for lossless networks).

By building upon the proposed convexification framework for the PF problem, we develop a penalized convex program for solving the PSSE problem. In addition to an ℓ_1 norm penalty that is robust to outliers in the measurements, the objective function of the penalized convex problem features a linear regularization term whose coefficient matrix can be systematically designed according to the meter placements. We present a theoretical result regarding the quality of the optimal solution of the convex program. It is shown that the obtained optimal solution has a dominant rank-one matrix component, which is formed by lifting the vector of true system state. The distance between the solution of the penalized convex problem and the correct rank-one component is quantified as a function of the noise level. An upper bound of the tail probability of this distance is further derived, which also implies the correlation between the quality of the estimation and the number of measurements.

The effort of this paper is mainly on the scenario where the measurements include nodal voltage magnitudes and branch active power flows. However, the developed mathematical framework is rather general and could be adopted to study the PSSE problem with other types of measurements.

D. Notations

Boldface lower (upper) case letters represent column vectors (matrices), and calligraphic letters stand for sets. The symbols $\mathbb{R}$ and $\mathbb{C}$ denote the sets of real and complex numbers, respectively. $\mathbb{R}^N$ and $\mathbb{C}^N$ denote the spaces of N-dimensional real and complex vectors, respectively. $\mathbb{S}^N$ and $\mathbb{H}^N$ stand for the spaces of $N \times N$ complex symmetric and Hermitian matrices, respectively. The symbols $(\cdot)^\top$ and $(\cdot)^*$ denote the transpose and conjugate transpose of a vector/matrix. $\mathrm{Re}(\cdot)$, $\mathrm{Im}(\cdot)$, $\mathrm{rank}(\cdot)$, $\mathrm{Tr}(\cdot)$, and $\mathrm{null}(\cdot)$ denote the real part, imaginary part, rank, trace, and null space of a given scalar or matrix. $\|\mathbf{a}\|_2$, $\|\mathbf{A}\|_F$, and $\|\mathbf{A}\|_*$ denote the Euclidean norm of the vector $\mathbf{a}$, the Frobenius norm and the nuclear norm of the matrix $\mathbf{A}$, respectively. The relation $\mathbf{X} \succeq \mathbf{0}$ means that the matrix $\mathbf{X}$ is Hermitian positive semidefinite. The (i, j) entry of $\mathbf{X}$ is given by $X_{i,j}$. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. The symbol $\mathrm{diag}(\mathbf{x})$ denotes a diagonal matrix whose diagonal entries are given by the vector $\mathbf{x}$, while $\mathrm{diag}(\mathbf{X})$ forms a column vector by extracting the diagonal entries of the matrix $\mathbf{X}$. The imaginary unit is denoted by j. The expectation operator and the probability measure are denoted by $\mathbb{E}(\cdot)$ and $\mathbb{P}(\cdot)$, respectively. The notations $\angle x$ and $|x|$ denote the angle and magnitude of a complex number x. The notation $\mathbf{X}[\mathcal{S}_1, \mathcal{S}_2]$ denotes the submatrix of $\mathbf{X}$ whose rows and columns are chosen from the given index sets $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively.

II. PRELIMINARIES

A. System Modeling

Consider an electric power network represented by a graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N} := \{1, \ldots, N\}$ and $\mathcal{L} := \{1, \ldots, L\}$

denote the sets of buses and branches, respectively. Let $v_k \in \mathbb{C}$ denote the nodal complex voltage at bus $k \in \mathcal{N}$, whose magnitude and phase angle are given as $|v_k|$ and $\angle v_k$. The net injected complex power at bus k is denoted as $s_k = p_k + q_k \mathrm{j}$. Define $s_{lf} = p_{lf} + q_{lf}\mathrm{j}$ and $s_{lt} = p_{lt} + q_{lt}\mathrm{j}$ as the complex power injections entering the line $l \in \mathcal{L}$ through the *from* and *to* ends of the branch. Note that the current $i_{l,f}$ and $i_{l,t}$ may not add up to zero due to the existence of transformers and shunt capacitors. Denote the admittance of each branch (s,t) of the network as y_{st}. The Ohm's law dictates that

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

where $\mathbf{Y} = \mathbf{G} + \mathrm{j}\mathbf{B} \in \mathbb{S}^N$ is the nodal admittance matrix of the power network, whose real and imaginary parts are the conductance matrix $\mathbf{G}$ and susceptance matrix $\mathbf{B}$, respectively. Furthermore, $\mathbf{Y}_f \in \mathbb{C}^{L \times N}$ and $\mathbf{Y}_t \in \mathbb{C}^{L \times N}$ represent the *from* and *to* branch admittance matrices. The injected complex power can thus be expressed as $\mathbf{p} + \mathbf{q}\mathrm{j} = \mathrm{diag}(\mathbf{v}\mathbf{v}^*\mathbf{Y}^*)$. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ denote the canonical vectors in $\mathbb{R}^N$. Define

$$\mathbf{E}_k := \mathbf{e}_k\mathbf{e}_k^\top, \quad \mathbf{Y}_{k,p} := \frac{1}{2}(\mathbf{Y}^*\mathbf{E}_k + \mathbf{E}_k\mathbf{Y}),$$
$$\mathbf{Y}_{k,q} := \frac{\mathrm{j}}{2}(\mathbf{E}_k\mathbf{Y} - \mathbf{Y}^*\mathbf{E}_k). \tag{2}$$

For each $k \in \mathcal{N}$, the quantities $|v_k|^2$, p_k and q_k can be written as

$$|v_k|^2 = \mathrm{Tr}(\mathbf{E}_k\mathbf{v}\mathbf{v}^*), \ p_k = \mathrm{Tr}(\mathbf{Y}_{k,p}\mathbf{v}\mathbf{v}^*), \ q_k = \mathrm{Tr}(\mathbf{Y}_{k,q}\mathbf{v}\mathbf{v}^*). \tag{3}$$

Similarly, the branch active and reactive powers for each line $l \in \mathcal{L}$ can be expressed as

$$p_{l,f} = \mathrm{Tr}(\mathbf{Y}_{l,p_f}\mathbf{v}\mathbf{v}^*), \quad p_{l,t} = \mathrm{Tr}(\mathbf{Y}_{l,p_t}\mathbf{v}\mathbf{v}^*)$$
$$q_{l,f} = \mathrm{Tr}(\mathbf{Y}_{l,q_f}\mathbf{v}\mathbf{v}^*), \quad q_{l,t} = \mathrm{Tr}(\mathbf{Y}_{l,q_t}\mathbf{v}\mathbf{v}^*), \tag{4}$$

where the coefficient matrices $\mathbf{Y}_{l,p_f}, \mathbf{Y}_{l,p_t}, \mathbf{Y}_{l,q_f}, \mathbf{Y}_{l,q_t} \in \mathbb{H}^N$ are defined over the l-th branch from node i to node j as

$$\mathbf{Y}_{l,p_f} := \frac{1}{2}(\mathbf{Y}_f^*\mathbf{d}_l\mathbf{e}_i^\top + \mathbf{e}_i\mathbf{d}_l^\top\mathbf{Y}_f) \tag{5a}$$

$$\mathbf{Y}_{l,p_t} := \frac{1}{2}(\mathbf{Y}_t^*\mathbf{d}_l\mathbf{e}_j^\top + \mathbf{e}_j\mathbf{d}_l^\top\mathbf{Y}_t) \tag{5b}$$

$$\mathbf{Y}_{l,q_f} := \frac{\mathrm{j}}{2}(\mathbf{e}_i\mathbf{d}_l^\top\mathbf{Y}_f - \mathbf{Y}_f^*\mathbf{d}_l\mathbf{e}_i^\top) \tag{5c}$$

$$\mathbf{Y}_{l,q_t} := \frac{\mathrm{j}}{2}(\mathbf{e}_j\mathbf{d}_l^\top\mathbf{Y}_t - \mathbf{Y}_t^*\mathbf{d}_l\mathbf{e}_j^\top), \tag{5d}$$

where $\{\mathbf{d}_1, \ldots, \mathbf{d}_L\}$ is the set of canonical vectors in $\mathbb{R}^L$.

So far, nodal and line measurements of interest have been expressed as quadratic functions of the complex voltage $\mathbf{v}$. The PF and PSSE problems will be formulated next.

B. Convex Relaxation of Power Flow Equations

The task of the PSSE problem is to estimate the complex voltage vector $\mathbf{v}$ based on M real measurements:

$$z_j = \mathbf{v}^*\mathbf{M}_j\mathbf{v} + \eta_j, \quad \forall j \in \mathcal{M} := \{1, 2, \ldots, M\}, \tag{6}$$

where $\{z_j\}_{j \in \mathcal{M}}$ are the known measurements, $\{\eta_j\}_{j \in \mathcal{M}}$ are the possible measurement noises with known statistical information, and $\{\mathbf{M}_j\}_{j \in \mathcal{M}}$ are arbitrary measurement matrices that could be any subset of the Hermitian matrices defined

in (2) and (5). The PF problem is a noiseless version of the PSSE problem. More specifically, given a total of M noiseless specifications z_j for $j = 1, 2, \ldots, M$, the goal of PF is to find the nodal complex voltage vector $\mathbf{v}$ satisfying all quadratic measurement equations, i.e.,

$$\text{find} \quad \mathbf{v} \in \mathbb{C}^N \tag{7a}$$
$$\text{subject to} \quad \mathbf{v}^*\mathbf{M}_j\mathbf{v} = z_j, \quad \forall j \in \mathcal{M}. \tag{7b}$$

After setting the phase of the voltage at the slack bus to zero, the problem reduces to M power flow equations with $2N - 1$ unknown real parameters. The classical PF problem corresponds to the case $M = 2N - 1$, where the measurements are specified at the PV, PQ, and slack buses such that:

- For each PV (generator) bus k, the active power p_k and the voltage magnitude $|v_k|$ are given.
- For each PQ (load) bus k, the active power p_k and the reactive power q_k are given.
- For the slack (reference) bus, the voltage magnitude $|v_{\mathrm{ref}}|$ and the phase angle $\angle v_{\mathrm{ref}}$ are given.

Instead of solving the feasibility problem (7) to obtain the voltage vector $\mathbf{v}$, consider the optimization problem

$$\underset{\mathbf{X} \in \mathbb{H}^N, \mathbf{v} \in \mathbb{C}^N}{\text{minimize}} \quad \mathrm{Tr}(\mathbf{M}_0\mathbf{X}) \tag{8a}$$
$$\text{subject to} \quad \mathrm{Tr}(\mathbf{M}_j\mathbf{X}) = z_j, \quad \forall j \in \mathcal{M} \tag{8b}$$
$$\mathbf{X} = \mathbf{v}\mathbf{v}^*, \tag{8c}$$

where its objective function is to be designed later. Note that the constraint $\mathbf{X} = \mathbf{v}\mathbf{v}^*$ can be equivalently replaced by the two conditions $\mathbf{X} \succeq \mathbf{0}$ and $\mathrm{rank}(\mathbf{X}) = 1$. The SDP relaxation of (8) is obtained by dropping the rank-one constraint as

$$\underset{\mathbf{X} \in \mathbb{H}^N}{\text{minimize}} \quad \mathrm{Tr}(\mathbf{M}_0\mathbf{X}) \tag{9a}$$
$$\text{subject to} \quad \mathrm{Tr}(\mathbf{M}_j\mathbf{X}) = z_j, \quad \forall j \in \mathcal{M} \tag{9b}$$
$$\mathbf{X} \succeq \mathbf{0}. \tag{9c}$$

This relaxation correctly solves (8) if and only if it has a unique rank-1 solution $\mathbf{X}^{\mathrm{opt}}$, in which case $\mathbf{v}$ can be recovered via the decomposition $\mathbf{X}^{\mathrm{opt}} = \mathbf{v}\mathbf{v}^*$. The dual of (9) can be obtained as

$$\underset{\boldsymbol{\mu} \in \mathbb{R}^M}{\text{maximize}} \quad -\mathbf{z}^\top\boldsymbol{\mu} \tag{10a}$$
$$\text{subject to} \quad \mathbf{H}(\boldsymbol{\mu}) \succeq \mathbf{0}, \tag{10b}$$

where the vector $\mathbf{z} := [z_1, \ldots, z_M]^\top$ collects all the available measurements, $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_M]^\top$ is the Lagrangian multiplier vector associated with the linear equality constraints (9b), and the dual matrix function $\mathbf{H} : \mathbb{R}^M \to \mathbb{H}^N$ is defined as

$$\mathbf{H}(\boldsymbol{\mu}) := \mathbf{M}_0 + \sum_{j=1}^M \mu_j\mathbf{M}_j. \tag{11}$$

If strong duality holds while the primal and dual problems both attain their solutions, then every pair of optimal primal-dual solutions $(\mathbf{X}^{\mathrm{opt}}, \boldsymbol{\mu}^{\mathrm{opt}})$ satisfies the relation $\mathbf{H}(\boldsymbol{\mu}^{\mathrm{opt}})\mathbf{X}^{\mathrm{opt}} = \mathbf{0}$, due to the complementary slackness. Hence, if $\mathrm{rank}(\mathbf{H}(\boldsymbol{\mu}^{\mathrm{opt}})) = N - 1$ holds, then we have the inequality $\mathrm{rank}(\mathbf{X}^{\mathrm{opt}}) \leq 1$ such that the SDP relaxation can recover a solution of the PF problem.
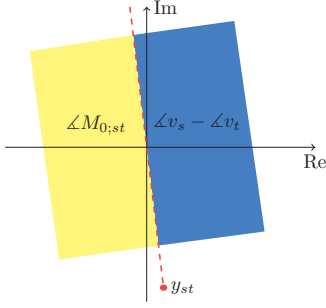
Fig. 1: The demonstration of the angle conditions (16) and (17). The acceptable regions for the voltage phase difference $\angle v_s - \angle v_t$ (blue open half-space) and the entry $M_{0;st}$ (yellow open half-space) are shown relative to the branch admittance y_{st} (red dot).

Definition 1 (SDP recovery). *It is said that the SDP relaxation problem (9) recovers the voltage vector $\mathbf{v} \in \mathbb{C}^N$ if $\mathbf{X} = \mathbf{v}\mathbf{v}^*$ is the unique solution of (9) for some input $\mathbf{z} \in \mathbb{R}^M$.*

Definition 2 (Dual certificate). *A vector $\boldsymbol{\mu} \in \mathbb{R}^M$ is regarded as a dual SDP certificate for the voltage vector $\mathbf{v} \in \mathbb{C}^N$ if it satisfies the following three properties:*

$$\mathbf{H}(\boldsymbol{\mu}) \succeq \mathbf{0}, \quad \mathbf{H}(\boldsymbol{\mu})\mathbf{v} = \mathbf{0}, \quad \mathrm{rank}(\mathbf{H}(\boldsymbol{\mu})) = N - 1. \quad (12)$$

Denote the set of all dual SDP certificates for the voltage vector $\mathbf{v}$ as $\mathcal{D}(\mathbf{v})$.

The SDP problem (9) can be further relaxed by replacing the high-order positive semidefinite constraint (9c) with second-order conic constraints on 2×2 principal sub-matrices of $\mathbf{X}$ corresponding to certain lines of the network. This yields the SOCP relaxation:

$$\underset{\mathbf{X} \in \mathbb{H}^N}{\text{minimize}} \quad \mathrm{Tr}(\mathbf{M}_0\mathbf{X}) \quad (13a)$$

$$\text{subject to} \quad \mathrm{Tr}(\mathbf{M}_j\mathbf{X}) = z_j, \qquad \forall j \in \mathcal{M} \quad (13b)$$

$$\begin{bmatrix} X_{s,s} & X_{s,t} \\ X_{t,s} & X_{t,t} \end{bmatrix} \succeq \mathbf{0}, \quad \forall (s,t) \in \overline{\mathcal{L}}, \quad (13c)$$

where $\overline{\mathcal{L}}$ denotes the set of those edges of the network graph for which the corresponding entry of $\mathbf{M}_j$ is nonzero for at least one index $j \in \{0, 1, \ldots, M\}$.

Definition 3 (SOCP recovery). *It is said that the SOCP relaxation problem (13) recovers the voltage vector $\mathbf{v} \in \mathbb{C}^N$ if there is some input $\mathbf{z} \in \mathbb{R}^M$ such that, for every solution $\mathbf{X}^{\mathrm{opt}}$ of (13), those entries of the matrix $\mathbf{X}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*$ on the diagonal or corresponding to the members of $\overline{\mathcal{L}}$ are all equal to zero.*

III. EXACT RECOVERY OF POWER FLOW SOLUTION

The objective of this section is to show that the SDP problem (9) is exact and the correct complex voltage vector $\mathbf{v}$ can be recovered for a class of nodal and branch noiseless measurements. Let $\mathcal{G}' = (\mathcal{N}, \mathcal{L}')$ denote an arbitrary subgraph of $\mathcal{G}$ that contains a spanning tree of $\mathcal{G}$. Throughout the rest of this section, we assume that the available measurements consist of: (i) voltage magnitudes at all buses, and (ii) active

power flow at the "from" end of each line of $\mathcal{G}'$. Note that whenever the SDP relaxation is exact for this set of measurements, it remains exact if more measurements are available. Please refer to Corollary 1 and Remark 1 for more details.

The SDP relaxation of (8) can be expressed as

$$\underset{\mathbf{X} \in \mathbb{H}^N}{\text{minimize}} \quad \mathrm{Tr}(\mathbf{M}_0\mathbf{X}) \quad (14a)$$

$$\text{subject to} \quad X_{k,k} = |v_k|^2, \qquad \forall k \in \mathcal{N} \quad (14b)$$

$$\mathrm{Tr}(\mathbf{Y}_{l,p_f}\mathbf{X}) = p_{l,f}, \qquad \forall l \in \mathcal{L}' \quad (14c)$$

$$\mathbf{X} \succeq \mathbf{0}. \quad (14d)$$

Moreover, the SOCP relaxation of (8) can be written as:

$$\underset{\mathbf{X} \in \mathbb{H}^N}{\text{minimize}} \quad \mathrm{Tr}(\mathbf{M}_0\mathbf{X}) \quad (15a)$$

$$\text{subject to} \quad X_{k,k} = |v_k|^2, \qquad \forall k \in \mathcal{N} \quad (15b)$$

$$\mathrm{Tr}(\mathbf{Y}_{l,p_f}\mathbf{X}) = p_{l,f}, \quad \forall l \in \mathcal{L}' \quad (15c)$$

$$\begin{bmatrix} X_{s,s} & X_{s,t} \\ X_{t,s} & X_{t,t} \end{bmatrix} \succeq \mathbf{0}, \quad \forall (s,t) \in \mathcal{L}'. \quad (15d)$$

Definition 4 (Sparsity graph). *Given a Hermitian matrix $\mathbf{W} \in \mathbb{H}^N$, the sparsity graph of $\mathbf{W}$, denoted by $\mathscr{G}(\mathbf{W})$, is a simple undirected graph with the vertex set $\{1, 2, \ldots, N\}$ such that every two distinct vertices i and j are connected to each other if and only if the (i, j) entry of $\mathbf{W}$ is nonzero.*

Assumption 1. *The edge set of $\mathscr{G}(\mathbf{M}_0)$ coincides with $\mathcal{L}'$ and in addition,*

$$-180° < \angle M_{0;st} - \angle y_{st} < 0, \quad \forall (s,t) \in \mathcal{L}', \quad (16)$$

where $M_{0;st}$ denotes the (s, t) entry of $\mathbf{M}_0$. Moreover, the solution $\mathbf{v}$ being sought satisfies the relations

$$0 < (\angle v_s - \angle v_t) - \angle y_{st} < 180°, \quad \forall (s,t) \in \mathcal{L}' \quad (17a)$$

$$(\angle v_s - \angle v_t) - \angle M_{0;st} \neq 0 \text{ or } 180°, \quad \forall (s,t) \in \mathcal{L}'. \quad (17b)$$

To reduce power losses, real-world transmission systems feature low R/X ratios (the ratio of line resistance to reactance). The angle of the line admittance $\angle y_{st}$ is therefore close to $-90°$ [52, Sec. 3.7]. Meanwhile, since the transferred real power is proportional to its corresponding voltage angle difference, the number $|\angle v_s - \angle v_t|$ is typically small due to thermal and stability limits [2], [53]. Hence, the angle condition (17a) is expected to hold. For lossless networks, (17a) requires each line voltage angle difference to be between $-90°$ and $90°$, which is a very practical assumption. The acceptable regions for $\angle v_s - \angle v_t$ and $M_{0;st}$ are shown in Figure 1. It can be observed that one convenient choice for the matrix $\mathbf{M}_0$ is to select its entries $M_{0;st}$ as complex numbers with negative real and imaginary parts.

Lemma 1. *Under Assumption 1, there exists a dual SDP certificate for the voltage vector $\mathbf{v} \in \mathbb{C}^N$.*

Proof. The proof is provided in Appendix A. □

Theorem 1. *Under Assumption 1, the SDP relaxation (14) and the SOCP relaxation (15) both recover the voltage vector $\mathbf{v} \in \mathbb{C}^N$.*
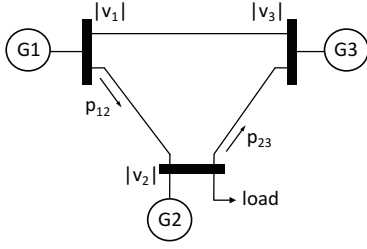
Fig. 2: A 3-bus power network with the voltage magnitude measurements $|v_1|$, $|v_2|$ and $|v_3|$, as well as the branch active power measurements p_{12} and p_{23}.

Proof. The proof is provided in Appendix B. □

To be able to recover a large set of voltage vectors, Theorem 1 implies that there are infinitely many choices for the objective function of the SDP relaxation, namely all matrices $\mathbf{M}_0$ satisfying Assumption 1. Now, consider the case with extra nodal measurements

$$\text{Tr}(\mathbf{Y}_{k,p}\mathbf{X}) = p_k, \qquad \forall k \in \mathcal{N}_p \tag{18a}$$
$$\text{Tr}(\mathbf{Y}_{k',q}\mathbf{X}) = q_{k'}, \qquad \forall k' \in \mathcal{N}_q. \tag{18b}$$

The next corollary shows that the property of the exact relaxation is preserved in presence of these arbitrary extra power injection measurements. As will be studied later in the paper, the availability of extra measurements seems unnecessary for the PF problem, but is instrumental in recovering the state of the system in the noisy setup.

Corollary 1. *Under Assumption 1, the SDP relaxation* (14) *and the SOCP relaxation* (15) *with the additional constraints of power injection measurements* (18) *both recover the voltage vector* $\mathbf{v} \in \mathbb{C}^N$.

Proof. With extra nodal power measurements, $\mathbf{X} = \mathbf{vv}^*$ still remains feasible for both problems. Therefore the corollary comes as a direct result of Theorem 1. □

A. Effect of Reactive Power Branch Measurements

In the preceding section, the exactness of the SDP and SOCP relaxations were studied in the case with the measurement of branch active power flows. In what follows, it will be shown that reactive power line flows do not offer the same benefits as active power measurements. Assume that, as opposed to the active power flow, the reactive power flow at the "from" end of each branch of $\mathcal{G}'$ is measured . In this case, Theorem 1 still holds if the conditions provided in Assumption 1 are replaced by:

$$\text{Re}(M_{0;st}y_{st}^*) \neq 0 \quad \text{and} \quad \text{Im}(v_s v_t^* M_{0;st}^*) \neq 0 \tag{19a}$$
$$\text{Re}(v_s v_t^* y_{st}^*)\text{Re}(M_{0;st}y_{st}^*) \leq 0. \tag{19b}$$

In contrast to the case with the measurements of $p_{l,f}$, the following two different scenarios must be considered for (19b)

(i): if $90° < (\measuredangle v_s - \measuredangle v_t) - \measuredangle y_{st} \leq 180°$, then $\text{Re}(v_s v_t^* y_{st}^*) < 0$ and $\text{Re}(M_{0;st}y_{st}^*) > 0$, which imply that

$$-90° \leq \measuredangle M_{0;st} - \measuredangle y_{st} \leq 90°, \tag{20}$$

(ii): if $0 \leq (\measuredangle v_s - \measuredangle v_t) - \measuredangle y_{st} < 90°$, then $\text{Re}(v_s v_t^* y_{st}^*) > 0$ and $\text{Re}(M_{0;st}y_{st}^*) < 0$, which imply that

$$90° \leq \measuredangle M_{0;st} - \measuredangle y_{st} \leq 270°. \tag{21}$$

As a result, $\measuredangle M_{0;st}$ must belong to one of the two complementary intervals $[\measuredangle y_{st}+90°, \measuredangle y_{st}+270°]$ and $[\measuredangle y_{st}-90°, \measuredangle y_{st}+90°]$, depending on the value of $\measuredangle v_s - \measuredangle v_t$. Therefore, it is impossible to design the matrix $\mathbf{M}_0$ in advance without knowing the phase angle difference $\measuredangle v_s - \measuredangle v_t$.

Remark 1. *We assume that available measurements include the voltage magnitude at each bus and active line flows over at least a spanning tree of the power network. Such an assumption is realistic in practical power systems since these two types of measurements are typically provided by the SCADA system with little incremental cost [54], while also used for conventional static state estimation algorithms [55]. Another source of voltage magnitude measurements comes from the increasing usage of PMUs. Moreover, the selection of line power flow measurements features several advantages [32], [34]:*

- *The spanning tree line flow measurements ensure the network observability [56], [57].*
- *The line flow measurements can be directly used for monitoring, which is of practical importance.*
- *Measurements at both ends of lines are very effective in detecting and identifying incorrect data.*
- *The numerical computation is fast and stable, while the results are less sensitive to measurement errors.*

Nevertheless, the above assumption on the types of measurements is not essential for the validity of the proposed convexification framework. In other words, this framework can be deployed for arbitrary measurements, but we study its performance under the above assumption. It is worth stressing that, similar to the aforementioned PF problem, additional measurements such as nodal power injections can be readily incorporated in our framework for PSSE.

B. Three-Bus Example

Consider the 3-bus power system shown in Figure 2. Suppose that the measured signals consist of the two active power line flows p_{12} and p_{23}, as well as the nodal voltage squared magnitudes $|v_1|^2$, $|v_2|^2$ and $|v_3|^2$. Theorem 1 states that the SDP and SOCP relaxation problems (14) and (15) are both able to find the unknown voltage vector $\mathbf{v}$, using an appropriately designed coefficient matrix $\mathbf{M}_0$. It turns out that $\mathbf{v}$ can also be found through a direct calculation. More precisely, one can write

$$p_{12} = \text{Re}(v_1(v_1 - v_2)^* y_{12}^*) = |v_1|^2 \text{Re}(y_{12})$$
$$- |v_1||v_2||y_{12}| \cos(\measuredangle v_1 - \measuredangle v_2 - \measuredangle y_{12}) \tag{22a}$$
$$p_{23} = \text{Re}(v_2(v_2 - v_3)^* y_{23}^*) = |v_2|^2 \text{Re}(y_{23})$$
$$- |v_2||v_3||y_{23}| \cos(\measuredangle v_2 - \measuredangle v_3 - \measuredangle y_{23}), \tag{22b}$$

which yields that

$$\angle v_1 - \angle v_2 = \arccos\left(\frac{p_{12} - |v_1|^2 \text{Re}(y_{12})}{|v_1||v_2||y_{12}|}\right) + \angle y_{12} \quad (23a)$$

$$\angle v_2 - \angle v_3 = \arccos\left(\frac{p_{23} - |v_2|^2 \text{Re}(y_{23})}{|v_2||v_3||y_{23}|}\right) + \angle y_{23}. \quad (23b)$$

Each phase difference $\angle v_1 - \angle v_2$ or $\angle v_2 - \angle v_3$ can have two possible solutions, but only one of them satisfies the angle condition (17a). Hence, all complex voltages can be readily recovered. This argument applies to general power networks. In other words, without resorting to the relaxed problems (14) and (15), the PF problem considered in this paper can be directly solved by the calculation of phase angles. However, once the measurements are noisy, the equations (23) cannot be used because the exact values of the quantities p_{12}, p_{23}, $|v_1|^2$, $|v_2|^2$ and $|v_3|^2$ are no longer available since they are corrupted by noise. In contrast, the proposed SDP and SOCP relaxations work in both noiseless and noisy cases. This will be elaborated in the next section.

As a byproduct of the discussion made above, one can obtain the following result.

Corollary 2. *The PF problem has a unique solution satisfying Assumption 1. Moreover, this solution can be recovered using the SDP relaxation* (14) *and the SOCP relaxation* (15).

IV. CONVEXIFICATION OF STATE ESTIMATION PROBLEM

Consider the PSSE as a generalization of the PF problem, where the measurements are subject to noise. As explained in Corollary 2, the unknown solution $\mathbf{v}$ is unique under Assumption 1. To find this solution, consider the optimization problem:

$$\underset{\mathbf{v}\in\mathbb{C}^N, \boldsymbol{\nu}\in\mathbb{R}^M}{\text{minimize}} \quad f(\boldsymbol{\nu}) \quad (24a)$$

$$\text{subject to} \quad z_j - \mathbf{v}^* \mathbf{M}_j \mathbf{v} = \nu_j, \quad \forall j \in \mathcal{M}, \quad (24b)$$

where $\boldsymbol{\nu} := [\nu_1, \ldots, \nu_M]^\top$ and the function $f(\cdot)$ quantifies the estimation criterion. Common choices of $f(\cdot)$ are the weighted ℓ_1 and ℓ_2 norm functions:

$$f_{\text{WLAV}}(\boldsymbol{\nu}) = \frac{|\nu_1|}{\sigma_1} + \frac{|\nu_2|}{\sigma_2} + \cdots + \frac{|\nu_M|}{\sigma_M} \quad (25)$$

$$f_{\text{WLS}}(\boldsymbol{\nu}) = \frac{\nu_1^2}{\sigma_1^2} + \frac{\nu_2^2}{\sigma_2^2} + \cdots + \frac{\nu_M^2}{\sigma_M^2}, \quad (26)$$

where $\sigma_1, \ldots, \sigma_M$ are positive constants.

Remark 2. *The above functions correspond to the weighted least absolute value (WLAV) and weighted least square (WLS) estimators, which arise as the maximum likelihood estimator when the noises have a Laplace or normal distribution, respectively. Note that possible outliers in the measurements can be better modeled by the Laplace distribution that features heavier tails than the normal. Consequently, the WLAV estimator is more robust to the outliers. On the contrary, the non-robustness of the WLS estimator is primarily attributed to the squared distance because outliers with large residuals can have a high influence to skew the regression.*

Due to the inherent quadratic relationship between the voltage vector $\mathbf{v}$ and the measured quantities $\{|v_i|^2, \mathbf{p}, \mathbf{q}, \mathbf{p}_l, \mathbf{q}_l\}$, the quadratic equality constraints (24b) make the problem (24) non-convex and NP-hard in general. To remedy this drawback, consider the penalized SDP relaxation

$$\underset{\mathbf{X}\in\mathbb{H}^N, \boldsymbol{\nu}\in\mathbb{R}^M}{\text{minimize}} \quad \rho f(\boldsymbol{\nu}) + \text{Tr}(\mathbf{M}_0\mathbf{X}) \quad (27a)$$

$$\text{subject to} \quad \text{Tr}(\mathbf{M}_j\mathbf{X}) + \nu_j = z_j, \quad \forall j \in \mathcal{M} \quad (27b)$$

$$\mathbf{X} \succeq \mathbf{0}, \quad (27c)$$

where $\rho > 0$ is a pre-selected coefficient that balances the data fitting cost $f(\boldsymbol{\nu})$ with the convexification term $\text{Tr}(\mathbf{M}_0\mathbf{X})$. The latter term is inherited from the SDP relaxation for the PF problem to deal with the non-convexity of the power flow equations. Similarly, a penalized SOCP relaxation problem can be derived as

$$\underset{\mathbf{X}\in\mathbb{H}^N, \boldsymbol{\nu}\in\mathbb{R}^M}{\text{minimize}} \quad \rho f(\boldsymbol{\nu}) + \text{Tr}(\mathbf{M}_0\mathbf{X}) \quad (28a)$$

$$\text{subject to} \quad \text{Tr}(\mathbf{M}_j\mathbf{X}) + \nu_j = z_j, \quad \forall j \in \mathcal{M} \quad (28b)$$

$$\begin{bmatrix} X_{s,s} & X_{s,t} \\ X_{t,s} & X_{t,t} \end{bmatrix} \succeq \mathbf{0}, \quad \forall (s,t) \in \overline{\mathcal{L}}, \quad (28c)$$

where $\overline{\mathcal{L}}$ denotes the set of edges of the network graph for which the corresponding entry of $\mathbf{M}_j$ is nonzero for at least one index $j \in \{0, 1, \ldots, M\}$. Based on the results derived earlier for the PF problem, we will next develop strong theoretical results on the estimation error for the PSSE.

A. Bounded Estimation Error

In this subsection, we assume that the function $f(\boldsymbol{\nu})$ corresponds to the WLAV estimator, and that the available measurements consist of the voltage magnitudes at all buses and the active power flow at the "from" end of each line of $\mathcal{G}'$. The results to be presented next hold true in presence of extra power measurements (see Remark 1). The penalized problem (27) can be expressed as

$$\min_{\mathbf{X}\succeq\mathbf{0}} \text{Tr}(\mathbf{M}_0\mathbf{X}) + \rho \sum_{j=1}^{M} \sigma_j^{-1} \left|\text{Tr}\left(\mathbf{M}_j(\mathbf{X} - \mathbf{v}\mathbf{v}^*)\right) - \eta_j\right|. \quad (29)$$

We aim to show that the solution of the penalized relaxation estimates the true solution of PSSE, where the estimation error is a function of the noise power. Define $\boldsymbol{\eta}$ as the vector of the noise values $\eta_1, .., \eta_M$.

Theorem 2. *Suppose that Assumption 1 holds. Consider an arbitrary dual SDP certificate $\hat{\boldsymbol{\mu}} \in \mathcal{D}(\mathbf{v})$, where $\mathbf{v}$ is the unique solution of the PSSE problem. Let $(\mathbf{X}^{\text{opt}}, \boldsymbol{\nu}^{\text{opt}})$ denote an optimal solution of the penalized convex program* (27) *with $f(\boldsymbol{\nu}) = f_{\text{WLAV}}(\boldsymbol{\nu})$ and a coefficient ρ satisfying the inequality*

$$\rho \geq \max_{j\in\mathcal{M}} |\sigma_j \hat{\mu}_j|. \quad (30)$$

There exists a scalar $\beta > 0$ such that

$$\zeta := \frac{\|\mathbf{X}^{\text{opt}} - \beta\mathbf{v}\mathbf{v}^*\|_F}{\sqrt{N \times \text{Tr}(\mathbf{X}^{\text{opt}})}} \leq 2\sqrt{\frac{\rho \times f_{\text{WLAV}}(\boldsymbol{\eta})}{N\lambda}}, \quad (31)$$

where λ is the second smallest eigenvalue of the matrix $\mathbf{H}(\hat{\boldsymbol{\mu}})$.

Proof. The proof is provided in Appendix C. ☐

Note that the numerator of ζ quantifies the distance between the optimal solution of the penalized convex program and the true PSSE solution. The denominator of ζ is expected to be around N since $\mathrm{Tr}(\mathbf{X}^{\mathrm{opt}}) \simeq N$ in the noiseless scenario. Hence, the quantity ζ can be regarded as a root-mean-square estimation error. Theorem 2 establishes an upper bound for the estimation error as a function of the noise power $f_{\mathrm{WLAV}}(\boldsymbol{\eta})$. In particular, the error is zero if $\boldsymbol{\eta} = 0$. This theorem provides an upper bound on the estimation error without using any statistical information of the random vector $\boldsymbol{\eta}$. In what follows, the upper bound will be further studied for Gaussian random variables. To this end, define κ as $\frac{M}{N}$. If M were the number of lines in the network, κ was between 1.5 and 2 for most real-world power systems [58].

Corollary 3. *Suppose that the noise $\boldsymbol{\eta}$ is a zero-mean Gaussian vector with the covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1^2, ..., \sigma_M^2)$. Under the assumptions of Theorem 2, the tail probability of the estimation error ζ is upper bounded as*

$$\mathbb{P}(\zeta > t) \leq \mathrm{e}^{-\gamma M} \qquad (32)$$

for every $t > 0$, where $\gamma = \frac{t^4 \lambda^2}{32\kappa^2 \rho^2} - \ln 2$.

Proof. The proof is given in Appendix D. ☐

Recall that the measurements used for solving the PSSE problem include one active power flow per each line of the subgraph $\mathcal{G}'$. The graph $\mathcal{G}'$ could be as small as a spanning tree of $\mathcal{G}$ or as large as the entire graph $\mathcal{G}$. Although the results developed in this paper work in all of these cases, the number of measurements could significantly vary for different choices of $\mathcal{G}'$. A question arises as to how the number of measurements affects the estimation error. To address this problem, notice that if it is known that some measurements are corrupted with high values of noise, it would be preferable to discard those bad measurements. To avoid this scenario, assume that there are two sets of measurements with similar noise levels. It is aimed to show that the set with a higher cardinality would lead to a better estimation error.

Definition 5. *Define $\omega(\mathcal{G}')$ as the minimum of $2\sqrt{\frac{\rho}{N\lambda}}$ over all dual SDP certificates $\hat{\boldsymbol{\mu}} \in \mathcal{D}(\mathbf{v})$, where $\rho = \max_{j \in \mathcal{M}} |\sigma_j \hat{\mu}_j|$ and λ denotes the second smallest eigenvalue of $\mathbf{H}(\hat{\boldsymbol{\mu}})$.*

In light of Theorem 2, the estimation error ζ satisfies the inequality

$$\zeta \leq \omega(\mathcal{G}')\sqrt{f_{\mathrm{WLAV}}(\boldsymbol{\eta})} \qquad (33)$$

if an optimal coefficient ρ is used in the penalized convex problem. The term $\sqrt{f_{\mathrm{WLAV}}(\boldsymbol{\eta})}$ is related to the noise power. If this term is kept constant, then the estimation error is a function of $\omega(\mathcal{G}')$. Hence, it is desirable to analyze $\omega(\mathcal{G}')$.

Theorem 3. *Consider two choices of the graph $\mathcal{G}'$, denoted as $\mathcal{G}'_1$ and $\mathcal{G}'_2$, such that $\mathcal{G}'_1$ is a subgraph of $\mathcal{G}'_2$. Then, the relation*

$$\omega(\mathcal{G}'_2) \leq \omega(\mathcal{G}'_1) \qquad (34)$$

holds.

Proof. The proof follows from the fact that the feasible set of the dual certificate $\hat{\boldsymbol{\mu}}$ for the case $\mathcal{G}' = \mathcal{G}'_1$ is contained in the feasible set of $\hat{\boldsymbol{\mu}}$ for $\mathcal{G}' = \mathcal{G}'_2$. ☐

The penalized convex program (27) may have a non-rank-1 solution in the noisy case. Whenever the optimal solution $\mathbf{X}^{\mathrm{opt}}$ is not rank 1, an estimated voltage vector $\hat{\mathbf{v}}$ can be obtained using a rank-1 approximation method, such as the following algorithm borrowed from [43]:

i) Set the voltage magnitudes via the equations

$$|\hat{v}_k| = \sqrt{\mathbf{X}^{\mathrm{opt}}_{k,k}}, \quad k = 1, 2, \ldots, N. \qquad (35)$$

ii) Set the voltage angles via the convex program

$$\angle\hat{\mathbf{v}} = \arg\min_{\angle\mathbf{v} \in [-\pi,\pi]^N} \sum_{(s,t)\in\mathcal{L}} |\angle\mathbf{X}^{\mathrm{opt}}_{s,t} - \angle v_s + \angle v_t| \qquad (36a)$$

$$\text{subject to} \quad \angle v_{\mathrm{ref}} = 0. \qquad (36b)$$

Note that $\hat{\mathbf{v}}$ is the true solution of the PSSE problem if $\mathbf{X}^{\mathrm{opt}}$ has rank 1.

B. Reduction of Computational Complexity

Due to the presence of the positive semidefinite constraint $\mathbf{X} \succeq \mathbf{0}$, solving the conic problems (9) and (27) is computationally expensive or even prohibitive for large-scale power systems. In this subsection, we deploy a graph-theoretic approach to replace the complicating constraint $\mathbf{X} \succeq \mathbf{0}$ with a set of small-sized SDP or SOCP constraints.

Definition 6. *The sparsity graph of the problem (9) or (27) is defined as the union of the sparsity graphs of the coefficient matrices $\mathbf{M}_j$ for $j = 0, 1, \ldots, M$. In other words, the sparsity graph of (9) or (27) denoted as $\tilde{\mathcal{G}} = (\mathcal{N}, \tilde{\mathcal{L}})$ is a simple undirected graph with N vertices, which has an edge between every two distinct vertices s and t if and only if the (s,t) entry of $\mathbf{M}_{j;st}$ is nonzero for some $j \in \{0, 1, \ldots, M\}$.*

Definition 7 (Tree decomposition). *A tree decomposition of $\tilde{\mathcal{G}}$ is a 2-tuple $(\mathcal{B}, \mathcal{T})$, where $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_Q\}$ is a collection of subsets of $\mathcal{N}$ and $\mathcal{T}$ is a tree whose nodes (called bags) are the subsets $\mathcal{B}_r$ and satisfy the following properties:*

- *Vertex coverage: Each vertex of $\tilde{\mathcal{G}}$ is a member of at least one node of $\mathcal{T}$, i.e., $\mathcal{N} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_Q$.*
- *Edge coverage: For every edge (s,t) in $\tilde{\mathcal{G}}$, there is a bag $\mathcal{B}_r$ that contains both ends s and t.*
- *Running intersection: For every two bags $\mathcal{B}_i$ and $\mathcal{B}_j$ in $\mathcal{T}$, every node on the path connecting $\mathcal{B}_i$ and $\mathcal{B}_j$ contains $\mathcal{B}_i \cap \mathcal{B}_j$. In other words, all nodes of $\mathcal{T}$ that contain a common vertex of $\tilde{\mathcal{G}}$ should form a subtree.*

Theorem 4. *The optimal objective values of the SDP problems (9) and (27) do not change if their constraint $\mathbf{X} \succeq \mathbf{0}$ is replaced by the set of constraints*

$$\mathbf{X}[\mathcal{B}_r, \mathcal{B}_r] \succeq \mathbf{0}, \quad \forall r \in \{1, 2, \ldots, Q\}. \qquad (37)$$

Proof. This theorem is a direct consequence of the matrix completion theorem and chordal extension [59]. ☐

As a by-product of Theorem 4, all off-diagonal entries of $\mathbf{X}$ that do not appear in the submatrices $\mathbf{X}[\mathcal{B}_r, \mathcal{B}_r]$ are redundant

and could be eliminated from the SDP relaxations. This significantly reduces the computational complexity for sparse power systems. As an example, consider the case where the sparsity graph $\tilde{\mathcal{G}}$ is acyclic. Then, $\tilde{\mathcal{G}}$ has a tree decomposition such that each bag contains only two connected vertices of $\tilde{\mathcal{G}}$. Hence, the decomposed constraints (37) boil down to positive semidefinite constraints on a set of 2×2 submatrices of $\mathbf{X}$. This special case is formalized below.

Corollary 4. *Suppose that the sparsity graph $\tilde{\mathcal{G}}$ is a spanning tree of $\mathcal{G}$. Then, the optimal objective value of the penalized SDP problem (27) is equal to the optimal objective value of the penalized SOCP problem (28).*

It can be readily shown that the number of scalar optimization variables associated with the SOCP relaxation (28) (after eliminating redundant variables) is $\mathcal{O}(N)$ as opposed to $\mathcal{O}(N^2)$ for the SDP relaxation (27).

V. NUMERICAL TESTS

In this section, numerical results are presented to verify the performance of the proposed convexification techniques for the PSSE problem. The tests are conducted on several benchmark power systems [60], where the admittance matrices and the underlying system states are obtained from MATPOWER [61]. Unless otherwise stated, the available measurements are assumed to be: (i) voltage magnitudes at all buses, and (ii) one active power flow per line of a spanning tree of the network. The tree is obtained by the function graphminspantree in the Matlab bioinformatics toolbox [62].

We first compare the proposed SOCP relaxation (28) with the conventional WLS estimator (by using the Matpower function run_se with flat start) for estimating the true complex voltage vector $\mathbf{v}$. The performance metric is the root-mean-square error (RMSE) of the estimated voltage $\hat{\mathbf{v}}$, which is defined as $\xi(\hat{\mathbf{v}}) := \|\hat{\mathbf{v}} - \mathbf{v}\|_2 / \sqrt{N}$. The simulation results tested on the IEEE 57-bus and 118-bus systems are shown in Figures 3(a) and 3(b), respectively. In each case, the measurements are under 100 randomly generated realizations of noise, which correspond to the voltage magnitudes for all buses and the active power flows at both ends of all lines. The zero-mean Gaussian noises have 0.002 and 0.001 per unit standard deviations for squared voltage magnitudes and line flows, respectively. In addition, 20% of randomly chosen line flow measurements are generated as bad data, which are contaminated by adding zero-mean Gaussian noises with 0.1 per unit standard deviation. The coefficient matrix $\mathbf{M}_0$ is chosen as a real symmetric matrix with negative values at entries corresponding to the line flow measurements and zero elsewhere. The penalty weight is set to $\rho = 1$ for all test cases. Clearly, the penalized SOCP method significantly outperforms the conventional Newton-based WLS estimator.

Furthermore, we evaluate the effect of different types of measurements and scaling of load demand on the performance of PSSE. The simulation results are shown in Figure 4. In Figure 4(a), voltage measurements are only given at the reference and load (PQ) buses. In addition to the active power flows at both ends of all lines, reactive power
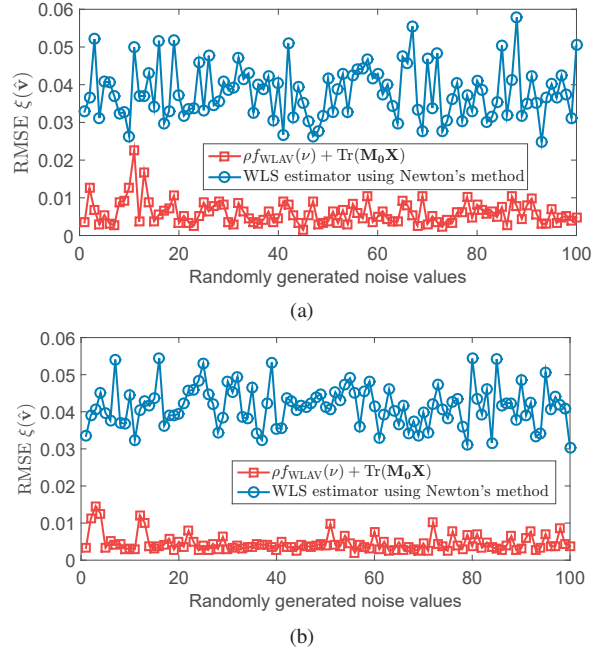

(a)

(b)

Fig. 3: The RMSEs of the estimated voltages obtained by the penalized SOCP method and the WLS-based Newton method: (a) IEEE 57-bus system, (b) IEEE 118-bus system.

TABLE I: Performance of the penalized SDP (27) with the noise level $c = 0.01$.

Cases	$\xi(\hat{\mathbf{v}})$	ζ	$\zeta^{\max}$	β	λ	f_{WLAV}	$\rho^{\min}$
9-bus	0.0111	0.0145	0.1535	0.9972	1.3417	14.768	0.0048
14-bus	0.0057	0.0078	0.2859	1.0005	0.3812	20.509	0.0053
30-bus	0.0060	0.0084	0.3728	0.9997	0.1094	51.479	0.0022
39-bus	0.0077	0.0083	0.8397	1.0009	0.7438	62.558	0.0817
57-bus	0.0092	0.0102	0.8364	1.0013	0.0912	88.434	0.0103
118-bus	0.0057	0.0079	1.2585	0.9992	0.0878	179.509	0.0228

available at "to" ends of half of the lines. Despite the fact that our assumption on voltage measurements does not hold in this case, the proposed approach still has much smaller RMSEs. Similarly, performance gains are observed in Figure 4(b), where all fixed loads are scaled up 10%.

The numerical results for the penalized SDP relaxation problem (27) performed on several benchmark systems are shown in Tables I and II. The following numbers shown in (31) are reported for each case:

- ζ: the RMSE of the obtained optimal SDP solution $\mathbf{X}^{\text{opt}}$.
- $\zeta^{\max}$: the upper bound of ζ.
- Other relevant quantities β, λ, f_{WLAV} and $\rho^{\min}$.

In this test, for each squared voltage magnitude $\{|v_k|^2\}_{k \in \mathcal{N}}$, the standard deviation of the zero-mean Gaussian noise is chosen c times higher than its noiseless value, where $c > 0$ is a pre-selected scalar quantifying the noise level. Likewise, the standard deviations for nodal and branch active/reactive power measurements are $1.5c$ and $2c$ times higher than the corresponding noiseless values, respectively. The entries of matrix $\mathbf{M}_0$ are set as $\mathbf{M}_{0;st} = -\mathbf{B}_{st}$ for all $(s,t) \in \mathcal{L}'$, and $\mathbf{M}_{0;ii} = \sum_{j=1}^N |\mathbf{B}_{i,j}|$ for $i = 1, 2, \ldots, N$. The penalty weight is set to $\rho^{\min} := \max_{j \in \mathcal{M}} |\sigma_j \hat{\mu}_j|$ as given in (30).

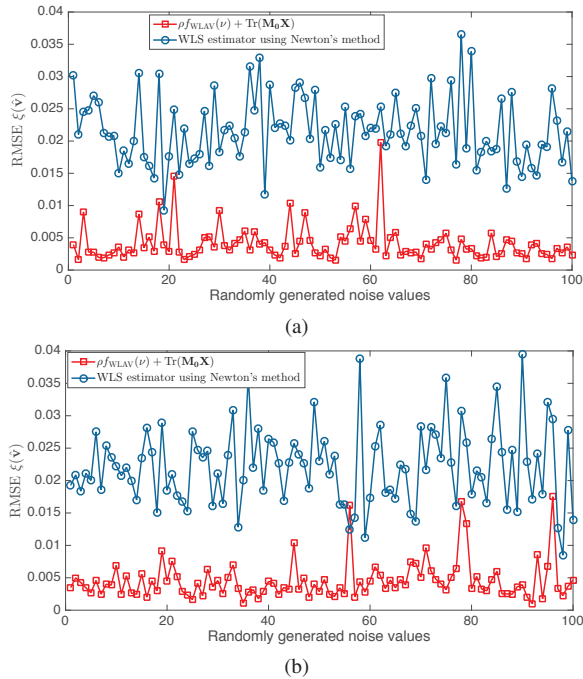For all test cases, it can be observed that the obtained

Fig. 4: The RMSEs of the estimated voltages obtained by the penalized SOCP method and the WLS-based Newton method for the IEEE 57-bus system: (a) voltage magnitude measurements are not available at generator (PV) buses, (b) active and reactive power of all loads are scaled up 10%.

TABLE II: Performance of the penalized SDP (27) with the noise level $c = 0.1$.

Cases	$\xi(\hat{\mathbf{v}})$	ζ	$\zeta^{\max}$	β	λ	f_{WLAV}	$\rho^{\min}$
9-bus	0.0357	0.0462	0.4237	0.9779	1.3417	11.250	0.0482
14-bus	0.0418	0.0537	0.8119	0.9682	0.3812	16.536	0.0532
30-bus	0.0297	0.0405	1.1734	0.9882	0.1094	50.993	0.0222
39-bus	0.0485	0.0676	2.4315	0.9840	0.7438	52.462	0.8173
57-bus	0.0907	0.1028	2.6937	1.0393	0.0912	91.724	0.1028
118-bus	0.0559	0.0743	4.0302	0.9871	0.0878	184.093	0.2284

optimal solutions of the penalized SDP method yield good estimates of the complex voltages featuring small RMSEs $\xi(\hat{\mathbf{v}})$ and ζ. These two error metrics are roughly on the same order as the corresponding noise levels. Furthermore, the value of $\zeta^{\max}$ is calculated using the quantities ρ and λ. As expected, this is a legitimate upper bound on ζ that corroborates our theoretical results in Theorem 2. The tightness of this upper bound depends on the second smallest eigenvalue of the dual matrix $\mathbf{H}(\hat{\boldsymbol{\mu}})$, which is a function of the true state $\mathbf{v}$ and the matrix $\mathbf{M}_0$. The discrepancy between ζ and $\zeta^{\max}$ is rooted in the fact that ζ corresponds to our realization of noise, but $\zeta^{\max}$ works for all realizations of the noise independent of its statistical properties. Moreover, the value of the scaling factor β (see (31)) is always very close to 1 for all scenarios. This implies that the optimal SDP solution $\mathbf{X}^{\text{opt}}$ is close to the true lifted state $\mathbf{v}\mathbf{v}^*$ without scaling this rank-one matrix.

To further show the merit of the proposed penalized SDP framework, we compare the performance of the convex problem (27) against two other estimation techniques. To this end, consider three convex programs that are obtained from

TABLE III: The average RMSEs of the estimated voltage vector $\hat{\mathbf{v}}$ obtained by the penalized SDP (27) for six different objective functions with the noise level $c = 0.1$.

Methods	$\rho f(\boldsymbol{\nu}) + \text{Tr}(\mathbf{M}_0\mathbf{X})$		$\rho f(\boldsymbol{\nu}) + \|\mathbf{X}\|_*$		$\rho f(\boldsymbol{\nu})$	
	WLAV	WLS	WLAV	WLS	WLAV	WLS
9-bus	0.0648	0.1293	1.2744	1.1483	1.1619	1.1633
14-bus	0.1307	0.1784	1.1320	1.3871	1.4233	1.4215
30-bus	0.2055	0.2543	1.4236	1.4306	1.4269	1.4268
39-bus	0.1324	0.1239	1.1317	1.3135	1.2764	1.2757
57-bus	0.2343	0.2809	1.2981	1.3004	1.3235	1.3098
118-bus	0.1136	0.1641	1.3620	1.3272	1.3445	1.3577

TABLE IV: Simulation times of the penalized conic relaxations with $f_{\text{WLAV}}(\boldsymbol{\nu})$ and the noise level $c = 0.1$ (the unit is second).

Cases	Solver time	Total time
9-bus	0.89	1.58
14-bus	1.23	2.54
30-bus	1.33	3.21
39-bus	1.56	3.28
57-bus	1.97	4.09
118-bus	2.38	5.63
1354-bus	4.55	9.48
2869-bus	13.17	24.44
9241-bus	58.00	109.14

(27) by changing its objective to: (i) $\rho f(\boldsymbol{\nu}) + \text{Tr}(\mathbf{M}_0\mathbf{X})$, (ii) $\rho f(\boldsymbol{\nu}) + \|\mathbf{X}\|_*$ (see [48] and [63]), and (iii) $\rho f(\boldsymbol{\nu})$ (see [46], [47], [49], [50]). Each of these methods is tested for both WLAV and WLS functions. Furthermore, 10% of the measurements are generated as bad data to show the robustness of WLAV compared with WLS. These bad data are simulated by adding uniformly distributed random numbers (over the interval $[0, 2]$) to the original measurements. Table III reports the RMSE $\xi(\hat{\mathbf{v}})$ averaged over 50 Monte-Carlo simulations for each test case, where the parameter ρ is set to 0.1. The penalized SDP method proposed in this work clearly outperforms the other techniques.

To show the scalability of the proposed approaches, we conduct simulations on large-scale systems by solving the penalized SDP or SOCP relaxations. Figure 5 shows the effect of additional measurements on reducing the estimation error. In Figures 5(a) and 5(b), the RMSEs of the estimated voltage vectors $\hat{\mathbf{v}}$ are depicted for four different objective functions with respect to the percentage of nodes having measured active power injections. The measurements are under two samples of the noise $\boldsymbol{\eta}$ corresponding to $c = 0.01$ and $c = 0.02$. It can be observed that the quality of the estimation improves with the increase of nodal active power measurements. Even in the case when the number of measurements is limited and close to the number of unknown parameters, the proposed approach can still produce good estimates. In contrast, the methods with no penalty yield very high errors that are out of the plot ranges.

In Figure 5(c) for all four curves, it is assumed that the voltage magnitudes at all buses and active power flows in one direction for all branches are measured. Moreover, different percentages of nodes are chosen at which nodal active and reactive power measurements are made simultaneously. The noise level is set to $c = 0.01$ and the weight is $\rho = 5$. It can be observed that the quality of the estimation improves by increasing the number of additional measurements. The
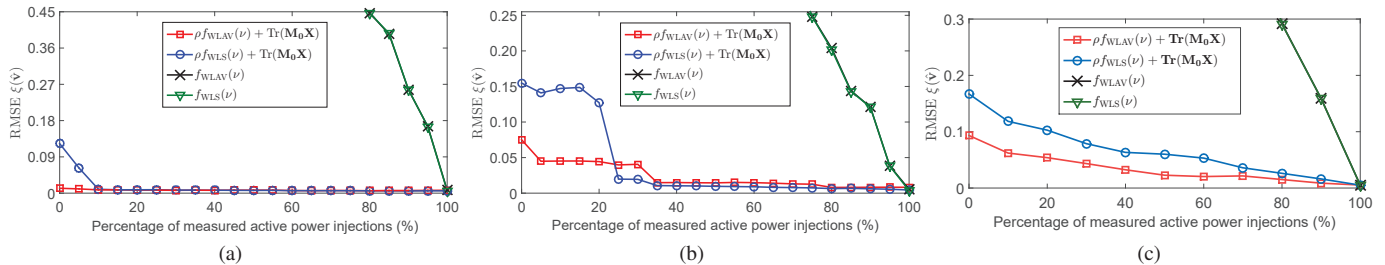
Fig. 5: The RMSEs of the estimated voltages obtained with additional nodal power measurements: (a) $c = 0.01$ and (b) $c = 0.02$. Both are tested on PEGASE 1354-bus system using the penalized SDP; (c) $c = 0.01$ with PEGASE 9241-bus system using the penalized SOCP.

RMSE value at each data point is the average over 10 Monte-Carlo simulations for different noise realizations and choices of nodes with measured power injections.

Finally, Table IV lists the simulation time of the proposed conic relaxations. The total time is obtained by the command `cvx_cputime`, which includes both CVX modeling time and solver time [64]. For all benchmark systems from 9-bus to 118-bus, the SDP relaxation problem is solved by `SDPT3 4.0` [65]. The simulation time is obtained by averaging over 50 Monte-Carlo simulations, which are tested on a macOS system with 2.7GHz Intel Core i5 and 8GB memory. For the last three large-scale test cases, we utilize the SOCP relaxation with the solver `MOSEK 7.0` [66]. The simulation time corresponds to a single run, which is tested on a Windows system with 2.20GHz CPU and 12GB RAM. Clearly, it only takes a few seconds for each case (except the last one) to yield an optimal solution. Even for the large-scale 9241-bus network, the solver time for the proposed SOCP is less than 1 minute, which is fairly practical in real-world applications.

VI. CONCLUSIONS

In this paper, a convex optimization framework is developed for solving the non-convex PF and PSSE problems. To efficiently solve these two problems, the quadratic power flow equations are lifted into a higher-dimensional space, which enables their formulation as linear functions of a rank-one positive semidefinite matrix variable. By meticulously designing an objective function, the PF feasibility problem is converted into a non-convex optimization problem and then relaxed to a convex program. The performance of the proposed convexification is studied in the case where the set of measurements includes: (i) nodal voltage magnitudes, and (ii) one active power flow per line for a spanning tree of the power network. It is shown that the designed convex problem finds the correct solution of the PF problem as long as the voltage angle differences across the lines of the network are not too large. This result along with the proposed framework is then extended to the PSSE problem. Aside from the well-designed objective function for dealing with the non-convexity of PF, a data fitting penalty based on the weighted least absolute value is included to account for the noisy measurements. This leads to a penalized conic optimization scheme. The distance between the optimal solution of the proposed convex problem and the unknown state of the system is quantified in terms of

the noise level, which decays as the number of measurements increases. Extensive numerical results tested on benchmark systems corroborate our theoretical analysis. Moreover, compared with the conventional WLS-based Newton's method as well as other convex programs with different regularizers, the proposed approaches have significant performance gains in terms of the RMSE of the estimated voltages.

APPENDIX

A. Proof of Lemma 1

Let $\mu_1, \ldots, \mu_N \in \mathbb{R}$ and $\mu_{N+1}, \ldots, \mu_M \in \mathbb{R}$ be the Lagrange multipliers associated with the constraints (14b) and (14c), respectively. A key observation is that each matrix $\mathbf{Y}_{l,pf}$ has only three possible nonzero entries:

$$\mathbf{Y}_{l,pf}(s,t) = \mathbf{Y}_{l,pf}^*(t,s) = -\frac{y_{st}}{2}, \quad \mathbf{Y}_{l,pf}(s,s) = \mathrm{Re}(y_{st}).$$

Hence, in order to design $\hat{\boldsymbol{\mu}} \in \mathcal{D}(\mathbf{v})$, we first construct a vector $\hat{\boldsymbol{\mu}}^{(l)}$ for each $l \in \mathcal{L}'$ and then show that the summation

$$\hat{\boldsymbol{\mu}} := \hat{\boldsymbol{\mu}}^{(1)} + \hat{\boldsymbol{\mu}}^{(2)} + \cdots + \hat{\boldsymbol{\mu}}^{(|\mathcal{L}'|)} \tag{38}$$

satisfies the requirements for being a dual SDP certificate. Let $\{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_M\}$ denote the canonical vectors in $\mathbb{R}^M$. Under Assumption 1, it is possible to set

$$\hat{\boldsymbol{\mu}}^{(l)} := \hat{\mu}_s^{(l)} \tilde{\mathbf{e}}_s + \hat{\mu}_t^{(l)} \tilde{\mathbf{e}}_t + \hat{\mu}_{N+l}^{(l)} \tilde{\mathbf{e}}_{N+l} \tag{39}$$

for each line l that connects node s to node t, such that

$$\hat{\mu}_{N+l}^{(l)} := \frac{2\mathrm{Im}(v_s v_t^* M_{0;st}^*)}{\mathrm{Im}(v_s v_t^* y_{st}^*)}, \quad \hat{\mu}_t^{(l)} := -\frac{|v_s|^2 \mathrm{Im}(M_{0;st} y_{st}^*)}{\mathrm{Im}(v_s v_t^* y_{st}^*)},$$

$$\hat{\mu}_s^{(l)} := \frac{|v_t|^2}{|v_s|^2} \hat{\mu}_t^{(l)} - \mathrm{Re}(y_{st}) \hat{\mu}_{N+l}^{(l)}. \tag{40}$$

In order to prove that $\hat{\boldsymbol{\mu}}$ satisfies the conditions in (12), define

$$\widehat{\mathbf{H}}^{(l)} := \mathbf{M}_0^{(l)} + \hat{\mu}_s^{(l)} \mathbf{E}_s + \hat{\mu}_t^{(l)} \mathbf{E}_t + \hat{\mu}_{N+l}^{(l)} \mathbf{Y}_{l,p_f} \tag{41}$$

for every $l = (s,t) \in \mathcal{L}_s$, where

$$\mathbf{M}_0^{(l)} := M_{0;st} \mathbf{e}_s \mathbf{e}_t^\top + M_{0;st}^* \mathbf{e}_t \mathbf{e}_s^\top. \tag{42}$$

It is easy to verify that the relations

$$\widehat{\mathbf{H}}^{(l)} \mathbf{v} = \mathbf{0}, \qquad \widehat{\mathbf{H}}^{(l)} \succeq \mathbf{0} \tag{43}$$

hold for every $l \in \mathcal{L}'$. Therefore, the matrix

$$\mathbf{H}(\hat{\boldsymbol{\mu}}) = \sum_{l \in \mathcal{L}'} \widehat{\mathbf{H}}^{(l)} \tag{44}$$

satisfies the first two conditions in (12).

It remains to prove that $\mathrm{rank}(\mathbf{H}(\hat{\boldsymbol{\mu}})) = N - 1$, or equivalently $\dim(\mathrm{null}(\mathbf{H}(\hat{\boldsymbol{\mu}}))) = 1$. Let $\mathbf{x}$ be an arbitrary member of $\mathrm{null}(\mathbf{H}(\hat{\boldsymbol{\mu}}))$. Since each matrix $\widehat{\mathbf{H}}^{(l)}$ is positive semidefinite, we have $\widehat{\mathbf{H}}^{(l)}\mathbf{x} = \mathbf{0}$ for every $l \in \mathcal{L}'$. As a result, it follows from (40) and (43) that $\frac{x_s}{x_t} = \frac{v_s}{v_t}$ for $l = (s,t)$. By the same reasoning, the relation $\frac{x_t}{x_a} = \frac{v_t}{v_a}$ holds true for each line $l' = (t,a)$. Upon defining $c := \frac{x_s}{v_s}$, one can write

$$\frac{x_s}{v_s} = \frac{x_t}{v_t} = \frac{x_a}{v_a} = c.$$

Repeating the above argument over the connected spanning subgraph $\mathcal{G}'$ through all nodes $k \in \mathcal{N}$ yields that $\mathbf{x} = c\mathbf{v}$, and subsequently $\mathrm{rank}(\mathbf{H}(\hat{\boldsymbol{\mu}})) = N - 1$.

B. Proof of Theorem 1

By choosing sufficiently large values for the Lagrange multipliers associated with the voltage magnitude measurements (14b), a strictly feasible point can be obtained for the dual problem (10). Therefore, strong duality holds between the primal and dual SDP problems. According to Lemma 1, there exists a dual SDP certificate $\hat{\boldsymbol{\mu}} \in \mathcal{D}(\mathbf{v})$ that satisfies (12). Therefore,

$$\mathrm{Tr}(\mathbf{H}(\hat{\boldsymbol{\mu}})\mathbf{v}\mathbf{v}^*) = 0, \qquad \mathbf{H}(\hat{\boldsymbol{\mu}}) \succeq \mathbf{0}. \tag{45}$$

This certifies the optimality of the point $\mathbf{X} = \mathbf{v}\mathbf{v}^*$ for the SDP relaxation problem (14). Moreover, the property $\mathrm{rank}(\mathbf{H}(\hat{\boldsymbol{\mu}})) = N - 1$ justifies the uniqueness of the primal SDP solution $\mathbf{X}^{\mathrm{opt}}$.

In order to prove recovery through the SOCP relaxation problem (15), it is useful to derive the dual SOCP problem:

$$\underset{\substack{\boldsymbol{\mu} \in \mathbb{R}^M \\ \mathbf{F}^{(1)}, \ldots, \mathbf{F}^{|\mathcal{L}'|} \in \mathbb{H}^2}}{\mathrm{maximize}} \quad -\mathbf{z}^\top \boldsymbol{\mu} \tag{46a}$$

$$\text{subject to} \sum_{l=(s,t) \in \mathcal{L}'} \begin{bmatrix} \tilde{\mathbf{e}}_s^\top \\ \tilde{\mathbf{e}}_t^\top \end{bmatrix}^\top \mathbf{F}^{(l)} \begin{bmatrix} \tilde{\mathbf{e}}_s^\top \\ \tilde{\mathbf{e}}_t^\top \end{bmatrix} = \mathbf{M}_0 + \sum_{j=1}^M \mu_j \mathbf{M}_j, \tag{46b}$$

$$\mathbf{F}^{(l)} \succeq \mathbf{0}, \qquad \forall\, l \in \mathcal{L}', \tag{46c}$$

where $\mu_1, \ldots, \mu_N \in \mathbb{R}$ and $\mu_{N+1}, \ldots, \mu_M \in \mathbb{R}$ are the Lagrange multipliers associated with the constraints (15b) and (15c) respectively, and each 2×2 matrix $\mathbf{F}^{(l)}$ is the Lagrange multiplier associated with its corresponding second-order cone constraint in (15d). Let $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}$, and set $\mathbf{F}^{(l)} := \widehat{\mathbf{H}}^{(l)}[(s,t),(s,t)]$ for every $l \in \mathcal{L}'$. As before, the primal and dual feasibility conditions are satisfied, and in addition complementary slackness

$$\mathrm{Tr}\left(\mathbf{F}^{(l)} \begin{bmatrix} |v_s|^2 & v_s v_t^* \\ v_t v_s^* & |v_t|^2 \end{bmatrix}\right) = 0, \quad \forall\, (s,t) \in \mathcal{L}' \tag{47}$$

holds between the dual feasible point $(\boldsymbol{\mu}, \{\mathbf{F}^{(l)}\}_{l=1}^M)$ and the primal feasible point $\mathbf{v}\mathbf{v}^*$. Let $\mathbf{X}^{\mathrm{opt}}$ be an arbitrary solution of (15). One can write:

$$\mathrm{Tr}\left(\mathbf{F}^{(l)} \begin{bmatrix} X_{s,s}^{\mathrm{opt}} & X_{s,t}^{\mathrm{opt}} \\ X_{t,s}^{\mathrm{opt}} & X_{t,t}^{\mathrm{opt}} \end{bmatrix}\right) = 0, \quad \forall\, (s,t) \in \mathcal{L}'. \tag{48}$$

Both matrices $\mathbf{F}^{(l)}$ and $\begin{bmatrix} X_{s,s}^{\mathrm{opt}} & X_{s,t}^{\mathrm{opt}} \\ X_{t,s}^{\mathrm{opt}} & X_{t,t}^{\mathrm{opt}} \end{bmatrix}$ are positive semidefinite and besides

$$\mathrm{null}(\mathbf{F}^{(l)}) = \{c \times [v_s\ v_t]^\top \mid c \in \mathbb{R}\}, \tag{49}$$

where $l = (s,t)$. As a result,

$$\begin{bmatrix} X_{s,s}^{\mathrm{opt}} & X_{s,t}^{\mathrm{opt}} \\ X_{t,s}^{\mathrm{opt}} & X_{t,t}^{\mathrm{opt}} \end{bmatrix} = c' \times \begin{bmatrix} |v_s|^2 & v_s v_t^* \\ v_t v_s^* & |v_t|^2 \end{bmatrix} \tag{50}$$

holds for some $c' > 0$. On the other hand, according to primal feasibility, we have $X_{s,s}^{\mathrm{opt}} = |v_s|^2$ and $X_{t,t}^{\mathrm{opt}} = |v_t|^2$, which means that $c' = 1$. Therefore, all diagonal and those off-diagonal entries of $\mathbf{X}^{\mathrm{opt}}$ associated with the members of $\mathcal{L}'$ are equal to their corresponding entries in $\mathbf{v}\mathbf{v}^*$.

C. Proof of Theorem 2

Observe that

$$\mathrm{Tr}(\mathbf{M}_0\mathbf{X}^{\mathrm{opt}}) + \rho \sum_{j=1}^M \sigma_j^{-1} \big| \mathrm{Tr}(\mathbf{M}_j(\mathbf{X}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*)) \big| - \rho f_{\mathrm{WLAV}}(\boldsymbol{\eta})$$

$$\overset{(a)}{\leq} \mathrm{Tr}(\mathbf{M}_0\mathbf{X}^{\mathrm{opt}}) + \rho \sum_{j=1}^M \sigma_j^{-1} \big| \mathrm{Tr}\left(\mathbf{M}_j(\mathbf{X}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*)\right) - \eta_j \big|$$

$$\overset{(b)}{\leq} \mathrm{Tr}(\mathbf{M}_0\mathbf{v}\mathbf{v}^*) + \rho f_{\mathrm{WLAV}}(\boldsymbol{\eta}), \tag{51}$$

where the relation (a) follows from a triangle inequality and the inequality (b) is obtained by evaluating the objective of (29) at the feasible point $\mathbf{v}\mathbf{v}^*$. Therefore, we have

$$\mathrm{Tr}\left(\mathbf{M}_0(\mathbf{X}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*)\right) + \rho \sum_{j=1}^M \sigma_j^{-1} \left| \mathrm{Tr}\left(\mathbf{M}_j(\mathbf{X}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*)\right) \right|$$

$$\leq 2\rho f_{\mathrm{WLAV}}(\boldsymbol{\eta}). \tag{52}$$

Recall that

$$\mathbf{M}_0 = \mathbf{H}(\hat{\boldsymbol{\mu}}) - \sum_{j=1}^M \hat{\mu}_j \mathbf{M}_j, \quad \mathbf{H}(\hat{\boldsymbol{\mu}})\mathbf{v} = \mathbf{0}.$$

Upon defining $\vartheta_j := \mathrm{Tr}\left(\mathbf{M}_j(\mathbf{X}^{\mathrm{opt}} - \mathbf{v}\mathbf{v}^*)\right)$, one can write

$$\sum_{j=1}^M (\rho \sigma_j^{-1} |\vartheta_j| - \hat{\mu}_j \vartheta_j) + \mathrm{Tr}(\mathbf{H}(\hat{\boldsymbol{\mu}})\mathbf{X}^{\mathrm{opt}}) \leq 2\rho f_{\mathrm{WLAV}}(\boldsymbol{\eta}). \tag{53}$$

Hence, it follows from (30) that

$$\mathrm{Tr}(\mathbf{H}(\hat{\boldsymbol{\mu}})\mathbf{X}^{\mathrm{opt}}) \leq 2\rho f_{\mathrm{WLAV}}(\boldsymbol{\eta}). \tag{54}$$

Now, consider the eigenvalue decomposition of $\mathbf{H}(\hat{\boldsymbol{\mu}}) = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^*$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$ collects the eigenvalues of $\mathbf{H}(\hat{\boldsymbol{\mu}})$ that are sorted in descending order. The matrix $\mathbf{U}$ is a unitary matrix whose columns are the corresponding eigenvectors. Define

$$\check{\mathbf{X}} := \begin{bmatrix} \tilde{\mathbf{X}} & \tilde{\mathbf{x}} \\ \tilde{\mathbf{x}}^* & \alpha \end{bmatrix} = \mathbf{U}^*\mathbf{X}^{\mathrm{opt}}\mathbf{U}, \tag{55}$$

where $\tilde{\mathbf{X}} \in \mathbb{H}_+^{N-1}$ is the $(N-1)$-th order leading principal submatrix of $\check{\mathbf{X}}$. It can be concluded from (54) that

$$2\rho f_{\mathrm{WLAV}}(\boldsymbol{\eta}) \geq \mathrm{Tr}(\mathbf{H}(\hat{\boldsymbol{\mu}})\mathbf{X}^{\mathrm{opt}}) = \mathrm{Tr}(\boldsymbol{\Lambda}\check{\mathbf{X}}) \geq \lambda_2 \mathrm{Tr}(\tilde{\mathbf{X}}),$$

where the last inequality follows from the fact that $\mathrm{rank}(\mathbf{H}(\hat{\boldsymbol{\mu}})) = N - 1$. Therefore, upper bounds for the trace and Frobenius norm of the matrix $\tilde{\mathbf{X}}$ can be obtained as

$$\|\tilde{\mathbf{X}}\|_F \leq \mathrm{Tr}(\tilde{\mathbf{X}}) \leq \frac{2\rho}{\lambda_2} f_{\mathrm{WLAV}}(\boldsymbol{\eta}).$$

By defining $\tilde{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|_2$, the matrix $\mathbf{X}^{\mathrm{opt}}$ can be decomposed as

$$\mathbf{X}^{\mathrm{opt}} = \mathbf{U}\check{\mathbf{X}}\mathbf{U}^* = \begin{bmatrix} \tilde{\mathbf{U}} & \tilde{\mathbf{v}} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{X}} & \tilde{\mathbf{x}} \\ \tilde{\mathbf{x}}^* & \alpha \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{U}}^* \\ \tilde{\mathbf{v}}^* \end{bmatrix}$$
$$= \tilde{\mathbf{U}}\tilde{\mathbf{X}}\tilde{\mathbf{U}}^* + \tilde{\mathbf{v}}\tilde{\mathbf{x}}^*\tilde{\mathbf{U}}^* + \tilde{\mathbf{U}}\tilde{\mathbf{x}}\tilde{\mathbf{v}}^* + \alpha\tilde{\mathbf{v}}\tilde{\mathbf{v}}^*. \quad (56)$$

Since $\check{\mathbf{X}}$ is positive semidefinite, the Schur complement dictates the relation $\tilde{\mathbf{X}} - \alpha^{-1}\tilde{\mathbf{x}}\tilde{\mathbf{x}}^* \succeq \mathbf{0}$. Using the fact that $\alpha = \mathrm{Tr}(\mathbf{X}^{\mathrm{opt}}) - \mathrm{Tr}(\tilde{\mathbf{X}})$, one can write

$$\|\tilde{\mathbf{x}}\|_2^2 \leq \alpha\mathrm{Tr}(\tilde{\mathbf{X}}) = \mathrm{Tr}(\mathbf{X}^{\mathrm{opt}})\mathrm{Tr}(\tilde{\mathbf{X}}) - \mathrm{Tr}^2(\tilde{\mathbf{X}}). \quad (57)$$

Therefore,

$$\|\mathbf{X}^{\mathrm{opt}} - \alpha\tilde{\mathbf{v}}\tilde{\mathbf{v}}^*\|_F^2 = \|\tilde{\mathbf{U}}\tilde{\mathbf{X}}\tilde{\mathbf{U}}^*\|_F^2 + 2\|\tilde{\mathbf{v}}\tilde{\mathbf{x}}^*\tilde{\mathbf{U}}^*\|_F^2 \quad (58\text{a})$$
$$= \|\tilde{\mathbf{X}}\|_F^2 + 2\|\tilde{\mathbf{x}}\|_2^2 \quad (58\text{b})$$
$$\leq \|\tilde{\mathbf{X}}\|_F^2 - 2\mathrm{Tr}^2(\tilde{\mathbf{X}}) + 2\mathrm{Tr}(\mathbf{X}^{\mathrm{opt}})\mathrm{Tr}(\tilde{\mathbf{X}}) \quad (58\text{c})$$
$$\leq 2\mathrm{Tr}(\mathbf{X}^{\mathrm{opt}})\mathrm{Tr}(\tilde{\mathbf{X}}) \quad (58\text{d})$$
$$\leq \frac{4\rho f_{\mathrm{WLAV}}(\boldsymbol{\eta})}{\lambda_2}\mathrm{Tr}(\mathbf{X}^{\mathrm{opt}}), \quad (58\text{e})$$

where (58a) follows from the fact that $\tilde{\mathbf{U}}^*\tilde{\mathbf{v}} = \mathbf{0}$, (58b) is due to $\tilde{\mathbf{U}}^*\tilde{\mathbf{U}} = \mathbf{I}_{N-1}$, and (58d) is in light of $\|\tilde{\mathbf{X}}\|_F \leq \mathrm{Tr}(\tilde{\mathbf{X}})$. The proof is completed by choosing β as $\alpha/\|\mathbf{v}\|_2^2$.

D. Proof of Corollary 3

Define $\tilde{\eta}_i := \eta_i/\sigma_i$ for $i = 1, \ldots, M$. Then, $\tilde{\boldsymbol{\eta}}$ is a standard Gaussian random vector and

$$f_{\mathrm{WLAV}}(\boldsymbol{\eta}) = \|\tilde{\boldsymbol{\eta}}\|_1. \quad (59)$$

Applying the Chernoff's bound [67] to the Gaussian vector $\tilde{\boldsymbol{\eta}}$ yields that

$$\mathbb{P}(\|\tilde{\boldsymbol{\eta}}\|_1 > t) \leq e^{-\psi t}\mathbb{E}\,e^{\psi\|\tilde{\boldsymbol{\eta}}\|_1} = e^{-\psi t}(\mathbb{E}\,e^{\psi|\tilde{\eta}_1|})^M$$
$$= e^{-\psi t}\left(\sqrt{\frac{2}{\pi}}\int_0^\infty e^{\psi x}e^{-x^2/2}\,\mathrm{d}x\right)^M$$
$$= e^{-\psi t}\left(e^{\frac{\psi^2}{2}}\mathrm{erfc}\left(\frac{-\psi}{\sqrt{2}}\right)\right)^M \leq 2^M e^{(M\psi^2 - 2\psi t)/2}, \quad (60)$$

which holds for every $\psi > 0$. Note that the complementary error function $\mathrm{erfc}(a) := \frac{2}{\sqrt{\pi}}\int_a^\infty e^{-x^2}\,\mathrm{d}x \leq 2$ holds for all $a \in \mathbb{R}$. The minimization of the upper bound (60) with respect to ψ gives the equation $\psi^{\mathrm{opt}} = \frac{t}{M}$. Now, it follows from Theorem 2 that

$$\mathbb{P}(\zeta > t) \leq \mathbb{P}\left(2\sqrt{\frac{\rho\|\tilde{\boldsymbol{\eta}}\|_1}{N\lambda}} > t\right) = \mathbb{P}\left(\|\tilde{\boldsymbol{\eta}}\|_1 > \frac{t^2 N\lambda}{4\rho}\right)$$
$$\leq \exp\left(M\ln 2 - \frac{t^4 N^2\lambda^2}{32M\rho^2}\right). \quad (61)$$

The proof is completed by substituting $N = \frac{M}{\kappa}$ into (61).

REFERENCES

[1] A. Wood, B. Wollenberg, and G. Sheblé, *Power generation, Operation, and Control*, 3rd ed. Hoboken, NJ: Wiley, 2013.

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

[3] R. Hartshorne, *Algebraic Geometry*. New York City, NY: Springer, 1977.

[4] D. Molzahn, D. Mehta, and M. Niemerg, "Toward topologically based upper bounds on the number of power flow solutions," in *Proc. of American Control Conf.*, Boston, MA, Jul. 2016, pp. 5927–5932.

[5] D. Bienstock and A. Verma, "Strong NP-hardness of AC power flows feasibility," Dec. 2015. [Online]. Available: http://arxiv.org/pdf/1512.07315v1.pdf

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

[7] A. Bergen and V. Vittal, *Power Systems Analysis*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2000.

[8] W. Tinney and C. Hart, "Power flow solution by Newton's method," *IEEE Trans. Power App. Syst.*, vol. PAS-86, no. 11, pp. 1449–1460, Nov. 1967.

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

[10] A. Trias, "The holomorphic embedding load flow method," in *Proc. of IEEE Power & Energy Society General Meeting*, San Diego, CA, Jul. 2012, pp. 1–8.

[11] T. Li, "Numerical solution of polynomial systems by homotopy continuation methods," in *Handbook of Numerical Analysis*, F. Cucker, Ed. Amsterdam, Netherlands: North-Holland, 2003, vol. 11, pp. 209–304.

[12] K. Dvijotham, M. Chertkov, and S. Low, "A differential analysis of the power flow equations," in *Proc. of IEEE Conf. on Decision and Control*, Osaka, Japan, Dec. 2015, pp. 23–30.

[13] K. Dvijotham, S. Low, and M. Chertkov, "Solving the power flow equations: A monotone operator approach," Jul. 2015. [Online]. Available: http://arxiv.org/pdf/1506.08472v2.pdf

[14] D. Mehta, D. Molzahn, and K. Turitsyn, "Recent advances in computational methods for the power flow equations," in *Proc. of American Control Conf.*, Boston, MA, Jul. 2016, pp. 1753–1765.

[15] R. Madani, J. Lavaei, and R. Baldick, "Convexification of power flow problem over arbitrary networks," in *Proc. of IEEE Conf. on Decision and Control*, Osaka, Japan, Dec. 2015, pp. 1–8.

[16] A. Monticelli, *State Estimation in Electric Power Systems: A Generalized Approach*, ser. Power Electronics and Power Systems. New York City, NY: Springer, 2012.

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

[18] Å. Björck, *Numerical methods for least squares problems*. Philadelphia, PA: SIAM, 1996.

[19] W. Mascarenhas, "The divergence of the BFGS and Gauss Newton methods," *Math. Program.*, vol. 147, no. 1, pp. 253–276, Oct. 2014.

[20] G. Zyskind and F. Martin, "On best linear estimation and a general Gauss-Markov theorem in linear models with arbitrary nonnegative covariance structure," *SIAM J. Appl. Math.*, vol. 17, no. 6, pp. 1190–1202, Nov. 1969.

[21] M. Amini, A. Sarwat, S. Iyengar, and I. Guvenc, "Determination of the minimum-variance unbiased estimator for DC power-flow estimation," in *40th Annual Conf. of the IEEE Ind. Electron. Society*, Dallas, TX, Oct. 2014, pp. 114–118.

[22] M. Amini, M. Ilić, and O. Karabasoglu, "DC power flow estimation utilizing Bayesian-based LMMSE estimator," in *IEEE PES General Meeting*, Denver, CO, Jul. 2015, pp. 1–5.

[23] M. Irving, R. Owen, and M. J. Sterling, "Power-system state estimation using linear programming," *Electrical Engineers, Proceedings of the Institution of*, vol. 125, no. 9, pp. 879–885, Sep. 1978.

[24] W. Kotiuga and M. Vidyasagar, "Bad data rejection properties of weighted least absolute value techniques applied to static state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-101, no. 4, pp. 844–851, Apr. 1982.

[25] H. Singh and F. Alvarado, "Weighted least absolute value state estimation using interior point methods," *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1478–1484, Aug. 1994.

[26] A. Abur and M. Celik, "A fast algorithm for the weighted least absolute value state estimation," *IEEE Trans. Power Syst.*, vol. 6, no. 1, pp. 1–8, Feb. 1994.

[27] M. Celik and A. Abur, "A robust WLAV state estimator using transformations," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 106–113, Feb. 1992.

[28] M. Irving, "Robust state estimation using mixed integer programming," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1519–1520, Aug. 2008.

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

[30] A. Conejo, S. de la Torre, and M. Canas, "An optimization approach to multiarea state estimation," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 213–221, Feb. 2007.

[31] A. Minot, Y. Lu, and N. Li, "A distributed Gauss-Newton method for power system state estimation," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3804–3815, Sep. 2016.

[32] J. Dopazo, O. Klitin, G. Stagg, and L. VanSlyck, "State calculation of power systems from line flow measurements," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 7, pp. 1698–1708, Sep./Oct. 1970.

[33] ——, "State calculation of power systems from line flow measurements, part II," *IEEE Trans. Power App. Syst.*, vol. PAS-91, no. 7, pp. 145–151, Jan. 1972.

[34] B. Porretta and R. Dhillon, "Performance evaluation of state estimation from line flow measurements on Ontario hydro power system," *IEEE Trans. Power App. Syst.*, vol. PAS-92, no. 5, pp. 1702–1712, Sep. 1973.

[35] S. Naka, T. Genji, T. Yura, and Y. Fukuyama, "A hybrid particle swarm optimization for distribution state estimation," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 60–68, Feb. 2003.

[36] K. Lee and M. El-Sharkawi, Eds., *Modern Heuristic Optimization Techniques: Theory and Applications to Power Systems*. Hoboken, NJ: Wiley-IEEE Press, 2008.

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

[38] J. Lavaei, "Zero duality gap for classical OPF problem convexifies fundamental nonlinear power problems," in *Proc. of American Control Conf.*, San Francisco, CA, Jun. 2011, pp. 4566–4573.

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

[40] S. Sojoudi and J. Lavaei, "Physics of power networks makes hard optimization problems easy to solve," in *Proc. of IEEE Power & Energy Society General Meeting*, San Diego, CA, Jul. 2012, pp. 1–8.

[41] ——, "Exactness of semidefinite relaxations for nonlinear optimization problems with underlying graph structure," *SIAM J. Optim.*, vol. 24, no. 4, pp. 1746–1778, Oct. 2014.

[42] R. Madani, S. Sojoudi, and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 199–211, Jan. 2015.

[43] R. Madani, M. Ashraphijuo, and J. Lavaei, "Promises of conic relaxation for contingency-constrained optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1297–1307, Mar. 2016.

[44] S. Low, "Convex relaxation of optimal power flow – part I: Formulations and equivalence," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 15–27, Mar. 2014.

[45] ——, "Convex relaxation of optimal power flow – part II: Exactness," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 2, pp. 177–189, Jun. 2014.

[46] H. Zhu and G. Giannakis, "Estimating the state of AC power systems using semidefinite programming," in *Proc. of North American Power Symp.*, Boston, MA, Aug. 2011, pp. 1–7.

[47] Y. Weng, Q. Li, R. Negi, and M. Ilić, "Semidefinite programming for power system state estimation," in *Proc. of IEEE Power & Energy Society General Meeting*, San Diego, CA, Jul. 2012, pp. 1–8.

[48] Y. Weng, M. Ilić, Q. Li, and R. Negi, "Convexification of bad data and topology error detection and identification problems in AC electric power systems," *IET Gener. Transm. Dis.*, vol. 9, no. 16, pp. 2760–2767, Nov. 2015.

[49] Y. Weng, Q. Li, R. Negi, and M. Ilić, "Distributed algorithm for SDP state estimation," in *Proc. of Innovative Smart Grid Tech.*, Washington, D.C., Feb. 2013, pp. 1–6.

[50] H. Zhu and G. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1039–1050, Dec. 2014.

[51] S. Kim, G. Wang, and G. Giannakis, "Online semidefinite programming for power system state estimation," in *Proc. of IEEE Conf. on Acoustics, Speech and Sigal Process.*, Florence, Italy, May 2014, pp. 6024–6027.

[52] B. Weedy, B. Cory, N. Jenkins, J. Ekanayake, and G. Strbac, *Electric Power Systems*, 5th ed. Hoboken, NJ: John Wiley & Sons, 2012.

[53] G. Andersson, "Modelling and analysis of electric power systems," Sep. 2008, [Online]. Available: http://www.eeh.ee.ethz.ch/uploads/tx_ethstudies/modelling_hs08_script_02.pdf.

[54] G. Korres and N. Manousakis, "State estimation and bad data processing for systems including PMU and SCADA measurements," *Electric Power Systems Research*, vol. 81, pp. 1514–1524, Apr. 2011.

[55] A. Phadke and J. Thorp, *Synchronized Phasor Measurements and Their Applications*. New York City, NY: Springer, 2008.

[56] A. Abur and F. H. Magnago, "Optimal meter placement for maintaining observability during single branch outages," *IEEE Trans. Power Syst.*, vol. 14, no. 4, pp. 1273–1278, Nov. 1999.

[57] Y. Wu, M. Kezunovic, and T. Kostic, "Cost minimization in power system measurement placement," in *Intl. Conf. on Power Syst. Tech.*, Chongqing, China, Oct. 2006, pp. 1–6.

[58] J. Chow, Ed., *Power System Coherency and Model Reduction*. New York City, NY: Springer, 2013.

[59] R. Grone, C. Johnson, E. Sá, and H. Wolkowicz, "Positive definite completions of partial hermitian matrices," *Linear Algebra Appl.*, vol. 58, pp. 109–124, Apr. 1984.

[60] C. Josz, S. Fliscounakis, J. Maeght, and P. Panciatici, "AC power flow data in MATPOWER and QCQP format: iTesla, RTE snapshots, and PEGASE," Mar. 2016, [Online]. Available: arxiv.org/abs/1603.01533.

[61] R. Zimmerman, C. Murillo-Sánchez, and R. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[62] MathWorks, "Bioinformatics toolbox," [Online]. Available: https://www.mathworks.com/help/bioinfo/ref/graphminspantree.html.

[63] S.-J. Kim, "Online power system state estimation using alternating direction method of multipliers," in *Proc. of IEEE Power & Energy Society General Meeting*, Denver, CO, Jul. 2015, pp. 1–5.

[64] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Jun. 2015.

[65] K. Toh, M. Todd, and R. Tutuncu, "SDPT3 – A MATLAB software package for semidefinite programming," *Optim. Method Softw.*, vol. 11, pp. 545–581, 1999.

[66] MOSEK ApS, *The MOSEK Release Notes. Version 7.0 (Revision 141)*, 2016. [Online]. Available: http://docs.mosek.com/7.0/toolsrelnotes/index.html

[67] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford, UK: Oxford University Press, 2013.