Nonlinear Least Absolute Value Estimator for Topology Error Detection and Robust State Estimation .

SangWoo Park, Reza Mohammadi-Ghazi, and Javad Lavaei

***Abstract*—Topology error, a modeling misrepresentation of the power system network configuration, can impair the quality of state estimation. In this paper, we propose a new technique for robust state estimation in the presence of a small number of topological errors for power systems modeled by AC power flow equations. The developed method leverages the availability of a large volume of SCADA measurements and minimizes the ℓ_1 norm of nonconvex residuals augmented by a nonlinear, but convex, regularizer. Noting that a power network can be represented by a graph, we first study the properties of the solution obtained by the proposed estimator and argue that, under mild conditions, this solution identifies a small subgraph of the network that contains the topological errors in the model used for the state estimation problem. Then, we propose a method that can efficiently detect the topological errors by searching over the identified subgraph. Furthermore, we develop a theoretical upper bound on the state estimation error to guarantee the accuracy of the proposed state estimation technique. The efficacy of the developed framework is demonstrated through numerical simulations on IEEE benchmark systems.**

***Index Terms*—Topological error, Nonlinear least absolute value, Local search method, State estimation**

I. Introduction

SAFEGUARDING energy infrastructures against progressive failures of stressed components is an important challenge in operating these systems and preventing blackouts [2], [3]. In doing so, the power system condition should be continuously monitored so that, if needed, required actions can be taken. This condition monitoring is performed through real-time state estimation that aims to recover the underlying system voltage phasors, given supervisory control and data acquisition (SCADA) measurements and a model that encompasses the system topology and specifications [4], [5]. In fact, state estimation not only helps prevent failures in the power network, but it also underpins every aspect of real-time power system operation and control. To ensure an accurate state estimation, it is essential to have the capability of detecting bad data. Assuming that the network parameters are known and the measurement devices are correctly calibrated, the main source of bad data is topological errors in the model. Topological errors refer to the inaccurate modeling of the current network configuration and are often initiated by the misconception of the system operator about the on/off switching status of a few lines in the network due to faults or unreported network reconfigurations. Due to their significant impact on the accuracy of state estimation, dealing with bad data and identifying topological errors have received considerable attentions in the past few years.

The authors are with the Department of Industrial Engineering and Operation Research, University of California Berkeley, CA 94710. Emails: {spark111, mohammadi, lavaei}@berkeley.edu

This work was supported by the ONR grant N00014-17-1-2933, DARPA grant D16AP00002, AFOSR grant FA9550- 17-1-0163, NSF grant 1552089 and ARO grant W911NF-17-1-0555

This manuscript is an extended version of the conference paper [1] with new additions including an updated Theorem 2 and comprehensive case studies.

A. Previous studies on topological error detection

The existing topological error detection methods take either a statistical or a geometric approach. Bayesian hypothesis testing [6], collinearity tests [7], and fuzzy pattern machine [8] are examples of statistical approaches for topology error detection. These methods usually need prior information about states and/or a decently-sized dataset from previous measurements.

The geometric approaches, on the other hand, aim to design state estimation techniques that are robust against topological errors and measurement noise. Using normalized Lagrange multipliers of the least-squares state estimation problem is one such technique that has been shown to be effective in some cases [9], although it is a heuristic method. Recent studies, such as the one proposed in [10], improve this technique; however, they may fail to detect certain scenarios called *critical parameter and measurement pairs*. Another important approach in this category is the least absolute value (LAV) estimator introduced in [11] for power systems. By minimizing the ℓ_1 norm of the linearized measurement residual vector, the LAV is capable of finding a minimum set of measurements free of large errors, thus rejecting bad data and yielding a robust estimate. Despite its robustness, the LAV is vulnerable to leverage points as explained in [12], [13]. Further investigation and suggestion of different methods to mitigate this issue have been presented in [14], [15]. In [16], the authors have shown that the effect of leverage points can be eliminated if measurements consist only of phasor measurement units (PMUs). The caveat of these methods is their reliance on a linearized DC model. Only few studies have addressed the fully nonlinear, non-convex problem with power measurements, e.g., [17] where a semidefinite programming (SDP) relaxation is proposed to convexify the nonlinear LAV state estimator; however, no theoretical guarantees have been developed to ensure the recovery of a high-quality solution. Moreover, the computational demand of solving the surrogate SDP problem may restrict the application of this method to small-sized problems in practice. These issues motivate further research on developing robust state estimation techniques with the capability of handling nonconvexities associated with various types of measurements.

B. Contributions

In light of the recently developed theoretical guarantees for the ℓ_2-norm to avoid spurious local solutions in nonconvex

optimization [18] and arising promises for the ℓ_1-norm [19], this study proposes a local search algorithm to find the global solution of the nonlinear LAV (NLAV) state estimator. The proposed method provides a robust approach for estimating the system's states in presence of a modest number of topological errors as well as detecting such errors. In doing so, the main contributions of this work can be summarized as: (1) proposing an algorithm for detecting modest topological errors and finding the state of the power system using an NLAV state estimator with local search algorithms, (2) formulating a regularized NLAV state estimator to handle severe nonconvexities, (3) finding error bounds and necessary properties for the regularization parameters. As explained later in this paper, local search algorithms would efficiently find global solutions of the underlying NLAV estimators given a sufficient number of noiseless measurements and a proper initialization of the algorithm. Also, many of the implications provided in this paper are all valid even if one uses an SDP relaxation of the proposed nonconvex estimators. The remainder of this paper is organized as follows. Preliminary materials are presented in Section II, followed by formulation of the algorithm in Section III. A comprehensive set of numerical simulations on the IEEE 57-bus system is presented in Section IV. Concluding remarks are drawn in Section V. The proofs are provided in the Appendix.

C. Notations

Throughout this paper, lower (resp. upper) case letters represent column vectors (resp. matrices) and calligraphic letters stand for sets and graphs. 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 sets of $N \times N$ complex symmetric matrices and Hermitian matrices, respectively. The symbols $(\cdot)^T$ and $(\cdot)^*$ denote the transpose and conjugate transpose of a vector or matrix. Re$(\cdot)$, Im$(\cdot)$, rank$(\cdot)$ and Tr$(\cdot)$ denote the real part, imaginary part, rank and trace of a given scalar or matrix. The notations $\|x\|_1, \|x\|_2$ and $\|X\|_F$ denote the ℓ_1-norm and ℓ_2-norm of vector x respectively, and the Frobenius norm of matrix X. The symbol $\langle X, Y \rangle$ denotes the Frobenius inner product of the matrices X and Y. The symbol $|\cdot|$ is the absolute value operator if the argument is a scalar, vector, or matrix; otherwise, it is the cardinality of a measurable set. The relation $X \succeq 0$ means that the matrix X is Hermitian positive semidefinite. The (i, j) entry of X is denoted by $X_{i,j}$. The notation $X[\mathcal{S}_1, \mathcal{S}_2]$ denotes the submatrix of X whose rows and columns are chosen from the index sets $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively. I_N shows the $N \times N$ identity matrix. The symbol diag(x) denotes a diagonal matrix whose diagonal entries are given by the vector x, whereas diag(X) forms a column vector by extracting the diagonal entries of the matrix X. The imaginary unit is denoted by $\mathbf{j} = \sqrt{-1}$. The symbol $\mathbb{1}$ denotes a vector of all ones with appropriate dimension. $\lambda_i(X)$ denotes the i-th smallest eigenvalue of the matrix X. Given a graph $\mathcal{G}$, the notation $\mathcal{G}(\mathcal{V}, \mathcal{E})$ implies that $\mathcal{V}$ and $\mathcal{E}$ are the vertex set and the edge set of this graph, respectively.

II. Preliminaries

Consider an electric power network represented by a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, \ldots, n\}$ and $\mathcal{E} := \{1, \ldots, r\}$ denote the sets of buses and branches, respectively. Also, suppose that the the slack bus is also the reference bus. Let $v_k \in \mathbb{C}$ denote the nodal complex voltage at bus $k \in \mathcal{V}$, whose magnitude and phase are given as $|v_k|$ and $\measuredangle v_k$. The net injected complex power at bus k is denoted as $s_k = p_k + q_k \mathbf{j}$. Define $s_{l,f} = p_{l,f} + q_{l,f}\mathbf{j}$ (resp. $i_{l,f}$) and $s_{l,t} = p_{l,t} + q_{l,t}\mathbf{j}$ (resp. $i_{l,t}$) as the complex power flows (resp. currents) entering the line $l \in \mathcal{E}$ through the 'from' and 'to' end of the branch. Note that the currents $i_{l,f}$ and $i_{l,t}$ may not add up to zero due to the existence of transformers and shunt capacitors. Let v and i be the vectors of nodal complex voltages and net current injections, respectively. The Ohm's law dictates that

$$i = Yv, \quad i_f = Y_f v, \quad \text{and} \quad i_t = Y_t v, \tag{1}$$

where $Y = G + B\mathbf{j} \in \mathbb{C}^{n \times n}$ is the admittance matrix of the power network, whose real and imaginary parts are the conductance matrix G and susceptance matrix B, respectively. Furthermore, $Y_f \in \mathbb{C}^{r \times n}$ and $Y_t \in \mathbb{C}^{r \times n}$ represent the 'from' and 'to' branch admittance matrices. The injected complex power can thus be expressed as $p + q\mathbf{j}$ = diag(vv^*Y^*). Let $\{e_1, \ldots, e_n\}$ denote the canonical vectors in $\mathbb{R}^n$. Define

$$\begin{gathered} E_k := e_k e_k^T, \quad Y_{k,p} := \frac{1}{2}(Y^* E_k + E_k Y), \\ Y_{k,q} := \frac{\mathbf{j}}{2}(E_k Y - Y^* E_k). \end{gathered} \tag{2}$$

Moreover, let $\{d_1, \ldots, d_r\}$ be the canonical vectors in $\mathbb{R}^r$. Define $Y_{l,p_f}, Y_{l,p_t}, Y_{l,q_f}$ and Y_{l,q_t} associated with the l-th branch from node i to node j as

$$\begin{gathered} Y_{l,p_f} := \frac{1}{2}(Y_f^* d_l e_i^T + e_i d_l^T Y_f), \; Y_{l,p_t} := \frac{1}{2}(Y_t^* d_l e_j^T + e_j d_l^T Y_t) \\ Y_{l,q_f} := \frac{\mathbf{j}}{2}(e_j d_l^T Y_f - Y_f^* d_l e_i^T), \; Y_{l,q_t} := \frac{\mathbf{j}}{2}(e_j d_l^T Y_t - Y_t^* d_l e_i^T) \end{gathered} \tag{3}$$

The traditional measurable quantities can be expressed as

$$|v_k|^2 = \text{Tr}(E_k vv^*) \tag{4a}$$

$$p_k = \text{Tr}(Y_{k,p} vv^*), \quad q_k = \text{Tr}(Y_{k,q} vv^*) \tag{4b}$$

$$p_{l,f} = \text{Tr}(Y_{l,p_f} vv^*), \quad p_{l,t} = \text{Tr}(Y_{l,p_t} vv^*) \tag{4c}$$

$$q_{l,f} = \text{Tr}(Y_{l,q_f} vv^*), \quad q_{l,t} = \text{Tr}(Y_{l,q_t} vv^*) \tag{4d}$$

These equations show that the nodal and line measurements can be expressed as simple quadratic functions of the complex voltage vector v. In this paper, we only focus on traditional voltage and power measurements. However, if we have linear PMU measurements (e.g., certain entries of v and i), they can be regarded as quadratic equations with a zero quadratic term and the results of this paper are all valid in this scenario as well. To proceed with the paper, we lay out several definitions. First, we define the *set of measurement matrices* based on (2)–(4):

Definition 1. *Given a power system model Ω characterized by the tuple (Y, Y_f, Y_t) and an index set of measurements $\mathcal{M} = \{1, \ldots, m\}$ specifying m measurements of the form* (4)*,*

the mapping from the measurement index set to the set of measurement matrices is defined as

$$C^{\Omega}(\mathcal{M}) \triangleq \{M_j(\Omega)\}_{j\in\mathcal{M}} \tag{5}$$

where each $M_j(\Omega)$ corresponds to one of the matrices E_k, $Y_{k,p}$, $Y_{k,q}$, Y_{l,p_f}, Y_{l,p_t}, Y_{l,q_f}, Y_{l,q_t} defined in (2) and (3), depending on the type of measurement j.

Second, we define the real-valued state vector and the corresponding real-valued matrices. This enables us to solve optimization problems involving complex voltages in the real-domain. The dimension of the real-valued state vector is $2n-1$ because the voltage angle at the slack/reference bus is fixed to be zero. Accordingly, the matrices also have $2n-1$ rows and columns.

Definition 2. *Given a complex-valued state vector $v \in \mathbb{C}^n$, define $\bar{v} \triangleq [Re\{v[\mathcal{Q}]^T\}\ Im\{v[\mathcal{O}]^T\}]^T \in \mathbb{R}^{2n-1}$ as the real-valued state vector of the power system's operating point with $\mathcal{Q}$ denoting the set of all buses and $\mathcal{O}$ indicating the set of all buses except for the slack bus. Furthermore, define $\bar{X} \in \mathbb{S}^{2n-1}$ as the real-valued symmetrization of $X \in \mathbb{H}^n$. To further clarify this notation, note that a general $n \times n$ Hermitian matrix can be mapped into a $(2n-1)\times(2n-1)$ real-valued symmetric matrix as follows:*

$$\bar{X} = \begin{bmatrix} Re\{X[\mathcal{Q},\mathcal{Q}]\} & -Im\{X[\mathcal{Q},\mathcal{O}]\} \\ Im\{X[\mathcal{O},\mathcal{Q}]\} & Re\{X[\mathcal{O},\mathcal{O}]\} \end{bmatrix} \tag{6}$$

Finally, we define an operator that maps the state vector to the vector of measurement values, and also its Jacobian:

Definition 3. *Given a system model Ω and a set of measurements $\mathcal{M}$, define the function $h^{\Omega}(\bar{v}) : \mathbb{R}^{2n-1} \to \mathbb{R}^m$ as the mapping from the real-valued state of the power system to the vector of noiseless measurement values:*

$$h^{\Omega}(\bar{v}) \triangleq [v^T M_1(\Omega) v \ \cdots \ v^T M_m(\Omega) v]^T \tag{7}$$

$$= [\bar{v}^T \bar{M}_1(\Omega)\bar{v} \ \cdots \ \bar{v}^T \bar{M}_m(\Omega)\bar{v}]^T \tag{8}$$

Furthermore, define $J^{\Omega}(\bar{v}) \in \mathbb{R}^{(2n-1)\times m}$ to be the Jacobian of $h^{\Omega}(\bar{v})$. In other words,

$$J^{\Omega}(\bar{v}) = 2[\bar{M}_1(\Omega)\bar{v} \quad \bar{M}_2(\Omega)\bar{v} \quad \cdots \quad \bar{M}_m(\Omega)\bar{v}] \tag{9}$$

III. Main Results

In this section, we first briefly discuss the most widely used nonlinear least-squares (NLS) state estimation formulation and its limitations. Then, we present the *NLAV* formulation and provide a theoretical upper bound on the state estimation error obtained by the NLAV problem. Finally, we uncover certain properties of the vector of residual errors and develop a novel algorithm that jointly performs state estimation and topology error detection.

A. Nonlinear least-squares state estimation

NLS is the most common state estimation technique, which was first proposed by Schweppe [21], [22]. Recent studies have shown that local search algorithms, such as Gauss-Newton, are able to find a global solution of this nonconvex problem in the case where the number of measurements is relatively higher than the degree of the freedom of the system and the measurements are noiseless [5], [18]. Similar to other estimators, this method requires that the system's measurement matrices (see Definition 1) be known. However, the model that power system operators use may be different from the true system due to the presence of topological errors arising from faults or recent changes in the switching status of some lines. The measurements at the vicinity of the incorrectly modeled lines are potentially the outliers, which can impact the solution of the state estimation problem over a large portion of the network. This is due to the incapability of the ℓ_2-norm in dealing with outliers and simulation results supporting this fact are shown in Figure 3 followed by further discussions in Section IV-B. Despite the drawbacks of *NLS*, the work [6] develops an effective tool for topology error detection using Bayesian-based hypothesis testing and the covariance matrix of the states. The method that we propose in this paper does not require the covariance information but takes advantage of the favorable aspects of ℓ_1-norm minimization.

B. Proposed NLAV formulation

A line whose presence in the system is misrepresented by the system operator is called *erroneous* in the remainder of this paper, and the set of all *erroneous* lines is denoted by Ξ. Let $C^{\widetilde{\Omega}}(\mathcal{M})$ be the set of measurement matrices corresponding to the model $\widetilde{\Omega}$ that the power system operator has access to, and $C^{\Omega}(\mathcal{M})$ be the set of measurement matrices corresponding to the true system Ω. Assume that Ω and $\widetilde{\Omega}$ are sparsely different in the sense that there is a small subset of lines in the system for which the operator misunderstands their on/off statuses. In this work, we only focus on sparse errors for two reasons. If Ω and $\widetilde{\Omega}$ are relatively different, then the state is not observable from a static set of measurements and dynamic time-stamped data is required. Second, topological errors often occur due to low probability events and it is unlikely that the operator's model be significantly different from the true model. To design an algorithm that jointly performs state estimation and sparse topological error detection, we propose the following optimization problem:

$$\min_{\bar{v}\in\mathbb{R}^{2n-1}} f(\bar{v}) \tag{10}$$

where

$$f(\bar{v}) = \bar{v}^T \bar{M}_0 \bar{v} + \rho \sum_{j=1}^{m} |\bar{v}^T \bar{M}_j(\widetilde{\Omega})\bar{v} - b_j| \tag{11}$$

and b_j is the j^{th} element of the measurement vector $b \in \mathbb{R}^m$ that is

$$b = h^{\Omega}(\bar{z}) + \eta. \tag{12}$$

Here, $\bar{z} \in \mathbb{R}^{2n-1}$ denotes the true underlying state of the system and η is the noise vector. Notice that the measurement values b are based on the true system Ω and the true system state $\bar{z}$. Also, $M_0 \in \mathbb{S}^n$ is a regularization matrix and ρ is a regularization coefficient. As will be discussed later, these two parameters assist with the convexification of the problem for finding a robust solution using local search algorithms. From

here on, we assume that the measurement set $\mathcal{M}$ is observable. A necessary condition for observability is that the Jacobian of the measurement equations (i.e. $J^{\Omega}(\bar{z})$) be full row rank [23]. Let $\bar{v}_*$ denote a globally optimal solution of (10), as in

$$\bar{v}_* = \text{argmin}_{\bar{v}\in\mathbb{R}^{2n-1}} f(\bar{v}) \tag{13}$$

Then, let $\epsilon \in \mathbb{R}^{2n-1}$ be the state estimation error vector and $r \in \mathbb{R}^m$ be the residual error vector, defined as

$$\epsilon = \bar{v}_* - \bar{z} \tag{14a}$$

$$r_j = |\bar{v}_*^T \bar{M}_j(\widetilde{\Omega})\bar{v}_* - b_j|, \quad \forall j \in \mathcal{M} \tag{14b}$$

The problem (10) aims to push the insignificant residual errors to hard zeros, while some of the r_j's associated with the outlier measurements are expected to remain nonzero. This phenomenon is supported empirically via an example in Figure 3(d). The performance of this estimator has a striking contrast with that of the ℓ_2 minimization (Figure 3(c)) where the residuals are spread out throughout all the measurements. In the remainder of this article, we use this intuition to design an efficient topological error detection algorithm. Note that, similar to the NLS state estimation, the objective function of (10) is nonlinear and nonconvex, which makes local search algorithms prone to falling into spurious local solutions. However, recent studies have shown that increasing the number of redundant measurements helps with reducing the non-convexity of NLS problems and hence, improves the likelihood of finding their global solutions using local search algorithms [5], [18]. Therefore, having access to many measurements is the key for real-world state estimation problems. This property is expected to hold for the NLAV estimator too, as partially proven in [19]. The risk of becoming stuck at a local optimum is further avoided by starting the algorithm close to the unknown state. This is possible because in power systems, voltage magnitudes are kept close to 1 and voltage angles are maintained to be small. Therefore, choosing the initial point to be the nominal point $\mathbb{1}$ would likely ensure that it is relatively close to the true state.

C. Estimation error

Given a design matrix M_0, we intend to prove a theoretical upper bound on the state estimation error obtained by the *NLAV* problem (10). To this end, it is useful to introduce the concept of dual certificate:

Definition 4. *Given a positive-semidefinite regularization matrix $M_0 \in \mathbb{S}^n$, a system model Ω, and a set of measurement matrices $C_\Omega(\mathcal{M})$, define $H_\mu^\Omega \triangleq M_0 + \sum_{j=1}^m \mu_j M_j(\Omega)$. A vector $\mu \in \mathbb{R}^m$ is called a* dual certificate *for the voltage vector $v \in \mathbb{C}^n$ of the system model Ω if it satisfies the following three conditions:*

$$H_\mu^\Omega \succeq 0, \quad H_\mu^\Omega v = 0, \quad \textit{rank}\{H_\mu^\Omega\} = n-1 \tag{15}$$

In essence, the existence of a dual certificate ensures that the second-smallest eigenvalue of H_μ^Ω is strictly positive, which enables us to derive an upper-bound of the form presented in the following theorem.

Theorem 1. *Suppose that the power system operator has a network model $\widetilde{\Omega}$ and a set $\mathcal{M}$ of measurement indices. For this model, assume that there exists a dual certificate μ for the true state vector z. Also, consider a parameter ρ satisfying the inequality $\rho \geq \max_{j\in\mathcal{M}} |\mu_j|$. Then, there exists a real-valued scalar β such that*

$$\frac{\|\bar{v}_* - \beta\cdot\bar{z}\|_2^2}{\|\bar{v}_*\|_2} \leq \sqrt{\frac{4n\cdot g(\bar{z},\eta,\rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}})}} \tag{16}$$

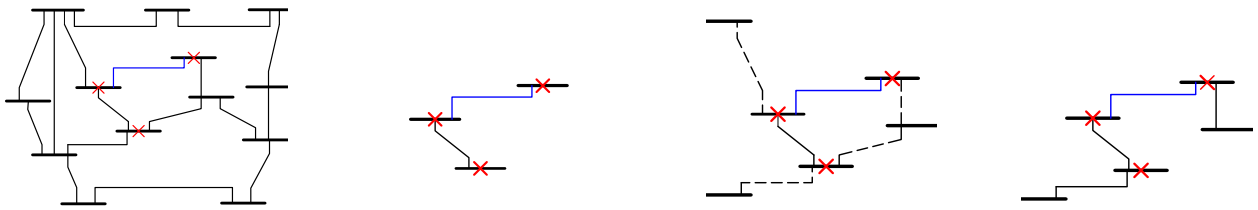
where $g(\bar{z},\eta,\rho)$ is equal to

$$\rho\left[\sum_{j\in\mathcal{N}} |\bar{z}^T(\bar{M}_j(\Omega) - \bar{M}_j(\widetilde{\Omega}))\bar{z}| + \sum_{j=1}^m |\eta_j|\right] \tag{17}$$

By recalling that $\bar{z}$ and $\bar{v}_*$ are, respectively, the true and recovered states of the system, the above inequality quantitatively bounds the error of the state estimation. There are several important characteristics of this bound. First, if there is no topology error and the measurements are noiseless, NLAV recovers a high-quality solution if not the actual state. On the other hand, if there are topology error and measurement noise, the upper bound for the state estimation error increases proportionally to the number of topology errors and the magnitude of noise. Furthermore, the upper-bound is a decreasing function of the second smallest eigenvalue of the matrix $H_\mu^{\widetilde{\Omega}}$, which acts as the Laplacian of a weighted graph obtained from the power network. The second smallest eigenvalue of this matrix, also called the algebraic connectivity [24], is a parameter that measures how well connected a weighted graph is. For example, a complete graph has the algebraic connectivity of n while this value is equal to 2 for a star graph and $2(1-cos\frac{\pi}{n})$ for a path graph (where n denotes the number of nodes in the graph). In the special case when M_0 reflects the connectivity of the original network $\mathcal{G}$ (i.e., $i \neq j$ and $(i,j) \notin \mathcal{E} \implies M_0(i,j) = 0$), the second smallest eigenvalue of $H_\mu^{\widetilde{\Omega}}$ represents the algebraic connectivity of the original network with different weights assigned to different edges. Finally, note that the bound in equation (16) does not guarantee a unique solution of the NLAV. For conditions that guarantee the uniqueness of solution, the reader is referred to Theorem 3.

D. Sparse suspect-subgraph

As shown above, the quality of the state estimation deteriorates under the presence of topological errors. Our approach for detecting and correcting these topological errors can be summarized as follows. First, we solve (10) and use the pattern of the nonzero residuals errors to find a (small) subset of lines that are potentially *erroneous* in the model. We call this subset the *suspect-subgraph*, which we then efficiently search through to identify the topological errors. This is followed by a correction of the model and a re-estimation of the system states. To formalize this approach, we first introduce some relevant subgraphs.



(a) (b) (c) (d)

Figure 1: (a) A power network, (b) the state estimation error graph $\mathcal{S}$, (c) the extended state estimation error graph $\widetilde{\mathcal{S}}$, and (d) the line residual graph $\mathcal{R}^L$. Unsolvable nodes are marked with red crosses, the blue line is the only erroneous line, and the dotted lines correspond to the edges added to $\mathcal{S}$ to obtain $\widetilde{\mathcal{S}}$. The graphic of the node residual graph $\mathcal{R}^N$ is omitted.

Definition 5. *A node $i \in \mathcal{V}$ is called* unsolvable *if ϵ_i is nonzero. On the other hand, if ϵ_i is zero, node i is called* solvable. *Define the following four induced subgraphs*[1] *of $\mathcal{G}$:*

1) *The* state estimation error graph $\mathcal{S}(\mathcal{V}_\mathcal{S}, \mathcal{E}_\mathcal{S})$ *is such that $\mathcal{V}_\mathcal{S}$ is the set of unsolvable nodes.*
2) *The* extended state estimation error graph $\widetilde{\mathcal{S}}(\mathcal{V}_{\widetilde{\mathcal{S}}}, \mathcal{E}_{\widetilde{\mathcal{S}}})$ *is such that $\mathcal{V}_{\widetilde{\mathcal{S}}}$ includes all nodes in $\mathcal{V}_\mathcal{S}$ and also those nodes that are adjacent to any node in $\mathcal{V}_\mathcal{S}$.*
3) *The* node residual graph $\mathcal{R}^N(\mathcal{V}_\mathcal{N}, \mathcal{E}_\mathcal{N})$ *is such that $\mathcal{V}_\mathcal{N}$ is the set of nodes whose associated entries in r are nonzero.*
4) *The* line residual graph $\mathcal{R}^L(\mathcal{V}_\mathcal{L}, \mathcal{E}_\mathcal{L})$ *is such that $\mathcal{V}_\mathcal{L}$ is the set of nodes that are either at the 'from' or 'to' end of a line whose associated entry in r is nonzero.*

In order to help the reader visualize the different subgraphs, we illustrate Definition 5 for a small system in Figure 1. In Theorem 2, we reveal how the *set of erroneous lines*, namely Ξ, relates to these subgraphs.

Theorem 2. *Suppose that the noise vector η is zero. In addition, assume that there do not exist any two distinct vectors of voltages resulting in the same measurement values, i.e.,*

$$\bar{x} \neq \bar{y} \Longrightarrow \|h^\Omega(\bar{x}) - h^\Omega(\bar{y})\|_1 \neq 0 \tag{18}$$

Then,

$$\mathcal{R}^L \subseteq (\Xi^c \cap \widetilde{\mathcal{S}}) \tag{19}$$

Moreover, if no two erroneous lines share the same node, the following statements hold:

$$\mathcal{R}^N = \widetilde{\mathcal{S}} \cup \Xi \tag{20}$$

The relationships between different subgraphs are illustrated in Figure 2. It is important to note that due to the sparsity of the state estimation error (as shown in Figure 3(b)) and the sparsity assumption on Ξ, most lines belong to the set $\widetilde{\mathcal{S}}^c \cap \Xi^c$. From the diagram, we can also easily infer that $\Xi \subseteq (\mathcal{R}^N \setminus \mathcal{R}^L) \subseteq (\widetilde{\mathcal{S}}^c \cap \Xi^c)^c$. Therefore, the practical benefit of Theorem 2 is that it enables us to develop a technique for efficiently detecting topological errors by searching over a small subgraph of the original power network. We call this small subgraph, namely $(\mathcal{R}^N \setminus \mathcal{R}^L)$, the *suspect-subgraph*.

[1]Note that, by definition, the edge set of an induced subgraph of $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of all edges in $\mathcal{E}$ that have both endpoints in the node set of the induced graph [25].

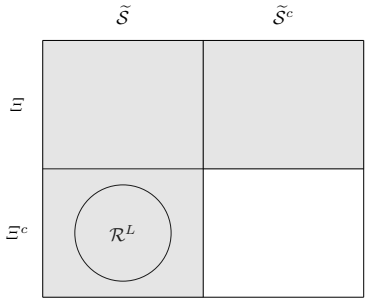


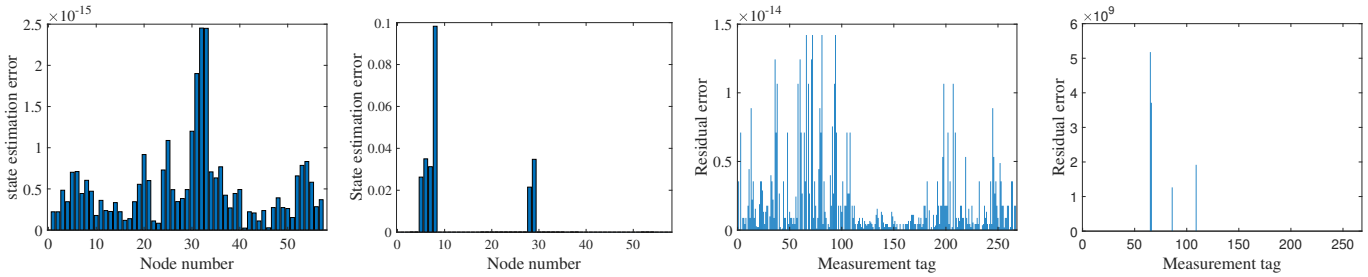
Figure 2: A diagram showing the relationship between different subgraphs. Each rectangle represents the intersection between two different sets. For example, the upper-left rectangle represents $\Xi \cap \widetilde{\mathcal{S}}$. $\mathcal{R}^N$ is equivalent to the gray-colored area.

E. Algorithm

Based on the above results, we propose Algorithm 1 for topological error detection. This algorithm initializes the set of detected erroneous lines, denoted by $\mathcal{D}_L$, with the empty set. Then, the algorithm searches over all branches in the *suspect-subgraph* $(\mathcal{R}^N \setminus \mathcal{R}^L)$, and evaluates the effect of the presence of each line on the accuracy of the solution. In doing so, the proposed method switches the line off if it is on in the model and vice versa, updates the model based on this change, and re-solves the NLAV problem with the updated model. If the norm of residual errors is decreased, the line is added to $\mathcal{D}_L$; otherwise, the change of line status is rejected and the algorithm proceeds to check the next line or terminates if all lines of $(\mathcal{R}^N \setminus \mathcal{R}^L)$ have already been evaluated. Justification for using such an algorithm is provided in the Appendix section D. Algorithm 1 summarizes the proposed topological error detection method.

F. Unpenalized NLAV estimator and unique solution

After all the topological errors have been detected and fixed, a final state estimation based on the correct network topology can be performed. However, this does not necessarily guarantee a recovery of the true state $\bar{z}$. In this subsection, we disregard the regularization term M_0 for simplicity and call this the *unpenalized NLAV* problem (in other words, we set M_0 to 0). Without prior knowledge of the state, designing a favorable M_0 penalty term could be difficult, in which case setting M_0 to zero makes logical sense. Theorem 3 provides



(a) (b) (c) (d)

Figure 3: Noiseless state estimation error for (a) NLS, (b) NLAV; and residuals for (c) NLS, (d) NLAV. Note that in (c) and (d), the x-axis shows the measurement tag, which is not the same as the node or line number due to a random selection of line measurements.

Algorithm 1 Subgraph search algorithm

Given: Hypothetical model $\widetilde{\Omega}$ and measurement vector b
Initialize: Set $\mathcal{D}_L = \emptyset$, $\epsilon > 0$, $\delta > 0$, $\mu > 0$
and calculate $\mathcal{C}_{\widetilde{\Omega}}(\mathcal{M})$ using Definition 1.
1. Solve NLAV problem (10) with $\widetilde{\Omega}$, $\mathcal{C}_{\widetilde{\Omega}}(\mathcal{M})$ and b, and calculate the residual r based on equation (14b).
2. Construct the suspect-subgraph $(\mathcal{R}^N \setminus \mathcal{R}^L)$.
3. Set $r^t \leftarrow r$.
while $\|r^t\|_2 > \delta$ **do**
  $\Omega^t \leftarrow \widetilde{\Omega}$.
  for line $l \in (\mathcal{R}^N \setminus \mathcal{R}^L)$ **do**
    Update Ω^t to $\Omega^{t'}$ by changing the on/off status of l.
    Re-solve (10) with $\Omega^{t'}$, $\mathcal{C}_{\Omega^{t'}}(\mathcal{M})$ and b to obtain the outputs $\bar{v}_*^{update}$ and r^{update}
    if $\|r^{update}\|_2 < \|r^t\|_2$ **then**
      Add l to $\mathcal{D}_L$ and set $\widetilde{\Omega} \leftarrow \Omega^{t'}$, $r^t \leftarrow r^{update}$.
    end if
  end for
end while
4. Return $\bar{v}_*^{update}$ and $\mathcal{D}_L$

a sufficient condition under which the *unpenalized NLAV* problem has a unique solution. Since without M_0, the state estimation error bound provided in Theorem 1 is no longer valid, Theorem 3 also provides a new bound.

Definition 6. *Given a system model Ω and a set of measurements $\mathcal{M}$, define the linear map $\mathcal{A}^\Omega : \mathbb{R}^{n\times n} \to \mathbb{R}^m$ as*

$$\mathcal{A}^\Omega(X) = [\langle M_1(\Omega), X\rangle \cdots \langle M_m(\Omega), X\rangle]^T \tag{21}$$

Theorem 3. *Given the true network model Ω and the measurement set $\mathcal{M} = \{1, \ldots, m\}$, let $\mathcal{A}^\Omega$ be the mapping as defined in Definition 6. Then, $\bar{v}^*$ obtained from solving the NLAV problem (10) with $M_0 = 0$ satisfies:*

$$\|\bar{v}_*\bar{v}_*^T - \bar{z}\bar{z}^T\|_F \le \frac{2}{t}\|\eta\|_1 \tag{22}$$

where t is defined as the optimal objective value of the following optimization problem:

$$\min_{K\in\mathbb{S}^n} \|\mathcal{A}^\Omega(K)\|_2$$
$$s.t. \quad rank(K) = 2,\ \|K\|_F = 1 \tag{23}$$

It is straightforward to verify that $t > 0$ if and only if there does not exist any set of noiseless measurements for the model Ω that leads to non-unique exact solutions. In other words, if $t > 0$, any global optimal of the NLAV is the true state that we wish to find (note that this applies when all topological errors have been detected and fixed). Therefore, t can be viewed as a quantification of the measurements' quality for finding a unique solution of the over-determined power flow equations. In addition, if $t > 0$, then condition(18) is implied.

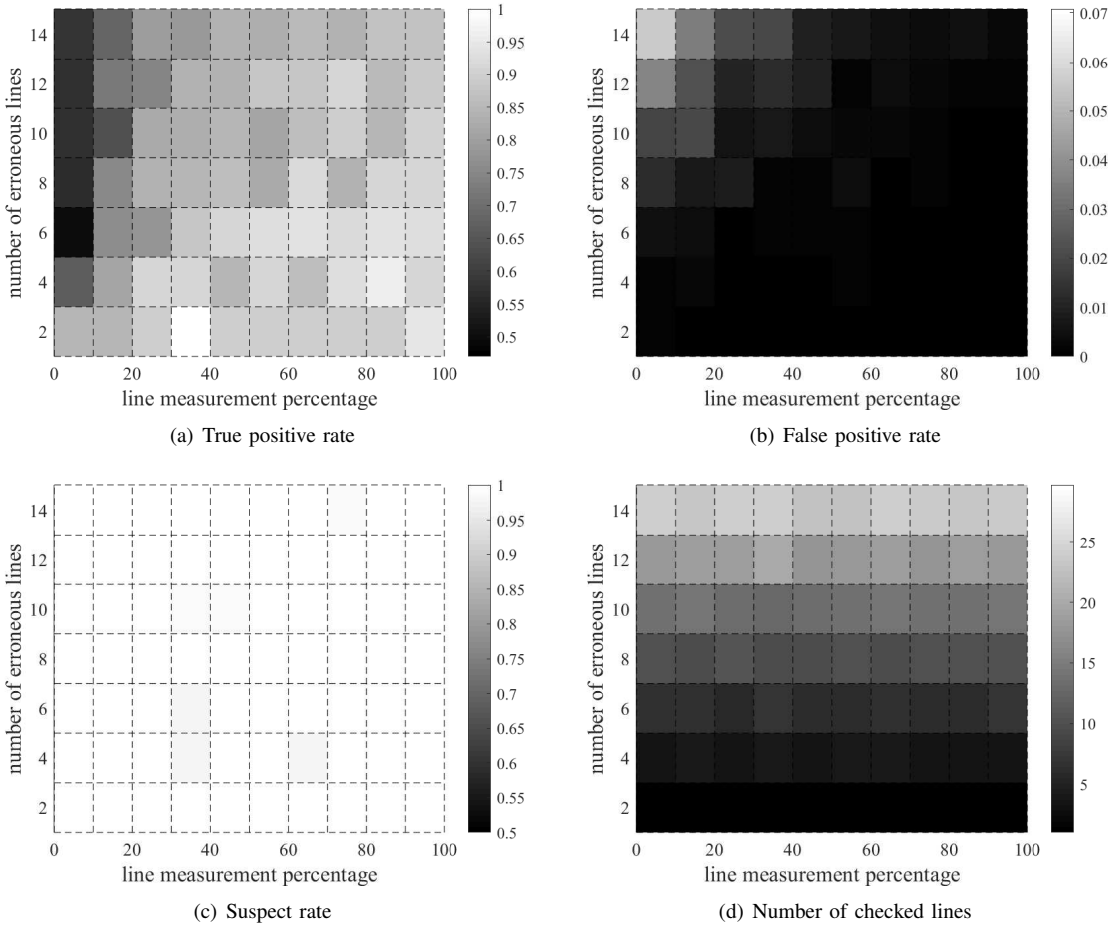
Recently, there has been some study on the connection between the property of *no spurious local minima* and the *restricted isometry property* (RIP). A linear map $\mathcal{A} : \mathbb{R}^{n\times n} \to \mathbb{R}^m$ is said to satisfy (r, δ_r)-RIP with constant $0 \le \delta_r < 1$ if there exists $p > 0$ such that for all rank-r matrices X: $(1-\delta_r)\|X\|_F^2 \le \frac{1}{p}\|\mathcal{A}(X)\|^2 \le (1+\delta_r)\|X\|_F^2$. If $\mathcal{A}$ satisfies $(2r, \delta_{2r})$-RIP with $\delta_{2r} < 1$, then finding a global optimum constitutes exact recovery of the state [26]. However, this does not exclude the existence of *spurious local minima* (local minima that are not globally optimal), which can be problematic when using local search algorithms. In order to guarantee no spurious local minima, $\mathcal{A}$ suffices to satisfy $(2r, \delta_{2r})$-RIP with $\delta_{2r} < 0.2$, which is a strict condition [27]. A milder condition on RIP for structured mappings (such as power subsystems) has been developed in [28]. The parameter t introduced above is clearly related to the RIP constant. In fact, $t > 0$ is equivalent to having $\delta_{2r} < 1$, which implies that there is a unique global solution.

IV. Simulation Results

In order to evaluate the efficacy of the proposed *NLAV* algorithm for detecting topological errors, this section presents numerical simulations on the IEEE 57-bus system. To run the simulations, we use MATPOWER data along with the MATLAB *fmincon* function as the local search algorithm.

A. Simulation setup

In this study we focus on two types of topological errors. **Type I error** is when a transmission line is switched off in the



(a) True positive rate

(b) False positive rate

(c) Suspect rate

(d) Number of checked lines

Figure 4: Simulation results on the IEEE 57-bus system. Each value represents the average over 20 simulations.

true system while it is switched on in the hypothetical model that is accessible to the power system operator; **Type II error** is when a branch is switched on in the true model while it is switched off in the hypothetical model. Our numerical evaluations consist of multiple cases where we vary the number of erroneous lines and the percentage of line measurements that are available. The procedure of running the simulations is as follows: (1) For a given number of erroneous lines and line measurement percentage, we run 20 simulations; (2) In each simulation the erroneous lines are randomly chosen and checked to ensure that they satisfy the system's observability and that they do not share common buses; (3) The type of topological error is also randomly assigned to each selected erroneous line; (4) In all simulations full nodal measurements (p_k, q_k and $|v_k|$) are considered; (5) The line measurements are randomly selected from the intact lines and no measurements are taken from the erroneous ones; (6) To generate a legitimate state, we assume that the voltage magnitudes are close to unity and the angles are small.

In order to assess the performance of the algorithm, we calculate the true/false positive rates and the suspect rate as:

$$\text{True positive rate} = |\mathcal{D}_L \cap \Xi|/|\Xi| \tag{24a}$$

$$\text{False positive rate} = |\mathcal{D}_L \cap \Xi^c|/|\Xi| \tag{24b}$$

$$\text{Suspect rate} = |(\mathcal{R}_N \setminus \mathcal{R}^L) \cap \Xi|/|\Xi| \tag{24c}$$

Finally, we also report the number of lines that the algorithm checks before termination, which is simply the cardinality of the set $(\mathcal{R}_N \setminus \mathcal{R}^L)$.

B. Example: Sparse residuals for NLAV

Before analyzing the bulk of simulations data, we focus on a single example to graphically illustrate the ideas discussed in Section III. This example is under a scenario with two erroneous lines (lines 8 and 67) and 30% line measurements. Figures 3(a) and 3(c) show the state estimation errors and residuals of *NLS* when topological errors exist. It follows from these plots that there is a lack of sparsity pattern, and the high peaks are not even associated with the end points of the *erroneous* lines. This implies that one needs to search over all possible combinations of lines to find the erroneous ones, which is numerically intractable for large systems. In contrast, the state estimation errors and the residuals after the first run of the *NLAV* are shown in Figure 3(b) and 3(d). The highest peaks of the residual vector in this plot correspond to the nodes/lines that are directly connected to the erroneous lines. This implies that the erroneous lines can be found by searching over only the lines that are associated with the highest peaks of the residual vector. By doing so, as stated in Algorithm 1,

both erroneous lines are correctly detected without any false positive detection. In the following subsection, we present a summary of the extensive simulations conducted on the IEEE 57-bus system.

C. 57-bus system

For the 57-bus system, we consider $\{1, 3, \ldots, 15\}$ as the discretized range for the possible number of erroneous lines and $\{0\%, 10\%, 20\%, \ldots, 100\%\}$ as the discretized range for the possible line measurement percentage. Combining these two sets gives the total of 88 scenarios for this system.

Figure 4 shows heat maps of the performance statistics for the above-mentioned 88 scenarios. Figure 4(c) shows that an erroneous line is in the suspect subgraph with high probability. In fact, all of the values are above 0.98, which illustrates that the assumptions made in Theorem 2 are reasonable. Figure 4(a) implies that Algorithm 1 is able to detect most of the erroneous lines given a sufficient number of measurements, and Figure 4(b) indicates that there is close to zero false positives. We can also see that detecting topological errors becomes more difficult as the number of such errors grows. However, note that the number of lines that need to be checked grows only linearly with respect to the number of erroneous lines. More specifically, Figure 4(d) shows that the number of lines to be checked is approximately twice the number of erroneous lines. These results imply that the proposed algorithm is capable of accurately detecting topological errors and therefore provides a tool for robust state estimation if the number of measurements is large enough.

V. Conclusion

This paper proposes a new technique to solve the state estimation problem for power systems in the presence of a modest number of topological errors and to detect such modeling errors. The developed method minimizes a nonconvex function of the ℓ_1-norm of the state estimation residual errors plus a convex quadratic penalty term. It is shown that, under mild conditions, the proposed method can efficiently detect the topological errors by searching over the lines of a (small) suspect-subgraph of the network inferred by the solution of the estimator. Two upper bounds are derived on the estimation errors, and the results are demonstrated on a benchmark system.

Appendix

A. Proof of Theorem 1

Before going into the proof, we impose the following two conditions for M_0:

Assumption 1. *The regularizer matrix M_0 satisfies the following properties:*

1) $M_0 \succeq 0$

2) $M_0 \cdot \mathbb{1} = 0$

Also, we define the set of erroneous measurements:

Definition 7. *Define $\mathcal{N} \in \mathcal{M}$ as the set of indices of the measurements that correspond to the erroneous lines.*

Consider the NLAV problem (10). One can create lower and upper bounds on the optimal objective value as follows:

$$\begin{aligned}
&\bar{v}_*^T \bar{M}_0 \bar{v}_* + \rho \sum_{j=1}^{m} |\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\Omega) \bar{z}| - \rho \sum_{j=1}^{m} |\eta_j| \\
&\overset{(a)}{\leq} \bar{v}_*^T \bar{M}_0 \bar{v}_* + \rho \sum_{j=1}^{m} |\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\Omega) \bar{z} - \eta_j| \\
&\overset{(b)}{\leq} \bar{z}^T \bar{M}_0 \bar{z} + \rho \sum_{j=1}^{m} |\bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z} - \bar{z}^T \bar{M}_j(\Omega) \bar{z} - \eta_j| \\
&\overset{(c)}{\leq} \bar{z}^T \bar{M}_0 \bar{z} + \rho \sum_{j \in \mathcal{N}} |\bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z} - \bar{z}^T \bar{M}_j(\Omega) \bar{z}| + \rho \sum_{j=1}^{m} |\eta_j|
\end{aligned}$$

where (a) is due to the triangle inequality and (b) is due to the optimality of v_*. The equality (c) follows from $\bar{M}_j(\widetilde{\Omega}) = \bar{M}_j(\Omega)$ whenever $j \notin \mathcal{N}$. Combining the above lower and upper bounds leads to

$$\begin{aligned}
&\bar{v}_*^T \bar{M}_0 \bar{v}_* - \bar{z}^T \bar{M}_0 \bar{z} + \rho \sum_{j=1}^{m} |\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\Omega) \bar{z}| \\
&\leq \rho \sum_{j \in \mathcal{N}} |\bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z} - \bar{z}^T \bar{M}_j(\Omega) \bar{z}| + 2\rho \sum_{j=1}^{m} |\eta_j| \qquad (25)
\end{aligned}$$

By adding and subtracting $\bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z}$ in the absolute value of the left-hand side, one can write:

$$\begin{aligned}
&\bar{v}_*^T \bar{M}_0 \bar{v}_* - \bar{z}^T \bar{M}_0 \bar{z} + \rho \sum_{j=1}^{m} |\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z}| \\
&\leq 2\rho \left\{ \sum_{j \in \mathcal{N}} |\bar{z}^T (\bar{M}_j(\widetilde{\Omega}) - \bar{M}_j(\Omega)) \bar{z}| + \sum_{j=1}^{m} |\eta_j| \right\} \\
&= 2g(\bar{z}, \eta, \rho) \qquad (26)
\end{aligned}$$

Now, consider the following optimization problem that serves as a tool for deriving a lower bound:

$$\min_{y} \ \bar{v}_*^T \bar{M}_0 \bar{v}_* - \bar{z}^T \bar{M}_0 \bar{z} + \rho \sum_{j=1}^{m} |\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z}|$$

Here y is a fictitious variable with a dimension of choice, and we call the objective of the above problem as f. By introducing a new variable $t \in \mathbb{R}^m$, an equivalent formulation can be written as

$$\begin{aligned}
\min_{t} \quad & \bar{v}_*^T \bar{M}_0 \bar{v}_* - \bar{z}^T \bar{M}_0 \bar{z} + \rho \sum_{j=1}^{m} t_j \\
s.t. \quad & \bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z} \leq t_j, \quad \forall j \in \mathcal{M} \\
& -\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* + \bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z} \leq t_j, \quad \forall j \in \mathcal{M}
\end{aligned} \qquad (27)$$

Let p_j^+'s and p_j^-'s be the nonnegative Lagrange multipliers for the first and second sets of constraints. The Lagrangian can be written as

$$\begin{aligned}
\mathcal{L}(t, p^+, p^-) &= \bar{v}_*^T \bar{M}_0 \bar{v}_* - \bar{z}^T \bar{M}_0 \bar{z} + \sum_{j=1}^{m} (\rho - p_j^+ - p_j^-) t_j \\
&+ \sum_{j=1}^{m} \{(p_j^+ - p_j^-)(\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z})\} \qquad (28)
\end{aligned}$$

By defining $d(p^+,p^-)=\min_t \mathcal{L}(t,p^+,p^-)$ and noting that $p_j^+ + p_j^- = \rho$ for every $j\in\mathcal{M}$ at optimality, we have

$$d(p^+,p^-) = \bar{v}_*^T\Big(\bar{M}_0 + \sum_{j=1}^{m}(p_j^+ - p_j^-)\bar{M}_j(\widetilde{\Omega})\Big)\bar{v}_* - \bar{z}^T\Big(\bar{M}_0 + \sum_{j=1}^{m}(p_j^+ - p_j^-)\bar{M}_j(\widetilde{\Omega})\Big)\bar{z} \tag{29}$$

Note that $d(p^+,p^-)$ gives a lower bound on f. By assumption, there exists a dual certificate $\mu\in\mathbb{R}^m$. We can find a set of vectors p_*^+ and p_*^- such that they satisfy the previous constraint $p_*^+ + p_*^- = \rho\cdot\mathbb{1}$ and also a new constraint $p_*^+ - p_*^- = \mu$. Then, $d(p_*^+,p_*^-)$ also gives a lower bound to f. Using the fact that $H_\mu^{\widetilde{\Omega}} z = 0$ and defining $X=\bar{v}_*\bar{v}_*^T$, we can establish the following:

$$d(p_*^+,p_*^-) = \bar{v}_*^T H_\mu^{\widetilde{\Omega}}\bar{v}_* - \bar{z}^T H_\mu^{\widetilde{\Omega}}\bar{z} \tag{30}$$

$$= \mathrm{Tr}\{H_\mu^{\widetilde{\Omega}}\bar{v}_*\bar{v}_*^T\} = \mathrm{Tr}\{H_\mu^{\widetilde{\Omega}}X\} \tag{31}$$

The rest of the proof can be adopted from [29] (Appendix, Proof of Theorom 2). Consider an eigen-decomposition of $H_\mu^{\widetilde{\Omega}} = U\Lambda U^T$, where $\Lambda = \mathrm{diag}(\lambda_{2n-1},...,\lambda_1)$ such that $\lambda_{2n-1}\geq\cdots\geq\lambda_1$ and U is a unitary matrix whose columns are the corresponding eigenvectors. Define

$$\breve{X} := \begin{bmatrix}\widetilde{X} & \widetilde{x}\\ \widetilde{x}^T & \alpha\end{bmatrix} = U^T X U \tag{32}$$

where $\widetilde{X}$ is the $(2n-2)^{\text{th}}$-order leading principle submatrix of $\breve{X}$, $\widetilde{x}$ is the $(2n-2)\times 1$ leftover vector and α is a scalar. It is known that

$$Tr\big(H_\mu^{\widetilde{\Omega}}X\big) = Tr\big(U\Lambda U^T U\breve{X}U^T\big) = Tr\big(\Lambda\breve{X}\big) \geq \lambda_2(H_\mu^{\widetilde{\Omega}})Tr\big(\widetilde{X}\big) \tag{33}$$

Combining (33) and (26) leads to

$$\mathrm{Tr}\big(\widetilde{X}\big) \leq 2\cdot g(\bar{z},\eta,\rho)/\lambda_2(H_\mu^{\widetilde{\Omega}}) \tag{34}$$

Define $\widetilde{z}=\bar{z}/\|\bar{z}\|_2$ and $\widetilde{v}_*=\bar{v}_*/\|\bar{v}_*\|_2$. Since $H_\mu^{\widetilde{\Omega}}$ is positive-semidefinite and the eigenvector corresponding to the smallest eigenvalue (i.e. zero) is $\bar{z}$, the matrix X can be decomposed as

$$X = U\breve{X}U^T = \begin{bmatrix}\widetilde{U} & \widetilde{z}\end{bmatrix}\begin{bmatrix}\widetilde{X} & \widetilde{x}\\ \widetilde{x}^T & \alpha\end{bmatrix}\begin{bmatrix}\widetilde{U}^T\\ \widetilde{z}^T\end{bmatrix} = \widetilde{U}\widetilde{X}\widetilde{U}^T + \widetilde{U}\widetilde{x}\widetilde{z}^T + \widetilde{z}\widetilde{x}^T\widetilde{U}^T + \alpha\widetilde{z}\widetilde{z}^T \tag{35}$$

Since $\breve{X}\succeq 0$, Schur complement dictates the relationship $\widetilde{X} - \alpha^{-1}\widetilde{x}\widetilde{x}^T \succeq 0$. Using the fact that $\alpha = \mathrm{Tr}(X) - \mathrm{Tr}(\widetilde{X})$, one can write

$$\|\widetilde{x}\|_2^2 \leq \alpha\mathrm{Tr}(\widetilde{X}) = \mathrm{Tr}(X)\mathrm{Tr}(\widetilde{X}) - \mathrm{Tr}^2(\widetilde{X}) \tag{36}$$

Therefore,

$$\begin{aligned}\|X-\alpha\widetilde{z}\widetilde{z}^T\|_F^2 &= \|\widetilde{U}\widetilde{X}\widetilde{U}^T + \widetilde{U}\widetilde{x}\widetilde{z}^T + \widetilde{z}\widetilde{x}^T\widetilde{U}^T\|_F^2\\ &\overset{(d)}{=} \|\widetilde{U}\widetilde{X}\widetilde{U}^T\|_F^2 + 2\|\widetilde{z}\widetilde{x}^T\widetilde{U}^T\|_F^2 \overset{(e)}{=} \|\widetilde{X}\|_F^2 + 2\|\widetilde{x}\|_2^2\\ &\leq \|\widetilde{X}\|_F^2 - 2\mathrm{Tr}^2(\widetilde{X}) + 2\mathrm{Tr}(X)\mathrm{Tr}(\widetilde{X})\\ &\overset{(f)}{\leq} 2\mathrm{Tr}(X)\mathrm{Tr}(\widetilde{X})\\ &\overset{(g)}{\leq} \frac{4g(\bar{z},\eta,\rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}})}\mathrm{Tr}(X)\end{aligned} \tag{37}$$

where (d) follows from the fact that $\widetilde{U}*\widetilde{z}=0$, (e) is due to $\widetilde{U}^T\widetilde{U}=I_{2n-2}$ and (f) comes from the fact that $\|\widetilde{X}\|_F\leq\mathrm{Tr}(\widetilde{X})$. Finally, (g) results from substituting equation (34). Plugging back in $X=\bar{v}_*\bar{v}_*^T$ yields that

$$\|X-\alpha\widetilde{z}\widetilde{z}^T\|_F^2 = \|\bar{v}_*\bar{v}_*^T - \frac{\alpha}{\|\bar{z}\|_2^2}\bar{z}\bar{z}^T\|_F^2 \leq \frac{4g(\bar{z},\eta,\rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}}(\bar{M}_0))}\mathrm{Tr}(\bar{v}_*\bar{v}_*^T) \tag{38}$$

By defining $\beta=\alpha/\|\bar{z}\|_2^2$ and realizing that $\mathrm{Tr}(\bar{v}_*\bar{v}_*^T)=\|\bar{v}_*\|_2^2$, the above inequality reduces to

$$\|\bar{v}_*\bar{v}_*^T - \beta\bar{z}\bar{z}^T\|_F^2 \leq \frac{4g(\bar{z},\eta,\rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}}(\bar{M}_0))}\|\bar{v}_*\|_2^2 \tag{39}$$

By notational simplicity, we denote $x(i)$ as the i-th element of a vector x. Notice that

$$\begin{aligned}\|\bar{v}_*\bar{v}_*^T - \beta\bar{z}\bar{z}^T\|_F &= \sqrt{\sum_{i,j}[\bar{v}_*(i)\bar{v}_*(j) - \beta\cdot\bar{z}(i)\bar{z}(j)]^2}\\ &\geq \sqrt{\sum_i[\bar{v}_*(i)^2 - \beta\cdot\bar{z}(i)^2]^2} \overset{(h)}{\geq} \sqrt{\sum_i[\bar{v}_*(i) - \beta\cdot\bar{z}(i)]^4}\\ &\overset{(i)}{\geq} \frac{1}{\sqrt{n}}\sum_i[\bar{v}_*(i) - \beta\cdot\bar{z}(i)]^2 = \frac{1}{\sqrt{n}}\|\bar{v}_* - \beta\cdot\bar{z}\|_2^2\end{aligned} \tag{40}$$

where (h) and (i) are due to Cauchy-Schwarz and Holder's inequality, respectively. Now combining this inequality with (39) leads to

$$\|\bar{v}_* - \beta\cdot\bar{z}\|_2^2 \leq \sqrt{n}\cdot\|\bar{v}_*\bar{v}_*^T - \beta\bar{z}\bar{z}^T\|_F \leq \sqrt{\frac{4g(\bar{z},\eta,\rho)\cdot n}{\lambda_2(H_\mu^{\widetilde{\Omega}})}}\|\bar{v}_*\|_2$$

which completes the proof.

B. Proof of Theorem 2

Define $N(k)$ to be the set of nodes adjacent to node k, including k itself. We will focus on a line $l\in\mathcal{E}$ that connects two nodes i and j.

(1) First, consider the case when $l\in\widetilde{\mathcal{S}}^c\cap\Xi$. The fact that $l\notin\widetilde{\mathcal{S}}$ implies that all nodes in the set $N(i)\cup N(j)$ are *solvable*. Also, since $l\in\Xi$, the nodal residual at nodes i and j are nonzero, which means that $i,j\in\mathcal{V}_\mathcal{N}$. Finally, noting that $l\notin\mathcal{R}^L$ because there is no line measurement for an erroneous line, we can conclude that $l\in\big(\mathcal{R}^N\setminus\mathcal{R}^L\big)$.

(2) Second, consider the case when $l\in\widetilde{\mathcal{S}}^c\cap\Xi^c$. Again, the fact that $l\notin\widetilde{\mathcal{S}}$ implies that all nodes in the set $N(i)\cup N(j)$ are

solvable. Also, since $l \in \Xi^c$, the nodal residuals at nodes i and j are zero, and the line residuals on line l is zero. Therefore, we can conclude that $l \notin \left(\mathcal{R}^N \cup \mathcal{R}^L\right)$.

(3) Third, consider the case when $l \in \widetilde{\mathcal{S}} \cap \Xi^c$. Since $l \in \widetilde{\mathcal{S}}$, at least one node in $N(i)$ and at least one node in $N(j)$ are *unsolvable*. From here, two different scenarios can happen. Scenario one is when at least one of nodes i and j is unsolvable. In this case, using the fact that there do not exist two distinct set of voltages that result in the same measurement values, we can easily conclude that $l \in \mathcal{R}^L \cap \mathcal{R}^N$. Scenario two is when both nodes i and j are solvable. In this scenario, the nodal residual at nodes i and j are nonzero but the line residual at l is zero. Therefore, $l \in \left(\mathcal{R}^N \setminus \mathcal{R}^L\right)$.

(4) Finally, consider the case when $l \in \widetilde{\mathcal{S}} \cap \Xi$. Since $l \in \widetilde{\mathcal{S}}$, at least one node in $N(i)$ and at least one node in $N(j)$ are *unsolvable*. Also, since $l \in \Xi$, the nodal residual at nodes i and j are nonzero, which means that $i, j \in \mathcal{V}_{\mathcal{N}}$. Finally, noting that $l \notin \mathcal{R}^L$ because there is no line measurement for an erroneous line, we can conclude that $l \in \left(\mathcal{R}^N \setminus \mathcal{R}^L\right)$.

From (1)–(4), we can deduce that $l \in \mathcal{R}^L \Longrightarrow l \in \left(\Xi^c \cap \widetilde{\mathcal{S}}\right)$, which proves the first part of the theorem. Furthermore, we can see that $l \in \widetilde{\mathcal{S}} \cup \Xi \Longrightarrow l \in \mathcal{R}^N$. Finally, from (2) specifically, we also know that if $l \in \widetilde{\mathcal{S}}^c \cap \Xi^c \Longrightarrow l \notin \mathcal{R}^N$. This concludes the fact that $\mathcal{R}^N = \widetilde{\mathcal{S}} \cup \Xi$.

C. Proof of Theorem 3

Proof. Consider equation (26) and set $\bar{M}_0 = 0$, $\rho = 1$. With some basic algebraic manipulations, one can write

$$\begin{aligned}&2\sum_{j\in N}|\bar{z}^T(\bar{M}_j(\widetilde{\Omega}) - \bar{M}_j(\Omega))\bar{z}| + 2\sum_{j=1}^{m}|\eta_j| \\ \geq &\|\mathcal{A}^{\Omega}(\bar{v}_*\bar{v}_*^T - \bar{z}\bar{z}^T)\|_1 \geq \|\mathcal{A}^{\Omega}(\bar{v}_*\bar{v}_*^T - \bar{z}\bar{z}^T)\|_2\end{aligned}$$

Therefore, if t is nonzero

$$\begin{aligned}&t \cdot \|\bar{v}_*\bar{v}_*^T - \bar{z}\bar{z}^T\|_F \\ \leq &2\sum_{j\in N}|\bar{z}^T(\bar{M}_j(\widetilde{\Omega}) - \bar{M}_j(\Omega))\bar{z}| + 2\sum_{j=1}^{m}|\eta_j| = 2g(\bar{z},\eta,1) \\ = &2\|\eta\|_1\end{aligned}$$

The last equality follows because all of the topological errors have been detected and fixed. This completes the proof. □

D. Theorem 4 and its proof

Theorem 4. *Denote $f^1(\cdot)$ as the objective function of an NLAV problem (i.e., equation (11)) with Ξ^1 as the set of erroneous lines and $\mathcal{M}$ as the index set of measurements. Similarly, denote $f^2(\cdot)$ as the objective function of another NLAV problem with Ξ^2 as the set of erroneous lines and $\mathcal{M}$ as the index set of measurements. Without loss of generality, suppose that $|\Xi^1| < |\Xi^2|$. Furthermore, assume that for any two vector of voltages, $\bar{x}$ and $\bar{y}$, and a measurement index j, the following holds:*

$$|\bar{x}^T\bar{M}_j(\widetilde{\Omega})\bar{x} - \bar{y}\bar{M}_j(\Omega)\bar{y}| > |\bar{x}^T\bar{M}_j(\Omega)\bar{x} - \bar{y}\bar{M}_j(\Omega)\bar{y}| \quad (41)$$

Then, $\min_{\bar{v}} f^1(\bar{v}) < \min_{\bar{v}} f^2(\bar{v})$

Proof. Let $\bar{v}_1$ and $\bar{v}_2$ be the global minimizer of $f^1(\cdot)$ and $f^2(\cdot)$, respectively. Also, let $\mathcal{M}^1$ and $\mathcal{M}^2$ be the set of measurement indices pertaining to the erroneous lines in Ξ^1 and Ξ^2, respectively. Then, the following inequalities hold:

$$\begin{aligned}f^2(\bar{v}_2) =&\bar{v}_2^T\bar{M}_0\bar{v}_2 + \rho\sum_{j=1}^{m}|\bar{v}_2^T\bar{M}_j(\widetilde{\Omega})\bar{v}_2 - \bar{z}^T\bar{M}_j(\Omega)\bar{z}| \\ \overset{(a)}{=}&\bar{v}_2^T\bar{M}_0\bar{v}_2 + \rho\sum_{j\in\mathcal{M}^2}|\bar{v}_2^T\bar{M}_j(\widetilde{\Omega})\bar{v}_2 - \bar{z}^T\bar{M}_j(\Omega)\bar{z}| \\ &+ \rho\sum_{j\in\mathcal{M}\setminus\mathcal{M}^2}|\bar{v}_2^T\bar{M}_j(\Omega)\bar{v}_2 - \bar{z}^T\bar{M}_j(\Omega)\bar{z}| \\ =&\bar{v}_2^T\bar{M}_0\bar{v}_2 + \rho\sum_{j\in\mathcal{M}^1}|\bar{v}_2^T\bar{M}_j(\widetilde{\Omega})\bar{v}_2 - \bar{z}^T\bar{M}_j(\Omega)\bar{z}| \\ &+ \rho\sum_{j\in\mathcal{M}\setminus\mathcal{M}^1}|\bar{v}_2^T\bar{M}_j(\Omega)\bar{v}_2 - \bar{z}^T\bar{M}_j(\Omega)\bar{z}| \\ &+ \rho\sum_{j\in\mathcal{M}^2\setminus\mathcal{M}^1}|\bar{v}_2^T\bar{M}_j(\widetilde{\Omega})\bar{v}_2 - \bar{z}^T\bar{M}_j(\Omega)\bar{z}| \\ &- \rho\sum_{j\in\mathcal{M}^2\setminus\mathcal{M}^1}|\bar{v}_2^T\bar{M}_j(\Omega)\bar{v}_2 - \bar{z}^T\bar{M}_j(\Omega)\bar{z}| \\ \overset{(b)}{\geq}&f^1(\bar{v}_1) + \rho\sum_{j\in\mathcal{M}^2\setminus\mathcal{M}^1}|\bar{v}_2^T\bar{M}_j(\widetilde{\Omega})\bar{v}_2 - \bar{z}^T\bar{M}_j(\Omega)\bar{z}| \\ &- \rho\sum_{j\in\mathcal{M}^2\setminus\mathcal{M}^1}|\bar{v}_2^T\bar{M}_j(\Omega)\bar{v}_2 - \bar{z}^T\bar{M}_j(\Omega)\bar{z}| \\ \overset{(c)}{>}&f^1(\bar{v}_1)\end{aligned}$$

where (a) follows from the fact that $\bar{M}_j(\widetilde{\Omega}) = \bar{M}_j(\Omega)$ if $j \notin \mathcal{M}^2$, (b) follows from the fact that $\bar{v}_1$ is the global minimum of $f^1(\cdot)$ and (c) follows from the assumption made in equation (41). □

REFERENCES

[1] S. Park, R. Mohammadi-Ghazi, and J. Lavaei, "Topology error detection and robust state estimation using nonlinear least absolute value," *2019 American Control Conference*, 2019.

[2] K. Clements and P. Davis, "Detection and identification of topology errors in electric power systems," *IEEE Transactions on Power Systems*, vol. 3, 1988.

[3] I. Dobson, B. A. Carreras, and D. E. Newman, "Probabilistic load-dependent cascading failure with limited component interactions," in *Proceedings of the International Symposium on Circuits and Systems*, 2004.

[4] A. MONTICELLI, "Electric power system state estimation," *Proceedings of the IEEE*, vol. 88, 2000.

[5] R. Zhang, J. Lavaei, and R. Baldick, "Spurious critical points in power system state estimation," in *Hawaii International Conference on System Sciences*, 2018.

[6] E. M. Lourenco, A. J. A. S. Costa, and K. A. Clements, "Bayesian-based hypothesis testing for topology error identification in generalized state estimation," *IEEE Transactions on power systems*, vol. 19, 2004.

[7] E. M. Lourenco, A. J. A. S. Costa, K. A. Clements, and R. A. Cernev, "A topology error identification method directly based on collinearity tests," *IEEE Transactions on power systems*, vol. 21, 2006.

[8] D. Singh, J. P. Pandey, and D. S. Chauhan, "Topology identification, bad data processing, and state estimation using fuzzy pattern matching," *IEEE Transactions on power systems*, vol. 20, 2005.

[9] K. A. Clements and A. S. Costa, "Topology error identification using normalized lagrange multipliers," *IEEE Transactions on power systems*, vol. 13, 1998.

[10] Y. Lin and A. Abur, "A computationally efficient method for identifying network parameter errors," in *Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2016.

[11] W.W.Kotiuga and M.Vidyasagar, "Bad data rejection properties of weighted least absolute value techniques applied to static state estimation," *IEEE Transactions on Power Apparatus and Systems*, vol. 101, pp. 844–853, 1982.

[12] P. Rousseeuw and A. Leroy, *Robust regression and outlier detection*. John Wiley and Sons, 1987.

[13] L. Mili, V. Phaniraj, and P. Rousseeuw, "Least median of squares estimation in power systems," *IEEE PES Summer Meeting*, pp. 493–497, 1990.

[14] M. Celik and A. Abur, "A robust wlav state estimator using transformations," *IEEE Transactions on Power Systems*, vol. 7, pp. 106–113, 1992.

[15] L. Mili, M. Cheniae, N. Vichare, and P. Rousseeuw, "Robust state estimation of power systems," *IEEE Transactions on Circuts and Systems*, vol. 41, pp. 349–358, 1994.

[16] M. Göl and A. Abur, "Lav based robust state estimation for systems measured by pmus," *IEEE Transactions on Smart Grid*, vol. 5, pp. 1808–1814, 2014.

[17] Y. Weng, M. D. Ilic, Q. Li, and R. Negi, "Convexification of bad data and topology error detection and identification problems in ac electric power systems," *IET Generation, Transmission & Distribution*, vol. 9, pp. 2760–2767, 2015.

[18] R. Zhang, J. Lavaei, and R. Baldick, "Spurious local minima in power system state estimation," *to appear in IEEE transactions on control of network systems (TCNS), 2019. Preprint available online at http://lavaei.ieor.berkeley.edu/SE_2018_1.pdf*, 2018.

[19] C. Josz, R. Zhang, Y. Ouyang, J. Lavaei, and S. Sojoudi, "A theory on the absence of spurious solutions for nonconvex and nonsmooth optimization," *Thirty-second annual conference on Neural Information Processing Systems (NIPS)*, 2018.

[20] S. Park, R. Mohammadi-Ghazi, and J. Lavaei, "Topology error detection and robust state estimation using nonlinear least absolute value," *archived and available at http://lavaei.ieor.berkeley.edu/SE_TC_1_2018.pdf*, 2018.

[21] F. C. Schweppe and J. Wildes, "Power system static-state estimationn, Part I: Exact model," *IEEE Transactions on Power Apparatus and Systems*, pp. 120–125, 1970.

[22] F. C. Schweppe, "Power system static-state estimationn, Part III: Implementation," *IEEE Transactions on Power Apparatus and Systems*, pp. 130–135, 1970.

[23] G.R.Krumpholz, K.A.Clements, and P.W.Davis, "Power system observability: A practical algorithm using network topology," *IEEE Transactions IEEE Transactions on Power Apparatus and Systems*, pp. 1534–1542, 1980.

[24] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak Math*, pp. 298–305, 1973.

[25] R. Diestel, *Graph Theory*. Springer, 2006.

[26] B.Recht, M. Fazel, and P. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimizations," *Siam review*, vol. 52, pp. 471–501, 2010.

[27] R. Ge, C. Jin, and Y. Zheng, "No spurious local minima in nonconvex low rank problems: A unified geometric analysis," *International Conference on Machine Learning*, pp. 1233–1242, 2017.

[28] I. Molybog, S. Sojoudi, and J. Lavaei, "No spurious solutions in nonconvex matrix sensing: structure compensates for isometry," *preprint available at https://lavaei.ieor.berkeley.edu/SRIP_2019_1.pdf*, 2019.

[29] Y. Zhang, R. Madani, and J. Lavaei, "Conic relaxations for power system state estimation with line measurements," *IEEE Transactions on Control of Network Systems*, vol. 5, pp. 1193–1205, september 2018.