Nonlinear Least Absolute Value Estimator for Power Systems State Estimation and Topological Error Detection

SangWoo Park, Reza Mohammadi-Ghazi, and Javad Lavaei

***Abstract*—This paper proposes 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

PROTECTING energy infrastructures against progressive failures of stressed components is an important challenge in operating these infrastructures and preventing blackouts [1], [2]. 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 [3], [4]. 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 measurement data and 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

A. Previous studies on topological error detection

The existing topological error detection methods take either a statistical or a geometric approach. Bayesian hypothesis testing [5], collinearity tests [6], and fuzzy pattern machine [7] 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 [8], although it is a heuristic method. Recent studies, such as the one proposed in [9], 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 [10] 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 [11], [12]. Further investigation and suggestion of different methods to mitigate this issue have been presented in [13], [14]. In [15], the authors show 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, nonconvex problem with power measurements, e.g., [16] where a semidefinite programming (SDP) relaxation is proposed to convexify the nonlinear LAV state estimator; however, no theoretical guarantees have been proposed 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 [17] and arising promises for the ℓ_1-norm [18], this study proposes a local search algorithm to find the global solution of the nonlinear LAV 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 a nonlinear LAV (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 propoer 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.

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$(\cdot)$, Tr$(\cdot)$ and null$(\cdot)$ denote the real part, imaginary part, rank, trace, and null space 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. $|\cdot|$ is the absolute value operator if the argument is a scalar, vector, or matrix; otherwise, it is the cardinality of a set if the argument is 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 $\mathbf{j} = \sqrt{-1}$. $\mathbb{1}$ 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. The operator vec$(\cdot)$ vectorizes its matrix argument.

II. Preliminaries

Consider an electric power network represented by a graph $\mathcal{G}(\mathcal{V},\mathcal{E})$, where $\mathcal{V} := \{1,\dots,n\}$ and $\mathcal{E} := \{1,\dots,r\}$ denote the sets of buses and branches, respectively. 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 injections (resp. currents) entering the line $l \in \mathcal{E}$ through the from and the 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 $\{e_1,\dots,e_n\}$ denote the canonical vectors in $\mathbb{R}^n$. Define

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

Moreover, let $\{d_1,\dots,d_r\}$ 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 make the following definitions:

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

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 $\bar{X}$ 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:*

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

Definition 3. *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*

$$\mathcal{A}^{\Omega}(X) = [\langle M_1(\Omega), X\rangle \cdots \langle M_m(\Omega), X\rangle]^T \tag{7}$$

Also, define the map $h^{\Omega}(\bar{v}) : \mathbb{R}^{2n-1} \to \mathbb{R}^m$ as the mapping from the complex-valued state of the power system to the vector of noiseless measurement values:

$$h^{\Omega}(\bar{v}) \triangleq [v^T M_1(\Omega) v, \dots v^T M_m(\Omega) v]^T \tag{8}$$

$$= [\bar{v}^T \bar{M}_1(\Omega)\bar{v}, \dots \bar{v}^T \bar{M}_m(\Omega)\bar{v}]^T \tag{9}$$

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,

$$J^{\Omega}(\bar{v}) = 2[\bar{M}_1(\Omega)\bar{v} \quad \bar{M}_2(\Omega)\bar{v} \quad \cdots \quad \bar{M}_m(\Omega)\bar{v}] \tag{10}$$

Definition 4. *Given a system model Ω and a set of measurements $\mathcal{M}$, define $\mathcal{J}_{\mathcal{M}} \in \mathbb{C}^n$ as the set of all voltage vectors v for which $J(\bar{v})$ has full row rank. A complex vector v is observable through the measurements $\mathcal{M}$ if it belongs to $\mathcal{J}_{\mathcal{M}}$.*

III. Main Results

In this section, we first state the most widely used state estimation formulation and discuss its properties. Then, we develop a novel algorithm that jointly performs state estimation and topology error detection based on ℓ_1-norm optimization, and analyze its features.

A. Nonlinear least-squares state estimation

Nonlinear least-squares (NLS) is the most common state estimation technique, which was first proposed by Schweppe [19], [20]. 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 measurements are noiseless. [4], [17]. Similar to other estimators, this method requires that the system's measurement matrices (see Definition 1) to 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. Thus, the algorithm should be revised in order to make a robust state estimation when topological errors are involved.

B. Proposed NLAV formulation

Let $\{\bar{M}_j(\widetilde{\Omega})\}_{j\in\mathcal{M}}$ be the set of measurement matrices corresponding to the model $\widetilde{\Omega}$ that the power system operator has access to, and $\{\bar{M}_j(\Omega)\}_{j\in\mathcal{M}}$ 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 would be significantly different from the true model. To design an algorithm that jointly performs both state estimation and sparse topological error detection, we propose the following optimization problem.

$$\min_{\bar{v}\in\mathbb{R}^{2n-1}} f(\bar{v}) = \bar{v}^T \bar{M}_0 \bar{v} + \rho \sum_{j=1}^{m} |\bar{v}^T \bar{M}_j(\widetilde{\Omega})\bar{v} - b_j| \tag{11}$$

where 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, z 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 z. We make the assumption that $z \in \mathcal{J}_{\mathcal{M}}$. 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. We impose the following conditions on M_0

Assumption 1. *The matrix M_0 satsifies the following properties:*

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

Let $\bar{v}_*$ denote a globally optimal solution of (11), i.e., $\bar{v}_* = \arg\min_{\bar{v}\in\mathbb{R}^{2n-1}} f(\bar{v})$. Finally, consider the vector of residuals $r = \{r_1, \dots, r_m\} \in \mathbb{R}^m$, where

$$r_j = |\bar{v}_*^T \bar{M}_j(\widetilde{\Omega})\bar{v}_* - b_j|, \quad \forall j \in \mathcal{M} \tag{13}$$

The problem (11) aims to push the insignificant residual errors to hard zero, while some of the r_j's associated with the outlier measurements are expected to remain non-zero. This phenomenon is supported empirically and is shown in Figure 2(e). The result has a striking constrast with the result of ℓ_2 minimization (Figure 2(d)) 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 (11) is nonlinear and nonconvex, which makes the local search algorithms prone to falling into spurious local solutions. However, recent studies have shown that increasing the number of redundant measurements helps with convexifying NLS problems and hence,

finding their global solutions using local search algorithms [4], [17]. Therefore, having access to high enough redundant measurements is the key for real-world state estimation problems. It is speculated that this property will carry on for the NLAV case and partial theoretical results have been developed in [18]. 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 one 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.

Our approach for solving the state estimation problem in the presence of topological error can be briefly summarized as follows. First, we solve (11) and use the pattern of the nonzero residuals errors to find a (small) subset of branches that are potentially erroneous in the model. Then, we use an efficient search method to find the topological errors followed by correcting the model accordingly, and re-estimating the underlying states of the system accurately. To formalize this approach, first we introduce the following definitions.

Definition 5. *A line is called erroneous if its presence in the system is misrepresented by the system operator in the model.*

Assumption 2. *There is no pair of erroneous lines that share a node.*

Definition 6. *Define $\mathcal{N} \in \mathcal{M}$ as the set of indices of the measurements that correspond to the erroneous lines.*

Definition 7. *Given a regularization matrix $M_0 \in \mathbb{S}^n \succeq 0$, 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:*

$$H_\mu^\Omega \succeq 0, \quad H_\mu^\Omega v = 0, \quad rank\{H_\mu^\Omega\} = n - 1 \tag{14}$$

Definition 8. *A node $i \in \mathcal{V}$ is called* unsolvable *if the i^{th} entry of $z - v_*$ is nonzero. On the other hand, if the i^{th} entry of $z - v_*$ is zero, node i is called* solvable*. Also, a node is called* isolated *if its degree is zero.*

Definition 9. *Define the following four subgraphs of $\mathcal{G}$:*

1) *The* state estimation error graph *$\mathcal{S}(\mathcal{V}_\mathcal{S}, \mathcal{E}_\mathcal{S})$ is an induced subgraph of $\mathcal{G}$ such that $\mathcal{V}_\mathcal{S}$ is the set of unsolvable nodes of $\mathcal{G}$.*
2) *The* extended state estimation error graph *$\widetilde{\mathcal{S}}(\mathcal{V}_{\widetilde{\mathcal{S}}}, \mathcal{E}_{\widetilde{\mathcal{S}}})$ is an induced subgraph of $\mathcal{G}$ 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 $\mathcal{V}_\mathcal{S}$. (therfore, $\mathcal{V}_\mathcal{S} \subseteq \mathcal{V}_{\widetilde{\mathcal{S}}}$)*
3) *The* residual graph *$\mathcal{R}(\mathcal{V}_\mathcal{R}, \mathcal{E}_\mathcal{R})$ is a subgraph of $\mathcal{G}$ where $\mathcal{E}_\mathcal{R}$ is the set of lines whose associated entries in the vector of residual errors, denoted as r, are nonzero. $\mathcal{V}_\mathcal{R}$ is the set of nodes that are either: 1) the* from *or* to *nodes of the lines in $\mathcal{E}_\mathcal{R}$, or 2) nodes whose associated entries in r are nonzero.*
4) *The* extended residual graph *$\widetilde{\mathcal{R}}(\mathcal{V}_{\widetilde{\mathcal{R}}}, \mathcal{E}_{\widetilde{\mathcal{R}}})$ is an induced subgraph of $\mathcal{G}$ such that $\mathcal{V}_{\widetilde{\mathcal{R}}}$ includes all nodes in $\mathcal{V}_\mathcal{R}$ and also those nodes that are adjacent to any node in $\mathcal{V}_\mathcal{R}$. (therfore, $\mathcal{V}_\mathcal{R} \subseteq \mathcal{V}_{\widetilde{\mathcal{R}}}$)*

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 [21]. For instance, for the state estimation error graph, $\mathcal{E}_\mathcal{S}$ is the set of all branches whose *from* and *to* nodes are both in $\mathcal{V}_\mathcal{S}$. Definition 9 is illustrated for a small system in Figure 1. Each unsolvable node is marked by a cross. The definitions introduced so far in this section will be utilized in the next section to prove various results.

C. Estimation error

In this section, first we provide a theoretical upper bound on the state estimation error obtained by the NLAV problem (11) and uncover certain properties of the vector of residual errors. Then, we use these results to develop an algorithm for state estimation and topological error detection.

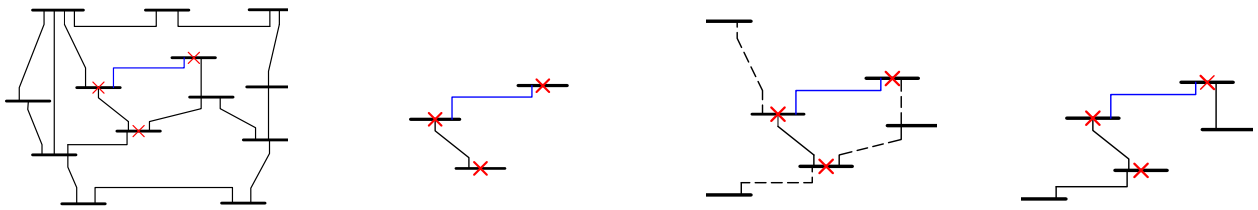
Theorem 1. *Suppose that the power system operator has a network model denoted by $\widetilde{\Omega}$ and a set of measurement indices $\mathcal{M}$. For this model, assume that there exists a dual certificate μ for the true state vector z. Also, consider a parameter ρ satisfying the inequality $\rho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

$$\frac{\|\bar{v}_* - \beta \cdot \bar{z}\|_2^2}{\|\bar{v}_*\|_2} \leq \sqrt{\frac{4n \cdot g(\bar{z}, \eta, \rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}})}} \tag{15}$$

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

$$\rho \left[\sum_{j \in \mathcal{N}} |\bar{z}^T (\bar{M}_j(\Omega) - \bar{M}_j(\widetilde{\Omega}))\bar{z}| + \sum_{j=1}^m |\eta_j| \right] \tag{16}$$

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 upperbound 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 [22], is a parameter that measures how well a weighted graph is connected. For example, a complete graph has algebraic connectivity of n while the value is 2 for a star graph and $2(1-cos\frac{\pi}{n})$ for a path (where n is 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., $M_0(i,j) = 0$ if line $(i,j) \notin \mathcal{E}$), 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 15 does not guarantee a unique solution of the NLAV. For conditions that



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

Figure 1: (a) A power network, (b) the state estimation error graph $\mathcal{S}$, (c) the extended state estmation error graph $\widetilde{\mathcal{S}}$, and (d) the residual graph $\mathcal{R}$. 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 $\widetilde{\mathcal{S}}$.

guarantee the uniqueness of solution, the reader is referred to Theorem 3.

Theorem 2. *Suppose that the noise vector η is zero. Under the setting of Theorem 1 and Assumption 2, the following statements hold:*

1) *The extended state estimation error graph $\widetilde{\mathcal{S}}$ does not have an isolated node and it contains all of the erroneous lines.*
2) *The residual graph $\mathcal{R}$ is a subset of $\widetilde{\mathcal{S}}$.*

For the special case where $|\mathcal{N}| = 1$, Theorem 2 implies that $\widetilde{\mathcal{S}}$ is connected. When $|\mathcal{N}| > 1$, each connected component of $\widetilde{\mathcal{S}}$ contains at least one erroneous line. The practical benefit of Theorem 2 is that it enables us to develop a technique for efficiently detecting topological errors by searching over the sparse graph $\widetilde{\mathcal{S}}$. Note that although $\widetilde{\mathcal{S}}$ is the ideal subgraph of $\mathcal{G}$ that always contains the erroneous line, it is impossible to form and search over this subgraph due to the unavailability of the true underlying states z of the system. To address this issue, we propose searching over $\widetilde{\mathcal{R}}$, which is a superset of $\mathcal{R}$ and is closer to $\widetilde{\mathcal{S}}$.

D. 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 $\widetilde{\mathcal{R}}$ 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-solve 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 the $\widetilde{\mathcal{R}}$ have already been evaluated. Algorithm 1 summarizes the proposed topological error detection method.

E. Unpenalized NLAV estimator

If the regularization is disregarded in (11), i.e., setting $M_0 = 0$, Algorithm 1 still works; however, the state estimation error bound provided in Theorem 1 is no longer valid. To obtain a new bound, we need the following theorem.

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 (11) with $\widetilde{\Omega}$, $\mathcal{C}_{\widetilde{\Omega}}(\mathcal{M}), b$ and calculate the residual r based on equation (13).
2. Construct the extended residual graph $\widetilde{\mathcal{R}}(\mathcal{V}_{\widetilde{\mathcal{R}}}, \mathcal{E}_{\widetilde{\mathcal{R}}})$.
3. Set $r^t \leftarrow r$.
while $\|r^t\|_2 > \delta$ **do**
  $\Omega^t \leftarrow \Omega$.
  for $l \in \mathcal{E}_{\widetilde{\mathcal{R}}}$ **do**
    Update Ω^t to $\Omega^{t'}$ by changing the on/off status of line l.
    Re-solve (11) 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 $\mathcal{D}_L$

Theorem 3. *Given the model $\widetilde{\Omega}$ and the measurement set $\mathcal{M} = \{1, \ldots, m\}$, let $\mathcal{A}^{\widetilde{\Omega}}$ be the mapping as defined in Definition 3. Then, the state estimation error is bounded above by the following inequality:*

$$\|\bar{v}_*\bar{v}_*^T - \bar{z}\bar{z}^T\|_F \leq \frac{2}{t} g(\bar{z}, \eta, 1) \tag{17}$$

where t is defined as

$$t = \min_{K \in \mathbb{S}^n} \|\mathcal{A}^{\widetilde{\Omega}}(K)\|_2 \quad s.t. \quad rank(K) = 2, \ \|K\|_F = 1 \tag{18}$$

It is straightforward to verify that $t > 0$ if and only if there does not exist any set of noiseless measurements for the model $\widetilde{\Omega}$ 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 for the over-determined power flow equations.

Recently, there has been some interesting research on the connection between the property of *no spurious local minima*

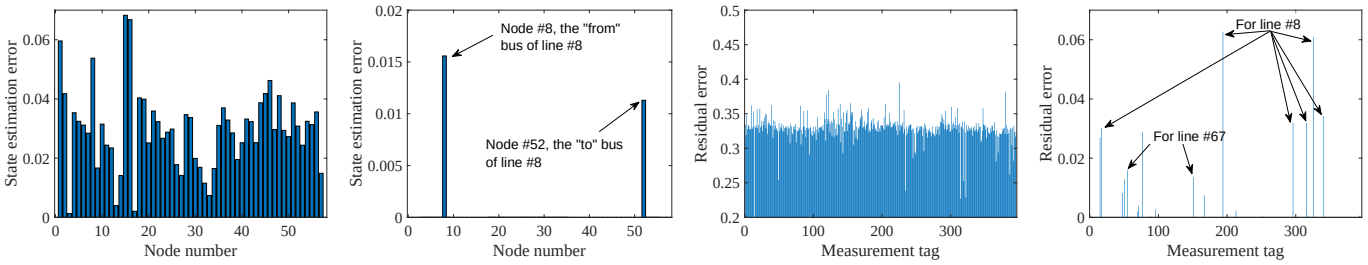


Figure 2: State estimation error for (a) NLS with noiseless measurements, (b) NLAV method with noiseless measurement; and residuals for (c) NLS with noiseless measurements, (d) NLAV method with noiseless measurement. Note that in (c) and (d), the x-axis shows the measurement tag, which is not the same as the node or line number due to random selection of line measurements.

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 $0 \leq \delta_r < 1$ if there exists $p > 0$ such that for all rank-r matrices X: $(1-\delta_r)\|X\|_F^2 \leq \frac{1}{p}\|\mathcal{A}(X)\|^2 \leq (1+\delta_r)\|X\|_F^2$. If $\mathcal{A}$ satisfies $(2r,\ \delta_{2r})$-RIP with $\delta_{2r} < 1$, then finding a global optimum constitutes exact recovery of the state [23]. However, this does not exclude the existence of spurious local minima, which can be problematic when using local search algorithms. In order to quarantee no spurious local minima, $\mathcal{A}$ must satisfy $(2r,\ \delta_{2r})$-RIP with $\delta_{2r} < 0.2$ which is a stricter condition [24]. 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

This section presents numerical simulations on the IEEE 57-bus and 118-bus systems for evaluating the efficacy of the proposed algorithm in estimating the underlying states of power systems in presence of topological errors and detecting such errors. To run the simulations, we use MATPOWER data along with the MATLAB *fmincon* function as the local search algorithm.

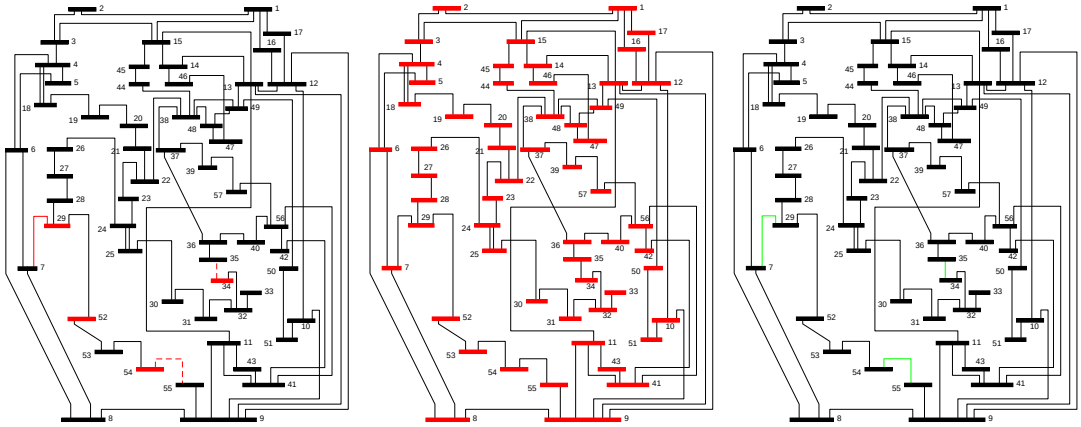
A. Simulation setup

In this study we focus on two types of topological errors as follows: **type I**: a transmission line is switched off in the true system while it is switched on in the hypothetical model that is accessible to the power system operator; **type II**: a branch is switched on in the true model while it is switched off in the hypothetical model. As stated in Definition 5, a branch that falls under either of these error types is called an *erroneous* line. Our numerical evaluations consist of multiple cases such that each is specified by a certain number of erroneous lines and the percentage of line measurements that are taken into account in state estimation, e.g., three erroneous lines and 20% of all possible line measurements. The procedure of running the simulations is as follows: (1) For a given number of erroneous lines and line measurement percentage, we run 30 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 use the true/false positive rates and the number of lines that the algorithm checks before termination as the evaluation criteria. Note that there is no unique way of defining true/false positive rates when it comes to monitoring infrastructures. In this study, we define true positive rate (%) as the number of erroneous lines whose topological error types are correctly detected divided by the total number of erroneous lines, times one-hundred. The false positive rate (%) is defined as the number of erroneous lines that are not detected correctly divided by the total number of lines, times one-hundred.

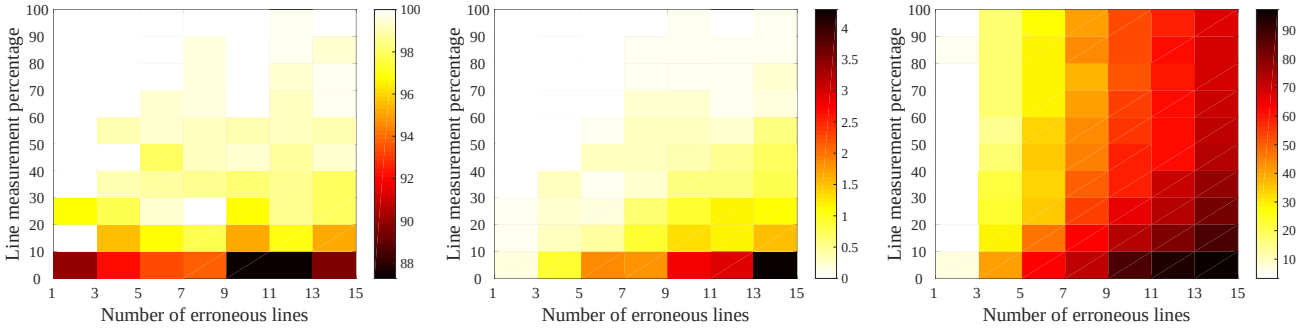
B. Example: Sparse residuals for NLAV

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 a scenario with two erroneous lines (line 8 and 67) and 70% line measurements and the results of solving the NLS and NLAV state estimations are depicted in Figure 2. Figures 2(a) and 2(c) show the state estimation errors and residuals of NLS when topological errors exist. It follows from these plots that the state estimation errors are considerably high, which means that the NLS lacks robustness in presence of topological errors. Also, there is no sparsity pattern in these plots, and the high peaks are not 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 2(b) and 2(d). The two only peaks in Figure 2(b) are associated with the state estimation errors at the ending nodes of the erroneous lines, and such error is zero at all other buses. Another important observation is the sparse pattern of the residual vector, shown in Figure 2(d). The



(a) (b) (c)

Figure 3: Case study: comparing the state estimation results of the NLS and the proposed NLAV estimator. (a) The true system where the hatched red lines are the erroneous lines with type I error; the solid red line is an erroneous line with type II error; the red buses are those where the first round state estimation result of the proposed NLAV is nonzero. (b) The result of NLS; the red buses are those where the state estimation error is not exact. (c) The result of the proposed NLAV; the green lines are the erroneous lines that are correctly detected.



(a) (b) (c)

Figure 4: Simulation results on the IEEE 57-bus system. (a) Average true positive rates, (b) average false positive rates, (c) average number of lines that are checked before the termination of the algorithm. All true and false positive rates are in percentage and between 0 and 100.

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. Comparing this result with that of NLS, one can observe the superiority of the NLAV estimator in finding the true states in the presence of topology error. In the following two subsections, we present extensive simulations conducted on the IEEE 57-bus and 118-bus systems.

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. One of these scenarios with three erroneous lines and 30% line measurements is depicted in Fig. 3. Fig. 3(a) shows the true system where the hatched red lines are those with type I topological error and the solid red line is the one with type II error. The red buses are the locations where the first round of state estimation by the proposed NLAV estimator along with the hypothetical model is inexact. Comparing this plot with the state estimation result of the NLS, Fig. 3(b), where none of the states were correctly estimated, shows the superiority of the proposed method over the traditional NLS. Following the iterative scheme presented in section III-D and by searching over only seven lines, we could accurately estimate the underlying states of the system as well as correctly detect the erroneous lines as shown in Fig. 3(c). In comparison, to find the erroneous lines via NLS we would need to search over all possible combinations of lines, which would result in an exponential time complexity.

Table I: Simulations results on the IEEE 118-bus system

	Average true positive (%)			**Average false positive (%)**			**Average number of lines checked**		
	Number of erroneous lines			Number of erroneous lines			Number of erroneous lines		
Line measurements	5	15	25	5	15	25	5	15	25
10%	98	96.44	94.40	0.466	1.039	2.258	61.46	154.13	173.7
40%	99.33	99.33	98.67	0.072	0.376	0.591	36.86	103.06	153.56
70%	100	99.11	99.07	0.072	0.107	0.179	40.6	93.56	138.43
100%	100	99.33	99.47	0.018	0.054	0.072	33	93.33	131.07

Fig. 4 shows the average value of the evaluation criteria of the 30 simulations for the above-mentioned 88 scenarios. The results imply that the proposed algorithm is capable of estimating the underlying states and finding the topological errors accurately if the number of redundant measurements is large enough. We can also see that detecting topological errors becomes more difficult as the number of such errors grows. Note that in some cases the algorithm may check more than 90% of the branches to find the topological errors if number of erroneous lines is large; however, this is still much smaller than the all possible combinations that a brute force algorithm would search over.

D. 118-bus system

To better assess the performance of the proposed NLAV estimator in a more realistic problem, we apply the proposed technique to analyze the IEEE 118-bus system. We pursue the procedure described in Section IV-A for numerical simulations, but consider $\{5, 15, 25\}$ and $\{10\%, 40\%, 70\%, 100\%\}$, respectively, as the candidate number of erroneous lines and line measurement percentages. TABLE I illustrates the state estimation and topological error detection results of these analyses, which are well matched with the ones for the IEEE 57-bus system.

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) 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. In particular, the proposed method can successfully detect up to 25 erroneous lines in the modeling of the power system for the IEEE 118-bus system.

Appendix

A. Proof of Theorem 1

Consider the NLAV problem (11). 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

$$\begin{aligned}
&\bar{v}_*^T \bar{M}_0 \bar{v}_* - \bar{z}^T \bar{M}_0 \bar{z} + \rho \sum_{j=1}^{m} |\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\Omega) \bar{z}| \\
&\leq \rho \sum_{j \in \mathcal{N}} |\bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z} - \bar{z}^T \bar{M}_j(\Omega) \bar{z}| + 2\rho \sum_{j=1}^{m} |\eta_j| \qquad (19)
\end{aligned}$$

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:

$$\begin{aligned}
&\bar{v}_*^T \bar{M}_0 \bar{v}_* - \bar{z}^T \bar{M}_0 \bar{z} + \rho \sum_{j=1}^{m} |\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z}| \\
&\leq 2\rho \left\{ \sum_{j \in \mathcal{N}} |\bar{z}^T (\bar{M}_j(\widetilde{\Omega}) - \bar{M}_j(\Omega)) \bar{z}| + \sum_{j=1}^{m} |\eta_j| \right\} \\
&= 2g(\bar{z}, \eta, \rho) \qquad (20)
\end{aligned}$$

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

$$\begin{aligned} \min_{t} \quad & \bar{v}_*^T \bar{M}_0 \bar{v}_* - \bar{z}^T \bar{M}_0 \bar{z} + \rho \sum_{j=1}^{m} t_j \\ s.t. \quad & \bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z} \le t_j, \quad \forall j \in \mathcal{M} \\ & -\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* + \bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z} \le t_j, \quad \forall j \in \mathcal{M} \end{aligned} \tag{21}$$

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

$$\begin{aligned} \mathcal{L}(t, p^+, p^-) &= \bar{v}_*^T \bar{M}_0 \bar{v}_* - \bar{z}^T \bar{M}_0 \bar{z} + \sum_{j=1}^{m} (\rho - p_j^+ - p_j^-) t_j \\ &+ \sum_{j=1}^{m} \{(p_j^+ - p_j^-)(\bar{v}_*^T \bar{M}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{M}_j(\widetilde{\Omega}) \bar{z})\} \end{aligned} \tag{22}$$

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

$$\begin{aligned} d(p^+, p^-) = \bar{v}_*^T \Big(\bar{M}_0 + \sum_{j=1}^{m} (p_j^+ - p_j^-) \bar{M}_j(\widetilde{\Omega}) \Big) \bar{v}_* \\ - \bar{z}^T \Big(\bar{M}_0 + \sum_{j=1}^{m} (p_j^+ - p_j^-) \bar{M}_j(\widetilde{\Omega}) \Big) \bar{z} \end{aligned} \tag{23}$$

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:

$$d(p_*^+, p_*^-) = \bar{v}_*^T H_\mu^{\widetilde{\Omega}} \bar{v}_* - \bar{z}^T H_\mu^{\widetilde{\Omega}} \bar{z} \tag{24}$$

$$= \text{Tr}\{H_\mu^{\widetilde{\Omega}} \bar{v}_* \bar{v}_*^T\} = \text{Tr}\{H_\mu^{\widetilde{\Omega}} X\} \tag{25}$$

The rest of the proof can be adopted from [25] (Appendix, Proof of Theorom 2). Consider an eigen-decomposition of $H_\mu^{\widetilde{\Omega}} = U \Lambda U^T$, where $\Lambda = \text{diag}(\lambda_{2n-1}, ..., \lambda_1)$ such that $\lambda_{2n-1} \ge \cdots \ge \lambda_1$ and U is a unitary matrix whose columns are the corresponding eigenvectors. Define

$$\breve{X} := \begin{bmatrix} \widetilde{X} & \widetilde{x} \\ \widetilde{x}^T & \alpha \end{bmatrix} = U^T X U \tag{26}$$

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

$$\begin{aligned} Tr\big(H_\mu^{\widetilde{\Omega}} X\big) &= Tr\big(U \Lambda U^T U \breve{X} U^T\big) = Tr\big(\Lambda \breve{X}\big) \\ &\ge \lambda_2(H_\mu^{\widetilde{\Omega}}) Tr\big(\widetilde{X}\big) \end{aligned} \tag{27}$$

Combining (27) and (20) leads to

$$\text{Tr}\big(\widetilde{X}\big) \le 2 \cdot g(\bar{z}, \eta, \rho) / \lambda_2(H_\mu^{\widetilde{\Omega}}) \tag{28}$$

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

$$\begin{aligned} X = U \breve{X} U^T &= \begin{bmatrix} \widetilde{U} & \widetilde{z} \end{bmatrix} \begin{bmatrix} \widetilde{X} & \widetilde{x} \\ \widetilde{x}^T & \alpha \end{bmatrix} \begin{bmatrix} \widetilde{U}^T \\ \widetilde{z}^T \end{bmatrix} \\ &= \widetilde{U} \widetilde{X} \widetilde{U}^T + \widetilde{U} \widetilde{x} \widetilde{z}^T + \widetilde{z} \widetilde{x}^T \widetilde{U}^T + \alpha \widetilde{z} \widetilde{z}^T \end{aligned} \tag{29}$$

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 $\alpha = \text{Tr}(X) - \text{Tr}(\widetilde{X})$, one can write

$$\|\widetilde{x}\|_2^2 \le \alpha \text{Tr}(\widetilde{X}) = \text{Tr}(X)\text{Tr}(\widetilde{X}) - \text{Tr}^2(\widetilde{X}) \tag{30}$$

Therefore,

$$\begin{aligned} \|X - \alpha \widetilde{z} \widetilde{z}^T\|_F^2 &= \|\widetilde{U} \widetilde{X} \widetilde{U}^T + \widetilde{U} \widetilde{x} \widetilde{z}^T + \widetilde{z} \widetilde{x}^T \widetilde{U}^T\|_F^2 \\ &\overset{(d)}{=} \|\widetilde{U} \widetilde{X} \widetilde{U}^T\|_F^2 + 2\|\widetilde{z} \widetilde{x}^T \widetilde{U}^T\|_F^2 \overset{(e)}{=} \|\widetilde{X}\|_F^2 + 2\|\widetilde{x}\|_2^2 \\ &\le \|\widetilde{X}\|_F^2 - 2\text{Tr}^2(\widetilde{X}) + 2\text{Tr}(X)\text{Tr}(\widetilde{X}) \\ &\overset{(f)}{\le} 2\text{Tr}(X)\text{Tr}(\widetilde{X}) \\ &\overset{(g)}{\le} \frac{4g(\bar{z}, \eta, \rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}})} \text{Tr}(X) \end{aligned} \tag{31}$$

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 $\|\widetilde{X}\|_F \le \text{Tr}(\widetilde{X})$. Finally, (g) results from substituting equation (28). Plugging back in $X = \bar{v}_* \bar{v}_*^T$ yields that

$$\begin{aligned} \|X - \alpha \widetilde{z} \widetilde{z}^T\|_F^2 &= \|\bar{v}_* \bar{v}_*^T - \frac{\alpha}{\|\bar{z}\|_2^2} \bar{z} \bar{z}^T\|_F^2 \\ &\le \frac{4g(\bar{z}, \eta, \rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}}(\bar{M}_0))} \text{Tr}(\bar{v}_* \bar{v}_*^T) \end{aligned} \tag{32}$$

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

$$\|\bar{v}_* \bar{v}_*^T - \beta \bar{z} \bar{z}^T\|_F^2 \le \frac{4g(\bar{z}, \eta, \rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}}(\bar{M}_0))} \|\bar{v}_*\|_2^2 \tag{33}$$

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

$$\begin{aligned} \|\bar{v}_* \bar{v}_*^T - \beta \bar{z} \bar{z}^T\|_F &= \sqrt{\sum_{i,j} [\bar{v}_*(i) \bar{v}_*(j) - \beta \cdot \bar{z}(i) \bar{z}(j)]^2} \\ &\ge \sqrt{\sum_i [\bar{v}_*(i)^2 - \beta \cdot \bar{z}(i)^2]^2} \overset{(h)}{\ge} \sqrt{\sum_i [\bar{v}_*(i) - \beta \cdot \bar{z}(i)]^4} \\ &\overset{(i)}{\ge} \frac{1}{\sqrt{n}} \sum_i [\bar{v}_*(i) - \beta \cdot \bar{z}(i)]^2 = \frac{1}{\sqrt{n}} \|\bar{v}_* - \beta \cdot \bar{z}\|_2^2 \end{aligned} \tag{34}$$

where (h) and (i) are due to Cauchy-Schwarz and Holder's inequality, respectively. Now combining this inequality with (33) leads to

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

which completes the proof.

B. Proof of Theorem 2

1) First, we prove that $\widetilde{\mathcal{S}}$ does not have any isolated nodes. By contradiction, assume that $\widetilde{\mathcal{S}}$ has an isolated node j. Then, j is unsolvable (i.e., $\bar{v}_*(j) - \bar{z}(j) \neq 0$) while its neighboring nodes are all solvable. However, this cannot be true because changing $\bar{v}_*(j)$ to $\bar{z}(j)$ will lower the residual values of this isolated neighbor while not interfering with the rest of $\widetilde{\mathcal{S}}$. In other words, this contradicts the fact that $\bar{v}_*$ is an optimal solution to the NLAV.

Next, we prove that $\widetilde{\mathcal{S}}$ always contains all erroneous lines. To prove by contradiction, assume that an erroneous line (i,j) is not contained in $\widetilde{\mathcal{S}}$, i.e. $v_*(k) = z(k) \quad \forall k \in N(i) \cup N(j)$. Here, $N(i)$ denotes the set of nodes adjacent to node i. Then, without loss of generality, the power injection at node i can be written as $s_i = \sum_{k\in N(i)} z(i)[z(i) - z(k)]^* y_{i,k}^* = \sum_{k\in N(i)} v_*(i)[v_*(i) - v_*(k)]^* \tilde{y}_{i,k}^*$, where $y_{i,k}$ is the actual admittance of the line (i,k) and $\tilde{y}_{i,k}$ is the modeled line admittance of the line (i,k). Because $j \in N(i)$ and $y_{i,j} \neq \tilde{y}_{i,j}$, the above equation cannot be satisfied. Therefore, at least one of the nodes from $N(i)$ should be an unsolvable node. In a similar fashion, we can also show that at least one of the nodes from $N(j)$ should be an unsolvable node. As a result, $i, j \in \mathcal{V}_{\widetilde{\mathcal{S}}}$ and the erroneous line should be contained in $\widetilde{\mathcal{S}}$.

2) Consider a line $(i,j) \notin \widetilde{\mathcal{S}}$. Due to Part 1 of Theorem 2, (i,j) is not erroneous. Combining this with the fact that both node i and node j are solvable, we arrive at the result that the residual corresponding to the nodal measurements at nodes i and j as well as the line measurments at line (i,j) all have zero values. Therefore, we have line $(i,j) \notin \mathcal{R}$ and consequently $\mathcal{R} \subseteq \widetilde{\mathcal{S}}$.

C. Proof of Theorem 3

Proof. Consider equation (20) and set $\bar{M}_0 = 0$, $\rho = 1$. With some basic algebraic manipulations, one can write

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

Therefore, if t is nonzero

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

This completes the proof. □

REFERENCES

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

[2] Y. Lin and A. Abur, "Probabilistic load-dependent cascading failure with limited component interactions," in *Proceedings of the International Symposium on Circuits and Systems*, 2004.

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

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

[5] 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.

[6] 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.

[7] 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.

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

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

[10] 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.

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

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

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

[14] 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.

[15] 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.

[16] 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.

[17] R. Zhang, J. Lavaei, and R. Baldick, "Spurious local minima in power system state estimation," *available online at http://lavaei.ieor.berkeley.edu/SE_2018_1.pdf*, 2018.

[18] C. Josz, R. Zhang, Y. Ouyang, J. Lavaei, and S. Sojoudi, "A theory on the absence of spurious optimality," *to appear in Neural Information Processing Systems (NIPS), available online at https://arxiv.org/pdf/1805.08204.pdf*, 2018.

[19] 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.

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

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

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

[23] 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.

[24] 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.

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