Digital Object Identifier 10.1109/ACCESS.2021.DOI

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

SANGWOO PARK[1], REZA MOHAMMADI-GHAZI[1], AND JAVAD LAVAEI.[1]
[1]Department of Industrial Engineering and Operation Research, University of California Berkeley, CA 94710 USA
Emails: {spark111, mohammadi, lavaei}@berkeley.edu

Corresponding author: SangWoo Park (e-mail: spark111@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.

ABSTRACT Topology error, a modeling misrepresentation of the power system network configuration, can undermine the quality of state estimation. In this paper, we propose a new methodology for robust power system state estimation (PSSE) modeled by AC power flow equations when there exists a small number of topological errors. The developed technique utilizes the availability of a large number of SCADA measurements and minimizes the ℓ_1 norm of nonconvex residuals augmented by a nonlinear, but convex, regularizer. Representing the power network by a graph, we first study the properties of the solution obtained from the proposed NLAV estimator and demonstrate 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 introduce a method that can efficiently detect the topological errors by searching over the identified subgraph. In addition, 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 power system infrastructures against cascading failures is a crucial challenge when operating these systems and if not managed properly, could lead to blackouts [1], [2]. In order to do so, the power network must be constantly overseen so that, if needed, appropriate actions can be taken. This monitoring is achieved via real-time state estimation that aims to recover the underlying system voltage phasors, given supervisory control and data acquisition (SCADA) measurements and a system model that encodes the network 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 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 quality of state estimation, coping with bad data and detecting topological errors have received considerable attention in the past few decades. In addition, recent research shows the impact of topology errors on real-time market operations such as locational marginal pricing [5].

A. LITERATURE SURVEY ON TOPOLOGICAL ERROR DETECTION

Bayesian hypothesis testing [6], collinearity testing [7], and fuzzy pattern machine [8] are examples of statistical approaches for detecting topological errors. These methods often require prior information on the states and/or a significant amount of historical data from past measurements. Other approaches venture to devise state estimators that are robust against topological errors and measurement noise. The work [9] used normalized Lagrange multipliers of the least-squares state estimation problem, and despite being a

heuristic method, has been shown to be effective in some cases. Later studies, such as [10], improved on this approach. Another noteworthy method in this category is the least absolute value (LAV) estimator, which was used in the context of power systems in [11]. By minimizing the ℓ_1 norm of the residual vector obtained from the linearized measurement equations, the LAV is able to find a minimum set of measurements untainted by gross errors, thus dismissing bad data and generating a robust state estimate. In spite of its strength, the LAV is susceptible to leverage points as discussed in [12], [13]. Therefore, further investigation and suggestion of various methods to resolve this issue have been made in [14], [15]. The work of [16] showed that the adverse effect of leverage points can be mitigated if measurements consist only of phasor measurement units (PMUs). The similarity of the methods mentioned so far is their reliance on linearized measurement equations. There is a limited number of research that have studied the fully nonlinear, non-convex problem with power measurements, for instance, [17] where a semidefinite programming (SDP) relaxation was proposed to convexify the nonlinear LAV state estimator; however, no theoretical guarantees have been developed to guarantee the recovery of a high-quality solution. Furthermore, the computational burden of solving the surrogate SDP problem may limit the use of this approach to relatively small-sized problems in practice. In [18], without considering topological errors in the model, the authors studied conditions under which a linearized iterative algorithm can recover the true state from nonlinear measurements that contain sparse bad data. The conditions are difficult to check and have not been verified for real-world power networks. The recent work [19] developed a modified least absolute value state estimator that is experimentally robust against both bad data and topological errors but may require theoretical guarantees and extended simulation results. These issues prompt further research on developing robust state estimation techniques with the capability of managing nonconvexities associated with various types of measurements.

B. CONTRIBUTIONS

Taking into account the theoretical guarantees recently developed for the ℓ_2-norm to avoid spurious local minimizers in nonconvex optimization [20] and arising promises for the ℓ_1-norm [21], our paper introduces a local search algorithm that can find the global solution of the nonlinear LAV (NLAV) state estimator with high probability. The proposed technique presents a robust approach for estimating the power system's voltages in the presence of a modest number of topological errors as well as detecting such errors. We summarize the main contributions of this work as follows: (1) introducing an algorithm for identifying topology errors and estimating the power system states using an NLAV state estimator combined 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 discussed in the later part of this paper, fast local search algorithms would efficiently find global solutions of the underlying NLAV estimators given a sufficient number of noiseless measurements and an adequate initialization of the algorithm. This manuscript is an extended version of the conference paper [22] with new additions including an updated Theorem 2, updated topology error detection algorithm (Algorithm 1), comprehensive case studies and a complete appendix with proofs. In Theorem 2 of [22], we derived a result stating that the *line residual graph* is a subset of the *extended state estimation error graph*, which contains the topological errors. The *extended state estimation error graph* is unattainable in practice and we resorted to searching over the *extended line residual graph*, which lacked strong theoretical support. To resolve this issue, in the updated Theorem 2 of this manuscript, we show that the topological errors are contained within a small subgraph of the power network, called the *suspect-subgraph*. This provides a solid theoretical foundation for our algorithm and improves on the efficiency of the detection algorithm since the search now iterates over a smaller subgraph than before. This change is also reflected in the algorithm, and therefore we present an updated version of Algorithm 1 in this manuscript.

The remainder of this paper is organized as follows. Preliminary materials such as notations and definitions are presented in Section II, followed by the formulation of the algorithm and the main theoretical results in Section III. A comprehensive set of numerical simulations on the IEEE 57-bus system and the 118-bus system is presented in Section IV. Finally, the summary and concluding remarks are drawn in Section V. The proofs are provided in the Appendix.

II. PRELIMINARIES

A. NOTATIONS

In this paper, lower case letters stand for column vectors, upper case letters stand for matrices and calligraphic letters represent sets and graphs. The sets of real and complex numbers are represented by symbols $\mathbb{R}$ and $\mathbb{C}$, respectively. Subsequently, $\mathbb{R}^N$ and $\mathbb{C}^N$ stand for the spaces of N-dimensional real and complex vectors, respectively. Next, $\mathbb{S}^N$ and $\mathbb{H}^N$ denote the sets of $N \times N$ complex symmetric matrices and Hermitian matrices, respectively. The transpose and conjugate transpose of a vector or matrix is denoted by the symbols $(\cdot)^T$ and $(\cdot)^*$, respectively. The notation $(\cdot)^c$ indicates the set complement. $\mathrm{Re}(\cdot)$, $\mathrm{Im}(\cdot)$, $\mathrm{rank}(\cdot)$ and $\mathrm{Tr}(\cdot)$ denote the real part, imaginary part, rank and trace of a given scalar or matrix. The notations $\|x\|_1$, $\|x\|_2$ and $\|X\|_F$ indicates 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 signifies 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 denotes the $N \times N$ identity matrix. For a given vector x, the symbol $\text{diag}(x)$ denotes its diagonalized matrix, whereas for a matrix X, $\text{diag}(X)$ denotes the vector consisting of the diagonal elements of X. The i-th smallest eigenvalue of the matrix X is denoted by $\lambda_i(X)$. 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 imaginary unit is denoted by $\mathbf{j} = \sqrt{-1}$. The symbol $\mathbb{1}$ denotes a vector of all ones with appropriate dimension.

B. POWER SYSTEM SCADA MEASUREMENTS

Let an electric power network be described by a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, \ldots, K\}$ and $\mathcal{E} := \{1, \ldots, L\}$ denote the sets of buses and lines (branches), respectively. We make the assumption that the slack bus is also the reference bus. Let $v_k \in \mathbb{C}$ represent the complex voltage at bus $k \in \mathcal{V}$, whose magnitude and phase angle are denoted as $|v_k|$ and $\measuredangle v_k$. The net apparent power injected at bus k is denoted by $s_k = p_k + q_k\mathbf{j}$. Given a fixed orientation on the branches, there are two complex power flows associated with each line. Define $s_{l,f} = p_{l,f} + q_{l,f}\mathbf{j}$ and $s_{l,t} = p_{l,t} + q_{l,t}\mathbf{j}$ as the complex power flows coming into the line $l \in \mathcal{E}$ through the 'from' and 'to' end of the branch. Let v and i be the vectors of nodal complex voltages and net current injections, respectively. Following Ohm's law, we know that

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

where $Y \in \mathbb{C}^{K\times K}$ symbolizes the nodal admittance matrix of the power network. Furthermore, $Y_f \in \mathbb{C}^{L\times K}$ and $Y_t \in \mathbb{C}^{L\times K}$ represent the 'from' and 'to' branch admittance matrices. Let $\{e_1, \ldots, e_K\}$ denote the canonical vectors in $\mathbb{R}^K$. Define the following three nodal measurement matrices:

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

Next, let $\{d_1, \ldots, d_L\}$ be the canonical vectors in $\mathbb{R}^L$. Define the following four line measurement matrices associated with branch l, which has a from node i and a to node j:

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

Then, the traditional measurable quantities can be expressed as the following seven equations, each of which is a simple quadratic function of the complex voltage vector v.

$$|v_k|^2 = \text{Tr}(E_k v v^*) \tag{4a}$$
$$p_k = \text{Tr}(Y_{k,p} v v^*), \quad q_k = \text{Tr}(Y_{k,q} v v^*) \tag{4b}$$
$$p_{l,f} = \text{Tr}(Y_{l,p_f} v v^*), \quad p_{l,t} = \text{Tr}(Y_{l,p_t} v v^*) \tag{4c}$$
$$q_{l,f} = \text{Tr}(Y_{l,q_f} v v^*), \quad q_{l,t} = \text{Tr}(Y_{l,q_t} v v^*) \tag{4d}$$

In a power system, measurements are acquired through the SCADA system. Available measurements consist a subset of the entire measurable quantities. Given a power system model Ω characterized by the tuple (Y, Y_f, Y_t) and an index set of measurements $\mathcal{M} = \{1, \ldots, M\}$ of the form (4), the mapping from the measurement index set to the set of measurement matrices can be defined as

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

where each $A_j(\Omega)$ represents one of the matrices defined in (2) and (3), depending on the type of measurement j.

Next, 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 $2K - 1$ because the voltage angle at the slack/reference bus is fixed to be zero. Accordingly, the matrices also have $2K - 1$ rows and columns.

Definition 1. *For the state vector $v \in \mathbb{C}^K$, define $\bar{v} \triangleq [Re\{v[\mathcal{V}]^T\}\ Im\{v[\mathcal{O}]^T\}]^T \in \mathbb{R}^{2K-1}$ to be the real-valued state vector where $\mathcal{O}$ denotes the set of all buses except for the slack bus. In addition, define $\bar{X} \in \mathbb{S}^{2K-1}$ to be the real-valued symmetrization of $X \in \mathbb{H}^K$. To elaborate, note that a general $K \times K$ Hermitian matrix can be mapped into a $(2K-1)\times(2K-1)$ real-valued symmetric matrix as follows:*

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

Finally, we define an operator that maps the state vector to the vector of measurement values.

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

$$h^\Omega(\bar{v}) \triangleq [v^T A_1(\Omega) v \ \cdots \ v^T A_M(\Omega) v]^T \tag{7}$$
$$= [\bar{v}^T \bar{A}_1(\Omega) \bar{v} \ \cdots \ \bar{v}^T \bar{A}_M(\Omega) \bar{v}]^T \tag{8}$$

In this paper, we disregard PMU measurements and only consider voltage magnitude and power measurements to streamline the presentation without loss of generality in our technique. More precisely, if we have access to PMU measurements, they can be viewed as quadratic equations with zero quadratic terms and can be easily incorporated in the current framework. To elaborate, we can append the real-valued state vector $\bar{v}$ with a scalar variable u and impose the condition $u^2 - 1 = 0$ so that its interaction with the other variables can create linear terms. Note that this can result in u taking the value of 1 or -1. If the solution that we obtain

results in $u = -1$, we can simply negate the rest of the values to obtain a meaningful solution.

III. MAIN RESULTS

In this section, we first briefly discuss the most commonly used nonlinear least-squares (NLS) state estimator and its limitations. Next, we present the *NLAV* formulation and derive a theoretical upper bound on the state estimation error obtained by the NLAV estimator. Finally, we uncover certain properties of the vector of residual errors and design a new algorithm that performs state estimation and topology error detection in a jointly fashion.

A. NONLINEAR LEAST-SQUARES STATE ESTIMATION

The most widely used state estimation technique is the nonlinear least-squares (*NLS*), first proposed by Schweppe [23], [24]. The objective of *NLS* is to minimize the ℓ_2-norm of the estimation residuals, which is often executed by local search algorithms such as the Gauss-Newton method. These methods however only guarantee a locally optimal solution, which could correspond to an estimate of the state that is significantly different from the true underlying voltages. Interestingly, recent research have shown that local search algorithms are capable of finding a globally optimal solution of this nonconvex problem when the number of measurements is relatively higher than the degree of the freedom of the system and the measurements are noiseless [4], [20]. As is the case with any other estimator, this method requires that the system's network topology (see Definition 5) be known. However, owing to the existence of topological errors resulting from simple faults or recent changes in the switching status of some lines, the model that the system operator has at hand may be different from the true network. The measurement data at the neighborhood of the incorrectly modeled lines are potential outliers, which can adversely impact the solution of the state estimation problem over a large fraction of the network. This is due to the fact that the ℓ_2-norm is inadequate 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] developed 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. As mentioned in Section I, there are existing works that have studied the nonlinear least absolute value estimator [18], [19]. In fact, [19] considered a wide variety of parameter errors and the presented numerical experiments are very promising. In this paper, we focus on topological errors in the model and develop theoretical guarantees along with a graph-based intuition for detecting those errors.

B. PROPOSED NLAV FORMULATION

For the remainder of this paper, a line whose presence in the network is misrepresented by the system operator is called *erroneous* and the set of all *erroneous* lines is denoted by Ξ. Let $\mathcal{A}^{\Omega}(\mathcal{M})$ denote the set of measurement matrices associated with the true system Ω, and $\mathcal{A}^{\widetilde{\Omega}}(\mathcal{M})$ denote the set of measurement matrices corresponding to the inaccurate model $\widetilde{\Omega}$ that the system operator possesses. In this work, we make the assumption that Ω and $\widetilde{\Omega}$ are sparsely different, in other words, the set of lines for which the operator misconceives their switch statuses constitutes only a small subset of the entire lines. It makes sense to only focus on sparse differences because topological errors often occur from low probability events and therefore it is unlikely that the operator's model be significantly different from the true model. We propose the following optimization problem as the first step to designing an algorithm that jointly performs state estimation and sparse topological error detection:

$$\min_{\bar{v}\in\mathbb{R}^{2K-1}} \bar{v}^T \bar{A}_0 \bar{v} + \rho \sum_{j=1}^{M} |\bar{v}^T \bar{A}_j(\widetilde{\Omega})\bar{v} - b_j| \tag{9}$$

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

$$b = h^{\Omega}(\bar{z}) + \eta. \tag{10}$$

In the above equation, $\bar{z} \in \mathbb{R}^{2K-1}$ symbolizes the true underlying state of the system and η denotes the noise vector. Note that the measurement values b are based on the true system Ω and $\bar{z}$. Also, $A_0 \in \mathbb{S}^K$ is a regularization matrix while ρ is a regularization coefficient. As discussed later in the paper, these two parameters help with deriving an upper bound on the state estimation error and also facilitate 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 be full row rank [25]. Let $\bar{v}_*$ denote a globally optimal solution of (9). Then, let $\epsilon \in \mathbb{R}^{2K-1}$ be the state estimation error vector and $r \in \mathbb{R}^K$ be the residual error vector, defined as

$$\epsilon = \bar{v}_* - \bar{z} \tag{11a}$$

$$r_j = |\bar{v}_*^T \bar{A}_j(\widetilde{\Omega})\bar{v}_* - b_j|, \quad \forall j \in \mathcal{M} \tag{11b}$$

By virtue of the ℓ_1 norm, the problem (9) attempts to push the insignificant residual errors to hard zeros, while the residuals r_j's associated with the outlier measurements are expected to remain nonzero. This phenomenon is observed empirically through 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 across all the measurements. In the sections that follow, we use this intuition to design an efficient topological error detection algorithm. Bear in mind that, similar to the NLS method, the objective function of NLAV is nonlinear and nonconvex, which makes local search algorithms prone to being stuck at 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 obtaining their global solutions using local search algorithms [4], [20]. Therefore, having access to many measurements is crucial for improving the quality of real-world state estimation problems. This property is expected to hold for the NLAV estimator too, as partially proven in [21]. The possibility of falling into a local optimum can be further avoided by initializing the algorithm close to the unknown state. This is achievable because in power systems, voltage magnitudes are maintained close to 1 and voltage angles are kept 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 A_0, we intend to prove a theoretical upper bound on the state estimation error obtained by the *NLAV* problem (9). To this end, it is useful to introduce the concept of dual certificate:

Definition 3. *Given a positive-semidefinite regularization matrix $A_0 \in \mathbb{S}^K$, a system model Ω and a set of measurement matrices $\mathcal{A}^{\Omega}(\mathcal{M})$, define $H_{\mu}^{\Omega} \triangleq A_0 + \sum_{j=1}^{M} \mu_j A_j(\Omega)$. A vector $\mu \in \mathbb{R}^M$ is called a dual certificate for the voltage vector $v \in \mathbb{C}^K$ of the system model Ω if it satisfies the following three conditions:*

$$H_{\mu}^{\Omega} \succeq 0, \quad H_{\mu}^{\Omega} v = 0, \quad rank\{H_{\mu}^{\Omega}\} = K - 1 \tag{12}$$

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

$$\frac{\|\bar{v}_* - \beta \cdot \bar{z}\|_2^2}{\|\bar{v}_*\|_2} \leq \sqrt{\frac{4K \cdot g(\bar{z}, \eta, \rho)}{\lambda_2(H_{\mu}^{\widetilde{\Omega}})}} \tag{13}$$

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

$$\rho \left[\sum_{j \in \mathcal{N}} |\bar{z}^T (\bar{A}_j(\Omega) - \bar{A}_j(\widetilde{\Omega})) \bar{z}| + \sum_{j=1}^{M} |\eta_j| \right] \tag{14}$$

with $\mathcal{M}' \subset \mathcal{M}$ being the set of measurement indices that correspond to the erroneous lines.

By recalling that $\bar{v}_*$ and $\bar{z}$ are, respectively, the recovered and true states of the system, inequality (13) quantitatively bounds the state estimation error. The bound has several important characteristics. First, if the measurements are noiseless and there is no topology error, the NLAV estimator recovers a high-quality solution if not the actual state. On the other hand, if there are measurement noise and topology error, the upper bound for the state estimation error increases proportionally to the magnitude of noise and the number of topology errors. Note that topology errors do not affect all the measurement matrices but only a subset, which is captured by the set $\mathcal{M}'$. Second, the upper bound is inversely proportional to the second smallest eigenvalue of the matrix $H_{\mu}^{\widetilde{\Omega}}$, which acts as the Laplacian of a weighted graph corresponding to the power network. The second smallest eigenvalue of this matrix is also called the algebraic connectivity [26] in graph theory, a parameter that gauges how well-connected the (weighted) graph is. For instance, a fully connected graph has the algebraic connectivity of K while this value is equal to 2 for a star-shaped graph and $2(1 - cos\frac{\pi}{K})$ for a path graph (where K denotes the number of nodes in the graph). In the special case when A_0 reflects the connectivity of the original network $\mathcal{G}$ (i.e., $i \neq j$ and $(i,j) \notin \mathcal{E} \Longrightarrow A_0(i,j) = 0$), the second smallest eigenvalue of $H_{\mu}^{\widetilde{\Omega}}$ represents the algebraic connectivity of the original network where different edges are assigned with different weights. As a final note, a unique solution of the NLAV is not guaranteed by the performance bound in equation (13). For conditions that guarantee the uniqueness of NLAV 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 outlined as follows. To start, we solve (9) and utilize the pattern of the nonzero residuals errors to identify 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.

Definition 4. *A node $k \in \mathcal{V}$ is called unsolvable if ϵ_k is nonzero. On the other hand, if ϵ_k is zero, node k is called solvable. Define the following four subgraphs of $\mathcal{G}$:*

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

In order to help the reader visualize the different subgraphs, we illustrate Definition 4 for a small system in

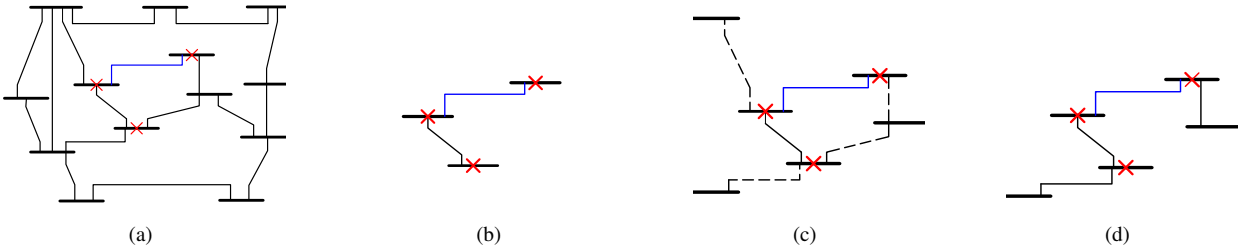

Figure 1. (a) A power system network, (b) the state estimation error graph $\mathcal{S}$, (c) the extended state estimation error graph $\widetilde{\mathcal{S}}$, and (d) the line residual graph $\mathcal{R}^L$. Nodes marked with red crosses are the unsolvable nodes, the blue line represents the erroneous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

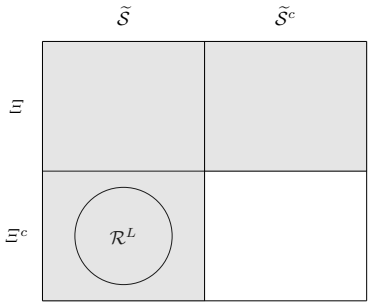

Figure 2. A diagram showing the relationship between different subgraphs. Each rectangle represents the intersection between two different sets. For example, the upper-left rectangle represents $\Xi \cap \widetilde{\mathcal{S}}$. $\mathcal{R}^N$ is equivalent to the gray-colored area.

Figure 1. In Theorem 2, we reveal how the *set of erroneous lines*, namely Ξ, relates to these subgraphs.

Theorem 2. *Suppose that the measurements are noiseless, i.e., $\eta = 0$. In addition, assume that there do not exist any two distinct vectors of voltages resulting in the same measurement values, i.e.,*

$$\bar{x} \neq \bar{y} \Longrightarrow \|h^{\Omega}(\bar{x}) - h^{\Omega}(\bar{y})\|_1 \neq 0 \tag{15}$$

Then,

$$\mathcal{R}^L \subseteq (\Xi^c \cap \widetilde{\mathcal{S}}) \tag{16}$$

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

$$\mathcal{R}^N = \widetilde{\mathcal{S}} \cup \Xi \tag{17}$$

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 Figure 2, it can also be inferred that $\Xi \subseteq (\mathcal{R}^N \setminus \mathcal{R}^L) \subseteq (\widetilde{\mathcal{S}}^c \cap \Xi^c)^c$. Henceforth, the pragmatic benefit of Theorem 2 is that it enables us to develop a method for efficiently detecting topology errors by probing over a small subgraph of the original power system model. We call this small subgraph, namely $(\mathcal{R}^N \setminus \mathcal{R}^L)$, the *suspect-subgraph*.

Algorithm 1 Suspect-subgraph search algorithm

Given: Current model $\widetilde{\Omega}$ and measurement vector b
Initialize: Set $\mathcal{D}_L = \emptyset$, $\epsilon > 0$, $\delta > 0$, $\mu > 0$
and calculate $\mathcal{A}_{\widetilde{\Omega}}(\mathcal{M})$ using (5).
1. Solve NLAV problem (9) with $\widetilde{\Omega}$, $\mathcal{A}_{\widetilde{\Omega}}(\mathcal{M})$ and b, and calculate the residual r based on equation (11b).
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 altering the on/off status of l.
    Re-solve (9) with $\Omega^{t'}$, $\mathcal{A}_{\Omega^{t'}}(\mathcal{M})$ 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 $\mathcal{D}_L$

E. ALGORITHM

Based on the results established so far, we propose Algorithm 1 for detecting topology errors while performing state estimation. Algorithm 1 begins by initializing the set of detected erroneous lines, denoted by $\mathcal{D}_L$, with the empty set. Then, the algorithm inspects all branches in the *suspect-subgraph* $(\mathcal{R}^N \setminus \mathcal{R}^L)$, and computes the effect that the existence of each line has on the accuracy of solution. Consequently, the method switches a line on if it is off in the model and vice versa, updates the model based on this change, and re-solves the NLAV problem with the modified model. If the objective value of NLAV goes down, the line is stored in $\mathcal{D}_L$; otherwise, the update of line status is dismissed and the algorithm proceeds to check another line until all lines of $(\mathcal{R}^N \setminus \mathcal{R}^L)$ are evaluated. The justification for using such a criteria is explained in the Appendix section D.

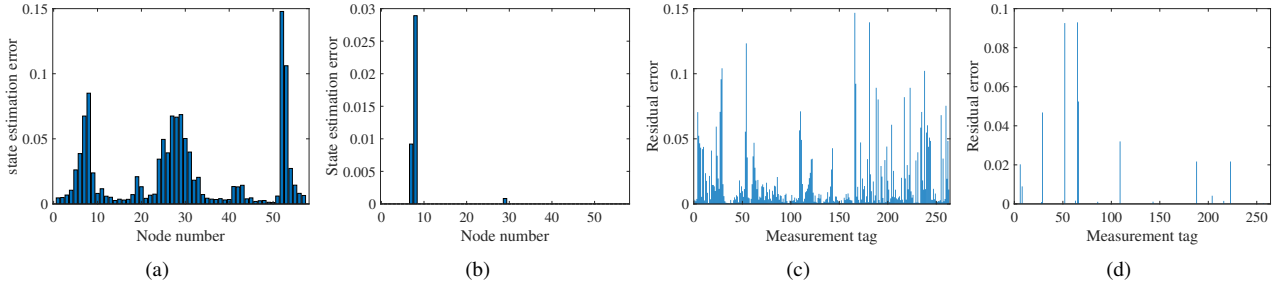

Figure 3. Noiseless state estimation error for (a) NLS, (b) NLAV; and residuals for (c) NLS, (d) NLAV. Note that in (c) and (d), the x-axis shows the measurement tag, which is not the same as the node or line number due to the concatenation of different types of measurements.

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 A_0 for simplicity and call this the *unpenalized NLAV* problem (in other words, we set A_0 to 0). Without prior knowledge of the state, designing a favorable A_0 penalty term could be difficult, in which case setting A_0 to zero makes logical sense. Theorem 3 provides a sufficient condition under which the *unpenalized NLAV* problem has a unique solution. Since without A_0, the state estimation error bound provided in Theorem 1 is no longer valid, Theorem 3 also provides a new bound.

Definition 5. *Given a system model Ω and a set of measurements $\mathcal{M}$, define the linear map $\mathcal{H}^{\Omega}: \mathbb{R}^{K\times K} \to \mathbb{R}^{M}$ as*

$$\mathcal{H}^{\Omega}(X) = [\langle A_1(\Omega), X\rangle \cdots \langle A_M(\Omega), X\rangle]^T \tag{18}$$

Theorem 3. *Given the true network model Ω and the measurement set $\mathcal{M} = \{1,\ldots,M\}$, let $\mathcal{H}^{\Omega}$ be the mapping as defined in Definition 5. Then, $\bar{v}^*$ obtained from solving the NLAV problem (9) with $A_0 = 0$ satisfies:*

$$\|\bar{v}_*\bar{v}_*^T - \bar{z}\bar{z}^T\|_F \le \frac{2}{t}\|\eta\|_1 \tag{19}$$

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

$$\begin{aligned} &\min_{X\in\mathbb{S}^n} \|\mathcal{H}^{\Omega}(X)\|_2 \\ &s.t. \quad rank(X) = 2,\ \|X\|_F = 1 \end{aligned} \tag{20}$$

One can easily verify that there does not exist any set of noiseless measurements for the model Ω that leads to non-unique exact solutions if and only if $t > 0$. That is to say, if $t > 0$, any global optimizer of the NLAV problem matches the true underlying state that we hope to find (note that this is for when all topological errors have been identified and corrected). Therefore, t can be interpreted as a quantification of the measurement set's capability to generate a unique solution of the over-determined power flow equations. In addition, if $t > 0$, then condition(15) 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 $\mathcal{H}: \mathbb{R}^{K\times K} \to \mathbb{R}^{M}$ 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: $(1-\delta_r)\|X\|_F^2 \le \frac{1}{p}\|\mathcal{H}(X)\|^2 \le (1+\delta_r)\|X\|_F^2$. If $\mathcal{H}$ satisfies $(2r,\delta_{2r})$-RIP with $\delta_{2r} < 1$, then finding a global optimum constitutes exact recovery of the state [27]. 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, $\mathcal{H}$ suffices to satisfy $(2r,\delta_{2r})$-RIP with $\delta_{2r} < 0.2$, which is a strict condition [28]. A milder condition on RIP for structured mappings (such as power subsystems) has been developed in [29]. 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 assess the efficacy of the proposed *NLAV* algorithm for detecting topological errors, this section presents numerical simulations on the IEEE 57-bus system and the 118-bus system. For running the simulations, we use MATPOWER data along with the MATLAB *fmincon* 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 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. More specifically, we select the unknown state by sampling each voltage magnitude from a normal distribution with mean equal to 1 and standard deviation equal to 0.1. This is more than a reasonable range since most voltage magnitudes lie within 5% of the nominal value. Voltage angles, computed in radians, are sampled from a normal distribution with mean equal to 0 and standard deviation equal to 0.1. In order to assess the performance of the algorithm, we calculate the true/false positive rates and the suspect rate as:

$$\text{True positive rate} = |\mathcal{D}_L \cap \Xi|/|\Xi| \tag{21a}$$

$$\text{False positive rate} = |\mathcal{D}_L \cap \Xi^c|/|\Xi| \tag{21b}$$

$$\text{Suspect rate} = |(\mathcal{R}_N \setminus \mathcal{R}^L) \cap \Xi|/|\Xi| \tag{21c}$$

In addition, we also report the number of lines that the algorithm checks before termination, which is simply the cardinality of the set $(\mathcal{R}_N \setminus \mathcal{R}^L)$. Note that for most of this section (with the exception of subsection IV-E), the measurement values are assumed to be noiseless.

B. EXAMPLE: SPARSE RESIDUALS FOR NLAV

Before analyzing the bulk of simulations data, we concentrate on a specific example to visually illustrate the ideas discussed in Section III. The example under scrutiny is for the 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* in the presence of topological errors. It can be observed from these plots that there is an absence of sparsity pattern, and the large peaks are not even related to the end points of the *erroneous* lines. This indicates that we need to scan over all realizable combinations of transmission lines to detect 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* (i.e. in the presence of all the topolgical errors) are shown in Figure 3(b) and 3(d). The largest peaks of the residual vector in this plot are associated with the nodes/lines that are directly connected to (or correspond to) the erroneous lines. This implies that the erroneous lines can be detected by searching over only those lines that are related to the largest peaks of the residual vector. Consequently, as stated in Algorithm 1, the two erroneous lines are correctly identified. 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. The computational time for each run ranges from 5 to 30 seconds (depending on the number of erroneous lines considered) on a laptop with 16GB RAM and an Intel i7-8750H processor.

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 1 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. The computational time for each run with 5 erroneous lines ranges from 1 to 3 minutes on a laptop with 16GB RAM and an Intel i7-8750H processor.

E. NOISY MEASUREMENTS

So far, the numerical simulations have been performed with noiseless measurements. When measurements are tainted with noise, this can affect the overall residual vector and in turn hinder the topology error detection. In order to analyze this more carefully, first consider Figure 5(a) which shows the nodal measurement residuals obtained by solving the initial NLAV on the IEEE-57 bus system with 30 percent line measurement, no noise in measurement values and the regularization matrix A_0 set to zero. The erroneous lines are line 8 and line 67, same as with the example in Figure 3. Note that this is a different instance of the problem with a distinct set of measurements and a distinct underlying true state. The

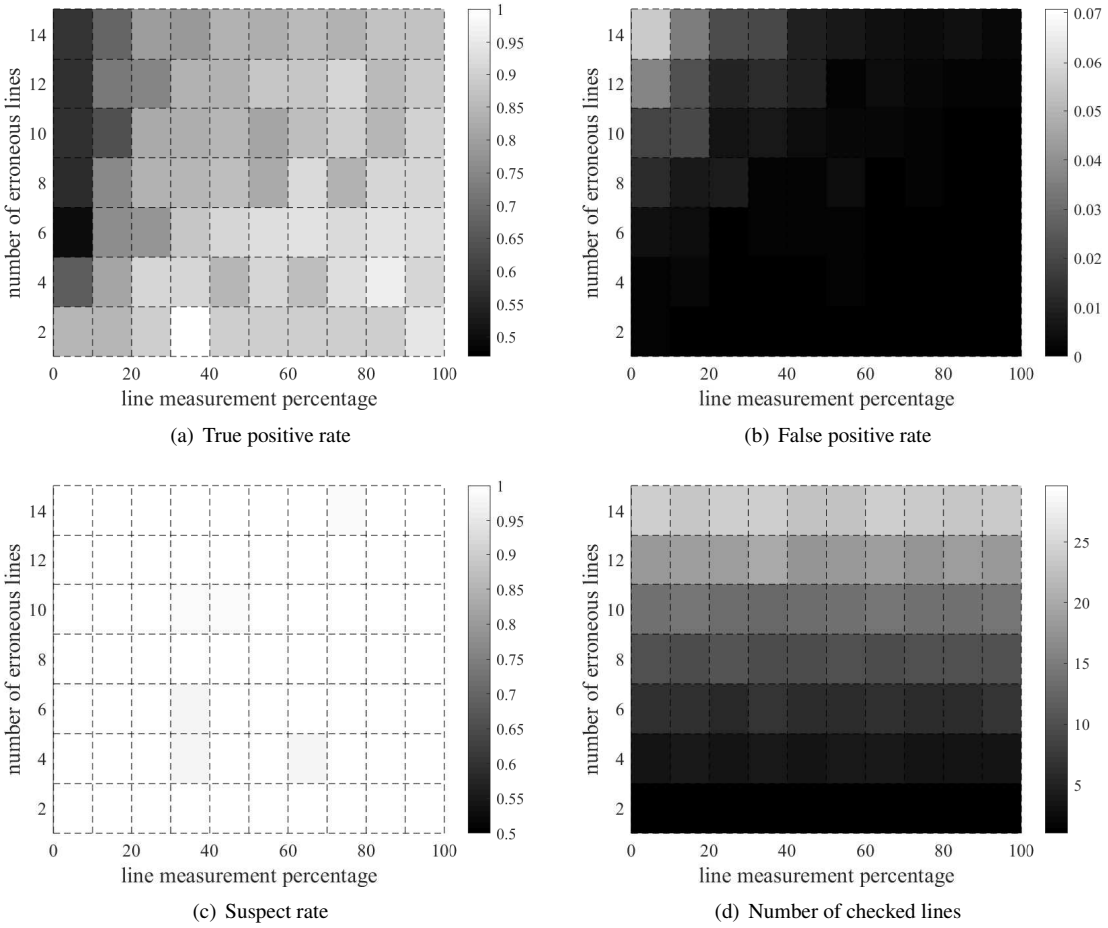

Figure 4. Simulation results on the IEEE 57-bus system. Each value represents the average over 20 simulations.

Table 1. 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

end-points of the two lines are buses $\{8, 9\}$ and $\{29, 52\}$, respectively. In Figure 5(a), we can see that these buses are identified, leading to the correct detection of the erroneous lines. However, once noise is added to the measurements, not only do we get a residual vector that is not sparse, but also the peaks of the residuals sometimes fail to identify the relevant buses. This is illustrated in Figure 5(b), where the peaks of the residuals do not capture bus 8 and bus 29, leading to a failure in topology error detection. An interesting result is observed once we add more line measurements and also a nonzero regularization matrix $A_0 = Y^*Y$. This result is shown in Figure 5(c), where we can observe that the residuals corresponding to bus 8 and bus 29 are shown peaking as opposed to being obsolete in the Figure 5(b). Therefore, we see that a high number of measurements combined with an appropriate regularizer offers a state estimator that is more robust to noise.

V. CONCLUSION

In this article we propose a novel methodology to solve the state estimation problem for power systems when there exists a modest number of topology errors and also to identify such modeling errors. The established technique minimizes a nonconvex function corresponding to the ℓ_1-norm of the nonconvex residual errors plus a convex quadratic regularization term. We show that, under mild assumptions, the presented

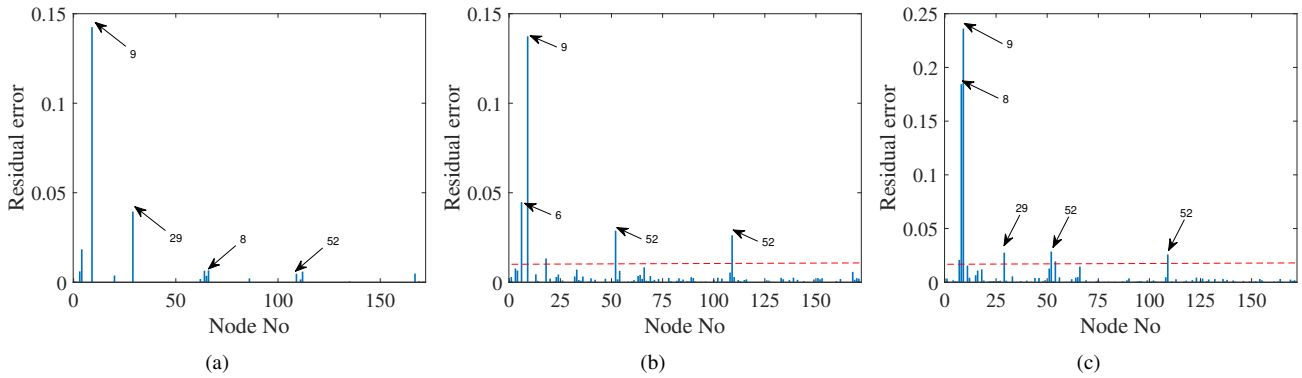

Figure 5. Residuals for NLAV on the IEEE-57 bus system with (a) 30 percent line measurement and no noise in measurement values, (b) 30 percent line measurement and noisy measurement values (c) 70 percent line measurement and noisy measurement values. Note that for all graphs, the x-axis shows the nodal measurement tag, which is not necessarily the same as the bus number due to the fact we have multiple types of measurements concatenated. The numbers next to the arrows show the actual bus numbers.

methodology can efficiently identify the topological errors by searching over the lines contained in a small suspect-subgraph of the network inferred by the initial solution of the NLAV estimator. Two theoretical upper bounds on the estimation errors are derived. The first upper bound shows that the state estimation error depends on the connectivity of the network and the aggregate quality of the measurement set. The second upper bound on the state estimation error is inversely related to a parameter t that is positive if and only if there do not exist two voltage profiles that can generate the same set of measurement values. Finally, the theoretical results are validated by testing the algorithm on several benchmark systems. The results imply that the proposed algorithm is capable of accurately detecting topological errors given that the number of measurements is large enough.

A. PROOF OF THEOREM 1

Before going into the proof, we impose the following two conditions for A_0:

Assumption 1. *The regularizer matrix A_0 satisfies the following properties:*

1) $A_0 \succeq 0$

2) $A_0 \cdot \mathbb{1} = 0$

Consider the NLAV problem (9). One can create lower and upper bounds on the optimal objective value as follows:

$$\begin{aligned}
&\bar{v}_*^T \bar{A}_0 \bar{v}_* + \rho \sum_{j=1}^{M} |\bar{v}_*^T \bar{A}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{A}_j(\Omega) \bar{z}| - \rho \sum_{j=1}^{M} |\eta_j| \\
&\overset{(a)}{\leq} \bar{v}_*^T \bar{A}_0 \bar{v}_* + \rho \sum_{j=1}^{M} |\bar{v}_*^T \bar{A}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{A}_j(\Omega) \bar{z} - \eta_j| \\
&\overset{(b)}{\leq} \bar{z}^T \bar{A}_0 \bar{z} + \rho \sum_{j=1}^{M} |\bar{z}^T \bar{A}_j(\widetilde{\Omega}) \bar{z} - \bar{z}^T \bar{A}_j(\Omega) \bar{z} - \eta_j| \\
&\overset{(c)}{\leq} \bar{z}^T \bar{A}_0 \bar{z} + \rho \sum_{j \in \mathcal{M}'} |\bar{z}^T \bar{A}_j(\widetilde{\Omega}) \bar{z} - \bar{z}^T \bar{A}_j(\Omega) \bar{z}| + \rho \sum_{j=1}^{M} |\eta_j|
\end{aligned}$$

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

$$\begin{aligned}
&\bar{v}_*^T \bar{A}_0 \bar{v}_* - \bar{z}^T \bar{A}_0 \bar{z} + \rho \sum_{j=1}^{M} |\bar{v}_*^T \bar{A}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{A}_j(\Omega) \bar{z}| \\
&\leq \rho \sum_{j \in \mathcal{M}'} |\bar{z}^T \bar{A}_j(\widetilde{\Omega}) \bar{z} - \bar{z}^T \bar{A}_j(\Omega) \bar{z}| + 2\rho \sum_{j=1}^{M} |\eta_j| \qquad (22)
\end{aligned}$$

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

$$\begin{aligned}
&\bar{v}_*^T \bar{A}_0 \bar{v}_* - \bar{z}^T \bar{A}_0 \bar{z} + \rho \sum_{j=1}^{M} |\bar{v}_*^T \bar{A}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{A}_j(\widetilde{\Omega}) \bar{z}| \\
&\leq 2\rho \left\{ \sum_{j \in \mathcal{M}'} |\bar{z}^T (\bar{A}_j(\widetilde{\Omega}) - \bar{A}_j(\Omega)) \bar{z}| + \sum_{j=1}^{M} |\eta_j| \right\} \\
&= 2g(\bar{z}, \eta, \rho) \qquad (23)
\end{aligned}$$

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

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

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

$$\begin{aligned}
\min_{t} \quad & \bar{v}_*^T \bar{A}_0 \bar{v}_* - \bar{z}^T \bar{A}_0 \bar{z} + \rho \sum_{j=1}^{M} t_j \\
s.t. \quad & \bar{v}_*^T \bar{A}_j(\widetilde{\Omega}) \bar{v}_* - \bar{z}^T \bar{A}_j(\widetilde{\Omega}) \bar{z} \leq t_j, \quad \forall j \in \mathcal{M} \\
& -\bar{v}_*^T \bar{A}_j(\widetilde{\Omega}) \bar{v}_* + \bar{z}^T \bar{A}_j(\widetilde{\Omega}) \bar{z} \leq t_j, \quad \forall j \in \mathcal{M}
\end{aligned} \qquad (24)$$

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

$$\mathcal{L}(t,p^+,p^-) = \bar{v}_*^T \bar{A}_0 \bar{v}_* - \bar{z}^T \bar{A}_0 \bar{z} + \sum_{j=1}^{M}(\rho - p_j^+ - p_j^-)t_j + \sum_{j=1}^{M}\{(p_j^+ - p_j^-)(\bar{v}_*^T \bar{A}_j(\widetilde{\Omega})\bar{v}_* - \bar{z}^T \bar{A}_j(\widetilde{\Omega})\bar{z})\} \quad (25)$$

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

$$d(p^+,p^-) = \bar{v}_*^T \left(\bar{A}_0 + \sum_{j=1}^{M}(p_j^+ - p_j^-)\bar{A}_j(\widetilde{\Omega})\right)\bar{v}_* - \bar{z}^T \left(\bar{A}_0 + \sum_{j=1}^{M}(p_j^+ - p_j^-)\bar{A}_j(\widetilde{\Omega})\right)\bar{z} \quad (26)$$

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:

$$d(p_*^+,p_*^-) = \bar{v}_*^T H_\mu^{\widetilde{\Omega}} \bar{v}_* - \bar{z}^T H_\mu^{\widetilde{\Omega}} \bar{z} \quad (27)$$

$$= \mathrm{Tr}\{H_\mu^{\widetilde{\Omega}} \bar{v}_* \bar{v}_*^T\} = \mathrm{Tr}\{H_\mu^{\widetilde{\Omega}} X\} \quad (28)$$

The rest of the proof can be adopted from [30] (Appendix, Proof of Theorom 2). Consider an eigen-decomposition of $H_\mu^{\widetilde{\Omega}} = U \Lambda U^T$, where $\Lambda = \mathrm{diag}(\lambda_{2K-1}, ..., \lambda_1)$ such that $\lambda_{2K-1} \geq \cdots \geq \lambda_1$ and U is a unitary matrix whose columns are the corresponding eigenvectors. Define

$$\breve{X} := \begin{bmatrix} \widetilde{X} & \widetilde{x} \\ \widetilde{x}^T & \alpha \end{bmatrix} = U^T X U \quad (29)$$

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

$$Tr\big(H_\mu^{\widetilde{\Omega}} X\big) = Tr\big(U\Lambda U^T U \breve{X} U^T\big) = Tr\big(\Lambda \breve{X}\big) \geq \lambda_2(H_\mu^{\widetilde{\Omega}}) Tr\big(\widetilde{X}\big) \quad (30)$$

Combining (30) and (23) leads to

$$\mathrm{Tr}\big(\widetilde{X}\big) \leq 2 \cdot g(\bar{z},\eta,\rho)/\lambda_2(H_\mu^{\widetilde{\Omega}}) \quad (31)$$

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

$$X = U\breve{X}U^T = \begin{bmatrix} \widetilde{U} & \widetilde{z} \end{bmatrix} \begin{bmatrix} \widetilde{X} & \widetilde{x} \\ \widetilde{x}^T & \alpha \end{bmatrix} \begin{bmatrix} \widetilde{U}^T \\ \widetilde{z}^T \end{bmatrix} = \widetilde{U}\widetilde{X}\widetilde{U}^T + \widetilde{U}\widetilde{x}\widetilde{z}^T + \widetilde{z}\widetilde{x}^T\widetilde{U}^T + \alpha\widetilde{z}\widetilde{z}^T \quad (32)$$

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

$$\|\widetilde{x}\|_2^2 \leq \alpha \mathrm{Tr}(\widetilde{X}) = \mathrm{Tr}(X)\mathrm{Tr}(\widetilde{X}) - \mathrm{Tr}^2(\widetilde{X}) \quad (33)$$

Therefore,

$$\begin{aligned} \|X - \alpha\widetilde{z}\widetilde{z}^T\|_F^2 &= \|\widetilde{U}\widetilde{X}\widetilde{U}^T + \widetilde{U}\widetilde{x}\widetilde{z}^T + \widetilde{z}\widetilde{x}^T\widetilde{U}^T\|_F^2 \\ &\overset{(d)}{=} \|\widetilde{U}\widetilde{X}\widetilde{U}^T\|_F^2 + 2\|\widetilde{z}\widetilde{x}^T\widetilde{U}^T\|_F^2 \overset{(e)}{=} \|\widetilde{X}\|_F^2 + 2\|\widetilde{x}\|_2^2 \\ &\leq \|\widetilde{X}\|_F^2 - 2\mathrm{Tr}^2(\widetilde{X}) + 2\mathrm{Tr}(X)\mathrm{Tr}(\widetilde{X}) \\ &\overset{(f)}{\leq} 2\mathrm{Tr}(X)\mathrm{Tr}(\widetilde{X}) \\ &\overset{(g)}{\leq} \frac{4g(\bar{z},\eta,\rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}})}\mathrm{Tr}(X) \end{aligned} \quad (34)$$

where (d) follows from the fact that $\widetilde{U} * \widetilde{z} = 0$, (e) is due to $\widetilde{U}^T\widetilde{U} = I_{2K-2}$ and (f) comes from the fact that $\|\widetilde{X}\|_F \leq \mathrm{Tr}(\widetilde{X})$. Finally, (g) results from substituting equation (31). Plugging back in $X = \bar{v}_*\bar{v}_*^T$ yields that

$$\|X - \alpha\widetilde{z}\widetilde{z}^T\|_F^2 = \|\bar{v}_*\bar{v}_*^T - \frac{\alpha}{\|\bar{z}\|_2^2}\bar{z}\bar{z}^T\|_F^2 \leq \frac{4g(\bar{z},\eta,\rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}}(\bar{A}_0))}\mathrm{Tr}(\bar{v}_*\bar{v}_*^T) \quad (35)$$

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

$$\|\bar{v}_*\bar{v}_*^T - \beta\bar{z}\bar{z}^T\|_F^2 \leq \frac{4g(\bar{z},\eta,\rho)}{\lambda_2(H_\mu^{\widetilde{\Omega}}(\bar{A}_0))}\|\bar{v}_*\|_2^2 \quad (36)$$

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

$$\begin{aligned} \|\bar{v}_*\bar{v}_*^T - \beta\bar{z}\bar{z}^T\|_F &= \sqrt{\sum_{i,j}[\bar{v}_*(i)\bar{v}_*(j) - \beta\cdot\bar{z}(i)\bar{z}(j)]^2} \\ &\geq \sqrt{\sum_i[\bar{v}_*(i)^2 - \beta\cdot\bar{z}(i)^2]^2} \overset{(h)}{\geq} \sqrt{\sum_i[\bar{v}_*(i) - \beta\cdot\bar{z}(i)]^4} \\ &\overset{(i)}{\geq} \frac{1}{\sqrt{K}}\sum_i[\bar{v}_*(i) - \beta\cdot\bar{z}(i)]^2 = \frac{1}{\sqrt{K}}\|\bar{v}_* - \beta\cdot\bar{z}\|_2^2 \end{aligned} \quad (37)$$

where (h) and (i) are due to Cauchy-Schwarz and Holder's inequality, respectively. Now combining this inequality with (36) leads to

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

which completes the proof.

B. PROOF OF THEOREM 2

Define $\mathcal{N}(k)$ 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 $\mathcal{N}(i) \cup \mathcal{N}(j)$ are *solvable*. Also, since $l \in \Xi$, the nodal residual at nodes i and j are nonzero, which means that $i, j \in \mathcal{V}_N$. Finally, noting that $l \notin \mathcal{R}^L$ because there is no line measurement for an erroneous line, we can conclude that $l \in \left(\mathcal{R}^N \setminus \mathcal{R}^L\right)$.

(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 $\mathcal{N}(i) \cup \mathcal{N}(j)$ 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 $\mathcal{N}(i)$ and at least one node in $\mathcal{N}(j)$ 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 $\mathcal{N}(i)$ and at least one node in $\mathcal{N}(j)$ are *unsolvable*. Also, since $l \in \Xi$, the nodal residual at nodes i and j are nonzero, which means that $i, j \in \mathcal{V}_N$. 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 (23) and set $\bar{A}_0 = 0$, $\rho = 1$. With some basic algebraic manipulations, one can write

$$\begin{aligned} &2 \sum_{j \in \mathcal{M}'} |\bar{z}^T (\bar{A}_j(\widetilde{\Omega}) - \bar{A}_j(\Omega))\bar{z}| + 2 \sum_{j=1}^{M} |\eta_j| \\ \geq &\|\mathcal{H}^{\Omega}(\bar{v}_* \bar{v}_*^T - \bar{z}\bar{z}^T)\|_1 \geq \|\mathcal{H}^{\Omega}(\bar{v}_* \bar{v}_*^T - \bar{z}\bar{z}^T)\|_2 \end{aligned}$$

Therefore, if t is nonzero

$$\begin{aligned} &t \cdot \|\bar{v}_* \bar{v}_*^T - \bar{z}\bar{z}^T\|_F \\ \leq &2 \sum_{j \in \mathcal{M}'} |\bar{z}^T (\bar{A}_j(\widetilde{\Omega}) - \bar{A}_j(\Omega))\bar{z}| + 2 \sum_{j=1}^{M} |\eta_j| = 2g(\bar{z}, \eta, 1) \\ = &2\|\eta\|_1 \end{aligned}$$

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 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:*

$$|\bar{x}^T \bar{A}_j(\widetilde{\Omega})\bar{x} - \bar{y}\bar{A}_j(\Omega)\bar{y}| > |\bar{x}^T \bar{A}_j(\Omega)\bar{x} - \bar{y}\bar{A}_j(\Omega)\bar{y}| \quad (38)$$

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:

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

where (a) follows from the fact that $\bar{A}_j(\widetilde{\Omega}) = \bar{A}_j(\Omega)$ 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 (38). □

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] 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.

[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] D.-H. Choi and L. Xie, "Impact of power system network topology errors on real-time locational marginal price," *Journal of Modern Power Systems and Clean Energy*, vol. 5, pp. 797–809, 2017.

[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] W. Xu, M. Wang, J.-F. Cai, and A. Tang, "Sparse error correction from nonlinear measurements with applications in bad data detection for power networks," *IEEE Transactions on Signal Processing*, vol. 61, pp. 6175–6187, 2013.

[19] Y. Lin and A. Abur, "Robust state estimation against measurement and network parameter errors," *IEEE Transactions on Power Systems*, vol. 33, pp. 4751–4759, 2018.

[20] R. Zhang, J. Lavaei, and R. Baldick, "Spurious local minima in power system state estimation," *IEEE Transactions on Control of Network Systems*, vol. 6, pp. 1086–1096, 2019.

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

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

[23] F. C. Schweppe and J. Wildes, "Power system static-state estimationn, Part I: Exact model," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, pp. 120–125, 1970.

[24] F. C. Schweppe, "Power system static-state estimationn, Part III: Implementation," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, pp. 130–135, 1970.

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

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

[27] 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.

[28] 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.

[29] I. Molybog, S. Sojoudi, and J. Lavaei, "No spurious solutions in nonconvex matrix sensing: structure compensates for isometry," *American Control Conference*, 2021.

[30] 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.