Topology Error Detection and Robust State Estimation Using Nonlinear Least Absolute Value

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 (small) subgraphs of the network that contain 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 subgraphs. 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 an IEEE benchmark system.**

***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]. Noting that the measurement noise and errors in modeling the system topology can significantly affect the accuracy of state estimation, dealing with bad data and identifying topological errors have received considerable attentions in the past few years.

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 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

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 [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 which are called *critical parameter and measurement pairs*. Another important approach in this category is the least absolute value (LAV) estimator that was 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 were 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, non-convex 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. Due to space restrictions, the proofs are moved to [19].

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. $\mathrm{Re}(\cdot)$, $\mathrm{Im}(\cdot)$, $\mathrm{rank}(\cdot)$, $\mathrm{Tr}(\cdot)$ and $\mathrm{null}(\cdot)$ denote the real part, imaginary part, rank, trace, and null space of a given scalar or matrix. 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 $\mathrm{diag}(x)$ denotes a diagonal matrix whose diagonal entries are given by the vector x, whereas $\mathrm{diag}(X)$ 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 $\mathrm{vec}(\cdot)$ vectorizes its matrix argument.

II. Preliminaries

Consider an electric power network represented by a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, \ldots, n\}$ and $\mathcal{E} := \{1, \ldots, r\}$ 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 $p + q\mathbf{j} = \mathrm{diag}(vv^*Y^*)$. Let $\{e_1, \ldots, e_n\}$ denote the canonical vectors in $\mathbb{R}^n$. Define

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

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

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

The traditional measurable quantities can be expressed as

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

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 $\mathcal{M} = \{1, \ldots, m\}$ specifying m measurements of the form* (4)*, the mapping from the measurement index set to the set of measurement matrices is defined as*

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

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

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 [20], [21]. 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 operators has, 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 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, picking the initial point to be the nominal point $\mathbb{1}$ would likely guarantee 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:*

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

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, is nonzero. $\mathcal{V}_\mathcal{R}$ is the set of nodes that are either: 1) the* from *or* to *nodes of the lines in $\mathcal{E}_\mathcal{R}$, 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 [22]. 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 using these results, we 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 number 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 [23], 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

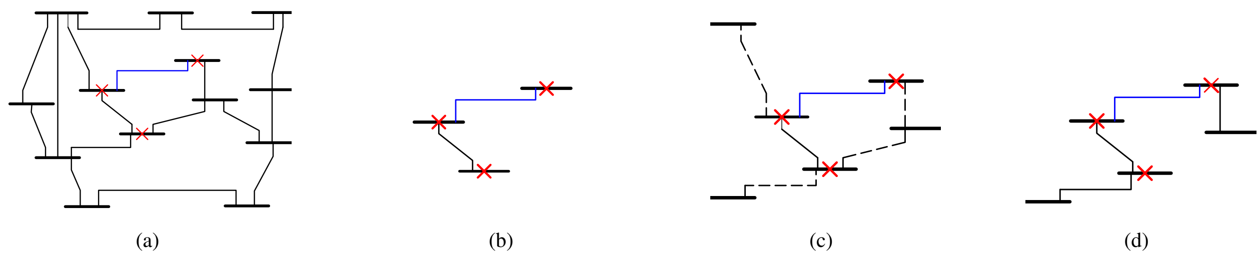

Fig. 1: (a) A power network, (b) the state estimation error graph $\mathcal{S}$, (c) the extended state estmation error graph $\widetilde{\mathcal{S}}$, and (d) the residual graph $\mathcal{R}$. Unsolvable nodes are marked with red crosses, the blue line is the only erroneous line, and the dotted lines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

(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 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 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.

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**
  $\Omega^t \leftarrow \widetilde{\Omega}$.
  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 $\mathcal{D}_L$

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.

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 only when all topological errors have been detected and fixed because the exact retrieval of the desired state is close to impossible if topological errors continue to exist). Therefore, t can be viewed as a quantification of the measurements' quality for finding a unique solution for the over-determined power flow equations.

IV. Simulation Results

In this section we show several simulation results on the IEEE 57-bus system to evaluate the efficacy of the proposed algorithm in estimating the state and detecting the topological errors. To run the simulations, we use MATPOWER data along with the MATLAB *fmincon* function as the local search algorithm.

A. Simulation setups

Consider the scenario where the model of the network (used by the power system operator) is different from the true system due to recent changes in the switching status of some of the lines. In particular, the model is yet to be updated about the fact that several lines recently went off. In order to simulate these topological errors, we randomly pick 2 out of 80 lines of the IEEE 57-bus system and switch them off by changing their resistance and reactance to a large number (e.g. 10^{10}). This ensures that no current flows through these lines. To satisfy the observability condition for the system, the two lines are picked in such a way that they do not share a common bus. We consider this system as the true system, with the associated measurement matrices $\{\bar{M}_j(\Omega)\}_{j\in\mathcal{M}}$ as mentioned in Section III-B. For the model that is accessible to the power system operators, we use the measurement matrices corresponding to the original benchmark system with no change to its line properties.

To generate the true system's state, we assume that the voltage magnitudes are close to unity and the phase angles are small. For data collection, we assume that: (1) All nodal measurements but only 70% of all possible line measurements are available; (2) no measurements are taken from the erroneous lines; (3) the lines with line measurements are selected at random such that the observability condition stated in Definition 4 is satisfied. The reason for considering 70% line measurements is to ensure that there are enough redundant measurements so that the proposed NLAV estimator is able to find the global optimum using a local search algorithm. This property, which is one of our main assumptions in formulating the proposed technique, has been proven for NLS in matrix completion [24] and power system state estimation problems [4], [17] but has not yet been proven for NLAV. However, the simulation results show that by considering such redundant measurements the NLAV estimator converges to the global optimum for the above-mentioned scenario. It should also be noted that to have consistent results for the comparisons, we fix the true system state vector and the lines selected for line measurements throughout the simulations. For numerical reasons, we also normalize each measurement matrix $\bar{M}_j(\Omega)$ with respect to the largest absolute value of its entries in all of the analyses.

B. Performance of the proposed NLAV estimator

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. The results of solving the NLS and NLAV state estimations are shown in Figure 2. Figures 2(a) and 2(b) 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 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 be numerically intractable for large systems. In contrast, the state estimation errors and the residuals after the first run of the NLAV are shown in Figures 2(c) and 2(d). The two only peaks in Figure 2(c) 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 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 clearly observe the superiority of the NLAV estimator in finding the true states in the presence of topology error. In the following subsection, we illustrate the robustness of this method even in the presence of noise.

C. Noisy measurements

Dealing with noisy measurements is an important challenge when it comes to implementing the state estimator in the real-world. In order to test the efficacy of the method in the presence of noise, we consider 1%, 3% and 5% measurement noise and repeat the previous analysis for the scenario with two missing lines (#8 and #67) without any regularization, i.e., setting $M_0 = 0$. The noise is calculated as a fixed percentage (i.e., 1%, 3% or 5%) of the true measurement values. For instance, when noise is 1%, $b = h^{\Omega}(\bar{z}) + \eta = 1.01 h^{\Omega}(\bar{z})$. Figures 3(a) and 3(b), respectively, show the state estimation errors and residuals with 1% and 5% measurement noise. With 1% noise, it is observed that despite increased state estimation errors and residuals (compared to Figures 2(c) and 2(d)), the peaks are still sufficiently sparse and correspond to the correct erroneous lines. This enables the operator to use Algorithm 1 to efficiently locate the topological errors. However, with 5% noise, many spurious peaks appear in the state estimation error

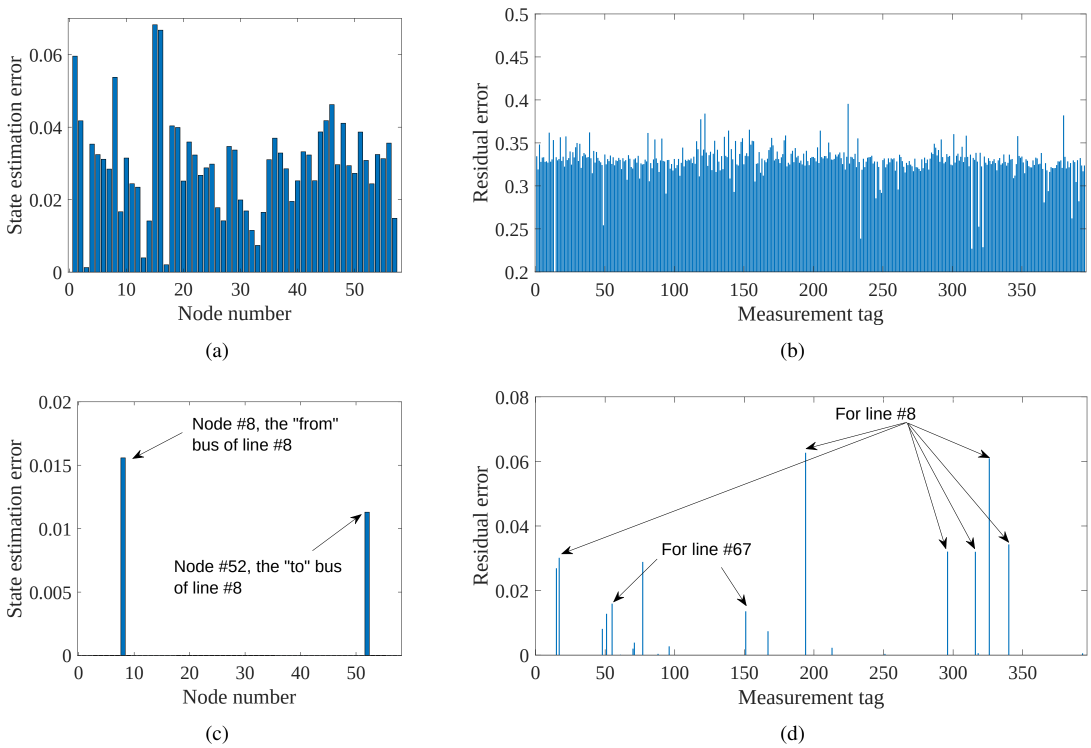

Fig. 2: State estimation error for (a) NLS with noiseless measurements, (c) NLAV method with noiseless measurements; and residuals for (b) NLS with noiseless measurements, (d) NLAV method with noiseless measurements. Note that in (b) and (d), the x-axis shows the measurement tag, which is not the same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

and the residuals. This increases the size of the associated search graph $\widetilde{\mathcal{R}}$ and hence requires a larger set of lines to search through before finding the erroneous lines.

Finally, to analyze the effect of the magnitude of topological error on the proposed method, scenarios with a higher number of erroneous lines (from 2 to 5) are simulated. The state estimation errors for the scenario with 5 misrepresented lines are shown in Figure 4. By comparing Figure 3(a) with Figure 4, it follows that by increasing the number of erroneous lines, the state estimation error vector becomes less sparse which again leads to the enlargement of the associated search graph $\widetilde{\mathcal{R}}$. It is also noteworthy to mention that in this fairly small system, the vector of state estimation errors loses sparsity after taking five lines out, but for larger systems more lines can be erroneous and still be efficiently detected by the proposed method. To support this claim, we have shown in the technical report [19] that up to 25 lines can be removed from the 118-bus system and yet the proposed method works.

V. Conclusion

This paper proposes a new technique to solve the state estimation problem for power systems in the presence of topological errors and to detect such modeling errors. The developed method aims to minimize the 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.

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.

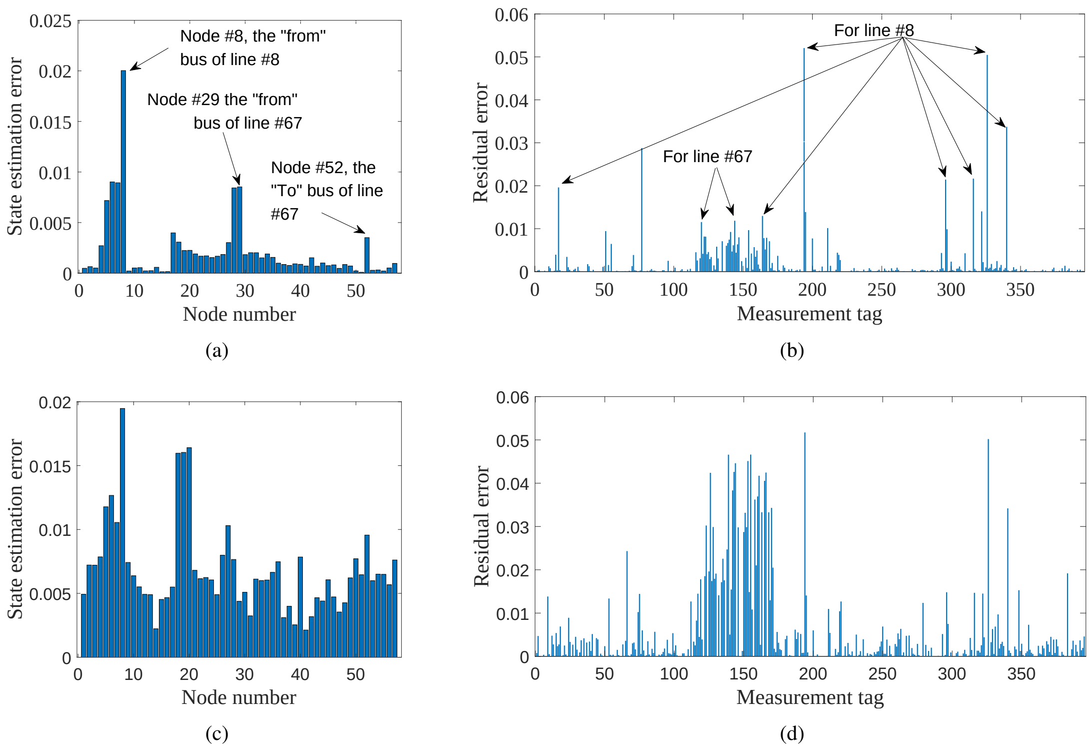

Fig. 3: State estimation error for (a) NLAV with 1% noisy measurements, (c) NLAV method with 5% noisy measurements; and residuals for (b) NLAV with 1% noisy measurements, (d) NLAV method with 5% noisy measurements. Note that in (b) and (d) the x-axis shows the measurement tag, which is not the same as the node or line number due to a random selection of line measurements.

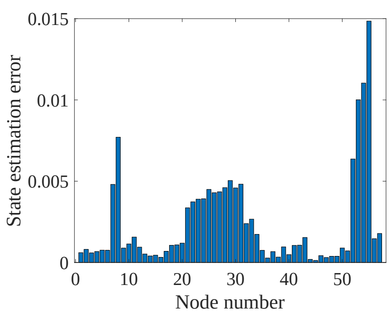

Fig. 4: State estimation error of the proposed method with 5 erroneous lines and 1% noise

[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] S. Park, R. Mohammadi-Ghazi, and J. Lavaei, "Nonlinear least absolute value estimator for power systems state estimation and topological error detection," *Technical report avilable at http://lavaei.ieor.berkeley.edu/SE_TC_1_2018.pdf*, 2018.

[20] F. C. Schweppe and J. Wildes, "Power system static-state estimationn, part i: Exact model," *IEEE T. Power. Ap. Syst.*, pp. 120–125, 1970.

[21] F. C. Schweppe, "Power system static-state estimationn, Part III: Implementation," *IEEE T. Power. Ap. Syst.*, pp. 130–135, 1970.

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

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

[24] R. Ge, J. D. Lee, and T. Ma, "Matrix completion has no spurious local minimum," in *Advances in Neural Information Processing Systems (NIPS)*, 2016.