Empirical analysis of ℓ_1-norm for state estimation in power systems

Reza Mohammadi-Ghazi and Javad Lavaei

***Abstract*—This paper provides an empirical study of the existence of spurious local optima for the nonlinear ℓ_1-norm state estimator for power systems. The linear least-absolute-value (LAV) estimator has attracted a lot of attentions over the past few years, especially in the machine learning community, mainly due to its convexity and robustness with respect to sparse noise. It is known that the linear LAV estimator outperforms the linear least-square estimator that is sensitive to outliers. With this presumption, recent studies have attempted to apply the nonlinear version of the LAV for problems such as topological error detection in power systems. It was shown in those studies that, given reliable prior knowledge about the system and an appropriate solver initialization, the global optima of the nonlinear least absolute value (NLAV) estimator can be found using local search algorithms. This study aims to analyze the success rate of solving such problems without sufficiently accurate prior knowledge about the system. In doing so, we perform empirical analyses on IEEE benchmark systems and show that the nonlinear NLAV estimator may have spurious local solutions even when all possible error-free measurements are used. An important observation about the spurious local solutions is that the buses with inaccurate estimated states form a localized and connected sub-network. This observation is justified using a decomposition technique.**

I. INTRODUCTION

Supplying economical energy is the lifeblood of modern civilization, and is highly related to the reliability and efficiency of energy infrastructures, such as power systems which provide clean and convenient energy for industrial and individual uses. An important challenge in operating these infrastructures is to protect them against progressive failures of stressed components that may lead to blackouts [1], [2]. To prevent such events, 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 bad data, sparse noise, and topological errors significantly affect the accuracy of state estimation, addressing these problems has received considerable attentions in the past few years.

A. State estimation for power systems

Given an N-bus power system and its associated vector of voltage phasors, denoted by $\mathbf{v} \in \mathbb{C}^N$, the state estimation problem is to recover $\mathbf{v}$ using a set of SCADA measurements $\{b_1, \cdots, b_m\}$, where m is the number of measurements. The i-th measurement, denoted by b_i, is

$$b_i = f_i(\mathbf{v}) + \epsilon_i \tag{1}$$

where $f_i(\cdot)$ is the nonlinear measurement function, which will be explained with more details in the preceding sections, and ϵ_i is an unknown measurement noise/error.

Given the above settings, recovering the true state of the system usually boils down to solving some optimization problem. In what follows, we will briefly review the main state estimation techniques.

B. Nonlinear least-square state estimator

One of the most common approaches for solving the aforementioned state estimation problem is to employ the nonlinear least-squares (NLS) technique proposed in [5], [6]. This method assume that the noise values ϵ_i are independent and identically distributed samples of a Gaussian distribution with zero mean and variances $1/w_1, \ldots, 1/w_m$. Under this assumption, the maximum likelihood estimation of $\mathbf{v}$ becomes equal to the solution to the least squares problem

$$\underset{u \in \mathbb{C}^N}{\text{minimize}} \ \frac{1}{2} \sum_{i=1}^{m} w_i [f_i(\mathbf{u}) - b_i]^2. \tag{2}$$

Starting from an arbitrary initial point $\mathbf{u}_0 \in \mathbb{C}^N$, one may use the Gauss-Newton method to find a stationary point of the above problem. Convergence to a stationary point is guaranteed under mild conditions [6] and convergence to the true state is expected if m is sufficiently large [4].

C. Nonlinear least-absolute-value (NLAV) state estimator

Due to its differentiability, the NLS state estimator has appealing features from both theoretical and computational perspectives. However, this estimator suffers from a lack of robustness to outliers and sparse noise/errors. Sparse errors could be due to sensor malfunctioning or adversarial attack. It could also cause by the topological errors in the model used by power operators (e.g., the operator may assume that a line switch is ON while it has been turned off recently) [7], [8], [9], [10], [11]. In this scenario, there are a set of functions $\tilde{f}_1(\cdot), \ldots, \tilde{f}_m(\cdot)$ modeling the system correctly, which are different from the functions $f_1(\cdot), \ldots, f_m(\cdot)$ used by the operator. The error ϵ_i is then equal to $\tilde{f}_i(\mathbf{v}) - f_i(\mathbf{v})$. Since topological errors often occur for a small part of the network, the vector $[b_1 \ b_2 \ \cdots \ b_m]$ is sparse. Although only a small number of measurement equations are wrong in this case, the solution of the NLS estimator (2) would be far away from the true state. The lack of robustness of the NLS estimator with respect to outliers have been well studied in statistical learning [12], and an analysis of this phenomenon in the context of power systems is reported in [11].

One approach for dealing with sparse errors is the deployment of synchronized point-on-wave measurements provided

This work was supported by the ONR grant N00014-17-1-2933, DARPA grant D16AP00002, AFOSR grant FA9550- 17-1-0163, NSF grant 1807260, and ARO grant W911NF-17-1-0555. The authors are with the Department of Industrial Engineering and Operation Research, University of California, Berkeley, CA 94710. Emails: {mohammadi, lavaei}@berkeley.edu

by phasor measurement units (PMUs). These measurements make the state estimation problem linear, and then it is possible to use a least-absolute-value (LAV) estimator that is known to provide a robust estimation in presence of sparsely occurring errors of arbitrary magnitudes [9], [13], [10]. However, this framework is inapplicable to the highly non-convex state estimation problem based on traditional nonlinear power flow measurements [11]. Inspired by the success of the LAV estimator, consider the nonlinear least-absolute-value (NLAV) estimator that can handle nonlinear power flow equations:

$$\underset{\mathbf{u}\in\mathbb{C}^N}{\text{minimize}} \sum_{i=1}^{m} |f_i(\mathbf{u}) - b_i|. \tag{3}$$

By the virtue of the notion of *global functions* proposed in [14] and by considering $\mathbf{u}_0$ to be the flat start vector $[1\ 1\cdots 1]^T$, it is shown in [11] that the above nonlinear estimator (3) is robust to modest topological errors and can recover the underlying state of the system correctly.

D. Spurious local solutions

A point is called a *spurious* solution of a nonlinear optimization problem if it satisfies the first-order and second-order necessary optimality conditions but is not a global solution. Numerical algorithms cannot often distinguish between spurious and correct solutions. Even in the absence of measurement noise and/or topological errors, it is still possible to become stuck in spurious local solutions of (2) and (3), due to the non-convexity of the objective function rooted in the nonlinearity of the power flow equations. So, there are two challenges: (i) could these estimators find a high-quality estimate of the true state of the system? (ii) is it easy to solve these estimation problems to global optimality? To address Question (i), there is a plethora of research in the literature suggesting that NLAV is significantly more robust than NLS if there are sparse errors of gross sizes. Question (ii) is related to the existence of spurious solutions and is far less explored since it is known that the problem is NP hard in the worst case [4], [15].

Recent studies have provided theoretical guarantees for convergence of the NLS estimator to the global solution in matrix completion and power system state estimation problems [16], [4], [15]. The main message of these studies is that all spurious solutions disappear if the amount of measured data is sufficiently large, in which case local search algorithms can find the global solution without ant prior knowledge about the solution. These studies have also analyzed the required number of measurements. In contrast, the theoretical analysis of the NLAV estimator is much more challenging due to its non-differentiability, and far less is known about the spurious solutions of this estimator. The existing theoretical guarantees for the NLAV estimator are limited to very special cases of the measurement functions.

E. Main results of this study

The common belief is that, due to its robustness with respect to outliers, the NLAV estimator always outperforms the NLS estimator. This has been empirically shown for problems such as the topological error detection in power systems [11]. However, the related results highly rely on the availability of some prior knowledge about the system, e.g., closeness of the initial guess to the true solution, which may not be always possible in practice. This study aims to analyze the performance of the NLAV estimator and the likelihood of finding the global solution in the absence of prior knowledge. In doing so, we follow an empirical approach based on random initialization along with using local search algorithms, which is a method for analyzing the existence of spurious solutions in optimization problems. After presenting the preliminaries, we review the related works on the analysis of spurious local optima for the NLAV estimator in Section III. Our findings are presented in Section IV followed by the interpretation of the results in Section V. In particular, we show that the NLAV estimator has spurious solutions even when all possible error-free measurements are deployed, and also observe that the set of buses with incorrect estimates form a connected sub-network. Concluding remarks are provided in Section VI.

F. Notations

The notations $\mathbb{R}^N$ and $\mathbb{C}^N$ denote the spaces of N-dimensional real and complex vectors, respectively. The symbols $(\cdot)^T$ and $(\cdot)^*$ denote the transpose and conjugate transpose of a vector/matrix. The imaginary unit is denoted by $\mathbf{i} = \sqrt{-1}$, and Re$(\cdot)$ and Im$(\cdot)$ denote the real part and imaginary part of a given scalar or matrix. The notations $\|\mathbf{x}\|_1$ and $\|\mathbf{x}\|_2$ denote the ℓ_1-norm and ℓ_2-norm of vector $\mathbf{x}$, respectively. The relation $\mathbf{X} \succeq 0$ means that the matrix $\mathbf{X}$ is Hermitian positive semidefinite. The (i, j) entry of $\mathbf{X}$ is denoted by $\mathbf{X}_{i,j}$. The notation $\mathbf{X}[\mathcal{S}_1, \mathcal{S}_2]$ denotes the submatrix of $\mathbf{X}$ whose rows and columns are chosen from the given index sets $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively.

II. Preliminaries

A. SCADA measurements

Consider an electric power network represented by a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, \ldots, N\}$ and $\mathcal{E} := \{1, \ldots, L\}$ denote the sets of buses and branches, respectively. Let $v_k \in \mathbb{C}$ denote the nodal complex voltage at bus $k \in \mathcal{V}$, whose magnitude and phase are given as $|v_k|$ and $\measuredangle z_k$. Then, the current flowing from bus k to bus l is

$$c_{k\to l} = Y_{k,l}(v_k - v_l) = [Y_{k,l}(\mathbf{e}_k - \mathbf{e}_l)]^T\mathbf{v}, \tag{4}$$

where $Y_{k,l} \in \mathbb{C}$ is the directional admittance of the line/transformer connecting the two buses; $\{\mathbf{e}_1, \ldots, \mathbf{e}_k, \ldots, \mathbf{e}_N\}$ are the canonical basis of $\mathbb{R}^N$; and $\mathbf{v} = [v_1, \ldots, v_N]^T$. The net current injection at bus k is

$$c_k = \left(Y_k\mathbf{e}_k + \sum_{l\in\mathcal{N}_k} Y_{k,l}(\mathbf{e}_k - \mathbf{e}_l)\right)^T \mathbf{v}, \tag{5}$$

where Y_k is the shunt admittance at bus k, and $\mathcal{N}_k$ is the set of neighboring buses of bus k. Given that complex power is the product of voltage and conjugate current, and that current is a linear function of voltages, complex power can be written as a quadratic function of $\mathbf{v}$. For instance, the complex power that is sent from bus k to bus l is

$$p_{k\to l} + q_{k\to l}\mathbf{i} = (\mathbf{v}^*\mathbf{P}_{k\to l}\mathbf{v}) + (\mathbf{v}^*\mathbf{Q}_{k\to l}\mathbf{v})\mathbf{i} \tag{6}$$

where $\mathbf{P}_{k\to l} = \frac{1}{2}(\mathbf{S}_{k\to l} + \mathbf{S}^*_{k\to l})$ and $\mathbf{Q}_{k\to l} = \frac{1}{2\mathbf{i}}(\mathbf{S}_{k\to l} - \mathbf{S}^*_{k\to l})$ are the Hermitian splitting for

$$\mathbf{S}_{k\to l} = Y^*_{k,l}(\mathbf{e}_k - \mathbf{e}_l)\mathbf{e}_l{}^T.$$

Therefore, the net complex power that is injected into or consumed at bus k is

$$p_k + q_k\mathbf{i} = c^*_k v_k = (\mathbf{v}^*\mathbf{P}_k\mathbf{v}) + (\mathbf{v}^*\mathbf{Q}_k\mathbf{v})\mathbf{i} \tag{7}$$

where p_k and q_k are the active and reactive injected/consumed powers, respectively; $\mathbf{P}_k = \frac{1}{2}(\mathbf{S}_k + \mathbf{S}^*_k)$ and $\mathbf{Q}_k = \frac{1}{2\mathbf{i}}(\mathbf{S}_k - \mathbf{S}^*_k)$ are the Hermitian splitting for

$$\mathbf{S}_k = Y^*_k\mathbf{e}_k\mathbf{e}_k{}^T + \sum_{l\in\mathcal{N}_k} Y^*_{k,l}(\mathbf{e}_k - \mathbf{e}_l)\mathbf{e}_k{}^T.$$

Based on the above equations, the power values at each bus are based on two quadratic functions $\mathbf{v}^*\mathbf{P}_k\mathbf{v}$ and $\mathbf{v}^*\mathbf{Q}_k\mathbf{v}$. Likewise, active and reactive powers over each line and the voltage magnitude at each bus are quadratic in $\mathbf{v}$. The nonlinearity of these functions is the reason for the non-convexity of the state estimation problems (2) and (3).

B. *Real-valued parameterization*

As shown in the previous section, active and reactive powers are naturally expressed in the complex domain. However, it is desirable to derive real-valued expressions for the power functions, as needed by local search algorithms. To this end, define $\mathcal{V}_r = \mathcal{V}\backslash\{\text{slack bus}\}$, i.e., the set of all buses except the slack bus. Consider the following real-valued symmetrization operator for complex-valued voltage vectors and measurement matrices:

$$\text{RS}(\mathbf{X}) = \begin{bmatrix} \text{Re}\{X[\mathcal{V},\mathcal{V}]\} & -\text{Im}\{X[\mathcal{V},\mathcal{V}_r]\} \\ \text{Im}\{X[\mathcal{V}_r,\mathcal{V}]\} & \text{Re}\{X[\mathcal{V}_r,\mathcal{V}_r]\} \end{bmatrix}, \tag{8a}$$

$$\text{RS}(\mathbf{x}) = \begin{bmatrix} \text{Re}\{\mathbf{x}[\mathcal{V}]\}^T & \text{Im}\{\mathbf{x}[\mathcal{V}_r]\}^T \end{bmatrix}^T. \tag{8b}$$

Basically, this operator maps an $N \times N$ Hermitian matrix into a $(2N-1)\times(2N-1)$ real-valued symmetric matrix, and a vector of length N into a real-valued vector of length $2N-1$ (by setting the phase at the slack bus to zero). Using this operator, define

$$\mathbf{M}_i \triangleq \text{RS}(\widetilde{\mathbf{M}}_i), \quad v \triangleq \text{RS}(\mathbf{v}), \quad x \triangleq \text{RS}(\mathbf{u}) \tag{9}$$

where $i \in \{1,\ldots,m\}$ indexes the SCADA measurements, and $\widetilde{\mathbf{M}}_i$ can be $\mathbf{P}_k$ or $\mathbf{Q}_k$ for a nodal measurement at bus k, $\mathbf{P}_{k\to l}$ or $\mathbf{Q}_{k\to l}$ for a line measurement, or $\mathbf{e}_k\mathbf{e}_k^T$ for a voltage magnitude measurement.

Using the definitions in (9) along with the SCADA measurement equations, the measurement functions can be written as

$$f_i(\mathbf{u}) = \mathbf{x}^T\mathbf{M}_i\mathbf{x} \tag{10}$$

and, hence, the estimation problems (2) and (3) can be cast as

$$\underset{\mathbf{x}\in\mathbb{R}^{2N-1}}{\text{minimize}} \sum_{i=1}^{m} \|\mathbf{x}^T\mathbf{M}_i\mathbf{x} - b_i\|_d^d \tag{11}$$

where d is either 1 or 2.

III. Related studies on spurious local solutions of the NLAV estimator

To numerically solve the non-smooth ℓ_1-norm minimization, we use the famous technique discussed in [17] for convex optimization to reformulate the problem (11) with $d = 1$ as a smooth non-convex quadratically-constrained quadratic program (QCQP). The reformulated problem is given as

$$\begin{aligned} \underset{z\in\mathbb{R}^m}{\text{minimize}} \quad & \sum_{i=1}^{m} z_i \\ \text{subject to} \quad & \mathbf{x}^T\mathbf{M}_i\mathbf{x} - b_i \le z_i, \qquad i = 1,\ldots,m \\ & \mathbf{x}^T\mathbf{M}_i\mathbf{x} - b_i \ge -z_i, \qquad i = 1,\ldots,m \end{aligned} \tag{12}$$

Nonlinear problems of the form (11) and (12) can be efficiently solved via different numerical algorithms, and some recent works have provided theoretical guarantees for the convergence of such methods to a stationary point [18]. However, the convergence analysis of the general form of these problems to their global minima has not been addressed in the literature. In a recent study, the work [14] has considered a very special form of the non-smooth problem (11) with $d = 1$ and analyzed the spurious solutions of the problem by introducing the notion of *global functions*. These are basically the functions that may have multiple local optima, but all of them are global. A simple, yet nontrivial, global function is $|x^2 - 1|$, as shown in Figure 1. This function has two local minima at $x = \pm 1$, both of which are global. No matter how it is initialized, a local search algorithm can find a global solution efficiently. In their study, they proved that if m is large and and each $\mathbf{M}_i$ is in the form of $\mathbf{e}_k\mathbf{e}_l^T$ for some l and k, then the objective function of the estimator (11) with $d = 1$ is a global function. This implies that finding a global minimizer of such function is guaranteed by the use of a local search algorithm. The objective function in (11) for a power system is related, but more complicated, than the one theoretically analyzed in [14]. In what follows, we will empirically analyze this function and its spurious local minima.

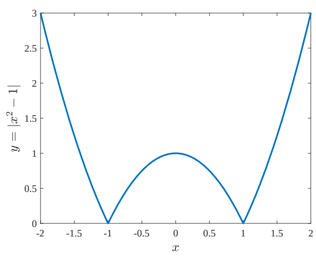


Fig. 1: Simple example of a global function $|x^2 - 1|$

IV. Empirical analysis of the NLAV estimator

To evaluate the performance of the NLAV estimator in the absence of reliable prior information, we pursue an empirical approach based on random initialization and analyze the effect of the number of measurements on the possibility of the existence of spurious local minima. In doing so, we first consider different levels of redundancy for the SCADA measurements of an N-bus system, namely $m \in [2N-1, 4L+3N-1]$. Note that the minimum number of measurements in this analysis is

$2N-1$ corresponding to all nodal power measurements except for the reactive power at the slack bus. For $m > 2N-1$, we randomly select $m-2N+1$ voltage magnitude and line flow measurements and add them to the base set of measurements with $2N-1$ elements. The true state of the system is produced by randomly selecting the voltage magnitudes and angles from the intervals $[0.9, 1.1]$ and $[-30°, 30°]$, respectively. To initialize the numerical algorithm, we generate a random point and scale it properly so that it is within a certain distance from the true solution, and vary this distance to construct different scenarios. More precisely, we consider the discrete distance set $\|\mathbf{v}_0 - \mathbf{v}\|_\infty \in \{0.1, 0.2, \ldots, 1\}$ for each m. Moreover, we run 50 simulations for each pair of m and $\|\mathbf{v}_0 - \mathbf{v}\|_\infty$. As discussed before, it is already known that the ℓ_1 estimator outperforms the ℓ_2 estimator in the linear case in presence of sparse errors. Since the goal of this paper is analyze the spurious solutions of these estimators, we assume that there is no measurement error/noise, i.e., $\epsilon_i = 0$ (note that spurious solutions are due to the nonlinearity of the problem rather than the measurement errors). We perform the experiments on the IEEE 39- and 57-bus benchmark systems, using the *fmincon* command of MATLAB with the interior point algorithm.

A. Simulation results

Figure 2 shows the effect of redundant measurements on existence of spurious solutions for the ℓ_1 and ℓ_2-norm state estimators for the 39- and 57-bus systems. The average rate of success for recovering the true solution in 50 simulations is color coded in these plots for each pair of m and $\|\mathbf{v}_0 - \mathbf{v}\|_\infty$. The y-axis of these plots shows the ratio m/n where $n = 2N-1$. Based on these figures, the NLS estimator empirically has no spurious solutions if $m/n \geq 3.3$ (respectively, 3.5) for the 39-bus (respectively, 57-bus) system. If the number of measurements exceeds this threshold, the NLS estimator would perfectly recover the underlying states no matter how far the initial guess is from the true solution within the described ranges. However, such a perfect recovery cannot be obtained for the NLAV estimator in some instances even if all possible measurements are considered. This confirms the existence of some spurious solutions for the ℓ_1-norm state estimator that are persistent and do not disappear.

To better investigate the nature of spurious local solutions, we select four cases from the regions in Figures 2(a) and 2(c) where the NLS estimator gives perfect recovery while the NLAV fails in some simulations. These cases are visualized in Figures 3 and 4. Note that there are more than one spurious local solution for some combinations of m and $\|\mathbf{v}_0 - \mathbf{v}\|_\infty$ (namely each block in Figure 2), but we visualize only one such solution for brevity. In these figures, the red colored buses are those where the state estimation error, for either the voltage magnitude or angle, is more than 10^{-3}. It follows from these plots that the buses with inaccurate state recovery are localized and form a connected cluster. In special cases, this cluster contains all buses, as shown in Figures 3(a) and 4(a). We will elaborate on this observation below.

V. Interpretation of the Simulation Results

In this section, we provide insights into the existence of spurious local solutions for the NLAV estimator based on the formulation in (12). Then, we discuss the locality and connectivity of the buses where the solution is inaccurate.

A. Spurious local solutions of the NLAV estimator

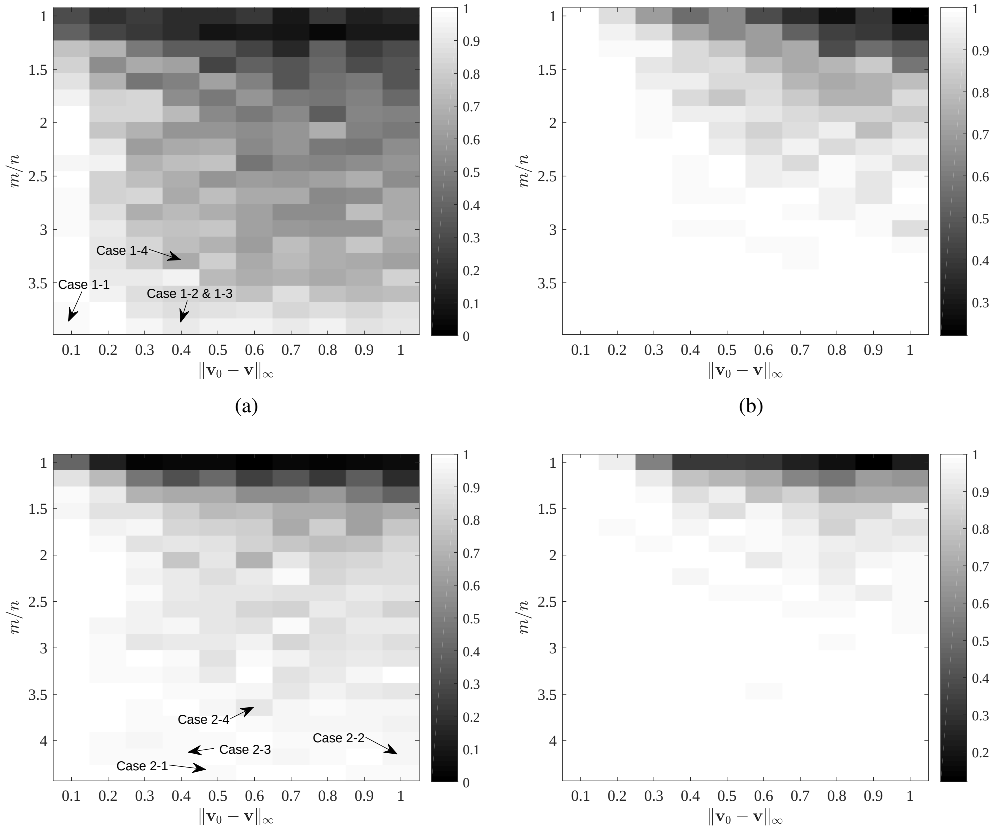
Consider the QCQP problem (12) and equivalently rewrite it as

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^m}{\text{minimize}} \quad & \sum_{i=1}^{m} |z_i| \\ \text{subject to} \quad & \mathbf{x}^T \mathbf{M}_i \mathbf{x} - b_i = z_i, \qquad i = 1, ..., m \end{aligned} \tag{13}$$

Note that z_i in the above formulation aims to estimate ϵ_i in (1). Basically, this problem imposes the hard constrains of (11) via a set of soft constraints. This reformulation is necessary because when the problem is overdetermined due to the volume of the measured data, it becomes infeasible as soon as any measurement has some nonzero error. If $m/n \to 1$, then the state estimation problem approaches the power flow problem and it could have an exponential number of solutions. All these solutions, except for the true state, act as spurious solutions of the state estimation problem. This explains the poor success rates of the first few rows of the plots in Figure 2 corresponding to the case where m/n is close to one, for both NLS and NLAV estimators. On the other hand, in the error-free scenario where ε_i are all zero but m/n is not close to one, the global solution of the problem (13) corresponds to $\mathbf{z} = 0$. However, since the objective function of the NLAV estimator is nonconvex, the above analyses show that the problem has some non-global local solutions. This is not the case for the NLS estimator in the simulations when m/n is large, and the reason for this discrepancy could be traced back to the differentiability of the NLS estimator and the non-smoothness of the NLAV estimator.

B. Locality and connectivity of the buses with inaccurate state estimation

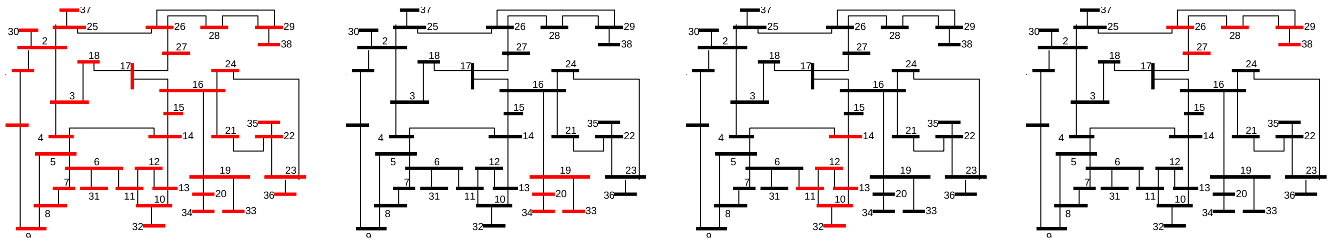
Consider a power network that is divided into (at least) three sub-networks as schematically shown in Figure 5. We denote these subgraphs as $\mathcal{G}_1$, $\mathcal{G}_2$ and $\mathcal{G}_3$, and their associated state variables as $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$, respectively. Assume that $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \mathbf{x}_3^T]^T$ (this may require relabeling the buses). Denote the objective function of (11) as $h(\mathbf{x})$. By noting that there is a cut that separates $\mathcal{G}_1 \cup \mathcal{G}_2$ from $\mathcal{G}_3$, we can decompose the objective function $h(\mathbf{x})$ as $h_1(\mathbf{x}_1, \mathbf{x}_2) + h_2(\mathbf{x}_2, \mathbf{x}_3)$, where $h_1(\cdot)$ is the summation of those components of $h(\cdot)$ that involve $\mathbf{x}_1$ and $\mathbf{x}_2$, and $h_2(\cdot)$ is similarly defined for $\mathbf{x}_2$ and $\mathbf{x}_3$. This decomposition happens in power systems because of the locality of the equations for complex powers and voltage magnitudes. Assume that $\mathbf{x}_2$ is fixed to its optimal solution $\mathbf{x}_2^{\text{opt}}$. Due to the above-mentioned decomposition of the objective function, the existence of a spurious local solution for $h_1(\mathbf{x}_1, \mathbf{x}_2^{\text{opt}})$, which depends only on $\mathbf{x}_1$, imposes a spurious solution for $h(\mathbf{x})$ (note that the minimum values of $h(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ and $h(\mathbf{x}_1, \mathbf{x}_2^{\text{opt}}, \mathbf{x}_3)$ are the same). In other words, if there is a local solution for a small part of the network (acting as a stand-alone system) under some settings, then it results into a spurious solution for the estimation of the state of the entire network. It is noteworthy to mention that since the decomposition of $h(\cdot)$ is not unique, various local solutions may exist.



(a) (b)

(c) (d)

Fig. 2: (a) Outcome of the ℓ_1 estimator for the 39-bus system, (b) outcome of the ℓ_2 estimator for the 39-bus system, (c) outcome of the ℓ_1 estimator for the 57-bus system, (d) outcome of the ℓ_2 estimator for the 57-bus system.

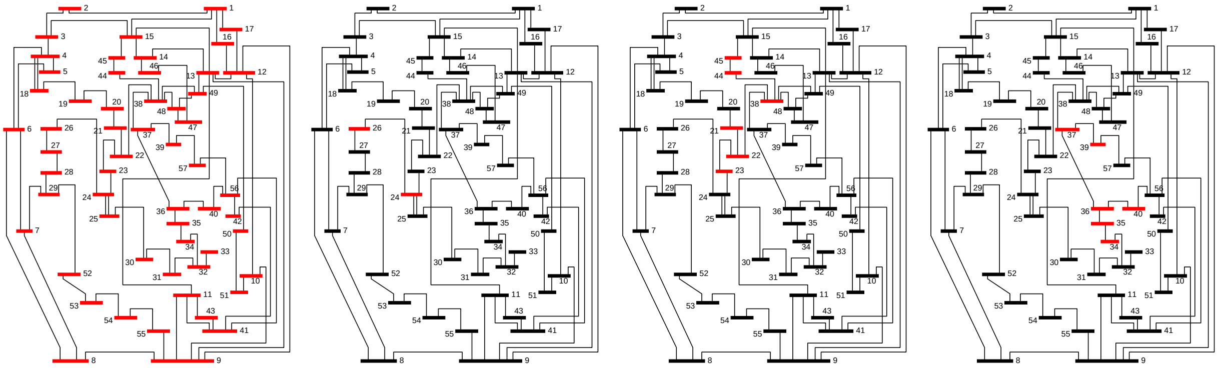


(a) (b) (c) (d)

Fig. 3: These pictures analyze the voltage mismatches for the cases shown in Figure 2(a). The red colored buses are those where the underlying states were not accurately estimated. (a) Case 1-1, (b) Case 1-2, (c) Case 1-3, (d) Case 1-4.

The numerical results of Section IV confirm the above argument regarding the connectivity and locality of the buses with wrong state recovery. Furthermore, they support the statement that multiple spurious local solutions could exist for the same combination of m and $\|\mathbf{v}_0 - \mathbf{v}\|_\infty$. For example, Figures 3(b) and 3(c) show two different spurious solutions for the same such combination. This implies that the aforementioned decomposition of the objective function is not unique.

In addition to the localized spurious local solutions, there are some rare cases such as the ones shown in Figures 3(a) and 4(a) where the underlying state recovery is inaccurate for all buses. To calculate the odds of the occurrence of such cases, we ran 2000 simulations, in addition to the previous 50 simulations, on the 39-bus system based on the full measurement set for $\|\mathbf{v}_0 - \mathbf{v}\|_\infty = 0.1$, which corresponds to the block for the case 1-1. There was only one case where all states were estimated incorrectly. Similarly, we ran 1000 simulations on the 57-bus system with the full measurement set and $\|\mathbf{v}_0 - \mathbf{v}\|_\infty = 0.3$, which corresponds to the block for the case 2-1. There were only two other scenarios where the state recovery was inaccurate for all buses. Although the interpretation of such cases is not easy due to the NP-hardness of the problem, the simulation results show that the likelihood of these events is extremely low.



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

Fig. 4: These pictures analyze the voltage mismatches for the cases shown in Figure 2(c). The red colored buses are those where the underlying states were not accurately estimated. (a) Case 2-1, (b) Case 2-2, (c) Case 2-3, (d) Case 2-4.

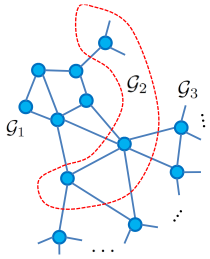


Fig. 5: A schematic power network that is divided into three sub-networks $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_3$.

VI. Conclusion

An empirical analysis of the ℓ_1 and ℓ_2-norm state estimators is performed for power systems in this work. It is known that the NLAV estimator outperforms the NLS estimator for linear systems in presence of sparse measurement errors of arbitrary magnitudes. However, the properties of the NLAV estimator in the nonlinear case are not known due to the high nonlinearity of the problem and its non-smoothness. In this work, we empirically study the existence of spurious solutions for the NLAV estimator and compare it with the NLS estimator. It is shown that no matter how many measurements are used, the state estimation problem may have a non-global local solution that corresponds to an incorrect state even when all measurements are error free. It is observed that the set of buses with incorrect state estimates form a small connected sub-network. This observation is theoretically justified via a decomposition technique and implies that a small part of the network could be the culprit for the failure of the NLAV state estimation problem. The findings of this paper suggest that a weighted estimator obtained by combining the ℓ_1 and ℓ_2 norms would offer a more balanced performance with respect to handling spurious solutions and dealing with sparse measurement errors/noise.

References

[1] K. Clements and P. Davis, "Detection and identification of topology errors in electric power systems," *IEEE Transactions on Power Systems*, vol. 3, no. 4, pp. 1748–1753, 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, Vancouver, BC, Canada, May 2004*.

[3] A. Monticelli, "Electric power system state estimation," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 262–282, 2000.

[4] R. Zhang, J. Lavaei, and R. Baldick, "Spurious critical points in power system state estimation," in *51st Hawaii International Conference on System Sciences*, 2018.

[5] F. C. Schweppe and J. Wildes, "Power system static-state estimation, Part I: Exact model," *IEEE Transactions on Power Apparatus and Systems*, no. 1, pp. 120–125, 1970.

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

[7] K. A. Clements and A. S. Costa, "Topology error identification using normalized lagrange multipliers," *IEEE Transactions on power systems*, vol. 13, no. 2, pp. 347–353, 1998.

[8] 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, no. 16, pp. 2760–2767, 2015.

[9] M. Korkali and A. Abur, "Robust fault location using least-absolute-value estimator," *IEEE Transactions on power systems*, vol. 28, no. 4, pp. 4384–4392, 2013.

[10] Y. Lin and A. Abur, "Robust state estimation against measurement and network parameter errors," *IEEE Transactions on power systems*, vol. 33, no. 5, pp. 4751–4759, 2018.

[11] S. Park, R. Mohammadi-Ghazi, and J. Lavaei, "Nonlinear least absolute value estimator for power systems state estimation and topological error detection," *(preprint)*, 2018. [Online]. Available: http://www.ieor.berkeley.edu/~lavaei/SE_TC_1_2018.pdf

[12] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer Series in Statistics, 2009.

[13] B. Donmez and A. Abur, "Sparse estimation based external system line outage detection," in *Power Systems Computation Conference (PSCC), Genoa, Italy, June 2016*.

[14] C. Josz, Y. Ouyang, R. Y. Zhang, J. Lavaei, and S. Sojoudi, "A theory on the absence of spurious optimality," in *Annual Conference on Neural Information Processing Systems (NIPS)*, 2018. [Online]. Available: https://arxiv.org/pdf/1805.08204.pdf

[15] R. Zhang, J. Lavaei, and R. Baldick, "Spurious local minima in power system state estimation," *preprint*, 2018. [Online]. Available: http://www.ieor.berkeley.edu/~lavaei/SE_2018_1.pdf

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

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.

[18] K. Khamaru and M. J. Wainwright. (2018) Convergence guarantees for a class of non-convex and non-smooth optimization problems. [Online]. Available: https://arxiv.org/pdf/1804.09629.pdf