Sensitivity-based Vulnerability Assessment of State Estimation

Gonzalo E. Constante-Flores, *Student Member*, *IEEE*, Antonio J. Conejo, *Fellow*, *IEEE*, and Jiankang Wang, *Member*, *IEEE*

***Abstract*—We propose a technique to assess the vulnerability of the power system state estimation. We aim at identifying the measurements that have a high potential of being the target of false data injection attacks. From the perspective of the adversary, such measurements have the following characteristics: ① being influential on the variable estimates; ② corrupting their measured values is likely to be undetected. Additionally, such characteristics should not change significantly with the system operation condition. The proposed technique provides a systematic way of identifying the measurements with such characteristics. We illustrate our methodology on a 4-bus system, the New England 39-bus system, and the IEEE 118-bus test system, respectively.**

***Index Terms*—False data injection attack, state estimation, sensitivity analysis, singular value decomposition.**

NOMENCLATURE

A. Sets

$\mathcal{N}$	Set of buses
$\mathcal{N}_i$	Set of buses connected to bus i
$\mathcal{V}$	Set of buses with voltage magnitude measurement
$\mathcal{P},\mathcal{Q}$	Set of buses with active and reactive power measurements
$\mathcal{P}_f$, $\mathcal{Q}_f$	Set of branches with active and reactive power flow measurements
$\mathcal{Z}$	Set of buses with zero injection

B. Parameters

$\boldsymbol{a}$	Vector of parameters, $\boldsymbol{a}\in\mathbb{R}^q$, $\boldsymbol{a}=[w_i^V, w_i^P, w_i^Q, w_{i,j}^P, w_{i,j}^Q, G_{i,j}, B_{i,j}, b_{i,j}^{\text{sh}}]$
$b_{i,j}^{\text{sh}}$	Shunt susceptance of line ij
$G_{i,j}, B_{i,j}$	Real and imaginary parts of entry of admittance matrix of line ij
P_i^m, Q_i^m	Measurements of active and reactive power injection at bus i
$P_{i,j}^m, Q_{i,j}^m$	Measurements of active and reactive power flow measurement on line ij
v_i^m, θ_i^m	Voltage magnitude and angle measurement at bus i
w_i^x	Weighting factor for a measurement at bus i, where superindex $x=V,P,Q$ refers to voltage, active power, and reactive power, respectively
$w_{i,j}^x$	Weighting factor for a measurement on line ij, where superindex $x=P,Q$ refers to active power and reactive power flows, respectively
$\boldsymbol{z}$	Vector of measurements, $\boldsymbol{z}\in\mathbb{R}^p$, $\boldsymbol{z}=[v_i^m, \theta_i^m, P_i^m, Q_i^m, P_{i,j}^m, Q_{i,j}^m]$

C. Variables

v_i, θ_i	Voltage magnitude and angle at bus i
P_i, Q_i	Active and reactive power injections at bus i
$P_{i,j}, Q_{i,j}$	Active and reactive power flows of line ij
$\boldsymbol{x}$	Vector of optimization variables, $\boldsymbol{x}\in\mathbb{R}^n$, $\boldsymbol{x}=[v_i, \theta_i, P_i, Q_i, P_{i,j}, Q_{i,j}]$

D. Dual Variable

$\boldsymbol{\lambda}$	Lagrange multiplier vector, $\boldsymbol{\lambda}\in\mathbb{R}^r$

E. Constants

n	Number of optimization variables
p	Number of measurements
q	Number of parameters
r	Number of equality constraints
$\mathbf{1}_t$	Vector of t-dimensional all-ones column

F. Functions

$\boldsymbol{c}(\cdot)$	Equality constraints representing pseudo-measurements, power flows, and power injections
$J(\cdot)$	Measurement error function

I. INTRODUCTION

ONE of the key functions of energy management systems (EMSs) is state estimation, which aims at finding

Manuscript received: September 4, 2020; accepted: January 11, 2021. Date of CrossCheck: January 11, 2021. Date of online publication: February 17, 2021.

This work was supported in part by the National Science Foundation (No. EPCN 1808169, No. ECCS 1711048).

G. E. Constante-Flores and J. Wang are with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, USA (e-mail: constanteflores.1@osu.edu; wang.6536@osu.edu).

A. J. Conejo (corresponding author) is with the Department of Integrated Systems Engineering and the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, USA (e-mail: conejo.1@osu.edu).

DOI: 10.35833/MPCE.2020.000658

the most likely estimate of the system state (i.e., voltage phasors) given the network topology and parameters, and a set of real-time measurements from telemetry and meters [1]. Although such measurements typically contain small errors due to the accuracy of the corresponding meters, they may also contain gross errors due to the failures in telemetry and/or meters. Gross errors can also be intentionally injected to deceive the control decisions of the EMS functions (i.e., security assessment, automatic generation control, and economic dispatch) by exploiting the vulnerabilities of the telemetry systems to cyber attacks. Such cyber attacks, known as data-integrity attacks, aim at altering the breaker status and/or measurements while remaining undetected [2]-[5]. In particular, we focus on a class of attacks known as false data injection attacks (FDIAs) where an adversary compromises a small subset of analog measurements (e. g., voltage magnitudes and angles, power flows, and power injections) to conceal a particular goal, e. g., deceive the system operator into changing generation dispatch, congesting transmission lines, or producing cascading failures [2].

Within the context above, we aim at analyzing the vulnerabilities of the state estimation against FDIAs based on sensitivity analysis. Such vulnerabilities are characterized in terms of the chance of an attack to significantly influence (if perturbed) the optimal estimates while remaining undetected.

The theoretical framework for sensitivity analysis in nonlinear optimization used in this paper is stated in [6], [7], following the pioneering works of [8]-[10]. The applications of this sensitivity analysis framework in power system include state estimation [11], [12] and pricing [13]. Other engineering applications include reliability analysis [14], calculus of variations [15], and optimal design of civil infrastructure [16].

Given the characteristics of the nonlinear state estimation [4], its vulnerability has been studied only in a few works. The vulnerability of the state estimation has been quantified by the minimum number of sensors that have to be compromised to stage a stealthy FDIA, which can be formulated as a minimum cardinality problem [17], [18]. Reference [19] proposes a graph-based algorithm to find the set of compromised sensors needed to stage an unobservable attack assuming that the adversary has perfect information of the system. Reference [20] extends the work in [19] by considering incomplete information of the system. Reference [21] proposes a framework to analyze the vulnerability of the nonlinear state estimation from the perspective of the system operator and presents countermeasures. Reference [22] uses the influence function, which measures the sensitivity of state estimation to an infinitesimal fraction of contamination of the measurements, to identify the influential measurements and parameters. Reference [23] formulates a framework based on a semi-definite convexification of the FDIA to find a near-optimal attack strategy and analyzes the attack stealthiness. They provide the theoretical guarantees of sparsity and unobservability. However, this formulation depends on the adversaries' objective, which is not necessarily always available to the system operator.

In this paper, we tailor the sensitivity analysis methodology in [7] to efficiently analyze vulnerabilities in the state estimation problem with respect to FDIAs. The proposed methodology is an off-line assessment to identify vulnerable measurements in the state estimation rather than to identify corrupted measurements during the system operation. Unlike the existing literature, our methodology neither depends on the adversary's objective nor quantifies the vulnerability by the minimum number of sensors needed to be compromised to stage an unobservable FDIA. From the perspective of the system operator, we rather focus on the vulnerability based on endogenous factors of the state estimation and the power grid such as the measurement configuration, system topology, and network parameters. The main contributions of this paper are threefold.

1) The vulnerabilities of the state estimation with respect to FDIAs based on sensitivity analysis are analyzed. We identify such vulnerabilities in terms of the stealthiness and impactfulness characteristics of an FDIA when it targets a particular measurement. The sensitivity analysis methodology allows us to compute both characteristics of all the measurements simultaneously.

2) Three scores to quantify and rank the vulnerability of each measurement to FDIAs are proposed, which can help identify vulnerable areas of the system and improve its security.

3) The variations of the sensitivities with respect to different operating conditions based on a singular value decomposition (SVD) approach are assessed. We aim at identifying whether the vulnerabilities of the state estimation vary with respect to the operating condition of the system or they remain almost invariant. The latter case would imply that the vulnerabilities are mainly dependent on the network topology and its parameters, and the configuration of the measurements.

Although we illustrate our methodology in the weighted least squared (WLS) state estimator, such methodology can also be implemented using other estimators (e.g., robust estimators) as long as they can be stated as a continuous optimization problem and their solution holds the Karush-Kuhn-Tucker (KKT) optimality conditions.

The remainder of this paper is organized as follows. In Section II, we present the characterization of vulnerable measurements, the state estimation formulation, and the analytical expressions to compute the sensitivities. The method to identify whether such sensitivities change with the operating conditions of the system is described in Section III. The proposed methodology is validated through numerical experiments in two test illustrative systems in Section IV. The effectiveness of the proposed methodology is verified using the IEEE 118-bus test system in Section V. The main conclusions of the paper are summarized in Section VI.

II. Vulnerability Analysis

In this section, we characterize the vulnerability of the measurements against FDIA. Also, we present the state estimation formulation and derive the analytical expressions to compute the sensitivities of the objective and estimated variables with respect to parameters and measurements.

A. Characterization of Vulnerable Measurements

The goal of an FDIA is to stealthily modify measurements to introduce gross errors in the variable estimates, which are then used in other control applications (e. g., security-constrained optimal power flow and security analysis) [4]. This goal shows two main characteristics as follows.

1) Stealthiness

Once the solution of the state estimation is computed, gross errors are detected by comparing the sum of squared errors with a bad data detection (BDD) flag. In the case of the WLS estimation, the widely adopted criterion for this flag comes from a χ^2 distribution [1], [24]. Note that if the state estimation is formulated as an optimization problem, the sum of squared errors will be the value of the objective function.

An adversary aims at modifying measurements without triggering the BDD flag, which could hinder the successful staging of the attack. Thus, an attacker would like to corrupt the measurements that do not change significantly the objective function when they are perturbed, which means that the rate of change in the objective function with respect to the measurement is small.

Although the vulnerability of a measurement can be induced by a low redundancy level around that measurement (critical measurement is an extreme example of this), this is not the only reason for high vulnerability. For example, a leverage measurement, which shows a small rate of changes in the objective function with respect to its perturbation, is also highly vulnerable. The vulnerability of such measurements is not caused by a lack of local redundancy, but by other factors such as system topology and network parameters [25], [26]. We use the χ^2 test as a detection criterion, which allows us to use the sensitivities as a metric to know whether changing a measurement will result in a large change in the objective function.

2) Impactfulness

Besides remaining undetected, an adversary aims at causing a large change in the variable estimates without significantly modifying the measurement under attack, i.e., the rate of change in the variable estimate as a measurement change has to be large. Since the state estimation can be also regarded as a nonlinear regression problem, this characteristic turns out to be the definition of leverage point in regression analysis [27]. Measurements with high leverage have three important characteristics as follows. Firstly, they have a significant influence on the variable estimate when they are perturbed. Secondly, they can be eliminated without losing system observability unless they are critical measurements [1], [25]. Finally, such measurements can also affect the convergence of the estimator [28].

A measurement with both characteristics is a high-potential target for cyber attack as an adversary can stage an impactful attack while remaining likely undetected. We note that both characteristics can be described in terms of the sensitivities of the objective function and the variable estimates with respect to the measurements. The proposed sensitivity analysis allows us to identify any vulnerable measurement of the cause of the vulnerability, e.g., low local redundancy, system topology, and/or network parameters.

We note that since our perspective is that of the system operator, it is a conservative assumption to consider that the attacker has full knowledge of the system. The system hardware parameters, e.g., line parameters, system topology, generator operating limits, and capacity of transmission lines, can be obtained by probing the supervisory control and data acquisition (SCADA) system [29]. The attacker needs to get the appropriate credentials to read these data. Mainly, such credentials can be obtained via implanting malware or compromising the firewall [30]. On the other hand, the measurement weights depend on the accuracy of the measurements. Generally, sensors must comply with a certain level of accuracy (accuracy class), which is generally public information. Additionally, system measurements, e.g., voltages, power injections, and power flows, can be obtained by compromising remote terminal units (RTUs), or launching a man-in-the-middle attack to the communication link between the field devices and the SCADA system. Possible attack methods include address resolution protocol (ARP) poisoning attack [31] and dynamic host configuration protocol (DHCP) starvation attack [32].

The remainder of this section presents a technique to systematically compute both sensitivities for all the measurements simultaneously solely using state estimation information.

B. State Estimation Formulation

The WLS state estimation can be formulated as an equality-constrained optimization problem as follows.

$$\min_{x} \sum_{i\in\mathcal{V}} w_i^V (v_i^m - v_i)^2 + \sum_{i\in\mathcal{P}} w_i^P (P_i^m - P_i)^2 + \sum_{i\in\mathcal{Q}} w_i^Q (Q_i^m - Q_i)^2 + \sum_{(i,j)\in\mathcal{P}_f} w_{i,j}^P (P_{i,j}^m - P_{i,j})^2 + \sum_{(i,j)\in\mathcal{Q}_f} w_{i,j}^Q (Q_{i,j}^m - Q_{i,j})^2 \tag{1}$$

s.t.

$$P_i = v_i \sum_{j\in\mathcal{N}_i} v_j (G_{i,j}\cos\theta_{i,j} + B_{i,j}\sin\theta_{i,j}) \quad i\in\mathcal{P} \tag{2}$$

$$Q_i = v_i \sum_{j\in\mathcal{N}_i} v_j (G_{i,j}\sin\theta_{i,j} - B_{i,j}\cos\theta_{i,j}) \quad i\in\mathcal{Q} \tag{3}$$

$$P_{i,j} = v_i v_j (G_{i,j}\cos\theta_{i,j} + B_{i,j}\sin\theta_{i,j}) - G_{i,j} v_i^2 \quad (i,j)\in\mathcal{P}_f \tag{4}$$

$$Q_{i,j} = v_i v_j (G_{i,j}\sin\theta_{i,j} - B_{i,j}\cos\theta_{i,j}) + v_i^2 (B_{i,j} - b_{i,j}^{\mathrm{sh}}/2) \quad (i,j)\in\mathcal{Q}_f \tag{5}$$

$$0 = v_i \sum_{j\in\mathcal{N}_i} v_j (G_{i,j}\cos\theta_{i,j} + B_{i,j}\sin\theta_{i,j}) \quad i\in\mathcal{Z} \tag{6}$$

$$0 = v_i \sum_{j\in\mathcal{N}_i} v_j (G_{i,j}\sin\theta_{i,j} - B_{i,j}\cos\theta_{i,j}) \quad i\in\mathcal{Z} \tag{7}$$

The objective (1) is to minimize the weighted sum of squared errors. Constraints (2) and (3) represent the active and reactive power injections of the buses with available injection measurements, respectively. Constraints (4) and (5) represent the active and reactive power flows of the lines with available flow measurements, respectively. Constraints (6) and (7) correspond to the zero-injections, i.e., exact pseudo-measurements.

The above problem can be expressed in compact form as:

$$\min_{\boldsymbol{x}} J(\boldsymbol{x},\boldsymbol{a},\boldsymbol{z}) \tag{8}$$

s.t.

$$\boldsymbol{c}(\boldsymbol{x},\boldsymbol{a})=\boldsymbol{0}: \quad \boldsymbol{\lambda} \tag{9}$$

Note that the equality constraints only depend on the optimization variables and the parameters, not the measurements.

In the following subsection, the feasible perturbations and sensitivity analysis are derived assuming that we have a clean set of measurements, i.e., there are not bad data. Therefore, neither the objective function value nor the normalized residuals would trigger any flag.

C. Feasible Perturbations and Sensitivity Analysis

Let $\boldsymbol{x}^*$ be a local optimal solution of (8) and (9), and assume that $\boldsymbol{x}^*$ is regular, i.e., the constraint gradients $\nabla_x c_k(\boldsymbol{x}^*,\boldsymbol{a})$, $k=1,2,\ldots,r$ are linearly independent [33]. Then, the KKT first-order optimality conditions are formulated as [33]:

$$\nabla_x J(\boldsymbol{x}^*,\boldsymbol{a},\boldsymbol{z})+\sum_{k=1}^{r}\lambda_k^*\nabla_x c_k(\boldsymbol{x}^*,\boldsymbol{a})=0 \tag{10}$$

$$c_k(\boldsymbol{x}^*,\boldsymbol{a})=0 \quad k=1,2,\ldots,r \tag{11}$$

The conditions in (11) are the primal feasibility ones.

To determine the sensitivity equations with respect to the parameters and measurements, we perturb $\boldsymbol{x}^*,\boldsymbol{\lambda}^*,J^*,\boldsymbol{a},\boldsymbol{z}$ in such a way that the KKT conditions still hold [7]. Thus, we differenciate the objective function (8) and the optimality conditions (10) and (11) as follows.

$$(\nabla_x J(\boldsymbol{x}^*,\boldsymbol{a},\boldsymbol{z}))^{\mathrm{T}}\mathrm{d}\boldsymbol{x}+(\nabla_a J(\boldsymbol{x}^*,\boldsymbol{a},\boldsymbol{z}))^{\mathrm{T}}\mathrm{d}\boldsymbol{a}+(\nabla_z J(\boldsymbol{x}^*,\boldsymbol{a},\boldsymbol{z}))^{\mathrm{T}}\mathrm{d}\boldsymbol{z}-\mathrm{d}J=\boldsymbol{0} \tag{12}$$

$$\left(\nabla_{xx} J(\boldsymbol{x}^*,\boldsymbol{a},\boldsymbol{z})+\sum_{k=1}^{r}\lambda_k^*\nabla_{xx} c_k(\boldsymbol{x}^*,\boldsymbol{a})\right)\mathrm{d}\boldsymbol{x}+\left(\nabla_{xa} J(\boldsymbol{x}^*,\boldsymbol{a},\boldsymbol{z})+\sum_{k=1}^{r}\lambda_k^*\nabla_{xa} c_k(\boldsymbol{x}^*,\boldsymbol{a})\right)\mathrm{d}\boldsymbol{a}+\nabla_{xz} J(\boldsymbol{x}^*,\boldsymbol{a},\boldsymbol{z})\mathrm{d}\boldsymbol{z}+\nabla_x c(\boldsymbol{x}^*,\boldsymbol{a})\mathrm{d}\boldsymbol{\lambda}=\boldsymbol{0} \tag{13}$$

$$(\nabla_x c(\boldsymbol{x}^*,\boldsymbol{a}))^{\mathrm{T}}\mathrm{d}\boldsymbol{x}+(\nabla_a c(\boldsymbol{x}^*,\boldsymbol{a}))^{\mathrm{T}}\mathrm{d}\boldsymbol{a}=\boldsymbol{0} \tag{14}$$

The above system of equations can be expressed in matrix form as:

$$\begin{bmatrix}\boldsymbol{J}_x & \boldsymbol{J}_a & \boldsymbol{J}_z & \boldsymbol{0} & -1\\ \boldsymbol{J}_{xx} & \boldsymbol{J}_{xa} & \boldsymbol{J}_{xz} & \boldsymbol{C}_x^{\mathrm{T}} & 0\\ \boldsymbol{C}_x & \boldsymbol{C}_a & \boldsymbol{0} & \boldsymbol{0} & 0\end{bmatrix}\begin{bmatrix}\mathrm{d}\boldsymbol{x}\\ \mathrm{d}\boldsymbol{a}\\ \mathrm{d}\boldsymbol{z}\\ \mathrm{d}\boldsymbol{\lambda}\\ \mathrm{d}J\end{bmatrix}=\boldsymbol{0} \tag{15}$$

The vectors and submatrices are defined in Appendix A.

Then, (12) can be written as:

$$\boldsymbol{T}\begin{bmatrix}\mathrm{d}\boldsymbol{x} & \mathrm{d}\boldsymbol{\lambda} & \mathrm{d}J\end{bmatrix}^{\mathrm{T}}=\boldsymbol{S}_a\mathrm{d}\boldsymbol{a}+\boldsymbol{S}_z\mathrm{d}\boldsymbol{z} \tag{16}$$

$$\begin{cases}\boldsymbol{T}=\begin{bmatrix}\boldsymbol{J}_x & \boldsymbol{0} & -\boldsymbol{1}\\ \boldsymbol{J}_{xx} & \boldsymbol{C}_x^{\mathrm{T}} & \boldsymbol{0}\\ \boldsymbol{C}_x & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\\ \boldsymbol{S}_a^{\mathrm{T}}=-\begin{bmatrix}\boldsymbol{J}_a & \boldsymbol{J}_{xa} & \boldsymbol{C}_a\end{bmatrix}\\ \boldsymbol{S}_z^{T}=-\begin{bmatrix}\boldsymbol{J}_z & \boldsymbol{J}_{xz} & \boldsymbol{0}\end{bmatrix}\end{cases} \tag{17}$$

Therefore, (16) be expressed as:

$$\begin{bmatrix}\mathrm{d}\boldsymbol{x} & \mathrm{d}\boldsymbol{\lambda} & \mathrm{d}J\end{bmatrix}^{\mathrm{T}}=\boldsymbol{T}^{-1}\boldsymbol{S}_a\mathrm{d}\boldsymbol{a}+\boldsymbol{T}^{-1}\boldsymbol{S}_z\mathrm{d}\boldsymbol{z} \tag{18}$$

It can be solved using the superposition principle by replacing $\mathrm{d}\boldsymbol{z}$ and $\mathrm{d}\boldsymbol{a}$ by the p- and q-dimension identity matrices, respectively. Then, we obtain the matrices with all sensitivities with respect to the parameters and measurements.

$$\begin{cases}\begin{bmatrix}\dfrac{\partial\boldsymbol{x}}{\partial\boldsymbol{a}} & \dfrac{\partial\boldsymbol{\lambda}}{\partial\boldsymbol{a}}\end{bmatrix}^{\mathrm{T}}=-\boldsymbol{H}_x^{-1}\boldsymbol{H}_a\\ \dfrac{\partial J}{\partial\boldsymbol{a}}=\boldsymbol{J}_a+\boldsymbol{J}_x\dfrac{\partial\boldsymbol{x}}{\partial\boldsymbol{a}}\end{cases} \tag{19}$$

$$\begin{cases}\begin{bmatrix}\dfrac{\partial\boldsymbol{x}}{\partial\boldsymbol{z}} & \dfrac{\partial\boldsymbol{\lambda}}{\partial\boldsymbol{z}}\end{bmatrix}^{\mathrm{T}}=-\boldsymbol{H}_x^{-1}\boldsymbol{H}_z\\ \dfrac{\partial J}{\partial\boldsymbol{z}}=\boldsymbol{J}_z+\boldsymbol{J}_x\dfrac{\partial\boldsymbol{x}}{\partial\boldsymbol{z}}\end{cases} \tag{20}$$

where $\boldsymbol{H}_x=\begin{bmatrix}\boldsymbol{J}_{xx} & \boldsymbol{C}_x^{\mathrm{T}}\\ \boldsymbol{C}_x & \boldsymbol{0}\end{bmatrix}$; $\boldsymbol{H}_a=\begin{bmatrix}\boldsymbol{J}_{xa}\\ \boldsymbol{C}_a\end{bmatrix}$; $\boldsymbol{H}_z=\begin{bmatrix}\boldsymbol{J}_{xz}\\ \boldsymbol{0}\end{bmatrix}$.

Clearly, the sensitivities of the objective and variable estimates with respect to the measurements, which allow us to define the vulnerability of each measurement, can be computed by (20).

D. Identifying Vulnerable Measurements

To better visualize the stealthiness and impactfulness of a measurement z_ℓ, we propose three scores to rank the vulnerability of z_ℓ: ① S-score, which quantifies how likely an FDIA is to be undetected; ② L-score, which quantifies the influence of an FDIA on the variables estimates; ③ V-score, which is a convex combination of the previous scores. The three scores are defined as (21)-(23), respectively.

$$S_{\text{score}}(z_\ell)=f\left(\gamma^{-\left|z_\ell\left(\partial J/\partial z_\ell\right)\right|\Big/\max\limits_{1\le k\le p}\left\{\left|z_k\left(\partial J/\partial z_k\right)\right|\right\}}\right) \tag{21}$$

$$L_{\text{score}}(z_\ell)=g\left(\left\|\frac{\partial\boldsymbol{x}}{\partial z_\ell}\right\|_2\right) \tag{22}$$

$$V_{\text{score}}(z_\ell)=\alpha S_{\text{score}}(z_\ell)+(1-\alpha)L_{\text{score}}(z_\ell) \tag{23}$$

where $\gamma>0$; $\alpha\in[0,1]$; and $f(\cdot)$ and $g(\cdot)$ are the non-decreasing functions with range and domain on $[0,1]$. It is noteworthy that in the computation of L-score, choosing different norms could result in different values of such a score. To score the leverage of z_ℓ, we consider the Euclidean norm of the sensitivities of all the variable estimates with respect to it. This allows us to take into account the influence of such measurement not only on its corresponding variable estimate (i. e., self-sensitivity), but also on the other variable estimates.

The proposed scores are closer to 1 when a measurement is more vulnerable. It is noteworthy that $f(\cdot)$ and $g(\cdot)$ and their arguments are user-defined. We suggest an S-shaped function for both scores such as:

$$f(\xi)=\begin{cases}0 & \xi\le 0\\ \dfrac{1}{1+\left(\dfrac{\xi}{1-\xi}\right)^{-\beta}} & 0<\xi<1\\ 1 & \xi\ge 1\end{cases} \tag{24}$$

where $\beta>0$. We use the S-shaped function as a mechanism

to visualize the vulnerability of the measurements. The parameter β can be understood as a way of controlling how conservative the identification of vulnerable measurements is, i.e., smaller values of β render more conservative scores because the function rapidly downweights the scores as they distance from 1, as depicted in Fig. 1. Note that such a parameter does not affect the order of the scores. For example, in the case of L-score, a measurement with the largest sensitivity of the estimates will always have the highest L-score independent of the choice of β.

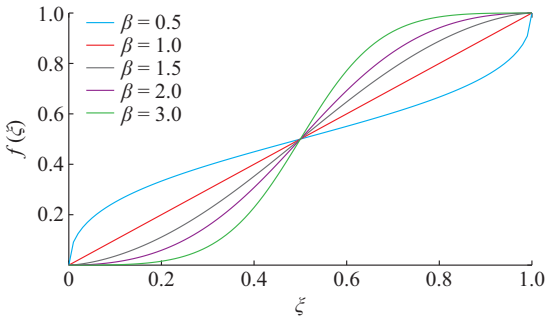

Fig. 1. Curves of S-shaped function.

Finally, the procedure to compute the sensitivities with respect to measurements and the proposed scores is summarized in Algorithm 1.

Algorithm 1: sensitivity analysis of state estimation

Input: optimal solution $(J^*, \boldsymbol{x}^*, \boldsymbol{\lambda}^*)$, parameters $\beta>0$, $\gamma>0$, and $\alpha\in[0, 1]$, and functions $f(\cdot)$ and $g(\cdot)$

Step 1: compute $\boldsymbol{H}_x$, $\boldsymbol{H}_z$, $\boldsymbol{J}_x$, $\boldsymbol{J}_z$

Step 2: compute $\partial J/\partial \boldsymbol{z}$ and $\partial \boldsymbol{x}/\partial \boldsymbol{z}$ by (20)

Step 3: for $\ell=1, 2, ..., p$ do

evaluate (21) - (23) for $S_{\text{score}}(z_\ell)$, $L_{\text{score}}(z_\ell)$, and $V_{\text{score}}(z_\ell)$, respectively

end for

Output: sensitivities $(\partial \boldsymbol{J}/\partial \boldsymbol{z}, \partial \boldsymbol{x}/\partial \boldsymbol{z})$, $S_{\text{score}}(z_\ell)$, $L_{\text{score}}(z_\ell)$, and $V_{\text{score}}(z_\ell)$, $\ell=1, 2, ..., p$

III. Robustness Analysis

In this section, we present a method to identify whether or not the sensitivities change with the system operating condition.

A. Preprocessing

To determine if the sensitivity vectors show significant changes with respect to the operating points, we consider t different operating conditions and compute their corresponding sensitivities. Then, we arrange these sensitivities in matrices $\boldsymbol{X}\in\mathbb{R}^{t\times np}$ and $\boldsymbol{J}\in\mathbb{R}^{t\times p}$ as follows.

$$\begin{cases}\boldsymbol{X}=[\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \cdots \quad \boldsymbol{x}_t]^{\mathrm{T}}\\ \boldsymbol{J}=[\boldsymbol{J}_1 \quad \boldsymbol{J}_2 \quad \cdots \quad \boldsymbol{J}_t]^{\mathrm{T}}\end{cases} \tag{25}$$

where $\boldsymbol{x}_k=\text{vec}(\partial \boldsymbol{x}/\partial \boldsymbol{z})$ and $\boldsymbol{J}_k=\partial J/\partial \boldsymbol{z}$ are the sensitivities at a given operating condition k. Each column of $\boldsymbol{X}$ and $\boldsymbol{J}$ corresponds to a particular sensitivity for all the operating conditions.

Note that in (25) we assume that every sensitivity vector $\boldsymbol{x}_k$ and $\boldsymbol{J}_k$ has the same dimension, i.e., the system topology and measurement configuration remain unchanged, which might not be always true. If the dimensions of the sensitivity vectors are different, it is necessary to only keep the sensitivities that are common for all the operating conditions.

SVD allows us to determine if such sensitivities significantly vary depending on the different operating points. Before computing the SVD of both matrices, it is necessary to subtract the mean of each column, i.e., the mean of each column is zero. We compute the row vectors containing the means of every column as:

$$\begin{cases}\bar{\boldsymbol{x}}=\dfrac{1}{t}\displaystyle\sum_{k=1}^{t}\boldsymbol{x}_k\\ \bar{\boldsymbol{j}}=\dfrac{1}{t}\displaystyle\sum_{k=1}^{t}\boldsymbol{J}_k\end{cases} \tag{26}$$

where $\bar{\boldsymbol{x}}\in\mathbb{R}^{1\times np}$ and $\bar{\boldsymbol{j}}\in\mathbb{R}^{1\times p}$. Then, we can compute the elements of the mean-centered matrices $\tilde{\boldsymbol{X}}$ and $\tilde{\boldsymbol{J}}$ as:

$$\begin{cases}\tilde{\boldsymbol{X}}=\boldsymbol{X}-\boldsymbol{1}_t\bar{\boldsymbol{x}}\\ \tilde{\boldsymbol{J}}=\boldsymbol{J}-\boldsymbol{1}_t\bar{\boldsymbol{j}}\end{cases} \tag{27}$$

B. SVD

SVD is one of the most ubiquitous methods for processing and compressing data as well as dimensionality reduction. Although SVD is considered as a computationally intensive matrix decomposition, significant efforts have been made to propose reliable and numerically efficient algorithms to compute or approximate such decomposition in the last two decades. In particular, the matrices with low-rank structures can be efficiently decomposed by modern randomized matrix algorithms [34].

SVD is helpful to determine if the sensitivities are significantly affected by the different operating points. We compute the SVD of both standardized matrices as:

$$\begin{cases}\tilde{\boldsymbol{X}}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}=\displaystyle\sum_{i=1}^{\min\{t,np\}}\sigma_i\boldsymbol{u}_i\boldsymbol{v}_i^{\mathrm{T}}\\ \tilde{\boldsymbol{J}}=\hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{V}}^{\mathrm{T}}=\displaystyle\sum_{i=1}^{\min\{t,p\}}\hat{\sigma}_i\hat{\boldsymbol{u}}_i\hat{\boldsymbol{v}}_i^{\mathrm{T}}\end{cases} \tag{28}$$

where the diagonal elements of $\boldsymbol{\Sigma}$ and $\hat{\boldsymbol{\Sigma}}$ are the singular values of $\tilde{\boldsymbol{X}}$ and $\tilde{\boldsymbol{J}}$, respectively, and they are ordered from the largest to smallest.

If the largest singular values are significantly larger than the smallest ones, the sensitivities are not strongly dependent on the system operating condition. Such a characteristic is key for a cyber-attack because it means that the sensitivities depend on the factors that do not change significantly over time, e.g., system topology, line parameters, and measurement locations and precisions. Thus, it allows the adversary to identify the target measurements off-line, and to stage an attack on one of these measurements without knowing other measurements.

To quantify the proportion of the variance of the mean-centered sensitivity matrices $\tilde{\boldsymbol{X}}$ and $\tilde{\boldsymbol{J}}$ captured by their first r singular values σ_i $(i=1, 2, \ldots, r)$, the cumulative energy (CE) is defined as:

$$\begin{cases} CE(\tilde{\boldsymbol{X}};r)=\dfrac{\sum_{i=1}^{r}\sigma_i}{\sum_{i=1}^{\min\{t,np\}}\sigma_i} \\ CE(\tilde{\boldsymbol{J}};r)=\dfrac{\sum_{i=1}^{r}\hat{\sigma}_i}{\sum_{i=1}^{\min\{t,p\}}\hat{\sigma}_i} \end{cases} \tag{29}$$

IV. Illustrative Examples

In this section, two case studies are analyzed considering a 4-bus system and the New England 39-bus system. The weights of the voltage measurements are assumed to be $w^V = 1\times10^4$, whereas the remaining measurements have the weight of $w = 2.5\times10^3$. For the sake of simplicity, we weigh the squared error of each measurement with the inverse of the variance of its meter. We note that more sophisticated weighting rules are possible [35]. We consider 24 operating conditions, which are generated by multiplying all the demands by the scale factors, and that the topology of the systems remains unchanged. Such scale factors are described in Appendix B. Also, we set $\alpha = 0.3$, and $\beta = 1$ and $\beta = 1.5$ for the S-score and L-score, respectively.

A. 4-bus System

The 4-bus system and its measurement configuration are depicted in Fig. 2, where P_2^m and Q_2^m are the zero-injection measurements. The presented measurement configuration provides a redundancy ratio of 1.71. The bus and branch data are detailed in Appendix C.

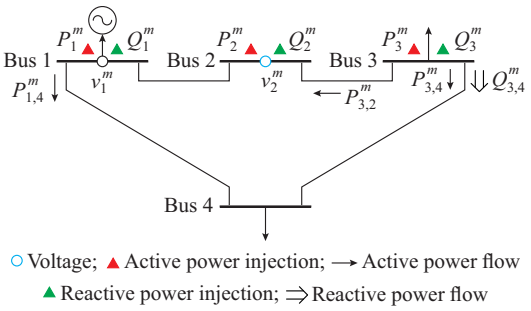

Fig. 2. Single-line diagram and measurement location of 4-bus system.

The sensitivity of the objective function with respect to the measurements is depicted in Fig. 3, where a darker color means that the measurement is more likely to be a target of undetected FDIAs. We note that the magnitudes of the sensitivities remain almost invariant with the operating conditions. P_1^m is the measurement with the largest normalized sensitivity $\left(\left|z_k\partial J/\partial z_\ell\right| \Big/ \max_k\left\{\left|z_k\partial J/\partial z_\ell\right|\right\}\right)$, thus is less vulnerable in terms of stealthiness as an FDIA against it is unlikely to remain undetected. Conversely, Q_3^m is the measurement with the smallest sensitivity followed by $Q_{3,4}^m$, Q_1^m, and $P_{3,2}^m$, respectively.

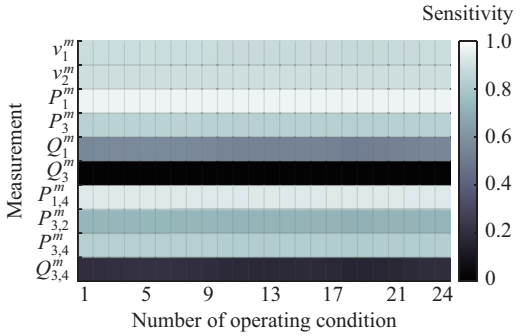

Fig. 3. Sensitivity of objective function with respect to measurement of 4-bus system.

Likewise, the sensitivity of the variable estimates with respect to the measurements at the maximum demand is depicted in Fig. 4. The sensitivities with larger absolute values are depicted with darker colors. The measurements with the highest self-sensitivities are Q_1^m, Q_3^m, P_3^m, and P_1^m, which are the most vulnerable ones in terms of impactfulness. Specifically, Q_1^m shows the largest self-sensitivity; v_1^m and v_2^m show the largest impact on the other variable estimates (v_3 and v_4). An FDIA compromising these measurements will have a significant impact on the corresponding variable estimates. Furthermore, it is convenient to analyze the dependence of the variable estimates with respect to each measurement. v_1^m and v_2^m have a significant influence on the estimates of v_3 and v_4, respectively. Similarly, P_1^m and Q_3^m ($P_{1,4}^m$ and $Q_{3,4}^m$) have a significant influence on the variable estimates of $P_{1,4}$ and $Q_{3,4}$ (P_1 and Q_3), respectively.

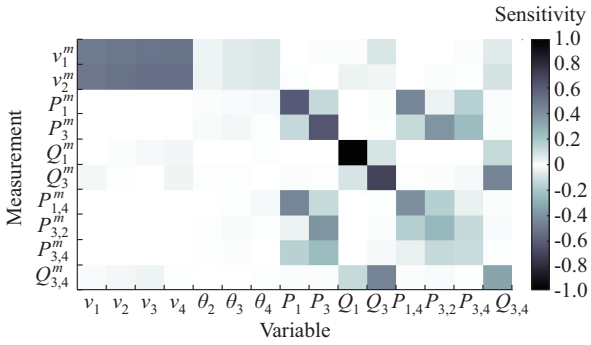

Fig. 4. Sensitivity of variable estimates with respect to measurement (scale factor is 1) of 4-bus system.

Table I provides the three proposed scores of each measurement when the scale factor is 1. It can be seen that Q_1^m is the most vulnerable measurement followed by Q_3^m. All these measurements are vulnerable due to the lack of redundancy of reactive power measurements near buses 1 and 3. These measurements exhibit the highest potential to be targeted for cyber-attacks. The best chance to stage a stealthy and impactful FDIA is corrupting any of these two measurements. To assess the impact of increasing the local redundancy of reactive power measurements near buses 1 and 3, we add two new measurements, $Q_{3,2}^m$ and $Q_{1,4}^m$, which provide a redundancy ratio of 2. As shown in Table II, the scores of

all the measurements decrease. That is, the measurements are less vulnerable. In particular, the V-score of Q_1^m, which is the most vulnerable measurement of the original measurement configuration, drops from 0.9366 to 0.6731.

TABLE I
VULNERABILITY SCORES OF 4-BUS SYSTEM

Measurement	S-score	L-score	V-score
v_1^m	0.4037	1.0000	0.8211
v_2^m	0.3506	1.0000	0.8052
P_1^m	0.0334	0.8692	0.6184
P_3^m	0.5092	0.8794	0.7683
Q_1^m	0.7913	0.9989	0.9366
Q_3^m	0.8927	0.9188	0.9109
$P_{1,4}^m$	0.1974	0.6919	0.5436
$P_{3,2}^m$	0.6747	0.5112	0.5603
$P_{3,4}^m$	0.5331	0.2678	0.3474
$Q_{3,4}^m$	0.8725	0.6076	0.6871

TABLE II
VULNERABILITY SCORES WITH HIGHER REDUNDANCY OF 4-BUS SYSTEM

Measurement	S-score	L-score	V-score
v_1^m	0.1382	1.0000	0.7415
v_2^m	0.2918	1.0000	0.7875
P_1^m	0.0035	0.8689	0.6092
P_3^m	0.3465	0.8791	0.7193
Q_1^m	0.2061	0.8733	0.6731
Q_3^m	0.7698	0.8761	0.8442
$P_{1,4}^m$	0.0764	0.6919	0.5072
$P_{3,2}^m$	0.5271	0.5110	0.5159
$P_{3,4}^m$	0.3714	0.2675	0.2986
$Q_{3,4}^m$	0.9370	0.6929	0.7661
$Q_{3,2}^m$	0.9805	0.5296	0.6649
$Q_{1,4}^m$	0.0061	0.2494	0.1764

SVD can be used to approximate matrices by keeping the most dominant singular vectors, which allows retaining their most relevant features. For example, in Fig. 5(a), the most dominant (largest) singular value is at least one order of magnitude greater than the second one and almost four orders of magnitudes greater than the third one, which means that the most dominant singular value captures most of the relevant features of matrices $\boldsymbol{X}$ and $\boldsymbol{J}$. Figure 5 shows that the leading singular values of $\boldsymbol{X}$ and $\boldsymbol{J}$ account for almost 95% and 96% of their variance, respectively. This means that the other 23 singular vectors provide only about 5% of the variance of the matrix, i.e., the matrix has a low-rank structure. Thus, both sensitivity vectors are almost invariant to the different operating points, which indicates that the vulnerabilities are mainly dependent on the network topology and its parameters, and the configuration of the measurements.

To validate the effectiveness of the proposed scores, we corrupt Q_1^m, which is the most vulnerable measurement, in such a way that it remains undetected.

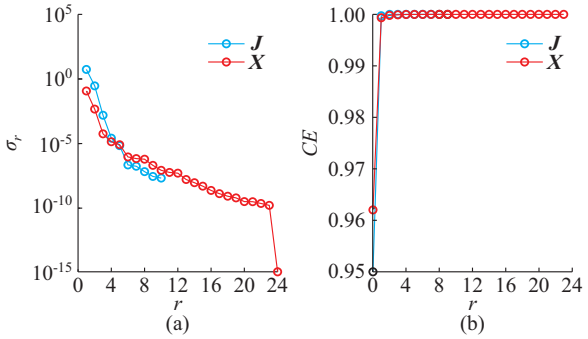

Fig. 5. Singular value σ_r and CE in the first r singular values of 4-bus system. (a) σ_r. (b) CE.

We modify the value of Q_1^m from 2.032432 p.u. to 1.702162 p.u., which represents a deviation of 16.25% from the original measured value. Figure 6 depicts three sets of values, namely the true values, the estimated values without corrupted measurements, and the estimated values with corrupted measurements. The estimated value of Q_1 without corrupted measurements is 2.031062 p.u., whereas the estimated value of Q_1 with corrupted measurements is 1.708243 p.u., i.e., having a deviation of 15.89%.

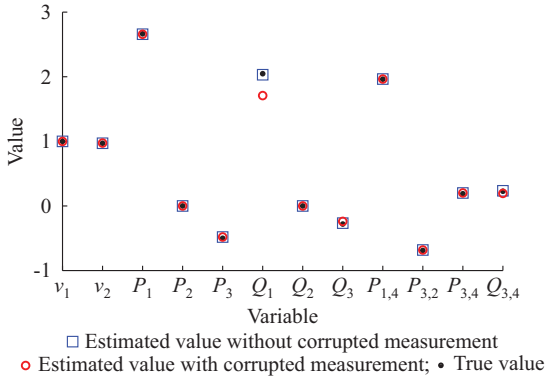

Fig. 6. Estimated values of 4-bus system.

B. New England 39-bus System

We consider that the New England 39-bus system has the following measurements: all the voltage magnitudes, 10 pairs of active and reactive power injections at all generation buses, and 46 pairs of active and reactive power flows at the sending ends of all lines, which results in a redundancy level of 1.96. The system data can be retrieved from MATPOWER [36].

The sensitivity of the variable estimates with respect to the measurements is depicted in Fig. 7. Each block in Fig. 7 represents the sensitivities of a certain set of state estimates with respect to a set of measurements. The voltage measurements are not leverage points since their self-sensitivities are small. Conversely, the majority of the active and reactive power flows and injection variables have high sensitivity with respect to their corresponding measurements. Note that some active and reactive power measurements show non-negligible mutual sensitivities with some state estimates.

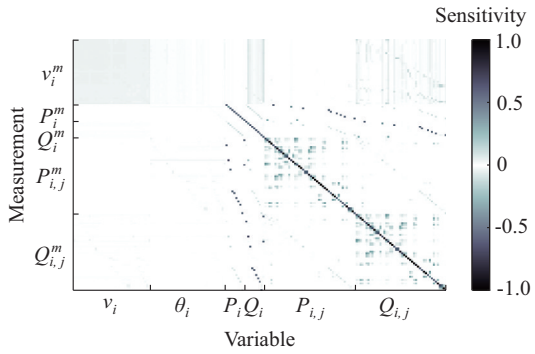

Fig. 7. Sensitivity of variable estimates with respect to measurements (scale factor is 1) of New England 39-bus system.

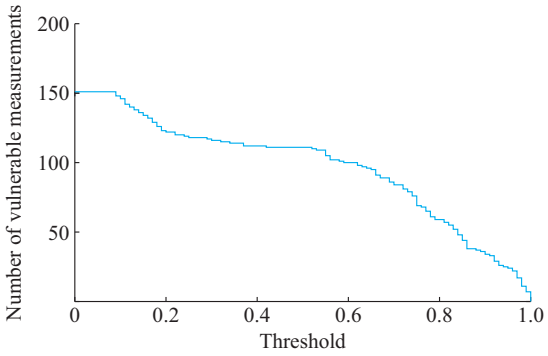

Fig. 9. Number of vulnerable measurements as a function of different threshold values of New England 39-bus system.

Additionally, Fig. 8 depicts the vulnerability scores of all the measurements. An important number of measurements has an S-score near 1. Hence, these measurements are attractive to an attacker in view of the stealthiness since they can be corrupted with gross errors without triggering the BDD flag. The L-score does not show the same distribution; however, there are 48 measurements whose L-score is greater than 0.8, and 22 active and reactive power flow measurements have a V-score greater than 0.95. These results show that the lack of redundancy of active and reactive power measurements is not localized in a certain area of the system, which may be due to the low redundancy ratio.

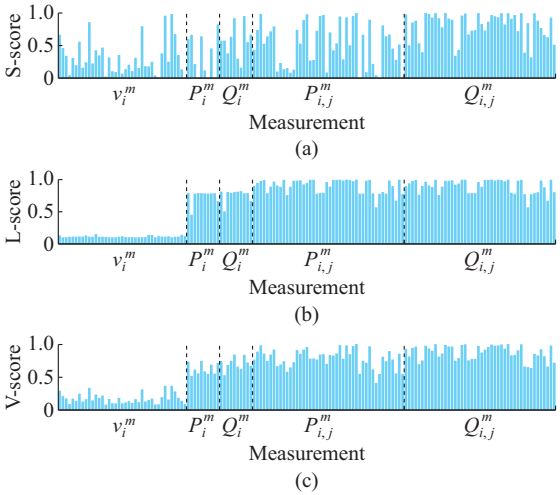

Fig. 8. Vulnerability scores of all measurements of New England 39-bus system. (a) S-score. (b) L-score. (c) V-score.

We also provide the number of vulnerable measurements as a function of different threshold values in Fig. 9. We consider that a measurement z_ℓ is vulnerable if $V_{\text{score}}(z_\ell)$ is more than the threshold. A smaller threshold implies higher conservativeness as it results in declaring a larger number of measurements as vulnerable. Table III lists the ten most critical measurements, $V_{\text{score}}(z_\ell) \geq 0.9836$, in descending order of their V-score. Clearly, these measurements are potential targets of FDIAs as their scores are close to 1, which means that if they are perturbed, they significantly influence their corresponding variable estimates.

TABLE III
TEN MOST CRITICAL VULNERABILITY SCORES OF NEW ENGLAND 39-BUS SYSTEM

Measurement	S-score	L-score	V-score
$P_{19,20}^m$	0.999176	1.000000	0.999753
$Q_{19,20}^m$	0.997233	0.999953	0.999137
$Q_{5,6}^m$	0.998304	0.996979	0.997377
$Q_{22,23}^m$	0.996568	0.995312	0.995689
$Q_{10,13}^m$	0.992317	0.997078	0.995649
$Q_{17,18}^m$	0.996833	0.992793	0.994005
$Q_{16,19}^m$	0.969506	0.999661	0.990615
$Q_{6,11}^m$	0.998550	0.980853	0.986162
$Q_{10,11}^m$	0.960126	0.997099	0.986007
$P_{16,19}^m$	0.945391	1.000000	0.983617

The leading singular values of $\boldsymbol{X}$ and $\boldsymbol{J}$, presented in Table IV, capture around 65% and 80% of their variance, respectively. In the case of $\boldsymbol{J}$, its first three singular values account for more than 97% of its variance. On the other hand, the four leading singular values of $\boldsymbol{X}$ capture around 90% of its variance. These results show the low-rank characteristic of both matrices.

TABLE IV
SVD RESULTS OF NEW ENGLAND 39-BUS SYSTEM

r	Singular value σ_r		CE	
	$\boldsymbol{J}$	$\boldsymbol{X}$	$\boldsymbol{J}$	$\boldsymbol{X}$
1	304.2029	3.4085	0.8010	0.6525
2	54.4586	1.2666	0.9443	0.8950
3	14.2156	0.4722	0.9818	0.9854
4	4.6080	0.0643	0.9939	0.9977
5	2.2604	0.0105	0.9999	0.9997
6	0.0398	0.0011	1.0000	0.9999

V. CASE STUDY

In this section, we verify the effectiveness of the proposed methodology using the IEEE 118-bus test system with the following measurements: all the voltage magnitudes, 54 pairs of active and reactive power injections at all generation buses, and 179 pairs of active and reactive power flows at the sending ends of all lines, which results in a redundancy

level of 2.49. The system data can be retrieved from MATPOWER [36]. The weights of the voltage measurements are assumed to be $w^V = 1\times10^4$, whereas the remaining measurements have the weight of $w = 2.5\times10^3$. We consider that the topology remains unchanged and there are 24 operating conditions, which are generated by multiplying all the demands by the scale factors presented in Appendix B. Additionally, we select $\alpha = 0.3$, and $\beta = 1$ and $\beta = 1.5$ for the S-score and L-score, respectively.

We present the scores of the most vulnerable measurements in Table V. We note that these measurements present attractive characteristics to be targeted by attackers. An attacker could corrupt any of these measurements without triggering the BDD routine and having a significant impact on the state estimates. We note that, even though the measurement configuration results in a reasonable redundancy level, there are 33 measurements, which represent 5.65% of the total number of measurements, with V-scores higher than 0.95.

TABLE V
VULNERABILITY SCORES OF IEEE 118-BUS SYSTEM

Measurement	S-score	L-score	V-score
$P^m_{76,118}$	0.999576	0.999734	0.999687
$P^m_{114,115}$	0.999961	0.999097	0.999357
$Q^m_{91,92}$	0.997856	0.994574	0.995559
$P^m_{49,51}$	0.998686	0.993842	0.995295
$P^m_{77,82}$	0.993701	0.993031	0.993232
$P^m_{55,59}$	0.991076	0.993085	0.992482
$Q^m_{55,59}$	0.998807	0.989771	0.992481
$Q^m_{49,51}$	0.999096	0.986709	0.990425
$Q^m_{12,14}$	0.999696	0.986201	0.990250
$Q^m_{86,87}$	0.992684	0.986354	0.988253
$P^m_{25,27}$	0.986790	0.985819	0.986110
$P^m_{91,92}$	0.955553	0.997469	0.984894
$Q^m_{77,82}$	0.970189	0.989875	0.983969
$Q^m_{25,27}$	0.999993	0.974777	0.982342
$P^m_{22,23}$	0.998919	0.974171	0.981595
$Q^m_{22,23}$	0.999266	0.971170	0.979599
$P^m_{12,16}$	0.984885	0.976930	0.979317

Figures 10-12 depict the distributions of the three proposed scores (scale factor is 1). Note that an important number of measurements have an S-score close to 1, which implies that changing those measurements will not cause to change the objective function significantly. In fact, more than 61% of the measurements have an S-score greater than 0.9. Conversely, the distribution of L-score shows that a smaller set of measurements has the potential to significantly change the state estimation. There are 64 measurements with L-score higher than 0.9. Figure 12 shows how V-score weighs both characteristics to provide an insight into the vulnerability of the measurements. We note that in this case study, the measurements with a high L-score also have a high S-score. The converse is not necessarily true.

Table VI presents the SVD results of $\boldsymbol{X}$ and $\boldsymbol{J}$. The leading singular values of $\boldsymbol{X}$ and $\boldsymbol{J}$ capture around 63% and 70%, respectively. The three leading singular values capture more than 96% of the variance of both matrices, whereas the 8 largest singular values capture all the variances. Note as well that the leading singular value of $\boldsymbol{J}$ is one order of magnitude greater than the second one and three orders of magnitudes greater than the sixth one, which shows the low-rank structure of $\boldsymbol{J}$.

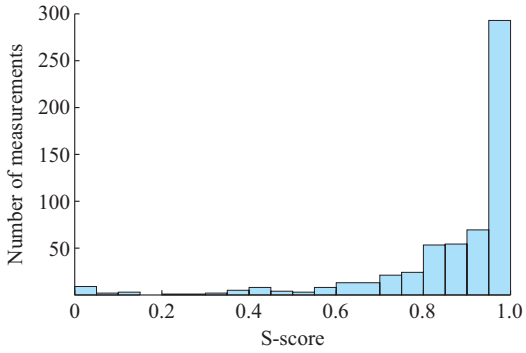

Fig. 10. S-score distribution of IEEE 118-bus system.

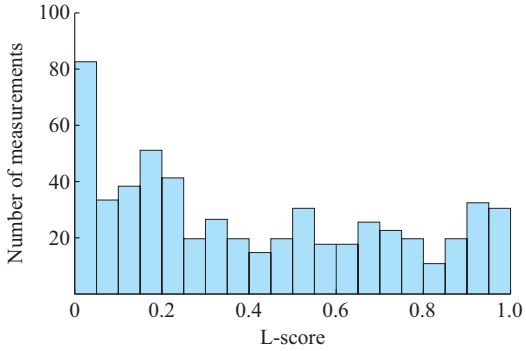

Fig. 11. L-score distribution of IEEE 118-bus system.

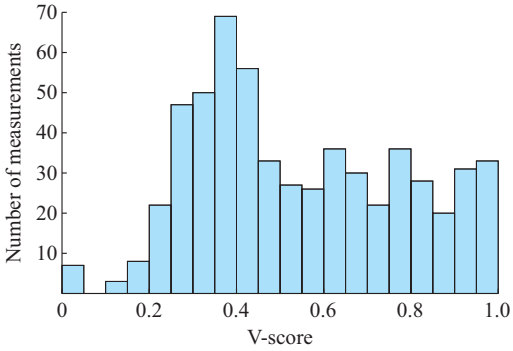

Fig. 12. V-score distribution of IEEE 118-bus system.

We also analyze the influence of an extreme operating condition in the proposed V-score. We assume that the system is operating close to voltage collapse. Figure 13 depicts the correlation between the V-scores of the measurements when the system is operating in nominal condition (scale factor is 1) and the ones when the system is operating close to voltage collapse (scale factor is 1.75). The measurements with higher V-scores are strongly correlated, i.e., if a V-score is high in nominal condition, it is also high in the heavy load condition.

TABLE VI
SVD RESULTS OF IEEE 118-BUS SYSTEM

r	Singular value σ_r		CE	
	$\boldsymbol{J}$	$\boldsymbol{X}$	$\boldsymbol{J}$	$\boldsymbol{X}$
1	121340.81	2.41	0.6983	0.6279
2	35982.14	1.11	0.9054	0.9172
3	10202.53	0.23	0.9641	0.9761
4	4099.27	0.08	0.9877	0.9957
5	1982.16	0.01	0.9991	0.9980
6	111.80	0.01	0.9997	0.9998
7	44.79	0	1.0000	0.9999
8	5.70	0	1.0000	1.0000

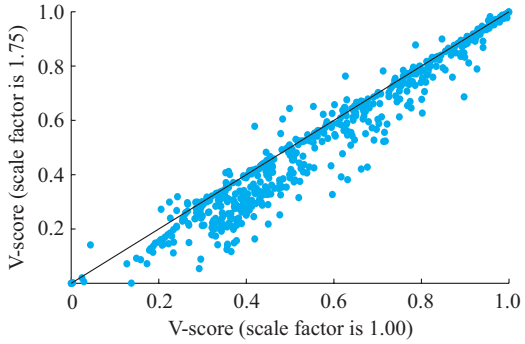

Fig. 13. V-score correlation in a heavy load condition of IEEE 118-bus system.

VI. CONCLUSION

This paper proposes a technique based on sensitivity analysis to identify the measurements with a high potential of being the target of FDIAs. We characterize the vulnerability of each measurement as a function of their potential to impact the variable estimates and to remain stealthy.

In our numerical studies, we demonstrate that there is a subset of measurements that shows both characteristics, thus being the most vulnerable to FDIAs. Furthermore, we numerically demonstrate that such vulnerabilities remain almost invariant to the system operating condition, which implies that they are mainly dependent on the network topology and its parameters, and the measurement configuration.

The proposed technique can be used to identify the most vulnerable measurements. Additionally, identifying such measurements can be used as an input to determine strategies to secure the state estimator, which is out of the scope of this work. Such strategies include: ① locating new measurements to improve local redundancy; ② securing the communication with a small but important subset of measurements; ③ implementing robust estimators.

APPENDIX A

The auxiliary submatrices and vectors in (15) necessary for computing the sensitivities are defined as:

$$\boldsymbol{J}_{x(1\times n)} = (\nabla_x J(\boldsymbol{x}^*, \boldsymbol{a}, \boldsymbol{z}))^{\mathrm{T}} \tag{A1}$$

$$\boldsymbol{J}_{a(1\times q)} = (\nabla_a J(\boldsymbol{x}^*, \boldsymbol{a}, \boldsymbol{z}))^{\mathrm{T}} \tag{A2}$$

$$\boldsymbol{J}_{z(1\times p)} = (\nabla_z J(\boldsymbol{x}^*, \boldsymbol{a}, \boldsymbol{z}))^{\mathrm{T}} \tag{A3}$$

$$\boldsymbol{J}_{xx(n\times n)} = \nabla_{xx} J(\boldsymbol{x}^*, \boldsymbol{a}, \boldsymbol{z}) + \sum_{k=1}^{r} \lambda_k^* \nabla_{xx} c_k(\boldsymbol{x}^*, \boldsymbol{a}) \tag{A4}$$

$$\boldsymbol{J}_{xa(n\times q)} = \nabla_{xa} J(\boldsymbol{x}^*, \boldsymbol{a}, \boldsymbol{z}) + \sum_{k=1}^{r} \lambda_k^* \nabla_{xa} c_k(\boldsymbol{x}^*, \boldsymbol{a}) \tag{A5}$$

$$\boldsymbol{J}_{xz(n\times p)} = \nabla_{xz} J(\boldsymbol{x}^*, \boldsymbol{a}, \boldsymbol{z})\mathrm{d}\boldsymbol{z} \tag{A6}$$

$$\boldsymbol{C}_{x(r\times n)} = \left[\nabla_x c(\boldsymbol{x}^*, \boldsymbol{a})\right]^{\mathrm{T}} \tag{A7}$$

$$\boldsymbol{C}_{a(r\times q)} = \left[\nabla_a c(\boldsymbol{x}^*, \boldsymbol{a})\right]^{\mathrm{T}} \tag{A8}$$

APPENDIX B

The scale factors of the 24 operating conditions are presented in Table BI [37].

TABLE BI
SCALE FACTORS OF 24 OPERATING CONDITIONS

Time (hour)	Demand factor	Time (hour)	Demand factor	Time (hour)	Demand factor
1	0.684511335	9	0.706039246	17	0.874071252
2	0.644122690	10	0.787007049	18	1.000000000
3	0.613069156	11	0.839016956	19	0.983615927
4	0.599733283	12	0.852733854	20	0.936368832
5	0.588874071	13	0.870642027	21	0.887597638
6	0.598018670	14	0.834254144	22	0.809297009
7	0.626786054	15	0.816536483	23	0.745856354
8	0.651743189	16	0.819394170	24	0.733473043

APPENDIX C

The data of the 4-bus system are presented in Tables CI and CII. The bus data correspond to the solution of the power flow at the demand factor of 1.

TABLE CI
BUS DATA OF 4-BUS SYSTEM

Bus	P_g (MW)	Q_g (Mvar)	P_d (MW)	Q_d (Mvar)
1	266.43	204.71		
2				
3			50	28
4			210	180

TABLE CII
BRANCH DATA OF 4-BUS SYSTEM

Line	$r_{i,j}$	$x_{i,j}$	$b_{i,j}^{\mathrm{sh}}$
(1,2)	0.01008	0.0504	0.1025
(1,4)	0.00744	0.0372	0.0775
(1,3)	0.00744	0.0372	0.0775
(1,4)	0.01272	0.0636	0.1275

REFERENCES

[1] A. Abur and A. Gómez-Expósito, "Power system state estimation: theory and implementation," in *Power Engineering*, New York: Marcel Dekker, 2004.

[2] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, pp. 1-33, May 2011.
[3] R. Deng, G. Xiao, R. Lu *et al.*, "False data injection on state estimation in power systems – attacks, impacts, and defense: a survey," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 411-423, Apr. 2017.
[4] G. Liang, J. Zhao, F. Luo *et al.*, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630-1638, Jul. 2017.
[5] G. Wang, G. B. Giannakis, and J. Chen, "Robust and scalable power system state estimation via composite optimization," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6137-6147, Nov. 2019.
[6] E. Castillo, A. J. Conejo, R. Mínguez *et al.*, "A closed formula for local sensitivity analysis in mathematical programming," *Engineering Optimization*, vol. 38, no. 1, pp. 93-112, Jan. 2006.
[7] E. Castillo, A. J. Conejo, C. Castillo *et al.*, "Perturbation approach to sensitivity analysis in mathematical programming," *Journal of Optimization Theory & Applications*, vol. 128, no. 1, pp. 49-74, Jan. 2006.
[8] A. V. Fiacco, "Introduction to sensitivity and stability analysis in nonlinear programming," in *Mathematics in Science and Engineering*, New York: Academic Press, 1983.
[9] I. Enevoldsen, "Sensitivity analysis of reliability-based optimal solution," *Journal of Engineering Mechanics*, vol. 120, no. 1, pp. 198-205, Jan. 1994.
[10] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, New York: Springer, 2000.
[11] R. Mínguez and A. J. Conejo, "State estimation sensitivity analysis," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1080-1091, Aug. 2007.
[12] E. Caro, A. J. Conejo, and R. Mínguez, "A sensitivity analysis method to compute the residual covariance matrix," *Electric Power Systems Research*, vol. 81, no. 5, pp. 1071-1078, May 2011.
[13] A. J. Conejo, E. Castillo, R. Mínguez *et al*., "Locational marginal price sensitivities," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 2026-2033, Nov. 2005.
[14] E. Castillo, A. J. Conejo, R. Mínguez *et al.*, "An alternative approach for addressing the failure probability-safety factor method with sensitivity analysis," *Reliability Engineering & System Safety*, vol. 82, no. 2, pp. 207-216, Nov. 2003.
[15] E. Castillo, A. J. Conejo, and E. Aranda, "Sensitivity analysis in calculus of variations: some applications," *SIAM Review*, vol. 50, no. 2, pp. 294-312, Jun. 2008.
[16] R. Mínguez, E. Castillo, C. Castillo *et al.*, "Optimal cost design with sensitivity analysis using decomposition techniques: application to composite breakwaters," *Structural Safety*, vol. 28, no. 4, pp. 321-340, Sept. 2006.
[17] O. Kosut, L. Jia, R. J. Thomas *et al.*, "Malicious data attacks on smart grid state estimation: attack strategies and countermeasures," in *Proceedings of 2010 IEEE International Conference on Smart Grid Communications*, Gaithersburg, USA, Oct. 2010, pp. 220-225.
[18] G. Dán and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *Proceedings of 2010 IEEE International Conference on Smart Grid Communications*, Gaithersburg, USA, Oct. 2010, pp. 1-6.
[19] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362-1370, Sept. 2012.
[20] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks against nonlinear state estimation in smart power grids," in *Proceedings of 2013 IEEE PES General Meeting*, Vancouver, Canana, Jul. 2013, pp. 1-5.
[21] J. Zhao, L. Mili, and M. Wang, "A generalized false data injection attacks against power system nonlinear state estimator and countermeasures," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4868-4877, Sept. 2018.
[22] J. B. Zhao, S. Fliscounakis, P. Panciatici *et al.*, "Robust parameter estimation of the french power system using field data," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5334-5344, Sept. 2019.
[23] M. Jin, J. Lavaei, and K. H. Johansson, "Power grid AC-based state estimation: vulnerability analysis against cyber attacks," *IEEE Transactions on Automatic Control*, vol. 64, no. 5, pp. 1784-1799, May 2019.
[24] A. Gómez-Expósito, A. J. Conejo, and C. Cañizares, *Electric Energy Systems: Analysis and Operation*, 2nd ed., Boca Raton: CRC Press, 2018.
[25] J. Zhao and L. Mili, "Vulnerability of the largest normalized residual statistical test to leverage points," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4643-4646, Jul. 2018.
[26] A. Majumdar and B. C. Pal, "Bad data detection in the con-text of leverage point attacks in modern power networks," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 2042-2054, May 2018.
[27] R. T. St. Laurent and R. D. Cook, "Leverage and superleverage in nonlinear regression," *Journal of the American Statistical Association*, vol. 87, no. 420, Dec. 1992, pp. 985-990.
[28] J. B. Zhao, L. Mili, and R. C. Pires, "Statistical and numerical robust state estimator for heavily loaded power systems," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6904-6914, Jun. 2018.
[29] M. Kezunovic, "Monitoring of power system topology in real-time," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, Kauai, USA, Jan. 2006, pp. 1-10.
[30] I. Nai Fovino, A. Carcano, M. Masera *et al.*, "An experimental investigation of malware attacks on SCADA systems," *International Journal of Critical Infrastructure Protection*, vol. 2, no. 4, pp. 139-145, Dec. 2009.
[31] S. Y. Nam, S. Jurayev, S.-S. Kim *et al.*, "Mitigating ARP poisoning-based man-in-the-middle attacks in wired or wireless LAN," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, p. 89, Mar. 2012.
[32] H. Mukhtar, K. Salah, and Y. Iraqi, "Mitigation of DHCP starvation attack," *Computers & Electrical Engineering*, vol. 38, no. 5, pp. 1115-1128, Sept. 2012.
[33] D. G. Luenberger and Y. Ye, "Linear and nonlinear programming," in *International Series in Operations Research & Management Science*, New York: Springer, 2008.
[34] N. B. Erichson, S. Voronin, S. L. Brunton *et al.*, "Randomized matrix decompositions using R," *Journal of Statistical Software*, vol. 89, no. 1, pp. 1-48, Jun. 2019.
[35] A. de la Villa Jaén, J. B. Martínez, A. Gómez-Expósito *et al.*, "Tuning of measurement weights in state estimation: theoretical analysis and case study," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4583-4592, Jul. 2018.
[36] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12-19, Feb. 2011.
[37] A. Soroudi, *Power System Optimization Modeling in GAMS*, New York: Springer, 2017.

Gonzalo E. Constante-Flores received the electrical engineering degree from Escuela Politécnica Nacional (EPN), Quito, Ecuador, in 2014, and the M. Sc. degree from The Ohio State University, Columbus, USA, in 2018, where he is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering. From 2013 to 2016, he was with the Department of Electrical Energy at EPN. His research interests include optimization and control of power systems, electricity markets, and integration of renewable energy and electric vehicles into electric power systems.

Antonio J. Conejo received the M.S. degree from the Massachusetts Institute of Technology (MIT), Cambridge, USA, in 1987, and the Ph.D. degree from the Royal Institute of Technology, Stockholm, Sweden, in 1990. He is currently a Professor in the Department of Integrated Systems Engineering and the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, USA. His research interests include control, operation, planning, economics and regulation of electric energy systems, as well as statistics and optimization theory and its applications.

Jiankang Wang received the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, USA, in 2009 and 2013, respectively. She joined the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, USA, as an Assistant Professor in 2014, and appointed as the Lead Technical Specialist of California Independent System Operator in 2018. Her research interests include power system cybersecurity, renewable energy integration, and electricity markets.