Coordinated Cyber-Attacks on the Measurement Function in Hybrid State Estimation

Yacine Chakhchoukh, *Member, IEEE*, and Hideaki Ishii, *Senior Member, IEEE*

***Abstract*—The paper assesses the cyber-security of power systems static state estimation (SE) in the possible presence of phasor measurement units (PMUs). Attacks are considered in the Jacobian matrix or the measurement function of the state estimation leading to the presence of coordinated leverage points. Leverage points, which are outliers, constitute a very challenging attack configuration even if randomly present. It is shown that coordinated cyber-attacks when applied to the Jacobian matrix raise major concerns about robust SE. The vulnerability of the least trimmed squares (LTS) estimator, which is robust towards leverage points, is shown. More generally, the weaknesses of robust regression equivariant estimators are discussed if attacks are developed and optimized based on a projection framework. Attack scenarios are outlined considering the number of attacked Jacobian elements, a decomposition of the system to maximize robustness, and whether a DC or AC formulation is used by the operator. Stealthy attacks that stay undetected with respect to the robust LTS are studied. Masked attacks are defined as well. Some possible solutions and remedial actions are proposed. Robust state estimation methods are evaluated and compared in the presence of different configurations of attacks through Monte Carlo simulations on the IEEE 14- and 30-bus test beds.**

***Index Terms*—Cyber-security, leverage points, PMUs, robust state estimation, smart grid.**

I. Introduction

THE need to study the cyber-security of power systems is becoming more crucial than ever before [1]. Power systems are becoming more complex with an increased uncertainty that justifies the involvement of more active real time diagnostic and monitoring. This will guarantee the stability of the future power grid and will minimize its operating costs. Among the few factors of future development for power systems are: 1) the integration of more renewable random generation sources such as solar and wind; 2) the rising connectivity and power exchange between different areas; 3) positive trend for more competitive and free electricity markets. These factors explain the necessity of having a cyber-system next to the power system in order to exchange data, communicate, analyze and control the grid. Furthermore, with the development of smart meters and phasor measurement units (PMUs), very large sets of data are available and need to be transferred and analyzed in near real time. These different tasks should be completed automatically, with high frequency and within short execution time intervals to operate efficiently the future power system. As a direct consequence of this evolution, the vulnerability and risks linked to cyber-attacks of operating the system will grow. An attacker can, for example, complete his adverse actions from remote places without even accessing the physical system. Since disturbing the normal operation of the power cyber-system has tremendous financial and security effects, it is vital to offer effective counter-measures to cyber-attacks.

One of the most important tasks for operating the grid is estimating the state of the system. A state estimator (SE) evaluates the voltage magnitudes and phase angles at different buses of interest [2], [3]. SE is important for contingency analysis to improve security, load forecasting, evaluating locational marginal pricing (LMP) for power markets, control and many other fundamental power applications. Different aspects of SE have been widely studied in power systems literature [2], [3]. Among those aspects, robust state estimation and outliers detection are of great importance [4], [5]. Outliers are deviations from an assumed model such as errors due to sensor failures, communication or human errors, certain connection or line conditions. The proposed robust methods consider outliers as occurring randomly due to some natural failures. The goal of robust methods is to fit a *reasonable* model to the majority of the data and tolerate some departures from the strict assumed parametric model without compromising the estimator performance.

Recently, several papers have raised the issue of cyber-attacks and how SE would be impacted by such disturbances [6]–[9]. In this case, the bad data present in certain meters is not accidental but generated by an intruder or attacker and coordinated in a fashion to avoid detection. Cyber-attacks can be viewed as special cases of outliers whose treatment is complex and of great importance. Improving cyber-security can be very challenging especially if attacks are generated by adversaries that have good knowledge of SE and robust methods used by the power operator. The knowledge and access degrees of attackers play a determining role in the detectability and possible actions against adverse disturbances. In [6], it is shown that an attacker can introduce a contamination called by false data injection attacks using the knowledge of the structure of the power system. Classical detection methods based on analyzing the residuals of the SE would not be able to detect such intrusions. In [7], the knowledge of the power system by the attacker was considered to be approximate and not exact which, in many cases, is more realistic. The worst configuration of attacks for the power

Manuscript received February 14, 2014; revised June 16, 2014, and August 14, 2014; accepted September 07, 2014. Date of publication October 13, 2014; date of current version July 17, 2015. This work was supported by Japan Science and Technology (JST) agency under the CREST program. Paper no. TPWRS-00225-2014

The authors are with the Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, Yokohama 226-8502, Japan (e-mail: yacine@sc.dis.titech.ac.jp, ishii@dis.titech.ac.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2014.2357182

operator was derived. The paper [8] generalizes the analysis to an alternative current (AC) SE. The properties of the AC representation present some advantages in detecting undetectable attacks over a simplified direct current (DC) representation. The paper [9] proposes a solution against stealthy attacks based on a Bayesian detection scheme. The approach exploits the distribution of the power state over time which is assumed to be multivariate Gaussian. Recently, [10] proposed an attack configuration on the measurement in the observation vector and the circuit breakers status readings to contaminate the estimated topology at the topology processor. Counter-measures have been proposed against such stealthy attacks by finding a set of sensors (PMUs) to be secured [11]. Reference [12] has studied the impact of cyber-attacks on power markets due to errors in the state estimates and topology. More generally, the topic of cyber-security is of interest in several research communities such as control, signal processing, communication and power systems [13]–[15].

In this paper, we are interested in possible attacks on the Jacobian matrix during SE and their impact on highly robust estimators resistant towards leverage points such as the least trimmed squares estimator (LTS). Even if the observations are clean from any modification, attacks in the Jacobian matrix can give rise to leverage points which are very challenging to treat in practice even when generated randomly. Scenarios that create an undetectable stealthy contamination are dangerous for robust SE. Stealthy attacks are derived theoretically for this context. Masked attacks are defined as well. The detection capabilities of the most popular diagnostic approaches used in power systems are studied theoretically and numerically.

The paper is organized as follows. Section II describes the power SE in the presence of outliers and coordinated cyber-attacks. Section III develops stealthy cyber-attacks on the Jacobian matrix considering leverage points-robust estimators. Conditions are derived to optimize attacks from the intruders' point of view on both AC and DC formulations. Masked attacks are defined. Robust regression equivariant estimators are investigated. Remedial actions are proposed as well. In Section IV, simulation results are run on the IEEE 14- and 30-bus test beds comparing several diagnostic methods. A variant of the least trimmed squares (LTS) with decomposition is proposed as a reference robust method towards leverage points. Both randomly generated outliers and coordinated cyber-intrusions scenarios are assessed. Finally, Section V concludes the paper.

II. State Estimation in the Presence of Cyber Attacks and Outliers

In this section, an overview is presented about the literature dealing with coordinated cyber attacks and random outliers in the SE context. A variant of least trimmed squares with decomposition is finally proposed as a benchmark evaluation method that is robust against leverage points.

The goal of SE is to estimate the bus voltage magnitudes and phase angles considered in the n-dimensional state vector $\boldsymbol{x}$ from the observation m-dimensional $\boldsymbol{z}(m > n)$ given by

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{e} \tag{1}$$

TABLE I
State Estimation Algorithm

- Start with $\hat{\boldsymbol{x}}^0$ such as the flat start or previously estimated state.
- Estimate $\hat{\boldsymbol{x}}^{k+1}$ by solving

$$\hat{\boldsymbol{x}}^{k+1} = \hat{\boldsymbol{x}}^k + \Delta\boldsymbol{x}^k$$

$$\Delta\boldsymbol{x}^k = G^{-1}(\hat{\boldsymbol{x}}^k)H^T(\hat{\boldsymbol{x}}^k)R^{-1}\left(\boldsymbol{z} - \boldsymbol{h}(\hat{\boldsymbol{x}}^k)\right)$$

where the gain matrix is $G(\hat{\boldsymbol{x}}^k) = H^T(\hat{\boldsymbol{x}}^k)R^{-1}H(\hat{\boldsymbol{x}}^k)$.

- Iterate by incrementing $k = k + 1$ until $\|\hat{\boldsymbol{x}}^{k+1} - \hat{\boldsymbol{x}}^k\| < \delta$.

where $\boldsymbol{h}(\cdot)$ is a nonlinear m-dimensional vector function, the error vector $\boldsymbol{e} \in \mathbb{R}^m$ is assumed to be Gaussian with zero mean $\mathbb{E}(\boldsymbol{e}) = \boldsymbol{0}$ and a covariance matrix $\mathbb{E}[\boldsymbol{e}\boldsymbol{e}^T] = R$. Boldface refers to vectors. The vector $\boldsymbol{z}$ can contain both SCADAs and PMUs. The nonlinear function $\boldsymbol{h}(\cdot)$ is linearized and solved iteratively using the weighted least squares (WLS), which is optimal under Gaussian noise. Table I illustrates the algorithm [2], where H is the Jacobian matrix of $\boldsymbol{h}(\cdot)$; δ and $\|\cdot\|$ are the chosen threshold and norm such as the infinite or Euclidean norm. The statistical model used in SE is evaluated and updated by the topology processor which tracks the system topology change over time.

A. Stealthy Cyber Attacks Introduced in Observations for SE

Several papers in the cyber-security literature considered stealthy attacks generated in the observation vector $\boldsymbol{z}$. The paper [6] discussed the possibility of stealthy attacks that remain undetected by any usual detection approach based on analyzing the residuals. An adverse can, for example, add a vector $\boldsymbol{a}$ to the observations contained in $\boldsymbol{z}$ to contaminate the estimated state $\hat{\boldsymbol{x}}$. This can be proven as follows. After neglecting the iterations effect (asymptotic case), consider the linearized regression model

$$\boldsymbol{z} = H\boldsymbol{x} + \boldsymbol{e} \tag{2}$$

where, for simplicity, the error covariance is assumed to be the identity matrix, i.e., $\mathrm{Cov}(\boldsymbol{e}) = I$. A simple matrix transformation of the data insures this condition. Assuming that the system is observable means that H is full rank, i.e., $\mathrm{rank}(H) = n$. The LS estimator is given by $\hat{\boldsymbol{x}} = (H^TH)^{-1}H^T\boldsymbol{z}$. Furthermore, the vector of residuals is simply

$$\boldsymbol{r} = \boldsymbol{z} - H\hat{\boldsymbol{x}} = \boldsymbol{z} - H(H^TH)^{-1}H^T\boldsymbol{z} = S\boldsymbol{z}$$

where $S = I - H(H^TH)^{-1}H^T$, and the matrix S corresponds to the projector on $\mathrm{Ker}(H^T)$ with $\dim(\mathrm{Ker}(H^T)) = m-n$ and $\dim(\mathrm{Im}(H)) = \mathrm{rank}(H) = n$; S is also known as the residual sensitivity matrix in power systems. If the observation vector is contaminated by an attack vector $\boldsymbol{a}$, the residual is given by

$$\boldsymbol{r}_c = S\boldsymbol{z}_c = S\boldsymbol{z} + S\boldsymbol{a} = \boldsymbol{r} + \boldsymbol{r}_a. \tag{3}$$

Note that the residuals vector in the absence of contamination is given by $\boldsymbol{r} = S\boldsymbol{z} = S\boldsymbol{e}$. If $\boldsymbol{a} = H\boldsymbol{c}$, the attack is in the space $\mathrm{Im}(H)$ which is orthogonal to $\mathrm{Ker}(H^T)$. This implies that $\boldsymbol{r}_a = S\boldsymbol{a} = \boldsymbol{0}$ with an estimated state of $\boldsymbol{x} + \boldsymbol{c}$, i.e., the state is contaminated by $\boldsymbol{c}$ with the estimated residuals kept the same. This is known as a *stealthy attack* since the detection

risk does not increase as studied in [6], [7]. Recently, the authors in [10] proposed to attack the observation vector $\boldsymbol{z}$ and the circuit breakers status readings to mislead the topology processor, which results in perturbing the updates of the topology impacting $\boldsymbol{h}$ and H.

In this paper, we consider attacks on the Jacobian matrix H. There are indeed possibilities of such attacks indirectly introduced, for example, by attacking the network topology estimates completed in the topology processor. It has been shown that coordinated attacks can disturb the updating of the grid topology [10]. More specifically, with the increased communication for real time operation and the often basic communication authentication in power systems that reduces time delays, an attacker can access the cyber-system at RTUs (Remote Terminal Units) and local data concentrators as a legitimate user. In [16], such access was considered to operate a *man-in-the middle* attack on the topology. Nowadays, energy management systems (EMS) are equipped with topology error, bad data and parameter errors detectors [2]. If the operator notices inconsistencies between estimated network topology (from breaker states and switches that are binary data) and meter data (power flows and injections, voltage magnitudes), he takes remedial actions to identify and correct the errors. This is effective especially if errors are randomly generated. An attack that changes both some breaker states and their corresponding meter data might be skipped by the bad data detection. In this case, an operator might validate a fictitious topology change or in the opposite does not consider a real topology change. This would generate undetected attacks on the topology and Jacobian. The impact of such attacks on time locational marginal price (LMP), for example, was recently studied [12], [16].

Another alternative is to attack the line parameter values directly. This option seems to be quite challenging in practice and its likelihood very low. However, since future power systems will be more automatized and considering the huge financial and security involvements, even internal attacks from within the control center should be considered and studied. Notice that in some cases, computers might be hacked or accessed in the control center [15]. For an AC model, an attack on the observations has an impact on the state estimates which will contaminate the Jacobian that depends on those estimates implying the possible presence of leverage points. Further experience in future practice might reveal further contamination possibilities.

B. Treatment of Random Outliers in the Jacobian Matrix H

More generally, if the observations obey exactly the statistical model in (1), the WLS offers the optimality of estimates. However, if a few outliers are present, the estimation performance can suffer severe degradation, i.e., an increased bias and inflated covariance matrix of the estimator.

The linearized regression model presented in (2) can be seen as fitting a hyperplane in an $(n+1)$-dimensional space containing m points $(z_i, \boldsymbol{H}_{i.})$, $i = 1, \ldots, m$, where $\boldsymbol{H}_{i.}$ is the ith row of the matrix H. An outlier is a point that does not follow exactly the model in (2). For example, it can obey a second model such as $\boldsymbol{z} = H\boldsymbol{x} + \boldsymbol{e} + \boldsymbol{b}$ or $\boldsymbol{z} = \boldsymbol{c}$ with $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^m$ or any different model. The clean points depart from the true hyperplane in the direction of the y-axis due to the Gaussian error. If a point $(z_i, \boldsymbol{H}_{i.})$ is very far (measured by the Euclidean distance of the residual which is much larger than $3\sqrt{R(i,i)}$) from the hyperplane in the direction of the y-axis, then this is said to be a y-axis outlier. Now if the n-dimensional space containing regressors, i.e., $\boldsymbol{H}_{i.}, i = 1, \ldots, m$, is analyzed and one $\boldsymbol{H}_{k.}$ is far from the bulk of the other $\boldsymbol{H}_{i.}$'s, then this is known as a leverage point. It can be a good leverage point if $\boldsymbol{z}_i$ corresponds to the true clean model or bad if $\boldsymbol{z}_i$ does not really obey the true model. Leverage points will have very high impacts on the estimation. Bad leverage points play a major adverse effect on the estimation and should be treated. Recent reviews of robustness concepts are available in more details in [17], [18]. Adaptation of some of these concepts to the power system SE is available in [19].

We conclude that outliers in power systems can be classified into coordinated cyber-attacks that can escape detection or be transparent to residuals analysis and random outliers such as errors and failures and random malicious actions. The most popular diagnostic approaches in power systems are based on analyzing the residuals of the WLS. Namely, the *chi-square* χ^2 *test* and *largest normalized residual test* [2] are used very actively. In the presence of leverage points, residuals obtained from more sophisticated methods should be analyzed [2]–[5]. The paper [19] applies the least median of squares (LMS), which is robust against leverage points, to the power systems SE context. The LMS minimizes the sample median of squared residuals. There, the authors propose to tune the algorithm in order to maximize the breakdown point (ε^*) taking into consideration a measure of redundancy (s^*). The breakdown point (ε^*) is the maximum fraction of outliers an estimator can resist while offering reliable estimates (before breaking down) [17], [18]. The redundancy s^* is the minimum number of measurements that, if removed, leave at least one remaining critical measurement. Disregarding a critical measurement makes the system unobservable [19].

Since power systems are very sparse, i.e., there could be a bus with very low redundancy, this would impose a constraint on the breakdown point (ε^*) of robust methods for large systems by making it very low. The solution would be to decompose the system into islands or subsystems with several ε^* [19]. Islands with high redundancy would have a high ε^* in the opposite to subsystems with low redundancy. This would augment the number of outliers treated by the algorithm and reduce the computation time especially if the latter is conducted in parallel. More specifically, the work [19] proposed a simplified systematic decomposition approach in two types, namely, radial and cyclic subsystems. A radial island is defined as a subset of buses and related measurements between lines where if one line is cut, the system is disconnected. It generally contains at least one bus with one connection line. The reference also considers minimal cycles or loops connecting buses that do not contain any sub-cycles (See Fig. 1).

A few approaches were proposed to detect leverage points in the power systems literature such as the projection statistics (PS) algorithm [5]. Down-weighting leverage points in the H matrix was also proposed [2] based on *matrix stretching*.

Both the LTS and LMS with decomposition are discussed in [4]. In [20], the least trimmed squares (LTS) was implemented

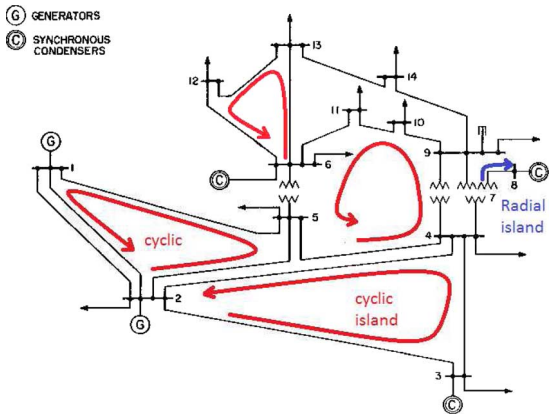


Fig. 1. Decomposition of the IEEE 14 bus system in subsystems for diagnostic (8 subsystems, 1 radial and 7 minimal cycles).

to estimate the power state. The decomposition in several subsystems was not considered. The least trimmed squares minimizes a trimmed percentage of the regression squared residuals [17]. If the squared residuals are sorted from the smallest to the largest $r^2_{(1)} \leq \cdots \leq r^2_{(m)}$, the LTS estimates the parameter $\boldsymbol{x}$ that minimizes the cost function

$$J(\boldsymbol{x}) = \sum_{i=1}^{\lfloor (1-\alpha)m \rfloor + 1} r^2_{(i)}(\boldsymbol{x}),\ \alpha \in [0,1) \tag{4}$$

where $\lfloor \cdot \rfloor$ is the floor function and α corresponds to the breakdown point (ε). In this paper, a variant of the LTS is used after decomposing the power system which is robust against leverage points. The number of treated outliers is maximized thanks to the decomposition. The operator is assumed to be very cautious where an observation is flagged as oultying even if detected in a single island. The algorithm is outlined in Table II and explained as follows.

First decompose the power system in radial and elementary cyclic sub-systems (N islands) as illustrated in Fig. 1. Run the SE estimation in each subsystem to diagnose outliers. Adapting power injections means simply that an injection at a given bus is corrected by removing the flows at cut-lines connected to the same bus. An example is provided in Section III-C where if the island {1,2,5} is considered (See Fig. 1) then the corrected injection is P_2^n. The weights are adapted as well where the corresponding new injection error covariance is $R_{ii} + R_{jj} + R_{kk}$; R_{ii}, R_{jj} and R_{kk} are the error covariances corresponding to the injection at bus 2 and the flows between buses {2-3} and {2-4}. The vector $\delta \boldsymbol{x}_i^k \in \mathbb{R}^{n_i}$, where n_i is the dimension of the state to be estimated in the ith island, is obtained from

$$\delta \boldsymbol{x}_i^k = \arg \min_{\Delta \boldsymbol{x}} \sum_{j=1}^{\lfloor m_i(1-\alpha_i) \rfloor + 1} \left(\hat{\boldsymbol{r}}_i^k(j) - \boldsymbol{H}_{j.}^{(i)} \Delta \boldsymbol{x} \right)^2$$

where m_i is the number of available measurements at the ith island, α_i is the trimming parameter and $\hat{\boldsymbol{r}}_i^k = \boldsymbol{z}^{(i)} - \boldsymbol{h}(\hat{\boldsymbol{x}}_i^k)$ is the ith island vector of residuals. The ith starting state is $\hat{\boldsymbol{x}}_i^1 = \mathbf{1} \angle \mathbf{0}$, i.e., flat start. $\boldsymbol{z}^{(i)}$ is the vector of observations in the ith island. Practically, the quantity $\delta \boldsymbol{x}_i^k$ is computed as explained in Table II. Reject the large residuals, i.e., large $|\hat{\boldsymbol{r}}_i^k(j) - \boldsymbol{H}_{j.}^{(i)} \delta \boldsymbol{x}_i^k|, j = 1, \ldots, m_i$. Normalized residuals can be compared to a chosen constant which was fixed to 7 in this practical implementation. It was noticed that the limited clean data is approximately Gaussian with slightly larger tails. The regression is re-evaluated by computing $\Delta \boldsymbol{x}_i^k$ to improve the efficiency after disregarding the detected outliers as

$$\Delta \boldsymbol{x}_i^k = \arg \min_{\Delta \boldsymbol{x}} \sum_{i=1}^{m_i^c} \left(\hat{\boldsymbol{r}}_{i,c}^k(j) - \boldsymbol{H}_{j.}^{(i)} \Delta \boldsymbol{x} \right)^2$$

where $m_i^c = m_i - n_o^k$ and n_o^k is the number of detected outliers at the kth iteration. Removing the elements corresponding to outliers in $\hat{\boldsymbol{r}}_i^k$ gives $\hat{\boldsymbol{r}}_{i,c}^k$. The state and residuals are updated, i.e., $\hat{\boldsymbol{x}}_i^{k+1}$ and $\hat{\boldsymbol{r}}^{k+1}$ are computed. The iterations are stopped once $\|\Delta \boldsymbol{x}_i^k\| < \delta$ and the set containing the detected outlier at the ith subsystem is $\mathcal{J}_i = \cup \mathcal{J}_i^k$, where $\mathcal{J}_i^k$ contains the flagged outliers at the kth iteration and $\cup$ is the union operator. Finally, the set containing all detected outliers is obtained by $\mathcal{J} = \cup \mathcal{J}_i, i = 1, \ldots, N$. Notice that a detected contaminated injection can be due to the injection itself or to the flows on the cut-lines. In this variant, if the outlier is due to the initial injection, then only

TABLE II
ALGORITHM FOR LEAST TRIMMED SQUARES (LTS) WITH DECOMPOSITION

- Decompose the system in several (N) minimal cyclic and radial subsystems to maximize the total number of outliers possibly treated [19].
  - A Radial subsystem consists of buses, lines, and related measurements such that if one line is cut, the subsystem is disconnected. It generally contains at least one bus with one connecting line.
  - A Minimal cycle subsystem is formed by buses forming a loop that does not contain any subloops.
- For each island diagnose outliers using the LTS-based SE.
  - Create all measurements (adapting power injections) and $H^{(i)}$, $\boldsymbol{h}^{(i)}$ for the i^{th} island by considering removed cut-lines.
  - If a PMU is absent choose a slack bus randomly in the subsystem.
  - Replace the WLS by LTS in each iteration to compute $\delta \boldsymbol{x}_i^k$. Each step is executed as follows.
  - Consider $n_s^{(i)}$ (=100) random samples of n rows of $H^{(i)}$ to get $H_s^{(i)}$, each sample ensures $\operatorname{rank}\left(H_s^{(i)}\right) = n_i$. (A heuristic search is conducted to avoid non-observable sets.)
  - For each sample run WLS on $\hat{\boldsymbol{r}}_s^k, H_s^{(i)}$, sort the residuals i.e. absolute elements of $\hat{\boldsymbol{r}}_s^k - H_s^{(i)} \delta \boldsymbol{x}_i^s$, and reject the α_i-largest fraction.
  - Re-estimate $\delta \boldsymbol{x}_i^s$ and repeat the last task n_u (=10) times. Finally, choose the estimates $(\delta \boldsymbol{x}_i^k)$ that minimize the residuals robust standard deviation $\hat{\sigma}_r\left(\hat{\boldsymbol{r}}_s^k - H_s^{(i)} \delta \boldsymbol{x}_i^s\right)$ among the $n_s^{(i)} = 100$ samples i.e.
    $$\delta \boldsymbol{x}_i^k = \arg \min_{\delta \boldsymbol{x}_i^s} \hat{\sigma}_r\left(\hat{\boldsymbol{r}}_s^k - H_s^{(i)} \delta \boldsymbol{x}_i^s\right)$$
  - Efficiency is improved by diagnosing residuals i.e. vector $\hat{\boldsymbol{r}}_i^k - H^{(i)} \delta \boldsymbol{x}_i^k$, detecting and removing outliers (their positions stored in $\mathcal{J}_i^k$) and computing $\Delta \boldsymbol{x}_i^k$ as in step 2 of Table I.
  - Compute $\hat{\boldsymbol{x}}_i^{k+1} = \hat{\boldsymbol{x}}_i^k + \Delta \boldsymbol{x}_i^k$ and $\hat{\boldsymbol{r}}_i^{k+1} = \boldsymbol{z}^{(i)} - \boldsymbol{h}^{(i)}\left(\hat{\boldsymbol{x}}_i^{k+1}\right)$.
  - Stop iterations for i^{th} island if $\|\Delta \boldsymbol{x}_i^k\| < \delta$.
  - Detected outliers positions at the i^{th} island stored in $\mathcal{J}_i = \cup \mathcal{J}_i^k$
- Run the SE algorithm on the whole system after removing the outliers detected in each island (i.e. $\mathcal{J} = \cup \mathcal{J}_i$) (See Table I).

the injection is flagged. If it is due to a cut-line flow, then both the flow and the injection are rejected. This variant is very safe where the union of all detected outliers is removed.

Notice that even leverage points resistant estimators such as LTS can not deal with stealthy attacks on the whole observation vector [6].

III. Stealthy Attacks on the Measurement Function

As mentioned above, we consider an attack configuration where the Jacobian is modified. This would create what is known as leverage points that can generate extra-complication for methods that analyze WLS residuals, such as the largest normalized residual, even when the residual caused by the attack is non-null ($\boldsymbol{r}_a \neq \mathbf{0}$) in (3). The point is that the largest residuals, in this case, do not necessarily correspond to the outliers (i.e., leverage points). This is called the *masking effect* in robust statistics literature and analyzing the residuals of the WLS is not really a good approach [17]. Considering this fact can motivate an intruder to attack H instead of the observation vector $\boldsymbol{z}$.

We consider two scenarios with different convergence of the SE following attacks.

- Scenario 1: convergence to an arbitrary state unknown to the attacker.
- Scenario 2: convergence to a state targeted by the attacker.

Another situation is where the attack causes divergence of the SE, which is not stealthy but still problematic to the operator. We now discuss the two scenarios separately in the following.

Scenario 1 can be achieved by generating more outliers than the breakdown point. If sophisticated methods robust towards leverage points are used, the intruder has to go beyond their limits. If the estimator breaks down, it can be very difficult to detect attacks and their positions post-estimation based on the residuals. Indeed, a possibly important bias is the consequence of the breakdown of a robust estimator. This implies the irrelevance of estimated residuals in order to detect the intrusions (masking effect). Since the WLS has a null-breakdown point in the presence of leverage points ($\varepsilon^* = 0$) [17], it can not be trusted to handle even one intrusion in H. We call these attacks masked attacks.

If the LTS with decomposition is considered, a possible approach for the attacker would be to detect the weakest points in the grid. The weak points are elements in the Jacobian linked to buses with low redundancy limiting the feasible breakdown point of the robust approach such as the LTS. Indeed, the breakdown point is bounded by the redundancy as in $\varepsilon_i^* < s_i^*/2\ell_i$ [19] where ℓ_i and s_i^* are the number of estimation points considered and the redundancy in the ith island respectively. This means that the weak points are found in the islands with minimum $n_{c,i}$; $n_{c,i}$ is the minimum number of contaminated rows of H for the ith subsystem that generates a masked attack and is given by

$$n_{c,i}^{(1)} = \lceil \ell_i \varepsilon_i^* \rceil \tag{5}$$

where $\lceil \cdot \rceil$ is the ceiling function. The breakdown point (ε^*) for a reasonable robust method is inferior to half. If the attacker has access to more than $\ell_i/2$ points of H in the concerned island, then there is no need to check for redundancy.

A. Coordinated Stealthy Attacks on the Jacobian Matrix

In scenario 2, a precise state is targeted to reach a predetermined contaminated state which can be useful for an intruder with possible financial gains, for example, after impacting the electricity markets. Creating an attack that does not disturb the residuals from WLS, which is very sensitive to leverage points and optimal in the outlier-free case, means that the attack obeys the regression structure. The development provides a very stealthy disturbance for a large spectrum of robust estimators. Diagnostic approaches are generally based on a majority of residuals which are obeying the true model and a minority of deviant outliers. If all residuals are not impacted and behaving the same, detecting the attack can not be achieved. This philosophy is followed to study the feasibility of stealthy attacks against leverage points robust estimators.

The contaminated WLS estimate is given by

$$\hat{\boldsymbol{x}}_c = \left(H_c^T H_c\right)^{-1} H_c^T \boldsymbol{z}$$

where H_c is the contaminated Jacobian matrix given by $H_c = H + \delta H$ with an additive attack δH on the Jacobian matrix. Notice that δH is an introduced attack and not an uncertainty in the knowledge of H. The state $\hat{\boldsymbol{x}}_c$ is obtained instead of $\hat{\boldsymbol{x}} = (H^T H)^{-1} H^T \boldsymbol{z}$. The observation vector $\boldsymbol{z}$ is assumed to be clean from contamination by default but can contain outliers. Similarly, the contaminated residuals are given by

$$\begin{aligned}\boldsymbol{r}_c &= \boldsymbol{z} - H_c \hat{\boldsymbol{x}}_c = \boldsymbol{z} - H_c (H_c^T H_c)^{-1} H_c^T \boldsymbol{z} = S_c \boldsymbol{z} \\ &= S_c(H\boldsymbol{x} + \boldsymbol{e}) = S_c(H_c \boldsymbol{x} - \delta H \boldsymbol{x} + \boldsymbol{e}) \\ &= S_c \boldsymbol{e} - S_c \delta H \boldsymbol{x}.\end{aligned}$$

The matrix S_c is a projector on $\mathrm{Ker}(H_c^T)$; since $\boldsymbol{e} \sim \mathcal{N}(\mathbf{0}, I)$ and $S_c^T S_c = S_c, \boldsymbol{r}_c \sim \mathcal{N}(\boldsymbol{b}_r, S_c)$, the created bias is $\boldsymbol{b}_r = -S_c \delta H \boldsymbol{x}$. In the absence of contamination, $\boldsymbol{r} \sim \mathcal{N}(\mathbf{0}, S)$. In this case, both the mean and the variance of the residuals vector are impacted by the attack, which was not the case with the previous attack configuration (i.e., an attack on $\boldsymbol{z}$ only).

In this scenario, the objective of the attacker is to reduce any detection post-estimation and ideally to eliminate it by keeping the residuals the same. This can be achieved by making the residuals bias $\boldsymbol{b}_r$ zero. Indeed, the distribution function of the residuals being symmetric (Gaussian distribution), reducing the bias in absolute value will reduce the chance to be detected.

Notice also that minimizing the bias in the residuals is crucial to mask an intrusion since it is known and expected by the operator to be null in the clean case (assuming $\mathbb{E}(\boldsymbol{e}) = \mathbf{0}$). The variance S_c, on the other hand, is estimated or computed at a given snapshot and it is difficult for the operator to check whether an attack occurred based on an inflated variance S_c. As an example, the normalized residual, which is expected to follow a standard Gaussian, is given by $r_i^n = r_i/\sqrt{S_c(i,i)}$. The covariance S_c is used instead of S by the operator since the Jacobian is contaminated. This means that $r_i^n \sim \mathcal{N}(\boldsymbol{b}_r(i)/\sqrt{S_c(i,i)}, 1)$ and the attack's impact can be detected from the bias only. The bias is generally much larger than the variance.

To reduce the bias, it is useful to introduce the projector on the image of δH : $\delta P = \delta H(\delta H^T \delta H)^{-1}\delta H^T$ where $\delta H = \delta P \delta H$. The Euclidean norm of the bias is given by

$$\|\boldsymbol{b}_r\| = \|S_c \delta H \boldsymbol{x}\| = \|S_c \delta P\| \|\delta H \boldsymbol{x}\|.$$

Using the definition of smallest principal angle between two spaces and its link to two orthogonal projectors as defined in [7] (*Definition 1 and Lemma 1*), we have

$$\|\boldsymbol{b}_r\| = \cos(\gamma)\|\delta H \boldsymbol{x}\|$$

where γ is the smallest principal angle between the spaces of the projectors δP and S_c, which are respectively $\mathrm{Im}(\delta H)$ and $\mathrm{Ker}(H_c^T) = \mathrm{Ker}((H+\delta H)^T)$. The term $\|\delta H \boldsymbol{x}\|$ is not available since it depends on the unknown true state. Even in the case where an estimate of $\boldsymbol{x}$ is available to the intruder by accessing PMUs for example, the term $\|\delta H \boldsymbol{x}\|$ should not be minimized since it reflects the impact of the attack on the estimation bias or the committed error on estimating the state. Indeed, the estimation is given by

$$\begin{aligned}\|\boldsymbol{b}_e\| = \|\boldsymbol{x} - \mathbb{E}(\hat{\boldsymbol{x}}_c)\| &= \left\| \left(H_c^T H_c\right)^{-1} H_c^T \delta H \boldsymbol{x} \right\| \\ &\leq \left\| \left(H_c^T H_c\right)^{-1} H_c^T \right\| \|\delta H \boldsymbol{x}\|.\end{aligned}$$

Decreasing $\|\delta H \boldsymbol{x}\|$ to zero reduces the estimation bias which should be avoided. For example, $\boldsymbol{x} \in \mathrm{Ker}(\delta H^T) \Rightarrow \|\boldsymbol{b}_e\| = 0$. Hence, the term that can be manipulated by an attacker to reduce $\|\boldsymbol{b}_r\|$ is $\|S_c \delta P\|$, i.e., γ should be made as near as possible to $\pi/2$. This implies that $\mathrm{Ker}(H^T + \delta H^T)$ is as closely orthogonal as possible to $\mathrm{Im}(\delta H)$ or, ideally, both spaces are orthogonal. This condition is satisfied if, for example, δH has an included or the same image space as H, i.e.,

$$\mathrm{Im}(\delta H) \subseteq \mathrm{Im}(H). \tag{6}$$

The diagnostic post-estimation of residuals where the WLS is executed will clearly not flag such an attack.

Moreover, the LTS, which is robust against leverage points as discussed earlier, will not be able to detect such an intrusion either. This is because the LTS can be seen as an orthogonal projection problem on a space of $\mathrm{Im}(H_n)$. Indeed, the LTS is equivalent to the WLS applied to a *clean* set containing a reduced number of points $(\lfloor(1-\alpha)m\rfloor + 1)$ having the smallest squared residuals and contained in H_n by keeping the rows of H corresponding to the clean set, i.e., $\mathrm{Im}(\delta H_n) \subseteq \mathrm{Im}(H_n)$.

Notice that these attacks can be seen theoretically as a contamination in $\boldsymbol{z}$ of a special kind even if, in practice, the robust estimation tools needed are very different to handle both situations. Asymptotically, the estimation problem is equivalent to $\boldsymbol{z} = (H + \delta H)\boldsymbol{x}_c + \boldsymbol{e}$. An attack can be seen as modifying the observation $\boldsymbol{z}$ by $-\delta H \boldsymbol{x}_c$. Such transformation will change the problem into an estimation with the regressors matrix H, i.e., $\widetilde{\boldsymbol{z}} = H\boldsymbol{x}_c + \boldsymbol{e}$. However, the introduced attack is linked to the obtained contaminated $\boldsymbol{x}_c$ which depends on the true state and δH as follows: $\boldsymbol{x}_c = (H_c^T H_c)^{-1} H_c^T H \boldsymbol{x}$. This can be viewed as a special constraint. The condition of an absolute stealthy attack when analyzing the residual consists in the same spirit, i.e., to create an additive element in the equation that is in the space spanned by columns of H to avoid inducing a bias in the residuals. It joins also the condition derived in [10], where $\boldsymbol{z}$ is attacked to modify $\boldsymbol{h}$ in the topology processor.

We believe important to mention that other robust approaches such as LMS and LAV do not have necessarily a projection interpretation. Considering this makes the problem seem complex theoretically. However, several of the developed robust methods insure *regression equivariance*: An estimator $\hat{\boldsymbol{x}}(\boldsymbol{z}, H)$ satisfies this property if [17]

$$\hat{\boldsymbol{x}}(\boldsymbol{z} + H\boldsymbol{\nu}, H) = \hat{\boldsymbol{x}}(\boldsymbol{z}, H) + \boldsymbol{\nu},\ \forall\, \boldsymbol{\nu} \in \mathbb{R}^n.$$

Regression equivariance by definition means that the estimator is not resistant against a linear translation on $\mathrm{Im}(H)$. In this case, this good property in the case of partly perturbed data can be exploited by the attacker. Thus, if a robust estimator is said to satisfy regression equivariance, then it does not stand the previously discussed attack (i.e., $\boldsymbol{r}_c = \boldsymbol{r}$).

Confidence intervals are also modified by the intruder following an attack, which is not the case for attacks on $\boldsymbol{z}$. Indeed, the estimates asymptotic covariance matrix is $\mathrm{Cov}(\hat{\boldsymbol{x}}_c) = (H_c^T H_c)^{-1}$ instead of $\mathrm{Cov}(\hat{\boldsymbol{x}}) = (H^T H)^{-1}$.

B. Scenario 2 Attacks on Practical Power Systems

There are a few important factors that might simplify further the task of an intruder when combined to (6). These factors are presented as follows.

As explained, the regression estimation considers m points in the $(n+1)$-dimensional space, i.e., $(z_i, \boldsymbol{H}_{i.})^T$ with $\boldsymbol{H}_{i.}$ being the ith row of the Jacobian H. A stealthy attack can be achieved as well if more row values in the Jacobian matrix than $n_c^{(2)}$ given by

$$n_c^{(2)} = \lfloor(1-\varepsilon^*)m\rfloor + 1 \tag{7}$$

are modified in a coordinated fashion around an erroneous state $\boldsymbol{x}_c$. Then the LTS (or any other robust estimator) will reject even the remaining good data detected as outliers since they are in minority. This is an interesting result since it shows that it is possible to transform the advantage provided by a high breakdown point robust estimator to a weak point. It allows the attacker to reduce further targeted positions in H.

Furthermore, since H is very sparse in real life power systems, the previous analysis can be completed on a subsystem of the power grid. Generating the previously derived condition can be completed by an attacker, to contaminate a few bus states of interest. A null bias for the observation residuals linked to these targeted states would be insured. This means that if (2) can be rewritten as follows:

$$\boldsymbol{z} = H(\boldsymbol{x}_{\mathcal{L}}^T\ \boldsymbol{x}_r^T)^T + \boldsymbol{e}$$

with

$$\boldsymbol{x} = \left(\boldsymbol{x}_{\mathcal{L}}^T\ \boldsymbol{x}_r^T\right)^T; \quad H = \begin{pmatrix} H_{\mathcal{L}} & H_{\mathcal{L},r} \\ \mathbf{0} & H_r \end{pmatrix}$$

then satisfying the condition (6) to contaminate $\hat{\boldsymbol{x}}_{\mathcal{L}}$ is equivalent to ensuring

$$\mathrm{Im}(\delta H_{\mathcal{L}}) \subseteq \mathrm{Im}(H_{\mathcal{L}}); \quad \delta H = \begin{pmatrix} \delta H_{\mathcal{L}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{8}$$

In order to contaminate $\boldsymbol{x}_{\mathcal{L}}$, $n_c^{(2)}$ points among all observations linked to this state portion are modified by the stealthy intrusion. If a decomposition of the power system is conducted and depending on the decision process followed by the operator to classify whether an observation is an outlier or not, the number of attacked elements can be reduced. If, for example, the operator is not very safe by flagging an overlapping observation as outlying only if detected in all islands, the attacker can reduce $n_c^{(2)}$. This can be achieved by considering the observations in one subsystem only (i.e., $n_{c,i}^{(2)}$) instead of modifying all observations linked to the bus state. In this case, ℓ_i and s_i^* are quantities corresponding to one island. Furthermore, the island with minimum redundancy (i.e., s_i^*) is a good candidate, if accessed, to reduce $n_c^{(2)}$.

An intruder might prefer scenario 1 over 2 since it needs less contaminated positions (breakdown is inferior to half). However, scenario 2 has the advantage of controlling the modified states and being stealthy.

To illustrate the development above, we consider a stealthy attack on $\hat{\boldsymbol{x}}(1)$. For (8) to be satisfied $\delta H_{\mathcal{L}}$ should belong to the space generated by the first column $\boldsymbol{H}_{.1}$, which implies that $\delta H_{\mathcal{L}} = (\eta - 1)\boldsymbol{H}_{.1}, \eta \in \mathbb{R}$. This means that the number of attacked elements in H should be equal to the number of non-zero elements in $\boldsymbol{H}_{.1}$ denoted by ℓ. Then, the obtained new state is $\boldsymbol{x}_c(1) = \boldsymbol{x}(1)/\eta$. A weaker constraint would be to generate $\delta H_{\mathcal{L}}^n = (\eta - 1)\boldsymbol{H}_{.1}^n$, where $\boldsymbol{H}_{.1}^n$ is constituted by any $n_c^{(2)}, n_c^{(2)} = \lfloor (1 - \varepsilon^*)\ell \rfloor + 1 < \ell$ dimensional set of elements from $\boldsymbol{H}_{.1}$. If the attacker accesses the breakdown point ε^*, then the number of necessary attacked positions $n_c^{(2)}$ to generate a stealthy intrusion is known ($n_c^{(2)} = \lfloor (1 - \varepsilon^*)\ell \rfloor + 1$). If the breakdown point is unknown, the attacker can still estimate an interval for $n_c^{(2)}$. Indeed, the breakdown point satisfies $0 < \varepsilon^* \leq s^*/2\ell$, where s^* is the redundancy measure of the island. Knowing ℓ and s^*, the attacker can find some interval where $n_c^{(2)}$ varies, i.e., $\lfloor \ell - s^*/2 \rfloor + 1 \leq n_c^{(2)} < \ell$.

This is particularly true for a DC formulation (linear regression problem). With an AC formulation, the linearization is not applied to updating the residuals (See Table II). Thus, there are two attack possibilities: 1) attacking the Jacobian matrix H without changing $\boldsymbol{h}$. 2) attacking $\boldsymbol{h}$ which implies also a contaminated H. If a stealthy attack is desired exactly on certain states without disturbing the rest of the buses, attacking H and modifying $\boldsymbol{h}$ by $\boldsymbol{h}_c(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x}) + \delta H \boldsymbol{x}$ was shown to be effective by simulation results (See Section IV).

If a stealthy attack is created where the observations are modified in order to make the estimate of θ_1 biased by a_1 (i.e., $\boldsymbol{z}_a = \boldsymbol{H}_{.1}a_1$), then the algorithm proposed in [8] flags the number of non-null elements in $\boldsymbol{H}_{.1}$. The same number of elements should be known in H to create $\boldsymbol{z}_a$. Since H for an AC SE depends on the state of the previous SE or even the previous iteration, such an attack requires quite a large access to the data of the system.

Both bus connection and redundancy in measurements play a determining role in the value taken by $n_c^{(2)}$. This information is included in H which means that accessing the Jacobian is important for generating stealthy attacks.

Furthermore, the attacker should target active power injections and flows if a phase angle is targeted. In this case, the n_c reduces where ℓ considers only the measurements linked to the phase angles. If voltage magnitudes are targeted, the reactive power should be attacked and $n_c^{(2)}$ reduces in the same fashion. This is important since the active power is preponderant in determining the phase angle estimates whereas reactive power is important for voltage magnitude.

C. Illustration on a Practical Power System

The previous discussion is illustrated further on the IEEE 14 bus system depicted in Fig. 1 where the decoupled SE combined to the LMS after decomposing the power system as in [19] is studied. The same results are obtained when considering the LTS with the same breakdown point. The obtained subsystems after decomposition [19] are, by bus number, grouped in

Cyclic islands $\{1,2,5\}, \{2,4,5\}, \{2,3,4\}, \{4,7,9\}$, $\{6,12,13\}, \{6,9,10,11,13,14\}, \{4,5,6,9,10,11\}$ and one radial island $\{7,8\}$.

The active power-phase angle iteration, for example, uses the Jacobian matrix given by [2]:

$$\frac{\partial P_i}{\partial \theta_i} = \sum_{j=1}^{n} V_i V_j \left[-G_{ij}\sin(\theta_{ij}) + B_{ij}\cos(\theta_{ij})\right] - V_i^2 B_{ii};$$

$$\frac{\partial P_{ij}}{\partial \theta_i} = V_i V_j [g_{ij}\sin(\theta_{ij}) - b_{ij}\cos(\theta_{ij})];$$

$$\frac{\partial P_{ij}}{\partial \theta_j} = -\frac{\partial P_{ij}}{\partial \theta_i};$$

where the parameters g_{ij}, G_{ij} and b_{ij}, B_{ij} reflect the system topology representing conductances and susceptances, i.e., system admittance. The state at bus i includes the voltage magnitude V_i, and phase angle θ_i. The difference between the phase angles at buses i and j is $\theta_{ij} = \theta_i - \theta_j$. The power flow between bus i and j is represented by P_{ij}, and P_i is the active power injection at bus i. The power values are measured.

The vulnerability of the cyclic subsystem composed by buses {1,2,5} can be analyzed as follows. The taken measurements are $P_{12}, P_{21}, P_{15}, P_{51}, P_{25}, P_{52}, P_1, P_2^n$ and P_5^n. Two new power injections are computed for this island after removing the power flows on the cut-lines, i.e., $P_2^n = P_2 - P_{24} - P_{23}$ and $P_5^n = P_5 - P_{56} - P_{54}$. The phase estimation problem at the island becomes $\boldsymbol{z}^{(1)} = H^{(1)}\boldsymbol{x}_1 + \boldsymbol{e}_1$ where $\boldsymbol{z}^{(1)}$ contains all the previous measurements, $\boldsymbol{x}_1 = (\theta_1, \theta_2, \theta_5)^T$, i.e., $m_1 = 9$ and $n_1 = 3$.

A scenario 2 attack is illustrated as follows. Notice that buses {2,5} are present in other subsystems whereas bus 1, which is connected to two buses only, is exclusively estimated in this subsystem. If an attacker targets θ_1, i.e., $\mathcal{L} = \{1\}$, the first column of H is analyzed showing seven non-null components. The breakdown point of the LMS is fixed to three. If four ($n_c = 7 - 3 = 4$) elements in H are modified as $\boldsymbol{\delta H}_{.1} = (\eta - 1)\boldsymbol{H}_{.1}$, then the contaminated estimate is $\theta_{1,c} = \theta_1/\eta$ (scenario 2, bus 1 is not the slack bus). A scenario 1 attack can be generated if four outliers are created randomly.

The first column of H shows interdependent elements. For example, $V_1V_2[g_{12}\sin(\theta_{12}) - b_{12}\cos(\theta_{12})]$ depends on the states at buses 1 and 2 as well as g_{12} and b_{12}. The element components are present in measurements P_{12}, P_{21}, P_2^n. Another contamination would completely break the LMS estimator even if the intrusion is completed in the observation vector $\boldsymbol{z}$.

To create a classical stealthy attack on observations, seven observations should be manipulated and seven should be known in H. Attacking H, in this example, reduces the minimum number of necessary targeted positions n_c.

D. Possible Remedial Actions for Power Operators

Estimators that analyze H independently of $\boldsymbol{z}$ can offer advantages in some cases. If some structure or confidence intervals are known about a few elements in H, this knowledge might provide a solution to stealthy intrusions targeting H. Such methods can not keep good leverage points in the estimation but a possible rejection of a good leverage point in order to insure robustness against cyber-attacks can constitute a good compromise. Combining two robust approaches and comparing their estimates might be a good solution.

Knowledge of the topology and placement of measurements by intruders is dangerous for power companies. Securing the topology processors and different communication channels is also of great importance. Furthermore, the topology can be masked in the cyber-information, which avoids finding weak points by an intruder. For example, keeping erroneous information that a bus is connected to more than one bus when in reality it is connected to a unique bus can help.

More specifically, this could be formulated as follows. If the knowledge of the Jacobian matrix H is approximative, i.e., $\widetilde{H} = H + \Delta H$, the attacker would disturb H by introducing δH while thinking that H is $\widetilde{H}$. The attacker action creates in reality $H_c = H + \delta H$. As in (6), the intruder would try to satisfy that $\mathrm{Ker}(\widetilde{H}^T + \delta H^T)$ is as closely orthogonal as possible to $\mathrm{Im}(\delta H)$. A solution would be to generate δH satisfying $\mathrm{Im}(\delta H) = \mathrm{Im}(\widetilde{H})$ implying that the real bias in Euclidean norm created by the attack is proportional to the cosine of the smallest principle angle between the spaces $\mathrm{Ker}(H_c^T) = \mathrm{Ker}(H^T + \delta H^T)$ and $\mathrm{Im}(\widetilde{H}) = \mathrm{Im}(H + \Delta H)$. A possible objective for the operator consists in creating an uncertainty ΔH expected to maximize this bias or to make $\mathrm{Im}(H + \Delta H)$ and $\mathrm{Ker}(H^T + \delta H^T)$ as closely parallel as possible. Such a solution was considered for attacked $\boldsymbol{z}$ [7].

The cyber-security is improved by increasing the total number of placed measurements. More specifically, PMUs can play a key role against attacks in the Jacobian. Indeed, it is known that the element in the Jacobian H corresponding to a PMU is fixed to 1. There is no real need to manipulate, compute or transmit this component which makes it safe against intrusion. A check that an element corresponding to a PMU in H is equal to 1 is a good way to detect an intrusion as well. A good approach to improve cyber-security could be to install PMUs at critical islands with low redundancy. Furthermore in [10], a placement of PMUs was proposed to secure topology processors against attacks. The authors assumed that the injections delivered from PMUs were safe.

IV. Simulation Results

We continue with the IEEE 14 bus system setting in Section III-C. The SE program available in MATPOWER [21] was used and modified to introduce the robust LTS with decomposition. An algorithm for multivariate LTS estimator is provided in [22] and adapted to handle sparsity in the grid (See Table II). The true state of the system $\boldsymbol{x}_T$ is known from the power flow solution. The slack bus was considered at bus 1 where voltage magnitude was forced to 1 p.u. and phase angle to zero. The system has been decomposed in eight subsystems. Each bus of the system is assumed to have a voltage magnitude measurement and both active and reactive power injections. Each line has two active and two reactive power flows at its ends. This is considered to be a high redundancy measurements configuration, which provides a benchmark study with a highly robust and rich case. Showing that even in this case attacks can be stealthy or masked to the operator emphasizes the danger of cyber-attacks. Indeed, the concern is increased since the redundancy is lower. The evaluation index of the estimation quality for the Monte Carlo simulation is

$$x_I^M = \frac{1}{M_c}\sum_{k=1}^{M_c}\left\|\hat{\boldsymbol{x}}^{[k]} - \boldsymbol{x}_T\right\| \tag{9}$$

where M_c is the number of considered Monte Carlo runs. The state $\hat{\boldsymbol{x}}^{[k]}$ is the estimate at the kth run and $\|\cdot\|$ is the Euclidean norm. Voltage phase angles and voltage magnitudes errors are evaluated separately, i.e., two indices are computed.

The number of introduced leverage points and observation outliers are n_l and n_z respectively. The number of detected leverage points that are truly present in the attack is denoted by n_T^l. The quantity n_T^z is its observation outliers counterpart. The number of outliers detected which are neither generated leverage points nor observation outliers is n_F. The estimated probabilities of detection for leverage points and observation outliers are: $P_d^l = \sum_{k=1}^{M_c} n_{T,k}^l/((n_{T,k}^l + n_{F,k})M_c), P_d^z = \sum_{k=1}^{M_c} n_{T,k}^z/((n_{T,k}^z + n_{F,k})M_c)$. The detection indices $d_I^l = \sum_{k=1}^{M_c} n_{T,k}^l/(n_l M_c), d_I^z = \sum_{k=1}^{M_c} n_{T,k}^z/(n_z M_c)$. The estimated probability of false detection is defined by $P_f = \sum_{k=1}^{M_c} n_{F,k}/((n_{T,k}^z + n_{T,k}^l + n_{F,k})M_c)$.

The first simulation analyzes the detection performance of three robust methods to randomly generated outliers. The outliers are introduced in the Jacobian matrix H and observation vector $\boldsymbol{z}$. The three robust methods are the LTS with decomposition (Table II), the popular largest normalized residual (LRr) and residual analysis (RA) [2]. RA is based on rejecting residuals obtained from WLS and departing from the Gaussian assumption (i.e., reject any residual $\{\hat{r}_i^2/\hat{S}_{ii}R_{ii}\} > 9$). The number of observations is 123 and $M_c = 100$.

At each MC run, a different set of quantities is generated for the observation noise, the locations and magnitudes of leverage points and observation outliers. Each island contains a leverage point placed randomly. One observation outlier is generated and placed randomly in the first seven islands. The island 8 was free from any observation outlier since it contains only 12 measurements. The LTS breakdown points were fixed to $\varepsilon^* = 0.1$ for each island. Each element in the observation outliers (n_z) is generated following a Gaussian $\mathcal{N}(6\sigma_i, \sigma_i^2)$, where σ_i is the

TABLE III
MONTE CARLO SE DETECTION PROBABILITIES AND ERROR NORM AVERAGE FOR AN IEEE 14 BUS SYSTEM TEST-BED

	P_d^l	P_d^z	P_f	d_I^l	d_I^z	x_I^M(pu)	x_I^M(deg)
LTS	0.98	0.59	0.01	0.57	0.35	0.025	0.155
RA	0.03	0.17	0.80	0.08	0.9	0.125	0.706
LRr	0.03	0.42	0.55	0.01	0.25	0.125	0.706

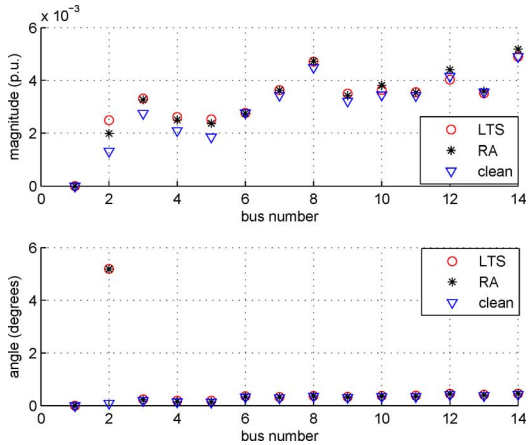


Fig. 2. Monte Carlo average absolute error in SE of the IEEE 14 bus system versus bus number: (a) Voltage magnitude (p.u.); (b) Voltage phase angle (deg).

standard deviation of the clean observation z_i. The leverage points, on the other hand, are obtained by adding a random element generated from $\alpha\mathcal{U}(2,12)$ to the ith row elements of H. The distribution $\mathcal{U}$ is the uniform distribution and α has an equal probability of half $(=1/2)$ to be 1 or -1. The standard deviation of the different SCADA measurements proposed in MATPOWER are 0.66% of the measured value plus a fixed value of 0.0016. Table III shows that the best detection for both leverage points is insured by the LTS after decomposition. The latter minimizes also the false detection. The popular LRr and RA behave badly as expected even when $\boldsymbol{h}$ is not attacked. Their errors are similar since the non-detected leverage points have a crucial impact on estimation bias. Even if $d_I^l = 0.57$ which is not near to 1, the efficiency of the LTS-based estimator is still high since its error norm average is close from the WLS run on the outliers-clean process (0.01 p.u. and 0.133 deg). This confirms the effectiveness of the LTS where the missed leverage points were not impacting heavily the SE. This is one possible outliers configuration which shows the need for robust methods against leverage points.

The second simulation illustrates the impact of the developed stealthy attack on the LTS and RA for the phase angle at bus 2. Fig. 2 depicts the Monte Carlo average absolute errors when the proposed condition (8) is satisfied with $\eta = 5$. Phase angle at bus 2 is modified in a stealthy fashion with the remaining estimates at other buses giving similar errors as the clean case. This confirms that the attack is concentrated on a unique bus. The average solution over 100 replications for LTS and RA are -1.28 deg and -1.29 deg. The Monte Carlo average estimate in the absence of the attack is -6.48 deg and the targeted phase value by the intruder is $-6.48/5 = -1.296$ deg. The latter value is very close to the obtained contaminated estimation. Analyzing neighboring buses by PMUs, for example, does not help the operator. A natural concern consist in analyzing the previous behavior if a PMU is installed at the same attacked bus 2.

TABLE IV
MONTE CARLO AVERAGE AND STANDARD DEVIATION OF PHASE ANGLE ESTIMATE AT BUS 2 OBTAINED WITH RESPECT TO n_s STEALTHY ATTACKS (LTS$_p$: LTS WITH A PMU, $n_c = 7, \theta_2 = -6.48$ deg)

	$n_s = 2$	$n_s = 4$	$n_s = 8$	$n_s = 17$
LTS	$-6.38(0.7)$	$-3.01(2.49)$	$\mathbf{-1.33(0.10)}$	$\mathbf{-1.25(0.09)}$
RA	$-1.92(1.14)$	$-4.03(2.07)$	$-1.25(0.29)$	$\mathbf{-1.10(0.16)}$
LTS$_p$	$-6.48(0.16)$	$-3.02(2.5)$	$\mathbf{-1.33(0.06)}$	$\mathbf{-1.29(0.04)}$
RA$_p$	$-6.05(0.8)$	$-5.16(1.32)$	$-2.77(2.02)$	$-2.48(1.84)$

Table IV provides the average phase angle estimate and standard deviation at bus 2 over 100 Monte Carlo replications considering an increased number of attacked elements in the Jacobian H where $\boldsymbol{h}_c(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x}) + \delta H\boldsymbol{x}$. The standard deviation for the PMU was fixed to 0.04 deg and 0.005 p.u. for phase angles and magnitude respectively. The attacker is assumed to concentrate the attack on one island (subsystem 1). Notice that bus 2 is present in two other islands and the operator is considered to be very safe where the observation is flagged as outlying even if detected in one island. Notice that an attack on the first island will generate outliers on overlapping or neighbor islands (injections, cut-lines) when evaluating the Jacobian. This could result in small additional errors in neighboring buses. The breakdown point of the estimates LTS is assumed to be 0.25 for island 1 and 0.1 for the remaining subsystems. There are 21 SCADA measurements in island 1; nine are active power flows and injections. It is assumed that the attacker has access to island 1. If the attacker follows the previous developed approach, he will target these by searching for the maximum absolute elements in the column H linked to bus 2 and related to active powers, after evaluating $n_{c,1}^{(1)} = \lceil 9 * 0.25 \rceil = 3$, he concludes that an attack of scenario 1 is possible when $n_s \geq 3$ whereas scenario 2 is obtained with $n_s \geq n_{c,1}^{(2)} = \lfloor 9 * 0.75 \rfloor + 1 = 7$. The n_s in this case should target the active power otherwise the value 9 should be replaced with 21 in the above which increases the needed number n_s. If $\varepsilon^* = 0.1$ for island 1, the standard deviation of the LTS reduces from 0.7 deg to 0.15 deg, the new result is -6.48 (0.15) deg; the WLS SE with clean signal gives -6.48 (0.1) deg.

Table IV shows that when $n_s \leq 2$, the LTS is resistant and can be trusted by the operator. At $n_s = 4$, both LTS and RA have broken-down and should not be trusted. This is a scenario 1 stealthy attack but the obtained state is not really controlled by the intruder (large standard deviation). Depending on the magnitude of the attack, the contaminated state changes but is unpredictable to the attacker who has partial knowledge. At $n_s = 8$, scenario 2 (highlighted in bold) is generated for LTS, where the stealthy attack controls the state. This is not the case for RA where scenario 1 seems still valid. At $n_s = 17$, both robust approaches are under scenario 2 attack. The LTS seems also to reject the PMU since it is not consistent with the majority of the data. This shows that the advantage of highly robust estimators can be changed into a weak point. The improvement brought by one PMU is clear for $n_s = 2$. However, beyond 2, the improvement is still limited. For bus 2, considering the value of the PMU on its own or considering the PMU after estimating

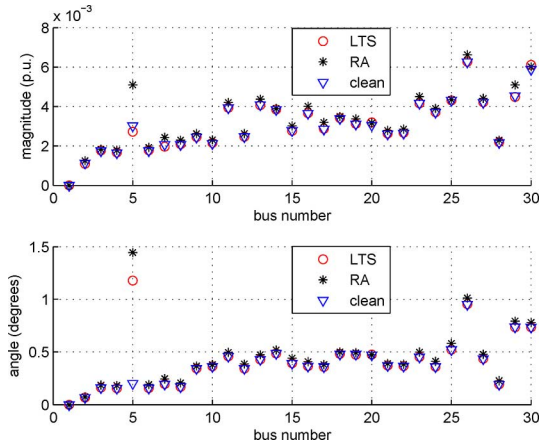


Fig. 3. Monte Carlo average absolute error in SE of the IEEE 30 bus system versus bus number: (a) voltage magnitude (p.u.); (b) voltage phase angle (deg).

the states from a SCADA SE (post-processing) when an attack is suspected might be a good solution.

The attack configuration satisfying (8) is also considered on the IEEE 30 bus systems where the phase angle of bus 5 is targeted. The system is decomposed into 12 subsystems [4] as follows: {2,4,5,6,7}, {2,5,6,7}, {6,8,28}, {6,9,10}, {27,29,30}, {10,21,22}, {12,14,15}, {1,2,3,4}, {10,12,15,16,17,18,19,20}, {6,8,10,22,24,25,27,28}, {10,15,18,19,20,21,22,23,24}, {4,6,9,10,12,16,17}. The attack's magnitude is $\eta_5 = 0.5$. Fig. 3 depicts the average absolute errors at each bus for both voltage magnitude and phase angle. Similar conclusions as in the case of the 14 bus system can be drawn.

The difficulty in generating the previous attacks in the Jacobian varies depending whether an AC or DC power flow is used. For the DC model, topology attacks seem easier to generate due to the linear model. For example, if the attacker wants to mislead the operator into thinking that a line is open, he needs to eliminate the measured flow (i.e., null) at that line and change the binary state of the circuit breaker or switch received by the operator to indicate open. The injections measured and linked to that line should be modified as well. In this case, the operator, after checking that the flows and injections correspond with the new state of the circuit breaker, validates that erroneous topology. Concerning the AC model, an attack could also target the switches states and measurements. However, it seems more complex practically to generate a stealthy attack since the linearized Jacobian depends on the unknown state which requires to read surrounding PMUs or a larger knowledge of the system by the intruder. More measurements should be contaminated as well and using the nonlinear measurement function ($\boldsymbol{h}_c$) in computing the residuals complicates the task of introducing such attack. Finally, this means that using an AC over a DC model by the operator offers the advantage of improving cyber-security. This is also true in the case where attacks target only the state of the grid, as observed in [8].

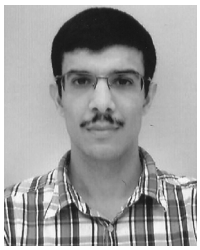
V. Conclusion

The paper analyzes the cyber-security of power systems state estimation considering both DC and AC formulations. The analysis focuses on attacks on the Jacobian matrix, i.e., regressors in the SE regression model which creates possibly coordinated leverage points. Conditions to generate stealthy attacks against robust estimators and diagnostic methods keeping good leverage points, and especially the least trimmed squares with decomposition, are derived. Factors simplifying stealthy attacks creation are highlighted. Some possible solutions and remedial actions are discussed. PMUs are effective against such stealthy attacks especially if used in post-processing. AC formulation offers also some additional complication for an intruder.

References

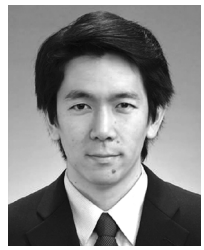
[1] M. Govindarasu and P. W. Bauer, "Special section on keeping the smart grid safe," *IEEE Power Energy Mag.*, vol. 10, no. 1, 2012.
[2] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. New York, NY, USA: CRC Press, 2004.
[3] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, 2000.
[4] L. Mili, M. G. Cheniae, and P. J. Rousseeuw, "Robust state estimation of electric power systems," *IEEE Trans. Circuits Syst. I: Fund. Theory Applicat.*, vol. 41, no. 5, pp. 349–358, 1994.
[5] L. Mili, M. G. Cheniae, N. S. Vichare, and P. J. Rousseeuw, "Robust state estimation based on projection statistics [of power systems]," *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 1118–1127, 1996.
[6] Y. Liu, M. K. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Security*, 2009, pp. 21–32.
[7] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S. S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *Proc. 49th IEEE Conf. Decision and Control (CDC)*, 2010, pp. 5991–5998.
[8] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012.
[9] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, 2011.
[10] J. Kim and L. Tong, "On topology attack of a smart grid: Undetectable attacks and countermeasures," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1294–1305, Jul. 2013.
[11] J. Kim and L. Tong, "On phasor measurement unit placement against state and topology attacks," in *Proc. IEEE Int. Conf. Smart Grid Communications (SmartGridComm) 2013*, Oct. 2013, pp. 396–401.
[12] L. Jia, J. Kim, R. J. Thomas, and L. Tong, "Impact of data quality on real-time locational marginal price," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 627–636, Mar. 2014.
[13] I. Watanabe, K. Masutomi, and I. Ono, "Robust meter placement against false data injection attacks on power system state estimation," in *ICONIP 2013, LNCS 8226*. New York, NY, USA: Springer, vol. 1, pp. 569–576.
[14] H. Nishino and H. Ishii, "Distributed detection of cyber-attacks and faults for power systems," in *Proc. 19th IFAC World Congress*, 2014.
[15] L. Chen-Ching, A. Stefanov, J. Hong, and P. Panciatici, "Intruders in the grid," *IEEE Power Energy Mag.*, vol. 10, no. 1, pp. 58–66, 2012.
[16] L. Tong, R. J. Thomas, and L. Xie, "Impacts of bad data and cyber attacks on electricity market operations," Power Syst. Eng. Res. Ctr. (PSERC), Tech. Rep., 2013.
[17] R. A. Maronna, R. D. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods, Wiley Series in Probability and Statistics*. Chichester, U.K.: Wiley, 2006.
[18] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts," *IEEE Signal Process. Mag.*, vol. 29, no. 4, pp. 61–80, Jul. 2012.
[19] M. G. Cheniae, L. Mili, and P. J. Rousseeuw, "Identification of multiple interacting bad data via power system decomposition," *IEEE Trans. Power Syst.*, vol. 11, no. 3, pp. 1555–1563, 1996.
[20] Y. Weng, R. Negi, Q. Liu, and M. D. Ilic, "Robust state-estimation procedure using a least trimmed squares pre-processor," *IEEE PES Innovative Smart Grid Technologies (ISGT)*, pp. 1–6, 2011.
[21] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER, a MATLAB power system simulation package," Jul. 2010 [Online]. Available: http://www.pserc.cornell.edu/matpower/manual.pdf

[22] J. Agullo, C. Croux, and S. Van Aelst, "The multivariate least-trimmed squares estimator," *J. Multivariate Analysis* vol. 99, no. 3, pp. 311–338, 2008 [Online]. Available: http://www.econ.kuleuven.be/public/NDBAE06/programs/mlts/mlts.txt



Yacine Chakhchoukh (M'10) received the Ph.D. degree in electrical engineering from Paris-Sud XI University, Paris, France, in 2010.

His industrial experience was with the French Electrical Transmission System Operator (RTE), France. He conducted research at the Technical University Darmstadt, Germany, from 2009 to 2011 and Arizona State University, Phoenix, AZ, USA, from 2011 to 2013. Currently, he is with the Tokyo Institute of Technology, Tokyo, Japan. His research interests are in cyber security of power systems, smart grid applications, time series analysis, and robust estimation for signal processing.



Hideaki Ishii (M'02–SM'12) received the M.Eng. degree in applied systems science from Kyoto University, Kyoto, Japan, in 1998, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2002.

He was a Postdoctoral Research Associate with the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign, Urbana, IL, USA, from 2001 to 2004, and a Research Associate with the Department of Information Physics and Computing, University of Tokyo, Tokyo, Japan, from 2004 to 2007. Currently, he is an Associate Professor in the Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, Yokohama, Japan. He is a co-author of the book *Limited Data Rate in Control Systems with Networks, Lecture Notes in Control and Information Sciences* (Springer, 2002). His research interests are in networked control systems, multiagent systems, hybrid systems, cyber security of power systems, and probabilistic algorithms.

Dr. Ishii has served as an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is an Associate Editor for *Automatica*. He is the Chair of the IFAC Technical Committee on Networked Systems.