Robust Dynamic State Estimator to Outliers and Cyber Attacks

Junbo Zhao, *Student Member, IEEE*, Lamine Mili, *Fellow, IEEE*, Ahmed Abdelhadi, *Senior Member, IEEE*

***Abstract*—To perform power system monitoring and control using synchrophasor measurements, various dynamic state estimators have been proposed in the literature, including the extended Kalman filte (EKF) and the unscented Kalman filte (UKF). However, they are unable to handle system model parameter errors and any type of outliers, precluding them from being adopted for power system real-time applications. In this paper, we develop a robust iterated extended Kalman filte based on the generalized maximum likelihood approach (termed GM-IEKF) for dynamic state estimation. The proposed GM-IEKF can effectively suppress observation and innovation outliers, which may be induced by model parameter gross errors and cyber attacks. We assess its robustness by carrying out extensive simulations on the IEEE 39-bus test system. From the results, we fin that the GM-IEKF is able to cope with at least 25% outliers, including in position of leverage.**

***Index Terms*—Dynamic state estimation, cyber attacks, model uncertainty, robust estimation, extended Kalman filte , unscented Kalman filte , phasor measurement unit, breakdown point.**

I. Introduction

TTH widespread deployment of synchrophasor measurement units (PMUs) on power transmission grids has made possible the real-time monitoring and control of power system dynamics. However, these functions cannot be reliably achieved without the development of a fast and robust dynamic state estimator (DSE). Indeed, the benefit of using a DSE are an improved system oscillation monitoring and an enhanced local and global system control, to cite a few [1], [2].

To date, a variety of dynamic state estimators have been proposed in the literature, including the extended Kalman filte (EKF), the iterated EKF (IEKF), and the unscented Kalman filte (UKF) [5]–[8], to cite a few. They can produce good results if the system model is well calibrated and the PMU measurements are reliable and secure. However, these assumptions may not hold true in practical power systems. Indeed, the system parameters may change with time and the load models are uncertain, among others, yielding uncertain dynamical system model. In addition, with the strong reliance of smart grid functions on communication networks, cyber attacks have become a major concern. The latter can be classifie as follows [11]:

- ***Bias injection attacks***, where an adversary attempts to corrupt the content of either the measurement or the control signals; for example, the man-in-the-middle attacks intercept the PMU measurement signals and corrupt them with large bias;
- ***Denial of Service (DoS) attack***, where the actuator and sensor data are prevented from reaching their respective destinations, resulting in the absence of data; for instance, this will be the case if the PMU metered values do not reach the phasor data concentrator (PDC);
- ***Replay attacks***, where a hacker firs performs a disclosure attack from a certain time period, gathering sequences of data, and then begins replaying the data during a certain period; for example, the current PMU measurements processed by a DSE are replaced by past values.

To handle observation and innovation outliers, we develop a robust iterated extended Kalman filte -based DSE using the generalized maximum likelihood approach, termed GM-IEKF for short. Specificall , we firs build a batch-mode regression form to enhance data redundancy, which allows our projection statistics to detect innovation and observation outliers. The innovation outliers can be induced by system model parameter errors while the observation outliers may be induced by impulsive communication noise or cyber attack, among others. Then, a GM-estimator using the convex Huber cost function is proposed to suppress the outliers. Next, a robust estimation error covariance matrix is updated by means of the total influenc function. Finally, the finite-sampl breakdown point of the GM-IEKF is evaluated to quantify the resistance of our DSE to cyber attacks. The rest of the paper is organized as follows. Section II deals with the problem formulation. Section III presents the proposed GM-IEKF, while Section IV discusses the results of some simulations carried out on the IEEE 39-bus test system. Section V concludes the paper.

II. Problem formulation

A discrete-time state space representation of a general nonlinear dynamical power system can be expressed as

$$\boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k\right) + \boldsymbol{w}_k, \tag{1}$$

$$\boldsymbol{z}_k = \boldsymbol{h}\left(\boldsymbol{x}_k, \boldsymbol{u}_k\right) + \boldsymbol{v}_k, \tag{2}$$

where $\boldsymbol{x}_k \in \mathbb{R}^{n\times 1}$ and $\boldsymbol{z}_k \in \mathbb{R}^{m\times 1}$ are the state vector and the measurement/observation vector at time sample k, respectively; $\boldsymbol{f}$ and $\boldsymbol{h}$ are vector-valued nonlinear functions; $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are the system process and observation noise, respectively, and they are assumed to be independent and identically distributed with zero mean and covariance matrices $\boldsymbol{W}_k$ and $\boldsymbol{R}_k$, respectively; $\boldsymbol{u}_k$ is the system input vector. In this paper, the detailed two-axis model with IEEE-DC1A exciter and TGOV1 turbine-governor is considered, yielding a 9th-order model [9]. It should be noted that in most literature, the four-order generator model is used for DSE, assuming that the fiel voltage is known and the mechanical power is a constant

The authors are with the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Northern Virginia Center, Falls Church, VA 22043, USA. (e-mail: zjunbo@vt.edu, lmili@vt.edu, aabdelhadi@vt.edu)

[8]. However, they may not hold as the fiel current and voltage are not measured in brushless excitation systems [4], and the mechanical power can vary significantl when control features such as fast valving or special protection schemes are used to limit the output of the steam driven generator during transients [10]. Therefore, it is of vital importance to keep track of the dynamic state variables of the exciter and the governor for controls and system stability analysis.

To estimate the system dynamic states using Kalman-type filters a two-step procedure is applied, namely a prediction step using (1), which is a Markov model, and a filtering/updat step using (2). Specificall , given a state estimate at time step $k-1$, $\widehat{\boldsymbol{x}}_{k-1|k-1}$, with its covariance matrix, $\boldsymbol{\Sigma}_{k-1|k-1}$, the predicted state is directly calculated from (1) or through a set of points drawn from the distribution that are following the probability distribution of $\widehat{\boldsymbol{x}}_{k-1|k-1}$. The latter is dependent on the assumed probability distributions of $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$. As for the filterin step, the predictions are used together with the observations at time sample k to estimate the state vector and its covariance matrix.

The Kalman-type filters including EKF, IEKF and UKF work well if the system model is well calibrated and the PMU measurements are reliable and secure. However, these assumptions may not hold true in practice; for instance, the parameter values of the synchronous generators may be inaccurate or the inputs may be unknown [12], yielding an uncertain system model; the PMU measurements may be attacked, resulting in strongly biased measurements (bias injection attacks); measurements may be lost (Dos attacks) or may be repeated from previous time samples (replay attacks), to cite a few [11], [13]. To address these problems, we propose to develop a robust GM-IEKF as described next.

III. Proposed GM-IEKF and Its Breakdown Point

A. Development of GM-IEKF

The proposed GM-IEKF consists of three major steps, which are the construction of batch-mode regression form, the robust prewhitening step, and the robust regression step with state and estimation error covariance matrix updating. The details for the development of the GM-IEKF are elaborated below.

First of all, we propose to build the batch-mode regression form to enhance the data redundancy by processing the observations and the predictions simultaneously. Note that this redundant regression form allows our GM-IEKF to bound the influenc of model uncertainties and various cyber attacks. To be specific given the filtere state vector $\widehat{\boldsymbol{x}}_{k-1|k-1}$ and its covariance matrix $\boldsymbol{\Sigma}_{k-1|k-1}$ at time step $k-1$, the predicted state and its covariance matrix can be obtained through

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1|k-1}), \tag{3}$$

$$\boldsymbol{\Sigma}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{\Sigma}_{k-1|k-1}\boldsymbol{F}_{k-1}^T + \boldsymbol{W}_k, \tag{4}$$

where $\boldsymbol{F}_{k-1} = {}^{\partial \boldsymbol{f}}/_{\partial \boldsymbol{x}}\big|_{\boldsymbol{x}=\widehat{\boldsymbol{x}}_{k-1|k-1}}$. Defin $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \boldsymbol{\eta}_k$, where $\boldsymbol{x}_k$ is the true state vector and $\boldsymbol{\eta}_k$ is the state prediction error vector with zero mean and covariance matrix $\boldsymbol{\Sigma}_{k|k-1}$. Then, by putting it with (2) into a compact form, we obtain

$$\begin{bmatrix} \widehat{\boldsymbol{x}}_{k|k-1} \\ \boldsymbol{z}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{h}(\boldsymbol{x}_k) \end{bmatrix} + \begin{bmatrix} -\boldsymbol{\eta}_k \\ \boldsymbol{v}_k \end{bmatrix}, \tag{5}$$

which can be rewritten as

$$\widetilde{\boldsymbol{z}}_k = \widetilde{\boldsymbol{h}}(\boldsymbol{x}_k) + \widetilde{\boldsymbol{e}}_k, \tag{6}$$

associated with the augmented error covariance matrix

$$\mathbb{E}\left[\widetilde{\boldsymbol{e}}_k\widetilde{\boldsymbol{e}}_k^T\right] = \begin{bmatrix} \boldsymbol{\Sigma}_{k|k-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_k \end{bmatrix} = \boldsymbol{S}_k\boldsymbol{S}_k^T, \tag{7}$$

where $\boldsymbol{S}_k$ is calculated by the Cholesky decomposition technique.

In the batch-mode regression form given by (6), a data prewhitening process is performed to uncorrelate the state prediction errors by using the matrix $\boldsymbol{S}_k$. This is done by premultiplying that nonlinear regression model by $\boldsymbol{S}_k^{-1}$, yielding

$$\boldsymbol{S}_k^{-1}\widetilde{\boldsymbol{z}}_k = \boldsymbol{S}_k^{-1}\widetilde{\boldsymbol{h}}(\boldsymbol{x}_k) + \boldsymbol{S}_k^{-1}\widetilde{\boldsymbol{e}}_k, \tag{8}$$

which is further rewritten as

$$\boldsymbol{y}_k = \boldsymbol{\varphi}(\boldsymbol{x}_k) + \boldsymbol{\xi}_k. \tag{9}$$

where $\mathbb{E}[\boldsymbol{\xi}_k\boldsymbol{\xi}_k^T] = \boldsymbol{I}_k$ and $\boldsymbol{I}_k$ is an identity matrix.

To bound the influenc of system parameter errors and suppress outliers when performing nonlinear regression on (9), a GM-estimator is proposed that minimizes an objective function given by

$$J(\boldsymbol{x}) = \sum_{i=1}^{m+n} \varpi_i^2 \rho(r_{S_i}), \tag{10}$$

where ϖ_i is the weight that will be elaborated later; $r_{S_i} = r_i/s\varpi_i$ is the standardized residual; $r_i = y_i - \varphi_i(\widehat{\boldsymbol{x}})$ is the residual; $s = 1.4826 \cdot b_m \cdot \text{median}_i\,|r_i|$ is the robust scale estimate; b_m is a correction factor for unbiasedness at the Gaussian distribution [15]; $\rho(\cdot)$ is the convex Huber ρ-function made up of a quadratic and a tangent function.

Calculating the weight ϖ_i: Note that applying $\boldsymbol{S}_k$ directly for prewhitening will cause a smearing effect to the estimation results when outliers occur. To address this issue, we firs detect and downweight the outliers in the derived batch-mode regression form by means of weights calculated using the projection statistics (PS) [14], [15] and a statistical test applied to them. The outliers are data points that are distant from the bulk of the point cloud. The PS values are some kinds of robust distances of a collection of data points, $\boldsymbol{\ell}_i$. Mathematically, the i-th PS value is define as

$$PS_i = \max_{\|\boldsymbol{l}\|=1} \frac{\left|\boldsymbol{\ell}_i^T l - med_j\left(\boldsymbol{\ell}_j^T l\right)\right|}{1.4826\; med_k\left|\boldsymbol{\ell}_k^T l - med_j\left(\boldsymbol{\ell}_j^T l\right)\right|}. \tag{11}$$

To detect outliers, we apply the PS to a 2-dimensional matrix $\boldsymbol{Z}$ that contains serially correlated samples of the innovations and the predicted state variables. Note that the innovation vector is define as the difference between the observations and their associated predicted values at the previous step. Formally, we have

$$\boldsymbol{Z} = \begin{bmatrix} \widehat{\boldsymbol{x}}_{k-1|k-2} & \widehat{\boldsymbol{x}}_{k|k-1} \\ \boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2}) & \boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1}) \end{bmatrix}. \tag{12}$$

Instead, we may apply the PS to higher dimensional samples, but we found that 2 dimensions are enough to identify outliers. Note that the PS values are calculated separately for the four subgroups of the matrix $\boldsymbol{Z}$, i.e., the predicted state vector and the innovation vectors of the real and the reactive power and of the voltage magnitudes and angles. Then, the PS values are compared to a threshold to identify the outliers. By conducting extensive Mote Carlo simulations and QQ-plots, we fin that the PS values follow a chi-square distribution with 2 degrees of freedom. Therefore, the threshold for the statistical test will be set to $\chi^2_{2,0.975}$ for a 97.5% significanc level. Those PS values that satisfy $PS_i > \chi^2_{2,0.975}$ will be flagge as outliers and downweighted via

$$\varpi_i = \min\left(1, d^2/PS_i^2\right), \tag{13}$$

where d is set to 1.5 to yield good statistical efficien y without increasing too much the bias induced by the outliers.

Robust batch-mode regression: To minimize (10), one takes its partial derivative and sets it equal to zero, yielding

$$\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}} = \sum_{i=1}^{m} -\frac{\varpi_i \boldsymbol{c_i}}{s}\psi(r_{S_i}) = \boldsymbol{0}, \tag{14}$$

where $\boldsymbol{c}_i$ is the i-th column vector of the matrix $\boldsymbol{C}^T$ given by $\boldsymbol{C} = {}^{\partial \boldsymbol{\varphi}}/_{\partial \boldsymbol{x}}|_{\boldsymbol{x}=\widehat{\boldsymbol{x}}}$; $\psi(r_{S_i}) = \partial\rho(r_{S_i})/\partial r_{S_i}$. Then, by dividing and multiplying by r_{S_i} on both sides of (14), we get the following equation in matrix form

$$\boldsymbol{C}^T\boldsymbol{Q}\left(\boldsymbol{y} - \boldsymbol{\varphi}(\boldsymbol{x})\right) = \boldsymbol{0}, \tag{15}$$

where $\boldsymbol{Q}$ =diag$(q(r_{S_i}))$ and $q(r_{S_i}) = \psi(r_{S_i})/r_{S_i}$.

Taking a first-orde Taylor series expansion of $\boldsymbol{\varphi}(\boldsymbol{x})$ about $\widehat{\boldsymbol{x}}_{k|k}$ and using the iteratively reweighted least squares algorithm [16], the state vector correction at the j-th iteration is calculated by

$$\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \left(\boldsymbol{C}^T\boldsymbol{Q}^{(j)}\boldsymbol{C}\right)^{-1}\boldsymbol{C}^T\boldsymbol{Q}^{(j)}\left(\boldsymbol{y} - \boldsymbol{\varphi}\left(\widehat{\boldsymbol{x}}_{k|k}^{j}\right)\right), \tag{16}$$

where $\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \widehat{\boldsymbol{x}}_{k|k}^{(j+1)} - \widehat{\boldsymbol{x}}_{k|k}^{(j)}$ and $\boldsymbol{C}$ is evaluated at $\widehat{\boldsymbol{x}}_{k|k}^{(j)}$. The algorithm converges when $\left\|\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)}\right\|_{\infty} \leq 10^{-2}$.

After the convergence of the iterative process, the estimation error covariance matrix $\boldsymbol{\Sigma}_{k|k}$ of the GM-IEKF needs to be updated so that the state prediction at the next time sample can be performed. Following the work from [14], [16], we derive the estimation error covariance matrix of our GM-IEKF as

$$\boldsymbol{\Sigma}_{k|k} = \frac{\mathbb{E}_{\Phi}\left[\psi^2(r_{S_i})\right]}{\left\{\mathbb{E}_{\Phi}\left[\psi'(r_{S_i})\right]\right\}^2}\left(\boldsymbol{C}_k^T\boldsymbol{C}_k\right)^{-1}\left(\boldsymbol{C}_k^T\boldsymbol{Q}_{\varpi}\boldsymbol{C}_k\right)\left(\boldsymbol{C}_k^T\boldsymbol{C}_k\right)^{-1}, \tag{17}$$

where $\mathbb{E}(\cdot)$ is the expectation operator; Φ is the standard normal probability distribution function; $\psi'(r_{S_i})$ is the derivative of $\psi(r_{S_i})$ with respect to r_{S_i}; $\boldsymbol{Q}_{\varpi} = \text{diag}(\varpi_i)$.

Remark: If a non-iterative EKF is considered such as the one in [17], the innovation outliers will corrupt the results of PS, yielding unreliable estimation results. In other words, it is unable to suppress any model parameter errors. By contrast, our GM-IEKF is able to effectively handle them.

B. Breakdown Point of GM-IEKF

Thanks to the statistical robustness of PS and the GM-estimator, GM-IEKF is resistant to outliers induced by cyber attacks. Then the remaining question is how many outliers GM-IEKF can handle without giving unreliable estimation results, that is, what is its breakdown point? This concept provides a general measure of the robustness of an estimator to outliers. Formally, it is define as [18]

$$\varepsilon\left(\widehat{\boldsymbol{x}}, \boldsymbol{\varphi}, T\left(\widehat{\boldsymbol{x}}\right)\right) = \min\left\{\varepsilon_-\left(\widehat{\boldsymbol{x}}, \boldsymbol{\varphi}, T\left(\widehat{\boldsymbol{x}}\right)\right), \varepsilon_+\left(\widehat{\boldsymbol{x}}, \boldsymbol{\varphi}, T\left(\widehat{\boldsymbol{x}}\right)\right)\right\}, \tag{18}$$

where the lower and upper breakdown are define as

$$\varepsilon_-\left(\widehat{\boldsymbol{x}}, \boldsymbol{\varphi}, T\left(\widehat{\boldsymbol{x}}\right)\right) = \begin{cases} \min\limits_{0\leq p\leq m+n}\left\{\frac{p}{m+n};\ \sup\limits_{p\in D^p}\varphi\left(\widehat{\boldsymbol{x}}, T\left(\widehat{\boldsymbol{x}}\right)\right) = \sup\limits_{T(\widehat{\boldsymbol{x}})}\boldsymbol{\varphi}\left(\widehat{\boldsymbol{x}}, T\left(\widehat{\boldsymbol{x}}\right)\right)\right\} \\ \quad if\ \sup\limits_{T(\widehat{\boldsymbol{x}})}\boldsymbol{\varphi}\left(\widehat{\boldsymbol{x}}, T\left(\widehat{\boldsymbol{x}}\right)\right) > \boldsymbol{\varphi}\left(\widehat{\boldsymbol{x}}, T\left(\widehat{\boldsymbol{x}}\right)\right) \\ 1 \quad \text{otherwise} \end{cases} \tag{19}$$

$$\varepsilon_+\left(\widehat{\boldsymbol{x}}, \boldsymbol{\varphi}, T\left(\widehat{\boldsymbol{x}}\right)\right) = \begin{cases} \min\limits_{0\leq p\leq m+n}\left\{\frac{p}{m+n};\ \sup\limits_{p\in D^p}\varphi\left(\widehat{\boldsymbol{x}}, T\left(\widehat{\boldsymbol{x}}\right)\right) = \sup\limits_{T(\widehat{\boldsymbol{x}})}\boldsymbol{\varphi}\left(\widehat{\boldsymbol{x}}, T\left(\widehat{\boldsymbol{x}}\right)\right)\right\} \\ \quad if\ \sup\limits_{T(\widehat{\boldsymbol{x}})}\boldsymbol{\varphi}\left(\widehat{\boldsymbol{x}}, T\left(\widehat{\boldsymbol{x}}\right)\right) < \boldsymbol{\varphi}\left(\widehat{\boldsymbol{x}}, T\left(\widehat{\boldsymbol{x}}\right)\right) \\ 1 \quad \text{otherwise} \end{cases} \tag{20}$$

where p is the number of outliers induced by either cyber attacks or other reasons; D^p represents the data sample space where p data points are replaced by outliers; $T(\cdot)$ is the proposed GM-IEKF estimator in the functional form.

IV. Numerical Results

The performance of the proposed GM-IEKF is evaluated on the IEEE 39-bus test system. The standard EKF and UKF are implemented for comparisons. In the simulations, the time-domain simulation results are used to generate a collection of samples of the nodal voltage magnitudes and phase angles as well as real and reactive power injections at the terminal buses of all the generators; a sampling rate of 48 samples/second is assumed; both system process and measurement noise are assumed to follow Gaussian distribution with zero mean and standard deviation 10^{-2}; the diagonal elements of the initial error covariance matrix of the UKF are set to 10^{-4}; the initial values of the state vector are arbitrarily chosen for all three estimators; the machine parameters are taken from [19]; at t=0.5s Line 15-16 is switched off to simulate system transients; the maximal number of iterations of GM-IEKF is 20. Due to the space limitation, only the estimated rotor speed, rotor angle, fiel voltage and mechanical power of Generator 5 are provided for illustration.

A. Case 1: Bias Injection Attacks

The real and reactive power measurements of Generator 5 are corrupted with 20% error from 3s to 4s to simulate the bias injection attacks. The test results are shown in Fig. 1. It is observed that both EKF and UKF are providing strongly biased state estimates due to their lack of robustness to gross

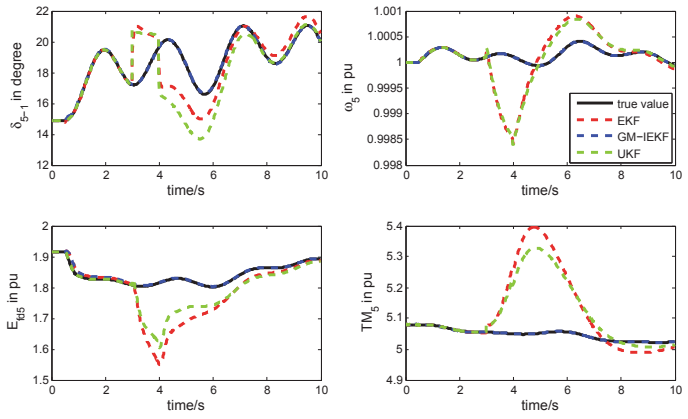

Fig. 1. Comparison results of three methods in terms of bias injection attacks on Generator 5 from 3s to 4s.

measurement errors. By contrast, the proposed GM-IEKF is able to suppress their influence leading to very reliable tracking results.

B. Case 2: DoS Attacks

All the PMU measurements of Generator 5 are lost due to the jammed communication between PMU to PDC or between PDC to the control center by means of DoS attacks. This absence of PMU measurements is modeled by considering them to be zeros or pure communication noise. The tracking performance of each method is presented in Fig. 2. It is found that EKF and UKF are mislead by those pure noise signals, yielding unreliable estimates. However, the proposed GM-IEKF is capable of suppressing them. This is achieved by firs identifying the pure noise signal as outliers and assigning them with very small weights, then suppressing their effects by using the GM-estimator.

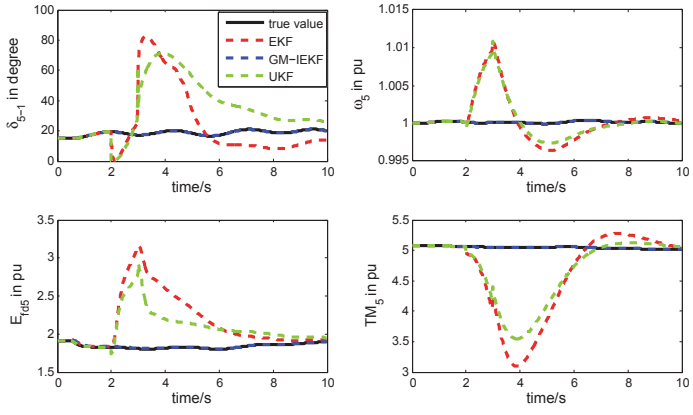

Fig. 2. Comparison results of three methods in terms of Dos attacks on Generator 5 from 2s to 4s.

C. Case 3: Replay Attacks

The received PMU measurements of Generator 5 from 6s to 8s are replaced by the past measurements from 4s to 6s. Fig. 3 shows the state estimates of each method. We observe that EKF and UKF have relative larger bias than the proposed GM-IEKF since they could not balance well the tradeoff between the predicted state vector and the incoming measurements. It

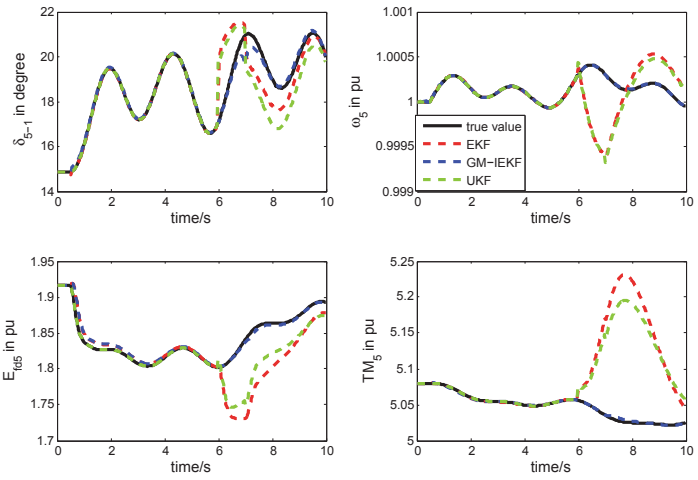

Fig. 3. Comparison results of three methods in terms of replay attacks on Generator 5 from 6s to 8s.

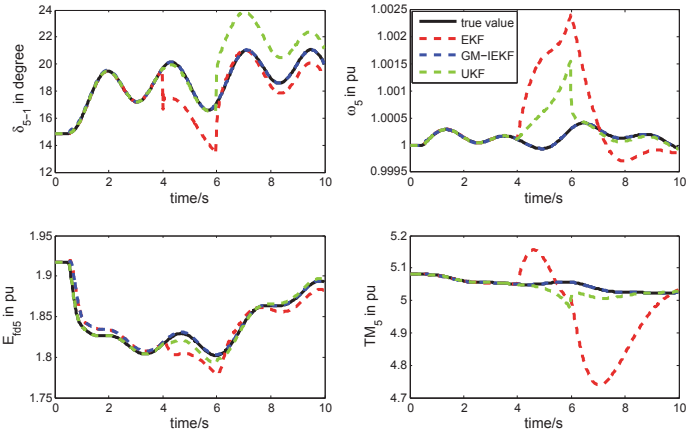

Fig. 4. Comparison results of three methods in terms of large model uncertainties on Generator 5 from 4s to 6s.

is also seen that GM-IEKF is biased slightly. This is because the large inconsistency between the model outputs and the measurements would be effectively identifie as outliers and downweighted. By contrast, if this inconsistency is small, it will not be identifie as outliers and as a result the GM-IEKF processes the predicted state vector and measurements equally somehow, yielding slightly biased estimation results.

D. Case 4: Model Uncertainties

To simulate the large generator model uncertainties, we assume that Generator 5 is not well calibrated, i.e., a generator parameter error occurs from 2s to 4s, leading to two incorrectly predicted dynamic states. The latter is represented by adding 20% errors to the predicted ω_5 and δ_5. The tracking performance of each method is shown in Fig. 4. It can be found that the UKF is less sensitive to model parameter error than the EKF, but its estimations of rotor angle and mechanical power are not acceptable. By contrast, the proposed GM-IEKF is able to balance the tradeoff between predicted state and the measurements through the projection statistics and the iterative robust regression, resulting in the best tracking performance among three methods. On the other hand, if the generator model is calibrated and only 1% error occurs on the predicted ω_5 and δ_5, the tracking results of both EKF and UKF are

TABLE I
AVERAGE COMPUTING TIMES OF THE THREE DSE METHODS FOR EVERY PMU SAMPLE

Cases	EKF	UKF	GM-IEKF
Case 1	5.24ms	5.28ms	9.56ms
Case 2	5.28ms	5.29ms	9.58ms
Case 3	5.33ms	5.38ms	9.50ms
Case 4	6.42ms	6.37ms	9.62ms

improved significantl as shown in Fig. 5. However, GM-IEKF still outperforms EKF and UKF.

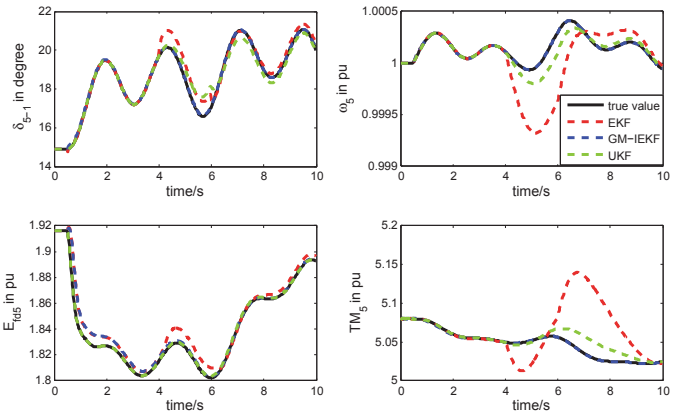

Fig. 5. Comparison results of three methods in terms of small model uncertainties on Generator 5 from 4s to 6s.

E. Breakdown Point to Cyber Attacks

To investigate the breakdown point of GM-IEKF to cyber attacks, we carry out extensive simulations on the IEEE 39-bus test system using the concept of finit sample breakdown in nonlinear regression define in (18). By replacing a varying number of observations by outliers in the vector $\boldsymbol{y}_k$ of dimension $m_t = m+n$, it is observed that the GM-IEKF can handle at least 25% outliers. Please note that the determination of the exact breakdown point of GM-IEKF is an open question that requires further investigations.

F. Computational Efficienc

To validate the applicability of the proposed GM-IEKF to online estimation with PMU sampling rate 30 or 60 samples per second, its computational efficien y is analyzed and compared to that of the EKF and UKF in the previous four cases. The test is performed on a PC with Intel Core i5, 2.50 GHz, 8GB of RAM. The average computing time of each method for every PMU sample is displayed in Table I. We observe from this table that EKF and UKF have comparative computational efficien y and their computing time are much lower than the PMU sampling period, which are 33.3ms and 16.7ms for 30 sample/s and 60 samples/s, respectively. Although GM-IEKF spends about 4 more millisecond, its execution time is still lower than the PMU sampling period, demonstrating its ability to track system real-time dynamic states.

V. CONCLUSION

This paper proposes a robust GM-IEKF method to track power system dynamic state variables using PMU measurements. The proposed GM-IEKF can effectively bound the influenc of a few parameter errors and various types of cyber attacks thanks to its statistical robustness, which is achieved by means of projection statistics and the GM-estimator. Comparison results with EKF and UKF on the IEEE 39-bus test system demonstrate the enhanced robustness of the proposed GM-IEKF. However, our GM-IEKF provides poor results when there are a large number of inaccurate parameters in the model. By contrast, the H-infinit filte can handle that case, but it lacks robustness to any type of outliers. As a future work, we will develop a hybrid robust DSE that integrates the H-infinit and the GM-IEKF filte , which will have the strengths of both methods.

REFERENCES

[1] I. Kamwa, R. Grondin, and Y. Hebert, "Wide-area measurement based stabilizing control of large power systems-a decentralized/hierarchical approach," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 136–153, Feb. 2001.
[2] H. Ni, G. T. Heydt, L. Mili, "Power system stability using robust wide area control," *IEEE Trans. Power Syst.*, Vol. 17, No. 4, pp. 1123–1131, Nov. 2002.
[3] L. Fan and Y. Wehbe, "Extended Kalman filterin based real-time dynamic state and parameter estimation using PMU data," *Elect. Power Syst. Res.*, vol. 103, pp. 168–177, Oct. 2013.
[4] E. Ghahremani, I. Kamwa, "Dynamic state estimation in power system by applying the extended Kalman filte with unknown inputs to phasor measurements," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2556–2566, Nov. 2011.
[5] E. Ghahremani, I. Kamwa, "Online state estimation of a synchronous generator using unscented Kalman filte from phasor measurements units," *IEEE Trans. Energy Convers.*, vol. 26, no. 4, pp. 1099–1108, Dec. 2011.
[6] S. Wang, W. Gao, A. P. S. Meliopoulos, "An alternative method for power system dynamic state estimation based on unscented transform," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 942–950, May 2012.
[7] A. K. Singh, B. C. Pal, "Decentralized dynamic state estimation in power systems using unscented transformation," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 794–804, Sep 2014.
[8] N. Zhou, D. Meng, Z. Huang, G. Welch, "Dynamic state estimation of a synchronous machine using PMU data: A comparative study," *IEEE Trans. Smart Grid.*, Vol. 6, no. 1, pp: 450–460, 2015.
[9] M. A. Pai, *Energy function analysis for power system stability*. Springer Science & Business Media, 1989.
[10] "Modeling of gas turbines and steam turbines in combined cycle power plants," CIGRE report, Task Force C4.02.25, 2003.
[11] A. Teixeira, D. Perez, H. Sandberg, K. H. Johansson, "Attack models and scenarios for networked control systems," *Proceedings of the 1st ACM international conference on High Confidenc Networked Systems*, pp: 55-64, 2012.
[12] Z. Huang, P. Du, D. Kosterev, S. Yang, "Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection," *IEEE Trans. Power Syst.*, Vol. 28, no. 2, pp. 1939–1949, 2013.
[13] A. F. Taha, J. Qi, J. Wang, J. H. Panchal, "Risk mitigation for dynamic state estimation against cyber attacks and unknown inputs," *IEEE Trans. Smart Grid*, in press. DOI: 10.1109/TSG.2016.2570546.
[14] M. Gandhi, L. Mili, "Robust kalman filte based on a generalized maximum-likelihood-type estimator, *IEEE Trans. Signal Processing*, vol. 58, no. 5, pp. 2509–2520, 2010.
[15] L. Mili, M. Cheniae, N. Vichare, and P. Rousseeuw, "Robust state estimation based on projection statistics," *IEEE Trans. Power Syst.*, vol. 11, no. 2, p. 1118–1127, 1996.
[16] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influenc Functions*. New York: John Wiley & Sons, Inc., 1986.
[17] N. Marcos, J. B. Zhao, L. Mili, "A robust extended Kalman filte for power system dynamic state estimation using PMU measurements," in *Proc. of IEEE PES General Meeting*, Boston, CO, pp. 1-5, July 2016.
[18] A. J. Stromberg, D. Ruppert, "Breakdown in nonlinear regression," *Journal of the American Statistical Association*, Vol. 87, no. 420, pp. 991–997, 1992.
[19] IEEE PES, "Dynamic models for turbine-governors in power system studies," Jan 2013.

State Estimation for Heavily Loaded System: A Comparative Study

Junbo Zhao, *Student Member, IEEE*, Lamine Mili, *Life Fellow, IEEE*

***Abstract*—Real time state information provided by the state estimator plays a major role in power system monitoring and control. As a result, the convergence of the estimator under various system operating conditions becomes one of the key requirements. This paper presents a robust state estimation framework that generalizes several well-known estimators. The statistical robustness of each estimator has been studied analytically through the total influenc function. Furthermore, the correlation between the statistical robustness of an estimator and the numerical robustness of the iterative algorithm is investigated as well. Numerical results carried out on the IEEE test system reveal that the Schweppe-type Huber generalized maximum-likelihood estimator works well in all simulated scenarios while the other alternatives have convergence problems or numerical instability issues under stressed system operating conditions.**

***Index Terms*—Power system state estimation, heavily loaded system, robust estimation, power system operation, convergence.**

I. Introduction

POWER system static state estimator is a basic operational tool in modern energy management system as it provides a complete, coherent and reliable real-time data-base for various applications, such as contingency analysis, voltage stability assessment, optimal power fl w, to cite a few. As a result, its convergence property under different operation conditions is a critical issue for system monitoring and control. Indeed, the divergence of the state estimator under stressed system condition is one of the most important factors that contributed to the 2003 northeast blackout [1].

Among all estimators proposed in the literature, the weighted least squares (WLS)-based state estimator is one of the most widely used methods by utilities. However, its convergence characteristic under different system operating conditions has not been well-investigated. Extensive simulations were carried out in [2]–[4] to study the convergence characteristic of the Gauss-Newton-based WLS estimator subject to topology errors and load changes. It is found that the WLS estimator yields highly biased state estimates or may suffer from convergence problems under these conditions. Note that the converged state estimates are extremely important for assessing system voltage stability margin and taking preventive control actions to avoid voltage collapse. Besides the WLS estimator, other statistical robust estimators have been proposed, including the least absolute value (LAV) estimator, the Huber Maximum-likelihood (M)-estimator and the Schweppe-type Huber generalized Maximum-likelihood (SHGM)-estimator [5]. However, their performances under stressed system conditions have not been studied. In addition, the relationship between statistical robustness and numerical robustness is rarely discussed. In this paper, we present a robust state estimation framework that generalizes all the aforementioned estimators. The statistical robustness of each estimator has been studied analytically through the total influ ence function. Furthermore, the correlation between the statistical robustness of an estimator and the numerical robustness of the iterative algorithm is investigated as well. It is found through extensive simulation results that the WLS estimator has convergence problem in presence of stressed system conditions while the LAV and the Huber M-estimator only present numerical stability issue for heavily loaded system. By contrast, the SHGM-estimator works well in all scenarios thanks to its statistical and numerical robustness.

The organization of this paper is as follows: Section II shows the proposed robust state estimation framework and investigates the statistical and numerical robustness of each estimator. Section III presents and analyzes the simulation results, and finall Section IV concludes the paper.

II. Generalized Maximum-likelihood Estimation Framework

A. Proposed Estimation Framework

For an N-bus power system, the relationship between the vector of measurements $\boldsymbol{z} \in \mathbb{R}^m$ and the state vector $\boldsymbol{x} \in \mathbb{R}^n$, $n = 2N - 1 < m$ is given by

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

where $\boldsymbol{x}$ contains the nodal voltage magnitudes and phase angles; $\boldsymbol{h}(\cdot) : \mathbb{R}^n \to \mathbb{R}^m$ is a vector-valued nonlinear function; $\boldsymbol{e} \in \mathbb{R}^m$ is the measurement error vector that is assumed to have zero mean and a covariance matrix $\boldsymbol{R} \in \mathbb{R}^{m \times m}$.

There exist several approaches to estimate $\boldsymbol{x}$ and among them, the WLS, the LAV, the Huber M-estimator and the SHGM-estimator are well investigated. This paper proposes to cast them into the generalized Maximum-likelihood (GM)-estimator framework by resorting to the robust statistics [6]. It aims to minimizes the following objective function

$$J(\boldsymbol{x}) = \sum_{i=1}^{m} \omega_i^2 \rho(r_{S_i}), \tag{2}$$

where ω_i is the weight to bound the influenc of bad data, including vertical outliers and bad leverage points; $\rho(\cdot)$ denotes the cost function; $r_{S_i} = r_i / \sigma_i \omega_i$ is the standardized residual; $r_i = z_i - h_i(\widehat{\boldsymbol{x}})$; σ_i is the standard deviation of ith measurement. Depending on the choice of $\rho(\cdot)$, different estimators can

This work is partially sponsored by U.S. National Science Foundation Award ECCS-1711191. The authors are with the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Falls Church, VA 22043, USA (e-mail: zjunbo@vt.edu, lmili@vt.edu).

be obtained. One of the popular choice that generalizes all the four estimators is the Huber convex cost function, which is define as

$$\rho\left(r_{S_i}\right)=\begin{cases} {r_{S_i}}^2/2 & \text{for } |r_{S_i}|\le\lambda \\ \lambda|r_{S_i}|-\lambda^2/2 & \text{for } |r_{S_i}|>\lambda \end{cases}, \quad (3)$$

where λ is the breakpoint that balances the trade-off between least squares and least absolute criterion. Then, we have the following conclusions:

1) The WLS-estimator: when λ tends to infinit and all the weights ω_i are equal to 1, the objective function becomes $J(\boldsymbol{x})=\sum_{i=1}^{m}{r_{S_i}}^2$, which is precisely the criterion used by the WLS estimator. If the Gauss-Newton iterative algorithm is adopted to solve for the state vector, we have

$$\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+\Delta\boldsymbol{x}^k, k=1,2,..., \quad (4)$$

$$\Delta\boldsymbol{x}^k=(\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{h}(\boldsymbol{x}^k)), \quad (5)$$

where $\boldsymbol{H}=\partial\boldsymbol{h}(x)/\partial\boldsymbol{x}|_{\boldsymbol{x}=\boldsymbol{x}^k}\in\mathbb{R}^{m\times n}$ is the Jacobian matrix. The algorithm converges once the norm of $\Delta\boldsymbol{x}^k$ is smaller than a pre-specifie threshold.

2) The LAV-estimator: if λ tends to zero and all the weights ω_i are equal to 1, the objective function reduces to the equivalent form $J(\boldsymbol{x})=\sum_{i=1}^{m}|r_{S_i}|$, which is precisely the criterion used by least absolute value (LAV) estimator.

3) The Huber M-estimator: all the weights ω_i are equal to 1 and the objective function reduces to $J(\boldsymbol{x})=\sum_{i=1}^{m}\rho(r_i/\sigma_i)$, yielding the well-known Huber M-estimator in robust statistics; note that λ can be any value between zero and infinit , but the widely adopted value is chosen between 1.5 and 3 to achieve high statistical efficien y under Gaussian noise [6]–[8];

4) The SHGM-estimator: the weight is calculated through $\omega_i=min[1,\chi^2_{\nu,0.975}/PS_i]$ and the objective function is $J(\boldsymbol{x})=\sum_{i=1}^{m}\omega_i^2\rho(r_i/\sigma_i\omega_i)$, where PS_i is calculated by applying the projection statistics (PS) [9] to the Jacobian matrix $\boldsymbol{H}$ evaluated at fla voltage profile the mathematical representation of PS is shown as follows:

$$PS_i=\max_{\|\boldsymbol{\ell}\|=1}\frac{\left|\boldsymbol{l}_i^T\boldsymbol{\ell}-med_j\left(\boldsymbol{l}_j^T\boldsymbol{\ell}\right)\right|}{1.4826\,med_{\varrho}\left|\boldsymbol{l}_{\varrho}^T\boldsymbol{\ell}-med_j\left(\boldsymbol{l}_i^T\boldsymbol{\ell}\right)\right|}, \quad (6)$$

where $i,j,\varrho=1,2,...,m$. The PS of the ith row vector, $\boldsymbol{l}_i$, of the Jacobian matrix $\boldsymbol{H}$ is define as the maximum of the standardized projections of all the $\boldsymbol{l}_i$'s on every direction $\boldsymbol{\ell}$ that originates from the coordinatewise medians of the $\boldsymbol{H}$ and that passes through every data point, and where the standardized projections are based on the sample median and the median-absolute-deviation [9]. Extensive Monte-Carlo simulations reveal that the PS values follow a χ^2 distribution with ν degree of freedom [9]. The latter is the number of none-zero elements of the ith row of $\boldsymbol{H}$. Note that if σ_i is unknown, it can be estimated by $s=1.4826{\cdot}b_m{\cdot}\text{median}\,|r_i|$, where b_m denotes a correction factor [9]. When the standardized residual of the ith measurement is smaller than the threshold λ, we have $\omega_i^2\rho(r_i/\sigma_i\omega_i)={r_i}^2/2\sigma_i^2$ and it will not be downweighted no matter it is a leverage point or not; otherwise the linear part of the ρ function is used and the measurement will be downweighted by ω_i. As a result, high statistical efficien y can be achieved under Gaussian and other non-Gaussian measurement noise.

Since the LAV-estimator, the Huber M-estimator and the SHGM-estimator are special cases of (2), the key is to fin its general solution. It should be noted that (2) is a convex objective function, its local optimal solution is thus the global one. Therefore, the necessary and sufficien condition that the minimum of (2) satisfie is given by

$$\frac{\partial J\left(\boldsymbol{x}\right)}{\partial\boldsymbol{x}}=\sum_{i=1}^{m}-\frac{\boldsymbol{c}_i\omega_i}{\sigma_i}\psi\left(r_{S_i}\right)=\boldsymbol{0}, \quad (7)$$

where $\boldsymbol{c}_i^T$ is the ith column vector of the Jacobian matrix $\boldsymbol{H}$; $\psi\left(r_{S_i}\right)=\partial\rho\left(r_{S_i}\right)/\partial r_{S_i}$. We multiply and divide both sides of (7) by r_{S_i}, yielding

$$\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{Q}\left(\boldsymbol{z}-\boldsymbol{h}(\boldsymbol{x})\right)=\boldsymbol{0}, \quad (8)$$

where $q\left(r_{S_i}\right)=\psi\left(r_{S_i}\right)/r_{S_i}$ and $\boldsymbol{Q}=\text{diag}\left(q\left(r_{S_i}\right)\right)$. By taking the first-orde Taylor series expansion of $\boldsymbol{h}(\boldsymbol{x})$ about $\widehat{\boldsymbol{x}}^{\ell}$ and using the iteratively reweighted least squares (IRLS) algorithm [6], we obtain the following iterative form:

$$\Delta\widehat{\boldsymbol{x}}^{(\ell+1)}=\left(\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{Q}^{(\ell)}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{Q}^{(\ell)}(\boldsymbol{z}-\boldsymbol{h}(\widehat{\boldsymbol{x}}^{\ell})), \quad (9)$$

where ℓ is the iteration counter. The algorithm converges if $\|\Delta\widehat{\boldsymbol{x}}^{(\ell+1)}\|_\infty\le10^{-3}$. It can be observed from (9) that the iterative process is very similar to that of the Gauss-Newton-based WLS estimator shown in (5) except for the presence of the weight matrix $\boldsymbol{Q}$ that changes at each iteration. Here, we would like to emphasize that ω_i is calculated by applying the PS to the Jacobian matrix assessed at the fla voltage profil and it does not require re-calculation if no topology or measurement configuratio changes.

B. Correlation Between Statistical Robustness and Numerical Robustness

The statistical robustness of an estimator represents its capability to be resistant to outliers or other departures from model assumptions. It can be studied by assessing the total influenc function (IF) of that estimator. As for numerical robustness, the principal concern is the instabilities caused by the proximity to singularities of the gain matrix. To be more general, the iterative algorithm of an estimator should not produce a wildly different result for very small change in the input data. It mainly refers to the effectiveness of inverting the gain matrix $(\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H})$ or $(\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{Q}\boldsymbol{H})$ without stability issues. In this section, we investigate the inherent correlation between the statistical robustness and the numerical robustness of an estimator.

We firs derive the total IF of each estimator and study their statistical robustness. To this end, consider the ϵ-contaminated model $G=(1-\epsilon)\,\Phi+\epsilon\Delta_r$, where Φ is the Gaussian distribution and Δ_r is the probability mass at r that is used to model bad data or the deviation from the assumption, and let the cumulative probability distribution of the residual vector

$\boldsymbol{r} = \boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})$ define as $\Phi(\boldsymbol{r})$. A GM-estimator provides an estimate of the state by solving the following implicit equation:

$$\sum_{i=1}^{m} \boldsymbol{\xi}_i(\boldsymbol{r}, \boldsymbol{x}) = \sum_{i=1}^{m} -\omega_i \frac{\partial \boldsymbol{h}_i(\boldsymbol{x})}{\partial \boldsymbol{x}} \psi(r_{S_i}) = \boldsymbol{0}. \quad (10)$$

By virtue of the Glivenko-Cantelli theorem, (10) can be written asymptotically as

$$\int \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T})\, dG = \boldsymbol{0}, \quad (11)$$

where $\boldsymbol{T}(G)$ is the functional form of the GM-estimator at G. Substituting G into (11) results in

$$\int \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(G)) d\Phi + \epsilon \int \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(G)) d(\Delta_r - \Phi) = \boldsymbol{0}. \quad (12)$$

Knowing that the GM-estimator is Fisher consistent at Φ, that is $\int \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(\Phi)) d\Phi = \boldsymbol{0}$, and taking the derivative of (12) with respect to ϵ, we obtain

$$\frac{\partial}{\partial \epsilon} \int \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(G)) d\Phi \big|_{\epsilon=0} + \int \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(G)) d(\Delta_r)|_{\epsilon=0} = \boldsymbol{0}. \quad (13)$$

Applying the sifting property of the Dirac impulse to the second term yields

$$\frac{\partial}{\partial \epsilon} \int \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(G)) d\Phi \big|_{\epsilon=0} + \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(\Phi)) = \boldsymbol{0}. \quad (14)$$

Using the interchangeability of differentiation and integration theorem to the firs term of (14), we get

$$\int \frac{\partial \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(G))}{\partial \boldsymbol{x}}\Big|_{\boldsymbol{T}(\Phi)} \cdot \frac{\partial \boldsymbol{T}(G)}{\partial \epsilon}\Big|_{\epsilon=0} d\Phi + \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(\Phi)) = \boldsymbol{0}. \quad (15)$$

By using the definitio of the asymptotic total influenc function [10], we derive the IF of $\boldsymbol{T}(G)$ from (14) as

$$\begin{aligned} \boldsymbol{IF}(\boldsymbol{r}, \Phi) &= \frac{\partial \boldsymbol{T}(G)}{\partial \epsilon}\Big|_{\epsilon=0} \\ &= -[\int \frac{\partial \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(G))}{\partial \boldsymbol{x}}\Big|_{T(\Phi)} d\Phi]^{-1} \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(\Phi)), \\ \frac{\partial \boldsymbol{\xi}(\boldsymbol{r}, \boldsymbol{T}(G))}{\partial \boldsymbol{x}} &= -\psi'(r_{S_i}) \boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H}, \end{aligned} \quad (16)$$

where $\boldsymbol{H}$ is the Jacobian matrix. Finally, IF is derived as follows:

$$\boldsymbol{IF}(\boldsymbol{r}, \Phi) = \frac{\psi(r_{S_i})}{\mathbb{E}_\Phi[\psi'(r_{S_i})]} \left(\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H}\right)^{-1} \boldsymbol{c}_i \omega_i, \quad (17)$$

where $\mathbb{E}[\cdot]$ is the expectation operator; $\psi'(r_{S_i}) = \partial \psi(r_{S_i}) / \partial r_{S_i}$; $\boldsymbol{c}_i$ is the ith column vector of $\boldsymbol{H}^T$.

Thus, the total IF of each estimator can be obtained as follows:

- If the weight $\omega_i = 1$ and $\psi(r_{Si}) = r_{Si}$, GM-estimator reduces to the WLS estimator and its IF is

$$\boldsymbol{IF}(\boldsymbol{r}_{Si}, \Phi) = r_{Si} \left(\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H}\right)^{-1} \boldsymbol{c}_i. \quad (18)$$

- If the weight $\omega_i = 1$ and λ tends to zero, GM-estimator reduces to the LAV estimator and its IF is

$$\boldsymbol{IF}(\boldsymbol{r}_{Si}, \Phi) = \frac{1}{2} \text{sign}(r_{Si}) \left(\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H}\right)^{-1} \boldsymbol{c}_i, \quad (19)$$

where sign$(\cdot)$ is the signum function;

- If the weight $\omega_i = 1$, GM-estimator reduces to the Huber M-estimator. In this case, its IF is

$$\boldsymbol{IF}(\boldsymbol{r}_{Si}, \Phi) = \frac{\psi(r_{S_i})}{\mathbb{E}_\Phi[\psi'(r_{S_i})]} \left(\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H}\right)^{-1} \boldsymbol{c}_i. \quad (20)$$

- If the weight ω_i is calculated through the projection statistics, GM-estimator is represented as the SHGM estimator with its IF shown as (17).

It can be easily verifie that both the residual r_{Si} and the position of leverage $\boldsymbol{c}_i$ of the WLS estimator is unbounded in the presence of outliers while the residuals of the LAV and the Huber M-estimator are bounded thanks to the bounded signum function and the ψ function, respectively. However, the position of leverage is unbounded for both LAV and Huber M-estimator if bad leverage points present, yielding unbounded IF. By contrast, thanks to the bounded ψ function as well as the weight ω_i, the IF of the SHGM-estimator is bounded, demonstrating its statistical robustness. Note that vertical outliers only affect the residual r_{Si} while bad leverage points affect both the residual and the position of leverage. In power system, a leverage point is a power injection measurement on a bus with a relatively large number of incident branches compared to the others or a power injection or a power fl w measurement associated with a line having a relatively small reactance compared to the others [9].

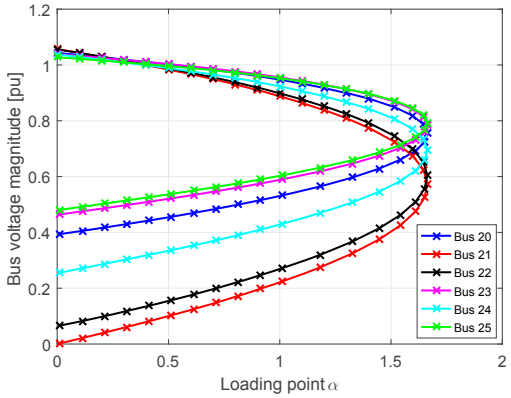

Fig. 1. PV curves of Buses 20-25 for the 30-bus system as the loading level at Bus 21 increases.

It is well-known that when the condition number of the gain matrix $(\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H})$ is very large, the Gauss-Newton approach suffers from numerical instability issue. To address that issue, many numerically stable approaches have been proposed in the literature, among them stands the Levenberg-Marquardt and its variants. These methods are widely used due to their well-proved performance [11], [12]. The key idea is to enhance the condition number of the gain matrix when updating the states through the following trail step:

$$\Delta \boldsymbol{x}^k = (\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H} + \eta \boldsymbol{I})^{-1} \boldsymbol{H}^T \boldsymbol{R}^{-1} (\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}^k)), \quad (21)$$

where $\eta > 0$ is the Levenberg-Marquardt parameter updated at each iteration.

We fin that the weight matrix $\boldsymbol{Q}$ in (9) plays the same role as $\eta \boldsymbol{I}$ in (21). The difference is that $\boldsymbol{Q}$ is adaptively adjusted at each iteration according to the system operation conditions and

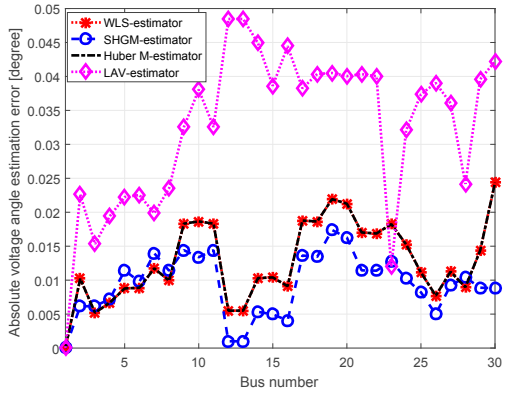

Fig. 2. Voltage angle estimation error of each bus for Case 1.

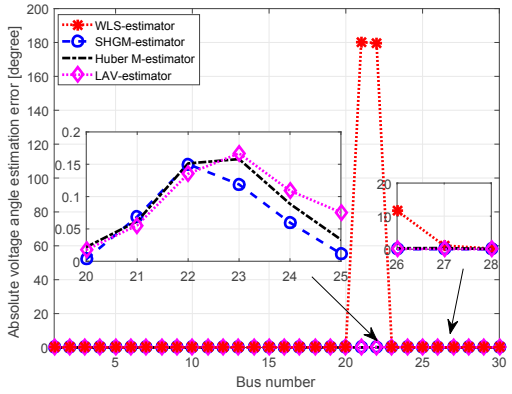

Fig. 3. Voltage angle estimation error of each bus for Case 2.

measurement quality instead of the trail and error step of the Levenberg-Marquardt approach. This provides the theoretical justificatio of the good convergence property of the algorithm that implements the SHGM-estimator to the right solution even when the system is highly stressed while the Gauss-Newton algorithm that implements the WLS (their $\boldsymbol{Q}$ is an identity matrix) fail. On the other hand, as demonstrated in [13], the type of measurements can significantl affect the condition number of the gain matrix. Therefore, the weight ω_i provides the SHGM-estimator better capability to handle stressed system operation conditions than the LAV estimator and the Huber M-estimator, such as leverage points. In summary, the weight ω_i and the bounded ψ function not only guarantee the statistical robustness of an estimator but also enhance the numerical robustness of the iterative algorithm. Here, we would like to emphasize that the statistical robustness of an estimator can typically yield enhanced numerical robustness but not the other way around.

III. Numerical Results

To evaluate the performance of each estimator under various operating conditions, the IEEE 30-bus test system is used as the benchmark. To simulate different stress levels of the system, we increase the load at Bus 21 continuously and use the continuation power fl w (CPF) approach to obtain the PV curves at Buses 20-25 as shown in Fig. 1. Note that Buses 20, 22-25 are adjacent to Bus 21. The obtained voltage magnitudes and angles at each bus are used to calculate the real and reactive power injections and fl ws, then Gaussian noise with zero mean and standard deviation 0.01 is added to simulate realistic measurements. Specificall , this system is measured by 93 SCADA measurements, including 18 pairs of active and reactive power injections, 28 of pairs power fl ws and voltage magnitude of Bus 1. The maximum iteration of all estimators are 30; the breakpoint of the ρ function is 1.5, a typical value used in the literature. 100 Monte Carlo simulations are performed and the average value of the absolute error is taken as the index to evaluate the performance of each estimator.

A. Case Studies

The following three cases are considered and tested:

- Case 1: Normal operation condition, where the load at Bus 21 changes slowly and reaches to the loading point α=0.3; only the estimation error of the voltage angle is shown in Fig. 2 due to space limitation;
- Case 2: Stressed operation condition, where the load at Bus 21 has reached to the loading point α=1.24, demonstrating a stressed operating condition; note that the maximum loading point is α=1.66; the estimation error of the voltage angle is shown in Fig. 3;
- Case 3: Heavily loaded condition, where the load at Bus 21 has reached to the loading point α=1.52, demonstrating a highly stressed operation condition. The simulation results are displayed in Figs. 4-5.

Based on the results, the following conclusions are drawn:

1) Under normal operation condition, the WLS, the Huber M-estimator and the SHGM-estimator have comparative performance while LAV shows lower statistical efficien y. However, this is not surprising because the LAV estimator is the maximum-likelihood estimator under Laplace noise and it only achieves 64% statistical efficien y in presence of Gaussian noise. With the increase of the loading level of the system, the WLS estimator is subject to numerical stability issues, yielding significantl biased estimates of the voltage angles while the LAV estimator, the Huber M-estimator and the SHGM-estimator are able to converge to the right solutions. When the system load level is further increased, the WLS, the LAV estimator and the Huber M-estimator suffer from numerical instabilities, yielding large estimation errors of voltage magnitudes and angles, in contrast with the SHGM-estimator. The reasons behind these observations are the following: the heavily loaded buses will have eigenvalues or singular values associated to them close to zeros, yielding ill-conditioning of the gain matrix. Thus, the Gaussian-Newton is unable to address this issue. By contrast, the SHGM-estimator will automatically detect the measurements associated with the stressed bus as outliers and downweight them, yielding enhanced numerical robustness as analyzed in the previous section. Then, during the iteration step, the effects of the ill-conditioned rows of the Jacobian matrix are eliminated. Note that the variances of the state estimates on a stressed bus may increase because of downweighting the measurements, leading to slightly decreased statistical efficien y. On the other hand, due to the lack of weights ω_i, the capability of the LAV estimator and the Huber M-estimator to downweight

TABLE I
COMPUTING TIMES AND ASSOCIATED NUMBER OF ITERATIONS FOR EACH ESTIMATOR.

Cases	WLS	LAV	Huber M	SHGM
Case 1	0.016s (8)	0.019s (6)	0.027 (7)	0.039s (10)
Case 2	0.032s (18)	0.053s (18)	0.046s (11)	0.068s (15)
Case 3	0.11s (25)	0.092s (24)	0.079s (17)	0.086s (19)

the measurements associated with the stressed bus is limited, preventing it from handling the heavily loaded condition;

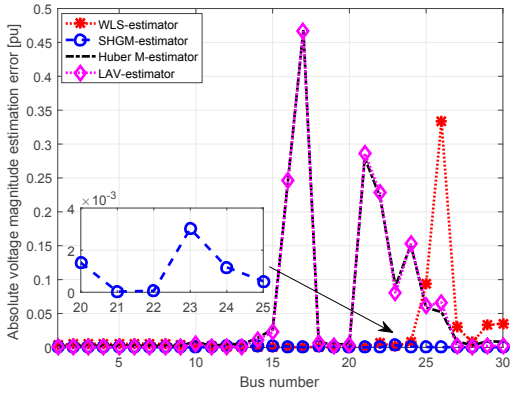

Fig. 4. Voltage magnitude estimation error of each bus for Case 3.

2) It is interesting to fin that under heavy loading conditions, except for the problem of estimating voltage magnitude and angle of Bus 21, its adjacent buses may encounter numerical instabilities as well (see Figs. 4-5). This is due to the fact their states are strongly correlated and if one has severe stability issue, it is likely that the others are affected significantl . On the other hand, it is observed that the leverage points at bus 16 cause numerical problem to the LAV estimator and the Huber M-estimator under highly stressed system condition. This is because their IFs are unbounded for the leverage points and the rounding errors of these measurements cause numerical problem of inverting the gain matrix.

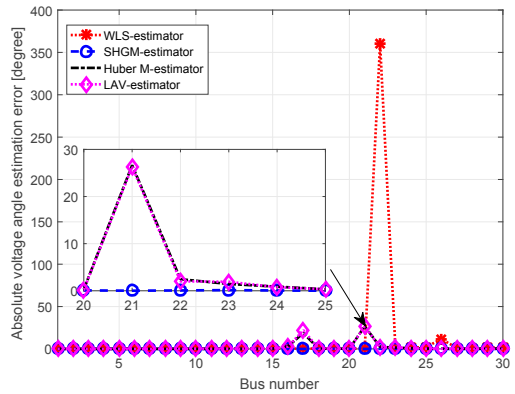

Fig. 5. Voltage angle estimation error of each bus for Case 3.

B. Computational Efficienc

The computing times as well as their associated number of iterations for each estimator to reach the solutions are shown in Table I. All the tests are performed on a PC with Intel Core i5, 2.50 GHz, 8GB of RAM. It is found that under normal operation condition, SHGM-estimator is the most time consuming approach among all four approaches because of the requirement of additional time to calculate measurement weights through PS. However, their difference is acceptable given that SHGM-estimator has robustness to both vertical outliers and bad leverage points. With the increase of system loading level, each approach needs more time and iterations to reach the solution. Interestingly, thanks to the enhanced numerical robustness by the weighting matrix $\boldsymbol{Q}$, the SHGM-estimator, the Huber M-estimator and the LAV-estimator spend less time than the WLS estimator under heavy loading conditions.

IV. CONCLUSION

This paper presents a robust state estimation framework that generalizes several well-known estimators. The statistical robustness of each estimator has been studied analytically through the total influenc function. Its correlation with numerical robustness of the iterative algorithm is investigated as well. The convergence characteristic of each estimator under different system operating conditions is compared. It is found that the SHGM-estimator works well in all scenarios while the other alternatives have convergence problems or suffer from numerical instabilities under stressed system conditions. Future work includes the test of all methods in large-scale systems under various conditions.

REFERENCES

[1] G. Andersson, *et al*., "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1922-1928, 2005.
[2] S. Lefebvre, J. Prevost, H. Horisberger, L. Mili, "Coping with multiple Q-V solutions of the WLS state estimator induced by shunt-parameter errors," *Proceedings of the 8th Conference on Probalistic Method Applied to Power Systems*, 2004.
[3] J. Chen, Y. Liao, B. Gou, K. Tocum, "Impacts of load levels and topology errors on WLS state estimation convergence," North American Power Symposium (NAPS), Champaign, IL, USA, 2012.
[4] J. Chen, Y. Liao, B. Gou, "Study of WLS state estimation convergence characteristics under topology errors," *Proceedings of IEEE Southeastcon*, Jacksonville, FL, USA, 2013.
[5] A. Abur, A. Gomez-Exposito, *Power System State Estimation: Theory and Implementation*, New York, NY: Marcel Dekker, 2004.
[6] P. J. Huber, *Robust Statistics*. New York: John Wiley & Sons, Inc., 1981.
[7] J. B. Zhao, M. Netto, L. Mili, "A robust iterated extended Kalman filte for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3205-3216, 2017.
[8] J. B. Zhao, L. Mili, "A framework for robust hybrid state estimation with unknown measurement noise statistics," *IEEE Trans. Industrail Informatics*, DOI: 10.1109/TII.2017.2764800, 2017.
[9] L. Mili, M. Cheniae, N. Vichare, and P. Rousseeuw, "Robust state estimation based on projection statistics," *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 1118–1127, 1996.
[10] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, W. A. Stahel, *Robust Statistics: The Approach Based on Influenc Functions*. New York, NY, USA: Wiley, 2011.
[11] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, *Q. Appl. Math.*, vol. 2, no. 2, pp. 164-166, 1944.
[12] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear inequalities, *SIAM J. Appl. Math.*, vol. 11, no. 2, pp. 431-441, 1963.
[13] R. Ebrahimian, R. Baldick, "State estimator condition number analysis," *IEEE Trans. Power Syst.*, vol. 16, no. 2, pp. 273-279, 2001.