

Robust Unscented Kalman Filter for Power System Dynamic State Estimation with Unknown Noise Statistics

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

Abstract—Due to the communication channel noises, GPS synchronization process, changing environment temperature and different operating conditions of the system, the statistics of the system process and measurement noises may be unknown and they may not follow Gaussian distributions. As a result, the traditional Kalman filter-based dynamic state estimator (DSE) may provide strongly biased state estimates. To address these issues, this paper develops a robust Generalized Maximum-likelihood Unscented Kalman Filter (GM-UKF). The statistical linearization approach is presented to derive a compact batch-mode regression form by processing the predicted state vector and the received measurements simultaneously. This regression form enhances the data redundancy and allows us to detect bad PMU measurements and incorrect state predictions, and filter out unknown Gaussian and non-Gaussian noises through the generalized maximum likelihood (GM)-estimator. The latter minimizes a convex Huber function with weights calculated via the projection statistics (PS). Particularly, the PS is applied to a proposed 2-dimensional matrix that consists of temporally correlated innovation vectors and predicted states. Finally, the total influence function is used to derive the error covariance matrix of the GM-UKF state estimates, yielding the robust state prediction at the next time instant. Extensive simulations carried out on the IEEE 39-bus test system demonstrate the effectiveness and robustness of the proposed method.

Index Terms—Dynamic state estimation, robust estimation, unscented Kalman filter, non-Gaussian noise, total influence function, bad data.

I. INTRODUCTION

RELIABLE and accurate dynamic state variables of the synchronous machines are of paramount importance for power system real-time monitoring and control. For example, the rotor speeds can be taken as inputs by power system stabilizers to enhance small-signal stability [1] while the rotor angles can be leveraged for the design of generator out-of-step relays [2], and so on. However, most of the machine dynamic state variables are not directly measured. Thanks to the widespread deployment of Phasor Measurement Units (PMUs) on power transmission grids, it becomes possible to utilize synchronized PMU measurements to estimate the full state variables of a synchronous generator by using a Kalman filter-based dynamic state estimator (DSE) [3]–[13].

In [3]–[6], the extended Kalman filter (EKF) is used to estimate the state variables of a synchronous generator with

different model orders. However, the first-order Taylor series expansion-based EKF will produce large estimation errors or even may diverge when the system nonlinearity is strong. The latter can be caused by severe disturbances or heavy loading conditions that significantly stress the system. To circumvent first-order approximation errors of the EKF, several other nonlinear filter have been proposed and applied to power system dynamic state estimation, such as the unscented Kalman filter (UKF) [7]–[11], the ensemble Kalman filter [12], the particle filter [13], to name a few.

The Kalman-type filter work well only under certain conditions [14]–[16]. Firstly, the means and covariance matrices of the system process noise w_k and measurement noise v_k are assumed to be known at each time instant, which are Q_k and R_k , respectively. This is because the Kalman-type filter require the knowledge of Q_k and R_k for the calculation of the prediction and filtering covariance matrices. Thirdly, the attractiveness of the Kalman filter stems from its optimality under the Gaussian assumption. Finally, the system model is assumed to be known exactly. However, for practical dynamical systems, these assumptions might not hold true given that 1) Q_k and R_k may be difficult to obtain; this is because both noises depend heavily on the actual operating conditions of the system and are changing from time to time; 2) the process and observation noise may not follow Gaussian distributions. This is demonstrated by two recent investigations conducted by Pacific Northwest National Laboratory (PNNL) [17], [18], where the PMU measurement errors of the voltage and current phasors obey non-Gaussian probability distributions; 3) the vector-valued state transition and measurement functions are with uncertainties, yielding unknown system process noise; 4) the received measurements can be biased significantly due to impulsive communication noise, cyber attacks, etc, inducing bad PMU measurements. Under these conditions, the aforementioned Kalman filter-based DSEs may obtain significantly biased estimation results.

To address these issues, this paper develops a robust Generalized Maximum-likelihood Unscented Kalman Filter (GM-UKF). It has the following salient features:

- The statistical linearization approach is presented to derive a compact batch-mode regression form by processing in a robust way the predicted state vector and the received measurements simultaneously, yielding a robust filter
- Projection statistics are extended to reliably detect observation and innovation outliers by analyzing the temporal correlations of a 2-dimensional innovation matrix;

This work is supported in part by the U.S. National Science Foundation under Grant ECCS-1711191.

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

- The proposed GM-UKF is able to filter out unknown Gaussian and non-Gaussian noises and to suppress observation and innovation outliers while achieving high statistical efficiency; note that the observation outliers (or bad data) refer to the received PMU measurements with large errors induced by impulsive communication noise or loss of communications, among others. As for the innovation outliers, they may be induced by the failure of the automatic voltage controller or the power system stabilizers or by impulsive system process noise, yielding incorrect predicted state variables;
- The total influence function is derived and the expression of the estimation error covariance matrix of the GM-UKF state estimates is calculated, allowing us to perform a robust state prediction at the next time step.

The rest of the paper is organized as follows. Section II presents the problem formulation. Section III develops the proposed GM-UKF. Section IV shows and analyzes the simulation results. Finally Section V concludes the paper.

II. PROBLEM FORMULATION

In this section, the nonlinear discrete-time state space model of a synchronous generator is first derived. Then, the statistical model of the system process and measurement noises is discussed. Finally, the problem statement is presented.

A. Nonlinear Discrete-Time State Space Model

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

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{w}_k, \quad (1)$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k, \quad (2)$$

where $\mathbf{x}_k$ is the state vector; $\mathbf{f}(\cdot)$ is the vector-valued function that relates $\mathbf{x}_k$ to $\mathbf{x}_{k-1}$; $\mathbf{u}_k$ represents the system input vector; $\mathbf{z}_k$ is the measurement vector; $\mathbf{h}(\cdot)$ is the vector-valued measurement function; $\mathbf{w}_k$ and $\mathbf{v}_k$ are system process and measurement noises, respectively, which are not necessarily assumed to be Gaussian; their covariance matrices are denoted by $\mathbf{Q}_k$ and $\mathbf{R}_k$, respectively. The above discrete-time state space form is usually derived from the differential and algebraic equations that govern the system dynamics. The reader is referred to Appendix A for details. The objective of this paper is to develop a robust filter method aimed at estimating the dynamic state variables of the synchronous generator.

B. Statistical Model of Noises

Due to the communication channel noises, GPS synchronization process, changing environment temperature and different operating conditions of the system, the statistics of the PMU measurement and system process noises may be unknown to the control center and they may not follow Gaussian distributions. For example, experiments conducted by PNNL [17], [18] reveal that the PMU measurement noise follows a non-Gaussian distribution. There are several models that can be applied to model deviations from Gaussian assumption. Among them, the Gaussian sum distribution is widely used because any non-Gaussian distribution $p(\mathbf{x})$ can be expressed as, or approximated sufficiently well, by a finite sum of known

Gaussian densities according to the Wiener approximation theorem [20], i.e.,

$$p(\mathbf{x}) = \sum_{i=1}^{N_A} a_i \mathcal{N}(\bar{\mathbf{x}}_i, \Sigma_i), \quad (3)$$

where a_i is the weight and $\sum_{i=1}^{N_A} a_i = 1$; N_A is the number of Gaussian components; the mean and covariance matrix associated with the i th Gaussian component are denoted by $\bar{\mathbf{x}}_i$ and Σ_i , respectively.

A popular variant of (3) used in robust statistics is the ϵ -contaminated model [21]. It models the full neighborhoods of a Gaussian distribution in the probability space and is given by the following functional form:

$$G = (1 - \epsilon)\Phi + \epsilon K, \quad (4)$$

where G denotes the probability distribution of the random vector $\mathbf{x}$; Φ denotes the Gaussian probability distribution; K denotes an unknown probability distribution, which can be a heavy-tailed distribution such as the Laplace distribution. Here, $\epsilon \in [0, 0.5)$ regulates the level of the non-Gaussian contamination, e.g., a small ϵ indicates a small fraction of non-Gaussian errors. Note that as long as $\epsilon \neq 0$, G is a non-Gaussian distribution, which is usually the case for practical power systems.

Remark: The ϵ -contaminated model is based on the assumption that the majority of the measurement error roughly follows a Gaussian distribution, a widely accepted assumption in the statistical literature. It is a convenient mathematical model that allows us to derive the bias-sensitivity of the estimator to infinitesimal contamination and to investigate the maximum fraction of outliers that the estimator can withstand, termed the breakdown point.

C. Problem Statement

To highlight the significance of the contributions of this paper, we list the following open problems related to power system dynamic state estimation that are pinpointed in the current literature:

- Experiments conducted by PNNL reveal that PMU noises follow non-Gaussian distributions, including Laplace and Cauchy distributions. However, existing power system dynamic state estimation approaches, including [7]–[10], assumes that both system process and measurement noises are Gaussian. As a result, the estimated states that they provide are significantly biased;
- Under the assumption that both system process and measurement noises follow a Gaussian distribution, a normalized innovation-based outlier detection test [8] may be applied. However, the detection threshold is system dependent and challenging to set in presence of unknown non-Gaussian noise distributions. In addition, it also supposes that the system model is exact, which is not true in presence of innovation outliers. It turns out that this type of outliers are rarely discussed in the literature despite the fact that they frequently occur in practice due to model parameter errors, controller failures, to cite a few;

- A robust state prediction covariance matrix to outliers along with a reformulation of the Kalman filter as a batch-mode regression problem is required to derive a robust dynamic state estimator in power systems.

In this paper, we propose a robust GM-UKF to address all the above challenges.

III. THE PROPOSED ROBUST UKF

In this section, the four main steps of the proposed robust UKF are presented, namely a batch-mode regression form step, a robust outlier detection step, a robust regression step, and a robust error covariance matrix updating step. Then, some practical implementation issues are discussed.

A. Batch-Mode Regression Step

Similar to the traditional UKF, $2n$ sigma points with weights $w_i, i = 1, \dots, 2n$ are generated to capture the statistics of the state estimate $\hat{\mathbf{x}}_{k-1|k-1} \in \mathbb{R}^{n \times 1}$ at time step $k-1$. Formally, we have [22]

$$\chi_{k-1|k-1}^i = \hat{\mathbf{x}}_{k-1|k-1} \pm \left(\sqrt{n \Sigma_{k-1|k-1}^{xx}} \right)_i, \quad (5)$$

and $w_i = \frac{1}{2n}, i = 1, \dots, 2n$; $\Sigma_{k-1|k-1}^{xx}$ is the covariance matrix of $\hat{\mathbf{x}}_{k-1|k-1} \in \mathbb{R}^{n \times 1}$. Note that the least number of sigma points is $n+1$, but $2n$ is the recommended number in the literature [22]. By substituting each sigma point into the nonlinear system process model (1), we obtain the following transformed sigma points:

$$\chi_{k|k-1}^i = \mathbf{f}(\chi_{k-1|k-1}^i). \quad (6)$$

Then, the weighted sample mean and sample covariance matrix of these transformed sigma points are used to approximate the predicted state vector and its covariance matrix by

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \sum_{i=1}^{2n} w_i \chi_{k|k-1}^i, \\ \Sigma_{k|k-1}^{xx} &= \sum_{i=1}^{2n} w_i (\chi_{k|k-1}^i - \hat{\mathbf{x}}_{k|k-1})(\chi_{k|k-1}^i - \hat{\mathbf{x}}_{k|k-1})^T + \mathbf{Q}_k. \end{aligned} \quad (7)$$

To construct the regression form, we define

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{x}_k - \mathbf{e}_k, \quad (8)$$

where $\mathbf{x}_k$ is the true state vector; $\mathbf{e}_k$ is the prediction error; and $\mathbb{E}[\mathbf{e}_k \mathbf{e}_k^T] = \Sigma_{k|k-1}^{xx}$. By applying the statistical linearization [23] to the nonlinear measurement equation around $\hat{\mathbf{x}}_{k|k-1}$, we obtain

$$\mathbf{z}_k = \mathbf{H}_k (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{v}_k + \boldsymbol{\varepsilon}_k, \quad (9)$$

where $\mathbf{H}_k = (\Sigma_{k|k-1}^{xx})^T (\Sigma_{k|k-1}^{zz})^{-1}$ is the statistical regression matrix; $\Sigma_{k|k-1}^{zz}$ is the cross-covariance matrix defined as

$$\Sigma_{k|k-1}^{zz} = \sum_{i=1}^{2n} w_i (\chi_{k|k-1}^i - \hat{\mathbf{x}}_{k|k-1})(\mathbf{z}_{k|k-1}^i - \hat{\mathbf{z}}_{k|k-1})^T,$$

where $\hat{\mathbf{z}}_{k|k-1} = \sum_{i=1}^{2n} w_i \mathbf{z}_{k|k-1}^i$ is the predicted measurement vector and $\mathbf{z}_{k|k-1}^i = \mathbf{h}(\chi_{k|k-1}^i)$; $\boldsymbol{\varepsilon}_k$ is the statistical linearization error used to compensate the system nonlinearity; its

covariance matrix is

$$\tilde{\mathbf{R}}_k = \mathbb{E}[\boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_k^T] = \Sigma_{k|k-1}^{zz} - (\Sigma_{k|k-1}^{xz})^T \Sigma_{k|k-1}^{xx} \Sigma_{k|k-1}^{xz}, \quad (10)$$

where the self-covariance matrix $\Sigma_{k|k-1}^{zz}$ is calculated by

$$\Sigma_{k|k-1}^{zz} = \sum_{i=1}^{2n} w_i (\mathbf{z}_{k|k-1}^i - \hat{\mathbf{z}}_{k|k-1})(\mathbf{z}_{k|k-1}^i - \hat{\mathbf{z}}_{k|k-1})^T + \mathbf{R}_k. \quad (11)$$

By processing the predicted state and the measurements simultaneously, we obtain the following batch-mode regression form:

$$\begin{bmatrix} \mathbf{z}_k + \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}) \\ \hat{\mathbf{x}}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_k \\ \mathbf{I} \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} \mathbf{v}_k + \boldsymbol{\varepsilon}_k \\ -\mathbf{e}_k \end{bmatrix} \quad (12)$$

which can be rewritten into a compact form as

$$\tilde{\mathbf{z}}_k = \tilde{\mathbf{H}}_k \mathbf{x}_k + \tilde{\mathbf{e}}_k. \quad (13)$$

The error covariance matrix is given by

$$\Sigma_k = \mathbb{E}[\tilde{\mathbf{e}}_k \tilde{\mathbf{e}}_k^T] = \begin{bmatrix} \Sigma_{k|k-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{k|k-1}^{xx} \end{bmatrix} = \mathbf{S}_k \mathbf{S}_k^T, \quad (14)$$

where $\Sigma_{k|k-1} = \mathbb{E}[(\mathbf{v}_k + \boldsymbol{\varepsilon}_k)(\mathbf{v}_k + \boldsymbol{\varepsilon}_k)^T] = \mathbf{R}_k + \tilde{\mathbf{R}}_k$; $\mathbf{I}$ is an identity matrix; $\mathbf{S}_k$ is calculated by the Cholesky decomposition technique.

Due to the nonlinear system state transition and measurement functions, the predicted state vector and measurements are correlated, yielding non-diagonal covariance matrix Σ_k . To uncorrelate them, we pre-multiply $\mathbf{S}_k^{-1}$ on both sides of (14), yielding

$$\mathbf{S}_k^{-1} \tilde{\mathbf{z}}_k = \mathbf{S}_k^{-1} \tilde{\mathbf{H}}_k \mathbf{x}_k + \mathbf{S}_k^{-1} \tilde{\mathbf{e}}_k, \quad (15)$$

which can be put into the compact form given by

$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \boldsymbol{\xi}_k, \quad (16)$$

where $\mathbb{E}[\boldsymbol{\xi}_k \boldsymbol{\xi}_k^T] = \mathbf{I}$ can be easily verified

B. Robust Outlier Detection Step

For practical power system, the received PMU measurements can be significantly biased because of impulsive communication noise, loss of communications, etc., yielding observation outliers. For the innovation outliers, they may be induced in several different ways, such as incorrect generator parameters due to saturations, failure of controllers (exciter, power system stabilizer), impulsive system process noise, yielding incorrect predicted state variables. To detect them, we apply the Projection Statistics (PS) to the following 2-dimensional matrix $\mathbf{Z}_k$:

$$\mathbf{Z}_k = \begin{bmatrix} \mathbf{z}_{k-1} - \mathbf{h}(\hat{\mathbf{x}}_{k-1|k-2}) & \mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}) \\ \hat{\mathbf{x}}_{k-1|k-2} & \hat{\mathbf{x}}_{k|k-1} \end{bmatrix}, \quad (17)$$

where $\mathbf{z}_{k-1} - \mathbf{h}(\hat{\mathbf{x}}_{k-1|k-2})$ and $\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1})$ are the innovation vectors while $\hat{\mathbf{x}}_{k-1|k-2}$ and $\hat{\mathbf{x}}_{k|k-1}$ are the predicted state vectors at time instants $k-1$ and k , respectively. The choice of $\mathbf{Z}_k$ is motivated by the fact that the innovation vector and the predicted state vector are time series samples of power system responses and have strong temporal correlations. If outliers occur, this relationship is violated. Thus, by checking

this statistical property of the matrix $\mathbf{Z}_k$, we are able to detect observation and innovation outliers. Note that in $\mathbf{Z}_k$, innovation vector consists of values around zeros while the values of the predicted states are close to the system state variables that are different from zeros. As a result, they are clustered around different points and need to be processed separately by the PS. The PS is define as [25]:

$$PS_j = \max_{\|\ell\|=1} \frac{|\mathbf{l}_j^T \ell - \text{medi}_i(\mathbf{l}_i^T \ell)|}{1.4826 \text{medi}_\kappa(|\mathbf{l}_\kappa^T \ell - \text{medi}_i(\mathbf{l}_i^T \ell)|)}, \quad (19)$$

for $i, j, \kappa = 1, 2, \dots, m+n$, where $\mathbf{l}_j^T$, $\mathbf{l}_i^T$ and $\mathbf{l}_\kappa^T$ are the j th, i th and κ th row vector of $\mathbf{Z}_k$, respectively. The detailed implementation steps of PS are shown in the Appendix B. Note that innovation vectors and predicted state vectors define above are random variables roughly obeying a Gaussian distribution, yielding a bi-variate Gaussian $\mathbf{Z}_k$. As a result, the PS values that are calculated based on the distribution of $\mathbf{Z}_k$ follow a chi-squares distribution with 2-degree of freedom as shown by extensive Monte Carlo simulations. This allows us to derive a statistical test to detect outliers. Formally, we have

$$\begin{cases} \mathcal{H}_0 : |\Pi| = 0 \\ \mathcal{H}_1 : 1 \leq |\Pi| \leq m+n \end{cases}, \quad (20)$$

where $\Pi = \{\text{PS}_i > \chi_{2,0.975}^2, i = 1, \dots, m+n\}$; $|\Pi|$ represents the cardinality of the set Π ; $\mathcal{H}_1$ and $\mathcal{H}_0$ correspond to occurrence of outliers and no outliers, respectively. Note that the detection threshold is set to be $\chi_{2,0.975}^2$ because the PS values follow a Chi-square distribution with 2 degree of freedom [5]. The detected outliers are downweighted via

$$\varpi_i = \min(1, d^2 / PS_i^2), \quad (21)$$

where the parameter d is usually set equal to 1.5 to yield good statistical efficiency at different distributions without increasing too much the bias induced by the outliers.

C. Robust Regression Step

After the detection of the outliers, the Weighted Least Squares (WLS) estimator can be applied to the regression model (17) to obtain the state estimates. However, it is well-known that the WLS is only optimal under Gaussian assumption with known variances, which is not the case for practical measurement noise. To filter out non-Gaussian measurement noise, we propose to use a robust Generalized-Maximum Likelihood (GM)-type estimator. It aims to minimize an objective function given by

$$J(\mathbf{x}_k) = \sum_{i=1}^{m+n} \varpi_i^2 \rho(r_{S_i}), \quad (22)$$

where ϖ_i is calculated by (21); $r_{S_i} = r_i / s\varpi_i$ is the standardized residual; $r_i = y_i - \mathbf{c}_i^T \hat{\mathbf{x}}$ is the residual, where $\mathbf{c}_i^T$ is the i th row vector of the matrix $\mathbf{C}_k$; $s = 1.4826 \cdot b_m \cdot \text{median}_i |r_i|$ is the robust scale estimate; 1.4826 is a correction factor chosen to ensure consistency of the proposed method under the Gaussian distribution [25]; b_m is a correction factor to achieve unbiasedness for a finite sample of size $m+n$ at a

given probability distribution [25]; $\rho(\cdot)$ is the convex Huber- ρ function [21] expressed as

$$\rho(r_{S_i}) = \begin{cases} \frac{1}{2} r_{S_i}^2, & \text{for } |r_{S_i}| < \lambda \\ \lambda |r_{S_i}| - \lambda^2/2, & \text{elsewhere} \end{cases}, \quad (23)$$

where the parameter λ is typically chosen between 1.5 to 3 in the literature.

Thanks to the convexity of the Huber- ρ function, the solution to an iterative algorithm will be a global minimum. Thus, the necessary and sufficient condition for optimality of objective function (22) is

$$\frac{\partial J(\mathbf{x}_k)}{\partial \mathbf{x}_k} = \sum_{i=1}^{m+n} -\frac{\varpi_i \mathbf{c}_i}{s} \psi(r_{S_i}) = \mathbf{0}, \quad (24)$$

where $\psi(r_{S_i}) = \partial \rho(r_{S_i}) / \partial r_{S_i}$.

As $\psi(\cdot)$ is a nonlinear function, an iterative algorithm has to be used to obtain the global minimum. In this paper, we advocate the use of the Iteratively Reweighted Least Squares (IRLS) algorithm [24], [26]. To this end, we divide and multiply the standardized residual r_{S_i} to both sides of (24), yielding the following matrix form

$$\mathbf{C}_k^T \mathbf{W} (\mathbf{y}_k - \mathbf{C}_k \mathbf{x}_k) = \mathbf{0}, \quad (25)$$

where $\mathbf{W} = \text{diag}(q(r_{S_i}))$ and $q(r_{S_i}) = \psi(r_{S_i}) / r_{S_i}$. Then, the state vector at the j iteration is calculated through

$$\Delta \hat{\mathbf{x}}_{k|k}^{(j+1)} = \left(\mathbf{C}_k^T \mathbf{W}^{(j)} \mathbf{C}_k \right)^{-1} \mathbf{C}_k^T \mathbf{W}^{(j)} \mathbf{y}_k, \quad (26)$$

where $\Delta \hat{\mathbf{x}}_{k|k}^{(j+1)} = \hat{\mathbf{x}}_{k|k}^{(j+1)} - \hat{\mathbf{x}}_{k|k}^{(j)}$. The algorithm converges when $\|\Delta \hat{\mathbf{x}}_{k|k}^{(j+1)}\|_\infty \leq 10^{-2}$.

D. Robust Error Covariance Matrix Updating Step

After the robust regression step, the error covariance matrix $\Sigma_{k|k}^{xx}$ is updated so that the state prediction for the next step can be performed. In our previous work [5], the total influence function-based error covariance matrix updating approach is used. In this paper, we follow that work and derive the asymptotic error covariance matrix of our GM-UKF at time sample k as

$$\Sigma_{k|k}^{xx} = \frac{\mathbb{E}_\Phi[\psi^2(r_{S_i})]}{\{\mathbb{E}_\Phi[\psi'(r_{S_i})]\}^2} (\mathbf{C}_k^T \mathbf{C}_k)^{-1} (\mathbf{C}_k^T \mathbf{Q}_\varpi \mathbf{C}_k) (\mathbf{C}_k^T \mathbf{C}_k)^{-1} \quad (27)$$

where $\mathbf{Q}_\varpi = \text{diag}(\varpi_i^2)$; the coefficient $\frac{\mathbb{E}_\Phi[\psi^2(r_{S_i})]}{\{\mathbb{E}_\Phi[\psi'(r_{S_i})]\}^2}$ can be shown to be 1.0369 if $\lambda = 1.5$.

E. Practical Implementation Issues of the GM-UKF

There are several parameters of the proposed GM-UKF that need to be determined, namely the settings of the breakpoint λ of the Huber ρ -function, the parameter d of the weighting function, and the convergence tolerance of the IRLS algorithm. λ determines the trade-off that we wish to achieve between a least-squares and a least-absolute-value fit. Indeed, when $\lambda \rightarrow 0$, the Huber ρ -function tends to the least-absolute-value ρ -function and when $\lambda \rightarrow \infty$, it tends to the least-squares ρ -function. Regarding the parameter d , it determines the statistical efficiency of the PS at the assumed probability distribution along with the robustness of the GM-estimator [24].

Decreasing this parameter too much shrinks the dimensions of the 97.5% confidence ellipse. As a result, good measurements may be unduly downweighted, which yields a decrease in the statistical efficiency. On the other hand, increasing d will increase the bias of the GM-estimator. Extensive simulations have shown that the parameters λ and d can be set to 1.5 to achieve a good statistical efficiency at the Gaussian and many other non-Gaussian distributions while achieving a good robustness to outliers. Regarding the convergence tolerance threshold of the IRLS algorithm, a typical value is 0.01; decreasing this value results in small incremental changes of the state estimates while increasing the computing time of the algorithm.

Another important problem that needs to be addressed is how to properly choose the covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$ for the system process and measurement noises. Since the classes of precisions of the PMU devices are provided, we can use them to derive the variances of the measurement noise, which are the diagonal elements of covariance matrix $\mathbf{R}_k$. It is interesting to note that the robust GM-UKF will identify those data points that are within the given standard deviations from the regression surface that fit the majority of the data points and mainly relies on them to provide the best state estimates while suppressing the outliers. Now, how to set up the diagonal elements of the covariance matrix $\mathbf{Q}_k$? If they are chosen smaller than those associated with $\mathbf{R}_k$, the Kalman filter-based DSE will automatically put more emphasis on the predicted state vector for state estimation; by contrast, if they are chosen larger, the filter will rely more on the measurements, and consequently, the estimation results may be biased under non-Gaussian noise. To address this issue, we propose to assign similar values to the diagonal elements of $\mathbf{Q}_k$ as those of $\mathbf{R}_k$ with appropriate dimensions and let the GM-UKF balance the tradeoff between state predictions and measurements for state estimation. As a result, our GM-UKF does not require the exact information of the noise statistics, which relaxes the general assumption of a Kalman filter-based DSE.

IV. NUMERICAL RESULTS

Extensive simulations are carried out on the 10-machine IEEE 39-bus New England system to assess the effectiveness and robustness of the proposed GM-UKF. It is assumed that a large system disturbance occurs at $t=0.5s$ by opening the transmission line between Buses 15 and 16. The time-domain simulation results are used to generate a collection of samples of the nodal voltage magnitudes and phase angles as well as of the real and reactive power injections at the terminal buses of all the generators. All parameter values of the two-axis generator model are taken from an IEEE report [27]. The PMU scan rate is assumed to be 50 samples/second. For the state initialization, the steady-state power flow solutions with 10% errors are used. Due to space limitation, not all the 9 state variables of each generator are shown; instead estimated values of the rotor angle and speed, the field voltage and the mechanical power of Generator 5 are utilized for illustration purposes. 100 Monte Carlo simulations at each PMU sample are conducted and the mean values are considered as the final

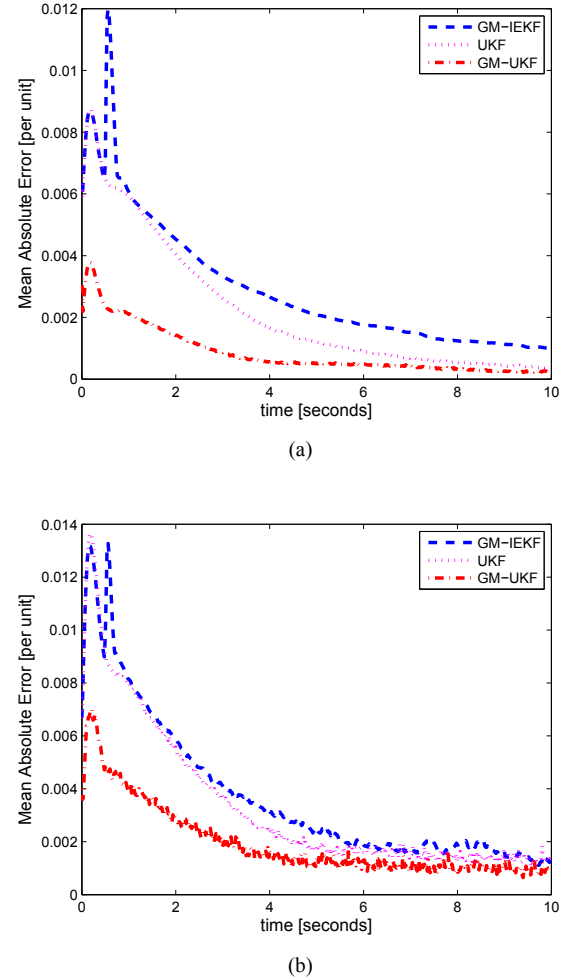
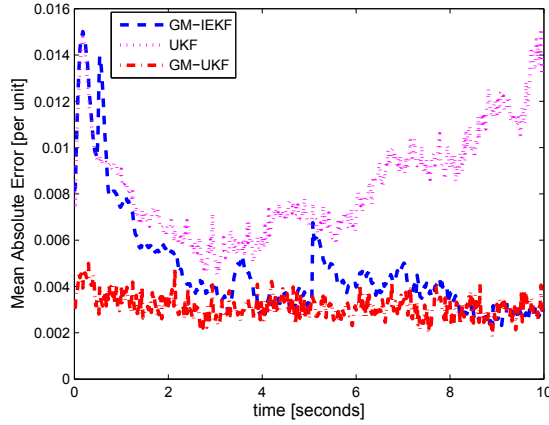


Fig. 1. Scenario 1: Mean absolute error of the GM-IEKF, the UKF, and the GM-UKF in presence of Gaussian noise; (a) small Gaussian noise with zero mean and covariance matrix $10^{-6}\mathbf{I}$; (b) large Gaussian noise with zero mean and covariance matrix $10^{-4}\mathbf{I}$.

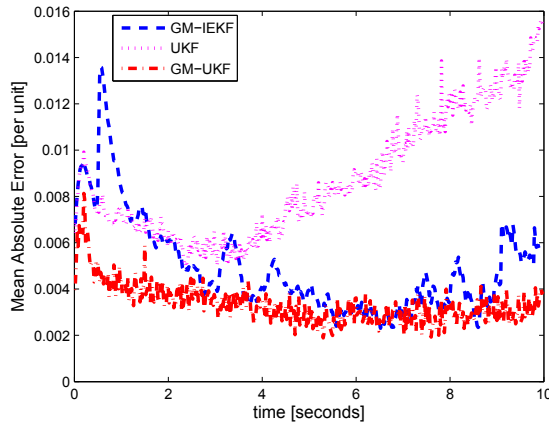
solution. The Mean Absolute Error (MAE) is utilized as the index to evaluate the overall performance of each method. The traditional UKF and the GM-IEKF [5] are included for comparisons. The diagonal elements of the system process and measurement noise covariance matrices are assumed to be 10^{-6} .

A. Scenario 1: Occurrence of Small and Large Gaussian Noise

In the scenario 1, we assume that both the system process and measurement noises follow a Gaussian distribution. Although this is a rare case for practical power system, we would like to compare the performances of the GM-IEKF, the UKF, and the GM-UKF under this ideal condition. Please note that in the existing literature about Kalman filter-based power system DSEs, Gaussian assumption is explicitly assumed for the system process and measurement noises. Specifically, a Gaussian random variable with zero mean and covariance matrix $10^{-6}\mathbf{I}$ is added to simulate system process and measurement noises with appropriate dimensions. The test results are displayed in Fig. 1-(a). It is observed from this figure that under this ideal condition, the UKF outperforms the GM-IEKF as the



(a)



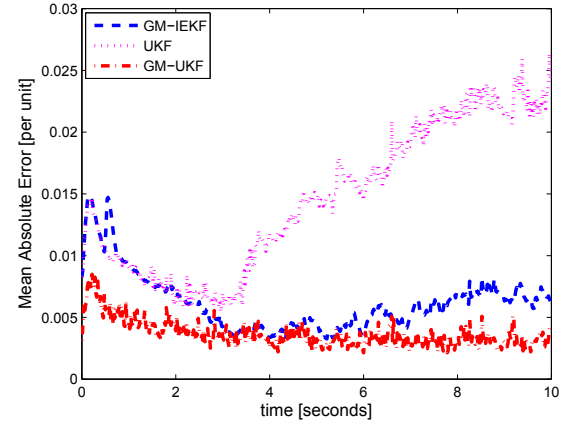
(b)

Fig. 2. Scenario 2: Mean absolute error of the GM-IEKF, the UKF, and the GM-UKF in presence of non-Gaussian noise that is represented by Gaussian mixture model; (a) 5% of the data is contaminated by another Gaussian component with zero mean and covariance matrix $10^{-5}\mathbf{I}$; (b) 10% of the data is contaminated by another Gaussian component with zero mean and covariance matrix $10^{-5}\mathbf{I}$. Note that the true covariance matrix is $10^{-6}\mathbf{I}$.

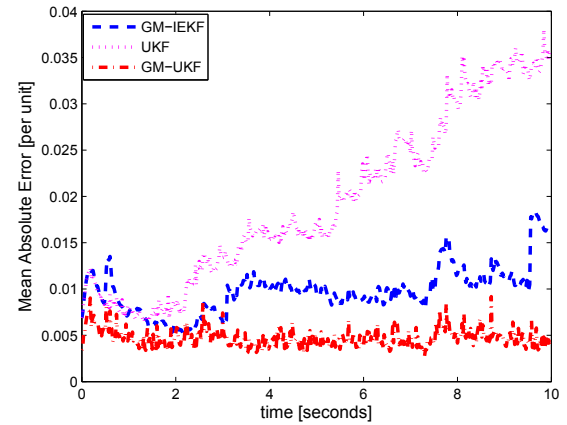
former is a better nonlinear filter than the EKF. However, their statistical efficiency is much lower than our GM-UKF. When the covariance matrices are increased from $10^{-6}\mathbf{I}$ to $10^{-4}\mathbf{I}$, the comparison results of three filter are shown in Fig. 1-(b). It should be noted that this is not the ideal condition anymore as the covariance matrices of system process and measurement noises are assumed to be $10^{-6}\mathbf{I}$ with appropriate dimensions while their true values are $10^{-4}\mathbf{I}$. It is observed from Fig. 1-(b) that the GM-IEKF achieves comparative performance as the UKF. Our GM-UKF outperforms the GM-IEKF and the UKF, yielding the highest statistical efficiency.

B. Scenario 2: Occurrence of Small Non-Gaussian Noise

In practice, both the system process and measurement noises are unknown and may deviate from the Gaussian assumption. This is because they depend on the operating condition of a system and may change from time to time. In this section, we test the scenario that the deviation from Gaussian assumption is small. To do that, a bimodal Gaussian mixture model with zero means, covariance matrices of $10^{-6}\mathbf{I}$ and $10^{-5}\mathbf{I}$ and weights of 0.95 and 0.05, are assumed for both the system



(a)



(b)

Fig. 3. Scenario 3: Mean absolute error of the GM-IEKF, the UKF, and the GM-UKF in presence of non-Gaussian noise that is represented by Gaussian mixture model; (a) 5% of the data is contaminated by another Gaussian component with zero mean and covariance matrix $10^{-4}\mathbf{I}$; (b) 10% of the data is contaminated by another Gaussian component with zero mean and covariance matrix $10^{-4}\mathbf{I}$. Note that the true covariance matrix is $10^{-6}\mathbf{I}$.

process and measurement noises with appropriate dimensions. The test results are shown in Fig. 2-(a). Compared with the Gaussian noises scenario, the state estimates provided by UKF are significantly biased as it is unable to filter out non-Gaussian noise. By contrast, both GM-IEKF and GM-UKF are capable of filtering out non-Gaussian noises thanks to their statistical robustness provided by the GM-estimator. However, GM-UKF achieves better statistical efficiency than GM-IEKF. When the deviation from Gaussian assumption increases slightly, that is, 10% of the data is contaminated by another Gaussian component with zero mean and covariance matrix $10^{-5}\mathbf{I}$ (this is simulated by changing the weights 0.95 and 0.05 to 0.9 and 0.1 for the Gaussian mixture model, respectively), the estimation results are displayed in Fig. 2-(b). It is found that with increased degree of the deviation from Gaussian assumption, estimation errors of all three filter increase. However, our GM-UKF still has a comparative performance as the previous case and it achieves the highest statistical efficiency.

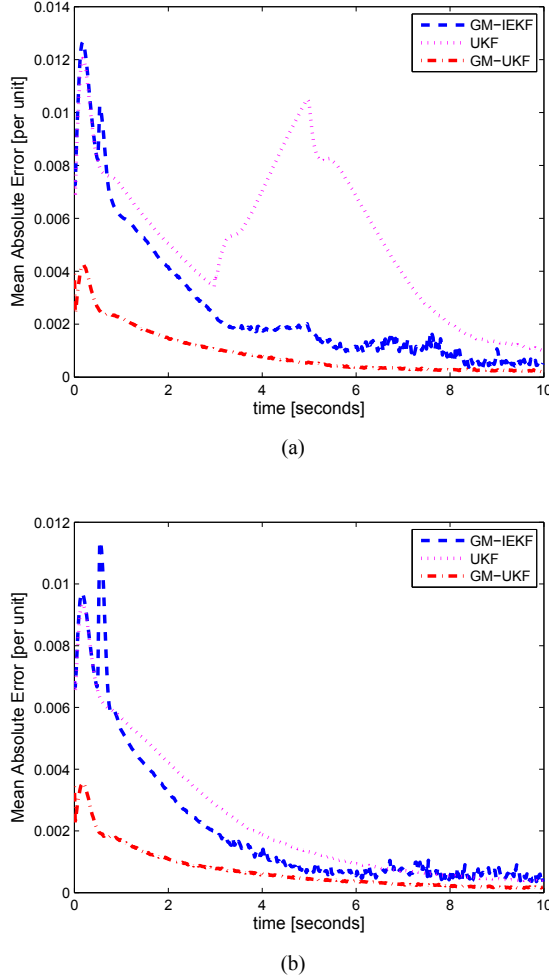


Fig. 4. Scenario 4: Tracking performance of the GM-IEKF, the UKF, and the GM-UKF in the presence of bad PMU measurements from $t=3s$ to $t=5s$, where the terminal real and reactive power measurements of Generator 5 are corrupted by 10% and 30% errors, respectively. (a) 10% errors of P_5 and Q_5 ; (b) 30% errors of P_5 and Q_5 .

C. Scenario 3: Occurrence of Large Non-Gaussian Noise

In Scenario 2, we have tested the case that both system process and measurement noises deviate slightly from the Gaussian assumption. Due to the changing operating status of the communication channels, the GPS synchronization process and the actual power systems, the deviation from Gaussian assumption can be large. Thus, it is important to investigate the robustness of the proposed GM-UKF to these scenarios. In this section, a bimodal Gaussian mixture with zero means, covariance matrices of $10^{-6}\mathbf{I}$ and $10^{-4}\mathbf{I}$ and weights of 0.95 and 0.05, are assumed for both the system process and measurement noises with appropriate dimensions. It should be noted that different combination of covariance matrices for the Gaussian mixture model will yield completely different non-Gaussian distributions. The mean absolute errors of each filter are plotted in Fig. 3-(a). From the results obtained in Scenario 2, it is found that all three filter have increased estimation errors. The UKF has the largest increased estimation errors among three filters followed by the GM-IEKF. Our GM-UKF has a slight increase of the estimation error, demonstrating

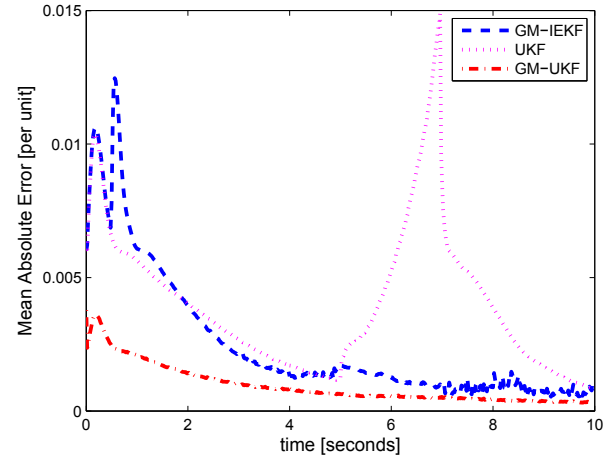


Fig. 5. Tracking performance of the GM-IEKF, the UKF, and the GM-UKF in the presence of innovation outliers from $t=5s$ to $t=7s$, where the transient reactance of the Generator 5 is corrupted with 30% errors.

its good robustness to unknown non-Gaussian noise. When the deviation from Gaussian assumption is further increased, that is, 10% of the data is contaminated by another Gaussian component with zero mean and covariance matrix $10^{-4}\mathbf{I}$ (this is simulated by changing the weights 0.95 and 0.05 to 0.9 and 0.1 for the Gaussian mixture model, respectively), the estimation results are displayed in Fig. 3-(b). A similar conclusion can be drawn from the results displayed in Fig. 3-(a).

D. Scenario 4: Occurrence of Outliers

Due to impulsive noise, loss of communication channels and saturations of the Potential Transformer (PT) and the Current Transformer (CT), bad PMU measurements can occur, yielding significant deviation from the Gaussian assumption of the measurement noise. To demonstrate the robustness of each filter, the following two cases are considered and tested:

Case 1: it is assumed that the received terminal real and reactive power measurements of Generator 5 are corrupted with 10% errors from $t=3s$ to $t=5s$.

Case 2: the received terminal real and reactive power measurements of Generator 5 are corrupted with 30% errors from $t=3s$ to $t=5s$.

The UKF-based normalized innovation statistical test [7], [8] is applied to detect observation outliers. Note that the detection threshold is system dependent and is challenging to set in presence of unknown non-Gaussian noise distributions. Following [7], [8], the detection threshold is set to be 8. Fig. 4 displays the mean absolute error of each of the three filter for these two scenarios. It can be found that the innovation statistical test is able to detect measurements with large gross errors while it fails to detect them when the gross error magnitudes are relatively smaller. This does not come as a surprise because the statistics of the system process and observation noise are unknown, yielding non-optimal detection threshold. Interestingly, while the GM-IEKF is able to withstand bad PMU measurements and provide reasonably good results, it produces large biases on the state estimates at the time when

TABLE I
AVERAGE COMPUTING TIMES OF THE THREE FILTERS AT EACH PMU
SAMPLE.

Scenarios	UKF	GM-IEKF	GM-UKF
Scenario 1	6.20ms	9.55ms	9.48ms
Scenario 2	6.23ms	9.58ms	9.50ms
Scenario 3	6.25ms	9.60ms	9.52ms
Scenario 4	6.36ms	9.66ms	9.62ms

the gross measurement errors occur. By contrast, our GM-UKF suppresses more effectively the bad PMU measurements, yielding much smaller biases on the state estimates.

In addition to observation outliers, the system model can be corrupted by incorrect parameters due to generator saturations, controller failures, among others, yielding innovation outliers. In this paper, we assume that the transient reactance of Generator 5 is corrupted by 30% errors from $t=5s$ to $t=7s$ due to saturations, yielding incorrect predicted state variables. The test results are displayed in Fig. 5. It is observed that the normalized innovation-based statistical test fails to detect the innovation outliers. This is expected because it assumes that the system model is reliable; in fact, it is unable to decide whether the predicted state variables or the observations are wrong. By contrast, our proposed GM-UKF will rely on the projection statistics to suppress the innovation outliers while exhibiting higher statistical efficiency than GM-IEKF.

E. Assessment of Computational Efficiency

The computing times of the UKF, the GM-IEKF and the GM-UKF in Scenarios 1-4 at each PMU sample are displayed in Table I. All the tests are performed on a PC with Intel Core i5, 2.50 GHz, 8GB of RAM. Note that all the approaches are implemented in a distributed manner for each generator. It can be concluded from this table that all three filter have comparative computational efficiency. The UKF has the fastest computing speed, followed by the GM-UKF. It is worth pointing out that although GM-UKF requires a slightly more execution time than the UKF, its computing time at each PMU sample is much lower than the actual PMU scan rate (20 ms). This demonstrates that our GM-UKF is suitable to track the dynamics of the system states in real-time. The major differences between GM-UKF and UKF that induce different computing times are the detection and processing outliers using projection statistics and the noise filtering using the iteratively reweighted least squares algorithm based GM-estimator.

V. CONCLUSION

In this paper, a robust GM-UKF is proposed to handle unknown statistics of system process and measurement noises as well as bad PMU measurements. A batch-mode regression form is first derived using the statistical linearization approach and by processing the predicted state vector and measurements simultaneously. This regression form enhances the data redundancy and allows us to detect bad PMU measurements and filter out unknown Gaussian and non-Gaussian noises through the GM-estimator. The error covariance matrix of the GM-UKF state estimates is derived from the total influence function of the GM-estimator. Numerical results carried out

on the IEEE 39-bus test system demonstrate the effectiveness and robustness of the proposed method. In the future work, we will extend our approach to develop a generalized GM-UKF for simultaneously estimating the system states and generator model parameters whose values are either inaccurate or incorrect.

APPENDIX A

TWO-AXIS GENERATOR MODEL

The differential and algebraic equations of the 9th-order two-axis generator model with IEEE-DC1A exciter and TGOV1 turbine-governor are represented as follows:

Differential equations:

$$T'_{do} \frac{dE'_q}{dt} = -E'_q - (X_d - X'_d) I_d + E_{fd}, \quad (28)$$

$$T'_{qo} \frac{dE'_d}{dt} = -E'_d - (X_q - X'_q) I_q, \quad (29)$$

$$\frac{d\delta}{dt} = \omega - \omega_s, \quad (30)$$

$$\frac{2H}{\omega_s} \frac{d\omega}{dt} = T_M - P_e - D(\omega - \omega_s), \quad (31)$$

$$T_E \frac{dE_{fd}}{dt} = -(K_E + S_E(E_{fd})) E_{fd} + V_R, \quad (32)$$

$$T_F \frac{dV_F}{dt} = -V_F + \frac{K_F}{T_E} V_R - \frac{K_F}{T_E} (K_E + S_E(E_{fd})) E_{fd}, \quad (33)$$

$$T_A \frac{dV_R}{dt} = -V_R + K_A (V_{ref} - V_F - V), \quad (34)$$

$$T_{CH} \frac{dT_M}{dt} = -T_M + P_{SV}, \quad (35)$$

$$T_{SV} \frac{dP_{SV}}{dt} = -P_{SV} + P_C - \frac{1}{R_D} \left(\frac{\omega}{\omega_s} - 1 \right), \quad (36)$$

Algebraic equations:

$$V_d = V \sin(\delta - \theta), V_q = V \cos(\delta - \theta), \quad (37)$$

$$I_d = \frac{E'_q - V_q}{X'_d}, I_q = \frac{V_d - E'_d}{X'_q}, \quad (38)$$

$$P_e = V_d I_d + V_q I_q, Q_e = -V_d I_q + V_q I_d, \quad (39)$$

where T'_{do} , T'_{qo} , T_E , T_F , T_A , T_{CH} and T_{SV} are time constants, in seconds; K_E , K_F and K_A are controller gains; V_{ref} and P_C are known control inputs; E'_q , E'_d , E_{fd} , V_F , V_R , T_M and P_{SV} are the q-axis and d-axis transient voltages, field voltage, scaled output of the stabilizing transformer and scaled output of the amplifier, synchronous machine mechanical torque and steam valve position, respectively; X_d , X'_d , X_q and X'_q are generator parameters; V and θ are the terminal bus voltage magnitude and phase angle, respectively; P_e and Q_e are the active and reactive electrical power outputs; I_d and I_q are the d and q axis currents, respectively.

By applying time discretization to (28)-(39), we get (1)-(2), yielding the state vector given by $\mathbf{x}_k = [\delta \ \omega \ E'_d \ E'_q \ E_{fd} \ V_F \ V_R \ T_M \ P_{SV}]$. The relationships given by (28)-(36) and (37)-(39) are represented in compact forms by the vector-valued function $\mathbf{f}(\cdot)$, which relates $\mathbf{x}_k$ to $\mathbf{x}_{k-1}$,

and by $\mathbf{h}(\cdot)$, respectively. The system input vector is denoted by $\mathbf{u}_k = [V_{ref} \ P_C]^T$. The measurement vector $\mathbf{z}_k$ contains a collection of voltage and current phasor measurements. For the case of the decentralized DSE, only metered values provided by the generator terminal PMUs are used; these PMUs measure the generator voltage and current phasors and the associated real and reactive power injections.

APPENDIX B PROJECTION STATISTICS ALGORITHM

The main steps of implementing the projection statistics algorithm are shown as follows:

- *Step 1:* For a point $\mathbf{l}_i$ in an n -dimensional space, calculate the coordinate-wise median given by

$$\mathbf{M} = \left\{ \text{med}_{j=1,\dots,m} (l_{j1}), \dots, \text{med}_{j=1,\dots,m} (l_{jn}) \right\}, \quad (40)$$

where m is the number of points;

- *Step 2:* Calculate the directions for projections $\mathbf{u}_j = \mathbf{l}_j - \mathbf{M}$, $j = 1, \dots, m$;
- *Step 3:* Normalize $\mathbf{u}_j$ to get

$$\boldsymbol{\ell}_j = \frac{\mathbf{u}_j}{\|\mathbf{u}_j\|} = \frac{\mathbf{u}_j}{\sqrt{u_{j1}^2 + \dots + u_{jn}^2}}; j = 1, \dots, m; \quad (41)$$

- *Step 4:* Calculate the standardized projections of the vectors $\{\mathbf{l}_1, \dots, \mathbf{l}_m\}$ on $\boldsymbol{\ell}_j$, which are given by

$$\zeta_{1j} = \mathbf{l}_1^T \boldsymbol{\ell}_j; \zeta_{2j} = \mathbf{l}_2^T \boldsymbol{\ell}_j; \dots, \zeta_{mj} = \mathbf{l}_m^T \boldsymbol{\ell}_j; \quad (42)$$

- *Step 5:* Calculate the median of $\{\zeta_{1j}, \dots, \zeta_{mj}\} = \zeta_{med,j}$;
- *Step 6:* Calculate the median absolute deviation (MAD) $MAD_j = 1.4826 \cdot b \cdot \text{med}_i |\zeta_{ij} - \zeta_{med,j}|$, where the correction factor is $b = 1 + 15/(m - n)$;
- *Step 7:* Calculate the standardized projections

$$P_{ij} = \frac{|\zeta_{ij} - \zeta_{med,j}|}{MAD_j} \text{ for } i = 1, \dots, m; \quad (43)$$

- *Step 8:* Repeat steps 4–7 for all vectors $\{\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_m\}$ to get the standardized projections $\{P_{i1}, \dots, P_{im}\}$ for $i = 1, \dots, m$;
- *Step 9:* Calculate the projection statistics

$$PS_i = \max \{P_{i1}, \dots, P_{im}\} \text{ for } i = 1, \dots, m. \quad (44)$$

REFERENCES

- [1] 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, 2002.
- [2] A. P. Meliopoulos, *et al.*, "Dynamic state estimation based protection: status and promise," *IEEE Trans. Power Delivery*, vol. 32, no. 1, pp. 320–330, 2017.
- [3] S. Akhlaghi, N. Zhou, Z. Huang, "A multi-step adaptive interpolation approach to mitigating the impact of nonlinearity on dynamic state estimation," *IEEE Trans. Smart Grid*, 2016.
- [4] E. Ghahremani, I. Kamwa, "Dynamic state estimation in power system by applying the extended Kalman filter with unknown inputs to phasor measurements," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2556–2566, 2011.
- [5] J. B. Zhao, M. Netto, L. Mili, "A robust iterated extended Kalman filter for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3205–3216, 2017.
- [6] J. B. Zhao, L. Mili, A. Abdelhadi, "Robust dynamic state estimator to outliers and cyber attacks," in *Proc. of IEEE PES General Meeting*, July 16–20, Chicago 2017.
- [7] 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, 2012.
- [8] 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, 2014.
- [9] A. Rouhani, A. Abur, "Linear phasor estimator assisted dynamic state estimation," *IEEE Trans. Smart Grid*, 2016.
- [10] G. Anagnostou, B. C. Pal, "Derivative-free Kalman filterin based approaches to dynamic state estimation for power systems with unknown inputs," *IEEE Trans. Power Syst.*, 2017.
- [11] J. B. Zhao, "Power system dynamic state estimation considering measurement correlations," *IEEE Trans. Energy Conversion*, DOI: 10.1109/TEC.2017.2742405, 2017.
- [12] 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.
- [13] N. Zhou, D. Meng, S. Lu, "Estimation of the dynamic states of synchronous machines using an extended particle filter," *IEEE Trans. Power Syst.*, vol. 28, no. 4, 4152–4161, 2013.
- [14] D. Simon, *Optimal State Estimation*. Hoboken, NJ: Wiley, 2006.
- [15] K. Zhou, J. C. Doyle, K. Glover, *Robust and Optimal Control*, Prentice-Hall, New Jersey, 1996.
- [16] J. B. Zhao, "Dynamic state estimation with model uncertainties using H-infinity extended Kalman filter," *IEEE Trans. Power Syst.*, DOI: 10.1109/TPWRS.2017.2688131, 2017.
- [17] Z. Huang, N. Zhou, R. Diao, S. Wang, S. Elbert, D. Meng, S. Lu, "Capturing real-time power system dynamics: opportunities and challenges," *Proc. IEEE Power Eng. Soc. General Meeting*, pp. 1–5, 2015.
- [18] S. Wang, J. B. Zhao, Z. Huang, R. Diao, "Assessing Gaussian assumption of PMU measurement error using field data," *IEEE Trans. Power Delivery*, 2017, in press.
- [19] P. W. Sauer, M. A. Pai, *Power System Dynamics and Stability*, Prentice-Hall, 1998.
- [20] I. Arasaratnam, S. Haykin, R. J. Elliott, "Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 953–977, 2007.
- [21] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.
- [22] S. Julier, J. Uhlmann, H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filter and estimators," *IEEE Trans. Autom. Control*, vol. 45, no. 3, pp. 477–482, 2000.
- [23] T. Lefebvre, H. Bruyninckx, J. De Schutter, "Comment on "a new method for the nonlinear transformation of means and covariances in filter and estimators,"" *IEEE Trans. Autom. Control*, vol. 47, no. 8, pp. 1406–1408, 2002.
- [24] M. Gandhi, L. Mili, "Robust Kalman filter based on a generalized maximum-likelihood-type estimator," *IEEE Trans. Signal Processing*, vol. 58, no. 5, pp. 2509–2520, 2010.
- [25] L. Mili, M. Cheniae, N. Vichare, P. Rousseeuw, "Robust state estimation based on projection statistics," *IEEE Trans. Power Syst.*, vol. 11, no. 2, p. 1118–1127, 1996.
- [26] L. Thomas, L. Mili, "A robust GM-estimator for the automated detection of external defects on barked hardwood logs and stems," *IEEE Trans. Signal Processing*, vol. 55, no. 7, pp. 3568–3576, 2007.
- [27] IEEE PES TF on Benchmark System for Stability Controls, "Benchmark systems for small-signal stability analysis and control," Aug. 2015.