Unscented Kalman Filter with Generalized Correntropy Loss for Robust Power System Forecasting-Aided State Estimation

Wentao Ma, *Member, IEEE,* Jinzhe Qiu, Xinghua Liu, *Member, IEEE,* Gaoxi Xiao, *Senior Member, IEEE,* Jiandong Duan, and Badong Chen, *Senior Member, IEEE*

***Abstract*—Due to the existence of various anomalies such as non-Gaussian process and measurement noises, gross measurement errors and sudden changes of system status, the robust forecasting-aided state estimation is pivotal for power system stability. This paper develops a novel unscented Kalman filter (UKF) with generalized Correntropy loss (GCL) (termed as GCL-UKF) to estimate power system state with forecasting-aid. The GCL is used to replace the mean square error loss in original UKF framework. Such an approach, as we shall show the strength of the GCL developed in robust information theoretic learning for addressing the non-Gaussian interference while benefiting from the strength of the UKF in handling strong model nonlinearities. In addition, we take into account the nontrivial influences of the bad data for the innovation vector. An enhanced GCL-UKF (EnGCL-UKF) method is established by introducing an exponential function of the innovation vector to adjust covariance matrix so as to improve the GCL-UKF based state estimation accuracy under the change of gain matrix caused by bad factors. Numerical simulation results carried out on IEEE 14-bus, 30-bus and 57-bus test systems validate the efficacy of the proposed methods for state estimation under various types of measurement.**

***Index Terms*—Generalized Correntropy Loss, unscented Kalman filter, power system forecasting-aided state estimation, non-Gaussian measurement noise.**

I. Introduction

Reliable and accurate state estimation techniques are of great benefits to the safe and stable operation of power systems [1]-[3]. In view of the real-time changes of the system load and to maintain the system balance and stability, the generator continuously adjusts the speed and frequency aiming to generate desired changes of the injection power and the branch power flow of the node. The forecasting-aided state estimation (FASE), which can forecast the state of the system via making full use of the prior information of the state variables and trace the changes of the node state, is becoming increasingly important for power system real-time modeling and control in modern energy management center [4]-[6]. The Kalman filtering (KF) approach is one of the most significant tools for power system state estimation due to its outstanding tracking ability. The iterated extended KF (EKF) based on the generalized maximum likelihood approach was developed to estimate state when subjected to disturbances [7]. Considering the difficulty in obtaining the process and measurement noise models, an adaptive EKF with inflatable noise variances is proposed for real time state estimation [8]. The traditional EKF, however, may produce large estimation errors or even diverge when the system nonlinearity is strong. Consequently, the unscented KF (UKF) algorithm with means of unscented transformation (UT) has been widely used in engineering, such as state of charge estimation [9] and power system state estimation [10]-[13]. In existing studies, the UKF was used to estimate the current system bus voltage magnitudes and phase angles under typical measurement condition [10]. An effective state estimation method was developed by using UT to calculate the mean and covariance of the nonlinear functions of state transition and observation models [11]. A constrained iterated UKF was proposed to estimate the state variables and unknown parameters [12]. In short, UKF with enhanced numerical stability was developed for state estimation of the multi-machine power system [13]. The UKF has been a promising approach for the state estimation due to its simple calculation process and the superior performance in highly nonlinear systems.

The performance of an estimator heavily depends on the accuracy of the measurements and the assumed estimation model [14]-[15]. A power system may suffer from unexpected noises in system sensing which may not necessarily follow Gaussian distribution [16]-[17]. However, the UKF with mean square error (MSE) loss can achieve accurate state estimation results only when the process and observation noises obey a Gaussian distribution; its good performance no longer holds when the Gaussianity assumption of the noises is violated, especially for impulsive noise. The state estimate vector obtained from the UKF hence may be significantly degraded due to the lack of statistical robustness of the filter to non-Gaussian noise[18]. To mitigate the impacts of non-Gaussian noise, Zhao proposed H-EKF to restrain the influence of system uncertainties [17]. A novel robust generalized maximum likelihood (GM)-estimator [6] and GM-UKF [18] were developed to enable suppressing

This work was supported in part by the National Natural Science Foundation of China (No.U1613219, 51877174), the Key Project of Natural Science Basic Research Plan in Shaanxi Province of China (No.2019JZ-05), the Natural Science Basic Research Program of Shaanxi (No.2017JM6033), the Thousand Talents Plan of Shaanxi Province for Young Professionals, and partially supported by Ministry of Education, Singapore (MOE2016-T2-1-119).

W. Ma, J. Qiu, X. Liu and J. Duan are with the School of Automation and Information Engineering, Xi'an University of Technology, Xi'an,710048 China (e-mail: mawt@xaut.edu.cn, qiujinzhe@163.com, liuxh@xaut.edu.cn, duanjd@xaut.edu.cn).

G. Xiao is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (e-mail: egxxiao@ntu.edu.sg).

B. Chen is with the School of Electronic and Information Engineering, Xian Jiaotong University, Xian, China (e-mail: chenbd@mail.xjtu.edu.cn)

observation and innovation outliers for FASE and dynamic state estimation. In recent years, Correntropy as a novel similarity function has been proposed in information theoretic learning methodology [19], and Correntropy based KFs have been developed in [20]-[21] which can obtain higher-order moments of error and improve the performance of the original KFs under non-Gaussian noise. However, it cannot change the shape of Correntropy freely due to the limit of dealing with only one particular type of noise with Gaussian function. The Gaussian kernel is, therefore, not always the best choice [22]-[23]. Consequently, the generalized Correntropy was defined by Chen [22], and it utilizes the generalized Gaussian density function instead of the traditional Gaussian function as a kernel function in Correntropy methodology. Thanks to its better more flexibility, the generalized Correntropy loss (GCL) has been used as the cost function to design robust adaptive filter through minimizing the GCL for different applications [24]-[25].

To enrich the methodology of robust power system state estimation and enhance the estimation accuracy, this paper mainly focuses on developing novel robust state estimation approaches along with the following contributions.

(i) A novel robust UKF with GCL (denoted as GCL-UKF) is established in this work. The GCL with a suitable kernel width achieves a robust performance under non-Gaussian noise interference.Thus it can improve the robustness of the UKF, which in turn ensures satisfactory performance of FASE under the non-Gaussian environments.

(ii) Considering the serious influence of the bad data for the innovation vector, an exponential function of the innovation vector is introduced into the GCL-UKF to adjust covariance matrix. Thus, the influence of bad data on gain matrix is suppressed, and further the robustness of the GCL-UKF method is enhanced.

(iii) The GCL-UKF and its enhancement version are used for power system forecasting-aided state estimation under non-Gaussian noise cases. Experiments and comparison analysis under various conditions validate the efficacy of the proposed methods in complex environment.

Rest of the paper is organized as follows. Section II describes the system model and the generalized Correntropy loss. Section III develops the UKF with GCL method and its enhanced version for robust power system estimation. Section IV shows and analyzes the simulation results under various conditions. Finally Section V concludes with a summary of the main findings of this paper.

II. SYSTEM MODEL AND GENERALIZED CORRENTROPY LOSS

A. Power System Dynamic Model

For FASE, it should establish a physical and mathematical model of the time-varying characteristics of the system between the state variables and the observation time first. Here, a power system model is given by a nonlinear discrete-time states and measurements equations as:

$$\mathbf{x}_i = \mathbf{f}(\mathbf{x}_{i-1}) + \mathbf{w}_i \tag{1}$$

$$\mathbf{y}_i = \mathbf{h}(\mathbf{x}_i) + \mathbf{v}_i \tag{2}$$

where $\mathbf{x}_i \in \mathbb{R}^n$ denotes the state vector consisting of magnitudes and angles of nodal voltage. The measurement vector $\mathbf{y}_i \in \mathbb{R}^m$ comprises of voltage magnitude measurements, real power injection measurements, reactive power injection measurements, real power flow measurements, and reactive power flow measurements. $\mathbf{w}_i$ and $\mathbf{v}_i$ are system process and measurement noises with covariance matrices $\mathbf{Q}_i \in \mathbb{R}^{n\times n}$ and $\mathbf{R}_i \in \mathbb{R}^{m\times m}$, respectively. The noises $\mathbf{w}_i$ and $\mathbf{v}_i$ are usually assumed to be Gaussian noise and independent of each other. However, the system may suffer from the interference of the noise with non-Gaussian distribution in practice.

We mainly focus on the development of novel robust SE method for power system. $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ represent the state-transition function and measurement function, where $\mathbf{f}(\cdot)$ can be obtained by the Holt's two-parameter linear exponential smoothing technique [26] as

$$\mathbf{x}_{i-1} = \mathbf{a}_{i-1} + \mathbf{b}_{i-1} \tag{3}$$

$$\mathbf{a}_{i-1} = \alpha_{i-1}\mathbf{x}_{i-1} + (1-\alpha_{i-1})\mathbf{x}_{i-1}^{'} \tag{4}$$

$$\mathbf{b}_{i-1} = \beta_{i-1}(\mathbf{a}_{i-1} - \mathbf{a}_{i-2}) + (1-\beta_{i-1})\mathbf{b}_{i-2} \tag{5}$$

where α_{i-1} and β_{i-1} are parameters at $(i-1)^{th}$ instant lying in the range from 0 to 1, and $\mathbf{x}_{i-1}^{'}$ denotes the predicted state vector at $(i-1)^{th}$ instant. To execute the state prediction step (time update) of the KF approach, the state forecasting function (3) is used in (1) to predict the state vector in advance when the state prediction step (time update) of the KF based approach is executed. To define the measurement function $\mathbf{h}(\cdot)$ for power system, the standard real and reactive power balance and line flow equations are used, which are given by

$$\mathbf{P}_s = \sum_{j=1}^{N} |\mathbf{V}_s||\mathbf{V}_j|(\mathbf{G}_{sj}\cos\theta_{sj} + \mathbf{B}_{sj}\sin\theta_{sj}) \tag{6}$$

$$\mathbf{Q}_s = \sum_{j=1}^{N} |\mathbf{V}_s||\mathbf{V}_j|(\mathbf{G}_{sj}\sin\theta_{sj} - \mathbf{B}_{sj}\cos\theta_{sj}) \tag{7}$$

$$\mathbf{P}_{sj} = \mathbf{V}_s^2(\mathbf{G}_{gs}+\mathbf{G}_{sj}) - |\mathbf{V}_s||\mathbf{V}_j|(\mathbf{G}_{sj}\cos\theta_{sj} + \mathbf{B}_{sj}\sin\theta_{sj}) \tag{8}$$

$$\mathbf{Q}_{sj} = -\mathbf{V}_s^2(\mathbf{B}_{gs}+\mathbf{B}_{sj}) - |\mathbf{V}_s||\mathbf{V}_j|(\mathbf{G}_{sj}\sin\theta_{sj} - \mathbf{B}_{sj}\cos\theta_{sj}) \tag{9}$$

where $|\mathbf{V}_s|$ is the voltage magnitude at bus s, θ_{sj} is the voltage angle between buses s and j, $\mathbf{P}_s$ and $\mathbf{Q}_s$ are the real power injection and reactive power injection at bus s, $\mathbf{P}_{sj}$ and $\mathbf{Q}_{sj}$ are the real power flow and reactive power flow between buses s and j, $\mathbf{G}_{sj}$ and $\mathbf{B}_{sj}$ are the conductance and susceptance of the line between buses s and j, $\mathbf{G}_{gs}$ and $\mathbf{B}_{gs}$ are the conductance and susceptance of the shunt at bus s.

B. Generalized Correntropy Loss

Considering two random variables $\mathbf{X}$ and $\mathbf{Y}$, the Correntropy is defined by

$$\mathbf{V}(\mathbf{X},\mathbf{Y}) = \mathbb{E}[\kappa_\sigma(\mathbf{X},\mathbf{Y})] = \int \kappa_\sigma(\mathbf{x},\mathbf{y})d\mathbf{F}_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) \tag{10}$$

where $\mathbb{E}$ denotes the expectation operator, $\kappa_\sigma(\cdot,\cdot)$ stands for a kernel function with kernel width $\sigma > 0$, and $\mathbf{F}_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})$ denotes the joint distribution function of $(\mathbf{X},\mathbf{Y})$.

The well-known generalized Gaussian density function with zero-mean is defined by

$$\begin{aligned}\mathbf{G}_{\alpha,\beta}(\mathbf{X},\mathbf{Y}) &= \frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})}\exp\{-(\frac{|\mathbf{X}-\mathbf{Y}|}{\beta})^\alpha\} \\ &= \tau_{\alpha,\beta}\exp\{-\gamma|\mathbf{X}-\mathbf{Y}|^\alpha\}\end{aligned} \tag{11}$$

where $\Gamma(\cdot)$ is the gamma function, $\alpha > 0$ represents the shape parameter, $\beta > 0$ is the scale parameter, $\gamma = \frac{1}{\beta^\alpha}$ is the kernel parameter, and $\tau_{\alpha,\beta} = \frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})}$ is the normalization constant. Then, the generalized Correntropy (GC) can be defined as [22]

$$\mathbf{V}_{\alpha,\beta}(\mathbf{X},\mathbf{Y}) = \mathbb{E}\big[\mathbf{G}_{\alpha,\beta}(\mathbf{X},\mathbf{Y})\big] \tag{12}$$

In practice, the data distribution is usually unknown and only a finite number of samples are available. Then the sample mean estimator of the GC is

$$\hat{\mathbf{V}}_{\alpha,\beta}(\mathbf{X},\mathbf{Y}) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{G}_{\alpha,\beta}(\mathbf{x}_i,\mathbf{y}_i) \tag{13}$$

Further, an estimator of the GC-loss (GCL) can be defined by

$$\begin{aligned}\hat{\mathbf{J}}_{GCL}(\mathbf{X},\mathbf{Y}) &= \mathbf{G}_{\alpha,\beta}(0,0) - \hat{\mathbf{V}}_{\alpha,\beta}(\mathbf{X},\mathbf{Y}) \\ &= \tau_{\alpha,\beta} - \frac{1}{N}\sum_{i=1}^{N}\mathbf{G}_{\alpha,\beta}(\mathbf{x}_i,\mathbf{y}_i)\end{aligned} \tag{14}$$

Similar to the Correntropy, the generalized Correntropy involves more higher-order absolute moments of data than those of MSE. The GCL function behaves like different norms (from l_α to l_0) of the data in different regions, and it is robust to outliers with a suitable kernel parameter. Generally, the non-Gaussian noise may occur when the power system is confronted with the loss of communications links, inaccurate phasor synchronization in measurements, or large biases caused by the saturation of metering current transformers [18].Therefore, a novel robust estimator based on GCL will be proposed in this work to mitigate the impacts of non-Gaussian noise with outliers.

III. GENERALIZED CORRENTROPY LOSS BASED UKF FOR ROBUST STATE ESTIMATION

Here, through the use of the GCL in UKF, the derivation process of the UKF with GCL is first presented based on the state-space equation of a nonlinear system represented in (1) and (2) which follows the structure of standard UKF.

A. UKF with generalized Correntropy loss (GCL-UKF)

1) Time update: According to the UT transform, a set of $2\mathbf{n}+1$ sigma points should be generated by (15) from the estimated sate $\mathbf{x}_{i-1|i-1}$ and estimate error covariance matrix $\mathbf{P}_{i-1|i-1}$ at time $i-1$.

$$\chi^s_{i-1|i-1} = \begin{cases}\hat{\mathbf{x}}_{i-1|i-1} & s=0 \\ \hat{\mathbf{x}}_{i-1|i-1} + \left(\sqrt{(\mathbf{n}+\lambda)\mathbf{P}_{i-1|i-1}}\right)_s & s=1,\cdots,\mathbf{n} \\ \hat{\mathbf{x}}_{i-1|i-1} + \left(\sqrt{(\mathbf{n}+\lambda)\mathbf{P}_{i-1|i-1}}\right)_{s-\mathbf{n}} & s=\mathbf{n}+1,\cdots,2\mathbf{n}\end{cases} \tag{15}$$

where $\left(\sqrt{(\mathbf{n}+\lambda)\mathbf{P}_{i-1|i-1}}\right)_s$ represents the sth column of $\sqrt{(\mathbf{n}+\lambda)\mathbf{P}_{i-1|i-1}}$, $\lambda = \delta^2(\mathbf{n}+\kappa) - n$ is a composite scaling factor, where the parameter $0 \le \delta \le 1$ determines the spread of the sigma points around, κ is set to $3-\mathbf{n}$ when the state variable is multivariable, which can be used to reduce the higher order errors of the mean and the covariance approximations.

Setting an initial state variable $\mathbf{x}_0$, then the initial state mean $\hat{\mathbf{x}}_{0|0} = \mathbb{E}[\mathbf{x}_0]$, and the initial state estimate error covariance matrix $\mathbf{P}_{0|0}$ can be represented as

$$\mathbf{P}_{0|0} = \mathbb{E}\big[(\mathbf{x}_0-\hat{\mathbf{x}}_0)(\mathbf{x}_0-\hat{\mathbf{x}}_0)^T\big] \tag{16}$$

Then, the prior state mean $\hat{\mathbf{x}}_{i-1|i-1}$ and covariance of the one step predicted state $\mathbf{P}_{i|i-1}$ are obtained by

$$\hat{\mathbf{x}}_{i|i-1} = \sum_{s=0}^{2\mathbf{n}} w^i_\eta \mathbf{f}(\chi^s_{i-1|i-1}) \tag{17}$$

$$\mathbf{P}_{i|i-1} = \sum_{s=0}^{2\mathbf{n}} w^i_c[\mathbf{f}(\chi^s_{i-1|i-1}) - \hat{\mathbf{x}}_{i|i-1}][\mathbf{f}(\chi^s_{i-1|i-1}) - \hat{\mathbf{x}}_{i|i-1}]^T + \mathbf{Q}_{i-1} \tag{18}$$

w^i_η, w^i_c are the weights of the sigma point mean, $w^0_\eta = \frac{\lambda}{\mathbf{n}+\lambda}$, $w^0_c = \frac{\lambda}{\mathbf{n}+\lambda} + 1 - \delta^2 + \zeta^2$ and

$$w^i_\eta = w^i_c = \frac{1}{2(\mathbf{n}+\lambda)},\ i=1,2,\cdots,2\mathbf{n} \tag{19}$$

where the parameter $\zeta \ge 0$ is related to the distribution of the state variable, and it is optimal when $\zeta = 2$ for Gaussian distribution.

2) Measurement update: During measurement update procedure, $2\mathbf{n}+1$ sigma points from $\hat{\mathbf{x}}_{i|i-1}$ and $\mathbf{P}_{i|i-1}$ should be first computed by

$$\chi^s_{i|i-1} = \begin{cases}\hat{\mathbf{x}}_{i|i-1} & s=0 \\ \hat{\mathbf{x}}_{i|i-1} + \left(\sqrt{(\mathbf{n}+\lambda)\mathbf{P}_{i|i-1}}\right)_s & s=1,\cdots,\mathbf{n} \\ \hat{\mathbf{x}}_{i|i-1} + \left(\sqrt{(\mathbf{n}+\lambda)\mathbf{P}_{i|i-1}}\right)_{s-\mathbf{n}} & s=\mathbf{n}+1,\cdots,2\mathbf{n}\end{cases} \tag{20}$$

Then the prior mean $\hat{\mathbf{y}}_{i|i-1}$ and the predicted measurement cross-covariance matrix $\mathbf{P}_{\mathbf{xy},i}$ can be computed by

$$\hat{\mathbf{y}}_{i|i-1} = \sum_{s=0}^{2\mathbf{n}} w^i_\eta \mathbf{h}(\chi^s_{i|i-1}) \tag{21}$$

$$\mathbf{P}_{\mathbf{xy},i} = \sum_{s=0}^{2\mathbf{n}} w^i_c[\chi^s_{i|i-1} - \hat{\mathbf{x}}_{i|i-1}][\mathbf{h}(\chi^s_{i|i-1}) - \hat{\mathbf{y}}_{i|i-1}]^T \tag{22}$$

To further finish the measurement update, a statistical linear regression model based on the GCL is employed in this procedure which is similar to the approach proposed in [27]. Specifically, a prior state estimation error is defined as

$$\eta(\mathbf{x}_i) = \mathbf{x}_i - \hat{\mathbf{x}}_{i|i-1} \tag{23}$$

Further, one can define a measurement slope matrix as

$$\bar{\mathbf{H}}_i = (\mathbf{P}^{-1}_{i|i-1}\mathbf{P}_{\mathbf{xy},i})^T \tag{24}$$

Then, the measurement equation (2) can be approximated by

$$\mathbf{y}_i \approx \hat{\mathbf{y}}_{i|i-1} + \bar{\mathbf{H}}_i(\mathbf{x}_i - \hat{\mathbf{x}}_{i|i-1}) + \mathbf{v}_i \tag{25}$$

Combining equations (1), (23) and (25) yields

$$\begin{bmatrix} \hat{\mathbf{x}}_{i|i-1} \\ \mathbf{y}_i - \hat{\mathbf{y}}_{i|i-1} + \bar{\mathbf{H}}_i\hat{\mathbf{x}}_{i|i-1} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \bar{\mathbf{H}}_i \end{bmatrix} \mathbf{x}_i + \boldsymbol{\xi}_i \tag{26}$$

where $\boldsymbol{\xi}_i = \begin{bmatrix} -\eta(\mathbf{x}_i) \\ \mathbf{v}_i \end{bmatrix}$ with

$$\mathbb{E}[\boldsymbol{\xi}_i\boldsymbol{\xi}_i^T] = \begin{bmatrix} \mathbf{S}_{p,i|i-1}\mathbf{S}_{p,i|i-1}^T & 0 \\ 0 & \mathbf{S}_{r,i}\mathbf{S}_{r,i}^T \end{bmatrix} = \mathbf{S}_i\mathbf{S}_i^T \tag{27}$$

where $\mathbf{S}_i$ represents the Cholesky decomposition factor of the matrix $\mathbb{E}[\boldsymbol{\xi}_i\boldsymbol{\xi}_i^T]$. Multiplying both sides of (26) by $\mathbf{S}_i^{-1}$, we have

$$\mathbf{D}_i = \mathbf{W}_i\mathbf{x}_i + \mathbf{e}_i \tag{28}$$

where

$$\mathbf{D}_i = \mathbf{S}_i^{-1} \begin{bmatrix} \hat{\mathbf{x}}_{i|i-1} \\ \mathbf{y}_i - \hat{\mathbf{y}}_{i|i-1} + \bar{\mathbf{H}}_i\hat{\mathbf{x}}_{i|i-1} \end{bmatrix} \tag{29}$$

$$\mathbf{W}_i = \mathbf{S}_i^{-1} \begin{bmatrix} \mathbf{I} \\ \bar{\mathbf{H}}_i \end{bmatrix} \tag{30}$$

$$\mathbf{e}_i = \mathbf{S}_i^{-1}\boldsymbol{\xi}_i \tag{31}$$

with $\mathbf{D}_i = [\mathbf{d}_{1,i}, \cdots, \mathbf{d}_{L,i}]^T$, $\mathbf{W}_i = [\mathbf{w}_{1,i}, \cdots, \mathbf{w}_{L,i}]^T$, $\mathbf{e}_i = [\mathbf{e}_{1,i}, \cdots, \mathbf{e}_{L,i}]^T$ and $\mathbf{L} = \mathbf{m}+\mathbf{n}$. According to (27) and (31), we know that $\mathbb{E}[\mathbf{e}_i\mathbf{e}_i^T] = \mathbf{I}$, where $\mathbf{I}$ is unity matrix.

For getting the optimal values of the state variables, GCL is employed as

$$\hat{\mathbf{J}}_{GCL}(\mathbf{x}_i) = \tau_{\alpha,\beta}\big[1 - \frac{1}{\mathbf{L}}\sum_{k=1}^{\mathbf{L}} \exp\{-\gamma|\mathbf{e}_{\mathbf{k},i}|^\alpha\}\big] \tag{32}$$

where $\mathbf{e}_{\mathbf{k},i}$ is the kth element of $\mathbf{e}_i$ given by

$$\mathbf{e}_{\mathbf{k},i} = \mathbf{d}_{\mathbf{k},i} - \mathbf{w}_{\mathbf{k},i}\mathbf{x}_i \tag{33}$$

The optimal estimate of $\mathbf{x}_i$ then can be obtained by minimizing (32) as

$$\hat{\mathbf{x}}_i = \arg\min_{\mathbf{x}_i} \tau_{\alpha,\beta}\big[1 - \frac{1}{\mathbf{L}}\sum_{k=1}^{\mathbf{L}} \exp\{-\gamma|\mathbf{e}_{\mathbf{k},i}|^\alpha\}\big] \tag{34}$$

Setting the gradient of the (32) regarding $\mathbf{x}_i$ to be zero yields

$$\begin{aligned} &\frac{\partial\hat{\mathbf{J}}_{GCL}(\mathbf{x}_i)}{\partial\mathbf{x}_i} \\ &= -\frac{\tau_{\alpha,\beta}}{\mathbf{L}}\sum_{\mathbf{k}=1}^{\mathbf{L}} \exp\{-\gamma|\mathbf{e}_{\mathbf{k},i}|^\alpha\}|\mathbf{e}_{\mathbf{k},i}|^{\alpha-1}\mathrm{sign}(\mathbf{e}_{\mathbf{k},i})\mathbf{w}_{\mathbf{k},i} \\ &= -\frac{\tau_{\alpha,\beta}\gamma}{\mathbf{L}}\sum_{\mathbf{k}=1}^{\mathbf{L}} \exp\{-\gamma|\mathbf{e}_{\mathbf{k},i}|^\alpha\}|\mathbf{e}_{\mathbf{k},i}|^{\alpha-2}(\mathbf{d}_{\mathbf{k},i} - \mathbf{w}_{\mathbf{k},i}\mathbf{x}_i)\mathbf{w}_{\mathbf{k},i} \\ &= 0 \end{aligned} \tag{35}$$

From (35), one can easily get

$$\begin{aligned} \mathbf{x}_i = &\big(\sum_{\mathbf{k}=1}^{\mathbf{L}} \exp\{-\gamma|\mathbf{e}_{\mathbf{k},i}|^\alpha\}|\mathbf{e}_{\mathbf{k},i}|^{\alpha-2}\mathbf{w}_{\mathbf{k},i}^T\mathbf{w}_{\mathbf{k},i}\big)^{-1} \\ &\times \big(\sum_{\mathbf{k}=1}^{\mathbf{L}} \exp\{-\gamma|\mathbf{e}_{\mathbf{k},i}|^\alpha\}|\mathbf{e}_{\mathbf{k},i}|^{\alpha-2}\mathbf{w}_{\mathbf{k},i}^T\mathbf{d}_{\mathbf{k},i}\big)^{-1} \end{aligned} \tag{36}$$

Equation (36) is actually a fixed-point equation with respect to $\mathbf{x}_i$, and it can be expressed as

$$\mathbf{x}_i = \mathbf{g}(\mathbf{x}_i) \tag{37}$$

Now, we obtain a fixed-point iterative equation from (37) as

$$\hat{\mathbf{x}}_{i,t+1} = \mathbf{g}(\hat{\mathbf{x}}_{i,t}) \tag{38}$$

where $\hat{\mathbf{x}}_{i,t}$ is the estimated state $\widehat{\mathbf{x}}_i$ at the tth fixed-point iteration. Here, (36) is further expressed in a matrix form as

$$\mathbf{x}_i = \big(\mathbf{W}_i^T\mathbf{C}_i\mathbf{W}_i\big)^{-1}\big(\mathbf{W}_i^T\mathbf{C}_i\mathbf{D}_i\big) \tag{39}$$

where

$$\mathbf{C}_i = \begin{bmatrix} \mathbf{C}_{\mu_1,i} & 0 \\ 0 & \mathbf{C}_{\mu_2,i} \end{bmatrix} \tag{40}$$

$$\begin{aligned} \mathbf{C}_{\mu_1,i} &= \mathrm{diag}\{\mathbf{g}_{\mathbf{C}}(\mathbf{e}_{1,i}), \cdots, \mathbf{g}_{\mathbf{C}}(\mathbf{e}_{\mathbf{n},i})\} \\ \mathbf{C}_{\mu_2,i} &= \mathrm{diag}\{\mathbf{g}_{\mathbf{C}}(\mathbf{e}_{\mathbf{n}+1,i}), \cdots, \mathbf{g}_{\mathbf{C}}(\mathbf{e}_{\mathbf{n}+\mathbf{m},i})\} \end{aligned} \tag{41}$$

with $\mathbf{g}_{\mathbf{C}}(\mathbf{e}_{k,i}) = \exp\{-\gamma|\mathbf{e}_{\mathbf{k},i}|^\alpha\}|\mathbf{e}_{\mathbf{k},i}|^{\alpha-2}$.

In addition, using the results (27) and (30), we have

$$\mathbf{W}_i = \begin{bmatrix} \mathbf{S}_{p,i|i-1}^{-1} & 0 \\ 0 & \mathbf{S}_{r,i}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \bar{\mathbf{H}}_i \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{p,i|i-1}^{-1} \\ \mathbf{S}_{r,i}^{-1}\bar{\mathbf{H}}_i \end{bmatrix} \tag{42}$$

Combining (40) and (42), we get

$$\big(\mathbf{W}_i^T\mathbf{C}_i\mathbf{W}_i\big)^{-1} = \big((\mathbf{S}_p^{-1})^T\mathbf{C}_{\mu_1}\mathbf{S}_p^{-1} + \bar{\mathbf{H}}_i^T(\mathbf{S}_r^{-1})^T\mathbf{C}_{\mu_2}\mathbf{S}_r^{-1}\bar{\mathbf{H}}_i\big)^{-1} \tag{43}$$

where $\mathbf{S}_p$, $\mathbf{S}_r$, $\mathbf{C}_{\mu_1}$ and $\mathbf{C}_{\mu_2}$ respectively represent $\mathbf{S}_{p,i|i-1}$, $\mathbf{S}_{r,i}$, $\mathbf{C}_{\mu_1,i}$ and $\mathbf{C}_{\mu_2,i}$ for simplicity. The variables substitution thus can be conducted as

$$\begin{aligned} &(\mathbf{S}_p^{-1})^T\mathbf{C}_{\mu_1}\mathbf{S}_p^{-1} \triangleq \mathbf{A}, \quad \bar{\mathbf{H}}_i^T \triangleq \mathbf{B}, \\ &(\mathbf{S}_r^{-1})^T\mathbf{C}_{\mu_2}\mathbf{S}_r^{-1} \triangleq \mathbf{C}, \quad \bar{\mathbf{H}}_i \triangleq \mathbf{D} \end{aligned} \tag{44}$$

Using the matrix inversion lemma, we yield

$$\begin{aligned} (\mathbf{W}_i^T\mathbf{C}_i\mathbf{W}_i)^{-1} &= (\mathbf{A} + \mathbf{BDC})^{-1} = \mathbf{S}_p\mathbf{C}_{\mu_1}^{-1}\mathbf{S}_p^T - \mathbf{S}_p\mathbf{C}_{\mu_1}^{-1}\mathbf{S}_p^T\bar{\mathbf{H}}^T \\ &\times (\bar{\mathbf{H}}\mathbf{S}_r^{-1}\mathbf{S}_p\mathbf{C}_{\mu_1}^{-1}\mathbf{S}_p^T\bar{\mathbf{H}}^T + \mathbf{S}_r\mathbf{C}_{\mu_2}^{-1}\mathbf{S}_r^T)^{-1}\bar{\mathbf{H}}\mathbf{S}_p\mathbf{C}_{\mu_1}^{-1}\mathbf{S}_p^T \end{aligned} \tag{45}$$

According to the definition of $\mathbf{S}_i^{-1}$, we further have

$$\begin{aligned} \mathbf{D}_i &= \mathbf{S}_i^{-1} \begin{bmatrix} \hat{\mathbf{x}}_{i|i-1} \\ \mathbf{y}_i - \mathbf{h}(\hat{\mathbf{x}}_{i|i-1}) + \bar{\mathbf{H}}_i\hat{\mathbf{x}}_{i|i-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{S}_{p,i|i-1}^{-1}\hat{\mathbf{x}}_{i|i-1} \\ \mathbf{S}_r^{-1}\big(\mathbf{y}_i - \mathbf{h}(\hat{\mathbf{x}}_{i|i-1}) + \bar{\mathbf{H}}_i\hat{\mathbf{x}}_{i|i-1}\big) \end{bmatrix} \end{aligned} \tag{46}$$

Using (40), (44) and (46), it can be derived that

$$\begin{aligned} \mathbf{W}_i^T\mathbf{C}_i\mathbf{D}_i &= (\mathbf{S}_p^{-1})^T\mathbf{C}_{\mu_1}\mathbf{S}_p^{-1}\hat{\mathbf{x}}_{i|i-1} \\ &+ \bar{\mathbf{H}}^T(\mathbf{S}_r^{-1})^T\mathbf{C}_{\mu_2}\mathbf{S}_r^{-1}\big(\mathbf{y}_i - \mathbf{h}(\hat{\mathbf{x}}_{i|i-1}) + \bar{\mathbf{H}}_i\hat{\mathbf{x}}_{i|i-1}\big) \end{aligned} \tag{47}$$

Combining (39) and (47), we have

$$\mathbf{x}_i = \hat{\mathbf{x}}_{i|i-1} + \tilde{\mathbf{K}}_i(\mathbf{y}_i - \hat{\mathbf{y}}_{i|i-1}) \tag{48}$$

where

$$\tilde{\mathbf{K}}_i = \tilde{\mathbf{P}}_{i|i-1}\bar{\mathbf{H}}_i^T\big(\bar{\mathbf{H}}_i\tilde{\mathbf{P}}_{i|i-1}\bar{\mathbf{H}}_i^T + \tilde{\mathbf{R}}_i\big)^{-1} \tag{49}$$

$$\tilde{\mathbf{P}}_{i|i-1} = \mathbf{S}_{p,i|i-1}\mathbf{C}_{\mu_1,i}^{-1}(\mathbf{S}_{p,i|i-1})^T \tag{50}$$

$$\tilde{\mathbf{R}}_{i|i-1} = \mathbf{S}_{r,i}\mathbf{C}_{\mu_2,i}^{-1}(\mathbf{S}_{r,i})^T \tag{51}$$

Meanwhile, the corresponding covariance matrix is updated by

$$\mathbf{P}_{i|i} = (\mathbf{I} - \tilde{\mathbf{K}}_i\bar{\mathbf{H}}_i)\mathbf{P}_{i|i-1}(\mathbf{I} - \tilde{\mathbf{K}}_i\bar{\mathbf{H}}_i)^T + \tilde{\mathbf{K}}_i\mathbf{R}_i\tilde{\mathbf{K}}_i^T \tag{52}$$

Remark 1: The gross errors caused by the outlier in the measurements may significantly affect the performance of the state estimation techniques [28], such as the traditional UKF with MSE loss, especially while under large disturbances caused by the impulsive noises in the measurements. The GCL is on the other hand local similarity function determined by the free parameter kernel width, that is the main reason why the proposed GCL-UKF appears to be insensitive to the outliers.

B. Enhanced GCL-UKF

In power systems, noises in measurements may affect the innovation vector in (48) while is the predicted state vector at time instants i. To avoid such performance degradation, it is important to restrain the change of innovation vector caused by the noise (especially for non-Gaussian noise). Inspired by the idea in [28], an exponential function of the innovation vector is introduced into the GCL-UKF to update the covariance matrix by adjusting innovation vector, which can be expressed as

$$\hat{\mathbf{R}}_i = \tilde{\mathbf{R}}_i \exp\{-|\psi_i|\} \tag{53}$$

$$\tilde{\mathbf{R}}_i = \hat{\mathbf{R}}_i^{-1} \tag{54}$$

From (53) and (54), one can see that once any raw measurements encounter a significant outlier that results in an increase of absolute residual vector, the inversion of the absolute of residual vector can help suppress the influence such that the estimation performance can be improved. In addition, when the residual vector is considerably small, the exponential output shall at most approach the value of one. And the suppression function can be added to the system model without too much change, so it is very simple and applicable. In addition, the modified GCL-UKF by adding all previous effect in absolute exponential form can reduce the effect of gain because the covariance matrix is included in the gain matrix (49). Then an enhanced GCL-UKF (EnGCL-UKF) method can be obtained when (53) and (54) are used in GCL-UKF to update the in (51). In addition, the enhanced versions of EKF and UKF can also be obtained when the covariance matrices of measurement noise is updated by (53) and (54), which are termed as EnEKF and EnUKF in this work, respectively.

C. Power System State Estimation Based On EnGCL-UKF

For the m-bus power system, 2m-1 states are required to be estimated that include m voltage amplitude and m-1 phase angle states. To estimate the 2m-1 states of the power system using EnGCL-UKF, it needs to assume the initial state vector and the corresponding covariance matrix . According to (15), a set of 2n+1 sigma points can be calculated. Now, we give the detailed procedure of the proposed EnGCL-UKF algorithm for state estimation as follows.

Step 1. Set proper free parameters, and choose a small positive ε; set an initial estimate $\hat{\mathbf{x}}_{0|0} = \mathbb{E}[\mathbf{x}_0]$ and the corresponding covariance matrix $\mathbf{P}_0$; let $i = 1$;

Step 2. Use Eqs (15)-(18) to obtain the prior estimate $\hat{\mathbf{x}}_{i|i-1}$ and $\mathbf{P}_{i|i-1}$, and calculate $\mathbf{S}_{p,i|i-1}$ by Cholesky decomposition;

Step 3. Use Eqs (20)-(22) to compute the prior measurement $\hat{\mathbf{y}}_{i|i-1}$ and use Eqs (18) and (22) to acquire the measurement slope matrix $\bar{\mathbf{H}}_i$; construct the statistical linear regression model (26);

Step 4. Transform (25) into Eq (28), let $t = 1$ and set the initial value $\mathbf{x}_{i|i,t}$ at time instant $t = 0$ as $\mathbf{x}_{i|i,0}$;

Step 5. Use the following equation (55) to compute $\hat{\mathbf{x}}_{i|i}$;

$$\hat{\mathbf{x}}_{i|i,t} = \hat{\mathbf{x}}_{i|i-1} + \tilde{\mathbf{K}}_i(\mathbf{y}_i - \hat{\mathbf{y}}_{i|i-1}) \tag{55}$$

where $\tilde{\mathbf{K}}_i$ can be obtained by (49)-(51). The only difference for the realization of the proposed methods and the derived procedure is the residual error in (41), which is now given by

$$\mathbf{e}_{k,i} = \mathbf{d}_{k,i} - \mathbf{w}_{k,i}\mathbf{x}_{i|i,t-1} \tag{56}$$

In addition, $\tilde{\mathbf{R}}_i$ is updated by (53) and (54) to further improve the performance of the proposed approach;

Step 6. Compare the estimation at the current step and the estimation at the last step. If $\frac{\|\hat{\mathbf{x}}_{i|i,t} - \hat{\mathbf{x}}_{i|i,t-1}\|}{\|\hat{\mathbf{x}}_{i|i,t-1}\|} \leq \varepsilon$ holds, set $\hat{\mathbf{x}}_{i|i,k}$ and continue to step (7); otherwise, $t + 1 \rightarrow t$, and go back to step 5;

Step 7. Update the posterior covariance matrix by Eq (52), $i + 1 \rightarrow i$ and go back to step (2).

The estimation of state vector (consisting of magnitudes and angles of nodal voltage) could be conducted with the preset initials of and . The inclusion of a prediction stage for estimators permits a better filtering of measurement noises so that the estimation results can be improved [28].

IV. NUMERICAL RESULTS

The performance of the proposed GCL-UKF and EnGCL-UKF methods is evaluated in the case study on the IEEE 14-bus, 30-bus and 57-bus test systems, respectively. To investigate the effectiveness and the robustness of the proposed methods, the three test scenarios including non-Gaussian noise with outliers in measurements, abnormal situation of measurement gross error, and sudden load change condition are mainly considered in the following simulations. Without mentioned otherwise, the process noises are all assumed to be of the Gaussian distribution with zero-mean and variance 0.01.In the following experiments, 50 time-sample intervals are set to simulate the dynamic changes of power systems. The EKF, UKF, EnEKF and EnUKF are included for comparisons. Test results are presented in terms of the performance index $J(\text{in } p.u)$ of the overall measurement, and the mean absolute error (MAE) of the phase angle and amplitude of each node voltage [28]. All the results are obtained by averaging over 200 independent Monte Carlo runs to achieve more reliable simulation results and higher statistical significance. All the tests are executed with MATLAB 2016b running on i7-4510U and 2.60 GHZ CPU.

A. Case 1: non-Gaussian noise with outliers in measurements

First, we consider the case with mixed-Gaussian noises in measurements generated by

$$(1-\theta)N(\mu_1, v_1^2) + \theta N(\mu_2, v_2^2) \tag{57}$$

where $N(\mu_i, v_i^2)$ $(i = 1, 2)$ denotes the Gaussian distributions with mean values μ_i $(i = 1, 2)$ and variances σ_i^2 $(i = 1, 2)$, and θ is the mixture coefficient. The Gaussian distribution with a larger variance can model stronger impulsive noises (outliers). In this simulation, the parameters value $(\mu_1, \mu_2, v_1^2, v_2^2, \theta)$ of the mixed Gaussian distribution is set at $(0, 0, 1, 80, 0.15)$ to model the measurement noises. The initial state error covariance matrix $\mathbf{P}_0$ is set to $10^{-4}\mathbf{I}$. The free parameter α is set as 1.6, and the kernel width parameter $\gamma = 0.0019$. The overall performance indexes for the different test systems at each time instant i are given in Fig.1 (a)-(c). It is observed that 1) the proposed GCL-UKF performs better than EKF and UKF because of due to the enhancement by GCL; 2) the performance of the proposed EnGCL-UKF algorithm has significant advantages over other methods in any test system. Such results illustrate that the EnGCL-UKF can achieve higher filtering capacities and more robust performance during slow dynamic changes under the Non-Gaussian measurement noises, because the absolute exponential form can restrain the change of innovation vector.

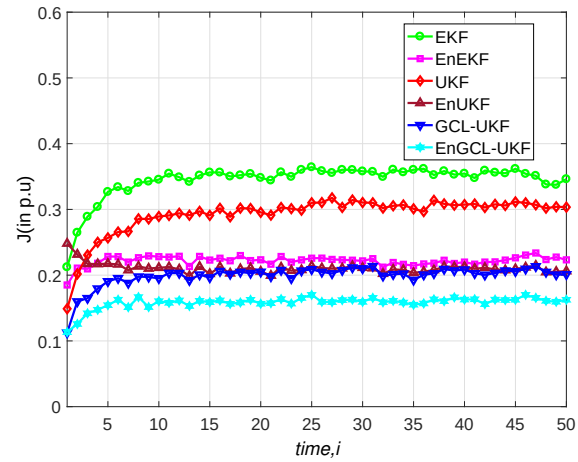
Since PMU measurement errors of real and reactive power flows calculated from voltage and phasors may follow Laplace distribution with outliers, the real and reactive power measurement noises generated by mixed Gaussian and Laplace distributions represented by (59) are considered in this case to further verify the feasibility of the proposed methods. We let the noises be presented as

$$0.95N(0, 100) + 0.15L(0, 1) \tag{58}$$

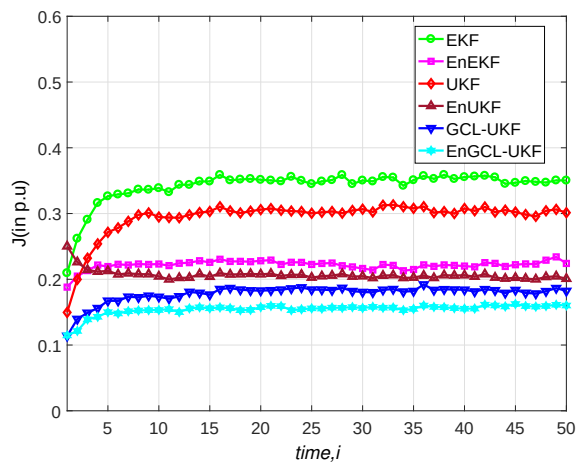
where $L(0, 1)$ denotes the noise drawn from the Laplace distribution with zero mean and scale 1. The average value of the overall performance index under different test systems are shown in Fig. 2. It is observed from the figure that the EnGCL-UKF performs better than other methods under this case in all test systems. Fig. 3 and Fig. 4 present the absolute error results of voltage amplitude and phase angle at the bus 3 in IEEE 14-bus test system. It is clear that results obtained by the proposed EnGCL-UKF are superior to those obtained by the other methods. In addition, the MAE and maximum value of absolute voltage amplitude and phase angle errors (denoted as MAE_θ, $MaxE_\theta$, $MaxE_v$ and MAE_v) at bus 3 in IEEE 30-test system are given in Table I. From this result, one can obtain the same conclusion above.

TABLE I
THE MAE AND MAXIMUM RESULTS OF VOLTAGE AMPLITUDE AND PHASE ANGLE

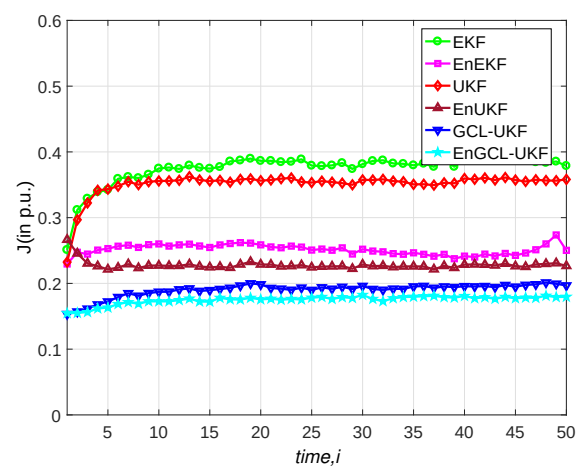
index	EKF	UKF	EnEKF	EnUKF	GCL-UKF	EnGCL-UKF
MAE_θ	0.0103	0.0092	0.0064	0.0058	0.0044	0.0039
$MaxE_\theta$	0.0131	0.0115	0.0083	0.0073	0.0056	0.0049
MAE_v	0.0057	0.0041	0.0036	0.0032	0.0029	0.0023
$MaxE_v$	0.0071	0.0049	0.0044	0.0043	0.0037	0.0027



(a) IEEE 14-bus test system



(b) IEEE 30-bus test system



(c) IEEE 57-bus test system

Fig. 1. Performance index $J(\text{in } p.u)$ under mixed-Gaussian measurement noise.

B. Case 2: Bad data Condition

The robustness of the proposed new estimator is further tested in the scenario where bad data are included in measurement. In this case, the real and the reactive power measurements are interrupted by some outliers which are modeled by expanding 20 at time index i =25. This test is only conducted in the IEEE

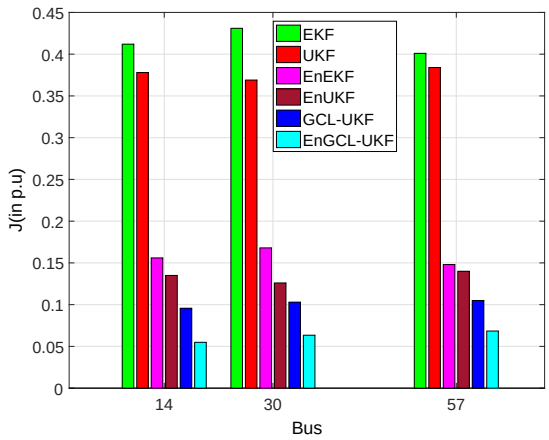


Fig. 2. The overall performance of EKF, UKF, GCL-UKF and EnGCL-UKF.

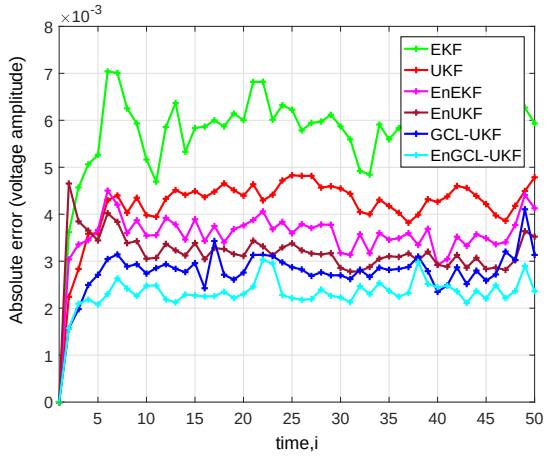


Fig. 3. Comparison error results of estimating the voltage amplitude at bus 3 in IEEE-14 bus system.

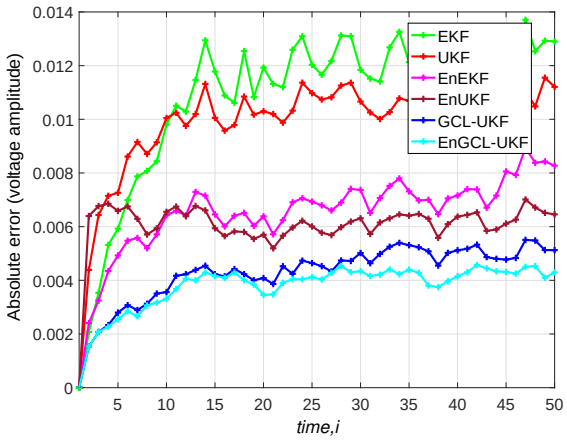


Fig. 4. Comparison error results of estimating voltage phase angle at bus 3 in IEEE-14 bus system.

30-bus system. Fig.5 presents the overall performance index for all time samples considered. From such results, we see that the proposed GCL-UKF and EnGCL-UKF perform better than EKF and UKF thanks to the robust performance of the GCL. Another observation we can make from these results is that the EnGCL-UKF still achieves best performance in comparison to GCL-UKF, EnEKF, and EnUKF. It demonstrates that the enhanced insensitivity of the proposed EnGCL-UKF partly coms from the GCL, and more significantly from the exponential updating mechanism in (53).

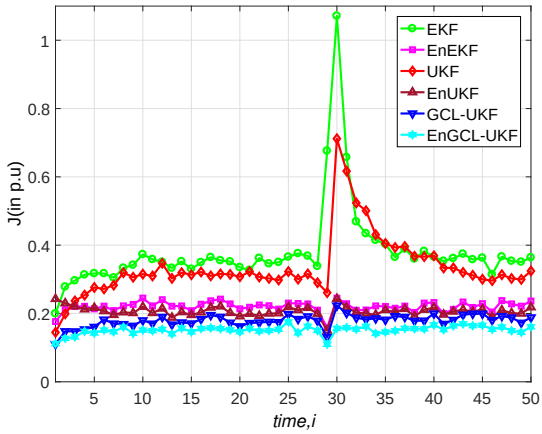


Fig. 5. The overall performance in IEEE 30-bus with large outliers at i = 25.

C. Case 3: Sudden load change condition

In this case, a sudden load changes scenario is simulated to investigate the robustness of the proposed algorithms. The estimation of stats at bus 3 in IEEE 30-bus test system at the 30th time sample is taken as an example, where the power of load at bus 3 has a 20 load drop at that time. The measurement noises are still generated by mixed-Gaussian model. The absolute voltage angle and amplitude error results at bus 3 are presented in Fig. 6 and Fig. 7. One can observe that the EKF, UKF, EnEKF, and EnUKF show obvious fluctuations at the time sample when the sudden load change occurs. The acceptable performance of the proposed EnGCL-UKF algorithm is easy to understand as it restrains the effects of load sudden change by the exponential updating mechanism. The estimation accuracy thus can be effectively improved.

V. Conclusion

To tackle the uncertainties caused by the non-Gaussian noises in power systems, this paper proposed a novel forecasting-aided state estimation method using GCL-UKF to estimate the state of the power system in realtime. To benefit from the high robustness of the generalized Correntropy loss (GCL), the GCL-UKF was derived by using the GCL to substitute the original cost function in the UKF framework, which manage to correctly estimate the state using the measurement results in the presence of non-Gaussian noises with outliers. To further improve the estimation performance of the proposed GCL-UKF, an exponential adjusting method is introduced to update the covariance matrix via adding all previous effect of the innovation vector in an absolute exponential form, which helps

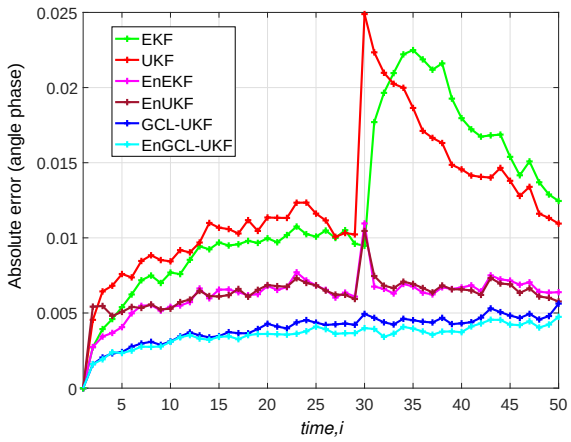


Fig. 6. The absolute voltage angle error at bus 3 in IEEE 30-bus test system.

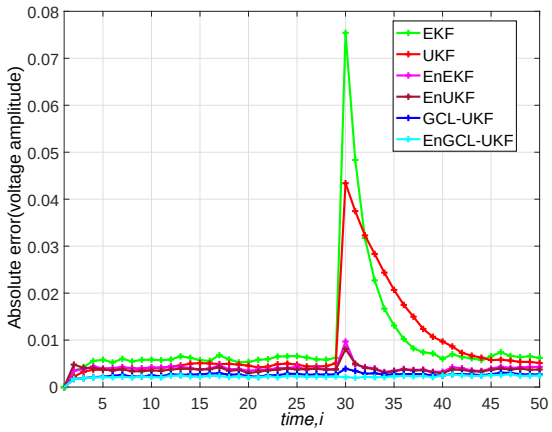


Fig. 7. The absolute voltage amplitude error at bus 3 in IEEE 30-bus test system.

reduce the effect of gain. Numerical simulations have been conducted for a variety of scenarios on IEEE 14-bus, 30-bus and 57-bus test systems, and the satisfactory efficiency and robustness of the proposed methods have been verified.

To further improve the estimation performance of the proposed methods, there are still some works to do, such as the optimization for the kernel width and the free parameter . Moreover, the proposed methods may be studied in the following aspects: 1) adopting the adaptive update method of the prediction error covariance [8-9]; 2) considering the unknown measurement noise statistics [17]. In addition, taking advantage of the particle filter (PF) approach for state estimation[29], a novel robust PF with GCL shall be developed in the future work.

References

[1] Y. Chen, F. Liu, S. Mei, J. Ma, "A Robust WLAV State Estimation Using Optimal Transformations," *IEEE Trans. Power Syst.*, vol.30, no.4, pp.2190-2191, July 2015.

[2] A. Ashok, M. Govindarasu, V. Ajjarapu. "Online Detection of Stealthy False Data Injection Attacks in Power System State Estimation,"*IEEE Trans. Smart Grid*, vol.9, no.3, pp.1636-1646, 2018

[3] R. Deng, G. Xiao, R. Lu, H. Liang, A. V. Vasilakos, " False Data Injection on State Estimation in Power Systems-Attacks, Impacts, and Defense: A Survey," *IEEE Trans. Ind. Inform.*, vol. 13, no. 2, pp. 411-423, April 2017

[4] J. Zhao, G. Zhang, L. A. Massimo, Z. Dong, C. Chen. "Short-Term State Forecasting-Aided Method for Detection of Smart Grid General False Data Injection Attacks," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1580-1590, July 2017.

[5] J. Zhao, G. Zhang, Z. Dong, K. P. Wong. "Forecasting-Aided Imperfect False Data Injection Attacks Against Power System Nonlinear State Estimation,"*IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 6-8, Jan 2016.

[6] J. Zhao, G. Zhang, Z. Dong, L. A. Massimo. "Robust Forecasting Aided Power System State Estimation Considering State Correlations,"*IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2658-2666, Jul 2018.

[7] J. 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, July 2017.

[8] L. Zanni, J. Le Boudec, R. Cherkaoui, M. Paolone, "A Prediction-Error Covariance Estimator for Adaptive Kalman Filtering in Step-Varying Processes: Application to Power-System State Estimation,"*IEEE Trans. Control Syst. Tech.*, vol. 25, no. 5, pp. 1683-1697, Sept. 2017.

[9] F. Sun, X. Hu, Y. Zou, S. Li., "Adaptive unscented Kalman filtering for state of charge estimation of a lithium-ion battery for electric vehicles,"*Energy*, vol.36, no.5, pp.3531-3540, May 2011.

[10] G. Valverde, V. Terzija, "Unscented Kalman Filter for power system dynamic state estimation," *IET Generation, Transmission and Distribution*, vol. 5, no.1, pp. 29-37, Jan. 2011.

[11] 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.

[12] A. Rouhani, A. Abur, "Constrained Iterated Unscented Kalman Filter for Dynamic State and Parameter Estimation,"*IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2404-2414, May 2018.

[13] J. Qi, K. Sun, J. Wang, H. Liu, "Dynamic State Estimation for Multi-Machine Power System by Unscented Kalman Filter With Enhanced Numerical Stability,"*IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1184-1196, March 2018.

[14] A. Abur, A. Gomez-Exposito , Power System State Estimation Theory and Implementation, New York, NY, USA: Marcel Dekker, 2004.

[15] N. Bretas, A. S. Bretas, A two steps procedure in state estimation gross error detection, identification, and correction, *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 484-490, 2015.

[16] J. Zhao, L. Mili, A Framework for Robust Hybrid State Estimation With Unknown Measurement Noise Statistics,*IEEE Trans. Ind. Inform.*, vol.14, no.5, pp.1866-1875, May, 2018.

[17] J. Zhao, "Dynamic State Estimation With Model Uncertainties Using H Extended Kalman Filter,"*IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1099-1100, Jan. 2018.

[18] J. Zhao, L. Mili, "A Robust Generalized-Maximum Likelihood Unscented Kalman Filter for Power System Dynamic State Estimation,"*IEEE J. Selected Topics in Signal Process.*, vol. 12, no. 4, pp. 578-592, Aug. 2018.

[19] W. Liu, P. P. Pokharel, J. C. Principe, "Correntropy: Properties and Applications in Non-Gaussian Signal Processing,"*IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286-5298, Nov. 2007.

[20] B. Chen, X. Liu, H. Zhao, J. Qin, J. Cao, "Maximum Correntropy Kalman Filter,"*Automatica*, vol. 76, pp. 70-77, Feb. 2017.

[21] L. Xi, B. Chen, B. Xu, Z. Wu, P. Honeine. "Maximum Correntropy unscented filter."*Int. J. Syst. Sci.*, vol. 48, no. 8, pp. 1607-1615, Jan. 2017.

[22] B. Chen, L. Xing, H. Zhao, N. Zheng and J. C. Principe, "Generalized Correntropy for Robust Adaptive Filtering,"*IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3376-3387, July, 2016.

[23] X. Luo, J. Sun, L. Wang, W. Wang, et al., "Short-term Wind Speed Forecasting via Stacked Extreme Learning Machine With Generalized Correntropy,"*IEEE Trans. Ind. Inform.*, vol.14, no.11, pp.4936-4971, Nov. 2018.

[24] S. Hakimi, G. Abed Hodtani. "Generalized maximum Correntropy detector for non-Gaussian environments."*Int. J. Adapt. Control and Signal Process.*, vol. 32, no.10, pp. 83-97, Aug. 2018.

[25] J. Zhao and H. Zhang, "Kernel Recursive Generalized Maximum Correntropy,"*IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1832-1836, Dec. 2017.

[26] A. M. Leite da Silva, M. B. Do Coutto Filho, J. F. de Queiroz, "State forecasting in electric power systems,"*IEE Proceedings C - Generation, Transmission and Distribution*, vol. 130, no. 5, pp. 237-244, September 1983.

[27] M. A. Gandhi and L. Mili, "Robust Kalman Filter Based on a Generalized Maximum-Likelihood-Type Estimator,"*IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2509-2520, May 2010.
[28] K. R. Shih, S. J. Huang, "Application of a robust algorithm for dynamic state estimation of a power system,"*IEEE Trans. Power Syst.*, vol. 17, no. 1, pp. 141-147, Feb. 2002.
[29] K. Emami, T. Fernando, H. H. Iu, H. Trinh, K. P. Wong, "Particle Filter Approach to Dynamic State Estimation of Generators in Power Systems,"*IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2665-2675, Sept. 2015.

Wentao Ma is currently a lecturer with the School of Automation and Information Engineering at the Xi'an University of Technology, Xi'an, China. He received the B.S. degree in Mathematics and Applied Mathematics from Shannxi University of Technology, in 2003, the M.S. degree in Computing Mathematics from Huazhong University of Science and Technology, in 2008, and the Ph.D. degree in Information and Communication Engineering from Xi'an Jiaotong University, in 2015. His research interests include machine learning, information theory, adaptive signal processing, and their applications in Electrical and Computer Engineering.

Jinzhe Qiu is currently a M.S student with the School of Electronic and Information Engineering, Xian University of technology, Xian, China. He received the B.S. degree in electrical engineering from Xuchang University, in 2017. His research interests include state estimation and frequency estimation of power system.

Xinghua Liu (M16)received the B.Sc. from Jilin University, Changchun, China, in 2009; and the Ph.D. degree in Automation from University of Science and Technology of China, Hefei, in 2014. From 2014 to 2015, he was invited as a visiting fellow at RMIT University in Melbourne, Australia. From 2015 to 2018, He was a Research Fellow at the School of Electrical and Electronic Engineering in Nanyang Technological University, Singapore. Dr.Liu has joined Xian University of Technology as a professor since September 2018. His research interest includes state estimation and control, intelligent systems, autonomous vehicles, cyber-physical systems, robotic systems, etc.

Gaoxi Xiao (M99-SM18) received the B.S. and M.S. degrees in applied mathematics from Xidian University, Xi'an, China, in 1991 and 1994 respectively. He was an Assistant Lecturer in Xidian University in 1994-1995. In 1998, he received the Ph.D. degree in computing from the Hong Kong Polytechnic University. He was a Postdoctoral Research Fellow in Polytechnic University, Brooklyn, New York in 1999; and a Visiting Scientist in the University of Texas at Dallas in 1999-2001. He joined the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2001, where he is now an Associate Professor. His research interests include complex systems and complex networks, communication networks, smart grids, and system resilience and risk management. Dr. Xiao serves/served as an Associate Editor or Guest Editor for IEEE Transactions on Network Science and Engineering, PLOS ONE and Advances in Complex Systems etc., and a TPC member for numerous conferences including IEEE ICC and IEEE GLOBECOM etc.

Jiandong Duan received the B.S. degree from Huazhong University of Science and Technology,Hubei, China, in 1995, the M.S. degree from Guangxi University, Nanning,China, in 1998, and the Ph.D. degree from Xian Jiaotong University, Xian,China, all in electrical engineering. He is currently a Professor with the School of Automation and Information engineering, Xian University of Technology. His research interests include relay protection, distributed energy supply network, and smart grid monitoring technology.

Badong Chen received the B.S. and M.S. degrees in control theory and engineering from Chongqing University, in 1997 and 2003, respectively, and the Ph.D. degree in computer science and technology from Tsinghua University in 2008. He was a Postdoctoral Researcher with Tsinghua University from 2008 to 2010, and a Postdoctoral Associate at the University of Florida Computational Neuro-Engineering Laboratory (CNEL) during the period October, 2010 to September, 2012. Currently he is a professor at the Institute of Artificial Intelligence and Robotics (IAIR), Xian Jiaotong University. His research interests are in signal processing, machine learning,and their applications to neural engineering. He has published 2 books,4 chapters,and over 200 papers in various journals and conference proceedings. Dr.Chen is an IEEE Senior Member, a Technical Committee Member of IEEE SPS Machine Learning for Signal Processing (MLSP) and IEEE CIS Cognitive and Developmental Systems (CDS), and an associate editor of IEEE Transactions on Cognitive and Developmental Systems,IEEE Transactions on Neural Networks and Learning Systems and Journal of The Franklin Institute, and has been on the editorial board of Entropy.