

# Improved Particle Filter for Non-Gaussian Forecasting-aided State Estimation

Lyuzerui Yuan, Jie Gu, Honglin Wen, and Zhijian Jin

**Abstract**—Gaussian assumptions of non-Gaussian noises hinder the improvement of state estimation accuracy. In this paper, an asymmetric generalized Gaussian distribution (AGGD), as a unified representation of various unimodal distributions, is applied to formulate the non-Gaussian forecasting-aided state estimation problem. To address the problem, an improved particle filter is proposed, which integrates a near-optimal AGGD proposal function and an AGGD sampling method into the typical particle filter. The AGGD proposal function can approximate the target distribution of state variables to greatly alleviate particle degeneracy and promote precise estimation, through considering both state transitions and latest measurements. For rapid particle generation from the AGGD proposal function, an efficient inverse cumulative distribution function (CDF) sampling method is employed based on the derived approximation of inverse CDF of AGGD. Numerical simulations are carried out on a modified balanced IEEE 123-bus test system. The results validate that the proposed method outperforms other popular state estimation methods in terms of accuracy and robustness, whether in Gaussian, non-Gaussian, or abnormal measurement errors.

**Index Terms**—State estimation, particle filter, asymmetric generalized Gaussian distribution, non-Gaussian noise.

## I. INTRODUCTION

AS a significant tool in power system monitoring and control, forecasting-aided state estimation (FASE, also referred to as dynamic state estimation [1]) has attracted wide attention [2]. It seeks to estimate the power system states, i.e., voltage magnitudes and phase angles, via system process trends and limited noisy measurements. In general, the FASE is described by the transition function of state variables and a function that maps the power system states to measurements. Both of them are blurred by uncertainties, which are referred to as process noise and measurement noise, respectively.

Since the FASE inevitably involves the aforementioned

noises, researchers often assume that the noises follow specific families of distributions such as Gaussian distribution and student- $t$  distribution. Although Gaussian distribution is widely used in existing works due to its nice properties in algebra, increasing works suggest that both measurement noise and process noise are non-Gaussian variables [3]. For instance, [4] presents that phasor measurement unit errors obey heavy-tailed distributions rather than a short-tailed Gaussian distribution. In cases of non-Gaussian noises, Gaussian assumptions may lead to significantly biased estimates [5], [6]. This has motivated many FASE methods applicable to non-Gaussian noise environments. It has been proposed to embed information theoretic criteria, e.g., minimum error entropy [7], into traditional FASE methods, in which the criteria can filter out non-Gaussian noises by Gaussian kernel functions [7]-[10]. However, the runtime of these improved methods are usually more than twice that of the original ones, due to high computational complexity of the criteria. Another route assumes specific distributions allowing for non-Gaussianity in noises. Although several non-Gaussian statistical distributions, e.g., student- $t$  distribution [11], have been validated to improve the robustness of FASE [11]-[13], they cannot deal with mixed noise systems, e.g., some noises follow the asymmetric Laplace distribution while the others follow student- $t$  distribution. Thus, research on FASE methods is needed with more generalized distribution assumptions of noises, to achieve accurate and robust estimation in complicated noise environments.

To develop such a method, particle filter (PF) is selected as a basic method, as it can cope with arbitrary noise distribution theoretically through the Monte Carlo simulation method [14]. Specifically, a sampling distribution (known as proposal distribution) is predefined to generate numerous particles. It is almost always Gaussian for convenient sampling. Each particle has a weight that is associated with the distributions of process noise, measurement noise, and the proposal. The overall weighted particles are regarded as a numerical approximation for the target distribution of state variables. Under conditions of non-Gaussian process and measurement noises, the target distribution also has non-Gaussianity [14]. In this situation, the Gaussian proposal distribution may greatly deviate from the target distribution, leading to particle degeneracy, i.e., a great majority of particles have negligible weights to cause unreliable or even divergent estimates [15]. The problem can be alleviated by constructing a non-Gaussian proposal distribution close to the target distribution and providing the corresponding sampling method.

Manuscript received: December 17, 2021; revised: March 23, 2022; accepted: May 13, 2022. Date of CrossCheck: May 13, 2022. Date of online publication: June 27, 2022.

This work was supported by the National Key Research and Development Program of China (No. 2016YFB0900100) and the Key Project of Shanghai Science and Technology Committee (No. 18DZ1100303).

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

L. Yuan, J. Gu (corresponding author), H. Wen, and Z. Jin are with School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: ylzr.z@sytu.edu.cn; gujie@sytu.edu.cn; linlin00@sytu.edu.cn; zjjin@sytu.edu.cn).

DOI: 10.35833/MPCE.2021.000805



These improvements for PF will yield a competitive FASE method applicable to general noise distributions. In this paper, we make an asymmetric generalized Gaussian distribution (AGGD) assumption for noise in the FASE problem, where the AGGD is a family of unimodal distributions and can represent many symmetric/asymmetric statistical distributions [16], e. g., Laplace, Gaussian, logistic, student- $t$ , and skewed distributions. Given this assumption, a generalized particle filter (GPF) method is proposed to solve the FASE problem, which integrates a near-optimal AGGD proposal distribution and a fast sampling method into the framework of PF. The AGGD proposal distribution, inspired by the GMapping algorithm in [15], is established by combining both state transition information and current measurements. It can approximate the target distribution to relieve the risk of particle degeneracy and then facilitate accurate estimation. For rapidly sampling from the AGGD proposal distribution, we select the inverse CDF sampling method because of its higher efficiency compared with other popular methods such as the Markov chain Monte Carlo sampling [17]. Considering that the inverse CDF of AGGD cannot be expressed explicitly [16], its approximate expression is derived based on a Gaussian piecewise function of the AGGD.

Experiments are performed on a balanced IEEE 123-bus test system with eight distributed generators (DGs) [18]. The results demonstrate that the estimation accuracy of GPF outperforms that of the seven benchmarks, i. e., weighted least squares [19], extended Kalman filter (EKF) [1], cubature Kalman filter (CKF), unscented Kalman filter (UKF) [3], UKF with generalized correntropy loss (GCL-UKF) [10], PF [14], and unscented PF (UPF) [20], in the presence of Gaussian and non-Gaussian noises. Thanks to the efficient inverse CDF sampling, the AGGD proposal distribution can facilitate superior estimates without sacrificing real time of GPF. Furthermore, it is verified that the GPF is robust under conditions of abnormal measurements. The main contributions are outlined as follows.

- 1) An asymmetric generalized Gaussian distribution (AGGD) is introduced for the assumptions of noises in FASE, which can characterize many popular statistical distributions.
- 2) A near-optimal AGGD proposal distribution close to the target distribution is established to alleviate particle degeneracy.
- 3) An accurate and robust PF method is proposed to address the non-Gaussian FASE problem, which embeds the AGGD proposal distribution and an efficient AGGD sampling method into the typical PF.

The rest of this paper is organized as follows. Section II reviews the related work. Section III presents the formulation of the FASE problem under AGGD assumptions of noises. The proposed GPF is presented in Section IV. Section V presents case study and illustrates the competitive performances of the GPF. Conclusion is drawn in Section VI.

## II. RELATED WORK

For better performances of FASE methods in non-Gaussian noise environments, two main routes are optimality criteria against noises, and appropriate distribution assumptions

of noises. Their pros and cons are shown in Table I.

TABLE I  
COMPARISON OF TWO ROUTES FOR DEALING WITH NON-GAUSSIAN NOISE

Route	Basic method	Criterion/distribution	Pros (+) and cons (-)
Robust criteria to filter out non-Gaussian noises	UKF, EKF	Minimum error entropy [7]	(+ Wide application (-) Heavy time consumption
		Maximum correntropy [8]	
		Cross-correntropy [9]	
		Generalized correntropy [10]	
	CKF	Inflatable noise variance [21]	
Non-Gaussian distribution assumptions of noises	UKF, PF	Square-root embedded cubature rule [22]	(+ Wide application (-) Heavy time consumption
		GMM [12], [23]	
	PF	Student- $t$ distribution [13]	(+ Time-efficiency (-) Limited application
		Asymmetric Laplace distribution [17]	
		Generalized Gaussian distribution [24]	(+ Wide application, time-efficiency (-) Particle degeneracy

The popular Kalman-based filters, i. e., Kalman filter (KF) and their variants, are generally improved through optimality criteria, as they cannot achieve optimal estimation with respect to the minimum mean square error under non-Gaussian noise assumptions [6]. The studies in [7]-[10] replace the mean square error loss of the Kalman-based filters with entropy criteria, e. g., the minimum error entropy [7], the maximum correntropy [8], the cross-correntropy [9], and the generalized correntropy [10]. These criteria use Gaussian kernel functions to filter out non-Gaussian noises. A non-entropy-based criterion named inflatable noise variance is designed in [21] to apply to the EKF, which compares actual noise covariances with the ideal Gaussian covariances to identify non-Gaussian noises. Reference [22] proposes a robust square-root embedded cubature rule dedicated to the CKF. Although the above criteria are sensitive to non-Gaussian disturbances, they make the improved methods consume more than twice computational time compared with the original ones.

Non-Gaussian noise assumptions are described by the Gaussian mixture model (GMM) or heavy-tailed statistical distributions. The GMM is a combination of Gaussian distributions to approximate any continuous distribution. It can be applied to improve Kalman-based filters, as well as combined with the PF, e. g., a GMM-based generalized CKF [12] and a GMM-based expectation-maximization PF [23]. However, parameter selection of the GMM is quite time-consuming. A more efficient way is to directly utilize specific distributions. Accordingly, PF becomes the most commonly used basic method [14]. Reference [13] assumes the student- $t$  distribution of measurement noise and addresses it by a variational Bayesian PF. Based on the asymmetric Laplace distribution assumption, an optimum proposal distribution of PF is derived in [17], which contains both measurement and process information to alleviate particle degeneracy. Nevertheless, these assumptions are only available for specific noise statistics. Reference [24] assumes that the noise follows a

more general distribution, i. e., generalized Gaussian function, which can describe multiple unimodal symmetric distributions. Under this non-Gaussian assumption, however, the Gaussian proposal distribution in [24] may cause particle degeneracy and large estimate errors. Thus, this paper will construct a near-optimal proposal distribution to form an accurate and robust PF method, under a generalized distribution assumption.

### III. FORMULATION OF FASE PROBLEM

This section introduces an AGGD. With noises satisfying the AGGD, the FASE problem is formulated and the typical PF algorithm is described to solve this problem.

#### A. AGGD

AGGD is a family of unimodal distributions to provide modeling of generic noise statistics [16]. It can describe a large group of symmetric/asymmetric statistical distributions, e.g., impulsive, Gaussian, logistic, student- $t$ , and skewed distributions. The probability density function (PDF) of AGGD is given by:

$$p(x|\mu, \alpha, \sigma_l^2, \sigma_r^2) = \begin{cases} \frac{\alpha \cdot \exp\left(-\left(\frac{-(x-\mu)}{\beta_l}\right)^\alpha\right)}{(\beta_l + \beta_r)\Gamma\left(\frac{1}{\alpha}\right)} & x - \mu < 0 \\ \frac{\alpha \cdot \exp\left(-\left(\frac{x-\mu}{\beta_r}\right)^\alpha\right)}{(\beta_l + \beta_r)\Gamma\left(\frac{1}{\alpha}\right)} & x - \mu \geq 0 \end{cases} \quad (1)$$

$$\begin{cases} \beta_l = \sigma_l \sqrt{\frac{\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{3}{\alpha}\right)}} \\ \beta_r = \sigma_r \sqrt{\frac{\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{3}{\alpha}\right)}} \end{cases} \quad (2)$$

where  $\Gamma(\cdot)$  is the Gamma function;  $\mu$  is the mode, i. e., the value that is most likely to be sampled; and  $\alpha$  is the shape parameter that represents the exponential rate of decay. This paper considers the range of  $\alpha$  as  $[1, +\infty)$  that is general enough to hold different sharpness of noise distributions [16], [25].  $\sigma_l^2$  and  $\sigma_r^2$  ( $\sigma_l, \sigma_r > 0$ ) control the skewness and width, named left variance and right variance, respectively. For convenience, we denote an AGGD as  $\mathcal{AGGD}(\mu, \alpha, \sigma_l^2, \sigma_r^2)$ . Many popular PDFs can be represented by setting parameters of AGGD, e.g.,  $\mathcal{AGGD}(\mu, 1, \sigma^2, \sigma^2)$  and  $\mathcal{AGGD}(\mu, 2, \sigma^2, \sigma^2)$  are the Laplace and Gaussian PDFs, respectively.

#### B. FASE

The states of power systems usually refer to voltage magnitudes and phase angles of buses. Although we hardly ob-

tain their true values because of the limited noisy measurements, FASE can provide an optimal estimate of states, according to state transition characteristics, measurements, and noise statistics. Concretely, the state transitions give a priori knowledge of system process trends. The measurements are used to acquire the likelihood that the estimates are equal to the true states. They include the data received from measurement devices such as phasor measurement unit (PMU) and supervisory control and data acquisition (SCADA) system, and the pseudo-measurements generated by algorithms. Noises are inevitably involved in state transitions and measurements, which should also be considered into the FASE.

Based on the above, the non-Gaussian FASE problem can be formulated. Given a power system with  $n$  nodes, we denote its state variables at timestep  $t$  as  $\mathbf{x}_t = [\theta_{1,t}, \theta_{2,t}, \dots, \theta_{n,t}, V_{1,t}, V_{2,t}, \dots, V_{n,t}]^T$ , where  $\theta_{i,t}$  and  $V_{i,t}$  represent the voltage phase angle and magnitude of node  $i$  at timestep  $t$ , respectively. To describe the process trends of  $\mathbf{x}_t$ , a discrete-time first-order Markov model is applied [26], i. e.,

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t) + \boldsymbol{\omega}_t \quad (3)$$

where  $f(\cdot)$  is a nonlinear state-transition function. For convenience, it is typically linearized by the Holt-Winters double exponential smoothing. More details are given in [26].  $\boldsymbol{\omega}_t$  represents the process noise. Its distribution is characterized by a known PDF called priori PDF  $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ , i. e.,  $\boldsymbol{\omega}_t \sim p(\mathbf{x}_{t+1}|\mathbf{x}_t)$ . In this paper,  $\boldsymbol{\omega}_t \in \boldsymbol{\omega}_t$  follows  $\mathcal{AGGD}(0, \alpha_\omega, \sigma_{\omega_l}^2, \sigma_{\omega_r}^2)$ , where  $\alpha_\omega$  is the shape parameter of the AGGD that  $\boldsymbol{\omega}_t$  follows; and  $\sigma_{\omega_l}^2$  and  $\sigma_{\omega_r}^2$  are the left variance and right variance of the AGGD that  $\boldsymbol{\omega}_t$  follows, respectively. Besides, we denote all measurements at timestep  $t$  as a measurement vector  $\mathbf{z}_t$ , and assume that the measurement noise variables are independent. The equation between  $\mathbf{z}_t$  and  $\mathbf{x}_t$  is given by:

$$\mathbf{z}_t = h(\mathbf{x}_t) + \mathbf{v}_t \quad (4)$$

where  $h(\cdot)$  is the nonlinear measurement function described in [26]; and  $\mathbf{v}_t$  is the measurement noise. Similar to  $\boldsymbol{\omega}_t$ ,  $\mathbf{v}_t$  is associated with a given PDF called likelihood PDF  $p(\mathbf{z}_t|\mathbf{x}_t)$ , i. e.,  $\mathbf{v}_t \sim p(\mathbf{z}_t|\mathbf{x}_t)$ , and  $\mathbf{v}_t \in \mathbf{v}_t$  follows  $\mathcal{AGGD}(0, \alpha_v, \sigma_{v_l}^2, \sigma_{v_r}^2)$ , where  $\alpha_v$  is the shape parameter of the AGGD that  $\mathbf{v}_t$  follows; and  $\sigma_{v_l}^2$  and  $\sigma_{v_r}^2$  are the left variance and right variance of the AGGD that  $\mathbf{v}_t$  follows, respectively. Note that  $\mathbf{v}_t$  and  $\boldsymbol{\omega}_t$  are independent.

The priori PDF  $p(\mathbf{x}_t|\mathbf{x}_{t-1})$  in (3) and the likelihood PDF  $p(\mathbf{z}_t|\mathbf{x}_t)$  in (4) provide priori and likelihood statistics of  $\mathbf{x}_t$ , respectively. According to the Bayesian theory [14], these two PDFs can be used to obtain full statistical information of  $\mathbf{x}_t$  that is referred to as posterior PDF  $p(\mathbf{x}_t|\mathbf{z}_{1:t})$ , where  $\mathbf{z}_{1:t} = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_t\}$  represents the all measurements up to timestep  $t$ . The posterior PDF  $p(\mathbf{x}_t|\mathbf{z}_{1:t})$  is considered as the target PDF of state variables, which is expressed as:

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}) = \frac{p(\mathbf{z}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{z}_{1:t-1})}{\int p(\mathbf{z}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{z}_{1:t-1})d\mathbf{x}_t} \quad (5)$$

$$p(\mathbf{x}_t|\mathbf{z}_{1:t-1}) = \int p(\mathbf{x}_{t-1}|\mathbf{z}_{1:t-1})p(\mathbf{x}_t|\mathbf{x}_{t-1})d\mathbf{x}_{t-1} \quad (6)$$

With the initial PDF  $p(\mathbf{x}_0)$  known,  $p(\mathbf{x}_t|\mathbf{z}_{1:t})$  can be ob-

tained by recursively calculating (5) and (6). We use it to calculate the optimal state estimate  $\hat{\mathbf{x}}_t$  via the maximum posterior probability estimator [27], i.e.,

$$\hat{\mathbf{x}}_t = E\{\mathbf{x}_t | \mathbf{z}_{1:t}\} = \int \mathbf{x}_t p(\mathbf{x}_t | \mathbf{z}_{1:t}) d\mathbf{x}_t \quad (7)$$

### C. PF

In practice, the integrals in (5) and (6) usually cannot be calculated analytically. To cope with the problem, the PF is proposed to approximate  $p(\mathbf{x}_t | \mathbf{z}_{1:t})$  using the Monte Carlo simulation method [14]. First,  $N_p$  particles  $\{\mathbf{x}_t^i\} (i=1, 2, \dots, N_p)$  are generated from a predetermined PDF called proposal PDF  $q(\mathbf{x}_t)$ , where each particle is regarded as an estimate of state variables and the superscript  $i$  represents the  $i^{\text{th}}$  particle. Then, their corresponding weights  $\{w_t^i\} \{\mathbf{x}_{c,t}^i\} (i=1, 2, \dots, N_p)$  ( $\sum_{i=1}^{N_p} w_t^i = 1$ ) are given by:

$$w_t^i = w_{t-1}^i \frac{p(\mathbf{z}_t | \mathbf{x}_t^i) p(\mathbf{x}_t^i | \mathbf{x}_{t-1}^i)}{q(\mathbf{x}_t^i)} \quad (8)$$

where  $w_{t-1}^i$  is the weight of the  $i^{\text{th}}$  particle at timestep  $t-1$ . Finally, the posterior PDF is given by:

$$p(\mathbf{x}_t | \mathbf{z}_{1:t}) \approx \sum_{i=1}^{N_p} w_t^i \delta(\mathbf{x}_t - \mathbf{x}_t^i) \quad (9)$$

where  $\delta(\cdot)$  is the standard Dirac delta function. In terms of (7) and (9), the optimal estimate of  $\mathbf{x}_t$  is:

$$\hat{\mathbf{x}}_t \approx \int \mathbf{x}_t \sum_{i=1}^{N_p} w_t^i \delta(\mathbf{x}_t - \mathbf{x}_t^i) d\mathbf{x}_t = \sum_{i=1}^{N_p} w_t^i \mathbf{x}_t^i \quad (10)$$

Typically, the proposal PDF  $q(\mathbf{x}_t)$  is set equal to the priori PDF  $p(\mathbf{x}_t | \mathbf{x}_{t-1})$  [14], which can simplify the weight equation in (8) to  $w_t^i = w_{t-1}^i p(\mathbf{z}_t | \mathbf{x}_t^i)$ . However, it may much differ from the likelihood PDF, resulting in particle degeneracy, as shown in Fig. 1.

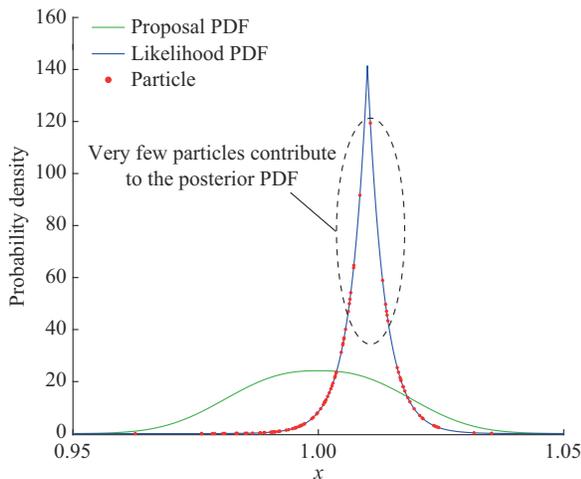


Fig. 1. An example of one-dimensional particle degeneracy in typical PF.

Moreover, the priori distribution in this paper follows an AGGD rather than a Gaussian distribution to make sampling difficult. To overcome the two issues, we should construct a

near-optimal AGGD proposal PDF that holds both priori knowledge and likelihood information [17], and develop an efficient AGGD sampling method.

## IV. GPF

In this section, a GPF method is proposed to solve the non-Gaussian FASE problem, which combines a near-optimal proposal PDF and the inverse CDF sampling method with the typical PF. Its block diagram is illustrated in Fig. 2.

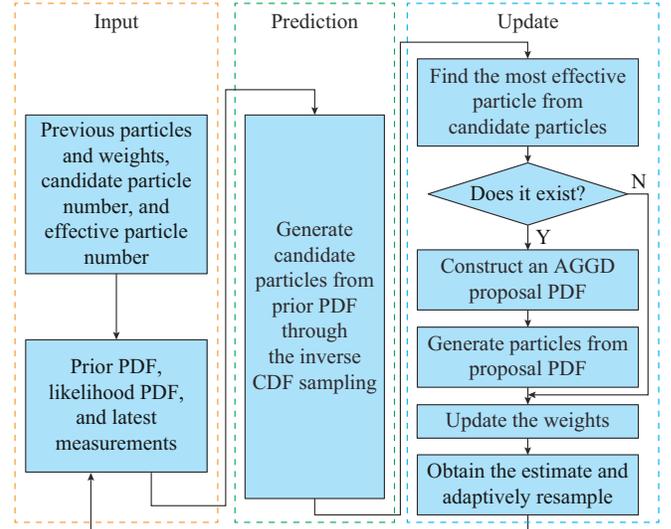


Fig. 2. Block diagram of GPF.

### A. AGGD Proposal Distribution

Referring to [15], [17], an optimum proposal PDF  $q^*(\mathbf{x}_t)$  should consider both process and measurement information to overcome particle degeneracy and facilitate high-precision estimation, i.e.,

$$q^*(\mathbf{x}_t) = p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{z}_t) = \frac{p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{x}_{t-1})}{\int p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{x}_{t-1}) d\mathbf{x}_t} = \frac{\hat{q}^*(\mathbf{x}_t)}{\lambda} \quad (11)$$

where  $\lambda = \int p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{x}_{t-1}) d\mathbf{x}_t$  is a constant; and  $\hat{q}^*(\mathbf{x}_t) = p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{x}_{t-1})$  is the unnormalized optimum proposal PDF. Although  $\hat{q}^*(\mathbf{x}_t)$  is the best choice to generate particles, we hardly acquire its exact expression. Therefore, an accessible near-optimal proposal PDF needs to be built.

In general, the likelihood PDF  $p(\mathbf{z}_t | \mathbf{x}_t)$  is more peaked compared with the priori PDF, due to the high precision of measurement devices but high volatility of state transitions. This infers that the sharpness of the optimum proposal PDF  $p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{z}_t)$  is similar to  $p(\mathbf{z}_t | \mathbf{x}_t)$ . Accordingly, we define an effective likelihood area, i.e.,  $L_t = \{\mathbf{x}_t | p(\mathbf{z}_t | \mathbf{x}_t) > \Delta\}$ , to judge whether the particles can be regarded as the feasible states, where  $\Delta$  is a fixed parameter to limit the scale of  $L_t$ . In this paper, we set  $\Delta = \max(p(\mathbf{z}_t | \mathbf{x}_t))/2$ . Inspired by the GMapping algorithm in [15],  $L_t$  is characterized via particles to establish a near-optimal proposal PDF.

First,  $N_p$  points  $\{\mathbf{x}_{c,t}^i\} (i=1, 2, \dots, N_p)$  named candidate particles are sampled from the priori PDF  $p(\mathbf{x}_t | \mathbf{x}_{t-1})$ . Under con-

dition of  $\exists \mathbf{x}_{c,t}^i \in L_r$ , the most effective candidate particle  $\tilde{\mathbf{x}}_t$  is selected by:

$$\tilde{\mathbf{x}}_t = \arg \max_x p(\mathbf{z}_t | \mathbf{x}_t) \quad \mathbf{x}_t \in \{\mathbf{x}_{c,t}^i\} \cap L_t \quad (12)$$

The solution to the optimization problem in (12), i.e.,  $\tilde{\mathbf{x}}_t$ , can be obtained through the traversal of likelihood values of all candidate particles, as shown in Algorithm 1.

---

**Algorithm 1:** find the most effective candidate particle at timestep  $t$

---

**Input:** candidate particles  $\{\mathbf{x}_{c,t}^i\} (i=1, 2, \dots, N_p)$ , measurements  $\mathbf{z}_t$ , likelihood PDF  $p(\mathbf{z}_t | \mathbf{x}_t)$ , and scale parameter  $\Delta$  of  $L_t$

**Output:** the most effective candidate particle  $\tilde{\mathbf{x}}_t$

```

1:  $L_{\max} = 0$ 
2: for  $j = 1, 2, \dots, N_p$  do
3: //Calculate the likelihood value of  $\mathbf{x}_{c,t}^j$ 
4:  $L^j = p(\mathbf{z}_t | \mathbf{x}_{c,t}^j)$ 
5: //Judge whether  $\mathbf{x}_{c,t}^j$  is the most effective candidate particle in  $\{\mathbf{x}_{c,t}^i\} (i=1, 2, \dots, j)$ 
6: if  $L^j > \Delta$  and  $L^j > L_{\max}$  then
7:    $\tilde{\mathbf{x}}_t = \mathbf{x}_{c,t}^j$ ,  $L_{\max} = L^j$ 
8: end if
9: end for

```

---

Since the traversal process only needs to calculate  $p(\mathbf{z}_t | \mathbf{x}_t)$ , there is no increase in the computational complexity compared with the weight calculation in (8) of typical PF. Next, we generate  $K$  effective particles  $\{\tilde{\mathbf{x}}_t^j\} (j=1, 2, \dots, k)$  around  $\tilde{\mathbf{x}}_t$  to provide a numerical approximation of  $L_t$ . Finally, according to (11), the effective particles  $\{\tilde{\mathbf{x}}_t^j\}$  can be utilized to obtain a near-optimal AGGD proposal  $AGGD(\boldsymbol{\mu}_t, \boldsymbol{\alpha}_t, \boldsymbol{\sigma}_{t,l}^2, \boldsymbol{\sigma}_{t,r}^2)$ , i.e.,

$$\boldsymbol{\mu}_t = \frac{1}{\lambda_t} \sum_{j=1}^K \tilde{\mathbf{x}}_t^j p(\tilde{\mathbf{x}}_t^j | \mathbf{x}_{t-1}) p(\mathbf{z}_t | \tilde{\mathbf{x}}_t^j) \quad (13)$$

$$\boldsymbol{\sigma}_{t,l}^2 = \frac{1}{\lambda_{t,l}} \sum_{\tilde{\mathbf{x}}_t^j < \boldsymbol{\mu}_t} (\tilde{\mathbf{x}}_t^j - \boldsymbol{\mu}_t)(\tilde{\mathbf{x}}_t^j - \boldsymbol{\mu}_t)^\top p(\tilde{\mathbf{x}}_t^j | \mathbf{x}_{t-1}) p(\mathbf{z}_t | \tilde{\mathbf{x}}_t^j) \quad (14)$$

$$\boldsymbol{\sigma}_{t,r}^2 = \frac{1}{\lambda_{t,r}} \sum_{\tilde{\mathbf{x}}_t^j \geq \boldsymbol{\mu}_t} (\tilde{\mathbf{x}}_t^j - \boldsymbol{\mu}_t)(\tilde{\mathbf{x}}_t^j - \boldsymbol{\mu}_t)^\top p(\tilde{\mathbf{x}}_t^j | \mathbf{x}_{t-1}) p(\mathbf{z}_t | \tilde{\mathbf{x}}_t^j) \quad (15)$$

$$\lambda_t = \sum_{j=1}^K p(\tilde{\mathbf{x}}_t^j | \mathbf{x}_{t-1}) p(\mathbf{z}_t | \tilde{\mathbf{x}}_t^j) \quad (16)$$

$$\lambda_{t,l} = \sum_{\tilde{\mathbf{x}}_t^j < \boldsymbol{\mu}_t} p(\tilde{\mathbf{x}}_t^j | \mathbf{x}_{t-1}) p(\mathbf{z}_t | \tilde{\mathbf{x}}_t^j) \quad (17)$$

$$\lambda_{t,r} = \sum_{\tilde{\mathbf{x}}_t^j \geq \boldsymbol{\mu}_t} p(\tilde{\mathbf{x}}_t^j | \mathbf{x}_{t-1}) p(\mathbf{z}_t | \tilde{\mathbf{x}}_t^j) \quad (18)$$

Each shape parameter  $\alpha_t \in \boldsymbol{\alpha}_t$  can be estimated according to the relationship between the Kurtosis ( $k_t > 1.865$ ) and  $\alpha_t$  [25], i.e.,

$$\begin{cases} \alpha_t \approx \sqrt{\frac{5}{k_t - 1.865}} - 0.12 \\ k_t = \frac{E((x - \mu_t)^4)}{E((x - \mu_t)^2)^2} \end{cases} \quad (19)$$

Considering the heavy computation of  $k_t$ , we simply approximate  $\alpha_t$  to the shape parameter  $\boldsymbol{\alpha}$  of  $p(\mathbf{z}_t | \mathbf{x}_t)$ , since the proposal PDF is based on  $L_t = \{\mathbf{x}_t | p(\mathbf{z}_t | \mathbf{x}_t) > \Delta\}$ .

## B. AGGD Inverse CDF Sampling

The inverse CDF sampling is a random sampling method applicable to arbitrary PDF with its inverse CDF known. It is more accurate and convenient than other methods such as the rejection sampling and the Markov chain Monte Carlo sampling [17]. Given the difficulty in expressing the inverse CDF of AGGD, we derive a Gaussian piecewise function from AGGD to facilitate acquisition of the inverse CDF. The AGGD with  $\alpha = 1$ , i.e., asymmetric/symmetric Laplace distribution, is not within the scope of this derivation, as an efficient sampling method for it has been presented in [28].

Consider the AGGD expression under condition of  $x - \mu < 0$  in (1), i.e.,

$$p_l(x) = \frac{\alpha}{(\beta_l + \beta_r) \Gamma\left(\frac{1}{\alpha}\right)} \exp\left(-\left(\frac{-(x - \mu)}{\beta_l}\right)^\alpha\right) = \frac{\alpha}{(\beta_l + \beta_r) \Gamma\left(\frac{1}{\alpha}\right)} \exp(A_l(x)) \quad (20)$$

Around  $x_0 = \varepsilon + \mu$  ( $\varepsilon < 0$ ), we expand  $A_l(x)$  to a second-order Taylor series expansion and ignore its third-order Taylor remainder. The quadratic expansion can substitute  $A_l(x)$  back into (20) to obtain:

$$p_l(x) \approx \frac{\alpha \sqrt{2\pi c_l^2(\varepsilon)} \exp(C_l(\varepsilon))}{(\beta_l + \beta_r) \Gamma\left(\frac{1}{\alpha}\right)} \mathcal{N}(x | b_l(\varepsilon), c_l^2(\varepsilon)) = a_l(\varepsilon) \mathcal{N}(x | b_l(\varepsilon), c_l^2(\varepsilon)) \quad x \in [\varepsilon + \mu - \eta, \varepsilon + \mu + \eta] \quad (21)$$

where  $\mathcal{N}(x | b_l(\varepsilon), c_l^2(\varepsilon))$  represents the Gaussian PDF with mean  $b_l(\varepsilon)$  and variance  $c_l^2(\varepsilon)$ .  $\varepsilon$  determines the Taylor expansion point and  $\eta$  ( $\eta > 0, \varepsilon + \eta < 0$ ) controls the expansion scale.  $C_l(\varepsilon)$  is a formula about  $\varepsilon$ , and its derivation process is the same as that of  $C_r(\varepsilon_r)$  in (A12) in Appendix A. Based on this, the domain of  $x < \mu$  is divided into multiple continuous but non-overlapping intervals, i.e.,  $x \in \bigcup_{i=1}^{\infty} (e_i^l + \mu - \eta_i^l, e_i^l + \mu + \eta_i^l) = \bigcup_{i=1}^{\infty} D(e_i^l, \eta_i^l)$ . In  $x \in D(e_i^l, \eta_i^l)$ ,  $p_l(x)$  is proportional to a Gaussian PDF. Similarly, the approximation of  $p_r(x | \mu, \alpha, \sigma_r^2, \sigma_r^2)$  ( $x \geq \mu$ ) is given in Appendix A, where the subscript  $r$  represents the case of  $x \geq \mu$ .

According to (21), the CDF of AGGD can be approximately expressed as:

$$P(x) = \int_{-\infty}^x p(x) dx \approx \begin{cases} a_l(e_i^l) \Phi(x; b_l(e_i^l), c_l^2(e_i^l)) + \phi(e_i^l, \eta_i^l) & x \in D(e_i^l, \eta_i^l) \\ a_r(e_r^l) \Phi(x; b_r(e_r^l), c_r^2(e_r^l)) + \phi(e_r^l, \eta_r^l) & x \in D(e_r^l, \eta_r^l) \end{cases} \quad (22)$$

where  $\phi(e_i^l, \eta_i^l)$  and  $\phi(e_r^l, \eta_r^l)$  are constants; and  $\Phi(x; b, c^2)$  is the CDF of  $\mathcal{N}(x | b, c^2)$ . We denote the inverse of  $\Phi(x; b, c^2)$  as  $\Phi^{-1}(x | b, c^2)$ , and then the inverse of  $P(x)$  is given by:

$$P^{-1}(y) = \begin{cases} \Phi^{-1}\left(\frac{y - \phi(\varepsilon_l^i, \eta_l^i)}{a_l(\varepsilon_l^i)}; b_l(\varepsilon_l^i), c_l^2(\varepsilon_l^i)\right) & y \in P(D(\varepsilon_l^i, \eta_l^i)) \\ \Phi^{-1}\left(\frac{y - \phi(\varepsilon_r^i, \eta_r^i)}{a_r(\varepsilon_r^i)}; b_r(\varepsilon_r^i), c_r^2(\varepsilon_r^i)\right) & y \in P(D(\varepsilon_r^i, \eta_r^i)) \end{cases} \quad (23)$$

Finally, the inverse CDF sampling for AGGD is as follows.

- 1) Determine appropriate intervals  $\bigcup_{i=1}^{\infty} D(\varepsilon_l^i, \eta_l^i)$  and  $\bigcup_{i=1}^{\infty} D(\varepsilon_r^i, \eta_r^i)$ .
- 2) Randomly sample a point  $u$  from the uniform distribution  $\mathcal{U}(0, 1)$ , i.e.,  $u \sim \mathcal{U}(0, 1)$ .
- 3) Calculate  $x = P^{-1}(u)$  in (23) to get a sample of AGGD.

### C. GPF

With embedding the AGGD proposal PDF and the AGGD sampling method into the typical PF, we propose a GPF method to the non-Gaussian FASE problem. It is assumed that the initial state  $\mathbf{x}_0$ , the priori PDF  $p(\mathbf{x}_i|\mathbf{x}_{i-1})$  and the likelihood PDF  $p(z_i|\mathbf{x}_i)$  are known and independent [20], [29]. At timestep  $t$ ,  $N_p$  candidate particles  $\{\mathbf{x}_{c,t}^i\}$  from  $p(\mathbf{x}_i|\mathbf{x}_{i-1})$  are generated by the inverse CDF sampling. Then, we find the most effective candidate particle  $\tilde{\mathbf{x}}_t$  within  $L_t = \{\mathbf{x}_i | p(z_i|\mathbf{x}_i) > \Delta\}$ . In case of the nonexistence of  $\tilde{\mathbf{x}}_t$ , the subsequent steps are the same as those of the typical PF described in Section III-C. Otherwise, we generate  $K$  effective particles around  $\tilde{\mathbf{x}}_t$  to construct an AGGD proposal PDF  $\mathcal{AGGD}(\boldsymbol{\mu}_t, \boldsymbol{\alpha}_t, \boldsymbol{\sigma}_{t,l}^2, \boldsymbol{\sigma}_{t,r}^2)$ . Next,  $N_p$  particles  $\{\mathbf{x}_t^i\}_{i=1}^{N_p}$  are sampled from  $\mathcal{AGGD}(\boldsymbol{\mu}_t, \boldsymbol{\alpha}_t, \boldsymbol{\sigma}_{t,l}^2, \boldsymbol{\sigma}_{t,r}^2)$  to update the weights based on (8) and (11), i.e.,

$$w_t^i \approx w_{t-1}^i \sum_{j=1}^{N_p} p(z_i|\tilde{\mathbf{x}}_t^j) p(\tilde{\mathbf{x}}_t^j|\mathbf{x}_{t-1}^j) \quad (24)$$

Finally, the optimal estimate of  $\mathbf{x}_t$  is given by:

$$\hat{\mathbf{x}}_t = \sum_{i=1}^{N_p} w_t^i \mathbf{x}_t^i \quad (25)$$

In addition, an adaptive resampling technique [15], which can keep a suitable variety of particles, is applied to the GPF for further decreasing the risk of particle degeneracy. The details of the GPF at timestamp  $t$  are described in Algorithm 2.

## V. CASE STUDY

In this section, numerical simulations are conducted on a modified balanced IEEE 123-bus test system with eight DGs, to demonstrate the accuracy and robustness of the proposed GPF. All experiments are run on a computer with Intel-i5-10400F CPU and 16 GB memory.

### A. Test System and Benchmarks

The balanced IEEE 123-bus network presented in [18], which has seven photovoltaic (PV) units and one wind farm, is selected as the test system.

---

### Algorithm 2: GPF at timestep $t$

---

**Input:** weighted particles  $S_{t-1} = \{(\tilde{\mathbf{x}}_{t-1}^i, w_{t-1}^i)\}_{i=1,2,\dots,N_p}$ , measurements  $z_t$ , priori PDF  $p(\mathbf{x}_i|\mathbf{x}_{i-1})$ , likelihood PDF  $p(z_i|\mathbf{x}_i)$ , and effective particle number  $K$

**Output:**  $S_t = \{(\tilde{\mathbf{x}}_t^i, w_t^i)\}$ ,  $\hat{\mathbf{x}}_t$

- 1: //Prediction step
- 2: **for**  $\tilde{\mathbf{x}}_{t-1}^i \in S_{t-1}$  **do**
- 3:  $\mathbf{x}_{c,t}^i \sim p(\mathbf{x}_i|\mathbf{x}_{i-1})$
- 4: **end for**
- 5: //Estimate the effective likelihood
- 6: Find the most effective candidate particle  $\tilde{\mathbf{x}}_t$  in (12)
- 7: **if**  $\tilde{\mathbf{x}}_t$  is nonexistent **then**
- 8:  $\tilde{\mathbf{x}}_t = \mathbf{x}_{c,t}^i$ ,  $w_t^i = w_{t-1}^i p(z_i|\tilde{\mathbf{x}}_t^i)$ ,  $\hat{\mathbf{x}}_t = \sum_{i=1}^{N_p} w_t^i \tilde{\mathbf{x}}_t^i$
- 9: **else**
- 10: **for**  $j = 1, 2, \dots, K$  **do**
- 11:  $r \sim p(z_i|\mathbf{x}_i)$ ,  $\tilde{\mathbf{x}}_t^j = \tilde{\mathbf{x}}_t + r$
- 12: **end for**
- 13: //Obtain AGGD proposal distribution
- 14: Calculate  $\boldsymbol{\mu}_t$ ,  $\lambda_t$ ,  $\lambda_{t,l}$ , and  $\lambda_{t,r}$  in (13), (16), (17), and (18), respectively
- 15: Calculate  $\boldsymbol{\sigma}_{t,l}^2$  and  $\boldsymbol{\sigma}_{t,r}^2$  in (14) and (15), respectively
- 16:  $\boldsymbol{\alpha}_t$  is set to be the  $\boldsymbol{\alpha}$  of  $p(z_i|\mathbf{x}_i)$
- 17: Obtain  $\mathcal{AGGD}(\boldsymbol{\mu}_t, \boldsymbol{\alpha}_t, \boldsymbol{\sigma}_{t,l}^2, \boldsymbol{\sigma}_{t,r}^2)$
- 18: //Generate new weighted particles
- 19: **for**  $i = 1, 2, \dots, N_p$  **do**
- 20:  $\mathbf{x}_t^i \sim \mathcal{AGGD}(\boldsymbol{\mu}_t, \boldsymbol{\alpha}_t, \boldsymbol{\sigma}_{t,l}^2, \boldsymbol{\sigma}_{t,r}^2)$
- 21: **end for**
- 22:  $\lambda = \sum_{j=1}^{N_p} p(z_i|\tilde{\mathbf{x}}_t^j) p(\tilde{\mathbf{x}}_t^j|\mathbf{x}_{t-1}^j)$
- 23:  $w_t = \lambda w_{t-1}$ ,  $\hat{\mathbf{x}}_t = \sum_{i=1}^{N_p} w_t^i \mathbf{x}_t^i$
- 24: **end if**
- 25: //Adaptive resampling
- 26: **if**  $\frac{1}{\sum (w_t^i)^2} < \frac{1}{2} N_p$  **then**
- 27:  $C(0) = 0$
- 28: **for**  $i = 1, 2, \dots, N_p$  **do**
- 29:  $C(i) = C(i-1) + w_t^i$
- 30: **end for**
- 31:  $\mathbf{x}_t^i = \tilde{\mathbf{x}}_t$
- 32: **for**  $j = 1, 2, \dots, N_p$  **do**
- 33:  $u \sim \mathcal{U}(0, 1)$
- 34: **if**  $C(m) > u$  and  $C(m+1) \leq u$  **then**
- 35:  $\mathbf{x}_t^j = \mathbf{x}_t^m$ ,  $w_t^j = \frac{1}{N_p}$
- 36: **end if**
- 37: **end for**
- 38:  $\tilde{\mathbf{x}}_t = \mathbf{x}_t^i$ ,  $w_t^i = \frac{1}{N_p}$ ,  $\hat{\mathbf{x}}_t = \sum_{i=1}^{N_p} w_t^i \tilde{\mathbf{x}}_t^i$
- 39: **end if**

---

The details of eight DGs in the modified balanced IEEE 123-bus test system for simulating power outputs of PVs and wind farms are illustrated in Table II, where  $Weibull(k, s)$  is the Weibull PDF with shape parameter  $k$  and scale parameter  $s$  [30], and  $Beta(u, v)$  is the Beta PDF with two parameters  $u$  and  $v$  [31]. Node 1 is selected as the slack bus with 1.05 p.u. voltage magnitude, and node 98 associated with the wind farm is regarded as the PV-type bus with 1.0 p.u. voltage magnitude. The bus loads are randomly changed within 80% to 120% of the base loads given in [18]. Additional network parameters are available in [18].

Based on the above settings, we simulate dynamic power flows of the test system over 100 timesteps. The voltage magnitudes and phase angles in the outcomes are considered as the true states. For comparison purposes, seven bench-

marks are chosen, including one traditional state estimation method, i.e., weighted least squares (WLS) [19], three popular FASE methods, i.e., EKF, UKF, and CKF, and three FASE methods applicable to non-Gaussian noises, i.e., UKF with generalized correntropy loss (GCL-UKF) [10], PF, and unscented PF (UPF) [20]. The root-mean-square error (RMSE), as a commonly used metric to evaluate the accuracy of FASE [14], [24], is applied in the experiments. Based on the average result from  $M=50$  Monte Carlo simulations, the RMSE within  $T=100$  timesteps can be calculated as:

$$RMSE = \sqrt{\frac{1}{M} \frac{1}{T} \sum_m \sum_t (\hat{x}_t^m - x_t)^2} \quad (26)$$

The RMSEs of voltage magnitudes and phase angles are denoted as  $RMSE_V$  and  $RMSE_\theta$ , respectively.

TABLE II  
DETAILS OF EIGHT DGs IN MODIFIED IEEE 123-BUS TEST SYSTEM

Node	DG type	Capacity (MVA)	Power output distribution
15	PV	0.2	<i>Beta</i> (4.2, 1.8)
25	PV	0.4	<i>Beta</i> (4.2, 1.8)
35	PV	0.6	<i>Beta</i> (4.0, 1.6)
68	PV	0.6	<i>Beta</i> (4.0, 1.6)
88	PV	0.4	<i>Beta</i> (4.2, 1.8)
98	Wind farm	1.0	<i>Weibull</i> (2, 10) (wind speed)
105	PV	0.4	<i>Beta</i> (4.0, 2.0)
114	PV	0.2	<i>Beta</i> (4.0, 2.0)

### B. Measurements

It is assumed that 46 PMUs are set in the test system for the acquisition of voltage phasors. Their locations are referred to the result in [32]. Also, voltage magnitudes of the remaining nodes, injection power of randomly selected 80 nodes, and power flows of randomly selected 60 branches can be obtained by SCADA. To simplify experimental settings, the imperfect synchronization of PMU and SCADA is neglected, as it has little impact on testifying the performances of state estimation algorithms in dealing with non-Gaussian noises [33], [34]. Gaussian noises, non-Gaussian noises, and some outliers are added into the simulation outcomes associated with the measurement devices, respectively, to create three cases of measurements. The details are described below.

1) Case 1: PMU measurements are contaminated by the additive Gaussian noise with the variance  $\mathbf{Q}_p = 2.5 \times 10^{-5} \mathbf{I}$ , i.e.,  $\mathcal{N}(0, \mathbf{Q}_p)$ , while the noise PDF of SCADA measurements is  $\mathcal{N}(0, \mathbf{Q}_s)$  with the variance  $\mathbf{Q}_s = 10^{-4} \mathbf{I}$ .

2) Case 2: noise PDFs of PMU and SCADA measurements are  $AGGD(0, 1.3, \mathbf{Q}_p, \mathbf{Q}_p)$  and  $AGGD(0, 1.6, \mathbf{Q}_s, \mathbf{Q}_s)$ , respectively. Since  $AGGD(0, 1.3, \mathbf{Q}_p, \mathbf{Q}_p)$  and  $AGGD(0, 1.6, \mathbf{Q}_s, \mathbf{Q}_s)$  are more concentrated near zero than  $\mathcal{N}(0, \mathbf{Q}_p)$  and  $\mathcal{N}(0, \mathbf{Q}_s)$ , respectively, the measurements in Case 2 are more precise than those in Case 1.

3) Case 3: the measurements of Case 2 are taken as the original measurements, and five nodes, i.e., nodes 30, 33, 55, 91, and 113, are randomly selected to generate outliers

in their associated measurements. It is assumed that the errors between the outliers and true values are 5%-10%. Each original measurement associated with the five nodes has a 20% probability of being replaced with an outlier. Thus, the measurements in Case 3 have bad data.

### C. Case 1

In this case, each estimate of WLS is obtained after 5 iterations. For particle-based filters, the particle number is usually determined via trial and error [35]. We conduct the repeated experiments with different particle numbers, and then select a suitable particle number with trading off runtime and accuracy. Finally, PF and UPF generate 600 particles for estimation. The numbers of candidate particles and effective particles in GPF are set to be 600 and 200, respectively. The process noises, RMSEs, and runtime of state estimation methods under Gaussian measurement noise are provided in Table III.  $\mathbf{Q} = (\mathbf{Q}_V \ \mathbf{Q}_\theta)$ , where  $\mathbf{Q}_V = 10^{-4} \mathbf{I}_{n \times n}$  and  $\mathbf{Q}_\theta = 2.5 \times 10^{-5} \mathbf{I}_{n \times n}$  are the noise covariance matrices of voltage magnitudes and phase angles, respectively.

TABLE III  
PROCESS NOISES, RMSEs, AND RUNTIMES OF STATE ESTIMATION METHODS UNDER GAUSSIAN MEASUREMENT NOISE

Method	Process noise	$RMSE_V$ ( $10^{-3}$ p.u.)	$RMSE_\theta$ ( $10^{-3}$ rad)	Runtime (s)
WLS		3.36	4.41	0.72
EKF	$\mathcal{N}(0, \mathbf{Q})$	3.17	4.30	0.97
CKF	$\mathcal{N}(0, \mathbf{Q})$	2.30	3.48	1.14
UKF	$\mathcal{N}(0, \mathbf{Q})$	2.25	3.33	1.03
GCL-UKF	$\mathcal{N}(0, \mathbf{Q})$	2.17	3.26	10.61
PF	$\mathcal{N}(0, \mathbf{Q})$	2.08	3.20	2.79
PF	$AGGD(0, 3, \mathbf{Q}, \mathbf{Q})$	2.01	3.16	2.80
UPF	$\mathcal{N}(0, \mathbf{Q})$	1.98	3.14	3.63
UPF	$AGGD(0, 3, \mathbf{Q}, \mathbf{Q})$	1.94	3.12	3.63
GPF	$\mathcal{N}(0, \mathbf{Q})$	1.70	3.09	3.26
GPF	$AGGD(0, 3, \mathbf{Q}, \mathbf{Q})$	1.64	3.06	3.26

#### 1) Efficiency of AGGD Inverse CDF Sampling

To demonstrate affect of the AGGD sampling method on real-time of particle-based filters, i.e., GPF, UPF, and PF, we compare the runtime of these methods under Gaussian and AGGD process noises. As shown in Table III, these particle-based filters consume almost the same time under conditions of the two process noise assumptions. This indicates that the AGGD inverse CDF sampling method has high efficiency to avoid extra computational burdens in particle generation.

#### 2) Accuracy Versus Runtime

The RMSEs and runtime of the benchmarks and the proposed GPF are shown in Table III. The traditional WLS has the best real-time performance but the worst precision, since the priori knowledge of system process trends is not considered in it. For the seven FASE methods, the particle-based filters outperform the EKF, UKF, and CKF in accuracy, whereas the reverse is true in runtime. In other words, they generate numerous weighted particles to achieve more precise approximation to the posterior PDF at the expense of real time. Furthermore, the particle-based filters outperform the GCL-UKF in both accuracy and runtime. The GCL-UKF

is developed to filter out non-Gaussian noises through the time-consuming generalized correntropy loss (GCL) criterion. In Case 1, however, the GCL criterion may contribute little to the accuracy because of the Gaussian measurement errors. As a result, the accuracy of GCL-UKF is only 2.80% better on average than that of UKF, while the runtime of it increases to 10 times that of UKF.

Within the three particle-based filters, the PF has the worst estimation accuracy but the shortest runtime, as its proposal PDF is the readily accessible process noise PDF that only contains priori knowledge. The  $RMSE_V$  and  $RMSE_\theta$  of the proposed GPF are 15.5% and 2.0% smaller than those of the UPF, respectively, and the runtime of it is 10.2% less than that of the UPF. This is inferred that the GPF can construct a proposal PDF closer to the target PDF in a shorter time, to efficiently achieve accurate estimation. However, the runtime of particle-based filters cannot satisfy the real-time estimation requirement in power systems, i.e., provide an optimal state estimate within 2 s [36]. To cope with it, the particle-based filters can be implemented in a distributed manner to significantly reduce the runtime [27].

#### D. Case 2

The measurement noise distributions in Case 2 are  $AGGD(0, 1.3, \mathbf{Q}_p, \mathbf{Q}_p)$  and  $AGGD(0, 1.6, \mathbf{Q}_s, \mathbf{Q}_s)$ . Since the WLS and the Kalman-based filters are unavailable to non-Gaussian FASE model, their measurement noise distributions are approximated as  $\mathcal{N}(0, \mathbf{Q}_p)$  and  $\mathcal{N}(0, \mathbf{Q}_s)$ , respectively. Other parameters are the same as those in Case 1.

##### 1) Accuracy Comparison

The RMSEs of state estimation methods under non-Gaussian measurement noise are provided in Table IV. Compared with the overall estimation accuracy in Case 1, those in Case 2 are improved due to the smaller measurement errors. Specifically, the  $RMSE_V$  and  $RMSE_\theta$  of the four FASE methods applicable to non-Gaussian noise environments, i.e., GCL-UKF, PF, UPF, and GPF, are reduced by 3.7% and 2.5%, 3.0% and 1.9%, 5.7% and 2.9%, 8.4% and 3.2%, respectively, while those of WLS, EKF, CKF, and UKF are 1.2% and 0.9%, 1.3% and 1.6%, 1.7% and 1.4%, 1.3% and 1.0% less, respectively. This indicates that the accuracy of FASE methods can be improved (through whether robust optimality criteria or appropriate distribution assumptions). Nevertheless, the appropriate distribution assumptions in the particle-based filters lead to more accurate estimates than the GCL criterion in the GCL-UKF. In addition, the RMSEs of GPF are the smallest and decreased the most, which validates that the proposed GPF can obtain high-accurate estimates against non-Gaussian noises.

##### 2) Effective Particle Number of GPF

The accuracy and runtime of GPF with different effective particle numbers are obtained, as shown in Table V, under conditions of  $AGGD(0, 3, \mathbf{Q}, \mathbf{Q})$  process noise and 300 candidate particles. The RMSEs decrease gradually with more effective particles, while the runtime increases by about 0.15 s for every 50 additional effective particles. Although the large number of effective particles sacrifices a little real-time performance of GPF, it narrows the gap between the proposal PDF and the posterior PDF, as shown in Fig. 3. In addition,

the estimates of GPF with 200 candidate particles and 400 effective particles have better accuracy but shorter runtime than those with 600 candidate particles and 200 effective particles (in Table IV). Thus, sufficient effective particles can significantly enhance the accuracy of GPF.

TABLE IV  
RMSEs OF STATE ESTIMATION METHODS UNDER NON-GAUSSIAN MEASUREMENT NOISE

Method	Process noise	$RMSE_V (10^{-3} \text{ p.u.})$	$RMSE_\theta (10^{-3} \text{ rad})$
WLS		3.32	4.37
EKF	$\mathcal{N}(0, \mathbf{Q})$	3.13	4.23
CKF	$\mathcal{N}(0, \mathbf{Q})$	2.26	3.41
UKF	$\mathcal{N}(0, \mathbf{Q})$	2.22	3.30
GCL-UKF	$\mathcal{N}(0, \mathbf{Q})$	2.09	3.18
PF	$\mathcal{N}(0, \mathbf{Q})$	2.00	3.12
PF	$AGGD(0, 3, \mathbf{Q}, \mathbf{Q})$	1.95	3.10
UPF	$\mathcal{N}(0, \mathbf{Q})$	1.87	3.05
UPF	$AGGD(0, 3, \mathbf{Q}, \mathbf{Q})$	1.83	3.03
GPF	$\mathcal{N}(0, \mathbf{Q})$	1.47	3.00
GPF	$AGGD(0, 3, \mathbf{Q}, \mathbf{Q})$	1.42	2.93

TABLE V  
RMSEs VERSUS RUNTIMES OF GPF WITH DIFFERENT EFFECTIVE PARTICLE NUMBERS AND 300 CANDIDATE PARTICLES

$K$	$RMSE_V (10^{-3} \text{ p.u.})$	$RMSE_\theta (10^{-3} \text{ rad})$	Runtime (s)
50	1.71	3.11	1.38
100	1.64	3.07	1.61
150	1.56	3.03	1.74
200	1.51	2.98	1.96
250	1.47	2.96	2.04
300	1.45	2.95	2.23
350	1.42	2.93	2.30
400	1.39	2.92	2.45

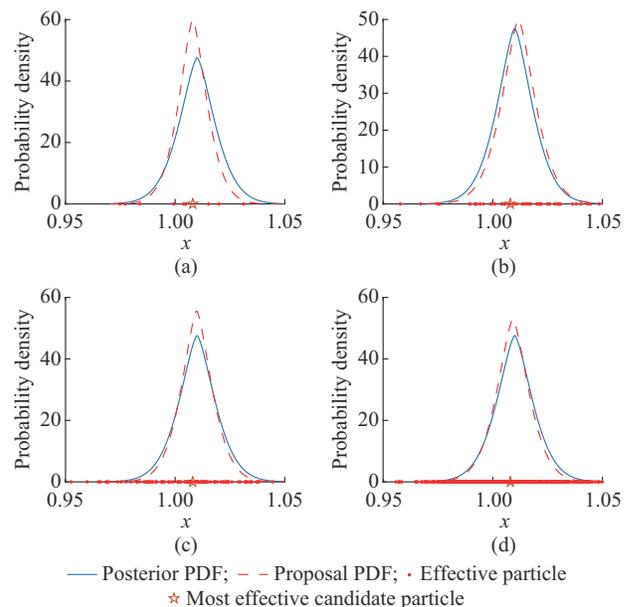


Fig. 3. Effect of effective particle number. (a)  $K=10$ . (b)  $K=50$ . (c)  $K=100$ . (d)  $K=500$ .

### E. Case 3

Using the abnormal measurements, we obtain the RMSEs of the benchmarks and the proposed GPF, where the noise distribution assumptions and other parameters of each method are the same as those in Case 2. Compared with the estimation errors under normal non-Gaussian measurements, the  $RMSE_V$  values of WLS, EKF, CKF, UKF, GCL-UKF, PF, UPF, and GPF, as shown in Table VI, increase by 17.16%, 14.70%, 10.12%, 9.90%, 7.18%, 8.20%, and 4.20%, respectively, and the  $RMSE_\theta$  values increase by 18.10%, 19.86%, 16.13%, 14.85%, 11.63%, 12.90%, 10.89%, and 7.17%, respectively. Among these methods, the GPF has the smallest increase of estimation errors. Figures 4 and 5 also demonstrate the robustness of the proposed GPF in cases of SCADA and PMU measurement anomalies. We infer that the priori knowledge in the AGGD proposal PDF can resist the severe disturbances in measurements.

TABLE VI  
RMSES OF STATE ESTIMATION METHODS UNDER ABNORMAL MEASUREMENTS

Method	Process noise	$RMSE_V (10^{-3} \text{ p.u.})$	$RMSE_\theta (10^{-3} \text{ rad})$
WLS		3.89	5.16
EKF	$\mathcal{N}(0, \mathbf{Q})$	3.59	5.07
CKF	$\mathcal{N}(0, \mathbf{Q})$	2.52	3.96
UKF	$\mathcal{N}(0, \mathbf{Q})$	2.44	3.79
GCL-UKF	$\mathcal{N}(0, \mathbf{Q})$	2.24	3.55
PF	$AGGD(0.3, \mathbf{Q}, \mathbf{Q})$	2.17	3.51
UPF	$AGGD(0.3, \mathbf{Q}, \mathbf{Q})$	1.98	3.36
GPF	$AGGD(0.3, \mathbf{Q}, \mathbf{Q})$	1.48	3.14

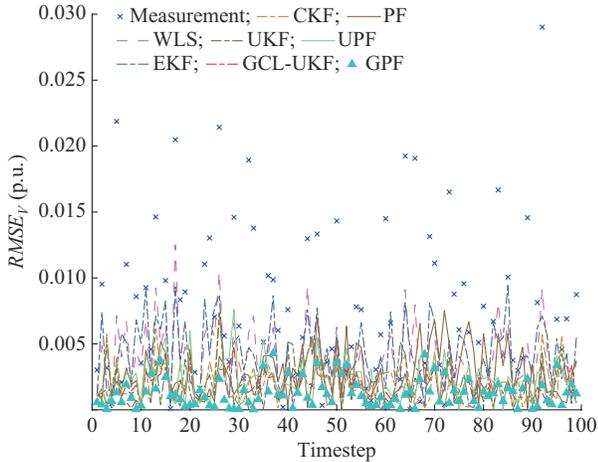


Fig. 4. Voltage magnitude errors of abnormal SCADA measurements and estimates at node 30.

## VI. CONCLUSION

This paper proposes to use a generic distribution assumption of noise, i.e., AGGD, to formulate the non-Gaussian FASE problem. To solve this problem, a novel GPF is presented, which improves the typical PF with a near-optimal AGGD proposal PDF and the corresponding rapid sampling method. Experiments are implemented on a balanced IEEE

123-bus test system with DGs. The results demonstrate that the proposed GPF is the most accurate and robust compared with the seven benchmarks, i.e., WLS, EKF, CKF, UKF, GCL-UKF, PF, and UPF, under conditions of the measurements with Gaussian noise, non-Gaussian noise, and outliers. Nevertheless, its computation is intense in the cases of large-scale power systems, and accuracy becomes a little worse under abnormal measurements. In future, we will study the distributed implementation and the bad data detection of the GPF to improve its real-time performance and robustness. Another focus of our future studies is the parameter estimation of AGGD to accurately characterize real non-Gaussian noises based on historical data.

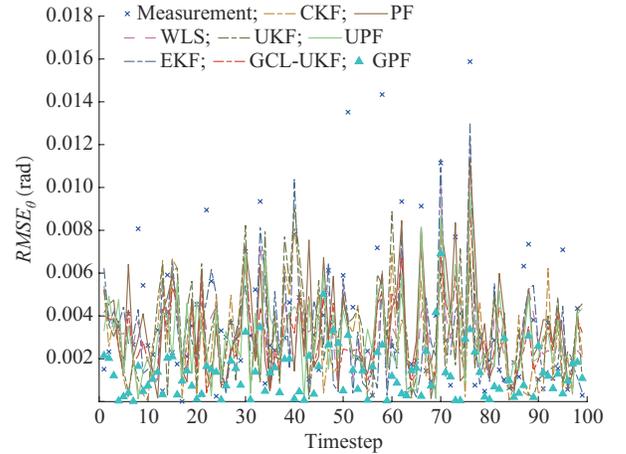


Fig. 5. Voltage phase angle errors of abnormal PMU measurements and estimates at node 55.

## APPENDIX A

Appendix A will present the detailed derivation of the Gaussian piecewise approximation of AGGD.

Around  $x_0 = \varepsilon + \mu$  ( $\varepsilon < 0$ ),  $A_l(x)$  in (20) is expanded to a second-order Taylor series expansion, i.e.,

$$A_l(x) = A_l(\varepsilon + \mu) + A_l'(\varepsilon + \mu)(x - \varepsilon - \mu) + \frac{A_l''(\varepsilon + \mu)}{2}(x - \varepsilon - \mu)^2 + R((x - \varepsilon - \mu)^3) \quad x \in [\varepsilon + \mu - \eta, \varepsilon + \mu + \eta] \quad (A1)$$

where  $\varepsilon$  determines the Taylor expansion point;  $\eta$  ( $\eta > 0, \varepsilon + \eta < 0$ ) controls the expansion scale;  $A_l'(\cdot)$  and  $A_l''(\cdot)$  are the first derivative and the second derivative of  $A_l(\cdot)$ , respectively;  $R((x - \varepsilon - \mu)^3)$  is the third-order Taylor remainder, which is equal to zero in the case of  $\alpha = 2$ . We ignore the remainder and merge the similar terms in (A1) to obtain:

$$A_l(x) \approx B_l(\varepsilon)x^2 + L_l(\varepsilon)x + C_l(\varepsilon) \quad (A2)$$

$$B_l(\varepsilon) = -\frac{1}{2} \alpha(\alpha - 1)(-\varepsilon)^{\alpha-2} \beta_l^{-\alpha} \quad (A3)$$

$$L_l(\varepsilon) = \alpha(\alpha\mu + \alpha\varepsilon - \mu - 2\varepsilon)(-\varepsilon)^{\alpha-2} \beta_l^{-\alpha} \quad (A4)$$

$$C_l(\varepsilon) = \left[ -\frac{\varepsilon^2(\alpha - 1)(\alpha - 2)}{2} + \frac{\mu^2\alpha(1 - \alpha)}{2} + \mu\alpha\varepsilon(2 - \alpha) \right] (-\varepsilon)^{\alpha-2} \beta_l^{-\alpha} \quad (A5)$$

The formula in (A2) can be rewritten as:

$$A_l(x) = -\frac{(x - b_l(\varepsilon))^2}{2c_l(\varepsilon)^2} + C_{l'}(\varepsilon) \quad (A6)$$

$$\begin{cases} b_l(\varepsilon) = -\frac{L_l(\varepsilon)}{2B_l(\varepsilon)} \\ c_l^2(\varepsilon) = -\frac{1}{2B_l(\varepsilon)} \\ C_{l'}(\varepsilon) = C_l(\varepsilon) - \frac{L_l^2(\varepsilon)}{4B_l(\varepsilon)} \end{cases} \quad (A7)$$

Substituting (A6) back into (20), we can obtain:

$$p_l(x) \approx \frac{\alpha \sqrt{2\pi c_l^2(\varepsilon)} \exp(C_{l'}(\varepsilon))}{(\beta_l + \beta_r) \Gamma\left(\frac{1}{\alpha}\right)} \mathcal{N}(x|b_l(\varepsilon), c_l^2(\varepsilon)) = a_l(\varepsilon) \mathcal{N}(x|b_l(\varepsilon), c_l^2(\varepsilon)) \quad (A8)$$

Accordingly,  $p_l(x)$  in  $x \in D(\varepsilon_l^i, \eta_l^i)$  is proportional to a Gaussian PDF.

The derivation of approximating AGGD under  $x \geq \mu$  is the same as the above. Consequently, the AGGD under condition of  $x \geq \mu$  can be expressed as:

$$p_r(x) \approx \frac{\alpha \sqrt{2\pi c_r^2(\varepsilon_r^i)} \exp(C_{r'}(\varepsilon_r^i))}{(\beta_l + \beta_r) \Gamma\left(\frac{1}{\alpha}\right)} \mathcal{N}(x|b_r(\varepsilon_r^i), c_r^2(\varepsilon_r^i)) = a_r(\varepsilon_r^i) \mathcal{N}(x|b_r(\varepsilon_r^i), c_r^2(\varepsilon_r^i)) \quad (A9)$$

$$x \in \bigcup_{i=1}^{\infty} [\varepsilon_r^i + \mu - \eta_r^i, \varepsilon_r^i + \mu + \eta_r^i] = \bigcup_{i=1}^{\infty} D(\varepsilon_r^i, \eta_r^i) \quad (A10)$$

$$\begin{cases} b_r(\varepsilon_r^i) = -\frac{L_r(\varepsilon_r^i)}{2B_r(\varepsilon_r^i)} \\ c_r^2(\varepsilon_r^i) = -\frac{1}{2B_r(\varepsilon_r^i)} \end{cases} \quad (A11)$$

$$C_{r'}(\varepsilon_r^i) = C_r(\varepsilon_r^i) - \frac{L_r^2(\varepsilon_r^i)}{4B_r(\varepsilon_r^i)} \quad (A12)$$

$$B_r(\varepsilon_r^i) = -\frac{1}{2} \alpha (\alpha - 1) (\varepsilon_r^i)^{\alpha-2} \beta_r^{-\alpha} \quad (A13)$$

$$L_r(\varepsilon_r^i) = \alpha (\alpha \mu + \alpha \varepsilon_r^i - \mu - 2\varepsilon_r^i) (\varepsilon_r^i)^{\alpha-2} \beta_r^{-\alpha} \quad (A14)$$

$$C_r(\varepsilon_r^i) = \left[ -\frac{1}{2} (\varepsilon_r^i)^2 (\alpha - 1)(\alpha - 2) + \frac{1}{2} \mu^2 \alpha (1 - \alpha) + \mu \alpha \varepsilon_r^i (2 - \alpha) \right] (\varepsilon_r^i)^{\alpha-2} \beta_r^{-\alpha} \quad (A15)$$

where  $\eta_r^i > 0$ ; and  $\varepsilon_r^i - \eta_r^i \geq 0$ .

## REFERENCES

- [1] M. B. D. C. Filho and J. C. S. D. Souza, "Forecasting-aided state estimation - Part I: panorama," *IEEE Transactions on Power Systems*, vol. 24, no. 4, pp. 1667-1677, Sept. 2009.
- [2] A. Primadianto and C. Lu, "A review on distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3875-3883, Dec. 2017.
- [3] J. Zhao, A. Gomez-Exposito, M. Netto *et al.*, "Power system dynamic state estimation: motivations, definitions, methodologies, and future work," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3188-3198, Jul. 2019.
- [4] S. Wang, J. Zhao, Z. Huang *et al.*, "Assessing Gaussian assumption of pmu measurement error using field data," *IEEE Transactions on Power Delivery*, vol. 33, no. 6, pp. 3233-3236, Dec. 2019.
- [5] T. Ahmad and N. Senroy, "Statistical characterization of PMU error for robust WAMS based analytics," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 920-928, Mar. 2019.
- [6] A. de Santis, A. Germani, and M. Raimondi, "Optimal quadratic filtering of linear discrete-time non-Gaussian systems," *IEEE Transactions on Automatic Control*, vol. 40, no. 7, pp. 1274-1278, Jul. 1995.
- [7] L. Dang, B. Chen, S. Wang *et al.*, "Robust power system state estimation with minimum error entropy unscented Kalman filter," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 11, pp. 8797-8808, Jun. 2019.
- [8] J. A. D. Massignan, J. B. A. London, and V. Miranda, "Tracking power system state evolution with maximum-correntropy-based extended Kalman filter," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 4, pp. 616-626, Jul. 2020.
- [9] T. Ahmad and N. Senroy, "An information theoretic approach to power-substation level dynamic state estimation with non-Gaussian noise," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1642-1645, Jan. 2020.
- [10] W. Ma, J. Qiu, X. Liu *et al.*, "Unscented Kalman filter with generalized correntropy loss for robust power system forecasting-aided state estimation," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 11, pp. 1642-1645, May 2019.
- [11] T. Chen, Y. Cao, X. Chen *et al.*, "A distributed maximum-likelihood-based state estimation approach for power systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-10, Sept. 2020.
- [12] Z. Jin, J. Zhao, S. Chakrabarti *et al.*, "A hybrid robust forecasting-aided state estimator considering bimodal Gaussian mixture measurement errors," *International Journal of Electrical Power & Energy Systems*, vol. 120, pp. 105962, Sept. 2020.
- [13] D. Xu, C. Shen, and F. Shen, "A robust particle filtering algorithm with non-gaussian measurement noise using student-t distribution," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 30-34, Jan. 2014.
- [14] M. S. Arulampalam, S. Maskell, N. Gordon *et al.*, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174-188, Feb. 2002.
- [15] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with Rao-Blackwellized particle filters," *IEEE Transactions on Robotics*, vol. 23, no. 1, pp. 34-46, Feb. 2007.
- [16] A. Tesi and C. S. Regazzoni, "The asymmetric generalized Gaussian function: a new HOS-based model for generic noise PDFs," in *Proceedings of 8th Workshop on Statistical Signal and Array Processing*, Corfu, Greece, Jun. 1996, pp. 210-213.
- [17] U. Andersson and S. Godsill, "Optimum kernel particle filter for asymmetric Laplace noise in multivariate models," in *Proceedings of 2020 IEEE 23rd International Conference on Information Fusion*, Rustenburg, South Africa, Jul. 2020, pp. 1-4.
- [18] Y. Chai, L. Guo, C. Wang *et al.*, "Network partition and voltage coordination control for distribution networks with high penetration of distributed PV units," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3396-3407, May 2018.
- [19] P. Risbud, N. Gatsis, and A. Taha, "Multi-period power system state estimation with PMUs under GPS spoofing attacks," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 4, pp. 597-606, Jul. 2020.
- [20] S. S. Yu, J. Guo, T. K. Chau *et al.*, "An unscented particle filtering approach to decentralized dynamic state estimation for DFIG wind turbines in multi-area power systems," *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 2670-2682, Jul. 2020.
- [21] J. Zhang, G. Welch, G. Bishop *et al.*, "A two-stage Kalman filter approach for robust and real-time power system state estimation," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 2, pp. 2670-2682, Apr. 2014.
- [22] X. Zhang, "A novel cubature Kalman filter for nonlinear state estimation," in *Proceedings of 52nd IEEE Conference on Decision and Control*, Florence, Italy, Dec. 2013, pp. 7797-7802.
- [23] A. Zia, T. Kirubarajan, J. P. Reilly *et al.*, "An EM algorithm for nonlinear state estimation with model uncertainties," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 921-936, Mar. 2008.
- [24] Z. Xiao, D. Xiao, V. Havyarimana *et al.*, "Toward accurate vehicle state estimation under non-Gaussian noises," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 921-936, Dec. 2019.
- [25] M. O. M. Mahmoud, M. Jaidane-Saidane, J. Souissi *et al.*, "Modeling

- of the load duration curve using the asymmetric generalized Gaussian distribution: case of the Tunisian power system," in *Proceedings of the 10th International Conference on Probabilistic Methods Applied to Power Systems*, Rincon, USA, Aug. 2008, pp. 1-7.
- [26] J. Yang, W. Zhang, and F. Guo, "Dynamic state estimation for power networks by distributed unscented information filter," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2162-2171, May 2020.
- [27] O. Hlinka, F. Hlawatsch, and P. M. Djuric, "Consensus-based distributed particle filtering with distributed proposal adaptation," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3029-3041, Jun. 2014.
- [28] S. Kotz, T. Kozubowski, and K. Podgorski, "The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance," in *Springer Science and Business Media*, Boston: Birkhäuser, 2001, pp. 22-29.
- [29] S. Yin and X. Zhu, "Intelligent particle filter and its application to fault detection of nonlinear system," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3852-3861, Jun. 2015.
- [30] L. Shi, C. Wang, L. Yao *et al.*, "Optimal power flow solution incorporating wind power," *IEEE Systems Journal*, vol. 6, no. 2, pp. 233-241, Jun. 2012.
- [31] A. Baharvandi, J. Aghaei, A. Nikoobakht *et al.*, "Linearized hybrid stochastic/robust scheduling of active distribution networks encompassing PVs," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 357-367, Jan. 2020.
- [32] K. Chauhan and R. Sodhi, "Placement of distribution-level phasor measurements for topological observability and monitoring of active distribution networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3451-3460, Jun. 2020.
- [33] P. A. Pegoraro, A. Angioni, M. Andrea *et al.*, "Bayesian approach for distribution system state estimation with non-Gaussian uncertainty models," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 11, pp. 2957-2966, Aug. 2017.
- [34] N. Kashyap, S. Werner, and Y. Huang, "Decentralized PMU-assisted power system state estimation with reduced interarea communication," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 607-616, Aug. 2018.
- [35] J. H. Kotecha and P. M. Djuric, "Gaussian sum particle filtering," *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2602-2612, Oct. 2003.
- [36] A. Monticelli, "Electric power system state estimation," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 262-282, Feb. 2000.

**Lyuzerui Yuan** received the B.E. degree in electrical engineering and automation from South China University of Technology, Guangzhou, China, in 2017. She is currently pursuing the Ph.D. degree in electrical engineering at Shanghai Jiao Tong University, Shanghai, China. Her current research interests include situational awareness, edge computing, and deep learning, with applications in power systems.

**Jie Gu** received the bachelor's, master's, and Ph.D. degrees in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1992, 1995, and 2000, respectively. She then worked with Shanghai Jiao Tong University, where she is currently an Associate Professor with the School of Electronic Information and Electrical Engineering. Her main research interests include power system planning, power markets, smart grids, microgrids, and intelligent algorithms.

**Honglin Wen** received the bachelor's degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2017, where he is currently working toward the Ph.D. degree with the Research Center for Big Data Engineering Technology. His research interests include machine learning applications in power systems, time series forecasting, uncertainty quantification, and decision-making under uncertainty.

**Zhijian Jin** received the Ph.D. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China. Since 1990, he has been with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, where he currently serves as a Tenured Full Professor and the Deputy Director of the National Smart Grid Research and Development Center. His research interests include smart grids and applied superconductor.