This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/TII.2024.3485791, IEEE Transactions on Industrial Informatics

Adaptive Decentralized State Estimation for Multi-Machine Power Grids Under Measurement Noises with Unknown Statistics

Bogang Qu, Zidong Wang, Bo Shen, Hongli Dong, and Daogang Peng

***Abstract*—This paper is concerned with the adaptive dynamic state estimation (DSE) problem for synchronous-generator-based multi-machine power grids under measurement noise with unknown statistics. The statistical properties of the measurement noises are efficiently revealed by utilizing limited measurement data contained in a sliding window, and such data is employed to establish the base distribution of the noises, with the aid of the Gaussian mixture model and the Kernel density estimation scheme. Subsequently, the component number of the base distribution of the measurement noises is reduced by designing a fuzzy c-means clustering algorithm with the Wasserstein distance criterion. An improved sliding-window-based adaptive cubature Kalman filtering scheme is then proposed, which leverages the already obtained statistical characteristics of the measurement noise and the concept of the Gaussian summation filter. Finally, the validity of the proposed adaptive DSE algorithm under various measurement noise statistics is illustrated by simulation studies conducted on the IEEE 39-bus system featuring three test scenarios.**

***Index Terms*—Adaptive state estimation, multi-machine power grids, unknown measurement noises, cubature Kalman filter, clustering algorithm.**

I. Introduction

Over the past few decades, the power grids have undergone significant evolution due to the high penetration of renewable power generations and widespread use of power electronics [12], [29], [48]. These developments frequently force the power grids to operate under extreme conditions, thereby increasing the demands for real-time control, reliable decision-making, and secure assessment of the grids [6]. It is worth noting that traditional control/optimization schemes may encounter numerous challenges in current power grid applications due primarily to their inherent limitations in addressing the inner states of the systems. In response, substantial research efforts have been dedicated to developing novel information perception techniques, such as situational awareness and state estimation (SE), see e.g. [6], [12].

Recently, there has been a significant surge in research interest focused on the SE problem in areas of target tracking [19], [37], fault diagnosis [49], and artificial intelligence [10], [20], [23]. Within the realm of power grids, SE techniques have garnered considerable research enthusiasm as highlighted by some notable works [12], [30], [42], [45]. For instance, in [42], a novel fully distributed unscented information filter has been designed for large-scale power networks. It is critical to acknowledge, however, that most existing SE methodologies for power systems are based on the fundamental assumption that measurement noises follow a *Gaussian* distribution with *known* statistics, but such an assumption may not hold due to the complex and variable operating status of power grids.

It has been reported in [2], [40] that the noise in phasor measurement units (PMUs) exhibits *non-Gaussian* characteristics (e.g. thick-tailed/multimodal distribution). These observations have highlighted the importance of studying the dynamic state estimation (DSE) problem for power systems under the influence of non-Gaussian noise. In response to this need, several key studies have addressed DSE issues in power systems with non-Gaussian noises. Representative works in this area include [1], [22]. For example, in [22], an event-based DSE algorithm has been developed for synchronous generators under the assumption of non-Gaussian measurement noises. Mostly recently, a robust particle filtering algorithm has been introduced in [1] to cater to non-Gaussian observation noises and various outlier types.

In many practical scenarios, the statistics of measurement noises are unknown and even time-varying [7], and this variability can pose challenges to the effective implementation of existing DSE methods. Specifically, most existing DSE methods for power systems are predicated on predetermined models such as system, measurement, and noise models, without fully considering the actual statistics of the measurement noises. As a result, there has been a rapidly growing interest in robust DSE schemes that operate effectively with unknown noise statistics [1], [8], [46]. Nevertheless, it is crucial to recognize that these robust DSE schemes are fundamentally passive in addressing non-Gaussian noises with unknown statistics, and their results may not be optimal. Therefore, a natural

This work was supported in part by the National Natural Science Foundation of China under Grants U21A2019, 61933007, 62273088, 62373241 and 62303301, the Shanghai Pujiang Program of China under Grant 23PJD040, the Royal Society of the UK, and the Alexander von Humboldt Foundation of Germany. *(Corresponding author: Daogang Peng)*

Bogang Qu and Daogang Peng are with the College of Automation Engineering, Shanghai University of Electric Power, Shanghai, 200090, China. (E-mail: bogangqu@shiep.edu.cn, pengdaogang@shiep.edu.cn)

Zidong Wang is with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom. (Email: Zidong.Wang@brunel.ac.uk)

Bo Shen is with the College of Information Science and Technology, Donghua University, Shanghai 200051, China, and also with the Engineering Research Center of Digitalized Textile and Fashion Technology, Ministry of Education, Shanghai 201620, China (e-mail: bo.shen@dhu.edu.cn).

Hongli Dong is with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing 163318, China, also with the Heilongjiang Provincial Key Laboratory of Networking and Intelligent Control, Northeast Petroleum University, Daqing 163318, China, and also with the Sanya Offshore Oil & Gas Research Institute, Northeast Petroleum University, Sanya 572024, China (e-mail: shiningdhl@vip.126.com).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TII.2024.3485791, IEEE Transactions on Industrial Informatics

progression is to develop "active" DSE schemes, which aim to enhance estimation performance by fully utilizing the real statistics of the measurement noises, thereby moving beyond the limitations of passive approaches.

As of now, adaptive DSE approaches have begun to draw initial research interest in the power system sector with some noteworthy contributions in [7], [18]. For instance, an adaptive Kalman filter has been developed in [18] to enhance the resilience of the SE algorithm in the face of step changes in AC/DC microgrids by employing a novel prediction-error covariance estimation method. Also, an adaptive SE scheme has been introduced in [7] to improve SE performance in power systems under unknown and time-varying measurement errors. This scheme operates under the expectation-maximization (EM) algorithm framework and also utilizes the EM algorithm to estimate the distribution of measurement errors. Unfortunately, in the context of multi-machine power grids, there has been limited focus on adaptive DSE problems that deal with unknown measurement noises.

TABLE I: Representative schemes and results

Schemes	Reference No.	Methodologies
Robust	[1], [22]	Improved particle filter
	[11], [29]	Improved Kalman filter
	[8], [46]	Robust unscented Kalman filter
	[31]	Robust information filter
Adaptive	[7]	Adaptive expectation maximization
	[18], [47]	Prediction-error covariance estimation

In light of the discussions presented earlier and the representative results given in Table. I, there is a noticeable gap in the development of an adaptive DSE algorithm for synchronous-generator-based (SG-based) power grids under measurement noises with unknown statistics. Specifically, 1) the measurement noise is inherently variable over time, which complicates the task of identifying their precise distributions; 2) the identification work requires extensive measurement data, and this poses higher demands on the flexibility and efficiency of the algorithms; and 3) the statistical identification result of measurement noises is dynamical and time-varying due to the evolving nature of power systems, which brings additional challenges in the development of state estimation algorithms.

As such, this paper aims to bridge the gap by introducing an innovative adaptive DSE algorithm with following key contributions.

1) A sliding-window-based scheme is proposed to construct the base Gaussian mixture model (GMM) of the measurement noise, which is achieved using the Kernel density estimation algorithm. The approach utilizes limited measurement data, and is therefore suitable for online applications.
2) A novel fuzzy c-means clustering algorithm that incorporates the Wasserstein distance criterion is developed. This method effectively reduces the component number of the base GMM (resulting in a reduced GMM), thereby enhancing computational efficiency in real-world applications.
3) An improved sliding-window-based DSE scheme is designed under the cubature Kalman filtering algorithm framework, which incorporates the concept of the Gaussian summation filter. Importantly, the parameters of this scheme are adaptively adjusted according to the reduced GMM obtained in the previous sliding window, ensuring dynamic responsiveness to changing grid conditions.

The rest of this paper is organized as follows. Section II formulates the decentralized model of power grids. In Section III, the sliding-window-based estimation method is designed for the distribution of the PMU measurement noises. Section IV investigates the sliding-window-based adaptive DSE algorithm design problem based on the already obtained statistics of the PMU measurement noises. In Section V, simulation studies and discussions are carried out on the IEEE 39-bus system. Finally, some conclusions are drawn in Section VI.

II. Problem Formulation

A. Model of Synchronous Generator

A multi-machine power system with M synchronous generators (SGs) is considered in this paper, where the 4-th order nonlinear discrete-time state-space model of the m-th SG is of the following form [34]:

$$\delta_{m,k+1} = \delta_{m,k} + (\omega_{m,k} - \omega_s)\Delta t, \tag{1a}$$

$$\omega_{m,k+1} = \omega_{m,k} + \frac{\omega_s}{2H_m}\big[T_{m,m,k} - P_{m,k} - D_i(\omega_{m,k} - \omega_s)\big] \times \Delta t, \tag{1b}$$

$$E'_{q,m,k+1} = E'_{q,m,k} + \frac{1}{T'_{d0,m}}\big[-E'_{q,m,k} - (X_{d,m} - X'_{d,m}) \times I_{d,m,k} + E_{fd,m,k}\big]\Delta t, \tag{1c}$$

$$E'_{d,m,k+1} = E'_{d,m,k} + \frac{1}{T'_{q0,m}}\big[-E'_{d,m,k} + (X_{q,m} - X'_{q,m}) \times I_{q,m,k}\big]\Delta t \tag{1d}$$

with

$$\begin{aligned} I_{d,m,k} &= \frac{1}{X'_{di}}(E'_{q,m,k} - V_{q,m,k}), \\ I_{q,m,k} &= \frac{1}{X'_{qi}}(-E'_{d,m,k} + V_{d,m,k}), \\ V_{d,m,k} &= V_{m,k}\sin(\delta_{m,k} - \theta_{m,k}), \\ V_{q,m,k} &= V_{m,k}\cos(\delta_{m,k} - \theta_{m,k}) \end{aligned} \tag{2}$$

where the subscript m is the index of SG $(m = 1, 2, \ldots, N)$; the discretization period and time instant are denoted as Δt and k, respectively; δ and ω are, respectively, the rotor angle and rotor speed of the SG; w_s is the nominal synchronous speed; P is the SG's terminal active power; $\frac{\omega_s}{2H_m}$ stands for the SG's inertia time; D and T_m represent the damping factor and mechanical torque input of SG, respectively; E_{fd} represents the excitation field voltage of SG; E'_d and E'_q are the d-axis and q-axis transient voltages of SG, respectively; X_d and X_q are, respectively, the d-axis and q-axis synchronous reactances of SG; X'_d and X'_q are the d-axis and q-axis transient synchronous reactances of SG, respectively; T'_{d0} and T'_{q0} are the d-axis and q-axis transient open-circuit time instants of SG, respectively; I_d and I_q are, respectively, the d-axis and q-axis currents of SG; V_d and V_q represent the d-axis and q-axis voltages of

Citation information: DOI: 10.1109/TII.2024.3485791, IEEE Transactions on Industrial Informatics

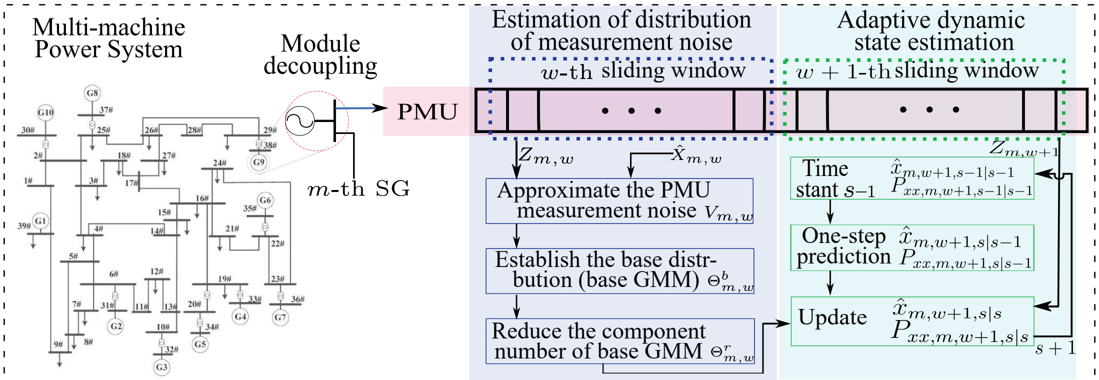


Fig. 1. Diagram of the adaptive DSE for multi-machine power grids with unknown PMU measurement noise statistics.

SG, respectively; and V and θ are the terminal bus voltage magnitude and phase angle of SG, respectively.

Based on (1)-(2), the discretized state-space model of the m-th SG can be written as

$$x_{m,k+1} = g_m(x_{m,k}, u_{m,k}) + w_{m,k} \tag{3}$$

where $x_{m,k} \triangleq \begin{bmatrix}\delta_{m,k} & \omega_{m,k} & E'_{q,m,k} & E'_{d,m,k}\end{bmatrix}^T \in \mathbb{R}^{n_x}$ is the state vector and $u_{m,k} \triangleq \begin{bmatrix}V_{m,k} & \theta_{m,k} & T_{m,m,k} & E_{fd,m,k}\end{bmatrix}^T \in \mathbb{R}^{n_u}$ is the input vector ($k\Delta t$ is simply denoted as k for brevity). The details of $g_m(\cdot)$ is given in (1)-(2), and $w_{m,k}$ represents the process noise satisfying the probability density function (PDF) $p_{w_{m,k}}$.

B. Measurement Model

To facilitate decentralized DSE in power systems, the approach adopted in this paper is inspired by the methodologies proposed in [46]. Specifically, the ideal terminal active and reactive power injections as well as the terminal frequency of the m-th SG are chosen as the PMU measurement, i.e.

$$f_{m,k} = f_0(w_{m,k} - w_s + 1), \tag{4a}$$
$$P_{m,k} = V_{d,m,k}I_{d,m,k} + V_{q,m,k}I_{q,m,k}, \tag{4b}$$
$$Q_{m,k} = -V_{d,m,k}I_{q,m,k} + V_{q,m,k}I_{d,m,k} \tag{4c}$$

where m $(m = 1, 2, \ldots, M)$ and k represent the SG index and the time instant, respectively; f and f_0 are the rotor base frequencies, respectively; P and Q are, respectively, the active and reactive power injections of the terminal bus. The definitions of V_d, V_q, I_d and I_q are all given in (2).

Taking the PMU measurement noise into account, a compact measurement model of the m-th SG can be arranged as

$$z_{m,k} = h_m(x_{m,k}, u_{m,k}) + v_{m,k} \tag{5}$$

where $z_{m,k} \triangleq [f_{m,k} \quad P_{m,k} \quad Q_{m,k}]^T \in \mathbb{R}^{n_y}$ represents the measurement vector, $h_m(\cdot)$ is determined by (4), and $v_{m,k}$ denotes the measurement noise with unknown statistics.

C. Problem Statement

In this paper, the aim is to develop an adaptive DSE algorithm for power grids with the capability to effectively address the challenges posed by non-Gaussian and time-varying measurement noises from PMUs. The algorithm is designed with three primary objectives given as follows.

1) The first objective is to uncover the statistical characteristics of PMU measurement noises, which is achieved by creating a base GMM that reflects these characteristics. Here, the base GMM is established using limited historical measurement data collected within a sliding window.
2) The second objective is to enhance computational efficiency, which is accomplished by reducing the number of components in the base GMM, while still retaining the main density features of the model.
3) The third objective is to estimate the states of a multi-machine power grid adaptively by utilizing the reduced GMM obtained from the previous steps.

III. Estimation for Distribution of Unknown PMU Measurement Noises

In this section, the base GMM is established to describe the statistics of the PMU measurement noises contained in a sliding window, and then the clustering-based algorithm is developed to reduce the component number of the base GMM. The diagram of the proposed approach is shown in Fig. 1.

A. The Establishment of Base GMM

Consider S consecutive scans of PMU measurements contained in the w-th sliding window, i.e.

$$\begin{aligned} Z_{m,w} &\triangleq \{z_{m,k-S}, z_{m,k-S+1}, \ldots, z_{m,k-1}\} \\ &\triangleq \{z_{m,w,1}, z_{m,w,2}, \ldots, z_{m,w,S}\} \end{aligned} \tag{6}$$

with $w = \lceil \frac{k}{S} \rceil$, the corresponding distribution of these PMU measurement noises can be characterized with the aid of previous estimates of measurements and the Kernel density estimation approach. To be specific, let $\hat{X}_{m,w}$ be the estimates of the m-th SG in the w-th sliding window, i.e.

$$\begin{aligned} \hat{X}_{m,w} &\triangleq \{\hat{x}_{m,k-S|k-S}, \hat{x}_{m,k-S+1|k-S+1}, \ldots, \hat{x}_{m,k-1|k-1}\} \\ &\triangleq \{\hat{x}_{m,w,1|1}, \hat{x}_{m,w,2|2}, \ldots, \hat{x}_{m,w,S|S}\}, \end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TII.2024.3485791, IEEE Transactions on Industrial Informatics

and then the corresponding PMU measurement noises of the w-th sliding window can be represented as

$$V_{m,w} \triangleq \{v_{m,k-S}, v_{m,k-S+1}, \dots, v_{m,k-1}\} \triangleq \{v_{m,w,1}, v_{m,w,2}, \dots, v_{m,w,S}\} \tag{7}$$

where $v_{m,w,i} \approx z_{m,w,s} - h_m(\hat{x}_{m,w,s|s}, u_{m,w,s})$ with $u_{m,w,s} = u_{m,w,k-S+s-1}$ $(s = 1, 2, \dots, S)$.

Letting $V_{m,w}$ be the set of the samples of $v_{m,w}$, the base Gaussian mixture density of $v_{m,w}$ can be estimated via

$$f^b(v_{m,w}|\Theta^b_{m,w}) = \frac{1}{S}\sum_{s=1}^{S} K_H(v_{m,w} - v_{m,w,s}) \tag{8}$$

where K_H is the Gaussian kernel function with bandwidth matrix H, and S represents the total number of samples. Note that the base PDF of $v_{m,w}$ with parameter $\Theta^b_{m,w}$ (obtained by applying the kernel density estimation approach) can also be represented by the GMM [35] in which each sample corresponds to a specific component of the GMM. Thus, we rewrite (7) as

$$f^b(v_{m,w}|\Theta^b_{m,w}) = \sum_{s=1}^{S} \alpha^b_{m,w,s}\mathcal{N}(v_{m,w}|\mu^b_{m,w,s}, \Sigma^b_{m,w,s}). \tag{9}$$

The detailed expression of parameter $\Theta^b_{m,w}$ is

$$\Theta^b_{m,w} = [(\theta^b_{m,w,1})^T, (\theta^b_{m,w,2})^T, \dots, (\theta^b_{m,w,S})^T]^T$$

where

$$\theta^b{}_{m,w,s} = [\alpha^b_{m,w,s}, \mu^b_{m,w,s}, \Sigma^b_{m,w,s}]^T,$$

$\mathcal{N}(v_{m,w}|\mu^b_{m,w,s}, \Sigma^b_{m,w,s})$ is the s-th component of the GMM with mean $\mu^b_{m,w,s} = v_{m,w,s}$ and covariance $\Sigma^b_{m,w,s} = H$ $(s = 1, 2, \dots, S)$, and the weight of each component of the base GMM is denoted as $\alpha^b_{m,w,s}$ with $\alpha^b_{m,w,s} = \frac{1}{S}$.

Remark 1: As it has been discussed in [3], any non-Gaussian distribution can be sufficiently represented or approximated by a finite number of Gaussian densities according to the Wienner approximation theorem. In addition, it has been reported in [38] that the GMM offers analytical advantages as it allows for the availability of the joint probability density function (PDF) of multiple random variables. These facts imply that the GMM can effectively encompasses various types of non-Gaussian noises encountered in practical PMU measurements such as logistic distributions and heavy-tailed distributions.

B. Design of Clustering-based Gaussian Mixture Reduction Algorithm

Recalling the base GMM given in (9), we can find that the number of components is large and this would raise the computational complexity issue in the following state estimator design. As such, the primary objective here is to derive a reduced Gaussian mixture density, denoted as $f^r(v_{m,w})$, that closely approximates the base GMM in terms of the Wasserstein distance. The detailed expression of $f^r(v_{m,w})$ is of the following form

$$f^r(v_{m,w}|\Theta^r_{m,w}) = \sum_{l=1}^{L} \alpha^r_{m,w,l}\mathcal{N}(v_{m,w}|\mu^r_{m,w,l}, \Sigma^r_{m,w,l}) \tag{10}$$

where the number of components $L \ll S$, and

$$\Theta^r_{m,w} = [(\theta^r_{m,w,1})^T, (\theta^r_{m,w,2})^T, \dots, (\theta^r_{m,w,L})^T]$$

with $\theta^r_{m,w,l} = [\alpha^r_{m,w,l}, \mu^r_{m,w,l}, \Sigma^r_{m,w,l}]^T$ $(l = 1, 2, \dots, L)$.

Remark 2: For the GMM, a common drawback is that the tendency for the number of components may grow without bound. Thus, if the state estimation algorithm simply follows the statistical model of Gaussian mixture, the number of components may increase exponentially over time. It is worth noting that in the GMM, some of the components have similar shapes or features, and these components can be merged or replaced by a single component. As such, it motivates us to develop algorithms to approximate the GMM with a lower number of components.

The clustering-based Gaussian mixture reduction (C-GMR) algorithm is developed as follows.

1) Initialization of C-GMR algorithm: At this stage, we aim to compute the initial parameter $\Theta^r_{m,w}$ of the reduced Gaussian mixture density shown in (10). To be specific, we follow the idea of [33] to find the two closet component of the base GMM first and then merge them by minimizing the Wasserstein-based average distance [4].

For sake of simplicity, we denote the i-th $(i = 1, 2, \dots, S)$ component of the base GMM shown in (9) as

$$p^b_i(v_{m,w}) \triangleq \mathcal{N}(v_{m,w}|\mu^b_{m,w,i}, \Sigma^b_{m,w,i}).$$

Then, the Wasserstein distance between $p^b_i(v_{m,w})$ and $p^b_j(v_{m,w})$ $(i, j = 1, 2, \dots, S, i \neq j)$ can be represented as

$$\begin{aligned} &W_D\big(p^b_i(v_{m,w}), p^b_j(v_{m,w})\big) \\ =&\mathrm{tr}\Big\{\Sigma^b_{m,w,i} + \Sigma^b_{m,w,j} - 2\big((\Sigma^b_{m,w,i})^{\frac{1}{2}}\Sigma^b_{m,w,j}(\Sigma^b_{m,w,i})^{\frac{1}{2}}\big)\Big\} \\ &+ \|\mu^b_{m,w,i} - \mu^b_{m,w,j}\|^2. \end{aligned} \tag{11}$$

Suppose that the Wasserstein distance given in (11) is minimal, the Gaussian density

$$q^{\mathrm{mer}}_i(v_{m,w}) \triangleq \mathcal{N}(v_{m,w}|\mu^{\mathrm{mer}}_{m,w,i}, \Sigma^{\mathrm{mer}}_{m,w,i}),$$

which is used to merge the candidate components $p_i(v_{m,w})$ and $p_j(v_{m,w})$, can be obtained by minimizing the Wasserstein-based average distance

$$\begin{aligned} J_{ij} =& \bar{\alpha}_i W_D\big(p^b_i(v_{m,w}), q^{\mathrm{mer}}_i(v_{m,w})\big) \\ &+ \bar{\alpha}_j W_D\big(p^b_j(v_{m,w}), q^{\mathrm{mer}}_i(v_{m,w})\big) \end{aligned} \tag{12}$$

with $\bar{\alpha}_i = \frac{\alpha^b_{m,w,i}}{\alpha^b_{m,w,i}+\alpha^b_{m,w,j}}$ and $\bar{\alpha}_j = \frac{\alpha^b_{m,w,j}}{\alpha^b_{m,w,i}+\alpha^b_{m,w,j}}$, and the definition of $W_D(\cdot)$ is given in (11).

Once $q^{\mathrm{mer}}_i(v_{m,w})$ is found, it can be used to replace $p^b_i(v_{m,w})$ and $p^b_j(v_{m,w})$ and thus, the number of the component of the base GMM reduces one. The solution to the problem (12) is given as follows [5]:

$$\alpha^{\mathrm{mer}}_{m,w,i} = \alpha^b_{m,w,i} + \alpha^b_{m,w,j}, \tag{13a}$$

$$\mu^{\mathrm{mer}}_{m,w,i} = \bar{\alpha}_i\mu^b_{m,w,i} + \bar{\alpha}_j\mu^b_{m,w,j}, \tag{13b}$$

$$\begin{aligned} \Sigma^{\mathrm{mer}}_{m,w,i} =& \bar{\alpha}_i^2\Sigma^b_{m,w,i} + \bar{\alpha}_j^2\Sigma^b_{m,w,j} + \bar{\alpha}_i^2\bar{\alpha}_j^2 \\ &\times \big((\Sigma^b_{m,w,i}\Sigma^b_{m,w,j})^{\frac{1}{2}} + (\Sigma^b_{m,w,j}\Sigma^b_{m,w,i})^{\frac{1}{2}}\big). \end{aligned} \tag{13c}$$

Repeating the aforementioned procedure till the desired number L is achieved, and the initialization of $\Theta^r_{m,w}$ is finished.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TII.2024.3485791, IEEE Transactions on Industrial Informatics

2) Computation of cluster centroids: After obtaining the initial value of $\Theta^r_{m,w}$, we would like to treat each component of the initial reduced GMM as one cluster first and then use the Wasserstein-distance-based objective function to find the new centroid of each cluster. For sake of simplicity, we denote the l-th ($l = 1, 2, \ldots, L$) component of the reduced GMM $f^r(v_{m,w}|\Theta^r_{m,w})$ as $q^r_l(v_{m,w}) \triangleq \mathcal{N}(v_{m,w}|\mu^r_{m,w,l}, \Sigma^r_{m,w,l})$.

The considered objective function is of the following form:

$$J_m = \sum_{l=1}^{L}\sum_{i=1}^{S} u^d_{il} W_D\big(p^b_i(v_{m,w}), q^r_l(v_{m,w})\big) \tag{14}$$

where u^d_{il} is the membership degree of $\theta^r_{m,w,l}$ for the l-th cluster with $d \in [1, \infty]$ being a parameter that controls the fuzziness, and $W_D\big(p^b_i(v_{m,w}), q^r_l(v_{m,w})\big)$ represents the Wasserstein distance between $p^b_i(v_{m,w})$ and $q^r_l(v_{m,w})$.

It is worth mentioning that the terms u^d_{il} and $W_D\big(p^b_i(v_{m,w}), q^r_l(v_{m,w})\big)$ in (14) rely on the positions of the new cluster centroids, and one has to find the cluster centroids in a recursive style till the differences between the current value and the previous value of the objective function J_m stay below a predefined threshold.

Taking partial derivation of J_m given in (14) with respect to u_{il} and letting the derivative be zero, we have

$$u_{il} = \frac{1}{\sum_{l'=1}^{L}\left(\frac{W_D\left(p^b_i(v_{m,w}), q^r_l(v_{m,w})\right)}{W_D\left(p^b_i(v_{m,w}), q^r_{l'}(v_{m,w})\right)}\right)^{\frac{1}{d-1}}} \tag{15}$$

where the terms $q^r_l(v_{m,w})$ and $q^r_{l'}(v_{m,w})$ can be obtained by using the cluster centroids of the last iteration.

Similarly, the new weight, mean and covariance of each cluster can also be computed, respectively, via

$$\alpha^{\text{new}}_{m,w,l} = \frac{\sum_{i=1}^{S} u_{il}}{\sum_{l=1}^{L}\sum_{i=1}^{S} u_{il}}, \tag{16a}$$

$$\mu^{\text{new}}_{m,w,l} = \frac{\sum_{i=1}^{S} u^m_{il} v_{m,w,i}}{\sum_{i=1}^{S} u^m_{il}}, \tag{16b}$$

$$\Sigma^{\text{new}}_{m,w,l} = \frac{\sum_{i=1}^{S} u^m_{il} \sum^r_{m,w,i}}{\sum_{i=1}^{S} u^m_{il}}. \tag{16c}$$

For ease of illustration, the pseudocode of our proposed C-GRM algorithm is outlined in Algorithm 1.

Algorithm 1 Estimation for distribution of unknown PMU measurement noises.

For the w-th sliding window:

Load the measurement data to the w-th sliding window via (6).
Approximate the measurement noise via (7).
Establish the base GMM via (8) and (9).

Initialization of C-GMR:

while $S > L$ **do**
  for $i = 1 : S-1$ **do**
    for $j = i+1 : S$ **do**
      Compute $W_D\big(p^b_i(v_{m,w}), p^b_j(v_{m,w})\big)$ via (12).
    end for
  end for
  Find (i, j) such that $\arg\min W_D\big(p^b_i(v_{m,w}), p^b_j(v_{m,w})\big)$.
  Compute the merged terms via (13).
  Obtain $q^{\text{mer}}_i(v_{m,w})$ to merge $p^b_i(v_{m,w})$ and $p^b_j(v_{m,w})$.
  Rebuilt the base GMM via

$$f^b(v_{m,w}|\Theta^b_{m,w}) = f^b(v_{m,w}|\Theta^b_{m,w}) - \alpha^b_{m,w,i}p^b_i(v_{m,w}) - \alpha^b_{m,w,j}p^b_j(v_{m,w}) + \alpha^{\text{mer}}_{m,w,i}q^{\text{mer}}_i(v_{m,w}).$$

  $S = S - 1$.
end while
Obtain the initial result of the C-GMR algorithm: $f^r_{ini}(v_{m,w}|\Theta^r_{m,w}) \leftarrow f^b(v_{m,w}|\Theta^b_{m,w})$.

Implementation of C-GMR:

while $f^r_{new}(v_{m,w}|\Theta^r_{m,w}) \neq f^r_{old}(v_{m,w}|\Theta^r_{m,w})$ **do**
  for $i = 1 : S$ **do**
    for $l = 1 : L$ **do**
      Compute $W_D\big(p^b_i(v_{m,w}), p^r_l(v_{m,w})\big)$.
    end for
    Compute the membership degree u_{il} via (15).
  end for
  for $l = 1 : L$ **do**
    Compute the new terms of each component of reduced GMM via (16).
  end for
  Obtain the newly reduced GMM $f^r_{new}(v_{m,w}|\Theta^r_{m,w})$.
end while

IV. Design of the Adaptive State Estimator

In this section, we aim to develop an adaptive state estimation algorithm for a class of multi-machine power grids under the framework of cubature Kalman filter (CKF), and the idea of the Gaussian summation filter is adopted to tackle the non-Gaussian PMU measurement noises.

A. State Prediction

1) Creation of the cubature points: For the $w+1$-th sliding window, suppose that the estimate $\hat{x}_{m,w+1,s-1|s-1}$ and the estimation error covariance matrix $P_{xx,m,w+1,s-1|s-1}$ are available at time instant $s-1$. Then, the τ-th ($\tau = 1, 2 \ldots, 2n_x$) cubature point can be generated via

$$\eta^\tau_{m,w+1,s-1|s-1} = S_{m,w+1,s-1|s-1}\chi^\tau + \hat{x}_{m,w+1,s-1|s-1} \tag{17}$$

where $S_{m,w+1,s-1|s-1}$ is the square root of $P_{xx,m,w+1,s-1|s-1}$ in the sense of Cholesky decomposition, i.e.

$$S_{m,w+1,s-1|s-1} = \left(\sqrt{P_{xx,m,w+1,s-1|s-1}}\right) \times \left(\sqrt{P_{xx,m,w+1,s-1|s-1}}\right)^T,$$

and χ^τ stands for the scalar parameter whose definition is given as

$$\chi^\tau = \begin{cases} \sqrt{n_x}\boldsymbol{e_\tau}, & \tau = 1, 2, \ldots, n_x \\ -\sqrt{n_x}\boldsymbol{e_{\tau-1}}, & \tau = n_x+1, 2, \ldots, 2n_x \end{cases}$$

with $\boldsymbol{e_\tau}$ being a unit vector.

2) Propagation of the new points: We now propagate each point through the state-transition function $g_m(\cdot)$ given in (3) to generate a new set of transformed cubature points $\xi^\tau_{m,w+1,s|s-1}$ ($\tau = 1, 2 \ldots, 2n_x$) with the following form

$$\xi^\tau_{m,w+1,s|s-1} = g_m(\eta^\tau_{m,w+1,s-1|s-1}, u_{m,w+1,s-1}). \tag{18}$$

3) Computation of the prediction and prediction error covariance: The prediction $\hat{x}_{m,w+1,s|s-1}$ and prediction error covariance $P_{xx,m,w+1,s|s-1}$ can be, respectively, calculated via

$$\hat{x}_{m,w+1,s|s-1} = \sum_{\tau=1}^{2n_x} \frac{1}{2n_x} \xi^{\tau}_{m,w+1,s|s-1} \quad (19)$$

and

$$\begin{aligned} P_{xx,m,w+1,s|s-1} =& \sum_{\tau=1}^{2n_x} \frac{1}{2n_x} \tilde{\xi}^{\tau}_{m,w+1,s|s-1} (\tilde{\xi}^{\tau}_{m,w+1,s|s-1})^T \\ &+ Q_{m,w+1,s} \end{aligned} \quad (20)$$

with $\tilde{\xi}^{\tau}_{m,w+1,s|s-1} \triangleq \xi^{\tau}_{m,w+1,s|s-1} - \hat{x}_{m,w+1,s|s-1}$.

B. Update

1) Creation of the new cubature points: The τ-th ($\tau = 1, 2 \ldots, 2n_x$) predicted cubature point, which is mapped through the measurement function $h_m(\cdot)$ in (5), is of the following form

$$\varepsilon^{\tau}_{m,w+1,s|s-1} = h_m(\varphi^{\tau}_{m,w+1,s|s-1}, u_{m,w+1,s-1}) \quad (21)$$

where $\varphi^{\tau}_{m,w+1,s|s-1} = S_{m,w+1,s|s-1}\chi^{\tau} + \hat{x}_{m,w+1,s|s-1}$ with $S_{m,w+1,s|s-1}$ denoting the square root of the matrix $P_{xx,m,w+1,s|s-1}$ in the sense of Cholesky decomposition.

2) Prediction of the measurement and its covariance matrices: The prediction of the measurement $\hat{z}^{l}_{m,w+1,s|s-1}$ can be computed by

$$\hat{z}^{l}_{m,w+1,s|s-1} = \sum_{\tau=1}^{2n_x} \beta^{\tau} \frac{1}{2n_x} \varepsilon^{\tau}_{m,w+1,s|s-1} + \mu^{r}_{m,w-1,l}. \quad (22)$$

where $\mu^{r}_{m,w-1,l}$ represents the l-th ($l = 1, 2, \ldots, L$) component of the reduced GMM of the PMU measurement noises which is obtained by using the previous measurements contained in the w-th sliding window. The corresponding measurement prediction error covariance matrix $P_{zz,m,w+1,s|s-1}$ and the state-measurement prediction error cross-covariance matrix $P_{xz,m,w+1,s|s-1}$ can be computed, respectively, as

$$\begin{aligned} P^{l}_{zz,m,w+1,s|s-1} =& \sum_{\tau=1}^{2n_x} \beta^{\tau} (\varepsilon^{\tau}_{m,w+1,s|s-1} - \hat{z}^{l}_{m,w+1,s|s-1}) \\ &\times (\varepsilon^{\tau}_{m,w+1,s|s-1} - \hat{z}^{l}_{m,w+1,s|s-1})^T \\ &+ \Sigma^{r}_{m,w-1,l} \end{aligned} \quad (23)$$

and

$$\begin{aligned} P^{l}_{xz,m,w+1,s|s-1} =& \sum_{\tau=1}^{2n_x} \beta^{\tau} (\varphi^{\tau}_{m,w+1,s|s-1} - \hat{x}_{m,w+1,s|s-1}) \\ &\times (\varepsilon^{\tau}_{m,w+1,s|s-1} - \hat{z}^{l}_{m,w+1,s|s-1})^T. \end{aligned} \quad (24)$$

3) State estimation: The estimation $\hat{x}_{m,w+1,s|s}$ and its corresponding estimation error covariance $P_{xx,m,w+1,s|s}$ can be calculated, respectively, via

$$\hat{x}_{m,w+1,s|s} = \sum_{l=1}^{L} \alpha^{r}_{m,w-1,l} \hat{x}^{l}_{m,w+1,s|s} \quad (25)$$

and

$$P_{xx,m,w+1,s|s} = \sum_{l=1}^{L} \alpha^{r}_{m,w-1,l} P^{l}_{xx,m,w+1,s|s} \quad (26)$$

where $\hat{x}^{l}_{m,w+1,s|s}$ and $P^{l}_{xx,m,w+1,s|s}$ are, respectively, of the following form:

$$\begin{aligned} \hat{x}^{l}_{m,w+1,s|s} =& \hat{x}_{m,w+1,s|s-1} + K^{l}_{m,w+1,s} \\ &\times (z_{m,w+1,s} - \hat{z}^{l}_{m,w+1,s|s-1}) \end{aligned}$$

and

$$\begin{aligned} P^{l}_{xx,m,w+1,s|s} =& P_{xx,m,w+1,s|s-1} - K^{l}_{m,w+1,s} P^{l}_{zz,m,w+1,s|s-1} \\ &\times (K^{l}_{m,w+1,s})^T \end{aligned}$$

with the gain matrix

$$K^{l}_{m,w+1,s} = P^{l}_{xz,m,w+1,s|s-1} (P^{l}_{zz,m,w+1,s|s-1})^{-1}. \quad (27)$$

Remark 3: In comparison with the conventional CKF, the features of the proposed method are that: 1) the state estimator is of parallel style and each CKF-based parallel branching is improved by the Gaussian summation filter; 2) every component of the reduced Gaussian mixture density can be fully utilized in the corresponding parallel branching of state estimator; and 3) the parameters of the proposed state estimator can be matched adaptively with the identification results of the PMU measurement noise statistics (i.e. the output of the C-GMR algorithm).

For ease of illustration, the pseudocode of our proposed adaptive DSE algorithm is outlined in Algorithm 2.

Algorithm 2 Adaptive DSE for multi-machine power grids.

Require: The estimate $\hat{x}_{m,w,S|S}$ and its covariance matrix $P_{xx,m,w,S|S}$ at time instant S of the w-th sliding window; the statistics of the PMU noises of the w-th sliding window (i.e. $\Theta^{r}_{m,w}$).

Adaptive state estimation for the $w+1$-th sliding window:

1: **for** $s = 1 : S$ **do**
2: Create cubature points $\eta^{\tau}_{m,w+1,s-1|s-1}$ ($\tau = 1, 2, \ldots, 2n_x$) via (17).
3: Generate transformed cubature points $\xi^{\tau}_{m,w+1,s|s-1}$ via (18).
4: Compute prediction state $\hat{x}_{m,w+1,s|s-1}$ and corresponding covariance matrix $P_{xx,m,w+1,s|s-1}$ in terms of (19)-(20).
5: Create predicted cubature points $\varepsilon^{\tau}_{m,w+1,s|s-1}$ by using (21).
6: Compute prediction measurement $\hat{z}^{l}_{m,w+1,s|s-1}$ and corresponding covariance matrices $P^{l}_{zz,w+1,s|s-1}$ and $P^{l}_{xz,w+1,s|s-1}$, respectively, via (21) and (23) with the distribution of PMU measurement noise $\mu^{r}_{m,w-1,l}$ and $\Sigma^{r}_{m,w-1,l}$ ($l = 1, 2, \ldots, L$) obtained in the previous w-th sliding window.
7: Update prediction values via (24)-(26) to derive $\hat{x}_{m,w+1,s|s}$ and $P_{xx,m,w+1,s|s}$.
8: **end for**

V. Simulation Example

To validate the effectiveness of the proposed algorithm, this section details the simulation experiments conducted on the IEEE 39-bus system [17]. In order to save space, only the 3-rd SG's states are considered. The distribution of the initial states is set as $x_{3,0} \sim \mathcal{N}([0.9\ 0.5\ 0.1\ 0.45]^T, \text{diag}_4\{0.01^2\})$.

Citation information: DOI: 10.1109/TII.2024.3485791, IEEE Transactions on Industrial Informatics

Moreover, 40 PMU scans (40% of the PMU scans collected in 1 second) contained in each sliding window are utilized in the proposed clustering-based GMR algorithm (labeled as C-GMR), and the number of the reduced Gaussian mixture component is set as 4. The average mean square error (AMSE) is adopted to assess the overall estimation performance of the 3-th SG, i.e. $\text{AMSE} = \frac{1}{n_x}\left(x_{3,k} - \hat{x}_{3,k}\right)^2$.

A. Scenario 1: Gaussian Distribution with Constant Parameters

In this subsection, it is assumed that the PMU measurement noise follows Gaussian distribution whose parameters are constant, i.e. $v_{3,k} \sim \mathcal{N}([0\ 0\ 0]^T, \text{diag}_4\{0.01^2\})$.

The results from the C-GMR algorithm are depicted in Fig. 2. For brevity, only the frequency measurement is considered. Specifically: 1) the first subfigure shows the actual distribution of the frequency measurement noise; 2) the second subfigure displays the measurement curve under this noise distribution; and 3) the third subfigure presents the estimation results achieved by the proposed C-GMR algorithm. Fig. 2 illustrates that the proposed C-GMR algorithm is effective, especially when the measurement noise follows a Gaussian distribution with constant parameters.

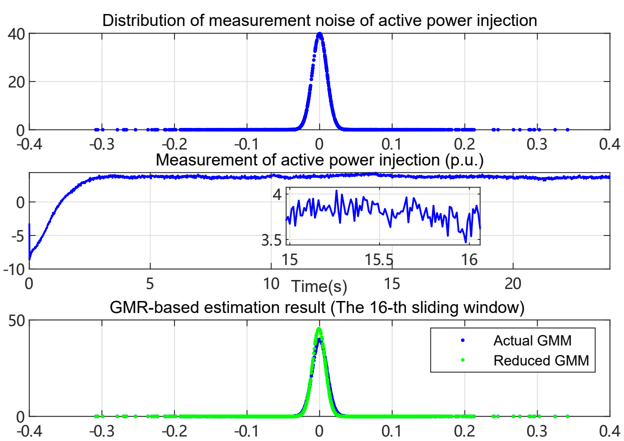
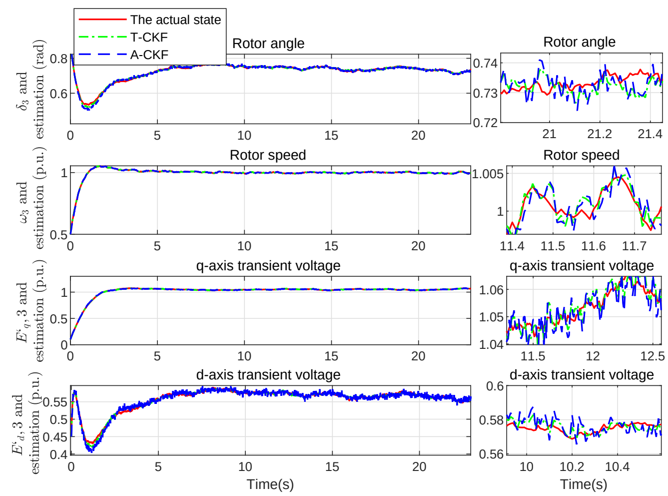


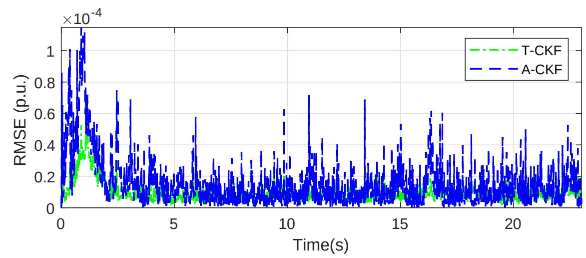
Fig. 2. Scenario 1: The estimation results of the Gaussian distribution characterized by constant parameters.

Based on the framework of our proposed adaptive state estimation algorithm, comparisons between the true distribution of the measurement noise (labeled as T-CKF) and the estimated distribution of the measurement noise (labeled as A-CKF) are carried out. The simulation results are given in Fig. 3 where: 1) the state trajectories and the corresponding estimates of SG 3 with the T-CKF and A-CKF are plotted in Fig. 3(a); and 2) the associated AMSEs are shown in Fig. 3(b).

From Fig. 3, it can be observed that under the Gaussian measurement noise with constant parameters: 1) the estimation performance of the A-CKF is close to T-CKF, which verifies that our proposed adaptive state estimation algorithm is effective; and 2) the proposed C-GMR algorithm is effective in revealing the distribution of the PMU measurement noise whose prior knowledge is totally unknown.



(a)



(b)

Fig. 3. Scenario 1: Results of SG 3 (a) Estimation. (b) AMSE.

B. Scenario 2: Non-Gaussian Distribution with Random Parameters

This subsection aims to verify the effectiveness of the proposed C-GMR algorithm and the proposed adaptive state estimation algorithm under the unknown and time-varying non-Gaussian PMU measurement noises. The GMM with 5 components is adopted to characterize the corresponding non-Gaussian measurement noises $v_{3,k}$, i.e. $p(v_{3,k}) = \sum_{n=1}^{N} \alpha_n \mathcal{N}_n(x|\mu_n, \Sigma_n)$ with $N = 5$. Specifically, the parameters of the main component are set as $\alpha_1 = 0.5$, $\mu_1 = 0$ and $\Sigma_1 = 0.06$, and the parameters of the rest components are randomly generated from the following ranges: $\alpha_n \in [0.1, 0.5]$, $\mu_n \in [0, 1]$ and $\Sigma_n \in [0.03, 0.07]$ with $n = 2, \ldots, 5$.

The measurement of frequency is again used for illustration, with the results of the C-GMR algorithm displayed in Fig. 4 where: 1) the first subfigure presents the actual distribution of the measurement noise for the frequency; 2) the second subfigure illustrates the curve corresponding to the frequency under the influence of the noise distribution; and 3) the third subfigure demonstrates the estimation result achieved using the proposed C-GMR algorithm. An observation from Fig. 4 reveals that the proposed C-GMR algorithm is capable of effectively estimating the distribution of measurement noise, particularly when the noise follows a multimodal distribution with randomly generated parameters.

Similar to Scenario 1, comparisons between the T-CKF and the A-CKF under Scenario 2 are illustrated in Fig. 5. Observa-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TII.2024.3485791, IEEE Transactions on Industrial Informatics

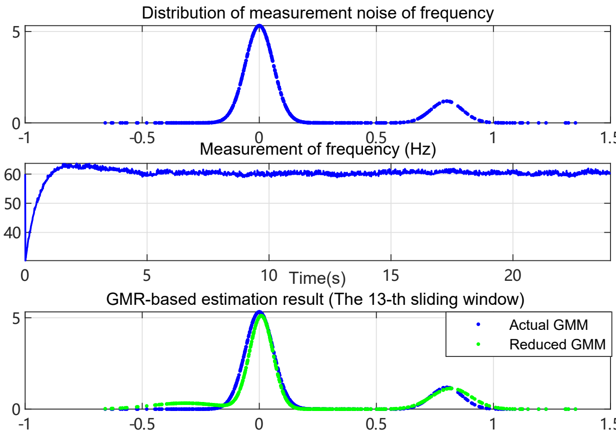
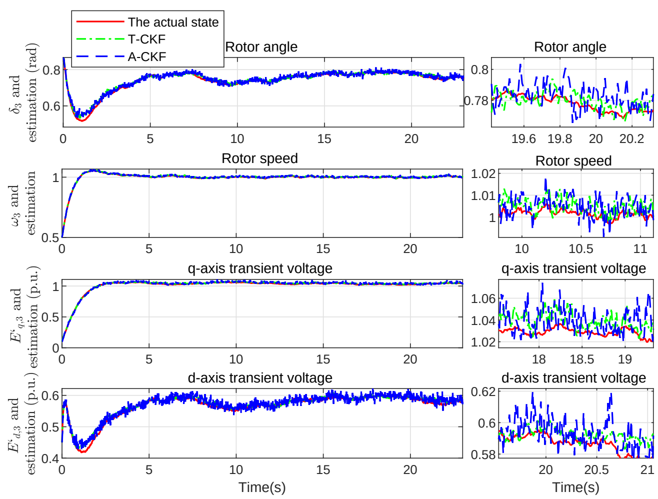
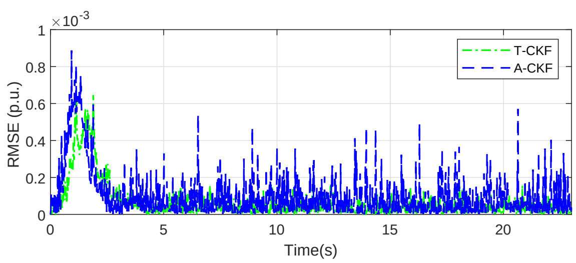


Fig. 4. Scenario 2: Estimation results of the measurement noise under the non-Gaussian distribution with random parameters.



(a)



(b)

Fig. 5. Scenario 2: Results of SG 3 (a) Estimation. (b) AMSE.

tions from this figure reveal several key points: 1) the proposed adaptive state estimation algorithm demonstrates effectiveness, particularly when the measurement noise follows a multimodal distribution with randomly generated parameters, which is evident from the accurate tracking of the real state trajectories and the close performance between the T-CKF and A-CKF; and 2) the results also underscore the effectiveness of the proposed C-GMR algorithm, and this is particularly notable in the context of parameter identification for measurement noise, which significantly influences the performance of state estimation.

C. Scenario 3: Complicated non-Gaussian Distribution with Random Parameters

In the context of clustering, one of the notable challenges arises when the data characteristics are closely similar. In this scenario, a more complex situation is considered to address this challenge, i.e. the measurement noise $v_{3,k}$ are generated randomly from the following ranges: $\alpha_n \in [0.1, 0.5]$, $\mu_n \in [0, 0.18]$ and $\Sigma_n \in [0.055, 0.06]$ with $n = 1, \ldots, 5$.

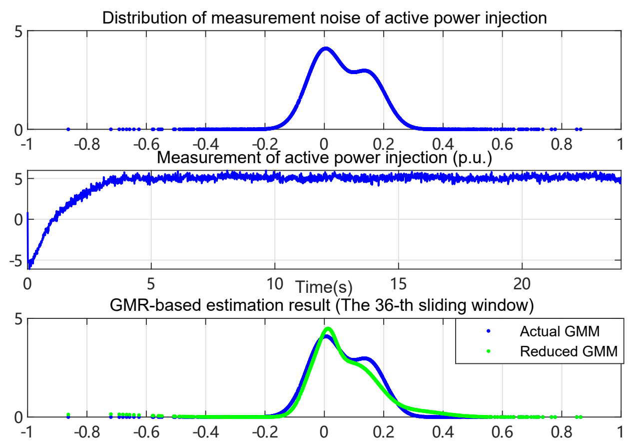


Fig. 6. Scenario 3: Estimation results of the measurement noise under the complicated non-Gaussian distribution.

To conserve space, the measurement of active power is used as an example. Fig. 6 displays the outcomes of the C-GMR algorithm. To be specific: 1) the first subfigure illustrates the actual distribution of the measurement noise for active power; 2) the second subfigure shows the measurement curve, depicting how the noise influences these measurements; and 3) the third subfigure presents the estimation results achieved using the C-GMR algorithm, highlighting its performance in parameter estimation. An analysis of Fig. 6 reveals that the proposed C-GMR algorithm is capable of effectively estimating the parameters of non-Gaussian measurement noise, even when the parameters of each component are very close.

Fig. 7 presents the comparative results between the T-CKF and the A-CKF under Scenario 3, where: 1) the proposed adaptive state estimation algorithm shows acceptable performance in tracking the true trajectories of the states, even under the influence of complicated measurement noise; 2) the figure also illustrates the differences between the T-CKF and A-CKF, providing insight into the adaptive capabilities and effectiveness of the A-CKF in comparison to the traditional approach; and 3) the estimation results validate the capability of the proposed C-GMR algorithm in accurately revealing the distribution of the measurement noise.

D. Discussions

In order to further discuss the overall performance of the proposed scheme under different scenarios, the average of the sum of AMSE (A-AMSE), defined by $\mathrm{A-AMSE} = \frac{1}{K}\frac{1}{n_x}\sum_{k=1}^{K}\left(x_{3,k} - \hat{x}_{3,k}\right)^2$, is selected as the index with K

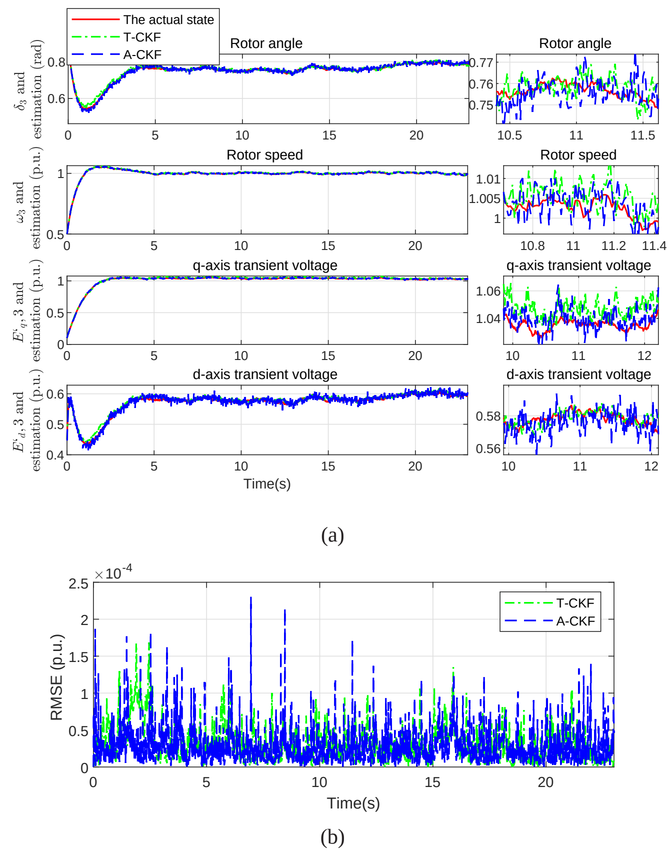


Fig. 7. Scenario 3: Results of SG 3 (a) Estimation. (b) AMSE.

being the total running steps, and the corresponding results are given in Fig. 8. It is worth noting that the T-CKF and A-CKF are all based on the proposed adaptive state estimation, and only the measurement noise distributions used are different (T-CKF is based on the real distributions and A-CKF is based on the identified distributions).

From Fig. 8, we can find that the state estimation performance of A-CKF is close to the one of T-CKF. Moreover, Fig. 8 also demonstrates that the statistic identification result for the unknown PMU measurement noises is close to its true value (since the differences between the T-CKF and A-CKF are only the distributions of the measurement noises used).

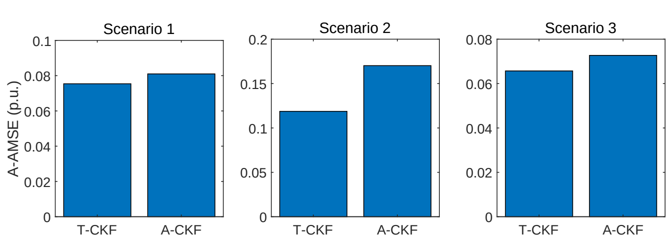


Fig. 8. A-AMSEs of T-CKF and A-CKF with three scenarios.

VI. Conclusion

In this paper, the adaptive state estimation problem has been addressed for a class of multi-machine power grids under PMU measurement noises with unknown statistics. A sliding-window-based algorithm has been developed to ascertain the statistical properties of the PMU measurement noises with aim to reduce computational costs. For adaptive state estimation, an enhanced cubature Kalman filtering algorithm has been introduced by incorporating the Gaussian summation filter concept. The parameters of this algorithm have been dynamically adjusted using the reduced GMM obtained from the latest sliding window. Finally, experiments on three test scenarios on the IEEE 39-bus system have been conducted to demonstrate the effectiveness of our proposed method. An interesting topic for future research would be to develop deep clustering schemes to further improve the identification results of the statistics of PMU measurement noises [9], [15], [24], [25], [43], [44]. Also, the main results of this paper can be extended to more general systems under network-induced phenomena [13], [14], [21], [26]–[28], [32], [36], [41].

References

[1] A. H. Abolmasoumi, A. Farahani and L. Mili, Robust particle filter design with an application to power system state estimation, *IEEE Transactions on Power Systems*, vol. 39, no. 1, pp. 1810–1821, 2024. .

[2] 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, 2020.

[3] I. Arasaratnam, S. Haykin and R. J. Ellioot, Discrete-time nonlinear filtering algorithm using Gauss-Hermite quadrature, in *Proceedings of IEEE*, vol. 95, no. 5, pp. 953–977, 2017.

[4] A. Assa and K. N. Plataniotis, Wasserstein-Distance-Based Gaussian mixture reduction, *IEEE Signal Processing Letters*, vol. 25, no. 10, pp. 1465–1469, 2018.

[5] A. N. Bishop, Information fusion via the Wasserstein barycenter in the space of probability measures: Direct fusion of empirical measures and Gaussian fusion with unknown correlation, in *Proceedings of 17th International Conference on Information Fusion*, Salamanca, Spain, 2014, pp. 1–7.

[6] T. Bi, M. Zhu and H. Liu, A powerful tool for power system monitoring: Distributed dynamic state estimation based on a full-view synchronized measurement system, *IEEE Power and Energy Magazine*, vol. 21, no. 1, pp. 26–35, 2023.

[7] G. Cheng, Y. Lin, Y. Chen and T. Bi, Adaptive state estimation for power systems measured by PMUs with unknown and time-varying error statistics, *IEEE Transactions on Power Systems*, vol. 36, no. 5, pp. 4482–4491, 2021.

[8] L. Dang, B. Chen, S. Wang, W. Ma and P. Ren, 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, 2020.

[9] J. Dou and Y. Song, An improved generative adversarial network with feature filtering for imbalanced data, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 4, art. no. 100017, Dec. 2023.

[10] J. Fang, W. Liu, L. Chen, S. Lauria, A. Miron and X. Liu, A survey of algorithms, applications and trends for particle swarm optimization, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 1, pp. 24-50, Mar. 2023.

[11] X. Fan, G. Wang, J. Han and Y. Wang, A background-impulse Kalman filter with non-Gaussian measurement noises, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 4, pp. 2434–2443, 2023.

[12] Z. Fang, Y. Lin, S. Song, C. Li, X. Lin and Y. Chen, State estimation for situational awareness of active distribution system with photovoltaic power plants, *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 239–250, 2021.

[13] S. Feng, X. Li, S. Zhang, Z. Jian, H. Duan and Z. Wang, A review: state estimation based on hybrid models of Kalman filter and neural network, *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2173682, 2023.

[14] H. Gao, Y. Li, L. Yu and H. Yu, Collaborative-prediction-based recursive filtering for nonlinear systems with sensor saturation under duty cycle scheduling, *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2247007, 2023.

[15] X. Guo, Z. Bi, J. Wang, S. Qin, S. Liu and L. Qi, Reinforcement learning for disassembly system optimization problems: A Survey, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 1, pp. 1–14, Mar. 2023.

[16] M. Goel and A. Abur, A hybrid state estimator for systems with limited number of PMUs, *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1511–1517, 2015.

[17] IEEE PES TF on Benchmark System for Stability Controls,(2015). Benchmark systems for small-signal stability analysis and control, IEEE Power Energy Society, Piscataway, New Jersey, USA, Technical Report, PES-TR18.

[18] W. Lambrichts and M. Paolone, Linear recursive state estimation of hybrid and unbalanced AC/DC micro-grids using synchronized measurements, *IEEE Transactions on Smart Grid*, vol. 14, no. 1, pp. 54–67, 2023.

[19] W. Li and F. Yang, Information fusion over network dynamics with unknown correlations: An overview, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 2, art. no. 100003, 2023.

[20] X. Li, S. Feng, N. Hou, R. Wang, H. Li, M. Gao and S. Li, Surface microseismic data denoising based on sparse autoencoder and Kalman filter, *Systems Science & Control Engineering*, vol. 10, no. 1, pp. 616-628, 2022.

[21] X. Li and D. Ye, Dynamic event-triggered distributed filtering design for interval type-2 fuzzy systems over sensor networks under deception attacks, *International Journal of Systems Science*, vol. 54, no. 15, pp. 2875–2890, 2023.

[22] X. Liu, L. Li, Z. Li, X. Chen, T. Fernando, H. H.-C. Iu and G. He, Event-trigger particle filter for smart grids with limited communication bandwidth infrastructure, *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6918–6928, 2018.

[23] Z. Lu and G. Guo, Control and communication scheduling co-design for networked control systems: a survey, *International Journal of Systems Science*, vol. 54, no. 1, pp. 189-203, 2023.

[24] X. Luo, H. Wu, Z. Wang, J. Wang and D. Meng, A novel approach to large-scale dynamically weighted directed network representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 9756–9773, Dec. 2022.

[25] X. Luo, Y. Yuan, S. Chen, N. Zeng and Z. Wang, Position-transitional particle swarm optimization-incorporated latent factor analysis, *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 8, pp. 3958–3970, Aug. 2022.

[26] S. Ma and Y. Li, Adaptive fuzzy fault-tolerant control for active seat suspension systems with full-state constraints, *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2153391, 2023.

[27] Z.-H. Pang, T. Mu, F.-Y. Hou, Y. Shi and J. Sun, Two networked predictive control methods for output tracking of networked systems with plant-model mismatch, *International Journal of Systems Science*, vol. 54, no. 10, pp. 2073–2088, 2023.

[28] W. Qian, D. Lu, S. Guo and Y. Zhao, Distributed state estimation for mixed delays system over sensor networks with multichannel random attacks and Markov switching topology, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 6, pp. 8623–8637, Jun. 2024.

[29] B. Qu, Z. Wang, B. Shen and H. Dong, Outlier-resistant recursive state estimation for renewable-electricity-generation-based microgrids, *IEEE Transactions on Industrial Informatics*, vol. 19, no. 5, pp. 7133-7144, 2023.

[30] B. Qu, Z. Wang, B. Shen and H. Dong, Decentralized dynamic state estimation for multi-machine power systems with non-Gaussian noises: Outlier detection and localization, *Automatica*, vol. 153, art.no. 111010, 2023.

[31] B. Qu, Z. Wang, B. Shen, H. Dong and H. Liu, Anomaly-resistant decentralized state estimation under minimum error entropy with fiducial points for wide-area power systems, *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 1, pp. 74–87, 2024.

[32] A. Rajagopal and S. Chitraganti, State estimation and control for networked control systems in the presence of correlated packet drops, *International Journal of Systems Science*, vol. 54, no. 11, pp. 2352–2365, 2023.

[33] A. R. Runnalls, Kullback-Leibler approach to Gaussian mixture reduction, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 989–999, 2007.

[34] P. W. Sauer and M. Pai, Power system dynamics and stability, Englewood Cliffs, NJ, USA: Prentice-Hall, 1997.

[35] B. W. Silverman, Density estimation for statistics and data analysis, FL, USA: CRC Press, 1986.

[36] Y. Wang, H.-J. Liu and H.-L. Tan, An overview of filtering for sampled-data systems under communication constraints, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 3, art. no. 100011, Sep. 2023.

[37] Y.-A. Wang, B. Shen, L. Zou and Q.-L. Han, A survey on recent advances in distributed filtering over sensor networks subject to communication constraints, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 2, art. no. 100007, 2023.

[38] Z. Wang, C. Shen, F. Liu and F. Gao, Analytical expressions for joint distributions in probabilistic load flow, *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2473–2474, 2017.

[39] Z. Wang, C. Shen, F. Liu, X. Wu, C.-C. Liu and F. Gao, Chance-constrained economic dispatch with non-Gaussian correlated wind power uncertainty, *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4880–4893, 2017.

[40] S. Wang, J. Zhao, Z. Huang and R. Diao, Assessing Gaussian assumption of PMU measurement error using field data, *IEEE Transactions on Power Delivery*, vol. 33, no. 6, pp. 3233–3236, 2018.

[41] Y. Xu, W. Lv, W. Lin, R. Lu and D. E. Quevedo, On extended state estimation for nonlinear uncertain systems with round-robin protocol, *Automatica*, vol. 138, Article 110154, 2022.

[42] J. Yang, W.-A. 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, 2020.

[43] Y. Yuan, G. Ma, C. Cheng, B. Zhou, H. Zhao, H.-T. Zhang and H. Ding, A general end-to-end diagnosis framework for manufacturing systems, *National Science Review*, vol. 7, no. 2, pp. 418–429, 2020.

[44] Y. Yuan, X. Tang, W. Zhou, W. Pan, X. Li, H.-T. Zhang, H. Ding and J. Goncalves, Data driven discovery of cyber physical systems, *Nature Communications*, vol. 10, no. 1, pp. 1–9, 2019.

[45] D. Yue, C. Deng, C. Wen and W. Wang, Finite-time distributed resilient tracking control for nonlinear MASs with application to power systems, *IEEE Transactions on Automatic Control*, vol. 69, no. 5, pp. 3128–3143, May. 2024.

[46] J. Zhao and L. Mili, Power system robust decentralized dynamic state estimation based on multiple hypothesis testing, *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4553–4562, 2018.

[47] L. Zanni, J.-Y. Le Boudec, R. Cherkaoui and M. Paolone, A prediction-error covariance estimator for adaptive Kalman filtering in step-varying processes: Application to power-system state estimation, *IEEE Transactions on Control Systems Technology*, vol. 25, no. 5, pp. 1683–1697, 2017.

[48] H. Zhang, D. Yue, C. Dou, X. Xie, K. Li and G. P. Hancke, Resilient optimal defensive strategy of TSK fuzzy-model-based Microgrids' system via a novel reinforcement learning approach, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 4, pp. 1921-1931, 2023.

[49] M. Zhong, X. Zhu, T. Xue and L. Zhang, An overview of recent advances in model-based event-triggered fault detection and estimation, *International Journal of Systems Science*, vol. 54, no. 4, pp. 929-943, 2023.

Bogang Qu received the B.Eng. degree in electrical engineering and automation and the M.Eng. degree in electrical engineering from University of Shanghai for Science and Technology, Shanghai, China, in 2014 and 2017, respectively, and the Ph.D. degree in control science and engineering from Donghua University, Shanghai, China, in 2022.

He is currently a Lecturer with the College of Automation Engineering, Shanghai University of Electric Power, Shanghai. From November 2021 to October 2022, he was a visiting Ph.D. Student with the Department of Computer Science, Brunel University London, Uxbridge, U.K. His current research interests include state estimation and information perception and their applications in smart grids. He is a very active reviewer for many international journals.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TII.2024.3485791, IEEE Transactions on Industrial Informatics

Zidong Wang (Fellow, IEEE) received the B.Sc. degree in mathematics in 1986 from Suzhou University, Suzhou, China, and the M.Sc. degree in applied mathematics in 1990 and the Ph.D. degree in electrical engineering in 1994, both from Nanjing University of Science and Technology, Nanjing, China.

He is currently Professor of Dynamical Systems and Computing in the Department of Computer Science, Brunel University London, U.K. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the UK. Prof. Wang's research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published more than 700 papers in international journals. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, William Mong Visiting Research Fellowship of Hong Kong.

Prof. Wang serves (or has served) as the Editor-in-Chief for *International Journal of Systems Science*, the Editor-in-Chief for *Neurocomputing*, the Editor-in-Chief for *Systems Science & Control Engineering*, and an Associate Editor for 12 international journals including IEEE Transactions on Automatic Control, IEEE Transactions on Control Systems Technology, IEEE Transactions on Neural Networks, IEEE Transactions on Signal Processing, and IEEE Transactions on Systems, Man, and Cybernetics-Part C. He is a Member of the Academia Europaea, a Member of the European Academy of Sciences and Arts, an Academician of the International Academy for Systems and Cybernetic Sciences, a Fellow of the IEEE, a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.

Bo Shen (Senior Member, IEEE) received the B.Sc. degree in mathematics from Northwestern Polytechnical University, Xi'an, China, in 2003, and the Ph.D. degree in control theory and control engineering from Donghua University, Shanghai, China, in 2011.

He is currently a Professor with the School of Information Science and Technology, Donghua University. From 2009 to 2010, he was a Research Assistant with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong. From 2010 to 2011, he was a Visiting Ph.D. Student with the Department of Information Systems and Computing, Brunel University London, London, U.K. From 2011 to 2013, he was a Research Fellow (Scientific Co-Worker) with the Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Duisburg, Germany. He has published around 80 articles in refereed international journals. His research interests include nonlinear control and filtering, stochastic control and filtering, as well as complex networks and neural networks.

Prof. Shen serves (or has served) as an Associate Editor or Editorial Board Member for eight international journals, including *Systems Science and Control Engineering*, *Journal of The Franklin Institute*, *Asian Journal of Control*, *Circuits, Systems, and Signal Processing*, *Neurocomputing*, *Assembly Automation*, *Neural Processing Letters*, and *Mathematical Problems in Engineering.* He is a program committee member for many international conferences.

Hongli Dong (Senior Member, IEEE) received the Ph.D. degree in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2012.

From 2009 to 2010, she was a Research Assistant with the Department of Applied Mathematics, City University of Hong Kong, Hong Kong. From 2010 to 2011, she was a Research Assistant with the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong. From 2011 to 2012, she was a Visiting Scholar with the Department of Information Systems and Computing, Brunel University London, Uxbridge, U.K. From 2012 to 2014, she was an Alexander von Humboldt Research Fellow with the University of Duisburg–Essen, Duisburg, Germany. She is currently a Professor with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing, China. She is also the Director of the Heilongjiang Provincial Key Laboratory of Networking and Intelligent Control, Daqing. Her current research interests include robust control and networked control systems.

Dr. Dong is a very active reviewer for many international journals.

Daogang Peng received his Ph.D. degree from Tongji University, Shanghai, China in 2009. Now he is a professor in the College of Automation Engineering of Shanghai University of Electric Power. His research interests include low carbon intelligent power generation, integrated energy system, virtual power plant and power inspection robot.