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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

Anomaly-Resistant Decentralized State Estimation Under Minimum Error Entropy With Fiducial Points for Wide-Area Power Systems

Bogang Qu, Zidong Wang, Bo Shen, Hongli Dong, and Hongjian Liu

***Abstract*—This paper investigates the anomaly-resistant decentralized state estimation (SE) problem for a class of wide-area power systems which are divided into several non-overlapping areas connected through transmission lines. Two classes of measurements (i.e., local measurements and edge measurements) are obtained, respectively, from the individual area and the transmission lines. A decentralized state estimator, whose performance is resistant against measurement with anomalies, is designed based on the minimum error entropy with fiducial points (MEEF) criterion. Specifically, i) an augmented model, which incorporates the local prediction and local measurement, is developed by resorting to the unscented transformation approach and the statistical linearization approach; ii) using the augmented model, an MEEF-based cost function is designed that reflects the local prediction errors of the state and the measurement; and iii) the local estimate is first obtained by minimizing the MEEF-based cost function through a fixed-point iteration and then updated by using the edge measuring information. Finally, simulation experiments with three scenarios are carried out on the IEEE 14-bus system to illustrate the validity of the proposed anomaly-resistant decentralized SE scheme.**

***Index Terms*—Wide-area power systems, decentralized state estimation, minimum error entropy, unscented Kalman filter, measurements with anomalies.**

I. Introduction

For a few decades, power systems have been undergoing dramatic changes due to the penetration of renewable power generations, the widespread use of demand-response devices as well as access to dynamic loads [1], [2]. In order to monitor power systems under complex working conditions, a new generation of infrastructure, namely, wide area monitoring systems (WAMS), has been increasingly deployed in modern power systems [3]. As a key part of the WAMS, the state estimation (SE) algorithm has been playing a vitally important role in raising situation awareness, facilitating real-time control and enhancing security assessment and protection [4]–[6].

So far, the SE algorithms based on the data collected from the phasor measurement unit (PMU) have spurred tremendous interest (see e.g. [7], [8]) because of PMU's merits of providing synchronized, accurate yet timely sensing data. Nevertheless, due primarily to unaffordable implementation costs and the limited communication resources, it is difficult to deploy the PMU widely in the foreseeable future [9], [10]. Therefore, one would need to make use of the measurements from both the advanced PMUs and conventional supervisory control and data acquisition (SCADA) in order to reach the tradeoff between communication/implementation cost and estimation accuracy.

There have been three typical SE frameworks, namely, centralized, hierarchical and decentralized frameworks, which have appeared in the literature [11]–[13]. In the centralized SE framework, the global estimates can be generated in the estimation center by utilizing the measurement data from the entire power grid [14], [15]. Such a centralized framework might suffer from computation costs and/or communication burdens, and this is especially true when the power grid is scaled up [16], [17]. The hierarchical SE framework, on the other hand, disperses the computation/communication burdens by dividing the power grids into several areas and then generating the local estimates of each area which are finally fused in the estimation center [9], [10]. Note that hierarchical frameworks are very much dependent on the efficiency and reliability of the SE schemes (executed at the estimation center), which might not be always guaranteed for large-scale power systems [17], [18]. Compared to its centralized and hierarchical counterparts, the decentralized SE framework aims to further save limited computation/communication resources by dividing the power grid into several areas and conducting state estimation in each individual area based on the information from the local area as well as the edges which are shared with its neighboring areas.

In the past decades, the decentralized SE problem for power systems has attracted increasing research interest, see [3],

This work was supported in part by the National Natural Science Foundation of China under Grants 61933007, U21A2019, 62273005, 62273088, the Program of Shanghai Academic/Technology Research Leader of China under Grant 20XD1420100, the Hainan Province Science and Technology Special Fund of China under Grant ZDYF2022SHFZ105, the Natural Science Foundation of Anhui Province of China under Grant 2108085MA07, and the Alexander von Humboldt Foundation of Germany. *(Corresponding author: Bo Shen.)*

B. Qu is with the College of Automation Engineering, Shanghai University of Electric Power, Shanghai, 200090, China (e-mail: bogangqu@163.com).

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

B. 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).

H. 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).

H. Liu is with the Key Laboratory of Advanced Perception and Intelligent Control of High-end Equipment, Ministry of Education, Anhui Polytechnic University, Wuhu 241000, China; and is also with the School of Mathematics and Physics, Anhui Polytechnic University, Wuhu 241000, China (e-mail: hjliu1980@gmail.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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

[16]–[19] for some representative works. For instance, in [3], a fully distributed static SE algorithm has been designed for wide-area power systems based on the weighted-least-square (WLS) method. In [18], an efficient distributed SE algorithm has been developed with the aid of the WLS method, where the Gauss-Newton step is achieved without inner iteration. It should be noted that the WLS-based SE method is inherently a *static* one, and is therefore incapable of capturing the dynamic behaviors most likely caused by sudden load changes or the penetration of renewable power generations.

As with the increasing demand of monitoring dynamic behaviors of large-scale power systems, the so-called decentralized dynamic state estimation (DDSE) problem has createdgreat enthusiasm from researchers leading to the development of three mainstream DDSE-based schemes, namely, the model-decoupling-based scheme [8], [20], consensus-based scheme [21], [22] and maximum-a-posteriori-based scheme [23]. Briefly speaking, the model-decoupling-based scheme is computationally efficient at the cost of sacrificing certain estimation performance since the neighboring information of each synchronous generator is not effectively utilized. For a given area, although the consensus-based DDSE scheme improves the estimation efficiency, it may not be able to fully utilize the available data since the information of the indirectly connected areas is largely ignored. The distributed maximum-a-posteriori-based DSE algorithm, on the other hand, can achieve an adequate tradeoff between information utilization efficiency and state estimation accuracy since the information of entire power grid can be accessed through finite steps of iteration.

It is worth pointing out that most existing SE algorithms for power systems have been developed in the minimum mean-square-error (MMSE) sense, and the corresponding estimation performances are therefore prone to contaminated measurement data due to anomalies (e.g. outliers, communication impulses, quantization errors and instrument failures) [24]–[28]. To enhance the resistance against anomalies, some novel SE schemes have been proposed based on the minimum error entropy (MEE) criterion and the maximum correntropy (MC) criterion, see e.g. [29] for the MEE-based unscented Kalman filter (UKF), [30] for the MC-based extended Kalman filter and [31] for the novel Cauchy-kernel-based MC filter. In many situations, the MEE-based SE scheme outperforms the MC-based one, but the shift-invariant of the MEE criterion introduces a bias to make sure that the estimation error goes to zero [32]. In hopes of combining the merits of both MC and MEE criteria, the so-called minimum error entropy with fiducial points (MEEF) criterion has been proposed in [32] without having to add a bias after training.

Based on the above discussions, it can be found that there is a lack of adequate investigation into the SE problem for a wide-area power system with anomaly, and the main purpose of this paper is therefore to shorten such a gap. In doing so, the technical challenges can be identified as threefold: 1) determining how to implement the SE algorithm in a decentralized style to decrease the computation/communication cost; 2) determining how to improve data redundancy in the case of a decentralized SE framework where only the local measurement is available; and 3) determining how to reduce the negative influences of the measurement with anomalies on the SE performance in the estimator design.

In connection with the challenges identified above, the main contributions are highlighted as follows. 1) For the purpose of implementing the SE in a decentralized manner, the wide-area power system is divided into several non-overlapping areas, and the state of a given area is first estimated by using local measuring information and then updated by using the edge measurements which are shared with its neighboring areas. 2) To improve the data redundancy, an augmented model which involves both the local prediction and the local measurement is constructed by resorting to the unscented transformation approach as well as the statistical linearization approach. 3) A local MEEF-based UKF is proposed based on the MEEF criterion so as to enhance resistance against measurement with anomalies.

The remainder of this paper is outlined as follows. Section II formulates the decentralized model of power systems. In Section III, a brief review of the error entropy criterion and the correntropy criterion is first given, and then the MEEF criterion is introduced. Section IV describes the proposed anomaly-resistant decentralized SE algorithm. In Section V, simulation studies and detailed discussions are carried out on the IEEE 14-bus power system. Finally, some conclusions are drawn in Section VI.

Notation The notation used here is fairly standard except where otherwise stated. $\mathbb{R}^n$ represents the n-dimensional Euclidean space. The superscript "T" represents the transpose operation. diag$\{\cdots\}$ represents the block-diagonal matrix. $\mathbb{E}\{x\}$ is the expectation of the stochastic variable x. $[A]_{ij}$ represents the element at the i-th row and the j-th column of the matrix A. $\|\cdot\|$ stands for the Euclidean norm of a vector.

II. Problem Formulation

A. Preliminaries

In this paper, an undirected acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is used to reveal the connections of different areas of a wide-area power system. Specifically, $\mathcal{V} = \{1, 2, \ldots, M\}$ represents the set of areas, and $\mathcal{E} \in \mathcal{V} \times \mathcal{V}$ is the set of edges. The edge $(m, n) \in \mathcal{E}$ represents the m-th area of the wide-area power system and is interconnected with the n-th area. The set of neighbors of area $m \in \mathcal{V}$ plus the node itself are denoted by $\mathcal{N}_m = \{m \in \mathcal{V} : (m, n) \in \mathcal{E}\}$.

B. Power System Model

Consider a power system which contains N buses (an example is shown in Fig. 1), where the dynamic equation of the i-th bus with the quasi-steady state is described by [10], [33]

$$x_{i,\sigma+1} = A_i x_{i,\sigma} + B_i \bar{x}_i + w_{i,\sigma}. \tag{1}$$

Here, $x_{i,\sigma} \in \mathbb{R}^2$ represents the state vector of bus i $(i = 1, 2, \ldots, N)$ given by

$$x_{i,\sigma} \triangleq \begin{bmatrix} V_{i,\sigma} & \theta_{i,\sigma}. \end{bmatrix}^T$$

with $V_{i,\sigma}$ and $\theta_{i,\sigma}$ representing the voltage magnitude and the voltage phase angle of the i-th bus, respectively; $A_i \in \mathbb{R}^{2\times 2}$ is a known matrix reflecting how fast the transitions between

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

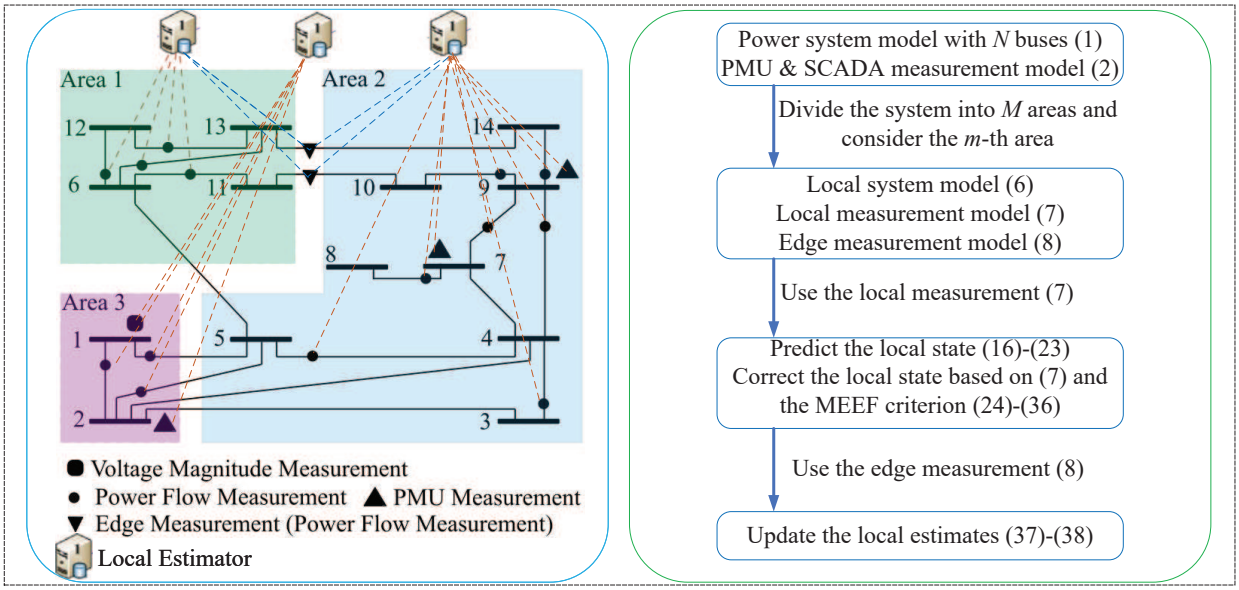

Fig. 1. Schematic view of the proposed decentralized SE for wide-area power systems.

states are; $\bar{x}_i$ is the expected steady state; B_i (commonly selected as $B_i = I - A_i$) is used to account for the trend behavior of the state trajectory; and $w_{i,\sigma} \in \mathbb{R}^2$ is Gaussian white noise with zero mean and covariance $W_{i,\sigma}$.

Remark 1: In this paper, the power system under consideration belongs to the widely used quasi-steady paradigm in which the loads and generators are assumed to operate stably without sudden changes in a short time interval [34], [35]. For the purpose of revealing the quasi-steady paradigm, a famous dynamic model (1) is used in which the bus magnitudes and phase angles are chosen as the state variables, the matrix A_i describes how fast the transitions between the states are and B_i represents the trend behavior of the state trajectory. It should be pointed out that the parameters of A_i and B_i can be identified in an online style and such a method is sufficient for the characterization of the quasi-steady behavior of the power systems [35].

C. SCADA and PMU Measurement Models

In this paper, the mixed SCADA and PMU measurements are used to monitor the power system (an example is shown in Fig. 1). The voltage magnitude of each bus and the power flows between two interconnected buses are measured by the SCADA units, and the voltage phasors of the selected buses are measured by the PMUs. It should be pointed out that all of the buses are installed with the SCADA units, and selected buses are installed with the PMUs since the implementation cost of the PMUs is much higher than that of the SCADA units [10].

The voltage magnitude of the i-th bus measured by the SCADA unit at time instant σ is expressed as

$$z^s_{i,\sigma} = C_s x_{i,\sigma} + v^s_{i,\sigma} \tag{2}$$

where $C_s = \begin{bmatrix} 1 & 0 \end{bmatrix}$, and $v^s_{i,\sigma}$ stands for the SCADA measurement noise that follows the Gaussian distribution with zero mean and covariance $V^s_{i,\sigma} > 0$.

The active and reactive power flows between the interconnected bus i and bus j measured by the SCADA unit at σ can be described by

$$\begin{aligned} P_{ij,\sigma} =& V^2_{i,\sigma}(g_{si} + g_{ij}) - V_{i,\sigma}V_{j,\sigma}\big(g_{ij}\cos(\theta_{i,\sigma} - \theta_{j,\sigma}) \\ &+ b_{ij}\sin(\theta_i - \theta_j)\big), \\ Q_{ij,\sigma} =& - V^2_{i,\sigma}(b_{si} + b_{ij}) - V_{i,\sigma}V_{j,\sigma}\big(g_{ij}\sin(\theta_{i,\sigma} - \theta_{j,\sigma}) \\ &- b_{ij}\cos(\theta_{i,\sigma} - \theta_{j,\sigma})\big) \end{aligned} \tag{3}$$

where $g_{si} + jb_{si}$ represents the admittance of the shunt branch connected to bus i, and $g_{ij} + jb_{ij}$ is the admittance of the series branch connecting buses i and j. Taking the measurement noise into account, a compact SCADA-based measurement model of the power flows is of the following form

$$z^s_{ij,\sigma} = h^s_{ij,\sigma}(x_{i,\sigma}, x_{j,\sigma}) + v^s_{ij,\sigma} \tag{4}$$

where

$$z^s_{ij,\sigma} \triangleq \begin{bmatrix} P_{ij,\sigma} & Q_{ij,\sigma} \end{bmatrix}^T,$$

$v^s_{ij,\sigma}$ is a Gaussian white noise on the SCADA unit with zero mean and covariance $V^s_{ij,\sigma} > 0$, and $h^s_{ij,\sigma}(\cdot)$ is determined by (3).

Letting the i-th bus be equipped with the PMU, the corresponding measurement model for the voltage phasor can be expressed as

$$z^p_{i,\sigma} = C_p x_{i,\sigma} + v^p_{i,\sigma} \tag{5}$$

where C_p is an identity matrix, and $v^p_{i,\sigma}$ is a zero-mean Gaussian white noise on the PMU with covariance $V^p_{i,\sigma} > 0$.

D. Decentralized Model

In this subsection, we divide the power system (1) into M non-overlapping areas and assume that the topology of such a divided system is acyclic. The system model and the corresponding measurement model for each area as well as the edge measurement model for two interconnected areas are given as follows:

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

1) Local System Model. The dynamic equation of the m-th area $(m = 1, 2, \cdots, M)$ can be given as

$$X_{m,\sigma+1} = \bar{A}_{m,\sigma} X_{m,\sigma} + \bar{B}_m \bar{X}_m + \bar{w}_{m,\sigma}, \quad m \in \mathcal{V}. \tag{6}$$

Similar to the descriptions of (1), $X_{m,\sigma}$ is the state vector of area m at time instant σ; $\bar{A}_m$ is the transition matrix of area m; $\bar{X}_m$ represents the expected steady state for area m; $\bar{B}_m = I - \bar{A}_m$ is related to the trend behavior of the state trajectory of area m; and the process noise is represented by $\bar{w}_{m,\sigma}$. The detailed descriptions of the variables in (6) are given as follows:

$$\begin{aligned}
X_{m,\sigma} &\triangleq \begin{bmatrix} x_{m_1,\sigma}^T & x_{m_2,\sigma}^T & \cdots & x_{m_{\tau_m},\sigma}^T \end{bmatrix}^T \in \mathbb{R}^{\mathbf{X}_m}, \\
\bar{A}_m &\triangleq \mathrm{diag}\{A_{m_1}, A_{m_2}, \cdots, A_{m_{\tau_m}}\} \in \mathbb{R}^{\mathbf{X}_m \times \mathbf{X}_m}, \\
\bar{B}_m &\triangleq \mathrm{diag}\{B_{m_1}, B_{m_2}, \cdots, B_{m_{\tau_m}}\} \in \mathbb{R}^{\mathbf{X}_m \times \mathbf{X}_m}, \\
\bar{X}_m &\triangleq \begin{bmatrix} \bar{x}_{m_1}^T & \bar{x}_{m_2}^T & \cdots & \bar{x}_{m_{\tau_m}}^T \end{bmatrix}^T \in \mathbb{R}^{\mathbf{X}_m}, \\
\bar{w}_{m,\sigma} &\triangleq \begin{bmatrix} w_{m_1,\sigma}^T & w_{m_2,\sigma}^T & \cdots & w_{m_{\tau_m},\sigma}^T \end{bmatrix}^T \in \mathbb{R}^{\mathbf{X}_m}
\end{aligned}$$

where, for $d = 1, 2, \ldots, \tau_m$, $x_{m_d,\sigma}$ represents the d-th state of the m-th area at time instant σ; A_{m_d} is the transition matrix of the d-th state of the m-th area; $\bar{x}_{m_d}$ represents the d-th expected steady state of area m; $B_{m_d} = I - A_{m_d}$ stands for the trend behavior of the d-th state trajectory of area m; and $w_{m_d,\sigma}$ is the process noise of the d-th state of the m-th area with zero mean and covariance matrix $W_{m_d,\sigma} > 0$. The covariance matrix of the process noise for the m-th area $\bar{w}_{m,\sigma}$ is $W_{m,\sigma} = \mathrm{diag}\{W_{m_1,\sigma}, W_{m_2,\sigma}, \cdots, W_{m_{\tau_m},\sigma}\}$. The number of the buses of the m-th area is denoted as τ_m and $\mathbf{X}_m = 2\tau_m$.

2) Local Measurement Model. The local measurement of the m-th area can be expressed as

$$z_{m,\sigma} = h_m(X_{m,\sigma}) + v_{m,\sigma}, \quad m \in \mathcal{V} \tag{7}$$

where

$$\begin{aligned}
z_{m,\sigma} \triangleq \big[&(z_{m_1,\sigma}^s)^T \ \cdots \ (z_{m_{\tau_m},\sigma}^s)^T \ \cdots (z_{m_p m_q,\sigma}^s)^T \\
&\cdots (z_{m_r,\sigma}^p)^T \ \cdots\big]^T \in \mathbb{R}^{\mathbf{Z}_m}, \\
v_{m,\sigma} \triangleq \big[&(v_{m_1,\sigma}^s)^T \ \cdots \ (v_{m_{\tau_m},\sigma}^s)^T \ \cdots (v_{m_p m_q,\sigma}^s)^T \\
&\cdots \ (v_{m_r,\sigma}^p)^T \ \cdots\big]^T \in \mathbb{R}^{\mathbf{Z}_m}.
\end{aligned}$$

Here, $z_{m_l,\sigma}^s$ $(l = 1, 2, \ldots, \tau_m)$ stands for the voltage magnitude of the l-th bus in the m-th area measured by the SCADA unit at time instant σ; $z_{m_p m_q,\sigma}^s$ represents the power flows between the interconnected bus p and bus q ($p, q \in \{1, 2, \ldots, \tau_m\}$ and $p \neq q$) in the m-th area measured by the SCADA unit at time instant σ; and $z_{m_r,\sigma}^p$ $(r \in \{1, 2, \ldots, \pi_m\})$ is the voltage phasor measurement of the r-th selected bus in the m-th area installed with PMU at time instant σ. The descriptions of every element of the measurement noise $v_{m,\sigma}$ are similar to that of the measurement vector $z_{m,\sigma}$ given above. The details of the nonlinear function $h_m(\cdot)$ rely on the states of the m-th area (i.e., $X_{m,\sigma}$) as well as the equations (2)–(5). The covariance matrix of the measurement noise $v_{m,\sigma}$ is $V_{m,\sigma} = \mathrm{diag}\{V_{m_1,\sigma}^s, \cdots, V_{m_{\tau_m},\sigma}^s, \cdots, V_{m_p m_q,\sigma}^s, \cdots, V_{m_r,\sigma}^p, \cdots\}$.

3) Edge Measurement Model. Since some buses of one area are often connected with other areas, the measurements related to these buses are treated as edge measurements in this paper. To be specific, for the m-th area, the edge measurement model can be represented by

$$\begin{aligned}
z_{mn,\sigma} =& h_{mn}(X_{m,\sigma}, X_{n,\sigma}) + v_{mn,\sigma}, \\
& m \in \mathcal{V}, n \in \mathcal{V}, m \neq n, n \in \mathcal{N}_m
\end{aligned} \tag{8}$$

where

$$\begin{aligned}
&h_{mn}(X_{m,\sigma}, X_{n,\sigma}) \\
&\triangleq \left[\cdots \ \left(h_{m_p n_q}^s(x_{m_p,\sigma}, x_{n_q,\sigma})\right)^T \ \cdots\right]^T \in \mathbb{R}^{\mathbf{Z}_{mn}}, \\
&v_{mn,\sigma} \\
&\triangleq \left[\cdots \ (v_{m_p n_q,\sigma}^s)^T \ \cdots\right]^T \in \mathbb{R}^{\mathbf{Z}_{mn}}.
\end{aligned}$$

Here, $h_{m_p n_q}^s(x_{m_p,\sigma}, x_{n_q,\sigma})$ represents the power flows between the interconnected bus p of area m and bus q of area n measured by SCADA unit, and the corresponding noise $v_{m_p n_q}^s(x_{m_p,\sigma}, x_{n_q,\sigma})$ follows Gaussian distribution with zero mean and covariance matrix $V_{m_p n_q,\sigma}^s > 0$. The details of $h_{mn}(X_{m,\sigma}, X_{n,\sigma})$ are dependent on the partition of the power system and the power flow measurements described in (4). The covariance matrix of the edge measurement noise $v_{mn,\sigma}$ is $V_{mn,\sigma} = \mathrm{diag}\{\ldots, V_{m_p n_q,\sigma}^s, \ldots\}$.

Remark 2: In actual power systems, the inducements of the measurement anomalies (e.g. stealthy attacks, communication impulses, instrument fails or packet losses) are numerous and the features of these anomalies (e.g. the type, occurring occasionality/probability, intermittency) are often unknown [7], [36]. As such, it is difficult and even impossible to reveal the measurement anomalies by using specific models, and this motivates us to improve the robustness of the estimator by adopting some novel methods.

E. Problem Statement

The aim of this paper is to develop an anomaly-resistant SE algorithm for power systems such that

1) the proposed SE algorithm is deployed in a decentralized manner by dividing the large-scale interconnected power system into several non-overlapping areas;
2) the decentralized state estimate is calculated iteratively by using the local measurement and edge measuring information; and
3) the resistance of the proposed algorithm against measurement with anomalies is enhanced by adopting the MEEF criterion.

III. Minimum Error Entropy with Fiducial Points Criterion

In this section, the error entropy criterion and the correntropy criterion are briefly reviewed, and then an error entropy with fiducial points criterion is developed.

A. Error Entropy Criterion

Let e be the error between the actual state x and the desired state y, (i.e. $e = x - y$). The Rényi's quadratic error entropy is defined as

$$H_{R,2}(e) \triangleq -\log \int p^2(e) de \tag{9}$$

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

where $p(e)$ is the probability density function (PDF) of the error variable e [37]. Based on the Parzen's window estimation method, the error PDF $p(e)$ can be estimated by drawing N samples $\{e_i\}_{i=1}^N$ from e, that is

$$\hat{p}(e) = \frac{1}{N}\sum_{i=1}^{N} G_\varsigma(e - e_i), \tag{10}$$

where $G_\varsigma(\cdot)$ is the kernel function and ς is the kernel bandwidth [38]. Without loss of generality, we choose the Gaussian kernel as the kernel function, i.e.,

$$G_\varsigma(e - e_i) = \exp\left(\frac{-(e-e_i)^2}{2\varsigma^2}\right). \tag{11}$$

Then, the Rényi's quadratic error entropy in (9) can be approximated as

$$\begin{aligned} H_{R,2}(e) &\approx -\log \int \hat{p}^2(e)de \\ &\triangleq -\log\big(IP(e)\big) \end{aligned} \tag{12}$$

with

$$IP(e) \triangleq \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_\varsigma(e_i - e_j)$$

representing the information potential of the error e [39]. Obviously, the MEE criterion is actually equivalent to the maximization of the information potential $IP(e)$.

B. Correntropy Criterion

Letting X and Y be stochastic variables, the correntropy between these two variables is defined as

$$C(X,Y) = \mathbb{E}\{K_\varsigma(X,Y)\} = \int\int K_\varsigma(x,y) f_{XY}(x,y)dxdy \tag{13}$$

where $K_\varsigma(\cdot,\cdot)$ denotes the Mercer kernel function and $f_{XY}(x,y)$ is the joint PDF. Without loss of generality, we still choose the Gaussian kernel in (11) as the kernel function. Note that the distribution of $f_{XY}(x,y)$ is usually unknown and we have to draw N samples $\{x_i, y_i\}_{i=1}^N$ from $f_{XY}(x,y)$ to estimate $C(X,Y)$, i.e.,

$$\hat{C}(X,Y) = \frac{1}{N}\sum_{i=1}^{N} G_\varsigma(x_i - y_i). \tag{14}$$

As such, the so-called MC criterion is actually equivalent to the maximization of the cost function given in (14).

C. Minimum Error Entropy with Fiducial Points Criterion

As discussed in [32], [37], the MEE is shift-invariant and one has to add a bias to guarantee that the estimation error goes towards zero. In order to overcome such a shortcoming, an improved approach called the MEEF criterion has been proposed in [32]. Specifically, the MEEF criterion is established by combining MEE and MC criteria through a cost function

$$J(E) = \frac{\mu}{N}\sum_{i=1}^{N} G_{\varsigma_1}(e_i) + \frac{1-\mu}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{\varsigma_2}(e_i - e_j) \tag{15}$$

where E is the error variable; $\mu \in (0,1)$ is the weight coefficient; ς_1 and ς_2 are the bandwidths of the Gaussian kernel functions $G_{\varsigma_1}(\cdot)$ and $G_{\varsigma_2}(\cdot)$, respectively; e_i (e_j) represents the i-th (j-th) sample drawn from E; and N is the number of samples.

Remark 3: Due to the unique merits in enhancing robustness against anomalies, the MEE and MC criteria have attracted a surge of research enthusiasm and some representative results have appeared, see e.g. [29]–[31]. For example, a novel MC-based filter has been designed in [31] where the inherent shortages of Gaussian kernels have been overcome with the aid of the Cauchy kernel. It is worth pointing out that the shift-invariant property of the MEE criterion makes it difficult to guarantee that the mean of the error PDF tends to zero [40]. As such, the MEEF criterion has been proposed in [40] where the MC term is used to fix the peak of error PDF at origin and the MEE term is adopted to minimize the error entropy [32], [41]. In this paper, the MEEF criterion will be used in the subsequent anomaly-resistant state estimator design.

IV. Anomaly-Resistant Decentralized State Estimator Design

In this section, we aim to study the anomaly-resistant SE problem for wide-area power systems in the presence of measurement with anomalies. First, a local UKF-based state estimation scheme is developed by using the local sensing data, where the MEEF criterion is utilized in the cost function design to enhance the resistance of the local state estimator against the measurement with anomalies. Then, the local estimate is updated by exploiting the edge measuring information. The schematic view of the proposed decentralized SE scheme is outlined in Fig. 1.

A. Design of the MEEF-based Local State Estimator

1) Prediction: For the m-th $(m = 1, 2, \ldots, M)$ area, suppose that the local estimate $\hat{X}_{m,\sigma-1|\sigma-1} \in \mathbb{R}^{\mathbf{X}_m}$ and its covariance matrix $P^{xx}_{m,\sigma-1|\sigma-1} \in \mathbb{R}^{\mathbf{X}_m \times \mathbf{X}_m}$ are available at $\sigma - 1$. Then, the $4\tau_m + 1$ sigma points can be generated via

$$\chi^j_{m,\sigma-1|\sigma-1} = \begin{cases} \hat{X}_{m,\sigma-1|\sigma-1}, \quad j = 0 \\ \hat{X}_{m,\sigma-1|\sigma-1} \\ \quad + \big(\sqrt{(2\tau_m + \lambda_m)P^{xx}_{m,\sigma-1|\sigma-1}}\big)_j, \\ \qquad\qquad j = 1, \ldots, 2\tau_m \\ \hat{X}_{m,\sigma-1|\sigma-1} \\ \quad - \big(\sqrt{(2\tau_m + \lambda_m)P^{xx}_{m,\sigma-1|\sigma-1}}\big)_j, \\ \qquad\qquad j = 2\tau_m + 1, \ldots, 4\tau_m \end{cases} \tag{16}$$

where $\big(\sqrt{(2\tau_m + \lambda_m)P^{xx}_{m,\sigma-1|\sigma-1}}\big)_j$ represents the j-th column of the square root of the matrix $(2\tau_m + \lambda_m)P^{xx}_{m,\sigma-1|\sigma-1}$ in the sense of Cholesky decomposition, i.e.,

$$\begin{aligned} &(2\tau_m + \lambda_m)P^{xx}_{m,\sigma-1|\sigma-1} \\ =&\big(\sqrt{(2\tau_m + \lambda_m)P^{xx}_{m,\sigma-1|\sigma-1}}\big)\big(\sqrt{(2\tau_m + \lambda_m)P^{xx}_{m,\sigma-1|\sigma-1}}\big)^T, \end{aligned}$$

and $\lambda_m \triangleq \alpha_m^2(2\tau_m + \kappa_m) - 2\tau_m$ stands for the scalar parameter with $0 \le \alpha_m \le 1$ and $\kappa_m \ge 0$.

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

Having obtained the sigma points, we now propagate each sigma point through the state-transition function (6) to generate a new set of transformed sigma points $\chi^j_{m,\sigma|\sigma-1}$ with the following form:

$$\chi^j_{m,\sigma|\sigma-1} = \bar{A}_{m,\sigma}\chi^j_{m,\sigma-1|\sigma-1} + \bar{B}_m\bar{X}_m, \quad j = 0, 1, \ldots, 4\tau_m. \tag{17}$$

Then, the prediction state $\hat{X}_{m,\sigma|\sigma-1}$ and the corresponding prediction error covariance $P^{xx}_{m,\sigma|\sigma-1}$ can be, respectively, calculated via

$$\hat{X}_{m,\sigma|\sigma-1} = \sum_{j=0}^{4\tau_m} \omega^j_m \chi^j_{m,\sigma|\sigma-1} \tag{18}$$

and

$$P^{xx}_{m,\sigma|\sigma-1} = \sum_{j=0}^{4\tau_m} \theta^j_m \tilde{\chi}^j_{m,\sigma|\sigma-1}(\tilde{\chi}^j_{m,\sigma|\sigma-1})^T + W_{m,\sigma} \tag{19}$$

where $\tilde{\chi}^j_{m,\sigma|\sigma-1} \triangleq (\chi^j_{m,\sigma|\sigma-1} - \hat{X}_{m,\sigma|\sigma-1})$, ω^j_m and θ^j_m are the weighted coefficients with the following form:

$$\omega^j_m \triangleq \begin{cases} \dfrac{\lambda_m}{2\tau_m + \lambda_m}, & j = 0 \\ \dfrac{1}{2(2\tau_m + \lambda_m)}, & j = 1, 2, \ldots, 4\tau_m \end{cases}$$

and

$$\theta^j_m \triangleq \begin{cases} \dfrac{\lambda_m}{2\tau_m + \lambda_m} + (1 - \alpha^2_m + \beta_m), & j = 0 \\ \dfrac{1}{2(2\tau_m + \lambda_m)}, & j = 1, 2, \ldots, 4\tau_m \end{cases}$$

with $\beta_m > 0$.

Similarly, the predicted sigma points, which are mapped through the measurement function $h_m(\cdot)$ in (7), are given as

$$\mathcal{Z}^j_{m,\sigma|\sigma-1} = h_m(\chi^j_{m,\sigma|\sigma-1}), \quad j = 0, 1, \ldots, 4\tau_m + 1. \tag{20}$$

Then, we can obtain the prediction of the local measurement $z_{m,\sigma}$ by the following weighted average operation:

$$\hat{z}_{m,\sigma|\sigma-1} = \sum_{j=0}^{4\tau_m} \omega^j_m \mathcal{Z}^j_{m,\sigma|\sigma-1}. \tag{21}$$

By denoting $\tilde{\mathcal{Z}}^j_{m,\sigma|\sigma-1} \triangleq (\mathcal{Z}^j_{m,\sigma|\sigma-1} - \hat{z}_{m,\sigma|\sigma-1})$, the measurement prediction error covariance matrix $P^{zz}_{m,\sigma|\sigma-1}$ and the state-measurement prediction error cross-covariance matrix $P^{xz}_{m,\sigma|\sigma-1}$ can be computed, respectively, as

$$P^{zz}_{m,\sigma|\sigma-1} = \sum_{j=0}^{4\tau_m} \theta^j_m \tilde{\mathcal{Z}}^j_{m,\sigma|\sigma-1}(\tilde{\mathcal{Z}}^j_{m,\sigma|\sigma-1})^T + V_{m,\sigma} \tag{22}$$

and

$$P^{xz}_{m,\sigma|\sigma-1} = \sum_{j=0}^{4\tau_m} \theta^j_m \tilde{\chi}^j_{m,\sigma|\sigma-1}(\tilde{\mathcal{Z}}^j_{m,\sigma|\sigma-1})^T. \tag{23}$$

2) Update: Based on the statistical linearization approach proposed in [42], the local measurement function (7) can be linearized around $\hat{X}_{m,\sigma|\sigma-1}$ as

$$\begin{aligned} z_{m,\sigma} =& H_{m,\sigma}(X_{m,\sigma} - \hat{X}_{m,\sigma|\sigma-1}) \\ &+ h_m(\hat{X}_{m,\sigma|\sigma-1}) + v_{m,\sigma} + \epsilon_{m,\sigma} \end{aligned} \tag{24}$$

with $H_{m,\sigma} \triangleq (P^{xz}_{m,\sigma|\sigma-1})^T(P^{xx}_{m,\sigma|\sigma-1})^{-1}$ being the pseudo measurement matrix, and $\epsilon_{m,\sigma}$ being the statistical linearization error with zero mean and covariance matrix

$$\begin{aligned} \bar{V}_{m,\sigma} \triangleq& \mathbb{E}\{\epsilon_{m,\sigma}\epsilon^T_{m,\sigma}\} \\ =& P^{zz}_{m,\sigma|\sigma-1} - (P^{xz}_{m,\sigma|\sigma-1})^T(P^{xx}_{m,\sigma|\sigma-1})^{-1}P^{xz}_{m,\sigma|\sigma-1} \end{aligned}$$

where the covariance matrices $P^{xx}_{m,\sigma|\sigma-1}$, $P^{zz}_{m,\sigma|\sigma-1}$ and $P^{xz}_{m,\sigma|\sigma-1}$ are calculated via (19), (22) and (23), respectively.

By integrating local prediction and local measurement into a compact form, we obtain the following augmented model

$$\bar{z}_{m,\sigma} = \bar{H}_{m,\sigma}x_{m,\sigma} + \bar{e}_{m,\sigma} \tag{25}$$

where

$$\bar{z}_{m,\sigma} \triangleq \begin{bmatrix} \hat{X}_{m,\sigma|\sigma-1} \\ z_{m,\sigma} + H_{m,\sigma}\hat{X}_{m,\sigma|\sigma-1} - \hat{z}_{m,\sigma|\sigma-1} \end{bmatrix} \in \mathbb{R}^{L_m},$$

$$\bar{H}_{m,\sigma} \triangleq \begin{bmatrix} I \\ H_{m,\sigma} \end{bmatrix} \in \mathbb{R}^{L_m}, \quad \bar{e}_{m,\sigma} \triangleq \begin{bmatrix} -\tilde{X}_{m,\sigma|\sigma-1} \\ v_{m,\sigma} + \epsilon_{m,\sigma} \end{bmatrix} \in \mathbb{R}^{L_m}$$

with $\tilde{X}_{m,\sigma|\sigma-1} \triangleq X_{m,\sigma} - \hat{X}_{m,\sigma|\sigma-1}$ and $L_m = \mathbf{X}_m + \mathbf{Z}_m$. The covariance matrix of $\bar{e}_{m,\sigma}$ is

$$\begin{aligned} \mathbb{E}\{\bar{e}_{m,\sigma}\bar{e}^T_{m,\sigma}\} &= \begin{bmatrix} P^{xx}_{m,\sigma|\sigma-1} & 0 \\ 0 & V_{m,\sigma} + \bar{V}_{m,\sigma} \end{bmatrix} \\ &= \begin{bmatrix} S^P_{m,\sigma}(S^P_{m,\sigma})^T & 0 \\ 0 & S^V_{m,\sigma}(S^V_{m,\sigma})^T \end{bmatrix} \\ &= S_{m,\sigma}S^T_{m,\sigma} \end{aligned} \tag{26}$$

where $S^P_{m,\sigma} \in \mathbb{R}^{\mathbf{X}_m}$, $S^V_{m,\sigma} \in \mathbb{R}^{\mathbf{Z}_m}$ and $S_{m,\sigma} \in \mathbb{R}^{L_m}$ are all calculated with the aid of Cholesky decomposition.

By multiplying $S^{-1}_{m,\sigma}$ on both sides of (25), we have

$$y_{m,\sigma} = F_{m,\sigma}X_{m,\sigma} + \xi_{m,\sigma} \tag{27}$$

where

$$\begin{aligned} y_{m,\sigma} \triangleq& S^{-1}_{m,\sigma}\bar{z}_{m,\sigma}, \\ F_{m,\sigma} \triangleq& S^{-1}_{m,\sigma}\bar{H}_{m,\sigma}, \\ \xi_{m,\sigma} \triangleq& S^{-1}_{m,\sigma}\bar{e}_{m,\sigma} \end{aligned} \tag{28}$$

with

$$\begin{aligned} y_{m,\sigma} &\triangleq \begin{bmatrix} y^1_{m,\sigma} & y^2_{m,\sigma} & \cdots & y^{L_m}_{m,\sigma} \end{bmatrix}^T \in \mathbb{R}^{L_m}, \\ F_{m,\sigma} &\triangleq \begin{bmatrix} F^1_{m,\sigma} & F^2_{m,\sigma} & \cdots & F^{L_m}_{m,\sigma} \end{bmatrix}^T \in \mathbb{R}^{L_m \times L_m}, \\ \xi_{m,\sigma} &\triangleq \begin{bmatrix} \xi^1_{m,\sigma} & \xi^2_{m,\sigma} & \cdots & \xi^{L_m}_{m,\sigma} \end{bmatrix}^T \in \mathbb{R}^{L_m}. \end{aligned}$$

Moreover, it can be seen that $\mathbb{E}\{\xi_{m,\sigma}\xi^T_{m,\sigma}\} = I$.

Remark 4: It has been proven in [43] that, for an arbitrary nonlinear function, the results obtained from the statistical linearization are equivalent to those from the unscented transformation. As such, for the purpose of improving data redundancy, an effective way is to construct an augmented model (i.e. (25)) to process the local prediction and local

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

measurement simultaneously with the aid of unscented transformation approach and statistical linearization approach.

According to the MEEF criterion given in (15), the cost function of the MEEF-based local estimator can be written as

$$\begin{aligned}J(x_{m,\sigma}) =&\frac{\lambda}{L_m}\sum_{i=1}^{L_m} G_{\varsigma_1}(\xi_{m,\sigma}^i)\\ &+\frac{1-\lambda}{L_m}\sum_{i=1}^{L_m}\sum_{j=1}^{L_m} G_{\varsigma_2}(\xi_{m,\sigma}^i-\xi_{m,\sigma}^j).\end{aligned}\tag{29}$$

By maximizing the cost function (29), we can obtain the local estimate of $X_{m,\sigma}$. To be specific, taking the partial derivative of $J(X_{m,\sigma})$ with respect to $X_{m,\sigma}$ and letting the derivative be zero, we have

$$\begin{aligned}\frac{\partial J(X_{m,\sigma})}{\partial X_{m,\sigma}} =&\frac{\lambda}{L_m}\sum_{i=1}^{L_m}\Big(G_{\varsigma_1}(\xi_{m,\sigma}^i)F_{m,\sigma}^i(y_{m,\sigma}^i-F_{m,\sigma}^{i^T}X_{m,\sigma})\Big)\\ &+\frac{1-\lambda}{L_m}\sum_{i=1}^{L_m}\sum_{j=1}^{L_m}\Big((\xi_{m,\sigma}^i-\xi_{m,\sigma}^j)\\ &\times G_{\varsigma_2}(\xi_{m,\sigma}^i-\xi_{m,\sigma}^j)(F_{m,\sigma}^i-F_{m,\sigma}^j)\Big)\\ =&\lambda F_{m,\sigma}^T\Pi_{m,\sigma}\xi_{m,\sigma}\\ &+\frac{1-\lambda}{L_m^2\varsigma_2^2}(F_{m,\sigma}^T\Omega_{m,\sigma}\xi_{m,\sigma}-F_{m,\sigma}^T\Gamma_{m,\sigma}\xi_{m,\sigma})\\ =&F_{m,\sigma}^T\Phi_{m,\sigma}\xi_{m,\sigma}\\ =&0\end{aligned}\tag{30}$$

where

$$\Phi_{m,\sigma}\triangleq\lambda\Pi_{m,\sigma}+\frac{1-\lambda}{L_m^2\varsigma_2^2}(\Omega_{m,\sigma}-\Gamma_{m,\sigma})$$

with

$$\begin{aligned}\Pi_{m,\sigma}&\triangleq\operatorname{diag}\big\{G_{\varsigma_1}(\xi_{m,\sigma}^1),\ldots,G_{\varsigma_1}(\xi_{m,\sigma}^{L_m})\big\},\\ \Omega_{m,\sigma}&\triangleq\operatorname{diag}\Big\{\sum_{i=1}^{L_m}G_{\varsigma_2}(\xi_{m,\sigma}^1-\xi_{m,\sigma}^i),\\ &\qquad\ldots,\sum_{i=1}^{L_m}G_{\varsigma_2}(\xi_{m,\sigma}^{L_m}-\xi_{m,\sigma}^i)\Big\},\\ [\Gamma_{m,\sigma}]_{ij}&\triangleq G_{\varsigma_2}(\xi_{m,\sigma}^i-\xi_{m,\sigma}^j).\end{aligned}$$

Based on the fixed-point iteration technique [41], the solution of (30) can be expressed as

$$\begin{aligned}(\check{X}_{m,\sigma|\sigma})_t =&\Big(F_{m,\sigma}^T(\Phi_{m,\sigma})_{t-1}F_{m,\sigma}\Big)^{-1}\\ &\times\Big(F_{m,\sigma}^T(\Phi_{m,\sigma})_{t-1}y_{m,\sigma}\Big)\end{aligned}\tag{31}$$

where t represents the t-th fixed-point iteration and $\check{X}_{m,\sigma|\sigma}$ stands for the local estimate of $X_{m,\sigma}$.

Remark 5: Note that, in the fixed-point iteration, the term $X_{m,\sigma}$ in (27) can be replaced by $(\check{X}_{m,\sigma|\sigma})_{t-1}$ (i.e. $\xi_{m,\sigma}=y_{m,\sigma}-F_{m,\sigma}(\check{X}_{m,\sigma|\sigma})_{t-1}$) with initial value $(\check{X}_{m,\sigma|\sigma})_0=\hat{X}_{m,\sigma|\sigma-1}$. Recalling the definitions of $\bar{e}_{m,\sigma}$, $\xi_{m,\sigma}$ and $\Phi_{m,\sigma}$, we can find that the calculation of $(\Phi_{m,\sigma})_{t-1}$ in (31) actually relies on $(\check{X}_{m,\sigma|\sigma})_{t-1}$, which means that the iterations between $(\check{X}_{m,\sigma|\sigma})_t$ and $(\check{X}_{m,\sigma|\sigma})_{t-1}$ are actually established by $(\Phi_{m,\sigma})_{t-1}$.

Next, we rewrite $\Phi_{m,\sigma}$ as

$$\Phi_{m,\sigma}\triangleq\begin{bmatrix}\Phi_{m,\sigma}^{xx} & \Phi_{m,\sigma}^{zx}\\ \Phi_{m,\sigma}^{xz} & \Phi_{m,\sigma}^{zz}\end{bmatrix}\in\mathbb{R}^{(\mathbf{X}_m+\mathbf{Z}_m)\times(\mathbf{X}_m+\mathbf{Z}_m)}\tag{32}$$

with

$$\begin{aligned}\Phi_{m,\sigma}^{xx}\triangleq&\lambda\operatorname{diag}\big\{G_{\varsigma_1}(\xi_{m,\sigma}^1),\ldots,G_{\varsigma_1}(\xi_{m,\sigma}^{\mathbf{X}_m})\big\}\\ &+\frac{1-\lambda}{L_m^2\varsigma_2^2}\Big((\Omega_{m,\sigma}^{ij})_{\mathbf{X}_m\times\mathbf{X}_m}-(\Gamma_{m,\sigma}^{ij})_{\mathbf{X}_m\times\mathbf{X}_m}\Big)\\ &(i=1,2,\ldots,\mathbf{X}_m;\ j=1,2,\ldots,\mathbf{X}_m),\\ \Phi_{m,\sigma}^{zx}\triangleq&\frac{1-\lambda}{L_m^2\varsigma_2^2}\Big((\Omega_{m,\sigma}^{ij})_{\mathbf{X}_m\times\mathbf{Z}_m}-(\Gamma_{m,\sigma}^{ij})_{\mathbf{X}_m\times\mathbf{Z}_m}\Big)\\ &(i=1,2,\ldots,\mathbf{X}_m;\\ &\quad j=\mathbf{X}_m+1,\mathbf{X}_m+2,\ldots,\mathbf{X}_m+\mathbf{Z}_m),\\ \Phi_{m,\sigma}^{xz}\triangleq&\lambda\operatorname{diag}\big\{G_{\varsigma_1}(\xi_{m,\sigma}^{\mathbf{X}_m+1}),\ldots,G_{\varsigma_1}(\xi_{m,\sigma}^{\mathbf{X}_m+\mathbf{Z}_m})\big\}\\ &+\frac{1-\lambda}{L_m^2\varsigma_2^2}\Big((\Omega_{m,\sigma}^{ij})_{\mathbf{Z}_m\times\mathbf{X}_m}-(\Gamma_{m,\sigma}^{ij})_{\mathbf{Z}_m\times\mathbf{X}_m}\Big)\\ &(i=\mathbf{X}_m+1,\mathbf{X}_m+2,\ldots,\mathbf{X}_m+\mathbf{Z}_m;\\ &\quad j=1,2,\ldots,\mathbf{Z}_m),\\ \Phi_{m,\sigma}^{zz}\triangleq&\frac{1-\lambda}{L_m^2\varsigma_2^2}\Big((\Omega_{m,\sigma}^{ij})_{\mathbf{Z}_m\times\mathbf{Z}_m}-(\Gamma_{m,\sigma}^{ij})_{\mathbf{Z}_m\times\mathbf{Z}_m}\Big)\\ &(i=\mathbf{X}_m+1,\mathbf{X}_m+2,\ldots,\mathbf{X}_m+\mathbf{Z}_m;\\ &\quad j=\mathbf{X}_m+1,\mathbf{X}_m+2,\ldots,\mathbf{X}_m+\mathbf{Z}_m),\end{aligned}$$

and then reformulate (31) as

$$(\check{X}_{m,\sigma|\sigma})_t=(\Delta_1+\Delta_2\Delta_3)^{-1}(\Delta_1\hat{X}_{m,\sigma|\sigma-1}+\Delta_2 y_{m,\sigma})\tag{33}$$

where

$$\begin{aligned}\Delta_1\triangleq&\Big((S_{m,\sigma}^P)^{-1^T}(\Phi_{m,\sigma}^{xx})_{t-1}+H_{m,\sigma}^T(S_{m,\sigma}^V)^{-1^T}\\ &\times(\Phi_{m,\sigma}^{xz})_{t-1}\Big)(S_{m,\sigma}^P)^{-1},\\ \Delta_2\triangleq&\Big((S_{m,\sigma}^P)^{-1^T}(\Phi_{m,\sigma}^{zx})_{t-1}+H_{m,\sigma}^T(S_{m,\sigma}^V)^{-1^T}\\ &\times(\Phi_{m,\sigma}^{zz})_{t-1}\Big)(S_{m,\sigma}^V)^{-1},\\ \Delta_3\triangleq&H_{m,\sigma}.\end{aligned}$$

By resorting to the matrix inversion lemma, (33) can be rearranged in the following recursive form:

$$(\check{X}_{m,\sigma|\sigma})_t=\hat{X}_{m,\sigma|\sigma-1}+K_{m,\sigma}\Big(z_{m,\sigma}-h_m(\hat{X}_{m,\sigma|\sigma-1})\Big)\tag{34}$$

with

$$\begin{aligned}K_{m,\sigma}=&\Big(\bar{P}_{m,\sigma|\sigma-1}^{xx}+\bar{H}_{m,\sigma}\bar{P}_{m,\sigma|\sigma-1}^{xz}+(\bar{P}_{m,\sigma|\sigma-1}^{xz}\\ &+\bar{H}_{m,\sigma}^T\bar{R}_{m,\sigma})\bar{H}_{m,\sigma}\Big)^{-1}(\bar{P}_{m,\sigma|\sigma-1}^{zx}+\bar{H}_{m,\sigma}\bar{R}_{m,\sigma})\end{aligned}$$

where

$$\begin{aligned}\bar{P}_{m,\sigma|\sigma-1}^{xx}&\triangleq\big((S_{m,\sigma}^P)^{-1}\big)^T(\Phi_{m,\sigma}^{xx})_{t-1}(S_{m,\sigma}^P)^{-1},\\ \bar{P}_{m,\sigma|\sigma-1}^{xz}&\triangleq\big((S_{m,\sigma}^V)^{-1}\big)^T(\Phi_{m,\sigma}^{xz})_{t-1}(S_{m,\sigma}^V)^{-1},\\ \bar{P}_{m,\sigma|\sigma-1}^{zx}&\triangleq\big((S_{m,\sigma}^P)^{-1}\big)^T(\Phi_{m,\sigma}^{zx})_{t-1}(S_{m,\sigma}^P)^{-1},\\ \bar{R}_{m,\sigma}&\triangleq\big((S_{m,\sigma}^V)^{-1}\big)^T(\Phi_{m,\sigma}^{zz})_{t-1}(S_{m,\sigma}^V)^{-1}.\end{aligned}$$

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

Then, the local estimation covariance of $\check{X}_{m,\sigma|\sigma}$ can be expressed as

$$\begin{aligned}\check{P}^{xx}_{m,\sigma|\sigma} =&(I - K_{m,\sigma}\bar{H}_{m,\sigma})P^{xx}_{m,\sigma-1|\sigma-1}(I - K_{m,\sigma}\bar{H}_{m,\sigma})^T\\ &+ K_{m,\sigma}(V_{m,\sigma} + \bar{V}_{m,\sigma})K^T_{m,\sigma}.\end{aligned} \tag{35}$$

For the convenience of the subsequent decentralized state estimator design, we transform the local estimate $\check{X}_{m,\sigma}$ and local estimation error covariance $\check{P}^{xx}_{m,\sigma|\sigma}$ into the information forms, i.e.

$$\begin{aligned}\check{\eta}_{m,\sigma} =&L_{m,\sigma}\check{X}_{m,\sigma|\sigma},\\ \check{L}_{m,\sigma} =&(\check{P}^{xx}_{m,\sigma|\sigma})^{-1}\end{aligned} \tag{36}$$

where $\eta_{m,\sigma}$ and $L_{m,\sigma}$ represent the information vector and the information matrix, respectively.

B. Design of Decentralized State Estimator

In this subsection, the maximum-a-posteriori state estimation method in [23] is extended to update the local estimate recursively by using the edge sensing information. For area $n \in \mathcal{N}_m$, the pseudo measurement matrix of the edge measurement $z_{mn,\sigma}$ can be expressed as

$$H_{mn,\sigma} \triangleq (P^{xz}_{mn,\sigma|\sigma-1})^T(P^{xx}_{m,\sigma|\sigma-1})^{-1} \tag{37}$$

where

$$\begin{aligned}P^{xz}_{mn,\sigma|\sigma-1} =&\sum_{j=0}^{4\tau_m}\theta^j_m(\chi^j_{m,\sigma|\sigma-1} - \hat{X}_{m,\sigma|\sigma-1})\\ &\times\left(h_{mn}(\chi^j_{m,\sigma|\sigma-1}, \hat{X}_{n,\sigma|\sigma-1}) - \hat{z}_{mn,\sigma|\sigma-1}\right)\end{aligned}$$

with

$$\hat{z}_{mn,\sigma|\sigma-1} = \sum_{j=0}^{4\tau_m}\omega^j_m h_{mn}(\chi^j_{m,\sigma|\sigma-1}, \hat{X}_{n,\sigma|\sigma-1}).$$

Then, for area $m \in \mathcal{V}$, the estimate and the estimation error covariance of $X_{m,\sigma}$ are of the following form

$$\begin{aligned}\hat{X}_{m,\sigma|\sigma} =&P^{xx}_{m,\sigma|\sigma}(\check{\eta}_{m,\sigma} + \sum_{n\in\mathcal{N}_m} H^T_{mn,\sigma}\bar{V}^{-1}_{nm,\sigma}\tilde{z}_{mn,\sigma}),\\ P^{xx}_{m,\sigma|\sigma} =&(\check{L}_{m,\sigma} + \sum_{n\in\mathcal{N}_m} H^T_{mn,\sigma}\bar{V}^{-1}_{nm,\sigma}H_{mn,\sigma})^{-1}\end{aligned} \tag{38}$$

where

$$\begin{aligned}\bar{V}_{nm,\sigma} \triangleq&V_{mn,\sigma} + H_{nm,\sigma}\check{L}^{-1}_{n,\sigma}H^T_{nm,\sigma},\\ \tilde{z}_{nm,\sigma} \triangleq&z_{mn,\sigma} - h_{mn}(\hat{X}_{m,\sigma|\sigma-1}, \hat{X}_{n,\sigma|\sigma-1}) + H_{mn,\sigma}\hat{X}_{m,\sigma|\sigma-1}\\ &+ H_{nm,\sigma}\hat{X}_{n,\sigma|\sigma-1} - H_{nm,\sigma}\check{L}^{-1}_{n,\sigma}\eta_{n,\sigma}.\end{aligned}$$

As discussed in [23], for each area $m \in \mathcal{V}$, the decentralized SE algorithm will converge at each σ after finite iterations. As such, an inner loop is added in the update of the local estimate. In summary, the pseudo code of the proposed anomaly-resistant decentralized SE for wide-area power systems is outlined in Algorithm 1.

Algorithm 1 Anomaly-Resistant decentralized SE algorithm for wide-area power systems.

Initialization: For each area $m \in \mathcal{V}$, initialize the parameters $\hat{X}_{m,0|0}$ and $P^{xx}_{m,0|0}$ and choose the weighted coefficients μ, ω^j_m and θ^j_m $(j = 0, 1, \ldots, 4\tau_m)$. In addition, set the kernel widths ς_1 and ς_2 and select the positive scalars ζ^* and t^*.

Recursion:

for $\sigma = 1, 2, \ldots$ **do**

Compute the sigma points $\chi^j_{m,\sigma-1|\sigma-1}$ $(j = 0, 1, \ldots, 4\tau_m)$ with known $\hat{X}_{m,\sigma-1|\sigma-1}$ via (16).

Compute the local prediction state $\hat{X}_{m,\sigma|\sigma-1}$ and the corresponding covariance matrices $P^{xx}_{m,\sigma|\sigma-1}$, $P^{zz}_{m,\sigma|\sigma-1}$ and $P^{xz}_{m,\sigma|\sigma-1}$ in terms of (18)-(23).

Linearize the local measurement function (7) around $\hat{X}_{m,\sigma|\sigma-1}$ via (24).

Construct the augmented model (25) and use the Cholesky decomposition to calculate $S^P_{m,\sigma}$, $S^V_{m,\sigma}$ and $S_{m,\sigma}$ in (26).

Compute $y_{m,\sigma}$, $F_{m,\sigma}$ and $\xi_{m,\sigma}$ via (28).

Set $t = 0$ and let $(\check{X}_{m,\sigma|\sigma})_t$ be the estimate at the fixed-point iterative step t. Denote $(\check{X}_{m,\sigma|\sigma})_0 = \hat{X}_{m,\sigma|\sigma-1}$ as the initial iteration value.

while $\zeta > \zeta^*$ or $t < t^*$ **do**

Update the iterative step $t = t + 1$.

Compute $(\check{X}_{m,\sigma|\sigma})_t$ via (34).

Compute the iteration accuracy

$$\zeta = \frac{\|(\check{X}_{m,\sigma|\sigma})_t - (\check{X}_{m,\sigma|\sigma})_{t-1}\|}{\|(\check{X}_{m,\sigma|\sigma})_{t-1}\|}.$$

end while

Denote $\check{X}_{m,\sigma|\sigma} = (\check{X}_{m,\sigma|\sigma})_t$ and update $\check{P}^{xx}_{m,\sigma|\sigma}$ via (35).

Calculate $\check{\eta}_{m,\sigma}$ and $\check{L}_{m,\sigma}$ via (36).

For each area $n \in \mathcal{N}_m$, area m receives $\hat{X}_{n,\sigma|\sigma-1}$ and calculates $H_{mn,\sigma}$ via (37).

Inner loop:

Set $h = 0$, for each area $n \in \mathcal{N}_m$, area m computes

$$\begin{aligned}(\upsilon^n_{m,\sigma})_0 =&H_{mn,\sigma}\check{X}_{m,\sigma|\sigma},\\ (\Upsilon^n_{m,\sigma})_0 =&H_{mn,\sigma}\check{P}^{xx}_{m,\sigma|\sigma}H^T_{mn,\sigma}\end{aligned}$$

and sends them to area n.

for $h = 1, 2, \ldots$ **do**

Each area $m \in \mathcal{V}$ updates the edge information once receive $(\upsilon^m_{n,\sigma})_{h-1}$ and $(\Upsilon^m_{n,\sigma})_{h-1}$ from its neighbors $n \in \mathcal{N}_m$ via

$$\begin{aligned}(\tilde{z}_{nm,\sigma})_h =&z_{mn,\sigma} - h_{mn}(\hat{X}_{m,\sigma|\sigma-1}, \hat{X}_{n,\sigma|\sigma-1})\\ &+ H_{mn,\sigma}\hat{X}_{m,\sigma|\sigma-1} + H_{nm,\sigma}\hat{X}_{n,\sigma|\sigma-1}\\ &- (\upsilon^m_{n,\sigma})_{h-1},\\ (\bar{V}_{nm,\sigma})_h =&V_{mn,\sigma} + (\Upsilon^m_{n,\sigma})_{h-1}.\end{aligned}$$

Area m computes the current state estimate and its covariance

$$\begin{aligned}(\hat{X}_{m,\sigma|\sigma})_h =&(L_{m,\sigma})^{-1}_h(\eta_{m,\sigma})_h,\\ (P^{xx}_{m,\sigma|\sigma})_h =&(L_{m,\sigma})^{-1}_h\end{aligned}$$

with

$$\begin{aligned}(\eta_{m,\sigma})_h =&\check{\eta}_{m,\sigma} + \sum_{n\in\mathcal{N}_m} H^T_{mn,\sigma}(\bar{V}_{nm,\sigma})^{-1}_h(\tilde{z}_{nm,\sigma})_h,\\ (L_{m,\sigma})_h =&\check{L}_{m,\sigma} + \sum_{n\in\mathcal{N}_m} H^T_{mn,\sigma}(\bar{V}_{nm,\sigma})^{-1}_h H_{mn,\sigma}.\end{aligned}$$

For each area $n \in \mathcal{N}_m$, area m computes

$$\begin{aligned}(\upsilon^n_{m,\sigma})_h =&H_{mn,\sigma}(L^n_{m,\sigma})^{-1}_h(\eta^n_{m,\sigma})_h,\\ (\Upsilon^n_{m,\sigma})_h =&H_{mn,\sigma}(L^n_{m,\sigma})^{-1}_h H^T_{mn,\sigma}\end{aligned}$$

with

$$\begin{aligned}(\eta^n_{m,\sigma})_h =&(\eta_{m,\sigma})_h - H^T_{mn,\sigma}(\bar{V}_{nm,\sigma})^{-1}_h(\tilde{z}_{nm,\sigma})_h,\\ (L^n_{m,\sigma})_h =&(L_{m,\sigma})_h - H^T_{mn,\sigma}(\bar{V}_{nm,\sigma})^{-1}_h H_{mn,\sigma}\end{aligned}$$

and sends $(\upsilon^n_{m,\sigma})_h$ and $(\Upsilon^n_{m,\sigma})_h$ to area n.

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

21: **end for**
22: **end for**

Remark 6: In this paper, the anomaly-resistant decentralized SE algorithm is designed for the wide-area power systems. For the purpose of improving computational efficiency, the SE for power system is executed in a decentralized manner where the whole power grid is divided into several non-overlapping areas which are interconnected through transmission lines. In the design of the anomaly-resistant SE algorithm, an augmented model is established by employing the unscented transformation technique and the statistical linearization technique. Then, an MEEF-based cost function is designed based on the augmented model where the local prediction errors of states and measurements are all reflected. Finally, the local estimate is obtained by iteratively solving an optimization problem with the MEEF-based cost function and such an estimate is then updated by using the edge measurement.

Remark 7: In comparison with the existing SE schemes for wide-area power systems, the main novelties of this paper lie in the following aspects: 1) the scheme we adopted is novel since the state is first estimated in a decentralized manner by using the local measurement and then updated with the aid of edge measurements shared with the neighboring areas; 2) the augmented model we developed is novel since it not only reflects the local prediction errors of the state and measurement but also improves the data redundancy; and 3) the MEEF criterion we adopted in the estimator design is novel since it is more comprehensive than the MEE and MC criteria in the sense of enhancing the resistance of the proposed SE scheme against the anomalies in the measurements.

V. Simulation Example

In this section, the proposed anomaly-resistant decentralized SE algorithm for wide-area power systems is tested on the IEEE 14-bus system. The simulation is implemented in MATLAB-R2018b with the aid of the Matpower package [44]. The whole test cases are performed on a PC with Intel Core CPU i7-7700HQ, 2.80GHz and 16 GB RAM. The IEEE 14-bus is modeled as (1) with $A_i = 0.89$, $B_i = 0.11$ and $W_i = 0.01^2$ $(i = 1, 2, \ldots, 14)$. The value of the expected steady state $\bar{x}_i$ can be found in Table I. The installation positions of the SCADA units and the PMUs remain the same as they were in [45]. For the SCADA measurement functions given in (2) and (4), the corresponding covariance matrices of the measurement noises are set to be $V^s_{i,\sigma} = 1 \times 10^{-3}$ and $V^s_{ij,\sigma} = 1 \times 10^{-3}$ $(i, j \in \{1, 2, \ldots, 14\}, i \neq j)$, respectively. For the i-th bus which is equipped with the PMU, the corresponding covariance matrix of the measurement noise is set as $V^p_{i,\sigma} = 1 \times 10^{-6}$.

As shown in Fig. 1, the IEEE 14-bus system is divided into 3 areas and the corresponding local measurements as well as the edge measurements are all labeled clearly. For the m-th $(m = 1, 2, 3)$ anomaly-resistant decentralized state estimator, the parameters of the unscented transformation are chosen as $\alpha_m = e^2$, $\kappa_m = 0.02$ and $\beta_m = 1$. Moreover, the parameters of the MEEF criterion are selected as $\mu = 0.2$, $\varsigma_1 = 25$ and $\varsigma_2 = 2$, respectively. The mean square error (MSE) criterion is adopted to assess the estimation performance of the i-th states, i.e.

$$\mathrm{MSE}_i = \frac{1}{T}\sum_{i=1}^{T}\left(x_{i,\sigma} - \hat{x}_{i,\sigma}\right)^2$$

where T represents the number of Monte Carlo simulations and $T = 100$.

A. Scenario 1: Centralized Scheme versus Decentralized Scheme

In this scenario, the standard UKF-based centralized SE scheme (labeled as C-UKF) is considered for the purpose of comparison with our proposed anomaly-resistant UKF-based decentralized SE scheme (labeled as RD-UKF). In order to save space, the state trajectories of bus 11 (in area 1) and the state trajectories of bus 10 (in area 2) are taken for illustration. The simulation results are given in Figs. 2-3. To be specific, Fig. 2(a) shows the real curves of the states and the associated estimates of bus 11 in area 1, the corresponding MSEs are plotted in Fig. 2(b). Similarly, the corresponding results for bus 10 in area 2 are shown in Fig. 3(a) and 3(b), respectively.

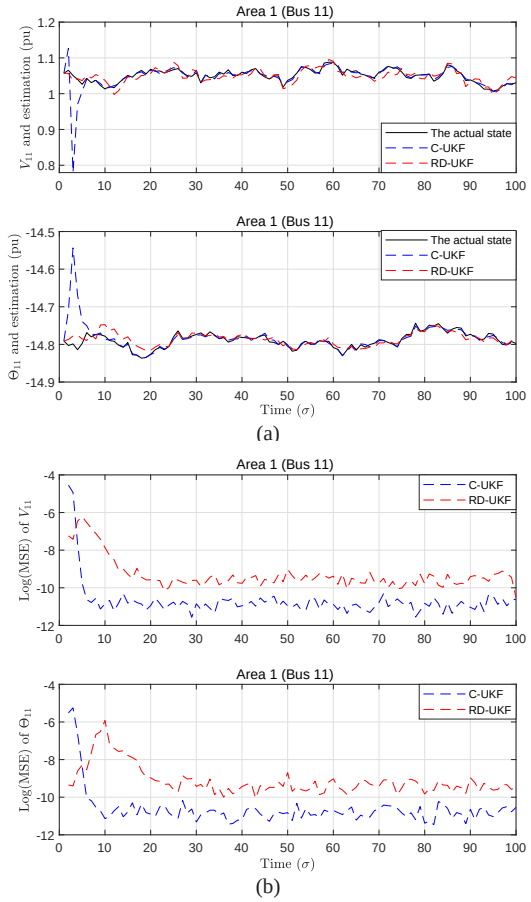

Fig. 2. Scenario 1: Simulation results of bus 11 for area 1: (a) Estimated states. (b) MSE.

It can be found from Figs. 2-3 that: 1) the estimation accuracy of the C-UKF is higher than that of the RD-UKF since the measurement redundancy of the former is higher than that of the latter; and 2) the estimation accuracy of our proposed RD-UKF is acceptable.

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

TABLE I: Expected States of IEEE 14-bus System

Bus	1	2	3	4	5	6	7
Voltage (p.u.)	1.060	1.045	1.010	1.018	1.020	1.070	1.062
Phasor (degree)	0.000	-4.983	-12.725	-10.313	-8.744	-14.211	-13.360
Bus	8	9	10	11	12	13	14
Voltage (p.u.)	1.090	1.056	1.051	1.057	1.055	1.050	1.036
Phasor (degree)	-13.360	-14.939	-15.097	-14.791	-15.076	-15.156	-16.034

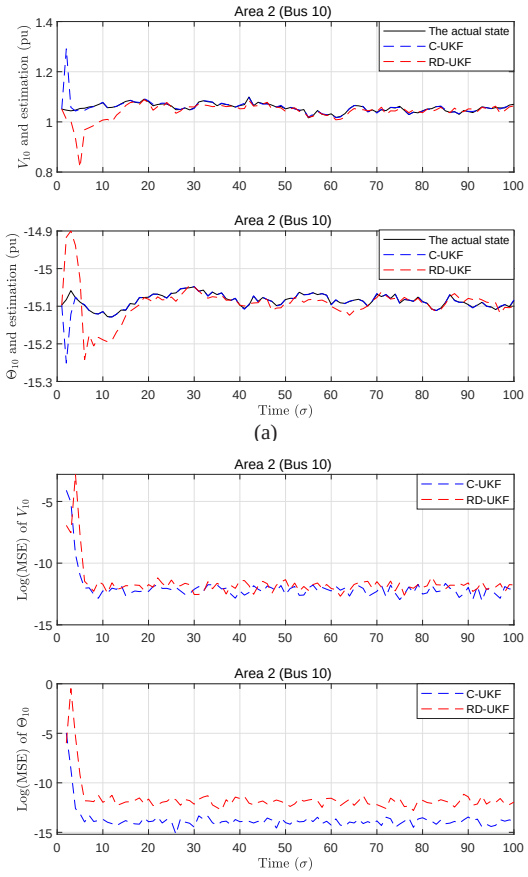

Fig. 3. Scenario 1: Simulation results of bus 10 for area 2: (a) Estimated states. (b) MSE.

B. Scenario 2: Resistance to Measurement Outliers Caused by Non-Gaussian Sequences

For the purpose of verifying the resistance of the developed anomaly-resistant decentralized SE scheme to the measurement outliers caused by the non-Gaussian noises, in this scenario, the following two cases are considered:

Case 1: The local measurements of 3 areas and their corresponding parts in the centralized measurement model are contaminated with consecutive measurement outliers. Specifically, from $\sigma = 30$ to $\sigma = 60$, a two-component Gaussian mixture sequence, which has zero means and covariance matrices of $0.1I$ and $10^{-2}I$ with weights 0.9 and 0.3, is used to model the outliers.

Case 2: In this case, the randomly occurring measurement outliers caused by the non-Gaussian sequence are considered. The detailed parameters of the non-Gaussian sequence are similar to the one in Case 1.

Similarly, only the states of bus 11 (in area 1) are taken for illustration in this scenario. Fig. 4 shows the simulation results under case 1. Specifically, Fig. 4(a) shows the true state trajectories and the associated estimates of bus 11 in area 1, and the corresponding measurement curves contaminated with consecutive measurement outliers are plotted in Fig. 4(b). Similarly, for case 2, the corresponding results are plotted in Fig. 5.

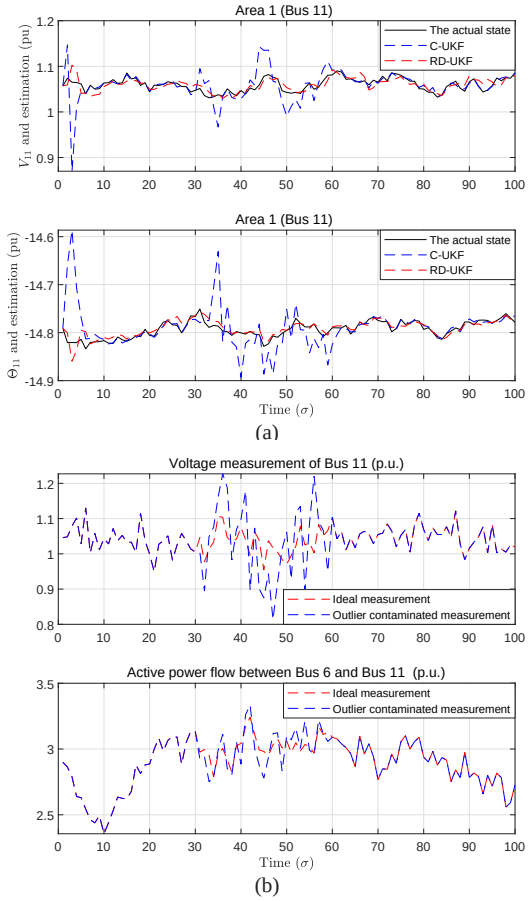

Fig. 4. Scenario 2: Simulation results of bus 11 for area 1 under Case 1: consecutive measurement outliers caused by non-Gaussian noises. (a) Estimated states. (b) Measurement curves.

From Figs. 4-7, it can be concluded that: 1) the estimation performance of the C-UKF degrades severely in the presence of consecutive measurement outliers, while our proposed RD-UKF is anomaly-resistant to such kind of measurement outliers; and 2) when the randomly occurring measurement outliers occur, the estimation performance of the C-UKF degrades severely, and the RD-UKF is still anomaly-resistant to these kinds of measurement outliers.

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

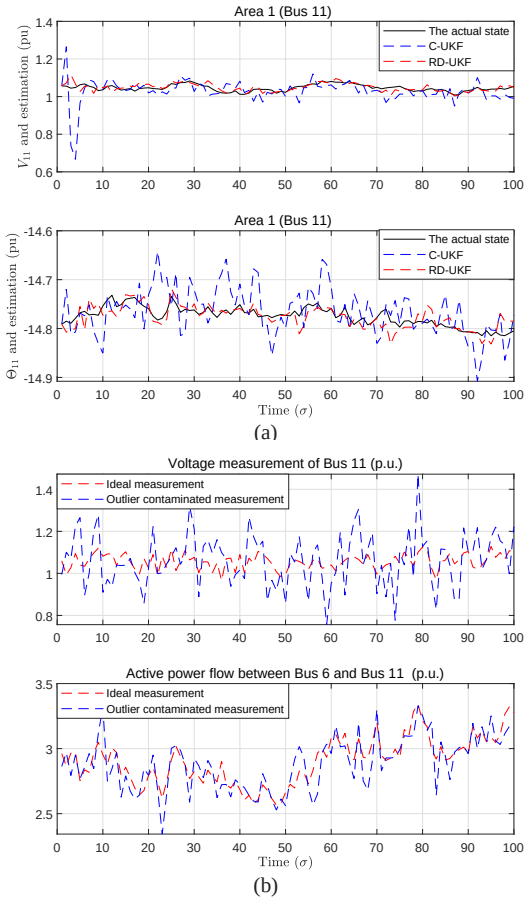

Fig. 5. Scenario 2: Simulation results of bus 11 for area 1 under Case 2: randomly occurring measurement outliers caused by non-Gaussian noises. (a) Estimated states. (b) Measurement curves.

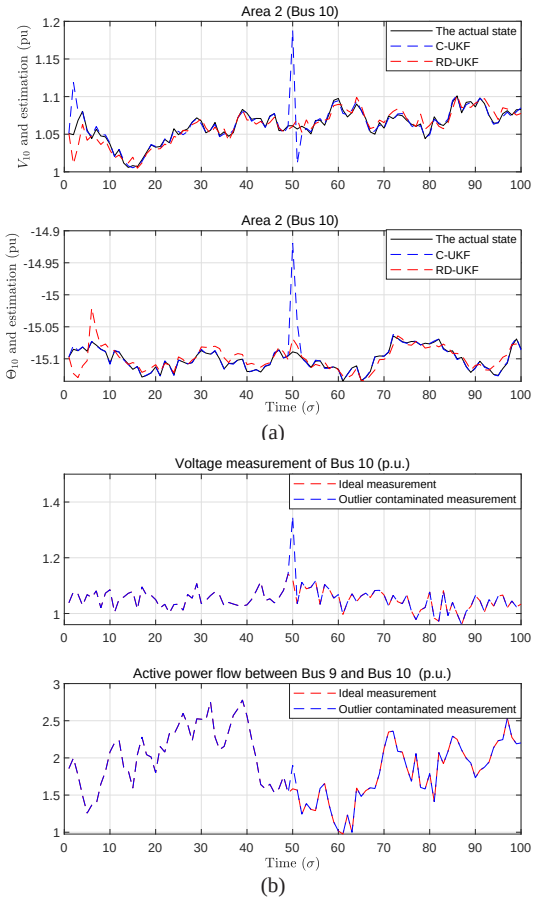

Fig. 6. Scenario 3: Simulation results of bus 10 for area 2 under Case 2: instantaneous bad measurements. (a) Estimated states. (b) Measurement curves.

C. Scenario 3: Resistance to Bad Measurement Data

Due to the presence of sensor saturations, gross errors and/or instrument failures, the measurements in power systems can be contaminated with bad measurements [46]. In order to assess the resistance of the developed SE algorithms, the following two cases are considered:

Case 1: The local measurements of 3 areas and their corresponding parts in the centralized measurement model are contaminated with instantaneous errors. Specifically, when $\sigma = 50$, the measurements are corrupted with 20% errors.

Case 2: In this case, the local measurements of 3 areas and their corresponding parts in the centralized measurement model are contaminated with consecutive errors. Specifically, from $\sigma = 30$ to $\sigma = 60$, the measurements are corrupted with 20% errors.

For the sake of saving space, only the states of bus 10 (in area 2) are taken for illustration in this scenario. The simulation results under case 1 are plotted in Fig. 6. To be specific, the trajectories of the states and the associated estimates of bus 10 in area 2 are plotted in Fig. 6(a), and the corresponds measurement curves contaminated with consecutive gross errors are plotted in Fig. 6(b). Similarly, the corresponding results for case 2 are plotted in Fig. 7.

It can be confirmed from Figs. 6-7 that: 1) when the instantaneous bad measurements occur, the estimation performance of the C-UKF degrades severely, and the proposed RD-UKF is anomaly-resistant to such kinds of gross errors; and 2) the estimates of the C-UKF are non-convergent when the consecutive bad measurements occur, while the proposed RD-UKF performs well under such kinds of gross errors.

D. Computational Efficiency

The average running times (ARTs) of the above three scenarios are given in Table II. It can be confirmed from Table II that the proposed decentralized SE has much higher computational efficiency than the centralized one.

TABLE II: Average Computing Time Under Three Scenarios At Each Iteration

Scenario	1	2		3	
Item	None	Case 1: consecutive measurement outliers	Case 2: randomly occurring measurement outliers	Case 1: instantaneous gross errors	Case 2: consecutive gross errors
AVT (ms) C-UKF/RD-UKF	27/15	28/18	27/17	28/15	26/15

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

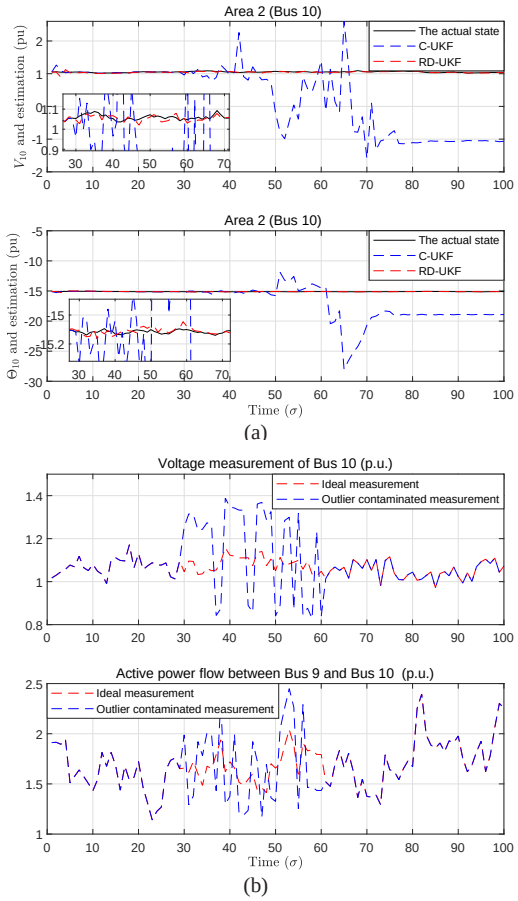

Fig. 7. Scenario 3: Simulation results of bus 10 for area 2 under Case 2: consecutive bad measurements. (a) Estimated states. (b) Measurement curves.

VI. Conclusion

In this paper, the anomaly-resistant decentralized SE problem has been investigated for wide-area power systems. For the purpose of lighting the computation costs and computation burdens, a model decomposition approach has been adopted to facilitate the SE in a decentralized manner. To enhance the resistance of the proposed decentralized SE scheme against measurement with anomalies, the MEEF criterion has been utilized in the estimator design. Specifically, i) by using the augmented model which reflects the information of the local prediction errors of state and measurement, a MEEF-based cost function has been developed; ii) by solving the MEEF-based cost function with the aid of the fixed-point iteration technique, local estimates have been obtained; and iii) by integrating the edge measuring information gradually with the iterative method, the decentralized estimates have been calculated. Finally, based on the IEEE 14-bus system, three test scenarios have been considered in the simulation to verify the effectiveness of our proposed anomaly-resistant decentralized SE scheme. One of the future research directions would be to extend the main SE approach developed in this paper to other more sophisticated systems such as networked systems [47]–[53].

References

[1] F. Han, X. Lao, J. Li, M. Wang and H. Dong, Dynamic event-triggered protocol-based distributed secondary control for islanded microgrids, *International Journal Electrical Power & Energy Systems*, vol. 137, art. no.107723, May 2022.

[2] Z. J. Wang, W. Wei, J. Z. F. Pang, F. Liu, B. Yang, X. P. Guan, and S. W. Mei, Online optimization in power systems with high penetration of renewable generation: Advances and prospects, *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 4, pp. 839–858, 2023.

[3] L. Xie, D-H. Choi, S. Kar and H. V. Poor, Fully Distributed State Estimation for Wide-Area Monitoring Systems, *IEEE Transactions on Smart Grid*, vol. 3, no. 3 pp. 1154–1169, 2012.

[4] A. Singh and S. K. Parida, Power system frequency and phasor estimation for a low-cost synchrophasor device using the nonlinear least-square method, *IEEE Transactions on Industry Applications*, vol. 58, no. 1, pp. 39–48, 2022.

[5] T. Meng, Z. Lin and Y. A. Shamash, Distributed cooperative control of battery energy storage systems in DC microgrid, *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 3, pp. 606-616, 2021.

[6] J. Wu, C. Peng, H. Yang and Y. L. Wang, Recent advances in event-triggered security control of networked systems: a survey, *International Journal of Systems Science*, vol. 53, no. 12, pp. 2624-2643, 2022.

[7] 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.

[8] M. Kooshkbagh, H. J. Marquez and W. Xu, Event-triggered approach to dynamic state estimation of a synchronous machine using cubature Kalman filter, *IEEE Transactions on Control Systems Technology*, vol. 28, no. 5, pp. 2013–2020, 2020.

[9] M. Göl 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.

[10] B. Qu, Z. Wang and B. Shen, Fusion estimation for a class of multi-rate power systems with randomly occurring SCADA measurement delays, *Automatica*, vol. 125, art.no. 109408, 2021.

[11] Z. H. Pang, L. Z. Fan, H. Guo, Y. Shi, R. Chai, J. Sun and G. Liu, Security of networked control systems subject to deception attacks: A survey, *International Journal of Systems Science*, vol. 53, no. 16, pp. 3577-3598, 2022.

[12] F. Qu, X. Zhao, X. Wang and E. Tian, Probabilistic-constrained distributed fusion filtering for a class of time-varying systems over sensor networks: A torus-event-triggering mechanism, *International Journal of Systems Science*, vol. 53, no. 6, pp. 1288-1297, 2022.

[13] 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.

[14] M. Ghosal and V. Rao, Fusion of multirate measurements for nonlinear dynamic state estimation of the power systems, *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 216–226, 2019.

[15] L. Hu, Z. Wang and X. Liu, Dynamic state estimation of power systems with quantization effects: A recursive filter approach, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1604–1614, 2016.

[16] M. N. Kurt, Y. Yılmaz and X. Wang, Secure distributed dynamic state estimation in wide-area smart grids, *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 800–815, 2020.

[17] D. E. Marelli and M. Fu, Distributed weighted least-squares estimation with fast convergence for large-scale systems, *Automatica*, vol. 51, pp. 27–39, 2015.

[18] K. Li and X. Han, A distributed Gauss-Newton method for distribution system state estimation, *International Journal of Electrical Power and Energy Systems*, vol. 136, art.no. 107694, 2022.

[19] M. Rostami and S. Lotfifard, Distributed dynamic state estimation of power systems, *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3395-3404, 2018.

[20] A. K. Singh and B. C. Pal, Decentralized robust dynamic state estimation in power systems using instrument transformers, *IEEE Transactions on Signal Processing*, vol. 66, no. 6, pp. 1541–1550, 2018.

[21] Z. Guo, S. Li, X. Wang and W. Heng, Distributed point-based Gaussian approximation filtering for forecasting-aided state estimation in power systems, *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2597–2608, 2016.

[22] M. Z. El-Sharfy, S. Saxena and H. E. Farag, Optimal design of islanded microgrids considering distributed dynamic state estimation, *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 1592–1603, 2021.

[23] Y. Sun, M. Fu, B. Wang, H. Zhang and D. Marelli, Dynamic state estimation for power networks using distributed MAP technique, *Automatica*, vol. 73, pp. 23–37, 2016.

[24] B. Wei, E. Tian, T. Zhang and X. Zhao, Probabilistic-constrained H_∞ tracking control for a class of stochastic nonlinear systems subject to DoS attacks and measurement outliers, *IEEE Transactions on Circuits and Systems I-Regular Papers*, vol. 68, no. 10, pp. 4381–4392, 2021.

[25] H. Tao, H. Tan, Q. Chen, H. Liu and J. Hu, H_∞ state estimation for memristive neural networks with randomly occurring DoS attacks, *Systems Science & Control Engineering*, vol. 10, no. 1, pp. 154-165, 2022.

[26] W. Xue, X. Luan, S. Zhao and F. Liu, A fusion Kalman filter and UFIR estimator using the influence function method, *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 4, pp. 709-718, Apr. 2022.

[27] B. Jiang, H. Dong, Y. Shen and S. Mu, Encoding-decoding-based recursive filtering for fractional-order systems, *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 6, pp. 1103-1106, Jun. 2022.

[28] F. M. Shakiba, M. Shojaee, S. M. Azizi and M. Zhou, Real-time sensing and fault diagnosis for transmission lines, *International Journal of Network Dynamics and Intelligence*, vol. 1, no. 1, pp. 36-47, 2022.

[29] 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.

[30] 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, 2020.

[31] H. Song, D. Ding, H. Dong and X. Yi, Distributed filtering based on Cauchy-kernel-based maximum correntropy subject to randomly occurring cyber-attacks, *Automatica*, vol. 135, art.no. 110004, 2022.

[32] W. Liu, P. P. Pokharel and J. C. Principe, Error entropy, correntropy and M-estimation, in *proceedings of 16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing*, Maynooth, Ireland, 2006, pp. 179–184.

[33] M. B. Do Coutto Filho and J. C. S. de Souza, Forecasting-aided state estimation–Part I: Panorama, *IEEE Transactions on Power Systems*, vol. 24, no. 4, pp. 1667–1677, 2009.

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

[35] A. Sinha and J. Mondal, Dynamic state estimator using ANN based bus load prediction, *IEEE Transactions on Power Apparatus and Systems*, vol. 14, no. 4, pp. 1219–1225, 1999.

[36] Q. Zhang and Y. Zhou, Recent advances in non-Gaussian stochastic systems control theory and its applications, *International Journal of Network Dynamics and Intelligence*, vol. 1, no. 1, pp. 111-119, 2022.

[37] J. C. Príncipe, Information theoretic learning: Rényi's entropy and kernel perspectives, Springer Press: New York, 2010.

[38] T. Hu, J. Fan, Q. Wu and D.-X. Zhou, Learning theory approach to minimum error entropy, *Journal of Machine Learning Research*, vol. 14, pp. 377–397, 2013.

[39] D. Erdogmus and J. C. Principe, An error-entropy minimization algorithm for supervised training of nonlinear adaptive systems, *IEEE Transactions on Signal Processing*, vol. 50, no. 7, pp. 1780–1786, 2002.

[40] R. J. Bessa, V. Miranda and J. Gama, Entropy and correntropy against minimum square error in offline and online three-day ahead wind power forecasting, *IEEE Transactions on Power Systems*, vol. 24, no. 4, pp. 1657–1666, 2009.

[41] Y. Xie, Y. Li, Y. Gu, J. Cao and B. Chen, Fixed-point minimum error entropy with fiducial points, *IEEE Transactions on Signal Processing*, vol. 68, pp. 3824–3833, 2020.

[42] T. Lefebvre, H. Bruyninckx and J. De Schuller, Comment on "A new method for the nonlinear transformation of means and covariances in filters and estimators", *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1406–1409, 2002.

[43] J. Zhao and L. Mili, A robust generalized-maximum likelihood unscented Kalman filter for power system dynamic state estimation, *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 578–592, 2018.

[44] R. D. Zimmerman, C. E. Murillo-Sánchez and R. J. Thomas, MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education, *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[45] G. N. Korres and N. M. Manousakis, State estimation and bad data processing for systems including PMU and SCADA measurements, *Electric Power Systems Research*, vol. 81, no. 7, pp. 1514–1524, Jul. 2011.

[46] J. Zhao, A. Gomez-Exposito, M. Netto, L. Mili, A. Abur, V. Terzija, I. Kamwa, B. Pal, A. K. Singh, J. Qi, Z. Huang and A. P. S. Meliopoulos, Power system dynamic state estimation: Motivations, definitions, methodologies, and future work, *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3188–3198, 2019.

[47] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, Networked fusion estimation with multiple uncertainties and time-correlated channel noise, *Information Fusion*, vol. 54, pp. 161–171, 2020.

[48] D. Ciuonzo, A. Aubry, and V. Carotenuto, Rician MIMO channel- and jamming-aware decision fusion, *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3866–3880, 2017.

[49] Y. S. Shmaliy, S. Zhao and C. K. Ahn, Unbiased finite impluse response filtering: an iterative alternative to Kalman filtering ignoring noise and initial conditions, *IEEE Control Systems Magazine*, vol. 37, no. 5, pp. 70-89, 2017.

[50] X.-M. Zhang, Q.-L. Han, X. Ge, and L. Ding, Resilient control design based on a sampled-data model for a class of networked control systems under denial-of-service attacks, *IEEE Transactions on Cybernetics*, vol. 50, no. 8, pp. 3616–3626, 2020.

[51] X. Ge, Q.-L. Han, and Z. Wang, A threshold-parameter-dependent approach to designing distributed event-triggered H_∞ consensus filters over sensor networks, *IEEE Transactions on Cybernetics*, vol. 49, no. 4, pp. 1148–1159, 2019.

[52] J. Hu, C. Jia, H. Yu and H. Liu, Dynamic event-triggered state estimation for nonlinear coupled output complex networks subject to innovation constraints, *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 5, pp. 941-944, 2022.

[53] P. Wen, X. Li, N. Hou and S. Mu, Distributed recursive fault estimation with binary encoding schemes over sensor networks, *Systems Science & Control Engineering*, vol. 10, no. 1, pp. 417-427, 2022.

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: DOI10.1109/JAS.2023.123795, IEEE/CAA Journal of Automatica Sinica

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.

Hongjian Liu (Member, IEEE) received his B.Sc. degree in applied mathematics in 2003 from Anhui University, Hefei, China, and the M.Sc. degree in detection technology and automation equipment in 2009 from Anhui Polytechnic University, Wuhu, China, and the Ph.D. degree in control science and engineering in 2018 from Donghua University, Shanghai, China. In 2016, he was a Research Assistant with the Department of Mathematics, Texas A&M University at Qatar, Doha, Qatar, for two months. From March 2017 to March 2018, he was a Visiting Scholar in the Department of Information Systems and Computing, Brunel University London, UK. He is currently a Professor in the School of Mathematics and Physics, Anhui Polytechnic University, Wuhu, China.

Dr. Liu's current research interests include filtering theory, memristive neural networks, and network communication systems. He is a very active reviewer for many international journals.