Power System Adaptive State Estimation in Unknown Measurement Environment

Gang Cheng, *Graduate Student Member, IEEE*, and Yuzhang Lin, *Member, IEEE*

***Abstract*— The complicated sensing and communication environment of power systems results in measurement errors with unknown, nonzero-mean, non-Gaussian, and time-varying statistics. Traditional state estimator designs are based on heuristic assumptions of measurement error distributions and are agnostic to the true error statistics, yielding suboptimal error filtering performances in reality. This article investigates the supervisory control and data acquisition (SCADA) and phasor measurement unit (PMU) measurement chain modeling and presents a new state estimation (SE) paradigm based on the concept of adaptive SE (ASE). Instead of ignoring or passively resisting the unknown measurement error statistics, it proactively captures this information and adapts the structure and parameters of the estimator online to optimize the accuracy of the state estimates. The proposed method can capture arbitrarily complex measurement error distributions, preserves high computational efficiency, adapts to abrupt gross errors, and also enables a sensor calibration approach for both PMUs and SCADA without the need for field experiments. The proposed method is validated on the IEEE 30-bus test system with complex and time-varying measurement errors generated by comprehensive SCADA and PMU measurement chain modeling.**

***Index Terms*— Adaptive state estimation (ASE), measurement error distribution, phasor measurement unit, sensor calibration.**

I. Introduction

COMBATING measurement errors is one of the main functionalities of power system state estimation (PSSE). As networked control systems with communication over vast geographical areas, electric power systems are exposed to complicated measurement environments with ambient noise, sensor drift, signal processing errors, time skews, communication delays, packet drops and failures, and even cyber-physical attacks [1], [2], resulting in the deviations of measurements from the true values of the measured variables, namely measurement errors. Although these errors are unknown and stochastic, it is possible to limit their impacts given their statistical knowledge as well as information redundancy. The core objective of PSSE is to minimize the impact of measurement errors on the recovery of system state variables.

The earliest proposed PSSE method is the weighted least squares (WLS) state estimation (SE) dating back to the 1970s [3]. In essence, it is a maximum likelihood estimator (MLE) assuming the following "benign" measurement error conditions: 1) Gaussian distributed; 2) zero-mean; and 3) known variance. Theoretically, if all these conditions are satisfied, WLS SE yields state estimates whose errors have zero means (i.e., unbiased) and minimum variances. However, the popularity of WLS SE actually resulted from a more practical reason: under the aforementioned conditions, the MLE criterion leads to an elegant unconstrained optimization problem with a quadratic-form objective function, which can be solved by Newton's method with modest computational costs [1]. This was a major advantage in the days when computing resources were of primary concern. However, a close examination reveals that these conditions are unlikely to hold in real-world power systems. Even though the noise in originally measured quantities may follow a close-to-Gaussian distribution, the signal processing algorithms (e.g., power calculation in supervisory control and data acquisition (SCADA) systems, fast Fourier transform in phasor measurement units (PMUs), etc.), and time skews induced by communication latencies will distort the distributions and make them non-Gaussian [4], [5], [6]. Sensor drift, improper calibration, and even cyber-attacks can easily make the measurement errors nonzero-mean [7], [8]. Moreover, the variances of errors are typically unknown and can vary with the operating points of the system and the aging condition of the sensors. The violation of the "benign" measurement error conditions may severely degrade the performance of WLS SE, producing state estimates with nonzero-mean errors and large variances.

In order to achieve more stable SE performances in realistic conditions, the concept of robust SE (RSE) is proposed. The earliest and most widely known RSE method is the weighted least absolute value (WLAV) SE [9]. In essence, WLAV SE is the MLE under measurement errors with Laplacian distributions with zero means and known scale parameters. The advantage of assuming Laplacian instead of Gaussian distributions is that they are "heavy-tailed" distributions encompassing large measurement errors (also known as gross errors) that commonly occur in power systems. Over the years, WLAV SE has been enhanced to handle leverage points [10], [11], model parameter errors [12], and to reduce computational costs [13]. The WLAV SE belongs to a larger family of M-estimators, which minimize the sum of various symmetric functions of residuals as the MLEs of zero-mean measurement error distributions of different shapes [14]. An effort has also been dedicated to reweighting M-estimators to suppress the effect of leverage points, referred to as the generalized M-estimators [15]. There have been other types of RSE methods that deviate from the MLE framework, e.g., the maximum

Manuscript received 9 December 2023; revised 20 April 2024; accepted 25 April 2024. Date of publication 20 May 2024; date of current version 5 June 2024. This work was supported by the National Science Foundation under Award 2348289. The Associate Editor coordinating the review process was Dr. Grazia Barchi. *(Corresponding author: Yuzhang Lin.)*

Gang Cheng is with the Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, MA 01854 USA (e-mail: gang_cheng@student.uml.edu).

Yuzhang Lin is with the Department of Electrical and Computer Engineering, New York University, Brooklyn, NY 11201 USA (e-mail: yuzhang.lin@nyu.edu).

Digital Object Identifier 10.1109/TIM.2024.3403203

normal measurement rate (MNMR) based SE [16], the least absolute shrinkage and selection operator (LASSO) based SE [17], the maximum exponential absolute value (MEAV) SE [18], and the data-fusion-based SE [19]. In summary, the RSE methods are designed either based on heuristic assumptions of heavy-tailed distributions of measurement errors or to have formulations that are insensitive to a wide variety of possible distributions, such that consistent results are obtained irrespective of the true distributions of measurement errors. As they are inherently agnostic to the true distributions of measurement errors, RSE methods can only stabilize the results, i.e., to avoid unacceptably large errors of state estimates, but cannot truly optimize the results, i.e., to minimize the means and variances of errors of state estimates.

In view of the limitations of conventional WLS SE and RSE methods, this article presents a fundamentally different paradigm to maneuver PSSE through a complicated and unknown measurement environment: adaptive SE (ASE). The ASE concept is motivated by the following visions: 1) in order to optimize the performance of PSSE, the true statistical knowledge of the measurement errors must be captured and exploited and 2) as the probability distributions of measurement errors are time-varying, the formulation of PSSE must be adapted consistently online to maintain near-optimal performance. Based on these visions, we will design an algorithm to capture the probability distributions of measurement errors, and then use this statistical knowledge to adapt the PSSE formulation online so as to enhance the statistical accuracy of state estimates. As a result, the performance of PSSE can always be optimized online under complicated measurement error conditions without relying on any prior knowledge of measurement error statistics. The contributions of the article are as follows.

1) We perform detailed measurement chain modeling to analyze and synthesize SCADA and PMU measurements in power systems. The sources of non-Gaussianity and time variance of measurement error distributions are identified, and realistic measurement errors are synthetized for the validation of PSSE performances.
2) We develop an effective algorithm to capture any arbitrarily complex probability distributions of measurement errors, jointly with power system state variables, using the expectation maximization (EM) algorithm and the Gaussian mixture model (GMM). As a byproduct, it also enables sensor calibration without the need for field experiments.
3) We develop an adaptive state estimator that proactively adapts its structure and parameters to the updated statistical knowledge of measurement errors to optimize the accuracy of state estimates under a complicated measurement environment. The proposed SE model achieves high computational efficiency. Through a special "gross error trap" design, it also maintains high accuracy even under abrupt gross errors that do not follow the captured statistical knowledge of measurement errors.

Note that the ASE concept is a well-established concept in the signal processing domain. However, *the concepts and algorithms proposed in this article constitute the first systematic framework, in the power system domain, to proactively optimize PSSE performance under unknown, non-Gaussian, nonzero-mean, and time-varying measurement error distributions with any arbitrary shape*. The preliminary effort was reported in our earlier work [20], where a mixture model of one Gaussian component and one Laplacian component, both with zero means, is used to model measurement errors. This article makes fundamental advancements over our earlier work [20] in multiple aspects, as listed in Table I.

TABLE I
COMPARISON: OUR EARLIER WORK [20] AND THE PROPOSED WORK

Our Earlier Work [20]	This Work
Assumes non-Gaussianity and time variation of measurement errors without explicit evidence	Performs power system measurement chain modeling and identifies sources of non-Gaussianity and time variation
Captures measurement error distributions with limited complexity (single peak, symmetric, zero-mean);	Captures measurement error distributions with arbitrary complexity (multiple peaks, asymmetric, nonzero-mean)
Assumes pre-calibrated sensors without permanent biases	Accounts for sensor permanent biases and enables sensor calibration
Estimator with a fixed structure and adaptive parameters	Estimator with both an adaptive structure and adaptive parameters
Applicable to PMU measurements only	Applicable to a mix of PMU and SCADA measurements
Low computational efficiency with l_2- l_1-mixed optimization formulation	High computational efficiency with l_2 optimization formulation
Validates the proposed method on synthetic measurements with heuristically designed error distributions	Validates the proposed method on synthetic measurements generated from realistically modeled PMU and SCADA measurement chains

II. MODELING AND ANALYSIS OF ERROR SOURCES AND CHARACTERISTICS IN MEASUREMENT CHAINS

In this section, we will model and analyze the SCADA and PMU measurement chains in detail and identify the sources of measurement error complexity, i.e., nonzero mean, non-Gaussianity, and time variance. In Table II, the measurement error sources of each component in the SCADA and PMU measurement chains are analyzed and summarized.

A. Measurement Chain Analysis

1) SCADA Measurement Chain: Modern SCADA systems rely on data concentrators and intelligent electronic devices (IEDs) in replacement of the conventional remote terminal units (RTUs) with their hardwired input and output (I/O) points [21]. In this article, three major components are modeled in the SCADA measurement chain, as shown in Fig. 1. The errors introduced by each component are elaborated as follows:

a) Instrument transformers: Instrument transformers comprise voltage transformers (VTs) and current transformers (CTs). The IEEE Standard C5713-2016 implies that VTs and CTs will introduce both *ratio* and *phase angle errors* [22], [23]. Accuracy classes, power factors of supply systems, percentages of rated currents, burdens, etc., are a few factors that may affect measurement errors [23]. First, for both *ratio* and *phase angle errors*, instrument transformers may introduce

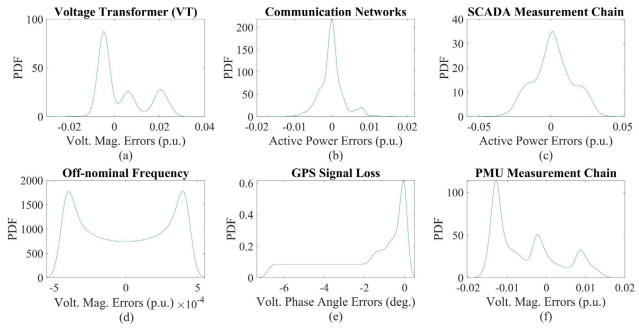

Fig. 1. Measurement error distributions introduced by different components in SCADA and PMU measurement chains: (a) voltage magnitude errors from VTs; (b) active power errors from communication networks; (c) active power errors from the entire SCADA chain; (d) voltage magnitude errors from off-nominal frequency; (e) voltage phase angle errors from GPS signal loss; and (f) voltage magnitude errors the entire PMU chain.

systematic errors dependent on the accuracy class of devices. In practice, VTs and CTs are usually calibrated by using ratio and phase shift angle correction factors when they are used in real-world power systems. Nevertheless, systematic errors introduced by instrument transformers may not be fully calibrated, potentially leading to measurement errors following *nonzero-mean* distributions. Second, for CTs, the magnitudes of systematic errors will be impacted by the percentage of rated currents. Typically, the smaller the percentage of rated current, the larger the systematic errors [23]. In power systems, power flow fluctuations may lead to time-varying currents, and thus *time-varying* systematic errors. Third, random noises are typically unavoidable in real-world power or electronic devices. However, the random noise introduced by instrument transformers has not been studied in existing literature. In this article, we consider assuming a *non-Gaussian* distribution as a more conservative and cautious choice for the random noise, since there is no literature demonstrating that random noises always follow a Gaussian distribution. This allows us to comprehensively explore the potential range of error distributions in future studies, avoiding oversights regarding possibilities. In summary, the factors discussed above may lead to both *ratio* and *phase angle errors* following *nonzero-mean*, *non-Gaussian*, and *time-varying* distributions.

b) Control cables and burdens: Control cables will introduce a time delay transformed into phase angle errors [24], resulting in *nonzero-mean* error distributions for phase angles. This delay depends on several characteristics of control cables, such as the cable length, material, and whether it is shielded. Moreover, burden resistances may also introduce systematic errors, resulting in *nonzero-mean* error distributions [25]. The impact of burden resistance combined with control cables can be significant, especially for longer cable lengths [24], [25]. In addition, instrumentation cables may also introduce random errors, called thermal noise or Johnson noise, that are caused by the motion of free electrons in a resistance due to temperature changes [4], [26]. In the real world, systematic errors introduced by control cables and burdens may not be fully calibrated, so they can lead to measurement errors following *nonzero-mean* distributions. Moreover, since the characteristics of control cables in different measurement channels may be different, the time delay introduced in different channels may also be different, resulting in time skew errors in measurements.

c) IEDs: The industry standard definition of an IED is "any devices incorporating one or more processors with the capability to receive or send data/control from or to an external source" [21]. In this article, it is assumed that the active/reactive power calculation is conducted within IEDs. The power measurements output from IEDs are expressed as follows:

$$P = V \cdot I \cdot \cos\left(\delta^V - \delta^I\right) \tag{1}$$

$$Q = V \cdot I \cdot \sin\left(\delta^V - \delta^I\right) \tag{2}$$

where P and Q represent the active and reactive power calculated within IEDs, respectively; V and I represent the voltage and current magnitudes input to IEDs, respectively; and δ^V and δ^I represent voltage and current phase angles input to IEDs, respectively. Based on (1) and (2), it can be concluded that power calculation will fuse both voltage and current magnitude and phase angle errors introduced by VTs, CTs, control cables, and burdens, further complicating the error distributions for active and reactive power measurements. Even if the errors of the original voltage and current quantities follow a Gaussian distribution, the power calculation within IEDs can transform them into *non-Gaussian* errors in their output power measurements, since the transformation of a random variable through a nonlinear function, such as taking the cosine, often leads to a more complicated distribution that may not have a simple closed-form expression [27].

d) Communication Networks: Communication networks may introduce unpredictable measurement errors due to both latency and power flow fluctuations [28], [29]. In communication networks, latency commonly arises from a variety of factors, such as propagation delay, transmission delay, processing delay, queueing delay, routing delay, network congestion, packet retransmission, etc. Essentially, the latency does not affect the measurement itself since it has already been computed and digitalized. However, since the communication latencies in different channels may be different, measurement data may not be simultaneously received in the control center, resulting in time skew errors in measurements, especially when they are not time-stamped (conventional SCADA measurements). Under the fluctuation of power flows, time skew errors can lead to discrepancies in the operating conditions of the grid compared to those present at the time the measurement was initially taken. Consequently, measurement data received in the control center will deviate from the current operating states of power grids, resulting in measurement errors that may follow *nonzero-mean* and *non-Gaussian* distributions. In addition, the changes in communication conditions and power flow fluctuations further complicate measurement errors, leading to *time-varying* error characteristics.

2) PMU Measurement Chain: In wide-area monitoring systems (WAMSs), measurement errors may stem from a variety of factors [7], [25], [30]. In this article, three major components, including instrument transformers, control cables and burdens, and PMUs, are considered to model the PMU

measurement chain. The measurement errors introduced by each component are elaborated as follows.

a) Instrument transformers: The measurement errors introduced by VTs and CTs in the PMU measurement chain are the same as those in the SCADA measurement chain.

b) Control cables and burdens: The measurement errors introduced by control cables and burdens are the same as those in the SCADA measurement chain.

c) PMUs: Phasor estimation is commonly conducted within PMUs, which aims to compute phasors from sampled values of the input signal [31]. Various phasor estimation algorithms have been proposed, such as discrete Fourier transform (DFT) methods, least squares methods, Kalman filter methods, and Prony methods. In phasor estimation, several factors can lead to estimation errors, such as sampling time errors [32], off-nominal frequency of input signals [33], and GPS signal loss [34]. Specifically, first, *sampling time errors* are usually introduced by the frequency drift of oscillators and the fact that the sampling clock is not precisely at a multiple of the power system frequency [32]. The sampling time error will be accumulated in one second and then cleared by the pulse per second (PPS) timing reference sent by the GPS. In DFT-based phasor estimation, sampling time errors will create inaccuracy in phasor estimates, yielding phase angle estimation errors following *nonzero-mean near-uniform* distributions [32]; Second, *off-nominal frequency* is the most common phenomenon in power system operation due to load and generation imbalances, generator inertia, frequency control, faults, and switching events. In DFT-based phasor estimation, the off-nominal frequency will result in phasor estimation errors with time-domain sinusoidal waveforms, making both magnitude and phasor angle estimation errors clearly *non-Gaussian* [33]. Even if the frequency deviation is very small (e.g., tens of mHz from 60 Hz), the magnitude of the phasor estimation error could be substantial, even though the time-domain sinusoidal waveform of the error will be a long-period, low-frequency signal. Moreover, the frequency of the sinusoidal waveform of phasor estimation errors is determined by the frequency deviation, which is time-varying with the change in the operating state of power systems. Hence, the frequency of the sinusoidal waveform will also change over time, leading to *time-varying* phasor estimation error distributions. Third, GPS signal losses can result from various uncontrollable and unpredictable factors, such as atmospheric disturbances, failure of the GPS antenna, extreme weather, etc. When the GPS signal is lost, the accurate PPS sent by the GPS will not be available in PMUs. Consequently, the sampling time error introduced by internal crystal oscillators will accumulate over time, resulting in time skew errors. This, in turn, leads to phase angle estimation errors following nonzero-mean, long-tailed, and asymmetric distributions [34].

B. Modeling and Simulation of Measurement Errors

In Section II-A, measurement error sources in both SCADA and PMU measurement chains have been comprehensively studied, revealing multiple sources of *nonzero-mean*, *non-Gaussian*, and *time-varying* measurement error distributions. In order to synthesize SCADA and PMU errors with realistic characteristics, we modeled the measurement chains in detail, which, due to the limited space of the article, was described in an online document [35]. In the article, we only illustrate a few simulation results from the developed measurement chain models.

In the SCADA measurement chain, systematic errors are set to follow a uniform distribution, where the limits are determined by the accuracy class and percentage of rated current [23]. As there is no existing study on *random errors* of instrument transformers, it is assumed that they follow a GMM distribution. The GMM serves as a tool to fit the non-Gaussian distribution, and various alternatives, such as the Laplacian distribution, Cauchy distribution, t-distribution, etc., exist for the same purpose. In this manuscript, the GMM is selected due to its flexibility, as it can effectively model arbitrary distributions when the parameters are appropriately selected. Also, we developed a procedure to allow user-defined properties of the GMM distributions when modeling the random errors, which is elaborated in [35]. *It allows users to customize the GMM distribution and determine the similarity between the designed GMM distribution and a Gaussian distribution*. The distribution of voltage magnitude measurement errors introduced by VTs is shown in Fig. 1(a). In communication networks, measurement errors resulting from latency [36] and power flow fluctuations (EPFL PMU dataset [37]) exhibit the *nonzero-mean* and *non-Gaussian* distribution [i.e., Fig. 1(b)]. Considering all components in the SCADA measurement chain [including those not shown in Fig. 1(a) and (b)], the final distribution of active power measurement errors is shown in Fig. 1(c), exhibiting *nonzero-mean* and *non-Gaussian* characteristics.

In the PMU measurement chain, VTs and CTs are modeled in the same way as those in the SCADA measurement chain. In phasor estimation, the distributions of measurement errors introduced by the off-nominal frequency and GPS signal loss are shown in Fig. 1(d) and (e), respectively. Clearly, the off-nominal frequency may result in errors exhibiting *multiple-peak* distributions, while the GPS signal loss may lead to errors with *long-tailed* and *asymmetric* distributions. Considering all components in the PMU measurement chain [including those not shown in Fig. 1(d) and (e)], the distribution of voltage magnitude measurement errors is shown in Fig. 1(f), which is, again, strongly *nonzero-mean* and *non-Gaussian*.

In addition, please note that the impact (contribution) of individual components on the final shape of measurement error distributions is dependent on the realistic measurement conditions, such as the accuracy class of instrument transformers, the control cable length, the communication delay, sampling time errors, frequency of input signals, etc.

III. Modeling of Unknown Measurement Errors

In view of the complexity of measurement error characteristics in SCADA and PMU measurement chains, the proposed ASE framework utilizes a generic and flexible GMM to model the unknown and time-varying measurement error distributions. Assume that a n-bus power system is measured by PMUs and SCADA. The numbers of PMU and SCADA

measurement channels are m_{pmu} and m_{scada}, respectively. The sampling rate of SCADA is one scan per 2–5 s. PMUs provide synchrophasor measurements with a sampling rate as fast as 60 scans per second [38].

As the article focuses on *static* SE under *steady-state* operation of power systems, two reasonable assumptions are made. It is assumed that state variables are approximately unchanged in a very short time period (e.g., within 0.2 s) containing multiple PMU measurement scans, and measurement error distributions are approximately unchanged within a median time interval (e.g., within 30 min). This is because the operating conditions of the real-world power grid, such as fluctuations of load profiles, operating states of field instruments, ambient temperatures, etc., usually have slight changes within a short time interval. To clearly describe the measurement data, all measurements sampled at exactly the same time instant are referred to as a *measurement scan*. All measurement scans within a very short time period with approximately unchanged state variables are referred to as a *measurement group*. All measurement groups within a time interval with approximately unchanged measurement error distributions are referred to as a *measurement set*. Note that SCADA measurements are usually asynchronous and accurate time stamps are unavailable. However, they can still be assigned to a measurement group using their time of arrival. The time skew errors are taken as part of the measurement errors whose statistics are to be collectively estimated with other parts of errors.

A measurement set including SCADA measurements and PMU measurements is shown as follows:

$$\boldsymbol{Z} = \left\{ z^{(1)}, z^{(2)}, \ldots, z^{(l)}, \ldots, z^{(L)} \right\}, \quad l = 1, 2, \ldots, L \tag{3}$$

$$z^{(l)} = \begin{cases} \left(z_{\text{pmu}}^{(l,1)T}, \ldots, z_{\text{pmu}}^{(l,s)T}, \ldots, z_{\text{pmu}}^{(l,S_{\text{pmu}})T} \right)^T, \\ \qquad \text{if the } l\text{th group does not contain SCADA data} \\ \left(z_{\text{scada}}^{(l,S_{\text{scada}})T}, z_{\text{pmu}}^{(l,1)T}, \ldots, z_{\text{pmu}}^{(l,s)T}, \ldots, z_{\text{pmu}}^{(l,S_{\text{pmu}})T} \right)^T, \\ \qquad \text{if the } l\text{th group contains SCADA data} \end{cases} \tag{4}$$

where $\boldsymbol{Z}$ represents the measurement set including SCADA measurements and PMU measurements within a time interval; $z^{(l)}$ represents the lth measurement group; $z_{\text{pmu}}^{(l,s)} \in \mathbb{R}^{m_{\text{pmu}} \times 1}$ is the sth PMU measurement scan in the lth group; $z_{\text{scada}}^{(l,S_{\text{scada}})} \in \mathbb{R}^{m_{\text{scada}} \times 1}$ is a SCADA measurement scan in the lth group; S_{pmu} is the number of PMU measurement scans in a group; S_{scada} is the number of SCADA measurement scan in a group, which is typically equal to 1 as the time between two SCADA scans is typically much longer than the length of a measurement group; and L is the number of measurement groups within the measurement set.

The relationship among measurement data, state variables, and measurement errors are as follows:

$$z^{(l)} = \boldsymbol{h}\left(\boldsymbol{x}^{(l)} \right) + \boldsymbol{e}^{(l)} \tag{5}$$

where $\boldsymbol{x}^{(l)} \in \mathbb{R}^{2n \times 1}$ represents the state variable vector of the lth measurement group; $\boldsymbol{h}(\cdot)$ represents the relation between the state variable vector and the measurement data; and $\boldsymbol{e}^{(l)}$ represents the measurement errors of the lth measurement group.

As GMM can be used to approximate probability distributions of any shape [39], the probability density functions (pdfs) of the measurement errors can be expressed as

$$p\left(e_{\kappa,i}^{(l,s)} \middle| \boldsymbol{\theta}_{\kappa,i} \right) = \sum_{k=1}^{K_{\kappa,i}} \phi_{\kappa,i,k} N\left(e_{\kappa,i}^{(l,s)} \middle| \mu_{\kappa,i,k}, \sigma_{\kappa,i,k}^2 \right) \tag{6}$$

where $e_{\kappa,i}^{(l,s)}$ represents the measurement error of the ith measurement channel of the sth scan of measurement type κ in the lth group; $\kappa = \text{pmu}, \text{scada}$; $i = 1, 2, \ldots, m_\kappa$; $s = 1, \ldots, S_\kappa$; $\boldsymbol{\theta}_{\kappa,i}$ represents the error distribution parameters of the ith measurement channel of measurement type κ; $K_{\kappa,i}$ is the number of components of the GMM distribution of the ith measurement channel of measurement type κ; $\phi_{\kappa,i,k}$, $\mu_{\kappa,i,k}$, and $\sigma_{\kappa,i,k}^2$ represent the mixture weight, the mean, and the variance of the kth component of the ith measurement channel of measurement type κ.

In reality, if a measurement channel is precalibrated, the total mean of the error probability distribution of the channel is close-to-zero. This additional information can be helpful in measurement error parameter estimation (EPE). If the measurement channel is not precalibrated, this condition does not hold, i.e., there may be a nonzero systematic bias in the measurement channel, which must be taken into account in PSSE.

The pdfs of the measurements can be expressed as

$$p\left(z_{\kappa,i}^{(l,s)} \middle| \boldsymbol{x}^{(l)}, \boldsymbol{\theta}_{\kappa,i} \right) = \sum_{k=1}^{K_{\kappa,i}} \phi_{\kappa,i,k} N\left(z_{\kappa,i}^{(l,s)} \middle| h_{\kappa,i}\left(\boldsymbol{x}^{(l)} \right) + \mu_{\kappa,i,k}, \sigma_{\kappa,i,k}^2 \right) \tag{7}$$

where

$$\boldsymbol{\theta}_{\kappa,i} = \left[\phi_{\kappa,i,1}, \mu_{\kappa,i,1}, \sigma_{\kappa,i,1}^2, \ldots, \phi_{\kappa,i,k}, \mu_{\kappa,i,k}, \sigma_{\kappa,i,k}^2, \ldots, \phi_{\kappa,i,K_{\kappa,i}}, \mu_{\kappa,i,K_{\kappa,i}}, \sigma_{\kappa,i,K_{\kappa,i}}^2 \right] \tag{8}$$

$$\boldsymbol{\theta} = \left\{ \boldsymbol{\theta}_{\text{pmu},1}, \boldsymbol{\theta}_{\text{pmu},2}, \ldots, \boldsymbol{\theta}_{\text{pmu},m_{\text{pmu}}}, \boldsymbol{\theta}_{\text{scada},1}, \boldsymbol{\theta}_{\text{scada},2}, \ldots, \boldsymbol{\theta}_{\text{scada},m_{\text{scada}}} \right\}. \tag{9}$$

$z_{\kappa,i}^{(l,s)}$ represents the measurement data of the ith measurement channel of the sth scan of measurement type κ in the lth group κ, and $h_{\kappa,i}(\cdot)$ is the relation between the state variables $\boldsymbol{x}^{(l)}$ and the measurement $z_{\kappa,i}^{(l,s)}$.

IV. Framework of ASE

In this section, the framework of the proposed PSSE paradigm will be presented. The framework consists of two parts: near-real-time EPE, i.e., Algorithm 2, and real-time ASE, i.e., Algorithm 3. The two parts are executed in different time intervals: the *previous interval* and the *current interval*. They refer to the time intervals where measurements are collected to perform EPE and ASE, respectively. The measurement data sampling scheme in the two different time intervals and the high-level framework of the proposed SE paradigm is shown in Fig. 2.

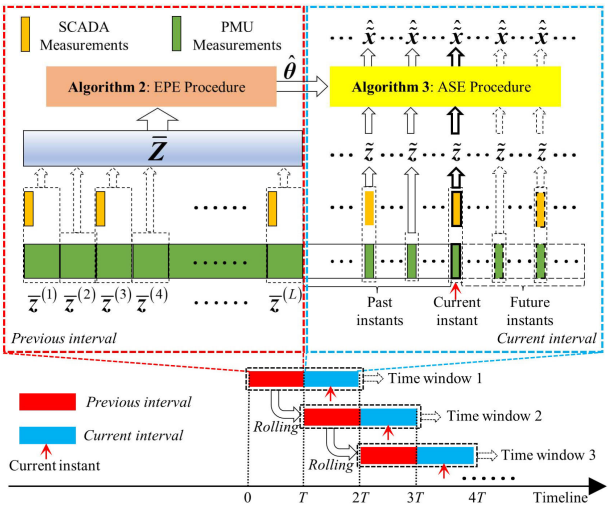

Fig. 2. Data sampling scheme and framework of the proposed paradigm.

Algorithm 1 High-Level Description of Proposed Method

Input: T, $\bar{\boldsymbol{Z}}$, and $\tilde{\boldsymbol{z}}$
Output: $\hat{\tilde{\boldsymbol{x}}}$
Set counter $\zeta = 1$ and time instant $t = T$
repeat
  Consolidate measurement data in the *previous interval* $[(\zeta - 1) \cdot T + 1, \zeta \cdot T]$ to form the measurement set $\bar{\boldsymbol{Z}}$
  Obtain parameter estimates of the GMM model representing measurement error distributions, $\hat{\boldsymbol{\theta}}$, by **Algorithm 2** // *Error Parameter Estimation (EPE) procedure described in Section IV*
  while $t \leq (\zeta + 1) \cdot T$ **do**
    Increase the time instant index t by 1, $t \leftarrow t + 1$
    Intake measurement data at the current instant t in the *current interval* $[\zeta \cdot T + 1, (\zeta + 1) \cdot T]$ to form the measurement scan $\tilde{\boldsymbol{z}}$
    Obtain state estimates $\hat{\tilde{\boldsymbol{x}}}$ of the current instant by **Algorithm 3** // *Adaptive State Estimation (ASE) procedure described in Section V*
  end
  Increase the counter ζ by 1, $\zeta \leftarrow \zeta + 1$
until
return state estimates $\hat{\tilde{\boldsymbol{x}}}$

Algorithm 2 EPE Procedure

Input: $\bar{\boldsymbol{Z}}$, ε_{SM}, and ε_θ
Output: $\hat{\boldsymbol{\theta}}$
Initialize the state variable $\bar{\boldsymbol{x}}_0$ by the WLS estimator
$\bar{\boldsymbol{x}} \leftarrow \bar{\boldsymbol{x}}_0$
for each measurement $i = 1, 2, \ldots, m_\kappa$;
$\kappa = scada, pmu$ **do**
  Set the initial number of Gaussian components $K_{\kappa,i} = 1$
  Set the initial similarity $SM_{\kappa,i}^{(K_{\kappa,i})} = 0$
  // *Outer loop: Determination of the proper number of Gaussian components described in Section V-A*
  while $SM_{\kappa,i}^{(K_{\kappa,i})} \leq \varepsilon_{SM}$ **do**
    Set $\left\|\Delta\hat{\boldsymbol{\theta}}\right\| = 10\varepsilon_\theta$
    Compute the initial measurement residuals

$$\bar{r}_{\kappa,i,0}^{(l,s)} = \bar{z}_{\kappa,i}^{(l,s)} - h_{\kappa,i}\left(\bar{\boldsymbol{x}}^{(l)}\right)$$

    Obtain the initial error distribution parameter $\boldsymbol{\theta}_{\kappa,i,0}^{(K_{\kappa,i})}$ by *k-means algorithm* with $\bar{r}_{\kappa,i,0}^{(l,s)}$ and $K_{\kappa,i}$
    $\boldsymbol{\theta}_{\kappa,i}^{(K_{\kappa,i})} \leftarrow \boldsymbol{\theta}_{\kappa,i,0}^{(K_{\kappa,i})}$
    // *Inner loop: EM algorithm for error distribution parameter estimation described in Section V-B*
    while $\|\Delta\hat{\boldsymbol{\theta}}\| > \varepsilon_\theta$ **do**
      // E step
      Compute $\bar{w}_{\kappa,i,k}^{(l,s)}$ by Eq. (20) with $\bar{\boldsymbol{x}}$, $\boldsymbol{\theta}_{\kappa,i}^{(K_{\kappa,i})}$, and $\bar{z}_{\kappa,i}^{(l,s)}$
      // M step
      Solve $\hat{\tilde{\boldsymbol{x}}}$ by Eq. (23) with $\bar{w}_{\kappa,i,k}^{(l,s)}$, $\boldsymbol{\theta}_{\kappa,i}^{(K_{\kappa,i})}$, and $\bar{z}_{\kappa,i}^{(l,s)}$
      Compute $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i})}$ by Eqs. (24) $\sim$ (28) with $\bar{w}_{\kappa,i,k}^{(l,s)}$, $\bar{\boldsymbol{x}}$, and $\bar{z}_{\kappa,i}^{(l,s)}$
      Compute $\Delta\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i})} - \boldsymbol{\theta}_{\kappa,i}^{(K_{\kappa,i})}$
      Update the values of parameters $\boldsymbol{\theta}_{\kappa,i}^{(K_{\kappa,i})} \leftarrow \hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i})}$
      Update the values of state variables $\bar{\boldsymbol{x}} \leftarrow \hat{\tilde{\boldsymbol{x}}}$
    end
    Compute the similarity $SM_{\kappa,i}^{(K_{\kappa,i})}$ by Eq. (16) with $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i})}$ and $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i}-1)}$ when $K_{\kappa,i} > 1$
    Increase the number of Gaussian components $K_{\kappa,i}$ by 1
    Update the values of state variables $\bar{\boldsymbol{x}} \leftarrow \hat{\tilde{\boldsymbol{x}}}$
  end
end
Derive the error distribution parameter estimates $\hat{\boldsymbol{\theta}} \leftarrow \hat{\boldsymbol{\theta}}^{(K_{\kappa,i}-1)}$
return $\hat{\boldsymbol{\theta}}$

The EPE aims to capture the measurement error distribution, i.e., the structure and parameters of the GMM in the *previous interval*. The ASE, on the other hand, aims to estimate the power system state variables at the current time instant lying within the *current interval* by using the up-to-date measurement error distribution information captured from the *previous interval* by EPE. To effectively estimate the measurement error distribution, the measurement redundancy required by EPE is much higher than that required by ASE [40]. The measurement set including L measurement groups in the *previous interval* (i.e., multiple measurement scans) is utilized in an aggregated fashion in one EPE execution. By contrast, only a single measurement scan, i.e., the most recent one collected in the *current interval*, is utilized in the ASE execution for the current instant.

Let $\bar{\boldsymbol{Z}}$ and $\hat{\boldsymbol{\theta}}$ represent the measurement set and the parameter estimates of the GMM model representing the measurement

Algorithm 3 ASE Procedure

Input: $\tilde{z}$, $\hat{\boldsymbol{\theta}}$, σ_{art}, ϕ_{art}, and ε_x
Output: $\hat{\hat{\boldsymbol{x}}}$
Initialize the state variables $\tilde{\boldsymbol{x}}_0$ by the WLS estimator
$\tilde{\boldsymbol{x}} \leftarrow \tilde{\boldsymbol{x}}_0$
Set $\left\|\Delta\hat{\hat{\boldsymbol{x}}}\right\| = 10\varepsilon_x$
while $\left\|\Delta\hat{\hat{\boldsymbol{x}}}\right\| > \varepsilon_x$ **do**
  // E step
  Compute $\tilde{w}'_{\kappa,i,k}$ and $\tilde{w}_{\kappa,i,art}$ by Eqs. (35) $\sim$ (36) with $\tilde{\boldsymbol{z}}$, $\tilde{\boldsymbol{x}}$, $\hat{\boldsymbol{\theta}}$, σ_{art}, and ϕ_{art}
  // M step
  Solve state estimates $\hat{\hat{\boldsymbol{x}}}$ by Eq. (34) with $\tilde{w}'_{\kappa,i,k}$, $\tilde{w}_{\kappa,i,art}$, $\tilde{\boldsymbol{z}}$, $\hat{\boldsymbol{\theta}}$, σ_{art}, and ϕ_{art}
  Compute the bias vector of state estimates $\Delta\hat{\hat{\boldsymbol{x}}} = \hat{\hat{\boldsymbol{x}}} - \tilde{\boldsymbol{x}}$
  Update the values of state variables $\tilde{\boldsymbol{x}} \leftarrow \hat{\hat{\boldsymbol{x}}}$
end
return $\hat{\hat{\boldsymbol{x}}}$

error distribution obtained from EPE, respectively, and let $\tilde{\boldsymbol{z}}$ and $\hat{\hat{\boldsymbol{x}}}$ represent the measurement scan and the state estimates obtained from ASE, respectively. A high-level overview of the proposed SE paradigm is provided in Algorithm 1, where T is the length of the time interval and ζ is a counting variable for the number of intervals passed in the algorithmic execution of the proposed procedure. The detailed procedures of EPE and ASE will be discussed in Sections V and VI, respectively.

From Algorithm 1, it can be clearly seen that the EPE procedure always consolidates all the measurement scans in the *previous interval* $[(\zeta-1)\cdot T+1, \zeta\cdot T]$ to estimate the parameters of measurement error statistics; while the ASE procedure always takes the latest received measurement scan at the current time instant t, which lies in the *current interval* $[\zeta\cdot T+1, (\zeta+1)\cdot T]$, to estimate the instantaneous state variables. The linkage between the EPE and the ASE is that the EPE procedure passes the updated knowledge about the measurement error statistics, i.e., the optimal structure and parameter estimates of GMM, to the ASE procedure to keep the near-optimal SE performance.

V. Capturing Measurement Error Statistics

The key to maintaining near-optimal SE performance is to keep track of the measurement error statistics. This section will discuss the EPE procedure for this purpose. As the outcomes of EPE will also reveal the systematic biases in the measurement channels, a byproduct of this procedure is the calibration of measurement channels for both SCADA and PMUs, which will also be discussed in this section.

To obtain the optimal structure and accurate parameters of GMM for the error distribution of each channel, $\boldsymbol{\theta}$, the joint MLE problem can be formulated in terms of state variables $\bar{\boldsymbol{x}}$ and parameters $\boldsymbol{\theta}$. The log-likelihood function is as follows:

$$
\begin{aligned}
&L\left(\bar{\boldsymbol{x}}, \boldsymbol{\theta} \middle| \bar{\boldsymbol{Z}}\right) \\
&= \log\left[\prod_{l=1}^{L}\prod_{\substack{\kappa=\text{pmu},\\ \text{scada}}}\prod_{s=1}^{S_\kappa}\prod_{i=1}^{m_\kappa} p\left(\bar{z}_{\kappa,i}^{(l,s)} \middle| \bar{\boldsymbol{x}}^{(l)}, \boldsymbol{\theta}_{\kappa,i}\right)\right] \\
&= \sum_{l=1}^{L}\sum_{\substack{\kappa=\text{pmu},\\ \text{scada}}}\sum_{s=1}^{S_\kappa}\sum_{i=1}^{m_\kappa} \log\left[\sum_{k=1}^{K_{\kappa,i}} \phi_{\kappa,i,k} N\left(\bar{z}_{\kappa,i}^{(l,s)} \middle| h_{\kappa,i}\left(\bar{\boldsymbol{x}}^{(l)}\right) + \mu_{\kappa,i,k}, \sigma^2_{\kappa,i,k}\right)\right].
\end{aligned} \tag{10}
$$

The goal of the MLE is to find the values of state variables $\bar{\boldsymbol{x}}$ and error distribution parameters $\boldsymbol{\theta}$ that maximize the log-likelihood function, which can be expressed as

$$
\left[\hat{\bar{\boldsymbol{x}}}, \hat{\boldsymbol{\theta}}\right] = \arg\max \hat{L}\left(\bar{\boldsymbol{x}}, \boldsymbol{\theta} \middle| \bar{\boldsymbol{Z}}\right). \tag{11}
$$

However, it is difficult to obtain the analytical solutions to the MLE problem in (11) because the logarithm of the sum of pdfs is present in (10). To solve the MLE problem, the EM algorithm is adopted and customized to obtain $\hat{\hat{\boldsymbol{x}}}$ and $\hat{\boldsymbol{\theta}}$. The EM algorithm is executed by iterating the so-called Q function between the expectation step (E step) and the maximization step (M step) to approach the true values of $\bar{\boldsymbol{x}}$ and $\boldsymbol{\theta}$.

As the first step of EPE, the initial values $\bar{\boldsymbol{x}}_0$ are estimated by using the WLS estimator; then, the initial measurement residuals $\bar{r}_{\kappa,i,0}^{(l,s)}$ can be computed by

$$
\bar{r}_{\kappa,i,0}^{(l,s)} = \bar{z}_{\kappa,i}^{(l,s)} - h_{\kappa,i}\left(\bar{\boldsymbol{x}}_0^{(l)}\right). \tag{12}
$$

The initial values of GMM parameters, $\boldsymbol{\theta}_0$, can be achieved by the k-means clustering algorithm [41] with $\bar{r}_{\kappa,i,0}^{(l,s)}$ and $K_{\kappa,i}$, where $K_{\kappa,i}$ is the number of Gaussian components of the GMM for the ith channel of measurement type κ. *First*, the k-means algorithm is adopted, where the input includes $\bar{r}_{\kappa,i,0}^{(l,s)}$ and $K_{\kappa,i}$ and the output is the cluster index variable *idx*; *Second*, the parameters of the GMM can be estimated based on the clustered measurement residuals via the following equations:

$$
\hat{\phi}_{\kappa,i,k,0} = \frac{\text{count}\,(idx=k)}{S_\kappa}, \quad k=1,2,\ldots,\ K_{\kappa,i} \tag{13}
$$

$$
\hat{\mu}_{\kappa,i,k,0} = \text{mean}\left(\bar{r}_{\kappa,i,0}^{idx=k}\right), \quad k=1,2,\ldots,\ K_{\kappa,i} \tag{14}
$$

$$
\hat{\sigma}^2_{\kappa,i,k,0} = \text{var}\left(\bar{r}_{\kappa,i,0}^{idx=k}\right), \quad k=1,2,\ldots,K_{\kappa,i} \tag{15}
$$

where count(.) is a counting function, which is used to count the number of measurement residuals that are clustered into the kth component; mean(.) is a function to calculate the mean of the clustered measurement residuals; var(.) is a function to calculate the variance of the clustered measurement residuals; S_κ is the number of measurement scans of the measurement type κ; $K_{\kappa,i}$ is the number of Gaussian components of the ith measurement channels of the measurement type κ; and $\bar{r}_{\kappa,i,0}^{idx=k}$ represents the measurement residuals that are clustered into the kth component. *Finally*, the initial parameter estimates of the GMM can be obtained, i.e., $\boldsymbol{\theta}_{\kappa,i,0}^{(K_{\kappa,i})} = \{\hat{\phi}_{\kappa,i,k,0}, \hat{\mu}_{\kappa,i,k,0}, \hat{\sigma}^2_{\kappa,i,k,0}\}$.

A. Determination of the Number of Gaussian Components

The number of Gaussian components (i.e., $K_{\kappa,i}$) determines the structure of GMM. It is critical for capturing the unknown measurement error distribution and should be properly selected.

The search for the appropriate number of Gaussian components $K_{\kappa,i}$ goes as follows. It starts from $K_{\kappa,i} = 1$, where there is only one Gaussian component and GMM is reduced to a regular Gaussian distribution. Then, $K_{\kappa,i}$ will be increased gradually, and the parameter estimates $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{k,i})}$ under each value of $K_{\kappa,i}$ will be obtained as described in Section V-B. Comparing the parameter estimates $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{k,i})}$ obtained under $K_{\kappa,i}$ components with those $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{k,i}-1)}$ obtained under $K_{\kappa,i}-1$ components, if there is a high similarity between the resulting pdfs, it implies that continuing the increase of $K_{\kappa,i}$ will not significantly improve the fitting performance; hence, the EPE can stop increasing $K_{\kappa,i}$ and output the final parameter estimates $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{k,i}-1)}$.

The cosine similarity between the two pdfs based on the parameter estimates under $K_{\kappa,i}$ Gaussian components and $K_{\kappa,i}-1$ Gaussian components can be quantified as follows [42]: (16) and (17), as shown at the bottom of the next page, where $\mathrm{SM}_{\kappa,i}^{(K_{\kappa,i})}$ represents the similarity of two pdfs based on $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i})}$ and $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i}-1)}$, respectively; $\boldsymbol{Y}$ is a vector consisting of a set of samples, which are utilized for computing the values of pdfs; D is the number of the samples; $P(y_j|\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i})})$ and $P(y_j|\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i}-1)})$ represent the values of the two probability densities with respect to y_i, $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i})}$ and $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i}-1)}$.

B. EPE Based on Customized EM Algorithm

In this section, the procedure for obtaining the parameter estimates $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{k,i})}$ given the structure of GMM, $K_{\kappa,i}$, will be presented. The method is based on a customized EM algorithm and can be divided into the E step and the M step.

1) E Step: The E step aims to develop the Q function based on the likelihood function. The measurement data $\bar{z}_{\kappa,i}^{(l,s)}$, referred to as the observation data, is known as the input. However, it is unknown which Gaussian component the measurement error $\bar{e}_{\kappa,i}^{(l,s)}$ comes from. Hence, the latent variable is defined as follows:

$$\bar{\gamma}_{\kappa,i,k}^{(l,s)} = \begin{cases} 1, & \text{if } \bar{e}_{\kappa,i}^{(l,s)} \text{ comes from the } k\text{th component} \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

In the E step, the values of state variables $\bar{\boldsymbol{x}}$ and error distribution parameters $\boldsymbol{\theta}$ will be used to compute the posterior probability. For the first iteration, the WLS estimator and the k-means algorithm are utilized for the initializations of $\bar{\boldsymbol{x}}$ and $\boldsymbol{\theta}$. For the future iterations, the values of $\bar{\boldsymbol{x}}$ and $\boldsymbol{\theta}$ are derived from the last iteration. The Q function is defined as follows:

$$Q = \sum_{l=1}^{L} \sum_{\substack{\kappa=\text{pmu,}\\ \text{scada}}} \sum_{s=1}^{S_\kappa} \sum_{i=1}^{m_\kappa} \sum_{k=1}^{K_{\kappa,i}} \times \left\{ \bar{w}_{\kappa,i,k}^{(l,s)} \cdot \log\left[p\left(\bar{z}_{\kappa,i}^{(l,s)}, \bar{\gamma}_{\kappa,i,k}^{(l,s)} = 1 \middle| \bar{\boldsymbol{x}}^{(l)}, \boldsymbol{\theta}_{\kappa,i} \right) \right] \right\} \tag{19}$$

$$\bar{w}_{\kappa,i,k}^{(l,s)} = p\left(\bar{\gamma}_{\kappa,i,k}^{(l,s)} = 1 \middle| \bar{z}_{\kappa,i}^{(l,s)}, \bar{\boldsymbol{x}}^{(l)}, \boldsymbol{\theta}_{\kappa,i} \right) = \frac{\phi_{\kappa,i,k} N\left(\bar{z}_{\kappa,i}^{(l,s)} \middle| h_{\kappa,i}\left(\bar{\boldsymbol{x}}^{(l)}\right) + \mu_{\kappa,i,k}, \sigma_{\kappa,i,k}^2 \right)}{\sum_{k=1}^{K_{\kappa,i}} \phi_{\kappa,i,k} N\left(\bar{z}_{\kappa,i}^{(l,s)} \middle| h_{\kappa,i}\left(\bar{\boldsymbol{x}}^{(l)}\right) + \mu_{\kappa,i,k}, \sigma_{\kappa,i,k}^2 \right)} \tag{20}$$

where $\bar{w}_{\kappa,i,k}^{(l,s)}$ is the posterior probability, which is the probability that the measurement error $\bar{e}_{\kappa,i}^{(l,s)}$ comes from the kth GMM component with respect to parameters $\boldsymbol{\theta}_{\kappa,i}$.

2) M Step: The M step aims to estimate the values of $\bar{\boldsymbol{x}}$ and $\boldsymbol{\theta}$ by using the posterior probability, i.e., $\bar{w}_{\kappa,i,k}^{(l,s)}$, computed in the E step. The objective function is shown as follows:

$$\left[\hat{\bar{\boldsymbol{x}}}, \hat{\boldsymbol{\theta}}\right] = \arg\max_{\bar{\boldsymbol{x}},\boldsymbol{\theta}} Q(\bar{\boldsymbol{x}}, \boldsymbol{\theta}) = \arg\max_{\bar{\boldsymbol{x}},\boldsymbol{\theta}} \sum_{l=1}^{L} \sum_{\substack{\kappa=\text{pmu,}\\ \text{scada}}} \sum_{s=1}^{S_\kappa} \sum_{i=1}^{m_\kappa} \sum_{k=1}^{K_{\kappa,i}} \left\{ \bar{w}_{\kappa,i,k}^{(l,s)} \cdot \log\left[p\left(\bar{z}_{\kappa,i}^{(l,s)}, \bar{\gamma}_{\kappa,i,k}^{(l,s)} = 1 \middle| \bar{\boldsymbol{x}}^{(l)}, \boldsymbol{\theta}_{\kappa,i} \right) \right] \right\} \tag{21}$$

where

$$\log\left[p\left(\bar{z}_{\kappa,i}^{(l,s)}, \bar{\gamma}_{\kappa,i,k}^{(l,s)} = 1 \middle| \bar{\boldsymbol{x}}^{(l)}, \boldsymbol{\theta}_{\kappa,i} \right) \right] = -\frac{\left[\bar{z}_{\kappa,i}^{(l,s)} - h_{\kappa,i}\left(\bar{\boldsymbol{x}}^{(l)}\right) - \mu_{\kappa,i,k} \right]^2}{2\sigma_{\kappa,i,k}^2} + \log\left(\frac{\phi_{\kappa,i,k}}{\sqrt{2\pi}\sigma_{\kappa,i,k}} \right). \tag{22}$$

The maximization of $Q(\bar{\boldsymbol{x}}, \boldsymbol{\theta})$ is divided into two steps: state regression and error distribution parameter estimation. This method divides the M step into partial minimization steps and is called expectation conditional minimization (ECM) [43].

The state regression step aims to estimate the values of state variables $\bar{\boldsymbol{x}}$ based on the error distribution parameters $\boldsymbol{\theta}$ obtained from the last iteration and the values of $\bar{w}_{\kappa,i,k}^{(l,s)}$ obtained from the E step. The maximization of (21) with respect to state variables $\bar{\boldsymbol{x}}$ leads to a convex optimization problem

$$\hat{\bar{\boldsymbol{x}}} = \arg\min_{\bar{\boldsymbol{x}}} \sum_{l=1}^{L} \sum_{\substack{\kappa=\text{pmu,}\\ \text{scada}}} \sum_{s=1}^{S_\kappa} \sum_{i=1}^{m_k} \sum_{k=1}^{K_{\kappa,i}} \times \left\{ \frac{\bar{w}_{\kappa,i,k}^{(l,s)}}{2\sigma_{\kappa,i,k}^2} \cdot \left[\bar{z}_{\kappa,i}^{(l,s)} - h_{\kappa,i}\left(\bar{\boldsymbol{x}}^{(l)}\right) - \mu_{\kappa,i,k} \right]^2 \right\}. \tag{23}$$

The optimization problem (23) can be converted into a least squares problem, which can be efficiently solved.

The error distribution parameters $\boldsymbol{\theta}$ can be estimated by differentiating (21) with respect to $\phi_{\kappa,i,k}$, $\mu_{\kappa,i,k}$, and $\sigma_{\kappa,i,k}$, and equating them to zero based on the values of $\bar{\boldsymbol{x}}$ obtained from the last iteration and the values of $\bar{w}_{\kappa,i,k}^{(l,s)}$ obtained from the E step. The results are shown as follows:

$$\hat{\phi}_{\kappa,i,k} = \left(\sum_{l=1}^{L} S_\kappa^{(l)} \right)^{-1} \cdot \sum_{l=1}^{L} \sum_{s=1}^{S_\kappa} \bar{w}_{\kappa,i,k}^{(l,s)} \tag{24}$$

$$\hat{\mu}_{\kappa,i,k} = \left(\sum_{l=1}^{L}\sum_{s=1}^{S_\kappa} \bar{w}_{\kappa,i,k}^{(l,s)}\right)^{-1} \cdot \sum_{l=1}^{L}\sum_{s=1}^{S_\kappa}\left[\bar{w}_{\kappa,i,k}^{(l,s)} \cdot \left(\bar{z}_{\kappa,i}^{(l,s)} - h_{\kappa,i}\left(\bar{\boldsymbol{x}}^{(l)}\right)\right)\right] \tag{25}$$

$$\hat{\sigma}_{\kappa,i,k}^2 = \left(\sum_{l=1}^{L}\sum_{s=1}^{S_\kappa} \bar{w}_{\kappa,i,k}^{(l,s)}\right)^{-1} \cdot \sum_{l=1}^{L}\sum_{s=1}^{S_\kappa}\left\{\bar{w}_{\kappa,i,k}^{(l,s)}\left[\bar{z}_{\kappa,i}^{(l,s)} - h_{\kappa,i}\left(\bar{\boldsymbol{x}}^{(l)}\right) - \mu_{\kappa,i,k}\right]^2\right\} \tag{26}$$

where $S_\kappa^{(l)}$ is the number of measurement scans of measurement type κ in the lth measurement group.

The detailed EPE procedure is provided in Algorithm 2. It should be noted that the procedure consists of two loops: an *inner loop procedure* and *an outer loop procedure*. The EPE given the number of Gaussian components, i.e., $K_{\kappa,i}$, is an *inner loop procedure* (Section V-B), while the determination of the number of Gaussian components is an *outer loop procedure* (Section V-A). The *inner loop procedure* aims to obtain the parameter estimates $\hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{k,i})}$ given the structure of GMM, i.e., $K_{\kappa,i}$. The *outer loop procedure* aims to determine the structure of GMM, i.e., to obtain the number of Gaussian components. The threshold for termination $\varepsilon_{\mathrm{SM}}$ can be set empirically; we recommend setting it in the range of 0.96–0.99, which will ensure that further incrementation of $K_{\kappa,i}$ brings about little benefit and the procedure can be terminated.

C. Sensor Calibration Based on EPE

From the sensor calibration perspective, the measurement errors can be viewed as the superposition of two components: a stochastic component with zero mean (stochastic error) and a deterministic component (systematic bias). The objective of sensor calibration is to change the internal settings of the sensor to offset the deterministic component (systematic bias) [8]. In this article, we propose a remote calibration method for measurement channels based on the EPE outcomes, which does not require traditional calibration processes. To distinguish between the traditional calibration and the proposed calibration based on EPE, *field calibration*, and *algorithmic calibration* are introduced in this section. *Field calibration* refers to the process of calibrating devices in the field, including calibrations of instrument transformers, control cables, burdens, PMUs, etc. *Algorithmic calibration* refers to the process of calibrating measurement channels by using the total mean of measurement error distributions estimated from the EPE procedure. For a particular EPE cycle, if a measurement channel has been field-calibrated in advance, it is referred to as *field precalibrated*; if it has gone through an *algorithmic calibration* via a previous EPE cycle, it is referred to as *algorithmic precalibrated*. For both *field precalibrated channels* and *algorithmic precalibrated channels*, systematic biases are commonly compensated, so they can be significantly reduced [44], [45], [46], [47], exhibiting relatively close-to-zero mean values for measurement error distributions. Therefore, they are collectively referred to as *precalibrated channels* for the given EPE cycle. If a measurement channel does not belong to either of the situations above, it is referred to as a *nonprecalibrated* channel.

As accurate mean values of the *precalibrated* channels are unavailable, but known to be close-to-zero, i.e., $\mu_{\kappa,i}^* \approx 0$, and significantly smaller than that of the *nonprecalibrated* channels, we approximate them as 0 to enhance the information redundancy, such that *nonprecalibrated* channels with large systematic errors can be calibrated more accurately via the EPE procedure. In the EPE procedure, the additional information that *precalibrated* measurement channels have close-to-zero means can be used to correct the estimate of the mean of each Gaussian component, i.e., $\hat{\mu}_{\kappa,i,k}$, by

$$\Delta\hat{\mu}_{\kappa,i} = \begin{cases} 0, & \text{non-pre-calibrated channel} \\ \mu_{\kappa,i}^* - \sum_{k=1}^{K_{\kappa,i}} \hat{\phi}_{\kappa,i,k} \cdot \hat{\mu}_{\kappa,i,k}, & \text{pre-calibrated channel} \end{cases}$$
$$\Downarrow \mu_{\kappa,i}^* \approx 0 \tag{27}$$
$$\Delta\hat{\mu}_{\kappa,i} = \begin{cases} 0, & \text{non-pre-calibrated channel} \\ 0 - \sum_{k=1}^{K_{\kappa,i}} \hat{\phi}_{\kappa,i,k} \cdot \hat{\mu}_{\kappa,i,k}, & \text{pre-calibrated channel} \end{cases}$$
$$\hat{\mu}'_{\kappa,i,k} = \hat{\mu}_{\kappa,i,k} + \Delta\hat{\mu}_{\kappa,i} \tag{28}$$

where $\hat{\mu}_{\kappa,i,k}$ and $\hat{\mu}'_{\kappa,i,k}$ are the uncorrected and corrected estimates of the mean of the kth component of the ith measurement channel of measurement type κ, respectively; and $\mu_{\kappa,i}^*$ is the actual mean of the *precalibrated* channels, which is unknown but close to 0 compared with those of the *nonprecalibrated* channels. The corrected estimates of the mean, i.e., $\hat{\mu}'_{\kappa,i,k}$, are used during iterations of the EPE procedure. In other words, once the mean is estimated by (25), it will be further corrected by (27) and (28); then the corrected mean will be used in the next iteration.

$$\mathrm{SM}_{\kappa,i}^{(K_{\kappa,i})} = \begin{cases} 0, & K_{\kappa,i} = 1 \\ \dfrac{\sum_{j=1}^{D} P\left(y_j \middle| \hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i})}\right) P\left(y_j \middle| \hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i}-1)}\right)}{\sqrt{\sum_{j=1}^{D}\left[P\left(y_j \middle| \hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i})}\right)\right]^2} \cdot \sqrt{\sum_{j=1}^{D}\left[P\left(y_j \middle| \hat{\boldsymbol{\theta}}_{\kappa,i}^{(K_{\kappa,i}-1)}\right)\right]^2}}, & K_{\kappa,i} > 1 \end{cases} \tag{16}$$

$$\boldsymbol{Y} = \left[y_1, y_2, \ldots, y_j, \ldots, y_D\right], \quad j = 1, 2, \ldots, D \tag{17}$$

As a result, the final estimated total mean of the GMM distribution of a *precalibrated* channel will be approximated by zero, while that of a *nonprecalibrated* measurement channel will be nonzero. The estimated total mean of the GMM distribution of the ith measurement channel without *precalibration* is given by

$$\hat{\mu}_{\kappa,i} = \sum_{k=1}^{K_{\kappa,k}} \hat{\phi}_{\kappa,i,k} \cdot \hat{\mu}_{\kappa,i,k}. \tag{29}$$

This information can be used to calibrate the *nonprecalibrated* channels. When these channels are properly calibrated by (29), in the next EPE cycle, they will be deemed *algorithmic precalibrated*, which will become part of the *precalibrated* measurement channel set.

As the proposed *algorithmic calibration* is carried out by (29), all channels will eventually become *precalibrated*. However, the calibration of measurement channels needs to be performed periodically. As time goes by, both *field precalibrated* channels and *algorithmic precalibrated* channels may experience sensor drafts due to operating condition changes, aging, or other factors. Therefore, calibration needs to be performed again if: 1) a measurement channel has not been calibrated for a long time and 2) if a measurement channel exhibits significant (or dubious) characteristic changes, for example, a significant shift of measurement residual distribution. In either of these two cases, the measurement channel may not be trustworthy anymore, and it should be put back into the *nonprecalibrated* set to undergo a new round of *algorithmic calibration* by (29) based on the proposed EPE procedure.

As will be shown in the simulation results in Section VII, the EPE can not only benefit the ASE that immediately follows, but also enable highly accurate calibration of SCADA and PMU channels to permanently remove the systematic biases.

VI. Online Adaptation of SE to Measurement Error Statistics

The structure $K_{\kappa,i}$ and parameter estimates of the developed GMM error distribution, $\hat{\boldsymbol{\theta}}$, are obtained by EPE in the *previous interval*. Then, parameter estimates will be used to perform the real-time SE of the power system in the *current interval*. In ASE, the procedure is also executed between the E step and the M step, and only the posterior probabilities $\tilde{w}_{\kappa,i,k}$ and state variables $\tilde{\boldsymbol{x}}$ need to be iteratively solved to achieve the near-optimal state estimates by using a single measurement scan that arrives mostly lately in the *current interval*.

A. ASE Based on Customized EM Algorithm

The ASE procedure based on the customized EM algorithm can be divided into the E step and the M step.

1) E step: The E step aims to develop the Q function and compute posterior probabilities $\tilde{w}_{\kappa,i,k}$ by using the error distribution parameter estimates $\hat{\boldsymbol{\theta}}$ obtained from EPE and the values of state variables $\tilde{\boldsymbol{x}}$ obtained from the last iteration. For the first iteration, the initial values of state variables, $\tilde{\boldsymbol{x}}_0$, are obtained from the WLS estimator. The Q function for ASE is as follows:

$$Q = \sum_{\substack{\kappa=\text{pmu,}\\ \text{scada}}} \sum_{i=1}^{m_\kappa} \sum_{k=1}^{K_{\kappa,i}} \times \left\{ \tilde{w}_{\kappa,i,k} \cdot \log \left[p\left(\tilde{z}_{\kappa,i}, \tilde{\gamma}_{\kappa,i,k} = 1 \mid \tilde{\boldsymbol{x}}, \hat{\boldsymbol{\theta}}_{\kappa,i} \right) \right] \right\} \tag{30}$$

where

$$\tilde{\gamma}_{\kappa,i,k} = \begin{cases} 1, & \text{if } \tilde{e}_{\kappa,i} \text{ comes from the } k\text{th component} \\ 0, & \text{otherwise} \end{cases} \tag{31}$$

$$\tilde{w}_{\kappa,i,k} = p\left(\tilde{\gamma}_{\kappa,i,k} = 1 \middle| \tilde{z}_{\kappa,i}, \tilde{\boldsymbol{x}}, \hat{\boldsymbol{\theta}}_{\kappa,i} \right) = \frac{\hat{\phi}_{\kappa,i,k} N\left(\tilde{z}_{\kappa,i} \middle| h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right) + \hat{\mu}_{\kappa,i,k}, \hat{\sigma}^2_{\kappa,i,k} \right)}{\sum_{k=1}^{K_{\kappa,i}} \hat{\phi}_{\kappa,i,k} N\left(\tilde{z}_{\kappa,i} \middle| h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right) + \hat{\mu}_{\kappa,i,k}, \hat{\sigma}^2_{\kappa,i,k} \right)}. \tag{32}$$

$\tilde{z}_{\kappa,i}$ and $\tilde{e}_{\kappa,i}$ are the ith measured value and its error of measurement type κ, respectively; $\tilde{\gamma}_{\kappa,i,k}$ is the latent variable defining whether the $\tilde{e}_{\kappa,i}$ comes from the kth component or not; and $\tilde{w}_{\kappa,i,k}$ is the posterior probability, which is the probability that $\tilde{e}_{\kappa,i}$ comes from the kth GMM component.

2) M step: The M step aims to estimate the state variables $\tilde{\boldsymbol{x}}$ by using the computed posterior probabilities $\tilde{w}_{\kappa,i,k}$ in the E step and error distribution parameter estimates $\hat{\boldsymbol{\theta}}$ obtained from the EPE procedure. The PSSE problem can be converted into a WLS problem, which is shown as follows:

$$\hat{\tilde{\boldsymbol{x}}} = \arg\min_{\tilde{\boldsymbol{x}}} \sum_{\substack{\kappa=\text{pmu,}\\ \text{scada}}} \sum_{i=1}^{m_\kappa} \sum_{k=1}^{K_{\kappa,i}} \left\{ \frac{\tilde{w}_{\kappa,i,k}}{2\hat{\sigma}^2_{\kappa,i,k}} \left[\tilde{z}_{\kappa,i} - h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right) - \hat{\mu}_{\kappa,i,k} \right]^2 \right\}. \tag{33}$$

Note that this problem is fundamentally different from the conventional WLS estimator, in that for the ith measurement of type κ, 1) there are $K_{\kappa,i}$ squared terms representing the $K_{\kappa,i}$ Gaussian components of its error; 2) the weight of each term, $\tilde{w}_{\kappa,i,k}/2\hat{\sigma}^2_{\kappa,i,k}$, is dynamically and adaptively computed; and 3) systematic errors (i.e., biases) of measurements are consistently corrected by $\hat{\mu}_{\kappa,i,k}$. However, as its unconstrained quadratic formulation resembles WLS SE, it is much more computationally efficient compared with the ASE formulation proposed in our previous work [20], which contains both square terms and absolute value terms in the objective function and must be converted into a constrained quadratic programming problem for solution.

B. Robustness Enhancement of Adaptive Estimator by Adding Gross Error Trap

In the developed SE paradigm, the parameter estimates of measurement error statistics in the *previous interval* are passed to the SE in the *current interval*. The underlying assumption is that the error distribution does not have an abrupt change between the *previous interval* and the *current interval*. In practice, however, abrupt gross error may appear in the case of sensor malfunctioning, communication interruption, or cyber-attacks [2], [7], [8], [28], [29], [30]. To address

this challenge, an additional artificial term, named the *gross error trap*, is introduced to the ASE. Abrupt gross errors are represented by an artificially introduced Gaussian component with a small mixture weight and a very large variance, which is not captured by EPE. The improved estimator is as follows: (34)–(36), as shown at the bottom of the next page, where ϕ_{art} is the mixture weight of the gross error trap; σ^2_{art} is the variance of the gross error trap; $\tilde{w}'_{\kappa,i,k}$ is the probability that the ith measurement error of measurement type κ comes from the kth mixture component; and $\tilde{w}_{\kappa,i,\text{art}}$ is the probability that the ith measurement error of measurement type κ comes from the gross error trap term. For example, if a gross error appears in the ith measurement channel of measurement type κ, the posterior probability $\tilde{w}'_{\kappa,i,k}$ will approach 0, and $\tilde{w}_{\kappa,i,\text{art}}$ will approach 1, i.e., the ASE will infer that the measurement comes from an artificial Gaussian component with a very large variance σ^2_{art}. As σ^2_{art} is much greater than the variances of other regular terms, $\hat{\sigma}^2_{\kappa,i,k}$, the contribution of the measurement with gross error to the objective function of (34) will be significantly suppressed. Therefore, the outcomes of ASE will be insensitive to measurements with large abrupt gross errors, even if they are not captured by the preceding EPE procedure. The detailed ASE procedure is provided in Algorithm 3.

VII. Simulation Results

To evaluate the performance of the proposed SE paradigm under a complicated measurement environment, comprehensive simulations are carried out on the IEEE 30-bus power system. About 24 PMUs are deployed across the system except for buses 3, 7, 14, 17, 19, and 26, which collect 24 voltage magnitude measurements, 24 voltage phase angle measurements, 71 real parts of current phasor measurements, and 71 imaginary parts of current phasor measurements, and 110 SCADA measurements including ten voltage magnitude measurements, 20 pairs of real and reactive power injection measurements, and 30 pairs of real and reactive power flow measurements are considered. Besides, it is assumed that 70% of measurement channels, including 133 PMU measurement channels and 77 SCADA measurement channels are *field-precalibrated*. In other words, the mean values of the error distributions of 70% measurement channels are close-to-zero, and those of the remaining 30% channels are significantly nonzero. Please note that the number of *field precalibrated* measurement channels hardly impacts the computational burden of the EPE procedure since the parameters of measurement error distributions of all measurement channels, including both *field precalibrated* and *field nonprecalibrated* channels, are estimated in the EPE procedure. The length of EPE and ASE time intervals are set as 30 min. The simulation is conducted for 24 h.

The effectiveness of the developed EPE and ASE will be verified under different measurement error conditions. In this section, measurement errors are synthesized based on the developed realistic SCADA and PMU measurement chains as discussed in Section II. The simulation is based on the real-world power flow trend derived from the ISO New England load data on June 1st, 2023 [48]. The details of the SCADA and PMU measurement chain models are described in a document online [35] due to the limited space of the article. The parameter settings for each component in the measurement chains are as follows.

1) In the SCADA measurement chain modeling, the parameter settings are tailored for VTs/CTs, IEDs, and communication networks. For VTs and CTs, the accuracy class is 0.6. Hence, the limits of uniform distributions for systematic errors can be obtained from the IEEE Std C57. 13[1]-2016 [23]. The random errors are set to follow GMM distributions with four components, with the total mean and total variance of random errors set to be 0 and 4×10^{-4} p.u. for ratio errors and 0 and 1° for phase angle error, respectively, and the Kullback-Leibler divergence (KLD) [42] with a Gaussian distribution, a "Gaussianity" metric, is set to follow a uniform distribution ranging from 0.8 to 1.0 for all measurement channels. These choices are based on the following rationales: 1) the errors have a high-to-median level of similarity to Gaussian distributions; 2) the number of Gaussian components, i.e., K, is set to 4 to mimic the causes of the multipeak and/or skewed measurement error distributions reported in [4], [5], and [6]; and 3) the total variance of the error is set as a fraction of the systematic errors, as for VTs and CTs, systematic errors play a more significant role compared with random errors. For control cables and burdens, typically, a 500 ft. RG-8 cable introduces a 0.4° phase angle error [24]. In some cases, the length of the cable can reach 3000 ft., which will cause an even larger phase angle error [24]. Hence, the ratio errors follow a Gaussian distribution with a mean of 0.002 p.u. and a standard deviation of 0.003 p.u., while the phase angle errors follow a Gaussian distribution with a mean of 0.2° and a standard deviation of 0.33°. For IEDs, the mean and variance of the Gaussian distribution for each type of measurement are set to be 0 and 1×10^{-6} p.u., respectively. For communication networks, the measurement errors are assumed to follow GMM distributions with four components and the Gaussian similarity is assumed to follow a uniform distribution ranging from 0.8 to 1.0 for all measurement channels. The total mean and total variance of the GMM distributions are set to be 0 and 1×10^{-4} p.u., respectively.
2) In the PMU measurement chain modeling, the parameter settings are tailored for VTs/CTs, control cables and burden, and the phasor estimation procedure. For VTs, CTs, and control cables and burdens, the parameters are set the same as those in the SCADA measurement chain modeling. For the phasor estimation procedure, the DFT algorithm with filters is used to estimate phasor measurements [31]. Specifically, it is assumed that the nominal frequency of input signals is 60 Hz. The off-nominal frequency of input signals follows a uniform distribution ranging from 59.8 to 60.2 Hz, which is used to mimic the realistic behavior of the power grid frequency. The reporting rate of PMUs is 60 Hz. The sampling rate of

[1]Trademarked.

the DFT algorithm is 16 sampled values per cycle, i.e., 960 Hz. The frequency of the oscillator is 20 MHz. The oscillator frequency drift is 0.15 μs/PPS.

Furthermore, abrupt gross errors are introduced in the 30th and 40th time intervals. For SCADA measurements, gross errors are introduced by abnormally long latency in communication networks. For PMU measurements, gross errors are introduced by GPS signal loss. The EPE tolerances for $\phi_{\kappa,i,k}$, $\mu_{\kappa,i,k}$, $\sigma_{\kappa,i,k}$, and $\mathrm{SM}_{\kappa,i}^{(K_{\kappa,i})}$ are set as $\varepsilon_\phi = 10^{-3}$, $\varepsilon_\mu = 10^{-4}$, $\varepsilon_\sigma = 10^{-4}$, and $\varepsilon_{\mathrm{SM}} = 0.98$, respectively. In ASE, the parameters of the gross error trap are set as $\phi_{\mathrm{art}} = 0.01$ and $\sigma_{\mathrm{art}} = 10$, respectively. The state regression tolerance in the EPE and ASE are set as $\varepsilon_x = 10^{-5}$.

In order to evaluate the performance of EPE, we need to quantify how close the estimated measurement error distributions are to the true distributions for all measurement channels. The Cosine similarity measure is used to evaluate the accuracy of outcomes of EPE. It evaluates the similarity between the pdfs of the true distribution and the estimated distribution of the errors from a measurement channel [42]

$$\mathrm{SM}_{\kappa,i}^{\mathrm{EPE}} = \frac{\sum_{j=1}^{U} P\left(o_j \middle| \hat{\boldsymbol{\theta}}_{\kappa,i}\right) \cdot P\left(o_j \middle| \boldsymbol{\theta}_{\kappa,i}\right)}{\sqrt{\sum_{j=1}^{U}\left[P\left(o_j \middle| \hat{\boldsymbol{\theta}}_{\kappa,i}\right)\right]^2} \cdot \sqrt{\sum_{j=1}^{U}\left[P\left(o_j \middle| \boldsymbol{\theta}_{\kappa,i}\right)\right]^2}} \tag{37}$$

$$\boldsymbol{O} = \left[o_1, o_2, \ldots, o_j, \ldots, o_U\right], \quad j = 1, 2, \ldots, U \tag{38}$$

where $\mathrm{SM}_{\kappa,i}^{\mathrm{EPE}}$ is the similarity between the pdf based on true error distribution parameters $\boldsymbol{\theta}_{\kappa,i}$ and the pdf based on estimated error distribution parameters $\hat{\boldsymbol{\theta}}_{\kappa,i}$; $P(o_j|\hat{\boldsymbol{\theta}}_{\kappa,i})$ and $P(o_j|\boldsymbol{\theta}_{\kappa,i})$ represent the probabilities of the error distributions with respect to $\hat{\boldsymbol{\theta}}_{\kappa,i}$, $\boldsymbol{\theta}_{\kappa,i}$, and o_j; $\boldsymbol{O}$ is a set of sample data, which is used to compute the probabilities of pdfs; and U is the number of sample data. The closer the similarity metric is to 1, the better the performance of the EPE procedure.

In order to evaluate the performance of ASE, we need to quantify how close the state estimates are to the true states. The mean absolute errors (MAEs) are used to evaluate the performance of the developed ASE, defined as follows:

$$\mathrm{MAE}_V = \frac{1}{n}\sum_{i=1}^{n}\left|V_i - \hat{V}_i\right| \tag{39}$$

$$\mathrm{MAE}_\delta = \frac{1}{n}\sum_{i=1}^{n}\left|\delta_i - \hat{\delta}_i\right| \tag{40}$$

where MAE_V and MAE_δ represent the MAEs of estimated voltage magnitudes and estimated phase angles, respectively; the variables with and without a hat represent the estimated and true values of state variables, respectively. The performance of the developed ASE is compared with three conventional methods, the WLS SE, the WLS SE with bad data correction (BDC), and the WLAV SE. The WLS SE implicitly assumes that the measurement errors follow pure zero-mean Gaussian distributions, and it is mostly widely used in industry today. In the realistic power system, the bad data detection, identification, and correction procedure is typically followed by the WLS SE since it is not robust against gross errors [1]. The WLAV SE assumes that the measurement errors follow pure zero-mean Laplacian distributions, and it is one of the most widely investigated RSE methods in literature. In this simulation, the weights of WLS SE and WLAV SE are set based on the variances of true measurement errors. Note that such knowledge is not often available in practice, and the weight settings of two estimators are typically less reliable than assumed in these simulations. Therefore, the presented simulations already mimic the upper bounds of the conventional estimators' performance in practice.

A. Performance Under Measurement Errors Synthesized From the Realistic Measurement Chain Models

First, the feasibility of the proposed method is demonstrated in the first two time intervals in the 24-h testing period. Then, the performances of EPE and ASE are illustrated over the entire 24-h period under *time-varying* error distributions and *abrupt gross error* conditions, respectively.

1) Feasibility Validation of the Proposed EPE and ASE Procedures: The measurements within the first 30-min time interval are used to perform EPE for capturing the parameters of measurement error distributions. The ASE is executed by using error distribution parameters obtained from the EPE and on a single measurement scan in the second time interval. The simulation is repeated 2000 times.

For EPE, the cosine similarity metric between the pdfs based on true parameters and estimated parameters of all measurement channels is shown in Fig. 3. For most measurement channels, the similarity metric is very close to 1, indicating that the estimated measurement error distributions are very close to the ground truths. To further illustrate the performance of EPE, the true and estimated pdfs of measurement errors from six different measurement channels in PMUs and SCADA are

$$\hat{\tilde{\boldsymbol{x}}} = \arg\min_{\tilde{\boldsymbol{x}}} \sum_{\substack{\kappa=\mathrm{pmu},\\ \mathrm{scada}}} \sum_{i=1}^{m_\kappa} \left\{ \sum_{k=1}^{K_{\kappa,i}} \left[\frac{\tilde{w}'_{\kappa,i,k}}{2\hat{\sigma}^2_{\kappa,i,k}} \left(\tilde{z}_{\kappa,i} - h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right) - \hat{\mu}_{\kappa,i,k}\right)^2 \right] + \left[\frac{\tilde{w}_{\kappa,i,\mathrm{art}}}{2\sigma^2_{\mathrm{art}}} \left(\tilde{z}_{\kappa,i} - h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right)\right) \right] \right\} \tag{34}$$

$$\tilde{w}'_{\kappa,i,k} = \frac{\hat{\phi}_{\kappa,i,k} N\left(\tilde{z}_{\kappa,i} \middle| h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right) + \hat{\mu}_{\kappa,i,k}, \hat{\sigma}^2_{\kappa,i,k}\right)}{\sum_{k=1}^{K_{\kappa,i}} \hat{\phi}_{\kappa,i,k} N\left(\tilde{z}_{\kappa,i} \middle| h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right) + \hat{\mu}_{\kappa,i,k}, \hat{\sigma}^2_{\kappa,i,k}\right) + \phi_{\mathrm{art}} N\left(\tilde{z}_{\kappa,i} \middle| h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right), \sigma^2_{\mathrm{art}}\right)} \tag{35}$$

$$\tilde{w}_{\kappa,i,\mathrm{art}} = \frac{\phi_{\mathrm{art}} N\left(\tilde{z}_{\kappa,i} \middle| h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right), \sigma^2_{\mathrm{art}}\right)}{\sum_{k=1}^{K_{\kappa,i}} \hat{\phi}_{\kappa,i,k} N\left(\tilde{z}_{\kappa,i} \middle| h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right) + \hat{\mu}_{\kappa,i,k}, \hat{\sigma}^2_{\kappa,i,k}\right) + \phi_{\mathrm{art}} N\left(\tilde{z}_{\kappa,i} \middle| h_{\kappa,i}\left(\tilde{\boldsymbol{x}}\right), \sigma^2_{\mathrm{art}}\right)} \tag{36}$$

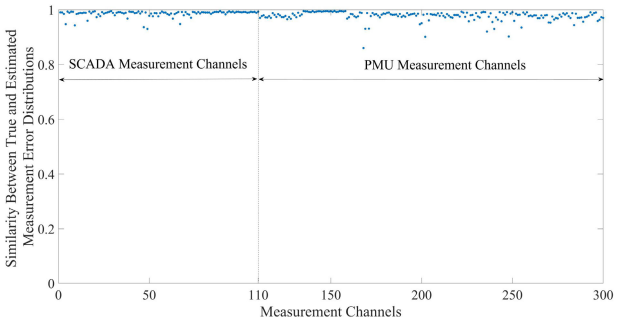

Fig. 3. Similarity between true and estimated measurement error distributions for all measurement channels.

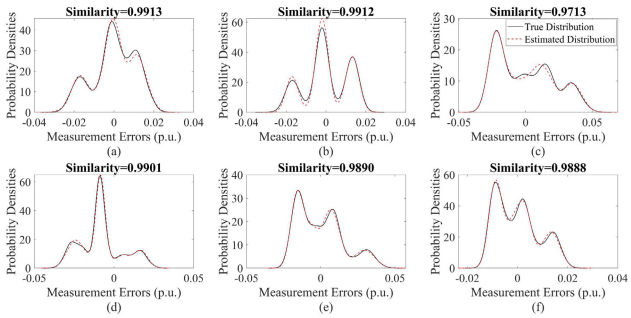

Fig. 4. Probability densities of measurement errors in selected channels: (a) 15th channel (P_{11}); (b) 41st channel (Q_{18}); (c) 112th channel (V_2); (d) 141st channel (δ_9); (e) 170th channel (real part of I_{5-2}); and (f) 234th channel (imaginary part of I_{2-6}).

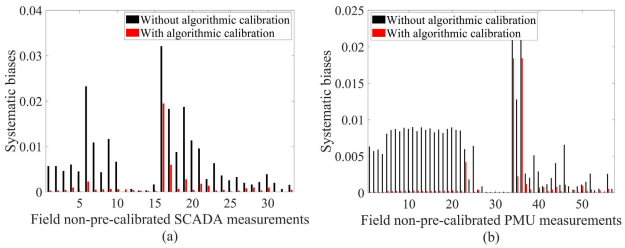

Fig. 5. Systematic biases of field nonprecalibrated measurement channels with and without algorithmic calibration using the EPE results: (a) SCADA channels and (b) PMU channels.

shown in Fig. 4. It is observed that the true and estimated pdfs are highly consistent, implying that the statistics of measurement errors can be accurately captured by the proposed EPE procedure. The accurate statistical knowledge of measurement error will support the subsequent ASE procedure to achieve accurate state estimates for power system operation.

As described in Section V-C, an important byproduct of EPE is the *algorithmic precalibration* of sensors with systematic errors. The measurement channels can be reset based on the estimated total mean of errors by (29) to remove systematic errors (biases) pertaining to the channel. In Fig. 5(a) and (b), the biases of the 30% *field nonprecalibrated* SCADA channels and PMU channels before and after *algorithmic precalibration* by EPE are presented, respectively. The biases are significantly reduced after *algorithmic precalibration* using the estimated total mean values given by EPE.

In order to evaluate the performance of ASE, the absolute errors of estimated voltage magnitudes and phase angles of

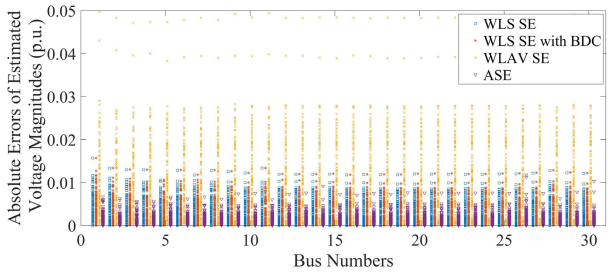

Fig. 6. Absolute errors of estimated voltage magnitudes.

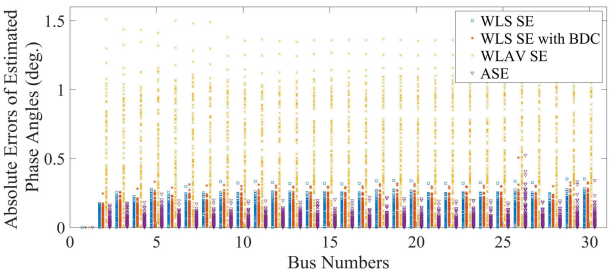

Fig. 7. Absolute errors of estimated phase angles.

all buses based on the WLS SE, the WLS SE with BDC, the WLAV SE, and the proposed ASE are shown in Figs. 6 and 7, respectively. They are scatter plots on the simulation results of 100 trials. Evidently, the distributions of absolute errors of all state estimates based on the proposed ASE are more concentrated around zero compared with WLS SE, WLS SE with BDC, and WLAV SE. The MAEs of the estimated voltage magnitudes and phase angles based on the developed ASE, the WLS SE, the WLS SE with BDC, and the WLAV SE are shown in Table III, along with the percentage improvements achieved by the ASE. It can be seen that the proposed ASE can achieve more accurate state estimates than the WLS SE, the WLS SE with BDC, and the WLAV SE by adopting the measurement error statistics captured by the proposed EPE. Note that the performance of the WLAV SE, which is well known for its capability of resisting heavy-tailed measurement errors, is inferior to the developed ASE. The reason is that WLAV SE is agnostic to the true statistics of measurement errors. It is the unbiased and minimum-variance estimator only when the measurement error follows a Laplacian distribution with a zero mean and a known scale parameter, which is not likely to be true in practice. The proposed ASE, on the other hand, does not require any prior knowledge of measurement error statistics. Instead, it is informed by the error statistics captured by EPE and adjusts its structure and parameters accordingly. Therefore, it achieves higher accuracy than robust estimators such as WLAV SE when the measurement error statistics are complex and unknown in advance.

The simulation results discussed above demonstrate the feasibility of the proposed EPE and ASE procedures in the first two time intervals. In practice, the measurement error conditions could be even more challenging: the error distributions could be time-varying and abrupt gross errors due to sensor/communication failures may also be present. The following two parts aim to evaluate the performance

TABLE II
ANALYSES OF MEASUREMENT ERROR SOURCES OF EACH COMPONENT IN SCADA AND PMU MEASUREMENT CHAINS

	Components	Impact Factors	Error Characteristics	Descriptions
SCADA Measurement Chain	Instrument Transformers	Systematic errors	Non-zero-mean	Systematic errors introduced by VTs and CTs may NOT be fully calibrated in real-world power systems.
		Power flow fluctuation	Time-varying	The magnitude of systematic errors introduced by CTs will be impacted by the percentage of rated current.
		Random noise	Likely Non-Gaussian	There is no literature or standard that study random noises introduced by instrument transformers, so no Gaussian assumption can be safely made.
	Control Cables and Burdens	Cable length and burden resistance	Non-zero-mean	Control cables can introduce a time delay, which will be transformed into phase angle errors.
	IEDs (power calculation)	Power calculation	Non-Gaussian	Power calculation will fuse voltage and current phasor measurements, resulting in non-Gaussian distributions even if the original errors follow Gaussian distributions.
	Communication Networks	Latency	Non-zero-mean, non-Gaussian, and time-varying	Latency will introduce biases for measurement data collected in control centers under power flow fluctuation.
		Power flow fluctuation		
PMU Measurement Chain	Instrument Transformers	The same as that in the SCADA measurement chain		
	Control Cables and Burdens	The same as that in the SCADA measurement chain		
	PMUs (phasor estimation)	Sampling time error	Non-zero-mean and non-Gaussian	In DFT-based phasor estimation, sampling time errors will create inaccuracy in phasor estimates, yielding phase angle estimation errors following *non-zero-mean near-uniform* distributions.
		Off-nominal frequency	Non-Gaussian and time-varying	In DFT-based phasor estimation, off-nominal frequency will result in phasor estimation errors following sinusoidal waveforms. Moreover, the frequency of the sinusoidal waveforms is impacted by the deviation of the off-nominal frequency.
		GPS signal loss	Non-zero-mean, non-Gaussian, and time-varying	Sampling time error will be accumulated over time and will not be cleared on one second by the PPS timing reference sent by the GPS.

TABLE III
SIMULATION RESULTS OF PSSE

	MAE_V (p.u.)	Improve. by ASE	MAE_δ (deg.)	Improve. by ASE
WLS SE	0.0042	59%	0.1100	55%
WLS SE w/ BDC	0.0042	59%	0.1135	56%
WLAV SE	0.0141	88%	0.5482	91%
ASE	0.0017	-	0.0497	-

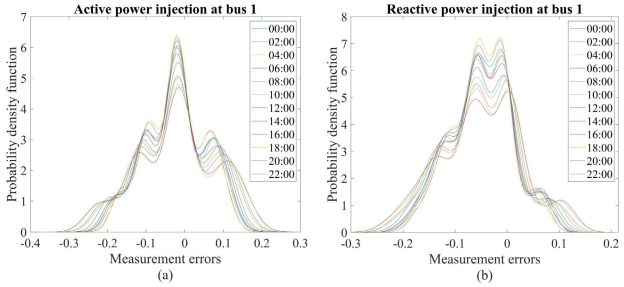

Fig. 8. Measurement error distributions in the SCADA measurement chain at different time instants: (a) active power errors and (b) reactive power errors.

of the proposed SE paradigm under such very challenging conditions. The simulation is executed continuously in a 24-h period. During this period, the power flow trend is derived from the actual load data from the ISO New England [48]. Consequently, the measurement error distributions of each measurement channel are time-varying due to the changes in operating states. Besides, the abrupt gross errors are introduced into 10% *field nonprecalibrated* measurement channels in the 30th and 40th time intervals as follows.

2) Normal Time-Varying Measurement Errors (24 h Excepting for the 30th and 40th Time Intervals): The measurement error distributions in the SCADA and PMU measurement chains at different time instants are illustrated in Figs. 8 and 9, respectively. The simulation is based on the real-world power flow trend derived from the ISO New England load data on June 1st, 2023 [48]. These results demonstrate that the statistics of measurement errors are indeed continuously varying due to the wide variety of varying conditions especially power system operating points, verifying the need to perform ASE. Meanwhile, the results also show that the variations of the measurement error statistics are relatively slow, and the difference between the 2 h is mild. Therefore, it can be inferred that the change of the measurement error distribution is very slight in a short time interval such as 30 min, which verifies the rationality of the assumption that measurement error distributions are approximately unchanged in the time window of the EPE procedure. In order to evaluate the performance of EPE under normal time-varying measurement error distributions, the average similarities of true and estimated error distributions of all measurement channels within different time intervals are shown in Fig. 10. The similarities between the true error distribution and the assumed error distributions by WLS SE

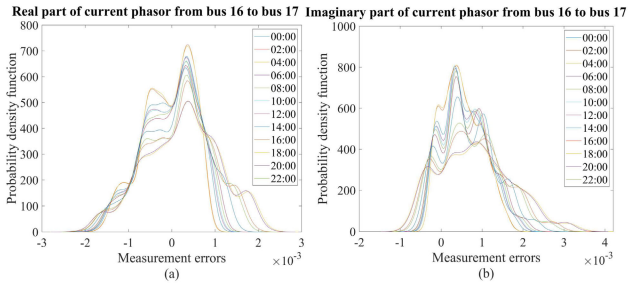

Fig. 9. Measurement error distributions in the PMU measurement chain at different time instants: (a) real part of current phasor errors and (b) imaginary part of current phasor errors.

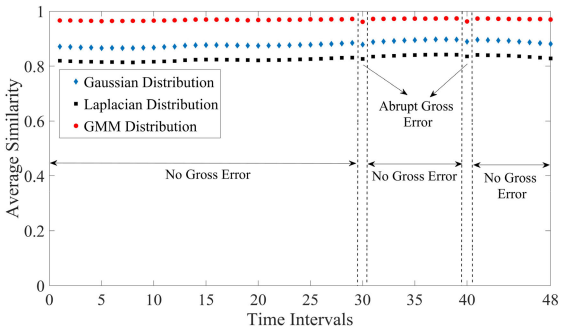

Fig. 10. Average similarities between captured and true error statistics.

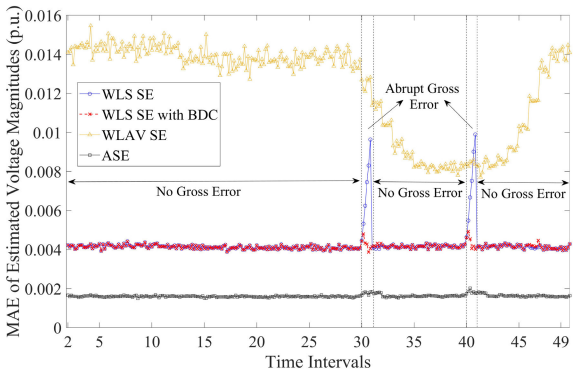

Fig. 11. MAEs of voltage magnitudes under complex and time-varying measurement errors.

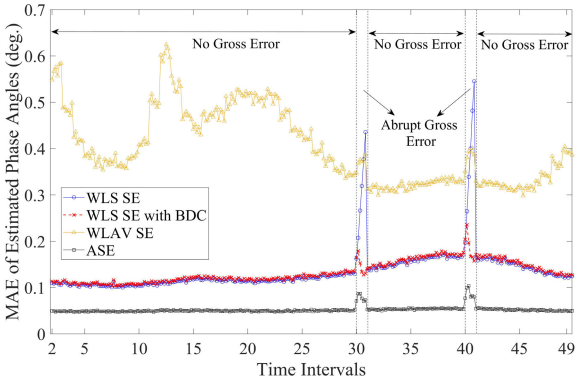

Fig. 12. MAEs of phase angles under complex and time-varying measurement errors.

and WLAV SE, which are a pure Gaussian distribution and a pure Laplacian distribution, are also shown comparatively. The results reveal that GMM used by the proposed EPE can fit the complex measurement error distribution much better than a pure Gaussian distribution or a pure Laplacian distribution. In order to evaluate the performance of ASE, the MAEs of estimated voltage magnitudes and phase angles of the four estimators from the second to the 49th interval are shown in Figs. 11 and 12, respectively. Note that the abscissas in Figs. 11 and 12 start at the second interval and end at the 49th interval. The reason is that parameter estimates obtained from a 30-min time window (i.e., the previous interval) by EPE are used by ASE in the subsequent 30-min time window (i.e., the current interval), as discussed in Section IV. Based on Figs. 11 and 12, it can be found that the performance of SE, particularly for the WLAV SE, exhibits time-varying characteristics as the underlying measurement error statistics vary with the change of operating states of power systems. Moreover, the proposed ASE outperforms the WLS SE, the WLS SE with BDC, and the WLAV SE regardless of the variation of measurement error statistics, manifesting strong adaptiveness under a complex and time-varying measurement environment.

3) Abrupt Gross Errors (the 30th and 40th Time Intervals): When gross errors abruptly occur due to sensor/communication failure, the error distribution may be severely distorted instantaneously, which the EPE procedure fails to capture in a prompt manner. In order to test the performance of the proposed ASE paradigm under abrupt gross errors, it is assumed that the communication delay and the GPS signal loss occurred in 10% *field nonprecalibrated* SCADA and PMU measurement channels, respectively. From Figs. 11 and 12, the developed ASE also remains the most accurate state estimate among the four estimators under the extreme situation. Especially, it still achieves better performances than the well-known robust WLAV estimator and the WLS with BDC. The reason is that although the ASE cannot exploit the statistics of measurement errors within the *current interval*, the M-step [i.e., (36)] will compute a large posterior probability for the gross error trap, $\tilde{w}_{\kappa,i,\text{art}}$ (i.e., $\tilde{w}_{\kappa,i,\text{art}}$ will be close to 1 and $\tilde{w}'_{\kappa,i,k}$ will be close to 0). In other words, in the E-step, the ASE will infer that there is a large probability that the measurement comes from the artificially introduced Gaussian component with a large variance σ^2_{art}, and then suppress its impact on the M-step estimation. As a result, the state estimates will not be heavily affected by the gross error. Hence, the developed ASE has the capability to suppress gross errors that occur instantaneously and are not captured promptly by EPE.

B. Computational Efficiency

The simulations are executed using a personal computer with Intel Core i7-9700K CPU, 32GB RAM, and Windows 10 64-bit operating system. The proposed EPE and ASE algorithms are implemented using MATLAB version R2020b. For the EPE procedure executed every 30 min (1800 s) for capturing the measurement error statistics, the average computational time is 436 s, which shows that the proposed EPE algorithm can be executed in the desirable frequency. For the ASE procedure executed at a much higher frequency

to estimate the state variables of the system, the average computational times for the WLS SE, the WLS SE with BDC, the WLAV SE, and the proposed ASE are 0.026, 0.043, 0.078, and 0.136 s, respectively. Evidently, the computational cost of the proposed ASE is only mildly higher than those of the conventional WLS SE and WLAV SE. Moreover, the computational burden for large power systems will not be an obstacle to the practical application of the proposed adaptive state estimator due to the partition and parallel computing approaches [49], [50], [51]. This is because the proposed EPE for a measurement channel and proposed ASE for a state variable largely rely on the information in the local area [1]. For example, for estimating the state variable of a given bus, the measurements that are three or four buses away may have a very little impact [1, Ch. 7.2]. In addition, in the EPE procedure, the parameter estimation, i.e., (20) and (24)–(26), are independent for different measurement channels and can be performed in parallel. In a control center where parallel computing is available, the computational time will be further reduced. Hence, the large power system can be partitioned into a few small-scale systems and executed in parallel with the EPE and ASE procedures. Therefore, the proposed SE paradigm is efficient enough for the online monitoring of power grids.

While partitioning and parallel computing methods can improve computational efficiency, they may also result in several issues, such as the deterioration of SE accuracy due to the global impact of voltage magnitude measurements, the uniqueness of phase angle solutions for multiple subsystems, the lower capability of bad data processing at the boundaries of subsystems, etc. Fortunately, the global impact of voltage magnitude measurements on the entire system can be mitigated since abundant voltage measurements are available in today's power systems [49], [50]. A unique phase angle solution for the entire system can be achieved by sharing the local SE results between neighboring estimators [49]. Moreover, the robustness of decentralized SE can be improved by overlapping the boundary buses in neighboring subsystems [50] and by implementing two-stage SE [51]. Although numerous methods have been proposed to tackle these challenges [49], [50], [51], careful consideration is still required when utilizing partitioning and parallel computing techniques to deal with large-scale systems.

VIII. Conclusion

The performance of PSSE is dependent on the degree of fitness between the formulation of estimators and the true measurement error statistics. The challenging reality is that the statistics of measurement errors are typically unknown, nonzero-mean, non-Gaussian, and time-varying, limiting the performances of estimators with stiff prior assumptions. This article proposes a comprehensive adaptive paradigm to maneuver PSSE through such an uncertain challenging measurement environment. Simulation results in the IEEE 30-bus test system demonstrate that the developed EPE method accurately captures complex probability distributions of measurement errors. It also enables an accurate sensor *algorithmic precalibration* approach without the need for field experiments. The ASE method, which updates the estimator formulation based on the captured statistical knowledge of measurement errors by EPE, yields more accurate state estimates compared with the mostly widely used WLS SE, WLS with BDC, and robust WLAV SE methods under unknown, time-varying, and highly complex measurement error conditions. Through a carefully designed gross error trap, the proposed ASE performs stably well even when abrupt gross errors occur. The proposed algorithms are applicable to a mix of PMU and SCADA measurements and are demonstrated to be computationally efficient for real-time applications. Although the measurement chain analysis in the article has sufficiently motivated and validated the proposed method, further investigation and documentation of the detailed error characteristics of individual components in the measurement chain and their combined effect in the real world is an important direction for the technical community. The authors' more detailed work in this aspect will be presented in follow-on IEEE publications.

References

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

[2] E. Shereen, M. Delcourt, S. Barreto, G. Dán, J.-Y. Le Boudec, and M. Paolone, "Feasibility of time-synchronization attacks against PMU-based state estimation," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 3412–3427, Jun. 2020.

[3] F. Schweppe, "Power system static-state estimation, Part III: Implementation," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 130–135, Jan. 1970.

[4] T. Ahmad and N. Senroy, "Statistical characterization of PMU error for robust WAMS based analytics," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 920–928, Mar. 2020.

[5] R. Martínez-Parrales, C. R. Fuerte-Esquivel, and B. A. Alcaide-Moreno, "Analysis of bad data in power system state estimation under non-Gaussian measurement noise," *Electr. Power Syst. Res.*, vol. 186, Sep. 2020, Art. no. 106424.

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

[7] C. Sitzia, C. Muscas, P. A. Pegoraro, A. V. Solinas, and S. Sulis, "Enhanced PMU-based line parameters estimation and compensation of systematic measurement errors in power grids considering multiple operating conditions," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.

[8] R. Szplet, R. Szymanowski, and D. Sondej, "Measurement uncertainty of precise interpolating time counters," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 11, pp. 4348–4356, Nov. 2019.

[9] A. Brameller and S. H. Karaki, "Power-system state estimation using linear programming," *Proc. Inst. Electr. Eng.*, vol. 126, no. 3, p. 246, 1979.

[10] M. K. Celik and A. Abur, "A robust WLAV state estimator using transformations," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 106–113, Feb. 1992.

[11] M. Dorier, G. Frigo, A. Abur, and M. Paolone, "Leverage point identification method for LAV-based state estimation," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[12] Y. Lin and A. Abur, "Robust state estimation against measurement and network parameter errors," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4751–4759, Sep. 2018.

[13] G. Wang, G. B. Giannakis, and J. Chen, "Robust and scalable power system state estimation via composite optimization," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6137–6147, Nov. 2019.

[14] M. Kabiri and N. Amjady, "A new hybrid state estimation considering different accuracy levels of PMU and SCADA measurements," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 9, pp. 3078–3089, Sep. 2019.

[15] L. Mili, M. G. Cheniae, N. S. Vichare, and P. J. Rousseeuw, "Robust state estimation based on projection statistics [of power systems]," *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 1118–1127, May 1996.

[16] G. He, S. Dong, J. Qi, and Y. Wang, "Robust state estimator based on maximum normal measurement rate," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2058–2065, Nov. 2011.

[17] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1617–1626, May 2013.
[18] Y. Chen, J. Ma, P. Zhang, F. Liu, and S. Mei, "Robust state estimator based on maximum exponential absolute value," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1537–1544, Jul. 2017.
[19] J. Zhao and L. Mili, "A framework for robust hybrid state estimation with unknown measurement noise statistics," *IEEE Trans. Ind. Informat.*, vol. 14, no. 5, pp. 1866–1875, May 2018.
[20] 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 Trans. Power Syst.*, vol. 36, no. 5, pp. 4482–4491, Sep. 2021.
[21] M. S. Thomas and J. D. McDonald, *Power System SCADA and Smart Grids*. Boca Raton, FL, USA: CRC Press, 2017.
[22] J. H. Harlow, *Electric Power Transformer Engineering*. Boca Raton, FL, USA: Taylor & Francis, 2012.
[23] *IEEE Standard Requirements for Instrument Transformers*, IEEE Standard C57.13-2016, 2016, pp. 1–96.
[24] J. Zhao, J. Tan, L. Wu, L. Zhan, W. Yao, and Y. Liu, "Impact of the measurement errors on synchrophasor-based WAMS applications," *IEEE Access*, vol. 7, pp. 143960–143972, 2019.
[25] *IEEE Guide for Synchronization, Calibration, Testing, and In-Stallation of Phasor Measurement Units (PMUs) for Power System Protection and Control*, IEEE Standard C37.242, 2013, pp. 1–107.
[26] *IEEE Guide for the Design and Installation of Cable Systems in Substations—Redline*, IEEE Standard 525-2016, Nov. 2016, pp. 1–534.
[27] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. London, U.K.: Oxford Univ. Press, 2020.
[28] P. Castello, G. Gallus, P. A. Pegoraro, and S. Sulis, "Measurement platform for latency characterization of wide area monitoring, protection and control systems," *IEEE Trans. Instrum. Meas.*, early access, Nov. 28, 2024, doi: 10.1109/TIM.2023.3334360.
[29] M. Bertocco, G. Gamba, A. Sona, and S. Vitturi, "Experimental characterization of wireless sensor networks for industrial applications," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 8, pp. 1537–1546, Aug. 2008.
[30] S. Chakrabarti, E. Kyriakides, and M. Albu, "Uncertainty in power system state variables obtained through synchronized measurements," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 8, pp. 2452–2458, Aug. 2009.
[31] *IEEE/IEC International Standard—Measuring Relays and Protec-tion Equipment—Part 118-1: Synchrophasor for Power Systems—Measurements*, IEC/IEEE Standard 60255-118-1:2018, Dec. 2018, pp. 1–78.
[32] W. Yao et al., "A novel method for phasor measurement unit sampling time error compensation," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1063–1072, Mar. 2018.
[33] J. Zhao, A. Goldstein, and Y. Liu, "Model of parameterized PMU estimation error," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Chicago, IL, USA, Jul. 2017, pp. 1–5.
[34] W. Yao et al., "Impact of GPS signal loss and its mitigation in power system synchronized measurement devices," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1141–1149, Mar. 2018.
[35] G. Cheng and Y. Lin, "Modeling of SCADA and PMU measurement chains," 2023, *arXiv:2312.02844*.
[36] G. Huang, D. Akopian, and C. L. P. Chen, "Measurement and characterization of channel delays for broadband power line communications," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 11, pp. 2583–2590, Nov. 2014.
[37] M. Pignati et al., "Real-time state estimation of the EPFL-campus medium-voltage grid by using PMUs," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf.*, Jul. 2015, pp. 1–5.
[38] K. E. Martin et al., "Exploring the IEEE standard C37.118-2005 synchrophasors for power systems," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 1805–1811, Oct. 2008.
[39] D. A. Reynolds, *Gaussian Mixture Models*. New York, NY, USA: Springer, 2009.
[40] Y. Lin and A. Abur, "Enhancing network parameter error detection and correction via multiple measurement scans," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2417–2425, May 2017.
[41] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, Jul. 2002.
[42] S.-H. Cha, "Comprehensive survey on distance/similarity measures between probability density functions," *Int. J. Math. Models Methods Appl. Sci.*, vol. 4, no. 1, pp. 300–307, 2007.
[43] X.-L. Meng and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: A general framework," *Biometrika*, vol. 80, no. 2, p. 267, Jun. 1993.
[44] A. Brandolini, M. Faifer, and R. Ottoboni, "A simple method for the calibration of traditional and electronic measurement current and voltage transformers," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 5, pp. 1345–1353, May 2009.
[45] K. V. Khandeparkar, S. A. Soman, and G. Gajjar, "Detection and correction of systematic errors in instrument transformers along with line parameter estimation using PMU data," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3089–3098, Jul. 2017.
[46] D. Shi, D. J. Tylavsky, and N. Logic, "An adaptive method for detection and correction of errors in PMU measurements," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1575–1583, Dec. 2012.
[47] H. Goklani, G. Gajjar, and S. A. Soman, "Instrument transformer calibration and robust estimation of transmission line parameters using PMU measurements," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 1761–1770, May 2021.
[48] (Jun. 2023). *System Load Graphs: ISO New England-Real-Time Maps and Charts*. [Online]. Available: https://www.iso-ne.com/isoexpress/web/charts
[49] J. Ban, J. Im, Y.-J. Kim, and J. Zhao, "Decentralization of phasor-aided state estimation using local state vector extension," *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4645–4659, Sep. 2021.
[50] C. Xu and A. Abur, "A massively parallel framework for very large scale linear state estimation," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4407–4413, Jul. 2018.
[51] C. Xu and A. Abur, "A fast and robust linear state estimator for very large scale interconnected power grids," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4975–4982, Sep. 2018.

Gang Cheng (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering and automation from Henan Polytechnic University, Jiaozuo, China, in 2016, the M.S. degree in control theory and control engineering from Guangxi University, Nanning, China, in 2019, and the Ph.D. degree in electrical engineering from the University of Massachusetts, Lowell, MA, USA, in 2023.

He is currently a Post-Doctoral Researcher at the University of Connecticut, Storrs, CT, USA. His current research interests include situational awareness, cyber security, data analysis, and machine learning of power systems.

Yuzhang Lin (Member, IEEE) received the bachelor's and master's degrees in electrical engineering from Tsinghua University, Beijing, China, in 2012 and 2014, respectively, and the Ph.D. degree in electrical engineering from Northeastern University, Boston, MA, USA, in 2018.

He was an Assistant Professor at the University of Massachusetts, Lowell, MA, USA, prior to joining NYU. He is currently an Assistant Professor at the Department of Electrical and Computer Engineering, New York University, New York, NY, USA. His research interests include modeling, situational awareness, data analytics, and cyber-physical resilience of smart grids.

Dr. Lin was a recipient of the NSF CAREER Award. He serves as the Co-Chair of the IEEE PES Task Force on Standard Test Cases of Power System State Estimation and the Secretary of the IEEE PES Distribution System Operation and Planning Subcommittee.