Synchrophasor-based Online State Estimated in Large-scale Power Grid

Junjie Lin, *Member*, *IEEE*, Chao Lu, *Senior Member*, *IEEE*, Wenchao Song, *Student Member*, *IEEE*, Yingduo Han, *Senior Member*, *IEEE*, Yinsheng Su, He Huang, and Licheng Li

***Abstract*—With integration of a larger amount of clean power sources and power electronic equipment, operation and dynamic characteristics of the power grid are becoming more and more complicated and stochastic. Therefore, it is necessary and urgent to obtain accurate real-time states, which is difficult from traditional state estimation. This paper systematically develops a phasor measurement unit (PMU) based real-time state estimator for a realistic large-scale power grid for the first time. The estimator mainly relies on three refined algorithms, i.e., an improved linear state estimation algorithm, a practical bad data identification method and a distributed topology check technique. Furthermore, a novel system architecture is designed and implemented for the China Southern Power Grid. Numerical simulations and extensive field operation results of the state estimator recorded under both normal and abnormal situations are presented. All the tests and field results demonstrate the advantages of the proposed algorithms in terms of online system monitoring and feasibility of refreshing the states of the whole system at intervals of tens of milliseconds.**

***Index Terms*—PMU-based state estimation, design and implementation, bad data identification, topology analysis.**

NOMENCLATURE

n	Number of buses.
m	Number of branches.
$\boldsymbol{I}_{\mathrm{b}}$	Vector of the branch current phasors.
$\boldsymbol{U}$	Vector of bus voltage phasors.
$\boldsymbol{Y}_{\mathrm{b}}$	Bus admittance matrix.
$\boldsymbol{I}_i$	Vector of nodal injection current phasors.
$\boldsymbol{B}$	Bus-branch association matrix.
$\boldsymbol{Y}_i$	Bus admittance matrix for injected bus.
$\boldsymbol{E}$	Identity matrix.
$\boldsymbol{Z}_{\mathrm{m}}$	Measurement vector.
$\boldsymbol{Z}$	Vector of true values.
$\boldsymbol{A}$	Coefficient matrix.
ε	Vector of measurement errors.
$\boldsymbol{W}$	Weighted matrix.
$\boldsymbol{G}$	Gain matrix.
$\hat{\boldsymbol{U}}$	Vectors of estimated voltage phasors.
$\boldsymbol{r}$	Vector of estimated residuals.
σ_{ii}	Covariance of residuals.
η_i	Specific thresholds for Topology Check.
τ	SE Qualified Rate.

Manuscript received November 16, 2021; revised January 3, 2022; accepted March 23, 2022. Date of online publication January 25, 2023; date of current version April 10, 2023. This work was supported by the National Natural Science Foundation of China (U1766214; U2066601).

J. J. Lin is with the College of Electrical Engineering and Automation, Fuzhou University, Fuzhou, China, and also with the Department of Electrical Engineering, Tsinghua University, Beijing, China.

C. Lu (corresponding author, email: luchao@thu.edu.cn), W. C. Song and Y. D. Han are with the State Key Lab of Power systems, Department of Electrical Engineering, Tsinghua University, Beijing, China.

Y. S. Su, H. Huang and L. C. Li are with China Southern Power Grid, Guangzhou, China.

DOI: 10.17775/CSEEJPES.2021.08550

I. INTRODUCTION

POWER system state estimation (SE) is a basic function of the energy management system (EMS) in a power grid. It serves as a cornerstone for many applications, such as power flow calculation, power dispatching, and online security and stability analysis. The primary purpose of SE is to maintain a complete and reliable real-time system state database that can provide real-time operating data (states and network models) for system operators and other power system analysis applications.

Conventional SE methods use measurement data from supervisory control and data acquisition (SCADA) system and generally adopt weighted least squares (WLS) algorithms to obtain system states. However, such methods relying on SCADA measurements suffer from a series of problems, such as long measurement intervals, large nonsynchronous measurement errors, and numerical instability due to nonlinear models [1].

Recently, with development of phasor measurement units (PMUs) and wide-area measurement systems (WAMSs), new opportunities have become available for obtaining real-time states of power systems [2], [3]. To date, more than 4000 substations have been deployed with installed PMUs in China's power grids, covering almost all 500 kV-and-above substations/plants and important 220 kV substations/plants [4]. WAMS uses the Global Positioning System (GPS) as a synchronous clock for PMU data acquisition at individual substations. Measurement data are stamped with highly accurate time values, and measurement data of a snapshot are highly synchronous. In addition, reporting rate of a PMU can be up to twice the frequency of the power grid, enabling efficient tracking of system-wide dynamic behaviors. Nevertheless, raw PMU measurements are likely to be somewhat contaminated with bad data due to device failure, communication interference, sensor error, etc. [5]. Therefore, development and implementation of SE systems based on PMU measurements can not only effectively improve quality of PMU measurement data but also promote maintenance and utilization of on-site PMUs. Real-time states and power flows of the system as obtained through PMU-based SE will also provide more accurate and refined information for system operation and dispatch.

Since invention of PMU technology, many studies have attempted to incorporate PMU measurements into SE [6]. Early studies simply introduced PMU measurements into traditional state estimators. Related efforts used PMU measurements to improve measurement redundancy and observability, but did not fully utilize the characteristics and advantages of PMU data [7], [8]. Later, hybrid SE using SCADA and PMU data was considered [9], [10], but a compatibility problem between the two different types of measurement data has not been effectively solved, and sometimes performance is even worse than of traditional state estimators due to time skew errors [3]. Prof. Phadke established a PMU-based measurement equation under the condition that PMU configuration satisfies observability and proposed the concept of linear state estimation. It also shows the WLS method can be used to solve the SE problem, but there is no further discussion on how to solve the complex domain WLS [3], [6]. Synchrophasors are used to solve the multiarea state estimation problem through a common LS method [11]. Current synchrophasors are converted to virtual voltage measurement in polar coordinates for traditional state estimation [12]. The CWLS algorithm [13] is proposed and applied to PMU linear state estimation [14]. CWLS-SE has achieved good results with faster calculation speed and robustness to phase angle errors [15]. New theories and methods concerning PMU-based SE have been developed, and many studies have achieved remarkable results [16], [17]. However, various problems, such as bad data detection and identification, real-time topology analysis have still not been completely addressed [18], [19]. Moreover, efficiency of some methods cannot be successfully adapted to high-density PMU data.

On the other hand, few practical PMU-based SE systems are implemented in the industry, although related research works have been widely reported. The China Southern Power Grid (CSG) conducted a PMU-based SE demonstration project in a small-scale partial power grid in 2010. Because of the nonlinear model, SE interval was 2–4 s, much longer than PMU measurement interval. Consequently, results did not reflect the essential characteristics of PMU-based SE methods. Several US utilities have carried out similar pilot projects on PMU-based SE in small-scale grids in the last few years [20], [21], and linear state estimators have been adopted and implemented in extra-high-voltage (EHV) power grids with tens of substations [22]. The paper [23] presented the architecture and testing of a real-time state estimator based on synchrophasors in a medium-voltage distribution network of the Swiss Federal Institute of Technology Lausanne (EPFL) campus. Most reported projects have been developed for and tested in small power systems, while applying of PMU-based SE on large-scale power grids puts forward higher requirements for computational efficiency and robustness of the algorithm, and real-time processing of massive PMU data. More efforts will be needed to pursue practical industrial application of PMU-based SE.

To address these problems in theoretical methods and industrial practice, this paper proposes a systematic SE framework based on PMU measurements and presents its design and implementation for use in the EHV transmission network of the CSG. The two main contributions are summarized as follows:

1) This paper proposes an improved PMU linear state estimation method that takes into account node injection constraints, incorporating practical bad data identification and distributed local topology check methods. The proposed methods can accurately and quickly track the real-time state of a power system.

2) Our work has first implemented and applied a PMU-based SE on a large-scale power grid consisting of hundreds of substation and plants. Field results show our work has significantly improved availability of PMU data and observability of the power grid, which can provide reliable real-time data for other system analysis applications.

The rest of the paper is structured as follows: Section II introduces the CSG system and illustrates design and implementation of the PMU-based state estimator for use in a control center of the CSG; practical issues facing deployment of the PMU-based state estimator are also presented in this section. Section III proposes key algorithms for the PMU-based SE system, including an improved linear SE model and solution and a bad data processing method. Section IV validates effectiveness of the proposed PMU-based state estimator through simulations. Section V presents field operation results under both normal and abnormal conditions. Section VI concludes the paper.

II. Architecture and Design

To ensure real-time performance of PMU-based SE on large-scale power grid, it is necessary to design flexible system architecture and efficient WAMS data processing techniques.

A. Overall Design of PMU-Based SE in the CSG

The CSG is composed of five provincial power networks in South China, with a total service area spanning 1 million square kilometers and serving a population of 252 million people. The 500 kV backbone network of the CSG includes a total of 234 substations and plants and 465 AC transmission lines. The power grid topological graph is shown in Fig. 3 below. Conventionally, the control center of the CSG uses a WLS-based SE based on SCADA measurements, as most utilities throughout the world do. However, SCADA Measurements of individual snapshots are weakly synchronized, and estimation intervals can be minute-long [24]. SCADA-based SE may suffer from divergence problems, and it might turn out not to be a global optimal solution because of nonlinearity of the estimation model.

Due to inherent shortcomings of the existing SE system, the PMU-based SE is studied and applied to improve real-time monitoring capabilities for a power grid. PMUs are deployed at all of the 500 kV substations in the CSG. Reporting rate of the PMUs in the CSG is 50 Hz, as described in the IEEE C37.118

standard [25]. A real-time PMU-based state estimator has been implemented that covers the whole EHV grid of the CSG. Some practical issues and concerns have been encountered while implementing this PMU-based state estimator. Detailed issues and the corresponding solutions adopted to tackle them are summarized as follows:

1) Although parts of the 220 kV voltage grid in the CSG are also covered by PMUs, since measurement data of the 220 kV substations are not uploaded to the control center, SE is currently performed only for a 500 kV grid. However, if necessary measurement data become available, the program can be easily extended to the 220 kV grid.

2) There are 8 ultrahigh-voltage DC transmission lines in the CSG, and PMUs are also installed on the AC lines at the AC/DC converter stations. The proposed PMU-based state estimator does not consider a DC line model, so power and current of DC lines are treated as injected power and current.

3) Adequate PMUs guarantee considerable measurement redundancy, so observability of the CSG will meet requirements for SE unless there is a large-scale communication failure or equipment failure. If some substations are unobservable because of bad data, they will be excluded from the main island, as in traditional SE.

4) PMUs can measure all three phase values and obtain positive sequence values. However, symmetry of the three phases in the EHV grid is sufficiently high, and the main grid model used by the operator is the positive sequence model. The number of matrix dimensions of the positive sequence model is only one third that of the three-phase model, meaning that calculation efficiency can be greatly improved by using a positive sequence model. Therefore, positive sequence modeling and measurements are used in the project.

5) Each substation/plant with a PMU in the CSG is also equipped with a phasor data concentrator (PDC) that is directly connected to the dispatch center. Communication delay between PDC and control center is less than 50 ms, in accordance with requirements of Chinese industrial standards. PMU data delay measured in the CSG is usually approximately 10 ms; accordingly, PMU-based SE result may lag by tens of milliseconds compared to actual time.

B. System Architecture

PMU-based SE system was developed and implemented based on Operation Smart System (OS2) of the CSG, independent of and parallel to an existing SE system. OS2 is the EMS of the CSG and provides various data interfaces to power system models, SCADA and PMU measurements, and other information. The system can be roughly divided into four parts as Fig. 1: platform layer, database layer, application layer, and graphical user interface (GUI) layer.

The database layer includes a static database, a real-time database (RTDB), and a WAMS measurement time series database. The static database is a relational database based on MYSQL that stores basic information of the power grid and configuration information required for system operation and communication. RTDB is a fast in-memory database developed based on Redis for use as a cache, message broker and queue while programs are running. WAMS database is designed to

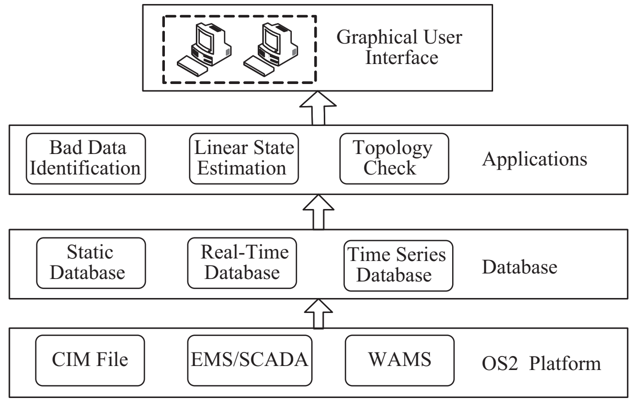


Fig. 1. PMU based SE System Architecture.

store/load WAMS real-time and historical data. Compared with traditional SE procedure, since PMU data, including SE results, are large in volume and high in resolution, it is necessary to develop a specialized time series database with higher performance for real-time access.

The application layer deploys the algorithms proposed in Section III, including bad data processing, linear state estimator, local topology check and other necessary system tools. These programs were developed in C/C++ for the best efficiency. In addition, to facilitate offline testing of the system, a tool that can send (historical or simulated) C37.118 PMU data files has been developed and integrated into the system.

C. WAMS Data Processing

The WAMS time series database is divided into two parts: a time series real-time database and a time series historical database. The real-time part of the database can be regarded as a database based on memory operations. It receives the latest data from the PMUs and performs data synchronization. A linked list and stack structure are used to store the latest PMU data in memory and in real time. Length of the saved data is fixed to 5 minutes. When data for a new time point come in, data from more than 5 minutes prior will be compressed and saved in the time series historical database. By storing the latest data in memory, it is ensured that applications can read these measurement data and SE result data at the fastest possible speed.

The historical database stores overflow data from the real-time database. The historical database specifically stores compressed floating-point data subjected to integer fixed-point and incremental processing. Processing procedures include the following steps, as shown in Fig. 2.

Conversion coefficients for various types of measurements are defined in accordance with accuracy requirements for WAMS data. All floating-point numbers are multiplied by corresponding conversion coefficients and converted by rounding into an integer fixed-point sequence; then, adjacent points in the fixed-point sequence are successively subtracted to obtain an incremental data sequence; and finally, incremental data are compressed via the Lempel-Ziv-Welch (LZW) algorithm [26]. When PMU data are to be used, decompression and incremental processing are performed in reverse order to restore original data.

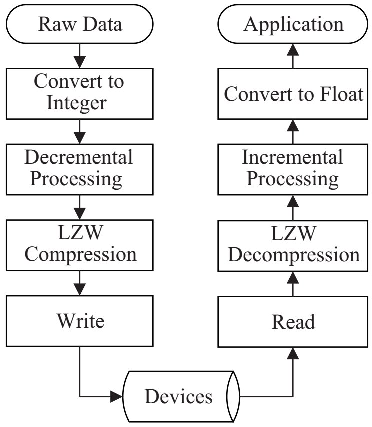


Fig. 2. PMU Data Processing.

III. Methods and Technologies

Data interval of the PMUs is only 20 ms, so the most important concern for PMU-based SE is to ensure accuracy while still achieving high calculation efficiency and robustness. This section introduces methods used to ensure accuracy and efficiency in realizing PMU-based real-time SE for a large-scale power grid.

A. Improved Linear State Estimator

The basic linear SE (LSE) algorithm has been well explored in recent decades [3]. As mentioned above, a PMU can directly measure current and voltage phasors. Therefore, an LSE model can be obtained in accordance with line impedance characteristics, as for a power system with n buses and m branches,

$$\boldsymbol{I}_{\mathrm{b}} = \boldsymbol{Y}_{\mathrm{b}}\boldsymbol{U} \tag{1}$$

where $\boldsymbol{I}_{\mathrm{b}} \in \mathbb{C}^{2m}$ and $\boldsymbol{U} \in \mathbb{C}^{n}$ are vectors of the branch current phasors and bus voltage phasors, respectively, and $\boldsymbol{Y}_{\mathrm{b}} \in \mathbb{C}^{2m\times n}$ is the admittance matrix, which represents the relationship between branch currents and node voltages.

Most PMU-based LSE methods are based on (1). However, such an LSE model is based only on branch characteristics, while constraints of the power network are not considered. In fact, a PMU can measure current phasors of all power lines at a substation, including windings of the transformers. Therefore, nodal equations can be derived as follows in light of Kirchhoff's current law (KCL):

$$\boldsymbol{I}_i = \boldsymbol{B}\boldsymbol{I}_{\mathrm{b}} \tag{2}$$

where $\boldsymbol{I}_i \in \mathbb{C}^{n}$ is the vector of nodal net injection current phasors, $\boldsymbol{B} \in \mathbb{R}^{n\times n}$ is bus-branch association matrix. According to the definition of an injection current, nodal injection current is the sum of currents of the loads and generators connected to the node. Therefore, an injection current of a 500 kV node is the sum of current measurements for high-voltage windings of the transformers, which are not counted as branches. There may be some nodes in the power grid that have no load or generation, that is, zero-injection nodes. Injection current value for such a node is zero.

Substituting equation (1) into (2) yields

$$\boldsymbol{I}_i = \boldsymbol{Y}\boldsymbol{U} \tag{3}$$

where

$$\boldsymbol{Y}_i = \boldsymbol{B}\boldsymbol{Y}_{\mathrm{b}} \tag{4}$$

represents the relationship between nodal injection currents and node voltages.

By combining equations (1) and (3), the SE model is derived:

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{U} \\ \boldsymbol{I}_{\mathrm{b}} \\ \boldsymbol{I}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{U} \\ \boldsymbol{Y}_{\mathrm{b}}\boldsymbol{U} \\ \boldsymbol{Y}_i\boldsymbol{U} \end{bmatrix} = \begin{bmatrix} \boldsymbol{E} \\ \boldsymbol{Y}_{\mathrm{b}} \\ \boldsymbol{Y}_i \end{bmatrix} \boldsymbol{U} = \boldsymbol{A}\boldsymbol{U} \tag{5}$$

where $\boldsymbol{Z} \in \mathbb{C}^{2m+2n}$ is the vector of true values corresponding to the PMU measurements. Thus, the measurement model is represented as the following:

$$\boldsymbol{Z}_{\mathrm{m}} = \boldsymbol{A}\boldsymbol{U} + \boldsymbol{\varepsilon} \tag{6}$$

where $\boldsymbol{Z}_{\mathrm{m}} \in \mathbb{C}^{2m+2n}$ is the measurement vector and $\boldsymbol{\varepsilon} \in \mathbb{C}^{2m+2n}$ is the vector of measurement errors, which are generally assumed to obey independent Gaussian distributions with mean values of zero. The objective function of the model expressed in (6) using a WLS estimator is as follows:

$$\min_{\boldsymbol{U}} \quad J_{\mathrm{WLS}}(\boldsymbol{U}) = (\boldsymbol{Z}_{\mathrm{m}} - \boldsymbol{A}\boldsymbol{U})^{\mathrm{H}}\boldsymbol{W}(\boldsymbol{Z}_{\mathrm{m}} - \boldsymbol{A}\boldsymbol{U}) \tag{7}$$

where $\boldsymbol{W} \in \mathbb{R}^{n\times n}$ is a diagonal weight matrix composed of reciprocals of the measured error variances. Covariance of measured data can be estimated by analyzing historical data, as in traditional SE method. Zero injections use very large weights. The solution to (7) is given by

$$\hat{U} = (\boldsymbol{A}^{\mathrm{H}}\boldsymbol{W}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{W}\boldsymbol{Z}_{\mathrm{m}} = \boldsymbol{G}\boldsymbol{Z} \tag{8}$$

This is the globally optimal unbiased estimate for problem (7) when measurement errors obey a Gaussian distribution. It is evident that only one matrix inversion operation is needed to obtain the solution, without any iteration. Thus, common problems of iterative algorithms, such as problems with convergence and numerical stability, are avoided. Furthermore, in the case that network topology and measurement configuration remain unchanged, gain matrix $\boldsymbol{G}$ is constant, which can significantly reduce computing time.

B. Implementation of LSE

To improve computational efficiency of LSE, the following techniques are further adopted.

First, the least squares method in the complex domain is used to perform the calculation. In our previous work [15], we proposed an LSE method based on the least squares approach in the complex domain. Compared with traditional methods, it shows better robustness to phase angle errors, and computational efficiency is also improved several-fold.

Second, because many of the matrices involved in the calculation process, such as branch admittance matrix and the weight matrix, are highly sparse, the PMU-based SE system uses a sparse matrix representation method to perform matrix inversion and other operations. This approach can also significantly increase speed of matrix operations.

Finally, a QR decomposition algorithm is adopted for solving linear equations. With use of the least squares approach in polar coordinates, the Q matrix in QR decomposition becomes a unitary matrix, while R is a complex non-singular upper triangular matrix. When topology and parameters of the power grid remain unchanged, admittance matrix also does not change; thus, QR decomposition does not need to be recalculated, which can greatly improve calculation efficiency.

C. Bad Measurement Identification

Bad data are inevitable in measurements; consequently, bad data identification is an essential part of SE. The largest normalized residuals (LNR) test [1] is a classical method for handling bad data. The basic idea is to estimate measurement errors from residuals and eliminate the measurement point with largest error. Definitions of residuals and normalized residuals are given as follows:

$$\boldsymbol{r} = \boldsymbol{Z}_{\mathrm{m}} - \boldsymbol{A}\hat{\boldsymbol{U}} \tag{9}$$

$$r_i^N = \frac{|r_i|}{\sqrt{\sigma_{ii}}} \tag{10}$$

where σ_{ii} is the covariance of r_i. Normalized residuals have standard normal distribution. Therefore, the largest normalized residual can be compared against a statistical threshold to determine whether it corresponds to bad data. However, the SE procedure needs to be repeated once such an identified bad measurement has been removed from the measurement set. For a large-scale power system or a system with a large amount of bad data, it may be necessary to repeat this process dozens of times to identify and remove all bad data, resulting in heavy computational burden.

A practical method of bad data identification is proposed to address this problem. A dual-pipeline processing method is adopted in the system. Specifically, the main pipeline performs the SE procedure, while the other pipeline identifies and processes bad measurements. The basic idea is to process a short period of measurements from before current time to identify bad data and obtain a set of measurement points with bad data during this period. Then, indexes of bad measurement points are stored and updated in the real-time database, where they can be accessed by the state estimator at any time. Detailed procedure of bad data identification method is described as follows.

In Fig. 3, t_{now} is always measurement time of the latest measured PMU data, t denotes the time at which a snapshot of the bad data is being processed, and S is the index set of bad measurement points. Steps are as follows:

Step 1: model and topology of the power grid are read. S is initialized as an empty set, and $t = t_{\mathrm{now}}$.

Step 2: PMU measurements at time t are read from the database, and bad measurements are removed according to S.

Step 3: WLS LSE is carried out using equation (8).

Step 4: normalized residuals are computed, and the maximum normalized residual $r_{\max}^N$ is determined.

Step 5: If $r_{\max}^N > c$, the corresponding measurement is eliminated from measurement set, index of this measurement point is added to S, and procedure returns to Step 3. Otherwise, it proceeds to Step 6. Here, c denotes the chosen threshold, for instance, 3.0.

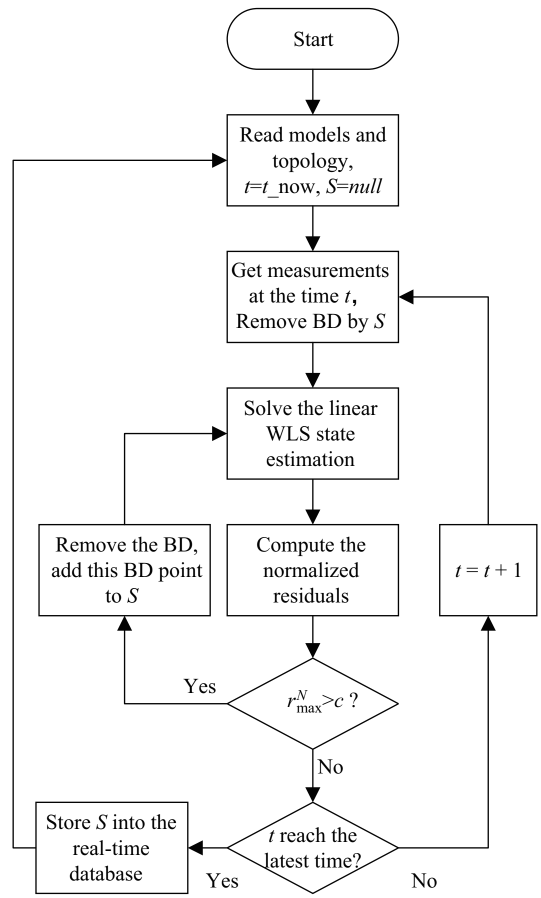


Fig. 3. Bad Data (BD) Identification.

Step 6: Check if t has reached latest time. If no, that is $t < t_{\mathrm{now}}$, t is increased by 1 time period, and procedure returns to Step 2 to continuously identify new bad measurements in the next data frame. If t has reached latest time, which means this round of identification has come to an end, then S is stored in the database and procedure returns to Step 1.

When a round of bad data detection procedure ends, measurement points marked as bad data in the previous round will be cleared, that is, S will be refreshed. At this time, if corrupted measurement point is no longer recognized as bad data, the point will be returned into the SE routine.

It should be noted this method is based on a simple idea that measurement points that experienced bad data are likely to produce bad data again, which has been verified by field data. Results obtained with this method are conservative and may exclude correct measurements.

D. Distributed Local Topology Check

Existing methods of topology analysis require measurements of switches and breakers from SCADA; however, the time interval of this system is too long to track topology change in real time. Although PMU data do not contain circuit-breaker/switch status information, voltage and current measurements can be used to assist in real-time monitoring of changes in system topology [20].

Transmission lines between substations in the power system are fixed; therefore, network topology depends on whether each of these lines is in "service" or "outage" state. Generally, an EHV substation has only a maximum of two busbars connected to the network. If the two busbars are electrically connected to each other, then substation is equivalent to a single node. If the two busbars are not connected to each other, then substation is equivalent to two nodes. In a check of local network topology, it is necessary to determine whether each substation is equivalent to one or two nodes, confirm whether each transmission line is in "service" or "outage" state, and then distinguish to which node each line belongs. It should be noted that because of installation positions of most potential transformers (PTs), the PMU will actually measure voltage phasors of each line connected to busbars instead of voltages of busbars themselves. Under the assumption there are no bad data in current measurements of all lines at a local plant/substation, the following rules can be established:

Rule 1: Sum of the currents on all lines, including the lines connected to transformers, should be zero. In practice, considering measurement error and the self-consumption of substation/plant, absolute value should be less than a certain threshold η_1 that is close to zero.

Rule 2: If the plant station is equivalent to 2 nodes (represented by nodes A and B), the sum of the currents on the lines belonging to node A (B) should be zero or, in practice, less than a certain threshold $\eta_2(\eta_3)$ that is close to 0.

Rule 3: voltages on all lines belonging to node A or node B should be equal. In practice, maximum amplitude (angle) difference of these voltage phasors should be less than a certain threshold $\eta_4(\eta_5)$ that is close to 0.

The above rules can be used to determine whether the number of equivalent nodes for a substation is one or two and connectivity between the lines and nodes. Operating state of a line is judged based on the following rule:

Rule 4: If current amplitude on the line is greater than a certain threshold η_6, then line is in "service" state; otherwise, it is in "outage" state.

Thresholds η_1–η_6 should be determined in accordance with rated values of the devices and PMU measurement errors. The proposed distributed topological check method can avoid topological errors and delays caused by use of SCADA digital measurements in traditional topological analysis and provides a practical way to quickly and accurately obtain network topology of a power system.

E. Overall Procedure

PMU-based SE procedure involves acquisition and storage of various types of data, and all the proposed modules must work together to complete the SE task. Operation procedure and the data flow are summarized below.

Step 1: grid model data are read and parsed from CIM files in static database, which is a very time-consuming process. Fortunately, this process needs to be executed only once when program is initialized.

Step 2: digital SCADA measurements of breakers and switches are obtained from the OS2 platform.

Step 3: Based on statuses of the switches and breakers, traditional breadth-first search method is used to carry out topological analysis to transform node-breaker model of system into a bus-branch model.

Step 4: latest PMU measurement data are read from the WAMS time series database, and latest set of bad measurement points is read from real-time database. Then, measurement points marked as bad data are removed from measurement set.

Step 5: A topological check is carried out using PMU measurements. If some measurements for a substation are excluded as bad data, then local topological analysis cannot be performed for this substation because there are not enough PMU measurements available.

Step 6: LSE is performed based on PMU measurements, and results are stored in the WRDB. If grid parameters, network topology and measurement points have not changed, then matrix $\boldsymbol{A}$ is constant, which makes calculations very efficient. The proposed bad data processing method can also avoid frequent changes in measurement points and improve efficiency.

Step 7: procedure returns to Step 4 when new PMU data arrive and returns to Step 2 when new SCADA data arrive. Note these two steps occur in parallel, which means that if PMU data and SCADA data arrive at the same time, the above operations will be performed simultaneously.

IV. Simulation Test

In this section, the accuracy and efficiency of the proposed improved SE method is verified through simulation tests.

A. Accuracy Validation

To verify accuracy of the proposed method, tests were carried out based on IEEE test cases and simplified CSG EHV power network, because we cannot obtain true values in the real power grid. The proposed improved linear state estimator was compared with classical linear weighted least squares method (WLS) and complex domain WLS (CWLS) [15]. To ensure statistically sound results, 1000 Monte Carlo simulations were performed. Simulated measurements were generated by adding zero-mean Gaussian noise to true values. Statistics of the noise are shown in Table I.

Mean total vector error (MTVE) index is defined for estimates as shown below to evaluate overall estimation accuracy in tests:

$$\mathit{MTVE}=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{L}\sum_{j=1}^{L}\frac{|x_{\mathrm{e},ij}-x_{i\mathrm{T}}|}{|x_{ij}-x_{i\mathrm{T}}|}\right) \tag{11}$$

where $x_{\mathrm{e},ij}$ and x_{ij} denote estimated and measured values, respectively, of i-th state variable in the j-th sample; $x_{i\mathrm{T}}$ is true value of the i-th state variable; L is number of samples; and n is number of state variables.

Simulation results for the CSG system are shown in Table II, which presents the MTVEs for voltage and current (MTVE_U and MTVE_I), sum of the absolute biases of active and reactive power for all buses (PB_P and PB_Q), maximum

TABLE I
STANDARD DEVIATIONS OF MEASUREMENT ERRORS

Error	Voltage		Current	
	Magnitude	Phase	Magnitude	Phase
Assumed	0.3%	0.4°	0.4%	0.7°

TABLE II
TESTS ON THE CSG SYSTEM

	WLS	CWLS	Proposed
MTVE_U	0.3472	0.2462	0.2279
MTVE_I	2.2446	0.7764	0.5620
PB_P (MW)	10427.77	9714.67	6246.24
PB_Q (MW)	4757.86	4016.69	2832.50
Max Error (p.u.)	0.0117	0.0087	0.0068
Time (s)	0.0451	0.0117	0.0169

errors among all estimated states, and average computation time for each algorithm. Algorithms are implemented in the same way and with the same measurement data.

It can be concluded from Table II the proposed method improves estimation accuracy for voltage phasors. Current phasors are also significantly improved, and maximum error among all states is reduced by more than 20% compared with other LSE methods. It can also be seen from power biases the proposed method has a significant advantage in power flow calculation.

B. Computational Efficiency Validation

As shown in Table II, computation time for LSE is on the level of milliseconds, longer than of the CWLS method but less than of the WLS method; therefore, it can still meet the requirements for real-time SE. Further simulations were performed on various test cases of different scales to compare computation times required for traditional SE based on SCADA data and proposed algorithm.

It can be seen from Table III the PMU-based SE algorithm is approximately one hundred times higher in computational efficiency than traditional SCADA-based SE algorithm, and the larger the grid size is, the more obvious the advantage of the proposed algorithm. This is mainly because traditional SE method requires iterative calculations, and the Jacobian matrix needs to be reconstructed in each calculation, causing this method to take more time. As size of the power grid increases, number of iterations also increases, further increasing computation time.

TABLE III
TESTS ON IEEE SYSTEMS

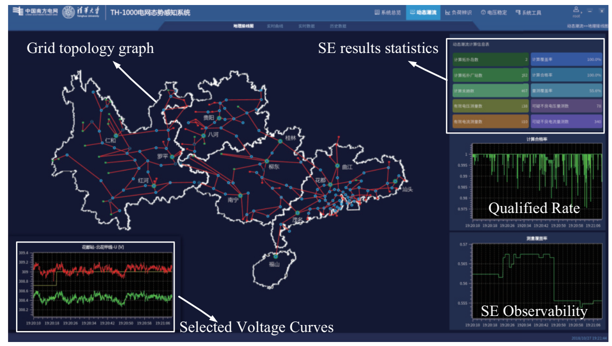
Case	Computation time/ms		Ratio
	SCADA-SE	PMU-LSE	
IEEE 14	7.2	0.21	34.2
IEEE 39	25	0.38	65.8
IEEE 57	53	0.57	93.0
IEEE 118	286	1.91	149.7
IEEE 300	1675	7.46	224.5

V. FIELD OPERATION RESULTS

In this section, field operation results of PMU-based SE system in CSG under both normal and abnormal circumstances are presented and analyzed, which have proven effectiveness and advantages of the PMU-based SE.

A. Field Results Under Normal Circumstances

The system was successfully put into operation at CSG in 2018. First, a snapshot of typical PMU-based SE results under normal circumstances is presented, as shown in Fig. 4. Key statistics of SE results are shown in Table IV.



(a)

(b)

Fig. 4. Screenshots of PMU-based SE system in CSG. (a) PMU-based SE system dashboard. (b) Comparison of SCADA based SE results, PMU based SE results and PMU measurements.

TABLE IV
STATISTICS OF SE RESULTS AT 15:48:25:000 ON AUG. 17, 2018

Statistic	Value
Number of Valid Topological Islands	2
Number of Topological Nodes	231
Number of Branches	464
State Estimation Observability	98.7%
Valid Measurement Coverage	67.8%
Number of Bad Measurements	291
Qualified Rate	97.9%

Here, Qualified Rate of SE [27] is adopted to evaluate performance of estimated results. Qualified Rate is the major indicator used to assess SE performance in the Chinese power industry and is defined as follows:

$$\tau = \frac{N_{\mathrm{q}}}{N} \times 100\% \tag{12}$$

where N is the total number of valid measurements and N_{q} is the number of qualified measurement points, whose residuals are less than threshold values. Qualified rate was empirically derived based on field experience without a theoretical basis, but it has proven to be a practical and effective criterion that provides good guidance in improving quality of SE [28]. Qualified rate of this snapshot is 97.9%, which means the estimation can be deemed successful. The number of measurement points identified as bad data is 291, accounting for 25% of all measurement points ($231 + 464 \times 2$). In addition, missing measurements caused by other problems, such as equipment failure and communication failure, cause total available measurements to account for only 67.8% of all measurement points, although all substations are equipped with PMUs. Fortunately, such effective measurement coverage still guarantees observability of almost the whole network, i.e., 98.7% of all substations.

Measured and estimated values of voltage phasors that are identified as bad measurements are compared in Fig. 5. As seen, most of the identified bad voltage measurements exhibit obvious errors in magnitude or phase. Besides, there are even some substations for which measured magnitudes and phases of the voltage are missing. PMU SE corrects most bad measurements in real time, and fills in missing data, increasing network observability from 67.8% to 98.7%.

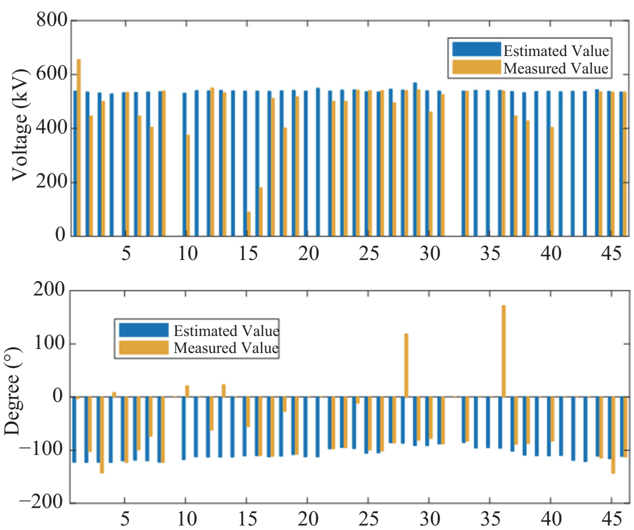


Fig. 5. Comparison between measured and estimated values of bad voltages.

Five-second traces around this snapshot for Xingshi substation, which is represented as 26$^{\text{th}}$ node in Fig. 5, are plotted in Fig. 6. Significant bad data appear in phase measurements. In addition, a slight disturbance occurred and quickly disappeared, and PMU-based estimator could effectively track system dynamics. However, current SCADA-based state estimator of EMS, which runs every minute, could not detect this disturbance because the SCADA system did not report measurements during the disturbance.

We continue by lengthening timeline to one week to examine corresponding SE results. Fig. 7 shows daily distributions of Qualified Rate from December 15th to December 21st, 2018, which are represented by box diagrams. In this figure, the red line in each box indicates the mean qualified rate, with the upper (lower) edge of the box representing the 95% (5%) quantile and the upper (lower) whisker representing the 99.5% (0.5%) quantile. Mean value of qualified rate is above 98%, and qualified rate is greater than 97% for more than 90% of the snapshots.

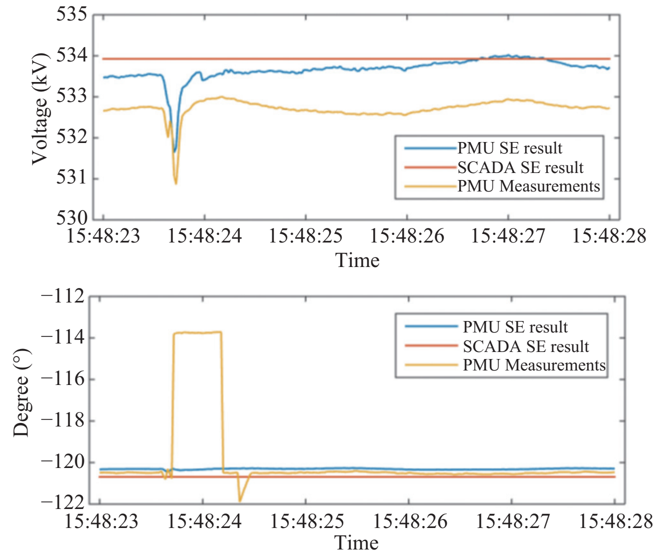


Fig. 6. Comparison between raw and estimated data at Xingshi substation.

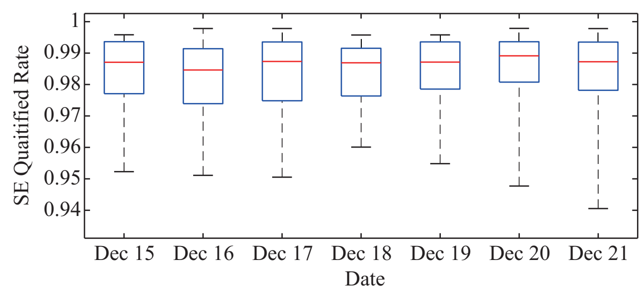


Fig. 7. Daily distributions of the SE qualified rate.

It is demonstrated that a PMU-based SE system can maintain normal and stable operation with satisfactory SE Qualified Rate for a long time.

In summary, PMU-based state estimator can successfully identify bad data and provide reliable system states in real time, which can be used for real-time monitoring and enhancing observability.

B. Field Results Under Fault Conditions

On August 19, 2018, at approximately 16:03:22, there was an abrupt outage in the 500 kV Line A between Shuixiang and Guancheng because of a short-circuit fault. There are two transmission lines, i.e., Line A and Line B, between Shuixiang and Guancheng. Fig. 8 shows the 5-minute traces of Line B's voltage and power during the fault as obtained via PMU-based SE and traditional SCADA-based SE. The short-circuit fault caused voltage to drop to below 400 kV while exhibiting continuous oscillations. PMU-based SE results were accurate and timely, whereas SCADA-based estimation results always remained near normal value, failing to reflect voltage drop and fluctuation. Because of the fault in Line A, power flow was transferred to Line B. Thus, Line B exhibited a large increase in power with multiple cycles of oscillations, as seen

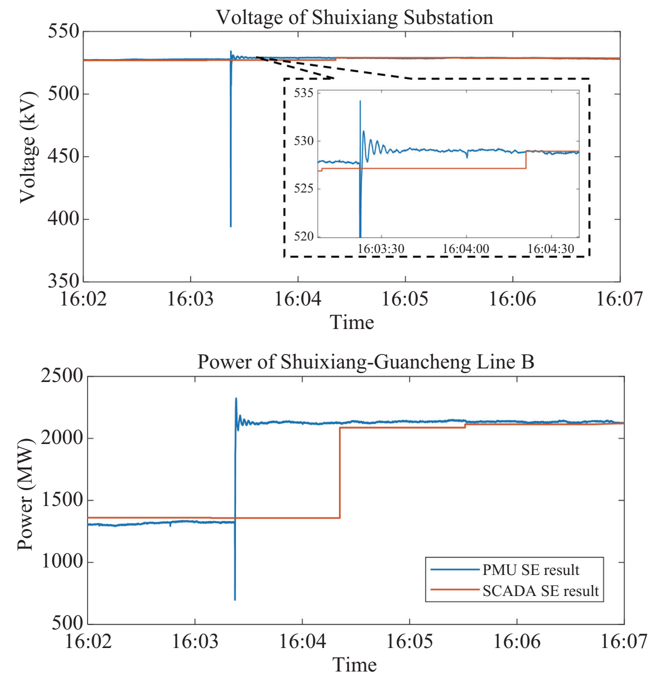


Fig. 8. Comparison of SE results during a fault.

from PMU-based SE results shown in Fig. 8, while traditional SE results were distorted and delayed. There was more than 1 minute of time when actual line transmission power had risen to more than 2100 MW but SCADA-based SE results were less than 1400 MW.

In general, short-circuit faults cause severe changes and oscillations in voltage and power. Results obtained via PMU-based SE are accurate and timely, but SCADA-based SE results cannot reflect rapid changes in voltage and power due to lower measurement frequency and longer SE interval. Thus, at the moment of a disturbance, the proposed PMU-based state estimator can quickly and accurately capture dynamic processes that traditional state estimator cannot track.

VI. Conclusion

As PMUs have been widely deployed in modern power grids, high-precision and real-time power system SE is becoming feasible. This paper develops a systematic solution for PMU-based SE in large-scale power grids. Specifically, this solution consists of three modules: an improved LSE algorithm, a practical technique for processing bad measurements and distributed local topology check method. Design and implementation of the proposed PMU-based SE method for real-world CSG are also illustrated. Extensive test and field results demonstrate PMU-based SE contributes to monitoring, protection, and operation of a power system in many respects, which include but are not limited to the following:

1) Providing real-time high-resolution state estimates for a power system, which cannot be obtained with a traditional state estimator.

2) Compensating for and mitigating loss of system observability caused by lack of measurement data. Furthermore, results for measurement points with bad data obtained via the PMU-based SE method can provide guidance for maintenance and conditioning of PMU devices.

3) Providing cleaned data for other applications. Many studies of wide-area analysis, protection and control rely on accurate PMU measurements. In fact, other advanced synchrophasor-based applications, e.g., transient stability assessment and online load parameter identification, have also been developed and deployed on the presented platform.

From the temporal perspective, the proposed PMU-based SE method can reduce delay for online monitoring and situational awareness of power systems from minutes to seconds or even tens of milliseconds. We believe that PMU-based SE will serve as an important cornerstone for future power system energy management.

References

[1] A. Abur and A. Gómez Expósito, *Power System State Estimation: Theory and Implementation*. Boca Raton: CRC Press, 2004.

[2] A. G. Phadke, “Synchronized phasor measurements in power systems,” *IEEE Computer Applications in Power*, vol. 6, no. 2, pp. 10–15, Apr. 1993.

[3] A. G. Phadke, J. S. Thorp, R. F. Nuqui, and M. Zhou, “Recent developments in state estimation with phasor measurements,” in *Proceedings of 2009 IEEE/PES Power Systems Conference and Exposition*, 2009, pp. 1–7.

[4] C. Lu, B. N. Shi, X. C. Wu, and H. B. Sun, “Recent developments of PMU/WAMS in China,” *IEEE Power & Energy Magazine*, vol. 13, no. 5, pp. 60–71, Sep./Oct. 2015.

[5] M. Wu and L. Xie, “Online detection of low-quality synchrophasor measurements: a data-driven approach,” *IEEE Transactions on Power Systems*, vol. 32, no. 4, 2817–2827, Jul. 2017.

[6] A. G. Phadke, J. S. Thorp, and K. J. Karimi, “State estimlatjon with phasor measurements,” *IEEE Transactions on Power Systems*, vol. 1, no. 1, pp. 233–238, Feb. 1986.

[7] G. Tian, Y. Gu, Z. Yu, Q. Zhang, D. Shi, and Q. Zhou, “Enhanced Denoising Autoencoder-aided Bad Data Filtering for Synchrophasor-based State Estimation.” *CSEE Journal of Power and Energy Systems*, vol. 8, no. 2, pp. 640–651, Mar. 2022.

[8] F. Chen, X. S. Han, Z. Y. Pan, and L. Han, “State estimation model and algorithm including PMU,” in *Proceedings of the 2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, 2008, pp. 1097–1102.

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

[10] J. B. Zhao, G. X. Zhang, K. Das, G. N. Korres, N. M. Manousakis, A. K. Sinha, and Z. Y. He, “Power system real-time monitoring by using PMU-based robust state estimation method,” *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 300–309, Jan. 2016.

[11] L. Zhao and A. Abur, “Multi area state estimation using synchronized phasor measurements,” *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 611–617, May 2005.

[12] S. Chakrabarti, E. Kyriakides, G. Ledwich, and A. Ghosh, “Inclusion of PMU current phasor measurements in a power system state estimator,” *IET Generation, Transmission & Distribution*, vol. 4, no. 10, pp. 1104–1115, Oct. 2010.

[13] X. Q. Gu, H. W. Kang, and H. X. Cao, “The least-square method in complex number domain,” *Progress in Natural Science*, vol. 16, no. 3, pp. 307–312, Mar. 2006.

[14] Y. F. Cheng, C. Lu, K. Men, and L. Tu, “Application of the complex algorithm in PMU-only state estimation,” in *Proceedings of 2014 International Conference on Power System Technology*, 2014, pp. 349–354.

[15] J. J. Lin, Y. S. Su, Y. F. Cheng, C. Lu, L. P. Zhu, H. Huang, and Y. S. Liu, “A robust complex-domain state estimator using synchrophasor measurements,” *International Journal of Electrical Power & Energy Systems*, vol. 96, pp. 412–421, Mar. 2018.

[16] M. Göl and A. Abur, “LAV based robust state estimation for systems measured by PMUs,” *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1808–1814, Jul. 2014.

[17] G. Y. Tian, Y. Z. Gu, Z. Yu, Q. B. Zhang, D. Shi, Q. Zhou, and Z. W. Wang, "Enhanced denoising autoencoder-aided bad data filtering for synchrophasor-based state estimation," *CSEE Journal of Power and Energy Systems*, vol. 8, no. 2, pp. 640–651, Mar. 2022.

[18] Y. Z. Lin and A. Abur, "Robust state estimation against measurement and network parameter errors," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4751–4759, Sep. 2018.

[19] X. M. Bian, X. R. Li, H. M. Chen, D. Q. Gan, and J. J. Qiu, "Joint estimation of state and parameter with synchrophasors—part II: parameter tracking," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1209–1220, Aug. 2011.

[20] K. D. Jones, J. S. Thorp, and R. M. Gardner, "Three-phase linear state estimation using phasor measurements," in *Proceedings of 2013 IEEE Power & Energy Society General Meeting*, 2013, pp. 1–5.

[21] L. Zhang, A. Bose, A. Jampala, V. Madani, and J. Giri, "Design, testing, and implementation of a linear state estimator in a real power system," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1782–1789, Jul. 2017.

[22] E. R. Fernandes, S. G. Ghiocel, J. H. Chow, D. E. Ilse, D. D. Tran, Q. Zhang, D. B. Bertagnolli, X. C. Luo, G. Stefopoulos, B. Fardanesh, and R. Robertson, "Application of a phasor-only state estimator to a large power system using real PMU data," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 411–420, Jan. 2017.

[23] M. Pignati, M. Popovic, S. Barreto, R. Cherkaoui, G. D. Flores, J. Y. Le Boudec, M. Mohiuddin, M. Paolone, P. Romano, S. Sarri, T. Tesfay, D. C. Tomozei, and L. Zanni, "Real-time state estimation of the EPFL-campus medium-voltage grid by using PMUs," in *Proceedings of 2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference*, 2015, pp. 1–5.

[24] J. B. Zhao, S. B. Wang, L. Mili, B. Amidan, R. K. Huang, and Z. Y. Huang, "A robust state estimation framework considering measurement correlations and imperfect synchronization," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4604–4613, Jul. 2018.

[25] *IEEE Standard for Synchrophasor Measurements for Power Systems*, C37.118.1-2011, 2011.

[26] J. T. Zhou, O. C. Au, X. Fan, and P. Hon-Wah Wong, "Secure Lempel-Ziv-Welch (LZW) algorithm with random dictionary insertion and permutation," in *Proceedings of 2008 IEEE International Conference on Multimedia and Expo*, 2008, pp. 245–248.

[27] Y. Guo, B. M. Zhang, W. C. Wu and H. B. Sun, "Reliability assessment of state estimation in practical power system," *Automation of Electric Power Systems*, vol. 46, no. 1, pp. 155–160, Jan. 2017.

[28] F. J. Bao, M. Wang, Q. Li, L. Wang, and X. C. Ma, "Evaluation of power grid data quality based on state estimation results," in *Proceedings of 2015 International Forum on Energy, Environment Science and Materials*, 2015.

Junjie Lin (S'16–M'23) received Bachelor and Ph.D. degrees from Tsinghua University, Beijing, China, in 2015 and 2020, respectively. He is currently an Associate Professor with the College of Electric Engineering and Automation, Fuzhou University, Fuzhou, China. His main research interests include power system operation and stability, synchrophasor measurement technologies.

Chao Lu (M'05–SM'12) received the B.E. and Ph.D. degrees in Electrical Engineering from Tsinghua University, Beijing, China, in 1999 and 2005, respectively. Currently he is an Associate Professor with the Department of Electrical Engineering, Tsinghua University. His research interests include data-driven power system analysis and control, and battery energy storage system control and optimization

Yingduo Han (SM'95) is an Academician with the Chinese Academy of Engineering. He has made many important contributions to several fields of power system research like flexible transmission and distribution, WAMS, control technology in China

Wenchao Song (S'20) received the B.S. degree in Electrical Engineering from Huazhong University of Science of Technology National, Wuhan, China, in 2019. He is currently pursuing the Ph.D. degree in Electrical Engineering at Tsinghua University, Beijing, China. His main research interests include power system state estimation, synchrophasor measurement technologies.

Yinsheng Su received the B.S. and M.S. degrees in Electrical Engineering from Shanghai Jiaotong University, Shanghai, China, in 1999 and 2002, respectively. He is currently a Senior Specialist with China Southern Grid Co. Ltd., Guangzhou, China. His research interests include electric power system operation and control.

He Huang received the B.E. and master degrees in Electrical Engineering from Tsinghua University, Beijing, China, in 1995 and 2003, respectively. Currently he is deputy general manager of CSG Power Dispatching Control Center. His research interests include power system analysis, protection and control.

Licheng Li received the B.S.E.E. degree from Tsinghua University, Beijing, China, in 1967. He is currently an Academician with the Chinese Academy of Engineering, Beijing. He is also the Secretary of the Expert Committee of the China Southern Power Grid, Guangzhou. His major research fields include HVDC/HVAC transmission networks, paralleling system operation and stability, and wide-area measurement and control.