This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.17775/CSEEJPES.2020.06860, CSEE Journal of Power and Energy Systems

Data-Driven Detection and Identification of Line Parameters with PMU and Unsynchronized SCADA Measurements in Distribution Grids

Jinping Sun, *Student Member*, *IEEE*, Qifang Chen, *Member*, *IEEE*, Mingchao Xia, *Senior Member*, *IEEE*

***Abstract*—Line parameters play an important role in the control and management of distribution systems. Currently, phasor measurement unit (PMU) systems and supervisory control and data acquisition (SCADA) systems coexist in distribution systems. Unfortunately, SCADA and PMU measurements usually do not match each other, resulting in inaccurate detection and identification of line parameters based on measurements. To solve this problem, a data-driven method is proposed. SCADA measurements are taken as samples and PMU measurements as the population. A probability parameter identification index (PPII) is derived to detect the whole line parameter based on the probability density function (PDF) parameters of the measurements. For parameter identification, a power-loss PDF with the PMU time stamps and a power-loss chronological PDF are derived via kernel density estimation (KDE) and the conditional PDF. Then, the power-loss samples with the PMU time stamps and chronological correlations are generated by the two PDFs of the power loss via the Metropolis-Hastings (MH) algorithm. Finally, using the power-loss samples and PMU current measurements, the line parameters are identified using the total least squares (TLS) algorithm. Hardware simulations demonstrate the effectiveness of the proposed method for distribution network line parameter detection and identification.**

***Index Terms*—Line parameter detection and identification, the time skew of PMU and SCADA measurements, distribution systems, probability density function, sampling algorithm.**

I. Nomenclature

K	The number of the Gaussian components in the Gaussian mixture models (GMM).
R	Resistance of the line.
X	Reactance of the line.
ΔP	Active power loss of the line.
ΔQ	Reactive power loss of the line.
A	Matrix of the current square measurements.
x_{RX}	Matrix of the line impedance.
y_{PQ}	Matrix of the power loss.
σ_{T+1}	The smallest singular value.
I_{ij}	Current phasor of line between *bus i* and *bus j*.
P_i	Active power flow of line from *bus i* to *bus j*.
Q_i	Reactive power flow of line from *bus i* to *bus j*.
U_i	Voltage phasor at *bus i*.
P_j	Active power flow of line from *bus j* to *bus i*.
Q_j	Reactive power flow of line from *bus j* to *bus i*.
$\|U_j\|$	Voltage amplitude at *bus j*.
$\|\mu_k^j\|'$	The mean of $\|U_j\|'$.
μ_k^j	The mean of $\|U_j\|$.
$(\omega_k, \mu_k, \sigma_k)$	The weight, mean, and standard deviation of the *k*th Gaussian component in the GMM probability density function (PDF).
$(\omega_k^{I_{ij}^2}, \mu_k^{I_{ij}^2}, \sigma_k^{I_{ij}^2})$	GMM PDF parameters of current square.
$(\omega_k^{P_i}, \mu_k^{P_i}, \sigma_k^{P_i})$	GMM PDF parameters of active power flow P_i.
$(\omega_k^{P_j}, \mu_k^{P_j}, \sigma_k^{P_j})$	GMM PDF parameters of active power flow P_j.
$(\omega_k^{\Delta P}, \mu_k^{\Delta P}, \sigma_k^{\Delta P})$	GMM PDF parameters of active power loss.
$t \in [1, tm]$	Time stamp.
$P(\Delta P_{t-1})$	PDF value of ΔP_{t-1} at time stamp *t*-1 for the time interval [1, *tm*].
$P(\Delta P_t, \Delta P_{t-1})$	Joint PDF value of ΔP_t and ΔP_{t-1} at time stamp *t* and *t*-1 for the time interval [1, *tm*].
$P(\Delta Q_{t-1})$	PDF value of ΔQ_{t-1} at time stamp *t*-1 for the time interval [1, *tm*].
$P(\Delta Q_t, \Delta Q_{t-1})$	Joint PDF value of ΔQ_t and ΔQ_{t-1} at time stamp *t* and *t*-1 for the time interval [1, *tm*].
$P(I_t^2)$	PDF value of I_{ij}^2 at time stamp *t* for the time interval [1, *tm*].
$P(\Delta P_t, I_t^2)$	Joint PDF value of ΔP_t and I_{ij}^2 at time stamp *t* for the time interval [1, *tm*].
$P(\Delta Q_t, I_t^2)$	Joint PDF value of ΔQ_t and I_{ij}^2 at time stamp *t* for the time interval [1, *tm*].
$q(x\|X)$	The proposal PDF with mean X and constant variance σ_q.

This work was supported by the National Key Research and Development Program under Grant 2017YFB0902900 and Grant 2017YFB0902902.
J. P. Sun, Q. F. Chen, M. C. Xia (corresponding author, e-mail: mchxia@bjtu.edu.cn) are with the School of Electrical Engineering, Beijing Jiaotong University (BJTU), Beijing 100044, China.

II. INTRODUCTION

In smart distribution systems, with a higher penetration of the distributed generation (DG), demand response (DR) enabled loads, renewable energy, and power electronics equipment, people have higher requirements for power distribution. Through careful monitoring, protection, and control of the power distribution, distribution systems can be ensured of efficient, reliable, and flexible operation [1]. In distribution grids, line parameter is one of the backbones of state estimation [2], fault location, reactive power optimization, and blackout management; therefore, a more accurate line parameter would improve power flow calculations and fault isolation. However, incorrect parameters may arise from poor line length estimation, slow updating of network changes in the database, aging, or environmental factors [3], [4]; therefore, the detection and identification of incorrect line parameters based on related measurements is of great importance. Furthermore, the accuracy of the line parameters can be directly affected by measurements.

Phasor measurement units (PMUs) are equipped with a global positioning system (GPS) receiver and render voltage and current phasors with exact time stamps [5], [6]. When a PMU is installed at a bus, the current phasors of the line connected to the other buses and the voltage phasor of this bus can be measured. The accuracy of PMU measurements is high, with a refresh rate of up to 50 (or even 100) times per second for a 50-Hz system [7]. In contrast, modern supervisory control and data acquisition (SCADA) measurements do not have exact time references and render voltage and current magnitude measurements. SCADA systems typically consist of remote terminal unit computers that can record real-time measurements and deliver this data to a control center with a communication system [8]. Owing to the SCADA architecture, its measurement value at present will not match the value of the moment when the SCADA measurement is taken, if the magnitude measurement is oscillates over time [9]. SCADA measurements taken at time t are delayed for ati + bti + cti, where ati is the period of the cyclic measurement gathering, bti is the time for the set of data to be received by the control center, and cti is the dead time between the arrival of measurements and their processing. For a given bus, the ati is different from one measurement to another with its value is between 0.1 s and 0.9 s. Furthermore, bti varies from one bus to another and its value is between 0.1 s and 0.5 s [10]. At a given time t, one SCADA measurement was taken at time $ts1<t$ that can be assigned to the time interval $[t-T, t]$ such that $0 \leq t - ts1 < T$; where T is approximately 1 s, and this delay appears reasonable [11]. In this study, the time skew of the SCADA measurements randomly delays in the range [0 s, 1 s].

Despite the advantages of the PMU system over that of SCADA, it still cannot replace the traditional SCADA system overnight, owing to the significant long-term investment and smooth operation of the SCADA system in existing power systems [12]. In the future, when a distribution system is capable of high PMU penetration, it will be possible to use sufficient PMU measurements to detect and identify line parameters. Therefore, at present, line parameter detection and identification of hybrid measurements is appealing when considering the fact that SCADA measurements do not match with PMU measurements in distribution grids.

Incorporating PMU measurements can improve the line parameter identification, especially when SCADA measurements incorrectly identify parameters [13], [14]. For transmission line parameters identification, authors [15] used an augmented state-parameter hybrid weighted least squares state estimator that was based on SCADA and PMU measurements to update the approach proposed in [16]. In [5], the line parameter estimation method required PMU measurements at one end of a given line and SCADA measurements at the other. During the short collection interval of the PMU and SCADA measurements, the magnitude of the voltage and current did not exhibit significant waveform changes; therefore, the time skew of the measurements was ignored. Considering the inconsistent sampling time of the PMU and SCADA measurements in a real system, the PMU and SCADA measurement systems were discussed separately in [17]. A method for estimating the self and mutual zero-sequence impedances for mutually coupled transmission lines was proposed using the unsynchronized three-phase current and voltage signals measured during the fault period. The unsynchronized data were aligned to complete the synchronization using a Time Time-transform (TT-transform) based on abrupt change detection [22]. Based on the magnitude measurements, a non-iterative estimation of the transmission line and transformer parameters was proposed using SCADA data [23]. Based on PMU measurements, a transmission line parameter estimation by at least one PMU was proposed by using several measurement snapshots from PMUs and estimated states provided by a hybrid state estimator, which used conventional magnitude measurements and available PMU synchronized measurements [4]. Furthermore, based on PMU measurements, an adaptive linear neuron (ADALINE) and IGG (Institute of Geodesy & Geophysics, Chinese Academy of Sciences) methods [24], robust identification methods [25]-[29], the least trimmed squares estimation method [30], a robust M-estimator method [31], and a maximum likelihood estimation method [32] were proposed to identify line parameters. These methods can address PMU measurement errors [24], PMU measurement outliers [25], [30], [31], the uncertainty in PMU measurements [32], instrument transformers errors [26]-[28], and PMU phase angle error [29]. In addition, many other methods for transmission line parameter identification [33]-[35] have been proposed using PMU measurements. Moreover, a measurement-based transmission line parameter estimation with an adaptive data selection scheme was proposed, which is applicable to both SCADA and wide area measurement system (WAMS) data [36].

The line parameter identification of hybrid measurements in the distribution and transmission networks is different. In transmission networks, line parameters include series resistance, series reactance, and shunt susceptance [4], [5], [15], [17]. However, in distribution networks the shunt susceptance is neglected [18]-[20]. A recent study used the precise time stamp of the PMU to align the SCADA measurement data at the same time section based on the instantaneous voltage values and the line parameter can be identified [37]. As is well known, SCADA renders voltage and current magnitude measurements; thus, we will investigate the SCADA magnitude measurements. In distribution grids, based on magnitude measurements, off-line tracking of series parameters [21] and regressing of the line impedance using the approximate relationship between the voltage drop amplitude and current [20] were used to identify line parameters. In addition, others have proposed a discrete dynamic Bayesian network method with advanced metering infrastructure (AMI) measurements and weather measurements [38] and a parameter estimation scheme that considers the influence of measurement and instrument transformer uncertainties with SCADA measurements [42]. Moreover, an automated determination of topology and line parameters with smart meters measurements [39], the identification of topology and line parameters with time-stamped voltage magnitude and injection samples of leaf nodes [40], and the identification of topology and line parameters without the information of voltage angles [41] can realize the parameter identification. Based on PMU measurements, iterations between the parameter estimation and topology identification were used for the joint line parameter and topology estimation, which combined PMUs and AMI measurements [18]. A PMU-based iterative line parameter estimation algorithm using only PMU data was presented, which included in the estimation model systematic measurement errors [43]. Our previous work [44] presented a classification identification method based on PMU measurements for line parameter identification under insufficient measurement conditions.

However, current approaches to the line parameter identification of hybrid measurements still pose serious limitations in distribution grids. For example, PMU measurements have exact time stamps, whereas SCADA measurements do not.

In this paper, we propose a measurement-based method for the detection and identification of parameters of a line with PMU and unsynchronized SCADA measurements in distribution systems. Owing to the nonsynchronization of the PMU and SCADA measurements, we analyze the probability distribution of PMU and SCADA measurements based on probabilistic approaches, respectively. The suspicious line parameter is detected using a probability parameter identification index (PPII). The PPII is the sum of relative errors of the means of the voltage magnitude measurements. Based on the probability distribution of measurements, we generate samples of the power loss with PMU time stamps and chronological correlations and identify the line parameter based on the current measurements from PMU device and the power-loss samples. The main contributions of this study are as follows:

1) The probability distribution of the measurements is used to solve the problem that SCADA measurements do not match with the PMU measurements. We take the SCADA measurements as the samples and the PMU measurements as the population. Then, based on probabilistic approaches, we realize line parameter detection and identification.

2) A PPII of the line parameter detection is proposed based on the probability density function (PDF) parameters of the measurements. The PPII contains the features of the line parameter and can be used to detect the line parameter. When comparing the PPII with a threshold many times over, we can determine whether the line parameter is correct. However, the proposed detection method cannot distinguish between the errors of the resistance and the reactance, but it can detect the errors as a whole.

3) A power-loss PDF with PMU time stamps is derived to model the correlation between the power loss and PMU time stamps, and a power-loss chronological PDF is derived to model the chronological correlation of the power loss. First, we build a joint PDF between the power loss and current measurements with the PMU time stamps, and a joint PDF of the power loss at the prior time stamp and the following time stamp. Then, using the conditional PDF, two PDFs of power loss can be obtained.

4) The power-loss samples with the PMU time stamps and chronological correlations are generated by the two PDFs of the power loss based on the sampling algorithm. The line parameter is identified by the power-loss samples and PMU current measurements. Hardware simulations show the proposed method is substantially effective for line parameter detection and identification when considering that the SCADA and PMU measurements do not match.

The remainder of this paper proceeds as follows: Section III details the method for dealing with the nonsynchronization of the PMU and SCADA measurements. Section IV considers the Gaussian mixture model (GMM) PDF and the Gaussian model PDF of the measurements, respectively. We derive a PPII for the line parameter detection based on the parameters of the PDFs. Section V describes the proposed approach for the line parameter identification. We derive two PDFs of the power loss, i.e., a power-loss PDF with PMU time stamps and a power-loss chronological PDF. We introduce two sampling algorithms to sample power-loss samples. Then, the line parameter is identified by the PMU current measurements and the power-loss samples with the features of the PMU time stamps and chronological correlations. Section VI presents the hardware simulation results of the proposed approach for line parameter detection and identification, and comparison analyses. Section VII presents our conclusions.

III. PMU Measurements Population And SCADA Measurements Samples

When a SCADA meter rather than a PMU meter is installed at one bus, we analyze how PMU measurements of this bus can be expressed by SCADA measurements. One measurement is produced by a SCADA meter every second [45], and 50 measurements are produced by a PMU meter every second with exact time stamps. In Fig. 1, $tx1$ denotes the time stamp of the SCADA measurement and $t1$, $t2$, … $t50$ denote the time stamps of the PMU measurements. In real distribution networks, we cannot obtain the real time of $tx1$. In this study, the PMU measurements of a fixed number in a fixed timeframe are taken as the population, and one SCADA measurement (the time stamp is $tx1$ in Fig. 1) is taken as a sample of the subpopulation of 50 PMU measurements (the time stamps are $t1$, $t2$, … $t50$ in Fig. 1) in the first second. Therefore, an unbiased estimator of the population mean of the PMU measurements can be obtained from the sample mean of the SCADA measurements in the abundant measurement snapshots. The population distribution of the PMU measurements can be obtained from the sample distribution of the SCADA measurements in the abundant measurement snapshots. This proposed solution is in accordance with the method of stratified random sampling [46], [47]. Therefore, when a SCADA is installed at one bus without installing a PMU, an unbiased estimator of the population mean of PMU measurements for this bus can be obtained by the sample mean of SCADA measurements. Moreover, the population distribution of PMU measurements for this bus can be obtained by the sample distribution of SCADA measurements.

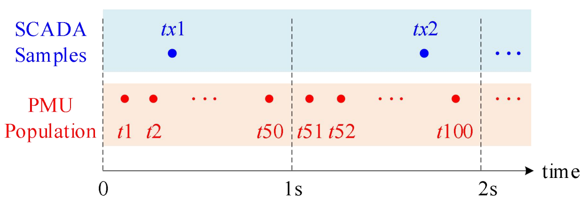


Fig. 1. Inconsistent sampling time of PMU and SCADA measurements

IV. Line Parameter Detection

A. PDF Parameters of Measurements

The GMM is the weighted finite sum of several Gaussian components. Mathematically, the PDF can be expressed as in [48]-[52].

We assume that the PDF of the measurement data is the Gaussian PDF or the GMM PDF. The advantage of the GMM approach is that the different types of the measurement distributions can be fairly represented by several normal distributions. The parameters of the GMM PDF can be determined using the expectation maximization (EM) algorithm [53]. The EM algorithm obtains the parameter set by iterating between the E-step and M-step until convergence has been reached [49], [54], [55]. However, the final solution is very sensitive to the given initial parameters when the EM algorithm converges [49]. In this study, we evaluate the performance of the EM algorithm based on the weights of each GMM, where GMMs of the measurements of the different types have the same weights for the same load fluctuation. Therefore, if the weights of different GMMs are different, then the GMM PDF parameters need to be recalculated using the EM algorithm. Through this method, we can obtain the PDF of the measurements based on the estimated parameter set.

B. PPII of Line Parameter Detection

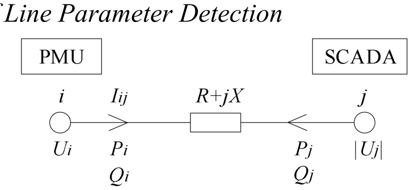


Fig. 2. PMU and SCADA measurements of the line

When a PMU meter is installed at one bus of the line and a SCADA meter is installed at another bus of the same line, the line parameter can be detected and identified. The location of the PMU and SCADA measurements in Fig. 2 is an example of the analysis. In Fig. 2, a PMU meter is installed at *bus i*, and a SCADA meter is installed at *bus j*. Note that when a PMU meter is installed at *bus j* and a SCADA meter is installed at *bus i*, the theory of the proposed method is similar to that of the locations in Fig. 2.

According to Section III, in the abundant measurement snapshots, U_i and I_{ij} are the population of the voltage phasor and current phasor, respectively, and $|U_j|$ is the samples of voltage magnitude. According to Ohm's law, the population of $|U_j|$ based on the line impedance is given as

$$|U_j|' = |U_i - I_{ij}(R + jX)| \tag{1}$$

If the line impedance is correct, the mean of the samples $|U_j|$ is the same as the mean of the population $|U_j|'$. However, if the line impedance is incorrect, the means of the samples and population differ. At *bus j*, the PDF of the voltage magnitude can be expressed using a GMM PDF or a Gaussian PDF. Comparing the mean of the samples $|U_j|$ with the mean of the population $|U_j|'$, we propose a PPII to realize incorrect line parameter detection:

$$\text{PPII} = \sum_{k=1}^{K} 100 \times \frac{\left| |\mu_k^j|' - \mu_k^j \right|}{\mu_k^j} \tag{2}$$

Here, K can be determined by the voltage magnitude, which is influenced by load fluctuation. When the PDF of the voltage magnitude is a Gaussian PDF, K is one. For a Gaussian distribution, if the PPII is greater than one threshold in most cases, this line parameter is incorrect. We determine this threshold experimentally. For the GMM distribution, this threshold of PPII is K multiplied by the threshold of the Gaussian PDF.

V. Line Parameter Identification

A. PDF Parameters of Active and Reactive Power Loss

In Fig. 2, the expression of the line parameter identification can be given as

$$I_{ij}^2(R + jX) = \Delta P + \Delta Q = (P_i - P_j) + j(Q_i - Q_j) \tag{3}$$

However, the active and reactive power loss cannot be precisely determined owing to the SCADA measurements that do not match with the PMU measurements. First, we analyze the PDF parameters of the active power loss. Next, we take the GMM PDF is an example of the analysis.

According to Section III, in abundant measurement snapshots, we can determine the GMM PDF parameters of the current square I_{ij}^2 population, the GMM PDF parameters of the active power flow P_i population, and the GMM PDF parameters of the active power flow P_j samples. Therefore, the PDF parameters of the active power loss ΔP is expressed as

$$\omega_k^{\Delta P} = \omega_k^{I_{ij}^2} = \omega_k^{P_i} = \omega_k^{P_j} \tag{4}$$

$$\mu_k^{\Delta P} = \mu_k^{P_i} - \mu_k^{P_j} \tag{5}$$

$$\sigma_k^{\Delta P} \approx \sigma_k^{I_{ij}^2} \tag{6}$$

Where the standard deviation $\sigma_k^{\Delta P}$ is equal to $\sqrt{R^2 \times (\sigma_k^{I_{ij}^2})^2}$, and the standard deviation $\sigma_k^{\Delta P}$ is approximately equal to $\sigma_k^{I_{ij}^2}$ when the resistance value in per-unit (p.u.) is approximately one.

We can determine the GMM PDF of the active power loss ΔP based on the PDF parameters. Using the same method, the GMM PDF and GMM PDF parameters of the reactive power loss ΔQ can be determined.

B. *Nonparametric Kernel Density Estimation*

The PDF of the measurements can be obtained from the nonparametric kernel density estimation (KDE), which is a data-driven method. In a univariate KDE [56]-[58], the type of kernel function has very little effect on the accuracy of the KDE. Furthermore, the Gaussian kernel function has a wide range of applications [57], [59]-[62]. In this study, a Gaussian kernel function is selected as the kernel function. For a Gaussian kernel, the optimal bandwidth can be determined as given in [58], [63].

In this study, a multivariate KDE with two-dimensional random variables [57] is used to develop a probability model with the PMU time stamps and a chronological probability model of the data. For a multivariate normal distribution kernel, the optimal bandwidth can be determined as given in [64].

C. *Probability Model of Power Loss*

1) *Chronological Probability Model*

The active power loss at time *t*-1 affects the loss at time *t*, and the active power loss at time *t* affects that at time *t*+1. In other words, the active power loss at the preceding time stamp affects that at the following time stamp. Consider the pair of active power losses at time stamp *t*-1 and at time stamp *t* as an example. The probability distribution of ΔP_{t-1} has an impact on the probability distribution of ΔP_t. The conditional PDF value of ΔP_t can be expressed as

$$P(\Delta P_t \mid \Delta P_{t-1}) = \frac{P(\Delta P_t, \Delta P_{t-1})}{P(\Delta P_{t-1})} \tag{7}$$

Where $P(\Delta P_{t-1})$ can be calculated based on the univariate KDE for the time interval [1, *tm*], and $P(\Delta P_t, \Delta P_{t-1})$ can be calculated based on the multivariate KDE for the time interval [1, *tm*]. Because the power-loss data are not time series samples, the samples of ΔP_t and ΔP_{t-1} are data within the time interval [1, *tm*] when $P(\Delta P_t, \Delta P_{t-1})$ and $P(\Delta P_{t-1})$ are calculated using the KDE, which neglects the effect of one sample in the abundant measurement data.

Similarly, we can obtain the conditional PDF value of ΔQ_t .

$$P(\Delta Q_t \mid \Delta Q_{t-1}) = \frac{P(\Delta Q_t, \Delta Q_{t-1})}{P(\Delta Q_{t-1})} \tag{8}$$

Therefore, PDFs (7) and (8) are the active and reactive power-loss chronological PDFs, respectively.

2) *Probability Model with PMU Time Stamps*

The chronological PDFs of ΔP_t and ΔQ_t simply reflect the relationship between the prior time stamp and the following time stamp and do not contain the time stamps of the PMU measurements. The PMUs provide accurate synchronized phasor measurements and have a high sampling rate. In this section, we obtain the PDF of ΔP_t with the time stamps of the PMU measurements. The current I_{ij} is the PMU measurements with exact time stamps, and thus, I_{ij}^2 is the measurements with the exact time stamps. The conditional PDF value of ΔP_t with the PMU time stamps can be obtained when the current I_{ij} is obtained from a PMU meter installed at *bus i*.

$$P(\Delta P_t \mid I_t^2) = \frac{P(\Delta P_t, I_t^2)}{P(I_t^2)} \tag{9}$$

$P(I_t^2)$ is calculated based on the univariate KDE for the time interval [1, *tm*], and $P(\Delta P_t, I_t^2)$ is calculated based on multivariate KDE for the time interval [1, *tm*].

Similarly, we can obtain the conditional PDF value of ΔQ_t .

$$P(\Delta Q_t \mid I_t^2) = \frac{P(\Delta Q_t, I_t^2)}{P(I_t^2)} \tag{10}$$

In this section, the conditional PDF values of ΔP_t and ΔQ_t reflect the effects of the exact time stamps of the PMU measurements. Therefore, PDFs (9) and (10) are power-loss PDFs with PMU time stamps.

D. *Sampling Algorithm*

Acceptance-rejection sampling (ARS) [65]-[66] is a technique for generating random variates from an alternative distribution, and then these variates are accepted or rejected according to a criterion. The Metropolis-Hastings (MH) algorithm [67]-[69] is a famous Markov chain Monte Carlo (MCMC) scheme. For MH sampling, the proposal distribution should generate random variates conveniently and can be any fixed distribution [70]. In this study, we choose a normal distribution as the proposal PDF $q(x|\mathrm{X})$ with mean X and constant variance σ_q .

E. *Line Parameter Identification Algorithm*

The expression of the power loss in (3) can be written as a matrix equation:

$$\begin{bmatrix} I_{ij}^2 & \\ & I_{ij}^2 \end{bmatrix} \begin{bmatrix} R \\ X \end{bmatrix} = \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} \tag{11}$$

By MH sampling, a lot of power-loss data with PMU time stamps and chronological correlations are generated from the PDFs of (7), (8), (9), and (10). Many current square measurements can be obtained. In abundant measurement snapshots, Equation (11) can be expressed as

$$Ax_{RX} = y_{PQ} \tag{12}$$

By using the total least squares (TLS) method, both the errors of the coefficient matrix A and observed data vector y_{PQ} can be taken into account simultaneously [18]. The line parameter can be obtained through the TLS method that utilizes singular value decomposition:

$$x_{RX} = (A^T A - \sigma_{T+1}^2 I)^{-1} A^T y_{PQ} \tag{13}$$

where σ_{T+1} is the smallest singular value of the expanded sample matrix $[A, y_{PQ}]$.

To show the proposed algorithm, the line parameter identification is shown in Algorithm 1 and in the flow chart in Fig. 3.

Algorithm 1 line parameter identification

Stage 1: (conditional PDF of active power loss for the time interval [1, *tm*])
if K=1
calculate $\mu_k^{P_i}$, $\mu_k^{P_j}$, $\sigma_k^{I_{ij}^2}$
calculate $\mu_k^{\Delta P}$ and $\sigma_k^{\Delta P}$ in (5) (6)
build PDF of power loss $P(\Delta P)$
generate random samples ΔP from $P(\Delta P)$
else
calculate $\mu_k^{P_i}$, $\mu_k^{P_j}$, $\omega_k^{I_{ij}^2}$, $\sigma_k^{I_{ij}^2}$ by EM
calculate $\omega_k^{\Delta P}$, $\mu_k^{\Delta P}$ and $\sigma_k^{\Delta P}$ in (4) (5) (6)
build PDF of power loss $P(\Delta P)$
generate random samples ΔP from $P(\Delta P)$ by ARS
end if
build PDF $P(\Delta P_{tm}, \Delta P_{tm})$ based on KDE
build PDF $P(\Delta P_{tm}, I_{tm}^2)$ based on KDE
generate PDF $P(\Delta P_{tm} \mid \Delta P_{tm})$
generate PDF $P(\Delta P_{tm} \mid I_{tm}^2)$
Stage 2: (population of power loss and $t \in [1, tm]$)
generate ΔP_t from $P(\Delta P_t \mid \Delta P_{t-1})$ and $P(\Delta P_t \mid I_t^2)$ by MH
output population of active power loss ΔP_t
generate population of reactive power loss ΔQ_t using the same method in stage 1 and stage 2
Stage 3: (line parameter identification)
identify line parameter x_{RX} by TLS in (13)

VI. HARDWARE SIMULATION RESULTS

A. Hardware Simulation Settings

In this study, the PDF of the measurement data is assumed to be the Gaussian PDF or the GMM PDF; however, the PDF of the real measurements in the distribution system is unknown. This section discusses the performance of the proposed method using the real measurements of the measuring equipment. Fig. 4 shows the hardware simulation in which two terminals of the line are deployed by measuring equipment with different accuracies. The specific network topology of the hardware simulation is illustrated in Fig. 5. The equipment that is used for the hardware simulation is presented in Tables I and II. Therefore, the measuring accuracies of the voltage and current of bus 1 are ±0.1% and ±(0.1%+0.06%), respectively, which are used to simulate the PMU measurements. The measuring accuracies of the voltage and current of bus 2 are ±(0.1%+0.5%) and ±(0.1%+1%), respectively, which are used to simulate the SCADA measurements.

A memory recorder can measure data and store data in real time. The measurements of bus 1 and bus 2 have the same refresh rate (i.e., 50 measurements per second) and the exact time references, but different measuring accuracies. The measurements of bus 2 have a lower measuring accuracy than that of bus 1. Therefore, we randomly select one measurement per second as one SCADA measurement per second. The PMU measurements of bus 1 and the SCADA measurements of bus 2 are used to detect and identify the line parameters using the proposed method.

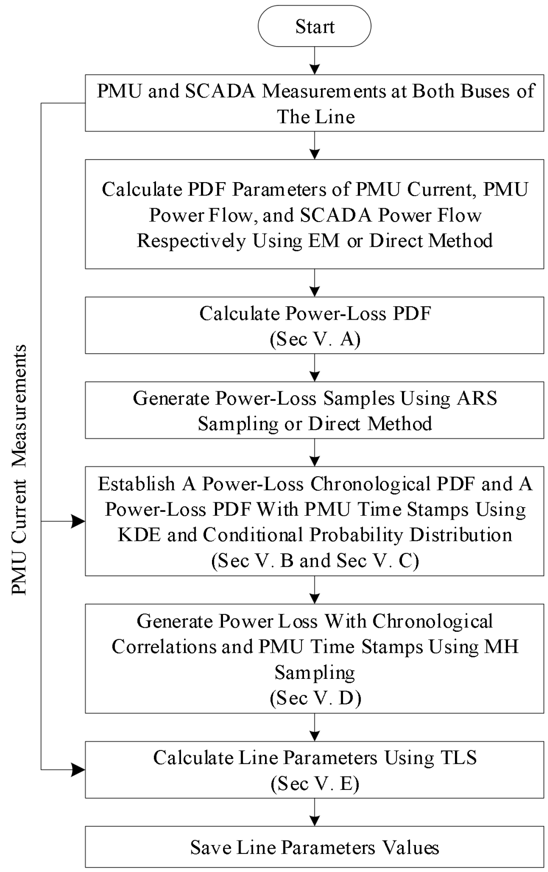


Fig. 3. Flow chart of line parameter identification

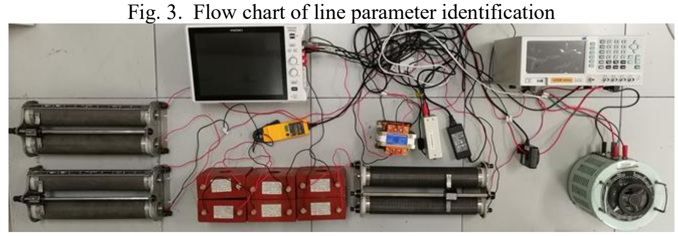


Fig. 4. Hardware simulation

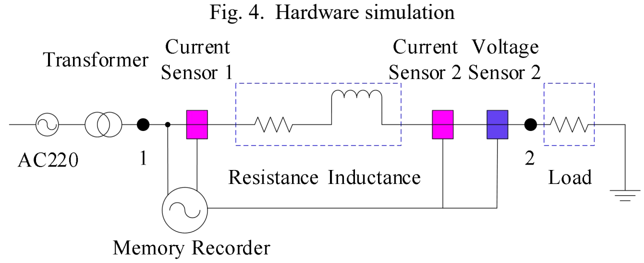


Fig. 5. Circuit of the hardware simulation

TABLE II
OTHER EQUIPMENT PARAMETERS OF THE HARDWARE SIMULATION

Name (or name in Fig. 5)	Equipment type	Parameter
Transformer	TDGC-2	220V/0-250V;50Hz
Resistance	BX7-22	6A;30Ω
Inductance	DKB-2	20A;5mH
Load	B16	6A;30Ω

TABLE I
THE MEASURING EQUIPMENT OF THE HARDWARE SIMULATION

Name (or name in Fig. 5)	Equipment type	Measuring accuracy	Ratio
Memory recorder	HIOKI MR6000	±0.1%	---
Current sensor 1	HIOKI CT6862	±0.06% (amplitude); 0.2° (phasor angle)	2V/50A
Current sensor 2	FLUKE i30s	±1%	100mV/A
Voltage sensor 2	CHNT JDG4	±0.5%	380/100V
Digital electric bridge	VC4092A	±0.3% (resistance); ±0.3% (inductance)	---

In this section, we identify the line parameters using Algorithm 1. The PDF of the measurement data is assumed to be a Gaussian PDF when K=1. In stage two of Algorithm 1, we obtain 15,000 power-loss samples with PMU time stamps and chronological correlations. However, the first sample may not follow the correct distribution and should be discarded [67]. Therefore, in this study, we choose the power-loss samples from 1,001 to 15,000. The PDF of the measurement data is assumed to be the GMM PDF when K is not one. In this GMM, the PDF of the measurements contains three Gaussian components, and the power-loss samples from 1 to 100 in each Gaussian component are discarded. Then, the remaining power-loss samples are used to identify the line parameters.

As the time skew of the SCADA measurement randomly delays, to simulate this situation, we randomly select one SCADA measurement each time. Therefore, the results of the line parameter detection and identification are different each time. However, to verify the effectiveness of the proposed method, the detection or identification of each line parameter will be performed five-times consecutively. Finally, the average value of the five results is regarded as the estimated impedance values.

B. Comparative Experiment

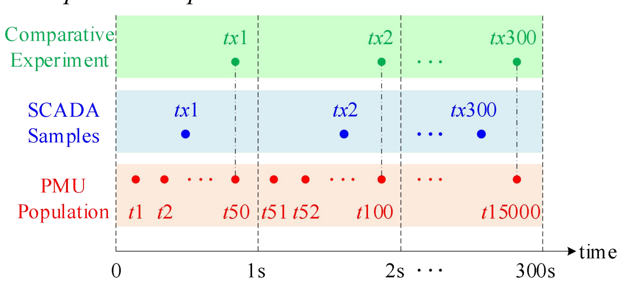


Fig. 6. Measurements of comparative experiment

To make a comparison, the tests are also performed under the condition that the delay of the SCADA measurement is neglected, i.e., the SCADA measurements obtained match with the closest PMU measurements. In Fig. 6, for each second, the time stamp of the SCADA measurement is the same as the fiftieth time stamp of the PMU measurements in the comparative experiment, i.e., the green $tx1$ is the same as $t50$. At one end of the line, there are 300 PMU measurements (the time stamps are $t50$, $t100$, … $t15000$). However, at the other end of the line, there are 300 SCADA measurements (the green time stamps are $tx1$, $tx2$, … $tx300$). The line parameter is directly identified by the TLS.

C. Experimental Results Analysis

In theory, the method proposed in this study can successfully detect line parameters and obtain more accurate line parameters. However, errors in detection and identification are inevitably caused by the following three factors:

1) The number of SCADA sample measurements: according to Section III, an unbiased estimator of the population mean can be given by the sample mean in the abundant measurement snapshots. When the number of samples is large, the probability distribution of the samples will be very close to that of the original population. In this study, 300 SCADA measurements are used as the samples of the population of 15,000 PMU measurements, which may lead to errors but can realize the detection and identification of line parameters.

2) The error of PMU and SCADA measurements: the PMU and SCADA measurements are provided by measuring equipment that possesses inherent measuring accuracy. When the measurements at both ends of the line have the same exact time stamps, the deviation of the estimated line parameter may be brought about by the related measurements.

3) The choice of the base power and base voltage: when the resistance value, in p.u., is approximately one, Equation (6) is used. However, the real value of the impedance is unknown. The impedance value, in p.u., can be approximately one, rather than exactly equal to one, using the appropriate choice of base power and base voltage. Therefore, the standard deviation of the power loss in the proposed algorithm deviates from the true standard deviation.

In this study, the real values of impedance are the measured resistance and inductance based on a digital electric bridge, i.e., VC4092A. The measuring accuracy of the VC4092A for the resistance and inductance is ±0.3% in Table I, which shows that the measured value of VC4092A can be believed. The initial value of the branch is known, but may have an error with the real value of this line. Because some real line parameters are unknown in the actual distribution network, we adjust the base power and base voltage such that the p.u. value of the initial parameter is close to one. The initial values of the line parameters are within a 20% error from their real values. Table III lists the impedance parameters, base power, and base voltage. In this study, the proposed method is analyzed using MATLAB 2018b installed on a computer with 8 GB of RAM, 2.60 GHz and 2.59 GHz processors, and an Intel Core i7-9750H CPU. The results of the line parameter detection and identification for five consecutive times are shown in Tables IV, V, and VI. We determine this threshold (0.5) based on many experiments of real measurements. For the GMM distribution, this threshold is 1.5.

As can be seen in Table IV, each PPII of the Gaussian PDF and GMM PDF of the measurements is larger than 0.5 and 1.5, respectively. Therefore, the initial impedance of the line parameters is different from the measured impedance using the VC4092A, that is, the initial impedance is incorrect. The detection results of hardware simulation are credible according to Table III, in which the initial impedance is different from the measured impedance of VC4092A. The average errors of the line parameter identification of the hardware simulation by the comparative experiment are 25.3233% and 20.6801% in Table V, and 21.9467% and 15.9152% in Table VI. However, those by the proposed method are -2.3084% and -1.1009% in Table V, and -3.7267% and -0.5621% in Table VI. It is clear from this data that the maximal error of the proposed method (3.7267%) is smaller than the minimal error of the comparative experiment (15.9152%), and the difference is more than 10%. The line parameter identification of the proposed method is close to the measured impedance based on a digital electric bridge. Therefore, in both Gaussian and GMM distributions, it is evident that for an unsynchronized issue of the PMU and SCADA measurements, the line parameter can be detected by comparing the PPII with a threshold, and the error of the line parameter identification is less than 3.7267% using the method proposed in this paper.

TABLE III
PARAMETERS SETTINGS OF HARDWARE SIMULATION

Parameters	Gaussian	GMM
Initial values of impedance (p.u.)	1+j×0.8	1+j×0.8
Measured impedance of VC4092A	11.18Ω;28.60mH	11.09Ω;28.66mH
Base power (VA)	90	90
Base voltage (V)	30	30

TABLE IV
DETECTION RESULTS OF HARDWARE SIMULATION

Distribution	PPII-1	PPII-2	PPII-3	PPII-4	PPII-5
Gaussian	2.1522	2.1907	2.2954	2.0197	1.5622
GMM	4.5543	6.5463	6.1470	4.3911	6.6003

VII. CONCLUSION

This paper presents a data-based method for line parameter detection and identification where PMU measurements have exact time stamps and SCADA measurements do not in distribution grids. Because the proposed method is a data-driven method, the PDFs obtained from the data do not need to be known previously. For the time skew of the PMU and SCADA measurements, we take SCADA measurements as the samples and PMU measurements as the population. This study derives a PPII to detect the line parameters. We generate samples of the power loss with PMU time stamps and chronological correlations from a power-loss PDF with PMU time stamps and a power-loss chronological PDF. Line parameter identification is realized by the power-loss samples and PMU current measurements. The primary conclusions of this study are as follows:

1) The time skew of the PMU and SCADA measurements for the line parameter detection and identification is solved using probabilistic approaches. Both the probability distribution of the measurements and the sampling algorithm play a crucial role in realizing the line parameter detection and identification.

2) A PPII is proposed to detect the line parameter. The proposed PPII is the sum of the relative errors of the voltage magnitude measurement means, derived from the PDF parameters of the measurements and line parameters. A PPII is derived based on the measurement-based method (i.e., the data-driven method). The detection of the accuracy of the line parameter depends on the PPII and threshold over many comparisons.

3) A power-loss chronological PDF and a power-loss PDF with PMU time stamps are derived based on the conditional probability and nonparametric KDE, which provides the PDF of the measurements based on the measurements-based method, i.e., the data-driven method. The power-loss chronological PDF is used to model the chronological correlation of the power loss at the prior time stamp and the following time stamp. Furthermore, the power-loss PDF with the PMU time stamps is used to model the correlation between the power loss and PMU time stamps.

4) The power-loss samples are sampled from two PDFs of the power loss by the MH algorithm. Using the PMU current measurements and power-loss samples with PMU time stamps and chronological correlations, we identify line parameters based on the data-driven and model-driven methods. For the Gaussian and GMM distributions, hardware simulations show that when SCADA measurements do not match with PMU measurements, the incorrect line parameter can be detected and the identification errors of the line parameters are between -3.7267% and -0.5621%.

TABLE V
IDENTIFICATION OF HARDWARE SIMULATION FOR GAUSSIAN DISTRIBUTION

Method type	Estimated impedance	Test 1	Test 2	Test 3	Test 4	Test 5	Average value
Identification of proposed method	Estimated values of impedance (p.u.)	1.0876+j×0.8908	1.1018+j×0.8883	1.1532+j×0.8790	1.0684+j×0.8756	1.0500+j×0.9068	1.0922+j×0.8881
	Relative errors of resistance (%)	-2.7207	-1.4476	3.1473	-4.4373	-6.0836	**-2.3084**
	Relative errors of reactance (%)	-0.8031	-1.0776	-2.1137	-2.4894	0.9795	**-1.1009**
Identification of comparative experiment	Estimated values of impedance (p.u.)	1.3801+j×1.0871	1.4140+j×1.0781	1.4132+j×1.0660	1.4447+j×1.1052	1.3536+j×1.0821	1.4011+j×1.0837
	Relative errors of resistance (%)	23.4432	26.4744	26.4004	29.2251	21.0736	**25.3233**
	Relative errors of reactance (%)	21.0630	20.0514	18.7069	23.0732	20.5060	**20.6801**

TABLE VI
IDENTIFICATION OF HARDWARE SIMULATION FOR GMM DISTRIBUTION

Method type	Estimated impedance	Test 1	Test 2	Test 3	Test 4	Test 5	Average value
Identification of proposed method	Estimated values of impedance (p.u.)	1.1084+j×0.8702	1.0277+j×0.8855	1.0180+j×0.8985	1.1189+j×0.9210	1.0654+j×0.8994	1.0677+j×0.8949
	Relative errors of resistance (%)	-0.0558	-7.3336	-8.2044	0.8926	-3.9323	**-3.7267**
	Relative errors of reactance (%)	-3.3082	-1.6094	-0.1646	2.3335	-0.0620	**-0.5621**
Identification of comparative experiment	Estimated values of impedance (p.u.)	1.3816+j×1.0415	1.3252+j×1.0508	1.3427+j×1.0551	1.3762+j×1.0249	1.3363+j×1.0439	1.3524+j×1.0432
	Relative errors of resistance (%)	24.5762	19.4949	21.0760	24.0893	20.4969	**21.9467**
	Relative errors of reactance (%)	15.7204	16.7588	17.2283	13.8765	15.9920	**15.9152**

The detection method presented here has its limitations: first, if the errors of the line parameters are too small, we cannot distinguish between the PPII and threshold; therefore, we may not detect the incorrect line parameters of the small errors. Second, despite the ability of the proposed detection method to detect errors that result from the resistance and reactance, it cannot detect whether the error is brought about by resistance or reactance. Nevertheless, the results of the hardware simulations are credible, and the detection method is feasible.

In our future work, we will attempt to remove some of the limitations of the detection method, and investigate the PMU deployment in distribution systems and the model of the line with a shunt capacitor to improve the scalability of the method.

ACKNOWLEDGMENT

Jinping Sun, Qifang Chen, and Mingchao Xia would like to thank Dahai Zhang and Pinghao Ni for providing part of the equipment of hardware experiment.

REFERENCES

[1] D. Ritzmann, P. S. Wright, W. Holderbaum, and B. Potter, “A Method for Accurate Transmission Line Impedance Parameter Estimation,” *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2204-2213, Oct. 2016.

[2] K. R. Mestav, J. Luengo-Rozas, and L. Tong, “Bayesian State Estimation for Unobservable Distribution Systems via Deep Learning,” *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4910-4920, Nov. 2019.

[3] A. M. Prostejovsky, O. Gehrke, A. M. Kosek, T. Strasser, and H. W. Bindner, “Distribution Line Parameter Estimation Under Consideration of Measurement Tolerances,” *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 726-735, Apr. 2016.

[4] M. Asprou, and E. Kyriakides, “Identification and Estimation of Erroneous Transmission Line Parameters Using PMU Measurements,” *IEEE Transactions on Power Delivery*, vol. 32, no. 6, pp. 2510-2519, Dec. 2017.

[5] S. S. Mousavi-Seyedi, F. Aminifar, and S. Afsharnia, “Parameter Estimation of Multiterminal Transmission Lines Using Joint PMU and SCADA Data,” *IEEE Transactions on Power Delivery*, vol. 30, no. 3, pp. 1077-1085, Jun. 2015.

[6] J. S. Li, S. D. Xu, H. Liu, and T. S. Bi, “A high-accuracy and low-complexity phasor estimation method for PMU calibration,” *CSEE Journal of Power and Energy Systems*, Jul. 2020. DOI: 10.17775/CSEEJPES.2020.01180

[7] M. Asprou, S. Chakrabarti, and E. Kyriakides, “A Two-Stage State Estimator for Dynamic Monitoring of Power Systems,” *IEEE Systems Journal*, vol. 11, no. 3, pp. 1767-1776, Sep. 2017.

[8] S. P. Carullo, and C. O. Nwankpa, “Experimental Validation of a Model for an Information-Embedded Power System,” *IEEE Transactions on Power Delivery*, vol. 20, no. 3, pp. 1853-1863, Jul. 2005.

[9] J. I. Escudero, J. Luque, and A. Carrasco, “Experimental Study on the Transmission of Measurements by Tolerance in SCADA Systems,” *IEEE Transactions on Power Delivery*, vol. 19, no. 2, pp. 590-594, Apr. 2004.

[10] B. A. Alcaide-Moreno, C. R. Fuerte-Esquivel, M. Glavic, and T. V. Cutsem, “Hybrid processing of SCADA and synchronized phasor measurements for tracking network state,” *2015 IEEE Power & Energy Society General Meeting, Denver*, CO, USA, 2015, DOI: 10.1109/PESGM.2015.7286333.

[11] B. A. Alcaide-Moreno, C. R. Fuerte-Esquivel, M. Glavic, and T. V. Cutsem, “Electric Power Network State Tracking From Multirate Measurements,” *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 33-44, Jan. 2018.

[12] J. Du, S. D. Ma, Y. C. Wu, and H. V. Poor, “Distributed Hybrid Power State Estimation Under PMU Sampling Phase Errors,” *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4052-4063, Aug. 2014.

[13] J. Zhu, and A. Abur, “Improvements in Network Parameter Error Identification via Synchronized Phasors,” *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 44-50, Feb. 2010.

[14] Y. Z. Lin, and A. Abur, “Strategic Use of Synchronized Phasor Measurements to Improve Network Parameter Error Detection,” *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5281 - 5290, Sep. 2018.

[15] M. R. M. Castillo, and J. B. A. London, “Off-line detection, identification and correction of branch parameter errors using SCADA and synchronized phasor measurements,” *2014 IEEE PES Transmission & Distribution Conference and Exposition - Latin America (PES T&D-LA), Medellin*, Colombia, 2014, DOI: 10.1109/TDC-LA.2014.6955189.

[16] M. R. M. Castillo, J. B. A. London, N. G. Bretas, S. Lefebvre, J. Prévost, and B. Lambert, “Offline Detection, Identification, and Correction of Branch Parameter Errors Based on Several Measurement Snapshots,” *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 870 - 877, May. 2011.

[17] H. B. Zhang, Z. W. Diao, and Y. F. Cui, “Identification of Power Network Branch Parameters Based on State Space Transformation,” *IEEE Access*, vol. 7, pp. 91720-91730, Jul. 2019.

[18] J. F. Yu, Y. Weng, and R. Rajagopal, “PaToPa: A Data-Driven Parameter and Topology Joint Estimation Framework in Distribution Grids,” *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4335 - 4347, Jul. 2018.

[19] J. F. Yu, Y. Weng, and R. Rajagopal, “PaToPaEM: A Data-Driven Parameter and Topology Joint Estimation Framework for Time-Varying System in Distribution Grids,” *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1682-1692, May. 2019.

[20] J. Peppanen, S. Grijalva, M. J. Reno, and R. J. Broderick, “Distribution system low-voltage circuit topology estimation using smart metering data,” *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, Dallas, TX, USA, 2016, DOI: 10.1109/TDC.2016.7519985.

[21] T. L. Williams, Y. N. Sun, and K. Schneider, “Off-line tracking of series parameters in distribution systems using AMI data,” *Electric Power Systems Research*, vol. 134, pp. 205-212, May. 2016.

[22] O. A. Gashteroodkhani, M. Majidi, and M. Etezadi-Amoli, “A Fault Data Based Method for Zero-Sequence Impedance Estimation of Mutually Coupled Transmission Lines,” *IEEE Transactions on Power Delivery*, Sep. 2020. DOI: 10.1109/TPWRD.2020.3026672

[23] A. S. Dobakhshari, M. Abdolmaleki, V. Terzija, and S. Azizi, “Online Non-iterative Estimation of Transmission Line and Transformer Parameters Using SCADA Data,” *IEEE Transactions on Power Systems*, Nov. 2020. DOI: 10.1109/TPWRS.2020.3037997

[24] A. C. Xue, H. Kong, Y. Z. Lao, Q. Xu, Y. H. Lin, L. Wang, F. Y. Xu, S. Leng, Z. Y. Yuan, and G. E. Wei, “A New Robust Identification Method for Transmission Line Parameters Based on ADALINE and IGG Method,” *IEEE Access*, vol. 8, pp. 132960-132969, Jul. 2020.

[25] J. J. Lin, J. Song, and C. Lu, “Synchrophasor Data Analytics: Transmission Line Parameters Online Estimation for Energy Management,” *IEEE Transactions on Engineering Management*, pp.1–11, Sep. 2019. DOI: 10.1109/TEM.2019.2939173

[26] H. Goklani, G. Gajjar, and S. A. Soman, “Instrument Transformer Calibration and Robust Estimation of Transmission Line Parameters using PMU Measurements,” *IEEE Transactions on Power Systems*, Nov. 2020. DOI: 10.1109/TPWRS.2020.3036605

[27] C. Wang, V. A. Centeno, K. D. Jones, and D. T. Yang, “Transmission Lines Positive Sequence Parameters Estimation and Instrument Transformers Calibration Based on PMU Measurement Error Model,” *IEEE Access*, vol. 7, pp. 145104-145117, Oct. 2019.

[28] M. Asprou, E. Kyriakides, and M. M. Albu, “Uncertainty Bounds of Transmission Line Parameters Estimated From Synchronized Measurements,” *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 8, pp. 2808-2818, Aug. 2019.

[29] A. C. Xue, F. Y. Xu, K. E. Martin, H. Y. You, J. S. Xu, L. Wang, and G. E. Wei, “Robust Identification Method for Transmission Line Parameters That Considers PMU Phase Angle Error,” *IEEE Access*, vol. 8, pp. 86962-86971, May 2020.

[30] A. Momen, Y. Chakhchoukh, and B. K. Johnson, “Series Compensated Line Parameters Estimation Using Synchrophasor Measurements,” *IEEE Transactions on Power Delivery*, vol. 34, no. 6, pp. 2152-2162, Dec. 2019.

[31] V. Milojević, S. Čalija, G. Rietveld, M. V. Ačanski, and D. Colangelo, “Utilization of PMU Measurements for Three-Phase Line Parameter Estimation in Power Systems,” *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 10, pp. 2453-2462, Oct. 2018.

[32] J. F. Fu, G. b. Song, and B. D. Schutter, "Influence of Measurement Uncertainty on Parameter Estimation and Fault Location for Transmission Lines," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 1, pp. 337-345, Jan. 2021.

[33] X. Y. Zhao, H. f. Zhou, D. Shi, H. S. Zhao, C. Y. Jing, and C. Jones, "On-line PMU-based transmission line parameter identification," *CSEE Journal of Power and Energy Systems*, vol. 1, no. 2, pp. 68-74, Jun. 2015.

[34] A. Wehenkel, A. Mukhopadhyay, J. L. Boudec, and M. Paolone, "Parameter Estimation of Three-Phase Untransposed Short Transmission Lines From Synchrophasor Measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6143-6154, Sep. 2020.

[35] P. X. Ren, H. Lev-Ari, and A. Abur, "Tracking Three-Phase Untransposed Transmission Line Parameters Using Synchronized Measurements," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4155-4163, Jul. 2018.

[36] C. G. Li, Y. P. Zhang, H. X. Zhang, Q. W. Wu, and V. Terzija, "Measurement-Based Transmission Line Parameter Estimation With Adaptive Data Selection Scheme," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5764-5773, Nov. 2018.

[37] A. D. Liu, Y. Li, W. Xie, C. G. Yang, S. R. Wang, and Z. X. Shi, ''Estimation Method of Line Parameters in Distribution Network Based on Multi-source Data and Multi-time Sections,'' *Automation of Electric Power Systems*, vol. 45, no. 2, pp. 46–54, Jan. 2021(in Chinese).

[38] W. Jiang, and H. B. Tang, "Distribution line parameter estimation considering dynamic operating states with a probabilistic graphical model," *Electrical Power and Energy Systems*, vol. 121, pp. 1-9, Oct. 2020.

[39] V. C. Cunha, W. Freitas, F. C. L. Trindade, and S. Santoso, "Automated Determination of Topology and Line Parameters in Low Voltage Systems Using Smart Meters Measurements," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5028-5038, Nov. 2020.

[40] S. Park, D. Deka, S. Backhaus, and M. Chertkov, "Learning With End-Users in Distribution Grids: Topology and Parameter Estimation," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 3, pp. 1428-1440, Sep. 2020.

[41] J. W. Zhang, Y. Wang, Y. Weng, and N. Zhang, "Topology Identification and Line Parameter Estimation for Non-PMU Distribution Network: A Numerical Method," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4440-4453, Sep. 2020.

[42] R. Dutta, V. S. Patel, S. Chakrabarti, A. Sharma, R. K. Das, and S. Mondal, "Parameter Estimation of Distribution Lines Using SCADA Measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, Sep. 2020.

[43] P. A. Pegoraro, K. Brady, P. Castello, C. Muscas, and A. V. Meier, "Line Impedance Estimation Based on Synchrophasor Measurements for Power Distribution Systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 4, pp. 1002-1013, Apr. 2019.

[44] J. P. Sun, M. C. Xia, and Q. F. Chen, "A Classification Identification Method Based on Phasor Measurement for Distribution Line Parameter Identification Under Insufficient Measurements Conditions," *IEEE Access*, vol. 7, pp. 158732-158743, Oct. 2019.

[45] 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.

[46] C. S. Ding, Q. Wu, C. T. Hsieh, and M. Pedram, "Stratified random sampling for power estimation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, no. 6, pp. 465-471, Jun. 1998.

[47] G. K. Singh, D. K. Rao, and M. G. M. Khan, "Calibration estimator of population mean in stratified random sampling," *Asia-Pacific World Congress on Computer Science and Engineering*, Nadi, Fiji, 2014, DOI: 10.1109/APWCCSE.2014.7053875.

[48] R. Singh, B. C. Pal, and R. A. Jabr, "Distribution system state estimation through Gaussian mixture model of the load as pseudo-measurement," *IET Generation, Transmission & Distribution*, vol. 4, no. 1, pp. 50-59, Jan. 2010.

[49] C. K. Reddy, H. D. Chiang, and B. Rajaratnam, "TRUST-TECH-based expectation maximization for learning finite mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 7, pp. 1146-1157, Jul. 2008.

[50] Q. C. Ding, J. D. Han, X. G. Zhao, and Y. Chen, "Missing-Data Classification With the Extended Full-Dimensional Gaussian Mixture Model: Applications to EMG Based Motion Recognition," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 8, pp. 4994-5005, Aug. 2015.

[51] L. Trailović, and L. Y. Pao, "Variance estimation and ranking of target tracking position errors modeled using Gaussian mixture distributions," *Automatica*, vol. 41, no. 8, pp. 1433-1438, Aug. 2005.

[52] M. Niknejad, H. Rabbani, and M. Babaie-Zadeh, "Image Restoration Using Gaussian Mixture Models With Spatially Constrained Patch Clustering," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3624-3636, Nov. 2015.

[53] R. Singh, B. C. Pal, and R. A. Jabr, "Statistical Representation of Distribution System Loads Using Gaussian Mixture Model," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 29-37, Feb. 2010.

[54] J. J. Qi, J. H. Wang, and K. Sun, "Efficient Estimation of Component Interactions for Cascading Failure Analysis by EM Algorithm," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3153-3161, May. 2018.

[55] P. Zeephongsekul, C. L. Jayasinghe, L. Fiondella, and V. Nagaraju, "Maximum-Likelihood Estimation of Parameters of NHPP Software Reliability Models Using Expectation Conditional Maximization Algorithm," *IEEE Transactions on Reliability*, vol. 65, no. 3, pp. 1571-1583, Sep. 2016.

[56] G. E. Constante-Flores, and M. S. Illindala, "Data-Driven Probabilistic Power Flow Analysis for a Distribution System With Renewable Energy Sources Using Monte Carlo Simulation," *IEEE Transactions on Industry Applications*, vol. 55, no. 1, pp. 174-181, 2019.

[57] Z. Y. Ren, W. Yan, X. Zhao, W. Y. Li, and J. Yu, "Chronological Probability Model of Photovoltaic Generation," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1077-1088, May. 2014.

[58] H. Nosratabadi, M. Mohammadi, and A. Kargarian, "Nonparametric Probabilistic Unbalanced Power Flow With Adaptive Kernel Density Estimator," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3292-3300, May. 2019.

[59] B. Khorramdel, C. Y. Chung, N. Safari, and G. C. D. Price, "A Fuzzy Adaptive Probabilistic Wind Power Prediction Framework Using Diffusion Kernel Density Estimators," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 7109-7121, Nov. 2018.

[60] Z. Y. Ren, K. Wang, W. Y. Li, L. M. Jin, and Y. Dai, "Probabilistic Power Flow Analysis of Power Systems Incorporating Tidal Current Generation," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 1195-1203, Jul. 2017.

[61] N. Soleimanpour, and M. Mohammadi, "Probabilistic Load Flow by Using Nonparametric Density Estimators," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3747-3755, Nov. 2013.

[62] M. L. Little, S. F. Rabbi, K. Pope, and J. E. Quaicoe, "Unified Probabilistic Modeling of Wind Reserves for Demand Response and Frequency Regulation in Islanded Microgrids," *IEEE Transactions on Industry Applications*, vol. 54, no. 6, pp. 5671-5681, Nov. 2018.

[63] B. W. Silverman, "Density Estimation for Statistics and Data Analysis," London, New York, Chapman and Hall, 1986.

[64] W. Härdle, M. Müller, S. Sperlich, and A. Werwatz, "Nonparametric and Semiparametric Models," New York, NY, USA: Springer, 2004.

[65] Christiane, Lemieux, "Monte Carlo and Quasi-monte carlo sampling," New York, NY, USA: Springer, 2009.

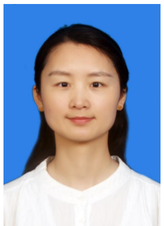
[66] L. Cheng, Z. Z. Lu, and L. G. Zhang, "Application of Rejection Sampling based methodology to variance based parametric sensitivity analysis," *Reliability Engineering & System Safety*, vol. 142, pp. 9-18, Oct. 2015.

[67] W. Hörmann, J. Leydold, and G. Derflinger, "Statistics and Computing: Automatic Nonuniform Random Variate Generation," Berlin, Germany: Springer, 2004.

[68] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, "Markov Chain Monte Carlo in Practice," London, Glasgow, Weinheim, New York, Tokyo, Melbourne, Madras: Chapman and Hall, 1996.

[69] Z. Wang, and C. Ling, "On the Geometric Ergodicity of Metropolis-Hastings Algorithms for Lattice Gaussian Sampling," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 738-751, Feb. 2018.

[70] Z. Wang, and C. Ling, "Lattice Gaussian Sampling by Markov Chain Monte Carlo: Bounded Distance Decoding and Trapdoor Sampling," *IEEE Transactions on Information Theory*, vol. 65, no. 6, pp. 3630-3645, Jun. 2019.



JINPING SUN (S'17) received the B.S. degree in Agricultural Electrification and Automation from Shenyang Agricultural University, Liaoning, China, in 2012, and the M.S. degree in Agricultural Engineering from Shenyang Agricultural University, Liaoning, China, in 2014. She is currently pursuing the Ph.D. degree in Electrical Engineering at Beijing Jiaotong University, Beijing, China. Her research interest includes distribution system line parameter identification.

QIFANG CHEN (S'13-M'17) received the B.S. and M.S. degrees in communication engineering and electric engineering from Xiangtan University, Hunan, in 2010 and 2013, respectively, and the Ph.D. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2017. Currently, he is an associate professor with the School of Electrical Engineering, Beijing Jiaotong University, Beijing, China. His research interests include micro-grid, renewable energy vehicles and integrated energy system.

MINGCHAO XIA (M'03-SM'17) received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in1998 and in 2003, respectively. He is currently a professor with the School of Electrical Engineering, Beijing Jiaotong University. His current research interests include energy internet, smart power distribution system control and optimization, power electronics in power distribution, and flexible load control.