Topology and Parameter Identification of Distribution Network Using Smart Meter and μPMU Measurements

Vedantham Lakshmi Srinivas, *Member, IEEE*, and Jianzhong Wu, *Senior Member, IEEE*

Abstract—**Incomplete and inaccurate information of network topology and line parameters affects state monitoring, analysis, and control of active distribution networks. To solve this issue, this article proposes a method for identifying distribution network topology and line parameters using the measurements obtained from smart meters (SMs) and microphasor measurement units (μPMUs) installed at various locations in a distribution network. A data-driven approach was developed, which uses a probabilistic method (unscented Kalman filter (UKF) based) and a deterministic method (Newton Raphson (NR) based) iteratively for accurate identification of network topology and parameters. The impact of the measurement noise with SMs and μPMUs is analyzed, and the acceptable noise levels are quantified. The impact of the identification algorithm on the network state estimation is examined. Moreover, optimal installation locations of the μPMU equipment are identified based on the estimation accuracy of the algorithm. The method is validated on benchmarked IEEE 33-bus and IEEE 123-bus test systems, while the impact of the renewable power injections at the different network nodes is studied as well. The qualitative and quantitative analysis is performed over the state-of-the-art methods, to highlight the effectiveness of the proposed methodology.**

Index Terms—**Active distribution network, microphasor measurement unit (μPMU), parameter identification, smart meters (SM), topology identification.**

I. INTRODUCTION

SMART meters (SMs) and microphasor measurement units (μPMUs) are being increasingly deployed in distribution networks for improved monitoring and control [1]. μPMUs offer multiple benefits to distribution network operators (DNOs), by facilitating an efficient and reliable network operation with increased penetration of renewable energy sources and unpredictable operating scenarios associated with connecting more active prosumers [2]. Various functions of the DNO, such as distributed state estimation, network planning, connection of renewable energy sources, and optimal network operation, require accurate network topology and line parameter information [3]. The network operator may refer to the blueprint, and however, due to incomplete and inaccurate information in the distribution network, the true topology and real line parameters may be unavailable or vary significantly compared to the blueprint [4]. This could significantly affect the state estimation and operation of the distribution network [5].

A. Related Work

To obtain the network topology information, Mayo-Maldonado *et al.* [7] developed a framework that uses the measurements data from μPMUs installed at various buses in a network. This framework, along with the topology estimation framework proposed by Deka *et al.* [8], is successful in the estimation of distribution network topology as well as the event detection (such as faults and line outages). However, these methods are not able to estimate the line parameters of each of the branches in the network. Moreover, the topology estimation is affected by an excessive penetration of renewable energy because of the renewable-derived uncertainty due to their stochastic nature (e.g., randomness generated by a gust of wind) and the lack of measurements associated with them in the distribution network. These frequently manifest as non-Gaussian noise in the measurements at the renewable penetrated buses in the network [6]. To tackle the noise in the measurements, He *et al.* [6] developed a hybrid framework methodology for topology identification, which proved robust under excessive penetration of renewable energy. While these methods are successful in determining the network topology, they lack providing the information of the exact branch parameters (e.g., line conductances and susceptances). With a view of providing complete network information, Gupta *et al.* [9] and Wehenkel *et al.* [10] have come up with methodologies for the estimation of topology as well as the branch parameters. Nevertheless, these methods are highly dependent on prior knowledge of the topology configuration. Notably, for the same purpose, expectation–maximization-based technique is proposed by Yu *et al.* [11], and sparse recovery-based techniques are proposed by Jafarian *et al.* [12] and Babakmehr *et al.* [13], while the data-driven techniques are also reported by Dua *et al.* [14].

Without any prior information of the topology, a novel Lasso-based algorithm is demonstrated by Ardakanian *et al.* [15]. This algorithm works well for estimating both topology

TABLE I
COMPARISONS BETWEEN EXISTING METHODS

Type of Methods	Topology identification	Robustness with respect to noise in measurements	Branch parameter estimation	Accuracy	Optimal location of synchronous device (e.g., μPMUs)	Identification with synchronous and non-synchronous devices
Data-driven Framework [7]	Considered	Capable	Incapable	Better	Not Considered	Not Considered
Graphical method [8]	Considered	Capable	Incapable	Good	Not Considered	Not Considered
Least square [10]	Not Considered	Capable	Capable	Better	Not Considered	Not Considered
Quadratic optimization [12]	Considered	Incapable	Incapable	Good	Not Considered	Not Considered
Group Lasso [23]	Considered	Capable	Incapable	Good	Not Considered	Not Considered
Optimization model [27]	Not Considered*	Incapable	Incapable	Better	Considered	Not Considered
Transfer learning model [28]	Not Considered*	Incapable	Incapable	Good	Not Considered	Not Considered

*Not capable of performing topology identification, however, capable of detecting a fault event for a given network configuration

and line parameters using μPMUs. However, an exhaustive number of data samples are required with high accuracy, and the algorithm is prone to erroneous results under the presence of biases in the measurements. An improvement adaptive Lasso is therefore proposed by the same authors (Ardakanian *et al.* [16]) to overcome the issue, and nevertheless, the identification performance is still affected as the regularization coefficient plays a vital role in the nonconvex optimization.

To use achieve the topology identification objective, using solely the SM measurements data, Zhang *et al.* [17] reported Newton Raphson (NR)-based iterative method. This is an excellent approach for balanced distribution networks. However, the method diverges if the initial condition is far away from the actual solution. Liang *et al.* [18] and Liao *et al.* [19] investigated power flow matching-based methods for topology detection using the SM measurements, which are able to identify the topology of medium voltage networks accurately. To deal with data quality issues and filter out the measurement noises effectively, Zhao *et al.* [20] investigated the deep belief network surrogate model for distribution network topology detection. This method achieved high topology identification accuracy while maintaining robustness to missing data and measurement noise under high renewable power penetration levels. Without any prior knowledge of network, Zhenyu *et al.* [21] identified distribution network topology and line impedances using SM data, while including multiple distributed generation units in the network. Furthermore, data-driven techniques for topology detection using the SM measurements data are reported by Pappu *et al.* [22] and Liao *et al.* [23], and however, the line parameters are not estimated.

B. Contributions

The approaches for the network topology identification in the distribution networks discussed in Section I-A use either μPMUs at all the buses or SMs at all the buses. However, μPMUs may not be installed at every bus due to high costs, while measurements from SMs could contain missing data and errors [24]. In a practical distribution network, the measurements data from both SMs and μPMUs may be available for the network operator. Therefore, the identification techniques using only μPMU measurements are less practical, while the techniques that rely only on SM measurements may fail when the measurements are inaccurate. Consequently, it is important to develop a method that could utilize both SM

and μPMU measurements for accurate network topology and parameters identification in a distribution network, complemented by some pseudo measurements. Some researchers, e.g., Shah and Zhao [25] and Ma and Wu [26], have contemplated a topology identification strategy using μPMU and SM measurements, and however, they lack the analysis of the method's impact on the state estimation when applied to a practical distribution network. Moreover, the methods are sensitive to measurement noise and do not consider the evaluation of the optimal locations for μPMUs installation. Jiang *et al.* [27] and Hai *et al.* [28] developed algorithms for optimal location of PMUs in context of identifying the network fault events. It is necessary to develop an algorithm for optimal location of μPMUs in context of obtaining an accurate estimation of network topology. A summary of comparison of the key existing methods on the network identification approaches is reported in Table I. In light of the discussion, the new contributions of this work are highlighted as follows.

1) A new method has been developed to accurately identify the distribution network topology information as well as the line parameters using the Kalman filtering method and NR based method iteratively.
2) Contrast to the various existing methods, this method leverages the measurements from both SMs and μPMUs, and thus, it offers an alternative to the network operator to choose μPMU measurements from the buses where SM data are missing or vice versa.
3) Moreover, the impact of various measurement errors on the identification algorithm, as well as the state-estimation performance run by the network operator, is analyzed and reported here.
4) Considering that μPMUs are expensive equipment compared to SMs, the critical points are evaluated for installation of μPMUs, which improves the method's accuracy.
5) Finally, the impact of the renewable power injection at certain network buses on the presented identification algorithm is also analyzed. Excellent identification results are observed, within 3% of the average error magnitude.

The rest of this article is organized as follows. Section II sets up the problem statement description. The complete framework for the network identification is reported in Section III. Section IV demonstrates the performance of the algorithm on an IEEE test system, with illustrative results. Finally, the conclusions are made in Section V.

II. PROBLEM FORMULATION

This work aims at precisely estimating the network topology information and the line parameters, which are uniquely determined by three sources of measurement information. The first source comes from the SMs installed at different buses in the network. The measurements from SM include the voltage magnitude, active power injection, and reactive power injection at a given bus. The second source of information comes from the μPMUs installed at a few buses, which include the voltage phasor information at a bus and optional current phasor information of the current flow in a line. The third source of information comes from some pseudo measurements of active and reactive power injections at certain buses where the SMs are not installed, considered with less accuracy (with error from $\sim$30% to 50% of the nominal value [29]). The different measurements are modeled as follows.

First, the information from the SMs, i.e., active power injection, reactive power injection, and the voltage magnitude at bus "i" at a specified instant "k," can be written as

$$P_i(k) = P_{i,n} + w_i^p; \quad \mathbb{E}\left[\left(w_i^p\right)^2\right] = \sigma_L^2 P_{i,n}^2 \quad (1a)$$

$$Q_i(k) = Q_{i,n} + w_i^q; \quad \mathbb{E}\left[\left(w_i^q\right)^2\right] = \sigma_L^2 Q_{i,n}^2 \quad (1b)$$

$$|V_i|(k) = |V_i|_n + w_i^V; \quad \mathbb{E}\left[\left(w_i^V\right)^2\right] = \sigma_V^2 |V_i|_n^2 \quad (1c)$$

where "$P_{i,n}$" and "$Q_{i,n}$" are the nominal (true) active and reactive power values in per unit (p.u.), respectively, "w_i^p" and "w_i^q" represent the corresponding measurement noises in p.u., and σ_L (unitless) is the measurement uncertainty (usually $0.01-0.02$ in case of SMs [24]). In case of pseudo measurements, σ_L holds a higher value, of about $0.3-0.5$ [29]. Similarly, $|V_i|_n$ represents the nominal value of voltage magnitude in p.u. and σ_V (unitless) represents its corresponding measurement uncertainty.

The measurements information from possible μPMUs that measure both the magnitude and the phase of voltage phasor at some buses and current phasor of the current flow in some lines at bus "i" at a specified instant "k" is

$$|V_i|(k) = |V_i|_n + w_i^V; \quad \mathbb{E}\left[\left(w_i^V\right)^2\right] = \sigma_{\text{PMU}}^2 |V_i|_n^2 \quad (2a)$$

$$\theta_i(k) = \theta_{i,n} + w_i^\theta; \quad \mathbb{E}\left[\left(w_i^\theta\right)^2\right] = \sigma_{\text{PMU}}^2 \quad (2b)$$

$$\left|I_{ij}\right|(k) = \left|I_{ij}\right|_n + w_{ij}^I; \quad \mathbb{E}\left[\left(w_{ij}^I\right)^2\right] = \sigma_{\text{PMU}}^2 \left|I_{ij}\right|_n^2 \quad (2c)$$

where $|V_i|_n$, $\theta_{i,n}$, and $|I_{ij}|_n$ represent the nominal value of voltage magnitude, voltage angle, and current flow in the line connecting buses "i" and "j" in p.u., while $\sigma_{\mu\text{PMU}}$ (unitless) represents the μPMU measurement uncertainty, which is a constant value depending on the μPMU technology (usually about $10^{-3}-\text{MU}^{-4}$ [29]). In the context of voltage magnitude measurements that are available with both SM and μPMUs at certain nodes, the μPMU measurements are considered due to high accuracy. With these different measurements in hand, mathematically, the work aims at finding the topology and line parameters (line conductances and susceptances) that best fit

the following problem for each dataset for "n" buses:

$$P_i = |V_i| \sum_{j=1}^{n} |V_j| \left(G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}\right) \quad (3a)$$

$$Q_i = |V_i| \sum_{j=1}^{n} |V_j| \left(G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}\right). \quad (3b)$$

In (3), "P_i" and "Q_i" are the active and reactive power injections of bus "i" in p.u., respectively, "$|V_i|$" is the corresponding voltage magnitude in p.u., while "G_{ij}," "B_{ij}," and "θ_{ij}" represent the per-unit line conductance, the per-unit line susceptance, and the voltage angle difference, respectively. As these parameters contain all the information about the topology of the network, this work is targeted at finding the optimal "G_{ij}" and "B_{ij}" of all the lines.

III. NETWORK IDENTIFICATION FRAMEWORK

This section describes the methodology for network identification, organized in three subsections. Section III-A describes a way to obtain the approximate conductances and susceptances of the network. Section III-B makes use of these approximate values and iteratively solves the probabilistic (using the unscented Kalman filter (UKF)-based method) and the deterministic (using the NR-based method) set of equations, to obtain more accurate conductances and susceptances. Section III-C proposes a method for finding optimal installation locations of costly μPMUs, based on the mean square errors in the estimated conductances and susceptances.

A. Obtain Approximate Conductance and Susceptance Matrices Associated With SM Nodes

Initially, the approximate conductances and susceptances are obtained, to serve as a good initial condition to the iterative problem discussed in the subsequent Section III-B. First, a regression problem is solved, which uses the active power injection, reactive power injection, and the voltage magnitude at each bus, obtained from SMs. Later, some μPMU measurements are utilized to obtain the better approximations.

1) Obtain Approximate Conductance and Susceptance Matrices Through Regression: Initially, the active and reactive power injections as shown in (3) are reformulated as follows:

$$\left[\begin{array}{cc} \dfrac{P_i}{|V_i|} & \dfrac{Q_i}{|V_i|} \end{array}\right]^T = \left[\begin{array}{cccc} G_{i1}^{\#} & G_{i2}^{\#} & \cdots & G_{in}^{\#} \\ B_{i1}^{\#} & B_{i2}^{\#} & \cdots & B_{in}^{\#} \end{array}\right] \left[\begin{array}{c} |V_1| \\ |V_2| \\ \vdots \\ |V_n| \end{array}\right] \quad (4)$$

where

$$G_{ij}^{\#} = \left(G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}\right)$$
$$B_{ij}^{\#} = -\left(G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}\right). \quad (5)$$

Considering that the SMs are installed at "r" buses ($r \in n$) and a total number of "K" consecutive data samples, it is possible to write the following expressions:

$$\left[\mathbf{PV^K}\right]_{r \times K} = \left[\mathbf{G^{\#}}\right]_{r \times n} \left[\mathbf{V^K}\right]_{n \times K} \quad (6a)$$

$$\left[\mathbf{QV^K}\right]_{r \times K} = \left[\mathbf{B^{\#}}\right]_{r \times n} \left[\mathbf{V^K}\right]_{n \times K}. \quad (6b)$$

The matrices in (6) are reported as follows:

$$[\mathbf{PV^K}] = \begin{bmatrix} \dfrac{P_1^{(1)}}{|V_1|^{(1)}} & \dfrac{P_1^{(2)}}{|V_1|^{(2)}} & \cdots & \dfrac{P_1^{(K)}}{|V_1|^{(K)}} \\ \dfrac{P_2^{(1)}}{|V_2|^{(1)}} & \dfrac{P_2^{(2)}}{|V_2|^{(2)}} & \cdots & \dfrac{P_2^{(K)}}{|V_2|^{(K)}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{P_r^{(1)}}{|V_r|^{(1)}} & \dfrac{P_r^{(2)}}{|V_r|^{(2)}} & \cdots & \dfrac{P_r^{(K)}}{|V_r|^{(K)}} \end{bmatrix};$$

$$[\boldsymbol{G^\#}] = \begin{bmatrix} G_{11}^\# & \cdots & G_{1n}^\# \\ \vdots & \ddots & \vdots \\ G_{r1}^\# & \cdots & G_{rn}^\# \end{bmatrix} \tag{7a}$$

$$[\mathbf{QV^K}] = \begin{bmatrix} \dfrac{Q_1^{(1)}}{|V_1|^{(1)}} & \dfrac{Q_1^{(2)}}{|V_1|^{(2)}} & \cdots & \dfrac{Q_1^{(K)}}{|V_1|^{(K)}} \\ \dfrac{Q_2^{(1)}}{|V_2|^{(1)}} & \dfrac{Q_2^{(2)}}{|V_2|^{(2)}} & \cdots & \dfrac{Q_2^{(K)}}{|V_2|^{(K)}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{Q_r^{(1)}}{|V_r|^{(1)}} & \dfrac{Q_r^{(2)}}{|V_r|^{(2)}} & \cdots & \dfrac{Q_r^{(K)}}{|V_r|^{(K)}} \end{bmatrix};$$

$$[\boldsymbol{B^\#}] = \begin{bmatrix} B_{11}^\# & \cdots & B_{1n}^\# \\ \vdots & \ddots & \vdots \\ B_{r1}^\# & \cdots & B_{rn}^\# \end{bmatrix} \tag{7b}$$

$$[\mathbf{V^K}] = \begin{bmatrix} |V_1|^{(1)} & |V_1|^{(2)} & \cdots & |V_1|^{(K)} \\ |V_2|^{(1)} & |V_2|^{(2)} & \cdots & |V_2|^{(K)} \\ \vdots & \vdots & \ddots & \vdots \\ |V_n|^{(1)} & |V_n|^{(2)} & \cdots & |V_n|^{(K)} \end{bmatrix}. \tag{7c}$$

The vector $[\mathbf{V^K}]$ includes the voltage magnitudes of all the buses. In this context, the voltage magnitudes at the grid supply point are considered as an approximate replacement for the voltage magnitudes of the buses where the SMs are not installed. This approximation is valid since the voltage magnitude difference between the grid supply point and any node in the distribution network is usually smaller than 5%–10% [25]. Thus, from the available measurement datasets of $[\mathbf{PV^K}]$, $[\mathbf{QV^K}]$, and $[\mathbf{V^K}]$, the approximate $[\boldsymbol{G^\#}]$ and $[\boldsymbol{B^\#}]$ matrices are obtained from (6) to (7) as

$$[\boldsymbol{G^\#}] = [\mathbf{PV^K}][\mathbf{V^K}]^T \left([\mathbf{V^K}][\mathbf{V^K}]^T \right)^{-1} \tag{8a}$$

$$[\boldsymbol{B^\#}] = -[\mathbf{QV^K}][\mathbf{V^K}]^T \left([\mathbf{V^K}][\mathbf{V^K}]^T \right)^{-1}. \tag{8b}$$

In case the SMs are installed at all the buses (i.e., $r = n$), then the $[\boldsymbol{G^\#}]$ and $[\boldsymbol{B^\#}]$ matrices obtained from (8) give the approximate information corresponding to all the lines. However, in case SM data are missing at some buses (i.e., $r < n$), the same procedure is still followed using the pseudo measurements of active and reactive power injections that are considered at the rest of "$r - n$" buses. For further denoising of $[\boldsymbol{G^\#}]$ and $[\boldsymbol{B^\#}]$ matrices, the following

optimization problem is solved [16]:

$$\min \left\| [\mathbf{PV^K}] - [\boldsymbol{G^\#}][\mathbf{V^K}] \right\|_2$$
$$\text{s.t. } [\boldsymbol{G^\#}] = [\boldsymbol{G^\#}]^T \,\&\, [\boldsymbol{G^\#}]\big|_{\mathrm{ND}} \le \boldsymbol{0} \tag{9a}$$
$$\min \left\| [\mathbf{QV^K}] + [\boldsymbol{B^\#}][\mathbf{V^K}] \right\|_2$$
$$\text{s.t. } [\boldsymbol{B^\#}] = [\boldsymbol{B^\#}]^T \,\&\, [\boldsymbol{B^\#}]\big|_{\mathrm{ND}} \ge \boldsymbol{0}. \tag{9b}$$

The constraints in (9) are valid since the $[\boldsymbol{G^\#}]$ and $[\boldsymbol{B^\#}]$ matrices are symmetric and the nondiagonal elements of $[\boldsymbol{G^\#}]$ matrix ($[\boldsymbol{G^\#}]|_{\mathrm{ND}}$) are negative, while the nondiagonal elements of $[\boldsymbol{B^\#}]$ matrix ($[\boldsymbol{B^\#}]|_{\mathrm{ND}}$) are positive. Besides, in an n-bus distribution network, the ratio of the nondiagonal element with the corresponding diagonal element is usually greater than "$1/(n-1)$," and therefore, the nondiagonal elements in $[\boldsymbol{G^\#}]$ and $[\boldsymbol{B^\#}]$ with lesser ratio are made zero.

2) Utilization of μPMU Measurements to Obtain Better Approximations: Being costly equipment, μPMUs are usually installed at only a few selected buses, which are small in number in comparison to SMs. The μPMU measurements obtained from certain buses are used to further improve the approximations obtained from the previous step using a Bayesian linear state estimator. The voltage angles are assumed zeros initially, at all the buses. The voltage phasors of all buses are given as

$$\bar{\mathbf{V}}^{\mathbf{K}} = \mathbf{V}^{\mathbf{K}}. \tag{10}$$

For "p" number of buses where μPMUs are installed

$$\bar{\mathbf{V}}^{\mathbf{K}}_{\mathrm{PMU}} = \mathbf{V}^{\mathbf{K}}_{\mathrm{PMU}} \angle \boldsymbol{\theta}^{\mathbf{K}}_{\mathrm{PMU}} \tag{11}$$

$$\mathbf{V}^{\mathbf{K}}_{\mathrm{PMU}} = \begin{bmatrix} |V_{m+1}|^{(1)} & |V_{m+1}|^{(2)} & \cdots & |V_{m+1}|^{(K)} \\ |V_{m+2}|^{(1)} & |V_{m+2}|^{(2)} & \cdots & |V_{m+2}|^{(K)} \\ \vdots & \vdots & \ddots & \vdots \\ |V_{m+p}|^{(1)} & |V_{m+p}|^{(2)} & \cdots & |V_{m+p}|^{(K)} \end{bmatrix} \tag{12a}$$

$$\boldsymbol{\theta}^{\mathbf{K}}_{\mathrm{PMU}} = \begin{bmatrix} \theta_{m+1}^{(1)} & \theta_{m+1}^{(2)} & \cdots & \theta_{m+1}^{(K)} \\ \theta_{m+2}^{(1)} & \theta_{m+2}^{(2)} & \cdots & \theta_{m+2}^{(K)} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{m+p}^{(1)} & \theta_{m+p}^{(2)} & \cdots & \theta_{m+p}^{(K)} \end{bmatrix}. \tag{12b}$$

From (2), (10), and (11)

$$\bar{\mathbf{V}}^{\mathbf{K}}_{\mathrm{PMU}} = \mathbf{C}_M \bar{\mathbf{V}}^{\mathbf{K}} + w_i^V \angle w_i^\theta \tag{13}$$

$$\mathbb{E}\left[\left(w_i^V / |V_i|_n \right)^2 \right] = \mathbb{E}\left[\left(w_i^\theta \right)^2 \right] = \sigma_{\mathrm{PMU}}^2 \mathbf{I}_{\mathrm{M}}. \tag{14}$$

In (13), "$\mathbf{C}_M$" is a matrix whose lth row has all zeros except for the mth entry, which is set to one. In practice, "$\mathbf{C}_M$" is a selection matrix that associates the μPMU measurement "$\bar{\mathbf{V}}^{\mathbf{K}}_{\mathrm{PMU}}$" with the corresponding voltage vector "$\bar{\mathbf{V}}^{\mathbf{K}}$." An appropriate voltage phasor estimate is then obtained by solving the following Bayesian linear state estimator:

$$\hat{\mathbf{V}}^{\mathbf{K}} = \bar{\mathbf{V}}^{\mathbf{K}} + \mathbf{K}(\bar{\mathbf{V}}^{\mathbf{K}}_{\mathrm{PMU}} - \mathbf{C}_M \bar{\mathbf{V}}^{\mathbf{K}}) \tag{15}$$

$$\mathbf{K} = \boldsymbol{\Sigma}_0 \mathbf{C}_M^T \left(\mathbf{C}_M \boldsymbol{\Sigma}_0 \mathbf{C}_M^T + 2 \left| V_1^{(1)} \right|^2 \sigma_{\mathrm{PMU}}^2 \mathbf{I}_{\mathrm{M}} \right)^{-1} \tag{16}$$

where "Σ_0" is the error covariance matrix of nominal voltage phasors. The voltage phasor estimate (15) gives information of both magnitudes and the phase angles at all the "K" instants. Finally, the approximate "G_{ij}" and "B_{ij}" are obtained from the $[G^{\#}]$ and $[B^{\#}]$ matrices as

$$G_{ij} = \left(G_{ij}^{\#} \cos \theta_{ij} - B_{ij}^{\#} \sin \theta_{ij}\right)\big|_{(i,j \in n)} \quad (17a)$$

$$B_{ij} = \left(G_{ij}^{\#} \sin \theta_{ij} + B_{ij}^{\#} \cos \theta_{ij}\right)\big|_{(i,j \in n)}. \quad (17b)$$

B. Obtain More Accurate Conductances and Susceptances

In order to obtain more accurate conductances and susceptances using the initial approximations obtained from (17), an UKF-based probabilistic method and NR-based deterministic method are iteratively invoked. This is explained as follows.

1) Obtaining the Initial Estimates in a Probabilistic Way: The probabilistic method gives the refined estimates of g, b, and θ, which serve as a convenient starting point for the deterministic approach in the succeeding step. The UKF-based probabilistic approach is employed here, using the measurements obtained from SMs as well as μPMUs. Initially, the state matrix and the measurement matrix are formulated as

$$\alpha^{(k)} = \begin{bmatrix} g & b & \theta^{(k)} \end{bmatrix}^{T} \quad (18)$$

$$\beta^{(k)} = \begin{bmatrix} P^{(k)} & Q^{(k)} & |V|^{(k)} & I_f^{(k)} & \theta_M^{(k)} \end{bmatrix} \quad (19)$$

where $\alpha^{(k)}$ and $\beta^{(k)}$ denote the state matrix and the measurement matrix, respectively, for the kth sample, with $k = 1, 2, \ldots, K$. The vectors "g" and "b" in the state matrix correspond to the nonzero elements "G_{ij}" and "B_{ij}" matrices, respectively, while "$\theta^{(k)}$" is the vector that contains the voltage angles of all the buses at the kth instant, and the rest of the vectors associated with the measurement matrix are given as follows:

$$P^{(k)} = \left[P_i^{(k)}\right]_{i \in [1,n]}; \quad Q^{(k)} = \left[Q_i^{(k)}\right]_{i \in [1,n]} \quad (20a)$$

$$|V|^{(k)} = \left[\left|V_i^{(k)}\right|\right]_{i \in [1,n]}; \quad \theta_M^{(k)} = \left[\theta_i^{(k)}\right]_{i \in [m,m+p]} \quad (20b)$$

$$I_f^{(k)} = \left[\left|I_{ij}^{(k)}\right|\right]_{i,j \in [m,m+p]}. \quad (20c)$$

In (20), as mentioned earlier, for the measurements of active and reactive power at unknown buses, the pseudo measurements are considered (with high error covariances). Similarly, for the voltage magnitude measurements, the measurements associated with the buses with μPMUs as well as SMs are considered in (19). Moreover, "I_f" corresponds to the magnitude of current flow in a line between buses "i" and "j," for which to be considered only if the buses to which the line is connected are known, as explained earlier. The measurement matrix is thus given as follows:

$$H\left(\alpha^{(k)}\right)$$
$$= \begin{bmatrix} P^{(k)} & Q^{(k)} & |V|^{(k)} & I_f^{(k)} & \theta_M^{(k)} \end{bmatrix}_{calc}$$

$$\equiv \begin{bmatrix} |V_i| \sum_{j=1}^{n} |V_j| \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}\right)\Big|_{i \in [1,n]} \\ |V_i| \sum_{j=1}^{n} |V_j| \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}\right)\Big|_{i \in [1,n]} \\ \left|V_i^{(k)}\right|\Big|_{i \in M} \\ \theta_i^{(k)}\Big|_{i \in M} \\ \sqrt{\left(G_{ij}^2 + B_{ij}^2\right)\left(|V_i|^2 + |V_j|^2 - 2|V_i||V_j|\cos \theta_{ij}\right)}\Big|_{i \in M} \end{bmatrix}.$$
$$(21)$$

The subscript "calc" in (21) indicates the calculated values of the measurements. The measurement matrix is assumed with finite noise vector, associated with finite Gaussian noise in each of the measurements, i.e.,

$$\beta^{(k)} = H\left(\alpha^{(k)}\right) + v^{(k)} \quad (22)$$

where $v^{(k)}$ is a white noise with covariance "R_k," which contains the covariances associated with the errors in SM, μPMU, and pseudo measurements. Given an initial state estimate $\alpha^{(k)}$ (with dimension "L"), having a covariance matrix given by $P_{\alpha}^{(k)}$, initially, sigma points are generated as [30]

$$A_0^{(k)} = \hat{\alpha}^{(k)}$$
$$A_i^{(k)} = \hat{\alpha}^{(k)} + \sqrt{(L + \lambda) P_{\alpha,i}^{(k)}}; \quad i = 1, 2, \ldots, L$$
$$A_i^{(k)} = \hat{\alpha}^{(k)} - \sqrt{(L + \lambda) P_{\alpha,i}^{(k)}}; \quad i = L + 1, \ldots, 2L \quad (23)$$

where "λ" $(=(\mu^2 - 1)L)$ is a scaling parameter and "μ" is a small positive value that determines the spread of the sigma points. These sigma vectors are then propagated through the nonlinear function (H) as

$$B_i^{(k)} = H\left(A_i^{(k)}\right); \quad i = 0, 1, 2, \ldots, 2L. \quad (24)$$

The vector $\alpha^{(k)}$ (which consists of the estimates of g, b, and θ) is then estimated for the $(k + 1)$th instant (which is used as the initial condition for deterministic approach in Section III-B.2) as

$$\hat{\alpha}^{(k+1)} = \sum_{i=0}^{2L} W_i^{(m)} A_i^{(k)} + G^{(k)}\left(\beta^{(k)} - \sum_{i=0}^{2L} W_i^{(m)} A_i^{(k)}\right). \quad (25)$$

In (25), "G" is the Kalman gain, which is evaluated as [30]

$$G^{(k)} = \left[P_{\alpha\beta}^{(k)}\right]\left[P_{\beta\beta}^{(k)}\right]^{-1} \quad (26a)$$

$$P_{\beta\beta}^{(k)} = \sum_{i=0}^{2L} W_i^{(c)}\left[B_i^{(k)} - \sum_{i=0}^{2L} W_i^{(m)} B_i^{(k)}\right]$$
$$\times \left[B_i^{(k)} - \sum_{i=0}^{2L} W_i^{(m)} B_i^{(k)}\right]^{T} + R_k \quad (26b)$$

$$P_{\alpha\beta}^{(k)} = \sum_{i=0}^{2L} W_i^{(c)}\left[A_i^{(k)} - \sum_{i=0}^{2L} W_i^{(m)} A_i^{(k)}\right]$$
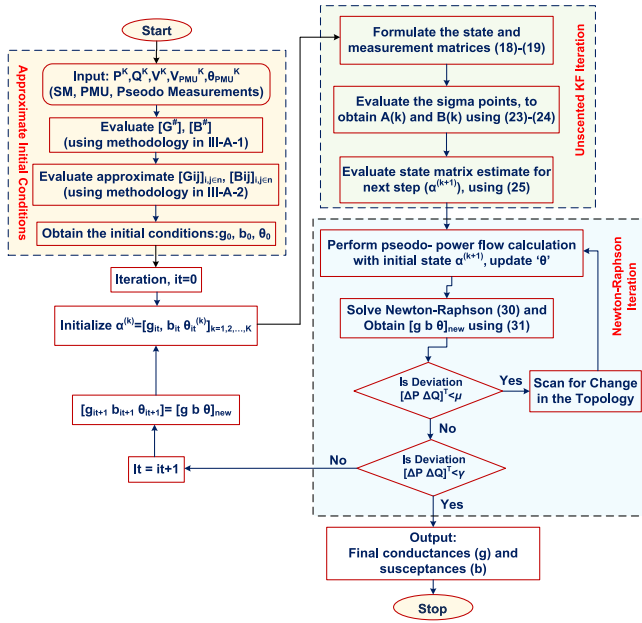$$\times \left[B_i^{(k)} - \sum_{i=0}^{2L} W_i^{(m)} B_i^{(k)}\right]^{T}. \quad (26c)$$

Fig. 1. Flowchart of the proposed algorithm.

"$W_i^{(m)}$" and "$W_i^{(c)}$" are weights, which are defined as

$$W_0^{(m)} = \lambda/L + \lambda; \; W_i^{(m)} = W_i^{(c)}\Big|_{i=1,\ldots,L} = 0.5/L + \lambda$$

$$W_0^{(c)} = \lambda/(L + \lambda) + \left(1 - \mu^2 + \gamma\right) \quad (27)$$

where γ is incorporated to account for the prior knowledge of the distribution of $\boldsymbol{\alpha}^{(k)}$ (here, γ is set to 2) [29]. The average of the estimated values of "g" and "b" for all the samples is considered as an initial start for the deterministic approach.

2) Obtain the Final Estimates Using in a Deterministic Way: An NR-based deterministic method is carried out to obtain the final accurate values of the conductances and susceptances. As a well-known fact, the NR method is prone to diverge (thereby causing large errors in the solution) in case the initial conditions are far apart from the true values. The approximate estimates as obtained from Section III-A might provide initial condition, which is far apart from the true values, due to the errors in the SM and pseudo measurements. Therefore, the estimates of g, b, and θ (with vector $\alpha^{(k)}$) evaluated in the previous step serve as a convenient initial condition for carrying forward the NR method. As reported in Fig. 1, the algorithm uses these two steps in an iterative way, i.e., step-1 is repeated before each step-2 iteration, while the estimates of step-2 are in-turn used in step-1. These iterations are carried out until the time NR converges. The convergence criterion is chosen as the sum of squares of the deviation between the measured and calculated values of active and reactive powers

is smaller than a threshold (γ). As per the NR theory [17], the following expression is formulated:

$$\begin{bmatrix} \Delta\mathbf{P}^K \\ \Delta\mathbf{Q}^K \end{bmatrix} = \begin{bmatrix} \dfrac{\partial\mathbf{P}}{\partial g} & \dfrac{\partial\mathbf{P}}{\partial b} & \dfrac{\partial\mathbf{P}}{\partial\theta} \\ \dfrac{\partial\mathbf{Q}}{\partial g} & \dfrac{\partial\mathbf{Q}}{\partial b} & \dfrac{\partial\mathbf{Q}}{\partial\theta} \end{bmatrix} \begin{bmatrix} \Delta g \\ \Delta b \\ \Delta\theta \end{bmatrix} \quad (28)$$

where (29), as shown at the bottom of the page.

The change in the active and reactive powers matrix on the left-hand side of (28) corresponds to the change between the measured values and the values calculated using the previous step estimates. The formulation (28) is then solved using

$$\begin{bmatrix} \Delta g & \Delta b & \Delta\theta \end{bmatrix}^T = \begin{bmatrix} \dfrac{\partial\mathbf{P}}{\partial g} & \dfrac{\partial\mathbf{P}}{\partial b} & \dfrac{\partial\mathbf{P}}{\partial\theta} \\ \dfrac{\partial\mathbf{Q}}{\partial g} & \dfrac{\partial\mathbf{Q}}{\partial b} & \dfrac{\partial\mathbf{Q}}{\partial\theta} \end{bmatrix}^\dagger \begin{bmatrix} \Delta\mathbf{P}^K \\ \Delta\mathbf{Q}^K \end{bmatrix}. \quad (30)$$

In (30), "$\dagger$" represents the pseudoinverse of the matrix. For the evaluation of the elements in the Jacobian matrix in (30), the pseudo inverse, and the threshold for convergence, the reader is directed to [17]. Update of the matrix with g, b, and θ is then performed at each iteration as

$$\begin{bmatrix} g & b & \theta \end{bmatrix}_{\text{new}}^T = \begin{bmatrix} g & b & \theta \end{bmatrix}^T + \begin{bmatrix} \Delta g & \Delta b & \Delta\theta \end{bmatrix}^T. \quad (31)$$

g, b, and θ thus obtained from (31) are fed as an initial condition for carrying out the Kalman filtering process in step-1. During the iteration, the branches with sufficiently small conductances and susceptances are considered as wrongly identified ones, and thus, a certain value (μ) is set as a topology modification threshold. Whenever a branch is less than this threshold, a scan for topology change is done [17] and the NR iteration is reconducted, while the precision of the voltage angle estimation is enhanced in each iteration using a pseudo power flow performed with the known states and measurement values. The complete methodology is highlighted in a flowchart in Fig. 1.

C. Evaluation of the Optimal Locations of the μPMUs

The optimal μPMU installation locations are evaluated based on the mean square errors in the estimated values of the conductances and susceptances. The performance metric considered is the averages mean square error (AMSE) averaged over the number of branches, which is defined as

$$\text{AMSE}(M) = \sqrt{\dfrac{1}{m}\sum_{x=1}^{m}\mathbb{E}\left[\left|\left|\begin{bmatrix} g_x \\ b_x \end{bmatrix}_{\text{true}} - \begin{bmatrix} g_x \\ b_x \end{bmatrix}_{\text{est}}\right|\right|^2\right]}. \quad (32)$$

In (32), the matrix $[g_x \; b_x]_{\text{est}}^T$ is the estimated branch conductance and susceptance by the proposed algorithm, and the matrix $[g_x \; b_x]_{\text{true}}^T$ depicts the true values of the conductances

$$\mathbf{P}^K = \begin{bmatrix} P_1^{(1)} & \cdots & P_1^{(K)} \\ \vdots & \ddots & \vdots \\ P_n^{(1)} & \cdots & P_n^{(K)} \end{bmatrix}_{n\times K} ; \quad \mathbf{Q}^K = \begin{bmatrix} Q_1^{(1)} & \cdots & Q_1^{(K)} \\ \vdots & \ddots & \vdots \\ Q_n^{(1)} & \cdots & Q_n^{(K)} \end{bmatrix}_{n\times K}$$

$$\theta = \begin{bmatrix} \theta_1^{(1)} & \cdots & \theta_1^{(K)}; \theta_2^{(1)} & \cdots & \cdots; \theta_n^{(1)} & \cdots & \theta_n^{(K)} \end{bmatrix}_{1\times nK} \quad (29)$$
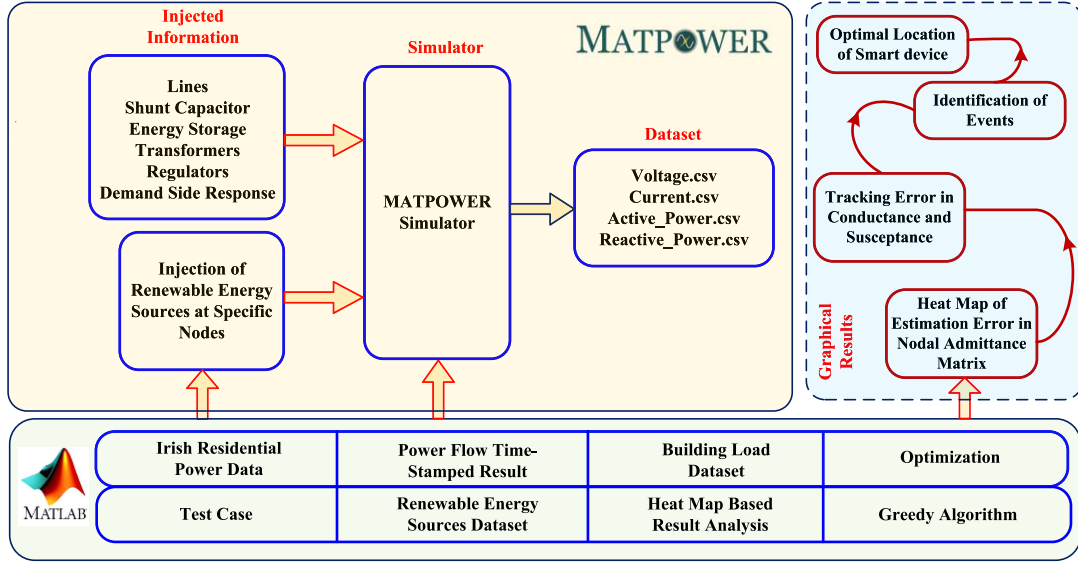
Fig. 2. Illustration of the interfacing between the MatPower package and the MATLAB coder function.

and susceptances, while "m" is the total number of branches in the network. This corresponds to the average prediction error. Using (32), the placement problem can be identified as

$$P^{op} = \text{argminAMSE}(P) \tag{33}$$

$$\text{AMSE}^{op}(P) = \text{AMSE}(P^{op}(P)) \tag{34}$$

where "P^{op}" is the optimal placement set of the μPMUs, with the number of μPMUs in a set fixed to the desired number of μPMUs. However, (33) is a combinatorial problem with its solution time-consuming when large grids are considered. For this reason, greedy μPMU placement procedure as reported in Algorithm 1 based on the sequential addition of one μPMU at a time could provide the fast performance.

Algorithm 1 Optimal Placement of μPMUs

Initialize: $P^{op}(0) := \varnothing, P^s(0) := \{1, 2, \ldots, P\}$
for $x = 1 : m$ **do**
$p^*(x) = min_{p \in P^s(x-1)} AMSE(P^{op}(x-1) \cup p)$
$P^{op}(x) = P^{op}(x-1) \cup p^*(x)$
$P^s(x) = P^s(x-1)p^*(x)$
$AMSE^{op}(x) = AMSE^{op}(P^{op}(x))$
end for

IV. RESULTS AND DISCUSSION

For validation purposes, the benchmarked IEEE 33-Bus and IEEE-123 bus distribution networks are considered and the MATLAB[1] software used for the simulation purpose while interfacing it with the MatPower toolbox package, as shown in Fig. 2. The input information for the MatPower is depicted as it processes measurement data, including voltage phasors, current phasors, and active and reactive powers. This measurement dataset is processed in the MATLAB coder to accomplish the identification objectives. The salient points are illustratively shown in Fig. 2. A schematic of the benchmarked IEEE 33-Bus distribution network is shown in Fig. 3, with the
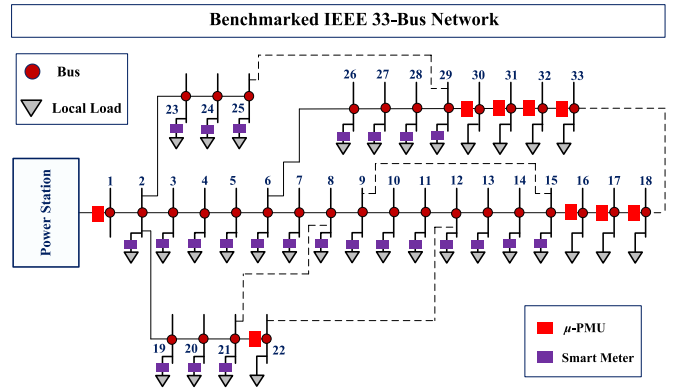
[1]Trademark.

Fig. 3. Schematic of the IEEE 33-bus distribution network.

SMs and μPMUs installed at different buses in the network, with the standard 12.66-kV and 100-MVA base values. The load data to assign loads at different buses of the network are generated using the hourly load data taken from the Ireland residents [31]. The data of a set of residents are grouped and assigned to each of the 32 buses (excluding the reference bus). The hourly data are linearly interpolated to have a data sampling time of every 6 min, which is equivalent to 240 samples for one day (i.e., "K" in (6) is 240). After the load data are assigned at different time instances of a day, the corresponding SM and μPMU measurements are recorded. The measurements from SM and μPMU at different buses, with various noise levels, are then used as an input to the identification algorithm. The results with various test cases are simulated and presented here, to highlight the performance of the proposed algorithm.

Initially, the performance of the algorithm was evaluated considering only SM measurements with the Gaussian distribution error of 1% in the power measurements (i.e., the deviations in SM power measurements scatter in $\pm 3\%$ three-sigma region) and 0.01% voltage magnitude error. Fig. 4 graphically reports the identification error of the algorithm. The expression of Fig. 4 is explained as follows.
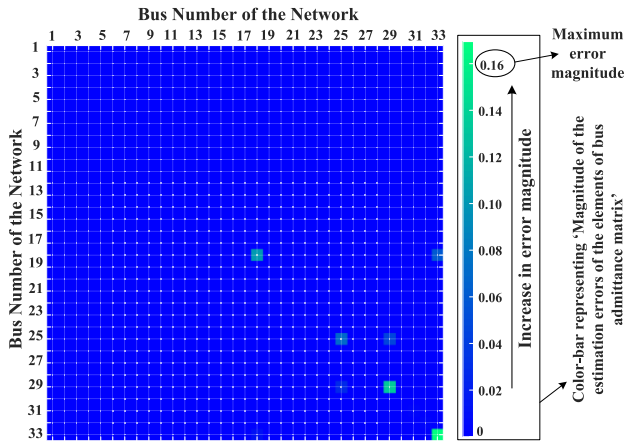
Fig. 4. Network identification error of the algorithm for IEEE 33 bus.
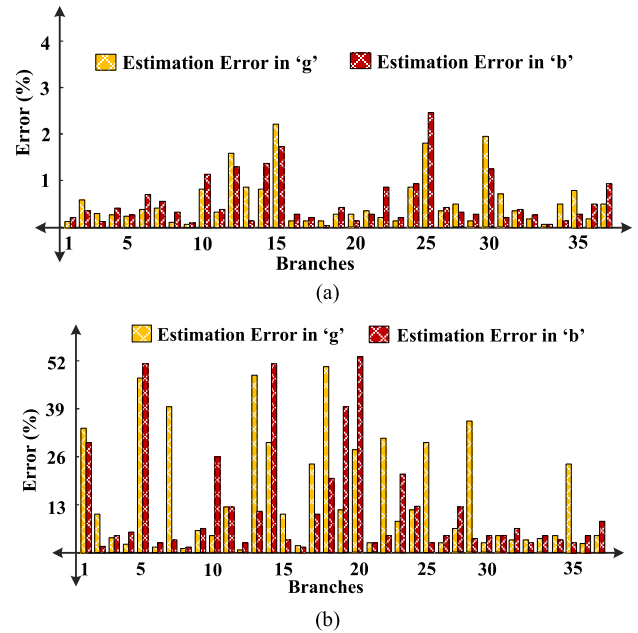

(a)

(b)

Fig. 5. Mismatch between the estimated and true conductances and susceptances in IEEE 33-bus network, with Gaussian distribution error in SM power measurements as (a) 1% and (b) 5%.

The bus admittance matrix for 33-bus system has a size of 33×33, thereby a total of 1089 elements. The estimation error (i.e., magnitude of the difference between the true admittance matrix and the admittance matrix estimated from the presented algorithm) of each element is quantified as per its error magnitude. These error magnitudes are represented as varying shades of the blue color, as shown in the colorbar in Fig. 4, where thicker shade of blue color represents less error, while the lighter shade represents more error. Accordingly, errors of different elements in the 33×33 admittance matrix are plotted in large square box with 33×33 mini-squares. The error magnitude is represented by the color from the colorbar. As it can be observed, most of the mini-square boxes have thicker shades of blue color, proving almost negligible error in the estimation of the admittance matrix elements. Even those of a few mini-squares, which have some lighter shades of the blue color, have a maximum per-unit identification error magnitude as less as 0.16. Therefore, it can be concluded that the estimation with presented identification algorithm is accurate. Consequently, the estimation error in conductances and susceptances of each of the 37 branches is observed low, as shown in Fig. 5(a). The mean of the percentage error (MPE) over all the branches in the network is as low as 2.2% for conductance and 1.8% for susceptance. However, with increase of the noise levels to 5% Gaussian distribution error in active and reactive power measurements, the MPE values increase significantly to 22.5% and 15.8% for conductance and susceptance, respectively, as shown in Fig. 5(b). It may be noted that the results correspond to the mean values obtained over 100 Monte Carlo simulation runs.

Moreover, to demonstrate the capability of the presented framework on large systems, the IEEE 123-bus system is used to demonstrate the efficacy of the presented framework. Fig. 6(a) shows the schematic of the benchmarked IEEE 123-bus test system. As explained earlier, the Irish load data [31] are used and assigned at different time instances of a day, and the corresponding SM measurements at various buses are obtained, while the Gaussian distribution error of 1% in the power measurements is then added to the SM measurements data. Fig. 6(b) shows the network identification error of the algorithm, where most of the mini-square boxes (123×123) have thicker shades of blue color

with negligible estimation error. Even those of a few mini-squares, which have some lighter shades, have a maximum per-unit identification error magnitude as less as 0.18. Thus, the framework exhibits satisfactory performance for the IEEE 123-bus benchmarked network, thereby depicting its scalability for the large test system.

A. Impact of Measurement Noise in SMs

The impact of measurement noise in SMs is further analyzed here. The impact of different noise levels in SM measurements is reported in Figs. 7 and 8. Fig. 7 shows how the estimation error of the branch susceptance varies for all the 37 branches, with increasing noise levels in the SM power measurements. The mean of the estimation errors in the line susceptances of all the branches, i.e., the MPE values, for different SM noise levels, is reported in Fig. 8(a) as red line, along with the MPE values of the conductance, is also reported as the blue line. To clearly understand the impact of these line parameter errors on the state estimation (usually conducted by the DNO), a least-square-based state estimation of the network is performed, using the method in [32], with the estimated values of "g" and "b," instead of the *true values*. A detailed description is reported in the Appendix, highlighting the principle of the state estimation method used with the estimated values of "*g*" and "*b*" and the hybrid measurements. The errors in the estimated voltage magnitudes and estimated angles are observed, with different noise levels in SM measurements. The mean of the estimated voltage magnitude error and the angle error of the 33 buses is plotted in Fig. 8(b). It is noted that the available SM measurements are provided as the measurement data input to the state estimation algorithm. From Fig. 8(b), it can be observed that the errors in the estimated states are increasing as the errors in the estimated conductances and susceptances increase. In a distribution network state

(a)

(b)

Fig. 8. Impact of noise in SM measurements. (a) MPE variation in "g" and "b" with various levels of noise in the SM measurements. (b) Impact of "g" and "b" errors on the network state estimation.
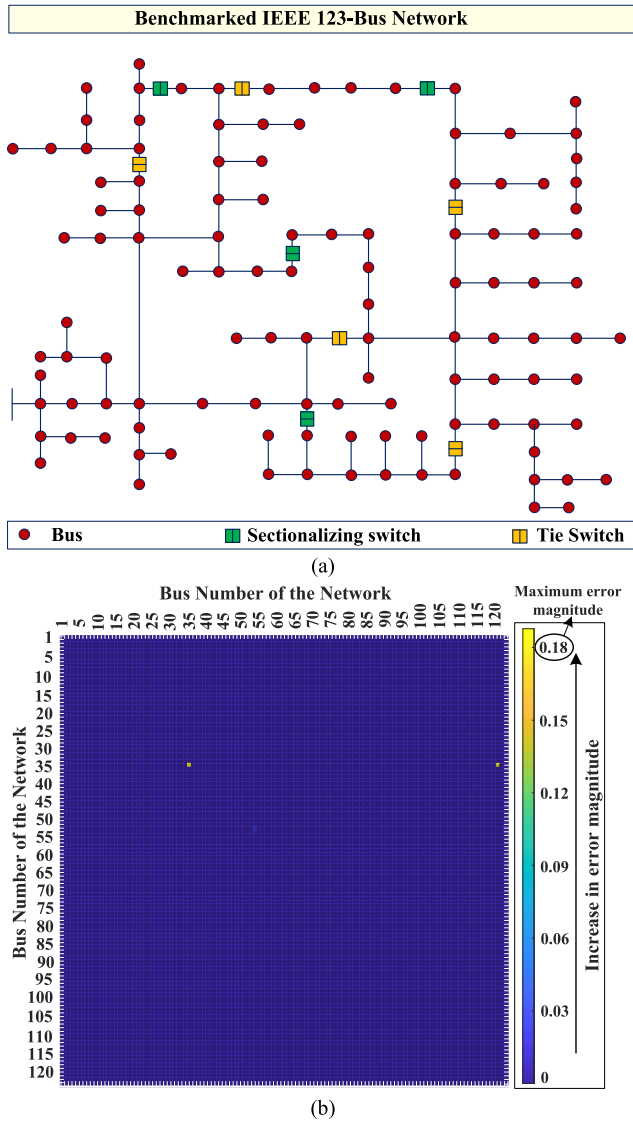

(b)

Fig. 6. (a) Schematic of the IEEE 123-bus distribution network. (b) Network identification error of the algorithm for IEEE 123-bus.
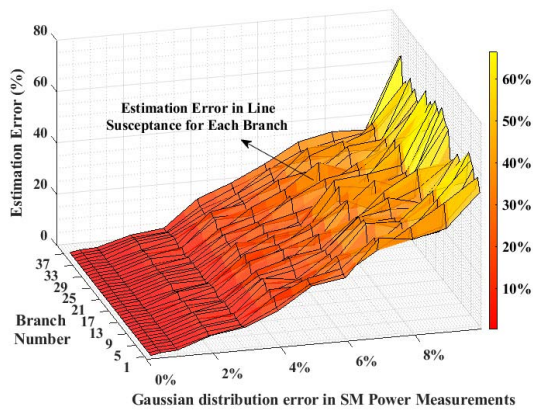

Fig. 7. Estimation error in each of the branches for various levels of noise in the SM power measurements.

estimation, the voltage magnitude error of up to 3% and the voltage angle error up to 0.03 radians are considered acceptable [33], and therefore, the Gaussian distribution error
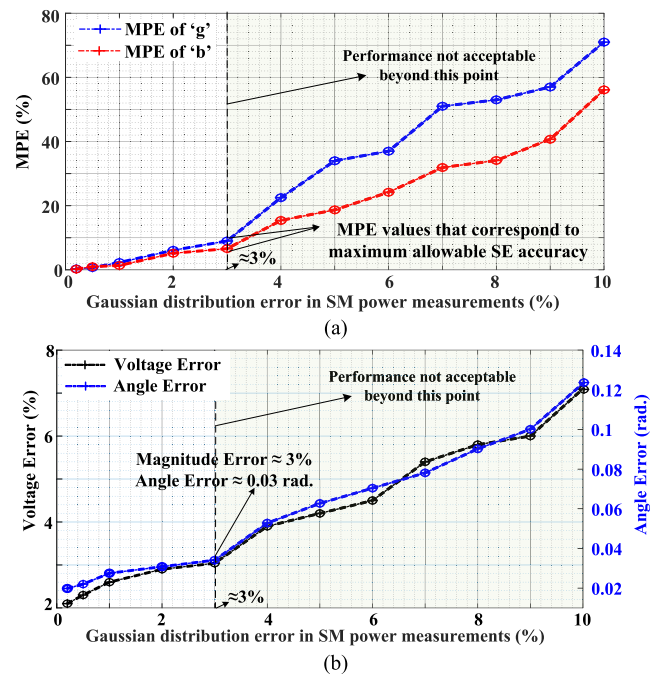
in SM power measurements (i.e., 3%) is marked in Fig. 8(b). The partitioned region is thus pointed out in Fig. 8(a) and (b), depicting that the measurement noise of up to 3% error in power measurements is acceptable. In case the noise levels go higher or there is missing SM data at certain buses, then the μPMUs need to be installed at different locations, as discussed next.

B. Impact of Measurement Noise in μPMUs Along With Noisy SM Measurements

Similar to the analysis performed with only SM measurement noise variation as in Fig. 8(a) and (b), an analysis is performed with the noise variation in the μPMU measurements coupled with noisy SM power measurements. It is considered that the μPMUs are installed at eight locations as highlighted in Fig. 3 (the chosen locations are the same as in [14]), while SMs are assumed absent at these locations. However, the pseudo measurements are considered at the locations where the SMs are absent, i.e., with high Gaussian error in the power measurements of 25%, while the SM power measurements at the rest of the buses are with 2% error. Fig. 9(a) and (b) shows the analysis with both μPMU and SM measurement noises. It can be seen in Fig. 9(a) and (b) that, with 2% Gaussian distribution error in the SM power measurements, as the Gaussian distribution error in the μPMU voltage and current phasor measurements increases, the algorithm's performance is affected, and the measurements within 0.18% error are observed to be acceptable values. Moreover, it is observed that the number of iterations (i.e., the value of "it" in Fig. 1) for convergence of the algorithm increases as the noise in the μPMU measurements increases, as shown in Fig. 10.
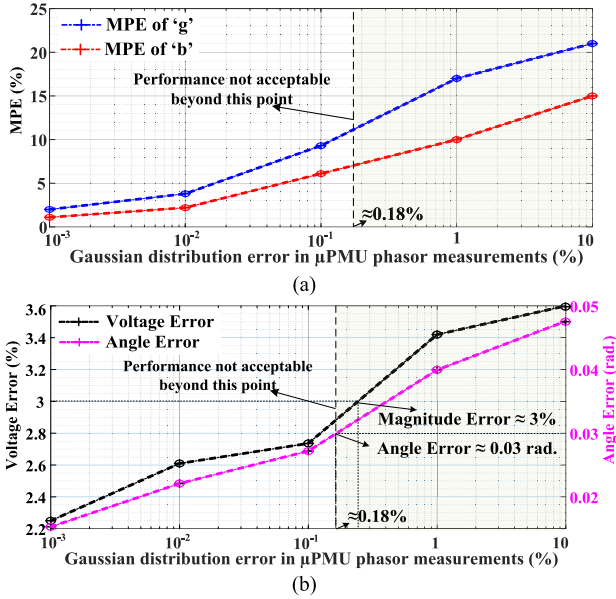
Fig. 9. Analysis of the noise in both μPMU and SM measurements. (a) MPE variation in "g" and "b" with variation in μPMU measurements noise. (b) Impact of "g" and "b" errors on the network state estimation.
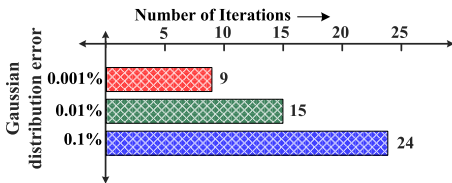

Fig. 10. Iterations for the convergence of the algorithm with different μPMU phasor measurement errors.

C. Impact of Renewable Power Injection on the Identification Results

The impact of the injection of renewable power at various buses in the network on the network identification algorithm is analyzed here. The renewable energy sources are coupled at certain buses (3, 8, 14, 25, 30, and 31) to study the impact of its stochastic nature on topology identification. The injected renewable power at these buses has been varied from zero to the nominal load power values at these buses. The graphical illustration of an identification error of admittance matrix is analyzed in Fig. 11. The expression of Fig. 11 is explained as follows.

The bus admittance matrix has a total of 33×33 elements for the 33-bus system. The estimation error (i.e., magnitude of the difference between *the true admittance matrix* and *the admittance matrix estimated from the presented algorithm*) with each of these elements is quantified as per its error magnitude. Thereby, it consists of 33×33 mini-*square boxes*. Before the connection of renewable energy sources, these error magnitudes are represented as varying shades of the green color, as depicted in the *left-side colorbar* in Fig. 11, where the thicker shade of green color represents less error, while the lighter shade represents more error. Most of the areas in the heat map are in green (i.e., related to left bar representation), which indicates that identification error is low under the nominal operating scenarios.

TABLE II
MPE OF CONDUCTANCES AND SUSCEPTANCES WITH DIFFERENT PMU LOCATIONS AND GAUSSIAN DISTRIBUTION ERRORS IN MEASUREMENT

No. of μPMUs	μPMU Locations	MPE of 'g' (%)		MPE of 'b' (%)	
		Error= 0.01%	Error= 0.1%	Error= 0.01%	Error= 0.1%
6	{16,17,18,22,31,33}	**4.09**	**10.91**	**4.16**	**9.09**
	{16,17,18,22,30,32}	7.98	21.97	7.36	16.57
	{16,18,22,30,31,33}	20.53	25.42	16.45	20.32
	{16,18,22,30,31,32}	22.51	27.63	20.14	24.06
7	{16,17,18,22,30,31,33}	**3.49**	**9.08**	**3.43**	**6.77**
	{16,17,18,22,30,31,32}	6.42	9.91	5.15	10.01
	{16,18,22,30,31,32,33}	10.47	14.68	9.24	12.94
	{16,17,22,30,31,32,33}	14.07	16.04	11.34	13.38

The *right colorbar*, however, represents a maximum *change in estimation error* upon the renewable power injection. The colors corresponding to this *right colorbar* are reflected as circles associated with each of 33×33 elements with varying radius. The radius of these circles varies with the color shades of *right colorbar*, and it represents the impact of renewable upon the particular element in the error matrix. For example, element (29, 29) has a red color circle with smaller radius, while the element at (2, 2) has a lighter shade circle with large radius. This implies that, upon the connection of renewables, the algorithm estimation accuracy is more prone to affect the parameters ("g" and "b") estimation of the lines connected to bus-2 more than bus-29. Thus, it can be observed from the right-hand bar since it has the maximum per-unit change of estimation error magnitude of only 0.06, the change in identification error is very low even upon renewable power injection at certain buses. As Kalman filter is an iterative algorithm, the presented algorithm provides excellent identification even under the presence of renewable power sources. Hence, the presented algorithm effectively accomplishes the topology identification objectives for the given network.

D. Impact of μPMUs Installation Locations

Fig. 12(a) shows the impact of the μPMU installation locations on the estimation error in "$\boldsymbol{g}$" and "$\boldsymbol{b}$." A total of six μPMUs are considered here and their different installation locations were selected out of the eight previously mentioned locations. The impact of errors in "$\boldsymbol{g}$" and "$\boldsymbol{b}$" with different μPMU locations is reported in Table II, along with their impact upon increasing the noise in μPMU phasor measurements (while keeping 2% Gaussian distribution error as a noise in SM power measurements). Table II shows the error variation, considering a total of six μPMUs as well as seven μPMUs. From the procedure highlighted in III-C, the optimal μPMU locations with six μPMUs are identified as {16, 17, 18, 22, 31, 33}, while the optimal μPMU locations with seven μPMUs are identified as {16, 17, 18, 22, 30, 31, 33}. The variation of estimation errors in "g" and "b" with this optimal configuration is shown in Fig. 12(a), and the corresponding impact on the state estimation algorithm
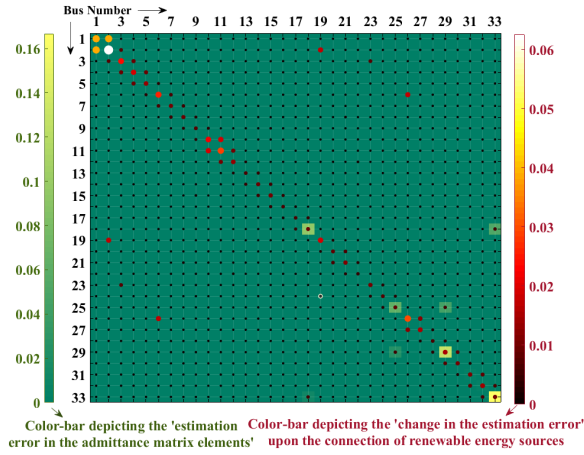
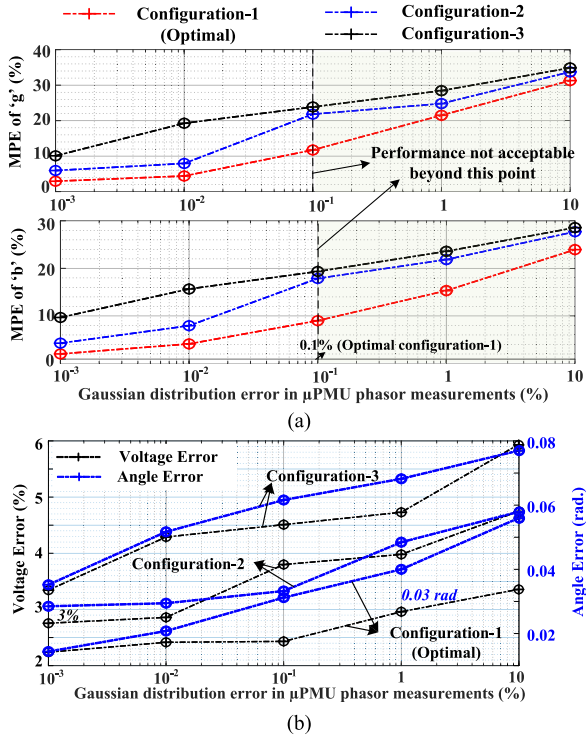Fig. 11. Performance of the presented method under renewable power penetration.

Fig. 12. Impact of different locations with six μPMUs. (a) MPE variation of "g" and "b." (b) Impact of "g" and "b" errors on the network state estimation.

is reported in Fig. 12(b). From Fig. 12(a), it is observed that as the Gaussian distribution error in the μPMU phasor measurements increases, the estimation error in "g" and "b" increases, while different configurations exhibit different MPE values. From Fig. 12(b), it is observed that the error of up to 0.1% is acceptable, in case six μPMUs installed at their optimal locations.

Another test case is considered, where the possible μPMU locations are not constrained to the set of eight locations highlighted in Fig. 3. The possible μPMU locations are chosen from all the 33 buses while assuming that one μPMU is present at the main substation reference bus (which is a practical consideration for active distribution networks).
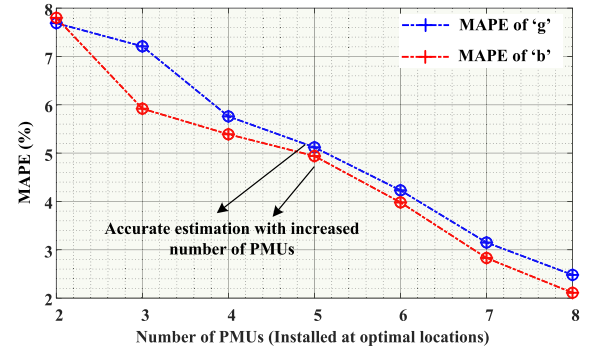

Fig. 13. Effect of a different number of μPMUs installed at their optimal locations.

TABLE III
NUMBER OF PMUS WITH CORRESPONDING OPTIMAL LOCATIONS

No. of PMUs	Optimal Locations
2	1, 21
3	1, 21, 8
4	1, 21, 8, 26
5	1, 21, 8, 26, 6
6	1, 21, 8, 26, 6, 10
7	1, 21, 8, 26, 6, 10, 11
8	1, 21, 8, 26, 6, 10, 11, 29

The optimal locations for a different number of μPMUs are then identified using Algorithm 1, and the corresponding locations are reported in Table III. Moreover, Fig. 13 shows the variation in MPE values of "g" and "b," with a different number of μPMUs, when they are installed at these optimal locations, with Gaussian distribution noise of μPMU phasor measurements and SM power measurements set to 0.01% and 2%, respectively. It is observed that as the number of μPMU locations increases, the MPE values in "g" and "b" reduce significantly.

E. Comparative Performance With Presence of Gaussian Noise in μPMU and SM Measurements

Fig. 14 shows the comparative performance to validate the effectiveness of the presented method over an existing method [16] based on a benchmarked IEEE 33-bus network. The Gaussian distribution error in μPMU measurements of 0.01% is considered in both approaches, while a total of six μPMUs (at their optimal locations) are considered for the presented approach. Fig. 14(a) and (b) shows the identification error between the estimated and the true admittance matrix using the existing method and the presented method, respectively. As shown in Fig. 14(b), noticeable reduction of the identification error between the estimated and the true admittance matrix is achieved using the presented approach. The individual "g" and "b" estimation errors are also reported in Fig. 14(c). Fig. 14(c) (top panel) shows that an estimation error of branch parameters using the existing method is obtained up to 4.5%, which may lead to erroneous results while performing the state estimation. However, the presented
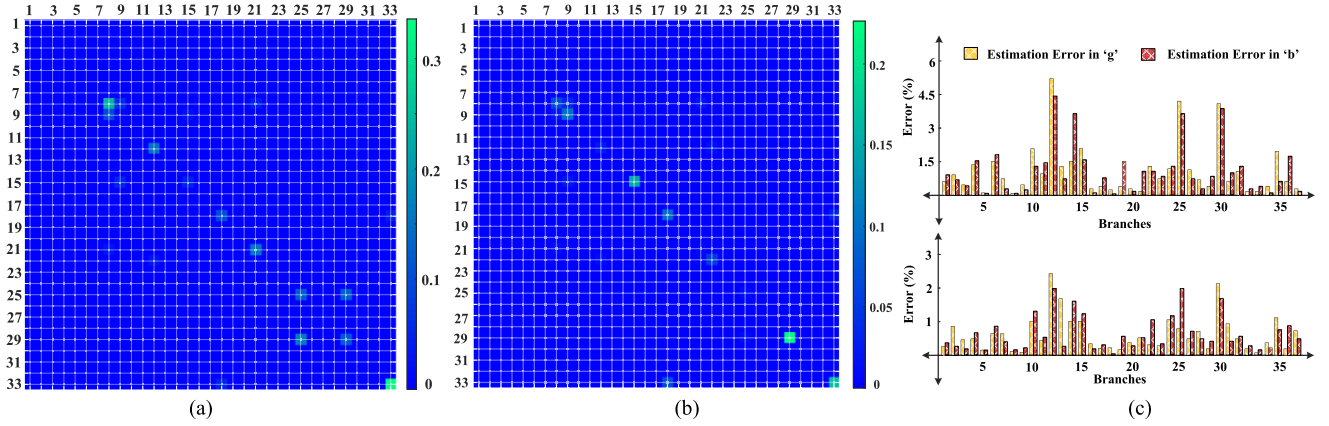
Fig. 14. Comparative performance under the presence of the Gaussian noise in μPMU measurements. Network identification with (a) state-of-the-art algorithm [16], (b) presented algorithm, and (c) estimation error in branch parameters using the state-of-the-art algorithm [16] (top) and the presented algorithm (bottom).
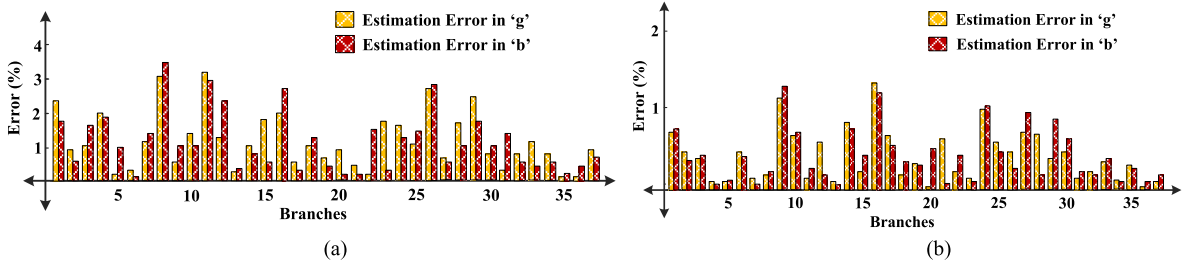

Fig. 15. Comparative performance in the presence of the Gaussian noise in SM measurements data. (a) Identification error in the estimated branch parameters with the state-of-the-art algorithm [17]. (b) Identification error in the estimated branch parameters with the presented algorithm.

TABLE IV
COMPARISON OF ATTRIBUTES WITH THE EXISTING WORK

	[8]	[13]	[16]	[22]	Presented
Algorithm	Graphical Approach	Sparse Recovery	Numerical	Graphical Approach	Numerical
Line parameter Estimation	No		Yes	No	Yes
Topology Identification	Yes				
Accuracy	Good		Better	Better	
Measurements	PMUs only			SMs only	SMs and PMUs
Measurement Noise Consideration	No	Yes			

algorithm limits an identification error within 2.5%, as shown in Fig. 14(c) (bottom panel). Moreover, Fig. 15(a) and (b) shows the response of the topology identification using the state-of-the-art framework [17] with noisy SM measurements data (the Gaussian distribution error of 2% in the SM power measurements). Since work [17] does not consider μPMUs, the data corresponding to the μPMU measurements are not utilized in this algorithm. However, the presented algorithm is able to improve the estimation using the μPMU measurements. The tracking performance of branch parameters with algorithm [17] is shown in Fig. 14(a). It is observed that the parameter estimation error is bounded within 4%, as shown in Fig. 15(a), while the improved performance

of network identification using the presented framework is shown in Fig. 15(b). The advantage of the hybrid method is that when the given network has data received from SMs and μPMUs, an improved update of the state variables from the UKF can be obtained, which means that the improved network identification can be obtained by leveraging data from various resources and scalable for large network. In addition, the proposed algorithm has better convergence as the initial conditions are made closer to the solution by UKF iteration. Thus, the tracking performance of branch parameters is shown in Fig. 15(b), where the error is accomplished within 2.5%. Finally, to highlight the contribution of the presented work, a comparison of different attributes with several key state-of-the-art studies on network identification is reported in Table IV.

V. CONCLUSION

Knowledge of the distribution system topology is crucial for most control applications and diagnostics, yet the information is often outdated in practice. This article explored a methodology for topology and parameter identification of distribution networks, which uses the measurements from SMs and μPMUs. The SM measurements include active power, reactive power, and voltage magnitude measurements at different buses, while the μPMU measurements give information of the voltage and current phasors associated with their installation locations. The modeling of these measurements with their errors is depicted. The network identification framework is reported in three main steps. First, the measurements from

SMs and μPMUs are utilized to obtain the approximate values of susceptances and the conductances of each line. Second, an UKF-based probabilistic approach is used, which uses the measurements at different time instants, to provide convenient initial conditions for the NR method for the third step. The third step solves for accurate "g" and "b," in a deterministic manner. Results for the benchmarked IEEE 33-bus and IEEE 123-bus networks are presented. The identification results with noisy measurements are presented, where the impact of measurement noises from both SMs and μPMUs are considered. Moreover, the impact of μPMU locations on the identification algorithm is analyzed and the algorithm is able to identify the optimal locations under different cases. A significant improvement in the topology estimation accuracy is observed when the μPMU measurements are utilized at the buses with missing/erroneous SM data. The improvement in estimation accuracy is observed to be proportional to the number of μPMUs installed, while it is also observed to be dependent on the installation locations. Solely with the Gaussian distribution error of 1% in SM power measurements in the IEEE 33-bus network, it is observed that the MPE in conductance is as low as 2.2% and that of susceptance is as low as 1.8%. With a state estimation analysis, it is observed that up to the error of 2% in the SM power measurements is acceptable for accurate network identification. Moreover, while considering limited SM measurements with the power measurements error of 2%, the acceptable limit of the Gaussian distribution error in the μPMU measurements from certain buses is observed to be 0.18%, for accurate network identification. It is also observed that the MPE in "g" and "b" decreased with the increase in μPMU locations, while the error is least with the optimal μPMU locations. Finally, the impact of the renewable power injection at certain buses is reported, while the comparative performance with some existing methods is also analyzed. The accurate identification results of the proposed approach under various test conditions show the viability and good performance of the approach for practical implementation purposes.

The future work can include the following research directions. Currently, the algorithm assumes that the power measurements received from the SMs from various locations have some measurement noise, and however, the effect of corrupt measurements (e.g., arising due to some phase misidentification) is not tackled. Even the μPMUs occasionally send corrupt data that are not suitable for identification. To tackle these issues, contemplations on appending some machine learning-based algorithms with the presented approach will be one of the future directions. Furthermore, the investigations to develop efficient algorithms for identifying the admittance matrix of radial distribution network with fewer measurement nodes and perform the sensitivity analysis of the identification results in nonstationary measurement errors.

APPENDIX

The least-square-based state estimation of the network is performed, with the *estimated values of "g" and "b,"* instead of the *true values*. The method is briefly highlighted as follows. The measurement vector consists of the measurements from SMs, as well as the measurements from μPMUs.

The vector of the SM measurements "$[z_1]$" can be formulated as nonlinear functions of the state vector as

$$[z_1] = h_1(x) + [\alpha_1] \tag{35}$$

while "α_1" denotes the measurement error vector with covariance ["W_1"]. Jacobian of h_1 is evaluated as

$$H_1(x) = \frac{\partial h_1}{\partial x}. \tag{36}$$

The weighted least-squares solution with given initial state vector ("x_k") and the measurements "$[z_1]$" is evaluated as

$$x_{k+1} = x_k + G_1(x_k)[H_1^T W_1^{-1}][z_1 - h_1(x_k)] \tag{37}$$

where $G_1(x_k) = [H_1^T(x_k)W_1^{-1}H_1(x_k)]^{-1}$. The appended measurement vector is obtained by appending the μPMU phasor measurements vector ("$[z_2]$"), which is given by

$$z = \begin{bmatrix} z_1 & z_2 \end{bmatrix}^T = \begin{bmatrix} h_1(x) & h_2(x) \end{bmatrix}^T + \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix}^T. \tag{38}$$

In (35) and (38), "h_1" and "h_2" are nonlinear functions of the network states dependent on the estimated "g" and "b" parameters with the relation as highlighted in [32]. "α_2" denotes the measurement error vector with covariance ["W_2"] corresponding to the μPMU measurements. The corresponding Jacobian matrix is evaluated as

$$H_2(x) = \frac{\partial h_2}{\partial x}. \tag{39}$$

Finally, the weighted least-squares solution with given initial state vector ($[x_k]$) and the measurements vector "$[z]$," is obtained as

$$\begin{aligned} x_{k+1} = x_k &+ G(x_k)[H_1^T W_1^{-1}][z_1 - h_1(x_k)] \\ &+ G(x_k)[H_2^T W_2^{-1}][z_2 - h_2(x_k)] \end{aligned} \tag{40}$$

where

$$G(x_k) = \left[H_1^T(x_k)W_1^{-1}H_1(x_k) + H_2^T(x_k)W_2^{-1}H_2(x_k)\right]^{-1}. \tag{41}$$

Thus, using the weighted least-squares solution, the errors in the estimated state vector (i.e., estimated voltage magnitudes and estimated angles at each bus) are observed. Since the solution is dependent on the network parameters, the estimation errors in voltage magnitudes and angles are reported with the estimated "g" and "b" parameters obtained from the proposed network identification algorithm.

REFERENCES

[1] P. A. Pegoraro, A. Meloni, L. Atzori, S. Sulis, and P. Castello, "PMU-based distribution system state estimation with adaptive accuracy exploiting local decision metrics and IoT paradigm," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 4, pp. 704–714, Apr. 2017.

[2] 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 Trans. Instrum. Meas.*, vol. 67, no. 10, pp. 2453–2462, Oct. 2018.

[3] E. W. S. Angelos, A. Abur, E. N. Asada, and O. R. Saavedra, "External topology tracking based on state estimation," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2623–2631, May 2021, doi: 10.1109/TPWRS.2020.3036038.

[4] D. Ritzmann, P. S. Wright, W. Holderbaum, and B. Potter, "A method for accurate transmission line impedance parameter estimation," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 10, pp. 2204–2213, Oct. 2016.

[5] W. Jiang and H. Tang, "Distribution line parameter estimation considering dynamic operating states with a probabilistic graphical model," *Int. J. Electr. Power Energy Syst.*, vol. 121, Oct. 2020, Art. no. 106133.

[6] X. He, R. C. Qiu, Q. Ai, and T. Zhu, "A hybrid framework for topology identification of distribution grid with renewables integration," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1493–1503, Mar. 2021.

[7] J. C. Mayo-Maldonado et al., "Data-driven framework to model identification, event detection, and topology change location using D-PMUs," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6921–6933, Sep. 2020.

[8] D. Deka, S. Talukdar, M. Chertkov, and M. V. Salapaka, "Graphical models in meshed distribution grids: Topology estimation, change detection & limitations," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4299–4310, Sep. 2020.

[9] R. K. Gupta, F. Sossan, J.-Y. L. Boudec, and M. Paolone, "Compound admittance matrix estimation of three-phase untransposed power distribution grids using synchrophasor measurements," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.

[10] A. Wehenkel, A. Mukhopadhyay, J.-Y.-L. Boudec, and M. Paolone, "Parameter estimation of three-phase untransposed short transmission lines from synchrophasor measurements," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6143–6154, Sep. 2020.

[11] J. Yu, Y. Weng, and R. Rajagopal, "Data-driven joint topology and line parameter estimation for renewable integration," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2017, pp. 1–5.

[12] M. Jafarian, A. Soroudi, and A. Keane, "Resilient identification of distribution network topology," *IEEE Trans. Power Del.*, vol. 36, no. 4, pp. 2332–2342, Aug. 2021.

[13] M. Babakmehr, F. Harirchi, A. Al-Durra, S. M. Muyeen, and M. G. Simoes, "Compressive system identification for multiple line outage detection in smart grids," *IEEE Trans. Ind. Appl.*, vol. 55, no. 5, pp. 4462–4473, Sep. 2019.

[14] G. S. Dua, B. Tyagi, and V. Kumar, "A novel approach for configuration identification of distribution network utilizing μPMU data," *IEEE Trans. Ind. Appl.*, vol. 57, no. 1, pp. 857–868, Feb. 2021.

[15] O. Ardakanian, Y. Yuan, R. Dobbe, A. von Meier, S. Low, and C. Tomlin, "Event detection and localization in distribution grids with phasor measurement units," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2017, pp. 1–5.

[16] O. Ardakanian et al., "On identification of distribution grids," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 3, pp. 950–960, Jan. 2019.

[17] J. Zhang, Y. Wang, Y. Weng, and N. Zhang, "Topology identification and line parameter estimation for non-PMU distribution network: A numerical method," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4440–4453, Sep. 2020.

[18] D. Liang, L. Zeng, H.-D. Chiang, and S. Wang, "Power flow matching-based topology identification of medium-voltage distribution networks via AMI measurements," *Int. J. Electr. Power Energy Syst.*, vol. 130, Sep. 2021, Art. no. 106938.

[19] Y. Liao, Y. Weng, G. Liu, Z. Zhao, C. Tan, and R. Rajagopal, "Unbalanced multi-phase distribution grid topology estimation and bus phase identification," *IET Smart Grid*, vol. 2, no. 4, pp. 557–570, Dec. 2019.

[20] L. Zhao et al., "Robust PCA-deep belief network surrogate model for distribution system topology identification with DERs," *Int. J. Electr. Power Energy Syst.*, vol. 125, Feb. 2021, Art. no. 106441.

[21] Z. Zhang, Y. Li, J. Zhang, J. Duan, and Y. Cao, "A topology identification and impedance estimation method for distribution network with distributed generations," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 13155–13160, 2020.

[22] S. J. Pappu, N. Bhatt, R. Pasumarthy, and A. Rajeswaran, "Identifying topology of low voltage distribution networks based on smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5113–5122, Sep. 2018.

[23] Y. Liao, Y. Weng, G. Liu, and R. Rajagopal, "Urban MV and LV distribution grid topology estimation via group lasso," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 12–27, Jan. 2019.

[24] (Jun. 2015). *Electricity Meters Disputed Meter Accuracy Report, National Measurement and Regulation Office, U.K.* [Online]. Available: https://assets.publishing.service.gov.U.K./government/uploads/ system/uploads/attachment_data/file/454144/Electricity_ Determination_Report_2003-2014.pdf

[25] P. Shah and X. Zhao, "Network identification using micro-PMU and smart meter measurements," *IEEE Trans. Ind. Informat.*, early access, Mar. 7, 2022, doi: 10.1109/TII.2022.3156652.

[26] L. Ma and L. Wu, "A general topology identification framework for distribution systems using smart meter and μ-PMU measurements," *Int. J. Elect. Power Energy Syst.*, vol. 139, Jul. 2022, Art. no. 108019.

[27] Q. Jiang, X. Li, B. Wang, and H. Wang, "PMU-based fault location using voltage measurements in large transmission networks," *IEEE Trans. Power Del.*, vol. 27, no. 3, pp. 1644–1652, Jul. 2012.

[28] A. A. Hai et al., "Transfer learning for event detection from PMU measurements with scarce labels," *IEEE Access*, vol. 9, pp. 127420–127432, 2021.

[29] L. Schenato, G. Barchi, D. Macii, R. Arghandeh, K. Poolla, and A. Von Meier, "Bayesian linear state estimation using smart meters and PMUs measurements in distribution grids," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2014, pp. 572–577.

[30] S. J. Julier, "The scaled unscented transformation," in *Proc. Amer. Control Conf.*, May 2002, pp. 4555–4559.

[31] *Commission for Energy Regulation, CER Smart Metering Project Electricity Customer Behaviour Trial*, Irish Social Science Data Archive, Dublin, Ireland, 2012, vol. 1.

[32] M. Zhou, V. Centeno, J. Thorp, and A. Phadke, "An alternative for including phasor measurements in state estimators," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1930–1937, Nov. 2006.

[33] A. Al-Wakeel, J. Wu, and N. Jenkins, "State estimation of medium voltage distribution networks using smart meter measurements," *Appl. Energy*, vol. 184, pp. 207–218, Dec. 2016.
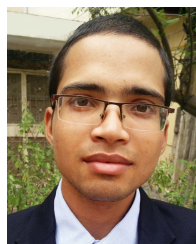
Vedantham Lakshmi Srinivas (Member, IEEE) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology (BHU) Varanasi, Varanasi, India, in 2015, after which he joined the Indian Institute of Technology Delhi, New Delhi, India, for M.Tech. degree in power systems and power electronics.

He joined Cardiff University, Cardiff, U.K., as a Post-Doctoral Research Associate, in March 2020, where he had been with the Centre for Integrated Renewable Energy Generation and Supply, School of Engineering, until November 2021. He is currently an Assistant Professor with IIT (ISM) Dhanbad, Dhanbad, India. His research interests include virtual power plants, electric vehicle drive and battery management systems, ac/dc microgrids, machine learning and its applications in power system state/parameter estimation, and stability and control.

Dr. Srinivas was a recipient of the POSOCO Power System National Award 2021 and the Research Excellence Travel Award and the International Research Grant from the Government of India.

Jianzhong Wu (Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Tianjin University, Tianjin, China, in 1999, 2002, and 2004, respectively.

From 2006 to 2008, he was a Research Fellow with The University of Manchester, Manchester, U.K. He joined Cardiff University, Cardiff, U.K., in June 2008, as a Lecturer, where he became a Senior Lecturer in 2013 and a Reader in 2014. He is currently a Professor of multivector energy systems and the Head of the School of Engineering, Cardiff University. He has coauthored books "*Smart Grid: Technology and Applications*" (2012, Wiley), "*Smart Electricity Distribution Networks*" (2017, CRC), and "*The Future of Gas Networks*" (2019, Springer). His research interests include integrated multienergy infrastructure and smart grid.

Prof. Wu is a member-at-large of the IEEE Technical Committee on Carbon Neutrality and the Coeditor-in-Chief of *Applied Energy*. He was the Deputy Editor-in-Chief of IET Energy Systems Integration and the Director of International UNiLAB on Synergies between Energy Networks. He is the Co-Director of the U.K. Energy Research Centre, an Associate-Director of EPSRC Supergen Energy Networks Hub, the Co-Chair of the INCOSE U.K. Energy Systems Interest Group, while he is a member of the BEIS U.K. Taxonomy Energy Working Group, the Wales Smart Energy System Group, the Northern Power Grid technical panel of ED2 business plan, and the Scottish Power Energy Networks Strategic Stakeholder Panel for England and Wales. He is also the President of the U.K. Branch of China Electrotechnical Society. He has contributed to more than 60 EC, EPSRC, and industry-funded projects as a Principal Investigator or a Co-Investigator.