Maximum Likelihood Estimation for Line Parameters in Distribution Grids Based on Expectation Maximization Algorithm

Shubhankar Kapoor, Adrian G. Wills, Johannes Hendriks, and Lachlan Blackhall

***Abstract*—This paper proposes a method for obtaining nonlinear models of distribution grid based on available measurements from the power grid. We formulate a maximum likelihood estimation (MLE) problem that estimates unknown line parameters—specifically, the impedance between nodes—using measured voltage magnitudes and net active and reactive power injections at each node. The nonlinear model for the distribution grid uses a nonlinear approximation of the DistFlow model, which includes line losses and is parameterized by the unknown line impedances. We solve the resulting MLE problem using an expectation maximization (EM) algorithm, tailored for the nonlinear model, and provide a numerically robust implementation. The proposed method is demonstrated on the IEEE 37-node test network, and we compare it with the state-of-the-art methods. The proposed method achieves a 70% reduction in voltage error and an error for state variables that is more than 10000 times smaller. A final comparison uses data from a real network, and the proposed method achieves parameter estimates with errors 100 times smaller than competing methods.**

***Index Terms*—Maximum likelihood estimation (MLE), distribution grid, parameter estimation, expectation maximization (EM).**

I. Introduction

A rapid transition to a carbon-neutral economy is required to limit global warming to 2 ℃ [1], [2]. This will require shifting from fossil fuel-based electricity generation to renewables-based ones [3]. With the shift towards renewable energy, the uptake of solar, batteries, controlled loads, and electric vehicles, also referred to as distributed energy resources (DERs), has been increasing in Australia [4]. Integrating DERs plays a pivotal role in the advancement of smart grids. However, the increase in DERs poses new challenges for the distribution grid (DG). For example, larger negative demand during sunny days results in negative power flows, which are not accommodated in the DG design [5]. This issue is known as the duck curve [6]. In 2023, the National Electricity Market achieved a record low demand, and the rooftop and grid-scale solar power contributed 57% of the total electricity supply [7]. This event marked one of the largest negative demand occurrences in the power grid. The scenarios characterized by the duck curve can result in instability and voltage issues on the DG [6]. To manage these issues, we need improved control capabilities for the DG that effectively utilizes DERs [8] for advanced distribution network management and operation strategies.

To improve the control of DERs, it is often required that the current state of DG is known [9], [10]. However, in many cases, the state information is not readily available and needs to be estimated [9], [11], [12], which is known as state estimation. Importantly, the majority of state estimation studies assume that the line parameters are known [13], [14]. However, these parameters are typically either unknown or inaccurate for a given DG [15], [16]. Therefore, it is crucial to estimate the line parameters for the success of state estimation approaches [10], [17] and maintaining grid stability [18]. This paper proposes a method independent of phasor measurement unit (PMU) data, aiming to fill a gap where such devices are not present.

There are several important studies that address the estimation of line parameters [10], [16], [19]-[22]. Some studies tackle the estimation of line parameters by assuming voltage phase angles are available [10], [16], [21]. However, the voltage phase angles are obtained using PMUs or micro-PMUs (μPMUs) that are expensive and cannot be found widely in the DG [20].

In [22], the active power p, reactive power q, and voltage magnitude V are used as measurements. However, this way still necessitates the derivation of voltage angle using a pseudo power flow analysis to accurately estimate the line parameters. By contrast, [20] does not use voltage angles but considers noise-free measurement data and estimates only the line lengths. Additionally, the line current measurements are required, which may be expensive to obtain [22]. In [18], the noisy data from smart meters are used to estimate line lengths and impedances. Reference [23] uses a linear power flow model that ignores line losses for estimating system

Manuscript received: August 17, 2024; revised: February 13, 2025; accepted: May 30, 2025. Date of CrossCheck: May 30, 2025. Date of online publication: June 27, 2025.

S. Kapoor (corresponding author) is with the Battery Storage Grid Integration Program at the Australian National University, Canberra, Australia (e-mail: Shubhankar.Kapoor@anu.edu.au).

A. G. Wills is with the School of Engineering, The University of Newcastle, Callaghan, Australia (e-mail: Adrian.Wills@newcastle.edu.au).

J. Hendriks is with Evergen, Newcastle, Australia (e-mail: johannes. n. hendriks@gmail.com).

L. Blackhall is with the Australian National University, Canberra, Australia (e-mail: Lachlan.Blackhall@anu.edu.au).

DOI: 10.35833/MPCE.2024.000910

state variables and line parameters. However, ignoring the line losses compromises the accuracy.

In this paper, we present a method by formulating a maximum likelihood estimation (MLE) problem to estimate the line parameters, which has the following distinct differences from previous works.

1) We propose a nonlinear approximation of the DistFlow model [1] that considers line losses, and we assume the availability of noisy measurements of active power p, reactive power q, and voltage magnitude V across the network [24], [25].

2) To solve the MLE problem, we employ an expectation maximization (EM) algorithm [26]-[29] and provide a numerically robust implementation of the resulting EM algorithm. To use this EM algorithm, we must adapt it to the nonlinear model, which is non-trivial and thus a major contribution of this paper. It is important to highlight that this is a significant departure from applying EM algorithm to the linear approximation of the DistFlow model used in [23], which cannot be applied to the nonlinear model adopted in this paper.

3) The line parameters can be accurately estimated with smart meter data of less than half a day without the need for expensive devices like PMUs or μPMUs and devices used to measure line currents/power. This makes the proposed method highly suitable and realistic for real-world applications. The impact lies in the significantly improved accuracy and robustness of the proposed method, which is validated on the IEEE 37-node test network [30] with noisy measurements using both simulated and real-world data. We consider noisy measurements at all nodes for testing, and this assumption is utilised in similar studies in [10], [16], [18], and [20]-[22]. The proposed method achieves a 70% reduction in voltage error and an error for state variables that is more than 10000 times lower. Furthermore, by comparing the results with [18], the robustness of the proposed method is demonstrated.

In summary, this paper presents the following significant contributions.

1) We propose a modified DistFlow model, i.e., nonlinear approximation of the DistFlow model, enhancing the accuracy by accounting for line losses.

2) We demonstrate that the proposed method relies only on commonly available measurements, such as voltage magnitude and power, without the need for phase angle.

3) We provide an EM algorithm tailored to the nonlinear model and detail a numerically stable implementation.

4) We demonstrate the performance of the proposed method using both simulated and real-world data.

In Section II, we introduce the power flow model and define the MLE problem. In Section III, we explain the use of the EM algorithm to solve the MLE problem, including computational implementation details. In Sections IV and V, we demonstrate the performance of the proposed method based on the simulated and real-world data, respectively. Section VI summarizes our findings and discusses potential future research.

II. Power Flow Model and MLE Problem Statement

A. Power Flow Model

The DG is represented as a tree, which is a valid assumption as most DGs are radial. $\mathcal{G}=(\mathcal{N}_0,\mathcal{E})$ is a special type of directed graph [13], where $\mathcal{N}_0$ is the set of nodes in the tree; and $\mathcal{E}$ is the set of edges in the tree. The nodes represent points of power injection/consumption to/from the grid or points of joining different branches, and the edges represent distribution lines between the nodes. Mathematically, the nodes are collected as integers in the set $\mathcal{N}_0=\{0,1,\ldots,i,j,k,\ldots,N\}$ with cardinality $|\mathcal{N}_0|=N+1$. The set $\mathcal{N}=\mathcal{N}_0\backslash\{0\}$ is also defined. The edges are integer pairs (i,j) that define a physical line connection between nodes i and j. Since the network is a tree, $|\mathcal{E}|=N$, and $\mathcal{G}=(\mathcal{N}_0,\mathcal{E})$ is acyclic and connected [13]. Node 0 is the slack node, which is connected to the upstream network defined as the ∞ node (also called the head node of the tree), as shown in Fig. 1. For node i, U_i is the square of voltage magnitude; and p_i and q_i are the net injected active and reactive power, respectively. For line $(i,j)\in\mathcal{E}$, P_{ij} and Q_{ij} are the active and reactive power flows, respectively; l_{ij} is the square of current magnitude flowing from node i to node j; and $z_{ij}=r_{ij}+\mathrm{j}x_{ij}$ is the impedance, where r_{ij} and x_{ij} are the resistance and reactance, respectively.

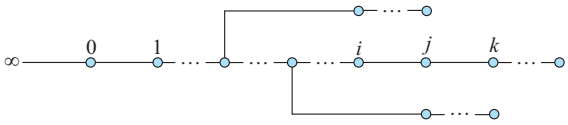

Fig. 1. Model of distribution network as a graph.

B. MLE Problem Statement

In this paper, we focus on estimating the impedance z_{ij} for all $(i,j)\in\mathcal{E}$ based on the following T measurements: ① U_i, $\forall i\in\mathcal{N}_0$; ② p_i and q_i, $\forall i\in\mathcal{N}_0$; and ③ P_{01} and Q_{01} of line (0, 1). Note that a measurement device obtains P_{01} and Q_{01} at the slack node, which are readily available in the DG rather than through a measurement device on the line.

To solve this line estimation problem, we require a relationship between the line impedances and measurements. We accomplish this using a mathematical model, i.e., a modified version of the Distflow model in [31] and [32], where the latter provides a relationship between impedance z_{ij} and measurements U_i, p_i, q_i, P_{ij}, Q_{ij} according to the nonlinear equations:

$$j\in\mathcal{N} \tag{1}$$

$$i=Pa(j) \tag{2}$$

$$U_j=U_i-2(r_{ij}P_{ij}+x_{ij}Q_{ij})+(r_{ij}^2+x_{ij}^2)l_{ij} \tag{3}$$

$$P_{ij}=p_j+r_{ij}l_{ij}+\sum_{k:(j,k)\in\mathcal{E}}P_{jk} \tag{4}$$

$$Q_{ij}=q_j+x_{ij}l_{ij}+\sum_{k:(j,k)\in\mathcal{E}}Q_{jk} \tag{5}$$

$$l_{ij} = \frac{P_{ij}^2 + Q_{ij}^2}{U_i} \tag{6}$$

where $i = Pa(j)$ represents that node i is the parent node of node j.

A major challenge in using this Distflow model is that U_j, P_{ij}, and Q_{ij} are implicitly defined and solving these equations often requires an iterative method, such as Newton's method. In this paper, we instead employ an approximation to (6) that yields explicit equations for U_j, P_{ij}, and Q_{ij} rather than implicit ones. The primary benefit is that the resulting model can be used directly for the purposes of estimation and does not require computationally intensive iterative solvers. In particular, the current l_{ij} from (6) is replaced with the following approximation:

$$l_{ij} \approx \hat{l}_{ij} = \frac{\left(\sum_{k \in N_{ij}^D} p_k\right)^2 + \left(\sum_{k \in N_{ij}^D} q_k\right)^2}{U_0} \tag{7}$$

where N_{ij}^D is the union of node j and nodes descendent from node j; and $\hat{l}_{ij}$ is the approximated value of l_{ij}. By definition, a node k is a descendent node of node j if there is a path [33] from node j to node k and the depth of node k is greater than that of node j with respect to the root node 0. This approximation is justified as follows.

1) The voltage U_i in (6) is approximated as $U_i \approx U_{i-1} \approx \ldots \approx U_0$, which is justified since the line losses are typically 1%-5% [34]. The acceptable nominal AC voltage tolerance in Australia is −6% to 10% of the nominal voltage. If there is a line loss of 5% with a maximum voltage error of 10%, this accounts to line loss error of 0.5%. With this small error, we get a tractable form for the line loss term rather than ignoring it.

2) The active and reactive power flows P_{ij} and Q_{ij} in (6) are approximated as:

$$\begin{cases} P_{ij} \approx \sum_{k \in N_{ij}^D} p_k \\ Q_{ij} \approx \sum_{k \in N_{ij}^D} q_k \end{cases} \tag{8}$$

This is based on the assumption that the line losses are small [35] and can be approximated as the sum of loads downstream of the line. Note that this assumption is only made for P_{ij} and Q_{ij} in the loss term in (6).

Replacing (6) with the approximation (7) results in a system of explicit equations, referred to as modified Distflow model:

$$U_j = U_i - 2(r_{ij}P_{ij} + x_{ij}Q_{ij}) + (r_{ij}^2 + x_{ij}^2)\frac{\left(\sum_{k \in N_{ij}^D} p_k\right)^2 + \left(\sum_{k \in N_{ij}^D} q_k\right)^2}{U_0} \tag{9}$$

$$P_{ij} = p_j + r_{ij}\frac{\left(\sum_{k \in N_{ij}^D} p_k\right)^2 + \left(\sum_{k \in N_{ij}^D} q_k\right)^2}{U_0} + \sum_{k:(j,k) \in \mathcal{E}} P_{jk} \tag{10}$$

$$Q_{ij} = q_j + x_{ij}\frac{\left(\sum_{k \in N_{ij}^D} p_k\right)^2 + \left(\sum_{k \in N_{ij}^D} q_k\right)^2}{U_0} + \sum_{k:(j,k) \in \mathcal{E}} Q_{jk} \tag{11}$$

Using a straightforward (but tedious) application of algebra, we can express these equations in the following compact form using functions $\boldsymbol{g_u}(\boldsymbol{\vartheta}, \boldsymbol{z})$, $\boldsymbol{g_P}(\boldsymbol{\vartheta}, \boldsymbol{z})$, and $\boldsymbol{g_Q}(\boldsymbol{\vartheta}, \boldsymbol{z})$.

$$\begin{cases} \boldsymbol{u} = \boldsymbol{g_u}(\boldsymbol{\vartheta}, \boldsymbol{z}) \\ \boldsymbol{P} = \boldsymbol{g_P}(\boldsymbol{\vartheta}, \boldsymbol{z}) \\ \boldsymbol{Q} = \boldsymbol{g_Q}(\boldsymbol{\vartheta}, \boldsymbol{z}) \end{cases} \tag{12}$$

$$\boldsymbol{z} = \begin{bmatrix} \boldsymbol{p} \\ \boldsymbol{q} \end{bmatrix} \tag{13}$$

$$\boldsymbol{\vartheta} = \begin{bmatrix} \boldsymbol{r} \\ \boldsymbol{x} \end{bmatrix} \tag{14}$$

$$\begin{cases} \boldsymbol{u} = [U_i] \\ \boldsymbol{P} = [P_{ij}] \\ \boldsymbol{Q} = [Q_{ij}] \end{cases} \tag{15}$$

where $\boldsymbol{p} = [p_i]$; $\boldsymbol{q} = [q_i]$; $\boldsymbol{r} = [r_{ij}]$; and $\boldsymbol{x} = [x_{ij}]$.

Returning to the problem we are addressing in this paper, we assume noisy measurements of $\boldsymbol{z}$, P_{01}, Q_{01}, and $\boldsymbol{u}$, and note that the models for P_{01} and Q_{01} are presented as the first elements of $\boldsymbol{g_P}(\boldsymbol{\vartheta}, \boldsymbol{z})$ and $\boldsymbol{g_Q}(\boldsymbol{\vartheta}, \boldsymbol{z})$, respectively. Recalling that t indicates a time instant, then we can stack all the measurements into a vector $\boldsymbol{y}_t$, and we have the following model for demonstrating how the impedances $\boldsymbol{\vartheta}$ and net injected power $\boldsymbol{z}_t$ relate to the available measurements at time t, reflecting that we measure P_{01} and Q_{01} instead of the remaining line power flows.

$$\boldsymbol{y}_t = \underbrace{\begin{bmatrix} \boldsymbol{z}_t \\ g_{P_{01}}(\boldsymbol{\vartheta}, \boldsymbol{z}_t) \\ g_{Q_{01}}(\boldsymbol{\vartheta}, \boldsymbol{z}_t) \\ \boldsymbol{g_u}(\boldsymbol{\vartheta}, \boldsymbol{z}_t) \end{bmatrix}_t}_{\boldsymbol{g}(\boldsymbol{\vartheta}, \boldsymbol{z}_t)} + \boldsymbol{e}_t \in \mathbb{R}^{n_y} \tag{16}$$

$$g_{P_{01}}(\boldsymbol{\vartheta}, \boldsymbol{z}_t) = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix} \boldsymbol{g_P}(\boldsymbol{\vartheta}, \boldsymbol{z}_t) \tag{17}$$

$$g_{Q_{01}}(\boldsymbol{\vartheta}, \boldsymbol{z}_t) = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix} \boldsymbol{g_Q}(\boldsymbol{\vartheta}, \boldsymbol{z}_t) \tag{18}$$

Further, $\boldsymbol{e}_t$ is a random variable that models uncertainty in the measurements and we employ the commonly used assumption (see [16], [18], [21], [22]) that it has a multivariate normal distribution and is independently and identically distributed (i.i.d.):

$$\boldsymbol{e}_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \boldsymbol{\Sigma}) \tag{19}$$

where $\boldsymbol{\Sigma} \in \mathbb{R}^{n_y \times n_y}$ is a positive-definite and symmetric covariance matrix. Since $\boldsymbol{\Sigma}$ is not generally known, we will estimate this matrix along with impedances $\boldsymbol{\vartheta}$.

We will henceforth refer to $\boldsymbol{z}_t$ as the state vector. Since $\boldsymbol{z}_t$ is not known exactly, we assume that it is a random variable following a multivariate normal distribution with mean $\boldsymbol{\mu}_t$ and covariance $\boldsymbol{\Pi}_t$, which is also a commonly employed assumption [14], [36].

$$\boldsymbol{z}_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\Pi}_t) \tag{20}$$

Since the mean $\boldsymbol{\mu}_t$ and covariance $\boldsymbol{\Pi}_t$ are typically not known, we will also estimate these terms along with $\boldsymbol{\vartheta}$ and $\boldsymbol{\Sigma}$. It will be convenient to collect all the unknown parameters into a single object $\boldsymbol{\theta}$ as:

$$\begin{cases} \boldsymbol{\theta} = [\boldsymbol{\beta}, \boldsymbol{\eta}] \\ \boldsymbol{\beta} = [\boldsymbol{\mu}_{1:T}, \boldsymbol{\Pi}_{1:T}] \\ \boldsymbol{\eta} = [\boldsymbol{\vartheta}, \boldsymbol{\Sigma}] \end{cases} \tag{21}$$

where $\boldsymbol{\mu}_{1:T} = \{\boldsymbol{\mu}_t\}_{t=1}^T$; and $\boldsymbol{\Pi}_{1:T} = \{\boldsymbol{\Pi}_t\}_{t=1}^T$.

Problem statement Given the measurements $\boldsymbol{y}_{1:T} = \{\boldsymbol{y}_t\}_{t=1}^T$, construct the MLE problem of unknown parameters $\boldsymbol{\theta}$ as:

$$\boldsymbol{\theta}_{ML} = \arg\max_{\boldsymbol{\theta}} \log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}) \tag{22}$$

where $p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T})$ is the probability distribution function parameterized by unknown $\boldsymbol{\theta}$. A similar notation is used for other distributions in this paper.

III. Proposed Solution

This section explains how the EM algorithm can be applied to solve the MLE problem (22). Additionally, this section provides the implementation details that maintain numerical robustness.

A. EM Algorithm

The main challenge in solving (22) arises from the unavailability of $\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T})$. The EM algorithm begins by inspecting MLE problem (22) to determine whether, hypothetically, extra data would render the problem tractable. These extra data are often called the "missing" data. In the current scenario, if we are additionally provided with the states $\boldsymbol{z}_{1:T}$, then the log-likelihood problem would become:

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T}) \tag{23}$$

According to our assumptions in (12)-(21), (23) can be expressed as a nonlinear least-squares problem and this can be solved using off-the-shelf optimisation routines. Of course, we do not have the desired "missing" data $\boldsymbol{z}_{1:T}$, but it seems tantalising to replace the missing data with an estimate so that we can update the unknown $\boldsymbol{\theta}$ variables by solving a tractable problem like (23). It may be tempting to replace the states $\boldsymbol{z}_{1:T}$ with the data from $\boldsymbol{y}_{1:T}$ related to these states (refer to (16)), but this is well-known to produce poor estimates [37].

Therefore, a natural question is how to choose an accurate estimation of the missing data to solve problem (22), which is our primary concern in this paper. To answer this, we note that there is an important link between the log-likelihood $\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T})$ and the joint log-likelihood $\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$, which are related via conditional probability as:

$$\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}) = \log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T}) - \log p_{\boldsymbol{\theta}}(\boldsymbol{z}_{1:T} | \boldsymbol{y}_{1:T}) \tag{24}$$

Further, we may take expectations of both sides with respect to any distribution over $\boldsymbol{z}_{1:T}$, i.e., $q(\boldsymbol{z}_{1:T})$ for the sake of exposition.

$$\int (\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T})) q(\boldsymbol{z}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T} = \int (\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})) q(\boldsymbol{z}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T} - \int (\log p_{\boldsymbol{\theta}}(\boldsymbol{z}_{1:T} | \boldsymbol{y}_{1:T})) q(\boldsymbol{z}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T} \tag{25}$$

Since $\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T})$ is not a function of $\boldsymbol{z}_{1:T}$, we have:

$$\int (\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T})) q(\boldsymbol{z}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T} = \log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}) \underbrace{\int q(\boldsymbol{z}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T}}_{=1} = \log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}) \tag{26}$$

$$\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}) = \int (\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})) q(\boldsymbol{z}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T} - \int (\log p_{\boldsymbol{\theta}}(\boldsymbol{z}_{1:T} | \boldsymbol{y}_{1:T})) q(\boldsymbol{z}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T} \tag{27}$$

If we make the careful choice for the distribution $q(\boldsymbol{z}_{1:T})$ as (28), then we have (29).

$$q(\boldsymbol{z}_{1:T}) = p_{\boldsymbol{\theta}'}(\boldsymbol{z}_{1:T} | \boldsymbol{y}_{1:T}) \tag{28}$$

$$\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}) = Q(\boldsymbol{\theta}, \boldsymbol{\theta}') - V(\boldsymbol{\theta}, \boldsymbol{\theta}') \tag{29}$$

where $\boldsymbol{\theta}'$ is a parameter estimate of unknown parameter $\boldsymbol{\theta}$; $Q(\boldsymbol{\theta}, \boldsymbol{\theta}') = \int (\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})) p_{\boldsymbol{\theta}'}(\boldsymbol{z}_{1:T} | \boldsymbol{y}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T}$; and $V(\boldsymbol{\theta}, \boldsymbol{\theta}') = \int (\log p_{\boldsymbol{\theta}}(\boldsymbol{z}_{1:T} | \boldsymbol{y}_{1:T})) p_{\boldsymbol{\theta}'}(\boldsymbol{z}_{1:T} | \boldsymbol{y}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T}$.

Therefore, the difference of the two log-likelihoods can be expressed as:

$$\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}) - \log p_{\boldsymbol{\theta}'}(\boldsymbol{y}_{1:T}) = Q(\boldsymbol{\theta}, \boldsymbol{\theta}') - Q(\boldsymbol{\theta}', \boldsymbol{\theta}') + \underbrace{V(\boldsymbol{\theta}', \boldsymbol{\theta}') - V(\boldsymbol{\theta}, \boldsymbol{\theta}')}_{\geq 0} \tag{30}$$

The inequality $V(\boldsymbol{\theta}', \boldsymbol{\theta}') - V(\boldsymbol{\theta}, \boldsymbol{\theta}') \geq 0$ holds because the difference in V is the Kullback-Leibler divergence, which is always non-negative. Importantly, since our primary aim is to maximise $\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T})$, one possible strategy is to find a $\boldsymbol{\theta}$ that increases $Q(\boldsymbol{\theta}, \boldsymbol{\theta}') - Q(\boldsymbol{\theta}', \boldsymbol{\theta}')$. This is the main strategy of the EM algorithm: compute the Q-function via expectation, and then maximise Q to obtain a new parameter vector that also guarantees the increase of log-likelihood. This process is repeated until no practical increase in the log-likelihood can be achieved.

To convert this idea into an algorithm, we note that $\boldsymbol{\theta}'$ can be replaced by the current parameter estimate $\boldsymbol{\theta}_k$ at the k^{th} iteration, and then $\boldsymbol{\theta}_{k+1}$ is obtained by maximising $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$ over $\boldsymbol{\theta}$. This is summarised as follows.

Step 1: let $k = 1$ and initialise $\boldsymbol{\theta}_1$.

Step 2: repeat the following steps for $k = 1, 2, \ldots$.

① Expectation step (E-step): form the Q-function $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$.

② Maximization step (M-step): solve $\boldsymbol{\theta}_{k+1} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$.

B. E-step

In this subsection, we form the Q-function $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$, which can be written as:

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k) = \int (\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})) p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_{1:T} | \boldsymbol{y}_{1:T}) \mathrm{d}\boldsymbol{z}_{1:T} = \sum_{t=1}^{T} \int (\log (p_{\boldsymbol{\theta}}(\boldsymbol{y}_t | \boldsymbol{z}_t) p_{\boldsymbol{\theta}}(\boldsymbol{z}_t))) p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t | \boldsymbol{y}_t) \mathrm{d}\boldsymbol{z}_t \tag{31}$$

where the second equality comes from the independence assumptions on $\boldsymbol{z}_t$ and $\boldsymbol{e}_t$ and their associations with $\boldsymbol{y}_t$. Exploiting properties of the logarithmic function, we can express $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$ as:

$$Q(\boldsymbol{\theta},\boldsymbol{\theta}_k)=Q_1(\boldsymbol{\beta},\boldsymbol{\theta}_k)+Q_2(\boldsymbol{\eta},\boldsymbol{\theta}_k) \tag{32}$$

$$Q_1(\boldsymbol{\beta},\boldsymbol{\theta}_k)=\sum_{t=1}^{T}\int(\log p_{\boldsymbol{\beta}}(\boldsymbol{z}_t))p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t|\boldsymbol{y}_t)\mathrm{d}\boldsymbol{z}_t \tag{33}$$

$$Q_2(\boldsymbol{\eta},\boldsymbol{\theta}_k)=\sum_{t=1}^{T}\int(\log p_{\boldsymbol{\eta}}(\boldsymbol{y}_t|\boldsymbol{z}_t))p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t|\boldsymbol{y}_t)\mathrm{d}\boldsymbol{z}_{t|t} \tag{34}$$

$p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t|\boldsymbol{y}_t)$ is generally intractable, and it is therefore not trivial to compute these expectations. Since the system is nonlinear, the linear regression techniques cannot be used to determine the $p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t|\boldsymbol{y}_t)$ distribution. Consequently, this makes the entire EM algorithm different from that in [23]. This paper investigates a multivariate normal approximation to $p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t|\boldsymbol{y}_t)$ based on a first-order Taylor expansion. Towards this, the first-order Taylor expansion of $\boldsymbol{g}(\boldsymbol{\vartheta}_k,\boldsymbol{z}_t)$ about the prior mean $\boldsymbol{\mu}_t$ is given by:

$$\boldsymbol{y}_t\approx\boldsymbol{g}(\boldsymbol{\vartheta}_k,\boldsymbol{\mu}_t)+\boldsymbol{C}_t(\boldsymbol{z}_t-\boldsymbol{\mu}_t)+\boldsymbol{e}_t \tag{35}$$

$$\boldsymbol{C}_t=\left.\frac{\partial\boldsymbol{g}(\boldsymbol{\vartheta}_k,\boldsymbol{z}_t)}{\partial\boldsymbol{z}_t^{\mathrm{T}}}\right|_{\boldsymbol{z}_t=\boldsymbol{\mu}_t} \tag{36}$$

It is well known [38], [39] that $p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t|\boldsymbol{y}_t)$ can be adequately approximated with a multivariate normal distribution with mean $\boldsymbol{\mu}_{t|t}$ and covariance $\boldsymbol{\Pi}_{t|t}$:

$$p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t|\boldsymbol{y}_t)\approx\mathcal{N}(\boldsymbol{\mu}_{t|t},\boldsymbol{\Pi}_{t|t}) \tag{37}$$

$$\boldsymbol{\mu}_{t|t}=\boldsymbol{\mu}_t+\boldsymbol{\Pi}_t\boldsymbol{C}_t^{\mathrm{T}}(\boldsymbol{C}_t\boldsymbol{\Pi}_t\boldsymbol{C}_t^{\mathrm{T}}+\boldsymbol{\Sigma})^{-1}(\boldsymbol{y}_t-\boldsymbol{g}(\boldsymbol{\vartheta}_k,\boldsymbol{\mu}_t)) \tag{38}$$

$$\boldsymbol{\Pi}_{t|t}=\boldsymbol{\Pi}_t-\boldsymbol{\Pi}_t\boldsymbol{C}_t^{\mathrm{T}}(\boldsymbol{C}_t\boldsymbol{\Pi}_t\boldsymbol{C}_t^{\mathrm{T}}+\boldsymbol{\Sigma})^{-1}\boldsymbol{C}_t\boldsymbol{\Pi}_t \tag{39}$$

This approximation can then be utilised within (31) to approximate the Q-function. This is summarised in Lemma 1.

Lemma 1 Assuming that $p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t|\boldsymbol{y}_t)\approx\mathcal{N}(\boldsymbol{\mu}_{t|t},\boldsymbol{\Pi}_{t|t})$ with mean and covariance given by (38) and (39), respectively, we have:

$$Q(\boldsymbol{\theta},\boldsymbol{\theta}_k)\approx\hat{Q}_1(\boldsymbol{\beta},\boldsymbol{\theta}_k)+\hat{Q}_2(\boldsymbol{\eta},\boldsymbol{\theta}_k) \tag{40}$$

$$\hat{Q}_1(\boldsymbol{\beta},\boldsymbol{\theta}_k)=c_1-\frac{1}{2}\sum_{t=1}^{T}tr\{\boldsymbol{\Pi}_t^{-1}(\boldsymbol{\mu}_t-\boldsymbol{\mu}_{t|t})(\boldsymbol{\mu}_t-\boldsymbol{\mu}_{t|t})^{\mathrm{T}}\}-\frac{1}{2}\log\left|\boldsymbol{\Pi}_t\right|+tr\{\boldsymbol{\Pi}_t^{-1}\boldsymbol{\Pi}_{t|t}\} \tag{41}$$

$$\hat{Q}_2(\boldsymbol{\eta},\boldsymbol{\theta}_k)=c_2-\frac{T}{2}\log|\boldsymbol{\Sigma}|-\frac{1}{2}tr\{\boldsymbol{\Sigma}^{-1}\Gamma(\boldsymbol{\vartheta})\} \tag{42}$$

where c_1 and c_2 are the constants; and $\Gamma(\boldsymbol{\vartheta})$ is expressed as:

$$\Gamma(\boldsymbol{\vartheta})=\sum_{t=1}^{T}(\boldsymbol{y}_t-\boldsymbol{g}(\boldsymbol{\vartheta},\boldsymbol{\mu}_{t|t}))(\boldsymbol{y}_t-\boldsymbol{g}(\boldsymbol{\vartheta},\boldsymbol{\mu}_{t|t}))^{\mathrm{T}}+\boldsymbol{C}_{t|t}(\boldsymbol{\vartheta})\boldsymbol{\Pi}_{t|t}\boldsymbol{C}_{t|t}^{\mathrm{T}}(\boldsymbol{\vartheta}) \tag{43}$$

$$\boldsymbol{C}_{t|t}(\boldsymbol{\vartheta})=\left.\frac{\partial\boldsymbol{g}(\boldsymbol{\vartheta}_k,\boldsymbol{z}_t)}{\partial\boldsymbol{z}_t^{\mathrm{T}}}\right|_{\boldsymbol{z}_t=\boldsymbol{\mu}_{t|t}} \tag{44}$$

Proof of Lemma 1 See Supplementary Material A.

Lemma 1 relies on the mean $\boldsymbol{\mu}_{t|t}$ and covariance $\boldsymbol{\Pi}_{t|t}$. These terms can be computed in a numerically robust manner using the square-root form [40], [41] as outlined in Lemma 2.

Lemma 2 Given the QR factorization in (45), the mean $\boldsymbol{\mu}_{t|t}$ and the Cholesky factor $\boldsymbol{\Pi}_{t|t}^{\frac{1}{2}}$ can be computed by (46) and (47), respectively.

$$\boldsymbol{Q}\begin{bmatrix}\boldsymbol{R}_1 & \boldsymbol{R}_2\\ \boldsymbol{0} & \boldsymbol{R}_3\end{bmatrix}=\begin{bmatrix}\boldsymbol{\Sigma}^{\frac{1}{2}} & 0\\ \boldsymbol{\Pi}_t^{\frac{1}{2}}\boldsymbol{C}_t^{\mathrm{T}} & \boldsymbol{\Pi}_t^{\frac{1}{2}}\end{bmatrix} \tag{45}$$

$$\boldsymbol{\mu}_{t|t}=\boldsymbol{\mu}_t+\boldsymbol{R}_2^{\mathrm{T}}\boldsymbol{R}_1^{-\mathrm{T}}(\boldsymbol{y}_t-\boldsymbol{g}(\boldsymbol{\vartheta}_k,\boldsymbol{\mu}_t)) \tag{46}$$

$$\boldsymbol{\Pi}_{t|t}^{\frac{1}{2}}=\boldsymbol{R}_3 \tag{47}$$

where $\boldsymbol{Q}$ is an orthogonal matrix; $\boldsymbol{R}_1$ and $\boldsymbol{R}_3$ are the upper triangular matrices; and $\boldsymbol{R}_2$ is a block matrix.

Proof of Lemma 2 See Supplementary Material B.

C. M-step

After approximating the Q-function and the distribution of system state variables $p_{\boldsymbol{\theta}_k}(\boldsymbol{z}_t|\boldsymbol{y}_t)$ in the E-step, we would maximize the Q-function approximation to obtain the unknown parameters:

$$\boldsymbol{\theta}^*=\arg\max_{\boldsymbol{\theta}}(\hat{Q}_1(\boldsymbol{\beta},\boldsymbol{\theta}_k)+\hat{Q}_2(\boldsymbol{\eta},\boldsymbol{\theta}_k)) \tag{48}$$

The optimal solution of (48) $\boldsymbol{\theta}^*$ can be provided according to Lemma 3.

Lemma 3 The optimal solution of (48) is provided by $\boldsymbol{\theta}^*=[\boldsymbol{\beta}^*,\boldsymbol{\eta}^*]$, where $\boldsymbol{\beta}^*$ is given by (49); and $\boldsymbol{\eta}^*=[\boldsymbol{\vartheta}^*,\boldsymbol{\Sigma}]$. Recalling the definition of $\Gamma(\boldsymbol{\vartheta})$ in (43), $\boldsymbol{\vartheta}^*$ and $\boldsymbol{\Sigma}^*$ are given in (50) and (51), respectively.

$$\boldsymbol{\beta}^*=[\boldsymbol{\mu}_{1|1},\ldots,\boldsymbol{\mu}_{T|T},\boldsymbol{\Pi}_{1|1},\ldots,\boldsymbol{\Pi}_{T|T}] \tag{49}$$

$$\boldsymbol{\vartheta}^*=\arg\min_{\boldsymbol{\vartheta}}\frac{T}{2}\log\left|\frac{1}{T}\Gamma(\boldsymbol{\vartheta})\right| \tag{50}$$

$$\boldsymbol{\Sigma}^*=\frac{1}{T}\Gamma(\boldsymbol{\vartheta}^*) \tag{51}$$

Proof of Lemma 3 See Supplementary Material C.

The optimisation problem (50) is differentiable and can be solved using any appropriate optimisation routine. For the experimental results in this paper, we use the fmincon function in MATLAB [42].

D. Process of EM Algorithm

In this subsection, we bring together all the components of the EM algorithm, as demonstrated in Algorithm 1.

IV. Experimental Results: Simulated Data

In this section, we evaluate the proposed method on the IEEE 37-node test network [30], which functions as a standard network for testing this problem (even smaller networks have been used for similar problems [16], [21]). The network topology is known, but the information about line impedances is unavailable.

In this case, the load profiles are simulated for the analysis. Subsequently, we obtain measurements by running the full AC power flow as described in [43]. We assume that these measurements are subject to Gaussian noise corruption with a standard deviation of 0.001, which is consistent with [10].

Algorithm 1: EM algorithm

Input: noisy measurements $\boldsymbol{y}_{1:T}$, initial parameter estimate $\boldsymbol{\theta}_1$, and the maximum number of iterations K
Output: $\boldsymbol{\theta}_{K+1}$
for $k=1$ to K **do**
  Begin E-step
  for $t=1$ to T **do**
    Obtain $\boldsymbol{\mu}_{t|t}$ and $\boldsymbol{\Pi}_{t|t}^{\frac{1}{2}}$ from Lemma 2 using $\boldsymbol{\theta}_k$
    Set $\boldsymbol{\Pi}_{t|t}=(\boldsymbol{\Pi}_{t|t}^{\mathrm{T}})^{\frac{1}{2}}\boldsymbol{\Pi}_{t|t}^{\frac{1}{2}}$
  end for
  end E-step
  Begin M-step
  Set $\boldsymbol{\beta}_{k+1}=[\boldsymbol{\mu}_{1|1},\ldots,\boldsymbol{\mu}_{T|T},\boldsymbol{\Pi}_{1|1},\ldots,\boldsymbol{\Pi}_{T|T}]$
  Solve (50) and set $\boldsymbol{\vartheta}_{k+1}=\boldsymbol{\vartheta}^*$
  Set $\boldsymbol{\Sigma}_{k+1}=\Gamma(\boldsymbol{\vartheta}_{k+1})/T$
  end M-step
end for

The introduced standard deviation represents a 10% relative error, in accordance with similar studies in [16] and [21]. Additionally, the added noise introduces a mean error of 4.5% following Gaussian distribution, exceeding the maximum values of 2% in [22] and 0.5% in [18].

All nodes have available measurements, consistent with similar studies in [10], [16], [18], [20], and [21]. Although we have assumed the measurement noise distribution to be the identical for each measurement, the proposed method does not require this assumption.

In this section, we will compare our experimental results with those presented in our recently published work [23], which employs an EM method with a linear power flow model (referred to as EM-LIN for simplicity, and the variables of which will be denoted with a subscript *lin*). Both studies use the same test setup. We will compare the results for estimating line parameters (r and x) and state variables (p_t and q_t). Additionally, we will analyze the voltage errors.

Furthermore, we will delve into the convergence analysis of the EM algorithm concerning line parameters. We conduct the test on $T=600$ timestamps, representing a dataset of less than half a day with 1-min resolution. $T=600$ has been chosen to represent a realistic case. In comparison, the study in [10] utilized measurements over three weeks to estimate line parameters. This paper randomly initializes the line parameters with values greater than 4 to 20 times the original, while [10] initialized them based on the utility database, which might not be easily available. The selection of the variable K in Algorithm 1 could be dynamically determined based on the difference in log-likelihood threshold at each iteration. We can set a threshold in the order of 10^{-7}. Another way to choosing the value of K is to set it sufficiently high to ensure the algorithm convergence. The relevance of Algorithm 1 lies in its ability to estimate new line parameters, namely resistance and reactance with limited data, a crucial aspect when changes occur in the line parameters.

A. Line Parameter Estimate

In this subsection, we discuss the results of the line parameter estimate. We present a comparison of the mean absolute error in line parameter estimates, as listed in Table I.

TABLE I
MEAN ABSOLUTE ERROR IN LINE PARAMETER ESTIMATES

Method	Mean absolute error in r	Mean absolute error in x
Proposed	1.45×10^{-4}	2.20×10^{-4}
EM-LIN	1.50×10^{-2}	5.33×10^{-4}

In Table I, the proposed method shows an improvement of 100 times on the mean absolute error in r and more than twice on the mean absolute error in x compared with the EM-LIN. The proposed method outperforms EM-LIN particularly when considering a limited number of measurements. This observation aligns with our focus on accounting for line losses, a factor overlooked by EM-LIN. While it is worth noting that the EM-LIN can achieve higher accuracy with an increased number of measurements, it still falls short of the level of accuracy achieved by the proposed method. The proposed method delivers accurate results even with fewer measurements and compensates for the systematic error introduced by accounting for losses. It is important to highlight that the proposed method is tested with various initializations of the estimates, and we consistently achieve accurate results in all cases.

B. Voltage and State Variable Estimates

In this subsection, we analyze the errors in voltage and state variable estimates p and q, as shown in Table II.

TABLE II
ERRORS IN VOLTAGE AND STATE VARIABLE ESTIMATES

Method	Voltage		p		q	
	MSE	MAE	MSE	MAE	MSE	MAE
Proposed	1.96×10^{-7}	1.66×10^{-3}	1.44×10^{-6}	4.63×10^{-3}	2.53×10^{-6}	5.66×10^{-3}
EM-LIN	7.14×10^{-7}	3.87×10^{-3}	1.75×10^{-2}	5.85×10^{-1}	1.53×10^{-2}	4.80×10^{-1}

From Table II, it should be noted that the mean square error (MSE) in voltage of the proposed method is 3.5 times lower than that of the EM-LIN, which highlights the superiority of the proposed method. This is equivalent to a reduction in error by 72%. Also, the maximum absolute error (MAE) in voltage of the proposed method is half as high as that of EM-LIN.

It is worth noting that the errors in voltage of EM-LIN are not relatively high when compared with those in the line parameters and state variables. This aligns with the expected behaviour for EM-LIN because the $\boldsymbol{\vartheta}_{lin}$ line estimates compensate for the line losses overlooked by the linear power flow model. In contrast, the proposed method incorporates the line losses into the power flow model, leading to im-

proved estimates without compensating for line losses in the estimates.

In Table II, the MSE in active power of the proposed method is lower by a factor of 10000 when compared with that of EM-LIN. A similar trend is observed for reactive power. The MAE of the proposed method is also lower by a factor of over 100 in both active power and reactive power. This demonstrates that the proposed method is well-suited for estimating system state variables p and q and performs exceptionally well in this context.

Figure 2 presents the voltage error distribution across all cases, with a maximum error of only 1.66×10^{-3}. These results further validate the accuracy and reliability of the proposed method.

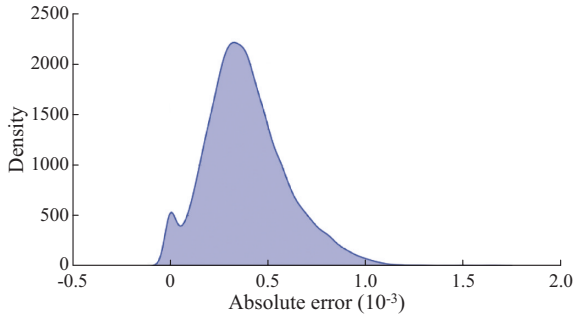

Fig. 2. Voltage error distribution.

For the IEEE 37-node test network, the proposed method requires approximately 5 hours on a single Intel Core Ultra 7 155H CPU. This remains within practical limits for line parameter estimation. Additionally, the proposed method primarily depends on the convergence behaviour of the EM algorithm and the solver used for parameter updates. The proposed method scales up with the number of nodes quadratically times a factor of 1/2 in a radial network.

C. Convergence of EM Algorithm

In this subsection, we demonstrate the convergence of EM algorithm by visually representing the log-likelihood values in Fig. 3 at each iteration.

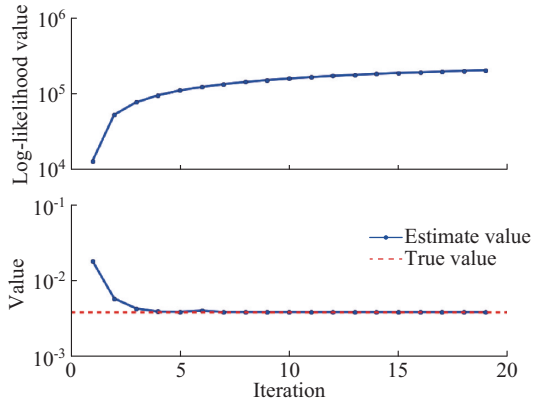

Fig. 3. Demonstration of convergence of EM algorithm. (a) Log-likelihood value. (b) Convergence of line parameter estimate.

As discussed in Section III-A, maximizing the Q-function corresponds to maximizing the log-likelihood. This behaviour is evident in Fig. 3, where the log-likelihood consistently increases with each EM iteration, confirming the convergence of EM algorithm. For an illustration, we provide one specific instance in Fig. 3 to showcase the convergence of the line parameter estimate, where it converges to the true value.

D. Comparison with a Latest Study

This subsection compares the proposed method with a recent related method [18]. We replicate the test setup from [18], which uses $T = 200$ timestamps and a measurement error of 0.5%. A difference exists in the state variables used: [18] assumes branch currents and nodal voltage in rectangular coordinates as state variables, and the proposed method uses p and q, which, as discussed earlier, are more readily available in the DG.

The accuracy of voltage estimates of the proposed method and the method in [18] is compared. The third quartile values of absolute voltage error with the method in [18] and the proposed method are around 0.001 p.u. and 0.0005 p.u, respectively, and the maximum values are 0.007 p. u. and 0.0012 p.u., respectively. This shows that the proposed method is highly robust.

V. Experimental Results: Real-world Data

In this section, we present the results of real-world loads at the consumer end collected in the Australian Capital Territory (ACT). This dataset is known as the NextGen dataset [44]. The customers here have solar panels and batteries, so this dataset represents the modern DGs. Measurements are available with a 5-min time resolution over multiple days. However, the timesteps across different nodes are not uniform. To capture the behaviour of the grid at a specific point in time, we filter the dataset to have common timesteps across different nodes. The measurement timesteps included are randomised to give a greater spread of real-world behaviour of loads. Additionally, to align with our previous test, we only consider 600 load values, even with access to data over multiple days. This corresponds to 600 timestamps. We use a limited dataset to demonstrate the robustness of EM algorithm under reduced data availability.

After obtaining the load data, we run the full AC power flow and introduce the same Gaussian noise, as detailed in Section IV, to generate the measurement set. We run the power flow on the IEEE 37-node test network, using the same test setup as described in the case study of Section IV. The line parameters are randomly initialized, with each value exceeding the true values by a factor ranging from 4 to 20.

Table III compares the mean absolute error in line parameter estimates, r and x, with the proposed method and EM-LIN using real-world consumer loads. The mean absolute error in r with proposed method is 100 times less than that with the EM-LIN and twice less for x.

The proposed method consistently delivers accurate results for line parameter estimates, while demonstrating its accurate performance across real-world consumer loads.

TABLE III
MEAN ABSOLUTE ERROR IN LINE PARAMETER ESTIMATES USING REAL-WORLD CONSUMER LOADS

Method	Mean absolute error in r	Mean absolute error in x
Proposed	1.45×10^{-4}	2.20×10^{-4}
EM-LIN	1.50×10^{-2}	5.33×10^{-4}

VI. CONCLUSION

In this study, we propose a state-of-the-art method to estimate line parameters using the EM algorithm. This method integrates the modified Distflow model to account for line losses, enhancing model accuracy. As the power system is nonlinear, we cannot estimate the state variables using Bayesian regression. To tackle this, we approximate the distribution of state variables given measurements using the first-order Taylor expansion and implement it computationally efficiently using the square root form. We further approximate the Q-function and maximize it to solve the parameters. Importantly, the proposed method uses measurements easily available in the DG: voltage magnitude and node power.

We demonstrate the accuracy of the proposed method through simulation on an IEEE 37-node test feeder and subsequently compare the results with a previous method that uses a linear power flow model. We demonstrate the performance of the proposed method using simulated and real-world data. The proposed method accurately estimates the line parameters, system state variables, and voltage magnitude. Additionally, we show the algorithm convergence by plotting the log-likelihood, which increases at each iteration as we maximize the Q-function. Accurate results can be achieved using data of less than half a day, showing the practical utility of the proposed method.

The proposed method accurately determines the DG line parameters. This information helps in having better state estimation techniques. This will improve control capabilities to mitigate issues from increasing DER integration. This will increase DER adoption in DGs, aiding the reduction of carbon emissions.

In future studies, we will examine the impact of missing measurements and device placement on estimating system state variables and line parameters. Additionally, the proposed method will be evaluated under non-Gaussian noise distributions to assess its robustness.

REFERENCES

[1] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Power Engineering Review*, vol. 9, no. 4, pp. 101-102, Apr. 1989.

[2] National Centers for Environmental Information. (2023, Jul.). Global Climate Report. [Online]. Available: https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202307

[3] D. Gielen, F. Boshell, D. Saygin *et al*., "The role of renewable energy in the global energy transformation," *Energy Strategy Reviews*, vol. 24, pp. 38-50, Apr. 2019.

[4] Clean Energy Council. (2022, May). Clean Energy Australia Report 2022. [Online]. Available: https://cleanenergycouncil.org.au/cec/media/background/resources/clean-energy-australia-report-2022.pdf

[5] Y. Lin, J. H. Eto, B. B. Johnson *et al*., "Research roadmap on grid-forming inverters," National Renewable Energy Lab (NREL), Golden, USA, Tech. Rep., 2020.

[6] P. Denholm, M. O'Connell, G. Brinkman *et al*., "Overgeneration from solar energy in california. a field guide to the duck chart," National Renewable Energy Lab (NREL), Golden, USA, Tech. Rep., 2015.

[7] Australian Energy Market Operator (AEMO). (2023, Sept.). NEM Operational Demand and Generation Mix on Sunday 17 September 2023. [Online]. Available: https://www. linkedin. com/posts/australian-energy-market-operator_yesterday-the-national-electricity-market-activity-7109379547325730816-oFi8/

[8] W. C. de Carvalho, E. L. Ratnam, L. Blackhall *et al*., "Optimization-based operation of distribution grids with residential battery storage: assessing utility and customer benefits," *IEEE Transactions on Power Systems*, vol. 38, no. 1, pp. 218-228, Jan. 2023.

[9] S. Radhoush, M. Bahramipanah, H. Nehrir *et al*., "A review on state estimation techniques in active distribution networks: existing practices and their challenges," *Sustainability*, vol. 14, no. 5, p. 2520, Mar. 2022.

[10] A. C. Varghese, A. Pal, and G. Dasarathy, "Transmission line parameter estimation under non-Gaussian measurement noise," *IEEE Transactions on Power Systems*, vol. 38, no. 4, pp. 3147-3162, Jul. 2023.

[11] M. Fotopoulou, S. Petridis, I. Karachalios *et al*., "A review on distribution system state estimation algorithms," *Applied Sciences*, vol. 12, no. 21, p. 11073, Nov. 2022.

[12] A. Primadianto and C. N. Lu, "A review on distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3875-3883, Sept. 2017.

[13] X. Zhou, Z. Liu, Y. Guo *et al*., "Gradient-based multi-area distribution system state estimation," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5325-5338, Nov. 2020.

[14] M. Antončič, I. Papič, and B. Blažič, "Robust and fast state estimation for poorly-observable low voltage distribution networks based on the Kalman filter algorithm," *Energies*, vol. 12, no. 23, p. 4457, Dec. 2019.

[15] V. Vittal, "The impact of renewable resources on the performance and reliability of the electricity grid," *The Bridge*, vol. 40, no. 1, pp. 5-12, Jan. 2010.

[16] J. 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.

[17] J. Ban, J. Y. Park, Y. J. Kim *et al*., "AMI data-driven strategy for hierarchical estimation of distribution line impedances," *IEEE Transactions on Power Delivery*, vol. 38, no. 1, pp. 513-527, Feb. 2023.

[18] M. Vanin, F. Geth, R. D'hulst *et al*., "Combined unbalanced distribution system state and line impedance matrix estimation," *International Journal of Electrical Power & Energy Systems*, vol. 151, p. 109155, Sept. 2023.

[19] A. M. Prostejovsky, O. Gehrke, A. M. Kosek *et al*., "Distribution line parameter estimation under consideration of measurement tolerances," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 726-735, Apr. 2016.

[20] S. Claeys, F. Geth, and G. Deconinck, "Line parameter estimation in multi-phase distribution networks without voltage angle measurements," *IET Conference Proceedings*, vol. 2021, no. 6, pp. 1186-1190, Nov. 2021.

[21] J. 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.

[22] J. Zhang, Y. Wang, Y. Weng *et al*., "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, Sept. 2020.

[23] S. Kapoor, J. Hendriks, and L. Blackhall, "System identification in distribution grid without phase angle using expectation maximization," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 10959-10964, Jan. 2023.

[24] S. Kapoor, B. Sturmberg, and M. Shaw. (2020, Jul.). A Review of Publicly Available Energy Data Sets, Wattwatchers' My Energy Marketplace (MEM). [Online]. Available: https://wattwatchers.com.au/wp-content/uploads/2020/07/ANU_literature_Review_of_Energy_Data_Sets.pdf

[25] S. Silwal, C. Mullican, Y. A. Chen *et al*., "Open-source multi-year power generation, consumption, and storage data in a microgrid," *Journal of Renewable and Sustainable Energy*, vol. 13, no. 2, p. 025301, Mar. 2021.

[26] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*: *Series B* (*Methodological*), vol. 39, no. 1, pp. 1-22,

Sept. 1977.
[27] T. Schön. (2009, May). An explanation of the expectation maximization algorithm. [Online]. Available: https://www.rt.isy.liu.se/research/reports/2009/2915.pdf
[28] T. B. Schön, A. Wills, and B. Ninness, "System identification of nonlinear state-space models," *Automatica*, vol. 47, no. 1, pp. 39-49, Jan. 2011.
[29] A. Wigren, J. Wågberg, F. Lindsten *et al*., "Nonlinear system identification: learning while respecting physical models using a sequential Monte Carlo method," *IEEE Control Systems Magazine*, vol. 42, no. 1, pp. 75-102, Feb. 2022.
[30] W. H. Kersting, "Radial distribution test feeders," in *Proceedings of 2001 IEEE PES Winter Meeting*, Columbus, USA, Jan. 2001, pp. 908-912.
[31] M. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 735-743, Jan. 1989.
[32] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 725-734, Jan. 1989.
[33] L. Stewart. (2011, Jan.). Basic graph theory definitions and notation. [Online]. Available: https://webdocs.cs.ualberta.ca/~stewart/c672/Definitions.pdf
[34] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," in *Proceedings of 52nd IEEE Conference on Decision and Control*, Firenze, Italy, Dec. 2013, pp. 4329-4334.
[35] L. Gan and S. H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial networks," in *Proceedings of 2014 Power Systems Computation Conference*, Wroclaw, Poland, Aug. 2014, pp. 1-9.
[36] M. Shafiei, G. Ledwich, G. Nourbakhsh *et al*. (2018, Apr.). Layered based augmented complex Kalman filter for fast forecasting-aided state estimation of distribution networks. [Online]. Available: https://arxiv.org/pdf/1804.08298
[37] Z. Griliches and V. Ringstad, "Error-in-the-variables bias in nonlinear contexts," *Econometrica*, vol. 38, no. 2, p. 368, Mar. 1970.
[38] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95-108, Mar. 1961.
[39] G. Welch and G. Bishop. (2006, Jan.). An introduction to the Kalman filter. [Online]. Available: https://perso. crans. org/club-krobot/doc/kalman.pdf
[40] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*. New York: Dover Publications, 2006.
[41] C. Thornton, "Triangular covariance factorizations for Kalman filtering," Ph.D. dissertation, University of California at Los Angeles, Los Angeles, USA, 1976.
[42] MathWorks, *Optimization Toolbox User's Guide*, Version R2023b. Natick: The MathWorks, Inc., 2023.
[43] J. J. Grainger and W. D. Stevenson Jr., *Power System Analysis*. New York: McGraw-Hill Education, 1994.
[44] M. Shaw, B. Sturmberg, L. Guo *et al*., "The NextGen energy storage trial in the ACT, Australia," in *Proceedings of 10th ACM International Conference Future Energy System*, Phoenix, USA, Jun. 2019, pp. 439-442.

Shubhankar Kapoor received the B.Tech. degree in electronics and communication engineering from Amity University, Noida, India, in 2016, and the M.Eng. degree in mechatronics from The Australian National University (ANU), Canberra, Australia, in 2018. He is currently pursuing the Ph.D. degree at ANU. His research interests include distribution grid state estimation, impedance estimation, grid management, mathematical optimization, signal processing, and machine learning application in power system.

Adrian G. Wills received the B.E. degree in electrical engineering and the Ph.D. degree from The University of Newcastle, Callaghan, Australia, in 1999 and 2003, respectively. He has held postdoctoral research positions at The University of Newcastle and spent three years working in industry. Since July 2015, he has been leading the Mechatronics Engineering Program at The University of Newcastle. His research interest includes Bayesian estimation with application in robotics and mechatronics.

Johannes Hendriks received the B.E. degree in mechatronics engineering and the Ph.D. degree in mechanical engineering from The University of Newcastle, Callaghan, Australia, in 2015 and 2020, respectively. He worked as a Postdoctoral Researcher at The University of Newcastle. He is currently a Data Scientist at Evergen, Newcastle, Australia. His research interests include machine learning application in engineering, statistical modeling, and predictive analytics.

Lachlan Blackhall is currently the Deputy Vice-chancellor (Research and Innovation) at The Australian National University (ANU), Canberra, Australia. Previously, he was an Entrepreneurial Fellow and Head of the Battery Storage and Grid Integration Program at ANU. His research interests include renewable energy integration, distributed energy resource, and energy system optimization.