Individual Load Model Parameter Estimation in Distribution Systems using Load Switching Events

Alireza Shahsavari, *Student Member, IEEE*, Mohammad Farajollahi, *Student Member, IEEE*,
Hamed Mohsenian-Rad, *Senior Member, IEEE*

Abstract—There currently exists a mature literature on modeling the *aggregate* load of a distribution feeder by making use of measurements at its feeder-head at substation. The primary application of such feeder-aggregated load models is in sub-transmission or transmission system analysis. However, there is a growing need in practice also to model each *individual* load across the feeder. If available, such individual load models have applications in power distribution system analysis; e.g., to better integrate distributed energy resources or to improve power quality and reliability. Motivated by this observation, in this paper, we propose a new method for individual load modeling in power distribution systems. It works by using the measurements only at the feeder-head. It takes an innovative approach to analyzing the load switching events across the distribution feeder itself, instead of or in addition to relying on upstream voltage events that are commonly used in feeder-aggregated load modeling. By tracking the downstream load switching events, the proposed method can make a robust estimation of the ZIP load model parameters for all individual loads. The proposed method is examined on small illustrative test-feeders as well as the IEEE 33-bus test system under various operating scenarios. The adverse impact of errors in measurements and system parameters are also investigated on the performance of the developed load modeling method.

Keywords: Load modeling, distribution system analysis, load switching events, ZIP model, nonlinear parameter estimation.

NOMENCLATURE

$[\cdot]^T$	Transpose of a vector
$[\cdot]^*$	Complex conjugate of a scalar or a vector
$\lvert\cdot\rvert$	Magnitude of a complex number
$\mathbb{I}(\cdot)$	A 0-1 indicator function
k	Sequence number of a load configuration
m_k	Load configuration at sequence number k
ξ	Number of measurement load configurations
c	Number of distinct load configurations
V	Voltage phasor measured at feeder-head
S, P, Q	Power measured at feeder-head
V_i	Voltage phasor at bus i
S_i^l, P_i^l, Q_i^l	Power of the individual load at bus i
n_p, n_q	Load model voltage exponents
Z_i	Impedance of line i
SW_i	Binary switching status of the load at bus i
B	Bus-injection to branch-flow matrix
e_i	Canonical basis of $\mathbb{R}^n$
r	Residue vector for function $F(\cdot)$
$\chi^2_{v,\alpha}$	Chi-square distribution function with probability α and with v degrees of freedom.

I. INTRODUCTION

Accurate load modeling is necessary for power system operation, control, and planning [1], [2]. Load modeling has applications in demand response [3], distributed energy resources (DERs) management [4], Volt-VAR control [5], voltage stability [6], and optimal power flow [7]. The existing load modeling methods in the power systems literature can be classified into two main categories [2]: *component-based* methods and *measurement-based* methods. Component-based methods make use of a-priori information or statistical assumptions on customer appliances and other load devices. In contrast, measurement-based methods use field measurements to learn and update the load model parameters in real-time. A recent CIGRE report in [2] has found that the majority of the utilities that were surveyed across 50 countries use measurement-based methods to estimate the parameters of their load models.

Measurement-based load modeling methods can be further classified into three broad categories: static load modeling, c.f., [7]–[12], composite load modeling, which is a combination of static load modeling and dynamic load modeling, c.f., [13]–[22], and component-based load modeling, c.f., [23]–[26]. Our focus in this study is on the first group; which is commonly used to estimate the parameters of the ZIP load model.

When it comes to modeling loads at distribution-level and based on measurements at feeder-head, the common approach in the literature is to obtain a ZIP model for the *aggregate* load of the distribution feeder. A feeder-aggregated load model provides sufficient information to conduct most practical upper-level analysis at sub-transmission and transmission systems. However, there is still a gap in the literature to model the individual loads at each medium to low voltage load transformer. We seek to address this open problem in this paper.

Our focus here is *not* on the trivial case where a measurement network such as a network of smart meters is available across the distribution feeder; because in that case the individual loads are monitored rather directly. We are instead interested in achieving individual load models by making use of limited measurements, mainly those at the feeder-head at the substation; thus to support the many utilities that are not yet equipped with a complete network of smart meters.

A. Related Work

There exists a rich literature on measurement-based load modeling methods that, similar to this paper, make use of the measurements at feeder-head to estimate the parameters of the load model. In [7], an approximate representation of the

ZIP model is proposed using semidefinite programming (SDP) relaxation of the optimal power flow (OPF) problem. In [8], a multi-state load model is developed for distribution system analysis. In [9], a measurement-based method is developed to estimate polynomial as well as exponential load models. In [10], [11], a measurement-based load model is developed for voltage stability analysis. In [12]–[14], a measurement-based feeder-aggregated load model is developed by using data from phasor measurements units (PMUs). In [15], a measurement-based load model is obtained using a multi-curve identification technique. In [16], measurement-based load modeling is done based on sensitivity analysis. In [17], a measurement-based per-phase load model is developed under unbalanced disturbances. Also, different methods are proposed in [18]–[22] to estimate composite load model parameters, while the static terms are mostly considered as feeder-head aggregate load model.

Despite the differences in methodologies and applications, the above studies on feeder-aggregated load models make use of a fundamental but similar Circuit Theory concept to analyze the feeder-head measurements during the *voltage events* that occur at the *up-stream* of the understudy distribution feeder at the sub-transmission or transmission networks, c.f. [7]–[22].

However, there are three limitations in measurements-based methods that rely on up-stream voltage events as the main enabler for load modeling. First, major up-stream voltage events may not occur frequently; therefore, online load modeling may not be possible for several hours until one such event occurs. Second, at every occurrence of an up-stream voltage event, there is practically a different combination of the individual loads across the distribution feeder that are switched on; therefore, the measurements obtained from different up-stream voltage events represent different snapshots of the understudy distribution feeder; thus, such measurements may not be directly comparable even for the purpose of modeling the aggregate feeder load. Third, the aggregate load models are often not useful to conduct distribution-level analysis, such as to study the impact of distributed energy resources or issues related to power quality across distribution feeders. An individual load model is needed in these cases.

B. Summary of Technical Contributions

This paper proposes a novel method for individual load modeling in power distribution systems. The main technical contributions in this paper can be summarized as follows:

1) Instead of or in addition to using up-stream voltage events as the enabler for load modeling, which is commonly used in the literature such as in [7]–[12], we make use of load switching events across the distribution feeder itself. In principle, once a load is switched, the switching event changes the voltage in the rest of the loads, which causes variation in their active power and reactive power usage; thus allowing us to estimate load parameters of the rest of the individual loads.

2) The proposed method can estimate load modeling parameters of individual loads using measurements only
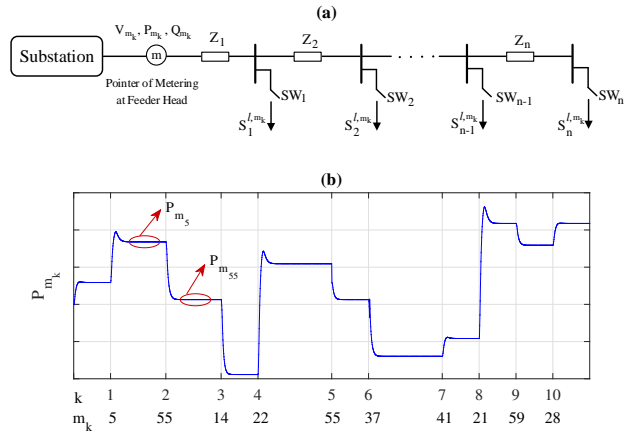

Fig. 1. An example distribution feeder: (a) the single line diagram of the feeder; (b) the total load that is measured at the feeder head.

at the feeder-head, i.e., at the distribution substation. No measurement is needed at individual loads.

3) We provide a theoretical foundation to determine the conditions on the extent of measurements needed to successfully achieve the individual load models.

4) The proposed individual load modeling can be done *sequentially* to obtain the load models for a subset of loads as more measurements become gradually available. Moreover, we proposed a variation of our method by using a forgetting factor so as to support estimating the parameters for time-varying individual loads.

5) By solving a non-liner least-squares problem, it is shown that the proposed methodology can be extended to utilize different types of *redundancy* in measurements in order to improve load modeling accuracy and robustness.

6) We develop a residue-based bad data detection and identification method to identify and drop load configuration measurements with erroneous switching status; thus to ensure the accuracy of the load models. This also helps with the cases with imperfect knowledge of line impedances, switch statuses, and other system parameters.

II. PROBLEM STATEMENT

Consider a distribution feeder with $n \geq 2$ buses. For now, and for the simplicity of discussion, suppose the distribution network does not have any lateral, as shown in Fig. 1(a). The case for distribution feeders with laterals is discussed in Section IV-C. Depending on which individual loads are turned on and which individual loads are turned off, there can be a total of $2^n - 1$ possible load configurations, excluding the no load situation. As time goes by, a variety of load configurations occur, changing the voltage and power that are measured at the feeder-head. The measuring can be done, for example, by using distribution-level phasor measurement unit (D-PMU), a.k.a., Micro-PMU, [27]–[33]. The measurements corresponding to a total of 10 load configurations are shown in Fig. 1(b), indexed as $k = 1, \ldots, 10$. The corresponding load configurations are denoted by $m_1, \ldots, m_{10}$, which take numbers between 1 to $2^6 - 1 = 63$; because $n = 6$ for

the network in this example. Each individual load can be an arbitrary combination of constant impedance, constant current, and constant power load components. The complex power that is drawn by the load at bus i under load configuration m_k is denoted by S_i^{l,m_k}. In this paper, we seek to answer the following question: *Can we model each of the n individual loads in Fig. 1(a) by studying a sequence of measurements at the feeder-head in Fig. 1(b)?*

Of course, there exist some special cases for which the above problem is somewhat trivial. For example, if at a load configuration, there is only a single load that is turned on and all other loads are turned off, then modeling that single load is relatively easy, because what is being monitored at the feeder head is the single load itself plus the power loss on distribution lines. However, beyond such relatively trivial special cases, answering the above question can be challenging.

III. Load Modeling Method

A. The First Set of Equations: Circuit Model

For a given k, consider the measurements that are obtained at the feeder-head during load configuration m_k. According to the law of complex power conservation, we have:

$$S_{m_k} = \sum_{i=1}^{n} \left(S_i^{l,m_k} SW_i^{m_k} \right) + \sum_{j=1}^{n} Z_j \left| \sum_{d=j}^{n} \left(\frac{S_d^{l,m_k}}{V_d^{m_k}} \right)^* SW_d^{m_k} \right|^2, \quad (1)$$

where the first term is the total load, and the second term is the total loss. Binary variable $SW_i^{m_k}$ is one if the individual load i is turned on during load configuration m_k; and zero otherwise. Note that, we could define a notation for current and replace (1) with an equation based on the Kirchhoff's Current Law (KCL). However, in this paper, we present the circuit model only in terms of complex power and voltage phasor.

Next, we write the Kirchhoff's Voltage Law (KVL) for the circuit in Fig. 1(a). Any *loop* can be used for this purpose. In fact, there are $n(n+1)/2$ loops in this circuit that can be used to write the KVL equations. However, as proved by using Graph Theory in [34], [35], one can write only n *independent* KVL equations in this circuit. This can be achieved by writing the KVL equations for the n loops between the substation and every node $i = 1, \cdots, n$, as shown below:

$$V_i^{m_k} = V_{m_k} - \sum_{j=1}^{i} Z_j \left(\sum_{d=j}^{n} \left(\frac{S_d^{l,m_k}}{V_d^{m_k}} \right)^* SW_d^{m_k} \right). \quad (2)$$

Together, the $n+1$ independent complex nonlinear equations in (1) and (2) provide the model for the circuit. The complex power S_{m_k} and the voltage phasor V_{m_k}, which are measured at the feeder-head, as well as the lines impedances Z_j for $j = 1, \ldots, i$, are the *known* parameters in these equations. In contrast, the individual load complex power S_i^{l,m_k} and the individual node voltage phasor $V_i^{m_k}$ are the *unknowns* to be determined for all $i = 1, \ldots n$. For any load configuration m_k, the number of unknowns can be counted as:

$$n + \sum_{i=1}^{n} SW_i^{m_k}, \quad (3)$$

where the first term counts $V_i^{m_k}$ for every node $i = 1, \ldots, n$; and the second term counts S_i^{l,m_k} for every load i that is turned on under load configuration m_k, for which $SW_i^{m_k} = 1$.

For every load configuration m_k, the system of non-linear equations in (1) and (2) can have a solution only if one single load is switched on, in which case the summation in the second term in (3) is one. In all other cases, the system of nonlinear equations in (1) and (2) is *under-determined*.

Suppose more load configurations occur as time goes by. This will provide new equations and also introduce new unknowns. Suppose the measurements are available for c distinct load configurations. The following upper bound always holds:

$$c \leq \min\{\xi, 2^n - 1\}. \quad (4)$$

In Fig. 1(b), we have $\xi = 10$ but $c = 9$; because load configurations m_2 and m_5 are the same.

Given the measurements at the feeder-head for c distinct load configurations, the number of unknowns becomes:

$$\text{Number of Unknowns} = c \times n + \sum_{k=1}^{c} \sum_{i=1}^{n} SW_i^{m_k}, \quad (5)$$

and the number of independent equations increases to:

$$\text{Number of Equations} = c \times (n+1). \quad (6)$$

The first term in (5) counts voltage phasor $V_i^{m_k}$ in every node $i = 1, \ldots, n$ and each of the c distinct load configurations $k = 1, \ldots, c$. The second term in (5) counts complex power S_i^{l,m_k} for every load $i = 1, \ldots, n$ that is turned on under load configuration m_k for each of the c distinct load configurations.

The system of non-linear equations in (1) and (2) for c distinct load configurations can have a solution *only if* a single load is switched on in every load configuration k. This occurs *only if* the following inequality holds:

$$\sum_{k=1}^{c} \sum_{i=1}^{n} SW_i^{m_k} \leq c. \quad (7)$$

In that case, the total number of unknowns in (1) and (2) would be less than or equal to the total number of independent equations in (1) and (2). In all other cases, we lack sufficient independent equations. This can potentially be resolved by deriving new equations from load models, as we discuss next.

B. The Second Set of Equations: Load Model

Any load with any combination of active and reactive power consumption can be expressed in the generic form of a ZIP load [2]. Specifically, for any two distinct load configurations m_k and m_h, the complex power consumption at the individual load at bus i can be modeled as:

$$S_i^{l,m_k} = P_i^{l,m_h} \left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|} \right)^{n_{p_i}} + j Q_i^{l,m_h} \left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|} \right)^{n_{q_i}}. \quad (8)$$

If $n_p = 0, 1, 2$, then the active power component of the load is constant power, constant current, and constant impedance, respectively. For all other values of n_p, the active power component is a combination of these three load elements. The reactive power component can be defined similarly using n_q.

Note that, for each individual load i, one can construct the load model in (8) only if there *do* exist at least two load configurations m_k and m_h in which load i is turned on. In other words, we cannot introduce the load model in (8) for individual load i unless the following inequality holds:

$$\sum_{k=1}^{c} SW_i^{m_k} \geq 2. \tag{9}$$

Thus, unless we state otherwise, we assume that the above condition holds for all loads $i = 1, \ldots, n$. The special case when (9) does *not* hold will be discussed in Section IV-B.

The load models of the form in (8) can provide additional equations that can be combined with the equations in (1) and (2) to obtain the unknowns that we identified in (5). However, the load equations in (8) also introduce new unknowns, because n_{p_i} and n_{q_i} for buses $i = 1, \ldots, n$, are not known. Therefore, the number of new unknowns becomes:

$$\text{Number of Unknowns} = n. \tag{10}$$

Here we count $n_p + j n_q$ as one unknown complex number.

Next, we need to identify how many of the new equations in the form of (8) are *independent*; and thus useful for identifying the unknown individual loads in our system. This matter is addressed in a Theorem as presented below.

Theorem 1. *Given c distinct load configurations, the number of independent complex equations in the form of (8) is:*

$$\text{Number of Equations} = \sum_{i=1}^{n}\sum_{k=1}^{c} SW_i^{m_k} - n. \tag{11}$$

Proof: For each individual load i, let us define:

$$c_i = \sum_{k=1}^{c} SW_i^{m_k}. \tag{12}$$

One can write a load model equation of the form in (8) for *any two distinct* load configurations m_k and m_h. Therefore, the total number of load model equations of the form in (8) that we can write for each individual load i is:

$$\binom{c_i}{2} = \frac{1}{2}\, c_i \times (c_i - 1). \tag{13}$$

However, these many equations are *not* independent; because the expression in (8) preserves *transitive relation* [36]. For instance, for a given individual load i, suppose we have $c_i = 3$. That is, suppose load i is turned on in three load configurations m_k, m_h, and m_g. From (8), we have:

$$P_i^{l,m_k}+jQ_i^{l,m_k}= P_i^{l,m_h}\left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|}\right)^{n_{p_i}}+ jQ_i^{l,m_h}\left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|}\right)^{n_{q_i}}, \tag{14}$$

$$P_i^{l,m_k}+jQ_i^{l,m_k}= P_i^{l,m_g}\left(\frac{|V_i^{m_k}|}{|V_i^{m_g}|}\right)^{n_{p_i}}+ jQ_i^{l,m_g}\left(\frac{|V_i^{m_k}|}{|V_i^{m_g}|}\right)^{n_{q_i}}, \tag{15}$$

$$P_i^{l,m_h}+jQ_i^{l,m_h}= P_i^{l,m_g}\left(\frac{|V_i^{m_h}|}{|V_i^{m_g}|}\right)^{n_{p_i}}+ jQ_i^{l,m_g}\left(\frac{|V_i^{m_h}|}{|V_i^{m_g}|}\right)^{n_{q_i}}, \tag{16}$$

However, the above equations are *not* independent due to their transitive relationship. Any one of the three equations in (14)-(16) can be obtained by applying proper non-linear operators to the other two equations. For instance, we can obtain (16)

by *dividing* the real and the imaginary parts of (14) by the real and the imaginary parts of (15), respectively, as shown below:

$$\frac{P_i^{l,m_k}}{P_i^{l,m_k}} = \frac{P_i^{l,m_h}}{P_i^{l,m_g}} \times \left[\left(\frac{|V_i^{m_k}|}{|V_i^{m_h}|}\right)^{n_{p_i}}\!\!\Big/ \left(\frac{|V_i^{m_k}|}{|V_i^{m_g}|}\right)^{n_{p_i}}\right]$$
$$\Rightarrow\quad P_i^{l,m_h} = P_i^{l,m_g}\left(\frac{|V_i^{m_h}|}{|V_i^{m_g}|}\right)^{n_{p_i}}. \tag{17}$$

The calculation for the imaginary part is similar and omitted for brevity. The above transitive relationship for the system of nonlinear equations in (8) is the generalization of the concept of linear dependency in system of linear equations.

In order to obtain the largest subset of independent load equations from the total number of equations in (13), first, for each individual load i, we consider one of the load configurations as the *reference configuration*. Next, we write $c_i - 1$ different load equations of the form in (8) based on the reference configuration together with each of the $c_i - 1$ remaining load configurations. All such $c_i - 1$ load configurations are, by-construction, independent. Once we repeat this procedure for all n loads, we can obtain several independent load equations, at a total number equal to the one in (11). ∎

C. Combining Circuit Model and Load Model

Suppose we have collected the measurements at the feeder head under c distinct load configurations. Also, suppose we combine the Circuit Model in Section III-A and the Load Model in Section III-B. We propose Algorithm 1 to obtain the individual loads model. This algorithm can be used as long as enough load configuration measurements are available. This raises the following question: What is the smallest c, denoted by $c_{\min}$, in order to solve the system of nonlinear equations in (6) and (11) to obtain the set of unknowns in (5) and (10)? This question is answered in a Theorem as presented below.

Theorem 2. *Suppose the inequality in (9) holds for all loads $i = 1, \ldots, n$, where $n > 2$. The minimum number of distinct load configurations in order to solve the system of equations in (6) and (11) to obtain the unknowns in (5) and (10) is:*

$$c_{\min} = 2\, n. \tag{18}$$

Proof: We need the number of unknowns to be less than or equal to the number of independent equations. From (6), (11), (5), and (10), this can be expressed as the following inequality:

$$n \times c + \sum_{i=1}^{n}\sum_{k=1}^{c} SW_i^{m_k} + n$$
$$\leq c \times (n+1) + \sum_{i=1}^{n}\sum_{k=1}^{c} SW_i^{m_k} - n. \tag{19}$$

Once we cancel out the common terms on both sides and reorder the rest of the terms, we can express the above inequality as $c \geq 2n$. Therefore, $c_{\min} = 2n$. ∎

Recall from Section II that our goal is to model each of the n individual loads in Fig. 1(a) by studying the sequence of measurements in Fig. 1(b). Theorem 2 indicates how far in the sequence of measurements we must go before we can obtain the load models. From Theorem 2, any arbitrary but

Algorithm 1 Individual Load Modeling

  Step 1: Obtain the Circuit Model
  for each load configuration m_k **do**
    Apply law of complex power conservation $\rightarrow$ (1)
    Apply KVL between substation and every node $i \rightarrow$ (2)
  end for
  Step 2: Obtain the Load Model
  for every node i **do**
    Calculate complex power consumption $\rightarrow$ (8)
  end for
  Step 3: Solve the system of equations in (1), (2), and (8)
  return Solutions

distinct $2n$ load configurations that satisfy (9) for all n loads can be used in Algorithm 1 to obtain the unknowns and model the loads.

There are several algorithms available in the literature to solve a system of non-linear equations [37], [38]. In this study, we use the Levenberg-Marquardt algorithm, which is iterative and commonly used in curve-fitting problems [38]. The initial guess for all unknown parameters in both circuit model and load model are set to 1 per-unit.

Note that, the unknowns that are of interest in this paper are n_{p_i} and n_{q_i}. The other unknowns, i.e., $V_i^{m_k}$ and S_i^{l,m_k}, act as auxiliary variables to help us identify n_{p_i} and n_{q_i}.

IV. REMARKS AND EXTENSIONS

A. Redundant Load Configurations

Recall from Section II that there exist $2^n - 1$ possible distinct load configurations for an n-bus system. From Theorem 2, as few as $2n$ of them is sufficient to solve the load modeling problem. But what if we continue collecting new distinct load configurations beyond $c_{\min}$? What can we do with the remaining $2^n - 1 - 2n$ distinct load configurations? Furthermore, is there any benefit to also look into the duplicate load configuration events? For example, recall that load configuration m_5 was the duplicate of load configuration m_2 in Fig. 1(b). Finally, can we also make use of the upstream voltage events? Note that, so far, we did not use such measurements.

To address the above questions, we collectively refer to the following items as redundant load configurations:

- Any additional distinct load configuration for $c > c_{\min}$;
- Any duplicate of an existing load configuration;
- Any major upstream voltage event.

We are interested in using the above redundant configurations to enhance load modeling accuracy in presence of errors in measurements. Note that, the above redundant load configurations introduce new unknowns in the Circuit Model, but they do *not* change the number of unknowns in the Load Model.

It is not difficult to construct the new equations and identify the new unknowns similar to (6)(11) and (5)(10). The details are omitted due to space limitation. Let $F(V_{m_k}, S_{m_k}, V_i^{m_k}, S_i^{l,m_k}, n_s) = 0$ denote the resulting system of equations. We can obtain the individual load parameters by solving the following non-linear least-squares problem:

$$\underset{V_i^{m_k}, S_i^{l,m_k}, n_s}{\text{minimize}} \left\| F(V_{m_k}, S_{m_k}, V_i^{m_k}, S_i^{l,m_k}, n_s) \right\|_2. \quad (20)$$

Again we can use a variation of the Levenberg-Marquardt algorithm, called the damped least-squares algorithm, to solve the problem in (20). As in Section III-C, our ultimate goal here is to obtain the values of n_{p_i} and n_{q_i} for all n loads. The difference compared to Section III-C is that we now have more equations and more auxiliary variables. Of course, if the redundant equations are removed and if there is no error in measurements, then solving problem (20) reduces to solving a system of nonlinear equations, just like in Section III-C.

B. Sequential Load Modeling

What if the inequality in (9) does *not* hold for all n individual loads? Let us define:

$$\hat{n} = \sum_{i=1}^{n} \mathbb{I} \left(\sum_{k=1}^{c} SW_i^{m_k} \geq 2 \right), \quad (21)$$

where $\mathbb{I}(\cdot)$ is a 0-1 indicator function. Here, $\hat{n} \leq n$ denotes the number of individual loads that *do* satisfy the inequality in (9). Let us also define $\hat{c}$ as a subset of c distinct load configurations in which all individual loads that are turned on do satisfy the inequality in (9). Of course, we have $\hat{c} \leq c$. We can now apply Theorem 2 to $\hat{n}$ and $\hat{c}$ and similarly achieve $\hat{c}_{\min} = 2\hat{n}$.

The above analysis can lead to developing a sequential load modeling approach, as shown in Algorithm 2. As time goes by, a variety of load configurations occur. Our algorithm must wait until such time that there exists an $\hat{n} > 2$ and $\hat{c} = 2\hat{n}$. It can then obtain the model for the $\hat{n}$ loads. As measurements from more load configurations become available, Algorithm 2 gradually and sequentially models all individual loads across the feeder.

Algorithm 2 Sequential Load Modeling

  while there is a new load configuration measurement **do**
    Consider $\hat{c}$ load configurations measured so far.
    Set $\hat{n}$ according to (21).
    if Theorem 2 is satisfied for $\hat{c}$ and $\hat{n}$ **then**
      Run Algorithm 1 for $\hat{c}$ and $\hat{n}$
      return Solutions
    end if
  end while

C. Distribution Feeder with Laterals

The feeder in Fig. 1(a) does not have any lateral. However, in practice, most distribution feeders do have laterals. In order to incorporate laterals in our model, we need to revise and replace the Circuit Model in (1) and (2) with

$$S_{m_k} = \sum_{i=1}^{n} \left(S_i^{l,m_k} SW_i^{m_k} \right) + \sum_{j=1}^{n} Z_j \left| \sum_{d=j}^{n} \left(\frac{S_d^{l,m_k}}{V_d^{m_k}} \right)^* SW_d^{m_k} e_j^T B e_d \right|^2, \quad (22)$$

$$V_i^{m_k} = V_{m_k} - \sum_{j=1}^{i} Z_j e_j^T B e_i \left(\sum_{d=j}^{n} \left(\frac{S_d^{l,m_k}}{V_d^{m_k}} \right)^* SW_d^{m_k} e_j^T B e_d \right), \quad (23)$$

where B is the *bus-injection to branch-flow matrix* [39], e_i is the canonical basis of $\mathbb{R}^n$, and $e_i^T B e_j$ is 1 if the power injection of bus j flows on branch i, and zero otherwise. An example to construct matrix B is given in Section V-A.

D. Imperfect Knowledge of Load Switching Status

So far, we have assumed that the switching status $SW_i^{m_k}$ of the individual loads is already known in each measured load configuration m_k. This can be achieved by using the existing methods that are designed to identify load switching on distribution networks. For example, in [30], the location of each load switching event along the distribution feeder is identified by applying the Compensation Theorem from Circuit Theory to the measurements that are available at the feeder-head and at the feeder-tail. No measurement is needed at individual loads.

We *do* recognize that, in practice, the methods such as the one in [30] are *not* entirely precise. As a result, our knowledge of the load switching status is imperfect and may carry errors. However, this issue can be resolved by the use of the redundant load configurations that we proposed in Section IV-A. Specifically, the use of redundant load configurations allows us to conduct *bad data detection* in order to identify and drop load configurations with erroneous switching status.

Let r denote the *residue vector* for function $F(\cdot)$ in the least-squares problem in (20). Using the Chi-squares test [40], we detect erroneous load configurations if:

$$\mathbb{I}\left(r^T r > \chi^2_{v,\alpha}\right) = 1, \tag{24}$$

where $\chi^2_{v,\alpha}$ is the Chi-square distribution function corresponding to a detection confidence level with probability α, and with $v = \xi - 2n$ degrees of freedom in solving problem (20) due to having $\xi - 2n$ redundant equations. Once the existence of an erroneous switch status is detected using the Chi-square test, we next apply the largest normalized residual (LNR) test to identify such erroneous load configurations and drop them from the analysis to assure accurate and robust load modeling.

E. Time-Varying Load Modeling

In practice, the parameters of the load models may change over time. In order to track time-varying load modeling parameters, we propose to apply *forgetting factors* to load configurations as well as using the Chi-squares test and the LNR test as in Section IV-D. Specifically, suppose the measurements corresponding to c load configurations are available, where $c > c_{\min}$ and $v = \xi - 2n$ is the number of redundant configurations. We define the forgetting factors as follows:

$$\beta(k) = \begin{cases} \lambda^{v-k+1} & k = 1, \ldots, v \\ 1 & k = v+1, \ldots, 2n \end{cases} \tag{25}$$

where $0 < \lambda \leq 1$. We use $\beta(k)$ as the *weight* corresponding to equation number k in the non-linear least-squares problem in (20). This allows us to estimate the load model parameters based on the most updated load configurations. Similar to the bad data detection method that we used in Section IV-D to detect error in load switching status, here, we use
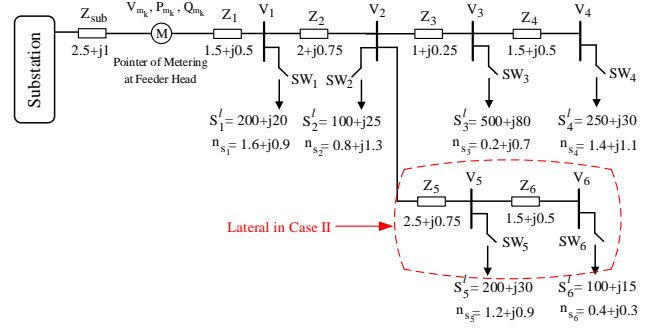

Fig. 2. Single-line diagram of the feeder in the two basic test cases in Section V-A. Case I does not include the lateral. Case II does include the lateral.

the Chi-squares test and the LNR test in order to detect, identify, and drop load configurations that impose error to the aforementioned weighted non-linear least-square problem. As a result, whenever a load changes, the first few subsequent load configuration measurements are dropped until the available load configuration measurements are sufficient to achieve an accurate load modeling result with low and consistent residues.

F. The Use of Multiple Sensors

The methodology that is proposed in this paper is intended to allow individual load modeling by making use of measurements from as few as only a single sensor that is installed at the feeder head. Furthermore, in principle, the proposed methodology is applicable to any distribution feeder with an arbitrary number of nodes and laterals. Nevertheless, as the number of nodes increases, one may start facing numerical issues in order to solve the formulated nonlinear system of equations due to its size. Therefore, when it comes to long distribution feeders with a large number of nodes, such as the case of the IEEE 123-bus test system in Section V-I, one needs to start installing additional sensors. This can be done by simply installing the additional sensors at the head of the long laterals. Importantly, the methodology that is needed to solve the load modeling problem remains the same. On one hand, each lateral whose aggregate load is measured at its head, can be seen as an independent load modeling problem, whose individual loads can be modeled using Algorithms 1 and 2. On the other hand, the measured aggregate load of such laterals can be used as known parameters when it comes to solving the combined circuit models and load models for the *rest* of the feeder, again using Algorithms 1 and 2. This essentially breaks down the load modeling problem into smaller problems of the same type.

V. Case Study and Results

All the case studies in this section are done in PSCAD [41] to construct the distribution feeder and all load configurations.

A. Basic Results With and Without Laterals

This section investigates the effectiveness of the proposed method based on using the sufficient number of load configurations as presented in Theorem 2. To do so, we apply the proposed method to two cases as defined in Fig. 2:

TABLE I
LOAD CONFIGURATIONS FOR CASE I IN SECTION V-A.

Configuration	SW_1	SW_2	SW_3	SW_4	Time
m_1	1	0	0	0	$[0\,,\,t_1]$
m_2	0	1	1	0	$[t_1\,,\,t_2]$
m_3	0	1	0	1	$[t_2\,,\,t_3]$
m_4	0	0	1	1	$[t_3\,,\,t_4]$
m_5	1	1	1	0	$[t_4\,,\,t_5]$
m_6	1	1	0	1	$[t_5\,,\,t_6]$
m_7	1	0	1	1	$[t_6\,,\,t_7]$
m_8	1	1	1	1	$[t_7\,,\,t_8]$

TABLE II
LOAD CONFIGURATIONS FOR CASE II IN SECTION V-A.

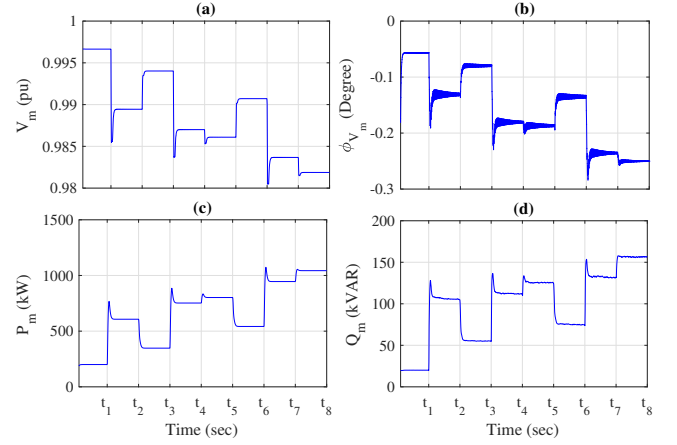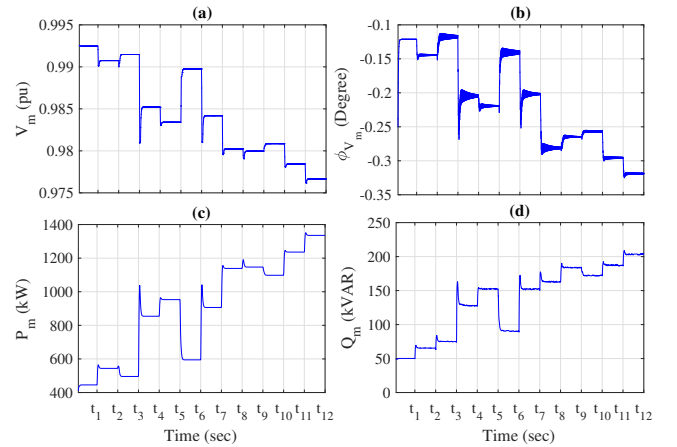| Configuration	SW_1	SW_2	SW_3	SW_4	SW_5	SW_6	Time
m_1	1	0	0	1	0	0	$[0\,,\,t_1]$
m_2	1	0	0	1	0	1	$[t_1\,,\,t_2]$
m_3	1	1	0	0	1	0	$[t_2\,,\,t_3]$
m_4	0	0	1	1	0	1	$[t_3\,,\,t_4]$
m_5	0	1	1	1	0	1	$[t_4\,,\,t_5]$
m_6	1	1	0	0	1	1	$[t_5\,,\,t_6]$
m_7	0	1	1	0	1	1	$[t_6\,,\,t_7]$
m_8	1	0	1	1	1	0	$[t_7\,,\,t_8]$
m_9	0	1	1	1	1	1	$[t_8\,,\,t_9]$
m_{10}	1	1	1	0	1	1	$[t_9\,,\,t_{10}]$
m_{11}	1	1	1	1	1	0	$[t_{10}\,,\,t_{11}]$
m_{12}	1	1	1	1	1	1	$[t_{11}\,,\,t_{12}]$

Fig. 3. Feeder-head measurement parameters for Case I: (a) voltage magnitude; (b) voltage angle; (c) active power; (d) reactive power.

Fig. 4. Feeder-head measurement parameters for Case II: (a) voltage magnitude; (b) voltage angle; (c) active power; (d) reactive power.

- Case I: Feeder *without* no lateral;
- Case II: Feeder *with* lateral.

Both cases are shown in Fig. 2. The load model parameters are considered based on the experimental studies in [42], [43]. For example, the load in bus 1 is a combination of residential appliances, including coffee maker, LED and tungsten lights, and LCD television, as described in [42]. Recall from Section II that there are $2^4 - 1$ and $2^6 - 1$ possible distinct load configurations in Case I and Case II, respectively. From Theorem 2, the individual load models for Case I and Case II can be obtained from any 8 and 12 distinct load configurations, that hold (9), respectively. The examples are given in Tables I and II, respectively. The resulting feeder-head measurements are shown in Figs. 3 and 4 for Case I and Case II, respectively.

For Case I, the Circuit Model includes 40 equations and 52 unknowns, see (5) and (6). The Load Model adds 16 additional equations and 4 new unknowns, see (10) and (11). Thus, the combined system of nonlinear equations has 56 equations and 56 unknowns. For Case II, the total number of equations is 126 and the total number of unknowns is also 126. Matrix B for Case II in Fig. 2 is obtained as:

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{26}$$

Figs. 5(a) and (b) shows the true and the estimated individual load models for Case I and Case II, respectively. We can see that the proposed method works well on both distribution

feeders with or without laterals. The error in individual load modeling is less than 0.01% for both Case I and Case II.

Figs. 6(a) and (b) show the true and estimated voltage phasors for Case I and Case II, respectively. The proposed method can also estimate auxiliary variables by 0.01% error. Similar results are obtained in estimating active and reactive power consumption, that are omitted due to brevity.

B. Performance Comparison

Before we get into the details in this section, we must emphasize that, to the best of our knowledge, this paper is the first study that provides *individual* load model parameters using feeder-head measurements. Therefore, we *cannot* compare the individual load modeling method in this paper with another individual load modeling method. But what we *can* do, is to aggregate the individual load models that we obtain here, and then compare the result with the result of applying the feeder-aggregated load modeling methods in the literature, such as in [7]–[12]. The test setup to conduct such performance comparison is similar to Case II in Section V-A. A total of 15 load configurations are considered, as given in Table III, out of which three cases are caused by upstream voltage events. They are denoted by m_3, m_7, and m_{12}. The remaining 12 cases are distinct load configurations.
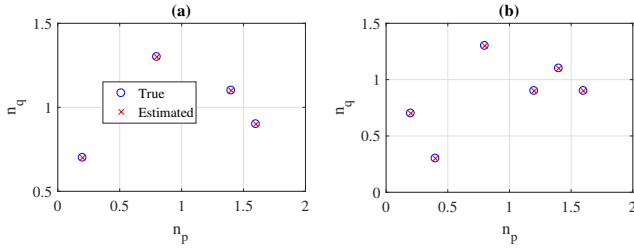
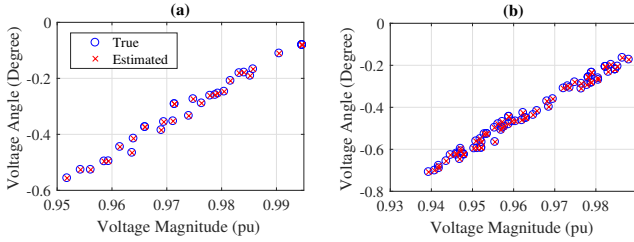Fig. 5. True and estimated load model for: (a) Case I; (b) Case II.

Fig. 6. True and estimated voltage phasor for: (a) Case I; (b) Case II.

The results are shown in Figs. 7(a) and (b) with respect to the feeder-aggregated load parameters n_p and n_q, respectively. Here, comparison is done with the method in [9]. We can make three key observations. First, one can use the feeder-aggregated load modeling method in [9] no sooner than time t_2, because time t_2 is the first time in this case study that an upstream voltage event occurs. In contrast, the method that is proposed in this paper can be used as early as time t_1, i.e., even before an upstream voltage event occurs. Second, as time goes by, the feeder-aggregated model using the method in [9] is not updated, except at time t_6 and time t_{11}, i.e., when the second and the third upstream voltage events occur, respectively. In contrast, the method that is proposed in this paper is updated much more frequently. Finally, and most importantly, we can see that every time that the feeder-aggregated model is updated using the methods in [9], it matches the aggregated model that is obtained based on the method in this paper.

All in all, we can conclude that the method in this paper is equally good compared to the methods in the literature as far as feeder-aggregated load modeling is concerned whenever an upstream voltage event occurs; with an added advantage that it can update such feeder-aggregated model more frequently. Of course, our method also provides individual load models, which is the primary goal of this paper.

C. Tackling Errors in Measurements

In practice, sensors, such as micro-PMUs, may *not* be precise, mostly due to errors in CTs and PTs. Several methods have been introduced in the literature to improve data quality in synchrophasors, c.f., [44], [45]. In this section, we assume that the measurments for load modeling applications still include some levels of error. Monte Carlo method is used to generate different scenarios for different levels of error in feeder-head measurements. The test setup is similar to Case II in Section V-A. However, this time, we also consider 12 redundant load configurations of all three types, as in Table IV. These scenarios are *in addition* to the 12 load configurations in Table II. The upstream events are 1.05 per unit step-up in feeder-head caused by an upstream voltage regulator. The

TABLE III
LOAD CONFIGURATIONS FOR CASE II IN SECTION V-B.

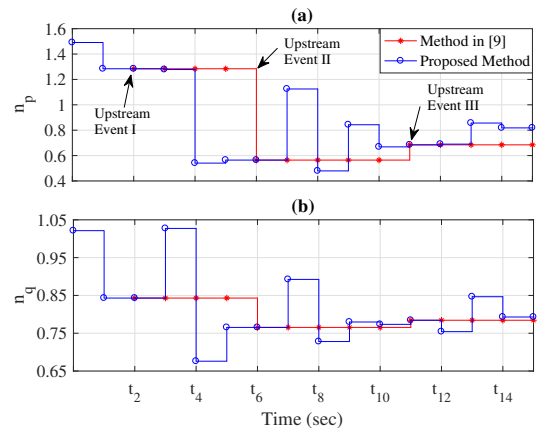| Configuration	V(pu)	SW_1	SW_2	SW_3	SW_4	SW_5	SW_6	Time
m_1	1.00	1	0	0	1	0	0	$[0 , t_1]$
m_2	1.00	1	0	0	1	0	1	$[t_1 , t_2]$
m_3	0.95	1	0	0	1	0	1	$[t_2 , t_3]$
m_4	0.95	1	1	0	0	1	0	$[t_3 , t_4]$
m_5	0.95	0	0	1	1	0	1	$[t_4 , t_5]$
m_6	0.95	0	1	1	1	0	1	$[t_5 , t_6]$
m_7	1.00	0	1	1	1	0	1	$[t_6 , t_7]$
m_8	1.00	1	1	0	0	1	1	$[t_7 , t_8]$
m_9	1.00	0	1	1	0	1	1	$[t_8 , t_9]$
m_{10}	1.00	1	0	1	1	1	0	$[t_9 , t_{10}]$
m_{11}	1.00	0	1	1	1	1	1	$[t_{10} , t_{11}]$
m_{12}	1.05	0	1	1	1	1	1	$[t_{11} , t_{12}]$
m_{13}	1.05	1	1	1	0	1	1	$[t_{12} , t_{13}]$
m_{14}	1.05	1	1	1	1	1	0	$[t_{13} , t_{14}]$
m_{15}	1.05	1	1	1	1	1	1	$[t_{14} , t_{15}]$

Fig. 7. Performance comparison with [9] in terms of feeder-aggregated load modeling: (a) load model voltage exponents for active power; (b) load model voltage exponents for reactive power.

total number of equations is 258 and the total number of unknowns is 246. Thus, the degrees of freedom in redundant measurements is $v = 258 - 246 = 12$.

Fig. 8 shows the feeder-head measurements for all the 24 load configurations in Tables II and IV. Problem (20) is solved to obtain individual load parameters. Tables V and VI show the load modeling error in percentage considering different error levels in the feeder-head complex power measurement

TABLE IV
REDUNDANT LOAD CONFIGURATIONS FOR CASE STUDY IN SECTION V-C.

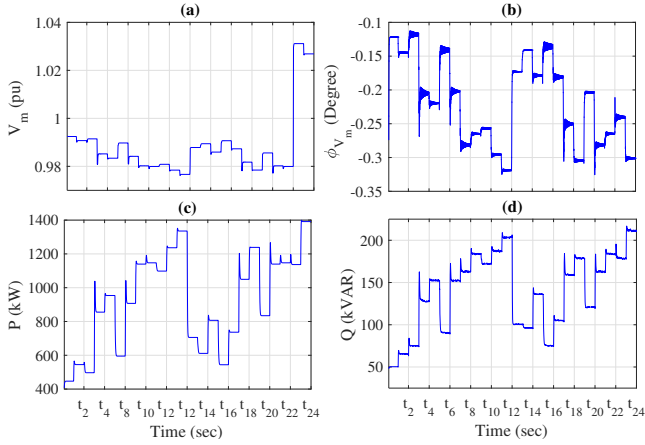| Configuration		SW_1	SW_2	SW_3	SW_4	SW_5	SW_6	Time
m_{13}	Additional	1	0	1	0	0	0	$[t_{12} , t_{13}]$
m_{14}	Additional	0	0	1	0	0	1	$[t_{13} , t_{14}]$
m_{15}	Additional	0	1	1	0	1	0	$[t_{14} , t_{15}]$
m_{16}	Additional	1	1	0	1	0	0	$[t_{15} , t_{16}]$
m_{17}	Additional	1	1	0	1	1	0	$[t_{16} , t_{17}]$
m_{18}	Additional	0	0	1	1	1	1	$[t_{17} , t_{18}]$
m_{19}	Additional	1	0	1	1	1	1	$[t_{18} , t_{19}]$
m_{20}	Additional	1	1	0	1	1	1	$[t_{19} , t_{20}]$
m_{21}	Duplicate	1	0	1	1	1	0	$[t_{20} , t_{21}]$
m_{22}	Duplicate	0	1	1	1	1	1	$[t_{21} , t_{22}]$
m_{23}	Upstream	1	1	1	0	1	1	$[t_{22} , t_{23}]$
m_{24}	Upstream	1	1	1	1	1	1	$[t_{23} , t_{24}]$

Fig. 8. Feeder-head measurements for the test case in Section V-C: (a) voltage magnitude; (b) voltage angle; (c) active power; (d) reactive power.

TABLE V
LOAD MODELING ESTIMATION ERROR IN PERCENTAGE CONSIDERING
ERROR IN COMPLEX POWER MEASUREMENT IN FEEDER-HEAD.

Error	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%
n_p	0.2317	0.4532	0.6446	0.8260	1.0638	1.2589
n_q	0.6991	1.3915	2.1482	2.5407	3.4625	4.0320

and voltage measurement, respectively. We see that, even in the presence of errors in measurements, the results demonstrate an overall satisfactory performance in load model estimation.

D. Impact of Error in Line Impedance

In this section, we investigate how the error in our knowledge about line impedances may affect the performance of the load modeling method. We use the Monte Carlo method to generate different scenarios for different levels of error in line impedances. The test setup is similar to Case II in Section V-A. Similar to Section V-C, we consider redundant load configurations, as in Tables II and IV. Again, the total number of equations is 258 and the total number of unknowns is 246. Accordingly, the degrees of freedom in the redundant measurements is $v = 258 - 246 = 12$. The results are shown in Table VII. We see that, even in the presence of some considerable errors in line impedances, the proposed load modeling algorithm can achieve an overall satisfactory performance in estimating the load modeling parameters.

E. Identifying Erroneous Switch Status

TABLE VI
LOAD MODELING ESTIMATION ERROR IN PERCENTAGE CONSIDERING
ERROR IN VOLTAGE MEASUREMENT IN FEEDER-HEAD.

Error	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%
n_p	0.4157	0.8201	1.2472	1.6862	2.1178	2.5517
n_q	1.3830	2.7107	4.0166	5.4983	6.9226	8.3686

TABLE VII
IMPACT OF ERROR IN OUR KNOWLEDGE ABOUT LINE IMPEDANCES.

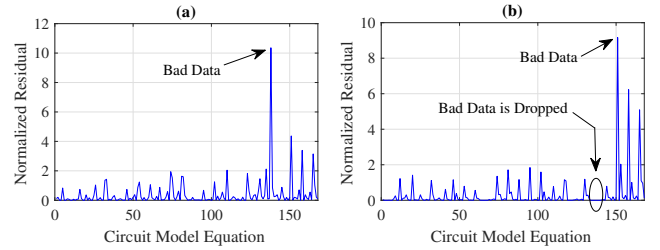| Error in Line Impedance	5%	10%	15%	20%	25%	30%
Error in Estimating n_p	0.09	0.93	1.30	1.98	2.54	3.38
Error in Estimating n_q	0.78	1.95	3.23	5.34	9.16	11.87

Fig. 9. Bad data identification based on the LNR test in two load configurations: (a) the first incorrect switch status data; (b) the second incorrect switch status data.

In this section, we examine the effectiveness of the proposed bad data detection and identification method to remove load configuration measurements with erroneous switch status. We use the Monte Carlo method to generate different scenarios with erroneous switch status for Case II, where the redundant load configurations are as in Tables II and IV. The Chi-squares test in (24) is performed by assuming v and α to be 12 and 0.005, respectively. The bad data detection method is able to detect erroneous switch status with 100% accuracy.

Next, we apply the LNR test to the residues to identify the load configuration measurements with erroneous switch status. We did so only for the residues of the Load Model equations; because the residues of the Circuit Model equations depend only on one load configuration. The load configuration measurements with erroneous switch status are then dropped. Afterwards, the load modeling is done on the remaining load configurations, and the Chi-squares test is applied again. If the Chi-squares test detects more bad data, then LNR test is applied again. This procedure continues until there is no outlier residue.

Fig. 9(a) shows the normalized residues in the presence of erroneous load configurations 20 and 22. Equation number 138 has the largest residue in this case. This equation corresponds to the Circuit Model in load configuration 20. Once the measurements for this load configuration are dropped, the normalized residues appear as in Fig. 9(b). The largest residue corresponds to equation 151, which corresponds to load configuration 22. Once the measurements for this load configuration are also dropped, then the remaining load configuration measurements will result in obtaining accurate load models.

F. Tracking Time-Varying Load Models

In this section, we examine the effectiveness of the proposed method in Section IV-E to track time-varying load modeling parameters. The test setup is similar to Case II. Load modeling is done based on the measurements from the most recent $c = 18$ load configurations. The degree of freedom in redundant measurements is $v = 6$. In total, suppose $k = 1, \ldots, 42$ load configurations occur. At load configuration number 19, the parameter of the load at bus 2 changes from $n_{p_2} = 0.8$ to $n_{p_2} = 1.8$. Fig. 10(a) shows the number of load configurations at each measurement point. Color black denotes the number of load configurations that occurred *before* the change in the load, while color red denotes the number of load configurations that occured *after* the change in the load. The purpose of
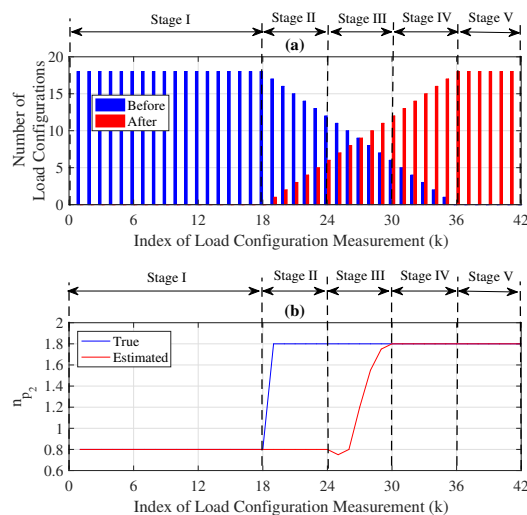
Fig. 10. An example to track the change in a time-varying load at bus 2 using the proposed method: (a) The number of load configurations corresponding to the cases *before* and *after* the change in the load; (b) the true and estimated load model parameters. Note that, Load modeling at each measurement point k is done over a window of the last 18 load configurations.

such color-coded distinction is to demonstrate the transition in the load modeling process from the measurement points that correspond to the old load model to the measurement points that correspond to the new load model.

Fig. 10(b) shows the true and estimated n_{p_2}. In Stage I, the estimated load model is accurate up until the moment when the load changes. In Stage II, the majority of the available measured load configurations correspond to the case *before* the change in the load. Therefore, either they do not trigger an update in the estimated load parameter, or the update in the estimated load parameter is not accepted due to a spike in the residues. In Stage III, the situation is reversed, and the majority of the available measured load configurations correspond to the case *after* the change in the load. Therefore, the estimated load parameter is updated until it gradually approaches the new true load parameter. In Stage IV, the available measured load configurations are sufficient to achieve accurate load modeling. Finally, in Stage V, the load modeling procedure is back to its initial stage where all the available measured load configurations correspond to the same load model.

G. Load Modeling in Presence of DERs

In this section, we examine the effectiveness of the proposed load modeling method in presence of DERs. The test setup is similar to Case II in Section V-A, where it is assumed that a PV unit is connected to bus number 6 through a line with $1.5 + j0.5$ ohms. The nominal active power and reactive power of the PV unit are 300 kW and -15 kVAR, respectively. Thus, the renewable energy penetration in this test setup is 32%. In this study, we assume that the output power of the PV unit is either directly measured or estimated, while the voltage phasor of the PV unit is estimated as an unknown in the load modeling estimation. Similar to Section V-C, we consider redundant load configurations, as in Tables II and IV.

TABLE VIII
IMPACT OF ERROR IN MEASURING OR ESTIMATING PV OUTPUT.

Error in PV Unit	0.5%	1.0%	1.5%	2.0%	2.5%	3.0%
Error in Estimating n_p	0.28	0.48	0.69	0.85	1.11	1.29
Error in Estimating n_q	0.75	1.42	2.29	2.73	3.80	4.45

It is assumed that the output power of the PV unit changes during these load configurations, with possible turn off events.

First, the load modeling is conducted under the assumption that the output power of the PV unit is measured accurately. As expected, the error in individual load modeling is less than 0.01%. Next, we use the Monte Carlo method to generate different scenarios for different levels of error in measuring or forecasting the output power of the PV unit. The results are shown in Table VIII. We see that, even in the presence of imperfect knowledge on output power of the PV unit, the results demonstrate an overall satisfactory performance in load modeling.

H. Sequential Load Modeling on IEEE 33-bus Test System

The detailed data for the IEEE 33-bus test feeder is available in [46]. We modified this feeder by adding one virtual infinite bus with line impedance 0.5+j0.25 Ohms to the first bus of the feeder. We consider the first bus as measuring bus, which has no load. Thus, we aim to find load models for $n = 32$ buses. There are $2^{32} - 1$ possible load configurations for this feeder. From Theorem 2, only 64 distinct load configurations that satisfy (9) are needed to obtain the individual load models.

Figs. 11(a)-(c) show the measured voltage, active power, and reactive power at the substation. We apply the sequential load modeling method from Section IV-B to $\hat{c}$ load configurations that hold (9) for $\hat{n}$ buses, where $\hat{c}_{\min} = 2\hat{n}$, to obtain the load models for $\hat{n}$ buses. Fig. 11(d) shows number of load models that are complete. Load modeling starts at time t_1 as soon as sufficient load configuration measurements become available.

Fig. 11(e) shows the computation time to run the load modeling algorithm. The computation time at each step depends on the number of equations and the number of unknowns in that step, and these numbers themselves depend on the new load configuration that occurs and measured at each step. For example, in the first execution of the algorithm at time t_1, the number of equations and unknowns are both 116; and the computation time is 62 seconds. As another example, in the second execution of the algorithm at time t_2, the number of equations and unknowns are both 23, and the computation time is 45 seconds. The total computation time across all steps to finish modeling all individual loads in this IEEE 33-bus test case is 38 minutes. This is the summation of the computation times across all 18 steps. That means, on average, each execution takes about two minutes.

I. Using Multiple Sensors on IEEE 123-bus Test System

In this section, we examine the effectiveness of the proposed load modeling method on the IEEE 123-bus test system, shown in Fig. 12. The detailed data for the IEEE 123-bus test feeder is available in [47]. We assume that five micro-PMUs are
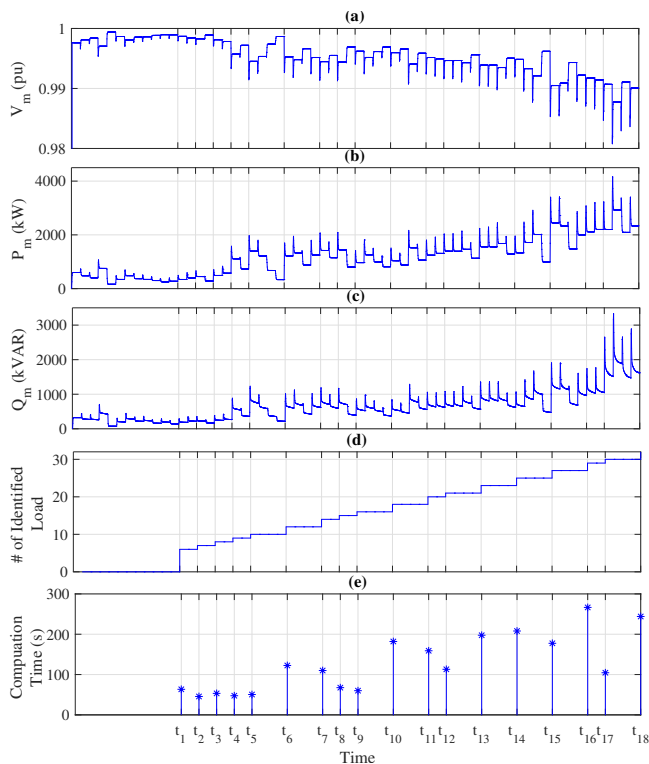
Fig. 11. Sequential load modeling for the IEEE 33-bus test feeder: (a) feeder-head voltage magnitude; (b) feeder-head active power; (c) feeder-head reactive power; (d) number of modeled load; (e) computation time of each sequence.

Fig. 12. Multiple sensor installations on the IEEE 123-bus test system.

deployed on the test feeder. This results in breaking down the load modeling problem into five zones, as shown in Fig. 12 by using different colors. Zone 1 includes the buses in downstream of sensor M_1 and upstream of sensors M_2, M_4, and M_5. Zone 2 includes the buses between sensors M_2 and M_3. Finally, Zones 3, 4, and 5 include the buses in downstream of sensors M_3, M_4, and M_5, respectively.

Recall from Section IV-F that each zone is essentially a separate load modeling problem. For instance, consider Zone 1 whose lines and buses are shown in black. This zone has 29 buses, out of which 19 buses are single phase. Thus, there exist 51 voltage phasors as unknowns in each load configuration. This zone also includes 21 single-phase loads. Note that, the load for the rest of the network is directly measured using sensors M_2, M_4, and M_5. From Theorem 2, the minimum number of distinct load configurations in order to estimate individual load model parameters for the loads in this zone is 42, i.e., twice the number of the loads in this zone. Here, we consider that 8 out of the 21 loads in Zone 1 are switched on during every load configuration. Thus, the Circuit Model includes $42\times(53+1) = 2,268$ equations and $42\times53+8\times42 = 2,562$ unknowns, see (5) and (6). The Load Model adds 315 additional equations and 21 new unknowns, see (10) and (11).

The combined system of nonlinear equations for load modeling in Zone 1 has 2,583 equations and unknowns. This number is 1,962 for Zone 2, 2070 for Zone 3, 456 for Zone 4, and 3,870 for Zone 5. To compare these numbers with the case when there is only one sensor installed, i.e., when sensors M_2, M_3, M_4, and M_5 are not available, the combined system of nonlinear equations has 47,405 equations and 47,405
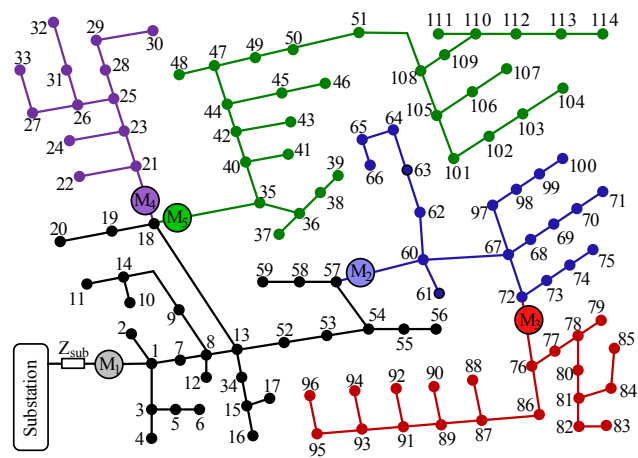
unknowns. Therefore, as expected, installing additional sensors can break down the load modeling problem into smaller problems, which accordingly reduces the computational complexity of the load modeling problem. Importantly, the resulting load modeling accuracy in each zone was 99.9% or more, demonstrating a very accurate load modeling performance. The computation-time of running Algorithm 1 on Zones 1, 2, 3, 4, and 5 was 31, 19, 21, 4, and 42 minutes, respectively.

VI. CONCLUSIONS

A new method was proposed for individual load modeling in power distribution systems. It works by analyzing the measurements at the feeder-head during load switching events as they occur across the feeder. The basic idea is that, once a load is switched, the switching event changes the voltage for the rest of the loads, which causes variation in their active and reactive power consumption; thus allowing us to estimate their load parameters. The proposed method can estimate load modeling parameters of individual loads. Therefore, it can support various practical use cases at distribution level, such as distribution-system optimal power flow analysis, DER management, Volt-VAR control, and voltage stability. Of course, if needed, the obtained individual load models can be aggregated to also provide a feeder-aggregated load model. A theoretical foundation was provided to determine the conditions for successful load modeling. The proposed load modeling can be done sequentially to model a subset of loads while more measurements become gradually available. The proposed method is extended to support also estimating the parameters for time-varying individual loads. The proposed method is extended also to utilize different types of redundancy in measurements to improve load modeling accuracy. Further, a residue-based bad data detection and identification method was developed to identify and drop load configuration measurements with erroneous switching status; thus to ensure the accuracy of the load models. Several test cases, including the IEEE 33-bus test feeder and IEEE 123-bus test feeder are studied to assess the performance of the proposed method.

References

[1] A. Arif, Z. Wang, J. Wang, B. Mather, H. Bashualdo, and D. Zhao, "Load modelinga review," *IEEE Trans. on Smart Grid*, vol. 9, no. 6, pp. 5986 – 5999, Nov. 2018.

[2] J. V. Milanovic, K. Yamashita, S. M. Villanueva, S. Ž. Djokic, and L. M. Korunović, "International industry practice on power system load modeling," *IEEE Trans. on Power Systems*, vol. 28, no. 3, pp. 3038–3046, Aug. 2013.

[3] M. Yao, D. K. Molzahn, and J. L. Mathieu, "The impact of load models in an algorithm for improving voltage stability via demand response," in *Proc. of IEEE Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Jan. 2017.

[4] M. Dabbaghjamanesh, S. Mehraeen, A. Kavousi-Fard, and F. Ferdowsi, "A new efficient stochastic energy management technique for interconnected ac microgrids," in *Proc. of IEEE PES General Meeting*, Portland, OR, Aug. 2018.

[5] K. P. Schneider and T. Weaver, "Volt-var optimization on american electric power feeders in northeast columbus," in *Proc. of IEEE PES T&D*, Orlando, FL, May 2012.

[6] B. Milosevic and M. Begovic, "Voltage-stability protection and control using a wide-area network of phasor measurements," *IEEE Trans. on Power Systems*, vol. 18, no. 1, pp. 121–127, Feb. 2003.

[7] D. K. Molzahn, B. C. Lesieutre, and C. L. DeMarco, "Approximate representation of ZIP loads in a semidefinite relaxation of the OPF problem," *IEEE Trans. on Power Systems*, vol. 29, no. 4, pp. 1864–1865, Jan. 2014.

[8] K. P. Schneider, J. C. Fuller, and D. P. Chassin, "Multi-state load models for distribution system analysis," *IEEE Trans. on Power Systems*, vol. 26, no. 4, pp. 2425–2433, Nov. 2011.

[9] A. S. Carneiro, L. F. Araujo, J. L. R. Pereira, P. A. Garcia, I. D. Melo, and M. B. Amaral, "Static load modeling based on field measurements," in *Proc. of IEEE PowerTech*, Manchester, UK, Jun. 2017.

[10] C. A. Baone, S. Veda, Y. Pan, W. Premerlani, J. Dai, and A. Johnson, "Measurement based static load model identification," in *Proc. of IEEE PES General Meeting*, Denver, CO, Jul. 2015.

[11] J. Zhao, Z. Wang, and J. Wang, "Robust time-varying load modeling for conservation voltage reduction assessment," *IEEE Trans. on Smart Grid*, vol. 9, no. 4, pp. 3304–3312, Nov. 2016.

[12] S. Pandey, A. K. Srivastava, P. Markham, M. Patel *et al.*, "Online estimation of steady-state load models considering data anomalies," *IEEE Trans. on Industry Applications*, vol. 54, no. 1, pp. 712–721, Jan. 2018.

[13] V. Vignesh, S. Chakrabarti, and S. Srivastava, "An experimental study on the load modelling using PMU measurements," in *Proc. of IEEE PES T&D*, Chicago, IL, Apr. 2014.

[14] Y. Ge, A. J. Flueck, D.-K. Kim, J.-B. Ahn, J.-D. Lee, and D.-Y. Kwon, "An event-oriented method for online load modeling based on synchrophasor data," *IEEE Trans. on Smart Grid*, vol. 6, no. 4, pp. 2060–2068, Mar. 2015.

[15] H. Renmu, M. Jin, and D. J. Hill, "Composite load modeling via measurement approach," *IEEE Trans. on Power Systems*, vol. 21, no. 2, pp. 663–672, Oct. 2006.

[16] S. Son, S. H. Lee, D. H. Choi, K. B. Song, J. D. Park, Y. H. Kwon, K. Hur, and J. W. Park, "Improvement of composite load modeling based on parameter sensitivity and dependency analyses," *IEEE Trans. on Power Systems*, vol. 29, no. 1, pp. 242–250, Sep. 2014.

[17] V. Vignesh, S. Chakrabarti, and S. C. Srivastava, "Load modeling under unbalanced disturbances," *IEEE Trans. on Power Systems*, vol. 31, no. 2, pp. 1661–1662, Mar. 2016.

[18] P. Regulski, D. Vilchis-Rodriguez, S. Djurovic, and V. Terzija, "Estimation of composite load model parameters using an improved particle swarm optimization method," *IEEE Trans. on Power Delivery*, vol. 30, no. 2, pp. 553–560, Feb. 2015.

[19] M. Jin, H. Dong, H. Ren-Mu, D. Zhao-Yang, and H. David, "Reducing identified parameters of measurement-based composite load model," *IEEE Trans. on Power Systems*, vol. 23, no. 1, pp. 76–83, Feb. 2008.

[20] K. Zhang, H. Zhu, and S. Guo, "Dependency analysis and improved parameter estimation for dynamic composite load modeling," *IEEE Trans. on Power Systems*, vol. 32, no. 4, pp. 3287–3297, Nov. 2017.

[21] D. Han, J. Ma, R.-m. He, and Z.-y. Dong, "A real application of measurement-based load modeling in large-scale power grids and its validation," *IEEE Trans. on Power Systems*, vol. 24, no. 4, pp. 1756–1764, Nov. 2009.

[22] B.-K. Choi and H.-D. Chiang, "Multiple solutions and plateau phenomenon in measurement-based load model development: issues and suggestions," *IEEE Trans. on Power Systems*, vol. 24, no. 2, pp. 824–831, May 2009.

[23] S. Shao, M. Pipattanasomporn, and S. Rahman, "Development of physical-based demand response-enabled residential load models," *IEEE Trans. on power systems*, vol. 28, no. 2, pp. 607–614, May 2013.

[24] A. J. Collin, G. Tsagarakis, A. E. Kiprakis, and S. McLaughlin, "Development of low-voltage load models for the residential load sector," *IEEE Trans. Power System*, vol. 29, no. 5, pp. 2180–2188, Sep. 2014.

[25] R. Torquato, Q. Shi, W. Xu, and W. Freitas, "A monte carlo simulation platform for studying low voltage residential networks," *IEEE Trans. on Smart Grid*, vol. 5, no. 6, pp. 2766–2776, Jul. 2014.

[26] K. McKenna, A. Keane *et al.*, "Residential load modeling of price-based demand response for network impact studies," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2285–2294, Sep. 2016.

[27] A. Shahsavari, M. Farajollahi, E. Stewart, E. Cortez, and H. Mohsenian-Rad, "Situational awareness in distribution grid using micro-PMU data: A machine learning approach," *IEEE Trans. on Power System*, to be published, doi: 10.1109/TSG.2019.2898676.

[28] H. Mohsenian-Rad, E. Stewart, and E. Cortez, "Distribution synchrophasors: Pairing big data with analytics to create actionable information," *IEEE Power and Energy Magazine*, vol. 16, no. 3, pp. 26 – 34, Apr. 2018.

[29] A. Shahsavari, A. Sadeghi-Mobarakeh, E. Stewart, E. Cortez, L. Alvarez, F. Megala, and H. Mohsenian-Rad, "Distribution grid reliability versus regulation market efficiency: An analysis based on micro-PMU data," *IEEE Trans. on Smart Grid*, vol. 8, no. 6, pp. 2916 – 2925, Nov. 2017.

[30] M. Farajollahi, A. Shahsavari, E. Stewart, and H. Mohsenian-Rad, "Locating the source of events in power distribution systems using micro-PMU data," *IEEE Trans. on Power Systems*, vol. 33, no. 6, pp. 6343 – 6354, May 2018.

[31] A. Shahsavari, M. Farajollahi, E. Stewart, A. von Meier, L. Alvarez, E. Cortez, and H. Mohsenian-Rad, "A data-driven analysis of capacitor bank operation at a distribution feeder using micro-PMU data," in *Proc. of the IEEE PES ISGT*, Washington D.C., Apr. 2017.

[32] A. Shahsavari, M. Farajollahi, E. Stewart, C. Roberts, F. Megala, L. Alvarez, E. Cortez, and H. Mohsenian-Rad, "Autopsy on active distribution networks: A data-driven fault analysis using micro-PMU data," in *Proc. of IEEE PES NAPS*, Morgantown, WV, Sep. 2017.

[33] A. Shahsavari, M. Farajollahi, E. Stewart, C. Roberts, and H. Mohsenian-Rad, "A data-driven analysis of lightning-initiated contingencies at a distribution grid with a PV farm using Micro-PMU data," in *Proc. of IEEE PES NAPS*, Morgantown, WV, Sep. 2017.

[34] P. Feldmann and R. A. Rohrer, "Proof of the number of independent kirchhoff equations in an electrical circuit," *IEEE Trans. on Circuits and Systems*, vol. 38, no. 7, pp. 681–684, Jul. 1991.

[35] M. Parodi and M. Storace, "Linear and nonlinear circuits: Basic & advanced concepts," *Springer*, pp. 230–231, 2018.

[36] D. Smith, M. Eggen, and R. S. Andre, "A transition to advanced mathematics," *Nelson Education*, pp. 147–148, 2014.

[37] T. M. Le, B. Fatahi, H. Khabbaz, and W. Sun, "Numerical optimization applying trust-region reflective least squares algorithm with constraints to optimize the non-linear creep parameters of soft soil," *Applied Mathematical Modelling*, vol. 41, pp. 236–256, Jan. 2017.

[38] M. K. Transtrum and J. P. Sethna, "Improvements to the levenberg-marquardt algorithm for nonlinear least-squares minimization," *arXiv preprint arXiv:1201.5885*, 2012.

[39] J. H. Teng, "A direct approach for distribution system load flow solutions," *IEEE Trans. on Power Delivery*, vol. 18, no. 3, pp. 882–887, Jul. 2003.

[40] E. Caro, A. J. Conejo, R. Mínguez, M. Zima, and G. Andersson, "Multiple bad data identification considering measurement dependencies," *IEEE Trans. on Power Systems*, vol. 26, no. 4, pp. 1953–1961, Nov. 2011.

[41] H. Manitoba, "Research centre," *PSCAD/EMTDC: Electromagnetic transients program including dc systems*, 1994.

[42] A. Bokhari, A. Alkan, R. Dogan, M. Diaz-Aguiló, F. De Leon, D. Czarkowski, Z. Zabar, L. Birenbaum, A. Noel, and R. E. Uosef, "Experimental determination of the ZIP coefficients for modern residential, commercial, and industrial loads," *IEEE Trans. on Power Delivery*, vol. 29, no. 3, pp. 1372–1381, Oct. 2014.

[43] E. Sortomme, A. Negash, S. Venkata, and D. Kirschen, "Voltage dependent load models of charging electric vehicles," in *Proc. of IEEE PES General Meeting*, Vancouver, BC, Canada, Nov. 2013.

[44] K. D. Jones, A. Pal, and J. S. Thorp, "Methodology for performing synchrophasor data conditioning and validation," *IEEE Trans. on Power Systems*, vol. 30, no. 3, pp. 1121–1130, May 2015.

[45] K. Kirihara, K. E. Reinhard, A. K. Yoon, and P. W. Sauer, "Investigating synchrophasor data quality issues," in *Proc. of IEEE Power and Energy Conference*, Champaign, IL, Feb. 2014.

[46] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, Apr. 1989.

[47] IEEE, Distribution System Analysis Subcommittee Report. [Online]. Available: http://sites.ieee.org/pes-testfeeders/resources/

Alireza Shahsavari (S17) received the M.Sc. degree in electrical engineering - Power systems from the University of Tehran, Tehran, Iran, in 2014. He is currently working toward the Ph.D. degree in electrical and computer engineering with the University of California, Riverside, CA, USA. He was also the recipient of NASA MIRO FIELDS fellowship in Big-Data analysis from NASA JPL, 2017. His research interests include data analysis in smart grid and distribution system planning, operation, and monitoring.

Mohammad Farajollahi (S15) received the B.Sc. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 2014, the M.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2016, and is currently working toward the Ph.D. degree with the University of California, Riverside, CA, USA. His research interests include power system planning, operation, reliability, as well as optimization. He is specifically working on applications of micro-PMUs and data analysis in distribution system monitoring.

Hamed Mohsenian-Rad (S'04-M'09-SM'14) received the Ph.D. degree in electrical and computer engineering from the University of British Columbia, Vancouver, BC, Canada, in 2008. He is currently an Associate Professor of electrical engineering with the University of California, Riverside, CA, USA. His research interests include modeling, data analysis, and optimization of power systems and smart grids. He was the recipient of the National Science Foundation CAREER Award, the Best Paper Award from the IEEE Power and Energy Society General Meeting, and the Best Paper Award from the IEEE Int. Conference on Smart Grid Communications. He was an Editor of the IEEE TRANSACTIONS ON SMART GRID and the IEEE POWER ENGINEERING LETTERS.