

# Unbalanced Three-Phase Distribution Grid Topology Estimation and Bus Phase Identification

Yizheng Liao, *Student Member, IEEE*, Yang Weng, *Member, IEEE*, Guangyi Liu, *Senior Member, IEEE*,  
Zhongyang Zhao, Chin-Woo Tan, Ram Rajagopal, *Member, IEEE*

**Abstract**—There is an increasing need for monitoring and controlling uncertainties brought by distributed energy resources in distribution grids. For such goal, accurate three-phase topology is the basis for correlating and exterminating measurements in unbalanced distribution networks. Unfortunately, such topology knowledge is often unavailable due to limited investment, especially for secondary distribution grids. Also, the bus phase connectivity information is inaccurate due to human errors or outdated records. For this challenge, we utilize smart meter data at different phases for an information-theoretic approach to learn the structures. Specifically, we convert the system of three unbalanced phasors into symmetrical components, namely the positive, negative, and zero sequences. Then, we prove that Chow-Liu algorithm can find the optimal topology by utilizing power flow equation and the conditional independence relationships implied by the radial three-phase structure of distribution grids with the presence of incorrect bus phase labels. At last, by utilizing Carson's equation, we prove that the bus phase connection can be correctly identified using voltage measurements. For validation, we extensively simulate on IEEE 37- and 123-bus systems using real data from PG&E, ADRES Project, and Pecan Street. We observe that the algorithm is highly accurate for finding three-phase topology in distribution grids even with strong load unbalancing condition and DERs. This ensures close monitoring and controlling DERs in distribution grids.

## I. INTRODUCTION

The power distribution system is currently undergoing a dramatic transformation in both form and function. A large-scale deployment of technologies such as rooftop solar, electric vehicles (EVs), and smart home management systems have the potential to offer cheaper, cleaner and more controllable energy to the customers. On the other hand, integrating these resources has been proven to be nontrivial, largely because of their inherent uncertainty and distributed nature.

For example, even a small-scale of DERs can affect the stability of the local distribution grid [1]. Such a problem will be aggravated by the unbalance situation in the distribution grid especially when uneven DER deployment happens. Furthermore, the more frequent bi-directional power flows easily leave the existing monitoring system with passive protective devices, insufficient for robust grid operation. In addition to the static connectivity, mobile components, such as EVs, can further jeopardize the grid due to their frequent plug-in [2]. Therefore, we need carefully designed three-phase grid monitoring tools for islanding and line work hazards in system operation with deep and uneven DER penetration. For such monitoring, grid topology information is a prerequisite.

For topology estimation, the transmission grid assumes a prior grid knowledge, which needs limited error correction. Also, it is assumed that infrequent reconfiguration happens,

identifiable by generalized state estimation [3]–[5]. Unfortunately, such assumptions do not hold in medium- and low-voltage distribution grids, where topology can change relatively more frequently with limited sensing devices. Furthermore, many urban distribution lines have been underground for decades, making prior knowledge of topology suspicious and expensive to verify [6].

For distribution grid topology identification, many methods have been proposed in recent years [7]–[11]. Unfortunately, most of them focus on the balanced systems or single phase systems. For utility practice, distribution grids are usually considered unbalanced and have multiple phases when power supplies to buildings and residential houses. Such unbalanced multi-phase systems include three-phase distributions and two-phase distributions. One reason is that the loads connected at different phases are unbalanced due to the uneven growth in each feeder territory [12], [13]. For example, surveyed by the American National Standards Institute (ANSI), 2% distribution grids in the USA have a significant undesirable degree of unbalance [14], [15].

In addition to the load unbalance problem in existing grids, such issue will become more frequent in future distribution grids due to growing renewable penetration. For example, the unbalance of the three-phase system appears more often because the installations and operations of many DER devices are not fully controlled by the utilities. This fact makes the requirement of balanced grids in previous works invalid. One way to address this problem is to apply the algorithms developed for single-phase or steady-state system for estimating connectivity of each phase. Then, one can merge the single-phase topology results to recover the three-phase system topology. However, this approach performs poorly in unbalanced three-phase systems due to the power line coupling between multi-phase buses. Finally, for many utilities, as high as 10% phase labeling are incorrect or unknown because of human errors or outdated records. This error makes identifying new topology based on existing methods not sound anymore.

For resolving the problems above, we propose a data-driven method that utilizes the smart meter data in different phases to estimate the topology of three-phase distribution grid systems. Two-phase and one-phase systems are special cases of our consideration. Building on the probabilistic graphical model formulation of distribution grids [16], we first expand from single-phase representation to three-phase systems with incorrect phase labels. In such model, a node represents the three-phase bus voltages and an edge between nodes indicates the statistical dependency among three-phase bus voltage measurements. Subsequently, we convert the system of three unbalanced phasors to three symmetrical components, namely the positive, negative, and zero sequences. Then, we prove that the Chow-Liu algorithm can identify the correct three-phase topology by utilizing power flow equation and the conditional independence relationships implied by the radial three-phase structure of distribution grids. As a highlight, the proposed method is robust to incorrect phase label, which is a big problem in distribution grid operation. This is due to the

Y. Liao, C-W.Tan, R. Rajagopal are with Department of Civil and Environmental Engineering, Stanford University, Stanford, CA, 94305 USA e-mail: ({yzliao, tancw, ramr}@stanford.edu). Y. Weng is with School of Electrical, Computing, and Energy Engineering, Arizona State University, Tempe, AZ, 85287 USA e-mail: yang.weng@asu.edu. G. Liu and Z. Zhao are with GEIRI North America, San Jose, CA, 95134, USA e-mail: (guangyi.liu@geirina.net, ecezhao@gmail.com)

label-invariant property of mutual information. Another major contribution we have is on phase identification. Specifically, we propose a data-driven approach to identify correct bus phase connection by utilizing Carson's equation [17], which is employed for deriving the primitive phase impedances of different lines. Finally, since phasor measurement units (PMUs) have not been widely deployed, we propose a refined algorithm that only needs the voltage magnitudes for topology estimation.

The performance of the proposed method is verified by simulations on the IEEE 37-bus and 123-bus distribution test cases [18]. We used 3 different PG&E residential household data sets in North California, ADRES project data set [19], [20] that contains 30 houses load profiles in Upper-Austria, and Pecan Street data set, which contains load data of 345 houses with PV panels in Austin, Texas. Simulations are conducted via GridLAB-D, an open source distribution grid simulator [21] for three-phase systems. Simulation results show that, provided with hourly measurements, the proposed algorithm can perfectly estimate the topology of three-phase distribution grids.

The rest of the paper is organized as follows: Section II introduces the modeling of the three-phase distribution system and the problem of data-driven topology estimation. In Section III, we firstly prove that topology estimation problem of three-phase distribution grid can be solved as a mutual information maximization problem and propose an algorithm to solve such a maximization problem in three-phase setup. Also, we propose a method to identify bus phase connection. In Section IV, to address the unbalance in distribution grids, we transform an unbalanced distribution grid to a symmetric system using sequence component frame and prove that the mutual information approach can still apply to grid topology estimation and phase identification. Section V evaluates the performance of our method using IEEE test cases and real data collected from different regions. Section VI concludes this paper.

## II. THREE-PHASE DISTRIBUTION GRID MODELING AND PROBLEM FORMULATION

In order to better formulate the topology estimation problem, we need to describe the three-phase distribution network topology and data. A distribution grid is modeled by a graph  $\mathcal{G} = (\mathcal{M}, \mathcal{E})$ , where the vertex set  $\mathcal{M} = \{0, 1, 2, \dots, M\}$  represents the set of buses and the unidirectional edge set  $\mathcal{E} = \{(i, k), i, k \in \mathcal{M}\}$  represents the branches. The branch between two buses is not necessary to be three-phase. In the distribution grid, bus 0 is the substation with a fixed voltage and is the root of the tree graph.  $\mathcal{M}^+$  denotes the set of buses excluding the substation, i.e.,  $\mathcal{M}^+ = \mathcal{M} \setminus \{0\}$ . If bus  $i$  and bus  $k$  are connected, i.e.,  $(i, k) \in \mathcal{E}$ , and bus  $i$  is closer to the root (substation) than bus  $k$ , bus  $i$  is the *parent* of bus  $k$  and bus  $k$  is the *child* of  $i$ . We use  $\text{pa}(i)$  to denote the parent bus of bus  $i$ . The root has no parent and all other buses in  $\mathcal{M}^+$  have exactly one parent. We use  $\mathcal{C}(i)$  to denote the set of child buses of bus  $i$  and use  $\mathcal{S}(i) = \mathcal{C}(\text{pa}(i)) \setminus \{i\}$  to denote the set of sibling buses of bus  $i$ .

Let  $a$ ,  $b$ , and  $c$  denote the three phases of the distribution grid. The vector  $\mathbf{V}_i^{abc} = [V_i^a, V_i^b, V_i^c]^T \in \mathbb{C}^3$  denotes the nodal voltages at bus  $i$ , where  $V_i^\phi$  denotes the line-to-ground complex voltage on phase  $\phi$  and  $T$  is the transpose operator. Similarly,  $\mathbf{I}_i^{abc} = [I_i^a, I_i^b, I_i^c]^T \in \mathbb{C}^3$  and  $\mathbf{S}_i^{abc} = [S_i^a, S_i^b, S_i^c]^T \in \mathbb{C}^3$  denote the vectors of current injections and injected complex powers at bus  $i$ , respectively. If a bus is only connected with one or two phases, the quantities of

the missing phase are zeros. For example, if bus  $i$  does not have phase  $c$ ,  $V_i^c = 0$ ,  $I_i^c = 0$ , and  $S_i^c = 0$ . For convenience, we use  $\mathbf{V}_i$ ,  $\mathbf{I}_i$ , and  $\mathbf{S}_i$  as the general notation of multi-phase quantities at bus  $i$ .

If bus  $i$  and bus  $k$  are connected, i.e.,  $(i, k) \in \mathcal{E}$ , the relationship between their nodal voltages and current injections can be expressed as follows [22], [23]:

$$\begin{bmatrix} \mathbf{I}_i \\ \mathbf{I}_k \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{ik} + \frac{1}{2}\mathbf{B}_{i,\text{shunt}} & -\mathbf{Y}_{ik} \\ -\mathbf{Y}_{ik} & \mathbf{Y}_{ik} + \frac{1}{2}\mathbf{B}_{k,\text{shunt}} \end{bmatrix} \begin{bmatrix} \mathbf{V}_i \\ \mathbf{V}_k \end{bmatrix},$$

where  $\mathbf{Y}_{ik} \in \mathbb{C}^{3 \times 3}$  denotes the admittance submatrix between bus  $i$  and bus  $k$  and  $\mathbf{B}_{i,\text{shunt}} \in \mathbb{C}^{3 \times 3}$  denotes the shunt capacitance at bus  $i$ . In a three-phase system,  $\mathbf{Y}_{ik}$  is not diagonal. The voltages at different phases are coupled. As we will show in Section V, this coupling property in the three-phase systems leads the existing steady-state methods to have poor performance in the unbalanced three-phase systems. Since  $\mathbf{B}_{i,\text{shunt}}$  is relatively small in distribution grids [24], we assume  $\mathbf{B}_{i,\text{shunt}} = \mathbf{0}$ . In the formulation above, the effect of the neutral wire is merged into the three-phase wires by applying Kron's reduction [23]. If bus  $i$  and bus  $k$  are not connected,  $\mathbf{Y}_{ik} = \mathbf{0}$ . Here, we do not need to assume the line to be transposed.

For bus  $i$ , the voltage measurement at time  $n$  is  $\mathbf{v}_i[n] = [v_i^a[n], v_i^b[n], v_i^c[n]]^T$ , where  $v_i^\phi[n] = |v_i^\phi[n]| \exp(j\theta_i^\phi[n])$  denotes the complex voltage measurement on phase  $\phi$  at time  $n$  and  $j = \sqrt{-1}$ . The magnitude  $|v_i^\phi[n]| \in \mathbb{R}$  is in volt and the phase angle  $\theta_i^\phi[n] \in \mathbb{R}$  is in degree. All measurements are assumed to be noiseless at first. In Section V, we will validate the proposed algorithm with noisy measurements.

With the modeling above, the multiphase distribution grid topology estimation and bus phase identification problem is defined as

- Problem: data-driven multiphase distribution grid topology and bus phase estimation using voltage measurements
- Given: the time-series voltage measurements with unknown bus phase labels  $\mathbf{v}_i[n]$ ,  $n = 1, \dots, N, i \in \mathcal{M}^+$
- Find: the unknown grid topology  $\mathcal{E}$  and bus phase  $\phi$ .

## III. THREE-PHASE DISTRIBUTION GRID TOPOLOGY ESTIMATION AND BUS PHASE IDENTIFICATION

In today's distribution grid, the smart meters and micro phase measurement unit ( $\mu\text{PMU}$ ) have been installed and provide highly accurate historical data about three-phase system operation. However, a practical issue that exists in many distribution grids, especially the low-voltage distribution grids, is that the smart meter phase connectivity information is inaccurate. In some countries, about 10% phase labels in low-voltage distribution grids are incorrect. Also, the phase connectivity can change over time when new customers and DER devices are connected to the grid [25]. As the correct phase connectivity information is critical to distribution grid planning, we want to estimate grid topology and identify phase connection at the same time. In this section, we will firstly extend our previous work [16] to estimate the topology of three-phase system with incorrect phase labels. Then, we will propose a method to identify phase connectivity by utilizing the statistical relationship between voltage measurements. The process of estimating topology and bus phase is summarized in Fig. 1.

The end-user measurements are time-series data. One way to represent these data is using a probability distribution. If we model the nodal three-phase voltage vector  $\mathbf{V}_i$  as a random vector, the joint distribution of voltage measurement  $P(\mathbf{V}_{\mathcal{M}^+})$

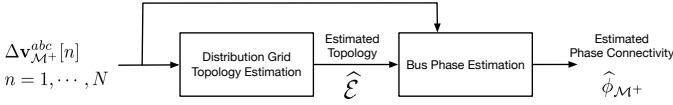


Fig. 1. Flow chart of topology estimation and bus phase identification process.

is  $P(\mathbf{V}_1)P(\mathbf{V}_2|\mathbf{V}_1)\cdots P(\mathbf{V}_M|\mathbf{V}_1, \dots, \mathbf{V}_{M-1})$ . Bus 0 is omitted because it is the slack bus with a fixed voltage.

In many previous works of distribution grid topology estimation [7], [9], [10], [16], only the steady-state voltages are required. However, with the presence of incorrect or unknown phase labels, we need to utilize all three phases voltage measurements for topology estimation. In the later part of this section, we will show that our method is invariant to phase labels and therefore, can estimate topology with incorrect or unknown phase labels.

In many medium- and low-voltage distribution grids, the distribution of voltage is irregular. To better formulate the topology estimation problem, we use the incremental change of measurements in this paper [26]. At bus  $i$ , the incremental change of voltage is  $\Delta \mathbf{v}_i[n] = \mathbf{v}_i^{abc}[n] - \mathbf{v}_i^{abc}[n-1]$  for  $n \geq 2$ . When  $n = 1$ ,  $\Delta \mathbf{v}_i[1] = 0$ . By using the incremental change  $\Delta \mathbf{V}$ , the joint probability becomes

$$P(\Delta \mathbf{V}_{\mathcal{M}+}) = P(\Delta \mathbf{V}_1)P(\Delta \mathbf{V}_2|\Delta \mathbf{V}_1)\cdots \times P(\Delta \mathbf{V}_M|\Delta \mathbf{V}_1, \dots, \Delta \mathbf{V}_{M-1}).$$

Since the nodal voltages are modeled as random vectors, the graph  $\mathcal{G}$  becomes a probabilistic graphical model with a tree structure. In a graphical model, the vertex represents a random vector (e.g.,  $\Delta \mathbf{V}_i$ ) and the edge between two vertices indicates the statistical dependency between bus voltages. Therefore, estimating distribution grid topology is equivalent to recovering the radial structure of the graphical model  $\mathcal{G}$ .

In the steady-state distribution grid, the nodal voltages only have statistical dependency with the nodal voltages of their parent bus [16]. Therefore, we extend such dependency in the steady-state system to approximate the three-phase system joint probability  $P(\Delta \mathbf{V}_{\mathcal{M}+})$  as

$$P(\Delta \mathbf{V}_{\mathcal{M}+}) \simeq \prod_{i=1}^M P(\Delta \mathbf{V}_i|\Delta \mathbf{V}_{\text{pa}(i)}). \quad (1)$$

If (1) holds, finding the structure of  $\mathcal{G}$  is equivalent to finding the parent of each bus for the same phase. Before proving the approximation in (1) holds with equality, we will make two assumptions and justify these assumptions using real data.

**Assumption 1.** In the three-phase distribution grid,

- the incremental change of the current injection  $\Delta \mathbf{I}$  at each non-slack bus is independent, i.e.,  $\Delta \mathbf{I}_i \perp \Delta \mathbf{I}_k$  for all  $i \neq k$ .
- the incremental changes of the current injection  $\Delta \mathbf{I}$  and bus voltage  $\Delta \mathbf{V}$  at each bus follow Gaussian distribution with zero means and non-zero covariances.

Fig. 2 shows the pairwise mutual information of the incremental changes of bus current injection using the real data from PG&E. The mutual information  $I(\mathbf{X}, \mathbf{Y})$  is a measure of the statistical dependence between two random vectors  $\mathbf{X}$  and  $\mathbf{Y}$ . When the mutual information is zero, these two random vectors are independent, i.e.,  $\mathbf{X} \perp \mathbf{Y}$  [27]. As shown in Fig. 2, most pairs of  $\Delta \mathbf{I}$  have small values. Therefore, we can assume that the current injections are independent with some approximation errors. This assumption has also been adopted in other works, such as [7] and [9]. To further validate the independence of  $\Delta \mathbf{I}$ , we plot the average auto-correlation of current injection increment of PG&E data in IEEE 123-bus

system in Fig. 3. The error bar is one standard deviation. We can observe that the auto-correlation of  $\Delta \mathbf{I}$  drops significantly as the lag increases. This observation justifies that the current injection increment is approximately independent over time.

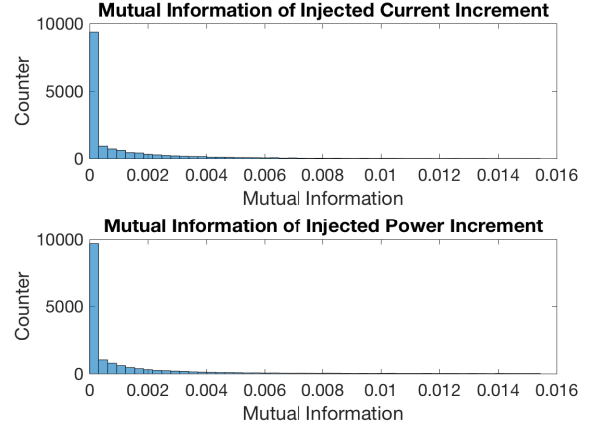


Fig. 2. Mutual information of pairwise current injection increment  $\Delta \mathbf{I}$  and power injection increment  $\Delta \mathbf{S}$  of PG&E data sets in the IEEE 123-bus system.

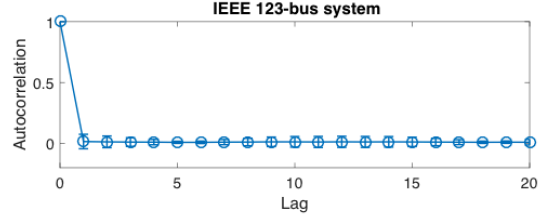


Fig. 3. Average auto-correlation of current injection increment  $\Delta \mathbf{I}$  of PG&E data sets in the IEEE 123-bus system. The error bar is one standard deviation.

In the existing distribution grid topology estimation works, both injected power increment independence and injected current increment independence are used. In [28], the authors use the real data to show that these two assumptions are equivalent in distribution grids. In Fig. 2, we illustrate that the mutual information of pairwise power injection increment  $\Delta \mathbf{S}$  and that of pairwise current injection increment  $\Delta \mathbf{I}$ . Both histograms are similar. In this paper, we prefer the assumption of current injection independence because it simplifies the proof of following theorems and lemmas.

With Assumption 1, we will use a two-stage proof to show that  $P(\Delta \mathbf{V}_{\mathcal{M}+}) = \prod_{i=1}^M P(\Delta \mathbf{V}_i|\Delta \mathbf{V}_{\text{pa}(i)})$ . In the first stage, we will prove that  $P(\Delta \mathbf{V}_{\mathcal{M}+}) = \prod_{i=1}^M P(\Delta \mathbf{V}_i|\Delta \mathbf{V}_{\{\text{pa}(i), \text{pa}(\text{pa}(i)), \text{S}(i)\}})$ . For convenience, we define  $\mathcal{C}^\infty(i)$  to denote all buses that are below bus  $i$ . For example, in Fig. 4,  $\mathcal{C}^\infty(1) = \{2, 3, 4, 5, 6, 7\}$  and  $\mathcal{C}^\infty(2) = \{4, 5\}$ .

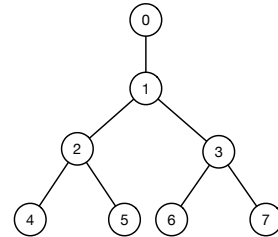


Fig. 4. An example of an 8-bus three-phase system. A node represents a bus, which can be single-phase or multi-phase. An edge represents a branch between two buses. The branch is unnecessary to be three-phase. Bus 0 is the substation (root).

**Lemma 1.** If the incremental change of current injection at each bus is approximately independent (i.e.,  $\Delta \mathbf{I}_i \perp \Delta \mathbf{I}_k$  for  $i \neq k$ ), given the incremental voltage changes of bus

$i$ 's parent ( $\Delta \mathbf{V}_{pa(i)}$ ), grandparent ( $\Delta \mathbf{V}_{pa(pa(i))}$ ), and siblings ( $\Delta \mathbf{V}_{S(i)}$ ), the incremental voltage changes of bus  $i$  and the buses that are not below bus  $i$  are conditionally independent, i.e.,  $\Delta \mathbf{V}_i \perp \Delta \mathbf{V}_k | \Delta \mathbf{V}_{\{pa(i), pa(pa(i)), S(i)\}}$  for  $k \notin \{pa(i), pa(pa(i)), S(i), C^\infty(i)\}$  and  $i \neq k$ .

We will use a simple example to demonstrate Lemma 1. A formal proof is given in Appendix section VII-A. For the example system in Fig. 4, the nodal admittance equation is  $\mathbf{Y}_{\mathcal{M}+} \Delta \mathbf{V}_{\mathcal{M}+} = \Delta \mathbf{I}_{\mathcal{M}+}$ , where

$$\mathbf{Y}_{\mathcal{M}+} = \begin{bmatrix} \mathbf{Y}_{11} & \mathbf{Y}_{12} & \mathbf{Y}_{13} & 0 & 0 & 0 & 0 \\ \mathbf{Y}_{21} & \mathbf{Y}_{22} & 0 & \mathbf{Y}_{24} & \mathbf{Y}_{25} & 0 & 0 \\ \mathbf{Y}_{31} & 0 & \mathbf{Y}_{33} & 0 & 0 & \mathbf{Y}_{36} & \mathbf{Y}_{37} \\ 0 & \mathbf{Y}_{42} & 0 & \mathbf{Y}_{44} & 0 & 0 & 0 \\ 0 & \mathbf{Y}_{52} & 0 & 0 & \mathbf{Y}_{55} & 0 & 0 \\ 0 & 0 & \mathbf{Y}_{63} & 0 & 0 & \mathbf{Y}_{66} & 0 \\ 0 & 0 & \mathbf{Y}_{73} & 0 & 0 & 0 & \mathbf{Y}_{77} \end{bmatrix},$$

$\mathbf{Y}_{ik} = \mathbf{Y}_{ki}$ , and  $\mathbf{Y}_{ii} = -\sum_{k=0, k \neq i}^7 \mathbf{Y}_{ik}$ . If  $\mathbf{Y}_{ik} = 0$ , there is no branch between bus  $i$  and  $k$ .

For bus 4,  $pa(4) = 2$ ,  $pa(pa(4)) = 1$ , and  $S(4) = \{5\}$ . Therefore, given  $\Delta \mathbf{V}_1 = \Delta \mathbf{v}_1$ ,  $\Delta \mathbf{V}_2 = \Delta \mathbf{v}_2$ , and  $\Delta \mathbf{V}_5 = \Delta \mathbf{v}_5$ , we have the following equations:

$$\Delta \mathbf{I}_1 = \mathbf{Y}_{11} \Delta \mathbf{v}_1 + \mathbf{Y}_{12} \Delta \mathbf{v}_2 + \mathbf{Y}_{13} \Delta \mathbf{V}_3, \quad (2)$$

$$\Delta \mathbf{I}_4 = \mathbf{Y}_{42} \Delta \mathbf{v}_2 + \mathbf{Y}_{44} \Delta \mathbf{V}_4, \quad (3)$$

$$\Delta \mathbf{I}_6 = \mathbf{Y}_{63} \Delta \mathbf{V}_3 + \mathbf{Y}_{66} \Delta \mathbf{V}_6, \quad (4)$$

$$\Delta \mathbf{I}_7 = \mathbf{Y}_{73} \Delta \mathbf{V}_3 + \mathbf{Y}_{77} \Delta \mathbf{V}_7. \quad (5)$$

Given  $\Delta \mathbf{I}_1 \perp \Delta \mathbf{I}_4$ , according to (2) and (3),  $\Delta \mathbf{V}_3$  and  $\Delta \mathbf{V}_4$  are conditionally independent given  $\Delta \mathbf{V}_1$ ,  $\Delta \mathbf{V}_2$ , and  $\Delta \mathbf{V}_5$ . In (2), we can rewrite  $\Delta \mathbf{V}_3$  as  $(\mathbf{Y}_{13})^{-1}(\Delta \mathbf{I}_1 - \mathbf{Y}_{11} \Delta \mathbf{v}_1 - \mathbf{Y}_{12} \Delta \mathbf{v}_2)$ . Then, we can substitute  $\Delta \mathbf{V}_3$  into (4). Since  $\Delta \mathbf{I}_4$  are independent with  $\Delta \mathbf{I}_1$  and  $\Delta \mathbf{I}_6$ ,  $\Delta \mathbf{I}_4$  and  $\Delta \mathbf{I}_6 - \mathbf{Y}_{63}(\mathbf{Y}_{13})^{-1} \Delta \mathbf{I}_1$  are independent. Therefore,  $\Delta \mathbf{V}_4$  and  $\Delta \mathbf{V}_6$  are conditionally independent. Similarly,  $\Delta \mathbf{V}_4$  and  $\Delta \mathbf{V}_7$  are conditionally independent.

For a non-leaf bus, bus 2, given  $\Delta \mathbf{V}_3 = \Delta \mathbf{v}_3$  and  $\Delta \mathbf{V}_1 = \Delta \mathbf{v}_1$ , we have the following equations:

$$\Delta \mathbf{I}_1 = \mathbf{Y}_{11} \Delta \mathbf{v}_1 + \mathbf{Y}_{12} \Delta \mathbf{v}_2 + \mathbf{Y}_{13} \Delta \mathbf{v}_3, \quad (6)$$

$$\Delta \mathbf{I}_6 = \mathbf{Y}_{63} \Delta \mathbf{v}_3 + \mathbf{Y}_{66} \Delta \mathbf{V}_6, \quad (7)$$

$$\Delta \mathbf{I}_7 = \mathbf{Y}_{73} \Delta \mathbf{v}_3 + \mathbf{Y}_{77} \Delta \mathbf{V}_7. \quad (8)$$

Given  $\Delta \mathbf{I}_1$  and  $\Delta \mathbf{I}_6$  are independent, according to (6) and (7),  $\Delta \mathbf{V}_2$  and  $\Delta \mathbf{V}_6$  are conditionally independent. Similarly,  $\Delta \mathbf{V}_2$  and  $\Delta \mathbf{V}_6$  are conditionally independent given  $\Delta \mathbf{I}_1$  and  $\Delta \mathbf{I}_7$  are independent. Our conclusion in Lemma 1 is similar to the results in [29].

**Assumption 2.** In a distribution grid, the mutual information between  $\Delta \mathbf{V}_i$  and its parent  $\Delta \mathbf{V}_{pa(i)}$  is much larger than the mutual information between  $\Delta \mathbf{V}_i$  and  $\Delta \mathbf{V}_{pa(pa(i))}$  and  $\Delta \mathbf{V}_{S(i)}$ .

Assumption 2 is inspired by the observation from real data. In Fig. 5, we plot the mutual information of voltage increments between each bus pair in IEEE 123-bus distribution system using the real load data from PG&E. The distribution grid configuration and simulation setup are described in Section V. In Fig. 5, the color in a square represents the mutual information of voltage increments of two buses. As discussed in [27], if the voltage increments of two buses are independent, their mutual information is zero (dark color). In Fig. 5, the circle refers to the bus neighbors (e.g. parent bus) and the crossing indicates the two-step neighbors (grandparent bus and sibling buses). If a square without any marker, it means the pair of buses is more than two-step away. As illustrated in Fig. 5, the mutual information between the voltages of

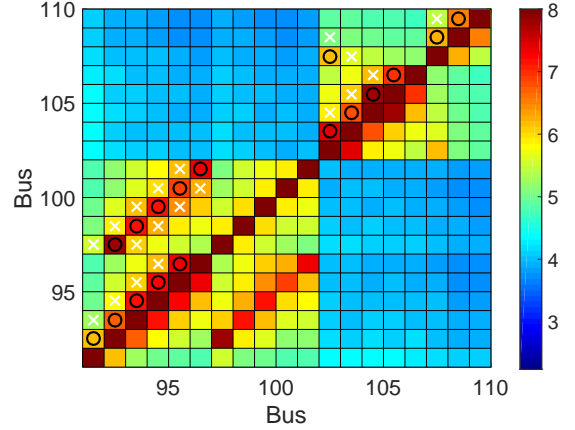


Fig. 5. Mutual information of pairwise buses in IEEE 123-bus system using PG&E data sets. The circle indicates the neighbors of bus  $i$ . The crossing indicates the two-step neighbor of bus  $i$ . The square without markers represents the bus pair that are more than two-step away.

two-step neighbors is higher than the mutual information of other bus pairs, but it is still much lower than the mutual information between two neighbors. The diagonal bus pairs have the highest mutual information because it is the self-information. With Assumption 2, we can simplify  $P(\Delta \mathbf{V}_{\mathcal{M}+})$  to depend on the voltages of parent buses. In Section V, we use numerical simulation to demonstrate that the approximation does not degrade the performance of topology and bus phase connectivity estimation.

**Lemma 2.** Given the incremental voltage changes of bus  $i$  in a multi-phase distribution grid, if the incremental change of current injection at each bus is approximately independent, the incremental voltage changes of every pair of bus  $i$ 's children are conditionally independent, i.e.,  $\Delta \mathbf{V}_k \perp \Delta \mathbf{V}_l | \Delta \mathbf{V}_i$  for  $k, l \in C(i)$  and  $k \neq l$ .

With Lemma 2, (1) holds with equality, i.e.,  $P(\Delta \mathbf{V}_{\mathcal{M}+}) = \prod_{i=1}^M P(\Delta \mathbf{V}_i | \Delta \mathbf{V}_{pa(i)})$ . Therefore, finding the distribution grid topology is equivalent to finding the parent of each bus. In next subsection, we will propose an information theoretical approach to estimate three-phase distribution grid topology with incorrect bus phase labels.

#### A. An Information Theoretical Approach to Estimate Three-phase Distribution Grid Topology

In order to find the parent of each bus, we can minimize the Kullback-Leibler divergence [27] of  $P(\Delta \mathbf{V}_{\mathcal{M}+})$  and  $Q(\Delta \mathbf{V}_{\mathcal{M}+}) = \prod_{i=1}^M P(\Delta \mathbf{V}_i | \Delta \mathbf{V}_{pa(i)})$ , i.e.,

$$\hat{\Theta} = \underset{\Theta \in \mathcal{M}+}{\operatorname{argmin}} D(P(\Delta \mathbf{V}_{\mathcal{M}+}) || Q(\Delta \mathbf{V}_{\mathcal{M}+}; \Theta)), \quad (9)$$

where  $\Theta$  denotes the collection of parent bus index of every bus, i.e.,  $\Theta = \{pa(1), \dots, pa(M)\}$ ,  $P$  denotes the joint distribution of all voltages, and  $Q$  denotes the distribution of voltage vectors with tree structure. When two distributions are identical, the KL divergence is zero. Therefore, as shown in Lemma 2, if we find a distribution  $Q(\Delta \mathbf{V}_{\mathcal{M}+}; \hat{\Theta})$  that is identical to  $P(\Delta \mathbf{V}_{\mathcal{M}+})$ ,  $\hat{\Theta}$  contains the parent bus index of every bus  $i$ . The associated structure of  $P_{CL}(\Delta \mathbf{V}_{\mathcal{M}+}) = Q(\Delta \mathbf{V}_{\mathcal{M}+}; \hat{\Theta})$  is the estimated topology of distribution grid. In Lemma 3, we will prove that (9) can be efficiently solved by utilizing the radial structure of distribution grids. In the following context,  $\Theta_i$  and  $pa(i)$  are used interchangeably.

**Lemma 3.** In a radial distribution grid, finding the topology is equivalent to solving the following optimization problem:

$$\hat{\Theta} = \underset{\Theta \in \mathcal{M}^+}{\operatorname{argmax}} \sum_{i=1}^M I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_{\Theta_i}), \quad (10)$$

where  $I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_{\Theta_i})$  denotes the mutual information.

The proof is in Appendix section VII-B. With such a lemma, we can use a mutual information-based maximum weight spanning tree algorithm, well-known as Chow-Liu algorithm [30], to find  $\hat{\Theta}$  and identify the three-phase distribution grid topology. This algorithm has been applied to single-phase system in [16]. In Theorem 1, we will prove that Chow-Liu algorithm can be applied to three-phase systems.

**Theorem 1.** In a radial three-phase distribution grid, the mutual information-based maximum weight spanning tree algorithm (Chow-Liu algorithm) estimates the best-fitted topology.

*Proof:* In this proof, we will show that the mutual information between connected buses is higher than those without a connection. If bus  $i$  is the parent of bus  $k$  and bus  $l$  and  $k \neq l$ , by utilizing the chain rule property of the mutual information [27], the joint mutual information is expressed as

$$\begin{aligned} & I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k, \Delta \mathbf{V}_l) \\ &= I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k) - I(\Delta \mathbf{V}_i, \Delta \mathbf{V}_k | \Delta \mathbf{V}_l), \\ &= I(\Delta \mathbf{V}_k; \Delta \mathbf{V}_l) - I(\Delta \mathbf{V}_k, \Delta \mathbf{V}_l | \Delta \mathbf{V}_i). \end{aligned}$$

Since  $\Delta \mathbf{V}_k | \Delta \mathbf{V}_i \perp \Delta \mathbf{V}_l | \Delta \mathbf{V}_i$ , the conditional mutual information  $I(\Delta \mathbf{V}_k, \Delta \mathbf{V}_l | \Delta \mathbf{V}_i)$  is zero. Then

$$I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k) = I(\Delta \mathbf{V}_k; \Delta \mathbf{V}_l) + I(\Delta \mathbf{V}_i, \Delta \mathbf{V}_k | \Delta \mathbf{V}_l).$$

Due to the fact that mutual information is always non-negative,  $I(\Delta \mathbf{V}_k; \Delta \mathbf{V}_i) \geq I(\Delta \mathbf{V}_k; \Delta \mathbf{V}_l)$ . Therefore, the mutual information between connected buses is larger than the mutual information between not connected buses. Then, by using the mutual information as the weight, the maximum weight spanning tree algorithm (Chow-Liu algorithm) solves (10) and estimates the distribution grid topology [16], [30]. ■

To compute mutual information  $I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k)$ , we can avoid integration by utilizing entropy, i.e.,

$$I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k) = H(\Delta \mathbf{V}_i) + H(\Delta \mathbf{V}_k) - H(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k), \quad (11)$$

where  $H(\Delta \mathbf{V}_i)$  denotes the entropy of  $\Delta \mathbf{V}_i$  and  $H(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k)$  denotes the cross-entropy of  $\Delta \mathbf{V}_i$  and  $\Delta \mathbf{V}_k$ . An advantage using (11) is that many distributions have a closed form of entropy. In Assumption 1, the incremental changes of voltages in distribution grid are assumed to follow Gaussian distribution approximately. Thus, if we assume  $\Delta \mathbf{V}_i$  follows a multivariate Gaussian distribution, its entropy is

$$H(\Delta \mathbf{V}_i) = \frac{r}{2} \log(2\pi \exp(1)) + \frac{1}{2} \log(\det \operatorname{Cov}(\Delta \mathbf{V}_i)), \quad (12)$$

where  $r$  denotes the dimension of the random vector  $\Delta \mathbf{V}_i$  and  $\operatorname{Cov}$  denotes the covariance. In some systems, the bus may not have all three phases. In this case, we need to exclude the disconnected phases in the computation of entropy.

As mentioned previously, a practical issue of using smart meter is that the phase connectivity information is usually inaccurate or missing. In some countries, about 10% phase labels in low-voltage distribution grids are incorrect or unknown. Also, the phase connectivity can change over time when new customers and DER devices are connected to the grid [25]. One may apply existing methods [25], [31], [32] to identify phase connectivity before estimating topology. Fortunately, the proposed topology estimation method does not require this preprocessing and is invariant to incorrect phase labels. Specifically,

when voltage phase is incorrectly labeled, the elements in the random variables in  $\Delta \mathbf{V}_i$  are permuted. This permutation does not affect the computation of  $\det \operatorname{Cov}(\Delta \mathbf{V}_i)$ , thus, does not change the values of  $H(\Delta \mathbf{V}_i)$  and  $H(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k)$ . Therefore, the mutual information  $I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k)$  is the same even the bus labels are incorrect. In Section V-C, we use numerical examples to show that our algorithm can recover the topology perfectly with the presence incorrect phase labels.

---

#### Algorithm 1 Multiphase Distribution Grid Topology Estimation

---

**Require:**  $\Delta \mathbf{v}_i[n]$  for  $i \in \mathcal{M}^+$ ,  $n = 1, \dots, N$

```

1: for  $i, k \in \mathcal{M}^+$  do
2:   Compute empirical mutual information  $I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k)$ 
   based on  $\Delta \mathbf{v}_i[n]$  and  $\Delta \mathbf{v}_k[n]$  using (11) and (12).
3: end for
4: Sort all possible bus pair  $(i, k)$  into non-increasing order
   by  $I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k)$ . Let  $\mathcal{T}$  denote the sorted set.
5: Let  $\hat{\mathcal{E}}$  be the set of nodal pair comprising the maximum
   weight spanning tree. Set  $\mathcal{E} = \emptyset$ .
6: for  $(i, k) \in \mathcal{T}$  do
7:   if cycle is detected in  $\hat{\mathcal{E}} \cup (i, k)$  then
8:     Continue
9:   else
10:     $\hat{\mathcal{E}} \leftarrow \hat{\mathcal{E}} \cup (i, k)$ 
11:   end if
12:   if  $|\hat{\mathcal{E}}| == M$  then
13:     break
14:   end if
15: return  $\hat{\mathcal{E}}$ 
16: end for

```

---

The proposed algorithm for three-phase distribution grid topology estimation is summarized in Algorithm 1. To efficiently build the maximum weight spanning tree (Steps 6 - 16), we can use the well-known Kruskal's minimum weight spanning tree algorithm [33], [34], which has a running time of  $O(M \log M)$  for a radial distribution network with  $M$  buses.

#### B. Bus Phase Identification

In the previous section, we have proved that even with incorrect phase labels, we can still estimate the three-phase distribution grid topology. In many field applications, accurate grid topology is not sufficient. The correct information of bus phase is also critical in grid planning and operation. In this subsection, we will propose a data-driven method to identify bus phase information.

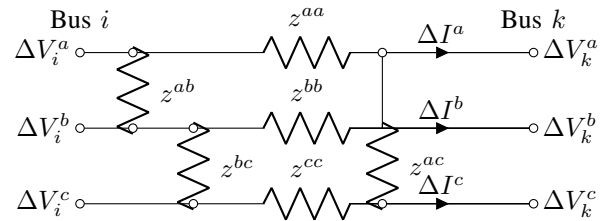


Fig. 6. An example of the two-port three-phase circuit.

**Lemma 4.** In a three-phase distribution grid, if two terminal buses of a branch are connected on the same phase, their phase voltage correlation is the largest.

*Proof:* Using the modified Carson's equation [17], the self impedance  $z^{aa}$  and mutual impedances  $z^{ab}$  and  $z^{ac}$  of a three-



phase power line can be computed as follows:

$$\begin{aligned} z^{aa} &= r_{ik} + 0.095 + j0.121 \times H_{ik}^a \Omega/\text{miles}, \\ z^{ab} &= 0.095 + j0.121 \times H_{ik}^{ab} \Omega/\text{miles}, \\ z^{ac} &= 0.095 + j0.121 \times H_{ik}^{ac} \Omega/\text{miles}, \end{aligned}$$

where  $H_{ik}^a$ ,  $H_{ik}^{ab}$ , and  $H_{ik}^{ac}$  are constants, and  $r_{ik}$  is resistance of branch  $i-k$  in  $\Omega/\text{miles}$ . In a distribution grid, the resistance is usually larger than reactance. Therefore, we can approximate that  $z^{aa} \simeq r_{ik}^a + 0.095$  and  $z^{ab} \simeq z^{ac} \simeq 0.095$ .

For bus  $i$  and bus  $k$ , the voltages and currents can be expressed as

$$\begin{bmatrix} \Delta V_i^a \\ \Delta V_i^b \\ \Delta V_i^c \end{bmatrix} = \begin{bmatrix} \Delta V_k^a \\ \Delta V_k^b \\ \Delta V_k^c \end{bmatrix} + \begin{bmatrix} Z^{aa} & Z^{ab} & Z^{ac} \\ Z^{ab} & Z^{bb} & Z^{bc} \\ Z^{ac} & Z^{bc} & Z^{cc} \end{bmatrix} \begin{bmatrix} \Delta I^a \\ \Delta I^b \\ \Delta I^c \end{bmatrix}, \quad (13)$$

where  $Z^{mn} = z^{mn} \times l$  and  $l$  is the line length. The equations above can be simplified to

$$\begin{aligned} \Delta V_i^a &= \Delta V_k^a + 0.095 \times l \times (I^a + I^b + I^c) + r_{ik} \times l \times I^a, \\ \Delta V_i^b &= \Delta V_k^b + 0.095 \times l \times (I^a + I^b + I^c) + r_{ik} \times l \times I^b, \\ \Delta V_i^c &= \Delta V_k^c + 0.095 \times l \times (I^a + I^b + I^c) + r_{ik} \times l \times I^c. \end{aligned}$$

The phase voltages at the two ends of a branch are in a linear relationship. Therefore, their correlation is the largest. ■

Table. I shows the corrections among bus 64, 65, and 66 in IEEE 123-bus system. Bus 64 and 65 are connected on phase  $b$ . Phase 65 and 66 are connected on phase  $c$ . We can observe that the correlation between bus 64 and 65 on phase  $b$  is much larger than other pairs. Similar observation holds for bus 65 and 66. Therefore, to identify bus phases in a distribution grid, we can apply the correlation check from the substation of the radial network down to the leaf buses. The reason is that we usually have reliable information about the bus labels of substation. Then, following the path of the estimated grid topology, we can find the phase of each bus. Note that, the metering device installed at each bus can provide the number of phases at each bus. Therefore, we can apply the method to all types of bus.

TABLE I. VOLTAGE CORRELATIONS AMONG BUS 64, 65, AND 66 IN IEEE 123-BUS SYSTEM.

	$\Delta V_{65}^a$	$\Delta V_{65}^b$	$\Delta V_{65}^c$
$V_{64}^b$	0.4956	0.9996	0.5332
$V_{66}^c$	0.9526	0.5479	1.0000

### C. Topology Estimation using Voltage Magnitude Only

In distribution grids, voltage angles  $\theta$  are hard to acquire because PMUs are not widely available. However, we can still find the distribution grid topology when only the voltage magnitudes  $|\Delta \mathbf{V}|$  are available. Using chain rule, the mutual information  $I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k)$  can be decomposed as

$$\begin{aligned} I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k) &= I(|\Delta \mathbf{V}_i|, \Delta \theta_i; |\Delta \mathbf{V}_k|, \Delta \theta_k) \\ &= \underbrace{I(|\Delta \mathbf{V}_i|, |\Delta \mathbf{V}_k|)}_{\text{term A}} + \underbrace{I(|\Delta \mathbf{V}_i|, \Delta \theta_i; |\Delta \mathbf{V}_k|)}_{\text{term B}} \\ &\quad + \underbrace{I(|\Delta \mathbf{V}_i|, \Delta \theta_k; |\Delta \mathbf{V}_k|, \Delta \theta_i)}_{\text{term C}}. \end{aligned}$$

In Fig. 7, we empirically plot the pairwise mutual information of term A, B, C using the IEEE 123-bus system and the real data from PG&E. We can see that the values of term A is much larger than term B and term C. The reason is that the changes of voltage angles are relatively small in distribution grids and thus, contain less information than voltage magnitudes. Therefore, based on our empirical observation, we can use  $I(|\Delta \mathbf{V}_i|; |\Delta \mathbf{V}_k|)$  to estimate distribution grid topology.

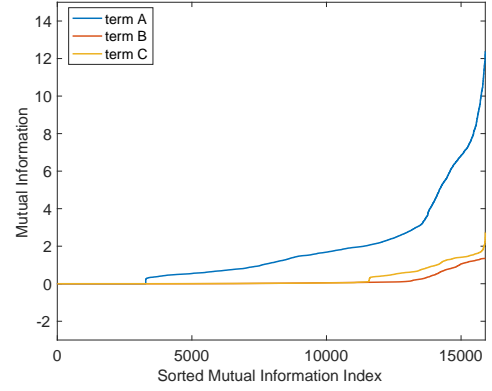


Fig. 7. Pairwise mutual information breakdown.

Many smart meters deployed can measure the voltages of all three phases. Therefore, we can still apply the proposed algorithm when only smart meter measurements are available.

### IV. UNBALANCED THREE-PHASE DISTRIBUTION GRID TOPOLOGY ESTIMATION WITH INCORRECT PHASE LABELS

The results in the previous section illustrate the topology estimation for balanced three-phase systems with incorrect phase labels. However, it is not directly expendable to unbalanced three-phase systems. As shown in Fig. 6, the voltages and currents are coupled cross different phases. Also, the unbalanced loads on each phase lead to the voltages angles are not separated by  $2\pi/3$ . To address these issues, we transform the grid using the sequence component framework. As a highlight, we do not require to know the correct phase labels. The voltage phasor is decomposed into three balanced phasors known as the positive sequence, the negative sequence, and the zero sequence. The three-phase voltage  $\Delta \mathbf{V}_i$  in phase frame can be decomposed as follows:

$$\Delta \mathbf{V}_i = \begin{bmatrix} 1 & 1 & 1 \\ h^2 & h & 1 \\ h & h^2 & 1 \end{bmatrix} \begin{bmatrix} \Delta V_a^p \\ \Delta V_a^n \\ \Delta V_a^z \end{bmatrix} = \mathbf{H} \Delta \mathbf{V}_a^{pnz}, \quad (14)$$

where  $h = \exp(j2\pi/3)$  and  $h^2 = \exp(-j2\pi/3)$ ,  $V_a^p, V_a^n, V_a^z$  denote the positive-sequence, negative-sequence, and zero-sequence voltage on phase  $a$ .  $\Delta \mathbf{V}_a^{pnz}$  is called the sequence voltage of phase  $a$ . Since each sequence component system is balanced, the sequence component voltages of phase  $b$  and phase  $c$  are the phase shifts of voltage on phase  $a$ ,  $\Delta \mathbf{V}_a^{pnz}$ . Thus, we do not need to compute them. In the following text, we will use  $\Delta \mathbf{V}_i^{pnz}$  to denote the sequence voltage vector of bus  $i$  on phase  $a$ . Since the smart meters can provide multi-phase measurements, we usually know the three-phase voltage measurements. Therefore, the sequence voltages can be computed as follows:

$$\Delta \mathbf{V}^{pnz} = \mathbf{H}^{-1} \Delta \mathbf{V} = \frac{1}{3} \mathbf{H}^H \Delta \mathbf{V}, \quad (15)$$

where the operator  $H$  denotes the Hermitian transpose. The same transformation can also be applied to the three-phase current phasors and admittance matrix, i.e.,

$$\Delta \mathbf{I}_i^{pnz} = \mathbf{H}^{-1} \Delta \mathbf{I}_i, \quad \mathbf{Y}_{ik}^{pnz} = \mathbf{H}^{-1} \mathbf{Y}_{ik} \mathbf{H}.$$

We want to highlight that the transformation above is applied to the three-phase voltage phasors, current phasors, and admittance matrix at a particular bus, not the entire system. Therefore, if two buses are not connected, e.g.,  $\mathbf{Y}_{ik} = 0$ ,  $\mathbf{Y}_{ik}^{pnz} = 0$ . Therefore, finding topology in a phase frame is equivalent to finding topology in a sequence component frame. To apply the mutual information-based maximum weight spanning tree algorithm (Chow-Liu algorithm), we want to prove

that the joint probability  $P(\Delta \mathbf{V}_{\mathcal{M}^+}^{pnz})$  can be factorized as  $\prod_{i=1}^M P(\Delta \mathbf{V}_i^{pnz} | \Delta \mathbf{V}_{\text{pa}(i)}^{pnz})$ .

**Lemma 5** (Data Processing Inequality [27]). *If random vectors  $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$  forms a Markov Chain, i.e.,  $\mathbf{X} \rightarrow \mathbf{Y} \rightarrow \mathbf{Z}$ ,  $I(\mathbf{X}; \mathbf{Y}) \geq I(\mathbf{X}; \mathbf{Z})$ . Also, for the function of  $\mathbf{Y}$ ,  $g(\mathbf{Y})$ ,  $I(\mathbf{X}; \mathbf{Y}) \geq I(\mathbf{X}; g(\mathbf{Y}))$ .*

**Lemma 6.** *Consider a multi-phase distribution grid and assume that the current injection increment at each bus is approximately independent, e.g.,  $\Delta \mathbf{I}_i \perp \Delta \mathbf{I}_k$  for  $i \neq k$ . Given the nodal bus voltage increment of bus  $i$  in sequence component frame, the nodal bus voltage increments of every pair of bus  $i$ 's children are conditionally independent, i.e.,  $\Delta \mathbf{V}_k^{pnz} \perp \Delta \mathbf{V}_l^{pnz} | \Delta \mathbf{V}_i^{pnz}$  for  $k, l \in \mathcal{C}(i)$  and  $k \neq l$ .*

*Proof:* Firstly, we want to prove that when the current injection increment at each bus is independent in phase frame, this independence also holds in sequence component frame. There are multiple ways to prove it. Here, we will use an information theoretical approach.

When two random vectors are independent, their mutual information is zero [27], e.g.,  $I(\Delta \mathbf{I}_i; \Delta \mathbf{I}_k) = 0$  if  $\Delta \mathbf{I}_i \perp \Delta \mathbf{I}_k$ . Since  $\Delta \mathbf{I}_k^{pnz} = \mathbf{H}^{-1} \Delta \mathbf{I}_k$  is a linear transformation of  $\Delta \mathbf{I}_k$ , we can form a Markov Chain, i.e.,  $\Delta \mathbf{I}_i \rightarrow \Delta \mathbf{I}_k \rightarrow \Delta \mathbf{I}_k^{pnz}$ , we can apply Lemma 5 to show that

$$I(\Delta \mathbf{I}_i; \Delta \mathbf{I}_k^{pnz}) \leq I(\Delta \mathbf{I}_i; \Delta \mathbf{I}_k) = 0.$$

Because the mutual information is non-negative,  $I(\Delta \mathbf{I}_i; \Delta \mathbf{I}_k^{pnz}) = 0$ .  $\Delta \mathbf{I}_i = \mathbf{H} \Delta \mathbf{I}_i^{pnz}$  is a function of  $\Delta \mathbf{I}_i^{pnz}$ . Thus, we can form another Markov Chain:  $\Delta \mathbf{I}_i^{pnz} \rightarrow \Delta \mathbf{I}_i \rightarrow \Delta \mathbf{I}_k^{pnz}$ . By applying Lemma 5 again, we have that

$$I(\Delta \mathbf{I}_i^{pnz}; \Delta \mathbf{I}_k^{pnz}) \leq I(\Delta \mathbf{I}_i; \Delta \mathbf{I}_k^{pnz}) = 0.$$

Hence,  $I(\Delta \mathbf{I}_i^{pnz}; \Delta \mathbf{I}_k^{pnz}) = 0$  due to the non-negativity of mutual information. Therefore, we prove that if the current injections are independent in phase frame, they are also independent in sequence component frame.

Next, we want to show that the conditional independence of nodal voltages holds in sequence component frame. We will use the example in Fig. 4 again to illustrate it. In sequence component frame, the nodal equation of the system in Fig. 4 is  $\mathbf{Y}_{\mathcal{M}^+}^{pnz} \Delta \mathbf{V}_{\mathcal{M}^+}^{pnz} = \Delta \mathbf{I}_{\mathcal{M}^+}^{pnz}$ , where

$$\mathbf{Y}_{\mathcal{M}^+}^{pnz} = \begin{bmatrix} \mathbf{Y}_{11}^{pnz} & \mathbf{Y}_{12}^{pnz} & \mathbf{Y}_{13}^{pnz} & 0 & 0 & 0 & 0 \\ \mathbf{Y}_{21}^{pnz} & \mathbf{Y}_{22}^{pnz} & 0 & \mathbf{Y}_{24}^{pnz} & \mathbf{Y}_{25}^{pnz} & 0 & 0 \\ \mathbf{Y}_{31}^{pnz} & 0 & \mathbf{Y}_{33}^{pnz} & 0 & 0 & \mathbf{Y}_{36}^{pnz} & \mathbf{Y}_{37}^{pnz} \\ 0 & \mathbf{Y}_{42}^{pnz} & 0 & \mathbf{Y}_{44}^{pnz} & 0 & 0 & 0 \\ 0 & \mathbf{Y}_{52}^{pnz} & 0 & 0 & \mathbf{Y}_{55}^{pnz} & 0 & 0 \\ 0 & 0 & \mathbf{Y}_{63}^{pnz} & 0 & 0 & \mathbf{Y}_{66}^{pnz} & 0 \\ 0 & 0 & \mathbf{Y}_{73}^{pnz} & 0 & 0 & 0 & \mathbf{Y}_{77}^{pnz} \end{bmatrix}$$

$\mathbf{Y}_{ik}^{pnz} = \mathbf{Y}_{ki}^{pnz}$ , and  $\mathbf{Y}_{ii}^{pnz} = -\sum_{k=1, k \neq i}^7 \mathbf{Y}_{ik}^{pnz}$ . If  $\mathbf{Y}_{ik}^{pnz} = 0$ , there is no branch between bus  $i$  and  $k$ . This equation is in the same format as (III). Since we have proved that  $\Delta \mathbf{I}_i^{pnz} \perp \Delta \mathbf{I}_k^{pnz}$  for all  $i \neq k$ , we can use the same method in the proof of Lemma 1 and Lemma 2 to show the conditional independence of nodal voltages in sequence component frame. ■

With Lemma 6, we prove that the conditional independence holds in the sequence component frame as well, e.g.,  $P(\Delta \mathbf{V}_{\mathcal{M}^+}^{pnz}) = \prod_{i=1}^M P(\Delta \mathbf{V}_i^{pnz} | \Delta \mathbf{V}_{\text{pa}(i)}^{pnz})$ . Since the sequence component system is a balanced three-phase system, we can apply the Chow-Liu algorithm to estimate topology in the sequence component frame.

**Theorem 2.** *In an unbalanced radial distribution grid, the topology can be estimated by solving the following problem*

$$\hat{\Theta} = \underset{\Theta \subset \mathcal{M}^+}{\operatorname{argmax}} \sum_{i=1}^M I(\Delta \mathbf{V}_i^{pnz}; \Delta \mathbf{V}_{\Theta_i}^{pnz}).$$

*Also, the mutual information-based maximum weight spanning tree algorithm (Chow-Liu algorithm) solves the problem above.*

The proof of Theorem 2 is omitted here because it is similar to the proofs of Lemma 3 and Theorem 1. The topology estimation algorithm for unbalanced three-phase distribution grid is summarized in Algorithm 2.

**Algorithm 2** Unbalanced Multiphase Distribution Grid Topology Estimation via Sequence Component Frame

**Require:**  $\Delta \mathbf{v}_i[n]$  for  $i \in \mathcal{M}^+$ ,  $n = 1, \dots, N$

- 1: Compute voltage phasor  $\Delta \mathbf{v}_i^{pnz}[n]$  using (15) for  $i \in \mathcal{M}^+$ ,  $n = 1, \dots, N$ .
- 2: **for**  $i, k \in \mathcal{M}^+$  **do**
- 3:   Compute empirical mutual information  $I(\Delta \mathbf{V}_i^{pnz}; \Delta \mathbf{V}_k^{pnz})$  based on  $\Delta \mathbf{v}_i^{pnz}[n]$  and  $\Delta \mathbf{v}_k^{pnz}[n]$ .
- 4: **end for**
- 5: Sort all possible bus pair  $(i, k)$  into non-increasing order by  $I(\Delta \mathbf{V}_i^{pnz}, \Delta \mathbf{V}_k^{pnz})$ . Let  $\mathcal{T}$  denote the sorted set.
- 6: Repeat Step 5 to Step 16 in Algorithm 1.

The grid topology is identical in phase frame and in sequence component frame. Therefore, once the unbalanced grid topology is estimated, we can apply Lemma 4 to identify the phases of all buses.

## V. SIMULATIONS AND NUMERICAL RESULTS

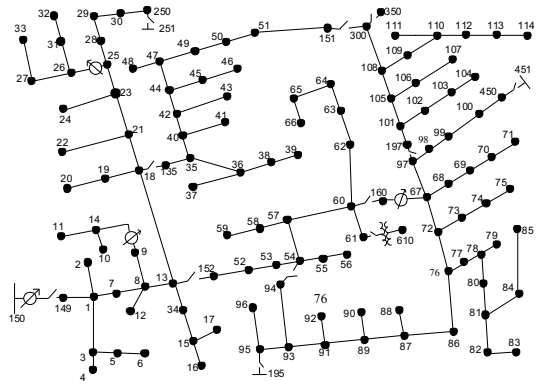


Fig. 8. IEEE 123-bus distribution test case.

The simulations are implemented on the IEEE PES distribution networks for IEEE 37-bus and 123-bus (Fig. 8) networks [18]. Both networks are three-phase systems. In each network, the feeder or substation is selected as the slack bus (bus 0). The historical data have been preprocessed by the GridLAB-D [21], an open source simulator for distribution grid. To simulate the power system behavior in a practical pattern, we use the load profile from PG&E in the subsequent simulation. This profile contains anonymized and secure hourly smart meter readings over 110,000 PG&E residential customers for a period of one year spanning from 2011 to 2012. Since both 37-bus and 123-bus systems are primary distribution grids, the real power at each bus is an aggregation of 10 – 100 customers. The load buses in both systems are unnecessary to be three-phase. The details of bus and branch phases are given in [18].

The PG&E data set does not contain the reactive power. The reactive power  $q_i^\phi[n]$  on phase  $\phi$  of bus  $i$  at time  $n$  is computed

according to a random lagging power factor  $pf_i^\phi[n]$ , which follows a uniform distribution, e.g.,  $pf_i^\phi[n] \sim \text{Unif}(0.8, 0.95)$ . To obtain voltage time-series, i.e.,  $\mathbf{v}_i^{abc}[n]$ , we run a power flow to generate the hourly states of the power system over a year.  $N = 8760$  measurements are obtained at each bus. The loads attached to each phase are unequal. Hence, the systems are unbalanced. Fig. 9 and Fig. 10 show the hourly aggregated real powers on each phase in 37-bus and 123-bus systems. Although each phase has the similar pattern over time, the magnitudes of real powers are different on each phase. Therefore, the testing systems are unbalanced.

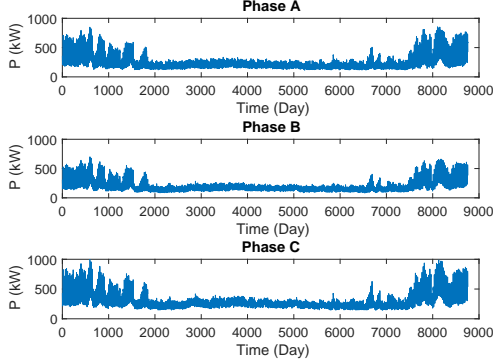


Fig. 9. Hourly aggregated real powers on each phase in 37-bus system.

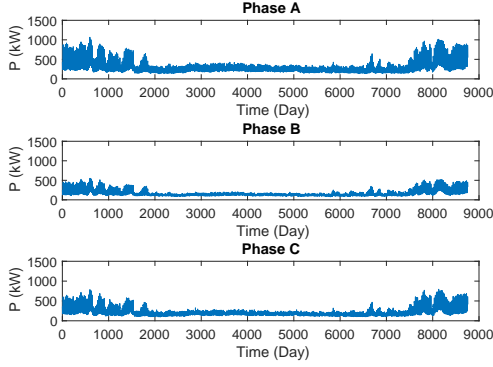


Fig. 10. Hourly aggregated real powers on each phase in 123-bus system.

#### A. Distribution Grid Topology Estimation Error Rate

In this section, we discuss the performance on grid topology estimation. We use the error rate (ER) as the performance evaluation metric, which is defined as

$$\text{ER} = \frac{1}{|\mathcal{E}|} \left( \underbrace{\sum_{(i,k) \in \hat{\mathcal{E}}} \mathbb{I}((i,k) \notin \mathcal{E})}_{\text{false estimation}} + \underbrace{\sum_{(i,k) \in \mathcal{E}} \mathbb{I}((i,k) \notin \hat{\mathcal{E}})}_{\text{missing}} \right) \%$$

where  $\hat{\mathcal{E}}$  denotes the edge set estimates,  $|\mathcal{E}|$  is the size of  $\mathcal{E}$ , and  $\mathbb{I}(\cdot)$  is the indicator function. The first and second terms represent the number of falsely estimated branches and the number of missing branches, respectively.

Table II summarizes the topology estimation error rates of unbalanced three-phase 37- and 123-bus systems. When phase angle data are available, our algorithm perfectly estimates the grid topology. Today, the PMUs have not been widely deployed. When only voltage magnitudes are available, our algorithm can still estimate the grid topology perfectly.

We also compare the proposed algorithm with a modified steady-state topology estimation in [16]. Specifically, we apply the steady-state topology estimator to each phase individually.

TABLE II. TOPOLOGY ESTIMATION ERROR RATE WITHOUT DERs.

System	Proposed Method		Modified Steady-State Method	
	$\Delta \mathbf{V}^{pnz}$	$ \Delta \mathbf{V}^{pnz} $	$\Delta \mathbf{V}^{pnz}$	$ \Delta \mathbf{V}^{pnz} $
37-bus	0.00%	0.00%	5.56%	8.33%
123-bus	0.00%	0.00%	1.64%	1.64%

Then, we combine the single-phase estimates to produce the three-phase system topology. As shown in Table II, the modified steady-state method has worse performance than the proposed algorithm. The key reason is that the modified steady-state method does not consider the voltage coupling across phases.

#### B. Distribution Grids with DER Integration

The penetration of DERs has grown significantly during last decade and will keep increasing in the future. As discussed earlier, the high penetration of DER will lead to a deeply unbalanced distribution grid. To evaluate the proposed algorithm with integrated DERs, we install several rooftop photovoltaic (PV) systems in the distribution networks. The profile of hourly power generation is obtained from NREL PVWatts Calculator, an online simulator that estimates the PV power generation based on weather history of PG&E service zone and the physical parameters of a 5kW PV panel in residential levels [35]. The power factor is fixed as 0.90 lagging, which satisfies the regulation of many U.S. utilities [36] and IEEE standard [37]. Totally, 20% of residents are installed with PV panels.

TABLE III. TOPOLOGY ESTIMATION ERROR RATE WITH DERs.

System	Proposed Method		Modified Steady-State Method	
	$\Delta \mathbf{V}^{pnz}$	$ \Delta \mathbf{V}^{pnz} $	$\Delta \mathbf{V}^{pnz}$	$ \Delta \mathbf{V}^{pnz} $
37-bus	0.00%	0.00%	8.33%	11.11%
123-bus	0.00%	0.00%	1.64%	1.64%

The error rates of grid topology estimation with the rooftop PVs integration are presented in Table III. Our algorithm does not have any performance degradation with DER integration. Also, the modified steady-state method still performs worse than the proposed method. Compared with the systems without DER, the modified steady-state method has performance degradation.

In order to further validate the proposed algorithm, we progressively increase the DER penetration level from 0% to 100%. In Fig. 11, we can observe that, besides 60% penetration of DERs, the error rates do not change with the growth of DER installation rate, which highlights the reliability of the proposed algorithm.

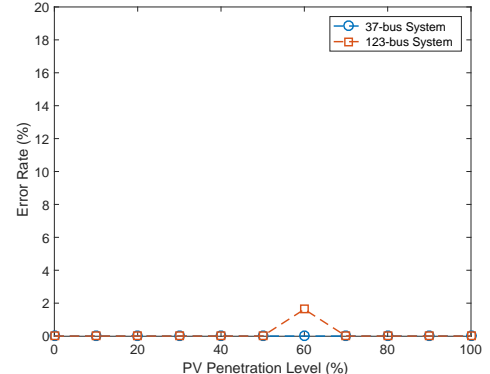


Fig. 11. Error rates with different levels of DER penetration.



### C. Distribution Grids with Incorrect Phase Labels

As we discussed in the previous section, in some distribution grids, up to 10% of the phase labels are incorrect or unknown. Therefore, in this section, we validate our algorithm on the 123-bus system with incorrect phase labels. To simulate the incorrect phase labels, we randomly pick up several buses and switch their phase  $a$  voltage measurements to either phase  $b$  voltage data or phase  $c$  voltage data.

Table IV shows the error rates with different percentages of incorrect phase labels using voltage magnitude only. We observe that our algorithm is insensitive to incorrect bus phase label. As we discussed previously, the incorrect phase labels is a permutation of random variables in  $|\mathbf{V}_i^{pnz}|$  and do not affect  $I(|\mathbf{V}_i^{pnz}|; |\mathbf{V}_k^{pnz}|)$ . If we use the steady-state approach, the error rate increases significantly because the mutual information is computed for incorrect bus pairs. For the 123-bus system, the error rate is 11.7% when 10% buses have incorrect phase labels.

TABLE IV. ERROR RATE WITH INCORRECT PHASE LABELS USING  $|\Delta \mathbf{V}^{pnz}|$ .

Percentage of Bus with Incorrect Phase Labels	Error Rate Average	Error Rate Standard Deviation
2%	0%	0%
6%	0%	0%
10%	0%	0%
14%	0%	0%
18%	0%	0%
20%	0%	0%

### D. Sensitivity Analysis

1) *Sensitivity to Data Pattern*: To understand our algorithm's sensitivity to load pattern, we validate the proposed algorithm on the "ADRES-Concept" Project load profile [19], [20]. This data set contains real and reactive power profiles of 30 houses in Upper-Austria. The data were sampled every second over 7 days in summer and 7 days in winter. We generate the voltage data using the 37-bus system. The load profiles are scaled to match the scale of power in the 37-bus system. The resulting three-phase system is unbalanced.

Fig. 12 compares the error rates using summer and winter load profiles. When there is only one measurement, the proposed algorithm has 200% error rate due to poor estimation of mutual information. The error rate is above 100% because all estimated branches are incorrect and none of the correct branch is found. As more measurements become available, the error rate reduces significantly. Also, our algorithm has consistent performance in winter and summer. Compared with the results in [16], the proposed algorithm perfectly estimates the grid topology with shorten time because more information is observed at each time step.

We also validate our algorithm using data set from Pecan Street which contains the hourly load measurements of 345 houses with PV integration in Austin, Texas. The measurements include both power consumption and renewable generation. In Fig. 13, our algorithm requires 16 hours' measurements to recover the topology of the 37-bus system, which is similar to the ADRES data set. This highlights the robustness of our algorithm.

2) *Sensitivity to Data Accuracy*: Our algorithm uses the smart meter data to estimate grid topology. In particles, the smart meter measurements are noisy. Thus, it is important to validate our algorithm under different levels of measurement noises. In the U.S., ANSI C12.20 standard (Class 0.5) requires

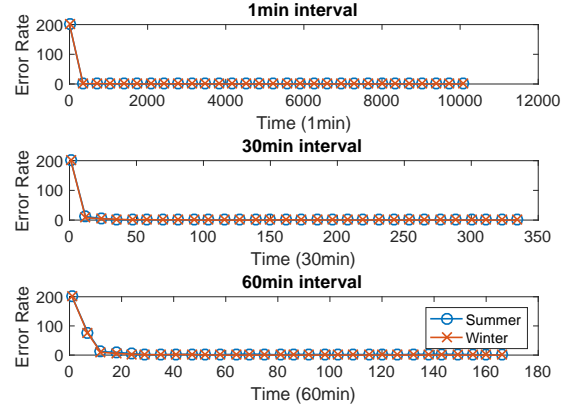


Fig. 12. Error rates with summer and winter load patterns.

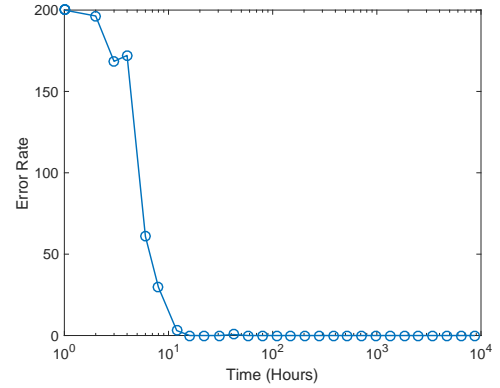


Fig. 13. Error rates on Pecan Street data set.

the smart meters to have an error less than  $\pm 0.5\%$  [38], [39]. Table. V shows the error rates with different noise levels over 20 iterations in the 123-bus system. Compared with the estimation results using perfect measurements, the error rates grow up as the increase in noise levels. These newly introduced errors are around the feeders. For example, bus 251 and 451 are incorrectly connected with the presence of noise. In real systems, we usually know the feeder buses. Therefore, we can post-process the topology estimate and remove these unnecessary branches from topology estimate. After performing post-processing, we can reduce the error rate to 1%. In Table. V, the standard deviation of error rate is very small and therefore, our algorithm can provide reliable and consistent results with noisy measurements.

TABLE V. ERROR RATES WITH DIFFERENT NOISE LEVELS.

Noise Level	Error Rate Average	Error Rate Standard Deviation
0.01%	2.95%	0.41%
0.05%	2.91%	0.50%
0.1%	2.99%	0.67%
0.2%	3.98%	0.77%

3) *Sensitivity to Data Length*: To explore the effect of data set size, we validate the proposed algorithm by using measurements from 1 up to 360 days. Fig. 14 illustrates the error rates of the 123-bus system, with and without DER, over different lengths of data. We can observe that with 20 days' measurements ( $24 \times 20 = 480$  data points), our algorithm can achieve zero error. This result is better than the single-phase system presented in [16], which requires 30 days' observations. The reason is that at time  $n$ , our proposed algorithm uses measurements from three phases, which contain more information than the single-phase system.

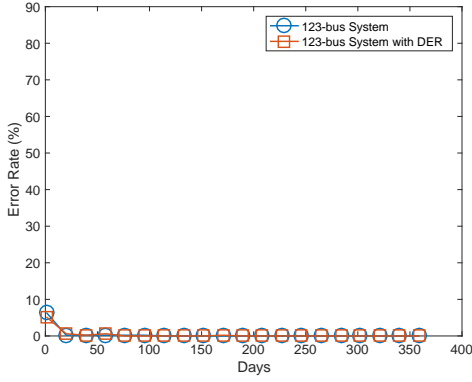


Fig. 14. Error rates with different data lengths.

4) *Sensitivity to Data Resolution:* Fig. 15 illustrates the performance of the proposed algorithm under different sampling frequencies using the ADRES data set. When the sampling period is 1 minute, we need about 6 hours' voltage profile to perfectly recover the system. The frequency of distribution grid reconfiguration is range from hours to weeks [40], [41]. Therefore, the proposed algorithm is suitable for the existing system and real-time operation. If the sampling period is 30 minutes, 35 data points ( $35 \times 30 \text{ minutes} = 17.5 \text{ hours}$ ) to recover the system topology. This estimation time is only half of the required time in [16].

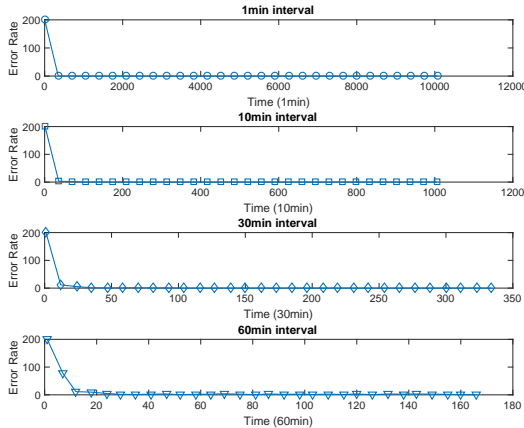


Fig. 15. Error rates with different data resolutions.

## VI. CONCLUSIONS

In order to estimate the three-phase distribution grid topology, we propose a data-driven approach based on newly available smart meter data. Unlike existing approaches, our method does not require the system to be balanced. We formulate the topology estimation problem as a joint distribution (voltage phasors) approximation problem under the probabilistic graphical model framework. Then, we prove that the distribution grid topology estimation is equivalent to the graphical model estimation problem and propose a mutual information-based maximum weight spanning tree algorithm, which is optimal and efficient. Moreover, we extend our algorithm to the case where only voltage magnitude is available. Finally, we simulate the proposed algorithm on IEEE 37- and 123-bus systems and compare with the existing steady-state method. Results show that the proposed algorithm outperforms the steady-state method and has robust performances when measurement phase labels are incorrect. We also validate our algorithms under different penetration levels of DERs and conduct the sensitivity

analysis. The numerical results are highly accurate and robust in various system configurations.

## VII. APPENDICES

### A. Proof of Lemma 1

*Proof:*

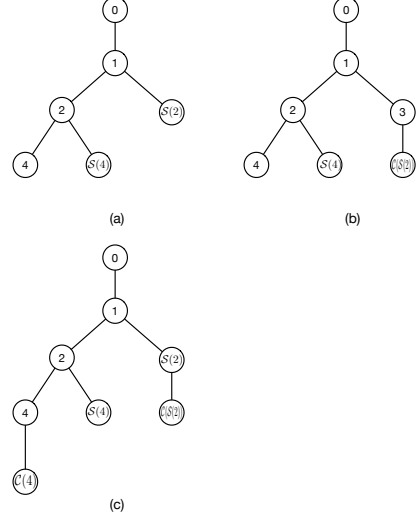


Fig. 16. Figure for the proof of Lemma 1.

We will use several cases illustrated in Fig. 16 to prove Lemma 1. We will start with the leaf node. In Fig. 16(a), for bus 4, given  $\Delta \mathbf{V}_2 = \Delta \mathbf{v}_2$ ,  $\Delta \mathbf{V}_{S(4)} = \Delta \mathbf{v}_{S(4)}$ , and  $\Delta \mathbf{V}_1 = \Delta \mathbf{v}_1$ , we have the following equations:

$$\begin{aligned} \Delta \mathbf{I}_4 &= \mathbf{Y}_{42} \Delta \mathbf{v}_2 + \mathbf{Y}_{44} \Delta \mathbf{V}_4, \\ \Delta \mathbf{I}_k &= \mathbf{Y}_{1k} \Delta \mathbf{v}_1 + \mathbf{Y}_{kk} \Delta \mathbf{V}_k \quad \forall k \in \mathcal{S}(2). \end{aligned}$$

Since  $\Delta \mathbf{I}_4 \perp \Delta \mathbf{I}_k$ , therefore,  $\Delta \mathbf{V}_4$  and  $\Delta \mathbf{V}_k$  are conditionally independent for  $k \in \mathcal{S}(2)$ . This results can be generalized to all leaf buses that share with same grandparent bus ( $\Delta \mathbf{V}_1$ ).

In Fig. 16(b), for bus 4, given  $\Delta \mathbf{V}_2 = \Delta \mathbf{v}_2$ ,  $\Delta \mathbf{V}_{S(4)} = \Delta \mathbf{v}_{S(4)}$ , and  $\Delta \mathbf{V}_1 = \Delta \mathbf{v}_1$ , we have the following equations:

$$\begin{aligned} \Delta \mathbf{I}_4 &= \mathbf{Y}_{42} \Delta \mathbf{v}_2 + \mathbf{Y}_{44} \Delta \mathbf{V}_4, \\ \Delta \mathbf{I}_3 &= \mathbf{Y}_{13} \Delta \mathbf{v}_1 + \sum_{k \in \mathcal{C}(3)} \mathbf{Y}_{3k} \Delta \mathbf{V}_k + \mathbf{Y}_{33} \Delta \mathbf{V}_3, \end{aligned} \quad (16)$$

$$\Delta \mathbf{I}_k = \mathbf{Y}_{3k} \Delta \mathbf{V}_3 + \mathbf{Y}_{kk} \Delta \mathbf{V}_k \quad \forall k \in \mathcal{C}(3). \quad (17)$$

Since  $\mathbf{Y}_{ii} = -\mathbf{Y}_{\text{pa}(i)i} - \sum_{k \in \mathcal{C}(i)} \mathbf{Y}_{ki}$  and  $\mathbf{Y}_{ik} = \mathbf{Y}_{ki}$ , if we combine (16) and (17), we have

$$\begin{aligned} &\Delta \mathbf{I}_3 + \sum_{k \in \mathcal{C}(3)} \Delta \mathbf{I}_k \\ &= \mathbf{Y}_{13} \Delta \mathbf{v}_1 + \mathbf{Y}_{33} \Delta \mathbf{V}_3 + \sum_{k \in \mathcal{C}(3)} \mathbf{Y}_{3k} \Delta \mathbf{V}_k \\ &\quad + \sum_{k \in \mathcal{C}(3)} (\mathbf{Y}_{3k} \Delta \mathbf{V}_3 + \mathbf{Y}_{kk} \Delta \mathbf{V}_k) \\ &= \mathbf{Y}_{13} \Delta \mathbf{v}_1 - (\mathbf{Y}_{13} \Delta \mathbf{V}_3 + \sum_{k \in \mathcal{C}(3)} \mathbf{Y}_{3k} \Delta \mathbf{V}_3) \\ &\quad + \sum_{k \in \mathcal{C}(3)} (\mathbf{Y}_{3k} \Delta \mathbf{V}_k + \mathbf{Y}_{3k} \Delta \mathbf{V}_3 - \mathbf{Y}_{3k} \Delta \mathbf{V}_k) \\ &= \mathbf{Y}_{13} \Delta \mathbf{v}_1 - \mathbf{Y}_{13} \Delta \mathbf{V}_3 \end{aligned} \quad (18)$$

Given  $\Delta \mathbf{I}_4 \perp (\Delta \mathbf{I}_3 + \sum_{k \in \mathcal{C}(3)} \Delta \mathbf{I}_k)$ ,  $\Delta \mathbf{V}_4$  and  $\Delta \mathbf{V}_3$  are conditionally independent. In (18), we can rewrite it as an equation of  $\Delta \mathbf{V}_3$ , i.e.,

$$\Delta \mathbf{V}_3 = (\mathbf{Y}_{13})^{-1} (\mathbf{Y}_{13} \Delta \mathbf{v}_1 - \Delta \mathbf{I}_3 - \sum_{k \in \mathcal{C}(3)} \Delta \mathbf{I}_k).$$

We can replace  $\Delta \mathbf{V}_3$  in (17) with the equations above. Specifically, for  $k \in \mathcal{C}(3)$ ,

$$\mathbf{Y}_{kk} \Delta \mathbf{V}_k + \mathbf{Y}_{3k} \Delta \mathbf{v}_1 = \Delta \mathbf{I}_k + \mathbf{Y}_{3k} (\mathbf{Y}_{13})^{-1} (\Delta \mathbf{I}_3 + \sum_{k \in \mathcal{C}(3)} \Delta \mathbf{I}_k).$$

Given  $\Delta \mathbf{I}_4$  and  $\Delta \mathbf{I}_k + \mathbf{Y}_{3k} (\mathbf{Y}_{13})^{-1} (\Delta \mathbf{I}_3 + \sum_{k \in \mathcal{C}(3)} \Delta \mathbf{I}_k)$  are independent and  $\Delta \mathbf{v}_1$  is a constant,  $\Delta \mathbf{V}_4$  and  $\Delta \mathbf{V}_k$  are conditionally independent for  $k \in \mathcal{C}(3)$ . When there are more child buses  $\Delta \mathbf{V}_{\mathcal{C}(3)}$ , we can apply the same induction method above to prove the conditional independence. Thus, the proof of Fig. 16(b) can be generalized to prove the conditional independence of a leaf bus and all other buses that are under the same grandparent bus.

Next, we will prove the lemma for non-leaf buses. In Fig. 16(c), for bus 4, given  $\Delta \mathbf{V}_2 = \Delta \mathbf{v}_2$ ,  $\Delta \mathbf{V}_{\mathcal{S}(4)} = \Delta \mathbf{v}_{\mathcal{S}(4)}$ , and  $\Delta \mathbf{V}_1 = \Delta \mathbf{v}_1$ ,

$$\Delta \mathbf{I}_4 = \mathbf{Y}_{42} \Delta \mathbf{v}_2 + \mathbf{Y}_{44} \Delta \mathbf{V}_4 + \sum_{k \in \mathcal{C}(4)} \mathbf{Y}_{4k} \Delta \mathbf{V}_k, \quad (19)$$

$$\Delta \mathbf{I}_k = \mathbf{Y}_{4k} \Delta \mathbf{V}_4 + \mathbf{Y}_{k4} \Delta \mathbf{V}_k, \quad \forall k \in \mathcal{C}(4), \quad (20)$$

$$\Delta \mathbf{I}_l = \mathbf{Y}_{1l} \Delta \mathbf{v}_1 + \sum_{m \in \mathcal{C}(l)} \mathbf{Y}_{lm} \Delta \mathbf{V}_m + \mathbf{Y}_{ll} \Delta \mathbf{V}_l, \quad (21)$$

$$\Delta \mathbf{I}_m = \mathbf{Y}_{lm} \Delta \mathbf{V}_l + \mathbf{Y}_{mm} \Delta \mathbf{V}_m, \quad (22)$$

where  $l \in \mathcal{S}(2)$  and  $m \in \mathcal{C}(l)$ . By combining (19) and (20), we have

$$\Delta \mathbf{I}_4 + \sum_{k \in \mathcal{C}(4)} \Delta \mathbf{I}_k = \mathbf{Y}_{42} \Delta \mathbf{v}_2 - \mathbf{Y}_{42} \Delta \mathbf{V}_4. \quad (23)$$

For every  $l$  in  $\mathcal{S}(2)$ , we can combine (21) and (22) and have the following equation:

$$\Delta \mathbf{I}_l + \sum_{m \in \mathcal{C}(l)} \Delta \mathbf{I}_m = \mathbf{Y}_{1l} \Delta \mathbf{v}_1 - \mathbf{Y}_{1l} \Delta \mathbf{V}_l. \quad (24)$$

We can apply the strategy in Fig. 16(b) to (23) and (24) and prove that  $\Delta \mathbf{V}_4$  and  $\Delta \mathbf{V}_l$  are conditionally independent. Also, we can prove that  $\Delta \mathbf{V}_4$  and  $\Delta \mathbf{V}_m$  are conditionally independent for  $m \in \mathcal{C}(l)$  by combining (22) and (24). The results in Fig. 16(c) can be generalized to all non-leaf buses. Using the results in Fig. 16, we can prove that Lemma 1 holds. ■

### B. Proof of Theorem 3

*Proof:* Recall the definition [27], KL divergence is expressed as

$$\begin{aligned} & D(P(\Delta \mathbf{V}_{\mathcal{M}^+}) \| Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta)) \\ &= E_{P(\Delta \mathbf{V}_{\mathcal{M}^+})} \log \frac{P(\Delta \mathbf{V}_{\mathcal{M}^+})}{Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta)} \\ &= \int P(\Delta \mathbf{V}_{\mathcal{M}^+}) \log \frac{P(\Delta \mathbf{V}_{\mathcal{M}^+})}{Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta)} \\ &= \int P(\Delta \mathbf{V}_{\mathcal{M}^+}) (\log P(\Delta \mathbf{V}_{\mathcal{M}^+}) - \log Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta)) \\ &= \int P(\Delta \mathbf{V}_{\mathcal{M}^+}) \log P(\Delta \mathbf{V}_{\mathcal{M}^+}) \\ &\quad - \int P(\Delta \mathbf{V}_{\mathcal{M}^+}) \log Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta). \end{aligned}$$

Because of Lemma 1, the radial structured PDF  $Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta)$  can be expressed as a conditional distribution

$$\begin{aligned} & \prod_{i=1}^M P(\Delta \mathbf{V}_i | \Delta \mathbf{V}_{\Theta_i}). \text{ Then,} \\ & D(P(\Delta \mathbf{V}_{\mathcal{M}^+}) \| Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta)) \\ &= \int P(\Delta \mathbf{V}_{\mathcal{M}^+}) \log P(\Delta \mathbf{V}_{\mathcal{M}^+}) \\ &\quad - \int P(\Delta \mathbf{V}_{\mathcal{M}^+}) \log \prod_{i=1}^M P(\Delta \mathbf{V}_i | \Delta \mathbf{V}_{\Theta_i}) \\ &= \int P(\Delta \mathbf{V}_{\mathcal{M}^+}) \log P(\Delta \mathbf{V}_{\mathcal{M}^+}) \\ &\quad - \int P(\Delta \mathbf{V}_{\mathcal{M}^+}) \sum_{i=1}^M \log P(\Delta \mathbf{V}_i | \Delta \mathbf{V}_{\Theta_i}), \end{aligned}$$

where  $P(\Delta \mathbf{V}_1 | \Delta \mathbf{V}_0) = P(\Delta \mathbf{V}_1)$  due to the fact that  $\Delta \mathbf{V}_0$  is a constant. By following the definition of conditional probability and adding  $P(\Delta \mathbf{V}_i)$  into the denominator, we can simplify the equation above as

$$\begin{aligned} & D(P(\Delta \mathbf{V}_{\mathcal{M}^+}) \| Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta)) \\ &= \int P(\Delta \mathbf{V}_{\mathcal{M}^+}) \log P(\Delta \mathbf{V}_{\mathcal{M}^+}) \\ &\quad - \int P(\Delta \mathbf{V}_i | \Delta \mathbf{V}_{\Theta_i}) \sum_{i=1}^M \log \frac{P(\Delta \mathbf{V}_i, \Delta \mathbf{V}_{\Theta_i})}{P(\Delta \mathbf{V}_i) P(\Delta \mathbf{V}_{\Theta_i})} \\ &\quad - \sum_{i=1}^M \int P(\Delta \mathbf{V}_i) \log P(\Delta \mathbf{V}_i) \\ &= -H(\Delta \mathbf{V}_{\mathcal{M}^+}) - \sum_{i=1}^M I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_{\Theta_i}) + \sum_{i=1}^M H(\Delta \mathbf{V}_i). \end{aligned}$$

The last equality is due to the definitions of entropy, i.e.,

$$H(\Delta \mathbf{V}_i) = - \int P(\Delta \mathbf{V}_i) \log P(\Delta \mathbf{V}_i), \quad (25)$$

and mutual information, i.e.,

$$\begin{aligned} & I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_{\Theta_i}) \\ &= \int P(\Delta \mathbf{V}_i, \Delta \mathbf{V}_{\Theta_i}) \log \frac{P(\Delta \mathbf{V}_i, \Delta \mathbf{V}_{\Theta_i})}{P(\Delta \mathbf{V}_i) P(\Delta \mathbf{V}_{\Theta_i})}. \quad (26) \end{aligned}$$

Thus, to minimize the KL-divergence between  $P(\Delta \mathbf{V}_{\mathcal{M}^+})$  and  $Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta)$ , we can choose the  $M - 1$  edges to maximize  $\sum_{i=1}^M I(P(\Delta \mathbf{V}_i); P(\Delta \mathbf{V}_{\Theta_i}))$ . The entropy term  $\sum_{i=1}^M H(\Delta \mathbf{V}_i) - H(\Delta \mathbf{V}_{\mathcal{M}^+})$  is irrelevant with the topology structure of distribution grid and we do not include it in the final optimization problem. Therefore, minimizing  $D(P(\Delta \mathbf{V}_{\mathcal{M}^+}) \| Q(\Delta \mathbf{V}_{\mathcal{M}^+}; \Theta))$  is equivalent to solving the following optimization problem:

$$\hat{\Theta} = \underset{\Theta \subset \mathcal{M}^+}{\operatorname{argmax}} \sum_{i=1}^M I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_{\Theta_i}). \quad \blacksquare$$

### REFERENCES

- [1] S. Dey, A. Jessa, and L. Gelbien, "Urban grid monitoring renewables integration," in *IEEE Conference on Innovative Technologies for an Efficient and Reliable Electricity Supply*, 2010, pp. 252–256.
- [2] K. Clement-Nyns, E. Haesen, and J. Driesen, "The impact of charging plug-in hybrid electric vehicles on a residential distribution grid," *Power Systems, IEEE Transactions on*, vol. 25, no. 1, pp. 371–380, 2010.
- [3] J. Huang, V. Gupta, and Y.-F. Huang, "Electric grid state estimators for distribution systems with microgrids," in *IEEE 46th Annual Conference on Information Sciences and Systems*, 2012, pp. 1–6.
- [4] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.

- [5] R. Lugtu, D. Hackett, K. Liu, and D. Might, "Power system state estimation: Detection of topological errors," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-99, no. 6, pp. 2406–2412, Nov 1980.
- [6] C. Rudin, D. Waltz, R. Anderson, A. Boulanger, A. Salieb-Aouissi, M. Chow, H. Dutta, P. Gross, B. Huang, S. Ierome, D. Isaac, A. Kressner, R. Passonneau, A. Radeva, and L. Wu, "Machine learning for the new york city power grid," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 2, pp. 328–345, Feb 2012.
- [7] D. Deka, M. Chertkov, and S. Backhaus, "Structure learning in power distribution networks," *IEEE Transactions on Control of Network Systems*, vol. PP, no. 99, pp. 1–1, 2017.
- [8] D. Deka, S. Backhaus, and M. Chertkov, "Estimating distribution grid topologies: A graphical learning based approach," in *Power Systems Computation Conference*. IEEE, 2016, pp. 1–7.
- [9] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," in *Conference on Decision and Control*, 2013, pp. 1659–1664.
- [10] Y. Liao, Y. Weng, and R. Rajagopal, "Urban distribution grid topology reconstruction via lasso," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, July 2016, pp. 1–5.
- [11] J. Peppanen, S. Grijalva, M. J. Reno, and R. J. Broderick, "Distribution system low-voltage circuit topology estimation using smart metering data," in *Transmission and Distribution Conference and Exposition*. IEEE, 2016, pp. 1–5.
- [12] K. L. Lo and C. Zhang, "Decomposed three-phase power flow solution using the sequence component frame," in *IEE Proceedings C (Generation, Transmission and Distribution)*, vol. 140, no. 3. IET, 1993, pp. 181–188.
- [13] N. Tleis, *Power systems modelling and fault analysis: theory and practice*. Newnes, 2007.
- [14] American National Standards Institute, "ANSI c84.1: Electric power systems and equipment voltage ratings (60 herz)," American National Standards Institute, Tech. Rep., 1995.
- [15] T. Routtenberg, Y. Xie, R. M. Willett, and L. Tong, "Pmu-based detection of imbalance in three-phase power systems," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1966–1976, 2015.
- [16] Y. Weng, Y. Liao, and R. Rajagopal, "Distributed energy resources topology identification via graphical modeling," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2682–2694, 2017.
- [17] W. H. Kersting, *Distribution system modeling and analysis*. CRC press, 2006.
- [18] —, "Radial distribution test feeders," in *IEEE Power Engineering Society Winter Meeting*, vol. 2, 2001, pp. 908–912.
- [19] A. Einfalt, A. Schuster, C. Leitingner, D. Tiefgraber, M. Litzlbauer, S. Ghaemi, D. Wertz, A. Frohner, and C. Karner, "Adres-concept: Konzeptentwicklung für adres-autonome dezentrale regenerative energiesysteme," *TU Wien, Institut für Elektrische Anlagen und Energiewirtschaft*, 2011.
- [20] Institute of Energy Systems and Electrical Drives, "Adres-dataset," Tech. Rep., 2016. [Online]. Available: [http://www.ea.tuwien.ac.at/projects/adres\\_concept/EN/](http://www.ea.tuwien.ac.at/projects/adres_concept/EN/)
- [21] D. P. Chassin, K. Schneider, and C. Gerkensmeyer, "Gridlab-d: An open-source power systems modeling and simulation environment," in *Transmission and distribution conference and exposition, 2008. t&d. IEEE/PES*. IEEE, 2008, pp. 1–5.
- [22] M. Laughton, "Analysis of unbalanced polyphase networks by the method of phase co-ordinates. part 1: System representation in phase frame of reference," in *Proceedings of the Institution of Electrical Engineers*, vol. 115, no. 8. IET, 1968, pp. 1163–1172.
- [23] T.-H. Chen, M.-S. Chen, K.-J. Hwang, P. Kotas, and E. A. Chebli, "Distribution system power flow analysis-a rigid approach," *IEEE Transactions on Power Delivery*, vol. 6, no. 3, pp. 1146–1152, 1991.
- [24] W. H. Kersting, *Distribution system modeling and analysis*. CRC press, 2012.
- [25] W. Wang, N. Yu, B. Foggo, J. Davis, and J. Li, "Phase identification in electric power distribution systems by clustering of smart meter data," in *Machine Learning and Applications (ICMLA), 2016 15th IEEE International Conference on*. IEEE, 2016, pp. 259–265.
- [26] Y. C. Chen, T. Banerjee, A. D. Domínguez-García, and V. V. Veeravalli, "Quickest line outage detection and identification," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 749–758, 2016.
- [27] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [28] Y. Liao, Y. Weng, G. Liu, and R. Rajagopal, "Urban mv and lv distribution grid topology estimation via group lasso," *IEEE Transactions on Power Systems*, pp. 1–1, 2018.
- [29] D. Deka, S. Talukdar, M. Chertkov, and M. Salapaka, "Topology estimation in bulk power grids: Guarantees on exact recovery," *arXiv preprint arXiv:1707.01596*, 2017.
- [30] C. Chow and C. Liu, "Approximating discrete probability distributions with dependence trees," *IEEE Transactions on Information Theory*, vol. 14, no. 3, pp. 462–467, 1968.
- [31] T. A. Short, "Advanced metering for phase identification, transformer identification, and secondary modeling," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 651–658, 2013.
- [32] V. Arya, D. Seetharam, S. Kalyanaraman, K. Dontas, C. Pavlovski, S. Hoy, and J. R. Kalagnanam, "Phase identification in smart grids," in *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*. IEEE, 2011, pp. 25–30.
- [33] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proceedings of the American Mathematical society*, vol. 7, no. 1, pp. 48–50, 1956.
- [34] T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein *et al.*, *Introduction to algorithms*. MIT press Cambridge, 2001, vol. 2.
- [35] A. P. Dobos, "Pvwatts version 5 manual," *National Renewable Energy Laboratory, September*, 2014.
- [36] A. Ellis, R. Nelson, E. Von Engeln, R. Walling, J. MacDowell, L. Casey, E. Seymour, W. Peter, C. Barker, B. Kirby *et al.*, "Review of existing reactive power requirements for variable generation," in *Power and Energy Society General Meeting, 2012 IEEE*. IEEE, 2012, pp. 1–7.
- [37] "IEEE guide for conducting distribution impact studies for distributed resource interconnection," *IEEE Std 1547.7-2013*, pp. 1–137, Feb 2014.
- [38] J. Zheng, D. W. Gao, and L. Lin, "Smart meters in smart grid: An overview," in *IEEE Green Technologies Conference*, April 2013, pp. 57–64.
- [39] National Electrical Manufacturers Association, "ANSI c12.20-2010: American national standard for electricity meter: 0.2 and 0.5 accuracy classes," American National Standards Institute, Tech. Rep., 2010.
- [40] R. A. Jabr, "Minimum loss operation of distribution networks with photovoltaic generation," *IET Renewable Power Generation*, vol. 8, no. 1, pp. 33–44, 2014.
- [41] M. R. Dorostkar-Ghamsari, M. Fotuhi-Firuzabad, M. Lehtonen, and A. Safdarian, "Value of distribution network reconfiguration in presence of renewable energy resources," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 1879–1888, 2016.