Unbalanced Multi-Phase Distribution Grid Topology Estimation and Bus Phase Identification

Yizheng Liao[1], Yang Weng[2], Guangyi Liu[3], Zhongyang Zhao[3], Chin-Woo Tan[1], Ram Rajagopal[1]

[1] *Department of Civil and Environmental Engineering, Stanford University, Stanford, CA, 94305 USA*
[2] *School of Electrical, Computing, and Energy Engineering, Arizona State University, Tempe, AZ, 85287 USA*
[3] *GEIRI North America, San Jose, CA, 95134, USA*
* *E-mail: yzliao@stanford.edu*

Abstract: There is an increasing need for monitoring and controlling uncertainties brought by distributed energy resources in distribution grids. For such goal, accurate multi-phase topology is the basis for correlating measurements in unbalanced distribution networks. Unfortunately, such topology knowledge is often unavailable due to limited investment, especially for low-voltage distribution grids. Also, the bus phase labeling information is inaccurate due to human errors or outdated records. For this challenge, this paper utilizes smart meter data for an information-theoretic approach to learn the topology of distribution grids. Specifically, multi-phase unbalanced systems are converted into symmetrical components, namely positive, negative, and zero sequences. Then, this paper proves that the Chow-Liu algorithm finds the topology by utilizing power flow equations and the conditional independence relationships implied by the radial multi-phase structure of distribution grids with the presence of incorrect bus phase labels. At last, by utilizing Carson's equation, this paper proves that the bus phase connection can be correctly identified using voltage measurements. For validation, IEEE systems are simulated using three real data sets. The simulation results demonstrate that the algorithm is highly accurate for finding multi-phase topology even with strong load unbalancing condition and DERs. This ensures close monitoring and controlling DERs in distribution grids.

1 Introduction

The power distribution system is currently undergoing a dramatic transformation in both forms and functions. Large-scale deployments 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, the integration of these resources has been proven to be nontrivial, largely because of their inherent uncertainty and distributed nature.

For example, even a small-scale of distributed energy resources (DERs) can affect the stability of distribution grids [1]. Such a problem will be aggravated by the unbalance situation in distribution grids 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 operations. In addition to the static connectivity, mobile components, such as EVs, can further jeopardize the grid stability due to their frequent plug-in [2]. Therefore, the multi-phase grid monitoring tools need to be carefully designed for islanding and line work hazards in system operation with deep and uneven DER penetrations. For such monitoring, grid topology information is a prerequisite.

For topology estimation, the transmission grid assumes a prior knowledge of grids, 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. For example, in [7], the correct topology is searched from a set of possible radial networks. Given the line parameters, [8] and [9] propose maximum likelihood methods to select the operational distribution grid topology. [10], [11], and [12] utilize the statistical correlation of single-phase voltages collected from smart meters to estimate distribution grid topology. Unfortunately, all of these methods focus on the balanced or single-phase systems. For utility practice, distribution grids for buildings and residential areas are usually unbalanced and multi-phase systems. One reason is that the loads connected at different phases are unbalanced due to the uneven growth in each feeder territory [13, 14]. For example, surveyed by the American National Standards Institute (ANSI), 2% of distribution grids in the USA have a significant undesirable degree of unbalance [15, 16]. With the growth of renewable penetration, the load unbalance problem will become more frequent in future distribution grids. For example, the unbalance of the multi-phase system appears more often because the installations and operations of many DER devices are not fully controlled by utilities. This fact makes the requirement of balanced grids in previous works invalid in field applications.

In order to find the topology of unbalanced multi-phase distribution grids, [17] and its follow-up work [18] formulate multi-phase measurements as vectors and apply the single-phase approach to estimate grid topology. In [19], the multi-phase power flow equations are linearized and the topology estimation is formulated as a statistical learning problem. For all these approaches, a prerequisite is installations of Phase Measurement Units (PMUs), which have not been widely available in distribution grids. In addition, these methods assume bus phase labelings are correct at each bus. For many utilities, as high as 10% phase labelings are incorrect or unknown because of human errors or outdated records. This high error rate makes identifying new topology based on existing methods not sound anymore.

For resolving the problems above, this paper proposes a data-driven method that utilizes the smart meter data in different phases to estimate the topology of multi-phase distribution grid systems. Building on our previous works on the probabilistic graphical model formulation of distribution grids [20], firstly, this paper expands

the method from the single-phase representation to multi-phase balanced systems with incorrect bus phase labels. In such model, a node represents the multi-phase bus voltages and an edge between nodes indicates the statistical dependency among multi-phase bus voltage measurements.

Subsequently, the system of three unbalanced phasors is converted to three symmetrical components, namely the positive, negative, and zero sequences. Then, the Chow-Liu algorithm is proved to be optimal for identifying the multi-phase topology by utilizing power flow equations and the conditional independence relationships implied by the radial multi-phase structure of distribution grids. As a highlight, the proposed method does not require PMUs and is robust to incorrect phase labels, which is a critical problem in distribution grid operations. This feature is due to the label-invariant property of mutual information. Another major contribution is bus phase correction and identification. Specifically, a data-driven approach is proposed to identify true bus phase connections by utilizing Carson's equation [21], which is employed for deriving the primitive phase impedances of different lines.

The performance of the proposed method is verified by simulations on the IEEE 37-bus, 123-bus, and 8500-bus distribution test cases [22]. Three different data sets are used for simulation: North California PG&E residential household data sets, ADRES project data set [23, 24] 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, Taxes. Simulations are conducted via GridLAB-D, an open source distribution grid simulator [25] for multi-phase systems. Simulation results show that, provided with hourly measurements, the proposed algorithm perfectly estimates the topology of multi-phase distribution grids with noiseless measurements.

The rest of the paper is organized as follows: Section 2 introduces the modeling of the multi-phase distribution system and the problem of data-driven topology estimation. Section 3 firstly proves the topology estimation problem of a multi-phase distribution grid can be solved as a mutual information maximization problem and proposes an algorithm to solve such a maximization problem in multi-phase setup. Also, a method is proposed to identify the bus phase connection. In Section 4, to address the unbalance in distribution grids, an unbalanced distribution grid is transformed 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 5 evaluates the performance of our method using IEEE test cases and real data collected from different regions. Section 6 concludes this paper.

2 Multi-phase Distribution Grid Modeling and Problem Formulation

A distribution grid is modeled by a graph $\mathcal{G} = (\mathcal{M}, \mathcal{E})$, where the vertex set $\mathcal{M} = \{0, 1, 2, \cdots, M\}$ 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 multi-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., $\mathcal{M}^+ = \mathcal{M} \backslash \{0\}$. 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. Let $\mathrm{pa}(i)$ denote the parent bus of bus i. The root has no parent and all other buses in $\mathcal{M}^+$ have exactly one parent. Let $\mathcal{C}(i)$ denote the set of child buses of bus i and use $\mathcal{S}(i) = \mathcal{C}(\mathrm{pa}(i)) \backslash \{i\}$ 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^ϕ denotes the line-to-ground complex voltage on phase ϕ 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, $\mathbf{V}_i$, $\mathbf{I}_i$, and $\mathbf{S}_i$ are used 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 currents can be expressed as follows [26, 27]:

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

where $\mathbf{Y}_{ik} \in \mathbb{C}^{3\times3}$ denotes the admittance submatrix between bus i and bus k and $\mathbf{B}_{i,\mathrm{shunt}} \in \mathbb{C}^{3\times3}$ denotes the shunt capacitance at bus i. In a multi-phase system, $\mathbf{Y}_{ik}$ is not diagonal. The voltages at different phases are coupled. As shown in Section 5, this coupling property in multi-phase systems leads the existing single-phase methods to have poor performance in unbalanced multi-phase systems. Since $\mathbf{B}_{i,\mathrm{shunt}}$ is relatively small in distribution grids [28], $\mathbf{B}_{i,\mathrm{shunt}}$ is assumed to be zeros, i.e., $\mathbf{0}$. In the formulation above, the effect of the neural wire is merged into the multi-phase wires by applying Kron's reduction [27]. If bus i and bus k are not connected, $\mathbf{Y}_{ik} = \mathbf{0}$.

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 $v_i^\phi[n] = |v_i^\phi[n]| \exp\left(j\theta_i^\phi[n]\right)$ denotes the complex voltage measurement on phase ϕ 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 5, the proposed algorithm will be validated with noisy measurements. In the following part, the upper-case letter denotes the symbol and the lower-case letter denotes the snapshot of symbol measurement. For example, $\mathbf{V}$ denotes the voltage symbol and $\mathbf{v}[n]$ denotes the voltage measurement at time n.

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

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

3 Multi-Phase Distribution Grid Topology Estimation and Bus Phase Identification

This section firstly extends our previous work [20] to estimate the topology of multi-phase system with incorrect phase labels. Then, a novel method is proposed to identify the true bus phase labels by utilizing the statistical relationship between voltage measurements. The method proposed in this section focuses on balanced distribution systems. When a distribution system is unbalanced, a modified algorithm is proposed in Section 4. Fig. 1 summarizes the criteria for the topology estimation method selection.

The end-user measurements are time-series data. One way to represent these data is using a probability distribution. If the nodal multi-phase voltage vector $\mathbf{V}_i$ is modeled as a random vector, the joint distribution of voltage measurement $P(\mathbf{V}_{\mathcal{M}^+})$ is $P(\mathbf{V}_1)P(\mathbf{V}_2|\mathbf{V}_1) \cdots P(\mathbf{V}_M|\mathbf{V}_1, \cdots, \mathbf{V}_{M-1})$. Bus 0 is omitted because it is the slack bus with a fixed voltage.

Many previous works of distribution grid topology estimation [7, 10, 20, 29] only require the single-phase voltages. However, with the presence of false or unknown phase labels, all three phases voltage measurements are needed for topology estimation. The latter part of this section will show that our method is invariant to phase label accuracy and therefore can estimate topology with false or unknown phase labels.

In many medium- and low-voltage distribution grids, the probability distribution of voltage is irregular. To better formulate the topology estimation problem, the incremental change of measurements is adopted in this paper [12, 30, 31]. At bus i, the incremental
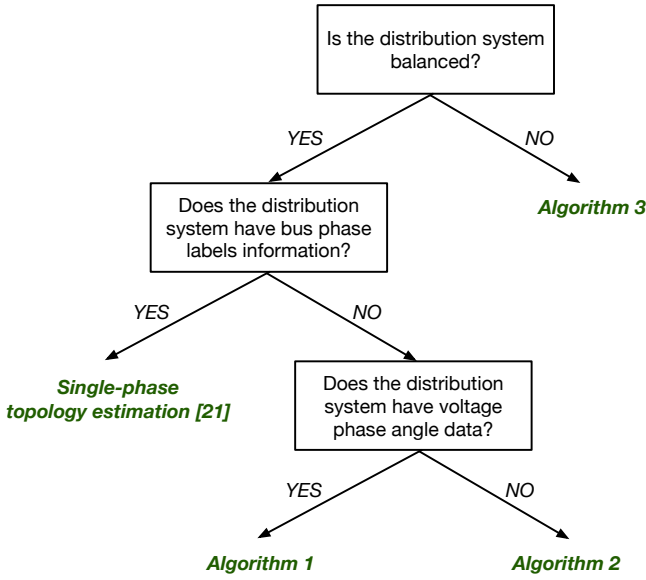
Fig. 1: Flow chart of topology estimation method selection.

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 is

$$
\begin{aligned}
P(\Delta\mathbf{V}_{\mathcal{M}+}) &= P(\Delta\mathbf{V}_1)P(\Delta\mathbf{V}_2|\Delta\mathbf{V}_1)\cdots \\
&\times P(\Delta\mathbf{V}_M|\Delta\mathbf{V}_1,\cdots,\Delta\mathbf{V}_{M-1}). \quad (2)
\end{aligned}
$$

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 a single-phase distribution grid, the nodal voltages only have statistical dependency with the nodal voltages of their parent bus [20]. In the next part, such dependency will be extended Revised on from single-phase systems to approximate multi-phase systems' joint probability $P(\Delta\mathbf{V}_{\mathcal{M}+})$ as

$$
P(\Delta\mathbf{V}_{\mathcal{M}+}) \simeq \prod_{i=1}^{M} P(\Delta\mathbf{V}_i|\Delta\mathbf{V}_{\mathrm{pa}(i)}). \quad (3)
$$

If (3) holds, finding the structure of $\mathcal{G}$ is equivalent to finding the parent of each bus. The next part uses a two-stage approach to prove the approximation in (3) holds with equality. In the first stage, bus voltages are proved to be conditionally independent, given their parents, grandparents, and siblings, i.e., $P(\Delta\mathbf{V}_{\mathcal{M}+}) = \prod_{i=1}^{M} P(\Delta\mathbf{V}_i|\Delta\mathbf{V}_{\{\mathrm{pa}(i),\mathrm{pa}(\mathrm{pa}(i)),\mathcal{S}(i)\}})$. Then, inspiring by the real data observation, (3) is shown to holds with equality.

Before starting the first stage proof, two assumptions are proposed and justified using real data.

Assumption 1. *In a multi-phase distribution gird,*

- *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., $\mathbf{X} \perp \mathbf{Y}$ [32]. In Fig. 2, most pairs of $\Delta\mathbf{I}$ have small values. Thus, the current injections are assumed to be independent with some approximation errors. This assumption has also been adopted in other works, e.g., [7, 10, 19]. To further validate the independence of $\Delta\mathbf{I}$, Fig. 3 plots the average auto-correlation of current injection increment of PG&E data in the IEEE 123-bus system. The error bar is one standard deviation. In Fig. 3, the auto-correlation of $\Delta\mathbf{I}$ drops significantly as the lag increases. This observation justifies that the current injection increments are approximately independent over time.
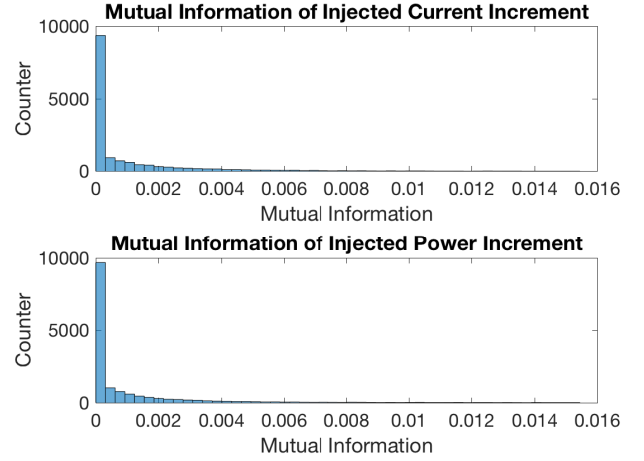

Fig. 2: Mutual information of pairwise current injection increment $\Delta\mathbf{I}$ and power injection increment $\Delta\mathbf{S}$ of PG&E data sets in the IEEE 123-bus system.
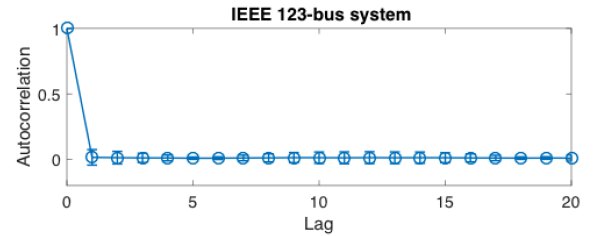

Fig. 3: Average auto-correlation of current injection increment $\Delta\mathbf{I}$ of PG&E data sets in the IEEE 123-bus system. The error bar is one standard deviation.

Both injected power increment independence and injected current increment independence are adopted in the existing works of distribution grid topology estimation. [12] uses the real data to show that these two assumptions are equivalent in distribution grids. Fig. 2 illustrates the mutual information of pairwise power injection increment $\Delta\mathbf{S}$ and pairwise current injection increment $\Delta\mathbf{I}$. Both histograms are similar. In this paper, the assumption of current injection independence is preferred because it simplifies the proof of following theorems and lemmas.

Fig. 4 illustrates the histograms of bus voltage $|\Delta\mathbf{V}|$ in IEEE 123-bus system using PG&E data. Hence, the voltage data approximately follow Gaussian distributions. With Assumption 1, $P(\Delta\mathbf{V}_{\mathcal{M}+})$ is proved to be $\prod_{i=1}^{M} P(\Delta\mathbf{V}_i|\Delta\mathbf{V}_{\{\mathrm{pa}(i),\mathrm{pa}(\mathrm{pa}(i)),\mathcal{S}(i)\}})$. For connivance, let $\mathcal{C}^{\infty}(i)$ denote all buses that are below bus i. For example, in Fig. 5, $\mathcal{C}^{\infty}(1) = \{2,3,4,5,6,7\}$ and $\mathcal{C}^{\infty}(2) = \{4,5\}$.

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*
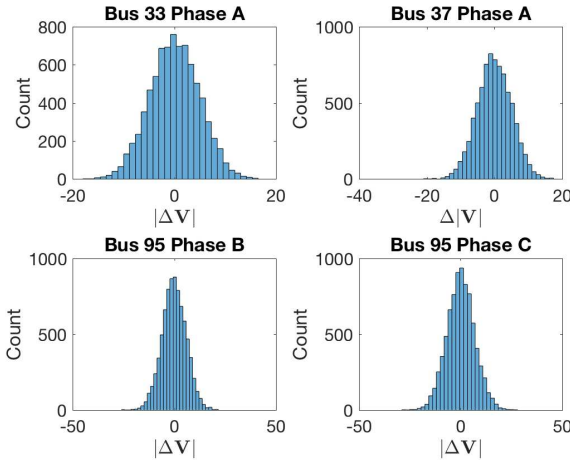
Fig. 4: Histograms of $|\Delta\mathbf{V}|$ of four buses in IEEE 123-bus system using PG&E data.
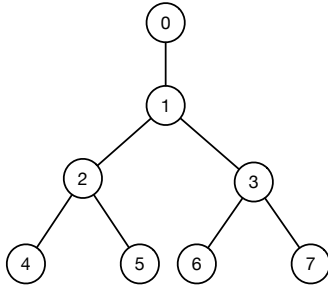

Fig. 5: An example of an 8-bus multi-phase system. A node represents a bus, which can be single-phase or multi-phase. An edge represents a branch between two buses. The branch is unnecessary to be multi-phase. Bus 0 is the substation (root).

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

Here, a simple example demonstrates Lemma 1. A formal proof is given in Appendix section 7.1. For the example system in Fig. 5, the nodal admittance equation is $\mathbf{Y}_{\mathcal{M}+}\Delta\mathbf{V}_{\mathcal{M}+} = \Delta\mathbf{I}_{\mathcal{M}+}$, where

$$\mathbf{Y}_{\mathcal{M}+} = \begin{bmatrix} \mathbf{Y}_{11} & \mathbf{Y}_{12} & \mathbf{Y}_{13} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{Y}_{21} & \mathbf{Y}_{22} & \mathbf{0} & \mathbf{Y}_{24} & \mathbf{Y}_{25} & \mathbf{0} & \mathbf{0} \\ \mathbf{Y}_{31} & \mathbf{0} & \mathbf{Y}_{33} & \mathbf{0} & \mathbf{0} & \mathbf{Y}_{36} & \mathbf{Y}_{37} \\ \mathbf{0} & \mathbf{Y}_{42} & \mathbf{0} & \mathbf{Y}_{44} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}_{52} & \mathbf{0} & \mathbf{0} & \mathbf{Y}_{55} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Y}_{63} & \mathbf{0} & \mathbf{0} & \mathbf{Y}_{66} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Y}_{73} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Y}_{77} \end{bmatrix}, \quad (4)$$

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

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

$$\Delta\mathbf{I}_1 = \mathbf{Y}_{11}\Delta\mathbf{v}_1 + \mathbf{Y}_{12}\Delta\mathbf{v}_2 + \mathbf{Y}_{13}\Delta\mathbf{V}_3, \quad (5)$$

$$\Delta\mathbf{I}_4 = \mathbf{Y}_{42}\Delta\mathbf{v}_2 + \mathbf{Y}_{44}\Delta\mathbf{V}_4, \quad (6)$$

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

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

Given $\Delta\mathbf{I}_1 \perp \Delta\mathbf{I}_4$, according to (5) and (6), $\Delta\mathbf{V}_3$ and $\Delta\mathbf{V}_4$ are conditionally independent given $\Delta\mathbf{V}_1$, $\Delta\mathbf{V}_2$, and $\Delta\mathbf{V}_5$. In (5), $\Delta\mathbf{V}_3$ can be rewritten as $(\mathbf{Y}_{13})^{-1}(\Delta\mathbf{I}_1 - \mathbf{Y}_{11}\Delta\mathbf{v}_1 - \mathbf{Y}_{12}\Delta\mathbf{v}_2)$. Then, $\Delta\mathbf{V}_3$ is substituted into (7). Since $\Delta\mathbf{I}_1$, $\Delta\mathbf{I}_4$, and $\Delta\mathbf{I}_6$ are independent, $\Delta\mathbf{I}_4$ 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 $\Delta\mathbf{V}_1 = \Delta\mathbf{v}_1$, there are the following equations:

$$\Delta\mathbf{I}_1 = \mathbf{Y}_{11}\Delta\mathbf{v}_1 + \mathbf{Y}_{12}\Delta\mathbf{V}_2 + \mathbf{Y}_{13}\Delta\mathbf{v}_3, \quad (9)$$

$$\Delta\mathbf{I}_6 = \mathbf{Y}_{63}\Delta\mathbf{v}_3 + \mathbf{Y}_{66}\Delta\mathbf{V}_6, \quad (10)$$

$$\Delta\mathbf{I}_7 = \mathbf{Y}_{73}\Delta\mathbf{v}_3 + \mathbf{Y}_{77}\Delta\mathbf{V}_7. \quad (11)$$

Given $\Delta\mathbf{I}_1$ and $\Delta\mathbf{I}_6$ are independent, according to (9) and (10), $\Delta\mathbf{V}_2$ 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 [33].

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 $\Delta\mathbf{V}_{\mathcal{S}(i)}$.*
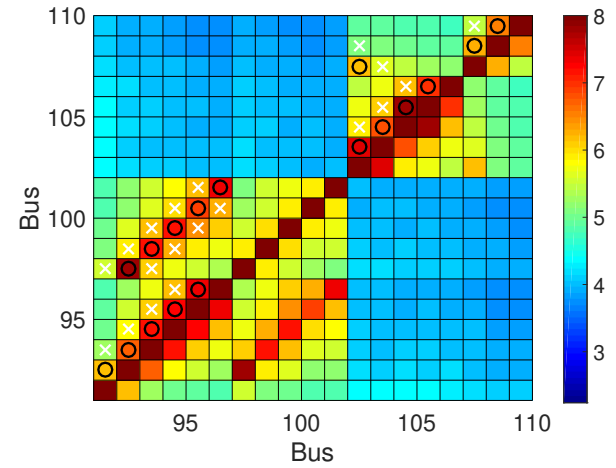

Fig. 6: Mutual information of pairwise buses in IEEE 123-bus system using PG&E data sets. The circle indicates the neighbors of bus i. The crossing indicates the two-step neighbor of bus i. The square without markers represents the bus pair that are more than two-step away.

Assumption 2 is inspired by the real data observations. Fig. 6 plots the mutual information of voltage increments between each bus pair in IEEE 123-bus distribution system using the PG&E data. The distribution grid configuration and simulation setup are described in Section 5. In Fig. 6, the color in a square represents the mutual information of voltage increments between two buses. If the voltage increments of two buses are independent, their mutual information is zero[32] (dark color). In Fig. 6, 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 does not have any marker, the corresponding pair of buses is more than two-step away. In Fig. 6, the mutual information between the voltages of 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, $P(\Delta\mathbf{V}_{\mathcal{M}_+})$ can be simplified to only depend on the voltages

of parent buses. Section 5 uses numerical simulations to demonstrate that this 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 $k, l \in \mathcal{C}(i)$ and $k \neq l$.*

With Lemma 2, (3) holds with equality, i.e., $P(\Delta \mathbf{V}_{\mathcal{M}+}) = \prod_{i=1}^{M} P(\Delta \mathbf{V}_i | \Delta \mathbf{V}_{\mathrm{pa}(i)})$. Thus, finding the distribution grid topology is equivalent to finding the parent of each bus. In the following subsections, an information theoretical approach is proposed to estimate the multi-phase distribution grid topology with incorrect bus phase labels.

3.1 An Information Theoretical Approach to Estimate Multi-phase Distribution Grid Topology

One way to find the parent of each bus is minimizing the Kullback-Leibler divergence [32] of $P(\Delta \mathbf{V}_{\mathcal{M}+})$ and $Q(\Delta \mathbf{V}_{\mathcal{M}+}) = \prod_{i=1}^{M} P(\Delta \mathbf{V}_i | \Delta \mathbf{V}_{\mathrm{pa}(i)})$, i.e.,

$$\widehat{\mathbf{\Theta}} = \underset{\mathbf{\Theta} \subset \mathcal{M}+}{\arg\min} \, D(P(\Delta \mathbf{V}_{\mathcal{M}+}) \| Q(\Delta \mathbf{V}_{\mathcal{M}+}; \mathbf{\Theta})), \quad (12)$$

where $\mathbf{\Theta}$ denotes the collection of parent bus index of every bus, i.e., $\mathbf{\Theta} = \{\mathrm{pa}(1), \cdots, \mathrm{pa}(M)\}$, 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 there exists a distribution $Q(\Delta \mathbf{V}_{\mathcal{M}+}; \widehat{\mathbf{\Theta}})$ that is identical to $P(\Delta \mathbf{V}_{\mathcal{M}+})$, $\widehat{\mathbf{\Theta}}$ contains the parent bus index of every bus i. The associated structure of $P_{CL}(\Delta \mathbf{V}_{\mathcal{M}+}) = Q(\Delta \mathbf{V}_{\mathcal{M}+}; \widehat{\mathbf{\Theta}})$ is the estimated topology of distribution grid. Lemma 3 proves that (12) can be efficiently solved by utilizing the radial structure of distribution grids. In the following context, $\mathbf{\Theta}_i$ and $\mathrm{pa}(i)$ are used interchangeably.

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

$$\widehat{\mathbf{\Theta}} = \underset{\mathbf{\Theta} \subset \mathcal{M}+}{\arg\max} \sum_{i=1}^{M} I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_{\mathbf{\Theta}_i}), \quad (13)$$

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

The proof is in Appendix 7.2. With Lemma 3, a mutual information-based maximum weight spanning tree algorithm, well-known as Chow-Liu algorithm [34], could find $\widehat{\mathbf{\Theta}}$ and identify the multi-phase distribution grid topology. This algorithm has been applied to single-phase system in [20]. Theorem 1 proves that Chow-Liu algorithm can be extended to multi-phase systems.

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

Proof: This proof shows 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 [32], the joint mutual information is expressed as

$$
\begin{aligned}
& I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k, \Delta \mathbf{V}_l) \\
= \; & I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k) - I(\Delta \mathbf{V}_i, \Delta \mathbf{V}_k | \Delta \mathbf{V}_l), \\
= \; & I(\Delta \mathbf{V}_k; \Delta \mathbf{V}_l) - I(\Delta \mathbf{V}_k, \Delta \mathbf{V}_l | \Delta \mathbf{V}_i). \quad (14)
\end{aligned}
$$

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

$$I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k) = I(\Delta \mathbf{V}_k; \Delta \mathbf{V}_l) + I(\Delta \mathbf{V}_i, \Delta \mathbf{V}_k | \Delta \mathbf{V}_l). \quad (15)$$

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 (13) and estimates the distribution grid topology [20, 34]. □

The mutual information $I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k)$ can be computed as

$$I(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k) = H(\Delta \mathbf{V}_i) + H(\Delta \mathbf{V}_k) - H(\Delta \mathbf{V}_i; \Delta \mathbf{V}_k), \quad (16)$$

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 (16) is that many distributions have closed forms of entropy. In Assumption 1, the incremental changes of voltages in distribution grids are assumed to follow Gaussian distribution approximately. Thus, the entropy of $\Delta \mathbf{V}_i$ is

$$H(\Delta \mathbf{V}_i) = \frac{r}{2} \log(2\pi \exp(1)) + \frac{1}{2} \log(\det \mathrm{Cov}(\Delta \mathbf{V}_i)), \quad (17)$$

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

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 false or unknown. Also, bus phase labels can change over time when new customers and DER devices are connected to grids [35]. As the correct bus phase connectivity information is critical to distribution grid plannings, the grid topology and phase connection should be estimated at the same time. To identify true bus phase labels, one may apply existing methods [35–37] to identify phase connectivity before estimating topology. Fortunately, our topology estimation method does not require this preprocessing step and is invariant to false phase labels. Specifically, when voltage phases are incorrectly labeled, the elements in random vector $\Delta \mathbf{V}_i$ are permuted. This permutation does not affect the computation of $\det \mathrm{Cov}(\Delta \mathbf{V}_i)$, 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. Section 5.3 uses numerical examples to show that our algorithm can recover the topology perfectly with the presence incorrect phase labels.

The proposed algorithm for multi-phase distribution grid topology estimation is summarized in Algorithm 1. The well-known Kruskal's minimum weight spanning tree algorithm [38, 39] can be applied to efficiently build the maximum weight spanning tree (Steps 6 - 16). The running time of the Kruskal's algorithm is $O(M \log M)$ for a radial distribution network with M buses.

3.2 Distribution Grid Topology Estimation using Voltage Magnitudes Only

Voltage phase angles are hard to acquire in distribution grids today because PMUs are not widely available. However, the proposed method can be extended to find the distribution grid topology only using voltage magnitudes $|\Delta \mathbf{V}|$. As presented in Lemma 3, the key step of the proposed method is computing the mutual information of each bus voltage pair. Using chain rule, the mutual

Algorithm 1 Multiphase Distribution Grid Topology Estimation

Require: $\Delta\mathbf{v}_i[n]$ for $i \in \mathcal{M}^+$, $n = 1, \cdots N$
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 (16) and (17).
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 $\widehat{\mathcal{E}}$ be the set of nodal pair comprising the maximum weight spanning tree. Set $\widehat{\mathcal{E}} = \emptyset$.
6: **for** $(i, k) \in \mathcal{T}$ **do**
7:     **if** cycle is detected in $\widehat{\mathcal{E}} \cup (i, k)$ **then**
8:         **Continue**
9:     **else**
10:         $\widehat{\mathcal{E}} \leftarrow \widehat{\mathcal{E}} \cup (i, k)$
11:     **end if**
12:     **if** $|\widehat{\mathcal{E}}| == M$ **then**
13:         **break**
14:     **end if**
15:     **return** $\widehat{\mathcal{E}}$
16: **end for**

information $I(\Delta\mathbf{V}_i; \Delta\mathbf{V}_k)$ can be decomposed as

$$
\begin{aligned}
I(\Delta\mathbf{V}_i; \Delta\mathbf{V}_k) &= I(|\Delta\mathbf{V}_i|, \Delta\boldsymbol{\theta}_i; |\Delta\mathbf{V}_k|, \Delta\boldsymbol{\theta}_k) && (18) \\
&= \underbrace{I(|\Delta\mathbf{V}_i|, |\Delta\mathbf{V}_k|)}_{\text{term A}} + \underbrace{I(|\Delta\mathbf{V}_i|, \Delta\boldsymbol{\theta}_i | |\Delta\mathbf{V}_k|)}_{\text{term B}} \\
&+ \underbrace{I(|\Delta\mathbf{V}_i|, \Delta\boldsymbol{\theta}_k | |\Delta\mathbf{V}_k|, \Delta\boldsymbol{\theta}_i)}_{\text{term C}}. && (19)
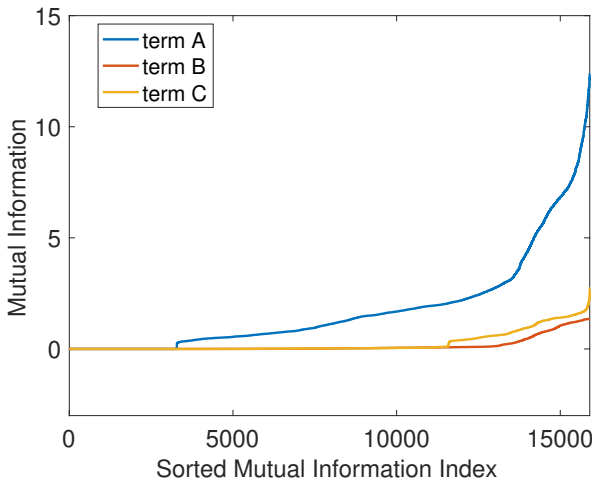\end{aligned}
$$

Fig. 7: Pairwise mutual information breakdown.

Fig. 7 empirically plots the pairwise mutual information of term A, B, C using the IEEE 123-bus system and the real data from PG&E. The mutual information computed in (19) is sorted by its value. The x-axis of Fig. 7 is the index of the sorted mutual information. The y-axis is the mutual information of each part in (19). The values of term A is much larger than term B and term C across all pairs of bus. The reason is that the changes of voltage angles are relatively small in distribution grids and thus, contain less information than voltage magnitudes. Based on our empirical observation in Fig. 7, $I(|\Delta\mathbf{V}_i|; |\Delta\mathbf{V}_k|)$ can be used to approximately estimate distribution grid structures. Specifically, the optimization problem in Lemma 3 is approximated as

$$
\widehat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta} \subset \mathcal{M}^+}{\arg\max} \sum_{i=1}^{M} I\left(|\Delta\mathbf{V}_i|; |\Delta\mathbf{V}_{\boldsymbol{\Theta}_i}|\right). \tag{20}
$$

Many smart meters deployed can measure the voltages of all three phases. Therefore, the proposed algorithm can still apply when only smart meter voltage magnitude measurements are available. Algorithm 2 summarizes the process for estimating topology using voltage magnitudes only.

Algorithm 2 Multiphase Distribution Grid Topology Estimation using Voltage Magnitudes

Require: $|\Delta\mathbf{v}_i[n]|$ for $i \in \mathcal{M}^+$, $n = 1, \cdots N$
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 (16) and (17).
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: Repeat Step 5 to Step 16 in Algorithm 1.

3.3 Topology Estimation of Weakly Mesh Distribution Grid

In the previous part, the multi-phase distribution grid topology estimation method is proposed for the radial system. In practice, with the increase penetration of DERs, more distribution grids become to mesh structure for robustness [12, 40, 41]. In mesh structures, a bus has more than one parents. Assuming only one bus has two parents, the joint distribution is rewritten as

$$
\begin{aligned}
P(\Delta\mathbf{V}_{\mathcal{M}^+}) &= P(\Delta\mathbf{V}_M | \Delta\mathbf{V}_{\text{pa}(M),1}, \Delta\mathbf{V}_{\text{pa}(M),2}) \\
&\times \prod_{i=1}^{M-1} P(\Delta\mathbf{V}_i | \Delta\mathbf{V}_{\text{pa}(i)}), && (21)
\end{aligned}
$$

where $\text{pa}(i), 1$ and $\text{pa}(i), 2$ represent the first and second parent of bus i. Following the same proof as Lemma 3, the KL distance $D(P(\Delta\mathbf{V}_{\mathcal{M}^+}) \| Q(\Delta\mathbf{V}_{\mathcal{M}^+}))$ can be written as

$$
\begin{aligned}
&D(P(\Delta\mathbf{V}_{\mathcal{M}^+}) \| Q(\Delta\mathbf{V}_{\mathcal{M}^+})) \\
&= -I(\Delta\mathbf{V}_M; \Delta\mathbf{V}_{\text{pa}(M),1}, \Delta\mathbf{V}_{\text{pa}(M),2}) \\
&- \sum_{i=1}^{M-1} I(\Delta\mathbf{V}_i; \Delta\mathbf{V}_{\text{pa}(i)}) + \text{constant.} && (22)
\end{aligned}
$$

Therefore, to find the topology for weakly meshed system (e.g., maximum number of parents is less than two), the optimization problem in Lemma 3 is approximated as

$$
\widehat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta} \subset \mathcal{M}^+}{\arg\max} \, I(\Delta\mathbf{V}_M; \Delta\mathbf{V}_{\boldsymbol{\Theta}_{M,1}}, \Delta\mathbf{V}_{\boldsymbol{\Theta}_{M,2}}) + \sum_{i=1}^{M-1} I(\Delta\mathbf{V}_i; \Delta\mathbf{V}_{\boldsymbol{\Theta}_i}).
\tag{23}
$$

Specifically, for each mutual information $I(\Delta\mathbf{V}_M; \Delta\mathbf{V}_{\boldsymbol{\Theta}_{M,1}}, \Delta\mathbf{V}_{\boldsymbol{\Theta}_{M,2}})$, $I(\Delta\mathbf{V}_M; \Delta\mathbf{V}_{\boldsymbol{\Theta}_{M,1}}, \Delta\mathbf{V}_{\boldsymbol{\Theta}_{M,2}}) + \sum_{i=1}^{M-1} I(\Delta\mathbf{V}_i; \Delta\mathbf{V}_{\boldsymbol{\Theta}_i})$ is computed by performing the maximum weighted spanning tree algorithm. Then, the one with the largest total mutual information is chosen to estimate system topology. The computational complexity is $\mathcal{O}(M(M-1)\log(M-1))$.

The method discussed above can be generalized to systems with more buses that contain more parents. However, the computational complexity also increases significantly. Therefore, for distributed systems that contain multiple loops, we recommend to adopt topology estimation methods that are designed for heavily mesh grids, such as [12, 33].

3.4 Bus Phase Identification and Correction

The previous section demonstrates that even with false phase labels, our method can correctly identify the multi-phase distribution

grid topology. In many field applications, accurate grid topology is not sufficient. The correct information of bus phases is also critical in grid plannings and operations. This subsection proposes a data-driven method to identify bus phase information and correct the false phase labels.
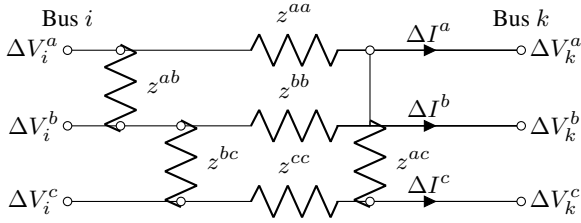

Fig. 8: An example of the two-port three-phase circuit.

Lemma 4. *In a multi-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 [21], the self impedance z^{aa} and mutual impedances z^{ab} and z^{ac} of a multi-phase power line can be computed as follows:

$$z^{aa} = r_{ik} + 0.095 + j0.121 \times H_{ik}^a \, \Omega/\text{miles}, \quad (24)$$

$$z^{ab} = 0.095 + j0.121 \times H_{ik}^{ab} \, \Omega/\text{miles}, \quad (25)$$

$$z^{ac} = 0.095 + j0.121 \times H_{ik}^{ac} \, \Omega/\text{miles}, \quad (26)$$

where H_{ik}^a, H_{ik}^{ab}, and H_{ik}^{ac} are constants, and r_{ik}^a is resistance of branch $i - k$ in Ω/miles. In a distribution grid, the resistance is usually larger than reactance [42]. Therefore, $z^{aa} \simeq r_{ik}^a + 0.095$ and $z^{ab} \simeq z^{ac} \simeq 0.095$.

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

$$\begin{bmatrix} \Delta V_i^a \\ \Delta V_i^b \\ \Delta V_i^c \end{bmatrix} = \begin{bmatrix} \Delta V_k^a \\ \Delta V_k^b \\ \Delta V_k^c \end{bmatrix} + \begin{bmatrix} Z^{aa} & Z^{ab} & Z^{ac} \\ Z^{ab} & Z^{bb} & Z^{bc} \\ Z^{ac} & Z^{bc} & Z^{cc} \end{bmatrix} \begin{bmatrix} \Delta I^a \\ \Delta I^b \\ \Delta I^c \end{bmatrix}, \quad (27)$$

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

$$\Delta V_i^a = \Delta V_k^a + C + r_{ik} \times l \times \Delta I^a, \quad (28)$$

$$\Delta V_i^b = \Delta V_k^b + C + r_{ik} \times l \times \Delta I^b, \quad (29)$$

$$\Delta V_i^c = \Delta V_k^c + C + r_{ik} \times l \times \Delta I^c, \quad (30)$$

where $C = 0.095 \times l \times (\Delta I^a + \Delta I^b + \Delta I^c)$. The phase voltages at the two ends of a branch are in a linear relationship. Therefore, their correlation is the largest. □

Table. 1 shows the voltage magnitude corrections among bus 64, 65, and 66 in IEEE 123-bus system. There are no PMUs in the system. Since PMUs can provide accurate phase measurements, the bus phase label identification problem is trivial with the presence of PMUs. In the 123-bus system, bus 64 and 65 are connected on phase b. Bus 65 and 66 are connected on phase c. In Table. 1, the correlation between bus 64 and 65 on phase b is much larger than other pairs. Similar observation holds for bus 65 and 66. Thus, to identify bus phases in a distribution grid, the correlation check can be applied from the substation of the radial network down to all leaf buses. The reason is that the substation bus label information is usually reliable. Then, the bus phase can be correctly identified, following the paths of estimated grid topology. Note that, the metering device installed at each bus can provide the number of phases at each bus. Therefore, this method is eligible for all types of bus. The same approach can also be applied to diagnose the correctness of the bus phase labels.

Table 1 Voltage Magnitude Correlations Between Bus 64, 65, and 66.

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

4 Unbalanced Multi-Phase Distribution Grid Topology Estimation with Incorrect Phase Labels

The results in the previous section illustrate the topology estimation for balanced multi-phase systems with incorrect phase labels. However, it is not directly expendable to unbalanced multi-phase systems. As shown in Fig. 8, 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, the grid is transformed using sequence component frameworks. The voltage phasor is decomposed into three balanced phasors known as positive sequence, negative sequence, and zero sequence. The multi-phase voltage $\Delta \mathbf{V}_i$ in phase frame is decomposed as follows:

$$\Delta \mathbf{V}_i = \begin{bmatrix} 1 & 1 & 1 \\ h^2 & h & 1 \\ h & h^2 & 1 \end{bmatrix} \begin{bmatrix} \Delta V_a^p \\ \Delta V_a^n \\ \Delta V_a^z \end{bmatrix} = \mathbf{H} \Delta \mathbf{V}_a^{pnz}, \quad (31)$$

where $h = \exp(j2\pi/3)$, $h^2 = \exp(-j2\pi/3)$, $\Delta V_a^p, \Delta V_a^n, \Delta V_a^z$ denote 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, the sequence components voltages of phase b and c are not required to compute. In the following text, $\Delta \mathbf{V}_i^{pnz}$ denotes the sequence voltage vector of bus i on phase a. The sequence voltages can be computed as follows:

$$\Delta \mathbf{V}^{pnz} = \mathbf{H}^{-1} \Delta \mathbf{V} = \frac{1}{3} \mathbf{H}^H \Delta \mathbf{V}, \quad (32)$$

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

$$\Delta \mathbf{I}_i^{pnz} = \mathbf{H}^{-1} \Delta \mathbf{I}_i, \quad (33)$$

$$\mathbf{Y}_{ik}^{pnz} = \mathbf{H}^{-1} \mathbf{Y}_{ik} \mathbf{H}. \quad (34)$$

A highlight is that the transformation above is applied to the multi-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., $\mathbf{Y}_{ik} = \mathbf{0}$, $\mathbf{Y}_{ik}^{pnz} = \mathbf{0}$. Therefore, finding topology in phase frame is equivalent to finding topology in sequence component frame. The proof of $P(\Delta \mathbf{V}_{\mathcal{M}^+}^{pnz}) = \prod_{i=1}^M P(\Delta \mathbf{V}_i^{pnz} | \Delta \mathbf{V}_{\text{pa}(i)}^{pnz})$ is required to apply the mutual information-based maximum weight spanning tree algorithm (Chow-Liu algorithm).

The transformation process in (32) does not require the correct phase labels in the phase frame. The reason is that when the bus phase labels are incorrect, the decomposition in (32) will become either the sequence component frame $\Delta \mathbf{V}_b^{pnz}$ or $\Delta \mathbf{V}_c^{pnz}$. Since both $\Delta \mathbf{V}_b^{pnz}$ or $\Delta \mathbf{V}_c^{pnz}$ are both balanced systems, the same method proposed for $\Delta \mathbf{V}_a^{pnz}$ can be applied to estimate system topology.

Lemma 5 (Data Processing Inequality [32]). *If random vectors $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ forms a Markov Chain, i.e., $\mathbf{X} \to \mathbf{Y} \to \mathbf{Z}$, $I(\mathbf{X}; \mathbf{Y}) \geq I(\mathbf{X}; \mathbf{Z})$. 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: The first step of the proof is showing that the current injection increment are independent in sequence component frame, given the current injection increment at each bus is independent in phase frame. There are multiple ways to prove it. Here, an information theoretical approach is adopted.

When two random vectors are independent, their mutual information is zero [32], 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 $\Delta\mathbf{I}_k$, these random vectors form a Markov Chain, i.e., $\Delta\mathbf{I}_i \to \Delta\mathbf{I}_k \to \Delta\mathbf{I}_k^{pnz}$. Applying Lemma 5,

$$I(\Delta\mathbf{I}_i; \Delta\mathbf{I}_k^{pnz}) \leq I(\Delta\mathbf{I}_i; \Delta\mathbf{I}_k) = 0. \tag{35}$$

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, another Markov Chain is formed: $\Delta\mathbf{I}_i^{pnz} \to \Delta\mathbf{I}_i \to \Delta\mathbf{I}_k^{pnz}$. Applying Lemma 5 again,

$$I(\Delta\mathbf{I}_i^{pnz}; \Delta\mathbf{I}_k^{pnz}) \leq I(\Delta\mathbf{I}_i; \Delta\mathbf{I}_k^{pnz}) = 0. \tag{36}$$

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

The second step of the proof is showing that the conditional independence of nodal voltages holds in sequence component frame. The example in Fig. 5 is adopted to illustrate it. In the sequence component frame, the nodal equation of the system in Fig. 5 is $\mathbf{Y}_{\mathcal{M}^+}^{pnz}\Delta\mathbf{V}_{\mathcal{M}^+}^{pnz} = \Delta\mathbf{I}_{\mathcal{M}^+}^{pnz}$, where $\mathbf{Y}_{\mathcal{M}^+}^{pnz}$ is

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

$\mathbf{Y}_{ik}^{pnz} = \mathbf{Y}_{ki}^{pnz}$, and $\mathbf{Y}_{ii}^{pnz} = -\sum_{k=1, k \neq i}^{7} \mathbf{Y}_{ik}^{pnz}$. If $\mathbf{Y}_{ik}^{pnz} = \mathbf{0}$, there is no branch between bus i and k. This equation is in the same format as (4). Since $\Delta\mathbf{I}_i^{pnz} \perp \Delta\mathbf{I}_k^{pnz}$ for all $i \neq k$, the same method used in the proof of Lemma 1 and Lemma 2 can show the conditional independence of nodal voltages in sequence component frame. □

With Lemma 6, the conditional independence of current injection is proved to hold 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 multi-phase system, the Chow-Liu algorithm can 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*

$$\widehat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta} \subset \mathcal{M}^+}{\arg\max} \sum_{i=1}^{M} I(\Delta\mathbf{V}_i^{pnz}; \Delta\mathbf{V}_{\boldsymbol{\Theta}_i}^{pnz}). \tag{38}$$

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 multi-phase distribution grids is summarized in Algorithm 3.

Algorithm 3 Unbalanced Multiphase Distribution Grid Topology Estimation via Sequence Component Frame

---

Require: $\Delta\mathbf{v}_i[n]$ for $i \in \mathcal{M}^+$, $n = 1, \cdots N$
1: Compute voltage phasor $\Delta\mathbf{v}_i^{pnz}[n]$ using (32) for $i \in \mathcal{M}^+$, $n = 1, \cdots N$.
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 phase angles of $\Delta\mathbf{V}$ are needed for performing the phase frame transformation in (32). However, as discussed in Section 3.2, PMUs have not been widely available in distribution grids. To only use voltage magnitudes to address the unbalance problem, the following approximation is proposed:

$$\Delta\mathbf{V}^{pnz} = \mathbf{H}^{-1}|\Delta\mathbf{V}|. \tag{39}$$

In this approximation, only the voltage magnitudes in phase frame are used to compute voltages in sequence component frame. As demonstrated in Section 5, this approximation does not introduce significant errors to topology estimation. In addition, the grid topology is identical in phase frame and in sequence component frame. Therefore, once the unbalanced grid topology is estimated, Lemma 4 is applied to identify the phases of all buses.

5 Simulations and Numerical Results

In this section, the proposed algorithms for balanced and unbalanced grid topology estimation are validated on on IEEE 37-bus, 123-bus (Fig. 10), and 8500-bus distribution networks [22, 43] using data from USA (California and Taxes) and Europe. Also, we validate the proposed algorithms on systems with different levels of DER penetration and the presence of incorrect phase labels. Furthermore, sensitivity analysis is conducted on data lengths, data accuracy, load patterns, and data resolutions.

37-bus, 123-bus, and 8500-bus systems are multi-phase. 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 [25], an open source simulator for distribution grid. The load profile from PG&E is used to simulate the power system behavior in a practical pattern. 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 multi-phase. The details of bus and branch phases are given in [22]. The voltage data at each bus are used for topology estimation. Fig. 9 summarizes the overall process of topology estimation and bus phase identification.
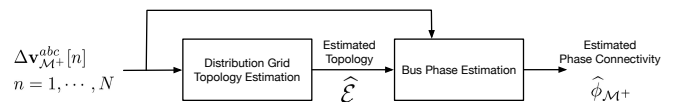

Fig. 9: Flow chart of topology estimation and bus phase identification process.

The PG&E data set does not contain the reactive power. The reactive power $q_i^\phi[n]$ on phase ϕ 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., $\mathbf{v}_i[n]$, the power flow analysis is run to generate the
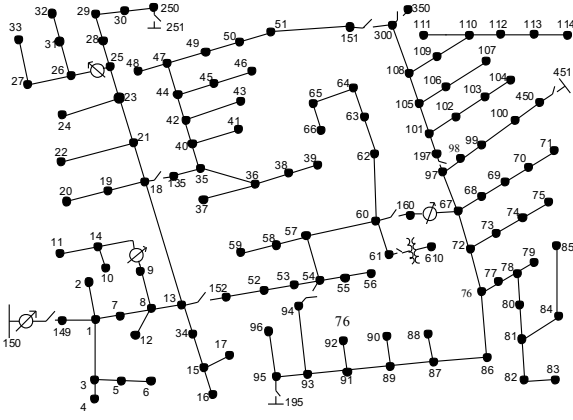
Fig. 10: IEEE 123-bus distribution test case.

hourly states of the power system over a year. $N = 8760$ measurements are obtained at each bus. Section 5.4.1 investigates the data length requirement for topology estimation. The loads attached to each phase are unequal. Hence, the systems are unbalanced. Fig. 11 and Fig. 12 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.
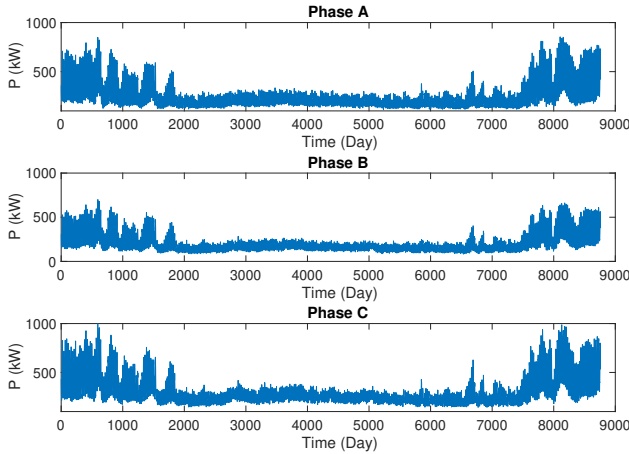

Fig. 12: Hourly aggregated real powers on each phase in 123-bus system.

When phase angle data are available, our algorithm perfectly estimates the grid topology. When only voltage magnitudes are available, our algorithm can still estimate the grid topology perfectly. This result also verifies that our approximation in (39) is sufficient for topology estimation.

Table 2 Topology Estimation Error Rate without DERs.

System	Proposed Method		Modified Single-Phase 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%

The proposed algorithm is also compared with a modified single-phase topology estimation in [20]. Specifically, the single-phase topology estimator is applied to each phase individually. Then, the single-phase topology estimates are combined to produce the multi-phase system topology. As shown in Table 2, the modified single-phase method has worse performance than the proposed algorithm. The key reason is that the modified single-phase method does not consider the voltage coupling across phases.

The proposed algorithm is compared with the method in [44], which is also based on minimizing the KL distance and searches the correct operational topology from all possible topology candidates. For 37-bus system, the error rate of [44] is 5.6%. For 123-bus system, the error rate is 8.2%. The high error rates are due to the DC approximation in [44]. For unbalanced distribution grids, the DC approximation does not hold generally.

In addition, we validate our algorithm on IEEE 8500-node distribution system [43], which contains both low-voltage and medium-voltage buses. The error rate that using $\Delta\mathbf{V}^{pnz}$ is 15.7%. Most incorrect identified branches are near the low-voltage grid feeders. In many systems, the locations and connectives of the low-voltage grid feeders are accurate. Therefore, with the prior knowledge of low-voltage grid feeders, the error rate is reduced to 3.8%.

Fig. 11: Hourly aggregated real powers on each phase in 37-bus system.

5.1 Distribution Grid Topology Estimation Error Rate

This section discusses the performance on grid topology estimation. The error rate (ER) is employed as the performance evaluation metric, which is defined as

$$\text{ER} = \frac{1}{|\mathcal{E}|} \Bigg(\underbrace{\sum_{(i,k) \in \widehat{\mathcal{E}}} \mathbb{I}\left((i,k) \notin \mathcal{E}\right)}_{\text{false estimation}} + \underbrace{\sum_{(i,k) \in \mathcal{E}} \mathbb{I}\left((i,k) \notin \widehat{\mathcal{E}}\right)}_{\text{missing}} \Bigg)\% \tag{40}$$

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

Table 2 summarizes the topology estimation error rates of unbalanced multi-phase 37- and 123-bus systems using noiseless data.
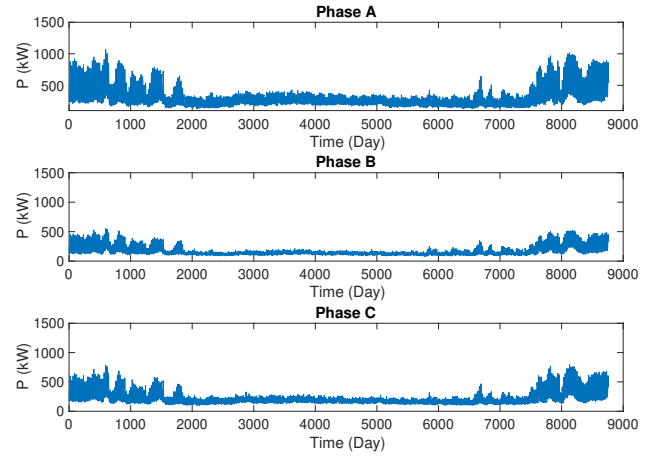
Table 3 Topology Estimation Computation Time (seconds)

System	Computation time of mutual information	Computation time of maximum weight spanning tree	Total time
37-bus	0.476	0.499	0.975
123-bus	2.978	3.114	6.092
8500-bus	244.345	478.111	722.456
8500-bus (parallel)	89.090	94.028	183.118

Table 3 summarizes the average computational time of the proposed algorithm on different systems over 1000 Monte Carlo simulation iterations. For 37-bus and 123-bus systems, our algorithm takes a few seconds to report the estimated topology, which makes it suitable for real-time monitoring. For 8500-bus system, the computational time of both mutual information and maximum weight spanning tree grows up. Though the computational time for a large system is high, some power system properties can help to speed up the topology estimation process. As mentioned above, the locations and connectives of low-voltage grid feeders are accurate. Hence, for large-scale system that has both low-voltage and medium-voltage systems, the topology estimation problem can be performed in two steps: 1) only identify the topology of low-voltage grids and 2) only estimate the topology of medium-voltage grids. Since each low-voltage grid operates independently, the topology estimation process can run in parallel. As indicated in Table 3, by decomposing a large-scale grid into multiple small sub-grids, the computational time of topology estimation is reduced by 75%. A highlight is that in the parallel computation, the maximum computational time is bounded by the largest low-voltage grid. If every low-voltage grid is small (e.g., similar size as the 123-bus system), the computational time can be much less. Another highlight is that the computational time is invariant to the integration of DERs and data lengths.

5.2 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, 20% of residents in the distribution networks are selected to install rooftop photovoltaic (PV) systems. The profiles of hourly power generation are 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 [45]. The power factor is fixed as 0.90 lagging, which satisfies the regulation of many U.S. utilities [46] and IEEE standard [47]. Similar to the simulations without DERs, we use one year's data (8760 samples) to estimate topology.

Table 4 Topology Estimation Error Rate with 20% PV Penetrations.

System	Proposed Method		Modified Single-Phase 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 4 using noiseless measurements. Our algorithm does not have any performance degradation with DER integration. Also, the modified single-phase method still performs worse than the proposed method. Compared with the systems without DER, the modified single-phase method has performance degradation.

In order to further validate the proposed algorithm, the DER penetration level is progressively increased from 0% to 100%. For each penetration level, Monte Carlo simulation is performed over 1000 iterations. Fig. 13 plots the error rate with different levels of DER penetration using the voltage magnitude $|\Delta\mathbf{V}^{pnz}|$ only. 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. 12 iterations of Monte Carlo simulation have errors when the DER penetration level is 60%. The incorrect identified branches are bus 57 - bus 58 and bus 58 - bus 59. The loads with PV integrations on these three buses are similar and the line impedances are identical. This causes that the voltage profiles of these buses are similar. Our algorithm is hard to identify the correct connectivity. However, such an instance requires the same

impedance and same voltage profiles. This rarely happens in practice. As the penetration level increase, this instance is not observed again and the proposed algorithm can correctly identify these two branches.
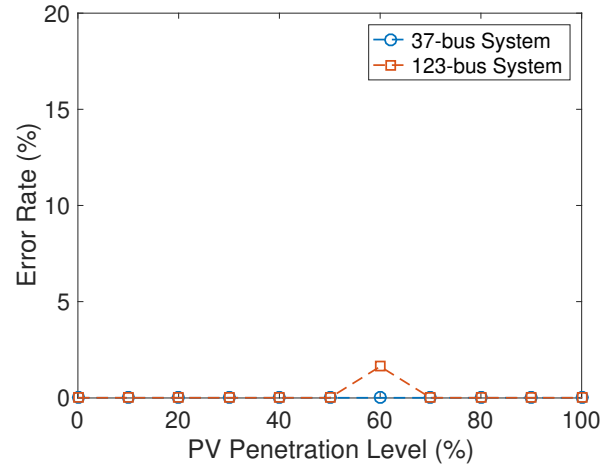

Fig. 13: Error rates with different levels of DER penetration using $|\Delta\mathbf{V}^{pnz}|$.

5.3 Distribution Grids with Incorrect Phase Labels

In some distribution grids, up to 10% of the phase labels are incorrect or unknown. Therefore, this section validates our algorithm on the 123-bus system with incorrect phase labels. To simulate the incorrect phase labels, several buses are randomly chosen and switch their phase a voltage measurements to data of either phase b voltage or phase c voltage.

Table 5 shows the error rates with different percentages of incorrect phase labels using voltage magnitude only and highlights that our algorithm is insensitive to incorrect bus phase label. As 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 the modified single-phase approach is used, 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 5 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%

5.4 Sensitivity Analysis

5.4.1 Sensitivity to Data Lengths: The proposed algorithm is validated with different data lengths, ranging from 1 to 360 days. Fig. 14 illustrates the error rates of the 123-bus system, with and without DER, over different lengths of the PG&E data set. With 20 days' measurements ($24 \times 20 = 480$ data points), the proposed method can achieve zero error. This result is better than the single-phase system presented in [20], 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. The frequency of distribution grid

reconfiguration ranges from hours to weeks [48, 49]. Section 5.4.4 demonstrates that the topology can still be estimated by increasing the sampling frequency of smart meters.
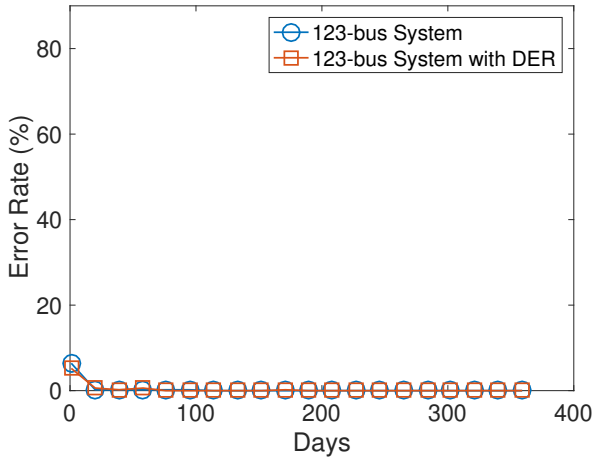

Fig. 14: Error rates with different data lengths.

5.4.2 Sensitivity to Data Accuracy: In particles, 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 the smart meters to have an error less than $\pm 0.5\%$ [50, 51]. Table. 6 shows the error rates with different noise levels over 20 iterations in the 123-bus system with PG&E data. 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 both feeders and incorrectly connected with the presence of noise. In real systems, the location of feeder buses are usually known. Therefore, a post-processing can be applied on the topology estimate and remove these unnecessary branches from topology estimate. The updated system is still a radial network. After performing post-processing, the error rate decreases to 1%. In Table. 6, the standard deviation of error rate is very small and therefore, our algorithm can provide reliable and consistent results with noisy measurements.

Table 6 Error Rates with Different Voltage Noise Levels in 123-bus System

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%

5.4.3 Sensitivity to Data Patterns: The "ADRES-Concept" Project load profile [23, 24] is employed to understand our algorithm's sensitivity to load patterns. 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. The voltage data are generated using the 37-bus system. The load profiles are scaled to match the scale of power in the 37-bus system. The resulting multi-phase system is unbalanced.

Fig. 15 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 a consistent performance in

winter and summer. Compared with the results in [20], the proposed algorithm perfectly estimates the grid topology with shorten time because more information is observed at each time step.
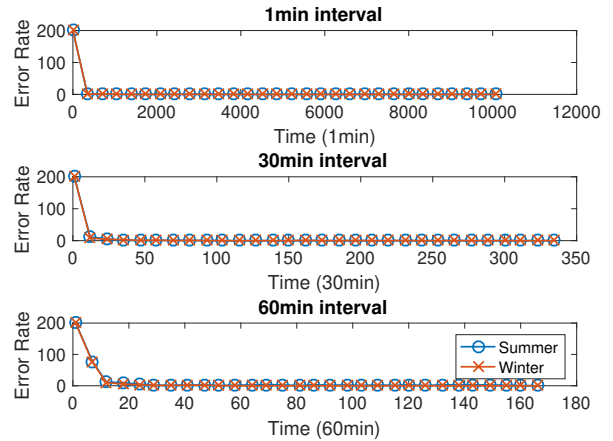

Fig. 15: Error rates with summer and winter load patterns.

Another validation of our algorithm is using data set from Pecan Street, which contains hourly load measurements of 345 houses with PV integrations in Austin, Taxes. The measurements include both power consumption and renewable generation. In Fig. 16, 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.
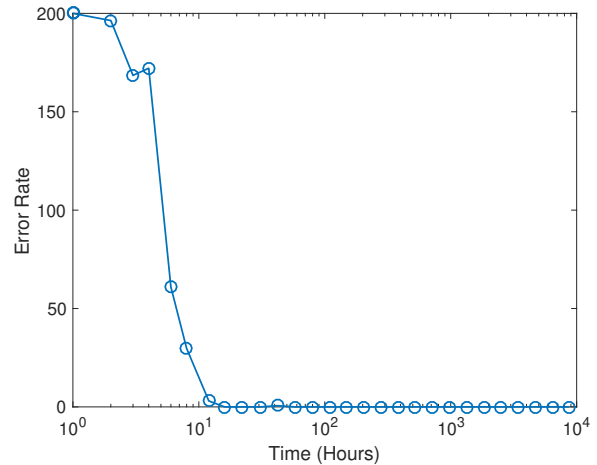

Fig. 16: Error rates on Pecan Street data set.

In order to better understanding the impacts of ZIP loads and high applicants on the topology estimation, the applicant/device simulation model [52] provided by Gridlab-D is adopted to generate load data. This simulator is based on the thermal data and device configurations. Therefore, compared with the real data provided from PG&E, a detailed setup of load patterns is possible. For each residential load, multiple devices and applicants (e.g., heating, electric hot water heaters, washer and dryers, cooking, electronic plugs and lights) are installed with various configurations and parameters. We run the simulation on IEEE 123-bus network with real temperature data from Palo Alto, CA. The simulation is performed on an hourly basis for one year's duration. By applying the proposed algorithm, the topology can still be correctly estimated and the required data lengths are consistent with the results in Section 5.4.1.

5.4.4 Sensitivity to Data Resolutions:

Fig. 17 illustrates the performance of the proposed algorithm under different sampling frequencies using the ADRES data set. When the sampling period is 1 minute, about 6 hours' voltage profile are required to perfectly recover the system. According to [49], some distribution grids reconfigure as fast as every 3 hours. Therefore, the proposed algorithm is suitable for existing systems and real-time operations. If the sampling period is 30 minutes, 35 data points (35×30 minutes = 17.5hours) to recover the system topology. This estimation time is only half of the required time in [20].
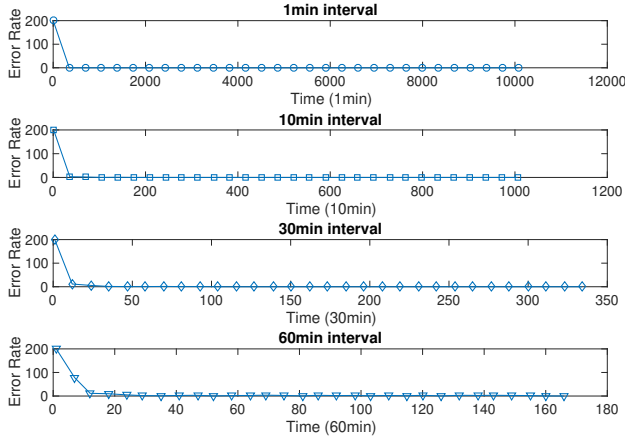

Fig. 17: Error rates with different data resolutions.

6 Conclusions

This paper proposes a data-driven approach to estimate multi-phase distribution grid topology by utilizing smart meter measurements. Unlike existing approaches, our method does not require the system to be balanced. Also, our method tolerates the errors of bus phase labels. Specifically, the topology estimation problem is formulated as a joint distribution (voltage phasors) approximation problem under the probabilistic graphical model framework. Then, the distribution grid topology estimation is proven to be equivalent to the graphical model estimation problem and propose a mutual information-based maximum weight spanning tree algorithm, which is optimal and efficient. Moreover, our algorithm is extended to the case where only voltage magnitude is available. In addition, as bus phase labels are critical to distribution grid plannings and operations, a simple approach is introduced to correct the error of bus phase labels by utilizing Carson's equations. Finally, the proposed algorithm is validated on IEEE 37- and 123-bus systems and compared with the existing single-phase method. Results show that the proposed algorithm outperforms the single-phase method and has robust performances when bus phase labels are incorrect. Our algorithms are also validated under different penetration levels of DERs and conduct the sensitivity analysis. The numerical results are highly accurate and robust in various system configurations.

7 Appendices

7.1 Proof of Lemma 1

Proof: Several cases illustrated in Fig. 18 are used to prove Lemma 1. The first step is proving the leaf nodes. In Fig. 18(a), 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 $\Delta \mathbf{V}_1 = \Delta \mathbf{v}_1$:

$$\Delta \mathbf{I}_4 = \mathbf{Y}_{42}\Delta \mathbf{v}_2 + \mathbf{Y}_{44}\Delta \mathbf{V}_4, \tag{41}$$

$$\Delta \mathbf{I}_k = \mathbf{Y}_{1k}\Delta \mathbf{v}_1 + \mathbf{Y}_{kk}\Delta \mathbf{V}_k \quad \forall k \in \mathcal{S}(2). \tag{42}$$
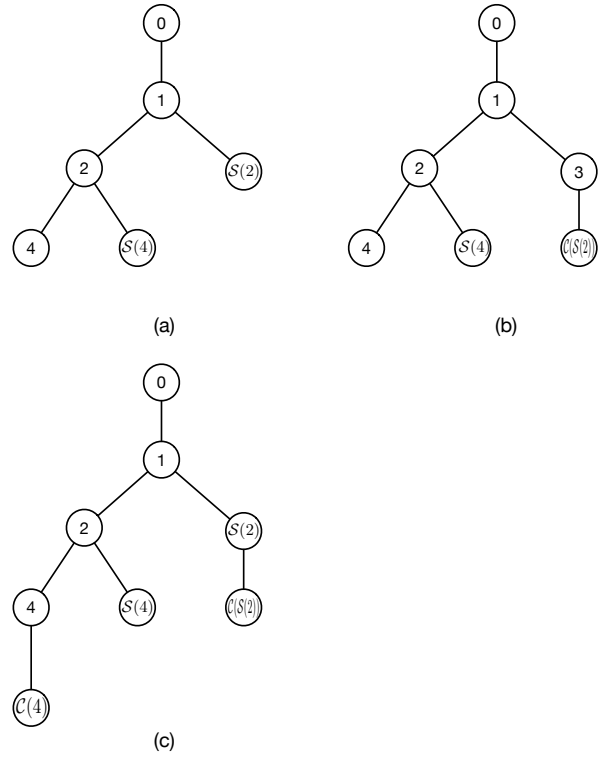

Fig. 18: Figure for the proof of Lemma 1.

Since $\Delta \mathbf{I}_4 \perp \Delta \mathbf{I}_k$, $\Delta \mathbf{V}_4$ 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. 18(b), 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 $\Delta \mathbf{V}_1 = \Delta \mathbf{v}_1$:

$$\Delta \mathbf{I}_4 = \mathbf{Y}_{42}\Delta \mathbf{v}_2 + \mathbf{Y}_{44}\Delta \mathbf{V}_4, \tag{43}$$

$$\Delta \mathbf{I}_3 = \mathbf{Y}_{13}\Delta \mathbf{v}_1 + \sum_{k \in \mathcal{C}(3)} \mathbf{Y}_{3k}\Delta \mathbf{V}_k + \mathbf{Y}_{33}\Delta \mathbf{V}_3, \tag{44}$$

$$\Delta \mathbf{I}_k = \mathbf{Y}_{3k}\Delta \mathbf{V}_3 + \mathbf{Y}_{kk}\Delta \mathbf{V}_k \quad \forall k \in \mathcal{C}(3). \tag{45}$$

Since $\mathbf{Y}_{ii} = -\mathbf{Y}_{\mathrm{pa}(i)i} - \sum_{k \in \mathcal{C}(i)} \mathbf{Y}_{ki}$ and $\mathbf{Y}_{ik} = \mathbf{Y}_{ki}$, combining (44) and (45), the equation becomes

$$
\begin{aligned}
&\Delta \mathbf{I}_3 + \sum_{k \in \mathcal{C}(3)} \Delta \mathbf{I}_k \\
&= \mathbf{Y}_{13}\Delta \mathbf{v}_1 + \mathbf{Y}_{33}\Delta \mathbf{V}_3 + \sum_{k \in \mathcal{C}(3)} \mathbf{Y}_{3k}\Delta \mathbf{V}_k \\
&\quad + \sum_{k \in \mathcal{C}(3)} (\mathbf{Y}_{3k}\Delta \mathbf{V}_3 + \mathbf{Y}_{kk}\Delta \mathbf{V}_k) \\
&= \mathbf{Y}_{13}\Delta \mathbf{v}_1 - (\mathbf{Y}_{13}\Delta \mathbf{V}_3 + \sum_{k \in \mathcal{C}(3)} \mathbf{Y}_{3k}\Delta \mathbf{V}_3) \\
&\quad + \sum_{k \in \mathcal{C}(3)} (\mathbf{Y}_{3k}\Delta \mathbf{V}_k + \mathbf{Y}_{3k}\Delta \mathbf{V}_3 - \mathbf{Y}_{3k}\Delta \mathbf{V}_k) \\
&= \mathbf{Y}_{13}\Delta \mathbf{v}_1 - \mathbf{Y}_{13}\Delta \mathbf{V}_3 \tag{46}
\end{aligned}
$$

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. (46) can be rewritten as an equation of $\Delta \mathbf{V}_3$, i.e.,

$$\Delta \mathbf{V}_3 = (\mathbf{Y}_{13})^{-1}(\mathbf{Y}_{13}\Delta \mathbf{v}_1 - \Delta \mathbf{I}_3 - \sum_{k \in \mathcal{C}(3)} \Delta \mathbf{I}_k). \tag{47}$$

Replacing $\Delta\mathbf{V}_3$ in (45) with the equations above, then, for $k \in \mathcal{C}(3)$,

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

(48)

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 $\Delta\mathbf{V}_{\mathcal{C}(3)}$, the same induction method above can be applied to prove the conditional independence. Thus, the proof of Fig. 18(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 part proves the lemma for non-leaf buses. In Fig. 18(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 $\Delta\mathbf{V}_1 = \Delta\mathbf{v}_1$,

$$\Delta\mathbf{I}_4 = \mathbf{Y}_{42}\Delta\mathbf{v}_2 + \mathbf{Y}_{44}\Delta\mathbf{V}_4 + \sum_{k\in\mathcal{C}(4)}\mathbf{Y}_{4k}\Delta\mathbf{V}_k, \quad (49)$$

$$\Delta\mathbf{I}_k = \mathbf{Y}_{4k}\Delta\mathbf{V}_4 + \mathbf{Y}_{kk}\Delta\mathbf{V}_k, \forall k \in \mathcal{C}(4), \quad (50)$$

$$\Delta\mathbf{I}_l = \mathbf{Y}_{1l}\Delta\mathbf{v}_1 + \sum_{m\in\mathcal{C}(l)}\mathbf{Y}_{lm}\Delta\mathbf{V}_m + \mathbf{Y}_{ll}\Delta\mathbf{V}_l, \quad (51)$$

$$\Delta\mathbf{I}_m = \mathbf{Y}_{lm}\Delta\mathbf{V}_l + \mathbf{Y}_{mm}\Delta\mathbf{V}_m, \quad (52)$$

where $l \in \mathcal{S}(2)$ and $m \in \mathcal{C}(l)$. Combining (49) and (50) yields

$$\Delta\mathbf{I}_4 + \sum_{k\in\mathcal{C}(4)}\Delta\mathbf{I}_k = \mathbf{Y}_{42}\Delta\mathbf{v}_2 - \mathbf{Y}_{42}\Delta\mathbf{V}_4. \quad (53)$$

For every l in $\mathcal{S}(2)$, combining (51) and (52) yields the following equation:

$$\Delta\mathbf{I}_l + \sum_{m\in\mathcal{C}(l)}\Delta\mathbf{I}_m = \mathbf{Y}_{1l}\Delta\mathbf{v}_1 - \mathbf{Y}_{1l}\Delta\mathbf{V}_l. \quad (54)$$

Applying the strategy in Fig. 18(b) to (53) and (54) could prove that $\Delta\mathbf{V}_4$ and $\Delta\mathbf{V}_l$ are conditionally independent. Also, $\Delta\mathbf{V}_4$ and $\Delta\mathbf{V}_m$ are proved to be conditionally independent for $m \in \mathcal{C}(l)$ by combining (52) and (54). The results in Fig. 18(c) can be generalized to all non-leaf buses. Using the results in Fig. 18, Lemma 1 is proved to hold. □

7.2 Proof of Theorem 3

Proof: Recall the definition [32], the KL divergence is expressed as

$$D(P(\Delta\mathbf{V}_{\mathcal{M}+})\|Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta}))$$
$$=E_{P(\Delta\mathbf{V}_{\mathcal{M}+})}\log\frac{P(\Delta\mathbf{V}_{\mathcal{M}+})}{Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta})}$$
$$=\int P(\Delta\mathbf{V}_{\mathcal{M}+})\log\frac{P(\Delta\mathbf{V}_{\mathcal{M}+})}{Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta})}$$
$$=\int P(\Delta\mathbf{V}_{\mathcal{M}+})(\log P(\Delta\mathbf{V}_{\mathcal{M}+}) - \log Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta}))$$
$$=\int P(\Delta\mathbf{V}_{\mathcal{M}+})\log P(\Delta\mathbf{V}_{\mathcal{M}+})$$
$$\quad - \int P(\Delta\mathbf{V}_{\mathcal{M}+})\log Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta}). \quad (55)$$

Because of Lemma 2, the radial structured PDF $Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta})$ can be expressed as a conditional distribution $\prod_{i=1}^{M}P(\Delta\mathbf{V}_i|\Delta\mathbf{V}_{\boldsymbol{\Theta}_i})$.

Then,

$$D(P(\Delta\mathbf{V}_{\mathcal{M}+})\|Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta}))$$
$$=\int P(\Delta\mathbf{V}_{\mathcal{M}+})\log P(\Delta\mathbf{V}_{\mathcal{M}+})$$
$$\quad - \int P(\Delta\mathbf{V}_{\mathcal{M}+})\log\prod_{i=1}^{M}P(\Delta\mathbf{V}_i|\Delta\mathbf{V}_{\boldsymbol{\Theta}_i})$$
$$=\int P(\Delta\mathbf{V}_{\mathcal{M}+})\log P(\Delta\mathbf{V}_{\mathcal{M}+})$$
$$\quad - \int P(\Delta\mathbf{V}_{\mathcal{M}+})\sum_{i=1}^{M}\log P(\Delta\mathbf{V}_i|\Delta\mathbf{V}_{\boldsymbol{\Theta}_i}), \quad (56)$$

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, onecan simplify the equation above as

$$D(P(\Delta\mathbf{V}_{\mathcal{M}+})\|Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta}))$$
$$=\int P(\Delta\mathbf{V}_{\mathcal{M}+})\log P(\Delta\mathbf{V}_{\mathcal{M}+})$$
$$\quad - \int P(\Delta\mathbf{V}_i|\Delta\mathbf{V}_{\boldsymbol{\Theta}_i})\sum_{i=1}^{M}\log\frac{P(\Delta\mathbf{V}_i,\Delta\mathbf{V}_{\boldsymbol{\Theta}_i})}{P(\Delta\mathbf{V}_i)P(\Delta\mathbf{V}_{\boldsymbol{\Theta}_i})}$$
$$\quad - \sum_{i=1}^{M}\int P(\Delta\mathbf{V}_i)\log P(\Delta\mathbf{V}_i)$$
$$=-H(\Delta\mathbf{V}_{\mathcal{M}+}) - \sum_{i=1}^{M}I(\Delta\mathbf{V}_i;\Delta\mathbf{V}_{\boldsymbol{\Theta}_i}) + \sum_{i=1}^{M}H(\Delta\mathbf{V}_i). \quad (57)$$

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

$$H(\Delta\mathbf{V}_i) = -\int P(\Delta\mathbf{V}_i)\log P(\Delta\mathbf{V}_i), \quad (58)$$

and mutual information, i.e.,

$$I(\Delta\mathbf{V}_i;\Delta\mathbf{V}_{\boldsymbol{\Theta}_i})$$
$$= \int P(\Delta\mathbf{V}_i,\Delta\mathbf{V}_{\boldsymbol{\Theta}_i})\log\frac{P(\Delta\mathbf{V}_i,\Delta\mathbf{V}_{\boldsymbol{\Theta}_i})}{P(\Delta\mathbf{V}_i)P(\Delta\mathbf{V}_{\boldsymbol{\Theta}_i})}. \quad (59)$$

Thus, to minimize the KL-divergence between $P(\Delta\mathbf{V}_{\mathcal{M}+})$ and $Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta})$, one can choose the $M-1$ edges to maximize $\sum_{i=1}^{M}I(P(\Delta\mathbf{V}_i);P(\Delta\mathbf{V}_{\boldsymbol{\Theta}_i}))$. 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 is excluded in the final optimization problem. Therefore, minimizing $D(P(\Delta\mathbf{V}_{\mathcal{M}+})\|Q(\Delta\mathbf{V}_{\mathcal{M}+};\boldsymbol{\Theta}))$ is equivalent to solving the following optimization problem:

$$\widehat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta}\subset\mathcal{M}+}{\operatorname{argmax}}\sum_{i=1}^{M}I(\Delta\mathbf{V}_i;\Delta\mathbf{V}_{\boldsymbol{\Theta}_i}). \quad (60)$$

□

References

1 Dey, S., Jessa, A., Gelbien, L. 'Urban Grid Monitoring Renewables Integration'. In: IEEE Conference on Innovative Technologies for an Efficient and Reliable Electricity Supply. (, 2010. pp. 252–256

2 Clement.Nyns, K., Haesen, E., Driesen, J.: 'The Impact of Charging Plug-in Hybrid Electric Vehicles on a Residential Distribution Grid', *Power Systems, IEEE Transactions on*, 2010, **25**, (1), pp. 371–380

3 Huang, J., Gupta, V., Huang, Y.F. 'Electric Grid State Estimators for Distribution Systems with Microgrids'. In: IEEE 46th Annul Conference on Information Sciences and Systems. (, 2012. pp. 1–6

4   Abur, A., Exposito, A.G.: 'Power System State Estimation: Theory and Implementation'. (CRC press, 2004)

5   Lugtu, R.L., Hackett, D.F., Liu, K.C., Might, D.D.: 'Power System State Estimation: Detection of Topological Errors', *Power Apparatus and Systems, IEEE Transactions on*, 1980, **PAS-99**, (6), pp. 2406–2412

6   Rudin, C., Waltz, D., Anderson, R.N., Boulanger, A., Salleb.Aouissi, A., Chow, M., et al.: 'Machine Learning for the New York City Power Grid', *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2012, **34**, (2), pp. 328–345

7   Deka, D., Chertkov, M., Backhaus, S.: 'Structure Learning in Power Distribution Networks', *IEEE Transactions on Control of Network Systems*, 2017, **PP**, (99), pp. 1–1

8   Cavraro, G., Kekatos, V., Veeramachaneni, S.: 'Voltage Analytics for Power Distribution Network Topology Verification', *IEEE Transactions on Smart Grid*, 2017,

9   Sharon, Y., Annaswamy, A.M., Motto, A.L., Chakraborty, A. 'Topology Identification in Distribution Network with Limited Measurements'. In: Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES. (IEEE, 2012. pp. 1–6

10   Bolognani, S., Bof, N., Michelotti, D., Muraro, R., Schenato, L. 'Identification of Power Distribution Network Topology via Voltage Correlation Analysis'. In: Conference on Decision and Control. (, 2013. pp. 1659–1664

11   Peppanen, J., Grijalva, S., Reno, M.J., Broderick, R.J. 'Distribution System Low-Voltage Circuit Topology Estimation using Smart Metering Data'. In: Transmission and Distribution Conference and Exposition. (IEEE, 2016. pp. 1–5

12   Liao, Y., Weng, Y., Liu, G., Rajagopal, R.: 'Urban MV and LV Distribution Grid Topology Estimation via Group Lasso', *IEEE Transactions on Power Systems*, 2018, pp. 1–1

13   Lo, K.L., Zhang, C. 'Decomposed Three-Phase Power Flow Solution using the Sequence Component Frame'. In: IEE Proceedings C (Generation, Transmission and Distribution). vol. 140. (IET, 1993. pp. 181–188

14   Tleis, N.: 'Power Systems Modelling and Fault Analysis: Theory and Practice'. (Newnes, 2007)

15   American National Standards Institute. 'ANSI C84.1: Electric power systems and equipment voltage ratings (60 Herz)'. (American National Standards Institute, 1995.

16   Routtenberg, T., Xie, Y., Willett, R.M., Tong, L.: 'PMU-based Detection of Imbalance in Three-Phase Power Systems', *IEEE Transactions on Power Systems*, 2015, **30**, (4), pp. 1966–1976

17   Yuan, Y., Ardakanian, O., Low, S., Tomlin, C.: 'On the inverse power flow problem', *arXiv preprint arXiv:161006631*, 2016,

18   Ardakanian, O., Yuan, Y., Dobbe, R., von Meier, A., Low, S., Tomlin, C. 'Event Detection and Localization in Distribution Grids with Phasor Measurement Units'. In: 2017 IEEE Power & Energy Society General Meeting. (IEEE, 2017. pp. 1–5

19   Deka, D., Chertkov, M., Backhaus, S.: 'Topology Estimation using Graphical Models in Multi-Phase Power Distribution Grids', *arXiv preprint arXiv:180306531*, 2018,

20   Weng, Y., Liao, Y., Rajagopal, R.: 'Distributed Energy Resources Topology Identification via Graphical Modeling', *IEEE Transactions on Power Systems*, 2017, **32**, (4), pp. 2682–2694

21   Kersting, W.H.: 'Distribution system modeling and analysis'. (CRC press, 2006)

22   Kersting, W.H. 'Radial Distribution Test Feeders'. In: IEEE Power Engineering Society Winter Meeting. vol. 2. (, 2001. pp. 908–912

23   Einfalt, A., Schuster, A., Leitinger, C., Tiefgraber, D., Litzlbauer, M., Ghaemi, S., et al.: 'ADRES-Concept: Konzeptentwicklung für ADRES-Autonome Dezentrale Regenerative EnergieSysteme', *TU Wien, Institut für Elektrische Anlagen und Energiewirtschaft*, 2011,

24   Institute of Energy Systems and Electrical Drives. 'ADRES-Dataset'. (, 2016. Available from: http://www.ea.tuwien.ac.at/projects/adres_concept/EN/

25   Chassin, D.P., Schneider, K., Gerkensmeyer, C. '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

26   Laughton, M. '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. (IET, 1968. pp. 1163–1172

27   Chen, T.H., Chen, M.S., Hwang, K.J., Kotas, P., Chebli, E.A.: 'Distribution System Power Flow Analysis-A Rigid Approach', *IEEE Transactions on Power Delivery*, 1991, **6**, (3), pp. 1146–1152

28   Kersting, W.H.: 'Distribution System Modeling and Analysis'. (CRC press, 2012)

29   Liao, Y., Weng, Y., Rajagopal, R. 'Urban distribution grid topology reconstruction via lasso'. In: 2016 IEEE Power and Energy Society General Meeting (PESGM). (, 2016. pp. 1–5

30   Chen, Y.C., Banerjee, T., Domínguez.García, A.D., Veeravalli, V.V.: 'Quickest Line Outage Detection and Identification', *IEEE Transactions on Power Systems*, 2016, **31**, (1), pp. 749–758

31   Deka, D., Backhaus, S., Chertkov, M. 'Estimating Distribution Grid Topologies: A Graphical Learning based Approach'. In: Power Systems Computation Conference. (IEEE, 2016. pp. 1–7

32   Cover, T.M., Thomas, J.A.: 'Elements of Information Theory'. (John Wiley & Sons, 2012)

33   Deka, D., Talukdar, S., Chertkov, M., Salapaka, M.: 'Topology Estimation in Bulk Power Grids: Guarantees on Exact Recovery', *arXiv preprint arXiv:170701596*, 2017,

34   Chow, C., Liu, C.: 'Approximating Discrete Probability Distributions with Dependence Trees', *IEEE Transactions on Information Theory*, 1968, **14**, (3), pp. 462–467

35   Wang, W., Yu, N., Foggo, B., Davis, J., Li, J. '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

36   Short, T.A.: 'Advanced Metering for Phase Identification, Transformer Identification, and Secondary Modeling', *IEEE Transactions on Smart Grid*, 2013, **4**, (2), pp. 651–658

37   Arya, V., Seetharam, D., Kalyanaraman, S., Dontas, K., Pavlovski, C., Hoy, S., et al. 'Phase Identification in Smart Grids'. In: Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on. (IEEE, 2011. pp. 25–30

38   Kruskal, J.B.: 'On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem', *Proceedings of the American Mathematical society*, 1956, **7**, (1), pp. 48–50

39   Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C., et al.: 'Introduction to Algorithms'. vol. 2. (MIT press Cambridge, 2001)

40   Liao, Y., Weng, Y., Wu, M., Rajagopal, R. 'Distribution grid topology reconstruction: An information theoretic approach'. In: North American Power Symposium. (, 2015. pp. 1–6

41   Cavraro, G., Kekatos, V., Veeramachaneni, S.: 'Voltage analytics for power distribution network topology verification', *IEEE Transactions on Smart Grid*, 2019, **10**, (1), pp. 1058–1067

42   Baran, M.E., Wu, F.F.: 'Network Reconfiguration in Distribution Systems for Loss Reduction and Load Balancing', *IEEE Transactions on Power delivery*, 1989, **4**, (2), pp. 1401–1407

43   Dugan, R., Arritt, R.: 'The ieee 8500-node test feeder', *Electric Power Research Institute, Palo Alto, CA, USA*, 2010,

44   Zhao, Y., Chen, J., Poor, H.V. 'Learning to Infer Power Grid Topologies: Performance and Scalability'. In: 2018 IEEE Data Science Workshop (DSW). (IEEE, 2018. pp. 215–219

45   Dobos, A.P.: 'PVWatts Version 5 Manual', *National Renewable Energy Laboratory, September*, 2014,

46   Ellis, A., Nelson, R., Von.Engeln, E., Walling, R., MacDowell, J., Casey, L., 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

47   IEEE: 'IEEE Guide for Conducting Distribution Impact Studies for Distributed Resource Interconnection', *IEEE Std 15477-2013*, 2014, pp. 1–137

48   Jabr, R.A.: 'Minimum Loss Operation of Distribution Networks with Photovoltaic Generation', *IET Renewable Power Generation*, 2014, **8**, (1), pp. 33–44

49   Dorostkar.Ghamsari, M.R., Fotuhi.Firuzabad, M., Lehtonen, M., Safdarian, A.: 'Value of Distribution Network Reconfiguration in Presence of Renewable Energy Resources', *IEEE Transactions on Power Systems*, 2016, **31**, (3), pp. 1879–1888

50   Zheng, J., Gao, D.W., Lin, L. 'Smart Meters in Smart Grid: An Overview'. In: IEEE Green Technologies Conference. (, 2013. pp. 57–64

51   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, 2010.

52   Chassin, D.P., Schneider, K., Gerkensmeyer, C. 'Gridlab-d: An open-source power systems modeling and simulation environment'. In: 2008 IEEE/PES Transmission and Distribution Conference and Exposition. (, 2008. pp. 1–5