Distribution Grid Impedance & Topology Estimation with Limited or No Micro-PMUs

Haoran Li[a], Yang Weng[a,*], Yizheng Liao[b], Brian Keel[c], Kenneth E. Brown[c]

[a] *Arizona State University, Tempe, AZ, 85281*
[b] *Stanford University, Stanford, CA, 94305*
[c] *Salt River Project, Phoenix, AZ, 85072*

Abstract

With more distributed energy resources (DERs) connected to distribution grids, better monitoring and control are needed, where accurate system topology and branch impedance are the prerequisites. However, this information is usually unknown or inaccurate due to limited observability in distribution grids. Therefore, the topology and branch impedance estimation methods are necessary for distribution grid operations. Among existed works, the regression-based methods have been frequently discussed, which leverages the estimated impedance matrix and graph theory to recover the radial topology. However, most of them assume that micro Phasor Measurement Units (micro-PMUs) have been placed at all load-nodes, which is unrealistic due to the sensor cost. In this paper, we target real cases where only nodes with high-variability loads have sensors, either micro-PMUs or smart meters. Firstly, we convert the micro-PMU-based or the smart meter-based impedance estimation to an Ordinary Least Square (OLS) estimation, where only measurements of observed nodes can be utilized for the estimation while hidden quantities cause errors. To quantify the error, we propose a Factorized Ordinary Least Squares (FOLS) method to decompose the error of OLS with the whitening matrix. Under Cholesky whitening, we theoretically provide the upper error bound and claim the error is negligible with

*This is to indicate the corresponding author.
Email addresses: lhaoran@asu.edu (Haoran Li), yang.weng@asu.edu (Yang Weng), yzliao@stanford.edu (Yizheng Liao), brian.keel@srpnet.com (Brian Keel), Ken.Brown@srpnet.com (Kenneth E. Brown)

our metering assumption. Finally, we introduce the Recursive Grouping algorithm to estimate the branch impedance & topology with hidden nodes. The numerical results demonstrate that the proposed algorithm achieves accurate results on real-world load data and IEEE standard test systems.

Keywords: impedance estimation, topology learning, distribution network, ordinary least square

---

1. Introduction

Distributed energy resources (DERs), including renewable energy sources, the electricity storage, and intelligent loads, have offered more controllability for system operators and more economic and sustainable choices for end-users. However, the large-scale deployment of DERs also brings numerous challenges to the stability and operation of distribution grids. For example, the rooftop PV panels can inject generated powers back to the distribution grids and cause the inverse power flows [1]. Also, the high penetration of PVs affects instant system power balancing [2]. For the low-voltage networks, DERs can cause the voltage rise and threaten the system reliability [3]. Thus, utilities and distribution grid operators need tools to monitor, control, and operate the systems under these profound changes.

As the foundation of the intelligent monitoring and control, the accurate impedance and most-updated system topology, are not always available due to limited instrumentation and low investment interest [4]. Further, the topology of the network may change frequently and cause the impedance matrix to have different structures and values. For example, in a city-level distribution network, the routine reconfiguration helps to achieve a good operational radial topology from a meshed network [5, 6]. The frequency of a topology change with PV panels [7] ranges from eight hours to once a month for medium-voltage grids [8]. In addition, the unpredictable events (e.g., power outages) and regular maintenance could result in a topology reconfiguration, which may not be well synchronized with all stakeholders. Finally, many devices, such as plug-and-play components, are not operated by the utilities and therefore, their connectives

and status are not always observed by the distribution grid operators.

In order to find the network topology, the growing coverage of smart meters and micro-PMUs offers new opportunities for data-driven approaches. For example, [9] uses terminal bus voltages and conducts conditional-independence tests to identify the connectivity. Similarly, [10] finds the topology using the bus voltage covariance. [11, 1, 12] formulate the grid topology as a probabilistic graphical model and learn the topology from the edge weights. [13] and [14] uses the second-order statistics of voltage and power injections to estimate grid topology. However, these works only estimate the topology but not the line impedance for further analysis and calculations.

With the emergence of distribution power flow analysis [15], the accurate line parameters also become important for system operations. In [16, 17], a joint-approach is proposed for estimating both topology and line parameters via a set of linear power flow equations with variable transformation. [18, 19] linearize the power flow equations and use the Ordinary Least Squares (OLS) approach to find line parameters. Similarly, [20, 21] use Ohm's Law to recover impedance. Though these works have demonstrated accurate estimation results in various data sets or simulations, some assumptions may be invalid in a real distribution system. For example, [16, 17] require all buses to have micro-PMU installed, which is unrealistic today due to high manufacturing cost. [18, 19] need a very special technique, inverter perturbations, to manually create data for the estimation. In [20, 21], the micro-PMUs are needed at buses with non-zero current injections. Though the requirement is less strict compared with other works, having micro-PMUs at every bus with power consumption still requires a significant amount of investments. In the real world, however, only partial power-consumed nodes have micro-PMUs or smart meters [4]. In this paper, we focus on the secondary distribution grid with partial nodes observable. However, if a subset of nodes in the primary distribution grid has meters, our algorithm can still work for that primary distribution system. Further, these nodes have high-variability loads and are of great importance to be monitored. The rest the of nodes, with either relatively stationary loads or zero loads, are usually without

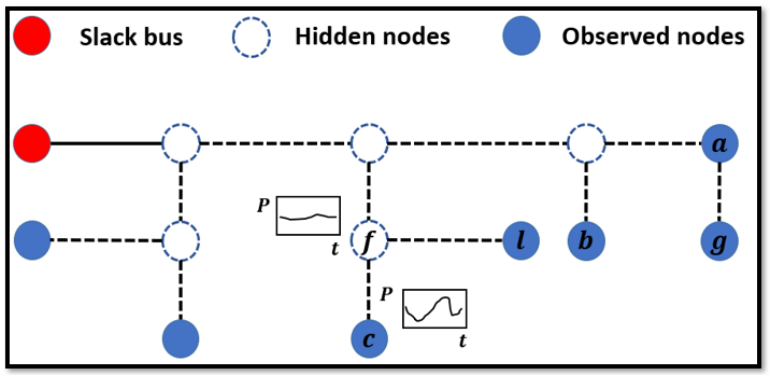


Figure 1: An illustration of the distribution network with observed nodes and hidden nodes.

sensors. Therefore, we denote the metered nodes as the observed nodes and the rest nodes as hidden nodes. For example, Fig. 1 illustrates the observability of a distribution network. In this figure, the power consumption of hidden nodes like node f changes relatively small while that of observed nodes like node c changes a lot. With this real-world observability assumption, the following question arises: can we estimate the impedance & topology with only data from observed nodes? To answer such a question, [21] clearly points out that only a radial grid can achieve the complete recovery of topology, given limited sensors measurements without other prior knowledge. Due to this limitation, we consider a radial grid for topology recovery in this paper. On the other hand, if we focus on the impedance estimation of observed nodes, the meshed network's impedance matrix can still be recovered via OLS-based estimation. Thus, considering loops in observed nodes as prior knowledge to recover the whole grid's topology can be a future research topic.

However, [21] still makes the unrealistic assumption that current injections of hidden nodes are zero. In this paper, we admit the relatively-stationary loads in hidden nodes and clearly quantify the error because of current injections from hidden nodes. Secondly, the other key contribution of our paper is the proposal of a unified estimation process for the following three scenarios of different grids and measurement types:

- Balanced distribution grids with micro-PMUs that have voltage and cur-

rent phasor.

- Balanced distribution grids with smart meters that have synchronized voltage magnitude, active and reactive power.

- Three-phase unbalanced distribution grids with micro-PMUs that have three-phase voltage and current phasor.

In such an estimation, OLS is formalized with linear representation, i.e., Ohm's law, linearized power flow equations [22], or linearized three-phase unbalanced power flow equations [23, 24, 25, 26], which completely covers the above three scenarios. Further, we partition the OLS into the observed part with measurements and the hidden part without measurements. Due to the missing quantities from the hidden nodes, the pure estimation of the observed part suffers errors. To quantify the error, we utilize the whitening matrix to factorize the OLS and achieve a Factorized Ordinary Least Square (FOLS) procedure, yielding a factorization of the error term with the whitening matrix. Then, we use Cholesky whitening to prove that the error is very small in real cases. Therefore, we prove that OLS with measurements of observed nodes delivers an accurate impedance estimation. Finally, with the input of the observed nodes' impedance, we introduce the so-called Recursive Grouping algorithm to estimate the hidden nodes' mutual impedance and topology. The performance of our algorithms is verified on multiple IEEE distribution systems with real-world load data from Pacific Gas and Electric Company (PG&E). The results show the high accuracy of our algorithms.

In summary, this paper has four main contributions:

- Proposal of OLS-based impedance estimation for balanced distribution grids with micro-PMUs or smart meters, and three-phase unbalanced distribution grids with micro-PMUs.

- Introduction of linearization methods for balanced networks with smart meters or three-phase unbalanced networks with micro-PMUs.

- Introduction of FOLS to quantify the error of the estimation.
- Comprehensive experiments with real-world load data and IEEE standard test systems.

The rest of the paper is organized as follows: Section 2 formulates the problem and introduces the OLS-based impedance matrix estimation. Section 3 illustrates FOLS for error quantification. Section 4 introduces the Recursive Grouping algorithm to estimate impedance among hidden nodes and identify the topology. Section 5 shows simulation and experimental results. Section 6 makes the conclusions.

2. Problem formulation and OLS estimation

In this section, we define the problem of data-driven impedance estimation. Then, the OLS-based impedance matrix estimation is illustrated as a foundation of our key contribution.

2.1. Problem formulation

A distribution system can be modelled as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is the bus set and $\mathcal{E}$ is the branch set. Due to the nodal measurement availability, let $\mathcal{O}$, $\mathcal{H}$ denote the observed node set and unobserved node set, respectively. Therefore, $\mathcal{O} \cup \mathcal{H} = \mathcal{V}$ and $\mathcal{O} \cap \mathcal{H} = \emptyset$. Correspondingly, we assume there are $L+1$ nodes (the slack bus is indexed with 0) for $\mathcal{V}$ and $M < L+1$ nodes for $\mathcal{O}$. For the sake of later derivation, M observed nodes have indices from $L-M+1$ to L.

For observed nodes in $\mathcal{O}$, there are either micro-PMUs or smart meters. Then, we presume voltage and current phasor data are available for micro-PMUs, and voltage magnitude and active/reactive power data are available for smart meters.

With such graph definition, for micro-PMUs in balanced distribution grids, we denote $\mathbf{V}, \mathbf{I} \in \mathbb{C}^{(L+1)\times N}$ as the nodal voltage and current phasor matrices, where N is the number of observations. Further, measurements from $\mathbf{V}_{\mathcal{O}}, \mathbf{I}_{\mathcal{O}}$ are available and from $\mathbf{V}_{\mathcal{H}}, \mathbf{I}_{\mathcal{H}}$ are hidden. Let $\mathbf{Y}$ be the admittance matrix and

Ohm's law is $\mathbf{I} = \mathbf{YV}$. Elimination of the slack bus-corresponded column and row in $\mathbf{Y}$, columns in $\mathbf{I}/\mathbf{V}$ makes $\mathbf{Y}$ invertible [27]. Further, if we consider the deviation of voltage and current from the sample mean, Ohm's equation after the subtraction from the sample mean and matrix elimination still holds. In the following derivation, we only consider deviation data from the sample mean in the eliminated version and keep the notation unchanged. With the data in network parameter estimation, [21] explains that due to the measurement correlations, estimation of impedance matrix $\mathbf{Z}$ will obtain a more robust result than that of admittance matrix $\mathbf{Y}$. Thus, we denote $\mathbf{Z} = \mathbf{Y}^{-1}$ and rewrite Ohm's law $\mathbf{V} = \mathbf{ZI}$. The block format of Ohm's law is as follows,

$$\begin{bmatrix} \mathbf{V}_{\mathcal{H}} \\ \mathbf{V}_{\mathcal{O}} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{\mathcal{H}\mathcal{H}} & \mathbf{Z}_{\mathcal{H}\mathcal{O}} \\ \mathbf{Z}_{\mathcal{O}\mathcal{H}} & \mathbf{Z}_{\mathcal{O}\mathcal{O}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{I}_{\mathcal{H}} \\ \mathbf{I}_{\mathcal{O}} \end{bmatrix}, \tag{1}$$

where we utilize subscripts $\mathcal{H}$ and $\mathcal{O}$ to represent the partition.

Similarly, for smart meters in balanced distribution grids, we denote $|\mathbf{V}|, \mathbf{P}, \mathbf{Q} \in \mathbb{R}^{(L+1)\times N}$ as the matrices of nodal voltage magnitude, active, and reactive power, where $|\cdot|$ for a complex matrix represents the element-wise operation to obtain the magnitude. These quantities can be constrained in a so-called Linear Coupled Power Flow (LC-PF) model [22], $|\mathbf{V}| = \mathrm{Re}(\mathbf{Z})\mathbf{P} + \mathrm{Im}(\mathbf{Z})\mathbf{Q}$, where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ represent the element-wise operation to obtain the real and imaginary part of the complex number. With observed and hidden nodes partitioning, the block format of LC-PF is:

$$\begin{aligned} \begin{bmatrix} |\mathbf{V}|_{\mathcal{H}} \\ |\mathbf{V}|_{\mathcal{O}} \end{bmatrix} = & \begin{bmatrix} \mathrm{Re}(\mathbf{Z}_{\mathcal{H}\mathcal{H}}) & \mathrm{Im}(\mathbf{Z}_{\mathcal{H}\mathcal{H}}) & \mathrm{Re}(\mathbf{Z}_{\mathcal{H}\mathcal{O}}) & \mathrm{Im}(\mathbf{Z}_{\mathcal{H}\mathcal{O}}) \\ \mathrm{Re}(\mathbf{Z}_{\mathcal{O}\mathcal{H}}) & \mathrm{Im}(\mathbf{Z}_{\mathcal{O}\mathcal{H}}) & \mathrm{Re}(\mathbf{Z}_{\mathcal{O}\mathcal{O}}) & \mathrm{Im}(\mathbf{Z}_{\mathcal{O}\mathcal{O}}) \end{bmatrix} \\ & \cdot \Big[\mathbf{P}_{\mathcal{H}}, \mathbf{Q}_{\mathcal{H}}, \mathbf{P}_{\mathcal{O}}, \mathbf{Q}_{\mathcal{O}}\Big]^T, \end{aligned} \tag{2}$$

where T represent the transpose operation. Similarly, we consider the measurements' deviation from the sample mean and keep the notation unchanged. Different from the partitioning in Kron Reduction [20, 21], we do not presume the current or power measurements in $\mathcal{H}$ to be zero or constant, which leads to the application feasibility in the most realistic cases.

Finally, for three-phase unbalanced distribution grids with micro-PMUs, we denote $|\mathbf{V}^{A,B,C}|, \mathbf{\Theta}^{A,B,C}, \mathbf{P}^{A,B,C}, \mathbf{Q}^{A,B,C} \in \mathbb{R}^{(L+1)\times 3N}$ as matrices of voltage magnitude, angle, active power and reactive power of three-phase A, B, C, where different phases' data of N columns are concatenated one after another. With such definitions, we employ the linearized three-phase unbalanced power flow equations from [23]:

$$\begin{bmatrix} |\mathbf{V}|_{\mathcal{H}}^{A,B,C} \\ \mathbf{\Theta}_{\mathcal{H}}^{A,B,C} \\ |\mathbf{V}|_{\mathcal{O}}^{A,B,C} \\ \mathbf{\Theta}_{\mathcal{O}}^{A,B,C} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{\mathcal{H}\mathcal{H}}^{A,B,C} & \mathbf{O}_{\mathcal{H}\mathcal{H}}^{A,B,C} & \mathbf{L}_{\mathcal{H}\mathcal{O}}^{A,B,C} & \mathbf{O}_{\mathcal{H}\mathcal{O}}^{A,B,C} \\ \mathbf{U}_{\mathcal{H}\mathcal{H}}^{A,B,C} & \mathbf{T}_{\mathcal{H}\mathcal{H}}^{A,B,C} & \mathbf{U}_{\mathcal{H}\mathcal{O}}^{A,B,C} & \mathbf{T}_{\mathcal{H}\mathcal{O}}^{A,B,C} \\ \mathbf{L}_{\mathcal{O}\mathcal{H}}^{A,B,C} & \mathbf{O}_{\mathcal{O}\mathcal{H}}^{A,B,C} & \mathbf{L}_{\mathcal{O}\mathcal{O}}^{A,B,C} & \mathbf{O}_{\mathcal{O}\mathcal{O}}^{A,B,C} \\ \mathbf{U}_{\mathcal{O}\mathcal{H}}^{A,B,C} & \mathbf{T}_{\mathcal{O}\mathcal{H}}^{A,B,C} & \mathbf{U}_{\mathcal{O}\mathcal{O}}^{A,B,C} & \mathbf{T}_{\mathcal{O}\mathcal{O}}^{A,B,C} \end{bmatrix} \cdot \left[\mathbf{P}_{\mathcal{H}}^{A,B,C}, \mathbf{Q}_{\mathcal{H}}^{A,B,C}, \mathbf{P}_{\mathcal{O}}^{A,B,C}, \mathbf{Q}_{\mathcal{O}}^{A,B,C}\right]^{T}, \tag{3}$$

where $\mathbf{L}, \mathbf{O}, \mathbf{U}, \mathbf{T}$ are the inverse of three-phase admittance matrix proposed in [23]. The linear format of Eqn. (3) also enables us to consider measurement deviation data and keep the notation unchanged.

Given the above measurement availability, we have the following problem formulation:

- Problem: distribution network impedance estimation and topology learning using measurements from limited sensors.
- Given: historical measurements of voltage/current phasor from micro-PMUs or voltage magnitude, active and reactive power from smart meters.
- Find: estimated impedance matrix $\widehat{\mathbf{Z}}$ and graph $\widehat{\mathcal{G}}$.

In general, Fig. 2 illustrates the flow chart of the impedance & topology estimation and the error quantification. Given the input data, the Ordinary Least Square (OLS) helps to find the observed-block of the impedance matrix. Then, the Recursive Grouping (RG) algorithm identifies the connectivity among all nodes. Further, RG estimates the line impedance with respect to the hidden nodes $\mathcal{H}$, yielding the estimation of the complete impedance matrix $\mathbf{Z}$. The above estimation, based on Eqn. (1), owns systematic error due to the missing

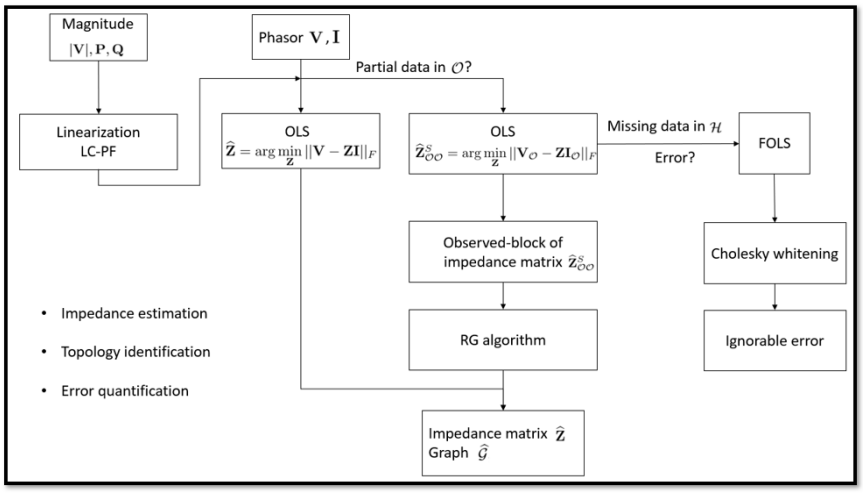


Figure 2: The flow chart of the impedance & topology estimation and the estimation error quantification.

data of $\mathcal{H}$. Therefore, we quantify the error with the whitening-based Factorized Ordinary Least Square (FOLS) and Cholesky whitening. Finally, we show the error is ignorable and the illustrated estimation yields an accurate result.

Remark: though the data types are different, Eqn. (1), Eqn. (2), and Eqn. (3) are all linear and have the same partitioned format with respect to $\mathcal{H}$ and $\mathcal{O}$. This convenience brings the same estimation and error-analysis format in the later derivation. Thus, we only consider Eqn. (1) for the sake of simplicity.

2.2. OLS-based impedance matrix estimation

[16, 17, 18, 19, 20, 21] provide different OLS models based on different linearized power flow equations. For Ohm's law, if $\mathcal{O} = \mathcal{V}$, the OLS problem becomes:

$$\widehat{\mathbf{Z}} = \arg\min_{\mathbf{Z}} ||\mathbf{V} - \mathbf{ZI}||_F = (\mathbf{II}^H)^{-1}\mathbf{IV}^H, \tag{4}$$

where H is the conjugate-transpose operation and $||\cdot||_F$ is the Frobenius norm.

If $\mathcal{O} \subset \mathcal{V}$, we target at estimating $\mathbf{Z}_{\mathcal{O}\mathcal{O}}$ in Eqn. (1) due to the fact that for a tree structure, the whole impedance matrix $\mathbf{Z}$ can be accurately estimated when we input $\mathbf{Z}_{\mathcal{O}\mathcal{O}}$ to a graph-based algorithm, Recursive Grouping (RG) [22, 20, 21]. For such a goal, the OLS in Eqn. (4) should be factorized into observed and

hidden parts. Given $\mathbf{V}_{\mathcal{O}}$ and $\mathbf{I}_{\mathcal{O}}$ data, the observed part of the OLS has an estimator,

$$\widehat{\mathbf{Z}}^{S}_{\mathcal{O}\mathcal{O}} = \arg\min_{\mathbf{Z}} ||\mathbf{V}_{\mathcal{O}} - \mathbf{Z}\mathbf{I}_{\mathcal{O}}||_F = (\mathbf{I}_{\mathcal{O}}\mathbf{I}^H_{\mathcal{O}})^{-1}\mathbf{I}_{\mathcal{O}}\mathbf{V}^H_{\mathcal{O}}, \tag{5}$$

where we utilize the superscript S to imply that the estimation uses only a subset of data. Thus, the estimated result is not the same as $\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}$, the block matrix of $\widehat{\mathbf{Z}}$ in Eqn. (4). For the sake of later derivations, S is added to the result whenever it is obtained via the observed subset of the input data.

The estimation in Eqn. (5) is biased since by Eqn. (1), $\mathbf{V}_{\mathcal{O}} = \mathbf{Z}_{\mathcal{O}\mathcal{H}}\mathbf{I}_{\mathcal{H}} + \mathbf{Z}_{\mathcal{O}\mathcal{O}}\mathbf{I}_{\mathcal{O}}$, causing the error term $(\mathbf{I}_{\mathcal{O}}\mathbf{I}^H_{\mathcal{O}})^{-1}\mathbf{I}_{\mathcal{O}}\mathbf{Z}_{\mathcal{O}\mathcal{H}}\mathbf{I}_{\mathcal{H}}$. Such a term can't be directly quantified due to the unknown quantities $\mathbf{Z}_{\mathcal{O}\mathcal{H}}\mathbf{I}_{\mathcal{H}}$. However, intuitively, the error is related to the ratio of observed-hidden current mutual-correlation $\mathbf{I}_{\mathcal{O}}\mathbf{I}^H_{\mathcal{H}}$ to the observed current self-correlation $\mathbf{I}_{\mathcal{O}}\mathbf{I}^H_{\mathcal{O}}$. Thus, in the next section, we decorrelate the current deviations with the whitening transformation. Then, the correlation-based estimator is converted to a whitening matrix-based estimator, where the structure of the whitening matrix leads to a clear error quantification.

3. Whitening-based error quantification for FOLS

In this section, we propose the Factorized Ordinary Least Square (FOLS) via rewriting the OLS with the whitening matrix, the factorization of the covariance matrix. Then, we prove that under the realistic data-correlation assumption, the error is negligible and the closed-form solution of the OLS in Eqn. (5) yields an accurate solution.

3.1. Whitening for FOLS

The whitening transformation is a linear transformation that helps to transform the sample covariance matrix to a identity matrix. Specifically, let $\mathbf{W}$ denote the whitening matrix and $\tilde{\mathbf{I}} = \mathbf{W}\mathbf{I}$ be the transformed measurements, we have: $\tilde{\mathbf{I}}\tilde{\mathbf{I}}^H = \mathbf{Id}$, where $\mathbf{Id}$ is the identity matrix. To achieve such a goal, the whitening matrix satisfies $\mathbf{W}^H\mathbf{W} = \mathbf{\Sigma}^{-1}$, where $\mathbf{\Sigma} = \mathbf{I}\mathbf{I}^H$ is the covariance matrix of the current deviations $\mathbf{I}$. In general, the whitening matrix is a decomposed matrix of $(\mathbf{I}\mathbf{I}^H)^{-1}$, enabling a factorized version of the OLS in (4) and (5).

Specifically, with the voltage $\mathbf{V}$ and the transformed current deviations $\tilde{\mathbf{I}}$, the FOLS is:

$$\begin{aligned}\widehat{\mathbf{K}} &= \arg\min_{\mathbf{K}} ||\mathbf{V} - \mathbf{K}\tilde{\mathbf{I}}||_F, \\ &= \mathbf{W}\mathbf{I}\mathbf{V}^H,\end{aligned} \tag{6}$$

where the second equality holds by the definition of whitening matrix $\mathbf{W}$. Combing Eqn. (6) and Eqn. (4), we have $\widehat{\mathbf{Z}} = \mathbf{W}^H\widehat{\mathbf{K}}$. Thus, estimation of $\mathbf{Z}$ is decomposed into 2 steps: 1) use FOLS in Eqn. (6) to estimate $\widehat{\mathbf{K}}$ and 2) use $\widehat{\mathbf{K}}$ and $\mathbf{W}^H$ to calculate $\widehat{\mathbf{Z}}$. Though the 2-step estimation result is theoretically the same as the result of the Eqn. (4), the specifically designed whitening matrix $\mathbf{W}$ makes the error quantification possible when only partial data is observed. For the observed OLS in Eqn. (5), the whitening-based 2-step estimation is:

$$\begin{aligned}\widehat{\mathbf{K}}^S_{\mathcal{O}\mathcal{O}} &= \arg\min_{\mathbf{K}} ||\mathbf{V}_{\mathcal{O}} - \mathbf{K}\tilde{\mathbf{I}}^S_{\mathcal{O}}||_F, \\ &= \mathbf{W}^S_{\mathcal{O}\mathcal{O}}\mathbf{I}_{\mathcal{O}}\mathbf{V}^H_{\mathcal{O}}, \\ \widehat{\mathbf{Z}}^S_{\mathcal{O}\mathcal{O}} &= (\mathbf{W}^S_{\mathcal{O}\mathcal{O}})^H\widehat{\mathbf{K}}^S_{\mathcal{O}\mathcal{O}},\end{aligned} \tag{7}$$

where we similarly employ the superscript S to show that $\widehat{\mathbf{K}}^S_{\mathcal{O}\mathcal{O}}$, $\tilde{\mathbf{I}}^S_{\mathcal{O}}$, and $\mathbf{W}^S_{\mathcal{O}\mathcal{O}}$ only calculated from the observed measurements and may not be the partitioned matrices of $\widehat{\mathbf{K}}$, $\tilde{\mathbf{I}}$, and $\mathbf{W}$. To distinguish them, we denote the partitioned matrices as $\widehat{\mathbf{K}}_{\mathcal{O}\mathcal{O}}$, $\tilde{\mathbf{I}}_{\mathcal{O}}$, and $\mathbf{W}_{\mathcal{O}\mathcal{O}}$, respectively.

Eqn. (7) enable the quantification of the error from $\widehat{\mathbf{Z}}^S_{\mathcal{O}\mathcal{O}}$ to $\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}$. For such a goal, finding a right structure of $\mathbf{W}$ with proper inner partitioning is the key part.

3.2. Choice of Whitening: Cholesky Whitening

Cholesky whitening transformation with an upper diagonal structure proves to be the right choice for the error quantification. Specifically, Cholesky decomposition implies $\mathbf{U}^H\mathbf{U} = \mathbf{\Sigma}^{-1}$, where $\mathbf{U}$ is the upper diagonal matrix. $\mathbf{U}$ is unique when Σ is Hermitian positive-definite, a condition easy to be satisfied when the number of samples is enough. By the definition of the whitening matrix $\mathbf{W}$, we have $\mathbf{W} = \mathbf{U}$ under Cholesky whitening. Further, the Cholesky whitening matrix is invertible [28] and let $\mathbf{M} = \mathbf{W}^{-1}$ ($\mathbf{M}^S_{\mathcal{O}\mathcal{O}} = (\mathbf{W}^S_{\mathcal{O}\mathcal{O}})^{-1}$) be

the inverse matrix that is also upper-diagonal. Then, we claim the following theorem.

Theorem 1. *With data partitioning in Eqn.* (1) *and for the Cholesky whitening of the current deviations,* $\mathbf{M}^S_{\mathcal{O}\mathcal{O}} = \mathbf{M}_{\mathcal{O}\mathcal{O}}$*,* $\mathbf{W}^S_{\mathcal{O}\mathcal{O}} = \mathbf{W}_{\mathcal{O}\mathcal{O}}$*, and* $\tilde{\mathbf{I}}^S_{\mathcal{O}} = \tilde{\mathbf{I}}_{\mathcal{O}}$*.*

The proof of Theorem 1 is in Appendix 7.1. Combining Theorem 1, Eqn. (6) and (7), we have:

$$\begin{aligned} \widehat{\mathbf{K}}_{\mathcal{O}\mathcal{O}} &= \mathbf{W}_{\mathcal{O}\mathcal{O}}\mathbf{I}_{\mathcal{O}}\mathbf{V}^H_{\mathcal{O}}, \\ &= \mathbf{W}^S_{\mathcal{O}\mathcal{O}}\mathbf{I}_{\mathcal{O}}\mathbf{V}^H_{\mathcal{O}}, \\ &= \widehat{\mathbf{K}}^S_{\mathcal{O}\mathcal{O}}, \end{aligned} \tag{8}$$

where the first equality holds by the upper diagonal structure of $\mathbf{W}$ matrix and the second equality holds by Theorem 1. Therefore, by Eqn. (7), $\widehat{\mathbf{Z}}^S_{\mathcal{O}\mathcal{O}} = \mathbf{W}^H_{\mathcal{O}\mathcal{O}}\widehat{\mathbf{K}}_{\mathcal{O}\mathcal{O}}$. Further, since matrix $\mathbf{M}$, instead of matrix $\mathbf{W}$, is the decomposed matrix of the covariance matrix. To analyze the effect of current-deviation correlations to the estimation error, we consider $\widehat{\mathbf{K}} = \mathbf{M}^H\widehat{\mathbf{Z}}$ and $\widehat{\mathbf{K}}_{\mathcal{O}\mathcal{O}} = \mathbf{M}^H_{\mathcal{O}\mathcal{O}}\widehat{\mathbf{Z}}^S_{\mathcal{O}\mathcal{O}}$, where the former equation yields the partitioned formula:

$$\widehat{\mathbf{K}}_{\mathcal{O}\mathcal{O}} = \mathbf{M}^H_{\mathcal{H}\mathcal{O}}\widehat{\mathbf{Z}}_{\mathcal{H}\mathcal{O}} + \mathbf{M}^H_{\mathcal{O}\mathcal{O}}\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}. \tag{9}$$

Consequently, the estimation error comes from off-diagonal block matrix $\mathbf{M}^H_{\mathcal{H}\mathcal{O}}$, determined by the mutual-correlations between the observed nodes and the hidden nodes.

The errors of the above FOLS-based impedance estimation can be categorized into (a) the systematic error due to the missing values of hidden nodes, (b) the error from measurement noise, and (c) the power flow linearization error of equations (2) and (3) (i.e., the goodness-of-fit of realistic data for equations (2) and (3)). In the experiments, we find that for balanced grids, errors from (b) and (c) are very small due to the limited noise and good performance of LC-PF model. For unbalanced grids, errors from (b) are small while errors from (c) are large due to the relatively bad performance of the three-phase linearized model in (3). Finding a better linearization model for the unbalanced grid can

be a promising topic to improve the impedance estimation in three-phase unbalanced grids. In this paper, we only focus on the systematic error caused by the missing data of hidden nodes. To quantify the error, we consider ρ_1 and ρ_2 as the Pearson correlation coefficient between two different observed nodes and one hidden node and one observed node, respectively. We assume such correlation coefficients to have unified values because we want to focus on how the magnitude of the coefficients can affect the line-wise impedance error. In reality, these correlation coefficients are calculated from a pair of nodes' current deviations or active/reactive power deviations. Though the calculated values are not unified like ρ_1 and ρ_2, our simulation results in Table 1 in Experiment show that the error bound based on ρ_1 and ρ_2 can safely limit the true error. Finally, we want the error ratio to be less than η, e.g. the percent error ratio to be less than 5%. Then, we have the following theorem.

Theorem 2. *Given the above definitions, for N_1 number of observed nodes and N_2 number of hidden nodes, the upper bound of the error ratio η is satisfied if:*

$$|\rho_2| \leq \frac{\eta|(1+\rho_1(N_1-1))|}{N_2}. \tag{10}$$

We have the following proof for this theorem.

Proof. Based on equations (6), (8) and Theorem 1 in my paper, the equation between true impedance matrix $\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}$ and the estimated impedance matrix $\widehat{\mathbf{Z}}^S_{\mathcal{O}\mathcal{O}}$ can be written as:

$$\begin{aligned} \widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}} &= \widehat{\mathbf{Z}}^S_{\mathcal{O}\mathcal{O}} - (\mathbf{M}^H_{\mathcal{O}\mathcal{O}})^{-1}\mathbf{M}^H_{\mathcal{H}\mathcal{O}}\widehat{\mathbf{Z}}_{\mathcal{H}\mathcal{O}}, \\ &= \widehat{\mathbf{Z}}^S_{\mathcal{O}\mathcal{O}} - (\mathbf{\Sigma}_{\mathcal{O}\mathcal{O}})^{-1}\mathbf{\Sigma}_{\mathcal{O}\mathcal{H}}\widehat{\mathbf{Z}}_{\mathcal{H}\mathcal{O}}, \end{aligned} \tag{11}$$

where the second equality is due to the definition of $\mathbf{M}$ matrix:

$$\begin{bmatrix} \mathbf{M}_{\mathcal{H}\mathcal{H}} & \mathbf{M}_{\mathcal{H}\mathcal{O}} \\ & \mathbf{M}_{\mathcal{O}\mathcal{O}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{M}^H_{\mathcal{H}\mathcal{H}} & \\ \mathbf{M}^H_{\mathcal{H}\mathcal{O}} & \mathbf{M}^H_{\mathcal{O}\mathcal{O}} \end{bmatrix} = \begin{bmatrix} \mathbf{\Sigma}_{\mathcal{H}\mathcal{H}} & \mathbf{\Sigma}_{\mathcal{H}\mathcal{O}} \\ \mathbf{\Sigma}_{\mathcal{O}\mathcal{H}} & \mathbf{\Sigma}_{\mathcal{O}\mathcal{O}} \end{bmatrix}.$$

Equation (11) indicates that the covariance between the hidden nodes and the observed nodes causes the error. To further investigate the error quantity

for each impedance matrix entry, we denote z to be the impedance of each entry in $\widehat{\mathbf{Z}}_{\mathcal{OO}}$ and $\widehat{\mathbf{Z}}_{\mathcal{HO}}$ for the sake of the derivation convenience. Recall that Lemma 1 in our paper claims that the entry of the impedance matrix is the sum of line impedances of the common path for two nodes, the entry in $\widehat{\mathbf{Z}}_{\mathcal{OO}}$ is usually larger than that in $\widehat{\mathbf{Z}}_{\mathcal{HO}}$. Thus, our simplification essentially enlarges the error, but the derivation is valid because we seek for the upper bound of the error ratio.

Then, equation (11) indicates the element-wise absolute error term err as:

$$err = \frac{N_2|\rho_2|}{|1+\rho_1(N_1-1)|}z \leq \eta z, \tag{12}$$

Thus, to satisfy the upper bound, we need the condition in equation (10). ■

This theorem shows that to guarantee a fixed upper bound of the error ratio (η), the only requirement for hidden nodes' current deviation lies in the correlation coefficients (ρ_2) between the hidden nodes and the observed nodes. Further, the larger the correlation coefficients of current deviations of observed nodes (ρ_1) is, the higher tolerance of ρ_2 is. Since the hidden nodes are either without loads [20, 21] or with relatively stationary conventional loads [4] while the observed nodes are usually with strong time-varying loads or generations, e.g., DERs, we can claim that $|\rho_1| \gg |\rho_2|$, causing the feasibility to claim for a small upper bound of the error ratio. The proposed theorem can help us quantify the condition we need for a small error and in practice, we can often achieve such a condition for a valid estimation.

Further, we discuss the impact of meter placement on the estimation accuracy with the following two aspects. First, if some hidden nodes are metered, the number of hidden nodes is decreasing and the absolute error caused by hidden nodes will be smaller. Thus, we will have a better impedance estimation result.

Second, if some observed (N_3 number of) nodes are not placed with meters,

we have the new requirements for ρ_2:

$$\frac{-\eta-\rho_1(\eta(N_1-1)+N_3(1-\eta))}{N_2}\leq\rho_2\leq\frac{\eta+\rho_1(\eta(N_1-1)-N_3(1+\eta))}{N_2}, \tag{13}$$

If $N_3=0$, the criteria in (13) will reduce to (10). Otherwise, the total range of ρ_2 in (13) is equal to the original range in (10) minus $\frac{2\rho_1 N_3}{N_2}$. In conclusion, the range of ρ_2 is reduced by $\frac{2\rho_1 N_3}{N_2}$.

With the estimated impedance matrix of the observed part, the Recursive Grouping (RG) algorithm can help to learn the structure.

4. Structure learning with RG algorithm

In this section, we introduce the so-called Recursive Grouping (RG) algorithm to learn the structure of a radial distribution grid. Specifically, RG algorithm inputs additive "distances" among all observed nodes to detect hidden nodes and calculate the "distance" between hidden nodes and the observed nodes. Thus, the detected hidden nodes are considered to be observed nodes since their connections and impedances are found. Then, the above process is repeated recursively to identify the structure. In our scenario, the "distance" is defined as the line impedance since it is physically additive for the series-connected lines.

To obtain the line impedance from the estimated impedance matrix $\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}$, we have the following Lemma 1[29, 27].

Lemma 1. *In a radial distribution grid,* $\mathbf{Z}(a,b)$ *is the sum for line impedances of the common path between nodes* a*,* b *to the slack bus,* $a,b\in\{1,2,3,\cdots,L\}$.

Therefore, the distance between every two observed nodes ($\forall a,b\in\mathcal{O}$) can be calculated as:

$$\hat{d}_{ab}=\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}(a,a)+\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}(b,b)-\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}(a,b)-\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}(b,a). \tag{14}$$

To illustrate the algorithm, we utilize the true value d_{ab} calculated from $\mathbf{Z}_{\mathcal{O}\mathcal{O}}$ to take place of the estimated value $\hat{d}_{ab}$ and prove that the true impedance matrix $\mathbf{Z}_{\mathcal{O}\mathcal{O}}$ can lead to the exact impedance and topology recovery among hidden nodes.

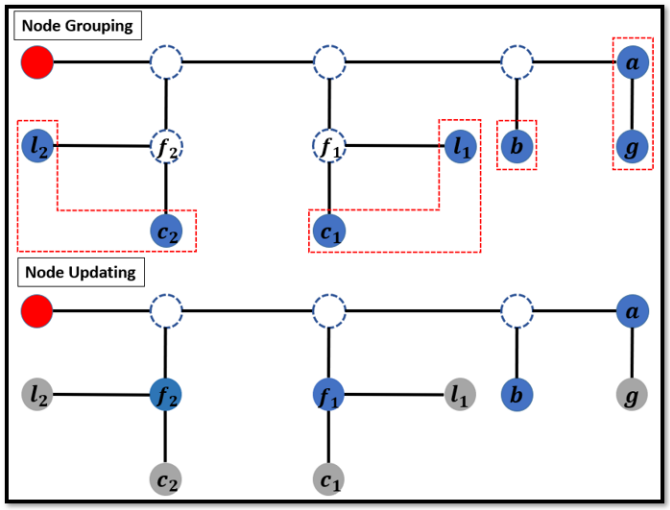


Figure 3: Illustration of the Recursive Grouping (RG) algorithm.

Specifically, RG algorithm includes the process of *Node Grouping* and *Node Updating*. In *Node Grouping*, the current observed-node set $\mathcal{Y}$, starting from $\mathcal{O}$, is partitioned is partitioned into different groups according to their connections with hidden nodes, indicating the location of hidden nodes. In *Node Updating*, the impedance between current observed nodes and hidden nodes is estimated, converting the detected hidden nodes to observed nodes and forming the new observed-node set $\mathcal{Y}_{new}$. In the meantime, we denote $\mathcal{Y}_{old} = \mathcal{Y}$ for later illustration. Subsequently, repeating these two processes generates the complete structure and impedance matrix. Before explaining the details, we define the following concepts for the sake of later illustrations. For a given node a, its **child node** g is the node 1-hop further from the root node (i.e., the slack bus), and we denote a as the **parent node** of g. If nodes c_1, l_1 share the same parent node f_1, we denote c_1 and l_1 as **siblings**. The set of siblings of node f_1 is denoted as $\mathcal{C}(f_1) = \{c_1, l_1\}$. Fig. 3 is an illustration of above definitions.

With the definitions, we show the details of the two processes as follows.

Node Grouping. To group $\mathcal{Y}$, we define $\Phi_{ijk} = d_{ik} - d_{jk}, \forall i, j \in \mathcal{Y}$. Then, [30] proves that if $\Phi_{ijk} = d_{ij}$ $(\forall k \in \mathcal{Y} \backslash \{i, j\})$, i is a child node and j is a parent of i, and vise versa. If $\Phi_{ijk} (\forall k \in \mathcal{Y} \backslash i, j)$ is constant but not equal to either

d_{ij} or $-d_{ij}$, i and j are observed nodes and they are siblings. For example, we find in Fig. 3, node a is the parent of node g since $\Phi_{agk} = d_{ak} - d_{gk} = -d_{ga}, \forall k \in \mathcal{Y}\backslash\{a, g\}$. *Node Grouping* categorizes the current-node set $\mathcal{Y}$ into different partitions $\{\Pi_q\}_{q=1}^{Q}$ with parent-child relationships like a, g, siblings like c_1, l_1 or a singleton like b in Fig. 3.

Node Updating. In this process, singleton, the parent node and the hidden node like node b, a, f_1 and f_2 are included to form the new current observed-node set $\mathcal{Y}_{new}$. For example, $\{a, b, f_1, f_2\}$ in Fig. 3. Then, the impedance among $\mathcal{Y}_{new}$ should be recalculated for the next iteration's *Node Grouping*.

For the hidden impedance estimation, we first illustrate how the impedance between a parent hidden node and a child observed node can be evaluated. Considering $c_1, l_1 \in \mathcal{C}(f_1)$, we denote $k \in \mathcal{Y}_{old} \setminus \{c_1, l_1\}$ be an arbitrary node. Knowing that $d_{c_1f_1} - d_{l_1f_1} = d_{c_1k} - d_{l_1k} = \Phi_{c_1l_1k}$ and $d_{c_1f_1} + d_{l_1f_1} = d_{c_1l_1}$, we calculate the distance between c_1 and f_1 as follows: $d_{c_1f_1} = \frac{1}{2}(d_{c_1l_1} + \Phi_{c_1l_1k})$.

With the above calculation and the impedance among $\mathcal{Y}_{old}$, we can calculate 3 types of impedance among nodes in $\mathcal{Y}_{new}$.

- Impedance between the observed parent node a and the detected hidden node f_1. $d_{af_1} = d_{ac_1} - d_{cf_1}$.
- Impedance between the detected hidden nodes f_1 and f_2. $d_{f_1f_2} = d_{c_1c_2} - d_{c_1f_1} - d_{c_2f_2}$.
- Impedance between the singleton b and the detected hidden node f_1. $d_{bf_1} = d_{bc_1} - d_{c_1f_1}$

The Recursive Grouping algorithm is introduced in Algorithm 1. We define RG algorithm as $\mathcal{G} = \mathcal{RG}(\mathbf{D}_{\mathcal{OO}})$, where $\mathbf{D}_{\mathcal{OO}}$ is the distance matrix such that $\mathbf{D}(a, b) = \mathbf{D}(b, a) = d_{ab}, \forall a, b \in \mathcal{O}$.

Finally, we form the joint impedance estimation and structure learning in Algorithm 2.

Remark: The choice of distance measure for the RG algorithm can be flexible to the realistic restrictions. For example, if the distribution grid contains

Algorithm 1 Recursive Grouping Algorithm for Structure Learning: $\mathcal{RG}$

Input: Distances matrix $D_{\mathcal{O}\mathcal{O}}$ of the observed nodes.

Output: Graph of the tree $\mathcal{G}$ and distance matrix $\mathbf{D}$.

Initialize $\mathcal{Y} := \mathcal{O}$.

while $|\mathcal{Y}| \geq 3$ **do**

    Compute $\Phi_{ijk} = d_{ik} - d_{jk}$ for all $i, j, k \in \mathcal{Y}$.

    Use "Node Grouping" to partition $\mathcal{Y}$ into $\{\Pi_q\}_{q=1}^{Q}$.

    Use "Node Updating" to update $\mathcal{Y}_{new}$.

    $\mathcal{Y}_{old} = \mathcal{Y}$ and $\mathcal{Y} = \mathcal{Y}_{new}$.

    For each hidden node $h \in \mathcal{Y}$, compute $d_{hp},\ \forall p \in \mathcal{Y}$.

end while

if $|\mathcal{Y}| = 2$ **then**

    Connect the two remaining nodes in $\mathcal{Y}$ with an edge and then stop.

else

    if $|\mathcal{Y}| = 1$ **then**

        Stop

    end if

end if

many underground cables with a comparatively large shunt susceptance, the estimated impedance is not additive. However, we can utilize the line resistance of observed nodes as input to the RG algorithm for topology learning.

Algorithm 2 Joint Structure Learning and Impedance Estimation

1: **Input**: The centralized current and voltage measurements of observed nodes $\mathcal{O}$, $\mathbf{V}_{\mathcal{O}}$ and $\mathbf{I}_{\mathcal{O}}$.

2: **Output**: The tree graph $\mathcal{G}$ of the distribution grid and the impedance matrix $\mathbf{Z}$.

3: Utilize Eqn. (5) to estimate the observed impedance matrix $\widehat{\mathbf{Z}}_{\mathcal{O}\mathcal{O}}$

4: Use Eqn. (14) to calculate the distance $\hat{d}_{ab},\ \forall a, b \in \mathcal{O}$, thus forming the estimated distance matrix $\widehat{\mathbf{D}}_{\mathcal{O}\mathcal{O}}$ of observed nodes.

5: $(\widehat{\mathcal{G}},\ \widehat{\mathbf{D}}) \leftarrow \mathcal{RG}(\widehat{\mathbf{D}}_{\mathcal{O}\mathcal{O}}, \mathcal{O})$.

6: Utilize the completed distance matrix $\mathbf{D}$ to reformulate the impedance matrix $\mathbf{Z}$ with Lemma 1.

5. Numerical experiments

We test IEEE distribution 8-, 19-, 33-bus, and 37-bus systems with partial nodes' data available, where 37-bus case is three-phase unbalanced. The unbalanced loads of the 37-bus system can be seen in [23, 31]. These measurements, either complex voltage/current or voltage magnitude and real/reactive power, are simulated from MATPOWER. Specifically, we employ the Pacific Gas and Electric Company's (PG&E's) hourly residential load profile (8760 samples for one year) in North California [1], the 15-min resolution residential data from Arizona, and the 1-second resolution household data from Upper Austria [32] for simulations. We assume smart meters have 1 measurement per 15 minutes and the micro-PMUs have 30 samples per second. Then, we interpolate PG&E's data and Austria's household data with the above sampling resolution to simulate data from smart meters and micro-PMUs. Specifically, we utilize cubic spline interpolation that conducts a piece-wise curve fitting with a set of 3-degree polynomial functions [33]. Thus, the interpolation can easily capture highly-nonlinear components of the load curve. As for hidden nodes $\mathcal{H}$, we

simulate the active power as multivariate Gaussian distribution $\mathcal{N} \sim (\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where we assume $\boldsymbol{\mu} = \mathbf{P}_{\mathcal{H}}^{base}$ to represent the base loads of hidden nodes from MATPOWER case file and $\boldsymbol{\Sigma}$ to be the covariance matrix with $(0.05)^2$ for the diagonal element and $(0.005)^2$ for the off-diagonal element. These nodes are essentially the stationary loads and the average Pearson correlation coefficient between the observed and the hidden nodes is 0.022. For each node $i \in \mathcal{V}$, the reactive power $q_i(n) = p_i(n)\sqrt{1 - pf_i(n)^2}/pf_i(n)$, where for each time n, we consider a random power factor, $pf_i(n) \sim Unif(0.85, 0.95)$. These settings can lead to realistic simulations based on MATPOWER.

The resulted voltage magnitude, active power, and reactive power form samples of smart meters, and voltage/current phasor form samples of micro-PMUs. In general, we summarize the data name, measurement types, and resolutions of the input data to our algorithms.

- $PG\&E, 1$. The dataset includes voltage and current phasor measurements with a resolution of 30 samples per second. These measurements are results of the simulation with interpolated $PG\&E$ load data.

- $PG\&E, 2$. The dataset includes voltage magnitude and power measurements with a resolution of 15 samples per minute. These measurements are results of the simulation with interpolated $PG\&E$ load data.

- *Aust*. The dataset includes voltage and current phasor measurements with a resolution of 30 samples per second. These measurements are results of the simulation with interpolated load data from Upper Austria.

- *Ariz*. The dataset includes voltage and power measurements with a resolution of 15 samples per minute. These measurements are results of the simulation with 15-minute resolution load data from Arizona.

Totally, we introduce 1440 samples from smart meter within 15 days and $18,000$ samples from micro-PMUs within 10 minutes. Subsequently, we vary the number of input samples with a ratio from 60% to 100%. Then, we add

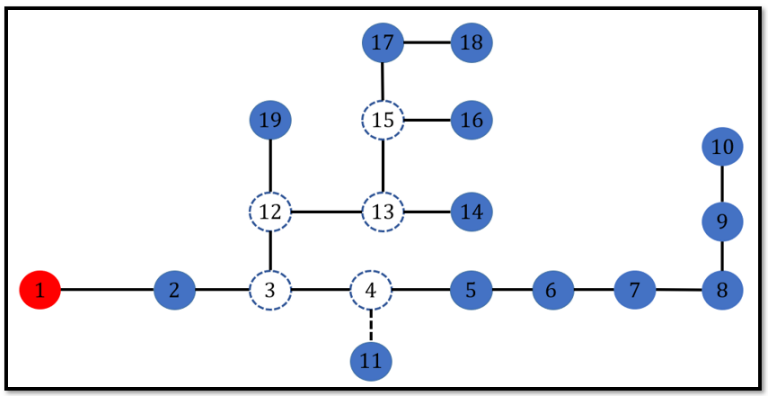


Figure 4: The topology of the 19-bus system. The red node, blue nodes, and white nodes are the slack bus, the observed nodes, and the hidden nodes, respectively.

noise with a standard signal-to-noise ratio (SNR) to be 125 [20] to test the robustness of the estimation algorithm.

After obtaining the measurements in $\mathcal{O}$, we can employ the OLS and RG algorithms to estimate the impedance and topology. To measure the goodness of estimation, we employ the normalized Total Vector Error ($nTVE\%$) [21] for impedance estimation. Specifically, we have:

$$nTVE(\mathbf{M}, \widehat{\mathbf{M}}) = 100 \times \frac{||\mathbf{M} - \widehat{\mathbf{M}}||_2}{||\mathbf{M}||_2}.$$

5.1. Impedance Estimation Result Display

In this section, we display the impedance estimation result over IEEE 8-, 19-, 33-bus, and 37-bus (three-phase unbalanced) systems. As an example, Fig. 4 is the topology of the IEEE 19-bus system, where the red node, blue nodes, and white nodes are the slack bus, the observed nodes, and the hidden nodes, respectively. Further, we change the percentage of the number of observed nodes in the 19-bus system to analyze the impact of system observability on the estimation error in the next subsection.

The results are shown in Fig. 5 to Fig. 8. We find that the estimation errors are stable given 864 samples (i.e., 9 days data) of voltage magnitude and power input or 10800 samples (i.e., 6 minutes data) of phasor input (we eliminate the presentation of high and unstable errors when the number of input

samples is small). Based on Theorem 2, this error is systematic and comes from the missing values of hidden nodes. For impedance estimation, Fig. 5 and Fig. 6 present the results using PG & E data with different resolutions by different interpolations, which simulates the scenarios of using phasor measurements and non-phasor measurements. Fig. 7 and Fig. 8 represent results using other two datasets, *Aust* with phasor measurements and *Ariz* with non-phasor measurements. Specifically, the average stable errors are 8%/10%/12% using $PG\&E, 1$ (phasor data), 8%/15%/15% using $PG\&E, 2$ (non-phasor data), 5%/5%/12% using *Aust* (phasor data) and 5%/10%/15% using *Ariz* (non-phasor data) for 8-, 19-, and 33-bus systems. Hence, we can find that with enough data, the general behavior with phasor measurements can outperform that with non-phasor measurements. The main reason is the linearization error of the LC-PF model. Namely, our proposed FOLS with phasor data is based on the accurate Ohm's law while the algorithm with non-phasor data can only employ the LC-PF that is a linearized approximation of the Ohm's law. The linearization error deteriorates the performance. However, it's noteworthy that the data needed from non-phasor measurements are less than the data needed from phasor-measurements to obtain a stable estimation. This is because the high resolution of micro-PMU measurements essentially brings a lot of similar measurements in the locally neighboring data streams. These measurements, though dense within a short time period, are not enough informative for parameter estimation. More precisely, compared to non-phasor measurements from sensors like smart meters, using limited data from micro-PMUs with high similarity can cause high correlations between observed nodes and hidden nodes, thus inducing (usually) higher and more unstable estimation errors based on Theorem 2. Even though, we find the required sampling time for the phasor data is much less than that of the non-phasor data, demonstrating the great potentials of information mining on the dense micro-PMU streams.

For unbalanced grids, we can only use phasor to estimate the impedance. As shown in Fig. 5 and Fig. 7, the stable estimation errors are 28% using $PG\&E, 1$ and 26% using *Aust*. They are much higher than the errors of balanced grids.

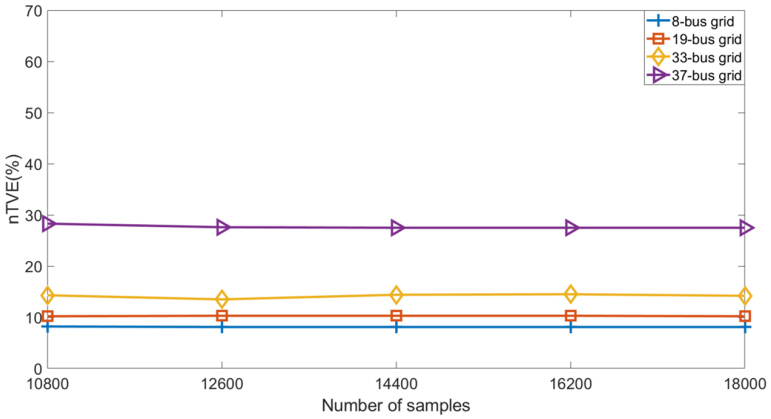


Figure 5: The impedance estimation results using data $PG\&E, 1$.

One reason is that the unbalanced grid has 37 buses and 20 of them are unobserved. This ratio is relatively larger than the balanced grid. Based on Theorem 2, the error is larger. Secondly, the linearization error of the three-phase power flow equations is very large, causing more extra errors for the estimation.

The above results can also verify the correctness of our Theorem 2 for the error upper bound. Specifically, we consider the case when we have enough input samples, i.e., 1296 samples from $PG\&E, 2$ and 16200 samples from $PG\&E, 1$. The results of the other two datasets are similar. Table 1 presents the values of the input's correlation coefficients, the number of hidden/observed nodes for the grid, the calculated error bound, and the estimated errors.

In conclusion, the error bound calculated based on Theorem 2 is larger than the true error during the estimation process for 8-, 19-, and 33-bus systems with both datasets. For the 37-bus unbalanced network, the true error exceeds the range of error bound. This is because for an unbalanced network, the linearization error for equation (3) is large and this part is not captured in Theorem 2. Finding a better-linearized equation and quantifying the linearization error for unbalanced networks are good topics for future research.

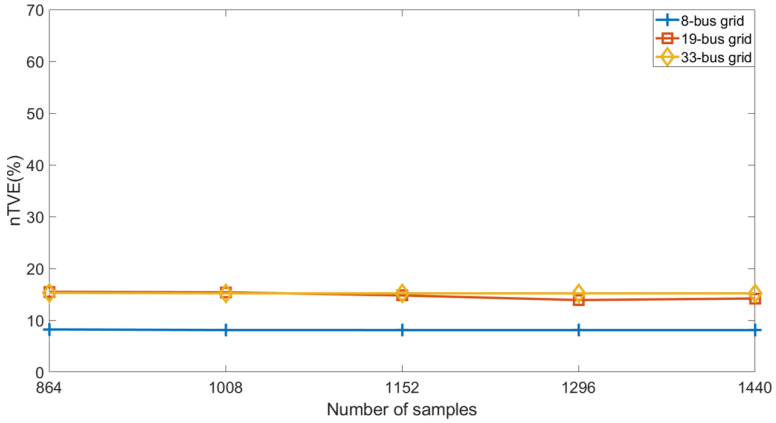


Figure 6: The impedance estimation results using data $PG\&E, 2$.

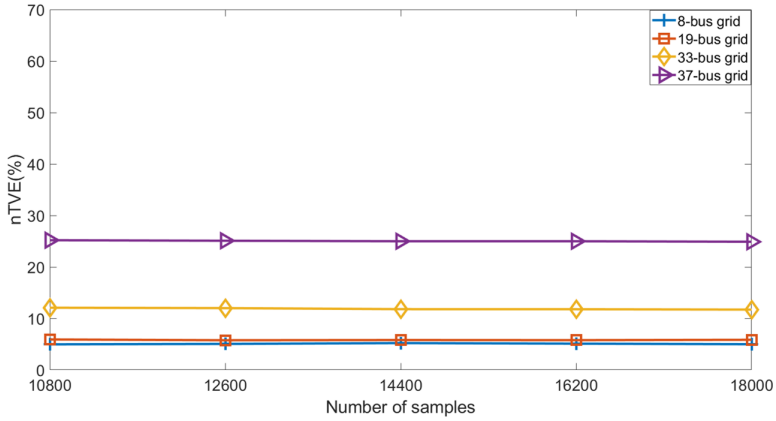


Figure 7: The impedance estimation results using data *Aust*.

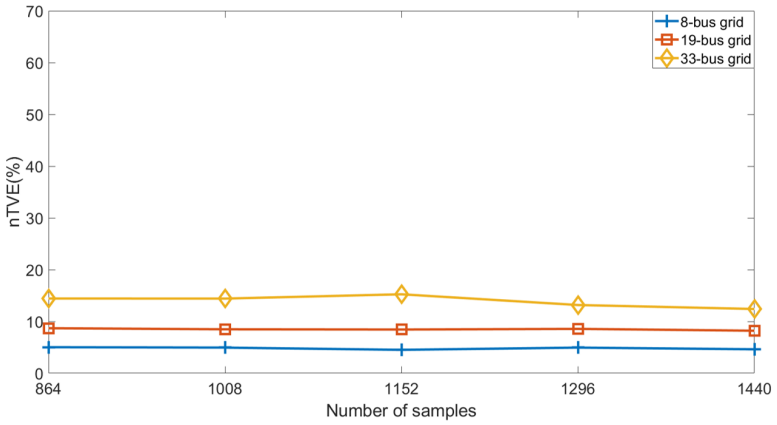


Figure 8: The impedance estimation results using data *Ariz*.

Table 1: The parameters to verify the relations of Theorem 2.

	Grid size	N_1	N_2	ρ_1	ρ_2	True error	Error bound
$PG\&E, 2$	8-bus	3	4	0.115	0.033	0.080	0.107
	19-bus	13	5	0.102	0.054	0.12	0.123
	33-bus	12	20	0.153	0.022	0.163	0.185
$PG\&E, 1$	8-bus	3	4	0.123	0.031	0.076	0.102
	19-bus	13	5	0.113	0.046	0.073	0.098
	33-bus	12	20	0.132	0.021	0.140	0.168
	37-bus	16	20	0.155	0.036	0.26	0.22

5.2. Sensitivity Analysis

To understand the impact of the system observable levels on the impedance estimation result, we conduct sensitivity analysis with respect to the percentage of observed nodes.

For the system observable levels, we consider smart meter data for 864 points, i.e., 9 days' data and also micro-PMU with 7200 points, i.e., 4 minutes' data. Then, we change the percentage of observed nodes in the 19-bus, shown in Fig. 4. The result of the estimation error with respect to the observed nodes' percentage is shown in Fig. 9. We find that the error is decreasing as we

increase the ratio of observed nodes, which corresponds to the analysis in our Theorem 2. Namely, the larger number of observed nodes will relax the upper bound required for the correlation between observed nodes and hidden nodes in Eqn. (10), thus yielding smaller errors for fixed measurement correlations. If the ratio of observed nodes is larger than 40%, the estimation error can be less than 20%. Then, micro-PMU's data perform better when the ratio is larger than 40%. This is because micro-PMU can provide phasor measurements that support the exact Ohm's law-based OLS, while smart meter can only provide voltage magnitude, active and reactive power data that bring linearization error when doing OLS.

However, it is noteworthy that the improvement of using phasor data is limited. We have the following reasons to explain this phenomenon. Firstly, our task is to conduct impedance estimation given unobserved nodes. Thus, the estimation error will always exist and only depends on the data correlations and the percentage of observed nodes, which is the core conclusion in our Theorem 2. Namely, given a limited number of observed nodes, even the micro-PMU data can't bring a perfect estimation result but only an approximated result with a fixed-range error. This error is systematic and can't be decreased to 0 unless the grid is completely observable or the unobservable quantities have 0 correlations to the observable measurements. Thus, the systematic error in Theorem 2 prevents the phasor data from far outperforming the magnitude-only data.

Secondly, in this subsection, we aim at demonstrating the sensitivity analysis with respect to the percentage of observed nodes rather than the comparison between phasor or non-phasor data. Thus, we employ 864 points (9 day's data) of magnitude-only measurements and 7200 points (4 minutes' data) of phasor measurements to test. Though phasor measurements have more samples, these samples have similar values within the 4 minutes and are therefore less informative than the 864 samples covering 9 days. Even though, they have a relatively better performance which shows the great potentials of micro-PMUs in parameter estimation.

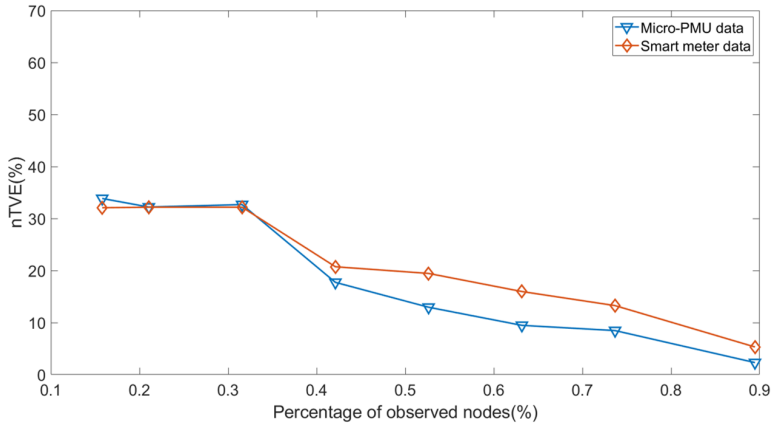


Figure 9: The performance of impedance estimation in the 19-bus system with respect to the ratio of observed nodes.

5.3. Topology Estimation Result Display

We utilize the hidden-node recovery rate ($NR\%$) and hidden-connection recovery rate ($CR\%$) to measure the performance of our algorithms on topology recovery. Specifically, we have:

$$\begin{aligned} NR &= 100 \times \frac{1}{T}\sum_{t=1}^{T}\frac{n_t}{n_0}, \\ CR &= 100 \times \frac{1}{T}\sum_{t=1}^{T}\frac{m_t}{m_0}, \end{aligned} \tag{15}$$

where T is the total number of trials due to different signal-to-noise ratios (SNRs) and input number of samples, n_t, n_0 are the number of recovered hidden nodes in the t^{th} trial and the true hidden nodes, respectively. Similarly, m_t, n_0 are the number of recovered hidden connections in the t^{th} trial and the true hidden connections, respectively. For calculations of m_t and n_t, we first count the detected hidden nodes/connections' numbers. Such a count will minus 1 if one false-positive scenario happens. Namely, if a hidden node/connection is identified but never exists in the true topology, we treat this case as mis-classification and the correct-classification number need to minus 1.

Based on the impedance estimation using smart meters with sample number larger than 864 or micro-PMU's data with sample number larger than 7200 (for

balanced grids) or 12600 (for unbalanced grids), we use the estimated impedance to do the topology learning and obtain the average results. The result of topology reconstruction is in Table 2.

Table 2: Average recovery rate result.

		8-bus	19-bus	33-bus	37-bus
$PG\&E, 1$	$NR(\%)$	100	100	100	87.23
	$CR(\%)$	100	100	100	84.32
$PG\&E, 2$	$NR(\%)$	98.07	95.37	90.11	
	$CR(\%)$	94.65	90.67	88.95	
$Aust$	$NR(\%)$	100	100	100	87.50
	$CR(\%)$	100	100	100	84.32
$Ariz$	$NR(\%)$	98.07	96.44	90.11	
	$CR(\%)$	94.65	92.33	88.95	

In the topology learning algorithm, i.e., Recursive Grouping (RG) algorithm, if the estimated impedance $\widehat{\mathbf{Z}}^S_{\mathcal{O}\mathcal{O}}$ is more accurate, the estimated graph $\mathcal{G}$ will be more accurate. On the other hand, the threshold in the RG algorithm increases the tolerance of the input impedance with some errors. In our Table 2, we show that the topology is perfectly recovered with phasor input ($PG\&E, 1$ and $Aust$) for balanced distribution grids, i.e., 100% for the $NR\%(CR\%)$ due to low impedance estimation errors. For the voltage magnitude ad power input ($PG\&E, 2$ and $Ariz$), the impedance estimation error is higher and we have around 2%(6%), 6%(10%), and 10%(12%) decreasing for the $NR\%(CR\%)$ accuracy. Finally, for three-phase unbalanced grid, the $NR\%(CR\%)$ accuracy decreases by 13%(16%) from 100%. This is because the estimated impedance has a larger error for the three-phase unbalanced grids due to the linearization errors of three-phase unbalanced power flow equations. Currently, such a recovery for unbalanced networks can only be implemented for phasor data in $PG\&E, 1$ and $Aust$ since equation (3) requires both voltage magnitude and angle values. For the 37-unbalanced network, the number of observed nodes is 16

and the number of hidden nodes is 20, and the slack bus is assumed to be known. In the simulation with $PG\&E, 1$, we utilize samples with total number to be $\{12600, 14400, 16200, 18000\}$ and under SNR to be $\{25, 50, 75, 100, 125, 150\}$. The detailed results of n_t (i.e., the correctly-identified number) for each scenario are listed in Table 3. Averagely, 2.5 hidden nodes are incorrectly recovered among the 20 hidden nodes.

Table 3: The detailed results of n_t for each scenario.

Sample number	25	50	75	100	125	150
12600	15	16	18	17	19	18
14400	16	16	16	18	18	19
16200	16	17	18	18	18	19
18000	17	18	18	18	18	18

5.4. Robustness Test

In this subsection, we test the robustness of our algorithm in face of short line segments that often appear in distribution grids. The small impedance of the short line will make the voltage difference between the terminal nodes small and even comparable to noise. The voltage drop along the short line segment, though small due to the small line impedance, is different from the noise since the voltage drop has some fixed patterns, where noise has a random pattern. Thus, our factorized ordinary least square (FOLS) estimation is still a good choice. As is claimed in [34], the OLS estimation is the best linear unbiased estimator without other information. To test the FOLS method's robustness to a small short line segment, we utilize 19-bus system as an example and lines $5-6$ and $7-8$ in this system have short line segments with small line resistance and reactance. The estimated impedance and the true impedance are shown in Fig. 10 using $PG\&E, 1$. We find that for both short line segments (lines $5-6$ and $7-8$) and the normal lines, the estimation errors are similar. Namely, our algorithm won't cause too much inaccuracy for the short line impedance

estimation.

Acknowledgement

The first two authors would like to acknowledge the support and work done by Salt River Project.

6. Conclusions

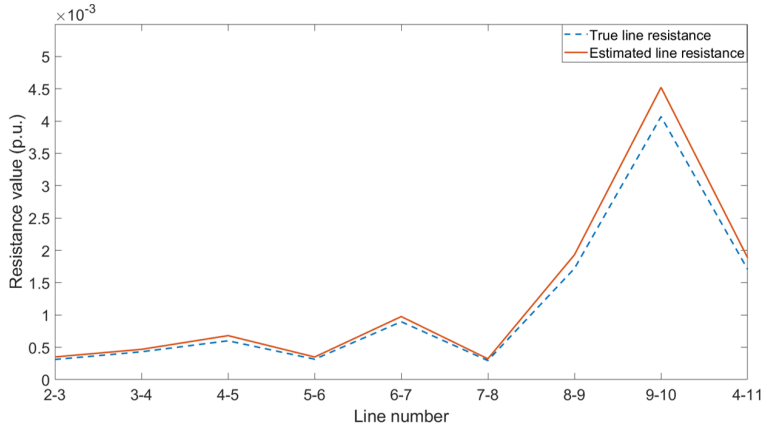
The distributed energy resources (DERs) are increasing in distribution grids, bringing high requirements for monitoring and control. To achieve these targets, the topology is needed. To find the topology and give an impedance estimation, this paper proposes a joint impedance estimation, topology identification, and error quantification procedure, where the input measurements can be phasor or magnitudes. In this procedure, the Ordinary Least Square (OLS) helps to estimate impedances of observed nodes. Inputting the estimated impedance to the Recursive Grouping algorithm can subsequently identify the topology and the impedance with respect to the hidden nodes. Since the missing data of hidden nodes' data brings errors, this paper quantifies the error via proposing the Factorized Ordinary Least Square (FOLS). Finally, we show that the error is very small with balanced IEEE standard test systems. For unbalanced systems, the error is relatively large due to the use of the linearized three-phase power flow in FOLS. Finding a better linearization method to fit the three-phase data in the unbalanced network is a promising topic to further decrease the impedance estimation error.

7. Appendix

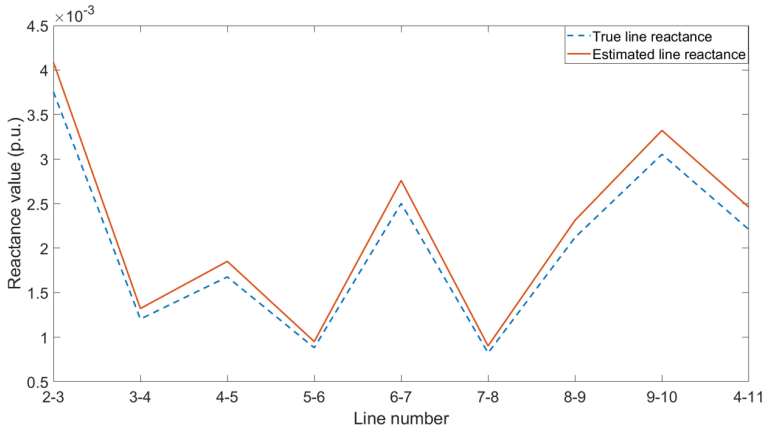
7.1. Proof of Theorem 1

Proof. For the whitening process, we have: $\mathbf{\Sigma} = \mathbf{M}\mathbf{M}^H$. Since the bus number of $\mathcal{O}$ ranges from k to L.

The observed covariance matrix calculated from $\mathbf{I}_{\mathcal{O}}$, termed as $\mathbf{\Sigma}^S_{\mathcal{O}\mathcal{O}}$, contains values $\Sigma_{kk}, \Sigma_{k,k+1}, \cdots, \Sigma_{LL}$ that can be directly calculated with the observed measurements. Therefore, $\mathbf{\Sigma}^S_{\mathcal{O}\mathcal{O}} = \mathbf{\Sigma}_{\mathcal{O}\mathcal{O}}$. $\mathbf{\Sigma}^S_{\mathcal{O}\mathcal{O}}$ is conjugate-symmetric



(a) The estimated resistance and the true resistance of different lines.



(b) The estimated reactance and the true reactance of different lines.

Figure 10: The estimated impedance and the true impedance of several lines for 19-bus system.

and *uniquely Cholesky-decomposable*. Let Cholesky decomposition of $\mathbf{\Sigma}^S_{\mathcal{OO}}$ be $\mathbf{\Sigma}^S_{\mathcal{OO}} = \mathbf{M}^S_{\mathcal{OO}}(\mathbf{M}^S_{\mathcal{OO}})^H$.

The upper-diagonal structure of $\mathbf{M}$ implies

$$\begin{bmatrix} \mathbf{M}_{\mathcal{HH}} & \mathbf{M}_{\mathcal{HO}} \\ & \mathbf{M}_{\mathcal{OO}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{M}^H_{\mathcal{HH}} & \\ \mathbf{M}^H_{\mathcal{HO}} & \mathbf{M}^H_{\mathcal{OO}} \end{bmatrix} = \begin{bmatrix} \mathbf{\Sigma}_{\mathcal{HH}} & \mathbf{\Sigma}_{\mathcal{HO}} \\ \mathbf{\Sigma}_{\mathcal{OH}} & \mathbf{\Sigma}_{\mathcal{OO}} \end{bmatrix}.$$

Thus, we have $\mathbf{M}^S_{\mathcal{OO}} = \mathbf{M}_{\mathcal{OO}}$ due to the uniqueness of the Cholesky decomposition of $\mathbf{\Sigma}^S_{\mathcal{OO}}$ (i.e., $\mathbf{\Sigma}_{\mathcal{OO}}$). Given that $\mathbf{W} = \mathbf{M}^{-1}$ is also an upper-diagonal matrix, we rewrite $\mathbf{MW} = \mathbf{Id}$ ($\mathbf{Id}$ is the identity matrix) in blocks form:

$$\begin{bmatrix} \mathbf{M}_{\mathcal{HH}} & \mathbf{M}_{\mathcal{HO}} \\ & \mathbf{M}_{\mathcal{OO}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{W}_{\mathcal{HH}} & \mathbf{W}_{\mathcal{HO}} \\ & \mathbf{W}_{\mathcal{OO}} \end{bmatrix} = \begin{bmatrix} \mathbf{Id}_{\mathcal{HH}} & \\ & \mathbf{Id}_{\mathcal{OO}} \end{bmatrix}.$$

Therefore, we have $\mathbf{W}_{\mathcal{OO}} = \mathbf{M}^{-1}_{\mathcal{OO}} = (\mathbf{M}^S_{\mathcal{OO}})^{-1} = \mathbf{W}^S_{\mathcal{OO}}$. Considering the whitening process:

$$\begin{bmatrix} \tilde{\mathbf{I}}_{\mathcal{H}} \\ \tilde{\mathbf{I}}_{\mathcal{O}} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{\mathcal{HH}} & \mathbf{W}_{\mathcal{HO}} \\ & \mathbf{W}_{\mathcal{OO}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{I}_{\mathcal{H}} \\ \mathbf{I}_{\mathcal{O}} \end{bmatrix},$$

We have $\tilde{\mathbf{I}}^S_{\mathcal{O}} = \mathbf{W}^S_{\mathcal{OO}}\mathbf{I}_{\mathcal{O}} = \mathbf{W}_{\mathcal{OO}}\mathbf{I}_{\mathcal{O}} = \tilde{\mathbf{I}}_{\mathcal{O}}$. ■

References

[1] Y. Weng, Y. Liao, R. Rajagopal, Distributed energy resources topology identification via graphical modeling, IEEE Transactions on Power Systems 32 (4) (2017) 2682–2694. doi:10.1109/TPWRS.2016.2628876.

[2] D. Lew, M. Asano, J. Boemer, C. Ching, U. Focken, R. Hydzik, M. Lange, A. Motley, The power of small: The effects of distributed energy resources on system reliability, IEEE Power and Energy Magazine 15 (6) (2017) 50–60. doi:10.1109/MPE.2017.2729104.

[3] P. D. F. Ferreira, P. M. S. Carvalho, L. A. F. M. Ferreira, M. D. Ilic, Distributed energy resources integration challenges in low-voltage networks: Voltage control limitations and risk of cascading, IEEE Transactions on Sustainable Energy 4 (1) (2013) 82–88. doi:10.1109/TSTE.2012.2201512.

[4] S. Bhela, V. Kekatos, S. Veeramachaneni, Enhancing observability in distribution grids using smart meter data, IEEE Transactions on Smart Grid 9 (6) (2018) 5953–5961. doi:10.1109/TSG.2017.2699939.

[5] C. Rudin, D. Waltz, R. N. Anderson, A. Boulanger, A. Salleb-Aouissi, M. Chow, H. Dutta, P. N. Gross, B. Huang, S. Ierome, D. F. Isaac, A. Kressner, R. J. Passonneau, A. Radeva, L. Wu, Machine learning for the new york city power grid, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (2) (2012) 328–345. doi:10.1109/TPAMI.2011.108.

[6] C. Rudin, S. Ertekin, R. Passonneau, A. Radeva, A. Tomar, B. Xie, S. Lewis, M. Riddle, D. Pangsrivinij, T. McCormick, Analytics for power grid distribution reliability in new york city, Interfaces 44 (2014) 364–383. doi:10.1287/inte.2014.0748.

[7] R. A. Jabr, Minimum loss operation of distribution networks with photovoltaic generation, IET Renewable Power Generation 8 (1) (2014) 33–44. doi:10.1049/iet-rpg.2012.0213.

[8] O. F. Fajardo, A. Vargas, Reconfiguration of mv distribution networks with multicost and multipoint alternative supply, part ii: Reconfiguration plan, IEEE Transactions on Power Systems 23 (3) (2008) 1401–1407. doi:10.1109/TPWRS.2008.926702.

[9] D. Deka, S. Backhaus, M. Chertkov, Estimating distribution grid topologies: A graphical learning based approach 23 (5) (2016) 1–7. doi:10.1109/PSCC.2016.7541005.

[10] G. Cavraro, V. Kekatos, S. Veeramachaneni, Voltage analytics for power distribution network topology verification, IEEE Transactions on Smart Grid 10 (1) (2019) 1058–1067. doi:10.1109/TSG.2017.2758600.

[11] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, L. Schenato, Identification of power distribution network topology via voltage correlation analysis 20 (15) (2013) 1659–1664. doi:10.1109/CDC.2013.6760120.

[12] Y. Liao, Y. Weng, G. Liu, R. Rajagopal, Urban mv and lv distribution grid topology estimation via group lasso, IEEE Transactions on Power Systems 34 (1) (2018) 12–27. doi:10.1109/TPWRS.2018.2868877.

[13] D. Deka, S. Backhaus, M. Chertkov, Structure learning and statistical estimation in distribution networks-part i, arXiv preprint arXiv:1501.04131.

[14] D. Deka, S. Backhaus, M. Chertkov, Structure learning and statistical estimation in distribution networks-part ii, arXiv preprint arXiv:1502.07820, 2015.

[15] M. J. Dolan, E. M. Davidson, I. Kockar, G. W. Ault, S. D. McArthur, Distribution power flow management utilizing an online optimal power flow technique, IEEE Transactions on Power Systems 27 (2) (2012) 790–799. doi:10.1109/TPWRS.2011.2177673.

[16] J. Yu, Y. Weng, R. Rajagopal, Patopa: A data-driven parameter and topology joint estimation framework in distribution grids, IEEE Transactions on Power Systems 33 (4) (2018) 4335–4347. doi:10.1109/TPWRS.2017.2778194.

[17] J. Yu, Y. Weng, R. Rajagopal, Patopaem: A data-driven parameter and topology joint estimation framework for time-varying system in distribution grids, IEEE Transactions on Power Systems 34 (3) (2019) 1682–1692. doi:10.1109/TPWRS.2018.2888619.

[18] G. Cavraro, V. Kekatos, Graph algorithms for topology identification using power grid probing, IEEE Control Systems Letters 2 (4) (2018) 689–694. doi:10.1109/LCSYS.2018.2846801.

[19] G. Cavraro, V. Kekatos, Inverter probing for power distribution network topology processing, IEEE Transactions on Control of Network Systems 6 (3) (2019) 980–992. doi:10.1109/TCNS.2019.2901714.

[20] Y. Yuan, S. Low, O. Ardakanian, C. Tomlin, Inverse power flow problem, arXiv preprint arXiv:1610.06631.

[21] K. Moffat, M. Bariya, A. Von Meier, Unsupervised impedance and topology estimation of distribution networks—limitations and tools, IEEE Transactions on Smart Grid 11 (1) (2020) 846–856. doi:10.1109/TSG.2019.2956706.

[22] S. Park, D. Deka, M. Chertkov, Exact topology and parameter estimation in distribution grids with minimal observability, arXiv preprint arXiv:1710.10727.

[23] Y. Wang, N. Zhang, H. Li, J. Yang, C. Kang, Linear three-phase power flow for unbalanced active distribution networks with pv nodes, CSEE Journal of Power and Energy Systems 3 (3) (2017) 321–324. doi:10.17775/CSEEJPES.2017.00240.

[24] H. Ahmadi, J. R. Martı, A. von Meier, A linear power flow formulation for three-phase distribution systems, IEEE Transactions on Power Systems 31 (6) (2016) 5012–5021. doi:10.1109/TPWRS.2016.2533540.

[25] K. J. G. Aligam, J. M. G. Dolot, E. S. Flores, J. R. C. Orillaza, Distribution system state estimator based on a linearized three-phase power flow 7 (3) (2018) 1–5. doi:10.1109/EEEIC.2018.8493653.

[26] A. Garces, A linear three-phase load flow for power distribution systems, IEEE Transactions on Power Systems 31 (1) (2016) 827–828. doi:10.1109/TPWRS.2015.2394296.

[27] R. B. Bapat, Graphs and matrices, Vol. 27, Springer, 2010.

[28] A. Kessy, A. Lewin, K. Strimmer, Optimal whitening and decorrelation, The American Statistician 72 (4) (2018) 309–314. doi:10.1080/00031305.2016.1277159.

[29] D. Deka, S. Backhaus, M. Chertkov, Learning topology of the power distribution grid with and without missing data, in: 2016 European Control Conference (ECC), 2016, pp. 313–320. doi:10.1109/ECC.2016.7810304.

[30] M. J. Choi, V. Y. F. Tan, A. Anandkumar, A. S. Willsky, Learning latent tree graphical models, J. Mach. Learn. Res. 12 (2011) 1771–1812. doi:10.5555/1953048.2021056.

[31] W. H. Kersting, Radial distribution test feeders, IEEE Transactions on Power Systems 6 (3) (1991) 975–985. doi:10.1109/PESW.2001.916993.

[32] Alfred Einfalt, Ea: Adres-conceptht tps://www.ea.tuwien.ac.at/projects/adres_concept/EN/.

[33] Lois Anne Leal, Numerical interpolation: Natural cubic splinehttps://towardsdatascience.com/numerical-interpolation-natural-cubic-spline-52c1157b98ac.

[34] S. Puntanen, G. P. Styan, The equality of the ordinary least squares estimator and the best linear unbiased estimator, The American Statistician 43 (3) (1989) 153–161. doi:10.1080/00031305.1989.10475644.