

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/306258254>

Learning Topology of Distribution Grids using only Terminal Node Measurements

Article · August 2016

CITATIONS

21

READS

144

3 authors:



Deepjyoti Deka

Los Alamos National Laboratory

111 PUBLICATIONS 790 CITATIONS

[SEE PROFILE](#)



Scott Backhaus

Los Alamos National Laboratory

136 PUBLICATIONS 5,101 CITATIONS

[SEE PROFILE](#)



Michael Chertkov

Los Alamos National Laboratory

291 PUBLICATIONS 7,037 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Optimal Topology Design for Power System Stability [View project](#)

Learning Topology of Distribution Grids using only Terminal Node Measurements

Deepjyoti Deka[†], Scott Backhaus[†], and Michael Chertkov[†]

[†]Los Alamos National Laboratory, USA

Email: deepjyoti@lanl.gov, backhaus@lanl.gov, chertkov@lanl.gov

Abstract—Distribution grids include medium and low voltage lines that are involved in the delivery of electricity from substation to end-users/loads. A distribution grid is operated in a radial/tree-like structure, determined by switching on or off lines from an underlying loopy graph. Due to the presence of limited real-time measurements, the critical problem of fast estimation of the radial grid structure is not straightforward. This paper presents a new learning algorithm that uses measurements only at the terminal or leaf nodes in the distribution grid to estimate its radial structure. The algorithm is based on results involving voltages of node triplets that arise due to the radial structure. The polynomial computational complexity of the algorithm is presented along with a detailed analysis of its working. The most significant contribution of the approach is that it is able to learn the structure in certain cases where available measurements are confined to only half of the nodes. This represents learning under minimum permissible observability. Performance of the proposed approach in learning structure is demonstrated by experiments on test radial distribution grids.

Index Terms—Distribution Networks, Power Flows, Tree learning, Voltage measurements, Missing data, Complexity

I. INTRODUCTION

The power grid is operationally divided hierarchically into transmission and distribution grids. While the transmission grid connects the generators and includes high voltage lines, the distribution grid comprises of medium and low voltage lines that connect the distribution substation to the end users/loads. Aside from low voltages, distribution grids are structurally distinguished from the transmission side by their operational radial structure. A distribution grid is operated as a tree with a substation placed at the root node/bus that is connected via intermediate nodes and lines to the terminal end nodes/households. This radial operational structure is derived from an underlying loopy network by switching on and off some of the lines (network edges) [1]. The specific structure may be changed from one radial configuration to another, by reversing the switch statuses. Fig. 1 presents an illustrative example of a radial distribution grid derived from an underlying graph.

Historically, the distribution grid has had limited presence of real time devices on the lines, buses and feeders [2]. Due to this, real-time monitoring of the operating radial structure and the states of the resident buses on the distribution side is not straightforward. These estimation problems, previously neglected, have become of critical importance due to introduction of new devices like batteries and electric

vehicles, and intermittent generation resources like rooftop solar panels. Optimal control in today's distribution grid requires fast topology and state estimation, often from limited real-time measurements. In this context, it needs to be mentioned that smart meters, micro-PMUs [3], frequency measurements devices (FNETs) [4] and advanced sensors (internet-of-things) capable of reporting real-time measurements are being deployed on the distribution side. However, in the current scenario, such devices are often limited to households/terminal nodes as their primary purpose for installation is services like price controllable demand and load monitoring. A majority of the intermediate lines and buses that connect the distribution substation to the terminal nodes do not have real-time measurement devices and are thus unobserved in terms of their structure and state. This hinders real-time topology and state estimation.

This paper is aimed at developing a learning framework that is able to overcome the lack of measurements at the intermediate nodes. Specifically, the primary goal of our work is to propose an algorithm to learn the operating radial topology using only real-time voltage magnitude measurements from the terminal nodes. The reliance only on measurements from terminal nodes is crucial as it makes our work applicable to realistic deployment of smart devices in today's distribution grids. Further, our learning algorithm is able to tolerate a much higher fraction of missing nodes (with unavailable data) compared to prior work in this area. Our approach is based on provable relations in *voltage magnitudes of triplets and pairs of terminal nodes* that are used to discover the operational edges iteratively. Computationally, the algorithm has polynomial complexity in the number of nodes in the system.

A. Prior Work

Topology learning in radial distribution grids has received attention in recent times. Learning techniques in this area vary, primarily based on the operating conditions and measurements available for estimation. The authors of [5] use a Markov random field model for nodal phases to identify faults in the grids. [6] uses conditional independence tests to identify the grid topology in radial grids. [7] uses signs of elements in the inverse covariance matrix of nodal voltages to learn the operational topology. Signature/envelope based identification of topology changes is proposed in [8]. Such comparison based schemes are used for parameter estimation

in [9], [10]. In contrast with nodal voltage measurements, line flow measurements are used in a maximum likelihood based scheme for topology estimation in [11]. In previous work [1], [12], authors have analyzed topology learning schemes that rely on trends in second moments of nodal voltage magnitudes. Further, a spanning tree based topology learning algorithm is proposed in [13]. This line of work [1], [12], [13] is close in spirit to machine learning schemes [14], [15] developed to learn the structure of general probabilistic graphical models [16]. A major limitation of the cited literature is that they assume measurement collection at most, if not all, nodes in the system. To the best of our knowledge, work that discuss learning in the presence of missing/unobserved nodes [12], [13] assume missing nodes to be separated by greater than two hops in the grid. As discussed earlier, distribution grids often have real-time meters only at terminal nodes and none at intermediate nodes that may be adjacent (one hop away).

B. Contribution of This Work

In this paper, we discuss topology learning in the radial distribution grid when measurements are limited to real-time voltage readings at the terminal nodes (end-users) alone. All intermediate nodes are unobserved and hence assumed to be missing nodes. We analyze voltages in the distribution grid using a linear lossless AC power flow model [1], [12] that is analogous to the popular [17], [18] LinDistFlow equations. For uncorrelated fluctuations of nodal power consumption, we construct functions of voltage magnitudes at a pair or triple of terminal nodes such that their values depend on the edges that connect the group of nodes. These functions provide the necessary tool to develop our learning scheme that iteratively identifies operational edges from the leaf onward to the substation root node. We discuss the computational complexity of the learning scheme and show that it is a third order polynomial in the number of nodes. In comparison to existing work, our approach is able to learn the topology and thereby estimate usage statistics in the presence of much greater fraction of missing/unobserved nodes. In fact, in limiting configurations, our learning algorithm is able to determine the true structure and statistics even when half of the nodes are unobserved/missing. We demonstrate the performance of our algorithm through experiments on test distribution grids.

The rest of the manuscript is organized as follows. Section II introduces notations used in the paper and describes the grid topology and power flow models used for analysis in later sections. Section III mentions assumptions made and describes important properties involving voltage measurements at terminal nodes. Our algorithm to learn the operating radial structure of the grid is discussed in Section IV with detailed examples. We detail the computational complexity of the algorithm in Section V. Simulation results of our learning algorithm on test radial networks are presented in Section VI. Finally, Section VII contains conclusions and discussion of future work.

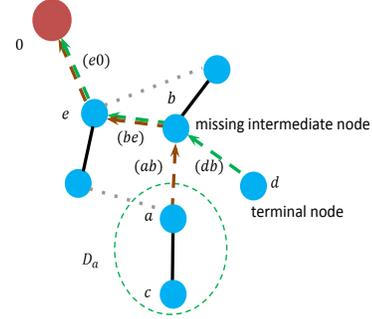


Fig. 1. Radial distribution grid tree with substation/root node colored in red. Dotted grey lines represent open switches. Terminal leaf nodes (d) represent end-users from which real-time measurements are collected. The intermediate missing nodes (b) are unobserved. Here, nodes a and c are descendants of node a . Dashed lines represent the paths from nodes a and d to the root node.

II. DISTRIBUTION GRID: STRUCTURE AND POWER FLOWS

Radial Structure: We denote the underlying graph of the distribution grid by the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the set of buses/nodes and \mathcal{E} is the set of all undirected lines/edges. We term nodes by alphabets (a, b, \dots). An edge between nodes a and b is denoted by (ab) . An illustrative example is given in Fig. 1. The operational grid consisting of one tree \mathcal{T} with nodes $\mathcal{V}^{\mathcal{T}}$ and operational edge set $\mathcal{E}^{\mathcal{T}} \subset \mathcal{E}$ as shown in Fig. 1. Our results are immediately extended to the case $K > 1$. In tree \mathcal{T} , \mathcal{P}_a denote the set of edges in the unique path from node a to the root node (reference bus). A node c is called a *descendant* of node a if the path from node c to the root passes through a , i.e., $\mathcal{P}_c \subset \mathcal{P}_a$. We use \mathcal{D}_a to denote the set of descendants of a and include node a in \mathcal{D}_a by definition. If edge $(ac) \in \mathcal{E}^{\mathcal{T}}$ and c is a descendant of a , we term a as *parent* and c as its *child node*. Further, as discussed in the Introduction, *terminal nodes/leaf nodes* in the tree represent end-users or households that are assumed to be equipped with real-time nodal meters. The remaining intermediate nodes (not terminal nodes) are not observed and hence termed as *missing nodes*. These definitions are illustrated in Fig 1. Next we describe the notation used in the power flow equations.

Power Flow Models: We use $z_{ab} = r_{ab} + ix_{ab}$ to denote the complex impedances of line (ab) ($i^2 = -1$) where r_{ab} and x_{ab} represent the line resistance and reactance respectively. Let real valued scalars, v_a , θ_a , p_a and q_a denote the voltage magnitude, voltage phase, active and reactive power injection respectively at node a . At each node a , Kirchoff's law of power flow relate them as follows:

$$P_a = p_a + iq_a = \sum_{b:(ab) \in \mathcal{E}^{\mathcal{T}}} \frac{v_a^2 - v_a v_b \exp(i\theta_a - i\theta_b)}{z_{ab}^*} \quad (1)$$

Note that Eq. (1) is nonlinear and non-convex. Under realistic assumption that losses of both active and reactive power losses on each line of tree \mathcal{T} is small, we can neglect second order terms in Eq. (1) to achieve the following linearized form [1], [7], [12]:

$$p_a = \sum_{b:(ab) \in \mathcal{E}^{\mathcal{T}}} (\beta_{ab}(\theta_a - \theta_b) + g_{ab}(v_a - v_b)), \quad (2)$$

$$q_a = \sum_{b:(ab) \in \mathcal{E}^{\mathcal{T}}} (-g_{ab}(\theta_a - \theta_b) + \beta_{ab}(v_a - v_b)) \quad (3)$$

$$\text{where } g_{ab} \doteq \frac{r_{ab}}{x_{ab}^2 + r_{ab}^2}, \beta_{ab} \doteq \frac{x_{ab}}{x_{ab}^2 + r_{ab}^2} \quad (4)$$

As shown in [1], Eqs. (2),(3) are equivalent to the LinDistFlow equations for power flow in distribution grid [17]–[19], when deviations in voltage magnitude are assumed to be small. Similar to LinDistFlow model, Eqs. (2),(3) are lossless with sum power equal to zero ($\sum_{a \in \mathcal{V}^{\mathcal{T}}} P_a = 0$). Further, note that both active and reactive power injections are functions of difference in voltage magnitudes and phases of neighboring nodes. Thus, the analysis of the system can be reduced by one node termed as reference node with voltage magnitude and phase at all other nodes being measured relative to this node. In our case, we take the substation/root node as the reference node with voltage magnitude 1 and phase 0 respectively. The reference node's injection also balances the power injections in the remaining network. Inverting Eqs. (2),(3) for the reduced system (without the reference node), we express voltages as a function of nodal power injections in the following vector form:

$$v = H_{1/r}^{-1}p + H_{1/x}^{-1}q \quad \theta = H_{1/x}^{-1}p - H_{1/r}^{-1}q \quad (5)$$

We term this as the **Linear Coupled Power Flow (LC-PF) model** where p , q , v and θ are the vectors of real power, reactive power injections, voltage magnitudes and phase angles respectively at the non-substation nodes. $H_{1/r}$ and $H_{1/x}$ are the reduced weighted Laplacian matrices for tree \mathcal{T} where reciprocal of resistances and reactances are used respectively as edge weights. The reduction is achieved by removing the row and column corresponding to the reference bus in the original weighted Laplacian matrix. We denote the mean of a random vector X by $\mu_X = \mathbb{E}[X]$. For two random vectors X and Y , the covariance matrix is denoted by $\Omega_{XY} = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)^T]$. Using Eq. (5), we can relate the means and covariances of voltage magnitudes with those of active and reactive injection as follows:

$$\mu_v = H_{1/r}^{-1}\mu_p + H_{1/x}^{-1}\mu_q \quad (6)$$

$$\begin{aligned} \Omega_v = & H_{1/r}^{-1}\Omega_p H_{1/r}^{-1} + H_{1/x}^{-1}\Omega_q H_{1/x}^{-1} + H_{1/r}^{-1}\Omega_{pq} H_{1/x}^{-1} \\ & + H_{1/x}^{-1}\Omega_{qp} H_{1/r}^{-1} \end{aligned} \quad (7)$$

Using these statistical quantities, we discuss useful identities that arise in radial distribution grids in the next section that form the basis of our learning algorithms.

III. PROPERTIES OF VOLTAGE MAGNITUDES IN RADIAL GRIDS

First, we make the following assumption regarding statistics of power injections at the non-substation grid nodes, under which our results hold.

Assumption 1: Fluctuations of active and reactive powers at different nodes are uncorrelated. Thus, $\forall a \neq b$ non-substation nodes, $\Omega_p(a, b) = \Omega_q(a, b) = \Omega_{qp}(a, b) = 0$.

As considered in prior literature [7], [12], this assumption is reasonable over short time-intervals where fluctuations in nodal power usage at households/terminal nodes are independent and hence uncorrelated. Intermediate nodes that do not represent end-users are uncorrelated if they have independent usage patterns. Specifically, for intermediate nodes involved in separation of power into downstream lines and without any major nodal usage, the net power injection is contributed by leakage or device losses and hence uncorrelated from the rest. Note that Assumption 1 does not specify the class of distributions that can model individual node's power injection. It is applicable when nodal injections are negative (loads), positive (due to local generation) or a mixture of both. In future work, we will relax this assumption and discuss learning in the presence of positively correlated end-user injection profiles.

Next, we mention an analytical statement relating the inverse of the reduced Laplacian matrix for a radial graph that we use in our later results.

Lemma 1. [1], [20] *The reduced weighted Laplacian matrix $H_{1/r}$ for tree \mathcal{T} satisfies*

$$H_{1/r}^{-1}(a, b) = \sum_{(cd) \in \mathcal{P}_a \cap \mathcal{P}_b} r_{cd} \quad (8)$$

In other words, the $(a, b)^{th}$ entry in $H_{1/r}^{-1}$ is equal to the sum of line resistances of edges common to paths from node a and b to the root. For example, in Fig. 1, $H_{1/r}^{-1}(a, d) = r_{be} + r_{e0}$. Using Eq. (8) it follows immediately that if node b is the parent of node a , then $\forall c$

$$H_{1/r}^{-1}(a, c) - H_{1/r}^{-1}(b, c) = \begin{cases} r_{ab} & \text{if node } c \in \mathcal{D}_a \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

Next, consider the function ϕ defined over two nodes a and b as $\phi_{ab} = \mathbb{E}[(v_a - \mu_{v_a}) - (v_b - \mu_{v_b})]^2$. ϕ_{ab} represents the variance of the difference in voltage magnitudes at nodes a and b . Using Eq. (7) we can write ϕ_{ab} in terms of power injection statistics in tree \mathcal{T} as follows

$$\begin{aligned} \phi_{ab} = & \Omega_v(a, a) - 2\Omega_v(a, b) + \Omega_v(b, b) \quad (10) \\ = & \sum_{d \in \mathcal{T}} (H_{1/r}^{-1}(a, d) - H_{1/r}^{-1}(b, d))^2 \Omega_p(d, d) \\ & + (H_{1/x}^{-1}(a, d) - H_{1/x}^{-1}(b, d))^2 \Omega_q(d, d) \\ & + 2 \left(H_{1/r}^{-1}(a, d) - H_{1/r}^{-1}(b, d) \right) \left(H_{1/x}^{-1}(a, d) - H_{1/x}^{-1}(b, d) \right) \Omega_{pq}(d, d) \end{aligned} \quad (11)$$

Note that Lemma 1 and Eq. (9) can be inserted in Eq. (11) to simplify it. In fact, doing so lets us derive properties of ϕ for terminal nodes in tree \mathcal{T} as discussed next.

Theorem 1. *Let node b be the parent of nodes a and c in \mathcal{T} such that (ab) and (bc) are operational edges (see Fig. 2(a)). Then*

$$\begin{aligned} \phi_{ac} = & \sum_{d \in \mathcal{D}_a} r_{ab}^2 \Omega_p(d, d) + x_{ab}^2 \Omega_q(d, d) + 2r_{ab}x_{ab} \Omega_{pq}(d, d) \\ & + \sum_{d \in \mathcal{D}_c} r_{bc}^2 \Omega_p(d, d) + x_{bc}^2 \Omega_q(d, d) + 2r_{bc}x_{bc} \Omega_{pq}(d, d) \quad (12) \\ = & \phi_{ab} + \phi_{bc} \quad (13) \end{aligned}$$

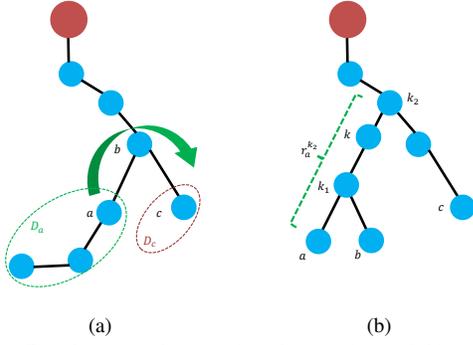


Fig. 2. (a) Distribution grid tree with nodes a and c as children of common parent node b . (b) Terminal nodes a and b have common parent k_1 . k_1 and terminal node c are descendants of k_2 . $r_a^{k_2}$ is the sum of resistance on lines (ak_1) , (k_1k) and (kk_2) connecting nodes a and k_1 .

Proof: Observe Lemma 1. As (ab) and (bc) are operational edges, the only nodes d such that $(H_{1/r}^{-1}(a,d) - H_{1/r}^{-1}(c,d)) \neq 0$ are either descendants of a (set \mathcal{D}_a) or of c (set \mathcal{D}_c). Further \mathcal{D}_c and \mathcal{D}_a are disjoint. When $d \in \mathcal{D}_a$, $(H_{1/r}^{-1}(a,d) - H_{1/r}^{-1}(c,d)) = r_{ab}$, while when $d \in \mathcal{D}_c$, $(H_{1/r}^{-1}(a,d) - H_{1/r}^{-1}(c,d)) = -r_{bc}$. Using this in the formula for ϕ_{ac} in Eq. (11) gives us the relation. The equality $\phi_{ac} = \phi_{ab} + \phi_{bc}$ is verified by plugging values in Eq. (11) for ϕ_{ab} , ϕ_{bc} and ϕ_{ac} . ■

A few points are in order. First, note that the only operational lines whose impedances appear on the right side of Eq. (12) are (ab) and (bc) . For the special case where a and c are terminal nodes with parent b , the relation reduces to the following:

$$\begin{aligned} \phi_{ac} = & r_{ab}^2 \Omega_p(a,a) + x_{ab}^2 \Omega_q(a,a) + 2r_{ab} x_{ab} \Omega_{pq}(a,a) \\ & + r_{bc}^2 \Omega_p(c,c) + x_{bc}^2 \Omega_q(c,c) + 2r_{bc} x_{bc} \Omega_{pq}(c,c) \end{aligned} \quad (14)$$

If the covariance of injections at a and b are known, the above condition can be checked for each potential parent node ‘ b ’ in linear time to identify the true parent. Second, the equality in Eq. (13) is true only when $\mathcal{P}_a \cap \mathcal{P}_c = \mathcal{P}_b$. It is replaced by a strict inequality for other configurations that we omit discussing as they are outside the scope of our learning conditions. The next theorem gives a result that relates the voltages at three terminal nodes.

Theorem 2. *Let terminal nodes a and b have common parent node k_1 . Let c be another terminal node such that $c, k_1 \in \mathcal{D}_{k_2}$ and $\mathcal{P}_{k_1} \cap \mathcal{P}_c = \mathcal{P}_{k_2}$ for some intermediate node k_2 (see Fig. 2(b)). Let $r_a^{k_2}$ and $x_a^{k_2}$ denote the sum of resistance and reactance respectively on lines on the path from node a to node k_2 , i.e., $r_a^{k_2} = \sum_{(ef) \in \mathcal{P}_a - \mathcal{P}_{k_2}} r_{ef}$, $x_a^{k_2} = \sum_{(ef) \in \mathcal{P}_a - \mathcal{P}_{k_2}} x_{ef}$. Define $r_b^{k_2}$, $r_{k_1}^{k_2}$ etc. in the same way. Then*

$$\begin{aligned} \phi_{ac} - \phi_{bc} = & \Omega_p(a,a)((r_a^{k_2})^2 - (r_{k_1}^{k_2})^2) + \Omega_q(a,a)((x_a^{k_2})^2 - (x_{k_1}^{k_2})^2) \\ & + \Omega_{pq}(a,a)(r_a^{k_2} x_a^{k_2} - r_{k_1}^{k_2} x_{k_1}^{k_2}) - \Omega_p(b,b)((r_b^{k_2})^2 - (r_{k_1}^{k_2})^2) \\ & + \Omega_q(b,b)((x_b^{k_2})^2 - (x_{k_1}^{k_2})^2) + \Omega_{pq}(b,b)(r_b^{k_2} x_b^{k_2} - r_{k_1}^{k_2} x_{k_1}^{k_2}) \end{aligned} \quad (15)$$

Proof: As a and b are terminal nodes with same parent k_1 , for each node $d \neq a \neq b$, the paths from d , a and b to

the root follow $\mathcal{P}_a \cap \mathcal{P}_d = \mathcal{P}_b \cap \mathcal{P}_d$. Using Lemma 1, we thus have $\forall d \neq a, b$, $H_{1/r}^{-1}(a,d) = H_{1/r}^{-1}(b,d)$ and

$$H_{1/r}^{-1}(a,d) - H_{1/r}^{-1}(c,d) = H_{1/r}^{-1}(b,d) - H_{1/r}^{-1}(c,d) \quad (16)$$

$$H_{1/r}^{-1}(a,a) - H_{1/r}^{-1}(a,b) = r_{ak_1}, H_{1/r}^{-1}(b,b) - H_{1/r}^{-1}(b,a) = r_{bk_1} \quad (17)$$

As k_1 and c are descendants of node k_2 and $\mathcal{P}_{k_1} \cap \mathcal{P}_c = \mathcal{P}_{k_2}$,

$$H_{1/r}^{-1}(a,a) - H_{1/r}^{-1}(c,a) = r_a^{k_2}, H_{1/r}^{-1}(b,b) - H_{1/r}^{-1}(c,b) = r_b^{k_2} \quad (18)$$

Further, using Eqs. (17, 18), we get

$$H_{1/r}^{-1}(b,a) - H_{1/r}^{-1}(c,a) = H_{1/r}^{-1}(a,b) - H_{1/r}^{-1}(c,b) = r_{k_1}^{k_2} \quad (19)$$

where $r_a^{k_2}$, $r_{k_1}^{k_2}$ etc. are defined in the statement of the theorem. Similar relations can be written for $H_{1/x}^{-1}$ terms as well. We now expand ϕ_{ac} and ϕ_{bc} using Eq. (11). Using Eq. (16,18,19) in the expression for $\phi_{ac} - \phi_{bc}$ gives Eq. (15). ■

Observe that the lines whose impedances appear on the right side of Eq. (15) are (ak_1) , (bk_1) and the ones on the path from node k_1 to k_2 . If this path is known till the penultimate node k before k_2 (see Fig. 2(b)), then Eq. (15) can be used to learn edge (kk_2) through a linear search among candidate nodes for k_2 . In the next section, we discuss the use of this relation to learn the path from terminal pairs with common parent (here a and b) to the root iteratively. Further, note that the value of $\phi_{ac} - \phi_{bc}$ is independent of injections at intermediate nodes and at terminal node c as long as c is a descendant of k_2 . Thus, Eq. (15) only helps identify c ’s relative position in the graph with respect to k_2 . As shown in the next section, we are able to locate c ’s parent using a post-order node traversal [21] in the grid graph.

IV. ALGORITHM TO LEARN TOPOLOGY USING TERMINAL NODE MEASUREMENTS

Using the relations described in the previous section, we discuss our algorithm to learn the operational topology of the radial distribution grid. As mentioned in the Introduction, voltage measurements are available only at the terminal/leaf nodes. The observer also has access to power injection statistics (variances) at the terminal nodes. These statistics are either known from historical data or computed empirically from power injection measurements collected at the terminal nodes. All intermediate nodes are missing and their voltage measurements and injection statistics are not observed/available. Further, we assume that impedances of all lines (open or operational) in \mathcal{E} for the underlying loopy graph \mathcal{G} are available as input. The task of the learning algorithm is to identify the set $\mathcal{E}^{\mathcal{T}}$ of operational edges in the radial grid \mathcal{T} . The substation root node in tree \mathcal{T} is assumed to be connected to a single intermediate node. Otherwise each sub-tree connected to the substation node can be considered as a disjoint tree. First, we make the following restriction regarding the degree of missing intermediate nodes.

Assumption 2: All missing intermediate nodes are assumed to have a degree greater than 2.

This assumption is necessary as without it, the solution to the topology learning problem will not be unique. As

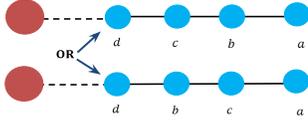


Fig. 3. Distribution grid tree where terminal node a is connected to node d via two unknown intermediate nodes of degree 2. Either configuration d, c, b, a or d, b, c, a for the intermediate nodes is a feasible structure given no measurements at nodes b and c .

an example, consider the test distribution grid given in Fig. 3 where the path from leaf node a to node d passes through two intermediate nodes b and c of degree 2 each. Both configuration A (operations edges $(ab), (bc), (cd)$) and configuration B (operations edges $(ac), (cb), (bd)$) are feasible operational topologies if voltage and power injection measurements are available only at the terminal node a . In other words, different values of injection covariances at nodes b and c can be taken such that relations between voltage and injections are satisfied in either configuration. Similar assumptions for uniqueness in learning general graphical models are mentioned in [22]. For radial configurations that respect Assumption 2, Algorithm 1 learns the topology using measurements of voltage magnitudes and information of injection statistics at terminal nodes.

Working: Algorithm 1 learns the structure of operational tree \mathcal{T} iteratively from leaf node pairs up to the root. Set \mathcal{L} represents the current set of terminal nodes/leaves with unknown parents, while set \mathcal{M} represents the set of all missing intermediate nodes. Note that for each node a , par_a denotes its identified parent while des_a denotes a set of two leaf nodes that are its descendants. Φ represents the empty set. The two sets (par and des) are used to keep track of the current stage of the learnt graph. There are three major stages in the edge learning process in Algorithm 1.

First, we build edges between leaf nodes pairs $a, c \in \mathcal{L}$ to their common missing parent node b in Steps 3 to 6. Here the relation for ϕ_{ab} derived in Theorem 1 and Eq. (14) is used to identify the true missing parent in set \mathcal{M} . Note that only parents with two or more leaf nodes as children can be identified at this stage. Parents that have at most one leaf node as child (other children being missing intermediate nodes) are not identified by this check.

Second, we remove nodes with discovered parents from the set of leaf nodes \mathcal{L} then iteratively learn the edges between intermediate nodes in Steps 10 to 24. At each stage of the iteration, \mathcal{M}_2 denotes the set of intermediate nodes with unknown parents whose descendants (des) were discovered in the previous iteration, while \mathcal{M}_1 denotes the set of all intermediate nodes with unknown parent (par) and known descendants. Clearly $\mathcal{M}_2 \subset \mathcal{M}_1$. The parent k_2 of each node k in \mathcal{M}_2 is identified by using relation Eq. (15) in Theorem 2 for its descendants a and b . There are two cases considered for candidate parent k_2 : (A) k_2 belongs to \mathcal{M}_1 and has known descendant c (Step 15), and (B) k_2 belongs to $\mathcal{M} - \mathcal{M}_1$ and has no discovered descendant (Step 18). For the second case, the algorithm tries to find some leaf node c that is potentially an undiscovered descendant of k_2 by

Algorithm 1 Topology Learning using Terminal Node Data

Input: Injection covariances $\Omega_p, \Omega_q, \Omega_{pq}$ at terminal nodes \mathcal{L} , Missing node set $\mathcal{M} = \mathcal{V}^{\mathcal{T}} - \mathcal{L}$, m voltage magnitude observations v for nodes in \mathcal{L} , set of all edges \mathcal{E} with line impedances.

Output: Operational Edge set $\mathcal{E}^{\mathcal{T}}$.

```

1:  $\forall$  nodes  $a, c \in \mathcal{L}$ , compute  $\phi_{ac} = \mathbb{E}[(v_a - \mu_{v_a}) - (v_c - \mu_{v_c})]^2$ 
2:  $\forall a \in \mathcal{V}^{\mathcal{T}}$ , define  $par_a \leftarrow \Phi, des_a \leftarrow \Phi$ 
3: for all  $a \in \mathcal{L}$  do
4:   if  $par_a = \Phi \& \exists c \in \mathcal{L}, b \in \mathcal{M}$  s.t.  $\phi_{ab}, c$  satisfy Eq. (14) then
5:      $\mathcal{E}^{\mathcal{T}} \leftarrow \mathcal{E}^{\mathcal{T}} \cup \{(ab), (bc)\}$ 
6:      $par_a \leftarrow b, par_c \leftarrow b, des_b \leftarrow a, c$ 
7:   end if
8: end for
9:  $\mathcal{L} \leftarrow \{a : a \in \mathcal{L}, par_a = \Phi\}, tp \leftarrow 1, \mathcal{M}_1 \leftarrow \Phi$ 
10: while  $tp > 0$  do
11:    $\mathcal{M}_2 \leftarrow \{k_1 : k_1 \in \mathcal{M} - \mathcal{M}_1, par_{k_1} = \Phi, des_{k_1} \neq \Phi\}$ 
12:    $\mathcal{M}_1 \leftarrow \{k_1 : k_1 \in \mathcal{M}, par_{k_1} = \Phi, des_{k_1} \neq \Phi\}$ 
13:   for all  $k \in \mathcal{M}_1$  with  $a, b \in des_k$  do
14:      $k_1 \leftarrow par_a$ 
15:     if  $\exists k_2 \in \mathcal{M}_1$  with  $\mathcal{M}_2 \cap \{k, k_2\} \neq \Phi$  with  $c \in des_{k_2}$ , s.t.
16:        $\phi_{ac} - \phi_{bc}$  satisfy Eq. (15) then
17:          $\mathcal{E}^{\mathcal{T}} \leftarrow \mathcal{E}^{\mathcal{T}} \cup \{(kk_2)\}, par_k \leftarrow k_2$ 
18:       else
19:         if  $k \in \mathcal{M}_1, \exists k_2 \in \mathcal{M} - \mathcal{M}_1, c \in \mathcal{L}$ , s.t.  $\phi_{ac} - \phi_{bc}$  satisfy
20:         Eq. (15) then
21:            $\mathcal{E}^{\mathcal{T}} \leftarrow \mathcal{E}^{\mathcal{T}} \cup \{(kk_2)\}, par_k \leftarrow k_2, des_{k_2} \leftarrow des_k$ 
22:         end if
23:       end if
24:     end for
25:    $tp \leftarrow |\{k_1 : k_1 \in \mathcal{M}_1, par_{k_1} \neq \Phi\}|$ 
26: end while
27: if then  $|\mathcal{M}_1| = 1$ 
28:   Join  $k \in \mathcal{M}_1$  to root
29: end if
30: Form a post-order traversal node set  $\mathcal{W}$  using  $par_a$  for
31:  $a : \forall a \in \mathcal{M}, des_a \neq \Phi$ 
32: for all  $c \in \mathcal{L}$  do
33:   for  $j \leftarrow 1$  to  $|\mathcal{W}|$  do
34:      $k_2 \leftarrow \mathcal{W}(j)$  with  $a, b \in des_{k_2}, k_1 \leftarrow par_a$ 
35:     if  $\phi_{ac} - \phi_{bc}$  satisfy Eq. (15) then
36:        $\mathcal{E}^{\mathcal{T}} \leftarrow \mathcal{E}^{\mathcal{T}} \cup \{(ck_2)\}, \mathcal{W} \leftarrow \mathcal{W} - \{k_2\}, j \leftarrow |\mathcal{W}|$ 
37:     end if
38:   end for
39: end for

```

checking for the relation $\phi_{ac} - \phi_{bc}$ specified in Eq. (15). If the relation holds, node k_2 is denoted as the parent of k_1 . As it is not clear if leaf c is an immediate descendant (child) of k_2 , no edge to c is added. The iterations culminate when no additional edge between intermediate nodes is discovered.

Third, in Steps 29 to 36, the parents of the remaining leaf nodes in \mathcal{L} are identified. Note that Eq. (15) also holds for non-immediate descendant c of k_2 , hence it cannot be used directly to identify c 's location. To overcome this, for each c in \mathcal{L} , we first check at descendants of k_2 before k_2 . This is ensured by creating a post-order traversal list \mathcal{W} [21] in Step 28. In post-order traversal \mathcal{W} , each potential parent is reached only after all its descendants have been reached. The node k_2 which satisfies Eq. (15) in Step 32 now gives the true parent of leaf c . The algorithm terminates after searching for all leaves in \mathcal{L} .

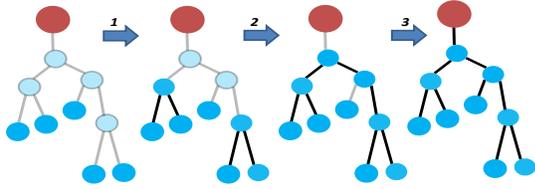


Fig. 4. Sequence in which edges are learnt by Algorithm 1 for an example radial grid using only measurements at leaf nodes. In the first step, leaf pairs with common parent are learnt, then intermediate nodes appear and finally connections to leaf nodes with no leaf sibling are discovered.

Note: Empirically computed second moments of voltages may differ from their true values and hence may not satisfy relations from Theorems 1 and 2 exactly. Hence we use tolerances τ_1 and τ_2 to check the correctness of Eq. (14) and Eq. (15) respectively in Algorithm 1. We defer the selection of optimal thresholds to future work.

We discuss the operation of Algorithm 1 through a test example given in Fig. 4. As described in the previous paragraph, in the first step, the common parents of leaf node pairs are discovered. In the next step, edges between intermediate nodes are discovered iteratively. Finally, positions of leaf nodes that do not have common parent with any other leaf node are located by a post-order traversal and then the substation connects to the node with no parent.

V. COMPUTATIONAL COMPLEXITY AND PERFORMANCE

Consider the original loopy graph \mathcal{G} to be complete (all node pairs belong to \mathcal{E}). We compute the computational complexity of operational tree in terms of the number of nodes $N = |\mathcal{V}^{\mathcal{J}}|$.

Algorithm Complexity: To measure the computational complexity, we focus on the iteration steps. Let the number of leaves in tree \mathcal{T} be l_0 . Detecting edges between leaf pairs and their common parent (Steps 3 to 6) takes $O(l_0^2(N - l_0))$ comparisons, which takes the form $O(N^3)$ in the worst case. Next we analyze the complexity of identifying edges between intermediate nodes. Let the number of ‘new’ intermediate nodes with unknown parents (set \mathcal{M}_2) in each iteration of (Steps 10 to 24) be l_i . Each ‘new’ node is compared with the set of all intermediate nodes with unknown parents ($\mathcal{M}_1 \supseteq \mathcal{M}_2$) first. As addition of a ‘new’ node leads to removal of its children (≥ 1) from \mathcal{M}_1 , the size of \mathcal{M}_1 never increases more than its initial value ($\leq l_0/2$) giving this step a worst case complexity of $O(l_i l_0)$. Comparison of nodes in \mathcal{M}_i with missing nodes in $\mathcal{M} - \mathcal{M}_1$ and leaf nodes in \mathcal{L} (Step 18) has worst case complexity $O(l_i l_0(N - l_0))$. Using l_0 as $O(N)$ and $\sum l_i = O(N)$, the computational complexity of all iterations in Steps 10 to 24 is thus $O(N^3)$. Finally composing the post-order tree traversal and checking at node location ($O(N)$) for each leaf ($O(N)$) has complexity less than or equal to $O(N^2)$. The overall complexity of the algorithm is thus $O(N^3)$.

Maximum Missing Node Fraction: Consider the two radial graphs given in Fig. 5 that satisfy the condition in Assumption 2. The first is a binary tree of depth d . All nodes till

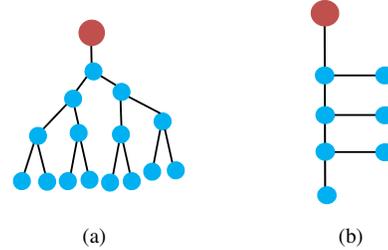


Fig. 5. Cases where fraction of missing nodes becomes 50% (a) A binary distribution grid tree with each intermediate node having degree 3. (b) Distribution grid tree with each intermediate node with one child

depth $d - 1$ are missing. The second graph has the structure of a line graph with one additional leaf node connected to each intermediate node on the line. In either graph, as the number of nodes increases, the fraction of missing nodes increases to 50%. In other words, *configurations exist such that Algorithm 1 is able to learn the grid structure with approximately half the nodes missing/unobserved.*

This is crucial because 50% fraction of missing nodes represents the threshold beyond which state estimation is not possible even in the presence of knowledge of the operational topology. This is because the LC-PF Eqs. (4) cannot be solved if less than half the nodes are observed.

VI. EXPERIMENTS

Here we demonstrate performance of Algorithm 1. We consider a radial network [23], [24] (Fig. 6) with 20 nodes and one substation. Out of those 20 nodes, 12 are terminal/leaf nodes while 8 are intermediate nodes which are missing. In each of our simulation runs, we construct complex power injection samples at the non-substation nodes from a multivariate Gaussian distribution that is uncorrelated between different nodes as per Assumption 1. Next we use LC-PF Eqs. (5) to generate nodal voltage magnitude measurements. To understand the performance of our algorithm, we introduce 30 additional edges (at random) to construct the loopy edge set \mathcal{E} . The additional edges are given impedances comparable to those of operational lines. The input to the observer consists of voltage magnitude measurements and injection statistics at the terminal nodes and impedances of all the lines within the loopy edge set \mathcal{E} . The average fraction errors (number of errors/size of the operational edge set) in reconstructing the grid structure is shown in Fig. 7(a) for different sizes of terminal voltage magnitude measurements used. Different curves in the plot depict different values of tolerances τ_1 and τ_2 to check the correctness of Eq. (14) and Eq. (15) in Algorithm 1. Note that the average fractional errors decreases steadily with increase in the number of measurements as empirical second moment statistics used in Algorithm 1 get more accurate. The values of tolerance to achieve the most accurate results are selected manually. We plan to develop a theoretical selection of the optimal tolerance values in future work.

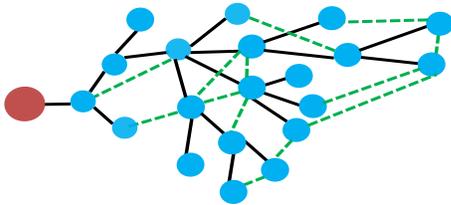


Fig. 6. Layouts of the grid tested with 20 non-substation nodes. Black lines represent operational edges. Some of the additional open lines (actual number 30) are represented by dotted green lines.

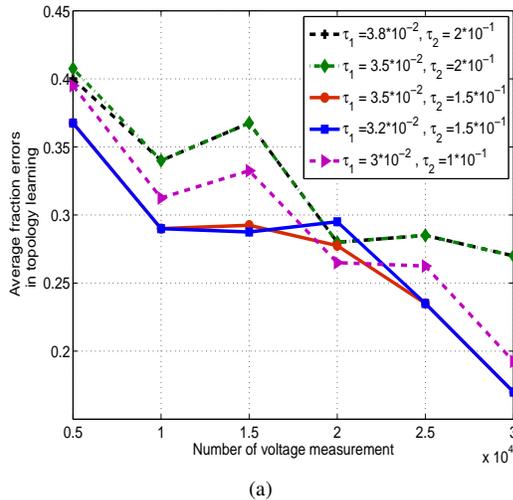


Fig. 7. Average fractional errors in learning operational edges vs number of samples used in Algorithm 1 for different values of tolerances τ_1 and τ_2 .

VII. CONCLUSIONS

Topology learning is an important problem in distribution grids as it influences several other control and managements applications that form part of the smart grid. The learning problem is complicated by the low placement of real-time measurements in the distribution grid, primarily being placed at terminal nodes that represent end-users or households. In this paper, we study the problem of learning the operational radial structure from a dense underlying loopy grid graph in this realistic regime where only terminal node (end-user) voltage measurements and injection statistics are available as input and all other nodes are unobserved/missing. We use a linear-coupled power flow model and show that voltages terminal node pairs and triplets satisfy relations that depend on their individual injection statistics and impedances on the lines connecting them. We use these properties to propose a learning algorithm that iteratively learns the grid structure from the leaves onward towards the substation root. We show that the algorithm has polynomial time complexity comparable to existing work, despite being capable of tolerating much greater and realistic fraction of missing nodes. For specific cases, this algorithm is capable of learning the structure with 50% unobserved nodes, beyond which state estimation is not possible even in the presence of topology information. We demonstrate performance of the learning algorithm through experiments on distribution grid test case. Computing the

sample complexity of the learning algorithm and optimizing the selection of tolerance values based on the number of samples are directions of our future research in this area. Further, we plan to expand this learning regime to cases with correlated injections.

REFERENCES

- [1] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning and statistical estimation in distribution networks - part i," *arXiv preprint arXiv:1501.04131*, 2015.
- [2] R. Hoffman, "Practical state estimation for electric distribution networks," in *Power Systems Conference and Exposition, 2006. PSCE'06. 2006 IEEE PES*. IEEE, 2006, pp. 510–517.
- [3] A. von Meier, D. Culler, A. McEachern, and R. Arghandeh, "Micro-synchphasors for distribution systems," in *Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2014, pp. 1–5.
- [4] Z. Zhong, C. Xu, B. J. Billian, L. Zhang, S.-J. S. Tsai, R. W. Conners, V. A. Centeno, A. G. Phadke, and Y. Liu, "Power system frequency monitoring network (fnet) implementation," *Power Systems, IEEE Transactions on*, vol. 20, no. 4, pp. 1914–1921, 2005.
- [5] M. He and J. Zhang, "A dependency graph approach for fault detection and localization towards secure smart grid," *Smart Grid, IEEE Transactions on*, vol. 2, no. 2, pp. 342–351, 2011.
- [6] D. Deka, S. Backhaus, and M. Chertkov, "Estimating distribution grid topologies: A graphical learning based approach," in *Power Systems Computation Conference (accepted)*, 2016.
- [7] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE, 2013, pp. 1659–1664.
- [8] G. Cavraro, R. Arghandeh, A. von Meier, and K. Poolla, "Data-driven approach for distribution network topology detection," *arXiv preprint arXiv:1504.00724*, 2015.
- [9] J. Peppanen, J. Grimaldo, M. J. Reno, S. Grijalva, and R. G. Harley, "Increasing distribution system model accuracy with extensive deployment of smart meters," in *PES General Meeting—Conference & Exposition, 2014 IEEE*. IEEE, 2014, pp. 1–5.
- [10] J. Peppanen, M. J. Reno, M. Thakkar, S. Grijalva, and R. G. Harley, "Leveraging ami data for distribution system model calibration and situational awareness," 2015.
- [11] R. Sevlian and R. Rajagopal, "Feeder topology identification," *arXiv preprint arXiv:1503.07224*, 2015.
- [12] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning and statistical estimation in distribution networks - part ii," *arXiv preprint arXiv:1502.07820*, 2015.
- [13] —, "Learning topology of the power distribution grid with and without missing data," in *European Control Conference (accepted)*, 2015.
- [14] P. Ravikumar, M. J. Wainwright, J. D. Lafferty *et al.*, "High-dimensional ising model selection using 1-regularized logistic regression," *The Annals of Statistics*, vol. 38, no. 3, pp. 1287–1319, 2010.
- [15] A. Anandkumar, V. Tan, and A. S. Willsky, "High-dimensional graphical model selection: tractable graph families and necessary conditions," in *Advances in Neural Information Processing Systems*, 2011, pp. 1863–1871.
- [16] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Foundations and Trends® in Machine Learning*, vol. 1, no. 1-2, pp. 1–305, 2008.
- [17] M. Baran and F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *Power Delivery, IEEE Transactions on*, vol. 4, no. 1, pp. 735–743, Jan 1989.
- [18] —, "Optimal capacitor placement on radial distribution systems," *Power Delivery, IEEE Transactions on*, vol. 4, no. 1, pp. 725–734, Jan 1989.
- [19] —, "Network reconfiguration in distribution systems for loss reduction and load balancing," *Power Delivery, IEEE Transactions on*, vol. 4, no. 2, pp. 1401–1407, Apr 1989.
- [20] J. Resh, "The inverse of a nonsingular submatrix of an incident matrix," *IEEE Transactions on Circuit Theory*, vol. 10, p. 131.
- [21] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. The MIT Press, 2001.
- [22] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 2014.

- [23] U. Eminoglu and M. H. Hocaoglu, "A new power flow method for radial distribution systems including voltage dependent load models," *Electric Power Systems Research*, vol. 76, no. 13, pp. 106 – 114, 2005.
- [24] "Reds: Repository of distribution systems," accessed: 2016-05-20. [Online]. Available: <http://www.dejazzer.com/reds.html>