Learning Distribution Grid Topologies: A Tutorial

Deepjyoti Deka®, Senior Member, IEEE, Vassilis Kekatos®, Senior Member, IEEE, and Guido Cavraro®, Member, IEEE

Abstract—Unveiling feeder topologies from data is of paramount importance to advance situational awareness and proper utilization of smart resources in power distribution grids. This tutorial summarizes, contrasts, and establishes useful links between recent works on topology identification and detection schemes that have been proposed for power distribution grids. The primary focus is to highlight methods that overcome the limited availability of measurement devices in distribution grids, while enhancing topology estimates using conservation laws of power-flow physics and structural properties of feeders. Grid data from phasor measurement units or smart meters can be collected either passively in the traditional way, or actively, upon actuating grid resources and measuring the feeder’s voltage response. Analytical claims on feeder identifiability and detectability are reviewed under disparate meter placement scenarios. Such topology learning claims can be attained exactly or approximately so via algorithmic solutions with various levels of computational complexity, ranging from least-squares fits to convex optimization problems, and from polynomial-time searches over graphs to mixed-integer programs. Although the emphasis is on radial single-phase feeders, extensions to meshed and/or multiphase circuits are sometimes possible and discussed. This tutorial aspires to provide researchers and engineers with knowledge of the current state-of-the-art in tractable distribution grid learning and insights into future directions of work.

Index Terms—Smart inverters, smart meter data, radial graph, recursive grouping, active sensing, voltage covariances, graph Laplacian matrix, linear distribution flow model.

I. INTRODUCTION

DISTRIBUTION grids constitute the final tier in the delivery of electricity to end-users. To ease protection and voltage control, most distribution grids are operated in a radial (tree-like) topology, which can be modified by changing breaker statuses on available lines [54]. In recent years, the growth of behind-the-meter distributed energy resources (DERs) and smart loads (e.g., air-conditioners, storage devices, electric vehicles) have brought distribution grids to the forefront of smart grid advancement [85]. Industrial and academic research on smart distribution grids has advocated the participation of distribution grid resources in wholesale electricity markets and ancillary services (such as demand response, frequency regulation, and transactive energy services [62]). Integrating renewables introduces new challenges for voltage regulation and calls for dispatching DERs without violating physical and operational grid ratings. This necessitates knowing the correct feeder models. Moreover, situational awareness requires accurate distribution grid state estimation (DSSE) [26], in which the operational topology is a critical component. Topology estimates are also important for ensuring the dynamic stability of inverter-interfaced DERs.

However, distribution utilities often have only partial knowledge of their primary and/or secondary networks and the associated line impedances. Similarly, even if the utility knows the line infrastructure and line impedances, it may not have information on which lines are currently energized. This is owing to the fact that distribution grids are frequently reconfigured for maintenance; load balancing; to improve voltage profiles, minimize losses, or alleviate faults; or rashly, while restoring service after extreme weather events. Such changes may not be logged into the distribution management system, and hence, need to be estimated or at least verified. In this context, grid topology learning can be broadly classified into topology detection and topology identification. In topology detection, the estimator or system operator knows the line infrastructure and their impedances and needs to determine the ones that are currently energized. Topology identification, on the other hand, aims to determine both the connections and line impedances. Without knowledge of line infrastructure, identification is a more challenging task.

Reliable topology estimation and DSSE in general are hindered by the limited placement of real-time metering devices on low-voltage grids. This is in stark contrast to high-voltage transmission systems that usually enjoy full observability and breaker statuses on lines are reported in real-time or identified jointly while estimating system states via the so-termed generalized power system state estimators; see [52], [46, Ch. 4]. On the other hand, smart meters and real-time sensors as part of the advanced metering infrastructure (AMI) have seen increased adoption at customer/load locations. In 2020, about 88% of the 102.9 million AMI installations in the United States have been at residential customers [92]. Such residential data at feeder buses, though non-ubiquitous, can be used for topology learning and other distribution grid tasks [7], [68].

Manuscript received 22 June 2022; revised 10 January 2023 and 2 March 2023; accepted 26 April 2023. Date of publication 1 May 2023; date of current version 26 December 2023. This work was supported in part by the National Science Foundation under Grant 1751085 and Grant 2034137; by the U.S. Department of Energy (DOE) under Contract DE-AC36-LLNL-88GO28308. The work of Deepjyoti Deka was supported by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract DE-AC36-98GO28308. The work of Deepjyoti Deka was supported by the National Nuclear Security Administration of the U.S. Department of Energy under Contract 89233218CNA000001. Paper no. TSG-00892-2022. The work of Deepjyoti Deka was supported by the National Nuclear Security Administration of the U.S. Department of Energy under Contract 89233218CNA000001. Paper no. TSG-00892-2022. (All the authors contributed equally to this work.) (Corresponding author: Vassilis Kekatos.) Deepjyoti Deka is with the Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545 USA (e-mail: deepjyoti@lanl.gov). Vassilis Kekatos is with the Bradley Department of ECE, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: kekatos@vt.edu). Guido Cavraro is with the National Renewable Energy Laboratory, Golden, CO 80401 USA (e-mail: guido.cavraro@nrel.gov).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2023.3271902.
Digital Object Identifier 10.1109/TSG.2023.3271902.
This tutorial systematically reviews approaches for topology learning in power distribution grids. Emphasizing on topology identification and topology detection, approaches are organized based on the type of measurements needed (namely smart meter data and synchrophasor data), and depending on whether injection and/or voltage readings are collected at all or a subset of buses. The focus is on topology estimates that are provably correct given a sufficient number of measurements. In the realistic setting of partial observability, several of the surveyed algorithms use constraints from the physics of power flows as well as topological properties emerging from the tree structure of feeders to ensure consistency of topology estimates. Depending on meter deployments, oftentimes complete topology recovery is not possible. Nonetheless, several of the presented schemes are able to estimate a reduced feeder model conveying some of the properties of the original grid, which may be sufficient in DER control applications. Identifiability and detectability claims are reviewed to explain when a topology can be unveiled given a particular meter placement. Such claims are important for understanding the limits in topology processing as well as informing future meter placement decisions.

Topology learning approaches are presented in the context of radial single-phase grids, although generalizations to multiphase and meshed settings are sometimes possible and discussed briefly in Section VI. The practical relevance of working with single-phase feeder models can be supported by two arguments: i) In countries where customers are connected to a three-phase main supply, distribution grids are relatively balanced, and working with the single-phase equivalent model may be a reasonable approximation; ii) In countries where residential customers are connected to a single-phase supply, the phase for each customer can be identified relatively easily by clustering voltage data [10], [12], [69], [94]. Subsequently, per-phase data can be processed assuming the approximate model of a single-phase feeder.

With advances in machine learning, there has been an influx of supervised topology learning schemes that train a classifier to detect one out of a finite number of topology configurations. Such classifiers are typically trained using labeled grid data (e.g., voltages) with labels being the associated topologies. Training data can be collected in the field, or synthetically generated using feeder models and simulated loads and renewables. Such approaches are the focus of a recent survey paper [77]. In contrast, our tutorial article focuses primarily on physics-aware and unsupervised (i.e., without labeled data) algorithmic solutions for topology learning.

Outline: The rest of this tutorial is organized as follows. Section II introduces some graph and feeder modeling preliminaries. Topology identification schemes relying on synchrophasor and non-synchrophasor data are presented in Sections III and IV, respectively. Topology detection schemes are reviewed in Section V. Extensions including generalizations to non-radial topologies, multi-phase grids, data from multiple feeders, and learning from dynamic grid data or non-electric grid data are discussed in Section VI. Section VII discusses research areas related to topology learning and possible directions of future work. The tutorial is concluded in Section VIII.

Notation: Lower- (upper-) case boldface letters denote column vectors (matrices). Calligraphic symbols are reserved for sets. Symbols $^\top$ and $^H$ stands for transposition and complex conjugate transpose. Vectors $\mathbf{0}$ and $\mathbf{1}$ are the all-zero and all-one vectors, while $\mathbf{I}_N$ denotes the $N\times N$ identity matrix. Symbol $\|\cdot\|$ denotes the Euclidean norm of the vector $\mathbf{x}$, and $\|\mathbf{A}\|_F$ denotes the Frobenius norm of matrix $\mathbf{A}$. Symbol $\mathcal{D}_\mathbf{x}$ denotes a diagonal matrix with the entries of vector $\mathbf{x}$ on its main diagonal. The expectation operator is defined as $\mathbb{E}[\cdot]$. Operator $\text{vec}(\mathbf{X})$ stacks the columns of matrix $\mathbf{X}$ into a vector.

II. Modeling Preliminaries

A. Preliminaries on Graphs

Because graph algorithms have been widely used in the literature to reconstruct grid topologies, this section reviews some basic concepts of graph theory; see Figure 1. Consider graph $\mathcal{G}_o = (\mathcal{N}, \mathcal{L})$. A rooted tree is a connected graph without loops with one node designated as the root and indexed by 0. In a tree graph, a path $\mathcal{P}_{n,m}$ is the unique sequence of edges connecting any two nodes $n$ and $m$. If node $n$ is along path $\mathcal{P}_{0,m}$, then node $m$ is a descendant of $n$, and $n$ is an ancestor of $m$. If there exists line $(m, n) \in \mathcal{L}$ and node $m$ is closer to the root, then node $m$ is the parent of $n$, and $n$ is the child of $m$. In radial grids, each node modulo the root has a unique parent. A node without children is called a leaf, while non-leaf nodes will be termed internal nodes.

B. Feeder Modeling

A single-phase power distribution grid having $N+1$ buses can be modeled by a graph $\mathcal{G}_o = (\mathcal{N}, \mathcal{L})$. The nodes in
Grid topology is captured by the branch-bus incidence matrix \( \mathbf{A} \in \{0, \pm 1\}^{L \times (N+1)} \), which can be partitioned into its first and the rest of its columns as \( \mathbf{A} = [a_0 \, \mathbf{A}] \). For a radial grid \( L = N \), the reduced incidence matrix \( \mathbf{A} \) is square and invertible [45]. Because \( a_0 = -\mathbf{A}^{-1} \).

Voltage and current phasors are related through the bus admittance matrix \( \mathbf{Y} := \mathbf{A}^T \mathbf{D}_y \mathbf{A} \), where \( \mathbf{D}_y \) is a diagonal matrix carrying the vector of line admittances \( \mathbf{y} = \mathbf{g} + j\mathbf{b} \) on its main diagonal; shunt admittances have been ignored. Upon partitioning to separate the substation, we get that

\[
\begin{bmatrix}
v_0 \\
v_0
\end{bmatrix} =
\begin{bmatrix}
Y_{00} & Y_{01} \\
Y_{01} & Y_{11}
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_0
\end{bmatrix} \quad \Rightarrow \quad i = \mathbf{Y}(u - u_0 \mathbf{1}) \tag{1}
\]

where \( \mathbf{Y} = \mathbf{A}^T \mathbf{D}_y \mathbf{A} \) is the reduced bus admittance matrix, and vectors \( (i, u) \) collect current and voltage phasors at all buses but the substation. The right-hand side (RHS) of (1) follows from \( \mathbf{Y}_0 = -\mathbf{Y}_1 \). We will oftentimes assume \( u_0 = 1 \) or that \( u_0 \) is known and can thus be subtracted from \( u \).

Note that ignoring the substation current in (1) does not incur any loss of information as \( i_0 = -1^T i \). Power injections depend non-linearly on voltages as \( \mathbf{s}_n = u_n \mathbf{v}_n^2 \). To arrive at simpler linear models, we usually linearize the nonlinear AC power flow equations. If linearization occurs at the so-called flat voltage profile where \( v_n = 1 \) and \( \theta_n = 0 \) for all \( n \), it can be shown that [14, 15, 30]

\[
\begin{align*}
p &\simeq \mathbf{A}^T \mathbf{D}_g \mathbf{A} (v - v_0 \mathbf{1}) + \mathbf{A}^T \mathbf{D}_b \mathbf{A} \theta \\
q &\simeq \mathbf{A}^T \mathbf{D}_b \mathbf{A} (v - v_0 \mathbf{1}) - \mathbf{A}^T \mathbf{D}_g \mathbf{A} \theta
\end{align*}
\tag{2a}
\]

where vectors \( (p, q, v, \theta) \) collect power injections, voltage magnitudes and angles at all buses excluding the substation.

The aforesaid model relates power injections linearly to voltages and holds for radial and meshed single-phase grids alike. The model can be inverted so that voltages are expressed as linear functions of power injections as

\[
\begin{align*}
v &\simeq \mathbf{R} p + \mathbf{X} q + v_0 \mathbf{1} \\
\theta &\simeq \mathbf{X} p - \mathbf{R} q
\end{align*}
\tag{3a}
\]

for appropriately defined matrices \( (\mathbf{R}, \mathbf{X}) \); see [15, 30]. The substation voltage is usually assumed to be \( v_0 = 1 \), or fixed at a known value.

For the special yet practically relevant case of radial single-phase grids, matrices \( (\mathbf{R}, \mathbf{X}) \) in (3) can be expressed as

\[
\begin{align*}
\mathbf{R} &= \mathbf{A}^{-1} \mathbf{D}_g \mathbf{A}^{-\top} \\
\mathbf{X} &= \mathbf{A}^{-1} \mathbf{D}_b \mathbf{A}^{-\top}
\end{align*}
\tag{4}
\]
TABLE I
KEY FEATURES OF THE PRESENTED FEEDER TOPOLOGY IDENTIFICATION AND VERIFICATION METHODS

| Data | Feeder Topology Identification | Feeder Topology Verification |
|------|--------------------------------|-----------------------------|
| Method | Type | Availability | Collection | Topology | Phases | Feeder | Recovery |
|       | AMI | PMU | Full | Partial | Passive | Active | Tree | Mesh | 1Φ | mΦ | One | Many | claims |
| [10], [32], [44], [59] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [8], [73], [83] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [13] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [22] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [23], [60] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [28], [29], [30], [31], [70], [71], [72] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [33], [34], [58], [60] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [57] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [55], [76], [99], [100], [101] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [5], [16], [96], [97] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [11] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [27], [61], [87], [95], [101] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [39], [48], [56] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [4], [35], [64] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [66] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [67], [98] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| [88] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

Also be derived for meshed single-phase [30], [24, Sec. IV A]; as well as radial multiphase feeders [51].

Having established commonly used feeder models, Sections III and IV deal with topology identification, Section V deals with topology detection, while extensions are discussed in Section VI. Table I summarizes the key features of the methods to be presented in this tutorial article.

III. TOPOLOGY IDENTIFICATION USING PHASOR DATA

Given voltage with/without injection data collected from a feeder, topology identification aims at finding the network connectivity with/without line impedances of the feeder. This section reviews methods relying on synchrophasor data, while Section IV reviews methods relying on smart meter data. For synchrophasor data, the data model of interest is \( \mathbf{i} = \mathbf{Yu} \) as the operator collects synchrophasor measurements of nodal voltages and/or currents on a subset of measured buses. Identifying the topology amounts to estimating the complex-valued grid Laplacian matrix \( \mathbf{Y} \) or at least its sparsity pattern (non-zero entries). Depending on the type and location of synchrophasor data, the ensuing setups may be identified.

A. Setup S1) Complete Data

This ideal setup presumes voltage and current phasors \( \{ (i_t, u_t) \} \) are collected at all buses. Data collected over \( T \) times are stacked as columns of \( N \times T \) matrices \( \mathbf{I} \) and \( \mathbf{U} \).

To account for measurement noise, we can postulate the model

\[
\mathbf{I} = \mathbf{YU} + \mathbf{N}. \tag{9}
\]

If \( \text{rank}(\mathbf{U}) = N \), matrix \( \mathbf{Y} \) can be found as the least-squares estimate \( \hat{\mathbf{Y}}_{\text{LS}} := \mathbf{I}^{\text{H}}(\mathbf{UU}^{\text{H}})^{-1} \), which is the minimizer of \( \| \mathbf{I} - \mathbf{YU} \|_2^2 \); see [56], [64]. Guaranteeing \( \text{rank}(\mathbf{U}) = N \) requires \( T \geq N \) and that the feeder undergoes sufficiently different loading conditions so that voltage sequences across buses are linearly independent. Instead of batch processing, [39] puts forth an adaptive scheme to deal with streaming data and updates \( \mathbf{Y} \) in an online fashion using a recursive least squares algorithm.

Either batch or adaptively computed, the least-squares (LS) estimate \( \hat{\mathbf{Y}}_{\text{LS}} \) will not have any zero entries due to numerical/noise errors, whereas the actual \( \mathbf{Y} \) is known to be sparse. To promote sparse estimates, the LS fit can be regularized by the \( \ell_1 \)-norm of the entries of \( \mathbf{Y} \). An \( \ell_1 \)-norm regularizer can be used to find meaningful estimates [5]. This is particularly useful when the sample size \( T \) is smaller than the number of unknown variables in \( \mathbf{Y} \). Additional regularization terms and constraints could be appended to the data fitting problem to enforce or promote the properties of the actual \( \mathbf{Y} \). Reference [48] aims at estimating \( \mathbf{Y} \) using voltage phasors and power rather than current injections based on an LS fit properly regularized and constrained to account for the Laplacian nature of \( \mathbf{Y} \). An alternative two-step heuristic to yield sparse estimates of \( \mathbf{Y} \) is suggested in [96]: After estimating \( \mathbf{Y} \) using an
LS fit, lines whose estimated conductances are relatively small are progressively removed. Heed that (9) models measurement noise on current phasors alone, even though both voltage and current readings are noisy. To deal with both sources of noise, a total least-squares (TLS) or error-in-variables approach is pursued in [96], [97]. A Bayesian version of TLS with sparsity-promoting priors is suggested in [16], [103].

References [2] and [3] consider the same topology identification setup for the case of a DC microgrid. To ensure are progressively removed. Heed that (9) models measurement noise on current phasors alone, even though both voltage and current readings are noisy. To deal with both sources of noise, a total least-squares (TLS) or error-in-variables approach is pursued in [96], [97]. A Bayesian version of TLS with sparsity-promoting priors is suggested in [16], [103].

References [2] and [3] consider the same topology identification setup for the case of a DC microgrid. To ensure are progressively removed. Heed that (9) models measurement noise on current phasors alone, even though both voltage and current readings are noisy. To deal with both sources of noise, a total least-squares (TLS) or error-in-variables approach is pursued in [96], [97]. A Bayesian version of TLS with sparsity-promoting priors is suggested in [16], [103].

B. Setup S2) Partial Injection Data

This setup assumes that voltage data are measured at all buses, but current data are only measured on a subset of buses \(S \subset \mathcal{N} \) with cardinality \( S = |S| \). Let \( \mathcal{S} = \mathcal{N} \setminus S \) denote its complement set. Then, the data model can be partitioned as

\[
\begin{bmatrix}
I_S \\
I_S^T \\
Y_{S_S} \\
Y_{S_S} \\
U_S \\
U_S
\end{bmatrix}
= \begin{bmatrix}
Y_{S_S} \\
Y_{S_S} \\
Y_{S_S} \\
Y_{S_S} \\
U_S \\
U_S
\end{bmatrix}
\begin{bmatrix}
I_S \\
I_S^T \\
Y_{S_S} \\
Y_{S_S} \\
U_S \\
U_S
\end{bmatrix}
(10)
\]

where metered and non-metered buses have been permuted accordingly without losing generality. Given \( (I_S, U_S, U_S) \), matrices \( Y_{S_S} \) and \( Y_{S_S} \) can be estimated using an LS fit.

C. Setup S3) Partial Data

This may be the most practically relevant setup. Voltage and current data are collected only at a subset of buses \( S \). Most methods further assume that all buses in \( S \) are zero-injection buses (ZIB). Plugging \( I_S = 0 \) into (10) yields \( U_S = -Y_{S_S}^{-1} Y_{S_S} U_S \). Substituting \( U_S \) into (10) provides the Kron-reduced data model [36]

\[
I_S = \tilde{Y}_S U_S
(11)
\]

where \( \tilde{Y}_S := Y_{S_S} - Y_{S_S} Y_{S_S}^{-1} Y_{S_S} \). The Kron-reduced or effective admittance matrix \( \tilde{Y}_S \) is still a (reduced) weighted Laplacian of a graph [67]. Nonetheless, this graph is in general non-radial despite the complete graph being radial. Topology recovery algorithms typically proceed in two steps. Matrix \( \tilde{Y}_S \) is first estimated using an LS fit on data \( (I_S, U_S) \). Given this estimate, graph algorithms are then used to infer a radial topology [67], [98]. Reference [67] proposes a complex-valued extension of the recursive grouping algorithm, detailed later in Section IV-C. The approach suggested in [66] aims at recovering the entire topology under the additional assumption that the number of ZIBs is known.

Remark 1: As noted earlier, a general matrix \( Y \) can be recovered from (9) if and only if \( \text{rank}(U) = N \). Unfortunately, such an assumption cannot be satisfied in the presence of ZIBs, as ZIB voltages are linearly dependent on non-ZIB voltages. If there are say \( K \) ZIBs, it holds that \( \text{rank}(U) \leq N - K \), and hence, the LS estimate \( \hat{Y}_{LS} \) is undefined even if the complete voltage and current data are available. However, not all is lost. The previous identifiability analysis treated \( Y \) as a general matrix. If one exploits the fact that \( Y \) is the Laplacian of a tree graph, there exist identifiability claims and algorithms that can recover \( Y \) using partial data. For example, [98] can unveil the actual radial topology from \( Y_S \) if each ZIB has at least three neighbors. References [23], [70], [71] provide a weaker assumption as they can recover the complete \( Y \) if all leaf nodes are non-ZIBs. The aforesaid references rely on a smart meter data model and they are presented in more detail in Section IV-C, yet they can be properly extended to the synchrophasor data model along the lines of [67].

IV. TOPOLOGY IDENTIFICATION VIA SMART METER DATA

Even though synchrophasor data have desirable monitoring features (high sampling frequency, accuracy, and time synchronization), they come at increased cost so widespread deployments are currently rare. Smart meters on the other hand are more widely available. A smart meter may be measuring active power injections, reactive power injections, and/or voltage magnitudes. The data model of choice is the linearized grid model (3a), which we repeat for time \( t \) as

\[
v_t = R p_t + X q + 1 + n_t,
(12)
\]

assuming the substation voltage is fixed at \( v_0 = 1 \) pu without loss of generality. The noise term \( n_t \) captures measurement noise and modeling (e.g., linearization and single-phase approximation) errors. Sample data across times \( t = 1, \ldots, T \), and stack them as columns of the \( N \times T \) matrices \( (V, P, Q) \) to arrive at

\[
V = RP + XQ + 11^T + N.
(13)
\]

Although the emphasis is on smart meter data, model (13) may be relevant to any non-synchronized grid data. Reference [84] arrives at a model similar to (13) upon successive linearization of the power flow equations and estimates the topology matrices using an unscented Kalman filter-based approach.

Different from the phasor data model, model (13) involves real-valued quantities. A key challenge is that one has to estimate two topology matrices \( (R, X) \) from \( NT \) equations. To overcome this issue, two types of approaches (grid probing and covariance-based methods) have been developed and will be reviewed later on. Similar to topology identification using phasor data, the ensuing setups can occur.

A. Setup M1) Complete Data

Similar to setup S1) in Section III-A, suppose one collects power and voltage data at all buses so that \( (V, P, Q) \) are known. Suppose we focus on finding \( R \). To eliminate the effect of reactive power on voltage data, one could try the technique of grid probing [22], [23]. Grid probing works with a data model on successive time differences

\[
\tilde{V} = RP + XQ + \tilde{N}
(14)
\]

where the \( t \)-th column of \( \tilde{P} \) is \( \tilde{p}_t := p_t - p_{t-1} \). In other words, matrix \( \tilde{P} \) contains the differences in active power injections between consecutive times. Matrices \( \tilde{V} \) and \( \tilde{Q} \) are defined similarly. Note that time differentiation amplifies (doubles) the variance in the entries of \( \tilde{N} \) compared to \( N \). Suppose for
now that the operator can control the injections at all buses. If reactive powers are kept or assumed constant across time, model (14) simplifies as

\[ \tilde{V} = R\tilde{P} + \tilde{N} \leftrightarrow \hat{P} = G\tilde{V} - G\tilde{N} \]  

(15)

which upon inversion yields the real-valued counterpart of (9). In theory, any method discussed under S1) is applicable here. In practice, the noise term \( \tilde{N} \) can be stronger than the noise term in (9), and \( G\tilde{N} \) is now correlated across buses.

An alternative idea to eliminate the effect of reactive power injections is a covariance-based approach. All grid data (voltages and powers) have been centered, i.e., for each signal, its mean (time-average) has been subtracted. Due to this pre-processing, the correlation between centered data coincides with the covariance between the original data. All processing is henceforth performed on centered data. Let us compute the correlation between voltage and power injection vectors. From (12), these matrices can be readily computed as

\[ \Sigma_{vp} = R \Sigma_{pp} + X \Sigma_{pq}^T \]  

(16a)

\[ \Sigma_{vp} = R \Sigma_{pp} + X \Sigma_{qq} \]  

(16b)

where \( \Sigma_{vp} := E[v_p p_q^T] \) can be practically estimated from data as \( \Sigma_{vp} = \frac{1}{T} \sum_{t=1}^{T} v_t p_t^T \). To simplify notation, we will henceforth denote \( \Sigma_{vp} \) simply by \( \Sigma_{vp} \). The remaining correlation matrices \( (\Sigma_{pq}, \Sigma_{pq}, \Sigma_{qq}) \) can be estimated from data similarly. In the absence of solar generation, it is reasonable to assume that matrices \( \Sigma_{pp}, \Sigma_{qq} \), and \( \Sigma_{pq} \) are diagonal. Then, it is not hard to verify from (16) that the \( (m,n) \)-th entries of matrices \( \Sigma_{vp} \) and \( \Sigma_{vq} \) satisfy the linear equations

\[ \begin{bmatrix} \mathbb{E}[v_m p_n] \\ \mathbb{E}[v_m q_n] \end{bmatrix} = \begin{bmatrix} \mathbb{E}[p_m] \\ \mathbb{E}[q_m] \end{bmatrix} \mathbb{R}(m,n) \]  

(17)

If \( \mathbb{E}[p_n^2], \mathbb{E}[q_n^2] > (\mathbb{E}[p_n q_n])^2 \), the matrix in (17) is invertible, and we can readily solve for \( \mathbb{R}(m,n) \) and \( \mathbb{X}(m,n) \) and then derive the topology by matrix inversion.

References [99], [100] rely also on complete data \( (V, P, Q) \) to recover grid topology using a two-step approach involving a linearization wherein ratios \( p_n/v_n \) and \( q_n/v_n \) are approximated as linear functions of voltage magnitudes and angles.

B. Setup M2) Partial Injection Data

Paralleling S2) in Section III-B, this setup assumes voltage data are collected at all buses, whereas power injection data are collected only at a subset of buses \( S \). The linearized models in (12)–(15) can be partitioned similarly to (10). Therefore, for setup M2) the given data is \( (V, P_S, Q_S) \). Based on these data, it is not hard to see that the covariance-based approach can estimate matrices

\[ R_S := \begin{bmatrix} R_{SS} \\ R_{SS}^T \end{bmatrix} \quad \text{and} \quad X_S := \begin{bmatrix} X_{SS} \\ X_{SS}^T \end{bmatrix}. \]  

(18)

Setup M2) may be more natural for the grid probing approach since it is unlikely an operator can control injections at every single bus. In that case, the set \( S \) is the set of buses that the operator can actually control. Power injections at \( S \) can be intentionally altered for short intervals. Building again on the difference model and assuming the operator perturbs only active injections at the buses in \( \tilde{S} \), we obtain

\[ \tilde{V} = R_S \tilde{P}_S + \tilde{N} \]  

(19)

The columns of \( R \) associated with \( \tilde{S} \) have been suppressed assuming injections in \( \tilde{S} \) are perturbed for short intervals during which injections in \( S \) remain roughly unchanged.

Despite the partial injection data, the complete matrix \( G = R^{-1} \) is identifiable if \( \text{rank}(\tilde{P}_S) = S \) and the grid is probed at all leaf buses [23]. This identifiability result carries over to the synchrophasor data model. Note that, thanks to the link established between setups M2) and S2) in Remark 1, the algorithms to be presented next are also applicable to M2).

Albeit the previous result asserts that the complete grid topology can be identified, it does not explain how. We next outline three alternative approaches for topology recovery under setup M2). The first group of approaches consists of the graph algorithms put forth in [22], [70], [71]. These algorithms use the knowledge of entries in \( R_{SS} \) to learn the complete leveraging properties of radial networks via a polynomial-complexity algorithm. These graph algorithms are discussed later under setup M3).

Despite their polynomial complexity, graph algorithms can be sensitive to noisy entries of \( R_S \). As an alternative, one can try recovering \( G = R^{-1} \) directly by fitting the available data \( (\tilde{V}, \tilde{P}_S) \); see [23]: Premultiplying (19) by \( G \) yields a model similar to the RHS of (15) with \( \tilde{P}^T = [\tilde{P}_S \ 0^T] \), which can be used to find an LS fit for \( G \). Nonetheless, the estimate \( \tilde{G}_{LS} \) would not satisfy the properties of the actual \( G \). In fact, ensuring the estimate \( \tilde{G} \) is the reduced Laplacian of a tree graph would entail a non-convex feasible set. To arrive at a convex data fitting problem, [23] suggests finding \( G \) as the minimizer of

\[ \min_{G \in \mathcal{M}} \| \tilde{P} - G\tilde{V} \|_F^2 + \lambda \text{Tr} \left[ G \left( I + 11^T \right) \right] - \mu \log |G| \]  

(20)

where \( \lambda, \mu > 0 \) are tunable parameters. Set \( \mathcal{M} \) ensures the minimizer is symmetric with non-positive off-diagonal entries. It can also incorporate prior information, such as \( G_{mn} = 0 \) if buses \( (m,n) \) are known not to be connected. The second term in the objective sums up the absolute off-diagonals of \( G \) to promote sparsity. The third term serves as a barrier keeping \( G \) within the interior of the positive definite cone. Problem (20) can be solved using the alternating direction method of multipliers (ADMM). As a post-processing step, the minimizer can be fed into a minimum spanning tree algorithm to recover a tree graph.

The third approach recovers \( G \) by fitting data while enforcing the estimated \( G \) is the reduced Laplacian of a tree graph via a mixed-integer linear program (MILP) [86]. To simplify notation, rewrite \( G \) from (6a) as \( G = A^T D_p A \) where \( \rho \) collects the inverse line resistances. If \( A \) is known, vector \( \rho \) can be found via an LS fit. To this end, vectorize both sides in the RHS of (19) to get

\[ \tilde{p} = H\rho + \tilde{n} \quad \text{where} \quad H := \tilde{V}^T A^T A^T \]  

(21)

and \( \tilde{n} := \text{vec}(\tilde{P}) \) and \( \tilde{n} := -\text{vec}(G\tilde{N}) \). This follows from the property that \( \text{vec}(X D_p Z^T) = (Z^T X y \) where symbol * denotes...
the Khatri-Rao matrix product. If \( \mathbf{H} \) is full column-rank, the LS estimate of \( \rho \) is [86]

\[
\rho^*_\Lambda = \arg \min_{\rho} \| \tilde{\mathbf{p}} - \mathbf{H} \rho \|_2^2 = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \tilde{\mathbf{p}}.
\]

This problem alone solves the impedance estimation task. Different from topology identification or topology detection, impedance estimation amounts to the task of estimating line impedances assuming the topology to be known. References [74, [75] provide impedance estimation solutions taking into account practical information, such as geographical/length constraints and line types a utility may be using.

For topology identification, however, we also need to find \( \mathbf{A} \). Let us plug \( \rho^*_\Lambda \) back into the LS fitting cost to get

\[
\| \tilde{\mathbf{p}} - \mathbf{H} \rho^*_\Lambda \|_2^2 = \| \tilde{\mathbf{p}} \|_2^2 - \tilde{\mathbf{p}}^T \mathbf{H} (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \tilde{\mathbf{p}}.
\]  (22)

and minimize it over \( \mathbf{H} \) that depends on \( \mathbf{A} \). Recall each row of \( \mathbf{A} \) corresponds to an energized line. Suppose the operator has a library of \( L > N \) lines that could be possibly energized. Candidate lines are encoded as rows of an \( L \times N \) reduced branch-bus incidence matrix \( \tilde{\mathbf{A}} \). Introduce binary variable \( \mathbf{w} \in \{0, 1\}^L \) whose \( \ell \)-th entry indicates whether candidate line \( \ell \) is energized. Finding \( \mathbf{A} \) now amounts to finding \( \mathbf{w} \). Minimizing (22) over \( \mathbf{A} \) can be shown to be equivalent to [86]

\[
\min_{\mathbf{w}, \mathbf{x}} \tilde{\mathbf{p}}^T \mathcal{D}_w \tilde{\mathbf{p}} + \tilde{\mathbf{p}}^T \mathcal{D}_w \mathbf{x}
\]

s.to \((\mathbf{C} - \mathcal{D}_w) \mathbf{x} = \mathcal{D}_w \tilde{\mathbf{p}}\)

radiality constraints on \( \mathbf{w} \in \{0, 1\}^L \).  (23)

where matrix \( \mathbf{C} \) depends on the line library and voltage data as \( \mathbf{C} := [(\tilde{\mathbf{A}} \mathbf{V} \tilde{\mathbf{A}}^T) \circ (\tilde{\mathbf{A}} \tilde{\mathbf{A}}^T) + \mathbf{I}_L]^{-1} \) and \( \circ \) denotes the Hadamard (entrywise) matrix product. Problem (23) can be reformulated as a MILP upon applying McCormick linearization on the products between continuous and binary variables \( \mathcal{D}_w \mathbf{x} \). The formulation is correct as long as \( \mathbf{b} \) yields a connected radial graph, which can be enforced via the radiality constraints discussed in [81, Corollary 1]. Although presented for smart meter data under probing, the approach is applicable to synchronized data setups (S1) and (S2).

To summarize setup M2), if data exhibit high signal-to-noise ratios (SNR), the graph approach is the method of choice as it ensures exact topology recovery at polynomial complexity. For moderate to low SNR, the operator may have to select between the convex heuristic of (20) and the MILP depending on the complexity/accuracy trade-off they can afford.

C. Setup M3) Partial Data

A more realistic setup is that non-synchronized data (voltages and injections alike) are collected only at a subset of buses \( S \). For data \( \{\mathbf{V}_S, \mathbf{P}_S, \mathbf{Q}_S\} \), the correlation-based approach of (16)–(17) can be used to compute the topology submatrices \( \mathbf{R}_{SS}, \mathbf{X}_{SS} \). It is worth noting that correlation-based estimation of \( \mathbf{R}_{SS}, \mathbf{X}_{SS} \) is viable even if injections in \( S \) are correlated (non-diagonal covariance matrices), as long as they remain uncorrelated from unobserved injections in \( \tilde{S} = \mathcal{N} \setminus S \) [72]. This is relevant when \( S \) contains buses with solar generations that can exhibit strong correlation, whereas \( \tilde{S} \) contains load buses alone that can be assumed uncorrelated over shorter intervals. References [22], [70], [71] provide provable algorithms to derive the correct radial topology as long as all unobserved nodes in set \( \tilde{S} \) have degree greater than 2. This implies that all leaf nodes are in \( S \). For all leaf node pairs \( \{m, n\} \), a well-known quantity termed effective resistance can hence be computed as [36]

\[
\mathbf{R}_{eff}(m, n) = \mathbf{R}(m, m) + \mathbf{R}(n, n) - 2\mathbf{R}(m, n).
\]

For radial grids and using (5), effective resistance can be proven to be additive along graph paths, meaning \( \mathbf{R}_{eff}(m, k) = \mathbf{R}_{eff}(m, n) + \mathbf{R}_{eff}(n, k) \) for any node \( k \in \mathcal{P}_{m,n} \). Additive nodal distances enable the construction of the radial graph using an iterative procedure termed recursive grouping (RG); see [25], [73]. Specifically in our grid setting, at the first iteration of RG, each leaf node pair \( \{m, n\} \) with a common parent can be identified by checking if \( \mathbf{R}_{eff}(m, k) - \mathbf{R}_{eff}(n, k) \) is fixed for all other nodes \( k \). If so, the parent of each pair \( \{m, n\} \) is discovered, and the effective resistances between the parents and other observed nodes are derived. During the next iteration, the discovered parents are treated as leaf nodes, and the algorithm uses their effective resistances to identify common parents and so on. A schematic is provided in Fig. 2.

To identify the grid topology from submatrices
$\mathbf{R}_{SS}$ or $\mathbf{X}_{SS}$ building on Prüfer sequences, which resemble recursive grouping.

D. Setup M4) Voltage-Only Data

A fourth setup that is of practical interest is when the operator accesses voltage data alone. This could be the case because the measurement is taken at the photovoltaic panel, and not at the meter of a home. In this case, the meter may be measuring the solar generation, but not the net injection. Another scenario is that the measurement is actually a smart meter recording the total (re)active injection for a house. From the voltage and the injection data, an operator can extrapolate to estimate the voltage at the pole transformer if the secondary line characteristics are known. In this way, the operator can estimate the voltage at the distribution transformer even if such transformer serves 5-20 houses and not all of them are instrumented with smart meters. Given data $\mathbf{V}$, the operator can only compute the centered correlation or covariance matrix

$$
\Sigma_{vv} = \mathbf{R}_{pp} \mathbf{R} + \mathbf{X}_{pq}^T \mathbf{R} + \mathbf{R} \mathbf{X}_{qp} \mathbf{X} + \Sigma_{nn}.
$$

(24)

We next review three approaches that rely on $\Sigma_{vv}$ to identify topologies using graph algorithms.

Signature in Inverse Voltage Covariance. Reference [13] uses $\Sigma_{vv}$ to unveil a radial grid topology under the assumptions that: a1) lines have identical resistance-to-reactance ratios ($r_t/x_t = y$ for all $\ell$) so that $\mathbf{X} = y \mathbf{R}$ from (6); a2) power injections are independent across buses (diagonal $\Sigma_{pp}$, $\Sigma_{qq}$, and $\Sigma_{pq}$); and a3) the measurement/modeling noise term $\Sigma_{nn}$ is negligible compared to other terms in (24). Under these assumptions, the voltage covariance turns out to be $\Sigma_{vv} = \mathbf{X} \Sigma \mathbf{X}$ for a diagonal matrix $\Sigma$ with positive diagonal entries that depend on injection covariances and $y$. Due to this, the inverse covariance (a.k.a. precision matrix in statistics) becomes $\Sigma_{vv}^{-1} = \mathbf{B} \Sigma^{-1} \mathbf{B}^T$. Thus, the precision matrix $\Sigma_{vv}^{-1}$ features the weighted structure of a squared graph Laplacian of a radial network. It can be shown that the $(m, n)$-th entry of $\Sigma_{vv}^{-1}$ is positive if $m = n$ or buses $(m, n)$ have a common neighbor; negative if buses $(m, n)$ are neighbors; and zero, otherwise. In other words, negative valued entries in $\Sigma_{vv}^{-1}$ mark true edges of the radial grid. Reference [13] suggests inverting the sample estimate of $\Sigma_{vv}$ and feeding its negative off-diagonal entries into a minimum-spanning algorithm to recover the grid topology. Numerical tests on a grid violating the common resistance-to-reactance ratio assumption demonstrate satisfactory topology identifiability probabilities for large $T$.

Unfortunately, the precision matrix $\Sigma_{vv}^{-1}$ can be estimated as the inverse of the sample covariance $\hat{\Sigma}_{vv} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{v}_t \mathbf{v}_t^T$ only if $T \geq N$ and the grid undergoes sufficiently rich conditions so that $\hat{\Sigma}_{vv}$ is invertible. Nevertheless, if $\Sigma_{vv}^{-1}$ is known to be sparse, it can be recovered even from singular $\hat{\Sigma}_{vv}$ with proper regularization. To be more specific, if voltage data are modeled as Gaussian random vectors that are independent across time, the precision matrix can be estimated via a sparsity-promoting maximum likelihood approach as the minimizer of the convex problem [91]

$$
\Sigma_{vv}^{-1} = \arg \min_{S} \text{Tr}\left( S \hat{\Sigma}_{vv} \right) + \lambda \| S \|_1 - \log | S |.
$$

(25)

This so-called Graphical Lasso method can estimate sparse precision matrices even with small $T$; see [93]. Its computational complexity scales as $O(N^3)$ [91]. Instead of using Graphical Lasso, each row of the precision matrix can also be estimated by $l_1$-norm regularized regression through the neighborhood Lasso [65], where the voltage sequence at each bus is projected on all other buses as described in [60].

Reference [27] generalized the previous approach to show that when working with voltage phasor data, the assumption of constant $r/x$ ratio at all lines can be waived. In this setting, one can compute covariances between voltage magnitudes and angles and invert the related matrix to get

$$
\Sigma_{vv}^{-1} = \begin{bmatrix} J_1 & J_2 \\ J_2^T & J_3 \end{bmatrix}
$$

where $J_{(v, o)} = \begin{bmatrix} \Sigma_{vv} & \Sigma_{vo} \\ \Sigma_{ov} & \Sigma_{oo} \end{bmatrix}$.

It can be shown that edge $(m, n)$ exists if and only if $J_1(m, n) + J_3(m, n) < 0$, and thus, edges can be identified by negative entries in the inverse covariance. Theoretically, non-adjacent nodes $(m, n)$ with common neighbors (two-hop neighbors) have a strictly positive non-zero entry at $J_1(m, n) + J_3(m, n)$. Entries corresponding to node pairs being more than two hops away are 0. Related works [58], [60] showed that for realistic distribution grids, the values of $J_1(m, n) + J_3(m, n)$ at two-hop nodes are much smaller in magnitude than entries corresponding to adjacent nodes. In other words, thresholding the magnitude of entries in the inverse covariance enables topology detection as

$$(m, n) \in \mathcal{L} \text{ if and only if } |J_1(m, n)| + |J_3(m, n)| > \lambda \tag{26}$$

where $\lambda$ is a user-specified threshold [58], [60]. The inverse covariance matrix $\Sigma_{vv}^{-1}$ bears extra properties that can be exploited to recover grid topology without resorting to sign-based rules as described next.

Conditional Independence Test for Voltages. If voltage data is modeled or approximated as Gaussian random variables, then zero entries in $\Sigma_{vv}^{-1}$ imply that the corresponding variables are conditionally independent; see [27], [93] and references therein. Mathematically, $\Sigma_{vv}^{-1}(m, n) = 0$ implies that $\mathbb{P}(v_m | v_n \forall k \notin \{m, n\}) = \mathbb{P}(v_m | u_k \forall k \notin \{m, n\})\mathbb{P}(v_n | u_k \forall k \notin \{m, n\})$. Conditional independence tests can be accomplished using the inverse covariance matrix or computing mutual information. The graph with edges for every pair of nodes that are not conditionally independent is called the conditional independence graph or graphical model [93]. It follows from the discussion in the previous paragraph that both adjacent nodes (one-hop neighbors), as well as two-hop neighbors in the grid $G_{\mathcal{C}}$, have edges in the conditional independence graph $G_{\mathcal{C}_1}$, as shown in Fig. 4.

Once the conditional independence graph of voltages has been derived, a conditional independence test [27] can be used to distinguish true from spurious grid edges in $G_{\mathcal{C}_1}$. Specifically, if voltages at two nodes $(m, n)$ are able to render voltages at some other nodes $(k, l)$ conditionally independent,
i.e., \( p(v_k, v_l | u_m, u_n) = p(v_k | u_m, u_n) p(v_l | u_m, u_n) \), then \((m, n) \in \mathcal{L}\). Two-hop neighbors do not satisfy this rule, and hence, can be identified and ruled out, while edges between non-leaf nodes (internal nodes) are identified. Once that is completed, edges between leaf nodes and their parents can be further determined.

References [49], [61], [95] present a similar but simpler conditional independence test for topology recovery in realistic grids by showing that two-hop neighbors are almost conditionally independent; see (26). As such, the true edges can be directly estimated by identifying node pairs \((m, n)\) for which their conditional dependence is above a threshold. A similar method using Markov random field (that characterizes conditional independence) is proposed and validated in [101]. While performance is often subpar, the advantage of combinatorial learning algorithms is that they can be used to provide sample guarantees on performance [27], [95]. They can also be easily extended to multiphase setups or setups derived from grid dynamics; see Section VI-D.

In recent works, inverse voltage covariances pertaining to partial data have been used for topology identification. These works extend the sparsity results on \( \Sigma^{-1}_{(v, \bar{v})} \). The observer has access only to the submatrix of \( \Sigma_{(v, \bar{v})} \) corresponding to set \( \tilde{S} \) and denoted as \( \Sigma_{(v, \bar{v})}(\tilde{S}, \tilde{S}) \). If the non-metered buses comprising set \( \tilde{S} \) are sufficiently less and non-adjacent, it has been shown in [4], [35] that

\[
[\Sigma_{(v, \bar{v})}(\tilde{S}, \tilde{S})]^{-1} = \Sigma^{-1}_{(v, \bar{v})}(\tilde{S}, \tilde{S}) + \mathbf{L}
\]

where \( \Sigma^{-1}_{(v, \bar{v})}(\tilde{S}, \tilde{S}) \) is a sub-matrix of the complete inverse covariance, and \( \mathbf{L} \) is a low-rank matrix. Upon computing \( \Sigma_{(v, \bar{v})}(\tilde{S}, \tilde{S}) \) using partial voltage data, regularized convex optimization schemes (expanding on (25)) can separate the sparse component \( \Sigma^{-1}_{(v, \bar{v})}(\tilde{S}, \tilde{S}) \) from the low-rank component \( \mathbf{L} \). While the former provides the edges between observed nodes (set \( \tilde{S} \)), the latter can be further decomposed to identify edges connected to \( \tilde{S} \).

**Greedy Algorithms using Variance of Voltage Differences.**

In addition to inverse voltage covariances, the statistic

\[
\phi_{mn} = \mathbb{E}\left[(v_m - v_n - \mathbb{E}[v_m - v_n])^2\right] \tag{27}
\]

has been utilized for topology identification, oftentimes with better performance at low sample regimes [29], [30], [31]. The statistic \( \phi_{mn} \) is shown to be monotonically increasing along the paths on a radial grid. In other words, if path \((m, n)\) passes through node \(k\), then both \( \phi_{mk} \) and \( \phi_{kn} \) are less than \( \phi_{mn} \). Such increasing trends enable a spanning tree-based topology recovery algorithm: 

1. **Algorithm:** Give each node pair \((m, n)\) a weight of \( \phi_{mn} \) and construct the minimum spanning tree from the set of weights greedily. This will coincide with the true radial topology.

2. **Extension:** As no inverse matrix operation is needed, the performance of the spanning tree algorithm at small numbers of voltage samples is often improved over other voltage-only learning algorithms. We next discuss topology detection schemes and draw connections to the identification schemes presented thus far.

**V. TOPOLOGY DETECTION**

Compared to topology identification, topology detection is a relatively simpler task. The operator now already knows the line infrastructure, namely the impedance of each line as well as the buses incident to it. The operator does not know the status of all or a subset of lines. Given data, topology detection amounts to deciding which of the existing lines are energized (see Fig. 5). This is useful during normal grid operation due to automated switching. Variants of topology detection schemes are also relevant after extreme weather events to inform grid restoration efforts for line outages and dispatch line crews.

A general formulation parallels the modeling after (22) by including the now available information on lines. Often, this information is available as a library of known lines \( \mathcal{L}_{\text{all}} \) with cardinality \( \mathcal{L}_{\text{all}} \), and known impedances. Out of these lines, the operator would like to select a set of lines \( \mathcal{L} \) of size \( N \) that are energized. Each candidate line can be associated with a row of an augmented \( \mathcal{L}_{\text{all}} \times N \) branch-bus incidence matrix \( \mathbf{A} \). The status of line \( \ell \) is indicated by a binary variable \( w_\ell \). If vector \( \mathbf{w} \) and a diagonal matrix \( \mathbf{D} \) collect all \( L \) line statuses, the topology matrices of interest can be parameterized as

\[
Y(\mathbf{w}) = \mathbf{A}^\top \mathbf{D} \mathbf{y} \mathbf{A} \tag{28a}
\]

\[
R(\mathbf{w}) = \left(\mathbf{A}^\top \mathbf{D} \mathbf{x} \mathbf{A}^\top \right)^{-1} \tag{28b}
\]

\[
X(\mathbf{w}) = \left(\mathbf{A}^\top \mathbf{D} \mathbf{x} \mathbf{A}^\top \right)^{-1} \tag{28c}
\]
Apparently, if the status of line $\ell$ is known, variable $w_\ell$ can be set to a fixed value. As with topology identification, topology detection methods are organized next based on the type of measurements used. Before proceeding with the methods, it is worth adding a note on topology detectability. Sections III and IV discussed topology identifiability results, i.e., conditions under which the complete or reduced topology of a grid can be recovered from complete or partial data. Given that topology detection is easier than identification, detectability can be ensured with less strict conditions on the number and placement of meters as established in [21].

### A. Topology Detection Using Phasor Data

Analogous to setup $S1$ in Section III-A, if phasor data are collected at all buses, one can try detecting $w$ via an LS fit. Similar to (21), upon vectorizing the data model in (9) and exploiting the parameterization $Y(w)$, we arrive at the mixed-integer quadratic problem (MIQP):

$$\min_w \| \text{vec}(I) - Hw \|^2$$

s.t. to radiality constraints on $w \in \{0, 1\}^{\ell \text{all}}$. (29)

where $H := U^T \hat{A}^T \hat{D}_{\ell} \ast \hat{A}^T$.

If the operator has to select among a few possible topologies $\{w_i\}_{i=1}^S$, one may evaluate which topology yields the smallest fitting error $\min_i \| I - Y(w_i) U \|^2$, or the voltage reconstruction error $\min_i \| Y^{-1}(w_i) I - U \|^2$, as suggested in [6].

### B. Topology Detection Using Smart Meter Data

The topology identification schemes presented earlier can be tailored for topology detection. For example, if the operator collects data using probing under setup $M2$ [cf. Section IV-B], one can detect $w$ by modifying (20); see [23]. The regularization terms may not be needed anymore, and the formulation becomes similar to the one in (29). To avoid mixed-integer programs, one may also pursue a box relaxation on $w$.

Adjusting identification algorithms towards topology detection is straightforward for schemes under partial data setups $M2$–$M3$ described in Sections IV-B and IV-C. As the set $\ell \text{all}$ of possible lines is known, edge discovery in the combinatorial graph algorithms is restricted to $\ell \text{all}$ instead of considering all node pairs. For example, in the recursive grouping-based scheme in $M2$ and $M3$, known edges can be included beforehand and only permissible parent-child pairs are discovered.

Similarly, detection algorithms for voltage-only data can be designed by modifying the signature or conditional independence (CI)-based tests on inverse voltage covariances, discussed under setup $M4$ in Section IV-D, to be restricted to $\ell \text{all}$. In an alternate approach, the dependence of permissible edges $w$ on $\Sigma_{\ell \ell}$ can be directly parameterized through (28) and (24). Similar to (25), given the sample voltage covariance $\hat{\Sigma}_{\ell \ell}$, a maximum likelihood detector is put forth in [24] to select $N$ edges from $\ell \text{all}$:

$$\min_w \text{Tr}(\Sigma_{\ell \ell}^{-1}(w) \hat{\Sigma}_{\ell \ell}) - \log | \Sigma_{\ell \ell}^{-1}(w) |$$

s.t. $w \in \{0, 1\}^{\ell \text{all}}, \ 1^T w = N$.

Despite the box relaxation on $w$, the objective function is non-convex, and a stationary point is found through a projected gradient descent algorithm. In a related work, [80] uses a maximum likelihood framework for Gaussian voltage distributions to identify topology changes from a set of permissible edges.

**Detection with Partial Voltage and Injection Statistics.** The variances of voltage differences $\phi_{mn}$, defined in (27) under $M4$, can be used to detect topologies if available for all node pairs. Additionally, it was shown in [28], [29], [31] that $\phi$'s carry more information:

$p1)$ The statistic $\phi$ has additional ordering for triples of nodes $(m, n, k)$, that is, the sign of $\phi_{mn} - \phi_{mk} - \phi_{nk}$ depends on the relative position of $(m, n, k)$ on the grid; see Fig. 6.

$p2)$ The value of $\phi_{mn}$ at adjacent nodes $(m, n)$ with parent node $m$, depends on the admittance of edge $(m, n)$ and injections at all descendant nodes of $n$.

These ordering properties enable topology detection, where voltage and injection statistics can be computed over a subset of buses $S$. Under partial data, [31] first creates a spanning tree for observed nodes using available $\phi$’s, and then inserts missing nodes iteratively by checking for properties $p1)$–$p2)$. The algorithm is able to detect the true topology provided each unobserved node has degree greater than 2. This is similar in theme to the learning algorithms proposed under recursive grouping in $M3$, with the crucial distinction that the algorithm has access to injection statistics and line impedances in $\ell \text{all}$, but not injection samples. The latter means that $\Sigma_{\ell \ell}$ and hence $\mathbf{R}$ cannot be estimated.

If we are confined to a few candidate topology vectors $\{w_i\}_{i=1}^S$, [50] suggests a model validation process, wherein the power flow solution obtained using each candidate topology is compared with the collected grid data to determine the operating topology. A similar workflow is pursued in [63] using Bayesian networks.

### C. Topology Detection Using Line Flow Data

Previously, we assumed that sensors have been installed only at buses. However, some works cope with topology detection using also line flow measurements. The work [40] considers a setup where only line current measurements and load (pseudo) measurements are available. Thanks to a power
flow linearization, they cast topology detection as a MILP to minimize the mismatch between measured and computed line currents. Reference [79] exploits the fact that in radial networks and upon neglecting power losses, the power flow on a line \((m, n)\) equals the sum of powers of all downstream buses. Given line flow and voltage data at a few locations along with (pseudo) measurements of loads, [90] casts topology detection as a mixed-integer quadratic program.

VI. EXTENSIONS

The schemes discussed so far refer to balanced or single-phase and radial grids, presumed steady-state operation, and relied on readings of electric and particularly nodal quantities (voltages, injections). We next briefly discuss topology learning schemes that waive one or more of these assumptions.

A. Learning Loopy/Meshed Topology

While distribution grids are typically operated under radial topologies, loopy or meshed topologies can be found at heavily loaded urban distribution grids [78]. Despite complications related to protection schemes, meshed topologies could advance reliability, especially in micro-grid settings. In this context, the inverse voltage covariance-based topology identification methods discussed in Section IV-D, can also learn meshed grids under realistic restrictions on the minimum size of loops/cycles, as shown in [34], [49], [58]. Further, these methods also extend to meshed grids with ZIBs (zero-injection buses), where the voltage covariance matrix is non-invertible. In that case, [33] first identifies ZIBs and their neighbors using flow conservation-based voltage regression, and then, discovers the remaining topology using properties of the inverse voltage covariance at non-ZIBs.

B. Multiphase Topology Learning

The linearized power flow model can be extended to multiphase settings under similar assumptions of small angle differences (per phase) for neighboring grid buses [43], [51]. The primary complexity under a multiphase model arises from the fact that voltages and injections at each bus are no longer complex scalar quantities anymore, and each line admittance is a complex-valued \(3 \times 3\) or \(2 \times 2\) matrix with off-diagonal entries. Furthermore, the phase assignment of each bus may be unknown. While an extensive discussion on phase identification falls outside the scope of this article, we note that methods based on voltage clustering and local voltage-based distances have been employed to assign phases [10], [12], [69], [94]. If phases are correctly identified or known, the variance of nodal voltage differences (see the single-phase case under setup M4) in Section IV-D) can be used as edge weights to identify the topology as a spanning tree. Crucially, even without phase identification, the conditional independence structure of three-phase nodal voltages factorizes exactly as described for the single-phase case in Section IV-D and Fig. 4. Conditional independence tests for three-phase voltages have been proposed for topology identification in [32], [59], using properties akin to the single-phase setting.

A mixed-integer linear program (MILP) formulation for three-phase topology detection is presented in [44], that can be categorized as a three-phase extension of the partial observability setup M2) in Section IV-B. The MILP objective here minimizes the weighted error in three-phase linearized power flow equations involving state variables, smart meter data, and pseudo-measurements. Reference [57] considers a regime where three-phase voltages and injections at leaf nodes alone are available, similar to the single-phase setup M2) in Section IV-C. The topology identification uses recursive grouping but is applied to efficiently estimate in-phase effective impedances.

C. Learning With Multiple Feeders

Distribution networks are typically operated as a union of radial sub-networks, each fed by a different substation or substation transformer referred to as source point. However, only a few works can cope with topology learning if the knowledge of which bus is fed by what source point is not known in advance [24], [30]. Such information might be lacking because utilities may have limited information about secondary circuits, and loads are oftentimes transferred between feeders to alleviate overloads, or during restoration after extreme weather events. Hence, before performing any topology processing algorithm, one should cluster in advance buses based on the source point to which they are connected. The main idea here is that the power flowing through the source point matches, up to the power loss, to the sum of the served loads [8], [20], and that voltages across feeders can be approximated as uncorrelated [20], [30].

D. Learning With Grid Dynamics

All previous works considered the steady-state operation of the distribution grid. In a fast time-scale setting, [11] compares the voltage fluctuations experienced at each bus after events, such as transformer tap or capacitor bank switching. It is observed that the closer two buses are, the more similar their responses to such events are; hence this can be used as a metric for topology identification. In fact, as grid reconfiguration itself induces changes in voltage profiles, researchers have devised methods to detect topology changes by comparing voltage data with a library of signatures [17], [18], [19], or using a subspace perturbation model [102].

E. Topology Processing With Non-Electric Data

So far, we have discussed how grid topologies can be inferred by processing electric grid data. Nonetheless, grid topologies can also be estimated using non-electric data. For example, reference [38] uses power line communication (PLC) signals to identify neighboring nodes on a radial grid. It first measures the communication delay \(t_{nm}\) for a PLC signal to be propagated between any pair \((n, m)\) of buses. Based on pairwise delays, two buses \((n, m)\) are declared as neighbors if \(t_{nm} < t_{nk} + t_{km}\) for any other bus \(k \neq \{n, m\}\). The key principle here is that the communication delay scales with the line length.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
VII. TOPICS NOT COVERED AND FUTURE DIRECTIONS

This tutorial has not discussed the considerable amount of work on topology learning in transmission systems. For example, the classical task of generalized state estimation (GSE) aims at jointly estimating the system state along with its topology [1], [52]. Beyond GSE, there have been efforts towards learning the topology of transmission grids using various types of data. For example, [47] uses voltage angle data to recover the network topology using an optimization scheme similar to (25). Reference [53] exploits the structure of wholesale electricity markets to infer the topology of power transmission systems using locational marginal prices. References [87], [89] attempt to recover network topologies from voltage angle data exploiting the structure of swing dynamics. While phase angle samples now are temporally correlated, conditional independence-based tests, similar to ones in Section IV-D, can identify neighboring buses even in meshed grids. The key difference here is that instead of inverse covariance, the sparsity and signs of the inverse power spectral density matrix are exploited; power spectral density is the Fourier transform of auto-correlation $E[\theta_t \theta_{t+h}^T]$. Conditional independence-based tests on inverse power spectral density have been extended for consistent identification in grids with partial observability [88], as well.

Focusing again on distribution grids, there is extensive literature on other topology-processing topics that were only briefly discussed or not touched upon at all here. For example, the concept of GSE has been extended to distribution grids. As with topology detection, line impedances are assumed to be known. It is performed by solving weighted least-squares or least-absolute-value (WLAV) problems over continuous (system states) and binary (line or breaker statuses) variables [76]. If the measurement set does not ensure identifiability, pseudo-measurements are oftentimes introduced [55]. Reference [82] casts GSE as a mixed-integer quadratic program for unbalanced multiphase feeders with possibly missing/bad data.

In addition, there is a growing literature on model-agnostic and supervised topology processing techniques, that has not been covered herein. Such techniques typically proceed in two stages. During the training stage, they collect labeled grid data [37]. Labels may correspond to known outages, particular topologies, or fault locations. Such labeled data can be used either as a library of recorded events; or alternatively, used to train a machine learning model (e.g., deep neural network, DNN) in a supervised manner for classification or regression; see, e.g., [9] and references therein. During operation, the operator uses the current data and the learned model to infer grid topology information. Most of the methods reviewed in this article presumed customers are Wye-connected and hence may need modifications to deal with Delta or line-to-line connections in general. Fault localization has been an area of active research in power transmission systems, with increasing relevance in distribution grids due to several grid-initiated wildfires in recent years. Detecting and identifying such faults with limited phasor data is a crucial research direction.

Moving forward, there are several research pathways in learning distribution grid topologies that warrant further investigation. Research directions for distribution grid topology learning that have not been studied as extensively include but are not limited to nonlinear power flow models or setups where grid data exhibit high correlation (e.g., due to solar injections) or suffer from low signal-to-noise ratios. Moreover, learning microgrid topologies is a pertinent task given their potentially mobile and ad-hoc nature. Unveiling meshed topologies may become increasingly more relevant for densely populated urban distribution grids. On the methodological front, tensor processing techniques may prove fruitful towards unveiling grid topologies.

New metering technologies and the aggregation of large amounts of heterogeneous data offer unprecedented opportunities for topology learning. Collecting synchrophasor data at higher-order harmonics in addition to the phasors at the fundamental frequency may offer near opportunities for grid topology learning. New meters that can record phasors accurately at sub-cycle levels (termed point-on-wave measurements) have led to growing interest in using them to identify topologies, estimate impedances, and deal with affiliated tasks, despite limited placement. Localizing hidden sources of oscillations has also become pertinent given new types of loads and inverter-interfaced distributed energy resources (DERs). PMUs or sensors with synchrophasor functionality installed in DERs provide data on the grid edge. Smart inverters and other smart grid devices (e.g., smart transformers, remotely-controlled voltage regulators and capacitor banks, D-STATCOMs) offer additional opportunities to collect data and estimate voltage sensitivities upon actuation and probing. Flow measurements from transformers and line sensors add to the abundance of data. Non-electric data such as GPS (global positioning system) data from AMI as well as data from drones surveying lines bring in a spatial dimension to the topology learning task at hand. Thermal cameras sensing the temperature of lines from afar can potentially infer whether a line is energized or not. Fusing data from heterogeneous sources, such as smart meters, phasor measurement units, line sensors, or geographical data from city maps and street layouts, aerial/drone images could be of practical interest.

As topology estimation methods are data-driven, targeted adversarial corruption of the measurement set can lead to errors in topology recovery. This is within the purview of data injection and jamming based attacks in power grids and warrants future research. Meter placement to improve topology estimation [37] and robustness is a potential area of future work. Topology identification and detection tools may be linked to feedback loops for data-driven grid control, where changes in structure or parameter values are seamlessly incorporated into evolving control logic. Studying the optimality and uncertainty quantification of such feedback presents an interesting research opportunity.

VIII. CONCLUSION

This tutorial has presented a study of topology identification and detection methods in radial distribution grids. Focusing
on analytical methods, we have compared and contrasted topology learning algorithms based on measurement type, the extent of observability, and the solution approach proposed. Under realistic low-observability regimes, we have presented explanatory discussions on the use of power flow physics and topological constraints in different algorithms to ensure topology recovery as well as discussed limitations on exact estimation. Connections to learning schemes for meshed, unbalanced grids, and related problems in change detection and physics-agnostic learning methods have been elucidated. Finally, we have commented on potential directions where the current state-of-the-art methods may be extended in the future.

ACKNOWLEDGMENT

The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

REFERENCES

[1] A. Abur and A. Gómez-Expósito, Power System State Estimation: Theory and Implementation. New York, NY, USA: Marcel Dekker, 2004.
[2] M. Angelichinoski, A. Scaglione, P. Popovski, and C. Stefanović, “Decentralized DC microgrid monitoring and optimization via primary control perturbations,” IEEE Trans. Signal Process., vol. 66, no. 12, pp. 3280–3295, Jun. 2018.
[3] M. Angelichinoski, C. Stefanovic, P. Popovski, A. Scaglione, and P. Blaabjerg, “Topology identification for multiple-bus DC microgrids via primary control perturbations,” in Proc. Int. Conf. DC Microgrids, 2017, pp. 202–206.
[4] R. Anguluri, G. Dasarathy, O. Kosut, and L. Sankar, “Grid topology identification with hidden nodes via structured norm minimization,” IEEE Control Syst. Lett., vol. 6, pp. 1244–1249, 2021.
[5] O. Ardakanian et al., “On identification of distribution grids,” IEEE Trans. Control Netw. Syst., vol. 6, no. 3, pp. 950–960, Sep. 2019.
[6] R. Arghandeh, M. Gahr, A. van Meier, G. Cavraro, M. Ruh, and G. Andersson, “Topology detection in microgrids with micro-synchrophasors,” in Proc. IEEE Power Energy Soc. Gen. Meeting, 2015, pp. 1–5.
[7] R. Arghandeh and Y. Zhou, Big Data Application in Power Systems. Amsterdam, The Netherlands: Elsevier, 2017.
[8] V. Arya, T. Jayram, S. Pal, and S. Kalyanaraman, “Inferring connectivity model from meter measurements in distribution networks,” in Proc. ACM Int. Conf. Future Energy Syst., 2013, pp. 173–182.
[9] B. Azimian, R. S. Biswas, S. Moshtagh, A. Pal, L. Tong, and G. Dasarathy, “State and topology estimation for unobservable distribution systems using deep neural networks,” IEEE Trans. Instrum. Meas., vol. 71, pp. 1–14, 2022.
[10] M. Bariya, D. Deka, and A. von Meier, “Guaranteed phase & topology identification in three phase distribution grids,” IEEE Trans. Smart Grid, vol. 12, no. 4, pp. 3605–3612, Jul. 2021.
[11] M. Bariya, A. von Meier, A. Ostfeld, and E. Ratnam, “Data-driven topology estimation with limited sensors in radial distribution feeders,” in Proc. IEEE Green Tech. Conf. (GreenTech), 2018, pp. 183–188.
[12] L. Blakely, M. J. Reno, and W.-C. Feng, “Spectral clustering for customer phase identification using AMI voltage time-series,” in Proc. IEEE Power Energy Conf. Illinois (PECI), 2019, pp. 1–7.
[13] S. Bolognani, N. BoF, D. Michelotti, R. Muraro, and L. Schenato, “Identification of power distribution network topology via voltage correlation analysis,” in Proc. IEEE Conf. Decis. Control, 2013, pp. 1659–1666.
[14] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, “Distributed reactive power feedback control for voltage regulation and loss minimization,” IEEE Trans. Autom. Control, vol. 60, no. 4, pp. 966–981, Apr. 2015.
[15] S. Bolognani and F. Dörfler, “Fast power system analysis via implicit linearization of the power flow manifold,” in Proc. Allerton Conf., 2015, pp. 402–409.
[16] J.-S. Brouillon, E. Fabbiani, P. Nahata, F. Dörfler, and G. Ferrari-Trecate, “Bayesian methods for the identification of distribution networks,” in Proc. IEEE Conf. Decis. Control, 2021, pp. 3646–3651.
[17] G. Cavraro, R. Arghandeh, G. Barchi, and A. von Meier, “Distribution network topology detection with time-series measurements,” in Proc. IEEE Conf. Innov. Smart Grid Technol., 2015, pp. 1–5.
[18] G. Cavraro, R. Arghandeh, K. Poolla, and A. von Meier, “Data-driven approach for distribution network topology detection,” in Proc. IEEE Power Energy Soc. Gen. Meeting, 2015, pp. 1–5.
[19] G. Cavraro and R. Arghandeh, “Power distribution network topology detection with time-series signature verification method,” IEEE Trans. Power Syst., vol. 33, no. 4, pp. 3500–3509, Jul. 2018.
[20] G. Cavraro and A. Bernstein, “Bus clustering for distribution grid topology identification,” IEEE Trans. Smart Grid, vol. 11, no. 5, pp. 4080–4089, Sep. 2020.
[21] G. Cavraro, A. Bernstein, V. Kekatos, and Y. Zhang, “Real-time identifiability of power distribution network topologies with limited monitoring,” IEEE Control Syst. Lett., vol. 4, no. 2, pp. 325–330, Apr. 2020.
[22] G. Cavraro and V. Kekatos, “Graph algorithms for topology identification using power grid probing,” IEEE Control Syst. Lett., vol. 2, no. 4, pp. 689–694, Oct. 2018.
[23] G. Cavraro and V. Kekatos, “Inverter probing for power distribution network topology processing,” IEEE Trans. Control Netw. Syst., vol. 6, no. 3, pp. 980–992, Sep. 2019.
[24] G. Cavraro, V. Kekatos, and S. Veeramachaneni, “Voltage analytics for power distribution network topology verification,” IEEE Trans. Smart Grid, vol. 10, no. 1, pp. 1058–1067, Jan. 2019.
[25] M. J. Choi, V. Y. F. Tan, A. Anandkumar, and A. S. Willsky, “Learning latent tree graphical models,” J. Mach. Learn. Res., vol. 12, pp. 1771–1812, Nov. 2011.
[26] K. Dehghapour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, “A survey on state estimation techniques and challenges in smart distribution systems,” IEEE Trans. Smart Grid, vol. 10, no. 2, pp. 2312–2322, Mar. 2019.
[27] D. Deka, S. Backhaus, and M. Chertkov, “Estimating distribution grid topologies: A graphical learning based approach,” in Proc. Power Syst. Comput. Conf., 2016, pp. 1–7.
[28] D. Deka, S. Backhaus, and M. Chertkov, “Learning topology of distribution grids using only terminal node measurements,” in Proc. Int. Conf. Smart Grid Commun., 2016, pp. 205–211.
[29] D. Deka, S. Backhaus, and M. Chertkov, “Learning topology of the power distribution grid with and without missing data,” in Proc. Eur. Control Conf., 2016, pp. 313–320.
[30] D. Deka, S. Backhaus, and M. Chertkov, “Structure learning in power distribution networks,” IEEE Trans. Control Netw. Syst., vol. 5, no. 3, pp. 1061–1074, Sep. 2018.
[31] D. Deka, M. Chertkov, and S. Backhaus, “Joint estimation of topology and injection statistics in distribution grids with missing node labels,” IEEE Trans. Control Netw. Syst., vol. 7, no. 3, pp. 1391–1403, Sep. 2020.
[32] D. Deka, M. Chertkov, and S. Backhaus, “Topology estimation using graphical models in multi-phase power distribution grids,” IEEE Trans. Power Syst., vol. 35, no. 3, pp. 1663–1673, May 2020.
[33] D. Deka, D. Harsik, S. Misra, and M. V. Salapaka, “Tractable learning in under-excited power grids,” IEEE Trans. Control Netw. Syst., vol. 9, no. 2, pp. 763–774, Jun. 2022.
[34] D. Deka, S. Talukdar, M. Chertkov, and M. V. Salapaka, “Graphical models in meshed distribution grids: Topology estimation, change detection & limitations,” IEEE Trans. Smart Grid, vol. 11, no. 5, pp. 4299–4310, Sep. 2020.
[35] H. Doddi, D. Deka, and M. Salapaka, “Learning partially observed meshed distribution grids,” in Proc. Int. Conf. Probab. Methods Appl. Power Syst., 2020, pp. 1–6.
[36] F. Dörfler and F. Bullo, “Krone reduction of graphs with applications to electrical networks,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 60, no. 1, pp. 150–163, Jan. 2013.
[37] G. S. Dua, B. Tyagi, and V. Kumar, “A novel approach for configuration identification of distribution network utilizing microPMU data,” IEEE Trans. Appl. Ind., vol. 57, no. 1, pp. 857–868, Jan./Feb. 2021.
[87] S. Talukdar, D. Deka, D. Materassi, and M. V. Salapaka, “Exact topology reconstruction of radial dynamical systems with applications to distribution system of the power grid,” in *Proc. Amer. Control Conf. (ACC)*, 2017, pp. 813–818.

[88] S. Talukdar, D. Deka, M. Chertkov, and M. V. Salapaka, “Topology learning of radial dynamical systems with latent nodes,” in *Proc. Amer. Control Conf. (ACC)*, 2018, pp. 1096–1101.

[89] S. Talukdar, D. Deka, H. Doddi, D. Materassi, M. Chertkov, and M. V. Salapaka, “Physics informed topology learning in networks of linear dynamical systems,” *Automatica*, vol. 112, Feb. 2020, Art. no. 108705.

[90] Z. Tian, W. Wu, and B. Zhang, “A mixed integer quadratic programming model for topology identification in distribution network,” *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 823–824, Jan. 2016.

[91] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *J. Roy. Stat. Soc. B Methodol.*, vol. 58, no. 1, pp. 267–288, 1996.

[92] “Annual Electric Power Industry Report.” Jun. 2022. [Online]. Available: https://www.eia.gov/electricity/data/eia861/

[93] M. J. Wainwright and M. I. Jordan, “Graphical models, exponential families, and variational inference,” *Found. Trends Mach. Learn.*, vol. 1, nos. 1–2, pp. 1–305, Nov. 2008.

[94] W. Wang, N. Yu, B. Foggo, J. Davis, and J. Li, “Phase identification in electric power distribution systems by clustering of smart meter data,” in *Proc. IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, 2016, pp. 259–265.

[95] Y. Weng, Y. Liao, and R. Rajagopal, “Distributed energy resources topology identification via graphical modeling,” *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2682–2694, Jul. 2017.

[96] J. Yu, Y. Weng, and R. Rajagopal, “PaToPa: A data-driven parameter and topology joint estimation framework in distribution grids,” *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 1682–1692, May 2019.

[97] J. Yu, Y. Weng, and R. Rajagopal, “PaToPa-EM: A data-driven parameter and topology joint estimation framework for time-varying system in distribution grids,” *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 1682–1692, May 2019.

[98] Y. Yuan, S. H. Low, O. Ardakanian, and C. J. Tomlin, “Inverse power flow problem,” *IEEE Trans. Control Netw. Syst.*, vol. 10, no. 1, pp. 261–273, Mar. 2023.

[99] J. Zhang, P. Wang, and N. Zhang, “Distribution network admittance matrix estimation with linear regression,” *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4896–4899, Sep. 2021.

[100] J. Zhang, Y. Wang, Y. Weng, and N. Zhang, “Topology identification and line parameter estimation for non-FMU distribution network: A numerical method,” *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4440–4453, Sep. 2020.

[101] J. Zhao, L. Li, Z. Xu, X. Wang, H. Wang, and X. Shao, “Full-scale distribution system topology identification using Markov random field,” *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 4714–4726, Nov. 2020.

[102] N. Zhou, L. Luo, G. Sheng, and X. Jiang, “Power distribution network dynamic topology awareness and localization based on subspace perturbation model,” *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1479–1488, Mar. 2020.

[103] H. Zhu, G. Leus, and G. B. Giannakis, “Sparsity-cognizant total least-squares for perturbed compressive sampling,” *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2002–2016, May 2011.

Deepjyoti Deka (Senior Member, IEEE) received the undergraduate degree in electronics and communication engineering from IIT Guwahati, India, with an Institute Silver Medal as the best outgoing student of the department in 2009, and the M.S. and Ph.D. degrees in electrical and computer engineering from The University of Texas at Austin, TX, USA, in 2011 and 2015, respectively. He is a Staff Scientist with the Theoretical Division, Los Alamos National Laboratory (LANL), where he was previously a Postdoctoral Research Associate with the Center for Nonlinear Studies. At LANL, he serves as a PI/Co-PI for DOE and LDRD projects on machine learning in power systems, interdependent networks, and in cyber-physical security. He has served as an Editor for IEEE TRANSACTIONS ON SMART GRID from 2020 to 2021.

Vassilis Kekatos (Senior Member, IEEE) received the Diploma, M.Sc., and Ph.D. degrees in computer science and engineering from the University of Patras, Greece, in 2001, 2003, and 2007, respectively. He is an Associate Professor with the Bradley Department of ECE, Virginia Tech. In 2014, he was a Visiting Researcher with The University of Texas at Austin and the Ohio State University. He has been a Research Associate with the ECE Department, University of Minnesota. His research focus is on optimization and learning for future energy systems.

He is a recipient of the NSF Career Award in 2018 and the Marie Curie Fellowship. He received the Postdoctoral Career Development Award (honorable mention) from University of Minnesota. From 2015 to 2022, he served on the editorial board of the IEEE TRANSACTIONS ON SMART GRID.

Guido Cavraro (Member, IEEE) received the Ph.D. degree in information engineering from the University of Padova, Italy, in 2015. He was a Visiting Scholar with the California Institute for Energy and Environment, University of California at Berkeley in 2014. From 2015 to 2016, he was a Postdoctoral Associate with the Department of Information Engineering, University of Padova. From 2016 to 2018, he was a Postdoctoral Associate with the Bradley Department of Electrical and Computer Engineering, Virginia Tech, USA. He is currently a Senior Researcher with the Power Systems Engineering Center, National Renewable Energy Laboratory, USA. His research interests include control, optimization, and estimation applied to power systems and smart grids.