This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

Inverse Power Flow Problem

Ye Yuan, *Senior Member, IEEE,* Steven H. Low, *Fellow, IEEE,* Omid Ardakanian, *Member, IEEE,* Claire J. Tomlin, *Fellow, IEEE*

***Abstract*—This paper formulates an inverse power flow problem which is to infer a nodal admittance matrix (hence the network structure of a power system) from voltage and current phasors measured at a number of buses. We show that the admittance matrix can be uniquely identified from a sequence of measurements corresponding to different steady states when every node in the system is equipped with a measurement device, and a Kron-reduced admittance matrix can be determined even if some nodes in the system are not monitored (hidden nodes). Furthermore, we propose effective algorithms based on graph theory to uncover the actual admittance matrix of radial systems with hidden nodes. We provide theoretical guarantees for the recovered admittance matrix and demonstrate that the actual admittance matrix can be fully recovered even from the Kron-reduced admittance matrix under some mild assumptions. Simulations on standard test systems confirm that these algorithms are capable of providing accurate estimates of the admittance matrix from noisy sensor data.**

***Index Terms*—Inverse Power Flow Problem, System Identification, Kron Reduction, Phasor Measurement Units.**

I. Introduction

THE power industry has witnessed profound changes in recent years. These changes are mostly driven by the widespread adoption of distributed energy resources (DER), active participation of customers in emerging energy markets, and rapid deployment of measurement, communication, and control infrastructure resulting in an unprecedented level of visibility and controllability, especially for distribution grids. Despite the increased amount of uncertainty, these changes offer opportunities for system operators to improve power system stability and efficiency by leveraging advanced optimization and control techniques. Most of these techniques require the knowledge of the network topology in real time.

The inverse power flow (IPF) problem we define in this paper concerns the estimation of the nodal admittance matrix from synchronized measurements of voltage and current phasors (i.e., magnitudes and phase angles) which can be obtained from phasor measurement units (PMUs) or conventional supervisory control and data acquisition (SCADA) technology. The IPF problem underlies several crucial smart grid applications, affecting real-time system operation and long-term planning, the most important of which are:

1) State Estimation [1] combines the knowledge of the admittance matrix with a set of known state-variables to determine the unknown ones, e.g., voltage magnitude and phase angle of unobserved buses, thereby building a real-time model of the network for control.

2) Optimization & Control [2] techniques determine a sequence of operations that can transition the power system from one steady state to another steady state that meets certain stability and efficiency targets.

3) Event Detection [3] concerns identifying faults, line outages, and other critical events, such as transformer tap changes, capacitor and switching operations from changes in the real-time network model.

4) Cybersecurity [4] is the problem of identifying the potential vulnerabilities of a power system and designing strategies to protect it from the potential cyber attacks using telemetry data along with information about its topology.

In this paper, we lay out a theoretical framework for the IPF problem. Using the bus injection model (BIM), we propose efficient algorithms to identify the admittance matrix. In particular, we show that when the system has no hidden nodes, the admittance matrix can be uniquely identified from a sequence of complex voltage and current measurements corresponding to different steady states. Should there be some hidden nodes in the network, we show that a reduced admittance matrix (from Kron reduction [5]) can be determined; we develop a method based on graph decomposition, maximal clique searching and composition for identifying the admittance matrix of the original system for radial networks. Power flow simulations are performed on the IEEE 14-bus system to illustrate the theoretical results and evaluate their sensitivity to measurement noise introduced by transducers[1].

The paper is outlined as follows: after surveying related work in Section II, we formulate the IPF problem and propose a solution for the case that the system is fully observable in Section III. When the system has hidden nodes, we propose efficient algorithms to solve the IPF problem for radial networks with theoretical guarantee in Section IV. We conclude the paper by presenting directions for future work in Section V.

II. Related Work

The availability of high-precision, high-sample-rate measurements of transmission and distribution networks in recent years has given impetus to research on topology and admittance identification.

The IPF problem that identifies both topology and admittance matrix has been studied extensively in transmission networks [7], [8], [9], [10] as well as distribution grids [11], [12] using single-phase a.c. and d.c. power flow models. For example, the topology identification problem is cast as a sparse subspace learning problem in [7] and an efficient algorithm is

Ye Yuan is with School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China. Steven H. Low is with Department of Computer and Mathematical Sciences and Electrical Engineering, California Institute of Technology, Pasadena, United States. Omid Ardakanian is with Department of Computing Science, University of Alberta, Edmonton, Canada. Claire J. Tomlin is with Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, United States. Correspondence to slow@caltech.edu.

[1]Simulation results can be found in the arxiv version of this paper due to page limit [6]: https://arxiv.org/abs/1610.06631.

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

proposed to estimate the admittance matrix of the underlying power system from the measured power injection of different buses. In [13], the topology of an urban (mesh) distribution network is inferred from voltage magnitude, and real/reactive power measurements carried out by smart meters at the end-nodes. A graphical model is built to describe the probabilistic relationships between different voltage measurements using lasso. In [11] a graphical learning based approach is developed to estimate the radial grid topology from nodal voltage measurements. The learning algorithm is based on conditional independence tests for continuous variables over chordal graphs. An efficient algorithm for topology identification of a power system is also proposed in [9] drawing on ideas from compressive sensing and graph theory. The authors assume that power and phase angle measurements are available for all nodes.

Different algorithms have been developed in the literature to make this inference, using various techniques such as weighted least square, maximum-likelihood/maximum-a-posteriori estimation, minimum spanning tree, sparse recovery, Lasso/Group Lasso, blind identification, quickest change detection theory, as well as graphical model learning. There are three limitations in the current literature that we propose overcome. First, most of the literature focuses on topology identification or change detection, but there is not much work on joint topology and parameter identification, with notable exceptions of [7], [14], [15]. Second, most papers require measurements at every node in the network, with the exceptions of [16], [15], [17], [18], [19], [20]. In particular, [15] learns the topology with parameters from a stochastic perspective, the true topology can only be found in probability, even when the number of samples is large; [18], [19] assume that perturbed data are available (therefore a special inverter is assumed) to identify the network, which could be strong in practice; [20] proposed a method based on recursive grouping to estimate the topology and branch impedance for networks that may have hidden nodes, however, without guarantee.

III. IPF WITHOUT HIDDEN NODES

In this section we study the IPF problem when voltage and current phasor measurements are available at every bus in the system. We formulate the identification problem as a constrained least squares problem and then convert it to an equivalent unconstrained least squares problem. We note that Y has a certain structure that can be exploited when solving the IPF problem— (a) Y is a symmetric but not Hermitian complex matrix (i.e., $Y \in \mathbb{S}^N$) and (b) Y encodes the topology of a connected graph (or a connected tree for radial networks).

A. Problem formulation

Let $\mathbb{C}$ denote the set of complex numbers, $\mathbb{R}$ the set of real numbers, and $\mathbb{N}$ the set of integers. For $A \in \mathbb{C}^{n\times n}$, Re$(A)$ and Im(A) denote matrices with the real and imaginary parts of A, respectively. Let $\mathbb{S}^n \subseteq \mathbb{C}^{n\times n}$ be the set of all $n\times n$ complex symmetric (not necessarily Hermitian) matrices. The transpose of a matrix A is denoted A^T and its Hermitian (complex conjugate) transpose is denoted A^H. $A[i,j]$ denotes the element of A located at ith row and jth column. We define $\mathcal{I}$ as the identity matrix with an appropriate dimension and define $\mathbf{1}$ as an all-1 column vector with an appropriate dimension.

A power system can be modeled by an undirected connected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where $\mathcal{N} := \{1, 2, \dots, N\}$ represents the set of buses, and $\mathcal{E} \subseteq \mathcal{N}\times\mathcal{N}$ represents the set of lines, each connecting two distinct buses. A bus $j \in \mathcal{N}$ can be a load bus, a generator bus, or a swing bus. Let V_j be the complex voltage at bus j and s_j be the net complex power injection (generation minus load) at that bus. We use s_j to denote both the complex number $p_j + \mathbf{i}q_j$ ($\mathbf{i} \triangleq \sqrt{-1}$) and the real pair (p_j, q_j) depending on the context. For each line $(i,j) \in \mathcal{E}$, we denote its series admittance by y_{ij}. The bus admittance matrix of this system is denoted Y, which is an $N\times N$ complex-valued matrix whose off-diagonal elements are $Y_{ij} = -y_{ij}$ and diagonal elements are $Y_{ii} = -\sum_{j\neq i} Y_{ij}$, assuming that there is no shunt element (this assumption can be relaxed). Hence, the current injection vector can be expressed as $I = YV$.

The formulated IPF problem is: given voltage and current measurements of different steady-states, i.e., $V_i(k)$ and $I_i(k)$ for $k = 1, \dots, K$ and $i \in \mathcal{M} = \{1, \dots, N\}$, how to recover the true admittance matrix Y. Specially, when $\mathcal{M}$ is a subset of $\mathcal{N}$, what part of the true admittance matrix Y can be recovered under what condition.

B. Identification algorithm

In this section, we consider the case when $\mathcal{M} = \{1, \dots, N\}$ and propose a solution to the IPF. We will relax this full measurement condition, i.e., when $\mathcal{M} \subset \{1, \dots, N\}$ in the next section.

For $k \in \{1, \dots, K\}$, the Kirchhoff's laws for a given time index yields $I(k) = YV(k)$. Rewriting this formula in vector form for all time indices yields the following equation for a bus i:

$$\underbrace{\begin{bmatrix} I_i(1) \\ I_i(2) \\ \vdots \\ I_i(K) \end{bmatrix}}_{I_i^K} = \underbrace{\begin{bmatrix} V_1(1) & V_2(1) & \dots & V_N(1) \\ V_1(2) & V_2(2) & \dots & V_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ V_1(K) & V_2(K) & \dots & V_N(K) \end{bmatrix}}_{V^K} \underbrace{\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{iN} \end{bmatrix}}_{Y_i}. \quad (1)$$

The admittance matrix Y can be obtained from solving the optimization problem below:

$$\hat{Y}^{K,l_2} \triangleq \arg\min_{Y} \left\| V^K Y - I^K \right\|_F \quad (2)$$
$$\text{s.t.:} \;\; Y \in \mathbb{S}^N, \;\; Y_{ii} = -\sum_{j\neq i} Y_{ij}, \;\; \forall i.$$

in which I^K is a $K\times N$ matrix, i.e., $I^K = \begin{bmatrix} I_1^K & I_2^K & \dots & I_N^K \end{bmatrix}$. Define

$$\text{vec}(Y) = \begin{bmatrix} Y_{11} & Y_{21} & \dots & Y_{N1} & Y_{12} & Y_{22} & \dots & Y_{NN} \end{bmatrix}^T,$$

and apply the *vec* operator to the objective function, we obtain:

$$\min_{\text{vec}(Y)\in\mathbb{C}^{N^2\times 1}} \left\| \left(\mathcal{I} \otimes V^K\right) \text{vec}(Y) - \text{vec}(I^K) \right\|_2 \quad (3)$$
$$\text{s.t.:} \;\; Y \in \mathbb{S}^N, \;\; Y_{ii} = -\sum_{j\neq i} Y_{ij}, \; \forall i,$$

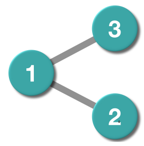


Fig. 1: An illustrative example of a three-node power system with two lines connecting bus 1 to buses 2 and 3 respectively.

where $\otimes$ is the Kronecker product. This holds because $\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$. Let $\text{svec}: \mathbb{S}^N \to \mathbb{C}^{(N^2-N)/2\times 1}$ be a mapping from a symmetric complex matrix to a complex vector defined as:

$$\text{svec}(Y) = [Y_{21} \quad Y_{31} \quad \dots \quad Y_{N1} \quad Y_{32} \quad Y_{42} \quad \dots \quad Y_{NN-1}]^T.$$

It can be readily seen that svec is a bijection for any matrix Y that satisfies a) $Y \in \mathbb{S}^N$ and b) $1^T Y = 0$. Based on this definition, we have $\text{vec}(Y) = \Gamma\text{svec}(Y)$, where $\Gamma \in \mathbb{R}^{N^2\times(N^2-N)/2}$ maps $\text{svec}(Y)$ to the vectorized admittance matrix as illustrated below.

Example 1. *For the network depicted in Figure 1, the Γ matrix has the following form*

$$\underbrace{\begin{bmatrix} Y_{11}\\ Y_{21}\\ Y_{31}\\ Y_{12}\\ Y_{22}\\ Y_{32}\\ Y_{13}\\ Y_{23}\\ Y_{33}\end{bmatrix}}_{vec(Y)} = \underbrace{\begin{bmatrix} -1 & -1 & 0\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 0 & 0\\ -1 & 0 & -1\\ 0 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & -1 & -1\end{bmatrix}}_{\Gamma} \underbrace{\begin{bmatrix} Y_{21}\\ Y_{31}\\ Y_{32}\end{bmatrix}}_{svec(Y)}.$$

Based on the definition of Γ in the above equation, the constrained optimization problem can be converted to an unconstrained one:

$$\min_{\text{svec}(Y)\in\mathbb{C}^{(N^2-N)/2\times 1}} \left\| \underbrace{\left(\mathcal{I}\otimes V^K\right)\Gamma}_{F} \text{svec}(Y) - \text{vec}(I^K) \right\|_2, \tag{4}$$

in which $\mathcal{I}$ denotes an identity matrix and $F \triangleq (\mathcal{I}\otimes V^K)\Gamma$. We define

$$\tilde{F} = \begin{bmatrix} \text{Re}(F) & -\text{Im}(F)\\ \text{Im}(F) & \text{Re}(F)\end{bmatrix}, \text{ and } \tilde{b} = \begin{bmatrix} \text{Re}(\text{vec}(I^K))\\ \text{Im}(\text{vec}(I^K))\end{bmatrix}.$$

The optimization problem (4) can be written as an unconstrained quadratic program in the real domain:

$$\min_{\tilde{f}(Y)\in\mathbb{R}^{(N^2-N)\times 1}} \left\|\tilde{F}\tilde{f}(Y) - \tilde{b}\right\|_2, \tag{5}$$

in which $\tilde{f}(Y) \triangleq [\text{svec}(Y_R)^T \ \text{svec}(Y_I)^T]^T$, $Y_R = \text{Re}(Y)$, and $Y_I = \text{Im}(Y)$. This least square problem yields a solution provided that $\tilde{M}$ has full column rank:

$$\tilde{f}(Y) = \left(\tilde{F}^T\tilde{F}\right)^{-1}\tilde{F}^T\tilde{b}. \tag{6}$$

We compute the solution of the original optimization problem (4) from the solution of the optimization problem (5) by taking the inverse map of $\tilde{f}$. A sufficient condition to guarantee the exactness of the solution is that V^K has full column rank. When V^K does not have full column rank, we can characterize the part of the admittance matrix that is identifiable (see [3]).

Proposition 1 (Exactness)**.** *If V^K has full column rank, the optimization problem (5) has a unique solution given by (6).*

Proof: Since $\Gamma \in \mathbb{R}^{N^2\times(N^2-N)/2}$ and has full column rank (this can be checked easily), there exists a matrix $\Gamma^\dagger$ such that $\Gamma^\dagger\Gamma = \mathcal{I}$. For the Kronecker product $\mathcal{I}\otimes V^K \in \mathbb{C}^{KN\times N^2}$, $\mathcal{I}\otimes V^K$ has full column rank when V^K has full column rank; therefore, $\tilde{F}$ and F have full column rank given the fact that $\text{rank}(\tilde{F}) = 2\text{rank}(F)$.

Finally, we prove by contradiction that if $\tilde{F}$ has full column rank, the solution to the optimization problem (5) is unique. Suppose there exists $\tilde{f}(Y_1)$ and $\tilde{f}(Y_2)$ ($\tilde{f}(Y_1) \neq \tilde{f}(Y_2)$) such that $\tilde{F}\tilde{f}(Y_1) = \tilde{b}$ and $\tilde{F}\tilde{f}(Y_1) = \tilde{b}$, then

$$\tilde{F}\left(\tilde{f}(Y_1) - \tilde{f}(Y_2)\right) = 0,$$

which contradicts the full column rank assumption. ∎

Remark 1. *The approach can be easily extended to the case of nonzero shunt elements where $\mathbf{1}^T Y \neq 0$ by changing the definitions of svec(Y), Γ and $\tilde{f}(\cdot)$. Specifically if Y includes shunt elements then svec(Y) will include diagonal elements $(Y_{11},\dots,Y_{NN})$.*

We can add the element-wise positivity constraint to this problem if the conductance and susceptance of each line are positive[2].

$$\min_{\tilde{f}(Y)\geq 0} \left\|\tilde{F}\tilde{f}(Y) - \tilde{b}\right\|_2. \tag{7}$$

The above problem is known as nonnegative least squares, which is a convex optimization problem and a global minimizer can be solved using different methods, such as the active set method [21].

IV. IPF WITH HIDDEN NODES

In the previous section we solve the IPF problem when voltage and current measurements are available at all buses. In this section we consider the case when voltage and current measurements are available only at a subset of the buses. In a distribution system, for example, measurements are typically available at the substation and customer meters but not throughout the grid.

In Section IV-A, we show that, in the presence of hidden nodes, the algorithm presented in Section III can identify a Kron-reduced admittance matrix $\bar{Y}$, defined in (11) below, for the nodes where measurements are available (Corollary 1). In Section IV-B we show that when the network is a tree then it is indeed possible to uniquely identify the original admittance matrix Y from its Kron reduction under reasonable assumptions.

A. Kron-reduced admittance matrix $\bar{Y}$

We call a bus/node a *measured bus/node* if measurements of its voltage and current injection are available for identification.

[2]The conductance of a line is always positive, the susceptance can be negative or positive depending on whether the line is inductive or capacitive.

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

We call it a *hidden bus/node* otherwise. Let $\mathcal{M}$ and $\mathcal{H}$ represent the set of measured and hidden nodes respectively.

Assumption 1. *We make the following assumptions:*

1.1 *The underlying graph $\mathcal{G}$ is connected.*
1.2 *Hidden nodes have zero injection $I_i(k) = 0$ for all $i \in \mathcal{H}$.*

Note that the second assumption, i.e., the injected current is zero at every hidden node, is reasonable because if generators or loads are connected to a node, then the current they inject or draw is typically measured.

Let H be the number of hidden nodes. Without loss of generality we label the buses so that the first $N - H$ buses are measured and the last H buses are hidden, i.e., $\mathcal{M} := \{1, \ldots, M \triangleq N - H\}$ and $\mathcal{H} := \{N - H + 1, \ldots, N\}$. We partition the bus admittance matrix Y into four sub-matrices:

$$Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} + \mathbf{i} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = G + \mathbf{i}B. \tag{8}$$

Here $Y_{11} \in \mathbb{S}^{N-H}$ describes the connectivity among the measured nodes, $Y_{12} = Y_{21}^T \in \mathbb{C}^{(N-H)\times H}$ the connectivity between the measured and the hidden nodes, and $Y_{22} \in \mathbb{S}^H$ the connectivity among the hidden nodes. For $i \in \mathcal{M}$, $(I_i(k), V_i(k), k = 1, \ldots, K)$ denote the current and voltage measurements at bus i at time k. To simplify notation, we index the entries of Y_{22}, not by $i, j = 1, \ldots, H$, but by $i, j = N - H + 1, \ldots, N$. We index the entries of Y_{12} by $i = 1, \ldots, N - H$ and $j = N - H + 1, \ldots, N$ and similarly for $Y_{21} = Y_{12}^T$, as well as submatrices G_{ij}, B_{ij} in (8).

For $i \in \mathcal{H}$, $I_i(k) = 0$, $\forall k$, and $V_i(k)$ is the voltage at bus i but is not available for identification. To simplify notation, We abuse $V_1(k)$ to denote both the voltage at bus 1 at time k and the vector of all voltages at measured buses at time k, depending on the context; similarly for $V_2(k)$ and $I_1(k)$. Then

$$\begin{bmatrix} I_1(k) \\ 0 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} V_1(k) \\ V_2(k) \end{bmatrix}, \qquad \forall k \tag{9}$$

If Y_{22} is invertible then eliminating V_2 from (9) yields a relation

$$I_1(k) = \bar{Y} V_1(k), \quad \forall k \tag{10}$$

between currents and voltages at measured nodes through the *Kron-reduced admittance matrix* $\bar{Y} \in \mathbb{S}^m$ defined as:

$$\bar{Y} \triangleq Y_{11} - Y_{12} Y_{22}^{-1} Y_{12}^T \tag{11}$$

for the set of measured nodes. In the rest of this subsection we first justify the invertibility of Y_{22} and hence the definition of $\bar{Y}$. Proposition 1 then implies that $\bar{Y}$ can be identified from voltage and current measurements. Moreover $\bar{Y}$ is the best we can identify for general networks because multiple admittance matrices Y may reduce to the same $\bar{Y}$.

Assumption 2. *The admittance matrix $Y \triangleq G + \mathbf{i}B$ defined in (8) satisfies:*

2.1 *Series impedances of the lines are resistive and inductive: $G[i,j] \le 0$ and $B[i,j] \ge 0$ for any $i \ne j$;*
2.2 *Diagonal dominance: $G_{22}[i,i] \ge -\sum_{j\ne i} G_{22}[i,j]$ and $-B_{22}[i,i] \ge \sum_{j \ne i} B_{22}[i,j]$ hold for any i.*

Lemma 1. *Under Assumption 1 and 2, $G_{22} \succeq 0$, $-B_{22} \succeq 0$ and $G_{22} - B_{22} \succ 0$, when $H < N$.*

Proof: From the Gershgorin Theorem and Assumption 2.2, all eigenvalues of the submatrix G_{22} lie in the right-half plane including the origin and all eigenvalues of B_{22} lie in the left-half plane including the origin. Together with the fact that G_{22} and B_{22} are symmetric, we have $G_{22} \succeq 0$ and $-B_{22} \succeq 0$.

This implies that $G_{22} - B_{22} \succeq 0$. We now show that indeed $G_{22} - B_{22} \succ 0$. Suppose for the sake of contradiction that there exists a nonzero $x \in \mathbb{R}^H$ such that $x^T(G_{22} - B_{22})x = 0$. Denote by $A_{22} := G_{22} - B_{22}$ and $A_{21} := G_{21} - B_{21}$ so that

$$A_{22}[i,i] = \sum_{i,j\in\mathcal{H}:j\ne i} (-A_{22}[i,j]) + \sum_{i\in\mathcal{H},j\in\mathcal{M}} (-A_{21}[i,j])$$

Then

$$\begin{aligned} x^T(G_{22} - B_{22})x = &\sum_{i,j\in\mathcal{H}:i\ne j} \left(A_{22}[i,j]x_i x_j - A_{22}[i,j]x_i^2\right) \\ &+ \sum_{i\in\mathcal{H},j\in\mathcal{M}} (-A_{21}[i,j])x_i^2 \\ = &\sum_{i,j\in\mathcal{H}:i<j} (-A_{22}[i,j])(x_i - x_j)^2 \\ &+ \sum_{i\in\mathcal{H},j\in\mathcal{M}} (-A_{21}[i,j])x_i^2 \end{aligned} \tag{12}$$

By definition $A_{22}[i,j] = A_{21}[i,j] = 0$ if $i \not\sim j$ (i.e., i and j are not adjacent). For $i, j \in \mathcal{M} \cup \mathcal{H}$, if $i \sim j$ (i.e., i and j are adjacent), then $Y_{ij} = G_{ij} + \mathbf{i}B_{ij} \ne 0$, i.e., at least one of $G_{ij} = -g_{ij} \le 0$ or $B_{ij} = -b_{ij} \ge 0$ is nonzero. This implies that $-A_{22}[i,j] = -G_{22}[i,j] + B_{22}[i,j] > 0$ and $-A_{21}[i,j] = -G_{21}[i,j] + B_{21}[i,j] > 0$ for all $i \sim j$. Therefore, for $x^T(G_{22} - B_{22})x = 0$ in (12), we must have:

1. $x_i = x_j$ if $i \sim j$ is a connection between hidden nodes in $\mathcal{H}$;
2. $x_i = 0$ for any hidden node $i \in \mathcal{H}$ connected to at least one observed node $j \in \mathcal{M}$.

Since the network is connected, for every hidden node $i \in \mathcal{H}$, there is a path that connects the hidden node to an observed node $k \in \mathcal{M}$. For all nodes j on this path from i to k, the above properties implies that $x_j = 0$. Since this is true for all hidden nodes, we have $x = 0$, a contradiction. Hence $G_{22} - B_{22} \succ 0$. ■

Recall that the network is connected and has N nodes of which H are hidden.

Proposition 2. *Under Assumptions 1 and 2, if $H < N$ then Y_{22} is invertible.*

Proof: We prove that 0 is not an eigenvalue of Y_{22}. Suppose for the sake of contradiction that there exists a nonzero vector $(v + \mathbf{i}w)$ such that $(G_{22} + \mathbf{i}B_{22})(v + \mathbf{i}w) = 0$, i.e.,

$$G_{22}v - B_{22}w = 0 \quad \text{and} \quad G_{22}w + B_{22}v = 0,$$

This implies

$$\begin{aligned} &(G_{22} + B_{22})v + (G_{22} - B_{22})w = 0, \\ &(G_{22} + B_{22})w - (G_{22} - B_{22})v = 0. \end{aligned} \tag{13}$$

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

Since $G_{22} - B_{22} \succ 0$ by Lemma 1, we can eliminate v from (13) to get

$$\left((G_{22} - B_{22}) + (G_{22} + B_{22})(G_{22} - B_{22})^{-1}(G_{22} + B_{22})\right) w = 0.$$

Multiplying w^T on the left we have

$$w^T(G_{22} - B_{22})w \; + \; \hat{w}^T(G_{22} - B_{22})^{-1}\hat{w} \;\; = \;\; 0,$$

where $\hat{w} = (G_{22} + B_{22})w$. This contradicts $G_{22} - B_{22} \succ 0$ unless $w = 0$. But $w = 0$ implies that $(G_{22} - B_{22})v = 0$ from (13), meaning $v = 0$. This is a contradiction and hence Y_{22} is invertible. ∎

Proposition 2 guarantees that Y_{22} is invertible and hence the Kron-reduced admittance $\bar{Y}$ in (11) is well defined under Assumption 2. Moreover, because of (10), the algorithm in Section III can identify the Kron-reduced admittance matrix $\bar{Y}$ from voltage and current measurements (Proposition 1).

Corollary 1. *Suppose Assumptions 1 and 2 and the condition in Proposition 1 hold. If $H < N$, then the Kron-reduced admittance matrix $\bar{Y}$ can be identified from voltage V_1^K and current I_1^K measurements at the measured nodes.*

An admittance matrix $Y \in \mathbb{S}^N$ specifies a unique weighted undirected graph $\mathcal{G}(Y) = (\mathcal{N}(Y), \mathcal{E}(Y))$ with $\mathcal{N} := \{1, \dots, N\}$ and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ such that there is an edge (i, j) if and only if $Y[i, j] \neq 0$. Its Kron reduction $\bar{Y}$ specifies a unique weighted graph $\bar{\mathcal{G}} := \mathcal{G}(\bar{Y}) = (\mathcal{M}, \bar{\mathcal{E}})$ that can be obtained from $\mathcal{G}$ through Algorithm 1.

Algorithm 1 Graph Condensation Algorithm

1: **Input:** a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with N nodes and a set of measured nodes $\mathcal{M} = \{1, 2, \dots, M\}$ and admittance matrix Y
2: **for** $v = M + 1 : N$ **do**
3: Remove hidden node v from $\mathcal{N} = \mathcal{N} - \{v\}$ and all edges from $\mathcal{E}$ that are incident on v;
4: For all node pairs w and l that are neighbors of v, add an edge between w and l to $\mathcal{E}$;
5: Update the admittance matrix $Y = Y/Y[i, i]$ using eq. (15).
6: **end for**
7: **return** $\bar{\mathcal{G}} = \mathcal{G}$ and $\bar{Y} = Y$.

Each iterative step in the algorithm removes a hidden node $i \in \mathcal{H}$ and connects all its neighbors to each other. This step can be represented algebraically as an update on the admittance matrix to compute its Schur complement of $Y[i, i]$. Specifically we can partition possibly after permutation an admittance matrix Y of an appropriate dimension into the form:

$$Y = \begin{bmatrix} Y(i,i) & Y(i,i] \\ Y(i,i]^T & Y[i,i] \end{bmatrix},$$

where $Y[i, i] \in \mathbb{C}$ is the ith diagonal elements of Y, and $Y(i, i), Y(i, i]$ are shown in Eq. (14). Then each iterative step of Algorithm 1 updates the admittance matrix by ($Y[i, i]$ is always invertible due to Proposition 2):

$$Y/Y[i,i] = Y(i,i) - Y(i,i]Y^{-1}[i,i]Y(i,i]^T. \tag{15}$$

This step is repeated until all the hidden nodes are removed from the original graph, producing the Kron-reduced graph $\bar{\mathcal{G}} = \{\mathcal{M}, \bar{\mathcal{E}}\}$ [5] and its admittance matrix $\bar{Y}$ (shown in Fig. 2).

Given an admittance matrix Y, each partition (Y_{11}, Y_{12}, Y_{22}) in (8) defines uniquely a Kron-reduced matrix $\bar{Y} := \bar{Y}(Y_{11}, Y_{12}, Y_{22})$ given by (11). This mapping is clearly not injective in general, i.e., given an $M \times M$ symmetric matrix $\bar{Y} \in \mathbb{S}^M$ (possibly with $\bar{Y}\mathbf{1} = 0$) there are generally multiple $N \times N$ symmetric matrices Y that can be partitioned into (non-unique) (Y_{11}, Y_{12}, Y_{22}) whose Kron reductions are the given $\bar{Y}$, as long as $N \geq M$.

Example 2. *Consider a (Kron-reduced) admittance matrix for a two-node network ($\theta \neq 0$):*

$$\bar{Y} \;\; = \;\; \begin{bmatrix} \theta & -\theta \\ -\theta & \theta \end{bmatrix}.$$

The following 3×3 admittance matrice Y with the given partition has $\bar{Y}$ as its Kron reduction:

$$Y \; := \; \left[\begin{array}{c|c} Y_{11} & Y_{12} \\ \hline Y_{12}^T & Y_{22} \end{array}\right] \; = \; \left[\begin{array}{cc|c} \theta + \theta' & -\theta & -\theta' \\ -\theta & \theta & 0 \\ \hline -\theta' & 0 & \theta' \end{array}\right],$$

for arbitrary nonzero θ'. The underlying network is shown in Fig. 1 with Node 3 as the hidden node, so that $\bar{Y}$ corresponds to the Kron-reduced admittance matrix for Nodes 1 and 2 in Fig. 1. In this case the hidden node has degree 1. Another 3×3 admittance matrices $\tilde{Y}$ that also has $\bar{Y}$ as its Kron reduction is:

$$\tilde{Y} \; := \; \left[\begin{array}{c|c} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \hline \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{array}\right] \; = \; \left[\begin{array}{cc|c} \theta_1 & 0 & -\theta_1 \\ 0 & \theta_2 & -\theta_2 \\ \hline -\theta_1 & -\theta_2 & \theta_1 + \theta_2 \end{array}\right],$$

as long as (θ_1, θ_2) satisfies

$$\frac{\theta_1 \theta_2}{\theta_1 + \theta_2} \;\; = \;\; \theta.$$

The network underlying $\tilde{Y}$ is isomorphic to that in Fig. 1 with Node 1 being the hidden node, so that $\bar{Y}$ is the Kron-reduced admittance matrix for Nodes 2 and 3 in Fig. 1. In this case the hidden node has degree 2.

Example 2 shows that in general only the Kron-reduced admittance matrix $\bar{Y}$ is identifiable from measurements at the measured nodes. For arbitrary networks it is impossible to identify the original admittance matrix Y whose Kron reduction yields $\bar{Y}$. We next show the surprising result that, when the underlying network is a tree and every hidden nodes has a degree ≥ 3, then the original admittance matrix Y can indeed be discovered even in the presence of hidden nodes.

B. Radial networks: exact identification

Consider a radial network and suppose we have identified a Kron-reduced admittance matrix $\bar{Y}$ from partial voltage and current measurements. In this section we develop a novel algorithm to compute the original admittance matrix Y from $\bar{Y}$ under the following additional assumptions.

Assumption 3. *The admittance matrix Y satisfies:*

3.1 *The underlying graph $\mathcal{G}(Y)$ is a tree.*

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

$$Y(i,i) = \begin{bmatrix} Y[1,1] & \dots & Y[1,i-1] & Y[1,i+1] & \dots & Y[1,N] \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Y[i-1,1] & \dots & Y[i-1,i-1] & Y[i-1,i+1] & \dots & Y[i-1,N] \\ Y[i+1,1] & \dots & Y[i+1,i-1] & Y[i+1,i+1] & \dots & Y[i+1,N] \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Y[N,1] & \dots & Y[N,i-1] & Y[N,i+1] & \dots & Y[N,N] \end{bmatrix}, \; Y(i,i] = \begin{bmatrix} Y[1,i] \\ \vdots \\ Y[i-1,i] \\ Y[i+1,i] \\ \dots \\ Y[N,i] \end{bmatrix}. \tag{14}$$

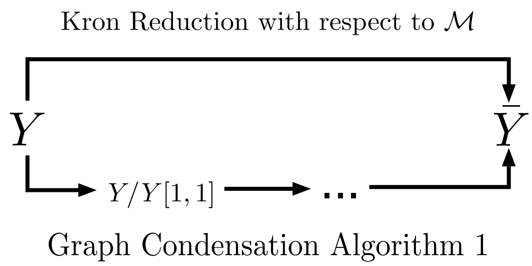


Fig. 2: Two equivalent schemes to compute the Kron reduced $\bar{Y}$.

3.2 *Every hidden node has a degree ≥ 3.*
3.3 *There is no shunt element in Y, i.e., $Y\mathbf{1} = 0$.*

Remark 2. *Assumption 3.2 is a necessary condition for identification as shown in Example 2 where the hidden node has a degree 1 or 2.*

Remark 3. *Assumption 3.3 can be further relaxed as demonstrated in the full version [6].*

We start with some definitions. Consider an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where $\mathcal{N} := \{1, \dots, N\}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of edges. A *complete* graph is one in which all nodes are adjacent. A subgraph of $\mathcal{G}$ is a graph $\mathcal{G}' = (\mathcal{N}', \mathcal{E}')$ with $\mathcal{N}' \subseteq \mathcal{N}$ and $\mathcal{E}' \subseteq \mathcal{E}$. A *clique* of $\mathcal{G}$ is a complete subgraph of $\mathcal{G}$. A *maximal clique* of $\mathcal{G}$ is a clique that is not a subgraph of another clique of $\mathcal{G}$. We say $\mathcal{G}$ is a *tree* if there is exactly one path between every two nodes. A *forest* is a disjoint union of trees.

For our purposes, an admittance matrix Y is a complex symmetric matrix (we usually assume Y satisfies Assumption 3.3). We sometimes refer to $\mathcal{G}(Y)$ as the *underlying graph of* Y and write $\mathcal{G} := (\mathcal{N}, \mathcal{E})$ when Y is clear from the context. Consider two $N \times N$ admittance matrices Y_1 and Y_2. We define two functions of (Y_1, Y_2) and their underlying graphs. First $Y_3 := Y_1 + Y_2$ is also an admittance matrix and its underlying graph $\mathcal{G}(Y_3) = (\mathcal{N}(Y_3), \mathcal{E}(Y_3))$ is the graph with the same set of nodes and edges in both graphs, i.e., $\mathcal{E}(Y_3) := \mathcal{E}(Y_1) \cup \mathcal{E}(Y_2)$. When the matrices are clear from the context, we also denote the function $Y_3 = Y_1 + Y_2$ by $\mathcal{G}_3 = \mathcal{G}_1 \oplus \mathcal{G}_2$. Note that if Y_1 and Y_2 satisfy Assumption 3.3, so does Y_3. Second define the $N \times N$ matrix $Y_4 := Y_1 \backslash Y_2$ as a function of (Y_1, Y_2) by:

$$Y_4[i,j] \;\; := \;\; \begin{cases} Y_1[i,j] & \text{if } i \sim j \text{ and } (i,j) \notin \mathcal{E}_2 \\ -\sum_j Y_4[i,j] & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

The underlying graph $\mathcal{G}(Y_4)$ is a subgraph of $\mathcal{G}(Y_1)$ where edges in $\mathcal{G}(Y_2)$ have been removed. When the matrices are clear from the context, we also denote the function $Y_4 = Y_1 \backslash Y_2$ by $\mathcal{G}_4 = \mathcal{G}_1 / \mathcal{G}_2$. Note that Y_4 satisfies Assumption 3.3 by definition.

Fix an (unknown) admittance matrix Y and assume its underlying graph $\mathcal{G} := \mathcal{G}(Y)$ is a tree. Suppose its Kron-reduced admittance matrix $\bar{Y}$ and its underlying graph $\bar{\mathcal{G}} := \mathcal{G}(\bar{Y})$ are given. For example $\bar{Y}$ is obtained according to Corollary 1 from partial voltage and current measurements at measured nodes in $\mathcal{M}$.

Next, we will propose a recursive algorithm to recover Y from $\bar{Y}$. Specifically, We can decompose $\bar{\mathcal{G}}$ to two graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ ($\bar{Y}_1$ and $\bar{Y}_2$ correspondingly) with distinct properties in Section IV-B1. Secondly, we further introduce a partition of Y in Section IV-B2 and a corresponding parameterization of Y in Section IV-B3. Thirdly, we can compute these parameters from known quantity in $\bar{Y}$ in Section IV-B4. Finally, the overall recursive algorithm to recover Y is proposed in IV-B5.

1) Decomposition of $\bar{\mathcal{G}}$: Let $\bar{\mathcal{E}}_1$ denote the subset of all edges of $\bar{\mathcal{G}}$ that are between measured nodes in the original graph $\mathcal{G}$, and $\bar{\mathcal{E}}_2$ denote the subset of all edges of $\bar{\mathcal{G}}$ that have been added by Step 4 of the graph condensation Algorithm 1 when hidden nodes are removed from $\mathcal{G}$. By definition of $\bar{\mathcal{E}}_1, \bar{\mathcal{E}}_2$, we have $\bar{\mathcal{G}} = (\mathcal{M}, \bar{\mathcal{E}}_1 \cup \bar{\mathcal{E}}_2)$.

Lemma 2. *Under Assumption 1 and 3, $\bar{\mathcal{E}}_1 \cap \bar{\mathcal{E}}_2 = \emptyset$.*

Proof: If there exists an edge $(i,j) \in \bar{\mathcal{E}}_1 \cap \bar{\mathcal{E}}_2$, then (i,j) must be an edge in the original graph $\mathcal{G}$ and nodes i and j must also be connected through a path consisting of only hidden nodes. This creates a loop and contradicts that $\mathcal{G}$ is a tree. Hence $\bar{\mathcal{E}}_1 \cap \bar{\mathcal{E}}_2 = \emptyset$. ■

Since $\bar{\mathcal{G}} = (\mathcal{M}, \bar{\mathcal{E}}_1 \cup \bar{\mathcal{E}}_2)$, Lemma 2 motivates decomposing $\bar{\mathcal{G}}$ into two subgraphs, $\mathcal{G}_1 := (\mathcal{M}, \bar{\mathcal{E}}_1)$ and $\mathcal{G}_2 := (\mathcal{M}, \bar{\mathcal{E}}_2)$, both defined on $\mathcal{M}$ of measured nodes but with disjoint edge sets. While the graph $\bar{\mathcal{G}} := \mathcal{G}(\bar{Y})$ is defined by the Kron-reduced admittance matrix $\bar{Y}$, at this point the graphs $\mathcal{G}_1$, $\mathcal{G}_2$ are only defined in terms of the graph $\bar{\mathcal{G}}$ (in fact in terms of $\mathcal{G}$) and are not associated with any admittance matrices. Define the matrices:

$$\bar{Y}_1 \;\; := \;\; Y_{11} \; - \; \text{diag}\{\mathbf{1}^T Y_{11}\} \tag{16a}$$
$$\bar{Y}_2 \;\; := \;\; \text{diag}\{\mathbf{1}^T Y_{11}\} \; - \; Y_{12} Y_{22}^{-1} Y_{12}^T. \tag{16b}$$

The key observation, stated in the next result, is that $\mathcal{G}_1$, $\mathcal{G}_2$ have simple structures, that the matrices defined in (16) are indeed admittance matrices, and that $\mathcal{G}_1, \mathcal{G}_2$ are the underlying graphs of these admittance matrices. Even though we do not know the submatrices Y_{11}, Y_{12}, Y_{22} of Y, the simple structures of $\mathcal{G}_1$, $\mathcal{G}_2$ allow us to compute $\bar{Y}_1$, $\bar{Y}_2$ as we will see.

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

Theorem 1 (Separability). *Suppose the admittance matrix Y satisfies Assumptions 1, 2 and 3. Then*

1. *$\mathcal{G}_1$ is a forest.*
2. *$\mathcal{G}_2 = \oplus_i \mathcal{C}_i$ for some $\mathcal{C}_i$ are edge-disjoint maximal cliques each with more than 2 nodes.*[3]
3. *$\mathcal{G}_1 = \mathcal{G}(\bar{Y}_1)$ and $\mathcal{G}_2 = \mathcal{G}(\bar{Y}_2)$ so that $\bar{\mathcal{G}} = \mathcal{G}_1 \oplus \mathcal{G}_2$.*

Proof: For the first assertion, $\mathcal{G}_1$ is a forest since it is a subgraph of the tree $\mathcal{G}$. For the second assertion $\mathcal{G}_2$ is a collection of maximal cliques $\mathcal{C}_i$ due to Step 4 of the graph condensation Algorithm 1. To show that the maximal clique (in each) $\mathcal{C}_i$ is of size at least 3, suppose $\mathcal{C}_i$ consists of m_i (measured) nodes and, in the original graph $\mathcal{G}$, these m_i measured nodes "surround" h_i hidden nodes, i.e., the neighbors of each of these hidden nodes are either hidden nodes or nodes in $\mathcal{C}_i$. Let d_j denote the degrees of hidden nodes $j = 1, \ldots, h_i$. These $m_i + h_i$ nodes form a (connected) subtree of $\mathcal{G}$ with exactly $m_i + h_i - 1$ edges. Since m_i of these edges are between measured and hidden nodes and $h_i - 1$ edges are between hidden nodes, we must have $\sum_{j=1}^{h_i} d_j = m_i + 2(h_i - 1)$ and hence $m_i = 2 + \sum_{i=1}^{h_i}(d_i - 2)$. Since $h_i \geq 1$ and $d_i \geq 3$ (Assumption 3.2), we have $m_i \geq 3$. To show that $\mathcal{C}_i$ and $\mathcal{C}_j$ are edge-disjoint, suppose for the sake of contradiction that there is an edge (k, l) in both $\mathcal{C}_i$ and $\mathcal{C}_j$. By the definition of $\mathcal{G}_2$, (k, l) is not an edge in the original graph $\mathcal{G}$. Since nodes k, l are both in $\mathcal{C}_i$, there is a path from k to l in $\mathcal{G}$ that consists of only hidden nodes connected to measured nodes in the maximal clique $\mathcal{C}_i$. Since nodes k, l are both in $\mathcal{C}_j$, there is disjoint path from k to l in $\mathcal{G}$ that consists of a set of hidden nodes connected to nodes in $\mathcal{C}_j$. These two paths form a loop in $\mathcal{G}$, a contradiction. Hence $\mathcal{C}_i$ and $\mathcal{C}_j$ do not share any edge in $\mathcal{G}_2$.

For the third assertion, note that the matrix $\bar{Y}_1$ defined in (16a) is symmetric and hence an admittance matrix. The diagonal entry $Y_{11}[i, i]$ of Y_{11} is the negative sum of the off-diagonal entries of the ith row/column of the original admittance matrix Y (plus any shunt element at bus i), so that the ith entry of $\mathbf{1}^T Y_{11}$ is equal to the ith row sum of Y_{12} (plus any shunt element at bus i). Hence $\bar{Y}_1$ satisfies Assumption 3.3 if Y does. Moreover, by the definition of $\mathcal{G}_1$, the edges in $\mathcal{E}_1$ correspond exactly to the off-diagonal entries of Y_{11} that are nonzero. This implies that the graph $\mathcal{G}(\bar{Y}_1)$ that underlies the admittance matrix in (16a) is indeed $\mathcal{G}_1$.

The matrix $\bar{Y}_2$ defined in (16b) is also symmetric and hence is an admittance matrix. If Y satisfies Assumption 3.3, then

$$\mathbf{1}^T Y_{11} = -\mathbf{1}^T Y_{12}^T, \qquad \mathbf{1}^T Y_{12} = -\mathbf{1}^T Y_{22}.$$

This implies

$$\text{diag}\{\mathbf{1}^T Y_{11}\} = \text{diag}\{\mathbf{1}^T Y_{12} Y_{22}^{-1} Y_{12}^T\},$$

i.e., $\bar{Y}_2$ defined in (16b) satisfies Assumption 3.3 when Y does.

Next we show that $\mathcal{G}_2 = \mathcal{G}(\bar{Y}_2)$. From (16)

$$\bar{Y} = Y_{11} - Y_{12} Y_{22}^{-1} Y_{12}^T = \bar{Y}_1 + \bar{Y}_2,$$

and hence $\bar{\mathcal{G}} := \mathcal{G}(\bar{Y}) = \mathcal{G}_1 \oplus \mathcal{G}_2$ with $\mathcal{G}(\bar{Y}_1) = \mathcal{G}_1$. Therefore we have $\mathcal{G}(\bar{Y}_2) = \mathcal{G}_2$. This concludes the proof. ∎

[3]Strictly speaking, each $\mathcal{C}_i$ is a subgraph of $\mathcal{G}_2$ with $\mathcal{M}$ as its node set. It consists of a single maximal clique and the remaining isolated nodes in $\mathcal{M}$. We will abuse notation and use $\mathcal{C}_i$ to both refer to this subgraph of $\mathcal{G}_2$ and to the maximal clique in $\mathcal{C}_i$.

Remark 4. *From the third assertion, we have shown that, once $\mathcal{G}_1$ and $\mathcal{G}_2$ are obtained from $\mathcal{G}$, $\bar{Y}_1$ and $\bar{Y}_2$ defined in (16) can be obtained.*

There are many algorithms for solving the clique problem, such as the Bron-Kerbosch algorithm, which we adopt in Algorithm 2.

Algorithm 2 Graph Decoupling Algorithm

Input: a condensed graph $\bar{\mathcal{G}}$
Set $\mathcal{G}' = \bar{\mathcal{G}}$, $i = 1$.
while $\mathcal{G}'$ has a clique with more than two nodes **do**
    Use Bron-Kerbosch Algorithm to find a clique (≥ 3 nodes, together with other isolating nodes) $\mathcal{C}_i$ in $\mathcal{G}'$,
    Let $\mathcal{G}' = \mathcal{G}'/\mathcal{C}_i$, $i = i + 1$,
end while
return $\mathcal{G}_2 = \oplus_i \mathcal{C}_i$, $\mathcal{G}_1 = \bar{\mathcal{G}}/\mathcal{G}_2$ and the corresponding $\bar{Y}_1$ and $\bar{Y}_2$.

2) Partition of Y: Next we propose an algorithm to obtain Y_{11}, Y_{22} and Y_{12}, and therefore the original admittance matrix Y.

The decomposition of $\bar{\mathcal{G}}$ into $\mathcal{G}_1$ and $\mathcal{G}_2$ guaranteed by Theorem 1 allows us to partition the set $\mathcal{M}$ into a subset of *internal* measured nodes that are not connected to any hidden nodes and a disjoint subset of *boundary* measured nodes that connect to some hidden nodes. We can similarly partition $\mathcal{H}$ into a subset of *internal* hidden nodes that are not connected to any measured nodes and the disjoint subset of *boundary* hidden nodes that connect to some measured nodes. The decomposition of $\bar{\mathcal{G}}$ into $\mathcal{G}_1$ and $\mathcal{G}_2$ identifies only the types of measured nodes, but not those of hidden nodes. We can hence arrange the original admittance matrix Y into the following structure (only the upper triangular submatrix is shown as Y is symmetric):

$$Y = \left[\begin{array}{c|c} Y_{11} & Y_{12} \\ \hline & Y_{22} \end{array}\right] =: \left[\begin{array}{cc|cc} Y_{11,11} & Y_{11,12} & 0 & 0 \\ & Y_{11,22} & Y_{12,21} & 0 \\ \hline & & Y_{22,11} & Y_{22,12} \\ & & & Y_{22,22} \end{array}\right]. \tag{17a}$$

Here, for Y_{11}, the submatrix $Y_{11,11}$ corresponds to connectivity among the internal measured nodes, $Y_{11,22}$ corresponds to connectivity among the boundary measured nodes, and $Y_{11,12}$ corresponds to connectivity between the internal and boundary measured nodes. Similarly, for Y_{22}, the submatrix $Y_{22,11}$ corresponds to connectivity among the boundary hidden nodes, $Y_{22,22}$ to that among the internal hidden nodes, and $Y_{22,12}$ to that between the internal and boundary hidden nodes. The submatrix $Y_{12,21}$ corresponds to connectivity between the set of boundary measured nodes and the set of boundary hidden nodes. Denote the inverse Y_{22}^{-1} by:

$$X_{22} := Y_{22}^{-1} =: \begin{bmatrix} X_{22,11} & X_{22,12} \\ X_{22,12}^T & X_{22,22} \end{bmatrix}.$$

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

We have

$$\bar{Y} = Y_{11} \;-\; Y_{12}Y_{22}^{-1}Y_{12}^T = \begin{bmatrix} Y_{11,11} & Y_{11,12} \\ Y_{11,12}^T & Y_{11,22} \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & Y_{12,21}X_{22,11}Y_{12,21}^T \end{bmatrix}, \tag{17b}$$

where $X_{22,11} = (Y_{22,11} - Y_{22,12}Y_{22,22}^{-1}Y_{22,12}^T)^{-1}$ from (17a) and the Woodbury formula. Specifically, given the definition of Schur complement

$$\det(Y_{22,11} - Y_{22,12}Y_{22,22}^{-1}Y_{22,12}^T)\det Y_{22,22} = \det Y_{22},$$

and from the invertibility of Y_{22} (shown in Proposition 2), the right-hand side of the above equation is nonzero. Therefore $\det(Y_{22,11}-Y_{22,12}Y_{22,22}^{-1}Y_{22,12}^T)$ cannot be zero, and as a result, the invertibility of $X_{22,11}$ can be guaranteed.

Since we can compute $\bar{Y}$ from partial voltage and current measurements, we can identify submatrices $Y_{11,11}$ and $Y_{11,12}$ for internal measured nodes from $\bar{Y}$ according to (17b). The edges in $\mathcal{E}_1$ correspond to the off-diagonal entries of $[Y_{11,11} \;\; Y_{11,12}]$ as well as $Y_{11,12}^T$, and they form a forest (Theorem 1). The edges in $\mathcal{E}_2$ correspond to the off-diagonal entries of $Y_{11,22} - Y_{12,21}X_{22,11}Y_{12,21}^T$, and they form a collection of cliques. Recall that both $\mathcal{G}_1$ and $\mathcal{G}_2$ have $\mathcal{M}$ as their node set; see the example in Fig. 3.

In the rest of this subsection we focus on identifying the remaining submatrices $Y_{11,22}$, $Y_{12,21}$ as well as Y_{22} (or specifically, $Y_{22,11}$, $Y_{22,12}$, $Y_{22,22}$) of Y. For this purpose we assume without loss of generality that all measured nodes are boundary measured nodes, i.e., the rows and columns corresponding to submatrices $Y_{11,11}$ and $Y_{11,12}$ as well as their contributions to the diagonal entries of $Y_{11,22}$ have been removed from Y. Then

$$Y = \left[\begin{array}{c|c} Y_{11} & Y_{12} \\ \hline & Y_{22} \end{array}\right] =: \left[\begin{array}{c|cc} Y_{11,22} & Y_{12,21} & 0 \\ \hline & Y_{22,11} & Y_{22,12} \\ & & Y_{22,22} \end{array}\right]. \tag{18}$$

Our goal is to identify Y in (18) given its Kron-reduction:

$$\bar{Y} \;=\; Y_{11,22} - Y_{12,21}X_{22,11}Y_{12,21}^T.$$

Theorem 1.2 implies that the underlying Kron-reduce graph $\mathcal{G}(\bar{Y})$ is a disjoint collection of maximal cliques $\mathcal{C}_i$ among boundary measured nodes. By *hidden nodes in a maximal clique* $\mathcal{C}_i$ of the Kron-reduced graph $\bar{\mathcal{G}}$, we mean the nonempty set of hidden nodes in the original graph $\mathcal{G}$ that are connected either to the measured nodes in $\mathcal{C}_i$ or other hidden nodes in $\mathcal{C}_i$. A measured node can be in multiple cliques $\mathcal{C}_i$ though $\mathcal{C}_i$ are edge-disjoint (Theorem 1.2).

Lemma 3. *Suppose the admittance matrix Y satisfies Assumptions 1, 2 and 3. A measured node can connect to only one hidden node in any cliques $\mathcal{C}_i$ of which it is a member.*

Proof: If a measured node connects to more than one hidden node in a maximal cliques $\mathcal{C}_i$, there exists a loop since there is a path between any two hidden nodes in $\mathcal{C}_i$, hence a contradiction. ■

We further assume, without loss of generality, that $\mathcal{G}(\bar{Y})$ consists of a single clique; otherwise, we can repeatedly apply Algorithm 3 below to each clique separately to determine the corresponding submatrices and then combine them to obtain Y_{22} and Y_{12}.

Remark 5. *With this further assumption, Lemma 3 guarantees that Y_{12} has exactly one nonzero element in each row.*

3) Parameterization of Y: Recall that there are M (boundary) measured nodes, indexed by $1,\dots,M$, so that $Y_{11,22}$ in (18) is $M\times M$. Suppose there are H_b boundary hidden nodes, indexed by $M+1,\dots,M+H_b$, and $H_i := H - H_b$ internal hidden nodes, indexed by $M+H_b+1,\dots,M+H$. Then $Y_{22,11}$ in (18) is $H_b\times H_b$ and $Y_{22,22}$ is $H_i\times H_i$. Suppose each measured node $i\in\{1,\dots,M\}$ is connected to the hidden node $h(i)\in\{M+1,\dots,M+H_b\}$ by a line with series admittance $y_{ih(i)}$. From Remark 5 we know there is a unique $h(i)$ for each i, but voltage and current measurements only identify the identity of each measured node i, but not the hidden node $h(i)$ it is connected to (nor the values of H, H_b, H_i). The next result suggests a method to identify all measured nodes that are connected to the same boundary hidden node.

Proposition 3. *Suppose the admittance matrix Y satisfies Assumptions 1, 2 and 3. Two measured nodes i and j are connected to the same hidden node if and only if the off-diagonal entries of rows i and j of $\bar{Y}_2$ are proportional, i.e., there exists $\gamma(i,j)\neq 0$ such that*

$$\frac{\bar{Y}_2[i,k]}{\bar{Y}_2[j,k]} \;=\; \gamma(i,j), \qquad k\neq i,j,\; k=1,\dots,M.$$

Proof: Application of Theorem 1 to the admittance matrix Y in (18) implies that

$$\bar{Y}_2 \;:=\; \text{diag}\{\mathbf{1}^T Y_{11,22}\} \;-\; Y_{12,21}X_{22,11}Y_{12,21}^T. \tag{19}$$

Remark 5 implies row i of $Y_{12,21}$ can be written as $-y_{ih(i)}\,u_{h(i)}^T$ where u_j is the H_b-dimensional column vector with "1" in its jth position and "0" elsewhere. Hence

$$Y_{12,21}X_{22,11}Y_{12,21}^T = \begin{bmatrix} -y_{1h(1)}u_{h(1)}^T \\ \vdots \\ -y_{Mh(M)}\,u_{h(M)}^T \end{bmatrix} \begin{bmatrix} X_{22,11} \end{bmatrix} \begin{bmatrix} -y_{1h(1)}\,u_{h(1)}^T \\ \vdots \\ -y_{Mh(M)}\,u_{h(M)}^T \end{bmatrix}^T.$$

Denote by β_{ij} the (i,j) entry of $X_{22,11}$. Then row i of $Y_{12,21}X_{22,11}Y_{12,21}^T$ is

$$y_{ih(i)}\cdot\left[\beta_{h(i)h(1)}\,y_{1h(1)} \quad \beta_{h(i)h(2)}\,y_{2h(2)} \quad \cdots \quad \beta_{h(i)h(M)}\,y_{Mh(M)}\right].$$

Consider two measured nodes $i,j\in\{1,\dots,M\}$. If they are connected to the same hidden node, then $h(i)=h(j)$ and hence row i and row j of $Y_{12,21}X_{22,11}Y_{12,21}^T$ are proportional:

$$\frac{\left(Y_{12,21}X_{22,11}Y_{12,21}^T\right)[i,k]}{\left(Y_{12,21}X_{22,11}Y_{12,21}^T\right)[j,k]} = \frac{y_{ih(i)}}{y_{jh(i)}} =: \gamma(i,j),\;\; \forall k\neq i,j.$$

The necessity of the proposition then follows from (19).

Conversely, suppose

$$\frac{\left(Y_{12,21}X_{22,11}Y_{12,21}^T\right)[i,k]}{\left(Y_{12,21}X_{22,11}Y_{12,21}^T\right)[j,k]} = \gamma(i,j), \qquad \forall k\neq i,j$$

for some $\gamma(i,j)\neq 0$. Then

$$\frac{y_{ih(i)}\beta_{h(i)h(k)}y_{kh(k)}}{y_{jh(j)}\beta_{h(j)h(k)}y_{kh(k)}} = \frac{y_{ih(i)}\beta_{h(i)h(k)}}{y_{jh(j)}\beta_{h(j)h(k)}} = \gamma(i,j),\;\; \forall k\neq i,j$$

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

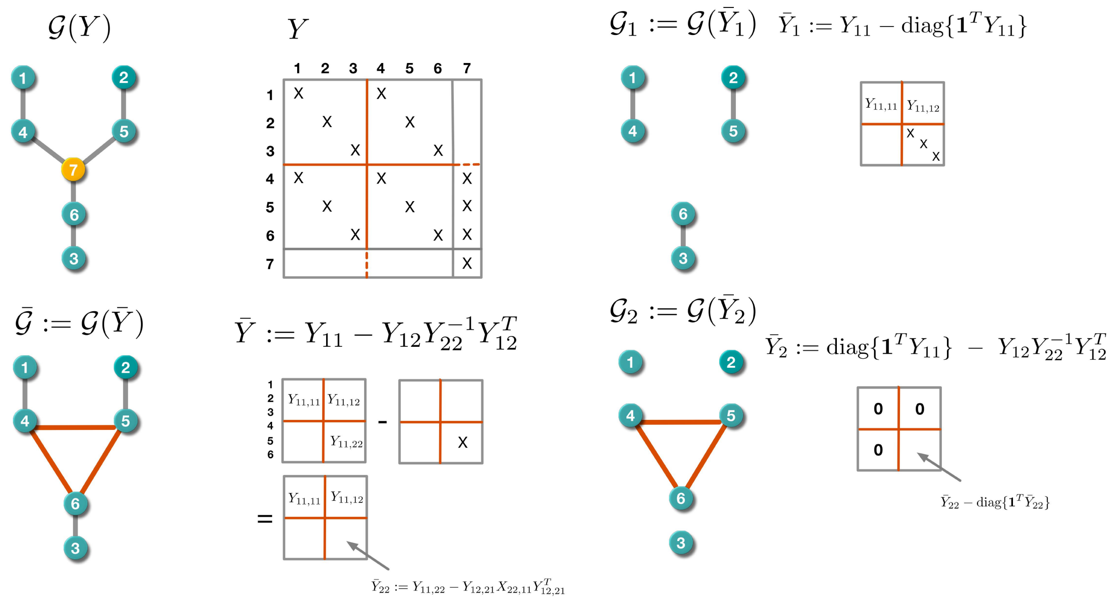


Fig. 3: Illustration of admittance matrices and their underlying graphs: $(Y, \mathcal{G}(Y))$, $(\bar{Y}, \bar{\mathcal{G}})$, $(\bar{Y}_1, \mathcal{G}_1)$, and $(\bar{Y}_2, \mathcal{G}_2)$.

and therefore

$$\frac{\beta_{h(i)h(k)}}{\beta_{h(j)h(k)}} =: \hat{\gamma}(i,j), \quad \forall k \neq i,j$$

Next we show that $(h(k), \forall k \neq i,j)$ span all columns of $X_{22,11}$, which is equivalent to showing

$$\frac{\beta_{h(i)h(j)}}{\beta_{h(j)h(j)}}, \frac{\beta_{h(i)h(i)}}{\beta_{h(j)h(i)}} \in \left\{\frac{\beta_{h(i)k}}{\beta_{h(j)k}}, k = 1, \ldots, H_b\right\}.$$

We prove the sufficiency by contradiction and consider the following three scenarios:

- If every hidden boundary node connects to at least two measured nodes. Therefore, there exist two indices i_1, j_1 such that $h(i_1) = h(i)$ and $h(j_1) = h(j)$. We can let $k = i_1$ and $k = j_2$ respectively and conclude the proof by combining the definition of hidden boundary nodes. This contradicts the invertibility of $X_{22,11}$.
- Secondly, we consider the case that $h(i)$ connects to exact one measured node i and $h(j)$ connects to at least two measured nodes. Without loss of generality, assume that $h(j)$ connects to two measured nodes j and j_1. Following the derivation above and noticing $h(j) = h(j_1)$, we have
$$\frac{\beta_{h(i)h(k)}}{\beta_{h(j)h(k)}} =: \hat{\gamma}(i,j), \qquad \forall k \neq i.$$
Note that $X_{22,11}^{-1} = Y_{22,11} - Y_{22,12}Y_{22,22}^{-1}Y_{22,12}^T$. Consider the $(h(i) - M)$th row of matrix $X_{22,11}^{-1}$, all the elements should be 0 except the edge between $h(i)$ and $h(j)$. This can be shown by computing the adjugate of $X_{22,11}$. This means $h(i)$ only connects $h(j)$ directly or through a path, implying $h(i)$ can maximally connect to one hidden node. On the other side, $h(i)$ only connects to one measured node and therefore this violates Assumption 3.2.
- Finally, if both node $h(i)$ and $h(j)$ only connect to one measured node. We can rearrange the matrix $X_{22,11}$ by combining the $h(i)$th and $h(j)$th rows and the $h(i)$th and $h(j)$th columns in a submatrix as
$$X_{22,11} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$
$$:= \left[\begin{array}{cc|ccc} \beta_{h(i)h(i)} & \beta_{h(i)h(j)} & \beta_{h(i)1} & \cdots & \beta_{h(i)H_b-2} \\ \beta_{h(j)h(i)} & \beta_{h(j)h(j)} & \beta_{h(j)1} & \cdots & \beta_{h(j)H_b-2} \\ \hline \beta_{1,h(i)} & \beta_{1,h(j)} & \beta_{11} & \cdots & \beta_{1,H_b-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_{H_b-2,h(i)} & \beta_{H_b-2,h(j)} & \beta_{H_b-2,1} & \cdots & \beta_{H_b-2,H_b-2} \end{array}\right].$$
A_{11} must be invertible, otherwise it will contradict the invertibility of $X_{22,11}$. Note that $\text{rank}(A_{12}) = 1$. Consider the submatrix $(X_{22,11}^{-1})_{12} = A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}$. Its rank is less or equal to 1. If $\text{rank}((X_{22,11}^{-1})_{12}) = 0$, this means $h(i)$ only connects to one hidden node $h(j)$. Also, $h(i)$ only connects to one measured node and therefore this violates Assumption 3.2. If $\text{rank}((X_{22,11}^{-1})_{12}) = 1$, then $h(i)$ and $h(j)$ connect to at least two identical hidden nodes. Yet it is not hard to see that the graph, in this case, must have a loop, and therefore violates the tree assumption (Assumption 3.1).

Combining all these cases, we have shown, by contradiction, sufficiency. ■

Note that if there are only $M = 3$ measured nodes then Assumptions 3.1 and 3.2 imply that all of them must be connected to the same boundary hidden node.

Given the Kron-reduced admittance matrix $\bar{Y}_2$, Proposition 3 allows us to group together the (boundary) measured nodes that are connected to the same (boundary) hidden node. This also identifies the number of boundary hidden nodes, even though we do not know (yet) the number or identity of internal hidden nodes nor the connectivity among the nodes. We can re-arrange the submatrix matrix $Y_{12,21}$ into a form easier for identification.

Specifically let measured nodes $1, \ldots, k_1$ be connected to hidden node $M+1$, measured nodes $k_1+1, \ldots, k_2$ to hidden node $M+2, \ldots$, measured nodes $k_{H_b-1}+1, \ldots, k_{H_b} := M$ to

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

hidden node $M+H_b$. Note that Proposition 3 yields the values for H_b and $(k_1, k_2, \ldots, k_{H_b} = M)$ even though it provides no information about the value of H, the total number of hidden nodes. To simplify notation, denote the series admittance $y_{ih(i)}$ of line $(i, h(i))$ by y_i. Then $Y_{12} = \begin{bmatrix} Y_{12,21} & 0 \end{bmatrix}$ where $Y_{12,21}$ is $M \times H_b$ and can be arranged into the following simple form:

$$Y_{12,21} = \begin{bmatrix} -y_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ -y_{k_1} & 0 & \cdots & 0 \\ 0 & -y_{k_1+1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & -y_{k_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -y_{k_{H_b}} \end{bmatrix} =: \begin{bmatrix} -\hat{y}_1 & 0 & \cdots & 0 \\ 0 & -\hat{y}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\hat{y}_{H_b} \end{bmatrix},$$

where for $i = 1, \ldots, H_b$, $\hat{y}_i$ is a $(k_i - k_{i-1})$-dimensional column vector corresponding to $k_i - k_{i-1}$ measured nodes that are connected to the hidden node $M+i$. Since Y has zero row sum by Assumption 3.3, the diagonal matrix $\text{diag}\{\mathbf{1}^T Y_{11}\} = \text{diag}\{\mathbf{1}^T Y_{11,22}\} = \text{diag}(y_i = y_{ih(i)}, i = 1, \ldots, M)$. We have

$$\begin{aligned} & Y_{12} Y_{22}^{-1} Y_{12}^T \\ =& \text{diag}(\hat{y}_j)\, X_{22,11}\, \text{diag}(\hat{y}_j^T) \\ =& \begin{bmatrix} \beta_{11}\, \hat{y}_1 \hat{y}_1^T & \beta_{12}\, \hat{y}_1 \hat{y}_2^T & \cdots & \beta_{1H_b}\, \hat{y}_1 \hat{y}_{H_b}^T \\ \beta_{21}\, \hat{y}_2 \hat{y}_1^T & \beta_{22}\, \hat{y}_2 \hat{y}_2^T & \cdots & \beta_{2H_b}\, \hat{y}_2 \hat{y}_{H_b}^T \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{H_b 1}\, \hat{y}_{H_b} \hat{y}_1^T & \beta_{H_b 2}\, \hat{y}_{H_b} \hat{y}_2^T & \cdots & \beta_{H_b H_b}\, \hat{y}_{H_b} \hat{y}_{H_b}^T \end{bmatrix}. \end{aligned}$$

Then the admittance matrix corresponding to the graph $\mathcal{G}_2$ in Theorem 1 is:

$$\begin{aligned} \bar{Y}_2 =& \begin{bmatrix} \text{diag}(\hat{y}_1) & 0 & \cdots & 0 \\ & \text{diag}(\hat{y}_2) & \cdots & 0 \\ & & \ddots & \vdots \\ & & & \text{diag}(\hat{y}_M) \end{bmatrix} \\ & - \begin{bmatrix} \beta_{11}\, \hat{y}_1 \hat{y}_1^T & \beta_{12}\, \hat{y}_1 \hat{y}_2^T & \cdots & \beta_{1H_b}\, \hat{y}_1 \hat{y}_{H_b}^T \\ & \beta_{22}\, \hat{y}_2 \hat{y}_2^T & \cdots & \beta_{2H_b}\, \hat{y}_2 \hat{y}_{H_b}^T \\ & & \ddots & \vdots \\ & & & \beta_{H_b H_b}\, \hat{y}_{H_b} \hat{y}_{H_b}^T \end{bmatrix}. \end{aligned} \tag{20}$$

Recall that we have already identified the Kron-reduced admittance matrix $\bar{Y}_2$, i.e., we know every entry of $\bar{Y}_2$ on the left-hand side of (20). We now explain how to identify $(\beta_{ij}, i, j = 1, \ldots, H_b)$ and $(y_i = y_{ih(i)}, i = 1, \ldots, M)$ on the right-hand side of (20). In particular, $(y_i = y_{ih(i)}, i = 1, \ldots, M)$ yields Y_{12} of the original admittance matrix Y.

4) Computation of parameters in Y: Let $\bar{Y}_{2,k_1}$ be the diagonal submatrix consisting of the first k_1 rows and columns of $\bar{Y}_2$ corresponding to the first k_1 measured nodes connected to the first hidden node $M+1$:

$$\begin{aligned} \bar{Y}_{2,k_1} :=& \text{diag}(\hat{y}_1) \quad - \quad \beta_{11}\, \hat{y}_1 \hat{y}_1^T \\ =& \begin{bmatrix} y_1 & \cdots & 0 \\ & \ddots & \vdots \\ & & y_{k_1} \end{bmatrix} - \beta_{11} \begin{bmatrix} y_1 \\ \vdots \\ y_{k_1} \end{bmatrix} \begin{bmatrix} y_1 & \cdots & y_{k_1} \end{bmatrix}. \end{aligned} \tag{21}$$

We first explain how to identify $(\beta_{11}, \hat{y}_1)$ on the right-hand side of (21) from the knowledge of $\bar{Y}_{2,k_1}$ on the left-hand side of (21). The identification of other $\beta_{ii}, \hat{y}_i$ corresponding to $k_i - k_{i-1}$ measured nodes connected to the hidden node $M+i$ from the diagonal blocks $\bar{Y}_{2,k_i} := \text{diag}(y_{k_{i-1}-1}, \ldots, y_{k_i}) - \beta_{ii}\, \hat{y}_i \hat{y}_i^T$ can be done similarly.

Case 1: $k_1 \geq 2$. In this case, hidden node $M+1$ is connected to two or more measured nodes indexed by $i = 1, \ldots, k_1$. Consider the first two measured nodes and the corresponding 2×2 principal submatrix of Y_{2,k_1}: for $i, j = 1, 2$

$$\bar{Y}_{2,k_1}[i,j] \quad = \quad \begin{cases} y_i - \beta_{11}\, y_i^2 & \text{if} \quad i = j \\ -\beta_{11}\, y_i y_j & \text{if} \quad i \neq j \end{cases} \tag{22}$$

This leads to the following equations in (β_{11}, y_1, y_2):

$$\begin{aligned} y_1 - \beta_{11} y_1^2 &= \bar{Y}_{2,k_1}[1,1] =: a_1 \\ -\beta_{11} y_1 y_2 &= \bar{Y}_{2,k_1}[1,2] =: a_2 \\ y_2 - \beta_{11} y_2^2 &= \bar{Y}_{2,k_1}[2,2] =: a_3 \end{aligned}$$

yielding:

$$\begin{aligned} y_1 &= \frac{a_1 a_3 - a_2^2}{a_2 + a_3}, \quad y_2 = \frac{a_1 a_3 - a_2^2}{a_1 + a_2}, \\ \beta_{11} &= -\frac{a_2 (a_1 + a_2)(a_2 + a_3)}{(a_1 a_3 - a_2^2)^2}. \end{aligned} \tag{23}$$

To identify other $(y_j, j > 2)$, note that

$$-\beta_{11}\, y_1\, y_j \quad = \quad Y_{2,k_1}[1,j], \qquad j = 3, \ldots, k_1$$

yielding

$$y_j \quad = \quad -\frac{Y_{2,k_1}[1,j]}{\beta_{11}\, y_1},$$

where β_{11} and y_1 are given by (23). Once $\hat{y}_1, \ldots, \hat{y}_{k_j}$ are found, we can calculate from off-diagonal entries of $\bar{Y}_2$ all β_{ij} from (20).

Case 2: Once we have recovered the coefficients for hidden boundary nodes with at least two connections to measured nodes in Case 1, next, we can treat these recovered hidden nodes as measured nodes and repeat the above procedure until no hidden node is left. A key step is to construct a new Kron reduced matrix once parts of the admittance matrix have been found. Let the original Y have the following partition as in (18):

$$Y = \left[\begin{array}{c|cc} Y_{11,22} & Y_{12,21} & 0 \\ \hline & Y_{22,11} & Y_{22,12} \\ & & Y_{22,22} \end{array}\right].$$

The Kron reduced admittance matrix can be decomposed to

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

$\bar{Y}_1$ and $\bar{Y}_2$. Specifically, $\bar{Y}_2$ has the following form:

$$\begin{aligned}\bar{Y}_2 =& \operatorname{diag}\{\mathbf{1}^T Y_{11,22}\} - Y_{12,21} X_{22,11} Y_{12,21}^T \\ =& \operatorname{diag}\{\mathbf{1}^T Y_{11,22}\} - \begin{bmatrix} Y_{12,21} & 0 \end{bmatrix} \begin{bmatrix} Y_{22,11} & Y_{22,12} \\ Y_{22,12}^T & Y_{22,22} \end{bmatrix}^{-1} \begin{bmatrix} Y_{12,21}^T \\ 0 \end{bmatrix}\end{aligned}$$

Based on the results in Case 1, one can recover $\operatorname{diag}\{\mathbf{1}^T Y_{11,22}\}$, $Y_{12,21}$ and $X_{22,11}$. Since $\bar{Y}_1$ is known from Algorithm 2, $\operatorname{diag}\{\mathbf{1}^T Y_{11,22}\}$ allows us to compute Y_{11} from the equality (16a) and the partition in (18):

$$Y_{11} = \bar{Y}_1 + \operatorname{diag}\{\mathbf{1}^T Y_{11,22}\}. \tag{24}$$

Hence the entire rows and columns of Y corresponding to the boundary measured nodes are known after (24). We can then focus on the submatrices $Y_{22,11}$, $Y_{22,12}$, $Y_{22,22}$ corresponding to only the boundary and internal hidden nodes, i.e., we can reduce the unknown admittance matrix Y to the new smaller admittance matrix below, which amounts to restricting attention to the subgraph without the boundary measured nodes.

$$Y = \left[\begin{array}{c|c} Y_{22,11} & Y_{22,12} \\ \hline & Y_{22,22} \end{array}\right].$$

The Kron reduced admittance matrix of this new (unknown) admittance matrix Y can then be obtained from the knowledge of $X_{22,11}$:

$$\bar{Y} := Y_{22,11} - Y_{22,12} Y_{22,22}^{-1} Y_{22,12}^T = X_{22,11}^{-1}.$$

Moreover, we have identified the set of boundary hidden nodes. Applying Theorem 1, Algorithm 2 and Proposition 3 to this new $\bar{Y}$ allows us to identify a set of internal hidden nodes to which this set of boundary hidden nodes are connected. Moreover, we can treat the set of boundary hidden nodes as boundary measured nodes and the newly identified internal hidden nodes as boundary hidden nodes. Therefore, even though we do not know the number or the identity of the *remaining* internal hidden nodes, we can partition the new (unknown) admittance matrix Y into the form at the beginning of Case 2 and therefore repeat the computation on this new (smaller) admittance matrix recursively, strictly reducing the number of internal hidden nodes in each iteration until the set of internal hidden nodes becomes null.

Case 3: For any hidden node that connects to one or zero measured node, these hidden nodes will eventually have more than one connection to measured nodes once the other hidden nodes have been recovered and therefore can be recovered. It is easy to show that there will never exist a scenario that all the hidden nodes have at most 1 connection to measured nodes for a tree graph. To see this, note that for any clique, $H \geq M$ as every hidden node connects to a different measured node. On one hand, the sum of all hidden nodes' degrees is greater than $3H$ under Assumption 3. On the other hand, it is at most $2(H-1)+M$, which is twice the sum of all edges between hidden nodes and the number of connections between hidden nodes and measured nodes. However, $2(H-1)+M < 3H$, a contradiction.

Case 4: If all hidden nodes are hidden boundary nodes, i.e., $Y_{22} = Y_{22,11}$, then $Y_{22} = X_{22,11}^{-1}$ and hence the entire admittance matrix Y can be identified. If there are hidden nodes that are not hidden boundary nodes, we can treat hidden boundary nodes as measured nodes now and repeat the above procedure based on Case 2.

5) Overall recursive algorithm: The overall identification procedure is summarized in Algorithm 3[4].

Algorithm 3 Recover Y from $\bar{Y}$

Input: $\bar{Y}_1$ and $\bar{Y}_2$
for each pair of nodes (j,k) **do**
  Compute $\gamma[j,k]$ from $\bar{Y}_2$.
end for
Solve for $\operatorname{diag}\{\mathbf{1}^T Y_{11,22}\}$, $Y_{12,21}$ and $X_{22,11}$ from (21), set $\hat{Y} = \begin{bmatrix} \bar{Y}_1 + \operatorname{diag}\{\mathbf{1}^T Y_{11,22}\} & Y_{12,21} \\ Y_{12,21}^T & X_{22,11}^{-1} \end{bmatrix} := \begin{bmatrix} \hat{Y}_{11} & \hat{Y}_{12} \\ \hat{Y}_{12}^T & \hat{Y}_{22} \end{bmatrix}$ and set $\bar{Y}_2 = X_{22,11}^{-1}$
if the graph corresponding to $\bar{Y}_2$, i.e., $\mathcal{G}(\bar{Y}_2)$ is not radial **then**
  for each pair of nodes (j,k) **do**
    Compute $\gamma[j,k]$ from $\bar{Y}_2$.
  end for
  Solve for $\operatorname{diag}\{\mathbf{1}^T Y_{11,22}\}$, $Y_{12,21}$ and $X_{22,11}$ from (21) and set $\hat{Y} = \begin{bmatrix} \hat{Y}_{11} & \hat{Y}_{12} & 0 \\ \hat{Y}_{12}^T & \operatorname{diag}\{\mathbf{1}^T Y_{11,22}\} & Y_{12,21} \\ 0 & Y_{12,21}^T & X_{22,11}^{-1} \end{bmatrix}$.
  Set
  $$\hat{Y}_{11} = \begin{bmatrix} \hat{Y}_{11} & \hat{Y}_{12} \\ \hat{Y}_{12}^T & \operatorname{diag}\{\mathbf{1}^T Y_{11,22}\} \end{bmatrix}, \ \hat{Y}_{12} = \begin{bmatrix} 0 \\ Y_{12,21} \end{bmatrix}.$$
  Set $\bar{Y} = X_{22,11}^{-1}$ and apply Algorithm 2 to obtain $\bar{Y}_1$ and $\bar{Y}_2$.
end if
return $Y = \hat{Y}$

V. Conclusions

This paper presents a framework for the inverse power flow problem which identifies the admittance matrix of a power system from synchronized voltage and current measurements pertaining to a subset of its buses. The algorithms proposed in this work can identify the graph topology together with its associated admittance matrix with guarantee for radial networks; and it can further jointly address state estimation and topology identification problems with theoretical guarantees, if certain conditions are met. These findings are supported by high-fidelity power flow simulations performed on standard test systems.

In future work, we plan to extend our framework to three phase power flow models, which takes the mutual coupling between phases into account, develop efficient algorithms for identifying the admittance matrix of radial distribution systems with few measurement nodes, and analyze the sensitivity of the identification results to non-stationary measurement errors.

VI. Acknowledgement

We thank Dr. Wei Zhou (Huazhong University of Science and Technology) for useful discussion. Ye Yuan was supported

[4]For notational simplicity, we assume without loss of generality that all measured nodes are boundary measured nodes. Yet, this assumption can easily be relaxed.

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3199084

by the National Natural Science Foundation of China under Grant 92167201.

REFERENCES

[1] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, "A survey on state estimation techniques and challenges in smart distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2312–2322, 2018.

[2] C. Le Floch, S. Bansal, C. J. Tomlin, S. J. Moura, and M. N. Zeilinger, "Plug-and-play model predictive control for load shaping and voltage control in smart grids," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2334–2344, 2017.

[3] O. Ardakanian, V. W. Wong, R. Dobbe, S. H. Low, A. von Meier, C. J. Tomlin, and Y. Yuan, "On identification of distribution grids," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 950–960, 2019.

[4] P. Zhuang, R. Deng, and H. Liang, "False data injection attacks against state estimation in multiphase and unbalanced smart distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6000–6013, 2019.

[5] F. Dorfler and F. Bullo, "Kron Reduction of Graphs With Applications to Electrical Networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, Jan 2013.

[6] Y. Yuan, O. Ardakanian, S. Low, and C. Tomlin, "On the inverse power flow problem," *arXiv preprint arXiv:1610.06631*, 2016.

[7] X. Li, H. V. Poor, and A. Scaglione, "Blind topology identification for power systems," in *Proc. IEEE SmartGridComm*, 2013.

[8] V. Kekatos, G. B. Giannakis, and R. Baldick, "Grid topology identification using electricity prices," in *IEEE PES General Meeting — Conference Exposition*, July 2014, pp. 1–5.

[9] M. Babakmehr, M. G. Simoes, M. B. Wakin, and F. Harirchi, "Compressive Sensing-Based Topology Identification for Smart Grids," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 532–543, April 2016.

[10] Y. Yuan, X. Tang, W. Zhou, W. Pan, X. Li, H.-T. Zhang, H. Ding, and J. Goncalves, "Data driven discovery of cyber physical systems," *Nature communications*, vol. 10, no. 1, pp. 1–9, 2019.

[11] D. Deka, S. Backhaus, and M. Chertkov, "Estimating distribution grid topologies: A graphical learning based approach," in *2016 Power Systems Computation Conference (PSCC)*, 2016, pp. 1–7.

[12] Y. Liao, Y. Weng, M. Wu, and R. Rajagopal, "Distribution grid topology reconstruction: An information theoretic approach," in *North American Power Symposium*, Oct 2015, pp. 1–6.

[13] Y. Liao, Y. Weng, and R. Rajagopal, "Urban Distribution Grid Topology Reconstruction via Lasso," in *IEEE PES General Meeting*, July 2016, pp. 1–6.

[14] J. Yu, Y. Weng, and R. Rajagopal, "Patopa: A data-driven parameter and topology joint estimation framework in distribution grids," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4335–4347, 2017.

[15] S. Park, D. Deka, S. Backhaus, and M. Chertkov, "Learning with end-users in distribution grids: Topology and parameter estimation," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 3, pp. 1428–1440, 2020.

[16] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2215–2224, November 2012.

[17] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning in power distribution networks," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1061–1074, September 2018.

[18] G. Cavraro and V. Kekatos, "Graph algorithms for topology identification using power grid probing," *IEEE control systems letters*, vol. 2, no. 4, pp. 689–694, 2018.

[19] G. Cavraro, V. Kekatos, and S. Veeramachaneni, "Voltage analytics for power distribution network topology verification," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 1058–1067, 2017.

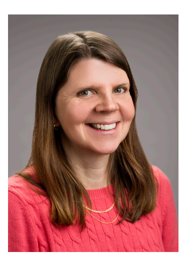
[20] H. Li, Y. Weng, Y. Liao, B. Keel, and K. E. Brown, "Distribution grid impedance & topology estimation with limited or no micro-pmus," *International Journal of Electrical Power & Energy Systems*, vol. 129, p. 106794, 2021.

[21] S. Wright and J. Nocedal, "Numerical optimization," *Springer Science*, vol. 35, pp. 67–68, 1999.

Ye Yuan (M'13, SM'20) received the B.Eng. degree from the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, in 2008, and the M.Phil. and Ph.D. degrees from the Department of Engineering, University of Cambridge, Cambridge, U.K., in 2009 and 2012, respectively. He has been a Full Professor at the Huazhong University of Science and Technology, Wuhan, China since 2016. Prior to that, he was a Postdoctoral Researcher at UC Berkeley, a Junior Research Fellow at Darwin College, University of Cambridge. His research interests include system identification and control with applications to cyber-physical systems.

Steven Low (F'08) is the F. J. Gilloon Professor of the Department of Computing & Mathematical Sciences and the Department of Electrical Engineering at Caltech and an Honorary Professor of the University of Melbourne. Before that, he was with AT&T Bell Laboratories, Murray Hill, NJ, and the University of Melbourne, Australia. He has held honorary/chaired professorship in Australia, China and Taiwan. He was a co-recipient of IEEE best paper awards, an awardee of the IEEE INFOCOM Achievement Award and the ACM SIGMETRICS Test of Time Award, and is a Fellow of IEEE, ACM, and CSEE. He was well-known for work on Internet congestion control and semidefinite relaxation of optimal power flow problems in smart grid. His research on networks has been accelerating more than 1TB of Internet traffic every second since 2014. His research on smart grid is providing large-scale cost effective electric vehicle charging to workplaces. He received his B.S. from Cornell and PhD from Berkeley, both in EE.

Omid Ardakanian (M'15) is an Assistant Professor at the University of Alberta, Canada. He received his B.Sc. from Sharif University of Technology in 2009, and M.Math. and Ph.D. from the University of Waterloo in 2011 and 2015. From 2015 to 2017, he was an NSERC Postdoctoral Fellow at UC Berkeley and the University of British Columbia. He received best paper awards at ACM e-Energy, ACM BuildSys, and IEEE PES General Meeting. His research focuses on the design and implementation of intelligent networked systems.



Claire J. Tomlin (F'10) is the Charles A. Desoer Professor of Engineering in the Department of Electrical Engineering and Computer Sciences (EECS), University of California Berkeley (UC Berkeley). She was an Assistant, Associate, and Full Professor in Aeronautics and Astronautics at Stanford University from 1998 to 2007, and in 2005, she joined UC Berkeley. Claire works in the area of control theory and hybrid systems, with applications to air traffic management, UAV systems, energy, robotics, and systems biology. She is a MacArthur Foundation Fellow (2006), an IEEE Fellow (2010), and in 2017, she was awarded the IEEE Transportation Technologies Award. In 2019, Claire was elected to the National Academy of Engineering and the American Academy of Arts and Sciences.