



Physics informed topology learning in networks of linear dynamical systems[☆]

Saurav Talukdar^a, Deepjyoti Deka^b, Harish Doddi^a, Donatello Materassi^c,
Michael Chertkov^b, Murti V. Salapaka^{a,*}

^a University of Minnesota Twin Cities, Minneapolis, USA

^b Los Alamos National Laboratory, Los Alamos, USA

^c University of Tennessee, Knoxville, USA

ARTICLE INFO

Article history:

Received 30 August 2018

Received in revised form 13 July 2019

Accepted 27 October 2019

Available online 25 November 2019

Keywords:

Networks

Topology learning

Graphical models

Structure learning of time series

ABSTRACT

Learning influence pathways in a network of dynamically related processes from observations is of considerable importance in many disciplines. In this article, influence networks of agents which interact dynamically via linear dependencies are considered. An algorithm for the reconstruction of the topology of interaction based on multivariate Wiener filtering is analyzed. It is shown that for a vast and important class of interactions, that includes physical systems with flow conservation, the topology of the interactions can be *exactly* recovered, even for colored exogenous inputs. The efficacy of the approach is illustrated through simulation and experiments on multiple important networks, including consensus networks, IEEE power networks and EnergyPlus based simulation of thermal dynamics of buildings.

© 2019 Published by Elsevier Ltd.

1. Introduction

Networks are widely used to represent the functioning of complex physical systems like the power grid (Deka, & Chertkov et al., 2017), brain (Bullmore & Sporns, 2009) as well as non-physical systems like social relationships (Scott, 2017), financial systems (Allen & Babus, 2009) and many others. In some applications, there is a clear and evident physical network of agents. For example, consider a network where agents comprise wireless hubs that receive and transmit signals according to a network topology. In other applications, a physical network is not evident,

for example, sensors measuring temperature at various zones in a building, do not admit easy identification of physical links that connect measurements. In either scenario, determining influences between important clusters presents insights into functioning of the complex system. The resulting influence pathways can suggest methods to steer the system toward a desired behavior. *Prior Work:* In many applications it is possible to manipulate the system to help identifying the structure of the network. Such an active approach for identifying the presence of influences between agents may entail removing agents from the network to evaluate their impact on other agents to infer influence pathways (Nabi-Abdolyousefi, 2014). We consider such approaches to be invasive. In Hayden, Chang, Goncalves, and Tomlin (2016) experiment design (or input design) for estimation of the network topology and transfer functions is discussed. Yuan, Stan, Warnick, and Goncalves (2011) present a Akaike Information Criteria (AIC) based invasive approach with a focus on robust network inference in the presence of measurement noise and non-linearities.

In many applications it is not possible to excite/perturb the system to decipher the influence of one variable on the rest. For instance, in the stock market, it is not possible to set the price of a particular stock to evaluate the impact on the other stocks. Thus, there is a clear need of ‘non-invasive’ approaches to infer the network representation of complex systems, which is the approach taken in this article. To be precise, non-invasive approaches do not entail the design of inputs or setting of nodal

[☆] The authors S. Talukdar, H. Doddi, D. Materassi and M.V. Salapaka acknowledge the support of ARPA-E for supporting this research through the project titled ‘A Robust Distributed Framework for Flexible Power Grids’ via Grant No. DEAR000071. Authors D. Deka and M. Chertkov acknowledge the support from the Department of Energy through the Grid Modernization Lab Consortium, and the Center for Non Linear Studies (CNLS) at Los Alamos National Laboratory. The material in this paper was presented at the Eighth ACM International Conference on Future Energy Systems, May 16–19, 2017, Shatin, Hong Kong, China; the 56th IEEE Conference on Decision and Control, December 12–15, 2017, Melbourne, Australia; the 57th IEEE Conference on Decision and Control, December 17–19, 2018, Miami Beach, Florida, USA. This paper was recommended for publication in revised form by Associate Editor Julien M. Hendrickx under the direction of Editor Christos G. Cassandras.

* Corresponding author.

E-mail addresses: sauravtalukdar@umn.edu (S. Talukdar), deepjyoti@lanl.gov (D. Deka), doddi003@umn.edu (H. Doddi), dmateras@utk.edu (D. Materassi), chertkov@lanl.gov (M. Chertkov), murtis@umn.edu (M.V. Salapaka).

states to a reference value for the purpose of network inference; the nodal inputs albeit present are treated as unmeasured disturbances (non-manipulated inputs).

Non-invasive methods of learning the network topology is an active area of research in multiple disciplines including control theory, computer science and statistics to name a few. Here an initial focus utilized modeling activity of agents via random variables; the graphical models approach (see [Friedman, Hastie, & Tibshirani, 2008](#); [Meinshausen & Bühlmann, 2006](#); [Pearl, 1988, 2009](#)) employs the random variable based abstraction extensively. More recently, network topology learning methods have utilized time series measurements of agents' activity modeled as stochastic processes ([Dahlhaus, 2000](#); [Gonçalves & Warnick, 2008](#)). The approach of using stochastic processes to model time-series is well-suited for scenarios where agents interact dynamically; where the past or present state of an agent can affect the present state of another agent. Such interdependence is particularly apt for modeling high resolution time-series measurements, which has become widespread due to the availability of smart sensors with high bandwidth (for example, Phase Measurement Units employed by the power grid ([Nuqui & Phadke, 2005](#))) in many domains including Internet of Things.

In [Giannakis, Shen, and Karanikolas \(2018\)](#), a kernel based regression approach for topology estimation in nonlinear dynamical systems is presented. In [Chiuso and Pillonetto \(2012\)](#) the authors focus on learning the network topology as well as the transfer functions in the high dimensional setting (more nodes and less observations per node) using kernel based estimator with regularization (l_1 type penalty). However, the results presented do not provide consistency guarantees. By consistency guarantees we mean that, given an associated adjacency matrix of the network generating the time series data, 'closeness' of the inferred adjacency matrix obtained using the proposed approach to that of the true underlying adjacency matrix is characterized rigorously.

A convex optimization based approach which endows a sparse adjacency matrix and graph Laplacian matrix under the assumption of weak stationarity of the nodal time series while preserving the spectral properties of the nodal time series is presented in [Egilmez, Pavez, and Ortega \(2017\)](#), [Segarra, and Marques et al. \(2017\)](#) and [Shafipour, Segarra, Marques, and Mateos \(2018\)](#). In [Segarra, and Schaub et al. \(2017\)](#), the ideas from [Segarra, and Marques et al. \(2017\)](#) are applied to linear consensus dynamics with the added requirement of several independent realizations of the multi-agent system being available. However, the results presented do not provide consistency guarantees and require independent and identically distributed samples. The algorithms proposed in this manuscript allow for the presence of correlated samples, provide consistency guarantees and do not have the requirement of independent realizations of the observations.

For consensus dynamics, a decentralized and distributed topology learning scheme is presented in [Kibangou and Commault \(2012\)](#) and [Morbidi and Kibangou \(2014\)](#) respectively. In [Shahrampour and Preciado \(2015\)](#), an active network learning approach based on spectral analysis is used for exact topology inference of Laplacian dynamics. The framework presented in this article is applicable to a linear dynamical system, which includes Laplacian dynamics.

In [Materassi and Salapaka \(2012\)](#) the authors present a multivariate Wiener filtering approach for inferring the network structure of agents that interact via linear dynamical relations. Here, filtering based methods with provable guarantees are devised that can determine the network topology within one hop of a true edge. In this case, no conclusions on causation can be reached (that is the directionality of the dependency cannot be ascertained). However, if the underlying system generating the

data is strictly temporally causal (the present of any process depends on the strict past of others) then causation information (that is direction of the dependency) as well as the network topology can be determined exactly. For nonlinear systems with a known bound on the in degree of a node, the authors in [Quinn, Kiyavash, and Coleman \(2015\)](#) use a directed information approach to infer directed graphs. The assumption of strictly causal dynamics is indeed a significant enabling factor to guarantee exact topology inference with edge directions in [Materassi and Salapaka \(2012\)](#) and [Quinn et al. \(2015\)](#). In [Dahlhaus \(2000\)](#) the authors use a power spectral analysis approach to infer the network topology. [Dahlhaus \(2000\)](#), [Materassi and Salapaka \(2012\)](#), [Quinn et al. \(2015\)](#) primarily use a mix of signal processing or optimization schemes to provide consistency guarantees on the inferred network structure. Without the strict causality assumption the inferred network structure is not guaranteed to be the exact network structure. In this article, we do not impose the requirement of strict causality for exact topology inference.

Our Contribution: We note that the above works do not utilize the governing equations of the system (possibly originating from simplified physics based models) toward passive network topology learning in the presence of colored exogenous inputs. We adopt a gray box approach for inference of the network structure. In this article, we focus on topology learning in physical dynamical systems like power networks ([Chowdhury & Crossley, 2009](#)) and thermal dynamic networks ([Reynders, Diriken, & Saelens, 2014](#)), where influences between agents, when present, are bi-directional, representing a form of mutual coupling and not necessarily a cause effect relationship. In these types of situations, the assumption of strict causality is not valid. We present a new algorithm for exact topology inference, which utilizes both the magnitude and phase response of multivariate Wiener filters for determining the presence/absence of links between two nodes. The main result establishes that the confounding effects of indirect relationship of an agent on another can be detected using the phase response of the multivariate Wiener filters. We provide provable guarantees for consistency of the inferred topology with the underlying topology, without resorting to structural restrictions like the topology being radial ([Talukdar, Deka & Materassi et al., 2017](#)) or bounded in-degree or strictly causal dynamics, and without relying on information of system parameters or exogenous inputs. The guarantees on exactness of network topology inference hold even in the presence of feedback loops unlike [Materassi and Salapaka \(2013\)](#) and [Talukdar, Deka and Lundstrom et al. \(2017\)](#). Moreover, the consistency guarantees are applicable even when the nodal exogenous inputs/noise are colored unlike prior work where assumptions of noise being white are necessary ([Pereira, Ibrahimi, & Montanari, 2010](#)). This article builds on the contributions of [Materassi and Salapaka \(2012\)](#), however, the contribution of this article is an algorithm with guaranteed exact inference of network topology for linear dynamical systems unlike [Materassi and Salapaka \(2012\)](#). We perform a detailed comparison through simulations of our algorithm with the one proposed in [Materassi and Salapaka \(2012\)](#) to illustrate the advantages.

An important application of topology learning is in real time monitoring of a network, which places a requirement of algorithms that can infer the topology with finite/limited samples of measurements from each node. We introduce group Lasso ([Bach, 2008](#)) based regularizers in the Wiener filtering optimization problem and demonstrate that our algorithm when used with regularized version of Wiener filtering provides accurate topology estimation even in the low sample regime. The effectiveness of the algorithms and theory presented is illustrated through simulations on power networks, thermal dynamics of buildings and directed consensus networks.

Preliminary work subsumed by this article instantiated to various application domains have appeared in Talukdar, Deka and Lundstrom et al. (2017) for power grid networks, Doddi, Talukdar, Deka, and Salapaka (2018) for thermal dynamics of buildings and Talukdar, Deka and Attree et al. (2017) for consensus networks. This article is a detailed version with complete proofs with a presentation from a general linear dynamical system perspective. In Section 2 we present the framework, define the problem of topology learning and provide motivating examples. Wiener filtering approach for topology inference is summarized in Section 3 followed by the derivation of the exact topology learning algorithm and its connections with physical conservation laws in Section 4. In Section 5, we illustrate the performance of the learning algorithm on power grid simulations, Energy plus based building simulation and Raspberry Pi based experimental data from consensus dynamics. Conclusions are presented in Section 6.

2. Preliminaries

Suppose x_i , $i = 1, \dots, n$ represent n measured time-series. Further assume that the dynamics generating the time-series x_i satisfy,

$$\sum_{m=1}^l a_{m,i} \frac{d^m x_i}{dt^m} + a_{0,i} x_i = \sum_{j=1, j \neq i}^n b_{ij} x_j(t) + p_i(t), \quad (1)$$

$i \in \{1, 2, \dots, n\}$, where, the exogenous forcing $p_i(t)$ is a zero mean wide sense stationary (WSS) process uncorrelated with $p_j(t)$ for $j \neq i$. Note that a linear transformation of $\{p_j(\cdot)\}_{j=1}^n$ results in $\{x_j(\cdot)\}_{j=1}^n$. We assume that the above dynamics is stable. Thus, $\{x_j(\cdot)\}_{j=1}^n$ are a collection of WSS processes and the collection $\{\{x_j(\cdot)\}_{j=1}^n, \{p_j(\cdot)\}_{j=1}^n\}$ are jointly wide sense stationary processes (JWSS) (Gubner, 2006). Here, $x_i(t) \in \mathbb{R}$, the i th state of the system is assumed to be measured for all i , and, the system parameters are assumed to be real numbers, that is, $a_{m,i} \in \mathbb{R}$, $b_{ij} \in \mathbb{R}$ for all $m \in \{0, 1, \dots, l\}$ and $i, j \in \{1, \dots, n\}$ with $i \neq j$. We remark that none of the system parameters are known. There are many systems which satisfy the dynamics represented by (1), some of which are discussed below.

- (1) Consensus dynamics: Distributed decision making methods in multi-agent systems often use the first order consensus protocol, where each agent updates its states based on the difference between other agents' present value and itself, described by

$$\frac{dx_i}{dt} = \sum_{j=1}^n c_{ij}(x_j - x_i) + p_j, \quad (2)$$

where, p_j denotes the receiver noise for agent j (Siami, Bolouki, Bamieh, & Motee, 2017). Inferring, the communication topology of a network of agents in a multi agent system is a relevant objective of a cyber attacker, and appropriate hardware/ software tools need to be designed to avert such attacks.

- (2) Thermal Resistance Capacitance(RC) networks: Tractable gray box modeling with lumped parameters prove effective for real-time control of buildings. Here RC models that assume discretized physical space are employed where every identified zone is represented by a common temperature. Such a lumped parameter model is described by

$$C_i \frac{dT_i}{dt} = \sum_{j=1}^n \frac{T_j - T_i}{R_{ij}} + p_j(t), \quad (3)$$

where, $C_j > 0$ is the capacitance of zone j , $R_{ij} \geq 0$ denotes the thermal resistance between nodes i and j , p_j

is the total internal heat generated in zone j . In thermal dynamics, $R_{ji} = R_{ij}$ and in the description corresponding to (1) measured variable x_i is identified with temperature T_i . Inference of the thermal RC network aids in developing network models for the thermal dynamics around an operating point.

- (3) Power Grid Network Dynamics: For small disturbances in the power grid, the dynamics of the deviation of voltage phase angle at bus j from the nominal value, denoted by θ_j , is modeled by the following linearized swing equation (Kundur, Balu, & Lauby, 1994),

$$M_i \ddot{\theta}_i + D_i \dot{\theta}_i = \sum_{j=1, j \neq i}^n c_{ij}(\theta_j - \theta_i) + p_i(t), \quad (4)$$

where, $M_i \geq 0$ denotes the inertia of the rotating mass, $D_i \geq 0$ denotes the damping and $p_i(t)$ is the power imbalance injections at the bus i . Here, $c_{ij}(\theta_j - \theta_i)$ gives the line flow from nodes j to i , where c_{ij} is the susceptance of the line. Under equilibrium conditions power balance is satisfied at each node. Note that, $c_{ij} = c_{ji} \geq 0$ for power system dynamics. Inference of the power grid network topology is important as it enables fault detection and optimizing network flows.

The Laplace transform of (1) leads to,

$$\sum_{m=0}^l a_{m,i} s^m X_i(s) = \sum_{j=1, j \neq i}^n b_{ij} X_j(s) + P_i(s), \quad (5)$$

where, $X_i(s)$ and $P_i(s)$ are the Laplace transform of $x_i(t)$ and $p_i(t)$ respectively for $i = \{1, \dots, n\}$. Using the bi-linear transformation of $s = \frac{2(1-z^{-1})}{\Delta t(1+z^{-1})}$ (Δt denotes the sampling time), the z transform of the discretization of continuous time dynamics (1) is given by,

$$S_i(z) X_i(z) = \sum_{j=1, j \neq i}^n b_{ij} X_j(z) + P_i(z), \quad (6)$$

where, $S_i(z) = \sum_{m=0}^l a_{m,i} \left(\frac{2(1-z^{-1})}{\Delta t(1+z^{-1})} \right)^m$, is the z domain operator determined by x_i and its derivatives along with their coefficients in (1). Here, $X_i(z)$ is the z transform of $x_i(k)$ for all $i = \{1, \dots, n\}$. Note that other transformation approaches from the s to z domain could be used as well and our results are independent of the discretization method used. Rewriting (6) we have,

$$X_i(z) = \sum_{j=1, j \neq i}^n H_{ij}(z) X_j(z) + E_i(z) \quad (7)$$

where, $H_{ij}(z) = \frac{b_{ij}}{S_i(z)}$, $E_i(z) = \frac{1}{S_i(z)} P_i(z)$. It can be shown that $e_i(\cdot)$ are uncorrelated with $e_j(\cdot)$ for $j \neq i$, where, $e_i(\cdot)$, is the inverse z transform of $E_i(z)$, and is a zero mean wide sense stationary sequence for all $i = 1, \dots, n$. Summarizing, the network dynamics is represented as,

$$\begin{aligned} X(z) &= H(z)X(z) + E(z), \text{ where } \\ X(z) &= [X_1(z) X_2(z) \dots X_n(z)]^T, \\ E(z) &= [E_1(z) E_2(z) \dots E_n(z)]^T, H(z)(i, j) = H_{ij}(z). \end{aligned} \quad (8)$$

Note that the diagonal entries of the matrix $H(z)$ are 0. Next, we associate a graphical representation derived from the transfer function matrix $H(z)$.

Graphical Representation: Consider a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{1, \dots, n\}$ being the nodes and $\mathcal{E} = \{(i, j) | H_{ij}(z) \neq 0\}$ as the edge set. Each node $i \in \mathcal{V}$ is a representation of the

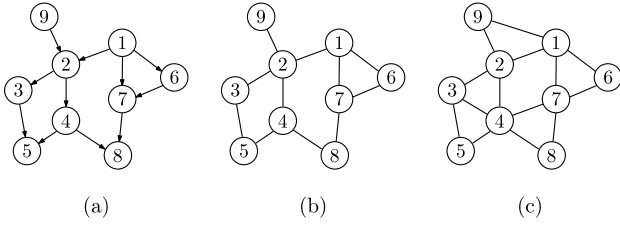


Fig. 1. A generative graph $\mathcal{G}$ is shown in (a), its topology $\mathcal{G}_T$ in (b), and its moral graph, $\mathcal{G}_M$ in (c) (Materassi & Salapaka, 2012).

measured time series $x_i(k)$. We refer to the directed graph $\mathcal{G}$ to be the *generative graph* of the measured time series.

Given a generative graph $\mathcal{G}$, its *topology* is defined as the undirected graph $\mathcal{G}_T = (\mathcal{V}, \mathcal{E}_T)$ obtained by removing the orientation on all its edges, while avoiding repetition. An example of topology of the generative graph in Fig. 1(a) is shown in Fig. 1(b).

Remark 1. Thermal dynamics and power grid dynamics are such that the physics of the system naturally lead to a bi-directed generative graph, where there is no clear notion of cause and effect among the nodes. Many other physical flow dynamics possess similar characteristics.

3. Topology learning

In this article, we address the problem of inferring the underlying topology, $\mathcal{G}_T$, of a linear dynamical system described by (1) with the generative graph $\mathcal{G}$, based on time series measurements $\{x_1(\cdot), \dots, x_n(\cdot)\}$, without the knowledge of the system parameters $\{a_{m,i}, b_{ij}\}$ and with exogenous inputs $\{p_i\}$ unmeasured.

Existing approaches in a dynamical setting with temporally correlated samples, model the exogenous inputs p_i to be Gaussian white noise and independent from p_j for $i \neq j$ (Pereira et al., 2010). In this work we show that, one can infer the exact topology $\mathcal{G}_T$ of the linear dynamical system described by (1), even where exogenous inputs are colored, that is, correlated across time. We begin by recalling the idea of power spectral density.

Definition 1 (Power Spectral Density(PSD) Matrix). For a n dimensional collection of WSS time series $x(\cdot) = \{x_1(\cdot), \dots, x_n(\cdot)\}^T$, the power spectral density matrix is defined as $\Phi_X(\omega) := \sum_{k=-\infty}^{\infty} \mathbb{E}(x(k)x(0)^T) e^{-j\omega k}$, where, $\mathbb{E}(\cdot)$ denotes the expectation operator.

Definition 2 (Cross Power Spectral Density(PSD) Matrix). For a n dimensional collection of WSS time series $X(\cdot) = \{x_1(\cdot), \dots, x_n(\cdot)\}^T$ and a m dimensional collection of WSS time series $Y(\cdot) = \{y_1(\cdot), \dots, y_m(\cdot)\}^T$, the power spectral density matrix is defined as $\Phi_{XY}(\omega) := \sum_{k=-\infty}^{\infty} \mathbb{E}(x(k)y(0)^T) e^{-j\omega k}$, where, $\mathbb{E}(\cdot)$ denotes the expectation operator.

We make the following key assumptions:

- (1) For well posedness we assume that $I - H(z)$ is invertible almost everywhere.
- (2) The power spectral density matrix of the exogenous inputs, $\Phi_P(\omega) = \text{diag}\{\Phi_{p_1}(\omega), \dots, \Phi_{p_n}(\omega)\}$ is assumed to be a positive definite matrix almost everywhere.

3.1. Multivariate Wiener filtering

Let $v(\cdot)$ and $x_1(\cdot), \dots, x_m(\cdot)$ be a collection of jointly wide sense stationary (JWSS) stochastic processes. Let $x(\cdot) := [x_1(\cdot), \dots,$

$x_m(\cdot)]^T$ and $\mathcal{X} := \text{span}\{x_1(k), \dots, x_m(k)\}_{k=-\infty}^{\infty}$. Consider the following least square optimization problem:

$$\hat{v}(k) := \arg \inf_{q \in \mathcal{X}} \mathbb{E}(v(k) - q)^2. \quad (9)$$

If $\Phi_X(\omega) \succ 0$ (that is $\Phi_X(\omega)$ is positive definite) almost surely, then the optimal solution $\hat{v}(k) \in \mathcal{X}$ exists, is unique and is given by

$$\begin{aligned} \hat{V}(z) &= \mathbf{W}(z)X(z), \mathbf{W}(z) = \Phi_{vX}(z)\Phi_X(z)^{-1} \\ &= [W_{v1}(z) \dots W_{vm}(z)]X(z) \\ &= \sum_{j=1}^m W_{vj}(z)X_j(z), \end{aligned}$$

where $\mathbf{W}(z) = \Phi_{vX}(z)\Phi_X(z)^{-1}$ is the Wiener filter. Here, $\hat{V}(z)$ is the z transform of $\hat{v}(\cdot)$. Refer Materassi and Salapaka (2012) for further details. Note that the Wiener filter is obtained using the power spectral density matrices, which can be estimated from the observed time series (Stoica, Moses, et al., 2005). The next result details the properties of Wiener filter for topology inference of linear dynamical systems described by (1) and is based on the results in Materassi and Salapaka (2012).

3.2. Learning the moral graph using multivariate Wiener filtering (Materassi & Salapaka, 2012)

First, we present some graph related notion which will be useful in the subsequent discussion.

In the generative graph $\mathcal{G}$, (i, j) denotes a directed edge from j to i if $H_{ij}(z) \neq 0$, where j is referred as the parent of i and i is referred as the child of j . We will refer to nodes having common children as spouses of each other, for example $i \in \mathcal{V}$ and $j \in \mathcal{V}$ are spouses if there exist a node $k \in \mathcal{V}$ such that $H_{ki}(z) \neq 0$ and $H_{kj}(z) \neq 0$ almost surely. Let C_j, P_j, K_j denote the set of nodes consisting of children, parents and spouses of node j in the generative graph $\mathcal{G}$. For illustration, in Fig. 1 we show a generative graph, where, nodes 1 and 9 are parents of node 2, while, node 2 is the child of nodes 1 and 9. Here, nodes 1 and 9 are spouses of each other.

Definition 3 (Moral Graph). The moral graph, $\mathcal{G}_M = (\mathcal{V}, \mathcal{E}_M)$ of the generative graph $\mathcal{G}$, is defined as the undirected graph obtained by removing the orientation on all its edges, avoiding repetition and adding an undirected edge between spouses. The moral graph of the generative graph in Fig. 1(a) is shown in Fig. 1(c).

Definition 4 (Path). A path between two nodes x_0, x_k in an undirected graph $\mathcal{G}_T = (\mathcal{V}, \mathcal{E}_T)$ is a set of unique nodes $\{x_0, x_1, \dots, x_k\} \subseteq \mathcal{V}$ where $\{(x_0, x_1), \dots, (x_{k-1}, x_k)\} \subseteq \mathcal{E}_T$. We will denote a path by $x_0 - x_1 - x_2 - \dots - x_{k-1} - x_k$. The length of a path is one less than the number of nodes in the path. For example: $1 - 2 - 3 - 5$ is a path of length three, between nodes 1 and 5 in the undirected graph of Fig. 1(b).

Definition 5 (m Hop Neighbor). In an undirected graph $\mathcal{G}_T = (\mathcal{V}, \mathcal{E}_T)$, $j \in \mathcal{V}$ is a m hop neighbor of $i \in \mathcal{V}$, if there is a path of length m between i and j in $\mathcal{G}_T$. For example: nodes 1 and 5 are three hop neighbors in the undirected graph in Fig. 1(b). If there is a path of length one between i and j then they are neighbors in $\mathcal{G}_T$. The set of m hop neighbors of a node $i \in \mathcal{V}$ is denoted by $\mathcal{N}_{i,m}$. The set of neighbors (that is 1 hop neighbor) of a node i in $\mathcal{G}_T$ is denoted as $\mathcal{N}_i$.

Remark 2. The network topology for thermal dynamics and power grid dynamics are such that, if i, j are neighbors in the

topology $\mathcal{G}_T$, then, i is a parent as well as a child of j in the generative graph $\mathcal{G}$. Furthermore, if i, j are two hop neighbors in $\mathcal{G}_T$, then, i and j are spouses in the generative graph. The exact inference of the topology in these physical flow networks is equivalent to exact network inference of the bi-directed generative graph.

Next, we present the main result from [Materassi and Salapaka \(2012\)](#) which relates multivariate Wiener filtering to the network topology and is the basis for Algorithm 1 which recovers the moral graph $\mathcal{G}_M$ of $\mathcal{G}$.

Theorem 1. Consider the linear dynamical system in (1) with topology $\mathcal{G}_T$ and $(H(z), E(z))$ specifying the network dynamics in (8). The nodal state measurements are given by $x(k) = [x_1(k), \dots, x_n(k)]^T$. Define the space $\mathcal{X}_j = \text{span}\{x_i(k)\}_{i \neq j, k=-\infty}^{k=\infty}$. The non-causal Multivariate Wiener filter estimate $\hat{x}_j(k) \in \mathcal{X}_j$ of the signal $x_j(k)$ is given by,

$$\hat{X}_j(z) = \sum_{i \neq j, i=1}^n W_{ji}(z) X_i(z), \quad (10)$$

where, $W_{ji}(z) \neq 0$ implies $i \in \mathcal{N}_j \cup K_j$ in $\mathcal{G}_T$.

Proof. The proof follows from the main result of [Materassi and Salapaka \(2012\)](#), which states that, $W_{ji}(z) \neq 0$ implies that j and i share a children-parent or spouse relationship in the generative graph $\mathcal{G}$. It follows that, in terms of the topology $\mathcal{G}_T$, if $W_{ji}(z) \neq 0$ then i and j are neighbors or spouses.

Remark 3. The above result does not guarantee that if $i \in \mathcal{N}_j \cup K_j$, then $W_{ji}(z) \neq 0$. However, such cases are pathological (see [Materassi and Salapaka \(2012\)](#)).

Thus the set of children, parents and spouses of each node in the network can be identified using non-zero entries in the corresponding multivariate Wiener filter. The moral graph, $\mathcal{G}_M$, can be obtained by adding a link between nodes with non zero entries in the corresponding multivariate Wiener filter. We summarize the procedure to obtain $\mathcal{G}_M$ from nodal time series measurement in Algorithm 1 below.

Algorithm 1 Learning Moral Graph using Wiener Filtering

Input: samples $x_i(k)$ for nodes $i \in \{1, 2, \dots, n\}$ from generative graph $\mathcal{G}$, threshold ρ

Output: Estimate of edges of $\mathcal{G}_M$, $\tilde{\mathcal{E}}_M$ (for large number of samples per node, $\tilde{\mathcal{E}}_M$ coincides with $\mathcal{E}_M$)

```

1: Edge set  $\tilde{\mathcal{E}}_M \leftarrow \{\}$ 
2: for all  $j \in \{1, 2, \dots, n\}$  do
3:   Compute Wiener filter  $W_j(e^{j\omega}) = [W_{j1}(e^{j\omega}) \dots W_{j,j-1}(e^{j\omega}), W_{j,j+1}(e^{j\omega}) \dots W_{jn}(e^{j\omega})]$ 
4: end for
5: for all  $i, j \in \{1, 2, \dots, n\}, i \neq j$  do
6:   if  $\sup_{\omega \in \Omega} \|W_{ji}(e^{j\omega})\| > \rho$  then
7:      $\tilde{\mathcal{E}}_M \leftarrow \tilde{\mathcal{E}}_M \cup \{(i, j)\}$ 
8:   end if
9: end for
```

Algorithm 1 results in the moral graph, which has spurious spouse edges apart from the edges present in the topology. For bi-directed generative graphs considered here, the number of spurious links in the moral graph is of the same order as that of true links; here, for every pair of two hop neighbor we get one edge in the moral graph which is not present in the generative graph. Indeed, consider the power grid topology in

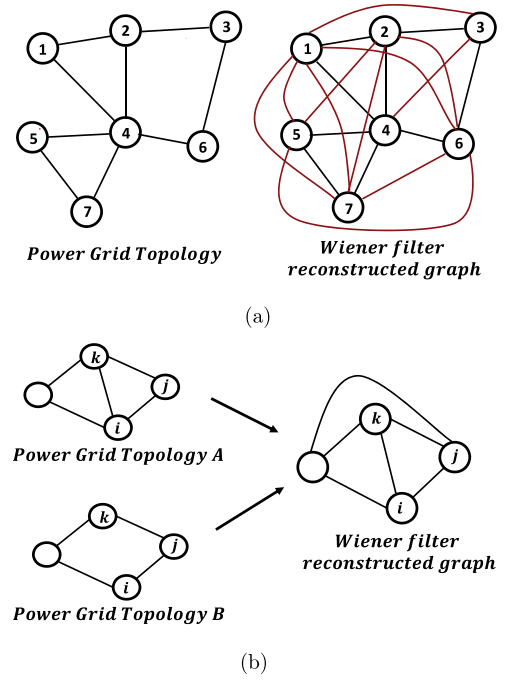


Fig. 2. (a) Number of spurious links (red edges) in the Wiener reconstructed graph are comparable with the number of true links (black edges) in the underlying power grid topology (b) Example of two power grid topologies which result in the same reconstructed moral graph using non-zero Wiener filters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

[Fig. 2\(a\)](#) and its moral graph; it is evident that the number of spurious links (links not present in the topology) is substantial. Furthermore, consider the example in [Fig. 2\(b\)](#), where both topologies A and B obtained from the generative graphs of two power grids, following the swing equation dynamics, result in the same reconstructed moral graph using multivariate Wiener filtering. The Wiener filtering based reconstructed edge between i and k in [Fig. 2](#) can thus represent a true edge or a spurious edge between spouses. Thus, elimination of spurious edges to recover the actual topology is a non trivial and important task owing to the significant presence of spurious links and presence of many candidate topologies for a given moral graph.

4. Exact reconstruction of the generative graph

The following theorem (based on results in [Materassi and Salapaka \(2012\)](#)) presents an explicit characterization of the contribution of neighbors $\mathcal{N}_j$ and two-hop neighbors $\mathcal{N}_{j,2}$ of node j to the multivariate non causal Wiener filter $\{W_{ji}(z)\}_{i=1, i \neq j}^n$.

Theorem 2. Consider the generative graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ described by (1), with $x(k) = (x_1(k) \dots x_n(k))^T$ as the output at time instant k . Let the z transform of the multivariate non causal Wiener filtering estimate $\hat{x}_j(k)$ of $x_j(k)$ be, $\hat{X}_j(z) = \sum_{i, i \neq j} W_{ji}(z) X_i(z)$. Then, $W_{ji}(z) = \hat{C}_{ji}(z) + \hat{P}_{ji}(z) + \hat{K}_{ji}(z)$, where,

$$\hat{C}_{ji}(z) = \frac{b_{ij} S_i(z) \Phi_{p_i}^{-1}(z)}{|S_j(z)|^2 \Phi_{p_j}^{-1}(z) + \sum_{l \in \mathcal{P}_j} b_{lj}^2 \Phi_{p_l}^{-1}(z)}, \quad (11)$$

$$\hat{P}_{ji}(z) = \frac{b_{ji} S_j^*(z) \Phi_{p_j}^{-1}(z)}{|S_j(z)|^2 \Phi_{p_j}^{-1}(z) + \sum_{l \in \mathcal{N}_j} b_{lj}^2 \Phi_{p_l}^{-1}(z)}, \text{ and,} \quad (12)$$

$$\hat{K}_{ji}(z) = -\frac{\sum_{k \in \mathcal{N}_i \cap \mathcal{N}_j} b_{kj} b_{ki} \Phi_{pk}^{-1}(z)}{|S_j(z)|^2 \Phi_{pj}^{-1}(z) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(z)}, \quad (13)$$

where, $S_i(z) = \sum_{m=0}^l a_{m,i} \left(\frac{2(1-z^{-1})}{\Delta t(1+z^{-1})} \right)^m$, $i \in \{1, \dots, n\}$.

Proof. It is shown in [Materassi and Salapaka \(2012\)](#) that, $W_{ji}(z) = \hat{C}_{ji}(z) + \hat{P}_{ji}(z) + \hat{K}_{ji}(z)$, where,

$$\hat{C}_{j*}(z) = \frac{\Phi_{ej}(z) H_{*j}^*(z) \Phi_e^{-1}(z)}{1 + |H_{*j}^*(z) \Phi_e^{-1}(z)| \Phi_{ej}(z)},$$

$$\hat{P}_{j*}(z) = (1 - \hat{C}_{j*}(z) H_{*j}(z)) H_{j*}(z),$$

$$\hat{K}_{ji}(z) = -\hat{C}_{j*}(z) H_{*i}(z),$$

with $H_{*j}(z)$ and $H_{j*}(z)$ representing the j th column and j th row of the $H(z)$ matrix respectively and $H_{*j}^*(z)$ represents complex conjugate transpose of the vector $H_{*j}(z)$. Here $\hat{C}_{ji}$, $\hat{P}_{ji}$ and $\hat{K}_{ji}$ are contributions of the i th time-series in the estimation of the j th time-series for being the child, parent and spouse of j respectively. Note that $\hat{C}_{ji}(z) = 0$ and $\hat{P}_{ji}(z) = 0$. In the context of the generative graph $\mathcal{G}$ with dynamics described by (1), $\hat{C}_{j*}(z)$ is a $1 \times n$ row of transfer functions described by,

$$\begin{aligned} \hat{C}_{j*}(z) &= \frac{\Phi_{ej}(z) H_{*j}^*(z) \Phi_e^{-1}(z)}{1 + |H_{*j}^*(z) \Phi_e^{-1}(z)| \Phi_{ej}(z)} \\ &= \frac{[H_{1j}^*(z) \Phi_{e1}^{-1}(z) H_{2j}^*(z) \Phi_{e2}^{-1}(z) \dots H_{nj}^*(z) \Phi_{en}^{-1}(z)]}{\Phi_{ej}^{-1}(z) + \sum_{l \in P_j} |H_{lj}(z)|^2 \Phi_{el}^{-1}(z)} \\ &= \frac{[\frac{b_{1j}}{S_1^*(z)} |S_1(z)|^2 \Phi_{p1}^{-1}(z) \dots \frac{b_{nj}}{S_n^*(z)} |S_n(z)|^2 \Phi_{pn}^{-1}(z)]}{|S_j(z)|^2 \Phi_{pj}^{-1}(z) + \sum_{l \in P_j} \frac{b_{lj}^2}{|S_l(z)|^2} |S_l(z)|^2 \Phi_{pl}^{-1}(z)} \\ &= \frac{[b_{1j} S_1(z) \Phi_{p1}^{-1}(z) \dots b_{nj} S_n(z) \Phi_{pn}^{-1}(z)]}{|S_j(z)|^2 \Phi_{pj}^{-1}(z) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(z)}. \end{aligned}$$

Thus, contribution to $W_{ji}(z)$ due to i being a child of j is given by,

$$\hat{C}_{ji}(z) = \frac{b_{ij} S_i(z) \Phi_{pi}^{-1}(z)}{|S_j(z)|^2 \Phi_{pj}^{-1}(z) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(z)}. \text{ Similarly, contribution from all}$$

parents of node j in the generative graph $\mathcal{G}$, $\hat{P}_{j*}(z)$ is a $1 \times n$ row of transfer functions described by,

$$\begin{aligned} \hat{P}_{j*}(z) &= (1 - \hat{C}_{j*}(z) H_{*j}(z)) H_{j*}(z) \\ &= (1 - \frac{\sum_{l \in N_j} b_{lj}^2 \Phi_{pl}^{-1}(z)}{|S_j(z)|^2 \Phi_{pj}^{-1}(z) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(z)}) \\ &\quad [H_{j1}(z) \dots H_{jn}(z)] \\ &= \frac{S_j^*(z) \Phi_{pj}^{-1}(z)}{|S_j(z)|^2 \Phi_{pj}^{-1}(z) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(z)} [b_{j1}(z) \dots b_{jn}(z)]. \end{aligned}$$

Thus, the contribution to $W_{ji}(z)$ due to i being a parent of j is

$$\hat{P}_{ji}(z) = \frac{b_{ij} S_j^*(z) \Phi_{pj}^{-1}(z)}{|S_j(z)|^2 \Phi_{pj}^{-1}(z) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(z)}. \text{ The net contribution to } W_{ji}(z)$$

due to $i \in \mathcal{N}_j$, is given by, $\hat{N}_{ji}(z) := \hat{C}_{ji}(z) + \hat{P}_{ji}(z)$. The contribution, $\hat{K}_{ji}$, to W_{ji} for i for being a spouse of j in the generative graph $\mathcal{G}$ with dynamics described by (1) is,

$$\begin{aligned} \hat{K}_{ji}(z) &= -\frac{\sum_{k \in C_j \cap C_i} b_{kj} S_k(z) \frac{b_{ki}}{S_k^*(z)} \Phi_{pk}^{-1}(z)}{|S_j(z)|^2 \Phi_{pj}^{-1}(z) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(z)} \\ &= -\frac{\sum_{k \in C_j \cap C_i} b_{kj} b_{ki} \Phi_{pk}^{-1}(z)}{|S_j(z)|^2 \Phi_{pj}^{-1}(z) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(z)}. \end{aligned} \quad (14)$$

This proves the theorem.

Next, we use the expressions derived in [Theorem 2](#) to distinguish between true and spurious edges in the moral graph $\mathcal{G}_M$ formed by non-zero entries of multivariate Wiener filter. The next theorem presents a result using the phase response of the non causal Wiener filter for spurious edges corresponding to strict spouse relationships, enabling them to be distinguished from true edges.

Theorem 3. Consider the generative graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ described by (1), with $x(k) = (x_1(k) \dots x_n(k))^T$ as the output at time instant k . If i and j are strict spouses in $\mathcal{G}$, that is, $C_i \cap C_j \neq \emptyset$ and $i \notin \mathcal{N}_j$, $j \notin \mathcal{N}_i$, then $\angle W_{ji}(e^{j\omega})$ is 0 or π for all $\omega \in [-\pi, \pi]$.

Proof. As $i \notin \mathcal{N}_j$ and $j \notin \mathcal{N}_i$, $b_{ij} = 0$, $b_{ji} = 0$. It follows from [Theorem 2](#) that, $\hat{C}_{ji}(z) = 0$ and $\hat{P}_{ji}(z) = 0$. Since, i and j are spouses, there exist $k \in C_i \cap C_j$ such that $b_{kj} \neq 0$ and $b_{ki} \neq 0$, implying $\hat{K}_{ji}(z) \neq 0$ (using (14)). Since, $\hat{K}_{ji}(e^{j\omega})$ is dependent on $\{\Phi_{pj}(e^{j\omega})\}_{j=1,2,\dots,n}$, which are positive real numbers for WSS processes ([Charbit, 2010](#)), it follows from (14) that, $\hat{K}_{ji}(e^{j\omega}) \in \mathbb{R}$ (no imaginary part) for all $\omega \in [-\pi, \pi]$. Thus, $\angle W_{ji}(e^{j\omega}) = \angle \hat{K}_{ji}(e^{j\omega}) = 0$ or π , for all $\omega \in [-\pi, \pi]$.

Remark 4. If $b_{ij} > 0$ whenever $(i, j) \in \mathcal{E}$ (true for power grid, thermal dynamics and consensus applications), then, $\angle(W_{ji}(e^{j\omega})) = \pi$ for all $\omega \in [-\pi, \pi]$ if the conditions in [Theorem 3](#) are satisfied.

We will use the above theorem to identify spurious edges between spouses in the moral graph $\mathcal{G}_M$ to recover the true topology $\mathcal{G}_T$. We now show that the above result does not hold for nodes that are neighbors. The next theorem lists a condition under which $\angle(W_{ji}(e^{j\omega})) = \pi$ or 0 for all $\omega \in [-\pi, \pi]$ for nodes i and j that are neighbors as well as spouses.

Theorem 4. Consider the generative graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ described by (1), with $x(k) = (x_1(k) \dots x_n(k))^T$ as the output at time instant k . Suppose i and j are such that $i \in \mathcal{N}_j$, $i \in K_j$. Then $\angle(W_{ji}(e^{j\omega})) = 0$ or π for all $\omega \in [-\pi, \pi]$, if and only if,

$$\text{Im}(b_{ij} S_i(e^{j\omega}) \Phi_{pi}^{-1}(e^{j\omega}) + b_{ji} S_j^*(e^{j\omega}) \Phi_{pj}^{-1}(e^{j\omega})) = 0,$$

for all $\omega \in [-\pi, \pi]$, where $\text{Im}(z)$ represent the imaginary part of the complex number z .

Proof. Using (11)–(13),

$$\begin{aligned} W_{ji}(e^{j\omega}) &= \frac{b_{ij} S_i(e^{j\omega}) \Phi_{pi}^{-1}(e^{j\omega}) + b_{ji} S_j^*(e^{j\omega}) \Phi_{pj}^{-1}(e^{j\omega})}{|S_j(e^{j\omega})|^2 \Phi_{pj}^{-1}(e^{j\omega}) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(e^{j\omega})} \\ &\quad - \frac{\sum_{k \in \mathcal{N}_i \cap \mathcal{N}_j} b_{kj} b_{ki} \Phi_{pk}^{-1}(e^{j\omega})}{|S_j(e^{j\omega})|^2 \Phi_{pj}^{-1}(e^{j\omega}) + \sum_{l \in P_j} b_{lj}^2 \Phi_{pl}^{-1}(e^{j\omega})} \end{aligned} \quad (15)$$

The second term in the above expression has no imaginary component and the denominators of both the terms are real.

($\Rightarrow$) Since, $\angle(W_{ji}(e^{j\omega})) = 0$ or π for all $\omega \in [-\pi, \pi]$, it follows that $(W_{ji}(e^{j\omega}))$ is a real number for all $\omega \in [-\pi, \pi]$. Thus, for all $\omega \in [-\pi, \pi]$, $\text{Im}(b_{ij} S_i(e^{j\omega}) \Phi_{pi}^{-1}(e^{j\omega}) + b_{ji} S_j^*(e^{j\omega}) \Phi_{pj}^{-1}(e^{j\omega})) = 0$.

($\Leftarrow$) If $\text{Im}(b_{ij} S_i(e^{j\omega}) \Phi_{pi}^{-1}(e^{j\omega}) + b_{ji} S_j^*(e^{j\omega}) \Phi_{pj}^{-1}(e^{j\omega})) = 0$ for all $\omega \in [-\pi, \pi]$, then, $\text{Im}(W_{ji}(e^{j\omega})) = 0$ for all $\omega \in [-\pi, \pi]$, which leads to $\angle(W_{ji}(e^{j\omega})) = 0$ or π for all $\omega \in [-\pi, \pi]$.

The next theorem lists a condition under which $\angle(W_{ji}(e^{j\omega})) = \pi$ or 0 for all $\omega \in [-\pi, \pi]$ for nodes i and j that are neighbors and not spouses.

Theorem 5. Consider the generative graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ described by (1), with $X(k) = (x_1(k) \dots x_n(k))^T$ as the output at time instant k . Nodes i and j are such that $i \in \mathcal{N}_j$ and $i \notin K_j$. Then, $\angle(W_{ji}(\hat{e}^{j\omega})) = 0$ or π for all $\omega \in [-\pi, \pi]$ if and only if $\text{Im}(b_{ij}S_i(\hat{e}^{j\omega})\Phi_{p_i}^{-1}(\hat{e}^{j\omega}) + b_{ji}S_j^*(\hat{e}^{j\omega})\Phi_{p_j}^{-1}(\hat{e}^{j\omega})) = 0$ for all $\omega \in [-\pi, \pi]$.

Proof. Since $i \notin K_j$, $\hat{K}_{ji}(z) = 0$. It follows that,

$$W_{ji}(\hat{e}^{j\omega}) = \hat{N}_{ji}(z) = \frac{b_{ij}S_i(\hat{e}^{j\omega})\Phi_{p_i}^{-1}(\hat{e}^{j\omega}) + b_{ji}S_j^*(\hat{e}^{j\omega})\Phi_{p_j}^{-1}(\hat{e}^{j\omega})}{|S_j(\hat{e}^{j\omega})|^2\Phi_{p_j}^{-1}(\hat{e}^{j\omega}) + \sum_{l \in \mathcal{N}_j} b_{lj}^2\Phi_{p_l}^{-1}(\hat{e}^{j\omega})}. \quad (16)$$

Using reasoning similar to the proof of Theorem 4 the rest of the proof follows.

Remark 5. The condition $\text{Im}(b_{ij}S_i(\hat{e}^{j\omega})\Phi_{p_i}^{-1}(\hat{e}^{j\omega}) + b_{ji}S_j^*(\hat{e}^{j\omega})\Phi_{p_j}^{-1}(\hat{e}^{j\omega})) = 0$ for all $\omega \in [-\pi, \pi]$, presented in Theorems 4 and 5, is to be satisfied for all $\omega \in [-\pi, \pi]$, such that for neighbor nodes i and j (irrespective of whether i and j are also spouses or not), for $\angle(W_{ji}(\hat{e}^{j\omega})) = 0$ or π $\omega \in [-\pi, \pi]$. This is an equality constraint on the space of system parameters and is a set of measure zero in the space of system parameter values. We refer to parameters in this measure zero set as pathological parameters. An example of pathological parameter set includes the case of static systems, that is, $\{a_{m,i} = 0\}_{m \in \{1, \dots, l\}, i \in \{1, \dots, n\}}$, implying $\text{Im}(S_i(\hat{e}^{j\omega})) = 0$ for all $i \in \{1, \dots, n\}$, $\omega \in [-\pi, \pi]$, thus satisfying the equality constraint.

Theorems 3–5 demonstrate that aside from pathological cases, the phase of the Wiener filter $\angle(W_{ji}(\hat{e}^{j\omega})) = 0$ or π for all $\omega \in [-\pi, \pi]$ only when i and j are **not** neighbors but are spouses in the generative graph $\mathcal{G}$. In particular, for such i and j , if the prior knowledge of $b_{ji} > 0$ exists then $\angle(W_{ji}(\hat{e}^{j\omega})) = \pi$ for all $\omega \in [-\pi, \pi]$. In other words, aside for the pathological cases as discussed in Remark 5, the converse of Theorem 3 holds and can be used as a criterion to differentiate between true edges and spurious edges in the moral graph $\mathcal{G}_M$.

We now present Algorithm 2 that estimates the topology of the generative graph $\mathcal{G}$ based on time-series of nodal measurements pertaining to dynamics expressed by (1). The algorithm consists of two parts. The first part (Steps 2–9) determines the multivariate Wiener filter $W_{ji}(z)$ to estimate the moral graph and is same as Algorithm 1. In the next part (Steps 10–16), we consider a representative set of frequency points Ω in the interval $[-\pi, \pi)$ and evaluate the phase angle of the Wiener filters for edges in $\mathcal{E}^w$. If the phase angle is within a pre-defined threshold τ of 0 or π , the algorithm designates them as spurious edges (see Theorem 3) and prunes them from $\mathcal{E}^w$ to produce edge set $\bar{\mathcal{E}}$ of the estimated true topology.

5. Simulation and experimental validation

In this section we present simulation and experimental validation of the topology inference algorithm presented in the previous section. First, we present the methods employed to compute the Wiener filter using nodal time series measurements, which is used in Algorithm 2.

Wiener filter computation: Let $x_i(k)$ be the nodal time series to be estimated from $\{x_i(k)\}_{i=1, i \neq j}^m$, $k \in \mathbb{Z}$. Consider the following least square optimization problem on the Hilbert space of $\mathcal{L}_2$ random variables,

$$\{h_{ji,o}\} = \arg \inf_{\{h_{ji}\}_{i=1, \dots, n, i \neq j}}$$

Algorithm 2 Topology Learning using Wiener Filtering with Pruning Step

Input: nodal time samples $x_i(k)$ for nodes $i \in \{1, 2, \dots, n\}$ in the generative graph $\mathcal{G}$, thresholds ρ, τ , frequency points $\Omega = \{\omega_1, \dots, \omega_m\}$ where $\omega_i \in [-\pi, \pi]$

Output: Estimate of Edges $\bar{\mathcal{E}}_T$ in the topology of $\mathcal{G}$. For large samples, $\bar{\mathcal{E}}_T$ coincides with $\mathcal{E}_T$

```

1: Edge set  $\mathcal{E}^w \leftarrow \{\}$ 
2: for all  $j \in \{1, 2, \dots, n\}$  do
3:   Compute Wiener filter  $W_j(\hat{e}^{j\omega}) = [W_{j1}(\hat{e}^{j\omega}) \dots W_{jn}(\hat{e}^{j\omega})]$ 
4: end for
5: for all  $i, j \in \{1, 2, \dots, n\}$ ,  $i \neq j$  do
6:   if  $\sup_{\omega_i \in \Omega} \|W_{ji}(\hat{e}^{j\omega_i})\| > \rho$  then
7:      $\mathcal{E}^w \leftarrow \mathcal{E}^w \cup \{(i, j)\}$ 
8:   end if
9: end for
10: Edge set  $\bar{\mathcal{E}}_T \leftarrow \mathcal{E}^w$ 
11: for all  $i, j \in \{1, 2, \dots, n\}$ ,  $i \neq j$  do
12:   if  $\pi - \tau \leq \angle W_{ji}(\hat{e}^{j\omega_i}) \leq \pi + \tau$  or
13:    $-\tau \leq \angle W_{ji}(\hat{e}^{j\omega_i}) \leq \tau, \forall \omega_i \in \Omega$  then
14:      $\bar{\mathcal{E}}_T \leftarrow \bar{\mathcal{E}}_T - \{(i, j)\}$ 
15:   end if
16: end for
```

$$\mathbb{E}(x_j(k) - \sum_{i=1, i \neq j}^n \sum_{L=-\infty}^{\infty} h_{ji}^L x_i(k-L))^2, \quad (17)$$

$$\approx \arg \inf_{\{h_{ji}\}_{i=1, \dots, n, i \neq j}} \frac{1}{N} \sum_{k=1}^N (x_j(k) - \sum_{i=1, i \neq j}^n \sum_{L=-F}^F h_{ji}^L x_i(k-L))^2, \quad (18)$$

where, $h_{ji} = [h_{ji}^{-F}, \dots, h_{ji}^0, \dots, h_{ji}^F]$ and N is the length of each observed time series. The above approximation follows from the mean ergodic theorem (Reed & Simon, 1980) and is exact in the limit of $N \rightarrow \infty$ and $F \rightarrow \infty$. From the implementation viewpoint, we consider lags up to a finite order F in (9) and a finite number of observations N . The solution to the above optimization problem is referred as the *finite impulse response multivariate non-causal Wiener filter* (Huang, Benesty, & Chen, 2006), $W_j(z) = [W_{j,1}(z), \dots, W_{j,j-1}(z), W_{j,j+1}(z), \dots, W_{j,m}(z)]$, where,

$$W_{ji}(z) = \sum_{L=-F}^F h_{ji,o}^L z^{-L}. \quad (19)$$

It is important to note that the multivariate Wiener filter, which is obtained from solving the above optimization problem is determined entirely from the measured time series without any knowledge of system parameters or the statistics of the exogenous inputs.

The above approach to Wiener filter computations gives accurate Wiener filter estimation when the length of the observed time series is ‘long enough’ (sufficient statistics or asymptotic observation regime). It results in poor accuracy of Wiener filter estimation in the limited number of observations and large network application (high dimensional statistics regime) due to over-fitting (Hastie, Tibshirani, & Wainwright, 2015).

Wiener filter with regularization: In order to account for limited samples of measurements at each node, a regularized version of multivariate Wiener filtering is obtained by minimizing the objective function in (18) along with the group Lasso regularization

penalty (Hastie et al., 2015).

$$\begin{aligned} \{h_{ji}, \gamma\} &= \arg \inf_{\{h_{ji}\}_{i=1, \dots, n, i \neq j}} \\ \mathbb{E}(x_j(k) - \sum_{i=1, i \neq j}^n \sum_{L=-\infty}^{\infty} h_{ji}^L x_i(k-L))^2 + \gamma \sum_{i=1, i \neq j}^n \|h_{ji}\|_2, \\ &\approx \arg \inf_{\{h_{ji}\}_{i=1, \dots, n, i \neq j}} \\ \frac{1}{N} \sum_{k=0}^N (x_j(k) - \sum_{i=1, i \neq j}^m \sum_{L=-F}^F h_{ji}^L x_i(k-L))^2 + \gamma \sum_{i=1, i \neq j}^n \|h_{ji}\|_2. \end{aligned} \quad (20)$$

Here $\gamma \geq 0$ is the regularization parameter and $h_{ji} := [h_{ji}^{-F}, \dots, h_{ji}^0, \dots, h_{ji}^F]$. The regularized Wiener filter is given as,

$$W_{ji}^\gamma(z) = \sum_{L=-F}^F h_{ji, \gamma}^L z^{-L}. \quad (21)$$

Choice of the threshold ρ , τ and regularization parameter γ : In practice, the choice of the threshold is very crucial as the data samples observed for each node is finite. Let $\Omega = \{\omega_1, \dots, \omega_s\}$ be a set of frequency points in $[0, \pi]$ rad/s. For the Wiener projection of time series of node j on $i \in \{1, \dots, n\}_{i \neq j}$, we compute, $s_{ji} := \frac{\sup_{\omega_k \in \Omega} \|W_{ji}(e^{j\omega_k})\|}{\max_{i \in \{1, \dots, n\}_{i \neq j}} \sup_{\omega_k \in \Omega} \|W_{ji}(e^{j\omega_k})\|}$ for every node $i \in \{1, \dots, n\}_{i \neq j}$. Note that the maximum value of s_{ji} is 1. The j th row of the adjacency matrix A of size $n \times n$ of the inferred topology is populated as follows:

$$A(j, i) = \begin{cases} 1, & \text{if } s_{ji} \geq \rho, i \neq j \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

For pruning step in the Algorithm 2, $|W_{ji}(e^{j\omega_i})|$ is computed for each of the frequency points in the set Ω and their average is computed. If the averaged value $\in (\pi - \tau, \pi + \tau)$, then the edge $j - i$ is inferred to be spurious. In the examples discussed below $\Omega = \{0, 0.1, 0.2, 0.3, \dots, 0.9\} \times \pi$.

Relative Error: Let the adjacency matrix of the true topology be denoted by A_T . The relative error percentage is defined as

$$\frac{\sum_{r=1}^{r=n} \sum_{t=1}^{t=n} \sum_{i \neq r} |A_T(r, t) - A(r, t)|}{\sum_{r=1}^{r=n} \sum_{t=1}^{t=n} \sum_{i \neq r} A_T(r, t)} \times 100.$$

It is ratio of total number of false edges inferred from observations (both true negatives and false positives) to that of the total number of edges in network generating the observations. Next, we present validation of Algorithm 2 for exact network topology inference through several examples. We would like to mention that in all the example cases discussed below, in the large sample limit, Algorithm 2 is able to detect all true edges in the underlying topological and no true link is pruned out as spurious, thereby reinforcing that pathological cases are rare.

5.1. Reconstruction of RC network structure from voltage time series with colored input noise

Here, we perform a comparative study of Algorithm 2 with the approach discussed in Algorithm 1 (Materassi & Salapaka, 2012) and the one lag regression approach in Pereira et al. (2010) using nodal voltage time series from the representative 20 node RC network as shown in Fig. 3(a) and its topology is shown in Fig. 3(b). Consider a node i with a neighbor j in the RC network. The node capacitance, C_i , exogenous input, p_i , and the resistance between i and j , R_{ij} , are chosen as,

$$R_{ij} = \frac{\log(i+j)}{i+j}, C_i = \frac{i^2}{100},$$

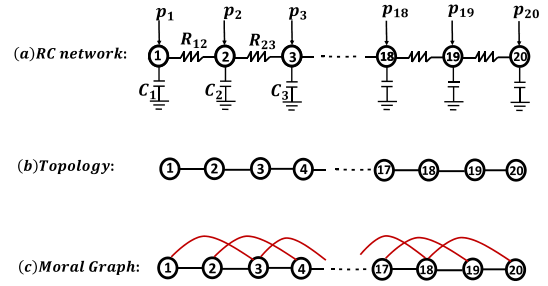


Fig. 3. RC network, its Moral graph and Topology. Here, p_j is the input current at node j . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

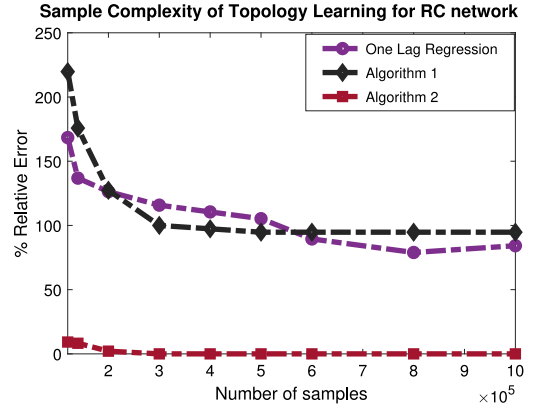


Fig. 4. Error percentage variation with number of samples per node for Algorithm 1, Algorithm 2 and one lag regression when inputs are colored.

$$T(z) = \frac{1 - 0.2z^{-1} - 0.3z^{-2}}{1 - 0.5z^{-1} + 0.4z^{-2} - 0.1z^{-3}}$$

$p_i = T(z)$ filtered 10 dB white noise.

The network dynamics is simulated in Matlab and the voltage time series obtained is used for inference of the unknown network topology. The thresholds for Algorithm 1 and 2 are chosen as mentioned below:

- Algorithm 1: $\rho = 0.1$
- Algorithm 2: $\rho = 0.1, \tau = 10^{-1}$

The topology inference accuracy in terms of the relative error using Algorithm 1, Algorithm 2 and the one lag regression approach is shown in Fig. 4. Applying Algorithm 1 (in the large sample limit, 6×10^5 samples per node onward) results in the moral graph as shown in 3(c), with 19 true edges (black), 18 spurious edges (red), and it is not known which of these inferred edges is true or spurious. It is seen in Fig. 4 that even in the high sample limit (5×10^5 samples per node onward) the errors in topology inference using the one lag regression approach of Pereira et al. (2010) and Algorithm 1 (Materassi & Salapaka, 2012) are significantly high, whereas Algorithm 2 results in zero inference errors from at least 2×10^5 samples per node. Also, it is interesting to note that in the low sample regime (10^4 samples per node), the topology inference errors using Algorithm 2 are less than 10%, which is significantly less than Algorithm 1 and one lag regression.

5.2. Comparative study and validation on energy plus simulations

Here, we will estimate the interaction topology amongst zonal temperatures in a building using Algorithm 2 and also highlight

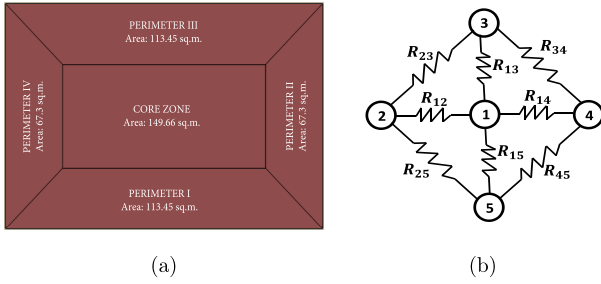


Fig. 5. (a) Topview of EnergyPlus building model consisting of 5 zones, where, the height of the building is 3.05 m, (b) Thermal resistor network of the building with thermal capacitance at each node derived from a CFD model (Radecki & Hency, 2012).

the limitations in inference of three node cycles using static approaches. Here, we illustrate that by applying Algorithm 2 to the temperature data of a five zone office building as shown in 5(a), the true thermal RC network topology is recovered exactly. The building envelope is created in Google SketchUp Make 2017 (Elis, Torcellini, & Crawley, 2008) (see Fig. 5) and its RC network model obtained using model reduction of CFD model (Radecki & Hency, 2012) are shown in Fig. 5(b). The building is located in Minneapolis, Minnesota, USA and the weather file used in EnergyPlus is obtained from <https://energyplus.net/weather>. The exogenous inputs to the building are heat gains from lights, electrical equipment and people. For our simulations, we consider the electrical and lighting loads as time-correlated wide sense stationary processes. The correlated inputs are generated by filtering white Gaussian noise through a first order digital filter in MATLAB. It is to be emphasized that EnergyPlus solves the nonlinear energy balance equation to obtain the zone temperatures, where the heat transfer coefficients are functions of zone temperatures (see DoE, 2010). The temperature time series data at 1 min granularity is generated using EnergyPlus (Crawley, Pedersen, Lawrie, & Winkelmann, 2000) and is used in topology learning algorithms to estimate the structure of the underlying RC network. For more details on EnergyPlus see DoE (2010).

Here, we perform a comparative study of four algorithms: graphical lasso (Friedman et al., 2008) (random variable framework for Gaussian graphical models), graphical lasso with sign based pruning (Bolognani, Bof, Michelotti, Muraro, & Schenato, 2013; Deka, & Talukdar et al., 2017), one lag regression approach (Pereira et al., 2010) (see Appendix A for details), Algorithm 2 with Wiener filtering computation without (18) and with regularizer (20). The relative errors for topology inference of the thermal dynamics of the building described above using the above mentioned algorithms are shown in Fig. 6. It is seen that in the low sample regime (around 10^4 samples), the error in inference by Algorithm 2 (with regularization) reduces by about half (in comparison to Algorithm 2 without regularization) due to use of regularization in Wiener filter computations. It is seen that, the graphical lasso as well as the graphical lasso with pruning step is unable to recover the exact topology of the underlying RC network, even in the large sample limit. This is a limitation of the random variable framework for networks with three node undirected cycles, which was also highlighted in the introduction of this article. Similarly, the one lag regression approach of Pereira et al. (2010) is unable to recover the exact topology of the RC network. However, the time series based approach of Algorithm 2 with and without regularizer has zero error in topology inference in the large sample limit. It is to be noted that the inferred topology matches with the RC network topology obtained by model reduction of a CFD model. It is seen in Fig. 6 that the relative error

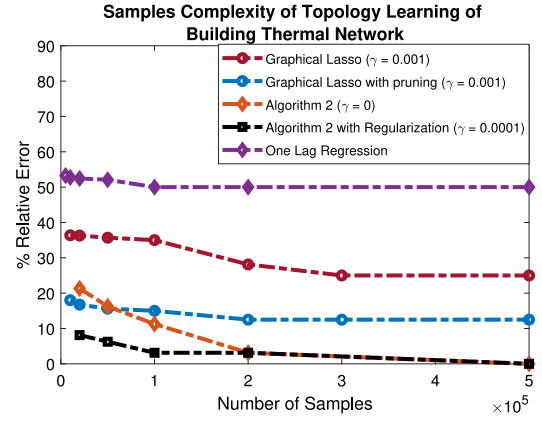


Fig. 6. Relative error percentage variation with number of samples per node for graphical lasso, graphical lasso with pruning, one lag regression and Algorithm 2 (with and without regularizers) with colored exogenous inputs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

for 10^4 samples per node is about 20% for Algorithm 2 with no regularizers while it is less than 10% for topology inference using Algorithm 2 with regularized Wiener filtering. Thus, the use of regularizers could significantly reduce the over-fitting errors in topology inference in the low sample regime.

5.3. Validation on consensus dynamics

Here, the performance of Algorithm 2 in estimating the topology of a network of 5 agents undergoing consensus iterations as depicted in Fig. 7(a) is presented. We present two sets of results, first based on MATLAB simulations, and second based on experimental results on a network of Raspberry Pis. The details of the experimental platform can be found in Prakash, Talukdar, Attree, Yadav, and Salapaka (2016). For our simulations, the receiver noise at each node is considered to be white. The trends from the nodal measurement time series are removed using the 'detrend' function in MATLAB. The reconstructed topology using Algorithm 1 with 10^7 samples from each node is shown in Fig. 7(c), where, the dashed edges denote the spurious links recovered. Fig. 8(a) and Fig. 8(b) show the frequency response of Wiener filters between node 2 and all other nodes that are derived using Algorithm 2 with 10^7 samples for each node. It is clear from Fig. 8(a) that the magnitude of the filter W_{25} is small across the frequency range and it is concluded that there exist no edge between nodes 2 and 5 in the moral graph. Using the pruning step, the absolute values of the phase response of $W_{21}(z)$, $W_{23}(z)$, $W_{24}(z)$ are analyzed as shown in Fig. 8(b). Here, the phase of the filter W_{24} remains close to π throughout the frequency range. It follows from Theorem 3 that the edge between 2 and 4 is spurious (phase response being close to π). The relative error percentage of Algorithm 2 with respect to number of samples for simulations as well as experiments is shown in Fig. 9. As expected, it is seen that as the number of samples per node increases, the error decreases.

5.4. Validation on power network

In this section, we demonstrate the effectiveness of Algorithm 2 on the IEEE 39 and IEEE 14 bus power network (Athay, Podmore, & Virmani, 1979) shown in Figs. 10(a) and 12 respectively with network dynamics as described by (4). Here p_i are modeled as filtered white Gaussian noise (colored noise unlike iid Gaussian in Deka, Backhaus, and Chertkov (2016)) to generate time series

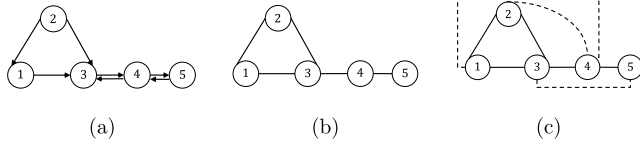


Fig. 7. (a) Generative graph of 5 node network undergoing consensus dynamics, (b) associated network topology, (c) reconstructed topology of the 5 node network of Fig. 7(a) obtained using multivariate Wiener filtering with 10^7 samples from each node. The dashed links are the spurious links due to spouse relationship.

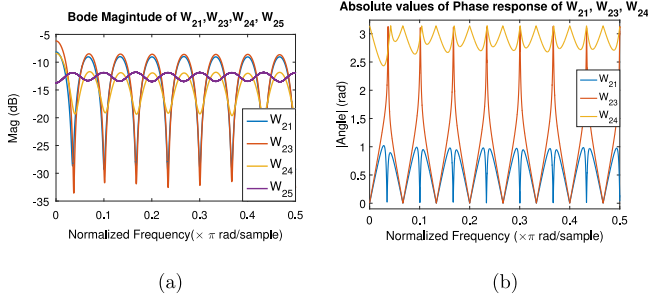


Fig. 8. (a) Bode magnitude plot of $W_{21}(z)$, $W_{23}(z)$, $W_{24}(z)$, $W_{25}(z)$ (b) absolute values of phase response of $W_{21}(z)$, $W_{23}(z)$, $W_{24}(z)$, $W_{25}(z)$. $W_{25}(z)$ has a phase response in the vicinity of π for all frequencies, hence, is eliminated by the pruning step.

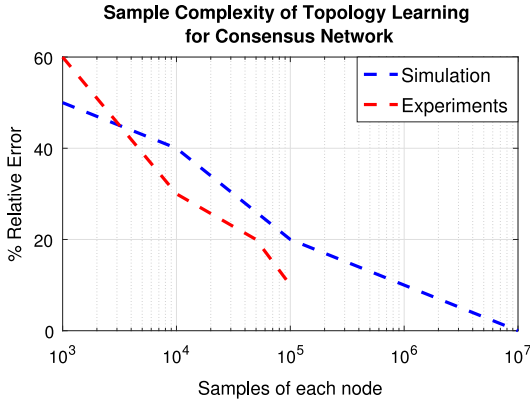


Fig. 9. Relative error percentage variation with number of samples per node in simulation as well as experiments.

data for evaluation of the proposed algorithm. The output at each node is sampled at 0.01s.

Consider the neighbors (green) and two-hop neighbors (red) of node 25 in the IEEE 39 bus system as shown in Fig. 10(b). Application of steps 1 to 9 of Algorithm 2 with a threshold ρ of 10^{-5} in step 6 and 6.5×10^6 samples per node leads to edges between node 25 and all nodes in Fig. 10(b). The absolute values of the phase response of the multivariate Wiener filters for node 25 and the nodes in its two-hop neighborhood are shown in Fig. 11(a). It is seen that the phase response of the Wiener filters corresponding to the nodes two hops away is close to π rad, while that of the neighbor nodes starts from 0 rad. Thus using the pruning step, all the two-hop neighbors can be removed, recovering the true physical topology. The relative error percentage in topology estimation for the IEEE 39 bus system as a function of sample size is shown in Fig. 11(b). The threshold ρ was chosen as 10^{-3} while $\tau = 0.2\pi$. In many cases, pruning step eliminated 58% of the total false positive edges obtained after step 9 of Algorithm 2.

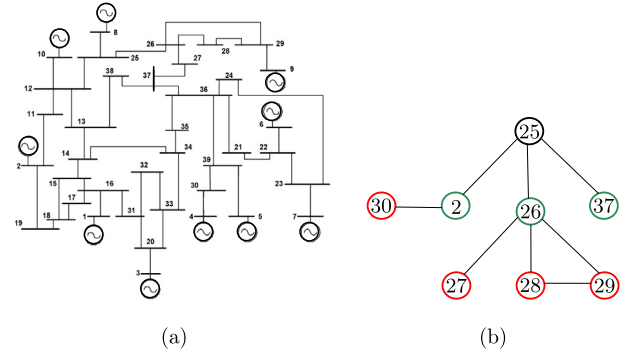


Fig. 10. (a) IEEE 39 bus system with generators at 10 buses (Athay et al., 1979) (b) The neighbors (green nodes) and strict two-hop neighbors (red nodes) of node 25 in the IEEE 39 bus system. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

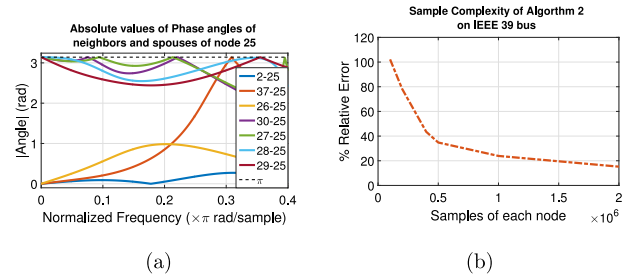


Fig. 11. (a) Absolute values of the phase response of the Wiener filters between node 25 and its two-hop neighborhood in the IEEE 39 bus system. The phase response begins from 0 rad for all three neighbors and from π for all spouses of node 25. (b) Relative error percentage of Algorithm 2 with samples per node for IEEE 39 bus systems.

Finally, we compare the performance of Algorithm 2 with and without regularizer for inference of the topology of the IEEE 14 bus system, using time series data from the simulation of the corresponding set of swing equations with colored exogenous inputs. As seen in Fig. 12, there are 4 three node cycles in the underlying topology. In a three node cycle, neighbors are also spouses. The performance of Algorithm 2 with and without regularizer is shown in Fig. 13, where it is seen that in the large sample limit both approaches converge to zero inference errors. However, in the low sample limit ($5e^4$ samples per node, about 10 min of data collected using a Phasor Measurement Unit (PMU)), the errors in topology inference without regularizer is about 40% while with regularizers it is about 7%. A typical grid undergoes topology changes about 5–10 times a day (Cavaro & Arghandeh, 2018), and hence, Algorithm 2 with appropriate regularizer is indeed applicable for monitoring of power grid network topology.

6. Conclusion

In this article, we presented an algorithm for exact topology inference of linear dynamical systems with a particular focus on physical flow networks. We utilized the phase response of multivariate Wiener filters to distinguish between true and spurious edges. The algorithm proposed in this article can handle exogenous noise that is colored, unlike some of the recent work on topology inference from time series measurements where noise is assumed to be white. Moreover, the proposed algorithm guarantees exact topology inference even in the presence of loops in the network. It is worth noting that, the proposed algorithm does not require any knowledge on statistics of the exogenous

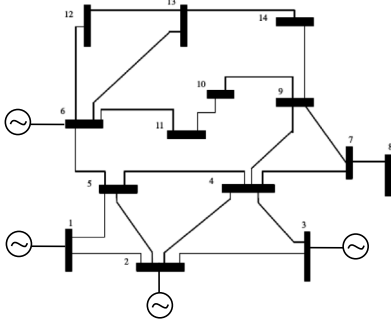


Fig. 12. IEEE 14 bus system with generators at 4 buses.

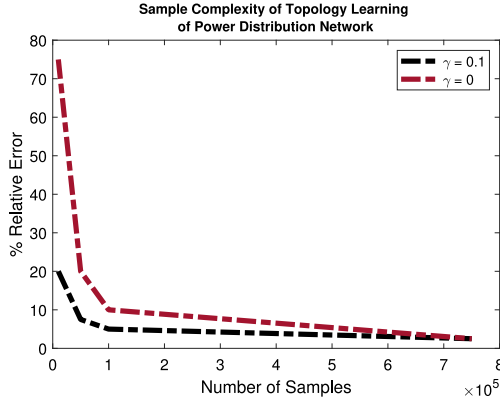


Fig. 13. Topology identification for Power network: Relative error percentage variation with number of samples per node, with and without regularization.

inputs or system parameters for topology inference. We believe that using phase response of filters to eliminate spurious links in topology learning can also be extended to other network topology learning approaches like directed information and spectral analysis, as connections of these approaches with the Wiener filtering approach are well understood (Etesami & Kiyavash, 2014; Materassi & Salapaka, 2012). The article, also highlighted the shortcomings of state of the art random variables approaches in dealing with networked dynamical systems, particularly in inferring three node cycles and the superiority of the algorithm proposed in such scenarios. We also illustrated the benefit of using regularizers in Wiener filter computations, for obtaining better topology inference performance in the low sample regime. This is particularly important for deploying the proposed algorithm for topology inference from finite size data windows of measurements across the network.

Appendix. One lag regression (Pereira et al., 2010)

$$\begin{aligned}
 \{R_{ji}, \gamma\} &= \arg \inf_{\{R_{ji}\}_{i=1, \dots, m, i \neq j}} \\
 \mathbb{E}[x_j(k) - \sum_{i=1, i \neq j}^n R_{ji} x_i(k-1)]^2 + \gamma \sum_{i=1, i \neq j}^n |R_{ji}|, \\
 &\approx \arg \inf_{\{R_{ji}\}_{i=1, \dots, m, i \neq j}} \\
 \frac{1}{N} \sum_{k=0}^N (x_j(k) - \sum_{i=1, i \neq j}^m R_{ji} x_i(k-1))^2 + \gamma \sum_{i=1, i \neq j}^n |R_{ji}|.
 \end{aligned} \tag{A.1}$$

Algorithm 3 Topology learning using One lag regression

Input: nodal time samples $x_i(k)$ for nodes $i \in \{1, 2, \dots, n\}$ in the generative graph $\mathcal{G}$, thresholds ϵ

Output: Estimate of Edges $\bar{\mathcal{E}}_R$ in the topology of $\mathcal{G}$

```

1: Edge set  $\bar{\mathcal{E}}_R \leftarrow \{\}$ 
2: for all  $j \in \{1, 2, \dots, n\}$  do
3:   Compute the Regression coefficient  $R_j = [R_{j1} \dots R_{jn}]$  using
   Eq. (A.1), where  $R_{ji}$  is the regression coefficient obtained by
   regressing  $x_j(k)$  on  $x_i(k-1)$ 
4: end for
5: for all  $i, j \in \{1, 2, \dots, n\}, i \neq j$  do
6:   if  $|R_{ji}| > \rho$  then
7:      $\bar{\mathcal{E}}_R \leftarrow \bar{\mathcal{E}}_R \cup \{(i, j)\}$ 
8:   end if
9: end for

```

Here $\gamma \geq 0$ is the regularization parameter.

References

- Allen, F., & Babus, A. (2009). Networks in finance. *The Network Challenge*, 367–382.
- Athay, T., Podmore, R., & Virmani, S. (1979). A practical method for the direct analysis of transient stability. *IEEE Transactions on Power Apparatus and Systems*, (2), 573–584.
- Bach, F. R. (2008). Consistency of the group lasso and multiple kernel learning. *Journal of Machine Learning Research (JMLR)*, 9(Jun), 1179–1225.
- Bolognani, S., Bof, N., Michelotti, D., Muraro, R., & Schenato, L. (2013). Identification of power distribution network topology via voltage correlation analysis. In *IEEE decision and control* (pp. 1659–1664). IEEE.
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3), 186.
- Cavaro, G., & Arghandeh, R. (2018). Power distribution network topology detection with time-series signature verification method. *IEEE Transactions on Power Systems*, 33(4), 3500–3509.
- Charbit, M. (2010). *Digital signal and image processing using MATLAB* (Vol. 666). John Wiley & Sons.
- Chiuso, A., & Pillonetto, G. (2012). A Bayesian approach to sparse dynamic network identification. *Automatica*, 48(8), 1553–1565.
- Chowdhury, S., & Crossley, P. (2009). *Microgrids and active distribution networks*. The Institution of Engineering and Technology.
- Crawley, D. B., Pedersen, C. O., Lawrie, L. K., & Winkelmann, F. C. (2000). Energyplus: energy simulation program. *ASHRAE Journal*, 42(4), 49.
- Dahlhaus, R. (2000). Graphical interaction models for multivariate time series1. *Metrika*, 51(2), 157–172.
- Deka, D., Backhaus, S., & Chertkov, M. (2016). Estimating distribution grid topologies: A graphical learning based approach. In *Power systems computation conference*.
- Deka, D., Chertkov, M., & Backhaus, S. (2017). Structure learning in power distribution networks. *IEEE Transactions on Control of Network Systems*.
- Deka, D., Talukdar, S., Chertkov, M., & Salapaka, M. V. (2017). Topology estimation in bulk power grids: Guarantees on exact recovery. In *Bulk power systems dynamics and control symposium (X) IREP symposium*.
- Doddi, H., Talukdar, S., Deka, D., & Salapaka, M. (2018). Data-driven identification of a thermal network in multi-zone building. In *IEEE 57th annual conference on decision and control*, 2018. IEEE.
- DoE, U. (2010). Energyplus engineering reference. In *The reference to energyplus calculations*.
- Egilmez, H. E., Pavez, E., & Ortega, A. (2017). Graph learning from data under Laplacian and structural constraints. *IEEE Journal of Selected Topics in Signal Processing*, 11(6), 825–841.
- Ellis, P. G., Torcellini, P. A., & Crawley, D. B. (2008). *Energy design plugin: An energyplus plugin for sketchup*; Preprint: Tech. Rep., National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Etesami, J., & Kiyavash, N. (2014). Directed information graphs: A generalization of linear dynamical graphs. In *2014 American control conference* (pp. 2563–2568). IEEE.

- Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3), 432–441.
- Giannakis, G. B., Shen, Y., & Karanikolas, G. V. (2018). Topology identification and learning over graphs: Accounting for nonlinearities and dynamics. *Proceedings of the IEEE*, 106(5), 787–807.
- Goncalves, J., & Warnick, S. (2008). Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks. *IEEE Transactions on Automatic Control*, 53(7), 1670–1674.
- Gubner, J. A. (2006). *Probability and random processes for electrical and computer engineers*. Cambridge University Press.
- Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity: the lasso and generalizations*. CRC press.
- Hayden, D., Chang, Y. H., Goncalves, J., & Tomlin, C. J. (2016). Sparse network identifiability via compressed sensing. *Automatica*, 68, 9–17.
- Huang, Y., Benesty, J., & Chen, J. (2006). *Acoustic MIMO signal processing*. Springer Science & Business Media.
- Kibangou, A. Y., & Commault, C. (2012). Decentralized Laplacian eigenvalues estimation and collaborative network topology identification. *IFAC Proceedings Volumes*, 45(26), 7–12.
- Kundur, P., Balu, N. J., & Lauby, M. G. (1994). *Power system stability and control* (Vol. 7). McGraw-hill New York.
- Materassi, D., & Salapaka, M. V. (2012). On the problem of reconstructing an unknown topology via locality properties of the wiener filter. *IEEE Transactions on Automatic Control*, 57(7), 1765–1777.
- Materassi, D., & Salapaka, M. V. (2013). Reconstruction of directed acyclic networks of dynamical systems. In *American control conference, 2013* (pp. 4687–4692). IEEE.
- Meinshausen, N., & Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *The Annals of Statistics*, 1436–1462.
- Morbidi, F., & Kibangou, A. Y. (2014). A distributed solution to the network reconstruction problem. *Systems & Control Letters*, 70, 85–91.
- Nabi-Abdolyousefi, M. (2014). Network identification via node knockout. In *Controllability, identification, and randomness in distributed systems* (pp. 17–29). Springer.
- Nuqui, R. F., & Phadke, A. G. (2005). Phasor measurement unit placement techniques for complete and incomplete observability. *IEEE Transactions on Power Delivery*, 20(4), 2381–2388.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann.
- Pearl, J. (2009). *Causality*. Cambridge university press.
- Pereira, J., Ibrahimi, M., & Montanari, A. (2010). Learning networks of stochastic differential equations. In *Advances in neural information processing systems* (pp. 172–180).
- Prakash, M., Talukdar, S., Attree, S., Yadav, V., & Salapaka, M. (2016). Max-min algorithm for distributed finite termination of consensus in presence of delays. ArXiv preprint arXiv:1701.00021.
- Quinn, C. J., Kiyavash, N., & Coleman, T. P. (2015). Directed information graphs. *IEEE Transactions on Information Theory*, 61(12), 6887–6909.
- Radecki, P., & Hencsey, B. (2012). Online building thermal parameter estimation via unscented kalman filtering. In *American control conference, 2012* (pp. 3056–3062). IEEE.
- Reed, M., & Simon, B. (1980). *Functional analysis (Vol. I)*. Academic Press, San Diego.
- Reynders, G., Diriken, J., & Saelens, D. (2014). Quality of grey-box models and identified parameters as function of the accuracy of input and observation signals. *Energy and Buildings*, 82, 263–274.
- Scott, J. (2017). *Social network analysis*. Sage.
- Segarra, S., Marques, A. G., Mateos, G., & Ribeiro, A. (2017). Network topology inference from spectral templates. *IEEE Transactions on Signal and Information Processing over Networks*, 3(3), 467–483.
- Segarra, S., Schaub, M. T., & Jadbabaie, A. (2017). Network inference from consensus dynamics. In *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on* (pp. 3212–3217). IEEE.
- Shafipour, R., Segarra, S., Marques, A. G., & Mateos, G. (2018). Identifying the topology of undirected networks from diffused non-stationary graph signals. ArXiv preprint arXiv:1801.03862.
- Shahrpour, S., & Preciado, V. M. (2015). Topology identification of directed dynamical networks via power spectral analysis. *IEEE Transactions on Automatic Control*, 60(8), 2260–2265.
- Siami, M., Bolouki, S., Bamieh, B., & Motee, N. (2017). Centrality measures in linear consensus networks with structured network uncertainties. *IEEE Transactions on Control of Network Systems*, 1.
- Stoica, P., Moses, R. L., et al. (2005). *Spectral analysis of signals*. NJ: Pearson Prentice Hall Upper Saddle River.
- Talukdar, S., Deka, D., Attree, S., Materassi, D., & Salapaka, M. V. (2017). Learning the exact topology of undirected consensus networks. ArXiv preprint arXiv:1710.00032.
- Talukdar, S., Deka, D., Lundstrom, B., Chertkov, M., & Salapaka, M. V. (2017). Learning exact topology of a loopy power grid from ambient dynamics. In *Proceedings of the eighth international conference on future energy systems* (pp. 222–227). ACM.
- Talukdar, S., Deka, D., Materassi, D., & Salapaka, M. V. (2017). Exact topology reconstruction of radial dynamical systems with applications to distribution system of the power grid. In *American control conference*.
- Yuan, Y., Stan, G.-B., Warnick, S., & Goncalves, J. (2011). Robust dynamical network structure reconstruction. *Automatica*, 47(6), 1230–1235.



Saurav Talukdar is a Control Systems and Machine Learning Engineer at Google and focuses on energy optimization for Google Data Center. Prior to joining Google in 2019, he was a Battery Algorithm Engineer at Apple working on system identification and thermal management of Lithium ion batteries in iPhones. He received the B.Tech and M.Tech degrees in Mechanical Engineering from Indian Institute of Technology, Bombay, India in 2013 and Ph.D. degree in Mechanical Engineering from the University of Minnesota, Minneapolis, USA in 2018.



Deepjyoti Deka is a staff scientist in the Applied Mathematics and Plasma Physics group of the Theoretical Division at Los Alamos National Laboratory (LANL), where he was previously a postdoctoral research associate at the Center for Nonlinear Studies. His research interests include data-analysis of power grid structure, operations and security, and optimization in social and physical networks. Before joining the laboratory he received the M.S. and Ph.D. degrees in electrical engineering from the University of Texas, Austin, TX, USA, in 2011 and 2015, respectively and his B.Tech degree in Electrical Engineering from IIT Guwahati, India in 2009.



Harish Doddi is a Ph.D. candidate in the Mechanical Engineering Department at the University of Minnesota, Minneapolis since 2016. He obtained his B.Tech degree from Indian Institute of Technology Madras, Chennai in 2014 and worked on Thermal Analysis of Commercial Buildings at the TCS Innovation Lab, Chennai, India till 2016. His research includes Information Thermodynamics, Energy limits in Computing and Structure learning of Graphical models.



Donatello Materassi holds a Laurea in “Ingegneria Informatica” and a “Dottorato di Ricerca” in Electrical Engineering/Nonlinear Dynamics and Complex Systems from Università degli Studi di Firenze, Italy. He has been a research associate at University of Minnesota (Twin Cities) from 2008 till 2011. He has been concurrently both a post-doctoral researcher at Laboratory for Information and Decision Systems (LIDS) at Massachusetts Institute of Technology and a lecturer at Harvard University FROM 2012–2014. Since 2019, he is an assistant professor at University of Minnesota in Minneapolis. Prior to that, he was an assistant professor at the University of Tennessee, Knoxville. He received the NSF CAREER award in 2015. His research interests include nonlinear dynamics, system identification and classical control theory with applications to atomic force microscopy, single molecule force spectroscopy, biophysics, statistical mechanics and quantitative finance.



Michael Chertkov's area of focus is mathematics, including statistics and data science, applied to physical, engineered and other systems. Dr. Chertkov received his Ph.D. in physics from the Weizmann Institute of Science in 1996, and his M.Sc. in physics from Novosibirsk State University in 1990. After his Ph.D., Dr. Chertkov spent three years at Princeton University as a R.H. Dicke Fellow in the Department of Physics. He joined Los Alamos National Lab in 1999, initially as a J.R. Oppenheimer Fellow in the Theoretical Division, and continued as a Technical Staff Member. In 2019, Dr.

Chertkov joined the University of Arizona as a Professor of Mathematics and leads the Interdisciplinary Graduate Program in Applied Mathematics. He is a fellow of the American Physical Society (APS) and a senior member of IEEE.



Murti V. Salapaka received the B.Tech. degree in Mechanical Engineering from the Indian Institute of Technology, Madras, in 1991 and the M.S. and Ph.D. degrees in Mechanical Engineering from the University of California at Santa Barbara, in 1993 and 1997, respectively. He was a faculty member in the Electrical and Computer Engineering Department, Iowa State University, Ames, from 1997 to 2007. Currently, he is the Director of Graduate Studies and the Vincentine Hermes Luh Chair Professor in the Electrical and Computer Engineering Department, University of Minnesota, Minneapolis.

His research interests include control and network science, nanoscience and single molecule physics. Dr. Salapaka received the 1997 National Science Foundation CAREER Award and is an IEEE fellow.