# On the Problem of Reconstructing an Unknown Topology via Locality Properties of the Wiener Filter

Donatello Materassi and Murti V. Salapaka

***Abstract*—Determining interrelatedness structure of various entities from multiple time series data is of significant interest to many areas. Knowledge of such a structure can aid in identifying cause and effect relationships, clustering of similar entities, identification of representative elements and model reduction. The majority of existing results are based on correlation based indices which effectively assume a static relationship between the time series data and are not suitable for detecting interrelatedness when the time series are dynamically related or when the time series involve loops. In this paper, a methodology for identifying the interrelatedness structure of dynamically related time series data is presented that also allows for the presence of loops in the connectivity structure. A linear dynamic graph model is presented where it is assumed that each time series data is the sum of an independent stochastic noise source and a dynamically weighted sum of other time series data. A link is assumed to be present between two time series if the weight of a time series, which is a linear time-invariant filter, is nonzero in the formation of the other. Reconstruction of the link connectivity structure under various scenarios is considered. It is shown that when the linear dynamic graph is allowed to admit non-causal weights, then the links structure can be recovered with the possibility of identifying spurious connections. However, it is shown that the spurious links remain local, where, a spurious link is restricted to be within one hop of a true link. Furthermore, strategies for exact reconstruction of the link structure when the weights are restricted to be causal are developed. The main tools for determining the network topology are based on variations of Wiener filtering. A significant insight provided by the article is that, in the class of network models considered in the paper, the Wiener filter estimating a stochastic process based on other processes remains local in the sense that the Wiener filter utilizes only measurements local to the node being estimated.**

***Index Terms*—Filtering, network analysis, system identification.**

## I. Introduction

THE interest on networks of dynamical systems is increasing in recent years, especially because of their capability of modeling and describing a large variety of phenomena and behaviors. Interconnections of simple systems are used to understand the emergence of complicated phenomena (see, for example, [1]–[4]) and have provided novel modeling approaches in many fields, such as economics (see, e.g., [5] and [6]), Biology (see, e.g., [7]–[9]), cognitive sciences (see, e.g., [10]), ecology (see, e.g., [11] and [12]) and geology (see, e.g., [13] and [14]), especially when the investigated phenomena are characterized by spatial distributions where multivariate analysis is involved [15]. While networks of dynamical systems are well studied and analyzed in physics [16]–[18] and engineering [19]–[21], there are fewer results that address the problem of reconstructing the topology of a network. Unraveling the interconnectedness of a set of processes is of significant interest in many fields, with the necessity for general tools rapidly increasing (see [22]–[24] and the bibliography therein for recent results). However, such a problem poses formidable theoretical as well as practical challenges (see [25]). Existing results derive a network topology from sampled data (see, e.g., [6], [22], [24], and [26]) or to determine the presence of substructures (see, e.g., [18], and [23]). The unweighted pair group method with arithmetic mean (UPGMA) [27] is one of the first techniques proposed to reveal an unknown topology. It has found widespread use in the reconstruction of phylogenetic trees, and is widely employed in other areas such as communication systems and for resource allocation [28]. UPGMA identifies a tree topology relying on the observation of leaf nodes, theoretically guaranteeing a correct identification only on the strong assumption that an ultrametric is defined among the leaves. Another well-known technique for the identification of a tree network is developed in [6] for the analysis of a stock portfolio. The authors identify a tree structure according to the following procedure: 1) a metric based on the correlation index is defined among the nodes; 2) such a metric is employed to extract the minimum spanning tree [29] which forms the reconstructed topology. Many improvements over [6] are now devised especially using shrinking techniques to estimate the correlation matrix [30], [31] or refining its estimate via high frequency sampling [32]. A reliability index for any link can also be defined using bootstrap techniques [31]. However, in [15] a severe limitation of this strategy is highlighted, where it is shown that, even though the actual network is a tree, the presence of dynamical connections or delays can lead to the identification of an incorrect topology. In [33], a similar strategy, where the correlation metric is replaced by a metric based on the coherence function, is numerically shown to provide an exact reconstruction for tree topologies. In [34], it is shown that a correct reconstruction can be guaranteed for a topology with no cycles.

In [23], different techniques for quantifying and evaluating the modular structure of a network are compared and a new

Manuscript received January 15, 2010; revised September 13, 2010 and April 01, 2011; accepted August 25, 2011. Date of publication January 06, 2012; date of current version June 22, 2012. This work was supported in part by the National Science Foundation (NSF) under Grant CMMI-0900113. Recommended by Associate Editor H. Zhang.

D. Materassi is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: donnie13@mit.edu).

M. V. Salapaka is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455-0167 USA (e-mail: murtis@umn.edu).

Digital Object Identifier 10.1109/TAC.2012.2183170

one is proposed trying to combine both the topological and dynamic information of the complex system. However, the network topology is only qualitatively estimated. In [22], a method to identify a network of dynamical systems is described. However, primary assumptions of the technique are the possibility to manipulate the input of every single node and the possibility of conducting experiments to detect the link connectivity. In [35], an interesting and novel approach based on auto-regressive models and Granger-causality [36] is proposed for reconstructing a network of dynamical systems. This technique relies on multivariate identification procedure to detect the presence of a link, but still no theoretical sufficient or necessary conditions are derived to check the correctness of the results. More recently, in [24] and [37] interesting equivalences between the identification of a dynamical network and a $l_0$ sparsification problem are highlighted, suggesting the difficulty of the reconstruction procedure [38]. Summarizing, to the best knowledge of the authors, apart from the results in [39], which are limited to tree topologies, no general theoretical guarantees about the correct reconstruction of network links are provided if there is no possibility of directly manipulating the inputs of the node.

This paper addresses the problem of reconstructing a network of dynamical systems where every node represents an observable scalar signal and the dynamics, which is linear, is represented by the connecting links. The problem, when analyzed from a systems theory point of view, provides a method for correctly identifying a topology that belongs to the pre-specified class of self-kin networks. Moreover, if the network does not belong to such a class, conditions about the optimality of the identified topology according to a defined criterion is established. From this perspective, sufficient conditions for the exact reconstruction of a large class of networks, which we name self-kin, are derived. Examples of self-kin networks are given by (but not limited to) rooted trees, and ring topologies [40]. In the case the network is not self-kin, the reconstructed topology is guaranteed to be the smallest self-kin network containing the actual one. The theory developed is not Bayesian and relies directly on Wiener filtering theory. Conditions derived for the detection of links are based on sparsity properties of the (non-causal) Wiener filter modeling the network. Indeed, conditions under which the Wiener filter smoothing a signal of the network is "local" are derived. From a different perspective, another important contribution of the paper are conditions for a local and distributed implementation of the Wiener filter. The results obtained bear a striking similarity to the ones developed in the area of machine learning for Bayesian networks (BNs) [41], [42] where the topology of a network of nodes that represent random variables is sought. One of the main results obtained in the BNs literature (see [43]) is that the probability distribution of a random variable conditioned on the rest of the random variables of the network is equal to the probability distribution of the random variable conditioned only on the random variables within the kin set of the random variable (Markov blanket). It is assumed, though, that the network has no loops. The problem considered in this paper is for a network of random processes and is not restricted to random variables as is the case for BNs. Evidently, issues concerning causality and stability do not arise for BNs which have to be addressed for a network of random processes. Moreover, in this paper no assumption on the absence of loops is made as is the case in [43].

The paper is organized as follows. In Section II, examples are illustrated to provide the basic intuition behind central ideas; in Section III definitions are provided based on standard notions of graph theory; in Section IV the main results are provided for non-causal Wiener filtering; in Section V the results are extended to causal Wiener filtering and Granger causality; in Section VI the implementation of algorithms for the detection of network topologies are discussed for different scenarios; in Section VII the robustness of the identification is addressed; eventually, in Section VIII numerical simulations illustrating the effectiveness of the methodology are presented.

*Notation:* The symbol $:=$ denotes a definition

| | |
|---|---|
| $\|x\|$ | 2-norm of a vector $x$. |
| $W^T$ | the transpose of a matrix or vector $W$. |
| $W^*$ | the conjugate transpose of a matrix or vector $W$. |
| $x_i$ or $\{x\}_i$ | the $i$th element of a vector $x$. |
| $W_{ji}$ | the entry $(j,i)$ of a matrix $W$. |
| $W_{j*}$ | $j$th row of a matrix $W$. |
| $W_{*i}$ | $i$th column of a matrix $W$. |
| $x_V$ | when $V=(v_1,\ldots,v_n)$ is a $n$-tuple of natural numbers denotes the vector $(x_{v_1}\ldots x_{v_n})^T$. |
| $\|A\|$ | cardinality (number of elements) of a set $A$. |
| $E[\cdot]$ | mean operator. |
| $R_{XY}(\tau):=E[X(t)Y^T(t+\tau)]$ | cross-covariance function of wide-sense stationary vector processes $X$ and $Y$. |
| $R_X(\tau):=R_{XX}(\tau)$ | autocovariance. |
| $\mathcal{Z}(\cdot)$ | Z-transform of a signal. |
| $\Phi_{XY}(z):=\mathcal{Z}(R_{XY}(\tau))$ | cross-power spectral density. |
| $\Phi_X(z):=\Phi_{XX}(z)$ | power spectral density. |
| $b_i$ | $i$th element of the canonical base of $\mathbb{R}^n$. |

In this section, a representation of dynamical networks in terms of oriented graphs is presented. According to this representation, every node $x_j$ is given by a scalar time-discrete wide-sense stationary stochastic process, while every directed arc form a node $x_i$ to a node $x_j$ represents a possibly non-causal transfer function $H_{ji}(z)\neq 0$. The absence of such an arc implies that $H_{ji}(z)=0$. Every node signal is also implicitly considered affected by an additive process noise $e_j$. Given the graphical representation, the dynamics of the network is described by

$$x_j=e_j+\sum_i H_{ji}(z)x_i. \tag{1}$$

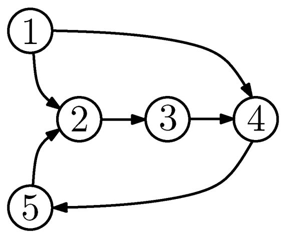


Fig. 1. Example of the graphical representation of dynamical networks. Every node represents a stochastic signal and the edge from a node $x_i$ to a node $x_j$ represents the transfer function $H_{ji}(z)$. Every node signal is also implicitly affected by a process noise.

For example, in Fig. 1 a network is represented, the dynamics of which corresponds to

$$\begin{aligned} x_1 &= e_1 \\ x_2 &= e_2 + H_{25}(z)x_5 \\ x_3 &= e_3 + H_{31}(z)x_1 + H_{32}(z)x_2 \\ x_4 &= e_4 + H_{43}(z)x_3 \\ x_5 &= e_5 + H_{51}(z)x_1 + H_{54}(z)x_4. \end{aligned}$$

## II. Illustrative Examples

In this section, special network configurations of dynamical systems are presented where the Wiener filter producing the estimate of a node signal $x_0$ from the other node signals $x_i$ has the characteristic of being sparse. In most cases, the presence of a non-null entry in the Wiener filter corresponds to the presence of a direct link between the signal $x_i$ and the signal $x_0$. In the following sections, it will be shown that this result holds in general for the class of self-kin networks, a class of networks defined later. Moreover, if the network is not self-kin it will be shown that the presence of a non-null entry in the Wiener filter identifies the presence of a link in the "smallest" (in the sense of number of edges) self-kin network containing the original one.

### A. Wiener Filtering of a Downstream Signal

Consider a network of four systems as represented in Fig. 2 where

$$\begin{aligned} x_3 &= e_3 \\ x_{i-1} &= e_{i-1} + H_{i-1,i}(z)x_i \end{aligned}$$

with $H_{i-1,i}(z) \neq 0$ three possibly non-causal SISO transfer functions for $i = 1, 2, 3$ and with signals $e_i$ mutually uncorrelated for $i = 0, 1, 2, 3$. Consider the Wiener filter that provides the estimate $\hat{x}_0$ of $x_0$ based on the other signals, $x_1$, $x_2$ and $x_3$. It can be shown that

$$\begin{aligned} \hat{x}_0 &= \begin{pmatrix} W_{01}(z) & W_{02}(z) & W_{03}(z) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} H_{01}(z) & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned}$$

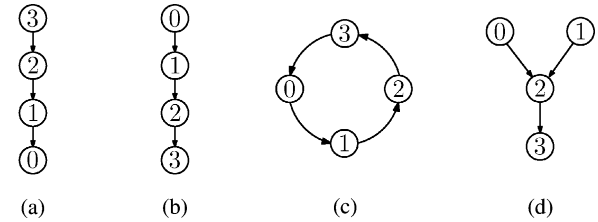


Fig. 2. Cascade network (a): the Wiener filter estimating $x_0$ from the other signals makes use only of the "child" signal $x_1$ which is directly linked to $x_0$. A cascade network (b): the Wiener filter estimating $x_0$ from the other signals makes use only of the "child" signal $x_1$ which is directly linked to $x_0$. A loop network (c): the Wiener filter estimating $x_0$ from the other signals makes use only of the "child" signal $x_1$ and the parent signal $x_3$ which are directly linked to $x_0$. A more general network (d): the Wiener filter estimating $x_0$ from the other signals makes use of the signal $x_2$ which is directly linked to it, but also of the signal $x_1$ which is not.

is the best estimate of $x_0$ in the least squares sense. Indeed,

$$E\left[(x_0 - \hat{x}_0)x_i\right] = E[e_0 x_i] = 0 \tag{2}$$

for $i = 1, 2, 3$. Thus, from the Hilbert projection theorem [44]

$$W_0(z) := \begin{pmatrix} W_{01}(z) & W_{02}(z) & W_{03}(z) \end{pmatrix} \tag{3}$$

is the Wiener filter providing the best estimate for $x_0$ [45]. In this case, the only non-null entry of $W_0$ correctly detects that there is a connection between the node 0 and the node 1. Note that the Wiener filter $W_0(z)$ can be computed using the time series of $x_0, x_1, x_2, x_3$ without the knowledge of the transfer functions of the network.

### B. Wiener Filtering of an Upstream Signal

Consider a network of four systems as represented in Fig. 2(a) with

$$x_0 = e_0 \tag{4}$$

$$x_{i+1} = H_{i+1,i}(z)x_i + e_{i+1} \tag{5}$$

where $H_{i+1,i}(z)$ are three possibly non-causal SISO transfer functions for $i = 1, 2, 3$ and with signals $e_i$ mutually not correlated for $i = 0, 1, 2, 3$. The Wiener filter that provides the estimate $\hat{x}_0$ of $x_0$ making use of the other signals $x_1$, $x_2$ and $x_3$ is such that

$$\begin{aligned} \hat{x}_0 &= \begin{pmatrix} W_{01}(z) & W_{02}(z) & W_{03}(z) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{\Phi_{x_0x_0}(z)H_{10}^*(z)}{|H_{10}(z)|^2\Phi_{x_0}(z) + \Phi_{e_1}} & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \end{aligned}$$

Again in this case, there is one only non-null entry in the Wiener filter corresponding to the link between the node 0 and the node 1.

### C. Wiener Filtering of a Loop

In the two previous cases the network configurations did not involve loops and the resulting Wiener filter had the property of having non-null entries corresponding to the node signals immediately connected to the node of interest. The identification of network topologies with loops is a challenging problem [6], [15]. Indeed, most techniques deal with networks with no cycles

[39]. The presence of a loop leads to complex relations between the node signals, especially in terms of the covariance function and (cross)-spectral densities. Thus, it is interesting to note that the absence or presence of loops does not seem to affect the sparsity of the Wiener filter as shown in the following example [39].

Consider a network of four systems as represented in Fig. 2(c) where

$$x_i = H_{i,[i-1]_{mod\ 4}}(z)x_{[i-1]_{mod\ 4}} + e_i$$

for $i = 0, \dots, 3$ and

$$[n]_{mod\ m} := \min\{q | q = n + km \geq 0 \text{ and } k \in \mathbb{Z}\}. \quad (6)$$

Assume also that the signals $e_i$ are mutually not correlated.

It is possible to show that the best least squares estimate of $x_0$, based on $x_1$, $x_2$ and $x_3$, is given by

$$\begin{aligned}\hat{x}_0 &= \begin{pmatrix} W_{01}(z) & W_{02}(z) & W_{03}(z) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \left( \frac{\Phi_{e_0} H_{10}(z)^*}{\Phi_{e_0}|H_{10}|^2 + \Phi_{e_1}} \quad 0 \quad \left(1 - \frac{\Phi_{e_0}|H_{10}|^2}{\Phi_{e_0}|H_{10}|^2 + \Phi_{e_1}}\right) H_{03}(z) \right) \\ &\quad \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.\end{aligned}$$

It is to be noted that the signal $x_2$, which is not directly connected to $x_0$, is not used in the optimal estimate of $x_0$. Thus, in all the examples presented so far the node signals actually employed by the Wiener Filter are the signals corresponding to nodes directly connected to the node signal being estimated. This example, along with the previous ones, leads to the conjecture that the Wiener filter can be used as a tool to identify which nodes are connected to a specific node of interest by checking whether the corresponding entries of the Wiener filter are null or not. Unfortunately, the situation is more complex as shown in the following example.

### *D. Wiener Filter Using a Not Directly Connected Signal*

Consider a network of four systems as represented in Fig. 2(d) where

$$\begin{aligned} x_0 &= e_0 \\ x_1 &= e_1 \\ x_2 &= H_{2,0}(z)x_0 + H_{2,1}(z)x_1 + e_2 \\ x_3 &= H_{3,2}(z)x_2 + e_3 \end{aligned}$$

and the noises $e_j$ are mutually not correlated. In this example, the Wiener filter providing the estimate $\hat{x}_0$ of $x_0$ using the other signals of the network is given by

$$\begin{aligned}\hat{x}_0 &= \begin{pmatrix} W_{01}(z) & W_{02}(z) & W_{03}(z) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \frac{\Phi_{e_0}(z) H^*_{20}(z) \begin{pmatrix} -H_{21}(z) & 1 & 0 \end{pmatrix}}{\Phi_{e_0}(z)|H_{20}(z)|^2 + \Phi_{e_2}(z)} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.\end{aligned}$$

In this case the Wiener filter estimating $x_0$ from the other signals makes use not only of the signal $x_2$ which is directly linked to it, but also of the signal $x_1$ which is not.

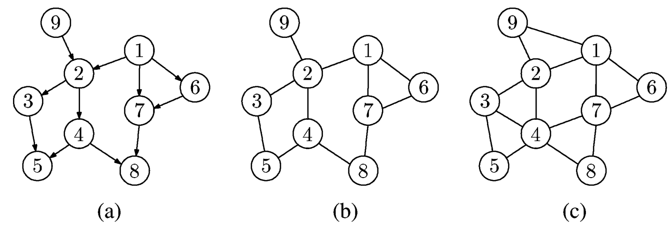


Fig. 3. (a) Directed graph, (b) its topology, and (c) its kin-topology.

After the analysis of these examples, it is natural to ask to what extent and under which assumptions it is possible to reconstruct the topology of a network of linear dynamical systems measuring the node signals and if the Wiener filter can be a useful tool for accomplishing this identification procedure.

## III. Preliminary Definitions

In this section, basic notions of graph theory, which are functional to the subsequent developments, will be recalled. For an extensive overview see [29]. First, the standard definition of undirected and oriented graphs is provided.

*Definition 1 (Directed and Undirected Graphs):* An undirected graph $G$ is a pair $(V, A)$ where $V$ is a set of vertices or nodes and $A$ is a set of edges or arcs, which are unordered subsets of two distinct elements of $V$.

A directed (or oriented) graph $G$ is a pair $(V, A)$ where $V$ is a set of vertices or nodes and $A$ is a set of edges or arcs, which are ordered pairs of elements of $V$.

In the following, if not specified, oriented graphs are considered.

*Definition 2 (Topology of a Graph):* Given an oriented graph $G = (V, A)$, its topology is defined as the undirected graph $G' = (V, A')$ such that $\{N_i, N_j\} \in A'$ if and only if $(N_i, N_j) \in A$ or $(N_j, N_i) \in A$, and $top(G) := G'$.

By removing the orientation on any edge of an oriented graph $G$, an undirected graph $G'$ is obtained that is its topology. An example of a directed graph is represented in Fig. 3(a) with its topology in Fig. 3(b).

*Definition 3 (Children and Parents):* Given a graph $G = (V, A)$ and a node $N_j \in V$, the children of $N_j$ are defined as $\mathcal{C}_G(N_j) := \{N_i | (N_j, N_i) \in A\}$ and the parents of $N_j$ as $\mathcal{P}_G(N_j) := \{N_i | (N_i, N_j) \in A\}$.

Extending the notation, children and the parents of a set of nodes are denoted as follows:

$$\mathcal{C}_G(\{N_{j_1}, \dots, N_{j_m}\}) := \cup_{k=1}^m \mathcal{C}_G(N_{j_k})$$
$$\mathcal{P}_G(\{N_{j_1}, \dots, N_{j_m}\}) := \cup_{k=1}^m \mathcal{P}_G(N_{j_k}).$$

*Definition 4 (Kins):* Given an oriented graph $G = (V, A)$ and a node $N_j \in V$, kins of $N_j$ are defined as

$$\begin{aligned}\mathcal{K}_G(N_j) := \{N_i | N_i \neq N_j \text{ and } N_i \in \mathcal{C}_G(N_j) \cup \\ \cup \mathcal{P}_G(N_j) \cup \mathcal{P}_G(\mathcal{C}_G(N_j))\}.\end{aligned}$$

Kins of a set of nodes are defined in the following way:

$$\mathcal{K}_G(\{N_{j_1}, \dots, N_{j_m}\}) := \cup_{k=1}^m \mathcal{K}_G(N_{j_k}).$$

*Definition 5 (Proper Parents and Proper Children):* Given an oriented graph $G = (V, A)$ and a node $N_j$, $N_i$ is a proper parent (child) of $N_j$ if it is a parent (child) of $N_j$ and $N_i \notin \mathcal{P}_G(\mathcal{C}_G(N_j))$. $N_i$ is a proper kin if it is a kin and $N_i \notin \mathcal{P}_G(N_j) \cup \mathcal{C}_G(N_j)$.

Note that the kin relation is symmetric, in the sense that $N_i \in \mathcal{K}_{\mathcal{G}}(N_j)$ if and only if $N_j \in \mathcal{K}_{\mathcal{G}}(N_i)$.

*Definition 6 (Kin-Graph):* Given an oriented graph $G = (V, A)$, its kin-graph is the undirected graph $\tilde{G} = (V, \tilde{A})$ where

$$\tilde{A} := \{\{N_i, N_j\} | N_i \in \mathcal{K}_G(N_j) \; for \; all \; j\}.$$

and it is denoted as$kin(G) = \tilde{G}$.

A directed graph and its kin-topology are represented in Fig. 3(a) and (c), respectively. Note that the kin-graph of $G$ is an undirected graph. It could be defined as a directed graph, but, because of the symmetry of the kin relation, a directed graph contains exactly the same information. Moreover, such a choice is motivated by the following definition.

*Definition 7 (Self-Kin Graph):* An oriented graph $G$ is self-kin if $top(G) = kin(G)$.

Many graphs are self-kin, such as oriented trees, rings, and triangular lattices [29].

*Definition 8:* Let $\mathcal{E}$ be a set containing time-discrete scalar, zero-mean, jointly wide-sense stationary random processes such that, for any $e_i, e_j \in \mathcal{E}$, the power spectral density $\Phi_{e_i e_j}(z)$ exists, is real rational with no poles on the unit circle and given by

$$\Phi_{e_i e_j}(z) = \frac{A(z)}{B(z)}$$

where $A(z)$ and $B(z)$ are polynomials with real coefficients such that $B(z) \neq 0$ for any $z \in \mathbf{C}$, with $|z| = 1$. Then, $\mathcal{E}$ is a set of rationally related random processes.

*Definition 9:* The set $\mathcal{F}$ is defined as the set of real-rational single-input single-output (SISO) transfer functions that are analytic on the unit circle $\{z \in \mathbf{C} | \; |z| = 1\}$.

*Definition 10:* Given a SISO transfer function $H(z) \in \mathcal{F}$, represented as

$$H(z) = \sum_{k=-\infty}^{\infty} h_k z^{-k} \tag{7}$$

the causal truncation operator is defined as

$$\{H(z)\}_C := \sum_{k=0}^{\infty} h_k z^{-k}. \tag{8}$$

*Lemma 11:* For every $H(z) \in \mathcal{F}$, it holds that $\{H(z)\}_C \in \mathcal{F}$.

*Definition 12:* The set $\mathcal{F}^+$ is defined as the set of real-rational SISO transfer functions in $\mathcal{F}$ such that

$$\{H(z)\}_C = H(z). \tag{9}$$

*Definition 13:* Let $\mathcal{E}$ be a set of rationally related random processes. The set $\mathcal{FE}$ is defined as

$$\mathcal{FE} := \left\{ x = \sum_{k=1}^{m} H_k(z) e_k | e_k \in \mathcal{E}, H_k(z) \in \mathcal{F}, m \in \mathbb{N} \right\}.$$

*Lemma 14:* The set $\mathcal{FE}$ is a vector space with the field of real numbers. Let

$$\langle x_1, x_2 \rangle := R_{x_1 x_2}(0) = \int_{-\pi}^{\pi} \Phi_{x_1 x_2}(e^{i\omega})$$

which defines an inner product on $\mathcal{FE}$ with the assumption that two processes $x_1$ and $x_2$ are considered identical if $x_1(t) = x_2(t)$, almost always for any $t$.

*Proof:* The proof is done by inspection checking the properties of vector space and of inner product. ■

For any $x \in \mathcal{FE}$, the norm induced by the inner product is defined as $\|x\| := \sqrt{\langle x, x \rangle}$.

*Definition 15:* For a finite number of elements $x_1, \ldots, x_m \in \mathcal{FE}$, $tf - span$ is defined as

$$tf - span\{x_1, \ldots, x_m\} := \left\{ x = \sum_{i=1}^{m} \alpha_i(z) x_i | \alpha_i(z) \in \mathcal{F} \right\}.$$

*Lemma 16:* The $tf - span$ operator defines a subspace of $\mathcal{FE}$.

*Proof:* The proof is left to the reader. ■

*Definition 17:* For a finite number of elements $x_1, \ldots, x_m \in \mathcal{FE}$, $c - tf - span$ is defined as

$$c{-}tf{-}span\{x_1, \ldots, x_m\} := \left\{ x = \sum_{i=1}^{m} \alpha_i(z) x_i | \alpha_i(z) \in \mathcal{F}^+ \right\}.$$

*Lemma 18:* The $c - tf - span$ operator defines a subspace of $\mathcal{FE}$.

*Proof:* The proof is left to the reader. ■

The following definition provides a class of models for a network of dynamical systems. It is assumed that the dynamics of each agent (node) in the network is represented by a scalar random process $\{x_j\}_{j=1}^n$ that is given by the superposition of a noise component $e_j$ and the "influences" of some other "parent nodes" through dynamic links. The noise acting on each node is assumed not related with the other noise components. If a certain agent "influences" another one a directed edge can be drawn and a directed graph can be obtained.

*Definition 19 (Linear Dynamic Graph):* A linear dynamic graph $\mathcal{G}$ is defined as a pair $(H(z), e)$ where:

- $e = (e_1, \ldots, e_n)^T$ is a vector of $n$ rationally related random processes such that $\Phi_e(z)$ is diagonal;
- $H(z)$ is a $n \times n$ matrix of transfer functions in $\mathcal{F}$ such that $H_{jj}(z) = 0$, for $j = 1, \ldots, n$.

The output processes $\{x_j\}_{j=1}^n$ of the LDG are defined as

$$x_j = e_j + \sum_{i=1}^{n} H_{ji}(z) x_i$$

or in a more compact way

$$x(t) = e(t) + H(z)x(t). \tag{10}$$

Let $V := \{x_1, \ldots, x_n\}$ and let $A := \{(x_i, x_j) | H_{ji}(z) \neq 0\}$. The pair $G = (V, A)$ is the associated directed graph of the LDG. Nodes and edges of a LDG will mean nodes and edges of the graph associated with the LDG.

Observe that (10) defines a map from a vector of rationally related processes $x$ to a vector of rationally related processes $e$.

Indeed, $e = (\mathcal{I} - H(z))x$ and each entry of $(\mathcal{I} - H(z))$ has no poles on the unit circle. If the operator $(\mathcal{I} - H(z))$ is invertible on the space of rationally related processes it can be guaranteed that, for any vector of rationally related processes $e$, a vector $x$ of processes in the space $\mathcal{F}e$ will be obtained. For this reason, the following definition is introduced.

*Definition 20:* A LDG $(H(z), e)$ is well-posed if each entry of $(\mathcal{I} - H(z))^{-1}$ belongs to $\mathcal{F}$. Thus, $x = (\mathcal{I} - H(z))^{-1}e$. can be written. A LDG $(H(z), e)$ is causally well-posed if all the entries of $(\mathcal{I} - H(z))$ and $(\mathcal{I} - H(z))^{-1}$ belong to $\mathcal{F}^+$.

Observe that if all the entries of $H(z)$ are strictly causal, then the LDG is well-posed.

A LDG is a complex interconnection of linear transfer functions $H_{ji}(z)$ connected according to a graph $G$ and forced by stationary additive mutually uncorrelated noise. The following definition will be useful for determining sufficient conditions for detection of links in a network.

*Definition 21:* A LDG $\mathcal{G} = (H(z), e)$ is topologically detectable if $\Phi_{e_i}(e^{i\omega}) > 0$ for any $\omega \in [-\pi, \pi]$ and for any $i = 1, \ldots, n$.

We formulate the problem tackled in this paper as follows.

*Problem 22:* Consider a well-posed LDG $\mathcal{G} = (H, e)$ where its associated graph $G$ is unknown. Given the power (cross-) spectral densities of $\{x_j\}_{j=1,\ldots,n}$, determine sufficient/necessary conditions under which it is possible to conclude that the edge $(x_i, x_j)$ belongs to $G$, its topology $top(G)$ or its kin-graph $kin(G)$.

## IV. Sparsity of the Non-Causal Wiener Filter

First, a lemma is provided that guarantees that any element in $\mathrm{tf} - \mathrm{span}\{x_i\}_{i=1,\ldots,n}$ admits a unique representation if the cross-spectral density matrix of its generating processes has full normal rank.

*Lemma 23:* Let $q$ and $x_1, \ldots, x_n$ be processes in the space $\mathcal{FE}$. Define $x = (x_1, \ldots, x_n)^T$. Suppose that $q \in tf - span\{x_i\}_{i=1,\ldots,n}$ and that $\Phi_x(e^{i\omega}) > 0$ almost for any $\omega \in [-\pi, \pi]$. Then there exists a unique transfer matrix $\Lambda(z) \in \mathcal{F}^{1\times n}$ such that $q = \Lambda(z)x$.

*Proof:* Note that if $\Lambda(z)$ is such that $q = \Lambda(z)x = 0$, then $\Phi_{qq}(e^{i\omega}) = 0 = \Lambda(e^{i\omega})\Phi_x(e^{i\omega})\Lambda^*(e^{i\omega})$. Since $\Phi_x(e^{i\omega}) > 0$ for any $\omega \in [-\pi, \pi]$, it holds that $\Lambda(e^{i\omega}) = 0$ almost everywhere which implies that $\Lambda(z) = 0$. Now, by contradiction assume that $q = \Lambda_1(z)x = \Lambda_2(z)x$, with $\Lambda_1(z) \neq \Lambda_2(z)$. Then $0 = [\Lambda_2(e^{i\omega}) - \Lambda_1(e^{i\omega})]\Phi_x(e^{i\omega})[\Lambda_2(e^{i\omega}) - \Lambda_1(e^{i\omega})]^*$ implying that $\Lambda_1(z) = \Lambda_2(z)$. ■

A specific formulation of the non-causal Wiener filter is introduced for the defined spaces.

*Proposition 24:* Let $v$ and $x_1, \ldots, x_n$ be processes in the space $\mathcal{FE}$. Define $x := (x_1, \ldots, x_n)^T$ and $X := tf - span\{x_1, \ldots, x_n\}$. Consider the problem

$$\inf_{q\in X} \|v - q\|^2. \tag{11}$$

If $\Phi_x(e^{i\omega}) > 0$, for $\omega \in [-\pi, \pi]$, then the solution $\hat{v} \in X$ exists, is unique and is given by $\hat{v} = W(z)x$ where

$$W(z) = \Phi_{vx}(z)\Phi_x(z)^{-1}.$$

Moreover, $\hat{v}$ is the only element in $X$ such that, for any $q \in X$

$$\langle v - \hat{v}, q\rangle = 0. \tag{12}$$

*Proof:* Observe that, since $q \in X$, the cost function satisfies

$$\|v - W(z)x\|^2 = \int_{-\pi}^{\pi} \Phi_v(e^{i\omega}) + W(e^{i\omega})\Phi_x(e^{i\omega})W^*(e^{i\omega}) + \\ -\Phi_{xv}(e^{i\omega})W^*(e^{i\omega}) - W(e^{i\omega})\Phi_{vx}(e^{i\omega}).$$

The integral is minimized by minimizing the integrand for all $\omega \in [-\pi, \pi]$. It is straightforward to find that the minimum is achieved for

$$W(e^{i\omega}) = \Phi_{vx}(e^{i\omega})\Phi_x(e^{i\omega})^{-1}. \tag{13}$$

Defining the filter $W(z) = \Phi_{vx}(z)\Phi_x(z)^{-1}$ a real-rational transfer matrix is obtained with no poles on the unit circle that has the specified frequency response given by (13). Thus, $\hat{v} = W(z)x \in X$ minimizes the cost (11). As a consequence of the Hilbert projection theorem (for pre-Hilbert spaces) (12) is satisfied for $\hat{v}$ if and only if it is the unique element of the subspace $X$ minimizing (11) [44]. If $\Phi_x(z) > 0$, the uniqueness of $W(z)$ follows from Lemma 23. ■

Observe that Proposition 24 allows one to replace the inf operator of (11) with a min operator.

In the following definition a notion of conditional non-causal Wiener-uncorrelation is given.

*Definition 25:* Let $v$, $x_1, \ldots, x_n$ be processes in the space $\mathcal{FE}$. Define $x := (x_1, \ldots, x_n)^T$ and $X := tf - span\{x_1, \ldots, x_n\}$. For any $i \in \{1, \ldots, n\}$, the process $v$ is conditionally non-causally Wiener-uncorrelated with $x_i$ given the processes $\{x_k\}_{k\neq i}$ if the $i$th entry of the Wiener filter to estimate $v$ from $x$ is zero, that is $\Phi_{vx}\Phi_x^{-1}b_i = 0$. where $b_i$ is the vector of $\mathbb{R}^n$ that has 1 as the $i$th entry and 0 in all other entries.

The following lemma provides an immediate relationship between non-causal Wiener-uncorrelation and the inverse of the cross-spectral density matrix. This result presents strong similarities with the property of the inverse of the covariance matrix for jointly Gaussian random-variables. Indeed, it is well-known that the entry $(i, j)$ of inverse of the covariance matrix of $n$ random variables $x_1, \ldots, x_n$ is zero if and only if $x_i$ and $x_j$ are conditionally independent given other variables. This also forms the reason for Definition 25.

*Lemma 26:* Let $\mathcal{E}$ be a set of rationally related processes and let $x_1, \ldots, x_n$ be processes in the space $\mathcal{FE}$. Define $x = (x_1, \ldots, x_n)^T$. Assume that $\Phi_x$ has full normal rank. The process $x_i$ is non-causally Wiener-uncorrelated with $x_j$ given the processes $\{x_k\}_{k\neq i,j}$, if and only if the entry $(i, j)$, or equivalently the entry $(j, i)$, of $\Phi_x^{-1}(z)$ is zero, that is, for $i \neq j$

$$b_j^T\Phi_x^{-1}b_i = b_i^T\Phi_x^{-1}b_j = 0. \tag{14}$$

*Proof:* Without any loss of generality, let $j=n$ and define $x_{\overline{n}} := (x_1,\ldots,x_{n-1})^T$. Suppose the non-causal Wiener filter estimating $x_n$ from $x_{\overline{n}}$ is $W_{n\overline{n}}$. Then

$$x_n = \varepsilon_n + W_{n\overline{n}}(z)x_{\overline{n}} \tag{15}$$

where, from (12), the error $\varepsilon_n$ has the property that $\Phi_{\varepsilon_n x_{\overline{n}}}(z)=0$. Define $r := (x_{\overline{n}}^T, \varepsilon_n)^T$ and observe that

$$r = \begin{pmatrix} \mathcal{I} & 0 \\ -W_{n\overline{n}}(z) & 1 \end{pmatrix} x; \qquad x = \begin{pmatrix} \mathcal{I} & 0 \\ W_{n\overline{n}}(z) & 1 \end{pmatrix} r.$$

It follows that

$$\begin{aligned}\Phi_x^{-1} &= \begin{pmatrix} \mathcal{I} & W_{n\overline{n}}(z)^* \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Phi_{x_{\overline{n}}}^{-1} & 0 \\ 0 & \Phi_{\varepsilon_n}^{-1} \end{pmatrix} \begin{pmatrix} \mathcal{I} & 0 \\ W_{n\overline{n}}(z) & 1 \end{pmatrix} \\ &= \begin{pmatrix} \Phi_{x_{\overline{n}}} + \frac{W_{n\overline{n}}^* W_{n\overline{n}}}{\Phi_{\varepsilon_n}} & W_{n\overline{n}}^* \Phi_{\varepsilon_n}^{-1} \\ \Phi_{\varepsilon_n}^{-1} W_{n\overline{n}} & \Phi_{\varepsilon_n}^{-1} \end{pmatrix}.\end{aligned}$$

The assertion is proven by pre-multiplying by $b_n^T$ and post-multiplying by $b_i$. ∎

The following theorem provides a sufficient condition to determine if two nodes in a LDG are kins.

*Theorem 27:* Consider a well-posed and topologically detectable LDG $(H(z),e)$ with associated graph $G$. Let $x=(x_1,\ldots,x_n)^T$ be its output. Define the space $X_j = tf-span\{x_i\}_{i\neq j}$. Consider the problem of approximating the signal $x_j$ with an element $\hat{x}_j \in X_j$ (non-causal Wiener filtering), defined as follows:

$$\min_{\hat{x}_j \in X_j} \|x_j - \hat{x}_j\|^2. \tag{16}$$

Then the optimal solution $\hat{x}_j$ exists, is unique and

$$\hat{x}_j = \sum_{i\neq j} W_{ji}(z)x_i \tag{17}$$

where $W_{ji}(z)\neq 0$ implies $\{x_i,x_j\}\in kin(G)$.

*Proof:* The LDG dynamics is given by $x=(\mathcal{I}-H(z))^{-1}e$ implying that

$$\Phi_x^{-1} = (\mathcal{I}-H)^*\Phi_e^{-1}(\mathcal{I}-H). \tag{18}$$

Consider the entry $(i,j)$ of $\Phi_x^{-1}$, with $i\neq j$. Recalling that $\Phi_e$ is diagonal, we have

$$\begin{aligned} b_j^T\Phi_x^{-1}b_i &= \left(b_j^T - (H_{*j})^*\right)\Phi_e^{-1}(\mathcal{I}-H)b_i \\ &= b_j^T b_i - b_j^T\Phi_e^{-1}H_{*i} - (H_{*j})^*\Phi_e^{-1}b_i + (H_{*j})^*\Phi_e^{-1}(H_{*i}) \\ &= -\Phi_{e_j}^{-1}b_j^T H_{*i} - (H_{*j})^* b_i \Phi_{e_i}^{-1} + \sum_{k=1}^n (H_{kj})^*\Phi_{e_k}^{-1}(H_{ki}) \\ &= -\Phi_{e_j}^{-1}H_{ji} - H_{ij}^*\Phi_{e_i}^{-1} + \sum_{k=1}^n \Phi_{e_k}^{-1}(H_{kj})^*(H_{ki}). \end{aligned}$$

The first two terms in the last expression are zero if $x_i$ and $x_j$ are not directly linked. The third term is zero if there is no $x_k$ that is a child of both $x_i$ and $x_j$. Thus, if $x_j$ and $x_i$ are not kins, the entry $(i,j)$ of $\Phi_x^{-1}$ is null. Using Lemma (26) the assertion is proven. ∎

The following result provides a sufficient condition for the reconstruction of a link in a LDG.

*Corollary 28:* Consider a well-posed and topologically detectable LDG $\mathcal{G}$ with associated graph $G$. Let $x=(x_1,\ldots,x_n)^T$ be its output. Let $W_{ji}(z)$ be the entry of the non-causal Wiener filter estimating $x_j$ from $\{x_k\}_{k\neq j}$ corresponding to the process $x_i$. If $\mathcal{G}$ is self-kin, then $W_{ji}(z)\neq 0$ implies $(x_j,x_i)\in top(G)$.

*Proof:* Since $\mathcal{G}$ is self-kin, $\mathcal{P}_G(x_j)\cup\mathcal{C}_G(x_j)\cup\mathcal{P}_G(\mathcal{C}_G(x_j)) = \mathcal{C}_G(x_j)\cup\mathcal{P}_G((x_j))$. Thus, from the previous theorem the assertion follows immediately. ∎

The following lemma is a key result to explicitly determine the expression of the Wiener filter for a LDG in the non-causal and in the causal scenarios.

*Lemma 29:* Consider a well-posed LDG $\mathcal{G}=(H(z),e)$ with associated graph $G$ and output $x=(x_1,\ldots,x_n)^T$. Fix $j\in\{1,\ldots,n\}$ and define the set

$$C := \{c | x_c \in \mathcal{C}_G(x_j)\} = \{c_1,\ldots c_{n_c}\}$$

containing the indexes of the $n_c$ children of $x_j$. Then, for $i\neq j$,

$$x_i \in tf-span\left\{\left\{\bigcup_{k\in C}(e_k+H_{kj}(z)e_j)\right\}\cup\left\{\bigcup_{k\notin C\cup\{j\}}\{e_k\}\right\}\right\}.$$

Furthermore, if $\mathcal{G}$ is causal,

$$x_i \in c-tf-span\left\{\left\{\bigcup_{k\in C}(e_k+H_{kj}(z)e_j)\right\}\cup\left\{\bigcup_{k\notin C\cup\{j\}}\{e_k\}\right\}\right\}.$$

*Proof:* Define

$$\begin{aligned} \varepsilon_j &:= 0 & & \\ \varepsilon_k &:= e_k + H_{kj}(z)e_j & if\ k\in C & \quad (19)\\ \varepsilon_k &:= e_k & if\ k\notin\{C\}\cup\{j\} & \\ \xi_k &:= \sum H_{ki}(z)x_i & if\ k=j & \\ \xi_k &:= x_k & if\ k\neq j & \quad (20) \end{aligned}$$

and, by inspection, observe that

$$[\mathcal{I}-H(z)]\,\xi = \varepsilon.$$

Since $\mathcal{G}$ is well posed, $[\mathcal{I}-H(z)]$ is invertible implying that the signals $\{\xi_i\}_{i=1,\ldots,n}$ are a linear transformation of the signals $\{\varepsilon_i\}_{i=1,\ldots,n}$. For $i\neq j$, we have

$$x_i = \xi_i \in \text{tf}-\text{span}\{\varepsilon_k\}_{k=1,\ldots,n} = \text{tf}-\text{span}\{\varepsilon_k\}_{k\neq j}$$

where the first equality follows from (20) and last equality follows from $\varepsilon_j=0$. The causality of $\mathcal{G}$ also implies that $x_i = \text{c}-\text{tf}-\text{span}\{\varepsilon_k\}_{k\neq j}$. This proves the assertion. ∎

In the case of a well-posed and topologically detectable LDG, the following theorem completes Theorem 27 providing an explicit expression for the non-causal Wiener filter $W$ since the solution of the minimization problem (16) is known to be unique.

*Theorem 30:* Consider a well posed and topologically detectable LDG $\mathcal{G}=(H(z),e)$ with associated graph $G$. Let $x=(x_1,\ldots,x_n)^T$ be the its output. Define, for $j=1,\ldots,n$

$$\hat{C}_{j*}(z) = \frac{\Phi_{e_j}(z)H_{*j}^*(z)\Phi_e^{-1}(z)}{1+\|H_{*j}^*(z)\Phi_e^{-1}(z)H_{*j}(z)\|\Phi_{e_j}(z)} \tag{21}$$

$$\hat{P}_{j*}(z) = \left(1 - \hat{C}_{j*}(z)H_{*j}(z)\right) H_{j*}(z) \tag{22}$$

$$\hat{K}_{ji}(z) = \begin{cases} 0, & \text{if } i = j \\ -\hat{C}_{j*}(z)H_{*i}(z), & \text{otherwise.} \end{cases} \tag{23}$$

Consider the estimate $\hat{x}_j$ for $x_j$ provided by the non-causal Wiener filter using the other signals $\{x_i\}_{i\neq j}$

$$\hat{x}_j = \sum_{i\neq j} W_{ji}(z)x_i.$$

Then, $W_{ji}(z) = \hat{C}_{ji}(z) + \hat{P}_{ji}(z) + \hat{K}_{ji}(z)$.

*Proof:* Fixed $j$, define the following set of indexes:

$$\begin{aligned} C &:= \{c | x_c \in \mathcal{C}_G(x_j)\} = \{c_1, \ldots c_{n_c}\} \\ P_l &:= \{p | x_p \in \mathcal{P}_G(x_l)\} / \{j\} = \{p_{l1}, \ldots p_{ln_{pl}}\} \\ K &:= \{k | x_k \in \mathcal{K}_G(x_j)\} = \{k_1, \ldots k_{n_k}\} \end{aligned}$$

for any $l = 1, \ldots, n$.

The set $C$ contains all the indexes of the $n_c$ children of $x_j$ while the set $K$ contains all the $n_k$ indexes of the kins of $x_j$. The set $P_l$ contains all the $n_{p_l}$ parents (except $x_j$) of a generic node $x_l$.

For any $i \neq j$, define

$$\varepsilon_i := x_i - \sum_{k\neq j} H_{ij}(z)H_{jk}(z)x_k - \sum_{k\neq j} H_{ik}(z)x_k. \tag{24}$$

Note that, for any $i \neq j$, $\varepsilon_i = e_i + H_{ij}(z)e_j$. Also note that

$$e_j := x_j - \sum_k H_{jk}(z)x_k \tag{25}$$

Consider an auxiliary process $\sigma_j$ such that has the same spectral density of $e_j$ (that is $\Phi_{\sigma_j}(z) = \Phi_{e_j}(z)$) but is independent from each component of $e$ (that is $\Phi_{\sigma_j e}(z) = 0$). Define $\sigma := (\sigma_1, \ldots, \sigma_n)^T$, where $\sigma_i := \varepsilon_i$ for $i \neq j$. Observe that

$$\begin{aligned} \Phi_\sigma(z) &= \Phi_{e_j}(z)H_{*j}(z)H^*_{*j}(z) + \Phi_e(z) \\ \Phi_{e_j\sigma} &= \Phi_{e_j}(z)H^*_{*j}(z). \end{aligned}$$

Consider the solution $\hat{e}_j$ to the minimization problem

$$\hat{e}_j = \arg \min_{q\in \text{tf}-\text{span}\{\sigma_i\}} \|e_j - q\|^2.$$

By applying Theorem 24, we find that

$$\begin{aligned} \hat{e}_j &= \Phi_{e_j\sigma}(z)\Phi^{-1}_\sigma(z)\sigma \\ &= \Phi_{e_j}(z)H^*_{*j}(z) \left[\Phi_{e_j}(z)H_{*j}(z)H^*_{*j}(z) + \Phi_e(z)\right]^{-1} \sigma. \end{aligned}$$

Since

$$\hat{C}_{j*}(z) \left[\Phi_{e_j}(z)H_{*j}(z)H^*_{*j}(z) + \Phi_e(z)\right] = \Phi_{e_j}(z)H^*_{*j}(z)$$

we obtain that

$$\hat{e}_j = \sum_{i=1}^n \hat{C}_{ji}(z)\sigma_i$$

where it can be observed that $\hat{C}_{jj}(z) = 0$ follows from the diagonal structure of $\Phi_e(z)$. Thus, we have that

$$\hat{e}_j = \arg \min_{q\in \text{tf}-\text{span}\{\varepsilon_i\}_{i\neq j}} \|e_j - q\|^2.$$

Now, let us consider the problem

$$\hat{x}_j = \arg \min_{q\in \text{tf}-\text{span}\{x_i\}_{i\neq j}} \|x_j - q\|.$$

From (25), its solution $\hat{x}_j$ satisfies

$$\begin{aligned} \hat{x}_j &= \sum_k H_{jk}(z)x_k + \arg \min_{q\in \text{tf}-\text{span}\{x_i\}_{i\neq j}} \|e_j - q\| \\ &= \sum_k H_{jk}(z)x_k + \arg \min_{q\in \text{tf}-\text{span}\{\varepsilon_i\}_{i\neq j}} \|e_j - q\| \end{aligned}$$

where the last equality has been obtained by using Lemma 29. Thus, we have

$$\hat{x}_j = \sum_{i\neq j} W_{ji}x_i = \sum_k H_{jk}(z)x_k + \sum_{i\neq j} \hat{C}_{ji}\varepsilon_i.$$

Substituting the expression of $\varepsilon_i$, $i \neq j$, as a function of $x_i$, $i \neq j$, the assertion is proven. ■

Theorem 30 gives the expression of the entry $W_{ji}(z)$ of the non-causal Wiener filter as the sum of the three components $\hat{C}_{ji}$, $\hat{P}_{ji}$, and $\hat{K}_{ji}$. The three components have a graphical interpretation: $\hat{C}_{ji}(z)$ is the contribution corresponding to the fact that $x_i$ is a child of $x_j$; $\hat{P}_{ji}(z)$ is the contribution when $x_i$ is a parent of $x_j$; and $\hat{K}_{ji}(z)$ is a term present when there is a "co-parent" relation.

*Lemma 31:* Given a well posed LDG, consider (21), (22), and (23). Then:

- $\hat{C}_{ji}(z) \neq 0$ if and only if $x_i \in \mathcal{C}_G(x_j)$.
- $\hat{P}_{ji}(z) \neq 0$ if and only if $x_i \in \mathcal{P}_G(x_j)$.
- $\hat{K}_{ji}(z) \neq 0$ implies $x_i \in \mathcal{P}_G(\mathcal{C}_G(x_j)) \setminus x_j$.

*Proof:* Since $\Phi_e(z)$ is diagonal it follows that $\hat{C}_{ji} \neq 0$ for any $x_i \in \mathcal{C}_G(x_j)$. Furthermore, it follows

$$\left\|\hat{C}_{j*}(e^{i\omega})H_{*j}(e^{i\omega})\right\| < 1 \tag{26}$$

implying, from (22), that $\hat{P}_{ji}(e^{i\omega}) \neq 0$, for any $x_i \in \mathcal{P}_G(x_j)$. ■

From Lemma 31, Theorem 27 can be recovered, since $\hat{C}_{ji}(z) = \hat{P}_{ji}(z) = \hat{K}_{ji}(z) = 0$ for all the nodes $x_i$ that are not kin of $x_j$. Lemma 31 also shows under which conditions a link could remain undetected. Indeed, both the expressions $\hat{C}_{ji}(z) + \hat{P}_{ji}(z) + \hat{K}_{ji}(z)$ and $\hat{C}_{ij}(z) + \hat{P}_{ij}(z) + \hat{K}_{ij}(z)$ must vanish. Examples under which a true link remains undetected are pathological, if $x_i$ and $x_j$ are simple kins.

## V. Sparsity of Causal Filtering Operators

First, we need to introduce the following lemma.

*Lemma 32:* Let $\mathcal{E}$ be a space of rationally related processes and let $v$ and $x_1, \ldots, x_n$ be processes in $\mathcal{FE}$. Define $x := (x_1, \ldots, x_n)^T$. Assume that $\Phi_{vx}(e^{i\omega}) = 0$ for all $\omega \in [-\pi, \pi]$. Then $\langle v, q\rangle = 0$ for all $q \in tf - span\{x_i\}_{i=1,\ldots,n}$.

*Proof:* As $q \in \mathrm{tf}-\mathrm{span}\{x_i\}_{i=1,\ldots,n}$, it follows that there exist $\alpha_i(z) \in \mathcal{F}$ such that

$$q = \sum_{i=1}^{n} \alpha_i(z)x_i =: \alpha(z)x$$

where $\alpha(z) = (\alpha_1(z), \ldots, \alpha_n(z))$ is a row vector of real-rational transfer functions. Then it follows that

$$\langle v, q\rangle = \int_{-\pi}^{\pi} \Phi_{vx}(e^{i\omega})\alpha(e^{i\omega})^* = 0.$$

■

Now, a specific formulation of the standard causal Wiener filter (see [46]) is introduced for the defined spaces.

*Proposition 33:* Let $v$ and $x_1, \ldots, x_n$ be processes in the space $\mathcal{FE}$. Define $x := (x_1, \ldots, x_n)^T$ and $X := c - tf - span\{x_1, \ldots, x_n\}$. Consider the problem

$$\inf_{q \in X} \|vs - q\|^2. \tag{27}$$

Let $S(z)$ be the spectral factorization of $\Phi_x(e^{i\omega}) = S(e^{i\omega})S^*(e^{i\omega})$. If $\Phi_x(e^{i\omega}) > 0$, for $\omega \in [-\pi, \pi]$, the solution $\hat{v}^{(c)} \in X$ exists, is unique and has the form $\hat{v}^{(c)} = W^{(c)}(z)x$, where $W^{(c)}(z) = \{\Phi_{vx}(z)\Phi_x(z)^{-1}S(z)\}_C S^{-1}(z)$. Moreover $\hat{v}^{(c)}$ is the unique element in $X$ such that, for any $q \in X$, satisfies $\langle v - \hat{v}^{(c)}, q\rangle = 0$.

*Proof:* This proof follows the standard derivation of the Wiener–Hopf filter (see [46]) and it is reported for the sake of completeness. Observe that $W^{(c)}(z)$ is rational and causal. Let us prove that, for any causal $H(z)$, we have

$$\langle v - \hat{v}^{(c)}, H(z)x\rangle = 0. \tag{28}$$

Define $r := S^{-1}(z)x$, and observe that $S(z), S^{-1}(z) \in \mathcal{F}^+$, since $S(z)$ is the spectral factor of $\Phi_x(z)$ that is real-rational and has full rank on the unit circle. The signal $r$ is white since $\Phi_r(z) = \mathcal{I}_n$. Observe that

$$\begin{aligned}\hat{v}^{(c)} &= \left\{\Phi_{vr}(z)\Phi_r^{-1}(z)\right\}_C r = \\ &= \left\{\Phi_{vx}(z)\Phi_x^{-1}(z)S(z)\right\}_C S^{-1}(z)x\end{aligned}$$

and observe that $\hat{v}^{(c)} \in X$ since it is obtained by composing transformations in $\mathcal{F}^+$. Also observe that $\Phi_{vr}(z)\Phi_r^{-1}(z)$ is the non-causal Wiener filter estimating $v$ from $r$. Let the strictly anti-causal component of such a filter be

$$\left\{\Phi_{vr}(z)\Phi_r^{-1}(z)\right\}_{SA} := \Phi_{vr}(z)\Phi_r^{-1}(z) - \left\{\Phi_{vr}(z)\Phi_r^{-1}(z)\right\}_C.$$

Now, for a $H(z) \in \mathcal{F}^+$, compute

$$\begin{aligned}&\langle v - \hat{v}, H(z)x\rangle = \left\langle v - W^{(c)}x, H(z)S(z)r\right\rangle = \\ &\left\langle x - \Phi_{vr}(z)\Phi_r^{-1}(z)r + \left\{\Phi_{vr}(z)\Phi_r^{-1}(z)\right\}_{AC} r, HSr\right\rangle = \\ &\left\langle\left\{\Phi_{xr}(z)\Phi_r^{-1}(z)\right\}_{SA} r, HSr\right\rangle = 0\end{aligned}$$

where Lemma 32 has been used to prove that $\langle x - \Phi_{xr}(z)\Phi_r^{-1}(z)r, H(z)S(z)r\rangle = 0$ and the last equality follows from the causality of $H(z)S(z)$, the strict anti-causality of $\{\Phi_{vr}(z)\Phi_r^{-1}(z)\}_{SA}$ and the whiteness of $r$. From the Hilbert projection theorem, $\hat{v}^{(c)}$ is the unique process minimizing the cost (27). ■

The following theorem proves the sparsity of the causal Wiener filter stating that the causal Wiener filter estimating $x_j$ from the signals $x_i$, $i \neq j$, has nonzero entries corresponding to the kin signals of $x_j$.

*Theorem 34:* Consider a well-posed, causal and topologically detectable LDG. Let $x_1, \ldots, x_n \in \mathcal{FE}$ be the signals associated with the $n$ nodes of its graph. Define $X_j = c - tf - span\{x_i\}_{i \neq j}$. Consider the problem of approximating the signal $x_j$ with an element $\hat{x}_j \in X_j$ (causal Wiener filtering), as defined as follows:

$$\min_{\hat{x}_j \in X_j} \|x_j - \hat{x}_j\|^2.$$

Then the optimal solution $\hat{x}_j$ exists, is unique and

$$\hat{x}_j = \sum_{i \neq j} W_{ji}(z)x_i$$

where $W_{ji}(z) \neq 0$ implies $(x_i, x_j) \in kin(G)$.

*Proof:* For any $i \neq j$, define $\varepsilon_i$ as in (24) and observe that $\varepsilon_i$ can be represented as in (19). Also note that

$$e_j := x_j - \sum_i H_{ji}(z)x_i. \tag{29}$$

Consider $\hat{e}_j$ defined as

$$\hat{e}_j := \arg \min_{q \in \mathrm{c-tf-span}\{\varepsilon_i\}_{i \neq j}} \|e_j - q\| = \sum_{i \neq j} C_{ji}^{(c)}(z)\varepsilon_i$$

where the transfer functions $C_{ji}^{(c)}(z)$ are given by the causal Wiener filter estimating $e_j$ from $\{\varepsilon_i\}_{i \neq j}$. Notice that, by (19), $C_{ji}^{(c)}(z)$ is equal to zero if $x_i$ is not a child of $x_j$. Now, let us consider the optimization problem

$$\hat{x}_j := \arg \min_{q \in \mathrm{c-tf-span}\{x_i\}_{i \neq j}} \|x_j - q\| = \sum_{i \neq j} W_{ji}(z)x_i$$

where $W_{ji}(z)$ are the entries of the associated causal Wiener filter. Its solution $\hat{x}_j$ satisfies

$$\begin{aligned}\hat{x}_j &= \sum_{i \neq j} H_{ji}(z)x_i + \arg \min_{q \in \mathrm{c-tf-span}\{x_i\}_{i \neq j}} \|e_j - q\| \\ &= \sum_i H_{ji}(z)x_i + \arg \min_{q \in \mathrm{c-tf-span}\{\varepsilon_i\}_{i \neq j}} \|e_j - q\|\end{aligned}$$

where the first equality derives from (29) and the last one has been obtained by using Lemma 29. Thus, we have

$$\hat{x}_j = \sum_{i \neq j} W_{ji}x_i = \sum_i H_{ji}(z)x_i + \sum_{i \neq j} C_{ji}\varepsilon_i.$$

Substituting the expression of $\varepsilon_i$, $i \neq j$, as a function of $x_i$, $i \neq j$, the assertion is proven. ■

The following theorem proves the sparsity of the one step prediction operator (or Granger-causal operator). Granger-causality is a widespread technique in econometrics to test the causal dependence of time series. If the stronger hypothesis of strictly causal transfer functions $H_{ji}(z)$ is met, one-step predictor provides an exact reconstruction of parent-child links in a LDG.

*Theorem 35:* Consider a well-posed, strictly causal and topologically detectable LDG. Let $x_1, \ldots, x_n \in \mathcal{FE}$ be the signals associated with the $n$ nodes of its graph. Define $X_j = c - tf - span\{x_1, \ldots, x_n\}$. Consider the problem of approximating the signal $zx_j$ with an element $\hat{x}_j \in X_j$, as defined below

$$\min_{\hat{x}_j \in X_j} \|zx_j - \hat{x}_j\|^2. \tag{30}$$

Then the optimal solution $\hat{x}_j$ exists, is unique and

$$\hat{x}_j = \sum_{i=1}^{n} W_{ji}(z) x_i \tag{31}$$

where $W_{ji}(z) \neq 0$ implies $i = j$ or $x_i$ is a parent of $x_j$.

*Proof:* Consider the minimization problem

$$\hat{e}_j := \arg \min_{q \in \mathrm{c-tf-span}\{e_i\}} \|ze_j - q\| = \sum_{i=1}^{n} C_{ji}^{(g)}(z) e_i$$

where the transfer functions $C_{ji}^{(g)}(z)$ are elements of $\mathcal{F}^+$. We have that $C_{ji}^{(g)}(z) = 0$ for any $i \neq j$. Indeed, since $\Phi_{e_i e_j}(e^{i\omega}) = 0$ for $i \neq j$, it holds that

$$\arg \min_{q \in \mathrm{c-tf-span}\{e_i\}_{i=1}^n} \|ze_j - q\| = \arg \min_{q \in \mathrm{c-tf-span}\{e_j\}} \|ze_j - q\|.$$

Now, let us consider the problem

$$\arg \min_{q \in \mathrm{c-tf-span}\{x_i\}_{i=1}^n} \|zx_j - q\|.$$

Noting that

$$e_j := x_j - \sum_i H_{ji}(z) x_i \tag{32}$$

we find that its solution $\hat{x}_j$ is

$$\begin{aligned}\hat{x}_j &= \sum_k zH_{jk}(z)x_k + \arg \min_{q \in \mathrm{c-tf-span}\{x_i\}_{i=1}^n} \|ze_j - q\| \\ &= \sum_k zH_{jk}(z)x_k + \arg \min_{q \in \mathrm{c-tf-span}\{e_i\}_{i=1}^n} \|ze_j - q\| \\ &= C_{jj}^{(g)}(z)x_j + \sum_{k \neq j} \left[zH_{jk}(z) - C_{jj}^{(g)}(z)H_{jk}(z)\right] x_k.\end{aligned}$$

This proves the assertion. ■

## VI. A Reconstruction Algorithm

The previous section provides theoretical results allowing for the reconstruction of a topology via Wiener filtering. It needs to be stressed that even in the case of sparse graphs, the reconstruction of the kinship topology can be considered a practical solution. The reasons are two-fold. In many situations it is possible to measure the outputs of many nodes, while it is important to identify a reduced number of possible interconnections among those nodes. For example, DNA-microarrays are devices that allow the measurement of gene expression of a cell. Such data can be useful in understanding which genes interact together and realize a specific metabolic pathway and how they are related [8]. Indeed, a cell can express tens of thousands of genes while only a few tens are typically involved in a gene regulatory network. The possibility of reducing the set of involved genes to test the presence of actual interactions with targeted experiments is of significant importance [47]. Analogously, in Finance, quantifying the strongest interconnections among a set of market stocks can suggest good strategies to balance a given portfolio [6]. Thus, it is important to have a quantitative tool to group together different stocks or, at least, to detect a limited set of possible dynamical connections. Similar problems are also present in neuroscience in order to understand neural interconnections [10]. A second reason why the presented analysis is of practical importance is that as a byproduct of the reconstruction an optimal model for the node dynamics which can be used for smoothing procedures (in the case non-causal Wiener filter are derived) or predictive ones (in the strictly causal case which is out of the scope of this paper) can be obtained.

The following algorithm is a pseudo-code implementation of the reconstruction technique that was developed in the previous section.

`Reconstruction algorithm`

0. Initialize the set of edges $A = \{\}$

1. For any signal $x_j$

2. Determine the Optimal filter entries $W_{ji}(z)$ (non-causal Wiener, causal Wiener or one-step predictor)

3. For any $W_{ji}(z) \not\cong 0$

4. add $\{x_i, x_j\}$ to $A$

5. end

7. end

8. return $A$

Under the assumption of ergodicity of the signals, there are a variety of techniques to perform step 2 using data. Most of them rely on estimating the spectral densities of the signal involved. For example an efficient technique based on Gram–Schmidt orthogonalization is described in [48]. In this case, for a fixed node $x_j$, the computational cost associated with determining a Wiener filter based on the other $n-1$ signals is $\mathcal{O}(n^2 m \log(m))$ where $m$ is the length of each measured sequence. Since the algorithm repeats the procedure for each of the $n$ nodes, the computational cost is given by $\mathcal{O}(n^3 m \log(m))$.

The theoretical results that we have derived rely on determining the nonzero components $W_{ji}(z)$ in order to detect a link between $x_i$ and $x_j$. When dealing with real data the identification of those entries can not be exact; thus, step 3 cannot be implemented as a strict condition $W_{ji}(z) \neq 0$. By $W_{ji}(z) \not\cong 0$, it is meant that $W_{ji}(z)$ needs to be "significantly" different from zero. This can be done in practice by implementing a condition of the form $\|W_{ji}(z)\| > \sigma_{thr}(x_i, x_j)$ where $\sigma_{thr}(x_i, x_j)$ is a threshold that might depend on both the signals $x_i$ and $x_j$, and a proper norm needs to be defined on the space of transfer functions. In the case of a simple model given for the measurement and estimation errors (in the non-causal scenario), a quantitative analysis to choose the value of the threshold for a given norm is carried on in the next section.

## VII. Discussion on the Robustness of the Reconstruction

In the previous sections, an ideal scenario was analyzed, where the node signals are affected by the processes $\{e_j\}$ which are driving the network. This has led to the result that the Wiener filter (in different formulations) is a useful tool for revealing the structure of a network. In certain situations, the reconstruction is exact, while in the general scenario only the smallest self-kin network that contains the original system is correctly identified.

All the presented techniques rely on determining the zero entries of an optimal estimator that needs to be computed from data. However, in practical situation optimal estimators cannot be computed exactly from realizations of the observed stochastic processes $\{x_j\}$.

Many factors such as numerical errors, limited amounts of data, measurement noise lead to non-exact computations.

Aim of this section is to show that the developed techniques offer a certain degree of robustness against an inaccurate computation of the Wiener filter. In particular, it will be shown that, in the presence of sufficiently small measurement noise (a simple model for all the inaccuracy sources mentioned before), the introduction of a threshold in the algorithm again provides theoretical guarantees about the reconstruction of the topology of the minimal self-kin network.

*Definition 36:* A Corrupted LDG (CLDG) is a pair $(\mathcal{G},\eta)$ where $\mathcal{G}=(H(z),e)$ is a LDG with output $\{x_1,\dots,x_n\}$ and $\eta=(\eta_1,\dots,\eta_n)^T$ is a set of noises with the property that they are mutually not correlated and not correlated with the signals $\{e_j\}$, either. The output of the CLDG are the signals $\{y_1,\dots,y_n\}$ defined as $y_j=x_j+\eta_j$.

The following important lemma about the inverse of a matrix is recalled.

*Lemma 37:* If $Q$, $\Delta$ and $Q+\Delta$ are invertible matrices the following equality holds:

$$(Q+\Delta)^{-1}-Q^{-1}=-Q^{-1}\left(Q^{-1}+\Delta^{-1}\right)^{-1}Q^{-1}.$$

The following result establishes a bound on the difference between the topological filter $W(z)$ which can be obtained by measuring directly the dynamics of a LDG and the corrupted topological filter $\hat{W}(z)$ obtained by measuring its output.

*Theorem 38:* Consider a CLDG $(\mathcal{G},\eta)$ with an $n$ nodes $\{x_1,\dots,x_n\}$. Fix $0<j\le n$ and let

$$\hat{x}_j=\sum_{i\neq j}W_{ji}(z)x_i,\qquad \hat{y}_j=\sum_{i\neq j}\hat{W}_{ji}(z)x_i \tag{33}$$

be respectively the minimum least square estimate of $x_j$ using the signals $\{x_i\}_{i\neq j}$ and the minimum least square estimate of $y_j$ using the signals $\{y_i\}_{i\neq j}$. Define $I_j:=(1,\dots,j-1,j+1,\dots,n)$. If, for any $z\in\mathbf{C}$

$$\frac{1}{\|\Phi^{-1}_{y_{I_j}}(z)\|}-\|\Phi_\eta(z)\|>0$$

then,

$$\left\|\hat{W}_{jI_j}(z)-W_{jI_j}(z)\right\|\le\frac{\left\|\Phi_{y_jy_{I_j}}(z)\right\|\left\|\Phi_{y_{I_j}}(z)^{-1}\right\|\|\Phi_\eta(z)\|}{\left(\frac{1}{\left\|\Phi^{-1}_{y_{I_j}}(z)\right\|}-\|\Phi_\eta(z)\|\right)}.$$

*Proof:* First note that $\Phi_{y_jy_{I_j}}(z)=\Phi_{x_jx_{I_j}}(z)$, thus

$$\begin{aligned}\hat{W}_{jI_j}(z)&-W_{jI_j}(z)\\&=\Phi_{y_jy_{I_j}}(z)\left[\Phi^{-1}_{y_{I_j}}(z)-\Phi^{-1}_{x_{I_j}}(z)\right]\\&=\Phi_{y_jy_{I_j}}(z)\left[\Phi^{-1}_{y_{I_j}}(z)-\left(\Phi_{y_{I_j}}(z)-\Phi_{\eta_{I_j}}(z)\right)^{-1}\right].\end{aligned} \tag{34}$$

By applying Lemma 37, it follows that

$$\begin{aligned}Q^{-1}-(Q-\Delta)^{-1}&=Q^{-1}\left(Q^{-1}-\Delta^{-1}\right)^{-1}Q^{-1} &(35)\\&=Q^{-1}(\Delta-Q)^{-1}\Delta, &(36)\end{aligned}$$

which implies

$$\begin{aligned}\left\|Q^{-1}-(Q-\Delta)^{-1}\right\|&\le\|Q^{-1}\|\|\Delta\|\left\|(Q-\Delta)^{-1}\right\| &(37)\\&=\frac{\|Q^{-1}\|\|\Delta\|}{\min sp(Q-\Delta)}\le\frac{\|Q^{-1}\|\|\Delta\|}{\min sp(Q)-\max sp(\Delta)} &(38)\\&=\frac{\|Q^{-1}\|\|\Delta\|}{\frac{1}{\|Q^{-1}\|}-\|\Delta\|}. &(39)\end{aligned}$$

The assertion is proven by applying this inequality to (34). ■

An immediate way to apply Theorem 38 is when a function $D(z)$ bounding $\|\Phi_\eta(z)\|$ is known as illustrated by the following corollary.

*Corollary 39:* Assume that there exists a real valued function $D(z)$ such that $\|\Phi_\eta(z)\|<D(z)$. If, for some $z\in\mathbf{C}$ and for any $j=1,\dots,n$,

$$\frac{1}{\left\|\Phi^{-1}_{y_{I_j}}(z)\right\|}-D(z)>0 \tag{40}$$

and

$$\left\|\hat{W}_{ji}(z)\right\|\ge\frac{\left\|\Phi_{y_jy_{I_j}}(z)\right\|\|\Phi_\eta(z)\|}{\left(\frac{1}{\left\|\Phi^{-1}_{y_{I_j}}(z)\right\|}-\|\Phi_\eta(z)\|\right)} \tag{41}$$

then $(x_j,x_i)\in kin(G)$.

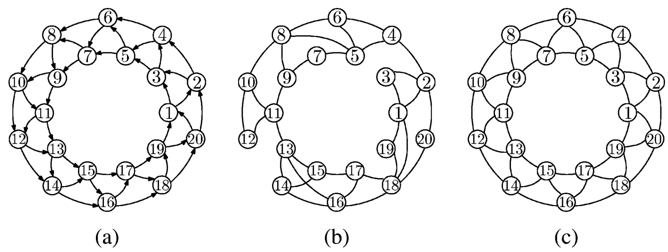

Fig. 4. Ring network of 20 nodes as the one considered in Example VIII-A (a). The reconstruction obtained using the partial correlation technique described in [25] which shows both false positives and false negatives (b). The reconstructed topology of Section VIII-A. Every single link has been correctly detected and, since the network is self-kin, the topology does not contain any false positive (c).

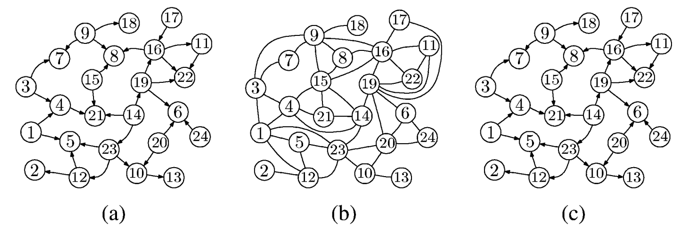

Fig. 5. A network of 24 nodes as the one considered in Section VIII-B (a), the reconstructed topology using non-causal and causal Wiener filters (b) and the reconstructed topology using the one-step predictor. In (b) every single link has been detected, but, since the network is not self-kin, as expected, the topology contains the additional links between the "kins." In (c) the reconstruction is exact (with detection of the link orientation) but this technique can be successfully applied only when the dynamic links of the network are strictly causal.

*Proof:* By contradiction, assume that $(y_j, y_i) \notin kin(G)$. Then $W_{ji}(z) = 0$. By Theorem 38 we have

$$|\hat{W}_{ji}(z)| \leq \frac{\left\|\Phi_{y_j y_{I_j}}(z)\right\| \|\Phi_\eta(z)\|}{\left(\frac{1}{\left\|\Phi^{-1}_{y_{I_j} y_{I_j}}(z)\right\|} - \|\Phi_\eta(z)\|\right)} < \frac{\left\|\Phi_{y_j y_{I_j}}(z)\right\| \|\Phi_\eta(z)\|}{\left(\frac{1}{\left\|\Phi^{-1}_{y_{I_j} y_{I_j}}(z)\right\|} - D(z)\right)}$$

which is a contradiction. ∎

Note that the inequality of Theorem 38 needs to hold for any $z$, so, in this sense, it constitutes a sharp criterion to detect the presence of a link in a CLDG.

## VIII. Numerical Examples

In this section, illustrative applications of the theoretical results are provided.

### A. Self-Kin Network

A ring network of 20 nodes is considered. The network structure is provided in Fig. 4(a) The dynamics of the links is given by randomly generated fifth-order causal finite impulse response (FIR) filters. The noise power is the same on every node. The network was simulated for 500 steps. A prevalent technique based on partial-correlation to infer the connectivity structure of a network is described in [25]. The reconstruction result provided by its implementation is reported in Fig. 4(b). As can be noticed, the reconstruction is not accurate. An implementation of the developed algorithms was applied to the data, as well. Specifically, we applied the two techniques based on non-causal and causal Wiener filtering. The reconstructed topology is the same in both cases and it is depicted in Fig. 4(c). Notice that, since the network is self-kin, every single link is detected, but no information about its orientation is recovered. Since the dynamics on the links is not strictly causal, the conditions to apply the technique based on one-step prediction are not met. Thus, an exact reconstruction of the topology cannot be expected.

### B. Generic Strictly Causal Network

Consider a network of 24 nodes for 500 steps as reported in Fig. 5(a). As a difference from the previous case, the dynamics of the links is given by randomly generated fifth-order *strictly* causal FIR filters. The conditions to apply all the three techniques reported in this article are met. The two techniques based on non-causal and causal Wiener filtering provide the same result shown in Fig. 5(b). As guaranteed by the theoretical results of this paper, the network in Fig. 5(b) is the smallest self-kin network containing the original topology. Since the link dynamics is strictly causal, the conditions to apply the technique based on one-step predictors are met and the reconstruction is exact as shown in Fig. 5(c).

## IX. Conclusion

This work has illustrated a simple but effective procedure to identify the general structure of a network of linear dynamical systems. The approach followed is based on Wiener Filtering. When the topology of the original graph is described by a self-kin network, the method developed guarantees an exact reconstruction. Self-kin networks provide a nontrivial class of networks since they allow the presence of loops, nodes with multiple inputs and disconnected sub-graphs. Moreover, the article also provides results about general networks. It is shown that, for a general graph, the developed procedure reconstructs the topology of the smallest self-kin graph containing the original one. Thus, the method is optimal in this sense. Numerical examples illustrate the correctness and also the reliability of the identification technique.

## Acknowledgment

D. Materassi would like to thank L. Ning for useful discussions.

## References

[1] A. Czirók, A.-L. Barabási, and T. Vicsek, "Collective motion of self-propelled particles: Kinetic phase transition in one dimension," *Phys. Rev. Lett.*, vol. 82, no. 1, p. 209, Jan. 1999.

[2] H. Levine, W.-J. Rappel, and I. Cohen, "Self-organization in systems of self-propelled particles," *Phys. Rev. E*, vol. 63, no. 1, p. 017101, Dec. 2000.

[3] A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1465–1476, Sep. 2004.

[4] J. Liu, V. Yadav, H. Sehgal, J. M. Olson, H. Liu, and N. Elia, "Phase transitions on fixed connected graphs and random graphs in the presence of noise," *IEEE Trans. Autom. Control*, vol. 53, no. 8, p. 1817, Sep. 2008.
[5] M. Naylora, L. Roseb, and B. Moyle, "Topology of foreign exchange markets using hierarchical structure methods," *Physica A*, vol. 382, pp. 199–208, 2007.
[6] R. Mantegna and H. Stanley*, An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[7] M. Eisen, P. Spellman, P. Brown, and D. Botstein, "Cluster analysis and display of genome-wide expression patterns," *Proc. Natl. Acad. Sci. USA*, vol. 95, no. 25, pp. 14863–14868, 1998.
[8] E. Ravasz, A. Somera, D. Mongru, Z. Oltvai, and A. Barabasi, "Hierarchical organization of modularity in metabolic networks," *Science*, vol. 297, p. 1551, 2002.
[9] D. Del Vecchio, A. Ninfa, and E. Sontag, "Modular cell biology: Retroactivity and insulation," *Nature Molec. Syst. Biol.*, vol. 4, p. 161, 2008.
[10] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura, and S. L. Bressler, "Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 26, pp. 9849–9854, Jun. 2004.
[11] A. Bunn, D. Urban, and T. Keitt, "Landscape connectivity: A conservation application of graph theory," *J. Environ. Manage.*, vol. 59, no. 4, pp. 265–278, 2000.
[12] D. Urban and T. Keitt, "Landscape connectivity: A graph-theoretic perspective," *Ecology*, vol. 82, no. 5, pp. 1205–1218, 2001.
[13] J.-S. Bailly, P. Monestiez, and P. Lagacherie, "Modelling spatial variability along drainage networks with geostatistics," *Math. Geol.*, vol. 38, no. 5, pp. 515–539, 2006.
[14] P. Monestiez, J.-S. Bailly, P. Lagacheriec, and M. Voltz, "Geostatistical modelling of spatial processes on directed trees: Application to fluvisol extent," *Geoderma*, vol. 128, pp. 179–191, 2005.
[15] G. Innocenti and D. Materassi, "A modeling approach to multivariate analysis and clusterization theory," *J. Phys. A*, vol. 41, no. 20, p. 205101, 2008.
[16] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, "Complex networks: Structure and dynamics," *Phys. Rep.*, vol. 424, no. 4–5, pp. 175–308, Feb. 2006.
[17] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proc. Nat. Acad. Sci.*, vol. 99, no. 12, 2002.
[18] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E*, vol. 69, no. 2, 2004, 026113 (pp. 1-15).
[19] H. Zhang, Z. Liu, M. Tang, and P. Hui, "An adaptive routing strategy for packet delivery in complex networks," *Phys. Lett. A*, vol. 364, pp. 177–182, 2007.
[20] R. Olfati-Saber, "Distributed kalman filtering for sensor networks," in *Proc. IEEE CDC*, New Orleans, LA, 2007, pp. 5492–5498.
[21] I. Schizas, A. Ribeiro, and G. Giannakis, "Consensus in ad hoc WSNs with noisy links-part i: Distributed estimation of deterministic signals," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 350–364, Jan. 2008.
[22] M. Timme, "Revealing network connectivity from response dynamics," *Phys. Rev. Lett.*, vol. 98, no. 22, p. 224101, 2007.
[23] S. Boccaletti, M. Ivanchenko, V. Latora, A. Pluchino, and A. Rapisarda, "Detecting complex network modularity by dynamical clustering," *Phys. Rev. E*, vol. 75, p. 045102, 2007.
[24] D. Napoletani and T. Sauer, "Reconstructing the topology of sparsely connected dynamical networks," *Phys. Rev. E*, vol. 77, p. 026103, 2008.
[25] E. Kolaczyk*, Statistical Analysis of Network Data: Methods and Models*. Berlin, Germany: Springer-Verlag, 2009.
[26] M. Ozer and M. Uzuntarla, "Effects of the network structure and coupling strength on the noise-induced response delay of a neuronal network," *Phys. Lett. A*, vol. 375, pp. 4603–4609, 2008.
[27] C. Michener and R. Sokal, "A quantitative approach to a problem of classification," *Evolution*, vol. 11, pp. 490–499, 1957.
[28] R. Freckleton, P. Harvey, and M. Pagel, "Phylogenetic analysis and comparative data: A test and review of evidence," *Amer. Natur.*, vol. 160, pp. 712–726, 2002.
[29] R. Diestel*, Graph Theory*. Berlin, Germany: Springer-Verlag, 2006.
[30] O. Ledoit and M. Wolf, "Honey, i shrunk the covariance matrix," *J. Portfolio Manage.*, pp. 110–119, 2004.
[31] M. Tumminello, F. Lillo, and R. N. Mantegna, "Shrinkage and spectral filtering of correlation matrices: A comparison via the Kullback–Leibler distance," *Acta Phys. Polonica B*, pp. 4079–4088, 2008.
[32] B. Toth and J. Kertesz, "Accurate estimator of correlations between asynchronous signals," *Physica A*, pp. 1696–1705, 2009.
[33] D. Materassi and G. Innocenti, "Unveiling the connectivity structure of financial networks via high-frequency analysis," *Phys. A: Statist. Mech. Its Applicat.*, vol. 388, no. 18, pp. 3866–3878, Jun. 2009.
[34] D. Materassi and G. Innocenti, "Topological identification in networks of dynamical systems," *IEEE Trans. Autom. Control*, vol. 55, no. 8, pp. 1860–1871, Aug. 2010.
[35] D. Marinazzo, M. Pellicoro, and S. Stramaglia, "Kernel method for nonlinear Granger causality," *Phys. Rev. Lett.*, vol. 100, p. 144103, 2008.
[36] C. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, pp. 424–438, 1969.
[37] E. Candès, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted l1 minimization," *J. Fourier Anal. Applicat.*, vol. 14, pp. 877–905, 2008.
[38] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
[39] D. Materassi and G. Innocenti, "Topological identification in networks of dynamical systems," in *Proc. IEEE CDC*, Cancun, Mexico, Dec. 2008, pp. 823–828.
[40] M. R. Jovanović and B. Bamieh, "On the ill-posedness of certain vehicular platoon control problems," *IEEE Trans. Autom. Control*, vol. 50, no. 9, pp. 1307–1321, Sep. 2005.
[41] L. Getoor, N. Friedman, B. Taskar, and D. Koller, "Learning probabilistic models of relational structure," *J. Mach. Learn. Res.*, vol. 3, pp. 679–707, 2002.
[42] N. Friedman and D. Koller, "Being Bayesian about network structure: A Bayesian approach to structure discovery in Bayesian networks," *Mach. Learn.*, vol. 50, pp. 95–126, 2003.
[43] J. Pearl*, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Waltham, MA: Morgan Kaufmann, 1988.
[44] D. G. Luenberger*, Optimization by Vector Space Methods*. Hoboken, NJ: Wiley, 1969.
[45] P. E. Caines*, Linear Stochastic Systems*. New York: Wiley, 1987.
[46] T. Kailath, A. Sayed, and B. Hassibi*, Linear Estimation*. Upper Saddle River, NJ: Prentice-Hall, 2000.
[47] D. Marinazzo, M. Pellicoro, and S. Stramaglia, "Kernel Granger causality and the analysis of dynamical networks," *Phys. Rev. E*, vol. 77, p. 056215, 2008.
[48] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inf. Theory*, vol. 44, pp. 2943–2959, 1998.

**Donatello Materassi** received the Laurea in "Ingegneria Informatica" and a "Dottorato di Ricerca" in electrical engineering/nonlinear dynamics and complex systems from Università degli Studi di Firenze, Florence, Italy.

He has been a Research Associate at University of Minnesota, Minneapolis. Currently, he is with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge. 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.

**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 a faculty member in the Electrical and Computer Engineering Department, University of Minnesota, Minneapolis. His research interests include nanotechnology, multiple-objective robust control, and distributed and structural control.

Dr. Salapaka received the 1997 National Science Foundation CAREER Award.