2010 American Control Conference
Marriott Waterfront, Baltimore, MD, USA
June 30-July 02, 2010

On the problem of reconstructing an unknown topology

Donatello Materassi Murti V. Salapaka
Department of Electrical and Computer Engineering,
University of Minnesota,
200 Union St SE, 55455, Minneapolis (MN)
mater013@umn.edu murtis@umn.edu

***Abstract*— The interest for networks of dynamical systems has been increasing in the past years, especially because of their capability of modeling and describing a large variety of phenomena and behaviors. Particular attention has been oriented towards the emergence of complicated phenomena from interconnections of simple models. In this paper we tackle, from a theoretical perspective, the problem of reconstructing the topology of an unknown network of dynamical systems. We propose a technique, based on Wiener filtering, which provides general theoretical guarantees for the detection of links in a network of dynamical systems. For a large class of network that we name "self-kin" sufficient conditions for a correct detection of a link are formulated. For networks not belonging to this class we give conditions for correct detection of links belonging to the smallest self-kin network containing the actual one.**

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: the i-th element of a vector x
W_{ji}: the entry (j,i) of a matrix W
x_v: when $v=(i_1,...,i_n)$ is a vector of natural numbers denotes the vector $(x_{i_1} \ldots x_{i_n})^T$
$|A|$: 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 stationary vector processes X and Y;
for the sake of simplicity, $R_{XY} := R_{XY}(0)$;
$R_X(\tau) := R_{XX}(\tau)$: autocovariance;
$\mathcal{Z}(\cdot)$: Zeta-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;
with abuse of notation, $\Phi_X(\omega) = \Phi_X(e^{i\omega})$.

I. Introduction

The interest for networks of dynamical systems is increasing in the past years, especially because of their capability of modeling and describing a large variety of phenomena and behaviors. Particular attention has been oriented towards the emergence of complicated phenomena from the interconnections of simple models.

The principal advantages provided by a networked system are three: a modular approach to design, the possibility of directly introducing redundancy and the realization of distributed and parallel algorithms. All these advantages have led to consider networked systems in the realization of many technological devices (see e.g. [1], [2]).

For the very same reason, it is not surprising that natural and biological systems tend to exhibit strong modularity, as well. Interconnected systems are successfully exploited to perform novel modeling approaches in many fields, such as Economics (see e.g. [3]), Biology (see e.g. [4]), Cognitive Sciences (see e.g. [5]), Ecology (see e.g. [6]) and Geology (see e.g. [7]), especially when the investigated phenomena are characterized by spatial distributions and a multivariate analysis technique is preferred [8].

Remarkably, while networks of dynamical systems were deeply studied and analyzed in physics [9], [10] and engineering [11], [12], there is a reduced number of results addressing the problem of reconstructing an unknown dynamical network, since it poses formidable theoretical and practical challenges [13]. However, the reconstruction of the link structure of a set of processes is a problem found in many other fields, and the necessity for general tools to address it is rapidly emerging (see [14], [15] and [16] and the bibliography therein for recent results).

The Unweighted Pair Group Method with Arithmetic mean (UPGMA) [17] is one of the first techniques proposed to reveal an unknown topology. It has been applied to the reconstruction of phylogenetic trees, but it has also been widely employed in other areas such as communication systems and resource allocations. UPGMA identifies a tree topology relying on the observation of the leaf nodes only, 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 network is developed in [3] for the analysis of a stock portfolio. The authors identify a tree structure which is obtained defining a correlation metric among the nodes and employing a Minimum Spanning Tree algorithm to obtain the final topology. In [18] a severe limit of this strategy is highlighted. In fact, it fails in identifying a tree network in the presence of dynamical connections or even simple delays among the processes.

An approach to the identification of more general topologies was developed in the area of Machine Learning for Bayesian dynamical networks [19]. However in this case a massive quantity of data needs to be collected in order to accurately evaluate conditional probability distributions.

More recently, in [16] and [20] interesting equivalences between the identification of a dynamical network and a l_0 sparsification problem are highlighted, suggesting the difficulty of the reconstruction procedure [21].

In this paper we rigorously formulate the problem of reconstructing a network of dynamical systems where every node represents an observable scalar signal and the dynamics is given by the connecting links. We derive sufficient conditions for the reconstruction of links in a large class of networks, which we name self-kin. Examples of self-kin networks are given by (but not limited to) trees, and ring topologies.

978-1-4244-7427-1/10/$26.00 ©2010 AACC

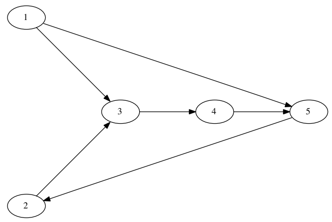


Fig. 1. An example of the graphical representation of dynamical networks. Every node represents a stochastic signal and the arc from a node N_i to a node N_j represents the transfer function $H_{ji}(z)$. Every node signal is also implicitly supposed to be affected by a process noise.

In the case the network is not self-kin, the detected links are guaranteed to belong to the self-kin network with the smallest number of links that contains the actual one. Thus, we can say that the reconstruction procedure provides an approximation which is optimal in this specific sense. The theory we develop is not Bayesian and relies on Wiener filtering theory. Indeed, a byproduct of our reconstruction procedure is a Wiener filter which models the network dynamics. The paper is organized as follows. In Section III we provide definitions recalling also standard notions of graph theory; in Section IV we cast our problem in a formal way; in Section V we develop the theory and present the main results; in Section VI we deal with implementation issues of our method; eventually, in Section VII we report numerical simulations illustrating the results of our technique.

II. Graphical Representation

In the following sections we use a convenient representation of dynamical networks in terms of oriented graphs [22] which is different from the usual graphical representation of dynamical systems.
In our representation, every node N_j represents a scalar time-discrete wide-sense stationary stochastic process x_j, while every directed arc form a node N_i to a node N_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

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

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

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

III. Preliminary definitions

In this section we will introduce a rigorous formulation of the main problem. To this aim some basic notions of graph theory, which are functional to the following developments, will be recalled as well. For an extensive overview see [22]. First we provide the standard definition of oriented graph.

Definition 1: An undirected graph G is a pair (V, A) comprising a finite set V of vertexes or nodes together with a set A of edges or arcs, which are unordered subsets of two distinct elements of V.

Definition 2: An oriented graph G is a pair (V, A) comprising a finite set V of vertexes or nodes together with a set A of edges or arcs, which are ordered pairs of elements of V.
In the following, if not specified, we will consider oriented graphs.

Definition 3: Given an oriented graph $G = (V, A)$, we define its topology 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 we write $top(G) := G'$.

Intuitively the topology of an oriented graph G is the undirected graph G' obtained removing the orientation from any edge.

Definition 4: Given a graph $G = (V, A)$ and a node $N_j \in V$, we define the children of N_j as $\mathcal{C}_G(N_j) := \{N_i | (N_j, N_i) \in A\}$ and the parents of N_j as $\mathcal{P}_G(N_j) := \{N_i | (N_i, N_j) \in A\}$.
Abusing the notation, we can consider the children and the parents of a set of nodes in the following way

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

Definition 5: Given an oriented graph $G = (V, A)$ and a node $N_j \in V$, we define the kins of N_j as

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

Abusing the notation, we can consider the kins of a set of nodes in the following way

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

Definition 6: Given an oriented graph $G = (V, A)$ and a node N_j, we say that N_i is a proper parent (child) if it is a parent (child) and $N_i \notin \mathcal{P}_\mathcal{G}(\mathcal{C}_G(N_j))$. N_i is a proper kin if it is a kin and $N_i \notin \mathcal{P}_\mathcal{G}(N_j) \cup \mathcal{C}_G(N_j)$.
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 7: Given an oriented graph $G = (V, A)$, we define its kin-graph as the undirected graph $\tilde{G} = (V, \tilde{A})$ in the following way $\tilde{A} := \{\{N_i, N_j\} | N_i \in \mathcal{K}_G(N_j)\}$ and we write $kin(G) = \tilde{G}$.

Note that the kin-graph of G is an undirected graph. It could have been defined as a directed graph, but, because of the symmetry of the kin relation, a directed graph contains exactly the same information. Moreover the choice of such a definition is motivated by the following definition

Definition 8: Given an oriented graph G, we say it is self-kin if $top(G) = kin(G)$.

Definition 9: Let $\mathcal{E}$ be a set containing time-discrete scalar, zero-mean, wide-sense, stationary random processes such that, for any entries e_i, e_j, the power spectral density $\Phi_{e_i e_j}(z)$ exists, it is real rational with no poles on the unit circle. Formally, we write

$$\Phi_{e_i e_j}(z) = \frac{A(z)}{B(z)} \tag{2}$$

with $A(z), B(z)$ real coefficient polinomials such that $B(z) \neq 0$ for any $z \in \mathbf{C}, |z| = 1$. We say that $\mathcal{E}$ is a set of rationally related random processes.

Definition 10: Define the set ${}^0\mathcal{F}$ as the set of real-rational SISO transfer function with no poles on $|z| = 1$.

Definition 11: Let $\mathcal{E}$ be a set of rationally related random processes. We define the set ${}^0\mathcal{F}\mathcal{E}$, as

$${}^0\mathcal{F}\mathcal{E} := \left\{ x = \sum_{k=1}^{m} H_k(z) e_k \mid e_k \in \mathcal{E}, H_k(z) \in {}^0\mathcal{F}, m \in \mathbb{N} \right\}.$$

The space ${}^0\mathcal{F}\mathcal{E}$ is a set of rationally related processes and, along with the inner product

$$< x_1, x_2 > := R_{x_1 x_2} = \int_{-\pi}^{\pi} \Phi_{x_1 x_2}(\omega),$$

constitutes a pre-Hilbert space with the technical assumption that two processes x_1 and x_2 are identical if $x_1(t) = x_2(t)$, almost always for any t.

For any $x \in {}^0\mathcal{F}\mathcal{E}$, we denote the natural norm induced by the inner product as $\|x\|^2 := < x, x >$. The following definition provides a class of models for a network of dynamical systems.

Definition 12: Consider the triplet $\mathcal{G} = (G, \mathcal{H}, e_I)$ where $G = (V, A)$ is a graph with n vertexes $\{N_j\}_{j=1,...,n}$; $\mathcal{H} : A \to {}^0\mathcal{F}$ is a function associating a transfer function $H_{ji}(z) \in {}^0\mathcal{F}$ to any edge $\{N_i, N_j\} \in A$; and $e_I = (e_1, ..., e_n)^T$ is a vector of n rationally related random processes in ${}^0\mathcal{F}\mathcal{E}$. We say that $\mathcal{G}$ is a Linear Dynamic Graph (LDG) if the following conditions are satisfied

- $\{N_i, N_j\} \in A$ implies $\{N_j, N_i\} \notin A$
- $\Phi_{e_i e_j}(\omega) = 0$ for $i \neq j$

We also define the output x_j of the dynamic of a LDG as

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

for any $j = 1, ..., n$. In a matrix form it is possible to write the network dynamics of an LDG as

$$x_I = H(z) x_I + e_I \tag{3}$$

where the entry (j, i) of the matrix $H(z)$ is given by the transfer function $H_{ji}(z)$. We say that the LDG is well-posed if the matrix $(I - H(z))$ is invertible, and both $(I - H(z))$ and $(I - H(z))^{-1}$ have full rank and no poles on the unit circle. This means that there is a transfer function $T(z) = [I - H(z)]^{-1}$ such that $x_I = T(z) e_I$. Interpreting $T(z)$ as a stable, possibly non-causal, linear transformation, we have that $x_i \in {}^0\mathcal{F}\mathcal{E}$ for any $i = 1, ..., n$.

Definition 13: For any finite number of elements $x_1, ..., x_m \in {}^0\mathcal{F}\mathcal{E}$ we define the operator

$$\text{span}\{x_1, ..., x_m\} := \left\{ x = \sum_{i=1}^{m} \alpha_i(z) x_i \mid \alpha_i(z) \in {}^0\mathcal{F} \right\}.$$

It is possible to prove that the span operator defines a subspace of ${}^0\mathcal{F}\mathcal{E}$.

Intuitively, 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 non correlated noises. The following definitions will be useful to determine sufficient conditions for the detection of links in a network.

Definition 14: A LDG $\mathcal{G} = (G, \mathcal{H}, e_I)$ is topologically detectable if $\Phi_{e_i}(\omega) > 0$ for any $\omega \in [-\pi, \pi]$ and for any $i = 1, ..., n$.

Definition 15: Consider a LDG $\mathcal{G} = (G, \mathcal{H}, e_I)$ with $G = (V, A)$ where $V = \{N_i\}_{i=1,...,n}$ and $X := \{x_i\}_{i=1,...,n}$ is the set of processes x_i corresponding to the nodes N_i. Define an auxiliary process x_0 that is zero-mean, white with unitary variance, such that $\Phi_{x_0} x_i(\omega) = 0$, for any ω. For any $j \in \{i = 1, ..., n\}$ compute the row vector

$$W_j(z) := \Phi_{x_j x_{I_j}}(z) \Phi^{-1}_{x_{I_j} x_{I_j}}(z). \tag{4}$$

We define the topological filter of $\mathcal{G}$ as

$$W(z) = \begin{pmatrix} \Phi_{x_1 x_{I_1}}(z) \Phi^{-1}_{x_{I_1}}(z) \\ \vdots \\ \Phi_{x_n x_{I_n}}(z) \Phi^{-1}_{x_{I_n}}(z) \end{pmatrix}. \tag{5}$$

IV. Problem Formulation

Problem 16: Consider a well-posed LDG $\mathcal{G} = (G, \mathcal{H}, e_I)$. Assume that the only available information is given by the Power (Cross-)Spectral Densities of $\{x_j\}_{i=1,...,n}$. Then, we intend to find sufficient conditions allowing one to reconstruct the unknown topology of the graph G of the LDG $\mathcal{G}$.

V. Main Results

We first introduce a specific formulation of the Wiener filter for the spaces we have defined.

Proposition 17: Let x and $x_1, ..., x_n$ be processes in the space ${}^0\mathcal{F}\mathcal{E}$. Define $I := (1, ..., n)$ and $X :=$ span$\{x_1, ..., x_n\}$. Consider the problem

$$\min_{q \in X} \|x - q\|^2. \tag{6}$$

If $\Phi_{x_I}(\omega) > 0$, the solution $\hat{x}$ exists, is unique and has the form $\hat{x} = W(z) x_I$, where $W(z) = \Phi_{x x_I}(z) \Phi_{x_I}(z)^{-1}$. Moreover, for any $q \in X$, it holds that

$$< x - \hat{x}, q > = 0. \tag{7}$$

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

$$\begin{aligned} &\|x - W(z) x_I\|^2 = \\ &= \int_{-\pi}^{\pi} \Phi_{xx}(\omega) + W(\omega) \Phi_{x_I x_I}(\omega) W^*(\omega) + \\ &\quad + \Phi_{x_I x}(\omega) W^*(\omega) + W(\omega) \Phi_{x x_I}(\omega). \end{aligned}$$

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

$$W(\omega) = \Phi_{x x_I}(\omega) \Phi_{x_I x_I}(\omega)^{-1}.$$

Defining the filter $W(z) = \Phi_{x x_I}(z) \Phi_{x_I x_I}(z)^{-1}$ we obtain a real-rational transfer matrix with no poles on the unit circle that has the specified frequency response. Thus $\hat{x} = W(z) x_i \in X$ minimizes the cost (6). Finally, since $W(z) x_I \in \text{span}\{x_i\}_{i=1,...,n}$, Equation (7) is an immediate consequence of the Hilbert projection theorem (for pre-Hilbert spaces). ■

Lemma 18: Let x and $x_1, ..., x_n$ be processes in the space ${}^0\mathcal{F}\mathcal{E}$. Define $I := (1, ..., n)$. Assume that $\Phi_{x x_I}(\omega) = 0$

for any $\omega \in [-\pi\pi]$. Then $< x, y >= 0$ for any $y \in \text{span}\{x_i\}_{i=1,...,n}$

Proof: The proof is left to the reader. ∎

Lemma 19: Let x and $x_1, ..., x_n$ be processes in the space ${}^0\mathcal{FE}$. Define $I := (1, ..., n)$ Assume that $x \in \text{span}\{x_i\}_{i=1,...,n}$ and that $\Phi_{x_I x_I}(\omega) > 0$ almost for any $\omega \in [-\pi, \pi]$. Then there exists exactly one transfer matrix $H(z)$ such that $x = H(z)x_I$.

Proof: First let us prove that $x = 0$ implies $H(z) = 0$. We have that $\Phi_{xx}(\omega) = 0 = H(\omega)\Phi_{x_I x_I}(\omega)H^*(\omega)$. Since $\Phi_{x_I x_I}(\omega) > 0$ almost everywhere, we have $H(\omega) = 0$ almost everywhere that implies $H(z) = 0$. Now, by contradiction, assume that $x = H_1(z)x_I = H_2(z)x_I$, with $H_1(z) \neq H_2(z)$. Then $0 = [H_2(\omega) - H_1(\omega)]\Phi_{x_I x_I}(\omega)[H_2(\omega) - H_1(\omega)]^*$ implying that $H_1(z) = H_2(z)$. ∎

Lemma 20: Consider a well-posed LDG $\mathcal{G} = (G, \mathcal{H}, e_I)$. Assume $V = \{N_i\}_{i=1,...,n}$ and let $X := \{x_i\}_{i=1,...,n}$ be the set of processes x_i corresponding to the nodes N_i. Fix $j \in \{1, ..., n\}$ and define the set

$$C := \{c | N_c \in \mathcal{C}_G(N_j)\} = \{c_1, ...c_{n_c}\}$$

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

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

Proof: Define

$$\begin{aligned}
&\varepsilon_j := 0 \\
&\varepsilon_k := e_i + H_{ij}(z)e_j && \text{if } k \in C \\
&\varepsilon_k := e_i && \text{if } k \notin \{C\} \cup \{j\} \\
&\xi_k := \sum H_{ki}(z)x_i && \text{if } k = j \\
&\xi_k := x_k && \text{if } k \neq j
\end{aligned} \tag{8}$$

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

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

where the first equality follows from (8) and last equality follows from the fact that $\varepsilon_j = 0$. This proves the assertion. ∎

Now we provide the main result of this work

Theorem 21: Consider a well-posed and topologically detectable LDG. Let $x_1, ..., x_n \in {}^0\mathcal{FE}$ be the signals associated with the n nodes of its graph. Define $X_j = \text{span}\{x_i\}_{i\neq j}$. Consider the problem of approximating the signal x_j with an element $\hat{x}_j \in X_j$

$$\min_{\hat{x}_j \in X_j} \|x_j - \hat{x}_j\|^2. \tag{9}$$

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

$$\hat{x}_j \in \text{span}\{x_{k_1}, ..., x_{k_{n_k}}\} \tag{10}$$

where $x_{k_1}, ..., x_{k_{n_k}}$ are the kin signals of x_j.

Proof: Fix j, and define the following set of indexes

$$\begin{aligned}
C &:= \{c | N_c \in \mathcal{C}_G(N_j)\} = \{c_1, ...c_{n_c}\} \\
P_l &:= \{p | N_p \in \mathcal{P}_G(N_l)\}/\{j\} = \{p_{l1}, ...p_{ln_{pl}}\} \\
K &:= \{k | N_k \in \mathcal{K}_G(N_j)\} = \{k_1, ...k_{n_k}\}
\end{aligned}$$

for any $l = 1, ..., n$. The set C contains all the indexes of the n_c children of N_j while the set K contains all the n_k indexes of the kins of N_j. The set P_l contains all the n_{p_l} parents (but N_j) of a generic node N_l.

For any $c \in C$, define

$$\varepsilon_c := x_c - \sum_{p\in P_j} H_{cj}H_{jp}(z)x_p - \sum_{p\in P_c} H_{cp}(z)x_p. \tag{11}$$

Thus ε_c can be expressed in terms of filtered versions of kins of N_j. Note that $\varepsilon_c = e_c + H_{cj}(z)e_j$, and it is given by a linear filtering of signals x_k, $k \in K$. In other words, there exist a matrix $M_1(z)$, such that $\varepsilon_C = M_1(z)x_K$. Note that M_1 is real-rational and has no poles on the unit circle. Also note that

$$e_j := x_j - \sum_{p\in P_j} H_{jp}(z)x_p,$$

so there exists a real-rational matrix $M_2(z)$ such that $e_j = x_j - M_2(z)x_K$. Observe that any entry of $M_2(z)$ is real rational and has no poles on the unit circle, as well.

Observe that $\Phi_{\varepsilon_C\varepsilon_C}(\omega) > 0$ for any $\omega \in [-\pi, \pi]$ since the LDG is topologically detectable. Let $M_3(z)$ be the real-rational tranfer matrix that solves the minimization problem

$$\min_{M_3(z)} \|e_j - M_3(z)\varepsilon_c\|^2.$$

According to Proposition 17, $M_3(z)$ exists; for Lemma 19, it is unique; and it has no poles on the unit circle. Moreover, defining $\hat{e}_j = M_3(z)\varepsilon_c$, we have that

$$< e_j - \hat{e}_j, \varepsilon >= 0 \tag{12}$$

for any $\varepsilon \in \text{span}\{\varepsilon_{c_1}, ..., \varepsilon_{c_{n_c}}\}$. Define the following estimate $\hat{x}_j$ of x_j

$$\hat{x}_j = [M_3(z)M_1(z) + M_2(z)]x_K.$$

Now, we prove that $< x_j - \hat{x}_j, \overline{x} >= 0$ for any $\overline{x} \in X_j$. First note that $x_j - \hat{x}_j = e_j - \hat{e}_j$. Define $\varepsilon_j = 0$ and, for $i \notin C \cup \{j\}$, define $\varepsilon_i = e_i$. From Lemma 20, we have that $x_j = \text{span}\{\varepsilon_i\}$. From Lemma 18 it suffices to prove that, for $i = 1, ..., n$

$$\Phi_{(e_j - \hat{e}_j)\varepsilon_i}(\omega) = 0. \tag{13}$$

The case $i = j$ is trivial because $\varepsilon_j = 0$. Now, consider $i \notin \{j\} \cup C$. Condition (13) follows by the fact that $e_j - \hat{e}_j \in \text{span}\{e_j, e_{c_1}, ..., e_{c_n}\}$ and $\varepsilon_i = e_i$.

For $i \in C$ Condition (13) is a consequence of Equation (12). Then, for the Hilbert projection theorem, $\hat{x}_j$ is the unique solution to the optimization problem. Morever, by its definition, it can be obtained by filtering only the "kin" signals $x_{k_1}, ..., x_{k_{n_k}}$. ∎

An immediate important consequence of Theorem 21 is that, for a well-posed and topologically detectable LDG, the solution to (9) can be obtained making use of the kin signals only. Since the minimizing $\hat{x}$ can be obtained also from the Wiener Filter, it means if the representation of $\hat{x}$ is unique in the span subspace, the Wiener Filter must have a null entry corresponding to any signal x_i which is not in its kin topology. This is stated in the following theorem.

Theorem 22: Consider a well-posed and topologically detectable LDG $\mathcal{G} = (G, \mathcal{H}, e_I)$ with $G = (V, A)$. Assume $V = \{N_i\}_{i=1,...,n}$ and let $X := \{x_i\}_{i=1,...,n}$ be the set of processes x_i corresponding to the nodes N_i. Compute the topological filter $W(z)$ of $\mathcal{G}$. Then

$$W_{ji}(z) \neq 0 \text{ implies } (N_i, N_j) \in kin(G)$$

Proof: We do not lose any generality if we prove the theorem only for $j = n$. Define $I := (1, ..., n)^T$, $I_n := (1, ..., n-1, 0)^T$ and $J = (1, ..., n-1)^T$. Since $\mathcal{G}$ is topologically detectable and well-posed, $\Phi_{x_I}(\omega) > 0$ for any $\omega \in [-\pi, \pi]$. Indeed, we have

$$\Phi_{x_I}(\omega) = [I - H(\omega)]^{-1} \Phi_{e_I}(\omega) [I - H(\omega)]^{-*},$$

where the right hand term is positive since $\Phi_{e_I}(\omega) > 0$ and $[I - H(\omega)]^{-1}$ has full rank for any $\omega \in [-\pi, \pi]$. Thus also $\Phi_{x_J}(\omega)$ is positive since it is a minor of $\Phi_{x_I}(\omega)$. The fact that $\Phi_{x_0 x_J}(\omega) = 0$ makes the matrix $\Phi_{x_{I_n}}(\omega)$ block diagonal and positive for any $\omega \in [-\pi, \pi]$, since x_0 is white. Note that the n-th row of $W(z)$ is defined as the non-causal Wiener filter estimating x_j from the processes $\{x_i\}_{i \neq j}$ and x_0. Thus $\hat{x} = \Phi_{x_n x_{I_n}}(z) \Phi^{-1}_{x_{I_n} x_{I_n}}(z)$ solves the problem

$$\min_{q \in \text{span}\{x_{I_n}\}} \|x_n - q\|^2$$

that can be rewritten as

$$\min_{q \in \text{span}\{x_J\}} \|x_n - q\|^2 + \min_{W_{nn}(z)} \|W_{nn}(z) x_0\|^2$$

because of the fact that $\Phi_{x_0 x_J}(\omega) = 0$ for any $\omega \in [-\pi, \pi]$. Since x_0 is white, the minimization requires $W_{nn}(z) = 0$. Then, from Theorem 22 $\hat{x}$ can be expressed in terms of the signals corresponding to the kins of the node N_n. From Lemma 19, $\hat{x}$ can be represented in a unique way as a linear transformation of $\{x_i\}_{i=1,...,n}$, thus $\Phi_{x_n x_{I_n}}(z) \Phi^{-1}_{x_{I_n} x_{I_n}}(z)$ has a null entry for any element that is not a kin of N_n. ■

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

Corollary 23: Consider a LDG $\mathcal{G} = (G, \mathcal{H}, e_I)$ with $G = (V, A)$ Assume $V = \{N_i\}_{i=1,...,n}$ and let $X := \{x_i\}_{i=1,...,n}$ be the set of processes x_i corresponding to the nodes N_i. Let $W(z)$ be the topological filter of $\mathcal{G}$. If $\mathcal{G}$ is self-kin, then $W_{ji}(z) \neq 0$ implies $(N_j, N_i) \in top(G)$.

Proof: Since $\mathcal{G}$ is self-kin, $\mathcal{P}_G(N_j) \cup \mathcal{C}_G(N_j) \cup \mathcal{P}_G(\mathcal{C}_G(N_j)) = \mathcal{C}_G(N_j) \cup \mathcal{P}_G((N_j))$. Thus, from the previous theorem we immediately have the assertion. ■

VI. A Reconstruction Algorithm

The previous section provides us with theoretical results for the detection of links in a self-kin LDG. We stress that even in the case we are dealing with 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 of practical importance even only estimating a reduced number of possible interconnections among those nodes. For example, DNA-microarrays are devices allowing to measure the complete gene expression of a cell. Those data can be useful to understand which genes interact together realizing a specific metabolic pathway and how they are related. Indeed, a cell can express tens of thousands of genes while we are expecting only few tens to be 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 absolute importance [23]. A method that succeeds in identifying the kin nodes of a target node in a sparse network is definitely helpful for these kind of applications. Analogously, in Economics quantifying the strongest interconnections among a set of market stocks can suggest good strategies to balance a given portfolio [3], 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 faced in neuroscience in order to understand neural interconnections [5]. A second reason why our analysis is of practical importance is that as a byproduct of the reconstruction we obtain an optimal model for the node dynamics which can be used for smoothing procedures (we are deriving non-causal Wiener filters). The developed theory shows an important property of the Wiener filters, namely that, for every node signal in an LDG, only the kin nodes are actually required for the computation of its output. This sparsity property, that can be easily deduced for causal Wiener filters as well, provides the theoretical backbone for local and distributed implementations of smoothing or predicting procedures. Of course, the theoretical conditions of a LDG are not expected to be met in practical scenarios (as any other conditions for identification or modeling methods), but the developed theory is helpful to justify the following algorithmic approach for the reconstruction of the kin topology of a LDG from data.

A reconstruction algorithm

The following algorithm is the natural pseudocode implementation of the reconstruction technique we have theoretically developed in the previous section.

```
Reconstruction Algorithm:
0. Initialize the set of edges A = {}
1. For any signal x_j
2.    solve the minimization problem (9)
3.    For any W_ji(z) ≠ 0
4.       add {N_i, N_j} to A
5.    end
7. end
8. return A
```

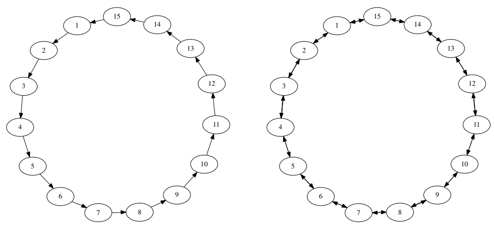
Step 2 can be computed efficiently for a large number of signals using techniques based on Gram-Schmidt orthogonalization as those described in [24]. When checking the condition at step 3, we have to take into account that we are dealing with real data. Consequently, the condition $W_{ji} \ncong 0$ can be implemented as $\|W_{ji}(z)\| > \sigma_{thr}(x_i, x_j)$ where $\sigma_{thr}(x_i, x_j)$ is a threshold depending on both the signals x_i and x_j and a norm must be defined on the space of transfer functions.

VII. Numerical Examples

In this section we introduce a suitable framework to illustrate the application of the previous theoretical results to numerical analysis. It is worth observing that the previous results have been developed for the general class of linear models. Indeed, no assumptions have been done on the order and causality property of the considered transfer functions. Moreover, let us highlight that our analysis has been realized "off-line", since the processes have been evaluated over their entire time span, even though "on-line" implementations are certainly possible.

A. Self-kin network

We have considered a ring network of 15 nodes where the dynamics of the links was given by 5-th order FIR filters and the noise power was the same on every node. The network topology is provided in Figure 2 The network



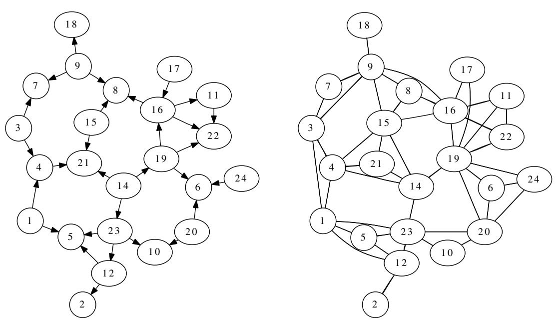
(a) (b)

Fig. 2. A ring network of 15 nodes as the one considered in Example VII-A (a). The reconstructed topology of Example VII-A. Every single links has been detected and, since the network is self-kin, the topology does not contain any spurious link (b).

has been simulated for 1000 steps and an implementation of the algorithm has been applied to the data providing the topology of Figure 2.

B. Generic network

We have also considered a generic (sparse) network and we have verified the reliability of our procedure. We simulated a network of 24 nodes for 1000 steps as reported in Figure 3 (a) and the reconstructed topology we obtained is depicted in Figure 3 (b).



(a) (b)

Fig. 3. A network of 24 nodes as the one considered in Example VII-B (a) and the reconstructed topology (b). Every single links has been detected, but, since the network is not self-kin, the topology contains the additional links between the "kins".

VIII. Conclusions

This work has illustrated a simple but effective procedure to identify links in the general structure of a network of linear dynamical systems. The followed approach is based on Wiener Filtering in order to detect the existing links of a network. When the topology of the original graph is described by a self-kin network, our method provides sufficient conditions for a correct detection of a link. Self-kin networks provide a non-trivial class of networks since they allow the presence of loops, nodes with multiple inputs and lack of connectivity. However, the paper also provides results about general networks. It is shown that, for a general graph, the developed procedure detects links belonging to the topology of the self-kin graph with the least number of nodes that contains the original graph. Thus, the method is optimal in this specific sense. Numerical examples illustrate the correctness and also the reliability of the identification technique.

References

[1] D. Geer, "Industry trends: Chip makers turn to multicore processors," *Computer*, vol. 38, no. 5, pp. 11–13, 2005.

[2] S. C. Kadu, M. Bhushan, and R. Gudi, "Optimal sensor network design for multirate systems," *Journal of Process Control*, vol. 18, no. 6, pp. 594 – 609, 2008.

[3] R. Mantegna and H. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge UK: Cambridge University Press, 2000.

[4] 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. 14 863–8, 1998.

[5] 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 Natl Acad Sci USA*, vol. 101, no. 26, pp. 9849–9854, June 2004.

[6] D. Urban and T. Keitt, "Landscape connectivity: A graph-theoretic perspective," *Ecology*, vol. 82, no. 5, pp. 1205–1218, 2001.

[7] J.-S. Bailly, P. Monestiez, and P. Lagacherie, "Modelling spatial variability along drainage networks with geostatistics," *Mathematical Geology*, vol. 38, no. 5, pp. 515–539, 2006.

[8] G. Innocenti and D. Materassi, "A modeling approach to multivariate analysis and clusterization theory," *Journal of Physics A*, vol. 41, no. 20, p. 205101, 2008.

[9] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175–308, February 2006.

[10] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, no. 2, 2004.

[11] H. Zhang, Z. Liu, M. Tang, and P. Hui, "An adaptative routing strategy for packet delivery in complex networks," *Physics Letters A*, vol. 364, pp. 177–182, 2007.

[12] R. Olfati-Saber, "Distributed kalman filtering for sensor networks," in *Proc. of IEEE CDC*, New Orleans, 2007, pp. 5492–5498.

[13] E. Kolaczyk, *Statistical Analysis of Network Data: Methods and Models*. Berlin, Germany: Springer-Verlag, 2009.

[14] M. Timme, "Revealing network connectivity from response dynamics," *Phys. Rev. Lett.*, vol. 98, no. 22, p. 224101, 2007.

[15] 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.

[16] D. Napoletani and T. Sauer, "Reconstructing the topology of sparsely connected dynamical networks," *Phys. Rev. E*, vol. 77, p. 026103, 2008.

[17] C. Michener and R. Sokal, "A quantitative approach to a problem of classification," *Evolution*, vol. 11, pp. 490–499, 1957.

[18] D. Materassi and G. Innocenti, "Unveiling the connectivity structure of financial networks via high-frequency analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 18, pp. 3866–3878, June 2009.

[19] N. Friedman and D. Koller, "Being bayesian about network structure: A bayesian approach tostructure discovery in bayesian networks," *Machine Learning*, vol. 50, pp. 95–126, 2003.

[20] E. Candès, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted l1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, 2008.

[21] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[22] R. Diestel, *Graph Theory*. Berlin, Germany: Springer-Verlag, 2006.

[23] D. Marinazzo, M. Pellicoro, and S. Stramaglia, "Kernel Granger causality and the analysis of dynamical networks," *Phys. Rev. E*, vol. 77, p. 056215, 2008.

[24] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Transactions on Information Theory*, vol. 44, pp. 2943–2959, 1998.