2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Reconstruction of topologies for acyclic networks of dynamical systems

Donatello Materassi

Abstract— **The paper deals with the problem of unveiling the link structure of a network of linear dynamical systems. A technique is provided guaranteeing the exact detection of the links for networks with no undirected cycles (Linear Dynamic Polytrees). The result extends previous work that was limited to a more restricted class (Linear Cascade Model Trees).**

I. INTRODUCTION

Recently, a significant interest for complex systems has been shown in many scientific fields, with a particular focus on the study of the emergence of complicated phenomena from the connection of simple models [1]. As a consequence, graph theory [2] has been successfully exploited to perform novel modeling approaches in several fields, such as Economics (see e.g. [3], [4]), Biology (see e.g. [5]) and Ecology (see e.g. [6]), especially when the investigated phenomena are characterized by spatial distribution and a multivariate analysis technique is preferred [7].

While networks of dynamical systems are well studied and analyzed in Physics [8], [9] and Engineering [10], [11], [12], there are fewer results that address the problem of reconstructing an unknown dynamical network from sampled data, since it poses formidable theoretical and practical challenges [13].

One of the most formidable complications in the reconstruction of a network of dynamical systems is given by the presence of cycles in the structure. This is the reason why most techniques focus on identifying acyclic structures (see for example [14], [5], [3], [15]). However, even though an acyclic topology may seem a quite reductive choice, given an intricate and connected link structure, one may be interested into "approximating" it with a tree scheme. Such an approximation could be considered "satisfactory" if the most important connections were captured.

For example, tree topologies have been successfully employed in [5] for the study of gene regulatory networks that have instead a more complicated structure.

Another well-known technique for the identification of a tree network in a complex scenario is developed in [3] for the analysis of a stock portfolio. The authors identify a tree structure according to the following procedure: i) a metric based on the correlation index is defined among the nodes; ii) such a metric is employed to extract the Minimum Spanning Tree [2] which forms the reconstructed topology. In [7] severe limitations of these strategies are 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 a wrong topology. In [16] 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. Finally, in [15] it is shown that the procedure theoretically guarantees a correct reconstruction for rooted tree topologies. Rooted tree topologies are good models for networks where propagation phenomena are present. However, they can still be considered a limited class of networks for other applications. Indeed assuming the presence of a single root node determines the orientation of all links in a unique way. Thus a single root process is assumed to drive the dynamics of all the other nodes. This makes all the network processes necessarily correlated.

In this paper we extend the theoretical guarantees that were provided in [15] for rooted tree networks to the more general case of connected networks with no undirected cycles.

Notation:
The symbol $:=$ denotes a definition
(v_1, v_2): ordered pair of two elements v_1 and v_2
W^*: the conjugate transpose of a matrix or vector W
$E[\cdot]$: mean operator;
$R_{xy}(\tau) := E[x(t)y^T(t+\tau)]$: cross-covariance function of wide-sense stationary processes x and y;
$R_x(\tau) := R_{xx}(\tau)$: 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;

II. PRELIMINARY RESULTS

In this section we give the necessary theoretical background in order to formulate the problem of reconstructing an acyclic network of dynamical systems.

A. Recall of Graph theory concepts

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 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.
Given a directed graph $G = (V, A)$ it is possible to define its undirected version $G' = (V', A')$ where $V' = V$ and $A' = \{\{v_1, v_2\} \subseteq V | (v_1, v_2) \in A\}$.

Definition 2: Given a graph $G = (V, A)$, the pair $C = (V_1, V_2)$ is a a *cut* of G if $V = V_1 \cup V_2$. Every edge $(N_1, N_2) \in A$ (or $\{N_1, N_2\} \in A$ for undirected graphs)

978-1-4577-0079-8/11/$26.00 ©2011 AACC

such that $N_1 \in V_1$ and $N_2 \in V_2$ is referred to as a an *edge of the cut* C and we write $(N_1, N_2) \in Cut_A(C)$ (or $\{N_1, N_2\} \in Cut_A(C)$ for undirected graphs).

The definition of "walk" in a graph will be widely used in the rest of the paper. In the literature of graph theory a "walk" is defined in a variety of way non always equivalent. Thus, we explicitely provide the definition that we will use.

Definition 3: Consider a directed graph $G = (V, A)$. We say that a finite ordered sequence $P = (N_{\pi_0}, \dots, N_{\pi_l})$ of elements of V is a *directed walk* from the node N_s to the node N_e if the following properties are met

1) $N_{\pi_0} = N_s$, $N_{\pi_l} = N_e$
2) $\left(N_{\pi_{k-1}}, N_{\pi_k}\right) \in A$ for $k = 1, ..., l$

We say that P is an *undirected walk* from the node N_s to the node N_e if this weaker property replaces 2)

2b) $\left(N_{\pi_{k-1}}, N_{\pi_k}\right) \in A$ or $\left(N_{\pi_k}, N_{\pi_{k-1}}\right) \in A$ for $k = 1, ..., l$.

If $G = (V, A)$ is an undirected graph we equally say that P is an *undirected walk* by replacing 2b) with

2c) $\left\{N_{\pi_{k-1}}, N_{\pi_k}\right\} \in A$ for $k = 1, ..., l$.

In all the cases above, we say that $l \geq 0$ is the *length* of the walk and that the walk is *simple* if all nodes are distinct.

Lemma 4: For any walk $P := (N_{\pi_0}, \dots, N_{\pi_l})$ that is not simple and with $N_{\pi_0} \neq N_{\pi_l}$, it is possible to find a walk P' from N_{π_0} to N_{π_l} with length $0 < l' < l$.

Proof: Since P is not simple assume that, for $1 \leq i < j \leq l$, we have $N_{\pi_i} = N_{\pi_j}$. Define $P' := (N_{\pi_0}, ..., N_{\pi_i}, N_{\pi_{j+1}}, ..., N_{\pi_l})$. It is straightforward to verify that P' is still a walk and that its length satisfies $0 < l' = l - j + i < l$. ∎

Lemma 5: For any walk $P := (N_{\pi_0}, \dots, N_{\pi_l})$ with $N_{\pi_0} \neq N_{\pi_l}$, there exists a simple walk from N_{π_0} to N_{π_l}

Proof: By contradiction, there is no simple walk from N_{π_0} to N_{π_l}. Since P is not simple there is a walk P_1 from N_{π_0} to N_{π_l} with length $l_1 \leq l - 1$. P_1 can not be simple, thus there must exist P_2 with length $l_2 \leq l - 2$. Iterating the argument we find that there must be a walk with non-positive length, that is a contradiction. ∎

We also introduce the concept of "connected graph".

Definition 6: A directed graph $G = (V, A)$ is directly (undirectly) *connected* if for any pair of distinct nodes $N_i,\ N_j \in V$ there is at least one directed (undirected) walk connecting them. Analogously, if $G = (V, A)$ is an undirected graph, we say it is connected if there is at least an undirected walk between each couple of distinct nodes $N_i,\ N_j \in V$.

Definition 7 (Directed and Undirected cycles): Given a graph $G = (V, A)$, a directed (undirected) cycle is a directed (undirected) walk from one node to itself.

Definition 8 (Ancestors, Descendants): Given a directed graph $G = (V, A)$, if there is a directed walk from N_i to N_j, we say that N_i is an ancestor of N_j or, equivalently, that N_j is a descendant of N_i.

Definition 9 (Common ancestor): Given a directed graph $G = (V, A)$, if N_k is an ancestor of both N_i and N_j, we say that N_k is a *common ancestor* of N_i and N_j.

Definition 10 (Related nodes): Given a directed graph $G = (V, A)$, two nodes N_i and N_j are related if one is a descendant of the other or if they have a common ancestor.

Definition 11 (Polytree): A polytree is a directed graph $G = (V, A)$ meeting the following two conditions

- for any two distinct nodes N_i and N_j, there is exactly one simple undirected walk linking them
- $\{N_i, N_j\} \in A$ implies $\{N_j, N_i\} \notin A$.

Any node for which there are no entering edges is called a *root* of the polytree.

Definition 12 (Rooted tree, undirected tree): A rooted tree is a polytree with exactly one root.
A tree is the undirected version of a polytree.
The following definition introduces the concept of Minimum Spanning Tree.

Definition 13: Given a connected undirected graph $G = (V, A)$, a *Spanning Tree* of G is a subgraph $T = (V, A_T)$ that is a tree. Given a weight function defined on the edges, $w : A \to \mathbb{R}$, T is also a *Minimum Spanning Tree* (MST) of G with respect to w if, for every Spanning Tree $T' = (V, A'_T)$, it satisfies

$$\sum_{a \in A_T} w(a) \leq \sum_{a' \in A'_T} w(a'). \tag{1}$$

Proposition 14 (Cut Property): Consider a connected and undirected graph $G = (V, A)$, a weight $w : A \to \mathbb{R}$ and a cut $C = (V_1, V_2)$. If, there exists an edge $\{N_1, N_2\} \in Cut_A(C)$ such that, for all $\{N_3, N_4\} \in Cut_A(C)$,

$$w(\{N_1, N_2\}) \leq w(\{N_3, N_4\}), \tag{2}$$

then there is a Minimum Spanning Tree of G with respect to w that contains the edge $\{N_1, N_2\}$.
Furthermore, if there is one edge $\{N_1, N_2\}$ with weight strictly smaller than every other edge of the cut, it belongs to any MST of G with respect to w.

Proof: It is a standard result in graph theory. See, for example, [2]. ∎

B. Rationally correlated processes

Definition 15: The set $\mathcal{F}$ is defined as the set of real-rational SISO transfer function with no poles on the unit circle $\{z \in \mathbf{C} |\ \ |z| = 1\}$.

Definition 16: Consider a transfer function $H(z) \in \mathcal{F}$ and let e be a wide sense stationary random process. Since $H(z)$ is analytical in a neighborhood of the unit circle, by the properties of the $\mathcal{Z}$-transform, there is a unique bi-infinite sequence h_k such that, for any $|z| = 1$,

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

By using the notation $y(t) = H(z)e(t)$ we denote the random process defined, for any t, as

$$y(t) = \sum_{k=-\infty}^{\infty} h_{t-k} e(k).$$

Definition 17: 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

$$\Phi_{e_i e_j}(z) = \frac{A(z)}{B(z)}, \tag{3}$$

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 correlated random processes.

C. Linear Dynamic Graphs, Polytrees and Rooted Trees

The following definition provides a class of models for a network of dynamical systems. We assume 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 correlated with the other noise components. If a certain agent "influences" another one, then a directed edge can be drawn and a directed graph can be obtained.

Definition 18 (Linear Dynamic Graph): A Linear Dynamic Graph is defined as a pair $(H(z), e)$ where

- $e = (e_1, .., e_n)^T$ is a vector of n rationally correlated random processes such that $\Phi_e(z)$ is diagonal
- $H(z)$ is a $n \times n$ matrix of transfer functions in $\mathcal{F}$.

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

$$x_j = e_j + \sum_{i=1}^{n} H_{ji}(z) x_i, \tag{4}$$

or in a more compact way $x(t) = e(t) + H(z)x(t)$.
A LDG is said topologically detectable if $\Phi_{e_j}(z) > 0$ for any $|z| = 1$.
Let $V := \{x_1, ..., x_n\}$ 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. Abusing the nomenclature we will refer to nodes, edges, cycles, walks, roots etc... of a LDG even though, formally, we should refer to them as nodes, edges, cycles, walks, roots etc... of its associated graph.

Definition 19 (Linear Dynamic Trees): The LDG $(H(z), e)$ is a Linear Dynamic Polytree (LDP) if the associated graph is a polytree.
The LDG $(H(z), e)$ is a Linear Cascade Model Tree (LCMT) if the associated graph is a rooted tree (see [15]).

As in [15] we define a pseudo-metric among the wide-sense stationary processes $x_1, ..., x_n$ in the following way

Definition 20: Given a LDG with nodes $\{x_1, ..., x_n\}$ we define the coherence pseudo-metric as

$$d(x_i, x_j) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(1 - \frac{|\Phi_{x_i x_j}(e^{i\omega})|^2}{\Phi_{x_i}(e^{i\omega}) \Phi_{x_j}(e^{i\omega})} \right) d\omega. \tag{5}$$

Finally, we recall the main result proved in [15].

Theorem 21: Consider a topologically detectable LCMT $(H(z), e)$ with associated directed graph $T = (V, A)$. Define the complete graph $Q = (V, E)$ where $E := \{\{x_i, x_j\} \in V^2 | x_i \neq x_j\}$ and consider the weight $w(\{x_i, x_j\}) := d(x_i, x_j)$ where $d(x_i, x_j)$ is the coherence pseudo-metric. The MST of Q with respect to w is unique and is the undirected version of T.

Proof: See [15]. ■

III. Problem Formulation

Given a LDP with output processes $\{x_j\}_{j=1}^n$, assume that only the (cross)-spectral densities $\Phi_{x_i x_j}(e^{i\omega})$ are known and determine, for any two processes x_i and x_j, if there is an edge linking them (disregarding the orientation).

IV. Preliminary Results

We start with the following lemma.

Lemma 22: Consider a LDG $(H(z), e)$ with nodes $\{x_1, ..., x_n\}$ and assume that $l < +\infty$ is the length of the longest directed walk (that is there are no directed cycles). Then, we have that

$$x = T(z)e = \left(I + \sum_{k=1}^{l} H^k(z) \right) e. \tag{6}$$

Furthermore, if there is no directed walk from x_i to $x_j \neq x_i$, $T_{ji}(z) = 0$.

Proof: Observe that a non-zero entry (j, i) of $H(z)$ represents the presence of a direct link from x_i to x_j, that is a walk of length 1. The entry (j, i) of $H^2(z)$ is given by

$$(H^2(z))_{ji} = \sum_{k=1}^{n} H_{jk}(z) H_{ki}(z) \tag{7}$$

that is zero if there is no direct walk of exactly length 2 from x_i to x_j. Iterating the reasonment we find that, if there is no direct walk of length q from x_i to x_j, then $(H^q(z))_{ji} = 0$. Then we have that $(H^q(z))_{ji}$, for any $q > l$. Now consider the relation $(I - H(z))x = e$. Since $H(z)$ is nilpotent of order $l + 1$, $(I - H(z))$ is invertible and $(I - H(z))^{-1} = I + \sum_{k=1}^{l} H^k(z)$. The statement follows immediately. ■

Lemma 23: Given a LDP (V, A) with nodes $V = \{x_1, ..., x_n\}$, each node that is not a root has a root ancestor

Proof: Without any loss of generality assume that x_1 is not a root. Then it has a parent. Let it be x_2. If x_2 is a root, the lemma is proven. If x_2 has a parent, then it must be different from x_1, otherwise we would have that both (x_1, x_2) and (x_2, x_1) belong to A, against the fact the graph is a polytree. Thus let x_3 be the parent of x_2. It holds that $x_3 \neq x_1, x_2$, and that (x_3, x_2, x_1) is a directed walk. Let us prove that if there is a direct walk $(x_k, x_{k-1}, ..., x_1$ and y is a parent of x_k, then $y \neq x_1, ..., x_k$.
If there exists $1 \leq j \leq k$ such that $y = x_j$, then the two walks (x_j, x_k) and $(x_j, x_{j+1}, ..., x_k)$ are two simple undirected walks from x_j to x_k. This is a contradiction since the graph is a polytree proving that $y \neq x_1, ..., x_k$. If x_3 has a parent then, it needs to be different from x_1, x_2. Let it be x_4. By iterating the argument we must eventually find a root or we will obtain the contradiction that the graph has more than n distinct nodes. ■

Lemma 24: Given a LDP with nodes $\{x_1, ..., x_n\}$, if x_i and x_j are related nodes, then they have a common ancestor that is a root or x_i is a root or x_j is a root.

Proof: If either x_i or x_j is a root, the assertion is true. So consider the case where neither of them is a root. Assume that x_i is an ancestor of x_j. By Lemma 23, x_i has a root ancestor that is also an ancestor of x_j.
Analogously, if x_j is an ancestor of x_i, by Lemma 23, x_j has a root ancestor that is also an ancestor of x_i.
If x_i and x_j have a common ancestor x_k, then it either can be a root or.have a root ancestor, proving the assertion even in this case. ∎

Lemma 25: Given a topologically detectable LDP with nodes $\{x_1, ..., x_n\}$, if x_i and x_j are not related nodes, then $\Phi_{x_i,x_j}(z) = 0$.

Proof: Let $T(z)$ be as in (6). Let x_k be a node of the LDP. First assume that $x_k \neq x_i, x_j$. Since x_i and x_j are not related nodes, it is not possible that there is at the same time a direct walk from x_k to x_i and a direct walk from x_k to x_j. Then, $T_{ik}(z)T^*_{jk}(z) = 0$. If $k = i$ (or $k = j$), then $T_{ii}(z)T^*_{ji}(z) = 0$ (or $T_{ij}(z)T^*_{jj}(z) = 0$) since one is not a descendant of the other and viceversa. This implies that

$$\Phi_{x_i x_j}(z) = T_{i*}(z)\Phi_e(z)T^*_{*j}(z) = 0 \tag{8}$$

where the last equality follows from the fact the $\Phi_e(z)$ is diagonal. ∎

Given a LDP $P = (H(z), e)$ with n nodes, it is possible to define a LCMT for each of its roots.

Proposition 26 (LCMT associated to a root of a LDP): Let $P = (H(z), e)$ be a LDP with n nodes. Without any loss of generality, let x_1 be a root of P and let $x_2, ..., x_m$ be all its descendants. Let $x_{m+1}, ..., x_n$ be all the other nodes. For any $i = 1, ..., m$ define

$$e_i^{(tree)} := e_i + \sum_{k=m+1}^{+\infty} H_{ik}x_k$$

$$H_{ji}^{(tree)}(z) := H_{ji}(z) \qquad \text{for } i, j = 1, ..., m.$$

Then, $(e^{(tree)}, H^{(tree)}(z))$ is a LCMT with nodes $\{x_1, ..., x_m\}$.

Proof: First we prove that $\{e_i^{(tree)}\}_{i=1}^m$ have null cross-spectral density making $(e^{(tree)}, H^{(tree)}(z))$ a LDG.

Consider $k_i > m$ and $k_j > m$ such that $H_{ik_i}(z) \neq 0$ and $H_{jk_j}(z) \neq 0$, for $i, j < m$ and $i \neq j$.

First, we show that x_{k_i} and x_{k_j} are not related nodes. By contradiction assume they are related. Let $P^{(i)} = (y_1^{(i)}, ..., y_{l_i}^{(i)})$ and $P^{(j)} = (y_1^{(j)}, ..., y_{l_j}^{(j)})$ the walks from x_1 to x_i and from x_1 to x_j respectively. They exists because x_i and x_j are descendants of x_1. $P' = (y_1^{(i)}, ..., y_{l_i}^{(i)}, y_{l_j}^{(j)}, ..., y_1^{(j)})$ is a walk from x_i to x_j. Let P be the simple version of P'. It is the unique simple walk connecting x_i and x_j. Observe that it contains nodes x_k with $k \leq m$ and none of them can be an ancestor for either x_{k_i} or x_{k_j} (otherwise they would be descendants of x_1). Build the walk from x_{k_i} to x_{k_j} by appending the edges x_{k_i} and x_{k_j} at the beginning and at the end of P respectively. This walk is simple and is not directed, thus x_{k_i} and x_{k_j} are not in a ancestor/descendant relation.
Thus, there must be a common ancestor for x_{k_i} and x_{k_j}. Use it to build a simple walk Q between them and observe that it can only contain nodes x_k with $k > m$. This is a contradiction since there must be a unique simple walk between two nodes. Since x_{k_i} and x_{k_j} are not related, from Lemma (25), their cross-spectral density is zero. Then, for $i \neq j$, we have that

$$\Phi_{e_i^{(tree)} e_j^{(tree)}}(z) = 0 \tag{9}$$

proving that $(H^{(tree)}(z), e^{(tree)})$ is a LDG. Define

$$x^{(tree)} = e^{(tree)} + H^{(tree)}(z)x^{(tree)} \tag{10}$$

and observe by inspection that $x_i^{(tree)} = x_i$ for $i \leq m$. The way the nodes have been chosen proves that the associate graph of $(H^{(tree)}(z), e^{(tree)})$ is a tree, thus the LDG is a LCMT. ∎

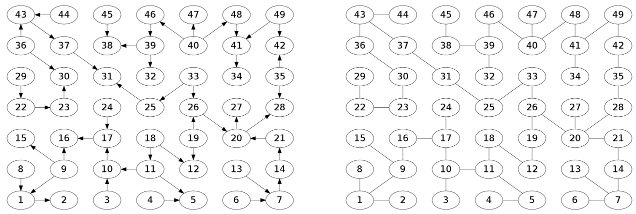
V. Main result

The following theorem is the main result of the paper. It states that, given a LDP, the MST associated with the coherence pseudo-metric of its nodes is unique and coincides with the undirected version of the LDP graph.

Theorem 27 (Topology reconstruction for LDP's): Consider a connected and topologically detectable LDP $(H(z), e)$ with associated graph $G = (V, A)$. Define the complete graph $Q := (V, E)$ where $E := \{\{x_i, x_j\} \in V^2 | x_i \neq x_j\}$, and the weight $w(\{x_i, x_j\}) = d(x_i, x_j)$. The MST of Q with respect to w is unique and coincides with the undirected version of G.

Proof: Let us consider an edge $(x_p, x_c) \in A$ and let us prove that it belongs to the MST. Consider the following two sets: V_1 contains x_p and all the nodes connected to it with an indirected walk that does not contain the edge $\{x_p, x_c\}$; and V_2 is its complement to V. Let us prove that, for any $x_1 \in V_1$ and any $x_2 \in V_2$ such that $x_1 \neq x_p$ or $x_2 \neq x_c$, we have $d(x_p, x_c) < d(x_1, x_2)$. First observe that $d(x_p, x_c) < 1$. Let us distinguish two cases.
If x_1 and x_2 are not related, from Lemma 25 their distance is 1, so the statement is true. If they are related, consider the LCMT T corresponding to one of their common roots. A common root exists because of Lemma 24. Let $x_{T_1}, ..., x_{T_m}$ be the nodes of T. Define $V_1^{(tree)} := \{x_{T_i} | x_{T_i} \in V_1\}$ and $V_2^{(tree)} := \{x_{T_i} | x_{T_i} \in V_2\}$. The sets $V_1^{(tree)}$ and $V_2^{(tree)}$ represent a cut of T. Observe that $x_1 \in V_1^{(tree)}$ and that $x_2 \in V_2^{(tree)}$. From Theorem 21, the edge $\{x_p, x_c\}$ belongs to the unique MST defined on the nodes $V_1^{(tree)} \cup V_2^{(tree)}$ and it is its only edge on the cut $(V_1^{(tree)}, V_2^{(tree)})$. Thus, the relation $d(x_p, x_c) < d(x_1, x_2)$ must be satisfied. ∎

Theorem 27 provides an immediate way to identify the network structure of a LDP from time series: under the assumption of ergodicity, estimate the coherence function and the coherence pseudo-metric to weight the edges, and then determine the MST to obtain the undirected network



(a) (b)

Fig. 1. The original network (a) and the reconstructed topology (b). Notice that while the edges have an orientation in the original network, the graph obtained using the proposed technique is not directed, thus the link orientation is not recovered. However, the detection of the presence/absence of the link between two nodes is guaranteed.

structure. Even though the technique does not reconstruct the link orientation it is guaranteed that all the connections are exactly detected.

VI. Numerical implementation and an example

The theory developed in the previous section relies on the knowledge of the coherence function associated with each pair of the network signals. An estimate of the coherence can be obtained from the observation of time realizations of the signals, under the assumption of ergodicity. To this aim, many techniques can be used. In the following example we have simply used a standard implementation based on the Welch algorithm (already provided in the software tool used for the numerical simulation).

A polytree of 49 nodes has been randomly generated (see Figure 1a). Each link of the network represents a third order transfer function that has been randomly generated as well. The noise signals acting on each node have been chosen to be white with unitary variance. The network has been simulated for 1000 steps and the coherence metric has been estimated form the time series. An implementation of the Prim algorithm [2] has been used to identify the MST and reconstruct the network topology (see Figure 1). Notice that while the technique is not capable of reconstructing the link orientation, every single link of the network structure has been exactly detected.

Identical results have been obtained by applying the same procedure to different randomly generated LDPs.

References

[1] 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.

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

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

[4] M. Naylora, L. Roseb, and B. Moyle, "Topology of foreign exchange markets using hierarchical structure methods," *Physica A*, vol. 382, pp. 199–208, 2007.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

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

[12] I. Schizas, A. Ribeiro, and G. Giannakis, "Consensus in ad hoc WSNs with noisy links-part i: Distributed estimation of deterministic signals," *IEEE Trans. on Signal Processing*, vol. 56, no. 1, 2008.

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

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

[15] D. Materassi and G. Innocenti, "Topological identification in networks of dynamical systems," *IEEE Trans. Aut. Control*, vol. 55, no. 8, pp. 1860–1871, August 2010.

[16] ——, "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.