A Control Theoretic Look at Granger Causality: Extending Topology Reconstruction to Networks With Direct Feedthroughs

Mihaela Dimovska and Donatello Materassi

***Abstract*—Reconstructing the causal structure of a network of dynamic systems from observational data is an important problem in many areas of science. One of the earliest and most prominent approaches to this problem is Granger causality. It has been shown that in a network with linear dynamics and strictly causal transfer functions, Granger causality consistently reconstructs the underlying graph of the network. On the other hand, techniques that allow for the presence of direct feedthroughs usually assume there are no feedback loops in the dynamics of the network. In this article, we develop an extension of Granger causality that provides theoretical guarantees for the reconstruction of a network topology even in the presence of direct feedthroughs and feedback loops. The only required assumption is a relatively mild condition of well-posedness named recursiveness, where at least one strictly causal transfer function needs to be present in every feedback loop. The technique is compared with other state-of-the-art methods on a benchmark example specifically designed to include several dynamic configurations that are challenging to reconstruct.**

***Index Terms*—Learning (artificial intelligence), stochastic processes, system identification, time series analysis.**

NOTATION

$y_i, i = 1, \ldots, n$: Wide-sense stationary processes.
$e_i, i = 1, \ldots, n$: External excitation processes.
A^T : Transpose of a matrix or vector A.
$\mathbf{C}$: Field of complex numbers.
$\mathcal{F}$: Set of potentially noncausal discrete-time finite-dimensional linear time-invariant systems.
$\mathcal{F}^+$: Set of causal discrete-time finite-dimensional linear time-invariant systems.
$H(z)$: Transfer matrix.
$\langle y_i, y_j \rangle$: Inner product between two elements y_i and y_j in a pre-Hilbert space.
$\mathcal{Z}(\cdot)$: Z-transform of a signal.
$\Phi_{y_i y_j}(z) := \mathcal{Z}[R_{y_i y_j}(\tau)](z)$: Cross-power spectral density.
$\frac{1}{z}$: Delay operator.
$W_{v,[y_i](y_1,\ldots,y_n)}(z)$: Component associated with y_i in the Wiener filter estimating v from $y_1, \ldots, y_n$.

Manuscript received September 7, 2019; revised January 27, 2020; accepted April 8, 2020. Date of publication April 20, 2020; date of current version January 28, 2021. This work was supported by NSF CAREER Award 1553504 and NSF SaTC 1816703. Recommended by Associate Editor G. Gu. *(Corresponding author: Mihaela Dimovska.)*

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: dimov003@umn.edu; mater013@umn.edu).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.
Digital Object Identifier 10.1109/TAC.2020.2989261

I. INTRODUCTION

NETWORKED and modular systems arise naturally in diverse fields such as biology [1], [2], neuroscience [3], economics [4], political science [5], [6], and social networks [7]. In many application scenarios the objective of the investigation is to infer how the individual components of the network affect and influence each other. While in some cases we have access to, or can directly manipulate the system under consideration, [8]–[13], there are more challenging scenarios, allowing only for noninvasive observations of the system. By noninvasive observations, we mean observations of the system outputs obtained under standard operating conditions and unknown external excitations. A wealth of different methodologies have been developed to recover the network graph from noninvasive observations [14]–[20]. These methods have different *a priori* assumptions about the system and achieve different reconstruction outcomes. On one end of the spectrum, some methods assume that the underlying graph structure is known and they pursue the identification of the dynamics that govern the direct couplings [18], [21]. On the other end of the spectrum, there are techniques whose very objective is to reconstruct the network structure, while the quantitative identification of the transfer functions might be a complementary goal [19], [22], [23], or not pursued at all [20]. The results developed in this article fall in the latter class of techniques, since an algorithm is provided with the primary goal of determining the underlying unknown topology (undirected structure) of a dynamic network.

To this end, most techniques make some assumptions regarding the dynamics between the variables in order to obtain theoretical guarantees. A typical case is given by Granger causality [24], where the consistency of the reconstruction is guaranteed in the linear case when all the transfer functions are strictly causal [20], [25], [26]. In [27], the authors provide an analogous result, but formulated in the continuous domain, where a by-product of the procedure is also the identification of the transfer functions in the network. Similarly, a recent result in [23], under the strict causality assumption, tackles the problem of reconstructing a network of heterogeneous data sets via linear regressions.

A fundamental challenge to these methods is the presence of transfer functions with direct feedthroughs.

Although there are methods posing no restrictions on the presence of direct feedthroughs, they typically require some assumptions on the underlying graph of the network. For example, the work in [28] provides an algorithm that recovers the exact topology for an acyclic network while the work in [29] specifically targets polytrees.

Reconstructing a network with direct feedthroughs, while making minimal assumptions on the network topology, is an active area of research [30]–[32]. Schiatti *et al.* [32] slightly modify Granger's test for causality attempting to capture contemporaneous influences, as well. In [30], the detection of links in a network is obtained via tests that make use of directed information. The work in [31] provides a two-test method with the goal of detecting strictly causal and nonstrictly causal operators in the network. However, these methods typically do not provide any theoretical guarantees.

In this article, we provide a method, with theoretical guarantees, for the reconstruction of the underlying topology of a large class of dynamic systems when direct feedthroughs are present. We do not impose any restrictions on the underlying graph structure nor on the individual transfer functions as long as there is at least one strictly causal transfer function in every feedback loop. In addition, we show that our technique naturally embeds Granger causality.

This article is organized as follows. In Section II, we introduce some preliminary notions including linear dynamic influence models (LDIMs) [33], the class of models which are the object of investigation in this article; Section III contains the main results, namely two theorems characterizing properties of predictors in LDIMs; these results lead to the derivation of the Mixed Delay (MD) algorithm described in Section IV; Section V makes use of a benchmark example specifically designed to include several dynamic configurations that are challenging to reconstruct and compares MD algorithm with other state-of-the-art methodologies. Finally, this article concludes in Section VI.

II. Preliminary Notions

In order to derive the main results, we first briefly recall some notions of graph theory and then introduce appropriate vector spaces of stochastic processes (see Section II-A) where causal Wiener filtering can be reinterpreted as a projection (see Section II-B) and the class of networks we are investigating (see Section II-C). We denote a directed (undirected) graph G as a pair (V, E) where V is a set of vertices or nodes and E is a set of directed (undirected) edges (see [34]). Following [35], we define the skeleton of a directed graph as the undirected graph obtained by removing all self-loops (edges from one node to itself) and the orientations from the remaining edges.

On a directed graph we also recall what "chains" (or "directed paths") are (see [33]–[35] for the formal definition). A *chain* starting from y_i and ending in y_j is an ordered set of contiguous edges in E

$$((y_{\pi_1}, y_{\pi_2}), (y_{\pi_2}, y_{\pi_3}), \ldots, (y_{\pi_{\ell-2}}, y_{\pi_{\ell-1}}), (y_{\pi_{\ell-1}}, y_{\pi_\ell}))$$

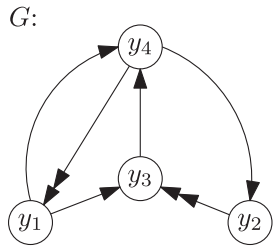

Fig. 1. Multi-arrowed graph G with single-headed set of edges $E_1 = \{(y_1, y_3), (y_3, y_4), (y_4, y_2), (y_1, y_4)\}$ and a double-headed set of edges $E_2 = \{(y_2, y_3), (y_4, y_1)\}$, $E_1 \cap E_2 = \emptyset$. Observe that this graph is recursive.

where $y_{\pi_1} = y_i$, $y_{\pi_\ell} = y_j$. We also use the standard notions of parents, children, ancestors, descendants in a directed graph G [33]. A vertex y_i is a *parent* of a vertex y_j if there is a directed edge from y_i to y_j. In such a case, we also say that y_j is a *child* of y_i. Furthermore, y_i is an *ancestor* of y_j if $y_i = y_j$ or there is a chain from y_i to y_j. In such a case, we also say that y_j is a *descendant* of y_i.

For the developments of this article, we also make use of a special instance of typed graphs, which is an extension of directed graphs. A typed graph is a directed graph which allows for multiple edges of different type between two vertices [36]. In this article, we consider edges of two types: single-headed edges and double-headed edges.

Definition 1 (Multi-arrowed graph): A multi-arrowed graph is a t-ple $G = (V, E_1, E_2)$ where E_1, the set of single-headed edges, and E_2, the set of double-headed edges, are disjoint subsets of $V \times V$.

We can graphically represent a multi-arrowed graph $G = (V, E_1, E_2)$ by drawing a single-headed edge from node y_i to y_j if $(y_j, y_i) \in E_1$ or a double-headed edge from node y_i to y_j if $(y_j, y_i) \in E_2$. An example of a multi-arrowed graph is given in Fig. 1.

Note that all the node relations extend naturally to multi-arrowed graphs by considering $E = E_1 \cup E_2$ as a set of standard edges. For example, vertex y_i is a parent of y_j if there exists an edge (y_i, y_j) that is either in E_1 or in E_2. Furthermore, if $(y_i, y_j) \in E_1$ we say that y_i is a single-headed parent of y_j, while if $(y_i, y_j) \in E_2$ we say that y_i is a double-headed parent of y_j. In particular, we will focus on recursive multi-arrowed graphs.

Definition 2 (Recursive multi-arrowed graph): We say that the multi-arrowed graph $G = (V, E_1, E_2)$ is recursive if in every directed loop there is at least one double-headed arrow.

Observe that the multi-arrowed graph in Fig. 1 is indeed recursive.

A. Transfer Function Span Spaces

We define the space of processes used in the development of our theoretical framework.

Definition 3 (Rationally related random processes): Let $\mathcal{E}$ be a set containing discrete-time scalar, zero-mean, jointly wide-sense stationary random scalar 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 $\Phi_{e_i e_j}(z) = A(z)^{-1}B(z)$, where $A(z)$ and $B(z)$ are polynomials with real coefficients such that $A(z) \neq 0$ for any $z \in \mathbf{C}$, with $|z| = 1$. Then, we say that $\mathcal{E}$ is a set of rationally related random processes.

We also introduce a class of linear operators $\mathcal{F}$ of discrete-time finite dimensional linear time-invariant systems, to transform sets of rationally related random processes.

Definition 4: The set $\mathcal{F}$ is defined as the set of real-rational transfer functions that are analytic and invertible on the unit circle $\{z \in \mathbf{C} |\ |z| = 1\}$. Given a transfer function $H(z) \in \mathcal{F}$, it can be uniquely represented in the time domain by a biinfinite sequence h_k (the impulse response of $H(z)$) satisfying

$$H(z) = \sum_{k=-\infty}^{\infty} h_k z^{-k} \tag{1}$$

for all $|z| = 1$. If $k < 0$ implies $h_k = 0$, then we say that the transfer function is causal. We define the space of causal transfer functions as $\mathcal{F}^+$. If $k \leq 0$ implies $h_k = 0$, then we say that the transfer function is strictly causal.

As a notation, we define $y = H(z)e$ as the process obtained by computing the convolution of e with the impulse response of $H(z)$, namely $y(t) = \sum_{-\infty}^{\infty} h_{t-k}e(k)$ for all $t \in \mathbb{Z}$. For example, $\frac{1}{z}e$ denotes the process e delayed by one time step. For a set of processes $S = \{y_1, \ldots, y_n\}$, we denote by $H(z)S$ the set $\{H(z)y_1, \ldots, H(z)y_n\}$. For example $\frac{1}{z}S$ denotes the set of all the processes in S delayed by one time step. It is possible to filter a finite number of rationally related processes with transfer functions in $\mathcal{F}$ or $\mathcal{F}^+$ to obtain a space of rationally related processes.

Definition 5 (Transfer function spans): Given a set $\mathcal{E}$ of rationally related random processes, we define the transfer function span and the causal transfer function span as

$$tfspan(\mathcal{E}) := \mathcal{F}\{\mathcal{E} = \left\{ y = \sum_k H_k(z)e_k \,|\, e_k \in \mathcal{E}, H_k \in \mathcal{F} \right\}$$

$$ctfspan(\mathcal{E}) := \mathcal{F}^+\mathcal{E} = \left\{ y = \sum_k H_k(z)e_k \,|\, e_k \in \mathcal{E}, H_k \in \mathcal{F}^+ \right\}.$$

The fact that $\mathcal{F}\mathcal{E}$ and $\mathcal{F}^+\mathcal{E}$ are spaces of rationally related processes is an immediate consequence of the Wiener-Khinchin Theorem (see [37]). Also, these two spaces equipped with the inner product

$$\langle y_i, y_j \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{y_i y_j}(e^{i\omega}) d\omega$$

are pre-Hilbert spaces [20]. Such an inner product induces the norm $\|y_i\|^2 = \langle y_i, y_i \rangle$ in the usual way.

Similarly to standard vector spaces, analogous properties of unique representation of an element and subspans inclusion properties hold in a $tfspan$ and $ctfspan$ as well, as summarized by the following proposition.

Proposition II.1: Let $Y = \{y_1, \ldots, y_n\}$, $V = \{v_1, \ldots, v_m\}$, and $W = \{w_1, \ldots, w_\ell\}$ be sets of rationally related processes. The following statements hold:

a) The transfer functions $H_k(z) \in \mathcal{F}$ (or $H_k \in \mathcal{F}^+$), $k = 1, \ldots, n$, are unique if and only if the power spectral density matrix $\Phi_{yy}(e^{i\omega})$ is positive definite for almost all $\omega \in [-\pi, \pi]$.

b) If $v_i \in ctfspan(w_1, \ldots, w_\ell)$, for $i = 1, \ldots, m$, then

$$ctfspan\,(y_1, \ldots, y_n, v_1, \ldots, v_m) \subseteq ctfspan\,(y_1, \ldots, y_n, w_1, \ldots, w_\ell)\ .$$

c) Let $q = \sum_{k=1}^{n} H_k(z)y_k$, with $H_k(z) \in \mathcal{F}$ (or $H_k(z) \in \mathcal{F}^+$) for $k = 1, \ldots, n$. If $H_k(z) \in \mathcal{F}$, for $k = 1, \ldots, n$, then

$$tfspan\,(q, y_2, \ldots, y_n) = tfspan\,(y_1, y_2, \ldots, y_n)$$

if and only if $H_1(z)^{-1} \in \mathcal{F}$. Similarly, if $H_k \in \mathcal{F}^+$ for $k = 1, \ldots, n$,

$$ctfspan\,(q, y_2, \ldots, y_n) = ctfspan\,(y_1, y_2, \ldots, y_n)$$

if and only if $H_1(z)^{-1} \in \mathcal{F}^+$.

Proof: See Appendix A– C. ■

B. Causal Wiener Filtering Formulation for Causal Transfer Functions Spans

The goal of this section is to reformulate the well-known notion of causal Wiener filtering in terms of causal transfer function spans.

Proposition II.2 (Causal Wiener Filter): Let v and $y_1, \ldots, y_n$ be processes in the space $\mathcal{F}\mathcal{E}$ for some set of rationally related processes $\mathcal{E}$. Let $y = (y_1, \ldots, y_n)^T$ and let $Y = ctfspan(y_1, \ldots, y_n)$. Consider the problem

$$\hat{v} = \arg\min_{q \in Y} \|v - q\|^2 .$$

The solution $\hat{v} \in Y$ exists and admits the form $\hat{v} = W(z)y$. If $\Phi_{yy}(e^{i\omega})$ is positive definite for $\omega \in [-\pi, \pi]$, the transfer function $W(z)$ is unique and known as the causal Wiener filter.

The proof of this proposition can be found in [20]. In general, the Wiener filter $W(z)$ estimating v from $y_1, \ldots, y_n$ is a transfer vector operating on the vector process $y = (y_1, \ldots, y_n)^T$. We introduce a notation to extract the individual components of $W(z)$.

Definition 6: Let v and $y_1, \ldots, y_n$ be processes in a space of rationally related processes. Let $\hat{v}$ be the causal Wiener estimate of v from $y_1, \ldots, y_n$

$$\hat{v} = W(z)y = \sum_{k=1}^{n} W_k(z)y_k.$$

We denote $W_{v,[y_k]|(y_1,\ldots,y_n)}(z) := W_k(z)$, namely the component of the Wiener filter $W(z)$ that operates on the process y_k to estimate v from $y_1, \ldots, y_n$.

We can use the uniqueness of the process $\hat{v} \in Y = ctfspan(y_1, \ldots y_n)$ that best approximates a process v in the least square sense to derive the following Lemma, which is going to be used in subsequent proofs.

Lemma II.3: (Hilbert parallel translation) Let v and $y_1, \ldots, y_n$ be processes in the space $\mathcal{FE}$. Let $Y = ctfspan(y_1, \ldots y_n)$. For all $w \in Y$

$$\arg\min_{q\in Y}||v-q||^2 = w + \arg\min_{q\in Y}||(v-w)-q||^2.$$

Proof: See Appendix D. ■

The next two notions help us express the conditional uncorrelation of stochastic processes using geometric concepts.

Definition 7 (Orthogonality to $ctfspans$): Let v and $y_1, \ldots, y_n$ be $n+1$ rationally related processes. We say that v is orthogonal to $ctfspan(y_1, \ldots, y_n)$, denoted as $v \perp ctfspan(y_1, \ldots, y_n)$, if and only if the causal Wiener filter estimating v from $y_1, \ldots, y_n$ is the zero transfer vector.

Proposition II.4: We have that $v \perp ctfspan(y_k)$ for all $k = 1, \ldots, n$, if and only if $v \perp ctfspan(y_1, \ldots, y_n)$.

Proof: See Appendix E. ■

Furthermore, orthogonality can be generalized considering a form of "conditional" orthogonality and the concept of Wiener separated processes [38] helps us express that.

Definition 8 (Wiener Separated Processes): Let $v, y_1, \ldots y_n$ be $n+1$ processes in the space $\mathcal{FE}$ and denote $y = (y_1, \ldots, y_n)$. We say that v and y_i are Wiener-separated given the set of processes $S = \{y_k | k \neq i\}$, if $W_{v,[y_i]|(y_1,\ldots,y_n)}(z) = 0$. We denote this relation as $wsep(v, S, y_i)$. If, instead, S is such that $W_{v,[y_i]|(y_1,\ldots,y_n)}(z) \neq 0$, we say that S does not Wiener separate v and y_i and we denote that as $\neg wsep(v, S, y_i)$.

Note that if the separating set S is empty, Wiener separation reduces to standard orthogonality.

C. Generative Models: LDIMs

The aim of this section is to introduce the class of LDIMs [33], that will be the focus of the network reconstruction methodologies developed in the main section.

The underlying intuition is to construct a class of models describing the input/output interconnectivity of components in a network of dynamical systems. It is assumed that the dynamics of each component (node) in the network is represented by a scalar random process y_j (for $j = 1, \ldots, n$) which is given by the superposition of its "autonomous behavior" e_j and the "influences" of some other "parent nodes" through dynamic links. The autonomous behavior on each node is assumed independent of the others. If a certain component "influences" another, a directed edge can be drawn and a directed graph can be obtained. This intuition can be formalized into the definition of LDIMs.

Definition 9 (Linear Dynamic Influence Model): An LDIM $\mathcal{G}$ is defined as a pair $(H(z), e)$ where

1) $e = (e_1, \ldots, e_n)^T$ is a vector of n rationally related random scalar processes $e_1, \ldots, e_n$, such that $\Phi_e(z)$ is diagonal, namely $\Phi_{e_i e_j} = 0$ for $i \neq j$.
2) $H(z)$ is an $n \times n$ transfer matrix with entries in $\mathcal{F}$. $H(z)$ is termed as the "dynamics" of the LDIM.

The output processes $\{y_j\}_{j=1}^n$ of the LDIM are defined as $y_j = e_j + \sum_{i=1}^n H_{ji}(z) y_i$, or in a more compact way

$$y = e + H(z)y \tag{2}$$

where $y = (y_1, \ldots, y_n)^T$.

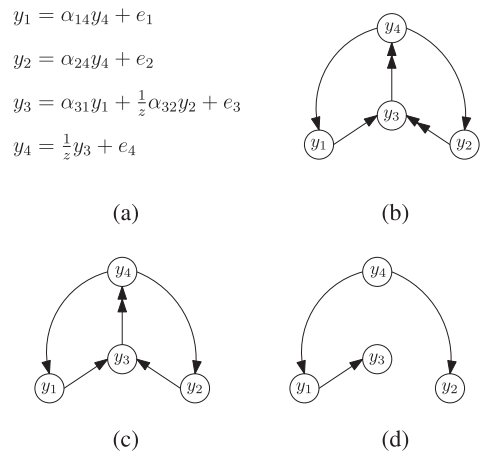

Fig. 2. (a) LDIM defined by the scalar parameters $\alpha_{14}, \alpha_{24}, \alpha_{31}, \alpha_{32}$. (b) Perfect graphical representation of the LDIM under the assumption that all the parameters are not zero. (c) Graphical representation of the LDIM. (d) Graph of instantaneous propagations of the perfect graphical representation of the LDIM (as per Definition 16).

Note that the term e in the LDIM $\mathcal{G} = (H(z), e)$ models an unknown external input and that if e is the zero vector, the output y is also zero.

For an LDIM $\mathcal{G} = (H, e)$, multi-arrowed graphs can be used to suggestively represent sparsity properties of the matrix $H(z)$ and the strict causality of some of its entries.

Definition 10: (Multi-arrowed graph representation of an LDIM) Let $\mathcal{G} = (H, e)$ be an LDIM with output processes $y_1, \ldots, y_n$. Let $V := \{y_1, \ldots, y_n\}$ and let E_1 and E_2 be two disjoint subsets of $V \times V$ such that

a) $(y_i, y_j) \notin E_1 \cup E_2$ implies $H_{ji} = 0$
b) $(y_i, y_j) \notin E_1$implies $H_{ji}(z)$ is strictly causal.

We say that the multi-arrowed graph $G = (V, E_1, E_2)$ is a graphical representation of the LDIM. Furthermore, if the implications (a) and (b) hold also in the opposite direction, we say that $G = (V, E_1, E_2)$ is a perfect graphical representation of the LDIM.

By extension, when considering a specific graphical representation of an LDIM, nodes and edges of an LDIM will mean nodes and edges of the specific graphical representation. Observe that the graphical representation of an LDIM provides partial information about the entries of the matrix $H_{ji}(z)$ that are identically zero or that are strictly causal. Indeed, if $(y_i, y_j) \notin E_1$, then $H_{ji}(z)$ is definitely strictly causal and if, furthermore, $(y_i, y_j) \notin E_2$, we also know that $H_{ji}(z) = 0$. However, unless it is known that the graphical representation is perfect, the presence of (y_i, y_j) in $E_1 \cup E_2$ has to be interpreted as $H_{ji}(z)$ is "not necessarily different from zero" and the presence of (y_i, y_j) in E_1 has to be interpreted as "not necessarily strictly causal." These concepts are illustrated in Fig. 2(a)–(c).

We introduce additional terminology related to an LDIM.

Definition 11: We say that an LDIM $\mathcal{G}$ is causal if every entry of $H(z)$ belongs to $\mathcal{F}^+$.

Definition 12: We say that an LDIM $\mathcal{G}$ is causally well-posed if, for every set of indices $I \subseteq \{1, \ldots, n\}$, $(I - H_{II}(z))^{-1}$ exists and is causal.

Definition 13: (Direct feedthrough) Let $\mathcal{G} = (H(z), e)$ be a causal LDIM. If there exists $H_{ji}(z)$ such that $\lim_{z\to\infty} H_{ji}(z) \neq 0$, we say that a nonzero direct feedthrough is present in the LDIM.

Definition 14: We say that an LDIM $\mathcal{G}$ is topologically detectable if $\Phi_{e_i e_i}(\omega) > 0$ for any $\omega \in [-\pi, \pi]$ and for any $i = 1, \ldots, n$.

The condition of topological detectability prevents pathological situations where a single node is not excited by an external input.

Definition 15: We say that an LDIM $\mathcal{G}$ is recursive if its perfect graphical representation is recursive.

Observation 1: Observe that in a recursive LDIM there is always a strictly causal transfer function in each feedback loop. Also if a graphical representation of an LDIM is recursive, then the LDIM is necessarily recursive, but the inverse implication does not hold in general.

III. Main Results

We consider the following problem.

Problem: Given a recursive LDIM, infer information about its perfect graphical representation from observational data.

The contribution of this article to this problem is an algorithm capable of determining the skeleton of the perfect graphical representation of a recursive LDIM. Such an algorithm can also infer the orientation of the double-headed arrows connecting distinct nodes in the graphical representation, while the single-headed arrows are left unoriented.

We start presenting a result (Theorem III.3) that fully extends Granger causality and enables, under some technical conditions, to correctly detect whether the past of one process directly influences the present of another, even in the presence of direct feedthroughs within the network. This result translates into a sufficient condition (Corollary III.5) that could be formulated as a necessary and sufficient condition if it were exactly known which links of the network contain a nonzero direct feedthrough. This knowledge is typically not available when reconstructing a network from observational data. For this reason, we provide a second result (Theorem III.6) that can detect the presence of direct feedthroughs between nodes. This second result again translates only into a sufficient condition (Corollary III.7), but, when combined with Corollary III.5, gives tighter conditions for the detection of a link. Corollary III.5 and Corollary III.7 are then merged to define an algorithm for the reconstruction of the skeleton of a recursive LDIM. After the application of Corollary III.7 and Corollary III.5, Theorem III.3 is used to determine the orientation of the double-headed arrows.

A. Extended Granger Causality

Given a graphical representation, we define a graph representing all the transfer functions in an LDIM with potentially nonzero feedthrough components.

Definition 16 (Graph of instantaneous propagations): Consider a well-posed, causal, and topologically detectable LDIM $(H(z), e)$ with outputs $y_1, \ldots, y_n$. Let G be a graphical representation of the LDIM. The graph of instantaneous propagations for the LDIM is the graph $G^{\prime}$ where there is an edge from y_i to y_j if and only if there is a single-headed link from y_i to y_j in G.

For example, assuming the LDIM of Fig. 2(a) and its graphical representation of Fig. 2(b), the associated graph of instantaneous propagations is represented in Fig. 2(d). The following proposition plays an important role in the development of the main results.

Proposition III.1: Consider a causal, causally well-posed and topologically detectable LDIM with outputs $y = (y_1, \ldots, y_n)^T$. Assume G is a recursive graphical representation of the LDIM and let $G^{\prime}$ be the associated graph of instantaneous propagations. Define y_{A_j} as the vector of processes containing all the ancestors of y_j in $G^{\prime}$. Let $y_{\overline{A}_j}$ be the vector of processes which are not in y_{A_j}. Then, $y_j \in ctfspan(e_{A_j}, \frac{1}{z}e_{\overline{A}_j})$.

Proof: First observe that $G^{\prime}$ is a directed acyclic graph (DAG), otherwise G would not be recursive. We prove the statement by induction on the number of ancestors of y_j in $G^{\prime}$. If y_j has itself as the only ancestor in $G^{\prime}$, we have that H_{ji} is strictly causal for all i. Thus, $y_j \in ctfspan(e_j, \frac{1}{z}y)$. Since the LDIM is causal $y \in ctfspan(e_j, e_{\overline{A}_j})$ where $\overline{A}_j$ denotes the set of nonancestors of y_j, implying that $\frac{1}{z}y \in ctfspan(\frac{1}{z}e_j, \frac{1}{z}e_{\overline{A}_j})$. By Proposition II.1(b), this gives

$$y_j \in ctfspan\left(e_j, \frac{1}{z}y\right) \subseteq ctfspan\left(e_j, \frac{1}{z}e_j, \frac{1}{z}e_{\overline{A}_j}\right) \subseteq ctfspan\left(e_j, \frac{1}{z}e_{\overline{A}_j}\right).$$

This proves the base step. Now assume that the statement holds if a node has a number of ancestors up to $K \geq 1$. If y_j has $K+1$ ancestors in $G^{\prime}$ for $K \geq 1$, we can write

$$y_j = e_j + \sum_{y_i \in y_{A_j}} H_{ji}(z) y_i + \sum_{y_i \in y_{\overline{A}_j}} H_{ji}(z) y_i.$$

For $y_i \in y_{\overline{A}_j}$, we have that H_{ji} is necessarily strictly causal, otherwise y_i would be parent of y_j in $G^{\prime}$. This yields

$$y_j \in ctfspan\left(e_j, y_{A_j}, \frac{1}{z}y_{\overline{A}_j}\right).$$

Since $G^{\prime}$ is a DAG, all elements in y_{A_j} have at most K ancestors and their ancestors are all ancestors of y_j. Thus, we have from the induction hypothesis that $ctfspan(y_{A_j}) \subseteq ctfspan(e_{A_j}, \frac{1}{z}e_{\overline{A}_j})$, where A_j denotes the set of ancestors of y_j. By inclusion of $ctfspans$ [see Proposition II.1(b)], we get

$$ctfspan\left(e_j, y_{A_j}, \frac{1}{z}y_{\overline{A}_j}\right) \subseteq ctfspan\left(e_j, e_{A_j}, \frac{1}{z}e_{\overline{A}_j}, \frac{1}{z}y_{\overline{A}_j}\right) \subseteq ctfspan\left(e_j, e_{A_j}, \frac{1}{z}e_{\overline{A}_j}, \frac{1}{z}e\right) \subseteq ctfspan\left(e_{A_j}, \frac{1}{z}e_{\overline{A}_j}, \frac{1}{z}e\right) \subseteq ctfspan\left(e_{A_j}, \frac{1}{z}e_{\overline{A}_j}\right).$$

■

The next Proposition gives an explicit expression for the Wiener filter predicting a node of an LDIM from its parents in a graphical representation.

Proposition III.2 (Prediction from parent set): Consider a causal, causally well-posed, and topologically detectable LDIM with outputs $y_1, \ldots, y_n$ and assume that G is a recursive graphical representation for it. Define

$$S_s = \{y_k | y_k \text{ is double-headed parent of } y_j, y_k \neq y_j\}$$
$$S_c = \{y_k | y_k \text{ is single-headed parent of } y_j, y_k \neq y_j\}.$$

The Wiener filter estimate $\hat{y}_j$ of y_j from $S = S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j\}$ is

$$\hat{y}_j = \left(H_{jj} + \frac{F_j(z)}{z}\right) y_j + \sum_{y_k \in S_c \cup S_s} \left(\mathcal{I} - \frac{F_j(z)}{z}\right) H_{jk} y_k.$$

where F_j is the causal Wiener filter estimating e_j from $\frac{1}{z}e_j$.

Proof: From Lemma II.3 and the fact that H_{jj} and H_{jk}, for $y_k \in S_s$, are strictly causal, the optimal estimate $\hat{y}_j$ is given by

$$\begin{aligned}\hat{y_j} &= \arg\min_{q \in ctfspan(S)} \|y_j - q\|^2 \\ &= zH_{jj}\frac{1}{z}y_j + \sum_{y_k \in S_c} H_{jk}y_k + \sum_{y_k \in S_s} zH_{jk}\frac{1}{z}y_k \\ &\quad + \arg\min_{q \in ctfspan(S)} \|e_j - q\|^2 \\ &= H_{jj}y_j + \sum_{y_k \in S_c \cup S_s} H_{jk}y_k + \arg\min_{q \in ctfspan(S)} \|e_j - q\|^2.\end{aligned}$$

Let $\epsilon_j = e_j - \hat{e_j}$, where $\hat{e_j} = F_j(z)\frac{1}{z}e_j$ is the one step ahead predictor for e_j, where e_j is the independent component of y_j. Observe that $\hat{e}_j \in ctfspan(S)$. Indeed, we have that $\hat{e}_j \in ctfspan(\frac{1}{z}e_j)$ and

$$e_j = y_j - H_{jj}y_j - \sum_{y_k \in S_c \cup S_s} H_{jk}y_k$$

implying that

$$\frac{1}{z}e_j \in ctfspan\left(\frac{1}{z}y_j, \frac{1}{z}S_c, \frac{1}{z}S_s\right).$$

Again, from Lemma II.3, we can write $\hat{y}_j$ as

$$\begin{aligned}\hat{y_j} &= H_{jj}y_j + \sum_{y_k \in S_c \cup S_s} H_{jk}y_k + \hat{e}_j \\ &\quad + \arg\min_{q \in ctfspan(S)} \|\epsilon_j - q\|^2.\end{aligned}$$

We observe that $\epsilon_j \perp ctfspan(\frac{1}{z}y_k)$ for all k. Indeed, since the LDIM is causal we have that $\frac{1}{z}y_k \in ctfspan(\frac{1}{z}e)$ and we also have that $\epsilon_j = e_j - \hat{e}_j$. Observe that ϵ_j is orthogonal to $ctfspan(\frac{1}{z}e_j)$ as ϵ_j is the one-step-ahead prediction error. Also, since $\epsilon_j = e_j - \frac{F_j(z)}{z}e_j$ and $\Phi_e(z)$ is diagonal, we have that ϵ_j is orthogonal to $ctfspan(\frac{1}{z}e_k)$ for $k \neq j$. From Proposition II.4, we have that $\epsilon_j \perp ctfspan(\frac{1}{z}e) = ctfspan(\frac{1}{z}y)$. Thus, in particular, $\epsilon_j \perp ctfspan(\frac{1}{z}y_k)$ for $y_k \in S_s$.

Now, we show that $\epsilon_j \perp ctfspan(y_k)$ for $y_k \in S_c$. For $y_k \in S_c$ let y_{A_k} be the set of ancestors of y_k in the graph of instantaneous propagations and $y_{\overline{A}_k}$ be the remaining nodes in the LDIM. By Proposition III.1, we have that $y_k \in ctfspan(e_{A_k}, \frac{1}{z}e_{\overline{A}_k})$. Since there is a single headed edge from y_k to y_j in G and G is recursive, there must be at least one double-headed edge on every chain from y_j to y_k in G. Thus, y_j is not an ancestor of y_k in the graph of instantaneous propagations, implying that for every $y_k \in S_c$, $y_k \in ctfspan(e_{A_k}, \frac{1}{z}e_{\overline{A}_k})$, where $e_j \notin e_{A_k}$. As Φ_e is diagonal, $\epsilon_j \perp ctfspan(e_k)$ for all $k \neq j$. Also, $\epsilon_j \perp ctfspan(\frac{1}{z}e_j)$ because ϵ_j is the error obtained from the one-step ahead prediction of e_j using $\frac{1}{z}e_j$. Thus, $\epsilon_j \perp ctfspan(e_{A_k}, \frac{1}{z}e_{\overline{A}_k})$ for all $y_k \in S_c$, leading to $\epsilon_j \perp ctfspan(y_k)$ for all $y_k \in S_c$.

As a consequence, by Proposition II.4, $\epsilon_j \perp ctfspan(S)$ yielding

$$\begin{aligned}\hat{y_j} &= H_{jj}y_j + \sum_{y_k \in S_c \cup S_s} H_{jk}y_k + \hat{e}_j \\ &= H_{jj}y_j + \sum_{y_k \in S_c \cup S_s} H_{jk}y_k + \frac{F_j(z)}{z} \\ &\quad \times \left(y_j - \sum_{y_k \in S_s \cup S_c} H_{jk}y_k\right) \\ &= \left(H_{jj} + \frac{F_j(z)}{z}\right) y_j + \sum_{y_k \in S_c \cup S_s} \left(\mathcal{I} - \frac{F_j(z)}{z}\right) H_{jk}y_k.\end{aligned}$$

■

Proposition III.2 provides us with an analytic expression for the least square estimate of a signal y_j from its parents in any given recursive graphical representation of the LDIM. Since least square estimates can also be computed from data (i.e., via linear regressions), this analytical expression enables us to obtain information about the graphical representation of the LDIM using observational time series. Specifically, we would like to infer the presence of a nonzero direct influence from a signal y_i to the signal y_j (that is an edge from y_i to y_j in every graphical representation or $H_{ji}(z) \neq 0$). This is precisely the kind of problem that Granger Causality tries to tackle. However, it is well-known that Granger Causality might detect spurious causal links in the presence of contemporaneous influences [39]. The following theorem is the first of two tools to address the problem of determining the presence of a direct form of influence from y_i to y_j even in presence of contemporaneous influences by using a test based on Wiener-separation. Furthermore, this result has the fundamental property of recovering Granger Causality as a special case.

Theorem III.3 (Extended Granger Causality): Consider a causal, causally well-posed, and topologically detectable LDIM $(H(z), e)$ with recursive graphical representation G. For distinct y_i and y_j, define

$$S_s = \{y_k | y_k \text{ is double-headed parent of } y_j \text{ and } y_k \neq y_i, y_j\}$$
$$S_c = \{y_k | y_k \text{ is single-headed parent of } y_j \text{ and } y_k \neq y_i, y_j\}.$$

We have that $wsep(y_j, S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j\}, \frac{1}{z}y_i)$ if and only if

a) $H_{ji}(z) = 0$; or

b) $H_{ji}(\infty) := lim_{z\to\infty} H_{ji}(z) \neq 0$ and

$$H_{ji}(\infty) W_{y_i,[\frac{1}{z}y_i]|(\frac{1}{z}y_i,\frac{1}{z}y_j,S_c,S_s)}(z) = -\left[zH_{ji} - zH_{ji}(\infty) - F_j(z)H_{ji}\right].$$

Proof: We prove the equivalence of the two conditions in the scenarios: (i) when H_{ji} is strictly causal; and (ii) when it is causal, but not strictly causal.

If H_{ji} is strictly causal and G has a single headed arrow from y_i to y_j, then define the graph G' as the graph obtained from G by having instead a double-headed arrow from y_i to y_j. If H_{ji} is strictly causal and G has a double headed arrow from y_i to y_j, then define the graph G' as G. In both cases G' is still a graphical representation of the LDIM. By applying Proposition III.2 to G' we get the expression

$$\hat{y}_j = \left(H_{jj} + \frac{F_j(z)}{z}\right) y_j + \left(\mathcal{I} - \frac{F_j(z)}{z}\right) H_{ji} y_i + \sum_{y_k \in (S_c \cup S_s)} \left(\mathcal{I} - \frac{F_j(z)}{z}\right) H_{jk} y_k$$

where $\hat{y}_j$ is the Wiener estimate of y_j from $S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j, \frac{1}{z}y_i\}$. Since $F_j(z)$ is causal, we have $wsep(y_j, S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j\}, \frac{1}{z}y_i)$ if and only if $H_{ji} = 0$.

If H_{ji} is causal, but not strictly causal, G necessarily has a single-headed arrow from y_i to y_j. If we were to apply Proposition III.2 to G, we would get

$$\hat{y}'_j = \left(H_{jj} + \frac{F_j(z)}{z}\right) y_j + \left(\mathcal{I} - \frac{F_j(z)}{z}\right) H_{ji} y_i + \sum_{y_k \in (S_c \cup S_s)} \left(\mathcal{I} - \frac{F_j(z)}{z}\right) H_{jk} y_k. \quad (3)$$

Namely, we get the expression of the Wiener estimate of y_j using the processes in $S' = S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j, y_i\}$. However, what we want is the Wiener estimate of y_j from the processes in $S = S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j, \frac{1}{z}y_i\}$. For this purpose, let us consider $\hat{y}_i$, the Wiener estimate of y_i from $S = S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j, \frac{1}{z}y_i\}$. Write $y_i = y_i^{\perp} + \hat{y}_i$, where $y_i^{\perp}$ is the prediction error when Wiener estimating y_i from the processes in S. Going back to the expression in (3) and substituting for y_i with $y_i^{\perp} + \hat{y}_i$, as well as writing $H_{ji}y_i = (zH_{ji} - zH_{ji}(\infty))\frac{1}{z}y_i + H_{ji}(\infty)y_i$ we get

$$\hat{y}'_j = \left(H_{jj} + \frac{F_j(z)}{z}\right) y_j + \sum_{y_k \in (S_c \cup S_s)} \left(\mathcal{I} - \frac{F_j(z)}{z}\right) H_{jk} y_k + H_{ji}(\infty) y_i^{\perp} + H_{ji}(\infty)\hat{y}_i + \left[zH_{ji} - zH_{ji}(\infty) - F_j(z)H_{ji}\right] \frac{y_i}{z}. \quad (4)$$

Let ϵ_j be the prediction error when predicting y_j from $S' = S_c \cup \frac{1}{z}S_s \cup \frac{1}{z}y_j, y_i$. Namely $\epsilon_j = y_j - \hat{y}'_j$. From (4), we can write an expression for the estimate $\hat{y}_j$ of y_j from S

$$\begin{aligned}\hat{y}_j &= \text{arg min}_{q \in ctfspan(S)} \|y_j - q\|^2 \\ &= \text{arg min}_{q \in ctfspan(S)} \left\|\hat{y}'_j + \epsilon_j - q\right\|^2 \\ &= \left(H_{jj} + \frac{F_j(z)}{z}\right) y_j + \sum_{y_k \in (S_c \cup S_s)} \left(\mathcal{I} - \frac{F_j(z)}{z}\right) H_{jk} y_k \\ &\quad + H_{ji}(\infty)\hat{y}_i + \left[zH_{ji} - zH_{ji}(\infty) - F_j(z)H_{ji}\right] \frac{y_i}{z} \\ &\quad + \text{arg min}_{q \in ctfspan(S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j, \frac{1}{z}y_i\})} \\ &\quad \times \left\|H_{ji}(\infty) y_i^{\perp} + \epsilon_j - q\right\|^2\end{aligned}$$

where the last step comes from Lemma II.3.

Note that $y_i^{\perp} \perp ctfspan(S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j, \frac{1}{z}y_i\})$. Also, $\epsilon_j \perp ctfspan(S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j, y_i\})$, which by Proposition II.1(c) implies that $\epsilon_j \perp ctfspan(S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j, \frac{1}{z}y_i\})$. Thus, we have that the Wiener estimate $\hat{y}_j$ of y_j from $S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j, \frac{1}{z}y_i\}$ is

$$\begin{aligned}\hat{y}_j = \text{arg min}_{q \in ctfspan(S)} \|y_j - q\| &= \left(H_{jj} + \frac{F_j(z)}{z}\right) y_j + H_{ji}(\infty)\hat{y}_i + \left[H_{ji} - H_{ji}(\infty) - F_j(z)H_{ji}\right] \frac{y_i}{z} \\ &\quad + \sum_{y_k \in S_c \cup S_s} \left(\mathcal{I} - \frac{F_j(z)}{z}\right) H_{jk} y_k.\end{aligned}$$

We have $wsep(y_j, S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j\}, \frac{1}{z}y_i)$ if and only there is no dependence on $\frac{1}{z}y_i$ in the expression of $\hat{y}_j$. The only two terms depending on $\frac{1}{z}y_i$ are $H_{ji}(\infty)\hat{y}_i$ and $[H_{ji} - H_{ji}(\infty) - F_j(z)H_{ji}]\frac{y_i}{z}$. The component in $\hat{y}_i$ that depends on $\frac{y_i}{z}$ is $W_{y_i,[\frac{1}{z}y_i]|S}(z)\frac{1}{z}y_i$. Thus, there is no dependence on $\frac{y_i}{z}$ if and only if b) holds. ■

Theorem III.3 has, as a special case, an already known result about Granger Causality, namely that if the transfer functions of an LDIM are all known to be strictly causal, Granger Causality provides a consistent reconstruction of the LDIM graph (see [40, Th. 35]). Now, this result becomes a straightforward corollary of Theorem III.3.

Corollary III.4: (Consistent LDIM graph reconstruction via Granger Causality)

Consider an LDIM $(H(z), e)$ with output $y = (y_1, \ldots, y_n)^T$ where each entry of $H(z)$ is strictly causal. We have that, for distinct y_i and y_j

$$H_{ji}(z) \neq 0 \Leftrightarrow W_{y_j,[\frac{1}{z}y_i]|(\frac{1}{z}y)}(z) \neq 0.$$

Proof: Let G be a graph where there is a double headed edge from each node to any other node and itself. As each entry of $H(z)$ is strictly causal, G is a graphical representation of the LDIM. Now, since all edges in G have a double headed arrow, the hypotheses of Theorem III.3 are met with $S_c = \emptyset$ and $S_s = \{y_k | k \neq i, j\}$. As a consequence, since $H_{ji}(z)$ is strictly causal, we have that $H_{ji}(z) \neq 0 \Leftrightarrow W_{y_j,[\frac{1}{z}y_i]|(\frac{1}{z}y)}(z) \neq 0$. ■

In other words, Corollary III.4 gives an immediate way to reconstruct the perfect representation of a strictly causal LDIM from data: the double headed link (y_i, y_j) is not present if and only if $\frac{1}{z}y_i$ (the past of y_i) is not needed to predict y_j, given the knowledge of all the nodes' past. This is nothing but the test proposed by Granger [24], that can be interpreted in our framework as testing if $wsep(y_j, S \cup \{\frac{1}{z}y_j\}, \frac{1}{z}y_i)$ with $S = S_c \cup \frac{1}{z}S_s$ and $S_c = \emptyset$ and $S_s = \{y_1, \ldots, y_n\} \setminus \{y_i, y_j\}$.

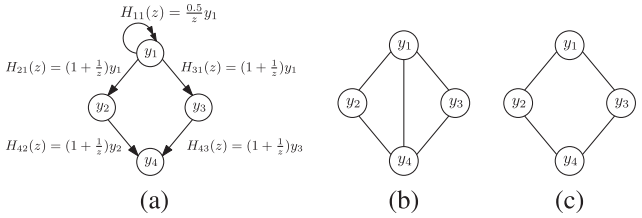

Fig. 3. (a) LDIM and its perfect graphical representation. (b) Skeleton of the perfect graphical representation as reconstructed by Granger causality, which detects a spurious link between y_1 and y_4. (c) Skeleton of the perfect graphical representation as reconstructed by using Corollary III.5.

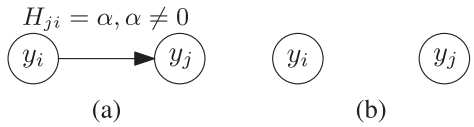

Fig. 4. (a) LDIM with white noise independent components and its perfect graphical representation. (b) Both Granger and Corollary III.5 cannot detect the link between y_i and y_j and the empty graph is inferred.

However, Theorem III.3 is a more general result, since it can provide information to reconstruct the topology of a recursive LDIM $(H(z), e)$ even when it is not known that all entries of $H(z)$ are strictly causal. Since a graphical representation of the LDIM is not known, we cannot apply Theorem III.3 by testing the condition $wsep(y_j, S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j\}, \frac{1}{z}y_i)$ since we cannot appropriately select the sets S_s and S_c. Indeed, this would require to know which parents of y_j have nonzero direct feedthrough components. A possible (naive) approach is to test $wsep(y_j, S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j\}, \frac{1}{z}y_i)$ for every possible choice of disjoint S_c and S_s, with $S_s, S_c \subseteq y \setminus \{y_i, y_j\}$. If we have that $\neg wsep(y_j, S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j\}, \frac{1}{z}y_i)$, for all possible choices of S_s and S_c, then necessarily we have $H_{ji}(z) \neq 0$. This is summarized by the following corollary.

Corollary III.5 (Link detection via extended Granger Causality): Consider a topologically detectable, causal, causally well-posed and recursive LDIM $(H(z), e)$ with output $y = (y_1, \ldots, y_n)^T$. Let $y_i \neq y_j$. If $\neg wsep(y_j, S_c \cup \frac{1}{z}S_s \cup \{\frac{1}{z}y_j\}, \frac{1}{z}y_i)$ for all disjoint S_c and S_s, with $S_s, S_c \subseteq y \setminus \{y_i, y_j\}$, then $H_{ji}(z) \neq 0$.

Proof: It is a direct consequence of Theorem III.3. ■

In the context of this article, we limit ourselves to searching among every possible choice of disjoint S_c and S_s, with $S_s, S_c \subseteq y \setminus \{y_i, y_j\}$ for finding a separating set for simplicity of exposition. Smarter searching strategies than an exhaustive search among all those subsets of nodes are possible as later highlighted in Observation 2 when discussing the fundamental algorithm derived in this article.

We provide an example that highlights the limitations of Granger Causality and the wider applicability of Corollary III.5 and Theorem III.3 when feedthrough components are present. Consider an LDIM with perfect graphical representation as shown in Fig. 3(a). Since feedthrough components are present, Granger causality is not guaranteed to correctly detect the network links. It can be verified that Granger causality is able to correctly detect the direct influence of y_1 on both y_2 and y_3, as well as the direct influences of both y_2 and y_3 on y_4. However, because of the direct feedthrough components in the chains $y_1 \to y_2 \to y_4$ and $y_1 \to y_3 \to y_4$, Granger causality also finds an incorrect link from y_1 to y_4, as shown in Fig. 3(b). The application of Corollary III.5 correctly detects the links $y_1 \to y_2, y_1 \to y_3, y_2 \to y_4$, and $y_3 \to y_4$, as Granger causality does. Yet, in this case, no link from y_1 to y_4 is detected since $wsep(y_4, (y_2, y_3, \frac{1}{z}y_4), \frac{1}{z}y_1)$ holds, as shown in Fig. 3(c).

This does not mean that we can conclude that there is no actual link from y_1 to y_4 because Corollary III.5 is not a necessary and sufficient condition. Indeed, instances where two processes $\frac{1}{z}y_i$ and y_j can be Wiener separated by a set S, but there is a nonzero transfer function between y_i and y_j, can occur in situations of practical relevance. A simple example is the following. Consider an LDIM consisting of two signals only, y_i and y_j, with white noise independent components e_i and e_j, as shown in Fig. 4(a). Let the only transfer function in the LDIM be a scalar $H_{ji}(z) = \alpha$, $\alpha \neq 0$. Namely, the only direct influence is a direct feedthrough from y_i to y_j without any strictly causal part. Granger Causality cannot detect the link because the past of y_i is not correlated with the present of y_j, namely $\frac{1}{z}y_i$ does not help predict $y_j(t)$. Thus Granger Causality infers the skeleton of Fig. 4(b). This example and the previous one illustrate that Granger causality can be susceptible to both false negatives and false positives. Corollary III.5, instead, is susceptible to false negatives only. Indeed, this two-node example is a case where Corollary III.5 is of no help because of a false negative situation. It is immediate to verify that $W_{y_j|\frac{1}{z}y_i}(z) = 0$ implying that $wsep(y_j, \emptyset, \frac{1}{z}y_i)$, but $H_{ji} \neq 0$. Thus, also Corollary III.5 infers the empty graph shown in Fig. 4(b).

Theorem III.3, being a necessary and sufficient condition, helps us understand this case in a better detail. Since we have $wsep(y_j, \emptyset, \frac{1}{z}y_i)$, but $H_{ji}(\infty) \neq 0$ we necessarily need to have condition b) verified. Indeed, since e_j is white, we have $F_j(z) = 0$. Also, $S = \emptyset$ and y_i is white, giving $W_{y_i|\frac{1}{z}y_i} = 0$. Thus, condition b) of Theorem III.3 becomes

$$\alpha \cdot 0 = [z\alpha - z\alpha - 0]$$

which is trivially verified.

Motivated by this simple example, we would like to find ways to supplement the test provided by Corollary III.5 to reduce the incidence of false negatives. If it were known that $H_{ji}(z)$ has a direct feedthrough component, we would immediately conclude that there is a link between y_i and y_j. On the other hand, if it were known that $H_{ji}(z)$ has no feedthrough component, we could apply Theorem III.3 avoiding the troublesome situations where condition b) is verified. This motivates the result developed in next section, which basically consists in a test to determine links with nonzero feedthrough component.

B. Test for Strict Causality

In the previous section, we saw that it would be useful to determine whether the transfer function $H_{ji}(z)$, indicating the direct influence from y_i to y_j, has a nonzero direct feedthrough

component. The following theorem, based on preliminary results described in [41], provides a test towards such a goal using the notion of Wiener filtering once again.

Theorem III.6 (Strict causality test): Consider a causal, causally well-posed, topologically detectable, and recursive LDIM with output signals $y = (y_1, \ldots, y_n)^T$. Let y_i and y_j be distinct, and such that the transfer functions H_{ji} and H_{ij} are strictly causal, namely $H_{ji}(\infty) = H_{ij}(\infty) = 0$. Then, there exist two disjoint sets of processes $S^+, S^- \subseteq \{y_1, \ldots, y_n\} \setminus \{y_i, y_j\}$, such that the component of the causal Wiener filter estimating y_j from $y_i, \frac{1}{z}y_j, S^+, \frac{1}{z}S^-$ associated with y_i is strictly causal, namely

$$\lim_{z\to\infty} W_{y_j[y_i]|\left(y_i, \frac{1}{z}y_j, S^+, \frac{1}{z}S^-\right)}(z) = 0.$$

Furthermore, the symmetrical statement

$$\lim_{z\to\infty} W_{y_i[y_j]|\left(y_j, \frac{1}{z}y_i, S^+, \frac{1}{z}S^-\right)}(z) = 0$$

holds as well.

Proof: See the Appendix F ■

Theorem III.6 leads to a corollary that allows one to detect the presence of a link by testing for strict causality.

Corollary III.7 (Link detection via strict causality test): Consider a causal, causally well-posed, topologically detectable, and recursive LDIM with output signals $y = (y_1, \ldots, y_n)^T$. Let y_i and y_j be distinct. If $\lim_{z\to\infty} W_{y_j[y_i]|(y_i, \frac{1}{z}y_j, S^+, \frac{1}{z}S^-)}(z) \neq 0$ or $\lim_{z\to\infty} W_{y_i[y_j]|(y_j, \frac{1}{z}y_i, S^+, \frac{1}{z}S^-)}(z) \neq 0$, for all disjoint S^+ and S^-, with $S^+, S^- \subseteq y \setminus \{y_i, y_j\}$, then either $H_{ji}(z)$ or $H_{ij}(z)$ have a nonzero direct feedthrough.

Thus, Corollary III.7 provides a sufficient condition for determining the presence of a direct feedthrough between two signals. The result of Theorem III.6 complements the result of Theorem III.3 and we can use both to derive an algorithm for link detection from observational time series.

IV. MD Algorithm to Recover Recursive Networks

Theorems III.3 and III.6 can be combined to derive an algorithm (that we name MD algorithm), testing, from observational data, the presence of a link connecting two distinct nodes y_i and y_j. The fact that Theorem III.3 is a necessary and sufficient condition capable of tackling scenarios with direct feedthroughs significantly strengthens the preliminary results of [41] which were instead based on the assumption that the transfer function $H_{ji}(z)$ had to be strictly causal.

Furthermore, with the exception of self-loops, MD can also infer the orientation of the double-headed arrows of the graphical representation, producing a partially-oriented graph containing information about which links have direct feedthroughs and which do not. The MD algorithm proceeds as follows. The first step is to use Theorem III.6 (or, better, Corollary III.7) to test if there is a direct feedthrough. For any possible choice of disjoint sets S^+, S^- compute the Wiener filter estimating y_j from $S = S^+ \cup \frac{1}{z}S^-$, y_i and $\frac{1}{z}y_j$. If there is a pair S^+, S^- for which the Wiener filter component associated with y_i is not strictly causal, then, we necessarily have a link between y_i and y_j in the skeleton of the perfect graphical representation. If, at the previous step, a link is not found, we proceed by using Corollary III.5, testing whether there is a set that Wiener separates $\frac{1}{z}y_i$ and y_j or a set that Wiener separates $\frac{1}{z}y_j$ and y_i. If either test fails, we conclude that there is a link between y_i and y_j in the skeleton of the perfect graphical representation. Here, we report the pseudocode of MD algorithm.

Test Link Presence(y_i, y_j)

0. Test 1: test if there exists S^+ and S^- as in Theorem III.6 that lead to W_{ji} strictly causal
1. If output of Test 1 is negative
2. The link between y_i and y_j is present in the skeleton
3. Else
4.1 Add a double-headed link from y_i to y_j
4.2 Add a double-headed link from y_j to y_i
4.3 Test 2.1: test if there exist S_c and S_s as in Theorem III.3 such that $y_j(t)$ is independent of $\frac{1}{z}y_i$ (Extended Granger causality)
5. If output of Test 2.1 is negative:
7. The double-headed link between y_i and y_j is present in the skeleton
8. Else:
9. Test 2.2: test if there exist S_c and S_s as in Theorem III.3 such that $y_i(t)$ is independent of $\frac{1}{z}y_j$ (Extended Granger causality)
10. If output of Test 2.2 is negative:
11. The double-headed link between y_i and y_j is present in the skeleton
12. Else:
13. No link between y_i and y_j is present in the skeleton

Observation 2: The statement of the algorithm in this form is mostly for simplicity of exposition. Indeed, preprocessing steps can be added to the algorithm in order to narrow down the search of separating sets. A possible preprocessing step might make use of the results in [20], where a polynomial method to infer the so-called Markov blanket of the node y_i is given. The Markov blanket of y_i consists of all the children, parents, and coparents of the node y_i in the underlying network of the linear dynamic system. As it is clear from the proofs of Theorem III.3 and Theorem III.6, once the Markov blanket of each node is computed, the search only needs to involve subsets of the union of the Markov blankets of y_i and y_j. Indeed, as the Markov blankets of the nodes y_i and y_j contain their parents, the search for the Wiener-separating set can be limited to the nodes in the union of their Markov blankets. If the network is sparse, and many naturally occurring large scale networks are, this is going to drastically reduce the number of subsets to be considered.

Also, note that Corollary III.5 and Corollary III.7 are sufficient conditions for the presence of a link in the skeleton of the perfect graphical representation. For this reason, MD algorithm gives, in the limit of infinite data, no false positives (all detected links are present in the skeleton of the perfect graphical representation), but it might be susceptible to false negatives (some links might pass undetected). However, a false negative for Corollary III.5 is not necessarily a false negative for Corollary III.7 or viceversa. Thus, the two tests strengthen each other when combined in the MD algorithm. A fundamental theoretical problem is the determination of the conditions leading to false negatives for

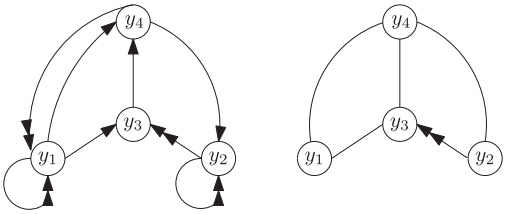

Fig. 5. (a) Perfect graphical representation of the LDIM used for the numerical comparisons when the parameters $a_{14}, a_{24}, a_{31}, a_{32}, a_{41}, a_{43}$ are all different from zero. (b) Partially directed graph that is output of the MD algorithm when applied to the LDIM with perfect graphical representation given in (a).

MD algorithm even when the data size is large. In this scenario, the occurrence of false negatives between y_i and y_j is difficult to quantify because it boils down to having a component of the Wiener filter which is strictly causal or null for at least one choice of separating sets S^+ and S^- or S_c and S_s, respectively. These conditions definitely depend on the specific transfer functions of the LDIM which directly determine the expression of the Wiener filter. In practical scenarios, this information is generally not available, thus it typically cannot be assessed.

The real fundamental issue in applications is how MD algorithm performs in the case of finite data where the Wiener filters or equivalent tools for testing conditional independence need to be computed from the observed time series. In this case, both false positives and negatives can occur because of the limited data size and the fact that the algorithm needs to be implemented via statistical tests. In next section, we numerically investigate the performance of the algorithm and compare it with other state-of-the art methods.

V. Numerical Example

The purpose of this section is to numerically compare the proposed network reconstruction method with other state-of-the-art techniques.

As a benchmark, we have considered the LDIM

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 0.4\frac{1}{z} & 0 & 0 & \frac{1}{z}a_{14} \\ 0 & 0.2\frac{1}{z} & 0 & a_{24} \\ a_{31}\left(1+\frac{1}{z}\right) & a_{32}\frac{1}{z} & 0 & 0 \\ a_{41} & 0 & a_{43} & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} + \begin{bmatrix} 0.4e_1 \\ 0.3e_2 \\ 0.4e_3 \\ 0.3e_4 \end{bmatrix}$$

where the noise components e_i, for $i = 1, 2, 3, 4$, are Gaussian independent processes with unit variance and the parameters have been chosen as $a_{14} = 0.35$, $a_{24} = -0.5$, $a_{32} = 0.45$, $a_{31} = 0.6$, $a_{43} = 0.6$, $a_{41} = 0.3$. The perfect graphical representation of this LDIM is given in Fig. 5(a).

The structure of this network has several challenging features which might complicate its identification: there are feedback loops, self-loops, and directed feedthroughs along with a so-called "V-structure" (or collider) [35], namely the presence of two nonconnected nodes (y_1 and y_2) sharing a child (y_3). In particular, the presence of such a collider makes it difficult to conclude that there is no link between y_1 and y_2 because the observation of the common child y_3 makes the two nodes conditionally correlated. For these reasons, this LDIM constitutes a simple, but extensive test bed to validate network reconstruction techniques.

In order to evaluate the performance of different techniques in terms of false positives and false negatives, we have simulated such a network for 3000 times, and for another 3000 times we have simulated the same network with exactly one of its parameters $a_{14}, a_{24}, a_{32}, a_{31}, a_{43}, a_{41}$ set to 0 with uniform probability. This is effectually equivalent to removing one of the links in the perfect graphical representation. These numerical experiments have been repeated for different simulation horizons, generating time series of 100 000, 10 000, 1000, and 500 data points. We have considered the following network reconstruction methods: a modern implementation of Granger causality [24], [42]; an alternative extension of Granger causality that takes into account direct feedthroughs [32]; a two-step method to detect lagged and contemporaneous dynamic relations [31]; a directed information method [30]; and a graphical model inspired methodology [20]. With the exception of Granger causality, all other methods have been designed to deal with direct feedthroughs.

The generated data has been used to evaluate the performance of each of the methods, asking them to detect the presence/absence of any of the links present in the skeleton of each perfect graphical representation and also the presence/absence of the link between y_1 and y_2 (absent in the skeleton of every perfect graphical representation). The comparison is performed on the skeleton of the graph because this is a structure containing common information that all these techniques are capable to infer. For example, the MD algorithm can differentiate in its output between strictly causal operators and operators with direct feedthorugh, but the other techniques in general cannot.

All of those methods rely on essentially determining the conditional dependence or independence of two variables given a conditioning set. For example, Corollary III.5 requires to determine if a certain component of the Wiener filter vanishes or not. This test does not necessarily demand the explicit computation of the Wiener filter. Instead, it can be implemented using statistical tests of conditional independence, linear regressions, or other parametric or nonparametric tools.

Since all the simulated systems are linear and Gaussian, our chosen method to test conditional independence is to apply an F-test [43]. Namely, in order to test if $\frac{1}{z}y_i$ and y_j are independent given the variables in a set $S_c \cup \frac{1}{z}S_s$, we compute the mean squared error ϵ_1 obtained when linearly estimating y_j using $S_c \cup \frac{1}{z}S_s$ and the mean squared error ϵ_2 obtained when linearly estimating y_j using $S_c \cup \frac{1}{z}S_s$ and $\frac{1}{z}y_i$. Obviously $\epsilon_2 \leq \epsilon_1$. However, if the introduction of $\frac{1}{z}y_i$ as an additional variable to explain y_j does not significantly reduce the mean squared error, we can conclude that $\frac{1}{z}y_i$ and y_j are independent given $S_c \cup \frac{1}{z}S_s$. In practice this translates in checking if $1 - \frac{\epsilon_2}{\epsilon_1} < \theta$ where θ is an appropriately chosen threshold (leading to an approximate F-test [43]). Similarly, the first test of MD algorithm (test for strict causality) can be interpreted

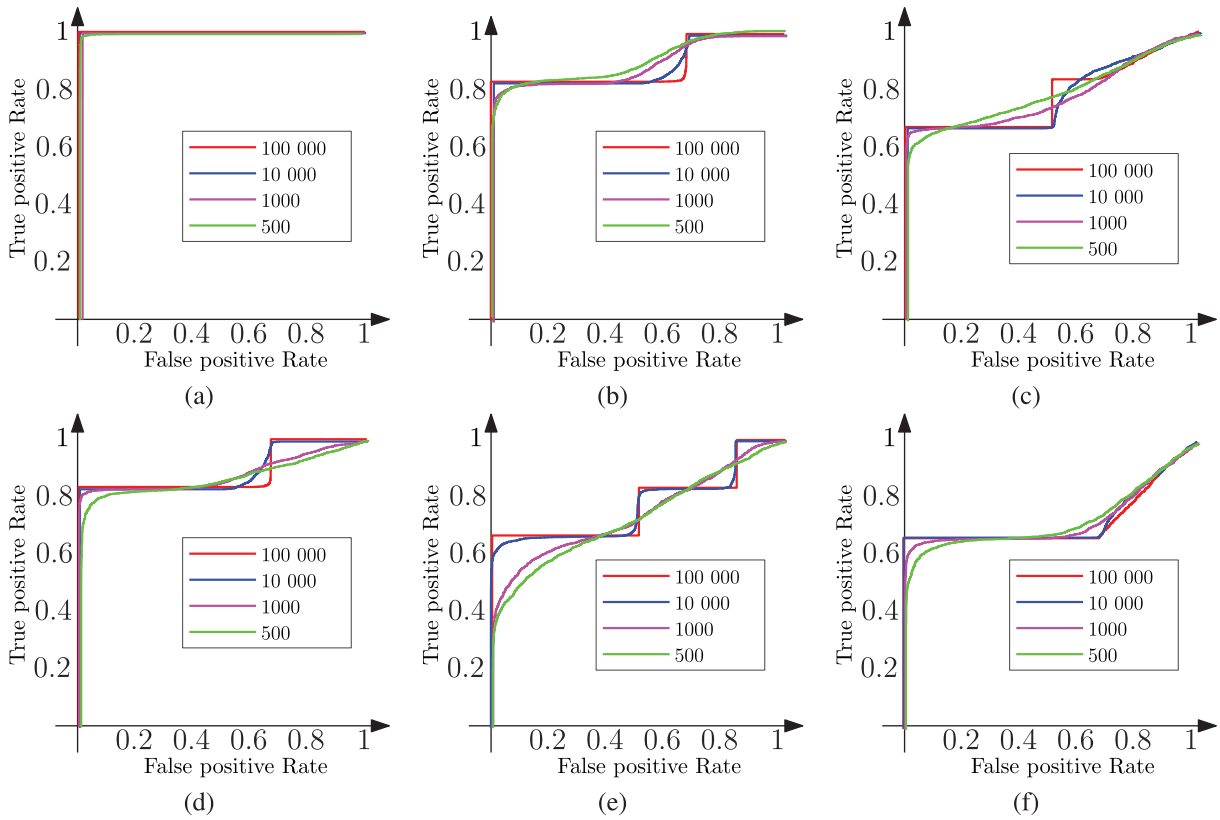

Fig. 6. Red is ROC on 100000 data points; blue is ROC on 10000 data points; magenta in ROC on 1000 data points; green is ROC on 500 data points. (a) ROC curve of our method. (b) ROC for Granger causality method. (c) ROC for the method in [30]. (d) ROC for the method in [31]. (e) ROC curve for the method in [32]. (f) ROC for the method in [20].

as a conditional independence statement: the Wiener filter estimating y_j from $S^+ \cup \frac{1}{z}S^- \cup \{y_i, \frac{1}{z}y_j\}$ has a strictly causal component for y_i if and only if the random variables $y_i(t)$ and $y_j(t)$ are independent given the variables in $S^+ \cup \frac{1}{z}S^- \cup \{\frac{1}{z}y_i, \frac{1}{z}y_j\}$. This last condition can again be translated into an F-test [43].

A common way to compare detection techniques is to plot their results on a receiver operating characteristic (ROC) curve [44]. Given a detection method defined by a threshold parameter, its ROC curve is the plot of the true positive rate versus the false positive rate for all possible values of the threshold. In our simulations, a true positive occurs when a link that is present in the skeleton of the perfect graphical representation of the input LDIM is inferred as an existing link by the methods. Similarly, a false positive occurs when a method infers the existence of a link in the skeleton of the network when such a link does not exist. A typical metric adopted to quantify the performance of a detection method is given by the area below the ROC curve. An ideal detection technique has such an area equal to one because it leads to perfect detection results for appropriate values of the threshold. The ROC curve for our identification method is depicted in Fig. 6(a) showing an underlying area of 1 for the horizons of 100 000 and 10 000 data points, about 0.9999 for the horizon of 1000 data points and about 0.9988 for the horizon of 500 data points. Furthermore, for the 3000 simulations with 100 000 data points in which all the parameters are different from zero [namely the perfect graphical representation is the one shown in Fig. 5(a)], we inspected the actual output of the MD algorithm. In all cases, the output was the partially directed graph shown in Fig. 5(b) with an oriented edge from y_2 to y_3. Given that the perfect graphical representation of the LDIM was the one in Fig. 5(a), what was found in the totality of the cases is the theoretical output of MD in the limit of infinite data. This additional numerical verification further validates the asymptotical consistency properties of the MD algorithm.

The ROC curve for Granger causality is reported in Fig. 6(b). As the plot clearly shows, this detection method is not ideal. This is to be expected since Granger causality has deficiencies when reconstructing a network with direct feedthroughs components which are present in the benchmark example.

The ROC curves for the detection methods in [30] and in [31] are reported in Fig. 6(c) and (d), respectively. Observe that these methods do not achieve a perfect identification even when the size of the data set becomes large.

The ROC curves for the detection methods in [32] and in [20] are reported in Fig. 6(e) and (f), respectively. Again, these methods do not approach the ideal curve even for large data sets.

From these numerical experiments, we notice that MD provides a better performance than the other methods even for small data sizes. Furthermore, in the limit of large data sizes MD achieves a perfect reconstruction of the skeleton, even though its theoretical guarantees are only about the absence of false positives. Other methods do not achieve a perfect reconstruction

of the skeleton even for large data sets and do not guarantee the absence of false positives.

VI. Conclusion

The main contribution of this article is the MD algorithm that is an extension of Granger causality for the reconstruction of the topology of a network. MD algorithm consists of two edge detection criteria that can be implemented via conditional independence tests. The first test checks if there is a relationship with direct feedthrough between two processes. The second test checks for the presence of a link associated with a strictly causal transfer function. Granger causality has well-known deficiencies when applied to the reconstruction of dynamic networks with direct feedthrough components. The proposed method overcomes most of these limitations. In particular, if the network of dynamic systems has at least one delay in each directed loop, the method provides no false positives in the limit of infinite data. We provide numerical simulations showing a favorable comparison of the MD algorithm with several other state-of-the art techniques. Future work could consider sample complexity analysis for the application of the MD algorithm to dynamically changing networks/topologies; and the implementation of distributed strategies on specific communication schemes.

APPENDIX

A. Proof of Proposition II.1(a)

Consider $q \in \mathcal{FE}$ (or $q \in \mathcal{F}^+\mathcal{E}$). By contradiction, suppose q has two distinct representations, namely

$$q = H_{11}(z)y_1 + \cdots + H_{1n}(z)y_n$$
$$q = H_{21}(z)y_1 + \ldots + H_{2n}(z)y_n.$$

Denote $H_i(z) = H_{1i}(z) - H_{2i}(z)$ for $i = 1, \ldots, n$ and $H(z) = (H_1(z), \ldots, H_n(z))$. Then, we also have a nontrivial representation for the null stochastic process (a stochastic process that is 0 at every time almost surely)

$$0 = \sum_{i=1}^{n} H_i(z)y_i = H(z)y$$

with at least one $H_i \neq 0$, where $y = (y_1, \ldots, y_n)^t$. Then, for the power spectral density of 0 we have

$$0 = H\Phi_{yy}H^*.$$

Thus, we have a unique representation for q if and only if $\Phi_{yy}(e^{i\omega})$ is positive definite for almost all $\omega \in [-\pi, \pi]$.

B. Proof of Proposition II.1(b)

For $q \in ctfspan(y_1, \ldots, y_n, v_1, \ldots, v_m)$, we have

$$q = \sum_{i=1}^{n} H_i^{(y)}(z)y_i + \sum_{i=1}^{m} H_i^{(v)}(z)v_i. \tag{5}$$

Also, for $i = 1, \ldots, m$ we have

$$v_i = \sum_{j=1}^{\ell} K_{ij}(z)w_j. \tag{6}$$

By substituting (6) in (5) we get the assertion.

C. Proof of Proposition II.1(c)

Let $H_k(z) \in \mathcal{F}^+$ for $k = 1, .., n$. The inclusion $ctfspan(q, y_2, \ldots, y_n) \subseteq ctfspan(y_1, y_2, \ldots, y_n)$ is trivial. Let us show that $ctfspan(y_1, y_2, \ldots, y_n) \subseteq ctfspan(q, y_2, \ldots, y_n)$. From $q = \sum_{k=1}^{n} H_k y_k$, $H_1 \neq 0$ we have that $H_1(z)y_1 = q - \sum_{k=2}^{n} H_k y_k$. Clearly, $y_1 = H_1(z)^{-1}(q - \sum_{k=2}^{n} H_k y_k)$ and $y_1 \in \mathcal{F}^+\{q, y_2, \ldots, y_n\}$ if and only if $H_1(z)^{-1} \in \mathcal{F}^+$. The proof is analogous for $H_k(z) \in \mathcal{F}$ for $k = 1, .., n$.

D. Proof of Lemma II.3

Denote

$$\hat{q}_0 = \arg\min_{q \in ctfspan(Y)} ||v - q||^2$$

and

$$\hat{q}_1 = \arg\min_{q \in ctfspan(Y)} ||(v - w) - q||^2.$$

By Proposition II.2, both $\hat{q}_0$ and $\hat{q}_1$ exist and are unique. We want to show that $\hat{q}_0 = w + \hat{q}_1$. As $\mathcal{F}^+\mathcal{E}$ is a pre-Hilbert space with the inner product $<, >$, we have that $\hat{q}_0$ and $\hat{q}_1$ are respectively the only elements in Y that satisfy

$$< v - \hat{q}_0, q >= 0$$
$$< (v - w) - \hat{q}_1, q >= 0$$

for all $q \in Y$. Observe that, for all $q \in Y$

$$< v - (w + \hat{q}_1), q >=< (v - w) - \hat{q}_1, q >= 0.$$

Thus, we necessarily have that $\hat{q}_0 = w + \hat{q}_1$.

E. Proof of Proposition II.4

Let $v \perp ctfspan(y_k)$ for all $k = 1, \ldots, n$. Then, it is true that $\langle v, H_k(z)y_k \rangle = 0$, for any $H_k(z) \in \mathcal{F}^+$. As that holds for all $k = 1, \ldots, n$ we have that

$$\left\langle v, \sum_{k=1}^{n} H_k(z)y_k \right\rangle = \sum_{k=1}^{n} \langle v, H_k(z)y_k \rangle = 0$$

for any $H_k(z) \in \mathcal{F}^+$. Thus, $\langle v - 0, q \rangle = 0$, for $q \in ctfspan(y_1, \ldots, y_n)$. Denote $y = (y_1, \ldots, y_n)$. By Proposition (II.2), $\hat{v}$, the Wiener Filter estimate of v from $q \in ctfspan(y_1, \ldots, y_n)$, is zero, which implies W_{ve} is zero. Thus, $v \perp ctfspan(y_1, \ldots, y_n)$.

Now let $v \perp ctfspan(y_1, \ldots, y_n)$. Then, $0 = \langle e - 0, q \rangle$ for any $q \in ctfspan(y_1, \ldots, y_n)$, $q = \sum_{k=1}^{n} H_k y_k$. Thus, for any $H_k(z) \in \mathcal{F}^+$ we have that $< v - 0, H_k(z)y_k >= 0$ ($y_k \neq v$), implying that $W_{vy_k} = 0$. Similarly, $W_{vy_k} = 0$ for $k = 1, \ldots, n$. Thus, $v \perp ctfspan(y_k)$ for $k = 1, \ldots, n$.

F. Proof of Theorem III.6

Choose a graph G that is a recursive representation of the LDIM. Given the graph G, partition the set of nodes $\{y_1, y_2, \ldots y_n\} \setminus \{y_i, y_j\}$ in the following subsets.

1) S^-_{ji}: The set of strictly causal parents for both y_j and y_i.
2) S^+_{ji}: The set of parents for both y_j and y_i that are not in S^-_{ji}.
3) $S^+_{j\bar{i}}$: The set of causal but not strictly causal parents of y_j, which are not parents of y_i.
4) $S^-_{j\bar{i}}$: The set of strictly causal parents of y_j, which are not parents of y_i.
5) $S^+_{i\bar{j}}$: The set of causal but not strictly causal parents of y_i, which are not parents of y_j.
6) $S^-_{i\bar{j}}$: The set strictly causal parents of y_i, which are not parents of y_j.
7) $S_{\bar{i}\bar{j}}$: All other nodes.

Given the above partition of the nodes, define

1) $S^+ := S^+_{i\bar{j}} \cup S^+_{j\bar{i}} \cup S^+_{ji}$ and
2) $S^- := S^-_{i\bar{j}} \cup S^-_{j\bar{i}} \cup S^-_{ji}$.

Let us compute the following MLSE (minimum least square estimate) for y_j

$$\hat{y}_j = \arg\min_q \|y_j - q\| \text{ , subject to } q \in ctfspan \times \left(y_i\,, \frac{1}{z}y_j, S^+, \frac{1}{z}S^-\right).$$

Observe that

1) $H_{jj}(z)$ is strictly causal since G is recursive
2) $H_{jk}(z)$ is strictly causal for $y_k \in S^-$
3) $H_{jk}(z) = 0$ for $y_k \in S^-_{i\bar{j}} \cup S^+_{i\bar{j}}$.

From these observations we can apply Lemma II.3 and get

$$\hat{y}_j = H_{ji}(z)y_i + H_{jj}(z)y_j + \sum_{y_k \in S^+ \cup S^-} H_{jk}(z)y_k + \arg\min_{q \in ctfspan(y_i, \frac{1}{z}y_j, S^+, \frac{1}{z}S^-)} \|e_j - q\|.$$

Let $\epsilon_j = e_j - \hat{e}_j$, where e_j is the independent component of y_j and $\hat{e}_j = F_j(z)\frac{1}{z}e_j$ is the one step ahead predictor for e_j from the past of e_j (see [45]).

From Proposition II.1(b), we have that

$$\hat{e}_j \in ctfspan\left(y_i\,, \frac{1}{z}y_j, S^+, \frac{1}{z}S^-\right).$$

Thus, applying again Lemma II.3, we have that

$$\hat{y}_j = H_{ji}(z)y_i + H_{jj}(z)y_j + \sum_{y_k \in S^+ \cup S^-} H_{jk}(z)\frac{1}{z}y_k + \hat{e}_j + \arg\min_q \|\epsilon_j - q\|$$

subject to

$$q \in ctfspan\left(y_i\,, \frac{1}{z}y_j, S^+, \frac{1}{z}S^-\right).$$

Similarly, let $\epsilon_i = e_i - \hat{e}_i$, where e_i is the independent component of y_i and $\hat{e}_i = F_i(z)\frac{1}{z}e_i$ is the one step ahead predictor for e_i from the past of e_i.

From Proposition II.1(c), we have that

$$ctfspan\left(y_i\,, \frac{1}{z}y_j\,, S^+, \frac{1}{z}S^-\right) = ctfspan\left(e_i\,, \frac{1}{z}y_i\,, \frac{1}{z}y_j\,, S^+, \frac{1}{z}S^-\right) = ctfspan\left(\epsilon_i\,, \frac{1}{z}y_i\,, \frac{1}{z}y_j\,, S^+, \frac{1}{z}S^-\right).$$

Thus, we can write

$$\hat{y}_j = H_{ji}(z)y_i + H_{jj}(z)y_j + \sum_{y_k \in S^+ \cup S^-} H_{jk}(z)y_k + \hat{e}_j + \arg\min_q \|\epsilon_j - q\|$$

subject to

$$q \in ctfspan\left(\epsilon_i\,, \frac{1}{z}y_i\,, \frac{1}{z}y_j\,, S^+, \frac{1}{z}S^-\right).$$

For all y_k in $S^+_{ji} \cup S^+_{j\bar{i}} \cup S^+_{i\bar{j}}$, define $y^\perp_k = y_k - \hat{y}_k$, where $\hat{y}_k$ is the following one-step ahead predictor

$$\hat{y}_k = \arg\min_q \|y_k - q\|$$

subject to

$$q \in ctfspan \times \left(\frac{1}{z}S^+_{ji}\,, \frac{1}{z}S^+_{j\bar{i}}\,, \frac{1}{z}S^+_{i\bar{j}}\,, \frac{1}{z}y_i\,, \frac{1}{z}y_j\,, \frac{1}{z}S^-_{ji}\,, \frac{1}{z}S^-_{j\bar{i}}\,, \frac{1}{z}S^-_{i\bar{j}}\right).$$

Define the set $S^\perp$ to be the collection of $y^\perp_k$. By Proposition II.1(c)

$$ctfspan\left(S^+_{ji}\,, S^+_{j\bar{i}}\,, S^+_{i\bar{j}}\,, \frac{1}{z}y_i\,, \frac{1}{z}y_j\,, \frac{1}{z}S^-_{ji}\,, \frac{1}{z}S^-_{j\bar{i}}\,, \frac{1}{z}S^-_{i\bar{j}}\right) = ctfspan\left(S^\perp\,, \frac{1}{z}y_i\,, \frac{1}{z}y_j\,, \frac{1}{z}S^+, \frac{1}{z}S^-\right).$$

Thus, we get

$$\hat{y}_j = H_{ji}(z)y_i + H_{jj}(z)y_j + \sum_{y_k \in S^+ \cup S^-} +\hat{e}_j + \arg\min_q \|\epsilon_j - q\|$$

subject to

$$q \in ctfspan\left(\epsilon_i\,, S^\perp\,, \frac{1}{z}y_i\,, \frac{1}{z}y_j\,, \frac{1}{z}S^+, \frac{1}{z}S^-\right).$$

Now, we have the following observations.

1) The innovation ϵ_j is orthogonal to $ctfspan(\frac{1}{z}S^+, \frac{1}{z}y_i\,, \frac{1}{z}y_j\,, \frac{1}{z}S^-)$ as $\epsilon_j \perp ctfspan(\frac{1}{z}e_j)$, and by the assumption that Φ_{ee} is diagonal, we have that $\frac{1}{z}e_j \perp ctfspan(\frac{1}{z}e_k)$, for all $k \neq j$;
2) Furthermore, $ctfspan(S^\perp, \epsilon_i)$ is orthogonal to $ctfspan(\frac{1}{z}S^+, \frac{1}{z}y_i\,, \frac{1}{z}y_j\,, \frac{1}{z}S^-)$ by construction.

Thus, we obtain

$$\hat{y}_j = H_{ji}(z)y_i + H_{jj}(z)y_j + \sum_{y_k \in S^+ \cup S^-} H_{jk}(z)y_k + \hat{e}_j + \arg\min_q \|\epsilon_j - q\|$$

subject to $q \in ctfspan(\epsilon_i, S^{\perp})$. Note that if we switched the roles of y_i and y_j, following analogous steps, we would get

$$\hat{y}_i = H_{ij}(z)y_j + H_{ii}(z)y_i + \sum_{y_k \in S^+ \cup S^-} H_{ik}(z)y_k + \hat{e}_i + \arg\min_q \|\epsilon_i - q\|$$

subject to $q \in ctfspan(\epsilon_j, S^{\perp})$. Now, at least one of these two statements holds

1) y_i is not a descendant of y_j in the graph of instantaneous propagations associated with G
2) y_j is not a descendant of y_i in the graph of instantaneous propagations associated with G

otherwise G would not be recursive. Without any loss of generality assume that y_j is not a descendant of y_i in the graph of instantaneous propagations, otherwise we can flip the roles of y_i and y_j in the following arguments.

We have that y_i is not an ancestor of any $y_k \in S^+ = S^+_{ji} \cup S^+_{j\bar{i}} \cup S^+_{i\bar{j}}$ in the graph of instantaneous propagations.

1) If $y_k \in S^+_{ji} \cup S^+_{j\bar{i}}$ there is an edge from y_k to y_j in the graph of instantaneous propagations. This means that there is no chain from y_i to y_k in the graph of instantaneous propagations, otherwise y_j would be a descendant of y_i in the graph of instantaneous propagations.
2) if $y_k \in S^+_{i\bar{j}}$ there is an edge from y_k to y_i in the graph of instantaneous propagations. This means that there is no chain from y_i to y_k in the graph of instantaneous propagations, otherwise we would have a loop in the graph of instantaneous propagations.

From Proposition III.1, we have that

$$y_k \in S^+ \Rightarrow y_k \in ctfspan\left(e_{A_k}, \frac{1}{z}e_{\overline{A}_k}\right)$$

where A_k is the set of indices of nodes that are ancestors of y_k in the graph of instantaneous propagations and $\overline{A}_k$ are all the other indices. Observe that $i \notin A_k$ for every $y_k \in S^+$. As $y_k^{\perp} \in ctfspan(y_k, \hat{y}_k)$, we also have that $y_k^{\perp} \in ctfspan(e_{A_k}, \frac{1}{z}e_{\overline{A}_k})$. From the previous observation we know that $e_i \in e_{\overline{A}_k}$ for every $y_k \in S^+$. Observe that $\epsilon_i = e_i - F_i\frac{1}{z}e_i$, thus, we have $\epsilon_i \perp ctfspan(e_\ell)$ for all $\ell \neq i$. Also $\epsilon_i \perp \frac{1}{z}e_i$. Thus, $\epsilon_i \perp ctfspan(y_k^{\perp})$ for all $y_k^{\perp} \in S^{\perp}$ and by using Proposition II.4, we have that $\epsilon_i \perp ctfspan(S^{\perp})$. It is also straightforward to conclude that $\epsilon_i \perp ctfspan(\epsilon_j)$, thus, we get

$$\hat{y}_j = H_{ji}(z)y_i + H_{jj}(z)y_j + \sum_{y_k \in S^+ \cup S^-} H_{jk}(z)y_k + \hat{e}_j + \arg\min_q \|\epsilon_j - q\|$$

subject to $q \in ctfspan(S^{\perp})$. For the expression of $\hat{y}_i$, we instead get

$$\hat{y}_i = H_{ij}(z)y_j + H_{ii}(z)y_i + \sum_{y_k \in S^+ \cup S^-} H_{ik}(z)y_k + \hat{e}_i.$$

Thus, we have that

$$\lim_{z\to\infty} W_{y_i[y_j]|\left(S^+, \frac{1}{z}S^-\right)}(z) = 0$$

which constitutes part of the theorem thesis.

The other part of the theorem thesis is obtained by manipulating the expression for $\hat{y}_j$.

$$\begin{aligned}\hat{y}_j &= H_{ji}(z)y_i + H_{jj}(z)y_j + \sum_{y_k \in S^+ \cup S^-} H_{jk}(z)y_k + \hat{e}_j \\ &\quad + \sum_{y_k^{\perp} \in S^{\perp}} W_{\epsilon_j, [y_k^{\perp}]|S^{\perp}}(z)y_k^{\perp} \\ &= H_{ji}(z)y_i + H_{jj}(z)y_j + \sum_{y_k \in S^+ \cup S^-} H_{jk}(z)y_k + \hat{e}_j \\ &\quad + \sum_{y_k \in S^+} W_{\epsilon_j, [y_k^{\perp}]|S^{\perp}}(z)y_k - \sum_{y_k \in S^+} W_{\epsilon_j, [y_k^{\perp}]|S^{\perp}}(z)\hat{y}_k\end{aligned}$$

where $\hat{y}_k$, the estimate of $y_k \in S^+$ from $\tilde{S} = S^+ \cup S^- \cup \{y_i, y_j\}$

$$\hat{y}_k = \sum_{y_\ell \in \tilde{S}} W_{y_k, [\frac{1}{z}y_\ell]|\frac{1}{z}\tilde{S}}(z)\frac{1}{z}y_\ell.$$

From the expression of $\hat{y}_j$, it is now immediate to verify that for $z \to \infty$ the transfer function operating on y_i is strictly causal giving the assertion.

REFERENCES

[1] D. Del Vecchio and R. M. Murray, *Biomolecular Feedback Systems*. Princeton, NJ, USA: Princeton Univ. Press, 2015.
[2] J. R. Pomerening, "Positive-feedback loops in cell cycle progression," *FEBS Lett.*, vol. 583, no. 21, pp. 3388–3396, 2009.
[3] M. Kamiński, M. Ding, W. A. Truccolo, and S. L. Bressler, "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance," *Biol. Cybern.*, vol. 85, no. 2, pp. 145–157, 2001.
[4] M. J. Naylor, L. C. Rose, and B. J. Moyle, "Topology of foreign exchange markets using hierarchical structure methods," *Phys. A, Statist. Mech. Appl.*, vol. 382, no. 1, pp. 199–208, 2007.
[5] J. R. Freeman, "Granger causality and the times series analysis of political relationships," *Amer. J. Political Sci.*, vol. 27, pp. 327–358, 1983.
[6] M. D. Ward, K. Stovel, and A. Sacks, "Network analysis and political science," *Annu. Rev. Political Sci.*, vol. 14, pp. 245–264, 2011.
[7] J. Scott, *Social Network Analysis*. Newbury Park, CA, USA: Sage, 2012.
[8] D. Hayden, Y. H. Chang, J. Goncalves, and C. J. Tomlin, "Sparse network identifiability via compressed sensing," *Automatica*, vol. 68, pp. 9–17, 2016.
[9] V. Chetty, D. Hayden, J. Goncalves, and S. Warnick, "Robust signal-structure reconstruction," in *Proc. IEEE 52nd Annu. Conf. Decis. Control*, 2013, pp. 3184–3189.
[10] S. Prabakaran, J. Gunawardena, and E. Sontag, "Paradoxical results in perturbation-based signaling network reconstruction," *Biophys. J.*, vol. 106, no. 12, pp. 2720–2728, 2014.
[11] M. Nabi-Abdolyousefi and M. Mesbahi, "Network identification via node knockout," *IEEE Trans. Autom. Control*, vol. 57, no. 12, pp. 3214–3219, Dec. 2012.
[12] S. Shahrampour and V. M. Preciado, "Topology identification of directed dynamical networks via power spectral analysis," *IEEE Trans. Autom. Control*, vol. 60, no. 8, pp. 2260–2265, Aug. 2015.
[13] F. Morbidi and A. Y. Kibangou, "A distributed solution to the network reconstruction problem," *Syst. Control Lett.*, vol. 70, pp. 85–91, 2014.
[14] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
[15] R. Mantegna and H. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[16] E. Kolaczyk, *Statistical Analysis of Network Data: Methods and Models*, (Springer Series in Statistics). New York, NY, USA: Springer, 2009.
[17] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Modern Phys.*, vol. 74, no. 1, 2002, Art. no. 47.

[18] H. H. Weerts, A. G. Dankers, and P. M. Van den Hof, "Identifiability in dynamic network identification," *IFAC-PapersOnLine*, vol. 48, no. 28, pp. 1409–1414, 2015.
[19] J. Gonçalves and S. Warnick, "Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks," *IEEE Trans. Autom. Control*, vol. 53, no. 7, pp. 1670–1674, Aug. 2008.
[20] D. Materassi and M. V. Salapaka, "On the problem of reconstructing an unknown topology via locality properties of the wiener filter," *IEEE Trans. Autom. Control*, vol. 57, no. 7, pp. 1765–1777, Jul. 2012.
[21] H. H. Weerts, P. M. Van den Hof, and A. G. Dankers, "Identifiability of linear dynamic networks," *Automatica*, vol. 89, pp. 247–258, 2018.
[22] V. Chetty, J. Eliason, and S. Warnick, "Passive reconstruction of non-target-specific discrete-time LTI systems," in *Proc. Amer. Control Conf.*, 2016, pp. 66–71.
[23] Z. Yue, J. Thunberg, W. Pan, L. Ljung, and J. Gonçalves, "Linear dynamic network reconstruction from heterogeneous datasets," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 10 586–10 591, 2017.
[24] C. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, pp. 424–438, 1969.
[25] C. A. Sims, "Macroeconomics and reality," *Econometrica: J. Econometric Soc.*, vol. 48, pp. 1–48, 1980.
[26] M. Gong, K. Zhang, B. Schölkopf, C. Glymour, and D. Tao, "Causal discovery from temporally aggregated time series," in *Proc. Conf. Uncertainty Artif. Intell.*, 2017, Art. no. 269.
[27] D. Hayden, Y. Yuan, and J. Gonçalves, "Network identifiability from intrinsic noise," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 3717–3728, Aug. 2017.
[28] D. Materassi and M. V. Salapaka, "Reconstruction of directed acyclic networks of dynamical systems," in *Proc. Amer. Control Conf.*, 2013, pp. 4687–4692.
[29] F. Sepehr and D. Materassi, "Inferring the structure of polytree networks of dynamic systems with hidden nodes," in *Proc. IEEE 55th Conf. Decis. Control*, 2016, pp. 4618–4623.
[30] C. J. Quinn, T. P. Coleman, N. Kiyavash, and N. G. Hatsopoulos, "Estimating the directed information to infer causal relationships in ensemble neural spike train recordings," *J. Comput. Neuroscience*, vol. 30, no. 1, pp. 17–44, 2011.
[31] J. Runge *et al.*, "Identifying causal gateways and mediators in complex spatio-temporal systems," *Nature Commun.*, vol. 6, 2015, Art. no. 8502.
[32] L. Schiatti, G. Nollo, G. Rossato, and L. Faes, "Extended Granger causality: A new tool to identify the structure of physiological networks," *Physiological Meas.*, vol. 36, no. 4, pp. 827–843, 2015.
[33] D. Materassi and M. V. Salapaka, "Signal selection for estimation and identification in networks of dynamic systems: A graphical model approach," *IEEE Trans. Autom. Control*, to be published, doi: 10.1109/TAC.2019.2960001.
[34] R. Diestel, *Graph Theory*. Berlin, Germany: Springer-Verlag, 2012.
[35] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA, USA: Morgan Kaufmann, 1988.
[36] T. Shafie, "A multigraph approachto social network analysis," *J. Social Struct.*, vol. 16, pp. 1–21, 2015.
[37] A. V. Oppenheim, A. S. Willsky, and S. Nawab, *Signals and Systems*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1997.
[38] D. Materassi and M. V. Salapaka, "Graphoid-based methodologies in modeling, analysis, identification and control of networks of dynamic systems," in *Proc. Amer. Control Conf.*, 2016, pp. 4661–4675.
[39] L. Barnett and A. K. Seth, "Granger causality for state-space models," *Phys. Rev. E*, vol. 91, no. 4, 2015, Art. no. 040101.
[40] D. Materassi and M. Salapaka, "On the problem of reconstructing an unknown topology via locality properties of the Wiener filter," *IEEE Trans. Autom. Control*, vol. 57, no. 7, pp. 1765–1777, Jul. 2012.
[41] M. Dimovska and D. Materassi, "Granger-causality meets causal inference in graphical models: Learning networks via non-invasive observations," in *Proc. IEEE 56th Annu. Conf. Decis. Control*, 2017, pp. 5268–5273.
[42] D. Marinazzo, M. Pellicoro, and S. Stramaglia, "Kernel method for nonlinear Granger causality," *Phys. Rev. Lett.*, vol. 100, 2008, Art. no. 144103.
[43] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*. Berlin, Germany: Springer, 2006.
[44] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve." *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.
[45] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, 2000.

Mihaela Dimovska received bachelor's degree in computer science and mathematics from the American University in Bulgaria, Blagoevgrad, Bulgaria, and the master's degree in computer science from the University of Tennessee, Knoxville, TN, USA. She is currently working toward the Ph.D. degree with the University of Minnesota, Minneapolis, MN, USA.

She was a Research Assistant with the University of Tennessee, from 2016 to 2018. In 2019, she was a Research Fellow with the Advanced Short Term Research Opportunity, Oak Ridge National Laboratory. She is currently a Research Assistant with the University of Minnesota. Her research interests include system identification, graphical models, machine learning, and stochastic modeling.

Donatello Materassi received a Laurea degree in ingegneria informatica and a Dottorato di Ricerca degree in electrical engineering/nonlinear dynamics and complex systems from the Università degli Studi di Firenze, Florence, Italy.

He has been a Research Associate with the University of Minnesota (Twin Cities), from 2008 to 2011. He has been concurrently both a Postdoctoral Researcher with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology and a Lecturer with Harvard University till 2014. From 2014 till 2019 he has been an Assistant Professor with the University of Tennessee, Knoxville, TN, USA and since 2019 he is an Assistant Professor at the University of Minnesota, Twin Cities. 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.

Dr. Materassi was the recipient of the NSF CAREER Award in 2015.