

Modeling and Topology Estimation of Low Rank Dynamical Networks[★]

Wenqi Cao^{*} Aming Li^{**}

^{*} Center for Systems and Control, School of Advanced Manufacturing and Robotics, Peking University, Beijing, China (e-mail: wenqicao@pku.edu.cn).

^{**} Center for Systems and Control, School of Advanced Manufacturing and Robotics, and Center for Multi-Agent Research, Institute for Artificial Intelligence, Peking University, Beijing, China (e-mail: amingli@pku.edu.cn).

Abstract: Conventional topology learning methods for dynamical networks become inapplicable to processes exhibiting low-rank characteristics. To address this, we propose the low rank dynamical network model which ensures identifiability. By employing causal Wiener filtering, we establish a necessary and sufficient condition that links the sparsity pattern of the filter to conditional Granger causality. Building on this theoretical result, we develop a consistent method for estimating all network edges. Simulation results demonstrate the parsimony of the proposed framework and consistency of the topology estimation approach.

Keywords: System identification, network topology, low rank stochastic processes, estimation and filtering, graphical models, Granger causality.

1. INTRODUCTION

Graphical models and dynamical networks provide a powerful framework for characterizing interaction patterns in multivariate systems, with broad applications across fields such as economics, ecological evolution and sociology. As system dimensionality increases, the corresponding graph topology often exhibits sparsity, implying each unit interacts directly with only a limited number of neighbors. This structural property not only provides critical functional insights but also renders the estimation of graph topology from data essential to systems identification and control. See, for example, Materassi and Salapaka (2012), Avventi et al. (2013), Dimovska and Materassi (2021), Pillonetto et al. (2025) and Zorzi (2025).

In the topology learning and parameter identification of dynamical networks, the invertibility of the spectral density is often a sufficient condition for a well-posed identification problem and a consistent estimate. In addition, an inverse spectral condition is commonly employed to determine the edges between nodes, i.e.,

$$[\Phi(e^{i\theta})^{-1}]_{kh} = 0 \Leftrightarrow \mathbf{Y}_{\{k\}} \perp \mathbf{Y}_{\{h\}} \mid \mathbf{Y}_{V \setminus \{k,h\}}, \quad (1)$$

meaning that the space $\mathbf{Y}_{\{k\}}$ and $\mathbf{Y}_{\{h\}}$ are conditionally independent given $\mathbf{Y}_{V \setminus \{k,h\}}$, where $\Phi(e^{i\theta})$ denotes the spectral density matrix of a stochastic vector process $y(t)$,

$$\mathbf{Y}_S := \text{span} \{y_{(i)}(t) : i \in S, t \in \mathbb{Z}\},$$

and S is an index set contained in the index set of all entries of $y(t)$.

However, as the complexity and dimensionality of the system increase, the underlying vector stochastic process exhibits a low-rank property, leading to a rank-deficient spectral density matrix (Vincent et al. (2024)). Such processes, referred to as low rank processes, have been studied in recent works like Cao et al. (2023a), Cao et al. (2023b) and Yu and Qin (2023). When the vector process underlying a graph or network is low-rank, the rank-deficient spectrum violates the well-posedness conditions assumed in existing methods, such as those in Materassi and Salapaka (2012) and Dimovska and Materassi (2021). Moreover, the inverse spectral condition (1) can no longer be applied to infer edges between nodes.

This paper addresses the lack of a general method for topology recovery of causal low-rank processes through three main contributions. First, we introduce a low rank dynamical network (LRDN) model that is computationally efficient and, crucially, ensures identifiability by guaranteeing a unique graphical representation. Second, we develop a corresponding topology learning approach based on causal Wiener filtering which is theoretically proven to be consistent. Third, we establish a fundamental one-one correspondence between conditional Granger causality and the entries of the causal Wiener filter (or the inverse spectral factor, as shown in (22)(27)), thereby resolving the failure of the conventional inverse spectral condition in (1). Notably, our method imposes no requirement for strictly causal edges in every chain—a key limitation of existing full-rank methods—thus generalizing consistent topology learning to a broader network class. Section 2 introduces preliminaries and the LRDN model.

[★] This work is supported by the National Natural Science Foundation of China under Grant Nos. T2525017, 62533002, T2421004 and 62503020, the National Key Research and Development Program of China under Grant No. 2022YFA1008400, the Jiangsu Provincial Scientific Research Center of Applied Mathematics under Grant No. BK20233002, the Beijing Natural Science Foundation under Grant No. 4254098, and the Postdoctoral Fellowship Program of CPSF under Grant No. GZC20240041.

Section 3 derives the causal Wiener filter. In Section 4, the relation between conditional Granger causality and Wiener filter is discussed, and then the topology estimation approach is proposed through establishing a sufficient and necessary condition for the existence of edges. An example and the conclusions are given in Sections 5 and 6.

2. LOW RANK GRAPHICAL MODEL

Consider an $(m + l)$ -dimensional low rank process,

$$y(t) = [y_{(1)}(t), y_{(2)}(t), \dots, y_{(m+l)}(t)]',$$

with spectral density matrix $\Phi(z)$ satisfying $\text{rank}(\Phi) = l$. Since $\Phi(z)$ is not invertible, the conditional dependencies among nodes cannot be directly recovered using the conventional inverse spectral condition (1). This paper thus addresses the problem of constructing a graphical structure for low-rank processes, where edges reflect inter-nodal relationships and the topology is identifiable.

We restrict our attention to causal processes and causal interdependencies, which underpin physically realizable identification and control schemes. We further assume no exogenous inputs, so that the graph must be identified solely from non-invasive observations of its internal dynamics. In the following, we review preliminaries on linear dynamical networks and low rank processes, and introduce the formal definition of a low rank dynamical network.

2.1 Linear Causal Dynamical Network

We recall the following foundational definition.

Definition 1. (Directed and undirected graphs). A directed (undirected) graph is a pair (V, E) , where V is the index set of nodes, and E is a set of edges, defined as ordered (unordered) pairs of elements in V .

Denote by $[M]_{ij}$ the (i, j) -th entry of a matrix. Building on the framework in Dimovska and Materassi (2021), we now define a linear causal dynamical network without exogenous inputs and its graph.

Definition 2. (Linear causal dynamical network model).

A linear causal dynamical network (LCDN) of a process $y(t)$ is defined as a pair $(G(z), w(t))$ where

- $G(z)$ is an $(m + l) \times (m + l)$ causal rational matrix, and $[G(\infty)]_{ii} = 0$ for $i = 1, \dots, m + l$;
- $w(t) = [w_{(1)}(t), \dots, w_{(m+l)}(t)]'$ is a vector of $m + l$ independent scalar white noises with a diagonal spectrum $\Sigma \succeq 0$.

The output processes $\{y_{(i)}(t)\}_{i=1}^{m+l}$ of a LCDN are given by

$$y_{(i)}(t) = w_{(i)}(t) + \sum_{j=1}^{m+l} [G(z)]_{ij} y_{(j)}(t), \quad (2)$$

or equivalently in vector form

$$y(t) = w(t) + G(z)y(t). \quad (3)$$

Definition 3. (Graph associated with a LCDN). Let $(G(z), w(t))$ be a LCDN with output processes $\{y_{(j)}(t)\}_{j=1}^{m+l}$. Let $V := \{1, \dots, m + l\}$ be the node index set, and let $E \subseteq V \times V$ such that

$$(i, j) \in E \Leftrightarrow [G(z)]_{ij} \neq 0,$$

meaning the edge from $y_{(j)}$ to $y_{(i)}$ exists. Then the directed graph (V, E) is the graphical representation of the LCDN.

This work extends the traditional linear dynamical influence model in Dimovska and Materassi (2021) by allowing self-loops in the associated graph. We impose strict causality on the diagonal entries of $G(z)$ by defining $[G(\infty)]_{ii} = 0$, thereby ensuring that the expansion of the right-hand side (RHS) in (2) contains no term for $y_{(i)}(t)$ at time t .

While existing methods like Materassi and Salapaka (2012), Dimovska and Materassi (2021) and Avventi et al. (2013), require $\Phi(z)$ to be full-rank for consistent topology identification, we relax this constraint to include the more general case of low rank processes and allow self-loops. Within this extended framework, we investigate the topology learning of low rank processes.

2.2 Special Feedback Model

Naturally, suppose $y(t)$ and $\Phi(z)$ can be partitioned as

$$y(t) = \begin{bmatrix} y_m(t) \\ y_l(t) \end{bmatrix}, \quad \Phi(z) = \begin{bmatrix} \Phi_m(z) & \Phi_{lm}^*(z) \\ \Phi_{lm}(z) & \Phi_l(z) \end{bmatrix}, \quad (4)$$

where $y_m(t) \in \mathbb{R}^m$, $y_l(t) \in \mathbb{R}^l$, and $\text{rank}(\Phi_l(z)) = l$ (i.e., y_l is a full rank process). This can be easily achieved by reordering the components of the original process $y^o(t)$.

Based on the rank-deficient property, from our previous works such as Cao et al. (2023a) and Cao et al. (2023b), a deterministic relation always exists between $\Phi_{lm}(z)$ and $\Phi_l(z)$. And the relationship between $y_m(t)$ and $y_l(t)$ can be described by the following special feedback model.

Definition 4. (Special feedback model). For a low rank vector process $y(t)$ as in (4) with $\text{rank}(\Phi_l(z)) = \text{rank}(\Phi(z)) = l$, there is a special feedback model,

$$y_m(t) = H(z)y_l(t), \quad (5a)$$

$$y_l(t) = F(z)y_m(t) + v_l(t), \quad (5b)$$

where $v_l(t)$ is an error process. The deterministic relation

$$H(z) = \Phi_{lm}(z)^* \Phi_l(z)^{-1}, \quad (6)$$

yields a unique $m \times l$ -dimensional casual function determined from $\Phi(z)$, and $F(z)$ is a (strictly) causal function.

The non-unique forward-loop function $F(z)$ may be determined from such as a minimal realization of $y(t)$ (Cao et al. (2023a)), or a one-step Wiener predictor (Cao et al. (2023b)).

2.3 Low Rank Dynamical Network

Though $y(t)$ is low-rank with a singular spectrum, its subprocess $y_l(t)$ has a nonsingular spectrum. Recalling the causal deterministic relation from y_l to y_m , a specific LCDN can be constructed for $y(t)$ via direct calculations.

Proposition 5. The low rank process $y(t)$ admits a LCDN representation $(G(z), w(t))$ with

$$G(z) = \begin{bmatrix} 0 & G_{ml}(z) \\ 0 & G_l(z) \end{bmatrix}, \quad w(t) = \begin{bmatrix} 0 \\ w_l(t) \end{bmatrix}, \quad (7)$$

where $G_{ml}(z), G_l(z)$ are causal rational, and $w_l(t)$ is l -dimensional with a full-rank spectrum.

From Proposition 5, the edge set E of the corresponding graph can be shrunk to a subset of $V \times \{m + 1, \dots, m + l\}$, because $[G(z)]_{ij} \equiv 0$ for $j = 1, \dots, m$. To simplify the representation and the subsequent topology estimation, we introduce the following model.

Definition 6. (Low rank dynamical network model). A low rank dynamical network (LRDN) for process $y(t)$ under partition (4) with $\text{rank}(\Phi_1(z)) = \text{rank}(\Phi(z)) = l$, is a triple $(G_{\text{ml}}(z), G_1(z), w_1(t))$, where

- $G_{\text{ml}}(z)$ is an $m \times l$ causal rational matrix;
- $G_1(z)$ is an $l \times l$ causal rational matrix, where $[G_1(\infty)]_{ii} = 0$ for $i = 1, \dots, l$;
- $w_1(t) = [w_{1(1)}, \dots, w_{1(l)}]'$ is a vector of l independent scalar white noises with a diagonal spectrum $\Sigma_1 \succ 0$.

The output processes are given by

$$y_m(t) = G_{\text{ml}}(z)y_1(t), \quad (8a)$$

$$y_l(t) = w_1(t) + G_1(z)y_1(t). \quad (8b)$$

Definition 7. (Graph associated with a LRDN). Let $(G_{\text{ml}}(z), G_1(z), w_1(t))$ be a LRDN of process $y(t)$ with outputs in (8). Let $V := \{1, \dots, m+l\}$ be the node index set and $V_1 := \{m+1, \dots, m+l\}$. Let $E = E_m \cup E_l$, $E_m \subset (V \setminus V_1) \times V_1$, $E_l \subset V_1 \times V_1$, such that for $j \in V_1$,

$$(i, j) \in E_m \Leftrightarrow [G_{\text{ml}}(z)]_{i, j-m} \neq 0, \text{ for } i \in (V \setminus V_1), \quad (9a)$$

$$(i, j) \in E_l \Leftrightarrow [G_1(z)]_{i-m, j-m} \neq 0, \text{ for } i \in V_1. \quad (9b)$$

Then the directed graph (V, E) is the graphical representation of the LRDN.

Definitions 6 and 7 establish a one-one correspondence between an LRDN and a low-rank process under a fixed partition (the proof of this identifiability is omitted for space). The LRDN is well-posed if $I - G_1(z)$ is invertible. The associated LRDN graph contains at most $(m+l) \times l$ edges, compared to $(m+l)^2$ for a full LCDN graph. This sparsity leads to significant computational advantages. In the following sections, we develop a Wiener filter-based method for estimating the unique graph of a LRDN.

3. CAUSAL WIENER FILTER

While the causal deterministic relation from y_l to y_m in (5a) determines the edges in E_m , this section, focuses on the internal topology within y_l . We derive a Wiener filter for each entry of $y_l(t)$ based on the other entries and its own strict past, and establish the uniqueness of this filter regardless of the specific spectral factor. Suppose the full-rank process $y_l(t)$ admits a minimum-phase spectral factorization

$$\Phi_l(z) = W(z)\Lambda W(z)^*, \quad (10)$$

where $W(z)$ is l -dimensional and square, with the form

$$W(z) = \sum_{k=0}^{n_W} W_k z^{-k}, \quad (11)$$

for some degree n_W , W_k are constant matrices, W_0 is invertible. Suppose the corresponding l -dimensional innovation process is $e(t)$ with spectrum Λ , we have

$$y_l(t) = W(z)e(t). \quad (12)$$

Under the above settings, $W(z)$ is unique up to right multiplication by a constant unitary matrix U satisfying $\Lambda U = U\Lambda$.

Denote the i -th entry of $y_l(t)$ by $y_{l(i)}(t) = y_{(i+m)}(t)$. For a subset $\mathcal{I} = \{i_1, i_2, \dots, i_r\} \subseteq \{1, \dots, l\}$ satisfying $i_1 < \dots < i_r$, denote by $B_{\mathcal{I}} = [b_{i_1}, b_{i_2}, \dots, b_{i_r}]'$, with b_k the k -th elementary (column) vector with 1 in the k -th entry and 0 elsewhere. Denote by

$$y_{\mathcal{I}}(t) := B_{\mathcal{I}}y_l(t), \quad W_{\mathcal{I}}(z) := B_{\mathcal{I}}W(z),$$

the entries of $y_l(t)$ indexed by $\mathcal{I}$, and the rows of $W(z)$ indexed by $\mathcal{I}$. Let $\overline{\mathcal{I}} := \{1, \dots, l\} \setminus \{\mathcal{I}\}$, and let $\{\cdot\}_c$ denote the orthogonal projection operator yielding the causal part of a function. Then we have the following results.

Lemma 8. Suppose the full rank process $y_l(t)$ has a spectral factorization (10). For process $y_{l(i)}(t)$, define subspace

$$\mathbf{Y}_{l(i)}^-(t) := \text{span}\{y_{l(i)}(k-1), y_{l(j)}(k) : j \in \overline{\mathcal{I}}, k \leq t\}, \quad (13)$$

hereafter $\mathbf{Y}_{l(i)}^-$, meaning the space generated by the strict past of $y_{l(i)}$ and the past of other entries in y_l . Consider the problem

$$\min_{q(t) \in \mathbf{Y}_{l(i)}^-} \|q(t) - y_{l(i)}(t)\|^2,$$

for a Wiener filter for $y_{l(i)}(t)$ on space $\mathbf{Y}_{l(i)}^-$. The solution exists, is unique regardless of the specific $W(z)$, and has the form

$$\hat{y}_{l(i)}(t) = (W_{\{\mathcal{I}\}}(z) - [W_0]_{ii}b'_i)W(z)^{-1}M_i(z)^*x_i^-(t), \quad (14a)$$

where

$$M_i(z) := I + (z^{-1} - 1)b_i b'_i, \quad (14b)$$

$$\begin{aligned} x_i^-(t) &:= \begin{bmatrix} B_{\{\mathcal{I}\}} \\ b_i \end{bmatrix}' \begin{bmatrix} y_{\overline{\mathcal{I}}}(t) \\ y_{l(i)}(t-1) \end{bmatrix} \\ &= [y_{l(1)}(t), \dots, y_{l(i)}(t-1), \dots, y_{l(l)}(t)]'. \end{aligned} \quad (14c)$$

Proof. From (14b), $M_i(z)$ is a diagonal causal function matrix with entry $[M_i(z)]_{ii} = z^{-1}$ and other diagonal entries equal to 1. It follows directly that

$$M_i(z)^* = I + (z - 1)b_i b'_i,$$

with $[M_i(z)]_{ii} = z$, and $M_i(z)M_i(z)^* = I$, i.e., $M_i(z)$ is unitary. Then $x_i^-(t)$ in (14c) is given by

$$x_i^-(t) = M_i(z)y_l(t). \quad (15)$$

Denote by

$$\varepsilon_i^-(t) := M_i(z)e(t).$$

Recall from (12) that an innovation process of $x_i^-(t)$ is given by $P\varepsilon_i^-(t)$ with an arbitrary $l \times l$ constant unitary matrix P satisfying $P^*P = I$. Consider the general solution $W(z)U$ to (10) satisfying $\Lambda U = U\Lambda$ and $UU^* = I$, such that

$$x_i^-(t) = M_i(z)W(z)UU^*M_i(z)^*P^*P\varepsilon_i^-(t).$$

Since

$$y_{l(i)} = W_{\{\mathcal{I}\}}(z)UU^*M_i(z)^*P^*P\varepsilon_i^-(t),$$

the Wiener filter for $y_{l(i)}(t)$ on $\mathbf{Y}_{l(i)}^-$ is given by

$$\mathbb{E}\{y_{l(i)}(t)|\mathbf{Y}_{l(i)}^-\} = \{W_{\{\mathcal{I}\}}M_i^*P^*\}_c (M_iWM_i^*P^*)^{-1}x_i^-(t),$$

which keeps the same for different spectral factors satisfying (10). We now prove this Wiener filter is unique independent of the specific choice of P . Since P^* is constant, the entries of $W_{\{\mathcal{I}\}}M_i^*P^*$ corresponds to l linear combinations of the entries in following row-vector,

$$W_{\{\mathcal{I}\}}M_i^* = [[W]_{i1}, \dots, z[W]_{ii}, \dots, [W]_{il}],$$

where only the i -th entry might be not causal because $W(z)$ is causal. Thus the causal projection and linear combination operations can commute, i.e., $\{W_{\{\mathcal{I}\}}M_i^*P^*\}_c = \{W_{\{\mathcal{I}\}}M_i^*\}_c P^*$, and

$$\begin{aligned}\mathbb{E}\{y_l(t)|\mathbf{Y}_{l(i)}^-\} &= \{W_{\{i\}}M_i^*\}_C P^* (M_i W M_i^* P^*)^{-1} x_i^-(t) \\ &= \{W_{\{i\}}M_i^*\}_C M_i W^{-1} M_i^* x_i^-(t),\end{aligned}$$

verifying the uniqueness.
From (11),

$$\begin{aligned}\{z[W(z)]_{ii}\}_C &= \{z[W_0]_{ii} + \sum_{k=1}^{n_W} [W_k]_{ii} z^{-k+1}\}_C \\ &= z[W(z)]_{ii} - z[W_0]_{ii}.\end{aligned}$$

Then we have

$$\begin{aligned}\{W_{\{i\}}M_i^*\}_C M_i &= (W_{\{i\}}M_i^* - z[W_0]_{ii}b'_i)M_i \\ &= W_{\{i\}}(z) - [W_0]_{ii}zb'_iM_i.\end{aligned}$$

Combined with (14b), the unique Wiener filter is simplified to (14a). $\square$

From (15), we have $M_i(z)^*x_i^-(t) = y_l(t)$. Then from (14) we have

$$\begin{bmatrix} \hat{y}_{l(1)}(t) \\ \vdots \\ \hat{y}_{l(l)}(t) \end{bmatrix} = \begin{bmatrix} (W_{\{1\}}(z) - [W_0]_{11}b'_1)W(z)^{-1}y_l(t) \\ \vdots \\ (W_{\{l\}}(z) - [W_0]_{ll}b'_l)W(z)^{-1}y_l(t) \end{bmatrix},$$

which can be concluded by the following theorem (with the proof omitted for the space).

Theorem 9. Suppose the full rank process $y_l(t)$ has a spectral factorization (10). Denote by

$$\hat{y}_l(t) = [\hat{y}_{l(1)}(t), \dots, \hat{y}_{l(l)}(t)]', \quad (16)$$

where $\hat{y}_{l(i)}(t)$ is the Wiener filter for $y_{l(i)}(t)$ on space $\mathbf{Y}_{l(i)}^-$ from Lemma 8 for $i = 1, \dots, l$. Then there is a mapping $S(z) : \mathbf{Y}_{V_l}^- \rightarrow \mathbf{Y}_{V_l}^-$ uniquely determined from the Wiener filters $\hat{y}_{l(i)}(t)$ for $i = 1, \dots, l$, such that

$$\hat{y}_l(t) = S(z)y_l(t), \quad (17a)$$

and $S(z)$ has the form

$$S(z) = I - DW(z)^{-1}, \quad (17b)$$

where D is a diagonal matrix satisfying

$$[D]_{ii} = [W_0]_{ii}, \quad \text{for } i = 1, \dots, l, \quad (17c)$$

with W_0 given in (11). Moreover, $S(z)$ is unique regardless of the specific $W(z)$, and $[S(\infty)]_{ii} = 0$ for $i = 1, \dots, l$.

Consequently, the causal Wiener filter for each node based on its strict past and the past of other nodes, is uniquely determined. The next section establishes the equivalence between this filter and conditional Granger causality, thereby enabling the reconstruction of the directed network topology that encodes these causal relationships.

4. TOPOLOGY LEARNING BY WIENER FILTER

Building on the unique Wiener filter derived above, we now formalize its connection to Granger causality and topology learning.

4.1 Conditional Granger Causality

We first recall the geometric definition of Granger non-causality (Cao et al. (2023a)): there is no causality in the sense of Granger from process β to α if and only if

$$\mathbf{A}^+ \perp \mathbf{B}^- | \mathbf{A}^{--}, \quad (18)$$

where $\mathbf{A}^{--}$, $\mathbf{A}^+$ are subspaces generated by the strict past and the future (including present) of $\alpha(t)$, respectively, and

$\mathbf{B}^-$ is the subspace generated by the past of $\beta(t)$. Similarly, strict non-causality is equivalent to $\mathbf{A}^+ \perp \mathbf{B}^{--} | \mathbf{A}^{--}$, where $\mathbf{B}^{--}$ denotes the strictly past subspace of $\beta(t)$. This geometric formulation leads us to the following definition of the conditional case.

Definition 10. For stochastic processes $\alpha(t), \beta(t), \gamma(t)$, we say there is no conditional Granger causality from β to α given γ , if and only if

$$\mathbf{A}^+ \perp \mathbf{B}^- | \mathbf{A}^{--} \vee \mathbf{C}^-. \quad (19)$$

or equivalently, for any $\lambda \in \mathbf{A}^+$,

$$\mathbb{E}\{\lambda | \mathbf{B}^- \vee \mathbf{A}^{--} \vee \mathbf{C}^-\} = \mathbb{E}\{\lambda | \mathbf{A}^{--} \vee \mathbf{C}^-\}, \quad (20)$$

where $\mathbf{C}^-$ is subspace generated by the past of $\gamma(t)$.

The equivalence between (19) and (20) follows from Lindquist and Picci (2015) (Proposition 2.4.2). The following theorem establishes a necessary and sufficient condition in terms of the zero entries of the Wiener filter, for conditional non-causality between different nodes.

Theorem 11. Define the subspaces

$$\mathbf{Y}_{l\mathcal{I}}^-(t) := \text{span}\{y_{l(h)}(k) : h \in \mathcal{I}, k \leq t\}, \quad (21a)$$

$$\mathbf{Y}_{l\{i\}}^-(t) := \text{span}\{y_{l(i)}(k), k < t\}, \quad (21b)$$

$$\mathbf{Y}_{l\{i\}}^+(t) := \text{span}\{y_{l(i)}(k), k \geq t\}, \quad (21c)$$

hereafter $\mathbf{Y}_{l\mathcal{I}}^-, \mathbf{Y}_{l\{i\}}^-, \mathbf{Y}_{l\{i\}}^+$. Let $S(z)$ be the Wiener filter matrix from Theorem 9. Then, for any $i \neq j$,

$$[S(z)]_{ij} = 0 \Leftrightarrow \mathbf{Y}_{l\{i\}}^+ \perp \mathbf{Y}_{l\{j\}}^- | (\mathbf{Y}_{l\{i\}}^- \vee \mathbf{Y}_{l\{i,j\}}^-), \quad (22)$$

that is, $[S(z)]_{ij} = 0$ if and only if there is no conditional Granger causality from $y_{l(j)}(t)$ to $y_{l(i)}(t)$ given the past of all other entries in $y_l(t)$.

Proof. We first prove that for any $i \neq j$,

$$[S(z)]_{ij} = 0 \Leftrightarrow y_{l(i)}(t) \perp \mathbf{Y}_{l\{j\}}^- | (\mathbf{Y}_{l\{i\}}^- \vee \mathbf{Y}_{l\{i,j\}}^-), \quad (23)$$

and then extends this to (22).

Let $\mathbf{Z}(t) := \mathbf{Y}_{l\{i\}}^-(t) \vee \mathbf{Y}_{l\{i,j\}}^-(t)$. The RHS of (23) is equivalent to

$$\mathbb{E}\{y_{l(i)}(t) | \mathbf{Y}_{l(i)}^-(t)\} = \mathbb{E}\{y_{l(i)}(t) | \mathbf{Z}(t)\}, \quad (24)$$

with $\mathbf{Y}_{l(i)}^-$ defined in (13). From Theorem 9, the left-hand side (LFS) of (24) is

$$\mathbb{E}\{y_{l(i)}(t) | \mathbf{Y}_{l(i)}^-(t)\} = \hat{y}_{l(i)}(t) = \sum_{h=1}^l [S(z)]_{ih} y_{l(h)}(t).$$

Denote by $e_{(i)}(t)$ the i -th entry of $e(t)$. Then from (17),

$$y_{l(i)}(t) - \mathbb{E}\{y_{l(i)}(t) | \mathbf{Y}_{l(i)}^-(t)\} = [D]_{ii} e_{(i)}(t).$$

Sufficiency of (23): Suppose $[S(z)]_{ij} = 0$. Since $\mathbf{Z}(t) \subseteq \mathbf{Y}_{l(i)}^-(t)$, we have

$$\begin{aligned}\|y_{l(i)}(t) - \mathbb{E}\{y_{l(i)}(t) | \mathbf{Z}(t)\}\|^2 \\ \geq \|y_{l(i)}(t) - \mathbb{E}\{y_{l(i)}(t) | \mathbf{Y}_{l(i)}^-(t)\}\|^2 = \|[D]_{ii} e_{(i)}(t)\|^2.\end{aligned} \quad (25)$$

Moreover, $\hat{y}_{l(i)}(t) \in \mathbf{Z}(t)$ from $[S(z)]_{ij} = 0$ and $y_{l(i)}(t) - \hat{y}_{l(i)}(t) = [D]_{ii} e_{(i)}(t) \perp \mathbf{Z}(t)$. Combining this with (25), it follows that $\hat{y}_{l(i)}(t)$ is also the Wiener filter of $y_{l(i)}(t)$ on $\mathbf{Z}(t)$, and thus (24) holds.

Necessity of (23): Suppose (24) holds. Then,

$$\sum_{h=1}^l [S(z)]_{ih} y_{l(h)}(t) \in \mathbf{Z}(t).$$

Since $\sum_{h \neq j} [S(z)]_{ih} y_{1(h)}(t) \in \mathbf{Z}(t)$, it follows that $[S(z)]_{ij} y_{1(j)}(t) \in \mathbf{Z}(t)$. This implies

$$[S(z)]_{ij} y_{1(j)}(t) = T(z) y_{1\{i,j\}}^-(t),$$

where $T(z)$ is a causal polynomial row vector. If $T(z) \neq 0$ and hence $[S(z)]_{ij} \neq 0$, the Wiener filter $\hat{y}_{1(i)}(t)$ would be non-unique, contradicting the uniqueness in Lemma 8. Hence, $[S(z)]_{ij} = 0$ is necessary.

Extension to (22): the RHS of (23) is equivalent to

$$y_{1(i)}(t + \tau) \perp \mathbf{Y}_{1\{j\}}^-(t + \tau) \mid \mathbf{Z}(t + \tau),$$

for any $\tau \geq 0$. Since $\mathbf{Z}(t) \subseteq \mathbf{Z}(t + \tau)$, and $\mathbf{Y}_{1\{j\}}^-(t) \subseteq \mathbf{Y}_{1\{j\}}^-(t + \tau)$ for $\tau \geq 0$, it follows that

$$y_{1(i)}(t + \tau) \perp \mathbf{Y}_{1\{j\}}^-(t) \mid \mathbf{Z}(t). \quad (26)$$

Now, for any $\lambda = \bar{T}(z) y_{1(i)}(t) \in \mathbf{Y}_{1\{i\}}^+$ with $\bar{T}(z) = \sum_{k=0}^{n_T} \bar{T}_k z^k$ anticausal, we have

$$\lambda = \sum_{k=0}^{n_T} \bar{T}_k y_{1(i)}(t + k),$$

where each term $\bar{T}_k y_{1(i)}(t + k)$ is conditionally orthogonal to subspace $\mathbf{Y}_{1\{j\}}^-(t)$ given $\mathbf{Z}(t)$ from (26). Therefore,

$$\lambda \perp \mathbf{Y}_{1\{j\}}^-(t) \mid \mathbf{Z}(t), \quad \forall \lambda \in \mathbf{Y}_{1\{i\}}^+,$$

which establishes (22). $\square$

From Theorem 11 and Equation (17b), the following result is ready concerning conditional Granger non-causality between nodes in a full-rank vector process, which provides a criterion for causal inference without explicit reference to the graph topology.

Corollary 12. For a full-rank process $y_1(t)$ with a minimum-phase spectral factor $W(z)$ in (10), and for $i \neq j$, we have

$$[W(z)^{-1}]_{ij} = 0 \Leftrightarrow \mathbf{Y}_{1\{j\}}^+ \perp \mathbf{Y}_{1\{j\}}^- \mid (\mathbf{Y}_{1\{i\}}^- \vee \mathbf{Y}_{1\{i,j\}}^-), \quad (27)$$

$\Leftrightarrow$ the absence of conditional Granger causality from $y_{1(j)}$ to $y_{1(i)}$ given the rest entries of $y_1(t)$.

In contrast to the symmetric relation of conditional independence characterized by (1), conditional Granger non-causality is directional and typically asymmetric. This directional property enables the direct learning of oriented edges. It bypasses the complexity inherent in inferring directed connections from symmetric measures, which often requires excluding indirectly connected ‘kin nodes’. We thus proceed to reconstruct the directed causal graph associated with a LRDN, where the presence of an edge (i, j) is equivalent to the existence of conditional Granger causality from $y_{1(j)}$ to $y_{1(i)}$.

4.2 LRDN Topology Learning

As established in Subsection 2.2, matrix $G_{ml}(z)$ in a LRDN (8) can be uniquely determined by the causal deterministic relation $H(z)$ in the special feedback model (5), which is itself a Wiener filter. We now demonstrate that the entire topology of an LRDN can be exactly reconstructed using Wiener filters between nodes. The associated graph admits both loops and strictly causal self-loops, offering greater generality than existing approaches. Moreover, the presence of a directed edge precisely corresponds to conditional Granger causality between the nodes.

Theorem 13. For a LRDN $(G_{ml}(z), G_1(z), w_1(t))$, suppose the $(m + l)$ -dimensional output process $y(t)$ with its spectrum are given in (4), satisfying $\text{rank}(\Phi(z)) = \text{rank}(\Phi_1(z)) = l$. Define subspaces $\mathbf{Y}_{\mathcal{I}}^-, \mathbf{Y}_{\{i\}}^-, \mathbf{Y}_{\{i\}}^+$ as in (21) by substituting $y_1(t)$ by $y(t)$.

Then, the edge set E of the graph (V, E) associated with the LRDN can be consistently reconstructed by the following sufficient and necessary condition. For $j \in V_1$,

$$(i, j) \in E \Leftrightarrow \begin{cases} [H(z)]_{i,j-m} \neq 0, & i \in (V \setminus V_1), \\ [S(z)]_{i-m,j-m} \neq 0, & i \in V_1, \end{cases} \quad (28)$$

where $V = \{1, \dots, m + l\}$, $V_1 = \{m + 1, \dots, m + l\}$, $H(z)y_1(t)$ is the causal Wiener filter for $y_m(t)$ given $y_1(t)$, the i -th row of $S(z)y_1(t)$, denoted by $S_{\{i\}}(z)y_1(t)$, is the causal Wiener filter of $y_{1(i)}(t)$ given its own strict past and the past of all other entries in $y_1(t)$.

Moreover, for a LRDN, there is no conditional Granger causality from node $y_{(j)}(t)$ to node $y_{(i)}(t)$ given all other entries in $y(t)$, if and only if $(i, j) \notin E$, i.e.,

$$(i, j) \notin E \Leftrightarrow \mathbf{Y}_{\{i\}}^+ \perp \mathbf{Y}_{\{j\}}^- \mid \mathbf{Y}_{\{i\}}^- \vee \mathbf{Y}_{V \setminus \{i,j\}}^-, \quad (29)$$

for $i \in V$, and $j \in V_1$.

Proof. The LRDN definition relies on the existence of a deterministic causal relation exists from y_1 to y_m . As shown in Subsection 2.2 and (6), $H(z)$ is the unique causal Wiener filter from y_1 to y_m . Therefore, in the LRDN we have

$$G_{ml}(z) = H(z), \quad (30)$$

which implies that the non-zeros entries of $H(z)$ correspond exactly to the edges in $E_m \subset (V \setminus V_1) \times V_1$.

With the same innovation spectrum $\Lambda = \Sigma_1 > 0$, it follows from (8b)(17b) that

$$I - S(z) = D(I - G_1(z)).$$

Consequently,

$$\begin{aligned} [S(z)]_{ij} &= [D]_{ii} [G_1(z)]_{ij}, \text{ when } i \neq j, \\ 1 - [S(z)]_{ii} &= [D]_{ii} (1 - [G_1(z)]_{ii}). \end{aligned}$$

Since $[G_1(\infty)]_{ii} = [S(\infty)]_{ii} = 0$, we have $[G_1(z)]_{ii} \neq 0 \Leftrightarrow [G_1(z)]_{ii}$ is strictly causal nonzero $\Leftrightarrow [S(z)]_{ii}$ is strictly causal and non-zero. Therefore, by (9), the edge condition (28) holds.

We now prove the necessary and sufficient condition for conditional Granger non-causality, considering $i \leq m$ and $i > m$ separately. Since both y_m and y_1 are determined by y_1 in (8), the conditional Granger non-causality from $y_{(j)}$ to $y_{(i)}$ for $j \in V_1$ is equivalent to

$$\mathbf{Y}_i^+ \perp \mathbf{Y}_j^- \mid \mathbf{Y}_{1\{j-m\}}^-, \text{ for } i \leq m, \quad (31a)$$

$$\mathbf{Y}_i^+ \perp \mathbf{Y}_j^- \mid \mathbf{Y}_{\{i\}}^- \vee \mathbf{Y}_{1\{i-m,j-m\}}^-, \text{ for } i > m. \quad (31b)$$

For $i > m$, by Theorem 11, (31b) holds $\Leftrightarrow [S(z)]_{i-m,j-m} = 0 \Leftrightarrow (i, j) \notin E_1$. For $i \leq m$, note that $\mathbf{Y}_{\{j\}}^- \vee \mathbf{Y}_{1\{j-m\}}^- = \mathbf{Y}_1^-$. From (8a)(30),

$$\mathbb{E}\{y_{1(i)}(t) \mid \mathbf{Y}_{\{j\}}^- \vee \mathbf{Y}_{1\{j-m\}}^- \} = \sum_{h=1}^l [H(z)]_{ih} y_{1(h)}.$$

Following a derivation similar to that in the proof of Theorem 11, we conclude that (31a) $\Leftrightarrow [H(z)]_{i,j-m} = 0 \Leftrightarrow (i, j) \notin E_m$. $\square$

The graphical structure of a LRDN estimated in Theorem 13, fully captures the conditional Granger causal

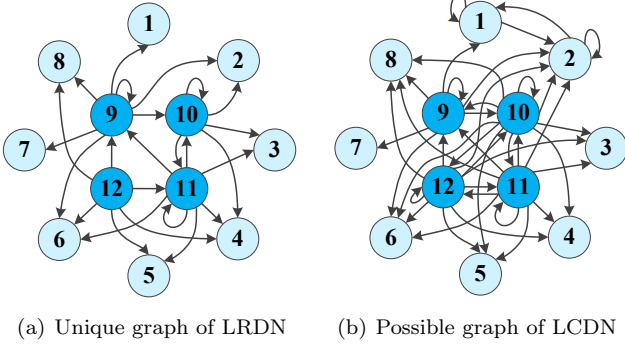


Fig. 1. Graphs associated to different network models of the low rank process in the simulation example. Panel (a) also corresponds to the estimated LRDN graph by Theorem 13.

relationships between different nodes in the low rank output process. This Wiener filter-based formulation offers significant identification advantages when W_0 is set to have unit diagonal entries. In this framework, $H(z)$ directly represents the deterministic relation, while $S(z)$ links to a minimum-phase spectral factor. Further elaboration is reserved for the journal version of this work.

5. SIMULATION EXAMPLE

We consider a 12-dimensional process $y(t)$ innovated by a 4-dimensional white noise with identify spectrum, generated as

$$y(t) = \begin{bmatrix} H(z)W(z) \\ W(z) \end{bmatrix} e(t).$$

To construct the graph, we first define the edge set E_1 by specifying the off-diagonal nonzero entries of $W(z)^{-1}$ according to Corollary 12, and set $y_{(12)}(t) = e_{(4)}(t)$ so that node 12 has no self-loop or incoming edge in E_1 . Next, we form E_m by choosing the nonzero entries of $H(z)$. The overall edge set is $E = E_m \cup E_1$. Finally, we randomly generate values of the non-zero entries of $W(z)$ and $H(z)$ consistent with E , and simulate $y(t)$ for $t = 1, \dots, 200$. The graph of the LRDN generated for this example is shown in Fig. 1 (a), where the nodes in y_1 are marked in darker blue. By modeling as a LRDN, there are 25 directed edges among the 12 nodes in its unique graph, pointing from V_1 to V . This graph represents the target topology to be estimated from the simulated data.

In this case, the singular spectrum prevents the construction of an interactive graph (i.e., moral graph) from conditional dependencies between different nodes. Ignoring the low-rank nature of the process during reconstruction leads to non-unique Wiener filter representations, which consequently prevents the unique recovery of a LCDN based on conditional Granger causality. A possible LCDN graph inferred under this omission is shown in Fig. 1 (b), which contains 41 edges and exhibits markedly more complex internal connections than the corresponding LRDN. Moreover, in an extreme scenario where all off-diagonal entries of $W(z)$ are constant, a complete LCDN graph with 132 edges may be obtained—a result that would severely undermine subsequent prediction or control tasks.

Using least square methods, we obtain unbiased estimates of the Wiener filters $S(z)$ in Theorem 9 and $H(z)$ in (5a). The existence of all edges in E of the LRDN graph is tested

based on Theorem 13. We find that the reconstructed graph under the LRDN model from the sampled data perfectly matches the original graph in Fig. 1 (a). The above results demonstrate that LRDN offers an efficient and well-posed modeling framework for low rank processes, while the employed Wiener filters provide a reliable and consistent approach for network topology learning.

6. CONCLUSIONS

In this work we have addressed the modeling and topology estimation of linear causal dynamical networks exhibiting low rank output processes. To this end, we have proposed the low rank dynamical network (LRDN) model, which provides a parsimonious and identifiable representation of the internal correlations. The topology of an LRDN is recovered using causal Wiener filters: one capturing the deterministic relation between the subprocesses in the special feedback model, and the other operating under the constraint of strictly causal self-loops. We establish a fundamental necessary and sufficient condition that links the presence of an edge to conditional Granger causality, or equivalently, to the non-zero entries of the corresponding Wiener filters. A simulation example demonstrates the parsimony of the LRDN framework and the consistent performance of our topology estimation method. Further details and network identification under output errors will be discussed in a forthcoming journal version.

REFERENCES

- Avventi, E., Lindquist, A., and Wahlberg, B. (2013). ARMA identification of graphical models. *IEEE Transactions on Automatic Control*, 58(5), 1167–1178.
- Cao, W., Lindquist, A., and Picci, G. (2023a). Modeling of low rank time series. *IEEE Transactions on Automatic Control*, 68(12), 7270–7285.
- Cao, W., Picci, G., and Lindquist, A. (2023b). Identification of low rank vector processes. *Automatica*, 151, 110938.
- Dimovska, M. and Materassi, D. (2021). A control theoretic look at Granger causality: Extending topology reconstruction to networks with direct feedthroughs. *IEEE Transactions on Automatic Control*, 66(2), 699–713.
- Lindquist, A. and Picci, G. (2015). *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification*. Springer, Berlin Heidelberg.
- Materassi, D. and Salapaka, M.V. (2012). On the problem of reconstructing an unknown topology via locality properties of the Wiener filter. *IEEE Transactions on Automatic Control*, 57(7), 1765–1777.
- Pillonetto, G., Aravkin, A., Gedon, D., Ljung, L., Ribeiro, A., and Schön, T. (2025). Deep networks for system identification: A survey. *Automatica*, 171, 111907.
- Vincent, T., Antoine, A., and Patrick, D. (2024). The low-rank hypothesis of complex systems. *Nature Physics*, 20(2), 294–302.
- Yu, J. and Qin, S.J. (2023). Low-dimensional latent state space identification with application to the shell control process. In *2023 IEEE Conference on Control Technology and Applications (CCTA)*, 1004–1009.
- Zorzi, M. (2025). On the identification of ARMA graphical models. *IEEE Transactions on Automatic Control*, 70(1), 403–414.