

University of Groningen

Relationship between Granger non-causality and network graph of state-space representations

Jozsa, Monika

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Jozsa, M. (2019). *Relationship between Granger non-causality and network graph of state-space representations*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 5

Granger causality and innovation transfer matrices

The results on LTI-SS representations presented in Chapters 2, 3 and 4 have direct implications on the transfer matrices between the input and output processes of LTI-SS representations. Note that in Chapters 2, 3, and 4 we studied Kalman representations, i.e., LTI-SS representations where the noise process is the innovation process of the output process. Therefore, to present these implications, we introduce a class of transfer matrices between processes and their innovation processes, called innovation transfer matrices. Among innovation transfer matrices, we focus on those that have transitive acyclic directed graphs (TADGs) as their network graphs. In the main results, we show that an innovation transfer matrix of a process has network graph G , where G is TADG, if and only if the components of this process satisfy conditional Granger causality conditions that are consistent with G and coincide with the conditions presented in Chapter 4 as G -consistent causality structure..

The paper (Caines and Wynn, 2007) is the closest one to the results in this chapter. The cited paper presents results on transfer matrices of Gaussian processes. These transfer matrices form a subclass of the transfer matrices discussed in the present chapter with additional assumptions on the covariance matrix of the noise process. Accordingly, the existence of these transfer matrices are formalized by conditional orthogonal conditions which are stronger than the conditional orthogonality condition that are counterparts of the conditional Granger causalities proposed in this chapter (and in Chapter 4). Note that (Caines and Wynn, 2007) do not provide proofs for the results on transfer matrices. The results of this chapter are based on the conference paper (Jozsa et al., 2017b).

First, transfer matrices and innovation transfer matrices of ZMSIR processes are introduced. Then, network graphs of innovation transfer matrices are related to a collection of conditional Granger causality conditions among the components of the output process. At last, we give an example to illustrate the results. The proofs of the statements can be found in Appendix 5.A. Note that this chapter presents its statements and proofs independently from Chapters 2, 3, and 4.

5.1 Innovation transfer matrix

To begin with, we introduce transfer matrices of ZMSIR processes. Consider an LTI-SS representation $(A, B, C, D, \mathbf{v},)$ of a ZMSIR process \mathbf{y} . Then the transfer matrix between \mathbf{y} and \mathbf{v} is the rational matrix function $G(z) = C(zI - A)^{-1}B + D$, see (Anderson and Moore, 1979, Appendix C & D). Consider the Laurent series expansion of $G(z)$, i.e. let $G_k \in \mathbb{R}^{n \times m}$, $k \geq 0$ be such that if $|z| \geq 1$ then $G(z) = \sum_{k=0}^{\infty} G_k z^{-k}$ for all $z \in \mathbb{C}$. Note that for any ZMSIR process \mathbf{z} , the expression $\sum_{k=0}^{\infty} G_k \mathbf{z}(t - k)$ converges in the topology of $\mathcal{H}_{t-}^{\mathbf{z}}$ (Anderson and Moore, 1979, Theorem 4.1). In the sequel, we will write

$$G(z)\mathbf{z}(t) = \sum_{k=0}^{\infty} G_k \mathbf{z}(t - k).$$

That is, $G(z)$ can be interpreted as a causal linear filter which transforms \mathbf{z} to the process $\{G(z)\mathbf{z}(t)\}_{t=-\infty}^{\infty}$.

From now on, we will assume that the processes considered in this chapter are ZMSIR and coercive (see (Lindquist and Picci, 2015, Definition 9.4.1)).

Assumption 5.1. *All the processes are ZMSIR and coercive.*

Below we define a class of transfer matrices, called innovation transfer matrices. To this end, consider a process \mathbf{y} and its covariance sequence $\Lambda_k = E[\mathbf{y}(t+k)\mathbf{y}^T(t)]$ for all $k \in \mathbb{Z}$. Then \mathbf{y} has a spectral density function $\Phi_{\mathbf{y}}(z) = \sum_{k=-\infty}^{+\infty} \Lambda_k z^k$ defined for all $z \in \mathbb{C}$ such that $|z| = 1$. It is well known (Lindquist and Picci, 2015; Anderson and Moore, 1979) that $\Phi_{\mathbf{y}}(z)$ admits a unique decomposition

$$\Phi_{\mathbf{y}}(z) = P(z)\Omega P^T(z^{-1}), \quad |z| = 1 \quad (5.1)$$

with the conditions: Ω is a positive definite symmetric matrix, $\lim_{z \rightarrow \infty} P(z) = I$, and $P(z)$ and its inverse $Q(z) = P^{-1}(z)$ are transfer matrices between the input \mathbf{v} and output \mathbf{y} processes of a finite-dimensional deterministic exponentially stable LTI-SS representation such that $\mathbf{y}(t) = P(z)\mathbf{v}(t)$ and $\mathbf{v}(t) = Q(z)\mathbf{y}(t)$ (Anderson and Moore, 1979, Section 9.4, Theorem 4.1).

In view of Assumption 5.1, note that coercivity of \mathbf{y} is equivalent to $P(z)$ being minimum phase, that is $P(z)$ has a stable rational proper inverse. Note that $P(z)Q(z) = Q(z)P(z) = I$ for all z in the domain of definition of P and Q .

Consider the innovation process $\mathbf{e}(t) = \mathbf{y}(t) - E_t[\mathbf{y}(t)|\mathcal{H}_{t-}^{\mathbf{y}}]$ of \mathbf{y} . It is known (Lindquist and Picci, 2015, Section 4.1.3) that \mathbf{e} is a white noise process for which

$\mathcal{H}_{t-}^e = \mathcal{H}_{t-}^y$. From the properties of $P(z)$ and $Q(z)$ the expressions

$$P(z)\mathbf{e}(t) = \sum_{k=0}^{\infty} P_k \mathbf{e}(t-k); \quad Q(z)\mathbf{y}(t) = \sum_{k=0}^{\infty} Q_k \mathbf{y}(t-k)$$

are well defined, where $P(z) = \sum_{k=0}^{\infty} P_k z^{-k}$ and $Q(z) = \sum_{k=0}^{\infty} Q_k z^{-k}$ are the Laurent series expansions of $P(z)$ and $Q(z)$. Moreover, for all $t \in \mathbb{Z}$,

$$\mathbf{y}(t) = P(z)\mathbf{e}(t) \quad \text{and} \quad \mathbf{e}(t) = Q(z)\mathbf{y}(t).$$

That is, $P(z)$ and $Q(z)$ define the relation between \mathbf{y} and its innovation process \mathbf{e} .

Definition 5.2. The transfer matrix $P(z)$ that satisfies (5.1) is called *the innovation transfer matrix of \mathbf{y}* .

The innovation transfer matrix $P(z)$ is a causal linear filter which transforms the innovation process \mathbf{e} into \mathbf{y} and $Q(z)$ is a causal linear filter which transforms \mathbf{y} into the innovation process \mathbf{e} . Therefore, the transfer matrix of a Kalman representation coincides with the innovation transfer matrix $P(z)$ of \mathbf{y} .

The innovation transfer matrix of \mathbf{y} has the following interpretation in terms of LTI-SS representations. Consider an LTI-SS representation (A, B, C, D, \mathbf{v}) of \mathbf{y} . Note that the transfer matrix $G(z) = C(zI - A)^{-1}B + D$ between \mathbf{y} and \mathbf{v} satisfies $\Phi_{\mathbf{y}}(z) = G(z)E[\mathbf{v}(t)\mathbf{v}^T(t)]G^T(z^{-1})$. Moreover, $G(z)$ has Laurent series expansion $G(z) = \sum_{k=0}^{\infty} G_k z^{-k}$ for any $z \in \mathbb{C}$, $\|z\| \geq 1$, where $G_0 = D$ and $G_k = CA^{k-1}B$, $k \geq 1$, thus

$$\mathbf{y}(t) = G(z)\mathbf{v}(t) = \sum_{k=1}^{\infty} CA^{k-1}B\mathbf{v}(t-k) + D\mathbf{v}(t).$$

If we consider an LTI-SS representation of a coercive process (see Assumption 5.1) then its transfer matrix $G(z)$ has a stable rational proper inverse (it is minimum phase). Then it can be shown that there exists an invertible matrix M , such that $\mathbf{v}(t) = M\mathbf{e}(t)$ for all $t \in \mathbb{Z}$. Therefore, $G(z)$ can be obtained from the innovation transfer matrix by multiplying it with M^{-1} from the right. Accordingly, the results presented in this chapter on innovation transfer matrices extend to transfer matrices of linear minimum phase systems.

5.2 Granger causality and innovation transfer matrix

Below, we present results on innovation transfer matrices and Granger causality that are direct implications of results in Chapter 4 for coercive processes. The general

question of that we try to answer is: what properties of an innovation transfer matrix of a process can mean conditional Granger non-causalities among the components of that process.

Partitioning a process \mathbf{y} into two components, such that $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$, it is well known that Granger non-causality from \mathbf{y}_1 to \mathbf{y}_2 is equivalent to the innovation transfer matrix of \mathbf{y} being in block triangular form ((Caines, 1976; Caines and Chan, 1975; Gevers and Anderson, 1982) and (Caines, 1988, Theorem 2.2)) or in other words, to the Wold decomposition of \mathbf{y} being in block triangular form. In the next theorem, we present this result for the class of ZMSIR processes, see e.g., (Caines, 1976, Theorem 2.2 (1) \iff (6)).

Theorem 5.3. Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$. Then \mathbf{y}_1 does not Granger cause \mathbf{y}_2 if and only if the innovation transfer matrix $P(z)$ of \mathbf{y} has the form of

$$P(z) = \begin{bmatrix} P_{11}(z) & P_{12}(z) \\ 0 & P_{22}(z) \end{bmatrix}, \quad (5.2)$$

where $P_{ij}(z)$ is an $r_i \times r_j$ block of $P(z)$, for $i, j = 1, 2$, and r_i is such that \mathbf{y}_i takes values in \mathbb{R}^{r_i} , $i = 1, 2$.

Remark 5.4. Assume that the innovation transfer matrix $P(z)$ is in the form of (5.2). Then, the inverse transfer matrix $Q(z) = P^{-1}(z)$ has the form of

$$Q(z) = \begin{bmatrix} Q_{11}(z) & Q_{12}(z) \\ 0 & Q_{22}(z) \end{bmatrix},$$

where $Q_{ij}(z)$ is an $r_i \times r_j$ dimensional transfer matrix for $i, j \in \{1, 2\}$.

Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ and let $\mathbf{e} = [\mathbf{e}_1^T, \mathbf{e}_2^T]^T$ be its innovation process such that $\mathbf{e}_i \in \mathbb{R}^{r_i}$ for $i = 1, 2$. Then (5.2) is equivalent to

$$\begin{aligned} \mathbf{y}_1(t) &= P_{11}(z)\mathbf{e}_1(t) + P_{12}(z)\mathbf{e}_2(t) \\ \mathbf{y}_2(t) &= P_{22}(z)\mathbf{e}_2(t) \end{aligned}$$

Next, we generalize Theorem 5.3 for processes partitioned into $n \geq 2$ components. More precisely, we generalize Theorem 5.3 for processes $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ with G -consistent causality structure, where $G = (V, E)$ is TADG, and, correspondingly, for a more complex zero structure of the transfer matrices. Recall that \mathbf{y} has G -consistent causality structure if $(i, j) \notin E$ implies that \mathbf{y}_i conditionally does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} , where $I_j = \{i \in V | (i, j) \in E\}$ denotes the set of parent nodes of j , see Definition 4.11.

Let $P(z)$ be the innovation transfer matrix of a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$. Assume that $\mathbf{y}_i \in \mathbb{R}^{r_i}$, $i = 1, \dots, n$ and consider the following decomposition of $P(z)$

$$P(z) = \begin{bmatrix} P_{11}(z) & P_{12}(z) & \cdots & P_{1n}(z) \\ P_{21}(z) & P_{22}(z) & \cdots & P_{2n}(z) \\ \vdots & \vdots & & \vdots \\ P_{n1}(z) & P_{n2}(z) & \cdots & P_{nn}(z) \end{bmatrix}, \quad (5.3)$$

where $P_{ij}(z)$ is a $(r_i \times r_j)$ block of $P(z)$. In other words, if $\mathbf{e} = [\mathbf{e}_1^T, \dots, \mathbf{e}_n^T]^T$ is the innovation process of \mathbf{y} such that $\mathbf{e}_i \in \mathbb{R}^{r_i}$ then, for $i = 1, \dots, n$ $\mathbf{y}_i(t) = \sum_{j=1}^n P_{ij}(z)\mathbf{e}_j(t)$.

Definition 5.5. We say that the transfer matrix $P(z)$ has G -zero structure if $(j, i) \notin E$ implies that the block $P_{ij}(z)$ of $P(z)$ is zero.

Roughly speaking, if the (j, i) -element of the adjacency matrix of G is zero then the (i, j) -block of the transfer matrix $P(z)$ is zero. Notice that if a transfer matrix has G -zero structure than its network graph, regarding the appropriate decomposition of $P(z)$, equals G . In fact, $P(z)$ can be viewed as a network of subsystems where each subsystem generates a component of $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$. Let \mathcal{S}_j , $j = 1, \dots, n$ be the subsystem generating \mathbf{y}_j , i.e.,

$$\mathbf{y}_j(t) = \sum_{i \in I_j} P_{ji}(z)\mathbf{e}_i(t). \quad (5.4)$$

Notice that if $(i, j) \in E$, i.e., i is a parent node of j , then subsystem \mathcal{S}_j takes \mathbf{e}_i as an inputs from subsystem \mathcal{S}_i . In contrast, if $(j, i) \notin E$, then \mathcal{S}_j does not take input from \mathcal{S}_i . Intuitively, it means that the subsystems communicate with each other as it is allowed by the edges of the graph G .

For instance, take the TADG graph $G = (\{1, 2, 3, 4\}, \{(4, 1), (4, 2), (3, 1), (2, 1)\})$ and a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T, \mathbf{y}_4^T]^T$ with innovation process $[\mathbf{e}_1^T, \mathbf{e}_2^T, \mathbf{e}_3^T, \mathbf{e}_4^T]^T$. Then the innovation transfer matrix $P(z)$ of \mathbf{y} with G -zero structure is given by

$$\begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \\ \mathbf{y}_3(t) \\ \mathbf{y}_4(t) \end{bmatrix} = \underbrace{\begin{bmatrix} P_{11}(z) & P_{12}(z) & P_{13}(z) & P_{14}(z) \\ 0 & P_{22}(z) & 0 & P_{24}(z) \\ 0 & 0 & P_{33}(z) & 0 \\ 0 & 0 & 0 & P_{44}(z) \end{bmatrix}}_{P(z)} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \mathbf{e}_3(t) \\ \mathbf{e}_4(t) \end{bmatrix}. \quad (5.5)$$

where $P_{ij}(z)$ is an $r_i \times r_j$ block of $P(z)$, $i, j = 1, 2, 3, 4$. The network graph of $P(z)$

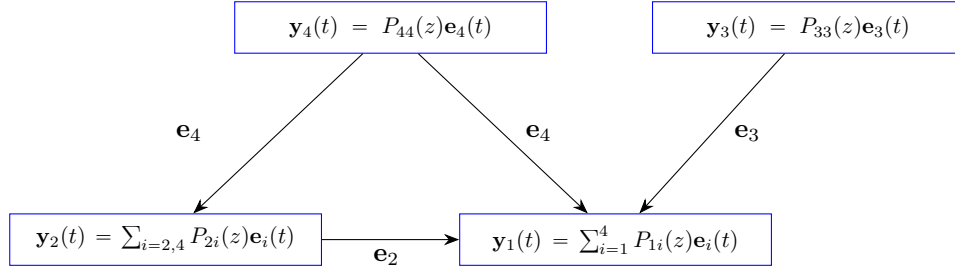


Figure 5.1: Network graph of the innovation transfer matrix (5.5) with G -zero structure

is the network of the subsystems $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \mathcal{S}_4$ defined in (5.4), generating y_1, y_2, y_3 and y_4 , respectively. See Figure 5.1 for illustration of this network graph.

The next theorem is a generalization of Theorem 5.3 for transfer matrices with TADG-zero structure.

Theorem 5.6. Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ and a TADG $G = (V, E)$ where $V = \{1, \dots, n\}$. The innovation transfer matrix $P(z)$ of \mathbf{y} has a G -zero structure if and only if \mathbf{y} has G -consistent causality structure.

The **proof** can be found in Appendix 5.A.

Informally, Theorem 5.6 can be explained as follows: knowing which components help to predict which components in a process \mathbf{y} , is equivalent to knowing how the information flows between \mathbf{y} and its innovation process.

Remark 5.7. It is worth mentioning that we started from a fixed partition of the output process and investigated the conditional non-causality structure between the chosen components. In fact, Theorem 5.6 holds for any partitioning of the output process for which the output has G -consistent causality structure with a TADG G , or equivalently, for which the innovation transfer matrix has G -zero structure. The more detailed partition we choose, the more information we have about the causality structure of the output process.

Theorem 5.6 relies on the property of transfer matrices having TADG-zero structure that they are closed under multiplication and inversion.

Lemma 5.8. Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ and a TADG $G = (V, E), V = \{1, \dots, n\}$. If the innovation transfer matrix $P(z)$ of \mathbf{y} has G -zero structure then the inverse transfer matrix $Q(z) = P^{-1}(z)$ has also G -zero structure.

The **proof** can be found in Appendix 5.A.

Note that if the graph-zero structure is non-TADG of a transfer matrix then in general, the inverse of the transfer matrix does not have the same zero structure as the transfer matrix itself. Moreover, the zero blocks of such transfer matrices do not necessarily define any non-causal relation among the components of \mathbf{y} , even if they are innovation transfer matrices. Next, we will show an example to illustrate the result of Theorem 5.6.

5.3 Example for non-TADG and TADG-zero structures

To illustrate the results of Theorem 5.6 and explain Remark 5.7 in more details, we will now present an example. In Remark 5.7 we mentioned that the components of the process for which we observe the causality relations can be chosen in several ways. Take the simplest non-transitive directed graph $G_1 = (\{1, 2, 3\}, \{(3, 2), (2, 1)\})$. Suppose that the innovation transfer matrix $P(z)$ of a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T$ has G_1 -zero structure as follows:

$$P(z) = \begin{bmatrix} P_{11}(z) & P_{12}(z) & 0 \\ 0 & P_{22}(z) & P_{23}(z) \\ 0 & 0 & P_{33}(z) \end{bmatrix}, \quad (5.6)$$

where $P_{ij}(z)$ is an $r_i \times r_j$ block of $P(z)$, $i, j = 1, 2, 3$. If \mathbf{y}_2 is further decomposed into $\mathbf{y}_2 = [\mathbf{y}_{21}^T, \mathbf{y}_{22}^T]^T$, then it can happen that the transfer matrix of $[\mathbf{y}_1^T, \mathbf{y}_{21}^T, \mathbf{y}_{22}^T, \mathbf{y}_3^T]^T$ has a TADG-zero structure, for example with the graph $G_2 = (\{1, 2, 3, 4\}, (4, 2), (3, 2), (3, 1))$:

$$P(z) = \begin{bmatrix} P_{11}(z) & 0 & P_{13}(z) & 0 \\ 0 & P_{22}(z) & P_{23}(z) & P_{24}(z) \\ 0 & 0 & P_{33}(z) & 0 \\ 0 & 0 & 0 & P_{44}(z) \end{bmatrix}, \quad (5.7)$$

where $P_{ij}(z)$ is a $q_i \times r_j$ block of $P(z)$ for $i, j = 1, 2, 3, 4$ such that $q_1 = r_1$, $q_2 + q_3 = r_2$ and $q_4 = r_4$. Naturally, it can also happen that the transfer matrix of $[\mathbf{y}_1^T, \mathbf{y}_{21}^T, \mathbf{y}_{22}^T, \mathbf{y}_3^T]^T$ has non-TADG structure (there might be for other partitioning of \mathbf{y}), for example with the graph $G_3 = (\{1, 2, 3, 4\}, (4, 2), (2, 3), (3, 1))$:

$$P(z) = \begin{bmatrix} P_{11}(z) & 0 & P_{21}(z) & 0 \\ 0 & P_{22}(z) & 0 & P_{24}(z) \\ 0 & P_{32}(z) & P_{33}(z) & 0 \\ 0 & 0 & 0 & P_{44}(z) \end{bmatrix}, \quad (5.8)$$

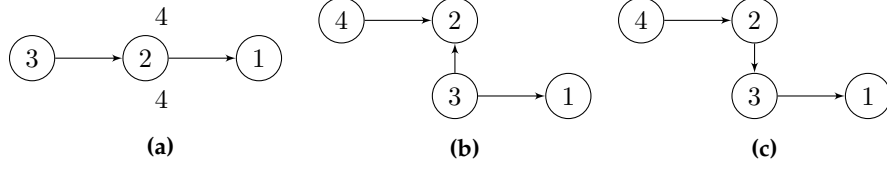


Figure 5.2: TADG (b) and non-TADG (a), (c) zero structures of $P(z)$ for different partitioning of $P(z)$ as in (5.6)–(5.8), respectively

where again $P_{ij}(z)$ is a $q_i \times r_j$ block of $P(z)$. Figure 5.2 illustrates the above-mentioned three cases: when $P(z)$ has G_1 -zero structure, G_2 -zero structure and G_3 -zero structure.

To illustrate the results of Theorem 5.6, we give an example for the case when $P(z)$ is as in (5.7), i.e., when the innovation transfer matrix of \mathbf{y} has TADG-zero structure. We define the innovation transfer matrix from the following Kalman representation of a $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T, \mathbf{y}_4^T]^T$ process:

$$\begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_2(t+1) \\ \mathbf{x}_3(t+1) \\ \mathbf{x}_4(t+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 0.8 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0 \\ 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0.7 \end{bmatrix}}_A \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \mathbf{x}_3(t) \\ \mathbf{x}_4(t) \end{bmatrix} + \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_K \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \mathbf{e}_3(t) \\ \mathbf{e}_4(t) \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \\ \mathbf{y}_3(t) \\ \mathbf{y}_4(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 0.9 & 0 & 0.7 & 0 \\ 0 & 0.5 & 0.4 & 0.3 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.8 \end{bmatrix}}_C \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \mathbf{x}_3(t) \\ \mathbf{x}_4(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \mathbf{e}_3(t) \\ \mathbf{e}_4(t) \end{bmatrix}.$$

The innovation transfer matrix $P(z) = C(Iz - A)^{-1}K + I$ of the process \mathbf{y} is then given by

$$P(z) = \begin{bmatrix} \frac{z+0.1}{z-0.8} & 0 & \frac{0.7}{z-0.7} & 0 \\ 0 & \frac{z-0.1}{z-0.6} & \frac{0.4}{z-0.7} & \frac{0.3}{z-0.7} \\ 0 & 0 & \frac{z-0.2}{z-0.7} & 0 \\ 0 & 0 & 0 & \frac{z+0.1}{z-0.7} \end{bmatrix}.$$

Notice that with the TADG $G = (\{1, 2, 3, 4\}, \{(4, 2), (3, 2), (3, 1)\})$ and the partitioning $[\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T, \mathbf{y}_4^T]^T$ of \mathbf{y} , $P(z)$ has G -zero structure. Note that from Lemma 5.8 we also know that the inverse transfer matrix $Q(z) = P(z)^{-1}$ has also G -zero structure.

Indeed,

$$Q(z) = \begin{bmatrix} \frac{z-0.8}{z+0.1} & 0 & \frac{-0.7z+0.56}{z^2-0.1z-0.02} & 0 \\ 0 & \frac{z-0.6}{z-0.1} & \frac{-0.4z+0.24}{z^2-0.3z+0.02} & \frac{-0.3z+0.18}{z^2-0.01} \\ 0 & 0 & \frac{z-0.7}{z-0.2} & 0 \\ 0 & 0 & 0 & \frac{z-0.7}{z+0.1} \end{bmatrix}.$$

From Theorem 5.6, the G -zero structure of $P(z)$ is equivalent to the following (conditional) Granger causality conditions:

- (i) y_1, y_2, y_3 do not Granger cause y_4
- (ii) y_1 conditionally does not Granger cause y_2 with respect to $[y_3^T, y_4^T]^T$
- (iii) y_1, y_2, y_4 do not Granger cause y_3
- (iv) y_2, y_4 conditionally do not Granger cause y_1 with respect to y_3 .

Remark 5.9. Defining $\mathbf{z}_1 = y_1$, $\mathbf{z}_2 = [y_2^T, y_3^T]^T$ and $\mathbf{z}_3 = y_4$, the process $\mathbf{z} = [\mathbf{z}_1^T, \mathbf{z}_2^T, \mathbf{z}_3^T]^T$ has the transfer matrix as in (5.6). Then, we can deduce the following from conditions (i)–(iv): (1) from (i) it follows that $\mathbf{z}_1, \mathbf{z}_2$ do not Granger cause \mathbf{z}_3 ; (2) from (ii) we can derive that \mathbf{z}_1 conditionally does not Granger cause \mathbf{z}_2 with respect to \mathbf{z}_3 and finally, (3) from (iv) we have that \mathbf{z}_3 conditionally does not Granger cause \mathbf{z}_1 with respect to \mathbf{z}_2 . Note that partitioning \mathbf{z} into two components such that the first one is $[\mathbf{z}_1^T, \mathbf{z}_2^T]^T$ and the second one is \mathbf{z}_3 , Theorem 5.3 implies (1). However, without looking at the more detailed partition of \mathbf{z}_2 , that is y_2 and y_3 , conditions (2) and (3) do not follow from the zero-structure of the 3×3 block transfer matrix in (5.6), see also Remark 5.7.

5.4 Conclusions

In this chapter, we have studied transfer matrices of ZMSIR process with certain network graphs and related them to conditional Granger causality conditions among the components of their output processes. In particular, considering a TADG G , we studied so-called innovation transfer matrices with G -zero structure whose zero blocks are determined by G . We then showed that an innovation transfer matrix has G -zero structure if and only if the output process has G -consistent causality structure.

5.A Proofs

Before the proof of Theorem 5.6 we present the proof of Lemma 5.8.

Proof of Lemma 5.8. Consider two transfer matrices

$$P(z) = \begin{bmatrix} P_{11}(z) & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1}(z) & \cdots & P_{nn} \end{bmatrix}, \quad \tilde{P}(z) = \begin{bmatrix} \tilde{P}_{11}(z) & \cdots & \tilde{P}_{1n} \\ \vdots & \ddots & \vdots \\ \tilde{P}_{n1}(z) & \cdots & \tilde{P}_{nn} \end{bmatrix},$$

where $P_{ij}(z)$ and \tilde{P}_{ij} are $\mathbf{r}_i \times \mathbf{r}_j$ dimensional transfer matrices. We start with proving that if $P(z)$ and $\tilde{P}(z)$ have G -zero structure with a TADG $G = (V, E)$ then $P(z)\tilde{P}(z)$ has also G -zero structure. Consider the ij component of $P(z)\tilde{P}(z)$ given by

$$(P(z)\tilde{P}(z))_{ij} = \sum_{k \in \{l \in V \mid (j,l), (l,i) \in E\}} P_{ik}(z)\tilde{P}_{kj}(z).$$

From transitivity, if $i, j, l \in V$ and $(j, l), (l, i) \in E$ then $(j, i) \in E$. It then follows that if $(j, i) \notin E$ then the set $\{l \in V \mid (j, l), (l, i) \in E\} = \emptyset$ and hence $(P(z)\tilde{P}(z))_{ij} = 0$. Consequently, $P(z)\tilde{P}(z)$ has G -zero structure. Next, we use induction to prove that if $P(z)$ has G -zero structure then any power of it has also G -zero structure. For the special case when $\tilde{P}(z) = P(z)$, the deduction above implies that $P^2(z)$ has G -zero structure. Assume that $P^r(z)$ has G -zero structure for $r = 1, \dots, m$ then it follows that $P^{m+1}(z) = P^m(z)P(z)$ is a product of two transfer matrices with G -zero structure thus if $(j, i) \notin E$ then

$$(P^{m+1}(z))_{ij} = \sum_{k \in \{l \in V \mid (j,l), (l,i) \in E\}} P_{ik}^m(z)P_{kj}(z) = 0.$$

In this way we proved that $P^r(z)$ has G -zero structure for all $r \geq 1$. Since $P^{-1}(z)$ can be expressed as a polynomial of $P(z)$, and taking linear combinations preserves the G -zero structure, the statement follows. \square

Proof of Theorem 5.6. Necessity: We start with proving that G -zero structure of innovation transfer matrices imply G -consistent causality structure in the output process. From Lemma 5.8, we know that the inverse of the innovation transfer matrix $P(z) = \sum_{k=0}^{\infty} P_k z^{-k}$, denoted by $Q(z) = \sum_{k=0}^{\infty} Q_k z^{-k}$ has G -zero structure. Take the decomposition of $P(z)$ and $Q(z)$ as in (5.3). From the G -zero structure of $P(z)$ and $Q(z)$ we obtain that for all $i, j \in V$ if $(i, j) \notin E$ (or equivalently $i \notin I_j$) then $P_{ji}(z) = Q_{ji}(z) = 0$. Therefore, $\mathbf{y}_j(t)$ and $\mathbf{e}_j(t)$, $j \in \{1, \dots, n\}$ can be written as

follows

$$\begin{aligned} \mathbf{y}_j(t) &= \mathbf{e}_j(t) + \sum_{i \in I_j} \underbrace{\sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k)}_{P_{ij}(z) \mathbf{e}_i} \\ \mathbf{e}_j(t) &= \mathbf{y}_j(t) + \sum_{i \in I_j} \underbrace{\sum_{k=1}^{\infty} (Q_k)_{ji} \mathbf{y}_i(t-k)}_{Q_{ij}(z) \mathbf{y}_i}. \end{aligned} \quad (5.9)$$

Note that since \mathbf{e} is the innovation process of \mathbf{y} it follows that $\mathcal{H}_{t-}^{\mathbf{y}} = \mathcal{H}_{t-}^{\mathbf{e}}$. Consider that $\mathbf{e}(t)$ is orthogonal to $\mathcal{H}_{t-}^{\mathbf{y}}$ and $\mathbf{e}(t-k) \in \mathcal{H}_{t-}^{\mathbf{y}}$ for $k > 0$. Then, taking the projection of $\mathbf{y}_j(t)$ onto $\mathcal{H}_{t-}^{\mathbf{y}}$ we obtain that

$$E_l[\mathbf{y}_j(t) | \mathcal{H}_{t-}^{\mathbf{y}}] = \sum_{i \in I_j} \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k) \in \mathcal{H}_{t-}^{\mathbf{e}_{I_j}}.$$

Note that $\mathbf{e}_{I_j}(t) := \mathbf{y}_{I_j}(t) - E_l[\mathbf{y}_{I_j}(t) | \mathcal{H}_{t-}^{\mathbf{y}}]$ is not the innovation process of \mathbf{y}_{I_j} unless the Granger causality condition that $\mathbf{y}_{\{1, \dots, n\} \setminus I_j}$ does not Granger cause \mathbf{y}_{I_j} hold. Notice that from (5.9) we have that $\mathbf{e}_j(t) \in \mathcal{H}_{(t+1)-}^{\mathbf{y}_{I_j}}$. From transitivity, if $i \in I_j$ then $I_i \subseteq I_j$ and thus $\mathbf{e}_i(t-1) \in \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}$ for all $i \in I_j$. This leads to the following:

$$E_l[\mathbf{y}_j(t) | \mathcal{H}_{t-}^{\mathbf{y}}] \in \mathcal{H}_{t-}^{\mathbf{e}_{I_j}} \subseteq \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}.$$

As a consequence, $E_l[\mathbf{y}_j(t) | \mathcal{H}_{t-}^{\mathbf{y}}] = E_l[\mathbf{y}_j(t) | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}]$ for $j \in \{1, \dots, n\}$. Likewise, the k -step prediction of \mathbf{y}_j based on $\mathcal{H}_{t-}^{\mathbf{y}}$ equals

$$E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}}] = \sum_{i \in I_j} \sum_{l=k+1}^{\infty} (P_l)_{ji} \mathbf{e}_i(t+k-l) \in \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}.$$

This implies that $E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}}] = E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}]$, thus that \mathbf{y}_i conditionally does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} for $i \notin I_j$, or equivalently, for $(i, j) \notin E$.

Sufficiency: It remains to show that if $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ has G -consistent causality structure with a TADG $G = (V = \{1, \dots, n\}, E)$ then its innovation transfer matrix has G -zero structure. From the conditions it follows that for any $(i, j) \notin E$ $E_l[\mathbf{y}_j(t) | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}] = E[\mathbf{y}_j(t) | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_i}]$. Since for every i , either $(i, j) \in E$ holds and hence $i \in I_j$, or $(i, j) \notin E$ holds and hence $i \notin I_j$ by transitivity, it follows that $E_l[\mathbf{y}_j(t) | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}] = E_l[\mathbf{y}_j(t) | \mathcal{H}_{t-}^{\mathbf{y}}]$, thus \mathbf{e}_{I_j} is the innovation process of \mathbf{y}_{I_j} . From

$\mathbf{y}_j(t) = \mathbf{e}_j(t) + \sum_{i=1}^n \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k)$ we can deduce that

$$E_l[\mathbf{e}_j(t) + \sum_{i=1}^n \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k) | \mathcal{H}_{t-}^{\mathbf{y}}] = \sum_{i=1}^n \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k) \in \mathcal{H}_{t-}^{\mathbf{y}_{I_j}} = \mathcal{H}_{t-}^{\mathbf{e}_{I_j}}.$$

Note that $\mathcal{H}_{t-}^{\mathbf{y}_{I_j}} = \mathcal{H}_{t-}^{\mathbf{e}_{I_j}}$ holds since \mathbf{e}_{I_j} is the innovation process of \mathbf{y}_{I_j} . Considering that $\sum_{i \in I_j} \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k) \in \mathcal{H}_{t-}^{\mathbf{e}_{I_j}}$ and that

$$\sum_{i \notin I_j} \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k) = \sum_{i=1}^n \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k) - \sum_{i \in I_j} \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k),$$

it follows that $\sum_{i \notin I_j} \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k) \in \mathcal{H}_{t-}^{\mathbf{e}_{I_j}}$. Considering that the one dimensional components of $\{E[\mathbf{e}_{I_j}(t)\mathbf{e}_{I_j}^T(t)]^{-1} \mathbf{e}_{I_j}(t-k)\}_{k=0}^{\infty}$ form an orthonormal basis for $\mathcal{H}_{t-}^{\mathbf{e}_{I_j}}$, there exist $\{\alpha_k\}_{k=0}^{\infty}$ such that

$$\sum_{i \notin I_j} \sum_{k=1}^{\infty} (P_k)_{ji} \mathbf{e}_i(t-k) = \sum_{k=1}^{\infty} \alpha_k \mathbf{e}_{I_j}(t-k). \quad (5.10)$$

Denote the complementary index set of I_j by $\hat{I}_j := \{1, \dots, n\} \setminus I_j$ and for $\hat{I}_j = \{i_1, \dots, i_m\}$ denote the transfer matrix $[(P_k)_{ji_1}, \dots, (P_k)_{ji_m}]$ from $\mathbf{e}_{\hat{I}_j}$ to \mathbf{y}_j by $(P_k)_{j\hat{I}_j}$. Notice that (5.10) can be written as

$$\sum_{k=1}^{\infty} [(P_k)_{j\hat{I}_j}, -\alpha_k] \mathbf{e}(t-k) = 0.$$

If we take the variance of the equation above and consider that \mathbf{e} is a white noise process, we have that

$$\sum_{k=1}^{\infty} [(P_k)_{j\hat{I}_j}, -\alpha_k] E[\mathbf{e}(t)\mathbf{e}^T(t)] [(P_k)_{j\hat{I}_j}, -\alpha_k]^T = 0.$$

Since \mathbf{y} is a weakly stationary full rank process, the variance of its innovation process at any time $t \in \mathbb{Z}$, $E[\mathbf{e}(t)\mathbf{e}^T(t)]$ is strictly positive definite. It then follows that $\alpha_k = 0$ and $(P_k)_{j\hat{I}_j} = 0$ for all $k \geq 0$. Therefore, for any $i \in \hat{I}_j$, or equivalently for any $i \in V$ such that $(i, j) \notin E$, $(P_k)_{ji} = 0$ for all $k \geq 0$ which, by definition, means that $P(z)$ has G -zero structure. \square