

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 4

Granger causality and Kalman representations with transitive acyclic directed graph zero structure

We have seen in Chapters 2 and 3 that the existence of Kalman representations whose network graphs are star graphs are equivalent to the lack of conditional and unconditional Granger causalities in the output process of those representations. In this chapter we present a generalization of these results by introducing Kalman representations with network graphs that are transitive acyclic directed graphs (TADG). We call these representations Kalman representations with TADG-zero structure. The existence of Kalman representations with TADG-zero structure is then associated to the lack of conditional and unconditional Granger non-causalities in the output process. These causalities are consistent with the network graph of the representation. We also present algorithms for constructing a Kalman representation with TADG-zero structure in the presence of the appropriate conditional and unconditional Granger non-causalities.

The paper (Caines and Wynn, 2007) is the closest one to the results in this chapter. The cited paper studies LTI-SS representations of Gaussian processes in a form that is a subclass of the Kalman representations with TADG-zero structure with additional assumptions on the covariance matrix of the noise process. The existence of these LTI-SS representations are formalized by conditional orthogonal conditions which are stronger than the conditional orthogonality condition that are counterparts of the conditional and unconditional Granger causalities proposed in this chapter. Note that in (Caines and Wynn, 2007) there are no detailed proofs or algorithms to calculate the representations. Furthermore, it does not deal with non-coercive or non-Gaussian processes. The results of this chapter are based on the conference papers (Jozsa et al., 2017a). However, several additional statements are presented here that were not in the cited paper.

This chapter is organized as follows: First, we introduce Kalman representations with TADG-zero structure. Then, we characterize their existence in terms of conditional and unconditional Granger causality. Next, the construction for calculating Kalman representations with TADG-zero structure and the corresponding al-

gorithms are presented. Finally, we provide an example to illustrate the results. The proofs of the statements can be found in Appendices 4.A, 4.B, and 4.C. If not stated otherwise, we assume throughout this chapter that $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ is a ZMSIR process where $n \geq 2$, $\mathbf{y}_i \in \mathbb{R}^{r_i}$, and $r_i > 0$ for $i = 1, \dots, n$.

4.1 Kalman representation with TADG-zero structure

In this section, we introduce Kalman representations whose network graph is a transitive acyclic directed graph (TADG) and discuss their properties. To begin with, we define the class of transitive acyclic graphs.

Definition 4.1 (TADG). A directed graph $G = (V, E)$, with set of nodes $V = \{1, \dots, k\}$ and set of directed edges $E \subseteq V \times V$ is called *acyclic* if there is no cycle i.e., closed directed path. Furthermore, it is *transitive* if for $i, j, l \in V$ the implication $(i, j), (j, l) \in E \implies (i, l) \in E$ holds. The class of transitive acyclic directed graphs is denoted by TADG.

For convenience we make the following assumption that applies for all ZMSIR processes throughout this chapter.

Assumption 4.2. For a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$, we assume that none of the components of \mathbf{y} is a white noise process, or equivalently, the dimension of a minimal Kalman representation of \mathbf{y}_i is strictly positive for all $i \in \{1, \dots, n\}$.

For a TADG $G = (V = \{1, \dots, n\}, E)$, the set of nodes V has a so-called *topological ordering*. By topological ordering we mean an ordering on V such that if $(i, j) \in E$ is a directed edge then $i > j$. Throughout this chapter we use integers to represent nodes of graphs and, without the loss of generality, we assume the following:

Assumption 4.3. Consider a TADG $G = (V, E)$ where $V = \{1, \dots, n\}$. Then $(i, j) \in E$ implies $i > j$.

Remark 4.4. Let $G = (V = \{a_1, \dots, a_n\}, E)$ be a TADG, then we can generate topological ordering on G as follows: Assume that the leaves of G are $(a_{i_1}, \dots, a_{i_{k_1}})$ where $k_1 > 1$, $i_j \in \{1, \dots, n\}$ for $j = 1, \dots, k_1$ and the leaves are enumerated in an arbitrary order. Then, delete the leaves of G and all the directed edges whose target node is a leaf. Call the new graph G_1 . Assume now that the leaves of G_1 are $(a_{i_{k_1+1}}, \dots, a_{i_{k_2}})$ where $k_2 > k_1$, $i_j \in \{1, \dots, n\}$ for $j = k_1 + 1, \dots, k_2$ and the leaves are again enumerated in an arbitrary order. Then, delete the leaves of G_1 and all the directed edges whose target node is a leaf of G_1 . Continue this until each of the nodes of G are enumerated. The new graph $\tilde{G} = (\{1, \dots, n\}, \tilde{E})$, where $(k, l) \in \tilde{E}$ if and only if $(a_{i_k}, a_{i_l}) \in E$ is isomorphic with G and has topological ordering.

The class of transitive acyclic directed graphs will be used to represent internal interconnection structure of Kalman representations. We will say that a Kalman representation has TADG-zero structure if its network graph is a TADG. To define this class of Kalman representations, we need to introduce some new terminology.

Notation 4.5 (parent and non-parent succeeding nodes). Let $G = (V = \{1, \dots, n\}, E)$ be a TADG and consider a node $j \in V$. The set of parent nodes $\{i \in V | (i, j) \in E\}$ of j is denoted by I_j . In addition, the set of non-parent succeeding (with respect to the topological ordering of V) nodes $\{i \in V | i > j, (i, j) \notin E\}$ of j is denoted by \bar{I}_j .

The topological ordering on the set of nodes of a TADG graph implies that $I_j, \bar{I}_j \subseteq \{j + 1, \dots, n\}$ for all $j \in \{1, \dots, n - 1\}$. Furthermore, from the definition of \bar{I}_j , we have that $I_j \cup \bar{I}_j = \{j + 1, \dots, n\}$. The next notation helps in referring to components of processes beyond the original partitioning of those processes.

Notation 4.6 (sub-process). Consider the finite set $V = \{1, \dots, n\}$ and a tuple $J = (j_1, \dots, j_l)$ where $j_1, \dots, j_l \in V$. Then for a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$, we denote the sub-process $[\mathbf{y}_{j_1}^T, \dots, \mathbf{y}_{j_l}^T]^T$ by $\mathbf{y}_{j_1, \dots, j_l}$ or by \mathbf{y}_J . By abuse of terminology, if J is a subset of V and not a tuple, then \mathbf{y}_J will mean process \mathbf{y}_α , where α is the tuple obtained by taking the elements of J in increasing order, i.e. if $J = \{j_1, \dots, j_k\}$, $j_1 < j_2 < \dots < j_k$, then $\alpha = (j_1, \dots, j_k)$. However, $\mathbf{y}_{\alpha, \beta}$ always means $[\mathbf{y}_\alpha^T, \mathbf{y}_\beta^T]^T$ regardless the topological order between the elements of α and β .

Next, we introduce what we mean by partition of matrices. Call the set $\{p_i, q_i\}_{i=1}^k$ a *partition* of (p, q) , where $p, q > 0$, if $\sum_{i=1}^k p_i = p$ and $\sum_{i=1}^k q_i = q$, where $p_i, q_i > 0$ for $i = 1, \dots, k$.

Definition 4.7 (partition of a matrix). Let $\{p_i, q_i\}_{i=1}^k$ be a partition of (p, q) for some $p, q > 0$. Then *the partition of a matrix* $M \in \mathbb{R}^{p \times q}$ with respect to $\{p_i, q_i\}_{i=1}^k$ is a collection of matrices $\{M_{ij} \in \mathbb{R}^{p_i \times q_j}\}_{i,j=1}^k$, such that

$$M = \begin{bmatrix} M_{11} & \cdots & M_{1k} \\ \vdots & \ddots & \vdots \\ M_{k1} & \cdots & M_{kk} \end{bmatrix}.$$

In Definition 4.7, the indexing of matrix M refers to the blocks of M and does not refer directly to the elements of M . It is parallel to the component-wise indexing of processes where the components can be multidimensional.

Notation 4.8 (sub-matrix). Consider the partition $\{M_{ij} \in \mathbb{R}^{p_i \times q_j}\}_{i,j=1}^k$ of a matrix $M \in \mathbb{R}^{p \times q}$ with respect to the partition $\{p_i, q_i\}_{i=1}^k$ of (p, q) . Furthermore, consider the

tuples $I = (i_1, \dots, i_n)$ and $J = (j_1, \dots, j_m)$ where $i_1, \dots, i_n, j_1, \dots, j_m \in \{1, \dots, k\}$. Then by the sub-matrix of M indexed by IJ we mean

$$M_{IJ} := \begin{bmatrix} M_{i_1 j_1} & \cdots & M_{i_1 j_m} \\ \vdots & \ddots & \vdots \\ M_{i_n j_1} & \cdots & M_{i_n j_m} \end{bmatrix}$$

We are now ready to defined Kalman representations which have a so-called TADG-zero structure:

Definition 4.9 (*G-zero structure*). Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ and a TADG $G = (V = \{1, \dots, n\}, E)$. Let (A, K, C, I, \mathbf{e}) be a p -dimensional Kalman representation of $\mathbf{y} \in \mathbb{R}^r$ and partition A with respect to $\{p_i, p_i\}_{i=1}^n$, K with respect to $\{p_i, r_i\}_{i=1}^n$ and C with respect to $\{r_i, p_i\}_{i=1}^n$ where $\{p_i, r_i\}_{i=1}^n$ is a partition of (p, r) . Then we say that (A, K, C, I, \mathbf{e}) has *G-zero structure* if $A_{ij} = 0$, $K_{ij} = 0$, $C_{ij} = 0$ whenever $(j, i) \notin E$. If, in addition, for all $j \in V$ the tuple $J := (j, \bar{I}_j, I_j)$ defines a Kalman representation $(A_{JJ}, K_{JJ}, C_{JJ}, I, [\mathbf{e}_j^T, \mathbf{e}_{\bar{I}_j}^T, \mathbf{e}_{I_j}^T]^T)$ of $[\mathbf{y}_j^T, \mathbf{y}_{\bar{I}_j}^T, \mathbf{y}_{I_j}^T]^T$ in causal coordinated form (see Definition 3.1), then we say that (A, K, C, I, \mathbf{e}) has *causal G-zero structure*.

Besides saying that a representation has *G-zero structure* or *causal G-zero structure*, we also say, *representation with G-zero structure* or *with causal G-zero structure*.

Consider the TADGs $G_1 = (\{1, 2\}, \{(2, 1)\})$ and $G_2 = (\{1, 2, \dots, n\}, \{(n, 1), (n, 2), \dots, (n, n-1)\})$. If the graph G in Definition 4.9 is G_1 , then Definition 4.9 coincides with Definition 2.1 in Section 2.1 considering ZMSIR processes that satisfy Assumption 4.2 (see Remark 2.2 in Section 2.1). In a similar manner, if the graph G in Definition 4.9 is G_2 then it coincides with Definition 3.1 in Section 3.1 considering ZMSIR processes that satisfy Assumption 4.2.

If a p -dimensional Kalman representation (A, K, C, I, \mathbf{e}) of $\mathbf{y} \in \mathbb{R}^r$ has causal *G-zero structure*, where $G = (V, E)$ is a TADG, then the partition $\{p_i, r_i\}_{i=1}^n$ of (p, r) in Definition 4.9 is uniquely determined by \mathbf{y} . It is equivalent of saying that the block dimensions of the partitioned matrices A, K and C are uniquely determined by \mathbf{y} . Indeed, for all nodes $j \in V$ the tuple $J := (j, \bar{I}_j, I_j)$ defines a Kalman representation $(A_{JJ}, K_{JJ}, C_{JJ}, I, [\mathbf{e}_j^T, \mathbf{e}_{\bar{I}_j}^T, \mathbf{e}_{I_j}^T]^T)$ of $[\mathbf{y}_j^T, \mathbf{y}_{\bar{I}_j}^T, \mathbf{y}_{I_j}^T]^T$ in causal coordinated form. Therefore, from Chapter 3 we know that the dimensions of (A_{lk}, K_{lk}, C_{lk}) for $k, l \in \{j, \bar{I}_j, I_j\}$ are uniquely determined by $[\mathbf{y}_j^T, \mathbf{y}_{\bar{I}_j}^T, \mathbf{y}_{I_j}^T]^T$. Then, using this for node $j = n-1, \dots, 1$ it is easy to see that all block dimensions of the partitioned matrices A, K and C are determined by \mathbf{y} .

A Kalman representation with TADG-zero structure can be viewed as consisting of subsystems where each subsystem generates a component of $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$.

More precisely, let $G = (V = \{1, \dots, n\}, E)$ be a TADG and $(A, K, C, I, \mathbf{e}, \mathbf{y})$ be a p -dimensional Kalman representation with G -zero structure where A, K and C are partitioned with respect to a partition $\{p_i, q_i\}_{i=1}^k$ of (p, q) . Furthermore, let $\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_n^T]^T$ be its state such that $\mathbf{x}_i \in \mathbb{R}^{p_i}$, $i = 1, \dots, n$. Then the representation with output \mathbf{y}_j , $j \in V$ is in the form of

$$\mathcal{S}_j \begin{cases} \mathbf{x}_j(t+1) = A_{jj}\mathbf{x}_j(t) + (A_{jI_j}\mathbf{x}_{I_j}(t) + K_{jI_j}\mathbf{e}_{I_j}(t)) + K_{jj}\mathbf{e}_j(t) \\ \mathbf{y}_j(t) = C_{jj}\mathbf{x}_j(t) + C_{jI_j}\mathbf{x}_{I_j}(t) + \mathbf{e}_j(t). \end{cases} \quad (4.1)$$

Notice that if $(i, j) \in E$, i.e., i is a parent node of j , then subsystem \mathcal{S}_j takes inputs from subsystem \mathcal{S}_i , namely the state and noise processes of \mathcal{S}_i . In contrast, if $(j, i) \notin E$, \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 directed paths of the graph G . Note that from transitivity, if there is a directed path from $i \in V$ to $j \in V$ then there is also an edge $(i, j) \in E$.

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

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_2(t+1) \\ \mathbf{x}_3(t+1) \\ \mathbf{x}_4(t+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ 0 & A_{22} & 0 & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & 0 & A_{44} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \mathbf{x}_3(t) \\ \mathbf{x}_4(t) \end{bmatrix} + \begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} \\ 0 & K_{22} & 0 & K_{24} \\ 0 & 0 & K_{33} & 0 \\ 0 & 0 & 0 & K_{44} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \mathbf{e}_3(t) \\ \mathbf{e}_4(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \\ \mathbf{y}_3(t) \\ \mathbf{y}_4(t) \end{bmatrix} &= \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ 0 & C_{22} & 0 & C_{24} \\ 0 & 0 & C_{33} & 0 \\ 0 & 0 & 0 & C_{44} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \mathbf{x}_3(t) \\ \mathbf{x}_4(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \mathbf{e}_3(t) \\ \mathbf{e}_4(t) \end{bmatrix}, \end{aligned} \quad (4.2)$$

where $A_{ij} \in \mathbb{R}^{p_i \times p_j}$, $K_{ij} \in \mathbb{R}^{p_i \times r_j}$, $C_{ij} \in \mathbb{R}^{r_i \times p_j}$ and $\mathbf{y}_i, \mathbf{e}_i \in \mathbb{R}^{r_i}$, $\mathbf{x}_i \in \mathbb{R}^{p_i}$ for some $p_i > 0$, $i, j = 1, 2, 3, 4$. The network graph of this representation is the network of the representations $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \mathcal{S}_4$ defined in (4.1), generating $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ and \mathbf{y}_4 , respectively. See Figure 4.1 for illustration of this network graph.

Motivation for Kalman representations with causal TADG-zero structure

If we consider a general LTI-SS representation of a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ with a TADG $G = (V = \{1, \dots, n\}, E)$ network graph, then the noise process could be any process. Such as for LTI-SS representations in coordinated form in Chapter 3, if the noise process were not the innovation process of \mathbf{y} , then it could happen that information flows through it in an implicit way that is not allowed by the directed

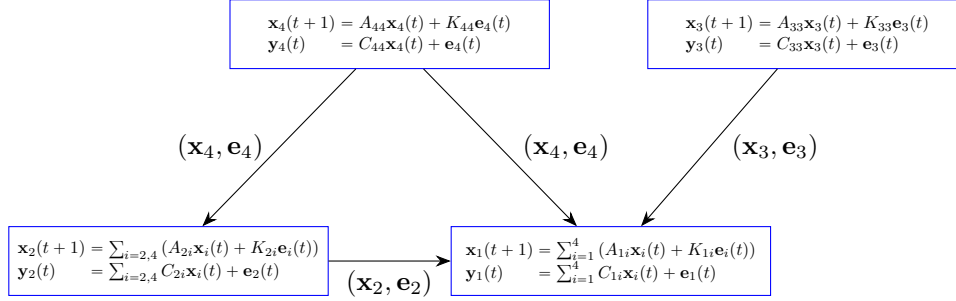


Figure 4.1: Network graph of the Kalman representation (4.2) with G -zero structure

paths (edges) of G . However, if we assume that $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is a Kalman representation with causal G -zero structure, then $[\mathbf{e}_j^T, \mathbf{e}_{I_j}^T]^T$ is the innovation process of $[\mathbf{y}_j^T, \mathbf{y}_{I_j}^T]^T$ and \mathbf{e}_{I_j} is the innovation process of \mathbf{y}_{I_j} for $j = 1, \dots, n$. Hence, the present value of \mathbf{e}_j depends only on the past and present values of $\mathbf{y}_j, \mathbf{y}_{I_j}$, whereas the present value of \mathbf{e}_{I_j} depends only on the past and present values of \mathbf{y}_{I_j} . Moreover, \mathbf{x}_{I_j} depends only on the past values of \mathbf{y}_{I_j} and \mathbf{x}_j depends only on the past values of $\mathbf{y}_j, \mathbf{y}_{I_j}$. That is, in case of Kalman representations with causal G -zero structure, information only flows from subsystems \mathcal{S}_{I_j} , generating \mathbf{y}_{I_j} , to the subsystem \mathcal{S}_j , generating \mathbf{y}_j , see (4.1). That is, the information flows according to the directed paths (edges) of G .

Kalman representations with causal TADG-zero structure have a number of desirable properties, e.g., as it is explained above, the block dimensions of the system matrices are determined by \mathbf{y} . Furthermore, in order to estimate a state \mathbf{x}_j , using a Kalman filter, only the output $[\mathbf{y}_j^T, \mathbf{y}_{I_j}^T]^T$ is necessary (if j is a root node of the TADG then only \mathbf{y}_j is necessary). Moreover, from Lemma 4.10 below, Kalman representations with causal G -zero structure are isomorphic (see Definition 1.11). Hence, if they represent the same output process, their properties are essentially the same. Note that as a consequence of Lemma 4.10, if a Kalman representation of a process \mathbf{y} with TADG-zero structure is not minimal then there does not exist a minimal Kalman representation of \mathbf{y} with TADG-zero structure.

Lemma 4.10. Consider a TADG $G = (V = \{1, \dots, n\}, E)$ and a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$. Then any two Kalman representations of \mathbf{y} with causal G -zero structure are isomorphic.

The **proof** can be found in Appendix 4.A.

4.2 Granger causality and Kalman representation with TADG-zero structure

Kalman representations of a process \mathbf{y} with causal TADG-zero structure determine causal relationships among the components of \mathbf{y} . In fact, we will show that the existence of a Kalman representation of \mathbf{y} with causal TADG-zero structure can be characterized by conditional Granger non-causalities among the components of \mathbf{y} .

To begin with, we define G -consistent causality structure in a process which involves a combination of conditional Granger non-causality conditions between the components of a process.

Definition 4.11 (G -consistent causality structure). Consider a TADG $G = (V, E)$, where $V = \{1, \dots, n\}$ and a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$. We say that \mathbf{y} has G -consistent causality structure if \mathbf{y}_i conditionally does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} for any $i, j \in V, i \neq j$ such that $(i, j) \notin E$.

If $G = (\{1, 2\}, \{(2, 1)\})$, then Definition 4.11 coincides with Definition 2.3. Furthermore, if $G = (\{1, 2, \dots, n\}, \{(n, 1), (n, 2), \dots, (n, n-1)\})$ then Definition 4.11 coincides with Definition 3.3.

Remark 4.12. Notice that if \mathbf{y}_i is a root node in the TADG graph, i.e., $I_j = \emptyset$ then none of the other components causes \mathbf{y}_i . In this case, the conditional Granger non-causality that for $(j, i) \notin E$ the process \mathbf{y}_j conditionally does not Granger cause \mathbf{y}_i with respect to \mathbf{y}_{I_i} simplifies to that \mathbf{y}_j does not Granger cause \mathbf{y}_i .

Lemma 4.13 below provides an equivalent reformulation of Definition 4.11.

Lemma 4.13. Consider a TADG $G = (V, E)$, where $V = \{1, \dots, n\}$ and a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$. Then \mathbf{y} has G -consistent causality structure if and only if

- \mathbf{y}_j does not Granger cause \mathbf{y}_{I_j}
- $\mathbf{y}_{\bar{I}_j}$ does not Granger cause \mathbf{y}_{I_j}
- \mathbf{y}_j does not Granger cause $[\mathbf{y}_{I_j}, \mathbf{y}_{\bar{I}_j}]$
- $\mathbf{y}_{\bar{I}_j}$ does not Granger cause $[\mathbf{y}_{I_j}, \mathbf{y}_j]$

for all nodes $j \in V$ of G .

The main result of this chapter includes a condition for the existence of minimal Kalman representations with G -zero structure. For this, we recall Definition 3.4, the definition of conditionally trivial intersection of two subspaces $U, V \subseteq \mathcal{H}$ in a Hilbert space \mathcal{H} with respect to a closed subspace $W \subseteq \mathcal{H}$.

Definition 4.14 (conditionally trivial intersection). Consider the subspaces $U, V, W \subseteq \mathcal{H}$ such that W is closed. Then U, V have a *conditionally trivial intersection* with respect to W denoted by $U \cap V|W = \{0\}$ if

$$\{u - E_l[u|W] \mid u \in U\} \cap \{v - E_l[v|W] \mid v \in V\} = \{0\},$$

i.e., the intersection of the projections of U and V onto the orthogonal complement of W in \mathcal{H} is the zero subspace.

Now we are ready to state the main result of this chapter:

Theorem 4.15. Consider the following statements for a TADG $G = (V = \{1, \dots, n\}, E)$ and a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$:

- (i) \mathbf{y} has G -consistent causality structure;
- (ii) (i) holds and for any node $j \in V$ in G

$$E_l[\mathcal{H}_{t+}^{\mathbf{y}_j} | \mathcal{H}_{t-}^{\mathbf{y}_j, \mathbf{y}_{I_j}}] \cap E_l[\mathcal{H}_{t+}^{\mathbf{y}_{I_j}} | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_{I_j}}] | E_l[\mathcal{H}_{t+}^{\mathbf{y}_{I_j}} | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}] = \{0\} \quad (4.3)$$

- (iii) there exists a minimal Kalman representation of \mathbf{y} with causal G -zero structure;
- (iv) there exists a Kalman representation of \mathbf{y} with causal G -zero structure;
- (v) there exists a Kalman representation of \mathbf{y} with G -zero structure;

Then, the following hold:

- (a) (ii) \iff (iii);
- (b) (i) \implies (v);
- (c) (iv) \implies (i).

If, in addition, \mathbf{y} is coercive, then we have

- (d) (i) \iff (iv) \iff (v).

The **proof** can be found in Appendix 4.C.

The intuition behind Theorem 4.15 is the following. If the information flows among subsystems $\{\mathcal{S}_i\}_{i=1}^n$ (see (4.1)) according to the topology of a TADG $G = (V = \{1, \dots, n\}, E)$, then the outputs of subsystems that are not connected by a directed path (edge) in G should not influence each other. For instance, there is no edge from a child node to its parent nodes, which implies that \mathbf{y}_j should not Granger cause \mathbf{y}_{I_j} . From the topological ordering of the nodes, it also follows that

$(i, k) \notin E$ if $i \in \bar{I}_j$ and $k \in I_j$, thus $\mathbf{y}_{\bar{I}_j}$ should not Granger cause \mathbf{y}_{I_j} . In addition, the succeeding non-parent nodes of a node j are disconnected from j , i.e., there is no edge from j to \bar{I}_j or from \bar{I}_j to j , they only can have some common parent nodes in I_j . In a similar manner to Chapter 3, it implies that $\mathbf{y}_{\bar{I}_j}$ and \mathbf{y}_j conditionally does not Granger cause each other with respect to \mathbf{y}_{I_j} . With the help of Lemmas 4.13 and 4.32 in Appendix 4.A, it can be seen that the discussion above supports the statement that a Kalman representation with causal G -zero structure implies G -consistent causality structure in its output process, see the proof of Theorem 4.15 for more details. The implication that a process with G -consistent causality structure always has a Kalman representation with G -zero structure is more involved. It is based on the construction of the Kalman representation with G -zero structure, see Section 4.3 and the proof of Theorem 4.15.

Condition (ii) for minimality in Theorem 4.15 can be explained as follows. It can be shown that a Kalman representation with causal G -zero structure is observable, so for minimality, we only have to ensure its reachability. This is equivalent to the components of the state, denoted by $\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_n^T]^T$, being linearly independent at each time. In a Kalman representation with causal G -zero structure, the space generated by the components of $\mathbf{x}_j(t)$ for a node $j \in V$ is $E_l[\mathcal{H}_{t+}^{\mathbf{y}_j} | \mathcal{H}_{t-}^{\mathbf{y}_j, \mathbf{y}_{I_j}}]$, the space generated by the component of $\mathbf{x}_{\bar{I}_j}(t)$ is $E_l[\mathcal{H}_{t+}^{\mathbf{y}_{\bar{I}_j}} | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_{\bar{I}_j}}]$ and the space generated by the components of $\mathbf{x}_{I_j}(t)$ is $E_l[\mathcal{H}_{t+}^{\mathbf{y}_{I_j}} | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}]$. It then can be seen that condition (ii) is a necessary and sufficient condition for the state \mathbf{x} to be linearly independent in a Kalman representation with causal G -zero structure (for more details see the proof of Theorem 4.15 in Appendix 4.C).

Using minimal Kalman representations with causal G -zero structure is desirable since they are isomorphic to any other minimal Kalman representation of the same process (see Proposition 1.12). Hence, any property derived for a minimal Kalman representation with causal G -zero structure, if it is invariant under isomorphism, remains valid for any other minimal Kalman representation of the same process. Theorem 4.15 gives a necessary and sufficient condition for existence of minimal Kalman representations with causal G -zero structure. Finding conditions on the output process that ensure existence of a minimal Kalman representation with (non-causal) G -zero structure is an open problem to the best of the author's knowledge.

4.3 Computing Kalman representations with TADG-zero structure

Assuming that a ZMSIR process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ has G -consistent causality structure for a TADG $G = (V = \{1, \dots, n\}, E)$, a Kalman representation of \mathbf{y} with G -

zero structure can be calculated algorithmically. In this section, we formulate two algorithms for this purpose: The first algorithm, Algorithm 11, takes the second order statistics of \mathbf{y} as input and calculates a Kalman representation of \mathbf{y} with G -zero structure. The second algorithm, Algorithm 12, calculates the same representation but takes an arbitrary LTI-SS representation of \mathbf{y} as its input.

In the rest of this chapter, we will use the following notation:

Notation 4.16. The restriction of a TADG $G = (V = \{1, \dots, k\}, E)$ to $I = \{i_1, \dots, i_p\} \subseteq V$ is the graph defined by $G|_I := (\{i_1, \dots, i_p\}, \{(i, j) \in E | i, j \in I\})$.

Remark 4.17. The restriction of a TADG to any subset of nodes is a TADG.

Consider a TADG $G = (V = \{1, \dots, n\}, E)$ and a process $\mathbf{y} = [\mathbf{y}_1^T \dots, \mathbf{y}_n^T]^T$ and recall that we assumed topological ordering on V , see Assumption 4.3. Then, the main idea of the procedure that calculates a Kalman representation of \mathbf{y} with G -zero structure is as follows: first, we take a minimal Kalman representation \mathcal{S}_0 of \mathbf{y}_n . Second, \mathcal{S}_0 is extended to a Kalman representation \mathcal{S}_1 of $[\mathbf{y}_{n-1}^T, \mathbf{y}_n^T]^T$ with $G|_{\{n-1, n\}}$ -zero structure. That is, if $(n, n-1) \in E$ then \mathcal{S}_1 is in block triangular form, otherwise it is in block diagonal form, i.e., the system matrices are block diagonal, see Lemma 4.23 below. Then, we continue it as follows: for $i = 1, \dots, n-2$ \mathcal{S}_i is extended to a Kalman representation \mathcal{S}_{i+1} of $[\mathbf{y}_{n-i}^T, \dots, \mathbf{y}_n^T]^T$ with $G|_{\{n-i, \dots, n\}}$ -zero structure. In general, the extension of \mathcal{S}_i can happen in three different ways depending on the edges of G : In the first case, when \bar{I}_{n-i} is empty, the extended representation has block triangular form; in the second case, when I_{n-i} is empty, it has block diagonal form and in the third case, when neither \bar{I}_{n-i} nor I_{n-i} is empty, then it has coordinated form. Note that $\bar{I}_{n-i} \cup I_{n-i} = \{n-i+1, \dots, n\}$, i.e., \bar{I}_{n-i} and I_{n-i} cannot be empty at the same time for $i = 1, \dots, n-2$. To ease the formulation of the procedure described above, we introduce some auxiliary results and algorithms on the above-mentioned extensions of Kalman representations.

4.3.1 Auxiliary results

To extend a Kalman representation of \mathbf{y}_2 to a Kalman representation of $[\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ in block triangular form, we will use the following definition:

Definition 4.18. Consider an observable Kalman representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ of \mathbf{y}_2 . An extension of $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in block triangular form is an ob-

servable Kalman representation of \mathbf{y} of the form

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_2(t+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} K_{11} & K_{12} \\ 0 & K_{22} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \end{bmatrix} &= \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix}. \end{aligned} \quad (4.4)$$

Next, we present Lemma 4.19, Algorithm 8 and Corollary 4.20 below on extensions of Kalman representations in block triangular form.

Lemma 4.19. *Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ and an observable Kalman representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ of \mathbf{y}_2 . If \mathbf{y}_1 does not Granger cause \mathbf{y}_2 then there exists an extension of $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in block triangular form. Moreover, if the representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ is minimal, then there exists an extension of $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in causal block triangular form which is a minimal Kalman representation of \mathbf{y} .*

The **proof** can be found in Appendix 4.B. Note that Lemma 4.19 can be seen as a consequence of Theorem 2.5. If $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ is minimal, then by the isomorphism between minimal Kalman representations, the minimal Kalman representation of \mathbf{y}_2 in Theorem 2.5 and in Lemma 4.19 are isomorphic. Using this isomorphism, the minimal Kalman representations of \mathbf{y} in Theorem 2.5 and in Lemma 4.19 are isomorphic as well. If $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ is not minimal, then the representation of \mathbf{y}_2 in Theorem 2.5 can be transformed to the representation of \mathbf{y}_2 in Lemma 4.19 with a non-singular state-space transformation. Also, the representation of \mathbf{y} in Theorem 2.5 can be transformed to an observable Kalman representation of \mathbf{y} that is the extension of the representation of \mathbf{y}_2 in Lemma 4.19. Therefore, Theorem 2.5 ensures the existence of the Kalman representations of \mathbf{y} in Lemma 4.19 provided that \mathbf{y}_1 does not Granger cause \mathbf{y}_2 .

The representation in Lemma 4.19 can be calculated using Algorithm 5. This is elaborated in Algorithm 8 below. Recall that for a tuple (A, C) of two matrices $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{m \times n}$, the finite observability matrix up to $N > 0$ is defined by

$$O_N = [C^T \ (CA)^T \ \dots \ (CA^{N-1})^T]^T. \quad (4.5)$$

Consider a ZMSIR process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ with covariance sequence $\{\Lambda_k^{\mathbf{y}}\}_{k=0}^{\infty}$. Let \mathbf{e} be the innovation process of \mathbf{y} and N be any number larger than or equal to the dimension of a minimal Kalman representation of \mathbf{y} . Assume that \mathbf{y}_1 does not Granger cause \mathbf{y}_2 and note that Algorithm 5 calculates a minimal Kalman representation in causal block triangular form (Remark 2.7). Apply Algorithm 8 with

Algorithm 8 Extension of an observable Kalman representation in block triangular form

Input $\{A_{22}, K_{22}, C_{22}\}$ and $\{\Lambda_k^y\}_{k=0}^{2N}$: System matrices of an observable Kalman representation of \mathbf{y}_2 and covariance sequence of $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$
--

Output $\{A, K, C\}$: System matrices of (4.4)
--

Step 1 Apply Algorithm 5 with input $\{\Lambda_k^y\}_{k=0}^{2N}$ and denote its output by $\{\hat{A}, \hat{K}, \hat{C}\}$, where

$$\hat{A} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{bmatrix}, \hat{K} = \begin{bmatrix} \hat{K}_{11} & \hat{K}_{12} \\ 0 & \hat{K}_{22} \end{bmatrix}, \hat{C} = \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12} \\ 0 & \hat{C}_{22} \end{bmatrix}.$$

Step 2 Define $T = \hat{O}_N^+ \mathcal{O}_N$ where \hat{O}_N^+ is the left inverse of the finite (up to N) observability matrix of $(\hat{A}_{22}, \hat{C}_{22})$ and \mathcal{O}_N is the finite (up to N) observability matrix of (A_{22}, C_{22}) .

Step 3 Define the following matrices

$$A = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12}T \\ 0 & A_{22} \end{bmatrix}, K = \begin{bmatrix} \hat{K}_{11} & \hat{K}_{12} \\ 0 & K_{22} \end{bmatrix}, C = \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12}T \\ 0 & C_{22} \end{bmatrix}.$$

input $\{A_{22}, K_{22}, C_{22}\}$ and $\{\Lambda_k^y\}_{k=0}^{2N}$, where $\{A_{22}, K_{22}, C_{22}\}$ are system matrices of an observable Kalman representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ of \mathbf{y}_2 . Denote the output by $\{A, K, C\}$. Then we have the following result:

Corollary 4.20 (Correctness of Algorithm 8). *The tuple $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is an observable Kalman representation and it is an extension of $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in block triangular form. Furthermore, if $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ is minimal then $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is a minimal Kalman representation in causal block triangular form.*

The **proof** can be found in Appendix 4.B.

Remark 4.21. Similar to Algorithm 8, Algorithms 5 and 4 in Chapter 2 also calculate Kalman representations in block triangular form provided that \mathbf{y}_1 does not Granger cause \mathbf{y}_2 . However, Algorithm 8 takes, besides the covariances of \mathbf{y} , the system matrices of a Kalman representation of \mathbf{y}_2 as its input and extends it in such a way that the input Kalman representation of \mathbf{y}_2 is a sub-system of the Kalman representation of \mathbf{y} that the output matrices of Algorithm 8 define. As a consequence, contrary to Algorithms 5 and 4, the Kalman representation that Algorithm 8 defines is not necessarily minimal or is in a causal block triangular form.

To extend a Kalman representation of \mathbf{y}_2 to a Kalman representation of $[\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ in block diagonal form, we will use the following definition:

Definition 4.22. Consider an observable Kalman representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ of \mathbf{y}_2 . An extension of $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in block diagonal form is an observable Kalman representation of \mathbf{y} of the form

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_2(t+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} K_{11} & 0 \\ 0 & K_{22} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \end{bmatrix} &= \begin{bmatrix} C_{11} & 0 \\ 0 & C_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix}. \end{aligned} \quad (4.6)$$

Next, we present Lemma 4.23, Algorithm 9, and Corollary 4.24 below on extensions of Kalman representations in block diagonal form.

Lemma 4.23. Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ and an observable Kalman representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ of \mathbf{y}_2 . If \mathbf{y}_1 and \mathbf{y}_2 mutually do not Granger cause each other then there exists an extension of $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in block diagonal form. Moreover, if the representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ is minimal, then there exists an extension of $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in block diagonal form which is a minimal Kalman representation of \mathbf{y} .

The **proof** can be found in Appendix 4.B.

An algorithm that calculates the representation in Lemma 4.23 is presented next: Consider $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$, its innovation process \mathbf{e} and the covariance sequence

Algorithm 9 Extension of an observable Kalman representation in block diagonal form

Input $\{A_{22}, K_{22}, C_{22}\}$ and $\{\Lambda_k^{\mathbf{y}_1}\}_{k=0}^{2N}$: System matrices of an observable Kalman representation of \mathbf{y}_2 and covariance sequence of \mathbf{y}_1
--

Output $\{A, K, C\}$: System matrices of (4.6)
--

Step 1 Apply Algorithm 1 with input $\{\Lambda_k^{\mathbf{y}_1}\}_{k=0}^{2N}$ and denote its output by $\{A_{11}, K_{11}, C_{11}\}$.

Step 2 Define the following matrices

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \quad K = \begin{bmatrix} K_{11} & 0 \\ 0 & K_{22} \end{bmatrix}, \quad C = \begin{bmatrix} C_{11} & 0 \\ 0 & C_{22} \end{bmatrix}.$$

$\{\Lambda_k^{\mathbf{y}_1}\}_{k=0}^{2N}$ of \mathbf{y}_1 where N is larger than or equal to the dimension of a minimal Kalman representation of \mathbf{y}_1 . Assume that \mathbf{y}_1 and \mathbf{y}_2 mutually do not Granger cause each other. Apply Algorithm 9 with input $\{A_{22}, K_{22}, C_{22}\}$ and $\{\Lambda_k^{\mathbf{y}_1}\}_{k=0}^{2N}$, where $\{A_{22}, K_{22}, C_{22}\}$ are system matrices of an observable Kalman representation

$(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ of \mathbf{y}_2 . Denote the output by $\{A, K, C\}$. Then we have the following result.

Corollary 4.24 (Correctness of Algorithm 9). *The tuple $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is an observable Kalman representation and it is an extension of $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in block diagonal form. Furthermore, if $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ is minimal then $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is also minimal.*

The **proof** can be found in Appendix 4.B.

Next, we discuss the extension of Kalman representations in coordinated form. To extend a Kalman representation of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ in block triangular form to a Kalman representation of $[\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T$ in coordinated form, we will use the following definition:

Definition 4.25. Consider an observable Kalman representation

$$\mathcal{S} = \left(\begin{bmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{bmatrix}, \begin{bmatrix} K_{22} & K_{23} \\ 0 & K_{33} \end{bmatrix}, \begin{bmatrix} C_{22} & C_{23} \\ 0 & C_{33} \end{bmatrix}, I, \begin{bmatrix} \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix}, \begin{bmatrix} \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix} \right)$$

of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ in block triangular form. An extension of \mathcal{S} for $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T$ in coordinated form is an observable Kalman representation of \mathbf{y} in the form

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_2(t+1) \\ \mathbf{x}_3(t+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \mathbf{x}_3(t) \end{bmatrix} + \begin{bmatrix} K_{11} & 0 & K_{13} \\ 0 & K_{22} & K_{23} \\ 0 & 0 & K_{33} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \mathbf{e}_3(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \\ \mathbf{y}_3(t) \end{bmatrix} &= \begin{bmatrix} C_{11} & 0 & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \mathbf{x}_3(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \mathbf{e}_3(t) \end{bmatrix}. \end{aligned} \quad (4.7)$$

Next, we present Lemma 4.26, Algorithm 10 and Corollary 4.27 on extensions of Kalman representations in coordinated form.

Lemma 4.26. *Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T$ and an observable Kalman representation \mathcal{S} of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ in block triangular form. If*

- (i) \mathbf{y}_1 does not Granger cause \mathbf{y}_3 ,
- (ii) \mathbf{y}_2 does not Granger cause \mathbf{y}_3 ,
- (iii) \mathbf{y}_1 conditionally does not Granger cause \mathbf{y}_2 with respect to \mathbf{y}_3 ,
- (iv) \mathbf{y}_2 conditionally does not Granger cause \mathbf{y}_1 with respect to \mathbf{y}_3 ,

then there exists an extension of \mathcal{S} for \mathbf{y} in coordinated form. Moreover, if \mathcal{S} is a minimal Kalman representation in causal block triangular form and for $i \neq j, i, j = 1, 2$

$$E_l[\mathcal{H}_{t+}^{\mathbf{y}_i} | \mathcal{H}_{t-}^{\mathbf{y}_i, \mathbf{y}_3}] \cap E_l[\mathcal{H}_{t+}^{\mathbf{y}_j} | \mathcal{H}_{t-}^{\mathbf{y}_j, \mathbf{y}_3}] \mid E_l[\mathcal{H}_{t+}^{\mathbf{y}_3} | \mathcal{H}_{t-}^{\mathbf{y}_3}] = \{0\},$$

then there exists an extension of \mathcal{S} for \mathbf{y} in coordinated form which is a Kalman representation of \mathbf{y} in causal coordinated form.

The **proof** can be found in Appendix 4.B.

Note that Lemma 4.26 can be seen as a consequence of Theorem 3.5 applied to a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T$. If \mathcal{S} is minimal and in causal block triangular form, then by the isomorphism between minimal Kalman representations, the representation of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ in Theorem 3.5 and in Lemma 4.26 are isomorphic. Otherwise, the representation of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ in Theorem 3.5 can be transformed to the representation \mathcal{S} in Lemma 4.26 with a non-singular state-space transformation. Also, the representation of \mathbf{y} in Theorem 3.5 can be transformed to an observable Kalman representation of \mathbf{y} that is the extension of the representation \mathcal{S} in Lemma 4.26. Therefore, Theorem 3.5 ensures the existence of the Kalman representations of \mathbf{y} in Lemma 4.19 provided the conditional and unconditional Granger non-causality conditions.

An algorithm that calculates the representation in Lemma 4.26 is presented next:

Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T$, its innovation process \mathbf{e} and its covariance sequence $\{\Lambda_k^{\mathbf{y}}\}_{k=0}^{2N}$, where N is larger than or equal to the dimension of a minimal Kalman representation of \mathbf{y} . Assume that \mathbf{y}_1 and \mathbf{y}_2 do not Granger cause each other with respect to \mathbf{y}_3 and $\mathbf{y}_1, \mathbf{y}_2$ do not Granger cause \mathbf{y}_3 . Apply Algorithm 10 with input $\{A_2, K_2, C_2\}$ and $\{\Lambda_k^{\mathbf{y}}\}_{k=0}^{2N}$, where $\{A_2, K_2, C_2\}$ are system matrices of an observable Kalman representation $(A_2, K_2, C_2, I, \mathbf{e}_2)$ of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ in block triangular form. Denote the output by $\{A, K, C\}$. Then we have the following result:

Corollary 4.27 (Correctness of Algorithm 10). *The tuple $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is an observable Kalman representation and it is an extension of $(A_2, K_2, C_2, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in coordinated form. Furthermore, if $(A_2, K_2, C_2, I, \mathbf{e}_2, \mathbf{y}_2)$ is a minimal Kalman representation in causal block triangular form, then $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is a Kalman representation in causal coordinated form.*

The **proof** can be found in Appendix 4.B.

Remark 4.28. Similar to Algorithm 10, Algorithms 7 and 6 in Chapter 3 also calculate Kalman representations in coordinated form. However, Algorithms 7 and 6 are formulated for a more general process class, where \mathbf{y} has $n \geq 3$ components. Furthermore, Algorithm 10 takes, besides the covariances of \mathbf{y} , the system matrices of a Kalman representation of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ as its input and extends it in a way that the

Algorithm 10 Extension of an observable Kalman representation in coordinated form

Input $\{A_2 = \begin{bmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{bmatrix}, K_2 = \begin{bmatrix} K_{22} & K_{23} \\ 0 & K_{33} \end{bmatrix}, C_2 = \begin{bmatrix} C_{22} & C_{23} \\ 0 & C_{33} \end{bmatrix}\}$ and $\{\Lambda_k^y\}_{k=0}^{2N}$: System matrices of an observable Kalman representation of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ and covariance sequence of $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T$

Output $\{A, K, C\}$: System matrices of (4.7)

Step 1 Apply Algorithm 7 with input $\{\Lambda_k^y\}_{k=0}^{2N}$ and denote its output by $\{\hat{A}, \hat{K}, \hat{C}\}$, where

$$\hat{A} = \begin{bmatrix} \hat{A}_{11} & 0 & \hat{A}_{13} \\ 0 & \hat{A}_{22} & \hat{A}_{23} \\ 0 & 0 & \hat{A}_{33} \end{bmatrix}, \hat{K} = \begin{bmatrix} \hat{K}_{11} & 0 & \hat{K}_{13} \\ 0 & \hat{K}_{22} & \hat{K}_{23} \\ 0 & 0 & \hat{K}_{33} \end{bmatrix}, \hat{C} = \begin{bmatrix} \hat{C}_{11} & 0 & \hat{C}_{13} \\ 0 & \hat{C}_{22} & \hat{C}_{23} \\ 0 & 0 & \hat{C}_{33} \end{bmatrix}$$

Step 2 Define $T = \hat{O}_N^+ \mathcal{O}_N$ where \hat{O}^+ is the left inverse of the finite (up to N) observability matrix of $(\hat{A}_{33}, \hat{C}_{33})$ and \hat{O} is the finite (up to N) observability matrix of (A_{33}, C_{33}) .

Step 3 Define the following matrices

$$A = \begin{bmatrix} \hat{A}_{11} & 0 & \hat{A}_{13}T \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}, K = \begin{bmatrix} \hat{K}_{11} & 0 & \hat{K}_{13} \\ 0 & K_{22} & K_{23} \\ 0 & 0 & K_{33} \end{bmatrix}, C = \begin{bmatrix} \hat{C}_{11} & 0 & C_{13}T \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{bmatrix}$$

input Kalman representation is a sub-system of the Kalman representation of \mathbf{y} that the output matrices of Algorithm 10 define. Therefore, contrary to Algorithms 7 and 6, the Kalman representation that Algorithm 10 defines is not necessarily in a causal coordinated form.

4.3.2 Algorithms for Kalman representation with causal TADG-zero structure

To formulate the algorithms that calculate a Kalman representation of \mathbf{y} with G -zero structure, we will use Algorithms 8, 9, and 10. Notice that these algorithms only calculate system matrices of Kalman representations if certain Granger causality conditions hold. We ensure these conditions by relying on the following result:

Lemma 4.29. Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ and a TADG graph $G = (V, E)$ with $V = \{1, \dots, n\}$, and assume that \mathbf{y} has G -consistent causality structure. Then for any $j \in \{1, \dots, n-1\}$ the following holds:

- (i) \mathbf{y}_j does not Granger cause \mathbf{y}_{I_j}
- (ii) $\mathbf{y}_{\bar{I}_j}$ does not Granger cause \mathbf{y}_{I_j}
- (iii) \mathbf{y}_j conditionally does not Granger cause $\mathbf{y}_{\bar{I}_j}$ with respect to \mathbf{y}_{I_j}
- (iv) $\mathbf{y}_{\bar{I}_j}$ conditionally does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} .

The **proof** can be found in Appendix 4.C.

Notice that for $j \neq n$, $\bar{I}_j = \emptyset$ and $I_j = \emptyset$ cannot happen at the same time since $\bar{I}_j \cup I_j = \{j+1, \dots, n\}$. When $\bar{I}_j = \emptyset$ then $I_j = \{j+1, \dots, n\}$ and the Granger causality conditions (i), (ii), (iii), and (iv) simplify to the condition that \mathbf{y}_j does not Granger cause \mathbf{y}_{I_j} . Hence, Algorithm 8 can be applied. On the other hand, if $I_j = \emptyset$ then $\bar{I}_j = \{j+1, \dots, n\}$ and the Granger causality conditions (i), (ii), (iii), and (iv) in Lemma 4.29 simplify to the conditions: $\mathbf{y}_{\bar{I}_j}$ does not Granger cause \mathbf{y}_j and \mathbf{y}_j does not Granger cause $\mathbf{y}_{\bar{I}_j}$. Therefore, Algorithm 9 can be applied. If neither I_j , nor \bar{I}_j is the empty set then from conditions (i)–(ii)–(iii) and (iv) Algorithm 10 can be applied.

Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ and a TADG graph $G = (V, E)$ with $V = \{1, \dots, n\}$, and assume that \mathbf{y} has G -consistent causality structure. The algorithms that calculate a Kalman representation of \mathbf{y} with G -zero structure are elaborated in Algorithms 11 and 12 below. Algorithm 11 takes the covariances of \mathbf{y} as its input and transforms it into a Kalman representation of \mathbf{y} with G -zero structure. Algorithm 12 calculates the same representation from an LTI-SS representation of \mathbf{y} . Note that by using empirical covariances Algorithm 11 can be applied to data.

Let $\{\Lambda_k^{\mathbf{y}}\}_{k=0}^{\infty}$ be the covariance sequence of \mathbf{y} and \mathbf{e} be the innovation process of \mathbf{y} . Furthermore, let N be any number larger than or equal to the dimension of a minimal Kalman representation of \mathbf{y} . Assume that \mathbf{y} has G -consistent causality structure and note that Algorithms 8, 9, and 10 calculate Kalman representations in causal block triangular, block diagonal and in causal coordinated form, respectively (see Remarks 4.20, 4.24, and 4.27). Apply Algorithm 11 with input $\{\Lambda_k^{\mathbf{y}}\}_{k=0}^{2N}$ and denote its output by $\{A, K, C\}$. Now apply Algorithm 12 with input $\{\bar{A}, \bar{B}, \bar{C}, \bar{D}, \Lambda_0^{\mathbf{v}}\}$ where $(\bar{A}, \bar{B}, \bar{C}, \bar{D}, \mathbf{v})$ defines an LTI-SS representation of \mathbf{y} and $\Lambda_0^{\mathbf{v}} = E[\mathbf{v}(t)\mathbf{v}^T(t)]$. Denote the output by $\{\hat{A}, \hat{K}, \hat{C}\}$. Then we have the following result.

Corollary 4.30 (Correctness of Algorithm 11 and Algorithm 12). *The tuples (A, K, C, I, \mathbf{e}) and $(\hat{A}, \hat{K}, \hat{C}, I, \mathbf{e})$ are observable Kalman representations of \mathbf{y} with G -zero structure. Furthermore, if for all nodes $j \in V$ in G*

$$E_l[\mathcal{H}_{t+}^{\mathbf{y}_j} | \mathcal{H}_{t-}^{\mathbf{y}_j, \mathbf{y}_{I_j}}] \cap E_l[\mathcal{H}_{t+}^{\mathbf{y}_{\bar{I}_j}} | \mathcal{H}_{t-}^{\mathbf{y}_{\bar{I}_j}, \mathbf{y}_{I_j}}] \mid E_l[\mathcal{H}_{t+}^{\mathbf{y}_{I_j}} | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}}] = \{0\},$$

then (A, K, C, I, \mathbf{e}) and $(\hat{A}, \hat{K}, \hat{C}, I, \mathbf{e})$ are minimal Kalman representations of \mathbf{y} with causal G -zero structure.

Algorithm 11 Kalman representation with causal G -zero structure based on output covariances

Input $\{\Lambda_k^y\}_{k=0}^{2N}$: Covariance sequence of $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$

Output $\{A, K, C\}$: System matrices of a Kalman representation of \mathbf{y} with G -zero structure

Step 1 Apply Algorithm 1 with input $\{\Lambda_k^y\}_{k=0}^{2N}$ and denote its output by $\{A_n, K_n, C_n, Q^{e_n}\}$.

Step 2

for $i = n, n-1, \dots, 2$

if $\bar{I}_{i-1} = \emptyset$ **then** apply Algorithm 8 with input $\{A_i, K_i, C_i\}$ and $\{\Lambda_k^z\}_{k=0}^{2N}$ where $\mathbf{z} = [\mathbf{y}_{i-1}^T, \mathbf{y}_{\{i, \dots, n\}}^T]^T$. Denote the output by $\{A_{i-1}, K_{i-1}, C_{i-1}\}$.

else if $I_{i-1} = \emptyset$ **then** apply Algorithm 9 with input $\{A_i, K_i, C_i\}$ and $\{\Lambda_k^{y_{i-1}}\}_{k=0}^{2N}$. Denote the output by $\{A_{i-1}, K_{i-1}, C_{i-1}\}$.

end if

if $I_{i-1} \neq \emptyset$ and $\bar{I}_{i-1} \neq \emptyset$ **then** apply Algorithm 10 with input $\{A_i, K_i, C_i\}$ and $\{\Lambda_k^z\}_{k=0}^{2N}$ where $\mathbf{z} = [\mathbf{y}_{i-1}^T, \mathbf{y}_{I_{i-1}}^T, \mathbf{y}_{\bar{I}_{i-1}}^T]^T$. Denote the output by $\{A_{i-1}, K_{i-1}, C_{i-1}\}$.

end if

end for

Step 3 Define $A = A_1, K = K_1$ and $C = C_1$.

Algorithm 12 Kalman representation with G -zero structure based on LTI-SS representation

Input $\{\bar{A}, \bar{B}, \bar{C}, \bar{D}, \Lambda_0^y\}, G = (V, E)$: System matrices of an LTI-SS representation $(\bar{A}, \bar{B}, \bar{C}, \bar{D}, \mathbf{v})$ of \mathbf{y} and variance matrix of \mathbf{v}

Output $\{A, K, C\}$: System matrices of a Kalman representation of \mathbf{y} with G -zero structure

Step 1 Find the solution Σ_x of the Lyapunov equation $\Sigma = \bar{A}\Sigma\bar{A}^T + \bar{B}\Lambda_0^y\bar{B}^T$.

Step 2 Define $G := \bar{C}\Sigma_x\bar{A}^T + \bar{D}\Lambda_0^y\bar{B}^T$ and calculate the output covariance matrices $\Lambda_k^y := \bar{C}\bar{A}^{k-1}G^T$ for $k = 0, \dots, 2n$, where n is such that $\bar{A} \in \mathbb{R}^{n \times n}$.

Step 3 Apply Algorithm 11 with input $\{\Lambda_k^y\}_{k=0}^{2n}$ and denote the output by $\{A, K, C\}$.

The **proof** can be found in Appendix 4.C.

4.4 Conclusions

In this chapter, we have studied Kalman representations whose network graphs are transitive acyclic directed graphs (TADGs), called Kalman representations with

TADG-zero structure. This class of Kalman representations have been related to conditional Granger causality conditions among the components of their output processes. More precisely, we have shown that there exists a Kalman representation with a TADG G -zero structure if and only if certain conditional Granger causality conditions hold that are determined by G . To construct the Kalman representations in question, we provided algorithms that take an arbitrary LTI–SS representation of the output process or the covariance sequence of that process as its input. In fact, the latter input can be substituted with empirical covariances, and thus the algorithm can be applied to data. Also, the results deal with the minimality of the representations and the so-called coercive property of the output processes.

4.A Proofs of Lemmas 4.10 and 4.13

Proof of Lemma 4.10. Consider a TADG $G = (V = \{1, \dots, n\}, E)$ and a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ where $\mathbf{y}_i \in \mathbb{R}^{r_i}$, for $r_i > 0, i = 1, \dots, n$. Let $S = (A, K, C, I, \mathbf{e})$ and $\hat{S} = (\hat{A}, \hat{K}, \hat{C}, I, \mathbf{e})$ be two Kalman representations of \mathbf{y} with causal G -zero structure. Then, by Definition 4.9, for $J := (1, \bar{I}_1, I_1)$ the tuples

$$\begin{aligned} S_{JJ} &= (A_{JJ}, K_{JJ}, C_{JJ}, I, [\mathbf{e}_1^T, \mathbf{e}_{\bar{I}_1}^T, \mathbf{e}_{I_1}^T]^T) \\ \hat{S}_{JJ} &= (\hat{A}_{JJ}, \hat{K}_{JJ}, \hat{C}_{JJ}, I, [\mathbf{e}_1^T, \mathbf{e}_{\bar{I}_1}^T, \mathbf{e}_{I_1}^T]^T) \end{aligned}$$

are Kalman representations of $[\mathbf{y}_1^T, \mathbf{y}_{\bar{I}_1}^T, \mathbf{y}_{I_1}^T]^T$ in causal coordinated form. By using Lemma 3.2, it follows that S_{JJ} and \hat{S}_{JJ} are isomorphic with a non-singular T transformation matrix, i.e., $A_{JJ} = T\hat{A}_{JJ}T^{-1}$, $K_{JJ} = T\hat{K}_{JJ}$, $C_{JJ} = \hat{C}_{JJ}T^{-1}$. Let the state processes of S and \hat{S} be $\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_n^T]^T$ and $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1^T, \dots, \hat{\mathbf{x}}_n^T]^T$, respectively. Assume that $\bar{I}_1 = \{i_1, \dots, i_k\}$ and that $I_1 = \{i_{k+1}, \dots, i_{n-1}\}$. Define the permutation matrices $P_{\mathbf{y}}$ and $P_{\mathbf{x}}$ such that

$$\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_n \end{bmatrix} = P_{\mathbf{y}} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_{\bar{I}_1} \\ \mathbf{y}_{I_1} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} = P_{\mathbf{x}} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_{\bar{I}_1} \\ \mathbf{x}_{I_1} \end{bmatrix}.$$

Note that $P_{\mathbf{y}}^{-1} = P_{\mathbf{y}}^T$ and $P_{\mathbf{x}}^{-1} = P_{\mathbf{x}}^T$. Then,

$$\begin{aligned} A &= P_{\mathbf{x}}^T A_{JJ} P_{\mathbf{x}} & K &= P_{\mathbf{x}}^T K_{JJ} P_{\mathbf{y}} & C &= P_{\mathbf{y}}^T C_{JJ} P_{\mathbf{x}} \\ \hat{A} &= P_{\mathbf{x}}^T \hat{A}_{JJ} P_{\mathbf{x}} & \hat{K} &= P_{\mathbf{x}}^T \hat{K}_{JJ} P_{\mathbf{y}} & \hat{C} &= P_{\mathbf{y}}^T \hat{C}_{JJ} P_{\mathbf{x}}. \end{aligned}$$

Therefore, by using $A_{JJ} = T\hat{A}_{JJ}T^{-1}$ we obtain that $P_{\mathbf{x}}AP_{\mathbf{x}}^T = TP_{\mathbf{x}}\hat{A}_{JJ}P_{\mathbf{x}}^T T^{-1}$, by using $K_{JJ} = T\hat{K}_{JJ}$ we obtain that $P_{\mathbf{x}}KP_{\mathbf{y}}^T = TP_{\mathbf{x}}\hat{K}_{JJ}P_{\mathbf{y}}^T$ and lastly, by using $C_{JJ} = \hat{C}_{JJ}T^{-1}$ we obtain that $P_{\mathbf{y}}CP_{\mathbf{x}}^T = P_{\mathbf{y}}\hat{C}_{JJ}P_{\mathbf{x}}^T T^{-1}$. It then follows that by the transformation matrix $\tilde{T} = P_{\mathbf{x}}TP_{\mathbf{x}}^T$ the representations S and \hat{S} are isomorphic. \square

For the proof of Lemma 4.13 we need auxiliary lemmas on the properties of Granger and conditional Granger causality: First, we recall Lemma 3.12 from Appendix 3.B and Lemma 3.6 from Chapter 3:

Lemma 4.31 (Lemma 3.12). *Consider a ZMSIR process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T, \mathbf{y}_4^T]^T$. Then \mathbf{y}_1 and \mathbf{y}_2 conditionally do not Granger cause \mathbf{y}_3 with respect to \mathbf{y}_4 if and only if $[\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ conditionally does not Granger cause \mathbf{y}_3 with respect to \mathbf{y}_4 .*

Lemma 4.32 (Lemma 3.6). *Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T$ and the following statements*

- (i) \mathbf{y}_1 does not Granger cause \mathbf{y}_3
- (ii) \mathbf{y}_2 does not Granger cause \mathbf{y}_3
- (iii) \mathbf{y}_1 conditionally does not Granger cause \mathbf{y}_2 with respect to \mathbf{y}_3
- (iv) \mathbf{y}_1 does not Granger cause $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$

Then (i)-(ii)-(iii) if and only if (ii)-(iv).

There are two additional results the we employ in the proof of Lemma 4.13. The first one is as follows:

Lemma 4.33. Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T, \mathbf{y}_4^T]^T$. If \mathbf{y}_1 and \mathbf{y}_2 conditionally do not Granger cause \mathbf{y}_3 with respect to \mathbf{y}_4 then \mathbf{y}_1 conditionally does not Granger cause \mathbf{y}_3 with respect to $[\mathbf{y}_2^T, \mathbf{y}_4^T]^T$.

Proof. Let $\alpha = \mathbf{y}_3(t+s) - E_l[\mathbf{y}_3(t+s)|\mathcal{H}_{t-}^{\mathbf{y}_3, \mathbf{y}_4}]$ for some $t, s \in \mathbb{Z}, s > 0$. Then from the conditional Granger non-causality conditions we obtain that $\alpha = \mathbf{y}_3(t+s) - E_l[\mathbf{y}_3(t+s)|\mathcal{H}_{t-}^{\mathbf{y}_1, \mathbf{y}_3, \mathbf{y}_4}]$ and $\alpha = \mathbf{y}_3(t+s) - E_l[\mathbf{y}_3(t+s)|\mathcal{H}_{t-}^{\mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4}]$. Therefore, α is orthogonal to $\mathcal{H}_{t-}^{\mathbf{y}_1, \mathbf{y}_3, \mathbf{y}_4}$ and to $\mathcal{H}_{t-}^{\mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4}$ and thus to $\mathcal{H}_{t-}^{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4}$. This implies that $E_l[\alpha|\mathcal{H}_{t-}^{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4}] = 0$ and thus $E_l[\mathbf{y}_3(t+s)|\mathcal{H}_{t-}^{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4}] = E_l[\mathbf{y}_3(t+s)|\mathcal{H}_{t-}^{\mathbf{y}_3, \mathbf{y}_4}]$. From the condition that \mathbf{y}_2 conditionally does not Granger cause \mathbf{y}_3 w.r.t. \mathbf{y}_4 , the latter is further equivalent to $E_l[\mathbf{y}_3(t+s)|\mathcal{H}_{t-}^{\mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4}]$. That is, $E_l[\mathbf{y}_3(t+s)|\mathcal{H}_{t-}^{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4}] = E_l[\mathbf{y}_3(t+s)|\mathcal{H}_{t-}^{\mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4}]$, which holds for any choice of $t, s \in \mathbb{Z}, s > 0$. This, by definition means that \mathbf{y}_1 conditionally does not Granger cause \mathbf{y}_3 w.r.t. $[\mathbf{y}_2^T, \mathbf{y}_4^T]^T$. \square

The last auxiliary lemma that helps us in proving Lemma 4.13 is presented below.

Lemma 4.34. Consider a process $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_n^T]^T$ and a TADG $G = (V = \{1, \dots, n\}, E)$. Then we state the following for the conditions below: for any node $i \in V$ (i), (ii), and (iii) hold if and only if (i), (iv), and (v) hold.

- (i) \mathbf{y}_i and $\mathbf{y}_{\bar{I}_i}$ do not Granger cause \mathbf{y}_{I_i}
- (ii) \mathbf{y}_i does not Granger cause $[\mathbf{y}_{I_i}, \mathbf{y}_{\bar{I}_i}]$
- (iii) $\mathbf{y}_{\bar{I}_i}$ does not Granger cause $[\mathbf{y}_{I_i}, \mathbf{y}_i]$
- (iv) \mathbf{y}_i conditionally does not Granger cause $\mathbf{y}_{\bar{I}_i}$ with respect to \mathbf{y}_{I_i}
- (v) $\mathbf{y}_{\bar{I}_i}$ conditionally does not Granger cause \mathbf{y}_i with respect to \mathbf{y}_{I_i}

Proof. Considering (i), (ii), and (iii), we can apply Lemma 4.32 to $\mathbf{y} = [\mathbf{y}_i^T, \mathbf{y}_{\bar{I}_i}^T, \mathbf{y}_{I_i}^T]^T$ and to $\mathbf{y} = [\mathbf{y}_{\bar{I}_i}^T, \mathbf{y}_i^T, \mathbf{y}_{I_i}^T]^T$. As a results, we obtain that the conditions (i), (ii), and (iii) hold if and only if the conditions (i), (iv), and (v) hold for any $i \in V$. \square

Remark 4.35. A Granger causality condition that \mathbf{y}_1 does not Granger cause $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ means by definition that

$$E_l \left[\begin{bmatrix} \mathbf{y}_2(t+k) \\ \mathbf{y}_3(t+k) \end{bmatrix} \middle| \mathcal{H}_{t-}^{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3} \right] = E_l \left[\begin{bmatrix} \mathbf{y}_2(t+k) \\ \mathbf{y}_3(t+k) \end{bmatrix} \middle| \mathcal{H}_{t-}^{\mathbf{y}_2, \mathbf{y}_3} \right]$$

for all $t, k \in \mathbb{Z}, k > 0$. By looking at the latter component-wise, an equivalent form is that \mathbf{y}_1 conditionally does not Granger cause \mathbf{y}_2 with respect to \mathbf{y}_3 and \mathbf{y}_1 conditionally does not Granger cause \mathbf{y}_3 with respect to \mathbf{y}_2 .

Also, it trivially holds for any $[\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]$ process that \mathbf{y}_1 conditionally does not Granger cause \mathbf{y}_2 with respect to $[\mathbf{y}_1, \mathbf{y}_3]$. That is, the conditional Granger non-causality holds automatically because \mathbf{y}_1 is in the condition of the conditional Granger non-causality.

Now we are ready to present the proof of Lemma 4.13.

Proof of Lemma 4.13. *Necessity:* We will prove that if the conditions (i), (ii), and (iii) in Lemma 4.34 hold for all $i \in V$ then \mathbf{y} has G -consistent causality structure. From Lemma 4.34 we know that (i), (ii), and (iii) imply (iv) and (v). By Lemma 4.31, (v) holds if and only if for all $i \in \bar{I}_j$ \mathbf{y}_i does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} . Recall that \mathbf{y} has G -consistent causality structure if $(i, j) \notin E$ implies that \mathbf{y}_i does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} . Therefore, considering (v), it remains to show that \mathbf{y}_i does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} for all $i \in V \setminus \bar{I}_j$ where $(i, j) \notin E$.

Define the set $S = \{i \in V | i < j\}$ and notice that $S = \{i \in V \setminus \bar{I}_j | (i, j) \notin E\}$. Therefore, to finish our proof, we have to show that for any $i \in S$, \mathbf{y}_i does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} . Fix an $s = j - L \in S$ and apply condition (ii) to $\mathbf{y}_{j-l}, l = L, \dots, j - 1$: for component \mathbf{y}_j it gives that \mathbf{y}_{j-l} conditionally does not Granger cause \mathbf{y}_j with respect to $[\mathbf{y}_{\bar{I}_{j-l}}^T, \mathbf{y}_{\bar{I}_{j-l}}^T]^T$, see also Remark 4.35. From $I_{j-l} \cup \bar{I}_{j-l} = I_{j-l+1} \cup \bar{I}_{j-l+1} \cup \{j-l+1\}$ the latter implies that

$$\begin{aligned} E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_{j-L}}, \mathbf{y}_{\bar{I}_{j-L}}, \mathbf{y}_{j-L}}] &= E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_{j-L}}, \mathbf{y}_{\bar{I}_{j-L}}}] = \\ E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_{j-L+1}}, \mathbf{y}_{\bar{I}_{j-L+1}}, \mathbf{y}_{j-L+1}}] &= E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_{j-L+1}}, \mathbf{y}_{\bar{I}_{j-L+1}}}] = \\ \dots &= E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_{j-1}}, \mathbf{y}_{\bar{I}_{j-1}}, \mathbf{y}_{j-1}}] = E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_{j-1}}, \mathbf{y}_{\bar{I}_{j-1}}}] = \\ &E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_{\bar{I}_j}, \mathbf{y}_j}] = E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_j}], \end{aligned} \quad (4.8)$$

where the last equality follows from (v). Projecting the first and last expressions of

the equation above to $\mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_j, \mathbf{y}_{j-L}}$ and considering that

$$\mathcal{H}_{t-}^{\mathbf{y}_{I_{j-L}}, \mathbf{y}_{\bar{I}_{j-L}}, \mathbf{y}_{j-L}} \supseteq \mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_j, \mathbf{y}_{j-L}} \supseteq \mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_j},$$

it follows that $E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_j, \mathbf{y}_{j-L}}] = E_l[\mathbf{y}_j(t+k) | \mathcal{H}_{t-}^{\mathbf{y}_{I_j}, \mathbf{y}_j}]$. This, by definition means that \mathbf{y}_{j-L} does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} . Since $s = j - L$ was an arbitrary element in S , this proves that for any $s \in S$ \mathbf{y}_s does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} which completes the necessity part of the proof.

Sufficiency: Below, we will show that G -consistent causality structure of \mathbf{y} implies (i), (ii), and (iii) in Lemma 4.34, respectively.

G -consistent causality implies (i): Consider a node $j \in V$ and let S be a subset of the set $\{l \in V \mid (l, j) \notin E\}$. Then, since for all $s \in S$, \mathbf{y}_s does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} , from Lemma 4.31 it follows that \mathbf{y}_S does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} . By definition it gives that

$$E_l[\mathbf{y}_j(t+k) | \mathcal{H}^{\mathbf{y}_{I_j}, \mathbf{y}_j, \mathbf{y}^S}] = E_l[\mathbf{y}_j(t+k) | \mathcal{H}^{\mathbf{y}_{I_j}, \mathbf{y}_j}]. \quad (4.9)$$

Next, by using (4.9), we show that for any $i \in V$ \mathbf{y}_i does not Granger cause \mathbf{y}_{I_i} . Let $j \in I_i$ and notice that since G is acyclic, $(i, j) \notin E$. Moreover, for any $l \in I_i \setminus I_j$, $(l, j) \notin E$, hence applying (4.9) to $S = i \cup (I_i \setminus I_j)$ and then to $S = I_i \setminus I_j$, it follows that

$$E_l[\mathbf{y}_j(t+k) | \mathcal{H}^{\mathbf{y}_{I_i}, \mathbf{y}_i}] = E_l[\mathbf{y}_j(t+k) | \mathcal{H}^{\mathbf{y}_{I_j}, \mathbf{y}_j}] = E_l[\mathbf{y}_j(t+k) | \mathcal{H}^{\mathbf{y}_{I_i}, \mathbf{y}_j}]$$

for every $j \in I_i$. In other words, \mathbf{y}_i does not Granger cause \mathbf{y}_{I_i} , see Remark 4.35.

For proving that $\mathbf{y}_{\bar{I}_i}$ does not Granger cause \mathbf{y}_{I_i} , we apply (4.9) for $S = \bar{I}_i$ and $j \in I_i$. Notice that $\bar{I}_i \subseteq \{l \in V \mid (l, j) \notin E\}$ since if $l \in \bar{I}_i$, then $(l, j) \notin E$ for any $j \in I_i$, otherwise $(l, j), (j, i) \in E$ would imply $(l, i) \in E$ by transitivity, which contradicts $l \in \bar{I}_i$.

G -consistent causality implies (ii): Let $s \in I_j \cup \bar{I}_j$. Then notice that since $s > j$ we know that $(j, s) \notin E$. Hence, from the fact that \mathbf{y} has G -consistent causality structure, \mathbf{y}_j conditionally does not Granger cause \mathbf{y}_s with respect to \mathbf{y}_{I_s} . Let $S = I_j \cup \bar{I}_j \setminus I_s$ and assume that $S = \{s_1, \dots, s_L\}$. To see that from the G -consistent causality structure of \mathbf{y} condition (ii) follows, we first show that \mathbf{y}_j and $\mathbf{y}_{s_{p+1}}$ conditionally do not Granger cause \mathbf{y}_s with respect to $[\mathbf{y}_{I_s}^T, \mathbf{y}_{s_1, \dots, s_p}^T]^T$ for all $p = 1, \dots, L - 1$.

Notice that $(s_k, s) \notin E$ for all $k = 1, \dots, L$, hence, \mathbf{y}_{s_1} and \mathbf{y}_{s_2} conditionally do not Granger cause \mathbf{y}_s with respect to \mathbf{y}_{I_s} . Considering the latter two conditional Granger causalities, we can apply Lemma 4.33 to $[\mathbf{y}_j^T, \mathbf{y}_{s_1}^T, \mathbf{y}_s^T, \mathbf{y}_{I_s}^T]^T$ and to $[\mathbf{y}_{s_2}^T, \mathbf{y}_{s_1}^T, \mathbf{y}_s^T, \mathbf{y}_{I_s}^T]^T$. Then, we obtain that \mathbf{y}_j and \mathbf{y}_{s_2} conditionally do not Granger cause \mathbf{y}_s with respect to $[\mathbf{y}_{I_s}^T, \mathbf{y}_{s_1}^T]^T$. Assume by induction that \mathbf{y}_j and $\mathbf{y}_{s_{i+1}}$ conditionally do not Granger cause \mathbf{y}_s with respect to $[\mathbf{y}_{I_s}^T, \mathbf{y}_{s_1, \dots, s_i}^T]^T$ for $i = 1, \dots, p - 1$

where p is smaller than the number of elements in S . From this, we can apply Lemma 4.33 to $[\mathbf{y}_j^T, \mathbf{y}_{s_p}^T, \mathbf{y}_s^T, [\mathbf{y}_{I_s}^T, \mathbf{y}_{s_1, \dots, s_{p-1}}^T]]^T$ and to $[\mathbf{y}_{s_{p+1}}^T, \mathbf{y}_{s_p}^T, \mathbf{y}_s^T, [\mathbf{y}_{I_s}^T, \mathbf{y}_{s_1, \dots, s_{p-1}}^T]]^T$. As a result, we obtain that \mathbf{y}_j and $\mathbf{y}_{s_{p+1}}$ conditionally do not Granger cause \mathbf{y}_s with respect to $[\mathbf{y}_{I_s}^T, \mathbf{y}_{s_1, \dots, s_p}^T]^T$, which completes the induction.

From the discussion above, \mathbf{y}_j and \mathbf{y}_{s_L} conditionally do not Granger cause \mathbf{y}_s with respect to $[\mathbf{y}_{I_s}^T, \mathbf{y}_{s_1, \dots, s_{L-1}}^T]^T$. By applying Lemma 4.33 to $[\mathbf{y}_j^T, \mathbf{y}_{s_L}^T, \mathbf{y}_s^T, [\mathbf{y}_{I_s}^T, \mathbf{y}_{s_1, \dots, s_{L-1}}^T]]^T$ we obtain that \mathbf{y}_j conditionally does not Granger cause \mathbf{y}_s with respect to $[\mathbf{y}_{I_s}^T, \mathbf{y}_s^T]^T$. Since s is an arbitrary element of $I_j \cup \bar{I}_j$, the latter is equivalent to condition (ii) if we look at the condition component-wise, see Remark 4.35.

G-consistent causality implies (iii): From Lemma 4.32 it follows that (i) and (v) is equivalent to (i) and (iii). Since (i) holds, in order to prove (iii) we will instead prove (v). Let $j \in V$ and $\bar{I}_j = \{\bar{i}_1, \dots, \bar{i}_p\}$ and notice that from $(\bar{i}_k, j) \notin E$ it follows that $\mathbf{y}_{\bar{i}_k}$ conditionally does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} , $k = 1, \dots, p$. Then, applying Lemma 4.31 to $[\mathbf{y}_{\bar{i}_1}^T, \mathbf{y}_{\bar{i}_2}^T, \mathbf{y}_j^T, \mathbf{y}_{I_j}^T]^T$ we obtain that $[\mathbf{y}_{\bar{i}_1}^T, \mathbf{y}_{\bar{i}_2}^T]^T$ conditionally does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} . Apply now Lemma 4.31 to $[\mathbf{y}_{\bar{i}_1, \bar{i}_2}^T, \mathbf{y}_{\bar{i}_3}^T, \mathbf{y}_j^T, \mathbf{y}_{I_j}^T]^T, \dots, [\mathbf{y}_{\bar{i}_1, \dots, \bar{i}_{p-1}}^T, \mathbf{y}_{\bar{i}_p}^T, \mathbf{y}_j^T, \mathbf{y}_{I_j}^T]^T$, respectively. It then implies that $\mathbf{y}_{\bar{I}_j}$ conditionally does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j} , i.e., (v) holds. \square

4.B Proof of auxiliary results in Section 4.3.1

In this section, we present the proofs of Lemmas 4.19-, 4.23, 4.26, and Corollaries 4.20, 4.24, and 4.27 from Section 4.3.1. These results are used later on in the proof of Theorem 4.15.

Proof of Lemma 4.19. Consider an observable Kalman representation $S_2 = (A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ with state process $\mathbf{x}_2 \in \mathbb{R}^{n_2}$. Then $E_l[\mathbf{y}_2(t+k) | \mathcal{H}_{t-}^{Y_2}] = C_{22} A_{22}^k \mathbf{x}_2(t)$ for all $k \geq 1$ and thus $E_l[Y_2(t) | \mathcal{H}_{t-}^{Y_2}] = \mathcal{O}_{n_2} \mathbf{x}_2(t)$ where \mathcal{O}_{n_2} is the finite observability matrix of (A_{22}, C_{22}) (up to n_2) and $Y_2(t) = [\mathbf{y}_2^T(t) \dots \mathbf{y}_2^T(t+n_2-1)]^T$.

Recall now the result (i) \iff (iii) of Theorem 2.5 in Chapter 2:

Corollary 4.36 (Theorem 2.5, (i) \iff (iii)). *Consider a ZMSIR 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 there exists a minimal Kalman representation of \mathbf{y} in block triangular form*

$$\begin{aligned} \begin{bmatrix} \hat{\mathbf{x}}_1(t+1) \\ \hat{\mathbf{x}}_2(t+1) \end{bmatrix} &= \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_1(t) \\ \hat{\mathbf{x}}_2(t) \end{bmatrix} + \begin{bmatrix} \hat{K}_{11} & \hat{K}_{12} \\ 0 & \hat{K}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \end{bmatrix} &= \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12} \\ 0 & \hat{C}_{22} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_1(t) \\ \hat{\mathbf{x}}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix}, \end{aligned} \tag{4.10}$$

where $(\hat{A}_{22}, \hat{K}_{22}, \hat{C}_{22}, I, \mathbf{e}_2)$ is a minimal Kalman representation of \mathbf{y}_2 .

Consider a minimal Kalman representation (4.10) of \mathbf{y} (by assumption \mathbf{y}_1 does not Granger cause \mathbf{y}_2). Then $(\hat{A}_{22}, \hat{K}_{22}, \hat{C}_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ is a minimal Kalman representation with state process $\hat{\mathbf{x}}_2$. Notice that $E_t[Y_2(t)|\mathcal{H}_t^{y_2}] = \hat{O}_{n_2} \hat{\mathbf{x}}_2(t)$ where \hat{O}_{n_2} is the finite observability matrix of $(\hat{A}_{22}, \hat{C}_{22})$ (up to n_2 , see (4.5)) and $Y_2(t) = [\mathbf{y}_2^T(t) \dots \mathbf{y}_2^T(t+n_2-1)]^T$. Since $(\hat{A}_{22}, \hat{K}_{22}, \hat{C}_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ is a minimal, thus observable Kalman representation, we have that $\hat{O}_{n_2}^+ E_t[Y_2(t)|\mathcal{H}_t^{y_2}] = \hat{\mathbf{x}}_2(t)$ where $\hat{O}_{n_2}^+$ is the left inverse of \hat{O}_{n_2} . Define now $T = \hat{O}_{n_2}^+ \hat{O}_{n_2}$ and notice that $\hat{\mathbf{x}}_2 = T \mathbf{x}_2$. Then

$$\begin{aligned} \begin{bmatrix} \hat{\mathbf{x}}_1(t+1) \\ \mathbf{x}_2(t+1) \end{bmatrix} &= \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12}T \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} \hat{K}_{11} & \hat{K}_{12} \\ 0 & K_{22} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \end{bmatrix} &= \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12}T \\ 0 & C_{22} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix} \end{aligned} \quad (4.11)$$

is a Kalman representation of \mathbf{y} . Furthermore, it is observable since the observability of (A_{22}, C_{22}) and $(\hat{A}_{11}, \hat{C}_{11})$ ensures the observability of (4.11). Hence, (4.11) is the extension of $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$ for \mathbf{y} in block triangular form.

If $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ was a minimal Kalman representation of \mathbf{y}_2 then the dimension of $\hat{\mathbf{x}}_2$ and \mathbf{x}_2 would be the same, i.e., (4.11) would be a minimal Kalman representation in causal block triangular form. \square

Proof of Corollary 4.20. Let S_2 in the proof of Lemma 4.19 be $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$, where \mathbf{e}_2 is the innovation process of \mathbf{y}_2 and A_{22}, K_{22}, C_{22} are the input matrices of Algorithm 8. Then the representation (4.11) coincides with the Kalman representation (A, K, C, I, \mathbf{e}) , where A, K, C are the output matrices of Algorithm 8. This completes the proof. \square

Proof of Lemma 4.23. If \mathbf{y}_1 and \mathbf{y}_2 mutually do not Granger cause each other then the innovation process \mathbf{e}_1 of \mathbf{y}_1 and the innovation process \mathbf{e}_2 of \mathbf{y}_2 together as $[\mathbf{e}_1^T, \mathbf{e}_2^T]^T$ form the innovation process of $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$. Then, putting together a minimal Kalman representation $(A_{11}, K_{11}, C_{11}, I, \mathbf{e}_1)$ of \mathbf{y}_1 and an observable Kalman representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ of \mathbf{y}_2 into a block diagonal representation such as

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_2(t+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} K_{11} & 0 \\ 0 & K_{22} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \end{bmatrix} &= \begin{bmatrix} C_{11} & 0 \\ 0 & C_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \end{bmatrix}, \end{aligned} \quad (4.12)$$

we obtain an observable Kalman representation of \mathbf{y} which is an extension of the Kalman representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ in block diagonal form. Note that the observability of (4.12) is ensured by the observability of (A_{11}, C_{11}) and (A_{22}, C_{22}) .

If the representation $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2)$ of \mathbf{y}_2 was minimal then (4.12) would also be minimal. Indeed, the controllability of (A_{11}, K_{11}) and (A_{22}, K_{22}) ensures the controllability of (4.12) (see Proposition 1.10). \square

Proof of Corollary 4.24. Let the observable Kalman representation of \mathbf{y}_2 in the proof of Lemma 4.23 be $(A_{22}, K_{22}, C_{22}, I, \mathbf{e}_2, \mathbf{y}_2)$, where \mathbf{e}_2 is the innovation process of \mathbf{y}_2 and A_{22}, K_{22}, C_{22} are the input matrices of Algorithm 8. Then the representation (4.12) coincides with the Kalman representation (A, K, C, I, \mathbf{e}) , where A, K, C are the output matrices of Algorithm 8. This completes the proof. \square

Proof of Lemma 4.26. Consider an observable Kalman representation

$$S = \left(\begin{bmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{bmatrix}, \begin{bmatrix} K_{22} & K_{23} \\ 0 & K_{33} \end{bmatrix}, \begin{bmatrix} C_{22} & C_{23} \\ 0 & C_{33} \end{bmatrix}, I, \begin{bmatrix} \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix} \right)$$

of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ in block triangular form, where $\dim(\mathbf{e}_i) = \dim(\mathbf{y}_i)$ for $i = 2, 3$. Denote the tuple $(A_{33}, K_{33}, C_{33}, I, \mathbf{e}_3)$ by S_3 . Notice that A_{33} is stable and because of (ii), the noise process \mathbf{e}_3 is the innovation process of \mathbf{y}_3 and hence S_3 is a Kalman representation of \mathbf{y}_3 . Furthermore, it is observable. Then, by using (i), we can apply Lemma 4.19 to obtain an observable Kalman representation for $[\mathbf{y}_1^T, \mathbf{y}_3^T]^T$ in block triangular form as follows:

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_3(t+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{13} \\ 0 & A_{33} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_3(t) \end{bmatrix} + \begin{bmatrix} K_{11} & K_{13} \\ 0 & K_{33} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_3(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_3(t) \end{bmatrix} &= \begin{bmatrix} C_{11} & C_{13} \\ 0 & C_{33} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_3(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_3(t) \end{bmatrix}, \end{aligned} \quad (4.13)$$

Combine the representation S of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ and the representation (4.13) of $[\mathbf{y}_1^T, \mathbf{y}_3^T]^T$ such that

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_2(t+1) \\ \mathbf{x}_3(t+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \mathbf{x}_3(t) \end{bmatrix} + \begin{bmatrix} K_{11} & 0 & K_{13} \\ 0 & K_{22} & K_{23} \\ 0 & 0 & K_{33} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \mathbf{e}_3(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \\ \mathbf{y}_3(t) \end{bmatrix} &= \begin{bmatrix} C_{11} & 0 & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \mathbf{x}_3(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \mathbf{e}_3(t) \end{bmatrix}. \end{aligned} \quad (4.14)$$

From conditions (iii) and (iv), it follows that \mathbf{e}_1 and \mathbf{e}_2 are the first and second components of the innovation process of $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \mathbf{y}_3^T]^T$. In addition, by using Lemma 3.12] that was recalled as Lemma 4.31 in Appendix 4.B, we obtain that the conditions (i) and (ii) are equivalent to that $[\mathbf{y}_1^T, \mathbf{y}_2^T]$ does not Granger cause \mathbf{y}_3 . The latter implies that $E_l[\mathbf{y}_3(t+k)|\mathcal{H}_{t-}^{\mathbf{y}_3}] = E_l[\mathbf{y}_3(t+k)|\mathcal{H}_{t-}^{\mathbf{y}}]$ for all $t, k \in \mathbb{Z}, k > 0$, i.e., $\mathbf{e}_3(t)$ is the third component of the innovation process of \mathbf{y} . As a consequence, (4.14) is a Kalman representation of \mathbf{y} in coordinated form. Furthermore, (4.14) is observable since the pairs $(A_{11}, C_{11}), (A_{22}, C_{22})$ and (A_{33}, C_{33}) are observable pairs.

Assume that S is a minimal representation in causal block triangular form and that for $i \neq j, i, j = 1, 2$

$$E_l[\mathcal{H}_{t+}^{\mathbf{y}_i} | \mathcal{H}_{t-}^{\mathbf{y}_i, \mathbf{y}_3}] \cap E_l[\mathcal{H}_{t+}^{\mathbf{y}_j} | \mathcal{H}_{t-}^{\mathbf{y}_j, \mathbf{y}_3}] | E_l[\mathcal{H}_{t+}^{\mathbf{y}_3} | \mathcal{H}_{t-}^{\mathbf{y}_3}] = \{0\}. \quad (4.15)$$

Then $(A_{33}, K_{33}, C_3, I, \mathbf{e}_3)$ is a minimal representation of \mathbf{y}_3 . Hence, when we apply Lemma 4.19, the representation (4.13) is minimal and in causal block triangular form. Therefore, (4.14) is a Kalman representation of \mathbf{y} in causal coordinated form. We know from Theorem 3.5 that the conditions (i), (ii), (iii), (iv) and (4.15) imply that there exist a minimal Kalman representation of \mathbf{y} in causal coordinated form. Since, by Lemma 3.2, Kalman representations of \mathbf{y} in causal coordinated form are isomorphic, we obtain that (4.14) is also minimal. \square

Proof of Corollary 4.27. Notice that the steps of the proof of Lemma 4.23 coincide with the steps of Algorithm 10. That is, if in the proof of Lemma 4.23 the initial observable Kalman representation S of $[\mathbf{y}_2^T, \mathbf{y}_3^T]^T$ was the Kalman representation $(A_2, K_2, C_2, I, [\mathbf{e}_2^T, \mathbf{e}_3^T]^T, [\mathbf{y}_2^T, \mathbf{y}_3^T]^T)$, where A_2, K_2, C_2 are the input matrices of Algorithm 10, then the system matrices A, K, C of the Kalman representation (4.14) in Lemma 4.23 would coincide with the output matrices of Algorithm 10. This completes the proof. \square

4.C Proofs of Lemma 4.29 , Theorem 4.15, and Corollary 4.30

For the proof of the Lemma 4.29 we will use two auxiliary lemmas. The first one is Lemma 3.12 from Chapter 3 that was recalled as Lemma 4.32 in Appendix 4.B and the second one is Lemma 4.13.

Proof of Lemma 4.29. By using Lemma 4.13 we obtain that \mathbf{y} has G -consistent causality structure if and only if

- (i) \mathbf{y}_j does not Granger cause \mathbf{y}_{I_j}

- (ii) $y_{\bar{I}_j}$ does not Granger cause y_{I_j}
- (iii) y_j does not Granger cause $[y_{I_j} y_{\bar{I}_j}]$
- (iv) $y_{\bar{I}_j}$ does not Granger cause $[y_{I_j} y_j]$

Then if we apply Lemma 4.32 to $\mathbf{y} = [y_j^T, y_{\bar{I}_j}^T, y_{I_j}^T]^T$ and to $\mathbf{y} = [y_{\bar{I}_j}^T, y_j^T, y_{I_j}^T]^T$ we obtain the statement of the Lemma. \square

Next, we present the proof of Theorem 4.15.

Proof of Theorem 4.15. Consider a TADG $G = (V = \{1, \dots, n\}, E)$ and a process $\mathbf{y} = [y_1^T, \dots, y_n^T]^T$. Then notice that any Kalman representation with causal G -zero structure is a Kalman representation with G -zero structure, hence (iv) \implies (v) follows. We continue with the proof of the remaining implications.

(i) \implies (v): Assume that \mathbf{y} has G -consistent causality structure. Using induction, we will show that with the help of Lemma 4.19, 4.23, and 4.26, an observable Kalman representation with G -zero structure can be constructed. In fact, we will show that for $\mathbf{y} = [y_{n-j}^T, \dots, y_n^T]^T$, $j = 1, \dots, n-1$ there exists an observable Kalman representation S_j with a $G|_{\{n-j, \dots, n\}}$ -zero structure such that S_j is an extension of S_{j-1} in block triangular, block diagonal or coordinated form.

Recall that the graph $G|_{\{n-j, \dots, n\}}$ is the restriction of G to the set of vertices $\{n-j, \dots, n\} \subseteq V$ and note that if G is TADG, then so is $G|_{\{n-j, \dots, n\}}$, see also Remark 4.17. For $j = 1$, $G|_{\{n-1, \dots, n\}}$ can be two types of graph: either $G|_{\{n-1, n\}} = (\{n-1, n\}, \{(n, n-1)\})$ or $G|_{\{n-1, n\}} = (\{n-1, n\}, \emptyset)$. Let $S_1 = (A_{nn}, K_{nn}, C_{nn}, I, \mathbf{e}_n)$ be a minimal Kalman representation of \mathbf{y}_n . If $G|_{\{n-1, n\}} = (\{n-1, n\}, \{(n, n-1)\})$, then by assumption, \mathbf{y}_{n-1} does not Granger cause \mathbf{y}_n . Hence, by using Lemma 4.19 for S_1 and $[y_{n-1}^T, y_n^T]^T$, we obtain a minimal Kalman representation of $[y_{n-1}^T, y_n^T]^T$ that is an extension of S_1 for $[y_{n-1}^T, y_n^T]^T$ in causal block triangular form.

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_{n-1}(t+1) \\ \mathbf{x}_n(t+1) \end{bmatrix} &= \begin{bmatrix} A_{(n-1)(n-1)} & A_{(n-1)n} \\ 0 & A_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{n-1}(t) \\ \mathbf{x}_n(t) \end{bmatrix} + \begin{bmatrix} K_{(n-1)(n-1)} & K_{(n-1)n} \\ 0 & K_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{e}_{n-1}(t) \\ \mathbf{e}_n(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_{n-1}(t) \\ \mathbf{y}_n(t) \end{bmatrix} &= \begin{bmatrix} C_{(n-1)(n-1)} & C_{(n-1)n} \\ 0 & C_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{n-1}(t) \\ \mathbf{x}_n(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{n-1}(t) \\ \mathbf{e}_n(t) \end{bmatrix}. \end{aligned} \quad (4.16)$$

Defining a partition $\{p_i, r_i\}_{i=n-1}^n$ where $p_i = \dim(\mathbf{x}_i)$ for $i = n-1, n$, it follows that the representation (4.16) has a causal $G|_{\{n-1, n\}}$ -zero structure.

If $G|_{\{n-1, n\}} = (\{n-1, n\}, \emptyset)$, then by assumption \mathbf{y}_n and \mathbf{y}_{n-1} do not Granger cause each other. Hence, by using Lemma 4.23 for S_1 and $[y_{n-1}^T, y_n^T]^T$, we obtain a minimal Kalman representation of $[y_{n-1}^T, y_n^T]^T$ that is an extension of S_1 for

$[\mathbf{y}_{n-1}^T, \mathbf{y}_n^T]^T$ in block diagonal form

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_{n-1}(t+1) \\ \mathbf{x}_n(t+1) \end{bmatrix} &= \begin{bmatrix} A_{(n-1)(n-1)} & 0 \\ 0 & A_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{n-1}(t) \\ \mathbf{x}_n(t) \end{bmatrix} + \begin{bmatrix} K_{(n-1)(n-1)} & 0 \\ 0 & K_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{e}_{n-1}(t) \\ \mathbf{e}_n(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_{n-1}(t) \\ \mathbf{y}_n(t) \end{bmatrix} &= \begin{bmatrix} C_{(n-1)(n-1)} & 0 \\ 0 & C_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{n-1}(t) \\ \mathbf{x}_n(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{n-1}(t) \\ \mathbf{e}_n(t) \end{bmatrix}. \end{aligned} \quad (4.17)$$

Defining a partition $\{p_i, r_i\}_{i=n-1}^n$ where $p_i = \dim(\mathbf{x}_i)$ for $i = n-1, n$, we can see that the representation (4.17) has a causal $G|_{\{n-1, n\}}$ -zero structure.

Suppose that we have an observable Kalman representation $S_j = (A, K, C, I, \mathbf{e})$ of $[\mathbf{y}_{n-j}^T, \dots, \mathbf{y}_n^T]^T$, $j \in \{1, \dots, n-2\}$ with a $G|_{\{n-j, \dots, n\}}$ -zero structure with respect to a partition $\{p_i, r_i\}_{i=n-j}^n$, i.e., the system matrices are

$$\begin{aligned} A &= \begin{bmatrix} A_{n-j, n-j} & A_{n-j, n-j+1} & \cdots & A_{n-j, n} \\ A_{n-j+1, n-j} & A_{n-j+1, n-j+1} & \cdots & A_{n-j+1, n} \\ \vdots & \vdots & & \vdots \\ A_{n, n-j} & A_{n, n-j+1} & \cdots & A_{n, n} \end{bmatrix} \\ K &= \begin{bmatrix} K_{n-j, n-j} & K_{n-j, n-j+1} & \cdots & K_{n-j, n} \\ K_{n-j+1, n-j} & K_{n-j+1, n-j+1} & \cdots & K_{n-j+1, n} \\ \vdots & \vdots & & \vdots \\ K_{n, n-j} & K_{n, n-j+1} & \cdots & K_{n, n} \end{bmatrix} \\ C &= \begin{bmatrix} C_{n-j, n-j} & C_{n-j, n-j+1} & \cdots & C_{n-j, n} \\ C_{n-j+1, n-j} & C_{n-j+1, n-j+1} & \cdots & C_{n-j+1, n} \\ \vdots & \vdots & & \vdots \\ C_{n, n-j} & C_{n, n-j+1} & \cdots & C_{n, n} \end{bmatrix} \end{aligned} \quad (4.18)$$

such that if $t, s \in \{n-j, \dots, n\}$, $(t, s) \notin E$ then $A_{st} = 0, K_{st} = 0, C_{st} = 0$. We will show that S_j can be extended to a representation of $[\mathbf{y}_{n-j-1}^T, \dots, \mathbf{y}_n^T]^T$ with a $G|_{\{n-j-1, \dots, n\}}$ -zero structure with a partition $\{p_i, r_i\}_{i=n-j-1}^n$. Note that the state process \mathbf{x} of S_j is partitioned by $\mathbf{x} = [\mathbf{x}_{n-j}^T, \dots, \mathbf{x}_n^T]^T$ where $\mathbf{x}_i \in \mathbb{R}^{p_i}$. For convenience, define $i := n-j-1$ and let the set of parent and non-parent nodes of i be $I_i = \{i_1, \dots, i_k\}$ and $\bar{I}_i = \{\bar{i}_1, \dots, \bar{i}_l\}$. Accordingly, we denote the subprocesses $\mathbf{x}_{I_i} = [\mathbf{x}_{i_1}^T, \dots, \mathbf{x}_{i_k}^T]^T$, $\mathbf{x}_{\bar{I}_i} = [\mathbf{x}_{\bar{i}_1}^T, \dots, \mathbf{x}_{\bar{i}_l}^T]^T$, $\mathbf{y}_{I_i} = [\mathbf{y}_{i_1}^T, \dots, \mathbf{y}_{i_k}^T]^T$, $\mathbf{y}_{\bar{I}_i} = [\mathbf{y}_{\bar{i}_1}^T, \dots, \mathbf{y}_{\bar{i}_l}^T]^T$ and $\mathbf{e}_{I_i} = [\mathbf{e}_{i_1}^T, \dots, \mathbf{e}_{i_k}^T]^T$, $\mathbf{e}_{\bar{I}_i} = [\mathbf{e}_{\bar{i}_1}^T, \dots, \mathbf{e}_{\bar{i}_l}^T]^T$.

Notice that because of $I_i \cup \bar{I}_i = \{i+1, \dots, n\}$, we can define permutation matrices $P_{\mathbf{y}}$ and $P_{\mathbf{x}}$ such that $[\mathbf{y}_{I_i}^T, \mathbf{y}_{\bar{I}_i}^T]^T = P_{\mathbf{y}}[\mathbf{y}_{i+1}^T, \dots, \mathbf{y}_n^T]^T$, $[\mathbf{e}_{I_i}^T, \mathbf{e}_{\bar{I}_i}^T]^T = P_{\mathbf{y}}[\mathbf{e}_{i+1}^T, \dots, \mathbf{x}_n^T]^T$ and $[\mathbf{x}_{\bar{I}_i}^T, \mathbf{x}_{I_i}^T]^T = P_{\mathbf{x}}[\mathbf{x}_{i+1}^T, \dots, \mathbf{x}_n^T]^T$. If neither I_i nor \bar{I}_i is the empty set then the

tuple $(P_x A P_x^{-1}, P_x K P_y^{-1}, P_y C P_x^{-1}, I, P_y \mathbf{e})$ is an observable Kalman representation of $[\mathbf{y}_{\bar{I}_i}^T, \mathbf{y}_{I_i}^T]^T$ with state process $P_x \mathbf{x}$ in the form of

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_{\bar{I}_i}(t+1) \\ \mathbf{x}_{I_i}(t+1) \end{bmatrix} &= \underbrace{\begin{bmatrix} A_{\bar{I}_i \bar{I}_i} & A_{I_i \bar{I}_i} \\ 0 & A_{I_i I_i} \end{bmatrix}}_{\tilde{A}} \begin{bmatrix} \mathbf{x}_{\bar{I}_i}(t) \\ \mathbf{x}_{I_i}(t) \end{bmatrix} + \underbrace{\begin{bmatrix} K_{\bar{I}_i \bar{I}_i} & K_{I_i \bar{I}_i} \\ 0 & K_{I_i I_i} \end{bmatrix}}_{\tilde{K}} \begin{bmatrix} \mathbf{e}_{\bar{I}_i}(t) \\ \mathbf{e}_{I_i}(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_{\bar{I}_i}(t) \\ \mathbf{y}_{I_i}(t) \end{bmatrix} &= \underbrace{\begin{bmatrix} C_{\bar{I}_i \bar{I}_i} & C_{I_i \bar{I}_i} \\ 0 & C_{I_i I_i} \end{bmatrix}}_{\tilde{C}} \begin{bmatrix} \mathbf{x}_{\bar{I}_i}(t) \\ \mathbf{x}_{I_i}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{\bar{I}_i}(t) \\ \mathbf{e}_{I_i}(t) \end{bmatrix}, \end{aligned} \quad (4.19)$$

where

$$A_{I_i} = \begin{bmatrix} A_{i_1, i_1} & \cdots & A_{i_1, i_k} \\ \vdots & \ddots & \vdots \\ A_{i_k, i_1} & \cdots & A_{i_k, i_k} \end{bmatrix}, \quad A_{\bar{I}_i} = \begin{bmatrix} A_{\bar{i}_1, \bar{i}_1} & \cdots & A_{\bar{i}_1, \bar{i}_l} \\ \vdots & \ddots & \vdots \\ A_{\bar{i}_l, \bar{i}_1} & \cdots & A_{\bar{i}_l, \bar{i}_l} \end{bmatrix}$$

and $K_{I_i}, K_{\bar{I}_i}, C_{I_i}$ and $C_{\bar{I}_i}$ are defined in a similar manner, see also Definition 4.7 and Notation 4.8. From Lemma 4.29, the conditions of Lemma 4.26 are satisfied for the process $[\mathbf{y}_i^T, \mathbf{y}_{\bar{I}_i}^T, \mathbf{y}_{I_i}^T]^T$ and for the Kalman representation $(\tilde{A}, \tilde{K}, \tilde{C}, I, [\mathbf{e}_{\bar{I}_i}^T, \mathbf{e}_{I_i}^T]^T)$ of $[\mathbf{y}_{\bar{I}_i}^T, \mathbf{y}_{I_i}^T]^T$. Then, by Lemma 4.26, there exist an observable Kalman representation that is an extension of (4.19) for $[\mathbf{y}_i^T, \mathbf{y}_{\bar{I}_i}^T, \mathbf{y}_{I_i}^T]^T$ in coordinated form

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_i(t+1) \\ \mathbf{x}_{\bar{I}_i}(t+1) \\ \mathbf{x}_{I_i}(t+1) \end{bmatrix} &= \begin{bmatrix} A_{ii} & 0 & A_{i I_i} \\ 0 & A_{\bar{I}_i \bar{I}_i} & A_{\bar{I}_i I_i} \\ 0 & 0 & A_{I_i I_i} \end{bmatrix} \begin{bmatrix} \mathbf{x}_i(t) \\ \mathbf{x}_{\bar{I}_i}(t) \\ \mathbf{x}_{I_i}(t) \end{bmatrix} + \begin{bmatrix} K_{ii} & 0 & K_{i I_i} \\ 0 & K_{\bar{I}_i \bar{I}_i} & K_{\bar{I}_i I_i} \\ 0 & 0 & K_{I_i I_i} \end{bmatrix} \begin{bmatrix} \mathbf{e}_i(t) \\ \mathbf{e}_{\bar{I}_i}(t) \\ \mathbf{e}_{I_i}(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_i(t) \\ \mathbf{y}_{\bar{I}_i}(t) \\ \mathbf{y}_{I_i}(t) \end{bmatrix} &= \begin{bmatrix} C_{ii} & 0 & C_{i I_i} \\ 0 & C_{\bar{I}_i \bar{I}_i} & C_{\bar{I}_i I_i} \\ 0 & 0 & C_{I_i I_i} \end{bmatrix} \begin{bmatrix} \mathbf{x}_i(t) \\ \mathbf{x}_{\bar{I}_i}(t) \\ \mathbf{x}_{I_i}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_i(t) \\ \mathbf{e}_{\bar{I}_i}(t) \\ \mathbf{e}_{I_i}(t) \end{bmatrix}. \end{aligned} \quad (4.20)$$

Notice that with the inverse permutation matrices P_x^{-1} and P_y^{-1} the following holds: $P_y^{-1}[\mathbf{y}_{\bar{I}_i}^T, \mathbf{y}_{I_i}^T]^T = [\mathbf{y}_{i+1}^T, \dots, \mathbf{y}_n^T]^T$, $P_y^{-1}[\mathbf{e}_{\bar{I}_i}^T, \mathbf{e}_{I_i}^T]^T = [\mathbf{e}_{i+1}^T, \dots, \mathbf{e}_n^T]^T$, and $P_x^{-1}[\mathbf{x}_{\bar{I}_i}^T, \mathbf{x}_{I_i}^T]^T = [\mathbf{x}_{i+1}^T, \dots, \mathbf{x}_n^T]^T$. Then, we can use P_x^{-1} and P_y^{-1} to transform (4.19) into a representation

$$(P_x^{-1} \tilde{A} P_x, P_x^{-1} \tilde{K} P_y, P_y^{-1} \tilde{C} P_x, I, [\mathbf{e}_{i+1}^T, \dots, \mathbf{e}_n^T]^T)$$

of $[\mathbf{y}_{i+1}^T, \dots, \mathbf{y}_n^T]^T$. Define the ordered set $J := (i+1, \dots, n)$. Then, combining

$$(P_x^{-1} \tilde{A} P_x, P_x^{-1} \tilde{K} P_y, P_y^{-1} \tilde{C} P_x, I, [\mathbf{e}_{i+1}^T, \dots, \mathbf{e}_n^T]^T)$$

and (4.20), we obtain a representation in the form of

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_i(t+1) \\ \mathbf{x}_J(t+1) \end{bmatrix} &= \underbrace{\begin{bmatrix} A_{ii} [A_{i\bar{I}_i} \ A_{iI_i}] P_{\mathbf{x}} \\ 0 \quad P_{\mathbf{x}}^{-1} \tilde{A} P_{\mathbf{x}} \end{bmatrix}}_{\hat{A}} \begin{bmatrix} \mathbf{x}_i(t) \\ \mathbf{x}_J(t) \end{bmatrix} + \underbrace{\begin{bmatrix} K_{ii} [K_{i\bar{I}_i} \ K_{iI_i}] P_{\mathbf{y}} \\ 0 \quad P_{\mathbf{x}}^{-1} \tilde{K} P_{\mathbf{y}} \end{bmatrix}}_{\hat{K}} \begin{bmatrix} \mathbf{e}_i(t) \\ \mathbf{e}_J(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_i(t) \\ \mathbf{y}_J(t) \end{bmatrix} &= \underbrace{\begin{bmatrix} C_{ii} [C_{i\bar{I}_i} \ C_{iI_i}] P_{\mathbf{x}} \\ 0 \quad P_{\mathbf{y}}^{-1} \tilde{C} P_{\mathbf{x}} \end{bmatrix}}_{\hat{C}} \begin{bmatrix} \mathbf{x}_i(t) \\ \mathbf{x}_J(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_i(t) \\ \mathbf{e}_J(t) \end{bmatrix}. \end{aligned} \quad (4.21)$$

Notice that $[\mathbf{e}_i^T, \mathbf{e}_J^T]^T$ is the innovation process of $[\mathbf{y}_i^T, \mathbf{y}_J^T]^T$. Furthermore, $A = P_{\mathbf{x}}^{-1} \tilde{A} P_{\mathbf{x}}$, $K = P_{\mathbf{x}}^{-1} \tilde{K} P_{\mathbf{y}}$, $C = P_{\mathbf{y}}^{-1} \tilde{C} P_{\mathbf{x}}$ and (A, C) are observable. Therefore, the representation (4.21) is an observable Kalman representation of $[\mathbf{y}_i^T, \mathbf{y}_J^T]^T$. Furthermore, $\hat{A}_{nm} = 0$, $\hat{K}_{nm} = 0$ and $\hat{C}_{nm} = 0$ for any $n \in \bar{I}_i$ and thus since S_j (see (4.18)) has $G|_{n-j, \dots, n}$ -zero structure, we have that for any $t, s \in \{i, \dots, n\}$, $(t, s) \notin E$, $\hat{A}_{st} = 0$, $\hat{K}_{st} = 0$, $\hat{C}_{st} = 0$. For conclusions, recall that $i = n - j - 1$. Then, (4.21) is an observable Kalman representation of $[\mathbf{y}_{n-j-1}^T, \dots, \mathbf{y}_n^T]^T$ with $G|_{\{n-j-1, \dots, n\}}$ -zero structure for the partitioning $\{p_l, r_l\}_{l=n-j-1}^n$ where $p_l = \dim(\mathbf{x}_l)$ and $r_l = \dim(\mathbf{y}_l)$ for $l = n - j - 1, \dots, n$. Note that if $\bar{I}_{n-j-1} = \emptyset$ or $I_{n-j-1} = \emptyset$ then (4.21) is an observable Kalman representation of $\mathbf{y}_{n-j-1, \dots, n}$ in block triangular or block diagonal form, respectively. With this, we showed that if \mathbf{y} has a G -consistent causality structure then an observable Kalman representation of $[\mathbf{y}_{n-j}^T, \dots, \mathbf{y}_n^T]^T$ with $G|_{\{n-j, \dots, n\}}$ -zero structure can be extended to an observable Kalman representation of $[\mathbf{y}_{n-j-1}^T, \dots, \mathbf{y}_n^T]^T$ in $G|_{\{n-j-1, \dots, n\}}$ -zero structure. Applying this to $j = 2, \dots, n - 2$, we obtain that there exists an observable Kalman representation of \mathbf{y} with G -zero structure. Hence, the implication (i) \implies (v) in Theorem 4.15 follows.

(ii) \implies (iii): Assume now that condition (ii) holds and notice that we know from the proof of (i) \implies (v) that there exists a minimal Kalman representation of $[\mathbf{y}_{n-1}^T, \mathbf{y}_n^T]^T$ in causal $G|_{n-1, n}$ -zero structure (see (4.16) and (4.17)). Then, by induction, we suppose that there exists a minimal Kalman representation $S_k = (A, K, C, I, \mathbf{e})$ of $[\mathbf{y}_{n-k}^T, \dots, \mathbf{y}_n^T]^T$ with a causal $G|_{\{n-k, \dots, n\}}$ -zero structure with respect to a partition $\{p_l, r_l\}_{l=n-k}^n$ for all $k = 1, \dots, j$ where $j \leq n - 2$. By applying the same permutation transformation on the system matrices and processes of S_j as in the proof of (i) \implies (v), we obtain that (4.19) is minimal. In addition, we can show that (4.19) is also in causal block triangular form: Assume indirectly that (4.19) is not in causal block triangular form. With the same notation as in the proof of (i) \implies (v), notice that \mathbf{e}_{I_i} is the innovation process of \mathbf{y}_{I_i} , thus $(A_{I_i I_i}, K_{I_i I_i}, C_{I_i I_i}, I, \mathbf{e}_{I_i})$ is a Kalman representation of \mathbf{y}_{I_j} . Therefore, it only can happen that (4.19) is not in causal block triangular form if $(A_{I_i I_i}, K_{I_i I_i}, C_{I_i I_i}, I, \mathbf{e}_{I_i})$ is not minimal. Take a minimal Kalman representation $(\hat{A}_{I_i I_i}, \hat{K}_{I_i I_i}, \hat{C}_{I_i I_i}, I, \mathbf{e}_{I_i})$ of \mathbf{y}_{I_j} with state pro-

cess $\tilde{\mathbf{x}}_{I_j}$. For the state process $\tilde{\mathbf{x}}_{I_j}$ it is then true that for any N larger than or equal to the dimension of $\tilde{\mathbf{x}}_{I_j}$, $\tilde{\mathbf{x}}_{I_j}(t) = \hat{\mathcal{O}}_N^+ E_l[Y_{I_j}(t) | \mathcal{H}_{t-}^{\mathbf{y}}]$, where $\hat{\mathcal{O}}_N^+$ is the left inverse of the finite (up to N) observability matrix of $(\tilde{A}_{I_j I_j}, \tilde{C}_{I_j I_j})$ and $Y_{I_j}(t) = [\mathbf{y}_{I_j}^T(t), \mathbf{y}_{I_j}^T(t+1), \dots, \mathbf{y}_{I_j}^T(t+N)]$. Define $T = \hat{\mathcal{O}}_N^+ \mathcal{O}_N$ where \mathcal{O}_N is the finite (up to N) observability matrix of $(A_{I_j I_j}, C_{I_j I_j})$ and note that $\mathbf{x}_{I_j}(t) = \mathcal{O}_N^+ E_l[Y_{I_j}(t) | \mathcal{H}_{t-}^{\mathbf{y}}]$. Then, we have that $\mathbf{x}_{I_j}(t) = T \tilde{\mathbf{x}}_{I_j}(t)$ and the representation

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_{\bar{I}_i}(t+1) \\ \tilde{\mathbf{x}}_{I_i}(t+1) \end{bmatrix} &= \begin{bmatrix} A_{\bar{I}_i \bar{I}_i} & A_{I_i \bar{I}_i} T \\ 0 & A_{I_i I_i} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{\bar{I}_i}(t) \\ \tilde{\mathbf{x}}_{I_i}(t) \end{bmatrix} + \begin{bmatrix} K_{\bar{I}_i \bar{I}_i} & K_{I_i \bar{I}_i} \\ 0 & \tilde{K}_{I_i I_i} \end{bmatrix} \begin{bmatrix} \mathbf{e}_{\bar{I}_i}(t) \\ \mathbf{e}_{I_i}(t) \end{bmatrix} \\ \begin{bmatrix} \mathbf{y}_{\bar{I}_i}(t) \\ \mathbf{y}_{I_i}(t) \end{bmatrix} &= \begin{bmatrix} C_{\bar{I}_i \bar{I}_i} & C_{I_i \bar{I}_i} T \\ 0 & C_{I_i I_i} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{\bar{I}_i}(t) \\ \tilde{\mathbf{x}}_{I_i}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{\bar{I}_i}(t) \\ \mathbf{e}_{I_i}(t) \end{bmatrix}. \end{aligned}$$

is a Kalman representation of $[\mathbf{y}_{\bar{I}_i}^T, \mathbf{y}_{I_i}^T]^T$ with smaller dimension than (4.19). This is a contradiction. Therefore, $(A_{I_i I_i}, K_{I_i I_i}, C_{I_i I_i}, I, \mathbf{e}_{I_i})$ is minimal and (4.19) is in causal block triangular form. By using this and that for $l \neq m, l, m = j, \bar{I}_j$

$$E_l[\mathcal{H}_{t+}^{\mathbf{y}^l} | \mathcal{H}_{t-}^{\mathbf{y}^l, \mathbf{y}^{I_j}}] \cap E_l[\mathcal{H}_{t+}^{\mathbf{y}^m} | \mathcal{H}_{t-}^{\mathbf{y}^m, \mathbf{y}^{I_j}}] | E_l[\mathcal{H}_{t+}^{\mathbf{y}^{I_j}} | \mathcal{H}_{t-}^{\mathbf{y}^{I_j}}] = \{0\},$$

we obtain that by Lemma 4.26, the representation (4.20) is a minimal Kalman representation in causal coordinated form. Since the dimension of the state processes in (4.20) and in (4.21) are the same, (4.21) is also minimal which implies that (ii) \implies (iii) in Theorem 4.15 holds.

(iv) \implies (i): Since \mathbf{y} has a Kalman representation with causal G -zero structure, if we define the tuple $J := (j, \bar{I}_j, I_j)$ for a node $j \in V$, then $(A_{JJ}, K_{JJ}, C_{JJ}, I, [\mathbf{e}_j^T, \mathbf{e}_{\bar{I}_j}^T, \mathbf{e}_{I_j}^T]^T)$ is a Kalman representation of $[\mathbf{y}_j^T, \mathbf{y}_{\bar{I}_j}^T, \mathbf{y}_{I_j}^T]^T$ in causal coordinated form. By Theorem 3.5, it is equivalent to the following conditions:

- \mathbf{y}_j does not Granger cause \mathbf{y}_{I_j}
- $\mathbf{y}_{\bar{I}_j}$ does not Granger cause \mathbf{y}_{I_j}
- \mathbf{y}_j conditionally does not Granger cause $\mathbf{y}_{\bar{I}_j}$ with respect to \mathbf{y}_{I_j}
- $\mathbf{y}_{\bar{I}_j}$ conditionally does not Granger cause \mathbf{y}_j with respect to \mathbf{y}_{I_j}

Then, by Lemma 4.13 these conditions are equivalent to that \mathbf{y} has G -consistent causality structure which completes the proof. \square

Proof of Corollary 4.30. Notice that the steps of the proof of Theorem 4.15 follow the steps of Algorithm 11. Assume that the minimal Kalman representation of \mathbf{y}_n in the proof of Theorem 4.15 is the minimal Kalman representation of \mathbf{y}_n that is calculated in Step 1 of Algorithm 12. Furthermore, notice that the extended Kalman representations in Lemmas 4.19, 4.23, and 4.26 are constructed based on Algorithms 8,

9, and 10, respectively. From this it follows that the representation constructed in the proof of Theorem 4.15 (i) \implies (v) coincides with the Kalman representation (A, K, C, I, e) where A, K, C are the output matrices of Algorithm 11. This completes the proof. \square

