

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 2

Granger causality and Kalman representations in block triangular form

Kalman representations, introduced in Section 1.2, are LTI-SS representations where the noise process is the innovation process of the output process. The noise and also the state processes of these representations can be expressed by using the output process, based on projections in Hilbert spaces generated by random variables of the output process. This interpretation of the noise and state processes helps us to connect the network graph of Kalman representations to properties of the output process. In this chapter we consider Kalman representations with a specific network graph which has two nodes and one directed edge. We call Kalman representations with this network graph, Kalman representations in block triangular form. Then we relate the existence of a Kalman representations in block triangular form to properties of their output processes as follows: Given a partition $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ of a process \mathbf{y} , there exists a Kalman representation of \mathbf{y} in block triangular form if and only if \mathbf{y}_1 does not Granger cause \mathbf{y}_2 . Kalman representations in block triangular form have several advantageous properties, e.g., they allow for distributed parameter and state estimation.

In (Granger, 1963), Granger causality has already been characterized by properties of VAR models. Also, the papers (Caines and Chan, 1975; Caines, 1976; Gevers and Anderson, 1982) relate Granger causality to network graphs of MA models. Therefore, the results of Chapter 2 can be viewed as a counterpart of the results in the cited papers for LTI-SS representations. The results presented here are not the first results on Granger causality in terms of LTI-SS representations. In (Barnett and Seth, 2015; Solo, 2016) Granger causality was characterized by properties of LTI-SS representation using transfer function approach. Contrary to (Barnett and Seth, 2015; Solo, 2016), we give a characterization for Granger non-causality by constructing Kalman representations in block triangular form. This chapter together with Chapter 3 is based on the journal paper (Jozsa et al., 2018b) and conference abstract (Jozsa et al., 2016).

The structure of this chapter is as follows: First, we introduce Kalman representations in block triangular form. Then, we characterize their existence in terms

of Granger causality. This will be followed by the presentation of two realization algorithms that calculate Kalman representations in block triangular form. Finally, we provide an example to illustrate the results. The proofs of the statements can be found in Appendix 2.A. If not stated otherwise, we assume throughout this chapter that $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ is a ZMSIR process where $\mathbf{y}_i \in \mathbb{R}^{r_i}$, and $r_i > 0$ for $i = 1, 2$.

2.1 Kalman representation in block triangular form

In this section we introduce Kalman representations in block triangular form and discuss their properties. To begin with, we define this class of representations:

Definition 2.1. A Kalman representation $(A, K, C, I, \mathbf{e} = [\mathbf{e}_1^T, \mathbf{e}_2^T]^T, \mathbf{y})$, where $\mathbf{e}_i \in \mathbb{R}^{r_i}$, $i = 1, 2$, is called a *Kalman representation in block triangular form*, if

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \quad K = \begin{bmatrix} K_{11} & K_{12} \\ 0 & K_{22} \end{bmatrix} \quad C = \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{bmatrix}, \quad (2.1)$$

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 $p_i \geq 0$ for $i, j = 1, 2$. If, in addition, $(A_{22}, K_{22}, C_{22}, I_{r_2}, \mathbf{e}_2)$ is a minimal Kalman representation of \mathbf{y}_2 , then $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is called a *Kalman representation in causal block triangular form*.

Remark 2.2. If $p_2 = 0$ in Definition 2.1, then the block matrices $A_{12}, A_{22}, K_{22}, C_{12}$ and C_{22} are absent, whereas if $p_1 = 0$, then $A_{11}, A_{12}, K_{11}, K_{12}$ and C_{11} are absent. Furthermore, if $p_2 = 0$, then $\mathbf{y}_2 = \mathbf{e}_2$ is a white noise process and if $p_1 = p_2 = 0$, then $\mathbf{y} = \mathbf{e}$ is a white noise process. In both cases, block triangular form implies causal block triangular form. A minimal Kalman representation (A, K, C, I, \mathbf{e}) of a white noise process \mathbf{y} has zero dimension (A, K, C are absent) and it is the trivial equation $\mathbf{y} = \mathbf{e}$.

If $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is a minimal Kalman representation in causal block triangular form satisfying (2.1) then the dimensions of the block matrices $A_{ij} \in \mathbb{R}^{p_i \times p_j}$, $K_{ij} \in \mathbb{R}^{p_i \times r_j}$, $C_{ij} \in \mathbb{R}^{r_i \times p_j}$, $i, j = 1, 2$ are uniquely determined by $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$. Indeed, $(A_{22}, K_{22}, C_{22}, I_{r_2}, \mathbf{e}_2)$ is a minimal Kalman representation of \mathbf{y}_2 thus p_2 is the dimension of a minimal LTI-SS representation of \mathbf{y}_2 . Furthermore, $p_1 = p - p_2$, where p is the dimension of a minimal LTI-SS representation of \mathbf{y} . That is, the dimension of a minimal LTI-SS representation of \mathbf{y}_2 and \mathbf{y} and the dimensions of \mathbf{y}_1 and \mathbf{y}_2 determine the dimensions of A_{ij}, K_{ij}, C_{ij} , $i, j = 1, 2$.

Kalman representations in block triangular form can be viewed as a cascade interconnection of two subsystems, see Figure 2.1. More precisely, let (A, K, C, I, \mathbf{e}) be a Kalman representation of \mathbf{y} in block triangular form satisfying (2.1) and let

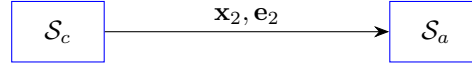


Figure 2.1: Network graph of a Kalman representation in block triangular form: \mathcal{S}_c is the coordinator (2.2), \mathcal{S}_a is the agent (2.3).

$\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T]^T$ be its state process where $\mathbf{x}_i \in \mathbb{R}^{p_i}$, $i = 1, 2$. Then we can define the dynamical systems \mathcal{S}_c and \mathcal{S}_a below.

$$\mathcal{S}_c \begin{cases} \mathbf{x}_2(t+1) = A_{22}\mathbf{x}_2(t) + K_{22}\mathbf{e}_2(t) \\ \mathbf{y}_2(t) = C_{22}\mathbf{x}_2(t) + \mathbf{e}_2(t) \end{cases} \quad (2.2)$$

$$\mathcal{S}_a \begin{cases} \mathbf{x}_1(t+1) = \sum_{i=1}^2 (A_{1i}\mathbf{x}_i(t) + K_{1i}\mathbf{e}_i(t)) \\ \mathbf{y}_1(t) = \sum_{i=1}^2 C_{1i}\mathbf{x}_i(t) + \mathbf{e}_2(t) \end{cases} \quad (2.3)$$

The subsystem \mathcal{S}_c , which generates \mathbf{y}_2 , will be called *coordinator*, and the subsystem \mathcal{S}_a , which generates \mathbf{y}_1 , will be called *agent*. The coordinator sends its state \mathbf{x}_2 and noise \mathbf{e}_2 to the agent while the agent does not send information to the coordinator. Accordingly, the network graph of $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is the two-node star graph with \mathcal{S}_c being the root node and \mathcal{S}_a being the leaf.

Motivation for Kalman representations in causal block triangular form

If we considered an LTI-SS representation of $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ without requiring it to be a Kalman representation, then in general the subsystems \mathcal{S}_1 with output \mathbf{y}_1 and \mathcal{S}_2 with output \mathbf{y}_2 could change information in both direction via the noise process. The fact that we require the LTI-SS representation to be a Kalman representation in causal block triangular form implies that this cannot be the case: notice that the second noise component \mathbf{e}_2 is the innovation process of \mathbf{y}_2 which implies that \mathbf{e}_2 and \mathbf{x}_2 depend only on the past and present values of \mathbf{y}_2 . In contrast, the first noise and state components \mathbf{e}_1 and \mathbf{x}_1 depend on the past and present values of both \mathbf{y}_1 and \mathbf{y}_2 . This ensures that, indeed, a Kalman representation in causal block triangular form means that there is no communication from the agent to the coordinator.

Kalman representations in causal block triangular form guarantee *the subsystem which corresponds to the coordinator to be minimal* and thus it is unique up to isomorphism (see Definition 1.11 and Proposition 1.12). An advantage of minimality is that it implies observability, and hence the state \mathbf{x}_2 can be estimated from \mathbf{y}_2 . This opens up the perspective of distributed estimation, and possibly, with the future inclusions of inputs, of distributed control. An example for not requiring the subsystem of the coordinator to be minimal: Let us consider a Kalman representation

$(A, K, C, I, \mathbf{e}, \mathbf{y})$ in block triangular form with state process $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T]^T$, where $\mathbf{y}_1, \mathbf{y}_2, \mathbf{e}_1, \mathbf{e}_2, \mathbf{x}_1 \in \mathbb{R}, \mathbf{x}_2 \in \mathbb{R}^2$ and the system matrices are given as below.

$$A = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.7 \end{bmatrix}, K = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

The coordinator is then the subsystem with system matrices

$$A_{22} = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.7 \end{bmatrix}, K_{22} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, C_{22} = [0 \ 1].$$

Since the coordinator is not minimal, this Kalman representation in block triangular form is not in causal block triangular form. What happens is that the coordinator contains dynamics which could also be made part of the agent. For this example, it is not enough to know y_2 to estimate the state of the coordinator, for that we have to use the output y_1 of the agent. Imposing minimality on the coordinator avoids such degenerate cases.

2.2 Characterization of Granger non-causality by Kalman representations in block triangular form

In this section we show that the existence of a Kalman representation in causal block triangular form characterizes Granger non-causality.

The next definition is a particular case of the concept of causality between stochastic processes defined in (Granger, 1963), if the latter is applied to ZMSIR processes, and if, using the terminology of (Granger, 1963), there is no external process.

Definition 2.3 (Granger non-causality). Consider a ZMSIR process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$. We say that \mathbf{y}_1 does not Granger cause \mathbf{y}_2 if for all $t, k \in \mathbb{Z}, k \geq 0$

$$E_l[\mathbf{y}_2(t+k) \mid \mathcal{H}_{t-}^{\mathbf{y}_2}] = E_l[\mathbf{y}_2(t+k) \mid \mathcal{H}_{t-}^{\mathbf{y}}].$$

Otherwise, we say that \mathbf{y}_1 Granger causes \mathbf{y}_2 .

Informally, \mathbf{y}_1 does not Granger cause \mathbf{y}_2 , if for all $k \geq 0$, the best k -step linear prediction of \mathbf{y}_2 based on the past values of \mathbf{y}_2 is the same as that of based on the past values of \mathbf{y} . Note that Definition 2.3 is equivalent to the weakly feedback free property of processes, see (Caines, 1988, Definition 2.1, Chapter 10). The notion of

strongly feedback free property is equivalent to that for all $k \geq 0$

$$E_l[\mathbf{y}_2(t+k) \mid \mathcal{H}_{t-}^{\mathbf{y}_2}] = E_l[\mathbf{y}_2(t+k) \mid \mathcal{H}_{t-}^{\mathbf{y}} + \mathcal{H}_t^{\mathbf{y}_1}], \quad (2.4)$$

i.e., when the k -step linear prediction of \mathbf{y}_2 based on the past values of \mathbf{y}_2 is the same as that of based on the past values of \mathbf{y} and the present values of \mathbf{y}_1 . In (Granger, 1963), causality was defined as in Definition 2.3, however, to a subclass of ZMSIR processes where Definition 2.3 implies (2.4). In the rest of the thesis we do not deal with causality as in (2.4).

Remark 2.4 (Related work). If \mathbf{y} is coercive, Granger non-causality from \mathbf{y}_1 to \mathbf{y}_2 is further equivalent to, (see (Barnett and Seth, 2015))

$$\forall k \geq 0 : (C(A - KC)^k K)_{21} = 0, \quad (2.5)$$

where $(\cdot)_{21}$ denotes the $r_2 \times r_1$ left lower block of a matrix and (A, K, C, I, \mathbf{e}) is a minimal Kalman representation of \mathbf{y} . The system matrices of a minimal Kalman representation $(A, K, C, I, \mathbf{e}, \mathbf{y})$ in block triangular form naturally satisfy (2.5). Moreover, they assume block triangular Wold decomposition, i.e., block triangular transfer matrix between the innovation process \mathbf{e} and \mathbf{y} , see (Hsiao, 1982; Caines, 1976; Barnett and Seth, 2015) and (Caines, 1988, Theorem 2.2, Chapter 10). Note that in the cited papers, the reader can find equivalent formulation of block triangular Wold decomposition, e.g., in terms of feedback free property of feedback systems. In contrast to the cited papers, Theorem 2.5 below covers the case when \mathbf{y} is non-coercive and in its proof the LTI-SS representation $(A, K, C, I, \mathbf{e}, \mathbf{y})$ that characterizes Granger non-causality in \mathbf{y} is constructed.

Below, we state the main result of this chapter :

Theorem 2.5. Consider the following statements for a ZMSIR process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$:

- (i) \mathbf{y}_1 does not Granger cause \mathbf{y}_2 ;
- (ii) there exists a minimal Kalman representation of \mathbf{y} in causal block triangular form;
- (iii) there exists a minimal Kalman representation of \mathbf{y} in block triangular form;
- (iv) there exists a Kalman representation of \mathbf{y} in block triangular form;

Then (i) \iff (ii). If \mathbf{y} is coercive, then (i) \iff (ii) \iff (iii) \iff (iv).

The **proof** can be found in Appendix 2.A. Intuitively, Granger non-causality in Theorem 2.5 means that for predicting \mathbf{y}_2 , there is nothing to be gained from the knowledge of the past of \mathbf{y}_1 . In parallel, considering a Kalman representation (2.1)

of \mathbf{y} in causal block triangular form the subsystem $(A_{22}, K_{22}, C_{22}, I_{r_2}, \mathbf{e}_2)$ of \mathbf{y}_2 depends only on \mathbf{e}_2 and thus only on the past values of \mathbf{y}_2 . On the other hand, the subsystem which generates \mathbf{y}_1 depends on the entire history of \mathbf{y} . Theorem 2.5 can also be interpreted as follows: Granger non-causality from \mathbf{y}_1 to \mathbf{y}_2 is equivalent to the process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ admitting a minimal Kalman representation with network graph depicted on Figure 2.1.

Theorem 2.5 relates Granger non-causality with existence of *minimal* Kalman representations in block triangular form. Since all minimal Kalman representations of \mathbf{y} are isomorphic, Theorem 2.5 not only guarantees that Granger non-causality translates into existence of a Kalman representation with a suitable network graph, but also guarantees that any minimal Kalman representation of \mathbf{y} is isomorphic to this particular one. Hence, most of the interesting dynamical properties of this Kalman representation are also valid for any other minimal Kalman representation. Furthermore, any minimal Kalman representation can be brought to this specific one via a linear state-space transformation. Since black-box identification algorithms, for example subspace methods, yield minimal Kalman representations, Theorem 2.5 is also interesting for deriving and interpreting network graphs based on data.

2.3 Computing Kalman representation in block triangular form

In this section we present a procedure for constructing a minimal Kalman representation in causal block triangular form.

Consider an LTI-SS representation $(\bar{A}, \bar{B}, \bar{C}, \bar{D}, \mathbf{v})$ of $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$. Then it can be transformed into a minimal Kalman representation $(\hat{A}, \hat{K}, \hat{C}, I, \mathbf{e})$ of \mathbf{y} using Algorithm 2. Now take the partition $\hat{C} = \begin{bmatrix} \hat{C}_1^T & \hat{C}_2^T \end{bmatrix}^T$ such that the number of rows of \hat{C}_i equals $r_i = \dim(\mathbf{y}_i)$ for $i = 1, 2$. Furthermore, define a non-singular matrix T such that $(T\hat{A}T^{-1}, \hat{C}_2T^{-1})$ is in the form (see, e.g., (Rosenbrock, 1970))

$$T\hat{A}T^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad \hat{C}_2T^{-1} = [0 \ C_{22}], \quad (2.6)$$

where (A_{22}, C_{22}) is observable and $A_{11} \in \mathbb{R}^{p_1 \times p_1}$, $A_{22} \in \mathbb{R}^{p_2 \times p_2}$ such that p_2 is the rank of the observability matrix of the pair (\hat{C}_2, \hat{A}) . Note that if (\hat{A}, \hat{C}_2) is observable, then $p_1 = 0$ and A_{11}, A_{12} are absent in (2.6). In addition, if the observability matrix of (\hat{A}, \hat{C}_2) has zero rank, then $p_2 = 0$, in which case A_{12}, A_{22} and C_{22} are absent in

(2.6). Note that such a matrix T always exists, (Rosenbrock, 1970). Define

$$A := T\hat{A}T^{-1}, \quad K := T\hat{K}, \quad C := \hat{C}T^{-1} \quad (2.7)$$

and consider the partition

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}, \quad C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}, \quad (2.8)$$

where $A_{ij} \in \mathbb{R}^{p_i \times p_i}$, $K_{ij} \in \mathbb{R}^{p_i \times r_j}$, $C_{ij} \in \mathbb{R}^{r_i \times p_i}$, $i, j = 1, 2$ and from (2.6), $A_{21} = 0$, $C_{21} = 0$. Based on Theorem 2.5 we can state the following result:

Corollary 2.6. *The following statements hold:*

- If \mathbf{y}_1 does not Granger cause \mathbf{y}_2 , then either K_{21} is absent or $K_{21} = 0$. Furthermore, (A, K, C, I, \mathbf{e}) is a minimal Kalman representation of $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ in causal block triangular form with A_{ij} , K_{ij} , C_{ij} , $i, j = 1, 2$ defined by (2.6), (2.7) and (2.8).
- If \mathbf{y} is coercive, then the absence of K_{21} or $K_{21} = 0$ implies that \mathbf{y}_1 does not Granger cause \mathbf{y}_2 .

The **proof** can be found in Appendix 2.A. Corollary 2.6 yields a method to calculate a minimal Kalman representation in causal block triangular form in the absence of Granger causality. This idea is elaborated in Algorithms 4 and 5 that rely on Algorithms 1 and 2. Algorithm 4 takes an LTI-SS representation as its input and transforms it into a minimal Kalman representation in causal block triangular form. Algorithm 5 calculates the same representation from output covariances. Hence, by using empirical covariances, it can be applied to data. Note that in Algorithms 4–5 the dimensions $r_i = \dim(\mathbf{y}_i)$, $i = 1, 2$ are predefined.

Algorithm 4 Minimal Kalman representation in causal block triangular form based on LTI-SS representation

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

Output $\{A, K, C\}$: System matrices of (2.1)

Step 1 Apply Algorithm 2 with input $\{\bar{A}, \bar{B}, \bar{C}, \bar{D}, \Lambda_0^{\mathbf{v}}\}$ and denote its output by $\{\hat{A}, \hat{K}, \hat{C}\}$.

Step 2 Let $\hat{C} = [\hat{C}_1^T \hat{C}_2^T]^T$ be such that $\hat{C}_i \in \mathbb{R}^{r_i \times n}$. Calculate a non-singular matrix T such that (2.6) holds and (A_{22}, C_{22}) is observable.

Step 3 Set $A := T\hat{A}T^{-1}$, $K := T\hat{K}$, $C := \hat{C}T^{-1}$.

Algorithm 5 Minimal Kalman representation in causal block triangular form based on output covariances

Input $\{\Lambda_k^y\}_{k=0}^{2N}$: Covariance sequence of $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$

Output $\{A, K, C\}$: System matrices of (2.1)

Step 1 Apply Algorithm 1 with input $\{\Lambda_k^y\}_{k=0}^{2N}$ and denote its output by $\{\hat{A}, \hat{K}, \hat{C}\}$.
Step 2 Steps 2–3 of Algorithm 4.

Remark 2.7 (Correctness of Algorithms 4–5). Consider a ZMSIR process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ with covariance sequence $\{\Lambda_k^y\}_{k=0}^\infty$ and an LTI–SS representation $(\bar{A}, \bar{B}, \bar{C}, \bar{D}, \mathbf{v})$ of \mathbf{y} . 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 LTI–SS representation of \mathbf{y} . Assume that \mathbf{y} satisfies condition (i) of Theorem 2.5 and note that Algorithms 1–2 calculate a minimal Kalman representation (Remark 1.8). Then from it follows Corollary 2.6 that if $\{A, K, C\}$ is the output of Algorithm 4 with input $\{\bar{A}, \bar{B}, \bar{C}, \bar{D}, \Lambda_0^y = E[\mathbf{v}(t)\mathbf{v}^T(t)]\}$, then (A, K, C, I, \mathbf{e}) is a minimal Kalman representation in causal block triangular form. Similarly, it follows from Corollary 2.6 that if $\{A, K, C\}$ is the output of Algorithm 5 with input $\{\Lambda_k^y\}_{k=0}^{2N}$, then (A, K, C, I, \mathbf{e}) is a minimal Kalman representation of \mathbf{y} in causal block triangular form.

Remark 2.8. In a similar fashion as in Remark 1.9, Algorithms 5 and 4 have polynomial complexity. Algorithm 4 is polynomial in the dimensions of the state, output and noise processes of the LTI–SS representation $(\bar{A}, \bar{B}, \bar{C}, \bar{D}, \mathbf{v})$. Algorithm 5 is polynomial in the number and size of the output covariances.

Remark 2.9 (Checking Granger non-causality). Algorithms 4 and 5 can be used to check Granger non-causality by looking whether the left lower block of matrix K is zero ($K_{21} = 0$) in the partition (2.8), where $\{A, K, C\}$ are the matrices returned by Algorithm 4 or 5. If $K_{21} \neq 0$, then \mathbf{y}_1 Granger causes \mathbf{y}_2 . If \mathbf{y}_2 is coercive and K_{21} is absent or $K_{21} = 0$, then \mathbf{y}_1 does not Granger cause \mathbf{y}_2 in the view of Corollary 2.6. If \mathbf{y} is non-coercive, then it should be checked if $(A_{22}, K_{22}, C_{22}, I_{r_2}, \mathbf{e}_2)$ is a minimal Kalman representation of \mathbf{y}_2 . This can be done by computing a minimal Kalman representation $(\tilde{A}_{22}, \tilde{K}_{22}, \tilde{C}_{22}, I, \tilde{\mathbf{e}}_2)$ of \mathbf{y}_2 using Algorithm 1 or 2. If the noise variance $E[\mathbf{e}_2(t)\mathbf{e}_2^T(t)]$ is equal to the new noise variance $E[\tilde{\mathbf{e}}_2(t)\tilde{\mathbf{e}}_2^T(t)]$, then from (Dufour and Renault, 1998, Proposition 2.3) we know that \mathbf{y}_1 does not Granger cause \mathbf{y}_2 and $\mathbf{e}_2 = \tilde{\mathbf{e}}_2$.

2.4 Example for block triangular representation

In this section we adopt a simplified version of a case study in (Kempker, 2012, Section 8.1) to illustrate the results of this chapter. The focus of this study is the dynamics of underwater vehicles that track a reference path in a fixed formation. Among the vehicles there is one acting as a coordinator that tracks a reference path and the others acting as agents that track the coordinator. For our purpose, we consider two underwater vehicles that track a reference path in a fixed distance, where one of them is the coordinator and the other is an agent.

In comparison with (Kempker, 2012, Section 8.1), besides the number of vehicles, we made the following changes:

- to ensure stationarity, the coordinator follows the zero position
- for convenience, we consider the movements of the vehicles along the first coordinate
- besides the position disturbance we include measurement noise.

We will show that the relative positions (concerning the fixed distance) of the vehicles are ZMSIR processes that can be modeled by a minimal Kalman representation in causal block triangular form. In fact, we reverse engineer the network graph from the observed process in the following way: Denote the agent vehicle by y_1 and the coordinator vehicle by y_2 . Then we verify that y_1 does not Granger cause y_2 . Based on Remark 2.7, this allows us to calculate a minimal Kalman representation in causal block triangular form using Algorithm 5.

Model description: Assume that we have two underwater vehicles V_1 and V_c where V_1 acts as an agent and V_c as the coordinator. For $j \in \{1, c\}$ denote the first coordinate at time $t \in \mathbb{Z}$ of the position, velocity, acceleration, position disturbance and measurement noise of V_j by $p_j(t)$, $s_j(t)$, $a_j(t)$, $w_j(t)$ and $\tilde{w}_j(t)$, respectively. Also, denote the first coordinate of the reference position and velocity of V_j by $p_j^R(t)$ and $s_j^R(t)$, respectively. Let

$$\begin{aligned} p_c^R(t) &= -(p_c(t) + \tilde{w}_c(t)) \\ p_1^R(t) &= (p_c(t) + \tilde{w}_c(t)) + \Delta_1. \end{aligned}$$

That is, V_c follows the zero position based on its own measured position and V_1 follows V_c in a distance Δ_1 based on the same information.

The dynamics of $[p_j, s_j]^T$, $j \in \{1, c\}$ is given by

$$\begin{bmatrix} p_j(t+1) \\ s_j(t+1) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & \frac{\tau-1}{\tau} \end{bmatrix} \begin{bmatrix} p_j(t) \\ s_j(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{\tau} \end{bmatrix} a_j + \begin{bmatrix} 1 \\ 0 \end{bmatrix} w_j(t) \quad (2.9)$$

where a_j is the control input and τ is a time constant. The reference signals $[p_j^R, s_j^R]^T$ are estimated by an observer with dynamics

$$\begin{bmatrix} \hat{p}_j^R(t+1) \\ \hat{s}_j^R(t+1) \end{bmatrix} = \begin{bmatrix} 1 - G_j^p & 1 \\ -G_j^s & \frac{\tau-1}{\tau} \end{bmatrix} \begin{bmatrix} \hat{p}_j^R(t) \\ \hat{s}_j^R(t) \end{bmatrix} + \begin{bmatrix} G_j^p \\ G_j^s \end{bmatrix} p_j^R(t), \quad (2.10)$$

where G_j^p, G_j^s are constant gains. The linear feedback control is then given by

$$a_j = \begin{bmatrix} F_j^p & F_j^s \end{bmatrix} \begin{bmatrix} p_j - \hat{p}_j^R \\ s_j - \hat{s}_j^R \end{bmatrix}.$$

Combining (2.9) and (2.10) we obtain the closed loop system

$$\begin{bmatrix} p_j(t+1) \\ s_j(t+1) \\ \hat{p}_j^R(t+1) \\ \hat{s}_j^R(t+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 \\ \frac{1}{\tau}F_j^p & \frac{\tau-1}{\tau} + \frac{1}{\tau}F_j^s & \frac{-1}{\tau}F_j^p & \frac{-1}{\tau}F_j^s \\ 0 & 0 & 1 - G_j^p & 1 \\ 0 & 0 & -G_j^s & \frac{\tau-1}{\tau} \end{bmatrix}}_{A_j} \begin{bmatrix} p_j(t) \\ s_j(t) \\ \hat{p}_j^R(t) \\ \hat{s}_j^R(t) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ G_j^p \\ G_j^s \end{bmatrix}}_{B_j} p_j^R(t) + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_E w_j(t).$$

Note that $\mathbf{x}_1 := [p_1 - \Delta_1, s_1, \hat{p}_1^R - \Delta_1, \hat{s}_1^R]^T$ has essentially the same dynamics as $[p_1, s_1, \hat{p}_1^R, \hat{s}_1^R]^T$, namely

$$\mathbf{x}_1(t+1) = A_1 \mathbf{x}_1(t) + B_1 p_c^R + E w_1(t).$$

Assuming that $\mathbf{v} := [w_1, w_c, \tilde{w}_1, \tilde{w}_c]^T$ is a white noise process we can define the following LTI-SS representation of the process $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_c]^T := [p_1 - \Delta_1, p_c]^T$:

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_c(t+1) \end{bmatrix} &= \underbrace{\begin{bmatrix} A_1 & [B_1 \ 0] \\ 0 & A_c - B_c F \end{bmatrix}}_A \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_c(t) \end{bmatrix} + \underbrace{\begin{bmatrix} E \ 0 \ 0 \ B_1 \\ 0 \ E \ 0 \ -B_c \end{bmatrix}}_B \mathbf{v}(t) \\ \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_c(t) \end{bmatrix} &= \underbrace{\begin{bmatrix} E^T & 0 \\ 0 & E^T \end{bmatrix}}_C \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_c(t) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \ 1 \ 0 \ 0 \\ 0 \ 0 \ 0 \ 1 \end{bmatrix}}_D \mathbf{v}(t). \end{aligned}$$

Parameter settings: Following the approach of (Kempker, 2012), we take $F_1^p = F_c^p$ and $F_1^s = F_c^s$ as the solution of the linear quadratic problem $\min_{a_c} \{\|\mathbf{z}_1^2\|^2 + \alpha \|a_c^2\|^2\}$ with respect to the dynamics

$$\begin{bmatrix} \mathbf{z}_1(t+1) \\ \mathbf{z}_2(t+1) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & \frac{\tau-1}{\tau} \end{bmatrix} \begin{bmatrix} \mathbf{z}_1(t) \\ \mathbf{z}_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{\tau} \end{bmatrix} a_c(t).$$

Accordingly, for $\tau = 2$ and $\alpha = 10$ the optimal solution is $F_1^p = F_c^p = -0.3$ and $F_1^s = F_c^s = -0.5$. The gain constants were chosen to be $G_1^p = 1.5$, $G_1^s = 0.3$, $G_{R_c}^p = 0.9$, and $G_{R_c}^s = 0.5$ for which the matrix A is stable. Finally, the joint noise process \mathbf{v} is chosen to be a normalized Gaussian white noise process.

Reverse engineering of the network graph: Assume that the output process \mathbf{y} of the LTI-SS representation $\mathcal{S} := (A, B, C, D, \mathbf{v}, \mathbf{y})$ is observed. Using the result of this chapter, we will calculate a minimal Kalman representation of \mathbf{y} in causal block triangular form. Note that we do not use prior knowledge of the structure of the network graph. In fact, this representation reconstructs the network graph of \mathcal{S} .

First, we check the Granger non-causal relations among the components of \mathbf{y} using the covariance sequence $\{\Lambda_k^{\mathbf{y}}\}_{k=0}^{2N}$ of \mathbf{y} where N is larger than or equal to the dimension of a minimal LTI-SS representation of \mathbf{y} . For this, we calculate a Kalman representation (A, K, C, I, \mathbf{e}) of \mathbf{y} and verify that \mathbf{y} is coercive by checking that $A - KC$ is invertible (see Section 1.2.2). In view of Corollary 2.6, a Granger non-causal relation can be verified by observing the output matrix K of Algorithm 5. More specifically, if the left lower block of the matrix K is zero, then an appropriate Granger non-causal relation holds (see Remark 2.9). Following this method, we apply Algorithm 5 choosing the coordinator to be \mathbf{y}_1 and \mathbf{y}_c , respectively. We obtain that \mathbf{y}_1 does not Granger cause \mathbf{y}_c the condition of Theorem 2.5 holds for $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_c]^T$.

Second, in order to calculate a Kalman representation of \mathbf{y} in block triangular form, we apply Algorithm 5 with the covariance sequence $\{\Lambda_k^{\mathbf{y}}\}_{k=0}^{2N}$ as its input. With our parameter settings, the Kalman representation, that the output matrices of Algorithm 5 defines, is given as follows:

$$\begin{bmatrix} \mathbf{x}_1(t+1) \\ \mathbf{x}_c(t+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 0.4 & -0.3 & -0.1 & 0.1 & | & 0.0 & 0.1 & 0.2 & -0.1 \\ 0.2 & 0.4 & 0.6 & 0.2 & | & 0.1 & 0.0 & 0.2 & 0.1 \\ 0.0 & -0.3 & 0.4 & -0.1 & | & 0.0 & 0.0 & 0.0 & -0.4 \\ -0.2 & 0 & -0.2 & 0.7 & | & 0.0 & -0.1 & 0.0 & 0.0 \\ \hline 0 & 0 & 0 & 0 & | & 0.2 & -0.9 & 0.1 & 0.0 \\ 0 & 0 & 0 & 0 & | & 0.6 & 0.3 & 0.3 & 0.1 \\ 0 & 0 & 0 & 0 & | & -0.1 & 0.2 & -0.4 & 0.4 \\ 0 & 0 & 0 & 0 & | & 0.1 & 0.0 & -0.5 & 0.3 \end{bmatrix}}_{A_k} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_c(t) \end{bmatrix} + \underbrace{\begin{bmatrix} 0.1 & | & 0.2 \\ -0.1 & | & -0.1 \\ 0.0 & | & 0.0 \\ -0.3 & | & -0.1 \\ \hline 0 & | & 0.1 \\ 0 & | & 0.2 \\ 0 & | & 0.2 \\ 0 & | & -0.1 \end{bmatrix}}_K \mathbf{e}_1$$

$$\begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_c(t) \end{bmatrix} = \underbrace{\begin{bmatrix} -0.4 & -0.3 & 0.4 & -1.8 & | & 0.2 & -0.2 & 0.0 & 0.2 \\ 0 & 0 & 0 & 0 & | & 0.5 & 0.0 & -0.2 & 0.0 \end{bmatrix}}_{C_k} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_c(t) \end{bmatrix} + \mathbf{e}.$$

In view of Remark 2.7, the representation $\mathcal{S}_k := (A_k, K, C_k, I, \mathbf{e}, \mathbf{y})$ is a Kalman

representation in causal block triangular form. The calculation of \mathcal{S}_k only requires the second order statistics of the output process and does not use prior knowledge of the network topology. Therefore, the construction of \mathcal{S}_k shows that the network graph of \mathcal{S} can be reverse engineered. Moreover, by Theorem 3.5, the reconstructed representation \mathcal{S}_k not only has block triangular structure, but also characterizes the causal relations in the observed process. The procedure can be repeated based on data, using empirical covariances which provide an estimation of \mathcal{S}_k . Note that \mathcal{S}_k could also be calculated in a distributed way which possibly reduces estimation error when applied to data.

2.5 Conclusions

In this chapter we characterized Granger non-causality between two components of a stationary process with the existence of a minimal Kalman representation in a so-called causal block triangular form that generates the process at hand. This representation not only characterizes Granger non-causality but has a network graph that illustrates the causal relationship between the components of its output process. We provided algorithms for calculating this Kalman representation, in particular, calculating it from the covariance sequence of the observed output process. The covariances can be estimated from data. Hence, our results open up the possibility of calculating this representation from output data. The results deal with coercive and non-coercive processes and the minimality of the representations.

2.A Proof of Theorem 2.5 and Corollary 2.6

Recall that $U+V := \{u+v|u \in U, v \in V\}$ denotes the sum of two subspaces $U, V \subseteq \mathcal{H}$, $V \ominus U$ denotes the orthogonal complement of U in V , $U \dot{+} V$ denotes the direct sum and $U \oplus V$ denotes the orthogonal direct sum of U in V . Furthermore, an LTI-SS representation (A, B, C, D, \mathbf{v}) of a white noise process \mathbf{y} has zero dimension, thus A, B, C are absent, and it is the trivial equation $\mathbf{y} = D\mathbf{v}$. A zero dimensional LTI-SS representation is minimal, observable and controllable by convention.

Proof of Theorem 2.5. First, we discuss the trivial implications: since any minimal Kalman representation in causal block triangular form is a Kalman representation in a causal block triangular form and any Kalman representation in a causal block triangular form is a Kalman representation in block triangular form (ii) \implies (iii) and (iii) \implies (iv) follow. In addition, the implication (ii) \implies (i) is easy to see; if (A, K, C, I, \mathbf{e}) is a minimal Kalman representation of \mathbf{y} in causal block triangular form (2.1), then $(A_{22}, K_{22}, C_{22}, I_{r_2}, \mathbf{e}_2)$ is a minimal Kalman representation of \mathbf{y}_2 , and hence $\mathbf{e}_2(t) = \mathbf{y}_2(t) - E_t[\mathbf{y}_2(t) | \mathcal{H}_{t-}^{\mathbf{y}_2}]$ equals the innovation process of \mathbf{y}_2 . By (Dufour and Renault, 1998, Proposition 2.3), the latter implies that \mathbf{y}_1 does not Granger cause \mathbf{y}_2 .

(iv) \implies (i) if \mathbf{y} is coercive: Since \mathbf{y} is coercive, Granger non-causality is equivalent to the transfer matrix of a Kalman representation of \mathbf{y} having a block triangular structure, see (Caines, 1976, Theorem 2.2.). Since the transfer function of a Kalman representation in block triangular form has a triangular structure described in (Hsiao, 1982; Caines, 1976), the implication (iv) \implies (i) follows.

(i) \implies (ii): To begin with, from Proposition 1.7 we know that the ZMSIR process $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T$ has a minimal Kalman representation $(\hat{A}, \hat{K}, \hat{C}, I, \mathbf{e})$. Assuming that (i) in Theorem 2.5 holds, we transform this representation into causal block triangular form. Consider the partition $\hat{C} = [\hat{C}_1^T, \hat{C}_2^T]^T$ such that $C_i \in \mathbb{R}^{r_i \times p}$ where $r_i = \dim(\mathbf{y}_i)$ is the dimension of \mathbf{y}_i , $i = 1, 2$ and p is the dimension of $(\hat{A}, \hat{K}, \hat{C}, I, \mathbf{e}, \mathbf{y})$. We assume that $p > 0$; if $p = 0$, then $\mathbf{y} = \mathbf{e}$ defines a minimal Kalman representation in causal block triangular form. Take the non-singular matrix T which brings (\hat{A}, \hat{C}_2) into observability staircase form, i.e., T is such that

$$T\hat{A}T^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad \hat{C}_2T^{-1} = [0 \quad C_{22}], \quad (2.11)$$

where (A_{22}, C_{22}) is observable and $A_{11} \in \mathbb{R}^{p_1 \times p_1}$, $A_{22} \in \mathbb{R}^{p_2 \times p_2}$ such that p_2 is the rank of the observability matrix of the pair (\hat{A}, \hat{C}_2) . Define $A := T\hat{A}T^{-1}$, $K := T\hat{K}$, $C := \hat{C}T^{-1}$ and notice that $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is a minimal Kalman representation since it is isomorphic to $(\hat{A}, \hat{K}, \hat{C}, I, \mathbf{e}, \mathbf{y})$. Note that if (\hat{A}, \hat{C}_2) is observable, then

$p_1 = 0$ and A_{11}, A_{12} are absent in (2.11). If the observability matrix of (\hat{A}, \hat{C}_2) has zero rank, then $p_2 = 0$ and A_{12}, A_{22}, C_{22} are absent. Moreover, if $p_2 = 0$, then $(A, K, C, I, \mathbf{e}, \mathbf{y})$ is already in causal block triangular form (see Remark 2.2). Hence, we can assume that $p_2 > 0$. Next, we show that $K_{21} = 0$ where

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}, \quad K_{ij} \in \mathbb{R}^{p_i \times r_j} \text{ for } i = 1, 2.$$

Denote the state of $(A, K, C, I, \mathbf{e}, \mathbf{y})$ by \mathbf{x} . Take the partition $\mathbf{e} = [\mathbf{e}_1^T, \mathbf{e}_2^T]^T$ and $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T]^T$ where $\mathbf{e}_i \in \mathbb{R}^{r_i}$ and $\mathbf{x}_i \in \mathbb{R}^{p_i}$, $i = 1, 2$. Notice that

$$C = \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{bmatrix}, \quad A^k = \begin{bmatrix} A_{11}^k & (A^k)_{12} \\ 0 & A_{22}^k \end{bmatrix},$$

where $C_{1i} \in \mathbb{R}^{r_1 \times p_i}$, $i = 1, 2$ and $(A^k)_{12} \in \mathbb{R}^{p_1 \times p_2}$ denotes the right upper block of A^k . It then follows that $C_2 A^k \mathbf{x}(t) = C_{22} A_{22}^k \mathbf{x}_2(t)$ and for $k > 0$

$$\mathbf{y}_2(t+k) = C_{22} A_{22}^k \mathbf{x}_2(t) + \sum_{l=0}^{k-1} M_l \mathbf{e}(t+k-l) \quad (2.12)$$

for some matrices M_0, \dots, M_{k-1} . Since \mathbf{e} is the innovation process of \mathbf{y} , it implies that $\mathbf{e}(t+k-l)$ is orthogonal to $\mathcal{H}_{t-}^{\mathbf{y}}$ and $\mathcal{H}_{t-}^{\mathbf{e}} = \mathcal{H}_{t-}^{\mathbf{y}}$, for $k-l \geq 0$, $t \in \mathbb{Z}$. Furthermore, from $\mathbf{x}(t) = \sum_{k=1}^{\infty} C A^{k-1} K \mathbf{e}(t-k)$, the components of $\mathbf{x}(t)$ belong to $\mathcal{H}_{t-}^{\mathbf{e}} = \mathcal{H}_{t-}^{\mathbf{y}}$. Using (2.12), it then follows that $E_l[\mathbf{y}_2(t+k)|\mathcal{H}_{t-}^{\mathbf{y}}] = C_{22} A_{22}^k \mathbf{x}_2(t)$. As \mathbf{y}_1 does not Granger cause \mathbf{y}_2 we know that $E_l[\mathbf{y}_2(t+k)|\mathcal{H}_{t-}^{\mathbf{y}_2}] = E_l[\mathbf{y}_2(t+k)|\mathcal{H}_{t-}^{\mathbf{y}}]$ for all $k \geq 0$, and thus $E_l[\mathbf{y}_2(t+k)|\mathcal{H}_{t-}^{\mathbf{y}_2}] = C_{22} A_{22}^k \mathbf{x}_2(t) \in \mathcal{H}_{t-}^{\mathbf{y}_2}$. Let O_2 be the observability matrix of (A_{22}, C_{22}) and denote its left inverse by O_2^+ . Then by defining $Y_2(t) = [\mathbf{y}_2^T(t), \dots, \mathbf{y}_2^T(t+n-1)]^T$, we have that $\mathbf{x}_2(t) = O_2^+ E_l[Y_2(t)|\mathcal{H}_{t-}^{\mathbf{y}_2}]$ and thus the elements of $\mathbf{x}_2(t)$ belong to $\mathcal{H}_{t-}^{\mathbf{y}_2}$. Note that since \mathbf{y}_1 does not Granger cause \mathbf{y}_2 , by (Dufour and Renault, 1998, Proposition 2.3) \mathbf{e}_2 is the innovation process of \mathbf{y}_2 , and hence $\mathcal{H}_{(t+1)-}^{\mathbf{y}_2} = \mathcal{H}_{t-}^{\mathbf{y}_2} \oplus \mathcal{H}_t^{\mathbf{e}_2}$. Therefore,

$$\mathbf{x}_2(t+1) = E_l[\mathbf{x}_2(t+1)|\mathcal{H}_{(t+1)-}^{\mathbf{y}_2}] = E_l[\mathbf{x}_2(t+1)|\mathcal{H}_{t-}^{\mathbf{y}_2}] + E_l[\mathbf{x}_2(t+1)|\mathcal{H}_t^{\mathbf{e}_2}].$$

From that $\mathbf{e}(t)$ is orthogonal to $\mathcal{H}_{t-}^{\mathbf{y}} \supseteq E_l[\mathbf{x}_2(t+1)|\mathcal{H}_{t-}^{\mathbf{y}_2}]$ and $E_l[\mathbf{x}_2(t+1)|\mathcal{H}_t^{\mathbf{e}_2}] \subseteq \mathcal{H}_t^{\mathbf{e}}$ we have that

$$E_l[\mathbf{x}_2(t+1)|\mathcal{H}_t^{\mathbf{e}}] = E_l[\mathbf{x}_2(t+1)|\mathcal{H}_t^{\mathbf{e}_2}] = \hat{R} \mathbf{e}_2(t)$$

for a suitable \hat{R} matrix. Then $\mathbf{x}_2(t+1) = A_{22} \mathbf{x}_2(t) + [K_{21} \ K_{22}] \mathbf{e}(t)$ and that $\mathbf{e}(t)$ is

orthogonal to $\mathcal{H}_t^{\mathbf{x}_2}$ implies that

$$E_l[\mathbf{x}_2(t+1)|\mathcal{H}_t^{\mathbf{e}}] = [K_{21} \ K_{22}] \mathbf{e}(t) = \hat{R}\mathbf{e}_2(t).$$

Using that \mathbf{y} is full rank, \mathbf{e}_1 and \mathbf{e}_2 are linearly independent, and hence $K_{21} = 0$, $K_{22} = \hat{R}$. That is, (A, K, C, I, \mathbf{e}) is a Kalman representation of \mathbf{y} in block triangular form.

In order to see that (A, K, C, I, \mathbf{e}) is in causal block triangular form, we need to show that $(A_{22}, K_{22}, C_{22}, I_{r_2}, \mathbf{e}_2)$ is a minimal Kalman representation of \mathbf{y}_2 . From Granger non-causality, \mathbf{e}_2 is the innovation process of \mathbf{y}_2 , hence we only need to prove minimality. Note that if $p_1 = 0$, then $A = A_{22}, K = K_{22}, C = C_{22}$ thus $(A_{22}, K_{22}, C_{22}, I_{r_2}, \mathbf{e}_2)$ is minimal. In view of Proposition 1.10, it is sufficient to show that (A_{22}, C_{22}) is observable and (A_{22}, K_{22}) is controllable. The former follows from the construction. Assume now indirectly that (A_{22}, K_{22}) is uncontrollable, i.e., that for some vector $\eta \neq 0$, $\eta^T A_{22}^k K_{22} = 0$ for all $k \geq 0$. However,

$$A^k K = \begin{bmatrix} A_{11}^k K_{11} & (A^k K)_{12} \\ 0 & A_{22}^k K_{22} \end{bmatrix},$$

where $(A^k K)_{12}$ denotes the right upper block of $A^k K$ with suitable dimensions. It follows that $[0 \ \eta^T] A^k K = 0$ for all $k \geq 0$, which implies that (A, K) is not controllable. Since (A, K, C, I, \mathbf{e}) is a minimal Kalman representation of \mathbf{y} , by Proposition 1.10 (A, K) is controllable, which is a contradiction. This implies that (A, K, C, I, \mathbf{e}) is a minimal Kalman representation in causal block triangular form which completes the proof. \square

Proof of Corollary 2.6. The construction of the Kalman representation (A, K, C, I, \mathbf{e}) of \mathbf{y} coincides with the one described in the proof of the implication (i) \implies (ii) in Theorem 2.5. Hence, if \mathbf{y}_1 does not Granger cause \mathbf{y}_2 , then the above-mentioned proof implies that either K_{21} is absent or $K_{21} = 0$ and that (A, K, C, I, \mathbf{e}) is a minimal Kalman representation of \mathbf{y} in causal block triangular form. Conversely, if K_{21} is absent or $K_{21} = 0$, and \mathbf{y} is coercive, then (A, K, C, I, \mathbf{e}) is a minimal Kalman representation of \mathbf{y} in block triangular form. Hence, by the implication (iii) \implies (i) of Theorem 2.5, \mathbf{y}_1 does not Granger cause \mathbf{y}_2 . \square

