



CrossMark  
click for updates

## Review

**Cite this article:** Eichler M. 2013 Causal inference with multiple time series: principles and problems. *Phil Trans R Soc A* 371: 20110613. <http://dx.doi.org/10.1098/rsta.2011.0613>

One contribution of 13 to a Theme Issue 'Assessing causality in brain dynamics and cardiovascular control'.

### Subject Areas:

statistics

### Keywords:

causal identification, Granger causality, causal effect, spurious causality, latent variables, impulse response function

### Author for correspondence:

Michael Eichler

e-mail: [m.eichler@maastrichtuniversity.nl](mailto:m.eichler@maastrichtuniversity.nl)

# Causal inference with multiple time series: principles and problems

Michael Eichler

Department of Quantitative Economics, Maastricht University,  
PO Box 616, 6200 MD Maastricht, The Netherlands

I review the use of the concept of Granger causality for causal inference from time-series data. First, I give a theoretical justification by relating the concept to other theoretical causality measures. Second, I outline possible problems with spurious causality and approaches to tackle these problems. Finally, I sketch an identification algorithm that learns causal time-series structures in the presence of latent variables. The description of the algorithm is non-technical and thus accessible to applied scientists who are interested in adopting the method.

## 1. Introduction

The learning of cause–effect relationships is an important task for our understanding of the functionality of the brain and the cardiovascular system. In recent years, an increasing number of studies have used time-series methods based on the notion of Granger causality [1–7]. This probabilistic concept of causality exploits the fact that causes must precede their effects in time. However, temporal precedence alone is not a sufficient condition for establishing cause–effect relationships, and it is well known that the omission of relevant variables can lead to the so-called spurious causalities, that is, falsely detected causal relationships. Consequently, one often meets reservations that Granger causality is not suited as a tool for causal inference [8]. For a controversy about the application of Granger causality in brain imaging, I refer to Roebroeck *et al.* [6] and the related papers in the same issue of *NeuroImage*.

While it is true that association—even though between lagged variables—does not necessarily constitute a causal link, the concept of Granger causality still remains a useful tool for causal learning. Given the nowadays widespread use of Granger causality, it is

therefore important to raise awareness of the limitations of Granger causality and to point out strategies for valid causal inference from time-series data. The objective of the present paper is to review the concepts and methods for causal inference and to outline the underlying assumptions on which the causal conclusion rests.

The paper is organized as follows. In §2, I review four theoretical concepts of causality for multivariate time series and point out the relations among them. The discussion shows under which circumstances Granger causality as well as Sims causality—another probabilistic concept of causality from econometrics that will be of interest for our discussion of causal inference—indeed correspond to a true causal link between the variables. In §3, I discuss the empirical versions of Granger and Sims causality further. In particular, I move that Sims causality corresponds to the total information flow from one variable to another; in §4, I illustrate its use for detecting certain types of spurious causalities. In §5, I briefly sketch the basic steps for identifying causal relationships from observational time-series data. Section 6 concludes the paper with a discussion of open problems and challenges for causal inference in the time-series case.

## 2. Concepts of causality for multiple time series

Among the many possible properties that have been hypothesized for cause–effect relationships, there are two properties that have been particularly fruitful in the context of time series.

- *Temporal precedence*: a cause precedes its effects in time.
- *Physical influence*: manipulations of the cause change the effects.

The second property expresses the common sense notion that causal relationships link variables fundamentally and continue to work if other parts of the system are changed. Thus, an enforced change of value in the causing variable will lead to corresponding changes in the caused variable. This aspect is central to the modern theory of causality [9–13].

In time-series analysis, most definitions of causality focus on the first property of temporal precedence. The concepts have the advantage that they provide readily empirical versions that can be used for inference. The most prominent concept is that of Granger causality [14–16], which recently has attracted much attention in the field of neuroscience [1,3,6,17]. However, similar to correlation in ordinary multivariate analysis, the concept of Granger causality is basically probabilistic and can lead to wrongly detected causal relations in the presence of latent variables. Nevertheless, it is an indispensable tool for causal inference if it is properly applied.

In the following, I review four approaches for defining causality for multiple time series and discuss the relations among them. Two of these concepts, Granger causality and Sims causality, have empirical versions, which will be the basis for the discussion of causal inference in the remaining sections of this paper.

### (a) Intervention causality

The most successful modern approach for describing causality is based on the concept of interventions [9–13]. The underlying principle of this approach is that a cause–effect relationship should persist if the cause is manipulated without directly affecting any other variables, whereas any non-causal associations should vanish. In the context of time series, this idea of defining a causal effect as the effect of an intervention in such a system was first proposed by Eichler & Didelez [18] and discussed in more detail in Eichler & Didelez [19] and Eichler [20].

More precisely, to describe the causal effect that one process  $X$  exerts on another process  $Y$ , I consider primitive interventions that set the value of  $X$  at some time  $t$  to a fixed value  $x^*$ , say. This yields a new probability measure, which I denote—in terms of Pearl’s do-terminology [10]—by  $\mathbb{P}(\cdot \mid \text{do}(X_t = x^*))$ . The causal effect of  $X$  on  $Y$  can now be quantified by any functional of

the post-intervention distribution  $\mathbb{P}(Y_{t'} | \text{do}(X_t = x^*))$  of  $Y_{t'}$  with  $t' > t$ . The most commonly used measure is the average causal effect (ACE) defined as the average increase or decrease in value caused by the intervention.

**Definition 2.1 (Average causal effect).** The ACE or causal effect in mean of an intervention setting  $X_t$  to some value  $x^*$  on the response variable  $Y_{t'}$  with  $t' > t$  is given by

$$\text{ACE}_{x^*} = \mathbb{E}(Y_{t'} | \text{do}(X_t = x^*)) - \mathbb{E}(Y_{t'}).$$

Here, the mean  $\mathbb{E}(\cdot)$  is taken with respect to the natural probability measure  $\mathbb{P}$  that describes the system without interventions. Depending on the specific application, other measures to describe the causal effect could be more appropriate. For instance, the causal effect in variance given by the difference  $\text{var}(Y_{t'} | \text{do}(X_t = x^*)) - \text{var}(Y_{t'})$  could be of interest in the analysis of financial time series, where an intervention might aim at reducing the volatility of the stock market.

Next, if the intervention in  $X_t$  shows no effect whatsoever on later instances of the process  $Y$ , we conclude that the process  $X$  does not cause the process  $Y$ .

**Definition 2.2 (Non-causality with respect to interventions).** The process  $X$  does not cause the process  $Y$  if an intervention in  $X$  setting  $X_t$  to a value  $x^*$  does not affect the distribution of future instances of  $Y$ , that is, if

$$\mathbb{P}(Y_{t'} \leq y | \text{do}(X_t = x^*)) = \mathbb{P}(Y_{t'} \leq y),$$

for all  $y \in \mathbb{R}$  and all  $t' > t$ .

I note that the process  $Y$  may be multi-dimensional to cover cases where an intervention in  $X$  does affect the relationship between two components of  $Y$ . Eichler [21] formulates a threshold autoregressive model as an example of such a process.

One common problem in time-series analysis is that controlled experiments cannot be carried out. Therefore, the effect of an intervention must be determined from data collected under the natural distribution  $\mathbb{P}$ . This is only possible if the intervention affects only the distribution of the target variable  $X_t$ , whereas all other conditional distributions remain the same as under the observational regime, i.e. without intervention in  $X_t$ . This invariance of distributions under different regimes is formalized by the following conditions: there exists a process  $V = (X, Y, Z)$  such that

- (I1) the distributions of  $X^{t-1}, Y^t, Z^t$  under the observational and the intervention regimes are the same, and
- (I2) conditional on  $V^t$  the joint distributions of any future instances  $V_{t+1}, \dots, V_{t'}$  for  $t' > t$  under the observational and the intervention regimes are the same.

Here,  $X^t = \{X_s | s \leq t\}$  denotes the history of the process  $X$  up to time  $t$ . Under these conditions, the distribution of the time series  $V = (X, Y, Z)$  under the intervention regime is fully specified by its natural distribution described by the measure  $\mathbb{P}$  and the intervention requiring that  $X_t = x^*$  with probability 1. Moreover, the conditions imply that any dependences between  $X_t$  and  $Y_{t'}$  for  $t' > t$  conditional on  $X^{t-1}, Y^t, Z^t$  found under the observational regime are causal. Thus, the conditions impose some causal Markov property on the system.

## (b) Structural causality

In a recent article, White & Lu [22] proposed a new concept of the so-called direct structural causality for the discussion of causality in dynamic structural systems. The approach is based on the assumption that the data-generating process has a recursive dynamic structure in which predecessors structurally determine successors. For ease of notation, the following definitions are slightly modified and simplified.

Suppose that we are interested in the causal relationship between two processes, the ‘cause of interest’  $X$  and the ‘response of interest’  $Y$ . We assume that  $X$  and  $Y$  are structurally generated as

$$X_t = q_{x,t}(X^{t-1}, Y^{t-1}, Z^{t-1}, U_{x,t})$$

and

$$Y_t = q_{y,t}(X^{t-1}, Y^{t-1}, Z^{t-1}, U_{y,t}),$$

for all  $t \in \mathbb{Z}$ . Here, the process  $Z$  includes all relevant observed variables, while the realizations of  $U_t = (U_{x,t}, U_{y,t})$  are assumed to be unobserved. The functions  $q_{x,t}$  and  $q_{y,t}$  are also assumed to be unknown.

**Definition 2.3 (Direct structural causation).** The process  $X$  does not directly structurally cause the process  $Y$  if the function  $q_{y,t}(x^{t-1}, y^{t-1}, z^{t-1}, u_{y,t})$  is constant in  $x^{t-1}$  for all admissible values for  $y^{t-1}$ ,  $z^{t-1}$  and  $u_{y,t}$ . Otherwise,  $X$  is said to directly structurally cause  $Y$ .

The idea of defining causality by assuming a set of structural equations is not new but has been used by a number of authors before [9,10,23,24]. However, in contrast to this strand of the literature, White & Lu [22] make no reference to interventions or to graphs.

It is clear from the definition that a primitive intervention on  $X_t$  results in replacing the generating equation by  $X_t = x^*$ . As the generating equation for the response variable  $Y_{t'}$ ,  $t' > t$ , is unaffected by this, we immediately obtain that direct structural non-causality implies non-causality with respect to interventions, that is,

$$\mathbb{P}(Y_{t+1} \leq y \mid \text{do}(X_t = x^*)) = \mathbb{P}(Y_{t+1} \leq y),$$

for all  $t' > t$  and all  $y$  whenever  $X$  does not directly structurally cause  $Y$ . White & Lu [22] also propose a definition of total structural causality. With this, the above result can be generalized to responses at arbitrary times  $t' > t$ . I omit the details and refer the reader to White & Lu [22].

Furthermore, I note that the converse implication is generally not true. As an example, suppose that  $Y_t$  is generated by  $Y_t = \text{sign}(X_{t-1}) U_{y,t}$ , where  $U_{y,t}$  has a symmetric distribution independent of  $X^{t-1}$  and  $Y^{t-1}$ . Then,  $X$  directly structurally causes  $Y$  but the distribution of  $Y_t$  does not depend on  $X^{t-1}$ . Such causal relationships are only relevant for counterfactual queries such as: ‘Suppose  $x_{t-1}$  has a positive value. What would have been the value of  $Y_t$  if  $x_{t-1}$  had been negative?’ The corresponding counterfactual realization  $y_{x^*,t}$  is given by  $y_{x^*,t} = \text{sign}(x^*) u_{y,t} = -y_t$ , that is,  $Y_t$  would have had the opposite sign. However, since  $Y_{x^*,t}$  is not observable, this causal dependence cannot be detected empirically.

### (c) Granger causality

In time-series analysis, inference about cause–effect relationships is commonly based on the concept of Granger causality [14,15]. Unlike the previous approach, this probabilistic concept of causality does not rely on the specification of a scientific model and thus is particularly suited for empirical investigations of cause–effect relationships. For this general definition of causality, Granger [14,15] evokes the following two fundamental principles:

- (i) the effect does not precede its cause in time;
- (ii) the causal series contains unique information about the series being caused that is not available otherwise.

The first principle of temporal precedence of causes is commonly accepted and has also been the basis for other probabilistic theories of causation [25–27]. By contrast, the second principle is more subtle, as it requires the separation of the special information provided by the former series  $X$  from any other possible information. To this end, Granger considers two information sets: first, the set of all information in the universe up to time  $t$ , denoted by  $\mathcal{I}_t^*$ ; and, second, the same information set reduced by the information of  $X$  up to time  $t$ , denoted by  $\mathcal{I}_{-X,t}^*$ . If the series

$X$  causes series  $Y$ , we expect by the above principles that the conditional distributions of  $Y_{t+1}$  given the two information sets  $\mathcal{I}_t^*$  and  $\mathcal{I}_{-X,t}^*$  differ from each other. Equivalently, letting  $X \perp\!\!\!\perp Y | Z$  denote that  $X$  and  $Y$  are conditionally independent given  $Z$ , we have the following:

**Definition 2.4 (Granger's definition of causality [14,15]).** The process  $X$  does not cause the process  $Y$  if  $Y_{t+1} \perp\!\!\!\perp X^t | \mathcal{I}_{-X,t}^*$  for all  $t \in \mathbb{Z}$ ; otherwise, the series  $X$  is said to cause the series  $Y$ .

The definition poses some theoretical problems such as the exact definition of the information set  $\mathcal{I}_{-X,t}^*$ . Leaving such problems aside, it is still clear that Granger's definition of causality cannot be used with actual data. In practice, only the background knowledge available at time  $t$  can be incorporated into an analysis. Therefore, the definition must be modified to become operational. Suppose that the process  $V = (X, Y, Z)$  has been observed. Substituting the new information sets consisting of  $X^t, Y^t, Z^t$  and  $Y^t, Z^t$  for  $\mathcal{I}_t^*$  and  $\mathcal{I}_{-X,t}^*$ , respectively, we obtain the following modified version of the above definition [15,16].

**Definition 2.5 (Granger causality).** The series  $X$  is Granger non-causal for the series  $Y$  with respect to  $V = (X, Y, Z)$  if  $Y_{t+1} \perp\!\!\!\perp X^t | Y^t, Z^t$  for all  $t \in \mathbb{Z}$ ; otherwise we say that  $X$  Granger-causes  $Y$  with respect to  $V$ .

Granger [15,16] used the term 'X is a prima facie cause of Y' to emphasize the fact that a cause in the sense of Granger causality must be considered only as a potential cause. It therefore can be (and indeed has been) debated whether 'Granger causality' is an appropriate term for this kind of association. A possible alternative is the equivalent notion of 'local independence' originally introduced in the context of time-continuous processes [28]. However, the notion of Granger causality is nowadays well established, and I will therefore continue to work with this term. Moreover, this restriction in causal interpretation does not mean that the concept is completely useless for causal inference. Indeed, as I discuss in §5, it can be an essential tool for recovering (at least partially) the causal structure of a time series  $X$  if used in the right way.

It is clear from the general definition given earlier that Granger intended the information to be chosen as large as possible such that, if available, all relevant variables are included. For example, if  $Y_{t+1} \perp\!\!\!\perp \mathcal{I}_t^* | V^t$  for some process  $V = (X, Y, Z)$ , Granger causality with respect to  $V$  coincides with the original general definition with respect to  $\mathcal{I}^*$ . As in that case  $V$  comprises all (distributional relevant) direct structural causes of  $Y$ , we have also that  $X$  is Granger non-causal for  $Y$  given  $V$  if  $X$  does not directly structurally cause  $Y$ . Again, since direct structural causes need not be distributional relevant, the converse implication does not hold in general.

Similar to direct structural causality, the concept of Granger causality aims to describe direct causal-effect relationships. Consequently, it can describe the effects of interventions in only a very limited way. More precisely, suppose that the process  $V = (X, Y, Z)$  is such that conditions (I1) and (I2) are satisfied for an intervention in  $X_t$ . Then, if  $X$  is Granger non-causal for  $Y$  with respect to  $V$ , there is no causal effect of intervening in  $X_t$  on  $Y_{t+1}$ , that is,

$$\mathbb{P}(Y_{t+1} \leq y | \text{do}(X_t = x^*)) = \mathbb{P}(Y_{t+1} \leq y).$$

I note that the results cannot be strengthened by considering interventions in all instances of  $X^t$ . For more information about such sequential interventions and relating their effects to quantities of the natural distribution, I refer to Eichler & Didelez [19].

#### (d) Sims causality

The econometric literature features other less well-known probabilistic notions of causality that are related to Granger causality. Among these, the concept introduced by Sims [29] seems of most interest. In contrast to Granger causality, it takes in account not only direct but also indirect causal effects.

**Definition 2.6 (Sims non-causality).** The process  $X$  is Sims non-causal for the process  $Y$  with respect to  $V = (X, Y, Z)$  if  $\{Y_{t'} | t' > t\} \perp\!\!\!\perp X_t | X^{t-1}, Y^t, Z^t$  for all  $t \in \mathbb{Z}$ ; otherwise, we say that  $X$  is Sims causal for  $Y$  with respect to  $V$ .

If the variables in  $Z$  include all confounders for the dependence of  $Y$  on  $X$ , Sims non-causality describes the absence of a causal effect of  $X$  on  $Y$  at any lag. More precisely, consider a process  $V = (X, Y, Z)$  such that conditions (I1) and (I2) are satisfied for an intervention in  $X_t$ . Then  $X$  is Sims non-causal for  $Y$  with respect to  $V$  if and only if  $\mathbb{P}(Y_{t'} \leq y | \text{do}(X_t = x^*)) = \mathbb{P}(Y_{t'} \leq y)$ , for all  $t' > t$  and all  $y$ , that is,  $X$  is non-causal for  $Y$  (with respect to an intervention in  $X_t$ ).

### 3. Empirical concepts: Granger and Sims causality

Of the four concepts of causality presented in §2, the first two concepts—intervention and structural causality—are more of a theoretical nature in the sense that they provide a formal setting to describe causality and allow conclusions about how the system behaves, for instance, under certain interventions. Working with these concepts typically requires some knowledge about the causal structure; a typical application is the derivation of conditions under which the effect of an intervention can be predicted from data taken under the observational regime [19]. They are less suitable for the identification of causal relationships.

In contrast, the concepts of Granger and Sims causality are formulated in terms of statistical independence and therefore provide readily operational versions for statistical inference. Although both Granger and Sims causality apply generally to multiple time series of any type [21], empirical applications usually are confined to the framework of vector autoregressive (VAR) models. In the following, I will briefly review the empirical versions of Granger and Sims causality in this setting. For this purpose, I consider throughout this paper stationary Gaussian processes  $X = (X_t)$  that have an infinite autoregressive representation of the form

$$X_t = \sum_{u=1}^{\infty} A_u X_{t-u} + \varepsilon_t, \quad (3.1)$$

where the innovation process  $\varepsilon = (\varepsilon_t)$  is white noise with mean zero and non-singular covariance matrix  $\Sigma$ . Furthermore, the  $d \times d$  matrices  $A_u$  ( $d$  being the dimension of the process  $X$ ) are assumed to be square summable, which ensures the existence of a moving average representation for  $X$ . In the following, recall that  $X^t = \{X_s | s \leq t\}$  denotes the past of  $X$  up to time  $t$ .

#### (a) Granger causality

As noted in §2, the series  $X_i$  does not Granger-cause another series  $X_j$  with respect to  $X$  if  $X_{j,t}$  does not depend on  $X_i^{t-1}$  conditional on the past of the remaining variables. In the above autoregressive representation,  $X_{j,t}$  is given by

$$X_{j,t} = \sum_{u=1}^{\infty} \sum_{k=1}^d A_{jk,u} X_{k,t-u} + \varepsilon_{j,t}.$$

Hence,  $X_{j,t}$  does not depend on  $X_i^{t-1}$  if and only if  $A_{ji,u} = 0$  for all lags  $u > 0$ . An alternative characterization of Granger causality is obtained by comparing the mean square prediction errors of predicting  $X_{j,t}$  from the full past  $X^{t-1}$  and past  $X_{-i}^{t-1}$  of all components except  $X_i$ :  $X_i$  does not Granger-cause  $X_j$  with respect to the full process  $X$  if and only if  $\text{var}(X_{j,t} | X^{t-1}) = \text{var}(X_{j,t} | X_{-i}^{t-1})$ , where the conditional variance is taken to be the variance about the linear projection on the set of conditioning variables [30].

In order to learn about the causal structure of a system, we need to determine which variables Granger-cause which other variables. This can be most easily accomplished by testing whether the corresponding autoregressive coefficients are zero. More precisely, let  $\hat{A}_u$  be the least-squares estimators obtained by fitting a VAR( $p$ ) model to the data. It is well known [31] that the coefficients  $\hat{A}_{ji,u}$ ,  $u = 1, \dots, p$ , are asymptotically jointly normally distributed with means  $A_{ji,u}$  and covariance matrix  $V = (V_{uv})$  with entries  $V_{uv} = \sigma_{jj} H_{ii,uv}$  for  $u, v = 1, \dots, p$ . Here,  $H_{ij,uv}$  is the entry at row  $d(u-1) + i$  and column  $d(v-1) + j$  of  $H_p = R_p^{-1}$ , the inverse of the covariance matrix of  $X_{t-1}, \dots, X_{t-p}$ .



Replacing  $\Sigma$  and  $R_p$  by estimates, the existence of a Granger causal effect of  $X_i$  on  $X_j$  can be tested by evaluation of the test statistic

$$S_{ji} = T \sum_{u,v=1}^p \hat{A}_{ji,u} (\hat{V}^{-1})_{uv} \hat{A}_{ji,v}.$$

Under the null hypothesis that  $X_i$  is Granger non-causal for  $X_j$  with respect to  $X$ , the test statistic  $S_{ji}$  is asymptotically  $\chi^2$ -distributed with  $p$  degrees of freedom. In the case of non-Gaussian innovations, data-driven procedures for hypothesis testing based on surrogate data have been suggested as an alternative [32].

In practice, one is not only interested in testing whether or not a variable  $X_i$  Granger-causes another variable  $X_j$  but also wants to quantify any deviations from the null hypothesis. As the autoregressive coefficients  $A_{ji,u}$  depend—as any regression coefficients—on the unit of measurement of  $X_i$  and  $X_j$ , they are less suited for measuring the strength of a causal relationship. An alternative are the partial directed correlations  $\pi_{ji,u}$  proposed by Eichler [3], which are defined as the correlation between  $X_{j,t}$  and  $X_{i,t-u}$  after removing the linear effects of  $X^{t-1} \setminus \{X_{i,t-u}\}$ , and thus possess the using scaling properties of correlations. They can be estimated by rescaling of the autoregressive estimates  $\hat{A}_{ji,u}$ ; for details, I refer to Eichler [3].

In many applications in neuroscience, the time series of interest are characterized in terms of their frequency properties. The frequency domain description of a process  $X$  and the relationships among its components are based on the spectral representation of  $X$ , which is given by

$$X_t = \int_{-\pi}^{\pi} e^{i\lambda t} dZ_X(\lambda),$$

where  $Z_X(\lambda)$  is a  $\mathbb{C}^d$ -valued random process on  $[-\pi, \pi]$  with mean zero and orthogonal increments [33]. In this representation, the complex-valued random increments  $dZ_{X_i}(\lambda)$  indicate the frequency components of the time series  $X_i$  at frequency  $\lambda$ . Similarly, we obtain  $dZ_\varepsilon(\lambda)$  for the innovation process  $\varepsilon$ . The autoregressive representation of  $X$  implies that the frequency components of the processes  $X$  and  $\varepsilon$  are related by the linear equation system

$$dZ_X(\lambda) = A(\lambda) dZ_X(\lambda) + dZ_\varepsilon(\lambda), \quad (3.2)$$

where  $A(\lambda) = \sum_{u \in \mathbb{N}} A_u e^{-i\lambda u}$  is the Fourier transform of the autoregressive coefficients  $A_u$ . The coefficient  $A_{ji}(\lambda)$  vanishes uniformly for all frequencies  $\lambda \in [-\pi, \pi]$  if and only if  $X_i$  is Granger non-causal for  $X_j$  with respect to the full process  $X$ . This means that the linear equation system (3.2) reflects the causal pathways by which the frequency components influence each other. More precisely, the complex-valued coefficient  $A_{ji}(\lambda)$  indicates how a change in the frequency component of the series  $X_i$  affects the frequency component of  $X_j$  if all other components are held fixed, that is,  $A_{ji}(\lambda)$  measures the direct causal effect of  $X_i$  on  $X_j$  at frequency  $\lambda$ .

There exist a number of frequency domain-based measures for Granger causality [17,34,35]. While they necessarily—in order to be correctly classified as Granger causality measures—vanish whenever the corresponding coefficients in the autoregressive representation are zero, they differ in the way deviations from zero are quantified. I also note that these measures usually are computed from the autoregressive estimates and therefore do not offer any advantage for testing for Granger causality.

Finally, I note that the autoregressive structure of  $X$  and hence the Granger causal relations among the components of  $X$  can be conveniently visualized by a path diagram [3,36]. In this graph—sometimes referred to as a *Granger causality graph* [20]—each node  $i$  represents one component  $X_i = (X_{i,t})$  while directed edges  $i \rightarrow j$  indicate that  $X_i$  Granger-causes  $X_j$  with respect to  $X$ . Additionally, dashed undirected edges are used to encode dependences among instances of the variables at the same point in time. More precisely, the edge  $i \dashrightarrow j$  entails that  $X_i$  and  $X_j$  are contemporaneously correlated, that is, after removing the linear effects of the past  $X^{t-1}$ . For a VAR process, this is equivalent to  $\sigma_{ij} = 0$ .

## (b) Sims causality

Unlike the concept of Granger causality, which has seen an increased interest in applications from neuroscience and other fields, the approach by Sims seems to be widely unknown to researchers outside economics. In the following, I will relate this concept to some measures known in time-series analysis, while, in the next section, I will show its usefulness for causal inference.

Recall that  $X_i$  does not Sims-cause  $X_j$  if  $X_{j,t+h}$  does not depend on  $X_{i,t}$  conditional on all variables in  $X^t$  except  $X_{i,t}$  for all  $h > 0$ . Thus, Sims causality can be expressed in terms of  $h$ -step predictions of  $X_j$ ,

$$\hat{X}_{j,t+h}(X^t) = \sum_{u=0}^{\infty} \sum_{k=1}^d A_{jk,u}^{(h)} X_{k,t-u}. \quad (3.3)$$

Thus,  $X_i$  does not Sims-cause  $X_j$  if and only if  $A_{ji,0}^{(h)} = 0$  for all  $h > 0$ . The coefficients  $A_k^{(h)}$  of the  $h$ -step predictor can be computed recursively from the autoregressive coefficients  $A_u$  by [37]

$$A_k^{(h)} = A_1 A_k^{(h-1)} + \dots + A_{h-1} A_k^{(1)} + A_{h+k}.$$

Expressing the coefficients  $A_0^{(h)}$  only in terms of  $A_u$ , we obtain

$$A_0^{(h)} = \sum_{u_1 + \dots + u_n = h} A_{u_1} \dots A_{u_n}.$$

This shows that  $A_{ji,0}^{(h)}$  accumulates the information flow from the direct link and all indirect links from  $X_i$  to  $X_j$ . This underlines our earlier remark that Sims causality can be viewed as a concept of total causality.

A different view on Sims causality can be obtained by noting that the  $h$ -step predictor can be derived by recursively substituting (3.1) for  $X_{t+h-1}, \dots, X_{t+1}$  in the autoregressive representation of  $X_{t+1}$ , which yields

$$X_{t+h} = \sum_{j=1}^{h-1} B_j \varepsilon_{t+h-j} + \sum_{j=0}^{\infty} A_j^{(h)} X_{t-j} + \varepsilon_{t+h},$$

where  $B_u$  are the coefficients in the moving average representation of  $X$ . Substituting once more  $X_t$  by (3.1), we find that  $B_h = A_0^{(h)}$ , that is, the coefficients  $A_0^{(h)}$  determining Sims causality are those of the moving average representation of  $X$ . Thus, we find that—less formally expressed— $X_i$  does not Sims-cause  $X_j$  if an independent shock in  $X_i$  does not affect  $X_j$  at later points in time and hence that the *impulse-response function*  $\text{IRF}_{ji}(u) = B_{ji,u}$  is zero at all lags  $u$ . More generally, the ACE of setting  $X_{i,t}$  to  $x^*$  on  $X_{j,t+h}$  can be expressed in terms of the IRF by

$$\text{ACE}_{x^*} = \mathbb{E}(X_{j,t+h} \mid \text{do}(X_{i,t} = x^*)) = B_{ji,h} x^* = \text{IRF}_{ji}(h) x^*.$$

I note that this causal interpretation of Sims causality and the IRF is based on assumptions (I1) and (I2) in §2. These hold if the process  $X$  contains all relevant variables such that all lagged associations are truly causal and, second, there are no instantaneous causal relationships among the variables at the same point in time. If this latter assumption is violated, the additional instantaneous links also need to be taken into account. In econometrics, this is achieved by considering structural VAR processes and the corresponding orthogonalized IRF. In the discussion, I will briefly touch upon the problems that such instantaneous causalities pose for causal inference.

In the frequency domain, the concept of Sims causality is related to the well-known directed transfer function (DTF) by Kamiński & Blinowska [38]. For this, let  $B(\lambda) = \sum_{u=0}^{\infty} B_u e^{i\lambda u}$  be the



frequency domain IRF. Then, the DTF measuring the influence of variable  $X_i$  on variable  $X_j$  at frequency  $\lambda$  is defined as

$$\gamma_{ji}^2(\lambda) = \frac{|B_{ji}(\lambda)|^2}{\sum_k |B_{jk}(\lambda)|^2}.$$

Thus, the DTF  $\gamma_{ji}^2(\lambda)$  vanishes at all frequencies  $\lambda$  if and only if  $X_i$  does not Sims-cause  $X_j$  with respect to the process  $X$ . This clarifies the causal interpretation of the DTF as a frequency domain measure of Sims causality and not—as has been claimed previously [39,40]—of Granger causality.

Finally, I note that estimates for the IRF and related quantities such as the DTF are usually computed from the autoregressive estimates  $\hat{A}_u$ . An alternative approach has been proposed by Jordà [41] based on (3.3). In contrast to the commonly used approach that uses the asymptotic  $\delta$ -method for statistical inference, the new method delivers estimates of the impulse response by regression methods.

## 4. Spurious causality

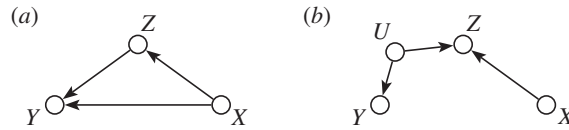
One problem with the empirical notion of Granger causality is its dependence on the set of variables included in the VAR model. It is well known that the omission of variables that affect two or more of the observed variables can induce conditional dependences among the observed variables that are wrongly interpreted as causal relationships by a Granger causality analysis. For instance, Porta *et al.* [7] have shown that respiration must be taken into account when analysing the causal relationships between heart period and systolic arterial pressure. In this section, I discuss criteria that help us to distinguish between such so-called *spurious causalities* and true cause–effect relationships.

A first step in this direction is the paper by Hsiao [42], who discussed causal patterns for vector time series of three variables  $X$ ,  $Y$  and  $Z$ , say. The general idea is that direct causes—described by Granger’s general definition—persist regardless of the background information used for the analysis, whereas indirect as well as spurious causes can be identified by either adding new variables to the analysis or removing already included variables. Comparing the Granger causal relations in the full trivariate model and a bivariate submodel, Hsiao distinguishes between the following cases:

- (a)  $X$  is a *direct cause* of  $Y$  if  $X$  Granger-causes  $Y$  in both the bivariate and the trivariate model;
- (b)  $X$  is an *indirect cause* of  $Y$  if (i)  $X$  Granger-causes  $Y$  in the bivariate model but not in the full model and (ii)  $X$  Granger-causes  $Z$  and  $Z$  Granger-causes  $Y$  in the trivariate model;
- (c)  $X$  is a *spurious cause of type II* for  $Y$  if (i)  $X$  Granger-causes  $Y$  in the bivariate but model not in the full model and (ii)  $Z$  Granger-causes  $X$  and  $Y$  in the trivariate model; and
- (d)  $X$  is a *spurious cause of type I* for  $Y$  if (i)  $X$  Granger-causes  $Y$  in the full model but not in the bivariate model and (ii)  $X$  Granger-causes  $Z$  and  $Z$  Granger-causes  $Y$  in the trivariate model.

A generalization of this classification approach to cases with more than three variables seems very complicated owing to the increasing number of configurations that need to be considered. Even for three variables, the above-mentioned four situations do not cover all possible situations; an example of such a system not covered by the above-mentioned cases will be studied later in this section.

Furthermore, it is clear that Hsiao’s definition of different types of causation suffers from the same defects as the empirical notion of Granger causality, since only the information of the three variables  $X$ ,  $Y$  and  $Z$  is taken into account. For instance, a ‘direct cause’ in a trivariate model might turn out to be indirect or even spurious (of type II) when adding further variables. Therefore, the term ‘cause’ in the above-mentioned definitions must be interpreted with some caution, and it would be more correct to speak of direct, indirect and spurious Hsiao causality.



**Figure 1.** Spurious causality of type I: (a) cancellation of direct and indirect effects—spurious causality is not encoded in the graph but requires additional constraints on the parameters; (b) confounding by unobserved variable  $U$ —the graph entails spurious causality without further constraints on the parameters. The graphs are Granger causality graphs in which each edge represents a Granger causal relationship between the adjacent variables.

Despite these shortcomings, Hsiao's definitions of spurious causality are of much interest for the final task of causal identification, as these phenomena provide important information about the underlying true causal structure. Therefore, I first discuss the two types of spurious causality in more detail.

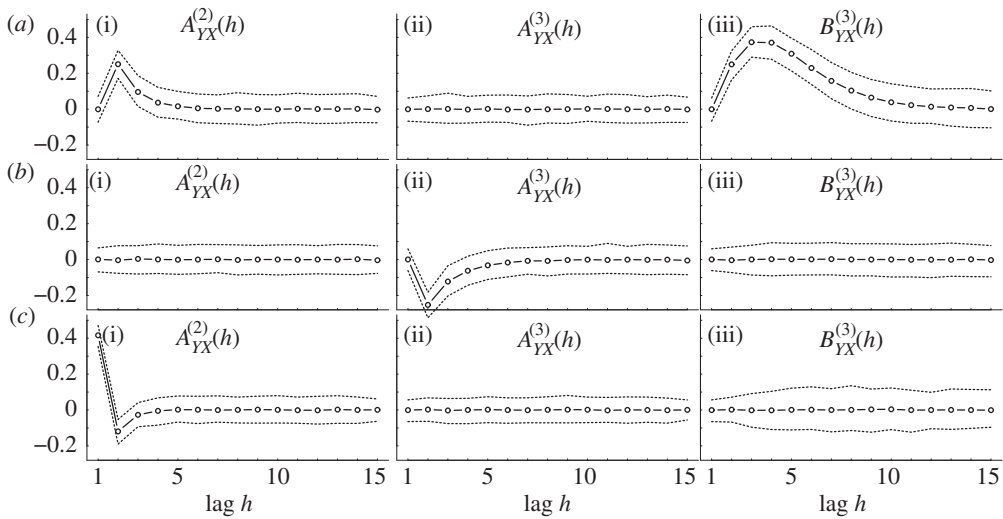
The intuitive understanding of spurious causalities is that omission of relevant variables induces associations among lagged variables that then—because of temporal precedence—are wrongly interpreted as causal relationships. If the relevant variables are added, the spuriously causal relationships vanish. This is expressed in the spurious causalities of type II, which can be detected by enlarging the set of variables. In order to avoid spurious causalities of this type, a multivariate VAR model with all available relevant variables included should be used. For spurious causalities that are induced by further latent variables, the above characterization however provides no help and they remain undetected.

In contrast, spurious causalities of type I are present in a large model and vanish only if the number of variables is reduced. According to the above-mentioned definition, a spuriously causal relationship of this type requires two causal pathways, namely one direct link and one indirect link via the third variable  $Z$  (figure 1a). However, as the causal relationship vanishes in the bivariate model, the effects of these two causal pathways on  $Y$  must cancel out and the total causal effect of  $X$  on  $Y$  is zero. This means that we can detect spurious causalities of type I by testing for Sims non-causality in the full model. This is illustrated in figure 2, which compares measures for bivariate and trivariate Granger causality as well as for trivariate Sims causality for three of the four cases considered by Hsiao. For the network with a spurious causality of type I, both the IRF and the bivariate autoregressive coefficients vanish.

Another far more plausible explanation for the phenomenon of spurious causality of type I is the network in figure 1b with an added fourth—unobserved—variable  $U$  that affects both  $X$  and  $Y$ . It has been shown elsewhere [3,20,43] that processes with autoregressive coefficients constrained by this graph show a spurious causality of type I without additional constraints on the parameters. In contrast, the cancellation of direct and indirect effects restricts the parameters to a lower-dimensional subset of the parameter space.

The main problem of causal inference is that in a larger network both types of spurious causalities might occur. Consider the network in figure 3a, with one latent variable  $U$ . Here  $X$  Granger-causes  $Y$  if  $Z$  is included in the analysis, owing to the spurious causality of type I. Because of the additional confounding by variable  $Z$ , variable  $X$  still Granger-causes  $Y$  if  $Z$  is omitted. According to Hsiao,  $X$  is a direct cause of  $Z$  even though the total causal effect is zero and manipulating  $X$  would be without consequences for  $Y$ . In contrast to Granger causality, which fails to detect this combination of spurious causalities of type I and II, the IRF indicates correctly that  $X$  is not Sims causal for  $Y$  with respect to  $(X, Y, Z)$ .

Finally, I note that for other networks correct causal identification requires the analysis of submodels. Figure 3b provides an example where there is no causal link from  $X$  to  $Y$  but  $X$  Granger-causes as well as Sims-causes  $Y$  with respect to the trivariate process. Only a bivariate analysis shows that  $X$  does not Granger-cause  $Y$  (and the same holds true for Sims causality).



**Figure 2.** Comparison of bivariate Granger causality (i), trivariate Granger causality (ii) and trivariate Sims causality (iii) for (a) indirect causality, (b) spurious causality of type I and (c) spurious causality of type II. The curves show means and standard deviations obtained from 400 repetitions of time series of length  $T = 1000$ .

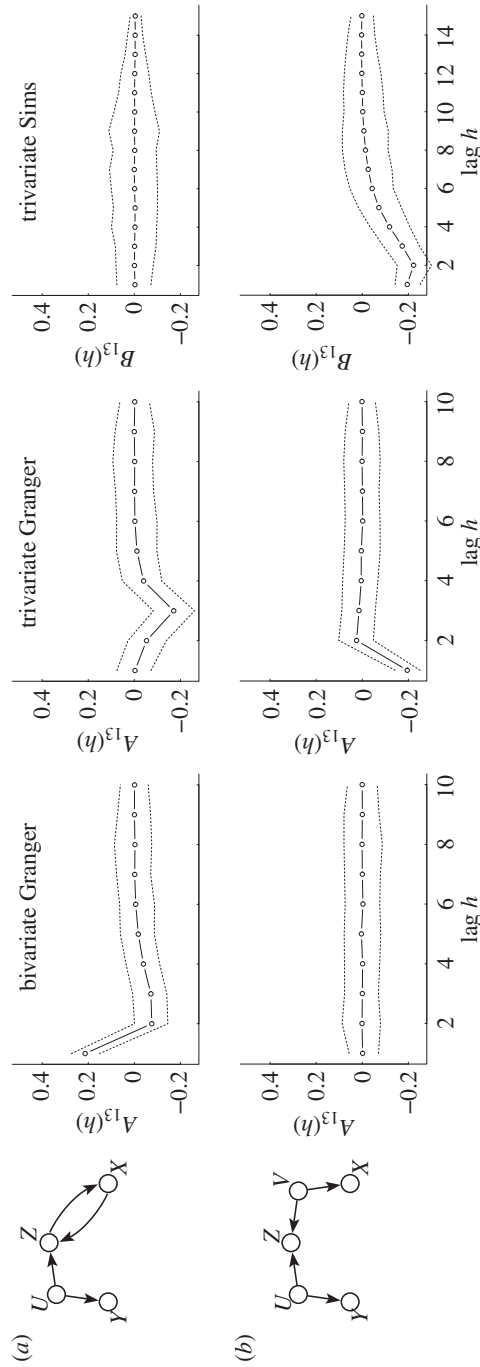
## 5. Causal identification

In the previous section, I described the problems in the empirical application of Granger and Sims causality. Despite the criticism that—in view of these problems—Granger causality is unsuited for causal inference, our examples have also shown that both Granger and Sims causality might contain some information about the network structure. In this section, I will review an approach for the (partial) identification of causal structures that uses these insights. For better readability, I avoid theoretical details such as graph-theoretic terminology and concentrate on the basic ideas; for a more detailed description of the approach, I refer the reader to the original articles [20,43].

The basic idea of the approach is to determine the set of all causal structures that are consistent with the empirically found Granger causal or Sims causal relationships. In contrast to the standard Granger causality analysis, we evaluate not only the full multivariate model, including all available variables, but also all possible submodels to detect spurious causalities of type I. Only if every causal structure that is consistent with the empirical findings from all these models features a causal link between two variables is this link identified as truly causal.

To motivate the approach, consider two variables  $X$  and  $Y$  such that  $X$  is Granger causal for  $Y$ . Without any further information, it is impossible to determine whether the empirically found association between  $X$  and  $Y$  is due to a true causal link or induced by some latent confounding variable. In contrast, if we have an additional variable  $Z$  such that  $Z$  is an indirect cause of  $Y$  in the sense of Hsiao, the link from  $X$  to  $Y$  must be truly causal. Similarly, if  $Z$  satisfies Hsiao's conditions for a spurious cause of type I for  $Y$ , the link from  $X$  and  $Y$  cannot be truly causal but must be induced by a confounding latent variable. I note that in both cases—without further information—nothing can be said about the nature of the link from  $Z$  to  $X$ .

Like similar approaches for ordinary multivariate data [10,13], the approach is based on two fundamental assumptions: the *causal Markov condition* and the *stability or faithfulness condition*. The causal Markov condition states that all dependences among the variables that are observed are due to the causal structure. This rules out time series that become dependent because of a common deterministic trend. The faithfulness condition states that also all independences among the variables are implied by the causal structure. In particular, this rules out that two or more causal links cancel each other out due to a particular choice of the parameters. This means that,



**Figure 3.** Comparison of direct and total causal effects of  $X$  on  $Y$  in two systems; variables  $U$  and  $V$  are latent. The curves show means and standard deviations obtained from 500 repetitions of time series of length  $T = 1000$ . The graphs are Granger causality graphs in which each edge represents a Granger causal relationship between the adjacent variables.

of the two possible explanations for spurious causalities of type I, we accept only the second with an additional latent confounding variable. Consequently, detection of spurious causalities of type I allows identifying spurious causalities of type II without knowing the confounding variable.

The identification algorithm itself consists of two steps. In the first step, all potential direct causal links among the components of a time series  $X$  are determined. Here, a variable  $X_i$  is said to be a potential cause of another variable  $X_j$  if  $X_i$  Granger-causes and Sims-causes  $X_j$  with respect to all processes  $X_S$ , where  $\{i, j\} \in S$ . More precisely, one also needs to consider certain combinations of Sims and Granger causality; below I give the precise condition but refer for details to the original work. The network of identified potential causes can be visualized by connecting two nodes  $i$  and  $j$  representing time series  $X_i$  and  $X_j$  by a dotted arrow  $i \cdots \rightarrow j$  whenever  $i$  is a potential cause of  $j$ .

The second step consists of trying to identify the exact nature of a potential causal link—whether it is a true or a spurious causal link. To achieve this, I make use of the fundamental difference between indirect causes and spurious causes of type I. As we have seen in §4, if  $X_i$  indirectly causes  $X_j$  with mediating variable  $X_k$ , then the causal link via  $X_k$  becomes broken by including  $X_k$  in the analysis. In the case of a spurious causality of type I, we find the opposite behaviour: the causal link from  $X_i$  to  $X_j$  induced by the variable  $X_k$  vanishes if  $X_k$  is omitted. Thus, depending on whether  $X_k$  needs to be included or omitted to make  $X_i$  Granger non-causal for  $X_j$ , we can label  $X_k$  as a potential cause of  $X_j$  as either a true cause or a spurious cause. Correspondingly, the edge  $k \cdots \rightarrow j$  in the graph is replaced by  $k \rightarrow j$  in the case of a true cause or by  $k \dashrightarrow j$  in the case of a spurious cause.

I have left out a few details such as contemporaneous associations, which are represented by dashed undirected edges ( $i \dashv j$ ). Furthermore, one needs to consider also certain combinations of Sims and Granger causality to identify potential causes. The exact condition is given below; for details, I again refer to the original papers.

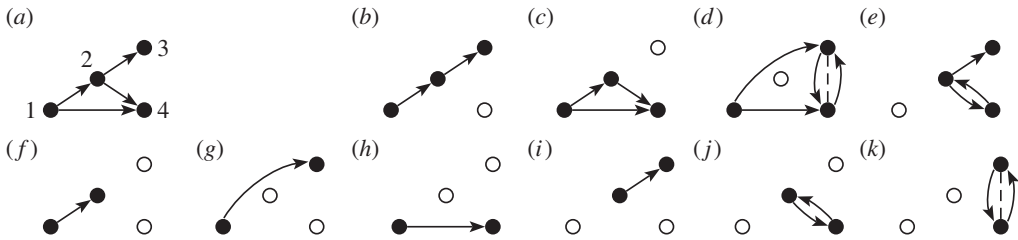
#### Definition 5.1 (Identification of adjacencies).

- (1) Insert  $a \dashv b$  whenever  $X_a$  and  $X_b$  are not contemporaneously independent with respect to  $X_V$ .
- (2) Insert  $a \cdots \rightarrow b$  whenever the following two conditions hold:
  - (i)  $X_a$  Granger-causes  $X_b$  with respect to  $X_S$  for all  $S \subseteq V$  with  $a, b \in S$ ;
  - (ii)  $X_{a,t-k}$  and  $X_{b,t+1}$  are not conditionally independent given  $X_{S_1}^t, X_{S_2}^{t-k}, X_a^{t-k-1}$  for some  $k \in \mathbb{N}$ , all  $t \in \mathbb{Z}$ , and all disjoint  $S_1, S_2 \subseteq V$  with  $b \in S_1$  and  $a \notin S_1 \cup S_2$ .

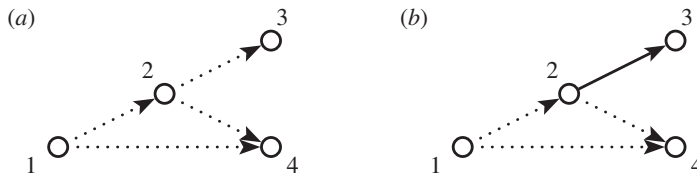
#### Definition 5.2 (Identification of tails).

- (1) *Spurious causality of type I*: Suppose that  $G$  does not contain  $a \cdots \rightarrow b$ ,  $a \rightarrow b$ , or  $a \dashrightarrow b$ . If  $a \cdots \rightarrow c \cdots \rightarrow b$  and  $X_a$  is Granger non-causal for  $X_b$  with respect to  $X_S$  for some set  $S$  with  $c \notin S$ , replace  $c \cdots \rightarrow b$  by  $c \dashrightarrow b$ .
- (2) *Indirect causality*: Suppose that  $G$  does not contain  $a \cdots \rightarrow b$ ,  $a \rightarrow b$ , or  $a \dashrightarrow b$ . If  $a \cdots \rightarrow c \cdots \rightarrow b$  and  $X_a$  is Granger non-causal for  $X_b$  with respect to  $X_S$  for some set  $S$  with  $c \in S$ , replace  $c \cdots \rightarrow b$  by  $c \rightarrow b$ .
- (3) *Ancestors*: If  $a \in \text{an}(b)$  replace  $a \cdots \rightarrow b$  by  $a \rightarrow b$ .
- (4) *Discriminating paths*: A fourth rule is based on the concept of discriminating paths. For details, I refer to Ali *et al.* [44].

Similar identification algorithms exist for the case of ordinary multivariate variables [10,13,45]. I further point out that the resulting graph is—even if all edges were classified as true causal or spurious causal—only identified up to Markov equivalence. This means that other causal structures, which imply the same set of Granger causal relations, are equally valid answers to the identification structure. As a consequence, it can be shown that directed edges in the graph cannot necessarily be interpreted as direct causal links, but that there must be at least one true causal pathway in the underlying causal structure.



**Figure 4.** Granger causality graphs for neuronal spike train data: (a) full model; (b–e) trivariate submodels; (f–k) bivariate submodels. Variables included in the model are represented by black nodes, omitted variables by white nodes. Directed edges represent Granger causal relationships with respect to the set of included variables; dashed (undirected) edges correspond to contemporaneous correlations.



**Figure 5.** Result of identification algorithm for neuronal spike train data: (a) identified adjacencies full model; (b) final graph showing identified true causal (solid edges) and spuriously causal (not present) as well as unidentified links (dotted edges).

For an illustration of the identification algorithm, I apply it to neuronal spike train data recorded from the lumbar spinal dorsal horn of a pentobarbital-anaesthetized rat during noxious stimulation. The firing times of 10 neurons were recorded simultaneously by a single electrode with an observation time of 100 s. The connectivity among the recorded neurons has been analysed previously by partial correlation analysis [46] and partial directed correlations [47].

For the analysis, I converted the spike trains of four neurons to binary time series. For the full as well as each trivariate and bivariate submodel, a VAR model of order  $p = 50$  was fitted and the associated Granger causality graph determined by a series of significance tests adjusted to have a false discovery rate of 5 per cent [48]. The resulting graphs are shown in figure 4. From these graphs, we first identify the causal adjacencies depicted in figure 5a; the adjacencies coincide with those of the Granger causality graph of the full model, which means that there are no spurious causalities of type I. Next, we try to determine the type of the edges by the rules of the identification algorithm. First, the graphs in figure 4a and 4b (or equivalently those in figure 4b and 4g) show that the first series causes the third series indirectly, which identifies the edge  $2 \rightsquigarrow 3$  as  $2 \rightarrow 3$ . For the edge  $2 \rightsquigarrow 4$ , a similar reasoning is not possible because the edge  $1 \rightsquigarrow 4$  prevents separation of nodes 1 and 4. The type of the remaining two edges cannot be identified as there are no edges with arrowheads at node 1 and hence none of the rules of the algorithm can be applied. Thus, we obtain the graph in figure 5b as the final result of the identification algorithm with only one true causal link identified.

## 6. Discussion of challenges and problems

Studying the applied literature on Granger causality, we are confronted with conflicting views favouring pairwise analysis or instead a full multivariate approach or claiming that Granger causality is not suited for causal inference at all as opposed to giving Granger causality a causal interpretation without much further thought.

This review has shown that the truth is somewhat in between and—as is usually the case—more complicated than the simple alternatives suggest. Causal inference based on Granger causality is indeed legitimate, but in many cases provides only sparse identification of true causal



relationships, that is, for most links among the variables it cannot be determined whether the link is truly causal or not. Moreover, instead of examining Granger causalities between pairs of variables or just in the full model, correct learning accumulates knowledge obtained from the large variety of possible submodels. This imposes feasibility constraints in the size of the networks that can be practically analysed. Summarizing our discussion shows that any analysis claiming full identification of the causal structure either must be based on very strong assumptions or prior information or—on closer inspection—turns out to be unwarranted.

Besides algorithmic and computation problems, there are further challenges to causal inference from time-series data:

- *Uncertainty in causal inference.* In most applications, causal inference is based on multiple tests to establish either multivariate or bivariante Granger causal relationships. A complete analysis of the causal structure requires the evaluation of a Granger causal relationship not only for the full multivariate process and pairs of series but also for all possible subprocesses. Consequently, the number of tests required to obtain the list of Granger non-causality relationships increases many times over. The wrong inclusion or exclusion of a directed edge at one level has not only a local effect but also a global effect owing to the classification of edges based on the global Markov property. I am not aware of any approaches that explicitly take into account this uncertainty, which limits the identification of causal structures (and hence effective connectivity) further.
- *Instantaneous causality.* In our discussion of causal inference, I have disregarded the possibility of the so-called instantaneous causality, that is, causal links between variables at the same point in time. The main problem with identification of instantaneous causality is the lack of directional information. In the case of Granger and Sims causality, the direction of the prima facie causal link is given by the temporal ordering. In the absence of such temporal precedence, the direction of the cause–effect relationship could be determined by additional subject matter information or by graph-based identification methods. So far, only a number of approaches exist for the identification of only the instantaneous causal structure [49–53].
- *Aggregation.* One possible explanation for instantaneous causality is the aggregation of variables over time (see the detailed discussion of instantaneous causality in [16]). For a correct evaluation of causal effects, it thus becomes necessary to include instantaneous causal structures in the modelling. A major problem arises in the case of feedback relationships between variables, as the standard causal models for multivariate observations are based on acyclic directed graphs, which must not contain cycles. It is not clear how these situations can be graphically represented and modelled.
- *Non-stationarity.* The standard framework for Granger causality analysis is stationary VAR models. Recordings taken to measure brain activity in behavioural situations are likely to violate the assumption of stationarity. Although there exist approaches for fitting VAR models with time-varying coefficients, tests for Granger causality require some kind of aggregation of the time-varying coefficients over time. Another problem arises if the causal structure itself is not stable in time. I am not aware of any approaches that deal with these problems.

## References

1. Goebel R, Roebroeck A, Kim D-S, Formisano E. 2003 Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn. Reson. Imaging* **21**, 1251–1261. (doi:10.1016/j.mri.2003.08.026)
2. Brovelli A, Ding M, Ledberg A, Chen Y, Nakamura R, Bressler SL. 2004 Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proc. Natl Acad. Sci. USA* **101**, 9849–9854. (doi:10.1073/pnas.0308538101)
3. Eichler M. 2005 A graphical approach for evaluating effective connectivity in neural systems. *Phil. Trans. R. Soc. B* **360**, 953–967. (doi:10.1098/rstb.2005.1641)

4. Deshpande G, Hu X, Stilla R, Sathian K. 2008 Effective connectivity during haptic perception: a study using Granger causality analysis of functional magnetic resonance imaging data. *NeuroImage* **40**, 1807–1814. (doi:10.1016/j.neuroimage.2008.01.044)
5. Havlicek M, Jan J, Brazdil M, Calhoun VD. 2010 Dynamic Granger causality based on Kalman filter for evaluation of functional network connectivity in fMRI data. *NeuroImage* **53**, 65–77. (doi:10.1016/j.neuroimage.2010.05.063)
6. Roebroeck A, Formisano E, Goebel R. 2011 The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. *NeuroImage* **58**, 296–302. (doi:10.1016/j.neuroimage.2009.09.036)
7. Porta A, Bassani T, Bari V, Pinna GD, Maestri R, Guzzetti S. 2012 Accounting for respiration is necessary to reliably infer Granger causality from cardiovascular variability series. *IEEE Trans. Biomed. Eng.* **59**, 832–841. (doi:10.1109/TBME.2011.2180379)
8. Cartwright N. 2007 *Hunting causes and using them*. Cambridge, UK: Cambridge University Press.
9. Pearl J. 1995 Causal diagrams for empirical research (with discussion). *Biometrika* **82**, 669–710. (doi:10.1093/biomet/82.4.669)
10. Pearl J. 2000 *Causality*. Cambridge, UK: Cambridge University Press.
11. Lauritzen SL. 2001 Causal inference from graphical models. In *Complex stochastic systems* (eds OE Barndorff-Nielsen, DR Cox, C Klüppelberg), pp. 63–107. London, UK: CRC Press.
12. Dawid AP. 2002 Influence diagrams for causal modelling and inference. *Int. Stat. Rev.* **70**, 161–189. (doi:10.1111/j.1751-5823.2002.tb00354.x)
13. Spirtes P, Glymour C, Scheines R. 2001 *Causation, prediction, and search*, 2nd edn (with additional material by D Heckerman, C Meek, GF Cooper and T Richardson). Cambridge, MA: MIT Press.
14. Granger CWJ. 1969 Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438. (doi:10.2307/1912791)
15. Granger CWJ. 1980 Testing for causality, a personal viewpoint. *J. Econ. Dyn. Control* **2**, 329–352. (doi:10.1016/0165-1889(80)90069-X)
16. Granger CWJ. 1988 Some recent developments in a concept of causality. *J. Econ.* **39**, 199–211. (doi:10.1016/0304-4076(88)90045-0)
17. Chen Y, Bessler SL, Ding M. 2006 Frequency decomposition of conditional Granger causality and application to multivariate neural field potentials. *J. Neurosci. Methods* **150**, 228–237. (doi:10.1016/j.jneumeth.2005.06.011)
18. Eichler M, Didelez V. 2007 Causal reasoning in graphical time series models. In *Proc. 23rd Conf. on Uncertainty in Artificial Intelligence*, 19–22 July, Vancouver, BC (eds R Parr, L van der Gaag). Corvallis, OR: AUAI Press.
19. Eichler M, Didelez V. 2010 On Granger-causality and the effect of interventions in time series. *Life Time Data Anal.* **16**, 3–32. (doi:10.1007/s10985-009-9143-3)
20. Eichler M. 2012 Causal inference in time series analysis. In *Causality: statistical perspectives and applications* (eds C Berzuini, AP Dawid, L Bernardinelli), pp. 327–354. Chichester, UK: Wiley.
21. Eichler M. 2012 Graphical modelling of multivariate time series. *Prob. Theory Related Fields* **153**, 233–268. (doi:10.1007/s00440-011-0345-8)
22. White H, Lu X. 2010 Granger causality and dynamic structural systems. *J. Fin. Econ.* **8**, 193–243. (doi:10.1093/jfinfec/nbq006)
23. Wright S. 1921 Correlation and causation. *J. Agric. Res.* **20**, 557–585.
24. Pearl J, Verma T. 1991 A theory of inferred causation. In *Principles of Knowledge Representation and Reasoning: Proc. 2nd Int. Conf., Cambridge, MA, 22–25 April* (eds JF Allen, RE Fikes, E Sandewall), pp. 441–452. San Mateo, CA: Morgan Kaufmann.
25. Good IJ. 1961 A causal calculus (I). *Br. J. Phil. Sci.* **11**, 305–318. (doi:10.1093/bjps/XI.44.305)
26. Good IJ. 1962 A causal calculus (II). *Br. J. Phil. Sci.* **12**, 43–51. (<http://www.jstor.org/stable/685537>)
27. Suppes P. 1970 *A probabilistic theory of causality*. Amsterdam, The Netherlands: North-Holland.
28. Schweder T. 1970 Composable Markov processes. *J. Appl. Prob.* **7**, 400–410. (doi:10.2307/3211973)
29. Sims CA. 1972 Money, income and causality. *Am. Econ. Rev.* **62**, 540–552.
30. Ramb R, Eichler M, Ing A, Thiel M, Weiller C, Grebogi C, Schwarzbauer C, Timmer J, Schelter B. 2013 The impact of latent confounders in directed network analysis in neuroscience. *Phil. Trans. R. Soc. A* **371**, 20110612. (doi:10.1098/rsta.2011.0612).

31. Lütkepohl H. 1993 *Introduction to multiple time series analysis*. New York, NY: Springer.
32. Faes L, Porta A, Nollo G. 2010 Testing frequency-domain causality in multivariate time series. *IEEE Trans. Biomed. Eng.* **57**, 1897–1906. (doi:10.1109/TBME.2010.2042715)
33. Brockwell PJ, Davis RA. 1991 *Time series: theory and methods*, 2nd edn. New York, NY: Springer.
34. Geweke JF. 1984 Measures of conditional linear dependence and feedback between time series. *J. Am. Stat. Assoc.* **79**, 907–915. (doi:10.1080/01621459.1984.10477110)
35. Baccalá LA, Sameshima K. 2001 Partial directed coherence: a new concept in neural structure determination. *Biol. Cybern.* **84**, 463–474. (doi:10.1007/PL00007990)
36. Eichler M. 2007 Granger causality and path diagrams for multivariate time series. *J. Econ.* **137**, 334–353. (doi:10.1016/j.jeconom.2005.06.032)
37. Box GEP, Jenkins GM, Reinsel GC. 1994 *Time series analysis, forecasting and control*, 3rd edn. Englewood Cliffs, NJ: Prentice-Hall.
38. Kamiński MJ, Blinowska KJ. 1991 A new method of the description of the information flow in the brain structures. *Biol. Cybern.* **65**, 203–210. (doi:10.1007/BF00198091)
39. Kamiński M, Ding M, Truccolo WA, Bressler SL. 2001 Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biol. Cybern.* **85**, 145–157. (doi:10.1007/s004220000235)
40. Eichler M. 2006 On the evaluation of information flow in multivariate systems by the directed transfer function. *Biol. Cybern.* **94**, 469–482. (doi:10.1007/s00422-006-0062-z)
41. Jordà O. 2005 Estimation and inference of impulse response by local projections. *Am. Econ. Rev.* **95**, 161–182. (doi:10.1257/0002828053828518)
42. Hsiao C. 1982 Autoregressive modeling and causal ordering of econometric variables. *J. Econ. Dyn. Control* **4**, 243–259. (doi:10.1016/0165-1889(82)90015-X)
43. Eichler M. 2009 Causal inference from multivariate time series: What can be learned from Granger causality. In *Logic, Methodology and Philosophy of Science, Proc. 13th Int. Congress, Beijing* (eds C Glymour, W Wang, D Westerståhl), pp. 481–496. London, UK: College Publications.
44. Ali RA, Richardson TS, Spirtes P. 2004 Markov equivalence for ancestral graphs. Technical Report no. 466, Department of Statistics, University of Washington, USA.
45. Zhang J, Spirtes P. 2005 A characterization of Markov equivalence classes for ancestral graphical models. Technical Report 168, Department of Philosophy, Carnegie Mellon University, USA.
46. Eichler M, Dahlhaus R, Sandkühler J. 2003 Partial correlation analysis for the identification of synaptic connections. *Biol. Cybern.* **89**, 289–302. (doi:10.1007/s00422-003-0400-3)
47. Dahlhaus R, Eichler M. 2002 Causality and graphical models for multivariate time series and point processes. In *IFMBE Proc. EMBEC 2002, Vienna*, vol. 3(2) (eds H Hutten, P Kroesl), pp. 1430–1431. Graz, Austria: Verlag Tu Graz.
48. Benjamini Y, Hochberg Y. 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B* **57**, 289–300. (<http://www.jstor.org/stable/2346101>)
49. Reale M, Tunnicliffe Wilson G. 2001 Identification of vector AR models with recursive structural errors using conditional independence graphs. *Stat. Methods Appl.* **10**, 49–65. (doi:10.1007/BF02511639)
50. Demiralp S, Hoover KD. 2003 Searching for the causal structure of a vector autoregression. *Oxford Bull. Econ. Stat.* **65**(Suppl.1), 745–767. (doi:10.1046/j.0305-9049.2003.00087.x)
51. Demiralp S, Hoover KD, Perez SJ. 2009 Still puzzling: evaluating the price puzzle in an empirically identified structural vector autoregression. Technical Report, Duke University, USA.
52. Moneta A. 2004 Graphical causal models and VAR-based macroeconometrics. PhD thesis, Sant'Anna School of Advanced Studies, Pisa, Italy.
53. Moneta A, Spirtes P. 2006 Graphical models for the identification of causal structures in multivariate time series models. In *Proc. 9th Joint Conf. on Information Sciences, Kaohsiung, Taiwan* (eds HD Cheng, SD Chen, RY Lin). Amsterdam, The Netherlands: Atlantis Press.