

Article

Causality in Reversed Time Series: Reversed or Conserved?

Jakub Kořenek ^{1,2}  and Jaroslav Hlinka ^{1,3,*} 

¹ Institute of Computer Science, Czech Academy of Sciences, Pod Vodarenskou Vezi 271/2, 182 07 Prague, Czech Republic; korenek@cs.cas.cz

² Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Brehova 7, 115 19 Prague, Czech Republic

³ National Institute of Mental Health, Topolova 748, 250 67 Klecany, Czech Republic

* Correspondence: hlinka@cs.cas.cz

Abstract: The inference of causal relations between observable phenomena is paramount across scientific disciplines; however, the means for such enterprise without experimental manipulation are limited. A commonly applied principle is that of the cause preceding and predicting the effect, taking into account other circumstances. Intuitively, when the temporal order of events is reverted, one would expect the cause and effect to apparently switch roles. This was previously demonstrated in bivariate linear systems and used in design of improved causal inference scores, while such behaviour in linear systems has been put in contrast with nonlinear chaotic systems where the inferred causal direction appears unchanged under time reversal. The presented work explores the conditions under which the causal reversal happens—either perfectly, approximately, or not at all—using theoretical analysis, low-dimensional examples, and network simulations, focusing on the simplified yet illustrative linear vector autoregressive process of order one. We start with a theoretical analysis that demonstrates that a perfect coupling reversal under time reversal occurs only under very specific conditions, followed up by constructing low-dimensional examples where indeed the dominant causal direction is even conserved rather than reversed. Finally, simulations of random as well as realistically motivated network coupling patterns from brain and climate show that level of coupling reversal and conservation can be well predicted by asymmetry and anormality indices introduced based on the theoretical analysis of the problem. The consequences for causal inference are discussed.

Keywords: causality; time reversal; temporal symmetry; reversed time series; vector autoregressive process; random networks; brain network; climate network



Citation: Kořenek, J.; Hlinka, J. Causality in Reversed Time Series: Reversed or Conserved? *Entropy* **2021**, *23*, 1067. <https://doi.org/10.3390/e23081067>

Academic Editors: Mohammad Reza Rahimi Tabar and Dimitris Kugiumtzis

Received: 19 May 2021

Accepted: 13 August 2021

Published: 17 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The temporal symmetry of physical processes, as well as the common fundamental lack of it, is among the most fascinating natural phenomena with deep theoretical consequences. Its characterization has been the topic of statistical physics including intriguing fundamental questions concerning quantum processes [1] as well as nonlinear time series analysis methods capturing the nonlinear aspects of brain electrophysiology [2]. Interestingly, it is intricately related to another fundamental concept used by humans to make sense of the processes in their natural environment—that of *causality*, with the notion of cause and effect as generalization of the naive observation that under (seemingly) equivalent circumstances, performing (or not) an action is followed by different event scenarios. While it is generally accepted that to uncover the causal principles ruling a given system one would need to be able to probe the systems responses experimentally [3], a range of methods were proposed that aim for estimation of the (network of) causal interactions just from the observed time series.

One of the prominent ideas behind a wide family of these methods is the principle that causality relates to predictability in that a variable can be considered causal with respect to

another, if using the current state of the former improves the prediction of the latter [4]. To avoid false inference due to possible shared information in the past or due to other variables, these other variables are in theory supposed to be also observed and used in both of the compared prediction models; of course in practice there are limitations on available observational data as well as estimability of the models that lead to the numerical methods providing only estimates of causal structure. However, the general definition represents an elegant and natural formalization of the intuitive understanding that the cause should precede the effect and carry some information about it (i.e., be at least statistically, if not deterministically related to the effect).

Notably, the relation between the two fundamental concepts of temporal symmetry and causality has recently become thematized [5–10]. One of the interesting conjectures, supported by observations from simulated processes, was that for linear processes, the direction of causal interaction is reversed upon the reversal of the direction of time, while it is conserved for nonlinear or chaotic time series [6,8]. The observation of causal direction reversal upon time-reversal in linear systems was previously noted in multiple studies, and previously used for construction of more robust causality scores [5,9,11], shown to potentially fail under the influence of an unobserved latent variable [8] (which is an important classical example of limitation of Granger causality and other time series analysis methods, which theoretically assume that all relevant variables are observed), and confirmed for Granger causality while not observed for a novel Compression-Complexity Causality measure [10].

In this work, we follow the observations in linear processes in more depth. In particular, we extend the analysis to multivariate systems, i.e., networks. For arbitrary network size, we start by building upon an almost forgotten result of Anděl [12] and derive the necessary and sufficient condition of causal structure reversal under time reversal for autoregressive processes of order 1; in this simplified model it turns out the key property is the normality of the causal interaction matrix, albeit for higher model orders the existence of such simplified conditions remains an open problem. In the previously studied bivariate case, we provide a simple analytic derivation illustrating that the causal structure reversal almost never happens perfectly, but only approximately (apart from two rather trivial cases with particular coupling symmetry), in line and further elucidating the previous reports of observations of approximate albeit not exact causal reversal. We consequently show that the minimal unidirectionally coupled network for which the exact causal structure reversal appears has 3 subsystems, and document this by an example. We finally turn to real-world data and simulation scenarios, studying the extent to which the causal structure reversal appears in linear vector autoregressive order one approximation of a brain and climate network under time reversal. The results show that already the linear approximations of both brain and climate systems show imperfect causal structure reversal under time reversal, while the extent to which this property is broken is closely predicted by deviation of the coupling matrix from normality (rather than from pure symmetry of the matrix) in both realizations of randomly connected networks as well as these more realistic connectivity structure scenarios. Finally, we discuss the relevance of these findings in a wider context, including how it problematizes the use of comparison with time-reversed version of a given process for inference of causal structures.

2. Theoretical Analysis

2.1. Reversed Autoregressive Process

In 1972, Anděl [12] proved a theorem describing a relation between a vector autoregressive process and its reversed process. First, we define the stationary autoregressive process and its reversed process according to his notation.

Definition 1 (Autoregressive process). Let $\{\mathbf{Y}_t\}_{t=-\infty}^{+\infty}$ be a series of uncorrelated n -dimensional random vectors such that $\mathbf{E}[\mathbf{Y}_t] = 0$ and covariance matrix of vector $\mathbf{Y}_t$, $\text{Var}[\mathbf{Y}_t] = \mathbb{I}$. Let $\mathbb{A}_0, \dots, \mathbb{A}_n$ be matrices from $\mathbb{R}^{n,n}$ such that:

1. $\det \mathbb{A}_0 \neq 0$
 2. $\mathbb{A}_n \neq 0$
 3. the equation $\det \left(\sum_{j=0}^p \mathbb{A}_j \lambda^{p-j} \right) = 0$ has all the roots smaller than 1 in absolute value
- and let $\{\mathbf{X}_t\}_{t=-\infty}^{+\infty}$ be a series defined by the recurrent formula

$$\sum_{j=0}^p \mathbb{A}_j \mathbf{X}_{t-j} = \mathbf{Y}_t \quad -\infty < t < +\infty, \quad (1)$$

or equivalently

$$\mathbf{X}_t = - \sum_{j=1}^p \mathbb{A}_0^{-1} \mathbb{A}_j \mathbf{X}_{t-j} + \mathbb{A}_0^{-1} \mathbf{Y}_t \quad -\infty < t < +\infty. \quad (2)$$

Then $\{\mathbf{X}_t\}$ is n -dimensional autoregressive process of order p (which is stationary due to conditions above) with the covariance function $R_{jk}(t) = \mathbf{E}[\mathbf{X}_t^j \mathbf{X}_0^k]$.

Please note that for an autoregressive process, the set of matrices $\mathbb{A}_i$ fully describes the causal interactions in the system.

Definition 2 (Reversed process). We say that $\{\mathbf{Z}_t\}$ is reversed (in time) with respect to an autoregressive process $\{\mathbf{X}_t\}$ with covariance function $R_{jk}(t)$, if $\{\mathbf{Z}_t\}$ has the matrix of the covariance functions $(R_{jk}(-t))_{j,k=1}^n$.

Anděl [12] proved the following theorem.

Theorem 1. Let $\{\mathbf{Z}_t\}$ be the n -dimensional autoregressive process defined as

$$\mathbf{Z}_t = - \sum_{j=1}^p \mathbb{B}_0^{-1} \mathbb{B}_j \mathbf{Z}_{t-j} + \mathbb{B}_0^{-1} \mathbf{Y}_t \quad -\infty < t < +\infty, \quad (3)$$

The series $\{\mathbf{Z}_t\}$ is reversed with respect to $\{\mathbf{X}_t\}$ if and only if its autoregressive matrices satisfy following conditions:

1. $\det \mathbb{B}_0 \neq 0$
 2. $\mathbb{B}_1 \neq 0$
 3. the equation $\det \left(\sum_{j=0}^p \mathbb{B}_j \lambda^{p-j} \right) = 0$ has all the roots smaller than 1 in absolute value
- and

$$\sum_{k=0}^{p-h} \mathbb{A}_{h+k}^\top \mathbb{A}_k = \sum_{k=0}^{p-h} \mathbb{B}_k^\top \mathbb{B}_{k+h} \quad 0 \leq h \leq p. \quad (4)$$

From the theorem, we know that the relation between the original process and the reversed process is given by $p + 1$ Equation (4). For simplicity, in the following we limit the analysis to the simplest but most commonly treated case of a system with delayed interactions with lag of one time step ($p = 1$) and with white noise, i.e., $\mathbb{A}_0^{-1} = \mathbb{I}$ (therefore $\mathbf{X}_t = -\mathbb{A}_1 \mathbf{X}_{t-1} + \mathbf{Y}_t$). In this case, the causal structure is given by a single matrix $\mathbb{A} := -\mathbb{A}_0^{-1} \mathbb{A}_1 = -\mathbb{A}_1$, which we shall further denote as the *coupling* matrix. Please note that non-zero elements of this matrix correspond to pairs of variables between which there is causal connection in the Granger sense, in particular, non-zero value of $\mathbb{A}_{ij}$ indicates non-zero input from $\mathbf{X}_j$ to $\mathbf{X}_i$, i.e., causal effect in the direction from j to i .

In this case the set of Equation (4) reduces to the following two matrix equations:

$$\mathbb{A}_1^\top = \mathbb{B}_0^\top \mathbb{B}_1 \quad (5)$$

$$\mathbb{I} + \mathbb{A}_1^\top \mathbb{A}_1 = \mathbb{B}_0^\top \mathbb{B}_0 + \mathbb{B}_1^\top \mathbb{B}_1. \quad (6)$$

We will derive the conditions under which the causality in the reversed process is reversed to the causality in the original process, i.e., when the coupling matrix of the reversed process $\mathbb{B} := -\mathbb{B}_0^{-1} \mathbb{B}_1$ is equal to the transpose of the coupling matrix of the original process $\mathbb{A}$.

Let us assume that $\mathbb{A} = \mathbb{B}^\top$; then $\mathbb{A}_1^\top = \mathbb{B}_0^{-1} \mathbb{B}_1$, i.e., $\mathbb{B}_1 = \mathbb{B}_0 \mathbb{A}_1^\top$. By substitution for $\mathbb{B}_1$ to (5) and (6) we obtain:

$$\mathbb{A}_1^\top = \mathbb{B}_0^\top \mathbb{B}_0 \mathbb{A}_1^\top \quad (7)$$

$$\mathbb{I} + \mathbb{A}_1^\top \mathbb{A}_1 = \mathbb{B}_0^\top \mathbb{B}_0 + \mathbb{A}_1 \mathbb{B}_0^\top \mathbb{B}_0 \mathbb{A}_1^\top = \mathbb{B}_0^\top \mathbb{B}_0 + \mathbb{A}_1 \mathbb{A}_1^\top. \quad (8)$$

Further, from Equation (8) we have that $\mathbb{B}_0^\top \mathbb{B}_0 = \mathbb{I} + \mathbb{A}_1^\top \mathbb{A}_1 - \mathbb{A}_1 \mathbb{A}_1^\top$. By substituting to Equation (7) we obtain

$$\mathbb{A}_1^\top = (\mathbb{I} + \mathbb{A}_1^\top \mathbb{A}_1 - \mathbb{A}_1 \mathbb{A}_1^\top) \mathbb{A}_1^\top, \quad (9)$$

which further simplifies to:

$$0 = (\mathbb{A}_1^\top \mathbb{A}_1 - \mathbb{A}_1 \mathbb{A}_1^\top) \mathbb{A}_1^\top. \quad (10)$$

Equation (10) thus constitutes a *necessary* condition for coupling matrix reversal. Please note that under the reasonably general assumption that $\mathbb{A}$ (and therefore $\mathbb{A}_1^\top$) is regular, the range of the operator $\mathbb{A}_1^\top$ is the full space $\mathbb{R}^n$ and thus Equation (10) holds *only* if $0 = \mathbb{A}_1^\top \mathbb{A}_1 - \mathbb{A}_1 \mathbb{A}_1^\top$, i.e., $\mathbb{A}_1$ is normal. On the other hand, any normal matrix $\mathbb{A}_1$ fulfils Equation (10), and together with an arbitrary choice of a unitary matrix $\mathbb{B}_0$ and $\mathbb{B}_1 = \mathbb{B}_0 \mathbb{A}_1^\top$ it fulfils the original equations Equations (5) and (6).

Overall, we proved that for VAR(1) process with normal coupling matrix $\mathbb{A}_1$ and white noise ($\mathbb{A}_0 = \mathbb{I}$), the coupling structure in the reversed process is exactly reversed to coupling matrix of the original process. We also proved that for a regular matrix $\mathbb{A}_1$, observation of the reversed coupling structure in the reversed process implies normality of the coupling matrix $\mathbb{A}_1$. Therefore, under the assumption of regularity of $\mathbb{A}_1$, normality of $\mathbb{A}_1$ is equivalent to the observation of reversed causality in the reversed process.

A specific subset of normal matrices are symmetric matrices. From the above results concerning normal matrices, we thus learn that for symmetric matrices the causality of time-reversed process is also reversed, which due to its symmetry means that it is *invariant* (conserved) on time reversal of the time series (as the coupling matrix and its transpose are identical). One may ask whether there is a broader family of coupling matrices for which the coupling matrix of the original process and the reversed process is identical. Assuming that $\mathbb{A}_1 = \mathbb{B}_0^{-1} \mathbb{B}_1$, then $\mathbb{B}_1 = \mathbb{B}_0 \mathbb{A}_1$ and from (5) and (6) follows:

$$\mathbb{A}_1^\top = \mathbb{B}_0^\top \mathbb{B}_0 \mathbb{A}_1 \quad (11)$$

$$\mathbb{I} + \mathbb{A}_1^\top \mathbb{A}_1 = \mathbb{B}_0^\top \mathbb{B}_0 + \mathbb{A}_1^\top \mathbb{B}_0^\top \mathbb{B}_0 \mathbb{A}_1. \quad (12)$$

Under assumption of the regularity of $\mathbb{A}_1$ then follows $\mathbb{B}_0^\top \mathbb{B}_0 = \mathbb{A}_1^\top \mathbb{A}_1^{-1}$ and

$$\mathbb{I} + \mathbb{A}_1^\top \mathbb{A}_1 = \mathbb{A}_1^\top \mathbb{A}_1^{-1} + \mathbb{A}_1^\top \mathbb{A}_1^\top, \quad (13)$$

from which we obtain the condition

$$\left(\mathbb{A}_1 - \mathbb{A}_1^{-1}\right) = \left(\mathbb{A}_1 - \mathbb{A}_1^{-1}\right)^{\top}. \quad (14)$$

In conclusion, we found normality and symmetry of the coupling matrix as sufficient conditions for reversal and conservation of the coupling under time-reversal, albeit in principle these may not be necessary depending on the existence of other eligible solutions for Equations (10) and (14)—we leave the characterization of the full set of solutions for these matrix equations under the additional requirements of Definition 1 and Theorem 1 as an open problem for further work, although based on preliminary analysis we conjecture the solutions might in fact be equivalent with the normality and symmetry condition, respectively.

2.2. Bivariate Case

Above, we discussed the conditions for the reversed process to have reversed, or conserved causality for multivariate autoregressive processes (for simplicity limiting the analysis to processes of order 1, with white noise). Let us analyze in more detail the bivariate case, i.e., model of the systems with two elements X and Y .

The convention:

$$\mathbb{A} = -\mathbb{A}_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (15)$$

leads after substitution to Equation (10) to the set of equations:

$$\begin{aligned} -ac^2 + abc + b^2d - bcd &= 0, \\ -c^3 + b^2c - cd^2 + acd + bd^2 - abd &= 0, \\ -b^3 - a^2b + bc^2 + abd + a^2c - acd &= 0, \\ -b^2d + bcd + ac^2 - abc &= 0, \end{aligned} \quad (16)$$

which has two types of solutions:

$$\mathbb{A} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}, \quad \text{and} \quad \mathbb{A} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad (17)$$

in the sense that all solutions have this form and any matrix of this form is a solution. These are specific cases of either a symmetric matrix ($c = b$) or a specific case of an off-diagonally antisymmetric matrix ($c = -b$, $d = a$).

A particularly interesting case is that of unidirectional coupling, without loss of generality one can focus on the case where Y affects X , but X does not affect Y . This model is given by the matrix:

$$\mathbb{A} = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}, \quad (18)$$

where $b \neq 0$, while $c = 0$.

Please note that this is the canonical model of unidirectional causality in a linear system, that was used to illustrate the conjecture that in a linear system (unlike in nonlinear, in particular chaotic, systems), time-reversal leads to reversal of the direction of causality [6].

Clearly, this is a specific case of the bivariate system considered above, and due to the condition $b \neq 0$, it is not among the solutions for Equation (10). Therefore we know that in this model a *perfect* reversal of causal structure with time reversal does not occur, albeit the *main/stronger/dominant*, direction of causality indeed typically reverts, as discussed by Paluš et al. [6] for simple linear systems, while giving counterexamples for other, typically

chaotic nonlinear systems. Please note that while Paluš et al. [6] discussed the coupling direction reversal in context of specific causal indices, in particular the conditional mutual information, the notion of comparing strength of causality in one and the other direction between two variables of the system is used in practice, and under some assumptions comparing the strength of some index of causality helps inference on causal direction from observed data [5,11]. However, the use of the comparison of arbitrary causal indices in both directions is not necessarily valid and might in principle provide different and indeed wrong answers. For instance, one can conceive scenarios where stronger coupling coefficient in the dynamical equations provides a weaker statistical significance due to different temporal scales of the dynamics or signal complexity [13]. Additionally, the value of the coupling coefficient can give a different result than of some causality index. We shall get back to the topic of conservation of the dominant direction of causality in the Discussion section, where we shall separately consider the values of the coupling coefficients, and of the Granger causality indices.

Let us study the bivariate system with unidirectional coupling in more detail. As shown earlier, in case that $\mathbb{A}$ is regular, perfect causal reversal occurs if and only if $\mathbb{A}$ is normal. Even without relying on the provided solutions of the general conditions (16) one can easily show that there is no normal matrix among regular matrices of type (18), i.e., among those that fulfill both $a \neq 0$ and $d \neq 0$. In particular, for the triangular coupling matrix $\mathbb{A}$, the condition of normality $\mathbb{A}\mathbb{A}^\top = \mathbb{A}^\top\mathbb{A}$ rewrites as:

$$\begin{pmatrix} a^2 + b^2 & bd \\ bd & d^2 \end{pmatrix} = \begin{pmatrix} a^2 & ab \\ ab & b^2 + d^2 \end{pmatrix} \quad (19)$$

and therefore $a^2 + b^2 = a^2$ which is in contradiction with $b \neq 0$.

It is also illustrative to consider the singular cases. While perfect causal reversal does not occur (unless in the trivial case of $b = 0$), in one of the cases at least the binary structure of the causal links is reversed (albeit not the exact coefficients). There are only two options such that $-\mathbb{A}_1$ is singular: $a = 0$ or $d = 0$.

If $d = 0$,

$$\mathbb{A} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}, \quad (20)$$

which leads to (See Appendix A for detailed derivation):

$$\mathbb{B} = \begin{pmatrix} a & 0 \\ \frac{b(1-a^2)}{1+b^2} & 0 \end{pmatrix}. \quad (21)$$

Please note that perfect causal reversal does not happen as $b = \frac{b(1-a^2)}{1+b^2}$ only if $a^2 + b^2 = 0$, which is in contradiction with the assumption $b \neq 0$.

If $a = 0$,

$$\mathbb{A} = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix}, \quad (22)$$

the reverse process has a coupling matrix

$$\mathbb{B} = \frac{1}{1+b^2} \begin{pmatrix} b^2d & bd^2 \\ b & d \end{pmatrix}. \quad (23)$$

Here, the reverted process shows *bidirectional* coupling, and perfect causal reversal happens again only in the trivial case of $b = \frac{b}{1+b^2}$, i.e., $b = 0$. See Appendix A for detailed derivation.

As we showed that perfect causal reversal does not happen for unidirectionally coupled bivariate AR(1) models with white noise, it is a natural question whether it can

happen at all in such systems, and in which dimension. The answer is that the smallest dimension in which a perfect causal reversal occurs in a nontrivial matrix is three. Consider the example system with three elements with a normal coupling matrix $\mathbb{A}$ the structure of which is depicted in Figure 1.

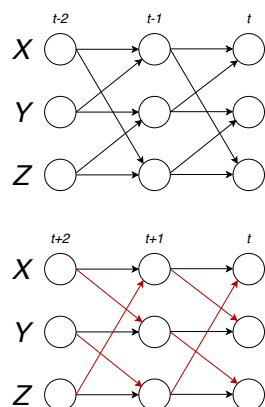


Figure 1. System with normal coupling matrix and its reversed system.

The binary structure of the coupling matrix of this system is:

$$\mathbf{S} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad (24)$$

and by rescaling it by a suitable coefficient $\frac{s}{\lambda_{\max}}$, where $\lambda_{\max}$ is the largest eigenvalue (in absolute value) of the matrix $\mathbf{S}$, and $s \in (0, 1)$ is an optional parameter—in this case $s = 0.8$ —we assure stationarity of the corresponding AR process. A minimal example of a unidirectionally coupled AR system exhibiting perfect coupling reversal upon time reversal would thus be given by for instance:

$$\mathbb{A} = \begin{pmatrix} 0.4 & 0.4 & 0 \\ 0 & 0.4 & 0.4 \\ 0.4 & 0 & 0.4 \end{pmatrix}. \quad (25)$$

As the matrix $\mathbb{A}$ is normal, the coupling matrix $\mathbb{B}$ of the reversed process is exactly the transpose of $\mathbb{A}$:

$$\mathbb{B} = \begin{pmatrix} 0.4 & 0 & 0.4 \\ 0.4 & 0.4 & 0 \\ 0 & 0.4 & 0.4 \end{pmatrix}. \quad (26)$$

3. Numerical Simulations

In the previous section, we discussed the conditions under which the time reversed process has reversed or conserved coupling matrix. In particular, we showed that the sufficient (and under regularity even necessary condition) for the former is the normality of the coupling matrix $\mathbb{A}_1$, while for the latter (conservation), the condition is given by Equation (14); which is clearly fulfilled for any symmetric matrix $\mathbb{A}_1$.

In real-world systems, one might expect that the coupling matrix does not fulfill the normality condition, and therefore even in linear auto-regressive systems, the causal structure of the time-reversed process may thus not be exactly reversed. Indeed, already the analysis of the bivariate case showed that strict reversal generally does not occur, and unidirectionally coupled system might show as bidirectionally coupled after time reversal; in fact this is the *generic* case in case the coupling source being autocorrelated due to self-coupling. Therefore it is natural to ask to what extent does the reversed time series causal structure differ from the exactly reversed in example real-world complex

systems, and whether this deviation from coupling reversal is somehow related to (and can be reasonably predicted from) the properties of the coupling matrix. In particular, we conjecture that it might be well approximated by the quantitative deviation from the normality condition that would enforce the exact coupling reversal.

3.1. Measures Introduction

In the following, we use numerical simulations to show that the level to which the causality is reverted (conserved) on time reversal is indeed well predicted by the deviation of the coupling matrix from normality (symmetry). In other words, that the difference between the matrices $\mathbb{A}$ and $\mathbb{B}^\top$ strongly depends on the level (a)normality of $\mathbb{A}$, and likewise the difference between $\mathbb{A}$ and $\mathbb{B}$ depends on the (a)symmetry of $\mathbb{A}$. To this end, we will first formally introduce indices of anormality and asymmetry. Please note that these indices are intended to characterize these particular properties of the matrix itself, rather than provide some information concerning the existence of specific causal links (so they are not to be confused with various causality indices used in causal inference). In particular, the analyses below suggest that the indices of deviation of the coupling matrix from symmetry and normality correlate substantially with the level of causal reversal/conservation, and can thus provide a proxy for whether causal reversal/conservation can be expected in particular (type of) causal structure. Despite their theoretical usefulness for understanding of the causal reversal/conservation behaviour, their practical application might be of course limited by the fact that many real-world systems are not well approximated by a VAR(1) process.

For a square matrix $\mathbb{A}$ we define the symmetric part of the matrix $\mathbb{A}_s = \frac{1}{2}(\mathbb{A} + \mathbb{A}^\top)$ and the anti-symmetric part $\mathbb{A}_a = \frac{1}{2}(\mathbb{A} - \mathbb{A}^\top)$. The deviation from symmetry is then defined as $\frac{\|\mathbb{A}_a\|}{\|\mathbb{A}_s\|}$, where $\|\cdot\|$ is some matrix norm. In this approach, for a symmetric matrix the value of the deviation from symmetry is equal to zero and for an antisymmetric it tends to infinity, and that for non-negative matrices it is normalized into the $[0, 1]$ interval.

In this paper, we use the Frobenius norm which is for a square matrix $\mathbb{A} \in \mathbb{R}^{n,n}$ defined as

$$\|\mathbb{A}\|_F = \left(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2}. \quad (27)$$

The resulting measure of asymmetry—the deviation from symmetry—is then defined as

$$d_{Sym}(\mathbb{A}) = \frac{\|\mathbb{A} - \mathbb{A}^\top\|_F}{\|\mathbb{A} + \mathbb{A}^\top\|_F}. \quad (28)$$

Based on the same idea, we define a deviation from normality (matrix is normal if $\mathbb{A}\mathbb{A}^\top = \mathbb{A}^\top\mathbb{A}$) as

$$d_{Norm}(\mathbb{A}) = \frac{\|\mathbb{A}\mathbb{A}^\top - \mathbb{A}^\top\mathbb{A}\|_F}{\|\mathbb{A}\mathbb{A}^\top + \mathbb{A}^\top\mathbb{A}\|_F}. \quad (29)$$

From the analysis outlined in Section 2, we know that a relation $\mathbb{A} = \mathbb{B}^\top$ holds for normal matrices ($\mathbb{A}\mathbb{A}^\top = \mathbb{A}^\top\mathbb{A}$) as well as $\mathbb{A} = \mathbb{B}$ holds for symmetric matrices ($\mathbb{A} = \mathbb{A}^\top$). In the following simulation we will show the relation between the deviation of matrix $\mathbb{A}$ from symmetry (28) and the difference between matrices $\mathbb{A}$ and $\mathbb{B}$. For this purpose we define the normalized difference of these matrices as

$$\overleftarrow{d}_{Sym}(\mathbb{A}) = \frac{\|\mathbb{A} - \mathbb{B}\|_F}{\|\mathbb{A} + \mathbb{B}\|_F}. \quad (30)$$

Additionally, we investigate the relation between anormality (29) and the normalized difference between $\mathbb{A}$ and $\mathbb{B}^\top$, defined as

$$\overleftarrow{d}_{Norm}(\mathbb{A}) = \frac{\|\mathbb{A} - \mathbb{B}^\top\|_F}{\|\mathbb{A} + \mathbb{B}^\top\|_F}. \quad (31)$$

Please note that we consider processes with white noise (zero mean, variance equals to one, that is VAR(1) process in the form

$$\mathbf{X}_t = \mathbb{A}\mathbf{X}_{t-1} + \mathbf{Y}_t, \quad (32)$$

with white noise $\mathbf{Y}_t$.

3.2. Simple Causal Structures

In this part we will support theoretical analysis with numerical simulations. According to the expression (32) the time series of length T are generated with different choices of matrix $\mathbb{A}$. These time series are then reversed in time, and the coupling matrix is subsequently estimated using the *arfit* function implemented in the MATLAB software.

As a validation example, we use the matrix $\mathbb{A}$ introduced in Equation (25). We obtain the following example estimates for time series of length $T \in \{10^2, 10^4, 10^6\}$ respectively, showing reasonably fast convergence of the estimates to the theoretically derived matrix $\mathbb{B}$ given in Equation (26):

$$\mathbb{B} \sim \begin{pmatrix} 0.290 & 0.119 & 0.363 \\ 0.212 & 0.397 & 0.005 \\ 0.152 & 0.336 & 0.286 \end{pmatrix}, \begin{pmatrix} 0.387 & -0.009 & 0.391 \\ 0.388 & 0.400 & 0.015 \\ -0.012 & 0.410 & 0.401 \end{pmatrix}, \begin{pmatrix} 0.400 & 0.002 & 0.400 \\ 0.402 & 0.399 & 0.001 \\ 0.000 & 0.399 & 0.401 \end{pmatrix}. \quad (33)$$

In the next simulation we explore to what level does the deviation from normality $d_{Norm}(\mathbb{A})$ determine the imperfection of causal structure reversal $\overleftarrow{d}_{Norm}(\mathbb{A})$, and similarly how does the deviation from symmetry $d_{Sym}(\mathbb{A})$ relate to the difference between the causal structure of the forward and time-reversed process $\overleftarrow{d}_{Sym}(\mathbb{A})$.

For this purpose, we use network topologies defined by a random binary directed graph given by a binary matrix of dimension $n \times n$ with value 1 of the diagonal elements (corresponding to loops) and with a total density D of nonzero elements. Each such realization is scaled by a constant $\frac{s}{\lambda_{\max}}$, where we set $s = 0.8$ to provide a 'stable' matrix $\mathbb{A}$; from this the indices $d_{Sym}(\mathbb{A})$ and $d_{Norm}(\mathbb{A})$ are calculated. Subsequently time series of length $T = 10^5$ are generated according to Equation (32). The time series are reversed, and the matrix $\mathbb{B}$ is estimated using the *arfit* function and finally $\overleftarrow{d}_{Sym}(\mathbb{A})$ and $\overleftarrow{d}_{Norm}(\mathbb{A})$ is calculated. For the robustness of calculations, the whole process is repeated for 1000 different realizations of random matrix $\mathbb{A}$. Two particular settings of the parameters are used to provide a comparison with subsequent analysis of complex real-world connectivity structures, namely $n = 42$, $D = 15\%$ and $n = 90$, $D = 5\%$.

The results concerning the prediction of conservation of coupling matrix under time series reversal from symmetry of the input matrix $\mathbb{A}$ are shown in Figure 2 Left and Figure 3 Left. It is apparent, that the asymmetry of the coupling matrix $d_{Sym}(\mathbb{A})$ provides a very good prediction of the conservation of the coupling structure under time reversal $\overleftarrow{d}_{Sym}(\mathbb{A})$, with the respective Pearson correlation coefficient attaining the value of $\rho = 0.99$ in both parameter settings. Similarly, Figure 2 Right and Figure 3 Right demonstrate that the deviation of the coupling matrix from normality quantified by $d_{Norm}(\mathbb{A})$ provides a very strong prediction of the coupling matrix reversal $\overleftarrow{d}_{Norm}(\mathbb{A})$, albeit weaker than in the case of the coupling conservation, with values $\rho = 0.89$ and $\rho = 0.88$ respectively.

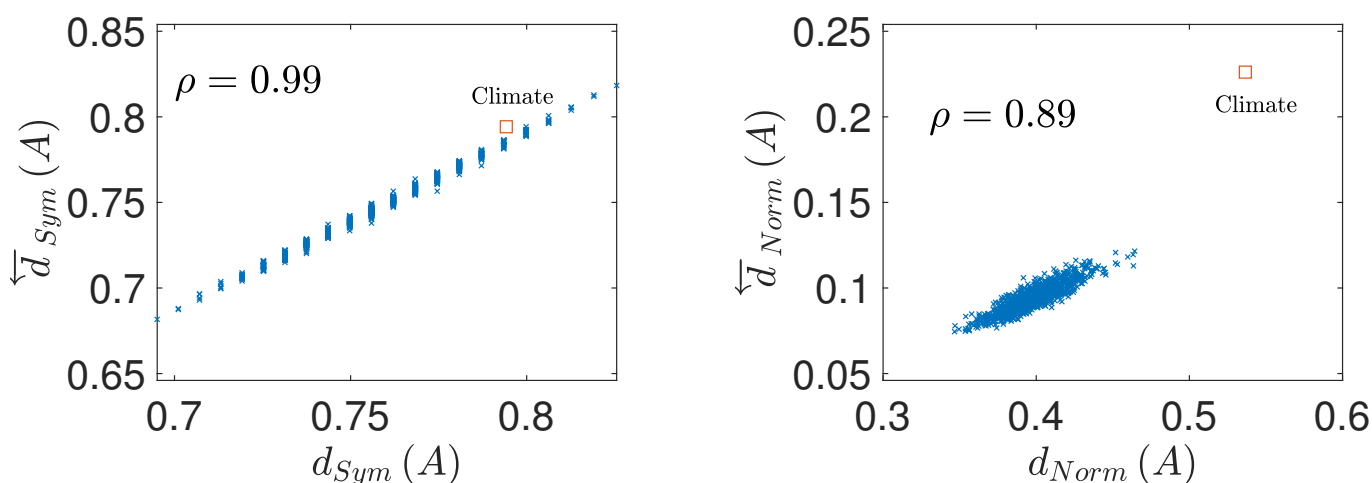


Figure 2. Prediction of the causal structure conservation and reversal from the coupling matrix properties, in a sample of random directed matrices with full diagonal (matrix size $n = 42$, density of nonzero elements $D = 15\%$). **(Left):** causal structure conservation under time reversal $\overleftarrow{d}_{Sym}(A)$ as function of the coupling matrix symmetry $d_{Sym}(A)$. **(Right):** causal structure reversal under time reversal $\overleftarrow{d}_{Norm}(A)$ as function of the coupling matrix normality $d_{Norm}(A)$. ρ denotes the correlation coefficient between the considered quantities. The red square denotes the values for the climate model.

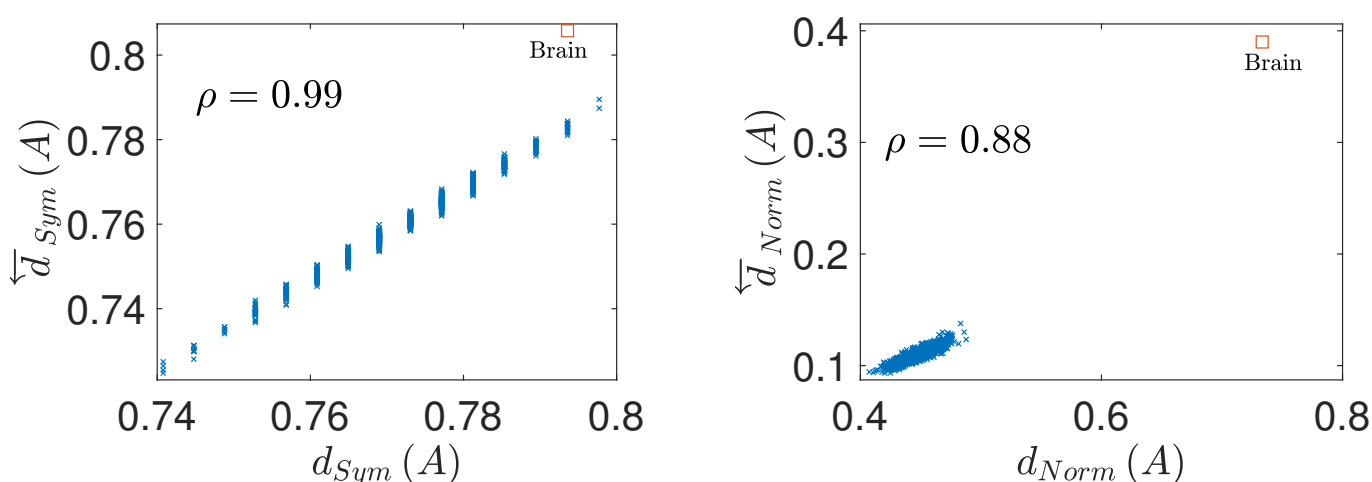


Figure 3. Prediction of the causal structure conservation and reversal from the coupling matrix properties, in a sample of random directed matrices with full diagonal (matrix size $n = 90$, density of nonzero elements $D = 5\%$). **(Left):** causal structure conservation under time reversal $\overleftarrow{d}_{Sym}(A)$ as function of the coupling matrix symmetry $d_{Sym}(A)$. **(Right):** causal structure reversal under time reversal $\overleftarrow{d}_{Norm}(A)$ as function of the coupling matrix normality $d_{Norm}(A)$. ρ denotes the correlation coefficient between the considered quantities. The red square denotes the values for the brain model.

3.3. Complex Connectivity Structures

So far, we showed that somewhat counter-intuitively, even in linear processes the time reversal has nontrivial effect on the coupling structure—in a general case it neither conserves nor reverses it, while these two special phenomena appear under quite specific conditions of normality (or more generally under fulfillment of the conditions given by Equation (10)) and symmetry (14), respectively. In the previous subsection, we demonstrated that indeed for random (ER) graphs, the coupling matrices of the time reversed process are relatively far from purely reversed or conserved causality, with the variability in the level of conservation/reversal being very well predicted by the symmetry/normality of the original (random) coupling matrix.

This motivates a question concerning the behaviour of more realistic models of real-world complex structures. To this end, we analyze the time-reversal effect on the coupling structure of linear vector autoregressive approximation of two archetypal examples of complex systems, the brain network and the climate network. We use the simple order 1 linear autoregressive process approximation of these processes, which is in line with the model order for which we derived that simplified condition for causal reversal/conservation. Therefore, the application of the corresponding indices of deviation from normality and symmetry is here well motivated. However, it should be kept in mind, that these model network examples, albeit realistically motivated, of course fall short of the complexity of the real brain and climate—both in terms of not capturing the nonlinear aspect of the brain/climate dynamics, and in simplifying the temporal dynamics by limiting the interactions to a single temporal lag. Indeed, even at the level of spatial and temporal sampling of the motivating datasets, the linear VAR(1) model is just an approximation, rather than an optimal model for the data—as model order criteria such as the Bayesian Information Criterion would suggest that a higher order VAR model would better describe the generating process. However, the quantification of the (deviation from) the causal reversal condition for higher-order VAR processes is not available in such a compact form; so we use only these order-one approximation of the real-world data coupling structures as motivating examples of somewhat more realistic connectivity matrices than provided by the ER random graph model.

3.3.1. Earth's Climate Network

The first example network is an approximation of a large-scale climate interaction network, constructed from regional daily temperature time series. The network has 42 nodes and was obtained by thresholding (binarizing) the interaction matrix in a data-fitted VAR(1) model to 15 percent density. In more detail, the original dataset contains the daily surface air temperature anomalies data obtained from the NECP/NCAR reanalysis dataset [14,15]. In particular, we used the daily air temperature fields at 1000 hPa level, spanning the period from 1 January 1948 to 31 December 2012 and sampled at $2.5^\circ \times 2.5^\circ$ angularly regular Gaussian grid. This original data contains more than 10,000 time series—a relatively dense grid covering the whole globe. The data were downsampled by remapping to a more suitable 42-point grid by applying an interpolation scheme based on the quasi-isotropic icosahedral grids [16] by a procedure implemented in the SCRIP package [17]. Details of the data origin and preprocessing are described in [18,19].

3.3.2. Brain Network

The other network example is an approximation of brain interaction network that has 90 nodes and was obtained by thresholding (binarizing) the interaction matrix in a data-fitted VAR(1) model to 5% density. We use data obtained as part of a study on healthy subjects brain activity. The activity in 90 brain regions of 84 subjects were measured by functional magnetic resonance imaging, and temporally concatenated in order to provide sufficiently long time series (20,160 time points in total) for estimation of the coupling matrix. Details of the data origin and preprocessing are described in [18,19].

3.3.3. Data Processing

As mentioned above, in preparing an approximate model of these real-world complex systems, we make the approximation that these datasets correspond to realizations of a VAR(1) process. Please note that albeit quite coarse and not capturing either the nonlinear aspects of the dynamics as well the plethora of time scales of these complex systems, the linear (autoregressive) approximation has been shown to reproduce sufficiently a range of properties of both the brain and climate network structure [20–23] and provide robust and interpretable estimate of causal interactions [22,24]. The estimated coupling matrix is then thresholded to predefined density (low values set to zero, only predefined amount of entries retained, in this case 15% for the climate matrix, 5% for the brain matrix),

binarized (non-zero values set to 1) and subsequently normalized by the constant $\frac{s}{\lambda_{\max}}$ for the stationarity of the resulting process (all simulations were provided with setting $s = 0.8$). This matrix defines the VAR(1) process from which we generate samples of arbitrary length T (in the simulations, we set $T = 10^5$). The sample is then reverted and a VAR(1) model fitted to it using the *arfit* MATLAB function, providing thus an estimate of the matrix $\mathbb{B}$. This matrix is then used to obtain an estimate of the values of $\overleftarrow{d}_{\text{Norm}}(\mathbb{A})$ and $\overleftarrow{d}_{\text{Sym}}(\mathbb{A})$.

3.3.4. Analysis

The resulting values of deviation from symmetry and normality of the random as well as brain and climate coupling matrices, and imperfection of their coupling conservation/reversal upon time series reversal are shown in Figures 2 and 3. How do the real-world matrices compare to the sets of matched random matrices? Clearly, both the brain and climate network show level of asymmetry relatively comparable to the directed random matrix of the same size and density. The level of coupling conservation upon time reversal is also similar to that of the asymmetric random matrices, suggesting that the approximate prediction of $\overleftarrow{d}_{\text{Sym}}(\mathbb{A})$ by $\overleftarrow{d}_{\text{Sym}}(\mathbb{A})$ works relatively well. To be more precise, both the climate and the brain matrix show slightly more asymmetry and less conservation under time reversal than random directed matrices of the same size and density, being safely above the 95-th percentile of the respective random matrix property distributions.

In contrast, in terms of normality, both the climate and brain coupling matrices show a pronounced deviation from normality, in particular substantially higher than directed random matrices of the same size and density. Notably, the deviation of the coupling matrix from perfect reversal under time reversal is also higher than for corresponding random matrices; as would be predicted by the high deviation of $\mathbb{A}$ from normality. These real-world complex system coupling matrices thus seem to substantially deviate from the intuition observed in two-dimensional linear systems, where the coupling tends to be approximately reversed upon time reversal; at least much more so that the realizations of randomly coupled linear networks do.

4. Discussion

The previous analysis has demonstrated, that even in very simple linear systems, the operation of time reversal has nontrivial effect on the coupling structure of the system, which only under very specific conditions corresponds to the reversal of the coupling structure, i.e., transposition of the matrix $\mathbb{A}$ defining the interactions in the case of the studied case of linear vector autoregressive process. For this particular case, we showed an explicit condition on the coupling matrix that needs to be fulfilled for the reversal to occur. Notably, this condition is far from fulfilled for a randomly coupled process, making perfect reversal an unlikely phenomenon; however we show that the deviation of the coupling matrix from normality, quantified by an intuitive index that we devised, seems to very well predict the level to which the reversal occurs. On the other hand, this index quantifying the deviation of the coupling matrix from normality is motivated by analytical derivation that showed normality of the coupling matrix as the key property for causal reversal for autoregressive processes of order one. For higher model orders, the condition for perfect reversal entails existence of solutions for a set of matrix polynomial Equation (4) that generally do not lend themselves to such an easy simplification as for order one that provided the normality condition; in particular, normality of all matrices $\mathbb{A}_1, \dots, \mathbb{A}_p$ is not sufficient for perfect causal reversal in the VAR(p) case.

Examples of first order multivariate linear autoregressive approximation of two real-world complex systems support the theoretical prediction of ubiquity of *imperfect reversal* of the coupling even for linear systems, and suggest that some if not many real-world systems, even when linear or linearly approximated, might be less prone to coupling reversal under time reversal than random coupled systems of the same size and density.

Of course, while perfect causal reversal seems reasonably predicted to be rare (as the coupling structure would have to fulfill very specific set of equations for it to occur), the

level of (lack) causal reversal may vary widely depending on the structure of the connectivity. The analysis of example VAR(1) processes, including with random connectivity, real-world systems approximations as well as carefully constructed or simulation-derived counterexamples, only scratch the surface of the richness of real-world processes here; we only used examples from the VAR(1) family here as the theoretical analysis provides the convenient simplification to the normality condition. As noted previously, the VAR(1) approximation of the brain or climate data does in fact not correspond to the optimal model fit—the Bayesian Information Criterion would suggest to use a higher order model to described more accurately the full interaction structure of the data.

Indeed, the VAR(1) approximation for fMRI data, albeit commonly used in practice, has been shown to be typically not optimal (although the correct model order widely varies depending on a range of properties of the fMRI datasets) [25]. Similarly, the climate temperature anomaly data are known to be only approximately linear and stationary (see [26]), and to contain nontrivial autocorrelation structure [27]. The fact that the VAR(1) is just an approximation—rather than the optimal model fit—for these datasets, however does not disprove the observation that even some more realistic rather than purely random connectivity patterns suggest that the deviation from causal reversal/conservation can be reasonably predicted from the deviation from the normality/symmetry. Indeed, qualitatively similar results are however obtained also on data downsampled by factor of 2 for the brain dataset and by factor of 3 for the climate dataset, for which indeed the Bayesian Information Criterion marks the order $p = 1$ as the optimal model order (see Figures 4 and 5 for the variant of the original Figures 2 and 3 using the downsampled data). Indeed, it may be interesting to collate a catalogue of behaviour of different real-world systems in terms of causal reversal; however it would also entail the difficult task of differentiating properties of the original system and of the observation process and data quality and specific sampling biases—an enterprise beyond the scope of the current manuscript.

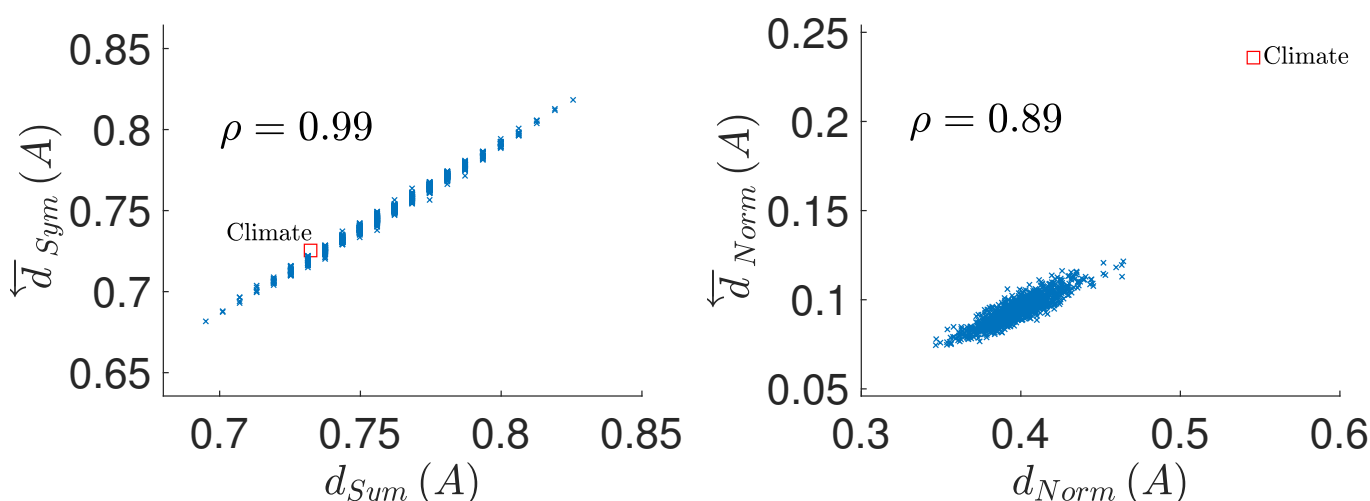


Figure 4. Prediction of the causal structure conservation and reversal from the coupling matrix properties, in a sample of random directed matrices with full diagonal (matrix size $n = 42$, density of nonzero elements $D = 15\%$). **(Left):** causal structure conservation under time reversal $\overleftarrow{d}_{Sym}(A)$ as function of the coupling matrix symmetry $d_{Sym}(A)$. **(Right):** causal structure reversal under time reversal $\overleftarrow{d}_{Norm}(A)$ as function of the coupling matrix normality $d_{Norm}(A)$. ρ denotes the correlation coefficient between the considered quantities. The red square denotes the values for the VAR(1) model of the climate data downsampled by factor of 3 so that the model order $p = 1$ provides the optimal model order according to the Bayesian Information Criterion.

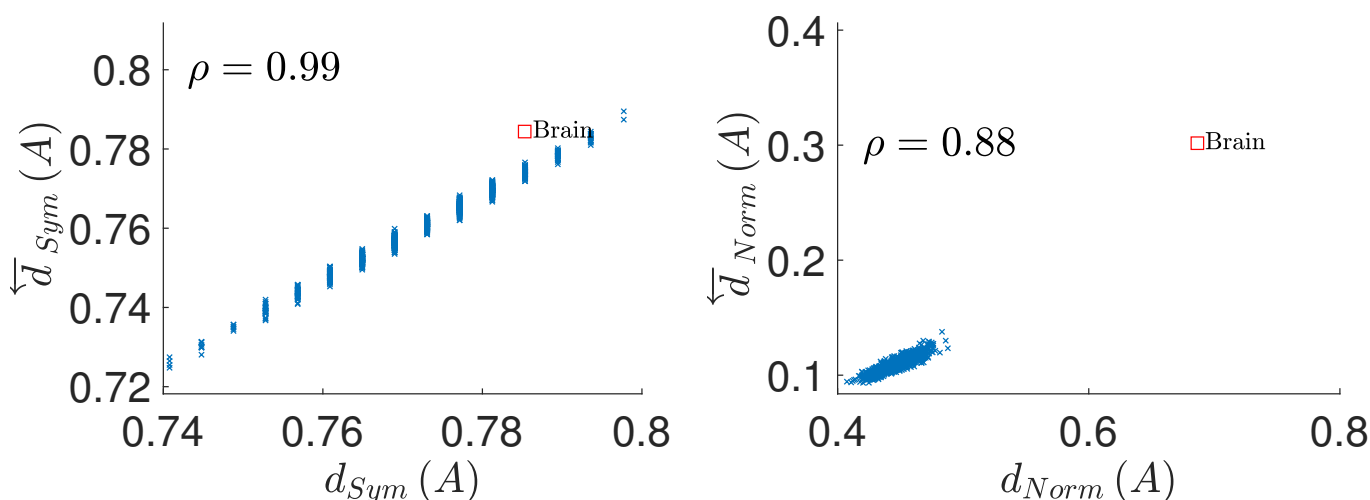


Figure 5. Prediction of the causal structure conservation and reversal from the coupling matrix properties, in a sample of random directed matrices with full diagonal (matrix size $n = 90$, density of nonzero elements $D = 5\%$). **(Left):** causal structure conservation under time reversal $\overleftarrow{d}_{Sym}(A)$ as function of the coupling matrix symmetry $d_{Sym}(A)$. **(Right):** causal structure reversal under time reversal $\overleftarrow{d}_{Norm}(A)$ as function of the coupling matrix normality $d_{Norm}(A)$. ρ denotes the correlation coefficient between the considered quantities. The red square denotes the values for the VAR(1) model of the brain data downsampled by factor of 2 so that the model order $p = 1$ provides the optimal model order according to the Bayesian Information Criterion.

These results may seem counter-intuitive in light of some earlier reports concerning the coupling in time-reversed systems. The discussion of behaviour of causality measures on time-reversed observations has generally followed two different albeit interdependent motivations. First, it was (more or less explicitly) used to provide a control condition or construct a statistical test for causality detection [5,7,11]. Secondly, it is a subject of deeper discussion of the nature of causality and the meaning of the arrow of time in both linear and nonlinear natural systems [6].

In general, while the (approximate) reversal of causal interactions under time reversal has motivated both formulation of new causality estimation methods as well as theoretical discussion concerning fundamental properties of causality in linear vs. nonlinear systems, it is increasingly recognized that the issue is far from binary reversal/non-reversal. For application in real-world problems, the generalizability of the previous reports to networks of larger size would be of key importance. We hope that a more detailed analysis of the simplest VAR(1) approximation of such networks, albeit far from the real-world complexity, would thus help make next steps to extending the applicability of the proposed concepts to real-world contexts.

4.1. Relevance for Causality Indices Using Time-Reversal

The use of time-reversed time series for statistical evaluation of causal interactions in observed systems has been motivated by some perceived limitations of other available methods for causal inference. While Haaga et al. [7] motivates their interest in time series reversal by arguing that other methods including Granger causality, transfer entropy and methods such as convergent cross-mapping can suffer from biases and low statistical power, others such as Winkler et al. [5] and Haufe et al. [11] note that the application of Granger-causal measures combined with standard significance testing leads to the detection of spurious connectivity when applied to electrophysiology data that commonly suffer from source mixing and observational noise. Thus, while motivations vary, the use of time series reversal seems potentially relevant both in combination with nonlinear indices as well as with the widely used Granger causality analysis approach framed in the context of linear autoregressive processes [11].

Let us remind the reader that Granger causality analysis, named after Sir Clive Granger, who proposed this approach to time series analysis in a classical paper [4] follows the basic idea can be traced back to Wiener [28], stating that if the prediction of one time series can be improved by incorporating the knowledge of a second time series, then the latter can be said to have a causal influence on the former. This idea was formalised by Granger in the context of linear regression models. In the following, we outline the methods of assessment of Granger causality, following the description given in [29–31]. Consider two stochastic processes X_t and Y_t and assume they are jointly stationary. Let further the autoregressive representations of each process be:

$$X_t = \sum_{j=1}^{+\infty} a_{1j} X_{t-j} + \epsilon_{1t}, \quad \text{var}(\epsilon_{1t}) = \Sigma_1, \quad (34)$$

$$Y_t = \sum_{j=1}^{+\infty} d_{1j} Y_{t-j} + \eta_{1t}, \quad \text{var}(\eta_{1t}) = \Gamma_1, \quad (35)$$

and the joint autoregressive representation be:

$$X_t = \sum_{j=1}^{+\infty} a_{2j} X_{t-j} + \sum_{j=1}^{+\infty} b_{2j} Y_{t-j} + \epsilon_{2t}, \quad (36)$$

$$Y_t = \sum_{j=1}^{+\infty} c_{2j} X_{t-j} + \sum_{j=1}^{+\infty} d_{2j} Y_{t-j} + \eta_{2t}, \quad (37)$$

where the covariance matrix of the noise terms is:

$$\Sigma = \text{Cov} \begin{pmatrix} \epsilon_{2t} \\ \eta_{2t} \end{pmatrix} = \begin{pmatrix} \Sigma_2 & Y_2 \\ Y_2 & \Gamma_2 \end{pmatrix}. \quad (38)$$

The causal influence from Y to X is then quantified based on the decrease in the residual model variance when we include the past of Y in the model of X , i.e., when we move from the independent model given by Equation (34) to the joint model given by Equation (36):

$$F_{Y \rightarrow X} = \ln \frac{\Sigma_1}{\Sigma_2}. \quad (39)$$

Similarly, the causal influence from X to Y is defined as:

$$F_{X \rightarrow Y} = \ln \frac{\Gamma_1}{\Gamma_2}. \quad (40)$$

Clearly, the causal influence defined in this way is always nonnegative. When the system consists of more than two processes, one can use the above described pairwise Granger Causality Analysis to assess pairwise interactions in each direction. However, this approach suffers from inherent limitations. To give an example, a system consisting of three processes X, Y, Z , where Z drives both X and Y , but with different temporal lags, may erroneously show causal influence between X and Y even if these were not directly coupled. To deal with such situations, one can work with so-called conditional Granger Causality, that allows taking into account the variance explained by third variable(s). Please note that these definitions can be generalised to multivariate variables, as done e.g., in [32].

A key practical question is what can one infer from the value of the estimated Granger causality index, in particular when can one reject the hypothesis that there is no causal interaction between the two subsystems. In general, one can depend on some (asymptotically) known distribution of the test statistic (causal index), or construct an empirical distribution that would represent well the null hypothesis of no causal interaction; and compare the observed value against this empirical distribution. Additionally, as the value

of the finite-sample estimate of the causality index might be affected by a bias depending on specific sample properties (e.g., length, autocorrelation or nonlinearity), various sophisticated heuristics have been proposed that would correct for this bias by subtracting the expected *baseline* value of the statistic.

In the context of Granger causality, Haufe et al. [11] suggested to contrast the Granger causality index $F_{X \rightarrow Y}$ with the one obtained from time-reversed time series $\tilde{F}_{\tilde{X} \rightarrow \tilde{Y}}$ and showed that this procedure, named time-reversed Granger causality (TRGC) robustly rejects causal interpretations on mixtures of independent signals. In 2016, this approach has been extended by Winkler et al. [5], who suggested three measures of causality based on the time reversal effect on the Granger causality index $F_{X \rightarrow Y}$. Let us define the net Granger causality score as $F_{X \rightarrow Y}^{(\text{net})} = F_{X \rightarrow Y} - F_{Y \rightarrow X}$. One can then define the time-reversed net Granger causality score as

$$\tilde{F}_{\tilde{X} \rightarrow \tilde{Y}}^{(\text{net})} = \tilde{F}_{\tilde{X} \rightarrow \tilde{Y}} - \tilde{F}_{\tilde{Y} \rightarrow \tilde{X}}, \quad (41)$$

the difference-based time-reversed Granger causality score as

$$\tilde{D}_{X \rightarrow Y} = F_{X \rightarrow Y} - \tilde{F}_{\tilde{X} \rightarrow \tilde{Y}}, \quad (42)$$

and the net difference-based time-reversed Granger causality score as

$$\tilde{D}_{X \rightarrow Y}^{(\text{net})} = F_{X \rightarrow Y}^{(\text{net})} - \tilde{F}_{\tilde{X} \rightarrow \tilde{Y}}^{(\text{net})}. \quad (43)$$

Please note that in the final score $\tilde{D}_{X \rightarrow Y}^{(\text{net})}$, Winkler et al. [5] essentially combined two useful tricks for controlling for bias of estimates of causality—they contrast to the other direction of interaction (swapping source and target variable) *and* to the time-reversed quantity. These heuristics have been previously used separately in various contexts [33,34], also in conjunction with nonlinear data analysis approaches; see for instance the work of Haaga et al. [7] for the time-reversal-based causal inference using information-theoretical quantities, although their use of the time-reversed surrogate differs from the current construction.

As a key result, Winkler proposed and proved non-negativity of $\tilde{D}_{X \rightarrow Y}$ and $\tilde{D}_{X \rightarrow Y}^{(\text{net})}$ in bivariate system in which causality flows from variable X to Y but not in the other direction i.e., system represented by triangular matrix of size 2×2 , i.e.,

$$\tilde{D}_{X \rightarrow Y} \geq 0 \quad (44)$$

$$\tilde{D}_{X \rightarrow Y}^{(\text{net})} \geq 0. \quad (45)$$

The authors have thus laid theoretical foundations to a more complex causal score, and in a range of numerical simulations showed its advantageous properties, particular in data confounded by observational (additive) noise, including scenarios where standard GC and other methods clearly fail, such as the presence of mixed autocorrelated noise. The (limitations) of applicability in nonlinear dynamical systems with a hidden influential variable [8] or with non-gaussian noise [9] were further studied.

Importantly, Winkler et al. [5] also state some open questions concerning the use of the net difference-based time-reversal Granger causality. In particular, its behaviour under bidirectional information flow and its extension to multivariate signals are considered to be important fields for further research. As we discuss below, a straightforward generalization of the appealing properties of coefficients $\tilde{D}_{X \rightarrow Y}$ and $\tilde{D}_{X \rightarrow Y}^{(\text{net})}$ in bidirectionally coupled or multivariate systems does unfortunately not hold in full generality; however we comment on conditions when it is applicable in its original form.

We envisage that in principle the knowledge of these conditions could either be used directly, in some form of pre-screening of these conditions as assumptions of the time-reversed Granger causality validity; or perhaps more efficiently, it may help develop a

variant of the original time-reversed Granger causality approach that would be valid also for multivariate and bidirectionally coupled systems, providing thus a more generally applicable causal inference method.

4.2. Case of Normally Coupled Unidirectionally Systems

Importantly, even in the multivariate scenario, for unidirectionally coupled systems, we can state some positive results using the earlier findings concerning systems with a *normal* coupling matrix A . In particular, we can prove the same inequalities (45) for a multivariate system represented by VAR(1) process whose coupling matrix is normal and in which information flows from variable X to Y but not in the other direction i.e., $F_{X \rightarrow Y} > 0$ but $F_{Y \rightarrow X} = 0$. As the coupling matrix is normal, the coupling in the time-reversed systems is perfectly reversed, that is $\mathbb{B} = \mathbb{A}^\top$, and therefore $F_{X \rightarrow Y} = \tilde{F}_{\tilde{Y} \rightarrow \tilde{X}}$. Then we have for the difference-based time-reversed Granger causality score in the causal direction:

$$\tilde{D}_{X \rightarrow Y} = F_{X \rightarrow Y} - \tilde{F}_{\tilde{X} \rightarrow \tilde{Y}} = F_{X \rightarrow Y} - F_{Y \rightarrow X} = F_{X \rightarrow Y} > 0. \quad (46)$$

In addition, similarly:

$$\tilde{D}_{X \rightarrow Y}^{(\text{net})} = F_{X \rightarrow Y}^{(\text{net})} - \tilde{F}_{\tilde{X} \rightarrow \tilde{Y}}^{(\text{net})} = F_{X \rightarrow Y} - F_{Y \rightarrow X} - (\tilde{F}_{\tilde{X} \rightarrow \tilde{Y}} - \tilde{F}_{\tilde{Y} \rightarrow \tilde{X}}) \quad (47)$$

$$= F_{X \rightarrow Y} - F_{Y \rightarrow X} - (F_{Y \rightarrow X} - F_{X \rightarrow Y}) = 2F_{X \rightarrow Y} > 0. \quad (48)$$

Based on previous numerical simulations we can conjecture that the coefficient $\tilde{D}_{X \rightarrow Y}$ and $\tilde{D}_{X \rightarrow Y}^{(\text{net})}$ are also non-negative for systems whose representing matrix is “not far from normal” i.e., whose deviation from normality is close to zero. This would provide a promise that these causality scores introduced by Winkler et al. [5] could be reasonably applied for inference of causal direction in such family of multivariate systems.

4.3. Counterexample for Unidirectionally Coupled Systems

However, one can also construct examples, where these convenient properties fail, motivating more detailed study and development of related inference techniques. Already for a trivariate system represented by the VAR(1) process with a coupling matrix $\mathbb{A}$ and corresponding coupling matrix of reversed process $\mathbb{B}$

$$\mathbb{A} = \begin{pmatrix} 0.10 & 0.10 & 0.90 \\ 0.00 & 0.70 & 0.90 \\ 0.00 & 0.00 & 0.90 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 0.20 & 0.35 & -0.00 \\ -0.11 & 1.13 & -0.37 \\ 0.33 & 0.08 & 0.37 \end{pmatrix}. \quad (49)$$

with a corresponding Granger causality matrices

$$F = \begin{pmatrix} 0.01 & 0.07 & 0.79 \\ 0.00 & 1.49 & 0.79 \\ 0.00 & 0.00 & 0.79 \end{pmatrix}, \quad \tilde{F} = \begin{pmatrix} 0.04 & 0.45 & 0.00 \\ 0.01 & 1.64 & 0.09 \\ 0.32 & 0.10 & 0.40 \end{pmatrix} \quad (50)$$

we obtain (using numerical simulations) matrices time-reversal Granger causality scores:

$$\tilde{D} = \begin{pmatrix} -0.02 & -0.38 & 0.79 \\ -0.00 & -0.15 & 0.70 \\ -0.32 & -0.10 & 0.39 \end{pmatrix}, \quad \tilde{D}^{(\text{net})} = \begin{pmatrix} 0.00 & -0.37 & 1.11 \\ 0.37 & 0.00 & 0.80 \\ -1.23 & -0.80 & 0.00 \end{pmatrix}. \quad (51)$$

As can be seen, there is a unidirectional causal effect from X_2 to X_1 (i.e., $F_{1,2} > 0$, but $F_{2,1} = 0$), but the time-reversed process has ‘stronger’ causality in both directions, and thus the values of the difference-based time-reversed Granger causality scores between X_1 and X_2 are negative: $\tilde{D}_{X_2 \rightarrow X_1} < 0$. Moreover, even $\tilde{D}_{X_2 \rightarrow X_1}^{(\text{net})}$ is negative; and thus inference depending solely on positivity of the scores would render wrong conclusion in this particular case.

4.4. Counterexample for Bidirectionally Coupled Systems

While for the unidirectionally coupled systems, Winkler et al. [5] proved analytically positivity of the proposed indices for the correct direction, the authors highlighted the interesting question of the behaviour of the time-reverted causality in the case of bidirectional causality as an open topic for further research. While we showed that in the general case the coupling matrix is not exactly reverted, for the application of the proposed causality scores it would be sufficient, if at least the dominant direction of causality was reverted—a behaviour that Paluš et al. [6] observed in example linear systems, in stark contrast to selected archetypal nonlinear systems in which causality indices typically conserved the dominant direction upon time reversal.

While we did not have a full analytical solution for the time-reversed coupling matrix and the resulting causality scores, we used here (similarly as in the three-dimensional unidirectionally coupled system) a systematic parameter search to test if any counterexamples can be numerically found. An example of such counterexample is shown below, already in dimension $n = 2$. Please note that the dominant coupling in both forward and backward time-series is from X_2 to X_1 , which is reflected both in the coupling matrices $\mathbb{A}$ and $\mathbb{B}$ as well as in the Granger causality matrices F and $\tilde{F}$, although the (net) difference-based time-reversed Granger causality matrices $\tilde{D}$, $\tilde{D}^{(\text{net})}$ still detect the correct direction of causality (and thus in this respect, the lowest dimension of counterexample stays at three):

$$\mathbb{A} = \begin{pmatrix} 0.10 & 0.70 \\ 0.50 & 0.60 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 0.17 & 0.64 \\ 0.59 & 0.53 \end{pmatrix} \quad (52)$$

$$F = \begin{pmatrix} 0.02 & 0.86 \\ 0.36 & 0.70 \end{pmatrix}, \quad \tilde{F} = \begin{pmatrix} 0.04 & 0.67 \\ 0.55 & 0.66 \end{pmatrix} \quad (53)$$

$$\tilde{D} = \begin{pmatrix} -0.02 & 0.19 \\ -0.19 & 0.03 \end{pmatrix}, \quad \tilde{D}^{(\text{net})} = \begin{pmatrix} 0.00 & 0.38 \\ -0.38 & 0.00 \end{pmatrix}. \quad (54)$$

In summary, while we have shown that a straightforward generalization of the useful concept of difference-based Granger causalities for causal inference in bidirectionally coupled or higher-dimensional systems is not available, the problem offers an interesting and rich set of open questions that may stimulate further development of this idea.

5. Conclusions

On top of the general results in the theoretical analysis section, the above discussed special cases demonstrated, that the situation concerning the coupling structure of the time-reversed systems is far from trivial already in the relatively simple case of linear autoregressive processes with white noise and a low dimension ($n = 2$ or $n = 3$). However, for some special cases (such as normality of the coupling matrix in the case of linear vector autoregressive model of order one), perfect causality reversal can be analytically shown, and for a broader family of high-dimensional random or complex structures, analytical tools seem to provide relevant clues concerning the time-reversed process coupling structure. We further illustrated that the coupling reversal or conservation does not follow strictly the dividing line between linear and nonlinear systems generally suggested by Paluš et al. [6], but deviate from it already within the realm of relatively simple linear systems; extending thus the problematic cases of linear systems with multiple delays and nonlinear systems in the synchronized regimes mentioned as possible exceptions to such rule of thumb in the original paper.

Notably, we used a theoretical framework that allows analytical derivation of conditions for causality reversal or conservation in linear systems. While this allowed us to derive the sufficient conditions for these phenomena in the simplest case of models with single time lag, further work may focus in particular on deriving necessary conditions, and

generalizing the family of processes for which they are observed; although this may prove increasingly technical given the nature of the conditions that need to be considered.

Author Contributions: Conceptualization, J.H.; methodology, J.H. and J.K.; software, J.K.; validation, J.H. and J.K.; formal analysis, J.H. and J.K.; investigation, J.H. and J.K.; resources, J.H.; data curation, J.H. and J.K.; writing—original draft preparation, J.H. and J.K.; writing—review and editing, J.H. and J.K.; visualization, J.K.; supervision, J.H.; funding acquisition, J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Czech Science Foundation projects No. 19-11753S, No. 19-16066S, No. 21-32608S, No. 21-17211S, by Ministry of Health Czech Republic—DRO 2021 (“National Institute of Mental Health—NIMH, IN: 00023752”) and by Grant Agency of the Czech Technical University in Prague, grant No. SGS20/183/OHK4/3T/14.

Acknowledgments: We thank Nikola Jajcay, David Hartman, and David Tomecek for valuable help with example data preparation.

Data Availability Statement: Example data presented in this paper are available upon request by email to the authors.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Bivariate Case

In this section, we will derive the explicit form of reverse process of two special cases of VAR(1) process with white noise and coupling matrix $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$.

- $d = 0$

$$\mathbb{A} = -\mathbb{A}_1 = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \quad (\text{A1})$$

The VAR(1) process with coupling matrix $\mathbb{A}$ is stationary if the equation $\det(\mathbb{A}_1 + \lambda \mathbb{I}) = 0$ has all the roots smaller than 1 in absolute value. Since

$$\det(\mathbb{A}_1 + \lambda \mathbb{I}) = \det \begin{pmatrix} -a + \lambda & -b \\ 0 & \lambda \end{pmatrix} = \lambda(\lambda - a), \quad (\text{A2})$$

the process is stationary for $|a| < 1$.

We will show that reverse process is given by coupling matrix

$$\mathbb{B} = -\mathbb{B}_0^{-1}\mathbb{B}_1 = \begin{pmatrix} a & 0 \\ \frac{b(1-a^2)}{1+b^2} & 0 \end{pmatrix} \quad (\text{A3})$$

and covariance matrix of the noise

$$\mathbb{B}_0^{-1}(\mathbb{B}_0^{-1})^\top = \begin{pmatrix} 1+b^2 & -ab \\ -ab & \frac{1+a^2b^2}{1+b^2} \end{pmatrix}. \quad (\text{A4})$$

Since $\det(\mathbb{B}_0^{-1}(\mathbb{B}_0^{-1})^\top) = 1$, hence $\det(\mathbb{B}_0^{-1}) \neq 0$ and also $\det(\mathbb{B}_0) \neq 0$. Process is stationary because

$$\det(\mathbb{B}_1 + \lambda \mathbb{B}_0) = \det \left(-\mathbb{B}_0 \begin{pmatrix} a & 0 \\ \frac{b(1-a^2)}{1+b^2} & 0 \end{pmatrix} + \lambda \mathbb{B}_0 \right) = \det(\mathbb{B}_0) \det \begin{pmatrix} \lambda - a & 0 \\ -\frac{b(1-a^2)}{1+b^2} & \lambda \end{pmatrix} = 0 \quad (\text{A5})$$

if $\lambda = 0 \vee \lambda = a$ and $|a| < 1$ from the stationarity of the original process. The last step is to prove the fulfillment of the Equations (5) and (6). From (A4) follows

$$\mathbb{B}_0^\top \mathbb{B}_0 = \begin{pmatrix} \frac{1+a^2b^2}{1+b^2} & ab \\ ab & 1+b^2 \end{pmatrix}, \quad (\text{A6})$$

the Equation (5) is then satisfied because

$$-\mathbb{A}_1^\top = \mathbb{B}_0^\top \mathbb{B}_1 = -\mathbb{B}_0^\top \mathbb{B}_0 \mathbb{B}_0^{-1} \mathbb{B}_1 = \begin{pmatrix} \frac{1+a^2b^2}{1+b^2} & ab \\ ab & 1+b^2 \end{pmatrix} \begin{pmatrix} a & 0 \\ \frac{b(1-a^2)}{1+b^2} & 0 \end{pmatrix} = \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} \quad (\text{A7})$$

and (6) is satisfied because

$$\mathbb{B}_1^\top \mathbb{B}_1 = \mathbb{B}_1^\top \mathbb{B}_0 \mathbb{B}_0^{-1} \mathbb{B}_1 = \mathbb{A}_1 \mathbb{B}_0^{-1} \mathbb{B}_1 \quad (\text{A8})$$

and

$$\begin{pmatrix} 1+a^2 & ab \\ ab & 1+b^2 \end{pmatrix} = \mathbb{I} + \mathbb{A}_1^\top \mathbb{A}_1 = \mathbb{B}_0^\top \mathbb{B}_0 + \mathbb{B}_1^\top \mathbb{B}_1 = \mathbb{B}_0^\top \mathbb{B}_0 + \mathbb{A}_1 \mathbb{B}_0^{-1} \mathbb{B}_1 = \quad (\text{A9})$$

$$\begin{pmatrix} \frac{1+a^2b^2}{1+b^2} & ab \\ ab & 1+b^2 \end{pmatrix} + \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & 0 \\ \frac{b(1-a^2)}{1+b^2} & 0 \end{pmatrix} = \begin{pmatrix} 1+a^2 & ab \\ ab & 1+b^2 \end{pmatrix}. \quad (\text{A10})$$

- $a = 0$

$$\mathbb{A} = -\mathbb{A}_1^\top = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} \quad (\text{A11})$$

The process is stationary if the equation $\det(\mathbb{A}_1 + \lambda \mathbb{I}) = 0$ has all the roots smaller than 1 in absolute value. Since

$$\det(\mathbb{A}_1 + \lambda \mathbb{I}) = \det \begin{pmatrix} \lambda & -b \\ 0 & -d + \lambda \end{pmatrix} = \lambda(\lambda - d), \quad (\text{A12})$$

the process is stationary for $|d| < 1$. We will show that reverse process is given by coupling matrix

$$\mathbb{B} = -\mathbb{B}_0^{-1} \mathbb{B}_1 = \frac{1}{1+b^2} \begin{pmatrix} b^2d & bd^2 \\ b & d \end{pmatrix} \quad (\text{A13})$$

and covariance matrix of the noise

$$\mathbb{B}_0^{-1} (\mathbb{B}_0^{-1})^\top = \frac{1}{1+b^2} \begin{pmatrix} (1+b^2)^2 + b^2d^2 & bd \\ bd & 1 \end{pmatrix}. \quad (\text{A14})$$

Similarly to previous case, since $\det(\mathbb{B}_0^{-1} (\mathbb{B}_0^{-1})^\top) = 1$, hence $\det(\mathbb{B}_0^{-1}) \neq 0$, $\det(\mathbb{B}_0) \neq 0$ and process is stationary because

$$\det(\mathbb{B}_1 + \lambda \mathbb{B}_0) = \det(\mathbb{B}_0) \det \begin{pmatrix} \lambda - \frac{b^2d}{1+b^2} & \frac{-bd^2}{1+b^2} \\ \frac{-b}{1+b^2} & \lambda - \frac{d}{1+b^2} \end{pmatrix} = \det(\mathbb{B}_0)(\lambda - d)\lambda = 0 \quad (\text{A15})$$

if $\lambda = 0 \vee \lambda = d$ and $|d| < 1$ from the stationarity of the original process. The last step is to prove the fulfillment of the Equations (5) and (6). From (A14) follows

$$\mathbb{B}_0^\top \mathbb{B}_0 = \frac{1}{1+b^2} \begin{pmatrix} 1 & -bd \\ -bd & (1+b^2)^2 + b^2d^2 \end{pmatrix}, \quad (\text{A16})$$

the Equation (5) is then satisfied because

$$-\mathbb{A}_1^\top = \frac{1}{(1+b^2)^2} \begin{pmatrix} 1 & -bd \\ -bd & (1+b^2)^2 + b^2d^2 \end{pmatrix} \begin{pmatrix} b^2d & bd^2 \\ b & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ b & d \end{pmatrix} \quad (\text{A17})$$

and (6) is satisfied because

$$\begin{pmatrix} 1 & 0 \\ 0 & 1+b^2+d^2 \end{pmatrix} = \mathbb{I} + \mathbb{A}_1^\top \mathbb{A}_1 = \mathbb{B}_0^\top \mathbb{B}_0 + \mathbb{B}_1^\top \mathbb{B}_1 = \mathbb{B}_0^\top \mathbb{B}_0 + \mathbb{A}_1 \mathbb{B}_0^{-1} \mathbb{B}_1 = \\ \frac{1}{1+b^2} \begin{pmatrix} 1 & -bd \\ -bd & (1+b^2)^2 + b^2d^2 \end{pmatrix} + \frac{1}{(1+b^2)^2} \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} \begin{pmatrix} b^2d & bd^2 \\ b & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1+b^2+d^2 \end{pmatrix}. \quad (\text{A18})$$

References

- Parrondo, J.M.; Van Den Broeck, C.; Kawai, R. Entropy production and the arrow of time. *New J. Phys.* **2009**, *11*, 073008.
- Paluš, M. Biological Cybernetics Nonlinearity in normal human EEG: Cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biol. Cybern.* **1996**, *75*, 389–396.
- Pearl, J. *Causality: Models, Reasoning and Inference*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2009.
- Granger, C.W. Investigating causal relations by econometric model and cross spectral methods. *Econometrica* **1969**, *37*, 424–438.
- Winkler, I.; Panknin, D.; Bartz, D.; Müller, K.R.; Haufe, S. Validity of Time Reversal for Testing Granger Causality. *IEEE Trans. Signal Process.* **2016**, *64*, 2746–2760.
- Paluš, M.; Krakovská, A.; Jakubík, J.; Chvosteková, M. Causality, dynamical systems and the arrow of time. *Chaos Interdiscip. J. Nonlinear Sci.* **2018**, *28*, 075307.
- Haaga, K.A.; Diego, D.; Brendryen, J.; Hannisdal, B. A simple test for causality in complex systems. *arXiv* **2020**, arXiv:stat.AP/2005.01860.
- Chvosteková, M. Granger Causality Inference and Time Reversal. In Proceedings of the 2019 12th International Conference on Measurement, Smolenice, Slovakia, 27–29 May 2019; pp. 110–113.
- Chvosteková, M.; Jakubík, J.; Krakovská, A. Granger Causality on forward and Reversed Time Series. *Entropy* **2021**, *23*, 409.
- Kathpalia, A.; Nagaraj, N. Time-Reversibility, Causality and Compression-Complexity. *Entropy* **2021**, *23*, 327.
- Haufe, S.; Nikulin, V.V.; Müller, K.R.; Nolte, G. A critical assessment of connectivity measures for EEG data: A simulation study. *NeuroImage* **2013**, *64*, 120–133.
- Anděl, J. Symmetric and Reversed Multiple Stationary Autoregressive Series. *Ann. Math. Statist.* **1972**, *43*, 1197–1203.
- Palus, M.; Vejmelka, M. Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections. *Phys. Rev. E* **2007**, *75*, 056211. doi:10.1103/PhysRevE.75.056211.
- Kistler, R.; Kalnay, E.; Collins, W.; Saha, S.; White, G.; Woollen, J.; Chelliah, M.; Ebisuzaki, W.; Kanamitsu, M.; Kousky, V.; et al. The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Am. Meteorol. Soc.* **2001**, *82*, 247–267.
- Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 437–471.
- Heikes, R.; Randall, D.A. Numerical integration of the shallow-water equations on a twisted icosahedral grid. Part I: Basic design and results of tests. *Mon. Weather Rev.* **1995**, *123*, 1862–1880.
- Jones, P.W. *A User's Guide for SCRIP: A Spherical Coordinate Remapping and Interpolation Package*; Los Alamos National Laboratory: New Mexico, NM, USA, 1997; Volume 1.
- Hlinka, J.; Hartman, D.; Jajcay, N.; Tomeček, D.; Tintêra, J.; Paluš, M. Small-world bias of correlation networks: From brain to climate. *Chaos* **2017**, *27*, 035812.
- Kořenek, J.; Hlinka, J. Causal network discovery by iterative conditioning: Comparison of algorithms. *Chaos* **2020**, *30*, 013117.
- Hlinka, J.; Palus, M.; Vejmelka, M.; Mantini, D.; Corbetta, M. Functional connectivity in resting-state fMRI: Is linear correlation sufficient? *NeuroImage* **2011**, *54*, 2218–2225.
- Hartman, D.; Hlinka, J.; Palus, M.; Mantini, D.; Corbetta, M. The role of nonlinearity in computing graph-theoretical properties of resting-state functional magnetic resonance imaging brain networks. *Chaos* **2011**, *21*, 013119.
- Hlinka, J.; Hartman, D.; Vejmelka, M.; Runge, J.; Marwan, N.; Kurths, J.; Paluš, M. Reliability of Inference of Directed Climate Networks Using Conditional Mutual Information. *Entropy* **2013**, *15*, 2023–2045.

23. Hlinka, J.; Jajcay, N.; Hartman, D.; Paluš, M. Smooth information flow in temperature climate network reflects mass transport. *Chaos* **2017**, *27*, 035811.
24. Runge, J.; Petoukhov, V.; Donges, J.F.; Hlinka, J.; Jajcay, N.; Vejmelka, M.; Hartman, D.; Marwan, N.; Palus, M.; Kurths, J. Identifying causal gateways and mediators in complex spatio-temporal systems. *Nat. Commun.* **2015**, *6*, 8502–8502.
25. Ting, C.M.; Seghouane, A.K.; Khalid, M.U.; Salleh, S.H. Is First-Order Vector Autoregressive Model Optimal for fMRI Data? *Neural Comput.* **2015**, *27*, 1857–1871. doi:10.1162/NECO_a_00765.
26. Hlinka, J.; Hartman, D.; Vejmelka, M.; Novotná, D.; Paluš, M. Non-linear dependence and teleconnections in climate data: sources, relevance, nonstationarity. *Clim. Dyn.* **2013**, *42*, 1873–1886. doi:10.1007/s00382-013-1780-2.
27. Paluš, M.; Hartman, D.; Hlinka, J.; Vejmelka, M.; Palus, M.; Hartman, D.; Hlinka, J.; Vejmelka, M.; Paluš, M.; Hartman, D.; et al. Discerning connectivity from dynamics in climate networks. *Nonlinear Processes. Geophys.* **2011**, *18*, 751–763. doi:10.5194/npg-18-751-2011.
28. Wiener, N. The theory of prediction. In *Modern Mathematics for Engineers*; McGraw-Hill: New York, NY, USA, 1956; pp. 165–190.
29. Ding, M.; Chen, Y.; Bressler, S.L. Granger Causality: Basic Theory and Application to Neuroscience. In *Handbook of Time Series Analysis*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2006; Chapter 17, pp. 437–460.
30. Geweke, J. Measurement of Linear Dependence and Feedback between Multiple Time Series. *J. Am. Stat. Assoc.* **1982**, *77*, 304–313.
31. Geweke, J.F. Measures of Conditional Linear Dependence and Feedback between Time Series. *J. Am. Stat. Assoc.* **1984**, *79*, 907–915.
32. Barnett, L.; Barrett, A.B.; Seth, A.K. Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. *Phys. Rev. Lett.* **2009**, *103*, 238701.
33. Nolte, G.; Ziehe, A.; Nikulin, V.V.; Schlögl, A.; Krämer, N.; Brismar, T.; Müller, K.R. Robustly Estimating the Flow Direction of Information in Complex Physical Systems. *Phys. Rev. Lett.* **2008**, *100*, 234101.
34. Nolte, G.; Ziehe, A.; Krämer, N.; Popescu, F.; Müller, K.R. Comparison of Granger Causality and Phase Slope Index. In Proceedings of the 2008 International Conference on Causality: Objectives and Assessment, Vancouver, BC, Canada, 8–10 December 2008; Volume 6, pp. 267–276.