Information Sciences 264 (2014) 279–290

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Causal conditioning and instantaneous coupling in causality graphs

Pierre-Olivier Amblard [a], Olivier Michel [b,*]

[a] Department of Mathematics and Statistics, University of Melbourne, Parkville, VIC 3010, Australia
[b] Gipsa-Lab UMR 5216, 961 rue de la Houille Blanche, BP.46, 38402 Saint Martin d'Hères Cedex, France

ARTICLE INFO

Article history:
Received 7 February 2013
Received in revised form 27 August 2013
Accepted 29 December 2013
Available online 7 January 2014

Keywords:
Directed information
Transfer entropy
Granger causality
Graphical model

ABSTRACT

This study aims at providing the definitive links between Massey and Kramer's directed information theory and Granger causality graphs, recently formalized by Eichler. This naturally leads to consider two concepts of causality that can occur in physical systems: dynamical causality and instantaneous coupling. Although it is well accepted that the former is assessed by conditional transfer entropy, the latter is often overlooked, even if it was clearly introduced and understood in seminal studies. In the bivariate case, we show that directed information is decomposed into the sum of transfer entropy and instantaneous information exchange. In the multivariate case, encountered for conditional graph modeling, such a decomposition does not hold anymore. We provide a new insight into the problem of instantaneous coupling and show how it affects the estimated structure of a graphical model that should provide a sparse representation of a complex system. This result is discussed and analyzed for the inference of causality graphs. Two synthetic examples are developed to illustrate the theoretical concepts. Practical issues are also briefly discussed on these examples and an extension of Leonenko's k-nearest neighbors based entropy estimators is used to derive a nonparametric estimator of directed information.

1. Introduction

1.1. Motivations

Graphical modeling has received major attention in many different domains, such as the neurosciences [18], econometry [8], and complex networks [26]. This proposes a representation paradigm to explain how information flows between the nodes of a graph. The graph vertices are in most cases, and particularly in the present study, associated to synchronous time series. To infer a graph thus requires that the edges or links between the vertices are defined. Granger [16,17] proposed a set of axiomatic definitions for the causality between, say, x and y (with a slight abuse of notation, each vertex will be named after its associated time series). Granger's definitions are based on the improvement that observations of x up to some time $t-1$ can allow the prediction of y at time t. The fundamental idea in Granger's approach is that *the past and the present may cause the future, but the future cannot cause the past* [16, Axiom A]. Granger's work also stresses the importance of side information, which accounts for the presence of all of the other vertices except x and y, in the assessing of the existence of a link between two nodes. This leads to what will be referred to as the bivariate case (an absence of side information) and the

* Corresponding author. Tel.: +33 476826333.
E-mail addresses: bidou.amblard@gipsa-lab.grenoble-inp.fr (P.-O. Amblard), olivier.michel@gipsa-lab.grenoble-inp.fr (O. Michel).

0020-0255/$ - see front matter
http://dx.doi.org/10.1016/j.ins.2013.12.037

multivariate case (a presence of side information) in what follows. Eichler and Dahlaus [7–9] were among the first to propose the use of Granger causality in the framework of graphical modeling.

In [9], precise definitions of Granger causality graphs are presented, and both the concepts of dynamical causality and instantaneous coupling (as we call them herein) are emphasized. Note that the concept of instantaneous coupling was present in the early works on Granger causality, although as this concept appeared quite weak compared to dynamical causality, it has been overlooked in modern studies on causality, and especially in the application of these. Instantaneous dependence in complex networks can have different origins. Here, if it is not easy to conceive of instantaneous information exchange between nodes, the recording process (which includes filters, sample and hold devices, converters) contains integrators over short time lags. Any information that flows between two nodes within a delay shorter than the integration time can then be seen as instantaneous. Such a case is often met in systems that require long integration times per sample, as for example in functional magnetic resonance imaging. Alternatively, instantaneous coupling can occur if the noise contributions in structural models are no longer independent, which is likely in real-world cases.

1.2. Aim of the paper and outline

The purpose of the present study is to provide new insight into the problems related to instantaneous coupling, and to show how the presence of such coupling can affect the estimated structure of a graphical model that should provide a sparse representation of a complex system. This report is focused on the interplay of two types of dependences: dynamical causality and instantaneous coupling, which have been largely overlooked in all other studies of directed information. The possibility to estimate directed-information-based measures with k nearest neighbors (k-NN)-based tools is illustrated. This novel application of k-NN circumvents the requirement for the definition of a parametric model, as has always been introduced so far in the estimation of directed information from continuous valued data.

We begin this report with a short review of the possible approaches to Granger causality. Section 3 then introduces a brief review and some definitions of Eichler and Dalhaus causality graphs [7,9], and presents an enlightening toy problem, where instantaneous coupling strongly affects the edge detection in a graphical model. Theoretical relations that demonstrate the link between directed information theory and Granger causality graphs are developed in the following section. The last section discusses some of the practical implementation issues, and provides the full treatment of a toy problem.

2. Approaches to Granger causality

2.1. Model-based approaches

In Geweke's pioneering work [11,12], an autoregressive modeling approach was adopted (for both the bivariate and multivariate cases) to provide practical implementation of Granger causality graphs. Such a model-based approach motivated further studies. Information theory tools were also added by Rissanen and Wax [30], to account for the complexity of the regression model. Directed transfer functions, which are frequency domain filter models for Granger causality, were derived in [16,17] for neuroscience applications. Nonlinear extensions have been proposed [14], with recent developments relying on functional estimation in a reproducing kernel Hilbert space [23,3]. All of these approaches are intrinsically parametric, and as such, they can introduce bias into the analysis.

2.2. Information-theory-based measures

An alternative to an assessment of the existence of a link between nodes was elaborated earlier in the bivariate case (see, for example, [31,15,32,29,33]). This consists of the adapting of information theory measures, such as mutual information or information divergence, to assess the existence and/or strength of a link between two nodes. The motivations for introducing such tools rely upon the ability of the information theory measure to account for the entire probability density function of the observations (provided that such a probability density exists), instead of only second-order characteristics, as for linear filter modeling approaches. Among these earlier studies, one of the oldest and maybe the least well known was developed by Gouriéroux et al. [15], where a generalization of Geweke's idea [11] using Kullback divergences was introduced, for the case of two time series. It is of note that the tools they introduced were later rediscovered by Massey [25] and Kramer [19] in their development of bivariate directed information theory.

The development of directionality-specific or causality-specific measures was initiated by a study on directed information of Marko [24], and extended by Massey [25], and later Kramer [19], who introduced causal conditioning by side information. This offers a means to account for side information, or to tackle the multivariate case. The first steps in the exploration of the relationships between the Geweke approach of Granger causality and the directed information theory tools were defined in [1] for the Gaussian case. with further insights then developed in [2], and in [27,28] in the absence of instantaneous dependence structure. In [2], a directed-information-based new definition was proposed for Granger causality. Eichler [9] recently studied this latter issue in a graph modeling framework, both from a theoretical point of view, with recourse to probability based definitions, and in a parametric modeling context.