Comenius University, Bratislava
Faculty of Mathematics, Physics and Informatics
Department of Applied Informatics

Quantification of causal interactions in complex systems

Master Thesis

Bc. Anton Vančo

2012

Katedra Aplikovanej Informatiky
Fakulta Matematiky, Fyziky a Informatiky
Univerzita Komenského, Bratislava

Kvantifikácia kauzálnych interakcií v komplexných systémoch

Diplomová práca

Študijný program:	Aplikovaná informatika
Študijný odbor:	2511 Aplikovaná informatika
Školiace pracovisko:	Aplikovaná informatika
Školiteľ:	doc. Ing. Igor Farkaš, PhD.

Bc. Anton Vančo

2012

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta:	Bc. Anton Vančo
Študijný program:	aplikovaná informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor:	9.2.9. aplikovaná informatika
Typ záverečnej práce:	diplomová
Jazyk záverečnej práce:	anglický
Sekundárny jazyk:	slovenský

Názov: Kvantifikácia kauzálnych interakcií v komplexných systémoch

Cieľ: Naštudujte si literatúru z oblasti formálnych neurovedne orientovaných prístupov k meraniu zložitosti a kauzálnych vzťahov v distribuovaných systémoch (napr. mozgu). Zamerajte sa na koncept Grangerovej kazuality a experimentálne analyzujte jej vyhodnocovanie v prípade vybraných simulovaných systémov.

Literatúra:
- Seth, A.K. (2010). Measuring autonomy and emergence via Granger causality Artificial Life. 16(2):179-196.
- ďalšie články na stránke http://www.informatics.sussex.ac.uk/users/anils/aks_newpapers2.htm
- Farkaš I.: Mental causation in a physical brain? In: Brain-Inspired Cognitive Systems, Madrid, 14-16 July, 2010.

Poznámka: Požiadavky: pasívna znalosť angličtiny, záujem o danú problematiku, relatívna samostatnosť, ochota pracovať priebežne.

Vedúci: doc. Ing. Igor Farkaš, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky

Dátum zadania: 04.10.2010

Dátum schválenia: 03.12.2010

doc. RNDr. Roman Ďurikovič, PhD.
garant študijného programu

I declare that I have developed this master thesis individually using only the cited literature.

.....

I would like to express my sincere gratitude to my mentor, doc. Ing. Igor Farkaš, PhD. for his support, helpful insights and his guidance throughout this project. I would also like to thank my family for their everlasting support throughout the years.

Abstract

The notion of causality has been tackled by philosophers for many centuries but only recently, several attempts have appeared to formally define causality in physical systems. In this master thesis we will take a look at the Wiener-Granger method for statistical analysis of causal interactions. We provide a theoretical introduction to this method defining its basic concepts and properties. Afterwards we first use this method to analyze artificially generated time series, and later we apply this method to measure emergent behavior in an artificial system of bird flocking.

keywords: Granger causality, time series, statistical analysis, autonomy, emergence

Abstrakt

Pojem kauzality bol predmetom skúmania mysliteľov po stáročia, ale až v dnešnej dobe boli predstavené postupy pre formálne popísanie kauzality vo fyzických systémoch. V tejto práci si priblížime metódu Wiener-Grangerovej kauzality, pomocou ktorej pristupujeme k skúmaniu kauzálnych vzťahov prostredníctvom štatistickej analýzy a predstavíme jej základné teoretické koncepty a vlastnosti. Ďalej aplikujeme túto metódu pre analýzu simulovaných časových radov a neskôr túto metódu použijeme na kvantifikáciu emergentných vlastností skúmaného systému.

kľúčové slová: Grangerova kauzalita, časové rady, štatistická analýza, autonómia, emergencia

Foreword

The topic of this document is the notion of causality applied in the context of a statistical analysis of interdependent time series carried out by taking advantage of the Wiener-Granger causality method. Using different experimental artificially generated data we test the accuracy of the Wiener-Granger causality and discuss a proposed measure for quantifying causal interactions – causal density, which is derived from this method.

During our testing we will specify a number of different sets of stochastic processes and analyze their behavior by both, the time-domain as well as the frequency-domain versions of the Wiener-Granger causality. Drawing from the ideas behind this method we will also describe and exemplify a complementary measure, termed G-autonomy, which will serve as a tool for assessing the independence of a process from the rest of a system.

We will also discuss the notion of different varieties of emergence and attempt to analyze weakly emergent behavior in the simulated environment of Boids, a textbook example of emergent behavior – flocking.

Contents

1 Introduction

Causality, the notion which is defined as the principle of the cause and its consequential effect, has been an important subject of study for philosophers, physicists and other thinkers throughout the history of civilization (Pearl, 2000). It is viewed as the cornerstone and one of the most basic requirements for *determinism* – the principle, which seems to govern the very laws of the universe which we exist in. Causality, or rather the causal influence, can also be interpreted as a rule by which a behavior of one entity is subject to and dependent on the behavior, state, energy and so on of another entity. If the temperature of the water in a pot reaches its boiling point, water changes its phase from liquid to gaseous, or when a falling ball hits a solid surface, it immediately bounces back.

One of the major aspects of determinism and, by proxy, of causality is that it appears to directly contradict the notion of *free will* (Dennett, 2003). Although this philosophical topic is not the subject of this particular thesis, we will discuss a bit more down-to-earth idea, connected to consciousness, that the causality in complex systems (such as the human brain) can be approached scientifically (Farkaš, 2007) and observed, tested and measured by the means of mathematical formalization and statistical analysis.

In this thesis we will draw upon the method of *Wiener-Granger causality* which has its roots in econometrics and has been first described by Clive Granger (Granger, 1969) in 1969, but which has in recent years become one of the favorite methods for analyzing the recorded neural activity observations by a ever growing number of scientists in the field of neuro- and cognitive sciences. This approach, based on the *time-series inference*, is becoming a promising tool for exploring yet unknown interactions between the functional

parts of the human brain. For the purpose of quantifying these causal interaction a new measure has been proposed – the *causal density* (Seth et al., 2011) metric which is designed to capture the underlying structure of the causal behavior sustained by a system.

The idea behind Granger causality (G-causality) is that if the prediction of one variable becomes more accurate if we consider, together with its previous values, past values of another variable, the second variable contains causal information about the first variable. In other words, the second variable contains influential information about the first one – a variable is said to *Granger cause* another. This idea is further extended for more realistic scenarios where we analyze the behavior of multivariate sets of processes and perform our analysis in the presence of latent (unmeasured) variables and other noisy influences.

Another aspect of the causality in complex systems is the presence of causal influences observed not only between the system's smallest units but also between groups of these units. These groups of variables form standalone regions with a specific causal behavior and thus create a sense of causality at different levels of the system. With this property in mind, another extension for G-causality, which operates at the multivariate level, has been proposed and is discussed in latter sections of this thesis.

Finally, by incorporating the intuitions about the phenomenon of emergence, we introduce the measure of *G-emergence* and elaborate on a rather metaphysical notion of *downward causation*. Emergence is a property which manifests itself as a macro-level property which becomes existent by the virtue of interactions between the system's micro-level elements (Bedau, 2003). These relationships, where micro-properties influence an emergent macro-property and in turn the macro influences the micro, bring about the

notion of quantifiable *inter-level causation* which is complementary to the *inter-level causation* observed between variables at the same level (Farkaš, 2010).

In this thesis we test this method and its extensions by applying them in the analysis of various artificially generated time series and simulated flocks of bird-like agents.

With the advent of the computer era, providing extremely fast and reliable calculations, complemented by the advances in the study of the neurobiological processes in the mammalian brain (human brain included) we are now able, provided that our G-causality analysis is useful, to shed new light on what lies behind the processes inside our thinking brains. Using vast amounts of data of neural activity obtained using functional magnetic resonance imaging (fMRI), electro-/magnetoencephalography (EEG/MEG), multielectrode arrays (MEA) and computational power of today's computers we can possibly broaden our insight and achieve a more detailed understanding of our own neurological processes and perhaps even understand the relationship between the physical brain and the abstract consciousness.

2 Measuring causality

In this section we will focus on the analysis of time series using the method of Wiener-Granger causality – *G-causality*. We will discuss the usage of time-domain and spectral analysis and introduce the *causal density* metric which will serve as a measure for causal interactions. Finally, we describe an extension of this method which finds its use in the scenarios when we analyze processes which are under the influence of other unknown (latent) processes.

2.1 Wiener-Granger causality

The basic principle of WCG is quite simple: Let us assume that $\boldsymbol{X}$ is a variable representing a stochastic process and, similarly, $\boldsymbol{Y}$ is a variable representing a different process. Using past values of $\boldsymbol{X}$ we try to predict its future values. If incorporating previous values of $\boldsymbol{Y}$ into our prediction model of $\boldsymbol{X}$ helps predicting its future values, $\boldsymbol{Y}$ is said to *Granger-cause* $\boldsymbol{X}$, i.e. $\boldsymbol{Y}$ contains information about the future values of $\boldsymbol{X}$ (Granger, 1969). This definition can further be extended to incorporate a third variable $\boldsymbol{Z}$ which will serve as a *conditioning* variable.

With respect to this definition we formalize the idea behind G-causality using *auto-regressive (AR) modelling*. First, let us describe the variables representing stochastic processes: We describe the variable $\boldsymbol{X}$ as a (potentially infinite) matrix of column vectors, where each vector $\boldsymbol{x}_t$ represents the values of variable $\boldsymbol{X}$ at a time t:

$$\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n), \boldsymbol{x}_i \in \mathbb{R}^d, d \in \mathbb{N}$$

analogously $\boldsymbol{Y}$:

$$\boldsymbol{Y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n), \boldsymbol{y}_i \in \mathbb{R}^d, d \in \mathbb{N}$$

With this description of variables we proceed to formalize our prediction models. Note that since we intend to use linear auto-regressive models, we assume that variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ are:

1. stochastic
2. wide-sense stationary (i.e. its mean value and variance do not change over time and/or space)

Let us define $\boldsymbol{X}_{t-1}^{(m)}$ as a vector matrix containing the previous (lagged) values of $\boldsymbol{X}$, bounded by the *model order* (m), formally:

$$\boldsymbol{X}_{t-1}^{(m)} = (\boldsymbol{x}_{t-1}, \boldsymbol{x}_{t-2}, \ldots, \boldsymbol{x}_{t-m})$$

where $\boldsymbol{x}_{t-i}$ is a column vector with the values of $\boldsymbol{X}$ at time $t-i$, m is a fixed number of lagged observations.
Now let us consider these two following AR models:

1. restricted:
$$\boldsymbol{X}_t = A \cdot \boldsymbol{X}_{t-1}^{(m)} + \epsilon_t$$
2. unrestricted:
$$\boldsymbol{X}_t = A' \cdot \left(\boldsymbol{X}_{t-1}^{(m)} \oplus \boldsymbol{Y}_{t-1}^{(m)}\right) + \epsilon_t'$$

Symbol $\oplus$ represents the concatenation of column vectors $\boldsymbol{X}_{t-1}^{(m)}$ and $\boldsymbol{Y}_{t-1}^{(m)}$. We can view this operation as a means to extend the parameter space by appending the lagged values of the variable $\boldsymbol{Y}$ to the values of $\boldsymbol{X}$. Matrices A and A' represent regression coefficients, ϵ and ϵ' are *residuals* (also called

'prediction errors'). Both are obtained by standard linear auto-regression methods, these include ordinary least squares and Yule-Walker equations (Kay, 1988).

If the variance of ϵ' (the residual of the unrestricted model) is *significantly* smaller than the variance of ϵ (the residual of the restricted model), there is an improvement in the prediction of the future values of $\boldsymbol{X}$ by incorporating the past values of $\boldsymbol{Y}$.

2.2 Time-domain analysis

In this part we will focus on the time-domain analysis of stochastic variables. First, we will describe the basic notion behind the comparison of the residuals of restricted and unrestricted models. In order to proceed, let us first make some adjustments to the previous definitions of variables and AR models. In this case we assume that the variable $\boldsymbol{X}$ is scalar (denoted as X), i.e. the value of observation at time t (X_t) is a single real number. Therefore, we update the notation of restricted and unrestricted models respectively:

$$X_t = A \cdot X_{t-1}^{(m)} + \epsilon_t$$
$$X_t = A' \cdot \left(X_{t-1}^{(m)} \oplus Y_{t-1}^{(m)} \right) + \epsilon'_t$$

$X_{t-1}^{(m)}$ now represents a vector of real numbers – the lagged observations from $t-1$ to $t-m$, analogously for $Y_{t-1}^{(m)}$. A and A' are the regression coefficient vectors and ϵ and ϵ' are scalar values.

Using the following equation, we obtain the *magnitude* of the causal influence of $\boldsymbol{Y}$ on scalar X (Geweke, 1984):

$$\mathcal{F}_{Y\to X} = \ln \frac{\mathrm{var}(\epsilon)}{\mathrm{var}(\epsilon')}$$

The variance of ϵ (and ϵ') is obtained from the set of every ϵ_t (and ϵ'_t) across all of the input data of the respective AR models. Thus, the following can be inferred from this equation:

1. it is possible, for scalar X and Y, that Y *is G-causal to* X and, at the same time, X *is G-causal to* Y – a feedback stochastic process (Geweke, 1984)

2. the G-causal magnitude will never be negative

3. the statistic significance can be determined using the F-statistic test (Greene, 2002)

$$F = \frac{\frac{\mathrm{RSS_r} - \mathrm{RSS_{ur}}}{m}}{\frac{\mathrm{RSS_{ur}}}{T-2m-1}}$$

where $\mathrm{RSS_r}$ and $\mathrm{RSS_{ur}}$ are the *Residual Sums of Squares* of restricted and unrestricted models

$$\mathrm{RSS} = \sum_{t=m+1}^{T} \epsilon_t^2$$

T is the total number of the observations used to create the unrestricted AR model and m and $(T - 2m - 1)$ are the degrees of freedom. With a significant result of this F-statistic, it is safe to assume that the unrestricted model provides better predictions than those obtained from the restricted one and, therefore, $\boldsymbol{Y}$ *exerts a G-causal influence on* X.

2.2.1 Conditional G-causality

It is possible to extend the original pairwise analysis of two variables to contain a third, *conditioning* variable. Let us assume that this third variable $\boldsymbol{Z}$ is also a wide-sense stationary stochastic process. Thus, we extend our original description of the restricted and unrestricted models as follows:

$$X_t = A \cdot \left(X_{t-1}^{(m)} \oplus Z_{t-1}^{(m)} \right) + \epsilon_t$$
$$X_t = A' \cdot \left(X_{t-1}^{(m)} \oplus Y_{t-1}^{(m)} \oplus Z_{t-1}^{(m)} \right) + \epsilon'_t$$

The magnitude of causal interaction is then defined(Geweke, 1984) as:

$$\mathcal{F}_{Y \to X|Z} = \ln \frac{\mathrm{var}(\epsilon)}{\mathrm{var}(\epsilon')}$$

This formulation can be interpreted as *the magnitude of G-causal influence of Y on X in the presence of Z.*

2.3 Spectral analysis

Another important perspective on G-causality is the counterpart of the time-domain analysis – *the frequency-domain decomposition.* In this spectral analysis, instead of focusing on each variable as a whole, we break down the signals of each variable into frequencies and measure the G-causal influence of these specific frequencies between variables. The sum of G-causalities of all frequency components (from zero to Nyquist frequency[1]) of a variable is equal to its total time-domain G-causality. We assume that X and Y are

[1] Nyquist frequency is the half of the sampling frequency of a signal. *Aliasing* of a sampled signal can be avoided if the Nyquist frequency is greater than the greatest component frequency.

both scalar and the magnitude of G-causal influence of frequency $\boldsymbol{\omega}$ is given by (Geweke, 1982, 1984):

$$\mathcal{F}_{Y\to X}(\boldsymbol{\omega}) = \ln \frac{S_{XX}(\boldsymbol{\omega})}{\tilde{H}_{XX}(\boldsymbol{\omega})\ \mathrm{var}(\epsilon'_t)\ \tilde{H}^*_{XX}(\boldsymbol{\omega})}$$

where $S_{XX}(\boldsymbol{\omega})$ is the autospectrum of X_t, $\tilde{H}_{XX}(\boldsymbol{\omega})$ is the (X, X) element of the normalized form of the transfer matrix $\tilde{H}(\boldsymbol{\omega})$ and $\tilde{H}^*_{XX}(\boldsymbol{\omega})$ is the complex conjugate of $\tilde{H}_{XX}(\boldsymbol{\omega})$. On a side note, $\tilde{H}_{XX}(\boldsymbol{\omega})$ is the inverse of the normalized spectral coefficient matrix $\tilde{A}(\boldsymbol{\omega})$, which is obtained as the Fourier transform of time-domain coefficient matrix of unrestricted models for X and Y (Ding et al., 2006).

In contrast to the time-domain analysis (which uses F-statistic testing), the statistical significance of frequency-domain interactions can be assessed by surrogate statistical tests, such as *bootstrapping* or *permutation testing* (Seth, 2010). For a more detailed description of these two methods, see Appendix B.

Spectral G-causality proved important in neuroscience, where causal interactions between neural populations often depend on the oscillatory synchrony (Bressler and Seth, 2010). Also, spectral analysis has an ambition to provide an insight into *mediated* causal interactions, as opposed to the time-domain analysis (illustrated in Figure 1).

2.4 Causal density

Causal density serves as a means to quantify the overall causal interactivity throughout a complex system. Using the magnitudes of G-causal interactions ($\mathcal{F}$) between the variables, it is possible to define the causal density as the

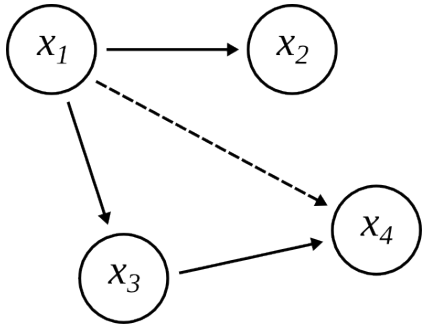

Fig. 1: Example of a system of interdependent variables. Influence of variable X_1 on variable X_4 can be mediated by variable X_3.

average of all pairwise G-causalities (Seth et al., 2011):

$$\mathrm{CD}(\boldsymbol{X}) = \frac{1}{n(n-1)} \sum_{i \neq j} \mathcal{F}_{X_i \to X_j | X_{[i,j]}}$$

where $\boldsymbol{X}$ is the set of all the measured variables, X_i and X_j are variables from $\boldsymbol{X}$ and $X_{[i,j]}$ denotes a subset of $\boldsymbol{X}$ with X_i and X_j omitted. This measure is also referred to as the *weighted* causal density, since every interaction is weighted by its magnitude. An unweighted, normalized version of causal density is defined as (Seth et al., 2011):

$$\mathrm{CD}_n(\boldsymbol{X}) = \frac{\alpha}{n(n-1)}$$

where α is the number of significant G-causal interactions.

Causal density provides such a measure of dynamical complexity of a system, in which a zero value is achieved only when the elements are completely independent from one another, or their behavior is completely integrated in their dynamics. Therefore, a non-zero result can be interpreted so that the behavior of elements is, to a certain degree, independent (so that novel information can be introduced into the predictions), but is, at the same time, globally integrated, i.e. the behavior of one variable exerts an influence on

another variable (Seth et al., 2011). An element with high causal density is called a *causal hub*.

2.5 Causal flow

Causal flow is a straightforward measure for individual variables. It is simply the difference between outward in inward G-causal influences of a given variable:

$$\mathrm{CF}(\boldsymbol{X}_n) = \deg_{\mathrm{out}}(\boldsymbol{X}) - \deg_{\mathrm{in}}(\boldsymbol{X})$$

where $\boldsymbol{X}_n$ is the n-th variable and $\deg_{\mathrm{out}}$ and $\deg_{\mathrm{in}}$ are the numbers of outward and inward interactions, respectively.

This measure is used to identify *causal sources* – the variables which have a strong influence throughout the system (high positive causal flow) and *causal sinks* – the variables which depend significantly on the other variables in the system (low negative causal flow).

2.6 Partial G-causality

One of the greatest hurdles in analysis of real-life data is the fact that we are *not aware* of the underlying structure of the analyzed system. Suppose we try to analyze the data acquired from fMRI or multielectrode arrays (MEAs) of a living brain during certain cognitive processes. In such a case, there are two important aspects which must be accounted for in order to produce relevant insights:

1. the underlying causal structure is generally unknown prior to our analysis, making our effort an *exploratory* one (in contrast with a confirmatory one)

2. in most cases we are bound to have only a partial record of the processes we observe – e.g. using the MEA approach we are capable to record only the behavior of a certain area of a brain

The problem which arises from the first issue is straightforward – if we are unaware of the underlying structure, we are unable to verify our measured causality results against the real causality structure. But since we are performing an exploratory analysis, such a lack of any previous knowledge is expected. The second issue brings a rather complicated problem. As stated earlier, we expect only partial information about the observed system, therefore it is very likely that these processes are under the influence of other parts of the system which we do not monitor – *exogenous inputs* and/or under the influence of the processes that we are not aware of at all – *latent variables*. The presence of such unaccounted variables may give way for unreliable causality results. For example, if two observed variables which are independent from one another are both influenced by a third, unknown, variable, G-causality analysis may (falsely) conclude that there is a causal relationship between these two variables, when there is only a correlation (caused by the third variable). There is also a possibility that the latent variables can create so much "noise" in our observed data that existing causal interactions (which would otherwise be correctly discovered) will remain invisible for the G-causality analysis.

Because of these issues a method based on the conditional G-causality has been proposed (Guo et al., 2008). It is based on the idea that the influences of latent variables may be detectable as the correlations between the residuals of AR models. This is analogous to the concept of *partial correlation* (Barnett et al., 2009) as we attempt to dampen the distortive influence of exogenous inputs by incorporating a third term. Note that this analogy is not quite

accurate since prior knowledge of the variances of the latent variables (which is not available in our case) is needed in order to use the partial correlation.

Now let us consider the definition of AR models used for conditional G-causality. We define the regressions of the conditional variable $\boldsymbol{Z}$ as follows:

$$\boldsymbol{Z_t} = B \cdot \left(\boldsymbol{X}_{t-1}^{(m)} \oplus \boldsymbol{Z}_{t-1}^{(m)} \right) + \eta_t$$
$$\boldsymbol{Z_t} = B' \cdot \left(\boldsymbol{X}_{t-1}^{(m)} \oplus \boldsymbol{Y}_{t-1}^{(m)} \oplus \boldsymbol{Z}_{t-1}^{(m)} \right) + \eta'_t$$

we can see that the roles of the 'predictee' and the conditioning variables have been switched. The partial G-causality is defined as:

$$\mathcal{F}^{p}_{\boldsymbol{Y} \to X | \boldsymbol{Z}} = \ln \frac{\sum (\epsilon_t | \eta_t)}{\sum (\epsilon'_t | \eta'_t)}$$

where $\sum$ represents the partial covariance:

$$\sum (\epsilon | \eta) = \operatorname{cov}(\epsilon) - \operatorname{cov}(\epsilon, \boldsymbol{\eta}) \operatorname{cov}(\boldsymbol{\eta})^{-1} \operatorname{cov}(\epsilon, \boldsymbol{\eta})^{\top}$$

Using this equation we obtain the value of the G-causal influence of variable $\boldsymbol{Y}$ on $\boldsymbol{X}$ in the context of $\boldsymbol{Z}$. Here, the $\boldsymbol{Z}$ variable can be viewed as a mediator by which we are able to capture the influences on the correlation of other unknown variables.

In theory, since we observe the changes in the correlations of residuals, partial G-causality will be useful only when the influence of exogenous and latent variables is equal on all observed variables. As it turns out, partial G-causality holds even in such cases when this exogenous influence differs among the observed variables. This is demonstrated in Section 4.1.5 where we analyze artificially generated data using this method. In the same section, we provide a performance comparison of both the conditional and partial G-

causality methods which clearly shows a superior performance of the partial G-causality over the conditional version.

2.7 Multivariate G-causality

Up to this point, we have been focusing on the univariate version of G-causality. One of the greatest shortcomings of this standard approach is that it only permits an univariate pairwise analysis of interactions of elements in a system.

Imagine the context of neuroanalysis in which we observe possibly a great number of variables in various different brain regions. For example, we observe the behavior of the neurons responsible for vision and at the same time the behavior of the neurons located in the prefrontal cortex (which has the executive function). In such setting we are not only interested in the causal influence of different neurons on one another, but also (and possibly interested even more) in the influence of one region on the second one, thus bringing a *scale* to our analysis. One of the possible approaches is to produce univariate time series of every region of interest, for example by averaging of all of the components (neurons), and then perform the standard G-causality analysis on these "reduced" time series.

In this section we will focus on the multivariate approach of G-analysis for obtaining causal interactions between multivariate groups of variables. Each group variable will consist of values of its component sub-variables at given time. For this purpose we extend the standard G-causality for evaluating

multivariate variables. First, we revise the original AR models:

$$\boldsymbol{X}_t = \boldsymbol{A} \cdot \left(\boldsymbol{X}_{t-1}^{(m)} \oplus \boldsymbol{Z}_{t-1}^{(m)} \right) + \boldsymbol{\epsilon}_t$$
$$\boldsymbol{X}_t = \boldsymbol{A}' \cdot \left(\boldsymbol{X}_{t-1}^{(m)} \oplus \boldsymbol{Y}_{t-1}^{(m)} \oplus \boldsymbol{Z}_{t-1}^{(m)} \right) + \boldsymbol{\epsilon}'_t$$

As we can see, the original concept remains intact as we are again comparing the restricted and unrestricted models. The difference here lies in the variables and residuals. $\boldsymbol{X}, \boldsymbol{Y}$ and $\boldsymbol{Z}$ now represent multivariate random variables (random vectors) as so do the residuals $\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}'$.

To clarify this notation, $\boldsymbol{X}$ (and all the other variables) is a potentially infinite matrix consisting of column vectors. Each of these column vectors represents the state of the sub-variables of which $\boldsymbol{X}$ consists.

$$\boldsymbol{X} = (\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n}) \qquad \boldsymbol{X_i} \in \mathbb{R}^d, d \in \mathbb{N}$$
$$\boldsymbol{X}_t = (x_1, x_2, \ldots, x_n)^\top \qquad x_i \in \mathbb{R}$$

Residuals $\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}'$ are defined in the same fashion.

Before we proceed to the definition of the equation for the multivariate version of G-causality, let us lay down a couple of notational conventions about the statistical operations (covariance and its extensions) used in this section. Suppose we have random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ which represent multivariate random variables – matrices of column vectors, each representing values of n sub-variables at time t, then:

- $\Sigma(\boldsymbol{X}, \boldsymbol{Y})$ represents the $n \times m$ cross-covariance matrix (covariances between sub-variables $\boldsymbol{X}_i$ and $\boldsymbol{Y}_j$
- $\Sigma(\boldsymbol{X}) = \Sigma(\boldsymbol{X}, \boldsymbol{X})$ represents the $n \times n$ covariance matrix (covariances between sub-variables $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$

- $\Sigma(\boldsymbol{X}|\boldsymbol{Y})$ represents the $n \times n$ *partial-covariance* matrix (Barrett et al., 2010) – the covariance matrix of the residuals of a linear regression of $\boldsymbol{X}$ on $\boldsymbol{Y}$, if $\Sigma(\boldsymbol{Y})$ is invertible it is then defined as:

$$\Sigma(\boldsymbol{X}|\boldsymbol{Y}) = \Sigma(\boldsymbol{X}) - \Sigma(\boldsymbol{X},\boldsymbol{Y})\,\Sigma(\boldsymbol{Y})^{-1}\,\Sigma(\boldsymbol{X},\boldsymbol{Y})^{\top}$$

- $\Sigma(\boldsymbol{X},\boldsymbol{Y}|\boldsymbol{Z})$ represents the $n \times m$ *partial-cross covariance* matrix (Barrett et al., 2010), if $\Sigma(\boldsymbol{Z})$ is invertible it is then defined as:

$$\Sigma(\boldsymbol{X},\boldsymbol{Y}|\boldsymbol{Z}) = \Sigma(\boldsymbol{X},\boldsymbol{Y}) - \Sigma(\boldsymbol{X},\boldsymbol{Z})\,\Sigma(\boldsymbol{Z})^{-1}\,\Sigma(\boldsymbol{Y},\boldsymbol{Z})^{\top}$$

2.7.1 Multivariate time-domain analysis

Building upon the previous definitions of G-causality we again try to measure the magnitude of the influence of one variable on another by comparing the changes in the accuracy of two prediction models:

$$\begin{aligned}\boldsymbol{X}_t &= A \cdot \left(\boldsymbol{X}_{\boldsymbol{t-1}}^{(\boldsymbol{p})} \oplus \boldsymbol{Z}_{\boldsymbol{t-1}}^{(\boldsymbol{r})}\right) + \boldsymbol{\epsilon}_t \\ \boldsymbol{X}_t &= A' \cdot \left(\boldsymbol{X}_{\boldsymbol{t-1}}^{(\boldsymbol{p})} \oplus \boldsymbol{Y}_{\boldsymbol{t-1}}^{(\boldsymbol{q})} \oplus \boldsymbol{Z}_{\boldsymbol{t-1}}^{(\boldsymbol{r})}\right) + \boldsymbol{\epsilon}'_t\end{aligned}$$

Here the variable $\boldsymbol{X}$ is first regressed on the previous p lags of itself and r lags of the conditioning variable $\boldsymbol{Z}$ and second also on the previous q lags of $\boldsymbol{Y}$.

Note that the notation $\boldsymbol{X}_t^{(p)} = (\boldsymbol{X}_t \oplus \boldsymbol{X}_{t-1} \oplus \cdots \oplus \boldsymbol{X}_{t-p})$ again represents the *vertical concatenation* of vectors $\boldsymbol{X}_t \ldots \boldsymbol{X}_{t-p}$, therefore $\boldsymbol{X}_t^{(p)}$ is a random vector of dimension np.

In the multivariate case of G-causality we are unable to compare the differences in the accuracy of the restricted and unrestricted models using simple

variance, since it can be obtained only from univariate data. Thus we are bound to make use of the *covariance* measure, which on the other hand produces *covariance matrices* (in contrast with variance which produces scalar values) and therefore we need to modify our G-causality measure (Barrett et al., 2010):

$$\mathcal{F}_{\boldsymbol{Y}\to X|\boldsymbol{Z}} = \ln\left(\frac{\operatorname{var}(\epsilon_t)}{\operatorname{var}(\epsilon'_t)}\right) = \ln\left(\frac{\Sigma(X|\boldsymbol{X}^{(p)}_{t-1}\oplus\boldsymbol{Z}^{(r)}_{t-1})}{\Sigma(X|\boldsymbol{X}^{(p)}_{t-1}\oplus\boldsymbol{Y}^{(q)}_{t-1}\oplus\boldsymbol{Z}^{(r)}_{t-1})}\right)$$

note that X here is scalar.

Two approaches for handling of the covariance matrices of the residuals have been proposed:

Total variance which is the trace of the covariance matrix – sum of values on the main diagonal of a square matrix $\operatorname{tr}(\boldsymbol{X}) = \sum_{i=1}^{n} X_{i,i}$.

The approach of total variance seems as a natural extension to the original G-causality (Ladroue et al., 2009) since the total variance is a common choice for assessing the fitness of a prediction model. At the same time, the total variance approach smoothly reduces to the original G-causality measure if the predictee variable is scalar. It is defined as follows (Barrett et al., 2010):

$$\mathcal{F}^{\mathrm{tr}}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}} = \ln\left(\frac{\operatorname{tr}\left[\Sigma(\boldsymbol{X}|\boldsymbol{X}^{(p)}_{t-1}\oplus\boldsymbol{Z}^{(r)}_{t-1})\right]}{\operatorname{tr}\left[\Sigma(\boldsymbol{X}|\boldsymbol{X}^{(p)}_{t-1}\oplus\boldsymbol{Y}^{(q)}_{t-1}\oplus\boldsymbol{Z}^{(r)}_{t-1})\right]}\right)$$

Generalized variance which is the determinant of the covariance matrix, denoted as $|\Sigma(\boldsymbol{\epsilon}_t)|$. This measure quantifies the *volume* in which the residuals lie (Barrett et al., 2010; Geweke, 1982). The generalized-variance approach for measuring G-causality is defined as follows (Barrett et al., 2010):

$$\mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}} = \ln\left(\frac{\left|\Sigma(\boldsymbol{X}|\boldsymbol{X}_{t-1}^{(p)} \oplus \boldsymbol{Z}_{t-1}^{(r)})\right|}{\left|\Sigma(\boldsymbol{X}|\boldsymbol{X}_{t-1}^{(p)} \oplus \boldsymbol{Y}_{t-1}^{(q)} \oplus \boldsymbol{Z}_{t-1}^{(r)})\right|}\right)$$

This measure, similarly as the total variance approach, will always be non-negative and will smoothly reduce to the original G-causality measure when the predictee variable is scalar. Barrett et al. (Barrett et al., 2010) concluded that the generalized variance approach is more suitable for the multivariate spectral G-causality compared to the total variance version, therefore this approach will be preferred in our thesis.

2.7.2 Multivariate causal density

The causal density measure which has been described in Section 2.4 is strictly univariate since it is the average of the pairwise G-causalities between the elements of a system (in the context of all the other elements of the same system). In the context of the multivariate G-causality a number of the multivariate extensions of the original causal density have been proposed (Barrett et al., 2010). These extensions provide a more detailed overview of the causal density on various *scales* of a complex system.

Group causal density is an extension of causal density which quantifies the causal density from size k to size r, in the context (conditioned on) of the rest of the system

$$\mathrm{gCD}_{k\to r}\left(\boldsymbol{X}\right) = \frac{1}{n_{k,r}} \sum_{k=1}^{n_{k,r}} \mathcal{F}_{\boldsymbol{V}_i^k \to \boldsymbol{U}_i^r | \boldsymbol{W}_i^{n-k-r}}$$

where $\boldsymbol{V}_i^k$, $\boldsymbol{U}_i^r$ and $\boldsymbol{W}_i^{n-k-r}$ are the ith tripartition of the system $\boldsymbol{X}$, with sizes k, r and $n-k-r$, respectively. The number $n_{k,r} \equiv \binom{n}{k}\binom{n-k}{r}$ is the

total number of all possible tripartitions of $\boldsymbol{X}$.

Bipartition causal density Using the group causal density, we can define another causal density measure which is defined as the average causal density sustained by a system divided into (every possible) two partitions

$$\mathrm{bCD}\left(\boldsymbol{X}\right) = \frac{1}{n-1}\sum_{k=1}^{n-1}\mathrm{gCD}_{k\to(n-k)}(\boldsymbol{X})$$

Scaled causal density Another interesting view on the causal density would be a comparison at different scales of the predictor and the predictee divisions

$$\mathrm{sCD}_s\left(\boldsymbol{X}\right) = \frac{1}{s-1}\sum_{k=1}^{s-1}\mathrm{gCD}_{k\to(s-k)}(\boldsymbol{X})$$

Note that $\mathrm{sCD}_2\left(\boldsymbol{X}\right)$ would be the same as the original pairwise causal density and $\mathrm{sCD}_n\left(\boldsymbol{X}\right)$ is equivalent to $\mathrm{bCD}\left(\boldsymbol{X}\right)$.

Tripartition causal density Finally, the average of all the causal densities at every scale of a system could provide an overall measure

$$\mathrm{tCD}\left(\boldsymbol{X}\right) = \frac{1}{n-1}\sum_{s=2}^{n}\mathrm{sCD}_s(\boldsymbol{X})$$

2.7.3 Partial multivariate G-causality

As described in Partial G-causality in our attempts to discover the causal interactions in a system we may come across the issue of exogenous and latent variables which could possibly render a correct analysis a rather difficult task. For such cases a method of *partial G-causality* has been proposed.

This method can be extended for multivariate data by combining the principle behind the original univariate case (based on the analogy with the partial correlation) with the approach of the generalized variance used in the standard multivariate G-causality.

Thus, we again switch the roles of the predictee and the conditioning variables in our AR models (Barrett et al., 2010)

$$\begin{aligned}\boldsymbol{Z_t} &= B \cdot \left(\boldsymbol{X}_{t-1}^{(p)} \oplus \boldsymbol{Z}_{t-1}^{(r)}\right) + \boldsymbol{\eta}_t \\ \boldsymbol{Z_t} &= B' \cdot \left(\boldsymbol{X}_{t-1}^{(p)} \oplus \boldsymbol{Y}_{t-1}^{(q)} \oplus \boldsymbol{Z}_{t-1}^{(r)}\right) + \boldsymbol{\eta}'_t\end{aligned}$$

with $\boldsymbol{Z}$ and $\boldsymbol{\eta}$ being random vectors and $\boldsymbol{X}_{t-1}^{(p)}$, $\boldsymbol{Y}_{t-1}^{(q)}$ and $\boldsymbol{Z}_{t-1}^{(r)}$ being the vertical concatenations of p, q and r lagged vector values, respectively.

Now we can extend the original partial G-causality measure based on the partial covariance

$$\mathcal{F}^p_{\boldsymbol{Y}\to X|\boldsymbol{Z}} = \ln\left(\frac{\Sigma\left(\boldsymbol{\epsilon}_t|\boldsymbol{\eta}_t\right)}{\Sigma\left(\boldsymbol{\epsilon}'_t|\boldsymbol{\eta}'_t\right)}\right)$$

to the multivariate case using Geweke's generalized variance (the determinant of the covariance matrices) (Barrett et al., 2010)

$$\mathcal{F}^p_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}} = \ln\left(\frac{\left|\Sigma\left(\boldsymbol{\epsilon}_t|\boldsymbol{\eta}_t\right)\right|}{\left|\Sigma\left(\boldsymbol{\epsilon}'_t|\boldsymbol{\eta}'_t\right)\right|}\right) = \ln\left(\frac{\left|\Sigma(X|\boldsymbol{X}_{t-1}^{(p)} \oplus \boldsymbol{Z}_{t-1}^{(r)})\right|}{\left|\Sigma(X|\boldsymbol{X}_{t-1}^{(p)} \oplus \boldsymbol{Y}_{t-1}^{(q)} \oplus \boldsymbol{Z}_{t-1}^{(r)})\right|}\right)$$

This switching of the role of $\boldsymbol{Z}$ to the position of a predictee can be viewed as using the $\boldsymbol{Z}$ variable as a sort of a "proxy" which helps us to capture the exogenous influences of the unobserved variables.

3 Emergence and downward causation

Emergence is a phenomenon by which a complex property arises from relatively simple interactions between a (large) number of elements. Such a macro-level property is not just a mere sum of its micro-level properties, it exhibits a behavior somewhat independent from its micro-element substrate. Examples of emergent processes can be found in the nature (snowflake forming, termites building their colonies) as well as in the context of artificial computer simulations (cyclic and repetitive behavior of the populations in Conway's game of life).

Depending on the conditions deemed necessary to be met such that a process can be considered emergent, three definitions of emergence have been proposed (Seth, 2008; Bedau, 2003)

Nominal emergence in the simplest terms, it is a property which can be possessed only by the macro-level entities and not by its micro-level constituents. For example a computer software system is emergent from the lines of code from which it is composed. This is the most trivial 'version' of emergence which is in principle fully predictable and therefore we will not consider it any further.

Strong emergence is probably the most controversial variety of emergence. It combines two crucial statements about the properties of a process. First, a macro-level property is unidentifiable from the observations of the micro-level properties alone. In consequence, a mechanistic reduction of a emergent process would be a direct violation of this condition. Second, the macro-level property exerts causal influences which are not explainable in terms of the micro-level interactions.

This brings forth a rather problematic notion of *downward causation*. Two main issues connected to the downward causation have been raised: First, such a causal influence of a macro-element on its micro-level constituents is rather contradictory to the statement that the macro-level property is the way it is due to those same micro-level elements which it influences. Second, if we accept such a causal influence we then have to resolve the conflicting nature of the macro- and micro-level causal influences in our analysis. An example of the strong emergence could be the emergence of conscious states (e.g. qualia) from neuro-biological processes, but this particular example can rather be viewed as a result of our lack of understanding of the consciousness as a whole rather than as a canonical demonstration of the strong emergence.

Weak emergence can be viewed as a "compromise" between the nominal and strong variations of emergence. Compared to the nominal one, weakly emergent properties cannot be inferred trivially from micro-level interactions. In contrast with strong emergence, macro-level properties are in principle identifiable from micro-level elements. In other words, the macro-level property can be in fact derived from the interactions of its micro-components, but these interactions are so complex that the macro-level property cannot be explained solely by the micro-level behavior. It has been postulated that the weakly emergent macro-properties have to be *epistemologically irreducible* (Bedau, 1997) – the macro-property is underivable from its micro properties *except by simulation* (Seth, 2008).

In this section we will build upon the notion of weak emergence, namely on the statement that a property is weakly emergent if it is not identifiable from the observations of the micro-level interactions alone. We will also

discuss a new measure which will, drawing from this intuition about weak emergence, serve as a tool for quantifying emergent behavior.

Building upon the previous principles of G-causality, we introduce a complementary measure – *G-autonomy* – and we will continue in this vein to define the measure of *G-emergence*. This will be illustrated on a well-know example of emergence: *Boid flocking*. Finally, we will discuss the metaphysical notion of *downward causation* – the causal influence of a macroscopic property on its microscopic elements.

3.1 G-autonomy

Autonomy can be viewed as a property of an element (of a complex system), which describes *how independent* is such an element from the rest of the system. That is, how little information about the behavior of such an autonomous element is contained in the behavior of other elements.

Intuitively, we will define such a measure upon the definition of G-causality. A process is G-autonomous if the introduction of its past values into prediction model gives better predictions of its future values than the predictions based solely on the past values of other processes (Seth, 2008). Thus, let is revise our original restricted and unrestricted models:

$$X_t = A \cdot \boldsymbol{Z}_{t-1}^{(m)} + \epsilon_t$$
$$X_t = A' \cdot \left(X_{t-1}^{(m)} \oplus \boldsymbol{Z}_{t-1}^{(m)} \right) + \epsilon'_t$$

where $\boldsymbol{Z}$ is a *set* of all the variables in a system with X omitted. We can see that the restricted model of X *does not* regress onto X itself. Then, the G-autonomy of X is given by:

$$\mathcal{A}_{\mathcal{X}|\boldsymbol{Z}} = \ln \frac{\text{var}(\epsilon)}{\text{var}(\epsilon')}$$

With G-autonomy at hand, we proceed to define the crucial measure of G-emergence.

3.2 G-emergence

As stated earlier, a property is G-emergent if its behavior is both dependent and autonomous from its micro elements. Let us assume that M is a macro-variable, emergent from a set of micro-variables $\boldsymbol{m} = \{m_1, m_2, \ldots, m_n\}$. Using the concepts of G-causality and G-autonomy, the definition of G-emergence is straightforward:

$$\mathcal{E}_{M|\boldsymbol{m}} = \mathcal{A}_{M|\boldsymbol{m}} \left(\frac{1}{N} \sum_{i=1}^{N} \mathcal{F}_{m_i \to M} \right)$$

It is noteworthy that this measure captures the basic requirements for emergence: if M has no autonomy from its micro-properties ($\mathcal{A}_{M|\boldsymbol{m}} = 0$), its G-emergence will also be zero. Similarly, if there is no influence of the micro-variables $\boldsymbol{m}$ on M ($\mathcal{F}_{\boldsymbol{m} \to M} = 0$), G-emergence will again be zero. Thus, a non-zero G-emergence hints at emergent property of M.

3.3 Downward causation

Finally, we elaborate on the notion of downward causation. The influence projected by a macro-level variable M on its micro-level constituents $m_1, m_2, \ldots, m_n$ can be defined as a pairwise G-causal influence of M on one (and possibly more) of the variables $\boldsymbol{x}$ (Seth, 2008; Bedau, 2003):

$$\mathcal{F}^{\text{dw}}_{M,m_i} = \mathcal{F}_{M \to m_i}, m_i \in \boldsymbol{m}$$

Two things should be noted. First, a G-emergent macro-variable, by the definition of weak emergence, is in principle not required to have a detectable G-causal influence on its micro-elements (although *strong emergence* does in fact require such a downward influence) (Seth, 2008). Second, there is a ongoing discussion about the metaphysical admissibility of the notion of downward causation, thus the $\mathcal{F}^{\mathrm{dw}}_{M,m_i}$ measure is at this point rather hypothetical and should be taken with caution (Seth, 2008).

4 Data analysis and results

In this section, we will present direct experimental results of G-causality analysis. In the first part we will analyze various artificially generated time series and examine differences between the time-domain and spectral versions of G-causality analysis. We will start off with a formal definition of our system of processes in which we will check whether the G-causality method discovered its inter-element interactions correctly. Furthermore, we will systematically investigate the effects of three parameters – level of noise, magnitude of the lagged values and the number of time lags on G-causality detection.

In the second part we will focus on the *partial* extension of G-causality, its application on a modified version of Baccala & Sameshima (Baccala and Sameshima, 2001) time series which will include latent variables and exogenous inputs. Again we will first describe these processes formally and then proceed to the actual analysis. We will compare the results obtained from the conditional and partial versions of G-causality analysis applied on different variations of this system – the influences whose different magnitudes of the exogenous and latent inputs exert on the overall correctness of the analysis.

Finally, we will proceed to the experimental application of the ideas and measures described in Section 3. First, we will provide a short example of the G-autonomy measure on a well-defined set of processes in a similar fashion as the G-causality example. Second, we will provide an example of a G-emergence measurement on a textbook example of an emergent behavior – *bird flocking*.

4.1 Time-series analysis

In this part, we will analyze the interactions in a system comprised of six variables whose design has been inspired by the system proposed by Baccala & Sameshima (Baccala and Sameshima, 2001). We will show that using the method of G-causality and F-statistic testing of the statistical significance we are able to obtain a correct overview of the interactions in the entire system based solely on the observations of different processes.

4.1.1 Formal description of variables

Here we provide a formal definition of the six variables which will be the subject of our experimental analysis:

$$
\begin{aligned}
x_1(t) &= 0.5025x_1(t-1) + 0.25x_1(t-3) + w_1(t) \\
x_2(t) &= 0.35x_1(t-2) + w_2(t) \\
x_3(t) &= -0.5x_1(t-1) + 0.95x_3(t-1) - 0.9025x_6(t-3) + w_3(t) \\
x_4(t) &= 0.3x_3(t-3) - 0.9025x_5(t-4) + w_4(t) \\
x_5(t) &= -0.3x_4(t-5) + 0.5x_6(t-1) + w_5(t) \\
x_6(t) &= -0.4x_5(t-3) + 0.4x_6(t-2) + w_6(t)
\end{aligned}
$$

Note that w_i is a white noise with a normal distribution ($\mu = 0$ and $\sigma^2 = 1$).

Figure 2 provides a visual overview of the dependence between these signals. Figure 3 displays the raw observations of our variables. We can see that x_3 has a visible influence on certain variables (e.g. x_4) by comparing the amplitudes of x_3 and other variables. There is a visible *echoing* (with a certain lag) of the peaks of x_3 by x_4, this is a sign that x_3 is very likely to G-cause x_4.

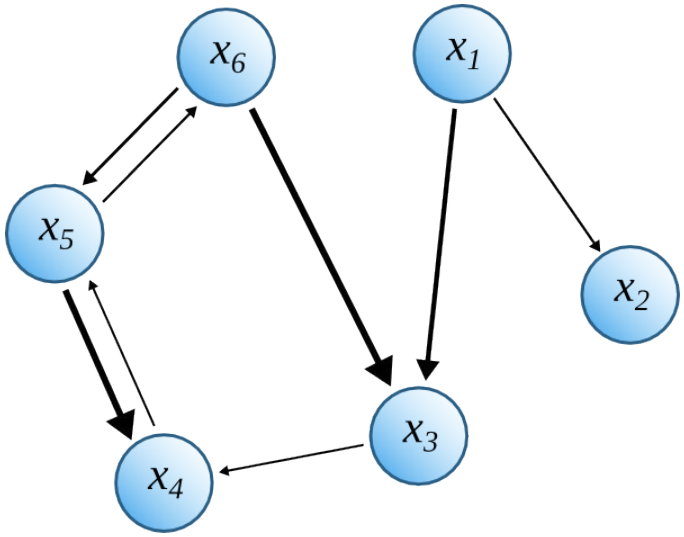

Fig. 2: Six variables representing interdependent processes. Different widths of the oriented lines correspond to the approximate magnitude of causal influences between processes.

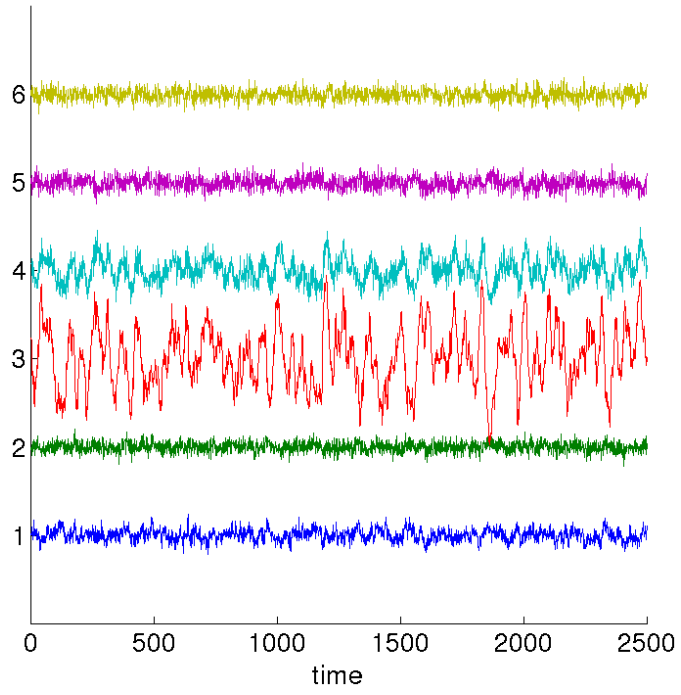

Fig. 3: A raw visualization of 2500 observations of the six variables.

4.1.2 Time-domain analysis

In order to analyze G-causality within our system we are first required to build our restricted and unrestricted AR models. We will generate six time series, each containing the values of our six processes at arbitrary time t. Each of these time series will contain 2500 raw observations. For the purpose of generating such data we provide our custom written MATLAB script on the DVD supplement as well as the sample files containing values generated by this script.

Before we proceed to build our actual AR models we need to specify the *model order* for our models. The model order is a value which specifies how many lagged observations are to be included in the prediction of a process' value at a certain time. This value needs to be selected carefully and *by hand* and its impact on the *fitness* of AR models will be extremely significant. If the value is too low, the AR models will most likely generate unreliable predictions of the values. This is due to the insufficient amount of the data needed to capture the internal properties of a process – the so called problem of *underfitting*. In contrast, if the model order value is too high the AR models' prediction capability will probably suffer by the virtue of the problem of *overfitting*. That is, the AR model will "predict" the values upon which it was generated almost perfectly but at the expense of failure to correctly predict the future values from the input outside of the model's training set.

Since the methods for assessing a perfectly fitting model order do not exist we have to take a different approach to find an estimate for the best model order. With an *a priori* knowledge of the analyzed system we can choose model order directly (e.g. in our six variable example, the model order corresponds to the highest time lag $t - 5$, the model order can thus be specified as 5). Another approach, given the input data, is to systematically

test the "correctness" of the AR models with a number of different model orders. This can be done using the Akaike information criterion (AIC) (Akaike, 1974; Bressler and Seth, 2010) or the Bayesian information criterion (BIC) (Schwartz, 1978; Bressler and Seth, 2010). Both of these methods attempt to find a balanced model order value depending on the fitness of the model together with the number of the parameters used in that model. For a more detailed description of these two methods, see Appendix A.

Using our input data, both AIC and BIC return an estimate of the best model order equal to 5, which can be deemed as the correct one.

With the input data and the desired model order at hand, we proceed to obtain the G-causalities using our method.

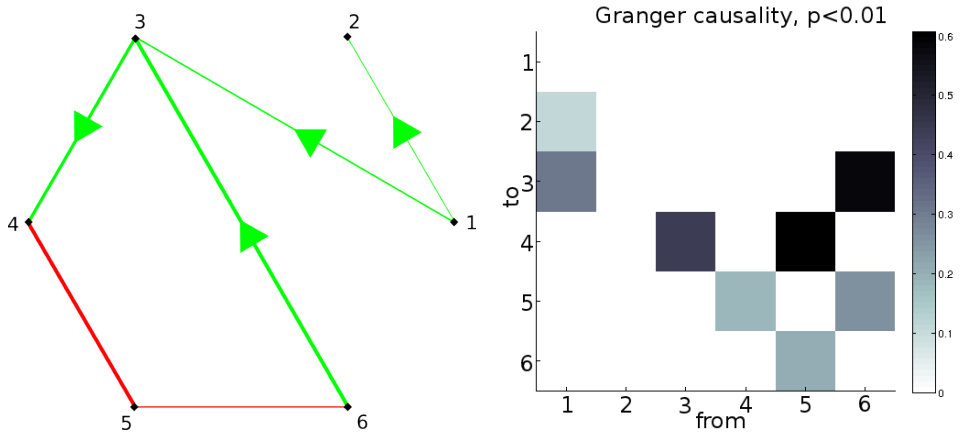

Fig. 4: Time-domain analysis in MATLAB toolbox. The graph represents the G-causal interactions identified by our method (green lines represent oriented one-way interaction, red lines represent two-way interactions). The width of these lines corresponds to the G-causal magnitudes. The matrix on the right represents pairwise interactions (a column G-causes a row), the darker is the color, the stronger is the influence.

We can see that the G-causality method identified both the unidirectional and bidirectional G-causal relations correctly (left). Green lines represent oriented one-way influences, red lines represent two-way influences. The line

width corresponds to the magnitude of the interaction (the width of the two-way lines corresponds to the greater magnitude of the two interactions). Matrix on the right visualizes the influences between the processes (a column causes a row) – darker color means a stronger influence, lighter color means a weaker influence. Again, we can see that the magnitudes as well as the orientations of the G-causal interactions are identified correctly.

In Table 1 we present the numerical results for every significant G-causal interaction. The statistical significance of a interaction is assessed by the F-test described in Section 2. Using the F-test evaluation we obtain the G-causality probability – denoted as the p-value (using the *cumulative distribution function*). The closer is the p-value to 0, the more significant is the respective G-causal interaction. In our analysis we use $p < 0.01$ as a threshold for identifying the significant G-causal interactions.

interaction	G-causality	F-test	p
$x_1 \to x_2$	0.0902	46.51	0
$x_1 \to x_3$	0.3340	195.53	0
$x_3 \to x_4$	0.4430	274.82	0
$x_4 \to x_5$	0.1761	94.94	0
$x_5 \to x_4$	0.6751	475.38	0
$x_5 \to x_6$	0.1866	101.15	0
$x_6 \to x_3$	0.6286	431.40	0
$x_6 \to x_5$	0.2521	141.37	0

Tab. 1: The properties of significant causal interactions between the processes (significance $p < 0.01$)

For a complete numerical overview of all the interactions in our system, see Table 7 on page 66.

We now proceed to the assessment of the *causal density* which describes the causal complexity sustained by our example system. Note that here we

differentiate between the *unit causal density* (unit cd), which is simply a ratio between the significant causal interactions and all the other interactions of a process, and *weighted unit causal density* (w-unit cd) which represents the unit causal density of the same process weighted by the magnitudes of the significant causal interactions.

variable	unit cd	w-unit cd	in-deg	out-deg	c. flow
x_1	0.3333	0.0707	0	2	2
x_2	0.1667	0.0150	1	0	-1
x_3	0.5000	0.2343	2	1	-1
x_4	0.5000	0.2157	2	1	-1
x_5	0.6667	0.2150	2	2	0
x_6	0.5000	0.1779	1	2	1
total	0.2667	0.0929	–	–	–

Tab. 2: G-causal properties of the processes

In Table 2 we provide the causal density as well as the causal flow properties of every variable. As stated in Section 2.5, the causal flow is the difference between outward and inward G-causal interactions of a given process. In our example, process x_1 has been correctly identified as the system's *causal source* (the highest positive causal flow) and process x_2 as a *causal sink* (the process does not exert any causal interactions whatsoever).
Figure 5 shows causal density (left) and causal flow (right) distribution between variables. Bars represent unweighted causal density and flow, blue lines show their weighted counterparts.

4.1.3 Spectral analysis

Continuing from the previous section, we now proceed to the spectral variety of G-causality analysis on the same data set used in Section 4.1.2. We analyze 2500 observations of our six variables' G-causalities in a frequency range of 0–150 Hz, with the sampling rate of 500 Hz and using the *bootstrapping*

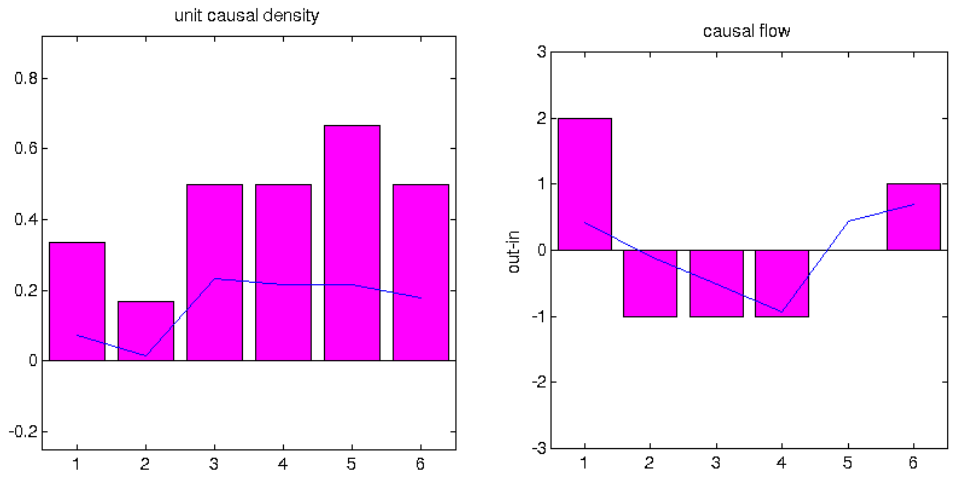

Fig. 5: Causal density (left) and causal flow (right) of the system. The unit causal density graph shows two sets of values: the bars represent the normalized causal densities of every variable; the blue line represents the weighted variety of causal density. The causal flow graph displays the causal flow of each variable, weighted and normalized varieties are displayed in the same fashion as in the causal density graph.

method (repeated 1000 times) to assess the statistical significance of the G-causal influence at every frequency. Our threshold for statistical significance will again be set to $p < 0.01$.

Figure 7 shows the average spectral G-causal influence of all variables in the system (peak at $\approx$ 5 Hz). And finally, Figure 8 shows G-causalities matrix where a column G-causes a row. G-causalities are again distributed among the frequencies in the range 0–150 Hz. Note that G-causalities of some frequencies seem to be "cut off". This is due to the fact that these particular frequencies and their influences were identified as not significant enough ($p \geq 0.01$).

Two interesting aspects of the spectral analysis of can be observed here: First, we can see that all the variables (except x_2) exhibit an influential behavior on processes to which they are not directly connected (if we consider

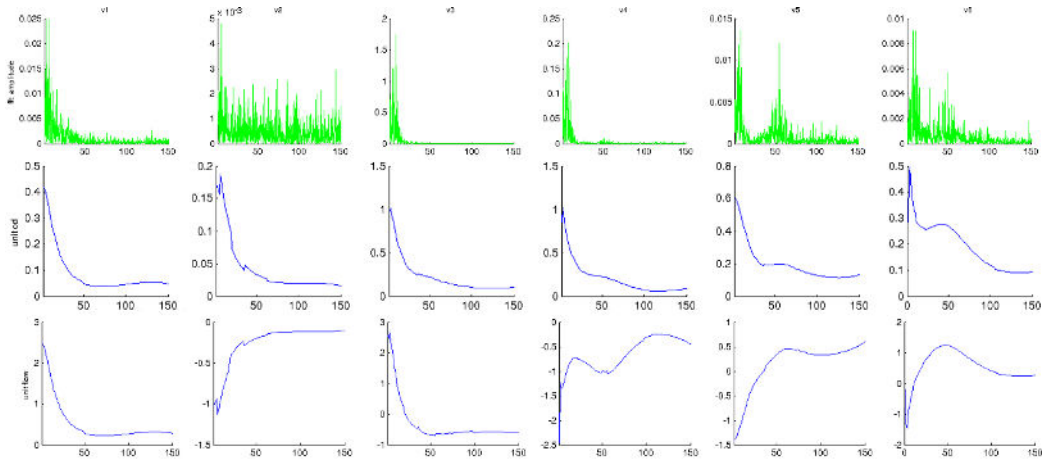

Fig. 6: Spectral decomposition of variables and causal densities/flows of frequencies. Top row displays the power spectrum of every variable (column represents variable) obtained using Fast Fourier Transform with sampling rate of 500 Hz. Center row contains per variable G-causalities of frequencies in range of 0–150 Hz. Bottom row contains causal flows distributed by frequencies, again, from 0 to 150 Hz.

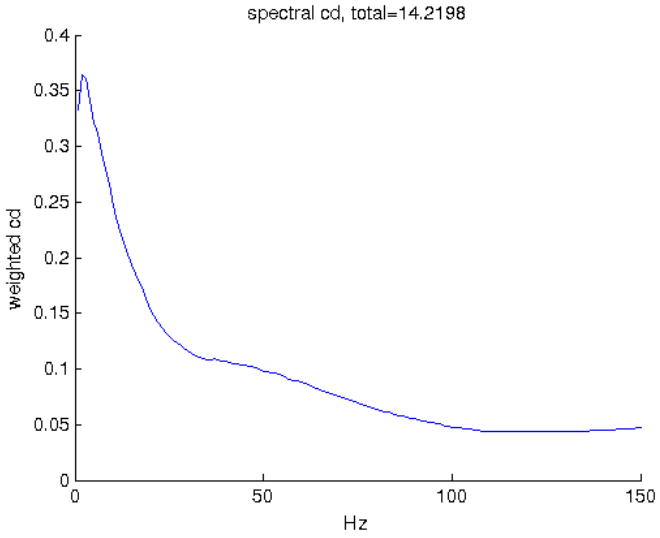

Fig. 7: The mean spectral G-causality of the system.

the formal description of variables). For example, take the x_6 variable. The influence of the frequencies at ≈ 50 Hz seem to "leak" through the variables x_3 and x_5 (to which x_6 is connected) on the variable x_4. Similar situation arises

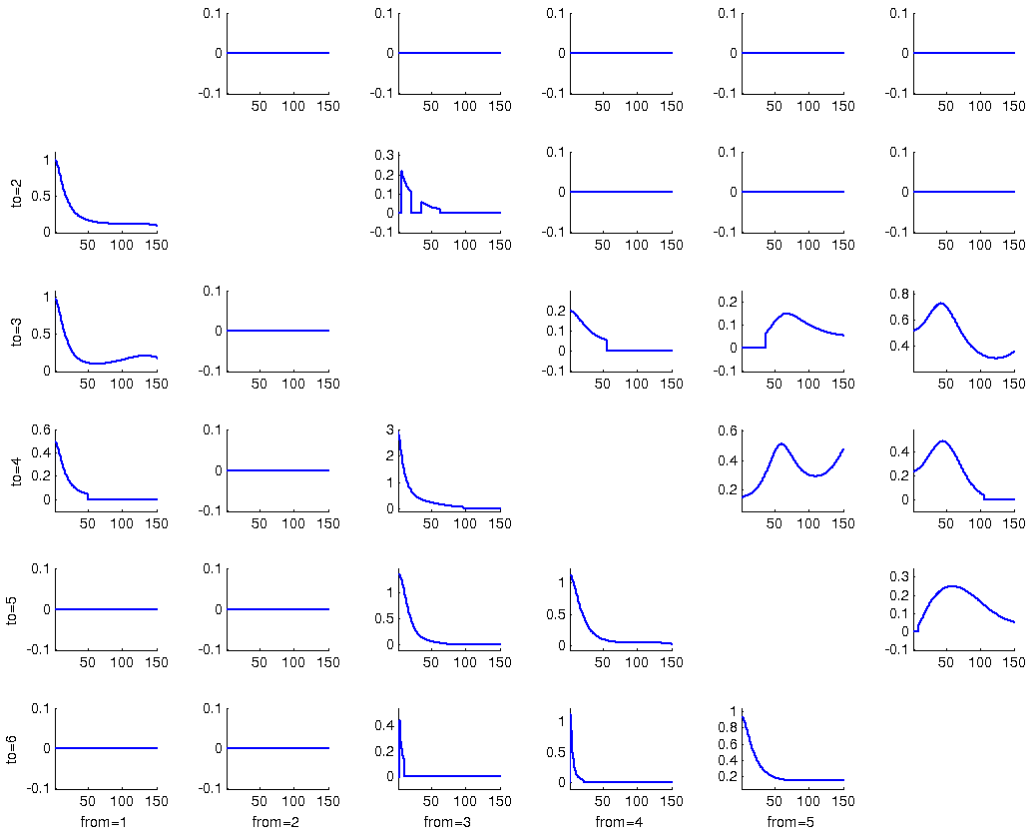

Fig. 8: Spectral G-causalities divided between variables (a column causes a row). The magnitudes of G-causal interactions are divided among the frequencies in the range 0–150 Hz.

in the case of the variable x_1 which seems to influence the variable x_4 (the frequencies at ≈ 5 Hz) by the virtue of a connection to the variable x_3. This can possibly be interpreted as an example of a *mediated G-causal influence* – an influence of one process through another on a third one. This assumption is also backed by the actual structure of our system. If we consider our system as an oriented graph whose vertices are our processes and edges are (oriented) causal influences, we can identify these mediated connections as *paths* consisting of more than two vertices.

A second, less favorable aspect is the presence of what seems as spurious

G-causal influences. An example of such an erroneous influence can be seen in the case of $x_3 \to x_2$. Although not strong at any of the 150 frequencies (with a portion of frequencies giving statistically insignificant influences altogether), we can safely conclude that this result does not reflect the actual causal substrate of our system. In order to exert a mediated influence on x_2, x_3 would need to influence x_1 (which is the only actual contributor to x_2). This is clearly not the case since x_1 is an autonomous process (its values are based solely on its previous values and white noise). Thus we can conclude that the notion of the *mediated* G-causality should be taken with caution when assessing possibly unforeseen interactions between the processes.

4.1.4 Impact of noise and time lags

In this section we modify our model in order to be able to adjust certain properties of the inter-variable interactions. Three properties will be subject to our testing:

1. the strength of white noise
2. the strength of lagged values
3. the length of time lags

We will observe the impact of these properties on the overall (weighted) causal density of the system. We first proceed to modify the formal definition

of our system:

$$
\begin{aligned}
x_1(t) &= \alpha\ 0.5025x_1(t-1-t_c) + \alpha\ 0.25x_1(t-3-t_c) + \beta\ w_1(t) \\
x_2(t) &= \alpha\ 0.35x_1(t-2-t_c) + \beta\ w_2(t) \\
x_3(t) &= -\alpha\ 0.5x_1(t-1-t_c) + \alpha\ 0.95x_3(t-1-t_c) \\
&\quad -\alpha\ 0.9025x_6(t-3-t_c) + \beta\ w_3(t) \\
x_4(t) &= \alpha\ 0.3x_3(t-3-t_c) - \alpha\ 0.9025x_5(t-4-t_c) + \beta\ w_4(t) \\
x_5(t) &= -\alpha\ 0.3x_4(t-5-t_c) + \alpha\ 0.5x_6(t-1-t_c) + \beta\ w_5(t) \\
x_6(t) &= -\alpha\ 0.4x_5(t-3-t_c) + \alpha\ 0.4x_6(t-2-t_c) + \beta\ w_6(t)
\end{aligned}
$$

where α is the coefficient modulating the strength of the lagged values, β is the coefficient of the strength of white noise and t_c is a positive natural number specifying an additional lag to the original one.

In our modification we observe the influence of α, β and t_c values set globally for the whole network in ranges $[0.5, 1.1]$, $[1.0, 2.0]$ and $[0, 10]$, respectively.

Figure 9 shows the change of the global weighted causal density (the darker the color, the higher the causal density) depending on the strength of the white noise and the strength of lagged values (left), and causal density's dependence on the noise and additional time lag (right). It is clear that noise has a strong distortive influence on the accuracy of the G-causality analysis, since in both left and right graphs, the causal density rapidly drops with the increasing interference of noise. The influence of the strength of lagged values is on the other hand the exact opposite – with an ascending α coefficient the overall causal density also rises. However, our experimental measurements point at a certain positive "indifference" of the G-causality analysis towards the increasing additional time lag, i.e. causal density remains steady even

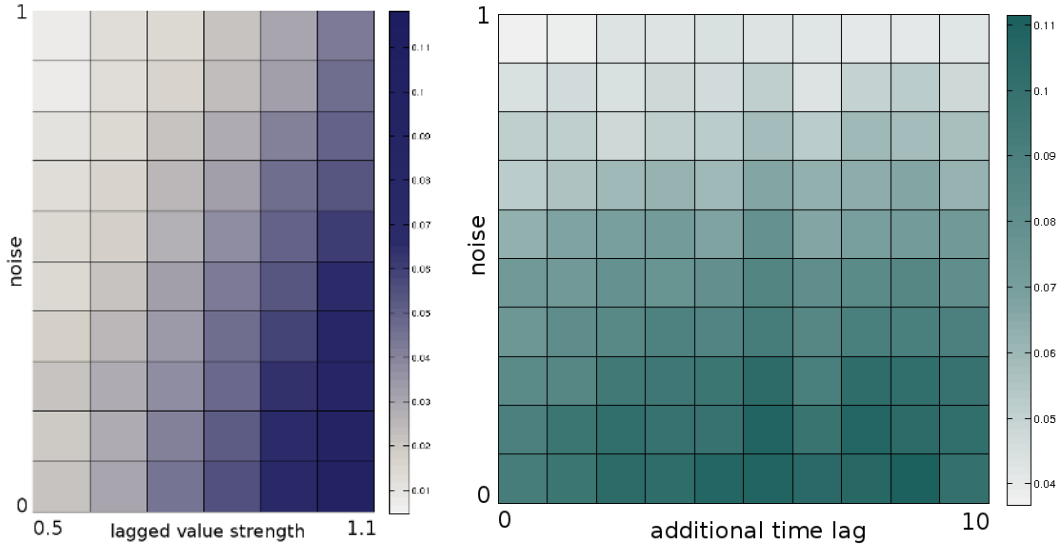

Fig. 9: The influence of the white noise and the strength of lagged values (left) on the overall causal density in the system and the influence of the white noise and additional time lags (right) on the overall causal density.

with different (longer) additional time lags. Note that during our analysis each time series underwent the assessment of the most fitting model order using AIC and BIC and therefore, AR models were generated with these additional time lags taken into account. Nonetheless, it is visible that G-causality holds even in the case of greater distances in time and is able to mitigate its influence.

4.1.5 Partial G-causality

We conclude the time-series analysis section by demonstrating the partial G-analysis, which can be potentially used to reduce the distortive effects of exogenous inputs and latent variables (Guo et al., 2008), by comparing its performance with the conditional G-causality. In our experiment we will analyze a simpler system (compared to the one described in Formal description

of variables), its formal definition is as follows:

$$x_1(t) = 0.3\ x_1(t-2) + 0.85\ x_3(t-1) + w_1(t) + \alpha_1 e_1 + \beta_1 e_2$$
$$x_2(t) = 0.7\ x_2(t-1) - 0.15\ x_2(t-3) + w_2(t) + \alpha_2 e_1 + \beta_2 e_2$$
$$x_3(t) = -0.75\ x_2(t-3) + 0.15\ x_4(t-2) + w_3(t) + \alpha_3 e_1 + \beta_3 e_2$$
$$x_4(t) = 0.5\ x_3(t-2) + 0.5\ x_4(t-1) + w_4(t) + \alpha_4 e_1 + \beta_4 e_2$$

where the additional terms e_1 and e_2 represent exogenous and latent inputs respectively, and α_i and β_i represent the magnitude coefficients of these additional influences. Note that in our experiment the e_1 variable is treated as an exogenous input and e_2 as a latent variable. In every test we will use 3σ as the confidence interval.

During our testing of both versions of G-causality, we are going to compare their performances on a set of three different configurations of α_i and β_i coefficients. The first configuration, with $\alpha_i = 0$ and $\beta_i = 0$ for all the variables, serves as a reference point for the rest of the coefficient configurations. Thus by excluding the influences of e_1 and e_2 altogether, we test whether these methods are able to correctly infer the underlying G-causal network in the most permissive of the cases. In the second configuration, we modify $\alpha_i = 2$, the same for all four variables while keeping $\beta_i = 0$ at its original value. In the third configuration, we modify both α_i and β_i coefficients. The coefficient α_i is now set to contain random values, selected from the range $[0, 1]$ with a different value for every variable while $\beta_i = 1$ for each of these variables.

Our results, visualized in Figure 10, can be interpreted as follows. The first referential case ($\alpha_i = 0$, $\beta_i = 0$) shows that both methods correctly reduce to the standard G-causality analysis and the underlying network of the

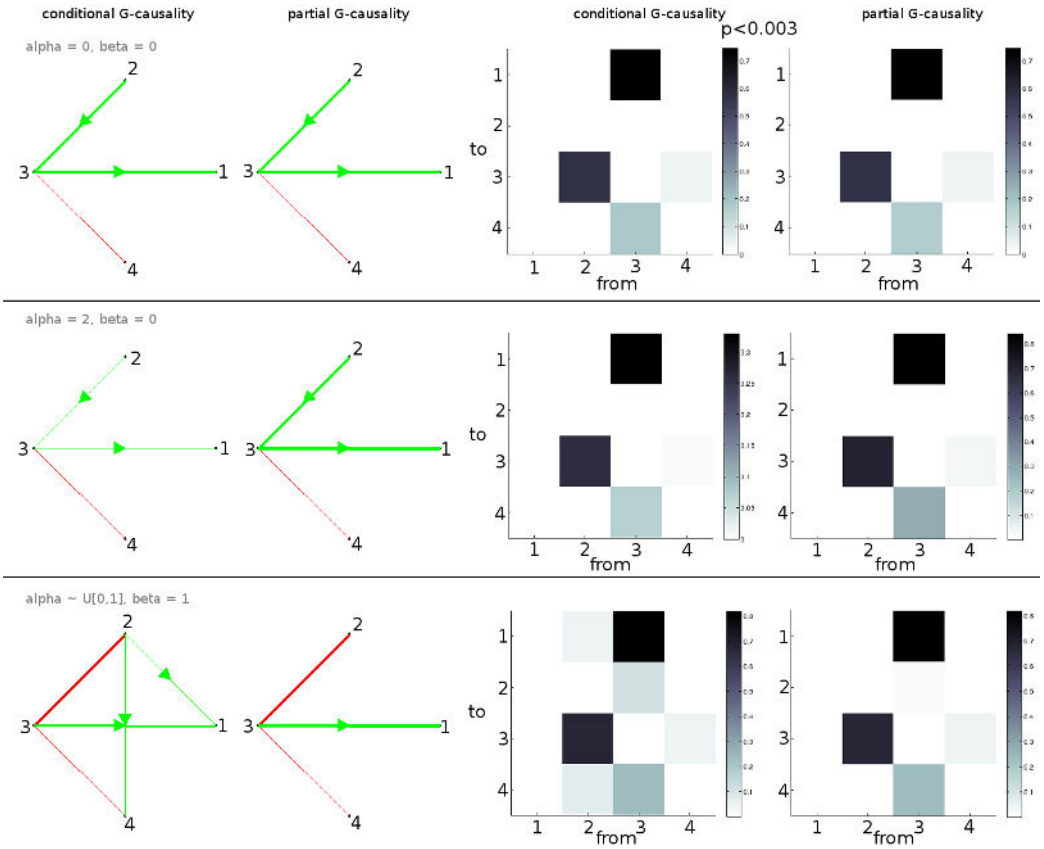

Fig. 10: Comparison of the partial and conditional versions of G-causality. Left side contains G-causal network models inferred by both versions, right side G-causality matrices (a column causes a row) of significant ($p < 0.003$) interactions with scales of their magnitudes. Each of the three rows represents a different configuration of α_i and β_i coefficients for e_1 and e_2.

causal influences is captured appropriately. Thus we proceed to the second case with the configuration $\alpha_i = 2$, $\beta_i = 0$. In this case we see a minor shift between these two versions of G-causality – although both versions have discovered the network correctly, the G-causal magnitudes measured by the conditional G-causality are comparatively lower, compared to those measured by the partial variety. In the third case, we can see that the conditional version failed to accurately capture the underlying network of causal interactions in our system by acquiring three additional spurious (false) influences

(as shown in the third row of Figure 10). However, the result from the partial G-causality, although containing one spurious influence of $x_3 \rightarrow x_2$ with a relatively small magnitude, fared quite well and discovered the main causal structure in the system.

4.2 G-autonomy

In this section, which will serve as an "interlude" between the G-causality and G-emergence measures' demonstrations, we will briefly demonstrate the behavior of the G-autonomy measure (see Section G-autonomy). G-autonomy is a complementary measure to the G-causality one in the sense that it quantifies the independence of a process from the rest of the system rather than its dependence on any other process.

To demonstrate this measure we will analyze a simple set of four variables, defined as follows:

$$
\begin{aligned}
x_1(t) &= w_1(t) \\
x_2(t) &= 0.5x_1(t-2) + 0.65x_2(t-2) + w_2(t) \\
x_3(t) &= 0.5x_2(t-3) + w_3(t) \\
x_4(t) &= 0.9025x_1(t-2) - 0.3x_4(t-2) + w_4(t)
\end{aligned}
$$

where w_i is a normally distributed white noise.

From this definition we can see that only x_2 and x_4 are autonomous processes since they contain auto-regressive components (the lagged values of themselves). The x_3 process contains the lagged values of another process (x_2) and thus is not autonomous. Finally, x_1 is completely random (and therefore cannot be predicted) and thus, it is not autonomous.

We now proceed to the actual results obtained by our G-autonomy ana-

lysis. In Table 3, we can clearly see that x_2 and x_4 were correctly identified as G-autonomous processes with a strong F-test significance of this property. Note that the significance threshold $p \leq 0.01$. x_1, a completely random process, is not G-autonomous because we are unable to build an appropriate AR model for its predictions. This is backed by the F-test result which gives an extremely low significance of its (very low) G-autonomy. A similar result has been obtained for the process x_3 since it contains only noise and lagged values of process x_2 and therefore is not G-autonomous.

variable	G-autonomy	F-test	p-value	signif.
x_1	0.0009	0.574	0.8	no
x_2	0.4757	377.67	0	yes
x_3	0.0021	1.28	0.25	no
x_4	0.0992	64.64	0	yes

Tab. 3: Properties of G-autonomy analysis of four processes (significance $p \leq 0.01$)

4.3 G-emergence of Boid flocking

In this final section, we demonstrate the usefulness of the G-emergence for analyzing emergent macro-properties of a complex system. Working under the constraint of weak emergence that a macro-property is underivable from its micro-properties except by simulation (Bedau, 1997, 2003) we simulate the behavior of a set of agents of a textbook example of the emergent behavior – flocking. The design of our simulation has been inspired by Reynolds' description of Boids (Reynolds, 1987) which names three rules guiding the behavior of each boid:

1. *grouping* – each boid tends to fly towards its relative (perceived) center of flock,

2. *avoidance* – during its flight, each boids tries to avoid collisions with nearby boids,

3. *matching* – each boid tries to align its velocity, both orientation and magnitude, with the velocity of other nearby boids.

In this example, the absolute center of flock (a sort of a center of mass) can be viewed as our potential macro-variable.

Our implementation of this simulation (in a 2D space) uses vector definitions of both velocity – x, y coordinates relative to the given agent describing its orientation and speed (magnitude), and position – absolute x and y values. These vectors' values are updated after every time step.

In order to comply with the aforementioned rules, we divide the update of the velocity vectors to three steps, afterwards we "move" the given boid's position by its updated velocity vector. Note that the velocities and positions of the boids are updated at a given time as a set – that means that all the

positions are changed only after all the velocities have been updated. This ensures the deterministic nature of the simulation.

Furthermore, we also introduce four variables $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in [0, 1]$ which will adjust the individual influence of these three rules on the behavior of a given boid. This enables us to perform a systematic analysis of a number of different configurations of the system and compare the results of G-emergence and its influence on the flocking behavior.

Before we proceed to the simulation itself, let us first describe the three rules, in terms of a pseudocode, for changing boid velocities.

Algorithm 4.1 Update of position and velocity of every boid at a given time step

```
for each boid do
    Vector r1 = rule1(boid);
    Vector r2 = rule2(boid);
    Vector r3 = rule3(boid);
    boid.velocity = boid.velocity + r1 + r2 + r3;
    boid.position = boid.position + boid.velocity;
end for
```

Algorithm 4.1 describes the operations performed at every time step where we first update given boid's velocity and then change its current position.

The first rule for updating velocities is based on moving the boid towards its *perceived (relative) center of mass*. The position of the relative center of mass is obtained by averaging the positions of other boids except the one whose velocity we are updating. From this position we obtain the relative orientation of a boid towards the center of mass by subtracting boids position from the position of the perceived center of mass. Note that the first coefficient α_1 modifies the magnitude of the resulting vector which will be added to the velocity vector. This operation is described in Algorithm 4.2.

Algorithm 4.2 Rule 1 - grouping

```
function RULE1(boid)
    Vector centerOfMass;
    for each Boid b do
        if boid ≠ b then
            centerOfMass = centerOfMass + b.position;
        end if
    end for
    centerOfMass = centerOfMass / N - 1;
    return (centerOfMass - boid.position) * α1;
end function
```

Algorithm 4.3 Rule 2 - avoidance

```
function RULE2(boid)
    Vector displacement;
    float maxDistance = (maxDimension / 10) * α2;
    for each Boid nearest do
        if boid ≠ nearest and distance(boid, nearest) ≤ maxDistance then
            orientation = nearest.position - boid.position;
            displacement = displacement - orientation;
        end if
    end for
    return displacement * α3;
end function
```

The second rule – the avoidance of nearby boids, operates by selecting every boid whose distance is within a specified arbitrary distance from the current boid, then obtaining the relative orientation towards the nearest boids and subtracting it from the displacement vector which will consequently be used to modify the given boid's velocity vector. This operation is described in Algorithm 4.3. The influence of this rule is managed by two coefficients: α_2 modifies the range within which are the other boids considered as 'near'. Note that the maximum value for this range ($\alpha_2 = 1$) is the tenth of the dimension of the 2D space in which the simulation takes place. The second coefficient α_3 (as in the case of α_1) modifies the magnitude of the displace-

ment vector and thus its influence on the resulting velocity vector of a boid.

Algorithm 4.4 Rule 3 - matching

```
function RULE3(boid)
    Vector meanVelocity;
    for each Boid b do
        if boid ≠ b then
            meanVelocity = meanVelocity + b.velocity;
        end if
    end for
    meanVelocity = meanVelocity / N - 1;
    return (meanVelocity - boid.velocity) * α4;
end function
```

The final rule, as shown in Algorithm 4.4, modifies the velocity vector of a boid to align with the *perceived velocity* of other boids. We proceed similarly as in the first rule by averaging the velocities of every boid except the one whose velocity is updated and then by obtaining the orientation towards this perceived velocity. Again, the strength of the influence of this rule is managed by α_4.

With the implementation of these rules at hand, we proceed to the actual simulation of the flocking behavior. First, we randomly place $N = 10$ boids in a toroidal continuous 2D space with dimensions 500×500. Each boid will possess its current position (x, y coordinates) and its current velocity (a vector with relative orientation and absolute magnitude). On a side note, in order to keep the model stable, we limit the speed (the length of the velocity vector) to the range $[3, 9]$. The simulation runs at arbitrary discrete time steps. After each step, the velocity and position of each boid is updated according to the three rules. During our experiment we analyze $4 \times 4 \times 4 \times 4 = 256$ different configurations of α_i coefficients:

coeff.	values			
α_1	0.01	0.05	0.07	0.1
α_2	0.1	0.3	0.5	1.0
α_3	0.2	0.3	0.5	0.8
α_4	0.1	0.125	0.25	0.5

Tab. 4: Tested values for α_i coefficients

4.3.1 G-emergence results

For each of the 256 test cases we analyzed 5000 observations (time steps) of the simulated flocking behavior of modified boids. Our analysis contains 10 + 1 observed processes, i.e. 10 boids plus the center of mass. Each observation (for boids and center of mass alike) is the distance from the absolute center of the simulation space. We use this reduction of dimensionality in order to avoid over-complicating of the AR models. The model order $m = 6$ was selected using the Akaike information criterion.

Three configurations for α_i were tested. Configuration c_R (random) which produced almost completely random behavior of the boids, configuration c_L (low) which created inadequate flocking behavior due to the strong influence of the matching of the perceived velocity, and configuration c_H (high) which produced a compelling flocking behavior.

config.	α_1	α_2	α_3	α_4
c_H	0.01	0.3	0.8	0.1
c_L	0.01	0.5	0.5	0.5
c_R	0.01	1.0	0.8	0.25

Tab. 5: Three configurations of α_i coefficients; c_H (high) produced a compelling flocking behavior, c_L evoked poor flocking and c_R produced a near-random flight of the boids.

Figure 11 shows the trajectories of boids under the given configuration of α_i coefficients. The top left part of this figure shows the comparison of

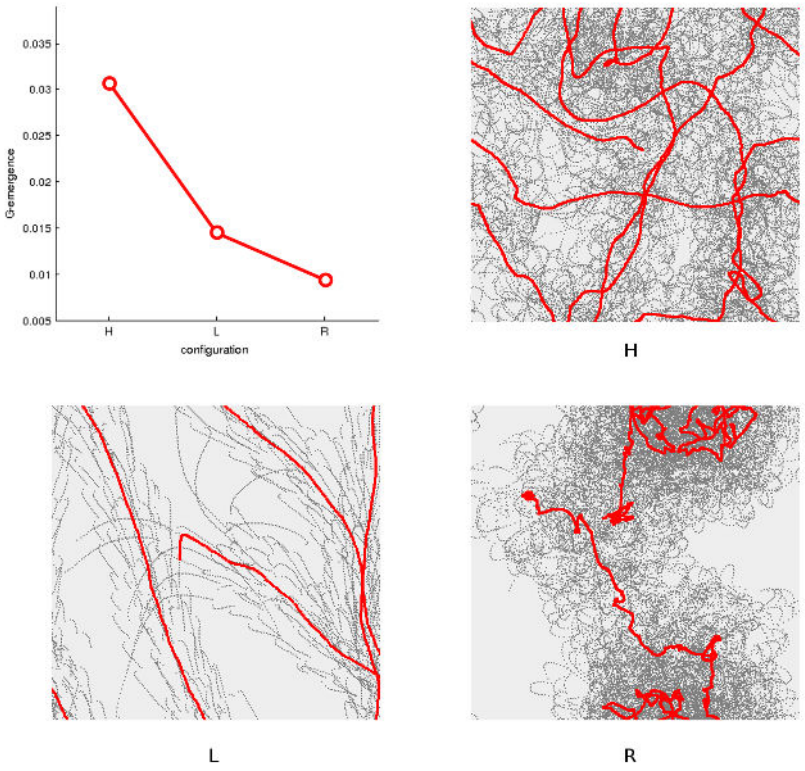

Fig. 11: G-emergence of the center of mass (top left). Example traces of the boids (dotted gray lines) and flock's center of mass (red line) under different coefficient configurations (c_H, c_L and c_R).

the G-emergence measurements for these three configurations. We can see that the c_H configuration provided both high G-emergence value and also a visually compelling flocking behavior.

Figure 12 represents the cross-sections through the four-dimensional coefficient space. Each cross-section contains the c_H configuration (the intersections of green lines). Gray scale shows G-emergence of flock's center of mass. Two aspects of the above visualizations should be noted. First, the "depth" of G-emergence changes smoothly in most regions of the parameter space, hinting at the robust nature of the G-emergence measure. Second, some regions exhibit steep transitions, for example in the top left cross-section we can see

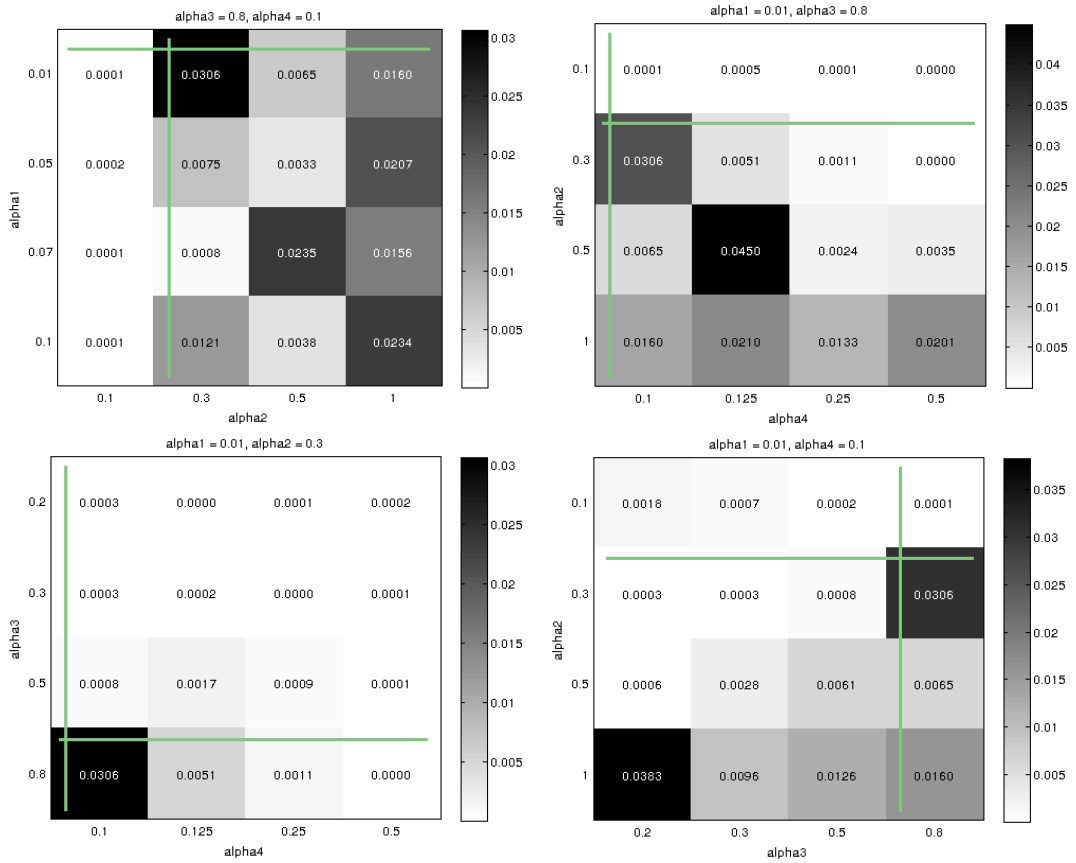

Fig. 12: Parameter space of the simulation model. Each of the four maps contains the G-emergence values (the darker the color, the higher the G-emergence value) for different combinations of α_i configurations. The intersection of green lines indicates the c_H configuration.

two "islands" of stronger G-emergence. The sensitivity of the G-emergence measure towards these sharp transitions hints at a possibility of its usefulness for identifying configuration regions in which a non-trivial weak emergence is present (Seth, 2008).

Finally, we elaborate on the notion of downward causation. Table 6 shows different values of G-causality of the flock's center of mass on the individual boids. Although we can see that the mean G-causality drops between the three configurations, it should be noted that these results show a rela-

boid	c_H	c_L	c_R
1	0.0030	0.0018	0.0003
2	0.0034	0.0008	0.0003
3	0.0020	0.0028	0.0009
4	0.0018	0.0012	0.0006
5	0.0011	0.0005	0.0015
6	0.0048	0.0025	0.0018
7	0.0069	0.0075	0.0025
8	0.0046	0.0010	0.0019
9	0.0041	0.0013	0.0033
10	0.0041	0.0027	0.0026
mean	0.0036	0.0022	0.0016

Tab. 6: G-causalities of the center of mass on the individual boids under configurations c_H, c_L and c_R.

tively small magnitude of these interactions. Also, our results showed little statistical significance of these magnitudes. It should be noted that these results differ from those reported by Seth (Seth, 2008). Two facts should be mentioned. First, our experiment involved a different simulation model than the model used in Seth's article. Second, weak emergence does not require downward causation of the macro-property to be present (but it is necessary for strong emergence) (Seth, 2008).

We thus can conclude that although our downward causation experiment did not produce expected results, the G-emergence measure served well in the case of the analysis of the flocking behavior of boids. Since the measured values were consistent with our expectations about flock's movements we recommend a further investigation of the G-emergence measure.

5 Conclusion

The measure of Wiener-Granger causality was first described in 1956 in econometrics with the intention to help the analysis of time series of economic processes. However, in recent years it has been proposed to put this measure to use for the analysis of time series recordings of neural activity. This was motivated by the idea that the understanding of the underlying causal interactions in neural systems (such as the human brain) can lead to the fundamental understanding of the set of phenomena commonly known as *consciousness* (Dennett, 1991).

In this thesis we discussed the G-causality analysis and its various extensions which we then put to the test by applying it to a number of experimental data. Our tests first confirmed the usefulness of this measure for discovering the causal interactions of processes solely based on the examination of time series representing the individual variables. We also confirmed that the spectral counterpart of this time-domain analysis produced satisfactory results consistent with the time-domain version, but also hinted at the possibility to discover *mediated* G-causal influences. Furthermore, the partial G-causality extension met out expectations about analysis in the context when latent and exogenous influences are present but not observed in terms of individual time series – which prevented direct knowledge of their behavior. This extension is especially useful in real life application for example in the case of analysis of data obtained by multielectrode array (MEA) equipment (used to directly measure neural activity of living brains).

In the later part we described a measure which combines G-causality analysis with the notion of weak emergence to formalize a means to quantify and measure emergent properties of complex systems. This measure was termed

G-emergence. We then constructed a simulation of a textbook example of emergence behavior – flocking, and tried to obtain results from the simulated data. G-emergence proved useful for identifying emergent behavior and its results were consistent with the visual aspect of the flocking behavior. However, we were unable to produce satisfactory results in the case of measuring the downward causation (influence of the flock on its members).

We conclude that the method of Wiener-Granger causality provides substantially useful results and has a valid potential to serve as tool for both exploratory and confirmatory analysis of time series in terms of causal interactions. The G-emergence measure, although not meeting all of our expectations, fared quite well and we conclude that this method, should it be further investigated, could serve as a meaningful method for discovering emergent behavior in complex systems.

A Determining AR Model Order

One of the most significant factors, which determine the effectiveness of an AR model, is its *model order*. The value of the model order denotes the number of lagged observations from which the AR model is built. Model order which is too low is likely to cause poor estimation performance of the AR model due to insufficient input for regression. Similarly, a too large model order value can render the AR model unusable due to the problem of *overfitting*.

Therefore, a method for evaluating model order's suitability is required to find a well-balanced value. Here, we will briefly describe two criteria capable of such evaluation.

Akaike Information Criterion (AIC)

$$\mathrm{AIC} = 2k - 2\ln(L)$$

where k represents the number of model parameters, L is the maximized value of the likelihood function for the estimated model (Akaike, 1974).

Bayesian Information Criterion (BIC)

$$-2 \cdot \ln P(x|k) \approx \mathrm{BIC} = -2 \cdot \ln L + k \ln(n)$$

where x are the observed values, n is the sample size of x, k is the number of regressors, $P(x|k)$ is the probability of the observed data with respect to the number of parameters k and L is the maximized value of the likelihood function (Schwartz, 1978).

For selection of the best model order, we proceed as follows:

1. we specify the range of tested model orders $m_i \in \mathbb{N}$
2. for each m_i we produce a respective AR model and evaluate it with AIC or BIC (or both)
3. AR model (and its model order), which scored the lowest AIC or BIC value, is the most suitable

B Surrogate statistical methods

Certain qualities (such as the probability distribution of the estimators) which are needed for assessing the statistical significance are in some cases not available by analytical means. This is the case, for example, of the spectral G-causality and the partial G-causality analysis (Seth, 2010). Thus, in order to obtain the statistical significance and confidence intervals of different G-causal interactions we need to employ certain surrogate statistical methods. This approach requires rather extensive and repeated calculations (up to a couple of thousands of repetitions). In this section we will shortly discuss two surrogate methods: *bootstrap* and *permutation* resampling. It should be noted that these methods preserve the G-causal interactions in observed data.

Bootstrap resampling The idea behind bootstrapping is that a single observation can stand for a distribution if it is resampled *with replacement* (Efron and Tibshirani, 1994). In our case of time-series analysis, it means that the input data which is to be analyzed is subdivided into multiple "windows" which are then repeatedly sampled and analyzed. The confidence intervals are then obtained by examining the empirical quantiles of the bootstrap distribution (Seth, 2010).

Permutation resampling The permutation resampling operates similarly as the bootstrap variety. Again we divide the input data into a number of windows but these windows are also rearranged for each of the analyzed variables. This creates matrices of input data which consist of different "parts" of the inputs. This contrasts with the bootstrapping approach in which the data windows remain internally unrearranged (Seth, 2010).

C Data analysis supplement

interaction	G-causality	F-test	p	signif.
$x_1 \to x_2$	0.0902	46.51	0	yes
$x_1 \to x_3$	0.3340	195.52	0	yes
$x_1 \to x_4$	0.0012	0.58	0.71	no
$x_1 \to x_5$	0.0014	0.68	0.64	no
$x_1 \to x_6$	0.0027	1.32	0.25	no
$x_2 \to x_1$	0.0029	1.44	0.21	no
$x_2 \to x_3$	0.0029	1.44	0.21	no
$x_2 \to x_4$	0.0012	0.58	0.71	no
$x_2 \to x_5$	0.0040	1.98	0.08	no
$x_2 \to x_6$	0.0047	2.30	0.04	no
$x_3 \to x_1$	0.0027	1.33	0.25	no
$x_3 \to x_2$	0.0047	2.31	0.04	no
$x_3 \to x_4$	0.4430	274.82	0	yes
$x_3 \to x_5$	0.0013	0.63	0.68	no
$x_3 \to x_6$	0.0039	1.93	0.09	no
$x_4 \to x_1$	0.0029	1.43	0.21	no
$x_4 \to x_2$	0.0006	0.31	0.91	no
$x_4 \to x_3$	0.0010	0.47	0.80	no
$x_4 \to x_5$	0.1761	94.94	0	yes
$x_4 \to x_6$	0.0008	0.38	0.86	no
$x_5 \to x_1$	0.0014	0.70	0.62	no
$x_5 \to x_2$	0.0008	0.39	0.86	no
$x_5 \to x_3$	0.0009	0.45	0.81	no
$x_5 \to x_4$	0.6751	475.38	0	yes
$x_5 \to x_6$	0.1866	101.15	0	yes
$x_6 \to x_1$	0.0025	1.21	0.30	no
$x_6 \to x_2$	0.0015	0.74	0.59	no
$x_6 \to x_3$	0.6286	431.40	0	yes
$x_6 \to x_4$	0.0032	1.59	0.16	no
$x_6 \to x_5$	0.2521	141.37	0	yes

Tab. 7: Complete overview of G-causal interactions in the system described in Time-series analysis

D DVD Supplement

All the actual numerical results presented in this thesis have been obtained either using the G-causality MATLAB toolbox developed by Seth (Seth, 2010) or our custom-made MATLAB software designed specifically for these operations. All of these scripts and other software, as well as digital versions of results presented here, are available on the DVD supplement of this thesis.

The directories contained on the DVD:

thesis/ contains the original LaTeX version of this document

thesis/images contains all the images, graphs and visualizations presented in this thesis

matlab/gcca contains the original GCCA toolbox for MATLAB developed by Seth (Seth, 2010)

matlab/causality contains our custom-made MATLAB tools for G-emergence assessment, generating time-series data, data visualization, etc.

emergence/ contains our implementation of Boids written in Java 1.6

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* , pp. 716–723.

Baccala, L. and K. Sameshima (2001). Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics 84*, pp. 463–474.

Barnett, L., A. Barrett, and A. Seth (2009). Granger causality and transfer entropy are equivalent for Gaussian variables. *Physical Review Letters 103*, pp. 238701.

Barrett, A., L. Barnett, and A. Seth (2010). Multivariate Granger causality and generalized variance. *Physical Review E 81*, pp. 041907.

Bedau, M. (1997). Weak emergence. *Philosophical Perspectives* , pp. 375–399.

Bedau, M. (2003). Downward causation and the autonomy of weak emergence. *Principia 6*, pp. 5–50.

Bressler, S. and A. Seth (2010). Wiener-Granger causality: A well established methodology. *NeuroImage 58(2)*, pp. 323–329.

Dennett, D. C. (1991). *Consciousness Explained*. Little, Brown and Co.

Dennett, D. C. (2003). *Freedom Evolves*. New York, USA: Viking Press.

Ding, M., Y. Chen, and S. Bressler (2006). Granger causality: Basic theory and application to neuroscience. In *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, pp. 437–460. Wiley-VCH.

Efron, B. and R. J. Tibshirani (1994). *Introduction to the Bootstrap*. Chapman & Hall.

Farkaš, I. (2007). Hľadanie kauzálnych vzťahov v probléme mysle a tela z pohľadu neredukcionistického fyzikalizmu. *Myseľ, inteligencia a život* , pp. 3–16.

Farkaš, I. (2010). Mental causation in a physical brain? *Brain-Inspired Cognitive Systems* .

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association 77*, pp. 304–313.

Geweke, J. (1984). Measures of conditional linear dependence and feedback between time series. *Journal of the American Statistical Association 79*, pp. 907–915.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37*, pp. 424–438.

Greene, W. H. (2002). *Econometric Analysis* (5th ed.). Prentice Hall.

Guo, S., A. Seth, K. Kendrick, C. Zhou, and J. Feng (2008). Partial Granger causality: eliminating exogenous inputs and latent variables. *Journal of Neuroscience Methods 172*, pp. 79–93.

Kay, S. (1988). *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall.

Ladroue, C., S. Guo, K. Kendrick, and J. Feng (2009). Beyond element-wise interactions: identifying complex interactions in biological processes. *PLoS One 4*.

Pearl, J. (2000). *Causality: Models, reasoning, and inference* (2nd ed.). The Edinburgh Building, Cambridge CB2 2RU, UK: Cambridge University Press.

Reynolds, C. (1987). Flocks, herds and schools: A distributed behavioral model. *Computer Graphics* , pp. 25–34.

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics* , pp. 461–464.

Seth, A. (2008). Measuring emergence via nonlinear granger causality. *Artificial Life: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems XI*, pp. 545–553.

Seth, A. (2010). A MATLAB toolbox for granger causal connectivity analysi. *Journal of Neuroscience Methods* , pp. 262–273.

Seth, A., A. Barrett, and L. Barnett (2011). Causal density and integrated information as measures of conscious level. *Philosophical Transactions of the Royal Society A* , pp. 3748–3767.