The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies

Wolfram Hesse, Eva Möller, Matthias Arnold, Bärbel Schack *

Institute of Medical Statistics, Computer Sciences and Documentation, Friedrich Schiller University of Jena, Jahnstr. 3, D-07740 Jena, Germany

Received 14 October 2002; received in revised form 25 November 2002; accepted 9 December 2002

Abstract

Understanding of brain functioning requires the investigation of activated cortical networks, in particular the detection of interactions between different cortical sites. Commonly, coherence and correlation are used to describe interrelations between EEG signals. However, on this basis, no statements on causality or the direction of their interrelations are possible. Causality between two signals may be expressed in terms of upgrading the predictability of one signal by the knowledge of the immediate past of the other signal. The best-established approach in this context is the so-called Granger causality. The classical estimation of Granger causality requires the stationarity of the signals. In this way, transient pathways of information transfer stay hidden. The study presents an adaptive estimation of Granger causality. Simulations demonstrate the usefulness of the time-variant Granger causality for detecting dynamic causal relations within time intervals of less than 100 ms. The time-variant Granger causality is applied to EEG data from the Stroop task. It was shown that conflict situations generate dense webs of interactions directed from posterior to anterior cortical sites. The web of directed interactions occurs mainly 400 ms after the stimulus onset and lasts up to the end of the task.

© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Granger causality; EEG; Connectivity; Directed interactions; Stroop task; Multivariate autoregressive models; Recursive least square procedure

1. Introduction

In cognitive operations large cortical networks are engaged, the investigation of which is usually carried out by correlation and coherence measures (cf. Bressler and Kelso, 2001) or phase synchronisation measures (cf. Varela et al., 2001) of the EEG data. These measures tell about the strengths of interactions between groups of neurons. But, these measures provide no insight into the directionality of information flow. Several recent works based on the structural analysis of signal deal with this problem. Causal relations between different components of a multi-dimensional signal can be analysed in the context of multivariate autoregressive modelling. The parameter matrices and transfer matrix of fitted vector autoregressive (VAR) models are mostly not symmetrical and thus are suitable for the detection of the direction of information transfer. The best-established approach in this context is the so-called Granger causality (Granger, 1969). The general idea of causality may be expressed in terms of predictability. If a signal $X$ causes a signal $Y$, the knowledge of the past of both $X$ and $Y$ should improve the prediction of the presence of $Y$ in comparison with the knowledge of the past of $Y$ alone. Within this concept of predictability, Geweke (1982) developed measures of linear feedback and proposed their decomposition by frequency. Bernasconi and König (1999) and Bernasconi et al. (2000) used these measures for the investigation of directional interactions between different areas of the cat visual system. Several further measures for the description of directed information transfer within the frequency
domain were developed currently. Baccala et al. (1998) and Sameshima and Baccala (1999) introduced directed coherence and partial directed coherence and applied them to the analysis of electrophysiological signals. Kaminski and Blinowska (1991) proposed the so-called directed transfer function (DTF). Kaminski et al. (2001) compared the properties of Granger causality and DTF. All these measures are based on the transfer matrix of a fitted VAR model and presuppose the stationarity of the signals in the time interval to be investigated. Ding et al. (2000) and Liang et al. (2000) developed the fit of VAR models for short-time windows using the information of multiple trials. This approach is used for constructing a short-time directed transfer function (STDTF) with a time resolution up to 100 ms. Ginter et al. (2001) used the STDTF for investigating short-time changes in the direction and spectral content of the propagation of EEG activity. Freiwald et al. (1999) generalised the Granger causality by applying local linear autoregressive models. In this study, the generalised Granger causality was used to detect both linear and nonlinear directed interactions between neural groups in the macaque inferotemporal cortex.

The present paper focuses on the recursive time-variant estimation of the Granger causality. Our approach overcomes the requirement of stationarity of the signals and thus permits the observation of transient directed neural networks. Ding et al. (2000) used a short-time windows technique, which require the stationarity of the signal within short-time windows only, and also enables the construction of a time-variant Granger causality. Our method is based on the adaptive recursive fit of a VAR model with time-dependent parameters by means of a generalised recursive least-square (RLS) algorithm, which also takes into consideration a set of EEG epochs as a whole. In contrast to short-window techniques, the multi-trial RLS algorithm involves the information of the actual past of the signal, whereby the influence of the past decreases exponentially with the time distance to the actual samples. Thus, adaptive filter algorithms enable the fit of VAR models with an arbitrary order. Properties of the time-variant Granger causality are demonstrated for simulated signals.

The applicability of the time-variant Granger causality is demonstrated on data from the Stroop task. In the standard color-word Stroop task, subjects must identify the color in which a word is written, while inhibiting the more automatic response of reading the word. The need for attentional selection is high in the so-called ‘incongruent condition’ in which the meaning of the word conflicts with the color in which it is written. This task is commonly employed in studies of selective attention and has been found to be sensitive to damages in prefrontal regions (see, e.g. Vendrell et al., 1995). For incongruent conditions, PET and fMRI studies have shown an increased activation of a widespread network of anterior brain regions. Most studies report on activation of the anterior cingulate cortex (ACC) and the frontal polar cortex (see, e.g. Pardo et al., 1990; Bench et al., 1993; Taylor et al., 1994; Carter et al., 1995, 2000; Banich et al., 2000; MacDonald et al., 2000). MacDonald et al. (2000) succeeded in discriminating the role of different areas of the frontal cortex in a network serving cognitive control. They found that the dorso-lateral prefrontal cortex (DLPFC) was selectively engaged during the preparatory period for color naming in particular and assigned the DLPFC a role in implementation of control. On the other hand, ACC was found to be selectively activated during the response period, more for incongruent color-naming tasks. This later activation speaks for the role of ACC in conflict monitoring. Several authors hint at changes of the regional cerebral blood flow (rCBF) in posterior cingulate and other posterior regions (see, e.g. Bench et al., 1993; Carter et al., 1995). In an EEG study, West and Bell (1997) have shown not only increased Alphal power (8–10 Hz) for medial (F3, F4) and lateral (F7, F8) frontal sites, but also for parietal regions (P3, P4). They suppose that greater activation of the parietal cortex may have resulted from the interaction between prefrontal and parietal regions during the suppression of the influence of the irrelevant word dimension of the stimulus. Ilan and Polich (1999) found in another EEG study that P300 latency did not vary across color/word congruence conditions. They reasoned that the reaction time difference between congruence conditions is originated after stimulus evaluation by the translation of stimulus code into response code. In our own previous EEG coherence study (Schack et al., 1999a, b), the authors ascertained increased fronto-parietal interactions beside increased coherences within the frontal area in the time interval from 400 ms up to the end of the task.

Based on our time-variant Granger causality approach, new aspects of the Stroop effect are detected pertaining to the role of the frontal areas of the cortex and the temporal interactions with posterior areas. The evaluation of the existence of temporal directed interdependencies is performed by the construction of temporal thresholds using surrogate data.

2. Material and methods

2.1. Experiment

Subjects. Ten healthy right-handed male volunteers (aged 20–30) participated in the study. Each subject was free from neurological or psychiatric disorders and had a normal EEG. None of the subjects was familiar with the aims of the research work.

Task. The subject sits in front of a 17” computer screen, where color words are presented written in
different colors. The task-order of ‘congruent’ cases (the word ‘red’ written in red color) and ‘incongruent’ cases (the word ‘red’ written in blue color) was random. Eight different colors were used. The subject has to name the color as quickly as possible. The number of trials was 21 for congruent and 29 for incongruent cases. The answers were continuously recorded with a microphone in order to detect erroneous answers and to register reaction time (RT). Only trials with correct answer were evaluated.

EEG session. The 19-channel EEG was recorded from the scalp by means of a non-polarisable Ag–AgCl electrode cap using the NeuroScan Medical System (international 10-20-system, linked ear-lobes reference, impedance <5 kΩ, sampling frequency at 250 Hz, bandpass of 0.5–45 Hz, 16-bit resolution). Simultaneously, the electrooculogram (EOG) was recorded on an additional channel. EOG artifacts were marked by a correlation procedure between the EOG and the EEG, and selected subsequently by a visual control. Only trials without artifacts were included in the further data processing. EEG epochs of 3072 ms duration, 1024 ms before and 2048 ms after the stimulus were prepared for the analysis.

2.2. Methods

2.2.1. Granger causality

The Granger causality is a fundamental tool for the description of causal interactions of two signals. Let \( X = \{x(n)\} \) and \( Y = \{y(n)\} \) be the two components of a multi-dimensional signal. In order to show the improvement of the prediction of one signal by taking into consideration the past of the second signal, univariate and bivariate AR models are fitted to the signals. In the case of univariate AR modelling

\[
x(n) = \sum_{k=1}^{p} a_{1k} x(n-k) + u_1(n),
\]

the prediction error depends only on the past of the own signal. For bivariate AR modelling

\[
x(n) = \sum_{k=1}^{p} a_{1k} x(n-k) + \sum_{k=1}^{p} c_{2k} x(n-k) + u_2(n),
\]

\[
y(n) = \sum_{k=1}^{p} b_{1k} y(n-k) + \sum_{k=1}^{p} d_{1k} x(n-k) + v_2(n),
\]

the prediction of a signal is based on the past of the own signal and additionally on the past of the second signal. In both cases, the accuracy of prediction may be expressed by the variance of the prediction errors for one-dimensional modelling

\[
\Sigma_{x|X} = \text{var}(u_1),
\]

\[
\Sigma_{y|Y} = \text{var}(v_1),
\]

and, respectively, for two-dimensional modelling

\[
\Sigma_{x|Y} = \text{var}(u_2),
\]

\[
\Sigma_{y|X} = \text{var}(v_2).
\]

If the signal \( Y \) causes the signal \( X \), the variance of the prediction error decreases for two-dimensional modelling, where the past of signal \( Y \) is taken into account for the prediction of \( X \). The Granger causality of \( Y \) to \( X \) as a measure of linear feedback between two signals (cf. Geweke, 1982) is defined by

\[
F_{Y \rightarrow X} = \ln \frac{\Sigma_{Y}}{\Sigma_{Y|X}}.
\]

Correspondingly, the Granger causality of \( X \) to \( Y \) is defined by

\[
F_{X \rightarrow Y} = \ln \frac{\Sigma_{X}}{\Sigma_{X|Y}}.
\]

The maximum of both terms

\[
F_{XY} = \max\{F_{Y \rightarrow X}, F_{X \rightarrow Y}\}
\]

represents a simple measure for the strength of directional and/or bi-directional interaction. Usually, Granger causality is estimated by the fit of VAR models. Several commonly used methods for the estimation of the correspondent model parameters are based on the so-called Yule-Walker equations. These approaches require the stationarity of the signals and result in time-invariant VAR models within the time interval to be analysed.

2.2.2. Time-variant Granger causality

The detection of short-time directed interactions requires the time-variant fit of VAR models. Ding et al. (2000) developed a short-time windows algorithm of VAR model fitting. A time-continuous fit of a VAR model is possible on the basis of adaptive filtering procedures as, e.g. Kalman filtering (Arnold et al., 1998), least mean squares (LMS) approach (Schack, 1999) or RLS approach. In Möller et al. (2001), the traditional RLS algorithm for single trials was generalised to a multi-trial procedure. Similarly to the method by Ding et al. (2000), the latter approach allows the simultaneous fit of one mean VAR model to a set of single trials, each of them representing the measurement of the same task. The inclusion of the whole ensemble of trials in the estimation procedure improves the accuracy of model fitting and allows the investigation of transient interaction processes. Details of the generalised RLS approach may be found in Möller et al. (2001). (For the interested reader, the generalised RLS algorithm was attached in Appendix A.) The time-variant VAR model fitting yields time-variant autoregressive parameters.
Consequently, Eqs. (1) and (2) are modified as follows:

\[ x(n) = \sum_{k=1}^{p} a_{ik}(n) x(n - k) + u_{1}(n), \]

\[ y(n) = \sum_{k=1}^{p} b_{1k}(n) y(n - k) + v_{1}(n), \]  

(8)

and

\[ x(n) = \sum_{k=1}^{p} a_{2k}(n) x(n - k) + \sum_{k=1}^{p} c_{2k}(n) y(n - k) + u_{2}(n), \]

\[ y(n) = \sum_{k=1}^{p} b_{2k}(n) y(n - k) + \sum_{k=1}^{p} d_{2k}(n) x(n - k) + v_{2}(n). \]  

(9)

The time-variant fit of VAR models leads to time-variant prediction errors. A general recursive computation according to

\[ \sigma^2(n + 1) = (1 - c)\sigma^2(n) + c r^2(n) \]  

(10)

with \(0 < c < 1\) yields for \(z(n) = u_1(n), v_1(n), u_2(n)\) and \(v_2(n)\) time-variant variances of the correspondent prediction errors:

\[ \Sigma_{X|X^{-}}(n), \quad \Sigma_{Y|Y^{-}}(n), \quad \Sigma_{X|X^{-},Y^{-}}(n) \quad \text{and} \quad \Sigma_{Y|Y^{-},X^{-}}(n). \]

(11)

Therefore, the calculation of momentary Granger causalities

\[ F_{Y \rightarrow X}(n) = \ln \frac{\Sigma_{X|X^{-}}(n)}{\Sigma_{X|X^{-},Y^{-}}(n)}, \]

\[ F_{X \rightarrow Y}(n) = \ln \frac{\Sigma_{Y|Y^{-}}(n)}{\Sigma_{Y|Y^{-},X^{-}}(n)} \]

(12)

analogously to Eqs. (5) and (6) is possible. The time-variant strength of interaction without the observance of direction may be quantified by the maximum at each time point of both directed Granger causalities:

\[ F_{XY}(n) = \max\{F_{Y \rightarrow X}(n), F_{X \rightarrow Y}(n)\}. \]

(13)

For EEG data analysis, VAR models were fitted individually for each subject according to the multi-trial algorithm (Appendix A) with \(c = 0.03\). The response involved vocalisation into a microphone. In order to avoid artifacts by EMG potentials generated by the speech muscles, individual time intervals from stimulus offset up to the mean individual reaction time were included in further analysis. For statistical comparisons, the time courses of Granger causality were interpolated to the mean reaction time of all subjects, separately for the congruent and the incongruent tasks.

2.2.3. Choice of the model order

For modelling signals by VAR models, the number of model parameters can be determined using the general AIC criterion

\[ \Phi_{M,N}(p) = N \ln[\det(\Sigma_{N}(p))] + 2pM^2, \]

(14)

where \(M\) denotes the dimension of the model and \(\Sigma_{N}(p)\) the variance of the prediction error for the model order \(p\). In the case of \(K\) trials of length \(n\), \(N = Kn\) is used. In practice, the maximal order \(p_0\) for all channel pairs and task conditions is fixed for bivariate modelling. In order to have the same number of parameters to be estimated, the order for univariate modelling was chosen \(2p_0\). This relation of order choices for the bivariate and univariate model fit seems to be suitable for fast changing signals, where the number of sample points for estimating momentary Granger causalities is strongly limited.

For EEG data, the order \(p_0\) was determined in the following way. For each subject and all channel pairs, model orders were calculated on the basis of Yule-Walker equations by the minimum of the AIC criterion (14) in the range \(1 \leq p \leq 30\). 80% of these orders were smaller than 14. Fig. 1 shows a typical example of the AIC criterion.

The model fit was performed for the time interval beginning with the stimulus onset up to the individual responses. The order \(p_0 = 13\) was fixed for all the following investigations of the EEG data.

2.2.4. Simulations

The following three simulation examples show the principal properties of the time-variant Granger causality.

2.2.4.1. The reflection of causal directed interaction

Generally, the most important property of the Granger causality is its positivity in the case, when a signal \(X\) causes a second signal \(Y\). The following first example

![Fig. 1. Typical example of the dependence of the Akaike information criterion according to (14) for one subject (electrode pair: Fz/T6, number of trials: 16).](image-url)
demonstrates the behaviour of the time-variant Granger causality in comparison with the time-invariant Granger causality in this context. Let \( X \) be an AR(2) process and \( Y \) the superposition of \( X \) with random noise:

\[
X = \text{AR}(2)
\]

with AR parameters \( a_1 = -0.07 \), \( a_2 = 0.32 \) and the variance of noise \( \text{var}(\epsilon) = 0.81 \) and

\[
Y = X + \xi, \quad \text{with} \quad \xi \sim N(0,1) \quad \text{and} \quad \text{cov}(\xi, \epsilon) = 0.
\]

Obviously, \( X \) causes \( Y \), but \( Y \) does not cause \( X \). Hundred realisations of \( X \) and \( Y \) each with 768 sample points were simulated. Fig. 2 shows the correspondent Granger causalities (\( p_0 = 2 \)).

The time-variant Granger causality \( F_{X \rightarrow Y}(n) \), showing the influence from \( X \) to \( Y \), is the whole time positive, and after a short adaptation period fluctuates around the time-invariant Granger causality \( F_{X \rightarrow Y} \). As expected, the Granger causality \( F_{Y \rightarrow X}(n) \) varies around the value zero indicating no influence from \( Y \) to \( X \).

### 2.2.4.2. Time-variant detection of directed causal interactions

The more interesting question is the ability of the time-variant Granger causality to react on changes in the directed influences between two signals. Therefore, 100 realisations of a two-dimensional AR(2) process with changes in dependencies are simulated according to Eq. (9) in the following way:

\[
a_{21}(n) = \begin{cases} 
-0.6, & 1 \leq n < 400, \\
0.3, & 400 \leq n \leq 768,
\end{cases}
\]

\[
b_{21}(n) = \begin{cases} 
0.3, & 1 \leq n < 400, \\
-0.6, & 400 \leq n \leq 768,
\end{cases}
\]

\[
a_{22}(n) = 0.1, \quad 1 \leq n \leq 768,
\]

\[
b_{22}(n) = 0.1, \quad 1 \leq n \leq 768,
\]

and

\[
\text{var}(u_x) = \begin{cases} 
0.81, & 1 \leq n < 400, \\
2.25, & 400 \leq n \leq 768,
\end{cases}
\]

\[
\text{var}(v_y) = \begin{cases} 
2.25, & 1 \leq n < 400, \\
0.81, & 400 \leq n \leq 768.
\end{cases}
\]

This means, for the first 200 sample points signal \( Y \) causes signal \( X \) and beginning with the sample point 400 signal \( X \) causes signal \( Y \). From sample point 201 up to sample point 399, no dependence exists between the two signals \( X \) and \( Y \). The time-variant Granger causalities (\( p_0 = 2 \)) are illustrated in Fig. 3.

The time-limited influence from \( X \) to \( Y \) beginning with the sample point 400 is detected by the positivity of \( F_{X \rightarrow Y}(n) \) (thick gray curve). The time-limited influence from \( Y \) to \( X \) for the first 200 sample points is identified by the positivity of \( F_{Y \rightarrow X}(n) \) (thick black curve). Both Granger causalities are nearly zero within the time interval \( 200 < n < 400 \) without any dependence between the two signal components. Further, time-dependent Granger causalities vary around the estimations of correspondent time-independent Granger causality (thin black and gray curves) within the stationary time intervals \( 0 < n \leq 200 \) and \( 400 \leq n \leq 768 \). The time behaviour of time-variant Granger causality for this example demonstrates the ability to react on changes in directed dependencies between two signals.

#### 2.2.4.3. Dependence of estimation accuracy on the number of trials

The estimation accuracy of the Granger causality improves with an increasing number of trials. The statistical properties of the multi-trial fit of an AR model on the basis of the generalised RLS algorithm were investigated in detail in Möller et al. (2001). The influence of the number of trials on the estimation of time-variant Granger causality is demonstrated in the following example. 20 trials of the following bivariate AR(2) with 768 samples were simulated:

\[
c_{21}(n) = \begin{cases} 
0.2, & 1 \leq n \leq 200, \\
0, & 200 < n \leq 768,
\end{cases}
\]

\[
d_{21}(n) = \begin{cases} 
0, & 1 \leq n < 400, \\
0.2, & 400 \leq n \leq 768.
\end{cases}
\]

\[
c_{22}(n) = \begin{cases} 
0.1, & 1 \leq n \leq 200, \\
0, & 200 < n \leq 768,
\end{cases}
\]

\[
d_{22}(n) = \begin{cases} 
0, & 1 \leq n < 400, \\
0.1, & 400 \leq n \leq 768,
\end{cases}
\]

and

\[
\text{var}(u_x) = 0.81,
\]

\[
\text{var}(v_y) = 2.25.
\]

This is the same AR(2) model as in the previous example for the first 200 samples. Fig. 4 shows the estimations of Granger causalities (\( p_0 = 2 \)) for a single trial, for 5 trials, for 10 trials and for 20 trials.
The thin gray constant line shows the estimated time-invariant Granger causality noise variance decreases with the number of trials and the variance of noise \(\text{var}(\varepsilon) = 0.81\) and \(Y = X + \xi\), with \(\xi \sim N(0,1)\) and \(\text{cov}(\xi, \varepsilon) = 0\). The thin gray constant line shows the estimated time-invariant Granger causality \(F_{X \rightarrow Y}\).

Obviously, the adaptively estimated Granger causality varies around the time-invariant estimation. Thereby, the variance decreases with the number of trials and the difference between the 10-trial estimation and 20-trial estimation is negligible. For EEG data, the number of trials varied between 13 and 19 in the congruent case and between 17 and 23 in the incongruent case.

2.2.5. Statistical evaluation of causal interactions by constructing a time-variant threshold for surrogate data

The distribution of the time-variant Granger causalities, calculated according to Eq. (12) of a fitted time-variant AR process, is unknown. In order to decide about the appearance of a causal influence between two signals, a time-variant threshold was constructed, representing the level of Granger causality above which values had less than a 5% probability of occurring by chance. For this purpose, the following surrogate data approach was used.

Let \(X\) and \(Y\) be two components of the 19-channel EEG. The signal \(X\) causes the signal \(Y\), if the knowledge of the past of \(X\) increases the prediction of \(Y\). This kind of causality is destroyed, when the \(Y\)-data are ordered randomly in time. For this purpose, the order of the data of the \(Y\) signal was randomised. This shuffle procedure saves the distributional properties of the \(Y\) signal, but destroys the regressive relationships within \(Y\) and between \(Y\) and \(X\). Shuffling was repeated 200 times. Afterwards, the 95% quantile was determined for each time point. Fig. 5 shows an example of the time courses of a Granger causality and its correspondent threshold.

The thick black curve represents the Granger causality \(F_{X \rightarrow Y}(n)\), where \(X\) is the EEG signal at electrode position T6 and \(Y\) is the EEG signal at electrode position F7 during the incongruent situation of the Stroop task for one subject. The thin black curve is the time course of the correspondent threshold, calculated with the original data at T6 and the shuffled data at F7. A directed causal influence from T6 to F7 is stated in the time inter vals, where the Granger causality exceeds the threshold and thus has less than 5% probability of occurring by chance. The set of gray curves represents the thresholds, calculated with the original data at T6 and the shuffled data at all other 17 electrode positions, which are different from T6 and F7. Obviously, the

\[ W. Hesse et al. / Journal of Neuroscience Methods 124 (2003) 27–44 \]
The reaction time was 711.2 ms. The data of T6 and the shuffled data from all other electrode positions were included in the analysis.

The subject in the incongruent situation of the Stroop task (16 trials) showed a reaction time of 672.1 ms. The EEG signal at electrode position F7 and O2 during the congruent situation of the Stroop task for one subject. The thin black and gray curves represent the thresholds calculated from the original data of T6 and the shuffled data from all other electrode positions apart from F7. The reaction time was 711.2 ms.

The example demonstrates the possible different levels of the thresholds for both directed interdependences and the different time courses of the Granger causalities $F_{X \rightarrow Y}(n)$ and $F_{Y \rightarrow X}(n)$. In this example, EEG at F7 causes the EEG at O2 in the time interval 500–700 ms, and the EEG at O2 causes the EEG at F7 in the interval 600–700 ms. Thus, there is a directed interaction between 550 and 600 ms and a bi-directional interaction between 600 and 700 ms.

3. Results

3.1. Behavioral data

The Stroop effect is confirmed by a longer reaction time (RT) for the incongruent situation. The mean RTs of the 10 subjects were, for the incongruent situation 793 ms (109 ms S.D.) and for the congruent situation 708 ms (114 ms S.D.). The difference of 85 ms in RT is significant (paired samples t-test at the significance level of 1%).

3.2. Granger causality

Individual time-variant Granger causalities in both directions and their maximum were calculated according to Eqs. (8)–(13) with $p_0 = 13$ for all 171 possible electrode pairs. The difference in strength of interactions for an electrode pair between the congruent and the incongruent situation is evaluated by the maximum (Eq. (13)) of both directed Granger causalities. A directed influence for a determined time interval is stated, if $F_{X \rightarrow Y}(n)$ is significantly larger than $F_{Y \rightarrow X}(n)$ in the time interval or vice versa. Fig. 7 exemplarily illustrates a typical result for the electrode pair Fp1/F3.

The upper panel A shows the mean (across all subjects) maximum of Granger causalities $F_{XY}(n) = \max\{F_{Y \rightarrow X}(n), F_{X \rightarrow Y}(n)\}$ for the congruent (black curve) and the incongruent (gray curve) situations. The strength of interaction is larger in the incongruent task beginning with 320 ms up to the end of the task. The strength of interaction differs significantly within the time interval 380–520 ms. The lower two panels show the directed interactions $F_{F3\rightarrow Fp1}(n)$ (black curves) and $F_{Fp1\rightarrow F3}(n)$ (gray curves). There is no significant difference in directedness for the congruent situation (middle panel). In opposite, the Granger causality $F_{F3\rightarrow Fp1}(n)$ is significantly larger in the time interval 200–620 ms than $F_{Fp1\rightarrow F3}(n)$ for the incongruent situation (lower panel). Thus, a superior influence from F3 to Fp1 is present within this time interval.

Whilst Fig. 7 demonstrates the directed information flow within the frontal area, the following example points at large-scale interaction.
For the chosen electrode pair F8/O2, the interaction is stronger in the congruent case (panel A). In this case, the interaction is directed from right frontal area (F8) to right occipital area (O2) beginning at 180 ms after the onset up to the mean end of the congruent task (panel B). During the incongruent task, the same superior directed interaction exists only in the time interval 240–330 ms (panel C).

In order to get a survey of the differences between congruent and incongruent tasks in strength of interactions, the time courses of the maximal Granger causalities were calculated and statistically compared for each time point (Wilcoxon test for paired samples, significance level 5%). For each time point, the number of electrode pairs with different interaction strengths (≤171) was determined. Time intervals with a high number of differences hint at time episodes being relevant for the Stroop effect.

The number of significant differences in strength of interactions ($F_{XY}^{\text{incongruent}} \neq F_{XY}^{\text{congruent}}$) at each time point is illustrated by the thick gray curve in both diagrams A and B of Fig. 9. Important time intervals appear at 120–240 ms with stronger interactions in the incongruent case (black curve in diagram A), at 360–408 ms with...
stronger interactions in the congruent case (black curve in diagram B) and within a later time segment between 440 and 520 ms, again with stronger interactions in the incongruent case (black curve in diagram A). In the following, undirected interactions during these time intervals are described in detail.

At first, the strength of interaction was compared between congruent and incongruent situations by means of maximum of Granger causalities \( F_{XY}(n) = \max \{F_{YX}(n), F_{XY}(n)\} \). Fig. 10 shows the topographies of the statistical results for the selected intervals.

In the case of significant differences within the selected time interval, the correspondent electrode pairs were linked by solid lines in the incongruent case and by dashed lines in the congruent case. For the time intervals 120–240 and 440–520 ms with stronger interactions for the incongruent case, stable connections appear within the frontal areas and between frontal and posterior areas mainly in the left hemisphere for the incongruent task, whereas stronger interactions for the congruent case appear within the right posterior area. For the interval 360–408 ms, strong interactions between temporal regions of both sides are found for the congruent situation. Whilst the duration of interactions is in a similar range for congruent and incongruent tasks during the early time interval 120–240 ms, they differ extremely for the later time intervals (see Table 1). Interactions with frontopolar sites are remarkably stable in time during incongruent tasks for the time interval 440–520 ms (see Table 1).

### Table 1
Electrode pairs with longest in time significant differences in strength of interactions (maximum of Granger causalities) for selected time intervals

<table>
<thead>
<tr>
<th>Electrode pair</th>
<th>Duration (ms)</th>
<th>Electrode pair</th>
<th>Duration (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differences in strength of interaction between congruent and incongruent tasks time interval: 120–240 ms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3/C3</td>
<td>80</td>
<td>F4/O2</td>
<td>76</td>
</tr>
<tr>
<td>F3/F4</td>
<td>56</td>
<td>F4/C4</td>
<td>72</td>
</tr>
<tr>
<td>T6/O1</td>
<td>56</td>
<td>Fp1/O2</td>
<td>44</td>
</tr>
<tr>
<td>T5/O2</td>
<td>52</td>
<td>C4/T4</td>
<td>28</td>
</tr>
<tr>
<td>F3/Cz</td>
<td>48</td>
<td>Fp2/T4</td>
<td>24</td>
</tr>
<tr>
<td>T5/O1</td>
<td>48</td>
<td>F3/O2</td>
<td>24</td>
</tr>
<tr>
<td>F3/T5</td>
<td>36</td>
<td>Fz/O2</td>
<td>24</td>
</tr>
</tbody>
</table>

| Differences in strength of interaction between congruent and incongruent tasks time interval: 360–408 ms | | | |
| Fp1/F3         | 24           | P3/T6          | 48           |
| F7/F3          | 16           | T6/O1          | 48           |
| T3/C3          | 44           | C3/T6          | 40           |
| C3/T5          | 40           | F8/T5          | 36           |
| C3/T6          | 36           | T3/T5          | 36           |
| T3/O2          | 32           | T6/O2          | 36           |

| Differences in strength of interaction between congruent and incongruent tasks time interval: 440–520 ms | | | |
| Fp1/F3         | 80           | F8/T5          | 20           |
| F7/Fz          | 71           | T4/O2          | 20           |
| Fp1/Cz         | 68           | F8/P3          | 12           |
| Fp1/F7         | 56           | Fp1/Fz         | 52           |
| Fp1/Hz         | 52           | Fp2/F7         | 52           |
| Fp1/C3         | 48           | Fp2/F8         | 36           |
| Fz/C4          | 32           |                  |              |

Statistical tests (Wilcoxon test with paired samples: for eight of 10 subjects yielded \( P < 0.05 \)) were performed for each time point. The columns ‘duration’ denote the duration of significant differences within the correspondent time intervals. In each case, electrode pairs with longest duration of significant differences are listed.

### Table 2
Electrode pairs with longest in time significant differences in strength of directed interactions for selected time intervals

<table>
<thead>
<tr>
<th>Electrode pair</th>
<th>Duration (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differences in strength of directed interaction for the time interval: 120–240 ms; ( F_{XY}(\text{congruent}) &gt; F_{XY}(\text{incongruent}) )</td>
<td></td>
</tr>
<tr>
<td>T4 → F8</td>
<td>88</td>
</tr>
<tr>
<td>F4 → Fp1</td>
<td>60</td>
</tr>
<tr>
<td>O1 → T5</td>
<td>56</td>
</tr>
<tr>
<td>F4 → Fp1</td>
<td>52</td>
</tr>
<tr>
<td>Cz → F8</td>
<td>28</td>
</tr>
<tr>
<td>Fz → F3</td>
<td>24</td>
</tr>
<tr>
<td>F3 → F8</td>
<td>20</td>
</tr>
<tr>
<td>T3 → T5</td>
<td>20</td>
</tr>
<tr>
<td>F4 → Fp2</td>
<td>20</td>
</tr>
<tr>
<td>F4 → F3</td>
<td>20</td>
</tr>
<tr>
<td>O2 → T6</td>
<td>20</td>
</tr>
</tbody>
</table>

| Differences in strength of directed interaction for the time interval: 360–408 ms; \( F_{XY}(\text{congruent}) > F_{XY}(\text{incongruent}) \) | | |
| T3 → T5        | 28           |
| T6 → P3        | 28           |
| Pz → O1        | 20           |
| T6 → O1        | 20           |
| T5 → P3        | 16           |
| T6 → O2        | 12           |

| Differences in strength of directed interaction for the time interval: 440–520 ms; \( F_{XY}(\text{congruent}) > F_{XY}(\text{incongruent}) \) | | |
| Fz → Fp1       | 52           |
| Cz → Fp2       | 52           |
| Cz → Fp1       | 48           |
| F4 → Fp1       | 44           |
| T5 → Cz        | 32           |
| F3 → Fp1       | 28           |
| F4 → Fp2       | 24           |
| T5 → P4        | 20           |

Only electrode pairs with values different from zero Granger causality (exceeding of 95% quantile threshold) are taken into consideration (McNemar test with \( P < 0.05 \)). The columns ‘duration’ denote the duration of significant differences within the correspondent time intervals. In each case, electrode pairs with longest duration of significant differences are listed.
Maximum of Granger causalities evaluates the strength of interaction only. In a second step directed interactions were compared between task situations. Thereby, only electrode pairs were taken into consideration with at least one Granger causality exceeding the 95% threshold for both tasks. In Fig. 11, those causal relationships are shown, which exceeded the calculated threshold and differ significantly for both situations (McNemar test, with \( P < 0.05 \)).

At each case of the selected time intervals, the dominant task with regard to the number of stronger interactions (cf. Fig. 9) is presented: directed network of stronger interactions with regard to the incongruent task for the time intervals 120–240 and 440–520 ms, and stronger interactions with regard to the congruent task for the time interval 360–408 ms. Stronger and longer in duration interactions directed towards frontopolar sites appear in the incongruent case, whereas stronger causal relationships are directed to posterior sites in the congruent situation. Further, the duration of directed influences is larger for the incongruent situation (see Table 2).

Up to now, Granger causalities were observed in the context of discovering differences between both situations. In a last step, the directions of interaction were compared separately for both tasks. Therefore, Granger causalities \( F_{Y \rightarrow X}(n) \) and \( F_{X \rightarrow Y}(n) \) were compared for each electrode pair (\( X, Y \)) and each time point. The number of electrode pairs with one superior direction of interaction (Wilcoxon test for paired samples, \( P < 0.05 \)) were counted for each time point both for congruent and incongruent cases. The diagram in Fig. 12 illustrates the absolute frequencies of Granger causalities with superior direction for each sample point for both tasks.

It can be seen that the number of directed interactions \( (F_{Y \rightarrow X}(n) \neq F_{X \rightarrow Y}(n)) \) is slightly increased within the time interval 120–240 ms for both situations and extremely increased at the end of the task for incongruent situation only.

Fig. 13 shows the networks of interactions with superior direction for congruent (A) and incongruent (B) situations in the early time interval 120–240 ms, marked in the diagram in Fig. 12 by the first dual arrow. The two networks differ only little in their topographies. But, in the mean, directed interactions are more stable in time for the incongruent situation (see Table 3). Electrode pairs with directed interactions, lasting longer than the half of the chosen time interval are listed in Table 3. Again, long in time interactions with superior direction towards frontopolar sites occur.

Because of the extremely high number of directed interactions, the late time interval 480–716 ms, marked in the diagram in Fig. 12 by the second dual arrow, is of special interest. Granger causalities with superior direction of at least 80 ms duration are shown in Fig. 14.

Whilst only a few long-lasting directed interactions occur for the congruent task (A), a very dense network of interactions with superior direction (B) appears for the incongruent situation. Map (B) in Fig. 14 shows this network of long-lasting interactions supplemented by directed connections with at least 30 ms duration for the short-time interval 600–660 ms. There are two general directions of long-lasting interactions for the incongruent task: from posterior to frontal sites and from left to right hemisphere. Within the time interval 600–660 ms, additional causal interactions appear, which are directed mainly to frontopolar sites.

### Table 3
Electrode pairs with longest in time significant differences (Wilcoxon test with paired samples) of interactions with respect to their directions within the time interval 120–240 ms for both congruent and incongruent situations

<table>
<thead>
<tr>
<th>Electrode pair</th>
<th>Duration (ms)</th>
<th>Electrode pair</th>
<th>Duration (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F4 → F8</td>
<td>120</td>
<td>Fz → T4</td>
<td>116</td>
</tr>
<tr>
<td>F4 → Fp2</td>
<td>120</td>
<td>Fz → F4</td>
<td>112</td>
</tr>
<tr>
<td>Fz → Fp1</td>
<td>116</td>
<td>Cz → T4</td>
<td>108</td>
</tr>
<tr>
<td>T6 → C4</td>
<td>116</td>
<td>F4 → T4</td>
<td>72</td>
</tr>
<tr>
<td>Fz → Fp2</td>
<td>112</td>
<td>P3 → F4</td>
<td>68</td>
</tr>
<tr>
<td>P4 → C4</td>
<td>100</td>
<td>Cz → F4</td>
<td>64</td>
</tr>
<tr>
<td>F4 → T4</td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4 → T4</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3 → Fz</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4 → F8</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4 → Fp1</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4 → Fp2</td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fz → F3</td>
<td>64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3 → C4</td>
<td>64</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Statistical tests were performed for each time point. The columns ‘duration’ denote the duration of significant differences within the time interval 120–240 ms. At each case, electrode pairs with longest duration of significant differences are listed.

### 4. Discussion

This study illuminates two essential aspects. First, the presented approach for time-variant estimation of Granger causality permits the detection of temporal causal interactions (see Figs. 5 and 6). Second, temporally directed interactions were detected successfully for electrophysiological data of the Stoop task on the basis of adaptive Granger causality (see Figs. 11–14).

The generalisation of the traditional RLS algorithm (Möller et al., 2001) to a multi-trial procedure lead to an estimation of Granger causality with high time resolution. Thereby, the increase of the number of trials enhances the estimation accuracy (cf. Fig. 4). Simulation in Fig. 3 demonstrates the reaction of the momentary Granger causality on changes in causality in a model process.
In Figs. 5–8, negative values of Granger causality occur. In the case of an optimal fit to the true autoregressive parameters of univariate and bivariate models, Granger causality is a non-negative value. In reality negative values appear at times. The occurrence of negative values may be due to three main reasons. Firstly, the RLS algorithm results in a tracking of the autoregressive parameters around their theoretical values. Because of the non-consistency of this estimation, fluctuations around zero and thus negative values of estimated Granger causality are possible. Second, in the case of event-related data some single trials may not be typical for the whole ensemble of trials. In this case, the estimation is not optimal. Third, the order \( p_0 = 13 \) was fixed for all further investigations in order to limit the computational effort. This order may be not optimal for single subjects or single-channel pairs.

The decision about the statistical validity of causal relations require the predetermination of thresholds. Because of unknown statistical properties of the momentary Granger causality, the construction of a threshold is necessary. The time-variant threshold was established by the use of surrogate data. The main condition for the construction of surrogate data is the preservation of the stochastic properties and the destruction of the causal relation between the causal signal and the surrogate. Because of the specific stochastic properties and time structure of the causal signal,
concrete thresholds for denoting the value being less than 5% probability of occurring by chance Granger causality has to be constructed. Thus the first problem of applying time-variant Granger causality to electrophysiological signals was the construction of specific thresholds for each pair of EEG signals (see Fig. 5), and even for both directions of causal interactions (see Fig. 6). Figs. 7 and 8 demonstrate the possibility of the detection of interactions with superior direction between two EEG signals for very short-lasting time intervals. In this respect, the method outplays classical estimations of Granger causality used, e.g. by Bernasconi and König (1999) and by Bernasconi et al. (2000). Beside the investigation of causal interactions using the above-mentioned threshold, a simple statistical comparison of Granger causalities for two task situations is used to find interactions with superior directions (see Figs. 12–14).

In our study, the strength and the direction of interactions was investigated during the Stroop task. The strength of interactions, determined by the maximum of both directed Granger causalities, hint at special time intervals with most differences in congruent and incongruent situations (see Fig. 9). The special role of the frontal and frontopolar cortex in selective attention is here expressed in diverse interactions for the incongruent tasks, in particular in the late time interval (compare map of undirected interactions for the time interval 440–520 ms in Fig. 10). Generally, interactions lasted longer for incongruent tasks in comparison with congruent tasks (Table 1). Thereby, causal interactions are directed to the frontal cortex in contrast to the congruent tasks (compare Fig. 11). Likewise, directed interactions turned out to be more stable in time in the incongruent case (Table 2).

The special comparison of interactions with superior direction \( F_{X\rightarrow Y}(n) \neq F_{Y\rightarrow X}(n) \) shows obvious differences between both tasks. While the number of directed interactions is similar at the beginning of the task (up to 360 ms), the number of directed interactions increases essentially in the late time interval for the incongruent situation (Fig. 12). MacDonald et al. (2000) determined different active frontal areas in the early time interval-activation of the DLPFC indicating the implementation of control and in the late time interval-activation of the ACC indicating conflict monitoring. The networks of directed interactions are similar with regard to their topography for the early time interval (120–240 ms), as
is shown in Fig. 13. Thereby, the network of causal (directed) interactions is more stable in time for incongruent tasks (Table 3).

In contrast, essential differences in directed interactions exist in the late time interval (480–716 ms). Whilst the number of stable (for at least 84 ms) causal interactions is very small in the congruent case, a widespread directed network of interactions exists in the incongruent case (Fig. 14). This observation supports the supposition of Ilan and Polich (1999) about the late time point of the origin in response time difference between both congruence conditions. Further, interactions with superior direction are directed from posterior to anterior sites of the cortex within both hemispheres towards frontal and from left posterior towards right anterior cortical areas (Fig. 14, panel B). Thereby, frontal polar sites are included in this web for the time interval with the highest number of interactions (Fig. 14, panel B). These results are consistent with the postulated special role of the anterior cortex for selective attention (see, e.g. Pardo et al., 1990; Bench et al., 1993; Taylor et al., 1994; Carter et al., 1995, 2000; Banich et al., 2000; MacDonald et al., 2000). On the other hand, the interaction between parietal and frontal sites could be an explanation for the activation of posterior areas (see, e.g. West and Bell, 1997; Bench et al., 1993; Carter et al., 1995). Possibly, the minimisation of the influence of the irrelevant word information originates directed interactions from parietal towards frontal sites. In our previous coherence study, increased coherences between parietal and frontal sites were also observed for the late time interval.

This study demonstrates the possibility of the detection and description of transient directed webs of interactions. Thus, the presented method seems to be suitable for the investigation of causal interactions between electrophysiological signals of different sites of the cortex, occurring during cognitive processes. In
Fig. 11. Topographies of directed interactions for selected time intervals. Causal interactions are shown, which exceed the 95% threshold and are significantly different for congruent and incongruent situations (McNemar test with significance level 5%). Topographies are presented for stronger causal interactions during the incongruent task for the time intervals 120–240 and 440–520 ms, and for stronger causal interactions during the congruent task for the time interval 360–408 ms.

Fig. 12. Absolute frequencies of Granger causalities with significant differences in direction for the congruent situation (thin black curve) and the incongruent situation (thick gray curve). For each electrode pair \((X, Y)\) Granger causalities \(F_{X \rightarrow Y}(n)\) and \(F_{Y \rightarrow X}(n)\) were compared for each time point (Wilcoxon test for paired samples; significance level 5%). \(rt_{con}\) denotes the mean response time of the congruent task and \(rt_{in}\) denotes the mean response time of the incongruent task. The dual arrows mark the time intervals of interest.
addition, this method could also be useful for connectivity analysis of other data, in particular from fMRI.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, project Scha 741/3-1). The authors thank A.C.N. Chen for data recording and Hellmuth Petsche for helpful discussions.

Appendix A

For the description of the RLS algorithm, the following model representation was used (cf. Möller et al., 2001). We considered $J$ trajectories of an $M$-dimensional vector $X_t$.
Let \( y_j^{(m)}(n) \) be the \( m \)-th component \((m = 1, \ldots, M)\) of the \( j \)-th trajectory \((j = 1, \ldots, J)\) at time \( t_n \) \((n = 1, 2, \ldots)\). The entire observation at this time point was composed into matrix \( Y_n \in \mathbb{R}^{J \times M} \):

\[
Y_n = \begin{pmatrix}
  y_1^{(1)}(n) & \cdots & y_1^{(M)}(n) \\
  y_2^{(1)}(n) & \cdots & y_2^{(M)}(n) \\
  \vdots & \ddots & \vdots \\
  y_J^{(1)}(n) & \cdots & y_J^{(M)}(n)
\end{pmatrix}
\]

Further, matrix \( W_n \in \mathbb{R}^{J \times Mp} \), \( W_n = (Y_{n-1}, \ldots, Y_{n-p}) \) collected the last \( p \) observations of the process. The ensemble of trajectories will be fitted by a VAR\((p)\) model with the model parameter \( \Theta = (A_1, \ldots, A_p) \), where \( A_k \in \mathbb{R}^{M \times M} \) is the \( k \)-th parameter matrix \((k = 1, \ldots, p)\). For the adaptive estimation \( \hat{\Theta}_n \) of the model parameter \( \Theta \) the following recursive computation formula:

\[
\hat{\Theta}_0 = 0 \quad \text{(starting values)},
\]

\[
C_0^{(k)} = I_{mp},
\]
C_n^{(0)} = (1 - c)^{-1} C_{n-1}^{(0)} (\text{iteration step: } n = 1, 2, 3, \ldots),
C_n^{(j)} = (1 - c)^{-1} \left( I_{Mp} - \frac{W_n^T(j, \cdot) W_n(j, \cdot) C_n^{(j-1)}}{W_n(j, \cdot) C_n^{(j-1)} W_n^T(j, \cdot) + 1} \right),
(j = 1, \ldots, J),
K_n = W_n C_n^{(j)};
Z_n = Y_n - W_n \hat{\Theta}_n^T Y_{n-1},
\hat{\Theta}_n = \Theta_n - 1 + Z_n^T K_n,

\text{where } I_{Mp} \text{ denotes the identity matrix of dimension } M p \times M p \text{ and } W_n(j, \cdot) \text{ is the } j\text{-th row of matrix } W_n. \text{ The factor } c, 0 \leq c < 1, \text{ is decisive for the adaptation capability of the estimation. More precisely, the adaptation speed and also die variance increase with } c \text{ and vice versa. For a fixed value of } c, \text{ the adaptation speed is the same for every dimension } M. \text{ In addition, the variance is reduced when the number of trajectories } J \text{ increased.}

\text{ Whilst modelling event-related EEG data, satisfactory balance between the adaptation speed and the variance of the estimation was obtained for } 0.01 \leq c \leq 0.05 (\text{cf. Möller et al., 2001}).

The sequence \{Z_n\}_n \text{ defined in the algorithm, is the instantaneous prediction error matrix, which gives the difference between the desired response } Y_n \text{ and the output } W_n \hat{\Theta}_n^T Y_{n-1} \text{ produced by the estimation. Thus, the sequence } \{\Sigma_n\}_n \text{, computed according to the formula}
\Sigma_n = \hat{\Sigma}_n + \frac{c}{J} Z_n^T Z_n, \quad n = 1, 2, \ldots,

\text{represents a time-variant (or an instantaneous) estimation of the variance of the prediction error.}

\text{ For the computation of the Granger causality, two univariate models were fitted, both with } M = 1 \text{ and } Y_n = (x_1(n) \cdots x_k(n))^T \text{ and } Y_n = (y_1(n) \cdots y_k(n))^T, \text{ respectively. Additionally, the bi-variate model with } M = 2 \text{ and}
Y_n = \begin{pmatrix} x_1(n) & y_1(n) \\ x_2(n) & y_2(n) \\ \vdots & \vdots \\ x_k(n) & y_k(n) \end{pmatrix}

\text{ was computed. In both cases, the same order } p \text{ was chosen. From the RLS algorithm, } \hat{\Theta}_n = (a_1(n), \ldots, a_p(n)) \text{ and } \hat{\Theta}_n = (b_1(n), \ldots, b_p(n)) \text{ were obtained for the univariate estimation. Thus, the corresponding instantaneous squared prediction errors were } \hat{\Sigma}_n = \text{var}(u_t(n)) \text{ and } \hat{\Sigma}_n = \text{var}(v_t(n)). \text{ In the bivariate case, the parameter was } \hat{\Theta}_n = (\hat{A}_1(n), \ldots, \hat{A}_p(n)) \text{ with}
\hat{A}_n = \begin{pmatrix} a_1(n) \\ b_1(n) \end{pmatrix}, \quad j = 1, \ldots, p,
\text{ and diag}(\hat{\Sigma}_n) = (\text{var}(u_t(n)), \text{var}(v_t(n))). \text{ The adaptation speed was sufficient for } c = 0.03 \text{ in order to show the dependence of the time-behaviour of the instantaneous Granger causality.}

\text{ References}


