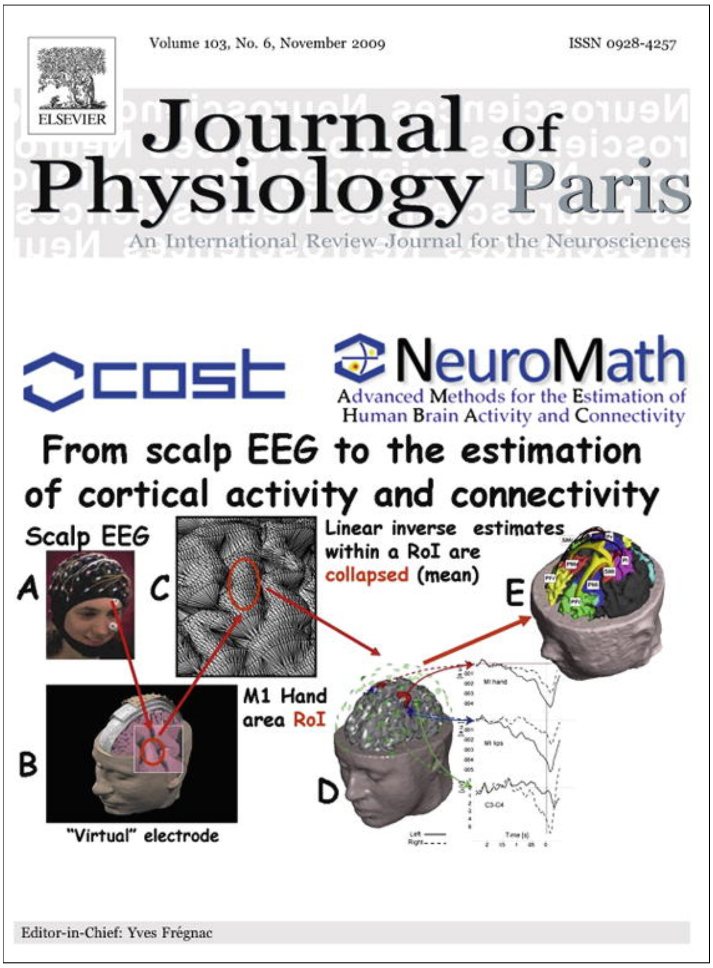
Volume 103, No. 6, November 2009
ISSN 0928-4257
ELSEVIER
Journal of Physiology Paris
An International Review Journal for the Neurosciences
cost
NeuroMath
Advanced Methods for the Estimation of Human Brain Activity and Connectivity
From scalp EEG to the estimation of cortical activity and connectivity
Scalp EEG
A
B
C
D
E
Linear inverse estimates within a RoI are collapsed (mean)
M1 Hand area RoI
"Virtual" electrode
Editor-in-Chief: Yves Frégnac

Journal of Physiology - Paris 103 (2009) 348–352

Contents lists available at ScienceDirect

Journal of Physiology - Paris

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

Time-variant estimation of directed influences during Parkinsonian tremor

Linda Sommerlade [a,c,e,*], Kathrin Henschel [a,c], Johannes Wohlmuth [a], Michael Jachan [a,c,d], Florian Amtage [d], Bernhard Hellwig [d], Carl Hermann Lücking [d], Jens Timmer [a,b,c,e], Björn Schelter [a,c,e]

[a] FDM, Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany
[b] Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Albertstr. 19, 79104 Freiburg, Germany
[c] Bernstein Center for Computational Neuroscience, Hansastr. 9A, 79104 Freiburg, Germany
[d] Neurology Department, University Hospital of Freiburg, Breisacherstr. 64, 79110 Freiburg, Germany
[e] Department of Physics, University of Freiburg, Hermann Herder Str. 3, 79104 Freiburg, Germany

ARTICLE INFO

Keywords:
Multivariate time series
Partial directed coherence
Time-variant estimation

ABSTRACT

The inference of interaction structures in multidimensional time series is a major challenge not only in neuroscience but in many fields of research. To gather information about the connectivity in a network from measured data, several parametric as well as non-parametric approaches have been proposed and widely examined. Today a lot of interest is focused on the evolution of the network connectivity in time which might contain information about ongoing tasks in the brain or possible dynamic dysfunctions. Therefore an extension of the current approaches towards time-resolved analysis techniques is desired. We present a parametric approach for time variant analysis, test its performance for simulated data, and apply it to real-world data.

1. Introduction

Since there is a tremendous number of questions considering the interplay of different parts of the brain, the multivariate analysis of network behavior measured through time series is an important challenge in neuroscience research. The underlying idea is to investigate interactions in the recorded activity to draw conclusions about the functional connectivity of different brain regions. If for example the influence of a process A on a process B is mediated by a third process C, i.e. A → C → B is a conduction, there is no direct influence from A to B. This scenario is often expected to be present in many real-world situations. A reliable conclusion is, however, only feasible if direct and indirect interactions could be distinguished. A pairwise analysis of the signals is usually not sufficient for the exploration of the network since direct and indirect influences cannot be distinguished. The distinction of direct and indirect influences can only be achieved by multivariate analysis techniques, for instance by partial coherence analysis, which corrects for the influences of third processes. Complementary to the detection of direct influences, an investigation towards the direction of influence is often important. If there is a connection between a process A and B it is of interest whether A influences B or vice versa, for instance to understand the physiological or pathological basis of many mechanisms. Since coherence and partial coherence are symmetric measures, they are not able to infer the direction of influence. Partial directed coherence (PDC), which is a parametric approach based on vector autoregressive modelling, has been introduced as a technique to determine directions of influence in multivariate systems (Baccala and Sameshima, 2001). Originally developed for linear stochastic systems, PDC was shown to reveal the underlying interaction structure also for nonlinear systems (Schelter et al., 2006a,b).

The methods mentioned above assume stationarity of the observed processes and thus, influences that are constant in time. Nevertheless, many problems in neuroscience are crucially concerned with changes of the connectivity with time. For instance, such a time varying behavior in the system is expected during a movement task or the perception of a stimulus. Therefore, we address the question of how a time dependent analysis of interaction structures can be performed.

The paper is organized as follows. We first discuss a parametric analysis of stationary time series which is extended to non-stationary processes afterwards. In Section 3 we use simulated data to illustrate the performance of the proposed method. We provide a real-world application to Parkinson tremor presented in Section 4.

2. Methods

In the following we briefly summarize concepts dedicated to the analysis of stationary time series. Afterwards an extension to

* Corresponding author. Address: FDM, Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany. Tel.: +49 761 203 7709; fax: +49 761 203 7700.
E-mail address: linda.sommerlade@fdm.uni-freiburg.de (L. Sommerlade).

0928-4257/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jphysparis.2009.07.005

non-stationary signals which is also capable to deal with observation noise is given.

2.1. Estimation of directed influence: stationary processes

Granger causality is the basic concept for the analysis of the direct directed influences between processes in a multivariate system (Granger, 1969). It formalizes the common sense conception that causes precede their effects in time. To quantify Granger causality in time series especially when accurate signal models are unavailable, it is a quite common approach to model the dynamics by a vector autoregressive processes (VAR-process) of order p

$$\vec{x}(t) = \sum_{r=1}^{p} \mathbf{a}(r)\vec{x}(t-r) + \vec{\epsilon}(t) \tag{1}$$

with $\vec{x}(t) \in \mathbb{R}^n$ denoting the n-variate process, $\mathbf{a}(r)$ the $n \times n$ coefficient matrices of the VAR-process and $\vec{\epsilon}(t) \in \mathbb{R}^n$ a white noise processes with covariance matrix $\mathbf{\Sigma}$. The information about the dynamics contained in the coefficient matrices can be analyzed in both the time as well as the frequency domain, whereas an inspection in the latter is favorable in the case of signals with predominant oscillatory activity. The Fourier transform of the coefficient matrices $\mathbf{a}(r)$ is given by

$$\mathbf{A}(\omega) = \mathbf{I} - \sum_{r=1}^{p} \mathbf{a}(r)e^{-\mathrm{i}\omega r} \tag{2}$$

with the $n \times n$ identity matrix $\mathbf{I}$. The direct directed linear influences in the multivariate system can be estimated by partial directed coherence (Baccala and Sameshima, 2001)

$$|\pi_{l\leftarrow j}(\omega)| = \frac{|\mathbf{A}_{lj}(\omega)|}{\sqrt{\sum_k |\mathbf{A}_{kj}(\omega)|^2}} \in [0,1] \tag{3}$$

If $|\pi_{l\leftarrow j}(\omega)| \neq 0$ a direct linear influence from process j to process l with respect to the other observed processes is indicated. A significance level for testing non-zero partial directed coherence can be derived from the statistical properties of the partial directed coherence, which are examined in Schelter et al. (2006a).

When dealing with measured data, the coefficient matrices $\mathbf{a}(r)$ are estimated. In the estimation procedure one has to deal with observational noise, i.e. the error-in-variables problem.

2.2. Error-in-variables problem

Contamination with observational noise is an almost inevitable characteristics of real-world data. Disregarding observational noise causes an underestimation of the process parameters. This issue is usually referred to as the error-in-variables problem.

If we consider a stationary AR[1]-process without observational noise

$$x(t) = a x(t-1) + \varepsilon(t) \tag{4}$$

the parameter a can be readily estimated by

$$\hat{a} = \frac{\sum_t x(t-1)x(t)}{\sum_t x(t-1)x(t-1)} \tag{5}$$

This follows from Eq. (4) by multiplication with $x(t-1)$. Whenever observational noise is present

$$y(t) = x(t) + \eta(t) \tag{6}$$

with $\eta \sim \mathcal{N}(0,R)$, the estimator for a is biased, when the observational noise is disregarded (Deuschl et al., 1996). Hence, Eq. (5) now reads

$$\langle \hat{a} \rangle = \frac{\langle y(t-1)y(t) \rangle}{\langle y(t-1)y(t-1) \rangle} \tag{7}$$

$$= \frac{\langle (x(t-1)+\eta(t-1))(x(t)+\eta(t)) \rangle}{\langle (x(t-1)+\eta(t-1))^2 \rangle} \tag{8}$$

$$= \frac{\langle x(t-1)x(t) \rangle}{\langle x^2(t-1)+R \rangle} \tag{9}$$

$$= a\frac{1}{1+\frac{R}{\langle x^2(t) \rangle}} \tag{10}$$

where $\langle . \rangle$ denotes the mean. For the above derivation we utilized the fact that $\eta(t)$ is an uncorrelated white noise process. Thus, Eq. (10) shows that as long as the variance R of $\eta(t)$ is non-zero the estimated parameter $\hat{a}$ is smaller than a. The amount of underestimation is quantified by the signal-to-noise ratio $\mathrm{SNR} = \frac{\mathrm{VAR(process)}}{\mathrm{VAR}(\eta(t))}$.

Similar to the above argument, it can be shown that disregarding the observational noise also induces an underestimation of the estimated strength of a directed influence between two interacting processes. To demonstrate the effect of the underestimation we simulated a 2-dimensional VAR[2]-process $\vec{x}(t)$

$$x_1(t) = 1.6x_1(t-1) - 0.96x_1(t-2) + \varepsilon_1(t) \tag{11}$$

$$x_2(t) = 1.8x_2(t-1) - 0.95x_2(t-2) + 0.1x_1(t-1) + \varepsilon_2(t) \tag{12}$$

in which $x_1(t)$ is driving $x_2(t)$, with $N = 5\,000$ datapoints. The simulation was afflicted with observational noise $\eta(t)$ of different variances quantified by the signal-to-noise ratios $\mathrm{SNR} = \frac{\mathrm{VAR}(x_1(t))}{\mathrm{VAR}(\eta(t))}$. Partial directed coherence was estimated for each signal-to-noise ratio using a model order of $p = 10$. In Fig. 1, we show the estimated PDC values $|\pi_{2\leftarrow 1}(\omega)|$ for different signal-to-noise ratios together with the analytical PDC. As the variance of the observational noise increases, i.e. the SNR decreases, the detection of the directed influence from $x_1(t)$ to $x_2(t)$ gets increasingly difficult. Based on the same simulated data, we also estimated the PDC values using a procedure that accounts for the observational noise. For the shown signal-to-noise ratios, the estimated PDC values at the peak frequency corresponded well with the analytical PDC.

2.3. State space model for stationary data

The estimation of time-varying interaction structures in multivariate systems is of particular interest for the understanding of ongoing brain activity. A first approach could be to segment the time series in reasonably short epochs, estimate the VAR-

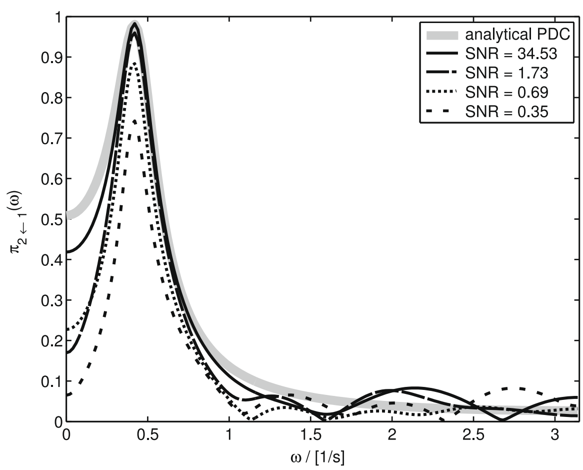

Fig. 1. Analytical $\pi_{2\leftarrow 1}(\omega)$ and estimated $\pi_{2\leftarrow 1}(\omega)$ for different signal-to-noise ratios. For increasing variance of the observational noise, the estimated PDC at the predominant frequency of the process decreases and the influence might not be detected.

coefficients and estimate the PDC values for each segment. Unfortunately, this method often fails to detect rapid changes in the interaction structure and will always be a trade-off between a desired high resolution and the necessity of data segments long enough for the analysis.

A procedure superior to windowing the time series is given by state space modelling, which also overcomes the challenge of poor signal-to-noise ratios.

For stationary time series, state space modelling combined with the Expectation–Maximization (EM) algorithm is a powerful tool to estimate the process parameters and thereby the interaction structure when observational noise cannot be neglected (Shumway and Stoffer, 2000). The general notation of a linear state space for a stationary process is given by

$$\vec{u}(t) = \mathbf{A}\vec{u}(t-1) + \vec{\varepsilon}_u(t) \tag{13}$$
$$\vec{y}(t) = \mathbf{C}\vec{u}(t) + \vec{\eta}(t). \tag{14}$$

Thereby Eq. (11) describes the dynamic of the hidden process $\vec{u}(t)$, whereas the unknown process parameters are incorporated in matrix **A**. The observed process $\vec{y}(t)$ contaminated with Gaussian white noise $\vec{\eta}(t)$ is given by Eq. (12). In many cases, an observation matrix **C** has to be taken into account which describes the way the state $\vec{u}(t)$ is transformed to the observation space. However, in the following we set **C** to the identity, assuming a direct observation of the states possibly contaminated with observation noise. Both, **A** and $\vec{u}(t)$ can be estimated via state space modelling and the EM-algorithm applying the Kalman filter (Wan and Nelson, 2001). Note that the state space model above addresses only a VAR[1]-process. This is sufficient because every n-dimensional VAR[p]-process $\vec{x}(t)$ can be rewritten as an np-dimensional VAR[1]-process $\vec{u}(t)$.

2.4. Extension to non-stationary processes

The extension of the algorithm to non-stationary data requires consideration of time dependent parameters $\mathbf{A}(t)$. For this purpose, a second state space for the parameters is introduced. The dual state space model incorporates two state spaces, one for the hidden variable $\vec{u}(t)$ and one for the parameters $\mathbf{A}(t)$. The process parameter matrix $\mathbf{A}(t)$ is assorted in a vector $\vec{a}(t)$ to make it accessible for state space modelling. The dual state space is given by

$$\begin{aligned}
\vec{u}(t) &= \vec{f}(\vec{u}(t-1)|\vec{a}(t-1)) + \vec{\varepsilon}_u(t),\\
\vec{y}(t) &= \vec{g}(\vec{u}(t)) + \vec{\eta}(t),\\
\vec{a}(t) &= \vec{a}(t-1) + \vec{\varepsilon}_a(t),\\
\vec{y}(t) &= \vec{g}(\vec{f}(\vec{a}(t-1)|\vec{u}(t-1)) + \vec{\varepsilon}_u(t)) + \vec{\eta}(t)
\end{aligned} \tag{15}$$

with multivariate, independent Gaussian noise $\vec{\eta}(t)$,

$$\vec{f}(\vec{u}(t-1)|\vec{a}(t-1)) = \vec{f}(\vec{a}(t-1)|\vec{u}(t-1)) = \mathbf{A}(t-1)\vec{u}(t-1) \tag{16}$$

and $\vec{g}(.)$ being the observation function. For all investigations and simulations it was selected such that the observed time series are the ones that are observable in the model. The function $\vec{g}(.)$ is thus basically the identity matrix with additional zeros to account for the fact that the delayed processes needed in the VAR[1] representation of any VAR[p]-process are not directly observable. The used notation of $\vec{f}(\cdot|\cdot)$ denotes that the second argument of $\vec{f}$ is a true parameter for the first argument. The expectation step of the EM-algorithm is carried out by a dual Kalman filter and can be divided in two main steps: first, the hidden variables $\vec{u}(t)$ are estimated based on the parameters $\vec{a}(t-1)$, which are considered as true in this step. In the above equations, denoted by $\vec{f}(\vec{u}(t-1)|\vec{a}(t-1))$ in Eq. (15). In the second step, the process parameters $\vec{a}(t)$ are estimated based on the process trajectory $\vec{u}(t-1)$ which has been estimated in the first step and is for the estimation of the parameters $\vec{a}(t)$ considered as true (Wan and Nelson, 1997, 2001). Details concerning the implementation of the Kalman filter can be found elsewhere (Shumway and Stoffer, 2000). The dual estimation procedure was used although others are possible. It is a numerically very efficient way of estimating the parameters.

Thus, the EM-algorithm applying the dual Kalman filter in its expectation step yields both the estimated parameters $\vec{a}(t)$ and the most likely trajectory of the process $\vec{u}(t)$. The parameters $\vec{a}(t-1)$ of the VAR[1]-process for each sampling point can be mapped to the parameters of a VAR[p]-process. Thereby $\mathbf{A}(t)$ is obtained. The direct application of Eq. (3) leads to a time resolved partial directed coherence $|\pi_{l\leftarrow j}(\omega,t)|$. This allows the quantification of the direct directed time dependent influence in multivariate systems contaminated with observational noise.

3. Simulations

To demonstrate the abilities of the method of time dependent partial directed coherence previously described, we simulated a 3-dimensional VAR[2]-process

$$\begin{aligned}
x_1(t) &= 0.59x_1(t-1) - 0.2x_1(t-2) + b(t)x_2(t-1)\\
&\quad + c(t)x_3(t-1) + \varepsilon_1(t)\\
x_2(t) &= 1.58x_2(t-1) - 0.96x_2(t-2) + \varepsilon_2(t)\\
x_3(t) &= 0.60x_3(t-1) - 0.91x_3(t-2) + \varepsilon_3(t)
\end{aligned} \tag{17}$$

with time variant interaction. The system can be considered as two damped stochastically driven oscillators (x_2 and x_3) and a stochastically driven relaxator (x_1). The time course of the parameters $b(t)$ and $c(t)$ which denote the intensity of the influences are depicted in Fig. 2. Although the presence of a directed influence from one process on another is a binary quantity, the influence itself can be weak or strong depending on the parameters. In the example, the process x_2 is influenced by x_1 through $b(t)$ with oscillating strength, whereas x_3 influences x_1 with increasing strength in the first half of the simulation and with decreasing strength in the second half. The estimated time variant PDC is shown in Fig. 3. It reproduces the simulated interaction structure. In matrix element $(1,2)$, which shows $\pi_{1\leftarrow 2}(\omega)$, the influence with oscillating strength of process x_2 on process x_1 is correctly identified. The increasing and deceasing influence of process x_3 on x_1 is also correctly revealed. The spectrum of x_1 reflects the driving influence of the other two processes. The order of the VAR-process fitted to the data was chosen with $p = 2$. For the choice of an appropriate order p for real-world data measures like Akaike's Information Criterion or Bayesian Information

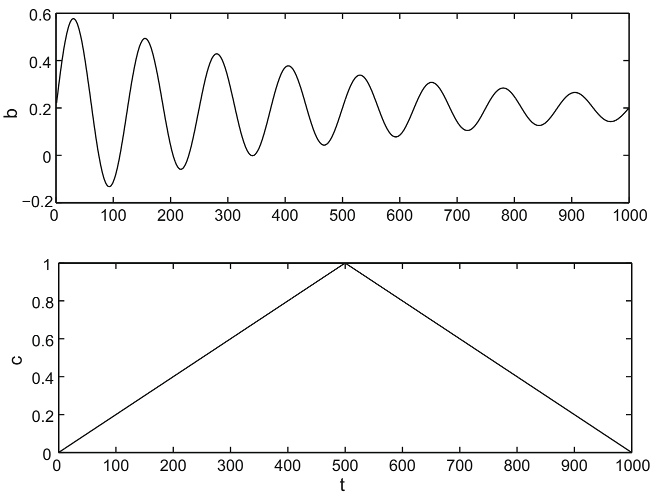

Fig. 2. The parameters of the strength of directed influence: $b(t)$ denotes the strength of influence between x_2 and x_1, $c(t)$ that between x_3 and x_1.

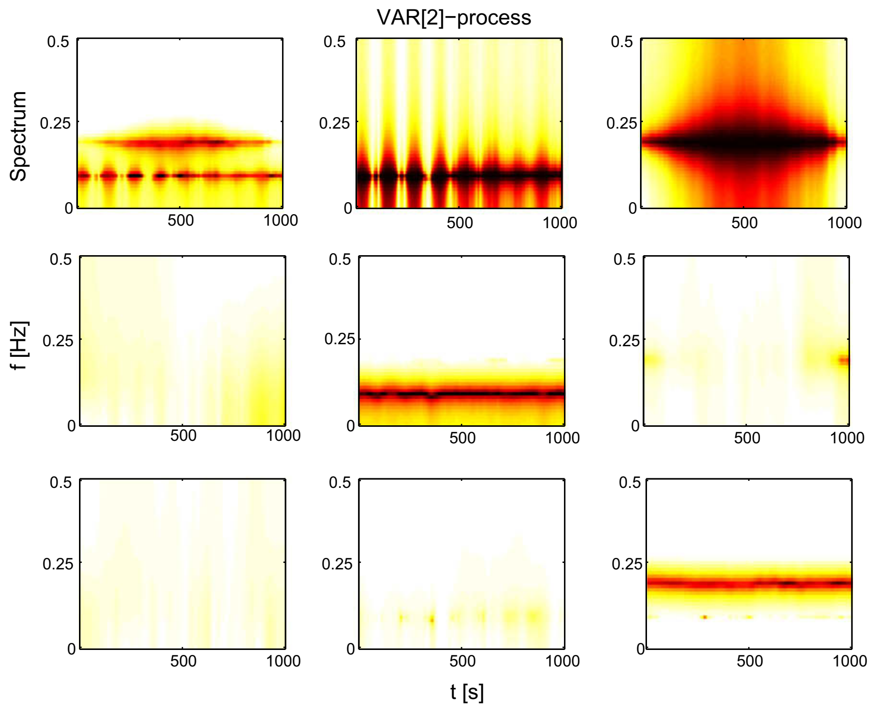

Fig. 3. Colour-coded plot of the time resolved interaction structure of the simulated VAR[2]-process: the spectra of the processes are shown on the diagonal. The off-diagonal matrix elements (i, j) contain the time variant PDC values $\pi_{i \leftarrow j}(\omega)$ which measure the influence of process j onto process i.

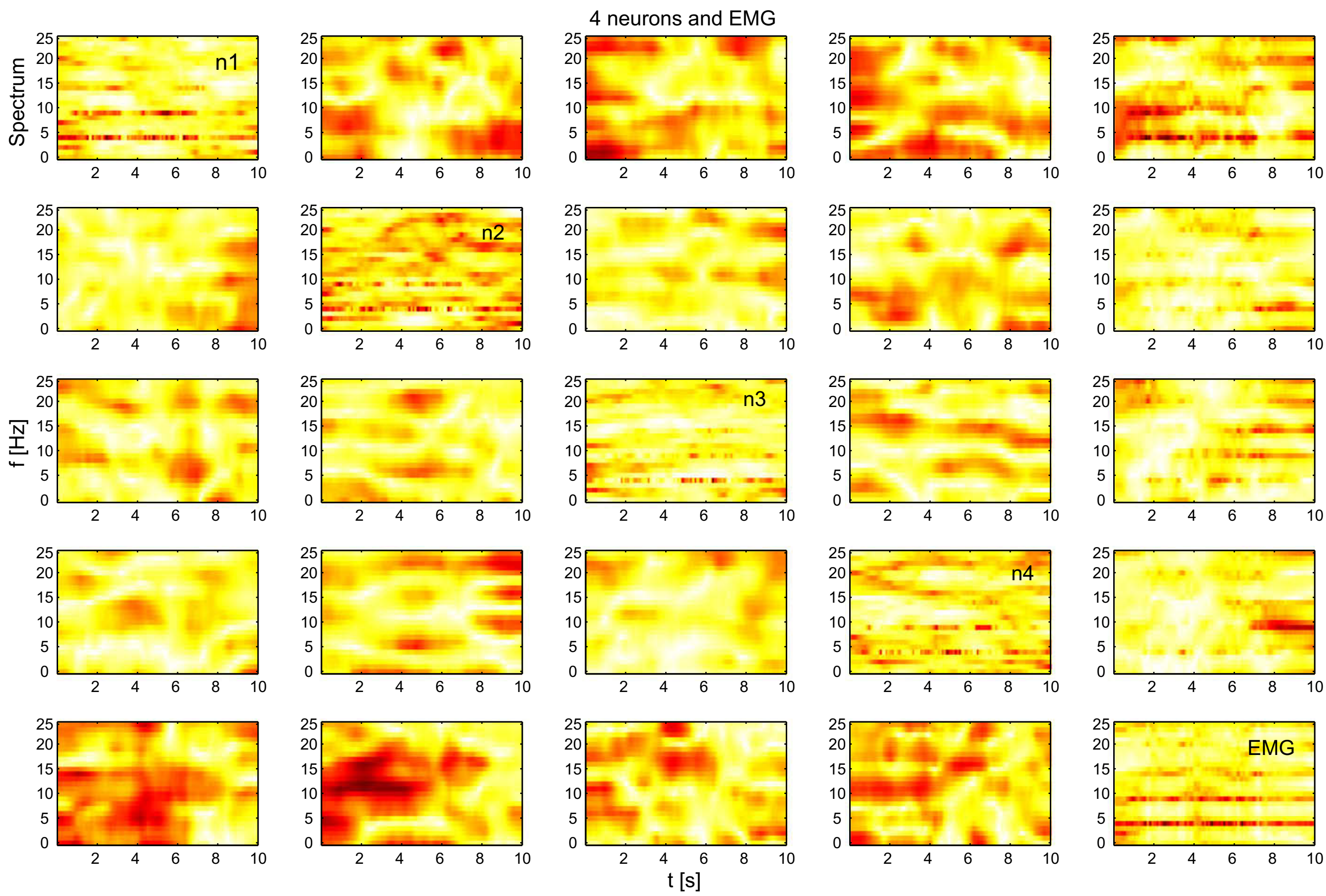

Fig. 4. Four STN neurons and flexor muscle of the contralateral forearm: spectra are shown on the diagonal. The time variant PDC coefficients are assorted as described in Fig. 3.

Criterion are available (Akaike, 1969; Schwarz, 1978). However, for non-linear processes a higher model order is required to describe the dynamics (Schelter et al., 2006a).

4. Application

One of the core symptoms of Parkinson's disease is tremor, which is an involuntary rhythmical movement predominantly of the upper limbs. The pathophysiology of the human tremor is still under debate. But it is very likely that an abnormal oscillatory activation within basal ganglia-cortical and cerebellar-cortical loops plays an important role in generating tremor. In Parkinson's disease the tremor frequency is mainly between 4 and 6 Hz (Timmer, 1998).

We investigated data from the subthalamic nucleus (STN), a part of the basal ganglia, and electromyography (EMG) recordings of trembling muscles, located at the forearm. Data was registered in the STN using microelectrodes during stereotactic neurosurgery. Afterwards the implantation of electrodes for deep brain stimulation was performed, which is an effective method if medicamentous therapy was not sufficient for treatment of tremor. The optimal position for the electrode implantation was determined by a fusion technique of MRI and stereotactic computerized tomography. The penetration of the STN as the target region was confirmed by detecting a specific neuronal spike activity. Micro electrode recordings with a sampling rate of 25 kHz were performed for different depths in the STN. During one recording session between 2 and 5 neurons could be identified via spike sorting (Quiroga and Nadasdy, 2004). Spike trains were down sampled to make them accessible for VAR-modelling. Simultaneously EMG electrodes were placed on the extensor- and flexor muscles of the contralateral wrist to measure the tremor activity at a sampling rate of 2.5 kHz. EMG recordings were rectified and corrected for the mean.

Tremor is a time dependent phenomenon (Hellwig et al., 2003), since it may change from strong to weak and vice versa. In addition, the signal-to-noise ratio of micro electrode data is not as good as that of EMG data. Hence, a time-variant estimation procedure that is able to deal with observational noise, is required. An example for ongoing tremor activity is shown in Fig. 4. The activity was analyzed over a duration of 10 s. The analysis reveals directed influences from the recorded neurons to the muscle as well as from the muscle to the neurons. The trembling frequency in this example is approximately 4.5 Hz. The analysis reveals in matrix element (1,5) influences from the muscle to neuron 1 at the tremor frequency and at its first higher harmonic. The influence at the first higher harmonic persists during the first 4 s, then it slowly vanishes while the influence at the tremor frequency persists over the whole 10 s. The PDC values for the opposite direction of influence, depicted in matrix element (5,1), are also identifiable, although characterized by a more wide-spread influence.

5. Conclusions

We have presented an approach to infer time dependent interaction structures in multivariate systems. The presented parametric approach consists of the estimation of the time dependent VAR[p]-coefficients, using a dual state space model, and the estimation of the PDC for each sampling point.

The performance of the method has been demonstrated by means of a simulated VAR-system with time variant interaction structure. The interaction structure could be inferred and the variation of the parameters defining the strength of influence in time was well captured.

We applied the method to an example of real-world data recorded from a patient suffering from tremor in Parkinson's disease. The analysis revealed changes in the dynamics as well as the interaction between brain structures and trembling muscles, indicating a dynamic interplay between both. In detail, we found a direct linear time-varying influence from the muscles to the STN at the tremor frequency or its higher harmonics in all four observed EMG-neuron combinations. The influence from the STN to the muscles could also be observed. This strongly suggests a participation of the STN in tremor generation.

Based on the first promising results of the time varying PDC analysis to real-world data, this technique suggests itself for further application.

Acknowledgements

We thank Jan Vesper for his contribution concerning data acquisition. This work was supported by the German Science Foundation (Ti 315/2-2; Ti 315/4-2; He1949/1-1) and the German Federal Ministry of Education and Research (BMBF grant 01GQ0420). This work was supported by the Excellence Initiative of the German Federal and State Governments.

References

Akaike, H., 1969. Fitting autoregressive models for prediction. Ann. Inst. Stat. Math. 21, 243–247.
Baccala, L.A., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. Biol. Cybern. 84, 463–474.
Deuschl, G., Krack, P., Lauk, M., Timmer, J., 1996. Clinical neurophysiology of tremor. J. Clin. Neurophysiol. 13, 110–121.
Granger, J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438.
Hellwig, B., Schelter, B., Guschlbauer, B., Timmer, J., Lücking, C., 2003. Dynamic synchronisation of central oscillators in essential tremor. Clin. Neurophysiol. 114, 1462–1467.
Quiroga, R.Q., Nadasdy, Z., 2004. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Computat. 16, 1661–1687.
Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., Lücking, C., Dahlhaus, R., Timmer, J., 2006a. Testing for directed influences among neural signals using partial directed coherence. J. Neurosci. Methods 152, 210–219.
Schelter, B., Winterhalder, M., Hellwig, B., Guschlbauer, B., Lücking, C.H., Timmer, J., 2006b. Direct or indirect? Graphical models for neural oscillators. J. Physiol. Paris 99, 37–46.
Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.
Shumway, R., Stoffer, D., 2000. Time Series Analysis and its Application. Springer, New York.
Timmer, J., 1998. Modeling noisy time series: physiological tremor. Int. J. Bifurcat. Chaos 8, 1505–1516.
Wan, E.A., Nelson, A.T., 1997. Neural dual extended Kalman filtering: applications in speech enhancement and monaural blind signal separation. In: IEEE Proceedings in Neural Networks for Signal Processing: VII. Proceedings of the 1997 IEEE Workshop, pp. 466–475.
Wan, E.A., Nelson, A.T., 2001. Dual extended Kalman filter methods. In: Kalman Filtering and Neural Networks. John Wiley, New York.