Analysis of multichannel biomedical data

Maciej Kamiński¹, Jarosław Żygierewicz¹, Rafał Kuś¹, and Nathan Crone²

¹ Department of Biomedical Physics, Warsaw University, 69 Hoża St., 00–681 Warsaw, Poland; ² The Johns Hopkins University School of Medicine, Dept. of Neurology, Meyer 2–147, 600 N Wolfe St., Baltimore, MD, USA

Abstract. Nowadays, there is a common practice in biomedical research to perform multiple time series recordings. In the first part of this paper, basic information about analysis of such multichannel biomedical data is given. A short overview of important differences between single-channel, two-channel and multichannel data sets is presented and various coherence functions are reported. Causal relations between channels are investigated by means of the Directed Transfer Function (DTF) and its dynamic version, the Short-Time Directed Transfer Function (SDTF). The introduced formalism was used to analyze 12-channel human electrocorticogram (ECoG) records. Preliminary results of a study of causal dependence in beta and gamma frequency bands in two patients performing a motor task are reported. Specific characteristics in activity propagation consistent for both subjects for different rhythms were found.

Key words: ERD, ERS, DTF, causal connections, multivariate AR model

The correspondence should be addressed to M. Kamiński, Email: Maciek.Kaminski@fuw.edu.pl
INTRODUCTION

The technological progress in data recording equipment, especially visible in the biomedical research field, has resulted in the production of multichannel sets of data. Today a typical dataset contains not only two or four but dozens of channels. This is especially true for electro- or magnetoencephalography (EEG/LFP/MEG) recordings. Analysis of multichannel data can give a better insight in the relations between the investigated sites and allows for estimation of network properties of the subject of interest. However, it may be challenging to extract the desired information from such datasets. Besides many experimental and computational difficulties, the problem quite often lies in the proper application of existing mathematical tools. Multichannel data sets require adequate handling in order to get proper results. For instance, the popular technique of the mapping of signal power on the head surface may seem at a first glance to be a multichannel method, because it uses information from all the channels. In fact, this technique uses values calculated for each channel separately, neglecting the covariance structure (interdependencies between channels) of the dataset and from this point of view it cannot be called a truly multichannel method (see Discussion section). In this paper an introduction of basic aspects of multichannel data processing will be presented. In the Experiment section a brief report of a new, ongoing study of human ECoG will be presented, which will additionally serve as an example of the introduced formalism. Unlike other studies, in the presented research we have focused on transmissions of activity between channels, not on the level of activation of each site separately.

METHODS

A typical neurobiological data recording consists of several time series: observations of the investigated variables at certain time intervals. For each time series, many quantities characterizing it can be calculated e.g., the average amplitude or variance in the time domain or power spectrum in the frequency domain. Such measures, applicable to separate data channels are called auto-quantities. However, a multichannel set of data contains more information than the values characterizing each channel separately. For instance, two data channels can be either correlated with each other or uncorrelated, while still having the same power spectra in each case. To describe inter-relations between channels we use cross-measures which simultaneously depend on two (or more) channels. Typically, for a pair of channels, cross-correlations in the time domain or coherences in the frequency domain are commonly used. For a k-channel system its power spectrum \( S(f) \) is a matrix with auto-spectra on the diagonal and cross-spectra outside the diagonal (formulas for calculating the spectral matrix can be found in the literature e.g., in Jenkins and Watts 1998):

\[
S(f) = \begin{pmatrix}
S_{11} & S_{12} & \cdots & S_{1k} \\
S_{21} & S_{22} & \cdots & S_{2k} \\
\vdots & \vdots & \ddots & \vdots \\
S_{k1} & S_{k2} & \cdots & S_{kk}
\end{pmatrix}
\]

(1)

The cross-spectral terms \( S_{ij} \) for \( i \neq j \) represent common (signals which are synchronized in phase) power appearing in two channels. The normalized version of the cross-spectrum is called (ordinary) coherence and is defined for a pair of channels \( i \) and \( j \) by the formula:

\[
K_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}
\]

(2)

The modulus of the coherence function lies in the range \([0, 1]\); zero means that the channels are not related with each other at the given frequency.

When a system consists of more than two channels, new possibilities of inter-dependency of the channels appear. They can be connected with several other channels, either simultaneously or in chain. To evaluate a set of more than two channels, making use of the previous observation, special multichannel measures should be used. The most popular are partial and multiple coherences (Jenkins and Watts 1998). The modulus value of each type of coherence lies in the range \([0, 1]\); values close to zero indicate a lack of relation.

Partial coherence is designed to describe only direct relations between channels. All relations which can be explained by linear combinations of other data channels will not be shown by this measure. It is defined as:

\[
D_{ij}(f) = \frac{M_{ij}(f)}{\sqrt{M_{ii}(f)M_{jj}(f)}}
\]

(3)

where \( M_{ij} \) represents a minor of spectral matrix \( S \) with the \( i \)-th row and the \( j \)-th column removed.
Multiple coherence

\[
G_t(f) = \sqrt{1 - \frac{\det(S(f))}{S_n(f)M_n(f)}}
\]

(4)

describes the similarity of the given signal to any other signal from the rest of the set. It helps detecting a situation when a signal is not really connected with the set.

Any measure describing causal relations between data channels obviously belongs to the “cross” quantities group. The first attempts to define such a measure date back to the 1950’s in social sciences. Since then many functions were proposed, but most of them were designed for application to a pair of channels only. The first definition of causality which can be used in time series analysis and which is now popular in biomedical data analysis was given by Granger (1969). His definition is based on predictability of time series. For two simultaneously measured signals \(X_i\) and \(X_j\), if we can better predict \(X_i\) by using the past information from both signals \(X_i\) and \(X_j\):

\[
X_i(t) = \sum_{j=1}^{p_1} A_{i1}(j)X_i(t-j) + \sum_{j=1}^{p_2} A_{i2}(j)X_j(t-j) + E'(t)
\]

(5)

then by only using the past information from signal \(X_i\) itself,

\[
X_i(t) = \sum_{j=1}^{p_1} A_{i1}(j)X_i(t-j) + E''(t)
\]

(6)

(e.g., when \(\text{var}(E') < \text{var}(E'')\)), then we call signal \(X_j\) causal to \(X_i\). \(E\) values can be called in this case error functions. Constants \(p_1\) and \(p_2\), determining the considered time lag of the relation, depend on the particular situation.

In this paper a parametric method of spectral data analysis will be used, namely multichannel (multivariate) autoregressive (MVAR) modeling. This well-known technique relies on the quality of the AR model fitting. It can be shown (Blinowska et al. 1988, Franaszczuk and Blinowska 1985) that that type of model describes well stochastic time series containing a set of damped sinusoids, so-called rhythms. This property makes it very useful in analysis of EEG and other biomedical signals, where specific rhythms of a certain frequency are embedded in a noisy background.

We assume that a sample of data (vector) \(X\) at a time \(t\) can be expressed by its \(p\) previous samples with certain (matrix) coefficients \(A(i)\) plus a (vector) noise component \(E(t)\) in the following way:

\[
X(t) = \sum_{i=1}^{p} A(i)X(t-i) + E(t)
\]

(7)

By fitting the model to the data we get a set of \(A(i)\) coefficients which describe properties of the original data. When transforming Eq. 7 to the frequency domain we get (Marple 1987):

\[
X(f) = (A(f))^{-1}E(f) = H(f)E(f)
\]

(8)

where \(X(f)\), \(A(f)\) and \(E(f)\) are Fourier transforms of the respective time domain variables \(X(t)\), \(A(i)\) and \(E(t)\). The MVAR model allows for easy calculation of power spectra and all types of coherences evaluated for the whole set of data simultaneously. Details of the model fitting procedure are presented in the literature (Jenkins and Watts 1998, Kamiński and Blinowska 1991, Marple 1987). Moreover, the MVAR model gives an opportunity to readily describe causal influences between data channels as well. The matrix \(H = A^{-1}\) is called the transfer matrix of the system. It contains frequency dependent information about all relations between channels. The Directed Transfer Function proposed by (Kamiński and Blinowska 1991) is a causal influence estimator based on the transfer matrix. DTF can be calculated in its normalized (Kamiński and Blinowska 1991) or non-normalized version. The non-normalized DTF is defined by simply using elements of the matrix \(H\):

\[
\theta_{ij}(f) = \left| H_{ij}(f) \right|
\]

(9)

Note that \(\theta_{ij}(f)\) describes the transmission of the signal from channel \(j\) to channel \(i\) at frequency \(f\).

The definition of Granger causality was originally given for a pair of channels. It can be extended for the multichannel case. In this situation we compare the variance of prediction error \(E'\) where the \(i\)-th channel is included into the prediction of the first channel (Eq. 10)

\[
X_i(t) = \sum_{m=1}^{i} \sum_{j=1}^{p_m} A_{im}(j)X_m(t-j) + E'(t)
\]

(10)

with the variance of prediction error \(E''\) where the \(i\)-th channel is not included (Eq. 11):
\[ X_1(t) = \sum_{m=1}^{k} \sum_{j=1}^{P_m} A_{jm} (j) X_m(t-j) + E'(t) \]  

(11)

If \( \text{var}(E') < \text{var}(E') \) we call the \( i \)-th signal causal to the first one in the multivariate sense.

The DTF function is a truly multivariate measure and it can be applied to an arbitrary number of channels. It was shown (Kamiński et al. 2001) that DTF is equivalent to Granger causality in the multivariate sense.

DTF is not a partial measure. To show only direct causal relations other functions were proposed and they can be found in the literature: the DC method, where the presence of a direct relation of channels \( i \) and \( j \) is indicated by a nonzero value of time domain MVAR coefficients \( A_{ji}(t) \) (Kamiński et al. 2001), the Partial Directed Coherence (PDC) method in (Sameshima and Baccalá 1999), or the direct DTF (dDTF) function where a version of DTF is combined with partial coherences (Korzeniewska et al. 2003). A comparison between different estimators is given in Kuś and coauthors (2004).

It should be stressed that the differences between a two-channel and a multi-channel \((k>2)\) dataset are very important. This is especially true for estimation of directions of influence. It can be shown that a pairwise approach may in this case lead to results which are likely to be misinterpreted (Blinowska et al. 2004, Kuś et al. 2004).

One of the advantages of an MVAR model spectral analysis is the fact that the model can be fitted to short data segments, typically of a length much shorter than the length required for e.g., Fourier transform methods, and still give reliable power spectrum estimates. Moreover, when multiple repetitions of an experiment are available, the whole ensemble of realizations can be utilized in calculations. Instead of averaging over time (which is implicitly done when analyzing long data epochs) we average the correlation matrix over the repetitions. This allows for processing even shorter data segments with acceptable statistical properties of the estimates. This idea led to the construction of the Short Time Directed Transfer Function (SDTF) (Kamiński et al. 2001). This function can be used to visualize the dynamics of transmissions during the investigated process. The whole time epoch is divided into short, overlapping time windows. The MVAR models are fitted for each window using all the realizations of the process. Based on the estimated MVAR parameters power spectra, coherences and DTFs are calculated. The details of the procedure can be found in Ding and coauthors (2000) or in Kamiński and coauthors (2001).

**The experiment**

The study concerned human ECoG data recorded at the Johns Hopkins Hospital in Baltimore. A selected group of epilepsy patients had electrode grids implanted over the brain cortex surface, which was part of the treatment. Some of those patients agreed to participate in the study. Each subject was asked to perform a specific task in response to a visual stimulus presented on a computer screen. For the presented analysis, two patients performing motor tasks were selected. They were required to sustain a muscle contraction of a specific body part (fist or tongue), indicated by a drawing which appeared on the screen for 3 seconds, and release the contraction at the moment the drawing disappeared. The task was repeated many times producing multiple realizations of the investigated process. Both patients had a subdural 64-electrode array grid implanted over similar areas of left frontoparietal cortex. The data sampling frequency was 1000 Hz. Because of our specific interest in the beta and gamma frequency bands the data were downsampled to 250 Hz. The ensemble of trials was aligned according to the stimulus onset time. Seven second long epochs were analyzed (3 s before to 4 s after the stimulus onset). The number of artifact free trials was around 50 (slightly different for each patient).

**RESULTS**

Because the beta (20–30 Hz) and gamma (35–45 Hz) bands were of primary interest, the data were bandpass filtered within the specified frequency ranges. The filtered data was analyzed by the SDF function. For the analysis, a sliding time window of 125 ms was chosen. Since the number of data points should be several times bigger than the number of fitted MVAR parameters, a subset of 12 out of original 64 signals were selected. To find the most relevant channels, changes of power in respect of the reference period (2.5 to 1 s before the stimulus) were calculated for 16 successive frequency bands from 10 to 50 Hz for every time window, giving the time-frequency maps of
Event Related Desynchronization and Synchronization (Zygierewicz et al. 2005) (Fig. 1). The channels with the biggest change in power were the first candidates for further selection; the final choice was subsequently consulted with a neurologist.

In the figures below, examples of time-frequency maps of the transmissions in the selected set of channels (SDTF function) for the fist clenching task are presented. The general observation is that patterns for beta and gamma bands differ. Beta transmissions (Fig. 2) are high for certain connections and at the time of the stimulus onset they rapidly vanish (35→53, 35→45, 53→61, 51→53, 43→53, 43→45 and others). In some cases the transmissions disappear at the end of the investigated epoch, sometimes at slightly higher frequency (35→45, 43→45, 35→44). Such behavior is common for both subjects, although the precise localizations differ.

Gamma activity transmissions (example in Fig. 3) were less organized, occurred rather in a form of bursts. Certain connections seem to be active during the whole investigated epoch (e.g., most transmissions from channel 43). For specific connections the bursts could be observed at the beginning and at the end of the motor activity (51→61, 51→45).

The matrices of SDTF time-frequency maps may be difficult to interpret. To solve this problem, the values of SDTF were integrated over the beta and gamma frequency bands respectively. The integrated transmission values are presented graphically in form of arrows pointing from source to destination sites. The values obtained by the integration of intensity of flows were coded as color (and transparency) of arrows (Fig. 4).

In this form of presentation we can see the topographical placement, strength and timing of each transmission in a more natural way. In Fig. 4 an example of transmissions for the first 10 channels presented in the form of maps in Fig. 2, is shown for two time points: 1.7 s before and 0.5 s after the stimulus onset (the last two channels from Fig. 2 were taken from a different
grid of electrodes and are not represented in this picture). The phenomenon of vanishing beta transmissions after the stimulus onset (time 0) is clearly visible, there are no strong flows present at 0.5 s after the stimulus.

**DISCUSSION**

Multichannel data sets contain not only information relevant to each data channel separately, but the mutual relations of the channels as well. We can fully utilize it only when all channels are analyzed simultaneously and, consequently, a multichannel measure, describing the covariance structure of a whole set, is obtained. Additionally, when directions of influence are of importance, a pairwise analysis of a bigger set of channels may lead to results which can be misinterpreted. The example below (Fig. 5) illustrates the fact that the correlation structure of a dataset contain different information than obtained from analysis of all the channels separately. For two human sleep scalp EEG segments maps of power were compared with DTF network analysis. In the first case, the position of the maximum on the map agrees with the position of the area of the main source of activity. For the second data segment, where two sources of activity are present, both methods apparently present different pictures. In fact, results presented by both methods
concern different quantities and the problem lies in the correct interpretation of the obtained pictures. In particular, a naive interpretation of a map can lead to a wrong conclusion concerning the localization of activity sources.

The MVAR modeling, which was widely used in biomedical data analysis, is a proposition of an approach allowing estimation of all coherence functions and causal connections estimators in the multichannel way. Although technological progress delivers new tools to investigate brain activity, like the functional magnetic resonance imaging (fMRI) technique, which can be combined with traditional EEG recordings (Sommer at al. 2003), we must stress the fact that such results may not be easily comparable with results of a causal relations pattern. Methods describing properties of each site separately give different information than multivariate analysis of transmissions.

By analyzing EEG or LFP data we observe collective activity of ensembles of neurons. Such activity propagates along anatomical tracts. In Fig. 5, where (in the second evaluated epoch) an alpha wave is analyzed, the results agree well with other neurophysiological evidence: the alpha activity is known to be generated in the parietal/occipital region of the cortex (visual) and then it propagates toward the front of the head. By applying a linear model tool to the EEG/LFP signals we detect the presence of causal relations in the system and we evaluate functional connectivity between sites (Büchel and Friston 1997, Kamiński and Liang 2005). We do not detect all possible anatomical connections, we find the ways of propagations active during the investigated phenomena.

The presented human ECoG study shows how the autoregressive modeling can be used to evaluate the dynamics of transmissions in EEG during motor activ-
Fig. 4. Propagations in the beta range (20–30 Hz) for a set of 10 electrodes for patient 1 for 1.7 s before (above), and 0.5 s after the stimulus onset (below). Direction is shown as a triangle-shaped arrow pointing from the source to the destination electrode. Intensity of transmissions is coded in color and transparency of arrows (scale bar on the left of each picture, identical for both pictures).
ity. It should be noted that the DTF method was already successfully applied to analyze EEG and LFP signals, namely to identify epileptogenic foci in humans (Franaszczuk et al. 1994), to determine a topographic pattern of transmissions during sleep from human scalp EEG data (Kamiński et al. 1997), and to find changes in interrelations between selected limbic system structures in rats performing locomotion tasks (Korzeniewska et al. 1997). In the present study for the first time the dynamic version of DTF was used to investigate causal relations in human multichannel ECoG data recorded during motor tasks. The inter-connections between channels are more complex and more difficult to visualize than quantities referring to separate signals, especially for bigger number of channels.

Integrating transmissions within given frequency bands reduces the amount of information to display and allows for the usage of more legible forms of presentation. Although there are differences between subjects (because of inter-subject variability and different placement of the electrode grids), some general systematic patterns of transmissions during the task are visible. The study is now in progress, extending the number of subjects in order to find common features and systematic trends in transmissions. In spite of the fact that the project is in a preliminary stage, we found that the most active regions identified in the study agree with those found by other methods to be important in motor control e.g., cortical stimulations and visual analysis of LFP traces by experts. It is especially important to evaluate the role of each of the rhythms during a motor task. Event-related desynchronization in the beta band is a well-known phenomenon related to motor activity (Gómez et al. 2004, Pfurtscheller and Arnibar 1979, Pfurtscheller and Lopes da Silva 1999).

We can see that this phenomenon is reflected in the pattern of transmissions between specific channels. Similarly, the role of gamma activity in synchronization of sites may find its confirmation and further development by analyzing specific transmissions in this frequency range (Crone et al. 1998).

ACKNOWLEDGMENT

We thank Professor K. J. Blinowska from the Department of Biomedical Physics at Warsaw University for helpful discussions. This paper was partly supported by the Polish Committee for Scientific Research grant 4T 11E 02823.

REFERENCES


Received 24 June 2005, accepted 3 November 2005