Biol. Cybern. 65, 203–210 (1991)

Biological Cybernetics

A new method of the description of the information flow in the brain structures

M. J. Kamiński and K. J. Blinowska

Medical Physics Laboratory, Institute of Experimental Physics, Warsaw University, ul. Hoża 69, Pl-00-681 Warszawa, Poland

Received February 10, 1991

Abstract. The paper describes the method of determining direction and frequency content of the brain activity flow. The method was formulated in the framework of the AR model. The transfer function matrix was found for multichannel EEG process. Elements of this matrix, properly normalized, appeared to be good estimators of the propagation direction and spectral properties of the investigated signals. Simulation experiments have shown that the estimator proposed by us unequivocally reveals the direction of the signal flow and is able to distinguish between direct and indirect transfer of information. The method was applied to the signals recorded in the brain structures of the experimental animals and also to the human normal and epileptic EEG. The sensitivity of the method and its usefulness in the neurological and clinical applications was demonstrated.

1 Introduction

The information about the flow of brain activity is frequently of primary importance in neurological investigations and also in clinical applications. This problem has been approached by different methods e.g.: momentary brain field maps of Lehman et al. (1987), multilag cross-covariance functions of Gevins (1989), comparison of coherences as a function of different electrode distances (Thatcher et al. 1986). Since EEG is usually characterized by means of its spectral properties, the frequency content as well as the direction of the information flow is of interest. Therefore in our recent investigations we have been aiming at the formulation of the estimator characterizing at the same time direction and spectral properties of the brain signals.

This problem has been approached by Saito and Harashima (1981) and Kamitake et al. (1984) who have defined the so-called *directed coherences* in the framework of the AR model and by Schnider et al. (1989). In the above mentioned works only two channels have been considered. The extension of the formalism to bigger number of channels causes the difficulties increasing very rapidly with their number. The problem can be solved only after making assumptions concerning the pattern of input noises. In the framework of the above mentioned approach we have developed a formula for calculation of *directed coherences* for 4 or more channels and we have obtained encouraging results both for simulated and for neurophysiological data (Kamiński 1988). Nevertheless, the method required solving a complicated set of non-linear equations and was not quite general. Therefore we continued our search for a good measure of the information flow in the brain. Below we shall describe the method of obtaining an estimator that we have found satisfactory in this respect.

2 The method

The method is formulated in the framework of the multichannel AR model. The inputs to all channels are uncorrelated white noise signals. All channels are interconnected (Fig. 1).

The multichannel AR model can be defined in the form:

$$\sum_{j=0}^{p} \hat{A}_j x_{t-j} = e_t \qquad (1)$$

where $x_t = (x_{1,t}, x_{2,t}, \ldots, x_{k,t})$ is the vector of EEG k-channel process, $e_t = (e_{1,t}, e_{2,t}, \ldots, e_{k,t})$ is the vector of multivariate zero mean uncorrelated white noise process, $\hat{A}_1, \hat{A}_2, \ldots, \hat{A}_p$ are the $k \times k$ matrices of model coefficients ($\hat{A}_0 = I$), p is a model order. This kind of model corresponds to the system of recursive filters acting on white noise inputs. As it is implied in the AR model formalism we assume independence of noise components and the signal. In our approach we treat every signal component, which is not the uncorrelated noise, as a signal. In the case of the EEG time series which have noisy character themselves such an approach seems to be plausible. In spite of the fact that the same model order is used for all channels the advantage of the method lies in the fact that all channels are treated as members of one system and their mutual influences (not limited to two by two relationships) are taken into account. The cofficients $\hat{A}_i$ can be obtained from (1) by multiplying its both sides by x_{t-s}^T (x^T – transposed vector) and taking expectation values. This leads to the equation:

$$\hat{R}(-s) + \hat{A}_1\hat{R}(1-s) + \cdots + \hat{A}_p\hat{R}(p-s) = 0 \qquad (2)$$

where $\hat{R}(s) = E[x_t, x_{t+s}^T]$ is the covariance matrix of lag s for the vector x_t (E means expectation value). Before calculation of the expectation values the data were normalized by subtracting the mean value and dividing by the variance. The covariance matrices were calculated in the range of the orders: 0–p. Coefficients of the model can be found by solving in respect to $\hat{A}_i$ (2) which is a multivariate analog of the Yule-Walker equation. The method of coefficients evaluation and model order choice for multichannel AR process was described in Franaszczuk et al. (1985). In our previous experience with multidimensional AR model we have found that AIC criterion gives an adequate model order estimation. The application of this criterion in the present investigation gave a satisfactory result as well.

In order to estimate the spectral properties of the investigated process the signal has to be transformed to the frequency domain. For a discrete signal this is performed by means of the application of the Z transform to both sides of (1). Denoting by capital letters the transformed functions, we obtain:

$$\hat{X}(z) = \hat{H}(z)\hat{E}(z) \qquad (3)$$

where $\hat{H}(z)$ is a *transfer function* of our system, z^{-1} is a unit delay operator which can be given in the form:

$$z^{-i} = \exp(-i2\Pi f\,\Delta t) \qquad (4)$$

$$\hat{H}(z) = \left(\sum_{j=0}^{p} \hat{A}_j z^{-j}\right)^{-1} \qquad (5)$$

The element $H_{ij}(f)$ of matrix $\hat{H}(f)$ describes the connection between the j-th input and i-th output of the system.

The spectral matrix of our process can be written in the form:

$$\hat{S}(f) = X(f)X(f)^* = \hat{H}(f)\hat{S}_\theta(f)\hat{H}^*(f) = \\ = \hat{H}(f)\hat{V}\hat{H}^*(f) \qquad (6)$$

(asterisk denotes transition and complex conjugation of vectors). $\hat{S}_\theta(f)$ is a spectral matrix of the input process. Since we have assumed that input noise is white, $\hat{S}_\theta(f)$ does not depend on frequency and we can write: $\hat{S}_\theta(f) = \hat{V}$.

The fact that $\hat{H}(s)$ is not a symmetric matrix means that the transmission from channel i to channel j is different than from channel j to channel i. This observation suggests that the elements of $\hat{H}(s)$ can be a measure of the directional information flow that we are seeking. Nevertheless, in order to be able to compare the results obtained for different data strings, a normalization is needed. This is performed by division of the elements $H_{ij}(f)$ by the squared sums of all the elements of the relevant row.

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^{k} |H_{im}(f)|^2} \qquad (7)$$

$\gamma_{ij}(f)$ – the directed transfer function has the properties similar to coherence. The normalization condition takes the form:

$$\sum_{n=1}^{k} \gamma_{in}^2(f) = \frac{\sum_{n=i}^{k} |H_{in}(f)|^2}{\sum_{m=1}^{k} |H_{im}(f)|^2} = 1 \qquad (8)$$

which means that the sum of the contributions from all the input channels $n = 1, \ldots, k$ to the i-th output channel is equal to 1. Function γ_{ij} defined in this way takes values from the interval $[0\ldots 1]$. γ_{ij} could be called *directed coherence*, but *directed transfer function* (DTF) seems to be be a more appropriate name. The value of this function depends on the transmission magnitude. If the AR model describes well the investigated time series, the value of γ_{ij} defined above should be a good estimator of direction and magnitude of the information flow. This can be confirmed by checking the values of matrix $\hat{V}$. Small values of the residual noise matrix $\hat{V}$ mean that this is indeed the case. From our previous experience (Blinowska et al. 1981; Franaszczuk et al. 1985; Blinowska et al. 1985) with AR model and also from the observed values of $\hat{V}$ matrix elements it follows that AR-model indeed describes the EEG time series well, providing that the signal is free of artifacts.

In our definition of DTF the noise matrix $\hat{V}$ is not present, therefore possible correlation of input noises among themselves, manifested in presence of non-diagonal elements in $\hat{V}$, should not influence our estimator, which is constructed solely from *transfer function* $\hat{H}(f)$. $\hat{H}(f)$ on the other hand is not sensitive to noise components since we have implied independence between signal and noise. Matrix $\hat{H}$ describes the filtering properties of the system only, not of its input, which is comprised in matrix $\hat{V}$ (3). Therefore, contrary to coherences which contain the contribution from the input noises, *directed transfer function* does not depend on the assumptions concerning the pattern of input noise sources e.g. existence of common noise source for two or more inputs. This statement was verified by the simulation experiments (Sect. 3).

3 The results and discussion

In order to verify the estimator found by us we have applied our method to the simulated and experimental signals. The calculations were performed by means of computer IBM PC AT-386. The signals used to check the method had the form of a white noise (taken from a random number generator) or they were made similar to the EEG. This was achieved by adding white noise to the AR process. In Figs. 2, 3, 4, 5 and 6 the results of the simulations are shown.

In channel 1 (Fig. 2a) the EEG signal was simulated by AR process of order 4. In channels 2 and 3 there were noise time series plus the contribution of a signal sent from channel 1 with some delay. The amplitude of the noise signal was 3 times higher than the amplitude of signal from channel 1. One can see in Fig. 2 the considerable value of DTF $(1 \rightarrow 2)$ and DTF $(1 \rightarrow 3)$. All other DTFs are equal zero. It is interesting that also DTF$(2 \rightarrow 3)$ and DTF$(3 \rightarrow 2)$ are zero. We can compare DTF with the values of *ordinary coherences*, *partial coherences* and *multiple coherences*. The last two estimators have been given in Franaszczuk et al. (1985) as:
partial coherence:

$$\kappa_{ij} = M_{ij}(f)/[M_{jj}(f)M_{ii}(f)]^{1/2}$$

multiple coherence:

$$\mu_i(f) = [1 - \det|\hat{S}(f)|/S_{ii}(f)M_{ii}(f)]^{1/2}$$

where $S_{ii}(f)$ is an element of spectral matrix $\hat{S}(f)$ and $M_{ij}(f)$ is its minor corresponding to the element $S_{ij}(f)$. *Partial coherence* is a measure of the joint variance at frequency f of two signals after the influence of all other signals of the set has been removed. *Multiple coherences* can be understood as the variance at frequency f due to regression on the all signals from the given set. In Fig. 3 *ordinary coherences* and *partial coherences* are shown for the same signals. In the case of *ordinary coherence* considerable values of all of them are observed although coherence between channel 2 and 3 is somewhat smaller. The *ordinary coherences* are unable to distinguish the direct from the indirect connection between

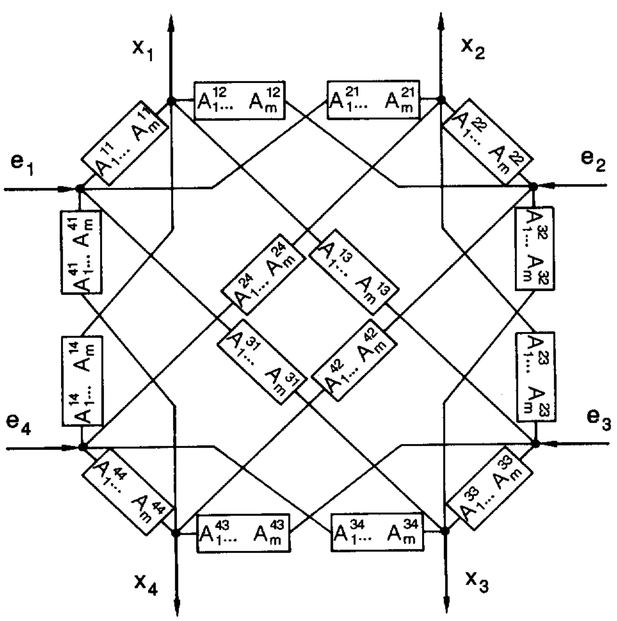

Fig. 1. Schematic representation of the 4-channel AR process. e_i is the input noise signal, x_i is the output EEG time series, and A_l^{ik} is the coefficient of the AR model. The subscript is the index of coefficient. The superscripts correspond to relevant inputs and outputs (From Franaszczuk et al. 1985 with permission from authors)

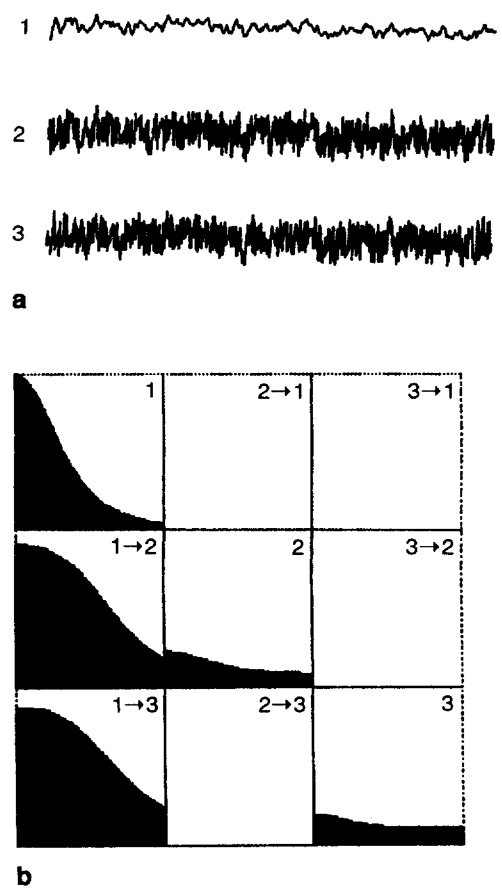

Fig. 2. **a** Simulated time series. *1)* AR process of order two, *2)* white noise plus delayed signal from channel *1*, *3)* white noise plus delayed signal from channel *1*. **b** On the diagonal – power spectra of the signals from **a**. Off-diagonal – *directed transfer functions*, the direction marked by arrows. Numbers in the right upper corners correspond to the relevant channels. The signals have been normalized by subtraction of mean value and division by variance

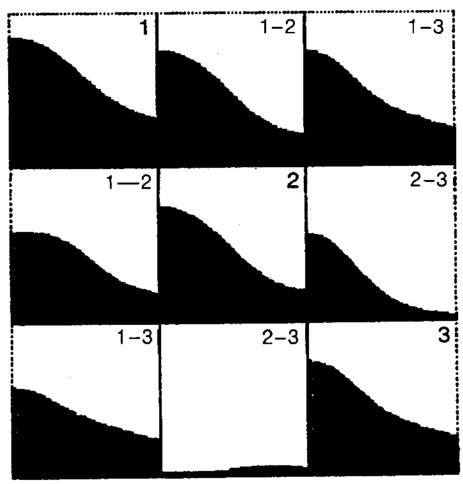

Fig. 3. *Ordinary coherences* above diagonal, *partial coherences* under diagonal, *multiple coherences* on the diagonal – for the signals shown on Fig. 2a. Numbers in the right upper corner correspond to the relevant channels

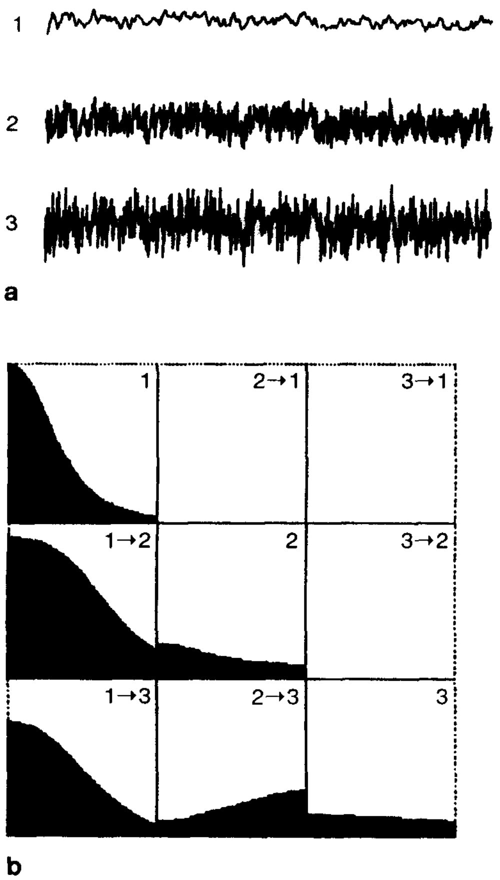

Fig. 4. **a** Simulated time series *1)* AR-process, *2)* white noise plus delayed signal from channel *1*, *3)* White noise plus delayed signal from channel *2*. **b** On the diagonal – power spectra of the signals from **a** normalized as those from Fig. 2b). Off – diagonal *directed transfer functions*

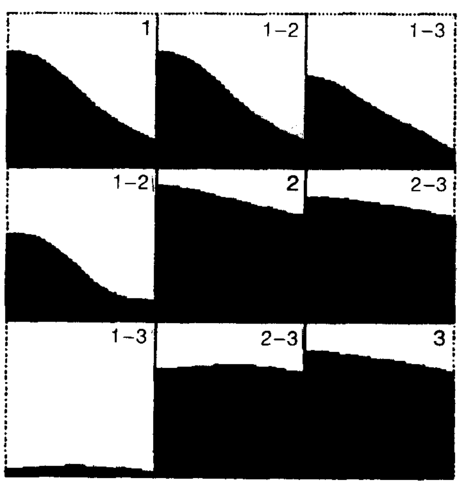

Fig. 5. *Ordinary coherences* above diagonal, *partial coherences* under diagonal, *multiple coherences* on the diagonal – for the signals shown on Fig. 4**a**. Numbers in the right upper corner correspond to the relevant channels

the channels. It is possible to make this distinction by observing *partial coherences* (lower part of Fig. 3). The property of distinction between direct and indirect flow of information is even clearer in the case of the *directed transfer function* (Fig. 2b).

The results of another simulation are shown in Fig. 4a, b. In channel 1 we have again an EEG-like signal. This signal is added with delay to the noise in channel 2, and the time series obtained in this way is sent with some delay to channel 3. We can observe a considerable DTF $(1 \rightarrow 2)$ value, and smaller value of DTF $(1 \rightarrow 3)$. All other DTF values are zero. The values of *ordinary coherences* and *partial coherences* for the same signals are given in Fig. 5. Although in this case *partial coherences* give more clear-cut information on the connections between channels, they don't show the direction of the information transfer. It seems that both estimators should be used in order to get information about the process.

By means of simulation experiment we have confirmed the fact that the input noise common to all channels doesn't influence DTF. In Fig. 6a the DTF pattern for signals simulated by means of AR model of order 5 is shown. The coefficients were estimated from experimental EEG time series; to the inputs of a model white noises were fed. Then to all inputs additional noise the same for all channels was added. Table 1 shows matrix $\hat{V}$ before and after the addition of a common noise components. One can see that off diagonal elements of $\hat{V}$ are increasing. The DTFs for this situation are shown in Fig. 6b. The amplitudes of the *directed transfer functions* in both cases were the same, since DTF does not depend on noise matrix $\hat{V}$ (6 and 7). This feature is very advantageous in the EEG evaluation, since it makes possible to estimate the direction of the brain activity flow in the presence of a common noisy background.

The described method was applied to the physiological signals namely: to the electrical activity registered in the brain structures of the experimental animals and to the human spontaneous EEG. The electrodes were chronically implanted in the following brain structures of a rat: hypocampus, lateral hypothalamus (right and left side), posterior hypothalamus (left side). The signals were registered during locomotion of the animal and when it was standing. The sampling rate was 200 Hz.

The power spectra of EEG signals and *directed transfer functions* are shown in Fig. 7a, b. The results have been averaged over 6–4 epochs of 5.12 s duration. Three animals were investigated and the results were

Table 1. Estimated $\hat{V}$ matrices

before addition of the common noise	after addition of the common noise
$\begin{pmatrix} 0.0195 & 0.0028 & -0.0012 \\ 0.0028 & 0.0170 & -0.0003 \\ -0.0012 & -0.0003 & 0.0072 \end{pmatrix}$	$\begin{pmatrix} 0.0271 & 0.0114 & 0.0067 \\ 0.0114 & 0.0141 & 0.0056 \\ 0.0067 & 0.0056 & 0.0068 \end{pmatrix}$

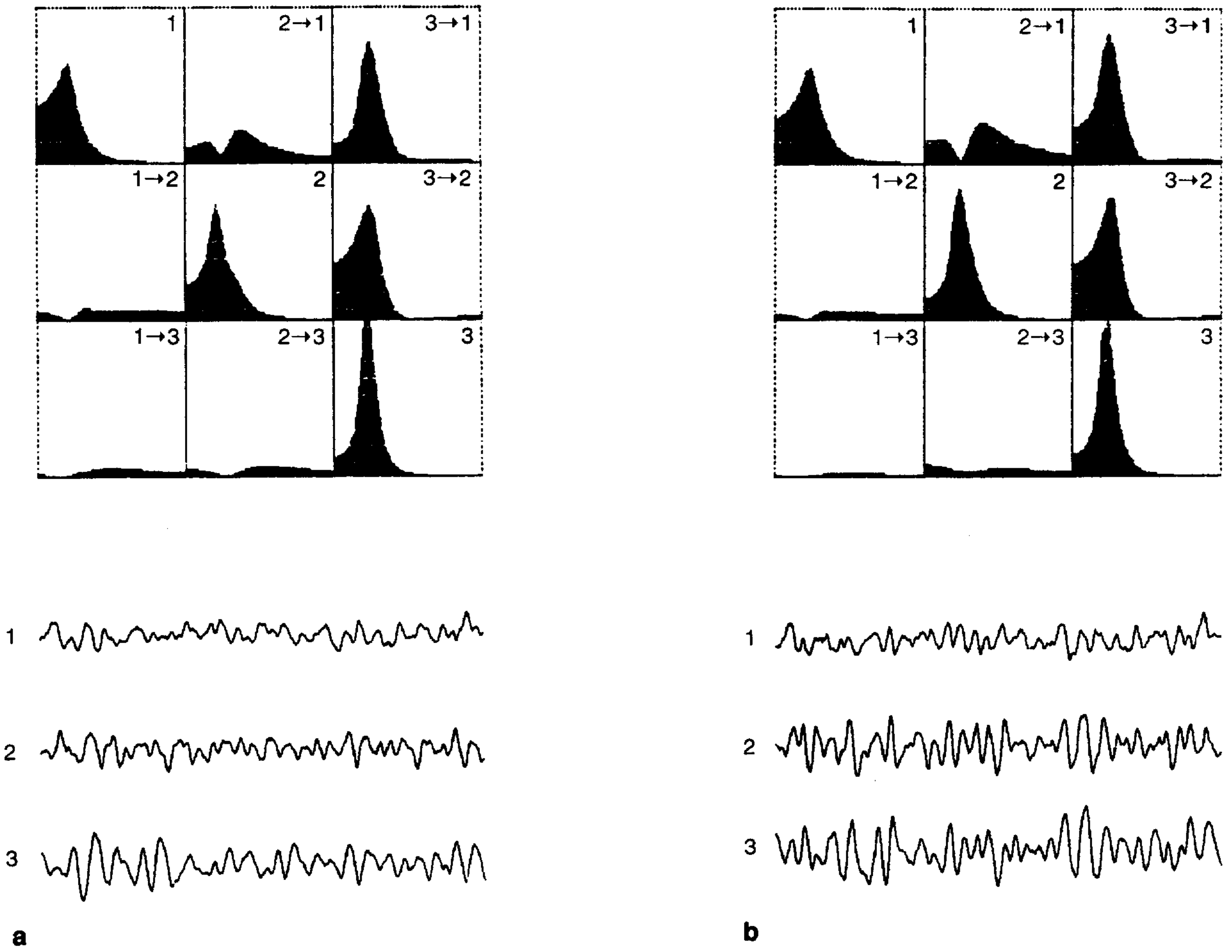

Fig. 6. a Directed transfer functions for the signals shown below. The signals were simulated from AR coefficients estimated from experimental EEG time series. **b** The same as **a**, except that additional common noise was added to the inputs of AR filters

very similar in all of them. When the animal was standing there was no clear difference between DTF($i \rightarrow j$) and DTF($j \rightarrow i$). During locomotion there was a pronounced increase of information flow from hypocampus to the other structures for all investigated animals, as one can judge from the increased values of DTFs in the last column on our picture. One can postulate that during locomotion the hypocampus is driving the lateral and posterior hypothalamus. In order to draw definite physiological conclusions further experimental studies are needed. They are in progress at present. Nevertheless, the obtained results indicate that the method is promising for neurophysiological investigations.

These results encouraged us to apply the method to the clinical data. We have analyzed spontaneous EEG activity (eyes closed) of young healthy males from four leads: F3-C3, F4-C4, P3-01, P4-02 for three patients. Presented in Fig. 8 values of DTF($i \rightarrow j$) are the averages of 4 data segments of 10.24 s duration for a typical subject. It seems that during the relaxed state with eyes closed a prevailing direction of EEG propagation is from back to the front and from the left to the right hemisphere. This follows from the asymmetries between values: DTF ($3 \rightarrow 1$) and DTF ($1 \rightarrow 3$), DTF ($4 \rightarrow 2$) and DTF ($2 \rightarrow 4$) (first ones of the pair being significantly greater) and differences between DTF ($1 \rightarrow 2$) and DTF ($2 \rightarrow 1$), DTF ($3 \rightarrow 4$) and DTF ($4 \rightarrow 3$). We shall not discuss this problem in detail here since the aim of our work is a methodical one.

The determination of the direction of brain electrical activity flow is important for the epilepsy study. In order to test the applicability of our method to the determination of the elileptic foci we have performed simulation experiment. We have described the shape of epileptic waveform by the appropriately chosen analytical functions (Fig. 9a). Then we have constructed an epileptic time series by means of the addition of the noise to the randomly distributed waveforms obtained in the previous step. The signals in channel 2 and 4 were delayed in respect to channel 1, and channel 3 was further delayed in respect to channel 2. The results indicated the following activity flow (Fig. 9c): DTF ($1 \rightarrow 2$), DTF ($2 \rightarrow 3$), DTF ($1 \rightarrow 4$), DTF ($1 \rightarrow 3$) but not vice versa, as was expected. The situation is similar to the ones shown in Fig. 2b and 4b.

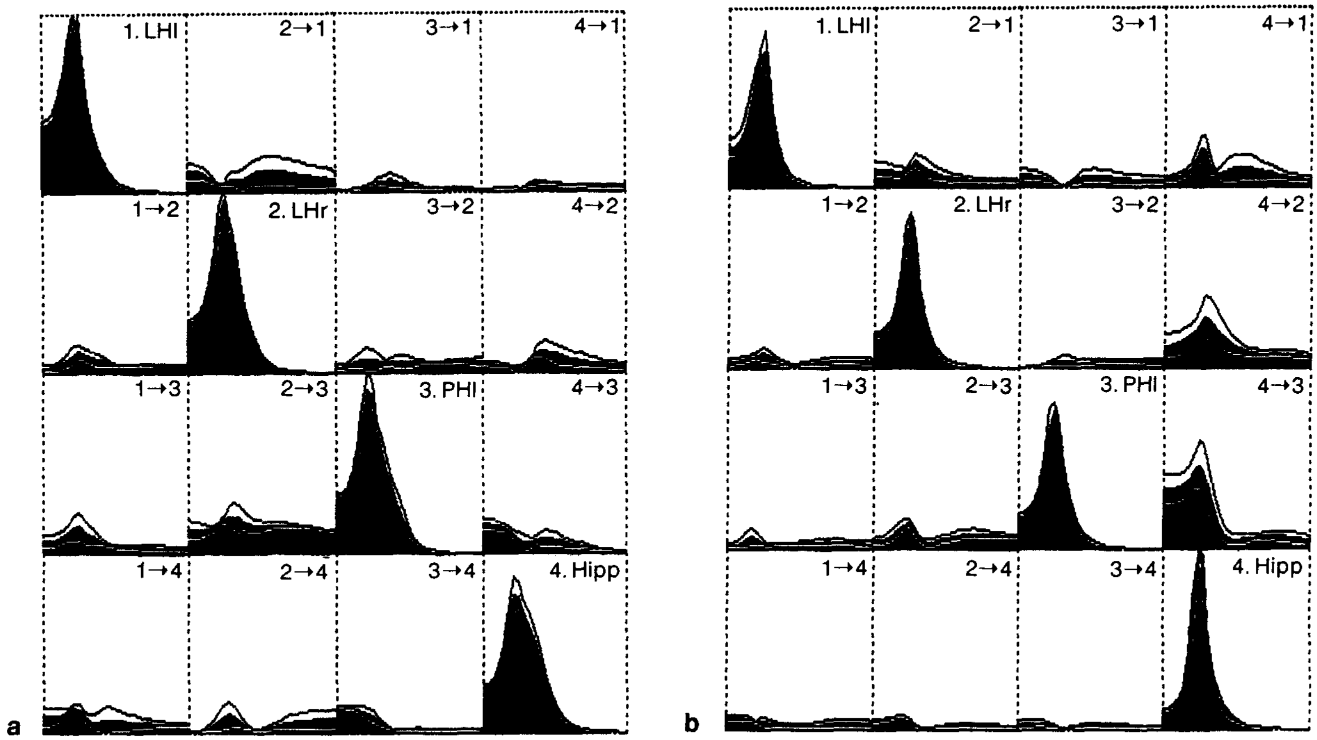

Fig. 7. **a** The *directed transfer functions* (off-diagonal) and power spectra (on the diagonal) for the signals registered in the brain structures of a standing rat. The numbers in the upper right corners correspond to the channels: *1)* Lateral Hypothalamus (*right*), *2)* Lateral Hypothalamus (*left*), *3)* Posterior Hypothalamus, *4)* Hypocampus. Horizontal scale: 0–30 Hz. Corridors of errors are shown. **b** The same as **a**, except that the animal is walking

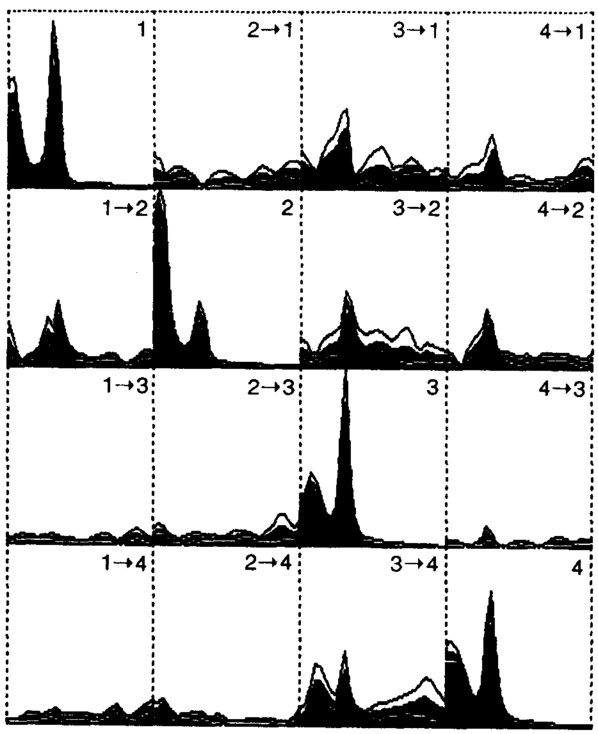

Fig. 8. The *directed transfer functions* (off-diagonal) and power spectra (on the diagonal) for human spontaneous EEG (eyes closed). Numbers in the right upper corners refer to the EEG leads: *1)* F3-C3, *2)* F4-C4, *3)* P3-O1, *4)* P4-O2. Horizontal scale 0–30 Hz. Corridors of errors are shown

In the next step of our analysis we have applied DTF method to the clinically registered epileptic signals. The results are shown in Fig. 10. The pattern of *directed transfer functions* reveals increased outflow of activity from electrode F3 and C4. In order to make any detailed conclusions further analysis is needed. Especially the influence of the position of reference electrode has to be investigated. We intend to study the pattern of *directed transfer functions* for the source of well known localization, which is possible to obtain for the animal study. It seems that there is a lot of interesting information in DTF, which can be extracted by the careful analysis.

The presented results demonstrate the usefulness of our approach for the analysis of the information flow in the brain. The method is being applied to a wide range of neurophysiological data from different brain structures and to the EEG of schizophrenics in comparison with the normal populations. The results will be presented elsewhere.

4 Conclusions

The material described in this paper offers new possibilities of determining information flow in brain structures. The direction of electrical activity spread and its frequency content are described in a quantitative way. The method seems to be a powerful one since it makes possible to distinguish the transmitted signal, even if it is embedded in the noise of the amplitude several times greater than the signal itself, as was demonstrated in the simulation experiments. *The directed transfer function* detects the connection between two structures only, if there is a link between them. It discriminates against cases when both channels are fed from the same source, but are not interconnected. Therefore it is an analog of *partial coherence* rather and not of *the ordinary coherence*.

The application of the described method to the electrodes implanted in the brain structures of the experimental animal had demonstrated its usefulness in the determination of connections between structures during physiological experiments. Its results are unequivocal in discriminating the ways of brain activity transmission. The sensitivity of the method was demonstrated in case of the analysis of data from scalp electrodes. In spite of the fact that the spread of the electrical activity is smeared by the volume conduction and attenuation in the head structures, clear and reproducible results were obtained. The presented examples of the application of the method were evaluated for 3–16 channels only but the algorithm works for any number of channels. The proposed method seems to be a new powerful tool in the investigation of brain processes. It is being applied now to the investigation of EEG activity in the neurophysiological as well as in clinical studies.

Acknowledgements. We are grateful to our colleague Dr P. J. Franaszczuk for the helpful discussions and M. Malinowski for the help in obtaining simulated epileptic signals. We would like to thank Prof. R. Tarnecki, Dr S. Kasicki and Dr W. Szelenberger for making their data available for us.

References

Blinowska KJ, Czerwosz LT, Drabik W, Franaszczuk PJ, Ekiert H (1981) EEG data reduction by means of autoregressive representation and discriminant analysis procedure. Electroencephalogr Clin Neurophysiol 51:650–658

Blinowska KJ, Franaszczuk PJ, Mitraszewski P (1988) A new method of presentation of the average spectral properties of the EEG time series. Int J Biomed Comput 22:97–106

Franaszczuk PJ, Blinowska KJ, Kowalczyk M (1985) The application of parametric multichannel spectral estimates in the study of electrical brain activity. Biol Cybern 51:239–247

Gevins AS (1989) Signs of model making by the human brain. In: Springer Series in Brain Dynamics, vol 4. Basar E, Bullock T (eds). Springer, Berlin Heidelberg New York, pp 408–415

Kamiński M (1988) The EEG signals mutual relationship investigation. MSc Thesis. Warsaw University

Kamitake T, Harashima H, Miyakawa H (1984) A time-series analysis method based on the directed transinformation. Electron Commun Jpn 6:103–110

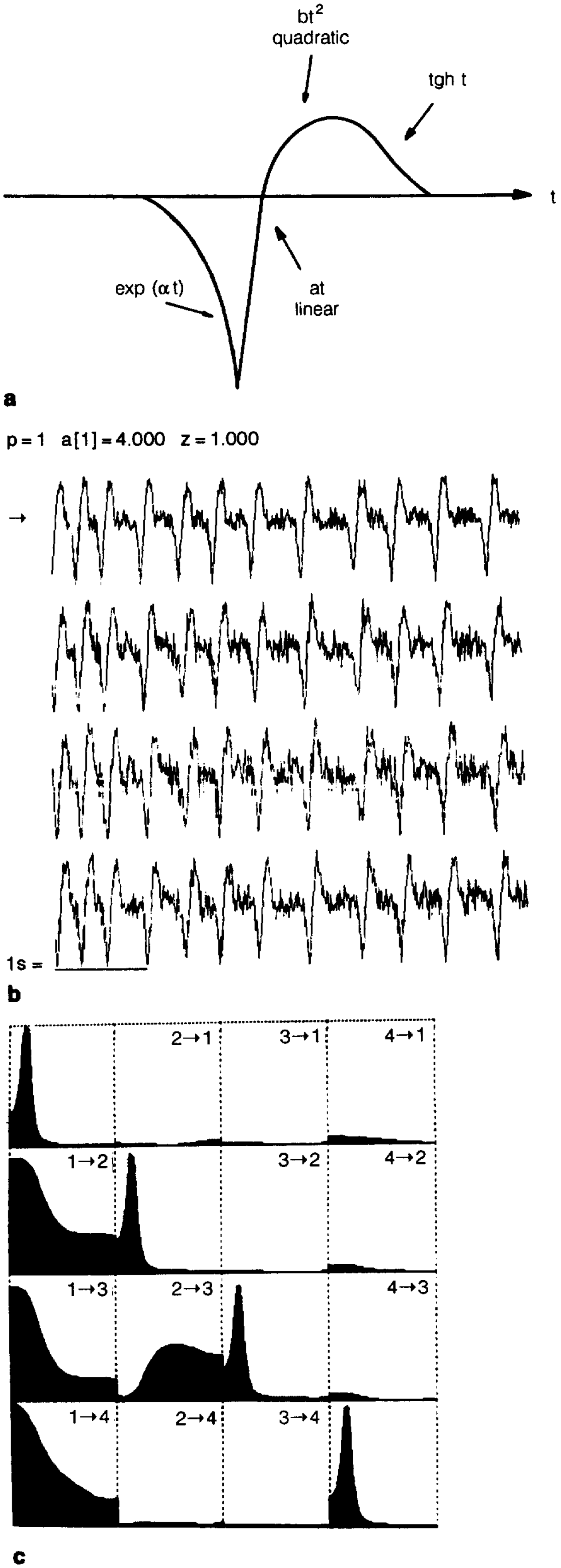

Fig. 9a–c. Simulation of the DTF for epileptic signal. **a** The shape of epileptic waveform described by means of analytical functions. **b** Trains of the epileptic waveforms. The simulated activity flow: 1→2→3 and 1→4 was obtained by adding delays. In each step noise component was also added. **c** The DTFs and the power spectra (on diagonal) for the simulated signals shown in **b**

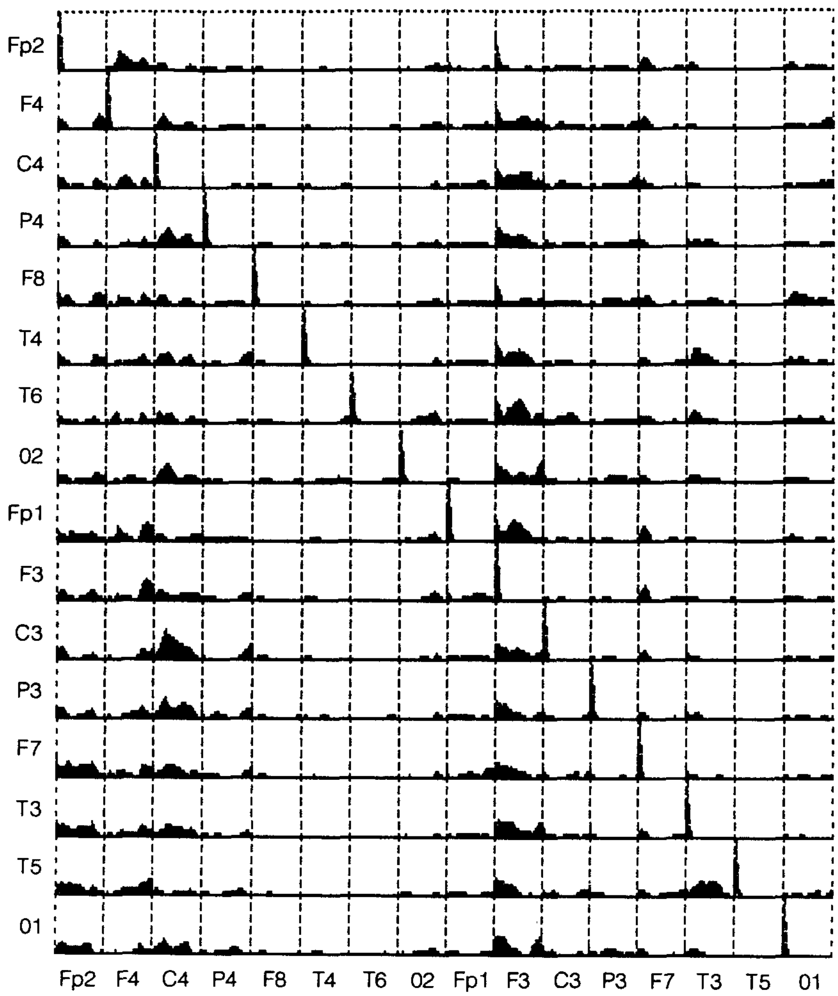

Fig. 10. *Directed transfer functions* and power spectra (on diagonal) for the human epileptic activity. Frequency scale: 0–64 Hz. DTFs in each column correspond to the outflow from electrode marked below that column to the electrode in the relevant row

Lehman D, Ozaki H, Pal I (1987) EEG alpha map series: brain microstates by space oriented adaptive segmentation. Electroencephalogr Clin Neurophysiol 67:271–288

Saito Y, Harashima H (1981) Tracking of information within multichannel EEG record-casual analysis in EEG. In: Yamaguchi N, Fujisawa K (eds). Recent advances in EEG and EMG data processing. Elsevier/North-Holland, Amsterdam, pp 133–146

Schnider SM, Kwong RH, Lenz FA, Kwan HC (1989) Detection of feedback in the central nervous system using system identification techniques. Biol Cybern 60:203–212

Thatcher RW, Krause PJ, Hrybyk M (1986) Cortico-cortical associations and EEG coherence: a two-compartamental model. Electroencephalogr Clin Neurophysiol 64:123–143

Professor Dr. K. J. Blinowska
Medical Physics Laboratory
Institute of Experimental Physics
Warsaw University
ul. Hoza 69
Pl-00-681 Warszawa
Poland