Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

Circular Correlation Methods for the Analysis of Oscillations in Dependent Time Series

Uwe Jentsch and Galina Ivanova, *Member*, *IEEE*

Abstract—**Phases of signals can be regarded as circular data. Time series of phases statistically depend in time. There are also dependencies between diverse circular time series. Instead of relative-phase-methods such as phase-coherence we use a circular correlation coefficient to quantify existing dependence. Based on this coefficient we built correlation matrices. After validation in simulations, we applied the method to phase time series of electroencephalographic data. The method, developed, was positively tested and can be used for oscillation detection and to find couplings between time series considering phase information.**

I. Introduction

MOTIVATION for this work is the analysis of brain oscillations, which are widely described in literature [1]. For this purpose the phases of electroencephalographic signals (EEG) are used. Phases are the oscillation state parameters of time series. By using an amplitude and phase modulation (AM-FM) decomposition of a signal phases can be estimated. The gist in handling phases (ϕ) of a signal is the definition of ϕ between *0* and 2π . One possible procedure is to unfold them by methods of unwrapping. This can be done by adding a manifold of 2π to the data. The problem of this approach is to find out the best estimation for unwrapping. To avoid such problems circular data can be analyzed as such. The phase response of a time series $x(t)$ is a circular time series $\phi(t)$. By default a circular time series has dependencies in time. Additionally dependencies between several time series could exist.

II. Methods

A. The Circular Correlation Coefficient

Dependency measurements for circular time series have a long history in sense of vector correlation calculations ([2], [3], [4], [5]). Phases can be regarded as vectors or complex values with unit radius/magnitude. Therefore many synonyms for phase correlations are used: vector, complex, circumplex, angular, directional and circular measures of dependency. The scientific fields, where such measures are applied are similarly broad: meteorology [6], geography [5], biology [7], physics [8], and chemistry [9]. The most commonly used methods are based on cross-phase or relative phase which is a construct of the bivariate phase (ϕ_X, ϕ_Y) based on the difference $\phi_X - \phi_Y$. Methods by phase-difference range from entropy based measures to first trigonometric moments.

Depending on the field of application the expectation of $e^{j\phi}$ with $j = \sqrt{-1}$ is called: first Fourier mode, mean resultant length, phase coherence, phase locking value, phase locking index, phase consistency, etc. If this measure is applied to the relative phase or cross phase it is not identical to the circular correlation coefficient by some requirements. One example could be that the mean phase coherence is not null for two unimodal distributed phases. The assumption of multimodality is hold by evoked potentials for example.

B. Statistical Properties

According to the target of our study the following requirements on the circular correlation coefficient (CCC) are defined:

I. Domain: The CCC(ϕ_X , ϕ_Y) is defined for all not trivial variables ϕ_X and ϕ_Y .
II. Symmetry: The CCC (ϕ_X , ϕ_Y) = CCC (ϕ_Y , ϕ_X).
III. Boundaries (Normalization): 0≤ CCC (ϕ_X , ϕ_Y) ≤ 1.
IV. Independence: CCC (ϕ_X , ϕ_Y)=0 for all independent ϕ_X and ϕ_Y .
V. Dependence: CCC (ϕ_X , ϕ_Y)=1 for all total, maximal, perfect, functional dependent ϕ_X and ϕ_Y .
VI. Invariance: There is no influence to CCC (ϕ_X , ϕ_Y) if ϕ_X and ϕ_Y are transformed by an orthogonal matrix or a linear vector transformation.
VII. Testing (Distribution): The distribution of CCC (ϕ_X , ϕ_Y) does not depend on the marginal distributions of ϕ_X or ϕ_Y .

If all these properties are satisfied by a CCC it is essential to know the marginal distribution of CCC under the null hypothesis and its quantiles to test for significant dependencies. This enables qualitative and quantitative analysis of circular dependence.

Manuscript received April 2, 2007.
U. Jentsch is with the Technische Universität Ilmenau, Institute of Biomedical Engineering and Informatics, D-98684 Ilmenau, Germany (e-mail:uwe.jentsch@tu-ilmenau.de).
G. Ivanova is with the Technische Universität Ilmenau, Institute of Biomedical Engineering and Informatics, D-98684 Ilmenau, Germany (++49-3677-691312; fax: ++49-3677-691311; e-mail: galina.ivanova@tu-ilmenau.de).

C. Similarity between Circular and Linear Methods

A measurement quantification of the phase correlation can be done by the basic approach of T-linearity/torodial-linearity (see figure 1) or other approaches of describing the relationship between two phases e.g. vector-linearity.
In the phase scatter plot the coil is the analogue to the regression line in a scatter plot of linear data. However scatter plots represent only one correlation coefficient and illustrate the T-linear dependencies.

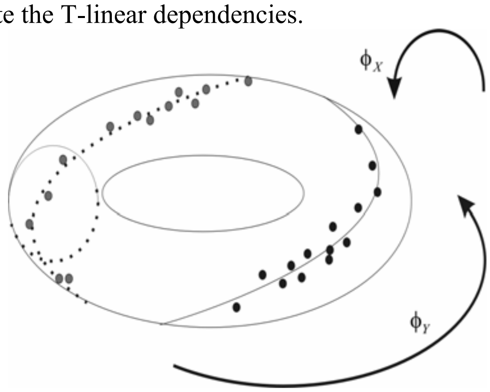

Fig. 1. Scatterplot of phase samples. The first sample ϕ_X is mapped on the smaller circumference and the second sample ϕ_Y is mapped on the larger one; the fitting coil is mapped as curve and dotted if behind the torus; dots behind the torus are in gray

Some formulas of circular correlation coefficients have structural similarity to known product-moment correlation coefficient for linear data. For example, the ordinary correlation coefficient can be rewritten as

$$\rho_{Pearson} = \sqrt{\sigma_{XX}^{-1}\sigma_{XY}\sigma_{YY}^{-1}\sigma_{YX}} \tag{1}$$

where $\sigma_{\bullet\bullet}$ are the variances and covariances of X and Y. This is structurally similar to the circular correlation coefficient described below.

D. Correlation Coefficient of Jupp and Mardia

Due to the high number of circular coefficients [6] and after an eligibility analysis the coefficient of Jupp and Mardia [10] was chosen for the succeeding analysis. This correlation coefficient has been developed by Hooper [11] and Cramer and Nicewander [12] previously, but is indicated in the literature as circular correlation coefficient of Jupp and Mardia.
The CCC of Jupp and Mardia is defined by:

$$\rho = \sqrt{\frac{1}{2} tr(\sum_{XX}{}^{-1}\sum_{XY}\sum_{YY}{}^{-1}\sum_{YX})} \tag{2}$$

Where $\Sigma = \begin{pmatrix} \sum_{XX} & \sum_{XY} \\ \sum_{YX} & \sum_{YY} \end{pmatrix}$ denotes the (4x4) covariance matrix of the transformed phases $X = \begin{pmatrix} \cos(\phi_X) \\ \sin(\phi_X) \end{pmatrix}$ and $Y = \begin{pmatrix} \cos(\phi_Y) \\ \sin(\phi_Y) \end{pmatrix}$ with its submatrices $\sum_{\bullet\bullet}$ of sizes (2x2). The trace of a matrix is denoted by tr().

E. Correlation Matrices

Comparing classically two sets of circular time series, we get a time series of correlations – the correlation function. The canonical way is to correlate all time samples with each other and to receive a matrix of correlations. In the case of autocorrelations the main diagonal becomes 'one' and a kind of waterfall diagram appears. The decay time of correlation defines the dependence in time in the circular time series. In the case of cross correlation between two time series (bivariate case) the highest values occurring build a kind of matrix crest, which is not equal to 'one' by default and does not necessarily lie on the main diagonal. The spread of the crest from the main diagonal describes the time delay between circular time series. The value of the crest quantifies the intensity of interaction of the two time series (maximum cross-correlation).

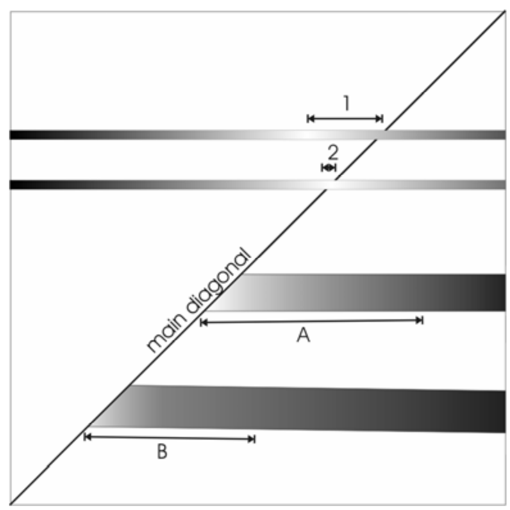

Fig. 2. Illustration of a correlation matrix. White denotes strong dependence (one) and black denotes weak dependencies (zero); long time decay (A) symbolizes durations of dependencies in time series (coherence time) and short decay time (B) denotes shorter latent dependencies; spread of the crest (1) shows a time delay and a crest equal to main diagonal (2) denotes no time delay.

III. DATA

A. Simulation

To identify oscillations and to characterize the set of circular time series univariate methods are applied. We used techniques for detection of the "coherence length" of a set of time series. Additionally to the dependencies in time, the trials have a distribution at each time. For the aim of finding links and bindings we use bivariate methods. Additionally multivariate methods e.g. (like in [13]) can be applied, but we do not want to dwell on these in this paper.

1) Univariate Case

We simulated 40 trials of 650 samples (500 Hz sample rate and time-base from -300 ms to 1000 ms) with an isotropic distribution from -300 ms to 0 ms to model the pre-stimulus series and from 0 ms to 1000 ms with a von Mises distribution of constant mean and concentration 0.8 to reflect phase-locked and time-locked characteristics of evoked activity. The latent dependencies are simulated by a circular moving average filter of order 10. To make contributions to induced characteristics we added a random phase-shift uniformly in the interval of 400 ms to 1000 ms and kept the phases constant.

Then we calculated the auto-correlation matrix for each time sample combinations with a time delay of 20 ms.

The main diagonal achieves the highest value. We defined the time delay needed for dropping below the significance threshold as "coherence time". For example, the plateau in the upper right corner of the matrix in figure 3 is characterized by long "coherence time" with the greatest value at the beginning and a corresponding linear decrease.

2) Bivariate Case

We simulated 28 channels corrsponding to the real data available. In this example all channels are independent and identically distributed. The result of the correlation matrix is visualized in figure 4.

Due to the simulation of independent signals, the cross-correlation observed is nowhere significant.

B. Real Data

In relation to the analysis of EEG-phases a P300 paradigm for cognitive potentials is used. The P300 is an event-related-potential (ERP) and is evoked through an "oddball"-paradigm. The principle of the paradigm, to receive the ERP, is done by the unexpected "target" stimuli. The „target" also defines the start-point in time-base and therefore gives the time-lock. The P300 is spatially expected in frontal and central-occipital areas. There also exist oscillations of induced activity which are non-phase-locked in sense of phases are different at a fixed time or non-time-locked in sense of phases are shifted by time delays in several trials. We try to use the characteristics of longer time-dependencies in induced activity than there are in latent dependencies or evoked activity. We used the 10-20 system of electrode placement with 28 electrodes to standardize our observations.

As assumed for simulated data, we had a sample rate of 500Hz. The real signals were pre-processed with filtering-, artefact-reduction-, and ICA-methods. The phases were estimated by a complex demodulation algorithm.

The obtained phases belong to a mean frequency of 10Hz and are derived from the EEG-alpha band.

Here we applied the methods to real EEG data with an analyzing time gap of 20 ms.

IV. RESULTS

A. Simulation

1) Univariate Case

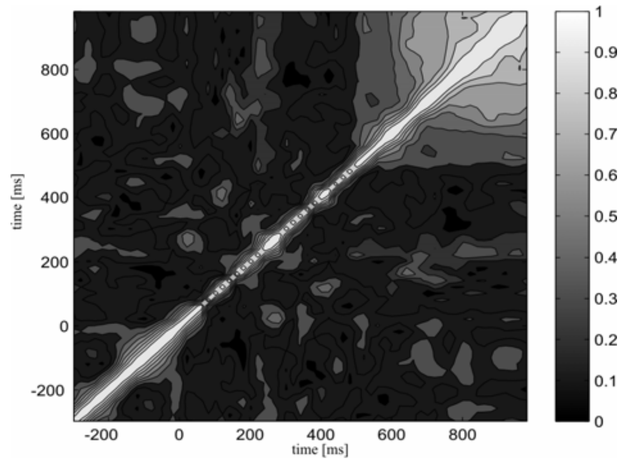

Fig. 3. Auto-correlation matrix of a simulated channel. The prestimulus interval amounts 300 ms and the poststimulus interval 1000 ms. 40 trials are realized (sample length)

In the univariate case (figure 3) we see the increasing time dependency (expected from 400ms to 1000ms) at about 600ms.

2) Bivariate Case

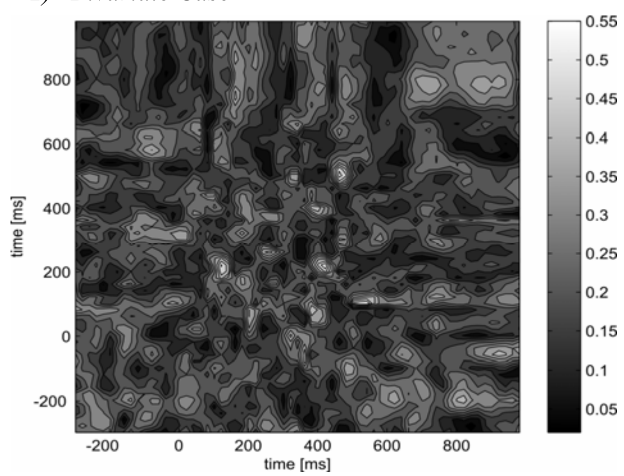

Fig. 4. Cross-correlation matrix of two independent simulated channels.

The bivariate case (figure 4) shows no significant dependency, but a low plateau at about 700ms to 1000ms, because of the nearly constant but independent trials. The first channel is mapped by the horizontal times and the second channel is mapped by its time vertical with the same time and sample length as in the univariate case.

B. Real Data

1) Univariate Case

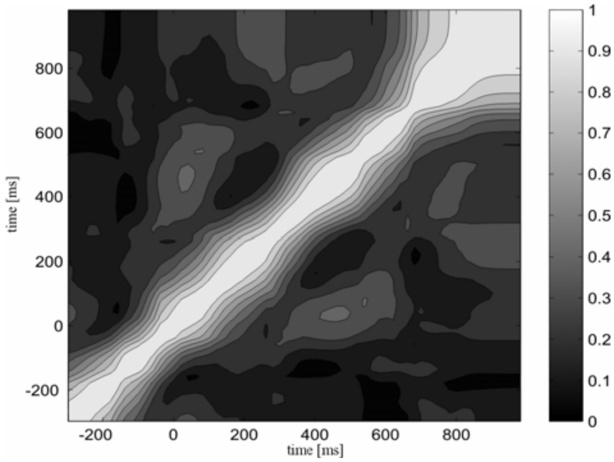

Fig. 5. Auto-correlation matrix of channel C3; with same time and sample length like in simulation

The results show a greater latent "coherence time" with a much clearer plateau (figure 5) as in the simulation (figure 3). The time-base and sample length is equal to simulations.

2) Bivariate Case

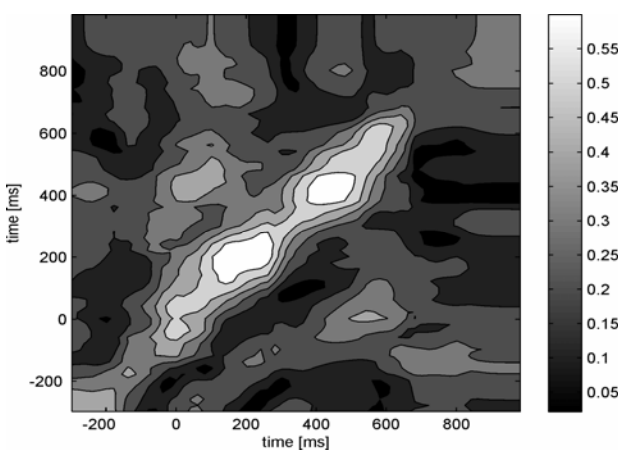

Fig. 6. Cross-correlation matrix of channel C3 and channel F4; where F4 is mapped by horizontal time; C3 is mapped by vertical time

In the bivariate case of correlation of the channels C3 and F4 the correlation matrix (figure 6) shows significant correlations (peaks) at time delays about 200ms and 400ms and a low plateau for about 600ms to 1000ms.

V. Discussion

The presented method, possesses the potential of application in detection of phase-dependencies in time series e.g. the EEG. The circular techniques are more laborious than the common phase coherence methods. The computational load for calculation of a circular correlation coefficient is higher than for procedures using only a phase difference.

While the methods presented here only use the phases, additional enhancements may be developed yielding both: "true" phase and magnitude dependencies. Coherence and phase coherence methods are state-of-the-art.

In case of bivariate methods usually the main diagonal of the presented correlation matrix is used with phase coherence as correlation coefficient. For example Allefeld and Kurths [13] use a sufficient number of such bivariate time series and map them over frequency.

Univariate methods for describing circular time series presented here are not known in the literature. The technique can be regarded as a contribution to the detection of evoked and induced components in EEG data. Even if these results do not describe the physiology underlying these non-linear circular dependencies results clearly demonstrate the presence of phase-dependent components in EEG signals. Additionally, extensive studies over different frequencies and clinical trials data go beyond the scope of this paper and are object of further publications.

VI. Conclusion

A method for multi-trial analysis was developed to point out the "true" phase correlations between and within time series e.g. EEG-signals. Comparing to other prevailing procedures, attention was paid not only to the uni- and bivariate phase time series but also to the circular nature of the data. The method can be applied to other similar directional problems, including directional time or space data. Finally while the here presented methods implicitly use the canonical correlation they can be enlarged to multivariate methods.

References

[1] E. Başar, *Brain function and oscillations*. Berlin: Springer-Verlag, 1999, vol. 2.

[2] R. Dietzius, "Ausdehnung der Korrelationsmethode und der Methode der kleinsten Quadrate auf Vektoren", Sitzungsberichte der kk. Akad. Wiss. Wien, 1916, no. 125(IIa), pp. 3-20.

[3] H.U. Sverdrup, "Über die Korrelation zwischen Vektoren mit Anwendungen auf Meteorologische Aufgaben", Meteorol. Zeit., 1917, vol. 34, pp. 285-291.

[4] H. Hotelling, "Relations between two sets of variates", Biometrika, 1936, vol. 28, pp. 321-377.

[5] M. Masuyama, "Correlation between tensor quantities", Proc. physic.-math. Soc. Japan, 1939, vol. 21, no. 3, pp. 638-647.

[6] J. Breckling, *The analysis of directional time series: Applications to wind speed and direction*, Berlin: Springer-Verlag, 1989.

[7] E. Batschelet, "Circular Statistics in Biology", London: Academic Press, 1981

[8] U. Parlitz, "Synchronisierte Fußgänger", Physik Journal, 2006, vol. 5, no. 1, pp.17-18

[9] Y., Kuramoto, *Chemical oscillations, waves, and turbulence*, Berlin: Springer-Verlag, 1984

[10] P.E. Jupp and K.V. Mardia, "A General Correlation Coefficient for Directional Data and Related Regression Problems", Biometrika, 1980, vol. 67, pp. 163-173.

[11] J.W. Hooper, "Simultaneous equations and canonical correlation theory", Econometrica, 1959, vol. 27, pp. 245-256.

[12] E.M. Cramer and W. A. Nicewander,"Some symmetric, invariant measure of multivariate association", Psychometrika, March 1979, vol. 44, no. 1, pp. 43-54.

[13] C. Allefeld and J. Kurths, "An Approach to Multivariate Phase Synchronization Analysis and its Application to Event-Related Potentials - Synchronization Cluster Analysis -", International Journal of Bifurcation and Chaos, 2004, vol. 14, no. 2, pp. 417-426.