



A LETTERS JOURNAL EXPLORING  
THE FRONTIERS OF PHYSICS

OFFPRINT

**Automated detection of time-dependent  
cross-correlation clusters in nonstationary time  
series**

C. RUMMEL, G. BAIER and M. MÜLLER

EPL, 80 (2007) 68004

Please visit the new website  
[www.epljournal.org](http://www.epljournal.org)

# TAKE A LOOK AT THE NEW EPL

*Europhysics Letters* (EPL) has a new online home at  
**www.epljournal.org**



Take a look for the latest journal news and information on:

- reading the latest articles, free!
- receiving free e-mail alerts
- submitting your work to EPL

**www.epljournal.org**

# Automated detection of time-dependent cross-correlation clusters in nonstationary time series

C. RUMMEL<sup>1(a)</sup>, G. BAIER<sup>1,2</sup> and M. MÜLLER<sup>1,3</sup>

<sup>1</sup> *Facultad de Ciencias, Universidad Autónoma del Estado de Morelos - 62209 Cuernavaca, México*

<sup>2</sup> *Manchester Interdisciplinary Biocentre, The University of Manchester - 131 Princess Street, Manchester, UK*

<sup>3</sup> *EML Research - Schloss-Wolfsbrunnengasse 33, 69118 Heidelberg, Germany*

received 7 August 2007; accepted in final form 19 October 2007  
published online 13 November 2007

PACS **87.19.La** – Neuroscience

PACS **05.45.Tp** – Time series analysis

PACS **89.75.Fb** – Structures and organization in complex systems

**Abstract** – A novel method for the detection of cross-correlation clusters in multivariate time series is suggested. It is based on linear combinations of the eigenvectors corresponding to the largest eigenvalues of the equal-time cross-correlation matrix. The linear combinations are found in a systematic way by maximizing an appropriate distance measure. The performance of the algorithm is evaluated with a flexible time-series-based test framework for cluster algorithms. Attribution errors are investigated quantitatively in model data and a comparison with three alternative approaches is made. As the algorithm is suitable for unsupervised online application we demonstrate its time-resolved use in the example of cluster detection in time series from human electroencephalogram.

Copyright © EPLA, 2007

**Introduction** . – One of the core problems in multivariate time series analysis is the identification and classification of clusters, *i.e.* groups of time series that carry similar information with respect to a given similarity measure. In general, the problem of clustering data has a long history and produced a variety of different solutions, see, *e.g.*, [1,2]. Still a reliable, computationally simple and parameter-free approach would find a wide range of applications for the analysis of physical, social, financial or medical systems. In the last decade various statistical physics methods ranging from Potts spin models [3,4] to maximum likelihood [5] or minimization of a cost function by simple rearrangement of the channel labels [6] have been developed. However, the multitude of the latter algorithms either suffers from an elevated computational effort or requires the definition of some kind of threshold parameter (occasionally fixed by the visual inspection of some statistics). In addition for nonstationary data a continuous adjustment of the parameters may become necessary.

In this letter we propose a method for hard partitional clustering of multivariate time series, where both, number and size of the clusters are determined in a self-contained manner, *i.e.* without (artificial) predefinition of any

(threshold-)parameter. Due to a much smaller parameter space used for optimization the algorithm is computationally faster than other statistical physics approaches to clustering. We restrict our discussion to the example of the equal-time cross-correlation matrix as similarity measure here and discuss possible generalizations [7] in the conclusion. The equal-time cross-correlation matrix is constructed by the prescription  $\mathbf{C} = 1/T \tilde{\mathbf{X}}\tilde{\mathbf{X}}^t$ , where the  $M \times T$  data matrix  $\tilde{\mathbf{X}}$  contains the  $M$  signal channels measured on  $T$  time steps after normalization to zero mean and unit variance. For any finite  $T$  the matrix elements  $C_{ij}$  have the deficiency that the genuine system-specific correlations are contaminated by spurious correlations of size  $\sim 1/\sqrt{T}$ . In the literature such correlations are often denoted as “random correlations”, see, *e.g.*, [8]. Consequently cluster detection algorithms that operate on the basis of the bivariate  $C_{ij}$  run the risk of mis-attributions of channels to clusters due to incidentally large matrix elements. A step towards the solution of this problem consists in using genuinely multivariate measures via diagonalization  $\mathbf{C}\mathbf{v} = \lambda\mathbf{v}$  of the  $C$ -matrix, see, *e.g.*, [8–11]. It has been found that genuine correlations manifest themselves via nonrandom level repulsion [12] between sub-groups of eigenvalues (ordered according to  $\lambda_l \geq \lambda_{l-1}$  in the following).

<sup>(a)</sup> E-mail: crummel@web.de

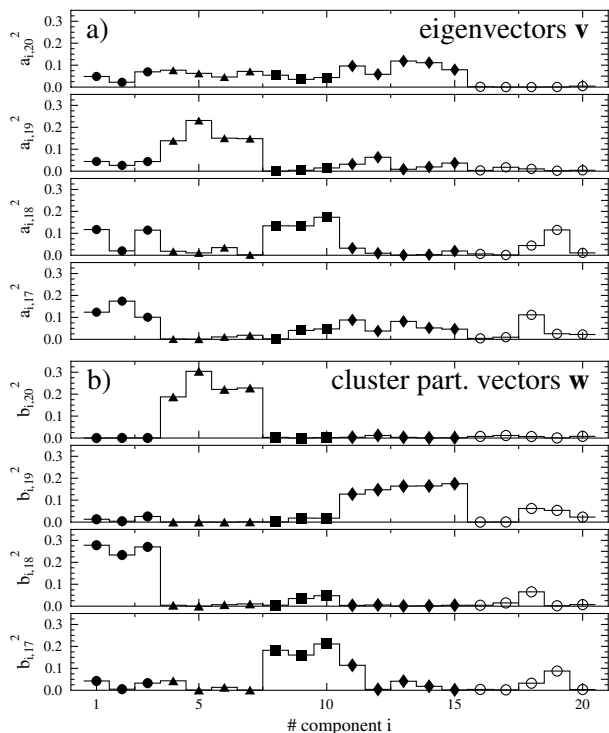


Fig. 1: Comparison of the squared components of the four largest EV  $\{\mathbf{v}_l\}$  of the  $C$ -matrix with those of their linear combinations  $\{\mathbf{w}_l\}$ . The channels that contribute to the  $K=4$  clusters of the model system are marked with full symbols ( $\bullet$ ,  $\blacktriangle$ ,  $\blacksquare$ ,  $\blacklozenge$ ) and the uncorrelated ones with open circles ( $\circ$ ).

Using a block diagonal model for the  $C$ -matrix it can easily be shown analytically that for systems containing  $K$  clusters with  $m_k > 1$  ( $k=1, \dots, K$ ) contributing channels (and sufficiently small inter-cluster correlations) for each cluster one eigenvalue is increased with respect to the uncorrelated situation and  $m_k - 1$  eigenvalues are decreased. Note that different from some other clustering techniques in the present context uncorrelated channels are *not* regarded as clusters of size  $m_k = 1$  but rather as unclustered channels. In this simplified situation the cluster structure can directly be deduced from the eigenvectors (EV) of the  $K$  largest eigenvalues (henceforth abbreviated as “largest EV”), whose nonzero components are restricted to the cluster sub-spaces where they have a flat distribution. In general, in the sense of principal component analysis (PCA) these EV point to the directions of maximal variance in the  $M$ -dimensional space, see, *e.g.*, [13].

Indeed, using  $C$ -matrices constructed from stock return time series, in [8,11] financial market sectors could be identified by looking for large components  $a_{il}^2$  of the  $K'$  largest EV  $\{\mathbf{v}_l\}$ . A refined algorithm that suppresses wrong attributions due to incidentally large components  $a_{il}^2$  was introduced in [14]. However, in many practical situations a clear separation is difficult due to the presence of spurious and inter-cluster correlations, which introduce a mixing of the EV components (see, *e.g.*, [6]). For

illustration we show in fig. 1a the example of a numerical test system (see below) where  $K=4$  groups of size  $m_1 = m_3 = 3$ ,  $m_2 = 4$  and  $m_4 = 5$  are correlated within a total of  $M=20$  time series. The correlation matrix is constructed from a noisy time series of  $T=1024$  data points length. The problem is that all channels corresponding to one of the clusters contribute to the largest EV  $\mathbf{v}_{20}$  with a comparable strength. In the context of financial time series this has been called the “market wide effect” [8,11]. Note that we obtain this behavior even though the separation of  $\lambda_{20} = 1.54$  from the next largest eigenvectors  $\lambda_{19} = 1.24$  and  $\lambda_{18} = 1.21$  is much less pronounced than for typical financial time series where the separation is often one order of magnitude. Based on  $\mathbf{v}_{20}$  a clear distinction is only possible between the total set of clustered channels and the unclustered ones whose squared components are smaller at least by a factor of 5 in the given example. Also the next largest EV  $\mathbf{v}_{19} \dots \mathbf{v}_{17}$  represent mixtures of the four clusters.

In [15,16] *participation indices* (PI)  $p_{il} = \lambda_l a_{il}^2$  were introduced as a step towards the solution of this kind of problems. The PI quantify the strength of a cluster by its eigenvalue  $\lambda_l > 1$  and the contribution of the channels by the components of the corresponding EV. Every channel “ $i$ ” is attributed to the cluster “ $k$ ” with the maximal PI:  $p_{ik} = \max_l p_{il}$ . Finally only clusters corresponding to the  $K'$  largest eigenvalues are accepted, where  $K'$  is generally chosen smaller than the number of eigenvalues larger than unity. This concept works well in situations with a pronounced cluster structure and a strong repulsion of the  $K'$  largest eigenvalues from the “bulk” that is affected by spurious correlations only. Furthermore, it provides a conceptually simple algorithm which is applicable in automated manner and real time. However, it tends to find “pseudo-clusters” [16] from channels that have incidentally large components in low EV, especially when  $K'$  is not fixed adequately in advance. In addition it was found in [16] that the PI are easily misled in situations of equal cluster size and large inter-cluster correlations.

**Cluster participation vectors.** – The basic idea of the approach presented in this letter is to find a set of orthonormal linear combinations  $\{\mathbf{w}_l\}$  —called *cluster participation vectors (CPV)* in the sequel— of the  $K' < M$  largest EV  $\{\mathbf{v}_l\}$  that have dominant entries exclusively for those components which correspond to the cluster “ $l$ ”. At the same time the CPV of clusters “ $l$ ” and “ $n$ ” should have a minimum of common components. Given  $K'$  the CPV can be found by maximizing the *distance measure*

$$D_{ln} = \sum_{i=1}^M |b_{il}^2 - b_{in}^2|, \quad (1)$$

where the  $b_{il}$  are the components of vector  $\mathbf{w}_l$ . It is easily checked that  $D_{ln}$  satisfies all the requirements of a *metric* (symmetry, positive semi-definiteness and the triangle inequality).  $D_{ln}$  is equal to zero for vectors where all components are identical to each other up to a sign and

assumes its maximum value  $D_{ln} = 2$  when the normalized  $\mathbf{w}_l$  and  $\mathbf{w}_n$  have no common components.

Starting from the full set  $\{\mathbf{v}_l\}$  of EV of the  $C$ -matrix the CPV are obtained by a rotation of the  $K' < M$  largest EV about the generalized Euler angles  $\alpha_\kappa$ ,  $\kappa = 1 \dots K'(K' - 1)/2$  such that the sum  $\Delta$  of the mutual distances (1) between the  $\{\mathbf{w}_l\}$  is maximized. An explicit recipe for the construction of the rotation matrix needed in  $K'$  dimensions is given in [17]. Due to the symmetry of the problem under reflection of the axes it suffices to use angles in the interval  $0 \leq \alpha_\kappa \leq \pi/2$  in our case. In order to obtain the global maximum of  $\Delta(\alpha_\kappa)$  in the  $(K'(K' - 1)/2)$ -dimensional space of the  $\alpha_\kappa$  we apply the great deluge algorithm (GDA) of [18] which is computationally fast and gets rarely stuck in local maxima of multidimensional landscapes.

An example for the CPV  $\{\mathbf{w}_l\}$  is illustrated in fig. 1b: Each rotated vector corresponds to one cluster and a clear distinction between clustering and uninvolved data channels becomes possible. On the average the mutual distances (1) have increased about 50% by the rotation. It still remains to answer the question of how the number  $K' < M$  of largest EV used for the described optimization can be determined. In the spirit of the discussion in the introduction it is given by the number of eigenvalues that are repelled from the bulk to larger values. Let us however mention, that a nonoptimal choice of  $K' > K$  does *not* affect the results dramatically. Rather the algorithm focuses automatically on the  $K$  largest EV that correspond to the  $K$  clusters. Qualitatively the CPV  $\{\mathbf{w}_k\}$  (where  $k = M - K, \dots, M - K' + 1$ ) have a structure very similar to the corresponding EV  $\{\mathbf{v}_k\}$  and mutual distances (1) noticeably below the maximum value 2 [7].

**Automated attribution to clusters.** – As we are aiming at a time-resolved application of the algorithm the channels have to be attributed to clusters automatically, once the CPV are found. For the separation of large from small components of  $\mathbf{w}_l$  it is suitable to first arrange the  $b_{il}^2$  in decreasing order. Two measures are candidates for a good separation algorithm: the largest ratio and the largest difference between neighboring ordered  $b_{il}^2$ . Either option alone fails in certain situations. However, using the test framework introduced below the product of both was found to develop a pronounced peak at the correct position in almost all cases. Consequently it allows one to deduce a robust threshold in a self-contained manner and can therefore be used for *automated* attribution. In rare cases this procedure assigns a channel “ $i$ ” to several clusters “ $l$ ”. To establish uniqueness in such cases the cluster with the largest  $b_{il}^2$  is selected.

**A test framework.** – In order to be able to properly evaluate and compare the performance of cluster detection algorithms we set up a time-series-based test framework:  $M$  time series of  $T$  data points length are sampled from independent white noise. The degree of correlations

between and within sub-groups of time series is controlled via common noise components of adjustable strength. To model a system that contains  $K$  clusters the following prescription is used:

$$X_{it} = \left( 1 - \sum_k \rho_{ik} - \sum_{(kk')} \sigma_{i(kk')} \right) \eta_{it} + \sum_k \rho_{ik} \xi_{kt} + \sum_{(kk')} \sigma_{i(kk')} \zeta_{(kk')t}. \quad (2)$$

The strength of the “intra-cluster correlations” is controlled via the parameters  $\rho_{ik}$  which have a nonzero value only if channel “ $i$ ” belongs to cluster “ $k$ ”. Similarly the parameters  $\sigma_{i(kk')}$  control the strength of the “inter-cluster correlations” and are finite only if channel “ $i$ ” belongs to one of the clusters  $k$  or  $k' \neq k$ . Finally  $\eta_{it}$ ,  $\xi_{kt}$  and  $\zeta_{(kk')t}$  denote the individual and common noise components that are drawn independently of a Gaussian distribution with zero mean and unit variance on every time step. In order to have a well-defined cluster structure in the sense that intra-cluster correlations must be stronger than inter-cluster correlations, the conditions  $\sigma_{i(kk')} \leq \max(\rho_{ik}, \rho_{ik'})$  must be satisfied. Moreover, in order to be able to switch between completely correlated and completely uncorrelated channels the value of the bracket in (2) must fall inside the interval  $[0,1]$  for all channels “ $i$ ”.

For a general situation of  $K$  clusters the attribution errors may be sorted into the following four categories: I) channels that contribute to a given cluster are not detected, II) channels are attributed to the wrong existing cluster, III) uncorrelated channels are attributed to an existing cluster, and IV) false clusters (“pseudo-clusters” [16]) are found within the uncorrelated channels. The attribution errors of either category can be measured by the average ratio of mis-attributions over  $N$  realizations of (2) with the same parameters.

**Statistical performance evaluation.** – In fig. 2 we compare the performance of three alternative approaches to the CPV algorithm. First, the method for automated attribution of channels to clusters is used on the basis of the largest EV instead of the CPV. This can be seen as an unsupervised version of the approach used in [8,11]. Second, we compare to the PI algorithm (including the trimming procedure of [15,16] but without the additional concept of bit strings that was introduced in [16] to filter out the most frequent cluster patterns). Third, we use the standard  $k$ -means algorithm for hard partitional clustering of data [1,2] with equal-time cross-correlation as proximity measure. As this algorithm is known to depend on the initial choice of the cluster centroids we checked the performance of the three possibilities beforehand: i) heuristic linear combination of  $K'$  disjoint groups of time series “ $i$ ” on the basis of large  $C$ -matrix elements, ii) use of  $K'$  randomly chosen time series and iii) use of  $K'$  randomly sampled time series. We used

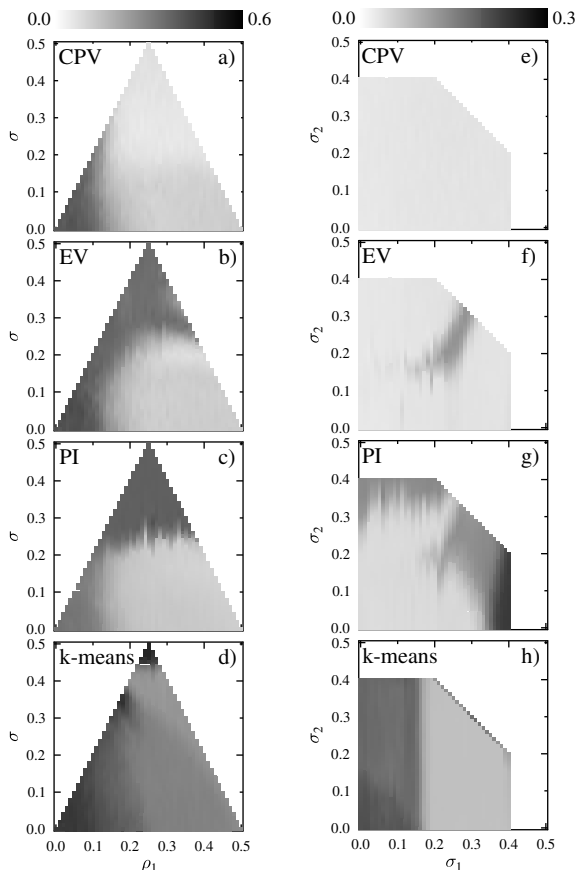


Fig. 2: Percentage of total mis-attributions in the test framework for  $N = 100$  repetitions. Shown is the equally weighted average over all error categories (I to IV). Left column: system with  $K = 2$  clusters of size  $m_1 = 8$ ,  $m_2 = 4$  and intra-cluster correlation given by  $\rho_1 = 0.5\rho_2$ . The  $K' = 3$  largest EV are taken into account. Right column: system with  $K = 4$  clusters of size  $m_1 = 2$ ,  $m_2 = 3$ ,  $m_3 = 4$ ,  $m_4 = 6$  and equal intra-cluster correlations resulting from  $\rho_k \equiv \rho = 0.4$ .  $\sigma_1$  controls the inter-cluster correlations between the first and last two clusters, whereas  $\sigma_2$  couples the clusters 2 and 3. The  $K' = 5$  largest EV are taken into account.

the best performing heuristic initial centroids i) in our comparison.

First we look at a two-cluster situation within  $M = 20$  channels and sample  $T = 1024$  data points from (2). The cluster sizes  $m_1 = 8$  and  $m_2 = 4$  are kept fixed with intra-cluster correlations according to  $\rho_1 = 0.5\rho_2$ . Eight channels do not belong to a cluster. In figs. 2a to d the percentage of total mis-attributions is shown as a function of  $\rho_1$  and the allowed  $\sigma$ . For all methods the same fixed number  $K' = 3 > K$  of large EV is taken into account in order to be able i) to compare the methods on the same footing, and ii) to quantify also the erroneously obtained clusters (error category IV). That is, in the present example we do not take advantage of the automated detection of the number of clusters of the CPV method. The limitation of  $K'$  to a relatively small value  $K' = K + 1$  is in favor of the PI because there the number of wrongly

obtained clusters increases rapidly with  $K' > K$ , whereas the CPV are less sensitive to the size of  $K'$  as mentioned above. For weak total correlations (*i.e.* small  $\rho_1$  and  $\sigma$ ) an increased number of errors is found for all algorithms and the cluster structure is not determined correctly. However, for large  $\sigma$ , *i.e.* large inter-cluster correlations, the CPV (fig. 2a) generate a smaller number of errors than the remaining algorithms. As compared to EV (fig. 2b) and PI (fig. 2c) the main reason is that due to the maximization of  $\Delta(\alpha_k)$  the CPV algorithm commits a greatly reduced amount of errors in category II, compare with fig. 1 and see [7] for a detailed examination of the error categories. Figure 2d displays the advantage of eigenvector-based clustering over our implementation of the simple  $k$ -means algorithm which commits a considerable amount of errors even for large  $\rho_1$  and small  $\sigma$ . The reason is twofold: First, the performance of  $k$ -means depends on the initial guess of the cluster centroids. Second, the algorithm operates directly on the bivariate elements of the  $C$ -matrix. However, it is known that the largest eigenvalues and corresponding eigenvectors made use of by the remaining algorithms are much less contaminated by noise and spurious correlations [8–12,14].

In figs. 2e to h we study  $K = 4$  clusters of size  $m_1 = 2$ ,  $m_2 = 3$ ,  $m_3 = 4$  and  $m_4 = 6$  and equal intra-cluster coupling parameters  $\rho_k \equiv \rho = 0.4$  within  $M = 20$  channels. The inter-cluster coupling parameters  $\sigma_1$  couple the first with the second and the third with the fourth cluster, whereas  $\sigma_2$  couples the second and the third cluster. Five channels do not belong to any of the clusters. Varying both parameters in the allowed region, again, the CPV algorithm (fig. 2e) performs clearly best. A direct operation on the largest EV instead of the CPV (fig. 2f) cannot separate the clusters correctly if  $\sigma_1$  and  $\sigma_2$  become large and comparable. As previously found in [16] the PI (fig. 2g) have problems if inter-cluster correlations become large also here. The  $k$ -means algorithm (fig. 2h) turns out to have problems mainly in separating the smallest clusters from larger ones they are coupled to.

The error of the CPV that remains for large inter-cluster correlations in figs. 2a and e stems exclusively from “pseudo-clusters” (error category IV). We have checked that it vanishes if  $K' = K$  is chosen correctly. Thus, if we allow the CPV algorithm to find the number of clusters automatically, almost no pseudo-clusters are detected. We have also tested other situations, obtaining very similar results.

#### Time-dependent application to human EEG. –

In real-world applications time series are often intrinsically nonstationary and it is not possible to create an ensemble of size  $N$  under stationary conditions that is sufficiently large to allow for error statistics as in the previous examples. In a running window application of cluster detection algorithms in such cases reliable results must be delivered on every time step. In order to test whether this requirement is met by the CPV we have produced artificial time series with a well-defined but rapidly changing

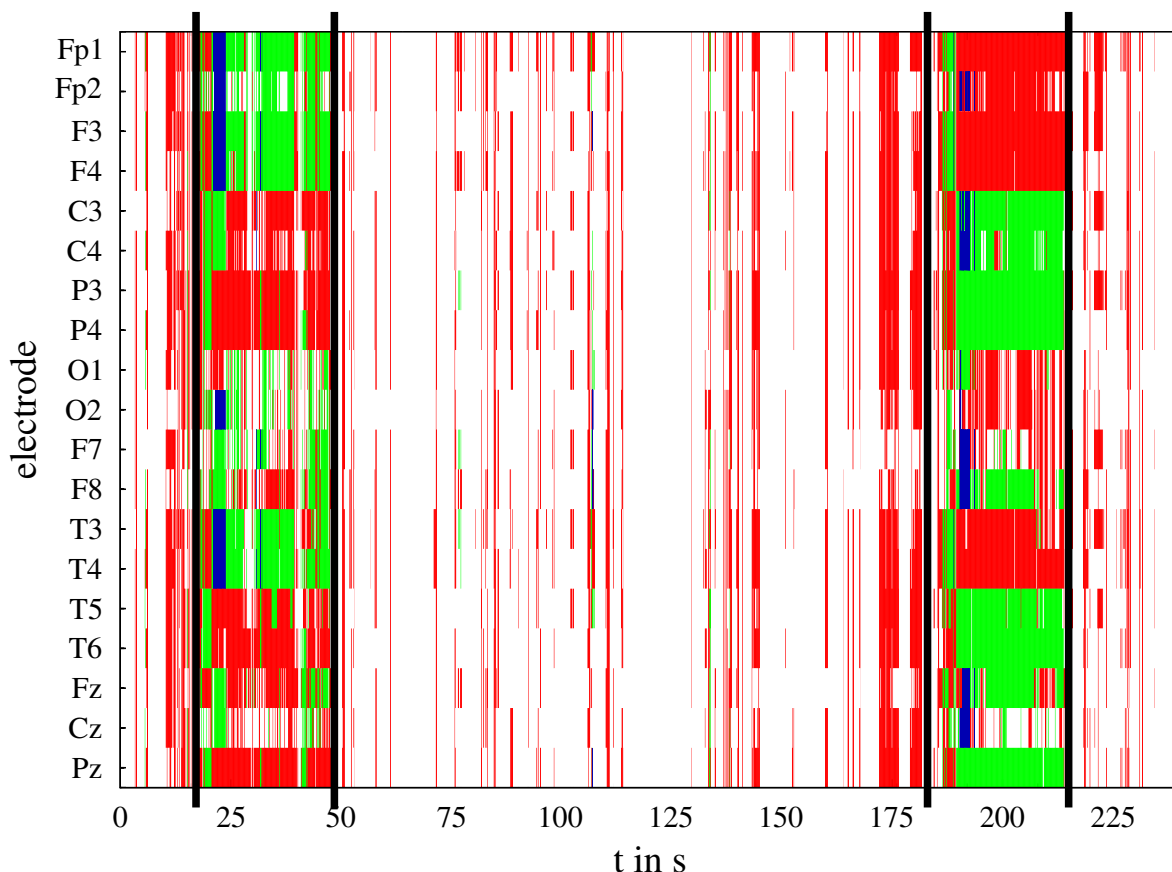


Fig. 3: (Color online) Application of the CPV algorithm to an EEG segment with two absence seizures. Electrophysiological onset and ending of the seizures are marked by thick vertical lines.

time-dependent correlation structure [7]. We found a convincing performance of the CPV in a wide range of window size  $T$ . Due to the reduction of spurious correlations between all channels the results become more significant for larger  $T$ . This improvement, however, has to be bought at the expense of a reduced temporal resolution.

Recently the PI have been applied to time-dependent cluster analysis of long-term recordings of intracranial electroencephalograms (EEG) of epilepsy patients [16]. Here we illustrate the use of the CPV algorithm by showing an application of the CPV to a highly nonstationary scalp EEG containing two primary generalized absence seizures, which start simultaneously in both hemispheres and show the 3/s spike-and-wave activity typical for this kind of seizures. The standard surface EEG with  $M = 19$  electrodes in Hjorth reference was recorded with sampling rate 256/s. The number  $K'$  of largest EV is found dynamically: An appropriately chosen reference interval without seizure activity (here:  $t = 110\text{--}140$  s) is taken to define the 95% confidence interval for each eigenvalue. Using the CPV algorithm the cluster analysis is then made on the basis of those  $K' < M$  eigenvalues that climb above the upper boundary of this interval. Due to the reduced amount of generalized Euler angles, this speeds up the search for the absolute maximum of  $\Delta$  considerably.

In fig. 3 the dynamical evolution of the cluster structure is shown. For the construction of the  $C$ -matrix window lengths in the range  $T = 256\text{--}1024$  data points (1 to 4 s) were tested but no qualitative changes of the results were found. The shown example corresponds to  $T = 512$  (2 s) and a displacement of the running window over the whole set by 32 data points (1/8 s). Channels that contribute to different clusters are marked by different colors.

The first observation from fig. 3 is that in the seizure-free intervals, clusters are detected only rarely and are not stable over time (*e.g.* between seconds 50 and 175). However, the CPV provide a clear and stable pattern of clusters during the seizures. In both cases, two well-pronounced clusters are formed to which the electrodes Fp1, Fp2, F3, F4, T3, T4, and C3, C4, P3, P4, T5, T6, Fz, Pz, respectively, contribute most. This cluster structure is symmetric and very similar in both seizures. Analogous results were found for other data sets from different patients containing absence seizures. In addition, it is remarkable that in both cases one large correlation cluster with involvement of almost all channels can be observed a few seconds before the electrophysiological seizure onset as defined by the method of [19] (black vertical lines in fig. 3). We have checked that this global correlation cluster before seizure onset is not caused by EEG artifacts.

This global cluster then breaks up with the beginning of the seizure. For a short period of time after the initiation of the seizures a third cluster is detected (near seconds 23 and 190). With the disappearance of the third cluster the stable two-cluster pattern described above is manifested.

**Discussion.** – In the present letter we suggested a conceptionally simple cluster detection algorithm that is based on the (multivariate) eigenvectors of matrices constructed from (bivariate) similarity measures for time series. As compared to other statistical physics approaches to clustering [3–6] the method is computationally fast due to the comparably small dimension of the spaces where maximization is carried out. As compared to cluster analysis based on the largest EV directly [8,11,15,16] the construction of CPV has the advantage of better separation of clusters in the presence of considerable inter-cluster correlations.

Although we restricted our presentation to the linear equal-time cross-correlation matrix the CPV algorithm is *not* limited to this measure. It can rather be applied directly to matrices constructed from symmetric and normalizable nonlinear similarity measures like, *e.g.*, mean phase coherence [15,16,20,21] or mutual information [22,23]. In many applications one is interested in the detection of clusters of time series that are correlated at finite time delays. Using the equal-time cross-correlation matrix as an input this is impossible. To the best of our knowledge a genuinely multivariate generalization of the bivariate cross-correlation function does not yet exist for finite time delays. However, mean phase coherence is not sensitive to delays and can therefore be used in such situations.

We checked the performance of the method on artificial data with a well-defined correlation structure. The framework (2) used for these numerical tests turns out useful for a quantitative analysis of the dependence of cluster detection algorithms on parameters like the strength of intra- and inter-cluster correlations or for comparison of different methods. Furthermore we could show that the CPV provide a clear picture of the cluster formation within nonstationary experimental data by using a running window approach. These observations recommend the CPV as a valuable tool for the study of a wide range of nonstationary complex systems. Finally we would like to comment that the generalized Euler angles  $\alpha_k$  can be used to reveal information about the strength of the inter-cluster correlations [7].

\*\*\*

The authors acknowledge fruitful discussions with C. ALLEFELD and S. BIALONSKI. This work was

supported by the Deutsche Forschungsgemeinschaft, Germany (grant RU 1401/1-1) and CONACyT, Mexico (Proj. Nos. 40885-F and 48500).

#### REFERENCES

- [1] JAIN A. K. and DUBES R. C., *Algorithms for Clustering Data* (Prentice Hall, Englewood Cliffs) 1988.
- [2] JAIN A., MURTY M. and FLYNN P., *ACM Comput. Surv.*, **31** (1999) 264.
- [3] BLATT M., WISEMAN S. and DOMANY E., *Phys. Rev. Lett.*, **76** (1996) 3251.
- [4] WISEMAN S., BLATT M. and DOMANY E., *Phys. Rev. E*, **57** (1998) 3767.
- [5] GIADA L. and MARSILI M., *Phys. Rev. E*, **63** (2001) 061101.
- [6] KIM D.-H. and JEONG H., *Phys. Rev. E*, **72** (2005) 046133.
- [7] RUMMEL C., in preparation (2007).
- [8] PLEROU V., GOPIKRISHNAN P., ROSENOW B., NUNES AMARAL L. A., GUHR T. and STANLEY H. E., *Phys. Rev. E*, **65** (2002) 066126.
- [9] PLEROU V., GOPIKRISHNAN P., ROSENOW B., NUNES AMARAL L. A. and STANLEY H. E., *Phys. Rev. Lett.*, **83** (1999) 1471.
- [10] LALOUX L., CIZEAU P., BOUCHAUD J.-P. and POTTERS M., *Phys. Rev. Lett.*, **83** (1999) 1467.
- [11] GOPIKRISHNAN P., ROSENOW B., PLEROU V. and STANLEY H. E., *Phys. Rev. E*, **64** (2001) 035106(R).
- [12] MÜLLER M., BAIER G., GALK A., STEPHANI U. and MUHLE H., *Phys. Rev. E*, **71** (2005) 046116.
- [13] JOLLIFFE I. T., *Principal Component Analysis* (Springer, Berlin) 1986.
- [14] UTSUGI A., INO K. and OSHIKAWA M., *Phys. Rev. E*, **70** (2004) 026110.
- [15] ALLEFELD C., MÜLLER M. and KURTHS J., to be published in *Int. J. Bifurcat. Chaos* (2007).
- [16] BIALONSKI S. and LEHNERTZ K., *Phys. Rev. E*, **74** (2006) 051909.
- [17] HOFFMAN D. K., RAFFENETTI R. C. and RUEDENBERG K., *J. Math. Phys.*, **13** (1972) 528.
- [18] DUECK G., *J. Comput. Phys.*, **104** (1993) 86.
- [19] SCHINDLER K., WIEST R., KOLLAR M. and DONATI F., *Clin. Neurophysiol.*, **112** (2001) 1006.
- [20] MORMANN F., LEHNERTZ K., DAVID P. and ELGER C. E., *Physica D*, **144** (2000) 358.
- [21] MORMANN F., ANDRZEJAK R. G., KREUZ T., RIEKE C., DAVID P., ELGER C. E. and LEHNERTZ K., *Phys. Rev. E*, **67** (2003) 021912.
- [22] QUIAN QUIROGA R., KRASKOV A., KREUZ T. and GRASSBERGER P., *Phys. Rev. E*, **65** (2002) 041903.
- [23] KRASKOV A., STÖGBAUER H. and GRASSBERGER P., *Phys. Rev. E*, **69** (2004) 066138.