Detection of transient generalized and mutual phase synchronization by clustering and application to single brain signals

Axel Hutt
Weierstrass-Institute for Applied Analysis and Stochastics
Mohrenstr.39, 10117 Berlin, Germany
email: hutt@wias-berlin.de

Michael Schrauf
DaimlerChrysler AG, Research and Technology
Information and Communication, 096 / T 728 - RIC/AP
Hedelfinger Strasse 10-14, 73734 Esslingen, Germany
email: michael.schrauf@daimlerchrysler.com

Short title: Detection of synchronization by clustering

Abstract

The present work introduces an analysis framework for the detection of metastable signal segments in multivariate time series. It is shown that in case of linear data these segments represent transient generalized synchronization, while metastable segments in circular data reflect transient mutual phase synchronization. We propose a single segmentation approach for both types of data considering the space-time structure of the data. Applications to both event-related potentials and single evoked potentials obtained from an auditory oddball experiment reveal the lack of the component P300 in an experimental condition, indicates attention effects in component N100 and shows dramatic latency jitters in single trials. A comparison of the proposed method to a conventional index of mutual phase synchronization demonstrates the superiority of considering space-time data structures.

1 Introduction

In the last decades synchronization has been found in various systems in biology, physics or medicine [Pikovsky et al. (2001)]. In neuroscience, synchronization has attracted much attention as a concept of information processing in the brain (see e.g. Singer & Gray (1995)). This approach is in the tradition of results found more than two decades earlier, which revealed cooperative, i.e. synchronized, activity in spatial cortical columns [Wilson & Cowan (1972); Luecke & von der Malsburg (2004)]. In addition, several studies have shown strong correlations between cooperative dendritic activity of neurons and electromagnetic activity on larger spatial scales, e.g. local field detectors or encephalographic potentials and fields [Freeman (2000); Nunez (1995)]. The present study focus to synchronization effects in evoked electroencephalographic potentials.

In neuropsychology, most experiments apply paradigms with several different conditions to gain information about a specific functionality of the brain, as e.g. processing of semantic differences [Kotz et al. (2001)] or prosody in stimuli [Schirmer et al. (2002)]. In order to gain significant results, single experimental conditions are repeated several times. The number of these repetitions depend on the complexity of the task and is typically in the range of $50 - 500$. In case of rather complex experimental paradigms, the number of trials is low. To extract significant results, trials of the same experimental condition are averaged. This procedure is reasonable under the assumption of seldom artifacts, as head movements or low attentiveness of the subject. However, in contrast to most experiments under controlled conditions in a labortaory, more and more experiments are carried out under less controlled every-day-life conditions [Schrauf & Kincses (2003)]. In these cases, only few repeated trials are acquired and artifacts play an important role. To extract significant results anyway, the analysis of averages over few trials or even of single trials is necessary. Several corresponding methods

have been proposed in the last years [Laskaris & Ioannides (2002); Ioannides et al. (2002); Karjalainen & Kaipio (1999)].

One of the major aims of multivariate analysis in neuropsychological research is the detection of functional components from observed data. Lehmann & Skrandies (1980) developed an algorithm to extract spatial activity maps from single data sets, which show metastable synchronized behaviour in time. These time segments are called microstates and reflect functional states in the brain [Brandeis et al. (1995)]. Subsequent work of Pascual-Marqui et al. (1995) and Wackermann (1999) extended this approach by a cluster algorithm and a classification scheme of the extracted components, respectively. The original approach compares the spatial distributions of successive time points and thus is succesfull mainly for single averaged data sets, which exhibit smooth behaviour. In case of less smooth data, the clustering extention is more robust and detects components in more noisy data. However, the method applies the cross-validation method to determine the number of clusters and, hence, fails for high dimensional data.

In the last years many studies analyzed both linear and phasic multivariate data. The former represents the observed data itsself while the latter represent the instantaneous phases extracted from the linear data. The two best-known definitions of instantaneous phases are given by the Hilbert- and the wavelet transformation. Recently, the analysis of phase synchronization between single time series attracted increased attention [Tass (1999); Haig et al. (2000); Lee et al. (2003)]. However, applications to typical encephalographic data need to consider a large set of spatially-distributed detectors as microscopic generators spread their activity on the scalp. Some methods have been developed to extract instantaneous mutual phase synchronization [Haig et al. (2000); Rosenblum et al. (2000)]. However, these methods neglect spatial distributions of phases. In addition, we mention the work of Allefeld & Kurths (2003), who recently developed a sophisticated method which extracts an instantaneous index for mutual phase synchronization by considering the space-time structure of data. However, most methods in this research field considers a high number of trials and, to our best knowledge, the detection of mutual phase synchronization in single data sets is still lacking.

The present work proposes a segmentation framework for both linear and phasic data sets. It extends both the detection of mutual phase synchronization to single trial analysis and the analysis of linear single trial analysis to the treament of phasic data. Our approach considers the spatiotemporal behaviour of multivariate brain signals and aims to extract segments of

metastable behaviour. The key point is the observation that all time series show a mutual change of their time scale, which yield clusters in the corresponding data space [Hutt & Kruggel (2001); Hutt et al. (2003)]. This observation is valid for both linear and phasic data, while clusters in linear data represent generalized synchronization and clusters in phasic data represent mutual phase synchronization.

Hence, a single cluster detection algorithm is applied, while its concrete implemementation depends on the data topology. Linear data behaves on a plane, while phasic data lives on a torus. The following sections show the application to auditory evoked potentials. The results reveal the effects of averaging and the latency jitter of components in single trial subsets. Further, the component N100 exhibits latency differences in two different experimental conditions which indicates early cognitive processing in the brain after 100ms from stimulus onset.

This article is structured as follows. The next section introduces the examined data, the cluster algorithm for both topologies and the applied statistical analysis. Application results follow in Sec. 3. The discussion in Sec. 4 closes the work.

2 Methods

2.1 Clustering of linear data

Let us consider two typical time series $Q_{Fz}(t)$, $Q_{Cz}(t)$ obtained during a cognitive experiment (Fig.1(a)). We observe mutual behaviour of the time series about 105ms, 276ms and 331ms, that is

$$Q_{Fz}(t) = F[Q_{Cz}(t)], \tag{1}$$

where F is a function independant from time. This relation is known in literature in the context of chaotic dynamics and defines generalized synchronization [Rulkov et al. (1995); Pyragas (1996)]. We extend this definition to the N-dimensional case and obtain the definition of global generalized synchronization

$$Q_i(t) = F_{ij}[Q_j(t)] \quad \forall i, j = 1, ..., N, . \tag{2}$$

Figure 2(b) shows both time series as a trajectory in data space. Obviously, the trajectory at the three time points exhibits turning points. Focussing to

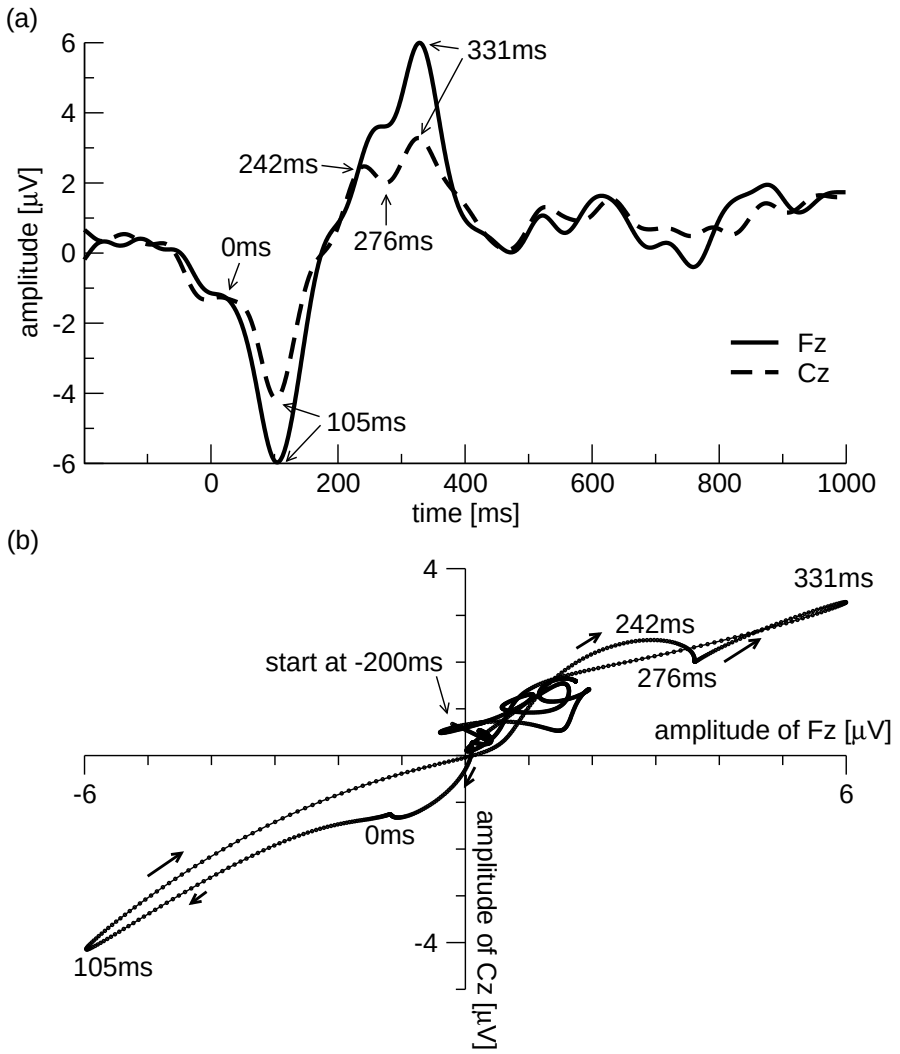

Figure 1: Two typical time series of observed electroencephalographic potentials, here taken at detectors Fz and Cz. They are plotted as single time series (top part) and trajectory in data space (bottom part). The arrows in the bottom part denote the temporal evolution direction of the signal.

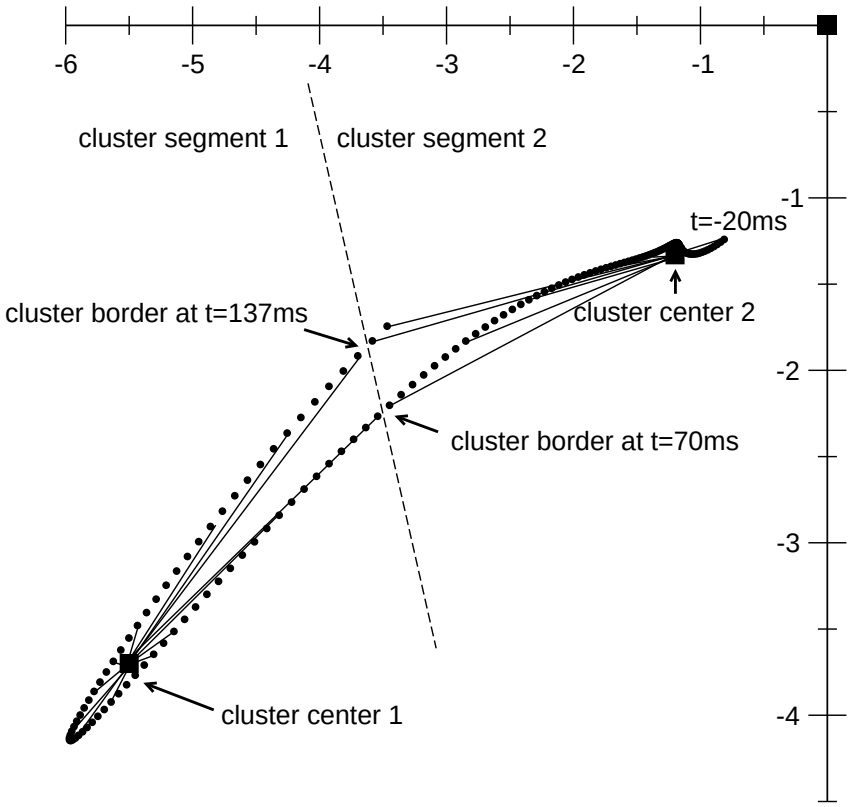

Figure 2: A trajectory segment in the time window $[-20\text{ms}; 139\text{ms}]$. The dashed line illustrates the border of both clusters.

these turning points, data are more dense than between these points. Since turning points of trajectories exhibit vanishing temporal derivatives, Eq.(2) yields

$$\frac{dQ_i}{dt} = \frac{\partial F_{ij}}{\partial Q_j}\frac{\partial Q_j}{\partial t} = 0 \tag{3}$$

$$\rightarrow \frac{dQ_i}{dt} = 0 \quad \forall i = 1, ..., N. \tag{4}$$

for turning points. Indeed, this inreased data density has been found in multivariate data in various studies [Pascual-Marqui et al. (1995); Hutt (2004); Hutt et al. (2000); Hutt & Kruggel (2001); Hutt & Riedel (2003)]. In case of event-related potentials/event-related fields (ERP/ERF), such metastable phenomena have been called differently in literature, e.g. microstates by Lehmann [Lehmann & Skrandies (1980)], quasi-stationary states [Hutt & Riedel (2003)], states of synchronization and desynchronization Pfurtscheller & da Silva (1999) or event-related components in many neuropsychological studies (see e.g. Rugg & Coles (1996)). In addition, we mention the notion of chaotic itinerancy Tsuda (2001); Kay (2003), which models the transients by phase transitions of first order Freeman (2003). Despite differences in these approaches, they describe the mutual decrease and subsequent increase in the time scale of data. In addition, all definitions classify such metastabilities by their latency shift from stimulus onset and their spatial activity distribution at the corresponding latency. In the following, we shall call these phenomena simply components. Subsequently, re-considering the previous discussion components reflect global synchronization.

In case of non-smooth data, mutual behaviour of time series is not that obvious anymore, however trajectory segments of components are assumed to still exhibit an increased data density. In mathematical terms, turning points subject to noise obey

$$Q_i(t) = \bar{Q}_i + \Gamma_i(t) \quad \forall i = 1, ..., N \tag{5}$$

where $\bar{Q}_i = const$ and Γ_i denote random fluctuations with $\langle \Gamma_i \rangle = 0$. Here $\langle ... \rangle$ denotes an average value. Hence, at a constant sampling rate trajectories near turning points obey $\langle Q_i(t) \rangle \approx \bar{Q}_i$ and, subsequently, components represent clusters in data space.

To detect these clusters, we apply the K-means cluster algorithm [Duda & Hart (1973)] which assumes a priorily a fixed number of clusters. Figure 2

shows a trajectory segment extracted from the data in Fig. 1. Two cluster centers have been guessed for illustration reasons. Here data between −20ms and 70ms and the two last data points are nearer to cluster center 2 than to cluster center 1, while the data between 71ms and 137ms belongs to cluster 1. This means the two cluster centers segment the data into three temporal segments, whose borders at 70ms and 138ms are determined by the distance from cluster centers. Now, we apply the K-means algorithm to the data segment of Fig. 2 for $K = 2, K = 3$ and $K = 5$ clusters, respectively. Figure 3 shows the computed squared Euclidean distances from cluster centers to data for the different number of clusters and the plots exhibit the change of nearest clusters and subsequently temporal segments.

The proposed method aims to find a reasonable quantity that distinguishes well-separated from intersecting clusters while taking into account errors by single outliers. This quantity represents the cluster quality of a data point at time t and is defined by the area $a_l(t)$ in Fig. 3, while $l = 1..N_K$ and N_K denotes the number of segments for a fixed number of clusters K. This area between the nearest and the second nearest cluster quantifies both the spatial separation of two segments and its cardinality.

In mathematical terms, the well-known global cost function for K-means clustering and K clusters reads

$$V_K = \sum_{l=1}^{K} \sum_{i \in C_l} (\mathbf{x}_i - \bar{\mathbf{x}}_l)^2 = \sum_{l=1}^{K} \sum_{i \in C_l} d_{ik}^2 \tag{6}$$

where $\bar{\mathbf{x}}_l$ denote cluster centers and C_l are the corresponding sets of members. V_K gives the mean distance of data to clusters and is minimal for the optimal choice of cluster centers. According to the previous discussion, the method extends this formulation to temporal segments S and also considers the distance to the second nearest cluster of each data point. That is

$$V'_K = \sum_{l=1}^{S} \sum_{i \in S_l} (e_{il}^2 - d_{il}^2) = \sum_{l=1}^{S} (N_l - 1)(\sigma_l^{sn} - \sigma_l^n) = \sum_{l=1}^{S} a_l \tag{7}$$

where d_{ik} and e_{ik} denote the Euclidean distance from the data point $\mathbf{x}_i$ to its corresponding nearest and second-nearest cluster center in segment l, respectively. N_l represents the number of data in segment l. Here, a_l is

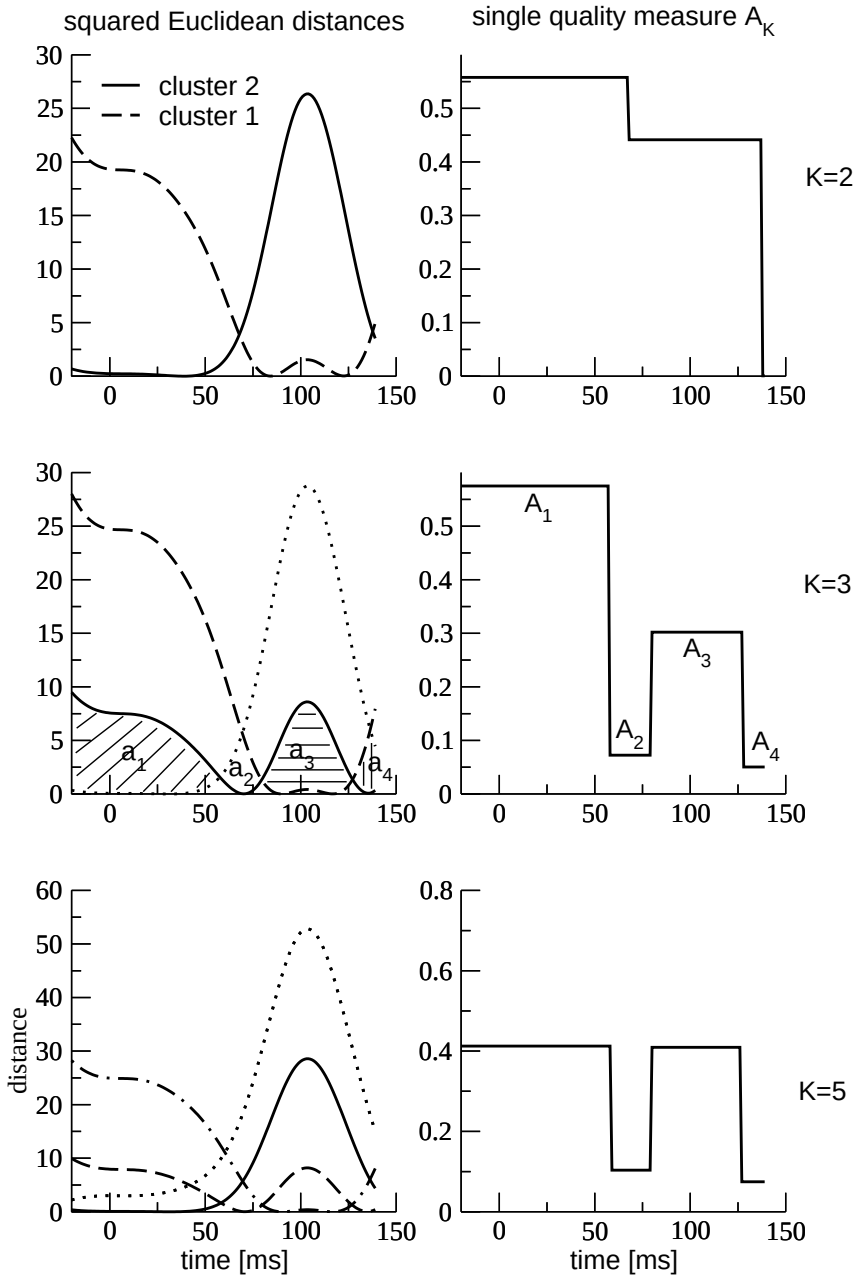

Figure 3: The basic elements of the introduced cluster quality illustrated for number of cluster $K = 2$, $K = 3$ and $K = 5$.

proportional to the difference of cluster variances σ_l^{sn} and σ_l^n between second-nearest and nearest cluster center in segment l,respectively. Now, in contrast to the global approach in (6),(7), the method associates each data point i to the cluster quality of its segment by $A_{li} = a_l I[i]$ with the indicator function $I[i \in S_l] = 1$, $I[i \notin S_l] = 0$. Finally, the normalization of A_{li} and averaging over increasing number of clusters, i.e.

$$\bar{A}_l(i) = \frac{A_{li}}{\sum_{l=1}^{S} A_{li}} \quad , \quad p(i) = \frac{1}{U-2} \sum_{K=2}^{U} \bar{A}_l$$

yields the mean cluster quality p and the present work fixes the maximum number of clusters $U = 20$. Previous studies [Hutt & Riedel (2003)] showed that results are robust towards the value of U if U exceeds the maximum number of expected clusters. According to this definition, large values of p give well-separated clusters, that is well-detected components, while falls and rises mark transitions between different clusters.

2.2 Clustering of circular data

In addition to the analysis of linear data, this section treats phasic or circular data. Several previous studies examined phase synchronization in evoked brain signals [Tass (1999); Allefeld & Kurths (2003); Haig et al. (2000); Breakspear (2002)]. Since a previous theoretical study has shown increased data densities in temporal segments of mutually phase-synchronized data [Hutt et al. (2003)], the extention of the derived cluster algorithm to circular data is straight forward.

Phasic data are physically reasonable only in an associated narrow frequency band. To obtain instantaneous phases from linear data, the present work applies a Gaussian filter in frequency space in combination with a complex Fourier transform [DeShazer et al. (2001)] obtaining

$$S(t) = 2 \int_{-\infty}^{\infty} e^{-(\nu-\nu_k)^2/\nu_\nu^2} \tilde{Q}(\nu) e^{-i\nu t} d\nu. \tag{8}$$

Here, $\tilde{Q}(\nu)$ denotes the Fourier transform of the signal $Q(t)$. Since $\nu_k > 0$, $S(t)$ is complex and the instantaneous spectral power and phase is given by

$$A(t) = \sqrt{\mathcal{I}(s(t))^2 + \mathcal{R}(s(t))^2} \quad , \quad \Phi(t) = \arctan \frac{\mathcal{I}(s(t))}{\mathcal{R}(s(t))} \tag{9}$$

for each frequency band about ν_k, respectively. Here, $s(t) = S(t) - \bar{S}$, $\bar{S}$ is the temporal average of $S(t)$ and $\mathcal{R}(s)$ and $\mathcal{I}(s)$ denote the real and imaginary part of s, respectively. The width of the frequency band is given by the variance of the filter σ_ν^2, which in turn determines the variance of the resulting temporal filter by $\sigma_t^2 = 1/\sigma_\nu^2$ according to the uncertainty principle. The corresponding standard deviation in the time domain represents an estimate for the number of correlated time points and we fix it to $2 \cdot 10$ oscillations, i.e. $\sigma_t = 10/\nu$. Subsequently, filtered data in low frequency bands exhibit higher temporal correlations than data for higher frequencies. In turn, the width of the frequency filter is proportional to the center frequency by $\sigma_\nu = \nu/10$. We mention the equivalence of this approach to the analysis by Morlet wavelets.

According to Pikovsky et al. (2000), mutual phase synchronization(MPS) exhibits bounded differences of phase pairs

$$|\Phi_k(t) - \Phi_l(t)| < \text{const} \quad \forall \quad k = 1, ..., N, \; l = k, ..., N. \tag{10}$$

Hence MPS yields data clusters in the extended space of all phase pairs defined by a new multivariate time series $\mathbf{y}(t) \in \mathcal{R}^M$ with $M = N(N-1)/2$ and $\{y_j(t)\} = \{\Phi_k(t) - \Phi_l(t) \; \forall \; k > l\}$.

There are just two more implementation differences to the linear case, namely the computation of circular distances and the computation of mean circular values. These computations obey basic rules in circular statistics and we refer the reader to Mardia & Jupp (1999) for more details. All subsequent computations of distances, averages and variances of circular data obey these rules.

Summarizing the proposed method for circular data, choose a narrow frequency band, then compute the circular time series by (8) and (9) and compute the new extended data set of phase differences before applying the cluster algorithm as proposed in the previous section.

Similar to the linear case, the obtained cluster quality exhibits large values in case of strong mutual phase synchronization, while sharp falls and rises, respectively, mark transitions between different clusters.

2.3 Statistical analysis

Since the K-means algorithm is iterative and the obtained cluster centers are sensitive to initial values, there is no guarantee that the algorithm converges to the optimal cluster results. Hence, the method repeats the computation of

$p(t)$ 10 times obtaining mean values $P(t)$ and corresponding variances $\sigma(t)$ for each time point t. To verify additionally the cluster results, surrogate data is generated by randomizing the data in time and the re-application of the cluster algorithm yields new mean cluster qualities $P_s(t)$ and corresponding variances $\sigma_s(t)$. The obtained surrogate data set exhibits a decorrelated temporal structure. Subsequently, no prominent cluster segment occurs and $P_s(t)$ is much smaller than in the original data. We shall verify the missing temporal struture by visual inspection, while the lower values of P_s are verified by the t-test for every time point t. The t-value reads

$$T(t) = \frac{P(t) - P_s(t)}{\sigma(t) + \sigma_s(t)} \sqrt{n}, \tag{11}$$

with the degrees of freedom $n = 19$. Equation (11) sets the null hypothesis such that P is indistinguishable from random cluster results P_s. For $T(t) > t_{\alpha,n}$ the test rejects the null hypothesis at an false positive error rate α and P is significantly different from P_s. Here $t_{\alpha,n}$ denotes the Students t-distribution.

In addition, the present work considers a mutual phase synchronization index motivated by Haig et al. (2000) and applied recently by Allefeld & Kurths (2003). It represents the global circular variance

$$R(t) = \frac{1}{L} \sum_{l=1}^{L} \sqrt{\left(\sum_{j=1}^{M} \sin y_{jl}(t) \right)^2 + \left(\sum_{j=1}^{M} \cos y_{jl}(t) \right)^2}, \tag{12}$$

where $\{y_{jl}\}$ are phase differences in trial $l = 1 \ldots L$. $R(t)$ gives a rough estimate of mutual phase synchronization for each frequency band. This index extracts information from trial ensembles and is not applicable for single trial analysis. However, we shall compare our results on single trial averages to results from Eq. (12) in a later section.

2.4 Data acquisition

Event-related potential (ERP) data are analyzed in two conditions of a 2-tone passive oddball paradigm. Tones used were a standard at a frequency of 1kHz and an occurence rate of 0.85 and a deviant tone at a frequency of 2Khz and an occurence rate of 0.15. The tones were presented at a level of 70 dBSL and had a rise and fall time of 10ms, a duration of 50ms with an

inter-stimulus-interval from 3.2s to 3.8s between the start of each stimulus. Tones were played through earphones. ERP recordings were made from 32 sites (electrocap, 10 : 20-system, impedance $< 5k\Omega$, linked mastoid reference) at a sampling rate of 1kHz and amplitude resolution $0.1\mu V$. Hardware filters were applied with the low cutoff at 0.5Hz, the high cutoff at 70Hz and the notch filter at 50Hz.
Topographical scalp current source density(CSD)-maps (order of splines:4, max. degree of Legendre polynomials: 10) were made for comparisons. The frontal (Fz), central (Cz), and parietal (Pz) midline electrode sites were used to facilitate correct identification of the P300 peak (Johnson (1993)). EOG artifact rejection was applied (Gratton et al. (1989)). Data were evaluated offline using a digital low-pass 25Hz filter (e.g. Polich (1998)). Driving tasks (with or without using an active cruse control named distronic) were alternated every 30 min to minimize effects of sequence and attention.
Recordings were analyzed from one physically and mentally healthy subject (male, 45 years, 25 years driving experience, about 50.000 km driven with the Mercedes Benz S-500 test car), with no history of neurological disorder, free of medication and corrected to normal vision. The test route was a 400 km stretch of a german highway (Stuttgart-Duesseldorf). Digital video of forward road scene was recorded for comparison of traffic density and to identify particular variations of traffic scenes.

3 Results

Now, we examine results from both linear and circular data for both experimental conditions. Since the present work proposes an algorithm to examine single data sets, we show results from averages over all trials, from averages over subsets of trials and results from single trials.

3.1 Application to linear data

Figure 4 presents time series from averages over all trials for both experimental conditions. Conventional methods identify the components N100 at $\sim$ 100ms, component P200 at $\sim$ 200ms and component P300 at $\sim$ 300ms for the distronic condition. For the non-distronic condition, we identify the components N100 at $\sim$ 100ms and P200 at $\sim$ 200ms.

The proposed method yields the mean cluster quality $P(t)$ which exhibits

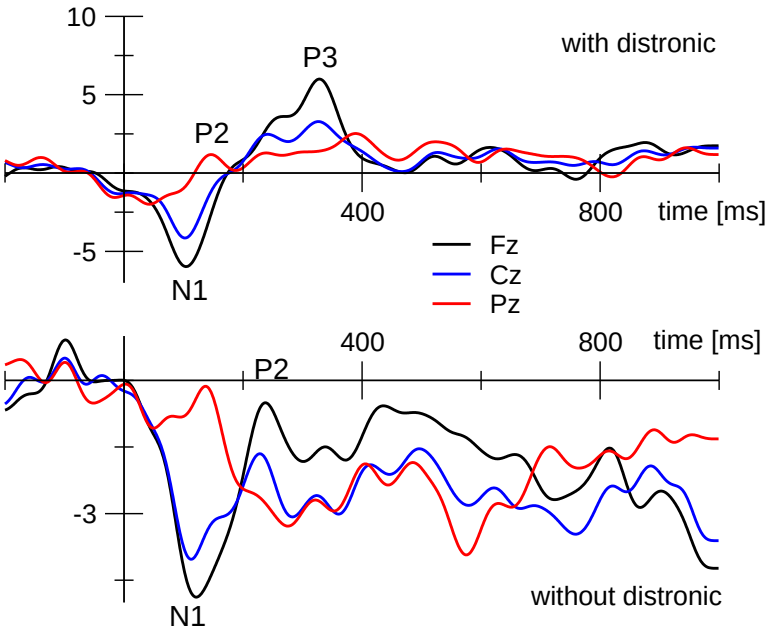

Figure 4: Observed time series at different detectors for both experimental conditions. Conventional methods classify the components N100, P200 and P300 for the distronic condition and N100 and P200 for the non-distronic condition at corresponding temporal latencies.

plateaus of constant values with sharp edges, as shown in Fig. 5a. We observe various clustered segments in time windows coinciding to the conventional results from Fig. 4. Focussing to the time window [0ms; 400ms] and re-applying the method, plateaus and edges in the same time windows occur, while subtle edges in results from the larger time window are more pronounced.

The difference in absolute values occur mainly by virtue of the normalization of P. Figure 5b presents the cluster results of randomized time series exhibiting poor temporal structure and much lower values of cluster quality. Here and in the following, the t-test gives p-values < 0.001 for all time windows and both experimental conditions. That means the cluster results $P(t)$ are significant. Figure 6 shows cluster results for averages over trial subsets in two different time windows for both conditions. It turns out that cluster segments in the data occur in similar time windows as in Fig 5, however slightly shifted, shortend or lengthend. This finding supports the hypothesis of latency shifts in single trials. Figure 7 presents components in single trials, which reveal the latency jitter as well.

Now, we focus to the shorter time window [0ms; 200ms] and classify components by their latencies and spatial distributions. Figure 8 presents results from averages over all trials and average current source density(CSD) maps corresponding to the detected time segments. We identify component the components N100 and P200 in both conditions. In addition, these results reveal a time shift of component N100 between both experimental conditions. Hence component N100 depends on the cognitive task, and thus reflects an endogeneous underlying process. This finding contrasts to the general hypothesis that N100 is an exogeneous component, i.e. independant from the cognitive task. Further results from trial subsets support this finding (Fig. 9), however not such obviously.

We point out that the previously detected components are identified by latency shift and duration only, as the corresponding spatial distributions are noisy and do not allow a clear classification. Future research shall apply spatial denoising of components and we refer the reader to forthcoming work.

3.2 Application to circular data

Now, we examine mutual phase synchronization in the data. Since the phases are defined in a corresponding narrow frequency band, the spectral density $A(t)$ and the global circular variance $R(t)$ are computed to indicate frequency

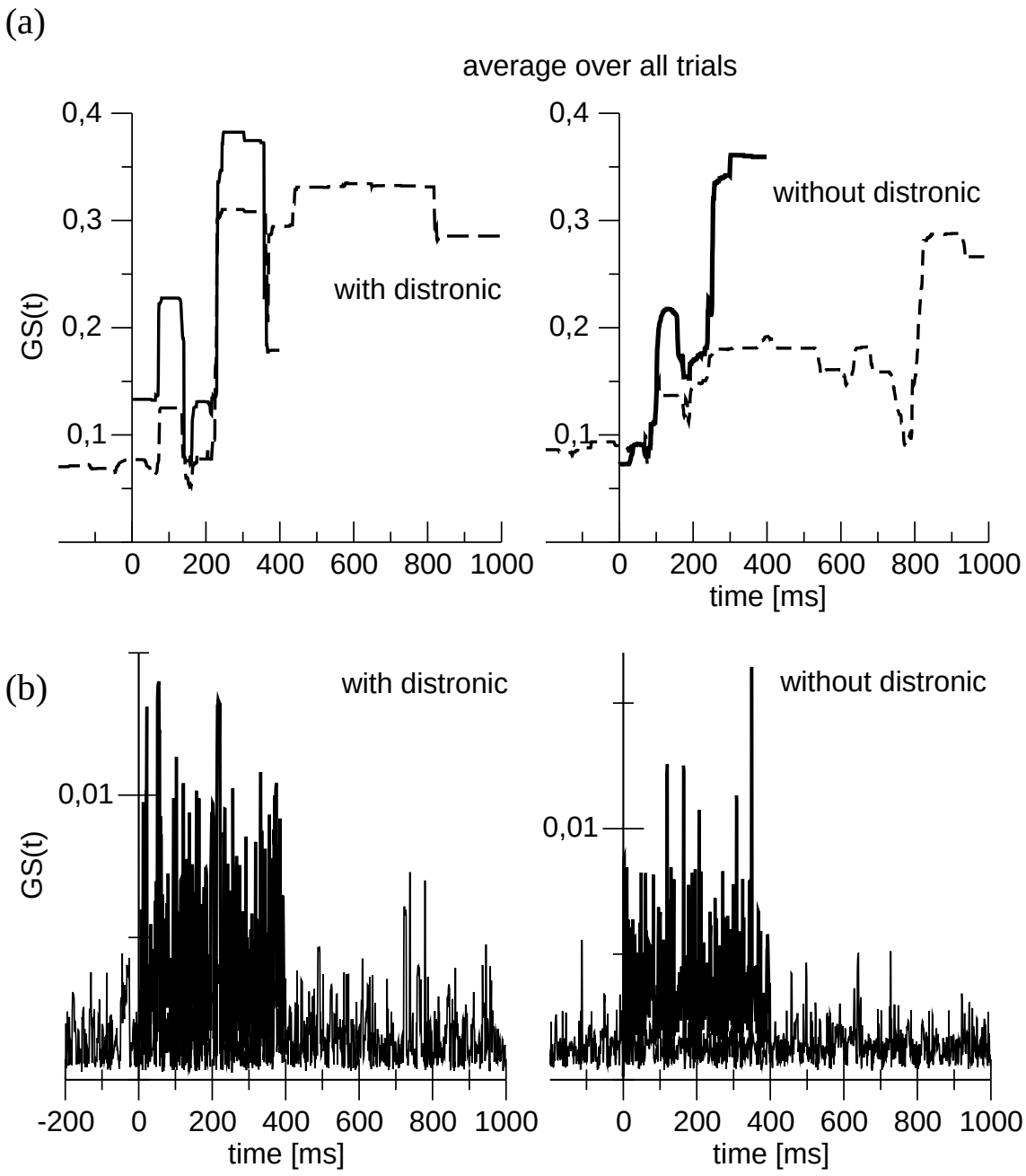

Figure 5: Cluster results for single averages over all trials. The cluster quality p quantifies the generalized synchronization $GS(t) = p(t)$ in time windows $[-200\text{ms}; 1000\text{ms}]$ (dashed line) and $[0\text{ms}; 400\text{ms}]$ for both experimental condition. The top part shows results from the original signal, where we observe a distinguished temporal structure. In contrast, the bottom part presents clustering results from the surrogate time-randomized data, which exhibits a poor temporal structure.

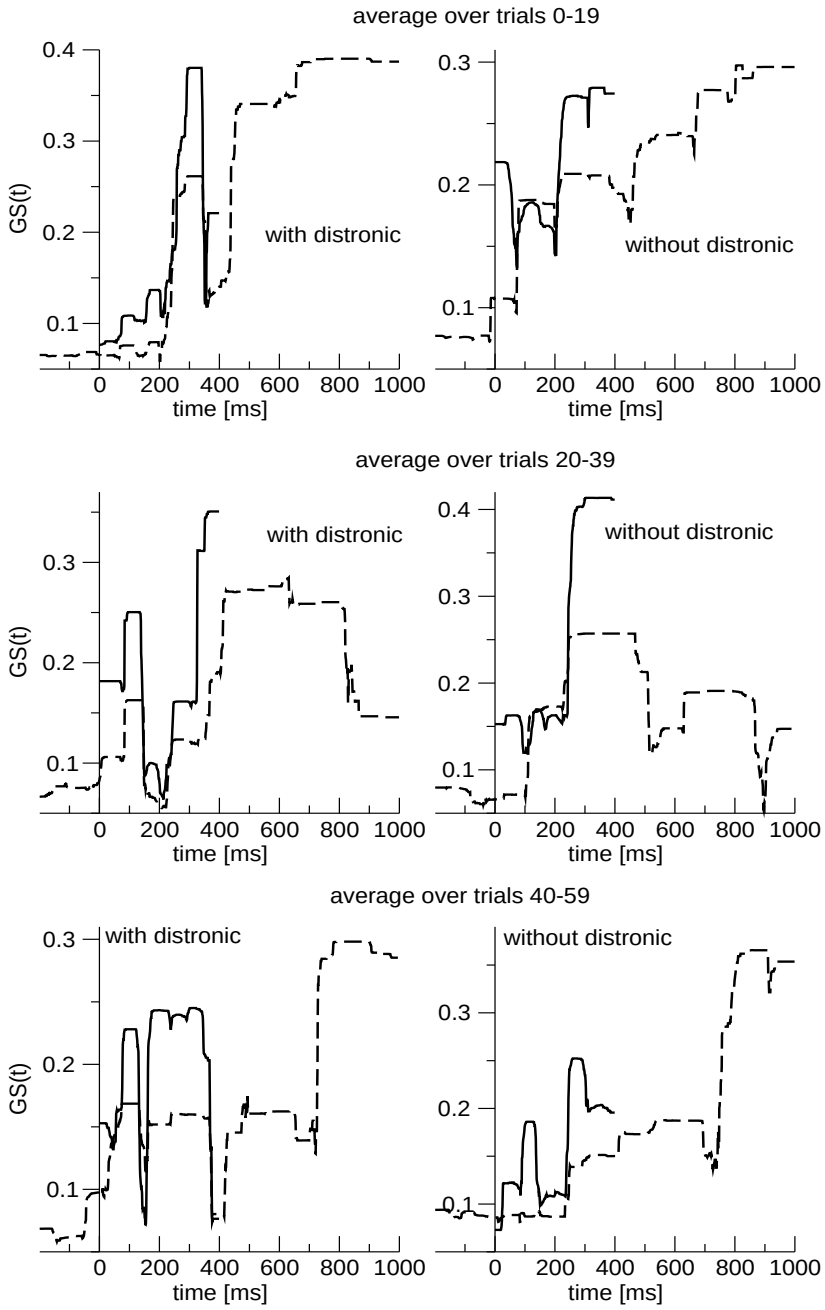

Figure 6: Cluster results from single data sets averaged over subsets of trials in two time windows for both experimental conditions. The cluster quality p quantifies the generalized synchronization $GS(t) = p(t)$ in time windows $[-200\text{ms}; 1000\text{ms}]$ (dashed line) and $[0\text{ms}; 400\text{ms}]$.

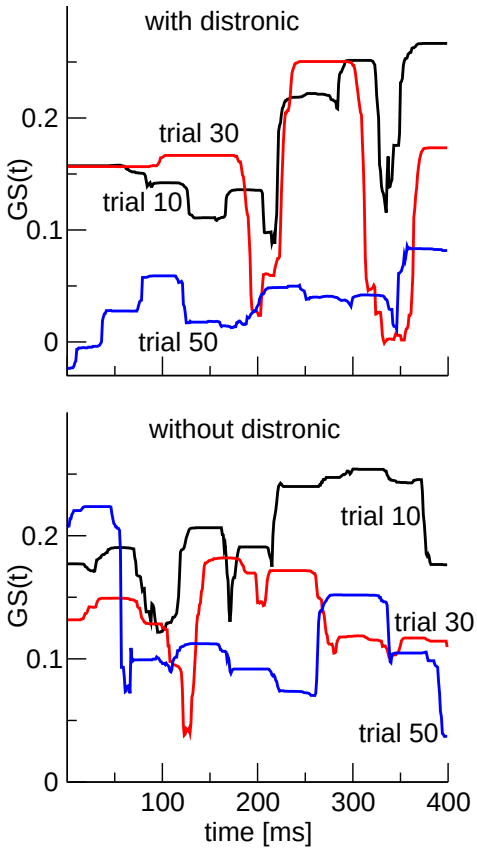

Figure 7: Cluster results from some single trials for both experimental conditions. The specific choice of trials is arbitrary. The cluster quality P quantifies the generalized synchronization $GS(t) = P(t)$. For illustration reasons, results from trials 30 and 50 have been shifted artificially to lower values in both experimental conditions.

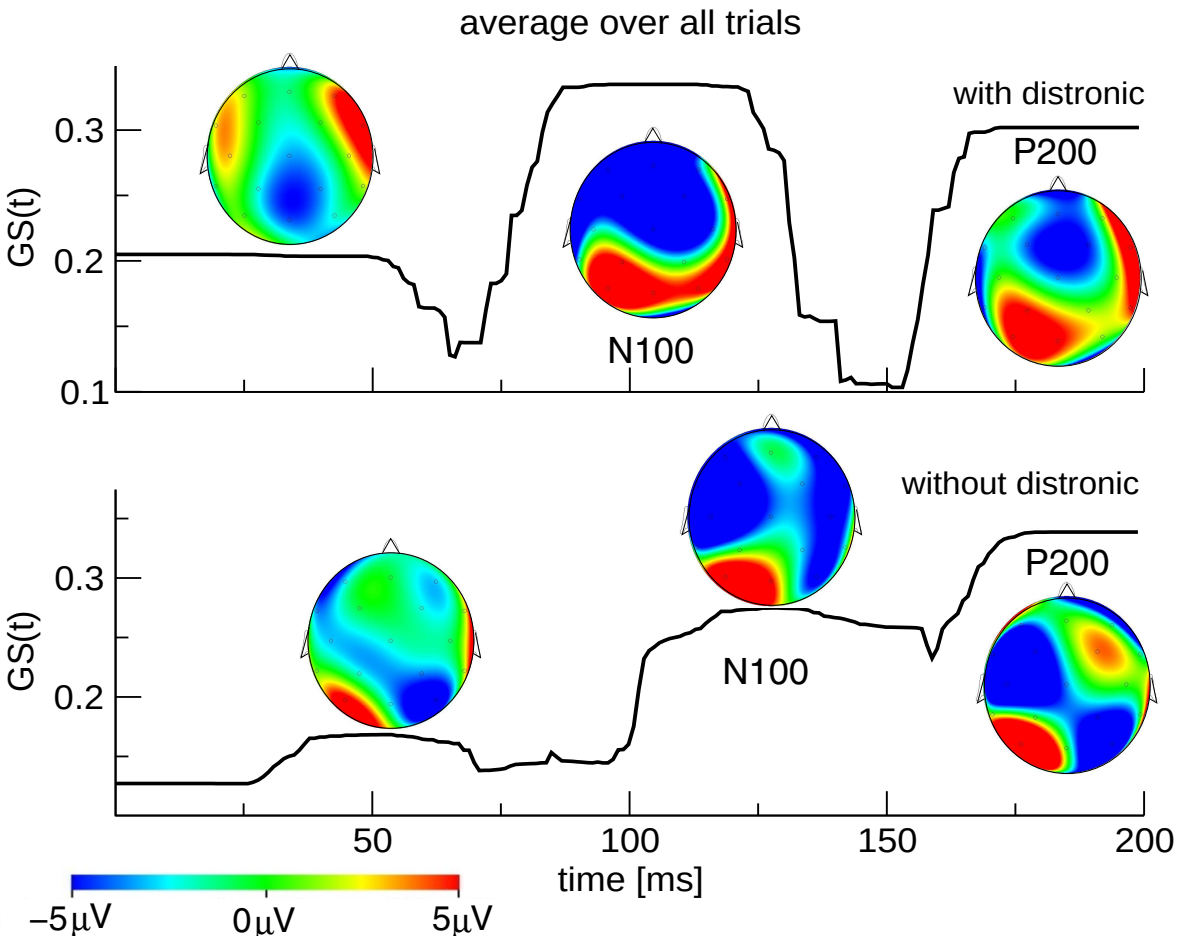

Figure 8: Cluster results from single data sets averaged over all trials for both experimental conditions and corresponding CSD-maps. The cluster quality p quantifies the generalized synchronization $GS(t) = p(t)$. Here, the focus to a shorter time window $[0\text{ms}; 200\text{ms}]$ increases the analysis resolution and reveals clear temporal segments. The CSD-maps represent averages over the corresponding time intervals.

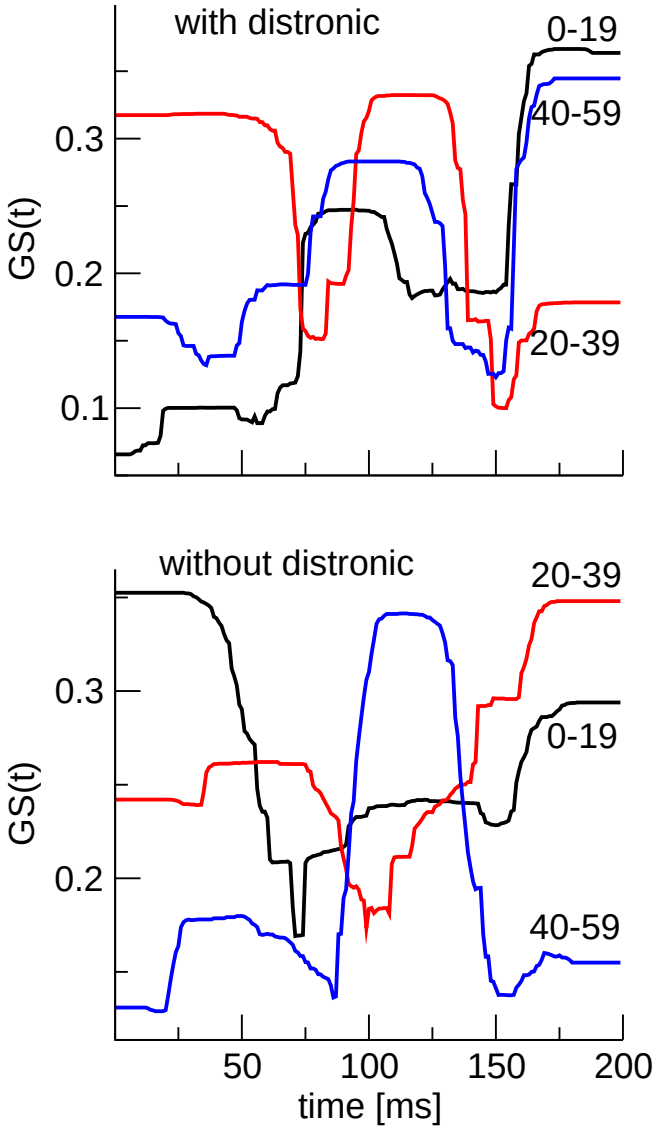

Figure 9: Cluster results from single data sets averaged over subsets of trials for both experimental conditions. The cluster quality p quantifies the generalized synchronization $GS(t) = p(t)$. Here, the focus to a shorter time window $[0\text{ms}; 200\text{ms}]$ increases the analysis resolution and reveals clear temporal segments. For illustration reasons, results $20 - 39$ and $40 - 59$ have been shifted artificially to lower values in both experimental conditions.

bands of functional relevance. Figure 10 reveals low spectral power beyond 15Hz, while the data exhibits increased global phase synchronization at 17Hz and 20Hz. In addition, both increased power spectral density and global phase synchronization occurs about $\nu = 6$Hz in the distronic condition and about $\nu = 5$Hz in the non-distronic condition, respectively. Hence, the analysis focus to the frequency bands $\nu = 6 \pm 0.6$Hz and $\nu = 5 \pm 0.5$Hz.

The subsequent analysis focus to the time window $[0\text{ms}; 400\text{ms}]$. The first examined datasets represent the averages over all trials and Fig. 11 shows short periods of increased MPS at $\sim$ 40ms, $\sim$ 80ms and $\sim$ 130ms in the distronic condition. Further, MPS is strong in $[240\text{ms}; 340\text{ms}]$ and even stronger after 340ms. In the non-distronic condition, the results reveal increased MPS from stimulus onset to $\sim$ 90ms, between 110ms and 185ms and between 190ms and 240ms. After a longer transition period, strong MPS emerges at 290ms and even stronger between 340ms and 400ms. Hence, the time segments of increased MPS are different in both conditions, while strong MPS coincide after $\sim$ 240ms. Here and in the following, the t-test gives p-values < 0.001 for both experimental conditions, i.e. all results are statistically significant.

Since these results reflect the average behaviour of all trials and might be smeared due to latency shifts in single trials, the focus to averages of trial subsets improves the temporal localization of components. Figure 12 shows latency shifts between all subset averages at rather early latencies at about 90ms in the distronic condition. Further, all trial averages reveal a synchronous plateau of MPS about 130ms, while averaged trials $0-19$ reveal retarded MPS at 260ms compared to the synchronous increase of MPS in subsets $20 - 39$ and $40 - 59$. This rather synchronous behaviour between different trials sets does not occur in the non-distronic condition, where only the promiment plateau of subset $40 - 59$ about 200ms coincides with the less prominent plateaus in $20 - 39$. Finally, in the distronic condition results from single trials reveal coincident components in [0ms;70ms], [70ms;140ms] and [250ms;300ms] in trials 30 and 50 (Fig. 13). In contrast, trial 10 exhibits components, which are shifted in latencies compared to trial 30 and 50. Less coincident behaviour is observed in the non-distronic condition.

3.3 Comparison to an existing method

The global circular variance $R(t)$ is only a rough quantity for mutual phase synchronization, as it smears out spatial inhomogeneities by averaging. In

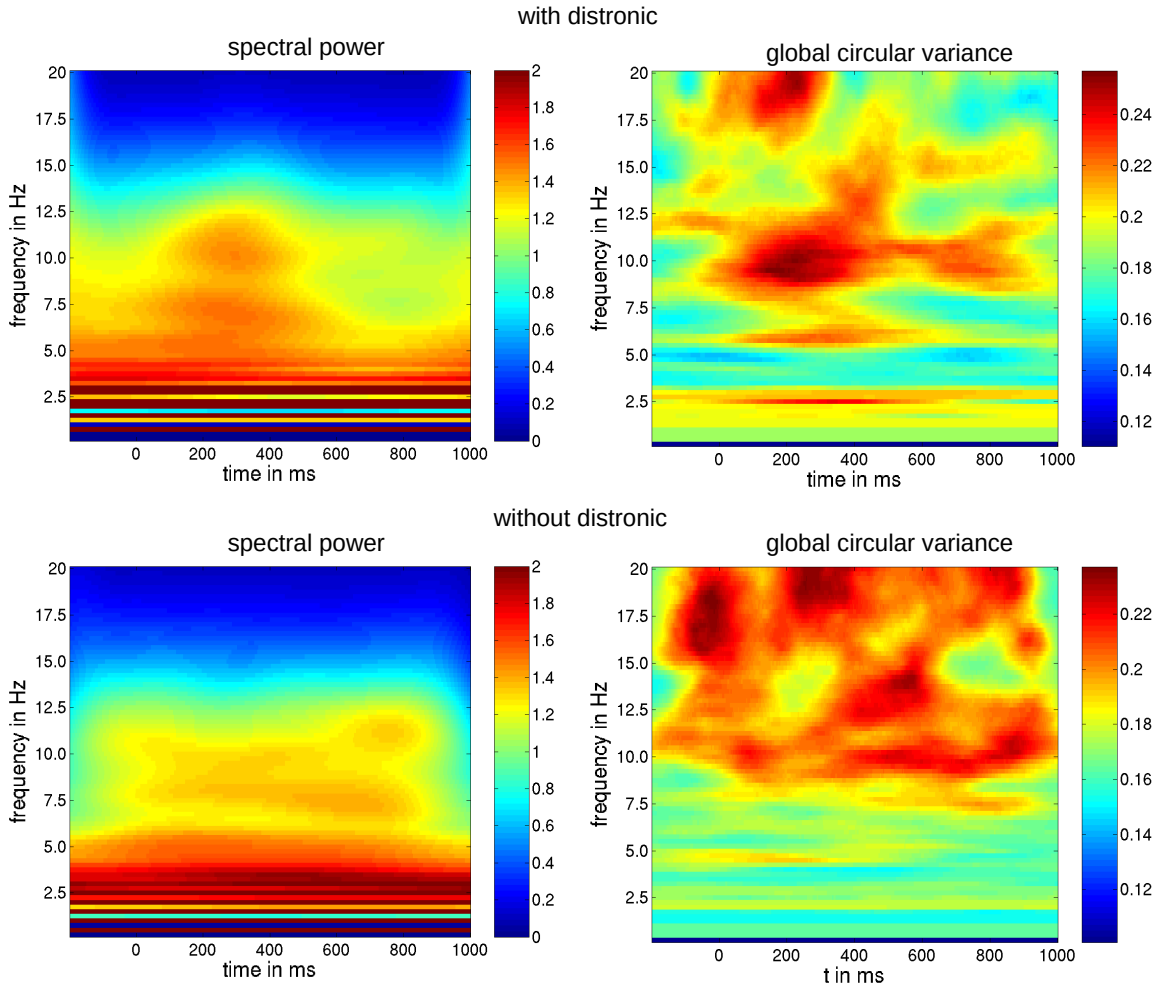

Figure 10: Spectral power from the average over all trials and global circular variance from all trials for both conditions. Spectral power contributions for frequencies larger than 20Hz are negligible.

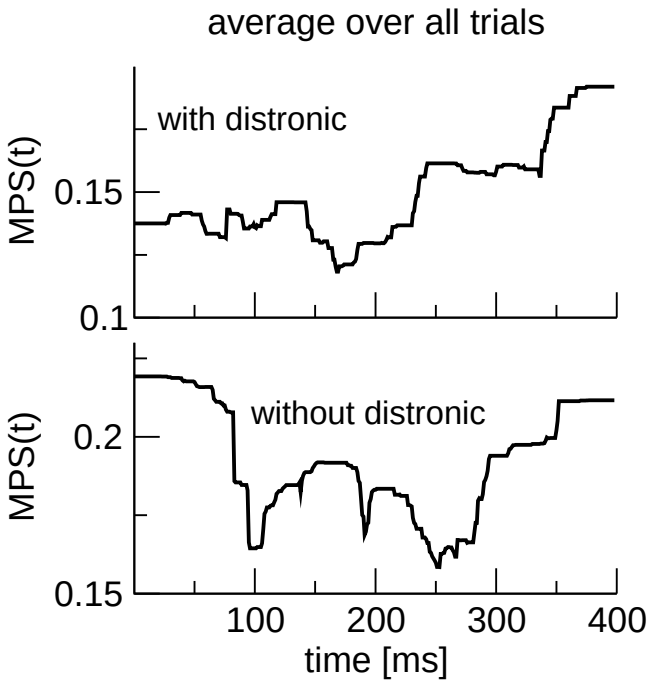

Figure 11: Cluster results from single phasic averages over all trials for both experimental conditions. The cluster quality p quantifies the mutual phase synchronization $MPS(t) = p(t)$. The original phasic signals are chosen in the freqency bands 6 ± 0.6Hz (with distronic) and 5 ± 0.5Hz (without distronic).

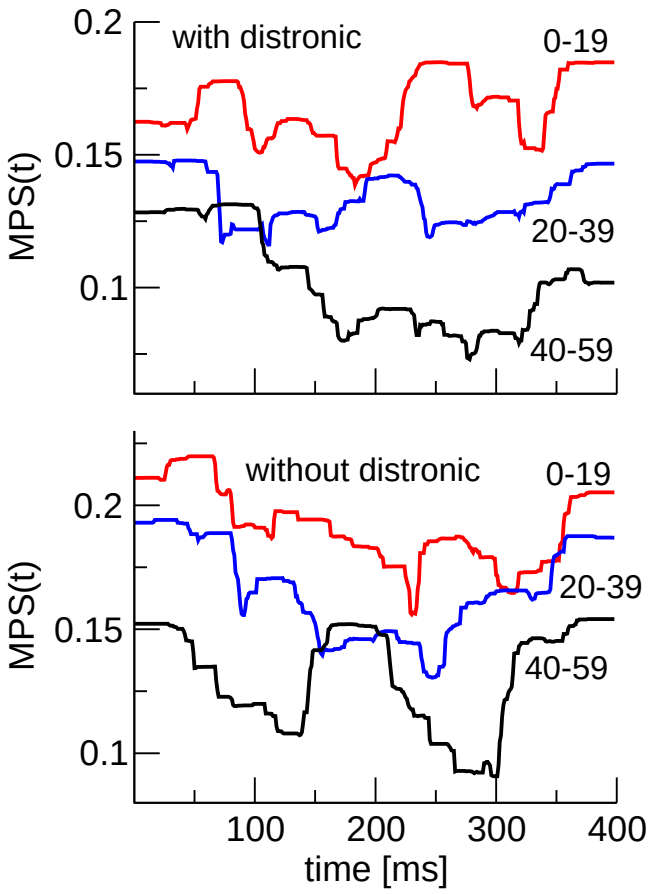

Figure 12: Cluster results from single phasic averages over subsets of trials for both experimental conditions. The cluster quality p quantifies the mutual phase synchronization $MPS(t) = p(t)$. Here, the phasic signals are chosen in the same freqency bands as in Fig. 11. For illustration reasons, results from $20-39$ and $40-59$ have been shifted artificially to lower values in both experimental conditions.

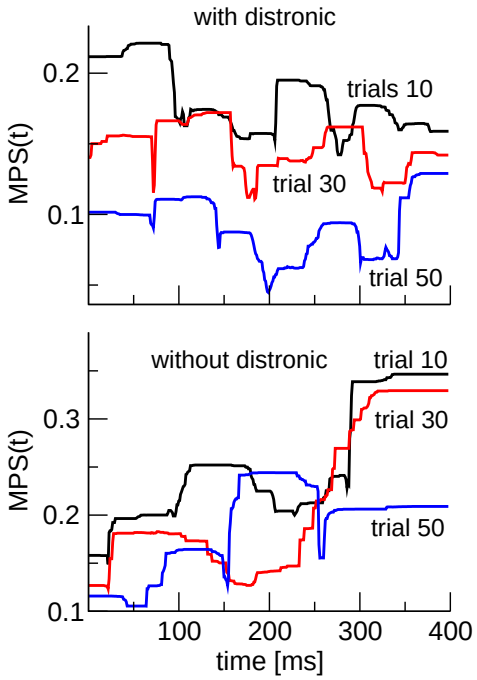

Figure 13: Cluster results from single trials for both experimental conditions. The specific choice of trials is arbitrary and the cluster quality p quantifies the mutual phase synchronization $MPS(t) = p(t)$. Here, the phasic signals are chosen in the same freqency bands as in Fig. 11. For illustration reasons, results from $20-39$ and $40-59$ have been shifted artificially to lower values in both experimental conditions.

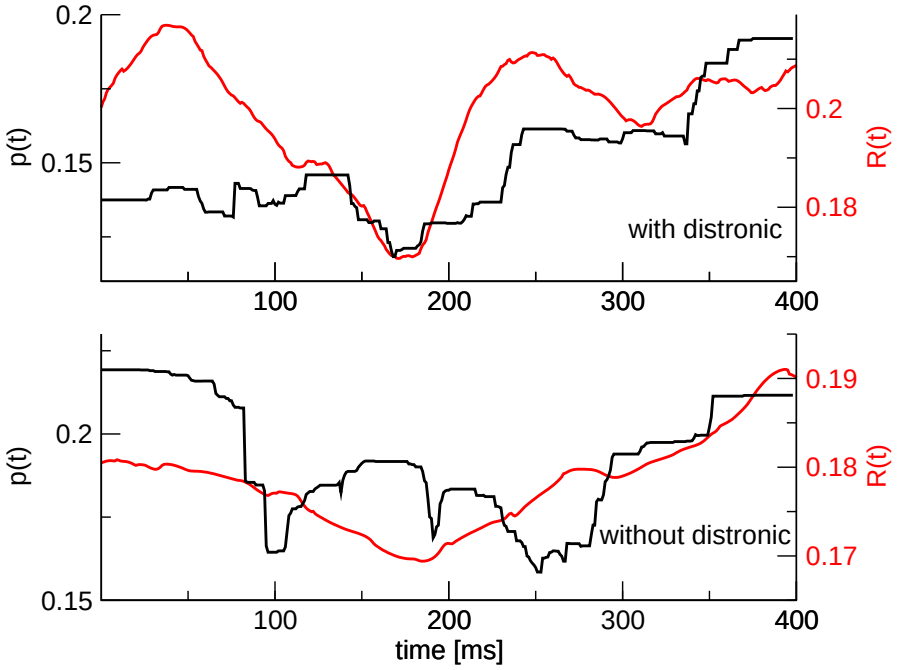

Figure 14: Comparison of cluster results to the conventional global circular variance for both experimental conditions.

contrast, the proposed cluster quality $P(t)$ takes into account the space-time structure of data. Figure 14 presents a direct comparison of both quantities for averages over all trials. In the distronic condition, the rough circular variance behaves in time similar to the cluster quality. More detailed, transients at $\sim$ 170ms coincide, while the transition from the component at 250ms to the component at 350ms occurs earlier in $R(t)$ than in $P(t)$. However, the most important difference between both quantities is the more detailed analysis of substructures by the proposed method. This is obvious in the non-distronic condition, where the substructure between 100ms and 250ms is lost in $R(t)$ and present in $P(t)$.

4 Discussion

The first part of the present work showed the relation of mutual space-time behaviour in brain signals to synchronization effects. Metastable behaviour

in linear data reflect transient generalized synchronization, while mutual metastability of circular data represent transient mutual phase synchronization. Considering these aspects, cluster analysis allows the segmentation of multivariate time series into metastable segments. The application to empirical linear and circular evoked potentials led to temporal segments, which show good accordance to cognitive components. Investigating subsets of trials revealed latency jitters between the sets. These latency shifts indicate that external stimuli do not reset the phase of brain activity to the same value at each stimulus onset. Hence, our findings attenuate the general assumption of fixed time delayed evoked response to the stimulus onset similar to previous studies (see e.g. Pfurtscheller & da Silva (1999)). That is event-related potentials do not represent a linear superposition of signal and uncorrelated noise and, subsequently, single trial averages have to be interpreted cautiously. In addition to the detection of latency jitters, we found a latency shift of component N100 between both experimental conditions in the averages over all trials. Against the conventional assumption, this novel result indicates an early attention effect. However, this shift is not such clear anymore in some trial subsets due to latency jitters in the component.

The major reason for the reason for the successful analysis of single data sets is the consideration of the space-time structure of the data. This becomes obvious by comparing our method to a conventional detection method for mutual phase synchronization treating only the temporal structure. It turns out that the conventional method looses important data structures, which are extracted by the proposed approach. In future work, we aim to develop a thorough single trial analysis with improved statistical assessment in order to gain further insights to the phase synchronization processes of underlying neural activity.

5 Acknowledgements

A.Hutt is supported by the DFG Research Center “Mathematics for key technologies” (FZT86) in Berlin, Germany.

References

Allefeld, C. & Kurths, J. [2003] "Multivariate phase synchronization analysis of eeg data" *IEICE Trans. Fundamentals* **E86-A**(9), 2218–2221.

Brandeis, D., Lehmann, D., Michel, C., & Mingrone, W. [1995] "Mapping event-related brain potential microstates to sentence endings" *Brain Topogrophy* **8**(2), 145–159.

Breakspear, M. [2002] "Nonlinear phase desynchronization in human electroencephalographic data" *Human Brain Mapping* **15**, 175–198.

DeShazer, D., Breban, R., Ott, E., & Roy, R. [2001] "Detecting phase synchronization in a chaotic laser array" *Phys. Rev. Lett.* **87**(4), 044101.

Duda, R. & Hart, P. [1973] Pattern Classification and Scene Analysis. (Wiley, New York).

Freeman, W. [2000] Neurodynamics: An Exploration in Mesoscopic Brain Dynamics (Perspectives in Neural Computing). (Springer-Verlag, Berlin).

Freeman, W. [2003] "Evidence from human scalp eeg of global chaotic itinerancy" *Chaos* **13**(3), 1069.

Gratton, G., Coles, M., & Donchin, E. [1989] "A procedure for using multielectrode information in the analysis of components of the event-related potential: Vector filter" *Psychophysiology* **26**, 222–232.

Haig, A., Gordon, E., Wright, J., Meares, R., & Bahramali, H. [2000] "Synchronous cortical gamma-band activity in task-relevant cognition" *Neuroreport* **11**, 669–675.

Hutt, A. [2004] "An analytical framework for modeling evoked and event-related potentials" *Int. J. Bif. Chaos* **14**(2), 653–666.

Hutt, A., Daffertshofer, A., & Steinmetz, U. [2003] "Detection of mutual phase synchronization in multivariate signals and application to phase ensembles and chaotic data" *Phys. Rev. E* **68**, 036219.

Hutt, A. & Kruggel, F. [2001] "Fixed point analysis: Dynamics of non-stationary spatiotemporal signals" in: S. Boccaletti, H. Mancini,

W. Gonzales-Vias, J. Burguete, & D. Valladares, eds., *Space-time Chaos: Characterization, Control and Synchronization* pp. 29–44 (World Scientific, Singapore).

Hutt, A. & Riedel, H. [2003] "Analysis and modelling of quasi-stationary multivariate time series and their application to middle latency auditory evoked potentials" *PhysicaD* **177**, 203–232.

Hutt, A., Svensen, M., Kruggel, F., & Friedrich, R. [2000] "Detection of fixed points in spatiotemporal signals by a clustering method" *Phys.Rev.E* **61**(5), R4691–R4693.

Ioannides, A., Kostopoulos, G., Laskaris, N., Liu, L., Shibata, T., Schellens, M., Poghosyan, V., & Khurshudyan, A. [2002] "Timing and connectivity in the human somatosensory cortex from single trial mass electrical activity" *Human Brain Mapping* **15**, 231–246.

Johnson, R. [1993] "On the neural generators of the p300 component of the event-related potential" *Psychophysiology* **30**, 90–97.

Karjalainen, P. & Kaipio, J. [1999] "Subspace regularization method for the single-trial estimation of evoked potentials" *IEEE Trans. Biomed. Eng.* **46**(7), 849–859.

Kay, L. [2003] "A challenge to chaotic itinerancy from brain dynamics" *Chaos* **13**(3), 1057–1066.

Kotz, S., Cappa, S., von Cramon, D., & Friederici, A. [2001] "Modulation of the lexical-semantic network by auditory semantic priming: An event-related functional mri study" *Neuroimage* **17**(4), 1761–1772.

Laskaris, N. & Ioannides, A. [2002] "Semantic geodesic maps: a unifying geometrical approach for studying the structure and dynamics of single trial evoked responses" *Clinical Neurophysiology* **113**, 1209–1226.

Lee, K., Williams, L., Breakspear, M., & E.Gordon [2003] "Synchronous gamma activity: a review and contribution to an integrative neuroscience model of schizophrenia" *Brain Research Reviews* **41**, 57–78.

Lehmann, D. & Skrandies, W. [1980] "Reference-free identification of components of checkerboard-evoked multichannel potential fields" *Electroenceph. clin. Neurophysiol* **48**, 609–621.

Luecke, J. & von der Malsburg, C. [2004] "Rapid processing and unsupervised learning in a model of the cortical macrocolumn" *Neural Computation* **16**(3), 501–533.

Mardia, K. & Jupp, P. [1999] Directional Statistics. (Wiley, New York).

Nunez, P. [1995] Neocortical dynamics and human EEG rhythms. (Oxfoord University Press, New York - Oxford).

Pascual-Marqui, R., Michel, C., & Lehmann, D. [1995] "Segmentation of brain electrical activity into microstates: Model estimation and validation" *IEEE Trans. Biomed. Eng.* **42**(7), 658–665.

Pfurtscheller, G. & da Silva, F. L. [1999] "Event-related eeg/meg synchronization and desynchronization: basic principles" *Clin Neurophysiol.* **110**(11), 1842–1857.

Pikovsky, A., Rosenblum, M., & Kurths, J. [2000] "Phase synchronization in regular and chaotic systems" *Int. J. Bif. Chaos* **10**(10), 2219.

Pikovsky, A., Rosenblum, M., & Kurths, J. [2001] Synchronization: A universal concept in nonlinear sciences. (Cambridge University Press).

Polich, J. [1998] "Clinical utility and control of variability" *J. Clin. Neurophysiol.* **15**(1), 14–33.

Pyragas, K. [1996] "Weak and strong synchronization of chaos" *Phys. Rev. E* **54**(5), R4508–4512.

Rosenblum, M., Pikovsky, A., Schafer, C., Tass, P., & Kurths, J. [2000] "Phase synchronization: from theory to data analysis" in: F. Moss & S. Gielen, eds., *Handbook of Biological Physics* vol. 4 of *Neuroinformatics* pp. 279–321 (Elsevier, New York).

Rugg, M. & Coles, M. [1996] Electrophysiology of Mind. (Oxford University Press, Oxford).

Rulkov, N., Sushchik, M., Tsimring, L., & Abarbanel, H. [1995] "Generalized synchronization of chaos in directionally coupled chaotic systems" *Phys. Rev. E* **51**(2), 980–994.

Schirmer, A., Kotz, S., & Friederici, A. [2002] "Sex differentiates the role of emotional prosody during word processing" *Cognitive Brain Research* **14**(2), 228–233.

Schrauf, M. & Kincses, W. [2003] "Imaging the driver's workload using eeg/erp" in: *Vision in Vehicles* pp. 13–14 (Elsevier Science, Amsterdam).

Singer, W. & Gray, C. [1995] "Visual feature integration and the temporal correlation hypothesis" *Annual Review Neuroscience* **18**, 555–586.

Tass, P. [1999] Phase resetting in medicine and biology : stochastic modelling and data analysis. (Springer, Berlin).

Tsuda, I. [2001] "Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems" *Behavioral and Brain Sciences* **24**(5), 793–847.

Wackermann, J. [1999] "Towards a quantitative characterisation of functional states of the brain: From the non-linear methodology to the global linear description" *Int. J. Psychophysiology* **34**, 65–80.

Wilson, H. & Cowan, J. [1972] "Excitatory and inhibitory interactions in localized populations of model neurons" *Biophys. J.* **12**, 1–24.