

A Comparative Study of Synchrony Measures for the Early Diagnosis of Alzheimer's Disease Based on EEG

J. Dauwels^{a,b,*}, F. Vialatte^c, T. Musha^d, A. Cichocki^c

^aLaboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA.

^bAmari Research Unit, RIKEN Brain Science Institute, Saitama, Japan.

^cLaboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Saitama, Japan.

^dBrain Functions Laboratory, Inc., Yokohama, Japan.

Abstract

It is well known that EEG signals of Alzheimer's disease (AD) patients are generally less synchronous than in age-matched control subjects. However, this effect is not always easily detectable. This is especially the case for patients in the pre-symptomatic phase, commonly referred to as Mild Cognitive Impairment (MCI), during which neuronal degeneration is occurring prior to the clinical symptoms appearance. In this paper, various synchrony measures are studied in the context of AD diagnosis, including the correlation coefficient, mean-square and phase coherence, Granger causality, phase synchrony indices, information-theoretic divergence measures, state space based measures, and the recently proposed stochastic event synchrony measures. Experiments with EEG data show that many of those measures are strongly correlated (or anti-correlated) with the correlation coefficient, and hence, provide little complementary information about EEG synchrony. Measures that are only weakly correlated with the correlation coefficient include the phase synchrony indices, Granger causality measures, and stochastic-event synchrony measures. In addition, those three families of synchrony measures are mutually uncorrelated, and therefore, they each seem to capture a specific kind of interdependence. For the data set at hand, only two synchrony measures are able to convincingly distinguish AD patients from age-matched control patients, i.e., Granger causality (in particular, full-frequency directed transfer function) and stochastic event synchrony. Those two measures are used as features to distinguish MCI patients from age-matched control subjects, yielding a leave-one-out classification rate of 83%. The classification performance may be further improved by adding complementary features from EEG; this approach may eventually lead to a reliable EEG-based diagnostic tool for MCI and AD.

Key words: Alzheimer's disease (AD), mild cognitive impairment (MCI),

Electroencephalography (EEG), synchrony, correlation coefficient, coherence function, corr-entropy coefficient, coh-entropy coefficient, wav-entropy coefficient, Granger causality, information-theoretic divergence measures, phase synchrony, state space based synchronization, S-estimator, omega complexity, non-linear interdependence, stochastic event synchrony, Hilbert transform, wavelet transform

1 Introduction

Many studies have shown that Alzheimer’s disease (AD) causes EEG signals to slow down: AD is associated with an increase of power in the low-frequency range (delta and theta) and a decrease of power in the high-frequency range (alpha and beta). In addition, the EEG signals of AD patients are generally less coherent than in age-matched control patients; see (Jeong, 2004; Uhlhaas & Singer, 2006) for an in-depth review. It is noteworthy, however, that those two effects are not always easily detectable: there tends to be large variability among AD patients. As a result, none of those phenomena allow at present to reliably diagnose AD at an early stage.

A considerable amount of research has recently been devoted to detecting fluctuations in EEG synchrony; more specifically, a large variety of measures has been proposed to quantify EEG synchrony, stemming from a wide spectrum of disciplines such as physics, information theory, statistics, and signal processing; we refer to (Pereda *et al.*, 2005; Uhlhaas & Singer, 2006; Pereda *et al.*, 2005; Breakspear, 2004; Kamiński *et al.*, 2005; Stam, 2005) for recent reviews on EEG synchrony measures.

In this paper, we try to systematically investigate EEG synchrony with special focus on the early diagnosis of AD. (A related but less exhaustive study has been presented in (Quiroga *et al.*, 2002; Sakkalis *et al.*, 2006) in the context of epilepsy.) We will consider EEG data from mild cognitive impairment (MCI) patients that later suffer from AD. We apply a large variety of synchrony methods to *the same* EEG data set recorded from MCI and control subjects. To our knowledge, those synchrony methods have so far only been applied

* Corresponding author.

Email addresses: `justin@dauwels.com` (J. Dauwels),
`fvialatte@brain.riken.jp` (F. Vialatte), `musha@bfl.co.jp` (T. Musha),
`cia@brain.riken.jp` (A. Cichocki).

¹ J.D. was in part supported by post-doctoral fellowships from the Japanese Society for the Promotion of Science (JSPS), the King Baudouin Foundation, and the Belgian American Educational Foundation (BAEF). Part of this work was carried out while J.D. was at the RIKEN Brain Science Institute, Saitama, Japan.

separately to *different* data sets from MCI/AD patients. As a result, it is hard to compare the effectiveness of those methods in the context of diagnosing MCI/AD from EEG.

Our objective is to apply synchrony measures to EEG data of MCI patients and control subjects, and to use those measures as *features* to separate MCI patients from control subjects. In other words, our study is mostly clinically motivated; in this paper, we will not try to identify the biophysical mechanisms that cause the observed effects. Such interpretation, albeit of great importance, is notoriously difficult, since fluctuations in EEG synchrony may be caused by a variety of phenomena, and may be affected by the choice of reference electrodes.

To understand some of those issues, it helps to divide the synchrony measures into two separate classes: *local* and *global* methods. Local measures establish links between pairs of signals. Global methods can in principle handle an arbitrary number of signals, and can therefore analyze signals from all EEG channels simultaneously. Therefore, global measures are applicable to processes that produce EEG signals distributed over the entire scalp. The problem with global measures is obviously that they usually cannot provide local information about synchronous activity. On the other hand, local measures can strictly speaking only be applied meaningfully when local activity has been identified; that essentially involves solving the inverse problem, which is known to be very hard. Solving those issues goes beyond the scope of this paper. Because of the above mentioned limitations of synchrony measures, it is hard to infer the precise biophysical mechanisms that cause loss in EEG synchrony. We will simply *conjecture* some plausible explanations, and will leave the in-depth investigation of those theories as subject of future research.

We wish to underline that although we consider a large variety of synchrony measure, it is virtually impossible to include all existing synchrony measures in this study. Novel synchrony measures are constantly being developed; the total number of proposed synchrony measures is probably close to one hundred or even higher.

This paper is structured as follows. In Section 2 we review the synchrony measures considered in this paper. In Section 3 those measures are applied to EEG data, in particular, for the purpose of diagnosing MCI: we describe our data set, elaborate on various implementation issues, and present our results. At the end of the paper (Section 4), we briefly relate our results to earlier work on this topic, and speculate about the neurophysiological interpretation of our results.

2 Synchrony Measures

We briefly review various families of synchrony measures: correlation coefficient and analogues in frequency and time-frequency domain, Granger causality, phase synchrony, state space based synchrony, information theoretic interdependence measures, and stochastic event synchrony.

2.1 Correlation Coefficient

The correlation coefficient r is perhaps one of the most well-known measures for (linear) interdependence between two signals x and y (Nunez & Srinivasan, 2006):

$$r = \frac{1}{N} \sum_{k=1}^N \frac{(x(k) - \bar{x})(y(k) - \bar{y})}{\sigma_x \sigma_y}, \quad (1)$$

where N is the length the signals, \bar{x} and \bar{y} is the (sample) mean of x and y respectively, and σ_x^2 and σ_y^2 is the (sample) variance of x and y respectively. The correlation coefficient r quantifies the linear correlation between x and y . If x and y are not linearly correlated, r is close to zero; on the other hand, if both signals are identical, then $r = 1$.

2.2 Coherence

The coherence function quantifies linear correlations in frequency domain (Nunez & Srinivasan, 2006). One distinguishes the magnitude square coherence function and the phase coherence function. In order to compute those quantities, the signals x and y are subdivided in M segments of equal length L . The coherence function is computed by averaging over those segments. In particular, the magnitude square coherence function $c(f)$ is computed as:

$$c(f) = \frac{|\langle X(f)Y^*(f) \rangle|^2}{|\langle X(f) \rangle| |\langle Y(f) \rangle|}, \quad (2)$$

where $X(f)$ and $Y(f)$ are the Fourier transforms of x and y respectively; Y^* is the complex conjugate of $Y \in \mathbb{C}$, $|Y|$ is the magnitude of Y , and $\langle X(f) \rangle$ stands for the average of $X(f)$ computed over the M segments, likewise $\langle Y(f) \rangle$ and $\langle X(f)Y^*(f) \rangle$.

The phase coherence function is defined as

$$\phi(f) = \arg \langle X(f)Y^*(f) \rangle, \quad (3)$$

where $\arg X$ is the argument of $X \in \mathbb{C}$. Note that both $c(f)$ and $\phi(f)$ depend on the frequency f . In our study, we will test various values for M (see Section 3.2.4).

2.3 Corr-Entropy Coefficient

The corr-entropy coefficient r_E is a recently proposed (Gunduz and Principe, 2009) non-linear extension of the correlation coefficient r :

$$r_E = \frac{\frac{1}{N} \sum_{k=1}^N \kappa(x(k), y(k)) - \frac{1}{N^2} \sum_{k,\ell=1}^N \kappa(x(k), y(\ell))}{\sqrt{K_X - \frac{1}{N^2} \sum_{k,\ell=1}^N \kappa(x(k), x(\ell))} \sqrt{K_Y - \frac{1}{N^2} \sum_{k,\ell=1}^N \kappa(y(k), y(\ell))}}, \quad (4)$$

where

$$K_X = \frac{1}{N} \sum_{k=1}^N \kappa(x(k), x(k)) \quad K_Y = \frac{1}{N} \sum_{k=1}^N \kappa(y(k), y(k)), \quad (5)$$

and where κ is a symmetric positive definite kernel function (Shawe-Taylor & Cristianini, 2004), for example, a Gaussian, sigmoidal, or polynomial kernel. In this paper (as in (Gunduz and Principe, 2009)), we will only consider the Gaussian kernel:

$$\kappa(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{|x-y|^2}{2\sigma^2}}, \quad (6)$$

with kernel width $\sigma = 0.4$ (as in (Gunduz and Principe, 2009)).

We wish to point out that the signals x and y need to be normalized before evaluating (4), by subtracting the mean and then dividing by the standard deviation. Such normalization is crucial, since x and y may have significantly different dynamic ranges; moreover, as a result of this normalization, both x and y are dimensionless.

Note that r_E is close to zero if x and y are independent (which is stronger than being uncorrelated); if both signals are equal, then $r_E = 1$.

2.4 Coh-Entropy and Wav-Entropy Coefficient

Similarly to (4), one can define a non-linear magnitude square coherence function, which we will refer to as ‘‘coh-entropy’’ coefficient $c_E(f)$:

$$c_E(f) = \frac{\langle \kappa(X(f), Y(f)) \rangle}{\sqrt{\langle \kappa(X(f), X(f)) \rangle} \sqrt{\langle \kappa(Y(f), Y(f)) \rangle}}, \quad (7)$$

where the averages $\langle \cdot \rangle$ are again computed over M segments of equal length L . Before computing $c_E(f)$, the Fourier-transforms $X(f)$ and $Y(f)$ need to be normalized by the mean and standard deviation, computed over all segments.

The coh-entropy is an extension of corr-entropy to the frequency domain. In addition, it can be viewed as a generalization of the magnitude square coherence function $c(f)$ (2).

One can also readily extend corr-entropy to the time-frequency domain by replacing $x(k)$ and $y(k)$ in (4) by the time-frequency (“wavelet”) transforms $X(k, f)$ and $Y(k, f)$ respectively, resulting in a “wav-entropy” coefficient $w_E(f)$. The (continuous) wavelet transform of x is obtained as:

$$X(k, s) = \sum_{\ell} x(\ell) \psi^* \left(\frac{\ell - k}{s} \right), \quad (8)$$

where $\psi(k)$ is the (complex) “mother” wavelet, and s is a scaling factor. In this paper, we use the complex Morlet wavelet

$$\psi(k) = A e^{-k^2/2\sigma_0^2} e^{2i\pi f_0 k}, \quad (9)$$

where the width σ_0^2 and frequency f_0 jointly determine the number of oscillations in the wavelet. The complex Morlet wavelet results in the optimal resolution in time and frequency; it has also proven to be well-suited for EEG signals (Herrmann *et al.*, 2005). The wavelet transform (8) is a time-frequency representation of x ; the scaled and shifted “daughter” wavelet in (8) has center frequency $f = f_0/s$. In the following, we will use the notation $X(k, f)$ instead of $X(k, s)$.

It is noteworthy that before computing the wav-entropy coefficient $w_E(f)$, the time-frequency transforms $X(k, f)$ and $Y(k, f)$ need to be normalized, for any frequency f , by subtracting the average (over time) and then dividing by the standard deviation (over time).

The two straightforward extensions $c_E(f)$ and $w_E(f)$ were not considered in (Gunduz and Principe, 2009); to our knowledge, they are novel. It is noteworthy that both coefficients are equal to one if the signals x and y are identical; if the latter signals are independent, both coefficients are close to zero.

2.5 Granger Causality

Granger causality² refers to a family of synchrony measures that are derived from *linear* stochastic models of time series; as the above linear interdependence measures, they quantify to which extent different signals are linearly interdependent (see (Granger, 1969; Kamiński *et al.*, 1991, 2005; Gourévitch *et al.*, 2006; Korzeniewska *et al.*, 2003; Eichler, 2006; Blinowska *et al.*, 2004; Ancona *et al.*, 2004; Astolfi *et al.*, 2004, 2005; Schelter *et al.*, 2005; Chen *et al.*, 2006) for detailed information about Granger causality). Whereas the linear interdependence measures of Section 2.1 to 2.4 are bivariate, i.e., they can only be applied to pairs of signals, Granger causality measures are multivariate, they can be applied to multiple signals *simultaneously*. Interestingly, non-linear extensions of Granger causality have been proposed recently (see, e.g., (Ancona *et al.*, 2004; Chen *et al.*, 2004)), but we will not consider such extensions in this paper, since they are less commonly used.

Suppose that we are given n signals $x_1(k), x_2(k), \dots, x_n(k)$, each stemming from a different channel; we will assume, without loss of generality, that those signals are normalized to have zero mean and standard deviation equal to one. We consider the multivariate autoregressive (MVAR) model:

$$x(k) = \sum_{\ell=1}^p \mathbf{A}(j)x(k-\ell) + e(k), \quad (10)$$

where $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$, p is the model order, the model coefficients $\mathbf{A}(j)$ are $n \times n$ matrices, and $e(k)$ is a zero-mean Gaussian random vector of size n . In words: Each signal $x_i(k)$ is assumed to linearly depend on its own p past values and the p past values of the other signals $x_j(k)$. The deviation between $x(k)$ and this linear dependence is modeled by the noise component $e(k)$. Model (10) can also be cast in the form:

$$e(k) = \sum_{\ell=0}^p \tilde{\mathbf{A}}(j)x(k-\ell), \quad (11)$$

where $\tilde{\mathbf{A}}(0) = \mathbf{I}$ (identity matrix) and $\tilde{\mathbf{A}}(j) = -\mathbf{A}(j)$ for $j > 0$. One can transform (11) into the frequency domain (by applying the z -transform and by substituting $z = e^{-2\pi i \Delta t}$, where $1/\Delta t$ is the sampling rate):

$$X(f) = \tilde{\mathbf{A}}^{-1}(f)E(f) = \mathbf{H}(f)E(f). \quad (12)$$

The power spectrum matrix of the signal $x(k)$ is determined as

$$\mathbf{S}(f) = \mathbb{E}[X(f)X(f)^\dagger] = \mathbf{H}(f)\mathbf{V}\mathbf{H}^\dagger(f), \quad (13)$$

² The Granger causality measures we consider here are implemented in the BioSig library, available from <http://biosig.sourceforge.net/>.

where \mathbf{V} stands for the covariance matrix of $e(k)$, and $X(f)^\dagger$ is the hermitian conjugate of $X(f)$. The Granger causality measures are defined in terms of the matrices \mathbf{A} , \mathbf{H} , and \mathbf{S} .

We now list the most common Granger causality measures.

2.5.1 Partial Coherence (PC)

One defines partial coherence (Kamiński *et al.*, 2005):

$$C_{ij}(f) = \frac{M_{ij}(f)}{\sqrt{M_{ii}(f)}\sqrt{M_{jj}(f)}} \in \mathbb{C}, \quad (14)$$

where M_{ij} is a minor (i, j) of \mathbf{S} , i.e., the determinant of \mathbf{S} with row i and column j removed; $|C_{ij}(f)| \in [0, 1]$ describes the amount of in-phase components in signals i and j at the frequency f when the influence (i.e., linear dependence) of the other signals is statistically removed. Note that this Granger causality measure is symmetric, i.e., $C_{ij} = C_{ji}$.

The following Granger causality measures capture causal relations, they are asymmetric or “directed”.

2.5.2 Directed Transfer Function (DTF)

The directed transfer function is defined in terms of \mathbf{H} (Kamiński *et al.*, 1991):

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{j=1}^m |H_{ij}(f)|^2} \in [0, 1], \quad (15)$$

where the (frequency-dependent) normalization is chosen so that $\gamma_{ij}^2(f)$ quantifies the fraction of inflow to channel i stemming from channel j .

2.5.3 Full Frequency Directed Transfer Function (ffDTF)

Full frequency directed transfer function is defined as (Kamiński *et al.*, 2005):

$$F_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_f \sum_{j=1}^m |H_{ij}(f)|^2} \in [0, 1], \quad (16)$$

which is a variation of $\gamma_{ij}^2(f)$ with a global normalization in frequency.

2.5.4 Partial Directed Coherence (PDC)

The previous two measures are causal extensions of coherence (2). One can also define causal extensions of partial coherence (14), i.e., partial directed coherence (Kamiński *et al.*, 2005):

$$P_{ij}(f) = \frac{\tilde{A}_{ij}(f)}{\sqrt{\sum_{i=1}^m |\tilde{A}_{ij}(f)|^2}} \in \mathbb{C}, \quad (17)$$

where $|P_{ij}(f)| \in [0, 1]$ represents the fraction of outflow from channel j to channel i due to the choice of normalization.

2.5.5 Direct Directed Transfer Function (dDTF)

At last, we consider the direct directed transfer function (Korzeniewska *et al.*, 2003):

$$\chi_{ij}^2(f) = F_{ij}^2(f)C_{ij}^2(f) \in [0, 1], \quad (18)$$

which is non-zero if the connection between channel i and j is causal (non-zero $F_{ij}^2(f)$) and direct (non-zero $C_{ij}^2(f)$).

It is noteworthy that DTF, dDTF, PDC, and fDTF are asymmetric measures, i.e., $\gamma_{ij}^2 \neq \gamma_{ji}^2$, $F_{ij}^2 \neq F_{ji}^2$, $P_{ij}^2 \neq P_{ji}^2$, and $\chi_{ij}^2 \neq \chi_{ji}^2$.

2.6 MVAR Coherence

By means of the matrix \mathbf{S} (13), one may also compute the coherence function (Kamiński *et al.*, 2005):

$$K_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)}\sqrt{S_{jj}(f)}} \in \mathbb{C}, \quad (19)$$

where $|K_{ij}(f)| \in [0, 1]$ describes the amount of in-phase components in signals i and j at the frequency f . The squared magnitude $|K_{ij}(f)|^2 \in [0, 1]$ is an alternative to the square magnitude coherence function $c(f)$ (2). The argument of $K_{ij}(f)$, i.e., $\arg K_{ij}(f)$, is an alternative to the phase coherence function $\phi(f)$ (3). The coherence functions $|K_{ij}(f)|^2$ and $\arg K_{ij}(f)$ may be more reliable than $c(f)$ (2) and $\phi(f)$ (3) respectively if the MVAR system (10) is a good model for the signals at hand.

Note that $K_{ij}(f)$ is usually *not* considered as a Granger measure, although it is also derived from an MVAR model (10); we will refer to $K_{ij}(f)$ it as ‘‘MVAR coherence’’.

2.7 Phase Synchrony

Phase synchrony refers to the interdependence between the instantaneous phases ϕ_x and ϕ_y of two signals x and y ; the instantaneous phases may be strongly synchronized even when the amplitudes of x and y are statistically independent. The instantaneous phase ϕ_x of a signal x may be extracted as (Lachaux *et al.*, 1999):

$$\phi_x^H(k) = \arg [x(k) + i\tilde{x}(k)], \quad (20)$$

where \tilde{x} is the Hilbert transform of x . Alternatively, one can derive the instantaneous phase from the time-frequency transform $X(k, f)$ (8) of x :

$$\phi_x^W(k, f) = \arg[X(k, f)]. \quad (21)$$

The phase $\phi_x^W(k, f)$ depends on the center frequency f of the applied wavelet. By appropriately scaling the wavelet, the instantaneous phase may be computed in the frequency range of interest.

The phase synchrony index γ for two instantaneous phases ϕ_x and ϕ_y is defined as (Lachaux *et al.*, 1999):

$$\gamma = \left| \left\langle e^{i(n\phi_x - m\phi_y)} \right\rangle \right| \in [0, 1], \quad (22)$$

where n and m are integers (usually $n = 1 = m$), and $\langle \cdot \rangle$ stands for the time average. We will use the notation γ_H and γ_W to indicate whether the instantaneous phases are computed by the Hilbert transform or time-frequency transform respectively. Note that γ_W depends on frequency f .

An alternative phase synchrony index is Global Field Synchronization (GFS) (Koenig *et al.*, 2001); whereas the index γ (22) is determined for *pairs* of signals x and y , GFS quantifies the synchrony of *multiple* signals $x_1(k), x_2(k), \dots, x_n(k)$. GFS is based on principal component analysis (PCA) applied in the complex plane; as a first step, one computes the Fourier transform $X_i(f) \in \mathbb{C}$ of the signals $x_i(k)$ ($i = 1, 2, \dots, n$). Next one constructs the vectors $X_R = (\text{Re}(X_1(f)), \dots, \text{Re}(X_n(f)))$ and $X_I = (\text{Im}(X_1(f)), \dots, \text{Im}(X_n(f)))$, and computes the covariance matrix $\mathbf{C} \in \mathbb{R}^{2 \times 2}$ for those two vectors. GFS is defined in terms of the normalized eigenvalues λ_1 and λ_2 of \mathbf{C} (with $\lambda_1 \geq \lambda_2$):

$$\text{GFS}(f) = \lambda_1 - \lambda_2. \quad (23)$$

It is noteworthy that GFS depends on the frequency f . If the phases of the signals $x_1(k), x_2(k), \dots, x_n(k)$ are strongly interdependent, the Fourier transforms $X_1(f), X_2(f), \dots, X_n(f)$ will lie on a straight line in the complex plane. The largest normalized eigenvalue of \mathbf{C} will be close to one, and as a result,

also GFS will tend to one. On the other hand, if the phases are not correlated, we have $\lambda_1 \approx 1/2 \approx \lambda_2$ and GFS will be close to zero.

In this paper, we will consider two additional phase synchrony indices, i.e., the evolution map approach (EMA) and the instantaneous period approach (IPA) (Rosenblum *et al.*, 2002). Due to space constraints, we will not describe those measures here, instead we refer the reader to (Rosenblum *et al.*, 2002)³; additional information about phase synchrony can be found in (Quiroga *et al.*, 2002).

2.8 State Space Based Synchrony

State space based synchrony (or “generalized synchronization”) evaluates synchrony by analyzing the interdependence between the signals in a state space reconstructed domain (see (Sauer *et al.*, 2005) for a review). The central hypothesis behind this approach is that the signals at hand are generated by an (unknown) deterministic, potentially high-dimensional, non-linear dynamical system:

$$ds/dt = f(s), \quad (24)$$

where $s \in \mathbb{R}^q$ is the state of the system, and q and f are usually unknown. In order to reconstruct such system from a signal x , one considers delay vectors $X(k) = (x(k), x(k - \tau), \dots, x(k - (m - 1)\tau))^T$, where m is the embedding dimension and τ denotes the time lag. If τ and m are appropriately chosen, and the signals are generated by a deterministic dynamical system with a smooth attractor, the delay vectors lie on a smooth manifold (“mapping”) in \mathbb{R}^m . The principle of state space reconstruction can be generalized from *smooth* attractors to generic sets of arbitrary box-counting dimension, including *fractal* attractors (Sauer *et al.*, 2005).

State space based synchrony measures assess the interdependence between two signals x and y by comparing local neighborhoods of their state space reconstructed mappings. More precisely, they quantify how local neighborhoods in one manifold are transformed into neighborhoods in the other manifold. If a neighborhood on one manifold is mapped unto a much larger neighborhood on the other manifold, the signals x and y are only weakly synchronized; on the other hand, if it is mapped to a neighborhood of about the same size, the two signals x and y are strongly coherent.

First we consider three non-linear interdependence measures (S^k , H^k , and N^k) that are based on this idea, then we consider a recently proposed state space based synchrony measure, i.e., the S-estimator. The latter may be considered as a state-space formulation of omega complexity, which is a classical

³ Program code is available at www.agnld.uni-potsdam.de/%7Emros/dircnew.m

similarity measure derived from principal component analysis (PCA). After describing the three non-linear interdependence measures S^k , H^k , and N^k , we will consider omega complexity and then proceed to the S-estimator.

We will again assume, without loss of generality, that all signals are normalized to have zero mean and standard deviation equal to one.

2.8.1 Nonlinear Interdependence

For each delay vector $X(k)$, one computes the average squared Euclidian distance $R_k^M(X)$ to its M nearest neighbors. In addition, one computes the mean squared Euclidian distance between delay vector $X(k)$ and the X -delay vectors corresponding to the M nearest neighbors of $Y(k)$, referred to as the Y -conditioned mean squared Euclidian distance $R_k^M(X|Y)$. Three measures of nonlinear independence, S^k , H^k , and N^k , are defined in terms of those two distances (Quiroga *et al.*, 2002)⁴:

$$S^k(X|Y) = \frac{1}{N} \sum_{k=1}^N \frac{R_k^M(X)}{R_k^M(X|Y)} \quad (25)$$

$$H^k(X|Y) = \frac{1}{N} \sum_{k=1}^N \log \frac{R_k(X)}{R_k^M(X|Y)} \quad (26)$$

$$N^k(X|Y) = \frac{1}{N} \sum_{k=1}^N \frac{R_k(X) - R_k^M(X|Y)}{R_k(X)}, \quad (27)$$

where $R_k(X)$ is the mean squared Euclidian distance between $X(k)$ and the other points $X(\ell)$ ($\ell \neq k$). The measure H^k is believed to be more robust against noise than S^k ; it is also easier to interpret, however, it is not normalized; N^k is a normalized version of H^k .

In our calculations (see Section 3), the number M of nearest neighbors was set equal to 10, as in (Quiroga *et al.*, 2002).

2.8.2 Omega Complexity and S-estimator

Omega complexity Ω is a measure derived from principal component analysis (PCA) (Saito *et al.*, 1998) (see (Yoshimura *et al.*, 2004) for an application to EEG). Suppose that we are given n signals $X_1(k), X_2(k), \dots, X_n(k)$; one computes the covariance matrix $\mathbf{C} \in \mathbb{R}^{n \times n}$. Omega complexity is defined in terms of the normalized eigenvalues λ_i of \mathbf{C} :

$$\Omega = \exp \left(- \sum_{i=1}^n \lambda_i \log \lambda_i \right). \quad (28)$$

⁴ Software is available from <http://www.vis.caltech.edu/~rodri/software.htm>

The argument of the exponential in (28) is the entropy of the normalized-eigenvalue distribution. If the signals $X_i(k)$ are identical, all eigenvalues except one are zero, and $\Omega = 1$. On the other hand, if the signals $X_i(k)$ are independent, all n eigenvalues are identical, and hence $\Omega = n$. Therefore, omega complexity Ω is a dissimilarity measure.

The S-estimator extends omega complexity to state space embedded signals (Carmeli *et al.*, 2005)⁵ (see (Celka *et al.*, 2006) for a recent refinement). Suppose again that we are given n signals $X_1(k), X_2(k), \dots, X_n(k)$; first one constructs m -dimensional delay vectors for each signal, then one computes the covariance matrix $\mathbf{C} \in \mathbb{R}^{mn \times mn}$ for the n sequences of delay vectors. The S-estimator is defined in terms of the normalized eigenvalues λ_i of \mathbf{C} :

$$S_{\text{est}} = 1 + \frac{\sum_{i=1}^{mn} \lambda_i \log(\lambda_i)}{\log(mn)}. \quad (29)$$

The numerator in the second term represents the negative entropy of the normalized eigenvalues λ_i , and the denominator is the negative entropy of the uniform distribution, which has maximum entropy; as a consequence, the second term in the RHS of (29) takes values in $[-1, 0]$, and hence, $S_{\text{est}} \in [0, 1]$. If the signals $X_i(k)$ are statistically independent, all normalized eigenvalues tend to be equal to $1/mn$, and as a result, S_{est} approaches 0. On the contrary, if the signals are well synchronized, only a few number of eigenvalues will remain prominent, and as a result, S_{est} is then close to 1.

2.9 Information-Theoretic Measures

Several interdependence measures have been proposed that have their roots in information theory (Cover & Thomas, 1991).

2.9.1 Mutual Information

Mutual information is perhaps the most well-known information-theoretic interdependence measure between two random variables X and Y :

$$I(X; Y) = H(X) + H(Y) - H(X, Y), \quad (30)$$

where $H(X)$ and $H(Y)$ is the Shannon entropy of X and Y respectively, and $H(X, Y)$ is the joint entropy of X and Y (Cover & Thomas, 1991). The mutual information quantifies the amount of information the random variable Y contains about random variable X (and vice versa); it is always positive, and it

⁵ A toolbox is available from <http://aperest.epfl.ch/docs/software.htm>

vanishes when X and Y are statistically independent. If X and Y are identical (with probability 1), then $I(X;Y) = H(X) = H(Y)$.

It is noteworthy that in the problem at hand (i.e., diagnosis of MCI from EEG), we are given *signals* and not random variables. Computing the mutual information between (stationary) signals is non-trivial: naively quantizing the signals and computing $I(X;Y)$ from the resulting histograms often leads to unreliable estimates of mutual information. Recently, a sophisticated and effective technique to compute mutual information was proposed (Kraskov *et al.*, 2004); we will use that method in this paper, with the same parameter settings⁶.

The method of (Kraskov *et al.*, 2004) computes mutual information in time-domain. Alternatively, this quantity may also be determined in time-frequency domain (Aviyente, 2005a,b) (see also (Quiroga *et al.*, 1999; Blanco *et al.*, 1995; Quiroga *et al.*, 2001)), more specifically, from the normalized spectrograms:

$$C_x(k, f) = \frac{|X(k, f)|^2}{\sum_{k,f} |X(k, f)|^2}, \quad (31)$$

where the summation in the denominator is carried out over the time window and frequency range of interest. The normalized spectrograms can be treated as probability distributions, and accordingly, mutual information can be determined in terms of the normalized spectrograms as:

$$I_W(C_x, C_y, C_{xy}) = \sum_{k,f} C_{xy}(k, f) \log \frac{C_{xy}(k, f)}{C_x(k, f)C_y(k, f)}, \quad (32)$$

where the (normalized) cross time-frequency distribution of x and y is defined as

$$C_{xy}(k, f) = \frac{|X(k, f)Y^*(k, f)|}{\sum_{k,f} |X(k, f)Y^*(k, f)|}. \quad (33)$$

In the following, we will list several information-theoretic measures that quantify the *dissimilarity* (or “distance”) between two random variables (or signals). In contrast to the previously mentioned measures, those divergence measures vanish if the random variables (or signals) are *identical*; moreover, they are not necessarily symmetric, and therefore, they can not be considered as distance measures in the strict sense. Divergences may be computed in time domain and time-frequency domain; in this paper, we will only compute the divergence measures in time-frequency domain, since the computation in time domain is far more involved.

⁶ Program code (in C) is available at www.klab.caltech.edu/~kraskov/MILCA/

2.9.2 Kullback-Leibler Divergence

The Kullback-Leibler divergence is a well-known distance measure:

$$K(C_x, C_y) = \sum_{k,f} C_x(k, f) \log \frac{C_x(k, f)}{C_y(k, f)}. \quad (34)$$

This measure diverges when the two distributions are disjoint. The Kullback-Leibler divergence is an asymmetric measure; it can be symmetrized, for example, by taking the average of $K(C_x, C_y)$ and $K(C_y, C_x)$ (Johnson & Sinanovic, 2001):

$$K(C_x; C_y) = \frac{K(C_x, C_y) + K(C_y, C_x)}{2}. \quad (35)$$

2.9.3 Rényi Divergence

Rényi divergence is a generalized formulation of Kullback-Leibler divergence:

$$D_\alpha(C_x, C_y) = \frac{1}{\alpha - 1} \log \sum_{k,f} [C_x(k, f)]^\alpha [C_y(k, f)]^{(1-\alpha)}, \quad (36)$$

where $\alpha \in [0, 1]$ is the order of the divergence. This measure is asymmetric (except if $\alpha = 0.5$), and converges to the Kullback-Leibler distance as $\alpha \rightarrow 1$.

2.9.4 Jensen-Shannon Divergence

An alternative divergence measure is Jensen-Shannon divergence:

$$J(C_x, C_y) = H\left(\frac{C_x + C_y}{2}\right) - \frac{H(C_x) + H(C_y)}{2}, \quad (37)$$

where H stands for Shannon entropy of normalized spectrograms (31):

$$H(C_x) = - \sum_{k,f} C_x(k, f) \log C_x(k, f). \quad (38)$$

Interestingly, $J(C_x, C_y)$ is symmetric.

2.9.5 Jensen-Rényi Divergence

Jensen-Rényi divergence extends Jensen-Shannon divergence from arithmetic to geometric mean, by using the Rényi entropy instead of the Shannon entropy. Rényi entropy can be defined for normalized spectrograms (31) as:

$$H_\alpha(C_x) = \frac{1}{1 - \alpha} \log \sum_{k,f} (C_x(k, f))^\alpha \quad (39)$$

Rényi entropy converges to Shannon entropy as $\alpha \rightarrow 1$. For two positive time-frequency distributions, Jensen-Rényi divergence is obtained as:

$$J_\alpha(C_x, C_y) = H_\alpha\left(\sqrt{C_x(k, f)C_y(k, f)}\right) - \frac{H_\alpha(C_x) + H_\alpha(C_y)}{2}, \quad (40)$$

which is a symmetric measure (as Jensen-Shannon divergence).

2.10 Stochastic Event Synchrony (SES)

Stochastic event synchrony (SES), a family of synchrony measures that we developed in recent work (Dauwels *et al.*, 2007, 2009a,b), quantifies the similarity between point processes. The latter may be defined in time, frequency, space or any other domain. In this paper, we extract point processes from the time-frequency representations of the EEG signals. We will in this section outline the main ideas behind stochastic event synchrony. In the Supplementary Material, we provide more details on the implementation, and we refer the interested reader to (Dauwels *et al.*, 2007, 2009a,b) for additional information; an other application of SES to EEG signals is presented in (Vialatte *et al.*, 2009a).

As a first step, the time-frequency transform of each signal is approximated as a sum of half-ellipsoid basis functions, referred to as “bumps”⁷ (see Fig. 1 and (Vialatte *et al.*, 2007, 2009b)); each bump may be considered as an event on the time-frequency plane, and the resulting bump models E and E' may be considered as two-dimensional point processes (“event sequences”), representing the most prominent oscillatory activity. Only bumps whose energy is larger than a threshold T are retained in the bump model (see Supplementary Material).

Next the two bump models are aligned (see Fig. 2 and 3): bumps in one time-frequency map may not be present in the other map (“non-coincident” bumps); other bumps are present in both maps (“coincident” bumps), but appear at slightly different positions on the maps.

The black lines in Fig. 3 connect the centers of coincident bumps, and hence, visualize the offset in position between pairs of coincident bumps. Stochastic event synchrony (SES) consists of five parameters that quantify the alignment of two bump models:

- ρ : fraction of non-coincident bumps,

⁷ An implementation of the bump extraction algorithm is available from http://www.bsp.brain.riken.jp/~fvialatte/bumptoolbox/toolbox_home.html.

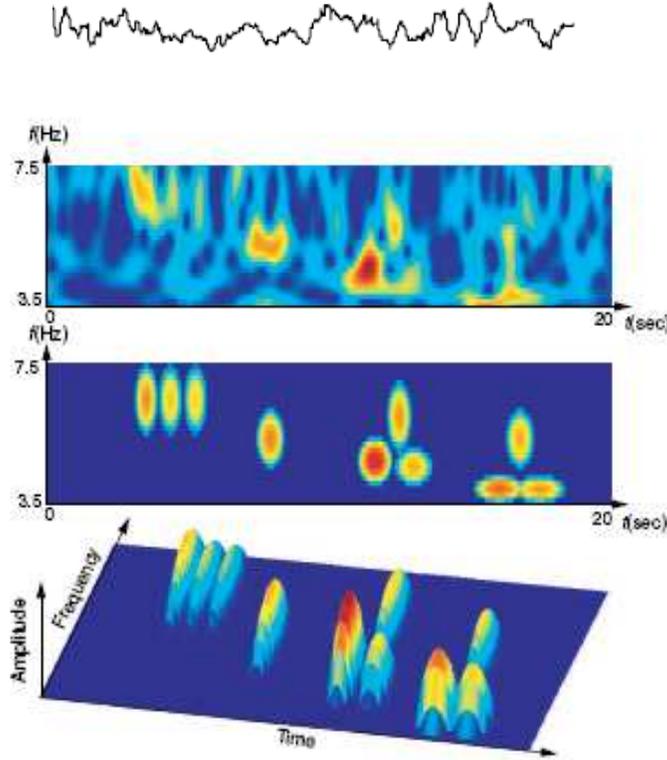


Fig. 1. Bump modeling. From top to bottom: EEG signal in time-domain; EEG signal in time-frequency domain; Bump model of the EEG signal (two-dimensional and three-dimensional view).

- δ_t and δ_f : average time and frequency offset respectively between coincident bumps,
- s_t and s_f : variance of the time and frequency offset respectively between coincident bumps.

The parameters ρ and s_t are the most relevant for the present study, since they quantify the synchrony between bump models and hence the original time-frequency maps; low ρ and s_t implies that the two time-frequency maps at hand are well synchronized.

The SES parameters are computed iteratively⁸. One starts with an initial guess $\hat{\delta}_t^{(0)}$, $\hat{\delta}_f^{(0)}$, $\hat{s}_t^{(0)}$, $\hat{s}_f^{(0)}$, and aligns the two bump models. Using this alignment, one updates the SES parameters $\hat{\delta}_t$, $\hat{\delta}_f$, \hat{s}_t , \hat{s}_f , and with those estimates, one re-aligns the bump models, etc., until convergence. From the final alignment, one can directly compute an estimate of ρ . More details on this procedure are provided in the Supplementary Material.

⁸ An implementation of SES for one-dimensional and multi-dimensional point processes is available from <http://www.dauwels.com/SESToolbox/SES.html>.

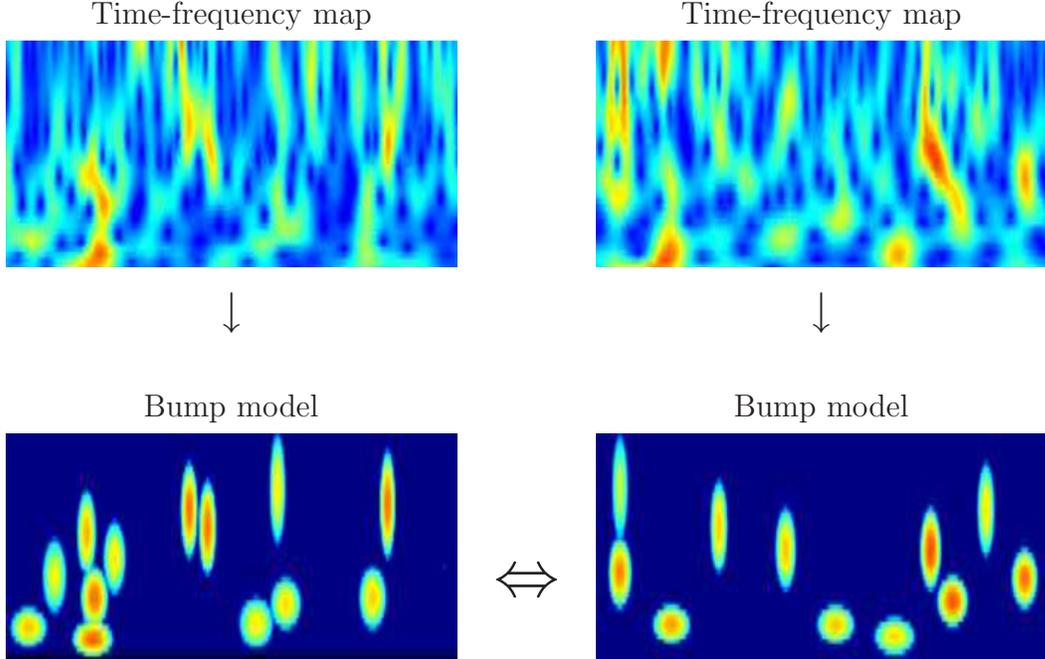


Fig. 2. Two-dimensional stochastic event synchrony. Top: two given EEG signals in time-frequency domain; Bottom: bump models extracted from those time-frequency maps. Stochastic event synchrony quantifies the similarity of two such bump models.

Important parameters in the SES method are the initial estimates $\hat{\delta}_t^{(0)}$, $\hat{\delta}_f^{(0)}$, $\hat{s}_t^{(0)}$, $\hat{s}_f^{(0)}$, and the threshold T . In addition, the SES model depends on a parameter β that controls the number of matches (see Supplementary Material). Indeed, it is easy to verify that the problem of aligning bump models is ambiguous. For example, one may only allow extremely small offsets, and as a result, virtually all bumps would be non-coincident; in the other extreme, one may allow unrealistically large offsets, in which case almost all bumps are coincident. The parameter β allows us to sweep between both extremes in order to obtain reasonable alignments and hence plausible estimates of the SES parameters.

2.11 Local vs. Global Synchrony

As we pointed out in the introduction, the measures we have reviewed in the above may be divided into two separate classes: *local* and *global* methods. The global methods in the above list are the Granger causality measures, MVAR coherence, GFS, Omega complexity, and the S-estimator, all other measures in the above are local.

Interestingly, the Granger causality measures and MVAR coherence provide also local synchrony information since from the underlying MVAR model,

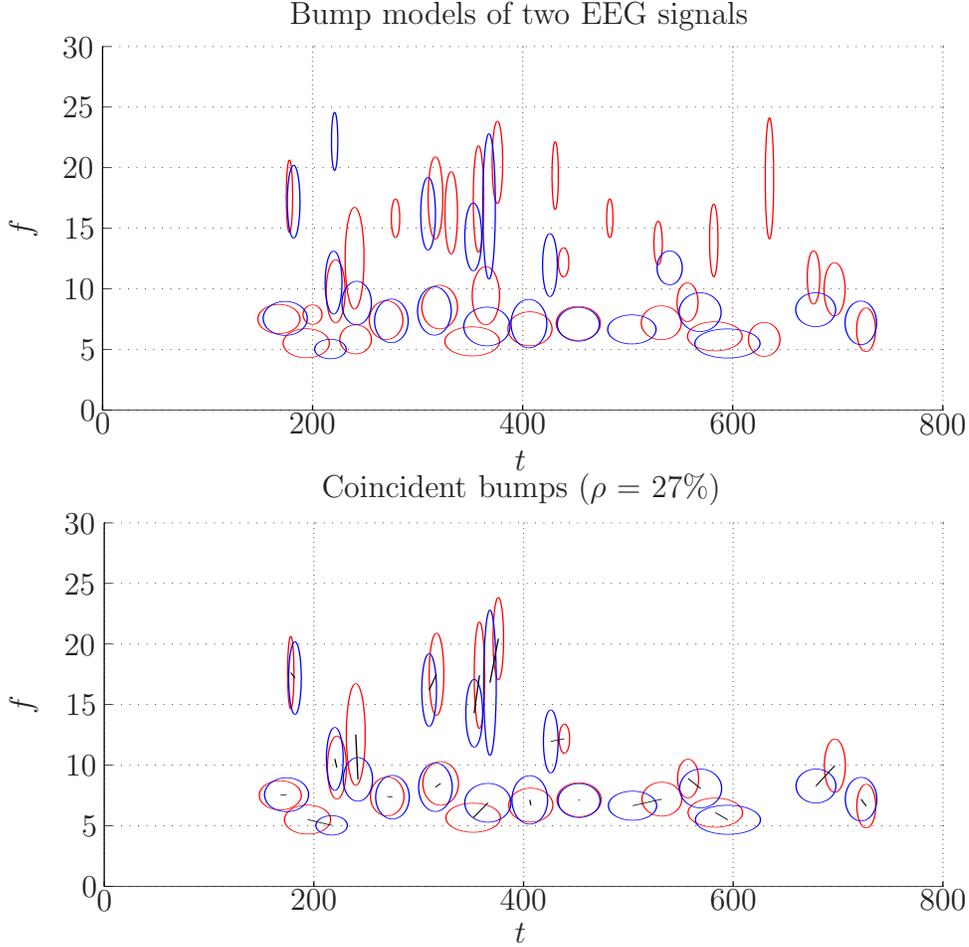


Fig. 3. Non-coincident and coincident activity (“bumps”). Top: Bump models of two signals (red and blue); one can observe pairs of bumps that are coincident (matched), other bumps are not overlapping and cannot be matched to bumps from the other bump model. Bottom: Coincident bumps; the black lines connect the centers of coincident bumps.

one can extract interactions between individual signals. In contrast, the other global measures considered in this paper, i.e., GFS, Omega complexity, and the S-estimator, do not provide any local information when applied to all EEG signals simultaneously. One may obtain local information by applying them to *subsets* of signals, as we will explain in Section 3.

We reiterate that we are fully aware of the drawbacks of both local and global synchrony measures. Our core objective is to use those synchrony measures as features to distinguish MCI patients from control subjects; to this end, we apply a large variety of synchrony methods to the *same* EEG data set of MCI and control subjects, a study that has not been carried out yet. Given the shortcomings and limitations of the synchrony measures, we will not try to unravel the biophysical mechanisms that cause the observed effects; however, we will suggest some possible theories.

3 Diagnosis of EEG Synchrony Abnormalities in AD Patients

In the following section, we describe the EEG data we analyzed. In Section 3.2 we address certain technical issues related to the synchrony measures, and in Section 3.3, we present and discuss our results.

3.1 EEG Data

The EEG data used here have been analyzed in previous studies concerning early diagnosis of AD (Chapman *et al.*, 2007; Cichocki *et al.*, 2005; Hogan *et al.*, 2003; Musha *et al.*, 2002; Vialatte *et al.*, 2005).

Ag/AgCl electrodes (disks of diameter 8mm) were placed on 21 sites according to 10–20 international system, with the reference electrode on the right ear-lobe. EEG was recorded with Biotop 6R12 (NEC San-ei, Tokyo, Japan) at a sampling rate of 200Hz, with analog bandpass filtering in the frequency range 0.5-250Hz and online digital bandpass filtering between 4 and 30Hz, using a third-order Butterworth filter. We used a common reference for the data analysis (right ear-lobe), and did not consider other reference schemes (e.g., average or bipolar references).

The subjects comprise two study groups. The first consists of 25 patients who had complained of memory problems. These subjects were diagnosed as suffering from mild cognitive impairment (MCI) when the EEG recordings were carried out. Later on, they all developed mild AD. The criteria for inclusion into the MCI group were a mini mental state exam (MMSE) score = 24, though the average score in the MCI group was 26 (SD of 1.8). The other group is a control set consisting of 56 age-matched, healthy subjects who had no memory or other cognitive impairments. The average MMSE of this control group is 28.5 (SD of 1.6). The ages of the two groups are 71.9 ± 10.2 and 71.7 ± 8.3 , respectively. Finally, it should be noted that the MMSE scores of the MCI subjects studied here are quite high compared to a number of other studies. For example, in (Hogan *et al.*, 2003) the inclusion criterion was MMSE = 20, with a mean value of 23.7, while in (Chapman *et al.*, 2007), the criterion was MMSE = 22 (the mean value was not provided); thus, the disparity in cognitive ability between the MCI and control subjects is comparatively small, making the present classification task relatively difficult.

All recording sessions were conducted with the subjects in an awake but resting state with eyes closed; the EEG technicians prevented the subjects from falling asleep (vigilance control). The length of the EEG recording is about 5 minutes, for each subject. After recording, the EEG data has been carefully inspected. Indeed, EEG recordings are prone to a variety of artifacts, for example due to

electronic smog, head movements, and muscular activity. For each patient, an EEG expert selected by visual inspection one segment of 20s artifact free EEG, blinded from the results of the present study. Only those subjects were retained in the analysis whose EEG recordings contained at least 20s of artifact-free data. Based on this requirement, the number of subjects in the two groups described above was further reduced to 22 and 38, respectively. From each subject, *one* artifact-free EEG segment of 20s was analyzed (for each of the 21 channels).

In the following sections, we will describe how we applied the synchrony measures to those EEG segments.

3.2 Methods

3.2.1 Different Approaches to Compute EEG Synchrony

In order to compute EEG synchrony, we first aggregated the EEG signals into 5 zones, as illustrated in Fig. 4. We quantified the EEG synchrony between those 5 different regions using each synchrony measure of Section 2. It is noteworthy that the EEG synchrony between those regions needs to be considered as *large-scale* synchrony, since each region spans several tens of millimeters. In the following, we detail how we applied the different synchrony measures to determine large-scale synchrony.

Local Synchrony Methods

Before applying the *local* synchrony measures, we first determine the arithmetic average of the signals within each zone. For example, we determine the average EEG in the left temporal zone by computing the arithmetic average of the EEG at channels F_7 , T_3 , and T_5 . Next we apply the local synchrony methods to each of the 10 pairs of average signals. We refer to this approach as “Local Approach 1”.

We also followed an alternative approach (Local Approach 2): in order to determine the synchrony between two zones, we first evaluated the synchrony between each EEG signal from one zone and each signal from the other. Next we averaged over all those signal pairs. For example, the synchrony between the left and right temporal region is determined by averaging the synchrony measures for the 9 pairs (F_7, F_8) , (F_7, T_4) , (F_7, T_6) , \dots , (T_5, T_6) . We apply this approach to all local measures except the information-theoretic measures (cf. Section 2.9), the wavelet-based phase synchrony index γ_W , the wav-entropy coefficient w_E , and SES. Local Approach 2 is not feasible for the information-theoretic measures, wavelet-based phase synchrony index γ_W and wav-entropy coefficient w_E , since it is too computationally complex. In the case of SES, we apply the following procedure: first we compute a bump model for each of the

21 EEG signals. We aggregate those 21 models into 5 models, one for each zone, by means of the algorithm described in (Vialatte *et al.*, 2007). We then compute the SES parameters for each pair of the resulting 5 bump models. We refer to that approach as Local Approach 2 for SES; it is a natural alternative to Local Approach 1.

Global Synchrony Methods

Global synchrony methods can in principle be applied to all 21 EEG signals simultaneously; this is a reasonable idea for GFS, Omega complexity, and the S-estimator. However, for the Granger measures and MVAR coherence, this would involve estimating a 21-dimensional MVAR model with 21^2 times p parameters (cf. (10)). The EEG signals at hand are too short (20s) to estimate such large number of parameters reliably. Moreover, EEG signals are non-stationary, and therefore, the MVAR models need to be trained from even shorter EEG segments (e.g., 1s or 5s).

Therefore, in order to reduce the dimensionality, we apply the Granger measures to the 5 average EEG signals and learn a 5-dimensional MVAR models (cf. (10)) from those averages. For the sake of consistency, we also compute GFS, Omega complexity, and the S-estimator from the 5 average EEG signals. We refer to this approach as “Global Approach 1”. Also in Local Approach 1, synchrony measures are applied to the 5 average EEG signals, and therefore, results from Local Approach 1 and Global Approach 1 are *directly* comparable.

In addition, we apply GFS, Omega complexity, and the S-estimator to all 21 EEG signals (Global Approach 2). We also apply those measures to all pairs of zones (Global Approach 3). More specifically, we evaluate Omega complexity and the S-estimator index between each pair of zones by applying PCA to the EEG signals and state space embedded EEG signals respectively within the two zones. Likewise, we compute GFS by applying PCA to the Fourier transformed EEG signals within each pair of zones. For example, in order to compute GFS index between the left and right temporal zone, we apply PCA to the Fourier transformed EEG signals recorded at channels F_7 , F_8 , T_3 , T_4 , T_5 , and T_6 .

3.2.2 Average EEG Synchrony

In this study, we solely consider spatially averaged EEG synchrony, we do not investigate the *spatial distribution* of EEG synchrony; after applying the synchrony measures, we average the synchrony values over the 5 regions (cf. Fig. 4) to obtain an average measure of EEG synchrony. More precisely, in Local Approach 1 and 2 and Global Approach 3, we average the pairwise synchrony values over the 10 pairs of regions; we also follow this averaging procedure in the case of Granger measures (Global Approach 1). Note that the other global

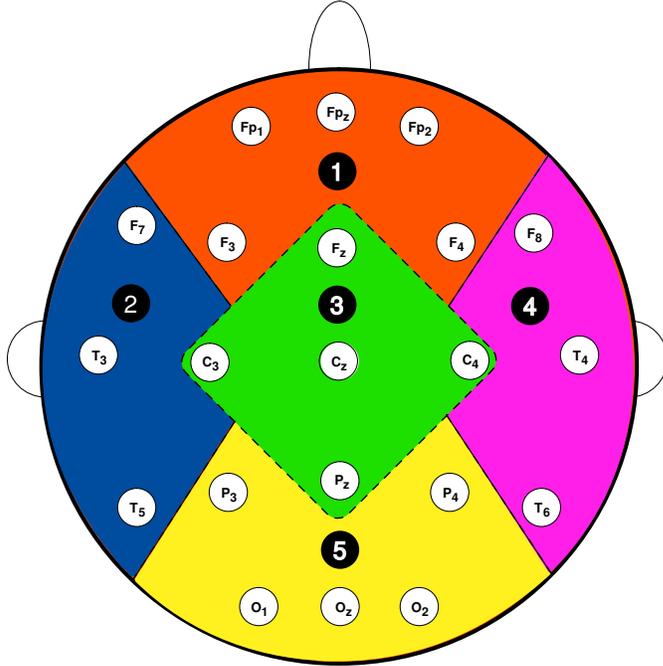


Fig. 4. The 21 channels used for EEG recording, distributed according to the 10–20 international placement system (Nunez & Srinivasan, 2006). The clustering into 5 zones is indicated by the colors and dashed lines (1 = frontal, 2 = left temporal, 3 = central, 4 = right temporal and 5 = occipital).

synchrony measures (GFS, Omega complexity, and the S-estimator) directly quantify average EEG synchrony in Global Approach 1 and 2, there is no need for additional averaging.

3.2.3 Frequency-Dependent Synchrony Measures

Several measures depend on the frequency f , e.g., the magnitude square coherence function $c(f)$ (2). We compute those measures for all (integer) frequencies within the frequency band at hand (4–30Hz). Next we average the measures over those frequencies. For the example, in case of the magnitude square coherence function $c(f)$, we compute:

$$c = \frac{1}{27} \sum_{f=4}^{30} c(f). \quad (41)$$

Alternatively, one may compute frequency-dependent measures at the center frequency of the frequency band at hand, for example at 17Hz for the 4–30Hz band. However, that frequency band is wide and the center frequency is not necessarily representative; one would neglect the contributions from other

frequencies in that band.

3.2.4 *Non-stationarity*

Spontaneous EEG is usually highly non-stationary, which is an important issue that needs to be taken into account before applying synchrony measures. The wavelet-based synchrony measures are directly applicable to non-stationary signals, in particular, the wavelet-based divergence measures (Kullback-Leibler, Rényi, Jensen-Shannon, and Jensen-Rényi divergence), wavelet-based mutual information I_W , the wavelet and Hilbert phase indices ϕ^W and ϕ^H respectively, wave-entropy w_E , and the SES measures (ρ and s_t). The other synchrony measures considered in this paper, however, are strictly speaking only applicable to stationary signals. Therefore, we computed those measures by averaging over M short segments of the EEG. All M segments have equal length L (cf. Section 2.2). Since the statistics of the EEG strongly fluctuates over time, the length L should be sufficiently small. On the other hand, in order to obtain reliable measures for synchrony, the length should be chosen sufficiently large. For example, in order to compute the Granger measures, one needs to determine the matrices \mathbf{H} , \mathbf{A} and \mathbf{S} . If the EEG signals were stationary, one should obviously use the entire EEG signals to estimate the coefficient matrices. However, since the EEG is non-stationary, those matrices may vary over time; it is then not necessarily meaningful to estimate them using the entire EEG signals (20s), instead one may divide the EEG signals in segments, and estimate separate matrices from each EEG segment. From each of those matrices, one computes Granger measures, and eventually one averages those measures over all EEG segments.

Since it is not a priori clear how to choose the length L , we tested several values, i.e., $L = 1\text{s}$, 5s , and 20s ; the latter is length of the entire EEG signal selected for the analysis at hand (see Section 3.1). In other words, we apply the measures to EEG segments of either 1 or 5s, and we also apply them to the *entire* EEG signal, of length 20s.

3.2.5 *Statistical Analysis*

We investigate whether there are statistically significant differences in EEG synchrony between MCI patients and age-matched control subjects. To this end, we apply the Mann-Whitney test to the EEG synchrony values of MCI patients and age-matched control subjects. In particular, we conduct a Mann-Whitney test for each individual synchrony measure, each approach to compute the synchrony measures (Local Approach 1 and 2, Global Approach 1, 2, and 3), and each parameter setting, e.g., different EEG segment length L . (The Supplementary Material contains detailed information about the param-

eter settings.) The Mann-Whitney test allows us to investigate whether the statistics at hand, in particular, the synchrony measures, take different values for the two subject populations. Low p-values indicate large difference in the medians of the two populations.

Since we consider many different synchrony measures, various approaches to compute those synchrony measures (Local Approach 1 and 2, Global Approach 1, 2, and 3), and many different parameter settings simultaneously, it is likely that a few p-values are small merely due to stochastic fluctuations and *not* due to systematic difference between MCI patients and control subjects; as a consequence, the p-values need to be appropriately corrected. One may try to control the probability that a false positive occurs by applying Bonferroni post-correction (Bonferroni, 1936). Since we are dealing with many different simultaneous tests (1270 tests in total), it is more natural to try to control the false discovery rate: the *fraction* of supposedly positive results that are in fact negative (Benjamini and Hochberg, 1995). In other words, instead of controlling the *absolute number* of false positives (which would be very conservative in our problem setting, see, e.g., (Benjamini and Hochberg, 1995; Storey, 2002; Vialatte and Cichocki, 2008)), we will bound the *fraction* of false positives within the supposedly positive results. The method of Benjamini-Hochberg (Benjamini and Hochberg, 1995) has recently be refined by Storey (Storey, 2002); the latter method is more powerful, however, it is not suitable for a small number of comparisons.

We consider two different schemes for statistical post-correction:

- (1) We apply the Storey method (Storey, 2002) to all 1270 comparisons simultaneously, that is all synchrony measures, the different approaches to compute those measures (Local Approach 1 and 2, Global Approach 1, 2 and 3), and their parameter settings. We retain the smallest p-value for each synchrony measure, and as a result, we obtain one (corrected) p-value for each synchrony measure.
- (2) We apply a two-step procedure: first we correct for the multiple approaches to compute each synchrony measure (Local Approach 1 and 2, Global Approach 1, 2 and 3) and the multiple corresponding parameter settings, then we correct for the multiple synchrony measures. In the first step, we treat each synchrony method separately, and apply the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) to the p-values for different parameter settings and computational approaches (Local Approach 1 and 2, Global Approach 1, 2 and 3). For each synchrony measure, we retain the smallest corrected p-value among the different parameter settings and computational approaches. As a result, we obtain one (corrected) p-value for each synchrony measure. Those p-values are further corrected in a second step by means of the Storey method (Storey, 2002); that correction accounts for the multiple synchrony measures.

In the first step of the latter scheme, we do not apply the Storey method since for some synchrony measures, there are only few different parameter settings; the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) is then more accurate.

Interestingly, as we will show later on, both statistical post-correction schemes lead to the same positive results. More details can be found in the Supplementary Material.

In addition to the tests of statistical significance, we investigate whether loss of EEG synchrony allows us to separate MCI patients from age-matched control subjects. Using the synchrony measures that yield statistically significant differences, we conduct linear and quadratic discriminant analysis with leave-one-out crossvalidation (Duda *et al.*, 1992).

3.3 Results and Discussion

We observed that most synchrony measures indicate decreased EEG synchrony in MCI patients. However, for only two measures, this effect is statistically significant. In the following we provide more detailed numerical results. Tables 1 to 3 show the sensitivity of the synchrony measures for detecting loss in EEG synchrony: they contain (uncorrected) p-values obtained by the Mann-Whitney test. We also indicate in those tables which synchrony measures yield significant differences after post-correction (cf. Section 3.2.5). In addition, we list (in brackets) the corrected p-values obtained through statistical post-correction scheme 1 (cf. Section 3.2.5). The results of statistical post-correction scheme 2 can be found in the Supplementary Material. Note that Tables 1 to 3 contain the corrected p-values for the individual computational approaches (Local/Global Approach 1 to 3); in the Supplementary Material, we list the smallest corrected p-values among the different computational approaches, we do not list corrected p-values for the different computational approaches separately.

Table 1 contains results for Local Approach 1 and Global Approach 1. Since in both approaches, the synchrony methods are applied to the same signals, i.e., the 5 local averages (cf. Fig. 4), differences in p-values are due to the measures themselves, and not to differences in preprocessing. As we pointed out earlier, for certain synchrony methods, alternative preprocessing techniques may be more adequate. Therefore we consider alternative preprocessing methods for local and global measures, i.e., Local Approach 2 and Global Approach 2 and 3 respectively (cf. Section 3.2.4). Results for Local Approach 2 and Global Approach 2 and 3 are presented in Table 2 and 3 respectively.

From Table 1 to 3, it becomes clear that several measures yield significant

results on the $p = 0.05$ level (without post-correction): correlation coefficient, magnitude coherence, phase coherence, corr-entropy, ffDTF, dDTF, N^k , H^k , Instantaneous Period, GFS, S-estimator, s_t , and ρ . Most of those measures, however, do not remain significant after statistical post-correction (see Supplementary Material); only two measures remain significant (with false discovery rate $\leq 5\%$) after correcting for the large number of measures and parameter settings: full-frequency DTF (ffDTF) and ρ .

Table 4 shows our results for discriminant analysis (DA) with as features full-frequency DTF (ffDTF) and ρ , the two synchrony measures that yield the most statistically significant differences; we analyzed those two measures separately and jointly, and applied linear as well as quadratic discriminant analysis (with diagonal covariance matrix estimates) (Duda *et al.*, 1992). The classification rates are obtained through leave-one-out crossvalidation (Duda *et al.*, 1992). From Table 4 we can see that the combination of full-frequency DTF and ρ is to some extent able to discriminate MCI patients from control subjects; this is illustrated in Fig. 5, for one particular parameter setting and classifier. Since the classifiers in linear/quadratic discriminant analysis depend on few parameters and the input space is only two-dimensional, it is not implausible that the classification results generalize to other data sets. In other words, this study seems to suggest that MCI-induced loss of EEG synchrony can be detected, as was reported earlier in the literature. We will expand on this issue in the following section.

As we pointed out earlier, the results we have presented so far are obtained after bandpass filtering the EEG signals in the frequency range 4–30Hz (see Section 3.1). However, we have also considered several components within that frequency range: 4–8Hz (theta), 8–10Hz (alpha 1), 10–12Hz (alpha 2), and 12–30Hz (beta). We did not observe any consistent and significant MCI induced perturbation in EEG synchrony for those narrower frequency bands (not shown here). Note also that by analyzing separate frequency bands, we introduce more degrees of freedom to the analysis; as a result, statistical post-correction becomes more severe, and positive results may vanish after post-correction. For those reasons and for the sake of conciseness, we decided not to include results for individual frequency components in this report; we limit ourselves here to the frequency range 4–30Hz.

Although the classification results of Table 4 are promising, the separation is not sufficiently strong to yield reliable early prediction of AD. For this purpose, the two features (ffDTF and ρ) need to be combined with complementary features of the EEG (e.g., spectral properties of the EEG, affected through the slowing effect of AD on EEG), or perhaps from different modalities such as PET, MRI, DTI, or biochemical indicators. On the other hand, we remind the reader of the fact that in the data set at hand, patients did not carry out any specific task; moreover, the recordings were short (only 20s). It is plausible that

Measure	Correlation	Coherence	Phase Coherence	Corr-entropy	Wave-entropy	
p-value	0.025* (0.059)	0.029* (0.064)	0.041* (0.077)	0.032* (0.067)	0.096 (0.11)	
References	(Nunez & Srinivasan, 2006)			(Gunduz and Principe, 2009)		
Measure	MVAR coherence	Partial Coherence	PDC	DTF	fDTF	dDTF
p-value	0.15 (0.13)	0.16 (0.14)	0.60 (0.32)	0.29 (0.19)	0.0012**† (0.0098*)	0.029* (0.064)
References	(Kamiński <i>et al.</i> , 2005)					
Measure	Kullback-Leibler	Rényi	Jensen-Shannon	Jensen-Rényi	I_W	I
p-value	0.065 (0.096)	0.067 (0.096)	0.069 (0.096)	0.074 (0.097)	0.052 (0.088)	0.060 (0.094)
References	(Aviyente, 2005a)					(Kraskov <i>et al.</i> , 2004)
Measure	N^k	S^k	H^k	S-estimator	Omega complexity	
p-value	0.029* (0.064)	0.045* (0.080)	0.052 (0.088)	0.042* (0.077)	0.079 (0.10)	
References	(Quiroga <i>et al.</i> , 2002)			(Carmeli <i>et al.</i> , 2005)	(Saito <i>et al.</i> , 1998)	
Measure	Hilbert Phase	Wavelet Phase	Evolution Map	Instantaneous Period	GFS	
p-value	0.96 (0.42)	0.082 (0.10)	0.64 (0.33)	0.73 (0.36)	0.031* (0.066)	
References	(Lachaux <i>et al.</i> , 1999)		(Rosenblum <i>et al.</i> , 2002)		(Koenig <i>et al.</i> , 2001)	
Measure	s_t	ρ				
p-value	0.012* (0.040*)	0.00044** (0.037*)				
References	(Dauwels <i>et al.</i> , 2007)					

Table 1

Sensitivity of average synchrony for prediction of MCI, following Local and Global Approach 1: uncorrected p-values for Mann-Whitney test; * and ** indicate $p < 0.05$ and $p < 0.005$ respectively; † indicates p-values that remain significant after post-correction. The corrected p-values (from post-correction method 1) are reported in brackets. Since Local and Global Approach 1 are based on the *same* 5 local averages, the p-values are directly comparable.

the sensitivity of EEG synchrony could be further improved by increasing the length of the recordings and by recording the EEG before, while, and after patients carry out specific tasks, e.g., working memory tasks. On the other hand, the present method might be applied to screen a population for AD, since it only requires an EEG recording system, which is a relatively simple and low-cost technology, at present available in most hospitals. Moreover, recently developed portable and wireless EEG systems make it possible to record EEG on virtually any location (see, e.g., (Jun *et al.*, 2005; Carmo *et al.*, 2006)). Alternative imaging technologies such as MRI and DTI are far less flexible in terms of mobility, and they are vastly more expensive.

In order to gain more insight in the relation between the different synchrony measures, we calculated the correlation between those measures (see Fig. 6). We calculated the correlation coefficient between all pairs of synchrony mea-

Measure	Correlation	Coherence	Phase Coherence	N^k	S^k	H^k
p-value	0.018* (0.052)	0.062 (0.094)	0.61 (0.32)	0.054 (0.090)	0.019* (0.054)	0.15 (0.13)
Measure	Hilbert Phase	Evolution Map	Instantaneous Period	Corr-entropy	s_t	ρ
p-value	0.15 (0.13)	0.072 (0.096)	0.020* (0.054)	0.036* (0.071)	0.0065* (0.027*)	0.00012***† (0.005**)

Table 2

Sensitivity of average synchrony for prediction of MCI, following Local Approach 2: uncorrected p-values for Mann-Whitney test; * and ** indicate $p < 0.05$ and $p < 0.005$ respectively, and † indicates p-values that remain significant after post-correction. The corrected p-values (from post-correction method 1) are reported in brackets.

	Global Approach 2	Global Approach 3
GFS	0.36 (0.22)	0.17 (0.14)
Omega complexity	0.47 (0.27)	0.46 (0.26)
S-estimator	0.36 (0.22)	0.33 (0.21)

Table 3

Sensitivity of average synchrony for prediction of MCI, following Global Approach 2 and 3: uncorrected p-values for Mann-Whitney test. The corrected p-values (from post-correction method 1) are reported in brackets.

	Linear DA	Quadratic DA
#DTF	70.0%	70.0%
ρ	68.3%	75%
#DTF and ρ	83.3%	83.3%

Table 4

Classification rates for discriminant analysis (DA) of full-frequency DTF and ρ , determined through leave-one-out crossvalidation.

asures:

$$r_{ij} = \frac{1}{N_{\text{subject}}} \sum_{k=1}^{N_{\text{subject}}} \frac{(m_i(k) - \bar{m}_i)(m_j(k) - \bar{m}_j)}{\sigma_i \sigma_j}, \quad (42)$$

where $m_i(k)$ and $m_j(k)$ is the average value of measure i and j respectively for subject k , the sum is computed over all 60 subjects ($N_{\text{subject}} = 60$), and \bar{m}_i , \bar{m}_j , σ_i , and σ_j are the mean and standard deviation of m_i and m_j respectively, also computed over all subjects. A similar study has been carried out in (Jellens *et al.*, 2008) for a few synchrony measures. From Fig. 6, it becomes strikingly clear that the majority of measures are strongly correlated (or anti-correlated) with each other. In other words, the measures can easily be classified in different families. More concretely, one can distinguish the following 9 families of measures:

- (1) correlation coefficient, cross-entropy, wav-entropy, state-space measures (N^k , H^k , S^k , Ω , S-estimator), mutual information (in time and frequency domain), divergence measures, MVAR coherence, Hilbert and wavelet phase,
- (2) GFS,
- (3) phase coherence,
- (4) EMA and IPA (phase measures),
- (5) partial coherence (PC) and direct DTF (Granger),
- (6) PDC and DTF (Granger),
- (7) full frequency DTF (Granger),
- (8) s_t (SES),
- (9) ρ (SES).

Interestingly, the first family contains the correlation coefficient; all measures from that family are strongly correlated (or anti-correlated) with the correlation measure, and therefore, provide little additional information about synchrony, at least for the EEG data set at hand. Measures that are only weakly correlated with the correlation coefficient include the phase synchrony indices, Granger causality measures, and stochastic event synchrony measures. Interestingly, those three families of synchrony measures are mutually uncorrelated, and as a consequence, they each seem to capture a specific kind of interdependence. Moreover, most Granger measures seem to be largely mutually uncorrelated, the same holds for the SES parameters s_t and ρ , and some phase synchrony measures.

In parallel work, Jellens *et al.* applied the correlation coefficient and the correlation dimension D2 (Grassberger and Procaccia, 1983) to EEG of mild to moderate AD patients and control subjects, and observed that both measures are strongly correlated (Jellens *et al.*, 2008). The correlation dimension D2 is a state space measure similar to the S^k , H^k , and N^k indices (Grassberger and Procaccia, 1983).

Fig. 6 seems also to suggest that in order to quantify EEG synchrony, it is not necessary to apply the whole set of synchrony measures considered in this paper. Instead it may suffice to evaluate only a few measures from each of the families, i.e., correlation coefficient, phase synchrony indices, Granger measures, and SES.

4 Conclusions

A variety of studies have investigated the effect of Alzheimer’s disease on EEG synchrony. However, it is not always obvious how to compare those studies to our work:

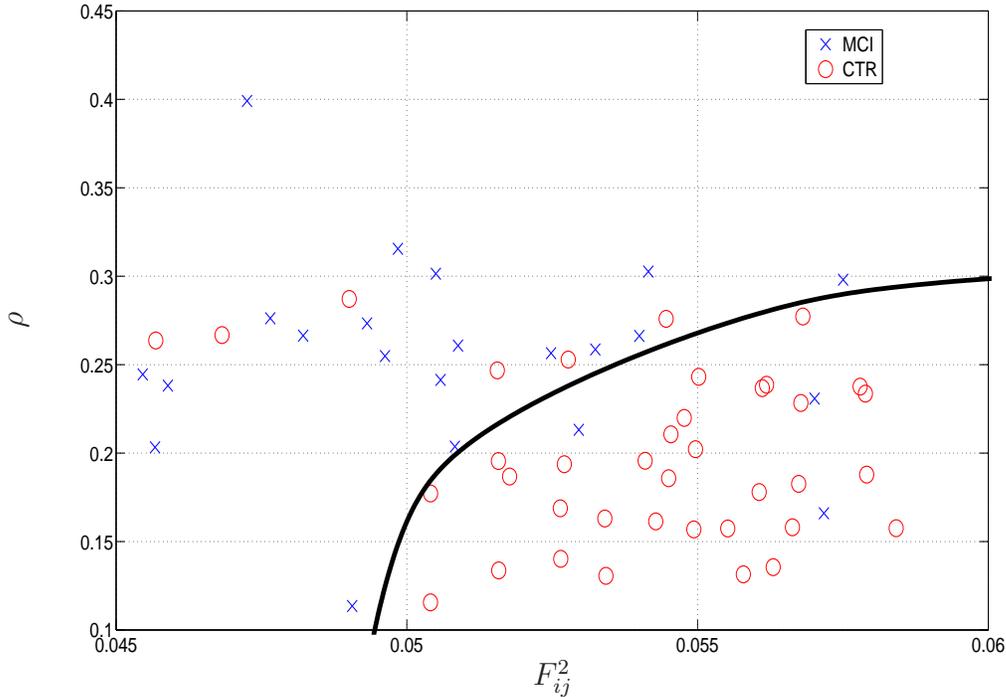


Fig. 5. ρ vs. ffDTF

- Several synchrony measures considered in this study have not been applied before to MCI or AD data,
- In some studies the EEG was recorded under different conditions, e.g., during working memory tasks instead of resting condition,
- Some studies consider different subject groups, for example, probable-MCI patients that do not necessarily evolve towards AD.

We first briefly comment on studies that are related but not directly comparable to our work. In some studies that investigate working memory tasks (Jiang, 2005a,b; Jiang & Zheng, 2006) or probable-MCI subjects (Rossini *et al.*, 2006), an *increase* of EEG synchrony was observed. This inverse effect is often interpreted as the result of a compensatory mechanism in the brain. In those studies, one applied magnitude coherence (cf. Section 2.2) as a measure for EEG synchrony. A similar effect was observed in (Pijnenburg *et al.*, 2004) by means of the synchronization likelihood measure, which is state-space based measure similar to the indices S^k , H^k , and N^k . In particular, in that study, one observed an increase of EEG synchrony during working memory task, but there was no significant effect in rest condition. We remind the reader of the fact that the data we investigated in this report were gathered in resting state, therefore, we could not study this inverse effect of increased EEG synchrony.

Now we proceed to studies that exclusively deal with subjects in resting condition. In most of those studies, brain dynamics in AD and MCI patients were

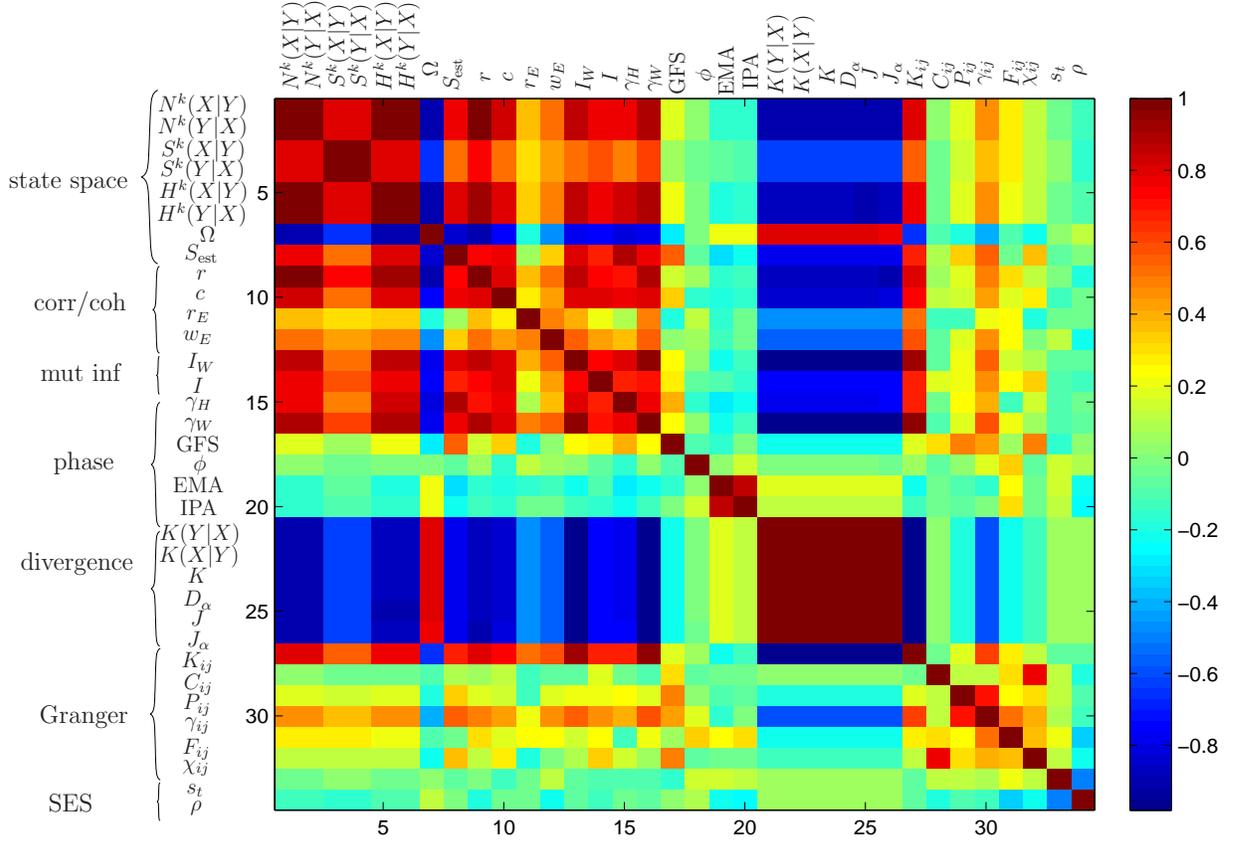


Fig. 6. Correlation between the synchrony measures (red and blue indicate strong correlation and anti-correlation respectively; Local and Global Approach 1).

mainly investigated using coherence (cf. Section 2.2) or state space based measures of synchrony (cf. Section 2.8). We provide a concise overview of existing studies and discuss them in the light of our own results:

- A large number of studies have reported a decrease in magnitude and phase coherence in the EEG of MCI and AD patients (Stevens *et al.*, 2001; Brassens *et al.*, 2004; Locatelli *et al.*, 1998; Wada *et al.*, 1998; Hogan *et al.*, 2003; Anghinah *et al.*, 2000; Besthorn *et al.*, 1994; Hidasi *et al.*, 2007; Jellens *et al.*, 2008; Stam, 2007; Güntekin *et al.*, 2008). In (Brassens *et al.*, 2004) those measures allowed to separate depressed MCI patients from control subjects. In some other studies, however, no effects on coherence were observed, neither between AD and controls (Stam *et al.*, 2002, 2003) nor between MCI and controls (Jiang, 2005a). We observed a weak decrease in magnitude coherence ($p = 0.029$; uncorrected) and phase coherence ($p = 0.041$; uncorrected).
- Koenig *et al.* (Koenig *et al.*, 2005) observed a general decrease of GFS (cf. Section 2.7) in correlation with cognitive decline and AD. In our study, GFS was lower for MCI patients, but this effect was not significant after statistical post-correction ($p = 0.031$; uncorrected). Recently, a decrease in phase synchrony has been reported in EEG of patients with mild AD (resting

condition), using three different measures: phase coherence, phase lag index (PLI), and imaginary component of coherence (Stam, 2007). Besides GFS, we analyzed five alternative phase synchrony measures: Hilbert and wavelet based phase synchrony, phase coherence, evolution map approach (EMA), and instantaneous period approach (IPA). The p-value of the latter is low ($p = 0.020$; uncorrected), in agreement with the results of (Koenig *et al.*, 2005), but this effect is non-significant after statistical post-correction.

- In a recent study, the state space based measure H^k decreased significantly in MCI and AD patients (Kramer *et al.*, 2007), however, the measure S^k did not decrease. Unfortunately, the measure N^k was not investigated in that study. Synchronization likelihood (Stam *et al.*, 2002), a state space based synchronization measure similar to the non-linear interdependence measures S^k , H^k , and N^k (cf. Section 2.8), is believed to be more sensitive than coherence to detect changes in AD patients (Stam *et al.*, 2003). Using state space based synchrony methods, significant loss of EEG synchrony was found in MCI (Stam *et al.*, 2003) and AD patients (Jeong, 2004; Stam *et al.*, 2003, 2005; Babiloni *et al.*, 2006; Pijnenburg *et al.*, 2004; Yagyu *et al.*, 1997; Wan *et al.*, 2008), which evolved consistently from MCI to AD stage (Stam *et al.*, 2003). We report here a low (uncorrected) p-value for the indices N^k ($p = 0.029$; uncorrected) and S^k ($p = 0.019$; uncorrected), and the S-estimator ($p = 0.042$; uncorrected); however, after post-correction, those effects are not statistically significant. Omega complexity was found to be greater in patients with mild AD, which corresponds to decreased synchronization (Yoshimura *et al.*, 2004). We also observed that effect in our data, but it was weak ($p = 0.079$; uncorrected).

In summary, it is not always straightforward to compare our results with previous reports due to significant differences in methodology and to inconsistencies caused by statistical fluctuations. Nevertheless, our results are generally consistent with most studies on the loss of average EEG synchrony in rest condition of MCI and AD patients. This decrease in synchrony is often attributed to a functional disconnection of the neocortex. If the decrease of EEG coherence were simply due to a loss of cortical neurons, it would be difficult to explain why all frequencies are not equally affected (Stam *et al.*, 2003). As a consequence, it is likely that the loss in EEG synchrony may be due to other mechanisms. For example it may result from both anatomical disconnections among different cortical regions and reduced cholinergic coupling interactions between cortical neurons (Jeong *et al.*, 2001). In particular, a common hypothesis, proposed about three decades ago, is that basal forebrain neurons may be severely affected in AD and result in a cerebral cholinergic deficit, underlying memory loss and other cognitive symptoms (Jeong *et al.*, 2001). In other words, our report supports the view that the cognitive disturbances associated with AD may not solely be due to the loss of neurons, but also due to impairments in the temporal coordination of distributed neuronal activity (Uhlhaas & Singer, 2006).

Our report also adds an interesting new element to these theories. The main effect observed in our study is a significantly higher degree of local asynchronous activity (quantified by ρ), more specifically, a high number of non-coincident, asynchronous oscillatory events ($p = 0.00012$; uncorrected, significant after post-correction). Interestingly, we did not observe a significant effect on the timing jitter variance s_t of the coincident, synchronous events. As a consequence, our results seem to indicate that there is significantly more non-coincident background activity, while the *coincident* activity remains well synchronized. On the one hand, this observation is in agreement with previous studies that report a general decrease of neural synchrony in MCI and AD patients; on the other hand, it goes beyond previous results, since it yields a more subtle description of EEG synchrony in MCI and AD patients. It shows that the loss of coherence is mostly due to an increase of (local) non-coincident background activity, whereas the locked (coincident) activity remains equally well synchronized. The interpretation of this effect, in regard to neurobiological theories of AD, needs further analysis, and will be the object of our future reports.

Interestingly, the observed MCI-induced increase in ρ allows us to discriminate MCI patients from age-matched control subjects, especially in conjunction with the observed decrease of ffDTF (cf. Table 4). The obtained classification rates are probably robust to overfitting, since (i) they were computed through leave-one-out crossvalidation; (ii) only two features are used; (iii) the classifiers are simple and depend only on a few parameters. It is hard to compare this result to existing studies: In the literature, one most often reports results obtained on the training set, without any crossvalidation; those results are most likely over-optimistic. Moreover, the few reported results obtained through crossvalidation are typically based on five or more features and use more complicated classifiers; those classification results are typically in the range of 80–90%, but they are probably more prone to overfitting than our results.

At last, we would like to make a critical remark regarding EEG synchrony. It is important to realize that EEG synchrony values may be significantly affected by brain events other than changes of synchrony, and by choices (like the reference electrode) that necessarily have to be made during the analysis. The scalp EEG signals that are being analyzed depend on the reference, and the synchrony of two signals may thus equally depend on the choice of reference. In this particular study, we compare the EEG synchrony of two different populations. We are not interested in the absolute values of EEG synchrony, but in differences between the two populations instead. Since we used the same reference electrodes for all subjects, it is less probable that difference in EEG synchrony are related to the choice of the reference electrode.

Furthermore, as a single active source in the brain may affect the EEG signals across the entire scalp, changes in synchrony, and especially simultaneity of some events across channels, may be observed when the activity of one source alone changes, which is remote from a change in neural synchrony. Therefore, the effects we observed in MCI patients concerning EEG synchrony may perhaps be described and explained in terms of the activity and distribution of equivalent brain sources. It might be that in MCI subjects, there are generally fewer active (equivalent) brain sources, or perhaps the activity level of those sources is lower, which indirectly can be observed as loss of EEG synchrony.

Obviously, the problem that measures of EEG synchrony and connectivity are often confounded by effects other than synchrony of brain activity is a general one, and solving this problem clearly goes beyond the scope of this paper. Here, we merely wish to underline that those issues need to be carefully taken into account when formulating and verifying theories that try to explain the loss in EEG synchrony in MCI patients. As a result, relating those perturbations in EEG synchrony to biophysical processes, which is a crucial step towards a deeper understanding of MCI and the development of potential therapies, may prove to be more challenging than commonly believed.

Acknowledgments

We wish to thank Dr. T. Koenig (University Hospital of Clinical Psychiatry, Berne) for providing us an implementation of GFS. J.D. is deeply indebted to Shun-ichi Amari and Andi Loeliger for continuing support and encouragement over the last years.

A Background information on SES

We provide here more information on the SES method outlined in Section 2.10.

We successively apply the following transformations to the EEG signals:

- (1) wavelet transform,
- (2) normalization of the wavelet coefficients,
- (3) bump modeling of the normalized wavelet representation,
- (4) aggregation of the resulting bump models in several regions,
- (5) computation of the SES parameters for each pair of aggregated bump models.

In the following, we elaborate on each of those five operations.

A.1 Wavelet Transform

In order to extract the oscillatory patterns in the EEG, we apply a wavelet transform. More specifically, we use the complex Morlet wavelets (Goupillaud *et al.*, 1984; Delprat *et al.*, 1992):

$$\psi(t) = A \exp\left(-t^2/2\sigma_0^2\right) \exp(2i\pi f_0 t), \quad (\text{A.1})$$

where t is time, f_0 is frequency, σ_0 is a (positive) real parameter, and A is a (positive) normalization factor. The Morlet wavelet (A.1) has proven to be well suited for the time-frequency analysis of EEG (Tallon-Baudry *et al.*, 1996; Herrmann *et al.*, 2005). The product $w_0 = 2\pi f_0 \cdot \sigma_0$ determines the number of periods in the wavelet (“wavenumber”). This number should be sufficiently large (≥ 5), otherwise the wavelet $\psi(t)$ does not fulfill the admissibility condition:

$$\int \frac{|\psi(t)|^2}{t} dt < \infty, \quad (\text{A.2})$$

and as a result, the temporal localization of the wavelet becomes unsatisfactory (Goupillaud *et al.*, 1984; Delprat *et al.*, 1992). In the present study, we choose a wavenumber $w_0 = 7$, as in earlier studies (Tallon-Baudry *et al.*, 1996; Vialatte *et al.*, 2007); this choice yields good temporal resolution in the frequency range considered in this study.

The wavelet transform $x(t, s)$ of an EEG signal $x(t)$ is obtained as:

$$x(t, s) = \sum_{t'=1}^K x(t') \psi^*\left(\frac{t' - t}{s}\right), \quad (\text{A.3})$$

where $\psi(t)$ is the Morlet “mother” wavelet (A.1), s is a scaling factor, and $K = f_s T$, with f_s the sampling frequency and T the length of the signal. For the EEG data at hand, we have $T = 20s$ and $f_s = 200\text{Hz}$ and hence $K = 4000$. The scaled and shifted “daughter” wavelet in (A.3) has center frequency $f = f_0/s$. In the following, we will use the notation $x(t, f)$ instead of $x(t, s)$.

Next we compute the squared magnitude $s(t, f)$ of the coefficients $x(t, f)$:

$$s(t, f) = |x(t, f)|^2. \quad (\text{A.4})$$

Intuitively speaking, the time-frequency coefficients $s(t, f)$ represents the energy of oscillatory components with frequency f at time instances t . It is noteworthy that $s(t, f)$ contains no information about the phase of that component.

It is well known that EEG signals have very non-flat spectra with an overall $1/f$ shape, besides state-dependent peaks at specific frequencies. Therefore, the map $s(t, f)$ contains most energy at low frequencies f . If we directly apply bump modeling to the map $s(t, f)$, most bumps would be located in the low-frequency range, in other words, the high-frequency range would be under-represented. Since relevant information might be contained at high frequency, we normalize the map $s(t, f)$ before extracting the bump models.

A.2 Normalization

The normalized coefficients $s(t, f)$ are obtained as:

$$\tilde{z}(t, f) = \frac{s(t, f) - m_s(f)}{\sigma_s(f)}, \quad (\text{A.5})$$

where $m_s(f)$ is obtained by averaging $s(t, f)$ over the whole length of the EEG signal:

$$m_s(f) = \frac{1}{K} \sum_{t=1}^K s(t, f). \quad (\text{A.6})$$

Likewise, $\sigma_s^2(f)$ is the variance of $s(t, f)$:

$$\sigma_s^2(f) = \frac{1}{K} \sum_{t=1}^K (s(t, f) - m_s(f))^2. \quad (\text{A.7})$$

In words: the normalized coefficients $\tilde{z}(t, f)$ encode fluctuations from the baseline EEG power at time t and frequency f . The normalization (A.5) is known as z-score (see, e.g., (Buzsáki, 2006)), and is commonly applied (Matthew *et*

al., 2002; Vialatte *et al.*, 2007). The coefficients $\tilde{z}(t, f)$ are positive when the activity at t and f is stronger than the baseline $m_s(f)$ and negative otherwise.

There are various approaches to apply bump modeling to the z-score $\tilde{z}(t, f)$. One may first set the negative coefficients to zero, and next apply bump modeling. The bump models in that case represent peak activity. Alternatively, one may first set the positive coefficients equal to zero, reverse the sign of the negative coefficients, and then apply bump modeling. In that case, the bump models represent dips in the energy maps $s(t, f)$.

In the present application of diagnosing MCI, we will follow yet another approach. In order to extract bump models, we wish to exploit as much information as possible from the \tilde{z} maps. Therefore we will set only a small fraction of the coefficients $\tilde{z}(t, f)$ equal to zero, i.e., the 1% smallest coefficients. This approach was also followed in (Vialatte *et al.*, 2007), and is equivalent to the following transformation: we shift the coefficients (A.5) in the positive direction by adding a constant α , the remaining negative coefficients are set to zero:

$$z(t, f) = \left[\tilde{z}(t, f) + \alpha \right]^+ = \left[\frac{s(t, f) - m_s(f)}{\sigma_s(f)} + \alpha \right]^+, \quad (\text{A.8})$$

where $[x]^+ = x$ if $x \geq 0$ and $[x]^+ = 0$ otherwise. The constant α is chosen such that only 1% of the coefficients remains negative after addition with α ; this corresponds to $\alpha = 3.5$ in the present application. (In the study of (Vialatte *et al.*, 2007), it corresponds to $\alpha = 2$.) The top row of Fig. 2 shows the normalized wavelet map z (A.8) of two EEG signals.

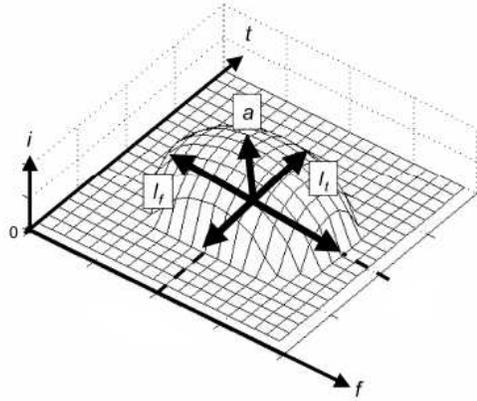
A.3 Bump Modeling

Next, bump models are extracted from the coefficient maps z (see Fig. 2 and (Vialatte *et al.*, 2007, 2009b)). We approximate the map $z(t, f)$ as a sum $z_{\text{bump}}(t, f, \theta)$ of a “small” number of smooth basis functions or “bumps” (denoted by f_{bump}):

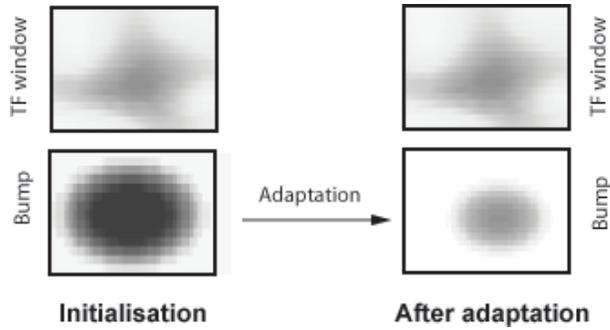
$$z(t, f) \approx z_{\text{bump}}(t, f, \theta) = \sum_{k=1}^{N_b} f_{\text{bump}}(t, f, \theta_k), \quad (\text{A.9})$$

where θ_k are vectors of bump parameters and $\theta = (\theta_1, \theta_2, \dots, \theta_{N_b})$. The sparse bump approximation $z_{\text{bump}}(t, f, \theta)$ represents regions in the time-frequency plane where the EEG contains more power than the baseline; in other words, it captures the most significant oscillatory activities in the EEG signal.

We choose half-ellipsoid bumps since they are well suited for our purposes (Vialatte, 2005; Vialatte *et al.*, 2007) (see Fig. A.1). Since we wish to keep the number



(a) Bump parameters: time t and frequency f , width Δt and height Δf , and amplitude w .



(b) Learning the bump parameters by minimizing the quadratic cost function (A.10); Top (left and right): a given patch of the time-frequency map. Bottom left: initial bump; Bottom right: bump obtained after adaptation.

Fig. A.1. Half ellipsoid bump.

of bump parameters as low as possible, the principal axes of the half ellipsoid bumps are restricted to be parallel to the time-frequency axes. As a result, each bump is described by five parameters (see Fig. 1(a)): the coordinates of its center (i.e., time t_k and frequency f_k), its amplitude $w_k > 0$, and the extension Δt_k and Δf_k in time and frequency respectively, in other words, $\theta_k = (t_k, f_k, w_k, \Delta t_k, \Delta f_k)$. More precisely, the ellipsoid bump function

$f_{\text{bump}}(t, f, \theta_k)$ is defined as:

$$f_{\text{bump}}(t, f, \theta_k) = \begin{cases} w_k \sqrt{1 - \kappa(t, f, \theta_k)} & \text{for } 0 \leq \kappa(t, f, \theta_k) \leq 1 \\ 0 & \text{for } \kappa(t, f, \theta_k) > 1, \end{cases} \quad (\text{A.10})$$

where

$$\kappa(t, f, \theta_k) = \frac{(t - t_k)^2}{(\Delta t_k)^2} + \frac{(f - f_k)^2}{(\Delta f_k)^2}. \quad (\text{A.11})$$

For the EEG data described in Section 3.1, the number of bumps N_b (cf. (A.9)) is typically between 50 and 100, and therefore, $z_{\text{bump}}(t, f, \theta)$ is fully specified by a few hundred parameters. On the other hand, the time-frequency map $z(t, f)$ consists of between 10^4 and 10^5 coefficients; the bump model $z_{\text{bump}}(t, f, \theta)$ is thus a sparse (but approximate) representation of $z(t, f)$.

The bump model $z_{\text{bump}}(t, f, \theta)$ is extracted from $z(t, f)$ by the following algorithm (Vialatte, 2005; Vialatte *et al.*, 2007, 2009b):

- (1) Define appropriate boundaries for the map $z(t, f)$ in order to avoid finite-size effects.
- (2) Partition the map $z(t, f)$ into small zones. The size of these zones depends on the time-frequency ratio of the wavelets, and are optimized to model oscillatory activities lasting 4 to 5 oscillation periods. Larger oscillatory patterns are modeled by multiple bumps.
- (3) Find the zone \mathcal{Z} that contains the most energy.
- (4) Adapt a bump to that zone; the bump parameters are determined by minimizing the quadratic cost function (see Fig. 1(b)):

$$\mathcal{E}(\theta_k) = \sum_{t, f \in \mathcal{Z}} \left(z(t, f) - f_{\text{bump}}(t, f, \theta_k) \right)^2. \quad (\text{A.12})$$

Next withdraw the bump from the original map.

- (5) The fraction of total intensity contained in that bump is computed:

$$F = \frac{\sum_{t, f \in \mathcal{Z}} f_{\text{bump}}(t, f, \theta_k)}{\sum_{t, f \in \mathcal{Z}} z(t, f)}. \quad (\text{A.13})$$

If $F < G$ for three consecutive bumps (and hence those bumps contain only a small fraction of the energy of map $z(t, f)$), stop modeling and proceed to (6), otherwise iterate (3).

- (6) After all signals have been modeled, define a threshold $T \geq G$, and remove the bumps for which $F < T$. This allows us to trade off the information loss and modeling of background noise: when too few bumps are generated, information about the oscillatory activity of the brain is lost. On the other hand, if too many bumps are generated, the bump model also contains low-amplitude oscillatory components; since the measurement pro-

cess typically introduces a substantial amount of noise, it is likely that the low-amplitude oscillatory components do not stem from organized brain oscillations but are instead due measurement noise. By adjusting the threshold T , we try to find an appropriate number of bumps.

We used a threshold $G = 0.05$. With this threshold, each bump model contains many bumps, and some of those bumps may actually model background noise. Therefore, we further pruned the bump models (cf. Step 6); we consider values of T .

We refer to (Vialatte, 2005; Vialatte *et al.*, 2007, 2009b) for more information on bump modeling. In particular, we used the same choice of boundaries (Step 1) and partitions (Step 2) as in those references.

Eventually, we obtain 21 bump models, i.e., one per EEG channel.

A.4 Aggregation

As a next step, we group the 21 electrodes into 5 regions, as illustrated in Fig. 4. From the 21 bump models obtained by sparsification (cf. Section A.3), we extract a single bump model for each of the zones by means of the aggregation algorithm described in (Vialatte *et al.*, 2007, 2009b).

A.5 Stochastic Event Synchrony

We compute the SES parameters ρ , δ_t , δ_f , s_t and s_f , for all pairs of regions. In addition, in order to obtain measures for *average* synchrony, we average the SES parameters over all region pairs, resulting in one set of average SES parameters per subject.

We now briefly describe how we compute the SES parameters; we refer to (Dauwels *et al.*, 2007, 2009a,b) for more details. The alignment of the two time-frequency maps (cf. Fig. 3, bottom) is cast as a statistical inference problem. The associated probabilistic model depends on the SES parameters $\theta = (\delta_t, \delta_f, s_t, s_f)$ besides the following two kinds of latent variables: (i) binary variables $C_{kk'}$, associated to each pair of events, where $C_{kk'} = 1$ indicates that event E_k of the first time-frequency map is coincident with event $E'_{k'}$ in the other time-frequency map, and where $C_{kk'} = 0$ otherwise; (ii) binary variables B_k which indicate whether the events E_k is coincident or not; likewise, binary variables $B'_{k'}$ for the events $E'_{k'}$. Note that one can straightforwardly infer ρ from B and

B' . The latent-variable model is of the form:

$$\begin{aligned}
p(e, e', b, b', c, \theta) \propto & \\
& \prod_{k=1}^n (\beta \delta[b_k - 1] + \delta[b_k]) \prod_{k'=1}^{n'} (\beta \delta[b'_{k'} - 1] + \delta[b'_{k'}]) \\
& \cdot \prod_{k=1}^n \prod_{k'=1}^{n'} \left(\mathcal{N}(t'_{k'} - t_k; \delta_t, s_t) \mathcal{N}(f'_{k'} - f_k; \delta_f, s_f) \right)^{c_{kk'}} \\
& \cdot \prod_{k=1}^n \left(\delta[b_k + \sum_{k'=1}^{n'} c_{kk'} - 1] \right) \prod_{k'=1}^{n'} \left(\delta[b'_{k'} + \sum_{k=1}^n c_{kk'} - 1] \right) \\
& \cdot p(\delta_t) p(s_t) p(\delta_f) p(s_f), \tag{A.14}
\end{aligned}$$

where β is a constant that serves as a knob to control the number of non-coincident events, n and n' is the total number of events in the two time-frequency maps, and $\mathcal{N}(x; m, s)$ stands for a univariate Gaussian distribution with mean m and variance s (Dauwels *et al.*, 2007, 2009a,b). For convenience, we choose improper priors $p(\delta_t) = p(\delta_f) = p(s_t) = p(s_f) = 1$.

The SES parameters $\theta = (\delta_t, \delta_f, s_t, s_f)$ and the latent variables C, B and B' are determined jointly by maximum a posteriori estimation (MAP) estimation. This may practically be carried out by cyclic maximization (Dauwels *et al.*, 2007, 2009a,b): for fixed θ , one maximizes $\log p$ (cf. (A.14)) w.r.t. C, B and B' and vice versa; those two steps are iterated until convergence. Conditional maximization w.r.t. θ is straightforward, and the conditional maximization w.r.t. C, B and B' is equivalent to a well-known problem in combinatorial optimization, i.e., bipartite max-weight matching. We solve that problem by applying the max-product algorithm on a graphical model corresponding to the latent-variable probabilistic model (A.14) (see Fig. A.2) (Dauwels *et al.*, 2007, 2009a,b). The edges correspond to variables, the nodes corresponds to factors in (A.14). The nodes \mathcal{N} corresponds to the Gaussian distributions in (A.14), the nodes denoted by $\bar{\Sigma}$ represent the factors $(\delta[b_k + \sum_{k'=1}^{n'} c_{kk'} - 1])$ (blue) and $(\delta[b'_{k'} + \sum_{k=1}^n c_{kk'} - 1])$ (red), and the nodes denoted by β correspond to the factors $(\beta \delta[b_k - 1] + \delta[b_k])$ and $(\beta \delta[b'_{k'} - 1] + \delta[b'_{k'}])$. The arrows in Fig. A.2 depict “messages”, which are probabilistic information about which pairs of events are coincident and which not. The messages are iteratively computed at each node according to the max-product computation rules (Loeliger *et al.*, 2007). Intuitively, the nodes may be viewed as computing elements that iteratively update their opinion about which events match and which do not, based on the opinions (“messages”) they receive from neighboring nodes. When the algorithm eventually has converged and the nodes have found a “consensus”, the messages are combined to obtain a decision on C, B and B' , and an estimate of ρ and the other SES parameters. We refer to (Dauwels *et al.*, 2007, 2009a,b) for a detailed description of the algorithm.

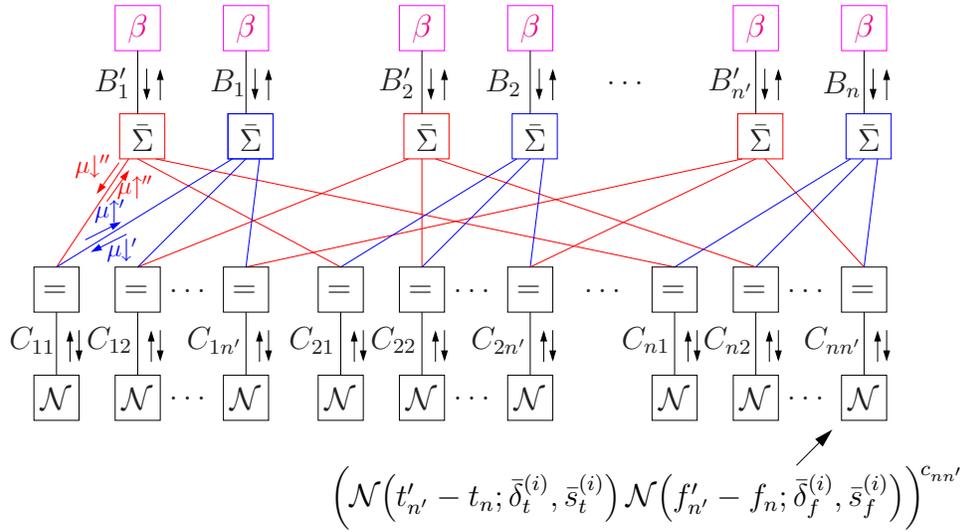


Fig. A.2. Max-product message passing; the arrows indicate the max-product messages, computed according to the max-product update rule.

B Parameter Settings for the Different Synchrony Measures

Many of the synchrony measures considered in this paper depend on one or more parameters, as pointed out in Section 3.2.5:

- We applied the stationary synchrony methods to EEG segments of length $L = 1\text{s}, 5\text{s},$ and 20s (see Section 3.2.4).
- For mutual information I (time domain) and the state space based measures $S^k, H^k, N^k,$ and $S_{\text{est}},$ the embedding dimension m and the time lag τ need to be chosen. In our experiments, we tested the values $m = 1, 2, \dots, 5;$ the time delay was each time set to $\tau = 1/30\text{s},$ which is the period of the fastest oscillations in the EEG signals at hand.
- In order to compute MVAR coherence and Granger causality measures, one needs to specify the order p of the MVAR model (10). Since it is not obvious which model order amounts to the best performance for diagnosing MCI, we have tried a range of model orders, i.e., $p = 1, 2, \dots, 10.$
- The Rényi and Jensen-Rényi divergence depend on the order parameter $\alpha.$ We computed the Rényi divergence with order parameter $\alpha = 0.1, 0.2, \dots, 1$ and the Jensen-Rényi divergence for $\alpha = 1, 2, \dots, 10.$
- We choose the parameters of the SES algorithm as follows. Since we are dealing with spontaneous EEG, it is unlikely that the EEG signals from certain channels are delayed w.r.t. other channels; moreover, systematic frequency offsets are unrealistic. Therefore, we choose the initialization $\hat{\delta}_t^{(0)} = 0 = \hat{\delta}_f^{(0)}$ (in Local Approach 1 and 2). In Local Approach 1 we investigated the parameter settings $\hat{\sigma}_t^{(0)} = \sqrt{\hat{s}_t^{(0)}} = 0.1, 0.15, 0.2, 0.25, \hat{\sigma}_f^{(0)} = \sqrt{\hat{s}_f^{(0)}} = 0.05, 0.1, 0.15, 0.2, \beta = 0.01, 0.001,$ and $T = 0.74, 0.76, \dots, 0.88;$ in Local Approach 2 we considered the parameter settings $\hat{\sigma}_t^{(0)} = \sqrt{\hat{s}_t^{(0)}} = 0.15, 0.175, \dots, 0.25, \hat{\sigma}_f^{(0)} = \sqrt{\hat{s}_f^{(0)}} = 0.025, 0.050, \dots, 0.15, \beta = 0.01, 0.001,$ and $T = 0.2, 0.21, \dots, 0.25.$

We list the parameter settings that yield the smallest (uncorrected) p-values. We first consider Local and Global Approach 1:

- Table B.1 to B.3 list the values of L that amounted to the strongest difference in (average) EEG synchrony between MCI patients and control subjects, referred to as “optimal” length $L.$ In particular, Table B.1 concerns the state space based measures $S^k, H^k, N^k,$ and $S_{\text{est}},$ and time-domain mutual information $I,$ whereas Table B.2 contains the optimal segment length L for the Granger causality measures. The optimal L for correlation coefficient, corr-entropy, magnitude and phase coherence, and the phase synchrony indices EMA, IPA, and GFS can be found in Table B.3.
- Table B.1 contains the optimal embedding dimension m for the state space based measures $S^k, H^k, N^k,$ and $S_{\text{est}},$ and time-domain mutual informa-

tion I .

- Table B.2 lists the optimal order p for the Granger causality measures. Note that not all Granger measures necessarily yield the smallest p-values with the *same* MVAR model. For example, we obtained the best results for DTF with a model of order 9, whereas for PDC the optimal order is 2.
- The p-values for the Rényi divergence did not seem to depend much on α ; in the case of Jensen-Rényi divergence, we obtained the smallest p-values for $\alpha = 2$.
- The optimal parameter settings for SES can be found in Table B.4.

In Local and Global Approach 1 the measures are applied to the 5 average EEG signals, as we pointed out earlier. We now consider Local Approach 2:

- Table B.5 lists the optimal parameters for the S , N , and H indices.
- Table B.6 contains the optimal length L for the remaining local measures (except time-domain mutual information and SES).
- The optimal parameter settings for SES can be found in Table B.7.
- The p-values for the Rényi divergence did not seem to depend much on α , the smallest p-value occur for $\alpha = 0.1$. In the case of Jensen-Rényi divergence, we obtained the smallest p-values for $\alpha = 1$.

At last we treat Global Approach 2 and 3:

- Table B.8 lists the optimal length L for Omega complexity and GFS.
- Table B.9 contains the optimal parameters for the S-estimator.

Measure	length L	embedding dimension m
S^k	20s	5
H^k	20s	1
N^k	20s	1
S_{est}	20s	1
I	20s	1

Table B.1

Optimal segment length L and embedding dimension m for the state space measures and mutual information I , applied to the 5 average EEG signals (Local and Global Approach 1).

C Dependency of Uncorrected p-values on Parameter Settings

We investigate how the (uncorrected) p-values depend on the parameters of the synchrony methods. This analysis will provide us some intuition about which

Granger measure	length L	model order p
MVAR Coherence	1s	8
Partial Coherence	1s	4
DTF	1s	9
ffDTF	1s	6
PDC	1s	2
dDTF	5s	6

Table B.2

Optimal segment length L and model order p for MVAR coherence and the Granger causality measures, applied to the 5 average EEG signals (Global Approach 1).

Measure	length L	Measure	length L
Correlation	20s	Corr-entropy	20s
Coherence	5s	Phase Coherence	5s
Evolution Map	20s	Instantaneous Map	1s
Hilbert Phase	5s	GFS	20s
Omega complexity	20s		

Table B.3

Optimal segment length L for various synchrony measures, applied to the 5 average EEG signals (Local and Global Approach 1).

Measure	β	$\hat{\sigma}_t^{(0)}$	$\hat{\sigma}_f^{(0)}$	T
s_t	0.001	0.2	0.02	0.86
ρ	0.001	0.1	0.02	0.76

Table B.4

Optimal parameters for SES (Local Approach 1); we list optimal parameter settings for s_t and ρ separately. We use the notation $\hat{\sigma}_t^{(0)} = \sqrt{\hat{s}_t^{(0)}}$ and $\hat{\sigma}_f^{(0)} = \sqrt{\hat{s}_f^{(0)}}$.

effects are statistically significant. Indeed, if the p-values are highly dependent on the specific choice of parameters, the effect is not really robust and therefore most probably not statistically significant. We will restrict ourselves to the synchrony measures whose smallest uncorrected p-values are below 0.05.

Table C.1 shows how the (uncorrected) p-values of correlation coefficient, (magnitude) coherence, phase coherence, corr-entropy, and Instantaneous Pe-

Measure	length L	embedding dimension m
S^k	20s	5
H^k	1s	1
N^k	1s	1

Table B.5

Optimal segment length L and embedding dimension m for the S , H , N indices, applied according to Local Approach 2.

Measure	length L	Measure	length L
Correlation	20s	Corr-entropy	20s
Coherence	1s	Phase Coherence	5s
Evolution Map	20s	Instantaneous Map	20s
Hilbert Phase	20s		

Table B.6

Optimal segment length L for various local synchrony measures, applied according to Local Approach 2.

Measure	β	$\hat{\sigma}_t^{(0)}$	$\hat{\sigma}_f^{(0)}$	T
s_t	0.001	0.175	0.025	0.22
ρ	0.001	0.225	0.050	0.22

Table B.7

Optimal parameters for SES (Local Approach 2); we list optimal parameter settings for s_t and ρ separately. We use the notation $\hat{\sigma}_t^{(0)} = \sqrt{\hat{s}_t^{(0)}}$ and $\hat{\sigma}_f^{(0)} = \sqrt{\hat{s}_f^{(0)}}$.

	GFS	Omega complexity
21 EEG signals	1s	20s
pairs of zones	20s	20s

Table B.8

Optimal segment length L for GFS and Omega complexity, applied (i) to all 21 EEG signals (Global Approach 2); (ii) to all pairs of zones (Global Approach 3).

riod Approach (IPA) depend on the length L of the EEG segments. Table C.2 and C.3 show how the (uncorrected) p-values of dDTF and fDTF respectively depend on L and the model order p . Table C.4, C.5, C.6, and C.7 show how the (uncorrected) p-values of N^k , S^k , the S-estimator, and GFS respectively

	length L	embedding dimension m
21 EEG signals	20s	1
pairs of zones	1s	4

Table B.9

Optimal segment length L and embedding dimension m for S-estimator (Global Approach 2 and 3).

depend on L and the embedding dimension m . Table C.9 to C.11 show how the (uncorrected) p-values of the SES parameters s_t and ρ depend on $\hat{s}_t^{(0)}$ and $\hat{s}_f^{(0)}$, β and T .

The (uncorrected) p-values of correlation coefficient are robust w.r.t. segment length L and computational approach (cf. Table C.1). This is less the case for corr-entropy, and not at all the case for (magnitude) coherence, phase coherence, and IPA (cf. Table C.1). The (uncorrected) p-values of dDTF are only below 0.05 for two specific choices of the model order p and segment length L (cf. Table C.2); the situation is similar for S-estimator and GFS (cf. Table C.6 and Table C.7). Also the measure N^k does not seem to yield small (uncorrected) p-values for most parameter settings (cf. Table C.4): only for $L = 5s$ and Local Approach 1 the (uncorrected) p-values are robust w.r.t. embedding dimension m . The (uncorrected) p-values of the measure S^k are even more sensitive to the parameters L and m . Table C.3 shows that for most parameter settings, the (uncorrected) p-values of fDTF are smaller than 0.005. Higher-order models seem to yield the best results. This is in agreement with our expectations: EEG signals have strong low-frequency components, and therefore, present values of the EEG signals depend on values from the past. Such long-term correlations can only be captured by higher-order MVAR models ($p > 1$), not by first-order MVAR models.

Under Local Approach 1 (cf. Table C.9), the (uncorrected) p-values for s_t and ρ are smaller than 0.05 for some values of $\hat{s}_t^{(0)}$, $\hat{s}_f^{(0)}$, β and T . However, the results for ρ are clearly more robust under Local Approach 2 (cf. Table C.11): The (uncorrected) p-values for ρ are smaller than 0.05 for almost all considered values of $\hat{s}_t^{(0)}$, $\hat{s}_f^{(0)}$, and β and for $T = 0.2, 0.21, 0.22$, and 0.23 ; for $T = 0.22$, the (uncorrected) p-values are all below 0.005. Note that for $T = 0.2$, practically the whole time-frequency map is filled with bumps, whereas for $T = 0.24$ and 0.25 , there are few bumps. We therefore a priori expect small p-values to occur for $0.21 < T < 0.24$, and it is not surprising that an optimum is achieved in terms of p-values at $T = 0.23$; there are sufficient bumps to reliably the SES parameters, yet at the same time, the bumps do not fill the entire time-frequency map. It is also noteworthy that Local Approach 2 is the most appropriate for SES: local averaging (Local Approach 1) typically reduces the high-frequency components in the EEG signals, and therefore, it

Local Approach 1	$L = 1s$	$L = 5s$	$L = 20s$	Local Approach 2	$L = 1s$	$L = 5s$	$L = 20s$
Correlation	0.028*	0.026*	0.025*	Correlation	0.025*	0.024*	0.0178*
Coherence	0.1949	0.0288*	NA	Coherence	0.062	0.0960	NA
Phase Coherence	0.5344	0.041*	NA	Phase Coherence	0.72	0.61	NA
IPA	0.7299	0.7415	0.8720	IPA	0.1795	0.1272	0.0201*
Corr-entropy	0.26	0.069	0.0324*	Corr-entropy	0.1795	0.0437*	0.0363*

Table C.1

Dependency of (uncorrected) p-values (Mann-Whitney test) on segment length L , following Local Approach 1 and 2: * indicates $p < 0.05$.

Model order p	$L = 1s$	$L = 5s$	$L = 20s$
1	0.4295	0.3150	0.1347
2	0.7766	0.9084	0.7185
3	0.3225	0.5651	0.7648
4	0.2112	0.2112	0.3534
5	0.1056	0.0668	0.1308
6	0.1056	0.0288*	0.1269
7	0.1308	0.0691	0.2595
8	0.1232	0.0715	0.2226
9	0.0991	0.0668	0.2057
10	0.1232	0.0421*	0.1845

Table C.2

Dependency of (uncorrected) p-values (Mann-Whitney test) for dDTF on segment length L and model order p , following Global Approach 1; * indicates $p < 0.05$.

reduces the number of high-frequency bumps on the time-frequency plane. In other words, the bump models obtained from locally averaged EEG may not always be a reliable representation of the time-frequency components. This seems to be in agreement with the results of Table C.11.

Model order p	$L = 1s$	$L = 5s$	$L = 20s$
1	0.4569	0.5861	0.3300
2	0.0582	0.2057	0.2864
3	0.0042**	0.0377*	0.0960
4	0.0024**	0.0201*	0.0542
5	0.0013**	0.0377*	0.0138*
6	0.0012**	0.0026**	0.0093*
7	0.0027**	0.0024**	0.0078*
8	0.0020**	0.0015**	0.0024**
9	0.0037**	0.0012**	0.0037**
10	0.0049**	0.0015**	0.0037**

Table C.3
Dependency of (uncorrected) p-values (Mann-Whitney test) for fDFTF on segment length L and model order p , following Global Approach 1; * and ** indicate $p < 0.05$ and $p < 0.005$ respectively.

Local Approach 1	$L = 1s$	$L = 5s$	$L = 20s$	Local Approach 2	$L = 1s$	$L = 5s$	$L = 20s$
$m = 1$	0.0324*	0.0324*	0.0288*	$m = 1$	0.0542	0.0691	0.0582
$m = 2$	0.0377*	0.0764	0.0645	$m = 2$	0.0764	0.1308	0.0991
$m = 3$	0.0336*	0.0872	0.0602	$m = 3$	0.0645	0.1515	0.1089
$m = 4$	0.0363*	0.1023	0.0790	$m = 4$	0.0645	0.2057	0.1651
$m = 5$	0.0383*	0.1159	0.0960	$m = 5$	0.0960	0.2661	0.2002

Table C.4
Dependency of (uncorrected) p-values (Mann-Whitney test) for N^k on segment length L and embedding dimension m , following Local Approach 1 and 2; * indicates $p < 0.05$.

Local Approach 1	$L = 1s$	$L = 5s$	$L = 20s$	Local Approach 2	$L = 1s$	$L = 5s$	$L = 20s$
$m = 1$	0.2864	0.0960	0.0715	$m = 1$	0.7532	0.1651	0.1308
$m = 2$	0.6181	0.2531	0.1347	$m = 2$	0.9816	0.6509	0.4477
$m = 3$	0.3377	0.1472	0.1023	$m = 3$	0.7300	0.3778	0.2468
$m = 4$	0.1698	0.0668	0.0602	$m = 4$	0.3077	0.1195	0.0764
$m = 5$	0.2737	0.0562	0.0453*	$m = 5$	0.2002	0.0487*	0.0193*

Table C.5

Dependency of (uncorrected) p-values (Mann-Whitney test) for S^k on segment length L and embedding dimension m , following Local Approach 1 and 2; * indicates $p < 0.05$.

m	$L = 1s$	$L = 5s$	$L = 20s$
1	0.0691	0.0505	0.0421*
2	0.4662	0.3077	0.2795
3	0.9939	0.8963	0.8002
4	0.8359	0.9816	0.9328
5	0.8720	0.9206	0.9450

Table C.6

Dependency of (uncorrected) p-values (Mann-Whitney test) for S-estimator on segment length L and embedding dimension m , following Global Approach 1; * indicates $p < 0.05$.

	$L = 1s$	$L = 5s$	$L = 20s$
1	0.51	0.031*	

Table C.7

Dependency of (uncorrected) p-values (Mann-Whitney test) for GFS on segment length L , following Global Approach 1; * indicates $p < 0.05$.

$T = 0.74, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.74, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.9694	0.4948	0.8841	0.9084	$\hat{\sigma}_f^{(0)} = 0.05$	0.3077	0.5446	0.9694	0.3300
$\hat{\sigma}_f^{(0)} = 0.1$	0.4385	0.6509	0.5344	0.8121	$\hat{\sigma}_f^{(0)} = 0.1$	0.0668	0.3534	0.3005	0.6181
$\hat{\sigma}_f^{(0)} = 0.15$	0.3377	0.0790	0.6509	0.7415	$\hat{\sigma}_f^{(0)} = 0.15$	0.5446	0.1698	0.5144	0.1698
$\hat{\sigma}_f^{(0)} = 0.2$	0.2595	0.0715	0.2112	0.2226	$\hat{\sigma}_f^{(0)} = 0.02$	0.4662	0.3005	0.3614	0.3300
$T = 0.76, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.76, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.4852	0.4477	0.5651	0.5651	$\hat{\sigma}_f^{(0)} = 0.05$	0.4477	0.8963	0.5966	0.2285
$\hat{\sigma}_f^{(0)} = 0.1$	0.7415	0.6731	0.3946	0.6289	$\hat{\sigma}_f^{(0)} = 0.1$	0.3150	0.3077	0.4477	0.5548
$\hat{\sigma}_f^{(0)} = 0.15$	0.4206	0.2345	0.6509	0.3077	$\hat{\sigma}_f^{(0)} = 0.15$	0.6181	0.2727	0.7300	0.3377
$\hat{\sigma}_f^{(0)} = 0.2$	0.3077	0.1698	0.2169	0.4295	$\hat{\sigma}_f^{(0)} = 0.02$	0.5144	0.8963	0.3695	0.4385
$T = 0.78, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.78, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.8002	0.2406	0.2285	0.2169	$\hat{\sigma}_f^{(0)} = 0.05$	0.1559	0.4206	0.8479	0.1559
$\hat{\sigma}_f^{(0)} = 0.1$	0.5966	0.3534	0.3005	0.5651	$\hat{\sigma}_f^{(0)} = 0.1$	0.2169	0.0691	0.0562	0.1897
$\hat{\sigma}_f^{(0)} = 0.15$	0.1269	0.0960	0.0991	0.1795	$\hat{\sigma}_f^{(0)} = 0.15$	0.2531	0.0715	0.1795	0.0930
$\hat{\sigma}_f^{(0)} = 0.2$	0.2002	0.0790	0.0624	0.1089	$\hat{\sigma}_f^{(0)} = 0.02$	0.2112	0.1949	0.1795	0.2112
$T = 0.8, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.8, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.4662	0.5144	0.3005	0.3695	$\hat{\sigma}_f^{(0)} = 0.05$	0.8963	0.5446	0.7766	0.5046
$\hat{\sigma}_f^{(0)} = 0.1$	0.9206	0.6399	0.5548	0.7070	$\hat{\sigma}_f^{(0)} = 0.1$	0.7070	0.1124	0.0668	0.2345
$\hat{\sigma}_f^{(0)} = 0.15$	0.5046	0.2285	0.0739	0.1269	$\hat{\sigma}_f^{(0)} = 0.15$	0.8359	0.1746	0.1651	0.1124
$\hat{\sigma}_f^{(0)} = 0.2$	0.6181	0.1429	0.0453*	0.0739	$\hat{\sigma}_f^{(0)} = 0.02$	0.7766	0.3377	0.1845	0.1559
$T = 0.82, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.82, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.9328	0.4206	0.2864	0.5755	$\hat{\sigma}_f^{(0)} = 0.05$	0.7532	0.3946	0.5966	0.8002
$\hat{\sigma}_f^{(0)} = 0.1$	0.6181	0.5244	0.2727	0.5548	$\hat{\sigma}_f^{(0)} = 0.1$	0.1897	0.1159	0.0453*	0.0523
$\hat{\sigma}_f^{(0)} = 0.15$	0.4569	0.1159	0.0324	0.0437*	$\hat{\sigma}_f^{(0)} = 0.15$	0.3946	0.1056	0.0237*	0.0349*
$\hat{\sigma}_f^{(0)} = 0.2$	0.5861	0.1023	0.0624	0.0901	$\hat{\sigma}_f^{(0)} = 0.02$	0.5344	0.1845	0.1347	0.0691
$T = 0.84, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.84, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.6181	0.8963	0.4295	0.5548	$\hat{\sigma}_f^{(0)} = 0.05$	0.8240	0.2112	0.1472	0.1604
$\hat{\sigma}_f^{(0)} = 0.1$	0.9816	0.4477	0.1698	0.1845	$\hat{\sigma}_f^{(0)} = 0.1$	0.5548	0.1845	0.1023	0.2406
$\hat{\sigma}_f^{(0)} = 0.15$	0.9084	0.4118	0.0844	0.1559	$\hat{\sigma}_f^{(0)} = 0.15$	0.4477	0.1651	0.0505	0.0764
$\hat{\sigma}_f^{(0)} = 0.2$	0.9816	0.1949	0.0256*	0.0349*	$\hat{\sigma}_f^{(0)} = 0.02$	0.5651	0.1795	0.0377*	0.0437*
$T = 0.86, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.86, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.2661	0.9328	0.8121	0.8121	$\hat{\sigma}_f^{(0)} = 0.05$	0.8479	0.3861	0.3946	0.1845
$\hat{\sigma}_f^{(0)} = 0.1$	0.7532	0.9450	0.3861	0.4031	$\hat{\sigma}_f^{(0)} = 0.1$	0.4118	0.0602	0.1124	0.2285
$\hat{\sigma}_f^{(0)} = 0.15$	0.9328	0.3225	0.0764	0.1388	$\hat{\sigma}_f^{(0)} = 0.15$	0.2864	0.0311*	0.0406*	0.0645
$\hat{\sigma}_f^{(0)} = 0.2$	0.6399	0.0790	0.0210*	0.0277*	$\hat{\sigma}_f^{(0)} = 0.02$	0.1429	0.0930	0.0111*	0.0300*
$T = 0.88, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.88, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.6399	0.5651	0.5446	0.2934	$\hat{\sigma}_f^{(0)} = 0.05$	0.7300	0.6181	0.3150	0.5046
$\hat{\sigma}_f^{(0)} = 0.1$	0.8720	0.7884	0.6289	0.4295	$\hat{\sigma}_f^{(0)} = 0.1$	0.3946	0.1195	0.0960	0.1056
$\hat{\sigma}_f^{(0)} = 0.15$	0.8963	0.3225	0.0715	0.1195	$\hat{\sigma}_f^{(0)} = 0.15$	0.2468	0.1559	0.0790	0.0437*
$\hat{\sigma}_f^{(0)} = 0.2$	0.3534	0.3005	0.1195	0.0739	$\hat{\sigma}_f^{(0)} = 0.02$	0.2531	0.1089	0.0453*	0.0470*

Table C.8

Dependency of (uncorrected) p-values (Mann-Whitney test) for s_t on $\hat{\sigma}_t^{(0)} = \sqrt{\hat{s}_t^{(0)}}$, $\hat{\sigma}_f^{(0)} = \sqrt{\hat{s}_f^{(0)}}$, β and T , following Local Approach 1; * indicates $p < 0.05$.

$T = 0.74, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.74, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.7648	0.2057	0.1472	0.1195	$\hat{\sigma}_f^{(0)} = 0.05$	0.0960	0.3614	0.2406	0.6620
$\hat{\sigma}_f^{(0)} = 0.1$	0.5144	0.0668	0.0901	0.0817	$\hat{\sigma}_f^{(0)} = 0.1$	0.2531	0.1472	0.2795	0.2864
$\hat{\sigma}_f^{(0)} = 0.15$	0.2661	0.1232	0.1651	0.1056	$\hat{\sigma}_f^{(0)} = 0.15$	0.0872	0.3778	0.0930	0.4385
$\hat{\sigma}_f^{(0)} = 0.2$	0.4385	0.1949	0.1089	0.1269	$\hat{\sigma}_f^{(0)} = 0.02$	0.0377*	0.1056	0.0505	0.4852
$T = 0.76, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.76, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.1089	0.0542	0.1308	0.0960	$\hat{\sigma}_f^{(0)} = 0.05$	0.0844	0.2169	0.1559	0.3377
$\hat{\sigma}_f^{(0)} = 0.1$	0.0645	0.0132*	0.2057	0.0715	$\hat{\sigma}_f^{(0)} = 0.1$	0.1429	0.1604	0.1269	0.1698
$\hat{\sigma}_f^{(0)} = 0.15$	0.1023	0.0764	0.1124	0.1232	$\hat{\sigma}_f^{(0)} = 0.15$	0.0157*	0.0764	0.0363*	0.1124
$\hat{\sigma}_f^{(0)} = 0.2$	0.1195	0.1269	0.0817	0.0790	$\hat{\sigma}_f^{(0)} = 0.02$	0.0044*	0.0336*	0.0288*	0.0901
$T = 0.78, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.78, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.5651	0.3455	0.3861	0.3695	$\hat{\sigma}_f^{(0)} = 0.05$	0.3150	0.4948	0.2795	0.8002
$\hat{\sigma}_f^{(0)} = 0.1$	0.2169	0.1746	0.5046	0.2864	$\hat{\sigma}_f^{(0)} = 0.1$	0.4569	0.4385	0.4757	0.4295
$\hat{\sigma}_f^{(0)} = 0.15$	0.6620	0.3946	0.2226	0.2595	$\hat{\sigma}_f^{(0)} = 0.15$	0.1515	0.6181	0.3455	0.6844
$\hat{\sigma}_f^{(0)} = 0.2$	0.4295	0.5344	0.4757	0.3534	$\hat{\sigma}_f^{(0)} = 0.02$	0.3300	0.2795	0.4948	0.3946
$T = 0.8, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.8, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.4031	0.0406*	0.1056	0.0715	$\hat{\sigma}_f^{(0)} = 0.05$	0.0336*	0.1023	0.0377*	0.0715
$\hat{\sigma}_f^{(0)} = 0.1$	0.1429	0.0470*	0.0668	0.0277*	$\hat{\sigma}_f^{(0)} = 0.1$	0.0739	0.0582	0.0266*	0.0201*
$\hat{\sigma}_f^{(0)} = 0.15$	0.2468	0.0391*	0.0311*	0.0138*	$\hat{\sigma}_f^{(0)} = 0.15$	0.0256*	0.0300*	0.0081*	0.0164*
$\hat{\sigma}_f^{(0)} = 0.2$	0.0624	0.0377*	0.0391*	0.0218*	$\hat{\sigma}_f^{(0)} = 0.02$	0.0157*	0.0324*	0.0047**	0.0106*
$T = 0.82, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.82, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.6399	0.0437*	0.2661	0.1651	$\hat{\sigma}_f^{(0)} = 0.05$	0.0377*	0.1845	0.1698	0.1949
$\hat{\sigma}_f^{(0)} = 0.1$	0.3150	0.0844	0.1515	0.1388	$\hat{\sigma}_f^{(0)} = 0.1$	0.0739	0.0901	0.1897	0.1388
$\hat{\sigma}_f^{(0)} = 0.15$	0.3946	0.1124	0.0960	0.1949	$\hat{\sigma}_f^{(0)} = 0.15$	0.0453*	0.1308	0.1232	0.0991
$\hat{\sigma}_f^{(0)} = 0.2$	0.4569	0.1845	0.1056	0.0764	$\hat{\sigma}_f^{(0)} = 0.02$	0.0237*	0.0562	0.0193*	0.0421*
$T = 0.84, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.84, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.3225	0.0406*	0.2345	0.2169	$\hat{\sigma}_f^{(0)} = 0.05$	0.0764	0.2795	0.3300	0.3225
$\hat{\sigma}_f^{(0)} = 0.1$	0.1651	0.1651	0.2795	0.2002	$\hat{\sigma}_f^{(0)} = 0.1$	0.0391*	0.0872	0.1897	0.1897
$\hat{\sigma}_f^{(0)} = 0.15$	0.0991	0.1056	0.0437	0.1429	$\hat{\sigma}_f^{(0)} = 0.15$	0.0377*	0.1195	0.1795	0.2169
$\hat{\sigma}_f^{(0)} = 0.2$	0.1195	0.0739	0.1089	0.1388	$\hat{\sigma}_f^{(0)} = 0.02$	0.0288*	0.1308	0.3300	0.4031
$T = 0.86, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.86, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.1845	0.0218*	0.0691	0.0817	$\hat{\sigma}_f^{(0)} = 0.05$	0.0277	0.0668	0.0960	0.2468
$\hat{\sigma}_f^{(0)} = 0.1$	0.1515	0.0150*	0.0324*	0.0453*	$\hat{\sigma}_f^{(0)} = 0.1$	0.0336*	0.0288*	0.0505	0.1515
$\hat{\sigma}_f^{(0)} = 0.15$	0.1089	0.0363*	0.0185*	0.0300*	$\hat{\sigma}_f^{(0)} = 0.15$	0.0505	0.0817	0.0739	0.4118
$\hat{\sigma}_f^{(0)} = 0.2$	0.2057	0.1056	0.0790	0.1124	$\hat{\sigma}_f^{(0)} = 0.02$	0.0523	0.1023	0.3455	0.6957
$T = 0.88, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.88, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.1$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.05$	0.3225	0.0991	0.0901	0.1232	$\hat{\sigma}_f^{(0)} = 0.05$	0.1159	0.0790	0.1746	0.1124
$\hat{\sigma}_f^{(0)} = 0.1$	0.3861	0.1023	0.0421*	0.0960	$\hat{\sigma}_f^{(0)} = 0.1$	0.1347	0.2285	0.1124	0.1897
$\hat{\sigma}_f^{(0)} = 0.15$	0.2661	0.1308	0.1559	0.0991	$\hat{\sigma}_f^{(0)} = 0.15$	0.0668	0.2934	0.1604	0.3455
$\hat{\sigma}_f^{(0)} = 0.2$	0.5446	0.2057	0.2002	0.2345	$\hat{\sigma}_f^{(0)} = 0.02$	0.2934	0.2661	0.3778	0.8002

Table C.9

Dependency of (uncorrected) p-values (Mann-Whitney test) for ρ on $\hat{\sigma}_t^{(0)} = \sqrt{\hat{s}_t^{(0)}}$, $\hat{\sigma}_f^{(0)} = \sqrt{\hat{s}_f^{(0)}}$, β and T , following Local Approach 1; * and ** indicate $p < 0.05$ and $p < 0.005$ respectively.

$T = 0.2, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.2, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.8240	0.6731	0.7185	0.5651	0.4852	$\hat{\sigma}_f^{(0)} = 0.025$	0.6399	0.3861	0.4295	0.8841	0.7532
$\hat{\sigma}_f^{(0)} = 0.05$	0.3225	0.3861	0.1089	0.1429	0.2795	$\hat{\sigma}_f^{(0)} = 0.05$	0.4852	0.2169	0.1347	0.0624	0.1269
$\hat{\sigma}_f^{(0)} = 0.075$	0.4569	0.4295	0.4569	0.4118	0.5548	$\hat{\sigma}_f^{(0)} = 0.075$	0.8240	0.4569	0.2864	0.1089	0.1651
$\hat{\sigma}_f^{(0)} = 0.1$	0.7070	0.3534	0.7300	0.7185	0.3614	$\hat{\sigma}_f^{(0)} = 0.1$	0.7415	0.5144	0.6399	0.1472	0.0901
$\hat{\sigma}_f^{(0)} = 0.125$	0.8002	0.7648	0.7185	0.9084	0.6509	$\hat{\sigma}_f^{(0)} = 0.125$	0.7648	0.9206	0.7648	0.4295	0.2345
$\hat{\sigma}_f^{(0)} = 0.15$	0.7766	0.6399	0.8841	0.8841	0.8240	$\hat{\sigma}_f^{(0)} = 0.15$	0.8600	0.9084	0.7415	0.5446	0.4477
$T = 0.21, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.21, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.4569	0.5651	0.8963	0.9084	0.4031	$\hat{\sigma}_f^{(0)} = 0.025$	0.0288*	0.1388	0.3300	0.3778	0.3005
$\hat{\sigma}_f^{(0)} = 0.05$	0.8600	0.8963	0.8600	0.6844	0.2406	$\hat{\sigma}_f^{(0)} = 0.05$	0.1897	0.4031	0.6073	0.9816	0.9572
$\hat{\sigma}_f^{(0)} = 0.075$	0.8002	0.9694	0.7415	0.5244	0.3778	$\hat{\sigma}_f^{(0)} = 0.075$	0.3695	0.3695	0.7185	0.9450	0.8121
$\hat{\sigma}_f^{(0)} = 0.1$	0.5966	0.7648	0.7532	0.5446	0.3861	$\hat{\sigma}_f^{(0)} = 0.1$	0.1604	0.5966	0.6731	0.8841	0.8720
$\hat{\sigma}_f^{(0)} = 0.125$	0.5244	0.9694	0.6731	0.6844	0.4118	$\hat{\sigma}_f^{(0)} = 0.125$	0.1515	0.2406	0.5651	0.9694	0.6399
$\hat{\sigma}_f^{(0)} = 0.15$	0.5344	0.9206	0.9939	0.6073	0.6181	$\hat{\sigma}_f^{(0)} = 0.15$	0.1746	0.2112	0.4948	0.7884	0.7532
$T = 0.22, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.22, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.0437*	0.0872	0.2406	0.1124	0.1023	$\hat{\sigma}_f^{(0)} = 0.025$	0.0127*	0.0065*	0.0300*	0.0288*	0.0582
$\hat{\sigma}_f^{(0)} = 0.05$	0.0739	0.1651	0.4385	0.2595	0.3455	$\hat{\sigma}_f^{(0)} = 0.05$	0.0421*	0.0582	0.1269	0.1949	0.3861
$\hat{\sigma}_f^{(0)} = 0.075$	0.0562	0.2864	0.2345	0.4385	0.2468	$\hat{\sigma}_f^{(0)} = 0.075$	0.0668	0.1023	0.1559	0.2727	0.3861
$\hat{\sigma}_f^{(0)} = 0.1$	0.0336*	0.2345	0.3150	0.5446	0.2595	$\hat{\sigma}_f^{(0)} = 0.1$	0.0844	0.0930	0.1159	0.2406	0.5548
$\hat{\sigma}_f^{(0)} = 0.125$	0.0624	0.2468	0.2795	0.4569	0.5446	$\hat{\sigma}_f^{(0)} = 0.125$	0.1472	0.1124	0.0960	0.2661	0.5144
$\hat{\sigma}_f^{(0)} = 0.15$	0.0470*	0.1269	0.2595	0.4757	0.7185	$\hat{\sigma}_f^{(0)} = 0.15$	0.1269	0.1429	0.0817	0.1429	0.7648
$T = 0.23, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.23, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.0624	0.1232	0.1269	0.1308	0.1232	$\hat{\sigma}_f^{(0)} = 0.025$	0.1347	0.0991	0.0582	0.0901	0.1347
$\hat{\sigma}_f^{(0)} = 0.05$	0.1472	0.1023	0.1269	0.1308	0.2934	$\hat{\sigma}_f^{(0)} = 0.05$	0.0991	0.0960	0.1023	0.1089	0.1308
$\hat{\sigma}_f^{(0)} = 0.075$	0.1056	0.1949	0.1949	0.1604	0.1559	$\hat{\sigma}_f^{(0)} = 0.075$	0.1651	0.1559	0.1023	0.1124	0.1949
$\hat{\sigma}_f^{(0)} = 0.1$	0.1347	0.2002	0.1472	0.2169	0.2112	$\hat{\sigma}_f^{(0)} = 0.1$	0.1308	0.1308	0.1308	0.1604	0.2285
$\hat{\sigma}_f^{(0)} = 0.125$	0.1429	0.2057	0.1897	0.1559	0.2169	$\hat{\sigma}_f^{(0)} = 0.125$	0.1746	0.1559	0.1746	0.2406	0.2406
$\hat{\sigma}_f^{(0)} = 0.15$	0.1195	0.1347	0.0960	0.1023	0.1089	$\hat{\sigma}_f^{(0)} = 0.15$	0.1429	0.1651	0.1897	0.2406	0.2406
$T = 0.24, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.24, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.0116*	0.0246*	0.0377*	0.0645	0.1056	$\hat{\sigma}_f^{(0)} = 0.025$	0.0150*	0.0081*	0.0201*	0.0171*	0.0336*
$\hat{\sigma}_f^{(0)} = 0.05$	0.0930	0.2595	0.1845	0.1472	0.1559	$\hat{\sigma}_f^{(0)} = 0.05$	0.0218*	0.0406*	0.0764	0.0764	0.0739
$\hat{\sigma}_f^{(0)} = 0.075$	0.0901	0.1159	0.2057	0.1347	0.1949	$\hat{\sigma}_f^{(0)} = 0.075$	0.0311*	0.0453*	0.1269	0.1604	0.1651
$\hat{\sigma}_f^{(0)} = 0.1$	0.1124	0.1089	0.1388	0.1698	0.2057	$\hat{\sigma}_f^{(0)} = 0.1$	0.0421*	0.0562	0.1515	0.1515	0.1604
$\hat{\sigma}_f^{(0)} = 0.125$	0.1388	0.1124	0.1124	0.1124	0.1897	$\hat{\sigma}_f^{(0)} = 0.125$	0.0542	0.0817	0.1515	0.1159	0.1515
$\hat{\sigma}_f^{(0)} = 0.15$	0.1089	0.1023	0.0960	0.1515	0.1651	$\hat{\sigma}_f^{(0)} = 0.15$	0.0562	0.0624	0.1559	0.0930	0.1429
$T = 0.25, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.25, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.0844	0.1159	0.1023	0.1269	0.1159	$\hat{\sigma}_f^{(0)} = 0.025$	0.0349*	0.0991	0.0523	0.1195	0.2169
$\hat{\sigma}_f^{(0)} = 0.05$	0.0453*	0.0960	0.1159	0.0505	0.0691	$\hat{\sigma}_f^{(0)} = 0.05$	0.0645	0.0406*	0.1347	0.1949	0.1949
$\hat{\sigma}_f^{(0)} = 0.075$	0.0237*	0.0715	0.0960	0.0872	0.1515	$\hat{\sigma}_f^{(0)} = 0.075$	0.0138*	0.0624	0.0790	0.1651	0.1195
$\hat{\sigma}_f^{(0)} = 0.1$	0.0624	0.1089	0.0739	0.0715	0.1232	$\hat{\sigma}_f^{(0)} = 0.1$	0.0193*	0.0406*	0.0764	0.1023	0.1159
$\hat{\sigma}_f^{(0)} = 0.125$	0.0582	0.0764	0.0624	0.1429	0.0645	$\hat{\sigma}_f^{(0)} = 0.125$	0.0210*	0.0668	0.1232	0.0930	0.0930
$\hat{\sigma}_f^{(0)} = 0.15$	0.0542	0.0668	0.1056	0.1559	0.1089	$\hat{\sigma}_f^{(0)} = 0.15$	0.0324*	0.1023	0.1347	0.0715	0.0739

Table C.10

Dependency of (uncorrected) p-values (Mann-Whitney test) for s_t on $\hat{\sigma}_t^{(0)} = \sqrt{\hat{s}_t^{(0)}}$, $\hat{\sigma}_f^{(0)} = \sqrt{\hat{s}_f^{(0)}}$, β and T , following Local Approach 2; * indicates $p < 0.05$.

$T = 0.2, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.2, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.0089*	0.0078*	0.0081*	0.0127*	0.0164*	$\hat{\sigma}_f^{(0)} = 0.025$	0.0132*	0.0132*	0.0074*	0.0059*	0.0054*
$\hat{\sigma}_f^{(0)} = 0.05$	0.0324*	0.0237*	0.0227*	0.0227*	0.0218*	$\hat{\sigma}_f^{(0)} = 0.05$	0.0237*	0.0324*	0.0218*	0.0201*	0.0171*
$\hat{\sigma}_f^{(0)} = 0.075$	0.0300*	0.0288*	0.0277*	0.0237*	0.0218*	$\hat{\sigma}_f^{(0)} = 0.075$	0.0246*	0.0237*	0.0237*	0.0288*	0.0288*
$\hat{\sigma}_f^{(0)} = 0.1$	0.0256*	0.0246*	0.0277*	0.0201*	0.0218*	$\hat{\sigma}_f^{(0)} = 0.1$	0.0164*	0.0210*	0.0237*	0.0227*	0.0171*
$\hat{\sigma}_f^{(0)} = 0.125$	0.0336*	0.0437*	0.0391*	0.0218*	0.0277*	$\hat{\sigma}_f^{(0)} = 0.125$	0.0171*	0.0210*	0.0201*	0.0256*	0.0227*
$\hat{\sigma}_f^{(0)} = 0.15$	0.0421*	0.0487*	0.0542*	0.0300*	0.0266*	$\hat{\sigma}_f^{(0)} = 0.15$	0.0111*	0.0157*	0.0201*	0.0218*	0.0218*
$T = 0.21, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.21, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.0037**	0.0049**	0.0056*	0.0054*	0.0062*	$\hat{\sigma}_f^{(0)} = 0.025$	0.0054*	0.0049**	0.0031**	0.0035**	0.0027**
$\hat{\sigma}_f^{(0)} = 0.05$	0.0059*	0.0059*	0.0116*	0.0193*	0.0138*	$\hat{\sigma}_f^{(0)} = 0.05$	0.0024**	0.0022**	0.0037**	0.0030**	0.0031**
$\hat{\sigma}_f^{(0)} = 0.075$	0.0047**	0.0093*	0.0089*	0.0111*	0.0157*	$\hat{\sigma}_f^{(0)} = 0.075$	0.0023**	0.0022**	0.0019**	0.0033**	0.0030**
$\hat{\sigma}_f^{(0)} = 0.1$	0.0081*	0.0132*	0.0185*	0.0150*	0.0171*	$\hat{\sigma}_f^{(0)} = 0.1$	0.0023**	0.0035**	0.0027**	0.0040**	0.0037**
$\hat{\sigma}_f^{(0)} = 0.125$	0.0116*	0.0121*	0.0164*	0.0132*	0.0193*	$\hat{\sigma}_f^{(0)} = 0.125$	0.0023	0.0024**	0.0031**	0.0049**	0.0093*
$\hat{\sigma}_f^{(0)} = 0.15$	0.0097*	0.0164*	0.0171*	0.0178*	0.0193*	$\hat{\sigma}_f^{(0)} = 0.15$	0.0033**	0.0042**	0.0059*	0.0074*	0.0144*
$T = 0.22, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.22, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.0003**	0.0002**	0.0002**	0.0003**	0.0001**	$\hat{\sigma}_f^{(0)} = 0.025$	0.0027**	0.0011**	0.0015**	0.0018**	0.0021**
$\hat{\sigma}_f^{(0)} = 0.05$	0.0004**	0.0004**	0.0003**	0.0002**	0.0002**	$\hat{\sigma}_f^{(0)} = 0.05$	0.0003**	0.0003**	0.0002**	0.0001**	0.0002**
$\hat{\sigma}_f^{(0)} = 0.075$	0.0004**	0.0008**	0.0006**	0.0011**	0.0008**	$\hat{\sigma}_f^{(0)} = 0.075$	0.0009**	0.0004**	0.0004**	0.0003**	0.0002**
$\hat{\sigma}_f^{(0)} = 0.1$	0.0006**	0.0005**	0.0008**	0.0009**	0.0010**	$\hat{\sigma}_f^{(0)} = 0.1$	0.0005**	0.0005**	0.0004**	0.0003**	0.0003**
$\hat{\sigma}_f^{(0)} = 0.125$	0.0011**	0.0011**	0.0016**	0.0024**	0.0017**	$\hat{\sigma}_f^{(0)} = 0.125$	0.0016**	0.0013**	0.0009**	0.0007**	0.0007**
$\hat{\sigma}_f^{(0)} = 0.15$	0.0015**	0.0015**	0.0020**	0.0035**	0.0040**	$\hat{\sigma}_f^{(0)} = 0.15$	0.0026**	0.0024**	0.0011**	0.0007**	0.0013**
$T = 0.23, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.23, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.0116*	0.0116*	0.0150*	0.0132*	0.0150*	$\hat{\sigma}_f^{(0)} = 0.025$	0.0542	0.0437*	0.0300*	0.0336*	0.0185*
$\hat{\sigma}_f^{(0)} = 0.05$	0.0138*	0.0116*	0.0121*	0.0150*	0.0193*	$\hat{\sigma}_f^{(0)} = 0.05$	0.0150*	0.0106*	0.0097*	0.0078*	0.0085*
$\hat{\sigma}_f^{(0)} = 0.075$	0.0102*	0.0062*	0.0106*	0.0138*	0.0210*	$\hat{\sigma}_f^{(0)} = 0.075$	0.0288*	0.0144*	0.0074*	0.0078*	0.0127*
$\hat{\sigma}_f^{(0)} = 0.1$	0.0377*	0.0363*	0.0391*	0.0739	0.0645	$\hat{\sigma}_f^{(0)} = 0.1$	0.0324*	0.0201*	0.0111*	0.0116*	0.0157*
$\hat{\sigma}_f^{(0)} = 0.125$	0.0453*	0.0377*	0.0523	0.0715	0.1056	$\hat{\sigma}_f^{(0)} = 0.125$	0.0715	0.0377*	0.0336*	0.0453*	0.0210*
$\hat{\sigma}_f^{(0)} = 0.15$	0.0542	0.0930	0.0602	0.0991	0.1269	$\hat{\sigma}_f^{(0)} = 0.15$	0.0487*	0.0487*	0.0363*	0.0668	0.0602
$T = 0.24, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.24, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.0068*	0.0266*	0.0277*	0.0421*	0.0391*	$\hat{\sigma}_f^{(0)} = 0.025$	0.0102*	0.0035**	0.0097*	0.0116*	0.0144*
$\hat{\sigma}_f^{(0)} = 0.05$	0.2169	0.2002	0.2468	0.2795	0.2864	$\hat{\sigma}_f^{(0)} = 0.05$	0.0991	0.1308	0.0960	0.1651	0.1746
$\hat{\sigma}_f^{(0)} = 0.075$	0.3455	0.4948	0.5861	0.3778	0.3455	$\hat{\sigma}_f^{(0)} = 0.075$	0.1651	0.1795	0.2531	0.2864	0.3300
$\hat{\sigma}_f^{(0)} = 0.1$	0.5861	0.6731	0.6620	0.5755	0.5244	$\hat{\sigma}_f^{(0)} = 0.1$	0.1845	0.2468	0.4662	0.3946	0.4206
$\hat{\sigma}_f^{(0)} = 0.125$	0.8359	0.8002	0.6957	0.7532	0.6731	$\hat{\sigma}_f^{(0)} = 0.125$	0.4031	0.4477	0.5548	0.5446	0.5548
$\hat{\sigma}_f^{(0)} = 0.15$	0.9572	0.8359	0.7648	0.8121	0.6620	$\hat{\sigma}_f^{(0)} = 0.15$	0.4948	0.6509	0.7415	0.7070	0.7300
$T = 0.25, \beta = 0.01$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$	$T = 0.25, \beta = 0.001$	$\hat{\sigma}_t^{(0)} = 0.15$	$\hat{\sigma}_t^{(0)} = 0.175$	$\hat{\sigma}_t^{(0)} = 0.2$	$\hat{\sigma}_t^{(0)} = 0.225$	$\hat{\sigma}_t^{(0)} = 0.25$
$\hat{\sigma}_f^{(0)} = 0.025$	0.1429	0.1746	0.3300	0.4757	0.4118	$\hat{\sigma}_f^{(0)} = 0.025$	0.0363*	0.0487*	0.0421*	0.0645	0.0562
$\hat{\sigma}_f^{(0)} = 0.05$	0.3225	0.4385	0.4031	0.3377	0.3861	$\hat{\sigma}_f^{(0)} = 0.05$	0.2934	0.3225	0.4385	0.4569	0.4477
$\hat{\sigma}_f^{(0)} = 0.075$	0.3455	0.4948	0.4662	0.5244	0.4948	$\hat{\sigma}_f^{(0)} = 0.075$	0.2934	0.3861	0.4206	0.3377	0.3150
$\hat{\sigma}_f^{(0)} = 0.1$	0.5861	0.7070	0.7415	0.6844	0.7185	$\hat{\sigma}_f^{(0)} = 0.1$	0.4569	0.4662	0.6399	0.6957	0.6844
$\hat{\sigma}_f^{(0)} = 0.125$	0.7415	0.7532	0.8240	0.9084	0.8359	$\hat{\sigma}_f^{(0)} = 0.125$	0.6844	0.7185	0.8002	0.7185	0.6957
$\hat{\sigma}_f^{(0)} = 0.15$	0.8963	0.8720	0.9816	0.9328	0.9816	$\hat{\sigma}_f^{(0)} = 0.15$	0.7766	0.8720	0.9084	0.8600	0.8479

Table C.11

Dependency of (uncorrected) p-values (Mann-Whitney test) for ρ on $\hat{\sigma}_t^{(0)} = \sqrt{\hat{s}_t^{(0)}}$, $\hat{\sigma}_f^{(0)} = \sqrt{\hat{s}_f^{(0)}}$, β and T , following Local Approach 2; * and ** indicate $p < 0.05$ and $p < 0.005$ respectively.

D Statistical Post-Correction

Since we consider many different synchrony measures, computational approaches, and parameter settings simultaneously, we need to correct the p-values of the single Mann-Whitney tests appropriately (cf. Section 3.2.5). We will try to control the false discovery rate: the fraction of supposedly positive results that are in fact negative (Benjamini and Hochberg, 1995). For convenience, we list our two different schemes for statistical post-correction once more:

- (1) We apply the Storey method (Storey, 2002) to all comparisons simultaneously, that is all synchrony measures, the different approaches to compute those measures (Local Approach 1 and 2, Global Approach 1, 2 and 3), and their parameter settings, comprising 1270 statistical tests. We retain the smallest p-value for each synchrony measure, and as a result, we obtain one (corrected) p-value for each synchrony measure.
- (2) We apply a two-step procedure: first we correct for the multiple approaches to compute each synchrony measure (Local Approach 1 and 2, Global Approach 1, 2 and 3) and the multiple corresponding parameter settings, then we correct for the multiple synchrony measures. In the first step, we treat each synchrony method separately, and apply the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) to the p-values for different parameter settings and computational approaches (Local Approach 1 and 2, Global Approach 1, 2 and 3). For each synchrony measure, we retain the smallest corrected p-value among the different parameter settings and computational approaches. As a result, we obtain one (corrected) p-value for each synchrony measure. Those p-values are further corrected in a second step by means of the Storey method (Storey, 2002); that correction accounts for the multiple synchrony measures.

The p-values resulting from the first and second post-correction scheme are summarized in Table D.1 and D.2 respectively; we list the lowest corrected p-value for each synchrony measure. In the first correction scheme, a significance level of 0.05 is too liberal: it would result in 199 positives out of 1270 tests, with 10 expected false positives. Instead we choose the significance level so that there is less than one expected false positive; that corresponds to a significance level of 0.01. Only full-frequency DTF and ρ are significant on the 0.01 level. Under the second correction scheme, a significance level of 0.05 is a natural choice; only full-frequency DTF and ρ are significant on the 0.05 level. Interestingly, those two measures remain significant under both post-correction schemes. Moreover, as we have shown earlier, the combination of both measures leads to good classification results (cf. Table 4).

Measure	Correlation	Coherence	Phase Coherence	Corr-entropy	Wave-entropy	
p-value	0.052	0.064	0.077	0.067	0.11	
References	(Nunez & Srinivasan, 2006)			(Gunduz and Principe, 2009)		
Measure	MVAR coherence	Partial Coherence	PDC	DTF	ffDTF	dDTF
p-value	0.13	0.14	0.32	0.19	0.0098*	0.064
References	(Kamiński <i>et al.</i> , 2005)					
Measure	Kullback-Leibler	Rényi	Jensen-Shannon	Jensen-Rényi	I_W	I
p-value	0.096	0.096	0.096	0.097	0.088	0.094
References	(Aviyente, 2005a)					(Kraskov <i>et al.</i> , 2004)
Measure	N^k	S^k	H^k	S-estimator	Omega complexity	
p-value	0.064	0.054	0.088	0.077	0.10	
References	(Quiroga <i>et al.</i> , 2002)			(Carmeli <i>et al.</i> , 2005)	(Saito <i>et al.</i> , 1998)	
Measure	Hilbert Phase	Wavelet Phase	Evolution Map	Instantaneous Period	GFS	
p-value	0.13	0.10	0.096	0.054	0.066	
References	(Lachaux <i>et al.</i> , 1999)		(Rosenblum <i>et al.</i> , 2002)		(Koenig <i>et al.</i> , 2001)	
Measure	s_t	ρ				
p-value	0.027	0.005*				
References	(Dauwels <i>et al.</i> , 2007)					

Table D.1

Sensitivity of average synchrony for prediction of MCI: corrected p-values (Correction Scheme 1); * indicates $p < 0.01$.

Measure	Correlation	Coherence	Phase Coherence	Corr-entropy	Wave-entropy	
p-value	0.079	0.091	0.12	0.079	0.079	
References	(Nunez & Srinivasan, 2006)			(Gunduz and Principe, 2009)		
Measure	MVAR coherence	Partial Coherence	PDC	DTF	ffDTF	dDTF
p-value	0.13	0.23	0.31	0.27	0.033*	0.13
References	(Kamiński <i>et al.</i> , 2005)					
Measure	Kullback-Leibler	Rényi	Jensen-Shannon	Jensen-Rényi	I_W	I
p-value	0.079	0.079	0.079	0.12	0.087	0.79
References	(Aviyente, 2005a)					(Kraskov <i>et al.</i> , 2004)
Measure	N^k	S^k	H^k	S-estimator	Omega complexity	
p-value	0.97	0.13	0.13	0.34	0.13	
References	(Quiroga <i>et al.</i> , 2002)			(Carmeli <i>et al.</i> , 2005)	(Saito <i>et al.</i> , 1998)	
Measure	Hilbert Phase	Wavelet Phase	Evolution Map	Instantaneous Period	GFS	
p-value	0.25	0.079	0.13	0.091	0.13	
References	(Lachaux <i>et al.</i> , 1999)		(Rosenblum <i>et al.</i> , 2002)		(Koenig <i>et al.</i> , 2001)	
Measure	s_t	ρ				
p-value	0.13	0.025*				
References	(Dauwels <i>et al.</i> , 2007)					

Table D.2

Sensitivity of average synchrony for prediction of MCI: corrected p-values (Correction Scheme 2); * indicates $p < 0.05$.

References

- Adler G., Brassens S., and A. Jajcevic, 2003. EEG coherence in Alzheimer's dementia. *Journal of neural transmission*. 110:1051–1058.
- Ancona N., Marinazzo D., and Stramaglia S., 2004. Radial basis function approach to nonlinear Granger causality of time series. *Physical Review E* 70.
- Anghinah R., Kanda P., Jorge M., De Lima E., Pascuzzi L., and De Paiva Melo A., 2000. Alpha band coherence analysis of EEG in healthy adult and Alzheimer's type dementia subjects. *Arq. Neuropsiquiatr.* 58, 272–275.
- Astolfi L., Cincotti F., Mattia D., Salinari S., Babiloni C., Basilisco A., Rossini P. M., Lei Ding, Ying Ni, Bin He, Marciani M. G., and F. Babiloni, 2004. Estimation of the effective and functional human cortical connectivity with structural equation modeling and directed transfer function applied to high-resolution EEG. *Magnetic Resonance Imaging* 22, 1457–1470.
- Astolfi L., Cincotti F., Mattia D., Babiloni C., Carducci F., Basilisco A., Rossini P. M., Salinari S., Lei Ding, Ying Ni, Bin He, and Babiloni F., 2006. Assessing cortical functional connectivity by linear inverse estimation and directed transfer function: simulations and applications to real data. *Clinical Neurophysiology* 116, 920–932.
- Aviyente S., 2005. A measure of mutual information on the time-frequency plane. *Proc. of ICASSP 2005*, 481–484.
- Aviyente S., 2005. Information-theoretic signal processing on the time-frequency plane and applications. *Proc. of EUSIPCO 2005*.
- Babiloni C., Ferri R., Moretti D. V., Strambi A., Binetti G., Dal Forno G., Ferreri F., Lanuzza B., Bonato C., Nobili F., Rodriguez G., Salinari S., Passero S., Rocchi R., Stam C. J., and Rossini P. M., 2004. Abnormal fronto-parietal coupling of brain rhythms in mild Alzheimer's disease: a multicentric EEG study. *European J. Neuroscience* 19, 2583–2590.
- Babiloni C., Ferri R., Binetti G., Cassarino A., Forno G. D., Ercolani M., Ferreri F., Frisoni G. B., Lanuzza B., Miniussi C., Nobili F., Rodriguez G., Rundo F., Stam C. J., Musha T., Vecchio F., and Rossini P. M., 2006. Fronto-parietal coupling of brain rhythms in mild cognitive impairment: a multicentric EEG study. *Brain Res. Bull.* 69, 63–73.
- Benjamini Y. and Hochberg Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57 (1), 289–300.
- Besthorn C., Forstl H., Geiger-Kabisch C., Sattel H., Gasser T., and Schreiter-Gasser U, 1994. EEG Coherence in Alzheimer's disease. *Electroencephalogr Clin Neurophysiol* 90, 242–5.
- Blanco S., Quiroga R. Q., Rosso O., and Kochen S., 1995. Time-frequency analysis of EEG series. *Physical Review E* 51.
- Blinowska K. J., Kuś R., and Kamiński M., 2004. Granger causality and information flow in multivariate processes. *Physical Review E* 70.
- Bonferroni C. E., 1936. *Teoria statistica delle classi e calcolo delle probabilità*. Pubblicazioni del R. Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8, 3–62.
- Brassens S., Braus D., Weber-Fahr W., Trost H., Moritz S., and Adler G., 2004.

- Late-onset depression with mild cognitive deficits: electrophysiological evidences for a preclinical dementia syndrome. *Dementia and Geriatric Cognitive Disorders* 18, 271–277.
- Breakspear M., 2004. Dynamic connectivity in neural systems: theoretical and empirical considerations. *Neuroinformatics* 2.
- Buzsáki G., 2006. *Rhythms of the brain*. Oxford University Press.
- Matthew B. and Cutmore T., 2002. Low-probability event-detection and separation via statistical wavelet thresholding: an application to psychophysiological denoising. *Clinical Neurophysiology* 113, 1403–1411.
- Brunovsky M., Matousek M., Edman A., Cervena K., and Krajca V., 2003. Objective assessment of the degree of dementia by means of EEG. *Neuropsychobiology* 48, 19–26.
- Calderon P., Parra M., Llibre J., Fernandez A., and Gongora E., 1997. The role of the cerebral coherence in the progress of the patient with Alzheimer’s disease. *Rev Neurol.* 25, 1393–8.
- Carmeli C., Knyazeva M. G., Innocenti G. M., and De Feo O., 2005. Assessment of EEG synchronization based on state-space analysis. *Neuroimage* 25, 339–354.
- Carmo J. P., Dias N., Mendes P. M., Couto C., and Correia J. H., 2006. Low-power 2.4-GHz RF transceiver for wireless EEG module plug-and-play. *Proc. 13th IEEE International Conference on Electronics, Circuits and Systems (ICECS 2006)*, 1144–1147.
- Celka P. and Kilner B., 2006. Carmeli’s index assesses motion and muscle artefact reduction in rower’s electrocardiograms. *Physiological Measurement* 27, 737–755.
- Chapman R., Nowlis G., McCrary J., Chapman J., Sandoval T., Guillily M., Gardner M., Reilly L., 2007. Brain event-related potentials: diagnosing early-stage Alzheimer’s disease. *Neurobiol. Aging* 28, 194–201.
- Chen Y., Rangarajan G., Feng J., and Ding M., 2004. Analyzing multiple nonlinear time series with extended Granger causality. *Physics Letters A* 324, 26–35.
- Chen Y., Bressler S., and Ding M., 2006. Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data. *Journal of Neuroscience Methods*, 150:2.
- Cichocki A., Shishkin S., Musha T., Leonowicz Z., Asada T., and Kurachi T., 2005. EEG filtering based on blind source separation (BSS) for early diagnosis of Alzheimer’s disease. *Clin. Neurophys* 116, 729–37.
- Cover T. M. and Thomas J. A., 1991. *Elements of Information Theory*. Wiley, New York.
- Dauwels J., Vialatte F., Rutkowski T., and Cichocki A., 2007. Measuring neural synchrony by message passing, *Advances in Neural Information Processing Systems* 20 (NIPS 20), in press.
- Dauwels, J., F. Vialatte, T. Weber, and A. Cichocki, 2009. Quantifying statistical interdependence by message passing on graphs, Part I: One-dimensional point processes. *Neural Computation* 21:8, in press.
- Dauwels, J., F. Vialatte, T. Weber, T. Musha, and A. Cichocki, 2009. Quantifying statistical interdependence by message passing on graphs, Part II: Multi-dimensional point processes. *Neural Computation* 21:8, in press.
- Delprat N., Escudié B., Guillemain P., Kronland-Martinet R., Tchamitchian P., and Torrèsani B., 1992. Asymptotic wavelet and Gabor analysis: extraction of

- instantaneous frequencies. *IEEE Trans. Information Theory* 38, 644–664.
- Duda R., Hart P., and Stork D., 2000. *Pattern Classification*, 2nd Edition, Wiley-Interscience.
- Eichler M., 2006. On the evaluation of information flow in multivariate systems by the directed transfer function. *Biological Cybernetics* 94, 469–482.
- Goupillaud P., Grossman A., and Morlet J., 1984. Cycle-octave and related transforms in seismic signal analysis. *Geoexploration* 23, 85–102.
- Gourévitch B., Le Bouquin-Jeannès R., and Faucon G., 2006. Linear and nonlinear causality between signals: methods, examples, and neurophysiological applications. *Biological Cybernetics* 93, 349–369.
- Granger C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37, pp. 424–438.
- Grassberger P. and Procaccia I., 1983. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena* 9 (1-2): 189–208.
- Gunduz A. and J.C. Principe J.C., 2009. Correntropy as a novel measure for non-linearity tests. *Signal Processing* 89(1): 14–23.
- Güntekin B., Saatçi E., Yener G., 2008. Decrease of evoked delta, theta and alpha coherences in Alzheimer patients during a visual oddball paradigm. *Brain Res.* 1235:109–116.
- Herrmann C. S., Grigutsch M., and Busch N. A., 2005. EEG Oscillations and Wavelet Analysis. In: Handy, T. (ed.), *Event-Related Potentials: a Methods Handbook*, Cambridge, MIT Press, pp. 229–259.
- Herrmann C. S. and Demiralp T., 2005. Human EEG gamma oscillations in neuropsychiatric disorders. *Clinical Neurophysiology* 116, 2719–2733.
- Hidasi Z., Czizler B., Salacz P., Csibri E., Molnár M., 2007. Changes of EEG spectra and coherence following performance in a cognitive task in Alzheimer’s disease. *Int. J. Psychophysiol.* 65(3):252–260.
- Hogan M., Swanwick G., Kaiser J., Rowan M. and Lawlor B., 2003. Memory-related EEG power and coherence reductions in mild Alzheimer’s disease. *Int. J. Psychophysiol.* 49, 147–163.
- Ikawa M., Nakanishi M., Furukawa T., Nakaaki S., Hori S., and Yoshida S., 2000. Relationship between EEG dimensional complexity and neuropsychological findings in Alzheimer’s disease. *Psychiatry Clin. Neurosci.* 54, 537–41.
- Jelic V., Shigeta M., Julin P., Almkvist O., Winblad B., and Wahlund L., 1996. Quantitative EEG power and coherence in Alzheimer’s disease and mild cognitive impairment. *Dementia* 7, 314–23.
- Jelles B., Scheltens P., van der Flier W.M., Jonkman E.J., da Silva F.H., and Stam C.J., 2008. Global dynamical analysis of the EEG in Alzheimer’s disease: frequency-specific changes of functional interactions. *Clin. Neurophysiol.* 119(4):837–841.
- Jeong J., Gore J., and Peterson B., 2001. Mutual information analysis of the EEG in patients with Alzheimer’s disease. *Clinical Neurophysiology* 112, 827–835.
- Jeong J., 2004. EEG Dynamics in patients with Alzheimer’s disease. *Clinical Neurophysiology* 115, 1490–1505.
- Jiang Z. Y., 2005. Abnormal Cortical Functional Connections in Alzheimer’s Disease: Analysis of Inter- and Intra-Hemispheric EEG Coherence. *J Zhejiang Univ Sci B.* 6, 259–264.

- Jiang Z. Y., 2005. Study on EEG Power and Coherence in Patients with Mild Cognitive Impairment During Working Memory Task. *J. Zhejiang Univ. Sci B.* 6, 1213–9.
- Jiang Z. Y. and Zheng L. L., 2006. Inter-and intra-hemispheric EEG coherence in patients with mild cognitive impairment at rest and during working memory task. *J. Zhejiang Univ Sci B.* 7, 357–364.
- Johnson D. and Sinanovic S., 2007. Symmetrizing the Kullback-Leibler distance. *IEEE Trans. on Information Theory* (submitted).
- Jordan M. I. (ed.), *Learning in Graphical Models*, MIT Press, 1999.
- Ji Jun, Yu Meng Sun, Zhou Yu Bin, and Jin Zhang Rui, 2006. A wireless EEG sensors system for computer assisted detection of alpha wave in sleep. *Proc. 27th Annual International Conference of Engineering in Medicine and Biology Society (IEEE-EMBS 2005)*, 5351–5353.
- Kamiński M. and Blinowska K. J., 1991. A new method of the description of the information flow in the brain structures. *Biol. Cybern.* 65, 203–210.
- Kamiński M. and Hualou Liang, 2005. Causal influence: advances in neurosignal analysis. *Critical review in biomedical engineering* 33, (4):347–430.
- Koenig T., Lehmann D., Saito N., Kuginuki T., Kinoshita T., and Koukkou M., 2001. Decreased functional connectivity of EEG theta-frequency activity in first-episode, neuroleptic-naive patients with schizophrenia: preliminary results. *Schizophrenia Res.* 1–2, 55–60.
- Koenig T., Prichep L., Dierks T., Hubl D., Wahlund L. O., John E. R., and Jelic V., 2005. Decreased EEG synchronization in Alzheimer’s disease and mild cognitive impairment. *Neurobiol. Aging.* 26, 165–71.
- Knott V., Mohr E., Mahoney C., and Ilivitsky V., 2000. EEG coherence in Alzheimer’s disease: comparisons with a control group and population norms. *J. Geriatr. Psychiatry Neurol.* 13, 1–8.
- Korzeniewska A., Mańczak M., Kamiński M., Blinowska K. J., and Kasicki S., 2003. Determination of information flow direction amongst brain structures by a modified directed transfer function (dDTF) method. *J. Neuroscience Methods* 125, 195–207.
- Kramer M., Chang F. L., Cohen M., Hudson D., and Szeri A., 2007. Synchronization measures of the scalp EEG can discriminate healthy from Alzheimer’s subjects. *International Journal of Neural Systems* 17, 1–9.
- Kraskov A., Stögbauer H., and Grassberger P., 2004. Estimating mutual information. *Phys. Rev. E* 69.
- Lachaux J.-P., Rodriguez E., Martinerie J., and Varela F. J., 1999. Measuring phase synchrony in brain signals. *Human Brain Mapping* 8, 194–208.
- Leuchter A., Newton T., Cook I., Walter D., Rosenberg-Thompson S., and Lachenbruch P., 1992. Changes in brain functional connectivity in Alzheimer-type and multi-infarct dementia. *Brain* 115, 1543–61.
- Locatelli T., Cursi M., Liberati D., Franceschi M., and Comi G., 1998. EEG coherence in Alzheimer’s disease. *Electroencephalogr. Clin. Neurophysiol.* 106, 229–37.
- Loeliger H.-A., Dauwels J., Hu J., Korl S., Li Ping, and Kschischang F., 2007. The factor graph approach to model-based signal processing. *Proceedings of the IEEE* 95(6), 1295–1322.
- Musha T., Asada T., Yamashita F., Kinoshita T., Chen Z., Matsuda H., Uno M.,

- Shankle W.R., 2002. A new EEG method for estimating cortical neuronal impairment that is sensitive to early stage Alzheimers disease. *Clin. Neurophysiol* 113, 1052–8.
- Nunez P. and Srinivasan R., 2006. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press.
- Pereda E., Quiroga R. Q., and Bhattacharya J., 2005. Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology* 77, 1–37.
- Pijnenburg Y. A., van de Made Y., van Cappellen van Walsum A. M., Knol D. L., Scheltens P., Stam C. J., 2004. EEG Synchronization likelihood in mild cognitive impairment and Alzheimer’s disease during a working memory task. *Clin. Neurophysiol.* 115, 1332–9.
- Quiroga Q. R., Rosso O., and Basar E., 1999. Wavelet-entropy: a measure of order in evoked potentials. *Electr. Clin. Neurophysiol. (Suppl.)* 49, 298–302.
- Quiroga Q. R., Rosso O., Schürmann M., and Basar E., 2001. Wavelet entropy in event-related potentials: a new method shows frequency tuning of EEG-oscillations. *Biological Cybernetics* 84, 291–299.
- Quiroga R. Q., Kraskov A., Kreuz T., and Grassberger P., 2002. Performance of different synchronization measures in real data: a case study on EEG signals. *Physical Review E* 65.
- Rosenblum M. G., Cimponeriu L., Bezerianos A., Patzak A., and Mrowka R., 2002. Identification of coupling direction: application to cardiorespiratory interaction. *Physical Review E* 65.
- Rossini P. M., Del Percio C., Pasqualetti P., Cassetta E., Binetti G., Dal Forno G., Ferreri F., Frisoni G., Chioventa P., Miniussi C., Parisi L., Tombini M., Vecchio F., and Babiloni C., 2006. Conversion from mild cognitive impairment to Alzheimer’s disease is predicted by sources and coherence of brain electroencephalography rhythms. *Neuroscience* 143, 793–803.
- Saito N., Kuginuki T., Yagyū T., Kinoshita T., Koenig T., Pascual-Marqui R. D., Kochi K., Wackermann J., and Lemann D., 1998. Global, regional and local measures of complexity of multichannel EEG in acute, neuroleptic-naive, first-break schizophrenics. *Society of Biological Psychiatry*. 43:794–802.
- Sakkalis V., Giurcăneacu C. D., Xanthopoulos P., Zervakis M., Tsiaras V., Yianghua Yang, and Micheloyannis S., 2006. Assessment of linear and non-linear EEG synchronization measures for evaluating mild epileptic signal patterns. *Proc. of ITAB 2006, Ioannina-Epirus, Greece*.
- Sauer T., Yorke J., and Casdagli M., 2005. Embedology. *Journal of Statistical Physics* 65:579-616.
- Schelter B., Winterhalder M., Eichler M., Peifer M., Hellwig B., Guschlbauer B., Lübbing C. H., Dahlhaus R., and Timmer J., 2005. Testing for directed influences among neural signals using partial directed coherence. *J. Neuroscience Methods* 152, 210–219.
- Shawe-Taylor J. and Cristianini N., 2004. *Kernel Methods for Pattern Analysis*, Cambridge University Press.
- Stam C. J. and van Dijk B. W., 2002. Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets. *Physica D* 163, 236–251.
- Stam C. J., van der Made Y., Pijnenburg Y. A. L., and Scheltens Ph., 2003. EEG

- synchronization in mild cognitive impairment and Alzheimer’s disease. *Acta Neurologica Scandinavica* 108, 90–96.
- Stam C. J., Montez T., Jones B. F., Rombouts S. A. R. B., van der Made Y., Pijnenburg Y. A. L., and Scheltens Ph., 2005. Disturbed fluctuations of resting state EEG synchronization in Alzheimer’s disease. *Clinical Neurophysiology* 116, 708–715.
- Stam C. J., 2005. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clinical Neurophysiology* 116, 2266–2301.
- Stam C.J., Nolte G., Daffertshofer A., 2007. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Hum. Brain Mapp.* 28(11):1178–1193.
- Stevens A., Kircher T., Nickola M., Bartels M., Rosellen N., and Wormstall H., 2001. Dynamic regulation of EEG power and coherence is lost early and globally in probable DAT. *Eur. Arch. Psychiatry Clin. Neurosci.* 251, 199–204.
- Storey J., 2002. A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B (Methodological)* 64 (3): 479–498.
- Tallon-Baudry C., Bertrand O., Delpuech C., and Pernier J., 1996. Stimulus specificity of phase-locked and non-phase-locked 40Hz visual responses in human. *Journal of Neuroscience* 16, 4240–4249.
- Uhlhaas P. and Singer W., 2006. Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron* 52, 155–168.
- Vialatte F., 2005. *Modélisation en bosses pour l’analyse des motifs oscillatoires reproductibles dans l’activité de populations neuronales : applications à l’apprentissage olfactif chez l’animal et à la détection précoce de la maladie d’Alzheimer*, PhD Thesis, Paris VI University, Paris.
- Vialatte F., Cichocki A., Dreyfus G., Musha T., Rutkowski T. M., and Gervais R., 2005. Blind source separation and sparse bump modelling of time-frequency representation of EEG signals: new tools for early detection of Alzheimer’s disease. *Proc. IEEE Workshop on Machine Learning for Signal Processing*, 27–32.
- Vialatte F., Martin C., Dubois R., Haddad J., Quenet B., Gervais R., and Dreyfus G., 2007. A machine learning approach to the analysis of time-frequency maps, and its application to neural dynamics. *Neural Networks* 20, 194–209.
- Vialatte F. and A. Cichocki, 2008. Split-test Bonferroni correction for QEEG statistical maps. *Biol Cybern.* 98(4):295–303.
- Vialatte F., J. Dauwels, M. Maurice, Y. Yamaguchi, and A. Cichocki, 2009. On the synchrony of steady state visual evoked potentials and oscillatory burst events. *Cogn Neurodynamics*, Epub ahead of print.
- Vialatte F., J. Sol-Casals, J. Dauwels, M. Maurice, and A. Cichocki, 2009. Bump time-frequency toolbox: a toolbox for time-frequency oscillatory bursts extraction in electrophysiological signals. *BMC Neuroscience* 10:46.
- Wada Y., Nanbu Y., Koshino Y., Yamaguchi N., and Hashimoto T., 1998. Reduced interhemispheric EEG coherence in Alzheimer’s disease: analysis during rest and photic stimulation. *Alzheimer Dis. Assoc. Disord.* 12, 175–81.
- Wan B., Ming D., Qi H., Xue Z., Yin Y., Zhou Z., Cheng L., 2008. Linear and nonlinear quantitative EEG analysis. *IEEE Eng. Med. Biol. Mag.* 27(5):58–63.
- Yagyu T., Wackermann J., Shigeta M., Jelic V., Kinoshita T., K. Kochi, P. Julin, O. Almkvist, L. O. Wahlund, I. Kondakor, and Lehmann D., 1997. Global di-

mensional complexity of multichannel EEG in mild Alzheimer's disease and age-matched cohorts. *Dement. Geriatr. Cogn. Disord.* 8, 343–7.

Yoshimura M., Isotani T., Yagyu T., Irisawa S., Yoshida T., Sugiyama M., Minami T., Sugimoto T., Nobuhara K., Okugawa G., and Kinoshita T., 2004. Global approach to multichannel electroencephalogram analysis for diagnosis and clinical evaluation in mild Alzheimer's disease. *Neuropsychobiology* 49:163–166.