International Journal of Bifurcation and Chaos, Vol. 20, No. 6 (2010) 1703–1721

DOI: 10.1142/S0218127410026770

SOURCE RECONSTRUCTION AND SYNCHRONY MEASUREMENTS FOR REVEALING FUNCTIONAL BRAIN NETWORKS AND CLASSIFYING MENTAL STATES

FRANÇOIS LAURENT[*,†], MICHEL BESSERVE[*,†], LINE GARNERO[*,†], MATTHIEU PHILIPPE[‡], GENEVIÈVE FLORENCE[‡] and JACQUES MARTINERIE[*,†,§]

[*]*UPMC Univ. Paris 06, Paris, France*

[†]*CNRS, UPR 640 LENA, Cognitive Neuroscience, and Brain Imaging Laboratory, Paris, France*

[‡]*Institut de Médecine Aérospatiale du Service, de Santé des Armées, Brétigny-sur-Orge, France*

[§]*jacques.martinerie@upmc.fr*

Received November 18, 2008; Revised August 8, 2009

We classified performance-related mental states from EEG-derived measurements. We investigated the usefulness of massively distributed source reconstruction, comparing scalp and cortical scales. This approach provides a more detailed picture of the functional brain networks underlying the changes related to the mental state of interest. Local and distant synchrony measurements (coherence, phase locking value) were used for both scalp measurements and cortical current density sources, and were fed into a SVM-based classifier. We designed two simulations where classification scores increased when our 60-electrode scalp measurements were reconstructed on 60 sources and on a 500-source cortex. Source reconstruction appeared to be most useful in these simulations, in particular, when distant synchronies were involved and local synchronies did not prevail. Despite the simplicity of the model used, certain flaws in accuracy were observed in the localization of informative activities, due to the relationship between amplitude and phase for mixed signals. Our results with real EEG data suggested that the phenomenon of interest was characterized merely by modulations in local amplitudes, but also in strength of distant couplings. After source reconstruction, classification rates also increased for real EEG data when seeking distant phase-related couplings. When reconstructing a large number of sources, the regularization coefficient should be carefully selected on a subject-by-subject basis. We showed that training classifiers using such high-dimension data is useful for localizing discriminating patterns of activity.

Keywords: EEG source reconstruction; coherence; phase locking values; classification of metal states.

[§]Address for correspondence: Laboratoire de Neurosciences Cognitives et Imagerie Cérébrale, CNRS UPR640 — LENA, Hôpital de la Salpêtrière, 47, bd de l'hôpital, 75651 Paris Cedex 13, France

1. Introduction

Most studies of neuronal activity in humans are based on either functional magnetic resonance imaging (fMRI) or electro- or magneto-electroencephalography (E/MEG). fMRI has a high spatial resolution and can therefore be used to study complex brain networks [Eguiluz *et al.*, 2005; Achard *et al.*, 2006]. However, the low temporal resolution limits the study of the dynamics. E/MEG displays the opposite trade-off: high temporal resolution but low spatial resolution. Nevertheless, many attempts have been made to reveal the structure of a functional network from E/MEG data, and interesting dynamic links between scalp sensors have been identified. Moreover, distributed brain source reconstruction can be used to increase the spatial resolution of E/MEG. Source reconstruction has been reviewed in [Baillet, 2007]. It makes use of a large number of current dipoles (or sources) located, for example, in grey matter in the brain, with the assistance of anatomical magnetic resonance image (MRI). In most currently used models, only the surface of the cortex is considered of interest because it is difficult to reconstruct deep sources correctly. Some studies have combined E/MEG source reconstruction and analyses of a functional brain network as a set of distant-site dynamic links [ten Caat *et al.*, 2008; Astolfi *et al.*, 2007; De Vico Fallani *et al.*, 2007]. These links are considered to be functional couplings, because they are quantitatively scored from observed or estimated activities. Many measurements can be used to quantify the synchrony between two signals.

This study aims to classify mental states from cerebral electrical activities quantified in different ways. Classification of mental states implies the definition of at least two classes, for example, two sets of time windows each related to a single mental state, and to find differences in the signal between classes. The idea is to process the signal a given way, and to assess the amount of thus-extracted information that would enable to discriminate between two or more mental states. In this paper, we investigate the discriminating power of two coupling measures — standard coherence and phase locking value (further details provided below) — for quantifying dynamic links between sources reconstructed on the cerebral cortex from EEG data. The data is such that 60 electrode signals gave rise to about 500 reconstructed source signals. The notion of "dynamic brain network" is therefore much closer to that in fMRI. We compared different levels of resolution (measured scalp potentials versus estimated cortex current densities) and distant synchrony measurements plus a conventional local synchrony measurement: mean amplitude (mean amplitude versus coherence versus phase locking value). We also considered mean amplitude, although this measurement estimates only local synchrony, because energy-like features have proved highly useful in many studies when calculated for frequency bands specific to E/MEG. This work is also applied in nature, borrowing tools and methods from the Brain Computer Interface (BCI) community (see [Wolpaw *et al.*, 2002] for an introduction). Here, spatial scales and functional features are assessed in terms of their ability to discriminate between two predefined mental states, using a classification/validation framework. Discrimination between selected mental states is a clear statistical paradigm with immediate applications. When mental states are known to have strong correlates in a specific region, the localization power of inverse methods alone may be sufficient to detect the occurrence of the mental state concerned, as shown by Qin *et al.* [2004], Congedo *et al.* [2006] and Kamousi *et al.* [2007] for motor imaging. However, less may be known about activity patterns underlying the mental state of interest, and it may therefore be necessary to extract several features from brain electrical activity and to feed these features into a multivariate statistical tool. The use of classification algorithms, such as the support vector machine (SVM, see [Burges, 1998] for an introduction) seems to be useful in situations in which the aim is to assess the amount of discriminating information encoded in several variables, such as energy-like features quantified in several frequency bands and at several spatial locations [Lotte *et al.*, 2007; Noirhomme *et al.*, 2008].

None of the BCI studies cited above made use of distant synchrony features. However, the observation of such features may be the best way to observe some phenomena common to many high-level cognitive states and consisting of synchronized oscillatory patterns simultaneously in several cortical regions. These phenomena result from the formation of "resonant neuronal assemblies" [Varela, 1995], which occurs during tasks involving memory, perception, attention or consciousness [Rodriguez *et al.*, 1999; Varela *et al.*, 2001]. This provides one

of the main justifications for high-complexity functional network methods. Such methods are suitable for BCI applications despite increasing computational costs.

In this study, virtual mental states were used to study, through simulations, the usefulness of source reconstruction and distant synchrony features for discrimination and localization purposes. Results for real data are presented, discriminating between two performance levels during a Simon task involving 12 subjects. This study was based on a previous investigation [Besserve *et al.*, 2008]. Two levels (low versus high mean reaction times) of performance in a spatial compatibility task (or Simon task) were to be discriminated from data from 60 EEG electrodes. Mean amplitude and phase locking value were determined for three frequency bands (theta: 3–7 Hz, alpha: 7–13 Hz, beta: 13–18 Hz). Mean amplitudes were determined for each electrode and phase locking values were determined for each electrode pair. Two classification/validation frameworks were compared.

2. Material and Methods

2.1. *Processing chain*

Figure 1 illustrates the main steps in the processing chain involved in the study for source reconstruction. Scalp measurements were complemented with anatomical information obtained from MRI segmentation items. In particular, as explained in more detail in the **Source reconstruction** section, a surface was extracted for the casting of a given number of dipoles corresponding to our sources. EEG electrode locations are also required for source reconstruction. Once activities were reconstructed, the signal was segmented, applying a 20 s sliding time window. A constant number of features were estimated from the signal in each time window (these features are described in the **EEG quantification** section). Each window is assigned to one of two classes, as described below, resulting in data for two sets of time windows in the form of two clouds of points in a multidimensional space (one dimension per feature). The classification/

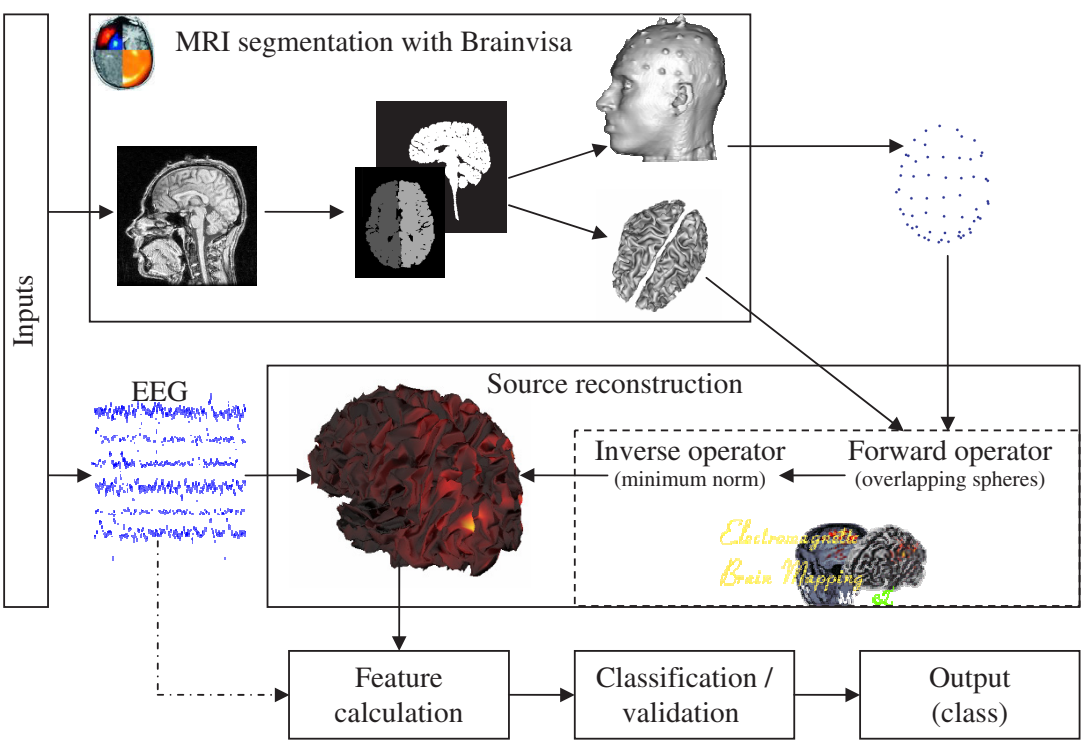

Fig. 1. Illustration of the processing steps involved in the study. Each subject was processed independently, using both T1 MRI and EEG recordings. MRI were segmented with Brainvisa software, to extract a cortical surface onto which sources were located, the inner skull surface and electrode locations. Forward and inverse operators were calculated with the BrainStorm MatLab toolbox. The projection of scalp measurements through the inverse matrix yielded estimates of current density time course. Reconstructed signals were then quantified. The resulting features were fed to a classification/validation process, which eventually yielded class estimates for the corresponding time windows. Features can be calculated directly from scalp measurements, as illustrated by the dashed line. Classification results with and without source reconstruction are compared.

validation step aims to identify the optimal surface separating these two sets of time windows (or points) and to assess the ability of this surface to predict a class of new (closely related but different) time windows. The procedure is explained in more detail in **Classification, validation and statistics** section.

The complete procedure was applied to generate classification scores in different situations. These situations were then compared by means of the classification scores obtained. We investigated three sets of conditions. The first concerned the type of data (one of the two simulations or real data). The second concerned the level at which activities were determined (scalp measurements, reconstructed sources with variable resolution) and the third concerned the feature used to quantify activities (mean amplitude, coherence and phase locking value, defined later).

2.2. *EEG quantification*

EEG quantification was carried out for both simulation data and real data. The EEG was acquired from or simulated for 60 EEG electrodes. Data was segmented into 20 s time windows. Each window was assigned to a class 1 or 2, the construction and significance of these classes being explained below.

Within a window, analytic signals were estimated for one or several frequency bands:

$$z_{[f_1,f_2]}(k,t) = a_{[f_1,f_2]}(k,t) \cdot e^{j\phi_{[f_1,f_2]}(k,t)}$$

for all sensors k, times t and frequency bands $[f_1, f_2]$.

The frequency bands explored for real data were 3–7 Hz (theta), 7–13 Hz (alpha) and 13–18 Hz (beta), as described by Besserve *et al.* [2008]. For simulations, only the 7–13 Hz frequency band was used, to keep the model simple.

The analytic signal was used to sift amplitude (a) and phase (ϕ) information. Mean amplitude depends on the analytic signal, so does phase locking value (PLV, [Lachaux *et al.*, 1999]):

$$\mathrm{PLV}_{[f_1,f_2]}(k,l,n) = \left| \frac{1}{T_n} \sum_{t=t_{n,0}}^{t_{n,0}+T_n} e^{j(\phi_{[f_1,f_2]}(k,t)-\phi_{[f_1,f_2]}(l,t))} \right|$$

for all different sensors k and l, windows n, and frequency bands $[f_1, f_2]$, with $t_{n,0}$ denoting the beginning time of the window and T_n its duration.

Standard coherence is also estimated, using fast Fourier transformations (FFTs) of 19 2 s Hanning subwindows with a 1 s overlap:

$$\mathrm{Coh}_{[f_1,f_2]}(k,l,n) = \sum_{m=1}^{19} \frac{|S^{n,m}_{[f_1,f_2]}(k,l)|}{S^{n,m}_{[f_1,f_2]}(k,k) \cdot S^{n,m}_{[f_1,f_2]}(l,l)}$$

for all different sensors k and l, windows n, and frequency bands $[f_1, f_2]$, with $S^{n,m}_{[f_1,f_2]}(k,l)$ denoting the cross-spectrum between sensors k and l, calculated on subwindow m of time window n.

2.3. *Source reconstruction*

T1 MRIs were segmented with Brainvisa software (http://brainvisa.info), to obtain the white matter/cortical gray matter surface and the interior surface of the skull as triangle meshes. A first processing step aims at computing the contribution to the scalp electric signal of activities from small cortical areas represented as dipoles. This step corresponds to the "forward problem". The location of dipoles were constrained to be the vertices of the cortical mesh, with directions expressed as the mean normal vectors of adjacent faces of the mesh, oriented toward the scalp and normalized. The number of dipoles was typically about 10 000. The contribution of the dipoles was modeled as a spatial transformation matrix A, with as many rows as electrodes and as many columns as sources. The method used here was known as the overlapping spheres method [Huang *et al.*, 1999] and was implemented in the BrainStorm MatLab toolbox (http://neuroimage.usc.edu/brainstorm).

A second step, or the "inverse problem", is the reconstruction of source signals from scalp signals and the forward matrix. We used the simplest available method, involved in the regularized pseudoinversion of the forward matrix and known as the minimum-norm inverse solution [Hämäläinen & Sarvas, 1989]:

$$G = A^t(AA^t + \alpha \mathrm{I})^{-1}$$

with A the forward matrix, α the regularization coefficient, generally equal to (λ=) 10% of the maximal singular value of A, and I the identity matrix in the electrode space (or measurement space).

The αI matrix may be interpreted as an estimate of measurement noise [Dale & Sereno, 1993]. We can therefore consider α as both a smoothing parameter and a threshold for the signal-to-noise ratio.

The final transformation was thus a time-independent spatial filter that projected the original

60 signals into a space of about 10 000 current density estimates. We could have estimated the source signal and computed the same measurements as considered for the scalp (mean amplitude, coherence and PLV in the frequency bands of interest). However, two problems arose at this step. First, there were computational limitations due to the combinatorial burden of coupling features such as PLV and coherence. Second, a number of sources higher than the number of electrodes would have made the number of features different between scalp and source levels; the "source" classifier would have had more input features and consequently, would have been more prone to overfit than its "scalp" counterpart. For these reasons, we compared classification scores using 60 sources only, i.e. as many sources as electrodes. For localization purposes, we drew statistical maps with sources reconstructed on a number of cortical dipoles reduced down to 500, leading to 1 24 750 pairs; this latter number must be multiplied by the number of frequency bands when estimating the total number of features yielded by each coupling measure.

The 60 dipoles used for classification purposes were selected the following way: for each electrode i, a subset S_i of 10 nearest neighbor dipoles were selected; then the winning dipole j^* was $j^* = \arg\max_{j \in S_i}(A_{j,i}/(\sum_k A_{j,k}))$, which can be thought as the most correlated source with the electrode signal; in the end, the final inverse matrix was the square matrix $G_{J,.}$, where J was the subset of all the selected dipoles. This inverse matrix was thus a kind of anatomy-aware Laplacian operator.

The 500 dipoles used for localization purposes were obtained by a two-step procedure. First, we spatially decimated the original 10 000-vertex mesh down to a 500-vertex mesh, using MathWorks' function "reducepatch". Then, for each new vertex, we selected the nearest one in the original mesh. As the reconstruction of a higher number of sources than that of electrodes may critically depend on the regularization coefficient of the minimum-norm inverse solution, we screened the following values of the λ parameter: 1%, 5%, 10%, 15% and 20%. This screening is not intended to be comprehensive.

2.4. *Classification, validation and statistics*

Features computed from segmented time windows were fed into a ten-fold cross-validation process. For real data, the cross-validation process was repeated ten times with different initializations of the random partitioning of labeled windows. This latter procedure was intended to reduce the inter-subject variability. The cross-validation process consisted of an iterative scheme in which two sets of time windows were defined. The first set, the learning set, allowed the algorithms to fit the classifier to the data, and the second set, the validation set, was used to assess the accuracy of the learnt classifier, comparing its output with the expected class labels.

Each subject or simulated data set was processed independently, yielding 100 classification estimates consisting of the ratio of correctly classified validation windows to the total number of validation windows.

The classification algorithm used was a binary linear support vector machine (SVM, see [Schölkopf & Smola, 2002] implemented in the simpleSVM MatLab toolbox (http://gaelle.loosli.fr/research/tools/simplesvm.html, [Vishwanathan *et al.*, 2003]). Support vector machines are a state-of-the-art classification tool. Only one hyperparameter, a regularization coefficient c, needed to be set. This coefficient can be set to infinity if the number of features in each time window exceeds the number of time windows provided that no feature is linearly dependent on any other feature. This situation appears to apply in all cases other than for the use of mean amplitudes with simulation data (due to the use of a single frequency band in this case). If there were too few features, we arbitrarily set c to 1.

Windows from the learning set overlapping windows from the validation set were discarded, during each learning step in the cross-validation iteration. This dramatically decreased the classification rate, but yielded more realistic estimates.

Classification scores for all subjects or simulated data sets were then digested using classical univariate statistical tests. A result was considered significant if the false positive rate of the test was no greater than 0.05.

In addition to higher classification scores, localization of key activities was one of our main concerns. We used the linear SVMs to build statistical maps. Indeed, a trained linear SVM is actually a separating plane in the feature space — a set of equal numbers of weights and variables. Thus, weights computed from mean amplitude, for example, can be mapped onto a scalp map or a cortical map, depending on the chosen level. These weights are calculated in a multivariate process and the result may therefore differ from a T-score map. The T-score we refer to was the Student's T-test

value, calculated for each feature, between windows from class 1 and windows from class 2, and then standardized based on the mean and standard deviation of T-test values from 100 permutations of windows (or class labels).

$$T_{\text{score},i} = \frac{T_i - \overline{T_{\text{surrogate},i}}}{\sigma(T_{\text{surrogate},i})} \quad \text{with}$$

$$T_i = \frac{\overline{X_{i,1}} - \overline{X_{i,2}}}{\sqrt{\frac{(n_1 - 1) \cdot \sigma^2(X_{i,1}) + (n_2 - 1) \cdot \sigma^2(X_{i,2})}{n1 + n2 - 2}}}$$

Average SVM weights were calculated using the following procedure. First, for each trained SVM, the set of weights (or orthogonal vector of the separating hyperplane) was normalized by dividing each weight by the norm of the vector. Then, these normal vectors were weighted-averaged across cross-validation iterations and cross-validation restarts, with corresponding classification rates as weightings; as each trained SVM was associated to a classification rate calculated using some validation data, multiplying the weights of the corresponding vector by this classification rate (divided by the sum of all the classification rates) yielded average weights taking into account generalizability of the trained SVM. Finally, vectors were averaged across subjects or simulation datasets.

We calculated the classification scores using features calculated from scalp measurements, and features calculated from reconstructed sources. As mentioned in the **Source reconstruction** subsection, different numbers of features could be responsible for a significant difference in the ability to overfit, between the trained classifiers, in particular between a "scalp" classifier based on features calculated from 60 signals and a "source" classifier using more than 60 sources. Therefore, only the displayed classification scores for signals reconstructed on 60 dipoles were suitable for statistical comparison with scalp measurements (using the 60 electrode signals). Classification scores were also provided for classifiers trained on signals reconstructed on 500 dipoles, to illustrate the impact of the regularization coefficient of the minimum-norm inverse solution, and to better localize discriminating patterns of activity.

Coupling measurements were to be folded so that we could derive maps and project them onto a representation of the scalp or cortical surface. To this aim, we designed the following procedure that applies both to T-scores and SVM weights: we ranked the set of values (one value per coupling); then for each source (or source/frequency band pair, or electrode, etc.), we calculated the average rank of all the couplings involving this source.

We hypothesized that multivariate maps could have some advantages over the corresponding univariate maps. In particular, using the 500 reconstructed sources, we compared how variable could be multivariate and univariate maps with respect to the regularization coefficient of the minimum-norm inverse solution. For coupling measurements, we implemented this comparison using the following procedure: we calculated one average map per value of the regularization coefficient and per type of maps (univariate/multivariate); then we calculated the standard deviation of each weight or source across all of the investigated values of the regularization coefficient; finally, we tested whether the distribution of "multivariate" standard deviations lay above the distribution of "univariate" standard deviations, using a one-sided Wilcoxon matched-pairs test. For mean amplitude, the maps to be compared were raw T-scores or SVM weights maps, therefore, we adapted the previous procedure dividing each map by its maximum. We performed this procedure on simulation data only.

2.5. *Simulations*

The MRI and electrode locations were obtained from one individual randomly chosen from those taking part in the real experiment. This anatomical data was used to construct simulation data which was processed in a similar way to real data. A 10 014-vertex mesh model (see Fig. 2) of the white matter/cortical gray matter interface was used to generate data, whereas a 504-vertex mesh of the same interface was used to reconstruct source activities from scalp signals. A high resolution mesh as the original model was expected to prevent "inverse crime" [Kaipio & Somersalo, 2005]. An inverse crime occurs every time some parameters (e.g. the source spatial distribution) are assumed to be perfectly known, which is never the case in reality, resulting in a bias in the study.

Two seeds on the cortical surface were selected and an area of active dipoles was defined around each selected seed. Two signals were generated and assigned to the dipoles of each area, such that the signal amplitude was maximal at the central source and decreased, following a Gaussian curve, along the paths of each area mesh.

The two signals were generated as random Gaussian processes (Mersenne Twister algorithm)

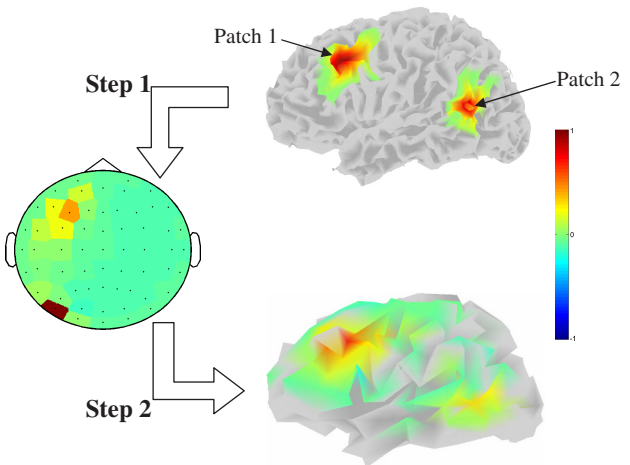

Fig. 2. Simulation model with two patches of activity related to two virtual mental states. Simulated cortical data are projected onto the scalp (step 1) and common processing methods are then used on either scalp data or source data after reconstruction (step 2). Grey areas on cortical maps display no activity and correspond simply to *gyri* and *sulci*.

and filtered through a band-pass filter centered around $f = 10\,\text{Hz}$ with transfer function:

$$H(z) = \frac{1 - z^{-1}}{1 - 2r\cos\left(\frac{2\pi f}{F_s}\right) z^{-1} + r^2 z^{-2}}$$

with $r = 0.95$, and F_s the sampling frequency.

For every second, the two signals had a feature in common. The duration of this pattern depended on the class of the corresponding window: 100 ms for class 1 (10% of total time) and 500 ms for class 2 (50% of total time). Thus, the differences between the two classes were only quantitative, and the classification task was difficult, so that the mean classification score never reached the maximal value of 100% correct classifications.

We designed two simulation experiments as illustrated in Fig. 3. In **simulation 1** (or amplitude-modulation condition), the amplitude of the two signals increased by 20% for 100 ms in every second in class 1 and 500 ms in every second in class 2. **Simulation 2** (or synchronization-modulation condition) corresponded to a situation in which the two signals were completely synchronized at the beginning of each second, for 100 ms in class 1 and 500 ms in class 2.

We expected mean amplitude to yield higher classification rates for validation data corresponding to **simulation 1**, and coupling measurements to perform better in **simulation 2**.

Gaussian random noise was then added to each source, such that all sources had similar energy levels in the 7–13 Hz band during each window. This noise was intended to simulate background

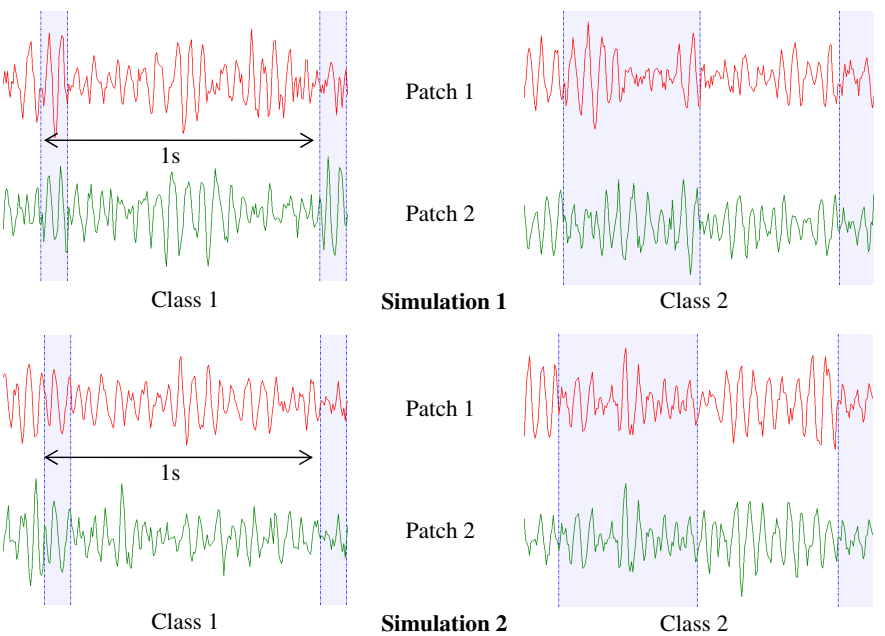

Fig. 3. Example of noise-free original signals for each patch. There is a feature common to both signals (top: **simulation 1**, concomitant increase in amplitude but no phase synchronization; bottom: **simulation 2**, complete synchronization) from the beginning of each second for a duration depending on the class of the 20 s time window. Left: example of original signals for the first class (100 ms common feature — 10% of time — in blue); right: example of original signals for the second class (500 ms common feature — 50% of time — in blue).

cortical activity not related to the mental state of interest.

The generated sources were projected onto the scalp and filtered in the 7–13 Hz frequency band, as this band was the only one used for the analysis of simulated data. By contrast to the procedure for real data, no average referencing was carried out at this step, as such referencing might have spread a mixture of the two informative signals everywhere.

We generated 20 data sets in this way for each simulation, to ensure that our statistical estimates were robust.

Statistical maps and classification rates were thus calculated as mean estimates over the 20 simulation data sets. Classification results were obtained with the two original noise-free signals only, with scalp data, with reconstructed data on the selected 60 sources with regularization coefficients of 10%, and with reconstructed data on the selected 500 sources with regularization coefficients of 1%, 5%, 10%, 15% and 20%.

In **simulation 1**, we expected mean amplitude to discriminate successfully between the windows of the two classes whereas coupling measurements were expected to be inefficient. **Simulation 2** should lead to the opposite situation between mean amplitude and coupling measurements.

2.6. *Real data*

Most of the experimental setting has been described elsewhere [Besserve *et al.*, 2008]. The experiment consisted of a spatial compatibility task. Subjects were asked to indicate as correctly and rapidly as possible the direction (to the left or to the right — relevant feature) of a displayed arrow. That arrow was located either on the left or the right of the screen (irrelevant feature), inducing a Simon effect. We investigated the decrement of performance.

Data was processed in an asynchronous way, segmenting the signal into sliding 20 s time windows (10 s overlap). The performance of the subject was quantified as a scalar index across windows, corresponding to the mean normalized reaction time over trials for each window. Normalization was also refined, with correction for the effect of stimulus congruency, expected answer, and previous trials ($2 \times 2 \times 2$ conditions). We independently labeled high-performance (first quartile of the performance index) and low-performance windows (last quartile of the performance index) for each subject. The remaining windows were ignored. High-performance windows accounted for our first class, and low-performance windows formed the second class.

The experiment consisted of four 10-min runs, with the first run discarded to eliminate learning effects. EEG recording was performed with a BrainAmps system, with 60 EEG, and 1 vertical EOG channel.

T1 MRI (61 slices of 256×256 pixels) were acquired for all subjects immediately after the EEG recording session, so that electrode locations could be determined from the patches of gel, visible as white parcels in the MRI.

The preprocessing of real data involved correction for ocular movements by a PCA-based method [Wallstrom *et al.*, 2004] and rejection of segments contaminated by muscle activity. Those steps were more refined than in the previous study.

Signals from the 60 EEG electrodes were referenced with respect to a common average reference calculated from the signals from all but most peripheral electrodes (the outer crown, omitting occipital electrodes, such as Fp1, F7, T7, etc.), in order to prevent contamination of the signal by muscle activity. Indeed, muscle activity is known to increase with time-on-task and mental effort, and we aimed to study only the cognitive pattern of performance decrement and mental fatigue. Moreover, source models for the forward problem do not take into account extracranial electrical activities.

Source reconstruction was carried out by calculating forward and inverse operators from individual MRIs.

Mean amplitude, coherence and phase locking value were calculated for three frequency bands (theta: 3–7 Hz, alpha: 7–13 Hz, beta: 13–18 Hz) from both scalp measurements and reconstructed source signals, in order to compare the amount of discriminating information available at these two levels.

3. Results for Simulations

3.1. *Mean amplitude*

In **simulation 1** (amplitude-modulation condition), mean amplitudes, from the 500-source signals in the correct frequency band, helped to localize the two patches of activity, with a source topography that appeared to be easier to interpret than that obtained for the scalp (Fig. 4).

The experimental setting illustrated in Fig. 4 led to more than 75% correct classifications, when

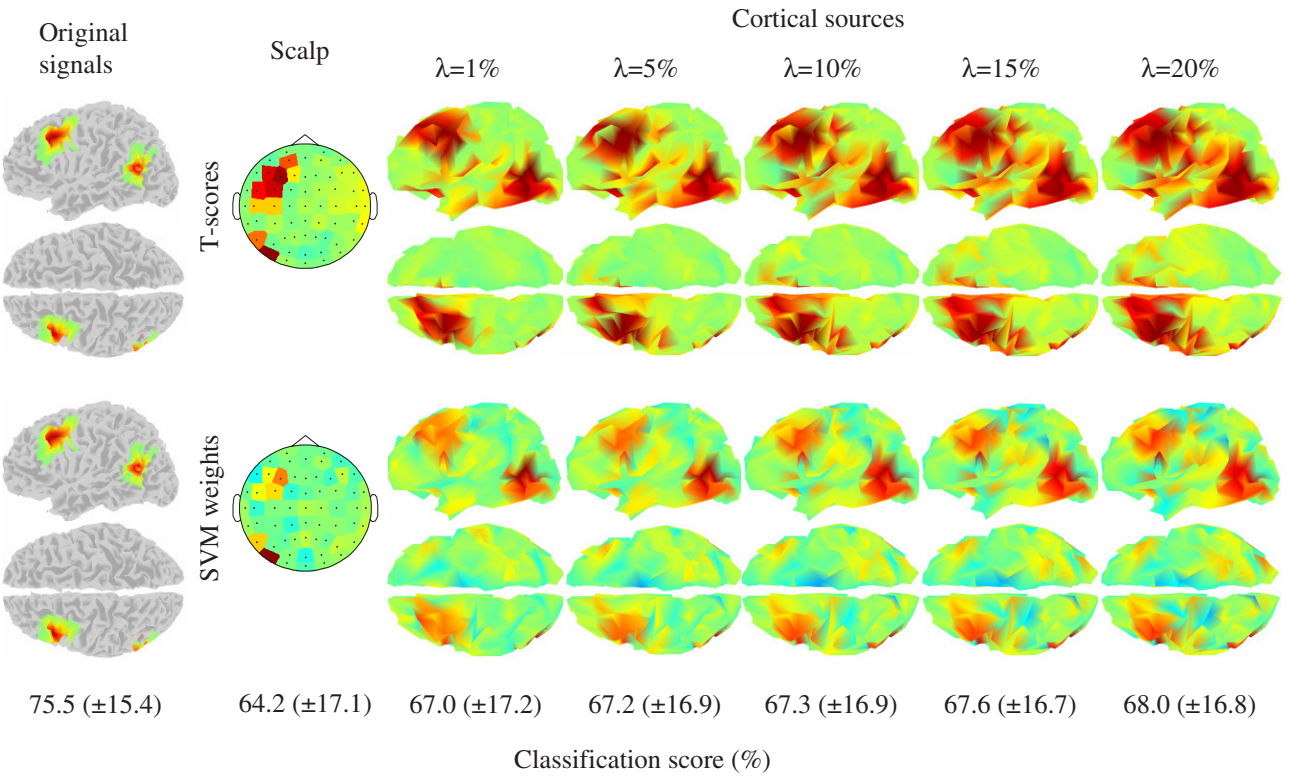

Fig. 4. **Simulation 1** — Classification scores (bottom row, in % correct classification with corresponding standard deviation in parentheses, chance level: 50%) and topographies (top: T-scores, bottom: SVM weights) for discrimination of amplitude modulations based on mean amplitudes at different levels, (from left to right): original signal pair, scalp, reconstructed sources with the regularization coefficient λ set to 1%, 5%, 10%, 15% and 20%, respectively.

classifying features that were calculated based on the two initial source signals without noise. This rate of correct classifications can be used as a sort of reference classification rate.

At the scalp level, original patches were identifiable, but the classification rate barely exceeded 64%. Classification rates were higher at the cortical level using 60 sources than at the scalp level, but this difference was actually not significant. At the 504-source level, the classification rates increased and were significantly higher than classification rates at the scalp level, whatever the features used.

Concerning mean amplitudes of the 504 reconstructed sources, the regularization parameter acted as a smoothing parameter. The smoothing effect was observed on both univariate (namely T-scores) and multivariate (SVM weights) statistical maps. The univariate depiction of a spatial discrimination pattern had the disadvantage of an overestimation for areas between the two patches. This might account for the artificial diffusion due to spatial smoothness of the inverse solution, and was surprisingly found to be more marked than for scalp data. SVM-based maps were significantly less sensitive to the regularization coefficient than were T-maps ($Z = 1.7$).

Classification scores based on mean amplitude increased significantly with the regularization coefficient (Friedman test).

In **simulation 2** (coupling-modulation condition), mean amplitudes, both of the original sources and at the scalp level, gave classification rates no higher than would be expected by chance (Fig. 5). However, at the source level, after reconstruction of 60 sources, the classification score was significantly higher. At the 504-source level, classification scores increased together with the regularization coefficient to reach levels sufficient for discrimination (about 61% with $\lambda = 20\%$). The sources capable of discriminating between the two classes seemed to be located half-way between the two patches.

3.2. *Coherence and phase locking value*

Figures 6 and 7 give classification scores for **simulation 1** (in the amplitude-modulation condition) and **simulation 2** (the synchronization-modulation

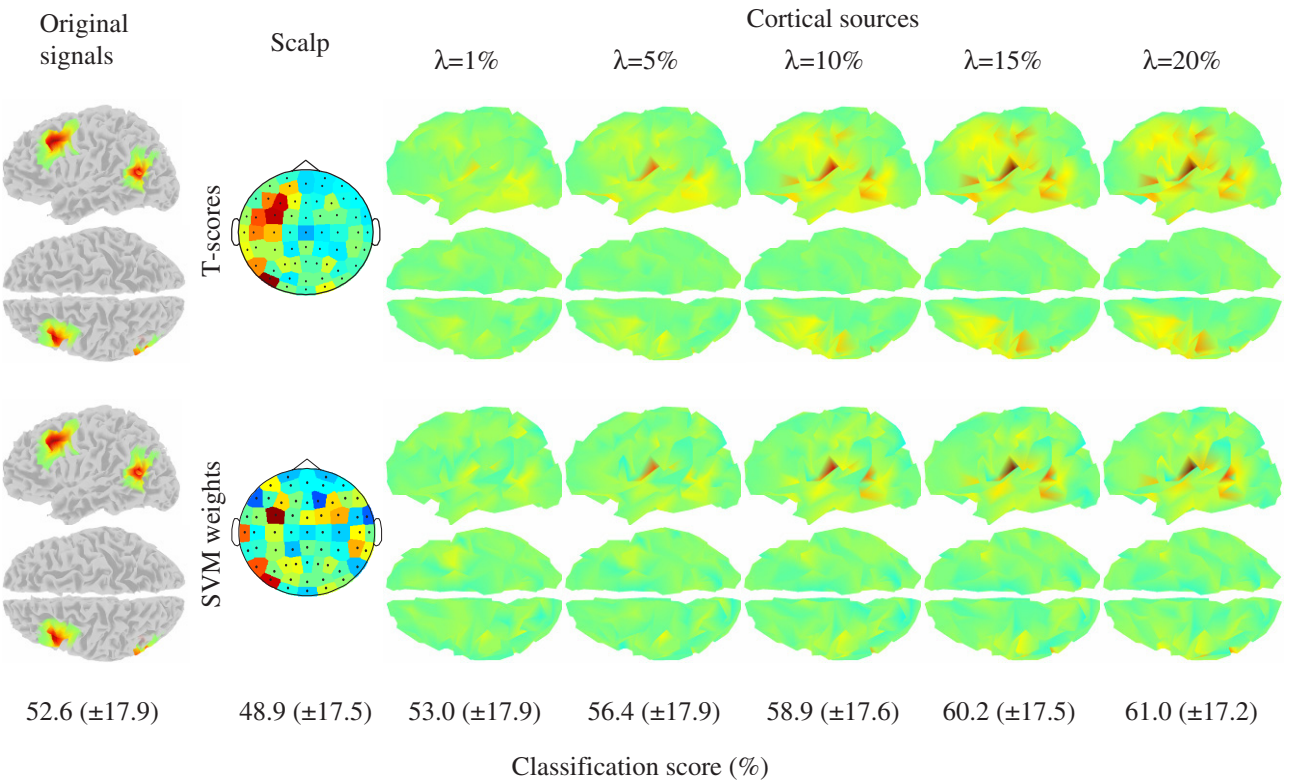

Fig. 5. **Simulation 2** — Classification scores (bottom row, in % correct classification, with corresponding standard deviation in parentheses, chance level: 50%) and topographies (top: T-scores, bottom: SVM weights) for discrimination of phase coupling modulations based on mean amplitudes at different levels, (from left to right) original signal pair, scalp, reconstructed sources with the regularization coefficient λ set to 1%, 5%, 10%, 15% and 20%, respectively.

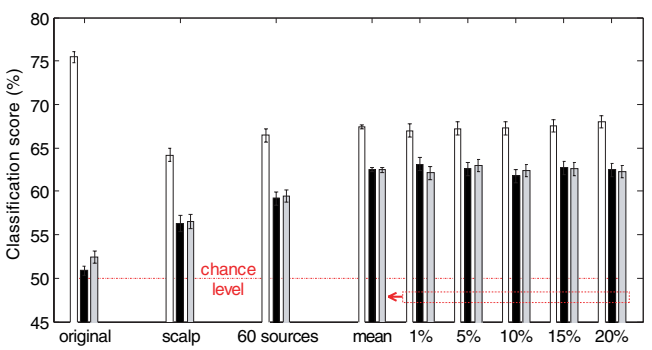

Fig. 6. Classification rates for **simulation 1**, based on mean amplitude (white bars), coherence (black bars) and phase locking value (grey bars) for different signals; *original*: pair of noise-free original signals; *scalp*: 60 electrode signals after forward projection; 60 *sources*: 60 reconstructed signals after forward and inverse projection, using $\lambda = 10\%$; *mean*, $1\%, 5\%, 10\%, 15\%$ and 20%: 504 source signals after forward and inverse projections with different values for λ. In particular, the group of columns labeled *mean* gives classification scores averaged across the five sets of results with the 504 reconstructed sources. The scores using 60 sources were compared with the scores obtained from scalp measurements. A star indicates that reconstructed source scores were higher than scalp measurement scores (200 samples).

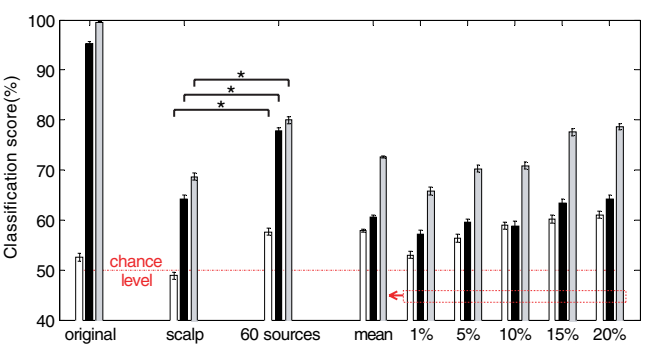

Fig. 7. Classification rates for **simulation 2**, based on mean amplitude (white bars), coherence (black bars) and phase locking value (grey bars) for different signals; *original*: pair of noise-free original signal; *scalp*: 60 electrode signals after forward projection; 60 *sources*: 60 reconstructed signals after forward and inverse projection, using $\lambda = 10\%$; *mean*, 1%, 5%, 10%, 15% and 20%: 504 source signals after forward and inverse projections with different values for λ. In particular, the group of columns labeled *mean* gives classification scores averaged across the five sets of results with the 504 reconstructed sources. The scores using 60 sources were compared with the scores obtained from scalp measurements. A star indicates that reconstructed source scores were higher than scalp measurement scores (200 samples).

condition), at the same spatial levels as above, using separate coherence and phase locking value. The results obtained with mean amplitudes are given to facilitate comparison.

Figure 6 relates to **simulation 1**. As expected, mean amplitude performed better than coherence or phase locking value for this simulation. When discriminating between the two simulated states using the two original signals, mean amplitude had a much higher discriminating power than coupling measurements. This base setting was designed to test whether, in this condition, mean amplitude would always give better results than coupling measurements, and to determine whether coupling measurements could be significantly more useful than chance alone.

Our results differed from our expectation: the classification rates for coherence and phase locking value were higher than chance alone, both at the scalp level and on reconstructed sources. Moreover, the classification rates increased together with the number of sources, revealing the presence of artificial couplings due to the spatial mixing of the original activities.

Figure 6 relates to **simulation 2**. Phase locking value appeared to be the most discriminating feature for coupling modulations, although coherence also seemed to give good results. The coherences calculated from scalp data (or better, from the 60 sources) outperformed those obtained after reconstruction of 504 sources, whatever the smoothing intensity used.

Classification rates for both coupling measures were higher with the 60 reconstructed sources than with scalp measurements. At the 504-source level, classification scores also increased together with the regularization coefficient.

Both coherence- and PLV-based SVM-derived maps, for the 504 sources, were more stable across the investigated values of the regularization coefficient than respectively coherence- and PLV-based T-score-derived maps (respectively $Z = 18.0$ and $Z = 16.3$).

We also investigated whether coherence or phase locking value could localize the expected pairs of sources in **simulation 2**, by defining a set of correct pairs of sources. The two patches were translated from the original 10 014-vertex mesh to the 504-vertex mesh used for the reconstruction of cortical electrical activities. All possible pairings between the two new regions of interest were labeled as correct (procedure illustrated in Fig. 8).

Links can be ranked on the basis of T-scores or SVM weights calculated from coherence or phase locking value. The capacity of a measurement to localize the discrimination pattern was quantified in a simple manner, by retrieving the rank of the best-ranked pair among correct pair and averaging this rank over data sets. This approach should be suitable for use in real situations, in which we often rely on the best-ranked link to infer the location of two distant coupled areas when it is hypothesized that there are at least two such areas. We compared the results obtained for coherence and phase locking value, based on both T-scores and SVM weights (Fig. 8). Phase locking value outperformed coherence in situations in which we expected to localize correctly a pair of distant areas intermittently displaying complete synchronization (see stars in Fig. 9). The best-ranked link was often found to be the correct one. Interestingly, PLV-based best ranks gave more reliable localization for smaller regularization coefficients. A positive relationship was observed (significance for both T-scores and SVM

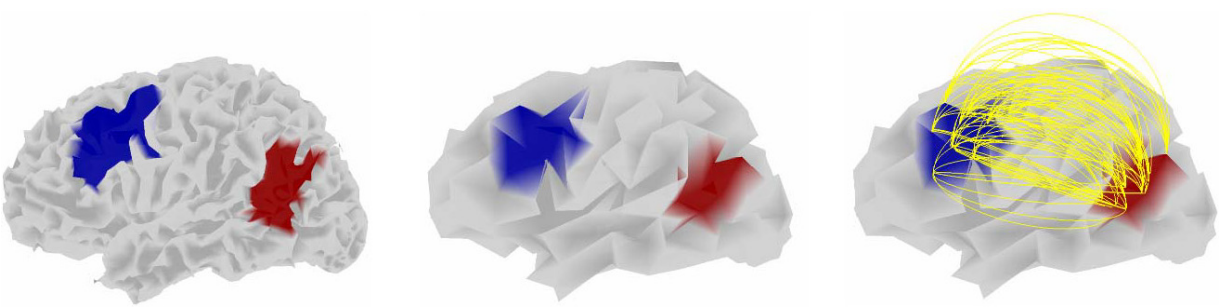

Fig. 8. The actual coupled pairs of sources (in yellow) in **simulation 2** were defined based on the original high-resolution mesh (left) and the two regions of interest (in blue and red), and adaptation to the source model (middle) for reconstruction by application of a simple nearest neighbor rule, defining the new regions of interest and then the set of correct links (90 of 1 26 756 links; right).

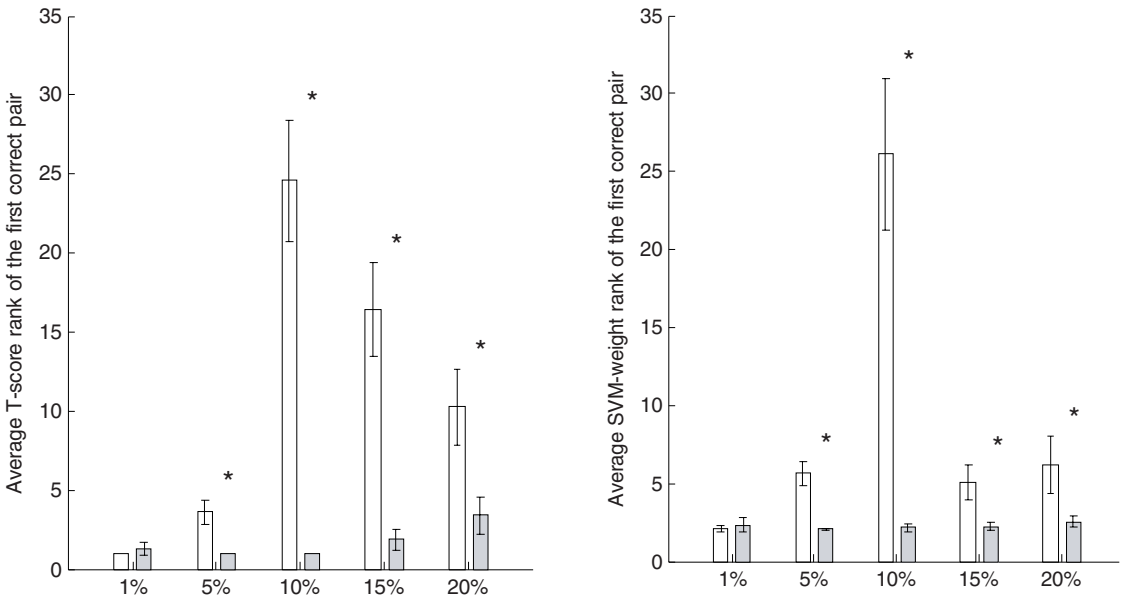

Fig. 9. **Simulation 2** — Assessment of the localizing power of coupling measurements. Rank of the best-ranked correct pair, averaged over simulation data sets; ranks are calculated from either T-scores (left) or SVM weights (right), themselves computed either from coherence (white) and phase locking value (grey) on cortical activities reconstructed with different regularization coefficient values. Confidence intervals relate here only to the variance across simulation data sets. Lower values indicate a better performance at localizing. A star indicates that PLV ranks were significantly lower than those obtained from coherence.

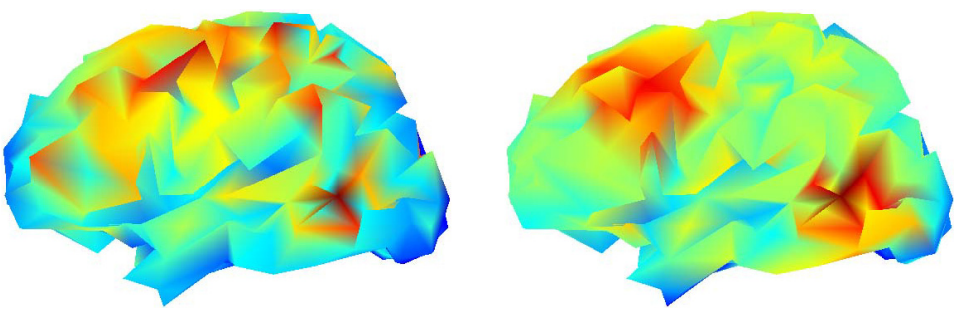

Fig. 10. Average rank maps using SVM weights of coherence (left) and PLV (right) from **simulation 2** data reconstructed with $\lambda = 1\%$. The value at each vertex is the average rank of adjacent links — links that share the specified vertex as one of their ends. Red indicates low average ranks and blue indicates high average ranks. Vertices with lower values had stronger connections with the other vertices.

weights; Friedman tests), with lower ranks associated with lower regularization coefficients, while classification scores showed the opposite pattern, with higher scores obtained for higher regularization coefficient values.

A cortical map (Fig. 10) was estimated by calculating, for each source, the average rank of links having this source as a common end. The top rank is coded as 1 (the lower the rank, the better).

4. Application to Real EEG Data

The Simon task experiment was designed to study mental fatigue as a performance decrement. We thus aimed to classify time windows as a function of mean normalized reaction time: class 1 (short reaction time, high performance) or class 2 (long reaction time, poor performance), with intermediate values ignored.

We show mean results from 12 subjects. Figure 11 summarizes classification scores for mean amplitude (white), coherence (black) and phase locking value (grey). The classification scores spanned an interval of about four points. This shows that there were few differences between types of features and observation levels.

At the scalp level, coherence gave slightly better results than mean amplitude (not significant)

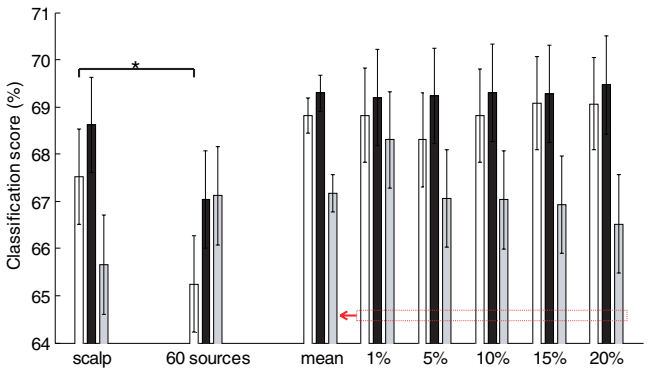

Fig. 11. Mean across 12 subjects (and 100 cross-validation estimates) of classification scores for every feature type (white: mean amplitude, black: coherence, grey: phase locking value) for scalp and reconstructed data as 60 sources (λ = 10%) and 500 sources (λ = 1%, 5%, 10%, 15% and 20%). 95% confidence intervals were calculated from the corresponding total variance, both across subjects and across validation iterations for each subject. Mean classification scores are given for reconstructed data on 500 sources (third group of bars). Classification scores for 60 sources are compared with classification scores for scalp measurements (1200 samples).

and classification scores using phase locking value were significantly above the chance level (one-sided Wilcoxon matched-pairs test against 50%). Physiological phenomena may have been more related to local patterns of activity in this case. However, as phase locking value did extract some information and coherence slightly outperformed mean amplitude, distant pair-wise activity patterns nonetheless contribute to the classification of performance levels.

Classification results for mean amplitude and coherence calculated from 60 reconstructed sources decreased. This decrease was even significant for mean amplitude. However, classification scores using PLV increased. Reconstructing activities on 500 sources led to higher classification rates. Average classification scores using PLV were significantly higher for 500 sources than for scalp measurements.

Topographical maps were generated (Fig. 12), showing only the theta band with λ = 1%, to save space. Coherence gave the highest classification scores; therefore the related map may be more relevant. A bi-lateralized patch appeared in the superior frontal cortices. Given the well-known difficulties of the minimum norm inverse method to correctly estimate deep sources and the lack of sensitivity of EEG electrodes to tangent-oriented dipoles, the related discriminating activity could originate from a superior area of the anterior cingulate cortices, which are known to play a role in response selection [Devinsky *et al.*, 1995] and conflict monitoring [van Veen & Carter, 2002]. Moreover, the right-lateralized pattern was quite similar to the so-called dorsal frontoparietal attentional network for salience detection [Corbetta & Shulman, 2002]. This latter pattern was common to all of the three types of features, but there were clear differences in the location of other implicated areas between the three types of features.

5. Discussion

We investigated the contribution of source reconstruction in several respects. One of our mere expectations was that the classification scores using features calculated from reconstructed sources would be greater than those derived from original scalp signal. We observed such an increase with simulations, but this was by far less clear with our real data. Indeed, concerning real data and reconstruction of 60 sources, we observed decreases with mean amplitude and coherence, which had given the highest scores at the scalp level. However, we noted two interesting points.

First, distant synchronies seemed to carry information on the two mental states to be discriminated. Despite the fact that informative patterns may be mainly local, classification scores using PLV were higher than the level of chance and classification scores using coherence were slightly higher than those using mean amplitude.

Second, source reconstruction could have actually helped at unveiling distant synchronies. Coherence is expected to reach a compromise between phase locking value, which extracts information from phase, and mean amplitude, which is related to amplitude. Here, we could hypothesize that coherence made profit from both components of the signal. We believe that the decrease in classification score for the coherence between the scalp level and the 60 reconstructed sources reflected mostly significant decrease in mean amplitude. Therefore, source reconstruction could favor distant synchronies at the expense of local phenomena. Another hypothesis is related to the contribution of electrical activities that current forward models had not yet taken into account. These could be muscle activity and deep cerebral activities. Both were likely to be involved in this experiment, as discussed again below. We believe that PLV is not much sensitive

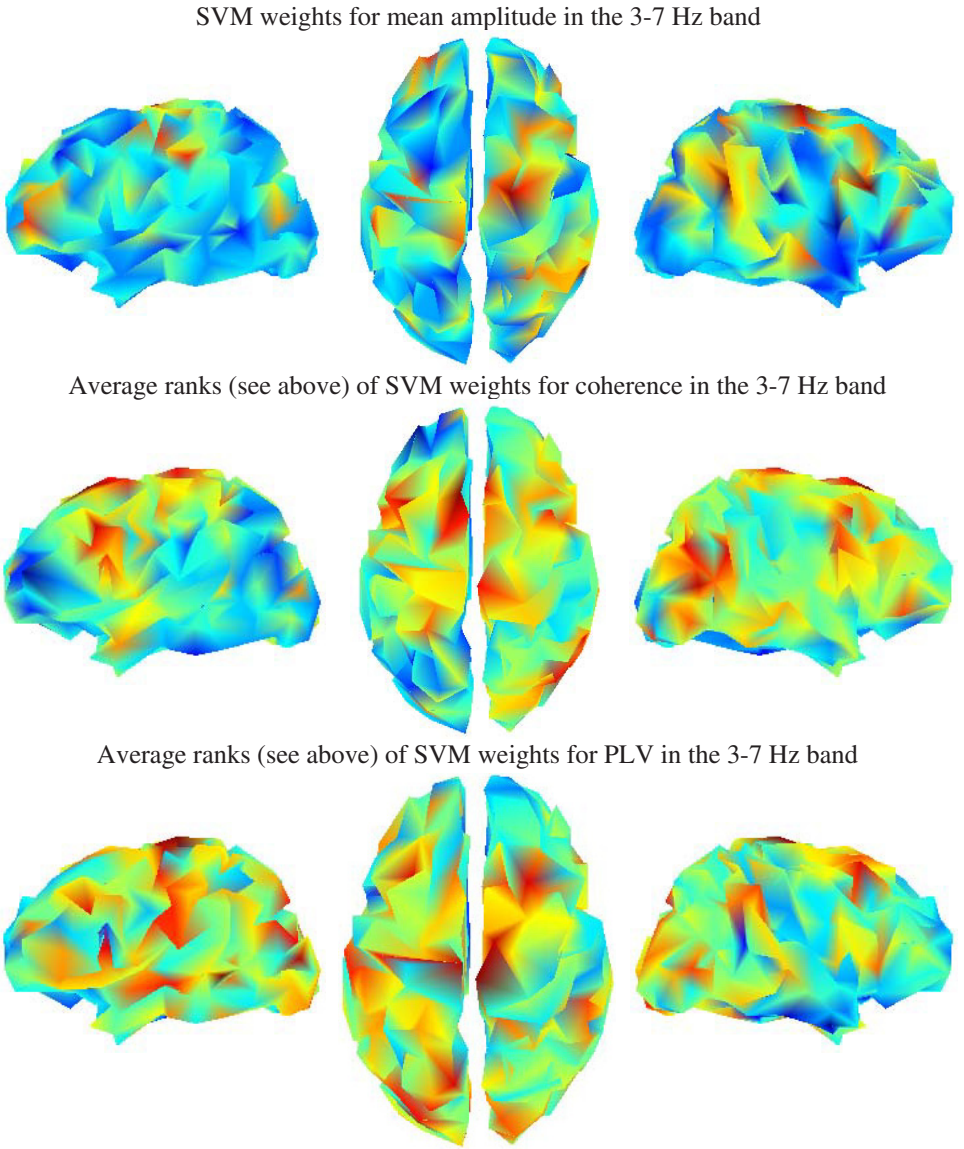

Fig. 12. Discrimination maps generated for activities reconstructed with $\lambda = 1\%$ in the 3–7 Hz band, with left, top and right views (from the left to the right); all maps are average estimates for the 12 subjects. Topmost pictures correspond to SVM weights for mean amplitude. Middle row pictures correspond to average ranks calculated based on SVM weights (as already described) for coherence. Bottom pictures correspond to average ranks calculated based on SVM weights for phase locking value. Here, averaging across subjects was made possible by matching source meshes with a reference mesh, and some defaults may have occurred. Precise values corresponding to colors do not help; the key point is that red indicates sources that affect the most the discrimination (high absolute SVM weight for mean amplitude-related map and low average rank for coherence and PLV-related maps), and blue indicates least discriminating sources.

to muscular noise, for example, because phase relationship is the hallmark of complex dynamics, and electrical activities from pericranial muscles might be less complex that cerebral activities. On the contrary, features based on signal amplitude are likely to catch information related to muscular noise. Hence, PLV could be more selective of cortical activities. Therefore, distant synchrony appears to be a useful concept. Coupled activities should therefore be investigated further.

Our results on simulations seem to support the notion that source reconstruction could refine signals that originated from the cerebral cortex, as higher classification scores were obtained for reconstructed sources than for scalp measurements.

Previous studies on real data showed small increases of (multivariate) classification scores when source activities were reconstructed and quantified with local synchrony measurements (power, mean amplitude, etc; [Lotte *et al.*, 2007; Noirhomme

et al., 2008]). In simulations, this result is thus confirmed and extrapolated to distant synchrony measurements.

We would also like to comment on the widespread inverse method known as sloreta (standardized low-resolution brain electromagnetic tomography [Pascual-Marqui, 2002]), and other standardized instantaneous inverse methods such as that described by Dale, with a constant noise covariance matrix [Dale & Sereno, 1993]. As these methods differ from the minimum-norm inverse method by a constant scaling of each source signal, identical classification results were obtained. Mean amplitude is linearly related to the scaling used in these methods, and coherence and phase locking value are unbiased. The differences were made irrelevant even with mean amplitude, as we used a linear classifier.

At the 500-source level, classification scores often increased when increasing the smoothness of the solution, whatever the feature used. This may be partly explained by the use of a multivariate classification method.

Indeed, discriminating activities might spread in larger patches with increasing regularization. The SVM, like other classification tools, makes use of the fact that barely detectable physiological phenomenon at a given location also occur in many other neighboring locations, and the cumulative discriminative power of all sources is potentially sufficient to separate two states with greater efficiency than chance alone.

Thus, the influence of the regularization (or smoothing) parameter may result from originally isolated activity being more widely distributed when the parameter has a high value, with more sources appearing to be informative in their activity. At low or moderate signal-to-noise ratios, these sources may not be completely redundant, depending on the instantaneous noise levels associated with each source.

For real data, some parameters could be fitted on a subject-by-subject basis, especially if getting higher classification scores is considered the sole objective. This is particularly true for the regularization coefficient of the minimum-norm inverse matrix, as suggested by the high level of intersubject variability in our experiment. For example, Noirhomme *et al.* [2008] used a cross-validation process to select the value of this regularization coefficient.

We did not optimize for this parameter using 60 sources, because the variability was much lower than when using 500 sources, and the comparison between the "scalp" and "source" classifiers would have been unfair if we had optimized for a parameter at the source level without having optimized for a similar parameter at the scalp level. The same explanation may also hold for the validity of the comparison of classification scores between the scalp and massively distributed source levels. Again, optimizing for different numbers of parameters might lead to a variable inclination for overfitting.

Moreover, as suggested by our simulated data, the features selected may play an important role in the processing chain. Coherence and phase locking value are more suitable to data in which distant synchronization patterns occur and are informative, whereas mean amplitude performs better when amplitude modulation is the key phenomenon.

As already mentioned, coherence is expected to reach a compromise between phase locking value, which extracts information from phase, and mean amplitude, which is related to amplitude. However, even phase locking value and mean amplitude may also provide similar information because of their respective indirect relationship to amplitude and phase. Indeed, phase alters amplitude and *vice versa*, as clearly shown by our simulations.

The spatial mixing of electrical activities from the brain to the sensors, and the blurring observed in reconstructed sources due to spatial transformation from scalp measurements to sources, and the presence of uninformative signal or noise, may result in misleading patterns. As shown, mean amplitude can be used to discriminate between purely phase-related dynamic states, because the amplitude of a linear combination of signals depends partly on the phase difference between these signals. Similarly, coupling measurements, even if related to phase only, may respond to simultaneous changes in amplitude.

In this study, we assessed the complete processing chain. Consequently, the observed effect associated to a given parameter may be made possible by another component in the chain and the impact of this other component is not always clear.

The most appropriate choice for the regularization parameter is not straightforward and, considering real data, clearly depended on the data for the concerned subject. The regularization parameter may depend on the sparseness of the informative patterns to be extracted, and the quality of source reconstruction (spatial fidelity, inverse method and assumptions concerning signal-to-noise ratio).

Most BCI studies reported higher classification scores than what we obtained. However, the classes defined in these studies always correspond to two different tasks, while the classes we designed did not qualitatively differ from each other. Errors rarely occurred, considering the whole session, and were not taken into account in our performance measurement, therefore errors could not account for any characteristic mental state in the poor-performance class.

The classification scores obtained on the basis of mean amplitude and phase locking value on the scalp presented here are slightly lower than those in our previous study [Besserve *et al.*, 2008]. This difference may result mainly from improvements in some preprocessing steps, especially the correction for ocular artifacts and the rejection of segments contaminated by muscle activity. Indeed, such extra-cerebral physiological variables have been correlated with mental fatigue or time-on-task [van Boxtel & Jessurun, 1993; Waterink & van Boxtel, 1994; Stern *et al.*, 1994]. Current forward models for source reconstruction take into account only electrical activity from the cerebral cortex. Models for other areas of grey matter in the human brain are emerging. However, no efforts have yet been made to take into account extracranial electrical sources, such as the facial and pericranial muscles and ocular globes, using assumptions concerning spatial location and spectral profile. This could be an explanation for subdued improvement of source reconstruction when classifying performance levels in our EEG experiment. In particular, muscle activity might be discriminating because of its link to mental effort, and a model that would ignore it may underestimate the contribution of this extra physiological variable.

In addition, the conventional features studied by EEG are based on rhythms and other time series patterns, as EEG displays a high level of temporal resolution. This suggests that dynamic inverse solutions might improve the performance of our method [Baillet & Garnero, 1997]. The inverse solution used in this study was the simplest available instantaneous solution.

Many studies in the BCI field are based on classification methods. Nevertheless, few BCI applications with multivariate classification methods made use of massively distributed source reconstruction methods at that time. MRIs are expensive to get and one could fear that source reconstruction would increase the amount of computation. Noirhomme *et al.* [2008] explained why source reconstruction does not necessarily increase computational burden and could be suitable for real-time applications, such as BCI. As long as convenient inverse methods are used, the projection of scalp measurements to sources is straight forward. The number of features derived from reconstructed activities should nevertheless grow at a similar rate to the number of sources. Coupling measurements are also suitable for real-time applications as they can be calculated rapidly, but at the expense of the number of sources, because of memory space limitations.

The use of features from reconstructed activities may increase in popularity as many studies suggested that it may be easier to interpret EEG signals if scalp measurements are translated into activities of cortical regions and coupling measurements between such regions are studied [ten Caat *et al.*, 2008; Astolfi *et al.*, 2007]. This may lead to a growing need for methods for the automatic segmentation of raw MRI to cortical areas and new standard functional topographies, such as that described by Tzourio-Mazoyer *et al.* [2002].

The procedure we described in this article is not to be considered as readily suitable for BCI application, because we trained classifiers on a subject-by-subject basis, using data from one session per subject. To expect a classifier to be stable enough over time as required by every application, one should test this classifier using the data of another session, another day for example, with the same subject. However, the described procedure is a useful framework to be adapted.

One of our aims was the detection of functional cortical networks. As we were interested in distant pair-wise measurements for this purpose, we developed a method for displaying an activity map for these features. This approach seemed to be successful for identifying key source locations for discrimination. Our results in the simulations suggested that phase locking value is suitable for the phenomenon of interest. More work is required to determine the robustness of this ability to discriminate when diffusion increases and informative coupling patterns become more randomly distributed across the cerebral cortex. Structural measurements, such as minimum path length or clustering coefficient, could also be used, for quantified network, so long as this does not increase overall computational costs.

Discriminating patterns seem to be better localized when extracted by multivariate methods,

such as SVM coefficients, as opposed to univariate scoring methods. To our knowledge, the suitability of such linear SVM coefficients for determination of the weight of input variables has not yet been demonstrated, but the notion has already been proposed (see for example [Mourão-Miranda *et al.*, 2005]). The weights of the trained SVM are the coefficients of a vector defining the axis of optimal separation between the two classes. When each feature for classification was first standardized, the weights can be compared to each other. The feature with the highest absolute weight is likely to be the most discriminating one. Moreover, we showed on simulation data that cortical maps derived from these weights were less sensitive to the value of the regularization coefficient. If one needs to save computation time and select an arbitrary value for this coefficient, or does not have any suitable procedure to select a value, SVM-based maps may be preferred. Therefore, linear SVM could be considered as a useful tool to derive SPM-like maps.

6. Conclusion

In this study, simulation data and real EEG data corresponded to various cerebral states differing in terms of type of informative activity (local or distant synchronization patterns) and complexity of informative patterns (clustered or sparse, low or high signal-to-noise ratio, etc). Simulation data was used to study the effect of the type of informative activity, whereas results from real EEG data showed some limitations of such simulations for studying mental state-related activity patterns.

Results from simulations suggested that source reconstruction may be useful when attempting to discriminate between two cerebral states using a suitable feature type, but care must be taken when interpreting results, due to the possible localization bias that may occur when focusing on a sole type of informative activity, such as local amplitude-related activity or distant phase-related activity. Indeed, once recorded signals are mixed in an unavoidably imperfect way to reconstruct the cortical electrical activity, phase-amplitude relationships may result in an above-than-chance discrimination, but the discriminating pattern may be partly misleading when the aim is to localize the pattern.

The proposed method involving the display of a full map based on a complete set of pair-wise quantifications seems to be useful, as we were able to locate the patches of interest in the second simulation. In addition, linear SVM provides interesting discrimination maps.

In our simulations, at the cortical level with increased spatial resolution, higher regularization parameters were associated with better classifications, possibly because informative pattern set for the simulation was spatially coarse and that the classifier was able to aggregate information across space. However, this strategy is clearly limited for real EEG applications, such as the classification of performance levels. In such cases, the regularization parameter should be determined on a subject-by-subject basis. With a close-to-optimality value, source reconstruction seems to be useful for discriminating between mental states with variable patterns of coupled activities. In our real EEG experiment, coherence gave the highest classification scores at the scalp and massively distributed source levels, and phase locking value gave results better than would be expected by chance, suggesting that performance-related mental states did yield phase information. Results with real EEG data may give a good picture of the complexity of the brain dynamics, with both amplitude and phase-related modulations. Simulations appeared as special cases that might have been combined in the real EEG dynamics. Thus, distant synchrony and phase relationships are useful concepts here, and should be investigated more closely.

Acknowledgments

This work has been supported by the French *Délégation Générale pour l'Armement.*

References

Achard, S., Salvador, R., Whitcher, B., Suckling, J. & Bullmore, E. [2006] "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *J. Neurosci.* **26**, 63–72.

Astolfi, L., Cincotti, F., Mattia, D., Marciani, M., Baccala, L., de Vico Fallani, F., Salinari, S., Ursino, M., Zavaglia, M., Ding, L., Edgar, J., Miller, G., He, B. & Babiloni, F. [2007] "Comparison of different cortical connectivity estimators for high-resolution EEG recordings," *Hum. Brain Mapp.* **28**, 143–157.

Baillet, S. & Garnero, L. [1997] "A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem," *IEEE Trans. Biomed. Eng.* **44**, 374–385.

Baillet, S., Mosher, J. & Leahy, R. [2001] "Electromagnetic brain mapping," *IEEE Sign. Process. Mag.* **18**, 14–30.

Besserve, M., Philippe, M., Florence, G., Laurent, F., Garnero, L. & Martinerie, J. [2008] "Prediction of performance level during a cognitive task from ongoing EEG oscillatory activities," *Clin. Neurophysiol.* **119**, 897–908.

Burges, C. [1998] "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.* **2**, 121–167.

Congedo, M., Lotte, F. & Lécuyer, A. [2006] "Classification of movement intention by spatially filtered electromagnetic inverse solutions," *Phys. Med. Biol.* **51**, 1971–1989.

Corbetta, M. & Shulman, G. [2002] "Control of goal-directed and stimulus-driven attention in the brain," *Nature Rev. Neurosci.* **3**, 201–215.

Dale, A. & Sereno, M. [1993] "Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach," *J. Cogn. Neurosci.* **5**, 162–176.

De Vico Fallani, F., Astolfi, L., Cincotti, F., Mattia, D., Tocci, A., Grazia Marciani, M., Colosimo, A., Salinari, S., Gao, S., Cichocki, A. & Babiloni, F. [2007] "Extracting information from cortical connectivity patterns estimated from high resolution EEG recordings: A theoretical graph approach," *Brain Topogr.* **19**, 125–136.

Devinsky, O., Morrell, M. & Vogt, B. [1995] "Contributions of anterior cingulate cortex to behaviour," *Brain* **118**, 279–306.

Eguiluz, V., Chialvo, D., Cecchi, G., Baliki, M. & Apkarian, A. [2005] "Scale-free brain functional networks," *Phys. Rev. Lett.* **94**, 018102.

Franaszczuk, P., Blinowska, K. & Kowalczyk, M. [1985] "The application of parametric multichannel spectral estimates in the study of electrical brain activity," *Biol. Cybern.* **51**, 239–247.

Hämäläinen, M. & Sarvas, J. [1989] "Realistic conductor geometry model of the human head for interpretation of neuromagnetic data," *IEEE Trans. Biomed. Eng.* **36**, 165–171.

Huang, M., Mosher, J. & Leahy, R. [1999] "A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG," *Phys. Med. Biol.* **44**, 423–440.

Kaipio, J. & Somersalo, E. [2005] *Statistical and Computational Inverse Problems*, eds. Antman, S., Marsden, J. & Sirovich, L. (Springer, NY), p. 160.

Kamousi, B., Amini, A. & He, B. [2007] "Classification of motor imagery by means of cortical current density estimation and Von Neumann entropy," *J. Neural Engin.* **4**, 17–25.

Lachaux, J., Rodriguez, E., Martinerie, J. & Varela, F. [1999] "Measuring phase synchrony in brain signals," *Hum. Brain Mapp.* **8**, 194–208.

Lotte, F., Lécuyer, A. & Arnaldi, B. [2007] "FuRIA: A novel feature extraction algorithm for brain-computer interfaces using inverse models and fuzzy region of interest," *Proc. 3rd IEEE-EMBS Int. Conf. Neural Engineering*, pp. 175–178.

Mourão-Miranda, J., Bokde, A. L., Born, C., Hampel, H. & Stetter, M. [2005] "Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data," *NeuroImage* **28**, 980–995.

Noirhomme, Q., Kitney, R. & Macq, B. [2008] "Single-trial EEG source reconstruction for brain-computer interface," *IEEE Trans. Biomed. Eng.* **55**, 1592–1601.

Pascual-Marqui, R. [2002] "Standardized low resolution brain electromagnetic tomography (sLORETA): Technical details," *Meth. Find. Exp. Clin. Pharmacol.* **24**, 5–12.

Qin, L., Ding, L. & He, B. [2004] "Motor imagery classification by means of source analysis for brain-computer interface applications," *J. Neural Eng.* **1**, 135–141.

Rodriguez, E., George, N., Lachaux, J., Martinerie, J., Renault, B. & Varela, F. [1999] "Perception's shadow: Long-distance synchronization of human brain activity," *Nature* **397**, 430–433.

Schölkopf, B. & Smola, A. [2002] *Learning with Kernels* (MIT Press, Cambridge, USA).

Stern, J., Boyer, D. & Schroeder, D. [1994] "Blink rate: A possible measure of fatigue," *Hum. Factors* **36**, 285–297.

ten Caat, M., Lorist, M., Bezdan, E., Roerdink, J. & Maurits, N. [2008] "High-density EEG coherence analysis using functional units applied to mental fatigue," *J. Neurosci. Meth.* **171**, 271–278.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B. & Joliot, M. [2002] "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *NeuroImage* **15**, 273–289.

van Boxtel, A. & Jessurun, M. [1993] "Amplitude and bilateral coherency of facial and jaw-elevator EMG activity as an index of effort during a two-choice serial reaction task," *Psychophysiol.* **30**, 589–604.

van Veen, V. & Carter, C. S. [2002] "The anterior cingulate as a conflict monitor: fMRI and ERP studies," *Physiol. Behav.* **77**, 477–482.

Varela, F. [1995] "Resonant cell assemblies: A new approach to cognitive functions and neuronal synchrony," *Biol. Res.* **28**, 81–95.

Varela, F., Lachaux, J., Rodriguez, E. & Martinerie, J. [2001] "The brainweb: Phase synchronization

and large-scale integration," *Nat. Rev. Neurosci.* **2**, 229–239.

Vishwanathan, S., Smola, A. J. & Murty, M. N. [2003] "SimpleSVM," *Proc. Int. Conf. Machine Learning*, eds. Fawcett, T. & Mishra, N. (AAAI press).

Wallstrom, G., Kass, R., Miller, A., Cohn, J. & Fox, N. [2004] "Automatic correction of ocular artifacts in the EEG: A comparison of regression-based and component-based methods," *Int. J. Psychophysiol.* **53**, 105–119.

Waterink, W. & van Boxtel, A. [1994] "Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information processing task," *Biol. Psychol.* **37**, 183–198.

Wolpaw, J., Birbaumer, N., McFarland, D., Pfurtscheller, G. & Vaughan, T. [2002] "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.* **113**, 767–791.