Cogn Process (2013) 14:357–369
DOI 10.1007/s10339-013-0568-y

RESEARCH REPORT

Optimal spatial filtering for brain oscillatory activity using the Relevance Vector Machine

P. Belardinelli · A. Jalava · J. Gross · J. Kujala · R. Salmelin

Received: 11 February 2013 / Accepted: 8 May 2013 / Published online: 1 June 2013

Abstract Over the past decade, various techniques have been proposed for localization of cerebral sources of oscillatory activity on the basis of magnetoencephalography (MEG) or electroencephalography recordings. Beamformers in the frequency domain, in particular, have proved useful in this endeavor. However, the localization accuracy and efficacy of such spatial filters can be markedly limited by bias from correlation between cerebral sources and short duration of source activity, both essential issues in the localization of brain data. Here, we evaluate a method for frequency-domain localization of oscillatory neural activity based on the relevance vector machine (RVM). RVM is a Bayesian algorithm for learning sparse models from possibly overcomplete data sets. The performance of our frequency-domain RVM method (fdRVM) was compared with that of dynamic imaging of coherent sources (DICS), a frequency-domain spatial filter that employs a minimum variance adaptive beamformer (MVAB) approach. The methods were tested both on simulated and real data. Two types of simulated MEG data sets were generated, one with continuous source activity and the other with transiently active sources. The real data sets were from slow finger movements and resting state. Results from simulations show comparable performance for DICS and fdRVM at high signal-to-noise ratios and low correlation. At low SNR or in conditions of high correlation between sources, fdRVM performs markedly better. fdRVM was successful on real data as well, indicating salient focal activations in the sensorimotor area. The resulting high spatial resolution of fdRVM and its sensitivity to low-SNR transient signals could be particularly beneficial when mapping event-related changes of oscillatory activity.

Keywords RVM · Beamformer · Magnetoencephalography · MEG inverse problem · Cortical rhythms · Source localization · Bayesian approaches

Abbreviations

RVM	Relevance vector machine
fdRVM	Frequency-domain RVM
MEG	Magnetoencephalography
EEG	Electroencephalography
DICS	Dynamic imaging of coherent sources
ECD	Equivalent current dipole
CSD	Cross-spectral density
MNE	Minimum-norm estimate
MCE	Minimum-current estimate
MVAB	Minimum variance adaptive beamformer
FWHM	Full width at half maximum
SNR	Signal-to-noise ratio
BEM	Boundary element method
EM	Expectation-maximization

Introduction

Magnetoencephalography (MEG) and electroencephalography (EEG) allow detection and tracking of cortical

P. Belardinelli · A. Jalava · J. Kujala · R. Salmelin
O.V. Lounasmaa Laboratory, Brain Research Unit, Aalto University, Espoo, Finland

P. Belardinelli (✉)
Neurosurgery Department, Tübingen University Hospital, Otfried-Müller Str. 45, 72076 Tübingen, Germany
e-mail: paolo.belardinelli@gmail.com

J. Gross
Department of Psychology, Centre for Cognitive Neuroimaging, University of Glasgow, Glasgow, UK

oscillatory activity. Its modulation by stimuli and tasks (Hari and Salmelin 1997; Pfurtscheller and Lopes da Silva 1999) suggests that rhythmic activity has an important role in brain function. However, reliable localization of the brain areas that generate discernible oscillatory power in MEG/EEG is not a trivial task (Belardinelli et al. 2012). The solution of the so-called inverse problem, i.e., transforming from the sensor (or electrode) level to the neural sources, has been estimated in several ways. A conceptually simple approach is to model an active cortical patch as an equivalent current dipole (ECD), which represents the centre of an active area as well as the mean direction and strength of current flow in that area (Baillet et al. 2001). ECDs have been used to model sources of rhythmic activity, within selected frequency ranges and separately at each time instant (Liljeström et al. 2005; Salmelin and Hari 1994). While this method can be very powerful in presence of a limited number of sources, it includes user-dependent choices and benefits from expertise. With the minimum-norm estimate (MNE) approach, the measured MEG signals are assumed to be generated by an electric current distribution that has the lowest total power, or lowest total current in the case of the spatially sharper minimum-current estimate (MCE) (Jensen and Vanni 2002; Jerbi et al. 2007).

Differently from MNE, beamformers estimate source activity in one voxel while suppressing contribution from activity in the other voxels. This procedure is repeated for all voxels by means of data covariance, yielding a 3D map of neural activity (Vrba and Robinson 2001). Dynamic imaging of coherent sources (DICS) is a frequency-domain implementation of a minimum variance adaptive beamformer (MVAB) (Van Veen et al. 1997) that has been successfully applied on MEG data to map both oscillatory power and coherence (Gross et al. 2001, 2002; Hirschmann et al. 2011; Kujala et al. 2007; Liljeström et al. 2005; Mazaheri et al. 2009; Osipova et al. 2006). The DICS spatial filter uses the cross-spectral density (CSD) of the Fourier-transformed sensor data as the basis of data processing, instead of the data covariance employed in the time-domain beamformer.

Minimum variance adaptive beamformer (MVAB) assumes source activities to be uncorrelated. However, this condition is obviously not always satisfied in real brain activity. Indeed, correlation, when present, is an essential marker of functional connectivity. The quality of the MVAB estimate deteriorates when the sources are correlated or the number of samples is small. In addition, if more than two sources are simultaneously active, the bias due to a mutual correlation among source activities is not predictable at a theoretical level (Sekihara et al. 2002; Zoltowski 1988). Furthermore, a high level of coherence between sources in one frequency band can affect the beamformer time course reconstruction and source localization also in other frequency ranges where the time courses are not necessarily correlated (Dalal et al. 2007).

The correlation bias in beamformer algorithms has been addressed by means of various approaches. One possible solution is the use of different spatial filters (correlation biased and not biased) on the same data. Unfortunately, the interpretation of the resulting set of mapping outputs remains subjective (Belardinelli et al. 2007). The use of a priori information has been considered in the case of two active sources (Brookes et al. 2004). One possible solution with a more general application is the modeling of correlation between sources by means of variational Bayesian methods (Dalal et al. 2006). Such an approach facilitates accurate recovery of source locations and their time courses of activation. This can be an effective procedure but remains a two-step solution to the problem: Firstly, the interfering sources need to be modeled appropriately and, subsequently, a facilitated region suppression of interfering sources is applied.

Bayesian variational methods based on the Laplace approximation like greedy search (GS) (Friston et al. 2008a) or automatic relevance determination (ARD) (Friston et al. 2008b) can take correlation effects into account by means of information priors. However, the defaults in the SPM software (http://www.fil.ion.ucl.ac.uk/spm/) where the algorithms are implemented, currently consider exclusively correlated sources for symmetrical brain areas in the two brain hemispheres (Belardinelli et al. 2012).

Recently, the relevance vector machine (RVM), a Bayesian algorithm for learning sparse models from possibly over-complete data sets, has been employed to remove the bias due to correlation between sources. A new "unbiased" covariance is calculated as a basis for MVAB to reach from the sensor level to the brain space (Wipf et al. 2009, 2010). This procedure has been applied to estimate the covariance structure of source activity by means of a maximum likelihood iterative method.

Here, building on the work by Wipf and colleagues, we introduce a frequency-domain approach that employs RVM (fdRVM) to map the power of oscillatory brain activity without relying on the cross-spectral density as in the MVAB-based procedure (such as DICS). Our approach separates the processing of real and imaginary parts of the Fourier-transformed time series by means of two distinct parallel loops and subsequently combines the real and imaginary source power components into a single mapping. This has remarkable computational advantages that are elucidated in the "Methods" and "Discussion" sections. Furthermore, the iterative learning procedure based on Bayesian assumptions removes the undesirable effects of

coherence and provides a reliable and noise-resistant tool for mapping electrophysiological brain rhythms. Here, we tested the performance of fdRVM and compared it with that of DICS. Two types of simulated MEG data were considered, as well as two real MEG data sets (continuous finger movement, resting state). The first simulation featured oscillatory sources that remained active throughout the time window of interest ('steady-state'). The second one consisted of sources that were active only transiently within the analyzed time window ('transient'). Effects of low and high mutual coherence on source localization were evaluated at different signal-to-noise ratios (SNRs). The power mapping results on the simulated data were quantified using three parameters: (a) number of sources correctly detected, (b) full width at half maximum (FWHM) of the source estimates and (c) cumulative mislocalization of the sources.

Methods

Dynamic imaging of coherent sources (DICS)

Dynamic imaging of coherent sources (DICS) utilizes a MVAB beamformer in the frequency domain (Gross et al. 2001). A linear transformation based on the cross-spectral density (CSD) matrix filters the source activity in a given frequency band in a certain voxel (grid point) with unit gain, while suppressing contribution from the other voxels. By means of this filter, it is possible to calculate either power or coherence (with a reference signal) of sources at the frequency of interest.

For every voxel, two orthogonal unit current dipoles are considered. In a spherical conductor, such dipoles span the space containing all possible source orientations that can be detected with MEG. For the more realistic boundary element method (BEM) model of the brain, this approximation still holds with satisfactory results. The third eigenvalue, corresponding to the pseudo-radial vector, is markedly smaller than the ones corresponding to the tangential directions in practically every voxel within the brain.

The source cross-power estimates between the two dipole components at the i-th voxel is a 2 by 2 matrix with singular values λ_1 and λ_2. If $\lambda_1 \gg \lambda_2$, the source can be considered to have a fixed orientation. Otherwise, the power estimate can be obtained by computing the trace of the matrix (Gross et al. 2001).

DICS regularization factor was obtained by adjusting the trade-off between spatial sharpness and noise sensitivity (Gross and Ioannides 1999). For both DICS and fdRVM, normalized leadfields were employed.

fdRVM, a frequency-domain RVM approach

Bayesian inversion schemes have been extensively considered for source localization in MEG/EEG (Auranen et al. 2005; Friston et al. 2002; Nummenmaa et al. 2007; Phillips et al. 2005; Sato et al. 2004; Wipf and Nagarajan 2009). Basically, the aim of Bayesian methods is the covariance estimation of the sources $\mathbf{S}$ that best explains the measured data $\mathbf{M}$. Once the probability distribution of $\mathbf{S}_i$ is known, a meaningful parameter (its mean, for example) can be used as an estimate of the source activity on the i-th voxel. fdRVM accommodates real and imaginary series obtained from Fourier-transformed data in two separate routines and processes them following the steps of Wipf and colleagues (Wipf et al. 2010) with 'MacKay' updates (MacKay 1992).

The original aim of RVM is the prediction of future data by employing an analogous set of data from the past (Tipping 2001). For this purpose, a kernel function defining a set of basis functions is estimated. These basis functions, in turn, are supposed to generate the new (future) data. RVM, by means of a probabilistic framework, makes predictions based on the function

$$f(\mathbf{y}, \mathbf{w}) = \sum_{i=1}^{N} w_i K(\mathbf{y}, \mathbf{y}_i) + w_0 \quad (1)$$

where f is a function defined over the input space upon which predictions are based, $\mathbf{y}$ is the input data array and w_i are adjustable weights. K is a kernel function defining a set of basis functions for each data sample x_i. In general, RVM's approach to problem (1) is similar to that of a support vector machine (Scholkopf et al. 2000), with the basic difference that RVM adopts an entirely probabilistic approach. A prior over the model weights is introduced. Such a prior is regulated by a set of hyperparameters associated with the weight set. Upon convergence of the iterative procedure, posterior probability distributions of a large part of the weights show a prominent peak around zero. In this way, sparsity is achieved, and the training vectors associated with the non-zero vectors are labeled *relevance vectors* (Tipping 2004).

Rather than predicting future biomagnetic data, we propose a particular use of RVM to localize rhythmic activity in the brain without the bias due to correlation among sources. Two different routines separately handle real and imaginary parts of Fourier-transformed MEG/EEG data. The final power mapping is obtained as a combination of the real and imaginary RVM outputs for each voxel. Separate handling of the two Fourier components is permitted by the linearity of the localization algorithm and by the independence of real and imaginary series. A physical

interpretation of the real and imaginary series obtained from FFT-transformed signals is not straightforward. The Fourier transform employs the complex values for a compact representation of sinusoidal and cosinusoidal components. Both imaginary and real series contribute in a different and non-negligible way to the mapping of rhythmic source activity.

Our procedure rests on the following steps: The time-domain data of each sensor are divided in N_w partially overlapping windows (75 % overlap in our case). The samples in each window (2048 in our case) are Fourier transformed. We assume a model of the type $\mathbf{Y} = \mathbf{LS} + \boldsymbol{\varepsilon}$, where $\mathbf{Y} = [\mathbf{y}(1),\ldots, \mathbf{y}(N_w)]$ is our data array at one frequency of interest (each time window yields one value at a given frequency) with dimensions $[N_s * N_w]$, where N_s is the number of sensors. $\boldsymbol{S} = [\boldsymbol{s}(1),\ldots, \boldsymbol{s}(N_w)]$ are the unknown source activities with dimensions $[N_v * N_w]$, where N_v is the number of voxels within the brain. The elements of $\mathbf{S}$ can be considered as the weights $\boldsymbol{w}$ in the traditional RVM formulation. $\boldsymbol{\varepsilon}$ denotes the error due to noise and correlation bias and is assumed to have a Gaussian distribution $p(\boldsymbol{\varepsilon}|\mathbf{Y}) = \mathcal{N}(0, \Sigma_\varepsilon)$. $\mathbf{L}$ is the lead-field matrix of dimensions $[N_s * N_v]$. The rows of $\mathbf{L}$ are the kernel components of K in Eq. (1); for the i-th sensor, $\mathbf{L}_i(\mathbf{Y}) = K(\mathbf{Y}, \mathbf{y}_i)$, where $\mathbf{y}_i = [\mathrm{y}_i(1),\ldots, \mathrm{y}_i(N_w)]$ is the data array for one frequency of interest and every time window. The real (R) and the imaginary (I) part of our FFT data are processed by two independent RVM routines under the following assumptions: a likelihood model for $\mathbf{Y}$ which is a complex Gaussian (Tipping 2001), given fixed sources:

$$p(\mathbf{Y}_{R,I}|\mathbf{S}_{R,I}) = (2\pi\Sigma_{\varepsilon R,I})^{-N_S * N_w/2} \exp\left(-\frac{1}{2}(\Sigma_{\varepsilon R,I}^{-1}\|\mathbf{Y}_{R,I} - \mathbf{LS}_{R,I}\|_{\Im}^2\right) \quad (2)$$

where $\boldsymbol{\Sigma}_{\varepsilon\ \mathrm{R,I}}$ is the error variance and $\|\cdot\|_{\Im}$ is the Frobenius norm. As additional constraint to avoid over-fitting in the estimation of $\mathbf{S}$ from Eq. (2), a zero-mean Gaussian prior distribution over real and imaginary source components is assumed:

$$p(\mathbf{S}_{R,I}|\gamma_{R,I}) = \prod_{i=1}^{N} \mathcal{N}(\mathbf{S}_{R,I_i}|0, \gamma_{R,I_i}^{-1}) \quad (3)$$

where $\boldsymbol{\gamma}_{R,I} = \left[\gamma_{R,I_{1,\ldots,}}\gamma_{R,I_N}\right]$ is the set of unknown hyperparameters controlling the prior covariance of each row of $\mathbf{S}_{\mathrm{R,I}}$.

The columns of $\mathbf{S}$ are assumed having independent probability distributions.

The hyperparameters can be estimated from the data by means of a maximum likelihood optimization

$$\boldsymbol{\Gamma}_{R,I} = \arg\max_{\gamma_{R,I}} p(\boldsymbol{\gamma}_{R,I}|\mathbf{Y}_{R,I})$$

We know from the Bayes theorem that as follows:

$$p(\boldsymbol{\gamma}_{R,I}|\mathbf{Y}_{R,I}) \propto p(\mathbf{Y}_{R,I}|\boldsymbol{\gamma}_{R,I})p(\boldsymbol{\gamma}_{R,I})$$

To obtain the best estimation of $\boldsymbol{\gamma}$, it is possible to integrate out $\mathbf{S}$ and then maximize the likelihood:

$$p(\mathbf{Y}_{R,I}|\boldsymbol{\gamma}_{R,I}) = \int p(\mathbf{Y}_{R,I}|\mathbf{S}_{R,I})p(\mathbf{S}_{R,I}|\boldsymbol{\gamma}_{R,I})\mathrm{d}\mathbf{S}_{R,I} \quad (4)$$

For the assumptions adopted in (2) and (3)

$$p\left(\mathbf{Y}_{R,I}|\boldsymbol{\gamma}_{R,I}\right) = \left|2\pi\Sigma_{YR,I}\right|^{-N_S * N_w/2} \times \exp\left(-\mathrm{trace}\left(\mathbf{Y}_{R,I}^H \Sigma_{YR,I}^{-1}\mathbf{Y}_{R,I}\right)\right)$$

where

$$\Sigma_{\mathrm{Y}R,I} = \Sigma_{\varepsilon R,I} + \mathbf{L}\Sigma_{\mathrm{R},I}\mathbf{L}^{\mathrm{T}} \quad (6)$$

are the real and imaginary model covariances of the FFT signals, respectively.

The estimation of the hyperparameters γ is obtained by means of a Type II expectation–maximization (EM) algorithm which considers $\mathbf{S}$ as hidden data (MacKay 1992). The marginalization provides a regularizing mechanism that shrinks the majority of the elements of $\boldsymbol{\gamma}$ to zero. For a generic mth voxel, if $\gamma_m = 0$ then $\mathbf{S}_m = 0$. Here, we assume hyperparameters with a uniform flat distribution. This assumption implies that no voxel is considered as a privileged site for a source due to prior information (automatic relevance determination) (Tipping 2001). Such non-informative distribution can be easily changed if reliable priors are available.

Instead of obtaining the maximum likelihood estimate of $\boldsymbol{\gamma}_{\mathrm{R,I}}$ by maximizing the probability distribution $\mathrm{p}(\mathbf{Y}_{R,I}|\boldsymbol{\gamma}_{R,I})$ as outlined in Eq. (4), we can minimize the negative logarithm of the cost function (Wipf and Nagarajan 2009):

$$\begin{aligned}\mathrm{L}(\boldsymbol{\gamma}_{R,I}) &= -\log p(\mathbf{Y}_{R,I}|\boldsymbol{\gamma}_{R,I}) \\ &= \mathrm{trace}\left(\mathbf{C}_{YR,I}\Sigma_{YR,I}^{-1}\right) + \log\left(|\Sigma_{YR,I}|\right)\end{aligned} \quad (7)$$

The expectation steps of the EM procedure involve the updates for the estimation of unknown $\mathbf{S}_R$ and $\mathbf{S}_I$ and their covariance:

$$E\left[\mathbf{S}_{R,I}|\mathbf{Y}_{R,I}, \boldsymbol{\Gamma}_{R,I(k)}\right] = \boldsymbol{\Gamma}_{R,I(k)}\mathbf{L}^T(\boldsymbol{\Gamma}_{\mathrm{Y}R,I})_{(k)}^{-1}\mathbf{Y}_{R,I} \quad (8)$$

$$\mathrm{Cov}\left[\mathbf{S}_{R,I}(n)|\mathbf{Y}_{R,I}(n), \boldsymbol{\Gamma}_{R,I(k)}\right] = \boldsymbol{\Gamma}_{R,I(k)} - \boldsymbol{\Gamma}_{R,I(k)}\mathbf{L}^T(\boldsymbol{\Gamma}_{\mathrm{Y}R,I})_{(k)}^{-1}\mathbf{L}\boldsymbol{\Gamma}_{R,I(k)}$$

where n is the window index and $\boldsymbol{\Gamma}_{R,I} = \mathrm{diag}(\boldsymbol{\gamma}_{R,I})$.

The maximization step involves the marginalization over $\boldsymbol{\gamma}$:

$$\gamma_{R,I(k+1)} = \left(\frac{1}{n}\left\|\gamma_{R,I(k-1)}\mathbf{L}_i^T(\boldsymbol{\Sigma}_{YR,I})_{(k-1)}^{-1}\mathbf{Y}\right\|_{\Im}^2 \times \left(\mathrm{tr}\left(\gamma_{R,I(k-1)}\mathbf{L}_i^T(\boldsymbol{\Sigma}_{YR,I})_{(k-1)}^{-1}\mathbf{L}_i\right)\right)\right)^{-1} \quad (9)$$

After the algorithm has converged to a maximum likelihood set $\gamma_{\mathrm{R,Iconv}}$, from Eq. (8), the RVM posterior mean estimates are obtained:

$$\hat{\mathbf{S}}_{R,I} = E\left[\mathbf{S}_{R,I}|\mathbf{Y}_{R,I},\gamma_{R,\mathrm{Iconv}}\right] = \boldsymbol{\Gamma}_{R,\mathrm{Iconv}}\mathbf{L}^T(\boldsymbol{\Sigma}_Y)_{R,\mathrm{Iconv}}^{-1}\mathbf{Y}_{R,I} \quad (10)$$

For source activity estimation, the posterior means $\hat{s}_{R_i} = E[\mathbf{S}_R(i)|\mathbf{Y}_R,\gamma_{\mathrm{Rconv}}]$ and $\hat{s}_{I_i} = E[\mathbf{S}_I(i)|\mathbf{Y}_I,\gamma_{\mathrm{Iconv}}]$ for the i-th voxel at a frequency of interest are considered. Then,

$$\hat{s}_i = \sqrt{\hat{s}_{Ri}^2 + \hat{s}_{Ii}^2} \quad (11)$$

provides the power estimate at that frequency. The power mapping is performed over the whole brain. The left term in Eq. (11) defines the fdRVM imaging output for each brain voxel.

It is worth noting that the relationship between estimated source activity and sensor signals in Eq. (10) is common to several Bayesian and classical schemes (Friston et al. 2008b; Mosher et al. 2004). For instance, DICS, in place of $\boldsymbol{\Gamma}_{\mathrm{R,I}}$ matrices, utilizes $\mathbf{L}_i^T\mathbf{C}_r(f)\mathbf{L}_i)^{-1}$ as a gain factor between source and sensor level, where $\mathbf{C}_r(f)$ is the CSD.

Construction of simulations

We designed a simulated set of three cortical sources to study the effect of correlation level and SNR, with the following guidelines: (a) exclusion of possible bias due to hemispheric symmetry effects (some Bayesian algorithms, due to prior information on bilateral correlated sources, tend to detect symmetric sources as false positives), (b) variation of mutual distances between sources to monitor the effect of proximity, and (c) variation of source depths from the cortical surface. Three neural sources, represented by current dipoles, were placed within one hemisphere only, here in the right hemisphere (Fig. 1a), asymmetrically with respect to the x axis that is directed from the left to the right ear. The intersource distances varied from 3 to 9 cm: Source S_3 was placed posteriorly in the brain ($y = -3$ cm; y axis from the back to the front of the head), 6 cm from source S_2 ($y = 3$ cm) and 9 cm from source S_1 ($y = 6$ cm). The other two coordinates were $x = 4$ cm and $z = 7$ cm for each of the three sources. The brain of a real subject was chosen as the source space. The BEM model generated for this brain was used for the solution to the forward problem. The grid step within the brain for both DICS and fdRVM was 6 mm. The distance between each simulated source location and the nearest grid point was $<10^{-2}$ cm. A Neuromag 306-channel whole-head MEG system (Elekta Oy, Finland) was used for both signal generation and localization. The device contains 102 triple sensor elements composed of two orthogonal planar gradiometers and one magnetometer. The planar gradiometers detect the maximum signal directly above an active cortical area. In this study, the analysis was performed using the gradiometers.

We simulated two 10-min data sets, one with 'steady-state' (with respect to the time-analysis window) sources

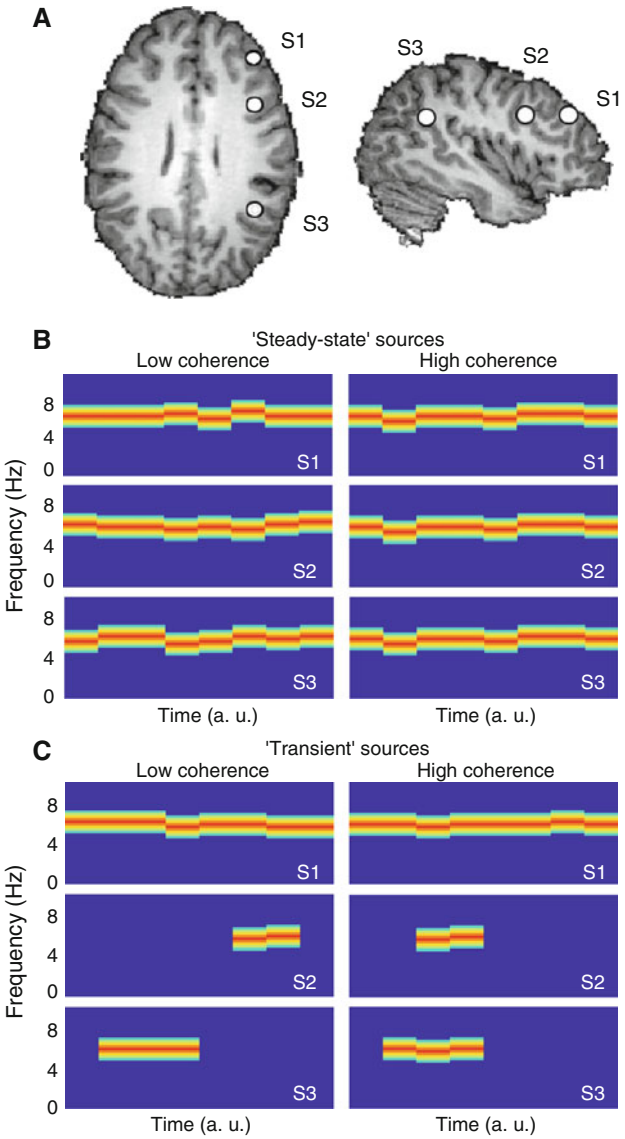

Fig. 1 Locations and time–frequency representations of the activity of each simulated source. **a** Three sources (anterior S_1 and S_2, posterior S_3,) were placed in the right hemisphere. **b** A schematic representation of how the simulated data with 'steady-state' and **c** 'transient' activity at high (>0.8) and low (<0.2) coherence behaved in time and frequency

and another with 'transient' sources, for the following reasons:

(a) Frequency-domain analysis methods typically assume 'steady-state' patterns using long stretches of data as the analysis window. Normally, several minutes and thousands of samples are considered (Fig. 1b). This implies that 'transient' sources (Fig. 1c) are detected at a lower SNR with respect to sources which display activation that spans the entire time window of analysis. In fact, even if a source is active only for a portion of the time stretch, the power localization in the frequency domain is biased by the whole amount of noise present in the entire time window. Therefore, a correct localization of this type of source tends to be problematic and requires particularly sensitive tools.

(b) In a simulation involving several sources, the degree of correlation among them can be manipulated in various ways. With 'steady-state' sources, the frequency or amplitude distribution over time can be adjusted to the desired level (Fig. 1b). With 'transient' sources, more leverage is available: Even if the time courses are very similar, their mutual correlations can be altered by adjusting their temporal overlap (Fig. 1c). A large overlap yields high coherence, whereas a small or nonexistent overlap assures low coherence. The use of different set-ups covers several realistic scenarios, as brain areas may have different activation times, time courses and phase lags.

In the first simulation ('steady-state' sources, Fig. 1b), the time courses of activation in the source areas were created firstly by setting instantaneous frequencies at each time-point and then generating the corresponding frequency-modulated signals. The instantaneous frequencies consisted of a base frequency and a random component. The base frequency was set to 6 Hz, and the random component for each source was set so that the instantaneous frequencies fell between 5 and 7 Hz at all time points. The random components were varied until all the sources showed either high (>0.8) or low (<0.2) coherence with each other.

In the second simulation, the amplitudes of the source time courses were modulated within the 10-min time window of interest so that they could be considered 'transient' with respect to that time interval. The base frequency f was 6 Hz, i.e., the same as in the 'steady-state' simulation. The three time courses S_1 (t), S_2 (t) and S_3 (t) were generated as follows:

$$S_k(t) = \exp((t - t_k)^2 / A_k^2) \cdot \sin(2\pi(f(t - t_k^*))$$

A, t_k and t_k* are parameters that determine the signal extent and shape and its position in the time window. In the high-coherence simulation, the values were as follows: $A_1 = 35$; $t_1 = 43$; t_1* $= 0.2$; $A_2 = 12$; $t_2 = 45$; t_2* $= 0.1$; $A_3 = 15$; $t_3 = 40$; t_3* $= 0.2$. In the low coherence simulation, the parameters assumed the values: $A_1 = 35$; $t_1 = 15$; t_1* $= 0.2$; $A_2 = 2$; $t_2 = 90$; t_2* $= 0.9$; $A_3 = 15$; $t_3 = 50$; t_3* $= 0.1$. The S_1 (t) time course can be considered a limit case of a 'transient' source. The relatively high value of A_1 attenuated the modulating action of the exponential. Therefore, source S_1 remained active for the whole span of the time window, although at the extremes of the time window, its amplitude values are remarkably narrower than in the central section (Fig. 1b, c). The sampling frequency for both simulations was 300 Hz.

Noise was generated at the sensor level by means of white noise (full frequency band: 0–100 Hz). SNRs were simulated with a range from 0 to −13 dB and a 1 dB step. Truly realistic brain noise has turned out to be difficult to simulate (Ghosh et al. 2008). As an alternative, we additionally tested resting-state data (eyes closed) as noise when using fdRVM to localize sources from 'steady-state' data at high coherence and 5 SNRs (0, −3, −6, −9, −12 dB).

Real data

This test exploited data from two earlier studies on healthy right-handed subjects: (a) Finger movements (Gross et al. 2002): Six subjects were asked to perform continuous sinusoidal flexion and extension movements in a horizontal plane with their right index finger. The data were recorded with a 122-channel Neuromag MEG system (Elekta Oy, Finland). (b) Resting state: Two subjects were asked to sit relaxed, with their eyes open (2 min) and closed (2 min). The data were recorded with the same MEG system that was used to generate the simulated signals.

Comparison of fdRVM and DICS performance

The localization performance of fdRVM and DICS on the parametrically varying simulations was evaluated by means of three parameters:

(a) The number of correctly detected sources. 'Correct detection' required that the local maximum fell on the correct grid point or on an immediately neighboring grid point. The cut-off threshold was set to 0.25 (normalized to the global maximum). Since all the simulated sources had the maximum unit amplitude, it was assumed that a reduction in the estimated amplitude by more than 75 % could not be considered correct.

(b) The spatial spread of the estimated source peak. As an expression of the spatial spread of the estimate, the full width at half maximum (FWHM) of source S_3 was computed for every condition (low/high coherence,

steady-state/transient time courses) at different SNR levels. By approximating the profile of the source intensity as a Gaussian function, FWHM was obtained as the number of grid steps between the two voxels at which the source intensity was equal to half of its maximum value. Only S_3 (Fig. 1a) was considered in this evaluation because S_1 and S_2 were often detected as one single source by DICS (and sometimes by fdRVM as well) in conditions of low SNR and/or high coherence.

(c) The cumulative distance of the misplaced peaks of activity with respect to the actual location.

Results

Simulations

Figure 2 illustrates source localization using fdRVM and DICS for the four different types of simulations ('steady-state' and 'transient' oscillatory activity, low and high coherence among sources at SNR = −4 dB). Figure 3 quantifies the number of detected sources as a function of SNR for the two techniques. When a source was detected, the localization was mostly accurate, i.e., the maximum of the power estimate fell on the correct grid point or on one of its immediate neighbors. The main cause of misplacement errors was the merging of the two anterior sources, resulting in a unique main peak between the two actual dipoles. Figure 4 depicts the spatial spread (FWHM) of the power estimate as a function of SNR (A, B) and the cumulative distance of mislocated sources from the actual ones (C, D). With decreasing SNR, the detection accuracy decreased more rapidly for DICS than fdRVM in each condition (Fig. 3). The first drop from 3 to 2 detected sources with DICS (at −4 dB for low-coherence and 0 dB for high-coherence 'steady-state' activity) was due to the merging of estimated source activity of the two frontal sources (Fig. 3a, c). The apparent merging was a result of the increasing spatial spread of the estimated activation with decreasing SNR (Fig. 2b, d). This increase had a markedly steeper slope for DICS than for fdRVM, in both low- and high-coherence simulations (Fig. 4a, b). In fdRVM, due to the algorithm action that prunes non-relevant components to zero (see "Methods" section), the two frontal sources remained separable down to SNR = −8 dB. The cumulative distance of misplaced sources from the actual locations plotted in (Fig. 4c, d) shows, as expected, that at low SNRs, the correlation level does not play a significant role, as the plots show no significant difference between conditions of high and low correlation for either DICS or fdRVM. To test the effect of different kinds of background noise on our performance comparison, we replaced the white noise with rest data in a subset of test conditions: continuous data, high coherence, 5 signal-to-noise ratios (0, −3, −6, −9 and −12). The results remained remarkably unchanged: the root mean squared deviation between source activity results with realistic noise and Gaussian distributed noise never exceeded 0.1. Replacing the white noise by rest data thus seemed to have little influence on the resulting maps.

In conditions of both low and high coherence, fdRVM detected sources S_1, S_2 and S_3 with the correct relative intensities. Since the simulated peak amplitude was the same for all sources (peak value = 1), the discriminating parameter is expected to be the duration of the source time course. Hence, the correct order of detected intensity should be $S_1 > S_3 > S_2$. This order was observed with fdRVM (Fig. 2c) while DICS did not detect the correct relationship between the effective source intensities (Fig. 2d).

Overall, the relationship between the performances of fdRVM and DICS as a function of SNR was similar in both the 'steady-state' and 'transient' source simulations. As expected, the decreasing SNR affected both algorithms, more when the sources were active only transiently than when they were active throughout the analysis time window ('steady-state'). fdRVM performed systematically better than DICS on these simulated data sets, especially when the coherence between sources was high (>0.8).

Figure 5 shows the S_R and S_I maps in the case of steady-state sources and high coherence with SNR = −4. The zero locations of the real and imaginary gamma maps are essentially the same, but their contributions to source intensities are different. Figure 5a displays the quantity:

$$\left(\left(\hat{\mathbf{S}}_I - \hat{\mathbf{S}}_R\right) - \min\left(\hat{\mathbf{S}}_I - \hat{\mathbf{S}}_R\right)\right) \Big/ \left(\max\left(\hat{\mathbf{S}}_I - \hat{\mathbf{S}}_R\right) - \left(\hat{\mathbf{S}}_I - \hat{\mathbf{S}}_R\right)\right) \quad \text{if} \left(\hat{\mathbf{S}}_I - \hat{\mathbf{S}}_R\right) > 0, \text{ otherwise } 0,$$

i.e., the contributions of the imaginary part that exceed those of the real part. Figure 5b depicts the opposite measure:

$$\left(\left(\hat{\mathbf{S}}_R - \hat{\mathbf{S}}_I\right) - \min\left(\hat{\mathbf{S}}_R - \hat{\mathbf{S}}_I\right)\right) \Big/ \left(\max\left(\hat{\mathbf{S}}_R - \hat{\mathbf{S}}_I\right) - \left(\hat{\mathbf{S}}_R - \hat{\mathbf{S}}_I\right)\right) \quad \text{if} \left(\hat{\mathbf{S}}_R - \hat{\mathbf{S}}_I\right) > 0, \text{ otherwise } 0,$$

i.e., the contributions of the real part that exceed those of the imaginary part. The anterior source receives a stronger contribution from the imaginary component. Figure 5c displays the combined effect $\hat{\mathbf{S}} = \sqrt{\hat{\mathbf{S}}_R^2 + \hat{\mathbf{S}}_I^2}$, similarly to Fig. 2. Accordingly, the real and imaginary contributions are non-negligible and complementary, and only the composition of the two provides the correct mapping.

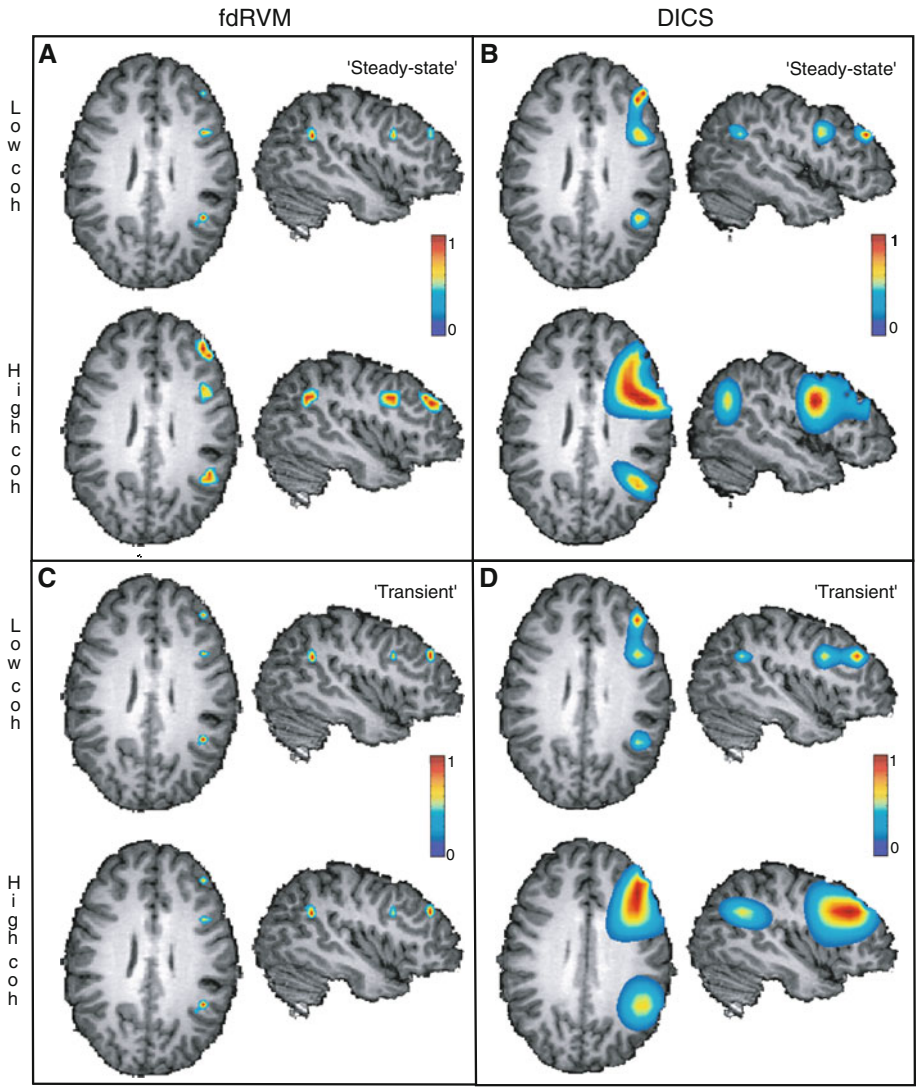

Fig. 2 Localization of simulated oscillatory sources. **a** fdRVM localization of sources in the 'steady-state' condition, for low and high coherence, at SNR = −4 dB. **b** DICS localization of sources in the 'steady-state' condition, for low and high coherence, at SNR = −4 dB. **c** fdRVM localization of sources in the 'transient' condition, for low and high coherence, at SNR = −2 dB. **d** DICS localization of sources in the 'transient' condition, for low and high coherence, at SNR = −2 dB. The activity mapping was normalized to the maximum peak. The *color bars* indicate normalized power (from 0 to 1). The cut-off threshold was set at 25 % of the maximum power value (color figure online)

Real data

In finger movements, a clear spectral peak at around 10 Hz was detectable at the sensor level over the sensorimotor area (Fig. 6a, upper panel; example of one subject's data). Oscillatory activity over the occipital area was evident as well. Localization of rhythmic activity at 8–12 Hz using fdRVM revealed a source in the right (ipsilateral) primary motor area (Fig. 6a, lower panel). It may reflect ipsilateral motor inhibition, which exerts an effect on the contralateral motor cortex (Daskalakis et al. 2002). Activity was detected in the occipital area as well. DICS localized the activity in the occipital area (Fig. 6a, lower panel); detection of the weaker sensorimotor sources requires that the signal from the strong occipital source first be removed (Liljeström et al. 2005).

In the resting-state data, as well, a strong power peak at about 10 Hz was evident at the sensor level (Fig. 6b, upper panel; example of one subject's data). Localization of 9–12 Hz activity using fdRVM displayed bilateral sources in the sensorimotor area (Fig. 6b, lower panel). Source of rhythmic activity in the occipital cortex was detected only when the cut-off threshold was lowered from 25 to 10 % of maximum power. DICS localized again the main activity in the occipital area, which would need to be removed in order to fully reveal the weaker sensorimotor sources (Fig. 6b, lower panel). Even in real data, fdRVM thus features very focal localizations, possibly penalizing sources which spread over a large number of sensors, whereas DICS mapping is spatially rather extended. The absence of occipital activity in the fdRVM localization in Fig. 6b is also probably influenced by the algorithm's sharp selection of the relevant frequency bins; the spectral peaks in the occipital area are 1–2 Hz higher in frequency than the peaks in the sensorimotor area, at the border of the frequency band of interest.

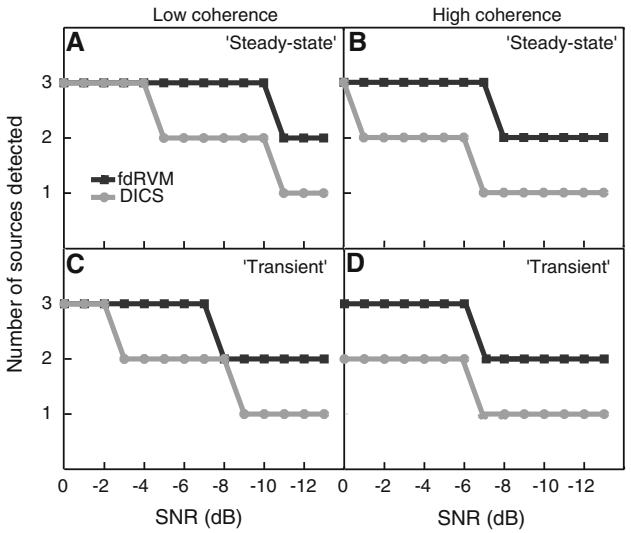

Fig. 3 Number of correctly detected sources as a function of decreasing SNR for fdRVM (*black line*) and DICS (*gray line*). **a** 'Steady-state' condition and low coherence among sources, **b** 'steady-state' condition and high coherence, **c** 'transient' condition and low coherence, **d** 'transient' condition and high coherence

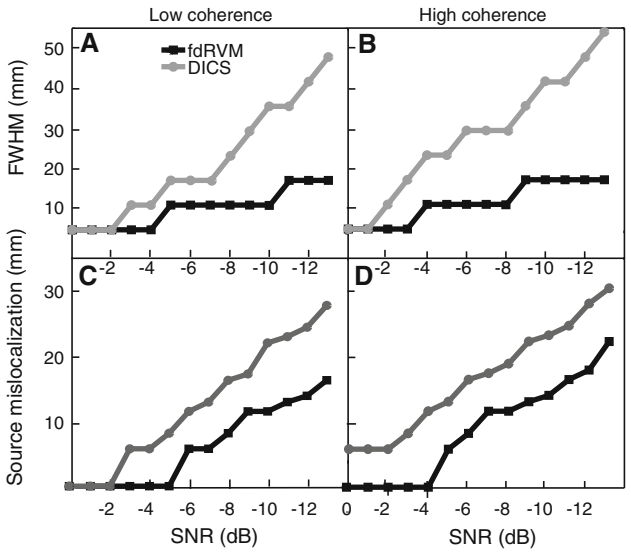

Fig. 4 Spatial accuracy of the two methods. Full width at half maximum of the source estimates from fdRVM (*black line*) and DICS (*gray line*) maps as a function of decreasing SNR for **a** low and **b** high coherence among sources. In conditions of high coherence, DICS detected the anterior sources S_1 and S_2 as merged together in both the 'steady-state' and 'transient' simulations (cf. Fig. 2b, d). For this reason, the FWHM results are plotted only for the posterior source S_3, as an average of the 'transient' and 'steady-state' estimates. Below, cumulative source mislocalization is plotted for both algorithms in case of **c** low and **d** high correlation

Discussion

In this study, an RVM-based algorithm was adapted for mapping brain rhythmic activity from Fourier-transformed biomagnetic signals. The signals in the time domain were partitioned in windows with an equal number of time samples. Subsequently, a Fast Fourier Transform was applied to each window, and one mean complex number was calculated for each frequency and time window, thus yielding a complex series in the time domain. The main idea of fdRVM consists of separately handling the resulting real and imaginary series by means of two distinct parallel loops and, consequently, combining the real and imaginary source power components in one single mapping. Critically, such a separation is permitted by the linearity of the RVM model and by the independence of the real and imaginary series.

Wipf and Nagarajan (2007) also considered the performance of RVM on complex data, they did this by simulating sources that showed complex activity (i.e., multi-dimensional sources). Thus, their approach was fundamentally different from the frequency-domain analysis conducted here, where the simulated and recorded activity consist of real-valued data (representing the type of activity and recordings that beamforming approaches are generally applied to) and the complex nature of the signals is due to a subsequent transformation (FFT) that aims to disentagle the content of the signals at different frequencies.

fdRVM has two useful computational properties: (a) It does not calculate the imaginary mixed terms produced by the multiplication of source CSD for its transposed conjugate. For example, in DICS, these terms are initially calculated but then discarded when the source-level power is estimated along the two orthogonal tangential directions, i.e., when only the real part of the 2×2 complex source activity matrix is considered for SVD evaluation (Gross et al. 2001). (b) fdRVM allows to process the real and imaginary contributions in parallel and simultaneous loops. This saves computation time and, with respect to the standard implementation of RVM, yields exclusively the terms needed for the power mapping (real and imaginary components, no spurious terms). The Fourier transform employs the complex numbers for a compact representation of sinusoidal and cosinusoidal components. Both imaginary and real series contribute in a different and non-negligible way to the mapping of rhythmic source activity.

Among various existing and fairly widely used methods for localization of rhythmic brain activity (minimum-norm estimates, beamformers and sequential dipole fitting), a beamformer-based technique (DICS) was found to be the most sensitive, and it could separate close-by sources better than other methods (Liljeström et al. 2005). Notably, however, the main assumption of beamformer algorithms is that the source time courses are uncorrelated. This assumption is no longer satisfied when SNR is increased and the degree of correlation among sources is high. At low SNR, an inaccuracy in source localization is mainly due to

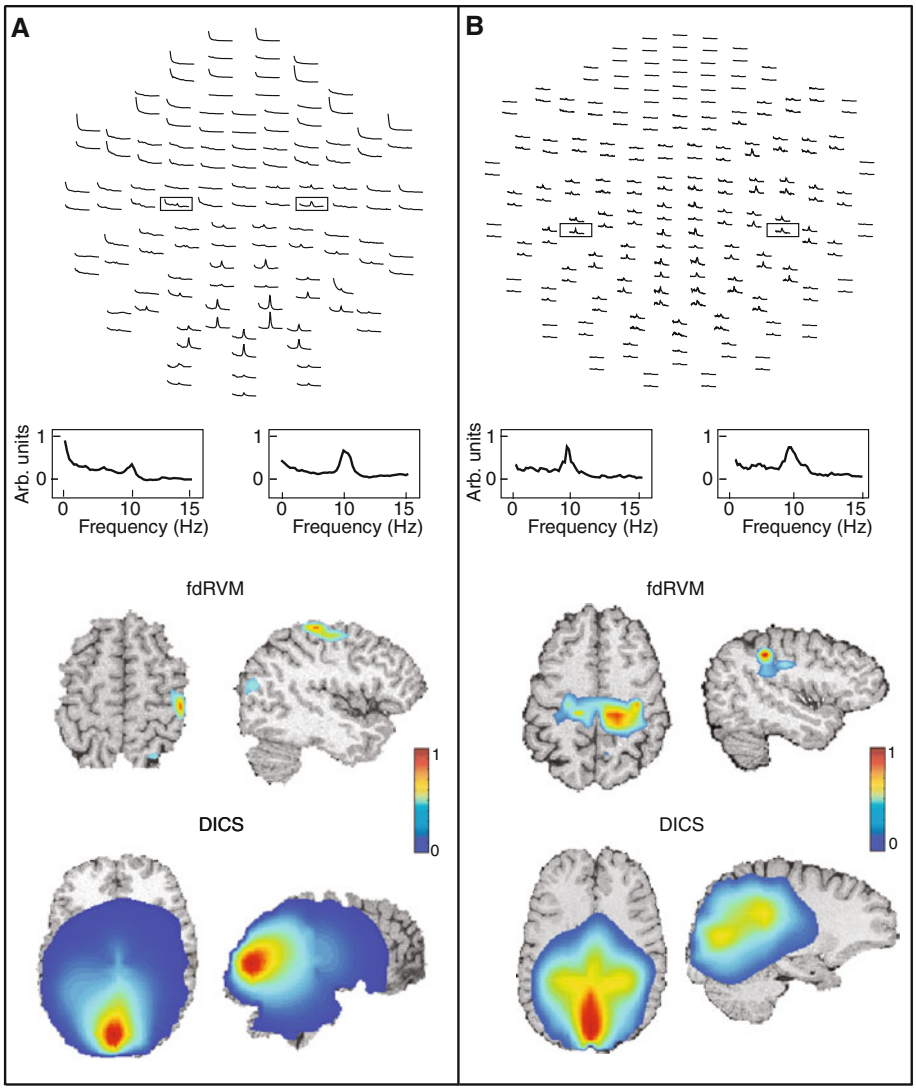

Fig. 5 Contribution of real and imaginary parts in fdRVM. Source maps for steady-state sources, high coherence and SNR = −4. **a** Localization of imaginary > real contribution. **b** Localization of real > imaginary contribution. **c** Final output of fdRVM that combines the real and imaginary contributions

an overestimation of the spatial spread of the detected source or to a missing detection rather than to a bias caused by correlation among the sources. In real brain activity, the observed coherence levels tend to be fairly low (Kujala et al. 2007; Timmermann et al. 2002) but the SNR is low as well. Since there is no fixed value for a 'safe' limit of power mapping with beamformers in real data, it would beneficial to have alternative—or supporting—methods that are less influenced by correlations. RVM-based approaches (Wipf and Nagarajan 2009) can provide this type of tools.

The fdRVM approach was compared with the beamformer method DICS using two types of simulated data as well as two different sets of real data. The simulated data comprised 'steady-state' oscillations and 'transient' sources, with both low and high degrees of mutual correlations. fdRVM showed three main advantages with respect to the beamformer-based technique: (a) It removed the undesirable effects of correlation when analyzing Fourier-transformed MEG/EEG data in the frequency domain. (b) The pruning of output components characteristic of the EM algorithm yielded localization results that were markedly more focal than with the beamforming approach. (c) fdRVM maintained its spatial resolution even at low SNR. Due to its resistance to correlation and noise effects, fdRVM can be used to reconstruct an unbiased CSD that shares the power-spectral properties of the original CSD. One might thus be able to obtain DICS coherence estimates using the unbiased CSD. Moreover, the imaginary component $\mathbf{S}_{\mathrm{I}}$ can be employed for functional connectivity between one source voxel and the rest of the brain. One could indeed use just the complex signal part to avoid leakage when calculating connectivity measures uniquely based on phase like imaginary coherence (Nolte et al. 2004) and Phase Lag Index (Stam et al. 2007).

An important contribution to the fdRVM correct performance is given by a proper estimation of the error variance Σ_ε in Eq. (10). Since we are analyzing rhythmic

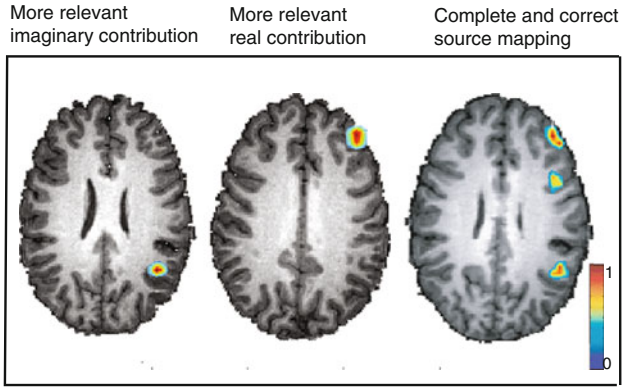

Fig. 6 Power mapping results on two real MEG data sets. **a** Slow horizontal movement of the right index finger, one subject. *On top* power spectra at the sensor level. The *helmet-shaped array* of sensors (Neuromag system, 122 gradiometers) is displayed flattened onto the plane, with the nose pointing upwards. Data of two sensors (*boxes*) are shown enlarged. Below, source-level maps at 10 Hz obtained with fdRVM and DICS. **b** Resting-state data of one subject. Sensor-level helmet view on *top* (204 gradiometers; Elekta), localization results below (centered at 10.5 Hz). The *color bars* indicate normalized power (from 0 to 1) (color figure online)

activity in the frequency domain, a noise estimation cannot be determined from baseline intervals as in (Zumer et al. 2007). The error variance represents the trade-off between data fit (small values of $\Sigma_{\varepsilon R,I}$ correspond to a noise-free estimation of the data) and mapping sparsity (high values of $\Sigma_{\varepsilon R,I}$ correspond to elevated confidence in the hyperpriors effect and, conversely, low trust in the data). We consider this term as an estimate of noise variance α mediated by a multiplicative constant λ which expresses our confidence in the hyperpriors: $\Sigma_{\varepsilon} = \lambda \cdot \alpha_{R,I}$. An adequate choice for λ in the case of noisy oscillating signals appears to be in the range between 10^{-4} and 10^{-1}. The noise variance was estimated taking into account the poor SNR of our data. Since in the simulations, the noise was of the same amplitude as the signal or higher, we approximated the noise variance to the largest eigenvalue of the signal variance $\mathbf{C}_{\mathbf{YR,I}}$. In practice, the λ factor implies a probable underestimation of the effects of noise. Still, an approach underestimating the noise variance (or even setting it to zero) appears to be a good practice for a better final mapping (Wipf et al. 2009). This kind of approach allows for more accurate results but requires a longer time for the algorithm to converge.

Frequency-domain RVM method (fdRVM) was more robust against poor SNR than DICS but, similarly to DICS, its performance was found to drop at lower SNR. The SNR may decrease via increasing noise level or shorter duration of the source activation. In the case of 'steady-state' sources, the algorithm yielded a less focal localization under condition of high than low coherence (Fig. 2a). Since correlation cannot affect the global minimum of convergence, we must assume that the algorithm became stuck in a local minimum very close to the global one. However, the influence of coherence level was clearly more marked on DICS than on fdRVM maps (Fig. 2b, d). fdRVM performed in a consistent manner on real data and it retained its focality advantage with respect to the beamformer-based technique. Since fdRVM detects amplitudes of the simulated sources in the correct order of intensity (as shown in the transient source simulation), whereas DICS does not, contrasting experimental conditions A and B should bring a further relative benefit to fdRVM results. This information is partly included already in the single condition simulation. If we consider our simulations as the active condition, our results show that while DICS noise-normalized mapping fails at low SNR, fdRVM "absolute" mapping shows the correct locations. If we were to add a baseline condition to this analysis, it could not improve DICS results as the original mapping has actually failed, rather than yielding insensitive results which could be sharpened by contrasting. For fdRVM, on the other hand, comparing to a baseline could further improve the results. The mapping in the baseline condition would probably be pruned to zero in each voxel (if baseline = homogeneous noise on the sensors) or to lower values than in the active condition (if baseline = less signal than in the active condition). When DICS yields "correct" mapping (high SNR), contrasting would of course be reasonable (at least in simulations).

At present, the main disadvantage of fdRVM is the computational time. The final power mapping file for one 10-min epoch requires 2–3 h on a single PC processor core. This inconvenience can be addressed both on algorithmic (faster convergence of the iterative EM procedure) and technical level (threaded programming).

Conclusions

Because of its sensitivity, robustness and steep localization mapping, fdRVM may afford some new insights to the study of brain rhythmic activity. It may be helpful, for example, in detecting sources of oscillatory activity and facilitating the study of their dynamics, even when the activity is strongly correlated with that of other brain areas. In particular, discrimination among 'transient' sources (i.e., not active for the whole duration of the time window of interest) appears to be possible with fdRVM. In conditions of equal source intensity, sources that had a longer duration were detected as more intense in the frequency-domain localization. A natural continuation of the present work would be the addition of time information in the algorithm. By means of a time–frequency resolved RVM, brain sources showing different intensities in the frequency domain could be correctly discriminated by their different intensities, durations, or both.

References

Auranen T, Nummenmaa A, Hämäläinen MS, Jääskeläinen IP, Lampinen J, Vehtari A, Sams M (2005) Bayesian analysis of the neuromagnetic inverse problem with lp-norm priors. Neuroimage 26:870–884

Baillet S, Mosher JC, Leahy RM (2001) Electromagnetic brain mapping. IEEE Signal Process Mag 18:14–30

Belardinelli P, Ciancetta L, Staudt M, Pizzella V, Londei A, Birbaumer N, Romani GL, Braun C (2007) Cerebro-muscular and cerebro-cerebral coherence in patients with pre- and perinatally acquired unilateral brain lesions. Neuroimage 37:1301–1314

Belardinelli P, Ortiz E, Barnes G, Noppeney U, Preissl H (2012) Source reconstruction accuracy of MEG and EEG Bayesian inversion approaches. PLoS ONE 7:e51985

Brookes MJ, Gibson AM, Hall SD, Furlong PL, Barnes GR, Hillebrand A, Singh KD, Holliday IE, Francis ST, Morris PG (2004) A general linear model for MEG beamformer imaging. Neuroimage 23:936–946

Dalal SS, Sekihara K, Nagarajan SS (2006) Modified beamformers for coherent source region suppression. IEEE Transact Biomed Eng 53:1357–1363

Dalal SS, Guggisberg AG, Edwards E, Sekihara K, Findlay AM, Canolty RT, Knight RT, Barbaro NM, Kirsch HE, Nagarajan SS (2007) Spatial localization of cortical time-frequency dynamics. Conf Proc IEEE Eng Med Biol Soc 2007:4941–4944

Daskalakis ZJ, Christensen BK, Fitzgerald PB, Roshan L, Chen R (2002) The mechanisms of interhemispheric inhibition in the human motor cortex. J Physiol 543:317–326

Friston KJ, Glaser DE, Henson RNA, Kiebel S, Phillips C, Ashburner J (2002) Classical and Bayesian inference in neuroimaging: applications. Neuroimage 16:484–512

Friston K, Chu C, Mourão-Miranda J, Hulme O, Rees G, Penny W, Ashburner J (2008a) Bayesian decoding of brain images. Neuroimage 39:181–205

Friston K, Harrison L, Daunizeau J, Kiebel S, Phillips C, Trujillo-Barreto N, Henson R, Flandin G, Mattout J (2008b) Multiple sparse priors for the M/EEG inverse problem. Neuroimage 39:1104–1120

Ghosh A, Rho Y, McIntosh A, Kötter R, Jirsa V (2008) Noise during rest enables the exploration of the brain's dynamic repertoire. PLoS Comput Biol 4:e1000196

Gross J, Ioannides A (1999) Linear transformations of data space in MEG. Phys Med Biol 44:2081–2097

Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R (2001) Dynamic imaging of coherent sources: studying neural interactions in the human brain. Proc Natl Acad Sci 98:694–699

Gross J, Timmermann L, Kujala J, Dirks M, Schmitz F, Salmelin R, Schnitzler A (2002) The neural basis of intermittent motor control in humans. Proc Natl Acad Sci 99:2299–2302

Hari R, Salmelin R (1997) Human cortical oscillations: a neuromagnetic view through the skull. Trends Neurosci 20:44–48

Hirschmann J, Özkurt T, Butz M, Homburger M, Elben S, Hartmann C, Vesper J, Wojtecki L, Schnitzler A (2011) Distinct oscillatory STN-cortical loops revealed by simultaneous MEG and local field potential recordings in patients with Parkinson's disease. Neuroimage 55:1159–1168

Jensen O, Vanni S (2002) A new method to identify multiple sources of oscillatory activity from magnetoencephalographic data. Neuroimage 15:568–574

Jerbi K, Lachaux JP, N'Diaye K, Pantazis D, Leahy RM, Garnero L, Baillet S (2007) Coherent neural representation of hand speed in humans revealed by MEG imaging. Proc Natl Acad Sci 104:7676–7681

Kujala J, Pammer K, Cornelissen P, Roebroeck A, Formisano E, Salmelin R (2007) Phase coupling in a cerebro-cerebellar network at 8–13 Hz during reading. Cereb Cortex 17:1476–1485

Liljeström M, Kujala J, Jensen O, Salmelin R (2005) Neuromagnetic localization of rhythmic activity in the human brain: a comparison of three methods. Neuroimage 25:734–745

MacKay DJC (1992) Bayesian interpolation. Neural Comput 4:415–447

Mazaheri A, Nieuwenhuis ILC, van Dijk H, Jensen O (2009) Prestimulus alpha and mu activity predicts failure to inhibit motor responses. Hum Brain Mapp 30:1791–1800

Mosher JC, Baillet S, Leahy RM (2004) Equivalence of linear approaches in bioelectromagnetic inverse solutions. IEEE Workshop on Statistical Signal Processing, pp 294–297

Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S, Hallett M (2004) Identifying true brain interaction from EEG data using the imaginary part of coherency. Clin Neurophysiol 115:2292–2307

Nummenmaa A, Auranen T, Hämäläinen MS, Jääskeläinen IP, Lampinen J, Sams M, Vehtari A (2007) Hierarchical Bayesian estimates of distributed MEG sources: theoretical aspects and comparison of variational and MCMC methods. Neuroimage 35:669–685

Osipova D, Takashima A, Oostenveld R, Fernández G, Maris E, Jensen O (2006) Theta and gamma oscillations predict encoding and retrieval of declarative memory. J Neurosci 26:7523–7531

Pfurtscheller G, Lopes da Silva FH (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol 110:1842–1857

Phillips C, Mattout J, Rugg M, Maquet P, Friston K (2005) An empirical Bayesian solution to the source reconstruction problem in EEG. Neuroimage 24:997–1011

Salmelin R, Hari R (1994) Characterization of spontaneous MEG rhythms in healthy adults. Electroencephalogr Clin Neurophysiol 91:237–248

Sato M, Yoshioka T, Kajihara S, Toyama K, Goda N, Doya K, Kawato M (2004) Hierarchical Bayesian estimation for MEG inverse problem. Neuroimage 23:806–826

Scholkopf B, Smola AJ, Williamson RC, Bartlett PL (2000) New support vector algorithms. Neural Comput 12:1207–1245

Sekihara K, Nagarajan SS, Poeppel D, Marantz A (2002) Performance of an MEG adaptive-beamformer technique in the presence of correlated neural activities: effects on signal intensity and time-course estimates. IEEE Trans Biomed Eng 49:1534–1546

Stam CJ, 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 28:1178–1193

Timmermann L, Gross J, Dirks M, Volkmann J, Freund HJ, Schnitzler A (2002) The cerebral oscillatory network of parkinsonian resting tremor. Brain 126:199–212

Tipping ME (2001) Sparse bayesian learning and the relevance vector machine. J Mach Learn Res 1:211–244

Tipping ME (2004) Bayesian inference: an introduction to principles and practice in machine learning. Lecture notes in computer science, pp 41–62

Van Veen BD, Van Drongelen W, Yuchtman M, Suzuki A (1997) Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans Biomed Eng 44:867–880

Vrba J, Robinson SE (2001) Signal processing in magnetoencephalography. Methods 25:249–271

Wipf D, Nagarajan S (2007) Beamforming using the relevance vector machine. ICML '07 Proceedings of the 24th international conference on Machine learning 1023–1030

Wipf D, Nagarajan S (2009) A unified Bayesian framework for MEG/EEG source imaging. Neuroimage 44:947–966

Wipf D, Owen J, Attias H, Sekihara K, Nagarajan S (2009) Estimating the location and orientation of complex, correlated neural activity using MEG. Adv Neural Inform Process Syst 21. http://www.goldenmetallic.com/research/nips08.pdf

Wipf D, Owen J, Attias H, Sekihara K, Nagarajan S (2010) Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. Neuroimage 49:641–655

Zoltowski MD (1988) On the performance analysis of the MVDR beamformer in the presence of correlated interference. IEEE Trans Acoust Speech Signal Process 36:945–947

Zumer JM, Attias HT, Sekihara K, Nagarajan SS (2007) A probabilistic algorithm integrating source localization and noise suppression for MEG and EEG data. Neuroimage 37:102–115