Inverse Problems and Imaging
Volume X, No. 0X, 200X, X–XX

Web site: http://www.aimSciences.org

PARTICLE FILTERING, BEAMFORMING AND MULTIPLE SIGNAL CLASSIFICATION FOR THE ANALYSIS OF MAGNETOENCEPHALOGRAPHY TIME SERIES: A COMPARISON OF ALGORITHMS

Annalisa Pascarella

Dipartimento di Informatica, Università di Verona, Strada le Grazie 15 37134 Verona, Italy
CNR - INFM LAMIA, via Dodecaneso 35 16146 Genova, Italy

Alberto Sorrentino

CNR - INFM LAMIA, via Dodecaneso 33 16146 Genova Italy

Cristina Campi

Dipartimento di Matematica, Università di Genova, via Dodecaneso 35 16146 Genova, Italy
CNR - INFM LAMIA, via Dodecaneso 33 16146 Genova, Italy

Michele Piana

Dipartimento di Matematica, Università di Genova, via Dodecaneso 35 16146 Genova, Italy
CNR - INFM LAMIA, via Dodecaneso 33 16146 Genova, Italy

(Communicated by the associate editor name)

Abstract. We present a comparison of three methods for the solution of the magnetoencephalography inverse problem. The methods are: a linearly constrained minimum variance beamformer, an algorithm implementing multiple signal classification with recursively applied projection and a particle filter for Bayesian tracking. Synthetic data with neurophysiological significance are analyzed by the three methods to recover position, orientation and amplitude of the active sources. Finally, a real data set evoked by a simple auditory stimulus is considered.

1. Introduction

Magnetoencephalography (MEG) [1] is a functional imaging modality which allows measurements of the magnetic fields associated with spontaneous or stimulus-induced neural currents by means of Superconducting Quantum Interference Devices (SQUID). Due to its formidable temporal resolution (the magnetic signal can be followed with a detail of some milliseconds) MEG is considered as an optimal tool for investigating the dynamical interplay of brain regions during information processing. Furthermore, modern scanners, characterized by a high number of superconducting sensors distributed to homogeneously cover the whole skull, and co-registration techniques onto high resolution structural maps, allow a centimeter-level spatial localization of the active neural areas. However the full potentiality of MEG has not yet been exploited, mainly owing to the complexity of MEG data analysis. In fact the troublesome issue in MEG investigation is the solution of the neuromagnetic inverse problem of determining the spatio-temporal evolution of the neural currents from the dynamical measurements of the magnetic field at different

2000 *Mathematics Subject Classification:* Primary: 65R32; Secondary: 65C05.
Key words and phrases: Inverse problems, magnetoencephalography, Bayesian methods.

locations outside the skull. This problem is difficult for many reasons. First, it is numerically unstable, the integral problem describing the data being ill-posed [2]. Second, the problem has to be solved at many different time points and this makes computational effectiveness a crucial issue in the data analysis process. Third, the input data is characterized by a very low signal-to-noise-ratio (SNR), MEG signal being contaminated by both neural background and sensors' noise.

The purpose of the present paper is to compare three different reconstruction methods for the solution of the dynamical neuromagnetic problem. The rationale of this analysis is that several methods exist for the solution of such a difficult inverse problem but the effectiveness of most of these approaches has not yet been comparatively and systematically assessed against synthetic time series of notable neuroscientific significance. The aim of our investigation is to provide the inverse problems community with some quantitative results which can point out optimal strategies for the analysis of real measurements.

In an effective method for the solution of the MEG inverse problem, the following features should in principle coexist. From a neurophysiological viewpoint the localization properties of this method should allow the reconstruction of very small sources (up to point sources). Furthermore, particularly in the case of complex tasks, the simultaneous or quasi-simultaneous activation of different cortical areas is a well-established issue and therefore the spatio-temporal reconstruction should be able to handle multi-target situations even with a high time correlation degree. Multi-target reconstructions typically require a notable computational effort, so that the method should be numerically effective and characterized by a high level of automation, inasmuch as approaches which require subjective choices of the computational parameters are typically time consuming. Finally the optimal method should be robust with respect to increasing noise amounts affecting the data and assure to some extent regularization properties which reduce the intrinsic numerical instability of the inverse problem. In the present paper we perform a comparative analysis of three approaches which, at least in principle, share all (or most of) these properties: a beamforming algorithm [3, 4, 5] which spatially filters the signal to focus the source as a weighted combination of the measurements; multiple signal classification [6, 7, 8, 9], which identifies the source positions by scanning the brain volume to find the solution of a non-linear optimization problem; a particle filter [10, 11, 12, 13, 14, 15], which realizes a Bayesian tracking of the sources by means of a sampling-resampling of the probability density functions involved. Our ultimate goal is to systematically verify whether and to what extent optimal performances are actually fulfilled by each method in the processing of synthetic data realized according to certain critical but neuroscientifically plausible conditions. Operatively, this will be done by comparing the performances in both localizing the original sources and reproducing their temporal course, in the framework of a spatio-temporal analysis. Eventually, an application to real data will be considered.

The plan of the paper is as follows. Section 2 will set up the MEG inverse problem. In Section 3 the three algorithms will be formulated while Section 4 will describe the comparison in both theoretical and computational terms. Our conclusions will be offered in Section 5.

2. The MEG inverse problem

Within the quasi-static approximation [2], the magnetoencephalography problem is defined by the Biot-Savart equation:

$$\mathbf{b}(\mathbf{r},t) = \frac{\mu_0}{4\pi}\int_\Omega \mathbf{j}^{tot}(\mathbf{r}',t)\times\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3}d\mathbf{r}' \tag{1}$$

where $\mathbf{b}(\mathbf{r},t)$ is the magnetic field produced at time t by the current density $\mathbf{j}^{tot}(\mathbf{r}',t)$ inside the conductor volume Ω representing the brain, and μ_0 is the magnetic permittivity of vacuum. In a MEG scanner, the magnetic field is sampled on a finite number M of sensors, each one measuring one component of the magnetic field; if $\mathbf{e}_{s_m}$ is the unit vector orthogonal to the sensor's surface at location $\mathbf{r}_{s_m}$,

$$b_m(t) = \mathbf{b}(\mathbf{r}_{s_m},t)\cdot\mathbf{e}_{s_m} \quad m=1,\dots,M \tag{2}$$

is the dynamical measurement on the sensor s_m. Recovering the electrical current density inside Ω from external measurements of the magnetic field is in general an ill-posed problem, because the Biot-Savart operator has a non-trivial kernel and is compact [16, 17]. Furthermore, the purpose of source estimation from MEG data is that of recovering the neural currents from measurements of the magnetic field, but the total current density $\mathbf{j}^{tot}$ inside the brain is in fact the sum of two terms [2]: the primary current flowing in the active neurons, and the passive current flowing in the whole conductor because of the electrical potential generated by the primary current itself. The presence of the volume currents further increases the difficulty in recovering the neural current $\mathbf{j}(\mathbf{r},t)$: in general, numerical solutions of integral equations are needed to account for their presence. However, one artificial but reliable assumption can notably simplify the problem. Indeed, when cortical regions other than the frontal lobe are involved in the analysis, it is reliable [18] to approximate the head as a spherical homogeneous conductor. This approximation, together with the discretization of $\mathbf{j}(\mathbf{r},t)$ realized by attaching a current dipole

$$\mathbf{j}_n(\mathbf{r},t) = \begin{cases} \mathbf{Q}_n(t) & \mathbf{r}=\mathbf{r}_n \\ 0 & \mathbf{r}\neq\mathbf{r}_n \end{cases} \quad n=1,\dots,N \tag{3}$$

to each point of a grid in the volume Ω, leads to

$$b_m(t) = \sum_{n=1}^{N}\mathbf{g}(\mathbf{r}_{s_m},\mathbf{r}_n)\cdot\mathbf{Q}_n(t)\ , \tag{4}$$

where

$$\mathbf{g}(\mathbf{r}_{s_m},\mathbf{r}_n) = \frac{\mu_0}{4\pi F^2(\mathbf{r}_{s_m})}\mathbf{r}_n\times[F(\mathbf{r}_{s_m})\mathbf{e}_{s_m} - (\nabla F(\mathbf{r}_{s_m})\cdot\mathbf{e}_{s_m})\,\mathbf{r}_{s_m}] \tag{5}$$

is named the lead-field vector and

$$F(\mathbf{r}) = |\mathbf{r}-\mathbf{r}_n|(|\mathbf{r}||\mathbf{r}-\mathbf{r}_n| + |\mathbf{r}|^2 - \mathbf{r}_n\cdot\mathbf{r})\ . \tag{6}$$

From (4)-(6) and in a matrix representation, the linear MEG inverse problem can be written as

$$\mathbf{B} = \mathbf{G}\mathbf{J} + \mathbf{N}\ , \tag{7}$$

where $\mathbf{B}$ is the $M\times K$ matrix whose entry B_{mk}, $m=1,\dots,M$; $k=1,\dots,K$ represents the magnetic field $b_m(t_k)$ (see equation (2)); $\mathbf{G}$ is the $M\times 3N$ matrix constructed by the N $M\times 3$ block matrices defined by the three components of the lead-field vector (5); $\mathbf{J}$ is the $3N\times K$ matrix constructed by the N $3\times K$ block matrices $\mathbf{J}_n$ defined by the three components of the vector $\mathbf{j}_n(\mathbf{r}_n,t_k)$; finally $\mathbf{N}$ is

the $M \times K$ matrix whose entries are the noise components affecting the measured magnetic field. In the following we will assume that the columns of $\mathbf{B}$, $\mathbf{N}$ and $\mathbf{J}$ are realizations of stochastic processes whose statistical estimate and variance can be empirically computed assuming that the averages over time can replace the averages over the realizations (ergodicity assumption). Furthermore we assume that the noise is white Gaussian and is not correlated with the signal, which implies, for the ergodicity assumption, that $\mathbf{N}\mathbf{N}^\top \simeq \sigma^2 I$, where σ^2 is the variance of a Gaussian distribution, and $\mathbf{G}\mathbf{J}\mathbf{N}^\top = \mathbf{N}(\mathbf{G}\mathbf{J})^\top \simeq 0$. We point out that the noise in real MEG measurements is not white Gaussian. However *pre-whitening* techniques allow to reduce the actual impact of this assumption at a pre-processing level.

Due to the organization of the human brain, primary currents elicited by external stimuli - or even due to spontaneous activity - usually concentrate in one or more small regions (few mm^2) of the cortex; therefore, it makes sense to assume that the measurements have been produced by a very small (< 10) set of current dipoles. Each one of these P dipoles is uniquely identified by the parameters $\{s_{\mathbf{Q}_p}(t), \mathbf{m}_{\mathbf{Q}_p}(t), \mathbf{r}_{\mathbf{Q}_p}(t)\}_{p=1}^P$ so that

$$\mathbf{j}(\mathbf{r}, t) = \sum_{p=1}^{P} s_{\mathbf{Q}_p}(t)\mathbf{m}_{\mathbf{Q}_p}(t)\delta(\mathbf{r} - \mathbf{r}_{\mathbf{Q}_p}(t)) \ , \tag{8}$$

where $s_{\mathbf{Q}_p}(t)$ is the amplitude of the dipolar signal, $\mathbf{m}_{\mathbf{Q}_p}(t)$ is the dipole orientation and $\mathbf{r}_{\mathbf{Q}_p}(t)$ is the dipole position. Assuming that both the position and the orientation of the dipoles are fixed with respect to the time course, it is straightforward to prove that the matrix representation

$$\mathbf{B} = \mathbf{A}\mathbf{S} + \mathbf{N} \tag{9}$$

holds, where $\mathbf{A}$ is the $M \times P$ matrix with entry

$$A_{mp} = \mathbf{g}(\mathbf{r}_{s_m}, \mathbf{r}_{\mathbf{Q}_p}) \cdot \mathbf{m}_{\mathbf{Q}_p} \ , \tag{10}$$

while $\mathbf{S}$ is the $P \times K$ matrix with entry

$$S_{pk} = s_{\mathbf{Q}_p}(t_k) \ . \tag{11}$$

Also here we assume that $\mathbf{N}\mathbf{N}^\top \simeq \sigma^2\mathbf{I}$ and $\mathbf{N}(\mathbf{A}\mathbf{S})^\top = \mathbf{A}\mathbf{S}\mathbf{N}^\top \simeq 0$. The parameter identification problem of estimating the dipole parameters $\mathbf{r}_{\mathbf{Q}_p}$, $\mathbf{m}_{\mathbf{Q}_p}$ and $s_{\mathbf{Q}_p}(t_k)$ for all p and k from equation (9) is clearly non-linear, since the dependence of A_{mp} on $\mathbf{r}_{\mathbf{Q}_p}$ is non-linear. However, once a method is formulated for determining $\mathbf{A}$, then $\mathbf{S}$ can be obtained through a linear least-squares-based approach.

3. Algorithms

Several algorithms have been applied for solving either the linear inverse problem (7) or the non-linear parameter identification problem (9). On a physiological basis such methods can be divided into two classes. Some approaches [19, 20], typically inspired by the regularization theory for linear ill-posed problems, address the MEG data analysis as an image restoration problem whereby the restored map solves a constrained minimization. The regularized primary current is typically very stable although its support is often too large with respect to a realistic dimension of a typical cortical active region, even when sparsity-enhancing constraints are applied. The second class contains methods which explicitly introduce in the reconstruction procedure the information that the neural currents are small, i.e. pointwise if compared to the sensors' dimension. In the present paper we will compare the

performances of three methods belonging to this second class: a beamformer, a MUltiple SIgnal Classification (MUSIC) algorithm and a particle filter for Bayesian tracking. While MUSIC and particle filters work in a multi-dipole setting and then solve the non-linear inverse problem, beamformers assume the more general model of a continuous current distribution; however, of the many beamformer implementations already available in the MEG community, we chose the "eigenspace projection vector beamformer" [5] which assumes that the number of active regions is very small. This assumption, which is adopted also for MUSIC, the additive representation of noise in (7) and (9) and the fact that noise and noise-free data are not correlated, imply that, in principle, the spectrum of $\mathbf{B}\mathbf{B}^\top$ contains a small number of eigenvalues significantly larger than the noise variance σ^2 and many eigenvalues of the same order of magnitude of σ^2.

In this section, we review the three methods considered in this analysis. Beamforming and MUSIC are well-known approaches to MEG data analysis. In the present paper, which focuses on a comparison of algorithms, only a schematic description of these two methods will be offered, their theoretical formulation being referred to already published papers. On the other hand, particle filter is a relatively new approach to the MEG inverse problem and is therefore described in more detail.

3.1. BEAMFORMERS. This method was originally developed in the radar and sonar signal processing community [22], but later it has been used in different fields, varying from the geophysical [23] to the biomedical area [24] with applications to the MEG inverse problem as well [25]. Beamformers are spatial filters discriminating the signals on the basis of their spatial location. The beamformer output is a weighted linear combination of the measurements, reflecting the source activity in a specified location over time. To introduce the Linearly Constrained Minimum Variance (LCMV) beamformer [4] we partition the $M \times 3N$ lead-field matrix $\mathbf{G}$ into N $M \times 3$ matrices

$$\mathbf{G}_n := [\mathbf{g}(\mathbf{r}_{s_1}, \mathbf{r}_n), \ldots, \mathbf{g}(\mathbf{r}_{s_M}, \mathbf{r}_n)]^\top \ , \tag{12}$$

$n = 1, \ldots, N$, and introduce N $M \times 3$ weight matrices $\mathbf{W}_n = \mathbf{W}(\mathbf{r}_n)$, $n = 1, \ldots, N$. The weight matrices are the key unknowns in beamforming and can be determined by solving the constrained minimum problem

$$\min_{\mathbf{W}_n} \mathrm{var}\hat{\mathbf{J}}_n \quad \text{subject to} \quad \mathbf{W}_n^\top \mathbf{G}_n = \mathbf{I}_3 \tag{13}$$

where

$$\hat{\mathbf{J}}_n = \mathbf{W}_n^\top \mathbf{B}, \tag{14}$$

is again a stochastic process with vectorial values. The motivation for (13), (14) is as follows. Denoting with $E\{\hat{\mathbf{J}}_n\}$ the $3 \times K$ matrix whose entry li is defined as

$$E\{\hat{\mathbf{J}}_n\}_{li} = \frac{1}{K} \sum_{k=1}^{K} (\hat{\mathbf{J}}_n)_{lk} \ \forall i = 1, \ldots, K, \ \forall l = 1, 2, 3 \ , \tag{15}$$

the empirical variance of $\hat{\mathbf{J}}_n$ is defined by

$$\mathrm{var}\hat{\mathbf{J}}_n = \mathrm{Tr}\left[\{(\hat{\mathbf{J}}_n - E\{\hat{\mathbf{J}}_n\})(\hat{\mathbf{J}}_n - E\{\hat{\mathbf{J}}_n\})^\top\}\right] \tag{16}$$

and can be interpreted as a measure of the variability of the stochastic vectorial process $\hat{\mathbf{J}}_n$. Now we partition the $3N \times K$ matrix $\mathbf{J}$ into N $3 \times K$ matrices $\mathbf{J}_n$,

$n = 1, \ldots, N$, and assume that

$$(\mathbf{J}_n - E\{\mathbf{J}_n\})(\mathbf{J}_{n'} - E\{\mathbf{J}_{n'}\})^\top = 0 \tag{17}$$

when $n \neq n'$ (neurophysiologically, this corresponds to assume that all neural currents in the brain are uncorrelated in time) and where $E\{\mathbf{J}_n\}$ is defined coherently with (15). Exploiting the constraint

$$\mathbf{W}_n^\top \mathbf{G}_n = \mathbf{G}_n^\top \mathbf{W}_n = \mathbf{I}_3 \ , \tag{18}$$

one can easily show that the relation between the variances of $\mathbf{J}_n$ and $\hat{\mathbf{J}}_n$ is

$$\mathrm{var}\hat{\mathbf{J}}_n = \mathrm{var}\mathbf{J}_n + \mathrm{var}\left(\mathbf{W}_n^\top \left[\sum_{l \neq n}^{N} \mathbf{G}_l \{(\mathbf{J}_l - E\{\mathbf{J}_l\})(\mathbf{J}_l - E\{\mathbf{J}_l\})^\top\} \mathbf{G}_l^\top \right] \mathbf{W}_n \right) \ . \tag{19}$$

Therefore, finding the weight matrix $\mathbf{W}_n$ which minimizes (19) corresponds to finding the source $\hat{\mathbf{J}}_n$ with the strength closest to the strength of $\mathbf{J}_n$ at the point $\mathbf{r}_n$ in the brain. Of course, a reliable estimate of the positions of the sources is obtained by computing the values of $\mathbf{r}_n$ for which the function $\mathrm{var}\hat{\mathbf{J}}_n$ has a maximum.

Some comments:

- Using Lagrange multipliers, it is easy to show that [4]

$$\mathbf{W}_n = (\mathbf{B}\mathbf{B}^\top)^{-1}\mathbf{G}_n(\mathbf{G}_n^\top(\mathbf{B}\mathbf{B}^\top)^{-1}\mathbf{G}_n)^{-1} \ . \tag{20}$$

- The presence of noise on the measurements can significantly affect the reconstructions giving non-negligible contributions to the estimated source strengths. In our algorithm these artefacts are reduced by means of two additional techniques. First [5], once computed through (20), each matrix $\mathbf{W}_n$ is used only after a projection onto the subspace associated to the eigenvalues of $\mathbf{B}\mathbf{B}^\top$ bigger than σ^2. We note that this projection procedure is meaningful only when the number of sources is small, much smaller than the number of sensors. Second, the inversion of the covariance matrix $\mathbf{B}\mathbf{B}^\top$ is regularized, i.e. $(\mathbf{B}\mathbf{B}^\top)^{-1}$ is replaced by $(\mathbf{B}\mathbf{B}^\top + \lambda\mathbf{I})^{-1}$ where the regularization parameter λ is chosen on the basis of the noise level by means of heuristic procedures.
- Condition (17) is a strong assumption which is often unrealistic from a neurophysiological viewpoint. A discussion of the quantitative impact of this hypothesis on the source reconstruction and a first approach to its relaxation are contained in [21]. Our beamforming algorithm implements this approach to the method.

3.2. MULTIPLE SIGNAL CLASSIFICATION. MUSIC was first developed in the array signal processing community [6]. Its application to the MEG inverse problem [7] is concerned with the non-linear framework of equation (9). Here the MUSIC algorithm provides an estimate of the forward operator $\mathbf{A}$, containing the stationary source parameters (dipole locations and orientations), while the source temporal behaviour $\mathbf{S}$ is estimated afterwards by linear least squares. The key element of the method is the use of the empirical data covariance matrix $\mathbf{B}\mathbf{B}^\top$. It follows easily from the already mentioned assumptions that

$$\mathbf{B}\mathbf{B}^\top = \mathbf{A}\mathbf{S}\mathbf{S}^\top\mathbf{A}^\top + \sigma^2\mathbf{I} \ ; \tag{21}$$

furthermore, the first term at the right side of (21) can be diagonalized:

$$\mathbf{A}\mathbf{S}\mathbf{S}^\top\mathbf{A}^\top = \mathbf{\Phi}\mathbf{\Lambda}_M\mathbf{\Phi}^\top, \tag{22}$$

where $\mathbf{\Lambda}_M$ contains R non-zero and $M-R$ zero eigenvalues, with $R = rank(\mathbf{S})$, and $\mathbf{\Phi} = (\mathbf{\Phi}_S, \mathbf{\Phi}_N)$ can be partitioned into the $M \times R$ matrix $\mathbf{\Phi}_S$ of the eigenvectors associated with the non-zero eigenvalues and the $M \times (M-R)$ matrix $\mathbf{\Phi}_N$ of the eigenvectors associated with the zero eigenvalues. As $\mathbf{S}$ contains the temporal functions of P dipoles, its rank is $R = P$ whenever all time courses are linearly independent, and $R < P$ otherwise. From equation (21) and (22) it follows that, in principle, the number of eigenvalues of $\mathbf{B}\mathbf{B}^\top$ bigger than σ^2 equals to R, the number of independent sources. To actually recover the parameters of the P dipoles, a second diagonalization is useful

$$\mathbf{S}\mathbf{S}^\top = \mathbf{U}\mathbf{\Lambda}_R\mathbf{U}^\top, \tag{23}$$

where $\mathbf{\Lambda}_R$ is the $R \times R$ diagonal matrix of the R non-zero eigenvalues of $\mathbf{S}\mathbf{S}^\top$ and $\mathbf{U}$ is the $P \times R$ matrix whose columns are the eigenvectors associated with the non-zero eigenvalues. Starting from (21), (22) and (23) it can be shown that

$$\mathcal{R}(\mathbf{A}\mathbf{U}) = \mathcal{R}(\mathbf{\Phi}_S), \tag{24}$$

where $\mathcal{R}(\cdot)$ denotes the range of a matrix. Equation (24) is the main result at the basis of the MUSIC algorithm. For independent sources, $\mathbf{U} = \mathbf{I}$ and the space spanned by the eigenvectors $\mathbf{\Phi}_S$ equals the range of the forward operator $\mathbf{A}$; for this reason, $\mathbf{\Phi}_S$ is often referred to as the *signal subspace*. Therefore: (i) the number of eigenvalues of $\mathbf{B}\mathbf{B}^\top$ bigger than σ^2 is an estimate of R; (ii) the subspace generated by the eigenvectors associated with these eigenvalues is an estimate of the signal subspace.

In the MUSIC framework, the number of independent sources is first estimated from the data covariance matrix; then the whole brain is scanned in order to find the locations and orientations which most likely produce the estimated signal subspace. If less than R single dipoles are found, the algorithm searches for linearly dependent sources in vector spaces which are cartesian products of the single dipole space. The RAP-MUSIC algorithm is a recursive application of MUSIC where, after each source is found, its lead field is projected out of the measurements. Two comments:

- in principle, the number of eigenvalues bigger than σ^2 could be determined automatically; however, in real applications this might be a rather difficult task, since the singular spectrum often does not present any clearly visible plateau; this also implies that, in order to keep track of any possible active source, it is often preferable to overestimate the value of R;
- in our implementation of RAP-MUSIC, the set of points and orientations satisfying (24) is determined by using the concept of correlation between subspaces described in [26] and computed by using a technique based on singular value decomposition. In actual applications this technique solves equation (24) only approximately, i.e. the subspace correlation is determined within a given threshold.

3.3. Particle Filter. Particle filters [10, 11] are a class of recently developed algorithms for the numerical implementation of Bayesian filtering [27]: the unknown and the measurements are modeled as stochastic processes and the aim is to sequentially compute the conditional probability density function (pdf) of the unknown, conditioned on the measurements and usually referred to as the posterior pdf. Particle filters have been mainly developed for tracking applications, and have been later applied for solving the MEG inverse problem [13, 14, 28] in a multi-dipole setting characterized by a minimal set of assumptions (for example, here the dipoles'

position and orientation are not kept fixed, as for RAP-MUSIC). Due to its sequential nature, in principle a computational effective particle filter should allow on-line tracking of neural currents from biomagnetic measurements.

In order to describe the formulation, in the following $\pi(\mathbf{x})$ is the probability density function of the random vector $\mathbf{X}$, $\pi(\mathbf{x}|\mathbf{y})$ is the probability density function of $\mathbf{x}$ conditioned on the realization $\mathbf{y}$ of the random vector $\mathbf{Y}$, and $\{\mathbf{B}_t\}_{t=1,\dots,K}$ and $\{\mathbf{J}_t\}_{t=1,\dots,K}$ are the measurements and primary current (current dipoles) processes, respectively. The main assumptions for applying Bayesian filtering are concerned with the Markovian nature of the two processes and can be synthesized in the following equations:

$$\pi(\mathbf{j}_{t+1}|\mathbf{j}_t, \mathbf{j}_{t-1}, ..., \mathbf{j}_1) = \pi(\mathbf{j}_{t+1}|\mathbf{j}_t) \tag{25}$$

$$\pi(\mathbf{b}_t|\mathbf{j}_t, \mathbf{j}_{t-1}, ..., \mathbf{j}_1) = \pi(\mathbf{b}_t|\mathbf{j}_t) \tag{26}$$

$$\pi(\mathbf{j}_{t+1}|\mathbf{j}_t, \mathbf{b}_t, \mathbf{b}_{t-1}, ..., \mathbf{b}_1) = \pi(\mathbf{j}_{t+1}|\mathbf{j}_t) \tag{27}$$

where, for sake of simplicity, notations like $\mathbf{j}_t$ mean $\mathbf{j}(t)$, which is here a collection of dipoles with corresponding strengths $\mathbf{q}(t)$ and positions $\mathbf{r}(t)$. Bayesian filtering is a sequential process which utilizes some initialization to provide the posterior density

$$\pi_{post}(\mathbf{j}_t|\mathbf{b}_{1:t}) := \pi(\mathbf{j}_t|\mathbf{b}_{1:t}) \tag{28}$$

for each time sample $t = 1, \dots, K$. The initializations consist of choosing three density functions: the prior pdf

$$\pi_{pr}(\mathbf{j}_1) := \pi(\mathbf{j}_1) = \pi_{pr}(\mathbf{r}_1)\pi_{pr}(\mathbf{q}_1) \ ; \tag{29}$$

the transition kernel

$$p(\mathbf{j}_{t+1}|\mathbf{j}_t) := \pi(\mathbf{j}_{t+1}|\mathbf{j}_t) \ ; \tag{30}$$

and the likelihood function

$$f(\mathbf{b}_t|\mathbf{j}_t) := \pi(\mathbf{b}_t|\mathbf{j}_t) \ . \tag{31}$$

In our implementation we will assume

$$\pi_{pr}(\mathbf{r}_1) = \frac{1}{m(V)} \tag{32}$$

and

$$\pi_{pr}(\mathbf{q}_1) = \frac{1}{\sqrt{2\pi}\sigma_q} e^{\frac{-|\mathbf{q}_1|^2}{\sigma_q^2}} \ , \tag{33}$$

where $m(V)$ is the measure of the volume V in which the dipoles lie and σ_q is an estimate of the dipole amplitude. Then, the transition kernel

$$p(\mathbf{j}_{t+1}|\mathbf{j}_t) = \frac{1}{\sqrt{2\pi}\det\Sigma} e^{-(\mathbf{j}_{t+1}-\mathbf{j}_t)^T\Sigma^{-1}(\mathbf{j}_{t+1}-\mathbf{j}_t)} \ , \tag{34}$$

where Σ is a diagonal covariance matrix: in the case of a single dipole, the values that define the evolution of the dipole position are stored in $\Sigma_{11} = \Sigma_{22} = \Sigma_{33}$ and the values for the evolution of the dipole moment are in $\Sigma_{44} = \Sigma_{55} = \Sigma_{66}$. Finally, the likelihood function

$$f(\mathbf{b}_t|\mathbf{j}_t) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{\frac{-|\mathbf{b}_t-\mathbf{b}(\mathbf{j_t})|^2}{\sigma_n^2}} \ , \tag{35}$$

where σ_n is estimated from the pre-stimulus.

The Bayesian filtering algorithm applied to the biomagnetic problem is the sequential application of [13, 14]

$$\pi_{post}(\mathbf{j}_t|\mathbf{b}_{1:t}) = \frac{f(\mathbf{b}_t|\mathbf{j}_t)\pi_{pr}(\mathbf{j}_t|\mathbf{b}_{1:t-1})}{\pi(\mathbf{b}_t|\mathbf{b}_{1:t-1})} \ , \tag{36}$$

$$\pi_{pr}(\mathbf{j}_{t+1}|\mathbf{b}_{1:t}) = \int p(\mathbf{j}_{t+1}|\mathbf{j}_t)\pi_{post}(\mathbf{j}_t|\mathbf{b}_{1:t})d\mathbf{j}_t \ , \tag{37}$$

where all the involved pdfs can be determined starting from the previous initializations. In particular, Bayes' theorem (36) determines the posterior pdf, i.e. the solution of the problem in the Bayesian framework, by combining the prior pdf with the likelihood function. The Chapman-Kolmogorov formula (37) is the tool for computing the next prior from the actual posterior.

However, in practice, these two equations cannot be solved analytically and particle filtering provides a computational scheme for their numerical solution. The simplest particle filter is the one known as Sampling Importance Resampling (SIR) particle filter [11]. It is a sequential Monte Carlo method where equation (36) is computed by means of an importance sampling strategy with the prior density $\pi_{pr}(\mathbf{j}_t|\mathbf{b}_{1:t-1})$ playing the role of proposal density at each time point.

Our SIR particle filter can be synthesized in the following steps:

1. Draw a set of N particles $\{\mathbf{j}_1^i\}_{i=1}^N$ from the prior pdf at a first time point, $\pi_{pr}(\mathbf{j}_1)$; we may write

$$\pi_{pr}(\mathbf{j}_1) \simeq \sum_{i=1}^{N} \frac{1}{N}\delta(\mathbf{j}_1 - \mathbf{j}_1^i) \ . \tag{38}$$

2. for $t \geq 0$, let $\{\tilde{\mathbf{j}}_t^i\}_{i=1}^N$ be a sample distributed according to $\pi_{post}(\mathbf{j}_t|\mathbf{b}_{1:t})$; then exploiting the Chapman-Kolmogorov equation (37) we may approximate the next prior pdf as follows:

$$\pi_{pr}(\mathbf{j}_{t+1}|\mathbf{b}_{1:t}) \simeq \frac{1}{N}\sum_{i=1}^{N} p(\mathbf{j}_{t+1}|\tilde{\mathbf{j}}_t^i) \ .$$

3. Sample the prior density by drawing a single particle from the transition kernel $p(\mathbf{j}_{t+1}|\tilde{\mathbf{j}}_t^i)$ for each i; we get

$$\pi_{pr}(\mathbf{j}_{t+1}|\mathbf{b}_{1:t}) \simeq \frac{1}{N}\delta(\mathbf{j}_{t+1} - \mathbf{j}_{t+1}^i) \ . \tag{39}$$

4. Apply the Bayes theorem, i.e. compute the relative weights of the particles:

$$w_{t+1}^i = f(\mathbf{b}_{t+1}|\mathbf{j}_{t+1}^i) \tag{40}$$

and then normalize the weights so that $\sum_{i=1}^N w_{t+1}^i = 1$; an approximation to the posterior pdf is given by

$$\pi_{post}(\mathbf{j}_{t+1}|\mathbf{b}_{1:t+1}) \simeq \sum_{i=1}^{N} w_{t+1}^i\delta(\mathbf{j}_{t+1} - \mathbf{j}_{t+1}^i) \ . \tag{41}$$

5. Resample the sample set representing the posterior density: extract a new set of particles $\{\tilde{\mathbf{j}}_{t+1}^i\}_{i=1}^N$ from the old set $\{\mathbf{j}_{t+1}^i\}_{i=1}^N$, such that the probability of drawing the i-th particle is w_{k+1}^i; we get

$$\pi_{post}(\mathbf{j}_{t+1}|\mathbf{b}_{1:t+1}) \simeq \sum_{i=1}^{N} \frac{1}{N}\delta(\mathbf{j}_{t+1} - \tilde{\mathbf{j}}_{t+1}^i) \ . \tag{42}$$

Then go back to step 2.

The output of the particle filter is a sequence of approximated pdfs. The final step of the method is to compute an estimate of the solution from these pdfs. Here we utilize the conditional mean, defined as

$$\mathbf{j}_t^{CM} = \int \mathbf{j}_t \pi_{post}(\mathbf{j}_t|\mathbf{b}_{1:t}) d\mathbf{j}_t \simeq \sum_{i=1}^{N} w_t^i \mathbf{j}_t^i \ . \tag{43}$$

In order to account for the simultaneous activation of several dipoles, we allow particles to belong to different spaces (single dipole space, double dipole space and so on). The birth or death of a dipole correspond to the transition of particles from a dipole space of dimension $6n$ to a dipole space of dimension $6(n+1)$ or $6(n-1)$.

Now some comments on the implementation:

- In our particle filter, the most time consuming step is the computation of the likelihood. In our implementation, we have notably reduced this cost by letting the particles move only between points of a predefined computational grid (with resolution comparable with the physical resolution of a typical MEG scanner). In this way the forward problem (which is addressed for determining the likelihood function) is computed only once, in correspondence with the points of this grid.
- In our implementation, σ_n is the standard deviation of the pre-stimulus; σ_q is an estimate of the dipole strength and is fixed according to neuroscientific considerations (see the conclusions); $\Sigma_{11} = \Sigma_{22} = \Sigma_{33} = 2$, which corresponds to let each particle move of two grid points at each time step; $\Sigma_{44} = \Sigma_{55} = \Sigma_{66} = \sigma_q/5$, which corresponds to let each dipole amplitude change of a neurophysiologically reasonable amount.
- At least in principle, an increase of the statistical effectiveness can be obtained by applying Rao-Blackwellization [29] to the filtering, i.e. by designing a hybrid process where a particle filter tracks the positions and a Kalman filter reconstructs the amplitudes (indeed a Kalman filter is what Bayesian filtering becomes when the problem is linear and Gaussian). We found that this Rao-Blackwellization procedure provides the same reconstruction accuracy with a smaller number of particles [28]. However, for each particle, the Kalman filter requires the inversion of several matrices, which may be rather time consuming. In this paper we employed a code which does not use Rao-Blackwellization.
- The Markovian properties (25)-(27) are not satisfied by a realistic cortical dynamics. However Bayesian filtering provides optimal solutions to the tracking problem even in the case of generalized Markovian processes of order higher than one, while in the case of more realistic models, it provides reliable approximations of the solution.

4. The comparison

4.1. Experimental conditions. The following experimental situations have been selected for the comparison. We first briefly discuss the neurophysiological motivation of each condition, together with the expected behavior of each algorithm, based on theoretical arguments.

Short-lasting sources. For somatosensory stimuli, activity in the primary somatosensory cortex usually happens about 25 milliseconds after the stimulus and

usually lasts for a few milliseconds [30]. Therefore short-lasting sources are a frequently met (and interesting) neurophysiological situation. We set up the following experimental situation: a single current dipole is placed in the left half of the helmet, its strength being non-zero for only 3, 9 and 15 samples. Gaussian noise is added to obtain a final SNR of around 6 dB at the peak. Both RAP-MUSIC and the beamformer diagonalize $\mathbf{B}\mathbf{B}^\top$ to estimate the eigenvalues associated to the true signal. However, in order that this is a reliable estimate, one needs a notable number K of time samples: using $K \approx 3M$ has been shown to give proper results [4] for independent Gaussianly distributed observations. Particle filters process the measurements in sequence, and indirectly exploit all the past behavior as prior information. All the methods are then expected to give better results as the duration of the source increases.

Low signal-to-noise ratio. MEG measurements are characterized by an extremely low SNR [1]; since the dominant noise source is the brain itself (the neural activity not of interest), it can be difficult or even impossible to get rid of noise. Indeed, while working with stimulus-induced activity one can average across different realizations of the stimulus; however, when this procedure is applied, the exact timing of single responses is lost. Therefore algorithms working with low SNRs are welcome. We set up the following experimental condition: a single current dipole with a given triangular waveform was placed in the right half of the helmet, and different Gaussian noise levels were added to obtain different SNRs in the range from 0.08 dB to 5.37 dB. The estimate of the signal subspace through $\mathbf{B}\mathbf{B}^\top$ deteriorates in the case of low SNR, which in principle leads to a worsening in the estimate of the source parameters from RAP-MUSIC. For the beamformer, which recovers distributed sources and not only single dipoles, the appearance of false positives is expected, although the use of the subspace projection surely reduces this drawback. Particle filters are expected to give worse results as well, although the use of the conditional mean for estimating the current dipole should reduce the variability of the estimate.

Quasi-dependent sources. From the neurophysiological viewpoint, the possibility of correctly recovering linearly dependent time courses is crucial since dependent sources show up in many, even simple experimental conditions (e.g. simple auditory stimuli [31]). We set up the following experimental situation: two current dipoles are used to produce the input data; one dipole is placed about 5 cm on the left with respect to the center of the helmet; the other dipole is placed symmetrically, about 5 cm on the right; the two amplitude waveforms are identical, but the left dipole appears 5, 10 and 15 samples before the right one, i.e., the two sources are not perfectly time-dependent. Noise is then added to the synthetic measurements, up to a final SNR of 12 dB at the peak. RAP-MUSIC and the particle filters are well suited for recovering time-dependent sources. Beamformers require as a technical assumption that the sources are not time correlated.

Orthogonal dipoles. When a current dipole is used to model the activity of a small brain region, its direction is determined by the underlying structure of the brain, and is usually perpendicular to the cortex (the direction of most axons); however, due to the presence of many sulci and gyri, this direction may change abruptly in a very short distance. We set up the following experimental situation: two current dipoles are placed at 1 mm distance in the right half of the MEG helmet; their dipole moments are orthogonal to each other; however, their amplitude waveforms are completely separated, in such a way that there is at most one active

dipole at each time sample. Gaussian noise is added to obtain a final SNR of 6 dB at the peak. Both the beamformer and especially RAP-MUSIC rely on an estimate of the signal subspace from $\mathbf{B}\mathbf{B}^\top$. Since two orthogonal sources produce opposite spatial correlation between the same pairs of sensors, in $\mathbf{B}\mathbf{B}^\top$ some contributions are canceled and the estimate of the signal subspace is wrong.

Multi-target case. External stimulations often evokes several activities in the same time range but in different cortical regions [31]. We have realized an experiment where four triangular waveforms are placed in four rather distant brain regions. Each waveform becomes active two time samples after the activation of the previous waveform. Gaussian noise is added to the computed biomagnetic field up to a SNR of 11 dB. As already discussed, beamformers are not effective in the case of correlated or quasi-correlated sources. In RAP-MUSIC the overestimate of the dimension of the signal subspace allows the reconstruction of several sources in the same time range. Finally, particle filters are specifically designed in a multi-target framework. In particular, the fact that the sources become active sequentially in time, assures the effectiveness of the conditional mean as an estimate of the solution.

Moving dipole. When nearby brain regions are active in sequence the resulting field can be interpreted as that of a single moving dipole. We set up the following experimental condition: a single dipole moves along a linear path with a speed of 1 millimeter per time sample; Gaussian noise is added to the measurements to a final SNR of 18 dB at the peak. RAP-MUSIC relies on the assumption of fixed dipole positions, and is expected to fail in recovering the moving source. The general beamformer algorithm is expected to correctly recover the activity of moving sources, although the use of the subspace projection (which may help in the low SNR case) may create problems here. Particle filters have been mainly developed for tracking moving objects, therefore they are expected to correctly track the moving source.

Real Data Set. We finally consider a real data set: we chose to work with auditory stimuli as they usually elicit activations in both hemispheres with a high temporal correlation. We recorded MEG measurements at the Low Temperature Laboratory, Helsinki University of Technology, Finland. Auditory responses were measured in a single healthy subject (male, 44 years old), who signed an informed consent before the MEG recording. The recording had a prior approval by the Helsinki-Uusimaa ethics committee. Auditory stimuli were 1-kHz tone pips within a 100-ms Hanning window, and they were randomly presented either to the left or right ear at a clearly audible yet comfortable level. The MEG data were acquired with a 306-channel whole-head neuromagnetometer (Elekta Neuromag Oy, Helsinki, Finland), which employs 102 sensor elements, each comprising one magnetometer and two planar gradiometers. Measurements were filtered to 0.1–170 Hz and sampled at 600 Hz. Prior to the actual recording, four small indicator coils attached to the scalp at locations known with respect to anatomical landmarks were energized and the elicited magnetic fields recorded to localize the head with respect to the MEG sensors and thus to allow subsequent co-registration of the MEG with anatomical MR images. Epochs with exceedingly large ($b > 3$ pT/cm) MEG signal excursions were rejected, and about 100 artifact-free trials for each stimulus category were collected and averaged on-line in the window $[-100, 500]$ ms with respect to the stimulus onset. Residual environmental magnetic interference was subsequently removed from the averages using the signal-space separation method. Prior to the analysis, the evoked responses were low-pass filtered at 40 Hz. An

3 time samples			
	position (cm)	orientation (rad)	amplitude (%)
beamformer	0.0	0.02	7.6
RAP-MUSIC	0.0	0.0	0.0
particle filter	0.4	0.05	18.6
9 time samples			
	position (cm)	orientation (rad)	amplitude (%)
beamformer	0.0	0.00	6.2
RAP-MUSIC	0.0	0.00	0.0
particle filter	0.5	0.04	14.9
15 time samples			
	position (cm)	orientation (rad)	amplitude (%)
beamformer	0.0	0.00	4.4
RAP-MUSIC	0.0	0.00	0.0
particle filter	0.2	0.03	6.2

Table 1. Reconstruction error of position, orientation and amplitude of the three algorithms in the case of a source lasting 3,9 and 15 time samples.

anatomical T1-weighted MR-image of the subject was obtained with a 1.5-Tesla Siemens Sonata MRI system using a standard 3D anatomical sequence with 1-mm cubical voxels.

4.2. Numerical results. For each experimental situation, we now quantitatively describe the performances of each algorithm by means of figures or tables. From a technical viewpoint, we notice that the beamformer and RAP-MUSIC do not depend on any initialization. In order to assess the robustness of the particle filter with respect to initialization of the particles' set, for each experiment ten runs of the algorithm are performed. Each particle filter reconstruction corresponds to the average over these ten realizations of position, orientation and strength of the conditional mean. Since we found that the conditional mean is very stable with respect to this initialization (the standard deviation is systematically smaller than 1% of the mean value), the error related to these realizations is not considered.

Short-lasting sources (Table 1). Despite the use of a small number of samples for estimating the signal subspace, the reconstructions of the beamformer and of RAP-MUSIC are good. The particle filter gives acceptable results as well, although less accurate. For all three methods the reconstruction accuracy increases if the sources are active for a larger time range. However this experiment utilizes a rather high SNR; the deterioration of the restoration accuracy with decreasing SNR speeds up for short-lasting sources, even for RAP-MUSIC and beamforming.

Low signal-to-noise ratio (Figure 1). We present here the reconstruction errors on position, orientation and amplitude for different amount of noise affecting the data. Once again, RAP-MUSIC and beamforming recover the source position and orientation highly accurately while the particle filter's reconstruction shows a certain spread around the true position. For all algorithms the source strength is highly corrupted by noise. In particular the beamformer notably underestimates

	5 samples	10 samples	15 samples
beamformer	0 samples	0 samples	0 samples
RAP-MUSIC	5 samples	9 samples	14 samples
particle filter	11 samples	7 samples	19 samples

TABLE 2. Distance between the peaks of the reconstructed sources for the three algorithms in the case where the theoretical distance is 5,10 and 15 samples.

the source peak. At high noise levels, the amplitude provided by the particle filter is certainly more accurate with respect to the other two methods.

Quasi-dependent sources (Table 2). In Table 2 the distances between the peaks of the two restored waveforms are measured when the distances between these peaks in the theoretical amplitudes are of 5, 10 and 15 samples. Beamforming appears to be highly ineffective in recovering time correlation while RAP-MUSIC results to be the most 'time-resolving' method. However, in order to obtain such accurate reconstructions the rank of the signal subspace in RAP-MUSIC must be systematically overestimated by means of some heuristic procedure. The particle filter, on the contrary, can reliably estimate the degree of time dependence in a more automatic fashion.

Orthogonal dipoles (Figure 2). As expected, RAP-MUSIC fails in estimating the source orientations. As a consequence of the wrong orientation, RAP-MUSIC also recovers wrong time-courses: instead of two sources active in sequence it recovers two contemporarily active dipoles (see Figure 2(g) and (h)). However this incorrect estimate of the signal subspace is not so dramatic to deteriorate the accuracy of the reconstructions provided by the beamformer. As expected, the particle filter is very reliable in this case.

Multi-target case (Table 3). Table 3 contains the reconstruction errors for the four sources, which show that RAP-MUSIC and the particle filter are particularly effective in this setting. In particular, the beamformer completely fails to detect one of the four dipoles and is not accurate in the case of two other dipoles. RAP-MUSIC and the particle filter detect all sources with a higher accuracy.

Moving dipole (Figure 3). Beamformers and RAP-MUSIC explain the field of a moving source as the superposition of the fields of two and three sources, respectively; on the contrary, the tracking nature of the particle filter algorithm allows to follow the movements of the source. However we observe that this experiment may be misleading, since the sources reconstructed by the beamformer and the MUSIC algorithm are able to reproduce the data with the same accuracy as the sources tracked by the particle filter. In fact, the result in this figure may be considered as an on-action demonstration of the inherent non-uniqueness of the MEG inverse problem.

Real Data Set (Figure 4 and Table 4). The two highly correlated sources activated by the auditory stimuli are correctly recovered by the particle filter and RAP-MUSIC; around the peak of activity (at a latency of about 100 ms after stimulus onset), the two algorithms provide approximately the same localization, which is also in accordance with the known brain topography; also the time courses of the reconstructed sources are similar; indeed, the reconstructions from the particle filter are allowed to move during the measurements' sequence, which explains

dipole 1			
	position (cm)	orientation (rad)	amplitude (%)
beamformer	1.0	0.27	39.4
RAP-MUSIC	0.5	0.22	12.9
particle filter	0.5	0.23	26.9
dipole 2			
	position (cm)	orientation (rad)	amplitude (%)
beamformer	not detected	not detected	not detected
RAP-MUSIC	0.8	0.39	49.2
particle filter	1.8	0.93	37.9
dipole 3			
	position (cm)	orientation (rad)	amplitude (%)
beamformer	0.0	0.54	30.8
RAP-MUSIC	0.63	0.16	19.9
particle filter	0.5	0.23	20.5
dipole 4			
	position (cm)	orientation (rad)	amplitude (%)
beamformer	2.0	0.54	29.5
RAP-MUSIC	0.45	0.83	38.3
particle filter	0.54	0.59	31.4

Table 3. Reconstruction error of position, orientation and amplitudes of the three algorithms in the case of four sources. The beamformer fails to reconstruct one of the dipoles.

the difference in the time courses for longer latencies. Differently, the source localization provided by the beamformer is slightly far from the expected position (about 1 cm); even more important, the amplitude of one of the two sources is notably underestimated probably due to the high correlation level between the two sources. To give a quantitative assessment of the reconstruction accuracy provided by the three methods, in Table 4 we compare the parameters provided by the beamformer, RAP-MUSIC and the particle filter with the ones provided by a standard dipole fitting procedure [1]. These reconstruction errors confirm the difficulty of the beamformer to restore an overall correct shape of one of the two dipoles and the fact that RAP-MUSIC and the particle filter work reasonably well, particularly in reconstructing the correct position and orientation.

5. Discussion

The effectiveness of beamforming, RAP-MUSIC and particle filtering in the analysis of both synthetic and real MEG time series can be assessed according to the following issues:

- **Reconstruction accuracy.** RAP-MUSIC is perhaps the most effective method as far as this aspect is considered. It is mostly accurate in the case of short-lasting sources, high noise levels (in the reconstruction of the source localization and orientation) and high time correlation degree. The particle filter is intrinsically a tracking algorithm for multi-target conditions and indeed,

dipole 1			
	position (cm)	orientation (rad)	amplitude (%)
beamformer	1.3	0.28	79.4
RAP-MUSIC	1.3	0.75	61.2
particle filter	1.1	0.21	46.2
dipole 2			
	position (cm)	orientation (rad)	amplitude (%)
beamformer	1.4	0.28	42.4
RAP-MUSIC	1.3	0.23	51.6
particle filter	1.1	0.28	45.1

TABLE 4. Reconstruction error of position, orientation and amplitudes of the three algorithms in the case of a real data (auditory stimuli). The errors are computed with respect to sources provided by a standard dipole fitting procedure.

together with RAP-MUSIC works fine in reconstructing four simultaneously active dipoles and outperforms the other two methods in reconstructing the two auditory real sources. In general, the beamformer's performance is critical in the case of multi-target time correlated constellations.

- **Computational effectiveness.** The computational cost of the three methods cannot be compared in a direct fashion. In fact the beamformer and RAP-MUSIC perform the analysis within a computational time which is independent of the length of the time series (while it depends on the number of sensors and on the number of grid points) while in the particle filter the computational cost depends on the number of particles utilized and on the number of time samples. For the helmet used at the Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology (which contains 306 channels) an analysis to detect a single source is performed by the beamformer within 26 s and by RAP-MUSIC within 22 s, with a 3 GHz CPU and for a computational grid of $36 \times 36 \times 36$ points. The particle filter employs around 2 s for the analysis of one time sample using 100000 particles, which means that a typical experimental time series of 400 time samples is processed in a computational time almost 40 times larger than for the other two methods (however, we notice that 100000 particles are certainly redundant and reliable reconstructions can be obtained even with 50000 particles). We point out that the computational cost for RAP-MUSIC increases when the method searches for time-correlated sources. In conclusion, it is clear that the beamformer and RAP-MUSIC are the most rapid methods. Despite this, the sequential nature of the particle filter makes it the best candidate for on-line tracking of neural sources. This is why investigation of effective implementations of the filter is worthwhile.
- **Automation degree.** All three methods need optimal a priori selection of some input parameters. The beamformer has two input parameters, i.e. the regularization parameter in the inversion of $\mathbf{R_b}$ and the dimension of the signal subspace. Heuristic analysis shows that reliable values for this parameter change with the noise level, but, to our knowledge, no automatic selection rule can be formulated for fixing it. An analogous situation holds for RAP-MUSIC:

again here there are two input parameters, the dimension of the signal subspace and the threshold for the subspace correlation. For the first parameter, applications show that heuristic recipes for the overestimation of this dimension must be looked for in the case of correlated or quasi-correlated signals. As for the threshold, in principle it should be one but this value must be tuned to a smaller number (down to 0.9) to obtain reasonable reconstructions. In the case of the particle filter, the input parameters are essentially two: the standard deviation in the likelihood and the standard deviation in the prior. For σ_n a reliable estimate is the standard deviation of the pre-stimulus. For σ_q, we observe that the strength of the dipole moment is proportional to the strength of the measured magnetic field and can be easily determined from it using an average position for neurophysiologically plausible source positions.

- **Generality degree.** The particle filter is probably the most general method. It can be applied with no a priori assumption on the number of sources, their time-correlatedness or position.

All three methods can be improved from the viewpoint of both the formulation and the implementation. In particular, the beamformer is not effective for reconstructing correlated sources, while, in the particle filter, some investigation is still necessary for estimating the solution from particles, since the conditional mean is less effective when the number of source dipoles increases.

Acknowledgments. Lauri Parkkonen of the Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology, is kindly acknowledged for useful discussions and for providing us the auditory data. This work is partly supported by a PRIN grant of the italian Ministero dell'Università e della Ricerca, by a INFM 'Seed Project' and by a grant of the Fondazione della Cassa di Risparmio di Verona.

References

[1] R. Hari, M. Hämäläinen, J. Knuutila, O.V. Lounasmaa, "Magnetoencephalography: theory, instrumentation and applications to non-invasive studies of the working human brain", *Rev. Mod. Phys.*, vol. 65, pp. 2-60, 1993.

[2] J. Sarvas, "Basic mathematical and electromagnetic concepts of the bomagnetic inverse problem" *Phys. Med. Biol.*, vol. 32, pp. 11-22, 1987.

[3] B. Van Veen and K. Buckley, "Beamforming: A versatile approach to spatial filtering", *IEEE ASSP Magazine*, vol. 5, pp. 4-24, 1988.

[4] B. Van Veen, W. van Drongelen, M. Yuchtman and A. Suzuki "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering" *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 867-880, 1997.

[5] K. Sekihara, S. Nagarajan, D. Poeppel, A. Marantz and Y. Miyashita " Application of an meg eigenspace beamformer to reconstructing spatio-temporal activities of neural sources", *Human Brain Mapping*, vol. 15, pp. 199-215, 2002.

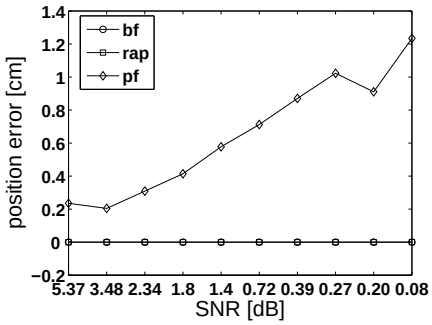
[6] R. O. Schmidt, "Multiple emitter location and signal parameter estimation", *IEEE Trans. Antenn. Propogat.*, vol. 34, pp. 276-280, 1986.

[7] J. Mosher, P. Lewis and R. Leahy, " Multiple dipole modeling and localization from spatio-temporal meg data" *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 541-557, 1992.

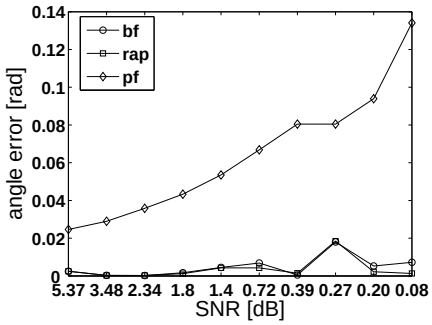
[8] J. Mosher and R. Leahy, "Recursive music: A framework for eeg and meg source localization", *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 1342-1354, 1998.

[9] J. Mosher and R. Leahy "Source localization using recursively applied and projected (rap) music", *IEEE Trans. Signal Proc.* vol. 47, pp. 332-340, 1999.

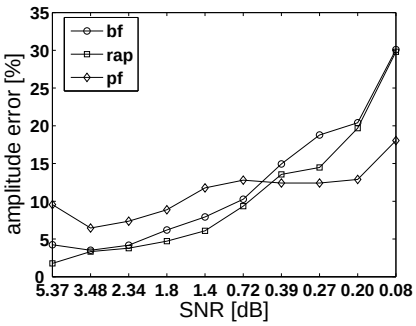
[10] A. Doucet, S. Godsill and C. Andrieu "On sequential Monte Carlo sampling methods for Bayesian filtering", *Statistics and Computing*, vol. 10, pp. 197-208, 2000.
[11] M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking" *IEEE Trans. Signal Proc.*, vol. 50, pp. 174-188, 2002.
[12] A. Doucet, "On sequential simulation-based methods for bayesian filtering", *Technical Report, Univ. of Cambridge*, 1998.
[13] E. Somersalo, A. Voutilainen and J. P. Kaipio, "Non-stationary magnetoencephalography by Bayesian filtering of dipole models", *Inverse Problems*, vol. 19, pp. 1047-1063, 2003.
[14] A. Sorrentino, L. Parkkonen and M. Piana " Particle filters: a new method for reconstructing multiple current dipoles from MEG data", in Proc. of the International Conference on Biomagnetism (Biomag 2006), vol. 1300, pp. 173-176, 2007.
[15] A. Pascarella, A. Sorrentino, M. Piana and L. Parkkonen, "Particle filters and rap-music in meg source modelling: a comparison", in Proc. of the International Conference on Biomagnetism (Biomag 2006), vol. 1300, pp. 161-164, 2007.
[16] R. Kress, L. Kuhn and R. Potthast, "Reconstruction of a current distribution from its magnetic field", *Inverse Problems*, vol. 18, pp. 1127-1146, 2002.
[17] J. Cantarella, D. De Turck and H. Gluck, "The Biot-Savart operator for application to knot theory, fluid dynamics and plasma physics", *J. Math. Phys.*, vol. 42, pp. 876-905, 2001.
[18] M. Hamalainen and J. Sarvas "Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data", IEEE *Trans. Biomed. Eng.*, vol. 36 , pp. 165-171, 1989.
[19] M. Hamalainen and R. Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates", *Med. Biol. Eng. Comput.*, vol. 32, pp. 35-42, 1994.
[20] K. Uutela, M. Hamalainen and E. Somersalo, "Visualization of magnetoencephalographic data using minimum current estimate", *Neuroimage*, vol. 10, pp. 173-180, 1999.
[21] S. S. Dalal, K. Sekihara and S. S. Nagarajan, "Modified beamformers for coherent source region suppression", *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 1357-1363, 2006.
[22] E. Brookner, "Phased-array radar", *Scientific American*, vol. 252, 94-102, 1985.
[23] S. Haykin, J. H Justice, N. L. Owsley,J. L. Yen and A. C. Kac, *Array Signal Processing.* New York: Prentice Hall, 1985.
[24] W. Gee, S. Lee, N. K. Bong, C. A. Cain, R. Mittra and R. L. Magin, "Focused array hyperthermia applicator: theory and experiment" *IEEE Trans. Biomed. Eng.*, vol. 31, pp. 38-45, 1984.
[25] B. Van Veen, J. Joseph and K. Hecox, "Localization of intra-cerebral sources via linearly constrained minimum variance spatial filtering" in Proc. of the IEEE Workshop on Statistical Signal and Array Processing, pp. 526-529, 1992.
[26] G. Golub and C. V. van Veen, *Matrix Computation*, Baltimora: The Johns Hopkins University Press.
[27] E. Somersalo and J. Kaipio, *Statistical and Computational Inverse Problems* Berlin: Springer Verlag, 2004.
[28] C. Campi, A. Pascarella, A. Sorrentino and M. Piana, "A Rao-Blackwellized particle filter for magnetoencephalography", *Inverse Problems*, vol. 24, 025023, 2008.
[29] G. Casella and C. Robert, "Rao-blackwellisation of sampling schemes", *Biometrika*, vol. 83, pp. 81-94, 1996.
[30] K. Iramina, H. Kamei, M. Yumoto and S. Ueno, "Effects of repetition rate of electric stimulation on meg and fmri signals", *IEEE Trans. Magnetic.*, vol. 37, pp. 2918-2920, 2001.
[31] R. Hari, *Magnetoencephalography in Clinical Neurophysiological Assessment of Human Cortical Funcions*, E. Niedermeyer, F. Lopes da Silva, Eds. Baltimore, MD: Lippincott, Williams and Wilkins, 2004, pp. 1165-1197.



(a)



(b)



(c)

FIGURE 1. Low signal-to-noise ratio: reconstruction error for the position (a), orientation (b) and amplitude (c) of the three algorithms in the case of the data with different signal-to-noise ratio.

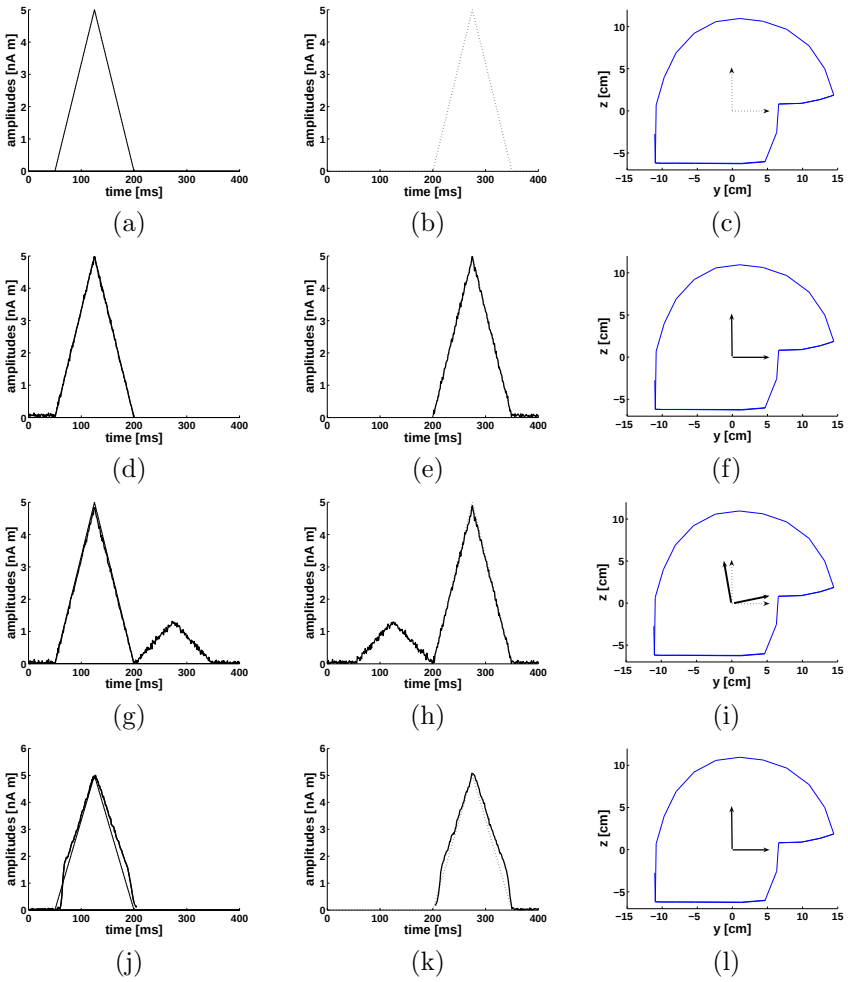


FIGURE 2. Orthogonal dipoles: amplitudes (a),(b) and projection on the sagittal plane (c) of the original dipoles; reconstructed amplitudes (d),(e) and projections (f), obtained with the beamformer, with RAP-MUSIC (shown in (g),(h) and (i)) and with the particle filter (shown in (j),(k) and (l)).

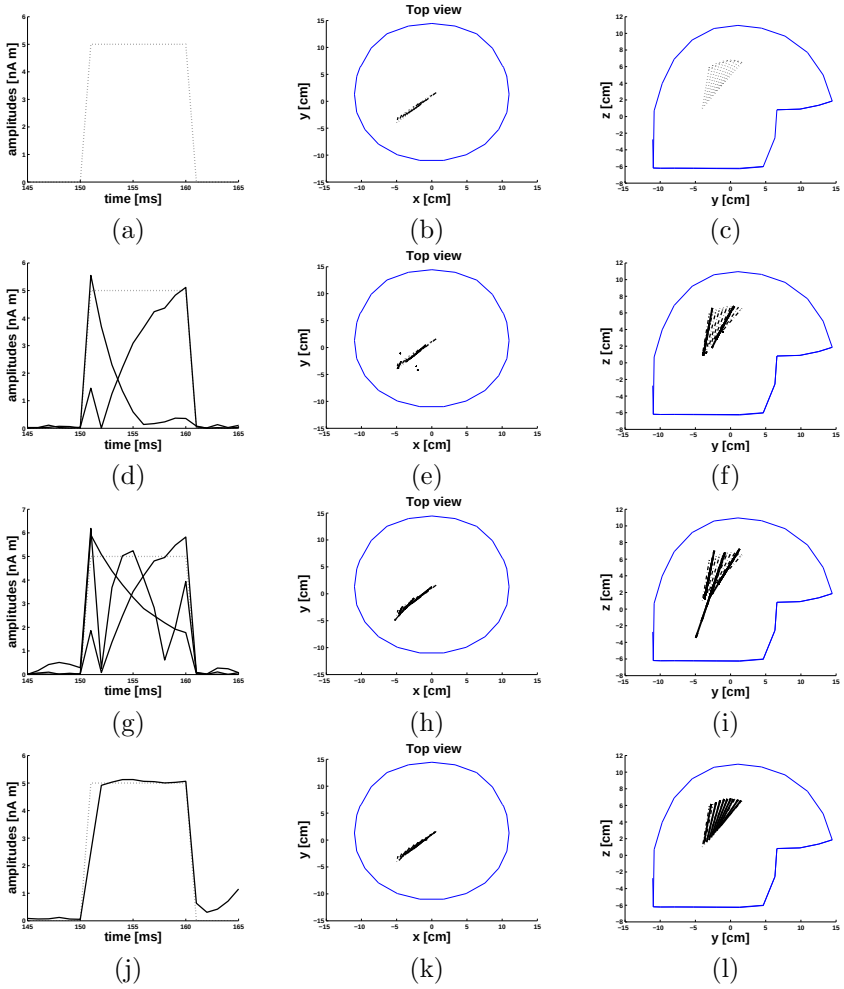


FIGURE 3. Moving dipole: amplitude (a) and projections on the axial (b) and sagittal planes (c) of the original dipoles; reconstructed amplitude (d) and reconstructed projections on the axial (e) and sagittal (f) planes, obtained with the beamformer, with RAP-MUSIC (shown in (g),(h) and (i)) and with the particle filter (shown in (j),(k) and (l)).

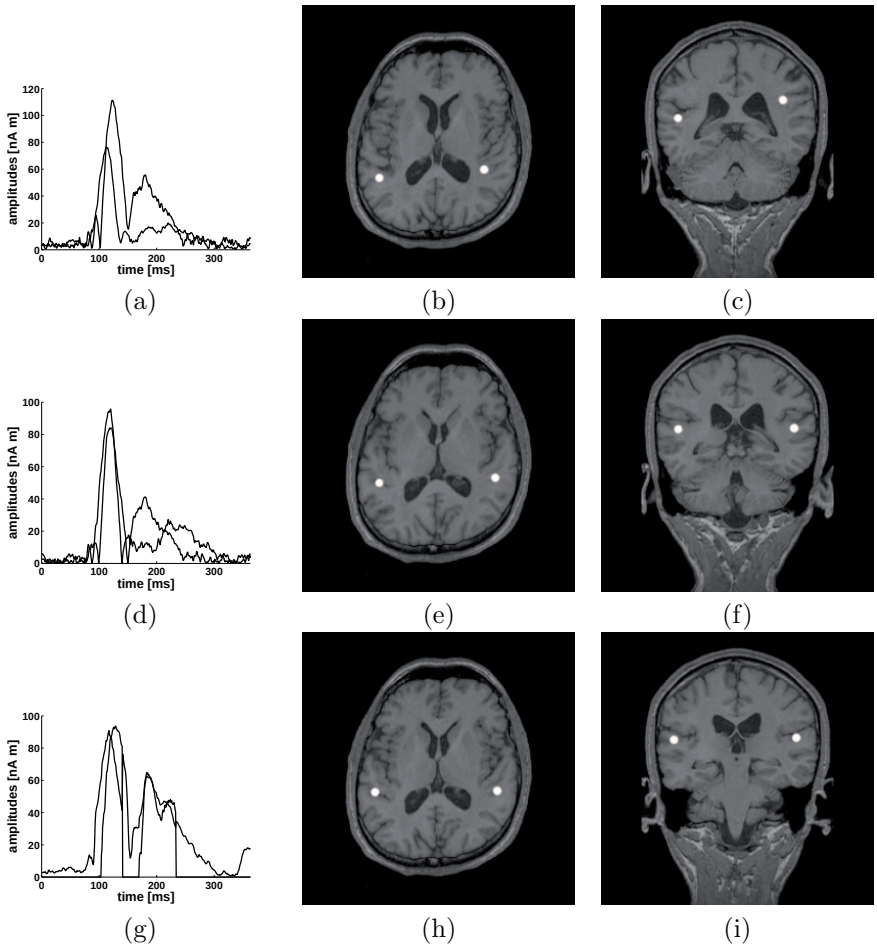


FIGURE 4. Real data: reconstructed amplitudes with beamformers (a), RAP-MUSIC (d) and particle filters (g); coregistation on a Magnetic Resonance high resolution axial view (b) and coronal view (c) of the sources obtained with the beamformer; (e),(f) and (h),(i) show the same objects for, respectively, RAP-MUSIC and the particle filter.

Title:
Particle filtering, beamforming and multiple signal classification for the analysis of magnetoencephalography time series: a comparison of algorithms

Running title:
Particle filtering, beamforming and multiple signal classification for MEG

Corresponding author:
Michele Piana
Dipartimento di Informatica, Università di Verona, Verona, Italy
CNR - INFM LAMIA, Genova, Italy
e-mail: michele.piana@univr.it