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

A. Pascarella§, A. Sorrentino‡, C. Campi§, M. Piana†
§Dipartimento di Matematica, Università di Genova, via Dodecaneso 35, I-16146 Genova, Italy
‡Dipartimento di Fisica, Università di Genova, via Dodecaneso 33, I-16146 Genova, Italy
†Dipartimento di Informatica, Università di Verona, Ca` Vignal 2, Strada le Grazie 15, 37134 Verona, Italy

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 time course of the active sources. Finally, a real data set evoked by a simple auditory stimulus is considered.

Introduction
Magnetoencephalography (MEG) (Hamalainen, Hari, Knuutila, & Lounasmaa, 1993) is a functional imaging modality which allows measurements of the magnetic fields associated with spontaneous or stimulus-induced neural currents by means of a gradiometer arrangement based on SQUID technology. 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 locations outside the skull. This problem is difficult for many reasons. First, it is numerically unstable, whereby such an instability is the direct consequence of the fact that the neuromagnetic inverse problem, as most inverse problems of interest in the applied sciences, is ill-posed in the sense of Hadamard (Sarvas, 1987). 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 nature of the MEG signal is peculiar: for example, the order of magnitude for the neuromagnetic field is some hundreds of femtoTesla, i.e. three orders of magnitude less than the magnetic signal associated with the cardiac activity and nine orders less than the magnetic field of the Earth; furthermore, the biomagnetic field recorded during experiments designed to investigate specific transients or oscillatory responses also contains components associated with spontaneous cortical activity, extraneous to the external stimulation; finally, the recording process introduces typical high frequency noise components with Gaussian distribution. Large spurious signals extraneous to the brain are eliminated by using shielded rooms and special arrangements for the gradiometers (Hamalainen et al., 1993); the uninteresting field components can be eliminated by appropriate averaging or pre-whitening (Sekihara, Poeppel, Marantz, Koizumi, & Miyashita, 1997) and the effects of the recording noise can be reduced applying regularization techniques in the inversion procedure (Somersalo & Kaipio, 2004). However it remains true that the neuromagnetic problem is a dynamical, numerically unstable, computationally heavy inverse problem whereby the input data is characterized by a low (or extremely low) signal-to-noise ratio (SNR).

The aim of the present paper is to provide some insights toward the identification of the optimal method for the solution of the neuromagnetic problem. In such an effective method the following features should in principle coexist. From a neurophysiological viewpoint the localization properties 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 proper-
ties: a beamforming algorithm (Van Veen & Buckley, 1988; Van Veen, Drongelen, Yuchtman, & Suzuki, 1997; Sekihara, Nagarajan, Poeppel, Marantz, & Miyashita, 2002) which spatially filters the signal to focus the source as a weighted combination of the measurements; multiple signal classification (Schmidt, 1986; Mosher, Lewis, & Leahy, 1992; Mosher & Leahy, 1998, 1999), which identifies the source positions by scanning the brain volume to find the solution of a non-linear optimization problem; a particle filter (Arulampalam, Maskell, Gordon, & Clapp, 2002; Doucet, 1998; Somersalo, Voutilainen, & Kaipio, 2003; Sorrentino, Parkkonen, & Piana, 2007; Pascalella, Sorrentino, Piana, & Parkkonen, 2007), 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 rigorous mathematical terms. 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.

The MEG inverse problem

Within the quasi-static approximation (Sarvas, 1987), the magnetoencephalography problem is defined by the Biot-Savart equation:

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

where $b(r, t)$ is the magnetic field produced at time $t$ by the current density $j^{tot}(r', t)$ inside the conductor volume $\Omega$ representing the brain, and $\mu_0$ is the magnetic permeability 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 $e_m$ is the unit vector orthogonal to the sensor’s surface at location $r_{sm}$,

$$b_m(t) = b(r_{sm}, t) \cdot e_m \quad m = 1, \ldots, M$$

is the dynamical measurement on the $m$-th sensor. Recovering the electrical current density inside $\Omega$ 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 (Kress, Kuhn, & Potthast, 2002; Cantarella, Turck, & Gluck, 2001). 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 $j^{tot}$ inside the brain is in fact the sum of two terms (Sarvas, 1987): the primary current $j$ flowing in the active neurons, and the passive current $j^{tot}$ 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 $j$: in general, numerical solutions of integral equations are needed to account for their presence. However, two artificial but reliable assumptions can notably simplify the problem. First, when cortical regions other than the frontal lobe are involved in the analysis, it is reliable (Hamalainen & Sarvas, 1989; Scherg, 1990) to approximate the head as a spherical homogeneous conductor. Second, a discretization of $j(r, t)$ can be obtained by attaching a current dipole $j_n(r, t) = Q_n(t) \delta(r - r_n)$, $n = 1, \ldots, N$ (3) to each point of a grid in the volume $\Omega$. Using these two approximations in (1) and (2) leads to

$$b_m(t) = \sum_{n=1}^{N} g(r_{s_m}, r_n) \cdot Q_n(t) \ ,$$

(4)

where

$$g(r_{s_m}, r_n) = \frac{\mu_0}{4\pi F^2(r_m)} r_n \times [F(r_m)e_m - (\nabla F(r_m) \cdot e_m) r_{s_m}]$$

(5)

is named the lead-field vector and

$$F(r) = |r - r_n| (|r| |r - r_n| + |r|^2 - r_n \cdot r) \ .$$

(6)

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

$$B = GJ + N \ ,$$

(7)

where $B$ is the $M \times K$ matrix whose entry $b(m, k)$, $m = 1, \ldots, M$; $k = 1, \ldots, K$ represents the magnetic field recorded by the SQUID placed at $r_{s_m}$ at time $t_k$; $G$ 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); $J$ is the $3N \times K$ matrix constructed by the $N 3 \times K$ block matrices defined by the three components of the vector $j_n(r_n, t_k)$; finally $N$ is the $M \times K$ matrix whose entries are the noise components affecting the measured magnetic field. In the following we will assume $NN^\top = \sigma^2 I$, where $\sigma^2$ is the variance of a Gaussian distribution, and $GJN^\top = N(GJ)^\top = 0$, i.e. noise and signal are not correlated.

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 
\( \{s_p(t), m_p(t), r_{Q_p}(t)\}_{p=1}^{P} \) so that

\[
j(r, t) = \sum_{p=1}^{P} s_p(t) m_p(t) \delta(r - r_{Q_p}(t)),
\]

where \( s_p(t) \) is the amplitude of the dipolar signal, \( m_p(t) \) is the dipole orientation and \( r_{Q_p}(t) \) 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

\[
B = AS + N
\]

holds, where \( A \) is an \( M \times P \) matrix whose entry \( a(m, p) \) only depends on the independent parameters \( r_{Q_p}, m_p \) while \( S \) is the \( P \times K \) matrix whose entry \( s(p, k) \) is defined by \( s(p, k) = s_p(t_k) \). Also here we assume that \( NN^\top = \sigma^2 I \) and \( N(AS)^\top = ASN^\top = 0 \).

The parameter identification problem of estimating the dipole parameters \( r_{Q_p}, m_p \) and \( s_p(t_k) \) for all \( p \) and \( k \) from equation (9) is clearly non-linear, since the dependence of \( a(m, p) \) on \( r_{Q_p} \) is non-linear. However, once a method is formulated for determining \( A \), then \( S \) can be obtained through a linear least-squares-based approach.

**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 (Hamalainen & Ilmoniemi, 1994; Uutela, Hamalainen, & Somersalo, 1999), typically inspired by the regularization theory for linear ill-posed problems, addresses 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. 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, a beamformer, a MUltiple SIgnal Classification (MUSIC) algorithm and a particle filter for Bayesian tracking, belonging to this second class. 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” (Dalal, Sekihara, & Nagarajan, 2006) which, in order to project on the signal subspace, assumes that the number of active regions is very small. This assumption is adopted also for MUSIC, and together with the additive representation of noise in (7) and (9) and with the fact that noise and noise-free data are not
correlated, implies that the spectrum of $\mathbf{B B}^\top$ contains a small number of eigenvalues significantly larger than the noise variance $\sigma^2$ and many eigenvalues of the same order of magnitude of $\sigma^2$.

In this section, we review the three methods which are able to face the most critical experimental situations, in principle with a notable degree of automation.

**Beamformers**

This method was originally developed in the radar and sonar signal processing community (Brookner, 1985), but later it has been used in different fields, varying from the geophysical (Haykin, Justice, Owsley, Yen, & Kak, 1985) to the biomedical area (Gee et al., 1984) with applications to the MEG inverse problem as well (Van Veen, Joseph, & Hecox, 1992). 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 describe the Linearly Constrained Minimum Variance (LCMV) beamformer (Van Veen et al., 1997) we decompose the $M \times 3 \times N$ lead-field matrix $\mathbf{G}$ into $N M \times 3$ matrices $\mathbf{G}_n = \mathbf{G}(\mathbf{r}_n)$, $n = 1, \ldots, N$ 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

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

(10)

where

$$
\hat{\mathbf{J}}_n = \mathbf{W}_n^\top \mathbf{B}.
$$

(11)

The motivation for (10), (11) is as follows. Denoting with $E\{\mathbf{J}_n\}$ the $3 \times K$ matrix whose entry $(l, i)$ is defined as

$$
E\{\hat{\mathbf{J}}_n\}(l, i) = \frac{1}{K} \sum_{k=1}^{K} \hat{\mathbf{J}}_n(l, k) \quad \forall i = 1, \ldots, K,
$$

(12)

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

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

(13)

and can be interpreted as a measure of the strength of the stochastic vectorial process $\hat{\mathbf{J}}_n$. Now we decompose the $3N \times K$ matrix $\mathbf{J}$ into $N 3 \times K$ matrices $\mathbf{J}_n$, $n = 1, \ldots, N$, and assume that

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

(14)

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 (12). Exploiting the constraint

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

(15)
one can easily show that the relation between the variances of $J_n$ and $\hat{J}_n$ is

$$\text{var}\hat{J}_n = \text{var}J_n + \text{var} \left( W_n^T \sum_{l \neq n} G_l \{(J_l - E\{J_l\})(J_l - E\{J_l\})^T\} G_l^T \right) W_n \right) . \tag{16}$$

Therefore, finding the weight matrix $W_n$ which minimizes (16) corresponds to finding the source $\hat{J}_n$ with the strength closest to the strength of $J_n$ at the point $r_n$ in the brain.

Some comments:

- Using Lagrange multipliers, it is easy to show that (Van Veen et al., 1997)

  $$W_n = (BB^T)^{-1} G_n (G_n^T (BB^T)^{-1} G_n)^{-1} . \tag{17}$$

- 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, once computed through (17), each matrix $W_n$ is used only after a projection onto an appropriate subspace. This is generated by the set of eigenvectors associated with the eigenvalues of $BB^T$ significantly bigger than the noise and therefore can be considered as the subspace where the noise-free measurements live. 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 $BB^T$ is regularized, i.e. $(BB^T)^{-1}$ is replaced by $(BB^T + \lambda I)^{-1}$ where the regularization parameter $\lambda$ is chosen on the basis of the noise level.

- Condition (14) 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 (Dalal et al., 2006). Our beamforming algorithm implements this approach to the method.

**Multiple signal classification**

MUSIC was first developed in the array signal processing community (Schmidt, 1986). Its application to the MEG inverse problem (Mosher et al., 1992) is concerned with the non-linear framework of equation (9). In its most general formulation (Mosher & Leahy, 1999) MUSIC explicitly account for possible time correlated sources by means of a formal approach based on linear algebra results. In fact, time correlation reflects in decreasing the rank of the matrix $S$, $\text{rank}(S)$, from its maximum value $P$ to $R \leq P$: $R$ represents the number of the active sources, where, in this setting, each cluster of correlated sources is counted only once.

Then two diagonalizations can be computed. In the first one,

$$\text{ASS}^T A^T = \Phi \Lambda_M \Phi^T , \tag{18}$$
\( \Lambda_M \) is the \( M \times M \) diagonal matrix of the eigenvalues of \( \mathbf{A S S}^\top \mathbf{A}^\top \) (\( R \) non-zero and \( M - R \) zero eigenvalues); \( \Phi = (\Phi_S, \Phi_N) \) can be decomposed into the \( M \times R \) matrix \( \Phi_S \) whose columns are the eigenvectors associated with the non-zero eigenvalues and the \( M \times (M - R) \) matrix \( \Phi_N \), whose columns are the eigenvectors associated with the zero eigenvalues. The column vectors of \( \Phi_S \) generate the signal subspace, while the column vectors of \( \Phi_N \) generate the noise subspace. In the second diagonalization,

\[
\mathbf{S S}^\top = \mathbf{U} \Lambda_R \mathbf{U}^\top, \tag{19}
\]

\( \Lambda_R \) is the \( R \times R \) diagonal matrix of the \( R \) non-zero eigenvalues of \( \mathbf{S S}^\top \) and \( \mathbf{U} \) is the \( P \times R \) matrix whose columns are the eigenvectors associated with the non-zero eigenvalues. The algebraic result at the basis of MUSIC for MEG is that

\[
\mathcal{R}(\mathbf{A U}) = \mathcal{R}(\Phi_S), \tag{20}
\]

where \( \mathcal{R}(\cdot) \) denotes the range of a matrix. In order to derive a reconstruction algorithm from (20), it is crucial to observe that

\[
\mathbf{B B}^\top = \mathbf{A S S}^\top \mathbf{A}^\top + \sigma^2 \mathbf{I}, \tag{21}
\]

Therefore: (i) the number of eigenvalues of \( \mathbf{B B}^\top \) bigger than \( \sigma^2 \) is an estimate of \( R \); (ii) the subspace generated by the eigenvectors associated with these eigenvalues is an estimate of the signal subspace. Then the MUSIC algorithm is realized in three steps: first, \( \mathbf{B B}^\top \) is diagonalized and the eigenvalues bigger than \( \sigma^2 \) are selected together with the corresponding eigenvectors; second, all the points in the brain and, for each point, all the orientations, are spanned to find the ones for which (20) is satisfied; third, given \( \mathbf{A} \) as determined in the previous two steps, the linear system (9) is solved with some least-squares method to determine \( \mathbf{S} \).

Some comments:

- In principle, the number of eigenvalues bigger than \( \sigma^2 \) can be determined automatically. However in real applications this may be a difficult task, since the singular spectrum often does not present any clearly visible plateau. This also implies that, in order not to loose any possible active source, it is often preferrable to overestimate the number of eigenvalues bigger than \( \sigma^2 \).

- The role of \( \mathbf{U} \) is to keep track of the clusters of correlated dipoles. Operatively, this is realized by looking for solutions of

\[
\mathcal{R}(\mathbf{A}) = \mathcal{R}(\Phi_S), \tag{22}
\]

in vector spaces which are cartesian products of spaces \( \mathbb{R}^5 \) (this is the space where we describe position and orientation of the source): if \( \mathbb{R}^5 \) describes all the parameters of one source, \( \mathbb{R}^5 \times \mathbb{R}^5 \) will describe all the parameters of two correlated sources and so on.
• In our implementation of MUSIC, which is called Recursively Applied Projection (RAP)-MUSIC (Mosher & Leahy, 1999), the set of points and orientations satisfying (20) is determined by using the concept of correlation between subspaces described in (Golub & Loan, 1984) and computed by using a technique based on singular value decomposition. In actual applications this technique solves equation (20) only approximately, i.e. the subspace correlation is determined within a given threshold.

Particle Filters

Particle filters (Doucet, Godsill, & Andrieu, 2000; Arulampalam et al., 2002) are a class of recently developed algorithms for the numerical implementaton of Bayesian filtering (Somersalo & Kaipio, 2004): 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 (Somersalo et al., 2003; Sorrentino et al., 2007) in a multi-dipole setting characterized by a minimal set of assumptions. In the following, \( \pi(x) \) is the probability density function of the random vector \( X \), \( \pi(x|y) \) is the probability density function of \( x \) conditioned on the realization \( y \) of the random vector \( Y \), and \( \{B_t\}_{t=1,...,K} \) and \( \{J_t\}_{t=1,...,K} \) are the measurements and primary current (current dipoles) processes, respectively. The main assumptions for applying the particle filter algorithm are concerned with the markovian nature of the two processes and can be synthesized in the following equations:

\[
\pi(j_{t+1}|j_t, j_{t-1}, \ldots, j_1) = \pi(j_{t+1}|j_t) \quad (23)
\]

\[
\pi(b_t|j_t, j_{t-1}, \ldots, j_1) = \pi(b_t|j_t) \quad (24)
\]

\[
\pi(j_{t+1}|j_t, b_t, b_{t-1}, \ldots, b_1) = \pi(j_{t+1}|j_t) \quad (25)
\]

where, for sake of simplicity, notations like \( j_t \) mean \( j(t) \). If (23)-(25) hold, one can assume knowledge of the two probability density functions \( f(b_t|j_t) := \pi(b_t|j_t) \), named ”likelihood function” and \( p(j_{t+1}|j_t) := \pi(j_{t+1}|j_t) \), named ”transition kernel”; the Bayesian filtering algorithm is the sequential application of:

\[
\pi_{\text{post}}(j_t|b_t) = \frac{f(b_t|j_t)\pi_{\text{pr}}(j_t)}{\pi(b_t)} \quad (26)
\]

\[
\pi_{\text{pr}}(j_{t+1}) = \int p(j_{t+1}|j_t)\pi_{\text{post}}(j_t|b_t)\,dj_t \quad (27)
\]

With Bayes theorem (26) the posterior pdf, i.e. the solution of the problem in the Bayesian framework, is obtained by combining the prior pdf, containing a priori information on the unknown primary current, with the likelihood function, which embodies the forward problem and assumptions on the noise statistics. The Chapman-Komogorov formula (27) provides the mean for computing the next prior pdf from
the actual posterior. Therefore, three objects are needed to solve the inverse problem in this framework: the prior pdf at the first sampled time point \( \pi_{pr}(j_1) \), the likelihood function \( f(b_t|j_t) \) and the transition kernel \( p(j_{t+1}|j_t) \).

Equations (26) and (27) form the optimal Bayesian solution, but the recursive propagation of the pdf is only a conceptual solution, because in many practical applications the pdfs cannot be computed analytically. Particle filters are a sequential Monte Carlo method where equation (26) is computed by means of an importance sampling strategy (Ripley, 1987) with the prior density \( \pi_{pr}(j_t|b_{t-1}) \) playing the role of proposal density: let \( \{\tilde{j}_i^t\}_{i=1,..,N} \) be a sample distributed according to \( \pi_{post}(j_t|b_t) \); each sample point is usually named particle; then, equation (27) guarantees that one can obtain a sample \( \{j_{i+1}^t\}_{i=1,..,N} \) which is distributed according to \( \pi_{pr}(j_{t+1}|b_t) \) simply by making each particle evolve randomly, according to the transition kernel; then the importance weights \( w_{t+1}^i \) can be computed through the Bayes theorem (26), \( w_{t+1}^i \propto f(b_{t+1}|\tilde{j}_{t+1}^i) \) and the weighted sample \( \{\tilde{j}_{i+1}^t, w_{t+1}^i\}_{i=1,..,N} \) can be used to compute estimates of the current dipoles parameters from the posterior density; finally, a resampling step is applied in order to obtain a new sample set \( \{\tilde{j}_{i+1}^t\}_{i=1,..,N} \) distributed according to the posterior density \( \pi_{post}(j_{t+1}|b_{t+1}) \), and so on.

In our implementation the first prior density \( \pi_{pr}(j_1) \) is uniform in the brain volume and Gaussian for the dipole moments; the transition kernel \( p(j_{t+1}|j_t) \) is a zero-mean Gaussian function with diagonal covariance matrix, whose entries reflect the expected evolution of the current dipoles parameters in a time step; the likelihood function is a zero-mean Gaussian density, the covariance matrix being that of noise, which can be variously estimated from the data. Finally, we use the mean values of the posterior densities as estimates of the current dipoles parameters.

Some comments:

- 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 is computed only once, in correspondence with the points of this grid.

- At least in principle, an increase of the statistical effectiveness can be obtained by applying Rao-Blackwellization (Casella & Robert, 1996) to the filtering. The idea starts from the observation that in (9) the dependence of the data on the unknown is non-linear with respect to the dipole position but linear with respect to the dipole amplitude. Therefore a hybrid filter can be designed, where a particle filter tracks the positions and a Kalman filter reconstructs the amplitudes. We found that this Rao-Blackwellization procedure provides the same reconstruction accuracy with a smaller number of particles. 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.

- Since the formulation of the method employs only the past observations to
obtain the solution at time $k$, particle filters are the only method which may allow, in principle, an on-line real-time reconstruction of the neural sources.

**The comparison**

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

**Experimental conditions**

**Quasi-correlated sources (Figure 1).** From the neurophysiological viewpoint, the possibility of correctly recovering time-correlated sources is crucial for two reasons: first, correlated sources show up in many, even simple experimental conditions (e.g. simple auditory stimuli); furthermore, time-correlatedness is among the key-questions when different brain regions interact. 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 10 samples before the right one, i.e., the two sources are not perfectly time-correlated. Noise is then added to the synthetic measurements, up to a final SNR of 16 dB at the peak.

RAP-MUSIC and the particle filters are well suited for recovering time-correlated sources. Beamformers require as a technical assumption that the sources are not time correlated. However, if the presence of correlated sources is known, a correction to standard beamforming is possible, which may partly restore such correlation (our beamformer does implement this correction).

**Short-lasting sources (Figure 2).** The duration of a stimulus-induced activation usually increases with its latency: during word reading the Wernicke area - where the word meaning is processed - usually activates about 300 milliseconds after the word has appeared and lasts even more than 100 milliseconds (Sorrentino et al., 2006); on the other hand, for somatosensory stimuli, activity in the primary somatosensory cortex usually happens about 25 milliseconds after the stimulus and usually lasts for a couple of milliseconds (Iramina, Kamei, Yumoto, & Ueno, 2001). 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 right half of the helmet, its strength being non-zero for only 9 samples. Gaussian noise is added to obtain a final SNR of 12 dB at the peak. Both RAP-MUSIC and beamformers need an estimate of the signal subspace for their implementation and realize this by diagonalizing $BB^\top$. 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 (Van Veen et al., 1997) 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.

**Orthogonal dipoles (Figure 3).** When a current dipole is used to model the activity of a small brain region, the direction of the current dipole 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 14 dB at the peak. Both RAP-MUSIC and beamformers rely on the estimate of the signal subspace from $BB^\top$: in particular, the estimate of the non-linear parameters in RAP-MUSIC directly exploits the estimated signal subspace; therefore, we expect that both beamformers and even more RAP-MUSIC encounter problems in recovering orthogonal sources. Particle filters are not expected to have any problem in this condition.

**Low signal-to-noise ratio (Figure 4).** MEG measurements are characterized by an extremely low SNR; 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 a strong Gaussian noise was added to obtain a final SNR of 0.86 dB at the peak. The estimate of the signal subspace through $BB^\top$ should deteriorate in the case of low SNR, which in turn should lead to a worsening in the estimate of the source parameters from RAP-MUSIC. For beamformers, which recover 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.
Moving dipole (Figure 5). 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 (auditory stimuli). 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 ($b > 3 \text{ pT/cm}$) MEG signal excursions were rejected, and about 100 artifact-free trials for each stimulus category were collected and averaged on-line in windows $[-100, 500]$ 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 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.

Numerical results

The source reconstructions provided by the three methods are compared by plotting the amplitude waveforms as well as superimposing the position and orientation
of the reconstructed and original dipoles. For all three methods we have represented
the position and orientation in correspondence with the peak of the amplitude signal.
For each experimental situation, we describe the reconstructions provided by each
algorithm.

**Quasi-correlated sources (Figure 1).** Although the two source waveforms are
not identical, the beamformer output computed by means of (17) is almost null
due to the cancelation effects caused by the use of the constrained minimum
problem (10). To produce this figure (and throughout all the other experiments,
too) we used the correction to the method described in (Dalal et al., 2006); the
order of magnitude of the original source amplitudes is correctly recovered, as
shown in Figure 1(d), but, on the other hand, the delay between the sources is
not recovered. As for RAP-MUSIC, the eigenvalue plot of $\mathbf{B}\mathbf{B}^\top$ shows a single
dominating eigenvalue, and therefore even RAP-MUSIC reconstructs the two
sources as being perfectly correlated in time (Figure 1(g) upper panel). The
time delay is correctly recovered by RAP-MUSIC only if the rank of the signal
subspace is overestimated (Figure 1(g) lower panel). Particle filters correctly
recover the activation hierarchy of the original sources, although the amplitude
waveforms are not as smooth as the original ones. All three methods recover
orientations and positions rather well.

**Short-lasting sources (Figure 2).** The use of a small number of samples for esti-
mating the signal subspace mainly affects the reconstruction of the beamform-
ers: the source strength is underestimated, as shown in Figure 2(d), although
orientation and position are well reconstructed. The reconstruction provided
by RAP-MUSIC is good: despite the small number of samples, the non-linear
parameters are correctly recovered and consequently the source strength is also
good. Particle filters give good results as well. 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 MUSIC and particle filters.

**Orthogonal dipoles (Figure 3).** RAP-MUSIC fails in estimating the source ori-
entations. The reason is that the two orthogonal sources produce opposite
correlation between the same pairs of sensor; therefore in $\mathbf{B}\mathbf{B}^\top$ some contri-
butions are canceled and the estimate of the signal subspace is wrong. 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 3(i) and (l)). Both beamformers and particle filter
correctly recover the parameters of the two sources.

**Low signal-to-noise ratio (Figure 4).** For all algorithms, the source strength
is highly corrupted by noise. The beamformer also presents an overall un-
derestimation of the peak while the overall shape of the signal reconstruction
provided by the particle filter, although less unstable, also presents some discontinuities. Position and orientation are reconstructed fairly well by the three methods, although the particle filtering reconstruction shows a certain spread of the estimated dipoles around the true position.

**Moving dipole (Figure 5).** 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 6).** The two highly correlated sources activated by the auditory stimuli are correctly recovered by particle filters 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 particle filters are allowed to move during the measurements sequence, which explains 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 the second source is notably underestimated probably due to the high correlation level between the two sources.

*Figure 1.* Quasi-correlated sources: amplitudes (a) and projections on the axial plane (b) and on the sagittal plane (c) of the original dipoles; reconstructed amplitudes (d) and projections (e),(f) obtained with beamformers, with RAP-MUSIC ((g),(h) and (i)) and with particle filters ((j),(k) and (l))

**Discussion**

We have applied three different reconstruction algorithms for the analysis of both synthetic and real MEG time series. Beamforming, multiple signal classification and particle filtering provide surprisingly accurate reconstructions of the dipoles’ orientation and position in almost all conditions, while the amplitude time courses
Figure 2. Short-lasting source: amplitude (a) and projections on the axial plane (b) and on the sagittal plane (c) of the original dipole; reconstructed amplitude (d) and projections (e),(f) obtained with beamformers, with RAP-MUSIC (shown in (g),(h) and (i)) and with particle filters (shown in (j),(k) and (l))

Figure 3. Orthogonal dipoles: amplitudes (a),(b) and projections on the axial plane (c) and on the sagittal plane (d) of the original dipoles; reconstructed amplitudes (e),(f) and projections (g),(h) obtained with beamformers, with RAP-MUSIC (shown in (i),(j),(k) and (l)) and with particle filters (shown in (m),(n),(o) and (p))

are restored in a less reliable way and often present numerical instabilities and artifacts. From a computational viewpoint, the automation degree of the three methods is comparable, since each algorithm requires an optimal choice of parameters like the noise variance (for the particle filter), the signal subspace dimension (for MUSIC and beamformer) and the regularization parameter (for beamformer). On the other hand, the particle filter is perhaps the most general of the three approaches, since it can be applied with adequate results under most general neurophysiological conditions, presenting, for example, a high degree of temporal correlatedness or measurements very deteriorated by noise. Finally, beamforming is clearly the fastest one of the three methods, while the particle filter we implemented is presently too slow for performing on-line analysis of the MEG time series and can be employed only for post-processing.

All three methods can be improved from the viewpoint of both the formulation and the implementation. In particular, the beamformer is still not highly 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.

Acknowledgements. Lauri Parkkonen of the Brain Research Unit, Low Temper-
Figure 5. Moving dipole: amplitudes (a) and projections on the axial plane (b) and on the sagittal plane (c) of the original dipoles; reconstructed amplitudes (d) and projections (e),(f) obtained with beamformers, with RAP-MUSIC (shown in (g),(h) and (i)) and with particle filters (shown in (j),(k),(l))

Figure 6. 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 beamformers. (e),(f) and (h),(i) show the same objects for, respectively, RAP-MUSIC and particle filters

ature Laboratory, Helsinki University of Technology, is kindly acknowledged. This work is partly supported by a PRIN grant of the italian Ministero dell’Università e della Ricerca.

References

Arulampalam, M., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Tr. Sig. Proc.*, 50(2), 174-188.

Brookner, E. (1985). Phased-array radar. *Scientific American*, 252, 94-102.

Cantarella, J., Turck, D. D., & Gluck, H. (2001). The biot-savart operator for application to knot theory, fluid dynamics, and plasma physics. *Journal of Mathematical Physics*, 42, 876-905.

Casella, G., & Robert, C. (1996). Rao-blackwellisation of sampling schemes. *Biometrika*, 83, 81-94.

Dalal, S. S., Sekihara, K., & Nagarajan, S. S. (2006). Modified beamformers for coherent source region suppression. *IEEE Transactions on Biomedical Engineering*, 53, 1357-1363.

Doucet, A. (1998). On sequential simulation-based methods for bayesian filtering. *Technical Report, Univ. of Cambridge*.

Doucet, A., Godsill, S., & Andrieu, C. (2000). On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, 10, 197-208.

Gee, W., Lee, S., Bong, N. K., Cain, C. A., Mittra, R., & Magin, R. L. (1984). Focused array hyperthermia applicator: theory and experiment. *IEEE Transactions on Biomedical Engineering*, 31, 38-45.

Golub, G., & Loan, C. V. (1984). *Matrix computation*. The Johns Hopkins University Press.

Hamalainen, M., Hari, R., Knuttila, J., & Loukasmaa, O. (1993). Magnetoencephalogra-
phy: theory, instrumentation and applications to non-invasive studies of the working human brain. *Reviews of Modern Physics*, 65, 413-498.

Hamalainen, M., & Ilmoniemi, R. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Med. & Biol. Eng. & Comput.*, 32, 35-42.

Hamalainen, M., & Sarvas, J. (1989). Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Transactions on Biomedical Engineering*, 36, 165-171.

Haykin, S., Justice, J. H., Owsley, N. L., Yen, J. L., & Kak, A. C. (1985). *Array signal processing*. Prentice Hall.

Iramina, K., Kamei, H., Yumoto, M., & Ueno, S. (2001). Effects of repetition rate of electric stimulation on meg and fmri signals. *IEEE Transactions of Magnetics*, 37, 2918-2920.

Kress, R., Kuhn, L., & Potthast, R. (2002). Reconstruction of a current distribution from its magnetic field. *Inverse Problems*, 18, 1127-1146.

Mosher, J., & Leahy, R. (1998). Recursive music: A framework for eeg and meg source localization. *IEEE Transactions on Biomedical Engineering*, 45.

Mosher, J., & Leahy, R. (1999). Source localization using recursively applied and projected (rap) music. *IEEE Transactions on Signal Processing*, 47, 332-340.

Mosher, J., Lewis, P., & Leahy, R. (1992). Multiple dipole modeling and localization from spatio-temporal meg data. *IEEE Transactions on Biomedical Engineering*, 39, 541-557.

Pascarella, A., Sorrentino, A., Piana, M., & Parkkonen, L. (2007). Particle filters and rap-music in meg source modelling: a comparison. In *Proceedings of the international conference on biomagnetism (biomag 2006)* (Vol. 1300, p. 161-164).

Ripley, B. D. (1987). *Stochastic simulation*. Wiley and Sons.

Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, 32, 11-22.

Scherg, M. (1990). Fundamentals of dipole source potential analysis. In *Auditory evoked magnetic fields and electric potential. advances in audiology* (Vol. 6, p. 40-69).

Schmidt, R. O. (1986). Multiple emitter location and signal parameter estimation. *IEEE Trans. Antenn. Propogat.*, AP-34, 276-280.

Sekihara, K., Nagarajan, S., Poeppel, D., Marantz, A., & Miyashita, Y. (2002). Application of an meg eigenspace beamformer to reconstructing spatio-temporal activities of neural sources. *Human Brain Mapping*, 15, 199-215.

Sekihara, K., Poeppel, D., Marantz, A., Koizumi, H., & Miyashita, Y. (1997). Noise covariance incorporated meg-music algorithm: a method for multiple dipole estimation tolerant of the influence of background brain activity. *IEEE Transactions on Biomedical Engineering*, 44, 839-847.

Somersalo, E., & Kaipio, J. (2004). *Statistical and computational inverse problems*. Springer Verlag.

Somersalo, E., Voutilainen, A., & Kaipio, J. (2003). Non-stationary magnetoencephalography by bayesian filtering of dipole models. *Inverse Problems*, 19, 1047-1063.

Sorrentino, A., Parkkonen, L., & Piana, M. (2007). Particle filters: a new method for reconstructing multiple current dipoles from meg data. In *Proceedings of the international conference on biomagnetism (biomag 2006)* (Vol. 1300, p. 173-176).

Sorrentino, A., Parkkonen, L., Piana, M., Massone, A. M., Narici, L., Carozzo, S., et al.
(2006). Modulation of brain and behavioural responses to cognitive visual stimuli with varying signal-to-noise ratio. *Clinical Neurophysiology, 117*, 1098-1105.

Uutela, K., Hamalainen, M., & Somersalo, E. (1999). Visualization of magnetoencephalographic data using minimum current estimates. *NeuroImage, 10*, 173-180.

Van Veen, B., & Buckley, K. (1988). Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine, 5*, 4-24.

Van Veen, B., Drongelen, W. van, Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering, 44*, 867-880.

Van Veen, B., Joseph, J., & Hecox, K. (1992). Localization of intra-cerebral sources via linearly constrained minimum variance spatial filtering. In *Proceedings of the ieee workshop on statistical signal and array processing* (p. 526-529).
Figure 1. Quasi-correlated sources: amplitudes (a) and projections on the axial plane (b) and on the sagittal plane (c) of the original dipoles; reconstructed amplitudes (d) and projections (e),(f) obtained with beamformers, with RAP-MUSIC ((g),(h) and (i)) and with particle filters ((j),(k) and (l))
Figure 2. Short-lasting source: amplitude (a) and projections on the axial plane (b) and on the sagittal plane (c) of the original dipole; reconstructed amplitude (d) and projections (e),(f) obtained with beamformers, with RAP-MUSIC (shown in(g),(h) and (i)) and with particle filters (shown in (j),(k) and (l)).
Figure 3. Orthogonal dipoles: amplitudes (a), (b) and projections on the axial plane (c) and on the sagittal plane (d) of the original dipoles; reconstructed amplitudes (e), (f) and projections (g), (h) obtained with beamformers, with RAP-MUSIC (shown in (i), (j), (k) and (l)) and with particle filters (shown in (m), (n), (o) and (p))
Figure 4. Low signal-to-noise ratio: amplitude (a) and projections on the axial plane (b) and on the sagittal plane (c) of the original dipole; reconstructed amplitude (d) and projections (e),(f) obtained with beamformers, with RAP-MUSIC (shown in (g),(h) and(i)) and with particle filters (shown in (j),(k) and (l))
Figure 5. Moving dipole: amplitudes (a) and projections on the axial plane (b) and on the sagittal plane (c) of the original dipoles; reconstructed amplitudes (d) and projections (e),(f) obtained with beamformers, with RAP-MUSIC (shown in (g),(h) and (i)) and with particle filters (shown in (j),(k),(l))
Figure 6. Real data: reconstructed amplitudes with beamformers (a), RAP-MUSIC (d) and particle filters (g); coregistration on a Magnetic Resonance high resolution axial view (b) and coronal view (c) of the sources obtained with beamformers. (e),(f) and (h),(i) show the same objects for, respectively, RAP-MUSIC and particle filters.