Dynamical MEG Source Modeling with Multi-Target Bayesian Filtering

Alberto Sorrentino,[1]* Lauri Parkkonen,[2,3] Annalisa Pascarella,[1,4] Cristina Campi,[1,5] and Michele Piana[1,4]

[1]*CNR–INFM LAMIA, Genova, Italy*
[2]*Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology, Finland*
[3]*Advanced Magnetic Imaging Center, Helsinki University of Technology, Finland*
[4]*Dipartimento di Informatica, Università di Verona, Verona, Italy*
[5]*Dipartimento di Matematica, Università di Genova, Genova, Italy*

Abstract: We present a Bayesian filtering approach for automatic estimation of dynamical source models from magnetoencephalographic data. We apply multi-target Bayesian filtering and the theory of Random Finite Sets in an algorithm that recovers the life times, locations and strengths of a set of dipolar sources. The reconstructed dipoles are clustered in time and space to associate them with sources. We applied this new method to synthetic data sets and show here that it is able to automatically estimate the source structure in most cases more accurately than either traditional multi-dipole modeling or minimum current estimation performed by uninformed human operators. We also show that from real somatosensory evoked fields the method reconstructs a source constellation comparable to that obtained by multi-dipole modeling. *Hum Brain Mapp 00:000–000, 2009*.

Key words: magnetoencephalography (MEG); random finite sets; particle filter; Bayesian filtering; source localization; inverse problem

INTRODUCTION

Magnetoencephalography (MEG) non-invasively measures, with excellent time resolution, the weak magnetic fields produced by the currents flowing in active neurons. Reconstructing the underlying neural currents from MEG measurements allows localizing the active brain regions with reasonable spatial accuracy and enables a variety of applications, both in clinical and basic research.

MEG source reconstruction is a difficult task and several methods have been applied to this ill-posed inverse problem. The existing methods can be divided into two classes based on the mathematical model employed to describe the neural sources: "imaging" methods assume a continuous current density, discretized with a dense set of current dipoles, and often result in regularization algorithms such as Minimum Norm Estimation [Hämäläinen and Ilmoniemi, 1994], Minimum Current Estimate (MCE) [Uutela et al., 1999], and various types of beamformers [Sekihara et al., 2002; Van Veen et al., 1997]; "parametric" methods employ a small set of current dipoles, and estimate their

Contract grant sponsors: Academy of Finland (National Centers of Excellence Programme 2006-2011) and Fondazione CaRiVerona.

*Correspondence to: Alberto Sorrentino, Via Dodecaneso 35, 16146 Genova (GE), Italy. E-mail: sorrentino@fisica.unige.it

Received for publication 31 October 2008; Revised 11 February 2009; Accepted 25 February 2009

DOI: 10.1002/hbm.20786

Published online in Wiley InterScience (www.interscience.wiley.com).

positions and magnitudes by non-linear optimization techniques [Aine et al., 2000; Hämäläinen et al., 1993; Mosher and Leahy, 1999; Uutela et al., 1998].

Both classes suffer from well-known shortcomings; among the "imaging" methods, the L2-norm estimates tend to be too wide-spread, and beamformers suppress temporally correlated sources; more recently developed algorithms, however, seem to be more effective, exploiting e.g., the temporal continuity of the solution [Ou et al., 2009]. Although several automatic algorithms have appeared for the parametric approach, manually-assisted dipole fitting is still the most widely used method due to its simplicity.

Recently, owing to the increase in the available computational power, Bayesian methods have become feasible. They cast the problem in a more general setting and estimate the whole posterior probability density function (pdf) instead of searching for a single optimal solution. Jun et al. [2005] consider the dipole parameter estimation as a Bayesian inference problem and apply Markov Chain Monte Carlo methods for sampling the posterior density. Galka et al. [2004] and Long et al. [2006], for EEG and MEG, respectively, consider the inverse problem as a dynamical one and apply Kalman filtering to a linear distributed source model.

In this article, we consider MEG source estimation as a dynamical Bayesian inverse problem, and use a multi-dipole model for the source distribution. Contrary to most multi-dipole methods, we allow the number of sources, as well as their positions and orientations, to vary over time. Due to the non-linearity of the model, a particle filter [Arulampalam et al., 2002; Doucet et al., 2000] is employed to explore the posterior densities.

Random Finite Sets (RFSs) [Matheron, 1975; Molchanov, 2005] provide a suitable mathematical framework for dealing with a time-varying number of sources in a Bayesian setting. RFSs are a generalization of Random Variables (RV), and they do not constrain the unknowns to a dimensionality known a priori. In the present study we exploit the theory of RFSs to construct a particle filter which solves the MEG inverse problem in a general manner, allowing a genuinely dynamical model comprising multiple sources. The mathematical details of the RFS theory will not be discussed in this article; see, e.g., Mahler [2003], Vihola [2004] and Vo et al. [2005] for the background.

We applied the multi-dipole particle filter to both simulated and real data. The first data sets were 1,000 Monte Carlo realizations of a three-source configuration, whereas the next two sets reflected complex activation sequences and they have been employed in a comparison of traditional dipole modeling and Minimum Current Estimate [Stenbacka et al., 2002]. Thus, we can directly compare the performance of the new algorithm with that of these two other methods. We evaluate the performance of the particle filter also by analyzing somatosensory evoked fields recorded in one subject.

METHODS

Bayesian Filtering for MEG Source Localization

In the Bayesian approach to inverse problems [Somersalo and Kaipio, 2004], the unknown and the measurements are modeled as Random Variables and the solution is the entire posterior probability density function of the unknown, obtained by the Bayes theorem. For dynamical inverse problems the sequential application of Bayes theorem, which requires a prior density at each time step, is mediated by the use of the Chapman-Kolmogorov equation, described below the result is a two-step algorithm known as Bayesian filtering.

In the MEG application, let j_t and b_t be the realizations of the Random Vectors $\mathbf{J}_t$ and $\mathbf{B}_t$, the primary current distribution and the magnetic field at time t, respectively. Assuming that the stochastic process $\mathbf{J}_t$ is a first-order Markov process, and that $\mathbf{B}_t$ is a first-order Markov process with respect to $\mathbf{J}_t$, the posterior density $\pi(j_t \mid b_{1:t})$ at time t, given the prior density $\pi(j_t \mid b_{1:t-1})$, is obtained by the Bayes theorem:

$$\pi(j_t|b_{1:t}) = \frac{\pi(b_t|j_t)\pi(j_t|b_{1:t-1})}{\pi(b_t|b_{1:t-1})} \tag{1}$$

where $\pi(b_t \mid j_t)$ is the likelihood function, determined by the forward model and the noise statistics; the denominator is the normalization constant $\pi(b_t|b_{1:t-1}) = \int \pi(b_t|j_t)\pi(j_t|b_{1:t-1})dj_t$. The prior density $\pi(j_{t+1} \mid b_{1:t})$ at time $t+1$, given the posterior density $\pi(j_t \mid b_{1:t})$ at time t, is estimated by the Chapman-Kolmogorov equation:

$$\pi(j_{t+1}|b_{1:t}) = \int \pi(j_{t+1}|j_t)\pi(j_t|b_{1:t})dj_t \tag{2}$$

where $\pi(j_{t+1} \mid j_t)$ is the transition kernel of the stochastic process underlying the data. Here, it reflects general assumptions, such as temporal continuity, on the dynamics of brain sources.

Equations (1) and (2) hold independently of the source and forward models; indeed, these equations have been applied to the MEG inverse problem in rather different frameworks: Long et al. [2006] utilized a Kalman filter for a distributed source model; Somersalo et al. [2003], Sorrentino et al. [2007] and Campi et al. [2008] employed particle filters to obtain the parameters of a set of current dipoles.

Bayesian Filtering of Random Finite Sets of Dipoles

We applied Bayesian filtering for solving the MEG inverse problem in a truly dynamical multi-dipole framework, where neural sources may appear, strengthen, move, weaken and disappear in the course of time. This approach is conceptually different from that of most other source modeling methods which merely represent the data as a collection of sources whose amplitudes vary over time. These methods cannot readily estimate, e.g., the number of sources at a particular time instant.

In general, the primary current at time t is a set of n_t dipoles

$$J_t = \{d_t^1, \ldots, d_t^{n_t}\} = (R_t, Q_t); \tag{3}$$

where $d_t^k = (r_t^k, q_t^k)$ is an abbreviated notation for a single dipole at location r_t^k and with dipole moment q_t^k. Similarly, collections of dipole locations and dipole moments are denoted respectively as

$$R_t = \{r_t^1, \ldots, r_t^{n_t}\} \quad \text{and} \quad Q_t = \{q_t^1, \ldots, q_t^{n_t}\}. \tag{4}$$

To apply Bayesian filtering, we model the primary current as a RFS of dipoles. Like RVs, RFSs have probability density functions. Furthermore, the integral of a set function is well-defined; Equations (1) and (2) of Bayesian filtering extend to RFSs, provided that the integrals involved are interpreted as set integrals [Vo et al., 2005].

We denote the RFS of dipoles at time t by J_t, and a realization of J_t by J_t. The RFS J_t contains a random number $\mathbf{N}_t$ of random dipoles: $\mathrm{J}_t = \{\mathbf{D}_t^1, \ldots, \mathbf{D}_t^{\mathbf{N}_t}\}$, and $J_t = \{d_t^1, \ldots, d_t^{n_t}\}$.

The equations of Bayesian filtering now read:

$$\pi(J_t|b_{1:t}) = \frac{\pi(b_t|J_t)\pi(J_t|b_{1:t-1})}{\pi(b_t|b_{1:t-1})} \tag{5}$$

and

$$\pi(J_{t+1}|b_{1:t}) = \int \pi(J_{t+1}|J_t)\pi(J_t|b_{1:t})\delta J_t. \tag{6}$$

Although these equations appear similar to Eqs. (1) and (2), there are differences: here J_t are realizations of an RFS, $\pi(J_t|\cdot)$ are probability densities of an RFS and the integral is a set integral.

One may question the rationale of using RFSs instead of RVs; for example, Somersalo et al. [2003], Sorrentino et al. [2007] and Campi et al. [2008] modeled the multi-dipole problem with RVs. However, the representation of physical source constellations with RVs is not unique: for example, the state formed by the two dipoles d^1 and d^2 has two representative vectors $(d^1, d^2) \in \mathsf{D} \times \mathsf{D}$ and $(d^2, d^1) \in \mathsf{D} \times \mathsf{D}$, where $\mathbf{D}$ is the single-dipole space. As a consequence, the posterior density is permutation-invariant and (artificially) multi-modal: it will peak at (d^1, d^2) and at (d^2, d^1). This bi-modality would prevent us from using standard estimators such as the conditional mean (which would give erroneous estimates) or the Maximum A Posteriori (which would not be unique). This permutation symmetry disappears with RFSs, where each constellation has a unique representative set $\{d^1, d^2\} \equiv \{d^2, d^1\}$, since order of the elements is not relevant in a set. Furthermore, as shown later, a suitable estimator is available for RFSs.

Belief Measures and Models

The actual application of Bayesian filtering requires knowledge of three probability density functions: the very first prior density to initialize the algorithm, the likelihood function, and the transition kernel of Eqs. (5) and (6). In the case of RFSs, the explicit form of the multi-target probability density function is often complicated; instead, it is preferable to use belief measures [Molchanov, 2005]. The belief measure of a RFS is often defined as the RFS-analogue of the cumulative distribution of RV [Vihola, 2004, p. 42]; it is fully determined by a finite set of probability densities, one defined on the single-dipole space $\mathbf{D}$, another on the double-dipole space $\mathbf{D} \times \mathbf{D}$, and so on. Since we assume that no more than $n_{\max}$ dipoles can be active simultaneously, $n_{\max}$ probability measures are enough to define the belief measure. Therefore, in our approach, the belief measures for the prior density, the likelihood function, and the transition kernel are computed starting from the more familiar densities over vector spaces.

Now we introduce the three pdfs used in the algorithm.

Initial prior density

Although the previous density at $t > 1$ is automatically computed by the algorithm, the first prior density needs to be defined. Since we have no a priori knowledge of the number of sources, we give a uniform distribution to the marginal probability $P(|J_1| = k) = 1/(n_{\max} + 1)$, $k = 0, 1, \ldots, n_{\max}$, where $|J_1|$ is the number of dipoles in the set J_1 and $n_{\max}$ is the maximum allowed number of simultaneous sources. Since we assume that simultaneous dipoles are independent, it is sufficient to define the prior density for a single dipole. For the locations r_1, we use a uniform distribution in the brain volume, and for the dipole moments q_1, we use a zero-mean Gaussian distribution $N(0, \gamma_q)$, where the standard deviation γ_q is of the order of magnitude of the expected sources.

Likelihood function

Since the dimension of the measurement vector is fixed, the data sequence can be modeled with Random Vectors $\mathbf{B}_t$. The forward operator, denoted as $F(\cdot)$, depends on the properties of the volume conductor. Additive noise $\mathbf{M}_t$ is included in the model:

$$\mathbf{B}_t = F(\mathrm{J}_t) + \mathbf{M}_t. \tag{7}$$

Assuming a zero-mean Gaussian noise, the likelihood function is $\pi(b_t|J_t) = N(b_t - F(J_t), \sigma_{\text{noise}})$ where σ_{noise} is the standard deviation of the noise and is determined from the pre-stimulus period. Such a noise model is an approximation; the statistical distribution of MEG noise is generally not known.

Transition kernel

The transition kernel in Eq. (6) is also a priori information inserted in the model: using a non-specific model can be interpreted as inserting less-informative priors. We

adopt a simple model for the dynamics of the sources; at each time point (i) a new dipolar source may appear with probability p_{new} and density π_{new}, (ii) an existing dipole may disappear with probability p_{dis}, or (iii) the existing dipole may survive, with probability $1 - p_{\text{dis}}$, and evolve according to a single-dipole evolution model $\pi(d_{t+1} \mid d_t)$. All these events are assumed to be independent. In terms of RFSs, the model equation is:

$$\mathsf{J}_{t+1} = \mathsf{S}(\mathsf{J}_t) \cup \mathsf{J}_{t+1}^{\text{new}} \tag{8}$$

where $\mathsf{J}_{t+1}^{\text{new}}$ is the RFS of the new sources, containing at most one new dipole, and $\mathsf{S}(\mathsf{J}_t)$ is the RFS of the survived and evolved dipoles.

Assuming that dipoles evolve independently, it can be shown that the probabilities p_{new} and p_{dis} together with the single-dipole evolution model $\pi(d_{t+1} \mid d_t)$ and the density $\pi_{\text{new}}(d_{t+1})$ fully determine the belief measure and therefore the transition kernel in Eq. (6). The densities we use in the applications are:

- the density $\pi_{\text{new}}(d_{t+1})$ of the new dipoles $d_{t+1} = (r_{t+1}, q_{t+1})$ which is uniform for the location and Gaussian for the dipole moment, i.e., the same as the initial prior density; the value of p_{new} is set to 0.5;
- the probability for a dipole to disappear at time $t + 1$ which is $p_{\text{dis}} = 1/(2n_t)$, where $n_t = |J_t|$. This implies that each particle loses, on average, half a dipole at each time point, which counterbalances the average number of newborn dipoles in each particle.
- the single-dipole evolution model which is a random walk constrained within the brain volume; the transition kernel $\pi(d_{t+1} \mid d_t)$ for a single dipole is characterized by Gaussian densities $\pi(r_{t+1} \mid r_t) = N(r_{t+1} - r_t, \gamma_r)$ and $\pi(q_{t+1} \mid q_t) = N(q_{t+1} - q_t, \gamma_q)$ where the standard deviations $\gamma_r = 1$ cm and $\gamma_q = 2$ nAm.

Multi-Dipole Particle Filter

Since the MEG forward problem is non-linear with respect to dipole position, Eqs. (5) and (6) can be solved only numerically. Here we applied a sequential Monte Carlo technique known as particle filtering [Arulampalam et al., 2002; Doucet et al., 2000], where a sample set ("particles") distributed according to the posterior density is obtained at each time step by discretizing Eqs. (5) and (6), and introducing a resampling step which reduces the number of unlikely sample points. The general scheme of the particle filter we apply is as follows:

- Initialization: Draw a sample $\{J_1^i\}_{i=1,\ldots,p}$ distributed according to the density $\pi(J_1)$. Each sample point J_1^i, or "particle", is a set of dipoles. Particles are sampled from a RFS distribution, and therefore different particles may contain a different number of dipoles.
- Observation: Apply the Bayes theorem, i.e., compute the forward solution for each particle J_1^i and its normalized weight (likelihood) $w_t^i = (1/k)\pi(b_t \mid J_t^i)$ with $k = \sum_{i=1}^p w_t^i$. The set of weighted particles $\{J_t^i, w_t^i\}$ is an approximation of the posterior density at time t;
- Resampling: Randomly select p particles from the set $\{J_t^i\}_{i=1,\ldots,p}$ in such a way that the probability of extracting J_t^i is equal to its likelihood w_t^i. The set of uniformly weighted particles $\{\tilde{J}_t^i\}$ is a new approximation of the posterior density at time t and in this set the particles which had larger weights appear more frequently.
- Evolution: Let each particle $\{\tilde{J}_t^i\}$ evolve according to the transition kernel by drawing a new particle J_{t+1}^i from each $\pi(J_{t+1} \mid \tilde{J}_t^i)$. Each surviving dipole of each particle evolves according to the single-dipole transition kernel described earlier. New dipoles may appear. The set of uniformly weighted particles $\{J_{t+1}^i\}$ is an approximation of the prior density at time $t + 1$.

Source Estimates

The posterior density of the current dipole set contains all the available information on the source constellation, but it is difficult to visualize as a whole; instead, different estimates can be computed in order to extract the relevant information.

The posterior density carries information on the number of sources: the marginal distribution

$$P(|J_t| = k) = \int_{\mathrm{D}(k)} \pi(J_t \mid b_{1:t}) \delta J_t, \tag{9}$$

where $\mathrm{D}(k)$ is the set of finite subsets with k dipoles, provides a time-varying estimate of the number of active sources

$$\hat{n}_t = \arg\max P(|J_t| = k). \tag{10}$$

For estimating the source parameters we used the RFS-analogue of the first moment of a RV: the Probability Hypothesis Density (PHD). Here, PHD is a function in the single-dipole space D, its integral over region R provides an estimate of the number of dipoles in R, and the peaks of PHD can be used as estimates of the active dipoles. In a particle filter

$$\mathrm{PHD}(d_t) = \sum_{i=1}^{p} w_t^i \left(\sum_{d \in J_t^i} \delta(d - d_t) \right). \tag{11}$$

The problem of estimating the source parameters is thus transformed to the problem of finding the local maxima of the PHD, which, in our case, is relatively easy since the number of peaks to be found is already given by the model selection [Eq. (10)].

To further reduce the computational complexity, we compute $\mathrm{PHD}(r_t)$ instead of $\mathrm{PHD}(d_t)$. Therefore, the estimate of the dipole positions is

$$\hat{R}_t = \{\hat{r}_t^1, \ldots, \hat{r}_t^{\hat{n}_t}\} = \arg\,\text{peaks}(\text{PHD}(r_t)) \qquad (12)$$

and the corresponding dipole moments are computed by the standard least-squares technique.

Implementation

The number of sources and particles

The number of sources which can be simultaneously recovered at any given time point is limited here to $n_{\max}$ for computational reasons; however, the global number of sources in an MEG data set is practically unlimited as the $n_{\max}$ sources can, in principle, be different for every time point. We observed that using a bigger $n_{\max}$ also required a larger number p of particles to properly approximate the posterior density in spaces with many dipoles; the convergence to the true posterior density is guaranteed only for p tending to infinity. Overestimating $n_{\max}$ does not change the reconstruction accuracy; on the other hand, if $n_{\max}$ is fixed smaller than the actual number of active sources, the algorithm would most likely recover $n_{\max}$ sources, missing (most likely) the source(s) producing the weaker SNR. To check that the particle filter had reached a reliable approximation of the true posterior density, we always ran the algorithm 10 times and verified the low variance of the reconstructions. The parameters $n_{\max} = 5$ and $P = 100{,}000$ provided stable results in all the cases considered in this study.

Source-point grid

The computational cost of the algorithm is largely determined by the observation step which involves computing the MEG forward solution for each dipole of each particle at each time step. The large number of particles may render the algorithm impractical even when resorting to the spherical conductor and its analytic forward solution. In order to reduce the computational cost, we constrained particles to a grid enclosing the brain volume, and precomputed the forward solution for each grid point. The grid is a 20-cm cube with 125,000 points in a 4-mm uniform lattice, requiring 375,000 forward solutions (the three orthogonal directions for each point). Employing 100,000 particles without the grid, assuming two simultaneous dipoles on average, and a typical 500-sample analysis window results in 100,000,000 forward solutions to compute; the cost thus reduces by a factor of 300, which makes the algorithm usable. On a standard PC (2 GHz CPU, 2 GB RAM), the time for analyzing a single time point with 100,000 constrained particles is about 1 s.

Such a reduction in computational cost could not be obtained by Rao-Blackwellization, which exploits the linear dependence of the measurements and the dipole moment [Campi et al., 2008; Casella and Robert, 1996]. In Rao-Blackwellization, the optimal solution is computed for the linear dipole moment parameters through a set of Kalman filters, while the non-linear source locations are sampled with a particle filter; this reduces the number of particles needed to obtain a given accuracy, however, a covariance matrix has to be computed by inversion at each time step for each particle. Although statistical efficiency increases, the cost of the matrix inversions is still higher than that of the particle filter we use in this study.

Clustering

The multi-dipole particle filter provides dynamical estimates of the active dipoles. The locations of the estimated sources at time t, $\{\hat{d}_t^1, \ldots, \hat{d}_t^{\hat{n}_t}\}$ may be slightly different from those at time $t-1$, $\{\hat{d}_{t-1}^1, \ldots, \hat{d}_{t-1}^{\hat{n}_{t-1}}\}$, and even the number of sources $\hat{n}_t$ may change. Thus, no particular dipole is continuously bound to a given neural source and it is not possible to readily provide the amplitude waveform for each source.

To overcome this problem, we search for clusters within the set of all estimated dipoles $\{\hat{d}_1^1, \ldots, \hat{d}_1^{\hat{n}_1}, \ldots, \hat{d}_T^1, \ldots, \hat{d}_T^{\hat{n}_T}\}$, each cluster representing a neural source. The clustering is performed in a 6-dimensional space: three dimensions for the location, two for the orientation, and one for time t when the source dipole was present. This temporal parameter ensures continuity of the source waveforms. We applied k-means [Spath, 1980] clustering algorithm, however, it requires the number of clusters specified a priori. Since the total number of neural sources active during the analysis epoch is unknown, we applied the following iterative procedure:

1. overestimate the number of possible sources s;
2. cluster the dipoles in s clusters with k-means;
3. test each pair of clusters by the Wilcoxon test [Weerahandi, 1995] for statistical difference; if two or more clusters are not significantly different, decrement s and go to step 2; otherwise, stop.

RESULTS

We applied the multi-dipole particle filter to synthetic and real MEG recordings.

In Simulation 1, we tested the particle filter with 1,000 different data sets; each case comprises three sources at random locations, two of them being perfectly correlated in time.

Simulations 2 and 3 present more challenging situations with multiple sources mimicking the activations evoked by a complex visual stimulus. First we show the estimates obtained by a single run of the particle filter. We also used these two data sets to investigate the robustness and statistical reliability of the algorithm by performing multiple runs.

We also applied the particle filter to somatosensory evoked fields recorded in one subject.

In all tests, $n_{\max}$ was set to 5, and the number of particles p to 100,000. According to our tests, these numbers guarantee a good compromise between stability and computational cost.

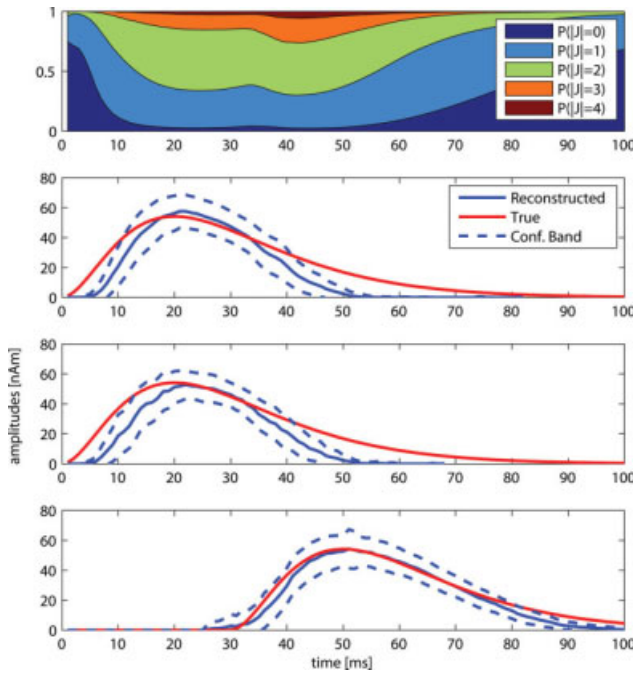

Figure 1.
Simulation 1. The marginal probability of dipole models with different number of dipoles, averaged across 1,000 runs, plotted as a function of time (top); each color represents a specific model with a different number of sources. The original (red solid) and average reconstructed (blue dashed) source waveforms (other panels), plus/minus the standard deviation (dotted).

Simulation 1

One thousand data sets were produced according to the following criteria. (1) Three dipolar sources were randomly located in the brain volume, with uniform distribution; source locations are not constrained to the grid points; the distance between any two sources could not be smaller than 2 cm. (2) Source orientations were randomly distributed on the plane orthogonal to the radius of the spherical head model. (3) The temporal waveforms of the three sources (see Fig. 1) are the same in all the 1,000 simulations: two sources (Sources 1 and 2) are totally correlated and have the peak at the 20th time point; the third source, Source 3, has the peak at the 50th time point. A spherical conductor model was used for the forward calculations, and Gaussian noise was added; the Signal-toNoise Ratio (SNR) is computed as $10\log_{10}$ ($|B|/|N|$), where $||$ is the Frobenius norm, B is the measurements matrix and N is the noise matrix; the SNR of the entire sequence ranged from 0.5 to 9.3 dB, the average being 3.8 dB; the SNR at the peak ranged between 1.1 and 23.9 dB, the average being 9.6 dB.

A source was considered to be correctly reconstructed if (i) the average localization error was less than 2 cm and (ii) the peak amplitude was within 10 ms from the true peak. The particle filter recovered 2,244 sources from the total of 3,000 (74.8%) sources; from 354 simulations (35.4%), three sources were reconstructed, from 536 simulations (53.6%) two sources were reconstructed and from 110 simulations (11.0%) only one source was recovered. The average localization error across the 1,000 simulations and 2,244 sources was 6.0 $\pm$ 2 mm; this localization error seemed not to depend on the SNR of the single source: sources producing an SNR between 0 and 2 were localized with an average error of 6.2 mm, for SNR between 2 and 4 the average error is 5.9 mm, for SNR higher than 4 the average error was still 5.9 mm. Source 1 was missed in 256 simulations, Source 2 in 251 simulations and Source 3 in 249 simulations; considering only the subset of simulations where 2 sources were recovered, the pair Source1-Source2 was found in 30% of cases, the pair Source1-Source3 in 34% and the pair Source2-Source3 in 36%; this suggests that the temporal correlation affects the detectability of sources only very slightly.

In Figure 1, we plot the average (across 1,000 simulations) model selection function and the average source waveforms, with confidence bands given by the standard deviation. On average, the algorithm was able to reconstruct the three source waveforms except for the very low-SNR tails, where the signal power was lower than the noise level.

Simulations 2 and 3

We utilized synthetic data from a previous study [Stenbacka et al., 2002] designed to evaluate and compare traditional multi-dipole modeling to MCE when performed by human operators unaware of the source structure of the data. Re-using the data enabled a direct comparison of the particle filter to these other methods.

Stenbacka et al. employed four simulations of increasing complexity. Here we present the results from the two most complex data sets (Simulation 3 and Simulation 4 in Stenbacka et al. [2002], hereafter referred to as Simulation 2 and 3, respectively) with a variable number of temporally overlapping sources within a volume comparable to a lobe of the brain. Both data sets are crafted to approximate hypothetical neural responses to a complex visual stimulus. Table I summarizes the locations, orientations and peak latencies of the 10 sources; six of them appeared in both Simulation 2 and 3, and the remaining four only in Simulation 3. The temporal waveforms of the sources are shown in Figures 2 and 3.

Realistically-shaped boundary element model of the brain was applied in the MEG forward calculation and signals corresponding to those from Vectorview^{TM} neuromagnetometer (Elekta Neuromag Oy, Helsinki, Finland), comprising 102 magnetometers and 204 planar gradiometers, were computed. Brain noise from a MEG experiment where the subject was silently resting was added to the simulated responses for a final SNR, computed as previously explained, of about 4 dB. For more details, see Stenbacka et al. [2002].

TABLE 1. Sources in simulations 2 and 3

Area	Location (mm) (x, y, z)	Orientation (x, y, z)	Sim 2 t_{peak} (ms)	Sim 3 t_{peak} (ms)
V1	(11.1,−53.4, 49.8)	$(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$	70	60
V2	(13.6,−60.2, 55.9)	(1, 0, 0)	90	80
V3	(17.3,−59.4, 59.8)	(0, 0, 1)	110	100
$V3_a$	(22.3,−54.8, 64.6)	$(-\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$	130	120
V4	(23.1,−47.3, 35.8)	$(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$	150	220
$V5_R$	(43.6,−36.8, 44.4)	$(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)$	170	160
$V5_L$	(−33.7, −48.5, 48.1)	$(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)$	–	190
POS	(3.0,−40.0, 83.0)	$(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$	–	200
STS_L	(52.0, 0, 48.0)	(0, 0, −1)	–	280
STS_R	(−52.0, −4.0, 48.0)	(0, 0, −1)	–	220

V1–V5 refer to the visual cortices, POS to parieto-occipital sulcus, and STS to superior temporal sulcus; Locations and orientations in "head coordinates" (x from left to right preauricular point; y towards nasion and perpendiculat to x; z upwards and normal to the xy-plane).

We evaluated the reconstructed sources using the same criteria as Stenbacka et al: for Simulation 2, a source was considered as correctly estimated if its location was within 2 cm from the true source and its peak latency within half the duration of the true source; for Simulation 3, a source was considered as correctly estimated if its location was again within 2 cm from the true source and the peak latency was either within 10 ms from the peak of the true source or the peak latency was within half of the time interval determined by the half maxima of the true source.

The particle filter was able to reconstruct all of the six sources in Simulation 2. The average localization error was 9.1 mm. The estimated time courses (see Fig. 2) never overlapped, i.e., at each time point the algorithm recovered just a single dipole since the true overlapping sources were within few millimeters from each other and thus a single dipole explained the measured field sufficiently well. The differing orientations of the sources enabled the algorithm to cluster them in six groups.

From Simulation 3, the particle filter was able to reconstruct six of the 10 sources (see Fig. 3) with average localization error of 7.6 mm. The estimated source classified as V2 accounted for the activity of V1 and V3, whereas sources V4, and $V5_L$ were missed. The other five sources were correctly recovered. The apparently worse performance of the algorithm with respect to Simulation 2 was due to two different reasons. First, the SNR produced by V3 in this case was weaker than in Simulation 2, as the source strength was lower. Second, the duration of V1 in Simulation 3 was shorter than in Simulation 2 and therefore there

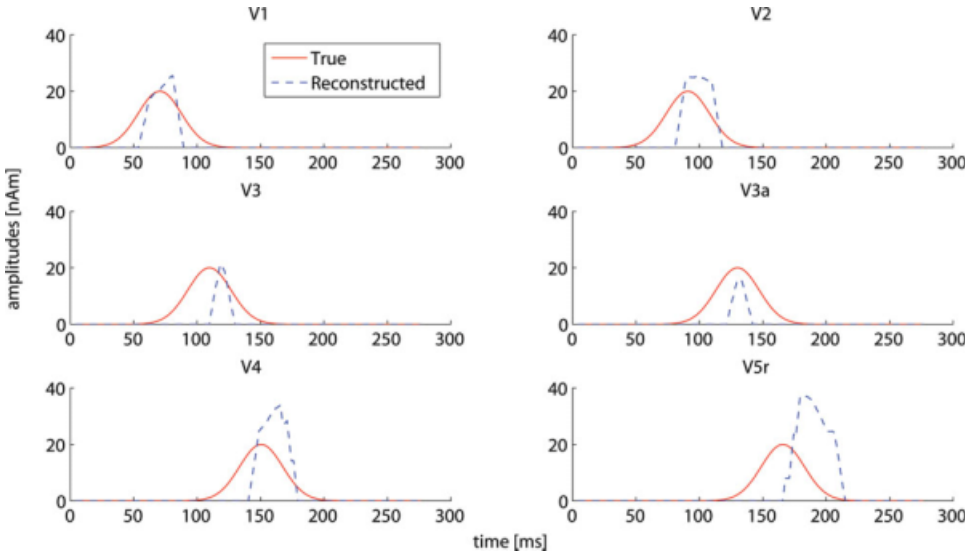

Figure 2.
Simulation 2. Time courses of the true (red solid) and estimated (blue dashed) sources.

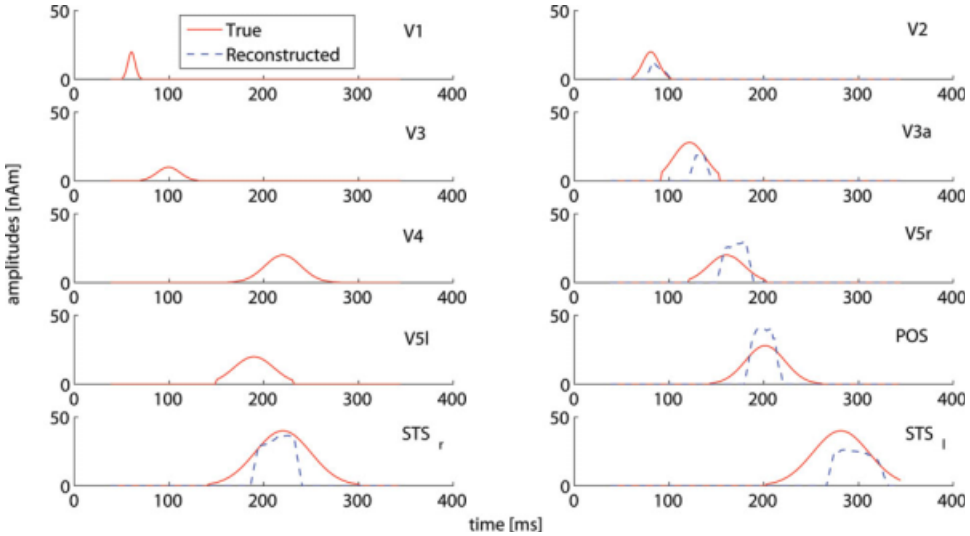

Figure 3.
Simulation 3. Time courses of the original (red solid) and estimated (blue dashed) sources.

were not enough sources to allow the clustering procedure to create two different groups.

Robustness and Reliability

The particle filter requires a priori information: the initial prior density, the transition kernel and the likelihood function. However, the initial prior density is usually uninformed, coding the lack of information on the initial configuration of the sources, and it does not affect the results substantially. We investigated the sensitivity of the reconstruction to the transition kernel parameters γ_r and γ_q in the ranges 1–2 cm and 1–5 nAm, respectively, and it proved to be negligible. The most relevant parameter is σ_{noise} in the likelihood function; it tunes the sensitivity of the algorithm. To investigate the statistical reliability of a single run and the robustness of the algorithm with respect to the parameter σ_{noise}, we performed 10 runs for each of five different values of σ_{noise} using the data of Simulations 2 and 3. Figure 4 shows the results; with both simulations the average number of recovered sources had a peak at $\sigma_{noise} = 4$; with higher values of σ_{noise}, the algorithm considered only stronger sources; at lower σ_{noise}, also weak sources were recovered but the algorithm was less stable and the localization error increased. Tuning σ_{noise} impacted also the variability across runs, which reached a minimum around the same value $\sigma_{noise} = 4$ and remained small above this value: the standard deviation of the source locations across 10 runs for $\sigma_{noise} \geq 4$ was 1.8 mm, the standard deviation of the peak amplitude was about 5%. Finally, the number of false positives decreased quickly for increasing values of σ_{noise}: for too low values the algorithm tries to model also noise.

Since σ_{noise} should reflect the noise level, it can be estimated from, e.g., the pre-stimulus baselines. Interestingly, the standard deviation of the baseline period was

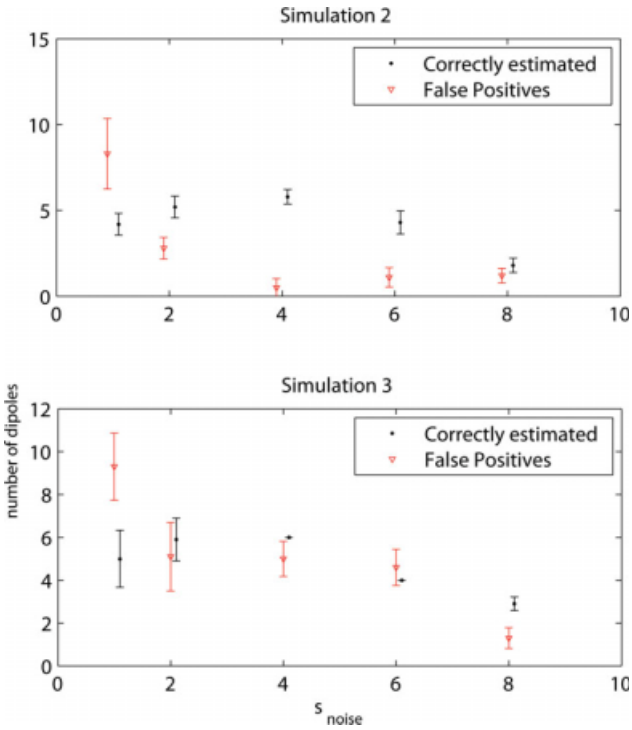

Figure 4.
Simulations 2 and 3. The average number of correctly estimated sources and of false positives. The pre-stimulus baseline variance was 4.3.

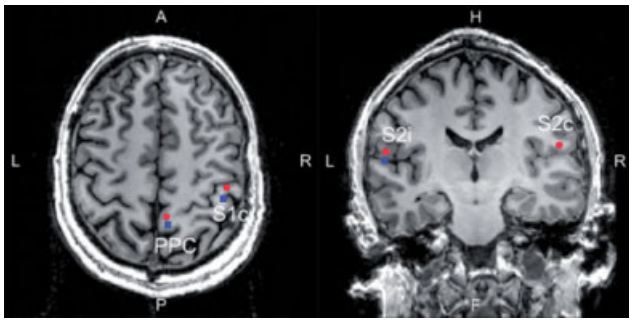

Figure 5.
Source reconstruction of somatosensory evoked fields. Axial (left) and coronal (right) views of the source models obtained by the particle filter (blue squares) and dipole modeling (red circles).

4.3 in Simulations 2 and 3, i.e., very close to the optimal $\sigma_{noise} = 4$.

Somatosensory Responses

We applied the particle filter to somatosensory evoked fields (SEF) measured in one healthy human. The recordings were performed after informed consent and had a prior approval by the local ethics committee.

The SEFs were acquired with a 306-channel MEG device (Elekta Neuromag Oy, Helsinki, Finland) comprising 204 planar gradiometers and 102 magnetometers in a helmet-shaped array. The left median nerve at wrist was electrically stimulated at the motor threshold with an interstimulus interval randomly varying between 7.0 and 9.0 s. The MEG signals were filtered to 0.1–200 Hz and sampled at 600 Hz. Trials with EOG or MEG exceeding 150 μV or 3 pT/cm, respectively, were excluded and 84 clean trials were averaged. To reduce external interference, signal-space separation method [Taulu et al., 2004] was applied to the average.

A 3D digitizer and four head position indicator coils were employed to determine the position of the subject's head within the MEG helmet with respect to anatomical MRIs obtained with a 3-Tesla MRI device (General Electric, Milwaukee, USA).

The SEFs (see e.g., Hari and Forss, 1999) were modeled with multiple dipoles whose anatomical locations were verified to be plausible: the N20m and P35m responses at 21 and 38 ms, respectively, localized in the assumed hand area of the S1 cortex contralateral to the stimulation, bilateral responses peaking around 90 ms in the S2 cortices, and a response at around 125 ms in the contralateral posterior parietal cortex (PPC). The N20m and P35m sources were only about 6 mm apart and had antiparallel orientations. To avoid spurious interaction, the N20m dipole was omitted from the multi-dipole model. The goodness-of-fit of the four-dipole model with respect to the data from all 306 channels ranged from 77 to 97% at the response peaks. This model served as the reference for evaluating the models obtained by the particle filter.

The particle filter was able to recover three of the four sources found by dipole modeling (see Fig. 5) and missed the source in the contralateral S2 cortex (S2c) probably due to its relatively weak field pattern compared with the other sources. The three reconstructed sources satisfied the spatial criterion of Simulation 3 (the average location was within 2 cm from the reference location). Figure 6 presents the source amplitude waveforms estimated by both methods. The particle filter clearly separates the activations of the different sources whereas the time courses given by the multi-dipole model suffer from leakage of residual activity and noise thus showing a non-zero dipole moment throughout the analysis period. Interestingly, the particle filter reconstructs two sources at 1.5 cm distance, separated in time and by orientation, for the ipsilateral S2 cortex. The two peaks of opposite polarity in the dipole waveform (S2i in Fig. 6) likely reflect these two sources at 1.5 cm distance.

DISCUSSION

The particle filter described in this study represents a notable improvement with respect to the previous implementations [Somersalo et al., 2003; Sorrentino et al., 2007].

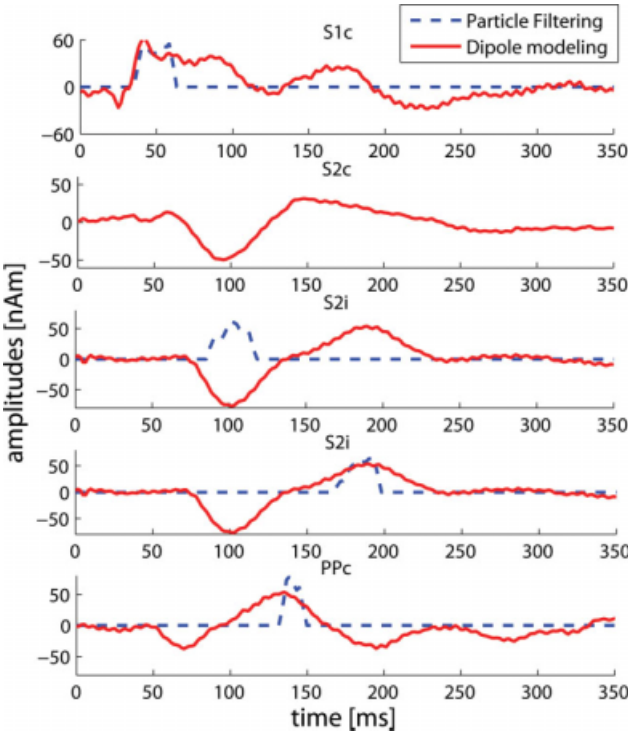

Figure 6.
Somatosensory evoked fields: Time courses estimated by the particle filter (blue dashed) and multi-dipole model (red solid).

The more sophisticated mathematical framework allowed a coherent description of the problem in terms of time-varying sets of dipoles, and provided a suitable estimator for multi-dipole states. The algorithm was able to localize a time-varying number of dipoles with no prior knowledge on their number. The use of a source-point grid made the algorithm fast enough to be practically usable.

The localization accuracy in time and space was investigated by testing the particle filter first against 1,000 Monte Carlo realizations of a 3-source configuration in a wide SNR range, and then against two challenging data sets which were previously analyzed by Stenbacka et al. [2002] with dipole modeling and MCE. In the 1,000 Monte Carlo simulations, the particle filter localized about 75% of the sources, with average localization error of 6.0 mm; the SNR does not seem to affect the localization error, rather, the SNR variation seems to produce an on/off effect on the detectability of dipolar sources; temporal correlation seems to affect the detectability of the sources only marginally. From the second data set (Simulation 2 in this study) the particle filter correctly localized 6 of 6 sources, with an average localization error of 9.1 mm, thus outperforming the other reconstruction methods applied by uninformed users, who recovered on average 2.4 sources with dipole modeling and 3.2 with MCE. In the third data set (Simulation 3), the particle filter correctly localized six of 10 sources, with an average localization error of 7.6 mm. Uninformed users obtained similar results, recovering on average 5.4 sources with dipole modeling and 6.0 sources with MCE, with an average localization error of 7.2 mm in both cases. Since the particle filter allows the sources to move and turn during the analysis epoch, the clustering step, which exploits both position and orientation, is able to distinguish sources that dipole modeling unavoidably lumps together. This difference likely explains why the particle filter outperformed the other methods in Simulation 2. In Simulation 3, the results by the particle filter are comparable to those obtained by uninformed users with either dipole modeling or MCE; however, the particle filter algorithm operated automatically while the other two methods required subjective decisions on what to consider a true source. From somatosensory evoked fields, previously analyzed by an informed human using multi-dipole modeling, the particle filter was able to reconstruct three of the four sources found by the human operator, only missing the weak contralateral S2 response. The particle filter properly localized the sources also in time, avoiding cross-talk between the dipoles.

In this study, we tested the particle filter under very general conditions, with as little a priori information as possible. However, such information is available and can be readily exploited. For example, the source space could be constrained to the cortical surface as the bulk of MEG responses are assumed to originate in the cortex. In addition, the noise covariance matrix could be estimated from the pre-stimulus intervals to further reduce the effect of non-white noise. The evolution model could also be replaced with a more realistic one. Optimal usage of the available prior information should be addressed in future work.

In conclusion, the algorithm presented here is a step towards an automatic MEG source modeling method which not only estimates source current distributions but provides a discrete set of significantly active sources for each time instant. Traditional multi-dipole modeling and minimum norm approaches require substantial post-processing to reach qualitatively comparable results. The multi-dipole particle filter we described directly provides instantaneous, time-varying estimates of the number of sources and of the dipole parameters.

ACKNOWLEDGMENT

The authors thank Linda Stenbacka for providing the simulated data used in her study and Riitta Hari for comments on an earlier version of the manuscript.

REFERENCES

Aine C, Huang M, Stephen J, Christner R (2000): Multistart algorithms for MEG empirical data analysis reliably characterize locations and time courses of multiple sources. NeuroImage 12:159–172.

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 50:174–188.

Campi C, Pascarella A, Sorrentino A, Piana M (2008): A Rao-Blackwellized particle filter for magnetoencephalography. Inverse Problems 24:025023.

Casella G, Robert CP (1996): Rao-Blackwellisation of sampling schemes. Biometrika 83:81–94.

Doucet A, Godsill S, Andrieu C (2000): On sequential Monte Carlo sampling methods for Bayesian filtering. Stat Comput 10:197–208.

Galka A, Yamashita O, Ozaki T, Biscay R, Valdes-Sosa P (2004): A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering. NeuroImage 23: 435–453.

Hari R, Forss N (1999): Magnetoencephalography in the study of human somatosensory cortical processing. Philos Trans R Soc Lond B Biol Sci 354:1145–1154.

Hämäläinen M, Hari R, Knuutila J, Lounasmaa O (1993): Magnetoencephalography: Theory, instrumentation and applications to non-invasive studies of the working human brain. Rev Mod Phys 65:413–498.

Hämäläinen M, Ilmoniemi R (1994): Interpreting magnetic fields of the brain: Minimum norm estimates. Med Biol Eng Comput 32:35–42.

Jun S, George J, Paré-Blagoev J, Plis S, Ranken D, Schmidt D, Wood C (2005): Spatiotemporal Bayesian inference dipole analysis for MEG neuroimaging data. NeuroImage 28:84–98.

Long C, Purdon P, Temeranca S, Desai N, Hämäläinen M, Brown E (2006): Large scale Kalman filtering solutions to the electrophysiological source localization problem—A MEG case study. In: Proceedings of the 28thIEEE EMBS Annual International Conference. Vol. 5. New York: The Printing House, Inc. pp. 4532–4535.

Mahler R (2003): Multitarget Bayes filtering via first-order multitarget moments. IEEE Trans Aerospace Electron Syst 39:1152–1178.

Matheron G (1975): Random Sets and Integral Geometry. New York: Wiley.

Molchanov I (2005): Theory of Random Sets. London: Springer-Verlag.

Mosher J, Leahy R (1999): Source localization using recursively applied and projected (RAP) MUSIC. IEEE Trans Signal Proc 47:332–340.

Ou W, Hämäläinen MS, Golland P (2009): A distributed spatio-temporal EEG/MEG inverse solver. NeuroImage 44:932–946.

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 Map 15:199–215.

Somersalo E, Kaipio J (2004): Statistical and Computational Inverse Problems. New York: 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 re constructing multiple current dipoles from MEG data. In: Cheyne D, Ross B, Stroink G, Weinberg H, editors. Proceedings of the International Conference on Biomagnetism (BIOMAG 2006), Elsevier. Vol. 1300. pp. 173–176.

Spath H (1980): Cluster Analysis Algorithms for Data Reduction and Classification of Objects. New York: Halsted Press.

Stenbacka L, Vanni S, Uutela K, Hari R (2002): Comparison of minimum current estimate and dipole modeling in the analysis of simulated activity in the human visual cortices. NeuroImage 16:936–943.

Taulu S, Kajola M, Simola J (2004): Suppression of interference and artifacts by the signal space separation method. Brain Topogr 4:269–275.

Uutela K, Hämäläinen M, Salmelin R (1998): Global optimization in the localization of neuromagnetic sources. IEEE Trans Biomed Eng 45:716–722.

Uutela K, Hämäläinen M, Somersalo E (1999): Visualization of magnetoencephalographic data using minimum current estimates. NeuroImage 10:173–180.

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

Vihola M (2004): Random sets for multitarget tracking and data fusion. Licentiate Thesis, Tampere University of Technology.

Vo B, Singh S, Doucet A (2005): Sequential Monte Carlo methods for multitarget filtering with random finite sets. IEEE Trans Aerospace Electron Syst 41:1224–1245.

Weerahandi S (1995): Exact Statistical Methods for Data Analysis. New York: Springer-Verlag.