NeuroImage 46 (2009) 168–176

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Technical Note

Selecting forward models for MEG source-reconstruction using model-evidence

R.N. Henson [a,*], J. Mattout [b], C. Phillips [c], K.J. Friston [d]

[a] MRC Cognition and Brain Sciences Unit, Cambridge, England, UK
[b] INSERM U821-Brain Dynamics and Cognition, Lyon, France
[c] Cyclotron Research Centre, University of Liège, Belgium
[d] Functional Imaging Laboratory, University College London, England, UK

ARTICLE INFO

Article history:
Received 6 November 2008
Revised 21 January 2009
Accepted 23 January 2009
Available online 11 February 2009

Keywords:
Magnetoencephalography
Inverse problem
Bayesian

ABSTRACT

We investigated four key aspects of forward models for distributed solutions to the MEG inverse problem: 1) the nature of the cortical mesh constraining sources (derived from an individual's MRI, or inverse-normalised from a template mesh); 2) the use of single-sphere, overlapping spheres, or Boundary Element Model (BEM) head-models; 3) the density of the cortical mesh (3000 vs. 7000 vertices); and 4) whether source orientations were constrained to be normal to that mesh. These were compared within the context of two types of spatial prior on the sources: a single prior corresponding to a standard L2-minimum-norm (MNM) inversion, or multiple sparse priors (MSP). The resulting generative models were compared using a free-energy approximation to the Bayesian model-evidence after fitting multiple epochs of responses to faces or scrambled faces. Statistical tests of the free-energy, across nine participants, showed clear superiority of MSP over MNM models; with the former reconstructing deeper sources. Furthermore, there was 1) no evidence that an individually-defined cortical mesh was superior to an inverse-normalised canonical mesh, but 2) clear evidence that a BEM was superior to spherical head-models, provided individually-defined inner skull and scalp meshes were used. Finally, for MSP models, there was evidence that the combination of 3) higher density cortical meshes and 4) dipoles constrained to be normal to the mesh was superior to lower-density or freely-oriented sources (in contrast to the MNM models, in which free-orientation was optimal). These results have practical implications for MEG source reconstruction, particularly in the context of group studies.

Introduction

There are various approaches to constructing a forward model that maps electrical activity at one or more sources within the brain to the electrical or magnetic field recorded by sensors outside the brain. Some models allow the sources to live anywhere within the three-dimensional brain volume, while others constrain the sources to a 2D manifold of the cortical surface, defined using MRI (Dale and Sereno, 1993). Another choice is whether the head-model, which uses Maxwell's equations to predict the electromagnetic field produced by the sources at a given sensor (the "leadfield"), is based on analytical solutions for spherical surfaces, or on numerical solutions for a Boundary Element Model (BEM) approximation to the head (Mosher et al., 1999). (Note that there are further numerical methods such as Finite Element and Finite Volume modelling, Pruist et al., 1993, but we do not consider these here.) Further choices concern the number of possible dipoles within the source space, and in the case of sources fixed on a 2D cortical surface, whether the orientation of those dipoles is free or constrained to be normal to the surface. The latter constraint reflects the assumption that MEG/EEG signals recorded over the scalp derive from synchronous activity in the pyramidal cells that are largely perpendicular to the cortical surface (Nunez, 1981).

We explored these choices, within the context of distributed norm inversions of different forward models, for MEG data recorded from nine participants. Within a Bayesian framework, the various choices for forward modelling constitute part of the generative model; therefore the Bayesian concept of "model-evidence" can be used to compare those choices (see Appendix). While there have been several previous formal comparisons of some of the models considered here, these have normally used simulated data (e.g. comparing the point-spread or crosstalk functions for various spherical vs. BEM head-models), for which at least one model is normally considered to be the truth (e.g., a BEM; Huang et al., 1999). Our empirical model comparisons provide an important complement to these simulations. Moreover, we are not aware of prior studies that have simultaneously explored the range of model attributes we now consider.

The choice of meshes

In the present work, we constrained the sources to lie within a tessellated mesh of the cortical mantle. Obtaining an accurate tessellation of the cortical surface via segmentation of an MRI is a

* Corresponding author. MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge, CB2 2EF, UK. Fax: +44 1223 359 062.
E-mail address: rik.henson@mrc-cbu.cam.ac.uk (R.N. Henson).

1053-8119/$ – see front matter.
doi:10.1016/j.neuroimage.2009.01.062

difficult problem, often requiring manual intervention (though see Fischl et al., 1999). One alternative that we proposed recently is to take a cortical mesh created carefully by hand from an MRI of a "template" brain, which has been transformed into a standard stereotactic space (Talairach and Tournoux, 1988). This template mesh can be warped to match an individual's MRI, using the inverse transformation of the spatial normalisation procedures that have been established in the MRI literature (Ashburner and Friston, 2005). When using simulated data, we previously found no evidence that this inverse-normalised, template mesh — which we called a "canonical mesh" - performed any worse than a mesh based on an individual's cortical surface; in terms of either the model-evidence or the localization error (Mattout et al., 2007). One key advantage of a canonical mesh is that it provides a one-to-one mapping between the individual's source-space and the template space, facilitating group analyses (Litvak and Friston, 2008) and the incorporation of spatial priors that live in the template space, such as group fMRI results (Flandin et al., 2009).

However, our previous results pertained only to the single individual, so it is unknown whether a canonical mesh would consistently be sufficient over a larger sample of individuals. Furthermore, our previous simulations used only a single-sphere head-model (aligned with the cortical mesh), whereas more complex head-models, such as BEMs, may be more sensitive to the choice of mesh (*viz* the use of canonical vs. individual inner skull or scalp meshes; see below).

Here, we used three meshes for each individual — one for the cortex, one for the inner skull and one for the outer scalp (see Fig. 1). Each mesh served a different function. The cortical mesh constrained the possible source locations (and their orientations in some models). The inner skull mesh was used to fit the single-shell head-model (i.e., a single sphere, overlapping spheres, or BEM; see below); the scalp mesh was used to coregister the MEG sensors with the meshes (that are defined in the individual's MRI space) via a set of digitized points on the scalp. We explored four combinations of meshes, depending on whether each corresponded to a template mesh, a canonical mesh, or was derived manually from an individual's MRI (see Fig. 1 and Results).

The choice of head-model

We considered three different head-models: a single-shell sphere (Sarvas, 1987), a sphere fitted separately to the local curvature below each sensor ("overlapping spheres", Huang et al., 1999), or a single-shell Boundary Element Model (BEM) (Mosher et al., 1999). All three were aligned to the same inner skull surface; since this tends to be the surface associated with the greatest change in conductivity. The single and overlapping sphere models can be solved analytically, using Sarvas's method (Sarvas, 1987), whereas the BEM requires numerical calculation for each face within the inner skull mesh. Note that these three head-models were considered for each of the four mesh-combinations above, since either the inner skull or cortical mesh differed within each set, creating a factorial model-space.

The choices of dipole density and orientation

We considered two cortical mesh densities: approximately 3000 or 7000 vertices. Both mesh-sizes were considered for fixed dipoles, with an orientation normal to the local curvature of the mesh, and free dipoles, where source magnitude was estimated for each of three orthogonal directions, effectively tripling the number of free

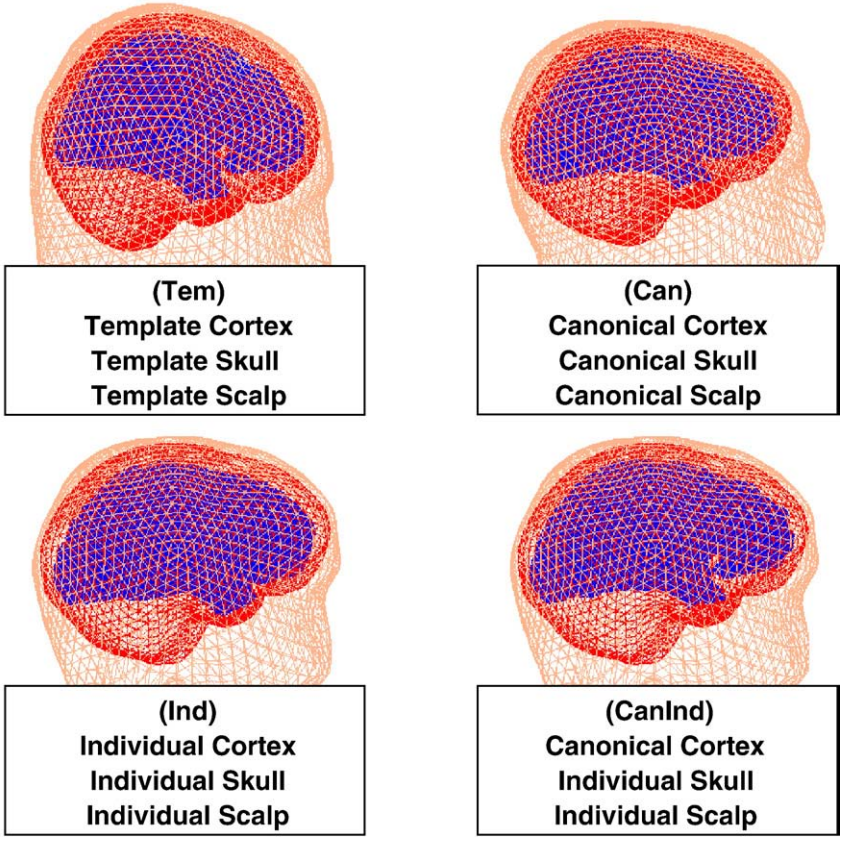

Fig. 1. The four combinations of three meshes considered for cortex, inner skull and scalp. **Tem** (Template) = created from a different (neurotypical) subject's MRI that had been warped to the Montreal Neurological Institute (MNI) template in Talairach space; **Can** (Canonical) = inverse-normalised Template, where normalisation parameters are determined by warping the individual's MRI to the template MRI; **Ind** (Individual) = created directly from an individual's MRI. **CanInd** = combination of canonical cortical mesh and individually-defined inner skull and scalp meshes.

parameters.[1] The precise orientation of dipolar sources often has a greater effect on leadfields than their precise localisation (e.g., Salayev et al., 2006). Given the convoluted nature of the cortical surface, and the ensuing errors in its segmentation and tessellation, one might expect better performance when the orientation is free to vary. This is particularly relevant when inverse-normalising a cortical mesh from a template brain, since there is no exact correspondence of sulci across brains. However, this must be considered in light of the massive under-determinacy of the inverse problem (i.e., estimating several thousand, or tens of thousands, parameters from only a few hundred, correlated sensor values). A more constrained source-space may actually produce more probable source estimates on average, even if it is less accurate. The best model is that which balances accuracy and complexity, as encapsulated in the "model evidence" (see Appendix).

The choice of source priors

The sources were estimated in two ways: either using Minimum (L2) Norm (MNM) or Multiple Sparse Priors (MSP). Whereas the above choices of mesh and head-model affect the form of the lead-field, the choice of source prior affects the prior covariance of the source parameters. These source priors also form part of the generative model within a Bayesian framework. The MNM inversion corresponds to a standard approach (Hauk, 2004) that can be expressed in terms of a single variance component. This spatial prior is an identity matrix over sources, reflecting the assumption that each source is independent and identically-distributed (effectively encouraging solutions with the minimal total energy). The hyper-parameter associated with this single source prior controls the relative weighting of the minimum-norm constraint relative to the fit to the data (the "regularisation"), and is here estimated by maximising the free-energy bound on the model log-evidence using an iterative Expectation-Maximisation (EM) algorithm. In brief, this entails optimizing the hyperparameters with respect to the free-energy, using conventional gradient ascent. By construction, the free-energy is always less than the log-evidence for a particular model (that is defined in terms of its covariance components). This means that when the free-energy is maximized, the hyperparameters are the most likely, given the data, and the free-energy becomes a bound approximation to the log-evidence that can then be used to compare models (see Friston et al., 2007, for full treatment).

The MSP source model is a more recent approach (Friston et al., 2008), in which the source-space is divided into a number of small patches (i.e., subsets of dipoles, weighted by their surface proximity to centre of each patch), typically resulting in several hundred spatial priors on the sources. This reflects the assumption that neural activity in the brain is sparse; i.e., typically occurs in a number of discrete regions (but presumably connected by long-range fibres). Here we used 768 patches: 256 for each hemisphere, and 256 bilateral (homologous) patches. The associated hyperparameters are estimated as above by optimising the free-energy. Simulations have shown that the MSP approach not only results in higher model-evidences than the MNM approach, but also produces more accurate localisations (Friston et al., 2008). It has also been shown to produce more plausible solutions for an EEG dataset, and circumvent the well-known bias of the MNM approach to produce widely-distributed, superficial solutions. However, MSP has not been compared to MNM on MEG data using a sample of individuals. We therefore thought it important to explore the effects of different MEG lead-fields on both a standard inversion prior (MNM) and a more recent approach (MSP).

The above four factors affecting the lead-field matrix (mesh-type, head-model, mesh-size and dipole-orientation), together with the fifth factor of source priors, define each model — resulting in a model-space of $4\times3\times2\times2\times2=96$ different models. To make exploration of this model-space more tractable we used a heuristic search by splitting the space into two, three-way factorial partitions: the first search considered the factors of mesh-type, head-model and source-priors (using the larger mesh of 7004, normally-oriented dipoles), whereas the second explored the factors of mesh-size, dipole-orientation and source-priors, using the best mesh-type and head-model from the first search (*viz* a canonical cortical mesh, individual skull and scalp meshes and a BEM head-model).

Test data

The above models were evaluated on MEG data recorded from 151 axial gradiometers from nine participants, while they made symmetry judgments on randomly intermixed trials of faces and scrambled faces. The 172 epochs in total (from −100 ms to +600 ms) were used to calculate the data covariance over sensors for each participant. These data were used to optimise the sensor and source covariance components required for model inversion. This produces both the free-energy approximation to the log-evidence and estimates of the source activity (see Appendix). We used the source estimates to illustrate the face validity of the models in terms of evoked responses. We focussed on the M170, a component around 150–200 ms post-stimulus that is greater for faces than non-face stimuli (such as scrambled faces), and for which there is good evidence from prior EEG and MEG experiments, in addition to fMRI and intracranial EEG, that it is generated by sources in mid-fusiform, lateral occipital and possibly lateral posterior temporal cortex (e.g., Allison et al., 1999; Henson et al., 2003; Watanabe et al., 2005). Thus, the reason for choosing the present dataset was not only that it has been used in the context of other methodological developments (Henson et al., 2007; Chen et al., 2009), but because the solution of each model could also be judged in terms of its plausibility.

Methods

The MEG data

The dataset is identical to that described in Henson et al. (2007). In brief, the data came from a single, eleven minute session in which participants saw 86 intact and 86 scrambled faces, subtending visual angles of approximately four degrees. Half of the faces were famous, and half were novel; the scrambled faces were phase-shuffled, Fourier-transformed versions of the faces. Participants made left–right symmetry judgments about each stimulus by one of two finger-presses (range of reaction times: 1031 ms–1798 ms). The MEG data were sampled at 625 Hz on a 151-channel axial gradiometer CTF Omega system at the Wellcome Trust Laboratory for MEG Studies, Aston University, England. Nine participants were tested, four female, ranging from young to middle-aged adults. Their involvement complied with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the standards established by a local review board.

MRI data, meshes and forward models

A T1-weighted MPRAGE-MRI scan was acquired for each participant with voxel-size of $1\times1\times1$ mm. These scans were segmented using SPM5 (http://www.fil.ion.ucl.ac.uk/spm), and the different partitions used to create meshes of 2002 vertices (4000 faces) for 1) the scalp and 2) the inner skull surface. These meshes were derived

[1] Note that for the single-sphere head-model, there was some redundancy among these parameters, because MEG cannot measure the purely radial component of the source orientation (given that radial sources produce no detectable magnetic field over the surface of a sphere). Note also that there are more sophisticated ways to accommodate orientation errors, such as scaling the tangential components with respect to the orthogonal one (Phillips et al, 2005), or using a "loose orientation constraint" (Lin et al, 2006).

from automated growing and eroding of binarized versions of the MRI (specifically, the sum of the gray, white and CSF partitions in the case of the inner skull), onto which a spherical mesh was projected and adjusted using an elastic correction. Each participant's scalp was also digitised using a Polhemus device, and the digitised head-points coregistered with the scalp mesh; so that the MEG sensor positions and orientations could be transformed into the MRI space.

BrainVISA/Anatomist (http://brainvisa.info) was used to create individual cortical meshes from each MRI of about 80,000 vertices, which were subsequently subsampled to between 7204 and 7211 vertices across participants. These meshes comprised a continuous triangular tessellation of the grey/white matter interface of the neocortex (excluding cerebellum). The mean inter-vertex spacing ranged from 4.3 mm to 5.3 mm across participants. The normal to the surface at each vertex was calculated from an estimate of the local curvature of the surrounding triangles (Dale and Sereno 1993). BrainVISA/Anatomist was also used previously to create the template cortical, inner skull and meshes (based on the MRI of a neurotypical male, normalised to a MNI template in Talairach space; Mattout et al., 2007). The template cortical meshes used here contained either 7204 vertices (14,400 faces) or 3004 vertices (6000 faces) (as available in the SPM5 software package).

Brainstorm (http://neuroimage.usc.edu/brainstorm) was used to fit a single-sphere, overlapping spheres or a BEM to the inner-skull mesh and to calculate lead-fields for sources normal to the cortical mesh or for three orthogonal directions. In the case of BEMs, a linear Galerkin method was used in which the inner skull mesh was reduced to 1000 vertices to reduce computational load.

Inversion

The MEG data were analysed using SPM5. The continuous data were epoched from −100 to +600 ms, and the data covariance calculated within a Hanning window across the epoch and a frequency-band of 1–44 Hz. The data were reduced to 6–8 temporal modes using singular-value decomposition (Friston et al., 2008), which typically captured over 93% of the data variance. This data covariance was then used to estimate the cortical sources using either Minimum Norm (MNM) or Multiple Sparse Priors (MSP), by maximising the free-energy approximation to the model-evidence (using greedy-search in the case of MSP). The remaining options were as default in SPM5, with the exception that no spatial dimension-reduction was performed, in order to compare forward models directly (see Appendix). For MSP, a fixed number of patches (256 per hemisphere) and smoothness (0.6) was used for all cortical meshes (note that higher density meshes entail smaller patches); for freely-oriented sources, each direction at each vertex had the same prior variance.

Reliable effects in the free-energy across participants were assessed using repeated-measures Analysis of Variance (ANOVA) with a Greenhouse–Geisser correction to the degrees of freedom. For subsequent evaluation of the source reconstructions, the difference in mean evoked energy across trials and participants, within a Gaussian window from 150 to 190 ms (Friston et al., 2006), was estimated for faces relative to scrambled faces.

Results

In what follows, we describe the results of our model-comparison and report the results of source reconstructions for the selected models identified by the heuristic search over model-space.

Analysis 1: effects of meshes, head-model and source-priors

In the first analysis, cortical meshes of approximately 7000 normally-oriented dipoles were used, and three factors were crossed: source-priors (2 levels), meshes (4 levels) and head-model (3 levels). The source-priors were the standard (L2) Minimum Norm (MNM) and Multiple Sparse Priors (MSP). The four meshes were 1: Template cortex, inner skull and scalp meshes (**Tem**), 2: Canonical (inverse-normalised Template) cortex, inner skull and scalp meshes (**Can**), 3: Canonical cortex mesh and individual skull and scalp meshes (**CanInd**), and 4: Individual cortex, skull and scalp meshes (**Ind**) (see Fig. 1). The three head-models (all fit to the inner-skull mesh) were 1: Single-sphere (**Sph**), 2: Overlapping-spheres (**OS**) and 3: Boundary Element Model (**BEM**).

The average free-energy across the nine participants for the resulting 24 models is shown in Figs. 2A and B. The largest effect size ($\eta^2 = 0.12$) in the $2\times4\times3$ ANOVA was the main effect of source-priors, $F(1,8) = 90.5$, $p<.001$, which reflected greater evidence for MSP relative to standard MNM (cf. panels A and B of Fig. 2). The next largest effect size ($\eta^2 = 0.09$) was the main effect of meshes, $F(1.03,8.26) = 12.8$, $p<.01$, which appeared to be driven by weaker evidence for Template meshes than the other three types of mesh (consistent with the absence of a reliable main effect of mesh when excluding Template meshes, $F(1.53,12.2)<1$).

The three-way interaction did not reach significance, $F(3.43,27.5) = 1.91$, $p = .14$, but there were reliable two-way interactions between

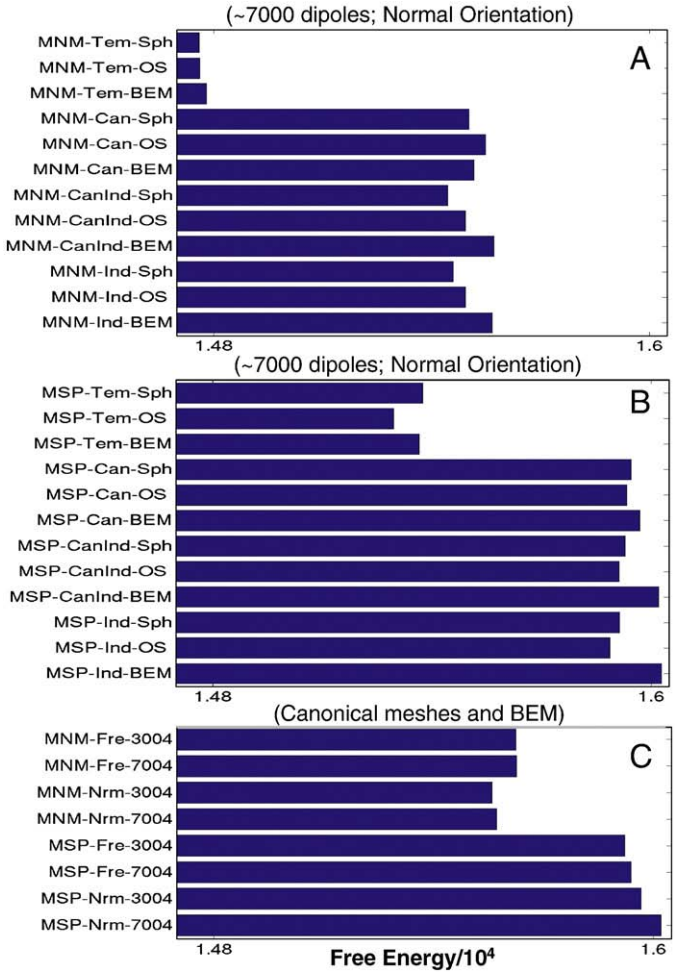

Fig. 2. Mean free-energy (arbitrary units) across the nine participants for: (A) the conditions explored in Analysis 1, when using a standard minimum norm (MNM) source-prior; (B) the conditions explored in Analysis 1 when using multiple sparse priors (MSP), and (C) the conditions explored in Analysis 2. **Tem** = template cortex, inner skull and scalp meshes; **Can** = canonical cortex, inner skull and scalp meshes; **CanInd** = canonical cortex mesh and individual skull and scalp meshes; **Ind** = individual cortex, inner skull and scalp meshes (see Fig. 1). **Sph** = single-sphere head-model; **OS** = overlapping-spheres head-model; **BEM** = Boundary Element (Head) Model; **Nrm** = dipoles normal to cortical mesh; **Fre** = dipole-orientation free to be estimated.

source-prior and head-model, $F(1.48,11.9) = 15.1$, $p < .001$, and between mesh and head-model, $F(2.28,18.3) = 3.58$, $p < .05$, and a trend for an interaction between Method and mesh, $F(1.24,9.94) = 3.99$, $p = .07$. These were explored in further ANOVAs for MSP and MNM source-priors separately. The 4×3 ANOVA for MNM source-priors showed a reliable interaction between head-model and meshes, $F(2.37,18.9) = 3.38$, $p < .05$ (even when excluding Template meshes, $F(1.47,11.8) = 6.15$, $p < .05$). This pattern appeared to reflect an advantage of BEM over sphere-based head-models, which became more pronounced for individual skull meshes (i.e., for **Ind** and **CanInd** versus **Can** conditions in Fig. 1A). Indeed, separate one-way ANOVAs on each set of meshes separately showed a reliable main effect of head-model for **Ind** meshes, $F(1.61,12.9) = 4.46$, $p < .05$, and a trend for **CanInd** meshes, $F(1.35,10.8) = 3.68$, $p = .073$ (but no trend for **Can** meshes, $F < 1$). This effect was confirmed by a reliable pairwise difference between BEM and Single-spheres for both **Ind** and **CanInd** meshes, $F(1,8)$'s > 5.14, $p < .05$ (though any improvement of BEMs over Overlapping-spheres did not reach significance, $F(1,8) < 2.60$, $p > .14$).

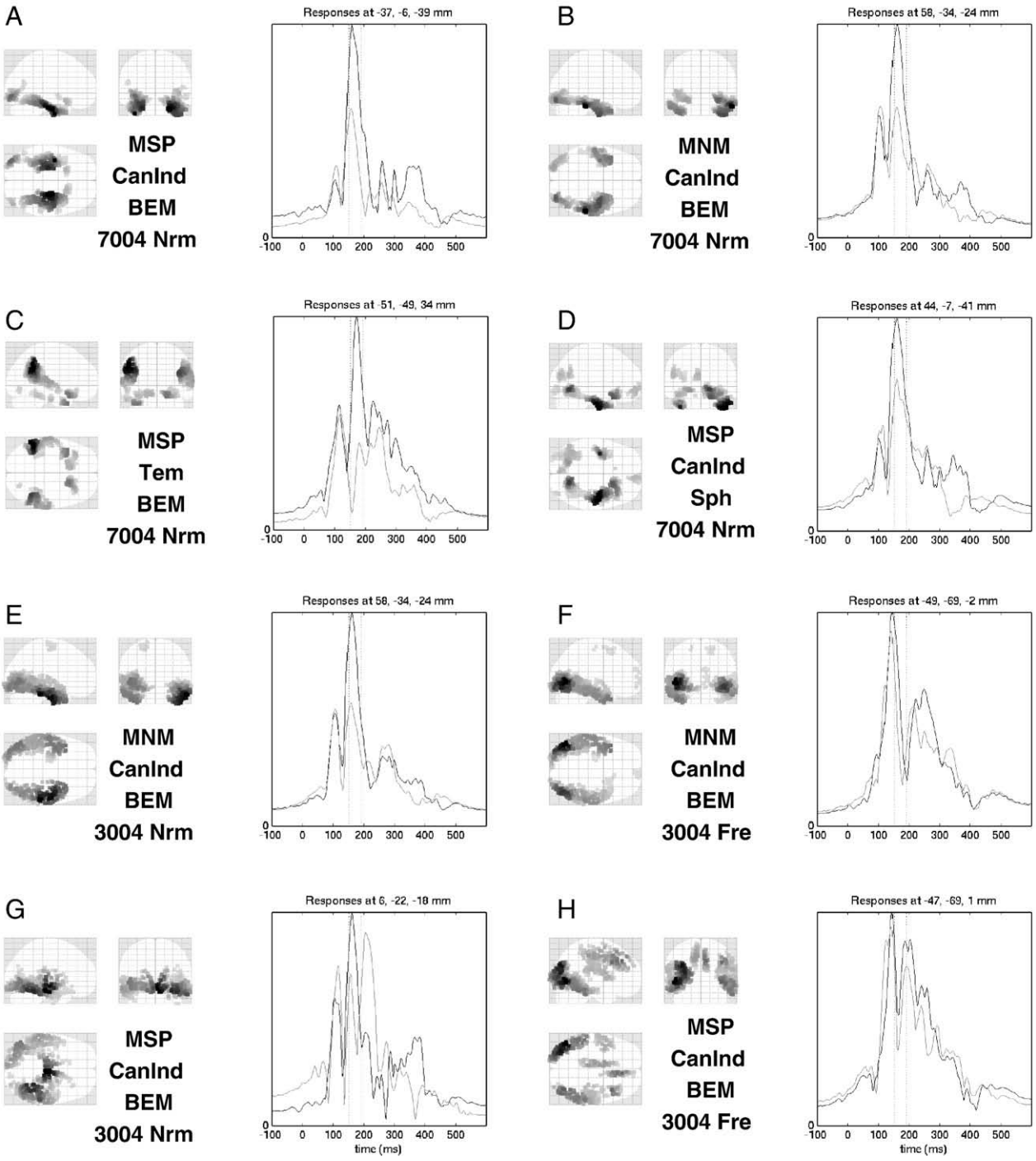

Fig. 3. Mean source solutions across participants for selected models from Analysis 1 and 2. The left part of each panel shows a Maximal Intensity Projection (MIP) of the 512 greatest source strengths within MNI space; the right part shows the magnitude of the evoked response to faces (dark lines) and scrambled faces (light lines) across the epoch for the dipole showing the biggest face-related response. For definition of acronyms, see Fig. 2 legend. Note only solutions with a Template or Canonical cortical mesh are shown, since only these have a one-to-one mapping with MNI space. For illustration purposes, the source estimates have been smoothed on the 2D mesh surface via 32 iterations of a graph Laplacian with adjacency ratio (autoregression coefficient) of 1/16.

The 4×3 ANOVA for MSP source-priors also showed a reliable interaction between head-model and meshes, $F(2.43,19.4)=3.55$, $p<.05$ (even when excluding Template meshes, $F(1.73,13.8)=5.96$, $p<.05$). This pattern again reflected an advantage of BEM over sphere-based head-models, which became more pronounced for Individual skull meshes (Fig. 1B). Indeed, separate one-way ANOVAs on each set of meshes separately showed a reliable main effect of head-model for **CanInd** and **Ind** meshes, $F>6.91$, $p<.05$ (but not for **Can** meshes, $F<1$), which were confirmed by reliable pairwise differences between BEM and both Single- and Overlapping-spheres for **CanInd** meshes, $F>6.83$, $p<.05$, and **Ind** meshes, $F>5.32$, $p<.05$.

To examine the choice of cortical mesh more closely, a final analysis was restricted to the **CanInd** and **Ind** conditions, which are matched in terms of the inner skull and scalp meshes that determine the head-model and data coregistration respectively. ANOVAs with the additional factor of head-model showed no reliable main effect or interaction involving canonical vs. individual cortical meshes, for either MNM or MSP source-priors, $F<1.98$, $p>.17$, nor was any reliable difference found between canonical and individual meshes when restricted to the best (i.e., BEM) head-model, $F<1$.

Figs. 3A–D show the resulting face-evoked responses around the latency of M170 component believed sensitive to face perception for selected models in MNI source space. The images on the left of each panel show Maximal Intensity Projections (MIPs) of the average source activity across participants for the 512 dipoles that show the greatest face-related activity; where the 'activity' reflects the difference of the evoked response magnitude for faces relative to scrambled faces within a Gaussian window between 150 and 190 ms. The plots on the right show the magnitude of activity across the whole epoch for faces (dark lines) and scrambled faces (light lines) for the dipole showing the biggest face-related response. Fig. 3A shows the solution for the optimal inversion (that with the highest free-energy); i.e., MSP with a canonical cortical mesh and a BEM defined on the individually-defined inner skull mesh.

Firstly, note the effect of source-priors (cf. Figs. 3A and B), in that the sparse solutions assumed by MSP encourage deeper sources (e.g., more medial in ventral temporal cortex) than the more superficial solutions characteristic of standard MNM (see also Friston et al., 2008). Secondly, note the effect of cortical mesh (cf. Figs. 3A and C), in the more realistic localisation of the M170 in ventral temporal regions using a Canonical cortical mesh than a Template cortical mesh. Thirdly, note the effect of head-model (when individually-defined; cf. Figs. 3A and D), in the differences between BEM and Single-sphere head-models, where the former appears to identify more occipital sources (possibly corresponding to the "OFA", Rossion et al., 2003), presumably because the single-sphere approximation is less accurate near the occipital pole.

Analysis 2: effects of source-priors, mesh-size and dipole-orientation

In the second analysis, three factors were explored: source-priors (MNM vs. MSP), mesh-size (7004 vs. 3004 dipoles) and dipole-orientation (Normal or Free). These eight models were based on a canonical cortical mesh, individual skull and scalp meshes and a BEM head-model (i.e., the optimal **CanInd-BEM** model from Analysis 1). The average free-energy across the nine participants is shown for each model in Fig. 2C.

As above there was a profound advantage of MSP over MNM, $F(1,8)=47.3$, $p<.001$ ($\eta^2=0.025$). The three-way interaction between source-priors, mesh-size and dipole-orientation was not significant, $F(1,8)=1.04$, $p=.34$, but the two-way interaction between source-priors and dipole-orientation was highly significant, $F(1,8)=29.5$, $p<.001$, with freely orientated sources increasing free-energy for MNM, but decreasing free-energy for MSP. The interaction between mesh-size and dipole-orientation was also significant, $F(1,8)=12.6$, $p<.01$, and the interaction between source-prior and mesh-size approached significance, $F(1,8)=5.23$, $p=.052$. These interactions were explored by separate ANOVAs on MSP and MNM source-priors. For MNM source-priors, there was only a highly reliable effect of dipole-orientation, $F(1,8)=47.9$, $p<.001$ (neither the main effect of mesh-size nor the interaction approached significance, $F<1.01$). The main effect of dipole-orientation reflected a greater free-energy for free, relative to normally-oriented sources, which was true in pairwise tests of orientation for both small and large mesh-sizes, $F(1,8)>31.1$, $p<.001$.

For MSP source-priors, the main effects of mesh-size and dipole-orientation were significant, $F(1,8)=7.34$, $p<.05$ and $F(1,8)=8.97$, $p<.05$ respectively and their interaction approached significance, $F(1,8)=4.40$, $p=.07$. This reflected the greatest free-energy for large meshes with normally-oriented dipoles. This pattern was clarified by reliable pairwise differences between the **MSP-Nrm-7004** model and each of the other three models, $F(1,8)>10.5$, $p<.05$, but no reliable differences between any of the other three models, $F(1,8)<2.56$, $p>.14$.

Figs. 3E–H shows the resulting face-evoked responses for selected models in MNI source-space. Firstly, note the small effect of mesh-size for standard minimum norm (cf. Figs 3E and B), yet a noticeable effect of free vs. normal orientation (cf. Figs. 3E and F), in that the maxima are more posterior without orientation constraints. For MSP, both smaller meshes (cf. Figs. 3G and A) and free-orientations (cf. Figs. 3H and A) result in less plausible solutions, consistent with their lower free-energy.

Discussion

Using a free-energy approximation to the Bayesian model-evidence and MEG data from nine participants, we compared different forward (generative) models within the same Parametric Empirical Bayesian framework (described in Appendix). We used a source-space in which several thousand dipoles were constrained to a tessellated neocortical manifold, and reconstructed the source activity over 172 epochs of 700 ms. We found greater model-evidence for MSP models that assumed multiple sparse sources (Friston et al., 2008), relative to MNM models that assumed a single uniform spatial prior across sources (corresponding to the standard Minimum Norm approach). Note that while both MSP and MNM had the same number of parameters (i.e., dipoles on the cortical mesh), MSP had many more hyperparameters (~750 vs 1), so is the more complex model. Importantly, the measure of model evidence penalizes model complexity, and yet MSP still had a higher model evidence than MNM, by virtue of being a more accurate model of the data covariance (see Appendix). This greater model-evidence was accompanied by a more realistic source reconstruction for the increase in evoked activity around 170 ms for faces relative to scrambled faces; namely in ventral temporal regions close to the fusiform gyrus, compared to the more superficial reconstructions that characterise the standard MNM approach. These MSP results confirm and extend prior conclusions from a single-participant EEG dataset (Friston et al., 2008).

Second, while we found evidence that the cortical mesh obtained from an individual's native MRI was superior to a "template" mesh (from a different brain in Talairach space), we found no reliable evidence that this individual cortical mesh was superior to a "canonical" mesh obtained by inverse-normalising the template cortical mesh (using normalisation parameters derived from warping the individual MRI to the template MRI). This was the case for both MSP and MNM source-priors. The lack of difference between individual and canonical cortical meshes held even when directly comparing our **Ind** and **CanInd** conditions, in which the inner-skull and scalp meshes were equivalent (individually-defined), and only the cortical mesh differed. This is an important result because it suggests that creating individual cortical meshes (and all the difficulties that this entails) can be an unnecessary exercise, in that

it does not necessarily improve the ensuing forward models relative to canonical mesh-based models.

These results concerning cortical meshes extend our prior claims from a single-participant analysis (Mattout et al., 2007). The inverse-normalisation of a template mesh can never be as accurate as an individually-defined cortical mesh (when the latter is done carefully), because normalisation typically only matches brains to a certain spatial scale (typically ~1 cm). Thus the inability to distinguish the two meshes empirically is likely to reflect the under-determinacy of the inverse problem (i.e., that there is simply insufficient information in the sensors to distinguish these two source spaces). Another perspective is that if the forward model is poor, or inversion assumptions are incorrect, then it does not matter which mesh is used. Nonetheless, the sufficiency of canonical meshes is important to establish, because construction of accurate cortical meshes directly from MRIs is time-consuming and often requires manual intervention. Moreover, the use of canonical meshes provides a one-to-one mapping between the source solutions in an individual's space and the template space, and hence a one-to-one mapping across individuals, which facilitates group analyses (Litvak and Friston, 2008). For example, it allows the individual source solutions to be written directly as a 3-dimensional image in the template space and then entered into a group-level SPM analysis (as is standard for summary measures of fMRI activity) (Henson et al. 2007). It also allows spatial constraints that live in the template space, such as fMRI results from group studies, to be easily applied to new MEG/EEG data (Flandin et al., 2009).

Third, we found that BEMs increase the model-evidence compared to single or overlapping spherical models, and led to more plausible reconstructions, suggesting that BEMs can justify the extra computation entailed. Note however, that the increase in model-evidence for BEMs was conditional on the use of individual skull and scalp meshes. This makes sense, given that spatial normalisation of MRIs (in SPM5) is based on matching grey- and white-matter segments to corresponding segments in template space (not on matching the skull or scalp). These normalisation parameters are therefore unlikely to be optimal for inverse-normalisation of a template skull and scalp. Thus both the BEM (aligned to the inner skull mesh) and the coregistration of MEG and MRI data (via aligning the digitised head shape to the scalp mesh) will be better for individual skull and scalp meshes.[2] Indeed, it would be informative to separate the relative contribution of these two effects (Lecaignard et al. 2008). Note also that the inner skull and scalp meshes are much easier to create automatically from MRIs by conventional shrink-wrap algorithms, because they are relatively smooth, unlike the highly convoluted surface of the cortex.

Fourth, we found interesting effects of cortical mesh size (3000 vs. 7000 vertices) and whether or not the dipoles in those meshes were constrained to be normal to the mesh surface. These results depended on the source-priors. For standard MNM, allowing the dipole component to be estimated for each orientation, rather than just the normal orientation, improved the model-evidence for both coarse and fine cortical meshes, while there was no reliable effect of mesh size. The effect on the source solutions was marked, moving the maximal signal magnitude for face-related activity more posteriorly in the brain. This may again reflect the bias towards superficial sources with MNM, if the greater flexibility in dipole orientation allows superficial sources to fit the data better.

For the MSP source-priors however, allowing free dipole-orientations or reducing the mesh-size both reduced the model-evidence, and led to less plausible source solutions. In other words, the use of multiple sparse priors works best when orientations are constrained and the cortical mesh density is closer to 7000 than 3000. The lack of a reliable increase in model evidence for free vs fixed orientations when using MSP priors may seem surprising, since there are bound to be errors in estimation of the surface normal, and small errors in dipole orientation can have large effects on the forward model (Phillips et al., 2005; Salayev et al., 2006). One reason for this may be that another location close to the true source has by chance an orientation close to that of the true source (i.e., orientation errors trade-off against small location errors). This would seem more likely to be the case for denser meshes, consistent with our results. But if this were not true, there may be much larger mislocalisations. While large mislocalisations were not obvious in the inversions of the present data (when assuming fixed orientations and the larger mesh), given the sources expected from previous studies, the only way to test this properly is with simulations of known sources.

The above observations speak to finding models of the optimum complexity. They suggest that MSP models cover optimal models; whereas MNM models do not. MSP is more flexible than MNM since it allows different variances for different locations in the brain. However it enforces neighbouring sources to covary, which allows MSP to emulate equivalent current dipole models, should they be the best explanation of the data. Critically, adding more parameters to MNM models improves them (either by adding more dipoles or more moments per dipole). Conversely, for MSP, increasing the number of dipoles improves the model, but reducing the degrees of freedom (by enforcing normal dipole orientation) makes these models even better. This makes sense when one considers MSP as a flexible model that can optimize source orientation locally by weighting the contribution of neighbouring dipoles with similar dynamics. Since dipoles in the same region have different orientations (under the normal constraint), they afford sufficient degrees of freedom to fit the regional source orientation.

Note that our inferences are based on the model evidence, which takes into account both the data fit and model complexity; if accuracy of localization were the sole criterion, more complex models (e.g, with free orientation, or higher mesh densities) might be justified in some cases. For example, while model-evidence is a principled metric in the present context, there are other important criteria, such as an inverse model's predictive validity and the reproducibility of its results across datasets. Of particular importance for future work will be to investigate further the role of fixed vs free orientations of distributed dipoles under sparse spatial priors using simulations. Note also that all of the above findings are restricted to the data we examined, and may not generalise to other datasets. For example, individually-defined cortical meshes may prove superior to canonical meshes for accurate localisation of very focal sources (e.g., for dipolar responses to the early response evoked by median-nerve stimulation; Wood et al., 1985). Nonetheless, the present data were chosen for their slightly later and more dispersed perception-based contrast (i.e., the M170 for faces vs. scrambled faces), and for group-level inferences in a normalised space, for which such precise localisation is less important. Future work may show whether the present findings generalise to other datasets, but in the absence of such tests, we expect that our findings will be a useful interim guide to MEG researchers when specifying their generative models.

Conclusion

Several recent methodological developments have been proposed for source reconstruction of MEG/EEG data, but the solutions furnished by these inversions are only as good as the generative model that is inverted. In relation to the present data and range of options explored, the optimal generative model was one that assumed Multiple Sparse Priors, a Boundary Element Model based on an individually-defined inner skull mesh, an individually-defined scalp

[2] One caveat when using a canonical cortical mesh with a BEM based on an individually-defined inner skull mesh is to check that the cortical mesh lies completely within the inner skull mesh (since this is not guaranteed when the cortical mesh is created by inverse-normalisation), and to ensure that the distance between each source and the closest part of the inner skull mesh is greater than the distance between vertices of the inner skull mesh (Mosher et al, 1999).

mesh to align the MEG data with the MRI, and approximately 7000 dipolar sources constrained within, and oriented normal to, a cortical mesh. This was regardless of whether the cortical mesh was defined individually or from inverse-normalising a template mesh (i.e., a canonical mesh). Thus, we have demonstrated again the superiority of multiple sparse priors over conventional minimum norm approaches to source reconstruction and the sufficiency of canonical meshes relative to individual cortical surface extraction.

Software note

All the inversion routines described in this paper are available freely as part of the SPM academic software (http://www.fil.ion.ucl.ac.uk/spm).

Acknowledgments

This work is funded by the UK Medical Research Council (WBSE U.1055.05.012.00001.01). We thank Jean-Francois Mangin for help with constructing individual cortical meshes, and Gareth Barnes, Krish Singh and Arjan Hillebrand for help with data acquisition.

Appendix

We assume a hierarchical linear model with Gaussian errors that can be formulated in a "Parametric Empirical Bayes" (PEB) framework (Phillips et al. 2005). This corresponds to a two-level model, with the first level representing the sensors and the second level representing the sources:

$$\mathbf{Y} = \mathbf{L}\mathbf{J} + E^{(1)} \qquad E_1 \sim N\left(0, V, C^{(1)}\right)$$

$$\mathrm{J} = 0 + E^{(2)} \qquad E_2 \sim N\left(0, V, C^{(2)}\right)$$

where $\mathbf{Y}$ is an n (sensors) $\times t$ (time points) matrix of sensor data; $\mathbf{L}$ is a $n \times p$ (sources) matrix representing the "forward model", and $\mathbf{J}$ is the $p \times t$ matrix of unknown dipole currents; i.e., the model parameters that we wish to estimate. E_1 and E_2 represent zero-mean, multivariate Gaussian distributions that assume a spatiotemporal factorisation into temporal covariance, V, and spatial covariances $C^{(1)}$ and $C^{(2)}$ (Friston et al., 2006).

The spatial covariance matrices are represented by a linear combination of N covariance components, Q_j:

$$C^{(i)} = \sum_{j=1}^{N} \lambda_j^{(i)} Q_j^{(i)}$$

where $\lambda_j^{(i)}$ is the "hyperparameter" for the j-th component of the i-th level. At the sensor level we assume white noise by setting $Q^{(1)} = I_{(n)} \Rightarrow C^{(1)} = \lambda^{(1)}\mathbf{I}_{(n)}$, where $\mathbf{I}_{(n)}$ is a $n \times n$ identity matrix. $C^{(2)}$ represents a spatial prior on the sources. It can be shown that the standard minimum norm solution corresponds to setting:

$$Q^{(2)} = I_{(p)} \Rightarrow C^{(2)} = \lambda^{(2)} I_{(p)}$$

Alternatively, in the "multiple sparse priors" (MSP) approach:

$$Q_j^{(2)} = q_j q_j^T \Rightarrow C^{(2)} = \sum_{j=1}^{N} \lambda_j^{(2)} q_j q_j^T$$

where q_j is the j-th column regularly sampled from a $p \times p$ matrix, $\mathbf{G}$, that codes the proximity of sources within the cortical mesh. $C^{(2)}$ therefore represents N cortical patches, where N is typically several hundred (see Friston et al., 2008, for more details).

The data, Y, are projected onto a small number (typically 6–8) temporal modes over the epoch using singular-value decomposition, in order to equate the temporal correlations at sensor and source levels within this subspace (see Friston et al., 2006, for more details). A similar projection can be performed onto a spatial subspace based on the lead-field matrix (Friston et al., 2008); however, this latter step was not performed for the present purposes of comparing different lead-field matrices.

We also add hyperpriors on the hyperparameters, for example to ensure positive covariance components. The latter is achieved by a log-normal hyper-prior, where $\alpha_i = \ln(\lambda_i) \Leftrightarrow \lambda_i = \exp(\alpha_i)$ and $p(\alpha) = N(\eta, \Omega)$ (Henson et al., 2007).

The generative model is then given by $M = \{\mathrm{L}, Q_j^{(i)}\}$. Because the priors factorise, maximising the model-evidence, $p(Y|M)$, is equivalent to maximising:

$$\ln p(Y|M) = \ln \int p(Y, J|M) dJ \approx F$$

where F is the variational "free-energy", and is equal to (bar a constant):

$$F = \frac{1}{2}\left(-tr\left(C^{-1}YY^T\right) - \ln|C| - (\alpha-\eta)^T\Omega^{-1}(\alpha-\eta) + \ln|\textstyle\sum\Omega^{-1}|\right)$$

where $C = LC^{(2)}L^T + C^{(1)}$, and $\sum$ is the posterior covariance of the hyperparameters (see Friston et al., 2007, for details). F can also be considered as the difference between the model accuracy (the first two terms) and the model complexity (the second two terms).

F can be maximised using standard variational schemes such as Expectation Maximisation (EM) to furnish a tight bound approximation to the log-evidence (given the linear, Gaussian model, Friston et al., 2007; see also Wipf and Nagarajan, 2009; Friston et al., 2008), which also yield posterior estimates of the hyperparameters and, in turn, the parameters:

$$\hat{\lambda} = EM\left\{YY^T, \widetilde{Q}\right\} \quad \widetilde{Q} = \left\{Q^{(1)}, LQ_1^{(2)}L^T, LQ_2^{(2)}L^T, \ldots\right\}$$

$$\hat{J} = \hat{C}^{(2)}L^T\hat{C}^{-1}Y \quad \hat{C} = \hat{C}^{(2)} + \lambda^{(1)}Q^{(1)} \quad \hat{C}^{(2)} = \textstyle\sum_j \hat{\lambda}_j^{(2)} LQ_j^{(2)}L^T.$$

References

Allison, T., Puce, A., Spencer, D.D., McCarthy, G., 1999. Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. Cereb. Cortex 9, 415–430.
Ashburner, J., Friston, K.J., 2005. Unified segmentation. NeuroImage 26, 839–851.
Chen, C.C., Henson, R., Stephan, K.E., Kilner, J.M., Friston, K.J., (2009). Forward and backward connections in the brain: a DCM study of functional asymmetries in face processing. Neuroimage. 45, 453-462.
Dale, A.M., Sereno, M., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI surface reconstruction: a linear approach. J. Cogn. Neurosci. 5, 162–176.
Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. 8, 272–284.
Flandin, G., Henson, R., Daunizeau, J., Friston, K., Mattout, J., 2009. A generic framework for fMRI-constrained MEG source reconstruction. Abstract at Human Brain Mapping Meeting.
Friston, K., Henson, R., Phillips, C., Mattout, J., 2006. Bayesian estimation of evoked and induced responses. Hum. Brain Mapp. 27, 722–735.
Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W., 2007. Variational free-energy and the Laplace approximation. NeuroImage 34, 220–234.
Friston, K.J., Harrison, L., Daunizeau, J., Kiebel, S.J., Phillips, C., Trujillo-Bareto, N., Henson, R.N., Flandin, G., Mattout, J., 2008. Multiple sparse priors for the M/EEG inverse problem. NeuroImage 39, 1104–1120.
Hauk, O., 2004. Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data. NeuroImage 21, 1612–1621.
Henson, R.N., Goshen-Gottstein, Y., Ganel, T., Otten, L.J., Quayle, A., Rugg, M.D., 2003. Electrophysiological and haemodynamic correlates of face perception, recognition and priming. Cereb. Cortex 13, 793–805.
Henson, R.N., Mattout, J., Singh, K.D., Barnes, G.R., Hillebrand, A., Friston, K.J., 2007. Population-level inferences for distributed MEG source localization under multiple constraints: application to face-evoked fields. NeuroImage 38, 422–438.
Huang, M.X., Mosher, J.C., Leahy, R., 1999. A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. Phys. Med. Biol. 44 (2), 423–440.

Lecaignard, F., Bouet, R.M., Mattout, J., 2008. Comparing models of cortical anatomy for MEG source reconstruction. Abstract, Biomag 2008.
Lin, F.-H., Belliveau, J.W., Dale, A.M., Hämäläinen, M.S., 2006. Distributed current estimates using cortical orientation constraints. Hum. Brain Mapp. 27, 1–13.
Litvak, V., Friston, K., 2008. Electromagnetic source reconstruction for group studies. NeuroImage 42, 1490–1498.
Mattout, J., Henson, R.N., Friston, K.J., 2007. Canonical source reconstruction for MEG. Computational Intelligence and Neuroscience, Article ID67613.
Mosher, J.C., Leahy, R.M., Lewis, P.S., 1999. EEG and MEG: forward solutions for inverse methods IEEE Trans. Biomed. Eng. 46, 245–259.
Nunez, P., 1981. Electric Fields Of The Brain: The Neurophysics Of Eeg. Oxford University Press, New York.
Phillips, C., Mattout, J., Rugg, M.D., Maquet, P., Friston, K.J., 2005. An empirical Bayesian solution to the source reconstruction problem in EEG. NeuroImage 24, 997–1011.
Pruist, G.W., Gilding, G.H., Peters, M.J., 1993. A comparison of different numerical methods for solving the forward problem in EEG and MEG. Physiol. Meas. 14, Al–A9.
Rossion, B., Caldara, R., Seghier, M., Schuller, A.-M., Lazeyrasm, F., Mayer, E., 2003. A network of occipito-temporal face sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. Brain 126, 1–15.
Sarvas, J., 1987. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys. Med. Biol. 32, 11–22.
Salayev, K.A., Nakasato, N., Ishitobi, M., Shamoto, H., Kanno, A., Iinuma, K., 2006. Spike orientation may predict epileptogenic side across cerebral sulci containing the estimated equivalent dipole. Clin. Neurophysiol. 117 (8), 1836–1843.
Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. George Thieme Verlag, Stuttgart, pp. 1–122.
Watanabe, S., Miki, K., Kakigi, R., 2005. Mechanisms of face perception in humans: a magneto- and electro-encephalographic study. Neuropathology 25, 8–20.
Wipf, D., Nagarajan, S., 2009. A unified Bayesian framework for MEG/EEG source imaging. NeuroImage. 44, 947-966.
Wood, C.C., Cohen, D., Cuffin, B.N., Yarita, M., Allison, T., 1985. Electrical sources in human somatosensory cortex: identification by combined magnetic and potential recordings. Science 227, 1051–1053.