

## MEG and EEG data fusion: Simultaneous localisation of face-evoked responses

Richard N. Henson<sup>a,\*</sup>, Elias Mouchlianitis<sup>a</sup>, Karl J. Friston<sup>b</sup>

<sup>a</sup> MRC Cognition & Brain Sciences Unit, Cambridge, England, UK

<sup>b</sup> Wellcome Trust Centre for Neuroimaging, University College London, England, UK

### ARTICLE INFO

#### Article history:

Received 12 January 2009

Revised 30 March 2009

Accepted 14 April 2009

Available online 3 May 2009

#### Keywords:

Inverse problem

Bayesian

### ABSTRACT

We present an empirical Bayesian scheme for distributed multimodal inversion of electromagnetic forward models of EEG and MEG signals. We used a generative model with common source activity and separate error components for each modality. Under this scheme, the weightings of error for each modality, relative to source components, are estimated automatically from the data, by optimising the model-evidence. This obviates the need for arbitrary user-defined weightings. To evaluate the scheme, we acquired three types of data simultaneously from twelve participants: total magnetic flux (as recorded by 102 magnetometers), orthogonal in-plane gradients of the magnetic field (as recorded by 204 planar gradiometers) and voltage differences in the electrical field (recorded by 70 electrodes). We assessed the relative precision of each sensor-type in terms of signal-to-noise ratio (SNR); using empirical sample variances and optimised estimators from the generative model. We then compared the localisation of face-evoked responses, using each modality separately, with that obtained by their “fusion” under the common generative model. Finally, we quantified the conditional precisions of the source estimates using their posterior covariance, confirming that EEG can improve MEG-based source reconstructions.

Crown Copyright © 2009 Published by Elsevier Inc. All rights reserved.

### Introduction

Reconstructing the generators of electroencephalographic (EEG) or magneto-encephalographic (MEG) data is an ill-posed ‘inverse problem’ (Nunez, 1981). Several studies have shown that MEG and EEG data provide supplementary information, both theoretically (Hämäläinen et al., 1988; Fokas, 2008), and practically, in the sense that simultaneous inversion of both affords more accurate reconstructions than unimodal inversions (e.g., Fuchs et al., 1998; Baillet et al., 1999; Liu et al., 2002; Bablioni et al., 2001, 2004; Huang et al., 2007; Sharon et al., 2007; Molins et al., 2008). Multimodal inversion requires some form of weighting that determines the relative contribution of each modality to the estimates of source activity. This weighting usually depends on estimates of the noise, or signal-to-noise ratio (SNR), of each modality (e.g., Fuchs et al., 1998; Lin et al., 2006; Molins et al., 2008). However, pure sensor-level noise estimates are sometimes difficult to obtain (e.g., for EEG), and SNR estimates from real data, which often use pre-stimulus baseline periods to estimate “noise”, confound sensor noise with “brain noise” (i.e., endogenous neural activity).

Recently, we proposed a variant of the L2-norm distributed source approach, in which the weightings of covariance components (or ‘hyperparameters’) are estimated automatically from the data by maximising the negative free-energy, which is a lower bound on the Bayesian model-evidence (Phillips et al., 2005; Mattout et al., 2006).

Importantly, this method can handle multiple covariance components (akin to regularisation terms), which play the role of priors in source space and error components in sensor space (Henson et al., 2007). Here, we extend this framework to model multiple error components, one for each type of sensor (with two sensor-types for MEG and one for EEG). This allows one to invert a single generative model, in which MEG and EEG data arise from the same underlying neural current sources, but are expressed at the sensors via different forward models and error components. We begin by outlining this method and then introduce the dataset on which it is evaluated.

#### *An empirical Bayesian approach to MEG/EEG fusion*

We use a hierarchical linear model with Gaussian errors that can be formulated in a Parametric Empirical Bayes framework (Friston et al., 2002). For a single-participant, this corresponds to a two-level model, where the first level represents the sensors and the second represents the sources:

$$\begin{aligned} Y &= LJ + E^{(1)} \\ J &= 0 + E^{(2)}. \end{aligned} \quad (1)$$

Here,  $Y$  is a  $n$  (sensors)  $\times$   $t$  (time points) matrix of sensor data;  $L$  is a  $n \times p$  (sources) ‘lead-field’ matrix, or ‘forward model’, and  $J$  is the  $p \times t$  matrix of unknown dipole currents; i.e., the model parameters that we wish to estimate. There are typically thousands of such parameters but only a few hundred sensors. The currents are assumed to be normal to a tessellated mesh representing the surface of the

\* Corresponding author. MRC Cognition & 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](mailto:rik.henson@mrc-cbu.cam.ac.uk) (R.N. Henson).

neocortex. The random terms  $E^{(i)} \sim N(0, V, C^{(i)})$  are sampled from zero-mean, multivariate Gaussian distributions whose covariance factorises into temporal,  $V$ , and spatial components;  $C^{(1)}$  and  $C^{(2)}$  at the sensor and source level respectively (Friston et al., 2006).

The assumption of equal temporal correlations at sensor and source levels is assured by projecting the data to a subspace of  $r$  temporal modes using singular-value decomposition (SVD) (see Friston et al., 2006, for more details). A similar projection is performed onto a spatial subspace of  $m$  spatial modes, defined by a SVD of  $LL^T$  (Friston et al., 2008). Thus  $Y$  becomes an  $m \times r$  matrix (and  $L$  an  $m \times p$  matrix).

The spatial covariance matrices are represented by a linear combination of covariance components,  $Q_j^{(i)}$ :

$$C^{(i)} = \sum_{j=1}^N \lambda_j^{(i)} Q_j^{(i)}, \quad (2)$$

where  $\lambda_j^{(i)}$  is the ‘hyperparameter’ for the  $j$ -th component of the  $i$ -th level.<sup>1</sup> At the source-level,  $C^{(2)}$  represents a spatial prior, and it can be shown that the standard ‘minimum norm’ solution corresponds to setting  $Q^{(2)} = I_p \Rightarrow C^{(2)} = \lambda^{(2)} I_p$ , where  $I_p$  is a  $p \times p$  identity matrix (Hauk, 2004). Here however, we employ multiple sparse priors (MSP):  $Q_j^{(2)} = q_j q_j^T \Rightarrow C^{(2)} = \sum_{j=1}^N \lambda_j^{(2)} q_j q_j^T$ , where  $q_j$  is the  $j$ -th column regularly sampled from a  $p \times p$  matrix that encodes the proximity of sources within the cortical mesh. This corresponds to the assumption that there are a number of local ‘patches’ of cortex in which activity co-occurs (see Friston et al., 2008, for details). In this paper, these spatial priors are the same for all models and we concentrate on the sensor-level error covariance components for the multiple sensor-types.

Given that lead-fields can be created for each sensor and encoded in a matrix for each sensor-type, the first level of the above hierarchical linear model can be extended to  $d$  sensor-types as:

$$\begin{bmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \\ \vdots \\ \tilde{Y}_d \end{bmatrix} = \begin{bmatrix} \tilde{L}_1 \\ \tilde{L}_2 \\ \vdots \\ \tilde{L}_d \end{bmatrix} J + \begin{bmatrix} E_1^{(1)} \\ E_2^{(1)} \\ \vdots \\ E_d^{(1)} \end{bmatrix} \quad (3)$$

where the data from the  $i$ -th sensor-type,  $\tilde{Y}_i$ , have been scaled (see below) and stacked into a single matrix: Similarly for the modality-specific lead-field matrices,  $\tilde{L}_i$ . Note that the  $p$  sources in  $J$  are unchanged. The temporal subspace is represented by concatenating the  $r_i$  temporal modes of each sensor-type (after projection to the  $m_i$  spatial modes), so that the concatenated data become a  $\sum m_i \times \sum r_i$  matrix. The spatial covariance matrix of the sensor error,  $C^{(1)}$ , is again estimated by a linear combination of variance components:

$$C^{(1)} = \begin{bmatrix} C_1^{(1)} & 0 & \cdots & 0 \\ 0 & C_2^{(1)} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & C_d^{(1)} \end{bmatrix} \quad (4)$$

To accommodate different scaling and measurement units across the different sensor-types, the data and the forward model are re-scaled (after projection to the spatial and temporal modes) as follows:

$$\tilde{Y}_i = \frac{Y_i}{\sqrt{\frac{1}{m_i} \text{tr}(Y_i Y_i^T)}} \quad \tilde{L}_i = \frac{L_i}{\sqrt{\frac{1}{m_i} \text{tr}(L_i L_i^T)}} \quad (5)$$

where  $\text{tr}(X)$  is the trace of  $X$ . In the case of multiple trials, the scaling factor for  $\tilde{Y}_i$  is averaged over trials. This effectively normalises the data

<sup>1</sup> In fact,  $\lambda$  is expressed through a scale factor  $\alpha = \exp(\lambda)$ , to ensure that the weighting of each component is positive, in order to constitute a proper covariance matrix (see Henson et al., 2007, for further details).

so that the average second-order moment (i.e., sample variance if the data are mean-corrected) on each spatial mode;  $\frac{1}{m_i} \text{tr}(Y_i Y_i^T)$ , is one for all sensor-types and, in the absence of sensor noise, the average variance expected under independent and identical sources with unit variance,  $\frac{1}{m_i} \text{tr}(L_i L_i^T)$ , is one. This rescaling means that the hyperparameters  $\lambda_i^{(1)}$  for the error components in each sensor-type can be compared quantitatively. We address the validity of this scaling in the Results.

Restricted Maximum Likelihood (ReML) estimates of the hyperparameters,  $\hat{\lambda}_{ij}^{(k)}$ , together with Maximum A Posteriori (MAP) estimates of source activity,  $\hat{J}$ , are obtained by maximising the negative free-energy using an iterative algorithm (see Friston et al., 2007, for details). This also furnishes estimates of the posterior covariance of the parameters, which quantify the conditional precision (inverse variance) of the estimated source activity.

### The multimodal data

We acquired data from three different types of sensor: magnetometers (MEG), gradiometers (MEG) and electrodes (EEG). All data were recorded and digitised simultaneously using a VectorView system (Elekta Neuromag, Helsinki). The EEG data were acquired from 70 electrodes placed according to the standard extended 10–10% system, and re-referenced to the average over electrodes. The magnetometers measure the magnetic flux passing through a single pick-up coil, whereas gradiometers measure the difference in flux along a given axis using two counter-wound coils. In the VectorView system, readings from a magnetometer and two, orthogonal, planar gradiometers are obtained from each of 102 locations above the head. Because magnetometers measure total flux (Tesla), while gradiometers measure a gradient (Tesla per meter), the two different types of MEG data are incommensurate (in the absence of a forward model). One solution is to rescale the data to common units of noise, based on factory-measured intrinsic noise levels (MaxFilter Users Guide, Version 2.0, Elekta Neuromag, Appendix B2), or on empirical estimates of noise (e.g., from empty-room data, or pre-stimulus baseline periods). Our approach represents a principled alternative that optimises the noise estimates for a given dataset and forward model.

## Methods

### Participants and paradigm

Twelve participants (six female) were tested, aged 20–36. The study protocol was approved by a local ethics review board (CPREC reference 2005.08). The paradigm was identical to that used previously with EEG and fMRI (Henson et al., 2003) and MEG (Henson et al., 2007). Here, we analyse new data from a single, 11 min session, in which participants saw intact or scrambled faces, subtending visual angles of approximately 4°. One half of the faces were famous; one half were non-famous. Given that famous and non-famous faces did not differ over the critical M170 time window in these data, or in previous EEG data (Henson et al., 2003), famous and non-famous faces were pooled. Scrambled versions of each face were created by phase-shuffling in Fourier space and masking by the outline of the original image (to match size). The scrambled faces were therefore approximately matched for spatial frequency power density. Participants made left–right symmetry judgments about each stimulus by pressing two keys with either their left and right index finger or left and right middle finger (mean reaction times were approximately 1 s). The first four trials were discarded to exclude practice effects, leaving 84 intact and 84 scrambled face trials.

### Data acquisition

The MEG data were collected with a VectorView system (Elekta-Neuromag, Helsinki, Finland), containing a magnetometer and two

orthogonal, planar gradiometers located at each of 102 positions within a hemispherical array situated in a light Elekta-Neuromag magnetically-shielded room. The position of the head relative to the sensor array was monitored continuously by feeding sinusoidal currents (293–321 Hz) into four Head-Position Indicator (HPI) coils attached to the scalp. The simultaneous EEG was recorded from 70 Ag–AgCl electrodes placed within an elastic cap (EASYCAP GmbH, Herrsching-Breitbrunn, Germany) according to the extended 10–10% system and using a nose electrode as the recording reference. Vertical and horizontal EOG were also recorded. All data were sampled at 1 kHz with a band-pass filter from 0.03–330 Hz.

A 3D digitizer (Fastrak Polhemus Inc., Colchester, VA) was used to record the locations of the EEG electrodes, the HPI coils and approximately 50–100 ‘head points’ along the scalp, relative to three anatomical fiducials (the nasion and left and right pre-auricular points). MRI images were obtained using a GRAPPA 3D MPRAGE sequence (TR = 2250 ms; TE = 2.99 ms; flip-angle = 9°; acceleration factor = 2) on a 3 T Trio (Siemens, Erlangen, Germany) with 1 mm isotropic voxels.

Twelve sets of empty-room MEG recordings of 1–5 min were also collected within a few days of each participant, to estimate intrinsic noise levels.

#### Data pre-processing

External noise was removed from the MEG data (including empty-room data) using the temporal extension of Signal-Space Separation (SSS; Taulu et al., 2005) as implemented with the MaxFilter software (Elekta-Neuromag). The MEG data were also compensated for movement every 200 ms. The total translation across the session ranged from 0.4–28.8 mm across participants (median = 2.8 mm). The rank of the data covariance after pre-processing with MaxFilter varied between 64–68 for magnetometers, and 64–69 for gradiometers.<sup>2</sup>

Manual inspection identified some bad channels (numbers ranged across participants from 0–11 in the case of MEG, and 0–7 in the case of EEG). These were recreated by MaxFilter in the case of MEG, but rejected in the case of EEG. The EEG data were re-referenced to the average over remaining channels. After uploading to SPM5 (<http://www.fil.ion.ucl.ac.uk/spm>), the continuous data were down-sampled to 200 Hz (using an anti-aliasing filter), further low-pass filtered to 40 Hz in both forward and reverse directions using a 5th-order Butterworth digital filter, and then epoched from –100 ms to 400 ms post-stimulus onset (removing the mean baseline from –100 ms to 0 ms). Epochs in which the EEG or EOG exceeded 100  $\mu$ V were rejected from both EEG and MEG datasets (number of rejected epochs ranged from 0 to 47 across participants, median = 5), leaving 79 face epochs and 79 scrambled face epochs, on average across participants.

MRI images of each participant were segmented and spatially normalised to an MNI template brain in Talairach space using SPM5 (Ashburner and Friston, 2005). The inverse of the normalisation transformation was then used to warp a cortical mesh from a template brain in MNI space to each participant's MRI space (see Mattout et al., 2007 for further details). This ‘canonical’ mesh was a continuous triangular tessellation of the grey/white matter interface of the neocortex (excluding cerebellum) created from a canonical MRI

**Table 1**

Results of uni- and multimodal inversion, using variable numbers of spatial and temporal modes across modalities and participants (based on SVD).

	MEG		EEG
	Mags	Grds	
Number of sensors (range across participants)	102	204	63–70*
Number of spatial modes ( $m_i$ ) (range across participants)	57–78	83–120	54–65
Number of temporal modes ( $r_i$ ) (range across participants)	6–9	7–11	4–9
Empirical (baseline) SNR (plus standard error across participants)	4.0 (0.3)	4.3 (0.5)	5.1 (0.5)
<i>Unimodal inversions</i>			
% Data variance explained (range across participants)	95–97	84–95	91–97
Conditional source precision (plus standard error across participants)	51.1 (12.9)	33.8 (4.1)	73.1 (7.7)
<i>Fused inversion</i>			
% Data variance explained (range across participants)		91–95	
Conditional SNR (plus standard error across participants)	23.4 (2.3)	6.8 (0.4)	5.1 (0.4)
Conditional source precision (plus standard error across participants)		56.8 (9.1)	

Grds = Planar gradiometers; Mags = Magnetometers. \*Number of EEG sensors varied over participants owing to removal of bad sensors.

MPRAGE sequence in MNI space using FreeSurfer (Fischl et al., 2001). The surface was inflated to a sphere and down-sampled using octahedra to achieve a mesh of 8196 vertices (4098 per hemisphere) with a mean inter-vertex spacing of ~5 mm. 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).

The tissue segments in the native space of each participant's MRI were used to create a binary image of the outer scalp, which was tessellated into a mesh of 2002 vertices using SPM. The MEG and EEG data were projected onto each participant's MRI space by a rigid-body coregistration based on minimising the sum of squared differences between the digitised head points (and electrode positions for EEG) and this scalp mesh.

Brainstorm (<http://neuroimage.usc.edu/brainstorm>) was used to fit a single sphere (for MEG) or three concentric spheres (for EEG) to the scalp mesh (using the Berg and Scherg, 1994, approximation for EEG). Lead-fields were then calculated for a dipole at each point in the cortical mesh, oriented normal to that mesh.

#### Model inversion

The  $m_i$  spatial modes for the  $i$ -th modality were based on a SVD of the outer-product of the lead-field matrix and a cut-off of  $\exp(-16)$  for the normalised eigenvalues (which retains 99.99% of the variance), producing 54–120 spatial modes depending on the sensor-type (Table 1). The data for each trial (once projected onto these spatial modes) were reduced to  $r$  temporal modes for frequencies above 1 Hz, using SVD with a normalised cut-off of 1 (Friston et al., 2006). This produced 4–11 temporal modes across participants and sensor-types (Table 1), which captured a minimum of 91% of the data variance over all sensor-types and participants (mean = 96%). The data covariance across sensors was then estimated by averaging the covariance for each trial (faces and scrambled faces), entailing approximately 15,800 time points in total (Friston et al., 2006). This data covariance was then fit using multiple sparse priors (MSP) and remaining default options in SPM5 (i.e., 256 cortical patches per hemisphere, plus 256 bilateral patches, with a spatial autoregression coefficient of 0.6, a Gaussian temporal autocorrelation function for the noise with a standard deviation of 4 ms, and a Greedy

<sup>2</sup> SSS removes spatial components from the data that derive from an ‘external’ space beyond a sphere enclosing the sensors (using a spherical harmonic expansion of the magnetic field). Assuming a full harmonic expansion and that quasistatic Maxwell's equations hold, signals from within the ‘internal’ sphere within the sensors, i.e., brain signal, do not leak into the external part and thus there is no need to correct (spherical-based) leadfields following SSS. In practice, the spherical harmonics are truncated when they fall below the level of sensor noise, but the effects of this truncation have been shown to be negligible (Taulu et al., 2005).



Search algorithm; see Friston et al., 2008, for more details), though without a Hanning window, in order to compare with empirical SNR estimates based on the pre-stimulus baseline. This was repeated for each modality separately and for the multimodal fusion model. A time-frequency contrast for each source was then computed as the mean evoked response for faces and scrambled faces for all frequencies within a Gaussian window centred between 150–190 ms. For  $T$ -tests of metrics across participants, two-tailed  $p$ -values are quoted.

## Results

The mean sensor topographies across participants for the differential evoked response to faces versus scrambled faces for each sensor-type are shown in Fig. 1, together with waveforms from sensors exhibiting the greatest difference (light lines = faces, heavy lines = scrambled faces). A deflection peaking around 100 ms can be seen for both stimulus-types (M/P100), together with a second deflection peaking around 150 ms, which is greater for faces than scrambled faces (maximal difference is around 170 ms, corresponding to the “M/N170”). Note that the magnetometer topography shows the size and direction of magnetic flux; the gradiometer topography shows the scalar magnitude (vector length) of the gradient in the two in-plane directions; the EEG topography shows the potential difference relative to the average over all (valid) electrodes.

We present our results in three sections: in the first section, we consider just the MEG magnetometer and gradiometer data. This allowed the ReML estimates of sensor noise for magnetometers and gradiometers to be validated in relation to estimates from the empty-room MEG recordings. In the second section, we used these empty-room noise covariance estimates to simulate magnetometer and gradiometer data, in order to validate our scaling assumptions in Eq. (5). In the final section, we combined the EEG data with the MEG data to evaluate fusion in terms of: (i) the relative SNR of each modality; (ii) the relative contribution to the source estimates from each modality; (iii) the relative precision of the unimodal and multimodal source estimates and (iv) the effects of changing the number of spatial and temporal modes.

## Fusion of magnetometers and gradiometers

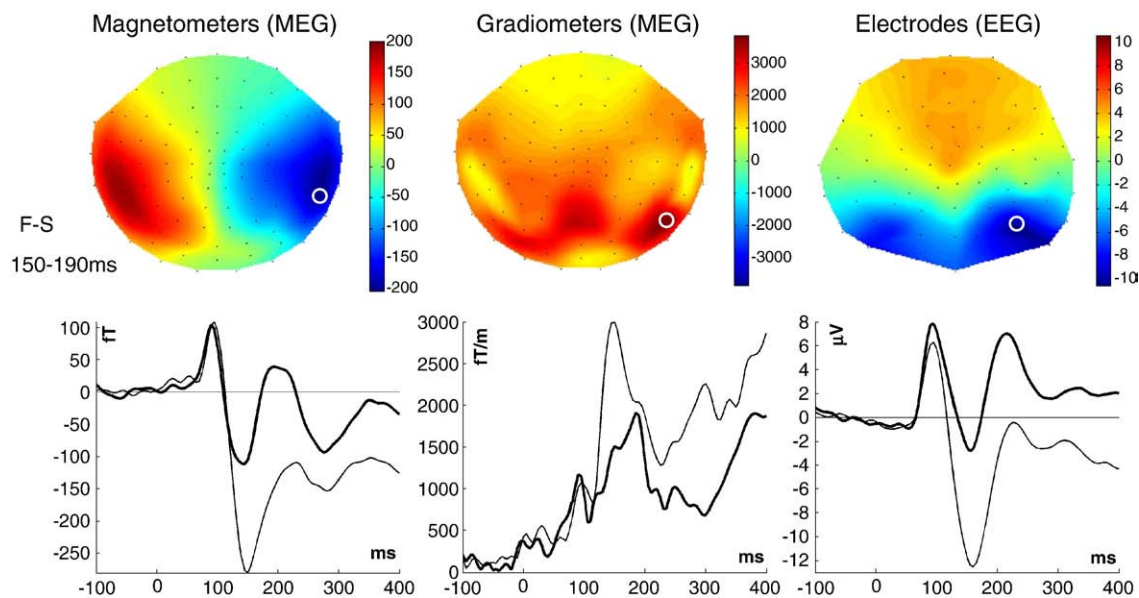
We begin by fusing the magnetometer and gradiometer data, assuming white noise at the sensor-level, i.e., a single variance component for each type of sensor,  $i = 1 \dots 2$ :

$$C_i^{(1)} = \lambda_i^{(1)} I_{m_i}, \quad (6)$$

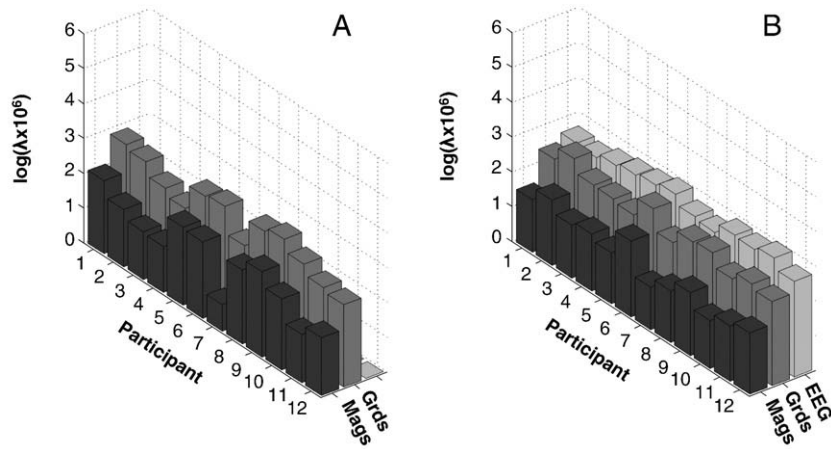
(see Eq. (4)), corresponding to independent and identically-distributed noise projected onto the  $m_i$  spatial modes associated with the lead-field matrix for each type of sensor. This resulted in larger sensor-level hyperparameter for gradiometers,  $\lambda_2^{(1)}$ , than for magnetometers,  $\lambda_1^{(1)}$ , which was consistent across participants,  $t(12) = 5.58$ ,  $p < .001$  (see Fig. 2A).

These estimates of sensor noise were compared with those estimated empirically from averaging across a number of empty-room recordings. Though the empirical noise covariance was not truly white (i.e., with some non-zero correlations between sensors, see Supplementary material for further details), the ratio of the standard deviation of the white component (i.e., square-root of the mean of the leading diagonal terms) for magnetometers relative to gradiometers ranged from 0.043 to 0.045 across twelve empty-room sessions. The mean ratio of 0.044 is close to the heuristic of 1/20 often quoted for the VectorView machine. The corresponding ratio for the hyperparameters across the real datasets, after undoing the scaling of the data in Eq. (5), ranged from 0.041 to 0.099 (with a mean of 0.079), which is in reasonable agreement with the empirical variance ratio. Note that a larger noise associated with the gradiometers than magnetometers is not a true reflection of differential sensitivity, but reflects simply their different physical units ( $T$  and  $T/m$  respectively).

The correspondence between the relative amplitude of noise in the presence of signal (hyperparameter estimates) and in their absence (sample variance of empty-room recordings) suggests that the covariance partitioning implicit in the inversion scheme is reasonably accurate. The source reconstructions for magnetometers, gradiometers and their fusion are shown in Figure S3 of Supplementary material. Also in the Supplementary material, we show how one can



**Fig. 1.** Top row: mean sensor-level topographies across participants (using a spherical projection and cubic spline interpolation) for each type of sensor for the difference between faces and scrambled faces between 150–190 ms. Bottom row: evoked waveforms from the sensor showing the maximal difference (highlighted by a white circle on topographies); light = faces, dark = scrambled faces. Note that for this figure (but not for source reconstruction), the gradiometer topography and waveform reflect the RMS value across the two in-plane, orthogonal gradiometers.



**Fig. 2.** ReML hyperparameter estimates for the sensor-level error variance component for each participant and each sensor-type when fusing magnetometer and gradiometer data (A) or when fusing magnetometer, gradiometer and EEG data (B). Mags = magnetometers; Grds = (planar) gradiometers.

add further covariance components at the sensor level – for example empirical estimates of the error covariance for each modality (such as those derived from independent empty-room data) – to capture dependencies between sensors of the same modality (i.e., non-white noise). In the next section, we combine the sources estimated from the above fusion of gradiometers and magnetometers with the empty-room estimates of noise covariance, to evaluate the validity of our scaling in Eq. (5).

*Evaluating our normalisation by simulation*

The normalisation of the lead-fields and data across modalities in Eq. (5) is essential for data fusion, because we need to weight the evidence in each modality when optimising the estimates of underlying source activity. For a single modality, the relative (within-modality) contribution of each channel is encoded by the columns of the lead-field matrix. However, the fused lead-field matrix in Eq. (3) contains lead-fields from different modalities, which may not have the same relative (between-modality) sensitivity or even the same physical units. Given that each modality is caused by the same source activity, a rough estimate of relative sensitivity is the standard deviation of modality-specific responses – which is essentially what we have used in Eq. (5)<sup>3</sup>. Note that this estimate preserves the relative sensitivity (and variations in sensitivity) of within-modality sensors. A problem with this approach is that we are assuming that the measured responses are caused entirely by source activity (i.e., that signal-to-noise is high; though not necessarily uniform). To lend face validity to this assumption, and the ensuing normalisation, we performed some simple simulations, to ensure the scaling estimate is within acceptable bounds.

We simulated responses for the gradiometers and magnetometers by using the conditional estimates of source activity estimated from their fusion above, and projected them through the *un-normalised* lead-field for each modality. We then added noise to each channel that was sampled randomly from a multivariate normal distribution with covariance derived from the empty-room measures,  $S_i$  (after projecting onto the  $m_i$  spatial modes):

$$\begin{aligned} Y_1 &= L_1 J + E_1 & E_1 &\sim N(0, S_1) \\ Y_2 &= L_2 J + E_2 & E_2 &\sim N(0, S_2) \end{aligned} \tag{7}$$

We repeated this 128 times for each participant to obtain an empirical distribution of the ratio:

$$\sqrt{\frac{\text{tr}(Y_1 Y_1^T)}{\text{tr}(Y_2 Y_2^T)}} \tag{8a}$$

which we then compared with the relative sensitivity of the lead-fields:

$$\sqrt{\frac{\text{tr}(L_1 L_1^T)}{\text{tr}(L_2 L_2^T)}} \tag{8b}$$

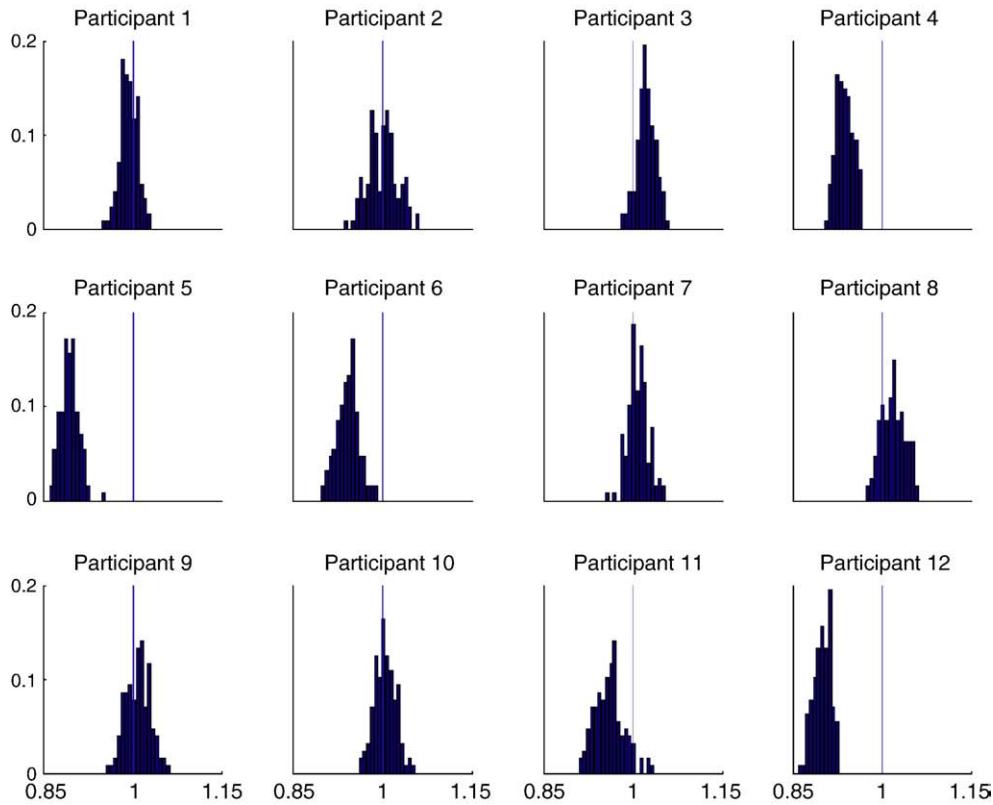
The results of these simulations (Fig. 3) show that the two ratios are nearly equal (i.e., their ratio is close to 1, with a range of 0.893 to 1.022 for the mean across participants, and 0.974 on average). In other words, there is a small bias of about 2.5% on average (and 11% at most). Note that these simulations are not trivial, because we projected the data onto temporal modes that are dominated by signal. It is therefore difficult to predict *a priori* how much the global power (variance),  $\text{tr}(Y_i Y_i^T)$ , is affected by noise. In summary, our simulation results suggest that our scaling normalisation is sufficient, at least for the present application (see Discussion). We now apply this normalisation and fusion scheme to multimodal data from both MEG and EEG.

*Fusion of magnetometers, gradiometers and EEG*

Having evaluated our method against empirically-derived measures of sensor noise for MEG and, in particular, our normalisation of between-modality sensitivity, we applied the method to simultaneously-recorded MEG–EEG data. It is more difficult to estimate sensor noise for EEG data: as we argue below, estimates based on, for example, a pre-stimulus baseline period are likely to conflate true sensor noise with brain “noise” reflecting ongoing (even if not stimulus-dependent) neural activity.

In this section, we consider the relative contribution of different modalities to the source estimates. We start by comparing SNRs from conventional measures (based on baseline sample variances) with optimised estimators (based on the conditional expectations of source activity). We then compare source reconstructions with and without multimodal constraints in terms of the spatial deployment of reconstructed activity. Importantly, we also compare unimodal and multimodal inversions in terms of the conditional precision of the source estimates to quantify the reduction in uncertainty about the amplitude of responses. Finally, we explore the effects of equating the number of spatial and temporal modes across modalities.

<sup>3</sup> Strictly speaking, Eq. (5) only pertains to the standard deviation if the data are mean-corrected, but the logic still holds, and the data are normally baseline-corrected and high-pass filtered in any case.



**Fig. 3.** Histograms of the ratio of the ratio of data power (Eq. (8a)) to the ratio of lead-field power (Eq. (8b)) for magnetometers relative to gradiometers from 128 simulated datasets (using noise covariance matrices derived from empty-room recordings), shown separately for each of the 12 participants. If the normalisation were perfect, the ratios should be distributed around 1 (see text for details).

#### Baseline empirical estimates of SNR

Empirical noise levels for each sensor-type were estimated by the sample standard deviation over the  $-100$  to  $0$  ms baseline period (after projection of the data onto the  $m_i$  spatial modes); signal levels were estimated by the standard deviation across all post-stimulus time points ( $0$  to  $400$  ms). These measures were averaged across spatial modes and trials, and converted into a signal to noise ratio (SNR) for each participant. The mean SNRs across participants are shown in the fourth row of Table 1. Two-tailed  $T$ -tests showed that the SNR for the EEG data was greater than for the MEG magnetometer or gradiometer data,  $T(11) > 2.50$ ,  $p < .05$ , but did not differ between magnetometer and gradiometer data,  $T(11) = 1.01$ ,  $p = .33$ . Note that this (conventional) estimate of SNR includes both sensor noise and endogenous or random fluctuations in signals from underlying neurophysiological sources. From the point of view of the forward model in Eq. (1), this is not a very useful estimate because “noise” includes source activity that fluctuates in the baseline period. Under the model in Eq. (1), “noise” is explicit sensor noise and cannot be estimated from the data unless there is no underlying brain activity. Nonetheless, these empirical estimates suggest that the SNR is quite similar over modalities, with EEG supervening slightly over MEG.

#### Conditional estimates of SNRs

A better estimate of modality-specific SNR can be obtained from the ReML estimates of the covariance components and the ensuing MAP estimates of the source activity. These estimates can be projected onto sensor space to quantify the variance explained by source activity (signal) relative to the residual variance (noise). To obtain the conditional mean and precision of source activity, we inverted each sensor-type separately and together. The models explained over 84% of the total data variance (or over 92% of variance in the spatiotemporal subspace) for all sensor-types and all participants, under all inversions; see fifth row (unimodal) and seventh (multimodal) row in Table 1.

The ReML estimates of the hyperparameters representing the sensor error variance when fusing the three sensor-types are shown in Fig. 2B. The relative values across sensor-types were remarkably consistent across participants. Indeed,  $T$ -tests across participants showed that the hyperparameter for the magnetometers was reliably smaller than that for the gradiometers,  $T(11) = 7.01$ ,  $p < .001$ , and that for the gradiometers was reliably smaller than that for the EEG data,  $T(11) = 4.38$ ,  $p < .005$ .

However, SNR also depends on the amount of data variance attributed to source activity. We therefore computed the source (signal) and error (noise) variance at the sensor level<sup>4</sup> using the conditional estimates of source activity,  $\hat{J}$ :

$$\begin{aligned} V_T &= \text{var}(\hat{Y}_i) \\ V_R &= \text{var}(\hat{Y}_i - L_i \hat{J}) \\ \text{SNR}_i &= \sqrt{\frac{V_T - V_R}{V_R}}. \end{aligned} \quad (9)$$

The resulting conditional SNR ratios (averaged over modes) are shown in the eighth row of Table 1. Consistent with the sensor-level noise hyperparameters, the conditional SNR estimates were greatest for the magnetometers and smallest for the EEG data, with pairwise  $T$ -tests significant in all cases,  $T(11) > 3.12$ ,  $p < 0.01$ . It is interesting that these optimised SNRs give exactly the opposite results when compared to the standard empirical estimates (cf. fourth and eighth rows of Table 1). One factor that may contribute to the relatively low SNR for EEG is the relatively high residual error,  $V_R$  (reflected in the sensor-level

<sup>4</sup> After projecting back from the temporal modes to the original time bins, in order to compare with the empirical SNR estimates from the pre- and post-stimulus periods.

hyperparameters) caused by the relatively poor forward model for EEG. This reflects the well-known fact that spherical head-model approximations are worse for EEG than MEG. A bigger approximation entailed by a single-sphere head-model for the gradiometers than for the magnetometers might also contribute to the worse conditional SNR estimate for gradiometers (see *Discussion*). Alternatively, the greater SNR for magnetometers could reflect the greater signal detectable by magnetometers, relative to gradiometers, for deeper sources that are typically engaged during face processing (see next section).

#### Source reconstructions

Source reconstructions representing the 128 sources showing the greatest increase in evoked responses for faces over scrambled faces for each sensor-type alone, together with the reconstruction under fusion, are shown in Fig. 4. One can see similar unimodal reconstructions of the two MEG sensor-types, with bilateral ventral temporal maxima (most likely corresponding to the mid-fusiform gyrus, or “fusiform face area”). The reconstruction of the EEG alone produced more posterior and lateral maxima in lateral occipito-temporal regions, particularly on the right (possibly including the “occipital face area”). The fused inversion (bottom right panel) shows aspects of both, with bilateral maxima in fusiform and a maximum in right lateral occipital, in addition to a maximum in left temporal pole. In short, fusion appears to produce a sensible balance between the modality-specific reconstructions.

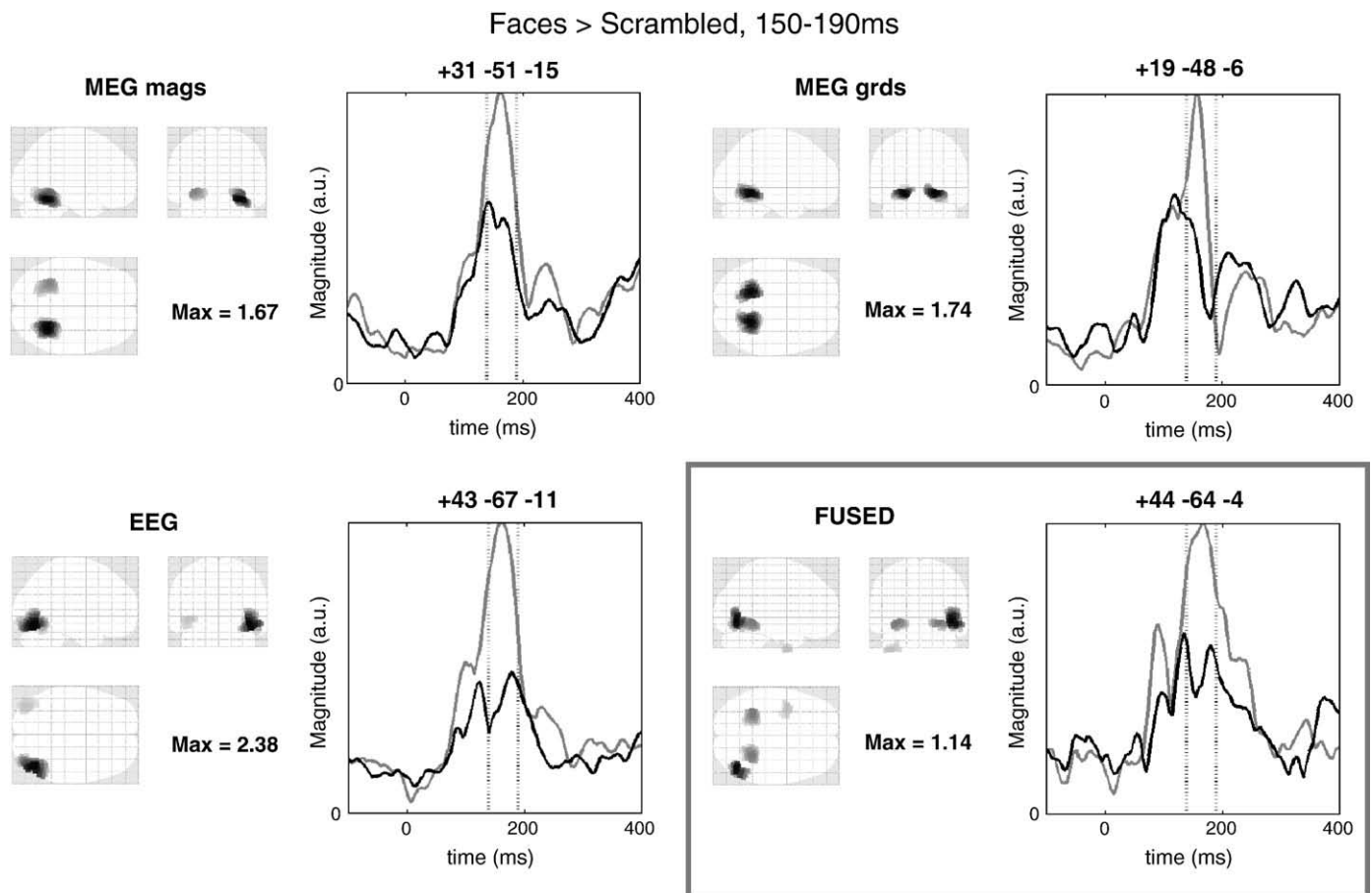
#### Conditional precisions

Finally, to quantify the improvement in precision of the source estimates that is afforded by fusion, we computed the average precision (inverse variance) over the 128 mesh dipoles with the highest responses; as measured by the sum of squared source activity over the whole epoch. The results are shown in the sixth row (for unimodal inversions) and the last row (for fused inversion) of Table 1. These show that fusing all modalities increases the conditional precision relative to that for inverting magnetometers or gradiometers alone. This increase was reliable in the case of gradiometers,  $T(11) = 2.64$ ,  $p < .05$ , though not for magnetometers,  $T(11) = 0.46$ ,  $p = .65$ , or EEG,  $T(11) = -1.61$ ,  $p = .14$ . We revisit this issue in the final analyses below.

#### Spatial and temporal modes

In the analyses above, the number of spatial and temporal modes differed across sensor-types, reflecting not only different properties of the data, but also the different number of sensors. To control for this, and ensure the robustness of the results, the analyses were repeated with a fixed number of  $m_i = 50$  spatial modes and  $r_i = 8$  temporal modes for all sensor-types. The use of 50 spatial modes also meant that, for MEG, their number did not exceed the rank of the data covariance after application of SSS (see *Methods*). The results are shown in Table 2.

The models still explained over 89% of the total data variance (with the decrease in spatial modes having negligible effect, and the increase in temporal modes increasing the fit slightly). The



**Fig. 4.** Mean source reconstructions across participants for the contrast of faces versus scrambled faces within a Gaussian window centred between 150 and 190 ms when inverting data from each sensor-type separately, and when inverting simultaneously (bottom right panel). The left part of each panel shows a Maximal Intensity Projection (MIP) of the 128 greatest source strengths within MNI space; the right part shows the magnitude of the evoked response to faces (light lines) and scrambled faces (dark lines) across the epoch for the maximum in each MIP. For illustration purposes, the source estimates have been smoothed on the 2D mesh surface via 32 iterations of a graph Laplacian with adjacency ratio (c.f., autoregression coefficient) of 1/16.



**Table 2**  
Results of uni- and multimodal inversion, using a fixed number of spatial and temporal modes.

	MEG		EEG
	Mags	Grds	
Number of sensors (range across participants)	102	204	63–70*
Number of spatial modes ( $m_i$ ) (range across participants)	50	50	50
Number of temporal modes ( $r_i$ ) (range across participants)	8	8	8
Empirical (baseline) SNR (plus standard error across participants)	3.8 (0.3)	4.3 (0.3)	4.6 (0.4)
<i>Unimodal inversions</i>			
% Data variance explained (range across participants)	94–98	89–95	93–99
Conditional source precision (plus standard error across participants)	59.4 (8.0)	76.1 (7.0)	95.3 (8.0)
<i>Fused inversion</i>			
% Data variance explained (range across participants)		91–96	
Conditional SNR (plus standard error across participants)	29.5 (2.5)	9.2 (0.9)	6.4 (0.6)
Conditional source precision (plus standard error across participants)		126.9 (15.3)	

Grds = Planar Gradiometers; Mags = Magnetometers. \*Number of EEG sensors varied over participants owing to removal of bad sensors.

conditional SNR estimates were again greatest for the magnetometers and smallest for the EEG data, with pairwise  $T$ -tests significant in all cases,  $T(11) > 2.31$ ,  $p < .05$ . The source reconstructions (not shown) were very similar to those in Fig. 4. However, the average conditional precision of the source estimates was now greatest for the fused inversion, being reliably greater than magnetometers alone,  $T(11) = 6.29$ ,  $p < .001$ , gradiometers alone,  $T(11) = 3.76$ ,  $p < .005$ , and also EEG alone, if one-tailed,  $T(11) = 2.05$ ,  $p < .05$ . These findings further support the value of fused inversion, and suggest that the differences in relative SNR (or conditional estimates) do not owe to different numbers of spatial or temporal modes across sensor-types.

As a final check, the analyses were repeated after splitting the gradiometers into two sets of 102 (approximately equally positioned across the helmet, with both gradient directions at a given location in the same set), i.e. giving four data-types in total, three of which (magnetometers and two sets of gradiometers) were matched in the number of sensors. The basic pattern of results was again replicated, with the hyperparameters for the two sets of gradiometers being, as expected, nearly identical.

## Discussion

We have described an empirical Bayesian approach to simultaneous distributed reconstruction, or “fusion”, of MEG and EEG data. The novelty of this approach rests on optimising modality-specific sensor-level error components by maximising the model-evidence. We evaluated the approach using three types of simultaneously-acquired data: MEG magnetometers, MEG planar gradiometers and EEG. Statistical comparisons across twelve participants of various metrics confirmed the face validity of this approach; particularly in that conditional estimates of source activity based on any one modality alone generally improved with the addition of others. More specifically, while sensor-level error was greatest for EEG, the inclusion of EEG data increased the conditional precision of the underlying source estimates. This supports prior claims that EEG can provide information that supplements the information in MEG (e.g., Fuchs et al., 1998; Baillet et al., 1999; Liu et al., 2002; Bablioni et al., 2001, 2004; Huang et al., 2007; Sharon et al., 2007; Molins et al., 2007). This is expected *a priori* from, for example, the ability of EEG to detect radial components of the electromagnetic field.

“Conventional” empirical estimates of SNR, using sample variance in the pre- and post-stimulus periods, showed similar SNRs for all sensor-types, with a small but reliable superiority for EEG. These estimates are of course specific to the present data (i.e., are not general statements about the sensitivities of different modalities). However, these estimates are not based on a generative model of signal versus sensor noise; indeed, the pre-stimulus baseline period is likely to include “signals” from the brain (e.g., endogenous oscillations), while the post-stimulus period certainly includes sensor noise. Our model-based (conditional) SNR estimates, in which the signal is estimated from the source activity, showed a different pattern, with much higher SNRs for magnetometers relative to gradiometers or EEG. While this SNR estimate depends on the model used (see below), it is likely to be a more relevant estimate for the comparison of different sensor-types.

One reason for the lower conditional SNR estimate for EEG than MEG gradiometers or magnetometers (in contrast to its higher empirical SNR estimate) is its greater sensor-level error variance hyperparameter. As intimated above, this may reflect the relatively poorer forward model for EEG relative to MEG, given that spherical approximations are known to be worse for EEG. A less accurate forward model would compromise the ability of the model to fit the data, increasing the residual error and reducing effective SNR. However, it is important to note that a worse SNR does not necessarily mean that a particular modality will be down-weighted (or regularised more heavily); this is because it may “see” sources that other modalities do not and provide important constraints on the reconstruction. Moreover, the addition of data with lower SNR (such as EEG in the present case), when simultaneously inverting the three sensor-types, still improved the conditional precision of the source estimates (relative to inverting the MEG data alone).

It is interesting to note that the conditional precision for the source estimates using unimodal EEG models was not always increased when fusing all modalities. It was increased when the number of spatial and temporal modes was fixed at 50 and 8 respectively, but not when they were based on modality- (and participant-) specific modes (which typically entailed more spatial modes for MEG, and fewer temporal modes for EEG). This lack of increase in conditional precision may seem counterintuitive, because adding data should increase the conditional precision. One reason for this paradoxical result is likely to be the sparsity constraints on the source reconstruction. This means irrelevant sources can be ‘switched off’ (where a source is a small patch of dipoles on the cortical surface), effectively changing the model. Differential ‘switching-off’ means we could not look at the same dipoles when comparing precisions under unimodal and multimodal inversions; because a dipole with a given precision in one inversion may not exist in another. This may explain why one does not always see an increase in conditional precision simply by adding data. Having said this, there is a clear increase in precision for MEG-based reconstructions when EEG is added; even if the dipoles with the most precise estimates change.

The source reconstructions for the face-related activity around the M170 revealed bilateral ventral temporal activations, close to where one might expect the “fusiform face area”, given prior fMRI studies on the same paradigm (e.g., Henson et al., 2003). The fused reconstruction (Fig. 3) appears to reflect a plausible combination of the individual reconstructions, possibly with separate maxima for FFA and OFA, at least on the right (e.g., Itier et al., 2007). The latter might even extend to a right lateral temporal generator near the superior temporal gyrus, which has previously been hypothesised to contribute to the N170 recorded with EEG (Henson et al., 2003) but to contribute little to the M170 recorded with MEG, given that it is likely to contain a large radial component (Watanabe et al., 2005). Regardless of the physiological plausibility of the reconstructions, it is important for present purposes to note that the multimodal reconstruction contains most of the sources that were seen in the unimodal equivalents. For example, the EEG recovered sources in the posterior and more superficial occipito-temporal regions, but not the anterior and deeper fusiform sources;



and conversely for MEG.<sup>5</sup> The fused reconstruction recovered both. Although one cannot generalise too much, this highlights the potential usefulness of multimodal data in characterising functional anatomy.

In the present framework, we assume that the majority of observed variance is signal, and use the global power to normalise the data to remove scaling differences across modalities. We tested this assumption for the MEG data by comparing the relative sensitivity of magnetometers to gradiometers, as predicted from their lead-field matrices, with the relative power of simulated magnetometer to gradiometer data. The data were simulated by using common cortical sources estimated from a fusion of the two sensor-types, plus sensor error generated randomly from the noise covariance estimated for each sensor-type from a number of empty-room recordings. Across all participants, these two ratios were close to equal, with a mean bias of only 2.5%. These simulations therefore suggest that our normalisation assumptions are realistic, at least for our MEG data. In future work, it may be possible to optimise the scaling of MEG and EEG data, by making their relative scaling a free parameter, again to be optimised through maximising the model-evidence.

There are, of course, other considerations pertinent to the fusion of EEG and MEG data. For example, there are more sophisticated ways to refine forward models for multiple modalities, such as the use of mutual information to weight EEG and MEG lead-fields for each source separately (Baillet et al., 1999), or use of Maxwell's equations to estimate radial and tangential components separately (Huang et al., 2007). It would also be interesting to repeat the analyses for other types of source priors, and for other forward models; such as Boundary Element Models (BEM), which one might expect to be superior (Henson et al., 2009), particularly for EEG (Ermer et al., 2001). Finally, it would be interesting to explore further choices of sensor-level covariance components ( $Q_j^{(i)}$ ). For most of the present results, we assumed independent and identically-distributed sensor noise (though see [Supplementary material](#) for an example of using noise covariance matrices obtained from empty-room recordings). All these issues may be fruitfully revisited within the context of the framework that we introduced above. In any case, the present results suggest that fusion can improve source solutions to the extent that future work on these multimodal generative models is likely to be worthwhile.

## Acknowledgments

We thank Samu Taulu for advice about Signal-Space Separation, Dan Wakeman for help with the cortical meshes, and Nelson Trujillo-Barreto and two reviewers for their helpful comments. This work is funded by the UK Medical Research Council (WBSE U.1055.05.012.00001.01) and the Wellcome Trust.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2009.04.063](https://doi.org/10.1016/j.neuroimage.2009.04.063).

## References

- Ashburner, J., Friston, K.J., 2005. Unified segmentation. *NeuroImage* 26, 839–851.
- Babloni, F., Carducci, F., Cincotti, F., Del Gratta, C., Pizzella, V., Romani, G.L., Rossini, P.M., Tecchio, F., Babloni, C., 2001. Linear inverse source estimate of combined EEG and MEG related to voluntary movements. *Hum. Brain Mapp.* 14, 197–209.

- Babloni, F., Babloni, C., Carducci, F., Romani, G.L., Rossini, P.M., Angelone, L.M., Cincotti, F., 2004. Multimodal integration of EEG and MEG data: a simulation study with variable signal-to-noise ratio and number of sensors. *Hum. Brain Mapp.* 22, 52–62.
- Baillet, S., Garnero, L., Marin, G., Hugonin, J.-P., 1999. Combined MEG and EEG source imaging by minimization of mutual information. *IEEE Trans. Biomed. Eng.* 46, 522–534.
- Berg, P., Scherg, M., 1994. A fast method for forward computation of multiple-shell spherical head models. *Electroencephalogr. Clin. Neurophysiol.* 90, 58–64.
- 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.
- Ermer, J.J., Mosher, J.C., Baillet, S., Leahy, T.M., 2001. Rapidly recomputable EEG forward models for realistic head shapes. *Phys. Med. Biol.* 46, 1265–1281.
- Fischl, B., Liu, A., Dale, A.M., 2001. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans. Med. Imaging.* 20, 70–80.
- Fokas, A.S., 2008. Electro-magneto-encephalography for a three-shell model: distributed current in arbitrary, spherical and ellipsoidal geometries. *J. R. Soc. Int.*
- Friston, K.J., Penny, W., Phillips, C., Kiebel, S., Hinton, G., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: theory. *Neuroimage* 16, 465–483.
- 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., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J., 2008. Multiple sparse priors for the M/EEG inverse problem. *Neuroimage* 39, 1104–1120.
- Fuchs, M., Wagner, M., Wischmann, H.-A., Kohler, T., Theissen, A., Drenckhan, R., Buchner, H., 1998. Improving source reconstructions by combining bioelectric and biomagnetic data. *Electroencephalogr. Clin. Neurophysiol.* 107, 93–111.
- Hämäläinen, M.S., Ilmoniemi, R.J., Sarvas, J., 1988. Interdependence of information conveyed by the magnetoencephalogram and the electroencephalogram. In: Haario (Ed.), *Theory and Applications of Inverse Problems*. Longman, Essex, England, pp. 27–37.
- 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., Mattout, J., Phillips, C., Friston, K.J., 2009. Selecting forward models for MEG source-reconstruction using model-evidence. *Neuroimage*. doi:10.1016/j.neuroimage.2009.01.062; <http://dx.doi.org/10.1016/j.neuroimage.2009.01.062>.
- 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., Barnes, G., 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., Song, T., Hagler Jr, D.J., Podgorny, I., Jousmaki, V., Cui, L., Gaa, K., Harrington, D.L., Dale, A.M., Lee, R.R., Elman, J., Halgren, E., 2007. A novel integrated MEG and EEG analysis method for dipolar sources. *Neuroimage* 37, 731–748.
- Itier, R.J., Herdman, A.T., George, N., Cheyne, D., Taylor, M., 2007. Inversion and contrast-reversal effects on face processing assessed by MEG. *Brain Res.* 1115, 108–120.
- 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.
- Liu, A., Dale, A.M., Belliveau, J.W., 2002. Monte Carlo simulation studies of EEG and MEG localization accuracy. *Hum. Brain Mapp.* 16, 47–62.
- Molins, A., Stufflebeam, S.M., Brown, E.N., Hamalainen, M.S., 2008. Quantification of the benefit from integrating MEG and EEG data in minimum L2-norm estimation. *Neuroimage* 42, 1069–1077.
- Mattout, J., Phillips, C., Penny, W.D., Rugg, M.D., Friston, K.J., 2006. MEG source localization under multiple constraints: an extended Bayesian framework. *Neuroimage* 30, 753–767.
- Mattout, J., Henson, R.N., Friston, K.J., 2007. Canonical source reconstruction for MEG. *Comput. Intell. Neurosci.* 2007.
- 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.
- Sharon, D., Hamalainen, M.S., Tootell, R.B.H., Halgren, E., Belliveau, J.W., 2007. The advantage of combining MEG and EEG: comparison to fMRI in focally stimulated visual cortex. *Neuroimage* 36, 1225–1235.
- Taulu, S., Simola, J., Kajola, M., 2005. Applications of the signal space separation method. *IEEE Trans. Signal. Process.* 53, 3359–3372.
- Watanabe, S., Miki, K., Kakigi, R., 2005. Mechanisms of face perception in humans: a magneto- and electro-encephalographic study. *Neuropathology* 25, 8–20.

<sup>5</sup> One might expect deeper sources to be recovered better by EEG than by MEG, given that the sensitivity of the latter decreases faster with depth, at least for gradiometers. The present results suggest instead that there was activity in posterior ventral occipito-temporal cortex, which was seen better by EEG than MEG; perhaps owing to it having an appreciable radial component.