M/EEG source analysis

Jérémie Mattout
Lyon Neuroscience Research Center

(with many thanks to Christophe Phillips, Rik Henson, Gareth Barnes, Guillaume Flandin, Jean Daunizeau, Stefan Kiebel, Vladimir Litvak and Karl Friston)
“Will it ever happen that mathematicians will know enough about the physiology of the brain, and neurophysiologists enough of mathematical discovery, for efficient cooperation to be possible”

Jacques Hadamard (French mathematician, 1865-1963)

- **ill-posed inverse problem**: no unique solution

- **usefulness of the Bayesian framework**:  
  - Explicit use of prior knowledge  
  - Principled inference on both model parameters and model themselves
Outline

1. The EEG/MEG forward model(s)
2. A variational Bayes \textit{dipolar} approach
3. An empirical Bayes \textit{imaging} approach
4. Multi-subject and Multi-modal integration
The EEG/MEG forward model(s) : physics

\[ Y = g(\vec{j}, \vec{r}) \]

Measures

\[ Y = V \quad \text{(EEG)} \]
\[ Y = \vec{B} \quad \text{(MEG)} \]

Quasi-static Maxwell’s Equations:

\[ \nabla \cdot \vec{j} = 0 \]

Current density

\[ \vec{j} \]
Orientation & amplitude

\[ \vec{r} \]
Location

Electrical potential

\[ \vec{E} = -\nabla V \]

Ohm’s law

\[ \nabla \cdot \vec{j} = \sigma \vec{E} \]

Kirkoff’s law

\[ \vec{E} = -\nabla V \]
The EEG/MEG forward model(s): *physics*

**Current density**
- \( \vec{j} \): Orientation & amplitude
- \( \vec{r} \): Location

**Measures**
- \( Y = V \) (EEG)
- \( Y = \vec{B} \) (MEG)

\[ Y = g(\vec{j}, \vec{r}) \]

\( g \) depends on:
- The type/location/orientation of sensors
- The conductivity of head tissues
- The geometry of the head

\( g \) can have analytic or numeric form
The EEG/MEG forward model(s): *head models*

**Concentric Spheres:**
- **Pros:** Analytic; Fast to compute
- **Cons:** Head not spherical; Conductivity is not isotropic, neither homogeneous

**Boundary Element Method (BEM):**
- **Pros:** Realistic geometry
  Homogeneous conductivity within boundaries
- **Cons:** Numeric; Slow
  Approximation Errors
The EEG/MEG forward model(s): *surfaces / meshes*

**Realistic head model:**
- Scalp (skin-air boundary)
- Outer Skull (bone-skin boundary)
- Inner Skull (CSF-bone boundary)

**Realistic source space:**
- Cortex (white-grey boundary)
The EEG/MEG forward model(s) : *deriving individual meshes*

**Canonical meshes**

Rather than extract surfaces from individual MRIs, why not warp Template surfaces from an MNI brain based on spatial (inverse) normalisation?
The EEG/MEG forward model(s): *deriving individual meshes*

**Inverse spatial normalization**

- Individual MRI
- Template
- Estimate Spatial Transform
- Deformation
The EEG/MEG forward model(s): *deriving individual meshes*

**Canonical meshes**

Rather than extract surfaces from individual MRIs, why not warp Template surfaces from an MNI brain based on spatial (inverse) normalisation?

Also provides a 1-to-1 mapping across subjects, so source solutions can be written directly to MNI space, and group-inversion applied

*Mattout et al (2007), Comp Int & Neuro*
The EEG/MEG forward model(s) : Bayesian form

\[ p(Y | \theta, m) \quad p(\theta | m) \]

Likelihood Prior

\[ p(\theta | Y, m) \quad p(Y | m) \]

Posterior Evidence

Forward Problem

Inverse Problem

Model

Data

Parameters
The EEG/MEG forward model(s) : dipolar vs. imaging

Likelihood

\[ Y = g(\vec{j}, \vec{r}) \]

Location

For small number of Equivalent Current Dipoles (ECD) anywhere in the brain:

- \( g \) is linear in \( \vec{j} \) but non-linear in \( \vec{r} \)

\[ Y = g(\vec{r}).\vec{j} \]

Orientation & amplitude

For large number of (Distributed) dipoles with fixed orientation and location:

- \( g \) is linear in \( \vec{r} \)

\[ Y = G([\vec{r}_1 \vec{r}_2 ... \vec{r}_N]).J \]

Prior

\( \vec{j} \) Orientation & amplitude

\( \vec{r} \) Location

\( Y \) Data
1. The EEG/MEG forward model(s)

2. A variational Bayes *dipolar* approach

3. An empirical Bayes *imaging* approach

4. Multi-subject and Multi-modal integration
A variational Bayes \textit{dipolar} approach

With a Bayesian framework, explicit priors can be put on the locations and orientations of the sources (e.g., symmetry constraints)

\begin{equation}
Y = g(\vec{r})j + e
\end{equation}

\begin{equation}
p(\vec{r}, j, \lambda_r, \lambda_j, \lambda_e | m) \propto p(Y | \vec{r}, j, \lambda_e, m) p(\lambda_e | m) p(\vec{r} | \lambda_r, m) p(\lambda_r | m) p(j | \lambda_j, m) p(\lambda_j | m)
\end{equation}

Like standard ECD approaches, the solution is obtained by iterating the optimization over location/orientation and is:

1. Left with the question of how many dipoles
2. Sensitive to the initial prior location

\textit{Kiebel et al (2008), Neuroimage}
Maximising the (free-energy approximation to the) model evidence $p(Y \mid m)$ offers a natural answer to such questions.
Outline

1. The EEG/MEG forward model(s)
2. A variational Bayes *dipolar* approach
3. An empirical Bayes *imaging* approach
4. Multi-subject and Multi-modal integration
The distributed or imaging source model

Given \( p \) sources fixed in location (e.g., on a cortical mesh), the forward model turns linear:

\[
\mathbf{Y} = \mathbf{GJ} + \mathbf{E}
\]

\( E \sim N(0, \mathbf{C}_e) \)

\( Y = \text{Data} \quad n \text{ sensors} \)
\( J = \text{Sources} \quad p \text{ sources (} \gg n) \)
\( G = \text{forward op.} \quad n \text{ sensors} \times p \text{ sources} \)
\( E = \text{Error} \quad n \text{ sensors} \ldots \)

\( \ldots \text{drawn from Gaussian covariance } \mathbf{C}_e \)

Since \( p \gg n \), regularization is needed such as in the classical L2-norm approach…
The classical $L_2$ or weighted minimum norm approach

$$Y = GJ + E \quad \text{E} \sim N(0, C_e)$$

$$J = \arg\min \left\{ \|C_e^{-1/2} (Y - GJ)\|^2 + \lambda \|W J\|^2 \right\}$$

$$= (W^T W)^{-1} G^T \left[ G(W^T W)^{-1} G^T + \lambda C_e \right]^{-1} Y$$

“Tikhonov”, weighted minimum norm or least-square solution

- $W = I$ “Minimum Norm”
- $W = DD^T$ “Loreta” ($D$=Laplacian)
- $W = \text{diag}(G^T G)^{-1}$ “Depth-Weighted”
- $W_p = \text{diag}(G_p^T C_y^{-1} G_p)^{-1}$ “Beamformer”

Phillips et al (2002), Neuroimage
Its Parametric Empirical Bayes (PEB) generalization

A 2-level hierarchical linear model:

\[
Y = GJ + E_e \quad E_e \sim N(0, C_e) \\
J = 0 + E_j \quad E_j \sim N(0, C_j)
\]

Likelihood \[ p(Y|J) = N(GJ, C_e) \]

Prior \[ p(J) = N(0, C_j) \]

Posterior \[ p(J|Y) \propto p(Y|J)p(J) \]

Maximum A Posteriori (MAP) estimate

\[
J_{MAP} = C_jG^T[G C_j G^T + C_e]^{-1}Y
\]

When compared to classical weighted minimum norm:

\[
(W^T W)^{-1} G^T [G (W^T W)^{-1} G^T + \lambda C_e]^{-1} \Rightarrow C_j = (W^T W)^{-1}
\]

Phillips et al (2005), Neuroimage; Mattout et al., (2006), Neuroimage
Its Parametric Empirical Bayes (PEB) generalization

Priors are specified in terms of covariance components

$$C = \sum \lambda_i Q^{(i)}$$

$\lambda$ = Hyper-parameters

$$C = \text{Sensor/Source covariance}$$

$$Q = \text{Covariance components}$$

1. Sensor components, $Q_{e}^{(i)}$ (error):

   - "IID" (white noise):
     - Empty-room (MEG):

2. Source components, $Q_{j}^{(i)}$ (priors/regularisation):

   - "IID" (min norm):
     - Multiple Sparse Priors (MSP):

When some $Q$'s are correlated, estimation of hyperparameters $\lambda$ can be difficult (e.g. local maxima), and they can become negative (improper for covariances)

To overcome this, one can:

1) impose positivity on hyperparameters:

$$\alpha_i = \ln(\lambda_i) \iff \lambda_i = \exp(\alpha_i)$$

2) impose weak, shrinkage hyperpriors:

$$p(\alpha) \sim N(\eta, \Omega) \quad \eta = -4 \quad \Omega = aI, a = 16$$

uninformative priors are then “turned-off” (cf. “Automatic Relevance Determination”)

$$\alpha \to -\infty \iff \lambda \to 0$$
When multiple $Q$'s are correlated, estimation of hyperparameters $\lambda$ can be difficult (e.g. local maxima), and they can become negative (improper for covariances).

To overcome this, one can:

1) impose positivity on hyperparameters:
\[
\alpha_i = \ln(\lambda_i) \iff \lambda_i = \exp(\alpha_i)
\]

2) impose weak, shrinkage hyperpriors:
\[
p(\alpha) \sim N(\eta, \Omega) \quad \eta = -4 \quad \Omega = aI, a = 16
\]

Useless priors are then “turned-off” (cf. “Automatic Relevance Determination”)
\[
\alpha \to -\infty \iff \lambda \to 0
\]
Full graphical representation

Source and sensor space

Standard Minimum Norm

Fixed
Variable
Data
Empirical Bayes

Full graphical representation

Source and sensor space

- $Q_j^{(1)}$, $Q_j^{(2)}$, ...
- $C_j$, $\lambda_j^{(i)}$
- $J$
- $Q_e^{(1)}$, $Q_e^{(2)}$, ...
- $\lambda_e^{(i)}$, $C_e$
- $E$
- $Y$

- Fixed
- Variable
- Data
Model estimation

1. Obtain Restricted Maximum Likelihood (ReML) estimates of the hyperparameters ($\lambda$) by maximising the variational “free energy” ($F$):

$$\hat{\lambda} = \max_{\lambda} p(Y \mid \lambda) = \max_{\lambda} F$$

2. Obtain Maximum A Posteriori (MAP) estimates of parameters (sources, $J$):

$$\hat{J} = \max_{J} p(J \mid Y, \hat{\lambda}) = \max_{J} F$$

3. Maximal $F$ approximates Bayesian (log) “model evidence” for a model, $m$:

$$\ln p(Y \mid m) = \ln \int \int p(Y, J, \lambda \mid m) dJ d\lambda \approx F(Y, \hat{\alpha}, \hat{\Sigma}) \quad m = \{G, Q, \eta, \Omega\}$$
Multiple Sparse Priors (MSP)

Hyperpriors allow the extreme of 100’s source priors

Multiple priors combined

\[ Q^c = I \]  
\[ Q^c = \{G, I\} \]  
\[ Q^c = \{q_1q_1^T, \ldots, q_Nq_N^T\} \]

Left patch  
Right patch  
Bilateral patches

Multiple Sparse Priors (MSP)

Hyperpriors allow the extreme of 100’s source priors

Summary

The empirical Bayesian approach...

- **Automatically** “regularises” in a principled fashion...
- …allows for multiple constraints (priors)...
- …to the extent that multiple (100’s) of sparse priors possible (MSP)…
- …(or multiple error components or multiple fMRI priors)…
- …furnishes estimates of model evidence, so can compare constraints
Outline

1. The EEG/MEG forward model(s)
2. A variational Bayes *dipolar* approach
3. An empirical Bayes *imaging* approach
4. Multi-subject and Multi-modal integration
Group inversion

Single subject

Source and sensor space

\[
\begin{align*}
Q_j^{(1)} & \rightarrow C_j \\
\lambda_j^{(i)} & \rightarrow J \\
J & \rightarrow \eta, \Omega \\
\eta, \Omega & \rightarrow Q_e^{(1)} \\
Q_e^{(1)} & \rightarrow C_e \\
\lambda_e^{(i)} & \rightarrow E \\
E & \rightarrow Y \\
Y & \rightarrow \text{Data}
\end{align*}
\]
Group inversion

Multiple subjects

\[ Q_j^{(1)} \rightarrow C_j \rightarrow \lambda_j^{(i)} \rightarrow J_1 \]

\[ \eta, \Omega \rightarrow Q_{e,1} \rightarrow \lambda_{e,1} \rightarrow C_{e,1} \]

\[ \eta, \Omega \rightarrow Q_{e,2} \rightarrow \lambda_{e,2} \rightarrow C_{e,2} \]

Source and sensor space

\[ \text{Fixed} \]

\[ \text{Variable} \]

\[ \text{Data} \]
Group inversion

Litvak & Friston (2008) Neuroimage
Multi-modal integration: EEG-MEG fusion

Single modality

$Q_j^{(1)} \cdots$

$C_j$

$\lambda_j^{(i)}$

$J$

Fixed

Variable

Data

Source and sensor space

$\eta, \Omega$

$Q_e^{(1)} \cdots$

$\lambda_e^{(i)}$

$C_e$

$E$

$\mathbf{Y}$

$\mathbf{J}$

$\mathbf{E}$

$\mathbf{Y}$
Multi-modal integration: EEG-MEG fusion

Henson et al. (2009) Neuroimage
Multi-modal integration: EEG-MEG fusion

IID noise for each modality; common MSP for sources

Henson et al (2009) Neuroimage
Multi-modal integration: fMRI priors

Fixed

Variable

Data

\[ Q_j^{(1)} \quad \ldots \quad Q_e^{(1)} \]

\[ C_j \quad \lambda_j^{(i)} \quad \lambda_e^{(i)} \quad C_e \]

\[ J \quad E \quad Y \]

Source and sensor space
Multi-modal integration: fMRI priors

\[ Y_{fMRI} \]

\[ \begin{align*}
    Q_j^{(1)} \rightarrow C_j \\
    \lambda_j^{(i)} \rightarrow J \\
    \end{align*} \]

\[ \begin{align*}
    \eta, \Omega \rightarrow \lambda_e^{(i)} \\
    Q_e^{(1)} \rightarrow C_e \\
    \end{align*} \]

Source and sensor space

\[ \begin{align*}
    Q_j^{(1)}, \ldots, Q_e^{(1)} \rightarrow Y \\
    \end{align*} \]

Fixed

Variable

Data
Multi-modal integration: fMRI priors

1. Thresholding and connected component labelling
2. Projection onto the cortical surface using the Voronoï diagram
3. Prior covariance components $Q_{ij}$

SPM(F) for faces versus scrambled faces, 15 voxels, p<.05 FWE

Multi-modal integration: fMRI priors

5 clusters from SPM of fMRI data from separate group of (18) subjects in MNI space

Multi-modal integration: fMRI priors

IID sources and IID noise (L2 MNM)

Magnetometers (MEG)

Gradiometers (MEG)

Electrodes (EEG)

None          Global          Local (Valid)

fMRI priors counteract superficial bias of L2-norm

Multi-modal integration: fMRI priors

Conclusion

1. SPM offers standard forward models (via FieldTrip)… (though with unique option of Canonical Meshes)

2. …but offers unique Bayesian approaches to inversion:
   
   2.1 Variational Bayesian ECD
   
   2.2 A PEB approach to Distributed inversion (eg MSP)

3. PEB framework in particular offers multi-subject and (various types of) multi-modal integration
Transition

Classical (static) source reconstruction

\[ Y = g(x, \theta) \]

Dynamic causal modelling

\[ \frac{dx}{dt} = f(x, \theta, u) \]

Auditory-Visual Stimulations (\(u\))

EEG/MEG Observations (\(Y\))