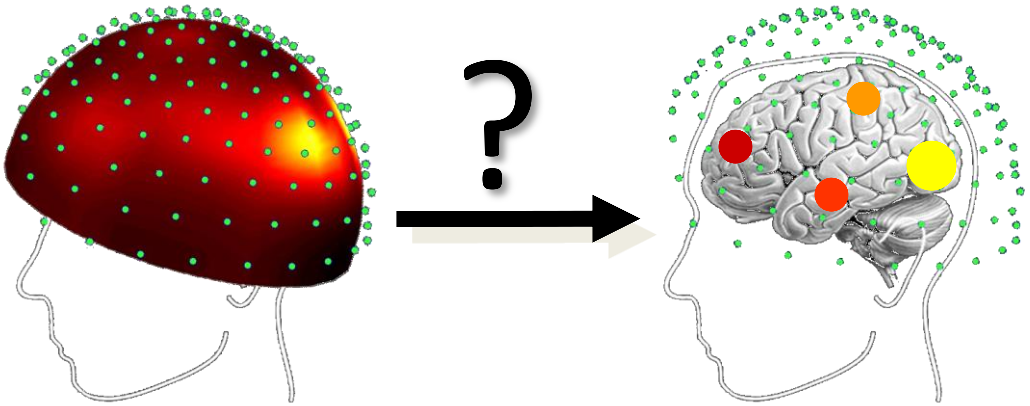
M/EEG source analysis

Jérémie Mattout

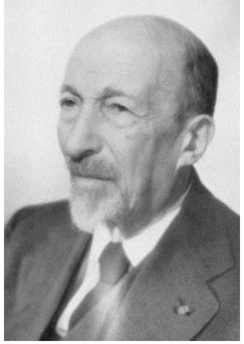
Lyon Neuroscience Research Center

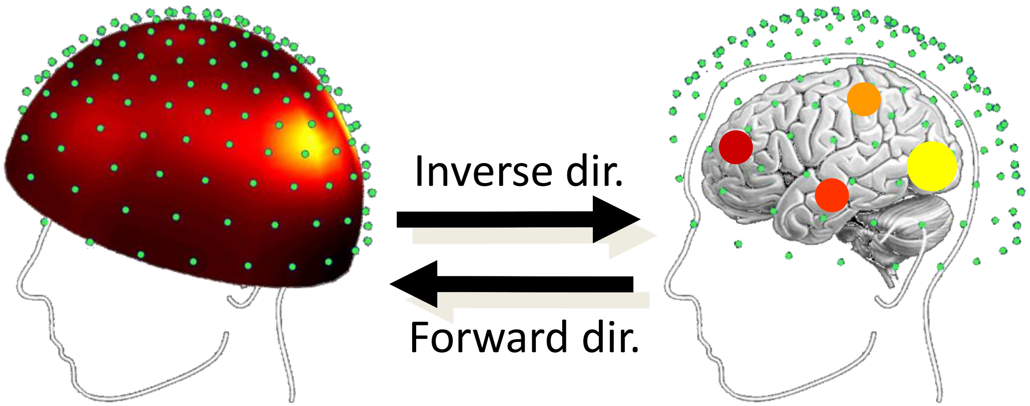


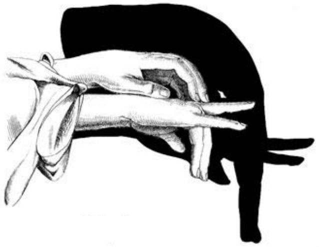
(with many thanks to Christophe Phillips, Rik Henson, Gareth Barnes, Guillaume Flandin, Jean Daunizeau, Stefan Kiebel, Vladimir Litvak and Karl Friston)

"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 *dipolar* approach
3. An empirical Bayes *imaging* approach
4. Multi-subject and Multi-modal integration

The EEG/MEG forward model(s) : *physics*

Current density

$\vec{j}$ Orientation & amplitude

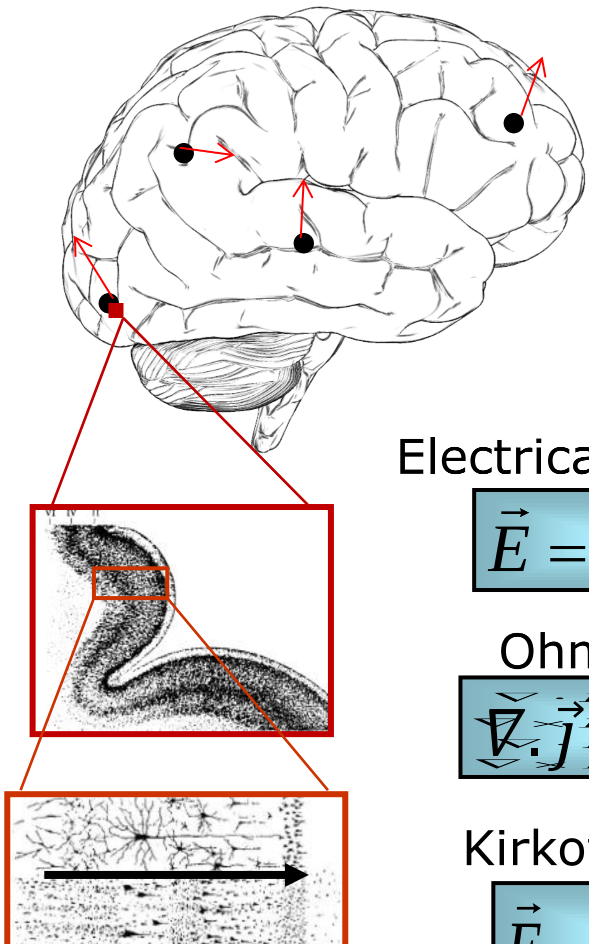
$\vec{r}$ Location

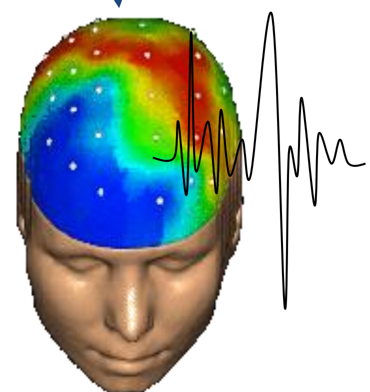
$$Y = g(\vec{j}, \vec{r})$$

Measures

$Y = V$ (EEG)

$Y = \vec{B}$ (MEG)





Electrical potential

$$\vec{E} = -\nabla . V$$

Ohm's law

$$\nabla . \vec{j} = \sigma \vec{E}$$

Kirkoff's law

$$\vec{E} = -\nabla . V$$

Quasi-static Maxwell's Equations:

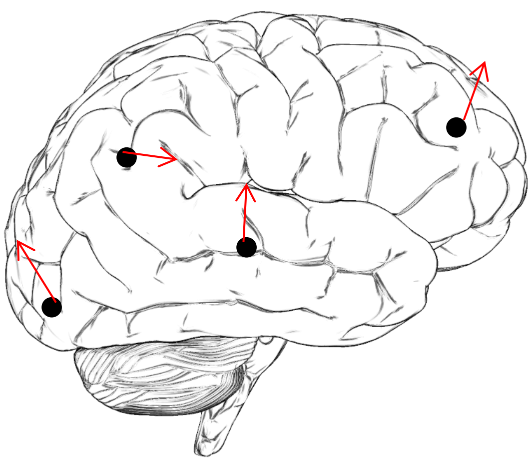
$$\vec{\nabla} \cdot \vec{j} = 0$$

The EEG/MEG forward model(s) : *physics*

Current density

$\vec{j}$ Orientation & amplitude

$\vec{r}$ Location

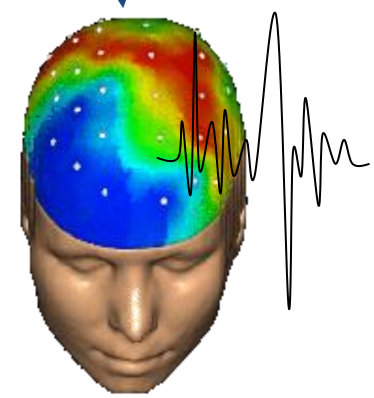


$$Y = g(\vec{j}, \vec{r})$$

Measures

$Y = V$ (EEG)

$Y = \vec{B}$ (MEG)

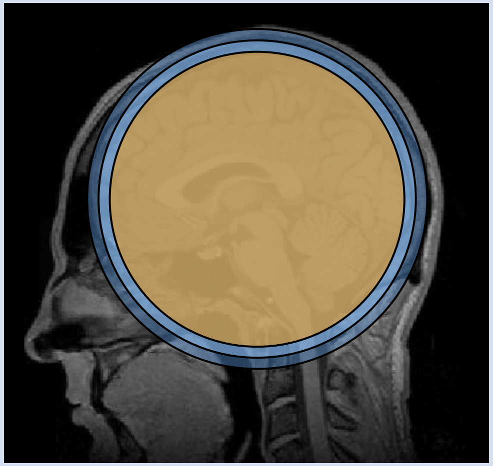


g **depends on:**

- The type/location/orientation of sensors
- The conductivity of head tissues
- The geometry of the head

g **can have analytic of numeric form**

The EEG/MEG forward model(s) : *head models*



Concentric Spheres:

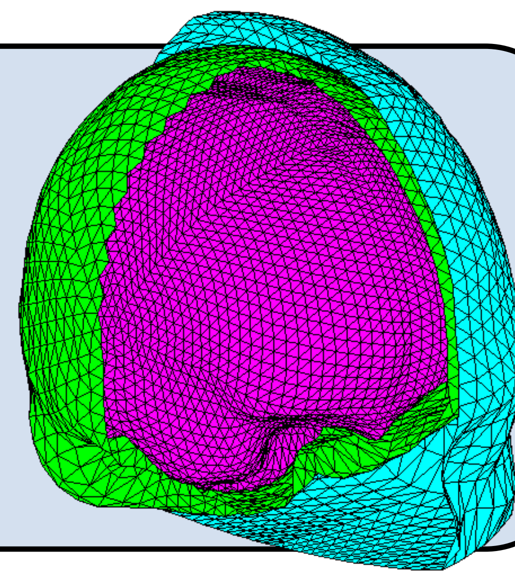
Pros: Analytic; Fast to compute

Cons: Head not spherical;
Conductivity is not isotropic,
neither homogeneous

Boundary Element Method (BEM) :

Pros: Realistic geometry
Homogeneous conductivity
within boundaries

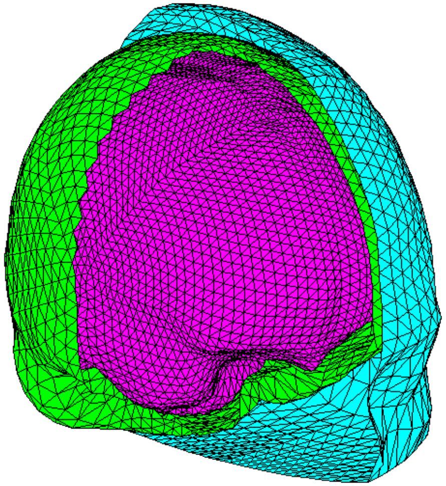
Cons: Numeric; Slow
Approximation Errors



The EEG/MEG forward model(s) : *surfaces / meshes*

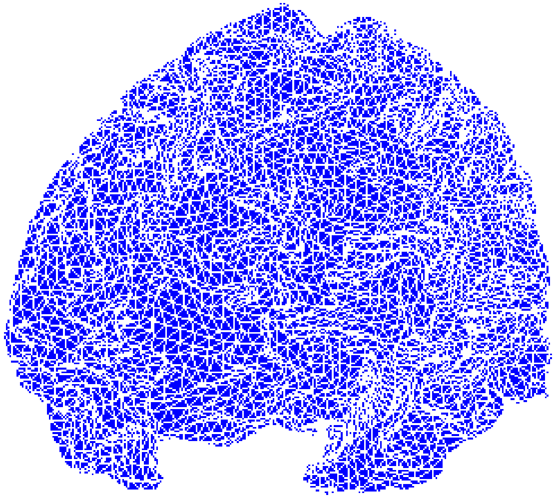
Realistic head model:

Scalp (skin-air boundary)
Outer Skull (bone-skin boundary)
Inner Skull (CSF-bone boundary)



Realistic source space:

Cortex (white-grey boundary)



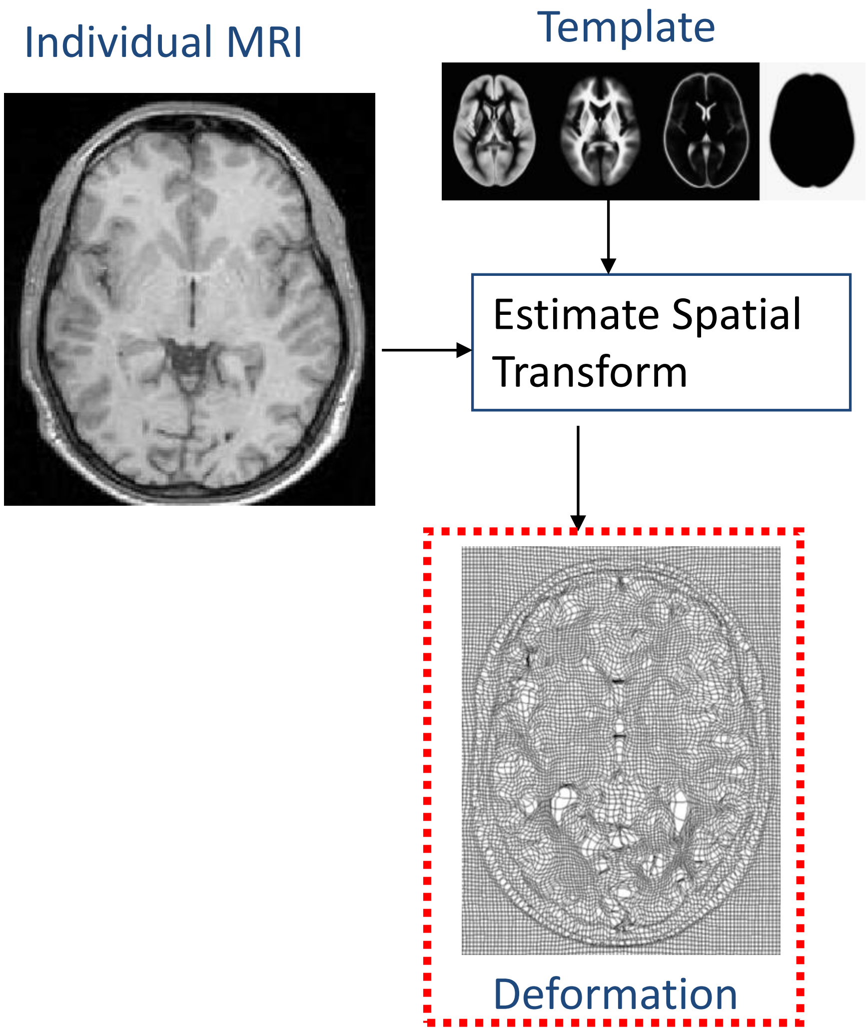
The EEG/MEG forward model(s) : *deriving individual meshes*

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*

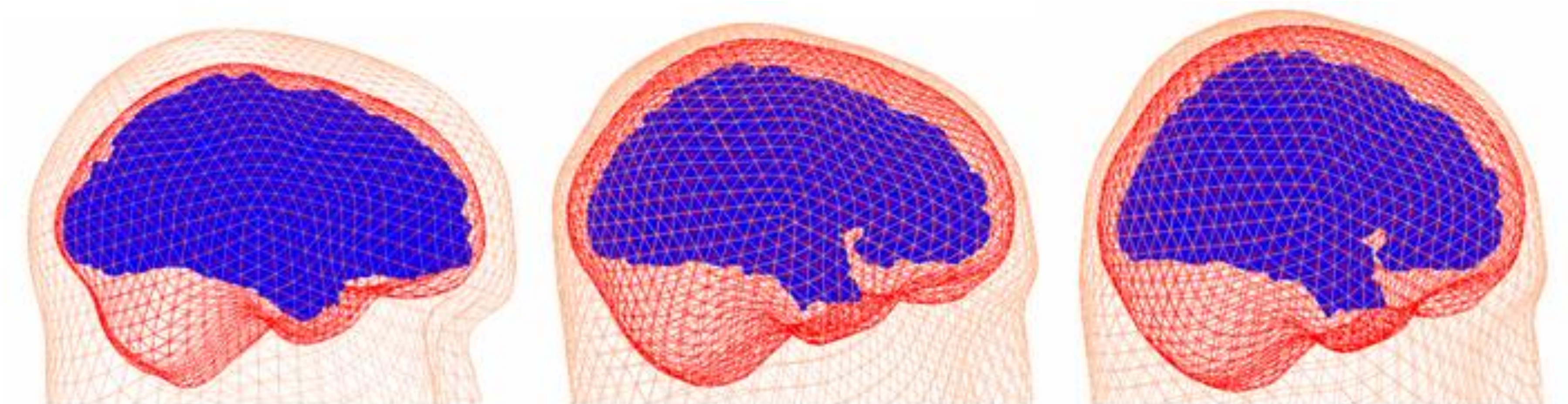
Inverse spatial normalization



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?



Individual Canonical (Inverse-Normalised) Template

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*

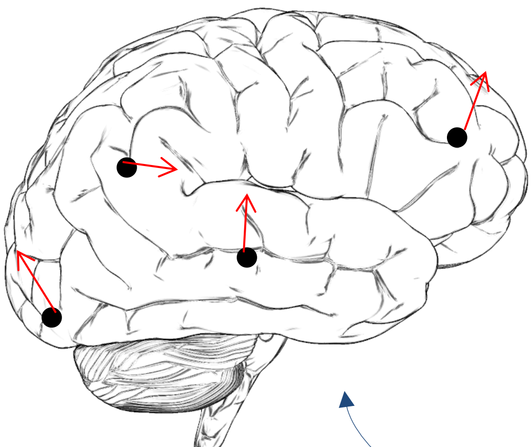
Forward Problem

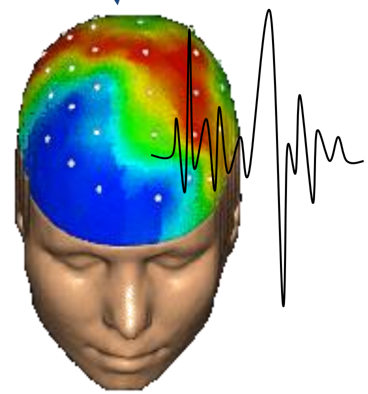
m Model

$$p(Y \mid \theta, m) \qquad p(\theta \mid m)$$

Likelihood   Prior

Y Data





Posterior   Evidence

$$p(\theta \mid Y, m) \qquad p(Y \mid m)$$

θ Parameters

Inverse Problem

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

(ECD) **(distributed)**

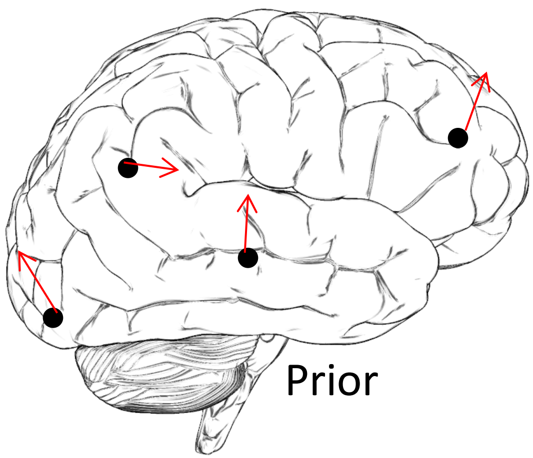
$\vec{j}$ Orientation & amplitude

$\vec{r}$ Location

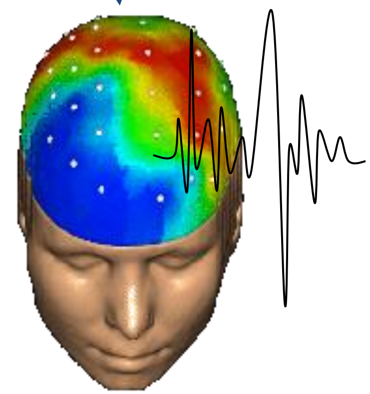
Likelihood

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

Y Data



Prior



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

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

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

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

g is linear in $\vec{r}$

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

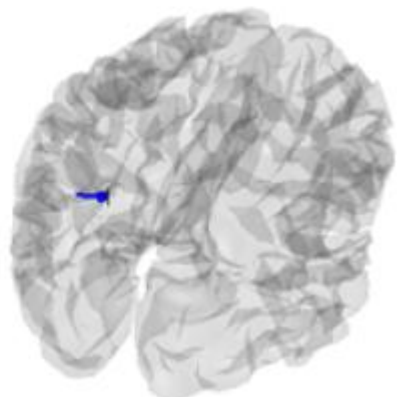
Outline

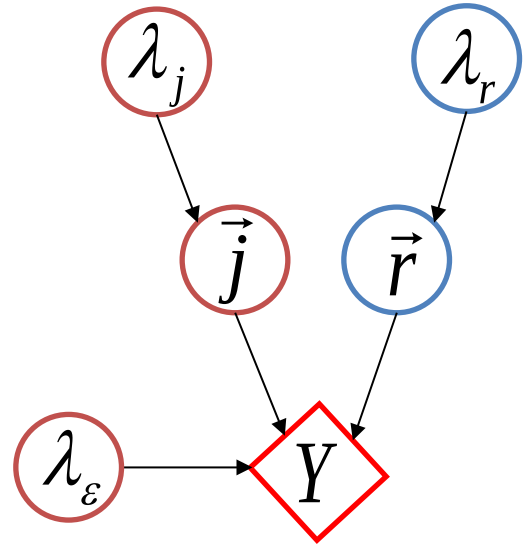
1. The EEG/MEG forward model(s)
2. **A variational Bayes *dipolar* approach**
3. An empirical Bayes *imaging* approach
4. Multi-subject and Multi-modal integration

A variational Bayes *dipolar* approach

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

$$Y = g(\vec{r})\vec{j} + e$$





$$p(\vec{r},\vec{j},\lambda_r,\lambda_j,\lambda_e \mid m) \propto p(Y \mid \vec{r},\vec{j},\lambda_e,m)\,p(\lambda_e \mid m)\,p(\vec{r} \mid \lambda_r,m)\,p(\lambda_r \mid m)\,p(\vec{j} \mid \lambda_j,m)\,p(\lambda_j \mid m)$$

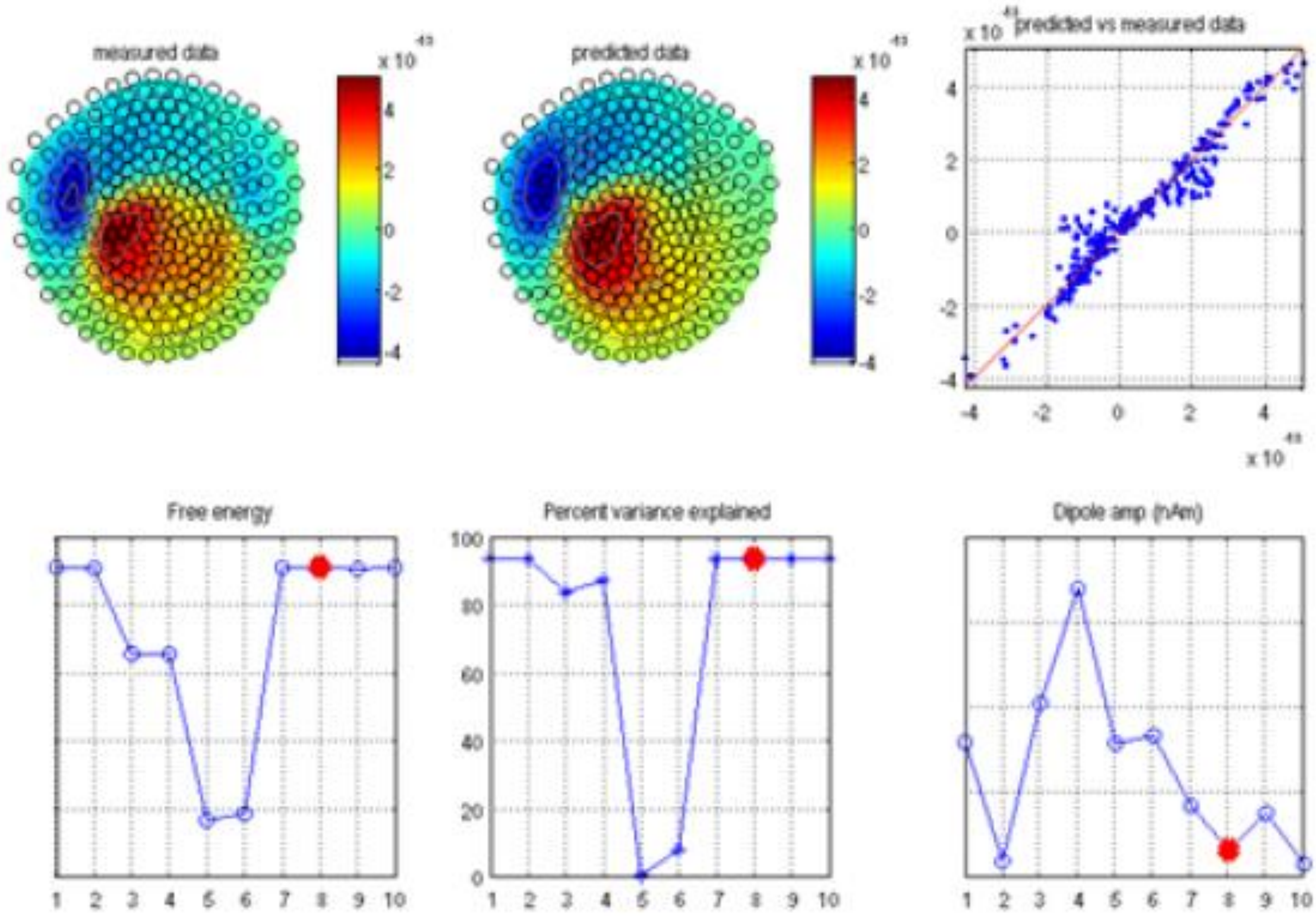
Like standard ECD approaches, the solution is obtained by iterating the optimization over location/orientation and is:

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

Kiebel et al (2008), Neuroimage

A variational Bayes *dipolar* approach

Maximising the (free-energy approximation to the) model evidence $p(Y \mid m)$ offers a natural answer to such questions



Kiebel et al (2008), Neuroimage

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:

$$\mathbf{Y} = \mathbf{GJ} + \mathbf{E}$$

$$\mathbf{E} \sim \mathbf{N}(\mathbf{0}, \mathbf{C_e})$$

Y = Data	*n* sensors
J = Sources	*p* sources (>> n)
G = forward op.	*n* sensors *x* *p* sources
E = Error	*n* sensors…
	…drawn from Gaussian covariance $\mathbf{C_e}$

Since p >> n, regularization is needed such as in the classical L2-norm approach…

The classical L2 or weighted minimum norm approach

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

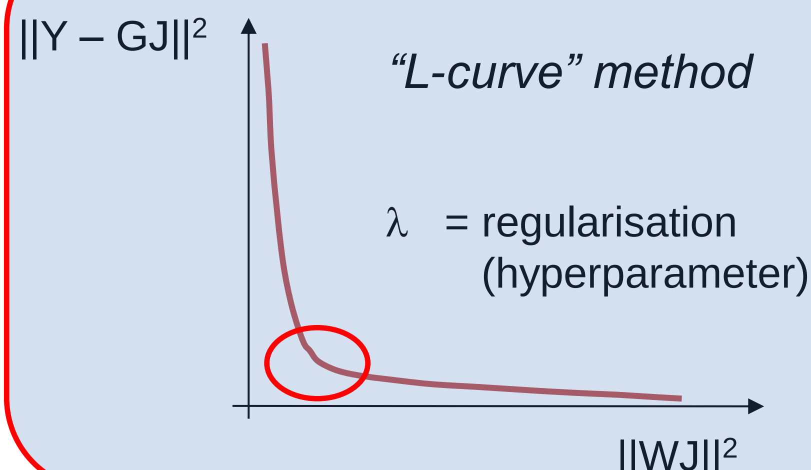
Regularization or Hyperparameter

$$\mathbf{J} = \mathbf{argmin}\left\{\left\|\mathbf{C_e}^{-1/2}.(\mathbf{Y} - \mathbf{GJ})\right\|^2 + \boldsymbol{\lambda}\|\mathbf{WJ}\|^2\right\}$$

Weighting or Constraint matrix

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

"Tikhonov", weighted minimum norm or least-square solution



$\mathbf{W} = \mathbf{I}$ "Minimum Norm"

$\mathbf{W} = \mathbf{DD}^T$ "Loreta" (**D**=Laplacian)

$\mathbf{W} = diag(G^T G)^{-1}$ "Depth-Weighted"

$\mathbf{W}_p = diag(G_p^T C_y^{-1} G_p)^{-1}$ "Beamformer"

$\mathbf{W} = \ldots$

Phillips et al (2002), Neuroimage

Its Parametric Empirical Bayes (PEB) generalization

A 2-level hierarchical linear model:

$$\mathbf{Y} = \mathbf{GJ} + \mathbf{E_e} \qquad \mathbf{E_e} \sim \mathbf{N}(\mathbf{0}, \mathbf{C_e})$$

$$\mathbf{J} = \mathbf{0} + \mathbf{E_j} \qquad \mathbf{E_j} \sim \mathbf{N}(\mathbf{0}, \mathbf{C_j})$$

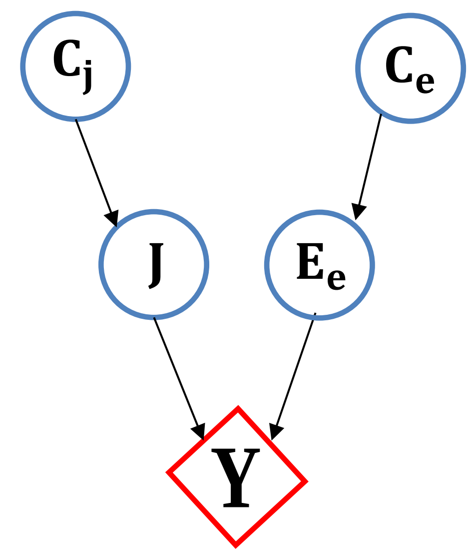
$\mathbf{C_e}$ = *n x n* Sensor (error) covariance

$\mathbf{C_j}$ = *p x p* Source (prior) covariance

Likelihood $\mathbf{p}(\mathbf{Y}|\mathbf{J}) = \mathbf{N}(\mathbf{GJ}, \mathbf{C_e})$

Prior $\mathbf{p}(\mathbf{J}) = \mathbf{N}(\mathbf{0}, \mathbf{C_j})$

Posterior $\mathbf{p}(\mathbf{J}|\mathbf{Y}) \propto \mathbf{p}(\mathbf{Y}|\mathbf{J})\mathbf{p}(\mathbf{J})$



Maximum A Posteriori (MAP) estimate

$$\mathbf{J_{MAP}} = \mathbf{C_j G^T}\left[\mathbf{G C_j G^T} + \mathbf{C_e}\right]^{-1}\mathbf{Y}$$

When compared to classical weighted minimum norm:

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

Phillips et al (2005), Neuroimage; Mattout et al., (2006), Neuroimage

Its Parametric Empirical Bayes (PEB) generalization

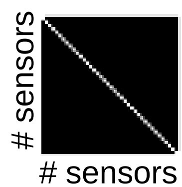
Priors are specified in terms of covariance components

$$\mathbf{C} = \sum_{\mathbf{i}} \boldsymbol{\lambda}_{\mathbf{i}} \mathbf{Q}^{(\mathbf{i})}$$

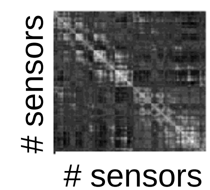
C = Sensor/Source covariance
Q = Covariance components
λ = Hyper-parameters

1. Sensor components, $\mathbf{Q_e}^{(\mathbf{i})}$ (error):

"IID" (white noise):

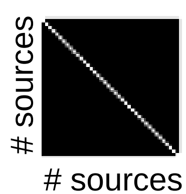


Empty-room (MEG):

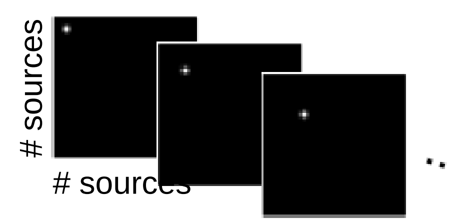


2. Source components, $\mathbf{Q_j}^{(\mathbf{i})}$ (priors/regularisation):

"IID" (min norm):



Multiple Sparse Priors (MSP):



Friston et al (2008) Neuroimage

Hyperpriors

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

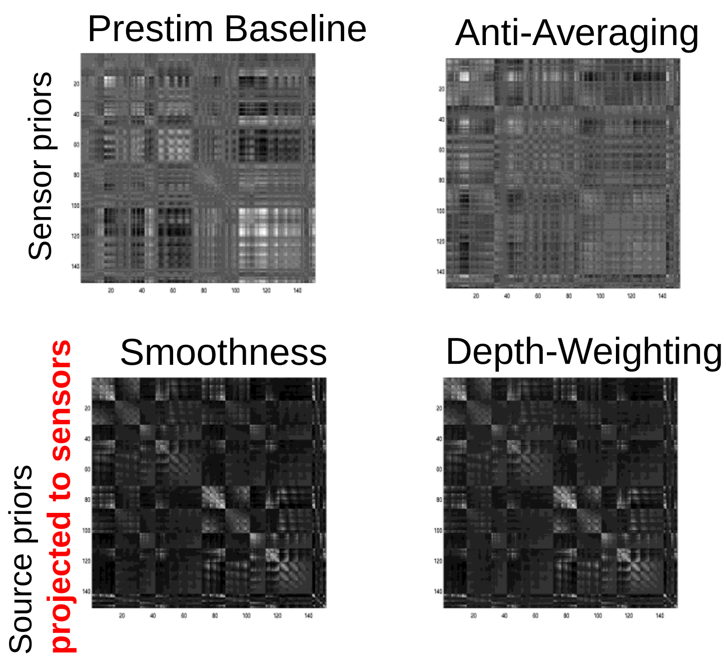
To overcome this, one can:

1) impose positivity on hyperparameters:

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

2) impose weak, shrinkage hyperpriors:

$$p(\boldsymbol{\alpha}) \sim N(\boldsymbol{\eta}, \boldsymbol{\Omega}) \qquad \boldsymbol{\eta} = -4 \qquad \boldsymbol{\Omega} = a\mathbf{I}, a = 16$$



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

$$\alpha \rightarrow -\infty \Leftrightarrow \lambda \rightarrow 0$$

Hyperpriors

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

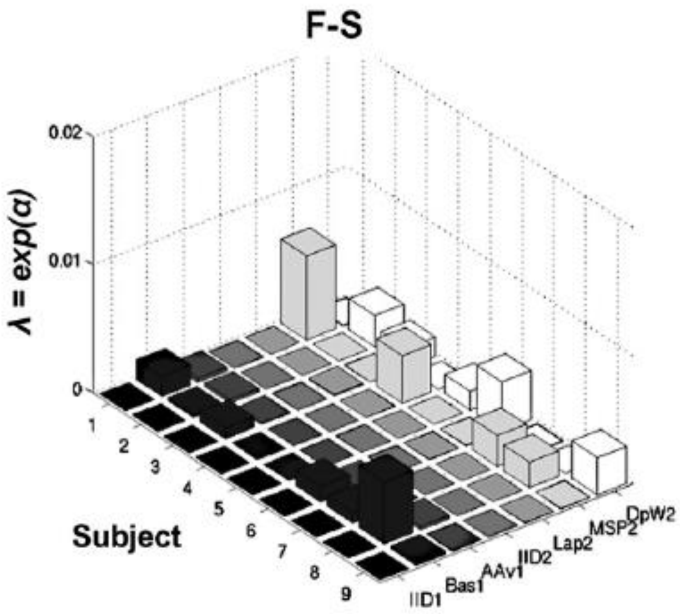
To overcome this, one can:

1) impose positivity on hyperparameters:

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

2) impose weak, shrinkage hyperpriors:

$$p(\boldsymbol{\alpha}) \sim N(\boldsymbol{\eta}, \boldsymbol{\Omega}) \qquad \boldsymbol{\eta} = -4 \qquad \boldsymbol{\Omega} = a\mathbf{I}, a = 16$$



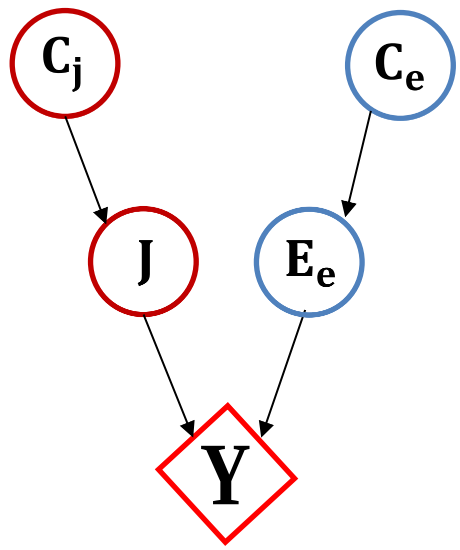
Useless priors are then "turned-off" (cf. "Automatic Relevance Determination")

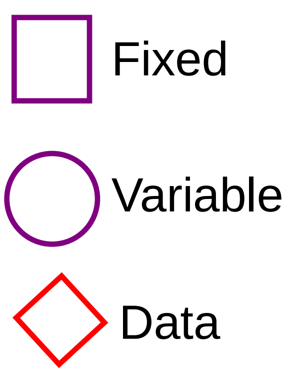
$$\alpha \to -\infty \Leftrightarrow \lambda \to 0$$

Full graphical representation

Source and sensor space

Standard Minimum Norm

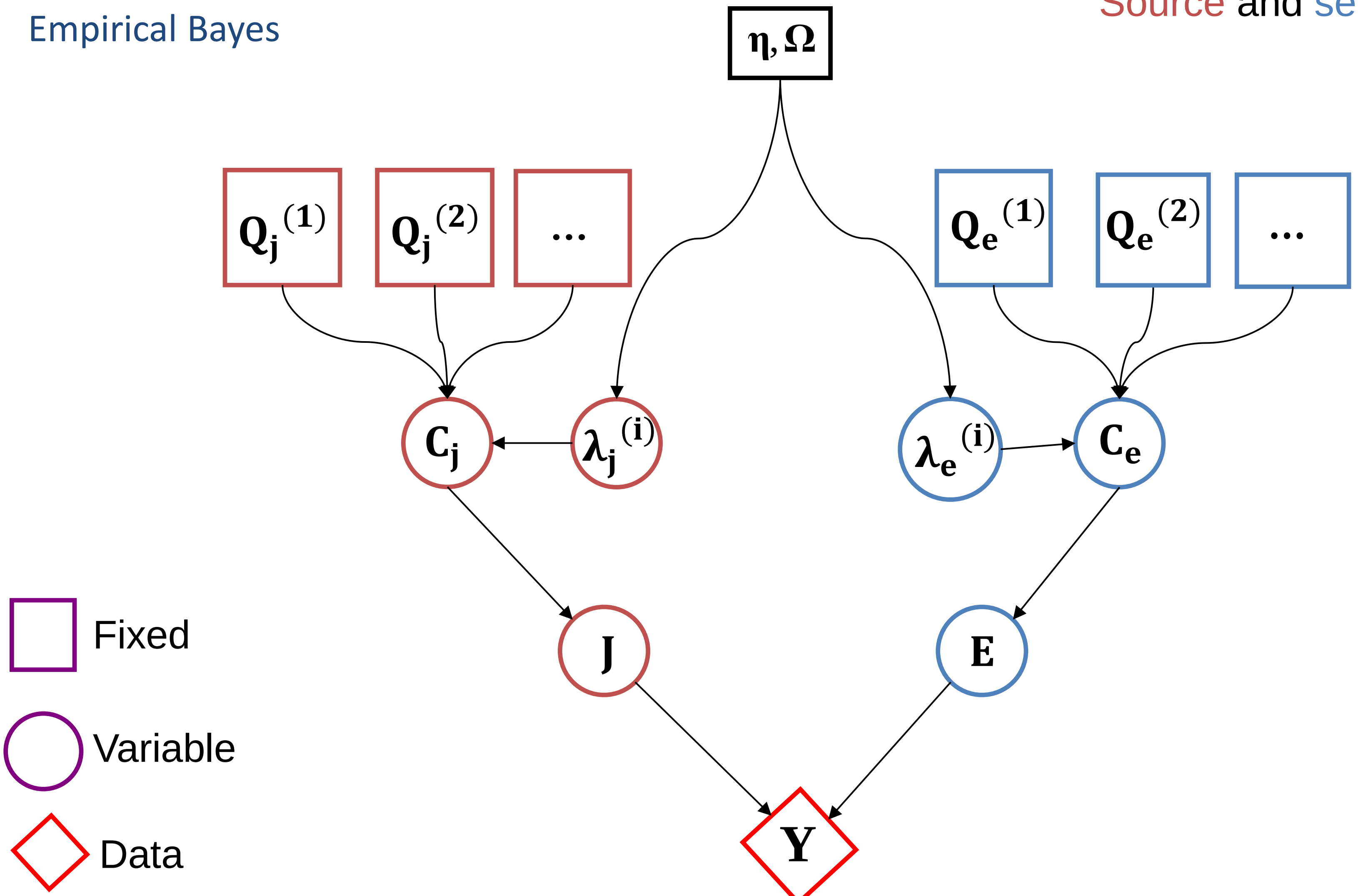




Full graphical representation

Empirical Bayes

Source and sensor space



Model estimation

1. Obtain Restricted Maximum Likelihood (ReML) estimates of the hyperparameters (λ) by maximising the variational "free energy" (F):

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

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

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

3. Maximal F approximates Bayesian (log) "model evidence" for a model, *m*:

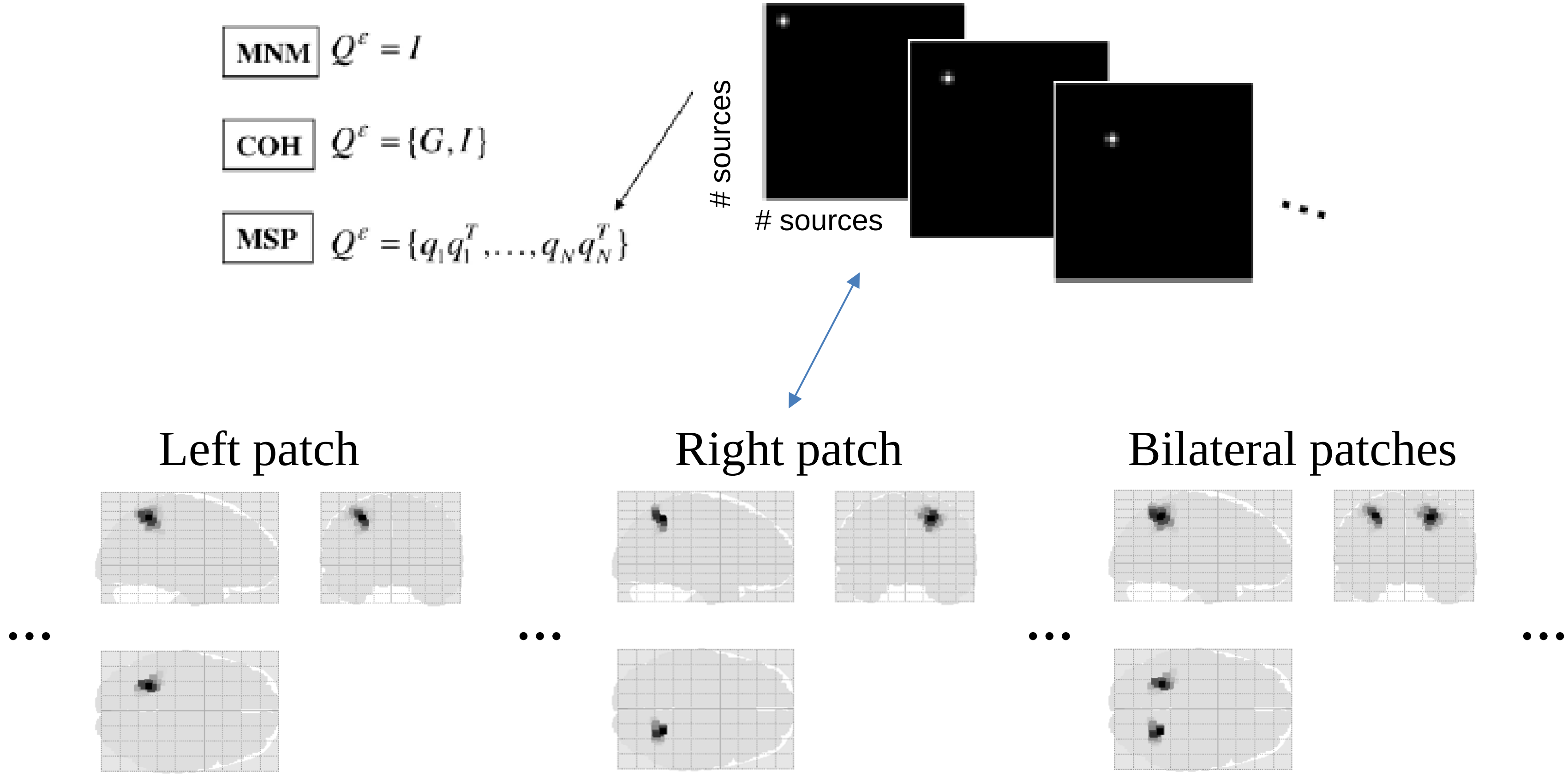
$$\ln p(\mathbf{Y} \mid m) = \ln \int\int p(\mathbf{Y}, \mathbf{J}, \boldsymbol{\lambda} \mid m) d\mathbf{J} d\boldsymbol{\lambda} \approx F(\mathbf{Y}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\Sigma}})$$

$$\boldsymbol{m} = \{\boldsymbol{G}, \boldsymbol{Q}, \boldsymbol{\eta}, \boldsymbol{\Omega}\}$$

Multiple Sparse Priors (MSP)

Hyperpriors allow the extreme of 100’s source priors

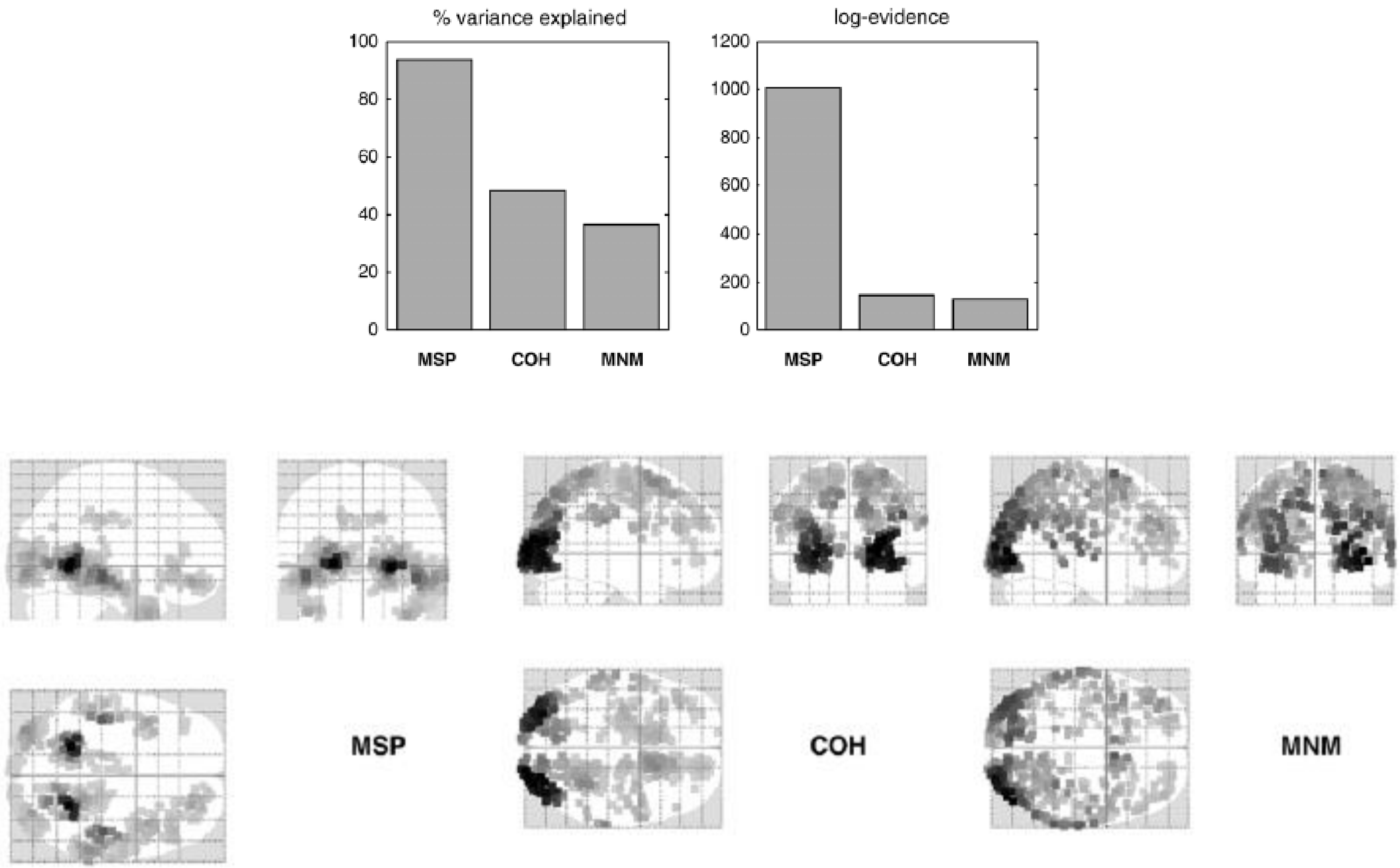
Multiple priors combined



Friston et al (2008) Neuroimage

Multiple Sparse Priors (MSP)

Hyperpriors allow the extreme of 100's source priors



Friston et al (2008) Neuroimage

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

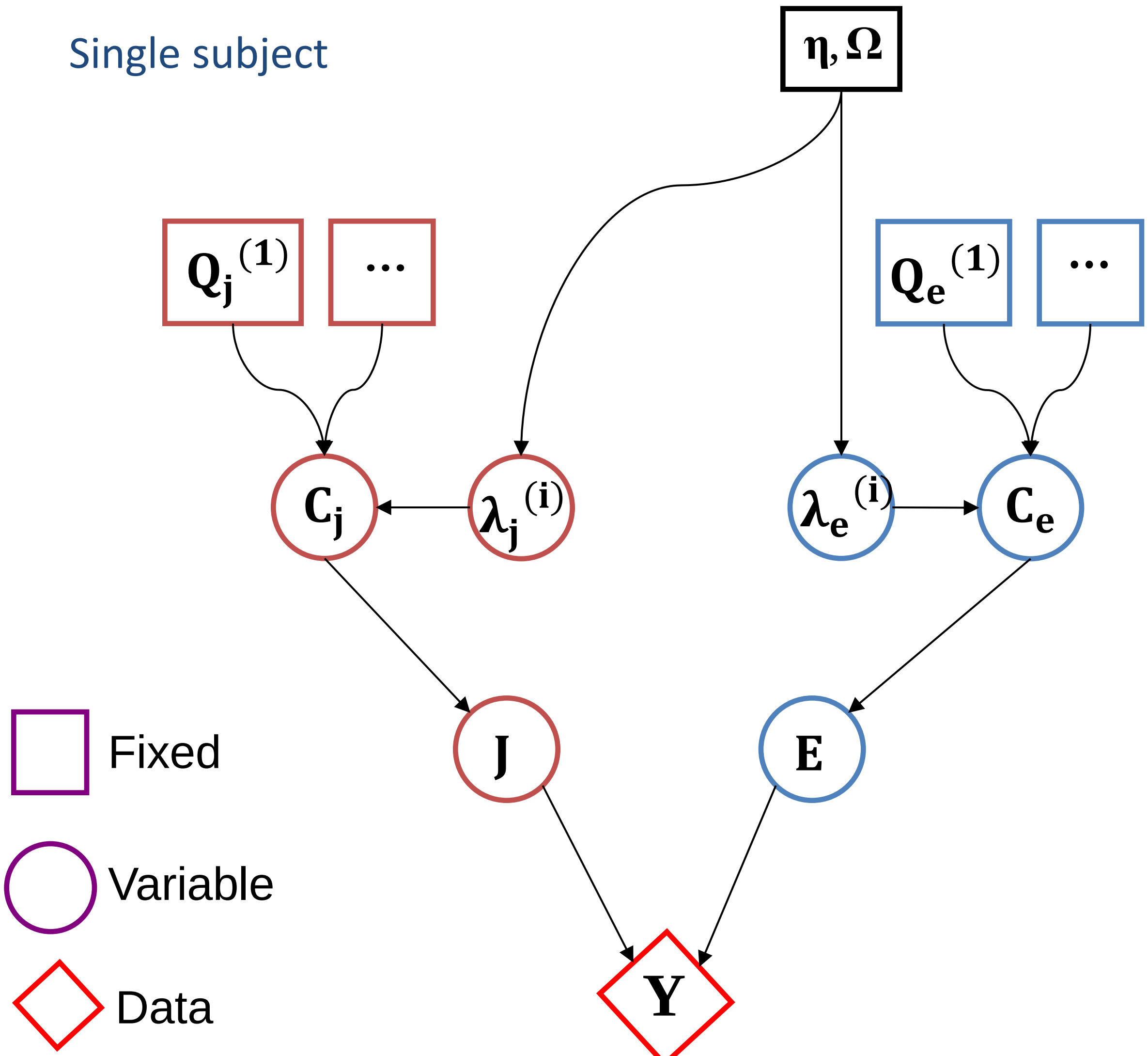
3. An empirical Bayes *imaging* approach

4. **Multi-subject and Multi-modal integration**

Group inversion

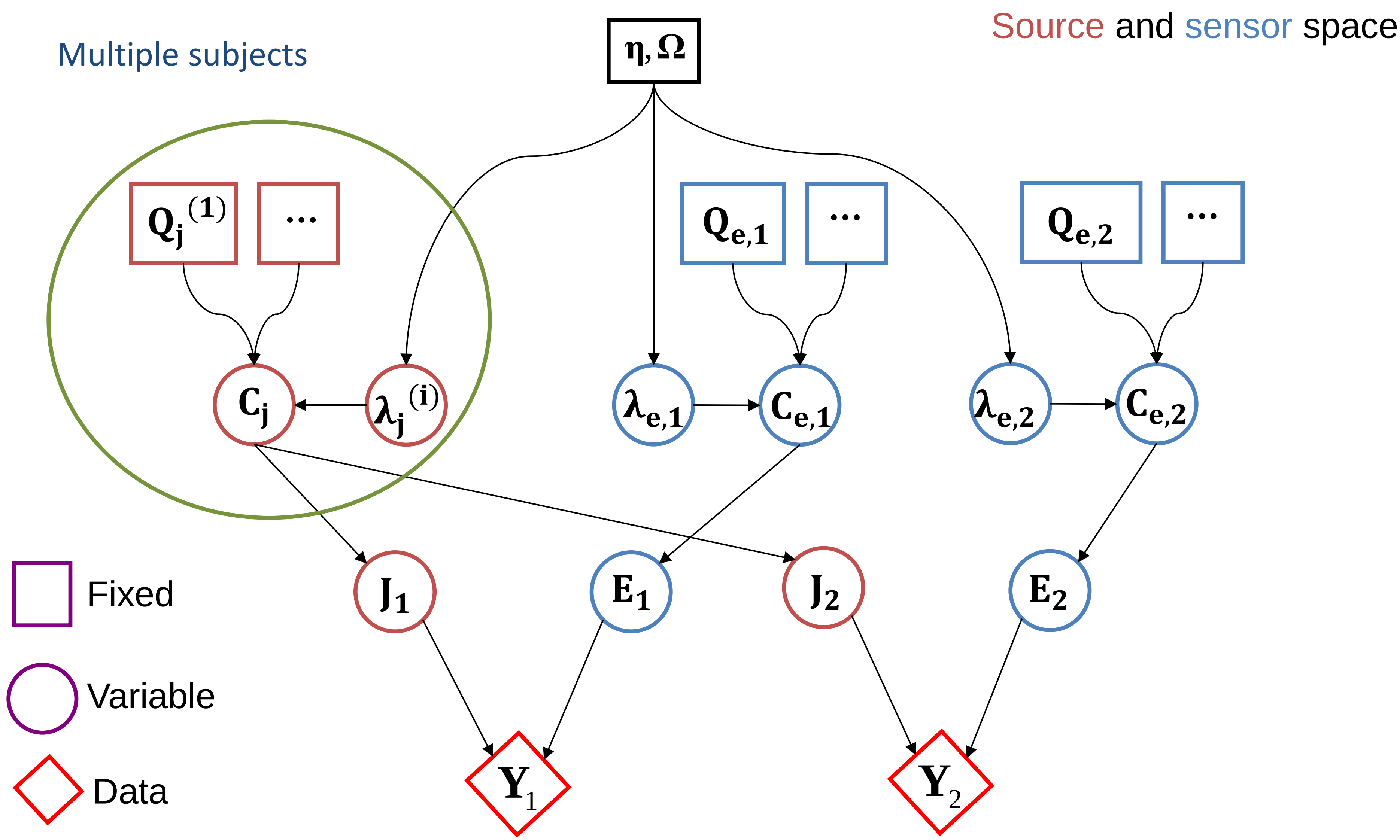
Single subject

Source and sensor space



Friston et al. (2008) Neuroimage

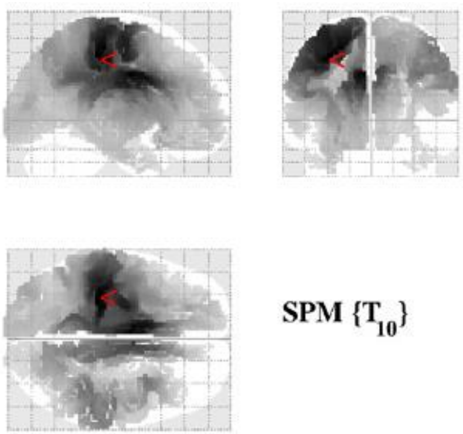
Group inversion



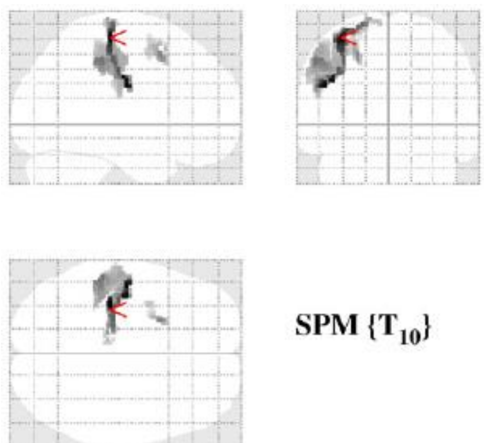
Litvak & Friston (2008) Neuroimage

Group inversion

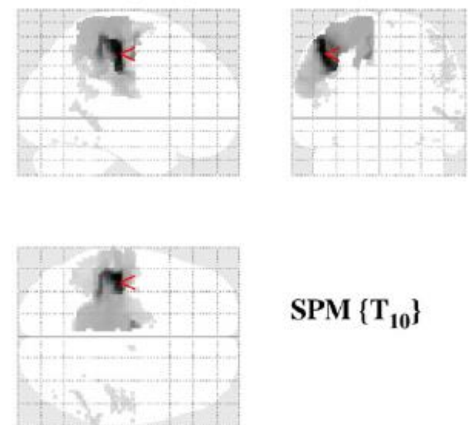
MMN



MSP



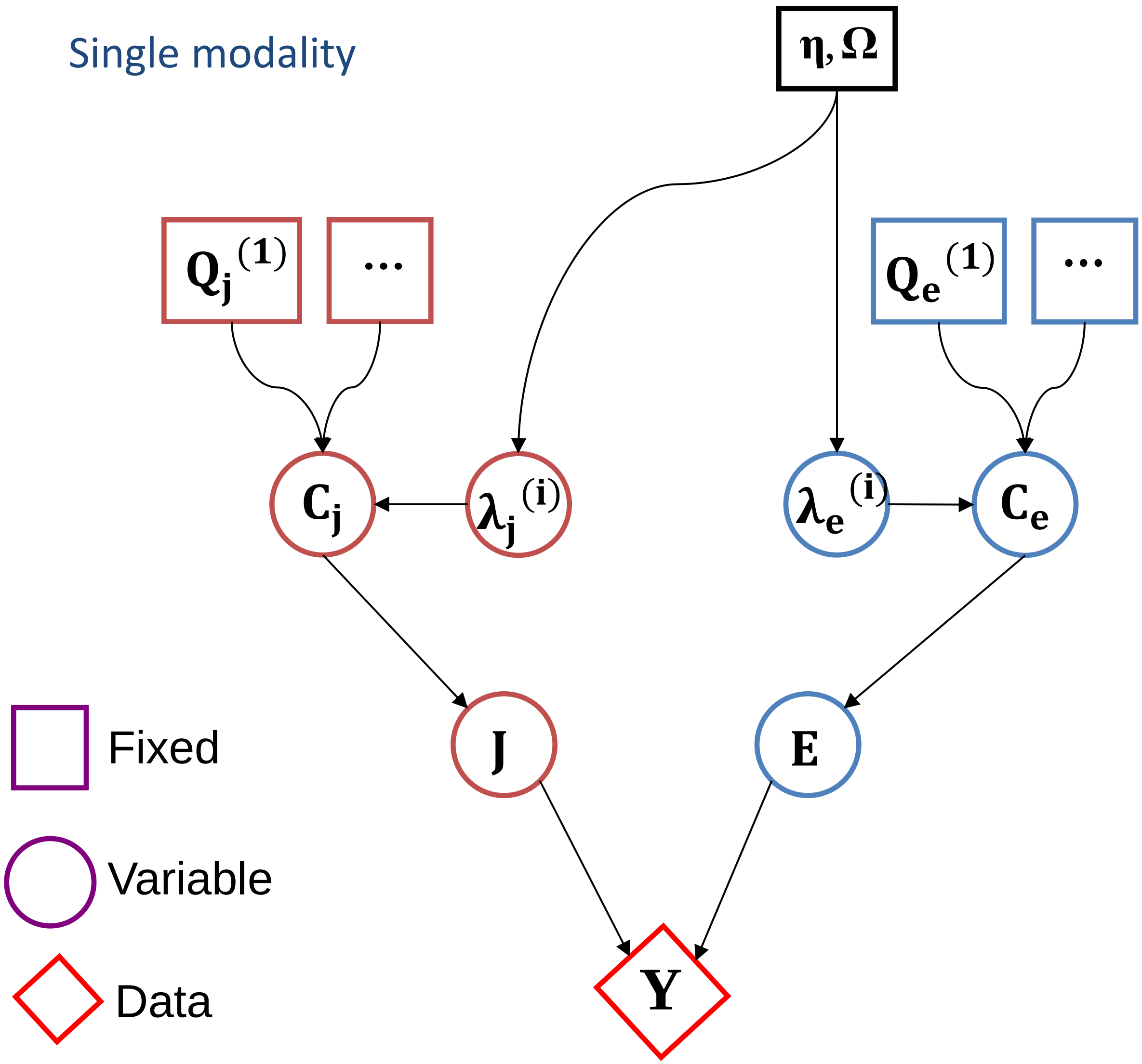
MSP (Group)



Litvak & Friston (2008) Neuroimage

Multi-modal integration: EEG-MEG fusion

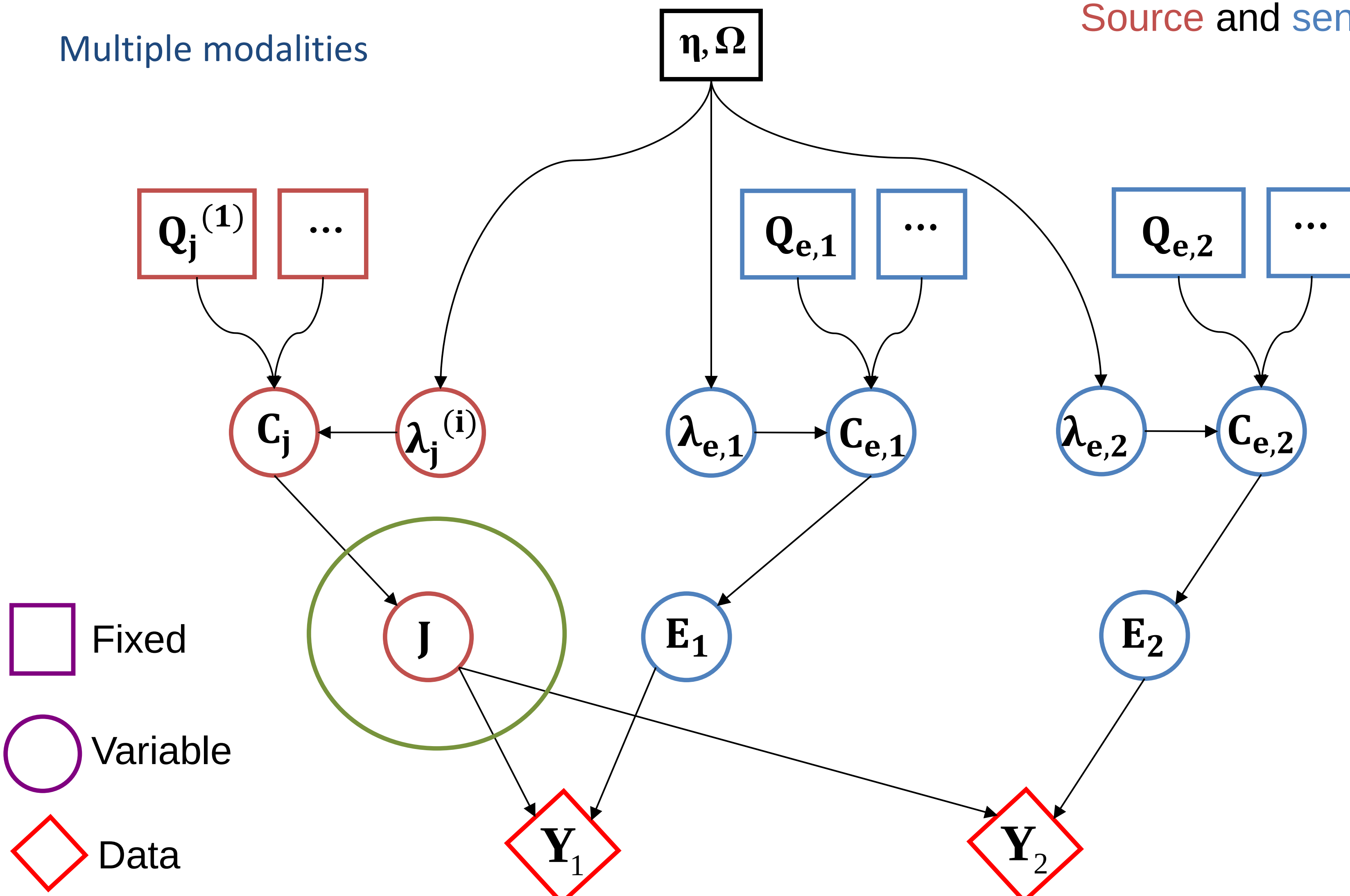
Source and sensor space



Multi-modal integration: EEG-MEG fusion

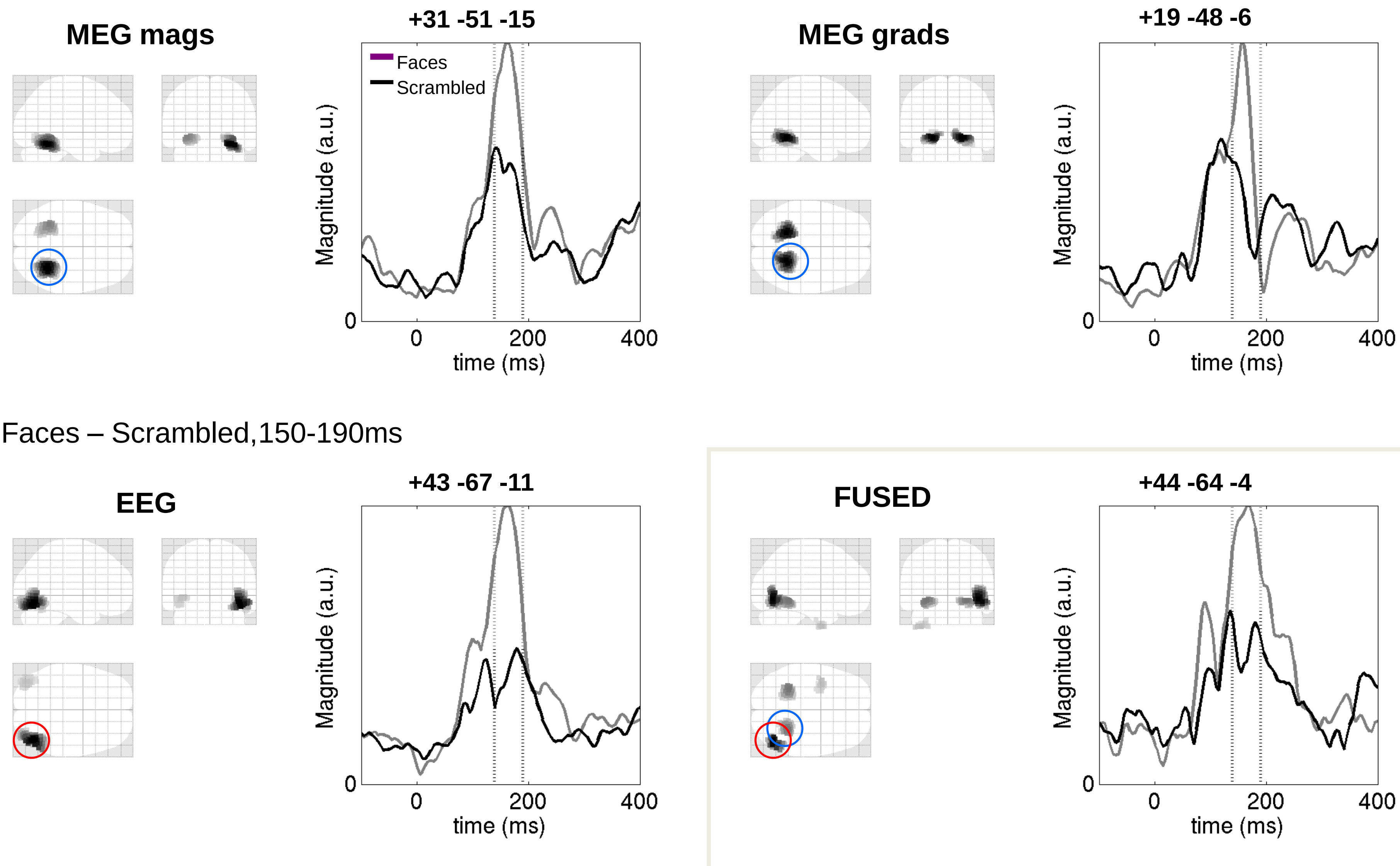
Multiple modalities

Source and sensor space



Henson et al. (2009) Neuroimage

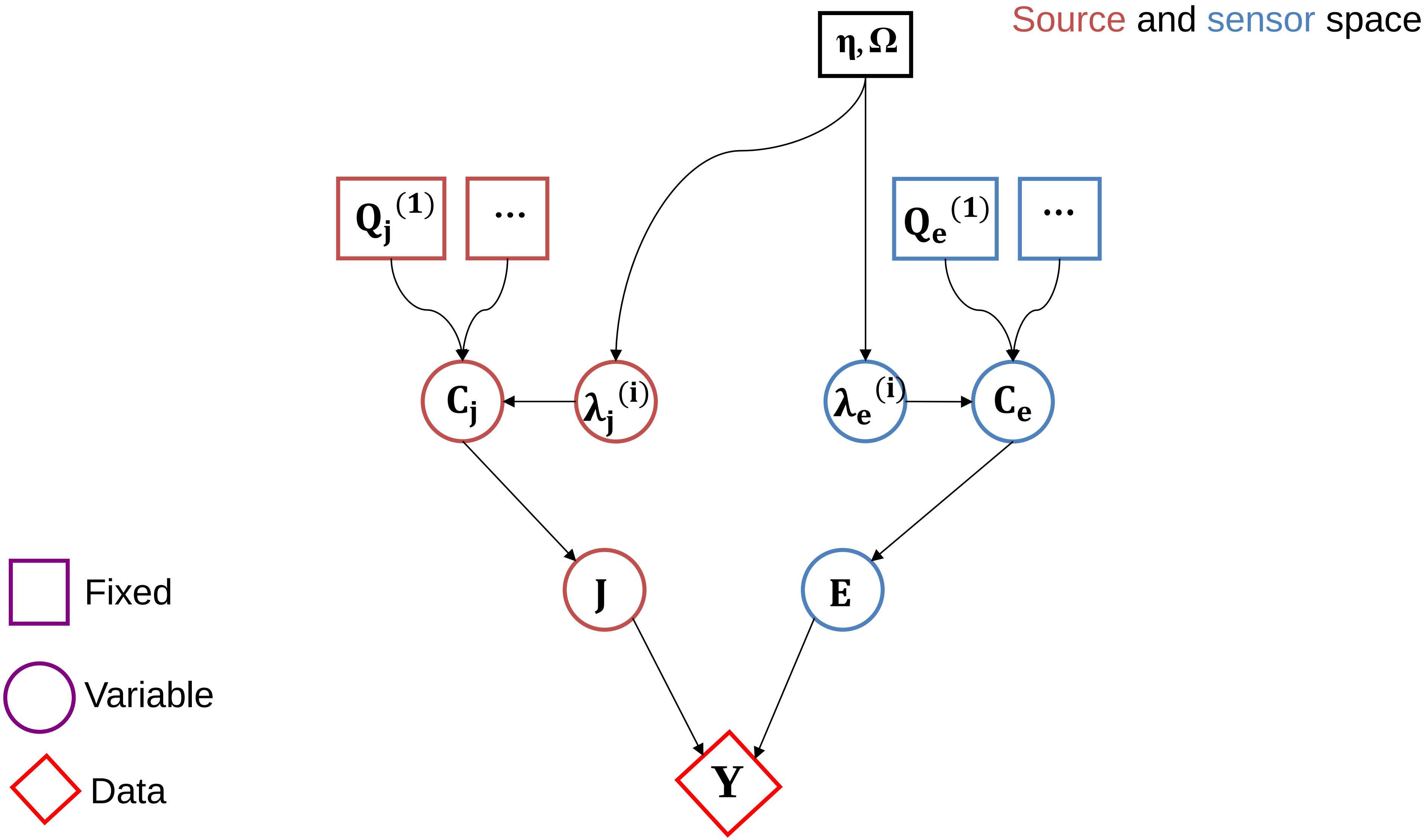
Multi-modal integration: EEG-MEG fusion



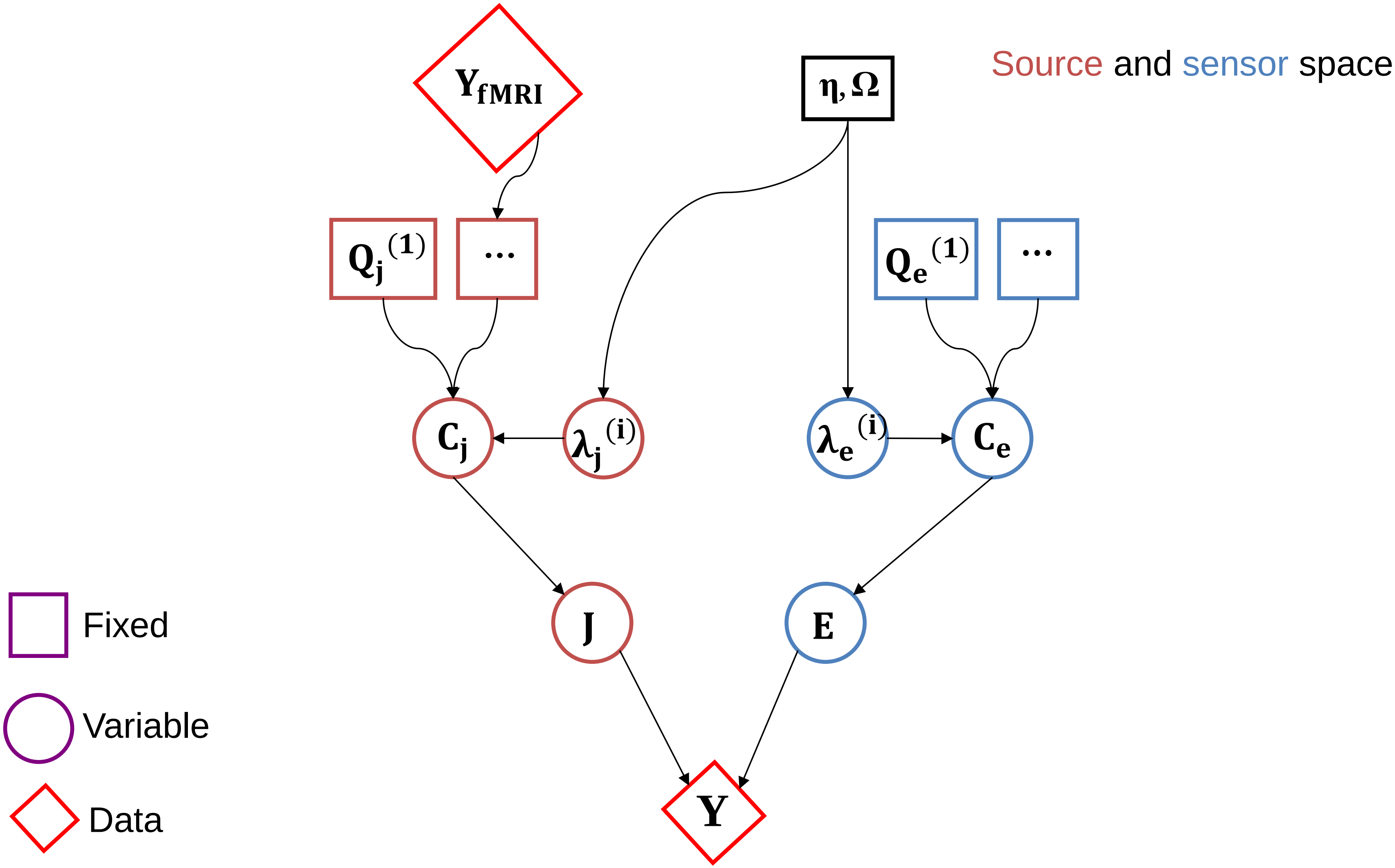
IID noise for each modality; common MSP for sources

Henson et al (2009) Neuroimage

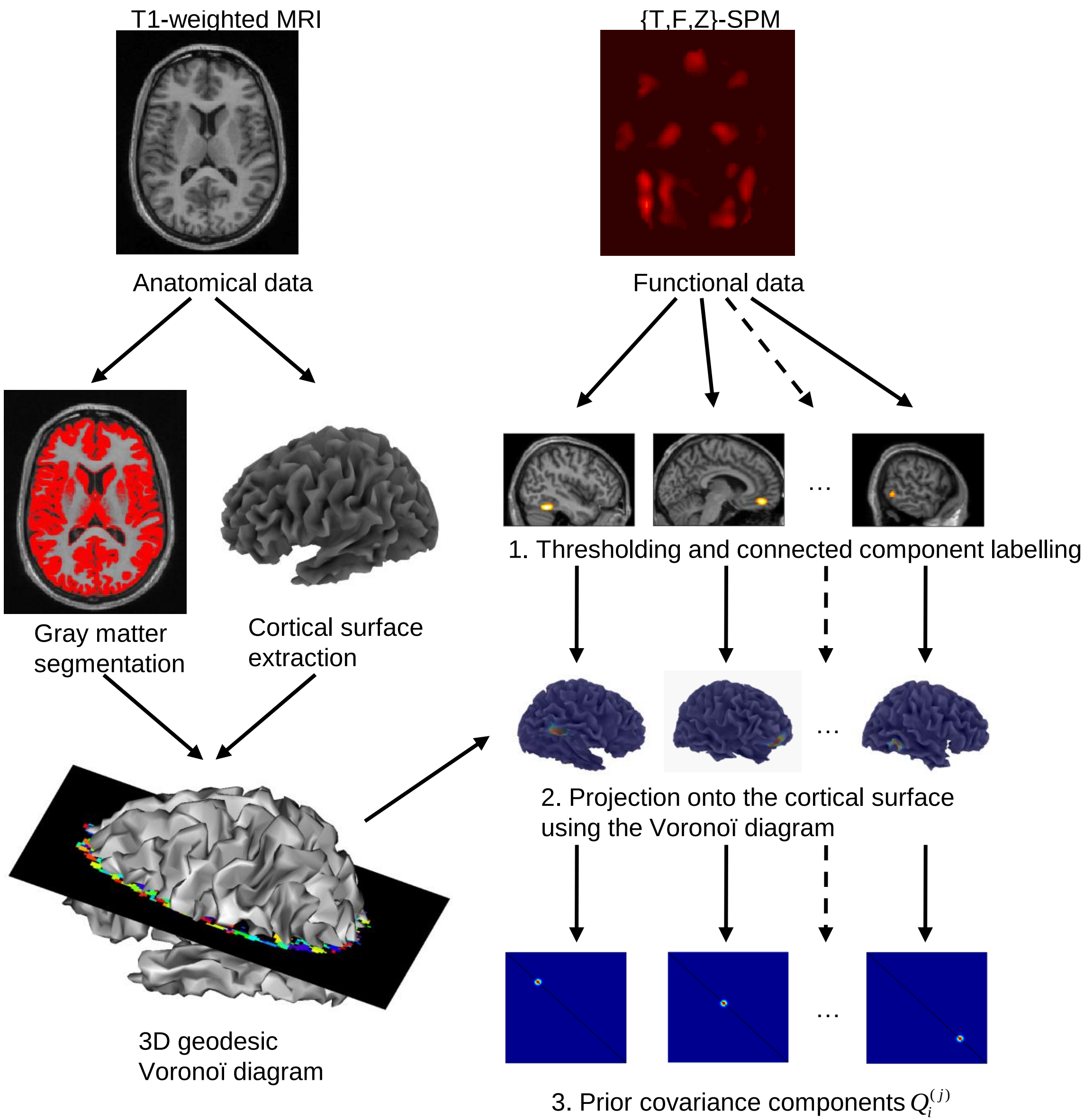
Multi-modal integration: fMRI priors



Multi-modal integration: fMRI priors

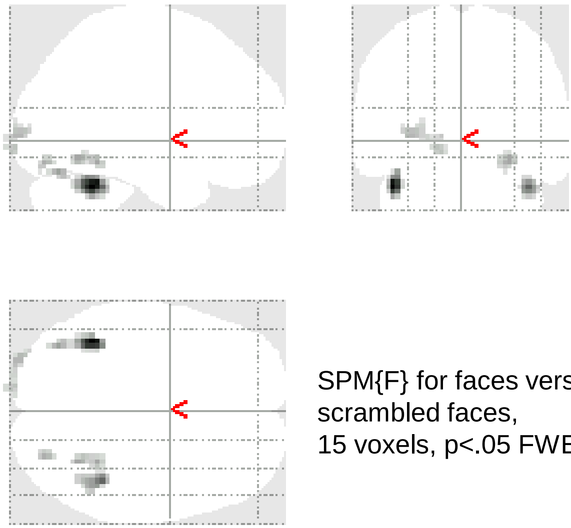


Multi-modal integration: fMRI priors

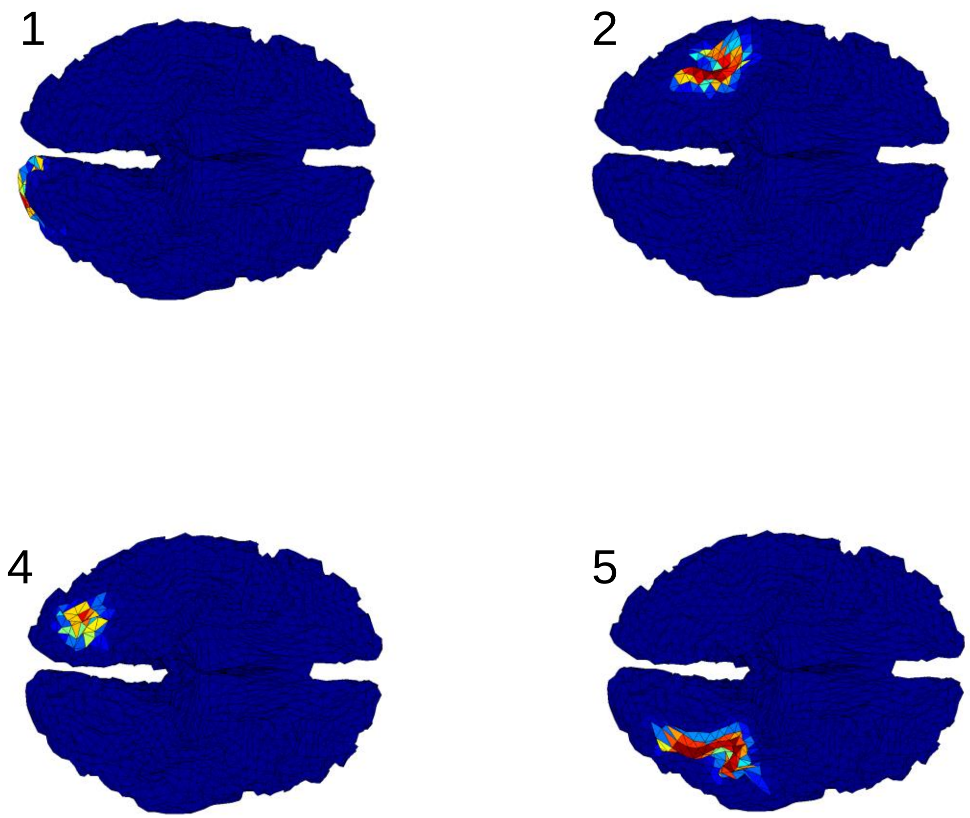


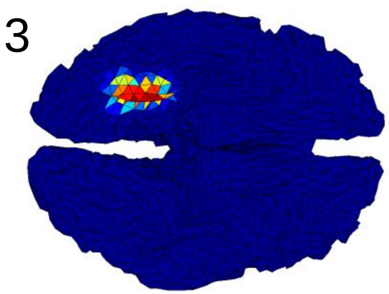
Henson et al (2010) Hum. Brain Map.

Multi-modal integration: fMRI priors



SPM{F} for faces versus
scrambled faces,
15 voxels, p<.05 FWE



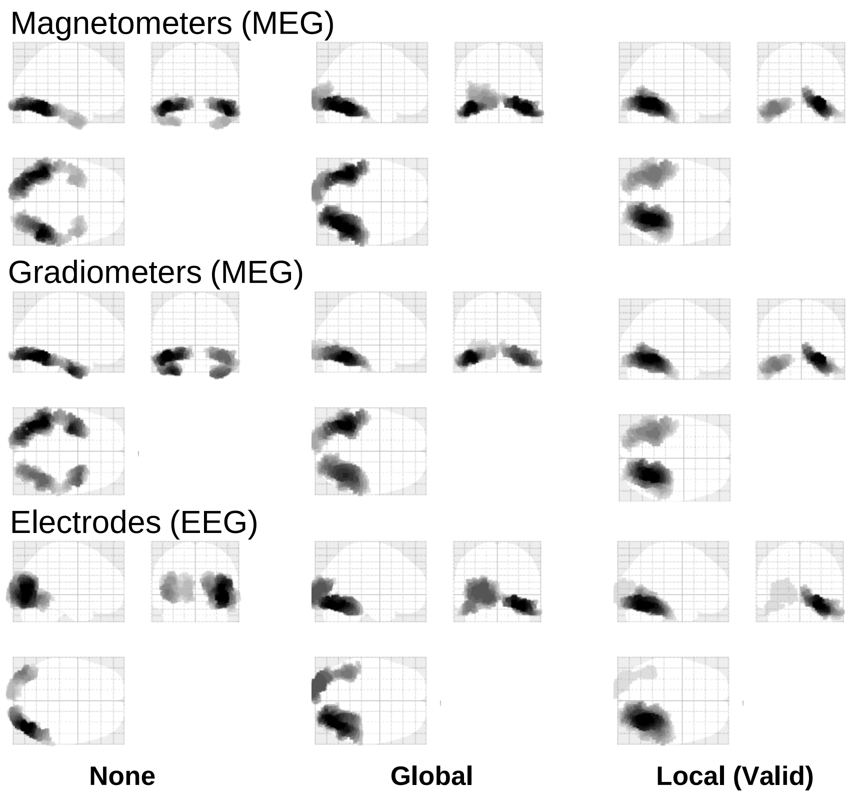


5 clusters from SPM of fMRI data from separate group of (18) subjects in MNI space

Henson et al (2010) Hum. Brain Map.

Multi-modal integration: fMRI priors

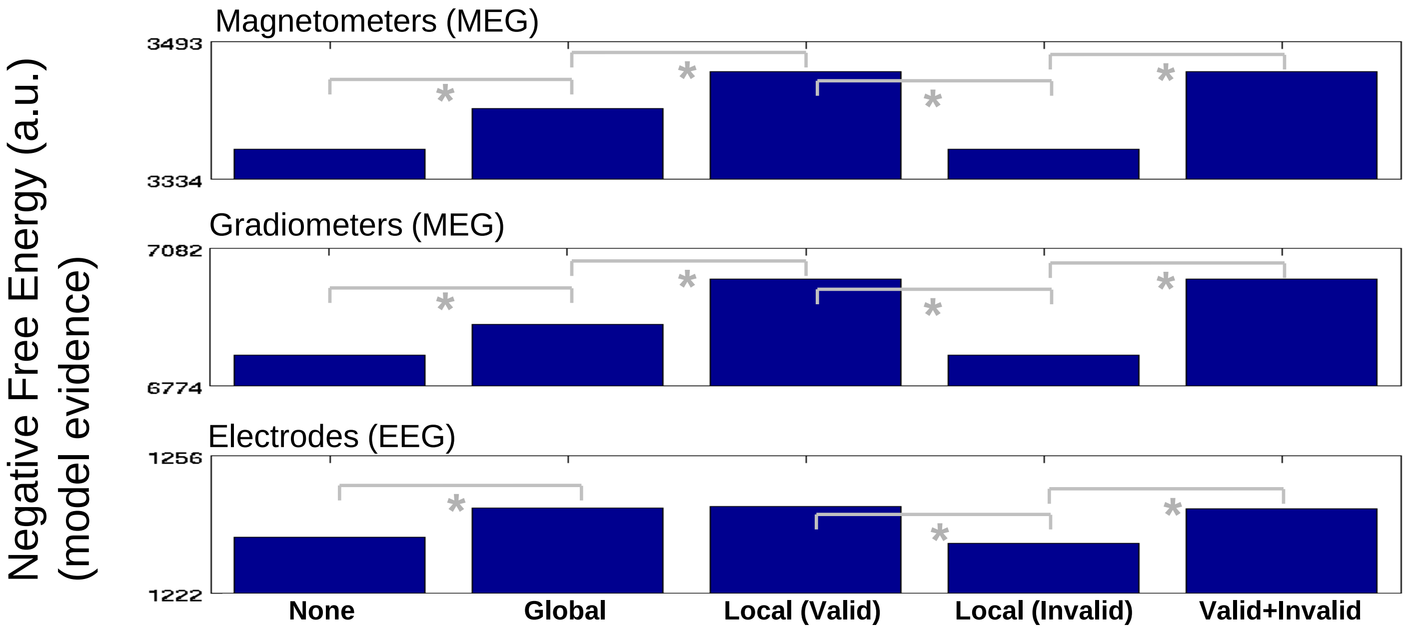
IID sources and IID noise (L2 MNM)



fMRI priors counteract superficial bias of L2-norm

Henson et al (2010) Hum. Brain Map.

Multi-modal integration: fMRI priors



Henson et al (2010) Hum. Brain Map.

Conclusion

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

2. …but offers unique Bayesian approaches to inversion:

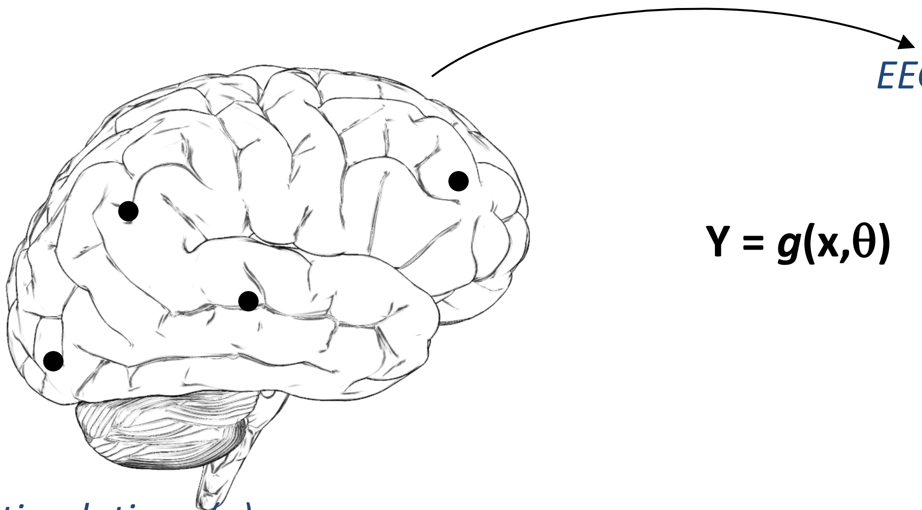
   2.1 Variational Bayesian ECD

   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

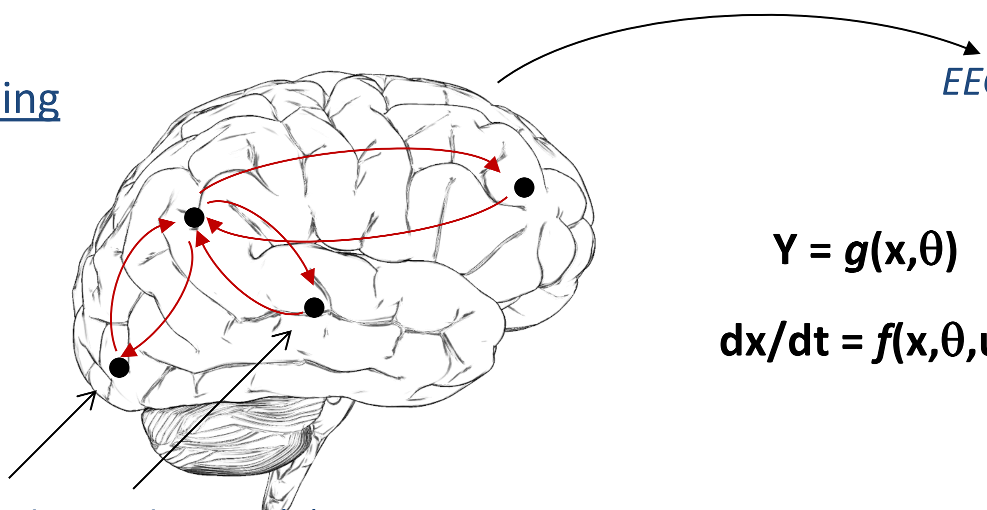


EEG/MEG Observations (Y)

$$Y = g(x,\theta)$$

Auditory-Visual Stimulations (u)

Dynamic causal modelling



EEG/MEG Observations (Y)

$$Y = g(x,\theta)$$

$$dx/dt = f(x,\theta,u)$$

Auditory-Visual Stimulations (u)