MRC | Cognition and Brain Sciences Unit

Multi-modal integration

Rik Henson

MRC CBU, Cambridge

Data-driven vs Model-driven

1. Data-driven in sense that no model that links specific features of one modality with features of another (still employ some form of statistical model), eg ICA, CCA

2. Model-driven in sense that either:

   2.1 test specific feature relations across modalities, eg SEM

   2.2 has a generative model of both modalities, eg PEB

Symmetric vs Asymmetric

MRC | Cognition and Brain Sciences Unit

1. Symmetric integration ("fusion") fits each modality simultaneously

2. Asymmetric integration uses one modality to model another modality

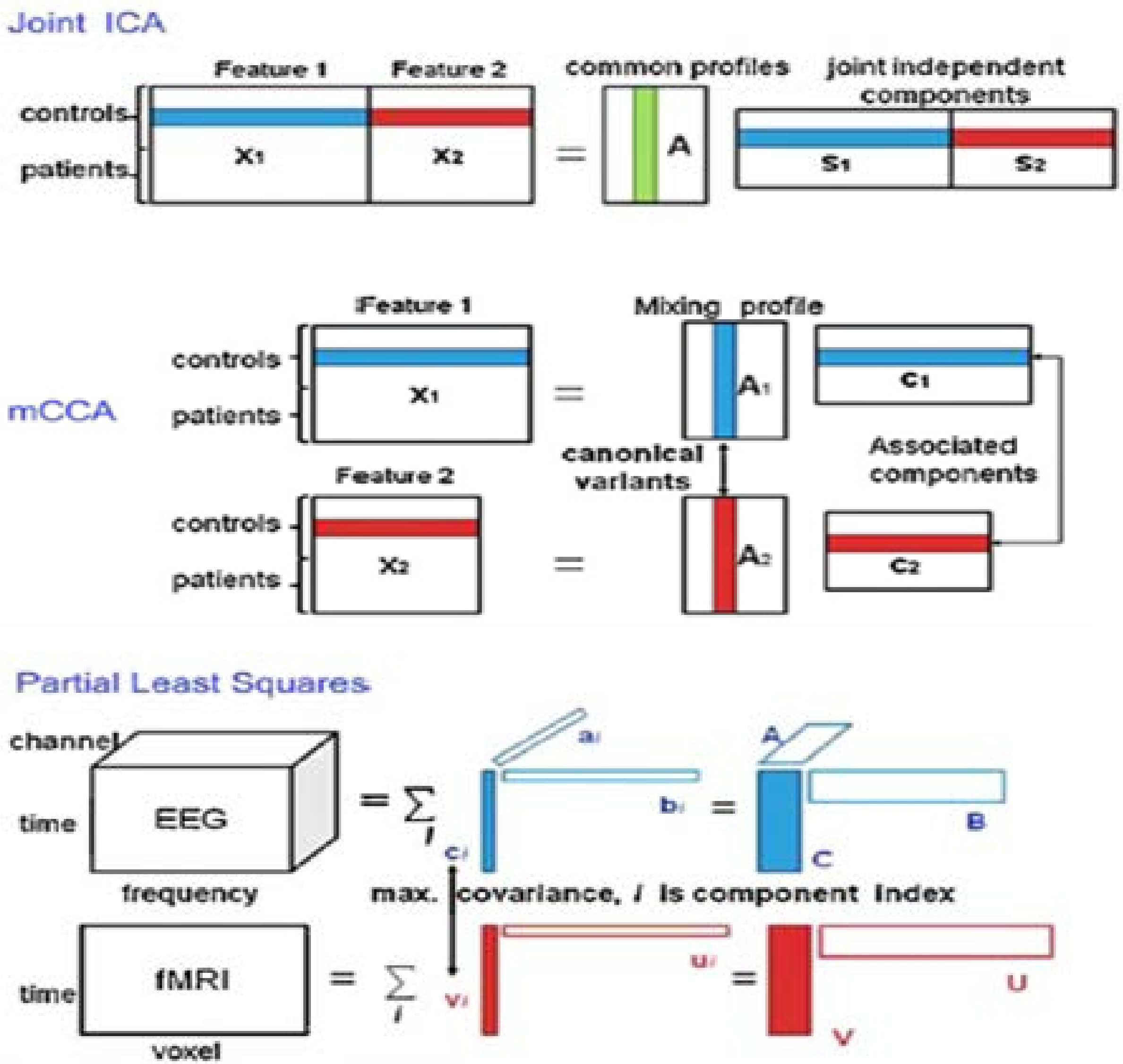
(Most data-driven approaches are symmetric; many, but not all, model-driven approaches are asymmetric)

Some Data-Driven Methods

1. Linked Matrix Factorisation methods (ICA, CCA, PLS)

2. Representational Similarity Analysis (RSA)

3. Graph Theory


Joint ICA
Feature 1
Feature 2
common profiles
joint independent components
controls
patients
X1
X2
A
S1
S2
mCCA
Feature 1
Mixing profile
controls
patients
X1
A1
C1
canonical variants
Associated components
Feature 2
X2
A2
C2
Partial Least Squares
channel
time
EEG
frequency
ai
bi
ci
A
B
C
max. covariance, i is component index
fMRI
voxel
ui
vi
U
V

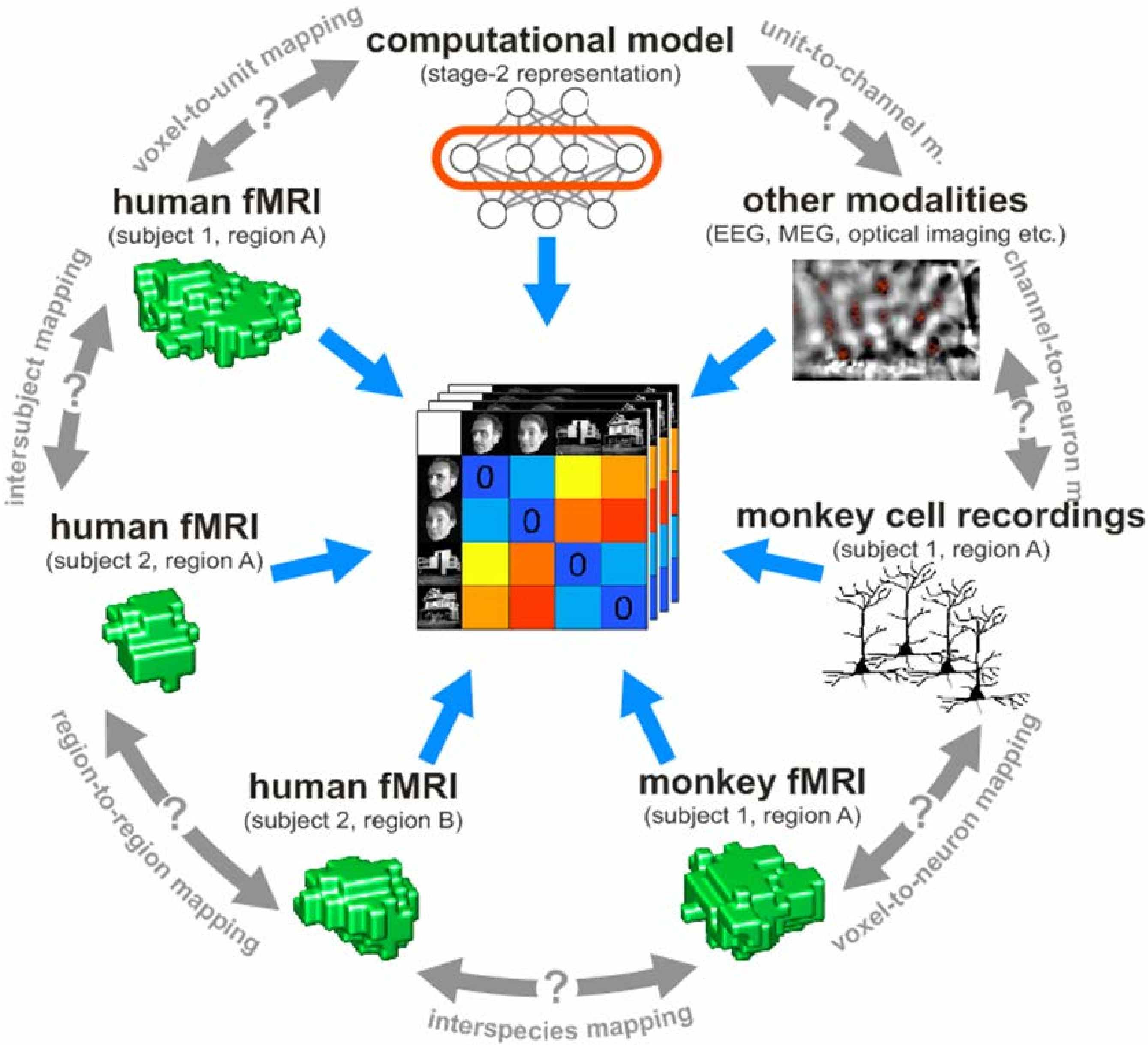
Some Data-Driven Methods

1. Linked Matrix Factorisation methods (ICA, PLS, CCA)

2. Representational Similarity Analysis (RSA)

3. Graph Theory

Kriegeskorte et al. 2008 *Frontiers in Systems Neuroscience*



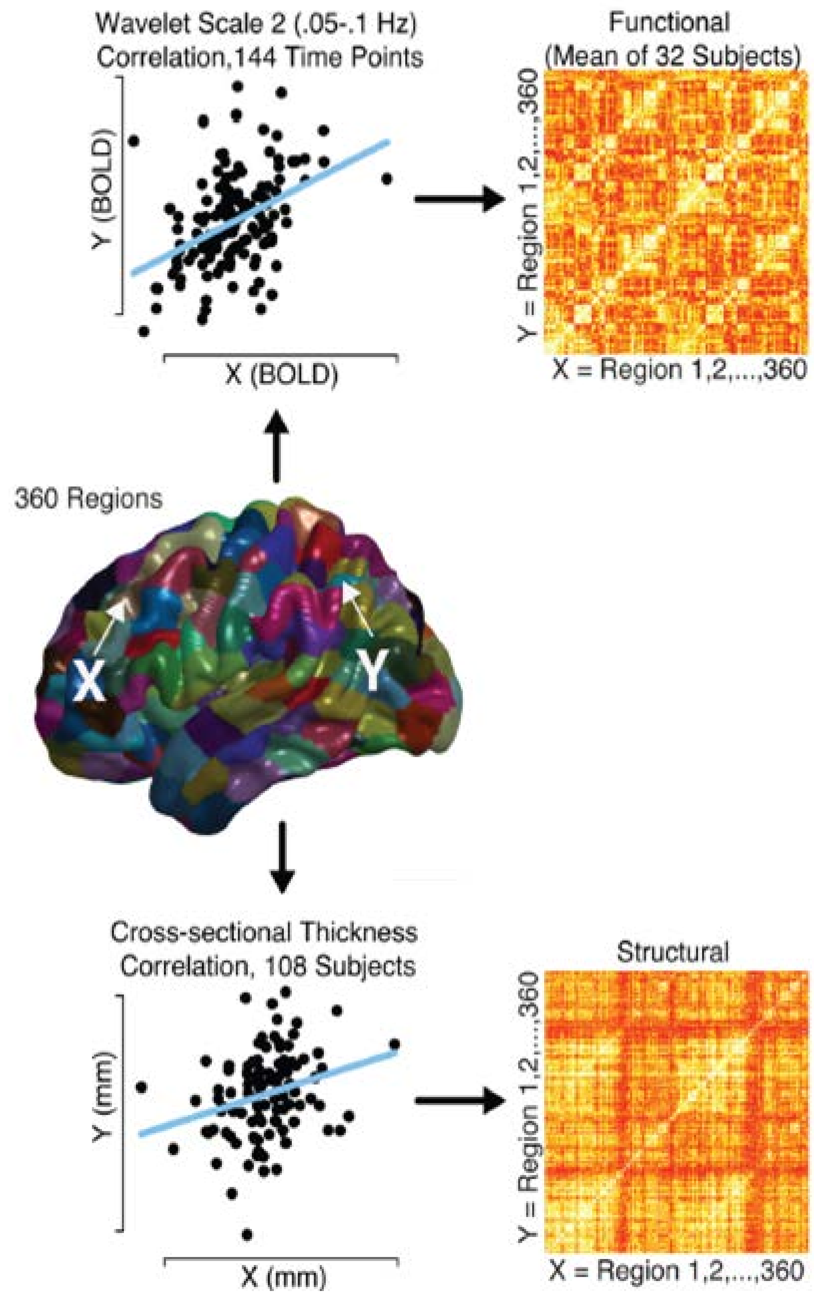
Some Data-Driven Methods

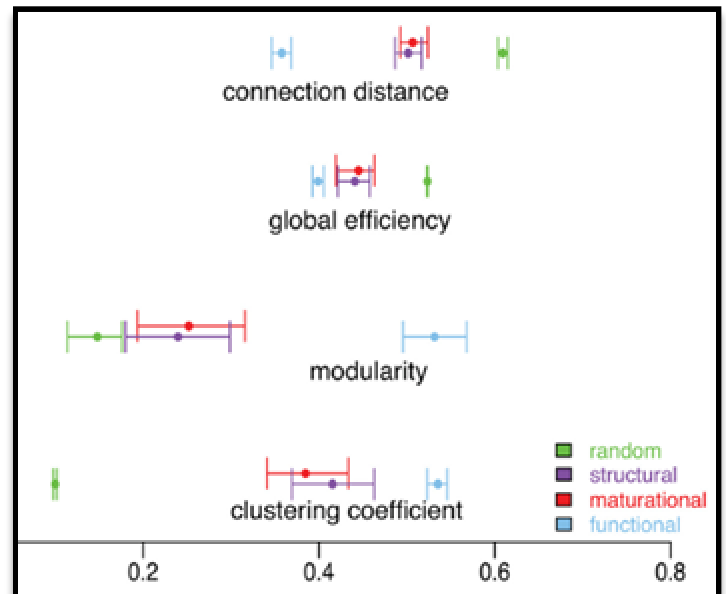
1. Linked Matrix Factorisation methods (ICA, PLS, CCA)

2. Representational Similarity Analysis (RSA)

3. Graph Theory

2890 • J. Neurosci., February 13, 2013 • 33(7):2889–2899 Alexander-Bloch et al. • Maturational Change and Structural Covariance





…or can even compare graphs with different nodes, eg fMRI ROIs and MEEG sensors…

Data-driven vs Model-driven

1. Data-driven in sense that no model that links specific features of one modality with features of another (still employ some form of statistical model), eg ICA, CCA

2. Model-driven in sense that either:

    2.1 test specific feature relations across modalities, eg SEM

    2.2 has a generative model of both modalities, eg PEB

Some Model-Based Examples (local CamCAN examples)

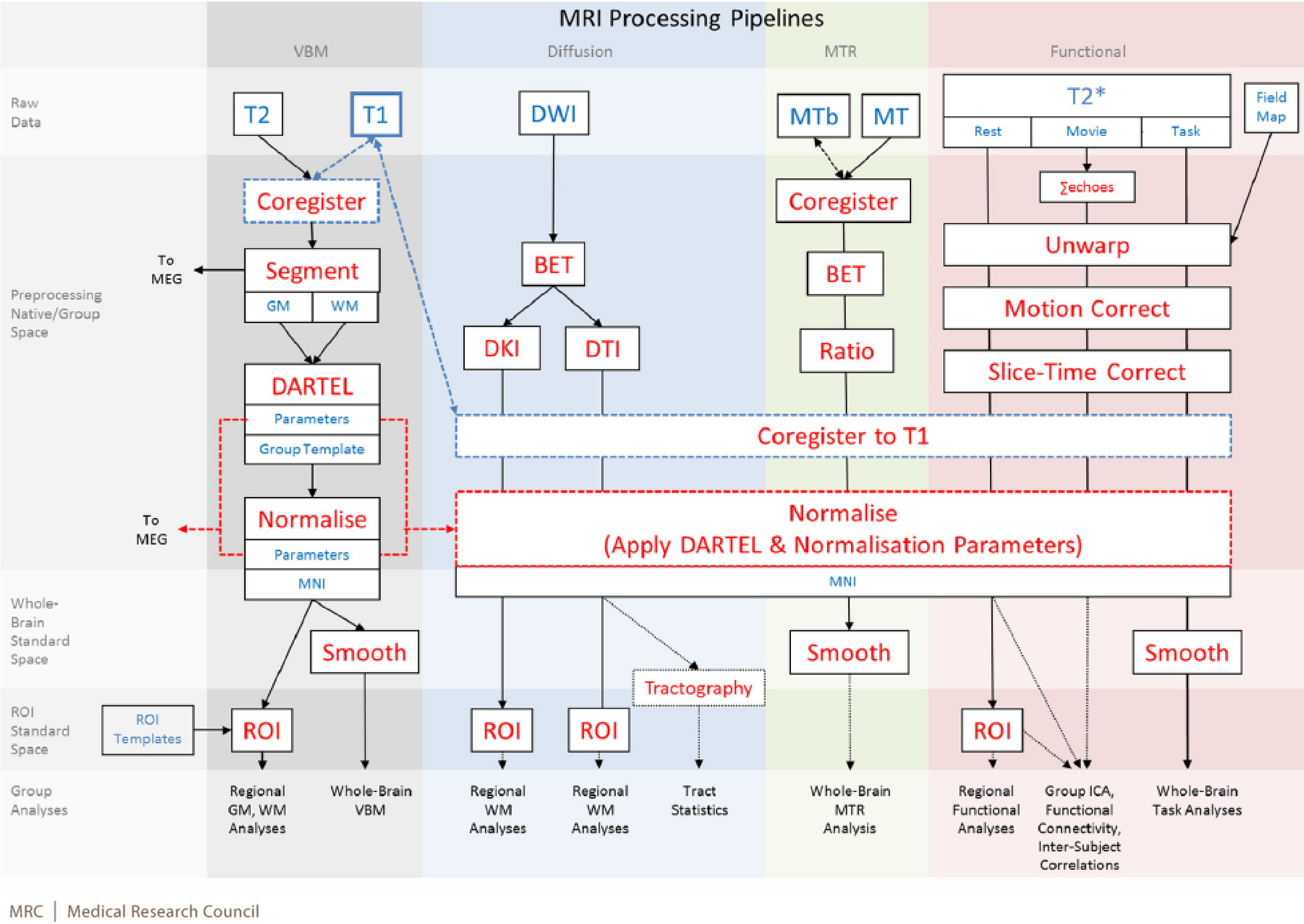
MRC Cognition and Brain Sciences Unit

1. Combining T1-, T2-, Diffusion- and MT-weighted images for segmentation and normalisation

2. Univariate mediation: Using MEG to separate effects of age on neural vs vascular responsivity in fMRI

3. Univariate mediation: Using DKI to investigate effects of age on MEG latency

4. Multivariate Structural Equation Modelling (SEM) to separate GM and WM contributions to executive function


MRI Processing Pipelines
VBM
Diffusion
MTR
Functional
Raw Data
Preprocessing Native/Group Space
Whole-Brain Standard Space
ROI Standard Space
Group Analyses
T2
T1
DWI
MTb
MT
T2*
Rest
Movie
Task
Field Map
Coregister
To MEG
Segment
GM
WM
DARTEL
Parameters
Group Template
Normalise
Parameters
MNI
To MEG
BET
DKI
DTI
Coregister
BET
Ratio
∑echoes
Unwarp
Motion Correct
Slice-Time Correct
Coregister to T1
Normalise
(Apply DARTEL & Normalisation Parameters)
MNI
Smooth
Smooth
Smooth
Tractography
ROI Templates
ROI
ROI
ROI
ROI
Regional GM, WM Analyses
Whole-Brain VBM
Regional WM Analyses
Regional WM Analyses
Tract Statistics
Whole-Brain MTR Analysis
Regional Functional Analyses
Group ICA, Functional Connectivity, Inter-Subject Correlations
Whole-Brain Task Analyses
MRC | Medical Research Council

Some Model-Based Examples (local CamCAN examples)

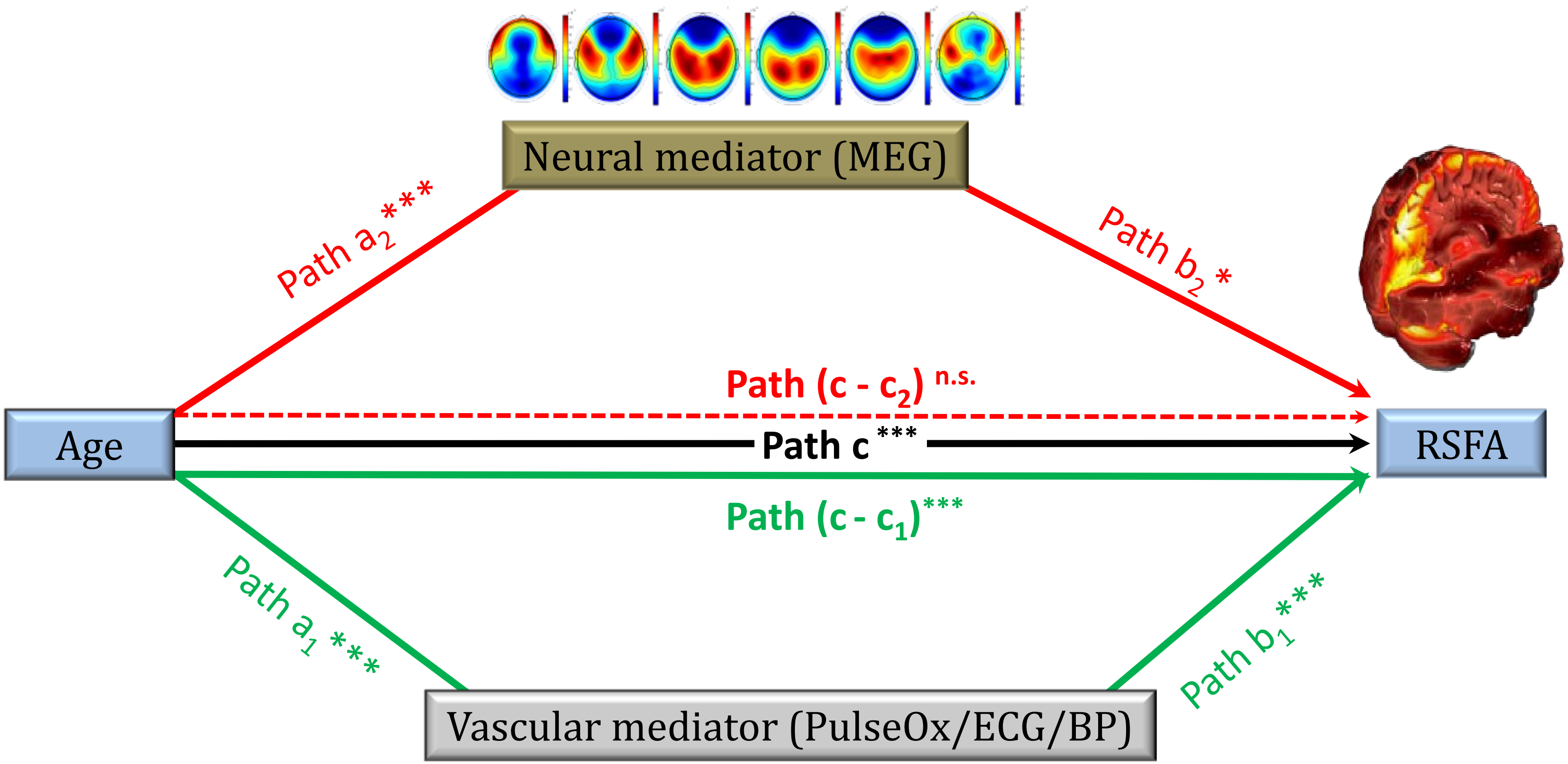
MRC | Cognition and Brain Sciences Unit

1. Combining T1-, T2-, Diffusion- and MT-weighted images for segmentation and normalisation

2. Univariate mediation: Using MEG to separate effects of age on neural vs vascular responsivity in fMRI

3. Univariate mediation: Using DKI to investigate effects of age on MEG latency

4. Multivariate Structural Equation Modelling (SEM) to separate GM and WM contributions to executive function



Tsvetanov et al (2015), Hum. Brain. Mapping

Some Model-Based Examples (local CamCAN examples)

MRC Cognition and Brain Sciences Unit

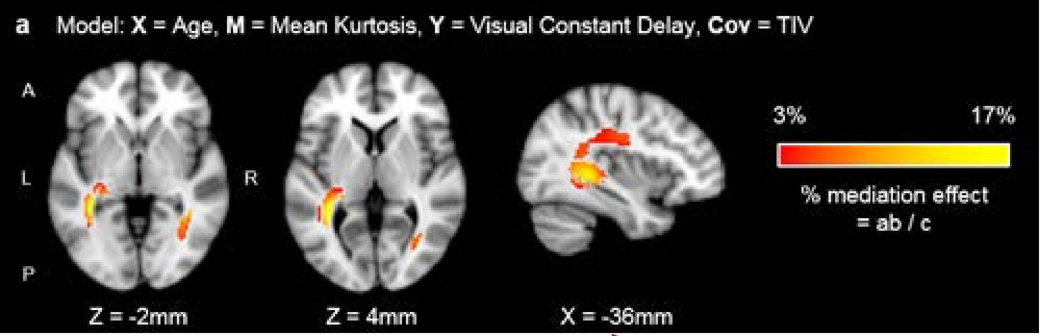
1. Combining T1-, T2-, Diffusion- and MT-weighted images for segmentation and normalisation

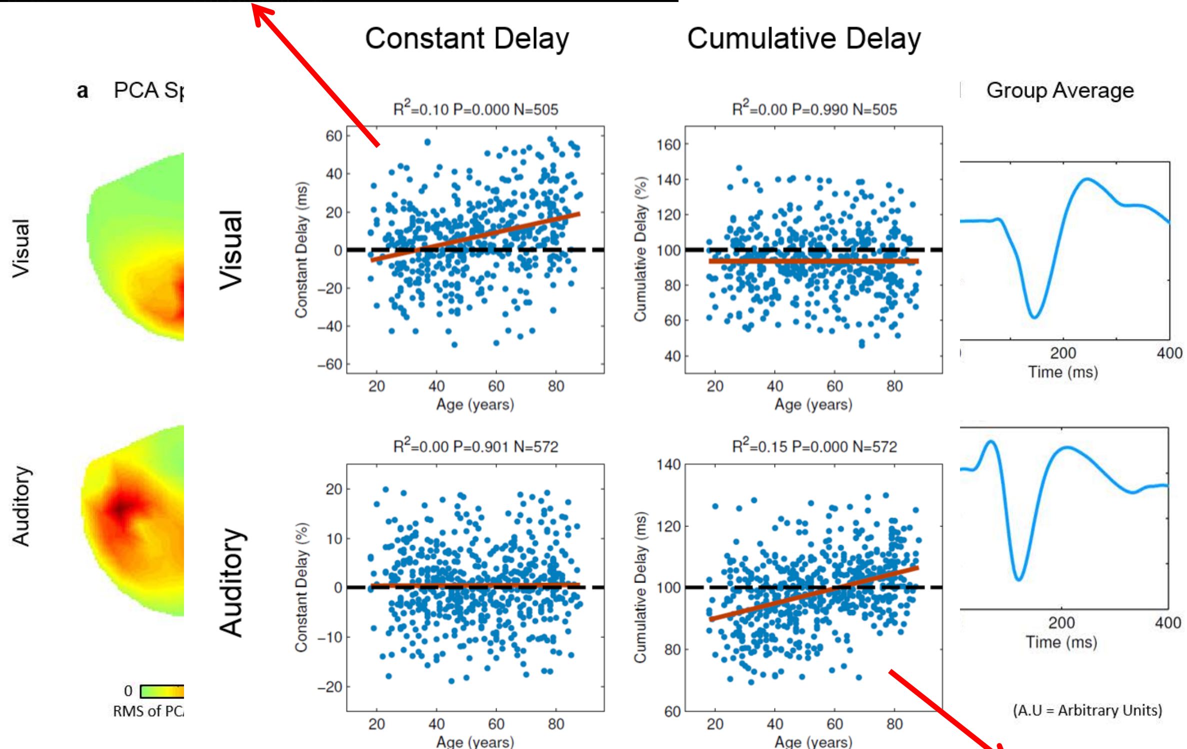
2. Univariate mediation: Using MEG to separate effects of age on neural vs vascular responsivity in fMRI

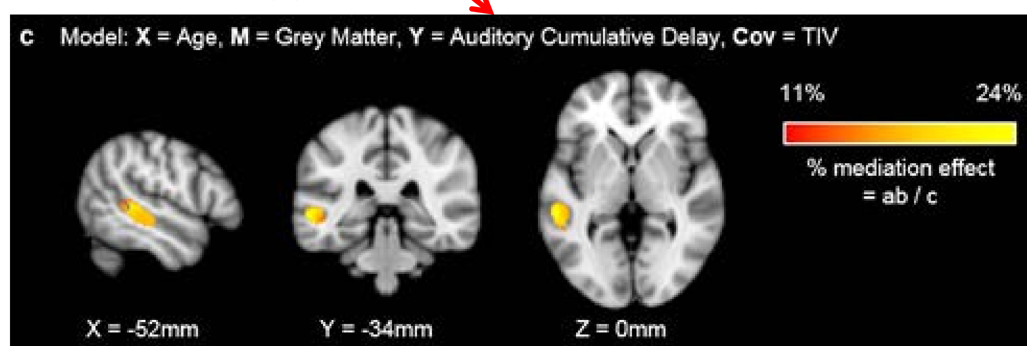
3. Univariate mediation: Using DKI to investigate effects of age on MEG latency

4. Multivariate Structural Equation Modelling (SEM) to separate GM and WM contributions to executive function

MRC Cognition and Brain Sciences Unit







Price et al (submitted)

Some Model-Based Examples (local CamCAN examples)

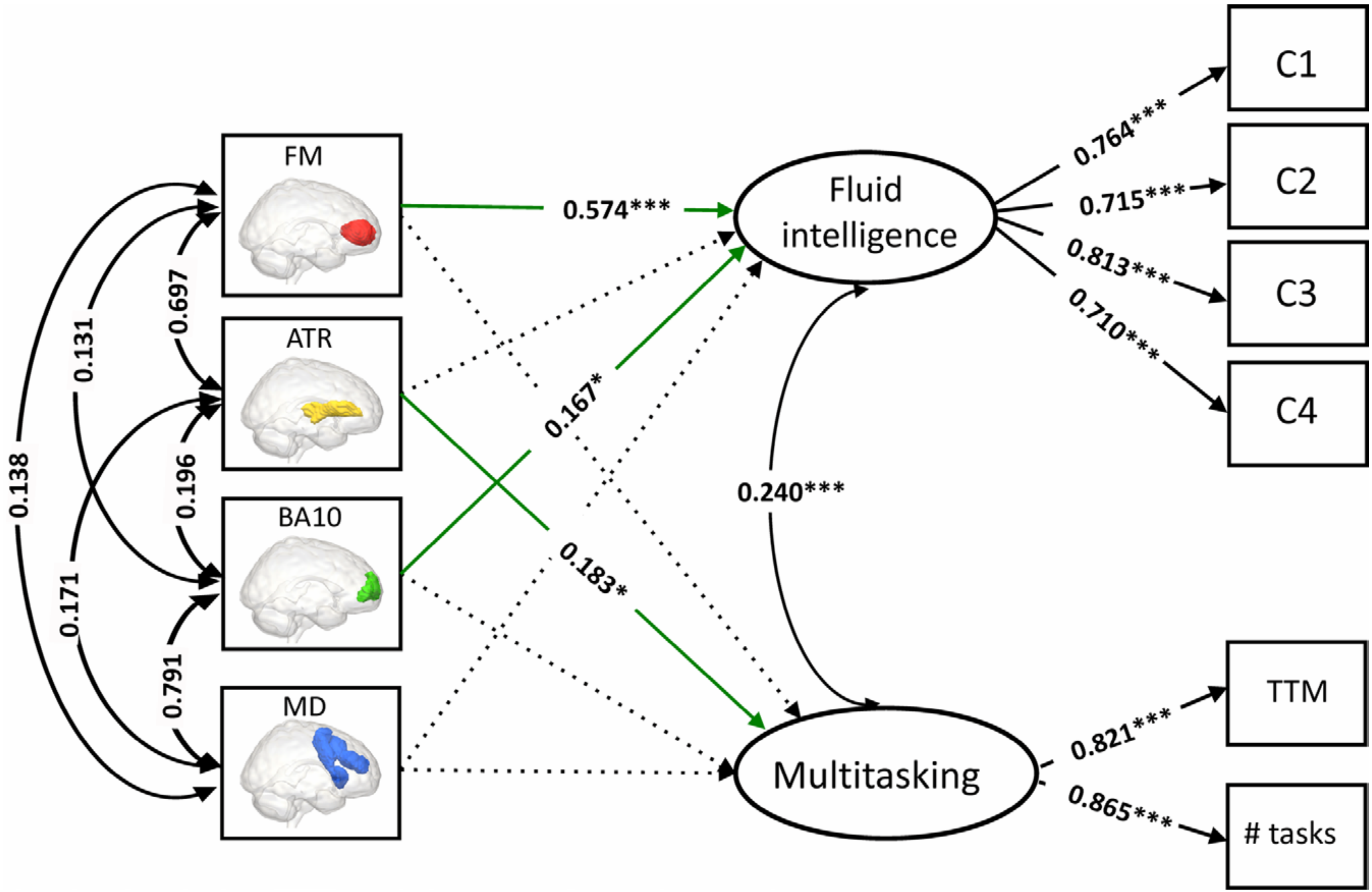
MRC Cognition and Brain Sciences Unit

1. Combining T1-, T2-, Diffusion- and MT-weighted images for segmentation and normalisation

2. Univariate mediation: Using MEG to separate effects of age on neural vs vascular responsivity in fMRI

3. Univariate mediation: Using DKI to investigate effects of age on MEG latency

4. Multivariate Structural Equation Modelling (SEM) to separate GM and WM contributions to executive function


FM
ATR
BA10
MD
Fluid intelligence
Multitasking
C1
C2
C3
C4
TTM
tasks
0.574***
0.167*
0.183*
0.240***
0.764***
0.715***
0.813***
0.710***
0.821***
0.865***
0.697
0.131
0.138
0.196
0.171
0.791

Data-driven vs Model-driven

MRC | Cognition and Brain Sciences Unit

1. Data-driven in sense that no model that links specific features of one modality with features of another (still employ some form of statistical model), eg ICA, CCA

2. Model-driven in sense that either:

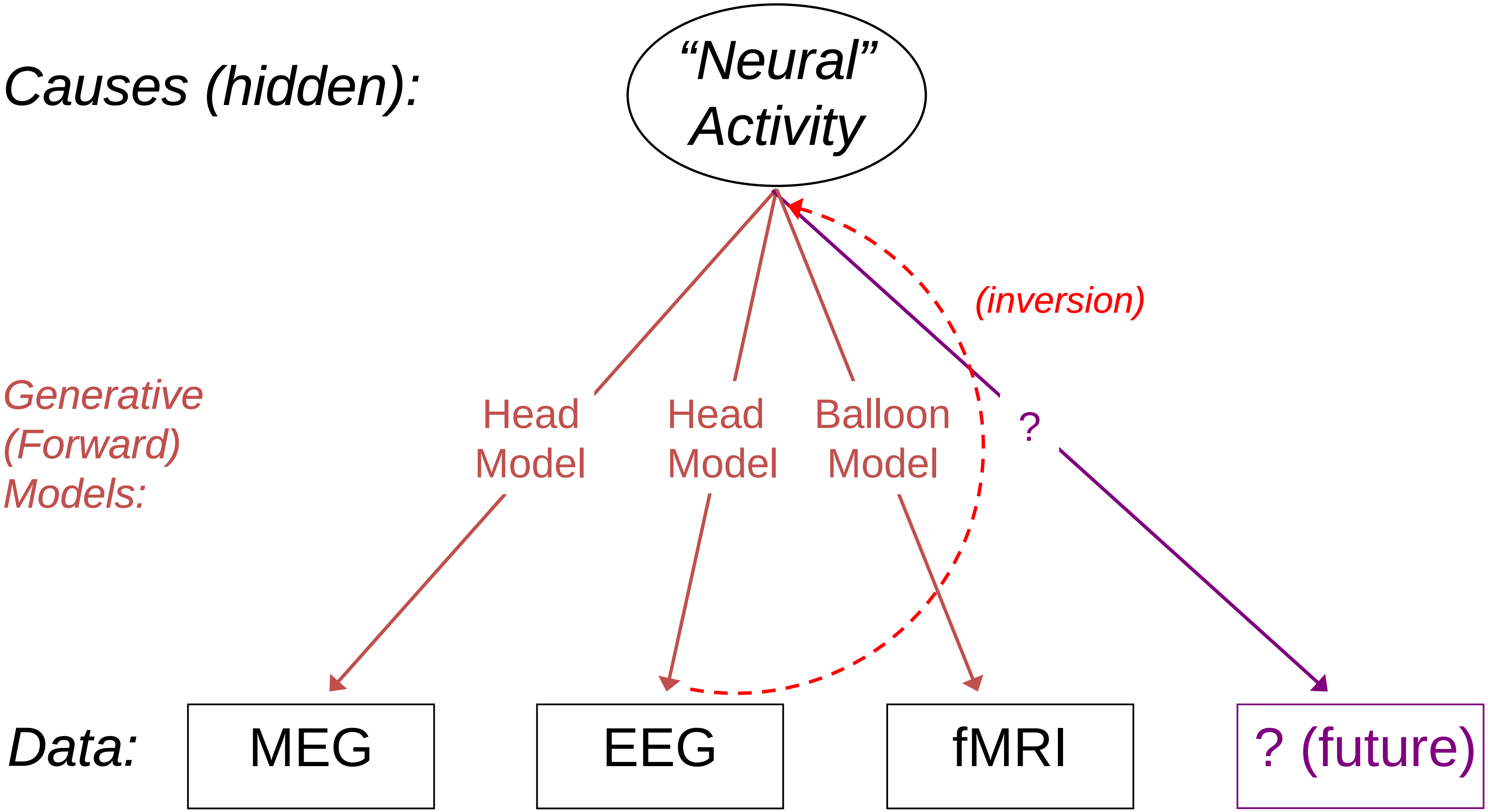
   2.1 test specific feature relations across modalities, eg SEM

   2.2 has a generative model of both modalities, eg PEB

Multi-modal integration of MEG, EEG & fMRI

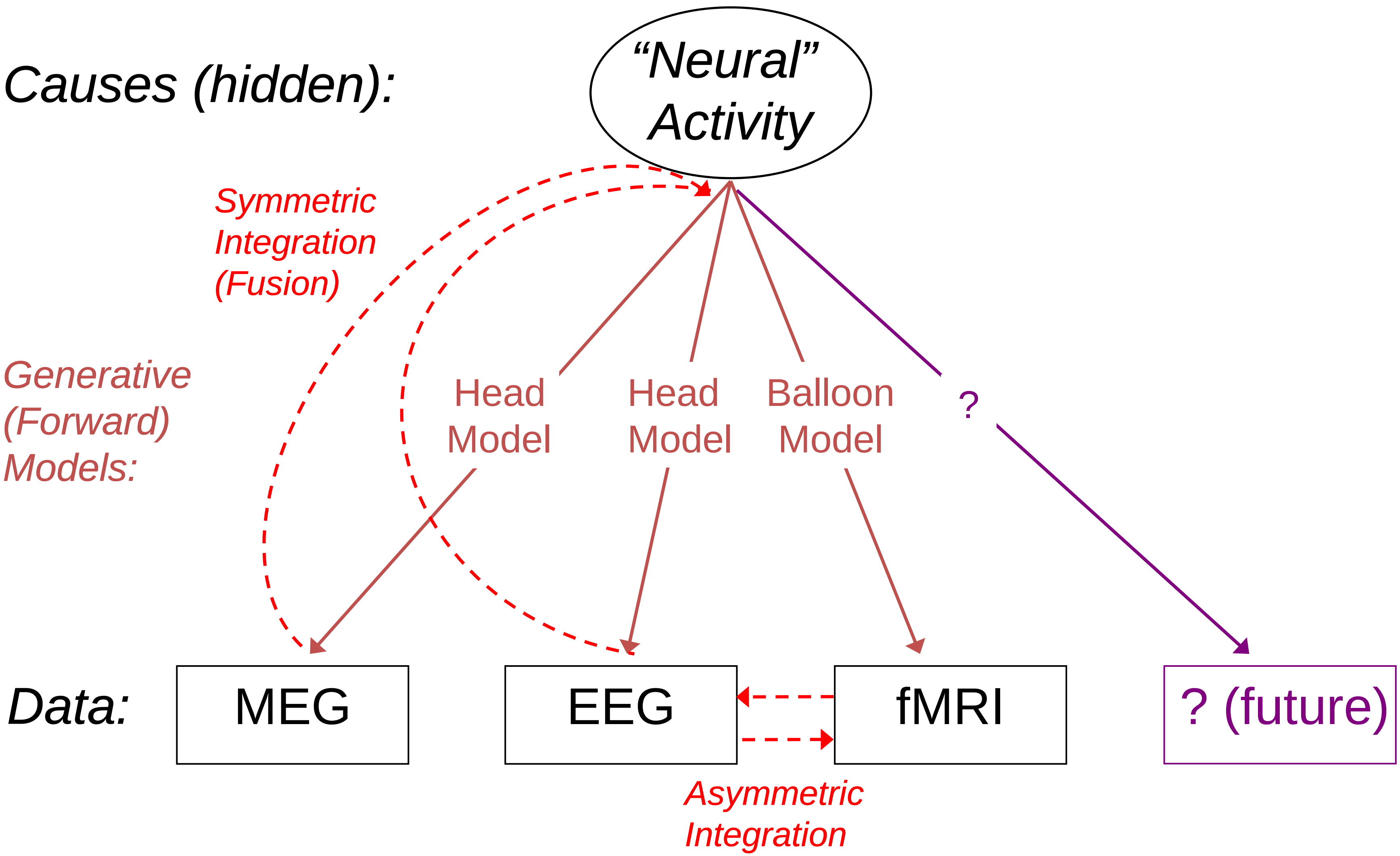
Multi-modal Integration

MRC | Cognition and Brain Sciences Unit



Multi-modal Integration

MRC | Cognition and Brain Sciences Unit



Daunizeau et al (2007), Neuroimage

Examples

1. EEG -> fMRI asymmetric integration

2. fMRI -> M/EEG asymmetric integration

3. MEG <-> EEG symmetric integration (fusion)

Examples

1. EEG -> fMRI asymmetric integration

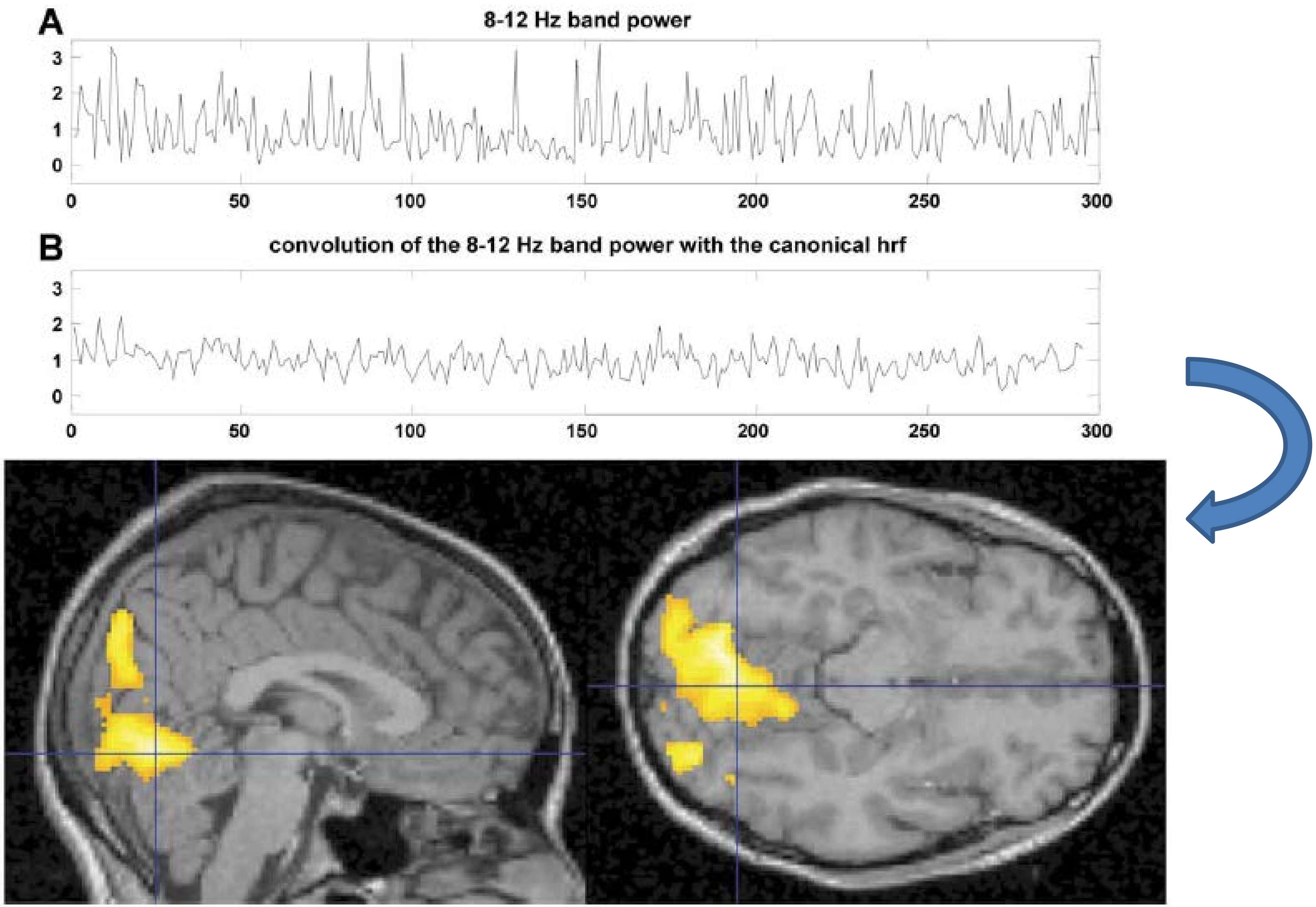
2. fMRI -> M/EEG asymmetric integration

3. MEG <-> EEG symmetric integration (fusion)

Concurrent EEG and fMRI

MRC Cognition and Brain Sciences Unit

H. Laufs et al. / NeuroImage 19 (2003) 1463–1476



Examples

1. EEG -> fMRI asymmetric integration

2. fMRI -> M/EEG asymmetric integration

3. MEG <-> EEG symmetric integration (fusion)

Examples

MRC | Cognition and Brain Sciences Unit

1. EEG -> fMRI asymmetric integration

(Background: The M/EEG inverse problem)

3. MEG <-> EEG symmetric integration (fusion)

M/EEG Linear Forward Model

Given *n* sensors and *p* sources fixed in location and orientation (e.g, on a cortical mesh), then linear Forward Model (for single timepoint):

$$\begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = \begin{bmatrix} L_{11} & L_{12} & \cdots L_{1p} \\ \vdots & \ddots & \vdots \\ L_{n1} & \cdots & \cdots L_{np} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_p \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$$

d = Data	*n* sensors
s = Sources	*p*>>*n* sources
L = Leadfields	*n* sensors *x p* sources
e = Error (noise)	*n* sensors…

Equivalent matrix format:

$$\mathbf{d} = \mathbf{L}\mathbf{s} + \mathbf{e}$$

Assume sensor noise is zero-mean Gaussian with error covariance $\mathbf{C}^{(e)}$:

$$e \sim N(0, \mathbf{C}^{(e)})$$

Assume sources similarly Gaussian with source covariance $\mathbf{C}^{(s)}$:

$$s \sim N(0, \mathbf{C}^{(s)})$$

M/EEG Linear Forward Model Assumptions to Solve

$$\mathbf{d} = \mathbf{L}\mathbf{s} + \mathbf{e}$$

$$\mathbf{e} \sim N(0, \mathbf{C}^{(e)})$$

$$\mathbf{s} \sim N(0, \mathbf{C}^{(s)})$$

d = Data — *n* sensors
s = Sources — *p*>>*n* sources
L = Leadfields — *n* sensors *x* *p* sources
e = Error (noise) — *n* sensors…

General solution is:

Hauk (2004), Neuroimage

$$\hat{\mathbf{s}} = \mathbf{C}^{(s)}\mathbf{L}^T(\mathbf{L}\mathbf{C}^{(s)}\mathbf{L}^T + \lambda\mathbf{C}^{(e)})^{-1}\mathbf{d}$$

λ = Regularisation (hyperparameter)

But how calculate $\mathbf{C}^{(e)}$ and $\mathbf{C}^{(s)}$?

MEG Linear Forward Model Assumptions to Solve

MRC Cognition and Brain Sciences Unit

One approach is to model sources and noise by variance components:

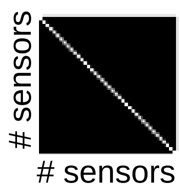
$$\mathbf{C} = \sum_i \lambda_i \mathbf{Q}_i$$

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

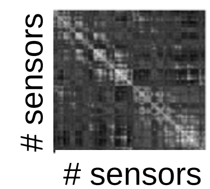
Friston et al (2008) Neuroimage

1. Sensor components, $\mathbf{Q}_i^{(e)}$ (error):

"IID" (white noise):

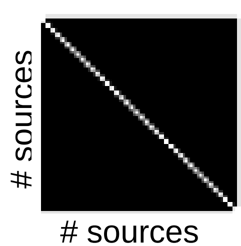


Empty-room:

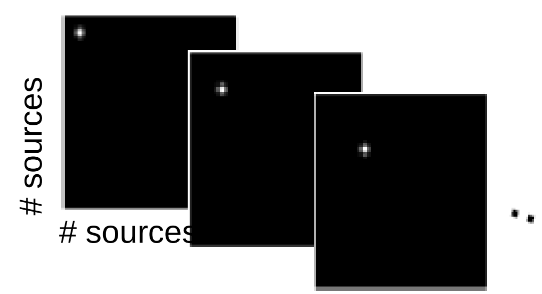


2. Source components, $\mathbf{Q}_i^{(s)}$ (priors/regularisation):

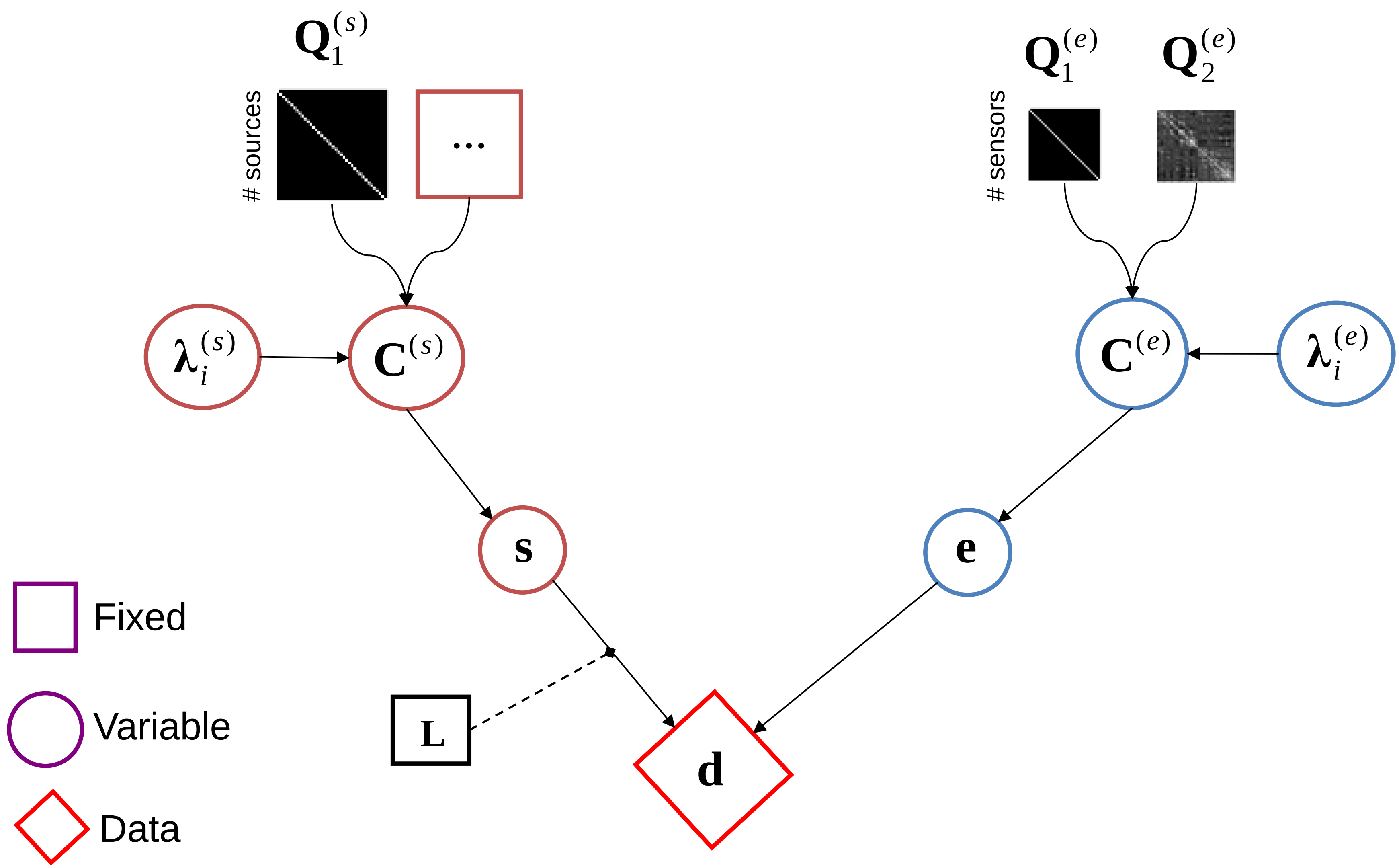
"IID" (min norm):



Multiple Sparse Priors (MSP):



MEG Generative Model



Henson et al (2009) Neuroimage

M/EEG Linear Forward Model
Assumptions to Solve

MRC | Cognition and Brain Sciences Unit

$$\mathbf{d} = \mathbf{Ls} + \mathbf{e}$$

$$\mathbf{e} \sim N(0, \mathbf{C}^{(e)})$$

$$\mathbf{s} \sim N(0, \mathbf{C}^{(s)})$$

d = Data — *n* sensors
s = Sources — *p*>>*n* sources
L = Leadfields — *n* sensors *x* *p* sources
e = Error (noise) — *n* sensors…

General solution is:

Hauk (2004), Neuroimage

$$\hat{\mathbf{s}} = \mathbf{C}^{(s)}\mathbf{L}^T(\mathbf{L}\mathbf{C}^{(s)}\mathbf{L}^T + \lambda\mathbf{C}^{(e)})^{-1}\mathbf{d}$$

λ = Regularisation (hyperparameter)

But how calculate $\mathbf{C}^{(e)}$ and $\mathbf{C}^{(s)}$?

Specify multiple (covariance) priors, and estimate their weighting (hyperparameters) by maximising **model evidence**
(using a variational Bayesian approach, eg EM algorithm)

Examples

1. EEG -> fMRI asymmetric integration
2. fMRI -> M/EEG asymmetric integration
3. MEG <-> EEG symmetric integration (fusion)

Asymmetric Integration of MEG+fMRI Background

- fMRI has superior spatial resolution (~mm) than M/EEG, but inferior temporal resolution (integrating over seconds)

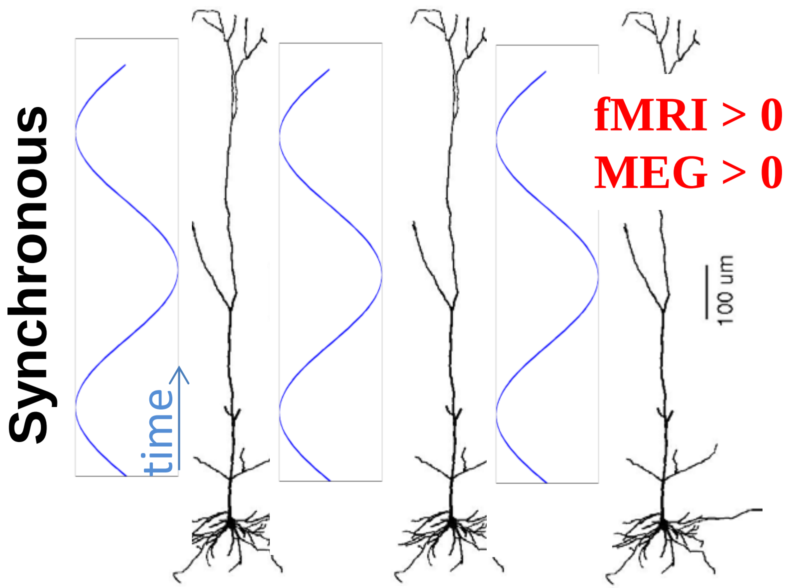
- fMRI and M/EEG measure similar, but not identical, neural activity

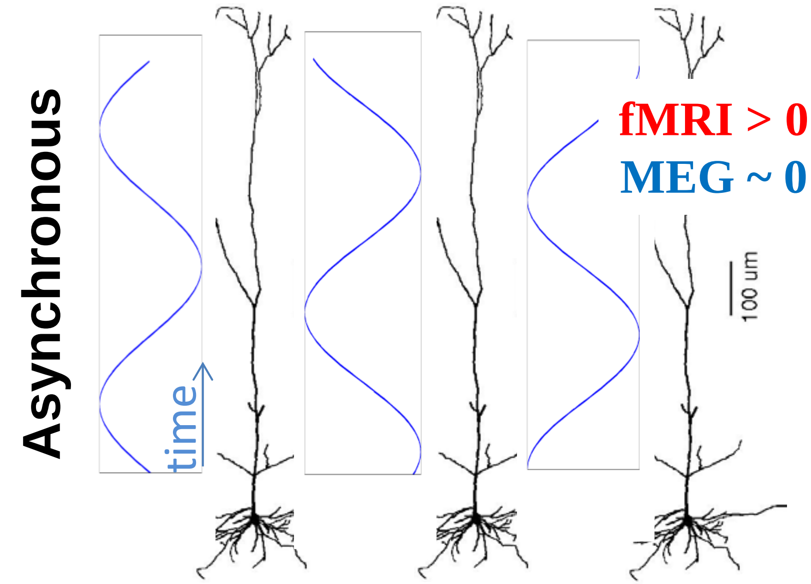
- Eg, some source configurations give detectable fMRI signal, but no detectable M/EEG signal…

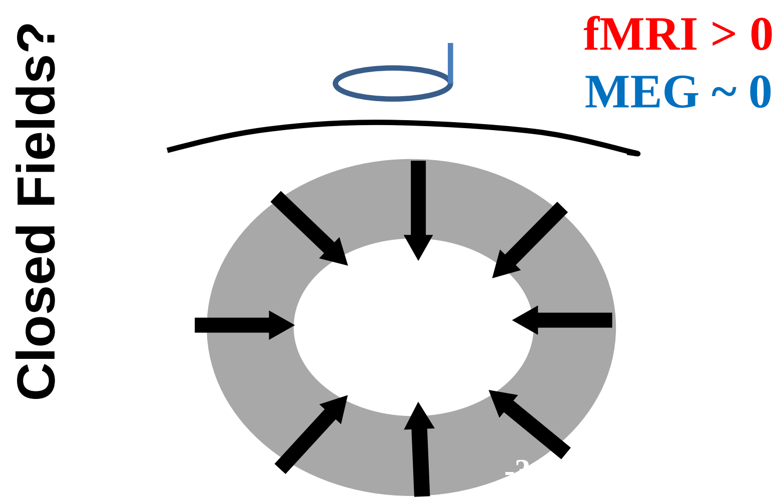
- …and vice versa

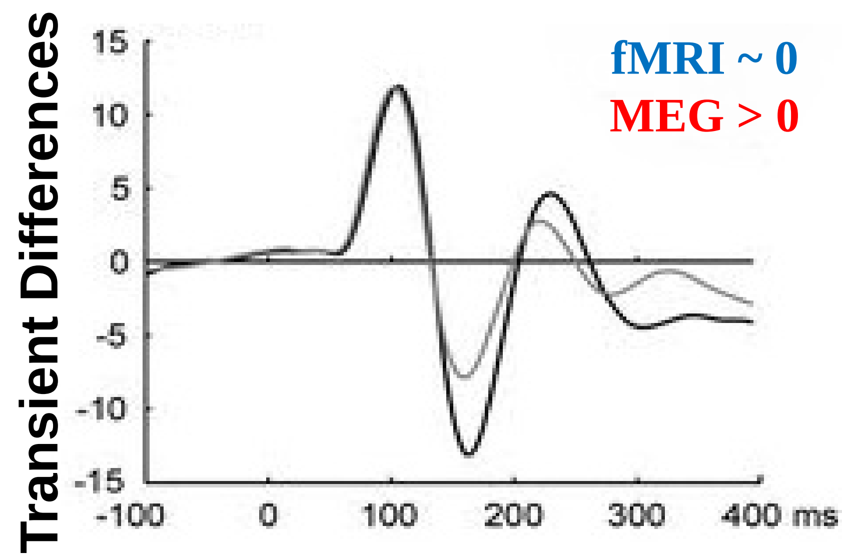
Asymmetric Integration of MEG+fMRI
Background

MRC | Cognition and Brain Sciences Unit









Asymmetric Integration of MEG+fMRI Background

MRC | Cognition and Brain Sciences Unit

- fMRI has superior spatial resolution (~mm) than M/EEG, but inferior temporal resolution (integrating over seconds)
- fMRI and M/EEG measure similar, but not identical, neural activity
- Eg, some source configurations give detectable fMRI signal, but no detectable M/EEG signal…
- …and vice versa
- Use fMRI as a soft, rather than hard, constraint on localisation of sources of M/EEG data…

Asymmetric Integration of MEG+fMRI

MRC | Cognition and Brain Sciences Unit

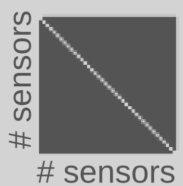
Specifying (co)variance components (priors/regularisation):

$$\mathbf{C} = \sum_i \lambda_i \mathbf{Q}_i$$

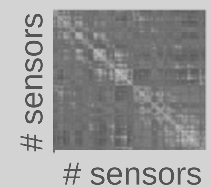
C = Sensor/Source covariance
Q = Covariance components
λ = Hyper-parameters

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

"IID" (white noise):

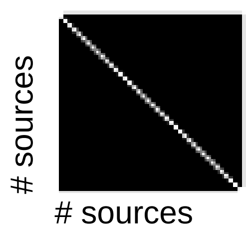


Empty-room:

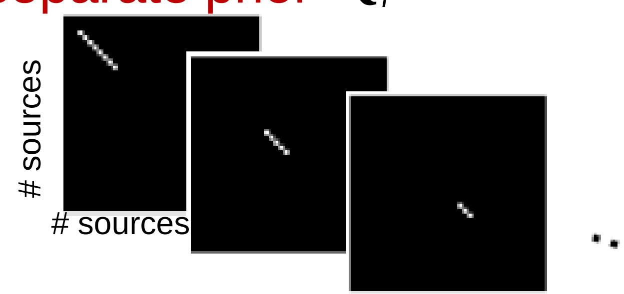


2. Each suprathreshold fMRI cluster becomes a separate prior $\mathbf{Q}_i^{(s)}$

"IID" (min norm):



fMRI Priors:



Henson et al (2010) Hum. Brain Map.

Asymmetric Integration of MEG+fMRI

MRC Cognition and Brain Sciences Unit

General solution again:

$$\hat{\mathbf{s}} = \mathbf{C}^{(s)}\mathbf{L}^T(\mathbf{L}\mathbf{C}^{(s)}\mathbf{L}^T + \lambda\mathbf{C}^{(e)})^{-1}\mathbf{d}$$

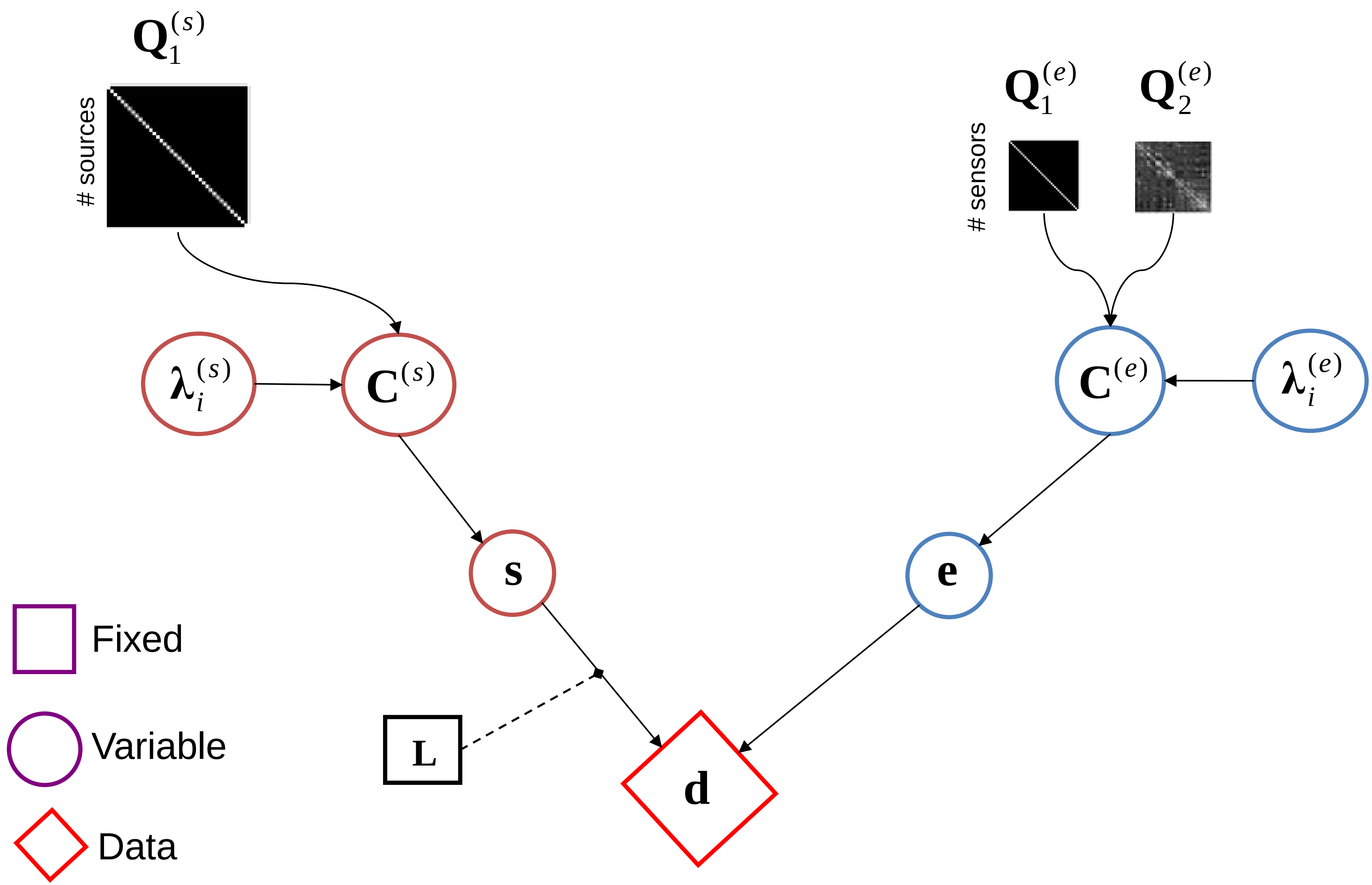
$$\mathbf{e} \sim N(0, \mathbf{C}^{(e)})$$
$$\mathbf{s} \sim N(0, \mathbf{C}^{(s)})$$

Now source covariance expressed as number of fMRI clusters:

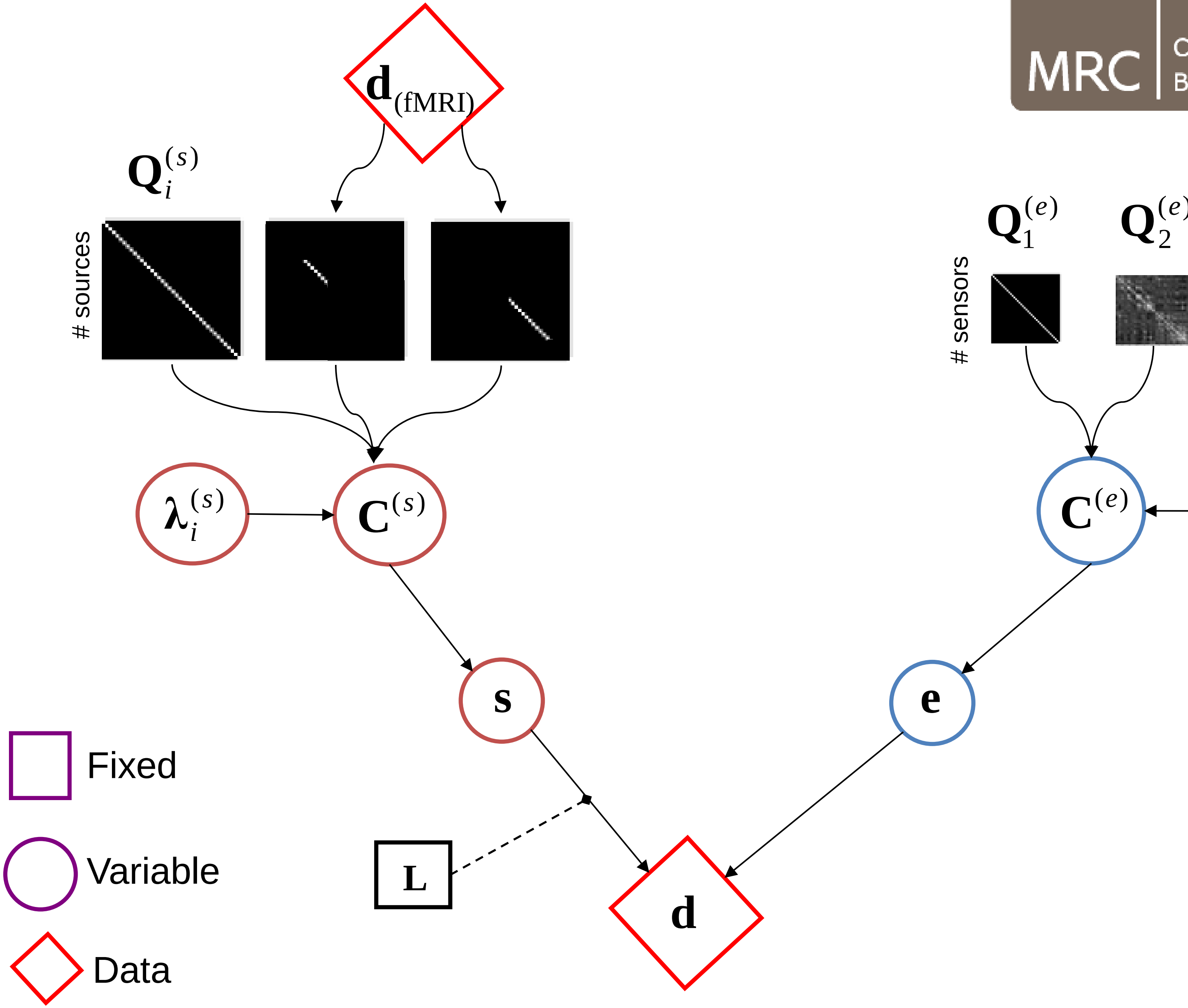
$$\mathbf{C}^{(s)} = \lambda_1^{(s)}\mathbf{Q}^{(s)}_{(fMRI\,1)} + \lambda_2^{(s)}\mathbf{Q}^{(s)}_{(fMRI\,2)} + \ldots$$

When $\mathbf{Q}_i^{(s)}$ does not help maximise model evidence, $\lambda_i^{(s)} \rightarrow 0$, i.e, constraints ignored…

…catering for situations where fMRI signal does not reflect same activity as in M/EEG signal (e.g, occurring later than time-window than M/EEG data)



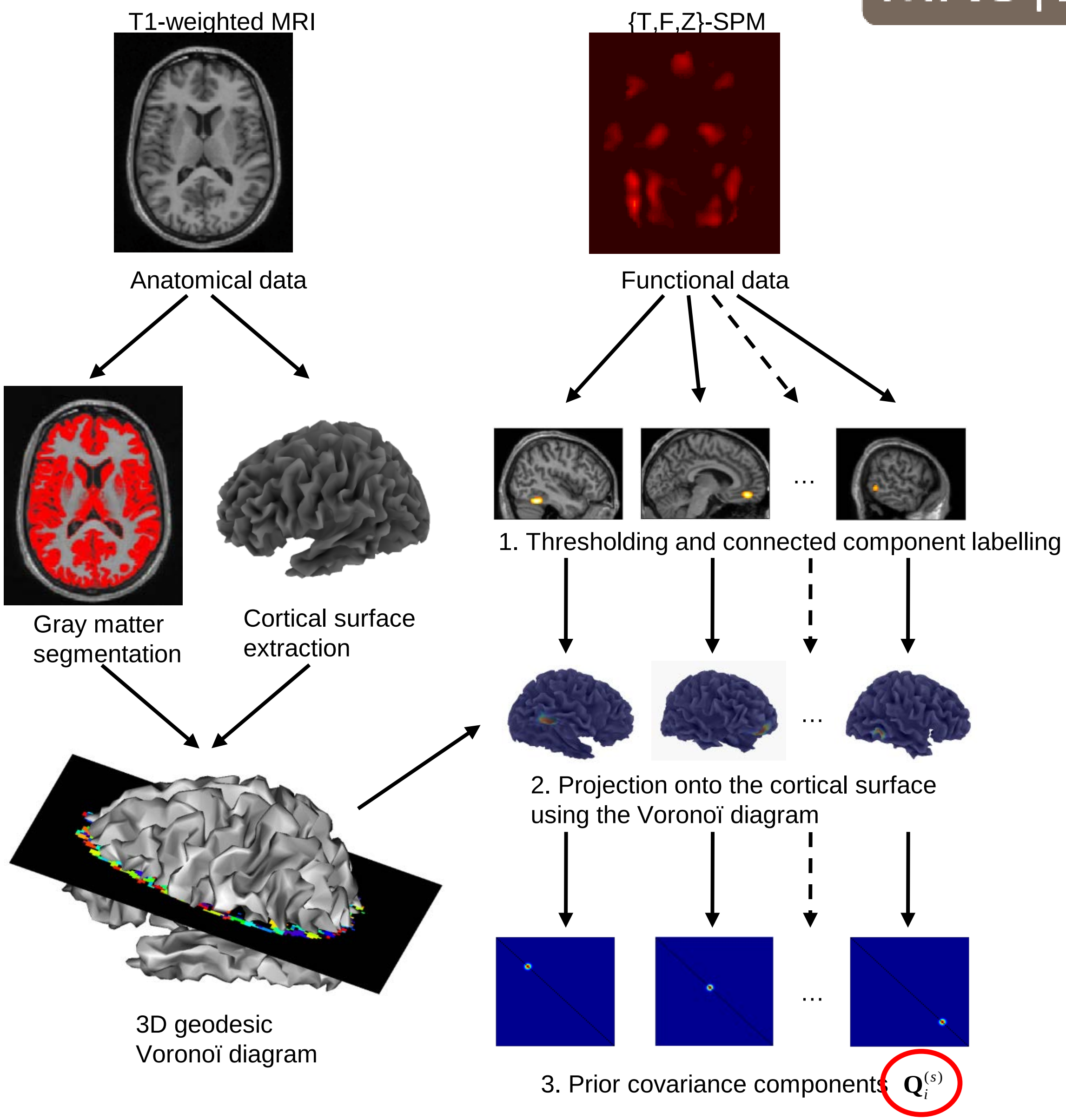
Henson et al (2010) Hum Brain Map



MRC | Cognition and Brain Sciences Unit

Henson et al (2010) Hum Brain Map

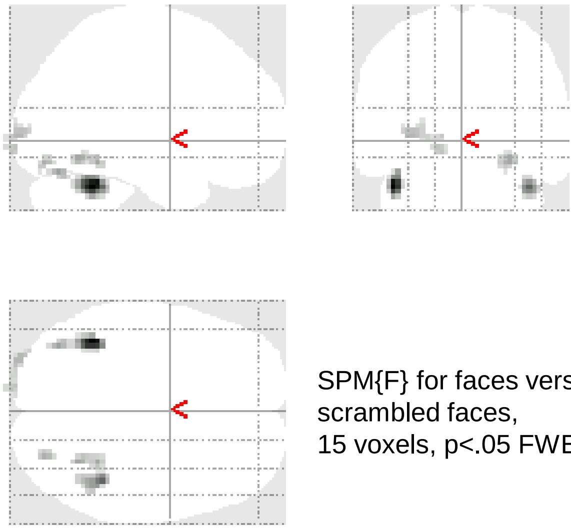
Asymmetric Integration of M/EEG+fMRI



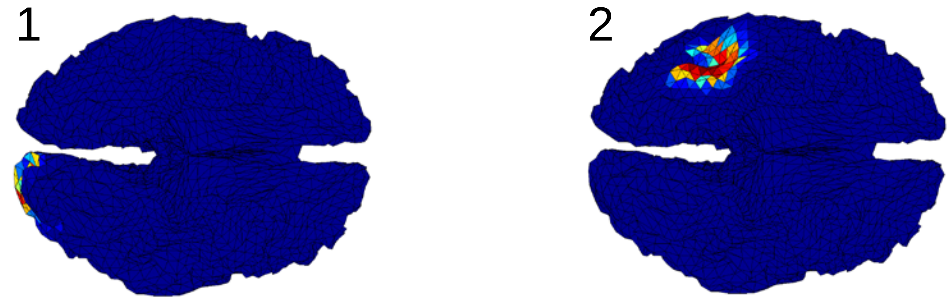
Henson et al (2010) Hum. Brain Map.

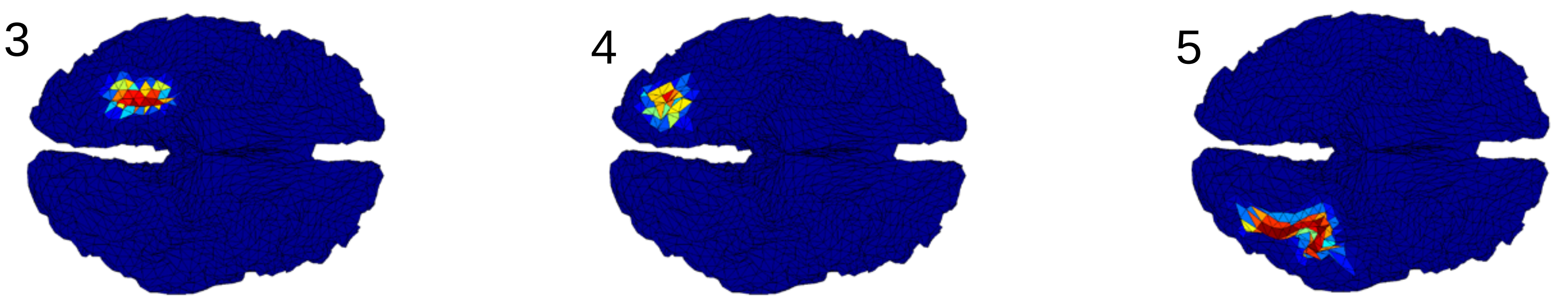
Asymmetric Integration of M/EEG+fMRI

MRC | Cognition and Brain Sciences Unit



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



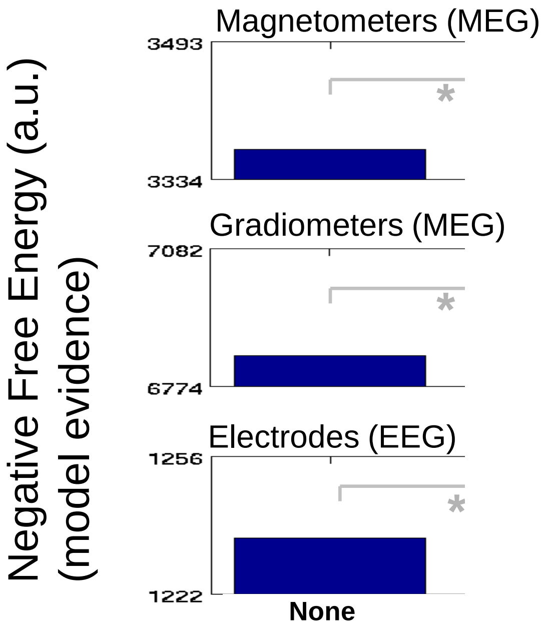


5 clusters from SPM of fMRI data from separate group of (18) subjects in MNI space

Henson et al (2010) Hum. Brain Map.

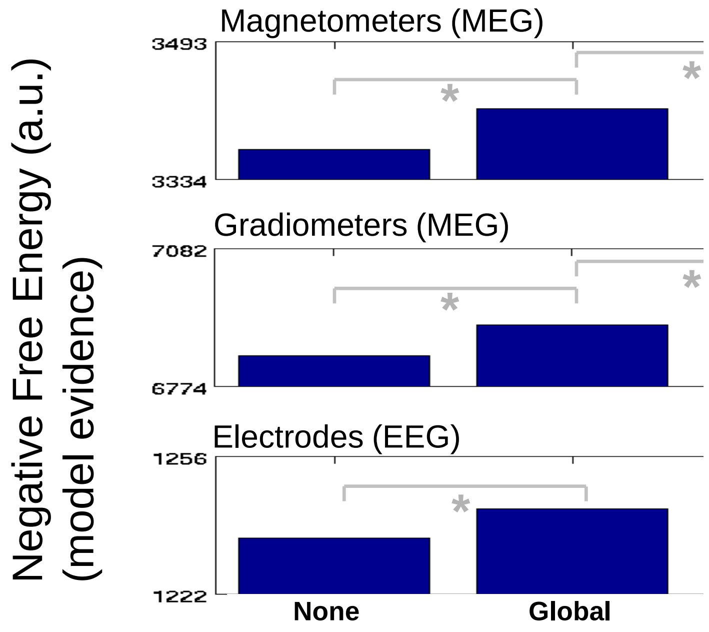
Asymmetric Integration of M/EEG+fMRI

MRC Cognition and Brain Sciences Unit



Henson et al (2010) Hum. Brain Map.

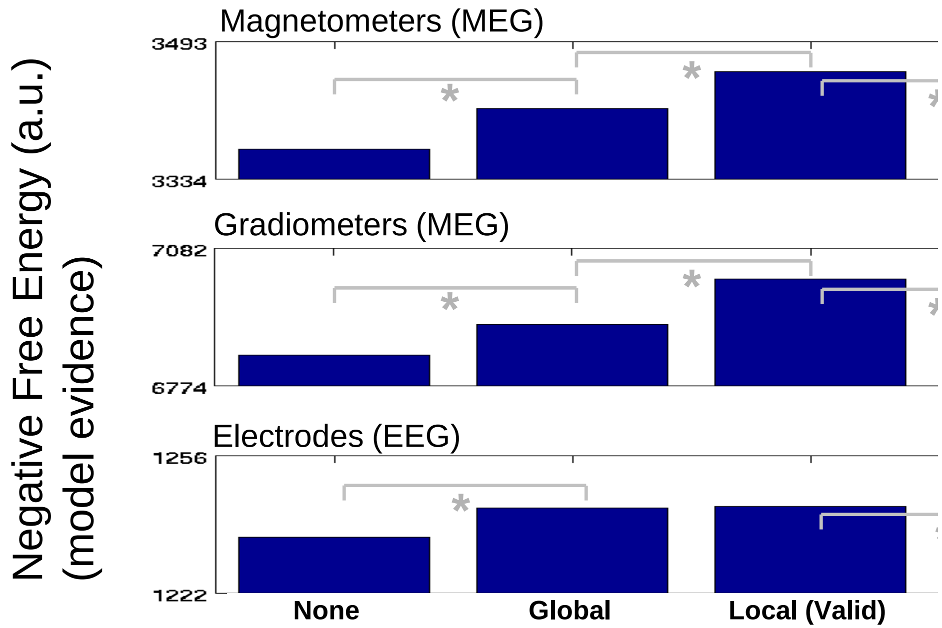
Asymmetric Integration of M/EEG+fMRI



Henson et al (2010) Hum. Brain Map.

Asymmetric Integration of M/EEG+fMRI

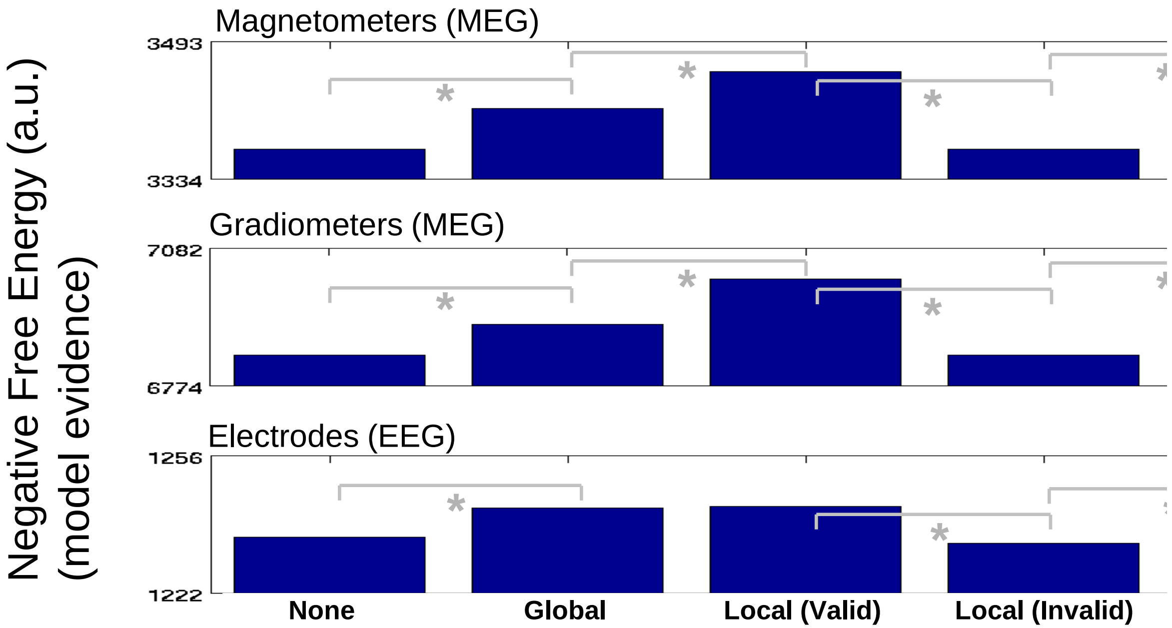
MRC | Cognition and Brain Sciences Unit



Henson et al (2010) Hum. Brain Map.

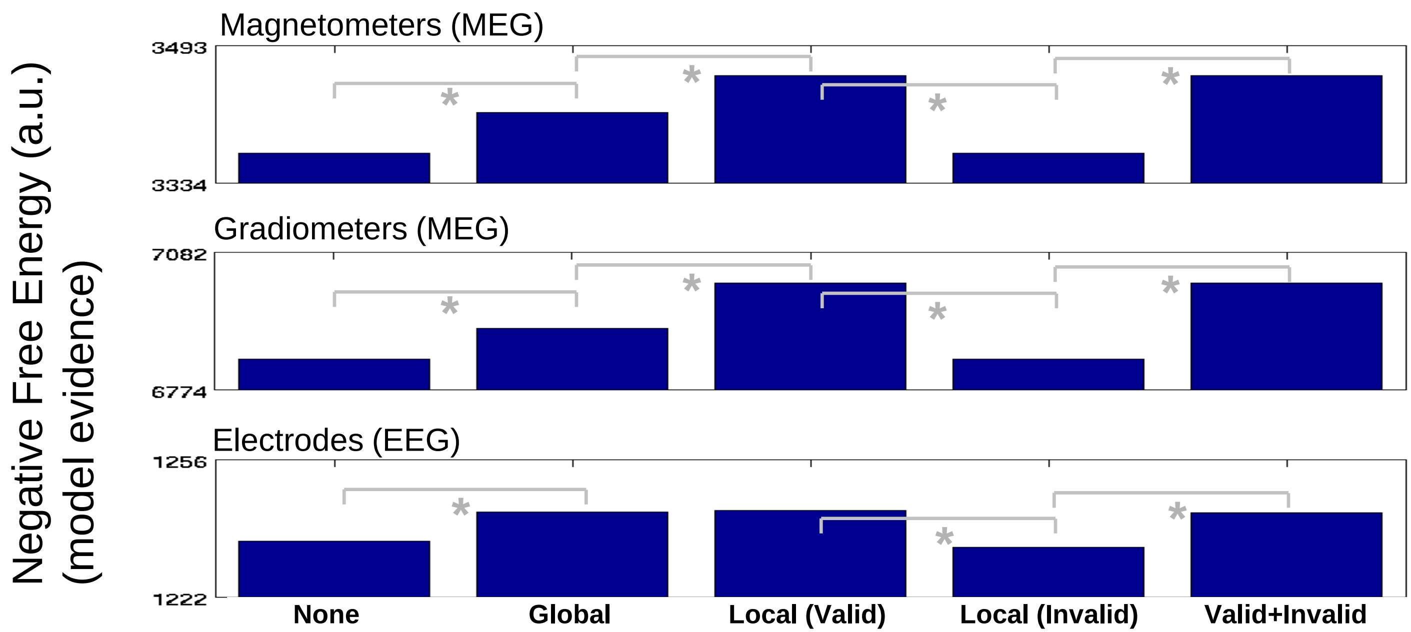
Asymmetric Integration of M/EEG+fMRI

MRC Cognition and Brain Sciences Unit



Henson et al (2010) Hum. Brain Map.

Asymmetric Integration of M/EEG+fMRI

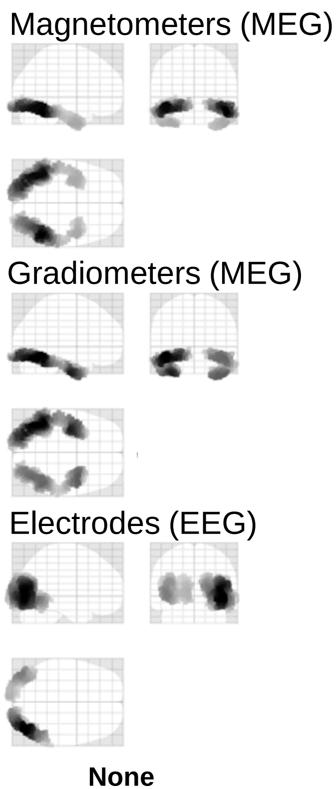


Henson et al (2010) Hum. Brain Map.

Asymmetric Integration of M/EEG+fMRI

MRC | Cognition and Brain Sciences Unit

IID sources and IID noise (L2 MNM)

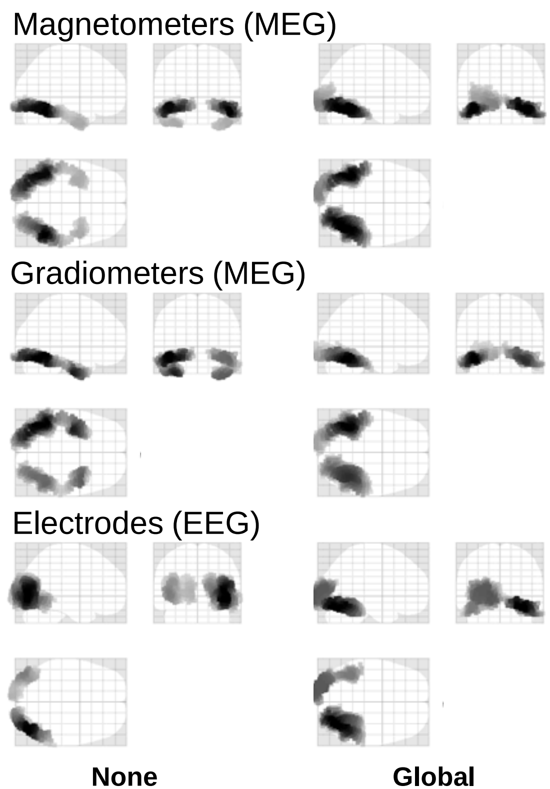


Henson et al (2010) Hum. Brain Map.

Asymmetric Integration of M/EEG+fMRI

MRC Cognition and Brain Sciences Unit

IID sources and IID noise (L2 MNM)

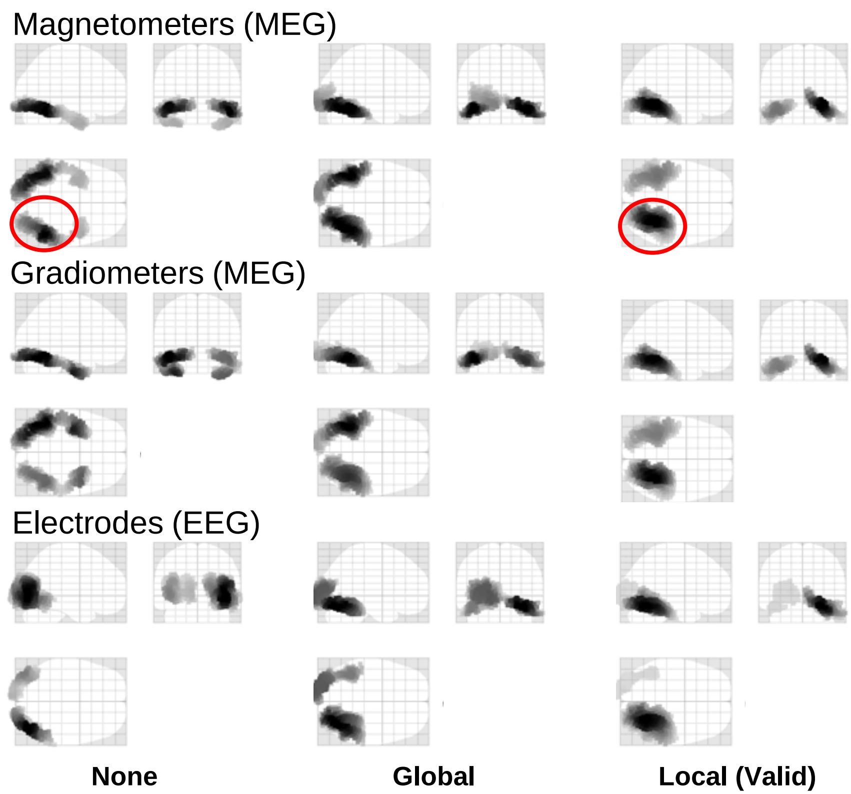


Henson et al (2010) Hum. Brain Map.

Asymmetric Integration of M/EEG+fMRI

MRC | Cognition and Brain Sciences Unit

IID sources and IID noise (L2 MNM)



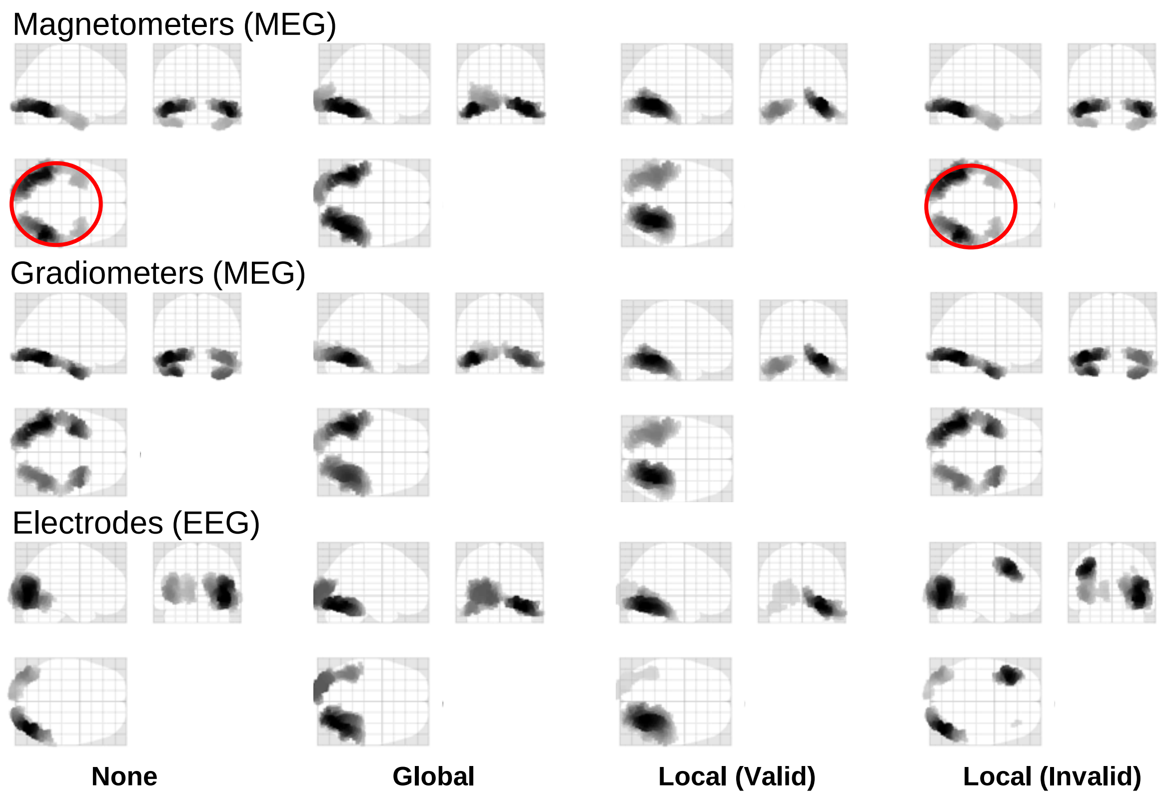
fMRI priors counteract superficial bias of Min Norm

Henson et al (2010) Hum. Brain Map.

Asymmetric Integration of M/EEG+fMRI

MRC | Cognition and Brain Sciences Unit

IID sources and IID noise (L2 MNM)



Invalid priors generally discounted (at least for MEG)

Henson et al (2010) Hum. Brain Map.

Asymmetric Integration of M/EEG+fMRI

- Adding a single, global fMRI prior increases model evidence
- Adding multiple valid priors increases model evidence further
- Adding invalid priors does not necessarily increase model evidence, particularly in conjunction with valid priors
  Helpful if some fMRI regions produce no MEG/EEG signal (or arise from neural activity at different times)
- Can counteract superficial bias of, e.g, minimum-norm
- Makes some allowance for different sensitivities of fMRI and M/EEG to certain types of neural activity

Henson et al (2010) Hum. Brain Map.

Examples

MRC Cognition and Brain Sciences Unit

1. EEG -> fMRI asymmetric integration
2. fMRI -> M/EEG asymmetric integration
3. MEG <-> EEG symmetric integration (fusion)

Symmetric Integration of MEG+EEG
Background

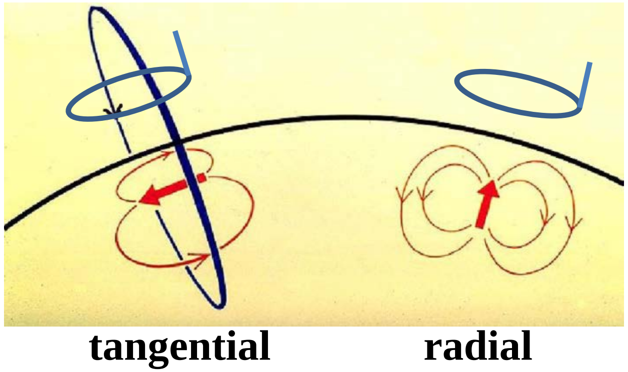
MRC | Cognition and Brain Sciences Unit

- MEG generally has superior spatial resolution than EEG (less blurred by skull/scalp)

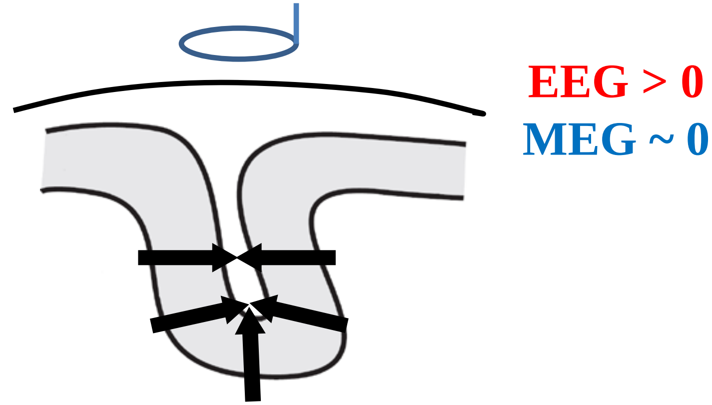
- MEG cannot detect radial component of current sources; EEG can!

Symmetric Integration of MEG+EEG Background

Dipolar Sources

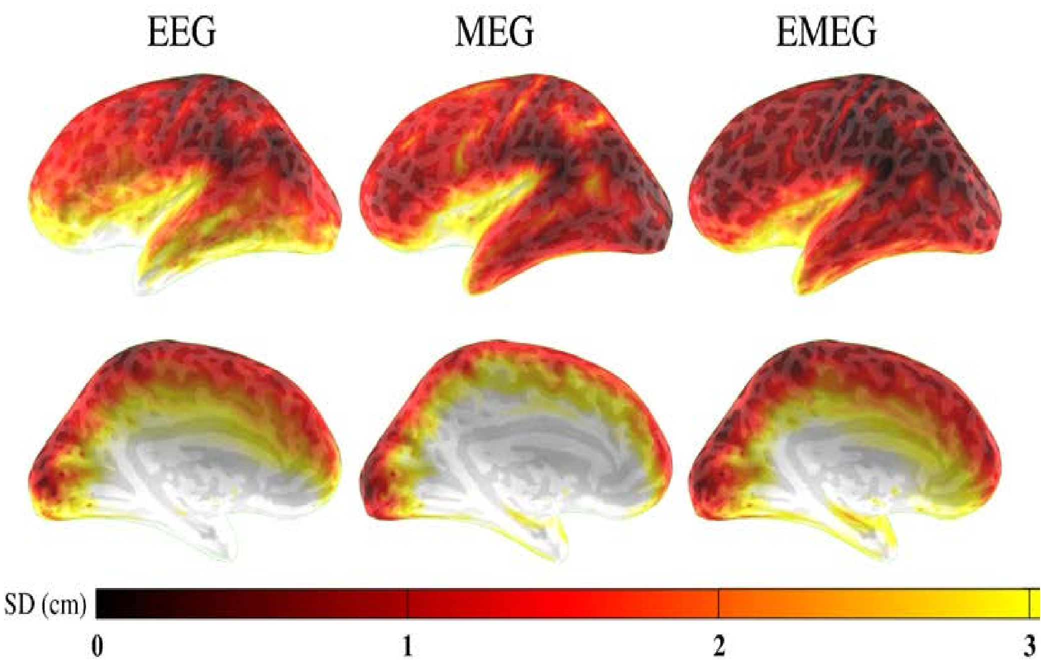


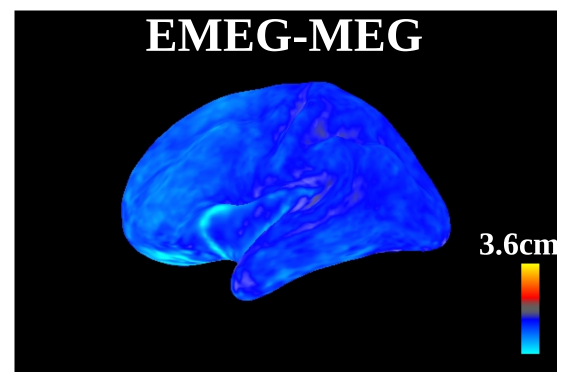
Extended Sources



Ahlfors et al., HBM 2010

Spatial Extent





Stenroos & Hauk, in prep

Molins et al., Neuroimage 2008

Symmetric Integration of MEG+EEG Background

MRC | Cognition and Brain Sciences Unit

- MEG generally has superior spatial resolution than EEG (less blurred by skull/scalp)
- MEG cannot detect radial component of current sources; EEG can!
- And few practical problems acquiring concurrent EEG (apart from extra time attaching electrodes)
- …but EEG data is more sensitive to head geometry and conductivity (potentially biasing any joint-localisation)…
- …and has different noise characteristics…

Symmetric Integration of MEG+EEG Generative Model

For fusing MEG and EEG, we can simply concatenate the MEG+EEG data:

$$\begin{bmatrix} \mathbf{d}_{(MEG)} \\ \mathbf{d}_{(EEG)} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{(MEG)} \\ \mathbf{L}_{(EEG)} \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{e}_{(MEG)} \\ \mathbf{e}_{(EEG)} \end{bmatrix}$$

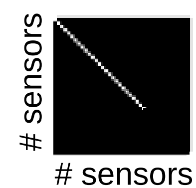
$$\mathbf{e} \sim N(0, \mathbf{C}^{(e)})$$

$$\mathbf{s} \sim N(0, \mathbf{C}^{(s)})$$

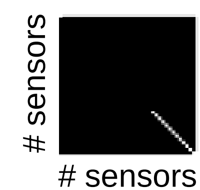
Noise expressed by MEG and EEG terms (e.g, white noise):

$$\hat{\mathbf{C}}^{(e)} = \lambda_1^{(e)} \mathbf{Q}^{(e)}_{(MEG)} + \lambda_2^{(e)} \mathbf{Q}^{(e)}_{(EEG)}$$

$$\mathbf{Q}^{(e)}_{(MEG)} =$$

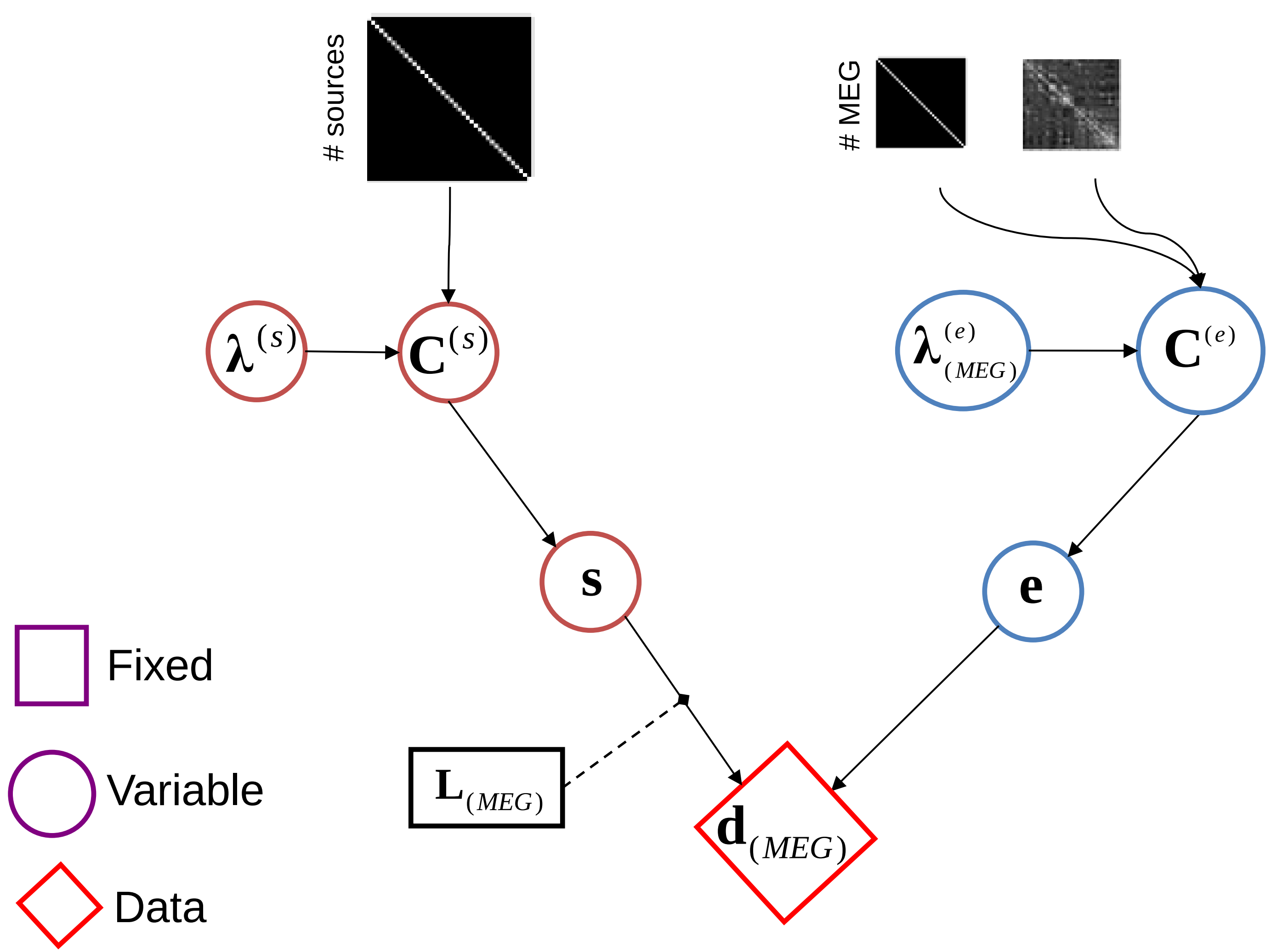


$$\mathbf{Q}^{(e)}_{(EEG)} =$$



The separate hyperparameters allow for different noise levels (SNR)

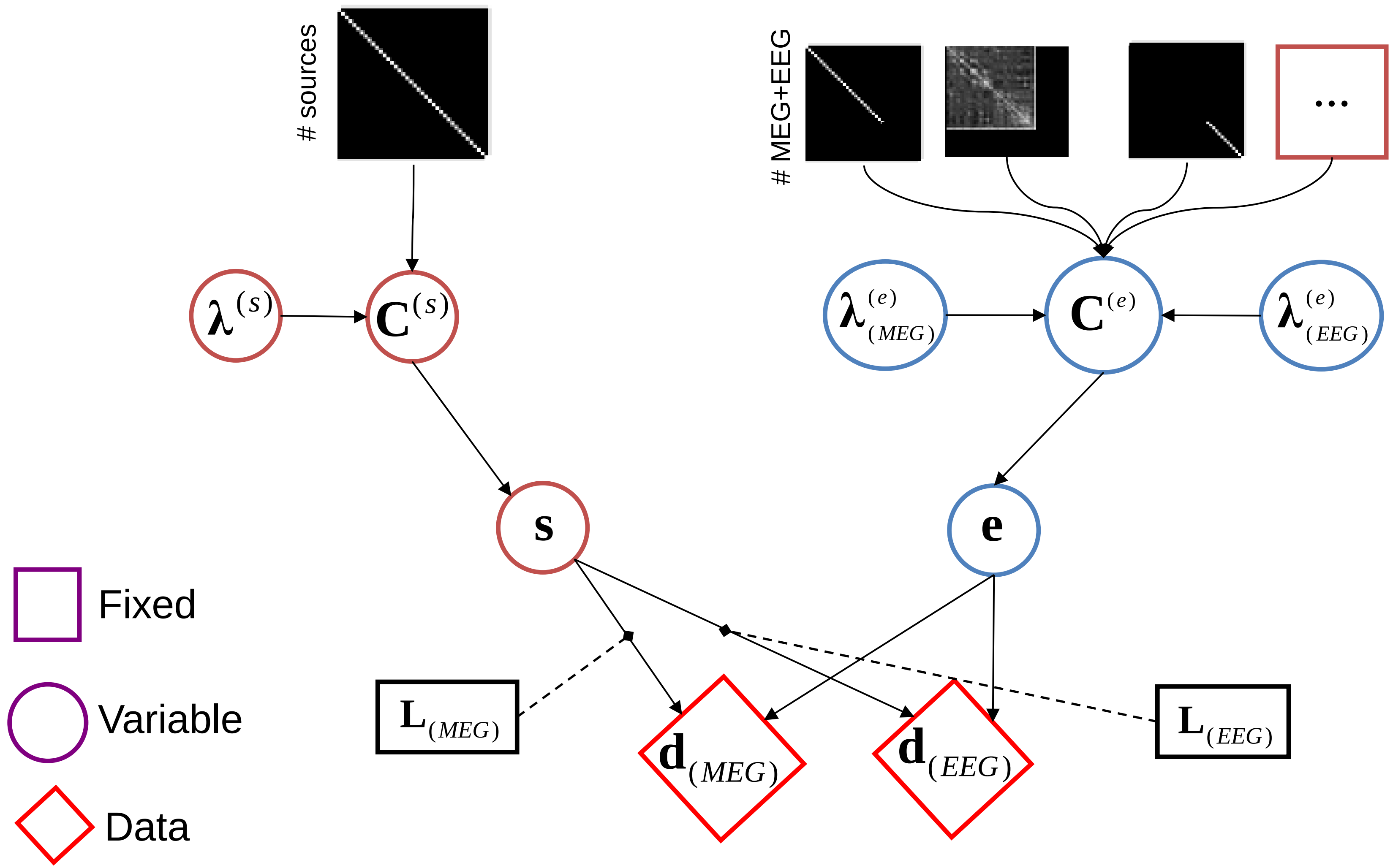
Symmetric Integration of MEG+EEG Generative Model



Henson et al (2009) Neuroimage

Symmetric Integration of MEG+EEG Generative Model

MRC Cognition and Brain Sciences Unit



Henson et al (2009) Neuroimage

One final problem…

MRC | Cognition and Brain Sciences Unit

- Though this allows for different additive noise levels in MEG and EEG…
- …we are assuming mapping from common electrical sources to sensor values (in terms of Telsa and Volts) is known precisely…
- …whereas in reality, this depends on several unknowns (e.g, precise conductivity of skull/scalp)
- One solution is to scale data/leadfields to have same variance:

$$\tilde{Y}_i = \frac{Y_i}{\sqrt{\frac{1}{n_i} tr(Y_i Y_i^T)}} \qquad \tilde{L}_i = \frac{L_i}{\sqrt{\frac{1}{n_i} tr(L_i L_i^T)}}$$

i = ith modality, ie MEG or EEG

n_i = Number of sensors for modality i

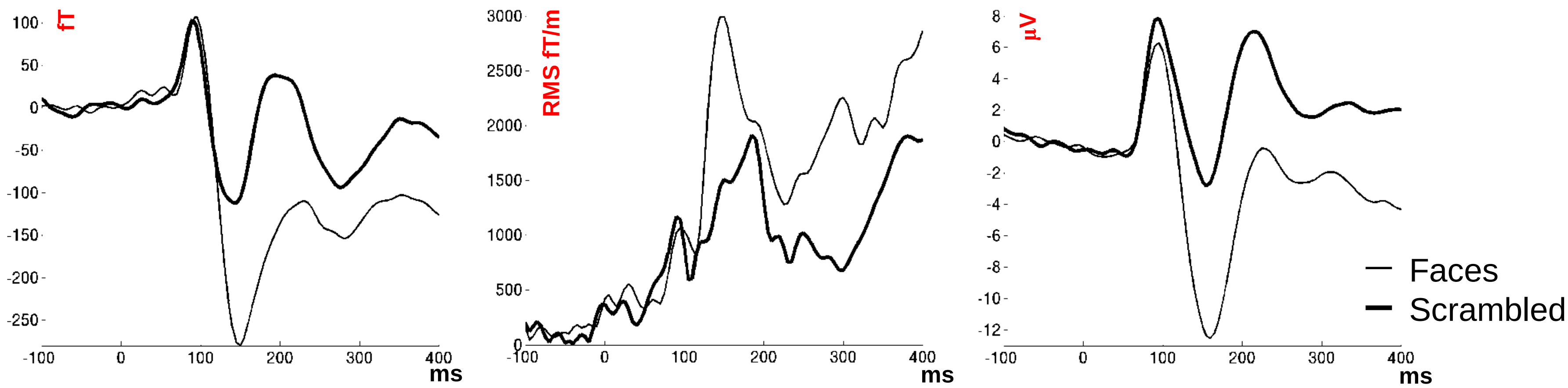
Henson et al (2009) Neuroimage

Symmetric Integration of MEG+EEG Example

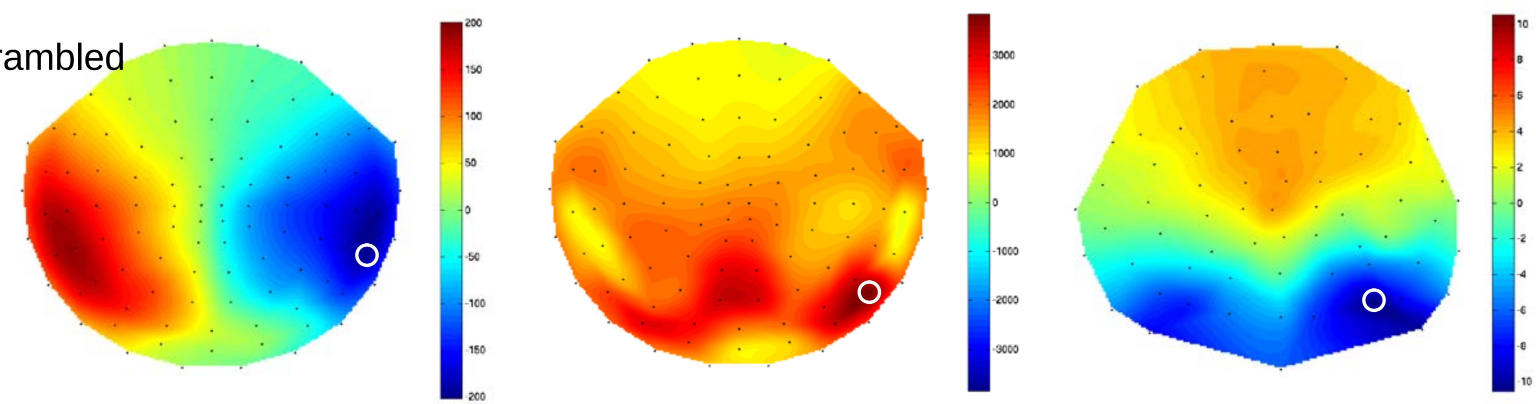
MRC Cognition and Brain Sciences Unit

ERs from 12 subjects for 3 simultaneously-acquired Neuromag sensor-types:

Magnetometers (MEG, 102) | (Planar) Gradiometers (MEG, 204) | Electrodes (EEG, 70)



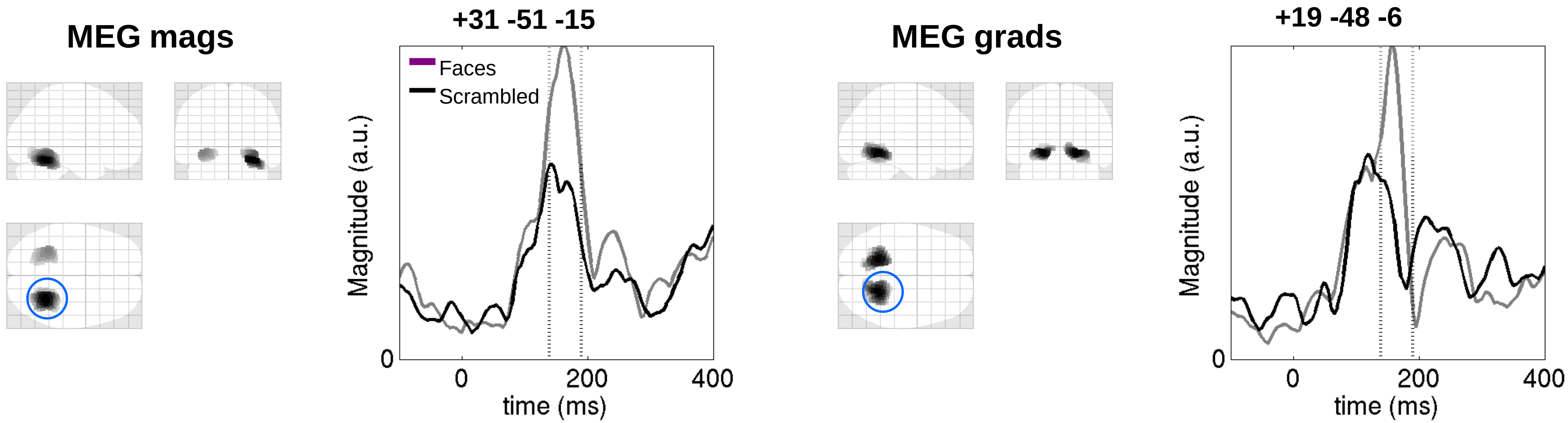
Faces - Scrambled
150-190ms



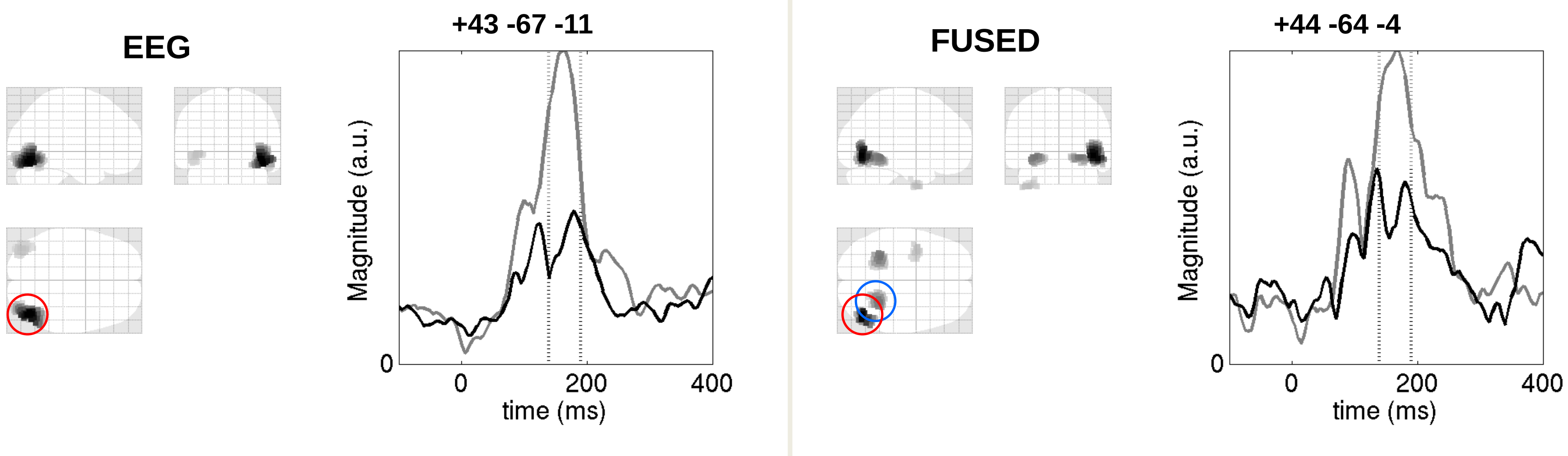
Henson et al (2009) Neuroimage

Symmetric Integration of MEG+EEG

MRC Cognition and Brain Sciences Unit



Faces – Scrambled,150-190ms



IID noise for each modality; common MSP for sources

Henson et al (2009) Neuroimage

Other Approaches to M/EEG fusion

MRC Cognition and Brain Sciences Unit

- Estimate noise covariance from pre-stimulus baseline (**b**):

$$\mathbf{C}^{(e)} = \begin{bmatrix} cov(\mathbf{b}_{(MEG)}) & \mathbf{0} \\ \mathbf{0} & cov(\mathbf{b}_{(EEG)}) \end{bmatrix}$$

Molins et al (2008), Neuroimage

(which can also be used to pre-whiten data and leadfields, scaling to noise units)…
…but downside is that **baseline contains source activity**, so not estimate of true sensor noise

- Maximise mutual information between MEG and EEG

Baillet et al (1999), IEEE

- Re-parameterise leadfields in terms of radial/tangential components

Huang et al (2007), Neuroimage

Examples

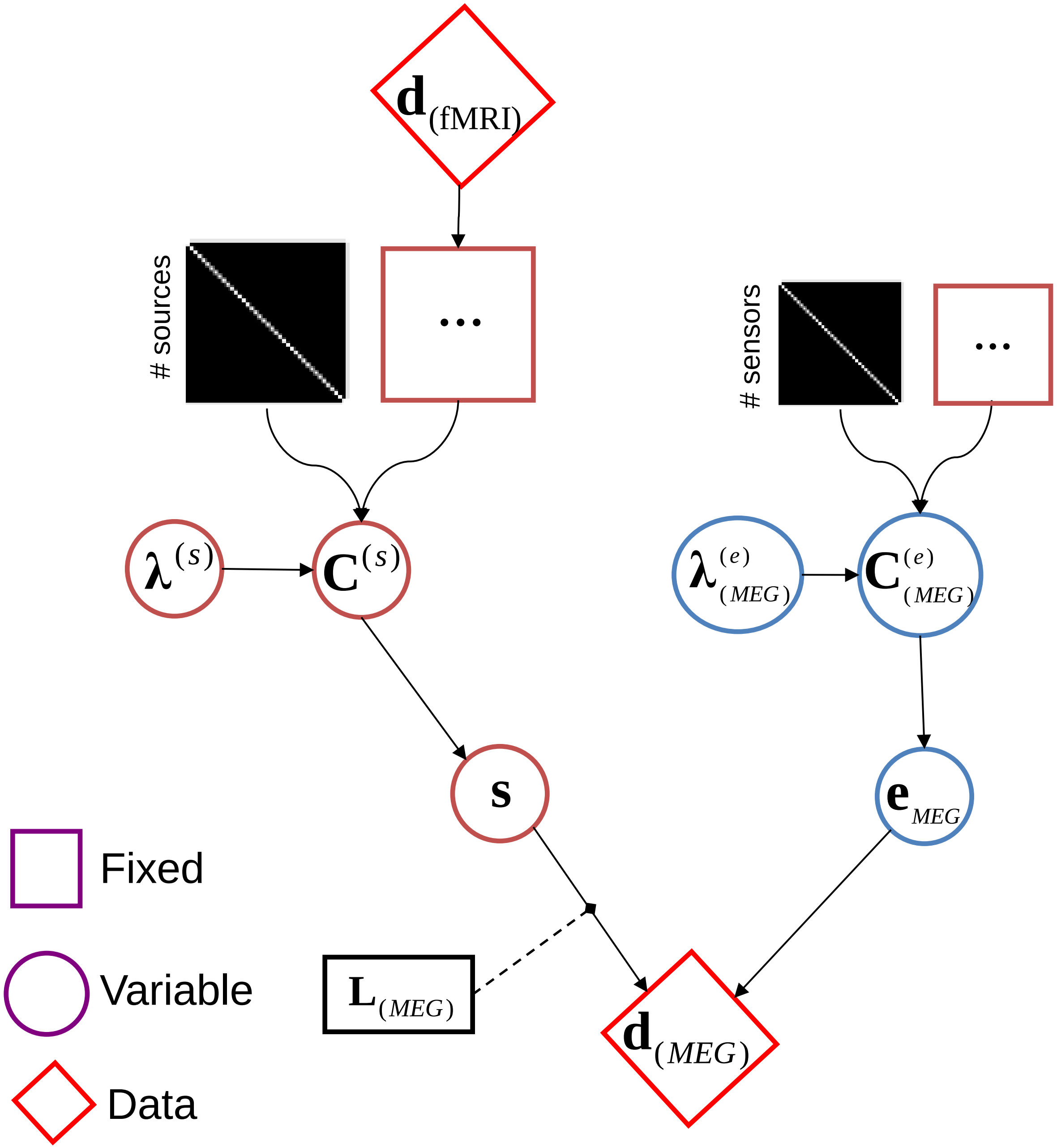
MRC | Cognition and Brain Sciences Unit

1. EEG -> fMRI asymmetric integration

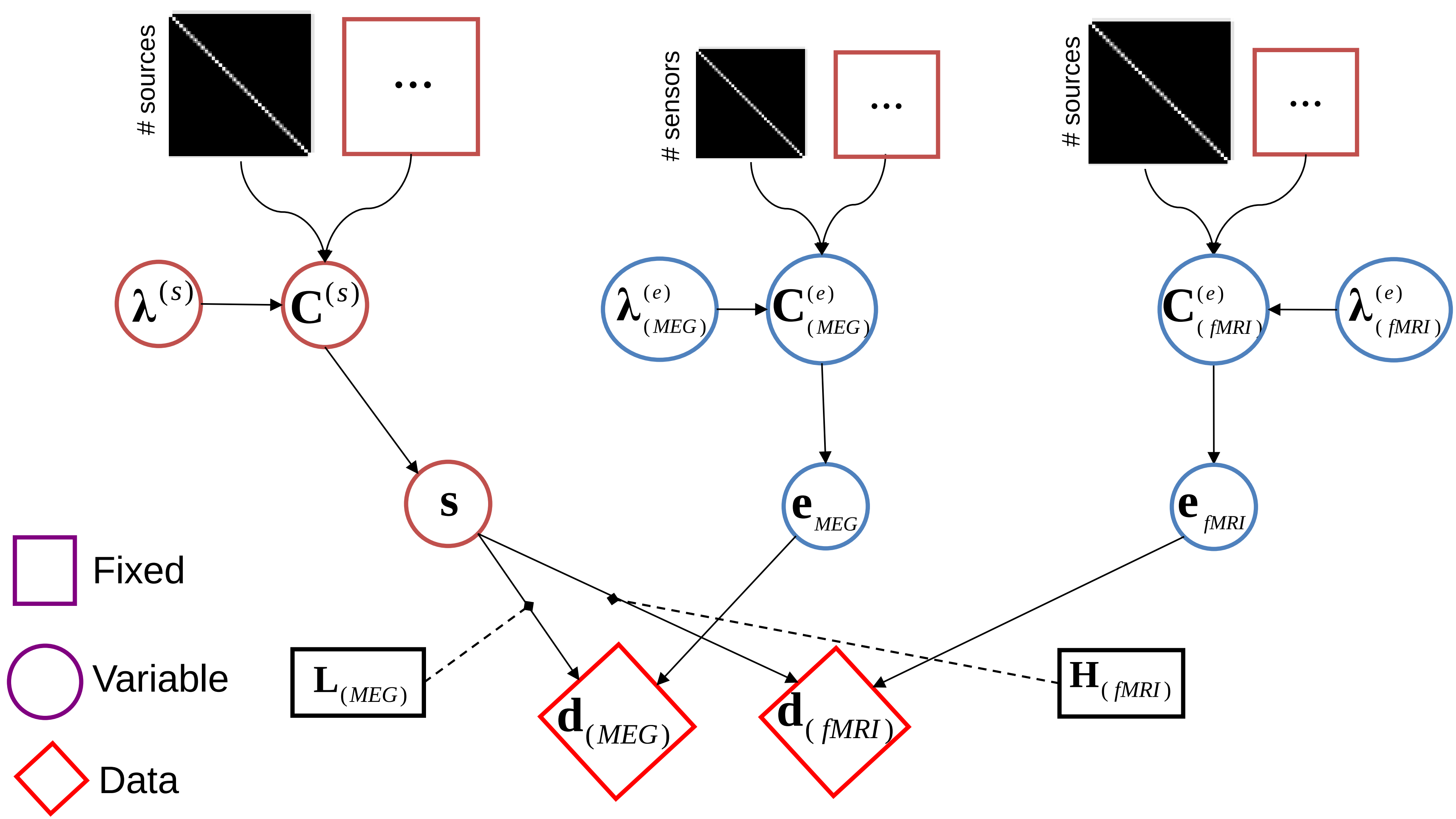
2. fMRI -> M/EEG asymmetric integration

3. MEG <-> EEG symmetric integration (fusion)

4. fMRI <-> MEG <-> EEG fusion?



Henson et al (2011) Frontiers



Henson et al (2011) Frontiers

Summary

MRC | Cognition and Brain Sciences Unit

Data-driven, symmetric approaches:

- Linked Matrix Factorisation methods (ICA, CCA, PLS)
- Representational Similarity Analysis (RSA)
- Graph Theory

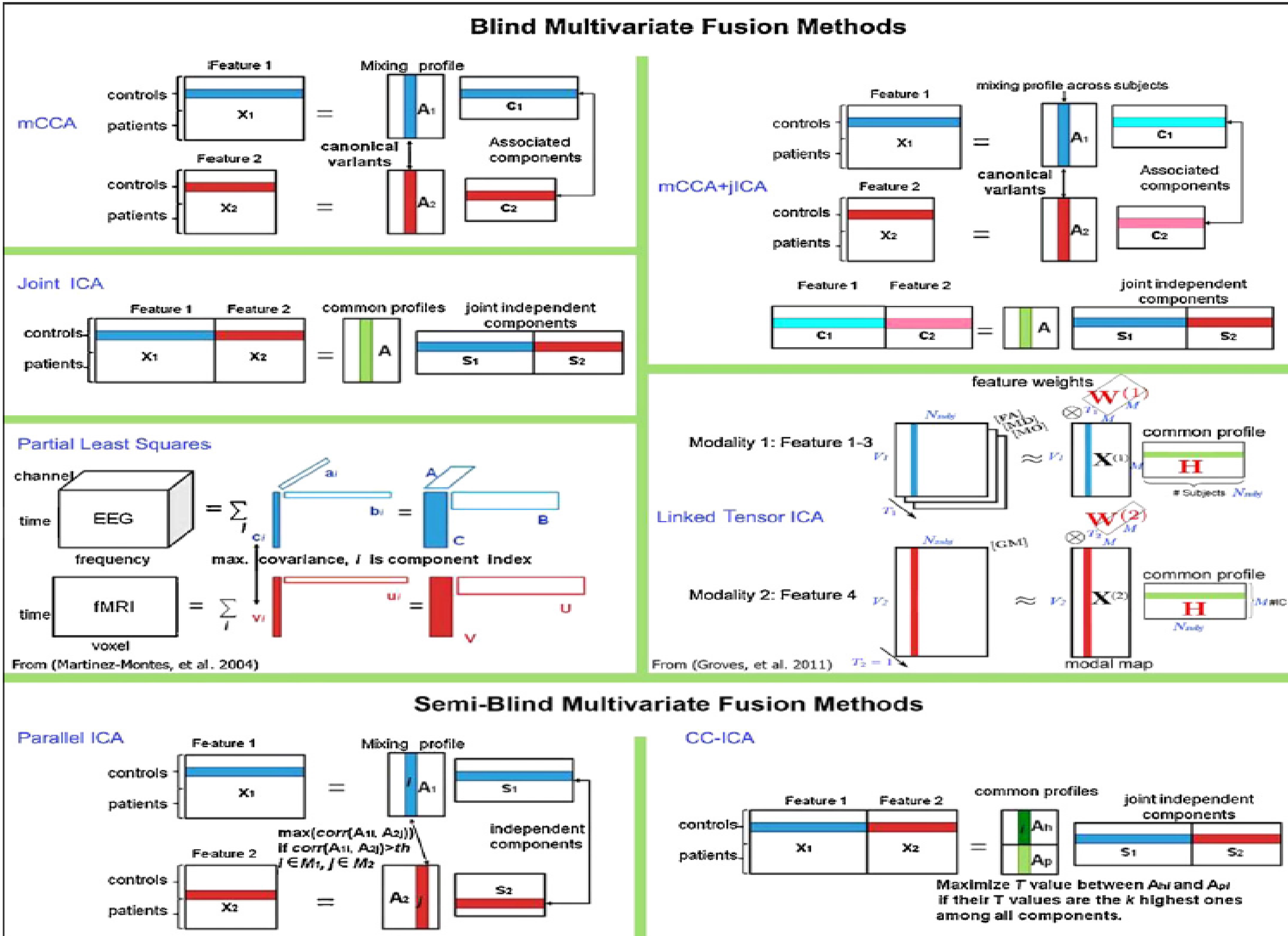
Model-driven, asymmetric approaches:

- Regession / Mediation / SEM

Model-driven, symmetric approaches:

- Generative models (eg PEB)

The End



Blind Multivariate Fusion Methods
mCCA
Feature 1
Mixing profile
controls
patients
X1
A1
C1
canonical variants
Associated components
Feature 2
X2
A2
C2
mCCA+jICA
mixing profile across subjects
joint independent components
S1
S2
A
Joint ICA
common profiles
Partial Least Squares
channel
time
EEG
frequency
max. covariance, i is component index
fMRI
voxel
From (Martinez-Montes, et al. 2004)
Linked Tensor ICA
feature weights
Modality 1: Feature 1-3
Modality 2: Feature 4
common profile
H
modal map
From (Groves, et al. 2011)
Semi-Blind Multivariate Fusion Methods
Parallel ICA
max(corr(A1i, A2j))
if corr(A1i, A2j)>th
i ∈ M1, j ∈ M2
independent components
CC-ICA
Maximize T value between Ahi and Api
if their T values are the k highest ones
among all components.

Inverse Problem: Standard L2-norm

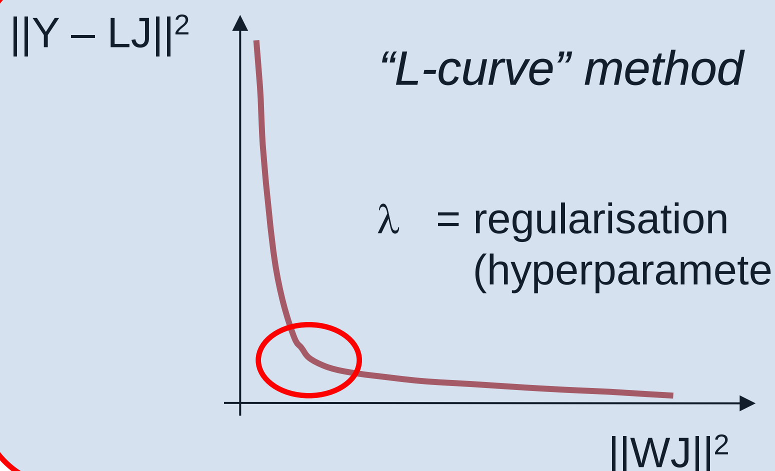
MRC Cognition and Brain Sciences Unit

$$\mathbf{Y} = \mathbf{LJ} + \mathbf{E} \qquad \mathbf{E} \sim N(\mathbf{0}, \mathbf{C}^{(e)})$$

$$\mathbf{J} = \arg\min\{\left\| \mathbf{C}^{(e)^{-1/2}}(\mathbf{Y} - \mathbf{LJ}) \right\|^2 + \lambda \left\| \mathbf{WJ} \right\|^2\}$$

$$= (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{L}^T[\mathbf{L}(\mathbf{W}^T\mathbf{W})^{-1}\mathbf{L}^T + \lambda\mathbf{C}^{(e)}]^{-1}\mathbf{Y}$$

"Tikhonov Solution"



"L-curve" method

λ = regularisation (hyperparameter)

$\mathbf{W} = \mathbf{I}$	"Minimum Norm"
$\mathbf{W} = \mathbf{DD}^T$	"Loreta" (**D**=Laplacian)
$\mathbf{W} = diag(\mathbf{L}^T\mathbf{L})^{-1}$	"Depth-Weighted"
$\mathbf{W}_p = diag(\mathbf{L}_p^T C_y^{-1} \mathbf{L}_p)^{-1}$	"Beamformer"
$\mathbf{W} = \ldots$	

Phillips et al (2002), Neuroimage

Inverse Problem: Equivalent PEB

MRC | Cognition and Brain Sciences Unit

Parametric Empirical Bayesian (PEB) 2-level hierarchical form:

$$\mathbf{Y} = \mathbf{L}\mathbf{J} + \mathbf{E}^{(e)} \qquad \mathbf{E}^{(e)} \sim N(0, \mathbf{C}^{(e)})$$

$$\mathbf{J} = 0 + \mathbf{E}^{(j)} \qquad \mathbf{E}^{(j)} \sim N(0, \mathbf{C}^{(j)})$$

$C^{(e)}$ = $n \times n$ Sensor (error) covariance

$C^{(j)}$ = $p \times p$ Source (prior) covariance

Likelihood:

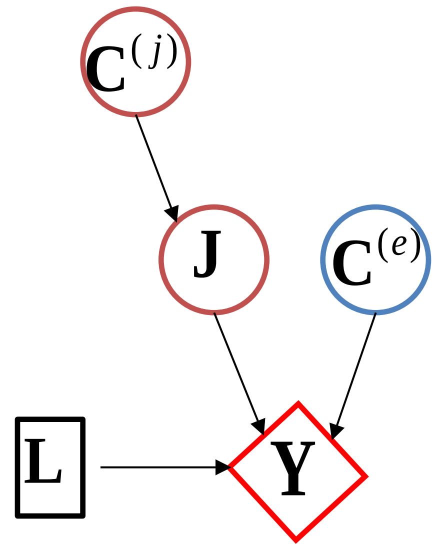
$$p(\mathbf{Y} \mid \mathbf{J}) = N(\mathbf{L}\mathbf{J}, \mathbf{C}^{(e)})$$

Prior:

$$p(\mathbf{J}) = N(0, \mathbf{C}^{(j)})$$

Posterior:

$$p(\mathbf{J} \mid \mathbf{Y}) \propto p(\mathbf{Y} \mid \mathbf{J})\, p(\mathbf{J})$$



Maximum A Posteriori (MAP) estimate:

$$\hat{\mathbf{J}} = \mathbf{C}^{(j)}\mathbf{L}^T[\mathbf{L}\mathbf{C}^{(j)}\mathbf{L}^T + \mathbf{C}^{(e)}]^{-1}\mathbf{Y}$$

cf Classical Tikhonov:

$$(\mathbf{W}^T\mathbf{W})^{-1}\mathbf{L}^T[\mathbf{L}(\mathbf{W}^T\mathbf{W})^{-1}\mathbf{L}^T + \lambda\mathbf{C}^{(e)}]^{-1}\mathbf{Y}$$

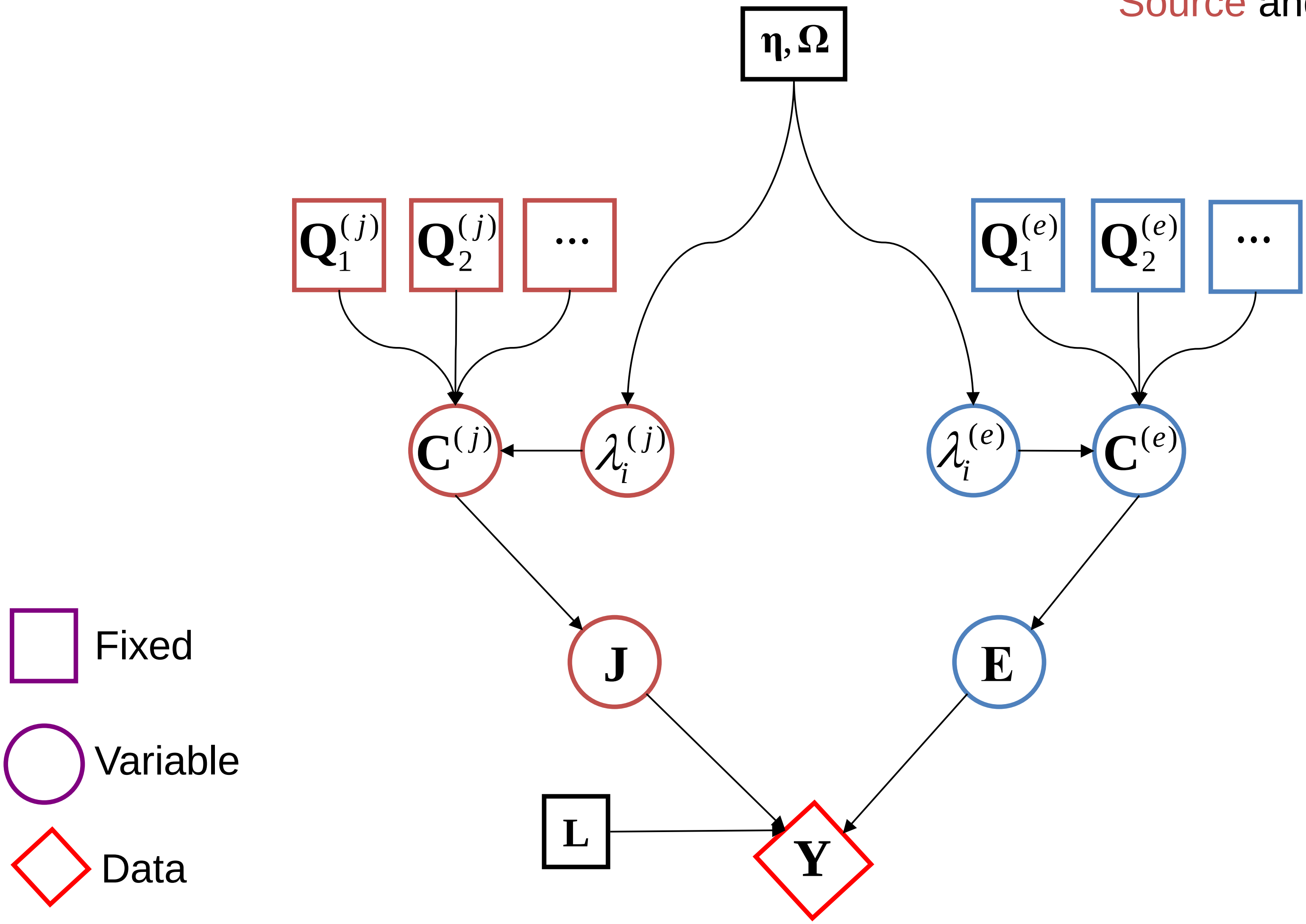
$$\Rightarrow \mathbf{C}^{(j)} = (\mathbf{W}^T\mathbf{W})^{-1}$$

Phillips et al (2005), Neuroimage

PEB: Full Generative Model (DAG)

MRC Cognition and Brain Sciences Unit

Source and sensor space



Friston et al (2008) Neuroimage

PEB: Estimation

MRC | Cognition and Brain Sciences Unit

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*:

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

$$F(\mathbf{Y}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\Sigma}}) \propto -tr(\mathbf{C}^{-1}\mathbf{Y}\mathbf{Y}^T) - \ln|\mathbf{C}| - (\hat{\boldsymbol{\alpha}} - \boldsymbol{\eta})^T \boldsymbol{\Omega}^{-1} (\hat{\boldsymbol{\alpha}} - \boldsymbol{\eta}) + \ln|\hat{\boldsymbol{\Sigma}}\boldsymbol{\Omega}^{-1}|$$

Accuracy Complexity

(…where $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\Sigma}}$ are the posterior mean and covariance of hyperparameters)

Friston et al (2002) Neuroimage

PEB: Multiple Sparse Priors

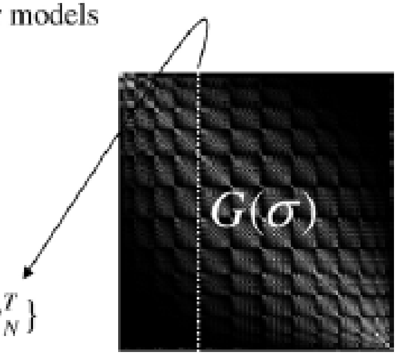
Hyperpriors allow the extreme of 100's source priors, or MSP

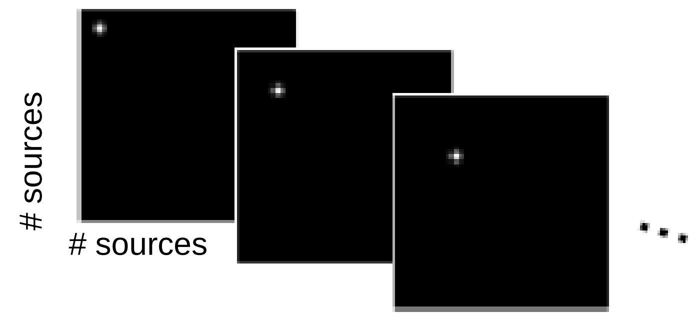
Three prior models

MNM $Q^{\varepsilon} = I$

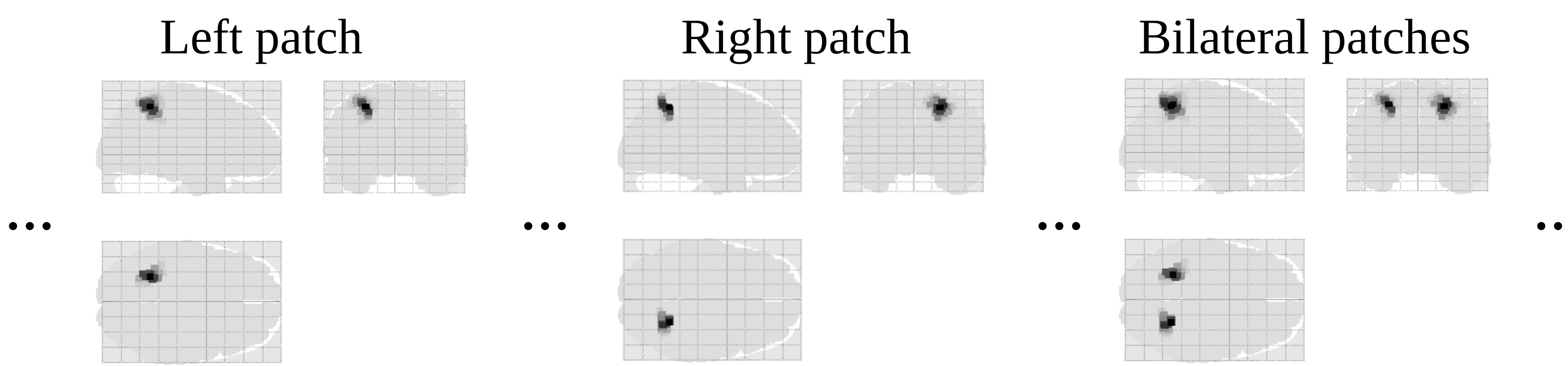
COH $Q^{\varepsilon} = \{G, I\}$

MSP $Q^{\varepsilon} = \{q_1 q_1^T, \ldots, q_N q_N^T\}$





$$G(\sigma) = [q_1, \ldots, q_N] = \sum_{i=0}^{8} \frac{\sigma^i}{i!} A^i \approx \exp(\sigma A)$$



Friston et al (2008) Neuroimage

PEB: Multiple Sparse Priors

MRC Cognition and Brain Sciences Unit

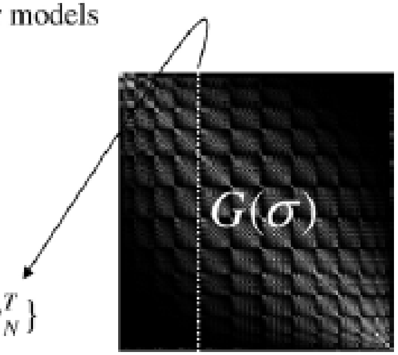
Hyperpriors allow the extreme of 100's source priors, or MSP

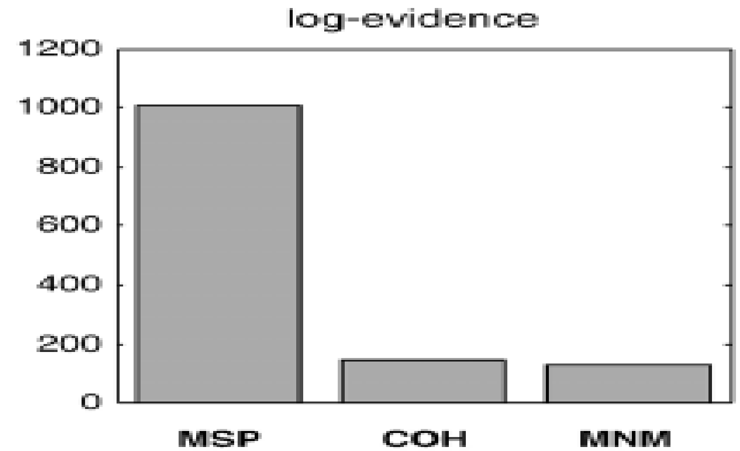
Three prior models

MNM $Q^{\varepsilon} = I$

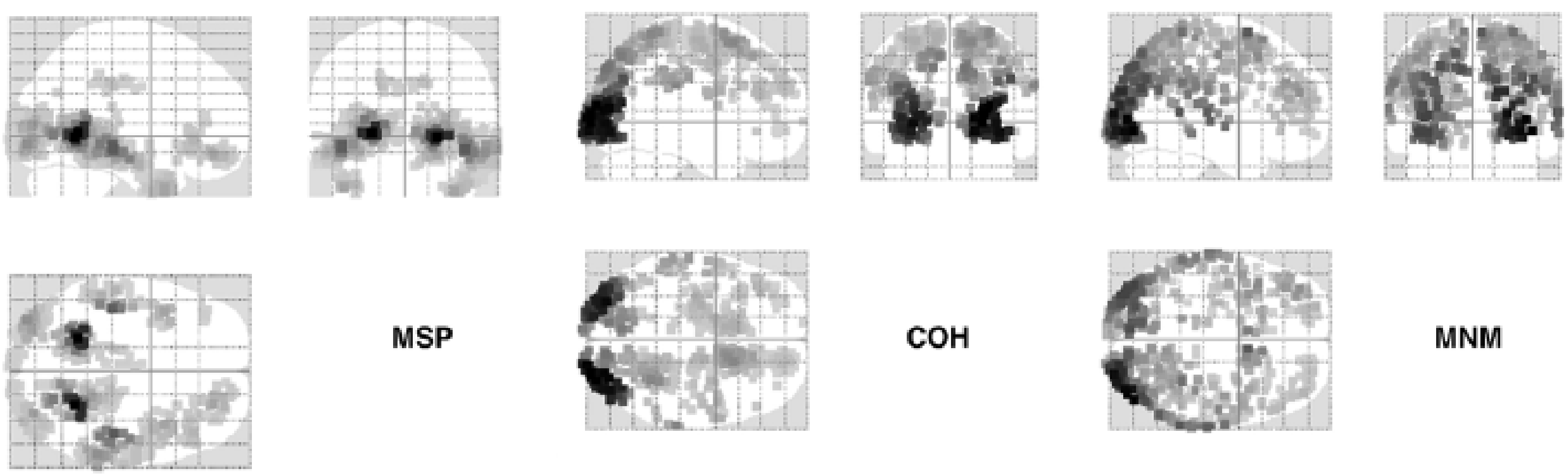
COH $Q^{\varepsilon} = \{G, I\}$

MSP $Q^{\varepsilon} = \{q_1 q_1^T, \ldots, q_N q_N^T\}$





$$G(\sigma) = [q_1, \ldots, q_N] = \sum_{i=0}^{8} \frac{\sigma^i}{i!} A^i \approx \exp(\sigma A)$$



Friston et al (2008) Neuroimage

PEB Summary

MRC | Cognition and Brain Sciences Unit

Summary:

- Automatically "regularises" in 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

Bayesian Perspective

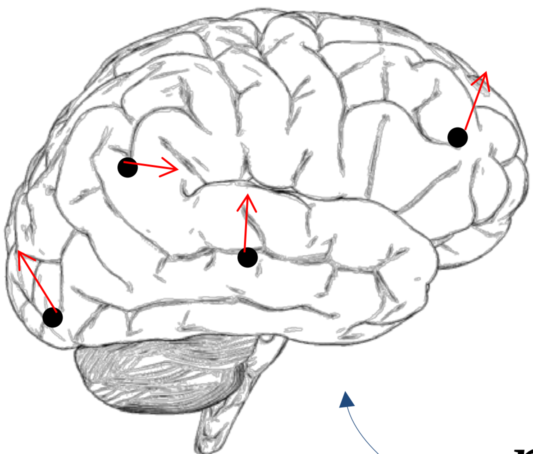
Forward Problem

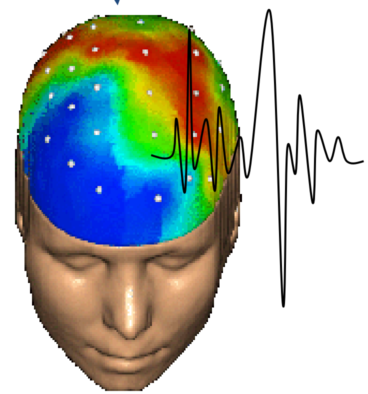
m Model

$p(Y \mid J, m)$ Likelihood

$p(J \mid m)$ Prior

Y Data





Posterior $p(J \mid Y, m)$

Evidence $p(Y \mid m)$

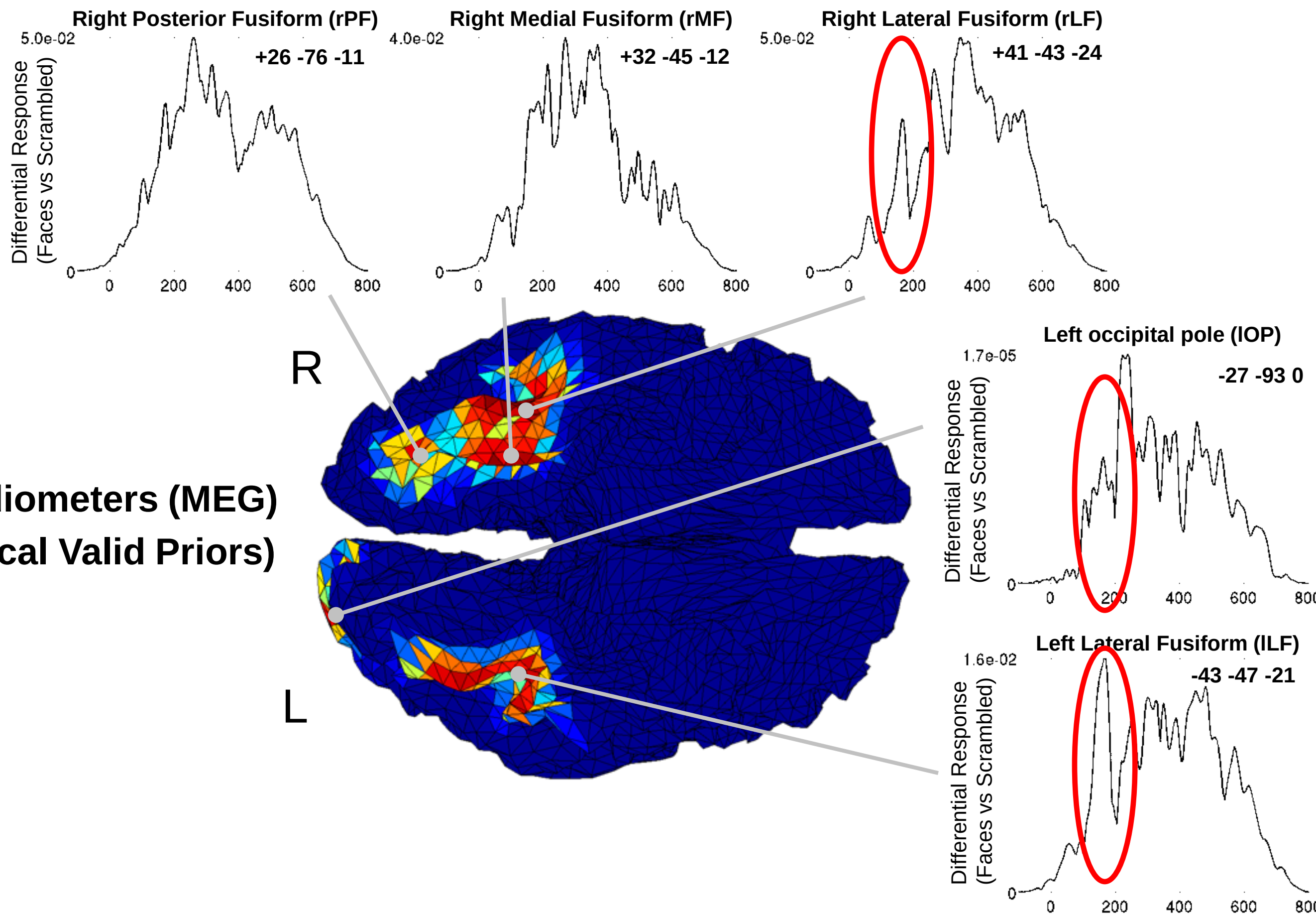
J Parameters

Inverse Problem

Asymmetric Integration of M/EEG+fMRI

MRC | Cognition and Brain Sciences Unit

Gradiometers (MEG)
(5 Local Valid Priors)



NB: Priors affect variance, not precise timecourse…

Henson et al (2010) Hum. Brain Map.