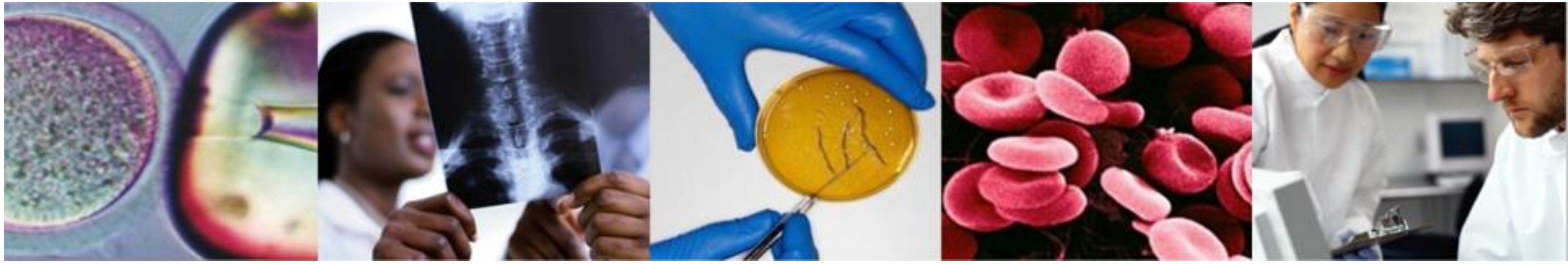
MRC Cognition and Brain Sciences Unit
75th ANNIVERSARY 1944 - 2019

UNIVERSITY OF CAMBRIDGE



EEG/MEG 2:

Head Modelling and Source Estimation

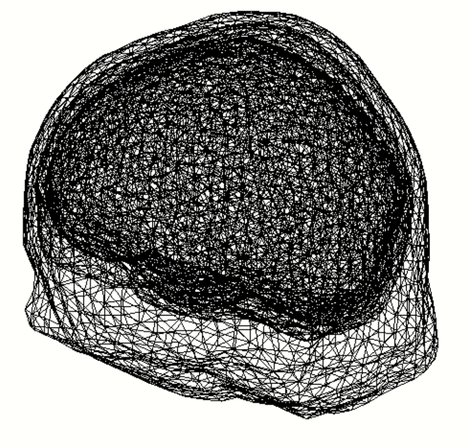
Olaf Hauk

olaf.hauk@mrc-cbu.cam.ac.uk

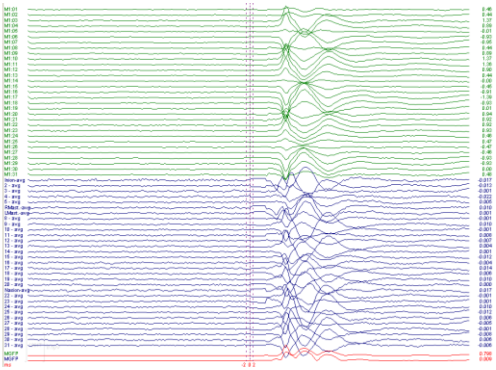
Introduction to Neuroimaging Methods, 4.2.2020

Ingredients for Source Estimation

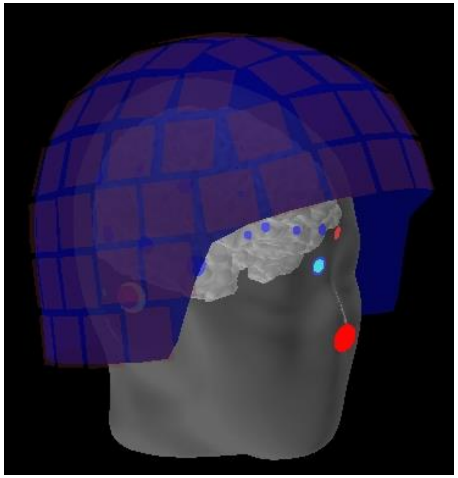
Volume Conductor/
Head Model



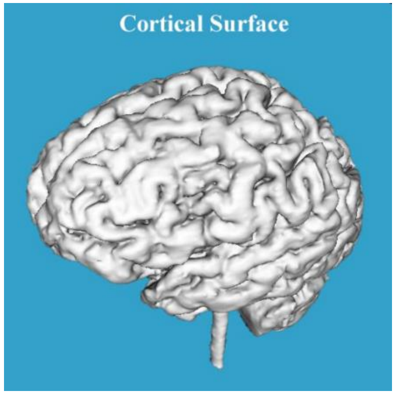
MEG data



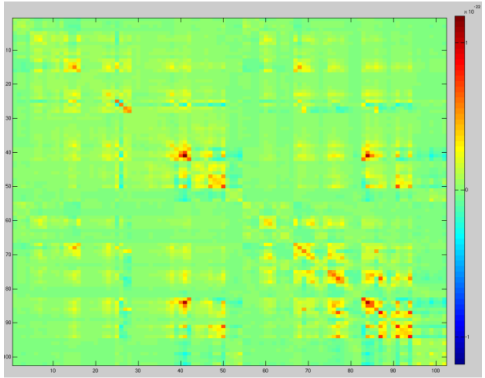
Coordinate
Transformation



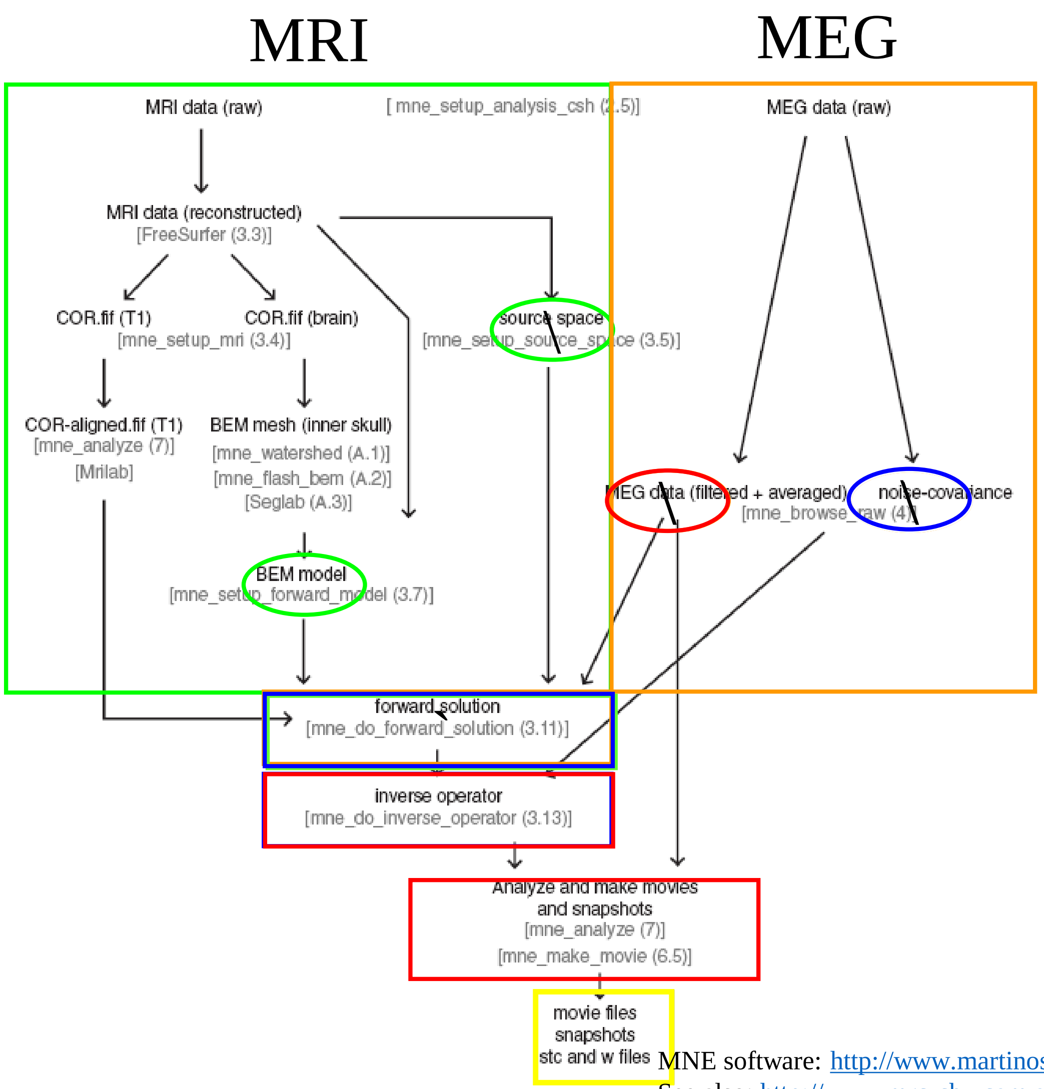
Source Space



Noise/Covariance Matrix



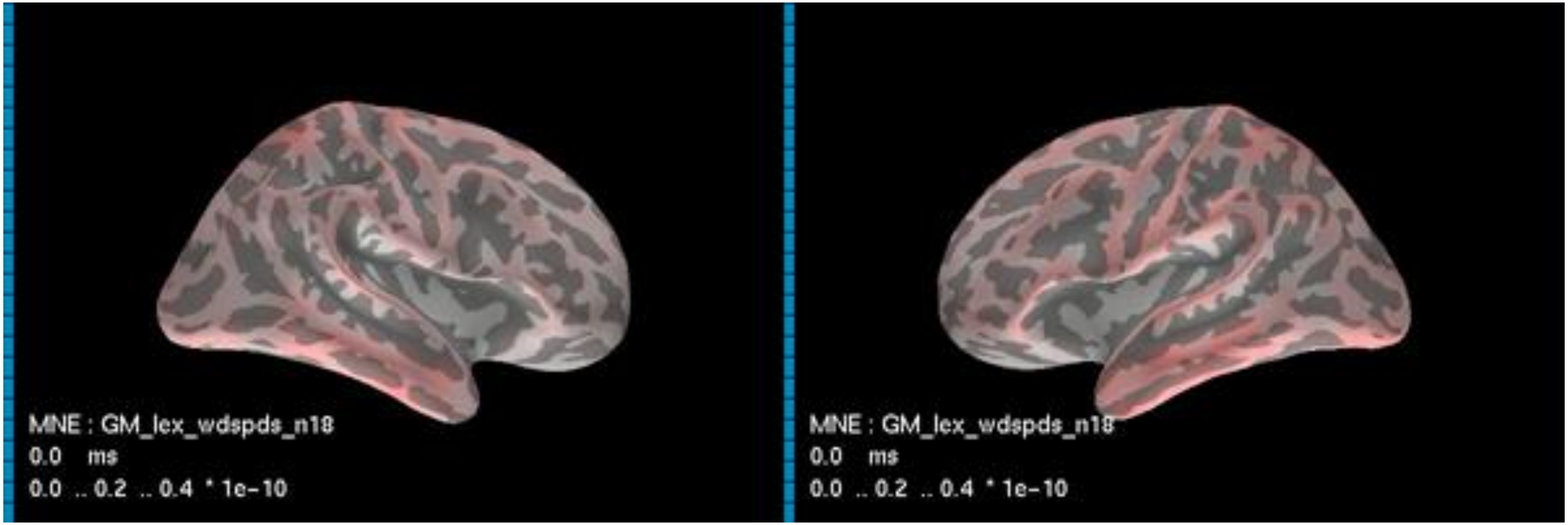
The Path to the Source



MNE software: http://www.martinos.org/mne/

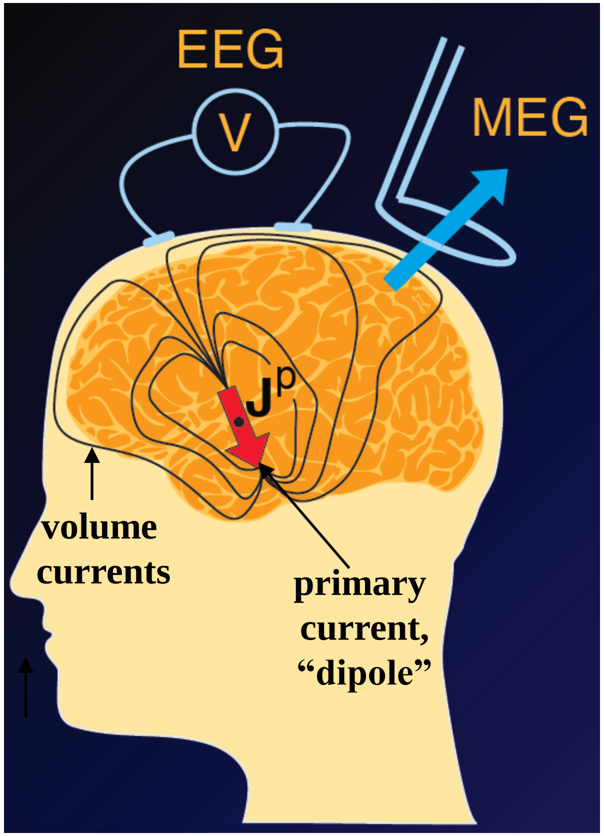
See also: http://www.mrc-cbu.cam.ac.uk/methods-and-resources/imaginganalysis/

Our Goal:
Brain Movies

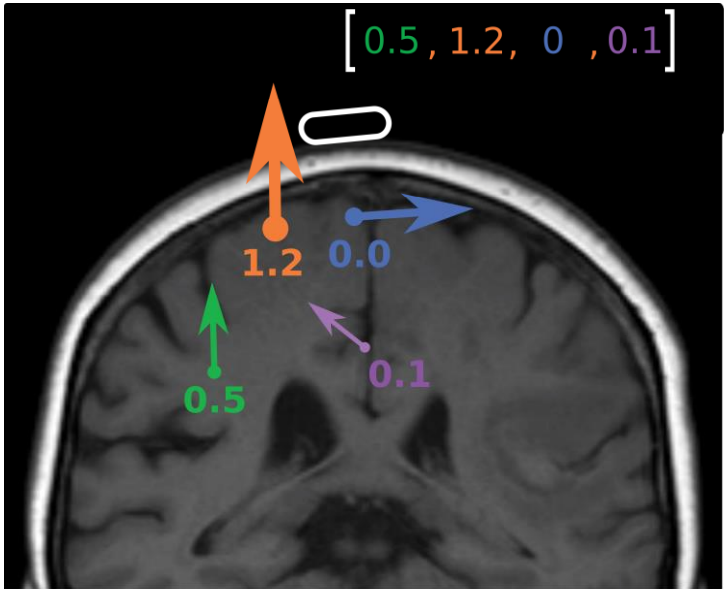


The EEG/MEG Forward Problem

EEG/MEG measure the primary sources indirectly



Sensors are differently sensitive to different sources

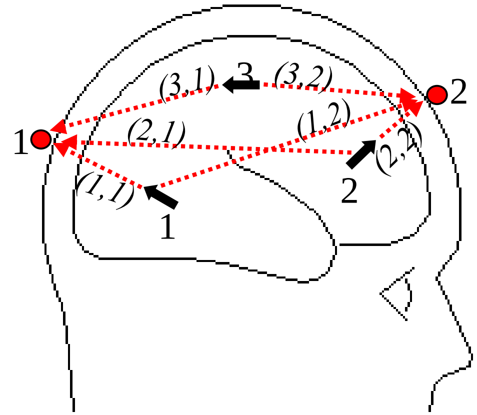


“Leadfield”

Hauk, Strenroos, Treder. In: Supek S, Aine C (edts), “Magnetoencephalography: From Signals to Dynamic Cortical Networks, 2nd Ed.”

The Goal:

Once We Have Stated the Forward Problem, We Are Ready Address the Inverse Problem



Inverse Operator

data "leadfield" dipoles

$$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 0.5 & 0 & 0.3 \\ 0 & 1 & -0.3 \end{pmatrix} \begin{pmatrix} j_1 \\ j_2 \\ j_3 \end{pmatrix}$$

?

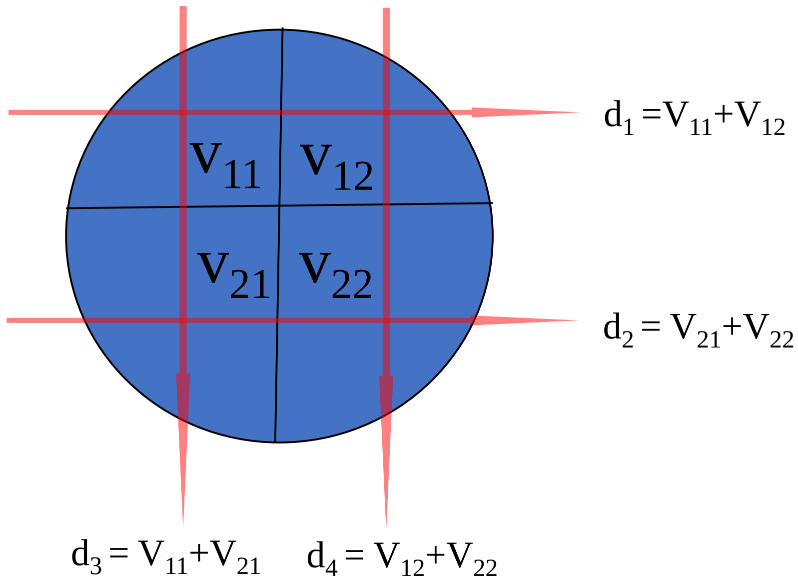
inversion

dipoles inverse data

$$\begin{pmatrix} j_1 \\ j_2 \\ j_3 \end{pmatrix} = \begin{pmatrix} 1.5034 & 0.1241 \\ 0.2483 & 0.9379 \\ 0.8276 & -0.2069 \end{pmatrix} * \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

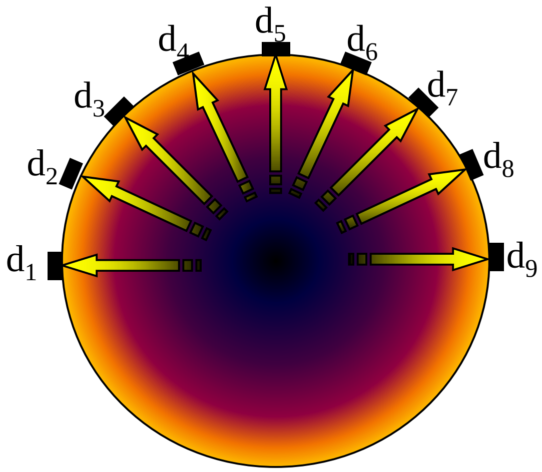
EEG/MEG "Scanning" is not "Tomography"

Tomography (CT, fMRI…)



$$d_1 = V_{11}+V_{12}$$
$$d_2 = V_{21}+V_{22}$$
$$d_3 = V_{11}+V_{21}$$
$$d_4 = V_{12}+V_{22}$$

EEG/MEG



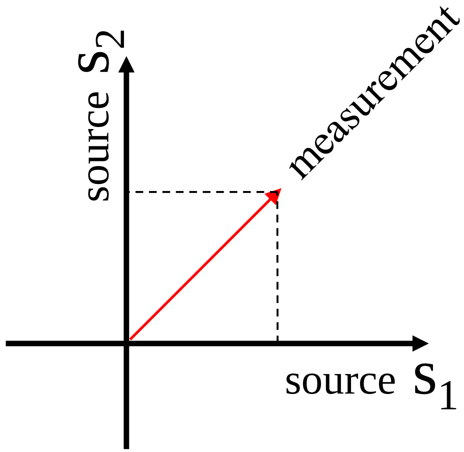
$$d_1 = V_{11}+V_{12}+V_{13}+V_{14} \ldots$$
$$d_2 = V_{21}+V_{22}+V_{23}+V_{24} \ldots$$

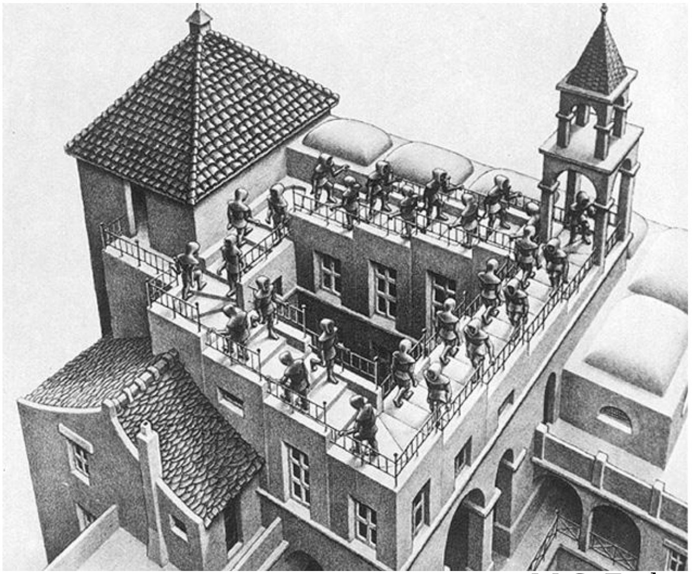
Information is lost during measurement

Cannot be retrieved by mathematics

Inherently limits spatial resolution

Why Inverse "Problem"?





M.C. Escher

In "signal space", we see a faint shadow of activity in "source space".

If you are not shocked by the EEG/MEG inverse problem…

… then you haven't understood it yet.

(freely adapted from Niels Bohr)

Non-Uniquely Solvable Problem

What is the solution to

$$x_1 + x_2 = 1$$

Maybe

$x_1 = 0$; $x_2 = 1$?

$x_1 = 1$; $x_2 = 0$?

$x_1 = 1000$; $x_2 = -999$?

$x_1 = \pi$; $x_2 = (1-\pi)$?

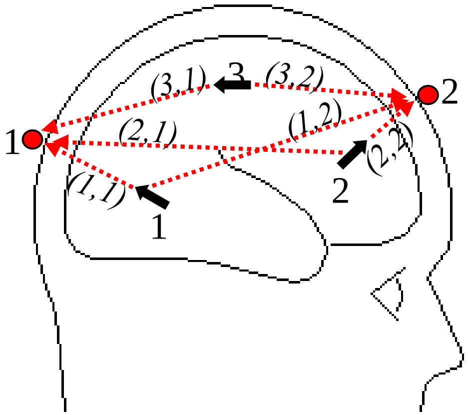
The minimum norm solution is:

$x_1 = 0.5$; $x_2 = 0.5$

with $(0.5^2 + 0.5^2)=0.5$ the minimum norm among all possible solutions.

The Goal:

Once We Have Stated the Forward Problem, We Are Ready Address the Inverse Problem



data "leadfield" dipoles

$$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 0.5 & 0 & 0.3 \\ 0 & 1 & -0.3 \end{pmatrix} \begin{pmatrix} j_1 \\ j_2 \\ j_3 \end{pmatrix}$$

? inversion

Inverse Operator

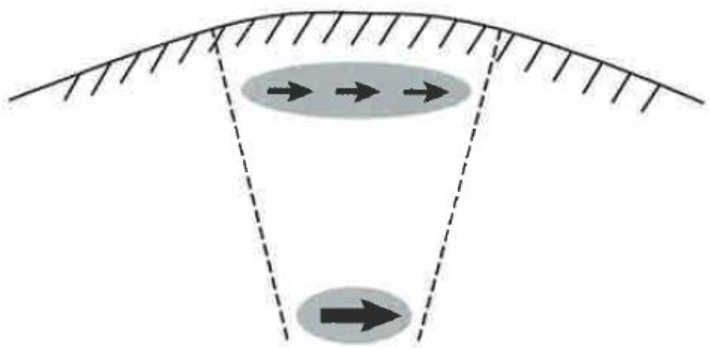
dipoles inverse data

$$\begin{pmatrix} j_1 \\ j_2 \\ j_3 \end{pmatrix} = \begin{pmatrix} 1.5034 & 0.1241 \\ 0.2483 & 0.9379 \\ 0.8276 & -0.2069 \end{pmatrix} * \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

MNE produces solution with minimal power or "norm":

$$\left(j_1^2 + j_2^2 + j_3^2\right)$$

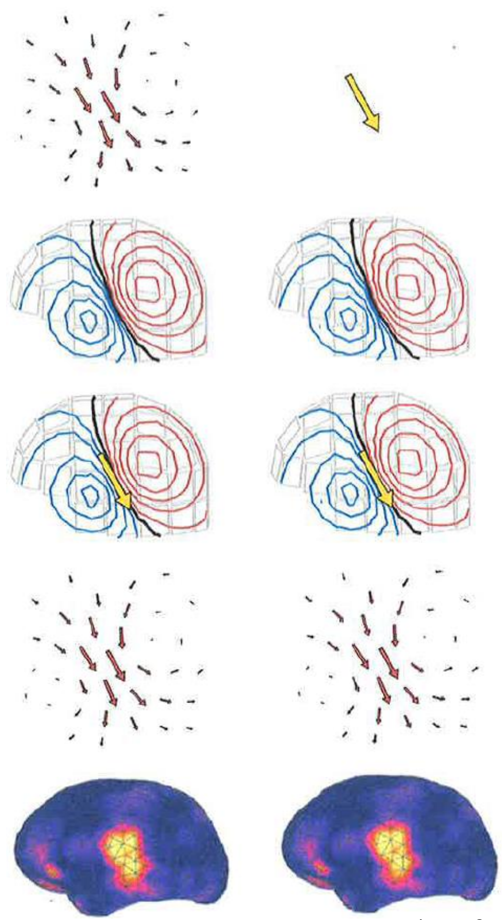
Examples for Non-Uniqueness



A distributed superficial distribution may be indistinguishable from a focal deep source.

Jensen & Hesse, chap. 7 in “MEG”, OUP 2010, Hansen/Kringelbach/Salmelin (edts.)
See also Krishnaswamy et al. PNAS 2017

Examples for Non-Uniqueness



Original Sources

Different Sources

Field Patterns

Same Field Patterns

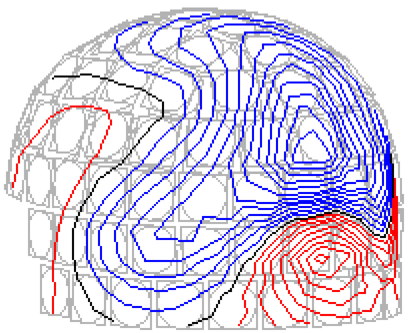
Dipole Model

Same Source Estimates

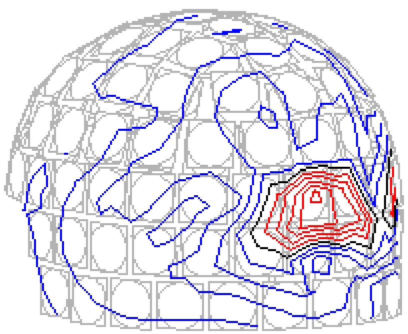
Minimum Norm Estimates

Hämäläinen & Hari, in Brain Mapping: The Methods (2nd), Elsevier 2002

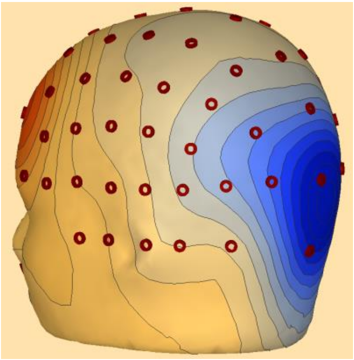
Example: Visually Evoked Activity ~100 ms



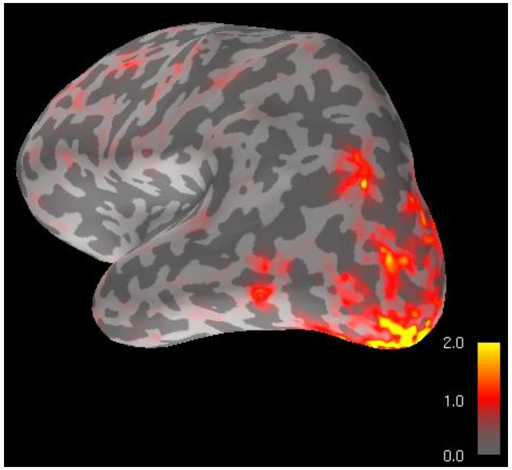
Magnetometers



Gradiometers

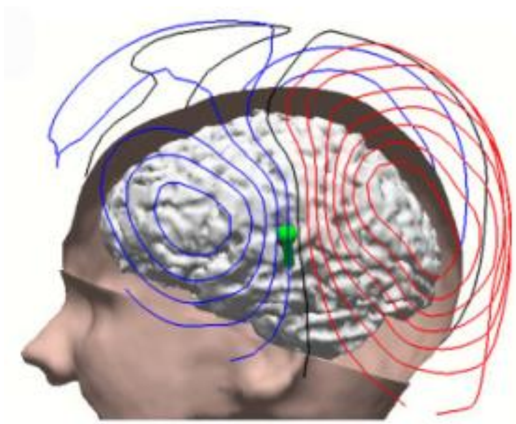


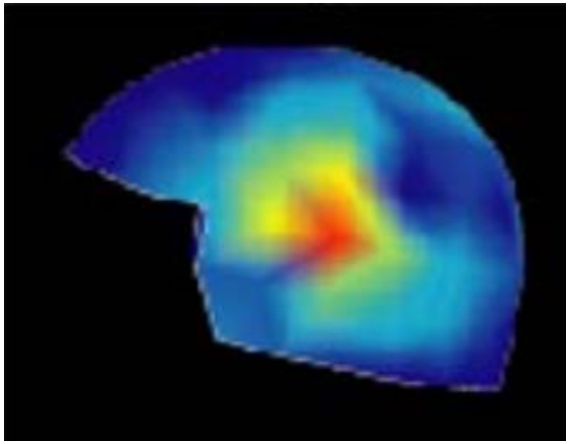
EEG

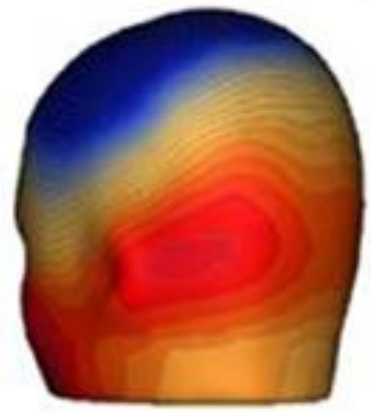


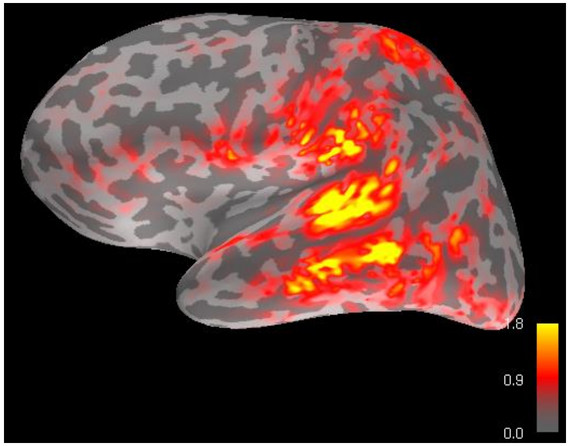
Minimum Norm Estimate

Example: Auditorily Evoked Activity









Minimum Norm Estimate

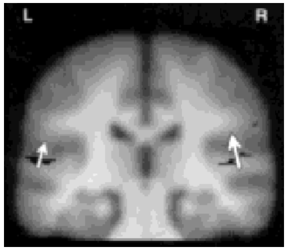
Source Estimation Approaches

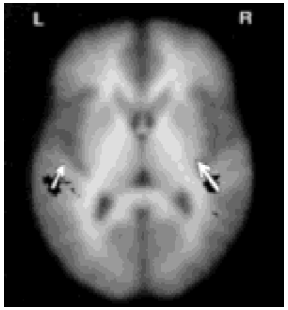
"Dipole Fitting"

1. Assume there are only a few distinct sources

2. Iteratively adjust the location, orientation and strength of a few dipoles…

3. …until the result best fits the data



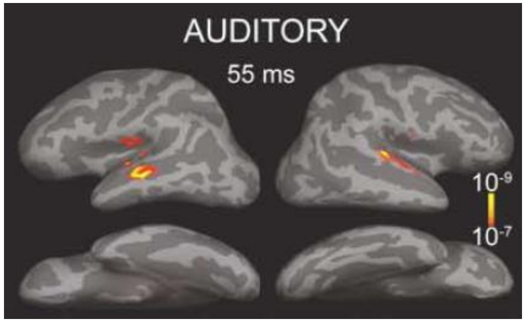


"Distributed Sources"

1. Assume sources are everywhere (e.g. distributed across the whole cortex)

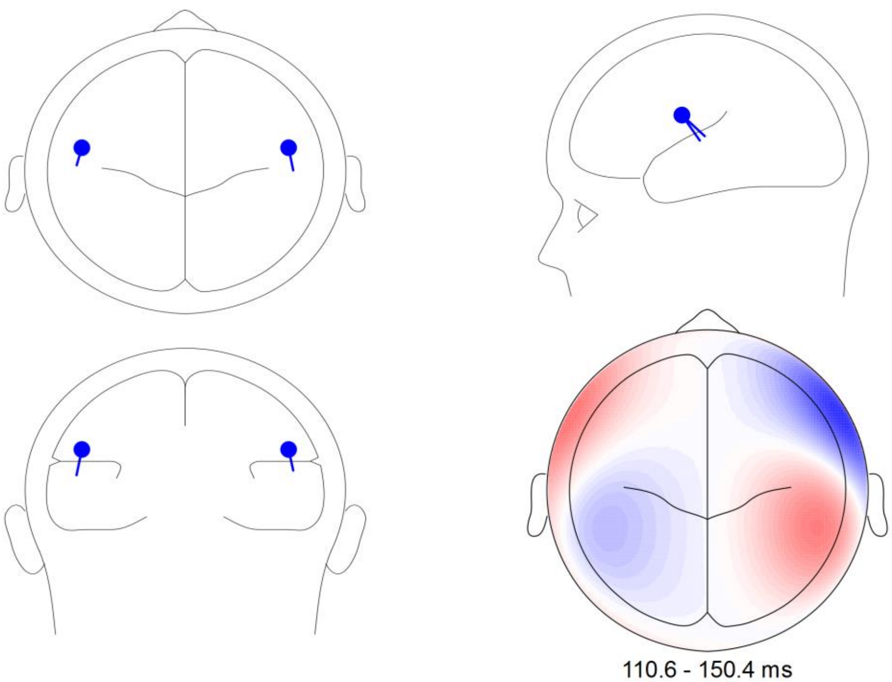
2. Find the distribution of source strengths that explains the data…

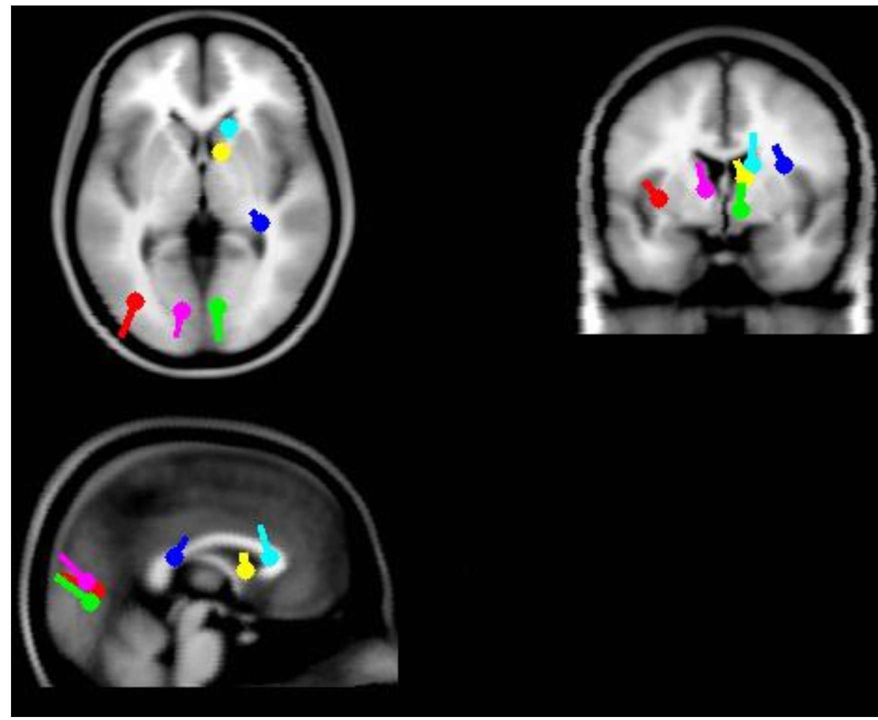
3. …AND fulfils other constraints



Isolated Dipoles

For dipole fits, the forward solutions are computed iteratively for every change of dipole location and orientation.

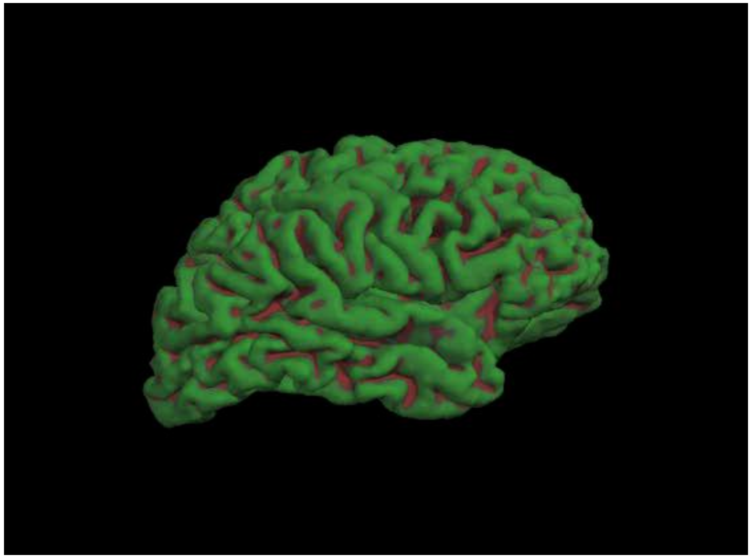




For dipole scans and distributed source methods, forward solutions are pre-computed for a large number of sources in a discretised volume or on a discretised surface ("leadfield matrix").

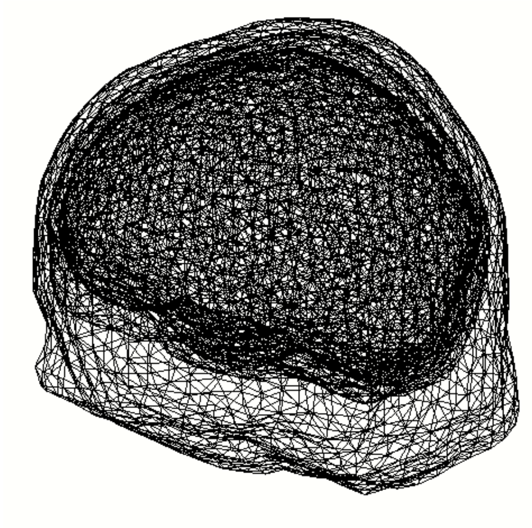
Distributed Sources: Source Space and Head Model

Source Space,
e.g. grey matter, 3D volume



http://www.cogsci.ucsd.edu/~sereno/movies.html

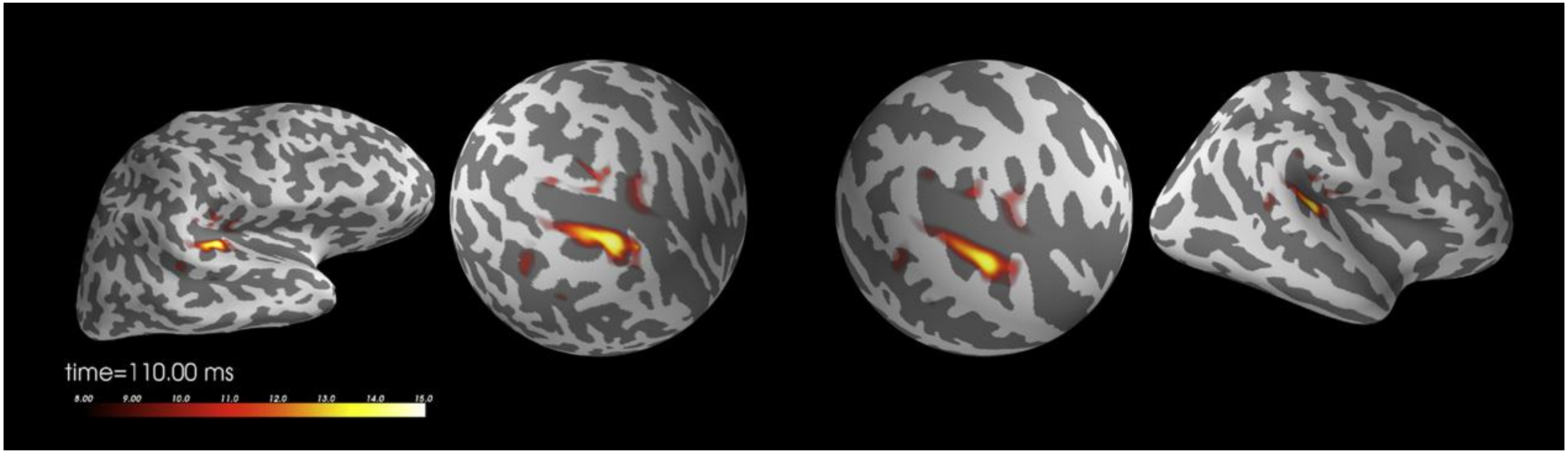
Volume Conductor/Head Model
e.g. sphere, 1- or 3-compartments from MRI



Sometimes "standard head models" are used, when no individual MRIs available.

SPM uses the same "canonical mesh" as source space for every subjects, but adjusts it individually.

Normalising (Morphing) Cortical Surfaces



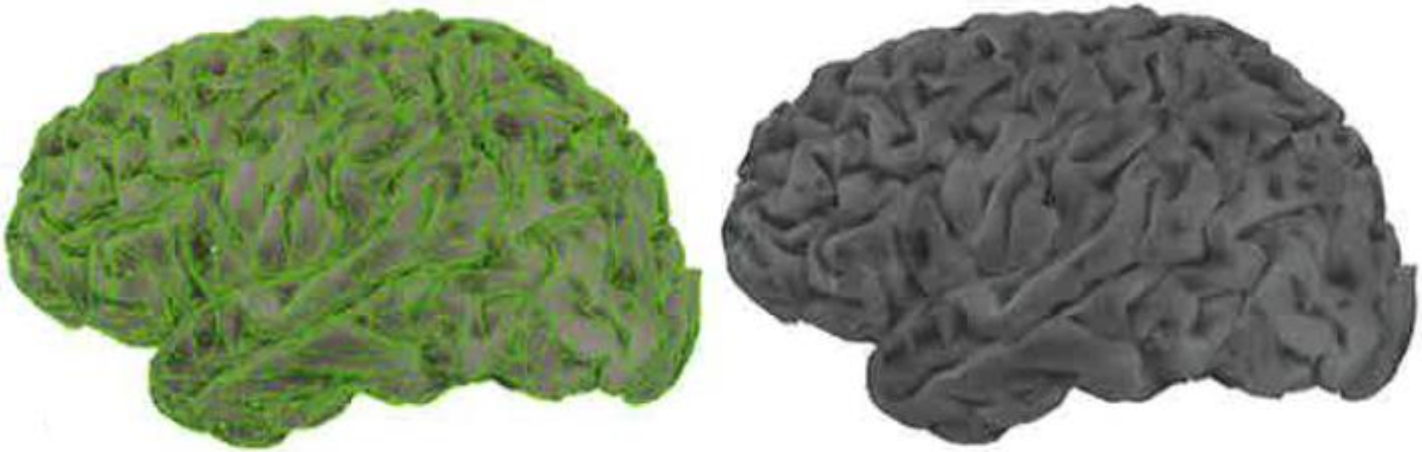
Gramfort et al., NI 2014

Sometimes “standard head models” are used, when no individual MRIs available.

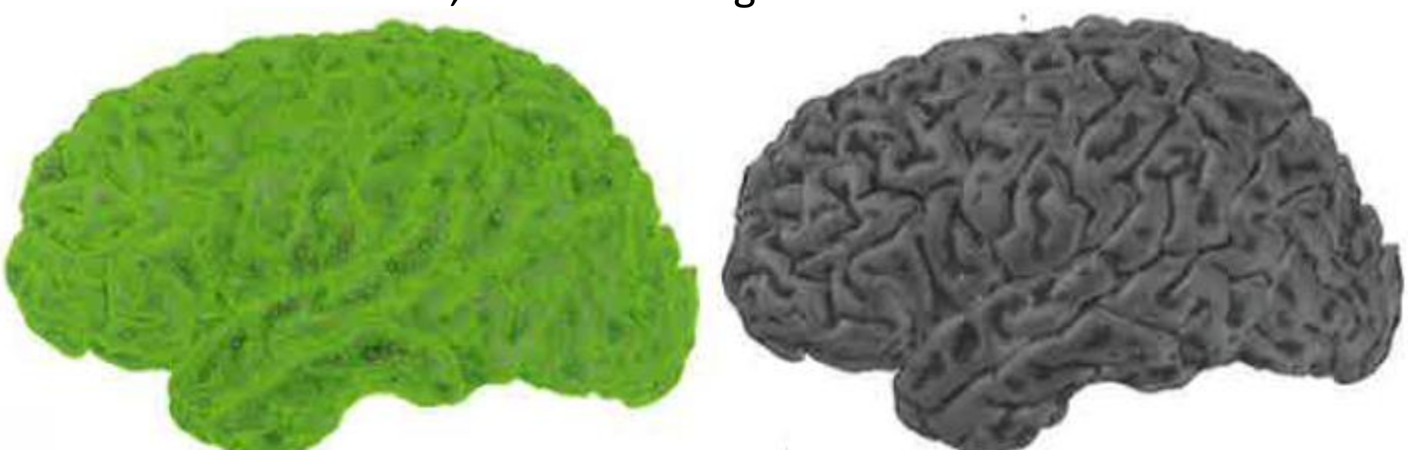
SPM uses the same “canonical mesh” as source space for every subjects, but adjusts it individually.

Spatial Sampling of Cortical Surfaces

10.034 vertices, 20.026 triangles of 10 mm^2 surface area

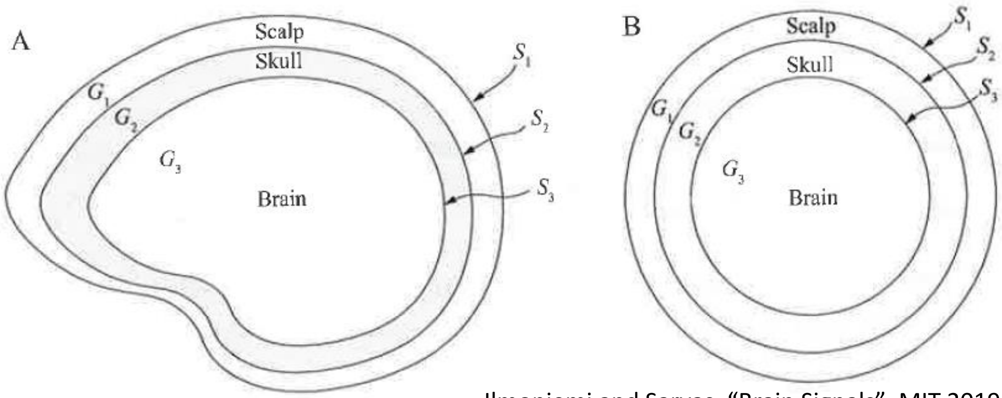


79.124 vertices, 158.456 triangles of 1.3 mm^2 surface area

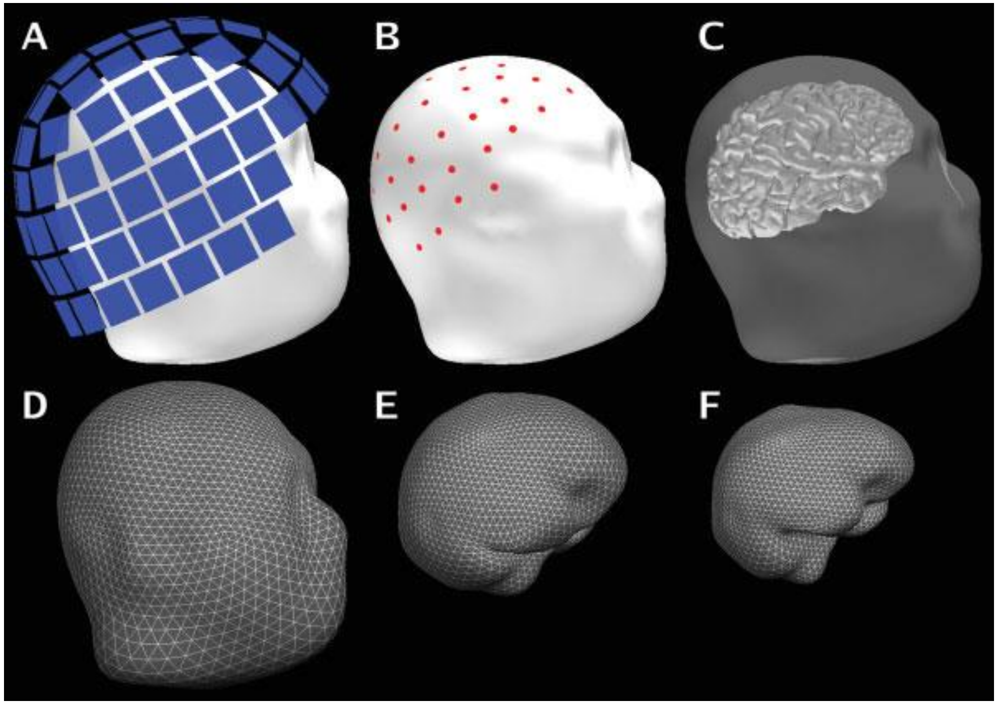


Baillet, chap. 5 in “MEG”, OUP 2010, Hansen/Kringelbach/Salmelin (edts.)

Head Modelling – Tissue Compartments



Ilmoniemi and Sarvas, “Brain Signals”, MIT 2019



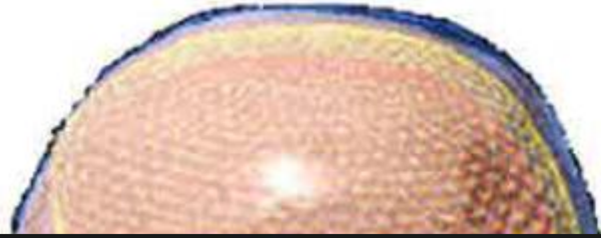
Goldenholz et al., HBM 2009

Head Models With Different Levels of Detail

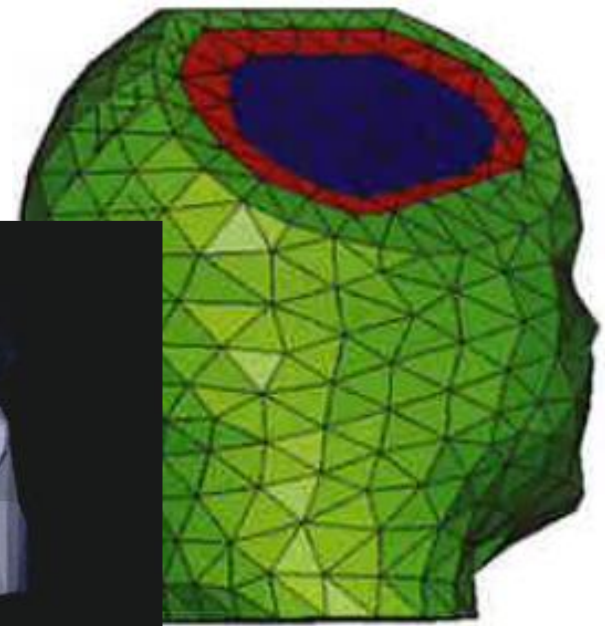
MRC
Cognition
and Brain
Sciences Unit

Spheres

Boundary Element Model
(BEM)



Finite Element Model
(FEM)



Baillet, chap. 5 in "MEG", OUP 2010, Hansen/Kringelbach/Salmelin (edts.)

Conclusion – Head Modelling

3-compartment BEM models are currently state-of-the-art for EEG/MEG source estimation.

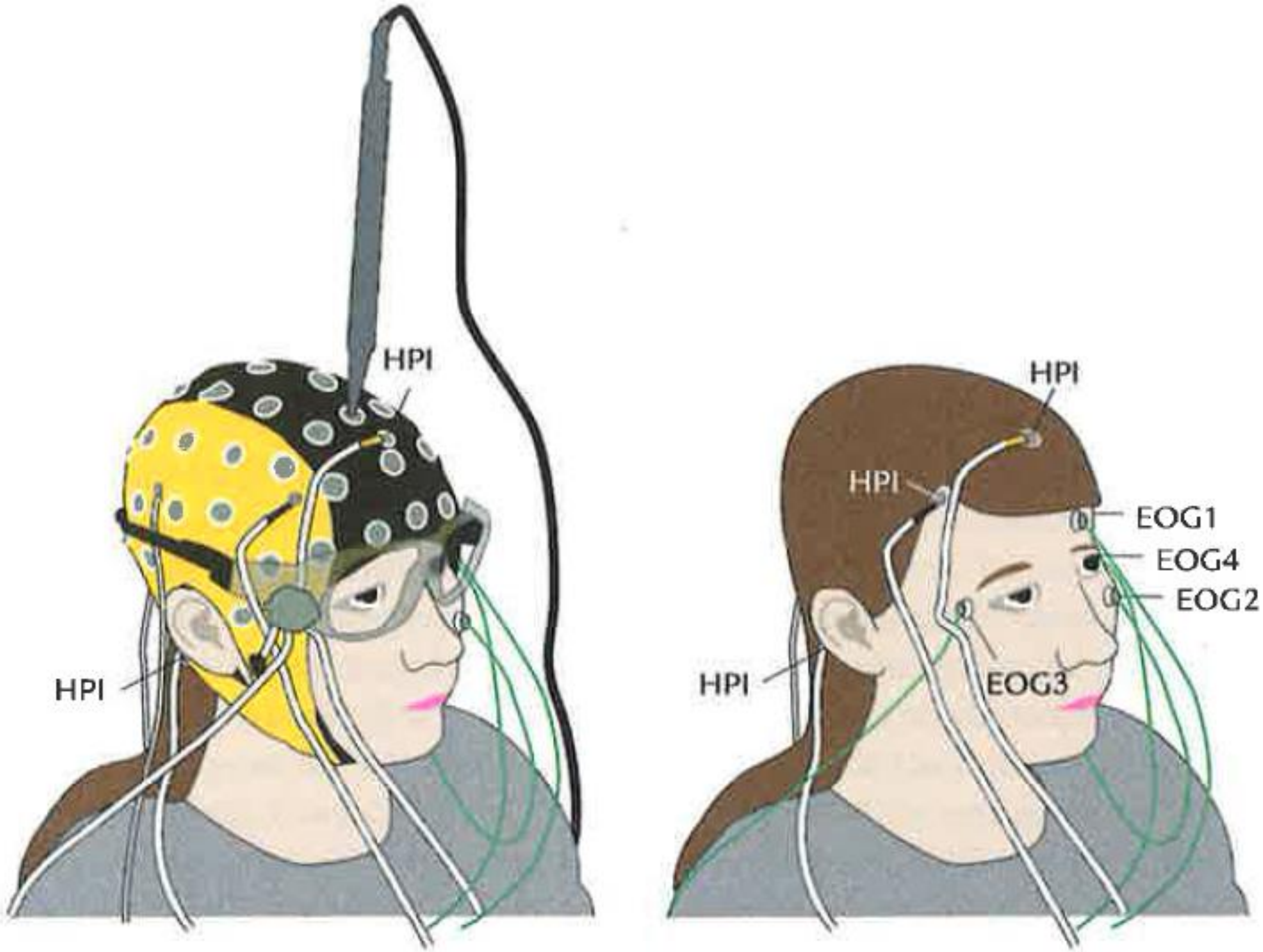
Single-shell approximations are common for MEG.

More detailed head models may increase accuracy, but require more accurate data and information, such as accurate MRI segmentations and conductivity values. (see e.g. Vorwerk et al., BioMeg Eng Online 2018) for Fieldtrip FEM pipeline)

There is no right or wrong, there are only different approximations – know your limits.

Practice

Coregistration of EEG/MEG and MRI Spaces

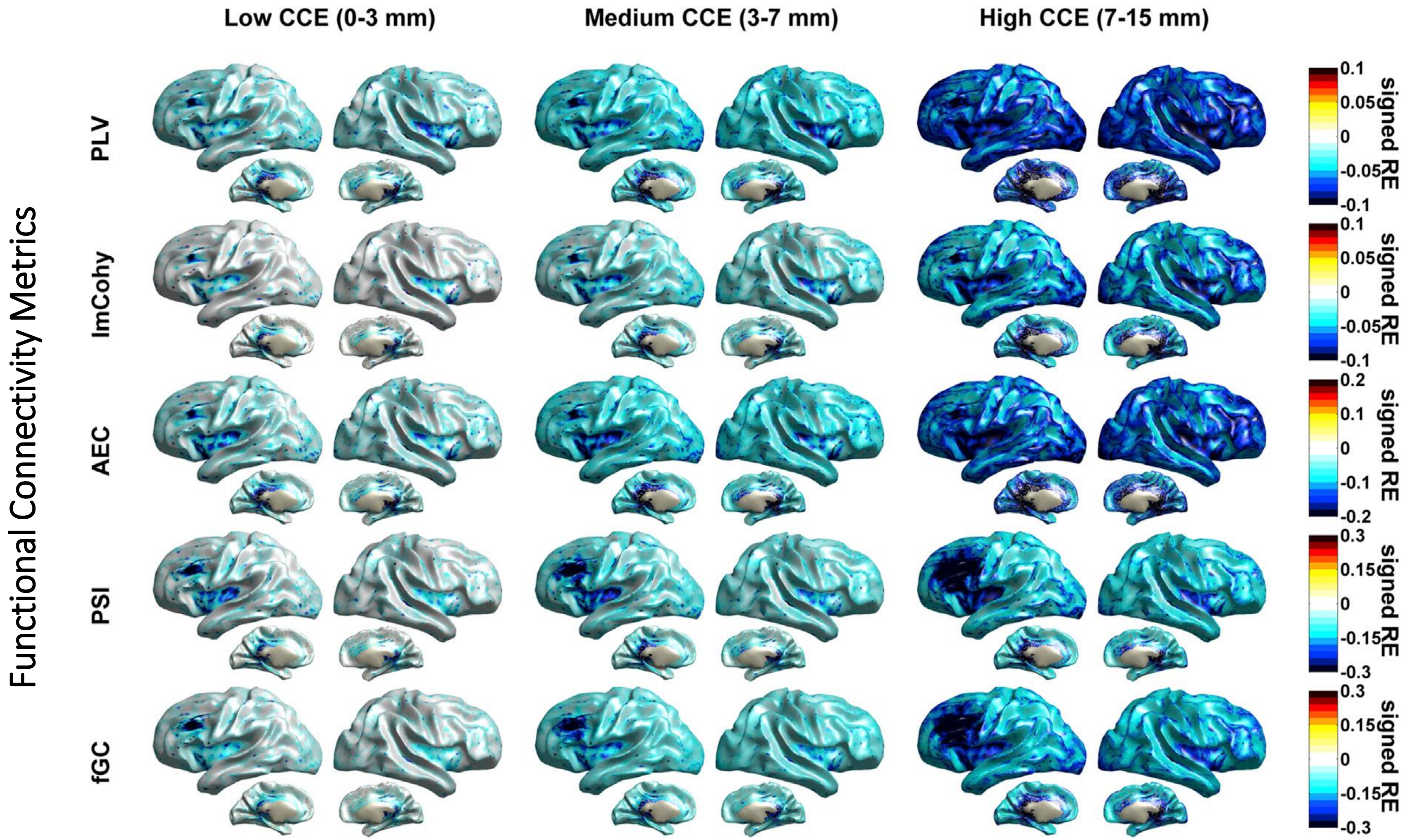


Hari & Puce, "MEG-EEG Primer", OUP 2017

Accurate Coregistration Is Important

Coregistration errors affect the forward model, and therefore everything that follows.
For example, connectivity analysis:

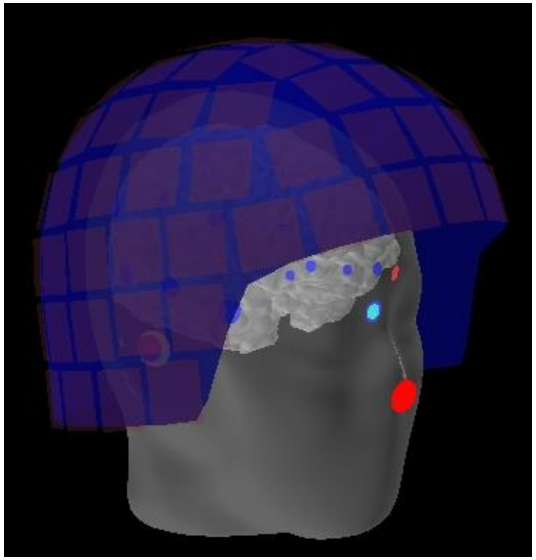
3 levels of coregistration error



Chella et al., NI 2019

Coregistration of EEG/MEG and MRI Spaces

Coordinate
Transformation



Spatial Resolution of Source Estimation

Spatial resolution depends on:

modeling assumptions

number of sensors (EEG/MEG or both)

source location

source orientation

signal-to-noise ratio

head modeling

=> difficult to make general statement

Spatial Resolution – A Naïve Estimate

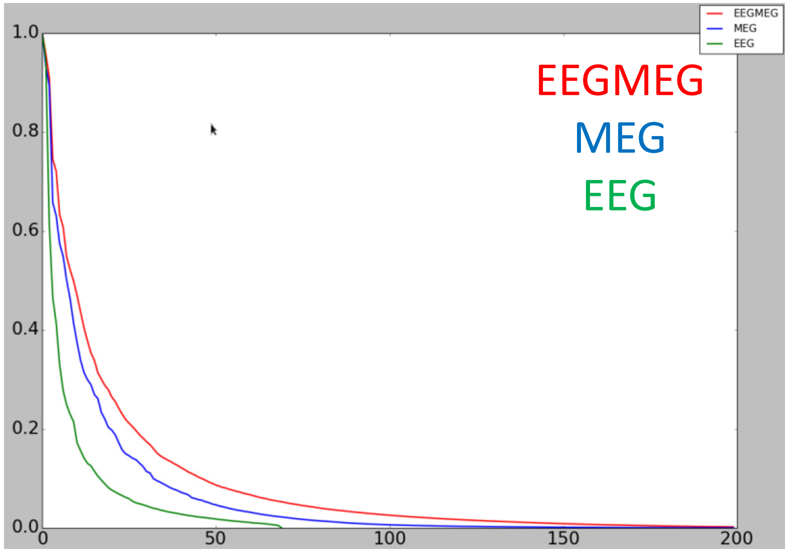
With ***n* sensors**:
-> *n* independent measurements
-> *n* independent parameters estimable
-> at best separate activity from *n* brain regions

Sensors are not independent, data are noisy: **~ 50 degrees of freedom**

Volume of source space:
Sphere 8cm minus sphere 4 cm: volume ~1877 cm^3

"Resel": 38 cm^3 -> $\underline{3.4}^3$ cm^3

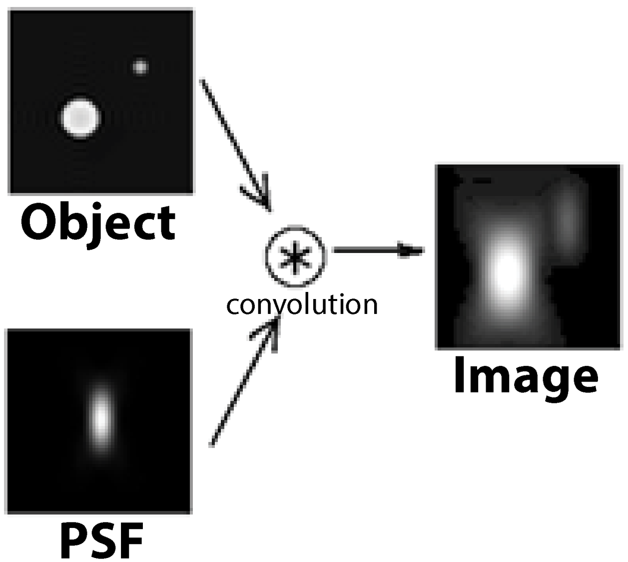
SVD of Leadfields



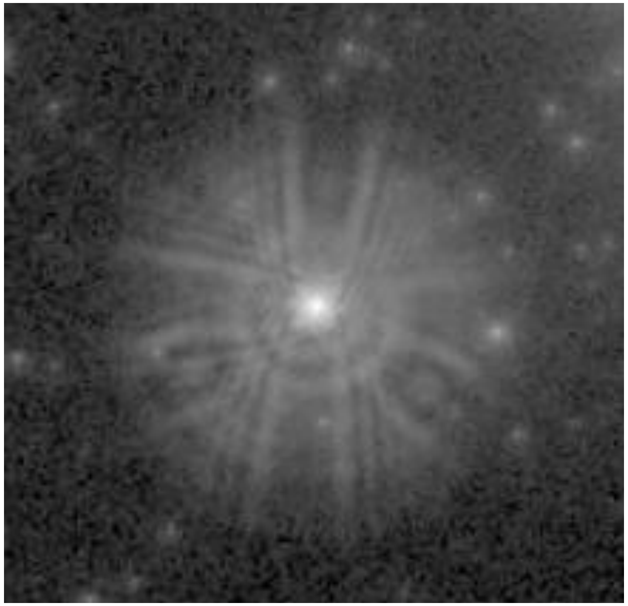
The “Blurry Image” Analogy

The Superposition Principle
An "Assumption-Free" Interpretation of Linear Methods

Microscopy



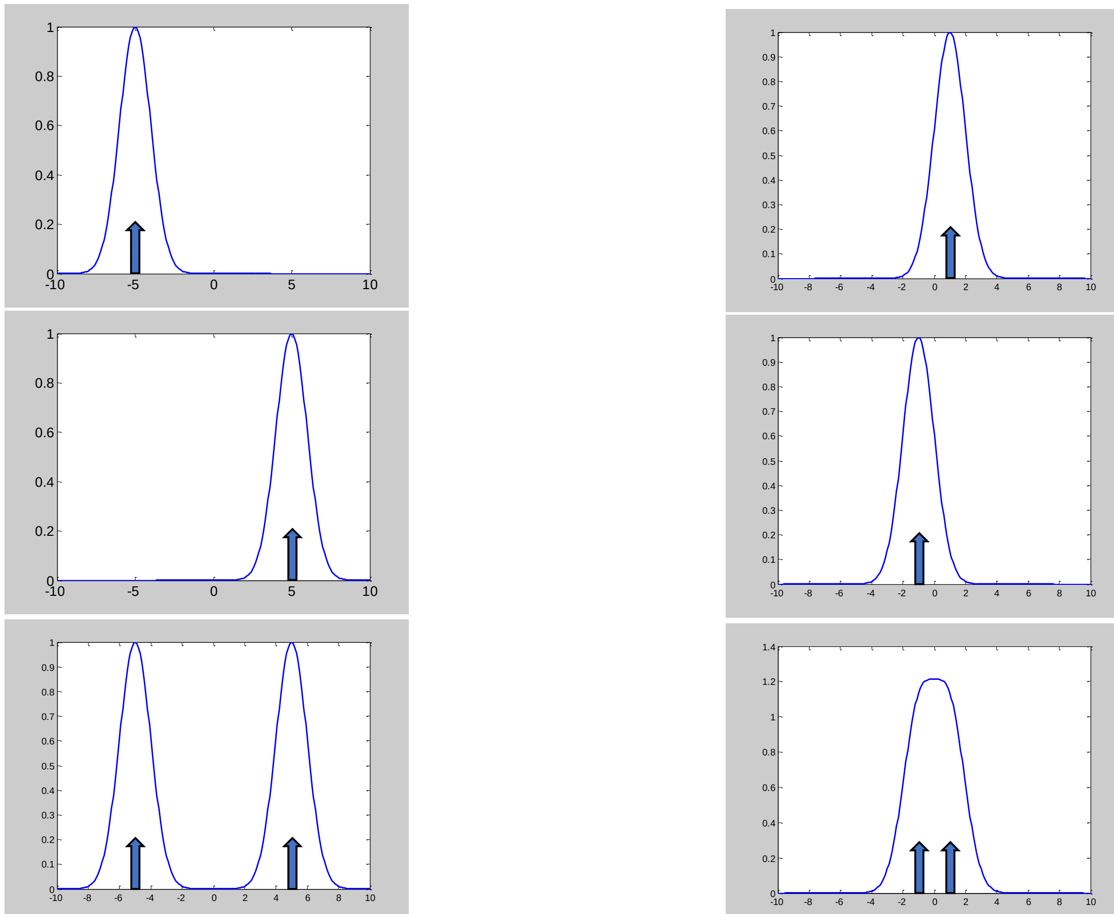
Astronomy



https://en.wikipedia.org/wiki/Point_spread_function

MRC
Cognition and Brain Sciences Unit

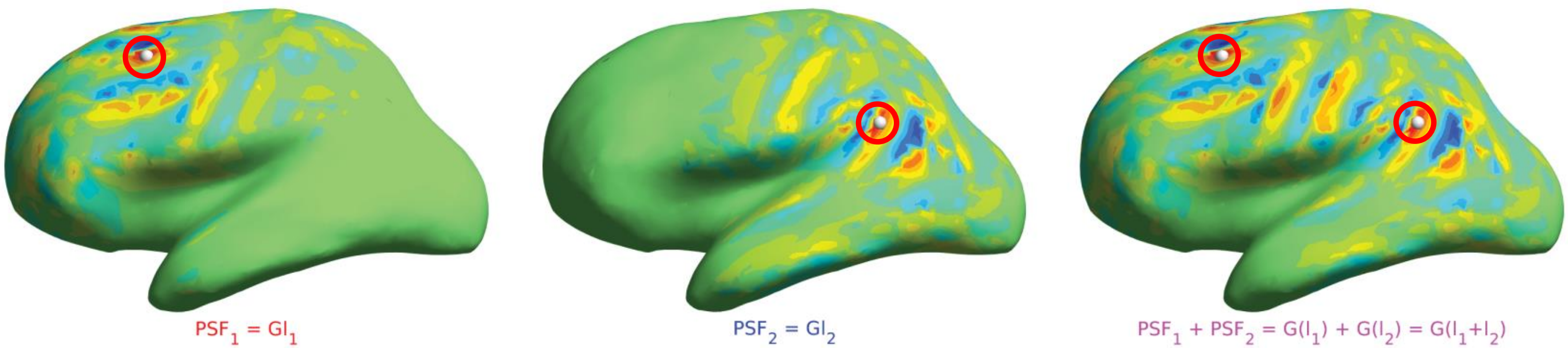
Linear Methods – Superposition Principle



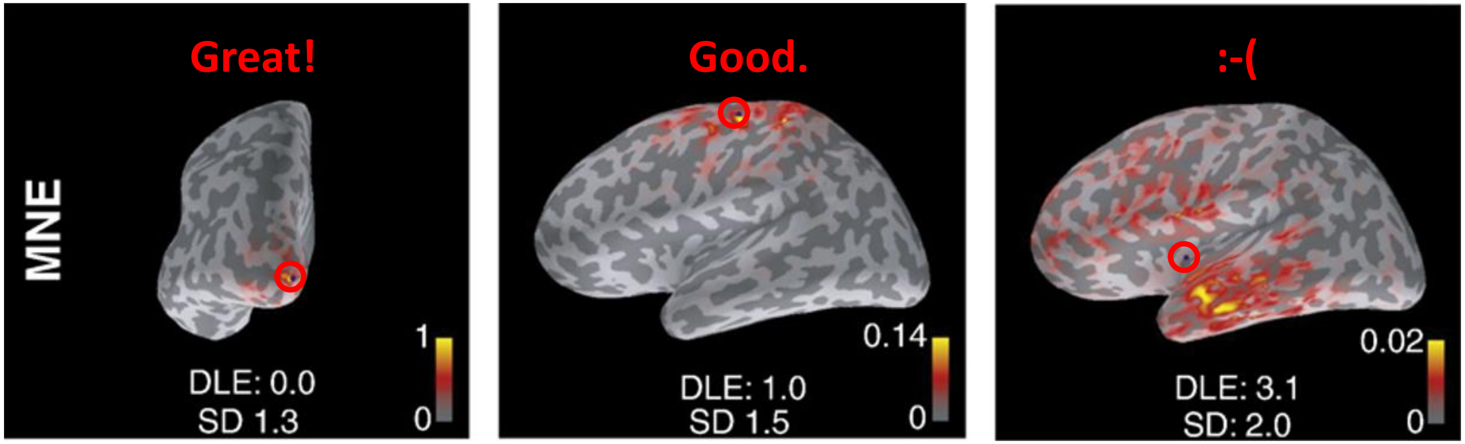
If you know the behaviour for point sources,
you can predict the behaviour for complex sources

Linear Methods – Superposition Principle

Superposition In Source Space



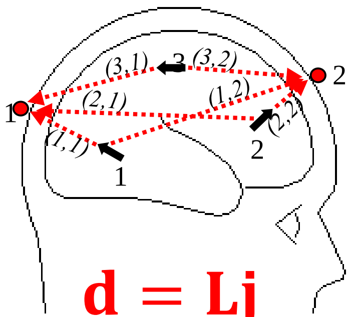
Example Point-Spread Functions



Hauk, Strenroos, Treder. In: Supek S, Aine C (edts), “Magnetoencephalography: From Signals to Dynamic Cortical Networks, 2nd Ed.”

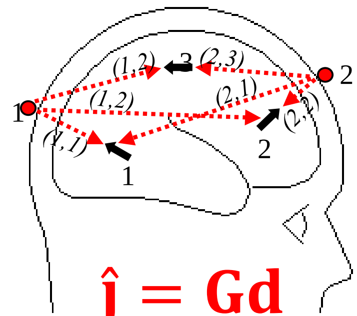
Resolution Matrix

Forward Problem



$$\mathbf{d} = \mathbf{Lj}$$

Linear Inverse Problem



$$\hat{\mathbf{j}} = \mathbf{Gd}$$

$$\hat{\mathbf{j}} = \mathbf{GLj} \stackrel{\text{def}}{=} \mathbf{Rj}$$

Relationship between estimated and true source distribution.

The Best Resolution Matrix

$$\hat{\mathbf{s}} = \mathbf{R}\boldsymbol{s}$$

The closer **R** is to the identity matrix, the closer our estimate is to the true source.

Therefore, let us minimise the difference between **R** and the identity matrix in the least-squares sense:

$$\|\boldsymbol{R} - \boldsymbol{I}\|_{\mathbf{2}} = min$$

Once again, we obtain the minimum-norm least-squares solution:

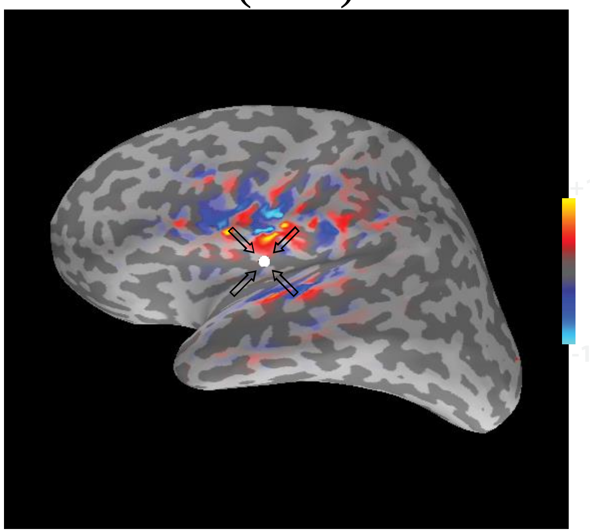
$$\boldsymbol{G}_{\boldsymbol{MN}} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1}$$

Its resolution matrix $\boldsymbol{R}_{\boldsymbol{MN}} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L}$ is symmetric.

Spatial Resolution:
Point-Spread and Cross-Talk/Leakage

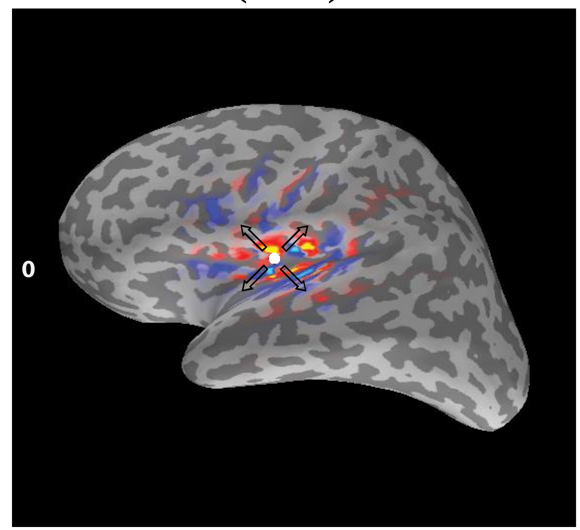
MRC
Cognition and Brain Sciences Unit

Cross-Talk Function (CTF)



How other sources may affect the estimate for this source

Point-Spread Function (PSF)



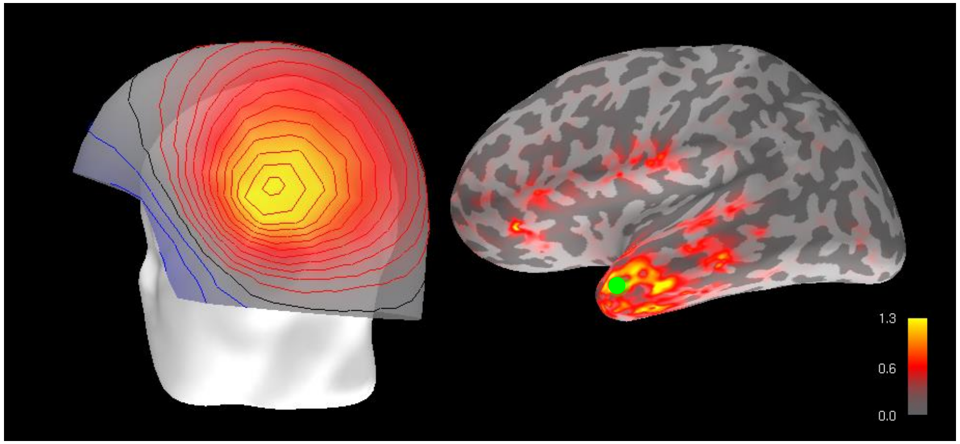
How this source affects estimates for other sources

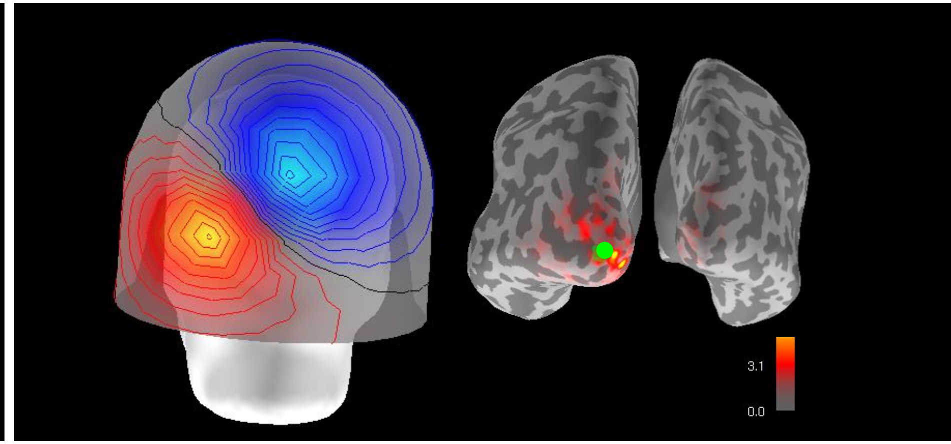
Hauk, Strenroos, Treder. In: Supek S, Aine C (edts), “Magnetoencephalography: From Signals to Dynamic Cortical Networks, 2nd Ed.”

MRC
Cognition and Brain Sciences Unit

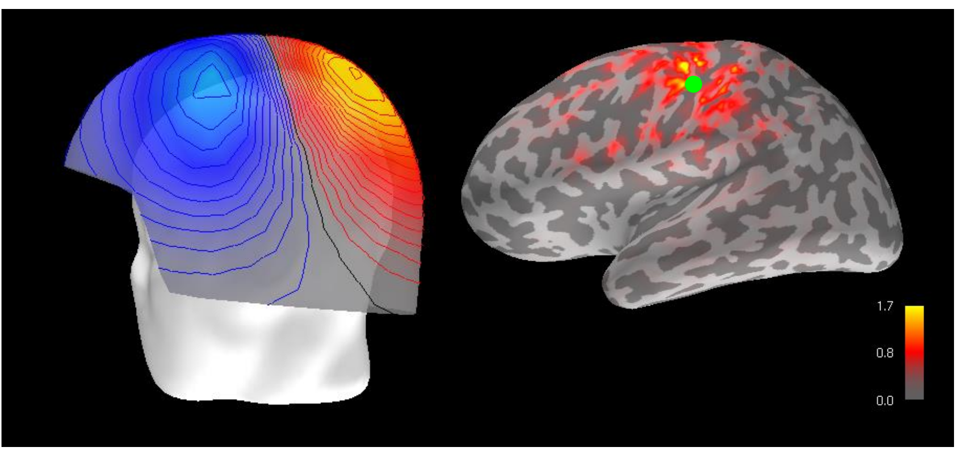
PSFs and CTFs for Some ROIs

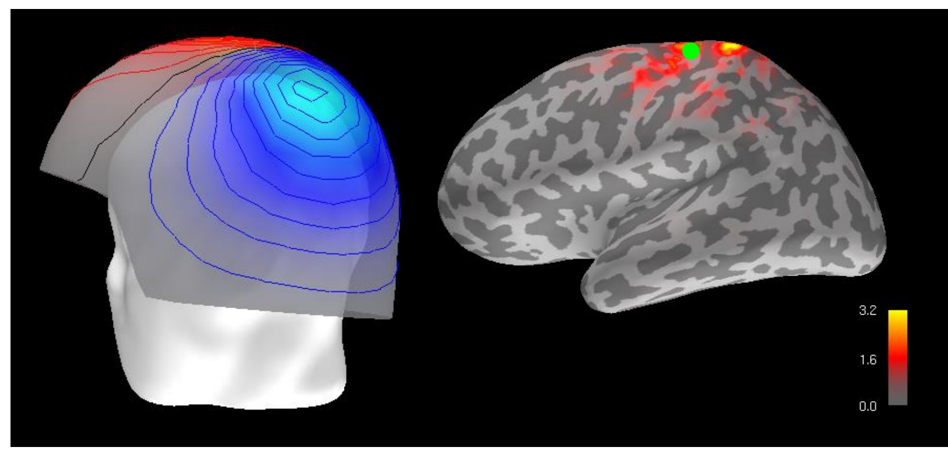
For MNE, PSFs and CTFs turn out to be the same





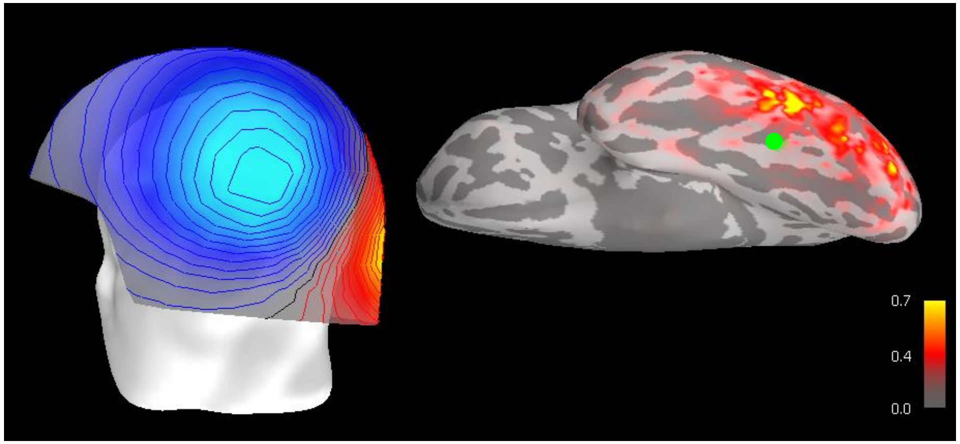
Good

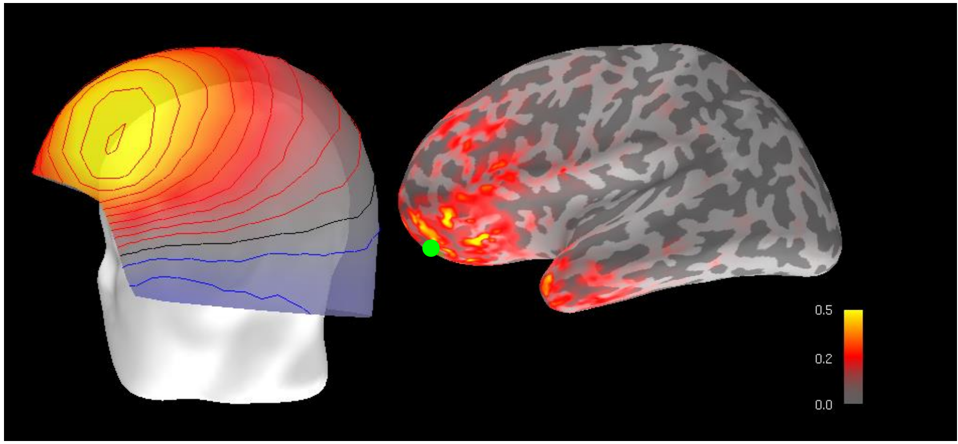




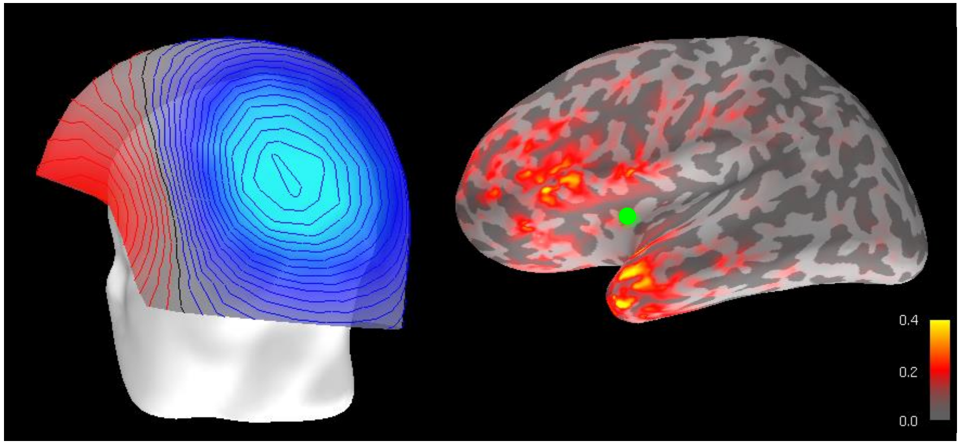
PSFs and CTFs for Some ROIs

For MNE, PSFs and CTFs turn out to be the same

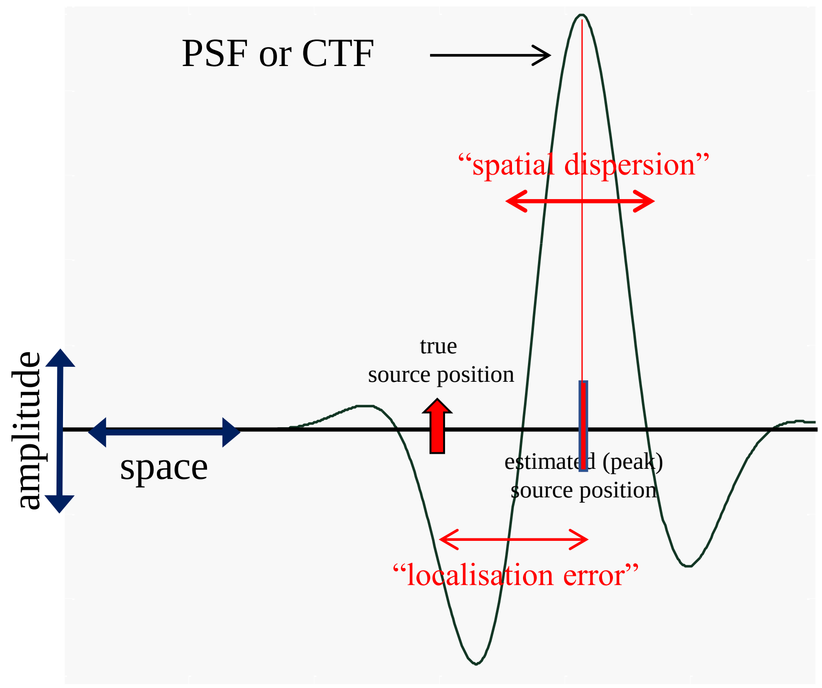


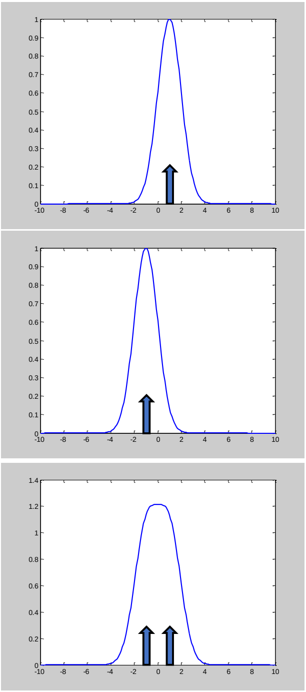


Less good



Quantifying Resolution From PSFs and CTFs





It's not just peak localisation that counts,
but also spatial extent of the distribution.

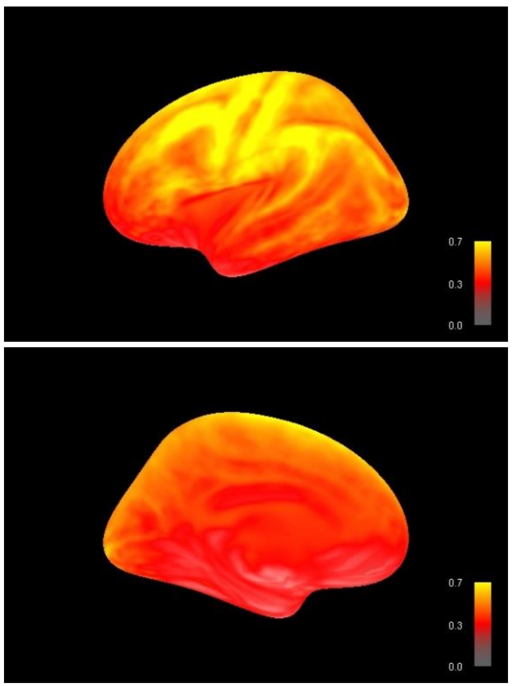
Resolution Metrics For PSFs/CTFs

- **MEG+EEG**: Elekta Vectorview (360+70 channels), Wakeman & Henson open data set
- **Whitened** leadfields and data to combine sensor types
- **Methods Comparison**:
  - L2-MNE
  - depth-weighted L2-MNE
  - dSPM
  - sLORETA
  - LCMV beamformer (noise covariance matrix from baseline intervals)
- **Resolution Metrics**:
  - Peak Localisation Error
  - Spatial Dispersion (extent)

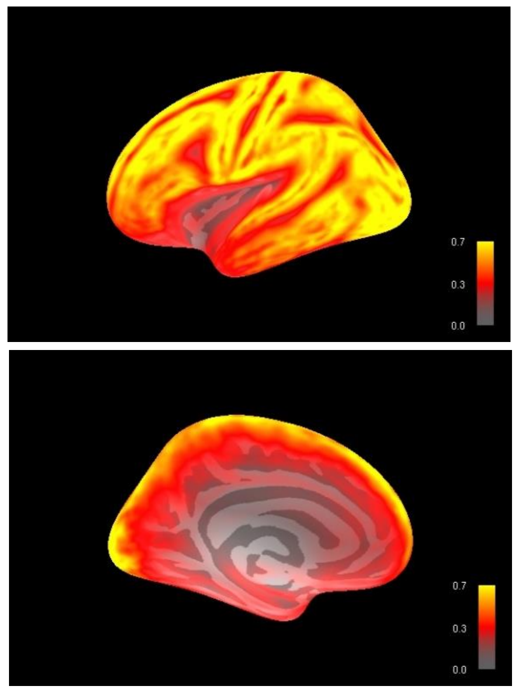
Sensitivity Maps

RMS of Leadfield Columns

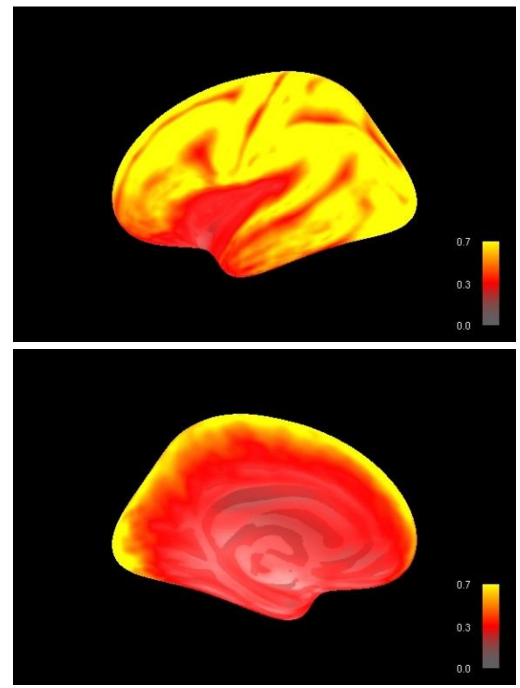
EEG
70 electrodes



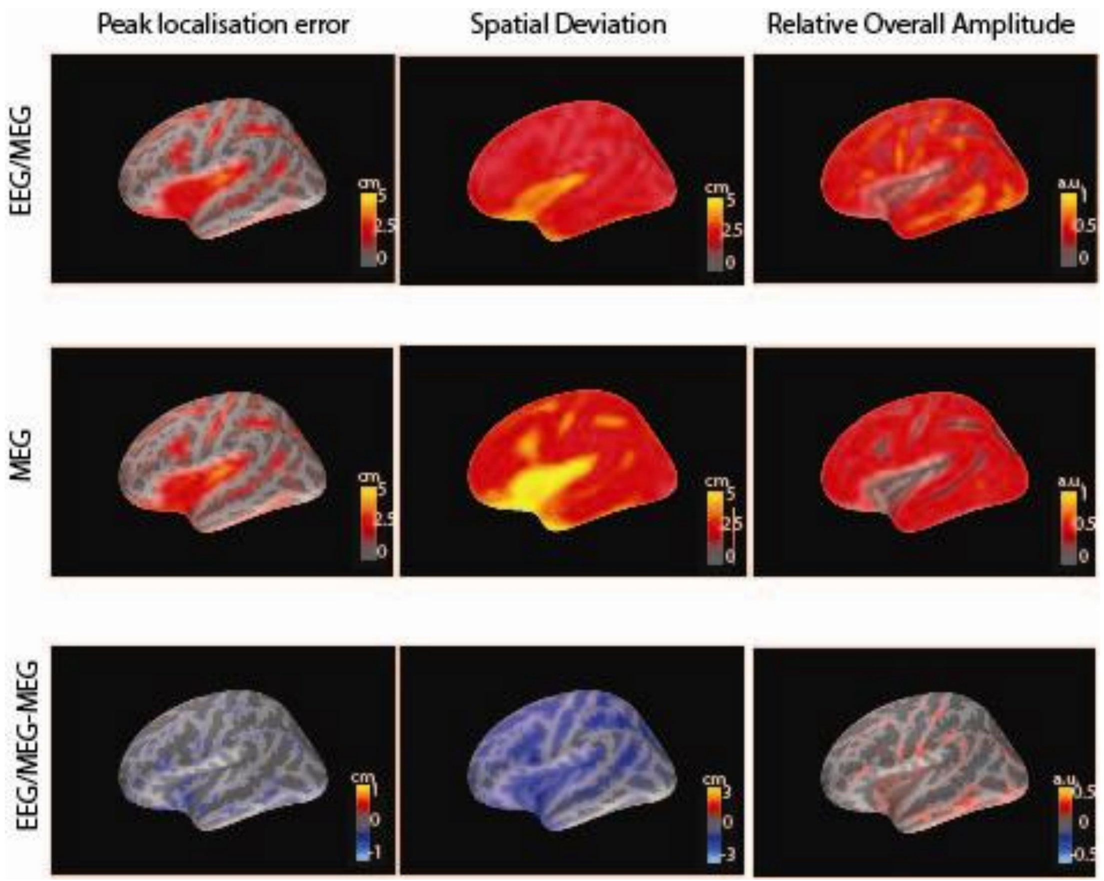
MEG
102 mags + 204 grads



EEG+MEG
102 mags + 204 grads

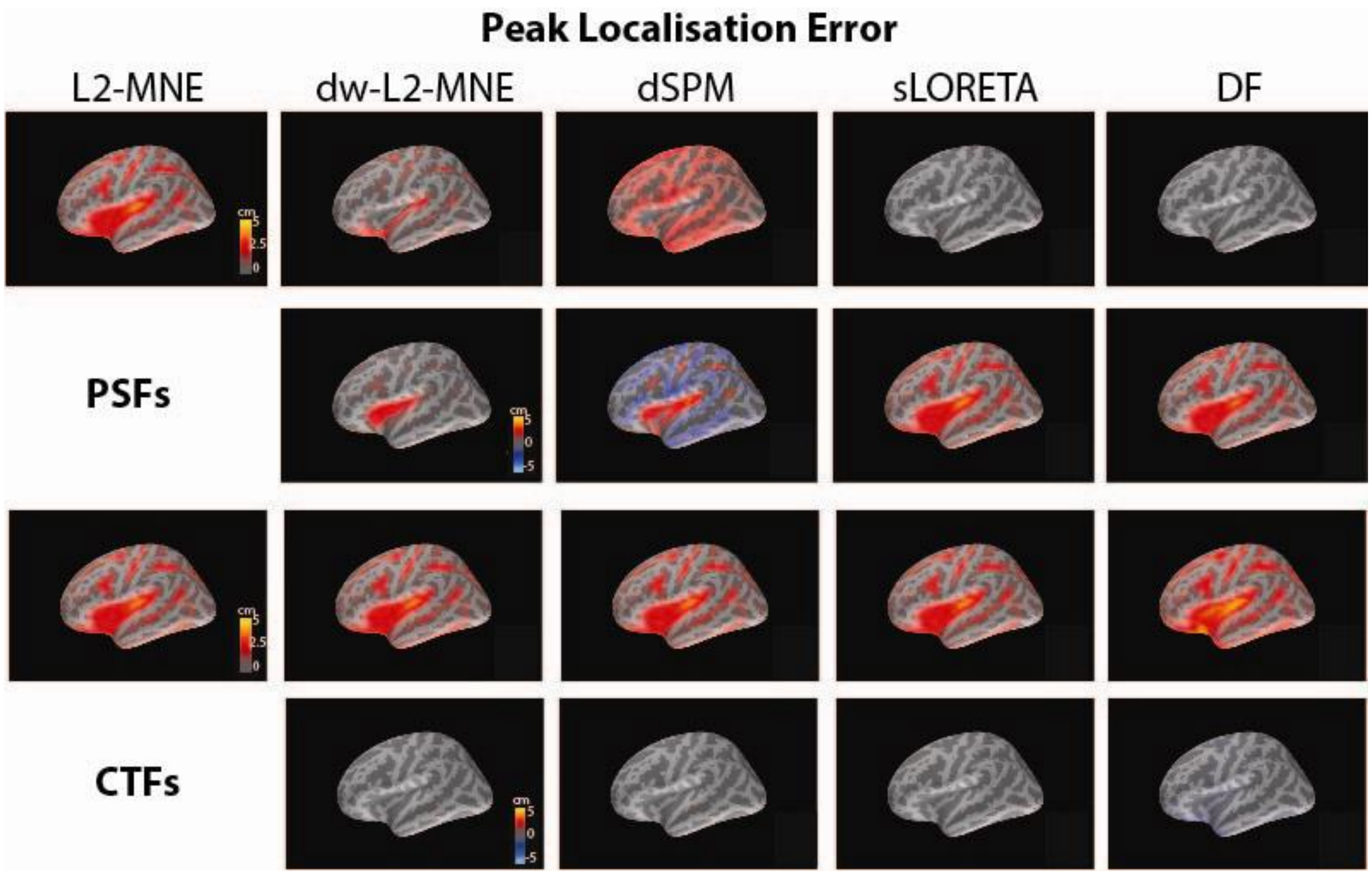


Comparing EEG+MEG and MEG-only



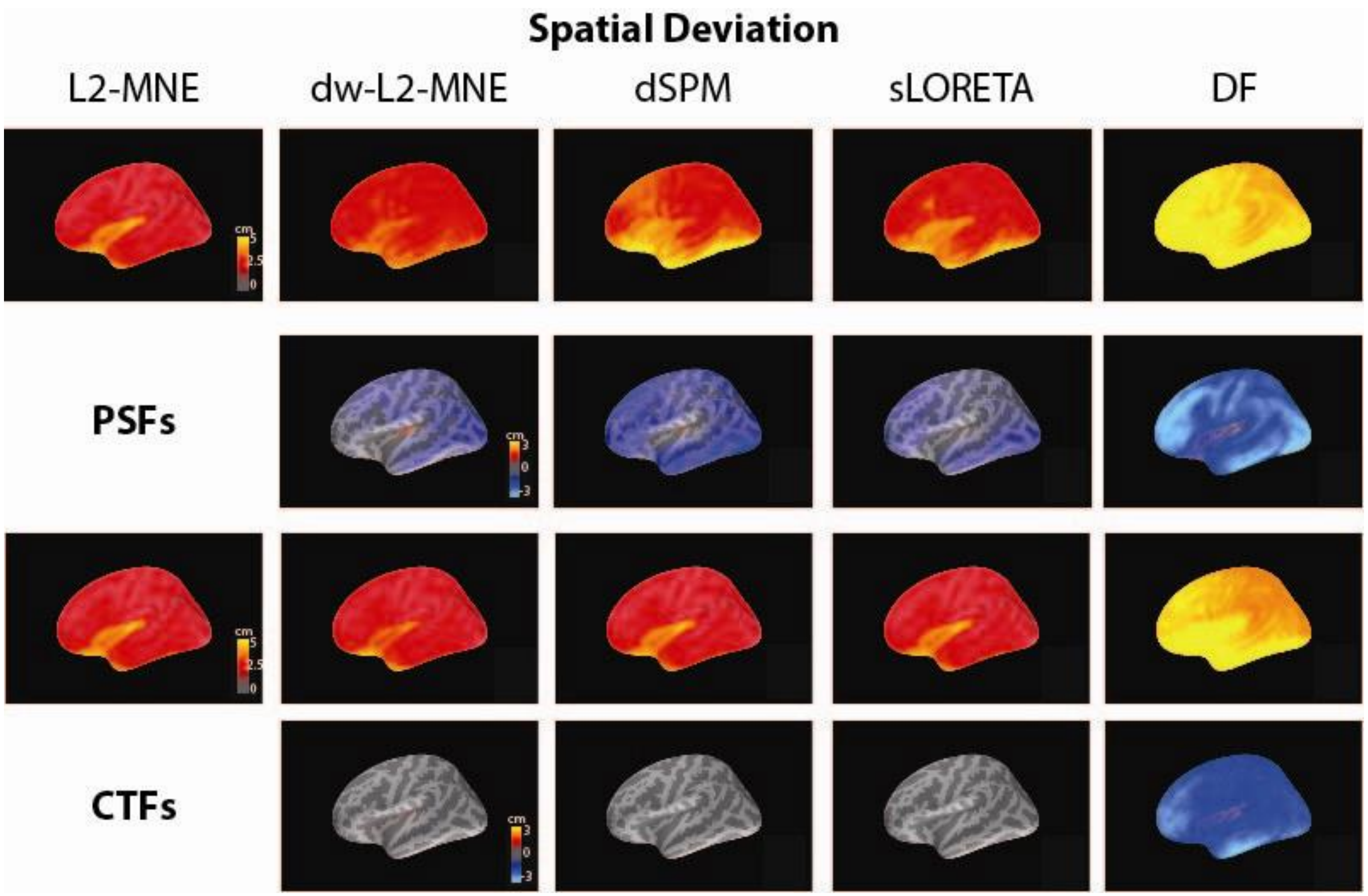
Hauk/Stenroos/Treder, bioRxiv 2019 | see also Molins et al., NI 2008

Comparing Estimators: Localisation Error



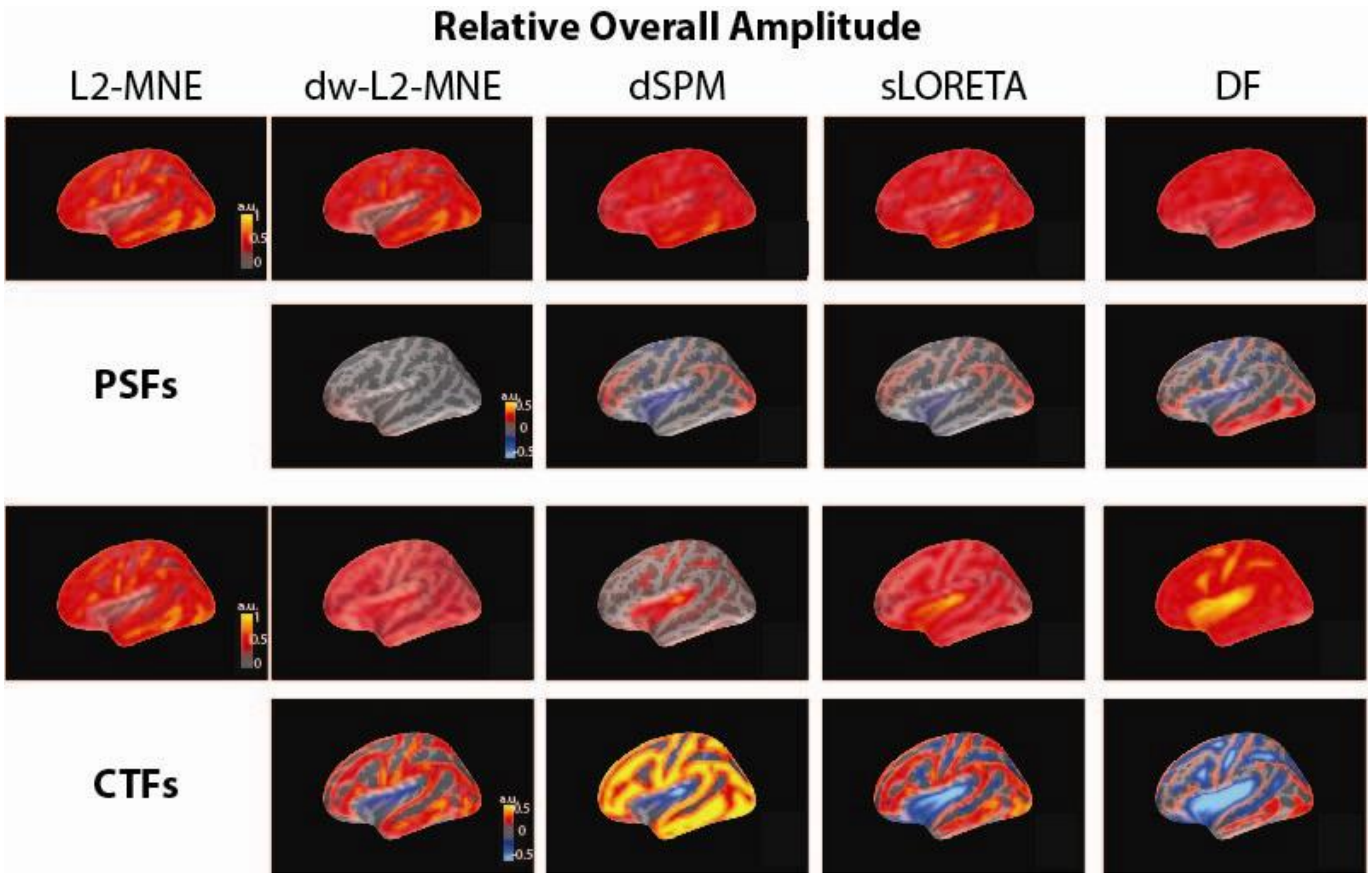
Hauk/Stenroos/Treder, bioRxiv 2019 | see also Hauk/Wakeman/Henson, NI 2011

Comparing Estimators: Spatial Extent



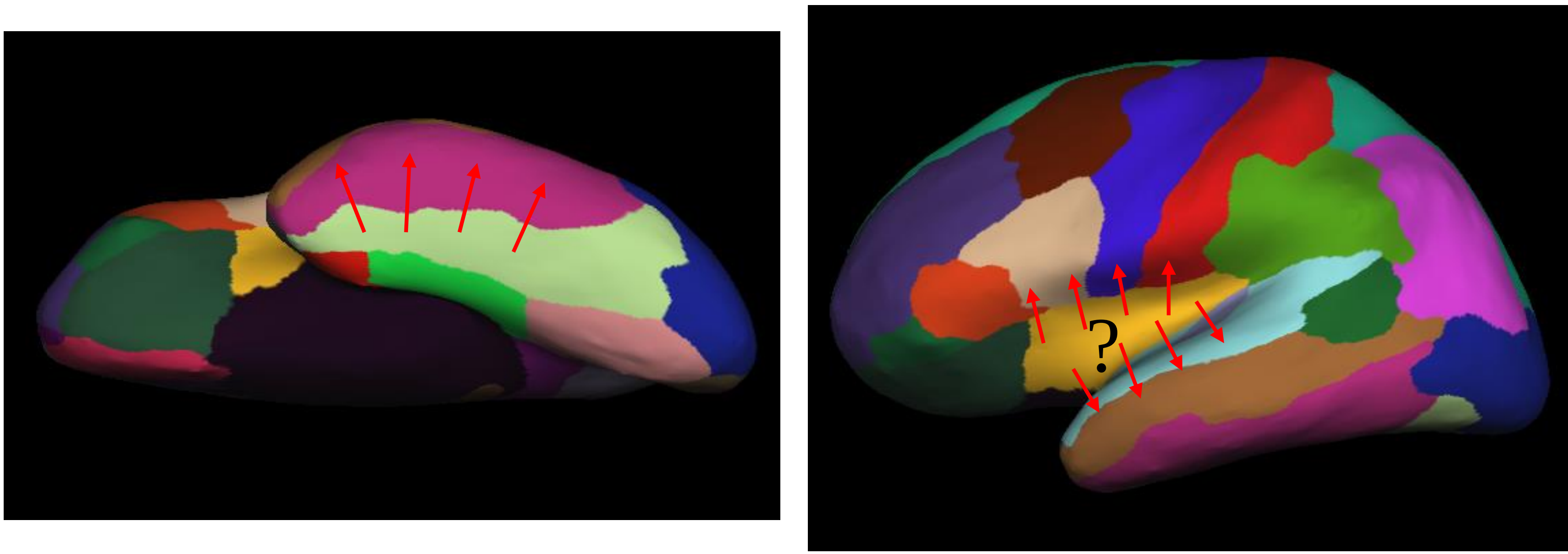
Hauk/Stenroos/Treder, bioRxiv 2019 | see also Hauk/Wakeman/Henson, NI 2011

Comparing Estimators: Relative Amplitude



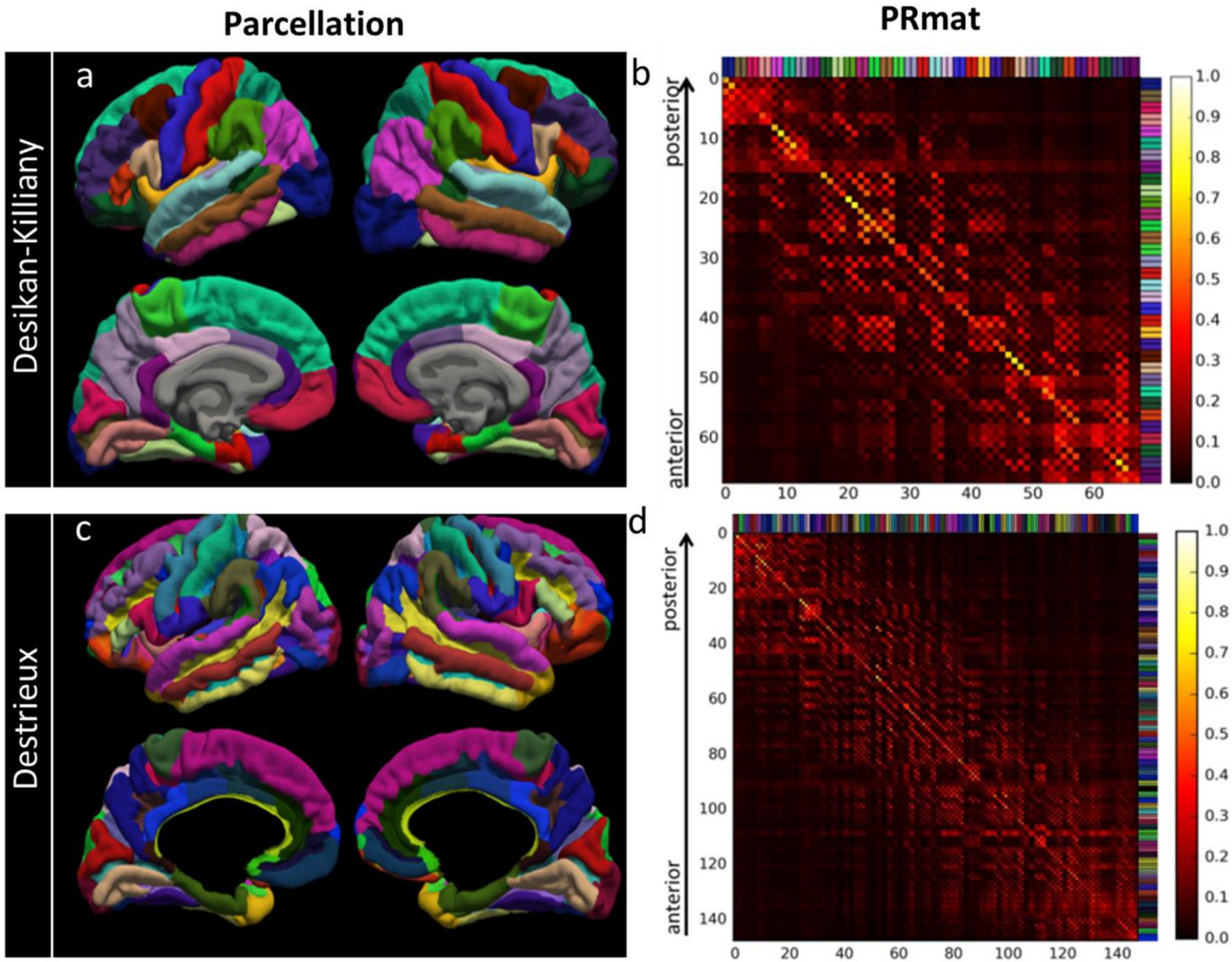
Hauk/Stenroos/Treder, bioRxiv 2019 | see also Hauk/Wakeman/Henson, NI 2011

Localisation Bias Has Consequences for ROI analysis



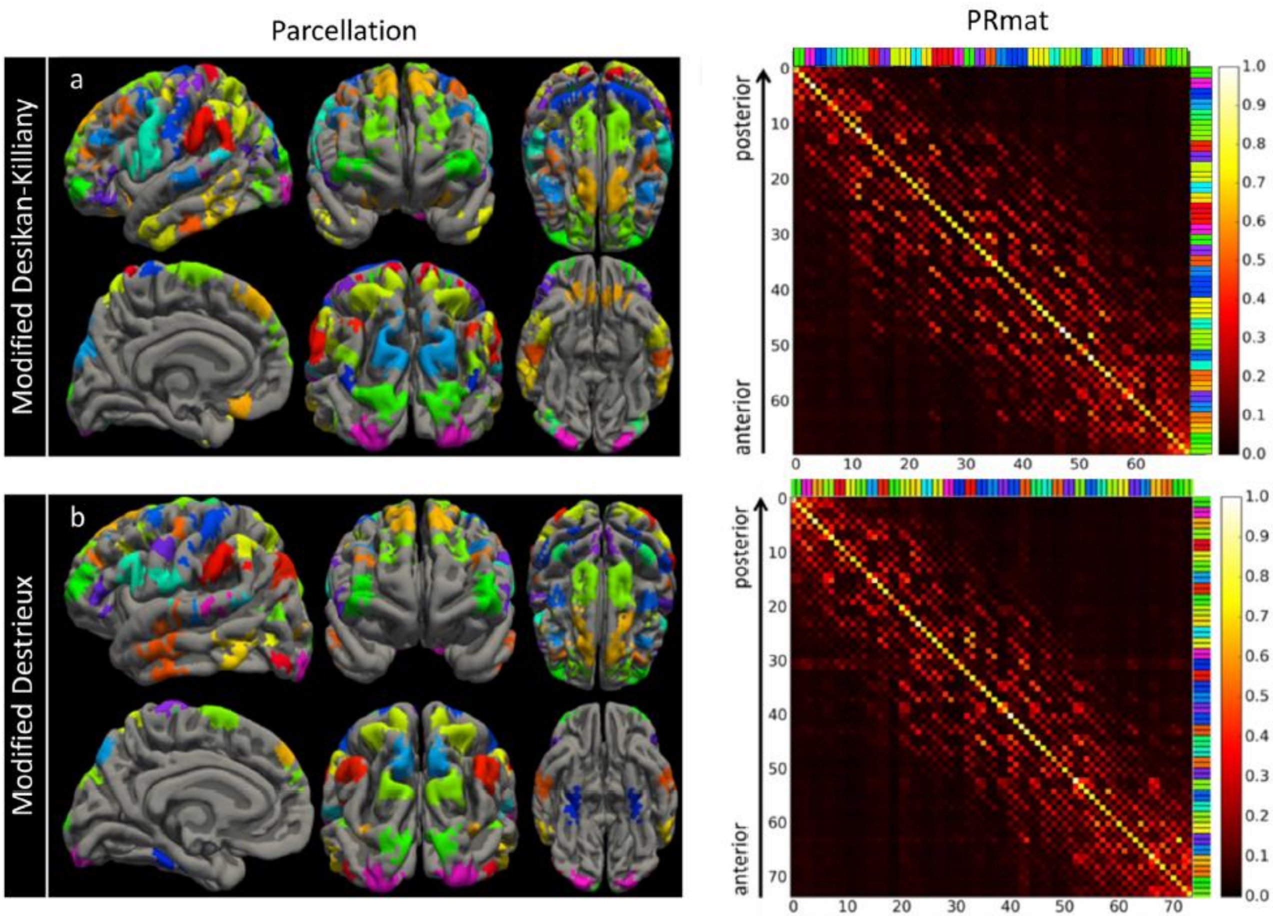
Desikan-Killiany Atlas parcellation

Anatomical Parcellations May Not Be Optimal For EEG/MEG



Farahibozorg, Henson, Hauk, NI 2018

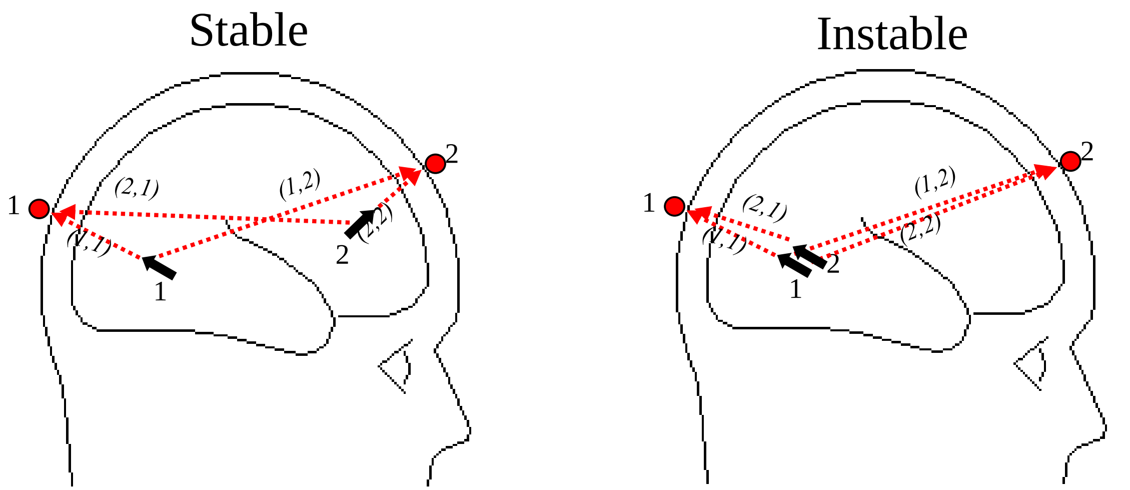
Adaptive Parcellations For EEG/MEG

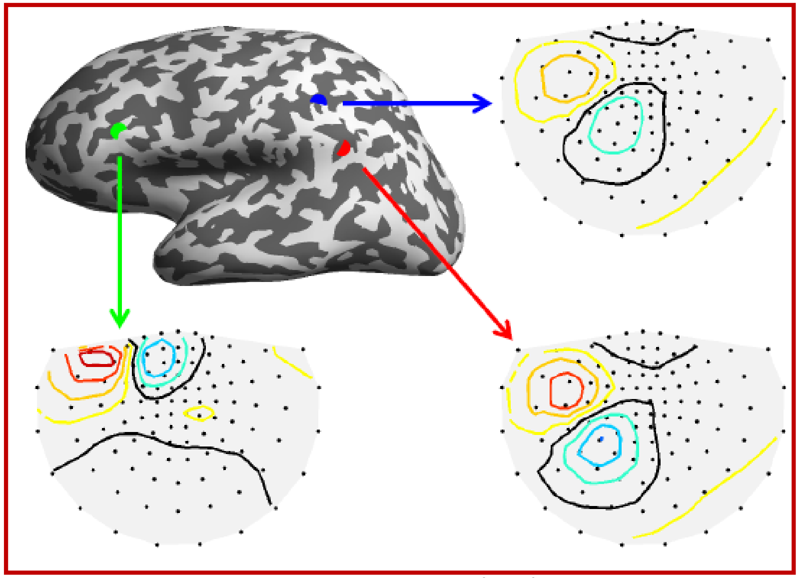


Farahibozorg, Henson, Hauk, NI 2018

Noise and Regularisation

(In)Stability – Sensitivity to Noise





Similar topographies are difficult to distinguish, especially in the presence of noise.

Thanks to Matti Stenroos.

Noise and Regularization

Noise: activity not accounted for by the model.
Hence it depends on the model.

Explaining the data 100% may not be desirable – some of the measured activity is not produced by sources in the model.

Explaining noise may require larger amplitudes in source space then the signal of interest:
Overfitting may seriously distort the solution ("variance amplification" in statistics/regression).

(In)Stability – Sensitivity to Noise

No linear dependence between rows/columns:

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \xrightarrow{Inversion} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

Some linear dependence:

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \xrightarrow{Inversion} \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix}$$

High linear dependence:

$$\begin{pmatrix} 2 & 1.999 \\ 1.999 & 2 \end{pmatrix} \xrightarrow{Inversion} \begin{pmatrix} 500.13 & -499.87 \\ -499.87 & 500.13 \end{pmatrix}$$

Noise covariance

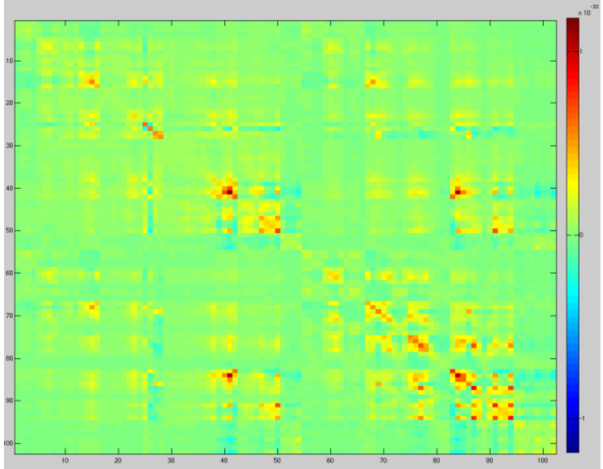
Some channels are noisier than others

⇒They should get different weights in your analysis

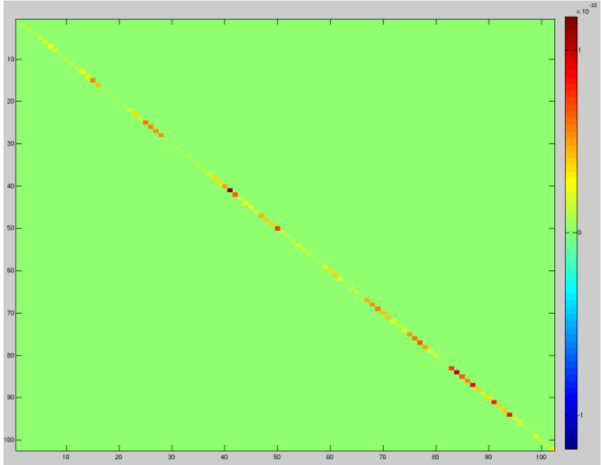
Sensors are not independent

=> Sensors that carry the same information should be downweighted relative to more independent sensors

(Full) Noise Covariance Matrix



(Diagonal) Noise Covariance Matrix
(contains only variance for sensors)



Leaving Variance Unexplained

$$\mathbf{Ls} = \mathbf{d} + \boldsymbol{\varepsilon} \Rightarrow \|\mathbf{Ls} - \mathbf{d}\|^2 <= \text{e, s.t. } \|\mathbf{s}\|_2 = min$$

This is equivalent to minimising the cost function

$$\|\mathbf{Ls} - \mathbf{d}\|^2 + \lambda\|\mathbf{s}\|^2, \lambda>0$$

We can give sensors different weightings,

e.g. based on their noise covariance matrix **C**:

$$\|\mathbf{C}^{-1}(\mathbf{Ls} - \mathbf{d})\|^2 = \|\mathbf{Ls} - \mathbf{d}\|_C^2 = \text{e}$$

$$\|\mathbf{Ls} - \mathbf{d}\|_C^2 + \lambda\|\mathbf{s}\|^2, \lambda>0$$

$$\boldsymbol{G_{MN}} = \mathbf{L}^T(\mathbf{LL}^T + \lambda\mathbf{C}^{-1})^{-1}$$

λ (Lambda) is the regularisation parameter that determines how much variance we want to leave unexplained.

Whitening and Choice of Regularisation Parameter

$$\boldsymbol{G}_{\boldsymbol{MN}} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T + \lambda\mathbf{C}^{-1})^{-1}$$

can also be written as

$$\boldsymbol{G}_{\widetilde{\boldsymbol{MN}}} = \tilde{\mathbf{L}}^T(\tilde{\mathbf{L}}\tilde{\mathbf{L}}^T + \lambda\mathbf{I})^{-1}$$

where $\tilde{\mathbf{L}}$ is the "whitened" leadfield $\mathbf{C}^{-1/2}\mathbf{L}$, and scaled such that trace($\tilde{\mathbf{L}}\tilde{\mathbf{L}}^T$)=trace($\mathbf{I}$).

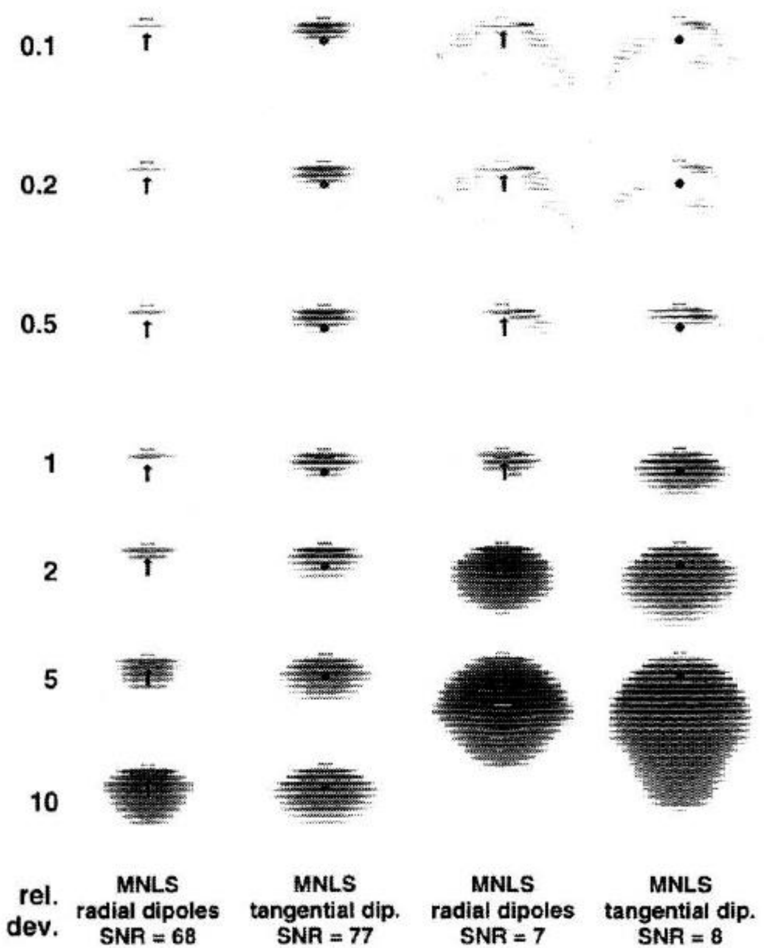
$\tilde{\mathbf{L}}$ and λ can now be interpreted in terms of signal-to-noise ratios.

A reasonable choice for λ is then the approximate SNR of the data (e.g. in MNE software).

Trade-off norm-variance, smoothness

Source at fixed excentricity 71% (60mm)

Regularisation parameter relative to optimal value



Fuchs et al., J Clin Neurophys 1999

Regularisation: Bayesian L2

Minimise cost function

$$F(\mathbf{s}) = \beta \|\mathbf{Ls} - \mathbf{d}\|_C^2 \text{-} 2\log\big(p(\mathbf{s})\big)$$

If we assume $p(\mathbf{s})$ is Gaussian

$$p(\mathbf{s}) = \left(\frac{\alpha}{2\pi}\right)^{\boldsymbol{N/2}} exp\left(-\frac{\alpha}{2}\|\mathbf{s}\|^{2}\right)$$

This leads to the cost function

$$\Rightarrow F(\mathbf{s}) = \beta \|\mathbf{Ls} - \mathbf{d}\|_C^2 + \alpha\|\mathbf{s}\|^2 \sim \|\mathbf{Ls} - \mathbf{d}\|_C^2 + \frac{\alpha}{\beta}\|\mathbf{s}\|^2$$

=> Equivalent to cost function for the L2 minimum-norm solution.

Regularisation: Bayesian L1

Minimise cost function

$$F(\mathbf{s}) = \beta\|\mathbf{Ls} - \mathbf{d}\|_C^2 - 2\log\left(p(\mathbf{s})\right)$$

If we assume $p(\mathbf{s})$ is Laplacian

$$p(\mathbf{s}) = \prod_{j=1}^{N} \frac{1}{2b} exp\left(-\frac{1}{b}\left|\mathrm{s_j}\right|\right)$$

this leads to the cost function

$$\Rightarrow F(\mathbf{s}) = \beta\|\mathbf{Ls} - \mathbf{d}\|_C^2 + \frac{2}{b}\|\mathbf{s}\|_1$$

=> Equivalent to cost function for the L1 minimum-norm solution.