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

EEG/MEG 2:

Head Modelling and Source Estimation

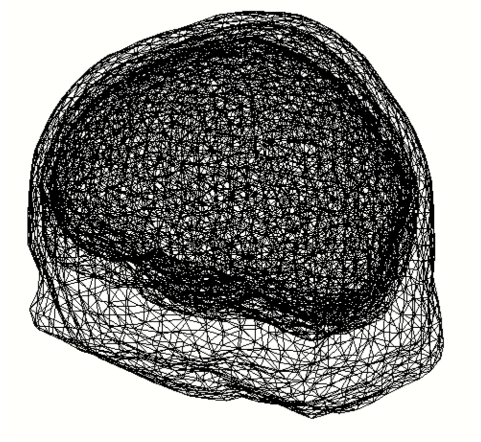
Olaf Hauk

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

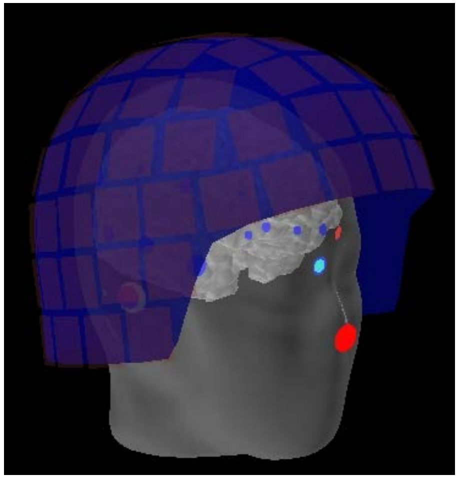
Introduction to Neuroimaging Methods, 25.1.2021

Ingredients for Source Estimation

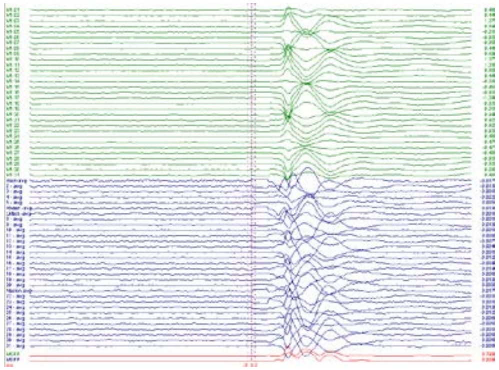
Volume Conductor/
Head Model



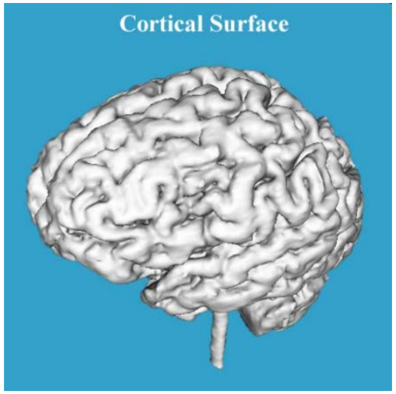
Coordinate
Transformation



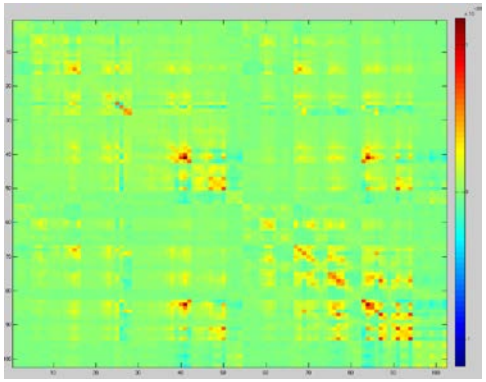
MEG data



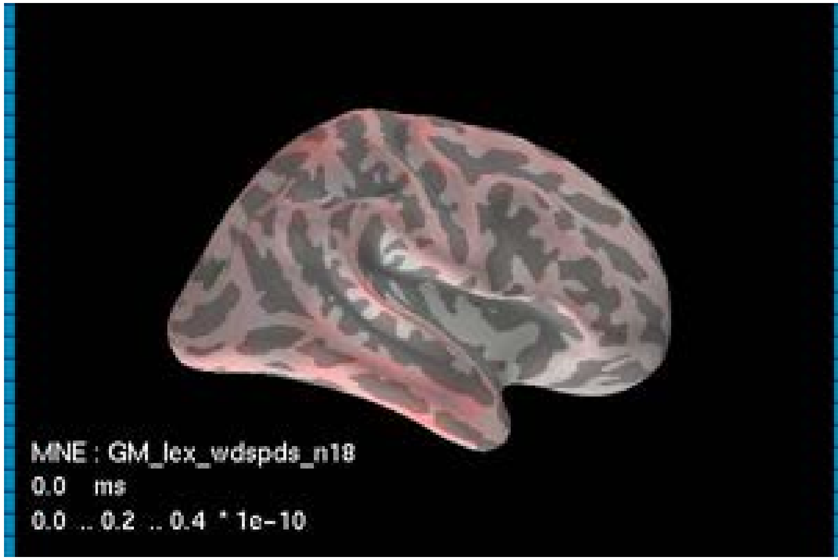
Source Space

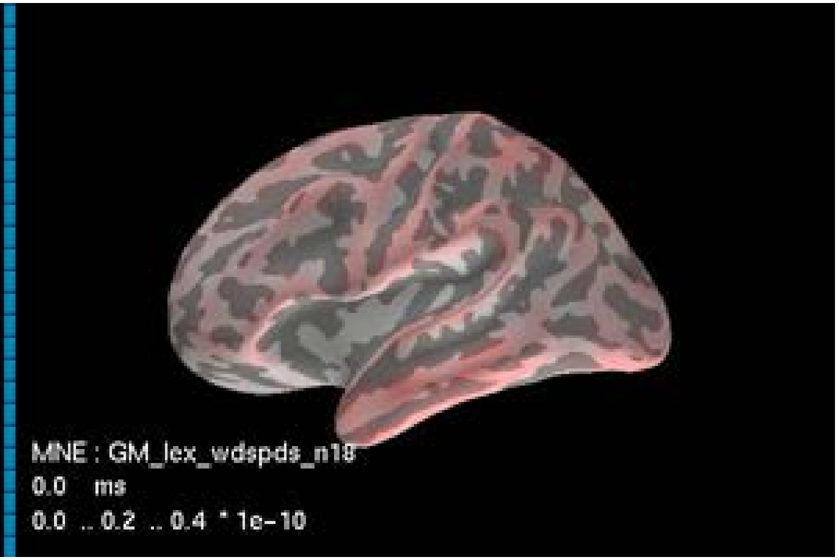


Noise/Covariance Matrix



Our Goal: Spatio-Temporal Brain Dynamics
“Brain Movies”



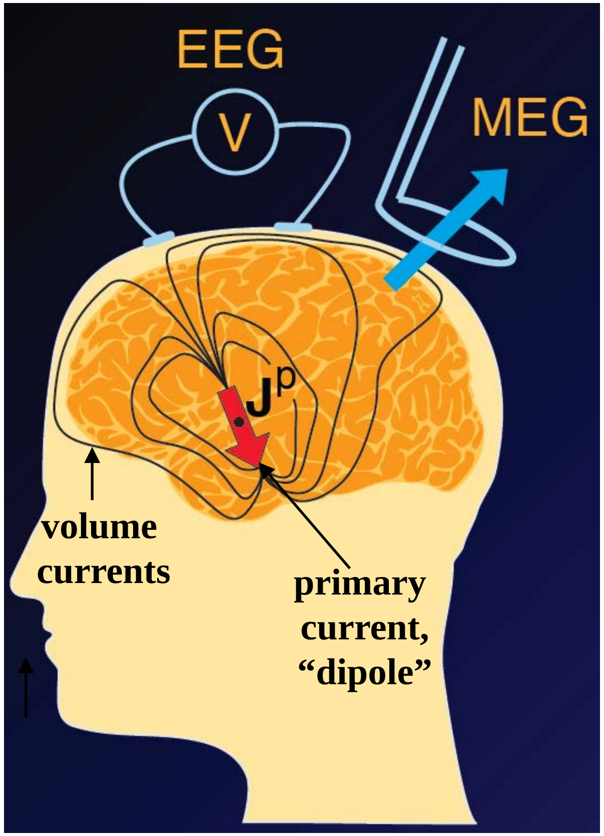


Forward And Inverse Problem

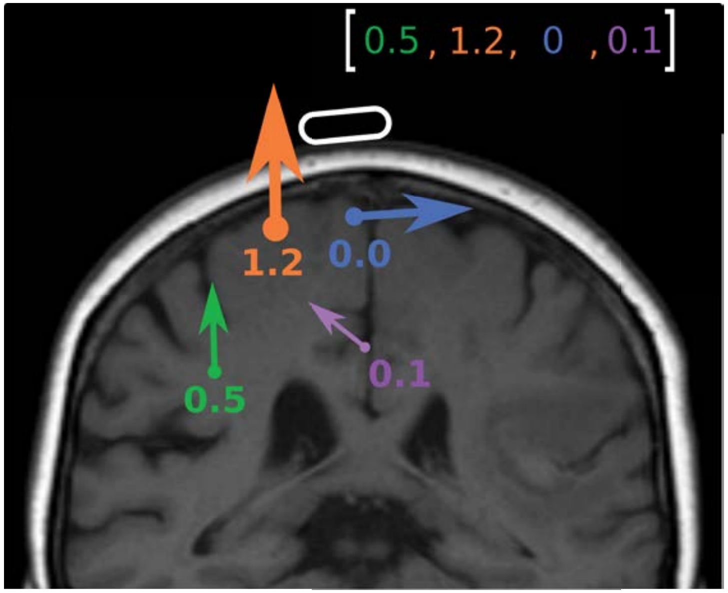
(and some solutions)

The EEG/MEG Forward Problem

EEG/MEG measure the primary sources indirectly



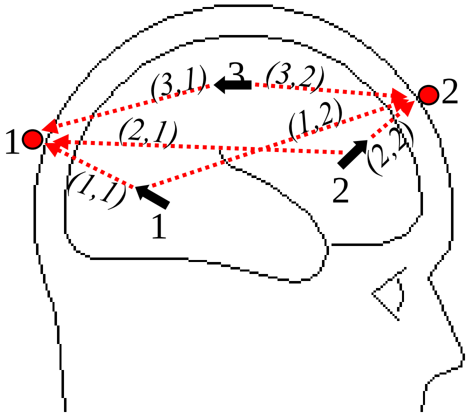
Sensors are differently sensitive to different sources



"Leadfield"

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

We Have To First State The Forward Problem In Order To Solve 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

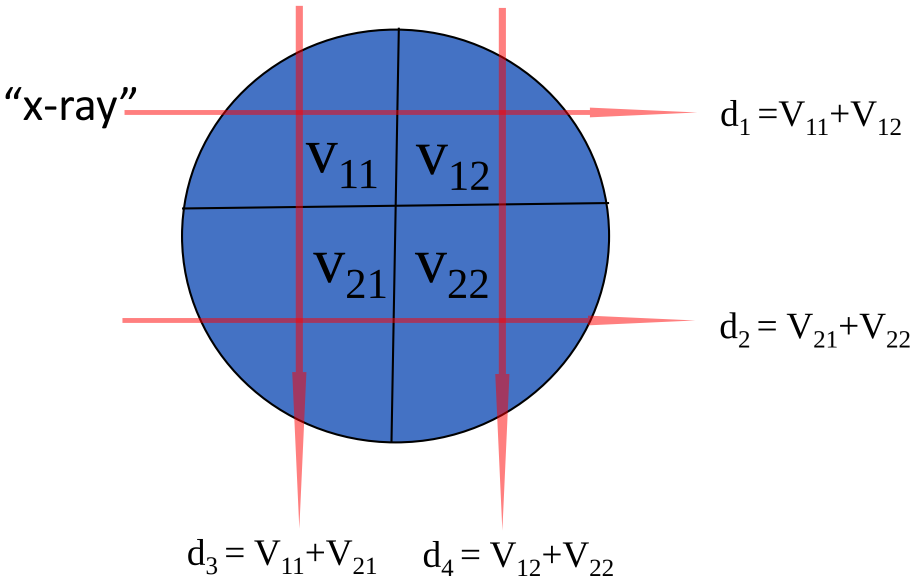
Inverse Operator

dipoles inverse data

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

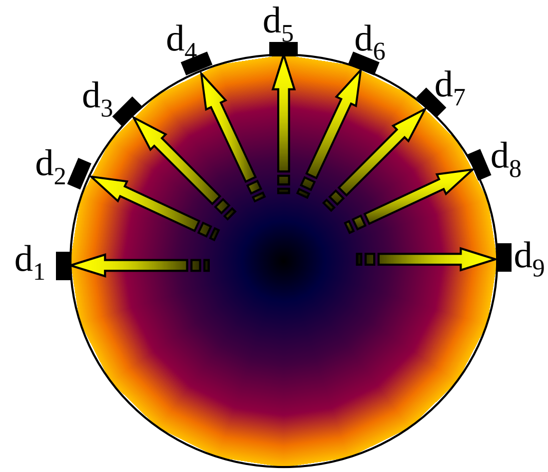
EEG/MEG "Scanning" is not "Tomography"

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



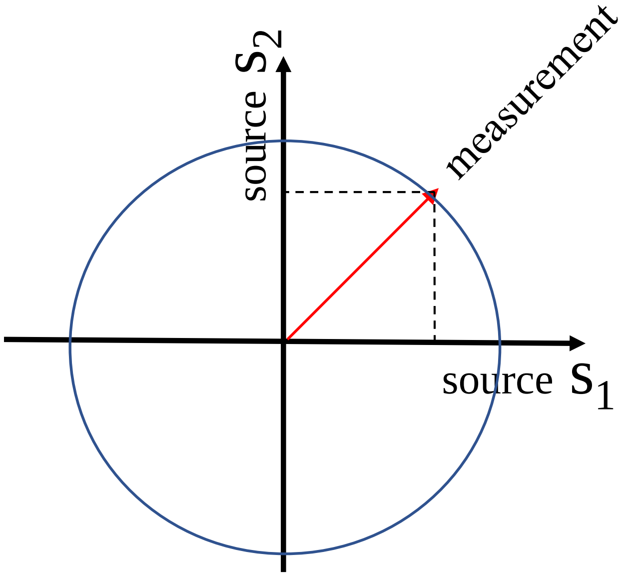
$$d_1 = V_{11}+V_{12}+V_{13}+V_{14} \ldots$$
$$d_2 = V_{21}+V_{22}+V_{23}+V_{24} \ldots$$

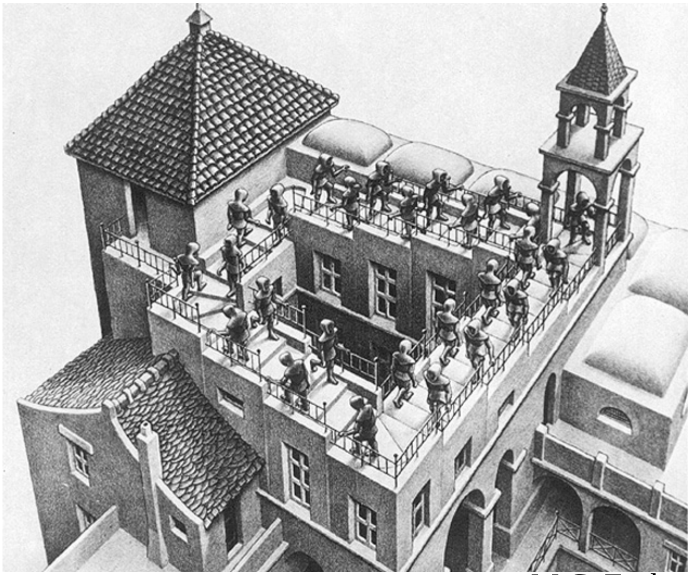
Information is lost during measurement

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

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

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

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

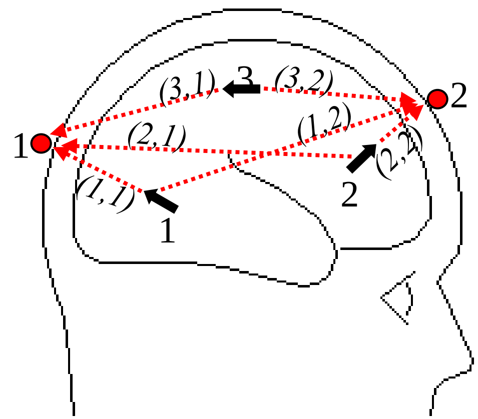
$x_1 = \pi$; $x_2 = (1-\pi)$?

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

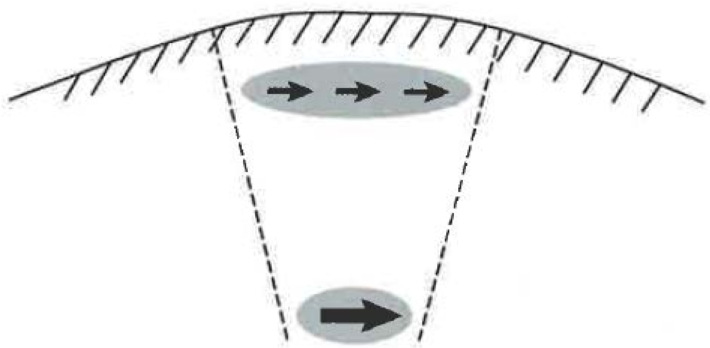
dipoles inverse data

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

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 for approaches for deep sources, https://www.pnas.org/content/114/48/E10465

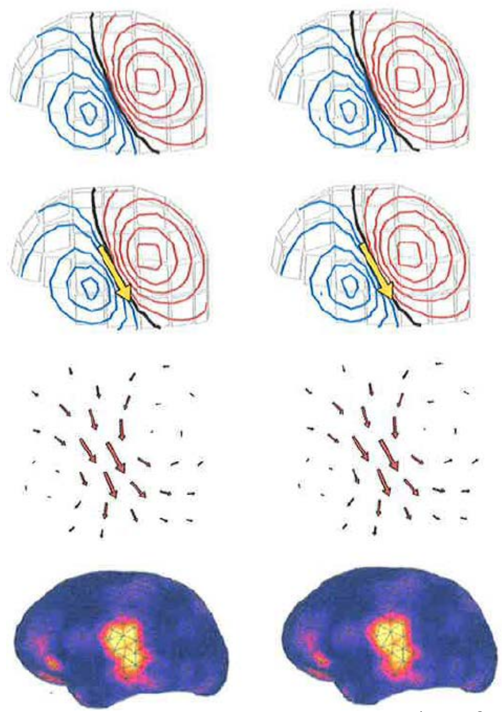
Examples for Non-Uniqueness

[redacted]

Field Patterns

Dipole Model

Minimum Norm Estimates

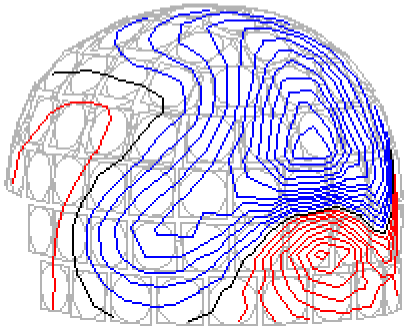


Same Field Patterns

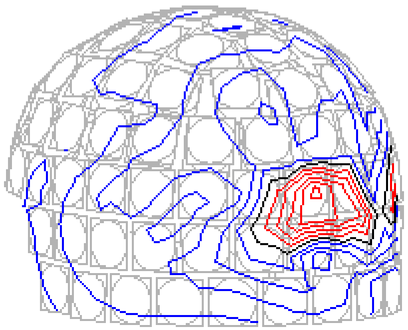
Same Source Estimates

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

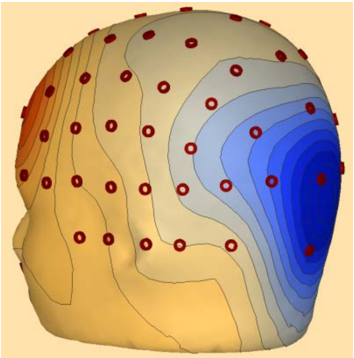
Example: Visually Evoked Activity ~100 ms



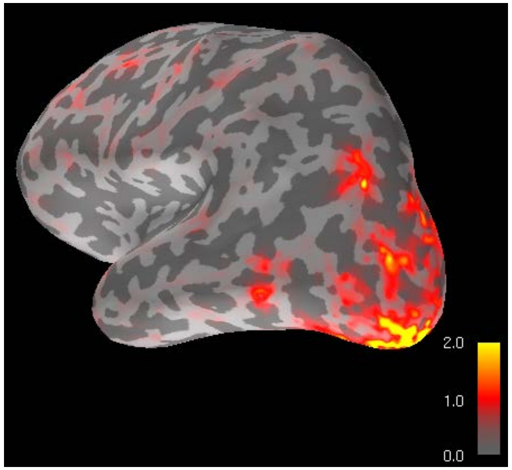
Magnetometers



Gradiometers

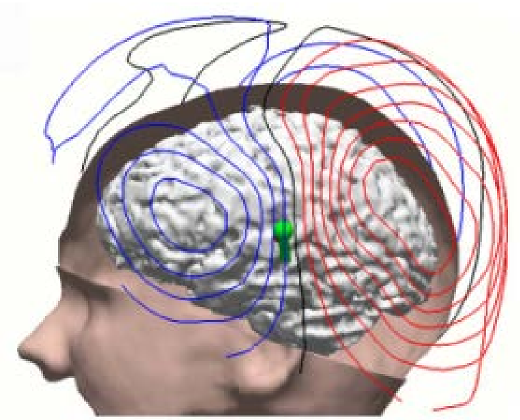


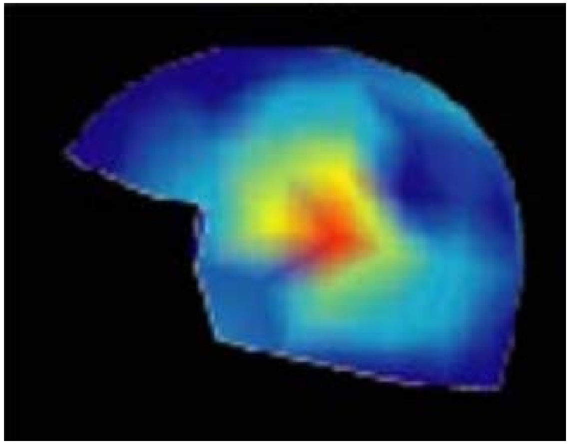
EEG

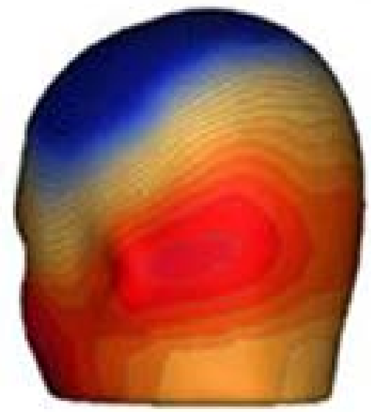


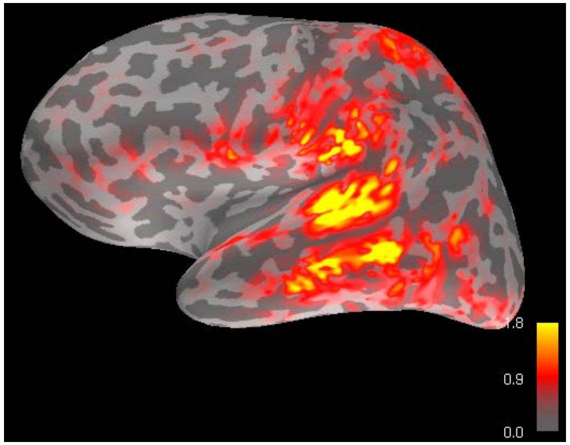
Minimum Norm Estimate

Example: Auditorily Evoked Activity









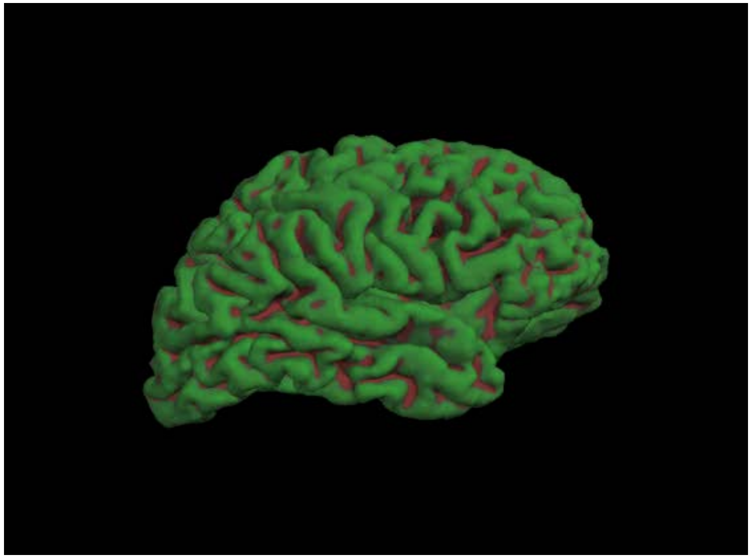
Minimum Norm Estimate

The Forward Problem and Head Modelling

Source Space and Head Model

Source Space

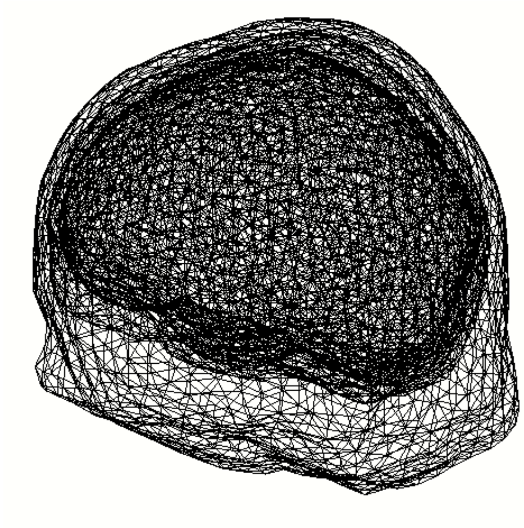
Which sources are modelled
e.g. grey matter, 3D volume



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

Volume Conductor/Head Model

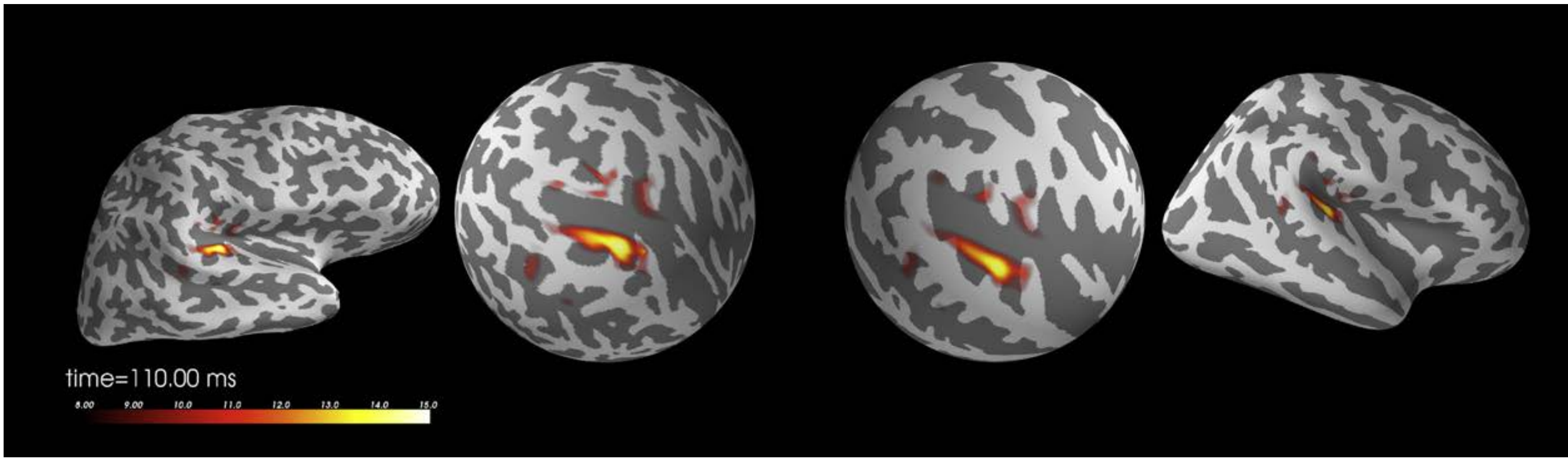
How we model conductivities/currents/potentials/fields in the head
e.g. sphere, 1- or 3-compartments from MRI



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

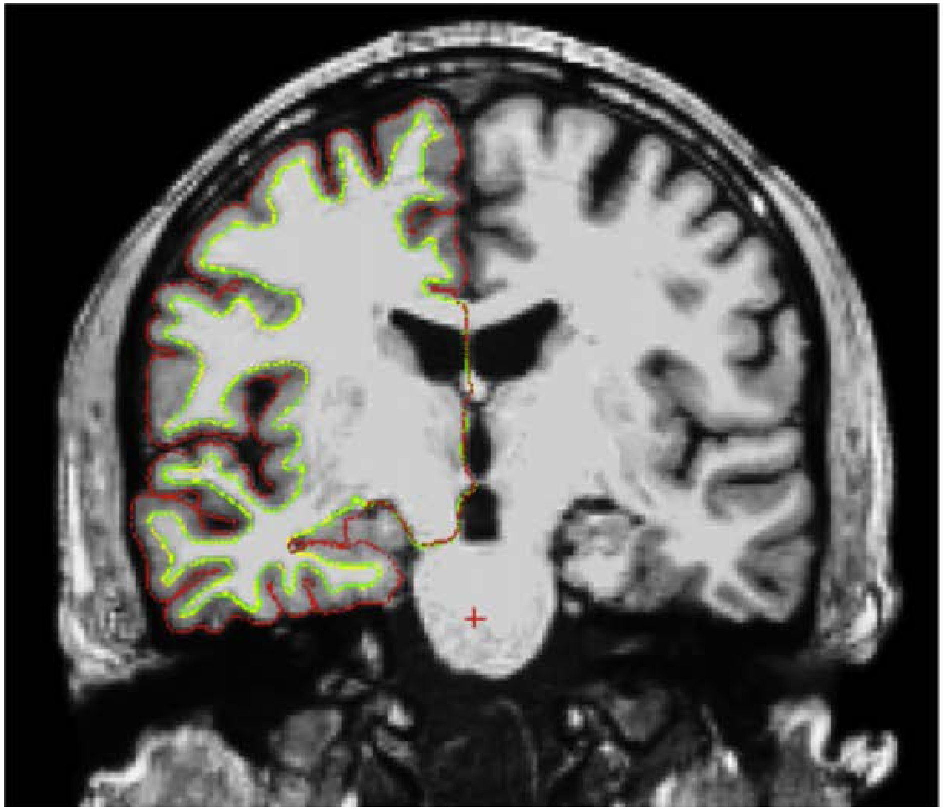
SPM uses the same "canonical mesh" as source space for every subjects, but adjusts it individually.

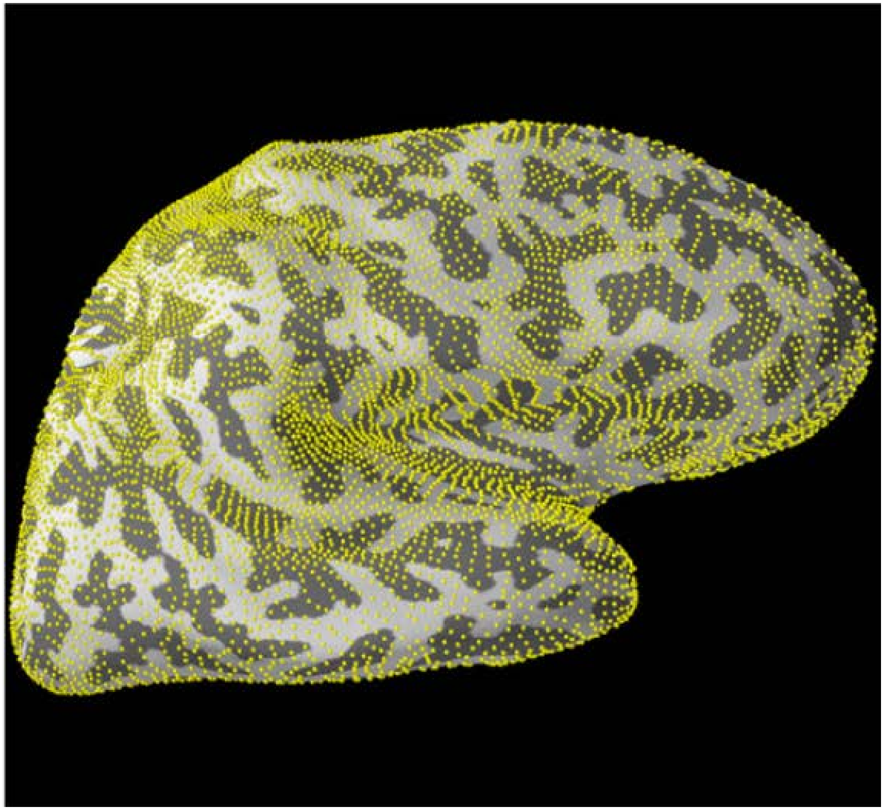
Normalising (Morphing) Cortical Surfaces



Gramfort et al., NI 2014

Source Spaces: Cortical Surface Segmentation

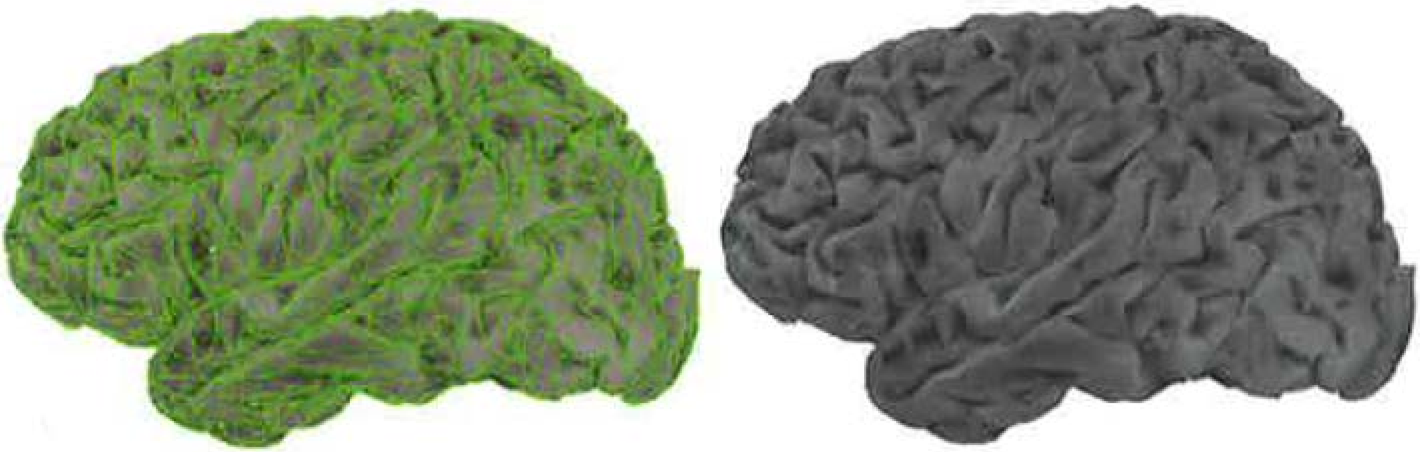




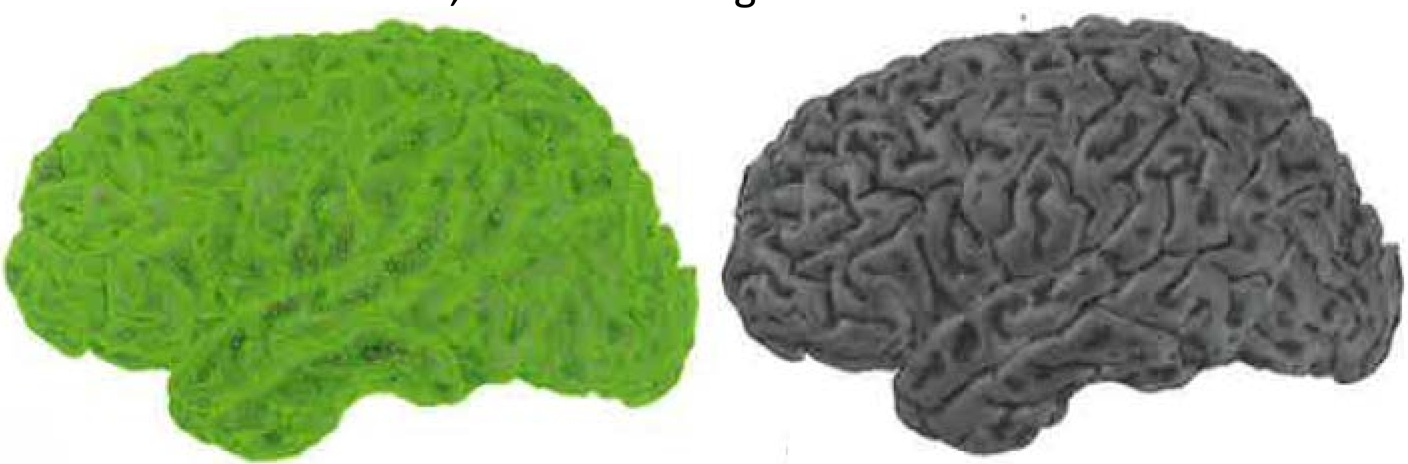
Gramfort et al., NI 2014

Spatial Sampling of Cortical Surfaces

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

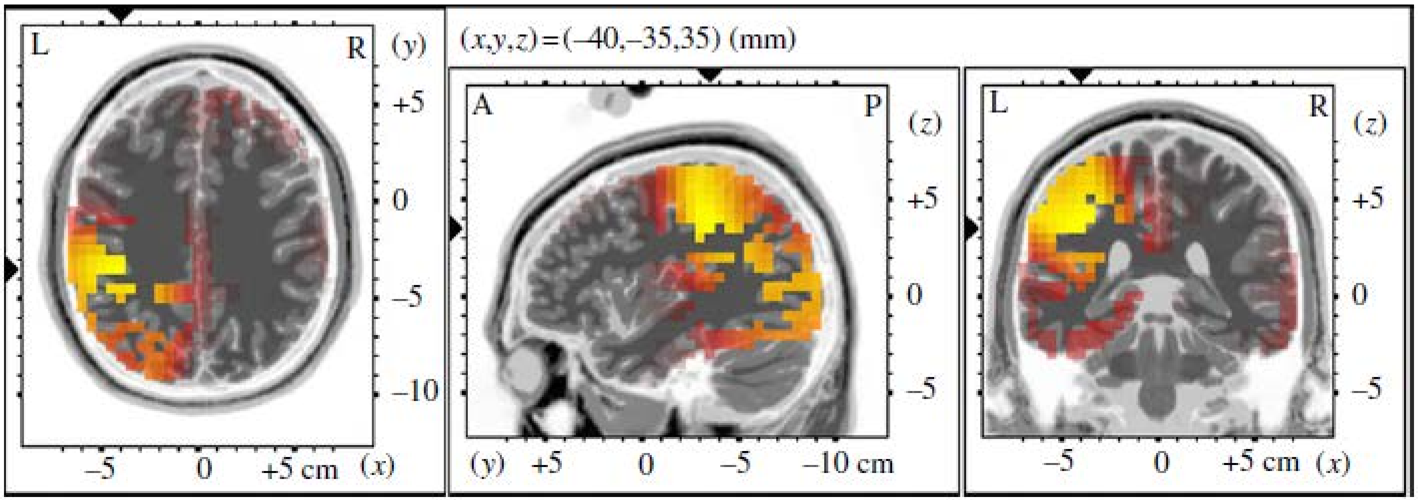


79.124 vertices, 158.456 triangles of 1.3 mm^2 surface area

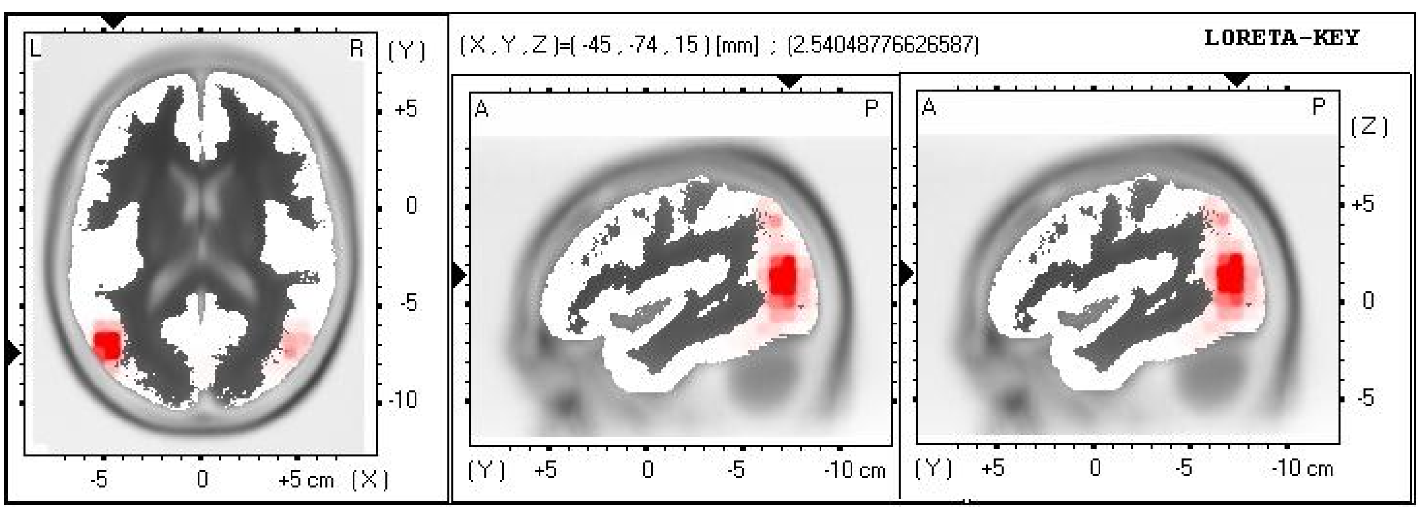


Baillet, chap. 5 in “MEG”, OUP 2010, Hansen/Kringelbach/Salmelin (edts.)

Volumetric Source Spaces

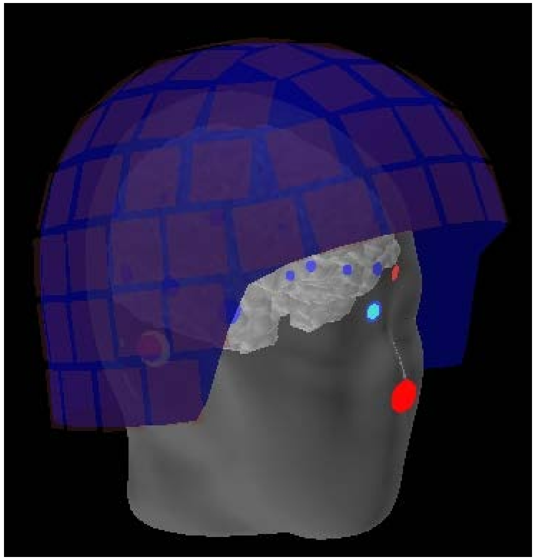


Pascqual-Marqui, PTRS-A 2011

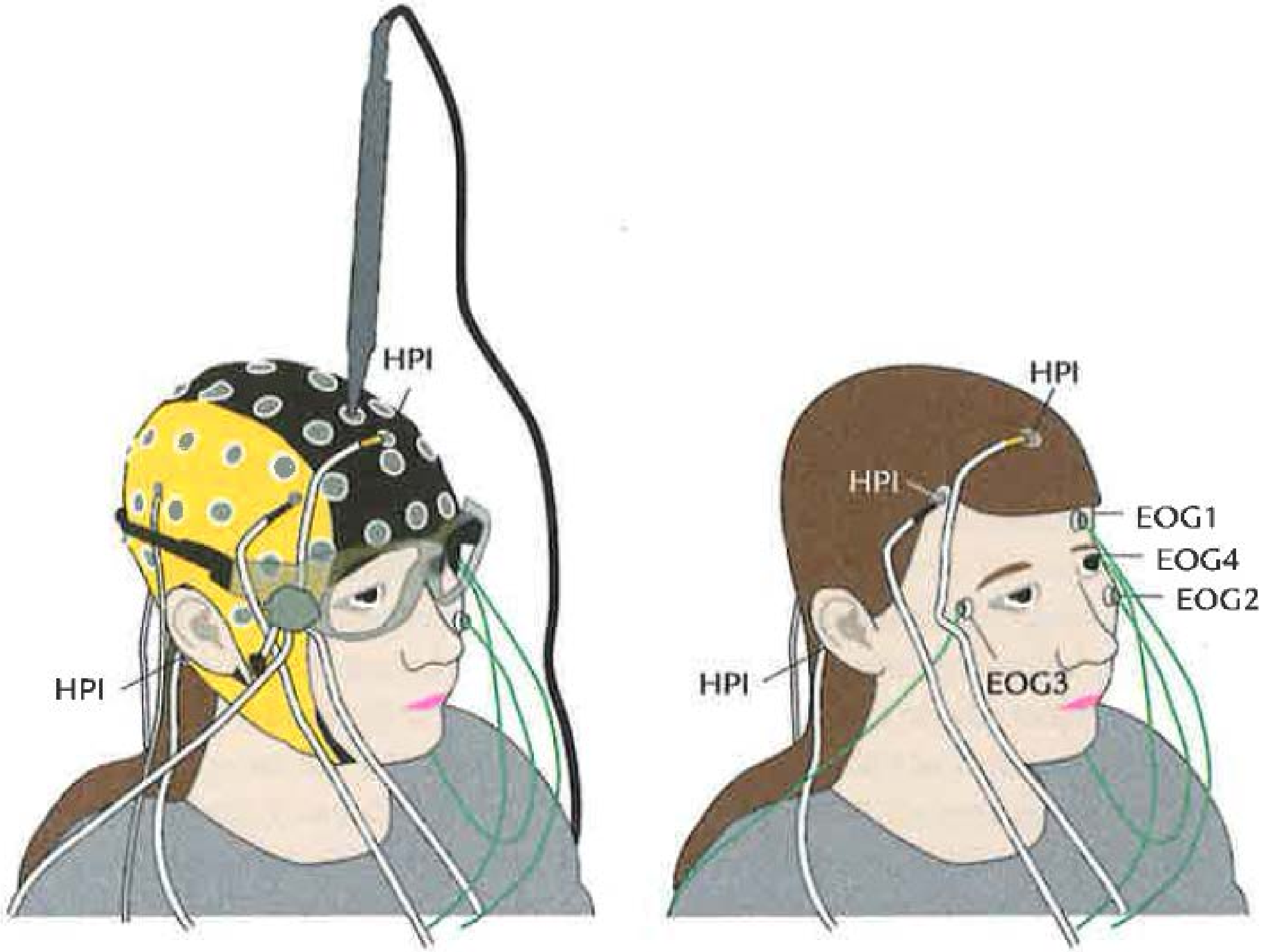


Coregistration of EEG/MEG and MRI Spaces

Coordinate
Transformation



Coregistration of EEG/MEG and MRI Spaces

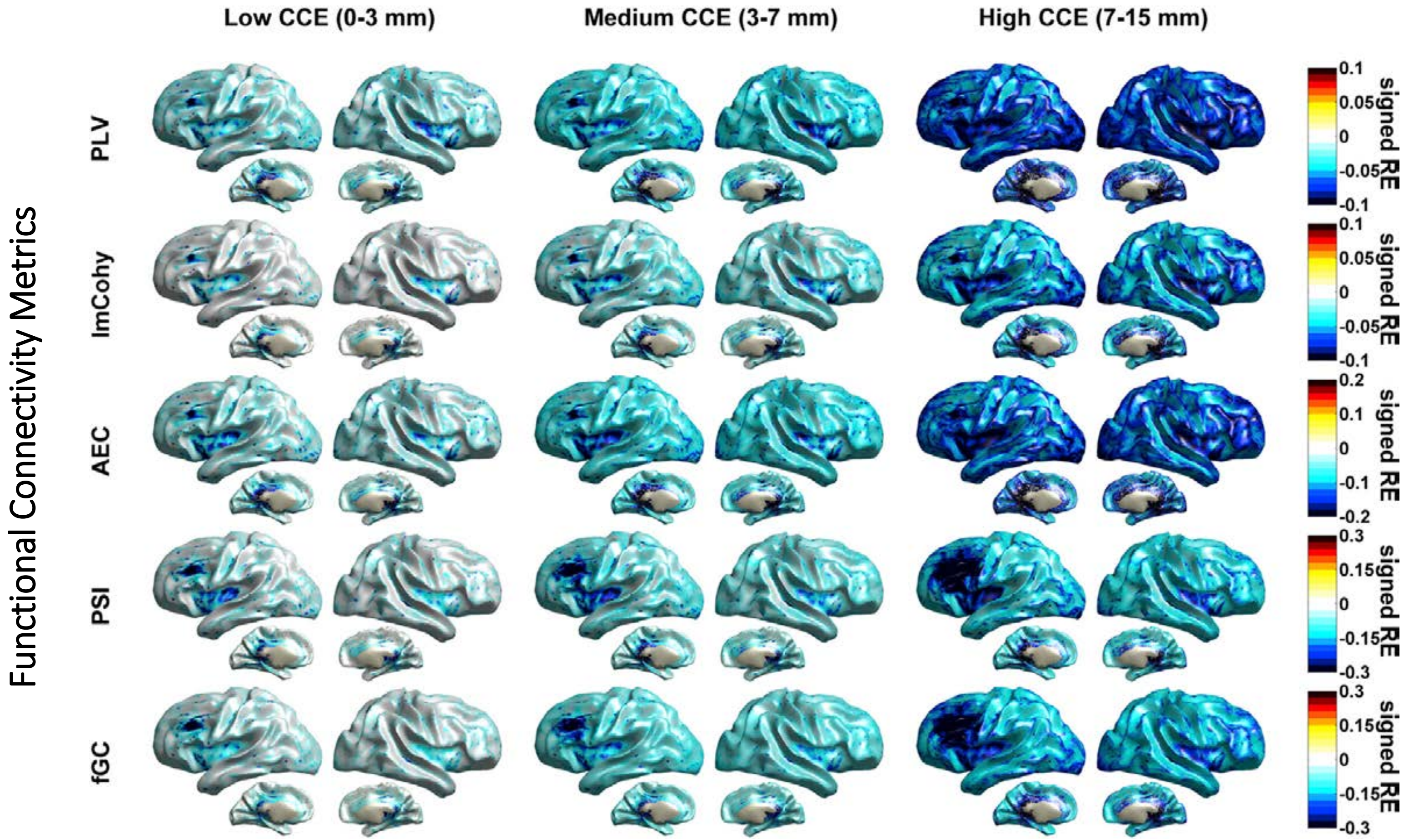


Hari & Puce, “MEG-EEG Primer”, OUP 2017

Accurate Coregistration Is Important

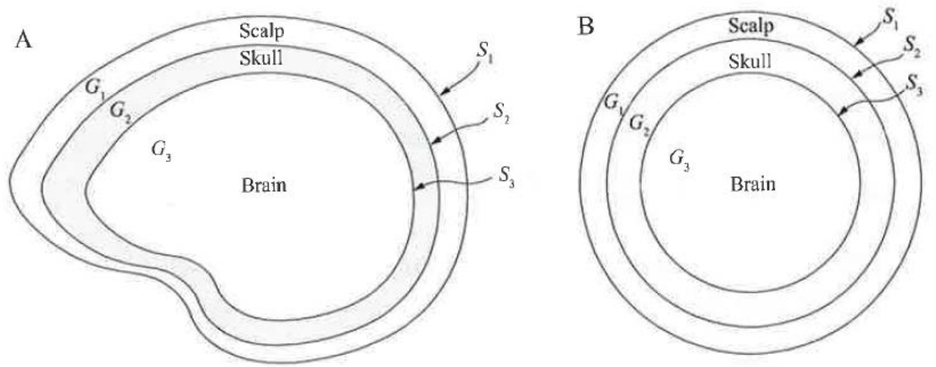
Coregistration errors affect the forward model, and therefore everything that follows.
For example, connectivity analysis:

3 levels of coregistration error



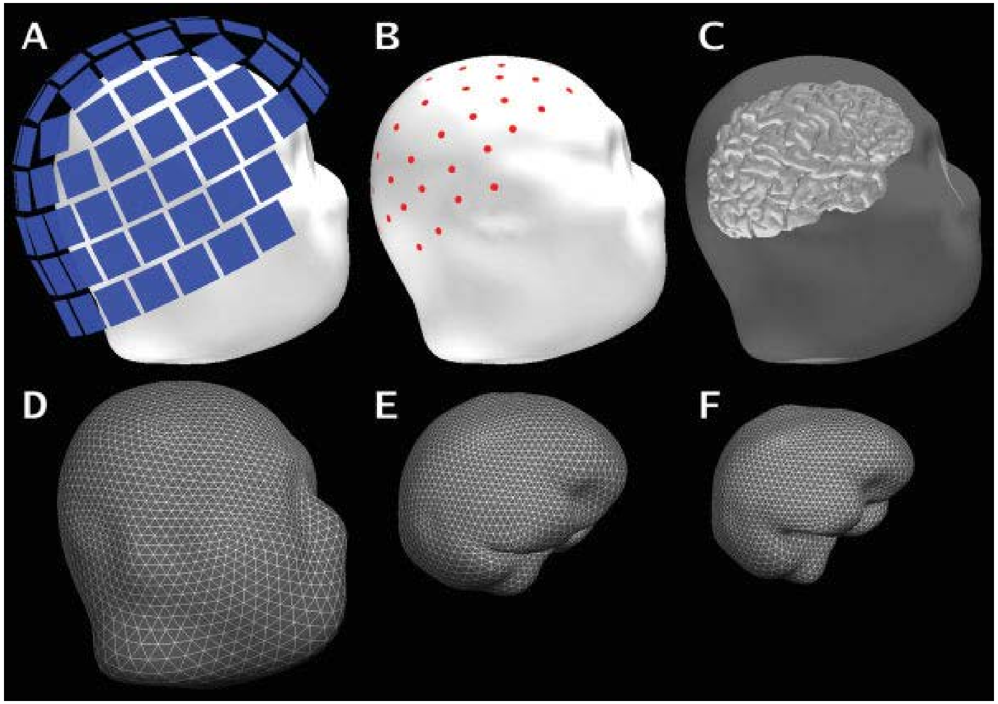
Chella et al., NI 2019

Head Modelling – Tissue Compartments



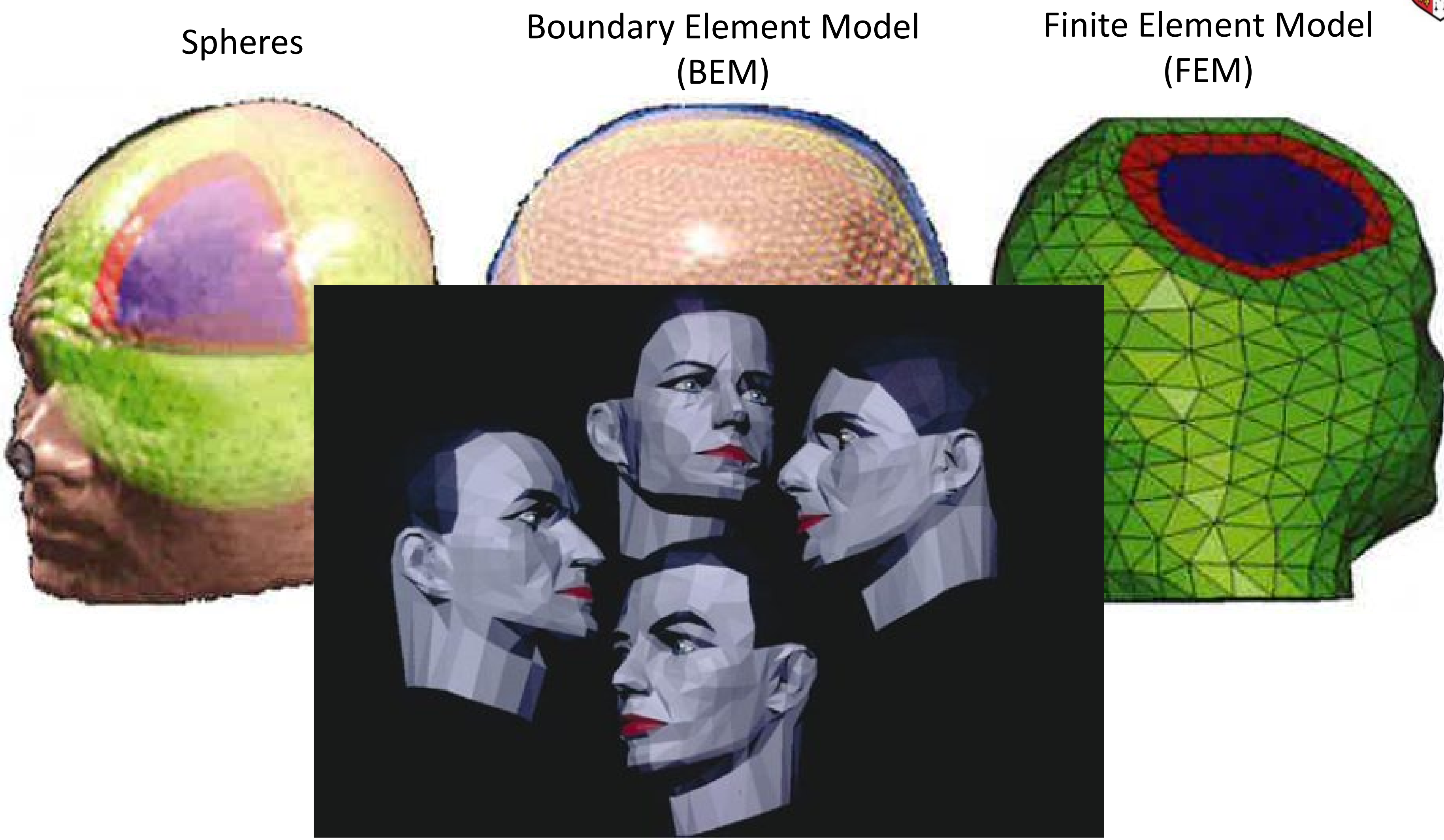
Ilmoniemi and Sarvas, “Brain Signals”, MIT 2019

Ingredients for a head model



Goldenholz et al., HBM 2009

Head Models With Different Levels of Detail



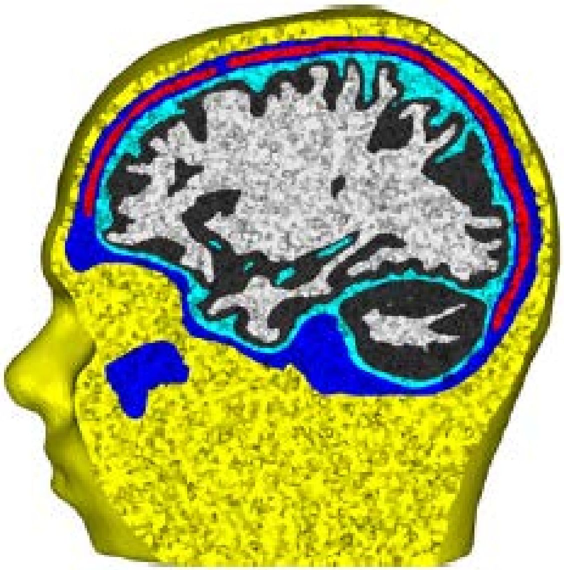
Baillet, chap. 5 in “MEG”, OUP 2010, Hansen/Kringelbach/Salmelin (edts.)

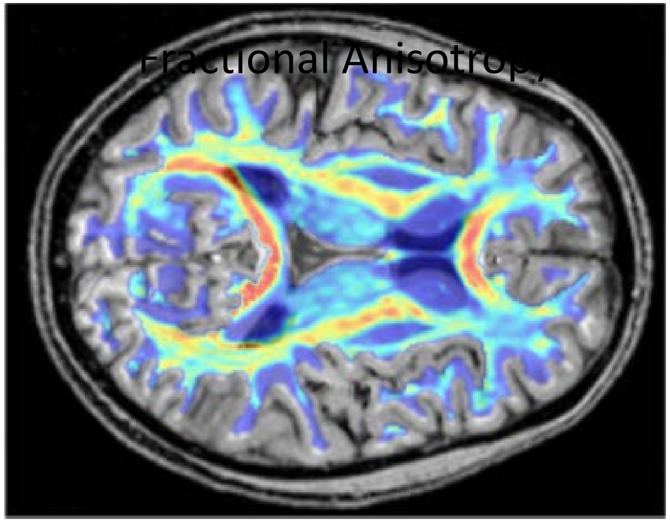
More Complex Head Models

The use of 3-layer (brain, skull, scalp) BEM models based on individual MRI images is recommended for accurate EEG/MEG source reconstruction.
For MEG-only, single shell BEMs and local/corrected sphere models can provide reasonable approximations.
These approaches are available in all major EEG/MEG software packages.

But heads are more complex:

White Matter
Gray Matter
CSF
Skull Compacta
Skull Spongiosa
Skin





Vorwerk et al., NI 2014

It is not obvious how to translate this into more accurate estimate for conductivity distributions.

Conductivities Of Tissues Can Only Be Approximated

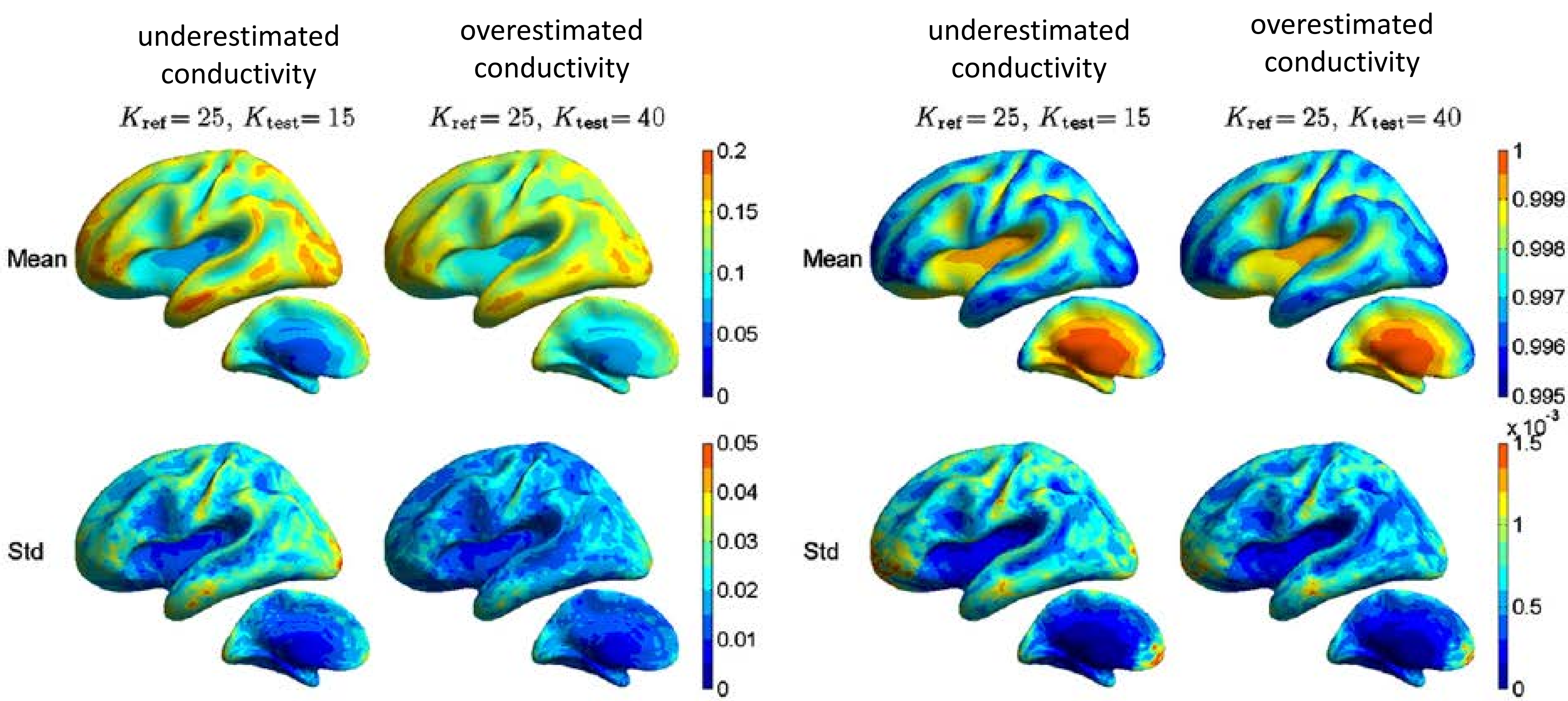
MRC
Cognition and Brain Sciences Unit

Table 2 Isotropic conductivity values of single tissue types used in human head volume conductor modeling

Tissue	Conductivity in S/m	Reference
Brain gray matter	0.45	Logothetis et al. 2007
Brain white matter	0.1	Akhtari et al. 2010
Spinal cord and cerebellum	0.16	Haueisen et al. 1995
Cerebrospinal fluid	1.79	Baumann et al. 1997
Hard bone (compact bone)	0.004	Tang et al. 2008
Soft bone (spongiform bone)	0.02	Akhtari et al. 2002
Blood	0.6	Gabriel et al. 2009
Muscle	0.1	Gabriel et al. 1996, 2009
Fat	0.08	Gabriel et al. 2009
Eye	1.6	Pauly and Schwan 1964; Lindenblatt and Silny 2001
Scalp	0.43	Geddes and Baker 1967
Soft tissue	0.17	Haueisen et al. 1995
Internal air	0.0001	Haueisen et al. 1995

Haueisen & Knösche, Chapter of “Magnetoencephalography” by Supek & Aine (edts.)

Boundary Element Models Are Relatively Robust Against Conductivity Errors

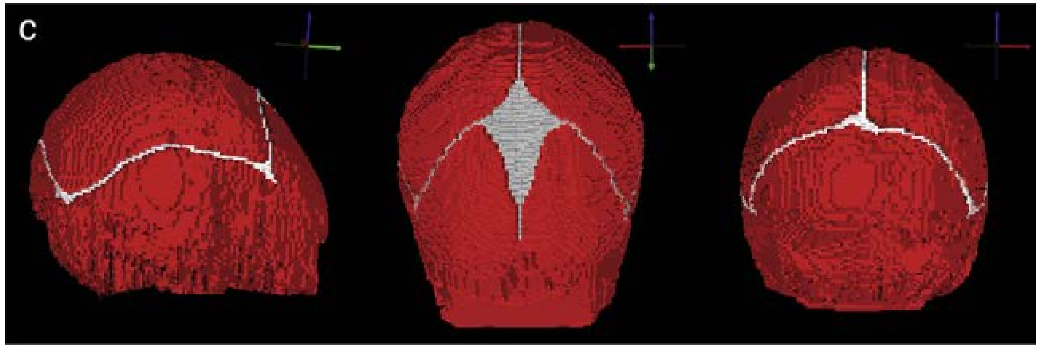


Relative Error

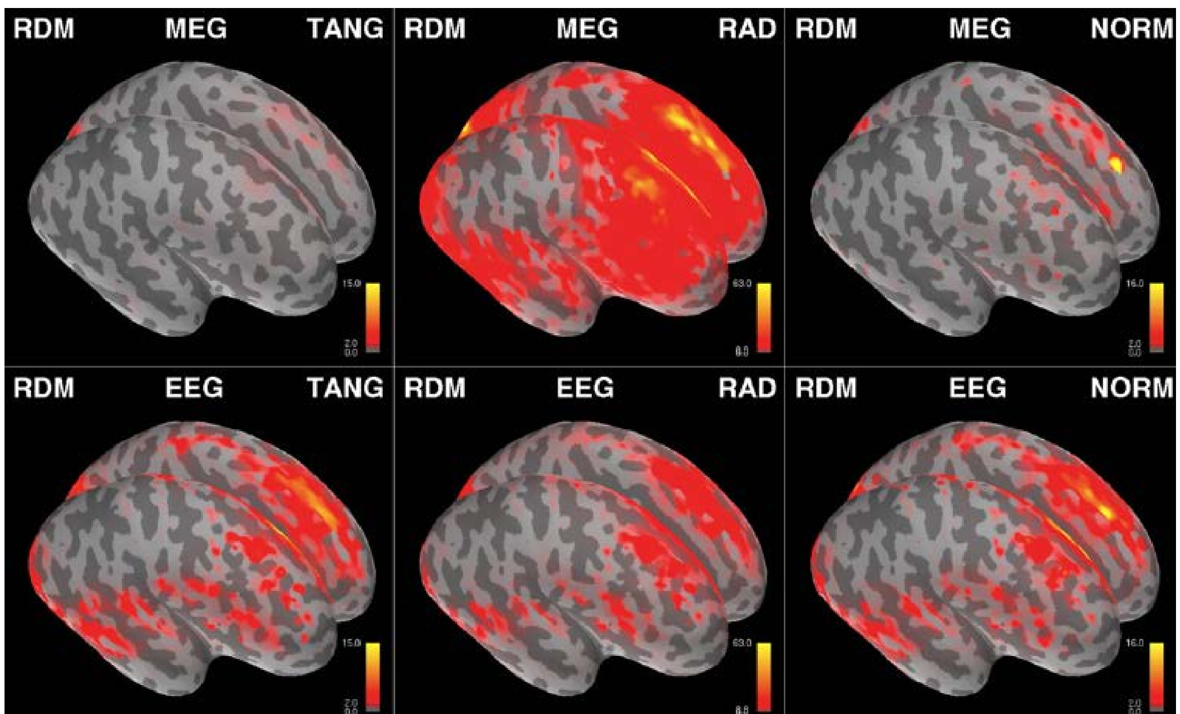
Correlation Coefficient

Stenroos & Hauk, NI 2013
https://pubmed.ncbi.nlm.nih.gov/23639259/

Infant Skulls – Fontanelles and Sutures



Relative error between models with and without fontanelles/sutures



Lew et al., NI 2013

Conclusion – Head Modelling

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

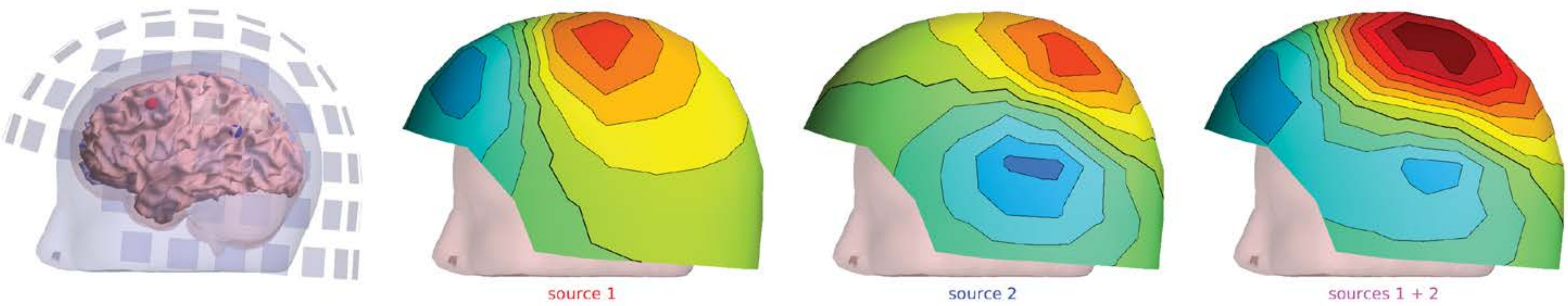
Single-shell approximations are still common for MEG.

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.

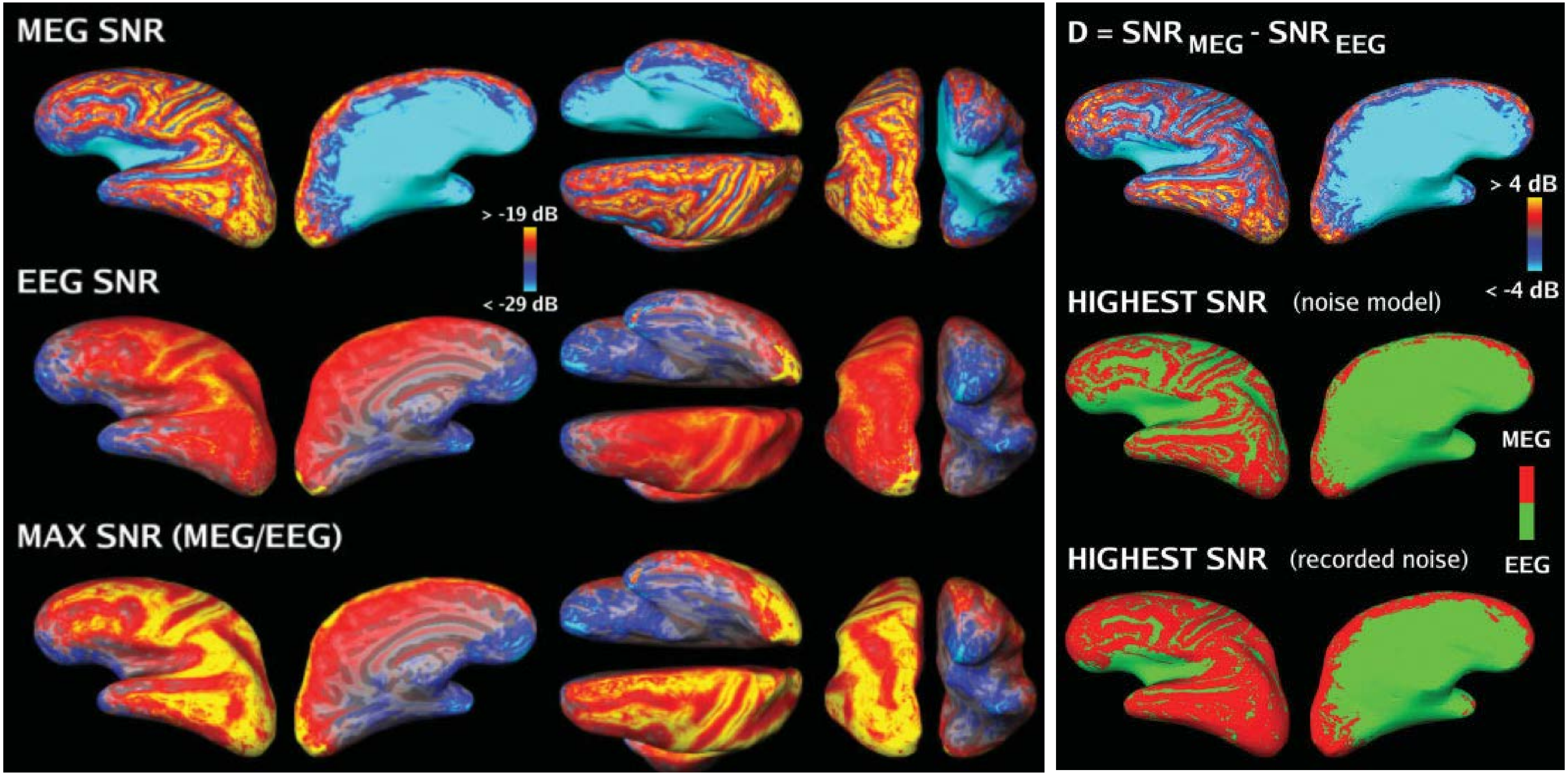
The Forward Problem Is Linear – Superposition Principle

Superposition In Sensor Space



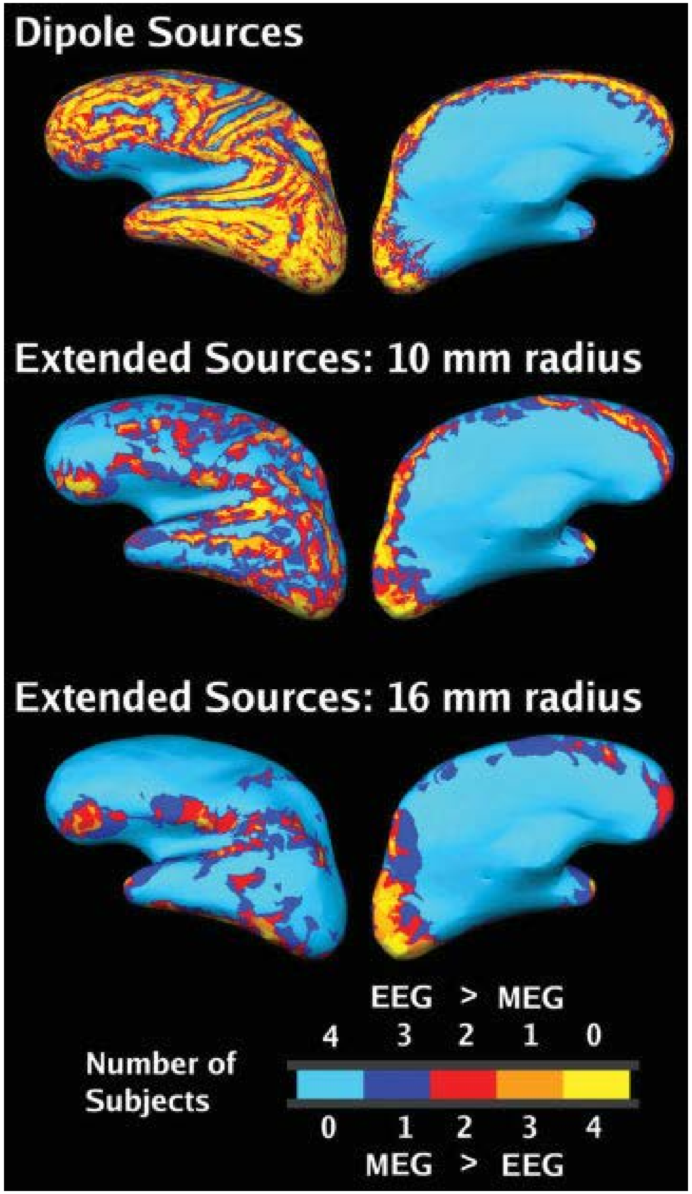
Hauk, Strenroos, Treder. In: Supek S, Aine C (edts), “Magnetoencephalography: From Signals to Dynamic Cortical Networks, 2nd Ed.”

Sensitivity of EEG and MEG



Goldenholz et al., HBM 2009

MEG Is Less Sensitive To Spatially Extended Sources Than EEG

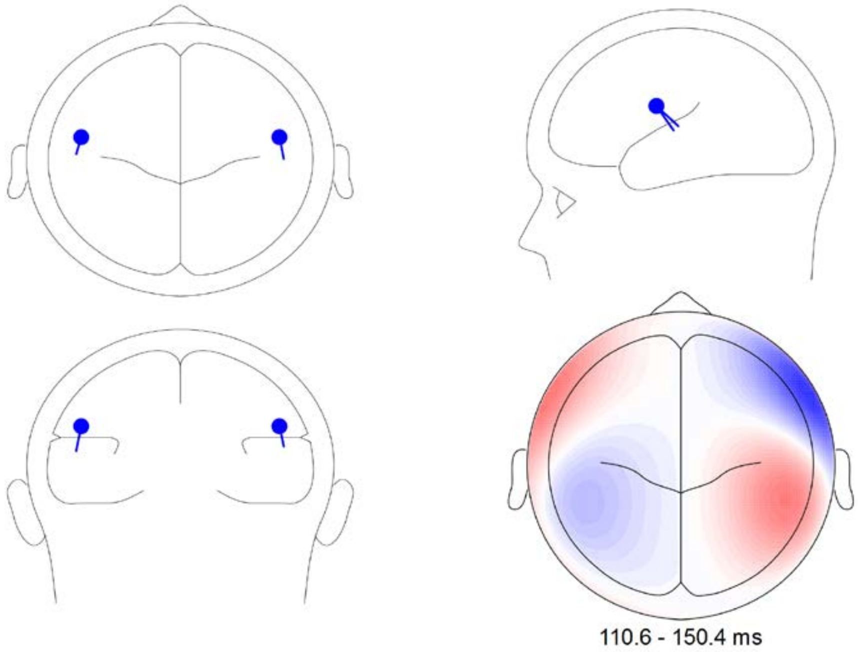


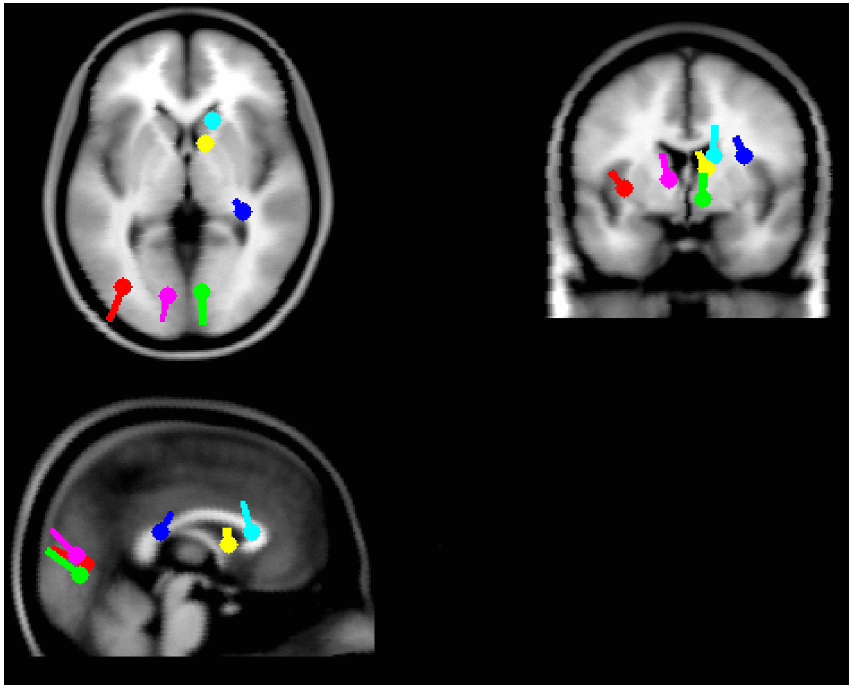
Goldenholz et al., HBM 2009
https://pubmed.ncbi.nlm.nih.gov/18465745/

Solutions To The Inverse Problem – Source Estimation

Hypothesis Testing - Dipole Fitting

Explicit assumptions about the number of focal sources (dipoles) are tested by fitting dipole models to the data.
The common criterion for the selection of models is the goodness-of-fit.

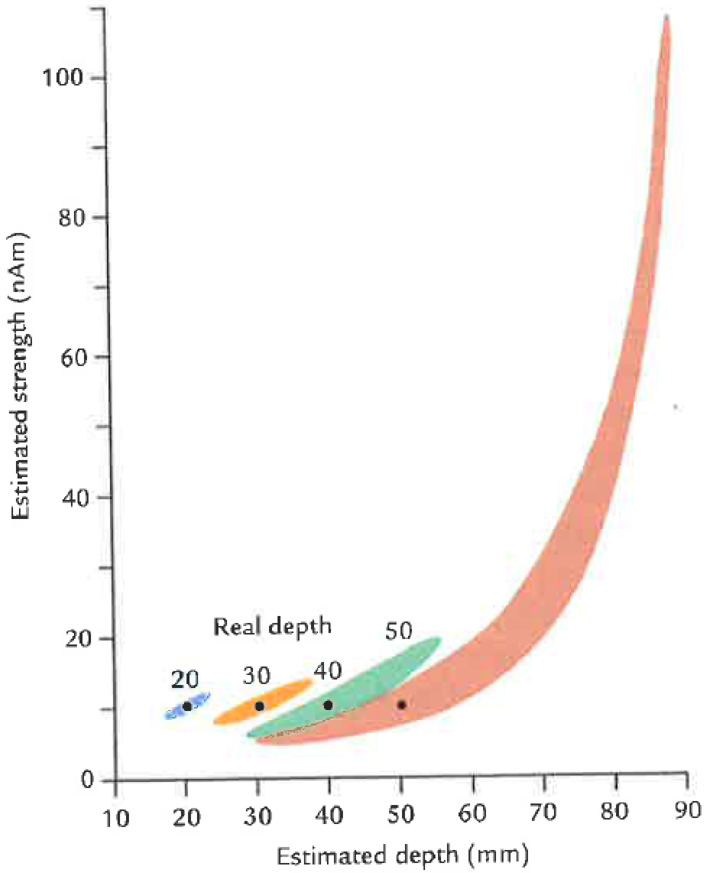




It can be hard to choose the appropriate number of dipoles – a priori knowledge is required.
Solutions for several/many dipoles can get stuck in local minima, and may not be robust to noise.

Assumptions Cannot Completely Remove Uncertainty

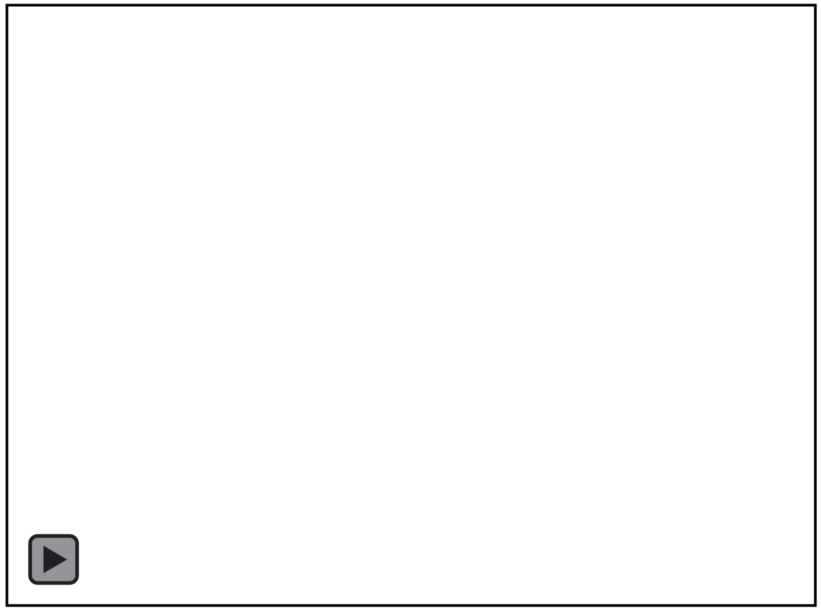
95% CIs for single dipole source



Hari et al., Electroencephalogr Clin Nph 1988

Dipole Scanning

We may have reasonable assumptions about possible locations for isolated dipole sources, e.g. on the cortical surface.



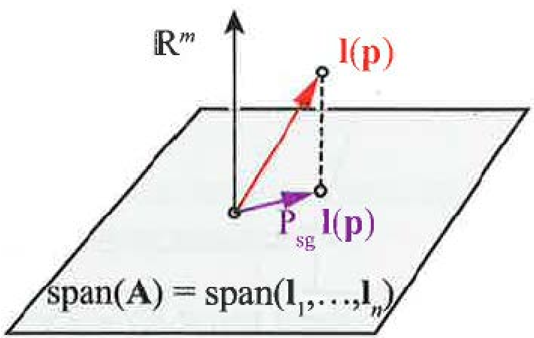
http://www.cogsci.ucsd.edu/~sereno/movies.html

Dipole scan: Fit dipoles vertex-by-vertex and plot the goodness-of-fit as a distribution.
The maxima in this distribution point to possible dipole locations.
The locations are reliable if there is only one dipole, or if multiple dipole topographies are mutually orthogonal (e.g. far apart).
This is not a “distributed source solution” (more on that later).

Multi-Dipole Scan: MUSIC

(Multiple Source Signal Classification)

Data and Noise Subspaces



Ilmoniemi & Sarvas, "Brain Signals", MIT 2019

Classical MUSIC

1) Obtain a spatio-temporal data matrix $\boldsymbol{F}$, comprising information from m sensors and n time slices. Decompose $\boldsymbol{F}$ or $\boldsymbol{F}\boldsymbol{F}^T$ and select the rank of the signal subspace to obtain $\hat{\Phi}_s$. Overspecifying the true rank by a couple of dimensions usually has little effect on performance. Underspecifying the rank can dramatically reduce the performance.
2) Create a relatively dense grid of dipolar source locations. At each grid point, form the gain matrix $\boldsymbol{G}$ for the dipole. At each grid point, calculate the subspace correlations subcorr$\{\boldsymbol{G}, \hat{\Phi}_s\}$.
3) As a graphical aid, plot the inverse of $\sqrt{1-c_1^2}$, where c_1 is the maximum subspace correlation. Correlations close to unity will exhibit sharp peaks. Locate r or fewer peaks in the grid. At each peak, refine the search grid to improve the location accuracy, and check the second subspace correlation. A large second subspace correlation is an indication of a "rotating dipole."

Mosher & Leahy, IEEE-TBME 1998

Recursively Applied (RAP) MUSIC

1) Estimate number of dipoles, e.g. using PCA/SVD.
2) Run MUSIC for one dipole.
3) Run MUSIC for 2nd dipole, partialling out dipole 1.
4) Repeat for estimated number of dipoles.

See e.g. for overview and recent updates of MUSIC algorithms: Ilmoniemi & Sarvas, "Brain Signals", MIT 2019;   Mäkelä et al., NI 2018 ("TRAP MUSIC", https://pubmed.ncbi.nlm.nih.gov/29128542/)

One problem with MUSIC algorithms: They don't give you source time courses.

“Spatial Filters”: Beamformers

Assumptions:

- All sources captured in data covariance matrix C (signal and noise)
- We are interested in one source *i* in many sources

Aim:

Design a spatial filter $\mathbf{w}_i$ which projects maximally on the source of interest and minimally on noise sources.

Project on source of interest: $\mathbf{w}_i^T \mathbf{f}_i$

Suppress noise: $min(\mathbf{w}_i^T \mathbf{C} \mathbf{w}_i)$

$$\mathbf{w}_i = \frac{\mathbf{f}_i^T \mathbf{C}^{-1}}{\mathbf{f}_i^T \mathbf{C}^{-1} \mathbf{f}_i}$$

Linearly-Constrained Minimum-Variance (LCMV) Beamformer

Van Veen et al., 1997, https://pubmed.ncbi.nlm.nih.gov/9282479/

Create and apply these spatial filters vertex-by-vertex (dipole-by-dipole) and plot the distribution (possibly normalised by noise variance).
Spatial filters can also produce time courses for every source.

e.g. Hauk&Stenroos, HBM 2013, https://pubmed.ncbi.nlm.nih.gov/23616402/ ,
Hauk et al., bioRxiv 2019, https://www.biorxiv.org/content/10.1101/672956v1

Beamformers

The "linearly-constrained maximum-variance" (LCMV) beamformer

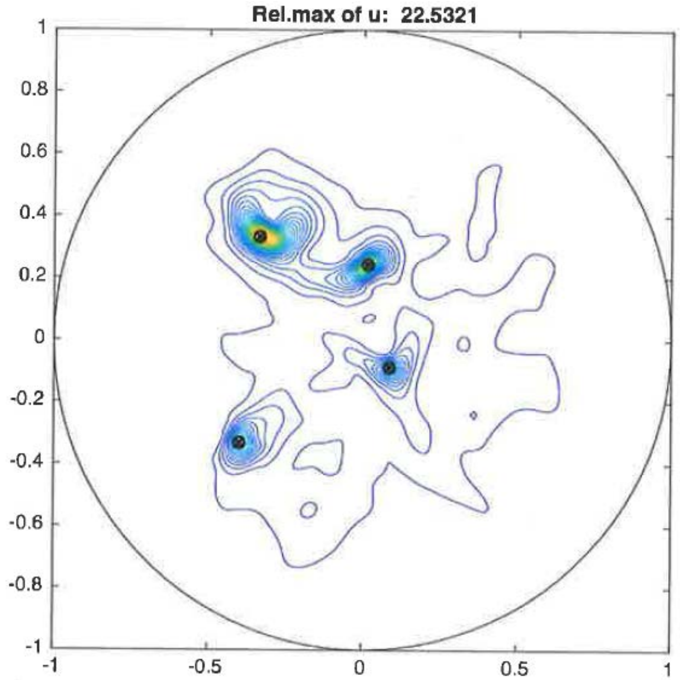
$$\mathbf{SF}_{LCMV}(\mathrm{i}) = \frac{\tilde{\mathbf{L}}_{.i}^{T}\mathbf{C}_{d}^{-1}}{\tilde{\mathbf{L}}_{.i}^{T}\mathbf{C}_{d}^{-1}\tilde{\mathbf{L}}_{.i}}$$

depends on the data covariance matrix ("adaptive").

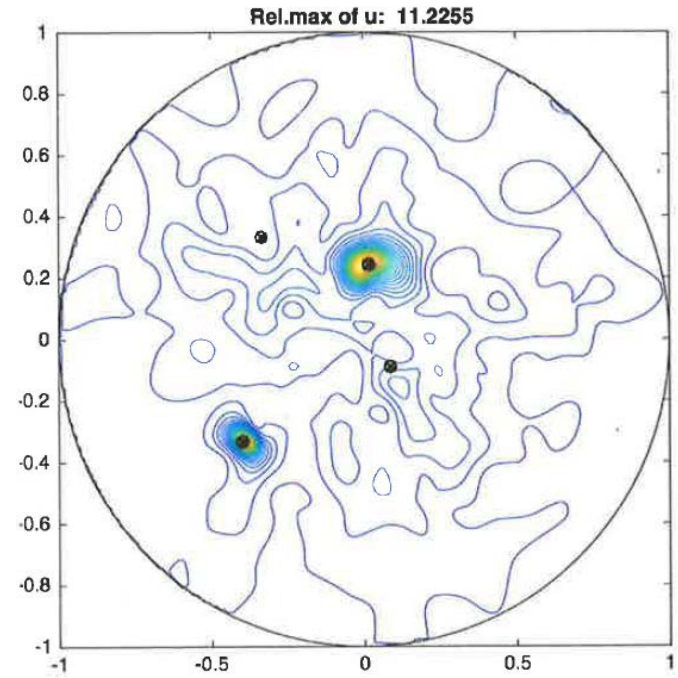
Beamformers result in linear transformations of the data ("spatial filters"), but those transformations strongly depend on the data of interest.

=> Beamformers are not linear with respect to the sources of interest.

Beamforming Is Problematic For Highly Synchronous Sources



4 non-synchronous sources

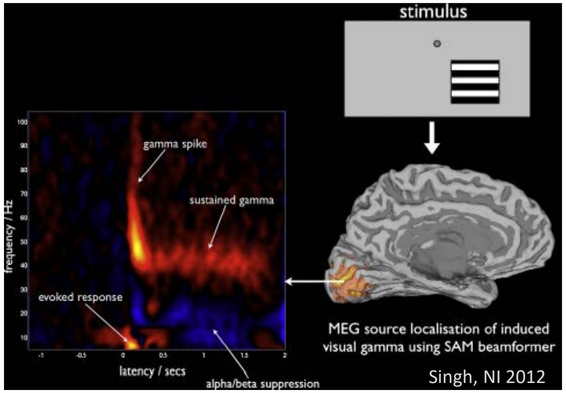


2 non-synchronous,
2 synchronous sources

Ilmoniemi and Sarvas, “Brain Signals”, MIT 2019

Beamformers Are Popular for Rhythmic Brain Activity and Resting State Activity

Visual Gamma Band Response

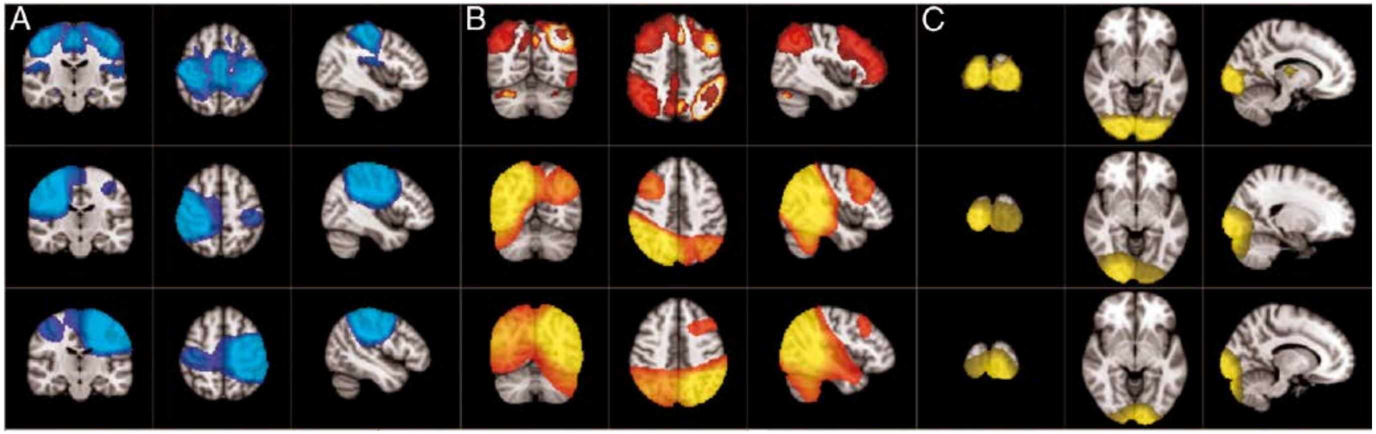


Resting State Networks

Motor

Fronto-Parietal

Visual



Brookes et al. PNAS 2011

Beamformers Are Popular for Rhythmic Brain Activity and Resting State Activity…

…but the choice of source estimation method should be based on knowledge (or its absence) about the source distribution.

Is there anything in rhythmic/oscillatory or resting state activity that favours some source distributions more than others
(e.g. number of sources, focality/sparsity, location)?

For example, visual gamma band sources may be focal, but resting state networks may be distributed.

Minimum Norm Estimation Of Distributed Sources

Minimum Norm Estimation Of Distributed Sources

$$\mathbf{L}\boldsymbol{s} = \mathbf{d} \quad \Rightarrow \|\mathbf{L}\boldsymbol{s} - \mathbf{d}\|^2 = \mathbf{0}$$

(ignore noise for now)

subject to constraint

$$\|\mathbf{s}\|_2 = min$$

yields the Minimum-Norm Least-Squares solution ("L2")

$$\hat{\boldsymbol{s}} = \boldsymbol{G}_{\boldsymbol{MN}}\mathbf{d}$$

with

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

But this is the result of mathematical desperation, and not based on physiology or what we want to know (e.g. localisation of sources).

There Are Many Norms, e.g. L1 vs L2 - Sparseness

Minimising the L2 norm, $\|\mathbf{s}\|_2 = |s_1|^2 + |s_2|^2 + \ldots + |s_N|^2$ penalizes large values in **s**
=> "smooth"

Minimising the L1 norm, $\|\mathbf{s}\|_1 = |s_1| + |s_1| + \ldots + |s_N|$ prefers large values in **s**
$\Rightarrow$ "sparse"

For example:

$$x_1 + 2x_2 = 1$$

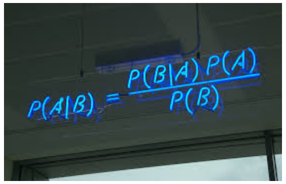
L2 solution: (0.2, 0.4)

L2-norm $0.2^2+0.4^2$~0.45, L1-norm 0.2+0.4=0.6

L1 solution: (0, 0.5)

L2-norm 0.5, L1-norm 0.5

There Are Different Optimisation Criteria: Bayesian Approach



Bayes' rule:

$$p(\mathbf{s}|\mathbf{d}) \sim p(\mathbf{d}|\mathbf{s}) * p(\mathbf{s})$$

posterior ~ likelihood * prior

Assume normal distribution for noise:

$$p(\mathbf{d}|\mathbf{s}) = \left(\frac{\beta}{2\pi}\right)^{\boldsymbol{M/2}} exp\left(-\frac{\beta}{2}\|\mathbf{Ls}-\mathbf{d}\|^{\mathbf{2}}\right)$$

Thus, minimise

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

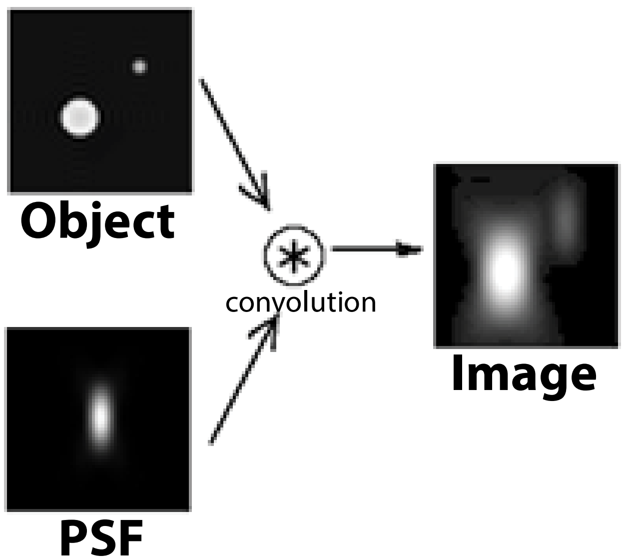
e.g. Henson et al., 2011, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3160752/

"Most likely" is still not what we want to know –
Does the method do what we want it to do?

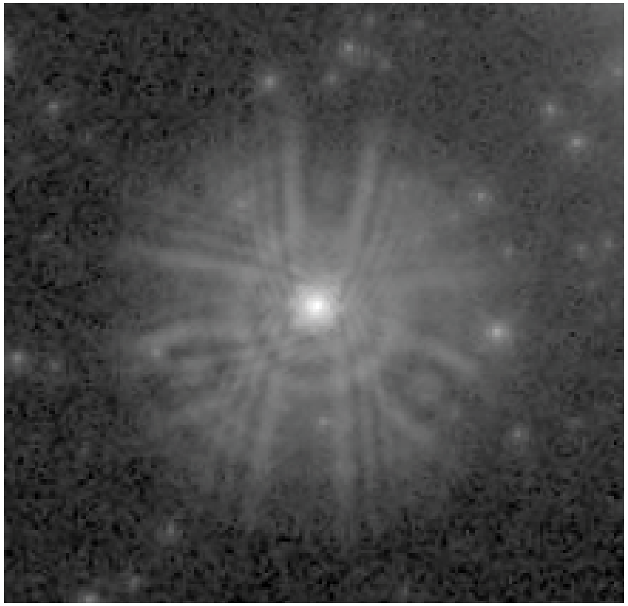
Let’s Start Again:
The “Blurry Image” Analogy

The Superposition Principle
A “Constraint-Free” Interpretation of Linear Methods

Microscopy



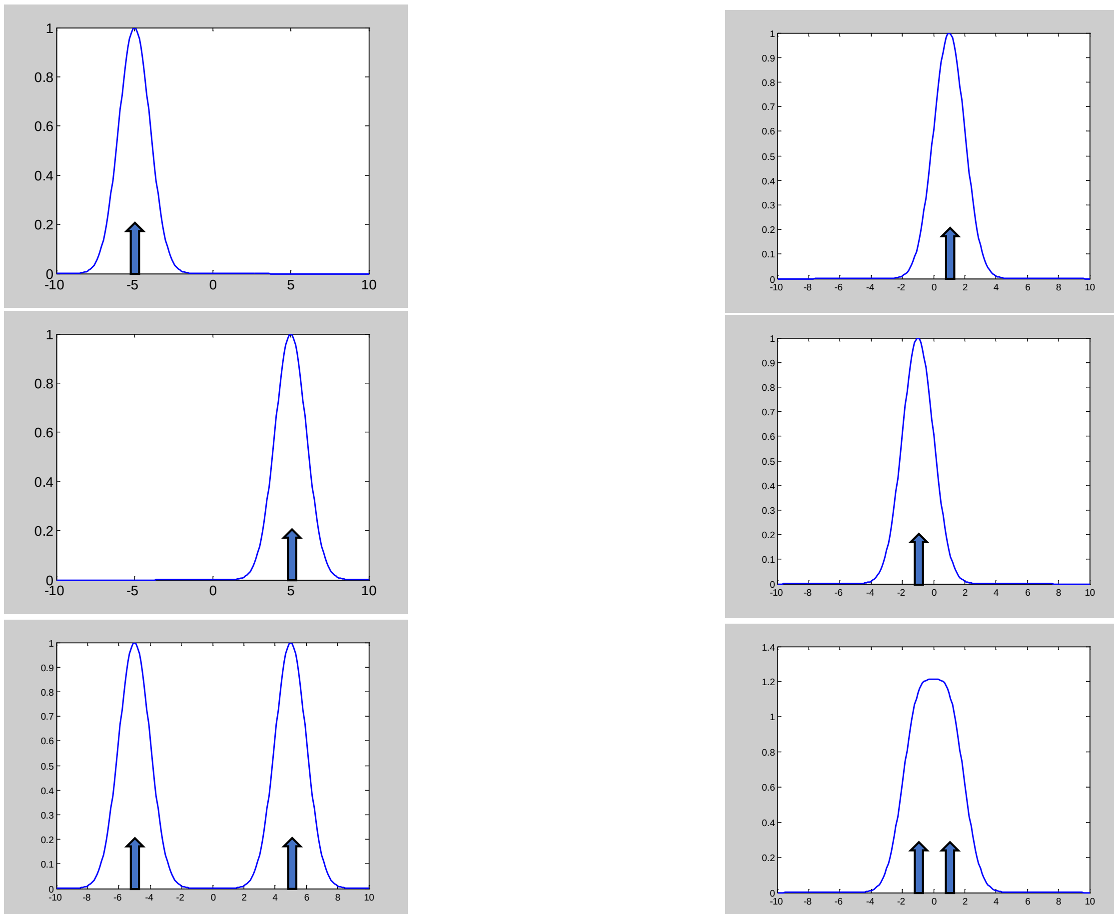
Astronomy



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

Linear Methods Can Easily Tell Us If They Do What We Want

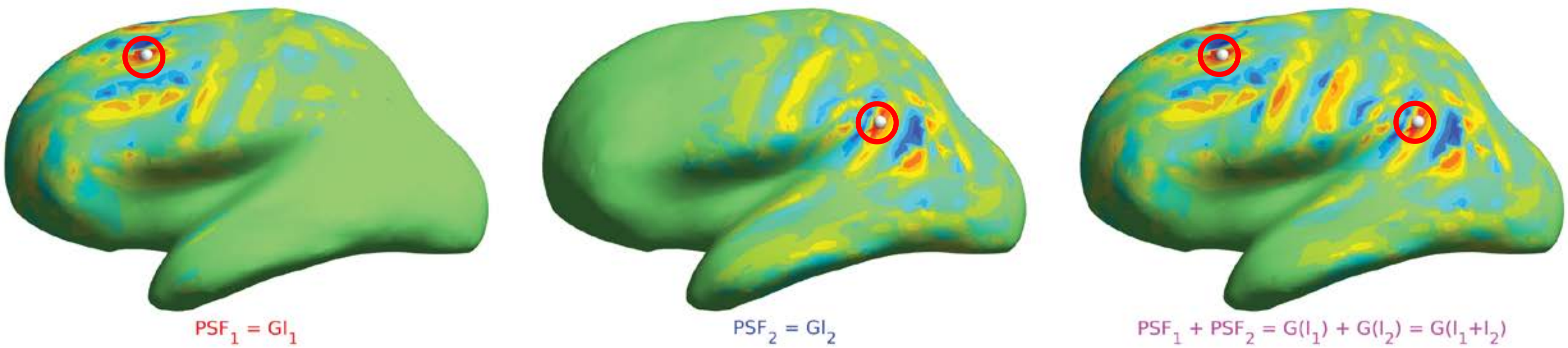
Superposition Principle



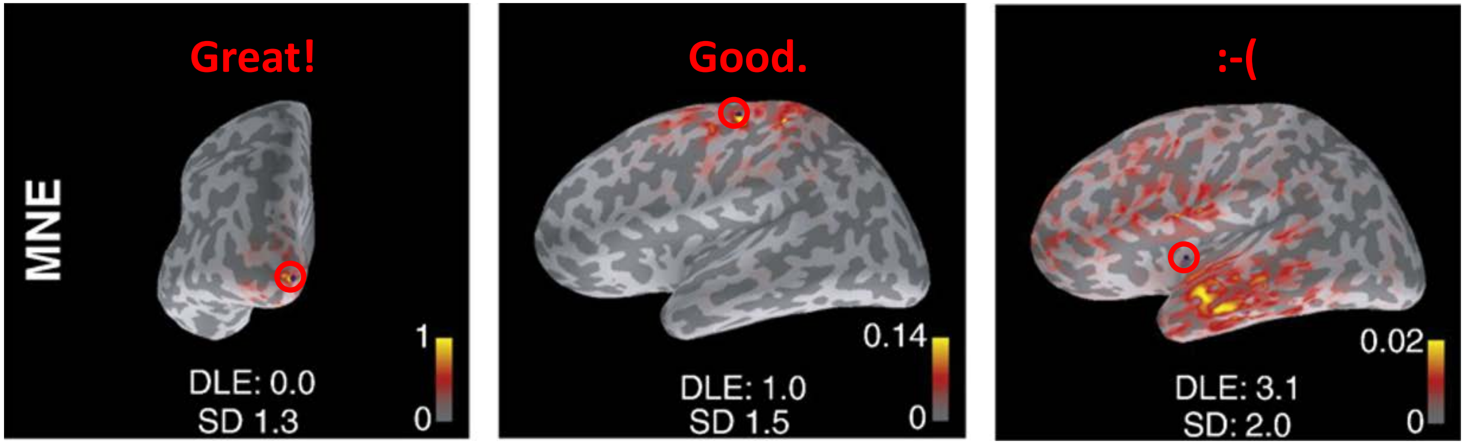
If you know the behaviour for point sources,
you can predict the behaviour for complex sources

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."

Spatial Resolution of Source Estimation Is Complex

Spatial resolution depends on:

modelling 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

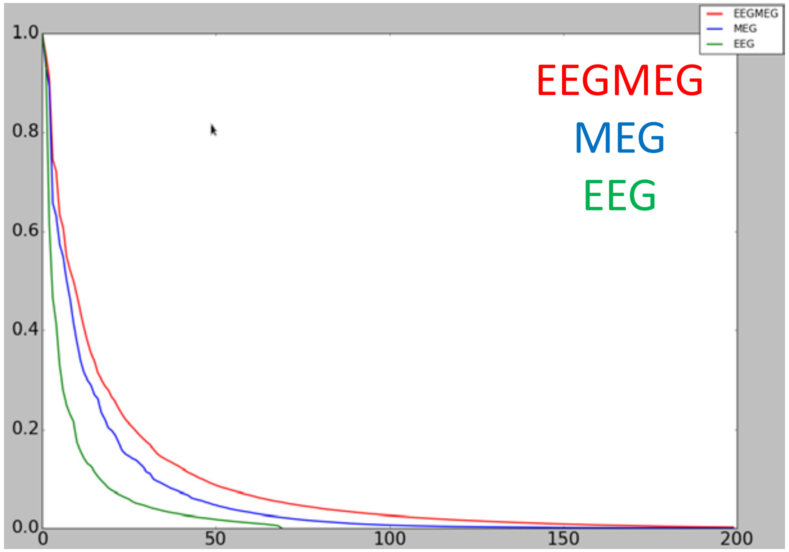
With ***n* sensors**:
-> *n* independent measurements
-> *n* independent parameters estimable
-> at best separate activity from *n* brain regions

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

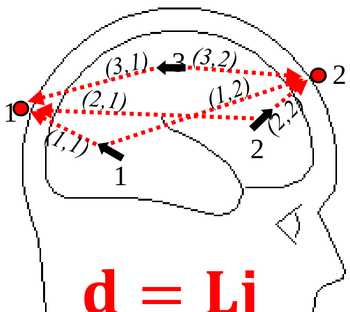
"Resel": 38 cm^3 -> $\underline{3.4}^3$ cm^3

SVD of Leadfields



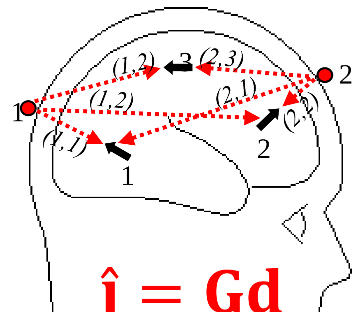
Resolution Matrix

Forward Problem

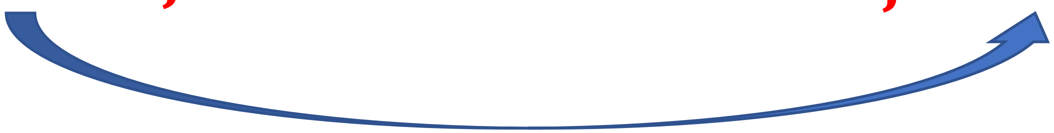


$$\mathbf{d} = \mathbf{L}\mathbf{j}$$

Linear Inverse Problem



$$\hat{\mathbf{j}} = \mathbf{G}\mathbf{d}$$



$$\hat{\mathbf{j}} = \mathbf{G}\mathbf{L}\mathbf{j} \stackrel{\text{def}}{=} \mathbf{R}\mathbf{j}$$

Relationship between estimated and true source distribution.

e.g. Hauk/Stenroos/Treder, bioRxiv 2019
https://www.biorxiv.org/content/10.1101/672956v1

MNE Has An Optimal Resolution Matrix

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

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:

$$\|\boldsymbol{R} - \boldsymbol{I}\|_2 = min$$

This leads to the **Minimum Norm Estimator (MNE)**:

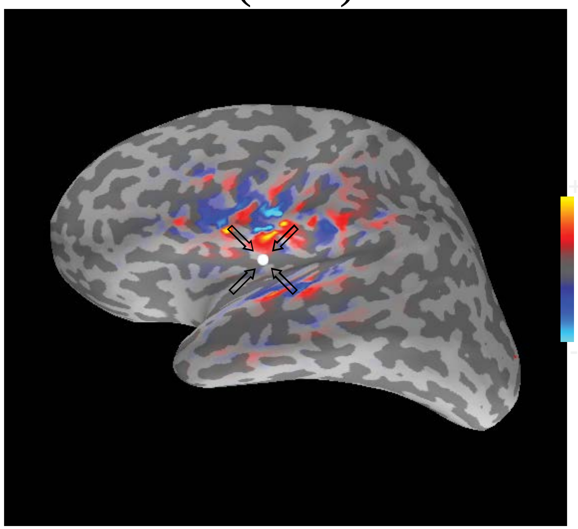
$$\boldsymbol{G}_{MN} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1}$$

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

Spatial Resolution:

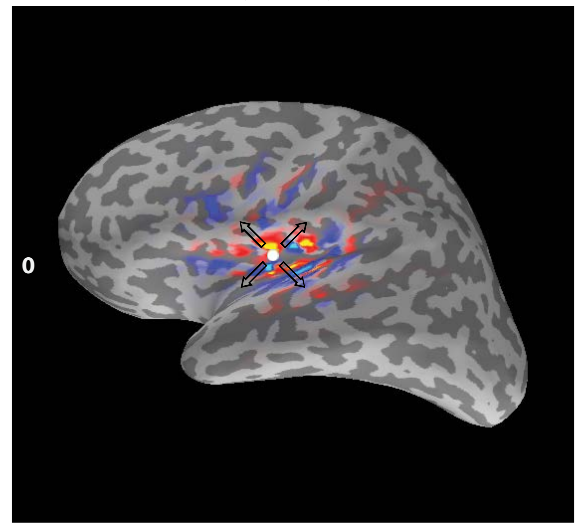
Point-Spread and Cross-Talk/Leakage

Cross-Talk Function (CTF)



How other sources may affect the estimate for this source

Point-Spread Function (PSF)

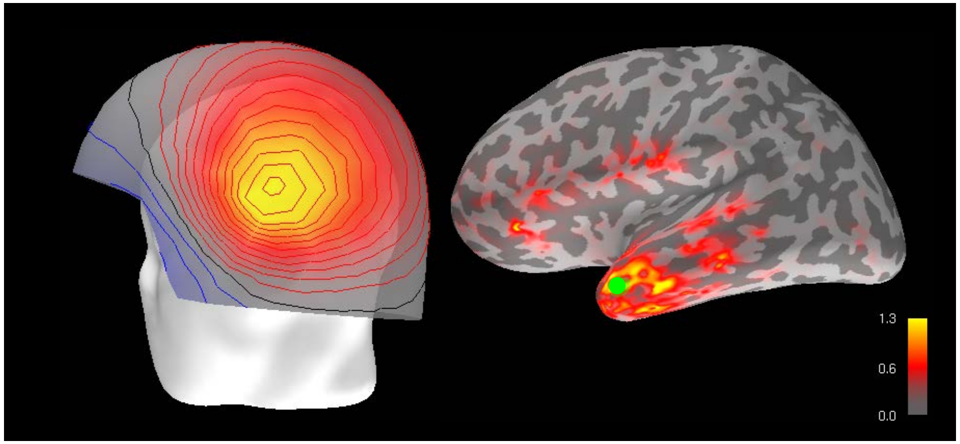


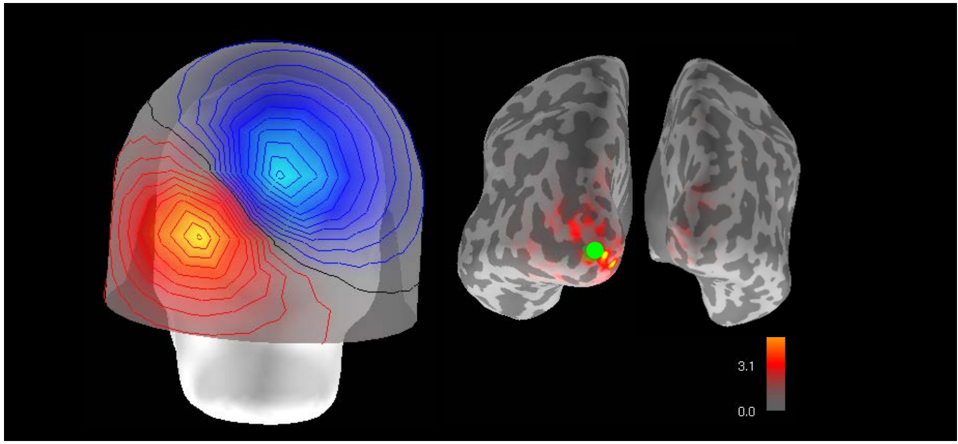
How this source affects estimates for other sources

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

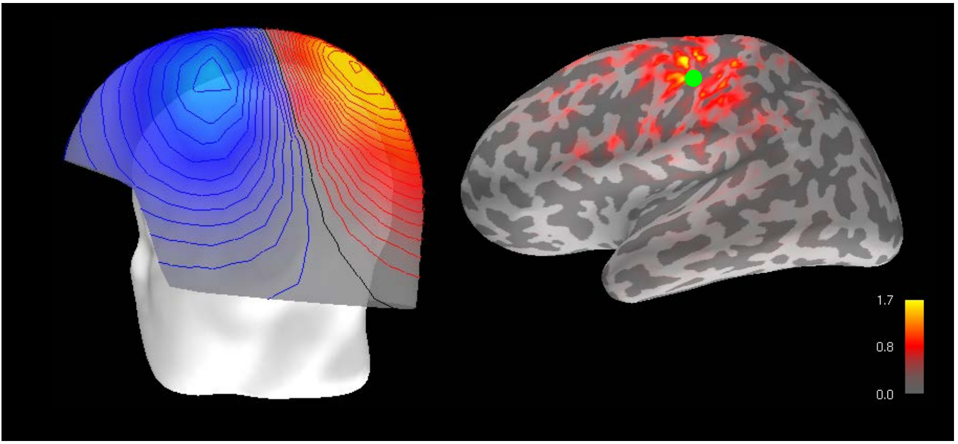
PSFs and CTFs for Some ROIs

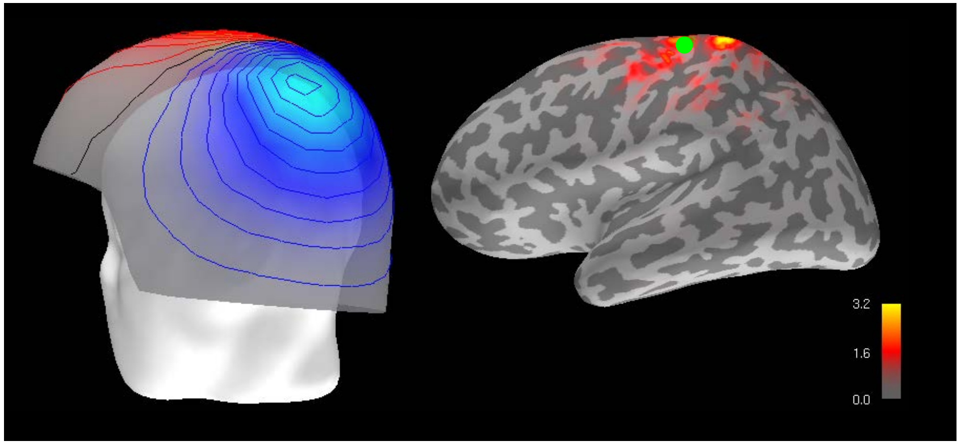
For MNE, PSFs and CTFs turn out to be the same





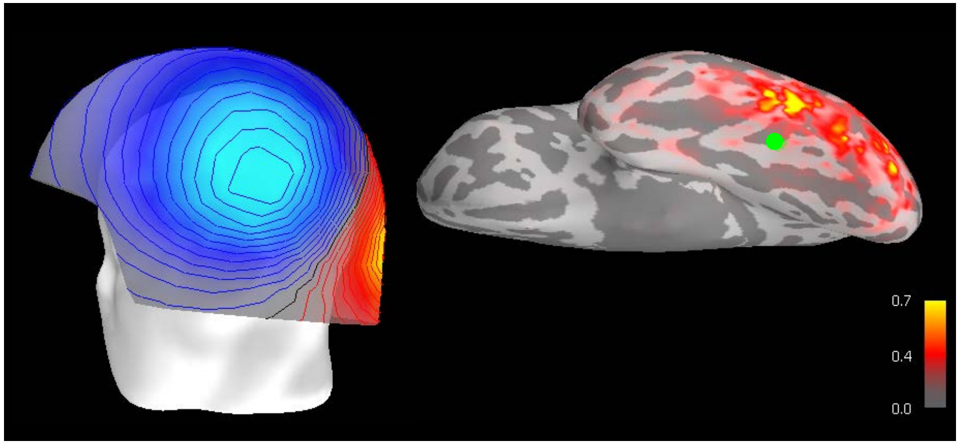
Good

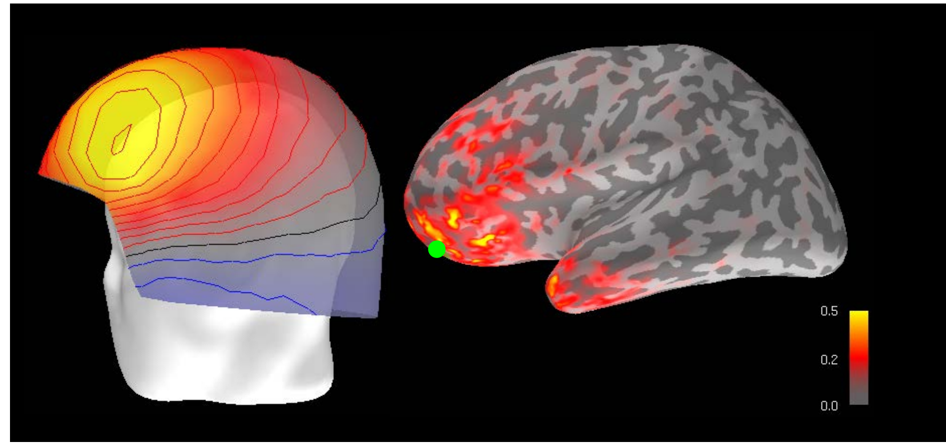




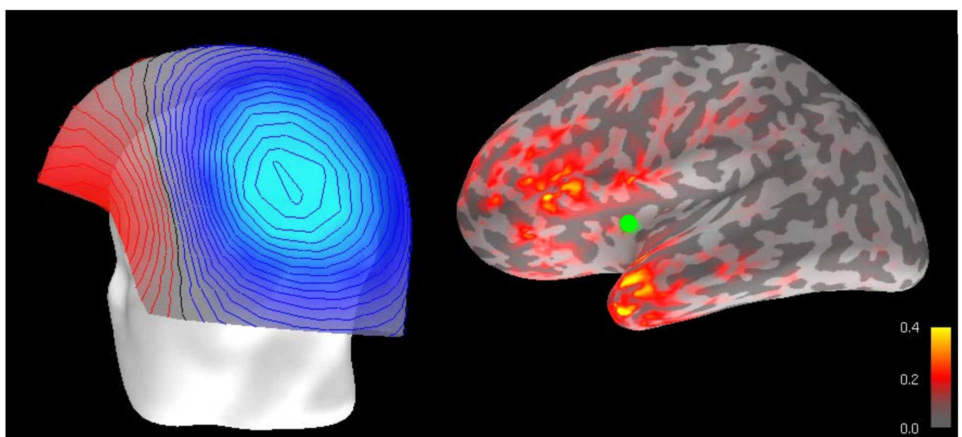
PSFs and CTFs for Some ROIs

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



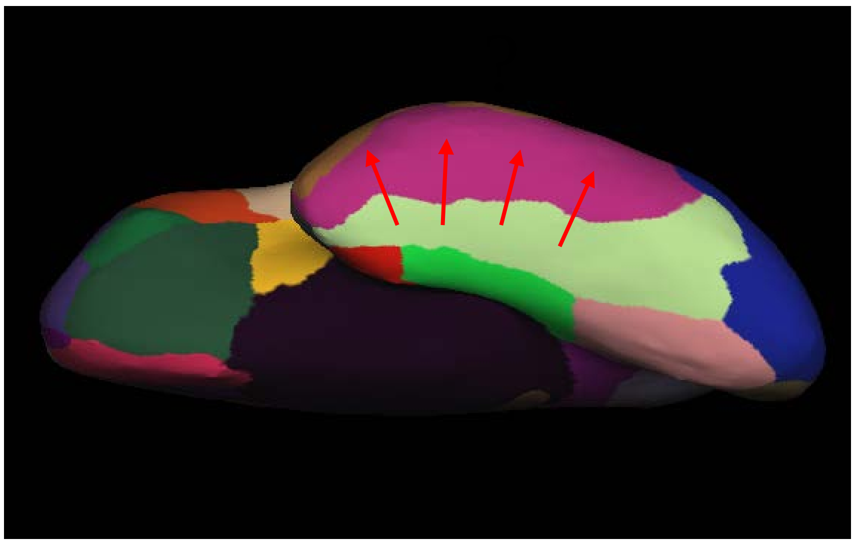


Less good

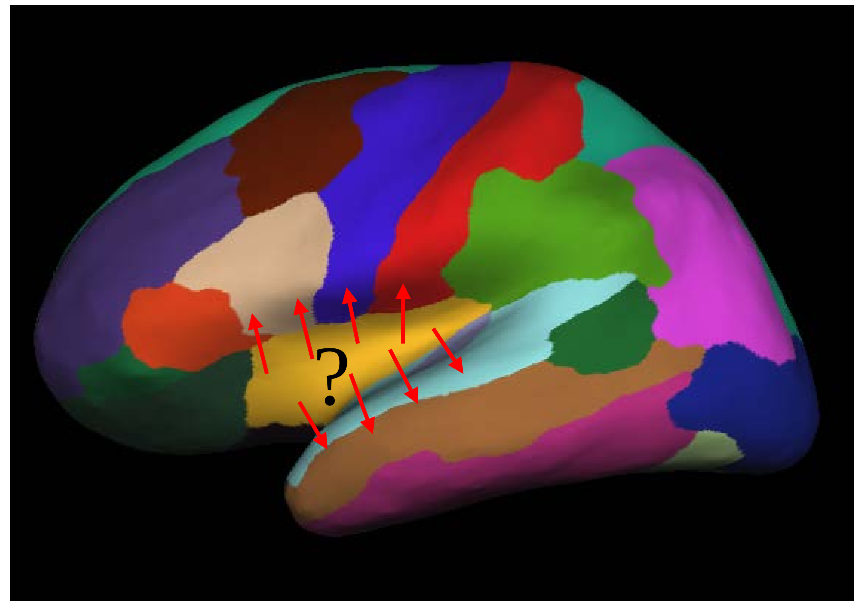


Localisation Bias Has Consequences for ROI analysis

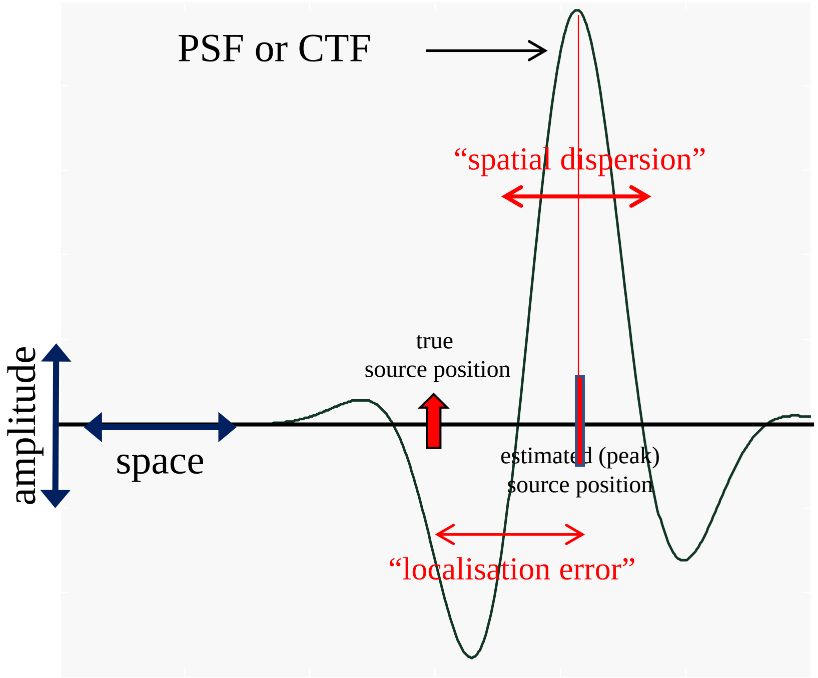
PSFs/CTFs Can Tell You How It Looks Like

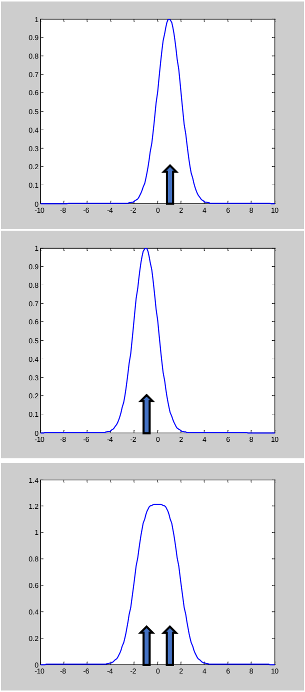


Desikan-Killiany Atlas parcellation



Quantifying Resolution From PSFs and CTFs





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

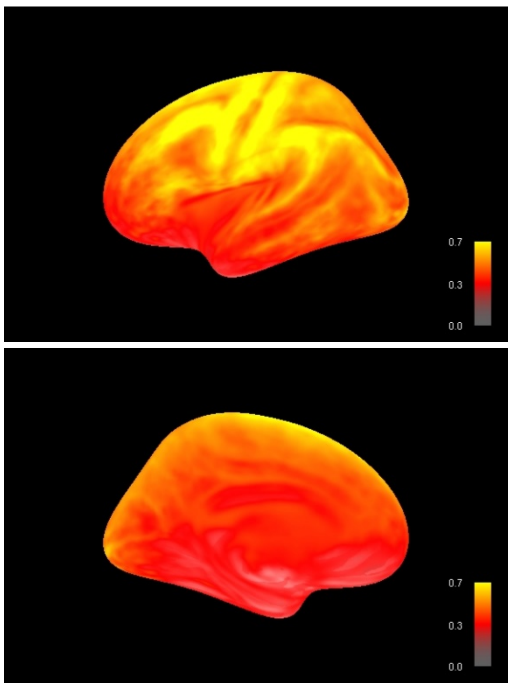
Resolution Metrics For PSFs/CTFs

- **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
  - 2 LCMV beamformers (pre- and post-stimulus covariance matrices)
- **Resolution Metrics**:
  - Peak Localisation Error
  - Spatial Dispersion (extent)

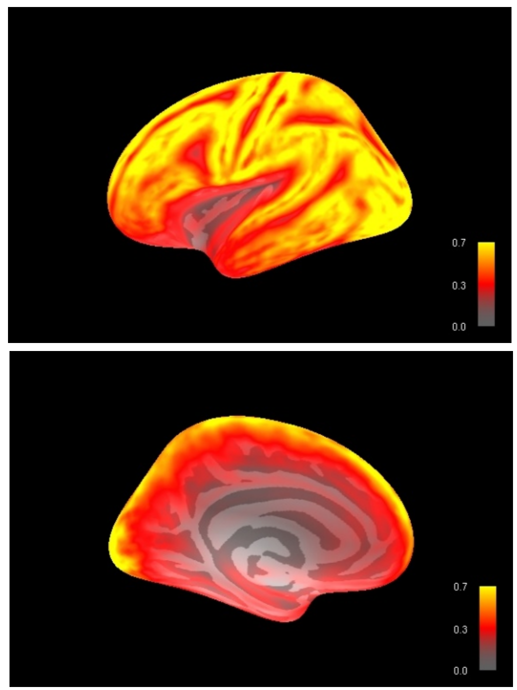
Sensitivity Maps

RMS of Leadfield Columns

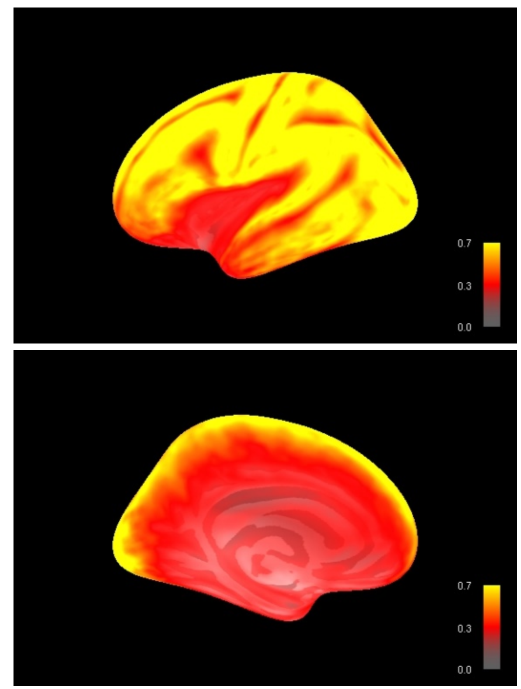
EEG
70 electrodes



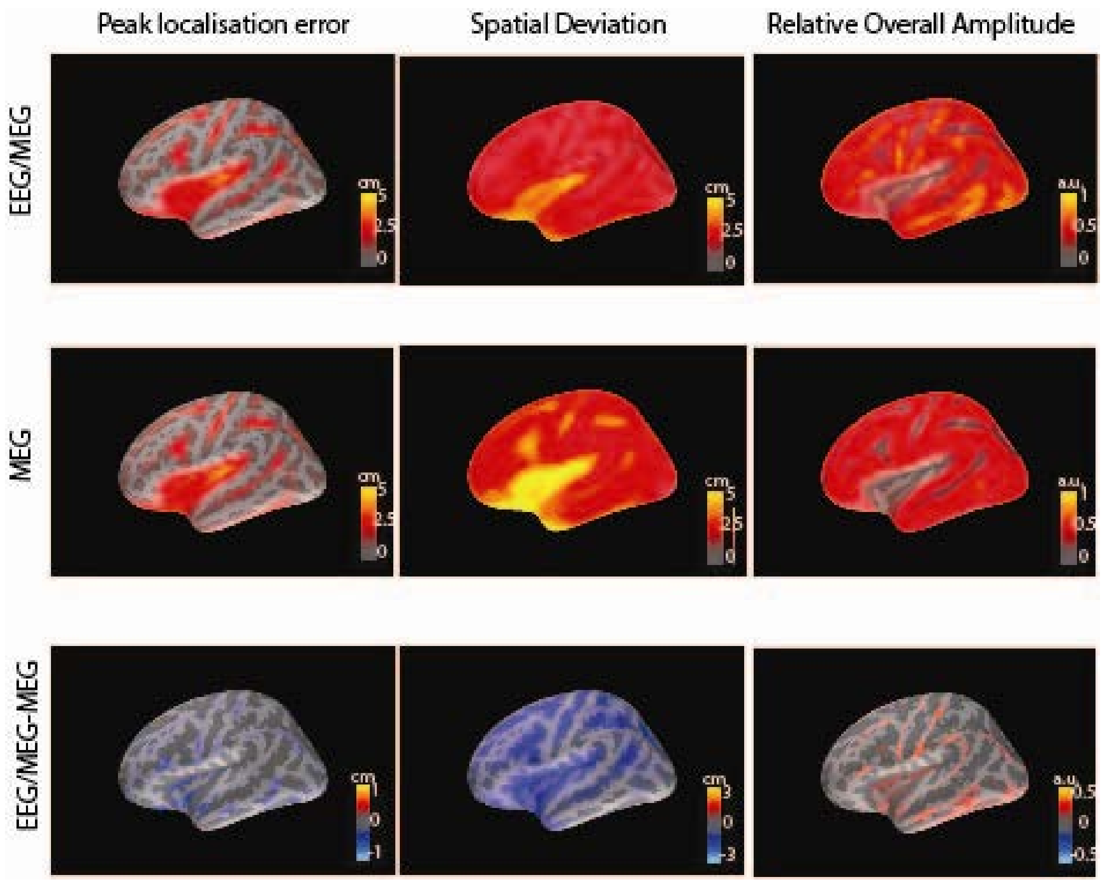
MEG
102 mags + 204 grads



EEG+MEG
102 mags + 204 grads



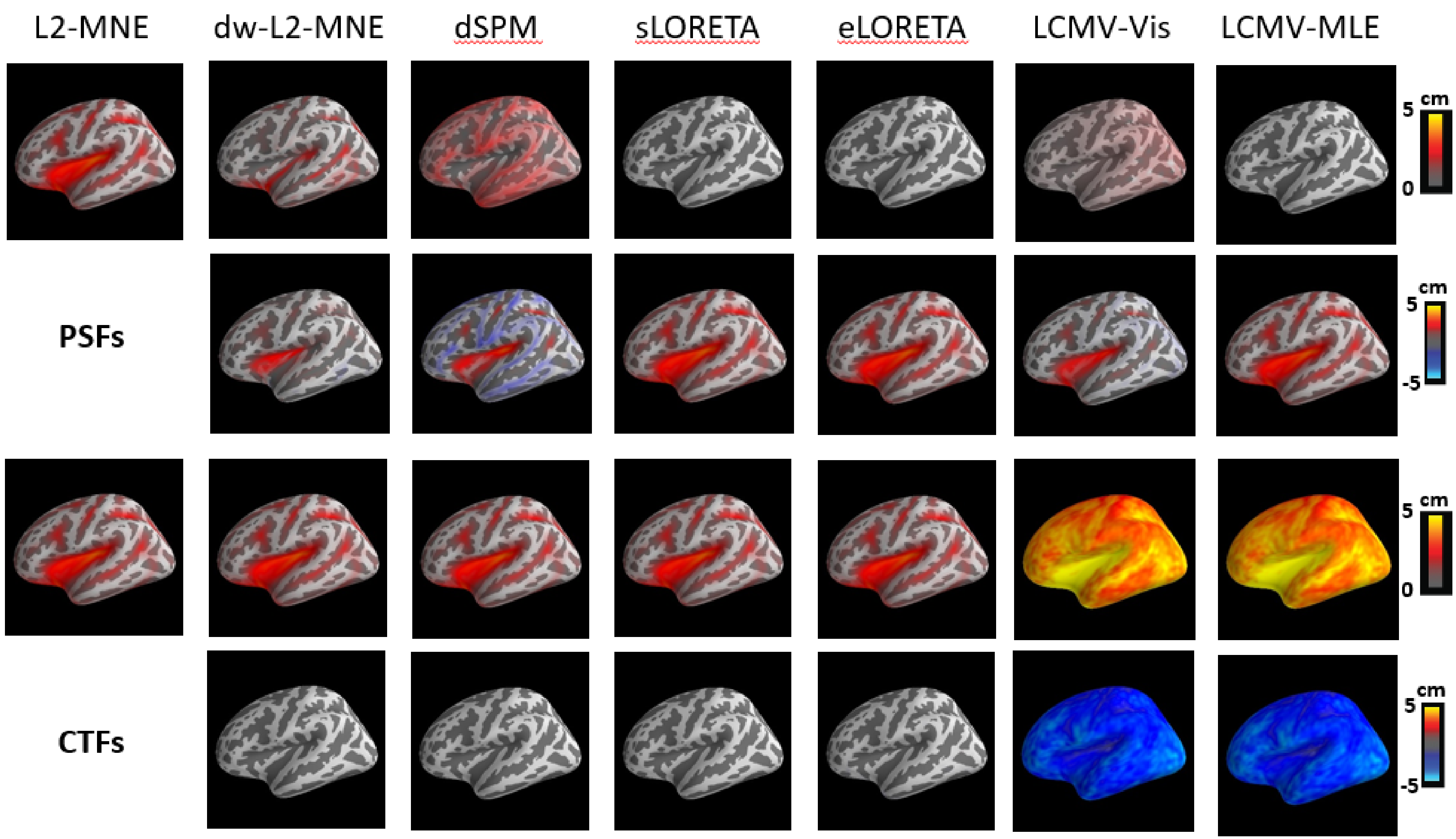
Combining EEG And MEG Improves Spatial Resolution



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

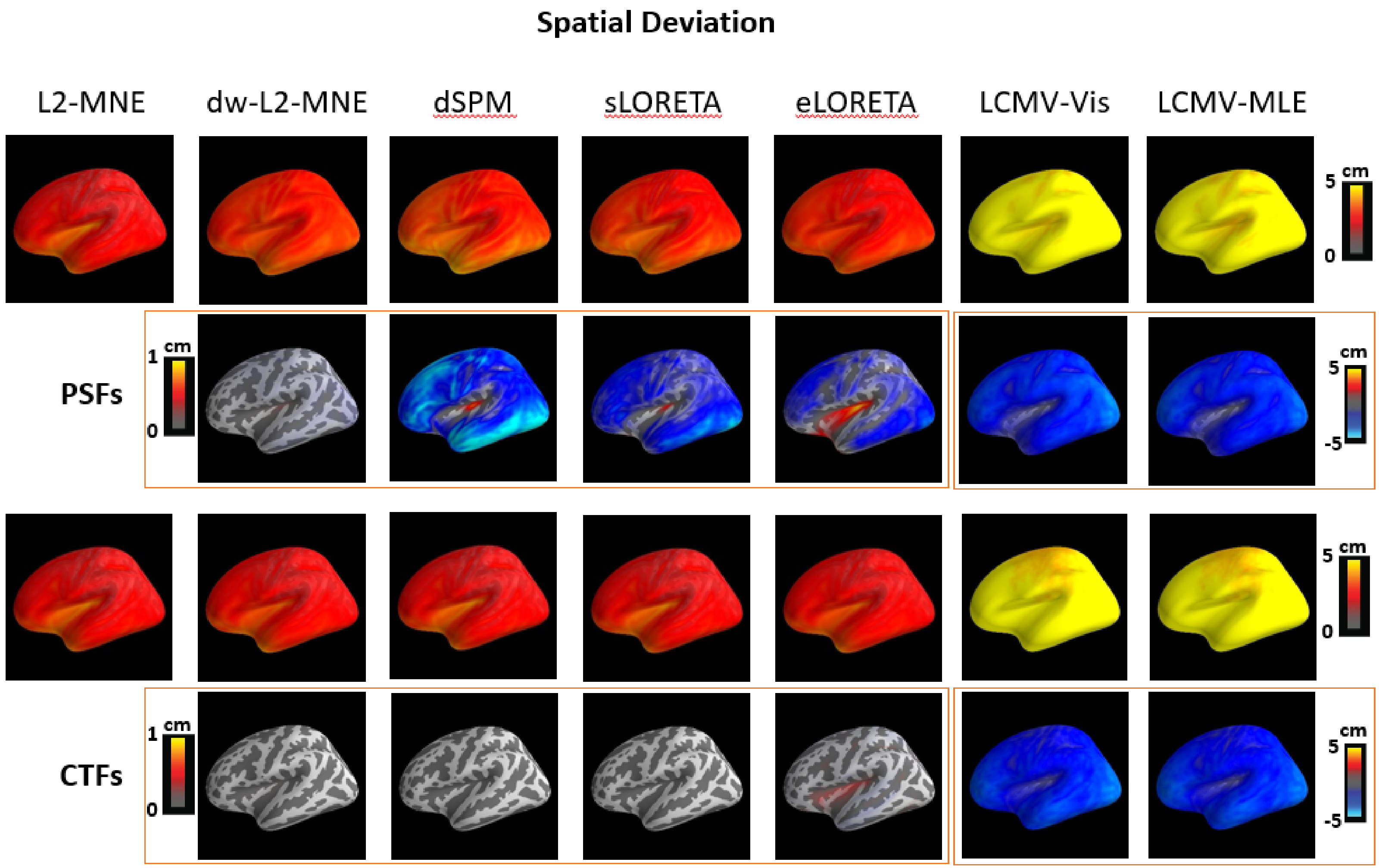
Comparing Estimators: Localisation Error

Peak Localisation Error



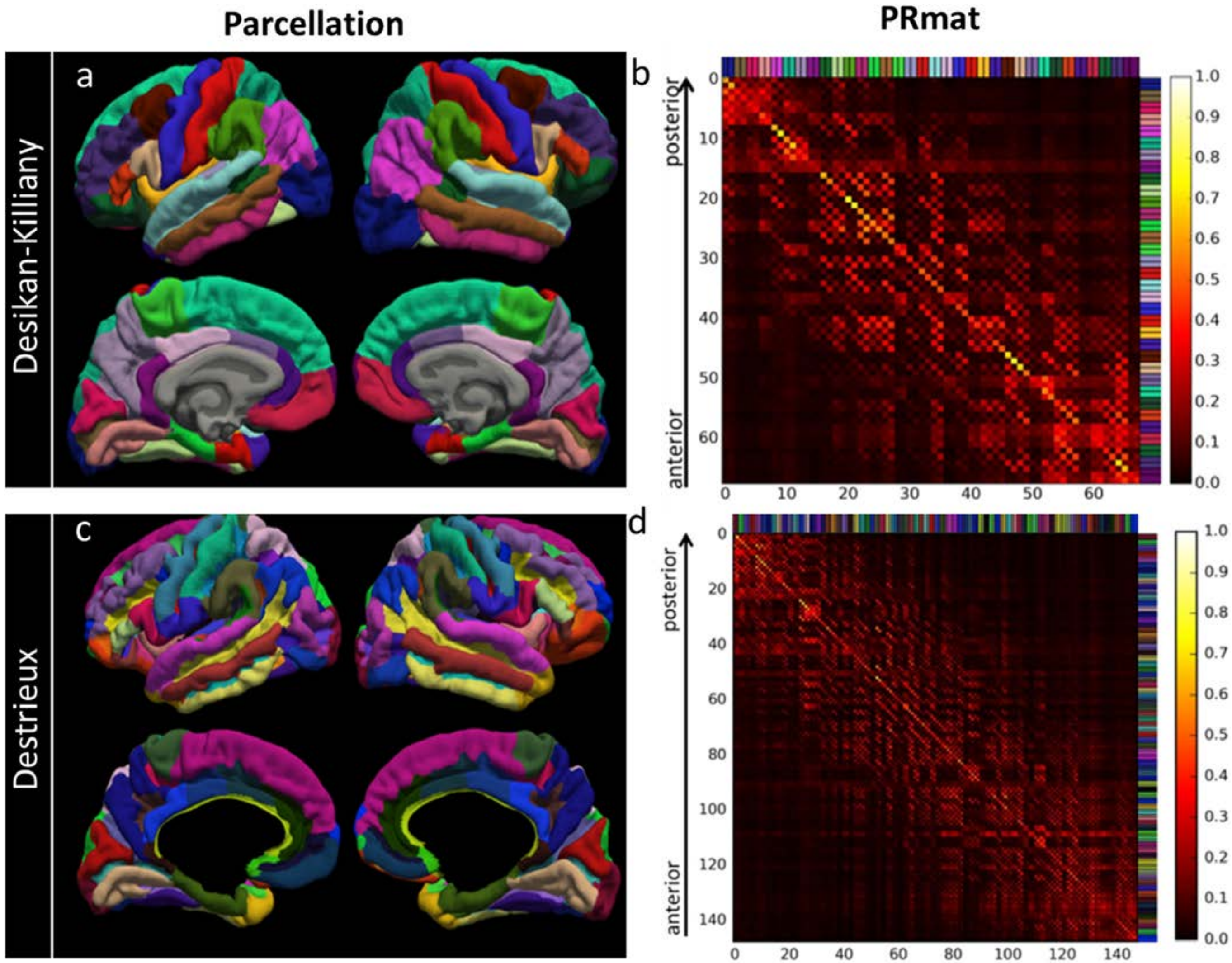
Hauk/Stenroos/Treder, bioRxiv 2019, https://www.biorxiv.org/content/10.1101/672956v1, see also Hauk/Wakeman/Henson, NI 2011

Comparing Estimators: Spatial Extent



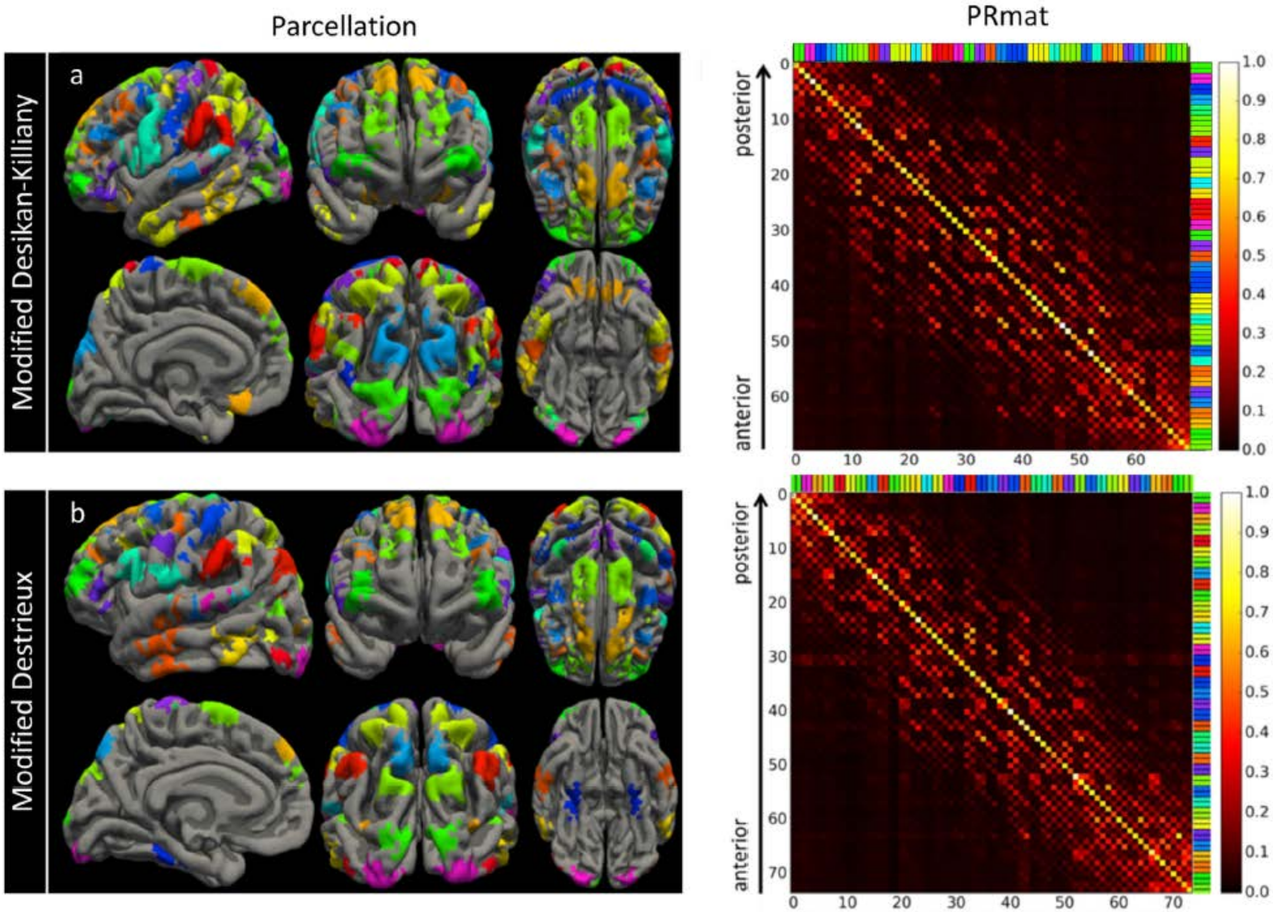
Hauk/Stenroos/Treder, bioRxiv 2019, https://www.biorxiv.org/content/10.1101/672956v1,
see also Hauk/Wakeman/Henson, NI 2011

Anatomical Parcellations May Not Be Optimal For EEG/MEG



Farahibozorg, Henson, Hauk, NI 2018

Adaptive Parcellations For EEG/MEG

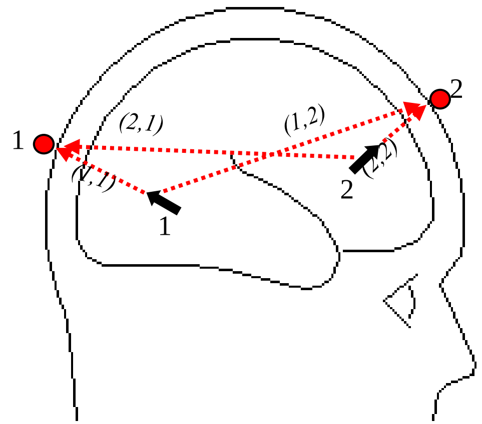


Farahibozorg, Henson, Hauk, NI 2018

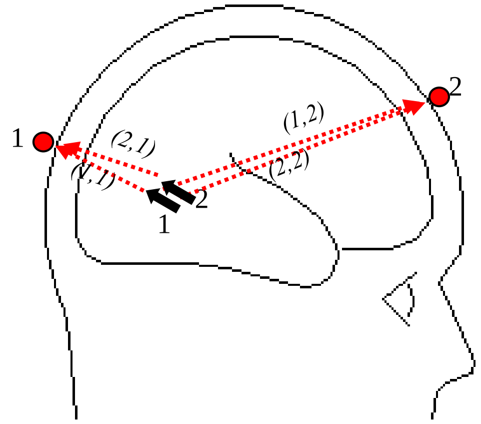
Noise and Regularisation

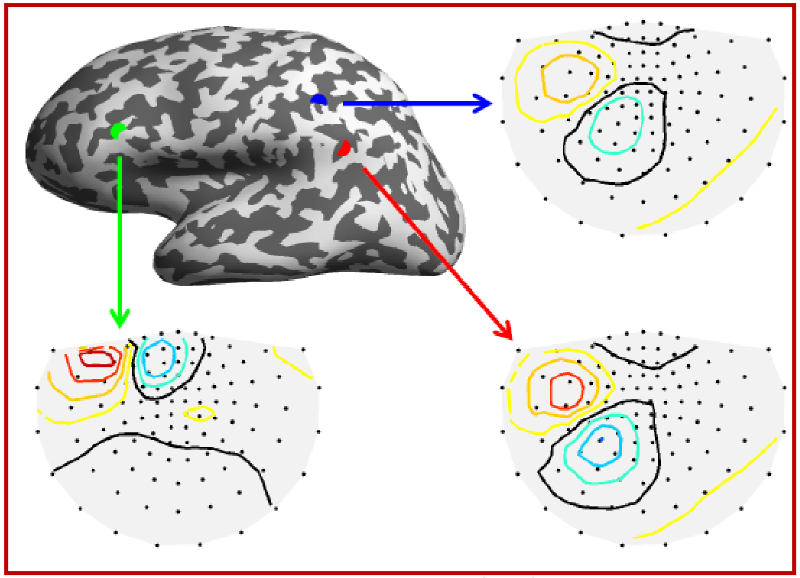
(In)Stability – Sensitivity to Noise

Stable



Instable





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

Thanks to Matti Stenroos.

Noise and Regularization

Over- And Under-Fitting

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 than the signal of interest:
Overfitting may seriously distort the solution ("variance amplification" in statistics/regression).

"Regularisation" results in a spatially smoother solution that is less affected by noise. The degree of smoothing depends on the "regularisation parameter" (also called "lambda").

Underfitting (over-smoothing) may waste spatial resolution.

Regularisation Can Take Into Account Noise covariance

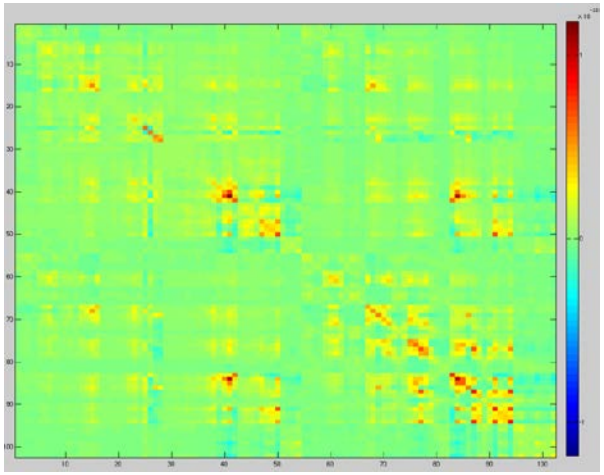
Some channels are noisier than others

⇒They should get different weights in your analysis

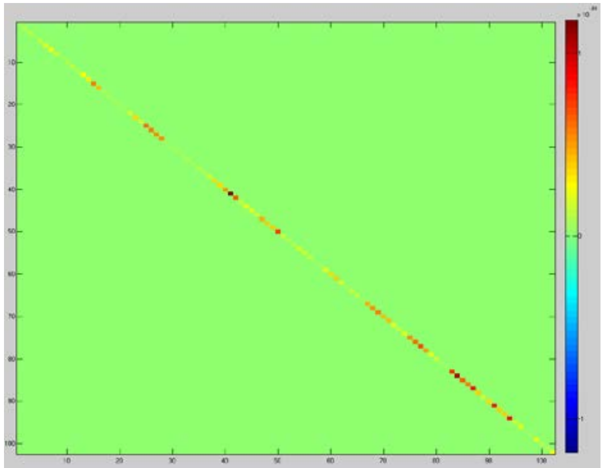
Sensors are not independent

=> 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

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

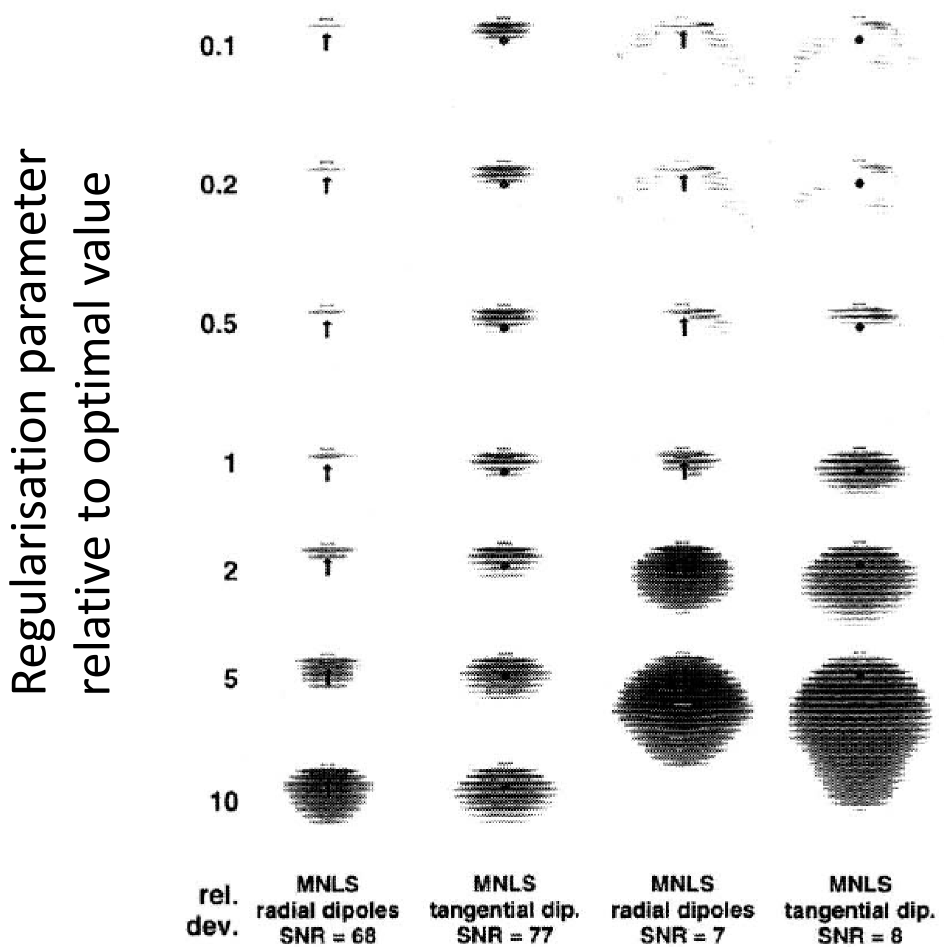
where $\tilde{\mathbf{L}}$ is the "whitened" leadfield $\mathbf{C}^{-1/2}\mathbf{L}$, and scaled such that trace($\tilde{\mathbf{L}}\tilde{\mathbf{L}}^T$)=trace($\mathbf{I}$).

$\tilde{\mathbf{L}}$ 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.

Trade-off norm-variance, smoothness

Source at fixed excentricity 71% (60mm)



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

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

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.

e.g. Henson et al., 2011, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3160752/