MRC Cognition and Brain Sciences Unit

UNIVERSITY OF CAMBRIDGE

EEG/MEG 2:

(Linear) Source Estimation

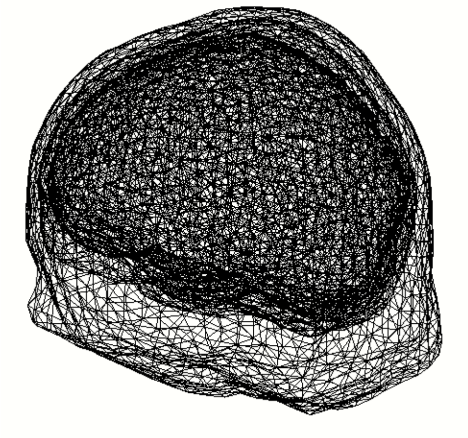
Olaf Hauk

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

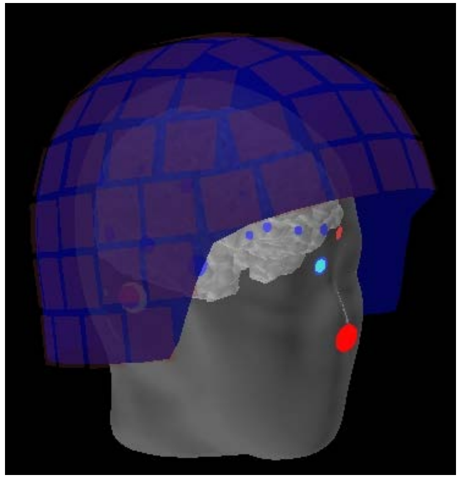
COGNESTIC 2022

Ingredients for Source Estimation

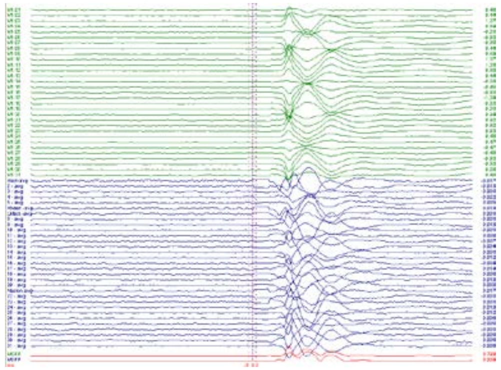
Volume Conductor/
Head Model



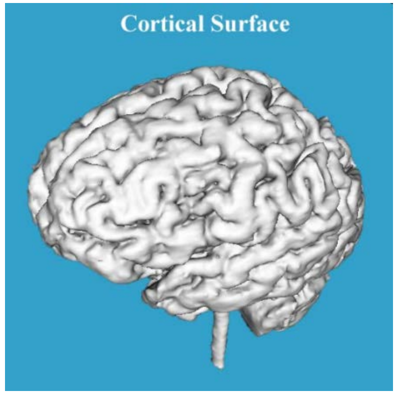
Coordinate
Transformation



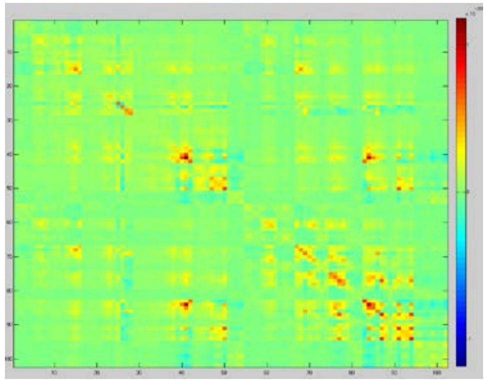
MEG data



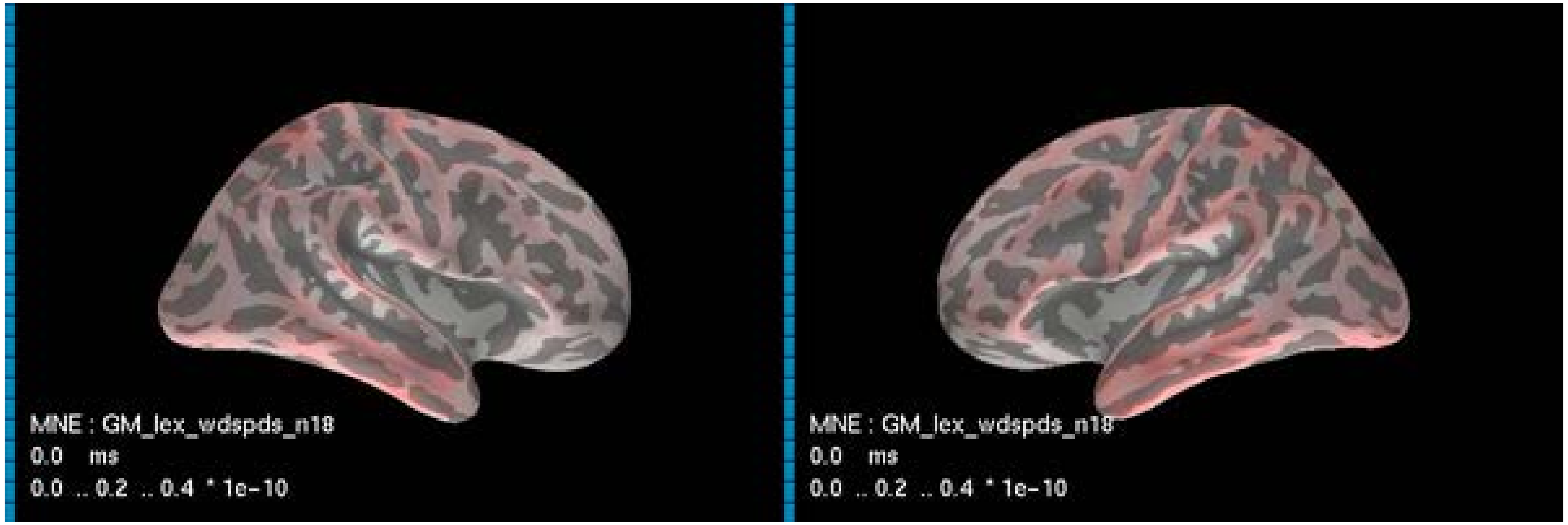
Source Space



Noise/Covariance Matrix



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

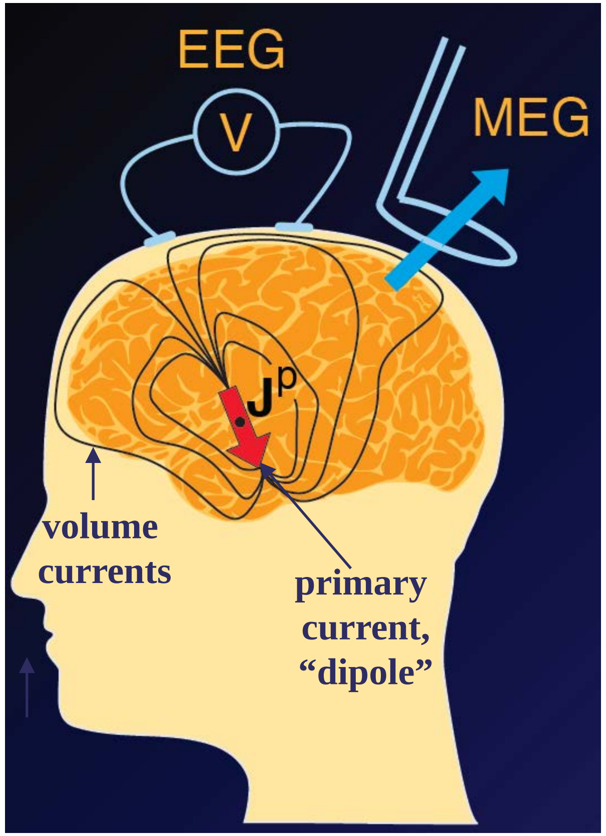


Forward And Inverse Problem

(and some solutions)

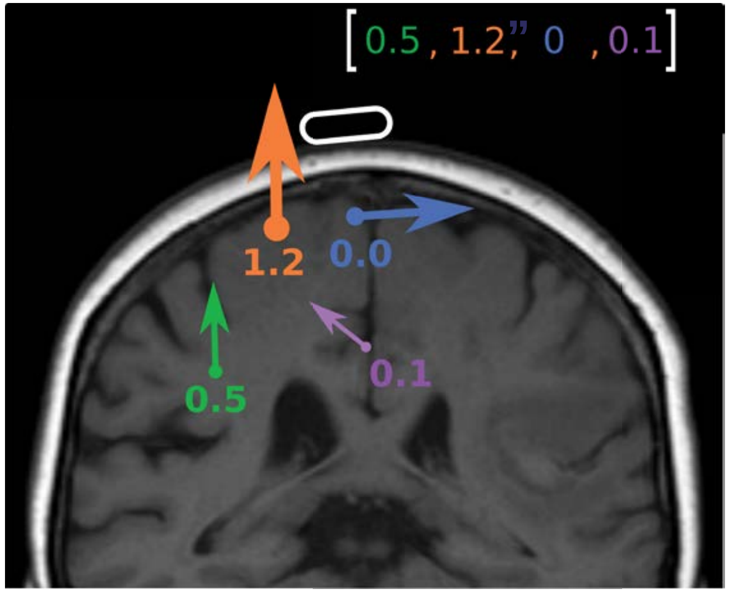
The EEG/MEG Forward Problem

EEG/MEG measure the primary sources indirectly



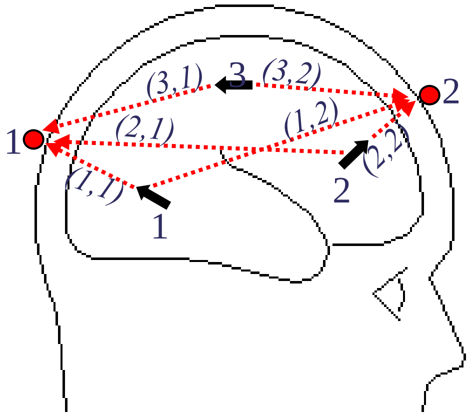
Sensors are differently sensitive to different sources

"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

$$\begin{array}{c} 1 \\ 2 \end{array} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 0.5 & 0 & 0.3 \\ 0 & 1 & -0.3 \end{pmatrix} \begin{pmatrix} j_1 \\ j_2 \\ j_3 \end{pmatrix} \begin{array}{c} 1 \\ 2 \\ 3 \end{array}$$

?

inversion

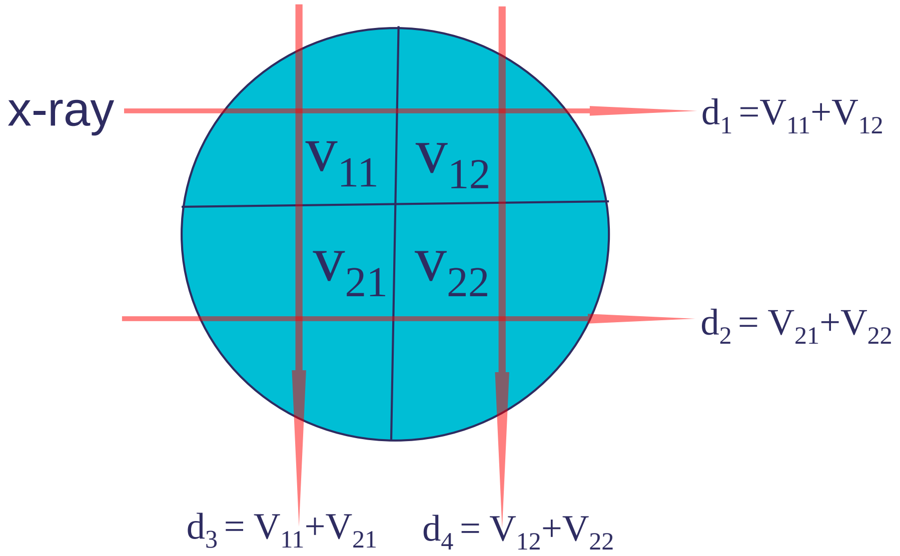
Inverse Operator

dipoles inverse data

$$\begin{array}{c} 1 \\ 2 \\ 3 \end{array} \begin{pmatrix} j_1 \\ j_2 \\ j_3 \end{pmatrix} = \begin{pmatrix} 1.5034 & 0.1241 \\ 0.2483 & 0.9379 \\ 0.8276 & -0.2069 \end{pmatrix} * \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \begin{array}{c} 1 \\ 2 \end{array}$$

EEG/MEG "Scanning" is not "Tomography"

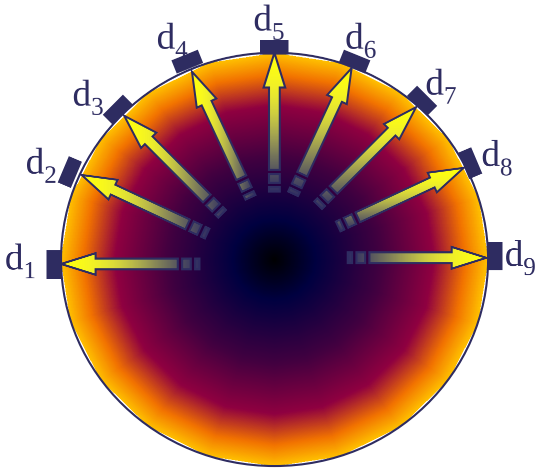
Tomography (CT, fMRI…)



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

Available information is determined by the equipment/experimenter

EEG/MEG



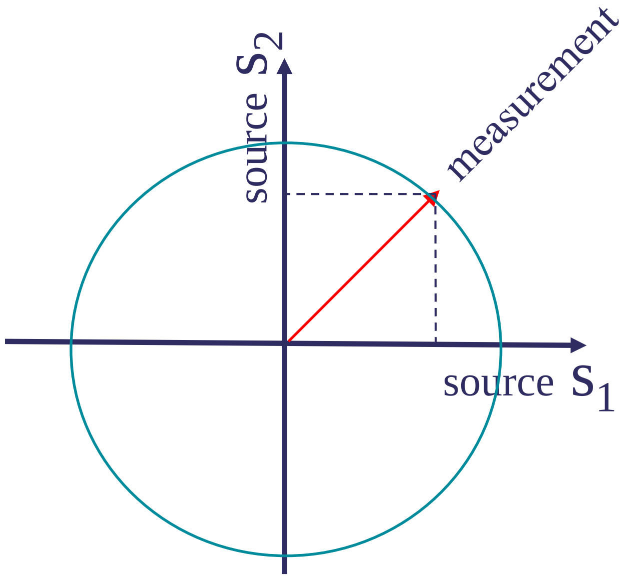
$$d_1 = V_{11}+V_{12}+V_{13}+V_{14}\ \ldots$$
$$d_2 = V_{21}+V_{22}+V_{23}+V_{24}\ \ldots$$

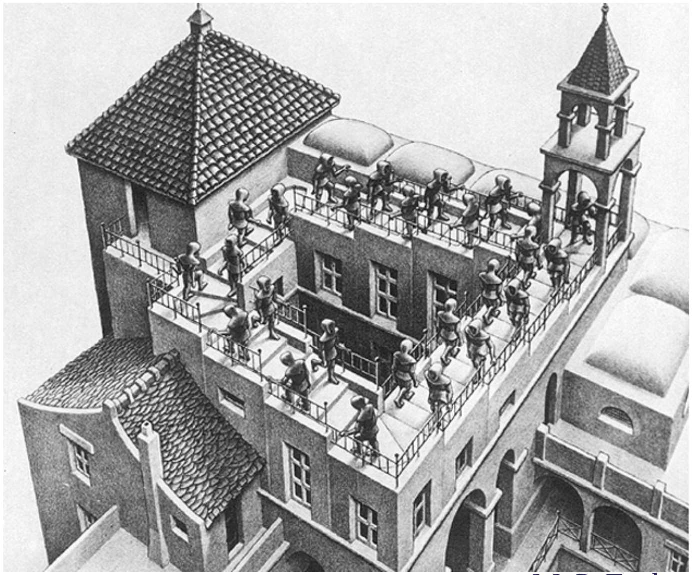
Information is lost during measurement

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)

Why Inverse “Problem”?

Without additional constraints the solution is non-unique, i.e. there are infinitely many solutions

What is the solution to

$x_1 + x_2 = 1$

Maybe

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

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

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

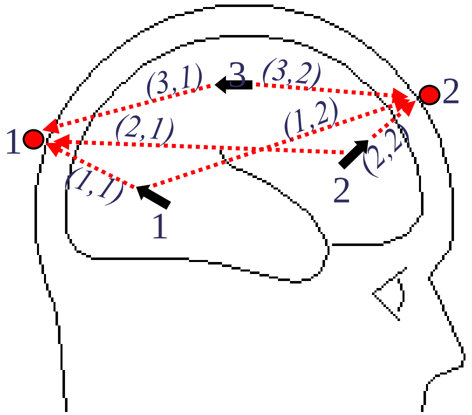
$x_1 = \pi$; $x_2 = (1-\pi)$?

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.

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



Forward Operation

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 Operation

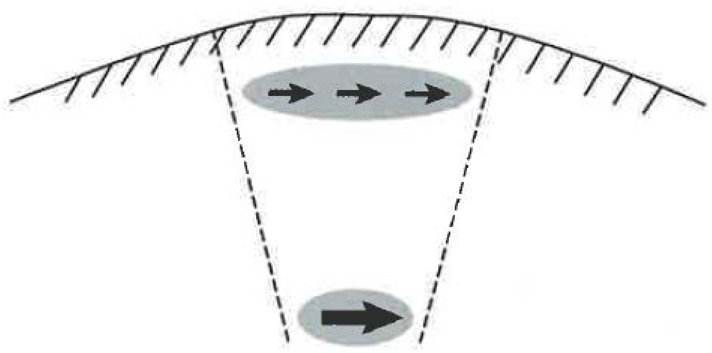
dipoles — inverse — data

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

E.g., 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

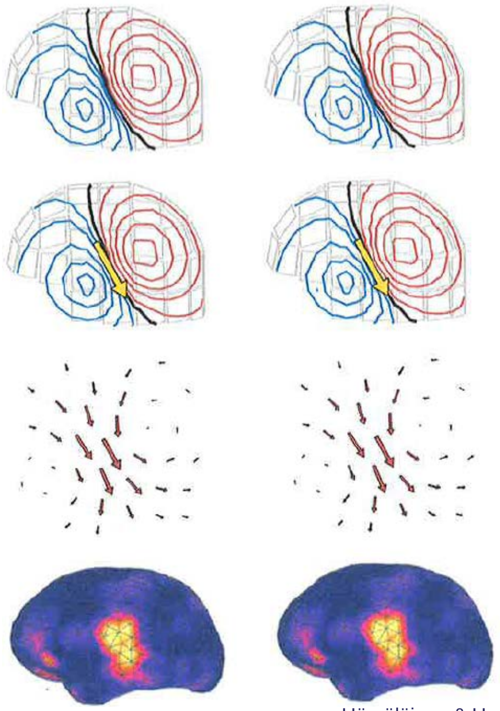
Field Patterns

Same Field Patterns

Dipole Model

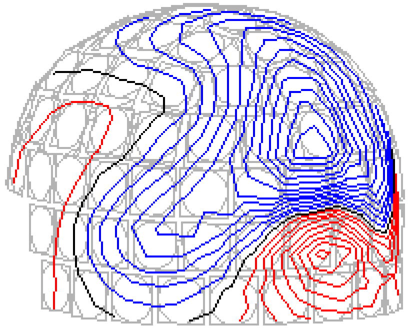
Same Source Estimates

Minimum Norm Estimates

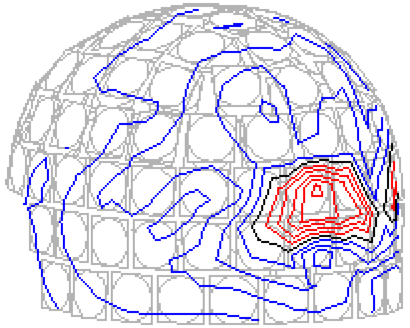


Hämäläinen & Hari, in Brain Mapping: The Methods (2nd), Elsevier 2002

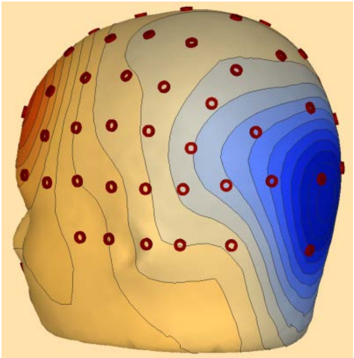
Example: Visually Evoked Activity ~100 ms



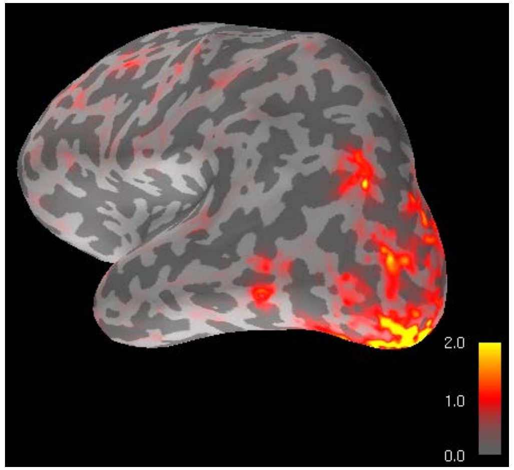
Magnetometers



Gradiometers

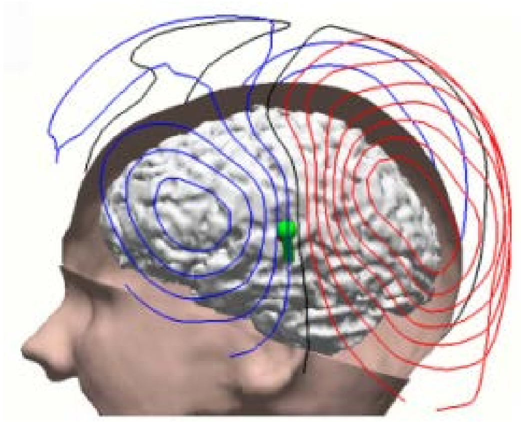


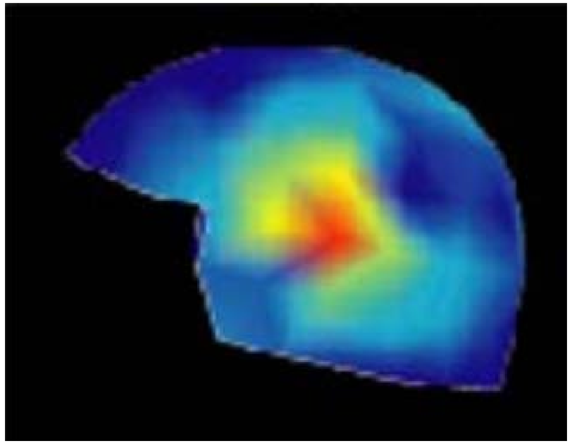
EEG

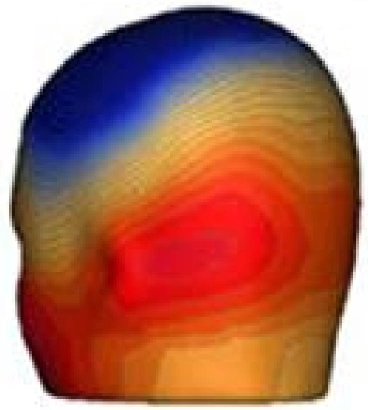


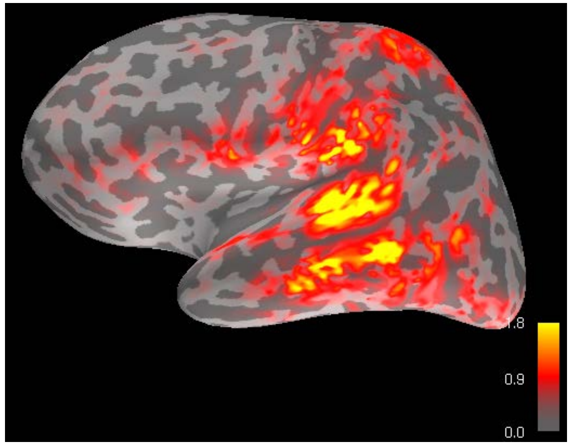
Minimum Norm Estimate

Example: Auditorily Evoked Activity





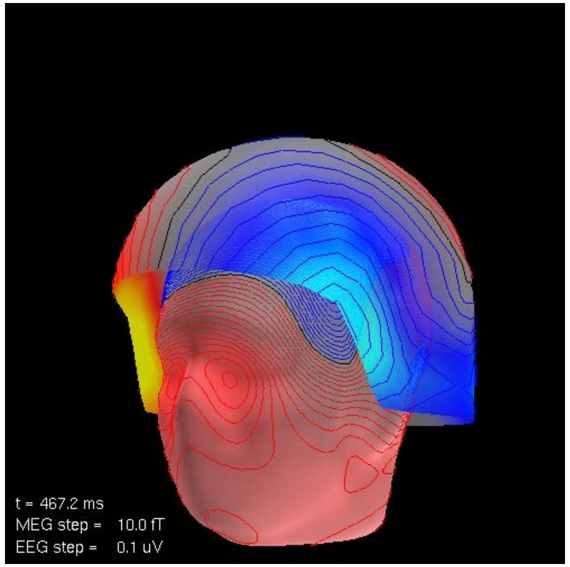


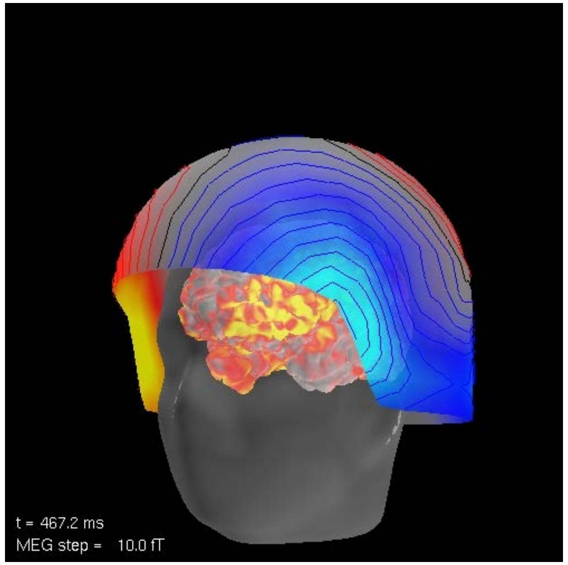


Minimum Norm Estimate

Reminder: Artefacts in EEG and MEG Will End Up in Source Space

Eye Blink





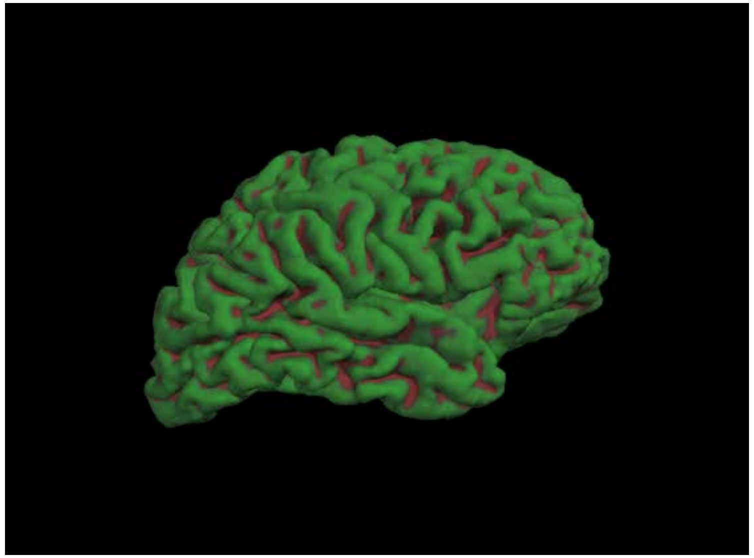
This will affect all source estimation methods –
get rid of your artefacts beforehand.

The Forward Problem and Head Modelling

Source Space and Head Model

Source Space

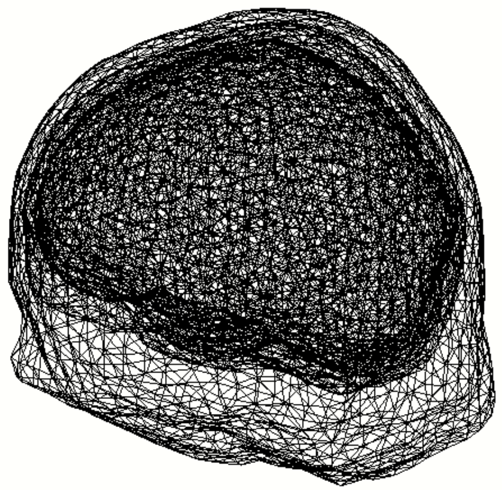
Where active sources may be located,
e.g. grey matter, 3D volume



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

Volume Conductor/Head Model

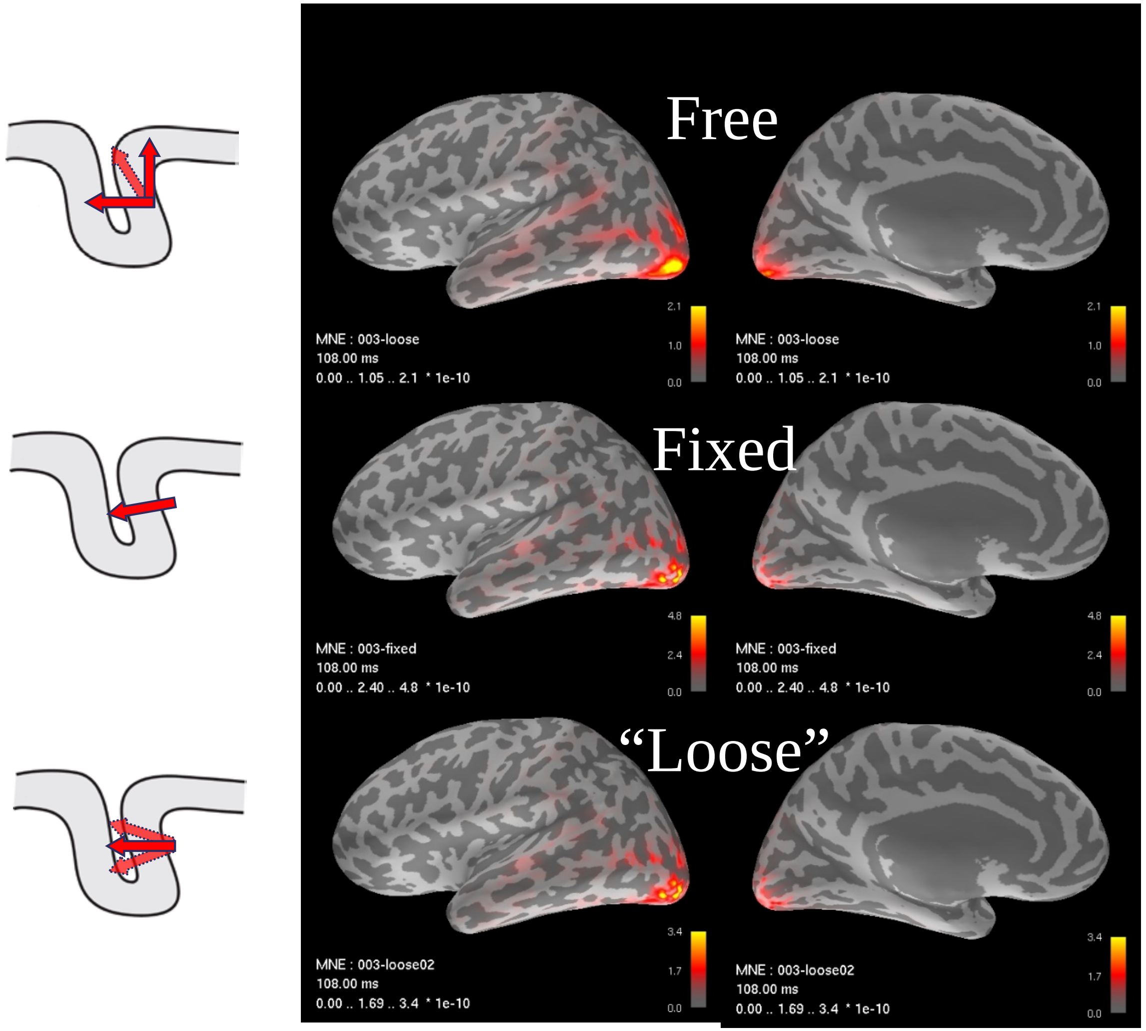
How we model conductivities/currents/potentials/fields in the head
e.g. sphere or realistic 1- or 3-compartments from MRI



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

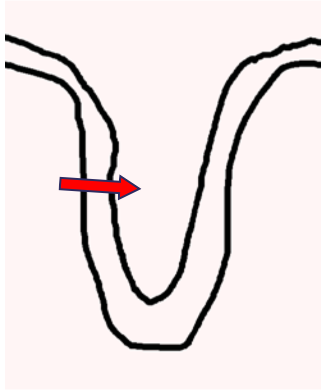
SPM uses the same "canonical mesh" as source space for every subjects, but adjusts it individually.

Source Orientation Constraints

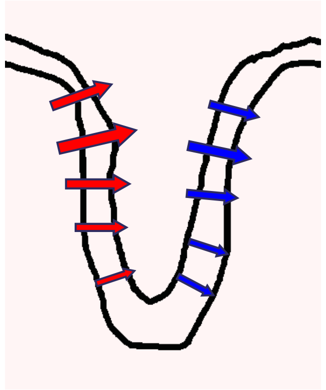


Direction of Current Flow

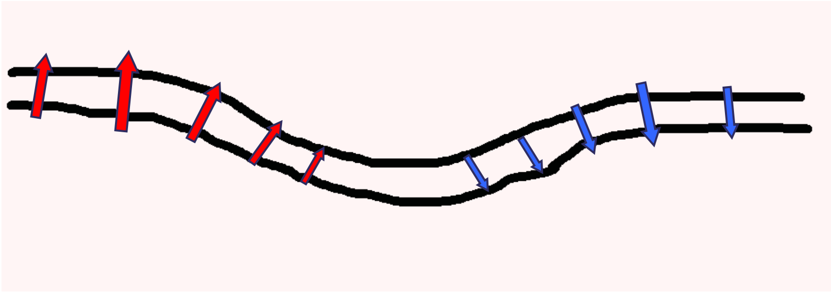
Dipole Source



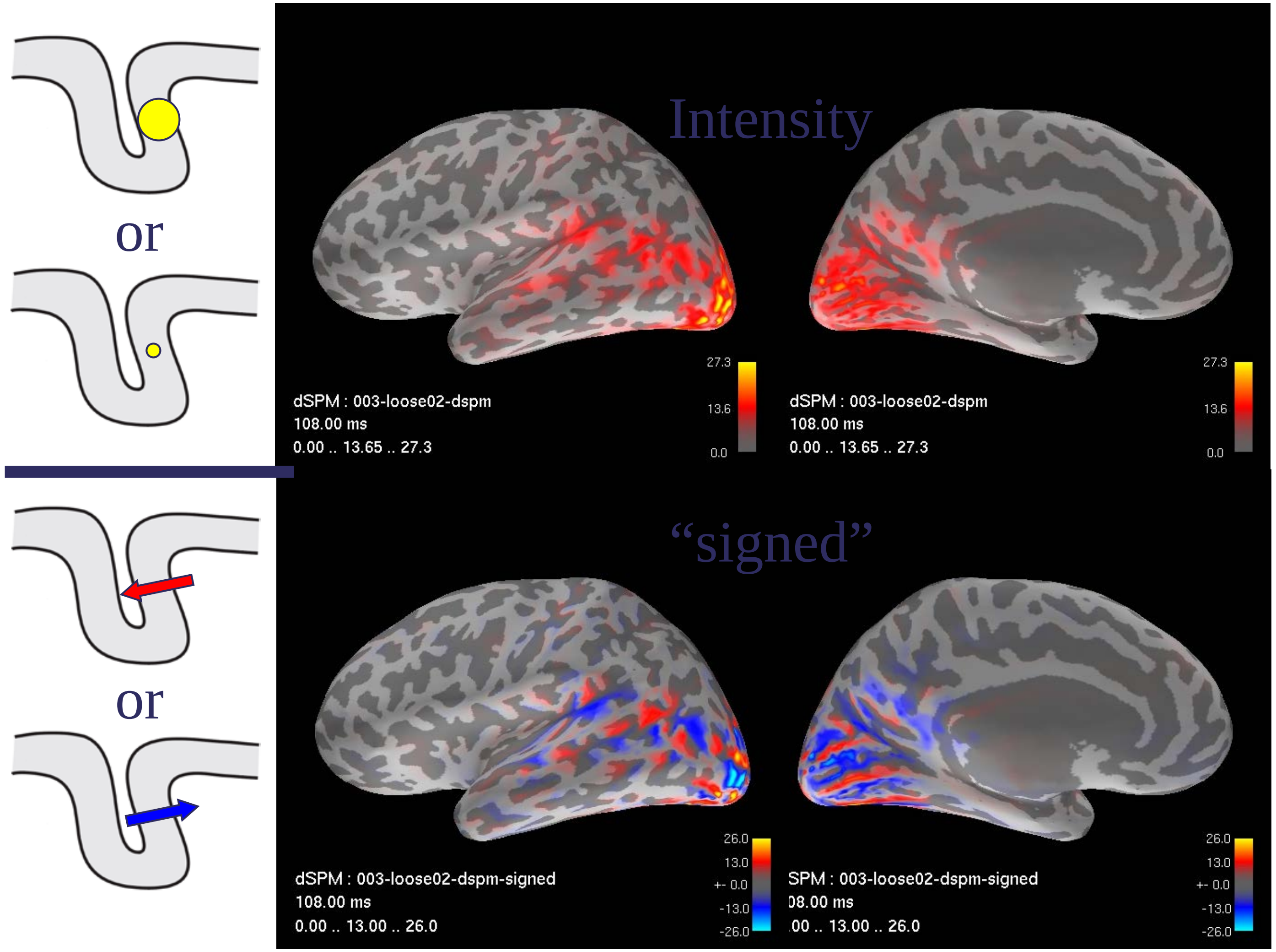
Distributed Source



Distributed Source, Inflated Surface



Direction of Current Flow



Solutions To The Inverse Problem – Source Estimation

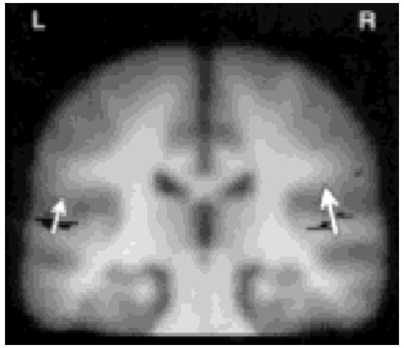
Paths To Uniqueness

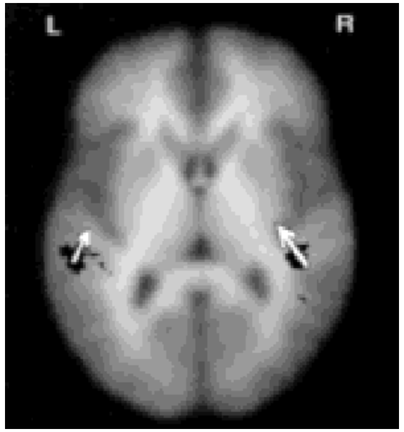
Dipole Fitting/Scanning

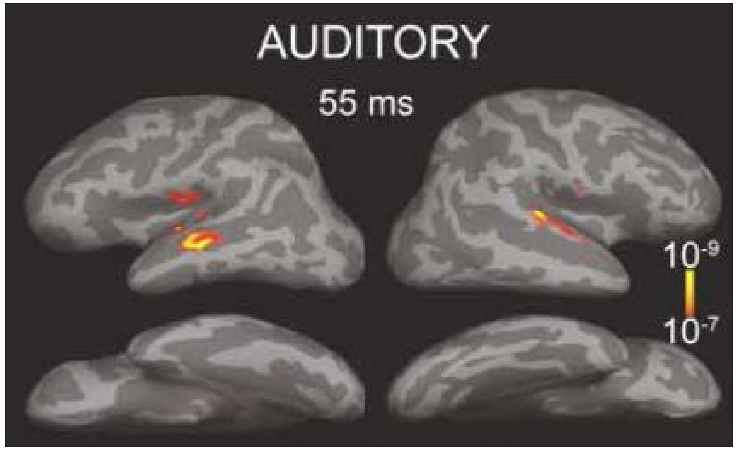
1. Assume there are only a few distinct sources
2. Iteratively adjust the location, orientation and strength of a few dipoles…
3. …until the result best fits the data

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

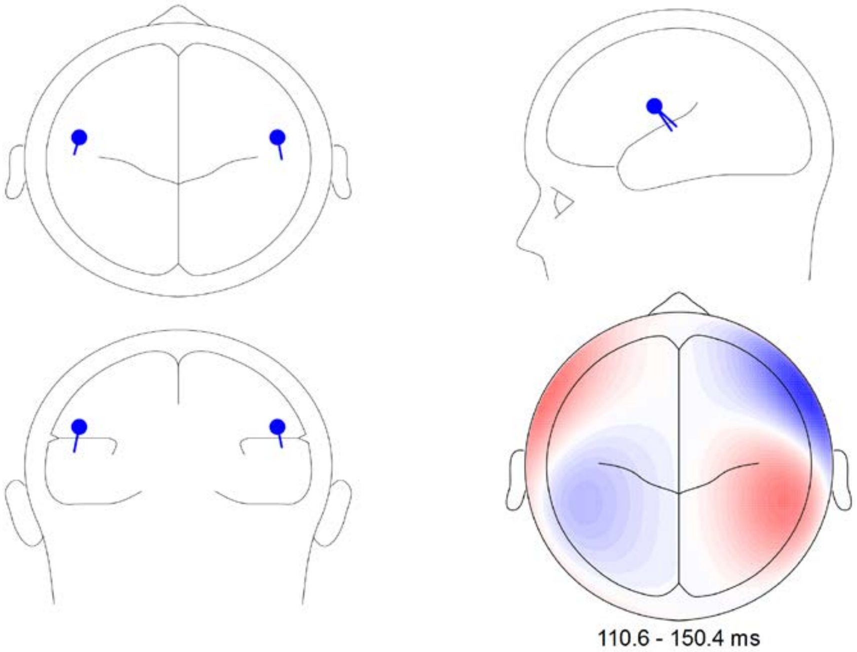


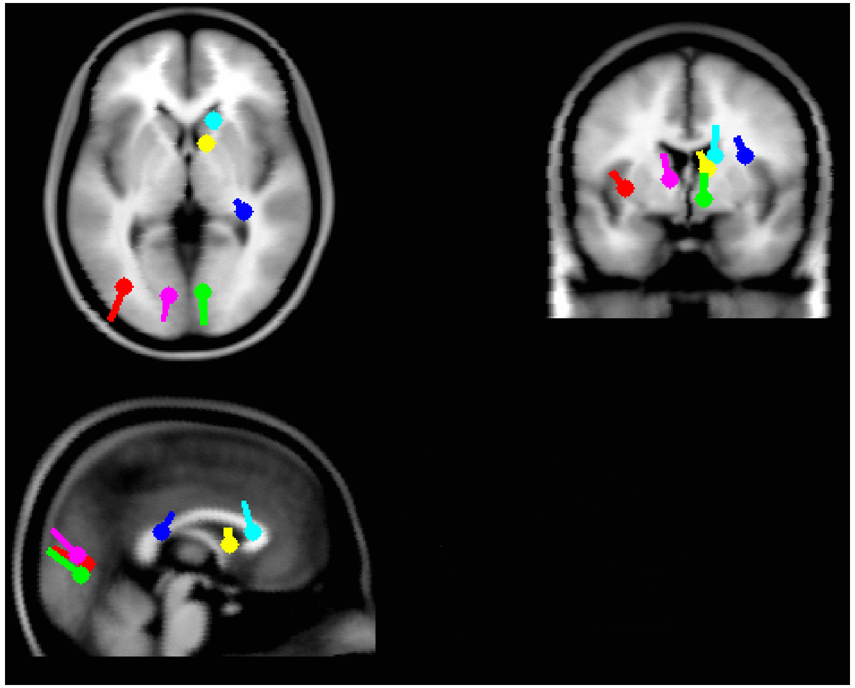




Hypothesis Testing - Dipole Fitting

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**.

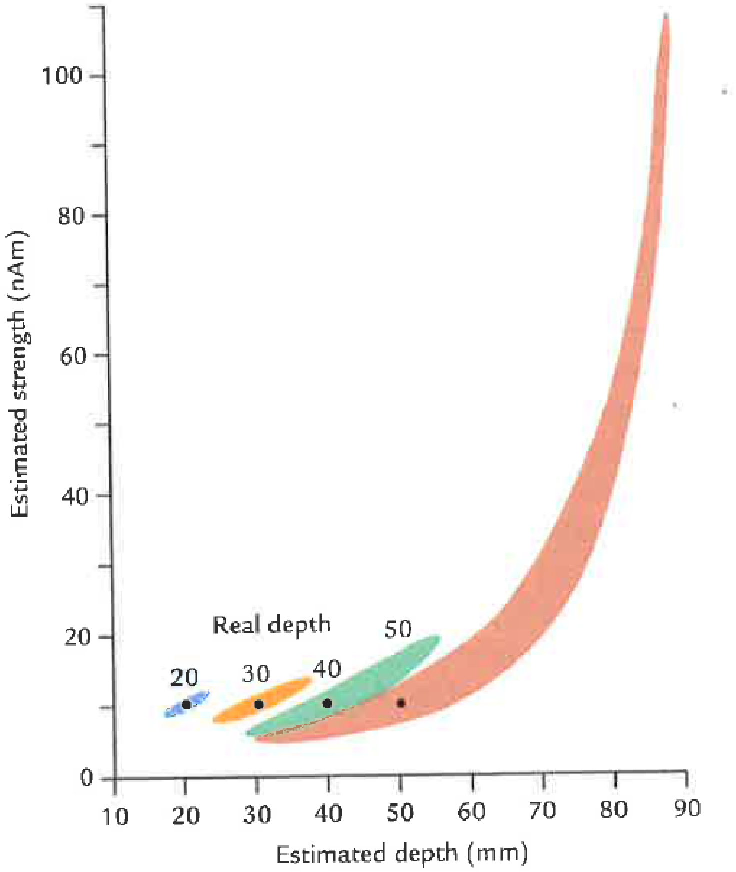




It can be hard to choose the appropriate number of dipoles – a priori knowledge is required.
Solutions for several/many dipoles can get stuck in local minima, and may not be robust to noise.

Assumptions Cannot Completely Remove Uncertainty

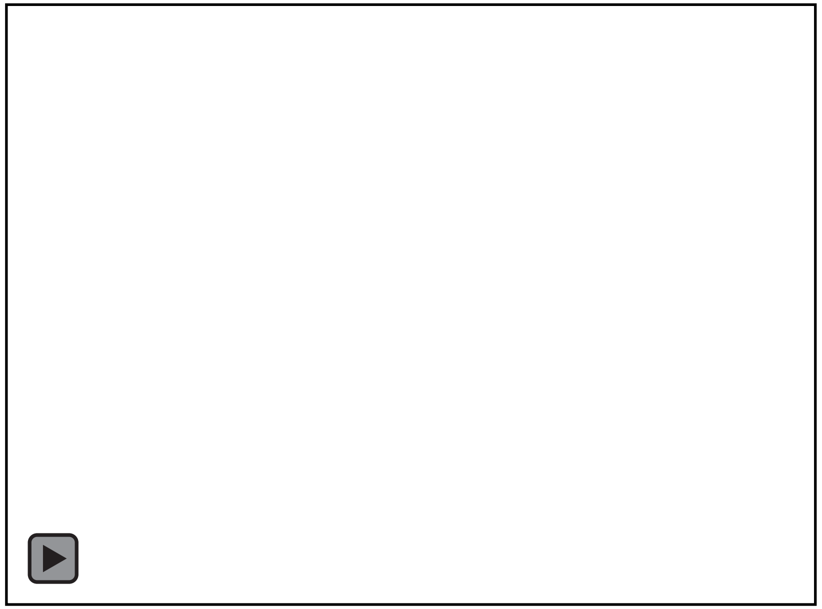
95% CIs for single dipole source



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.

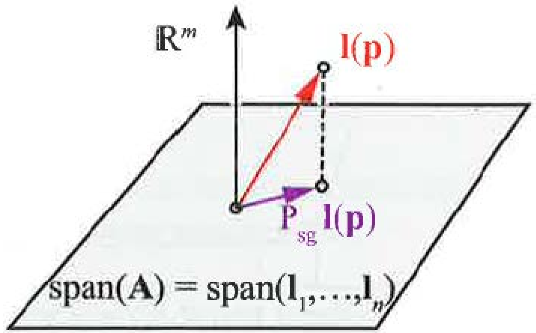


http://www.cogsci.ucsd.edu/~sereno/movies.html

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”.

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 $\mathrm{subcorr}\{\boldsymbol{G}, \hat{\Phi}_s\}$.
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 $\mathbf{w}_\mathrm{i}$ 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: $\boldsymbol{min}(\mathbf{w}_i^T \mathbf{C} \mathbf{w}_i)$

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

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.
But note: The “spatial filter” interpretation applies to all linear methods, including MNE-type methods.

e.g. Hauk et al., Neuroimage 2022, https://pubmed.ncbi.nlm.nih.gov/35390459/
Hauk&Stenroos, HBM 2013, https://pubmed.ncbi.nlm.nih.gov/23616402/ ,

Beamformers are adaptive - i.e. not strictly linear

The “linearly-constrained maximum-variance” (LCMV) beamformer

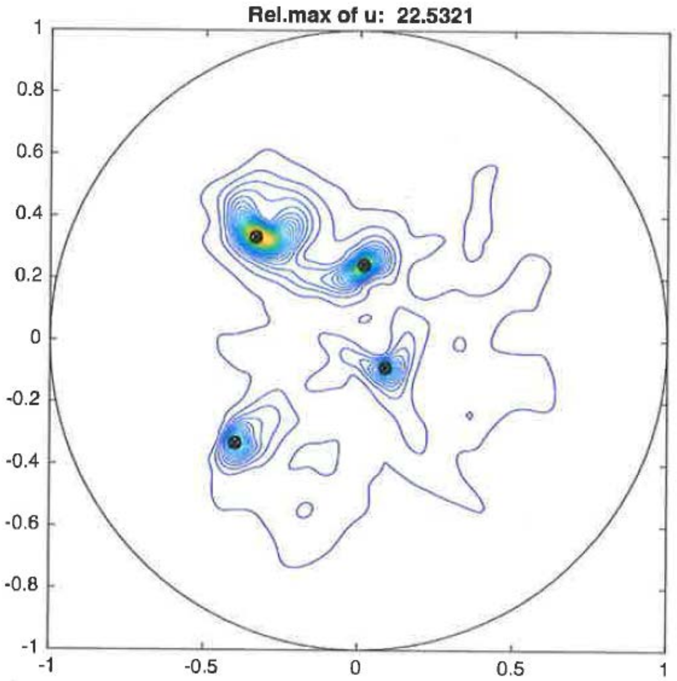
$$\mathbf{SF}_{LCMV}(\mathrm{i}) = \frac{\tilde{\mathbf{L}}_{.i}^{T}\mathbf{C}_{d}^{-1}}{\tilde{\mathbf{L}}_{.i}^{T}\mathbf{C}_{d}^{-1}\tilde{\mathbf{L}}_{.i}}$$

depends on the data covariance matrix (“adaptive”).

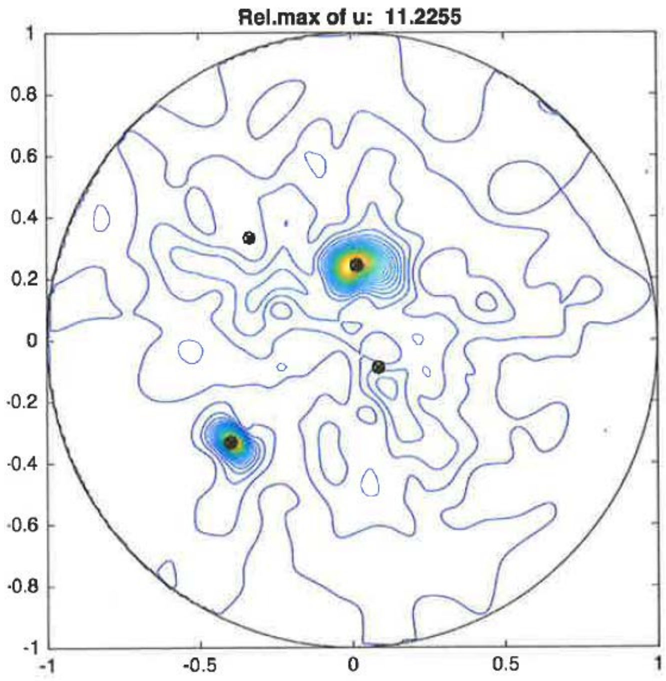
Beamformers result in linear transformations of the data (“spatial filters”),
but those transformations strongly depend on the data of interest.

=> Beamformers are data-dependent and not linear with respect to the sources of interest.

Beamforming Is Problematic For Highly Synchronous Sources



4 non-synchronous sources



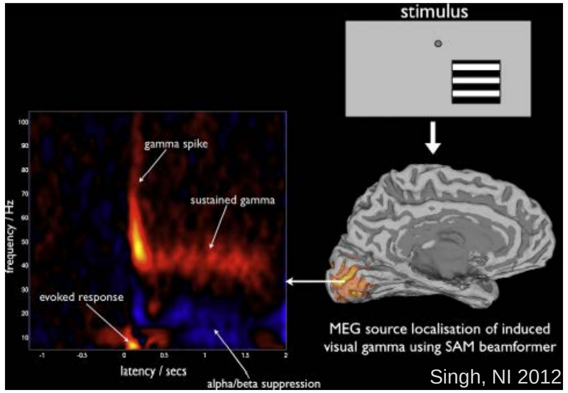
2 non-synchronous,
2 synchronous sources

Beamformers are designed for – and work best for – small numbers of focal sources with uncorrelated time courses.

Ilmoniemi and Sarvas, "Brain Signals", MIT 2019
Hauk et al., Neuroimage 2022, https://pubmed.ncbi.nlm.nih.gov/35390459/

Beamformers Are Popular for Rhythmic Brain Activity and Resting State Activity

Visual Gamma Band Response

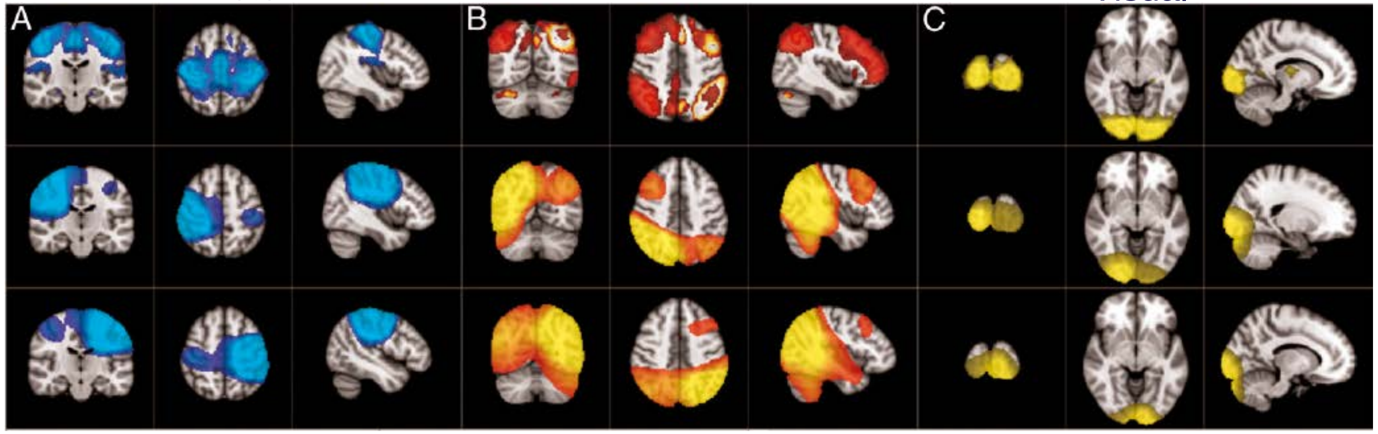


Motor

Fronto-Parietal

Visual

Resting State Networks



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

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

(ignore noise for now)

subject to constraint

$$||\boldsymbol{s}||_{\mathbf{2}} = \mathit{min}$$

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 multiple sources).

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

Minimising the L2 norm, $\|\boldsymbol{s}\|_{\boldsymbol{2}} = |s_{\boldsymbol{1}}|^{\boldsymbol{2}}+|s_{\boldsymbol{2}}|^{\boldsymbol{2}}+\ldots+|s_{\boldsymbol{N}}|^{\boldsymbol{2}}$ penalizes large values in **s**
=> “smooth”

Minimising the L1 norm, $\|\boldsymbol{s}\|_{\boldsymbol{1}} = |s_{\boldsymbol{1}}|+|s_{\boldsymbol{1}}|+\ldots+|s_{\boldsymbol{N}}|$ prefers large values in **s**
$\Rightarrow$ “sparse”

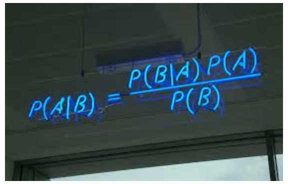
For example:

$$x_1 + 2x_2 = 1$$

L2 solution: (0.2, 0.4)
L2-norm $0.2^2+0.4^2 \sim 0.45$, 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/

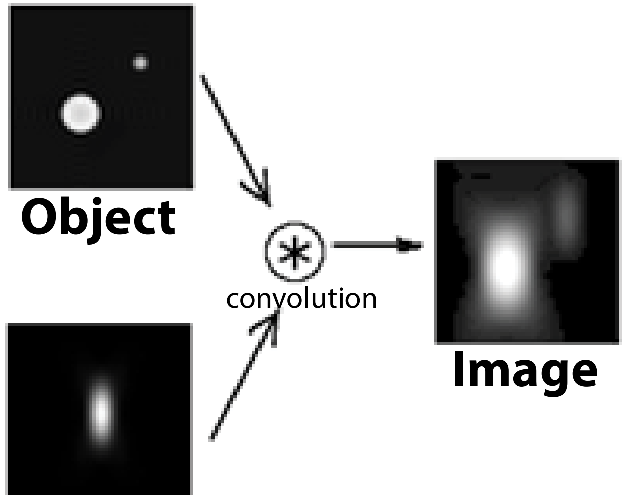
"Most likely" is still not what we want to know –
Does the method do what we want it to do?

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

Just because the brain is complicated doesn’t mean source estimation has to be complicated

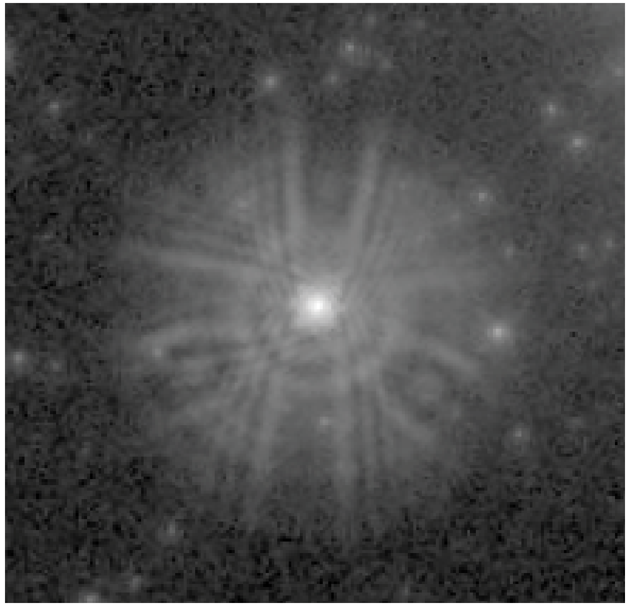
The Superposition Principle
A “Constraint-Free” Interpretation of Linear Methods

Microscopy



Point-Spread Function
PSF

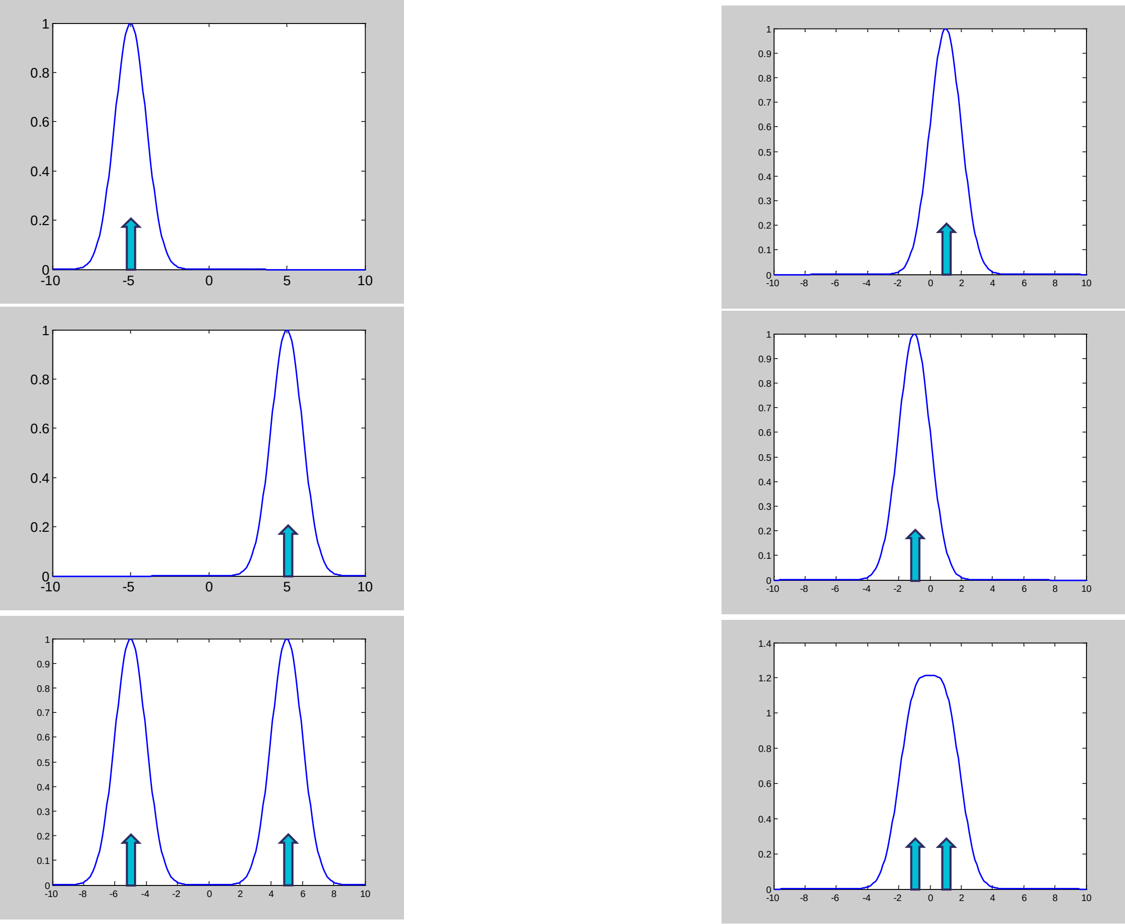
Astronomy



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

For Linear Methods We Can Find Out If They Do What We Want

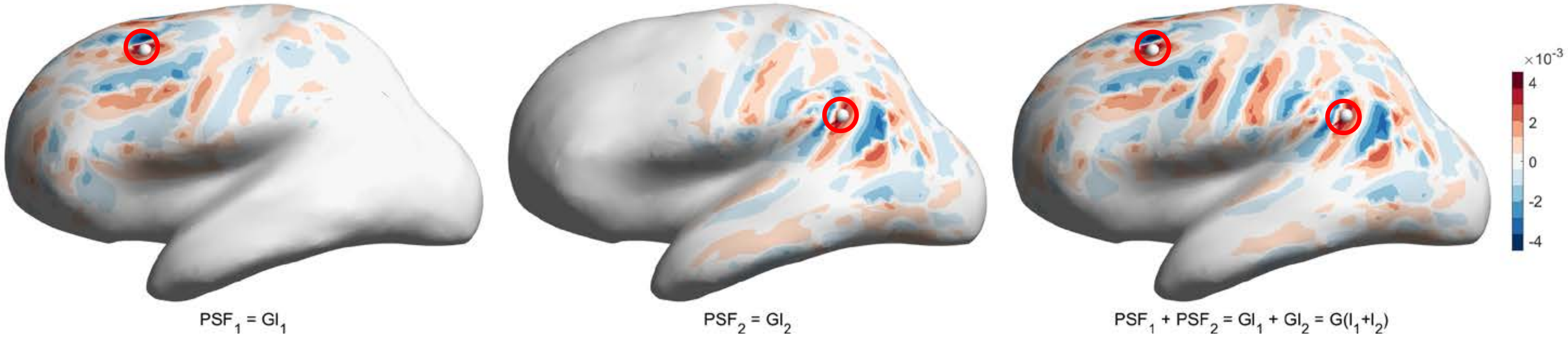
Superposition Principle



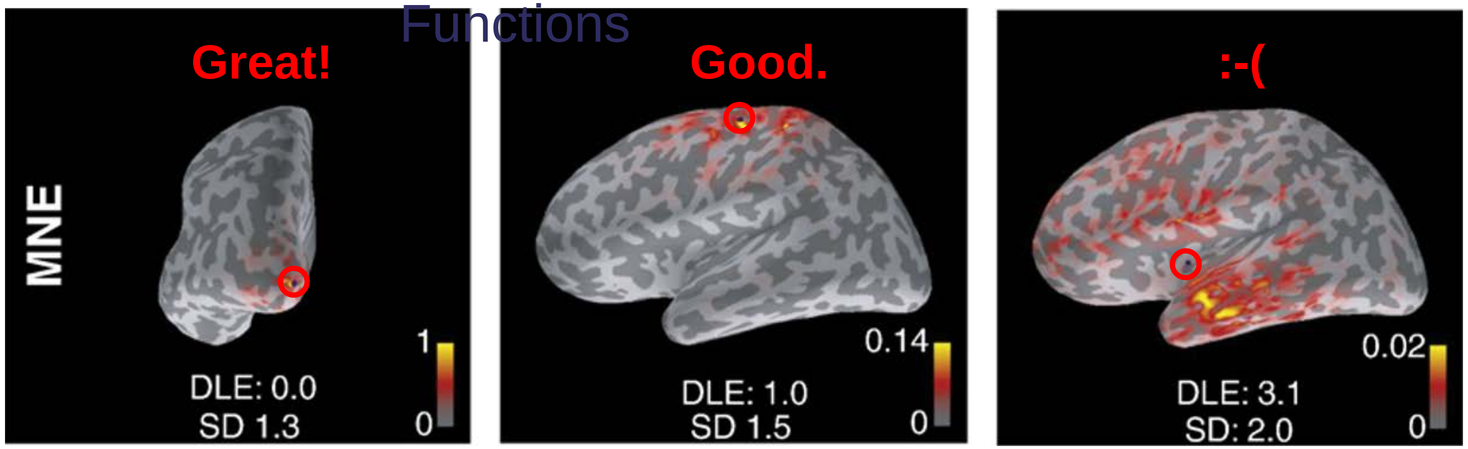
If you know the behaviour for point sources,
you can predict the behaviour for complex sources

Linear Methods – Superposition Principle

Superposition In Source Space



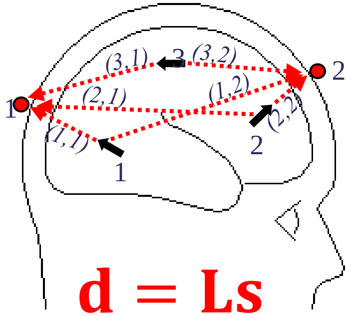
Example Point-Spread Functions



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

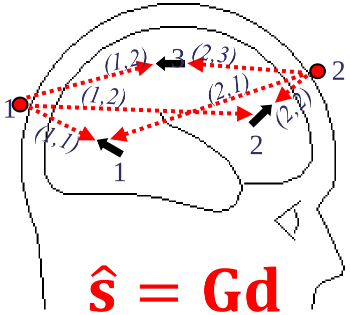
Resolution Matrix

Forward Problem



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

Linear Inverse Problem



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

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

Relationship between estimated and true source distribution.

e.g. Hauk et al., Neuroimage 2022, https://pubmed.ncbi.nlm.nih.gov/35390459/

Creating 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)**:

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

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

L2-MNE (non-weighted, non-normalised) has the ideal resolution matrix in the least-squares sense.

e.g. Hauk et al., Neuroimage 2022, https://pubmed.ncbi.nlm.nih.gov/35390459/

Weighted and Noise-Normalised MNE Methods

L2-MNE

$$\boldsymbol{K}_{MNE} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1}$$

(Depth-)Weighted MNE

$$\boldsymbol{K}_{wMN} = \mathbf{D}\mathbf{L}^T(\mathbf{L}\mathbf{D}\mathbf{L}^T)^{-1}$$

"Alleviating Depth Bias"

dSPM

$$\boldsymbol{K}_{dSPM} = \boldsymbol{W}_{dSPM}\boldsymbol{K}_{MNE} \quad \boldsymbol{W}_{dSPM} = \sqrt{\boldsymbol{diag}(\boldsymbol{K}_{MNE}\boldsymbol{C}\boldsymbol{K}_{MNE}^T)}^{\,-1}$$

eLORETA

$$\boldsymbol{K}_{eLOR} = \boldsymbol{D}_{eLOR}\mathbf{L}^T(\mathbf{L}\boldsymbol{D}_{eLOR}\mathbf{L}^T)^{-1}$$

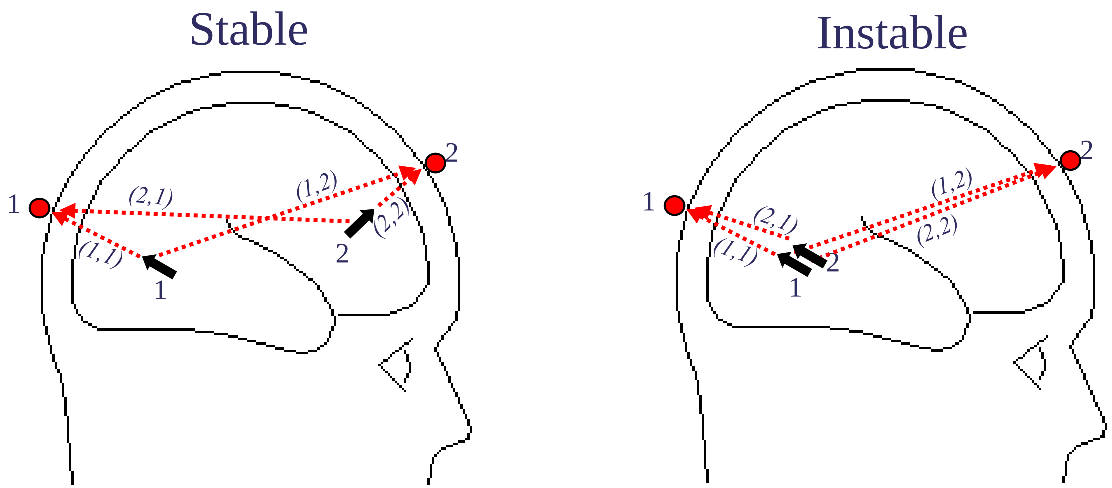
"Zero
Dipole Localization Error"

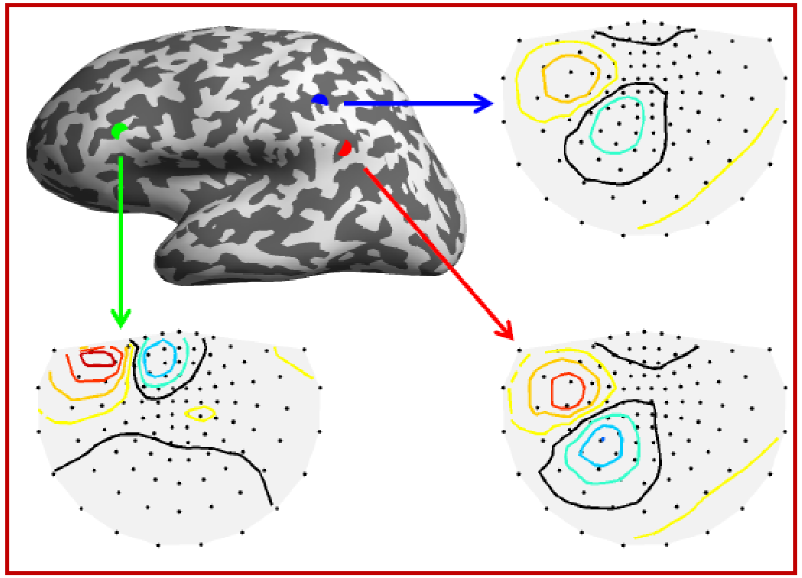
sLORETA

$$\boldsymbol{K}_{sLOR} = \boldsymbol{W}_{sLOR}\boldsymbol{K}_{MNE} \quad \boldsymbol{W}_{sLOR} = \sqrt{diag(\boldsymbol{R}_{MNE})}^{\,-1}$$

Noise and Regularisation

(In)Stability – Sensitivity to Noise





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

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

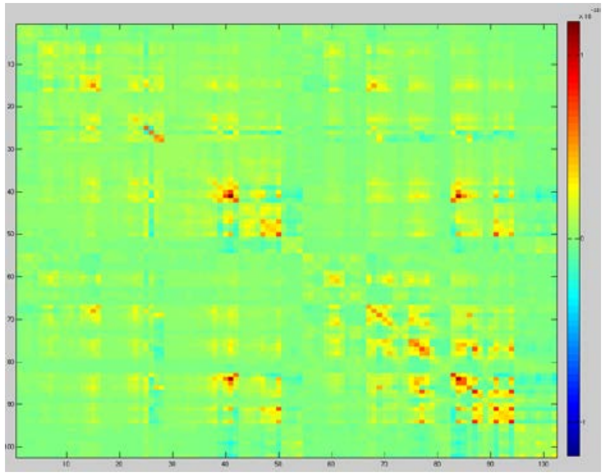
Some channels are noisier than others

⇒They should get different weights in your analysis

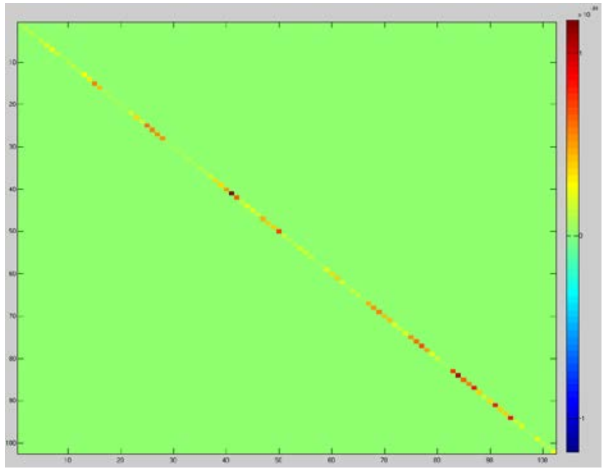
Sensors are not independent

=> 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

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

This is equivalent to minimising the cost function

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

We can give sensors different weightings,

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

$\|\mathbf{C}^{-1}(\mathbf{L}\boldsymbol{s} - \mathbf{d})\|^2 = \|\mathbf{L}\boldsymbol{s} - \mathbf{d}\|_C^2 = \mathrm{e}$

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

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

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

Whitening and Choice of Regularisation Parameter

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

can also be written as

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

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($\mathbf{I}$).

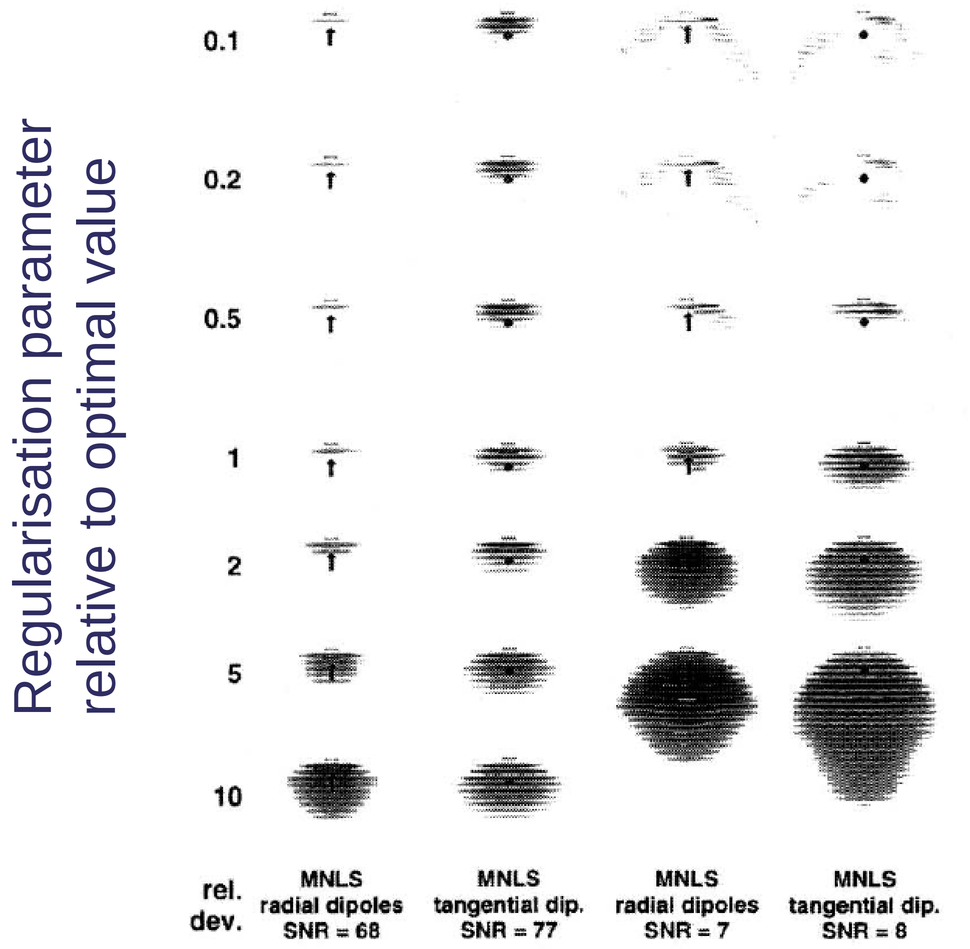
"Whitening" turns data and leadfield into signal-to-noise ratios.

$\tilde{\mathbf{L}}$ 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

MRC Cognition and Brain Sciences Unit

UNIVERSITY OF CAMBRIDGE

Thank you