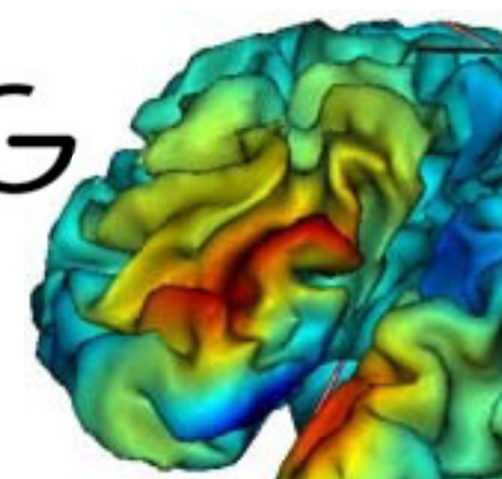
Source localization with: Minimum norm estimates (MNE)

Alexandre Gramfort

alexandre.gramfort@telecom-paristech.fr

Assistant Prof. Telecom ParisTech

CEA - Neurospin, France

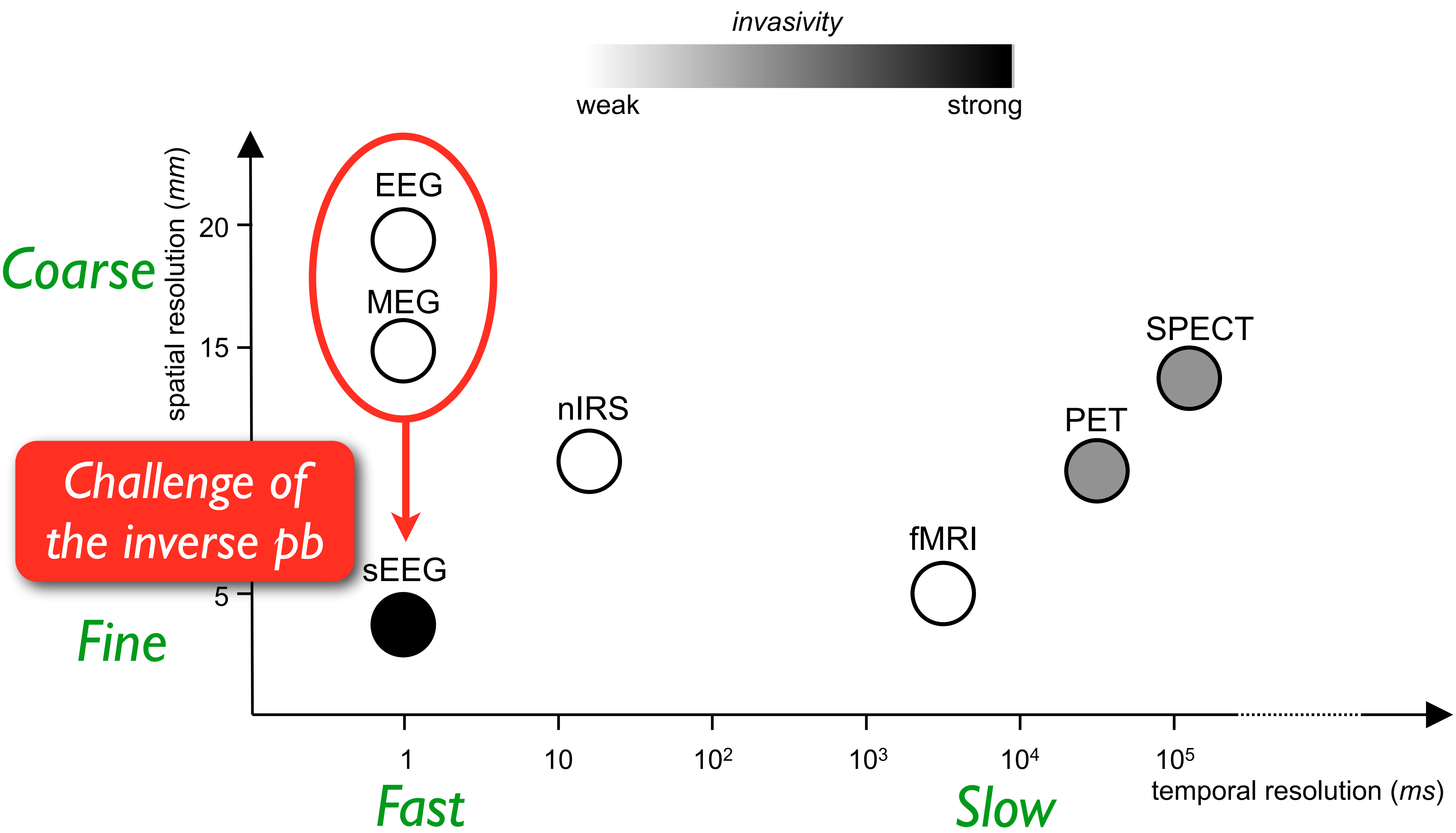


Natmeg - Jan. 2014

Relevant background on M/EEG

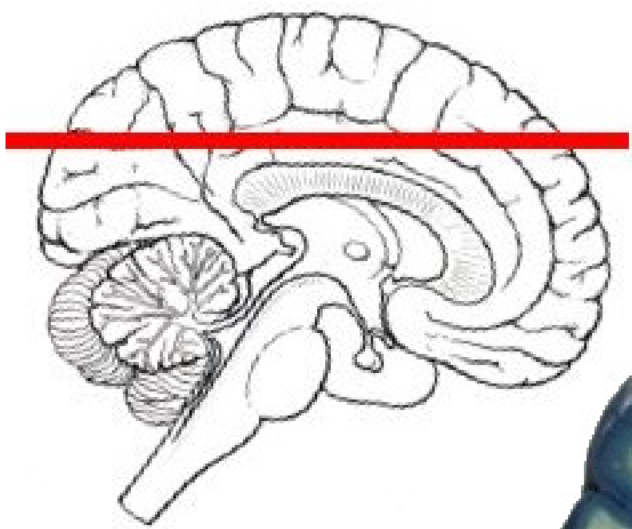
THM: Take home message (not Theorem)...

Functional neuroimaging



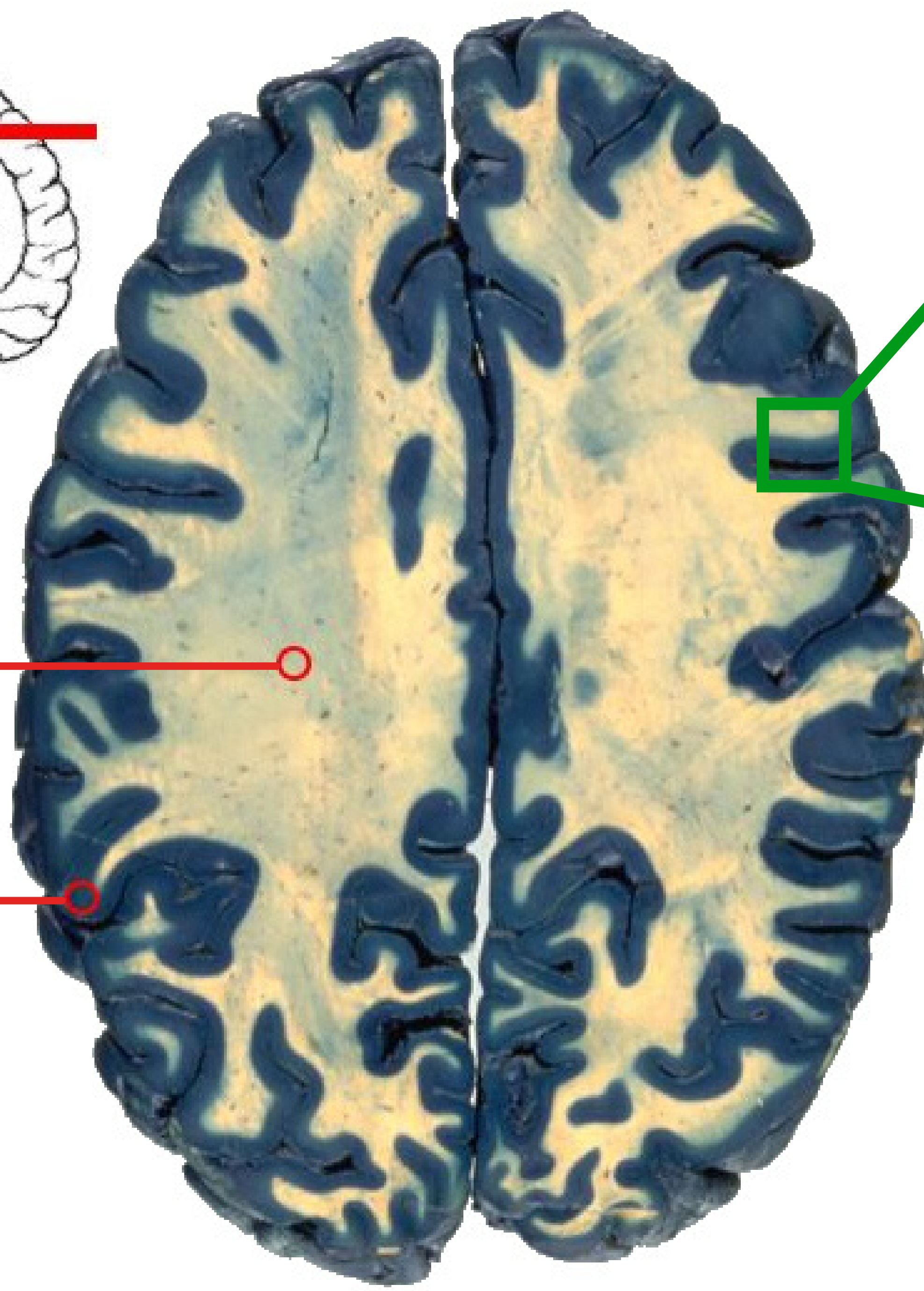
Brain anatomy

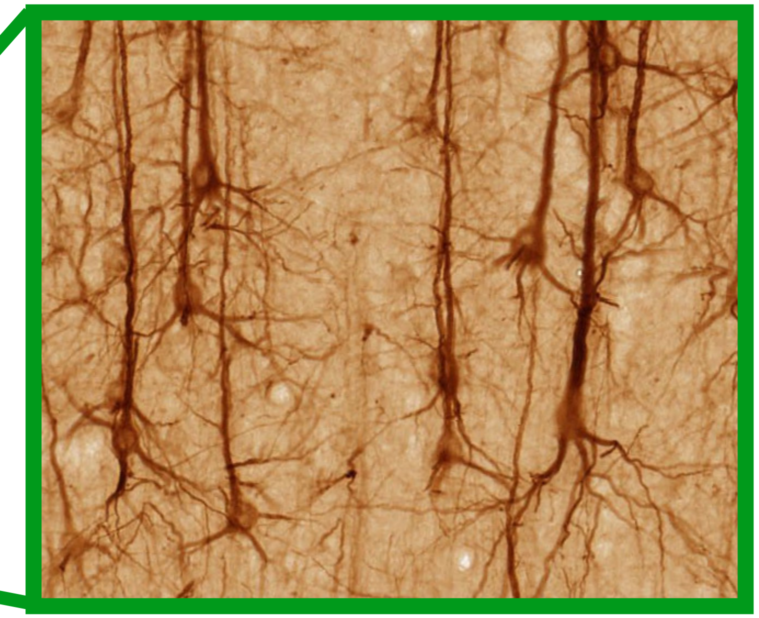
Axial slice



White matter

Gray matter





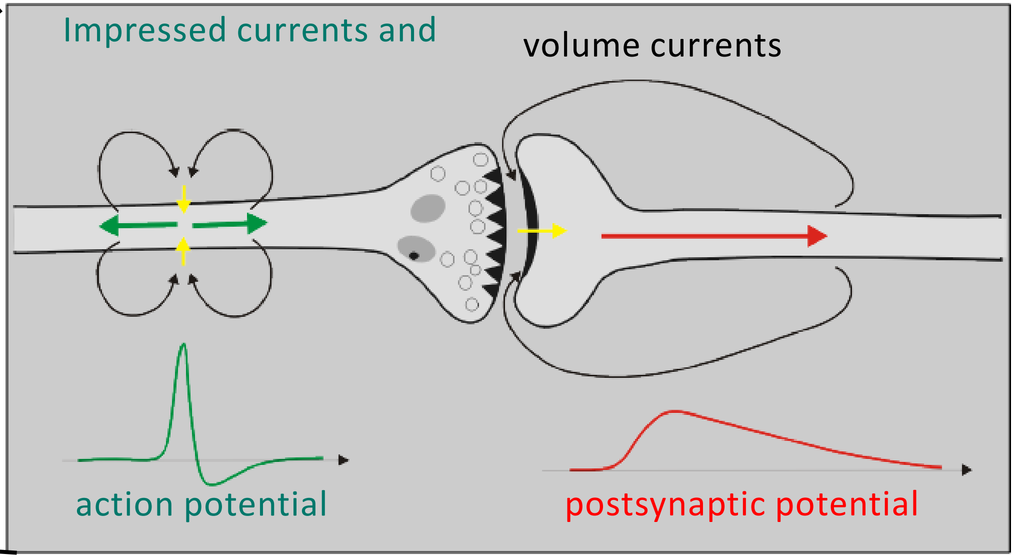
Neurons

in the gray matter

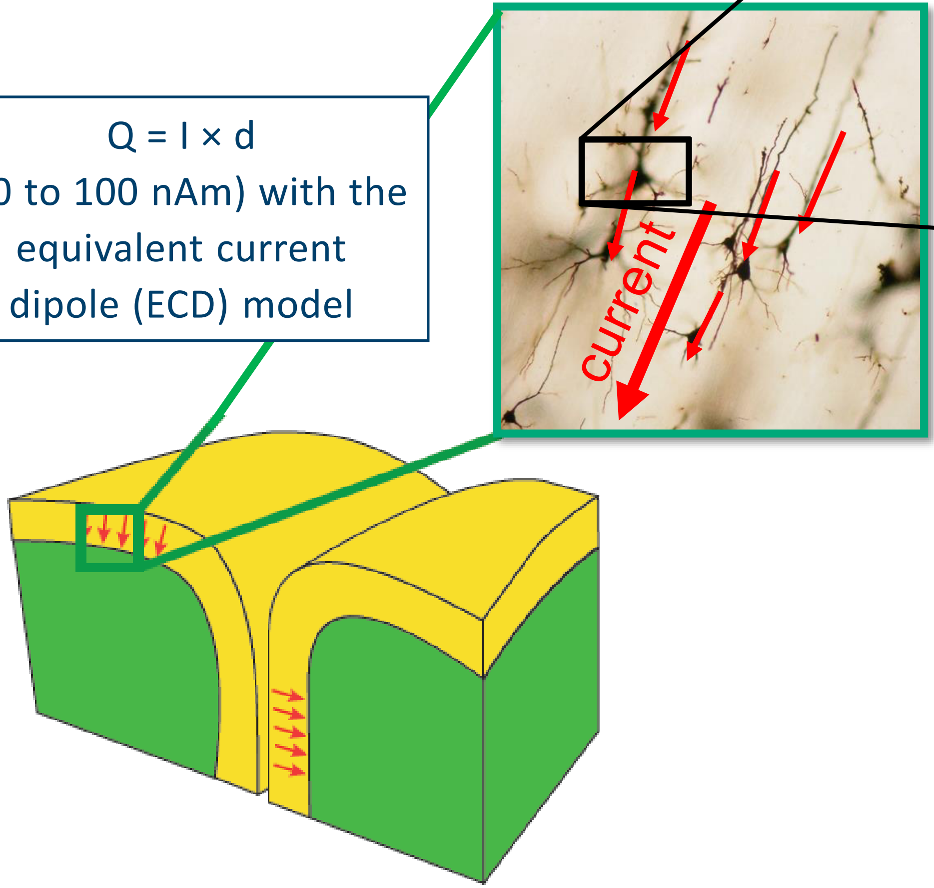
Source: dartmouth.edu

Neurons as current generators

Large cortical pyramidel cells organized in macro-assemblies with their **dendrites normally oriented to the local cortical surface**



Q = I × d
(10 to 100 nAm) with the equivalent current dipole (ECD) model



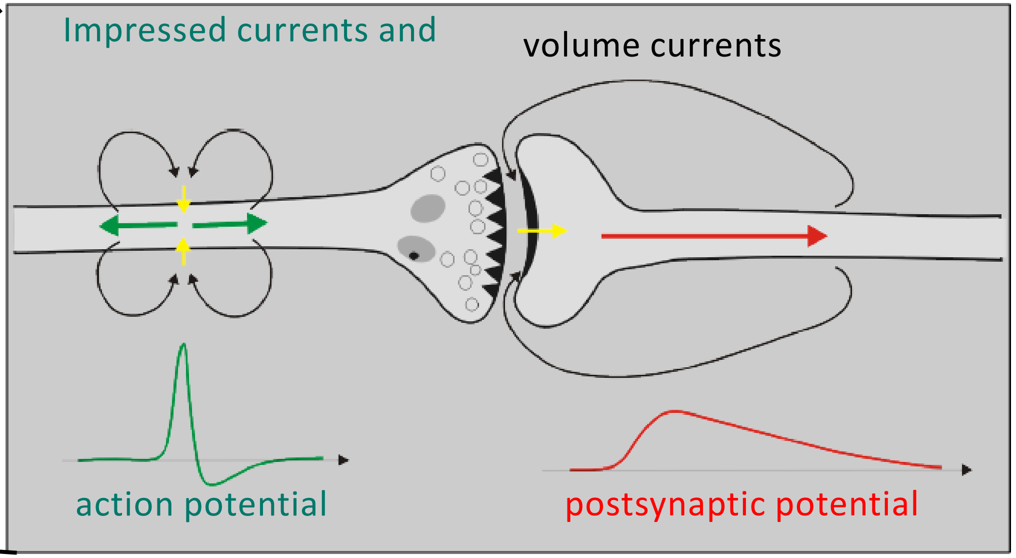
Impressed currents $\mathbf{J}_i(\mathbf{r})$
- due to electrochemical gradients and open ion channels across the cell membrane

Primary currents $\mathbf{J}_p(\mathbf{r})$
- due to impressed currents
- currents inside the dendrites and in their vicinity

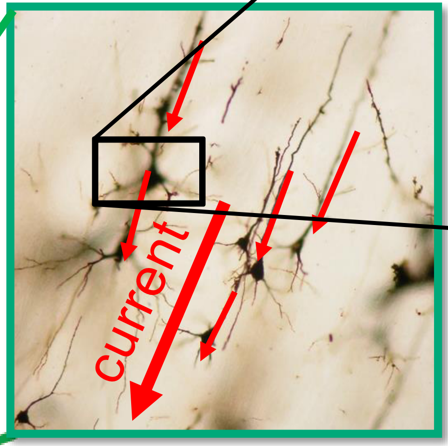
Volume currents $\mathbf{J}_v(\mathbf{r})$
- due to $\mathbf{J}_i$ and $\mathbf{J}_p$
- passive, ohmic current flow

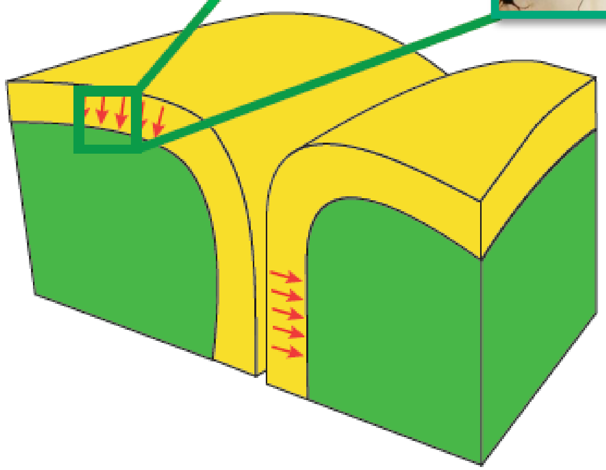
Neurons as current generators

Large cortical pyramidel cells organized in macro-assemblies with their **dendrites normally oriented to the local cortical surface**



Q = I × d
(10 to 100 nAm) with the equivalent current dipole (ECD) model





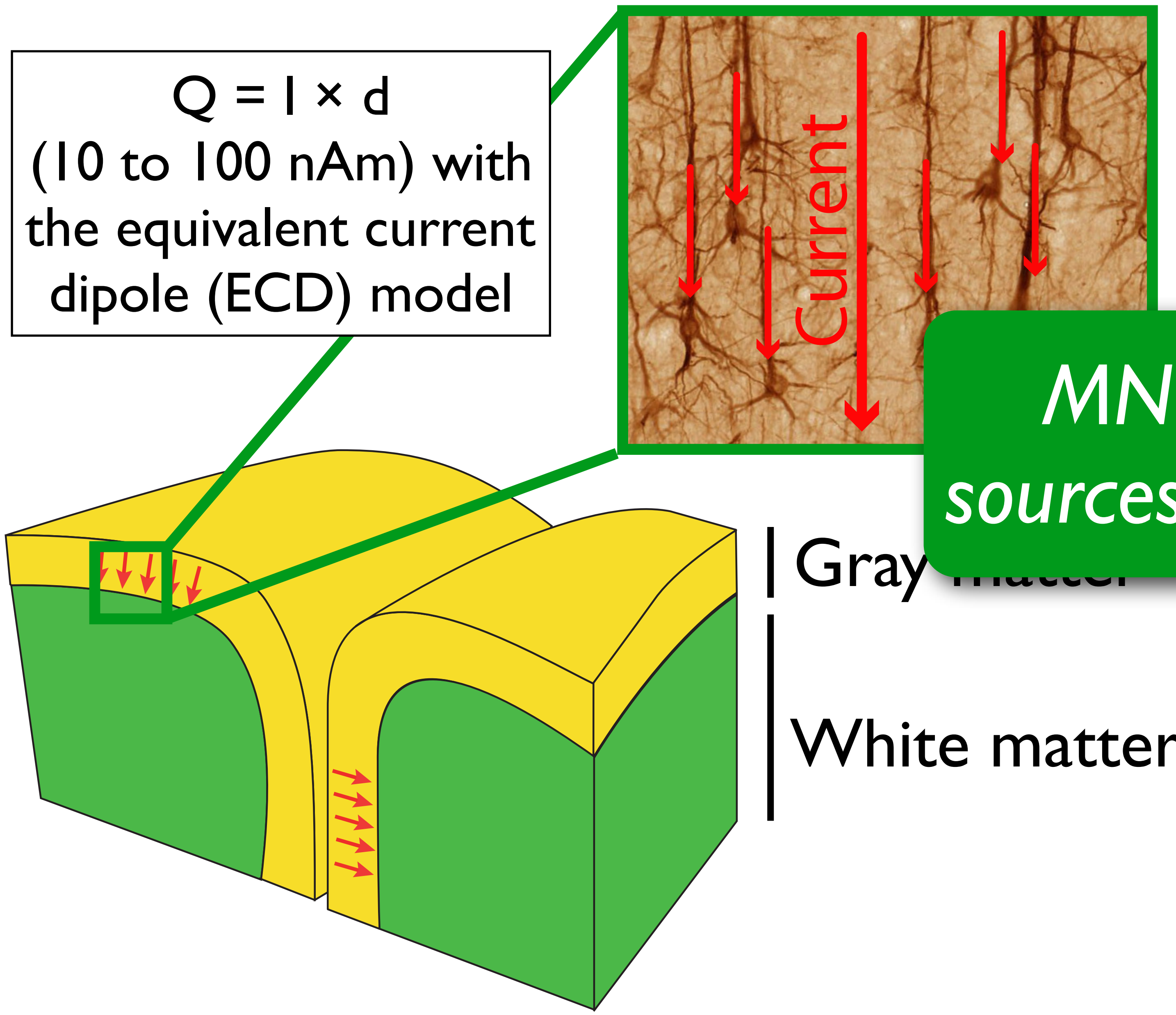
Generally, **all currents** generate a magnetic field!

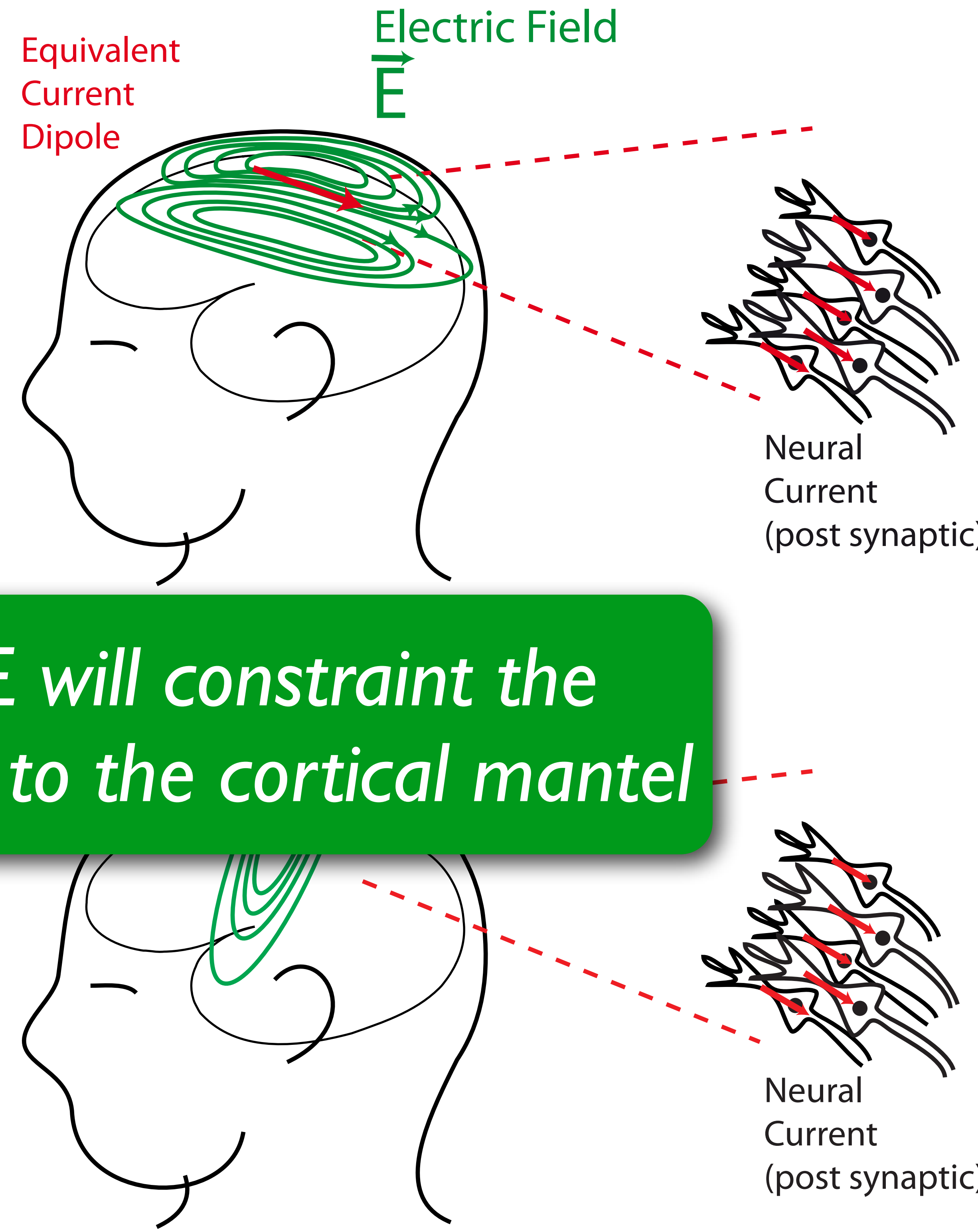
- Impressed currents $\mathbf{J}_i(\mathbf{r})$ can be omitted: short distance through the cell membrane
  ▶ negligible contribution to the total field

Neurons as current generators

Large cortical pyramidal cells organized in macro-assemblies with their **dendrites normally oriented to the local cortical surface**

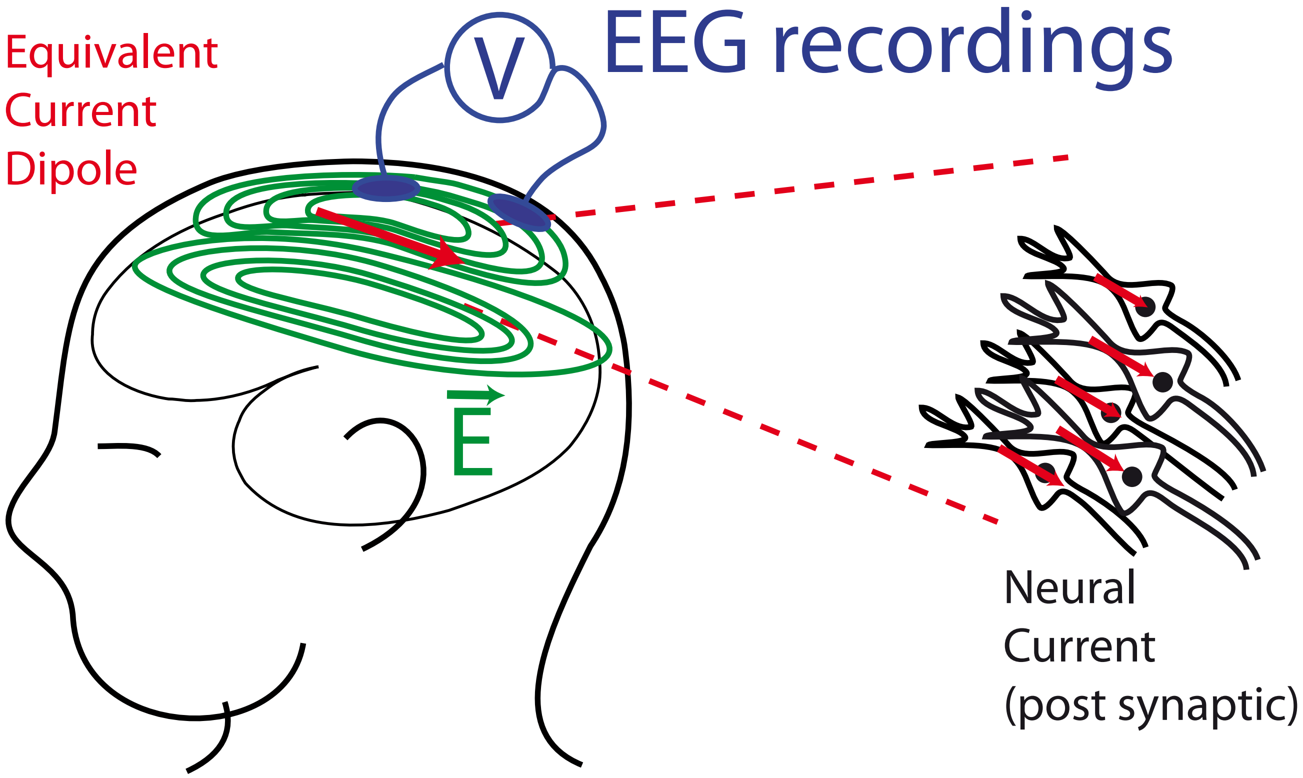
$Q = I \times d$
(10 to 100 nAm) with the equivalent current dipole (ECD) model

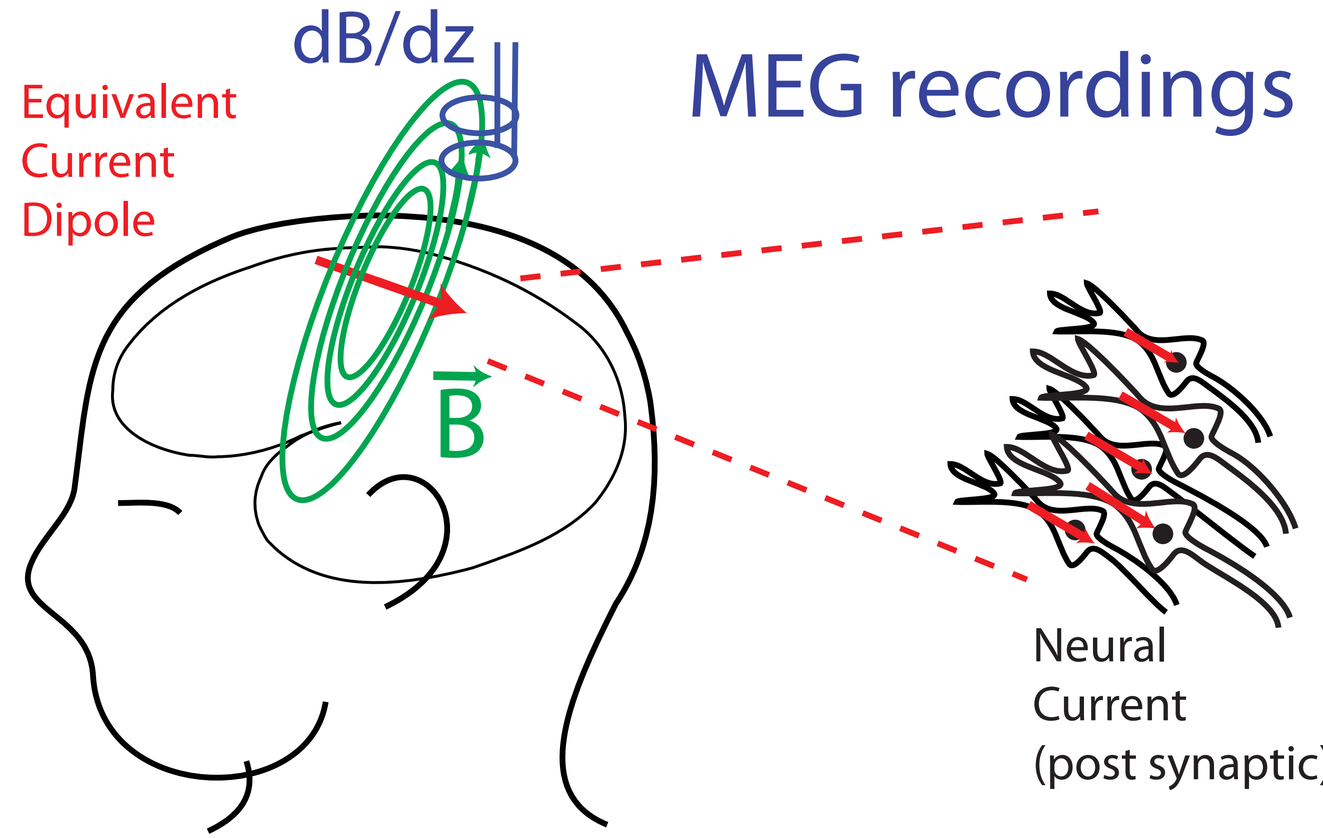




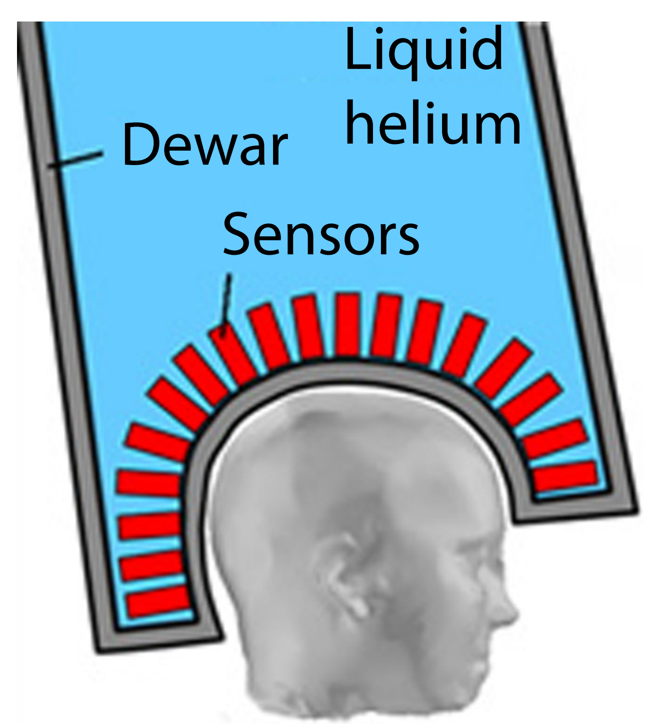
MNE will constraint the sources to the cortical mantel

EEG & MEG systems





First EEG recordings in 1929 by H. Berger



Hôpital La Timone
Marseille, France

MEG-lab

Measurement of electromagnetic activity of the brain:

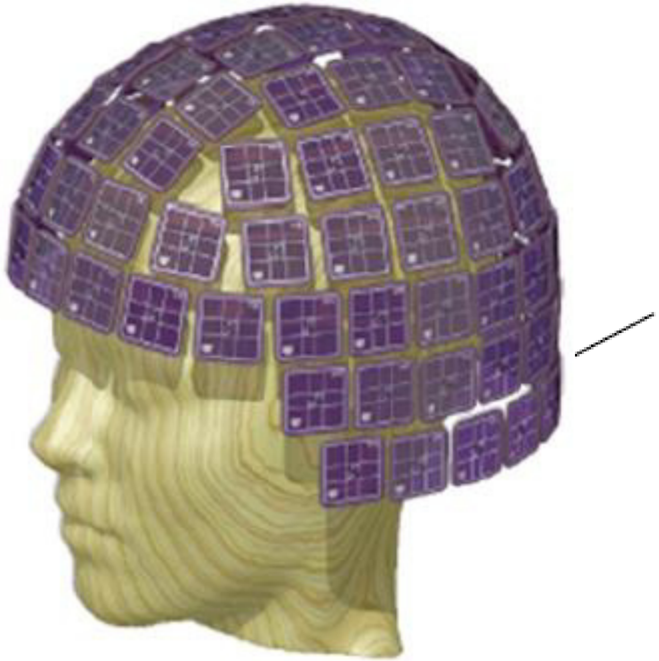
- **Magnetoencephalography (MEG)**
- Electroencephalography (EEG)

Magnetically shielded room with MEG

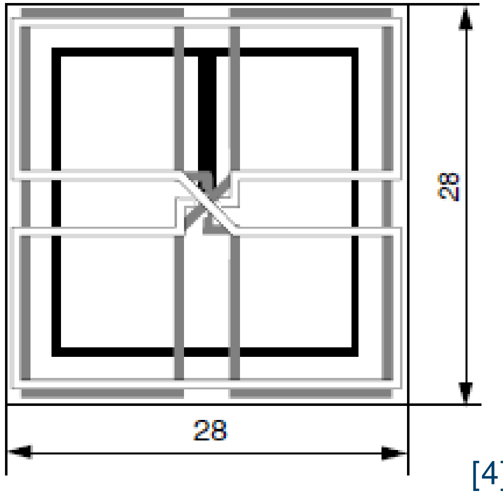
[1]

Room for preparation of subject + operation of MEG

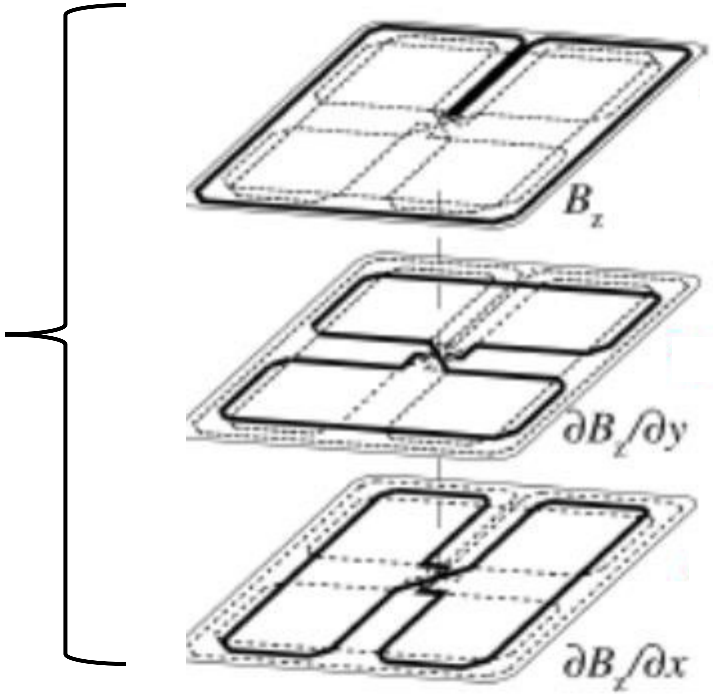
MEG sensors



[3]



[4]



[3]

Elekta Neuromag TRIUX MEG system

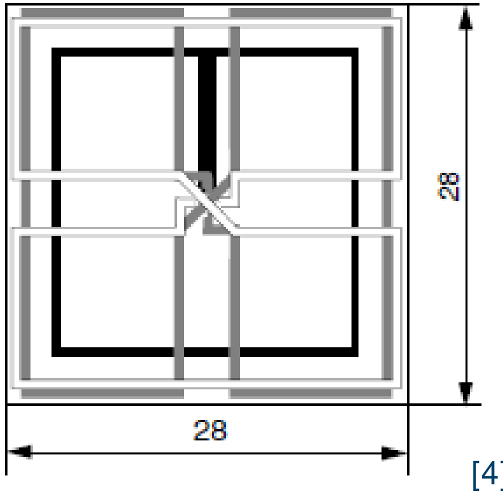
- 102 triple sensor detector units
  ► Σ 306 sensors

Per triple detector sensor unit

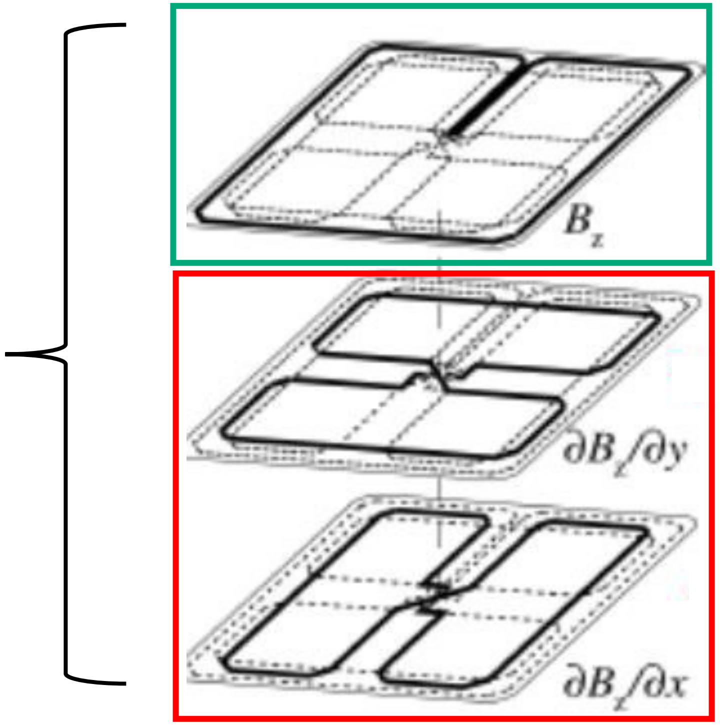
- 1x magnetometer, 2x planar gradiometers

MEG sensors

[3]



[4]



[3]

Elekta Neuromag TRIUX MEG system

- 102 triple sensor detector units
  ► Σ 306 sensors

Per triple detector sensor unit

- 1x magnetometer, 2x planar gradiometers

Magnetometer

- general magnetic fields
- very sensitive overall, noisy

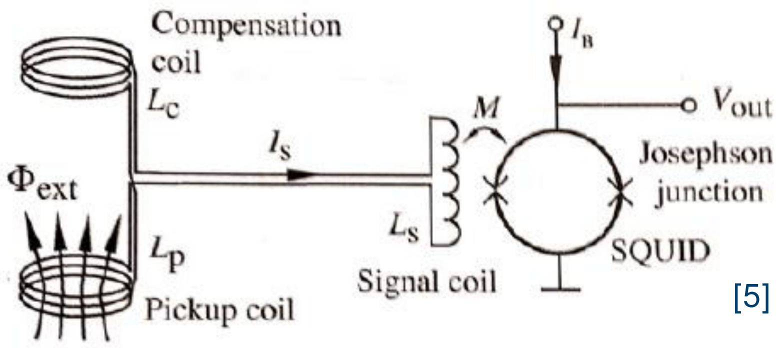
Planar gradiometer

- focal magnetic fields
- most sensitive to fields directly underneath

MEG sensors

MEG uses sophisticated sensing technology

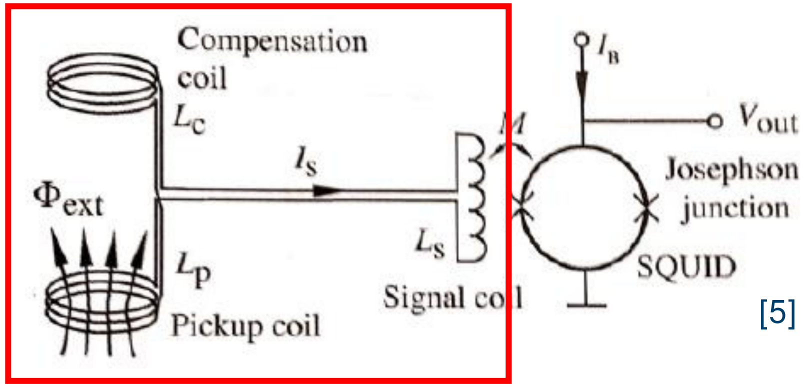
- SQUID (**S**uperconducting **Q**uantum **I**nterference **D**evice) sensor
- fully understanding of SQUIDs requires
    - quantum mechanical treatment
    - superconductivity



[5]

MEG sensors

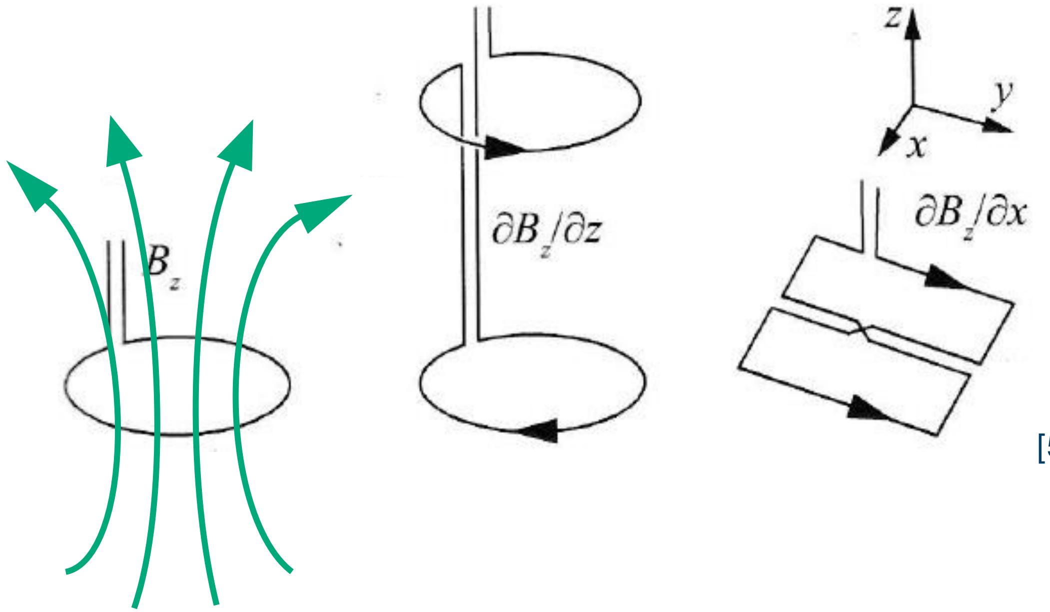
MEG uses sophisticated sensing technology
- SQUID (**S**uperconducting **Q**uantum **I**nterference **D**evice) sensor
- fully understanding of SQUIDs requires
  - quantum mechanical treatment
  - superconductivity



[5]

Different types of **flux transformers** are available which couple the signal to the SQUID

Magnetometer Axial gradiometer Planar gradiometer



[5]

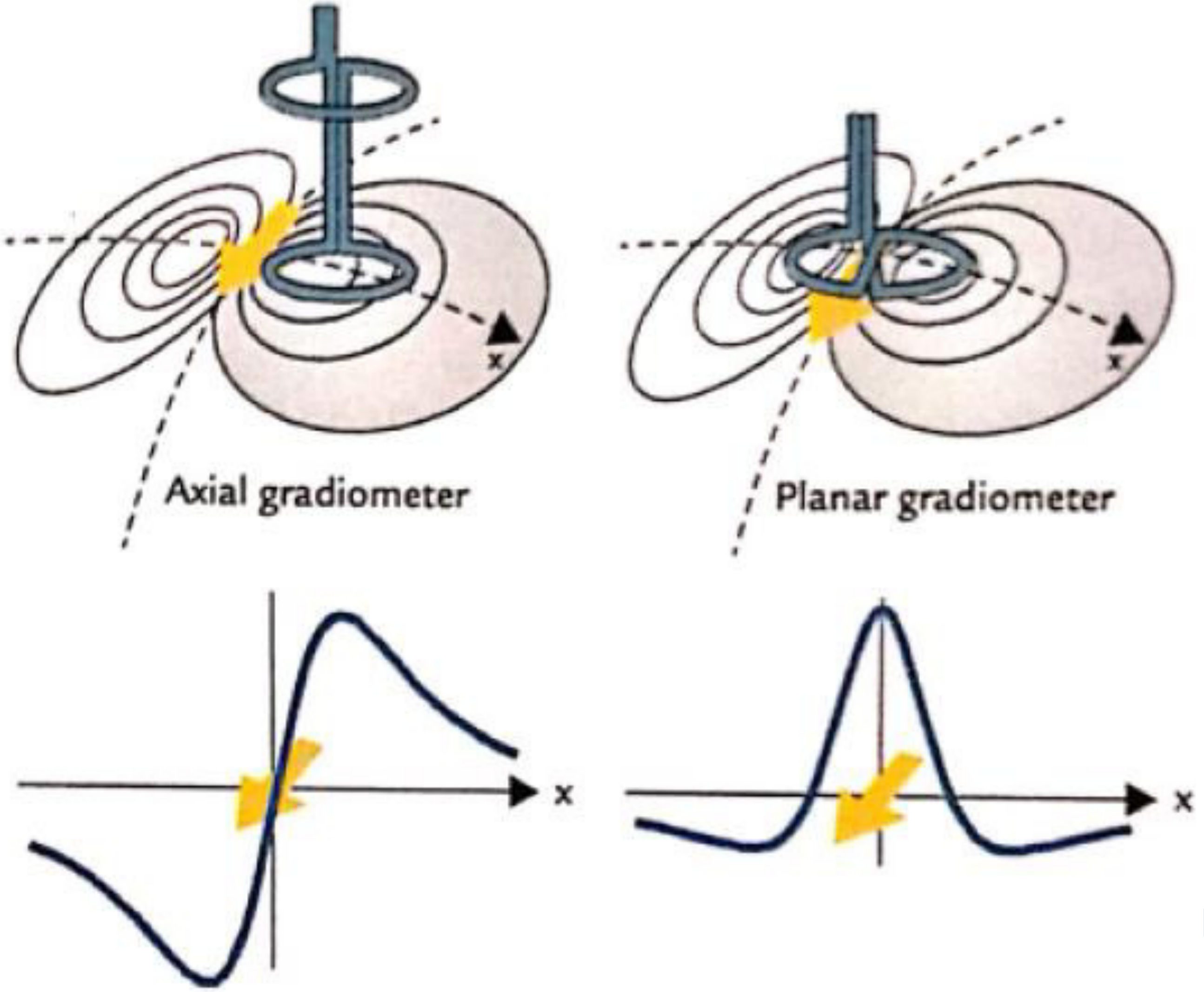
Basic principle

Magnetic flux across coil surface induces an electrical current in the coil wiring material, whose amplitude is instantaneously proportional to the magnetic field.

MEG sensors – sensitivity profile

Carl von Ossietzky Universität Oldenburg

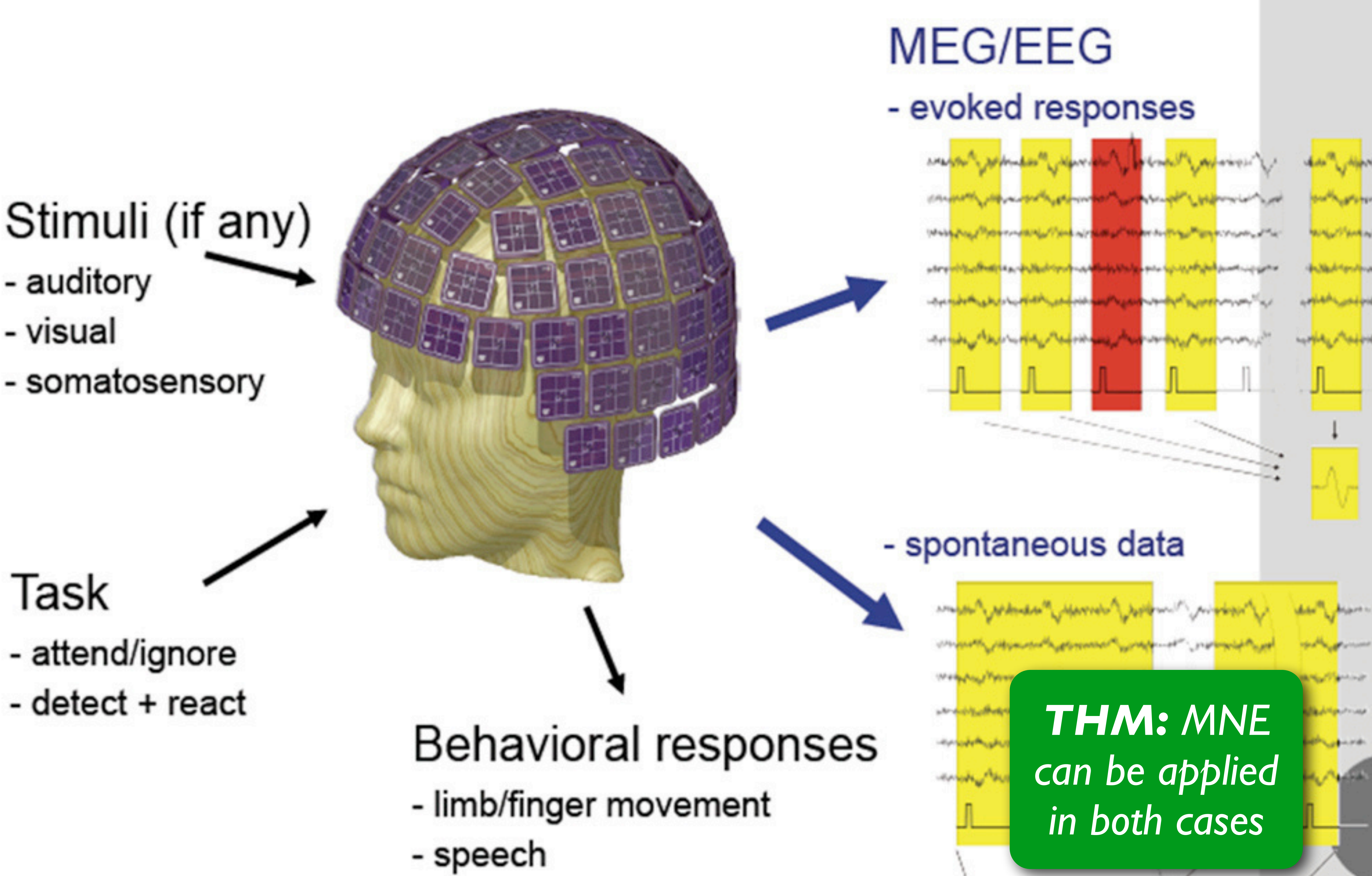
Influence of flux transformers on the measured **field distribution** and **signal strength**



[7]

- Gradiometers in general insensitiv to homogeneous fields
- **Axial gradiometer/ magnetometer** peaks for signals around the rim of the sensor
- **Planar gradiometer** gives maximum signals for sources right beneath them

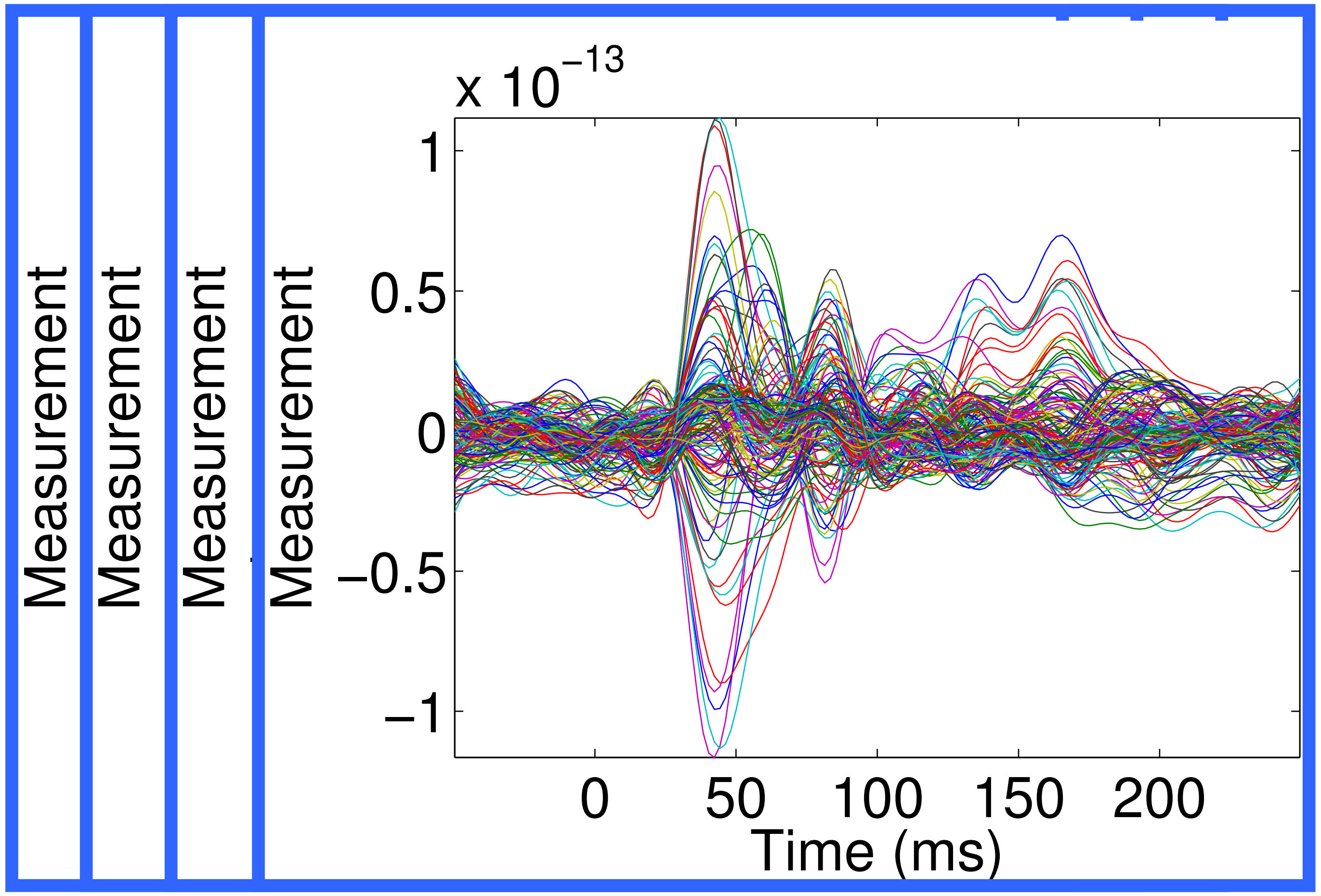
Data acquisition examples



***THM:** MNE can be applied in both cases*

Evoked response

Example of trial averaging:



Signal
on
151
MEG
Channels

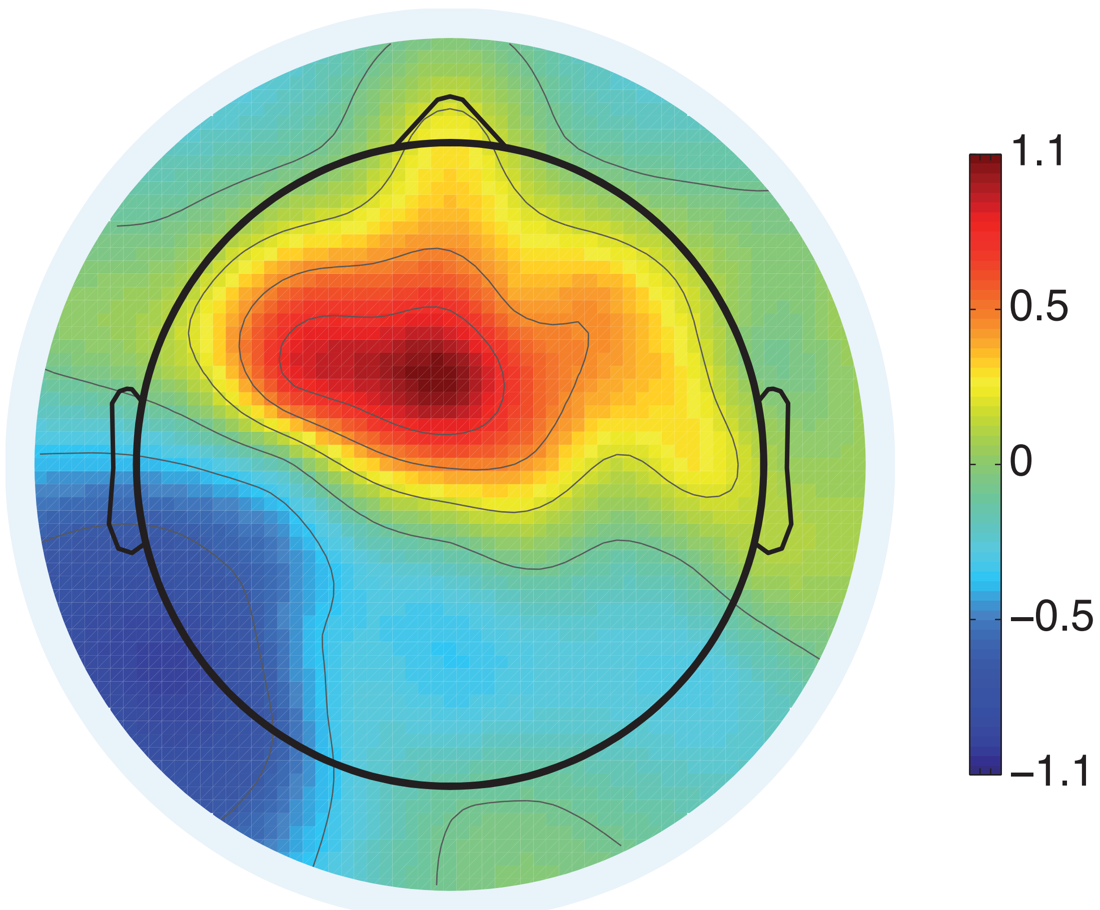
400 trials

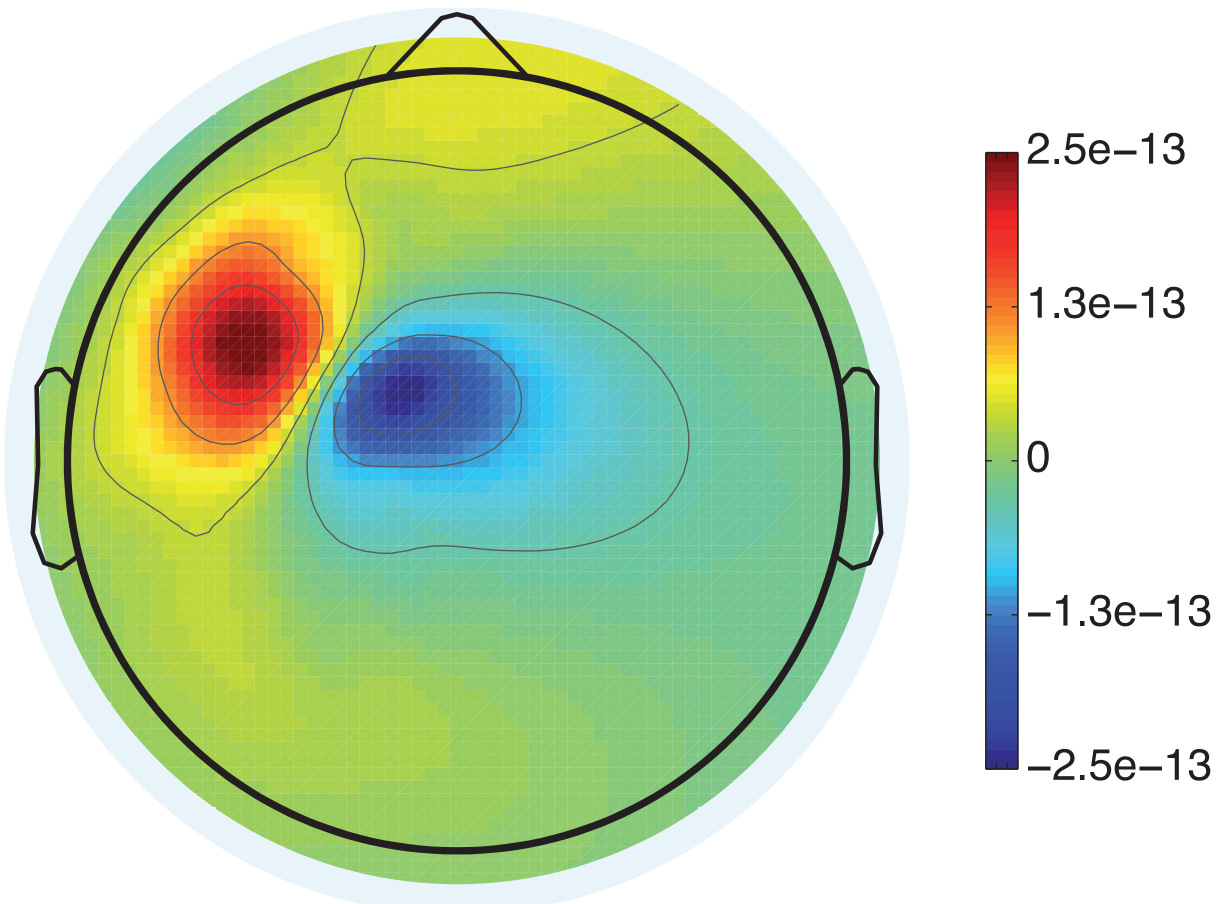
Clean M/EEG data can be obtained by **averaging multiple repetitions** (a.k.a. trials)

M/EEG Measurements

At **one time instant**:

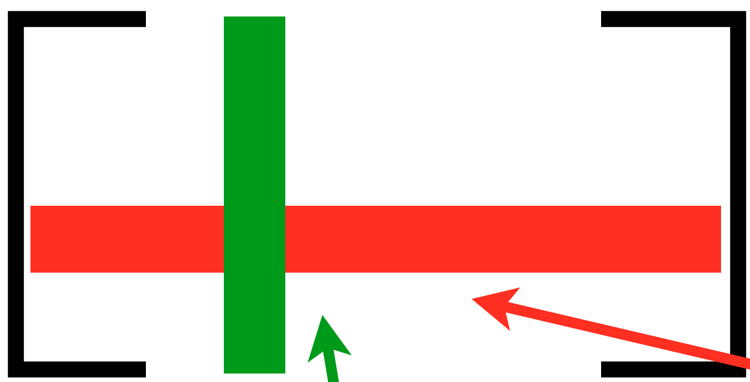
EEG topography vs. MEG topography





CTF system with 151 axial gradiometers

M/EEG Measurements: Notation



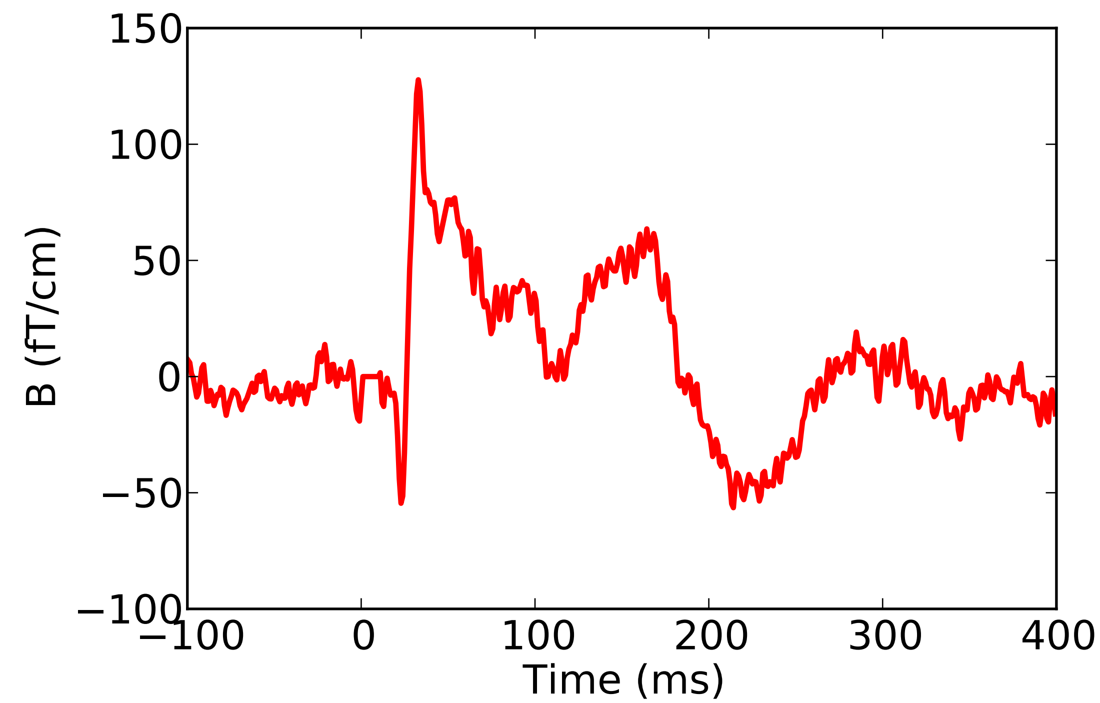
$$M = \left[\ \right] \in \mathbb{R}^{d_m \times d_t}$$

MEG and/or EEG

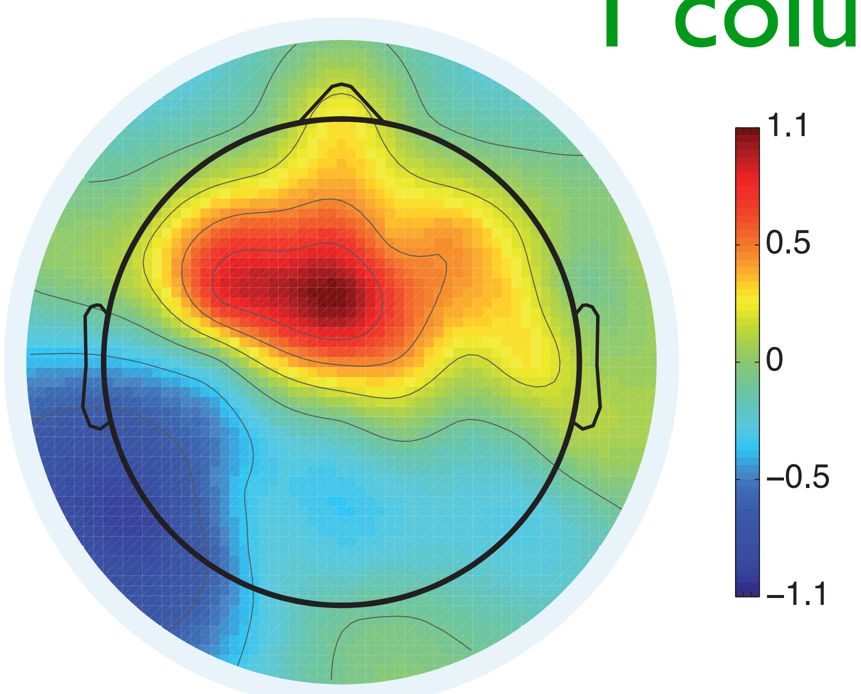
d_m : Nb of sensors

d_t : Nb of time points

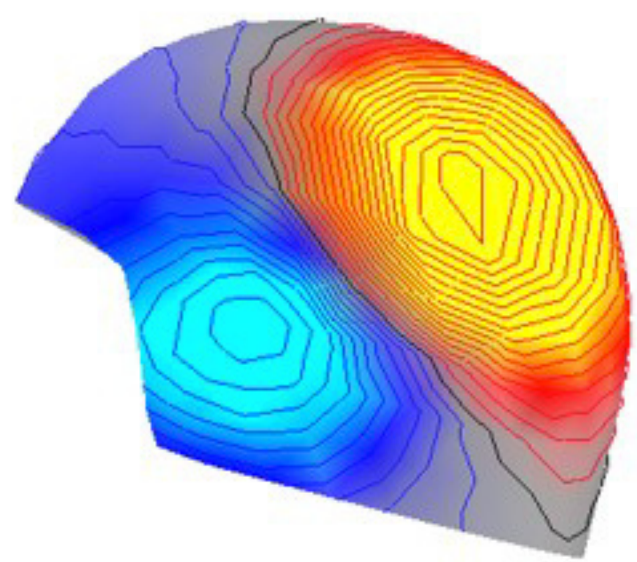
1 column



1 row = 1 time series on 1 sensor



2D topography

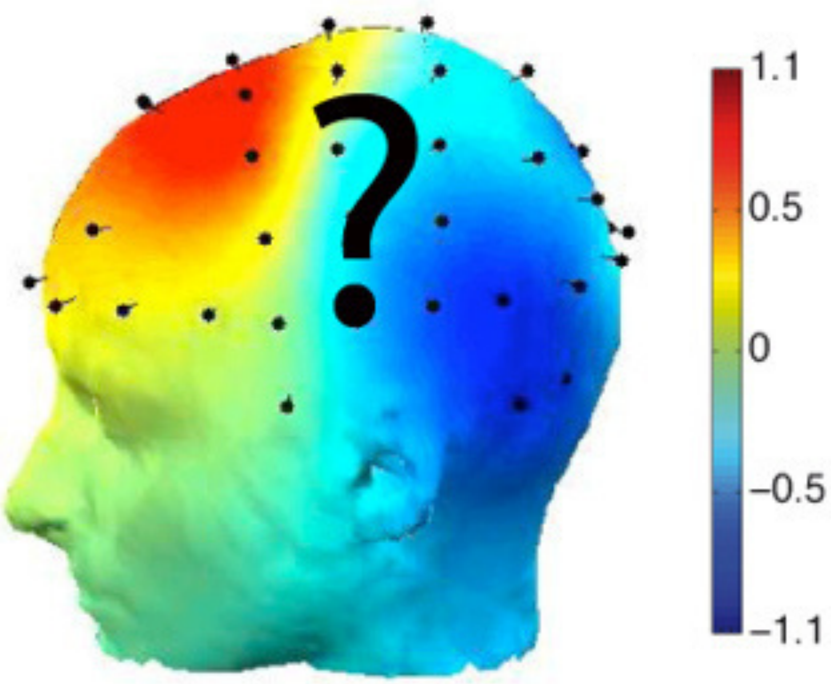


3D topography

The M/EEG inverse problem

Inverse problem: Objective

Find the current generators that produced the M/EEG measurements



Carl von Ossietzky Universität Oldenburg

Source modelling: overview

forward model

1

Physiological source
Electrical current

2

Body tissue
Volume conductor

3

Observed potential
or field

inverse model

[3]

Source modelling: overview

Solving *forward model* gives lead fields

1 Physiological source
Electrical current

2 Body tissue
Volume conductor

3 Observed potential
or field

inverse model

[3]

Linear forward problem: Maxwell

Maxwell Equations with **quasi-static** approximation

$$\begin{cases} \nabla \times \vec{E} = & 0 \\ \nabla \cdot \vec{B} = & 0 \\ \nabla \times \vec{B} = & \mu_0 \vec{J} \\ \nabla \cdot \vec{E} = & \frac{\rho}{\epsilon_0} \end{cases}$$

Remark: quasi-static implies no temporal derivatives and no propagation delay

Total currents: $\vec{J} = \vec{J}_p + \vec{J}_c$

Primary currents

Conduction currents

Ohm's law: $\vec{J}_c = -\sigma \nabla V$

V Electric potential

σ Tissue conductivity

Potential equation:

(relation btw. the potential and the sources)

$$\nabla \cdot \nabla \times \vec{B} = 0 \Rightarrow \nabla \cdot (\vec{J}_s + \vec{J}_c) = 0$$

$$\Rightarrow \nabla \cdot \vec{J}_p = \nabla \cdot (\sigma \nabla V)$$

THM: Instantaneous & Linear

Head models

Requires to **model the properties of the different tissues**: skin, skull, brain etc.

Hypothesis: The conductivities are **piecewise constant**

Sphere models

Analytical solutions fast to compute but very **coarse** head model (esp. for EEG)

EEG : [Berg et al. 94, De Munck 93, Zhang 95]
MEG : [Sarvas 87]

Realistic models

Boundary element method (BEM), i.e., numerical solver with approximate solution.

[Geselowitz 67, De Munck 92, Kybic et al. 2005, Gramfort et al. 2010]

2 Volume conduction model

- Describes electrical properties of tissue
- Describes geometrical model of the head
- Describes how the currents flow, not where they originate from

- Simplest volume conduction model is a **spherical volume conduction model**
- Spherical approximation of the geometry of the head tissue (e.g. brain, skull, scalp)
  - ➢ works reasonably well
- Has an analytical solution!

[6]

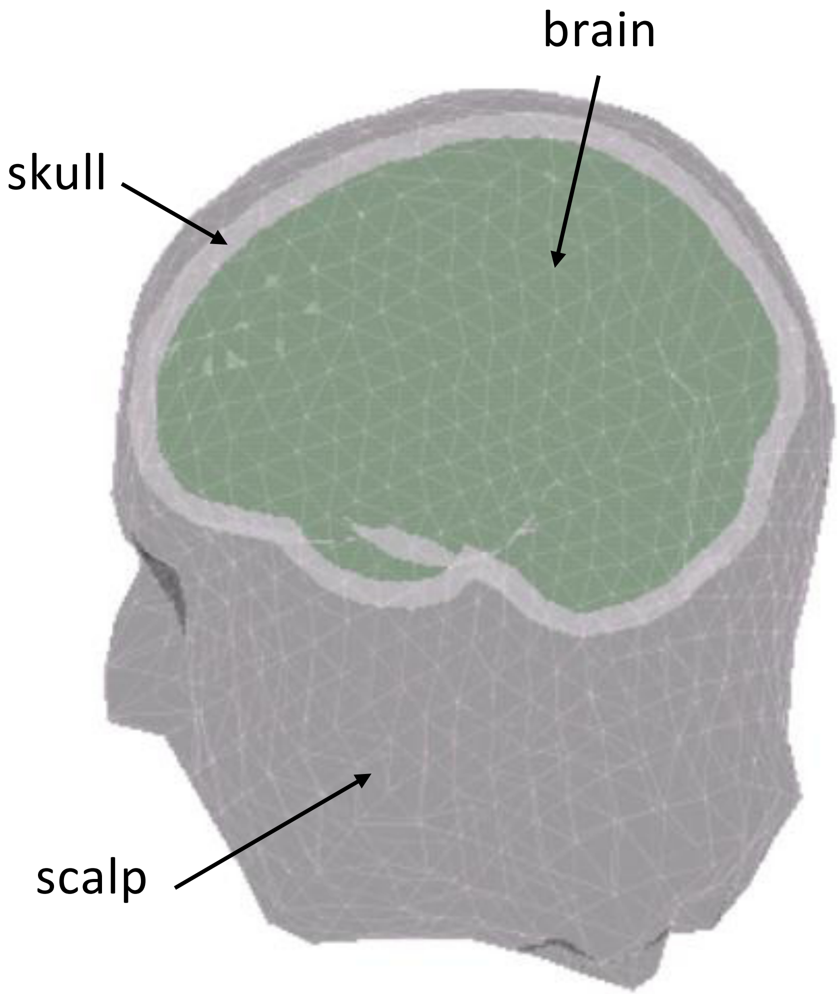
2 Volume conduction model

- Computational methods for volume conduction problem that allow for realistic geometries
  - Boundary Element Method (BEM)
  - Finite Element Method (FEM)
  - Finite Difference Method (FDM)

BEM: Geometrical description

- Triangulated surfaces describe boundaries between different tissues/compartments (e.g. brain, skull, scalp, (CSF))
- Conductivity of tissues is supposed to be homogeous and isotropic within each layer

➢ Derived by segmenting structural MRIs for different tissue types and then triangulating the resulting surfaces

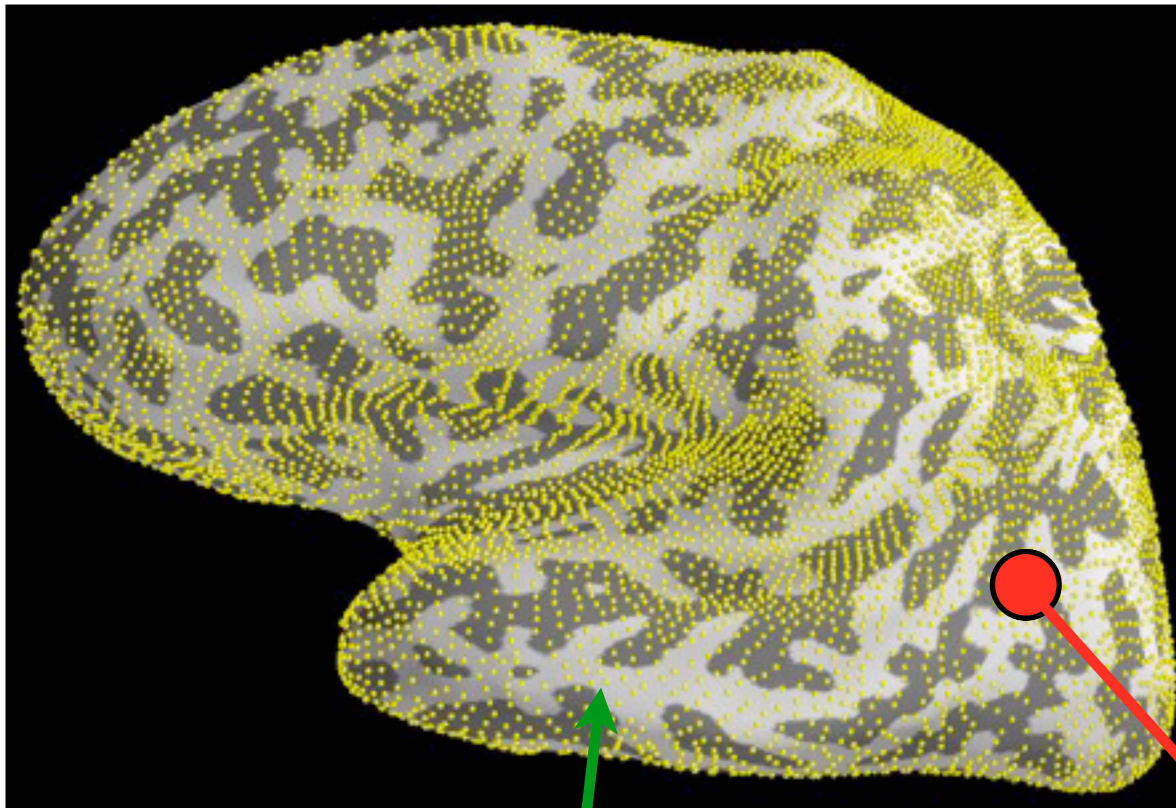


BEM volume conduction model

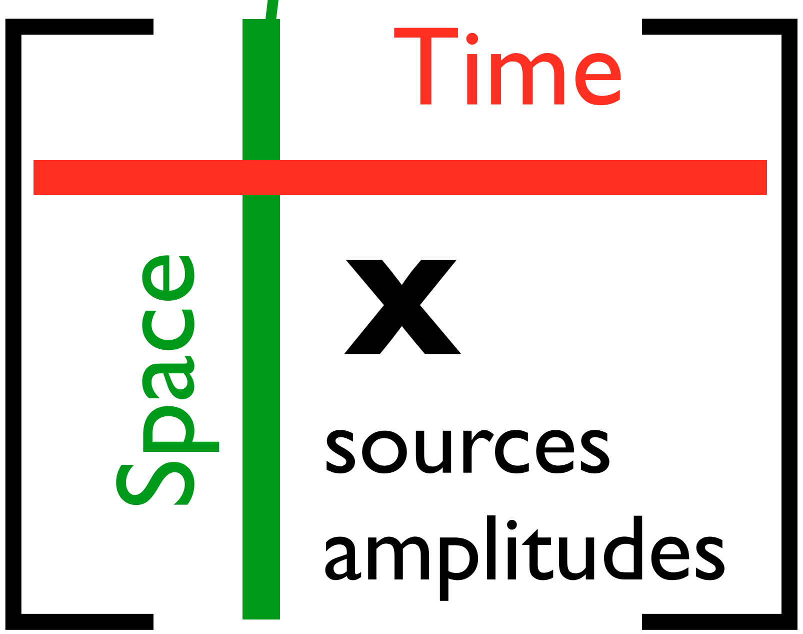
Inverse problem approaches

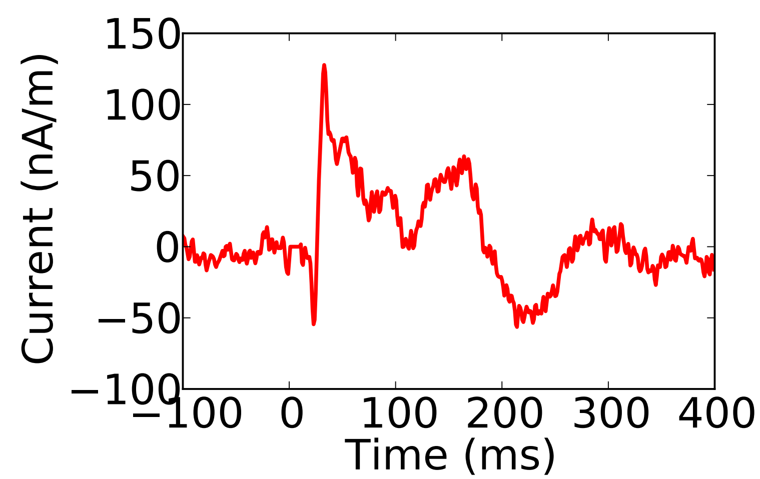
- Dipole fitting
- Scanning methods (Beamformers LCMV, DICS, SAM, MUSIC, RAP-MUSIC)
- Imaging methods with distributed models (MNE, dSPM, sLORETA, LORETA, MxNE, Gamma-Map/Champagne etc…)

Distributed model



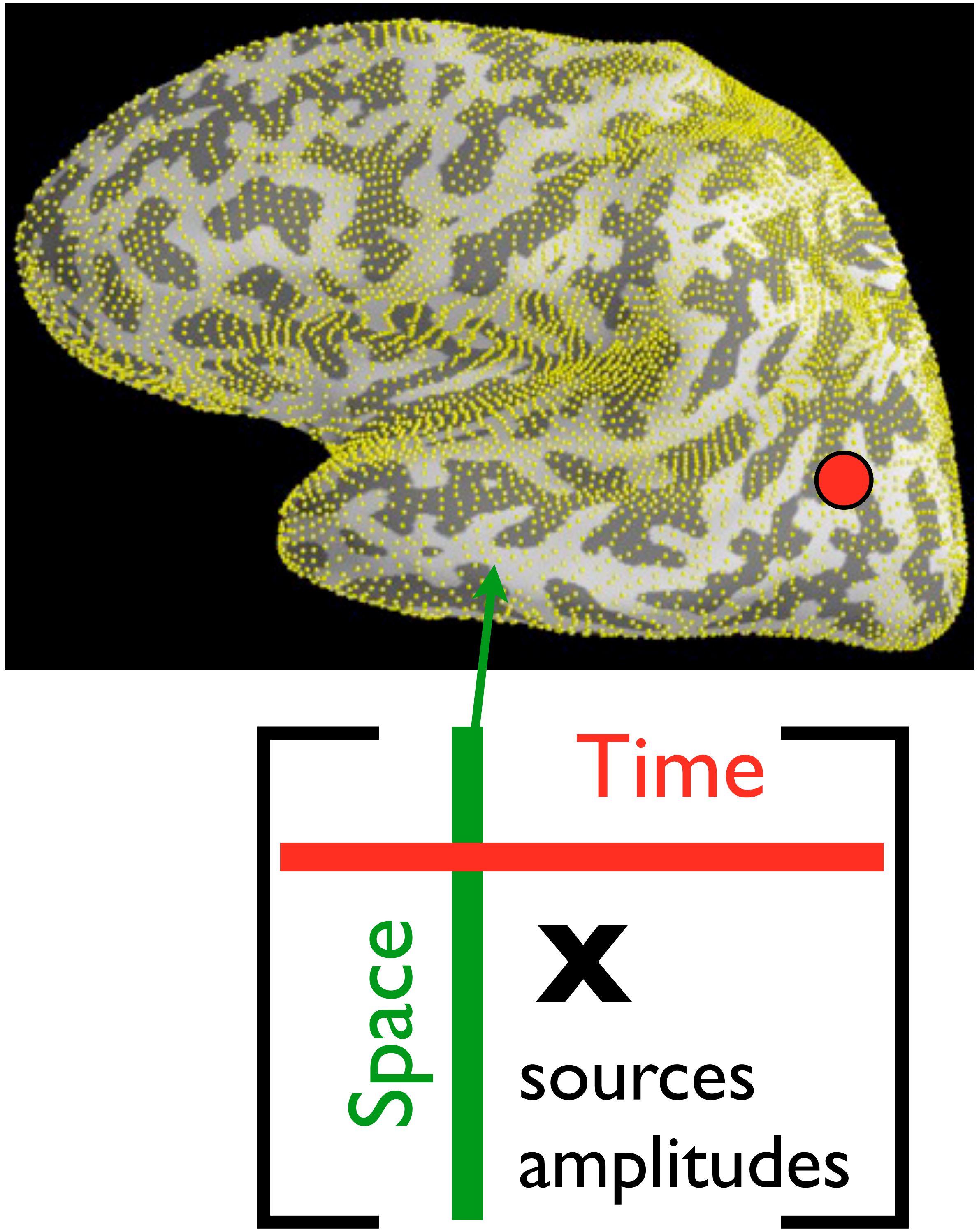
Position 5000 candidate sources over each hemisphere (e.g. every 5mm)



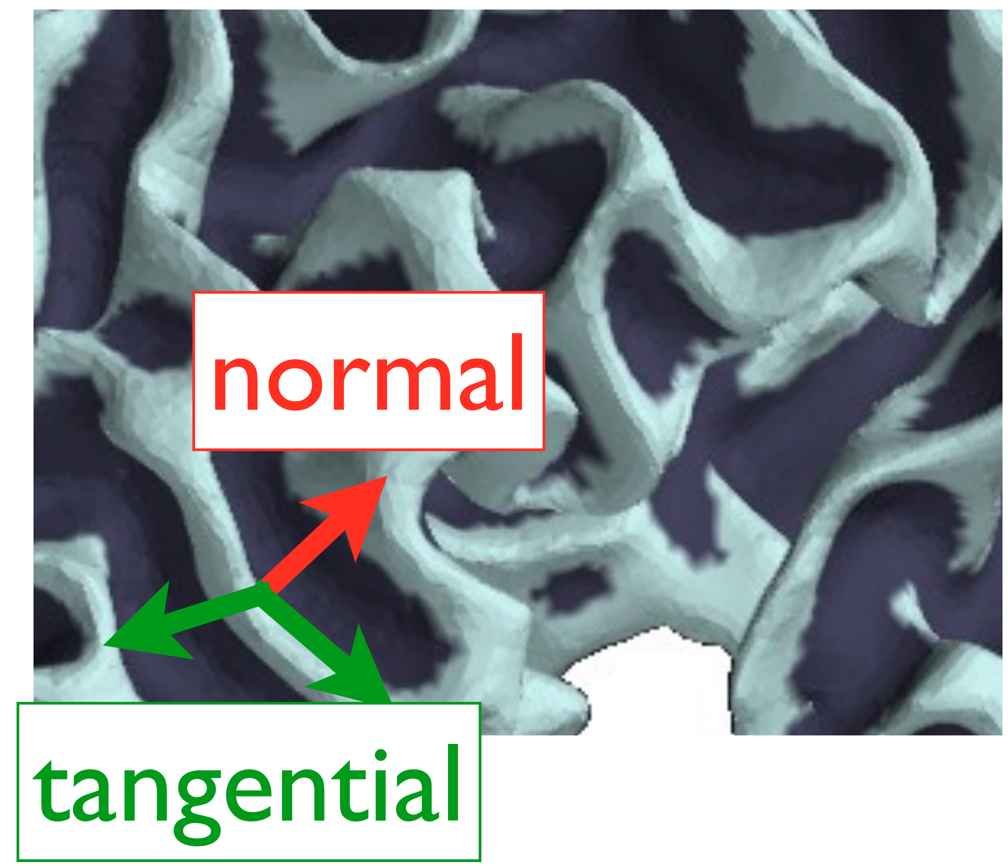


Scalar field defined over time

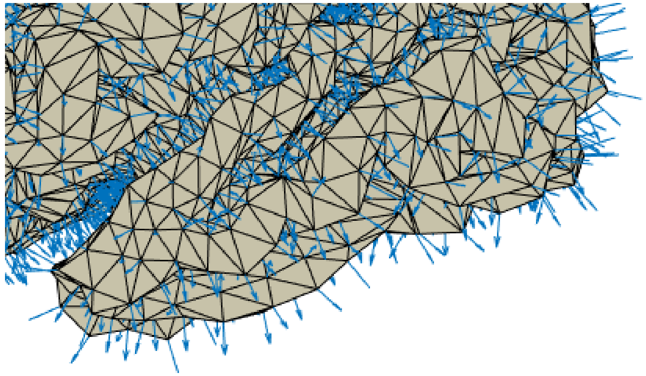
Distributed model



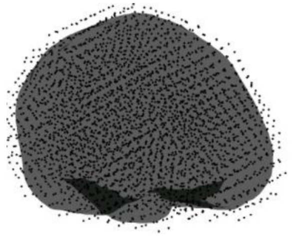
Or vector field (3 values per location)

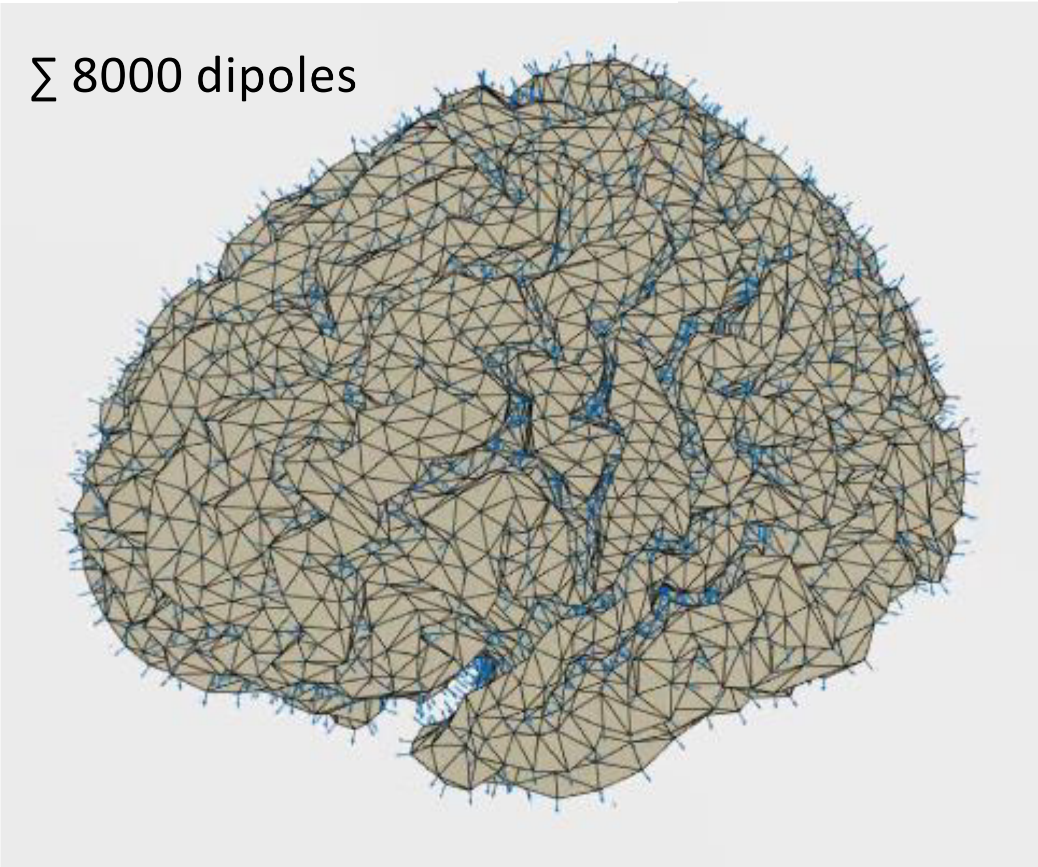


1 Source model



Surface-based source model based on a surface description of the cortical sheet
(volumetric source models are also possible)





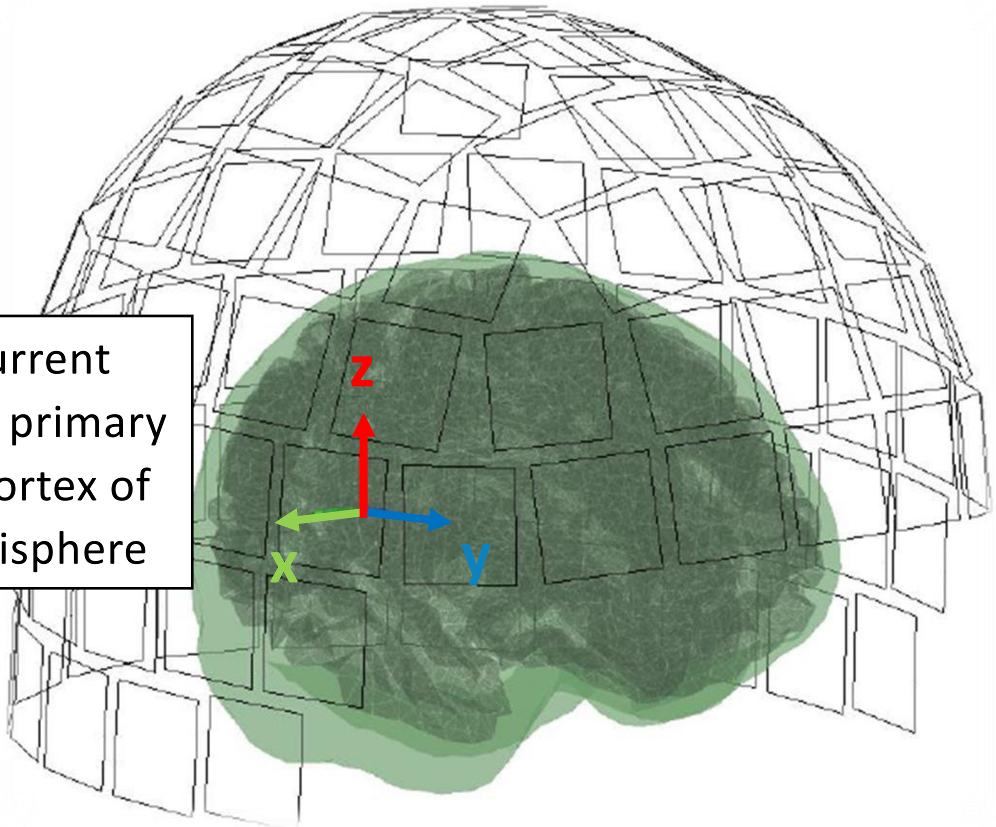
- Triangulated cortical mesh, ideally consisting of a number of triangles that form a topological sphere for each of the cerebral hemispheres
- Meshes have typically > 100000 vertices per hemisphere, but can be downsampled (~4000 vertices per hemisphere)
- Each vertex describes the location of a current dipole which dipole moment is to be estimated (blue vectors in figure)

Solution of forward model: Lead fields

- **lead fields** represent by definition the **field distribution** across all channels due to a **unitary current dipole** placed with a given **position** and **orientation**.

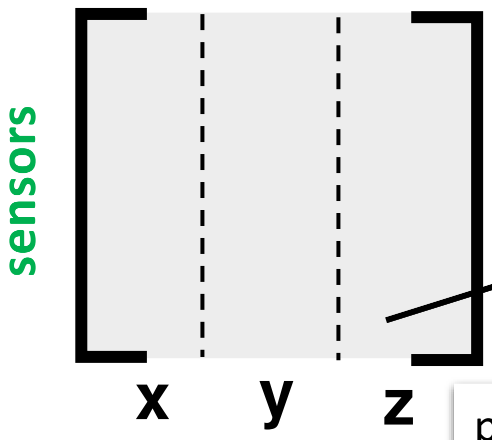
  (dipole moment $\boldsymbol{Q}$ is unit vector $|\mathbf{Q}| = 1$)

Unitary current dipoles in primary audtory cortex of right hemisphere

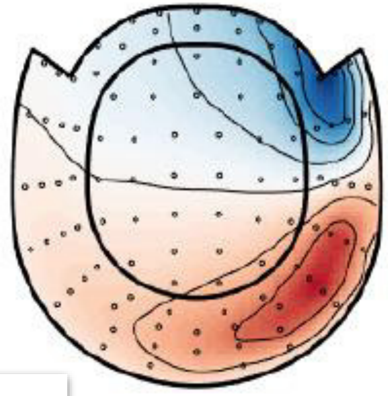


Lead field matrix for each dipole position

Lead field for magnetometers due to unitary dipole in z-direction

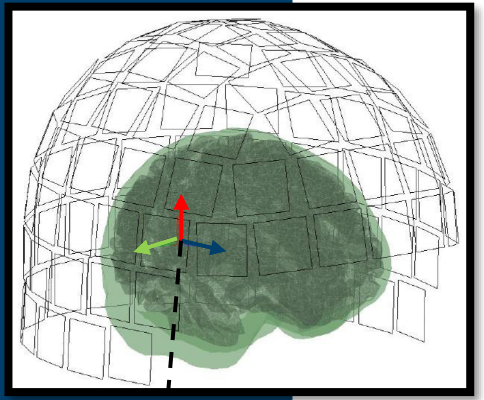


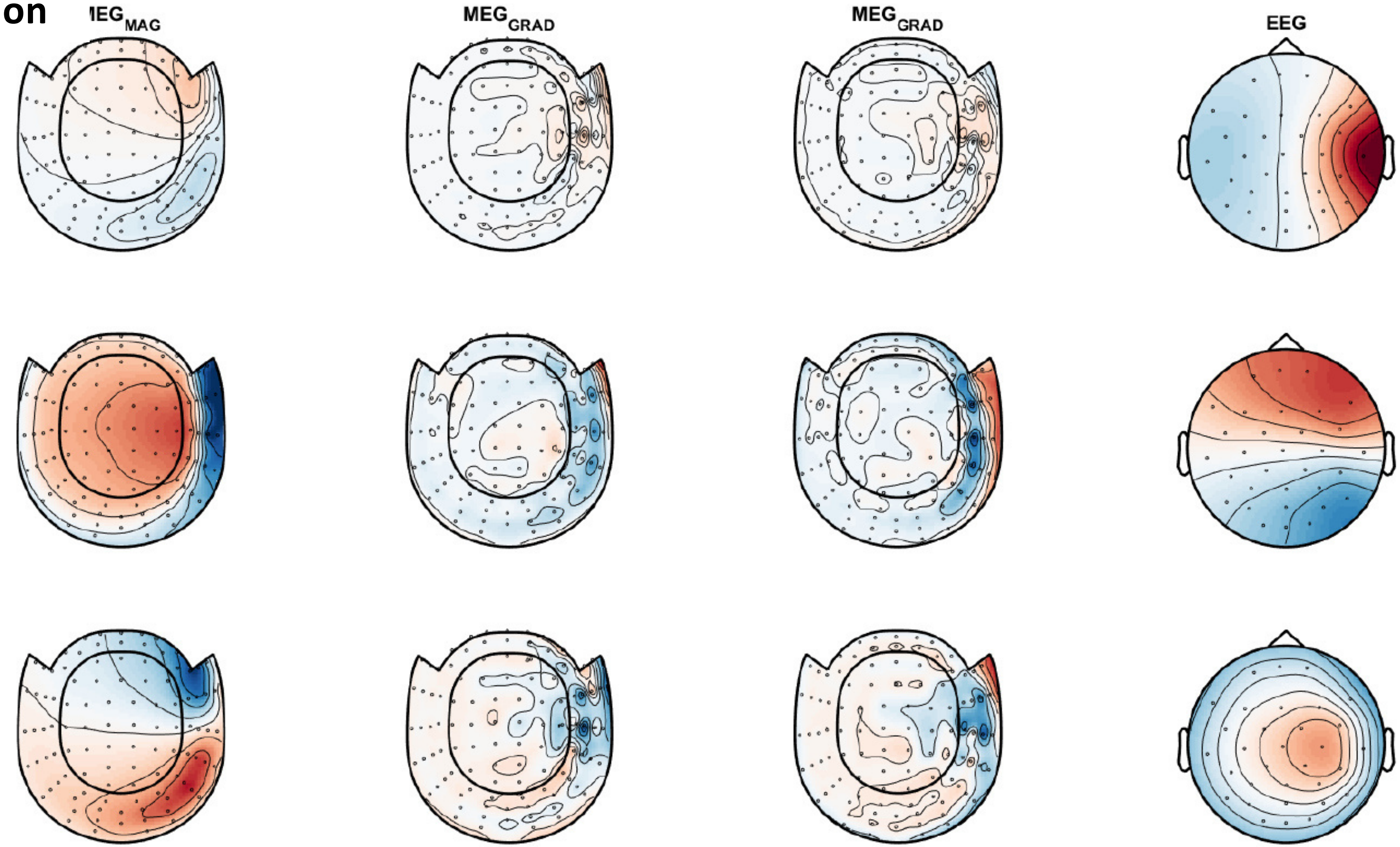
plot z-column



Lead fields

Dipole orientation

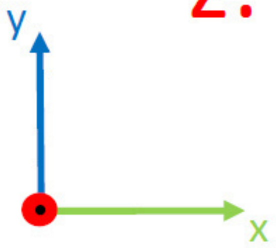




x:

y:

z:

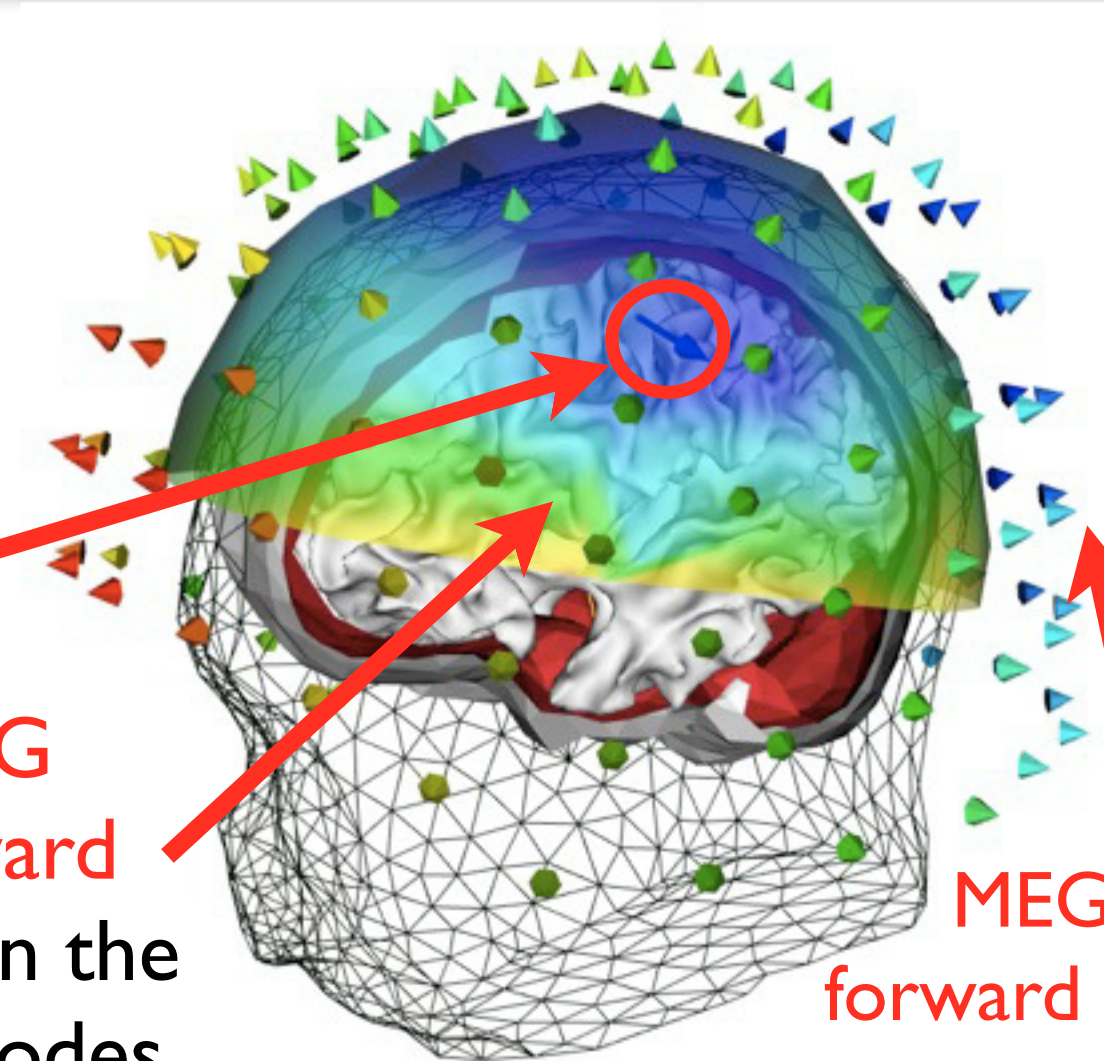


Distributed models

Dipoles sampled over the cortical surface extracted by MRI segmentation

[Dale and Sereno 93]

Current generator modeled as a **current dipole** (location, orientation and amplitude)



EEG forward field on the electrodes

MEG forward field on sensors

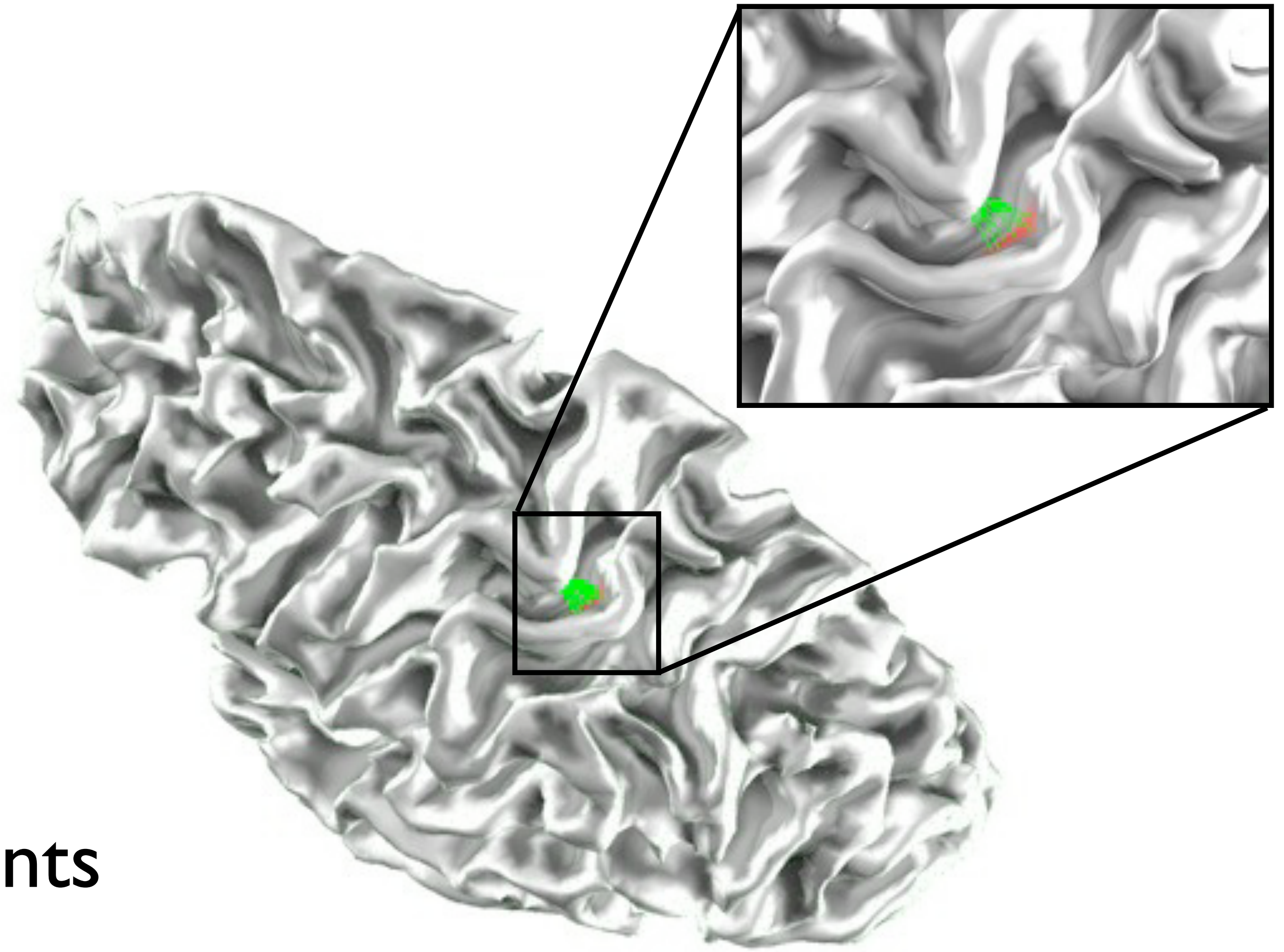
G =

is the lead field matrix obtained by concatenation of the forward fields

one column = Forward field of one dipole

Distributed source framework

$$\mathbf{M} = \mathbf{G}\mathbf{X} + \mathbf{E}$$



Linear forward model, i.e.,
M is the sum of the contributions of all the sources
(Superposition principle)

$\mathbf{M} \in \mathbb{R}^{d_m \times d_t}$: M/EEG Measurements

$\mathbf{X} \in \mathbb{R}^{d_x \times d_t}$: Source amplitudes (Unknowns)

$\mathbf{G} \in \mathbb{R}^{d_m \times d_x}$: Leadfield (or Gain) matrix

$\mathbf{E} \in \mathbb{R}^{d_m \times d_t}$: additive noise

Characteristic Features of Inverse Problems

Hadamard‘s definition of *well-posed* problems:

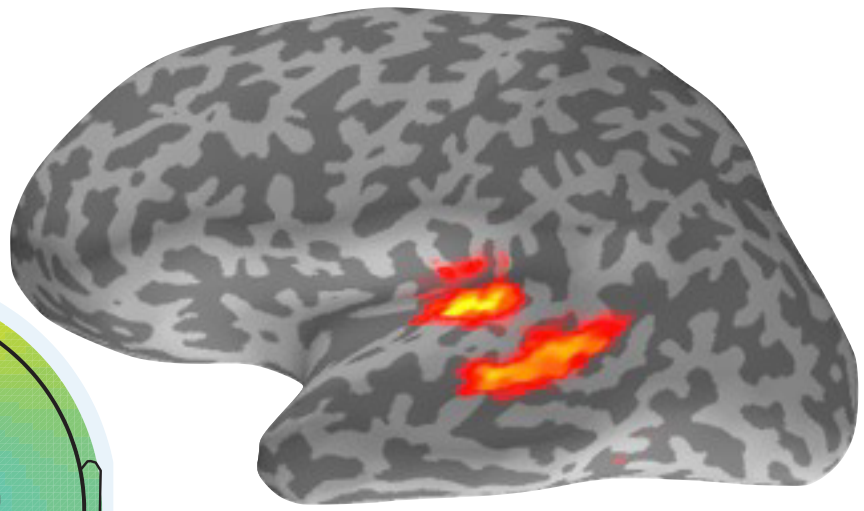
1. A solution exists.
2. The solution is unique.
3. The solution depends continuously on the data.

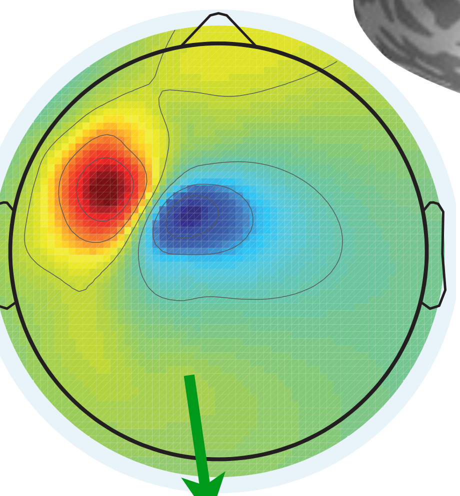
If one of the conditions does not hold, the problem is called ill-posed.

Inverse problems are typically ill-posed.

Jacques Salomon Hadamard
(1865-1963)

M = GX+E : An ill-posed problem

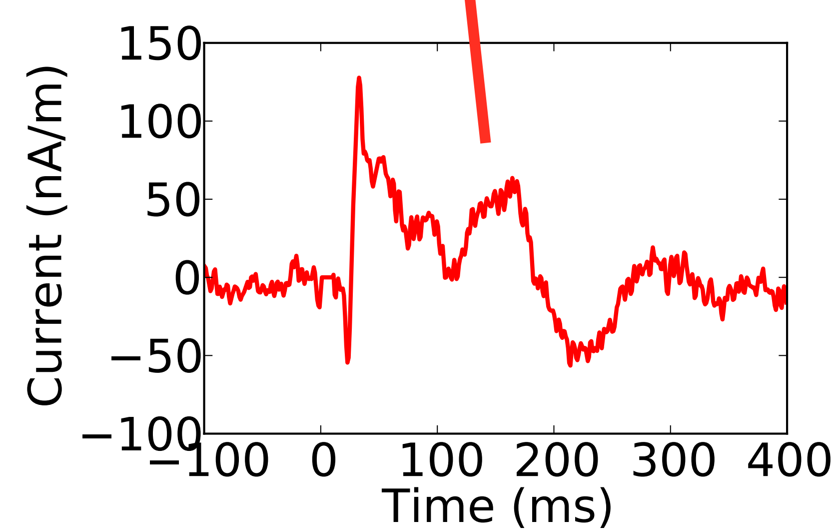




$$\mathbf{M} = \mathbf{G}\,\mathbf{X} + \mathbf{E}$$

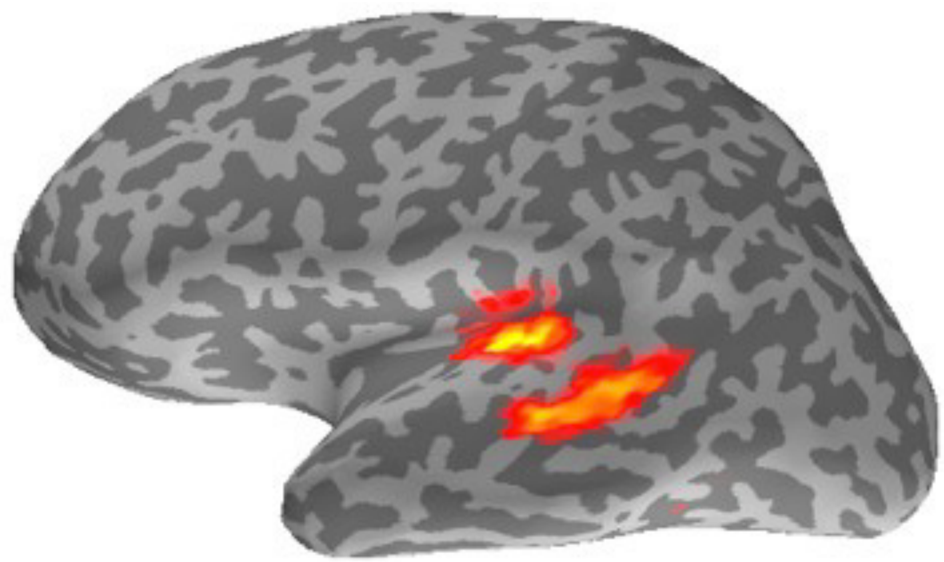
$d_x \approx 10000$

$d_m \approx 100$



Linear problem with more unknowns than the number of equations: it's ill-posed => Use prior

y = Xw+E : An ill-posed problem



Standard statistics notations

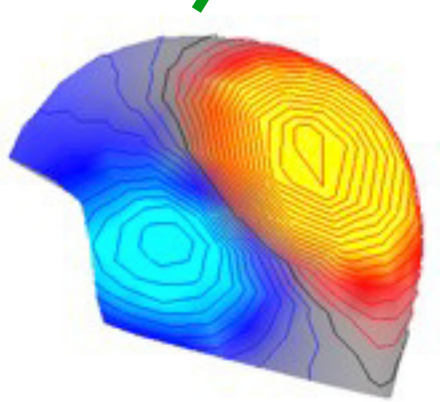
w (or β)
regression coefficients

$d_x \approx 10000$

$d_m \approx 100$

$$\mathbf{y} = \mathbf{X}\,\mathbf{w} + \mathbf{E}$$

X design matrix



THM: At **each time instant** the M/EEG inverse problem **IS** a **regression** with more variables than observations

Inverse problem framework

Penalized (variational) formulation (with whitened data):

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F^2 + \lambda\phi(\mathbf{X}), \lambda > 0$$

Data fit Prior

λ : Trade-off between the data fit and the prior

where $\|\mathbf{A}\|_F^2 = \mathbf{tr}(\mathbf{A}^T\mathbf{A})$

$\phi(\mathbf{X})$ is the prior.

Examples for $\phi(\mathbf{X})$: $\ell_1, \ \ell_2,$ Total-Variation ...

THM: when SNR goes UP λ goes DOWN.

Remark: will only work if all data are on the same scale

Variational/Tikhonov Regularization, Examples

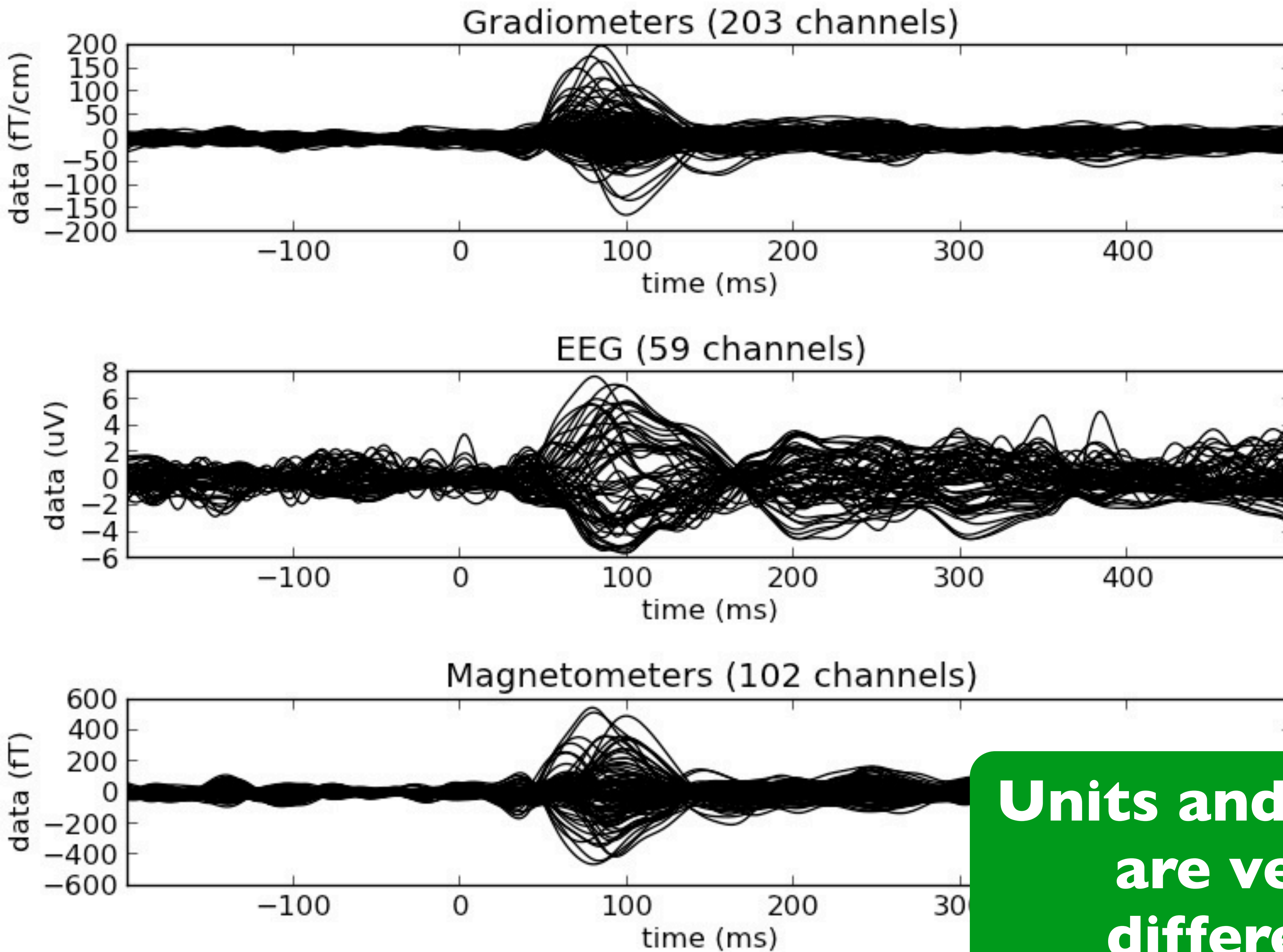
Classical minimum norm estimate:
$\mathcal{P}(s) = \|s\|_2^2 \quad \longrightarrow \quad s_\lambda = \operatorname{argmin} \ \|b - Ls\|_2^2 + \lambda \|s\|_2^2$

Weighted minimum norm estimate:
$\mathcal{P}(s) = \|Ws\|_2^2 \quad \longrightarrow \quad s_\lambda = \operatorname{argmin} \ \|b - Ls\|_2^2 + \lambda \|Ws\|_2^2$
Possible weightings: Depth weighting, spatial smoothness, fMRI/SPEC/PET.

Minimum current estimate. Let s_i^* denote the amplitude of the current vector at position i:
$\mathcal{P}(s) = \|s^*\|_1 \quad \longrightarrow \quad s_\lambda = \operatorname{argmin} \ \|b - Ls\|_2^2 + \lambda \|s^*\|_1$

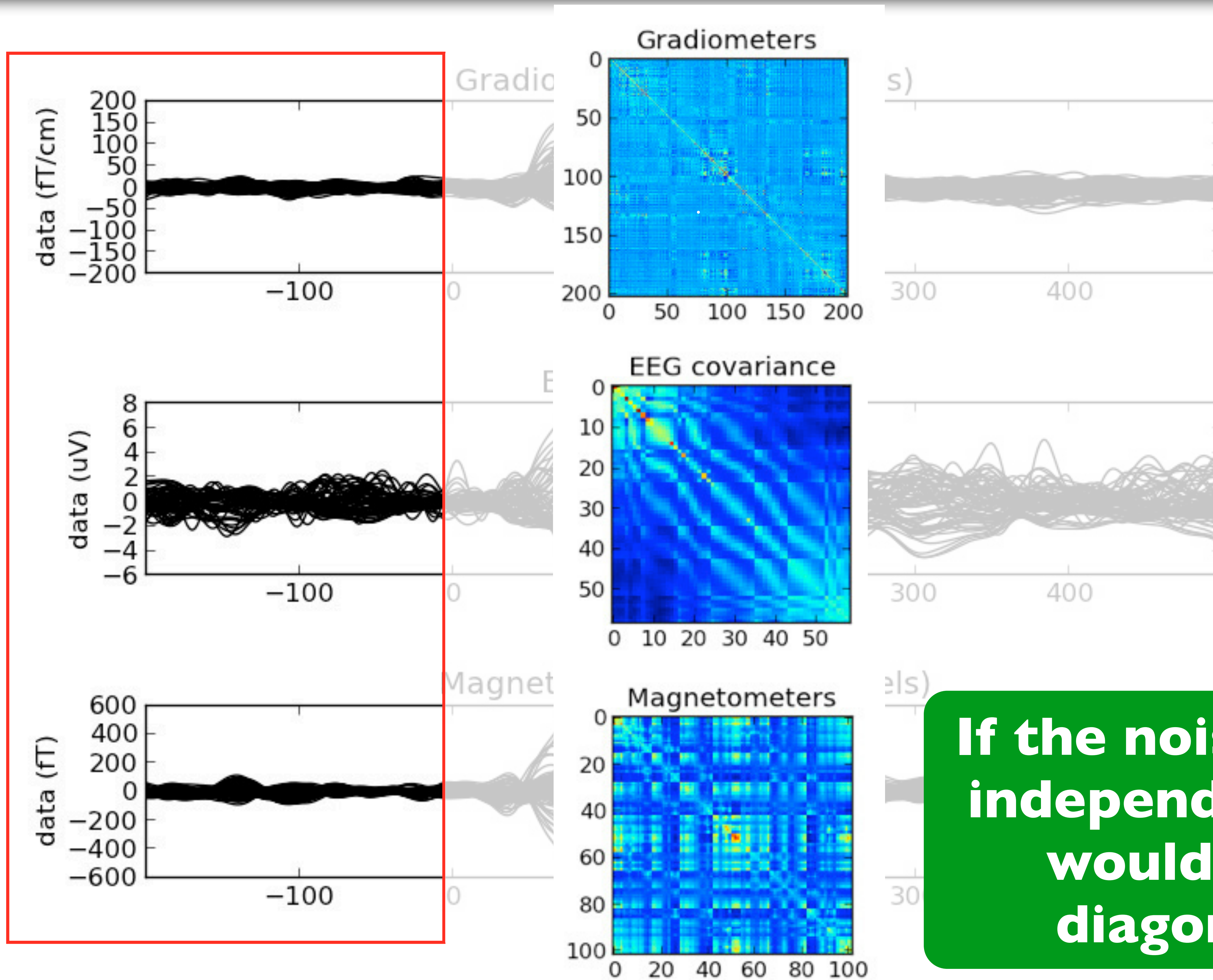
Weighted minimum current estimates:
$\mathcal{P}(s) = \|Ws^*\|_1 \quad \longrightarrow \quad s_\lambda = \operatorname{argmin} \ \|b - Ls\|_2^2 + \lambda \|Ws^*\|_1$
Possible weightings: Total Variation (TV, $W = \nabla$), depth weighting, fMRI/SPEC/PET

Spatial whitening



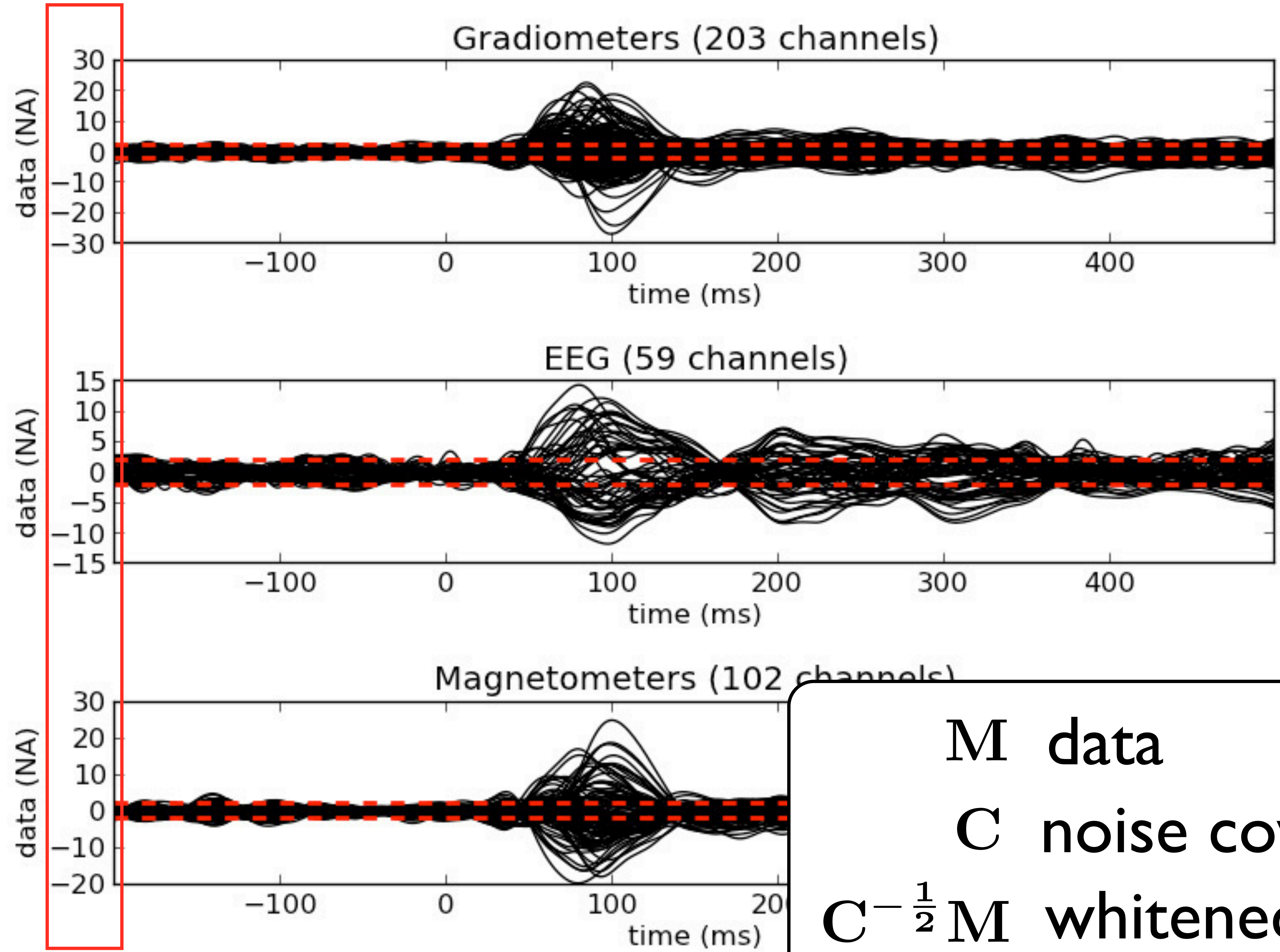
Units and scales are very different

Spatial whitening



If the noise was independent it would be diagonal

Spatial whitening



$\mathbf{M}$ data

$\mathbf{C}$ noise cov.

$\mathbf{C}^{-\frac{1}{2}}\mathbf{M}$ whitened data

L2 a.k.a. Minimum Norm Estimates (MNE)

$$\phi(\mathbf{X}) = \|\mathbf{W}\mathbf{X}\|_F^2 = \sum_{i,j} w_i^2 x_{ij}^2 = \|\mathbf{X}\|_{\mathbf{\Sigma},2}^2$$

$$\mathbf{W}^2 = \mathbf{\Sigma}$$ *source covariance*

Leads to a **closed form solution** (matrix multiplication):

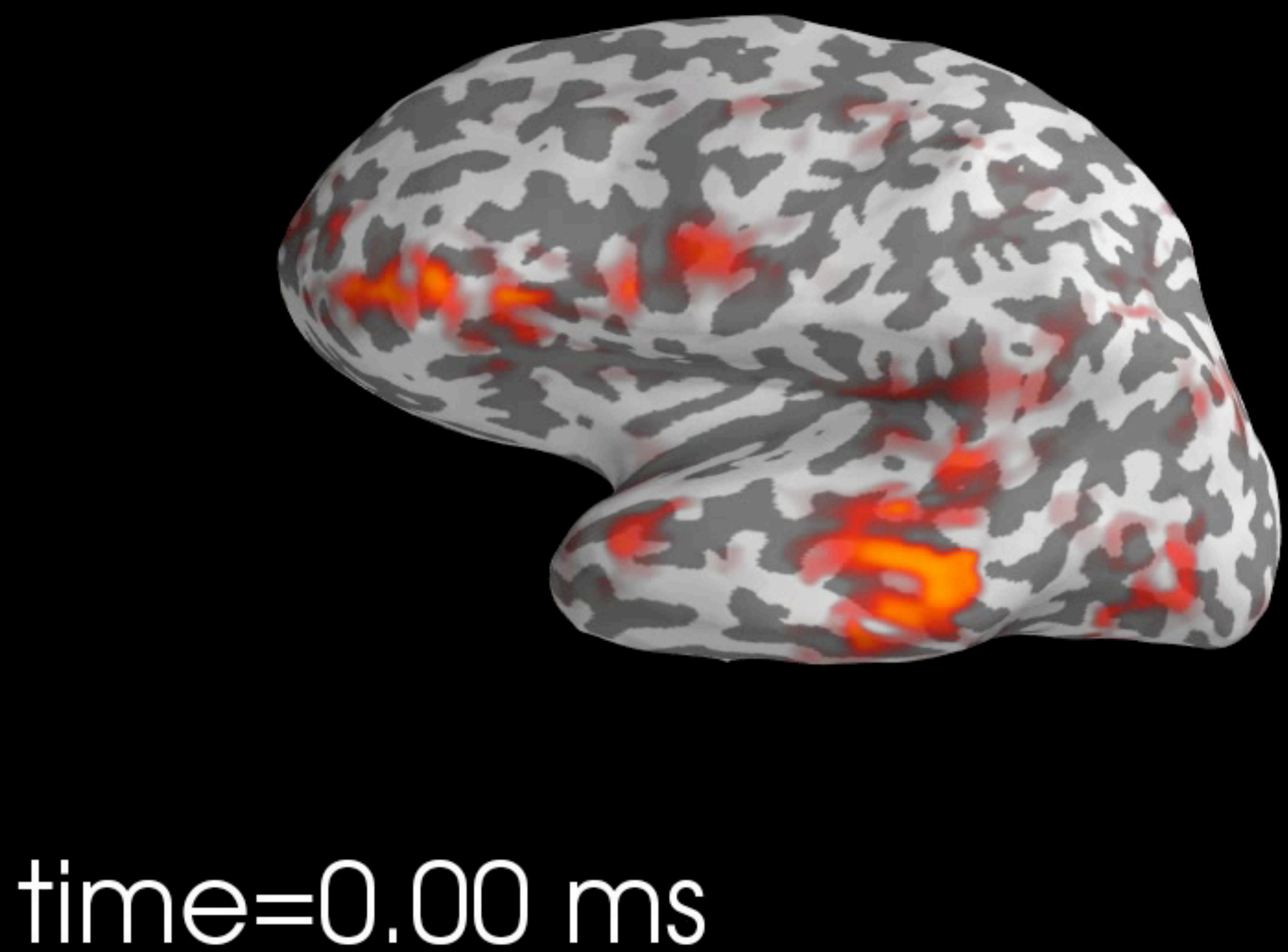
$$\mathbf{X}^* = \underbrace{\mathbf{\Sigma}^{-1}\mathbf{G}^T(\mathbf{G}\mathbf{\Sigma}^{-1}\mathbf{G}^T + \lambda\mathbf{Id})^{-1}}_{\text{Inverse operator}}\mathbf{M}$$

[Tikhonov et al. 77, Wang et al. 92, Hämäläinen et al. 94]

Remarks:

- **Really fast** to compute (SVD of **G**), hence very much used in the field.
- In practice, it's **much more complicated** (whitening data, correcting artifacts, channels with different SNRs, setting λ based on SNR, loose orientation, SNR varies with time...)

THM: A lot of domain knowledge to make it work http://youtu.be/Uxr5Pz7JPrs



Alexandre Gramfort, Telecom ParisTech, CEA/Neurospin

Demo