Functional Brain Imaging with MEG (Magnetoencephalography), EEG (Electroencephalography) and sEEG (Stereotaxic EEG)

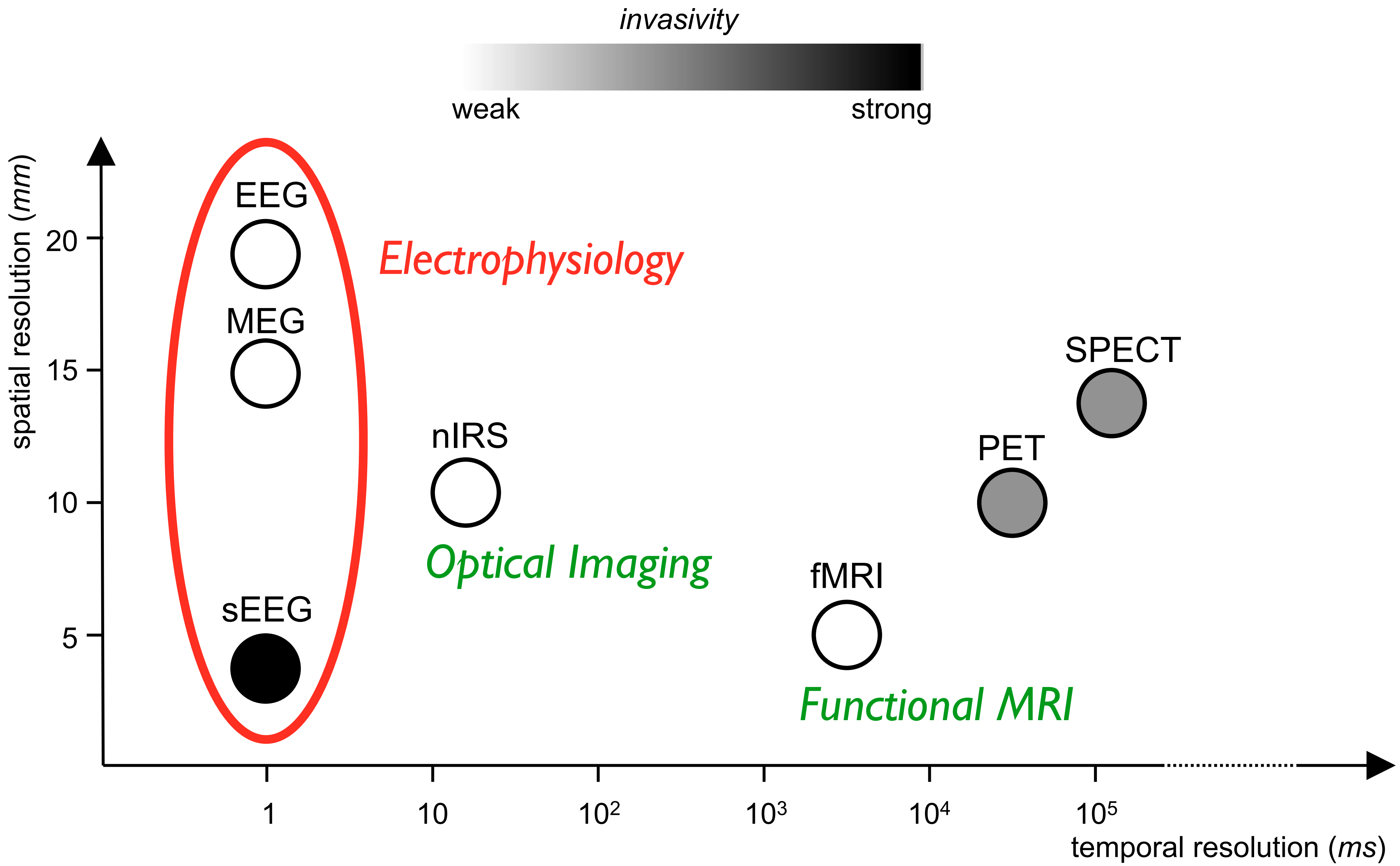
Alexandre Gramfort
alexandre.gramfort@telecom-paristech.fr

2013

TELECOM ParisTech

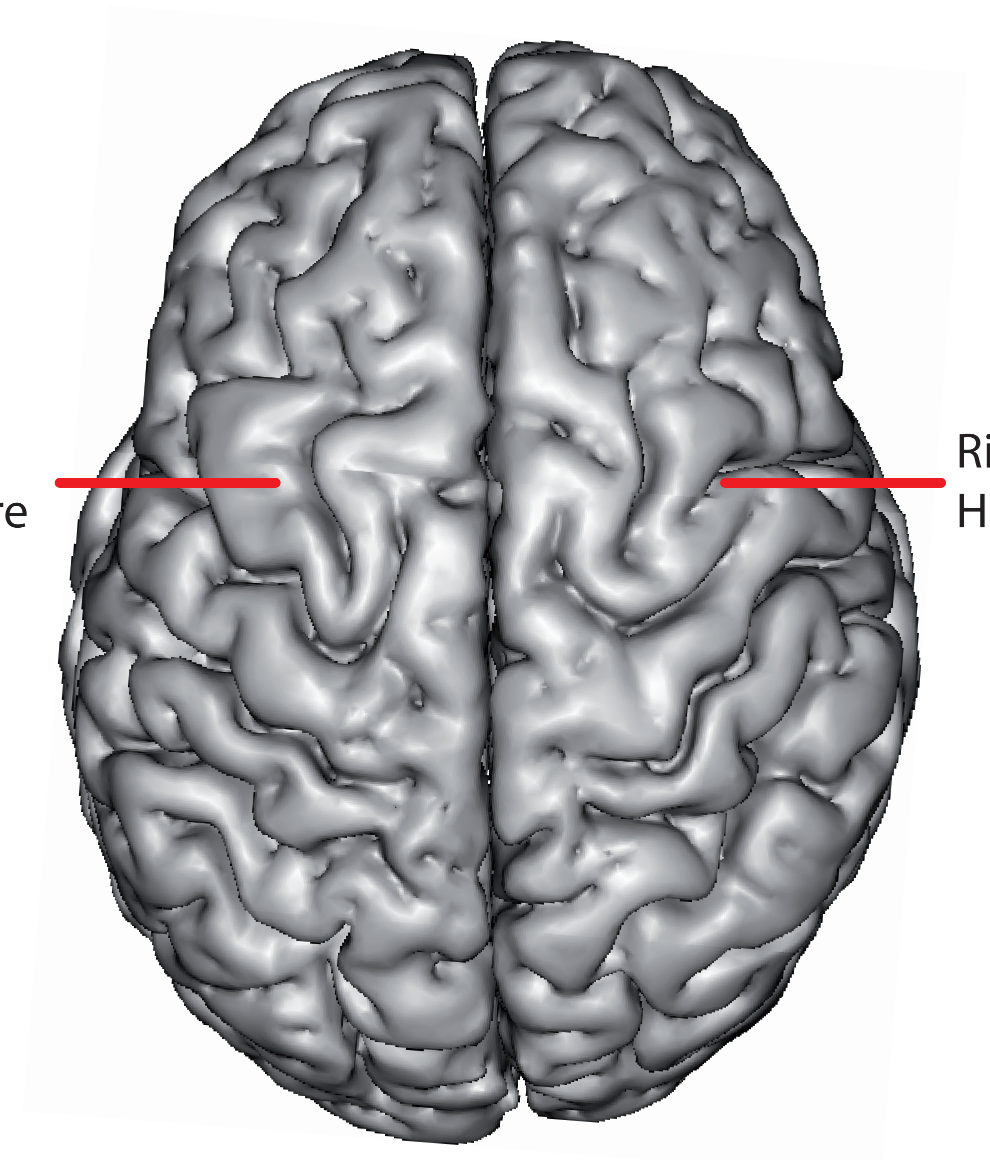
Functional neuroimaging

It's the study of the **brain activity** through **functional imaging devices**



Brain anatomy

Left Hemisphere

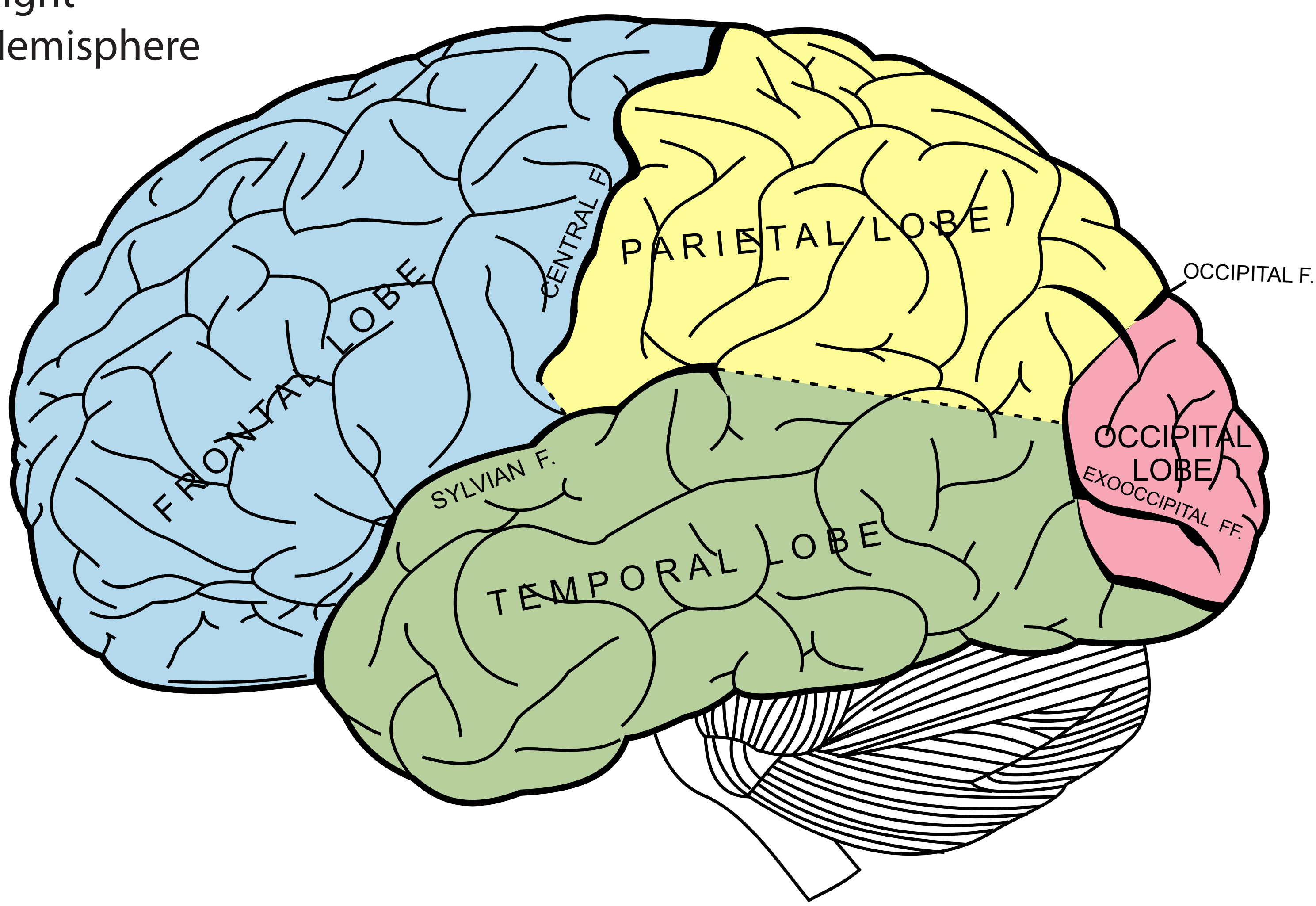


Right Hemisphere

Brain mesh obtained by MRI segmentation

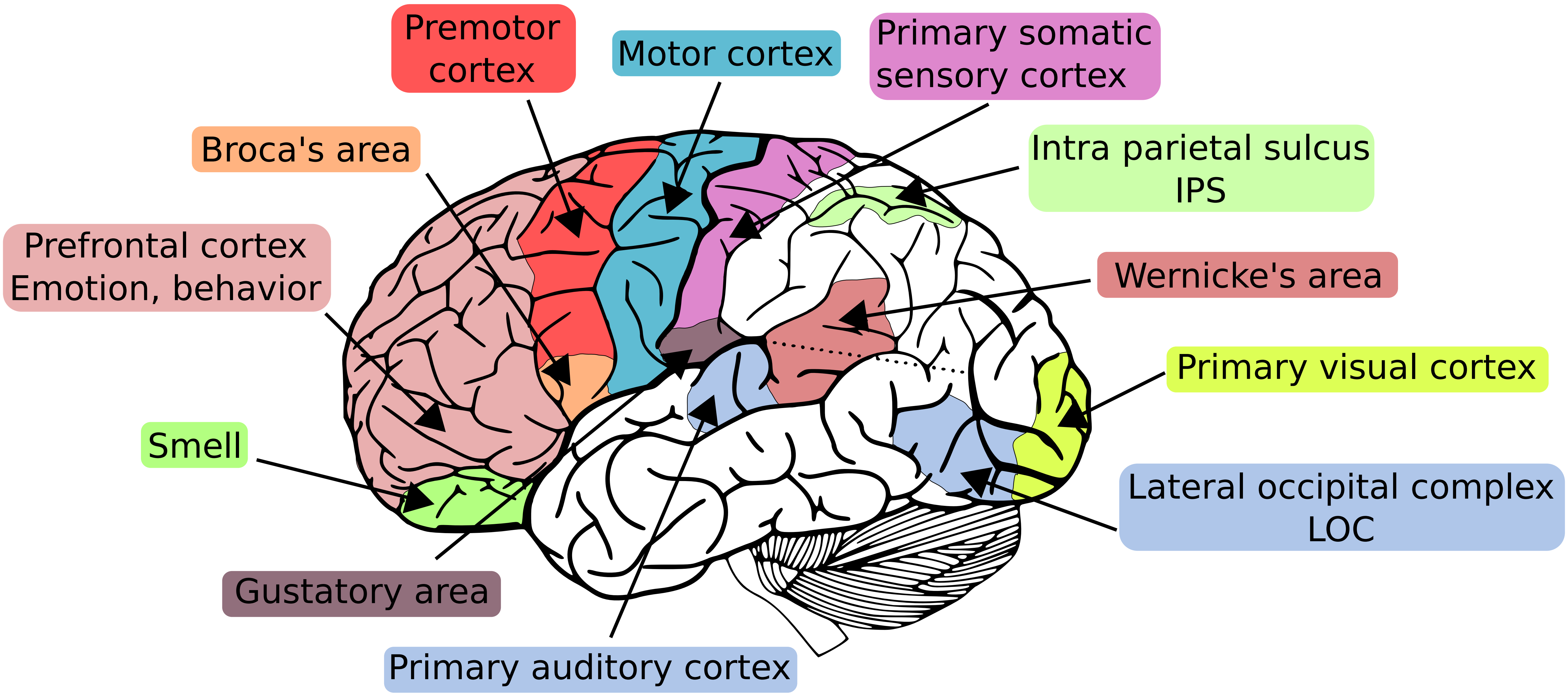
2 hemispheres

4 Lobes

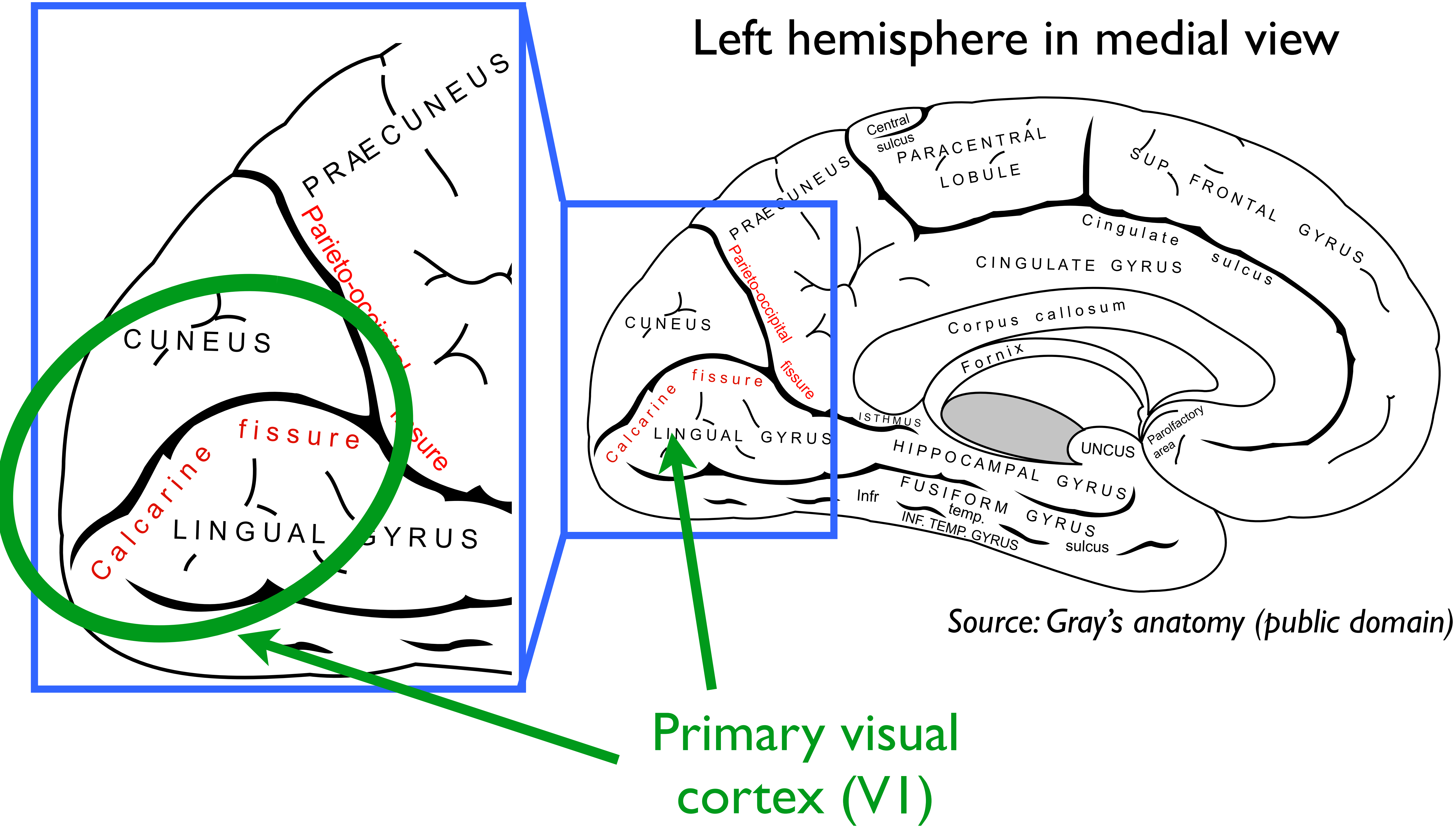


Source: Gray's anatomy (public domain)

Relation between location and function?

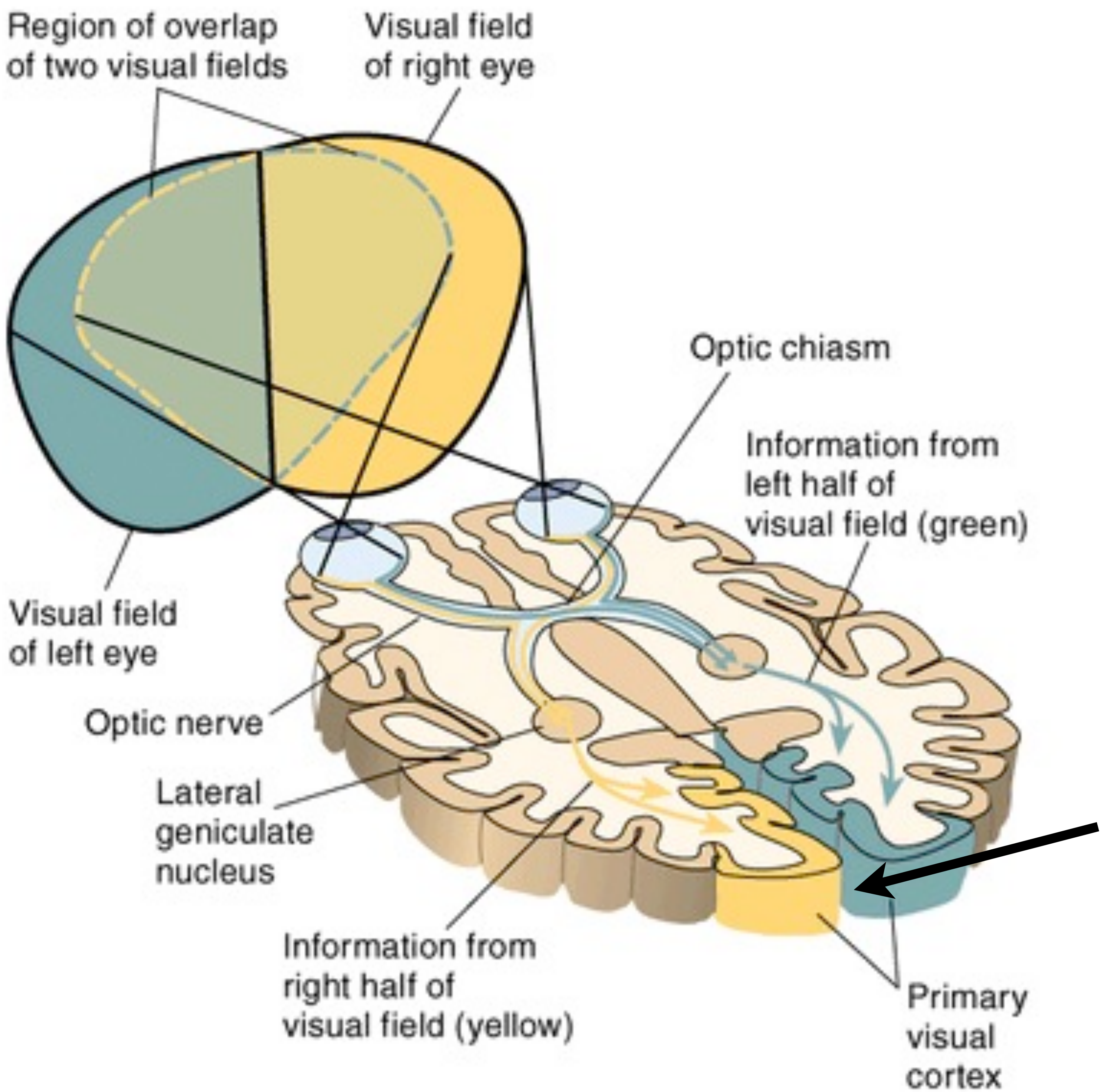


Relation between location and function?



Source: Gray's anatomy (public domain)

From the eye to the cortex

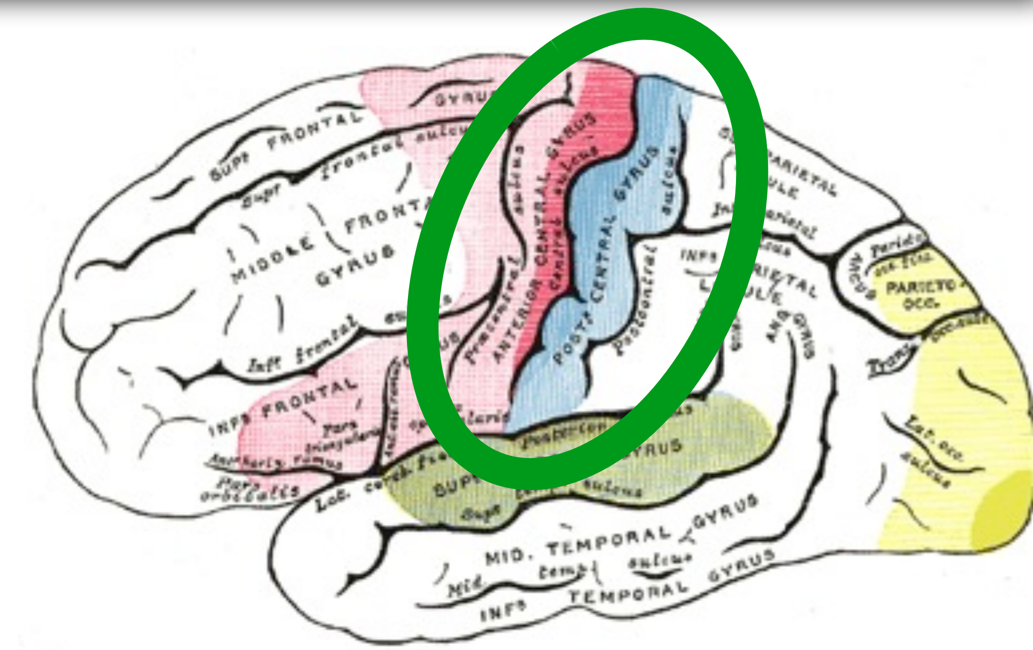


Left (resp. right) **visual field** is projected to the **right** (resp. left) **hemisphere** in the **primary visual cortex (V1)**

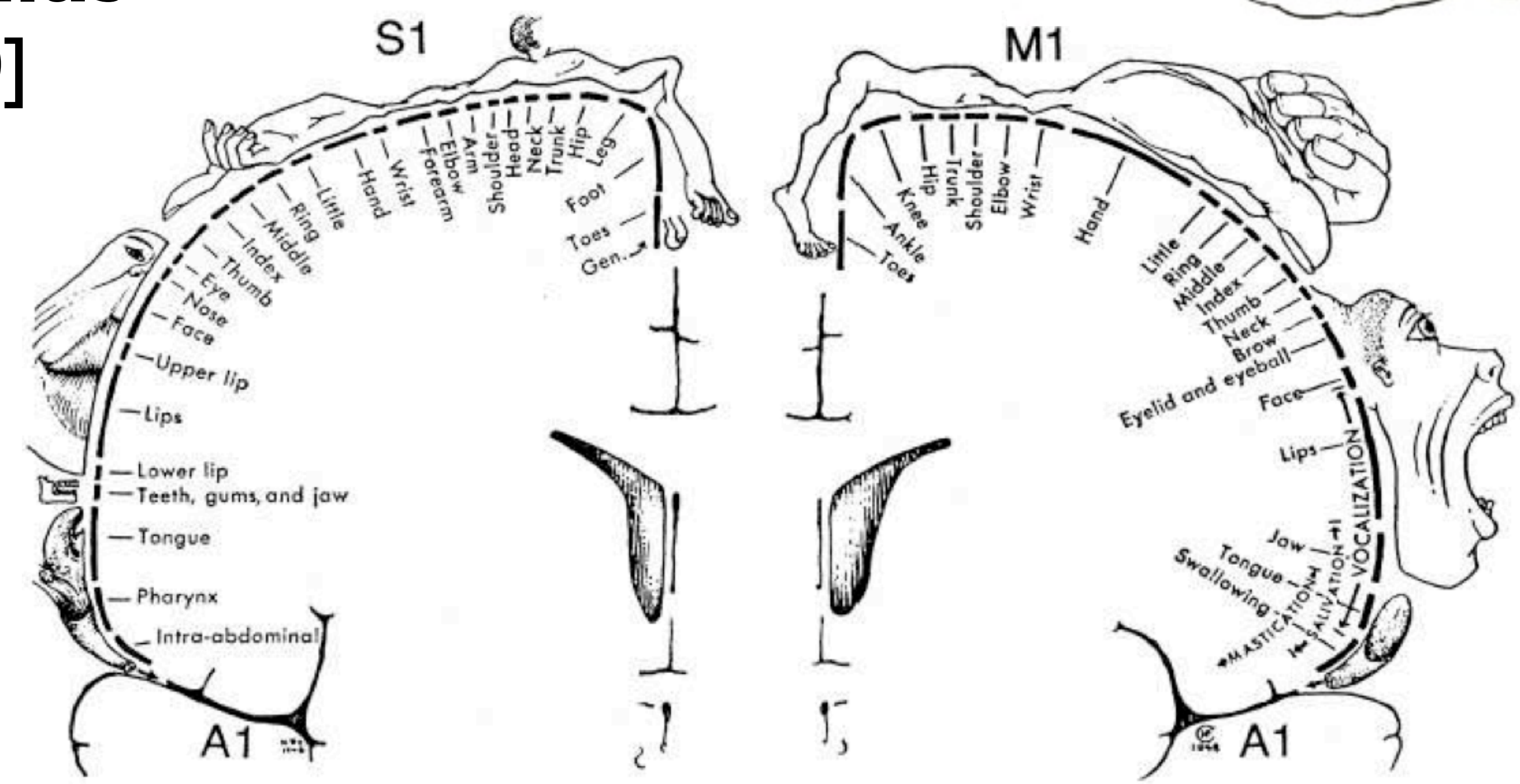
V1 stands in the occipital region around the **calcarine fissure**

Source: adapted from http://homepage.psy.utexas.edu/homepage/Class/Psy308/ Salinas/Vision/Vision.html).

Relation between location and function?



Homonculus
[Penfield 50]



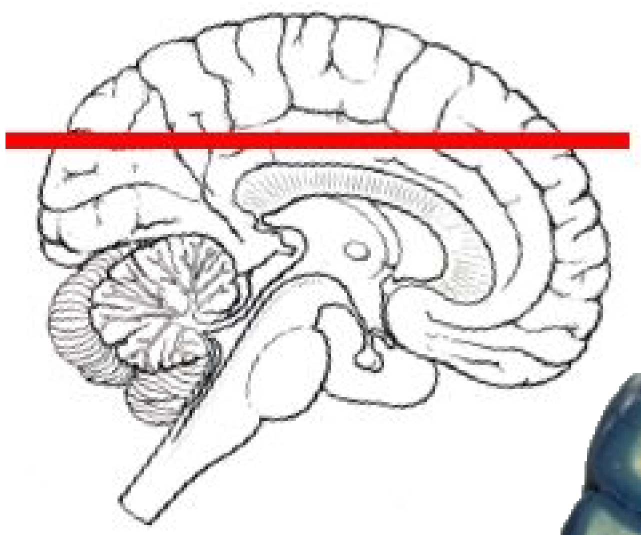
Primary Somatosensory Cortex (S1)

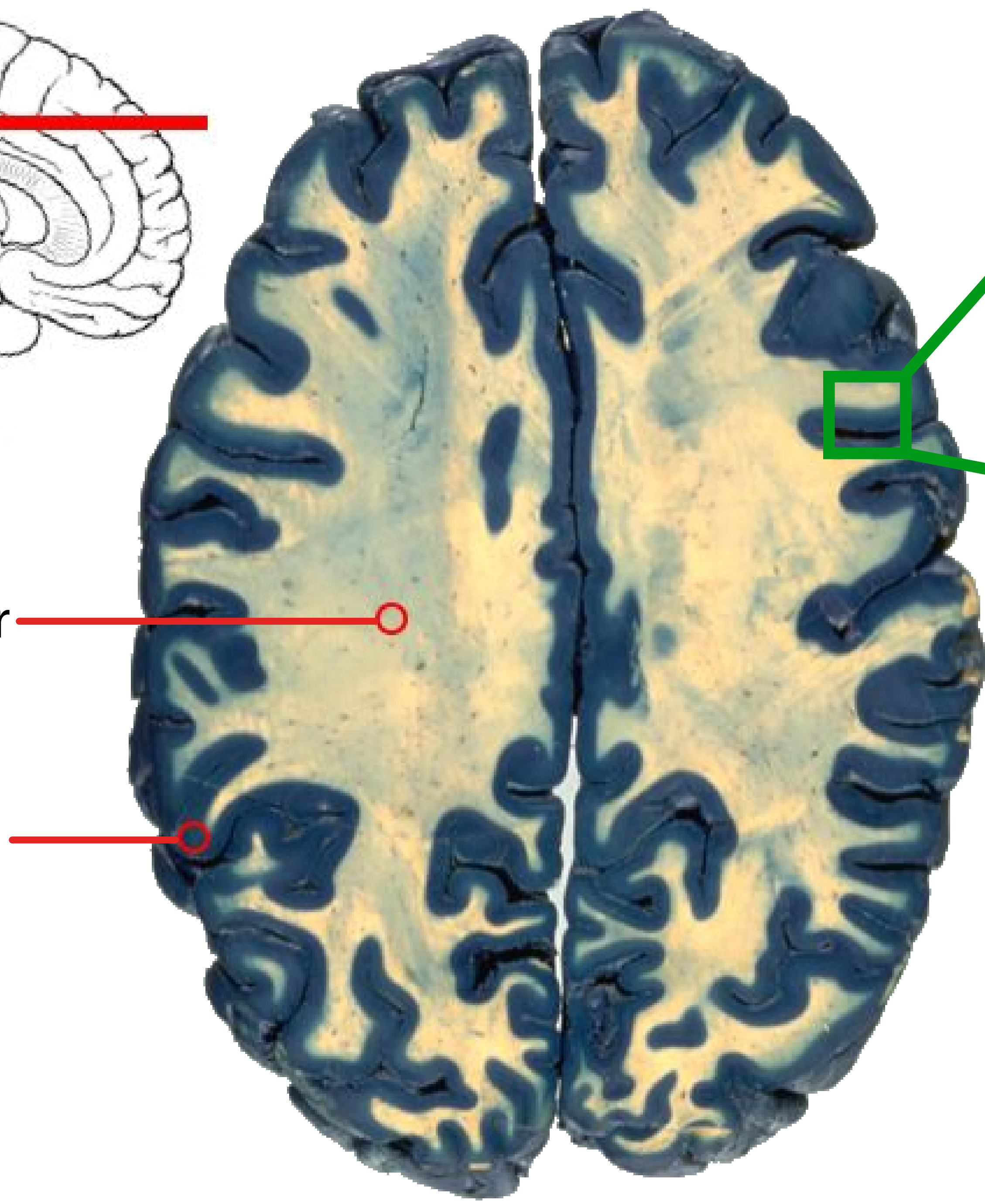
Primary Motor Cortex (M1)

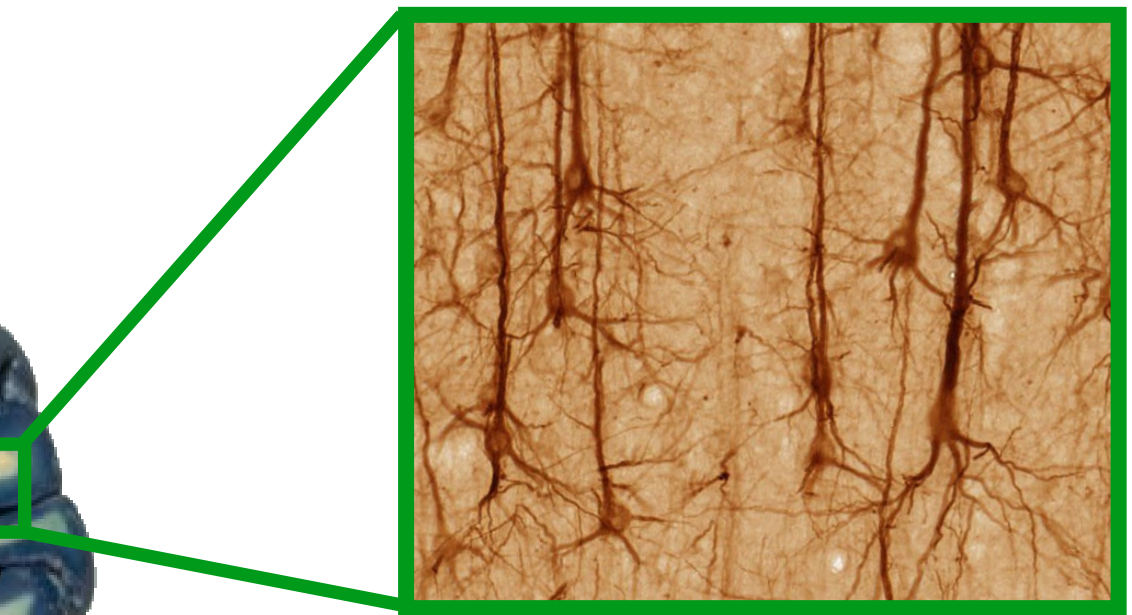
Electrophysiology: Origin of the signals

Brain anatomy

Axial slice



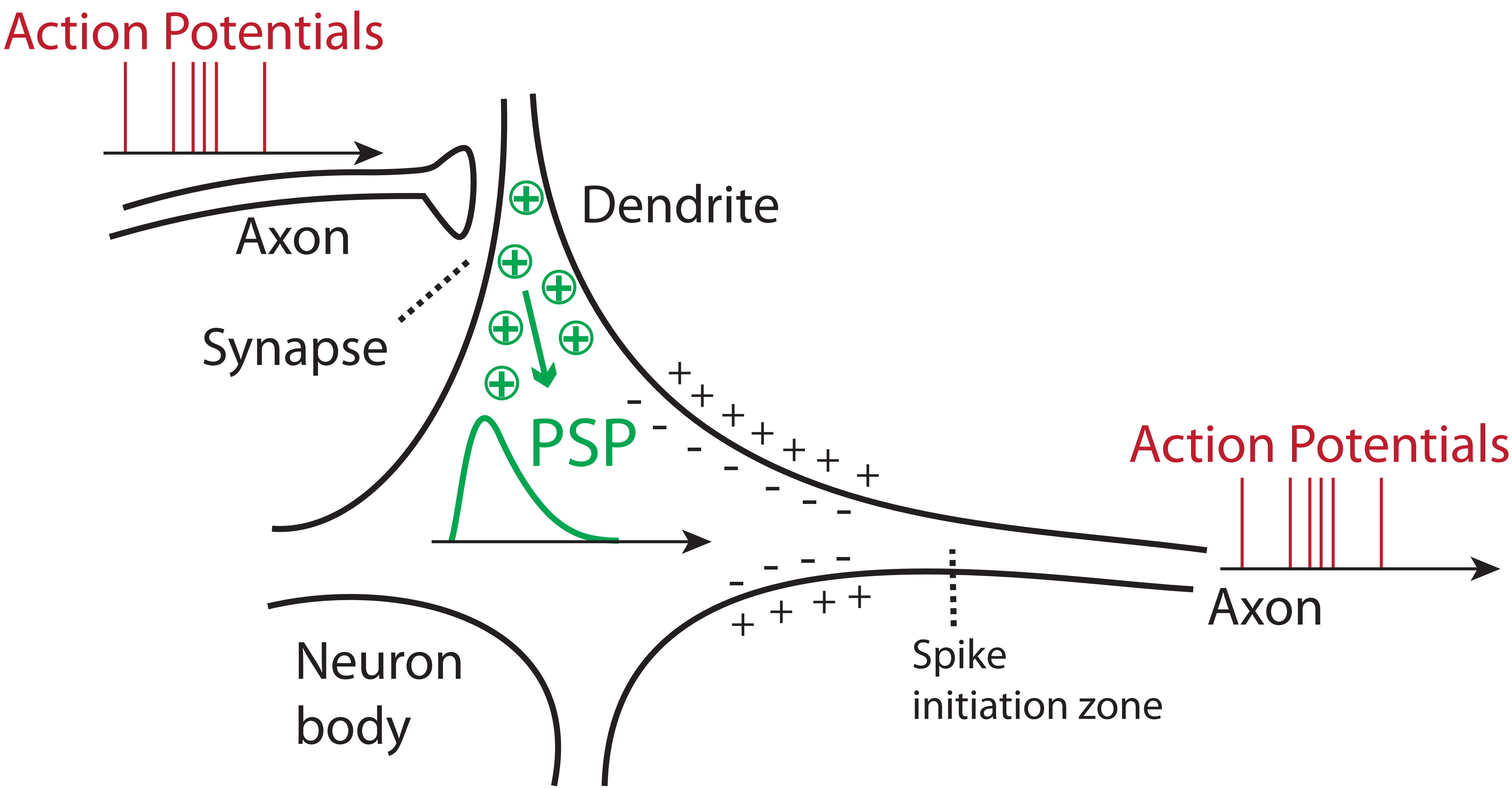




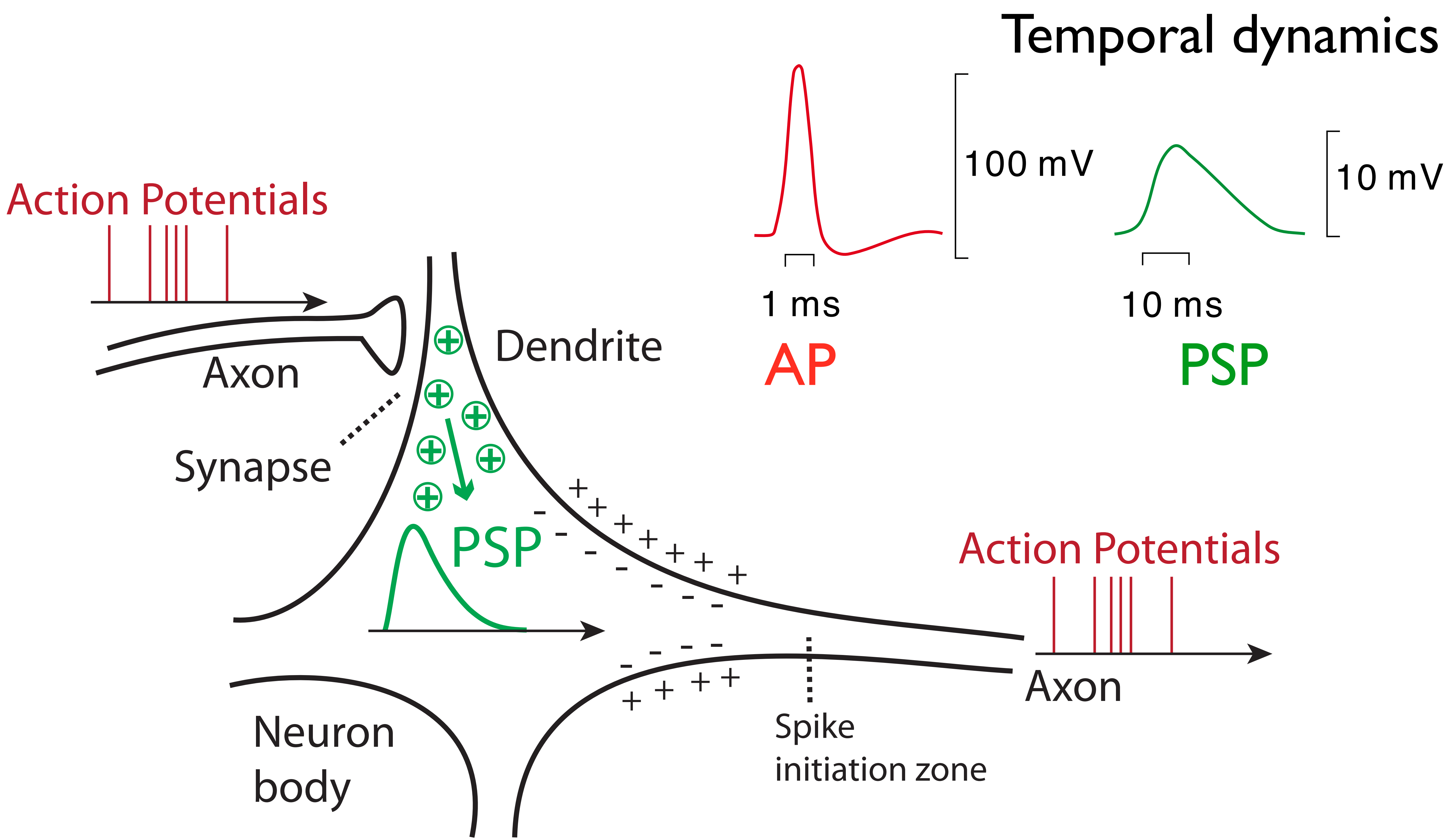
Neurons
in the gray matter

Source: dartmouth.edu

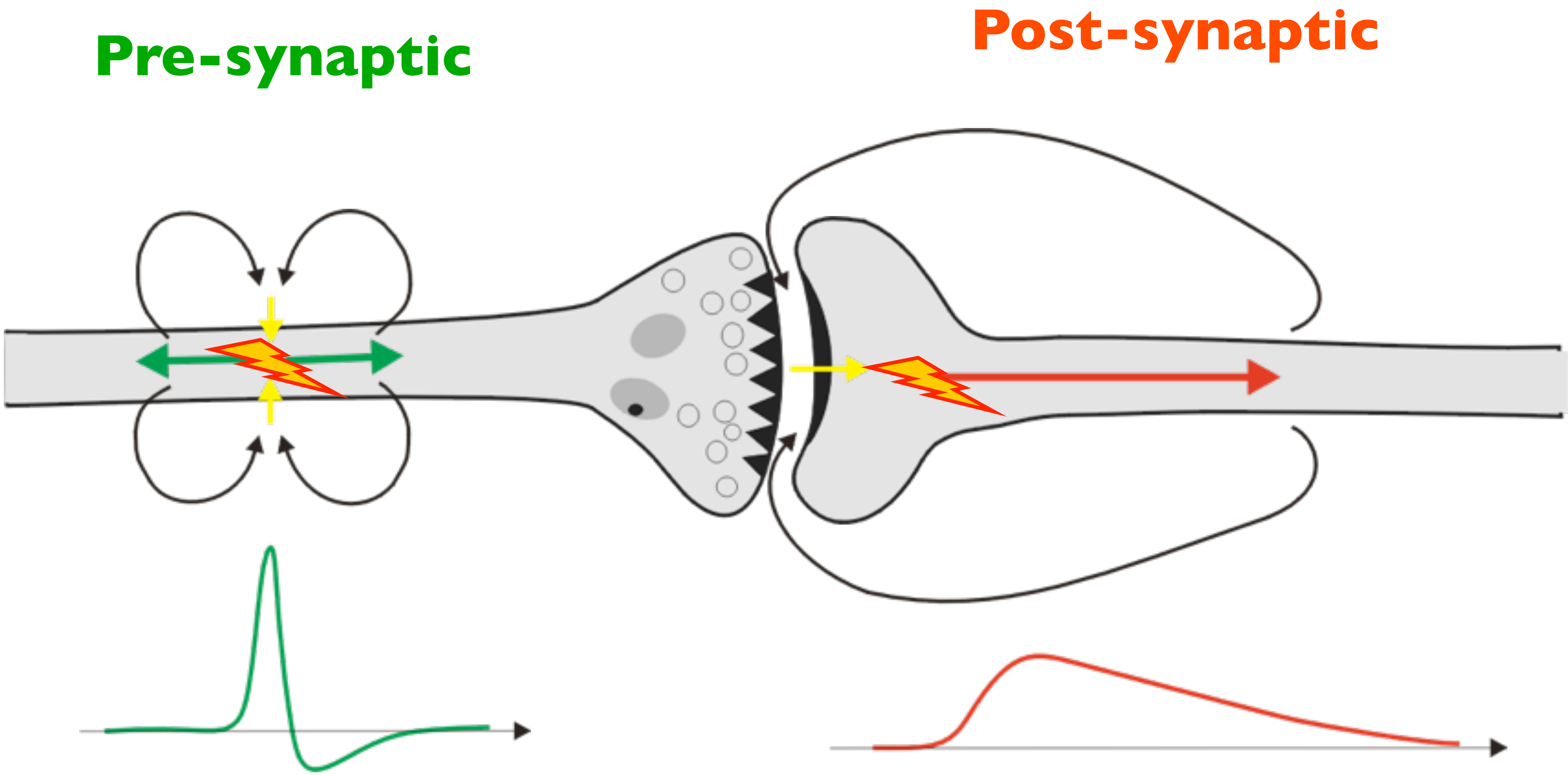
APs (action potentials) & PSPs (post-synaptic potentials)



APs (action potentials) & PSPs (post-synaptic potentials)



APs (action potentials) & PSPs (post-synaptic potentials)



Action potentials:

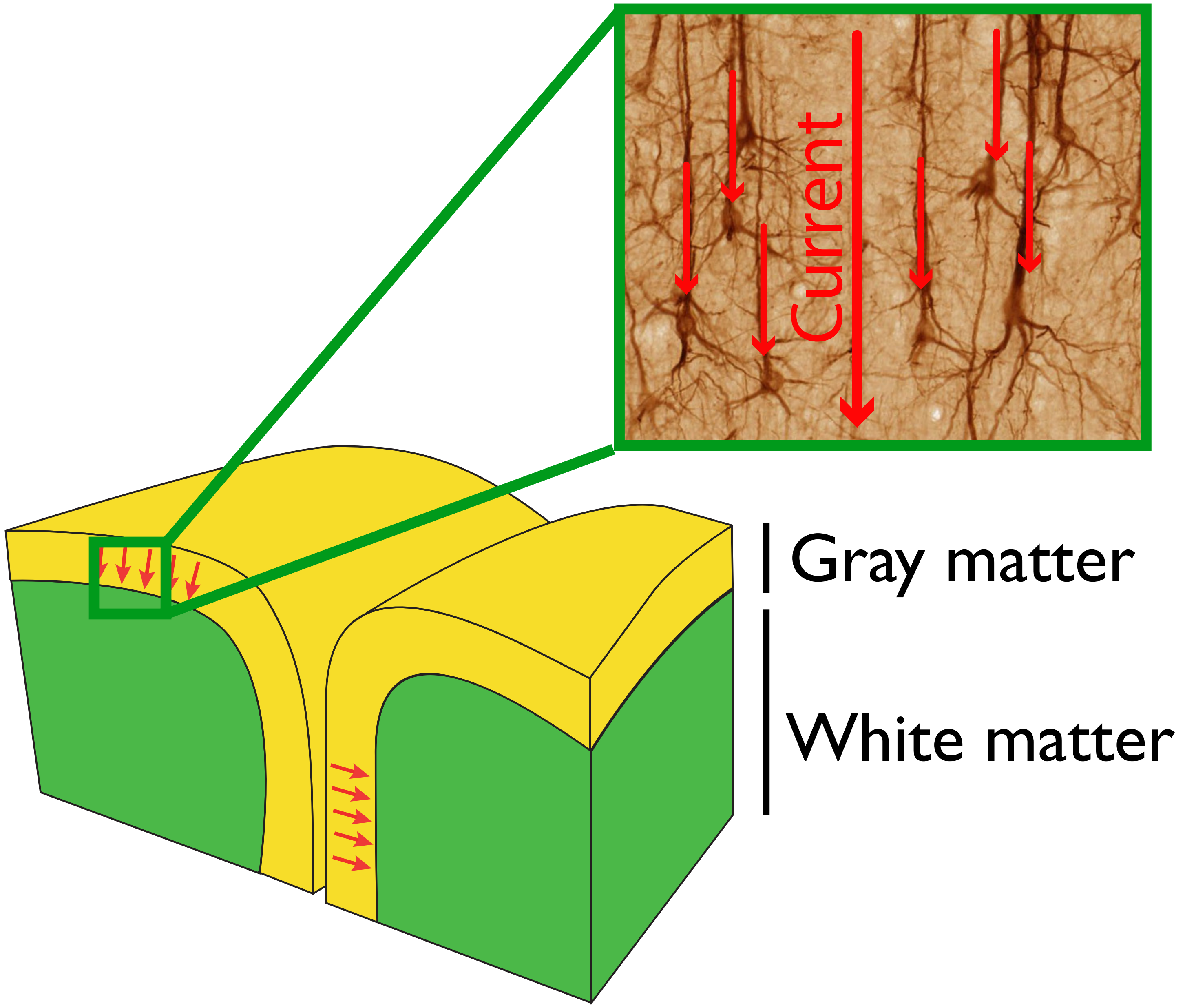
fields diminish too rapidly to sum

Postsynaptic currents:

fields diminish gradually

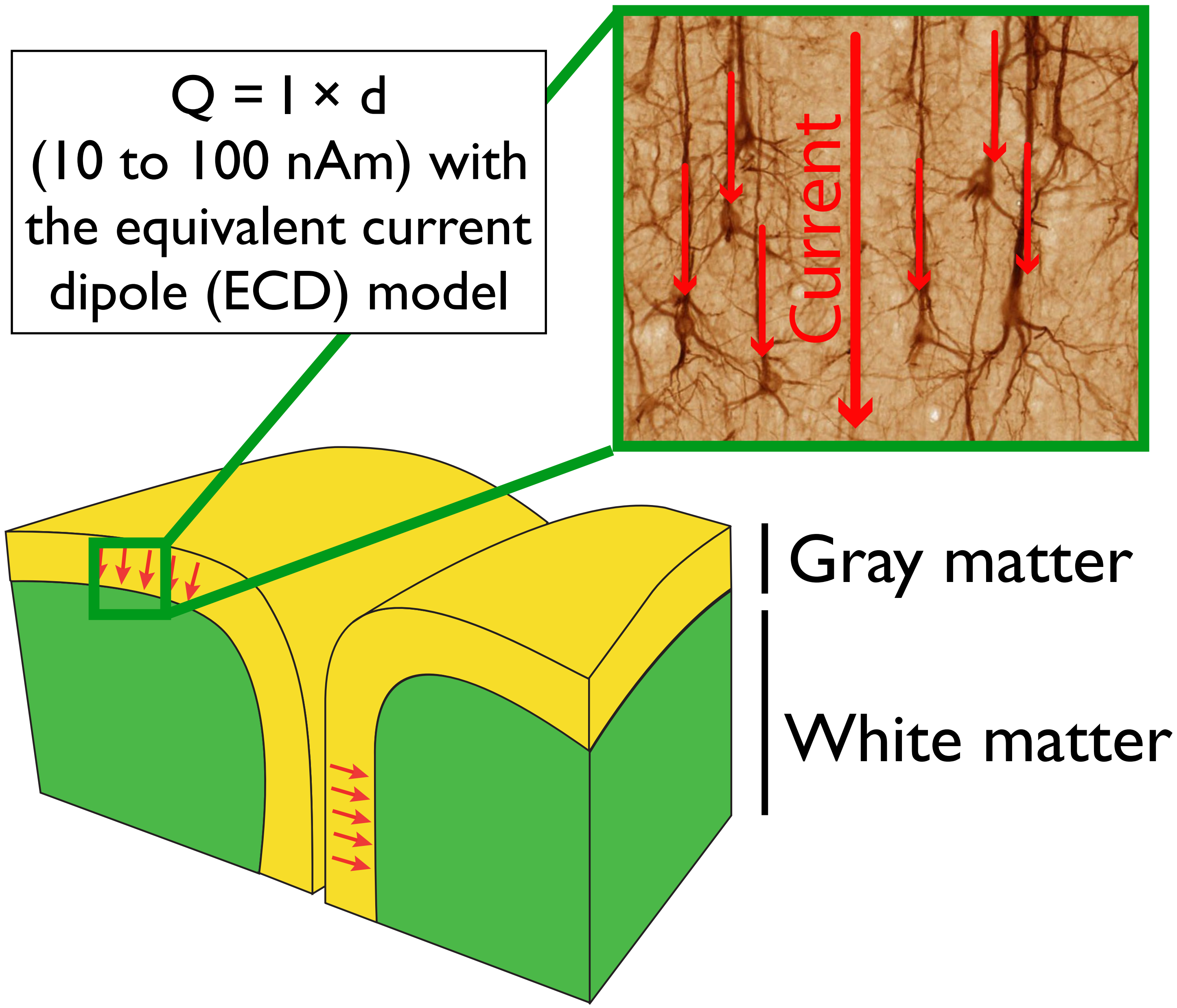
Neurons as current generators

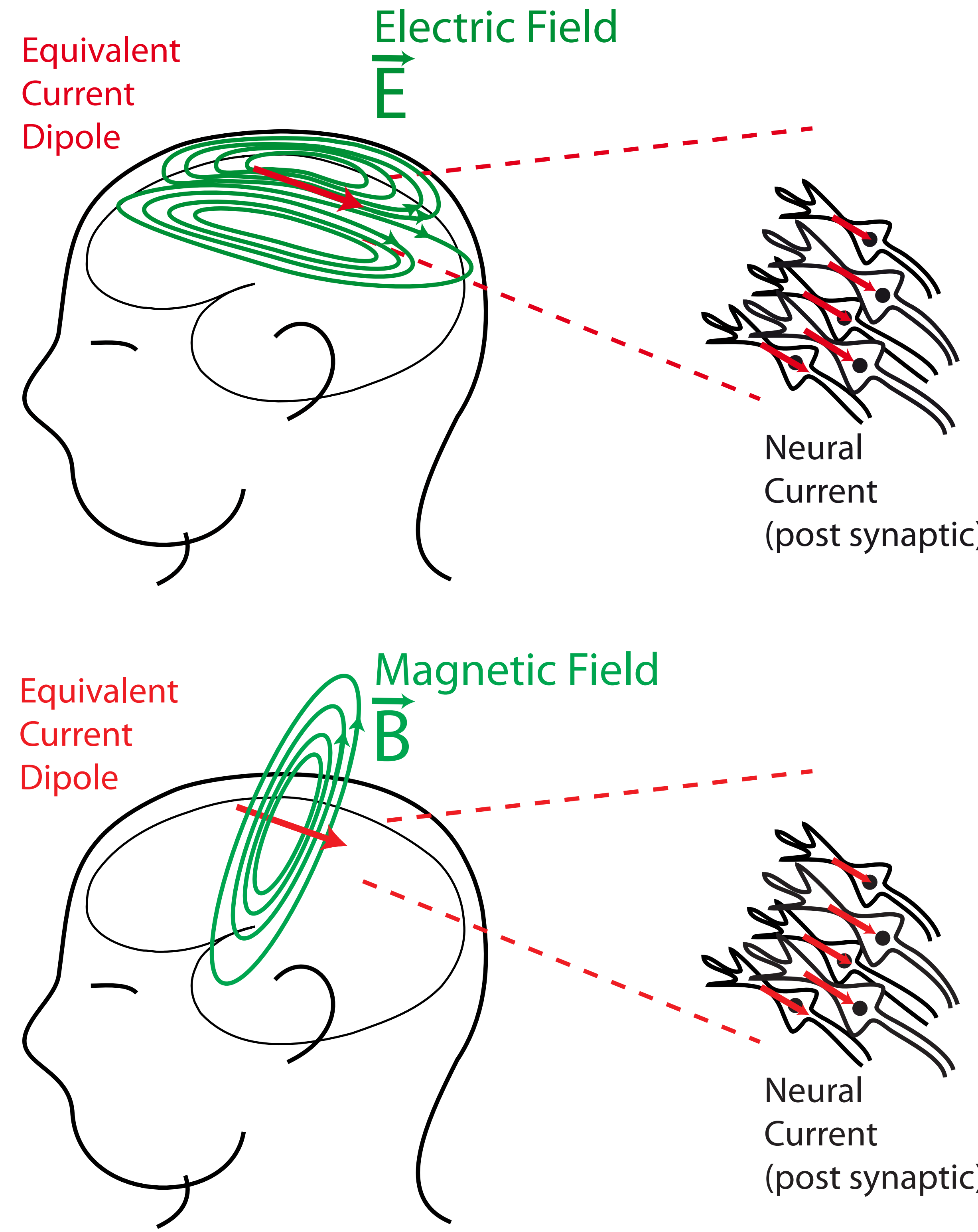
Large cortical pyramidal cells organized in macro-assemblies with their **dendrites normally oriented to the local cortical surface**



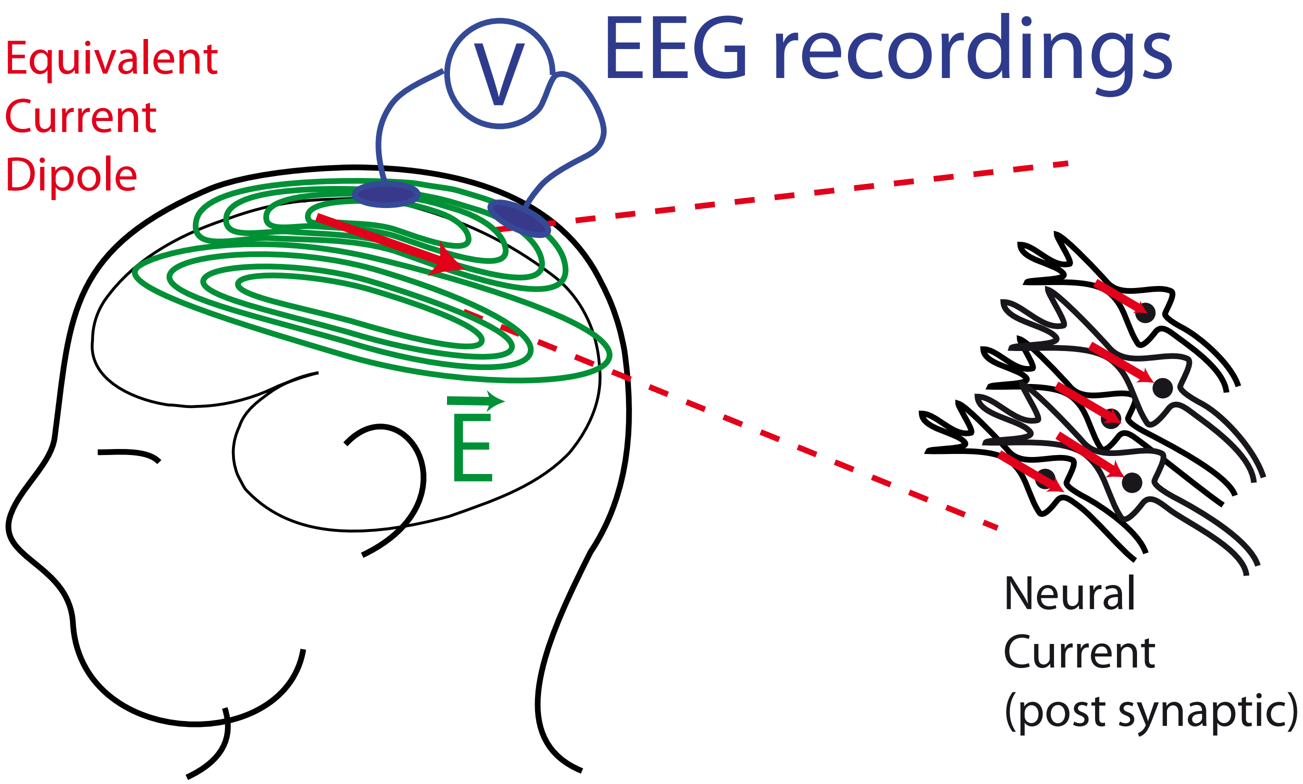
Neurons as current generators

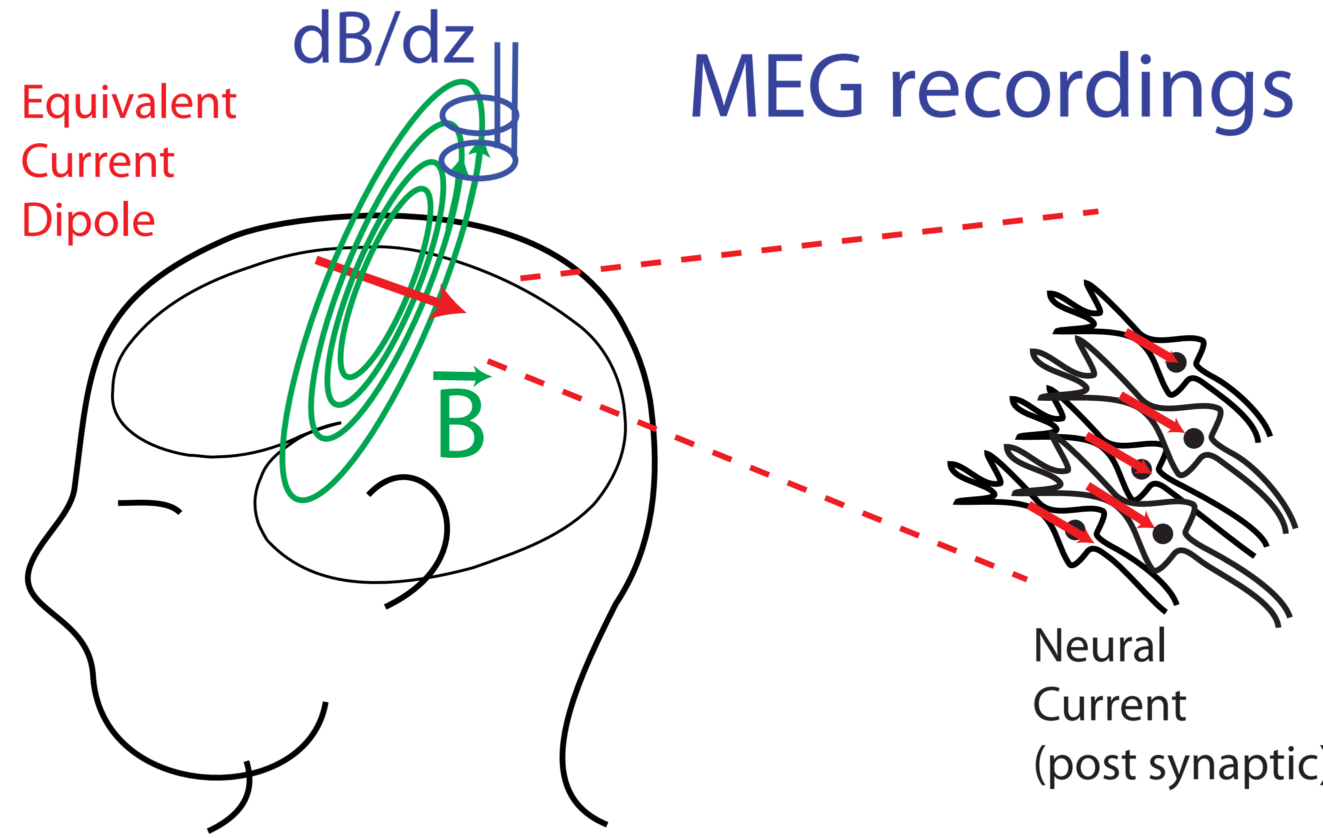
Large cortical pyramidal cells organized in macro-assemblies with their **dendrites normally oriented to the local cortical surface**



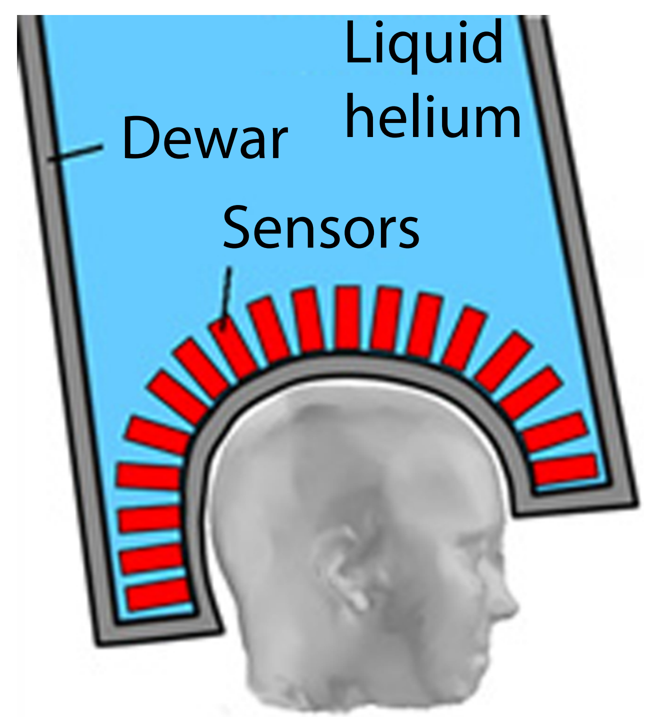


EEG & MEG systems



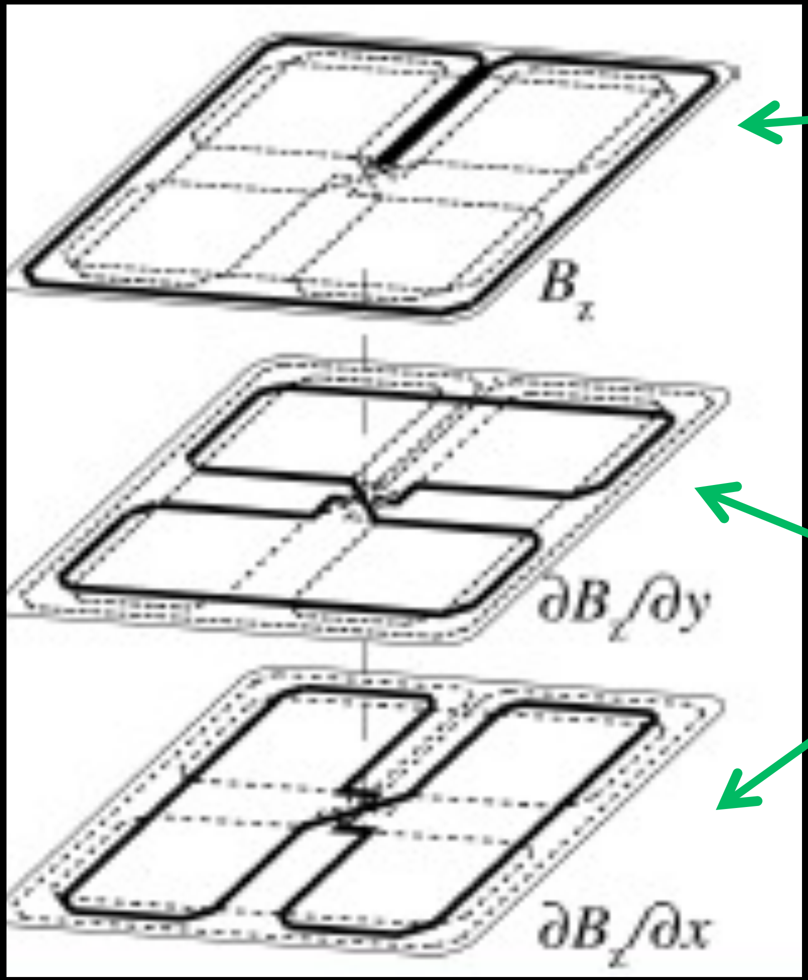


First EEG
recordings
in 1929
by H. Berger



Hôpital La Timone
Marseille, France

MEG sensors

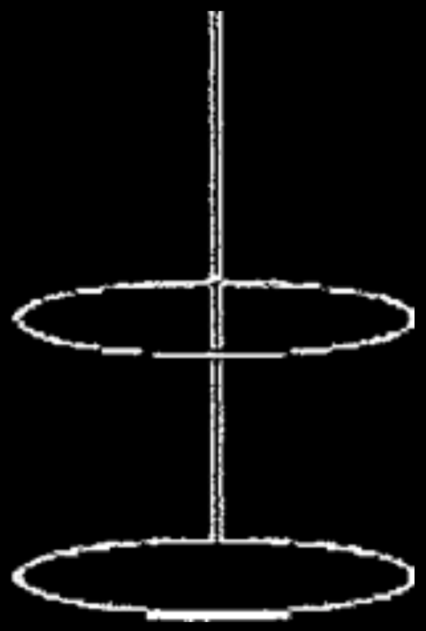


Magnetometer

- General magnetic fields
- Very sensitive overall, **noisy**

Planar Gradiometer

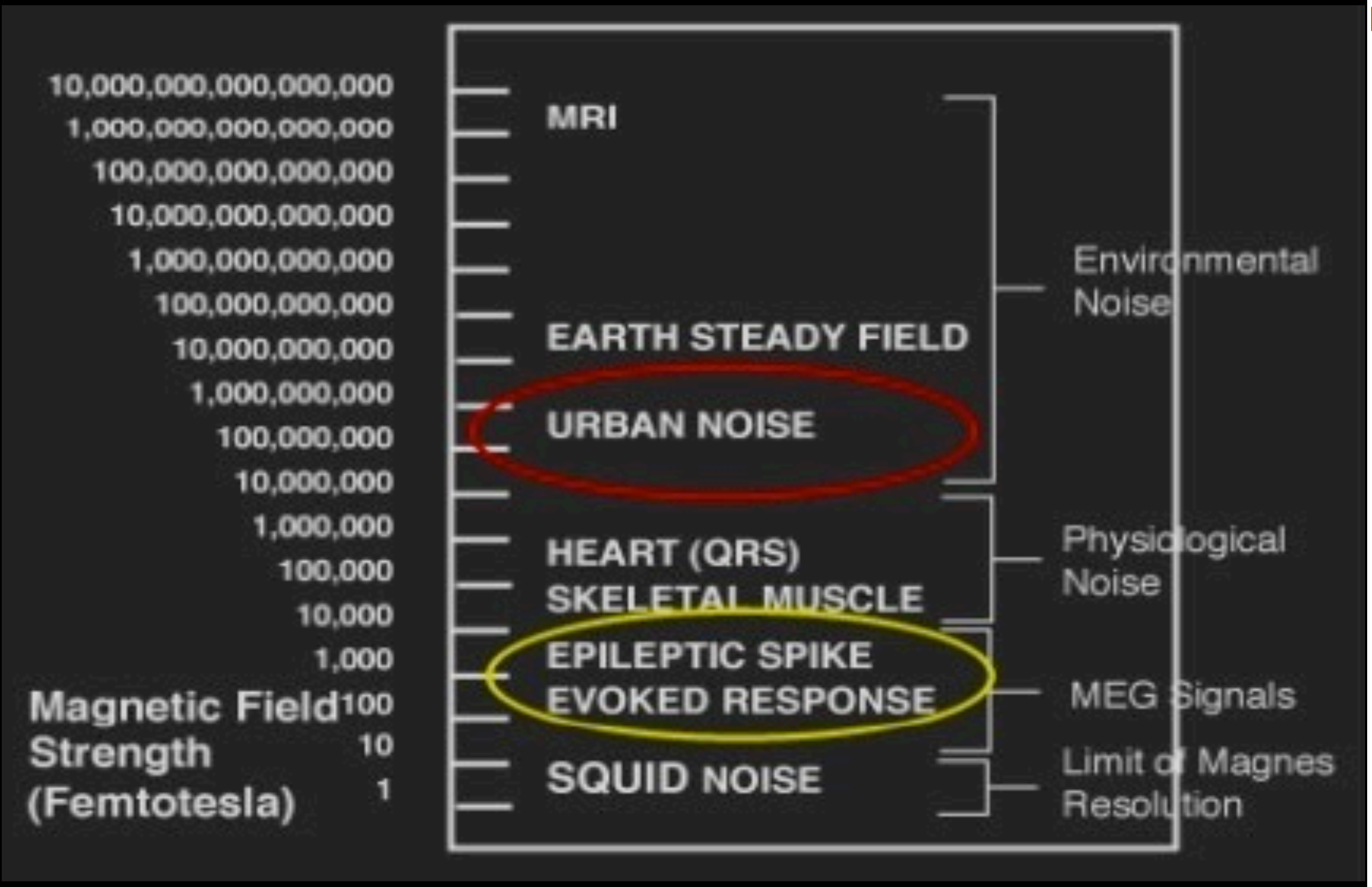
- Focal magnetic fields
- Most sensitive to fields directly underneath

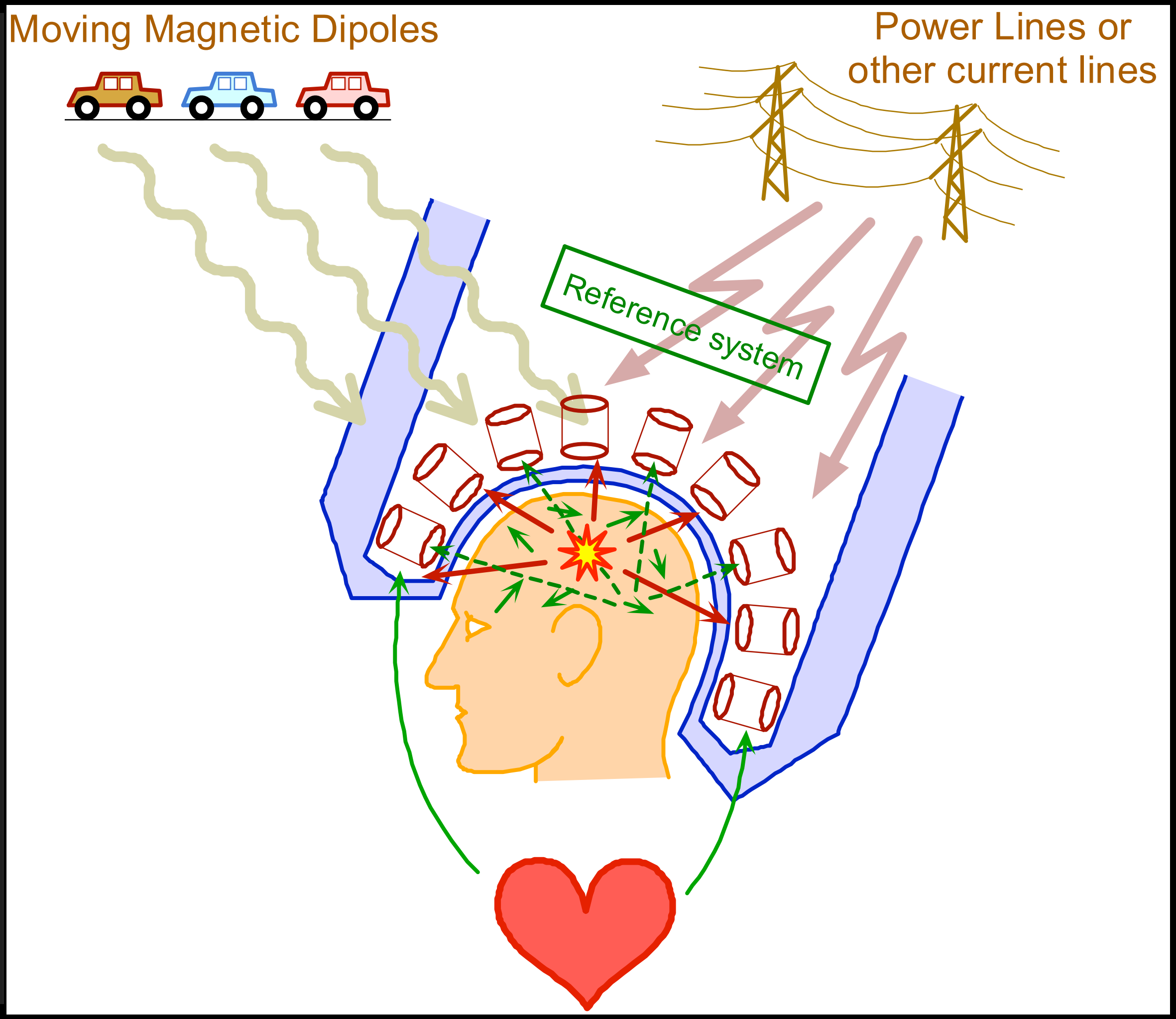


Axial Gradiometer

-Focal magnetic fields
- Most sensitive to fields directly underneath it

Magnetic shielding

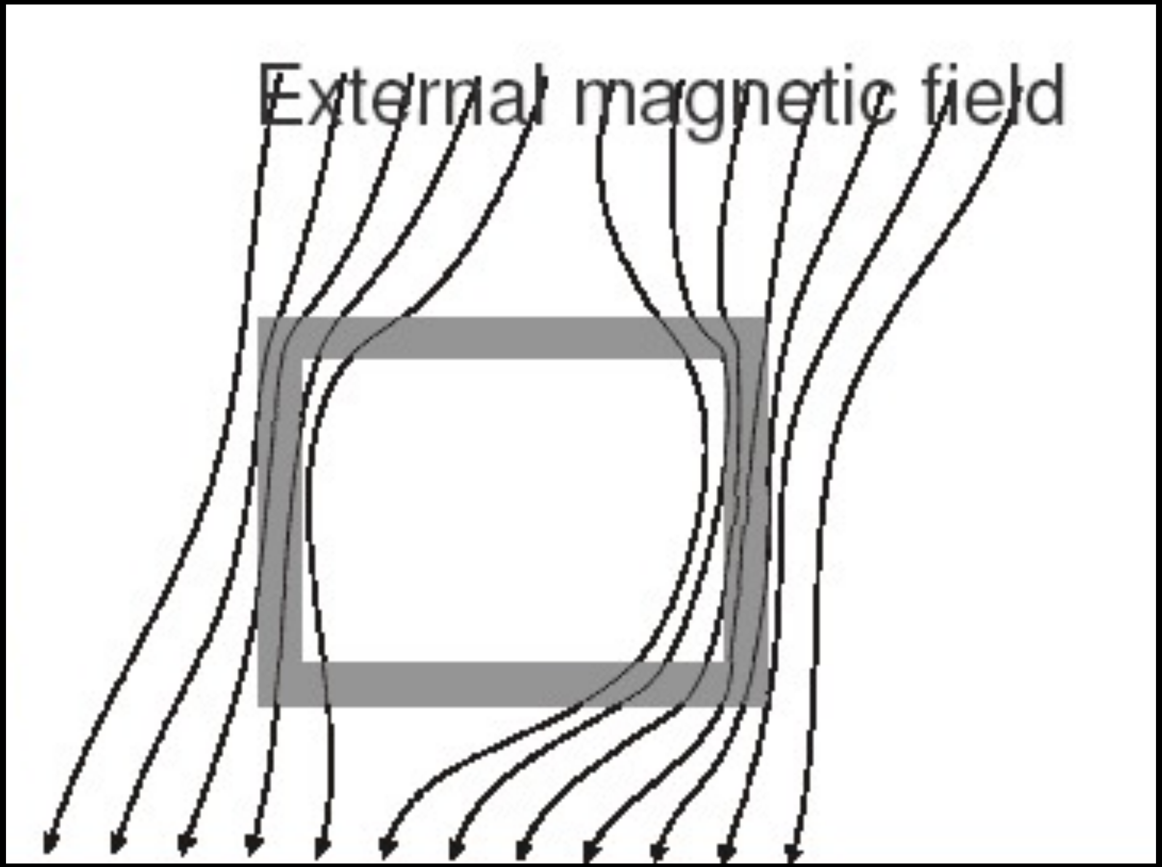




Hence the importance of shielding...

Magnetic shielding

Magnetically Shielded Room (MSR)



3-ply µ-metal room

A machine (Neromag vectorview)

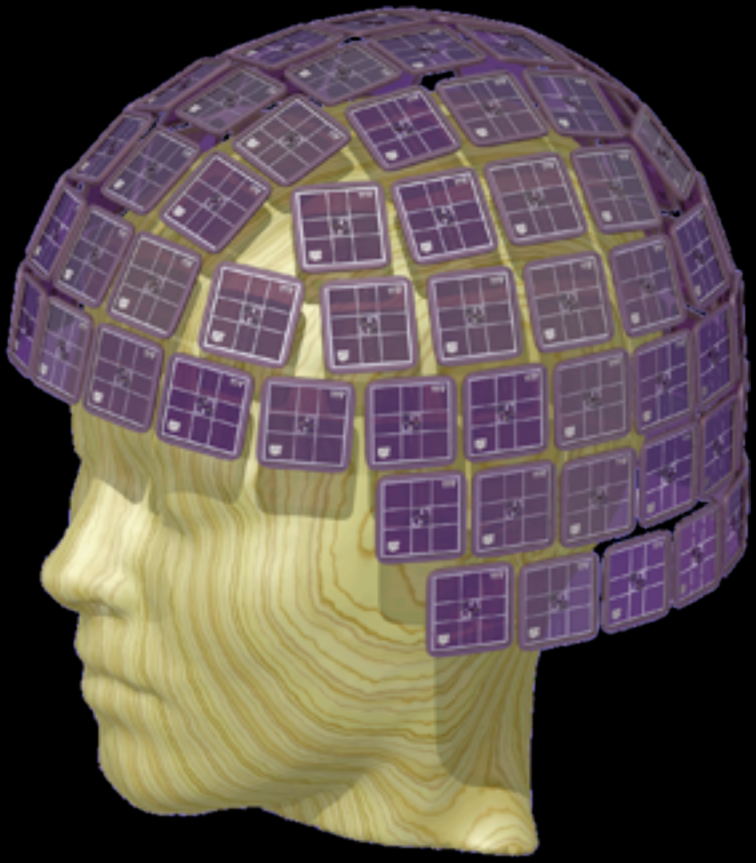
No Magnet

Quiet

Machine makes no noise

Participant can sit or lay down

Can record 128 EEG simultaneously

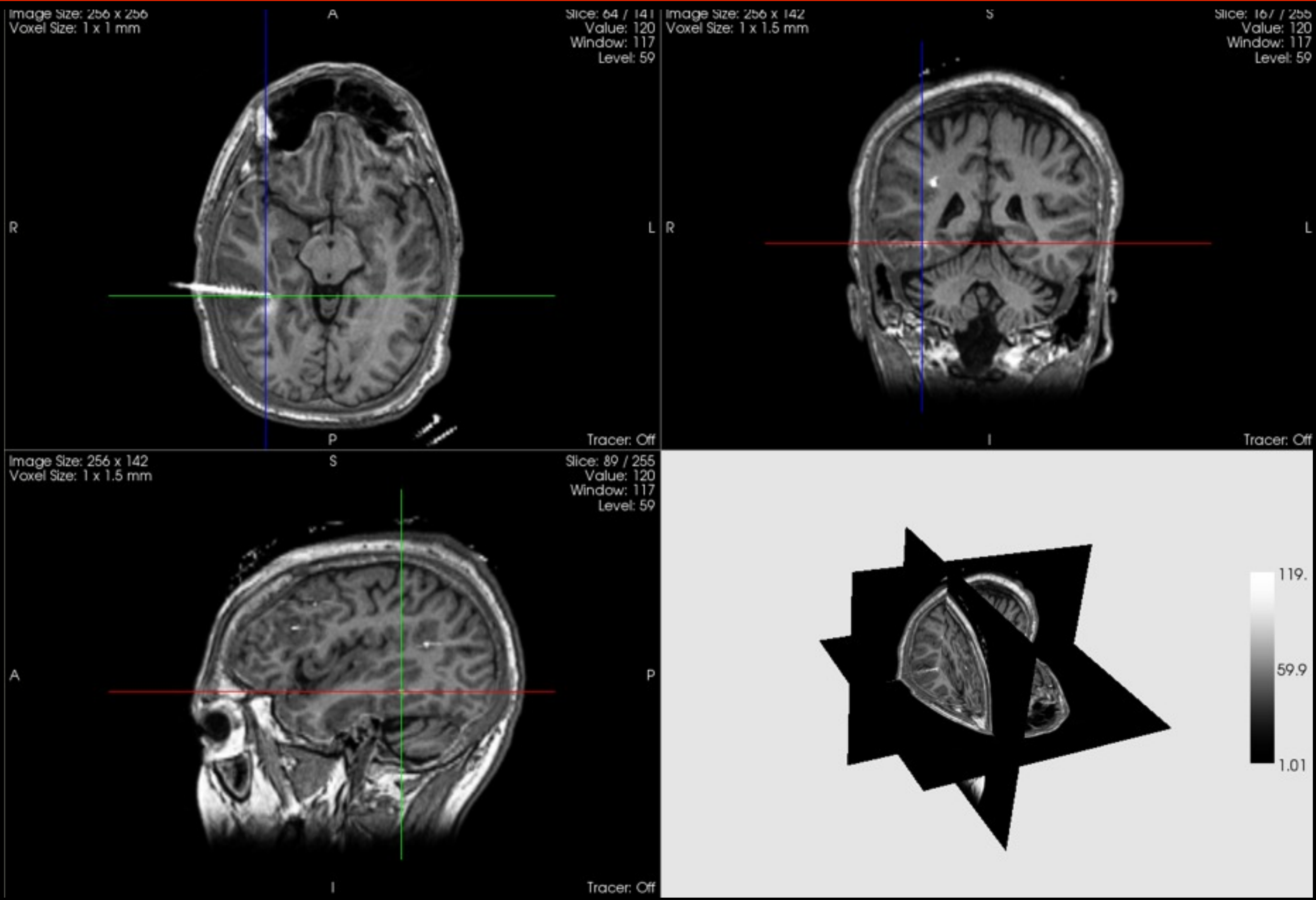


sEEG systems

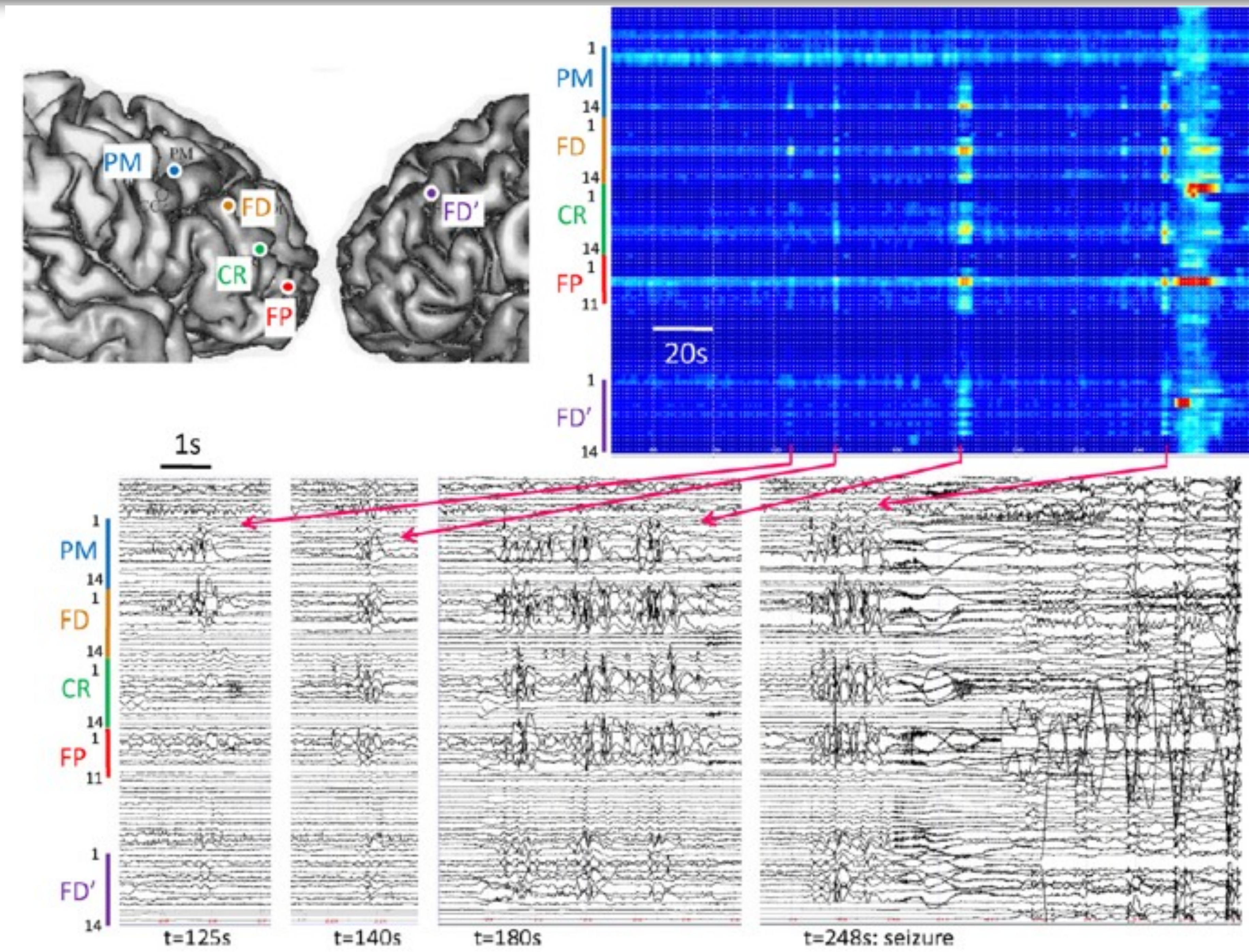
Intracranial electrodes;
5 to 15 contacts per electrode
Around 10 electrodes are implanted

Stereotaxic Implantation

sEEG systems



sEEG Measurements

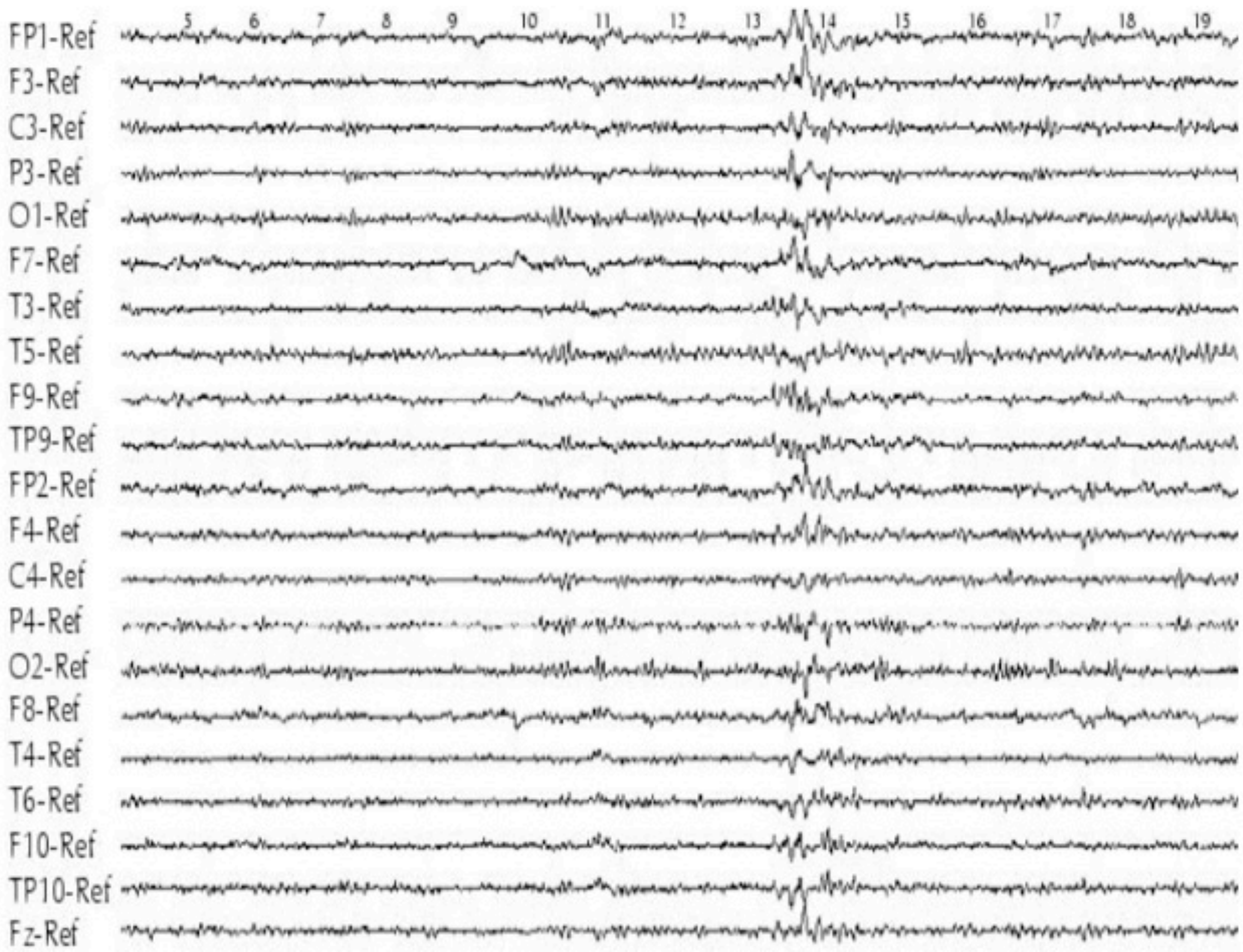


Interictal discharges involving multiple regions (a network)

Seizure onset

[Schwartz et al Epilepsy Res 2011]

M/EEG Measurements



Sample EEG measurements

EEG :

- ≈ 100 sensors

MEG :

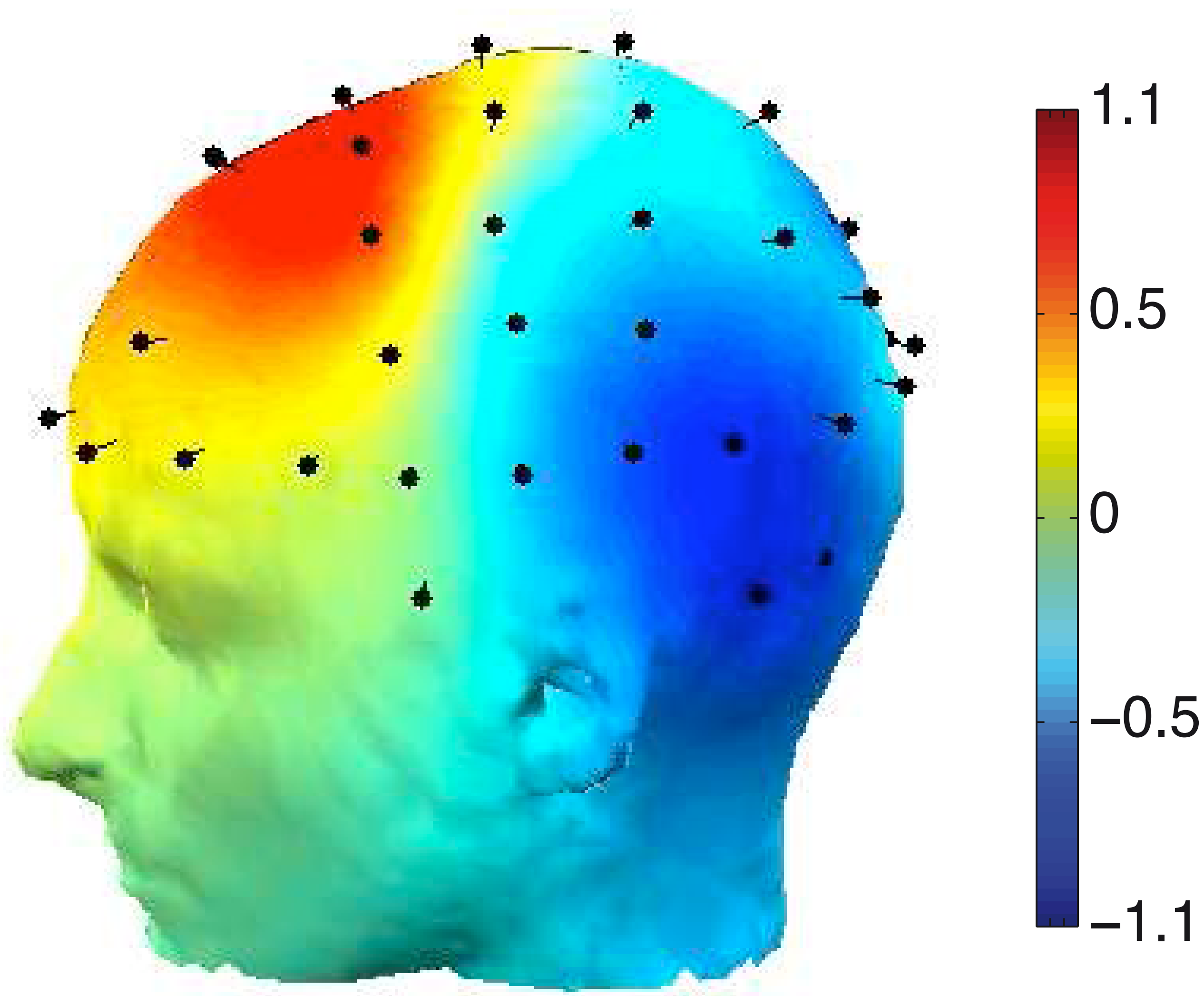
- ≈ 150 to 300 sensors

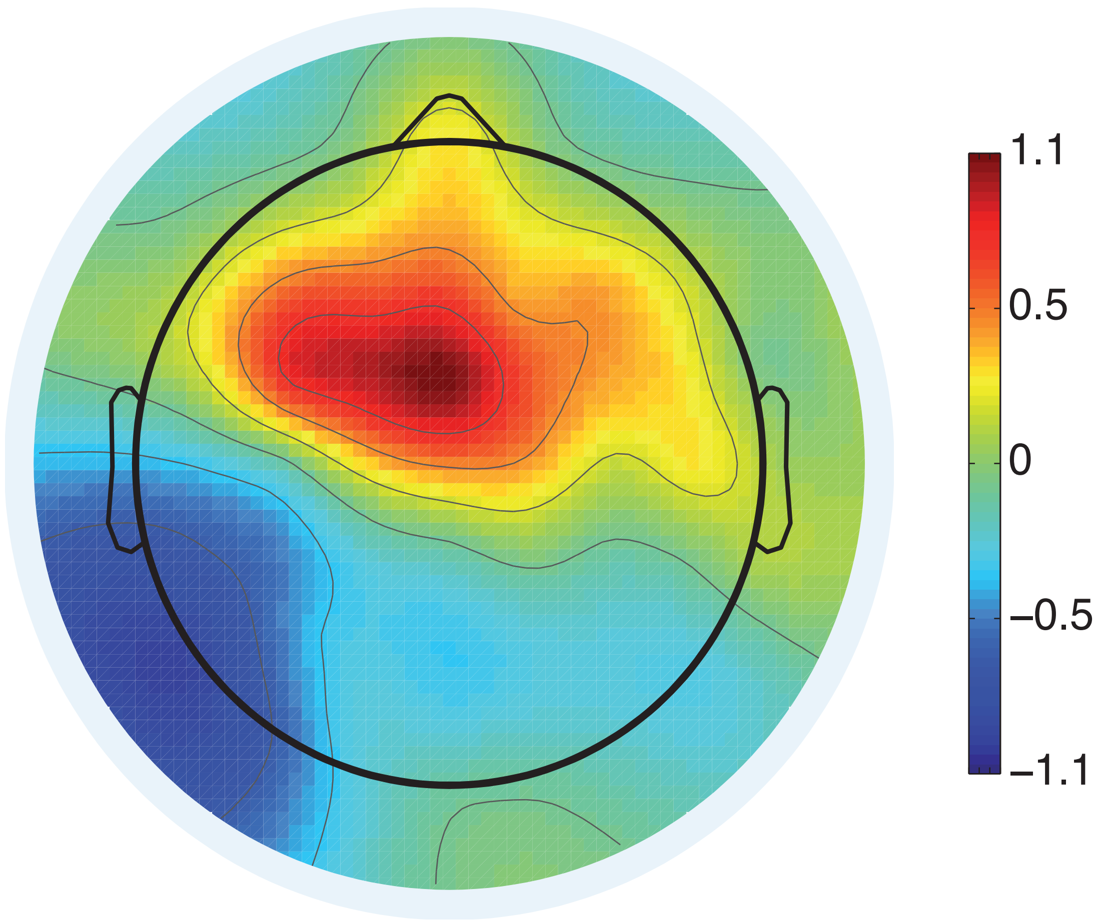
Sampling between 250 and1000 Hz

High temporal resolution but what about spatial resolution?

M/EEG Measurements

At **each time instant** EEG sensors measure a potential field



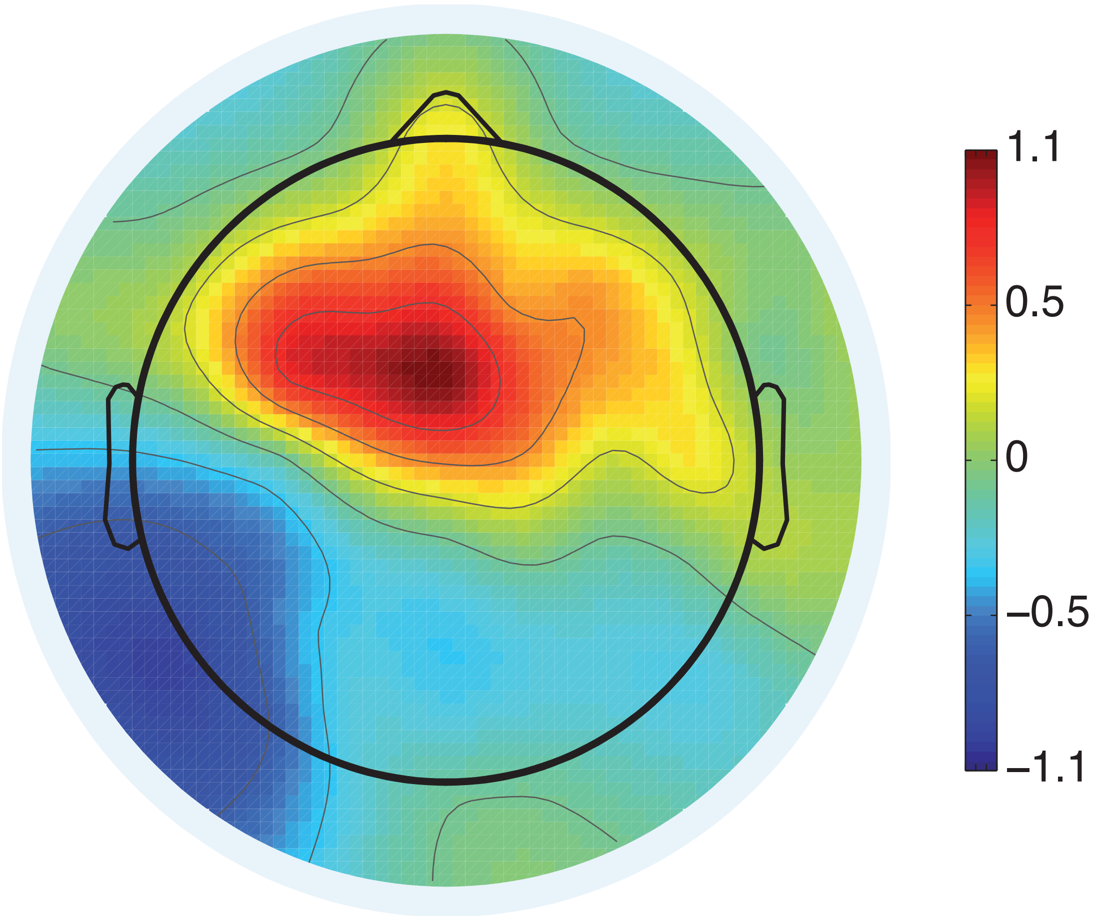


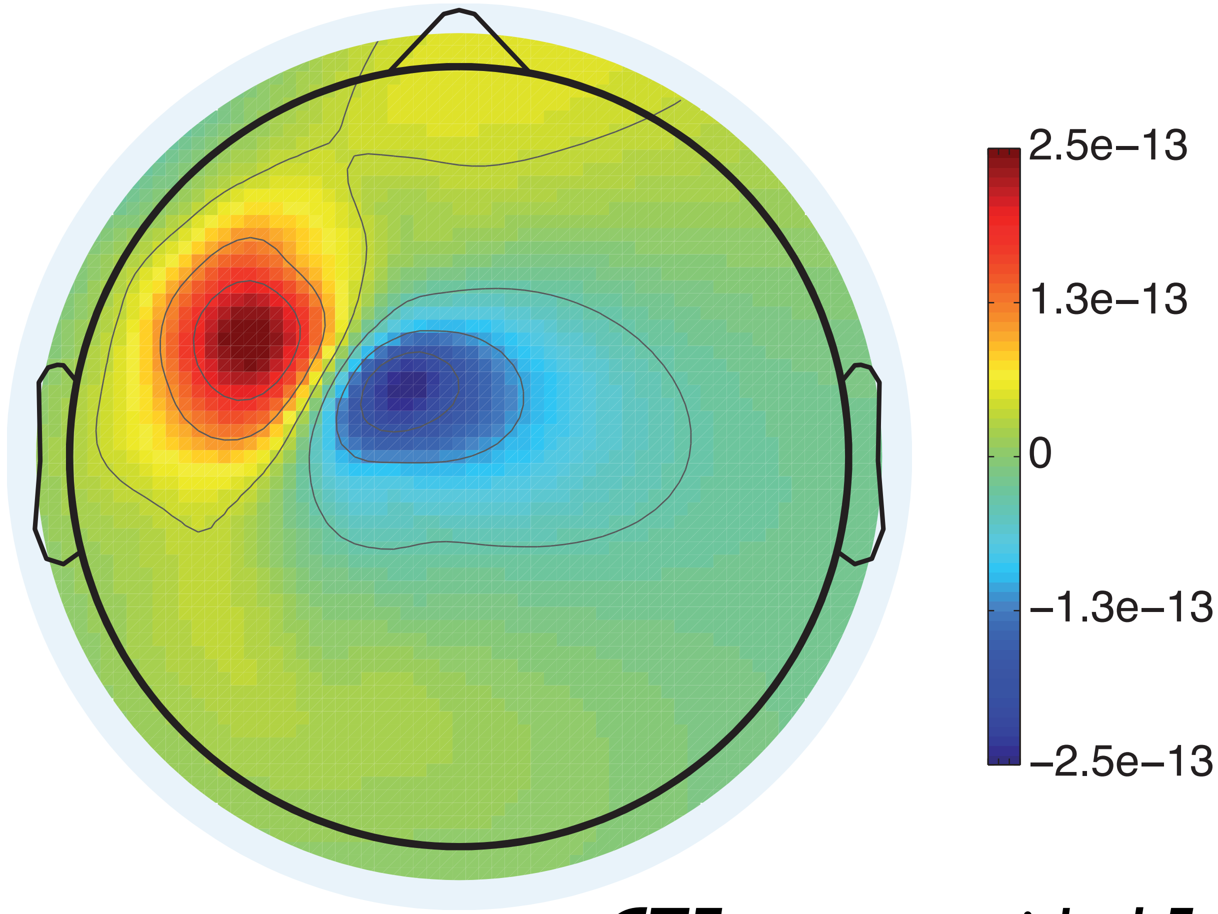
EEG topography

Remark: Such a smooth potential field confirms the presence of **current generators** within the head

M/EEG Measurements

EEG topography vs. MEG topography

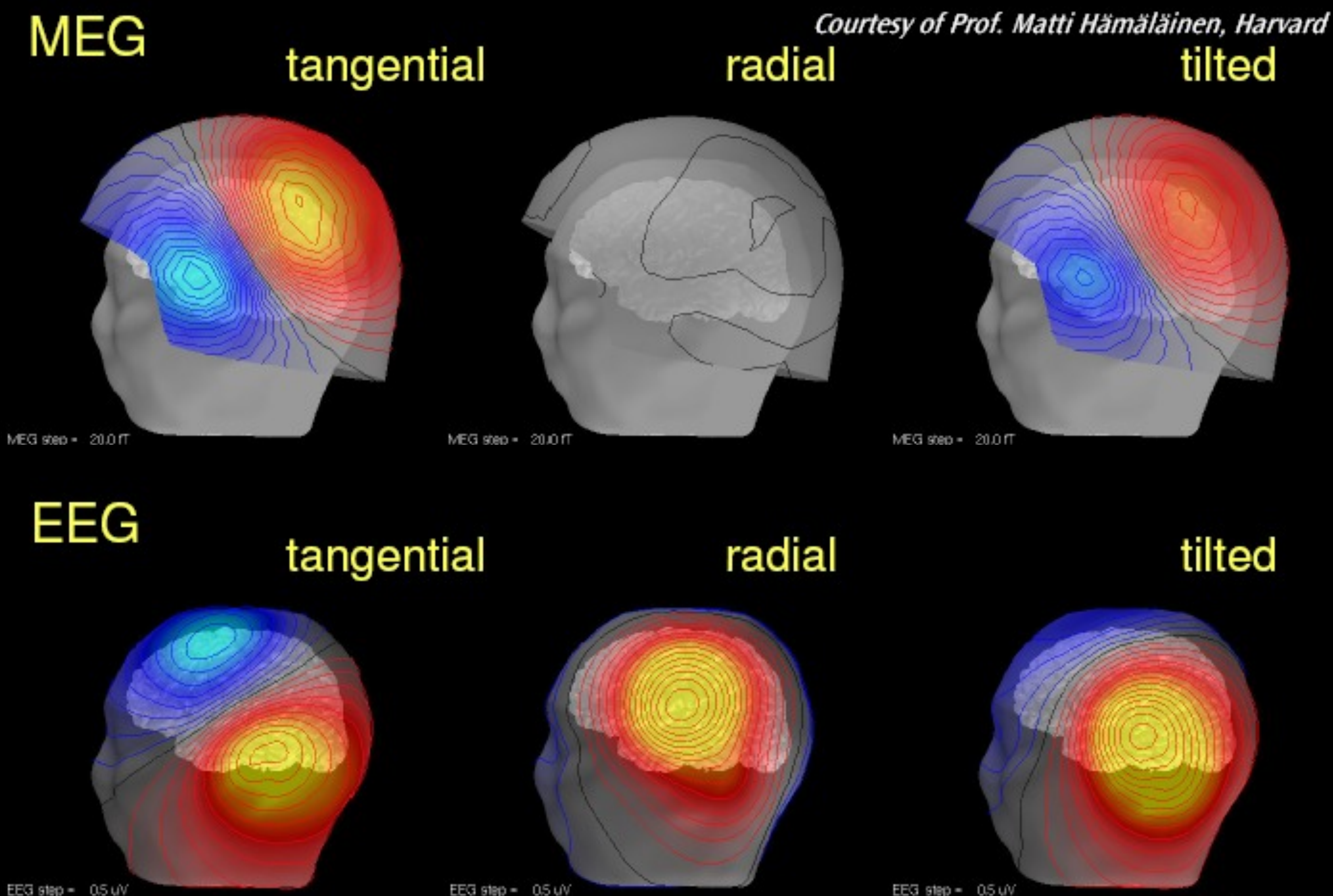




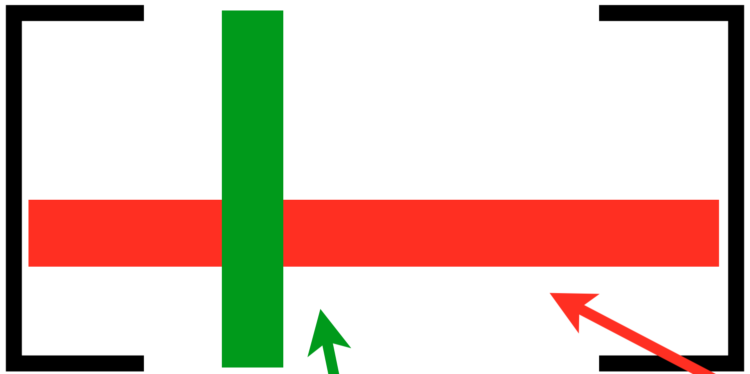
CTF system with 151 axial gradiometers

MEG topography exhibits also a dipolar field but MEG has a **better spatial resolution**

M/EEG Measurements



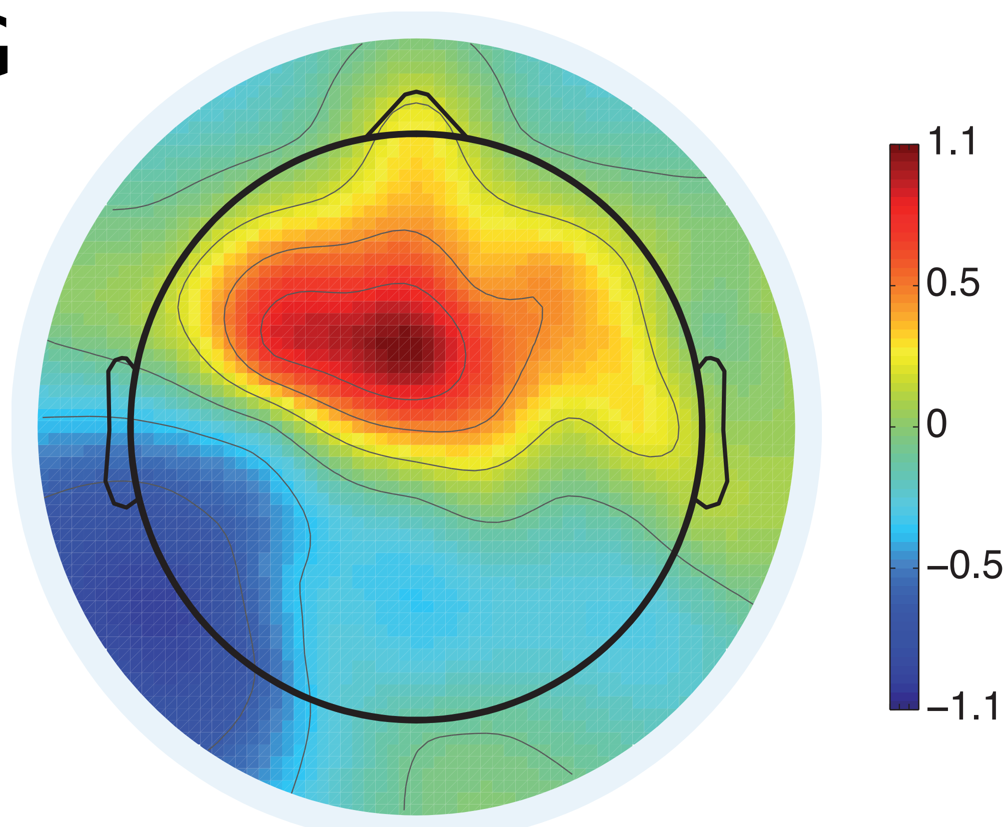
M/EEG Measurements: Notation

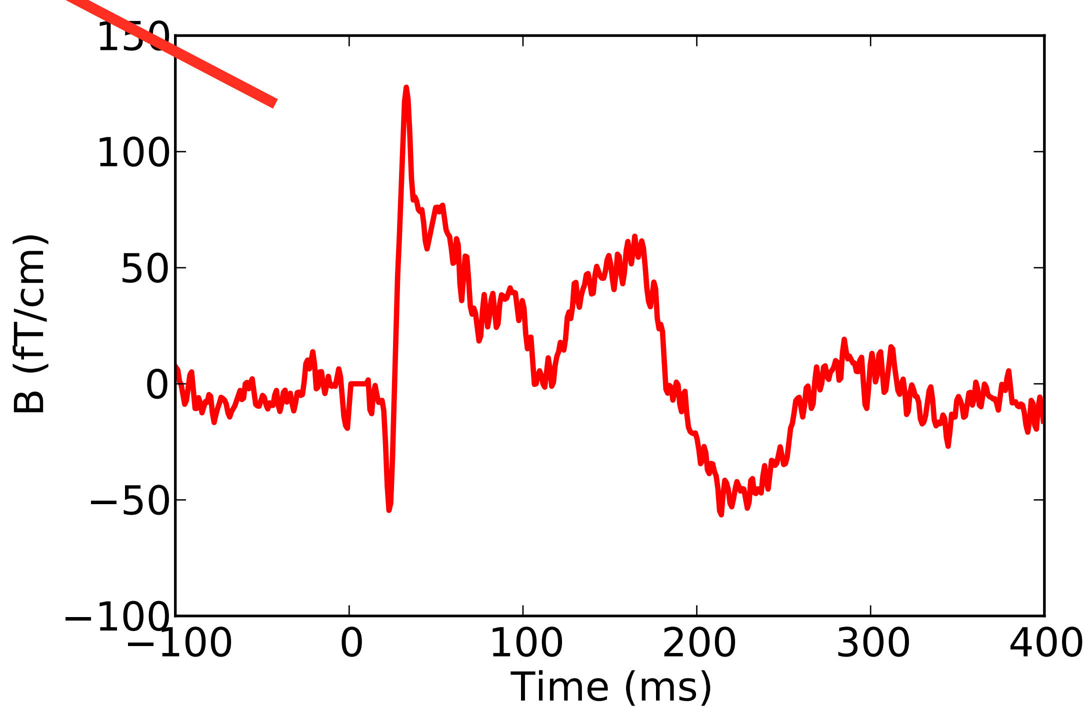
M =  $\in \mathbb{R}^{d_m \times d_t}$

MEG and/or EEG

d_m : Nb of sensors

d_t : Nb of time points





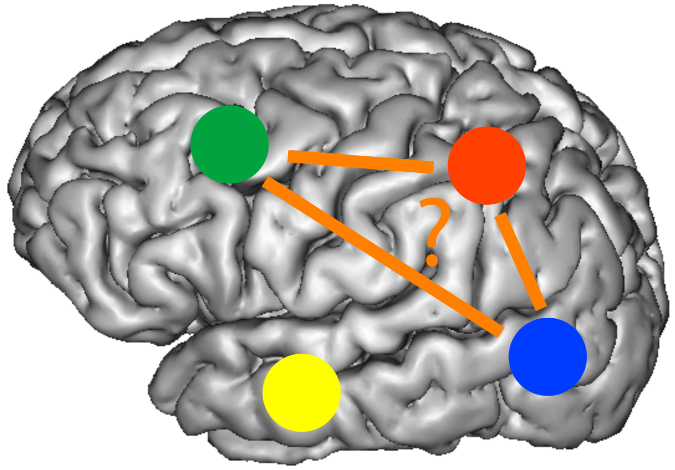
Time

1 column = 1 topography

1 row = 1 time series on 1 sensor

What can you do with M/EEG?

1. **Cognitive studies**
   - **Which areas** are activated during a given cognitive task? **When** are they active? **What is common** in a population of subjects?
2. **Therapy (Epilepsy)**
   - Where is the location of the **origin of epileptic seizures**?
   - Will my **patient be able to talk** if I remove this area of the cortex?
3. **Brain computer interfaces (BCI)**
   - How to extract in **real time** the signal of interest from EEG measurements in order to **control a computer?**



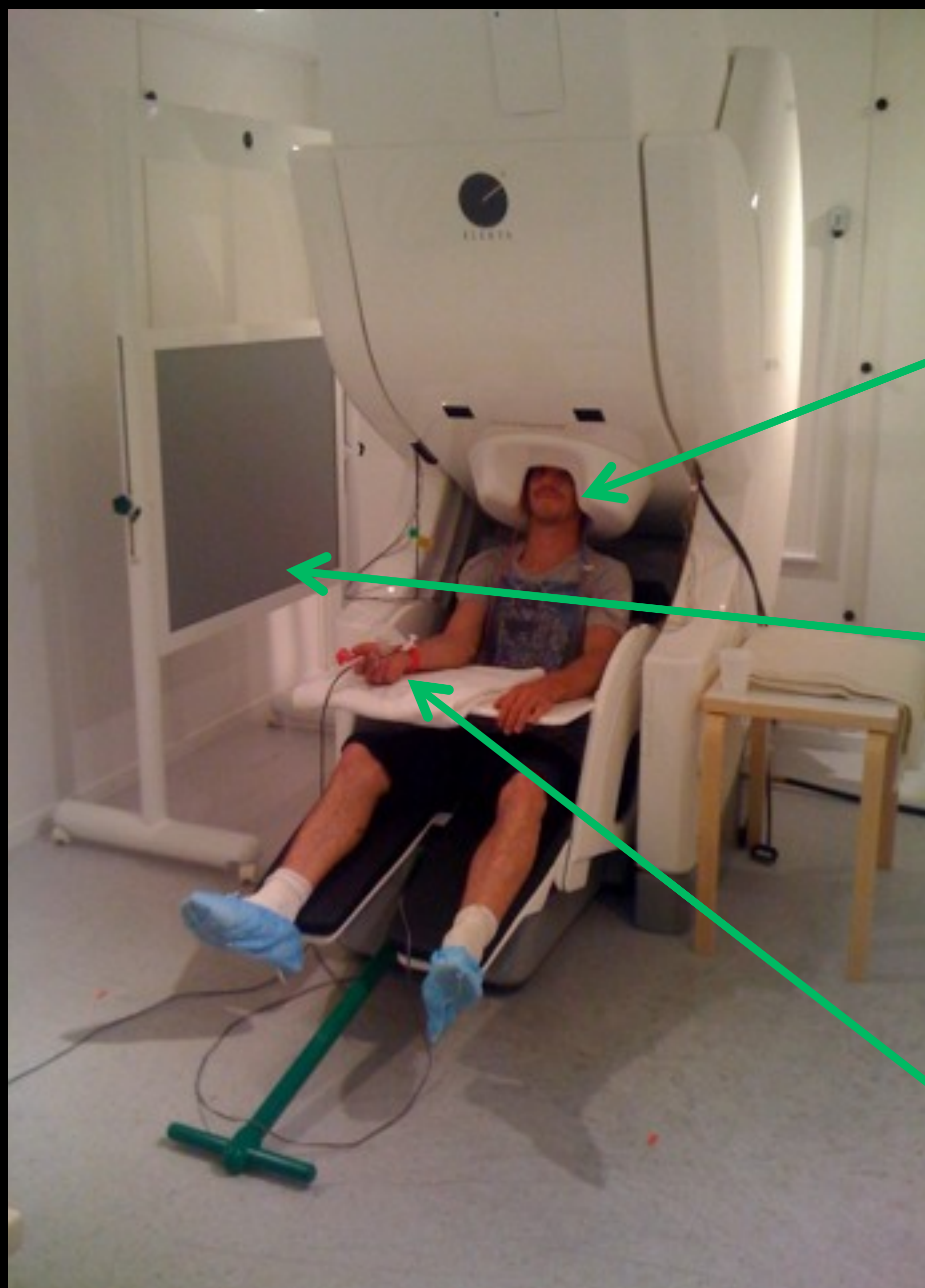
Source: life.com

Source: nih.gov

Data acquisition examples

Also:
Button Pads
Button Gloves
Manual Tapper

Stimulus delivered
by E-Prime,
PsychToolBox, etc.

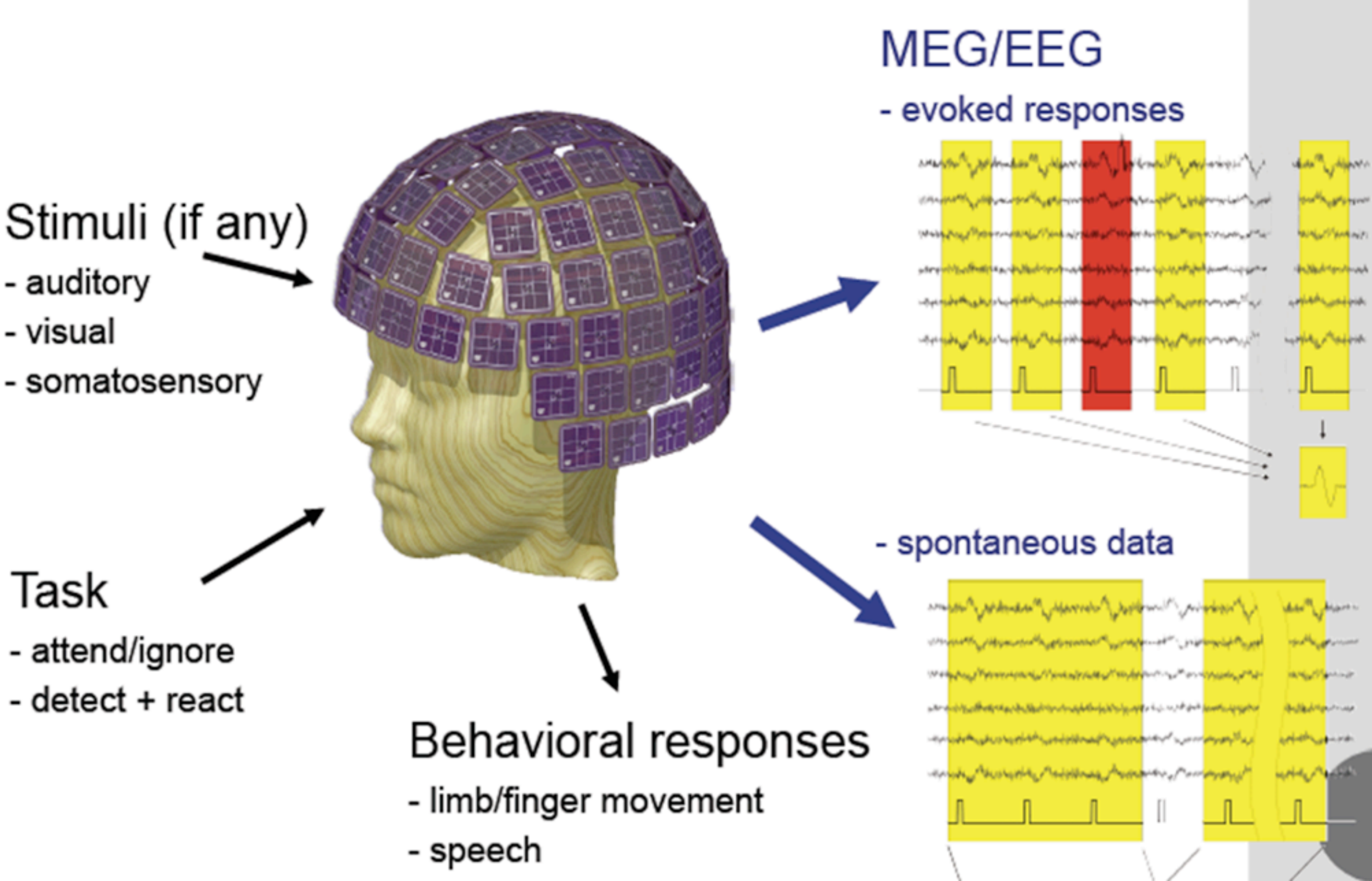


Earphones

Presentation Screen
(moved to front!)

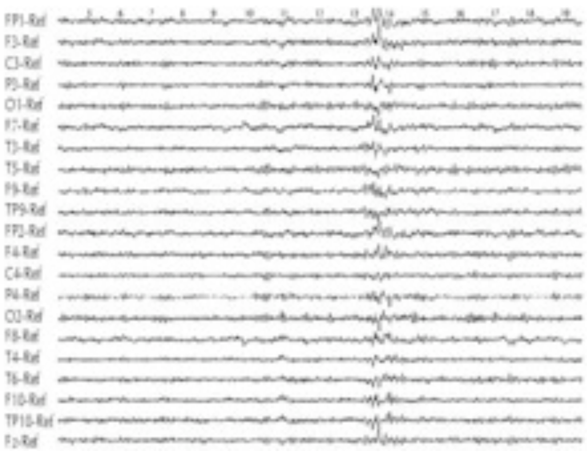
Electrical Stimulator

Data acquisition examples

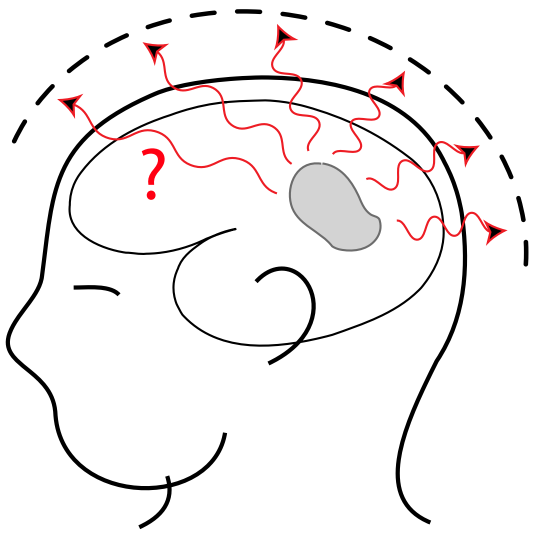


What are the challenges?

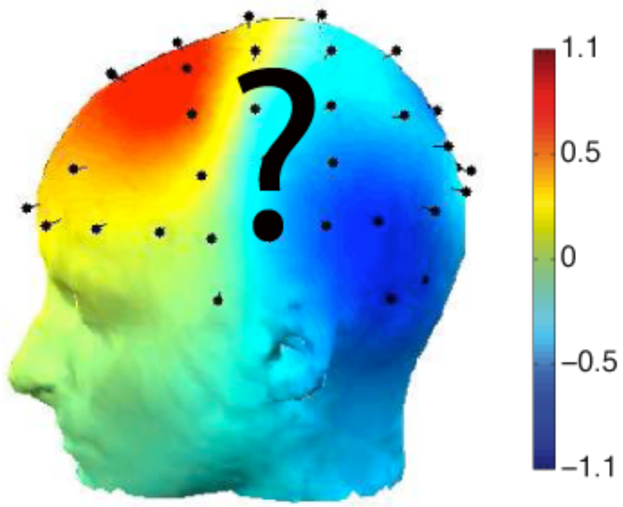
Signal Extraction:
Signal processing, Denoising, Artifact rejection, Single trial analysis.



Forward problem:
Maxwell Equations, Numerical solvers, Finite and Boundary Element Method (BEM & FEM), Image Segmentation and meshing for head modeling.

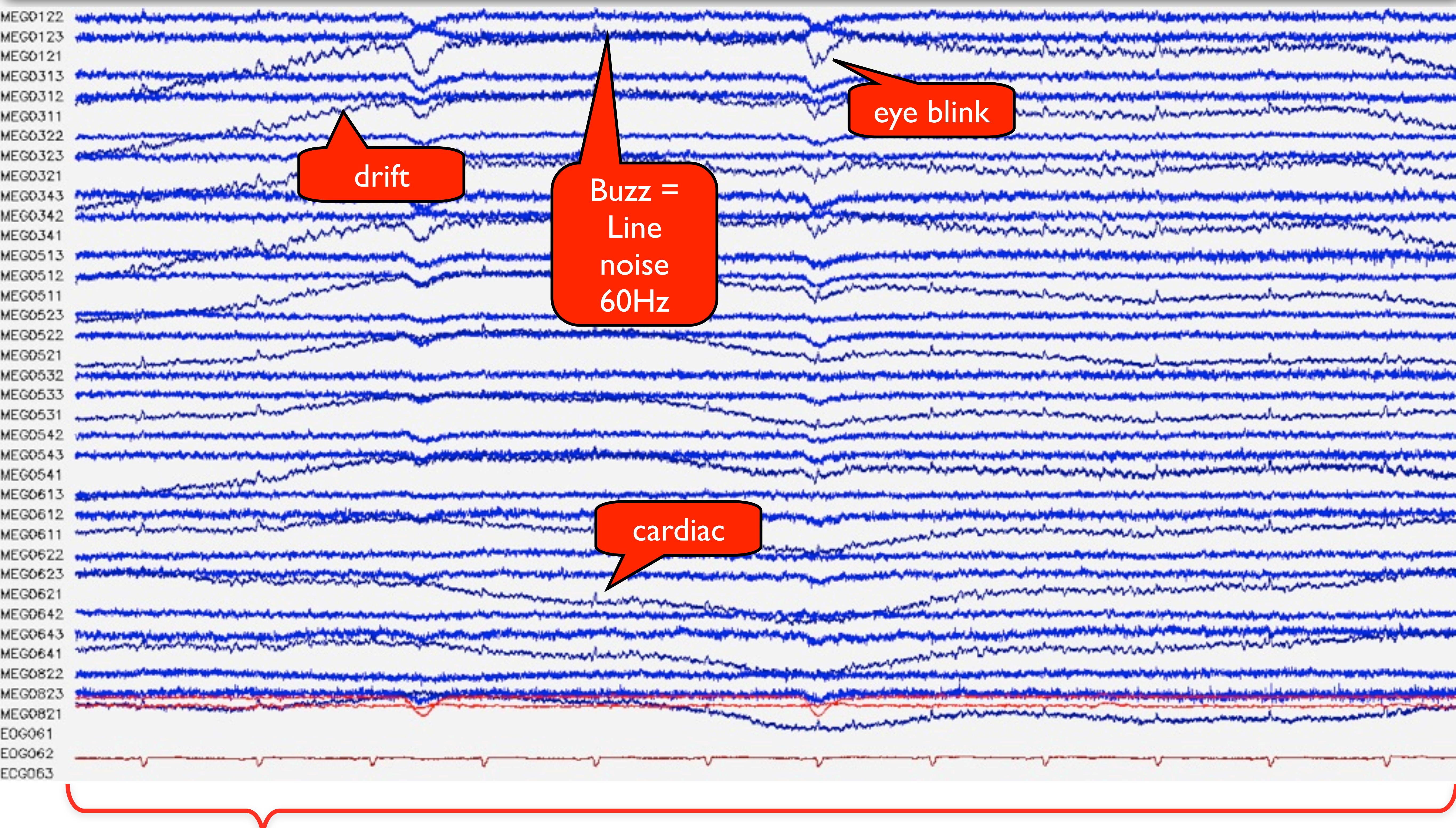


Inverse problem:
Deconvolution problem, Ill-posed problem, Requires efficient solvers to use different priors.



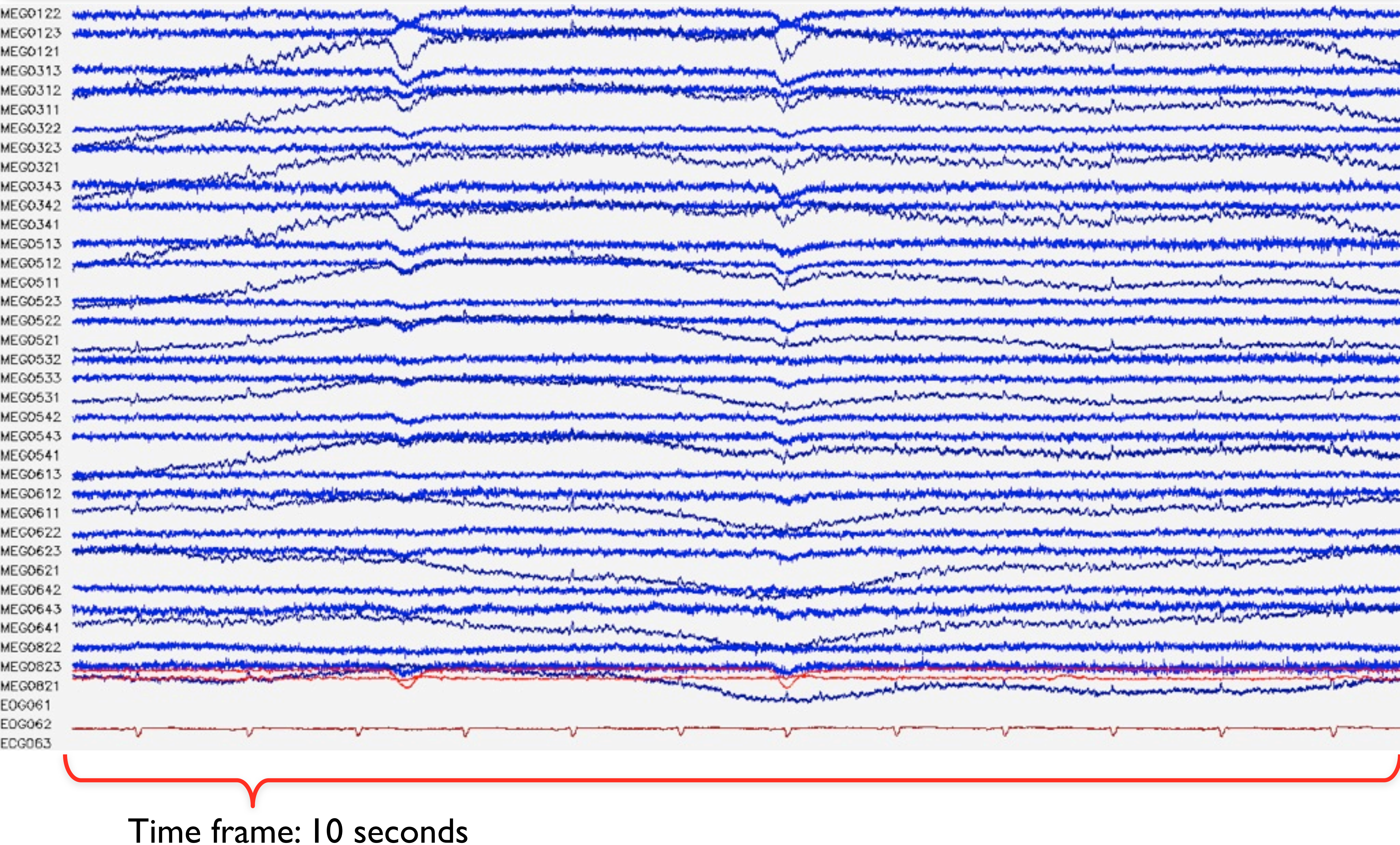
Preprocessing

Artifacts

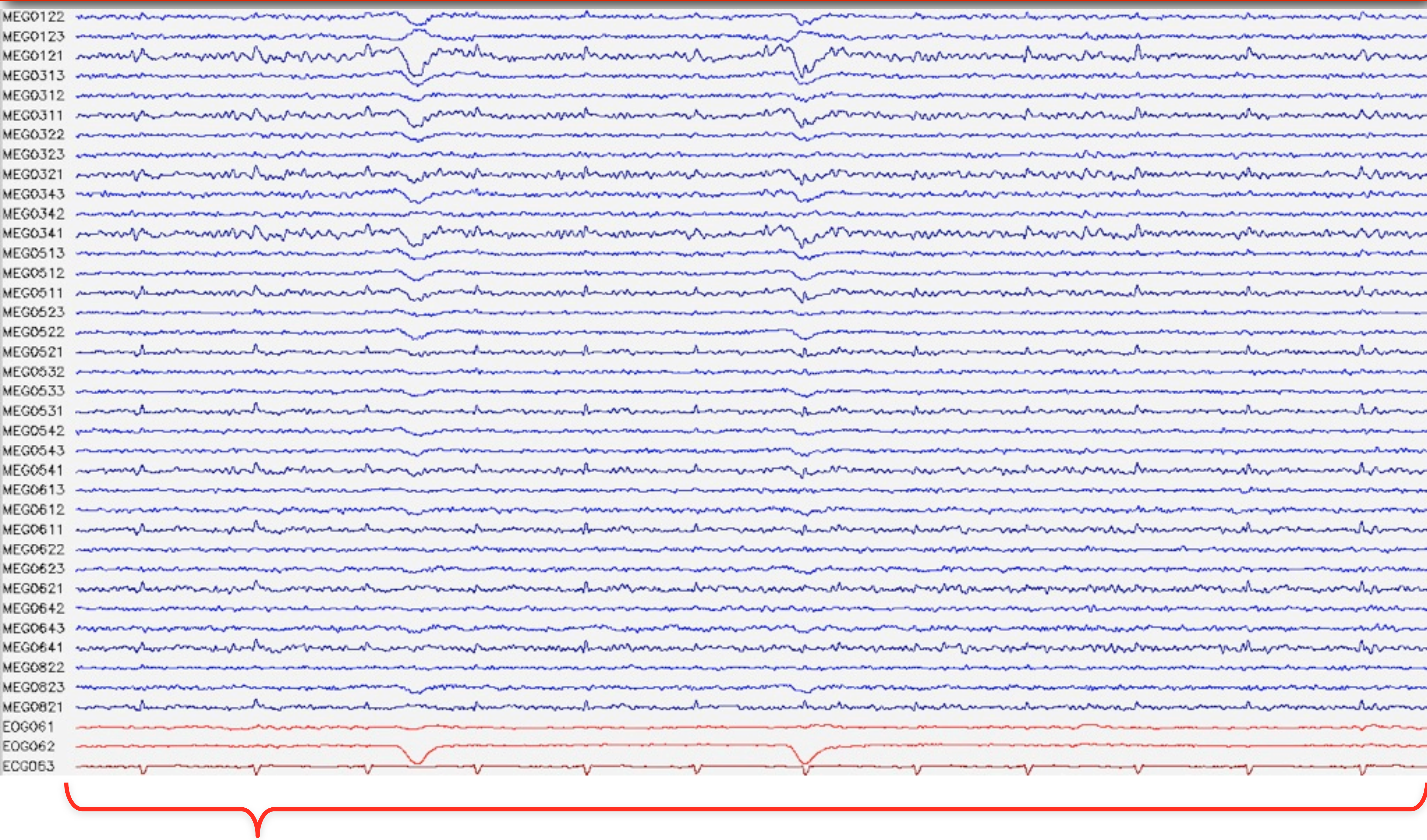


Time frame: 10 seconds

Raw continuous data



Filtered 1-40Hz

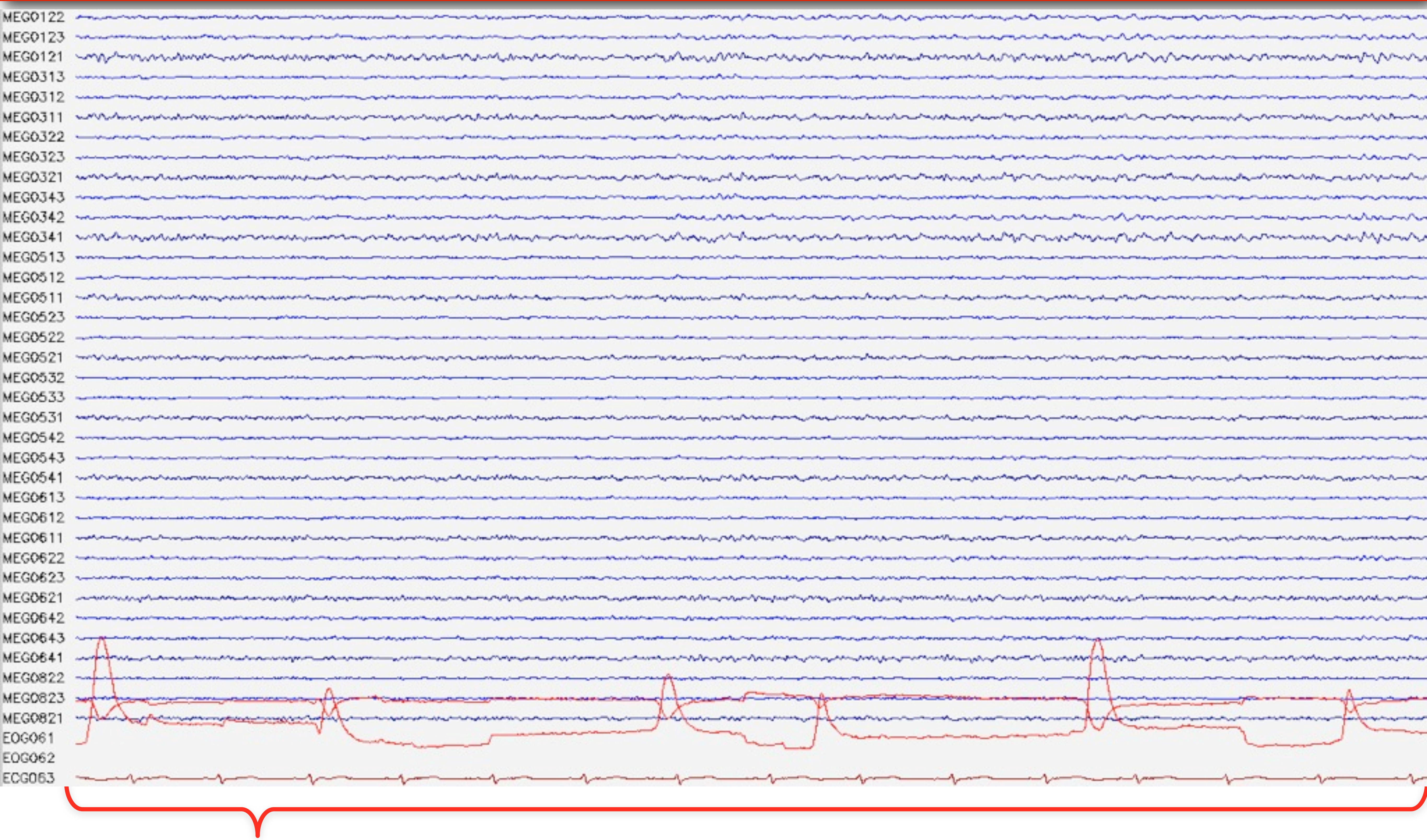


Time frame: 10 seconds

Artifact correction

- SSP - PCA correction
- Signal space projections
- Empty room correction
- Independent component analysis (ICA)
- ...

To get clean data...



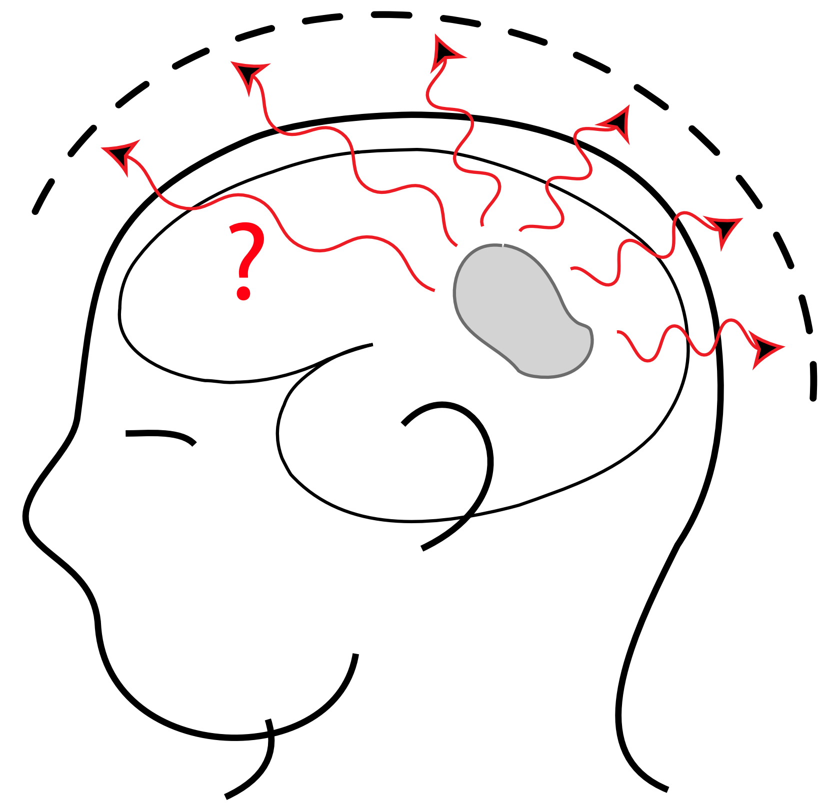
Time frame: 10 seconds

Source localisation with M/EEG:

The **forward** and **inverse** problems

Forward problem: Objective

Predict what is the **Electric Potential** or the **Magnetic Field** **produced by a current generator outside of the head**



How to do it?

- Find from **Maxwell equations** the equations adapted to the problem.
- Define a **model for the current generators** (e.g., sources modeled by equivalent current dipoles).
- **Solve numerically the differential equations** obtained for a real anatomy obtained by MRI.

Maxwell

Maxwell Equations with **quasi-static** approximation

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

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

Maxwell

Maxwell Equations with **quasi-static** approximation

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

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

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

Primary currents ($\vec{J}_p$)

Conduction currents ($\vec{J}_c$)

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

V Electric potential

σ Tissue conductivity

Maxwell

Potential equation (relation between the potential and the sources)**:**

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

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

Poisson Equation

Remark: Relation with Kirchoff's law

Magnetic field equation:

$$\vec{B} = \frac{\mu_0}{4\pi} \int \vec{J}(r') \times \frac{r - r'}{\|r - r'\|^3} dr'$$

Biot and Savart's law

$$\Rightarrow \vec{B} = \vec{B}_0 - \frac{\mu_0}{4\pi} \int \sigma \nabla V \times \frac{r - r'}{\|r - r'\|^3} dr'$$

where $\vec{B}_0 = \frac{\mu_0}{4\pi} \int \vec{J}_p \times \frac{r - r'}{\|r - r'\|^3} dr'$

Observations:

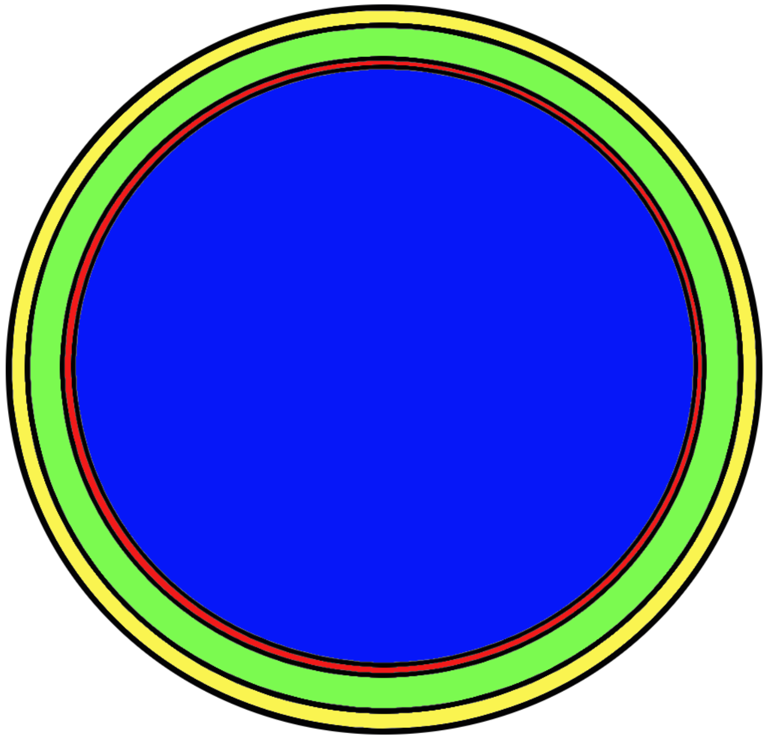
- B is obtained after V
- B decreases in $1/R^2$
- B is due both to primary currents and volume currents

Head models

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

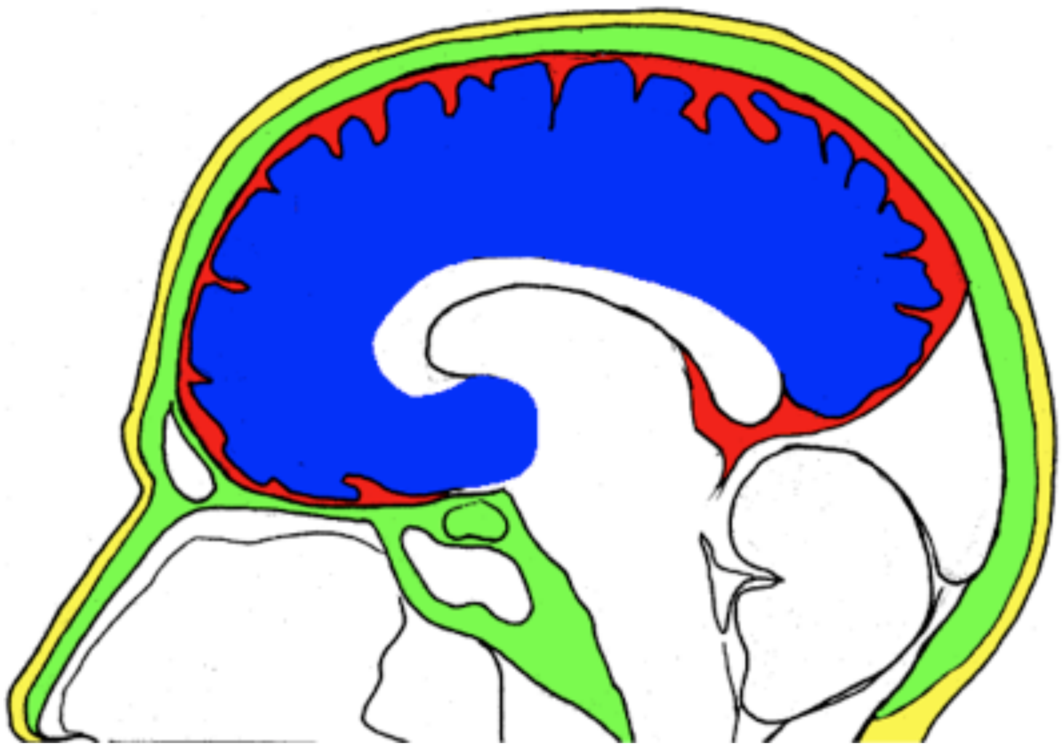
Hypothesis: The conductivities are **piecewise constant**

Sphere models



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

Realistic models



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

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

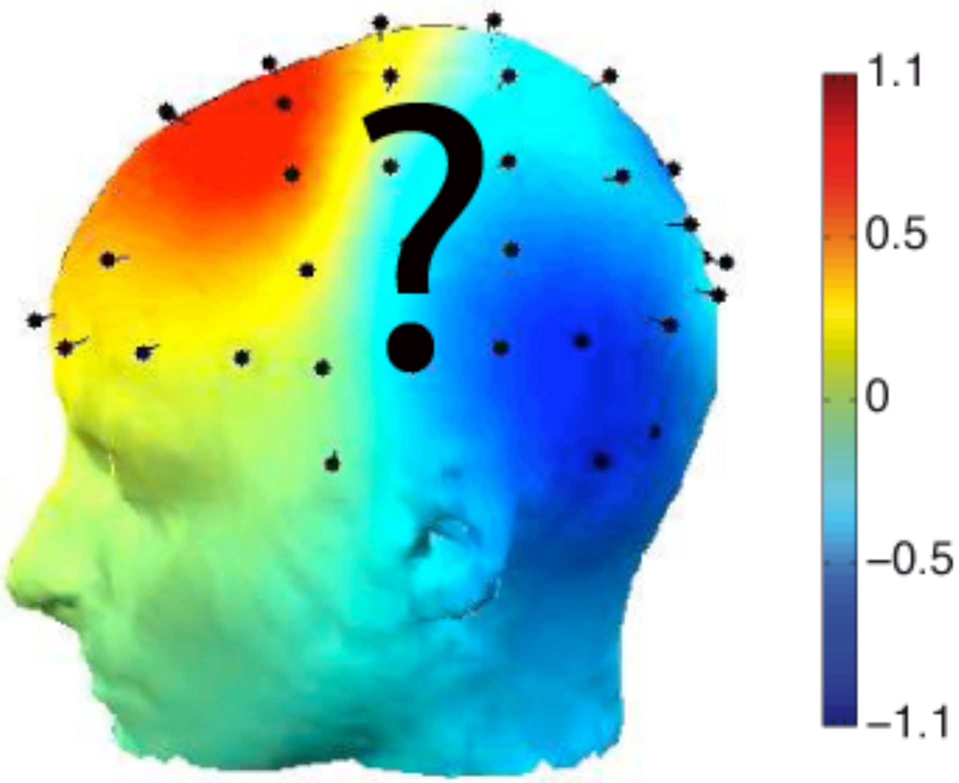
Boundary element method (BEM), i.e., numerical solver with approximate solution.

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

The M/EEG inverse problem

Inverse problem: Objective

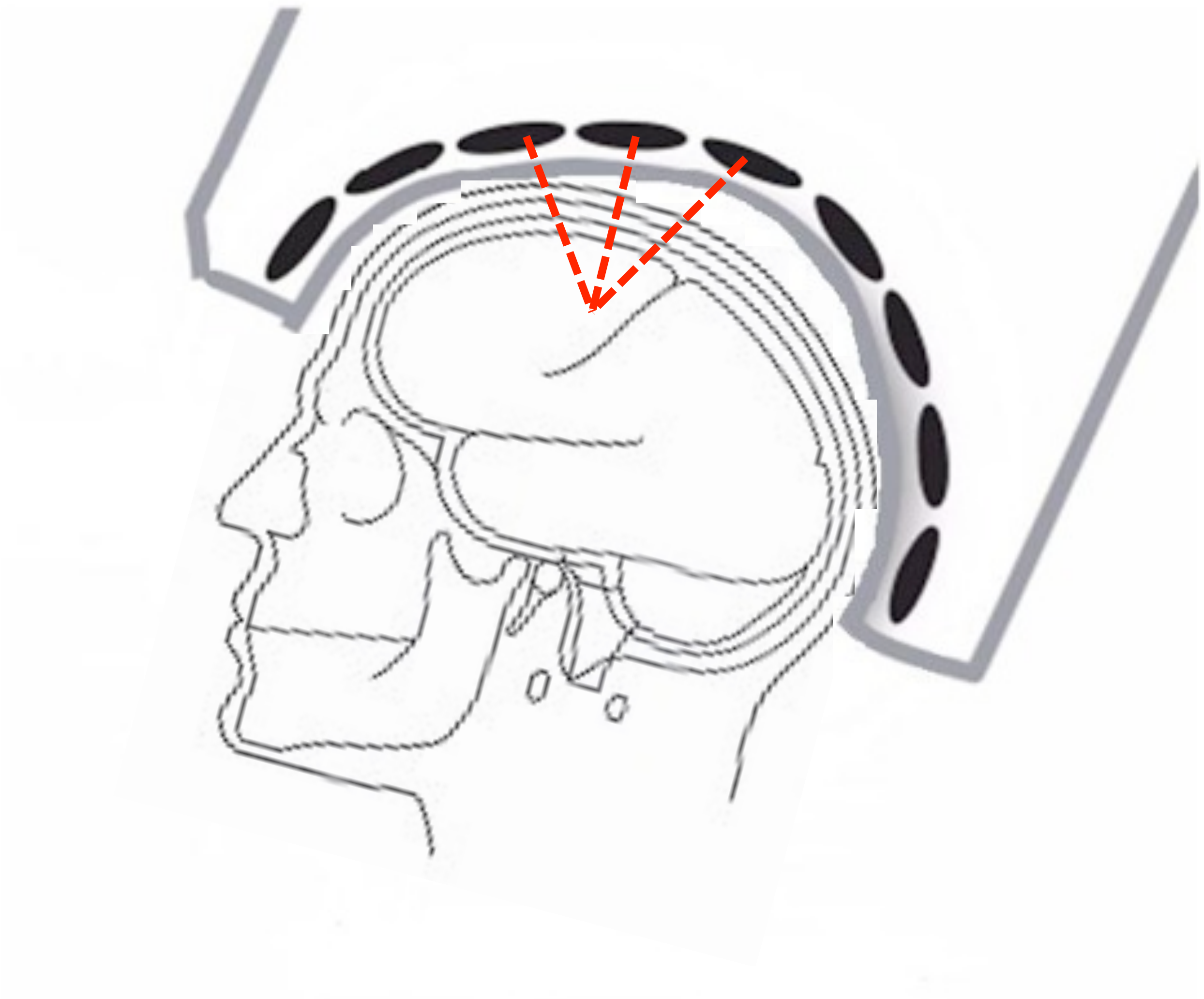
Find the current generators that produced the M/EEG measurements

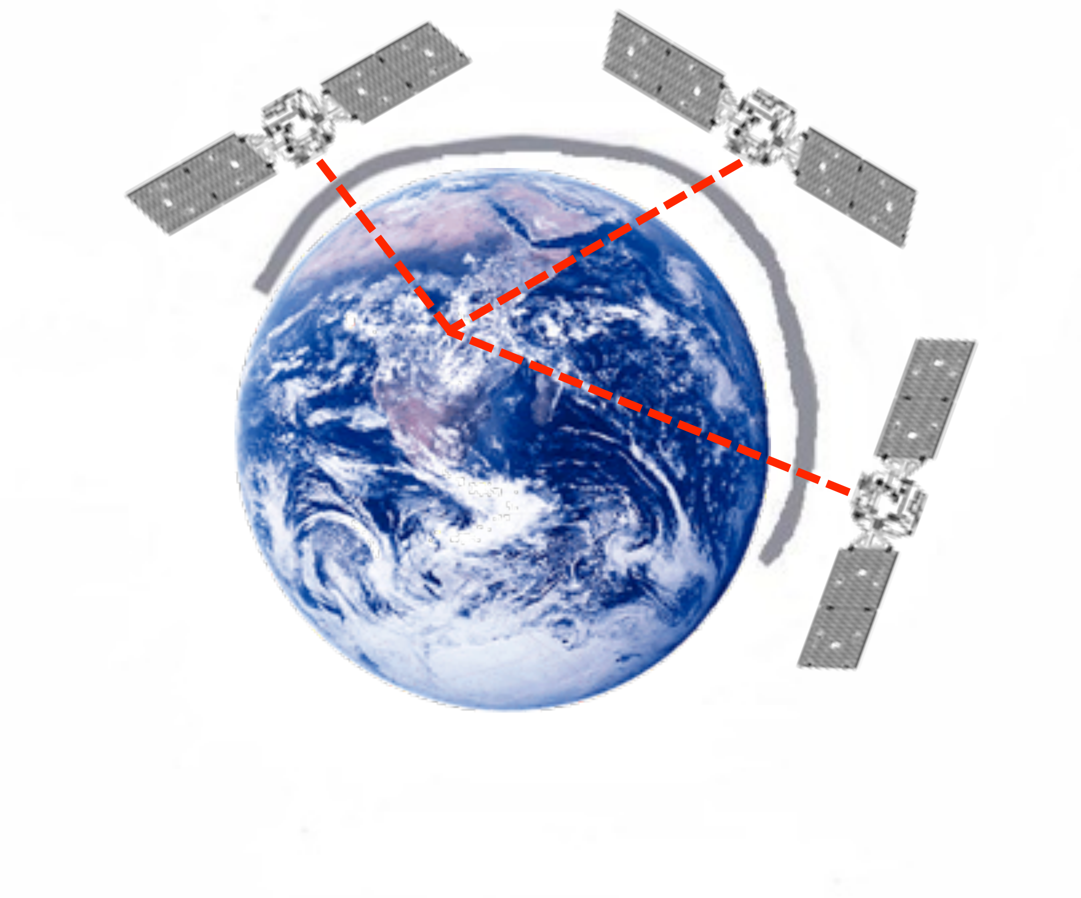


Inverse problem approaches

- Dipole fitting
- Scanning methods
- Distributed models

Dipole fitting

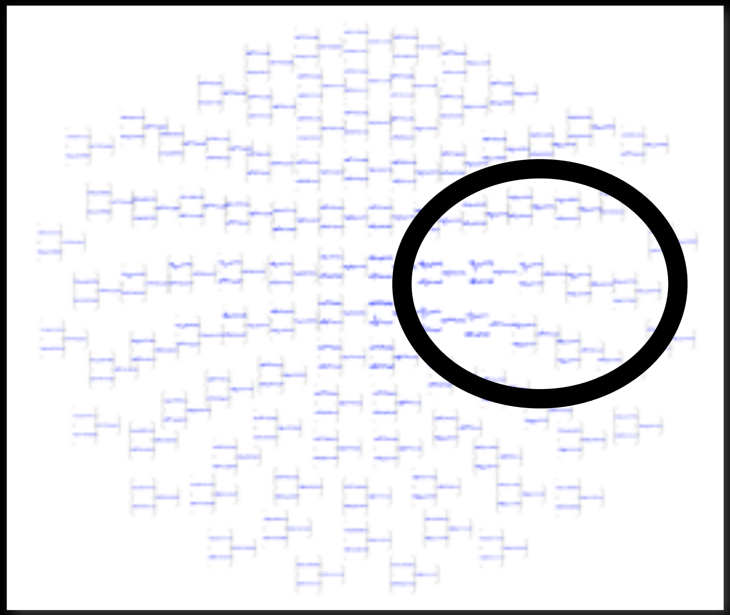




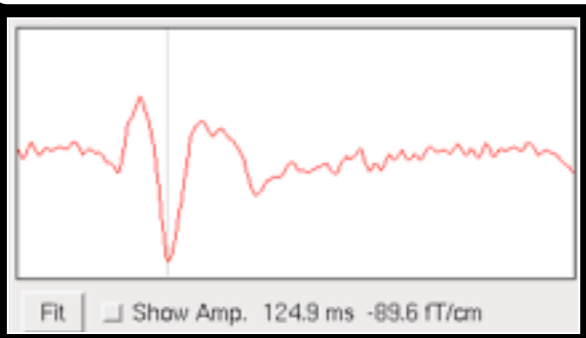
The equivalent of triangulation

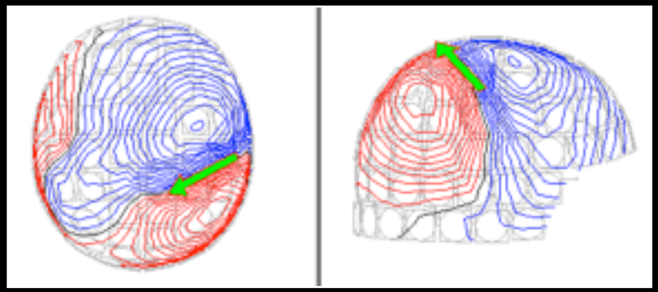
Dipole fitting: procedure

1) Pick subset of sensors w/ peak



2) Pick Time Point; Observe Mag Field



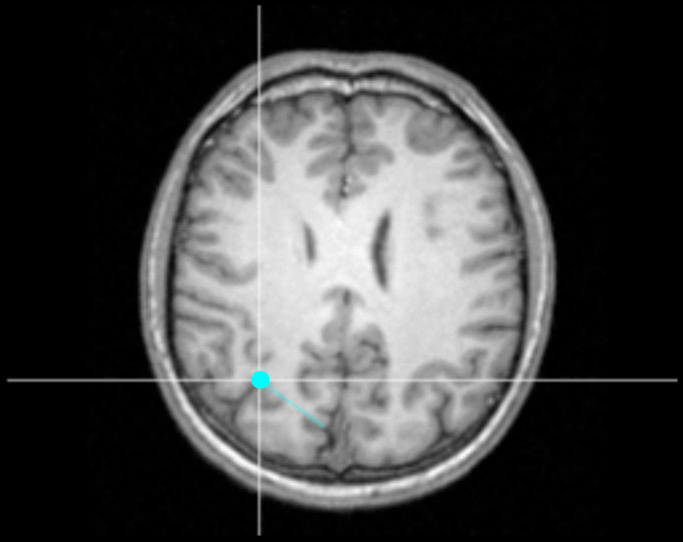


Equivalent Current Dipole Technique

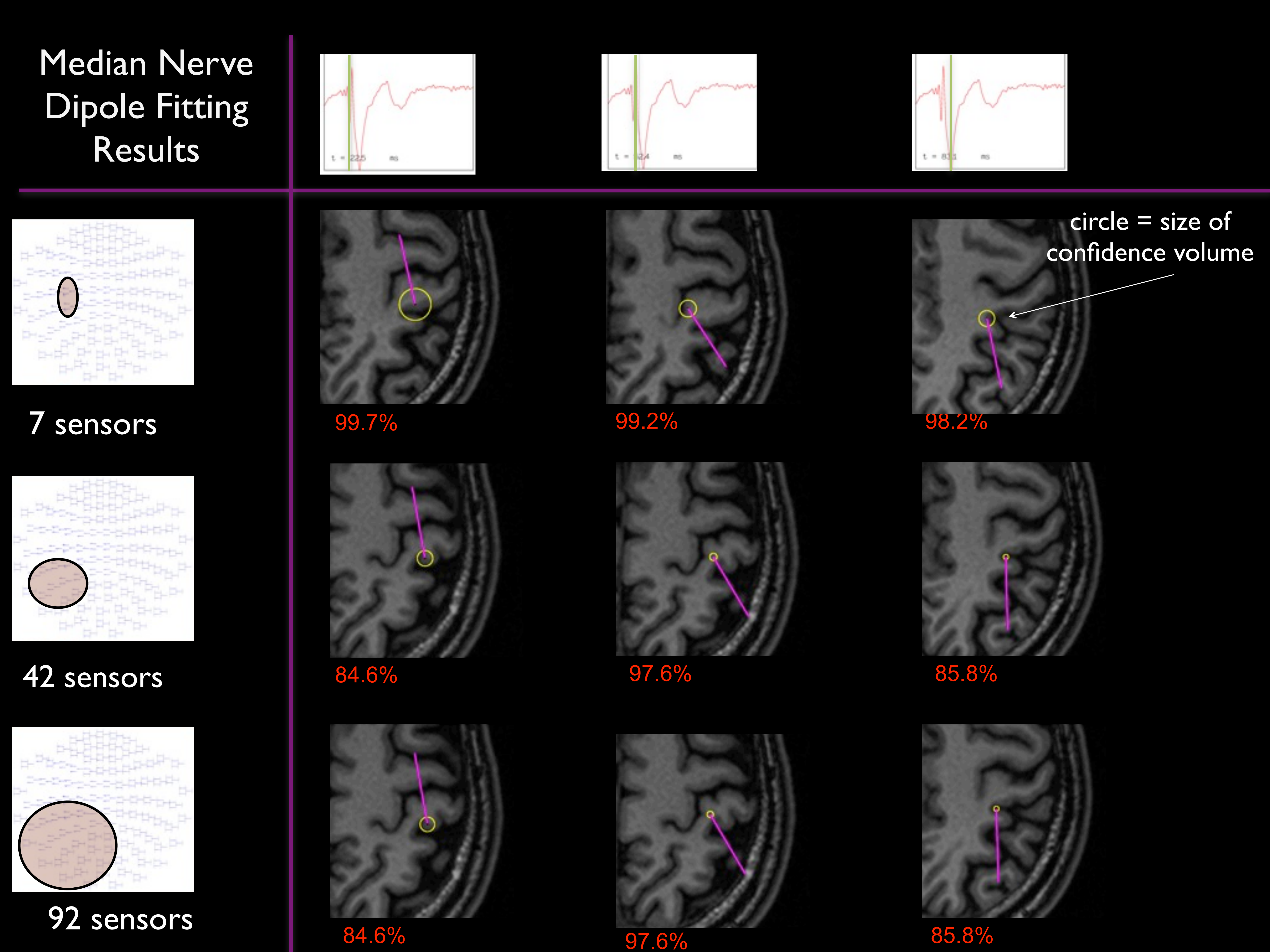
Goodness of Fit
% of activity explained by forward solution based on single dipole

Confidence Volume
volume within which you can be 95% confident that the dipole exists

3) Measures of Quality

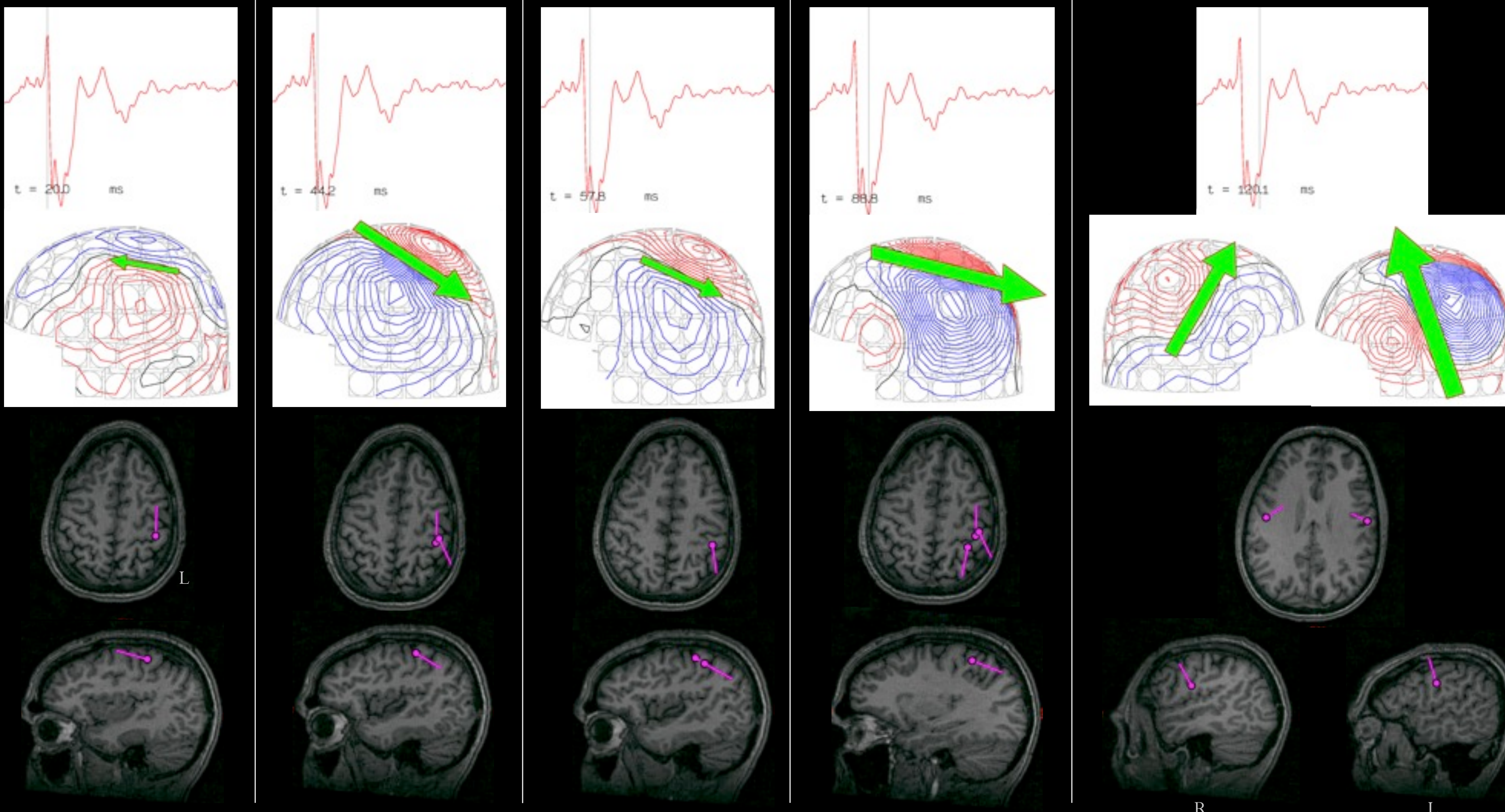


4) Map to MRI


Median Nerve
Dipole Fitting
Results
circle = size of
confidence volume
7 sensors
99.7%
99.2%
98.2%
42 sensors
84.6%
97.6%
85.8%
92 sensors
84.6%
97.6%
85.8%

Time course of SEF

20 ms	44 ms	57 ms	88 ms	120 ms

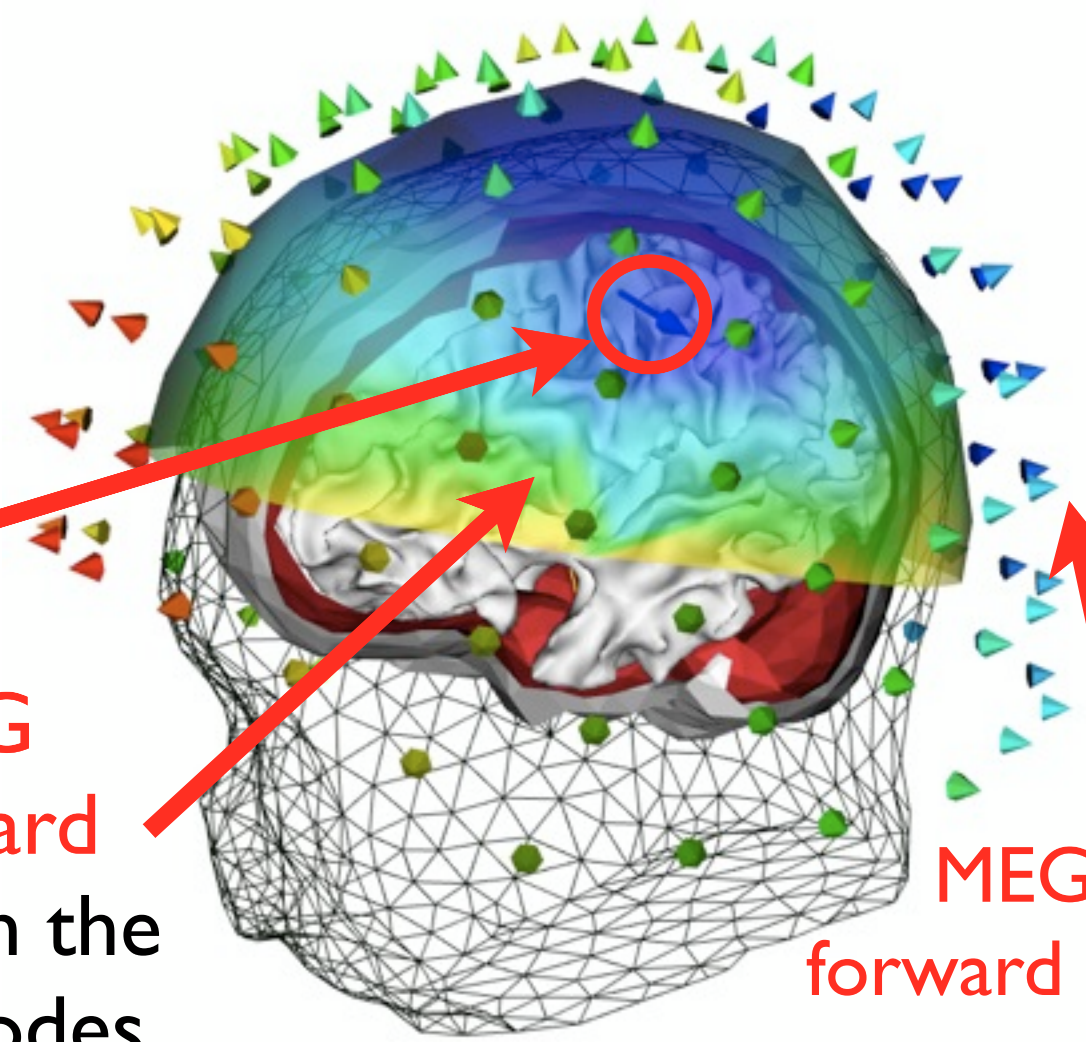


Distributed models

Dipoles sampled over the cortical surface extracted by MRI segmentation

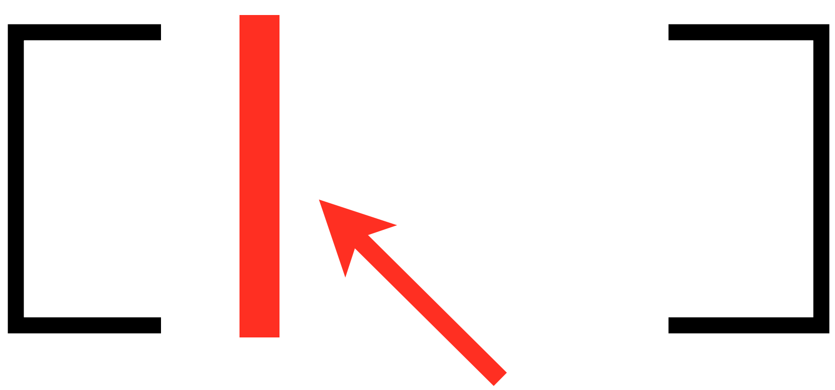
[Dale and Sereno 93]

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



EEG forward field on the electrodes

MEG forward field on sensors

G = 

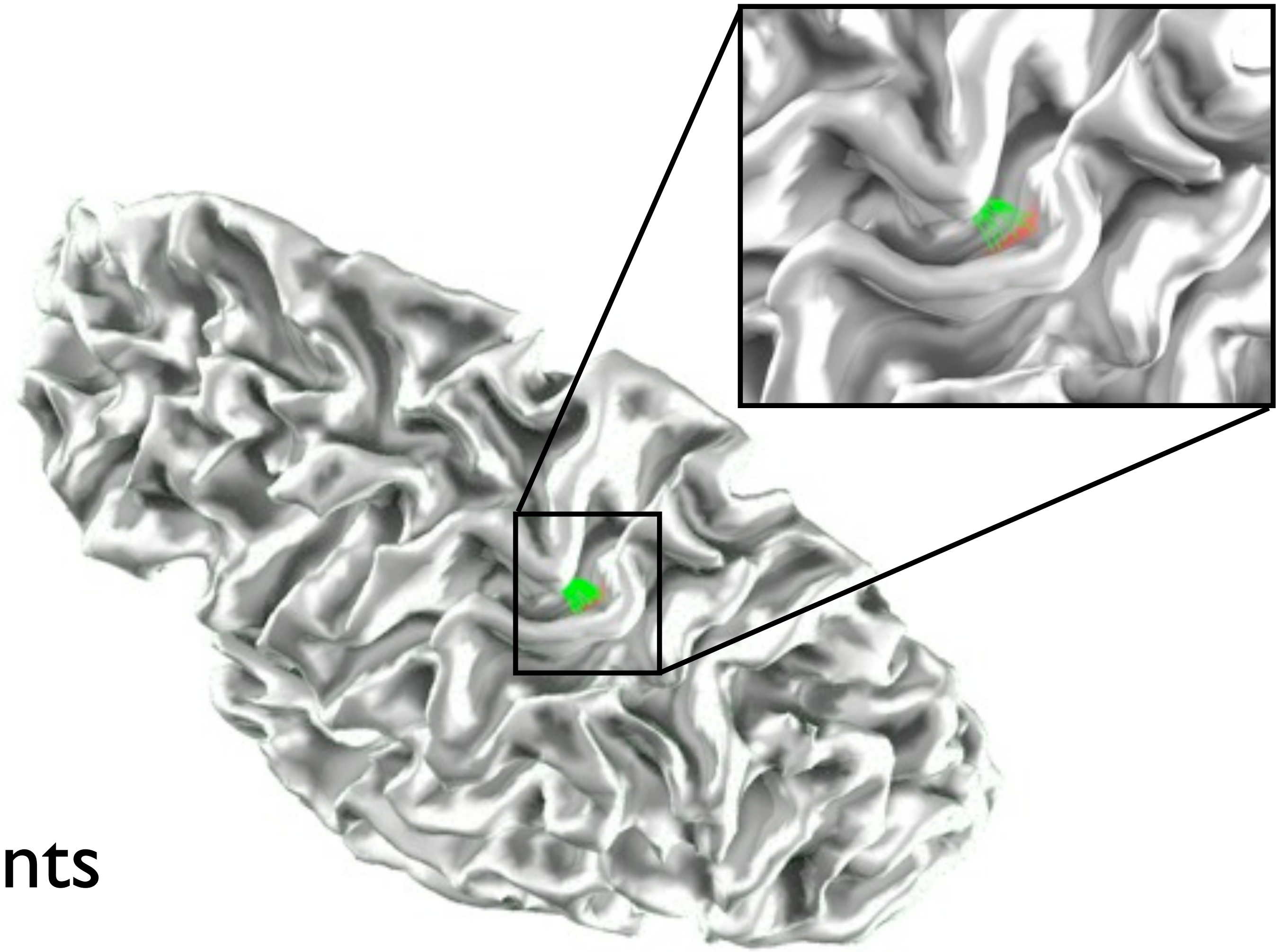
is the lead field matrix obtained by concatenation of the forward fields

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

Scanning methods

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

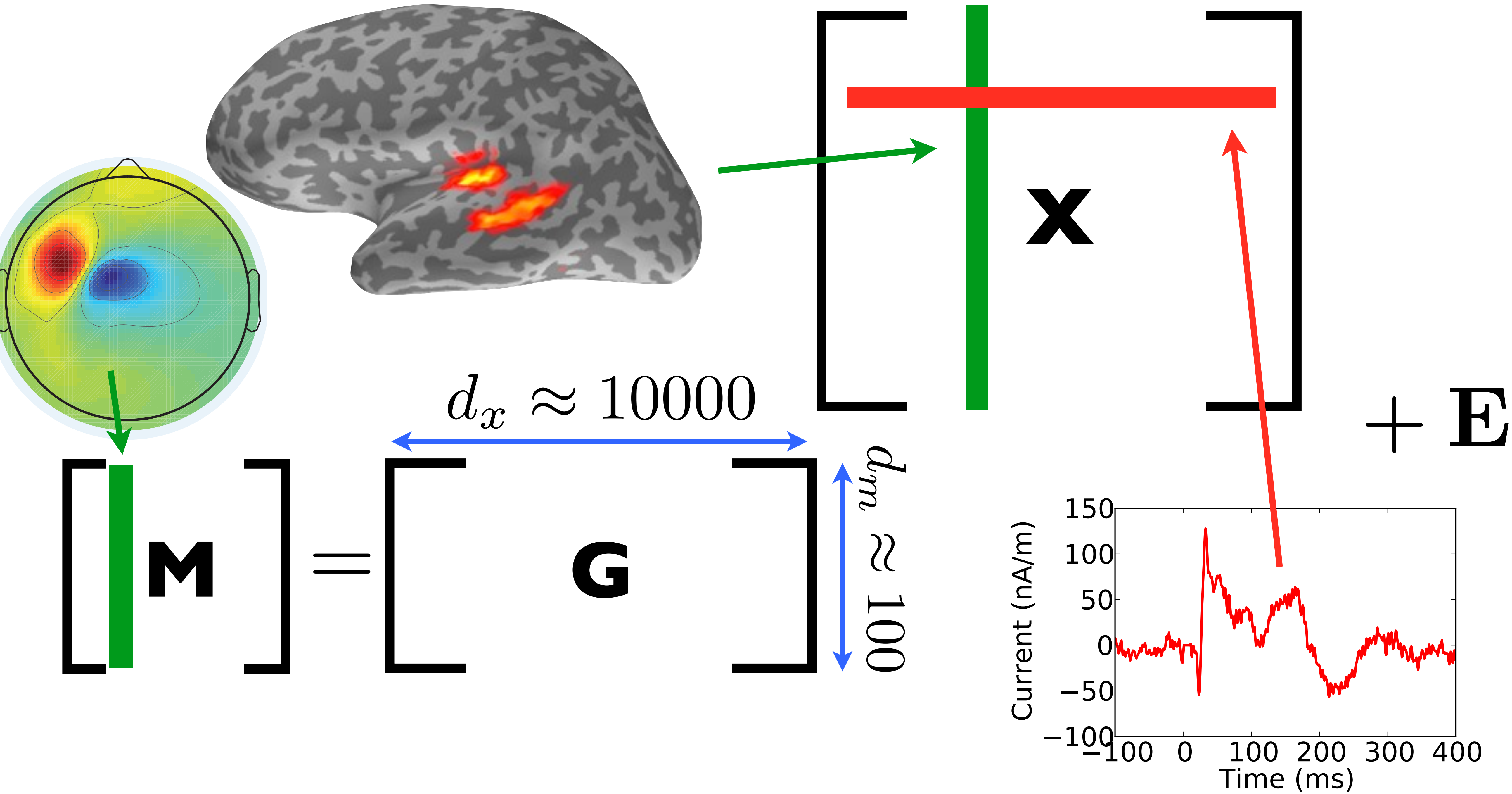
Scanning : One source at a time i.e. one column of **G** at a time

Idea: Find how well it can explain the data while trying to cancel what can come from other sources

Common methods: beamformers (LCMV) and MUSIC

But does not recover **X**...

M = GX+E : An ill-posed problem



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

Inverse problem framework

An optimization problem:

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

Data fit   Prior (penalization)

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

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

$\phi(\mathbf{X})$ Measures the *complexity* of X, it's the prior.

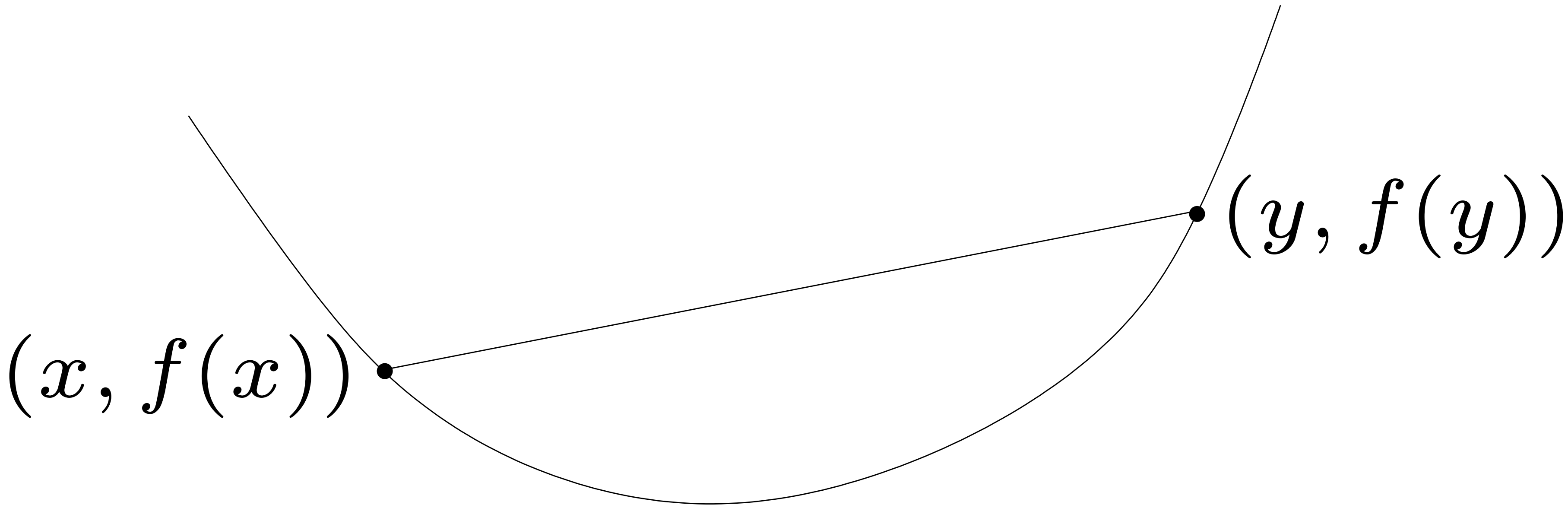
Examples for $\phi(\mathbf{X})$: $\ell_1,\ \ell_2,\ \ell_p$ with $p \geq 1$, entropy $\ldots$

Remark: If $\phi(\mathbf{X})$ is strictly convex we have a unique minimizer (sufficient but not a necessary condition)

Definition: Convex function

$f : \mathbb{R}^n \rightarrow \mathbb{R}$

is convex iff



$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

for all $0 \leq \theta \leq 1$

is strictly convex iff

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$$

for all $0 < \theta < 1$

Remark: The presentation is restricted to functions defined on $\mathbb{R}^n$

Inverse problem

Optimization problem:

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

Data fit

convex + convex

=

convex

- Data fit is **quadratic** hence **convex**
- If $\phi(\mathrm{X})$ is **convex**, then it's a **convex optimization problem**

Smooth or non-smooth

- **Smooth:**
  - L2 (regularized Least-squares, Tikhonov)

$$\phi(\mathbf{X}) = \|\mathbf{X}\|_2^2 = \sum_{i,j} x_{ij}^2$$

  - Entropy based methods
  - etc.
- **Non-smooth:**
  - L1

$$\phi(\mathbf{X}) = \|\mathbf{X}\|_1 = \sum_{i,j} |x_{ij}|$$

  - Total-Variation

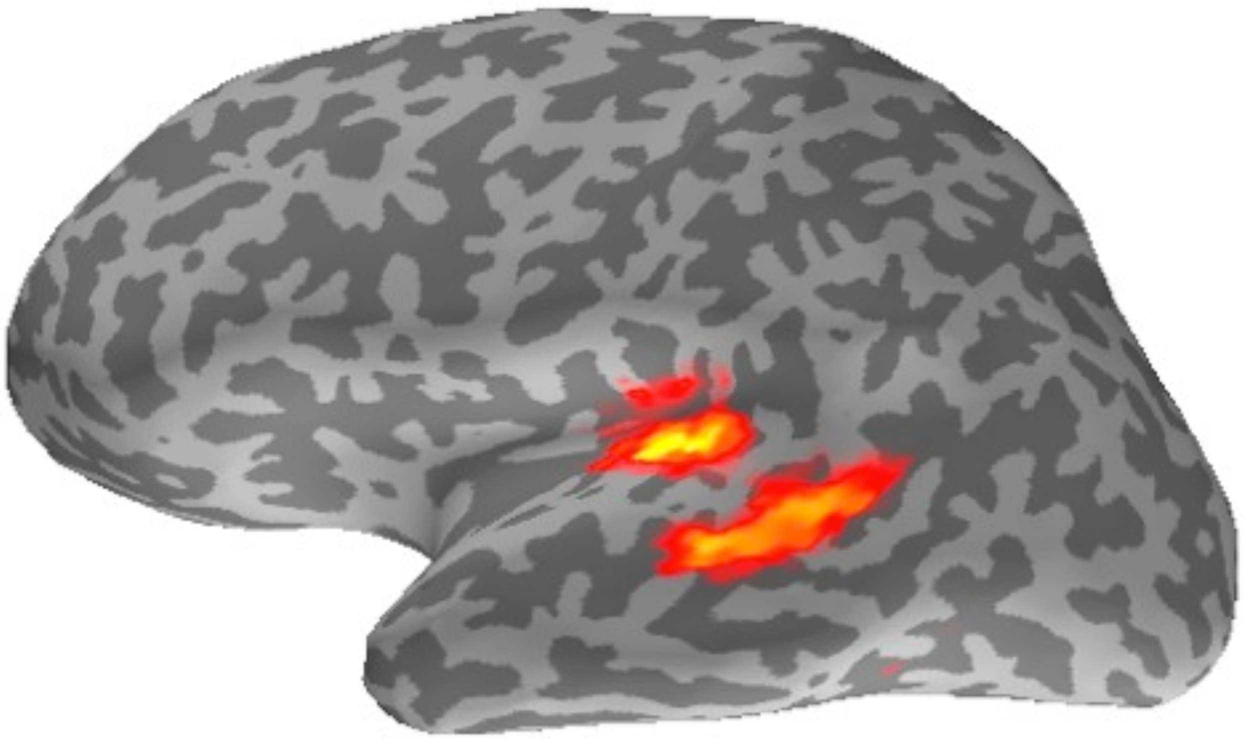
$$\phi(\mathbf{X}) = TV(\mathbf{X}) = \|\nabla_{surf}\mathbf{X}\|_1$$

  - etc.

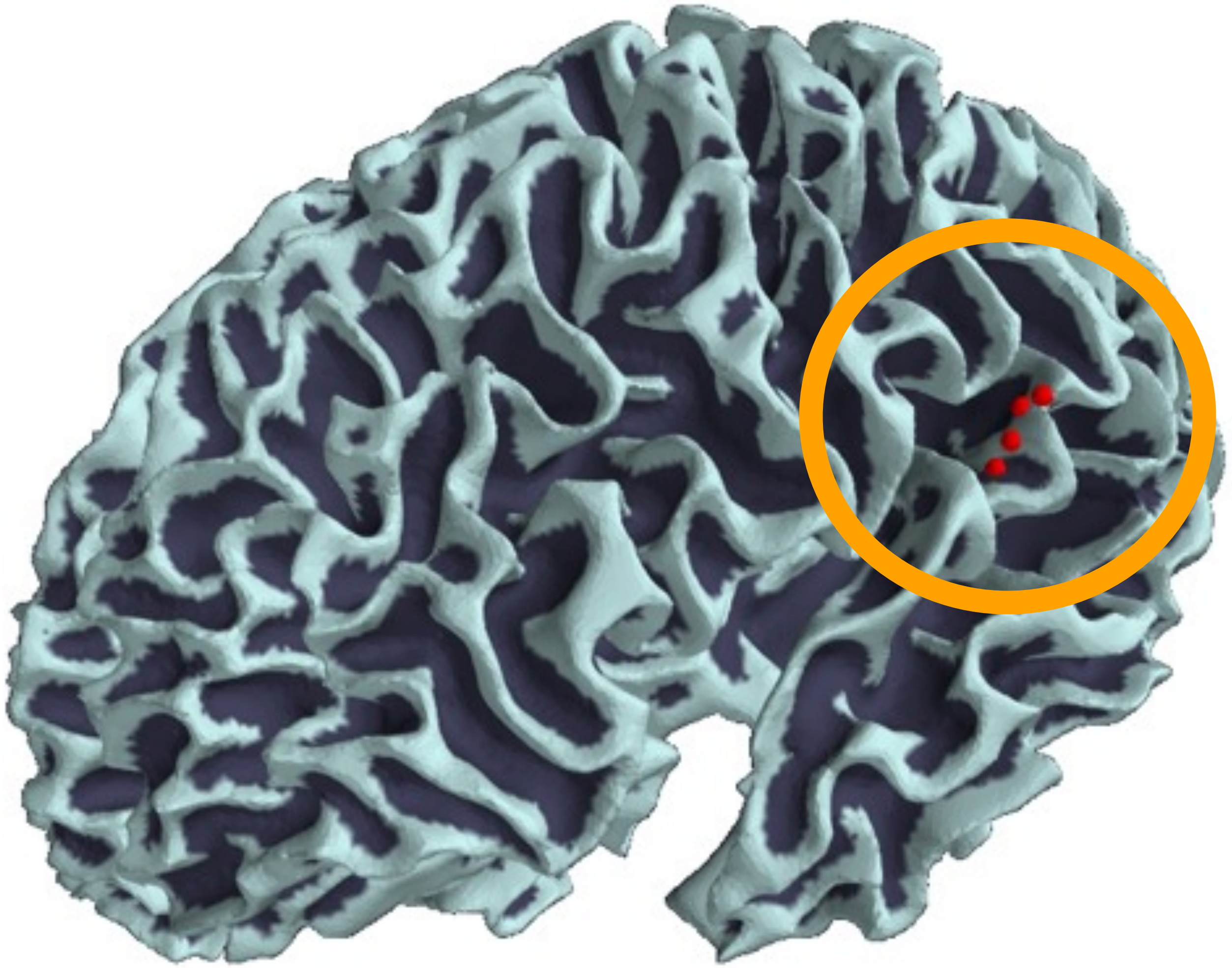
L1 vs L2 norms on combined M/EEG data

Activation in left-auditory cortex

L2 result



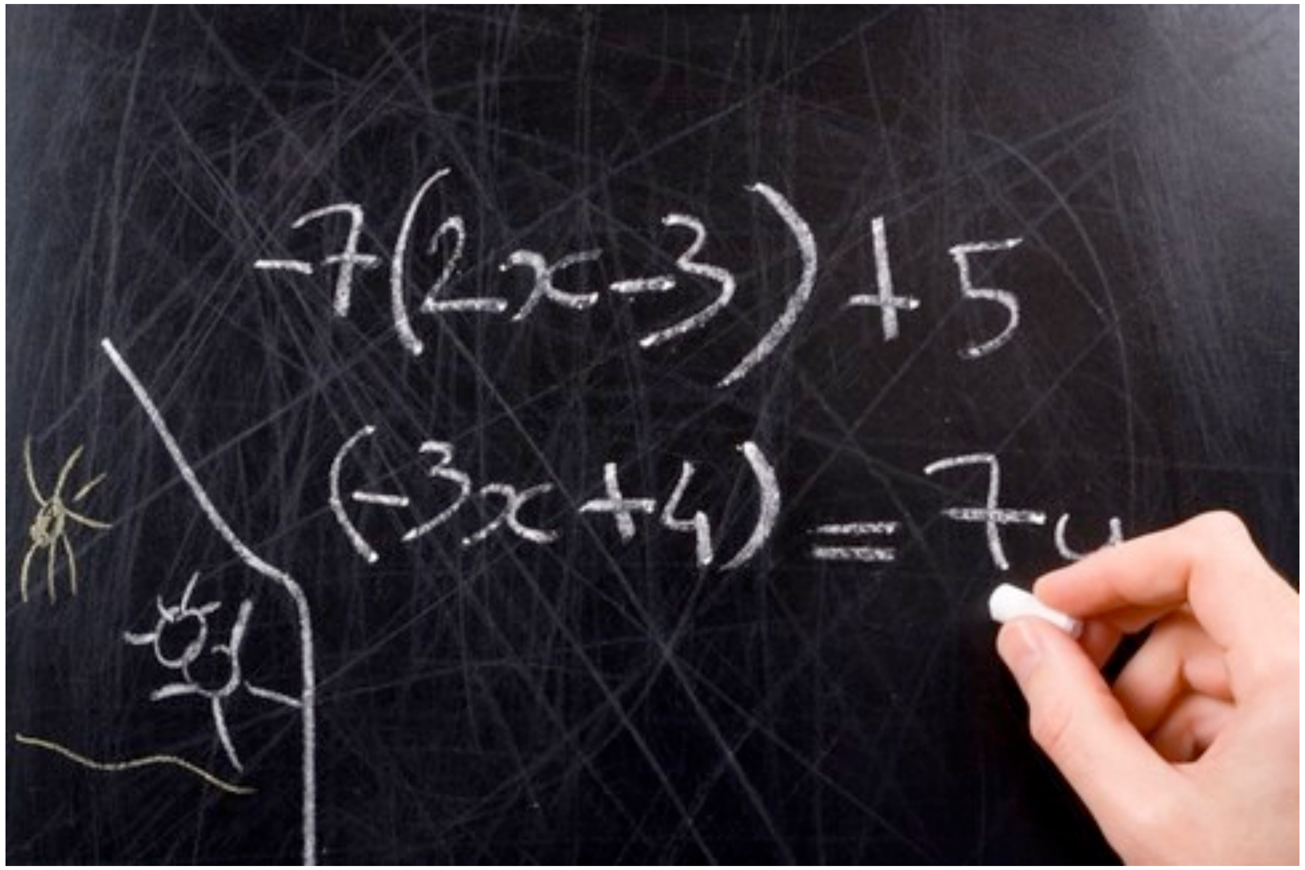
L1 result



$\phi(\mathbf{X})$ with M/EEG data: L2

Simple L2 (Tikhonov):

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



Quiz: Complexity and Computing times

- Complexity of matrix multiplication

$$GX \text{ with } G \in \mathbb{R}^{d_m \times d_x} \text{ and } X \in \mathbb{R}^{d_x \times d_t}$$

- Complexity of matrix inversion

$$(G^T G + \lambda I)^{-1}$$

- Resolution of a linear system: $Ax = b$
(when A is sparse or dense)
- Resolution of many linear system:

$$Ax_i = b_i, i = 1, ..., d_n$$

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):

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

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

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):

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

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

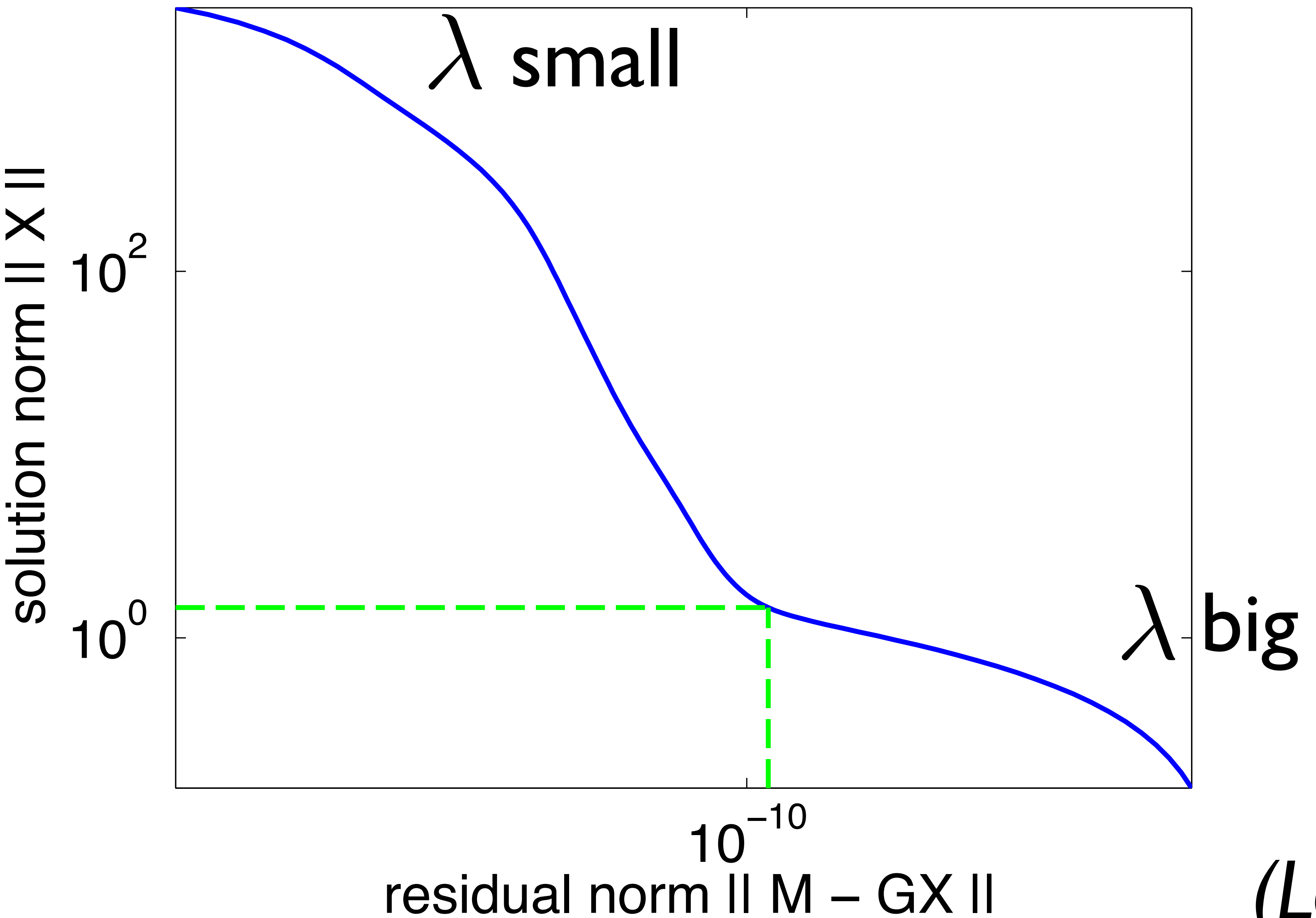
Remarks:

- **MNE** is known as **Ridge regression** in statistics.
- **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, ...)

THM: A lot of domain knowledge to make it work

How do I set the regularization parameter?

The L-curve



(Log-Log plot)

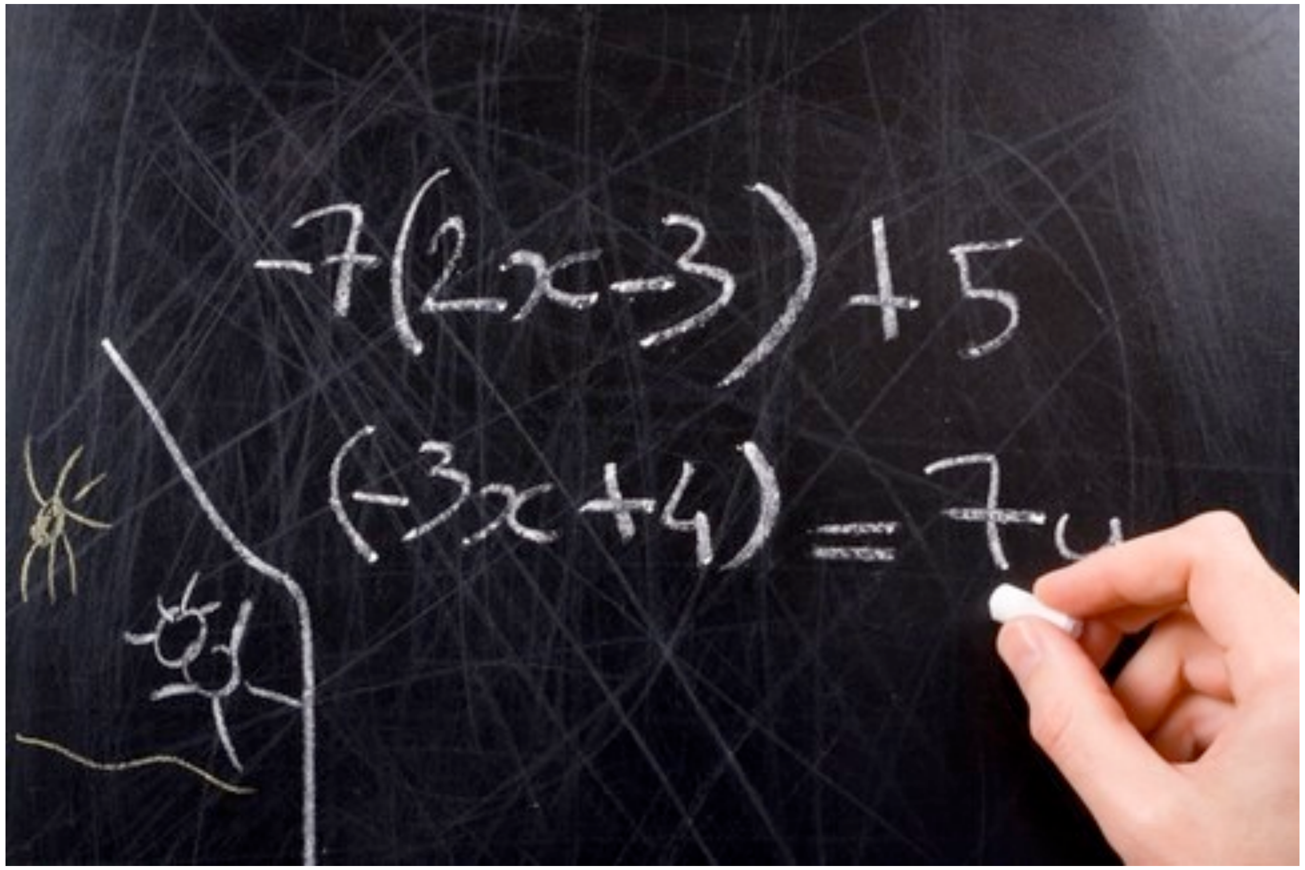
[Hansen 92]

A naïve but efficient approach

Compute the SVD (Singular Value Decomposition) of G:

$$\mathbf{G} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

with $\mathbf{U}\mathbf{U}^\mathbf{T} = \mathbf{U}^\mathbf{T}\mathbf{U} = \mathbf{I}$ $\mathbf{V}\mathbf{V}^\mathbf{T} = \mathbf{V}^\mathbf{T}\mathbf{V} = \mathbf{I}$ $\mathbf{S}$ diagonal + zeros



A naïve but efficient approach

Compute the SVD (Singular Value Decomposition) of G:

$$\mathbf{G} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

with $\mathbf{U}\mathbf{U}^\mathbf{T} = \mathbf{U}^\mathbf{T}\mathbf{U} = \mathbf{I}$ $\quad \mathbf{V}\mathbf{V}^\mathbf{T} = \mathbf{V}^\mathbf{T}\mathbf{V} = \mathbf{I}$ $\quad \mathbf{S}$ diagonal + zeros

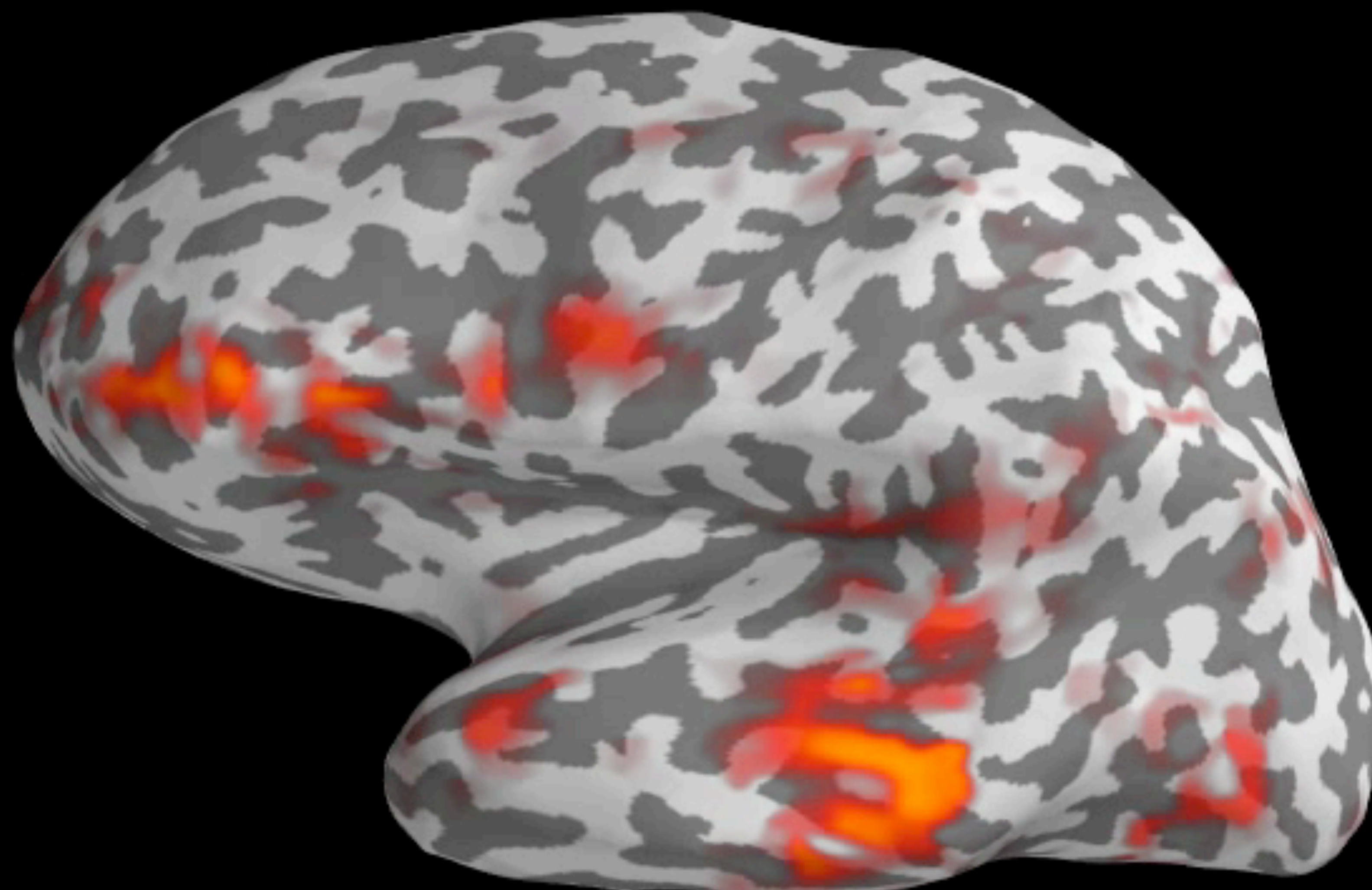
Replace the SVD in:

$$\mathbf{X}^* = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T + \lambda\mathbf{I})^{-1}\mathbf{M}$$

$$\begin{aligned}
\mathbf{X}^* &= \mathbf{G}^T(\mathbf{U}\mathbf{S}^2\mathbf{U}^T + \lambda\mathbf{I})^{-1}\mathbf{M} \\
&= \mathbf{G}^T(\mathbf{U}(\mathbf{S}^2 + \lambda\mathbf{I})\mathbf{U}^T)^{-1}\mathbf{M} \\
&= \mathbf{G}^T\mathbf{U}(\mathbf{S}^2 + \lambda\mathbf{I})^{-1}\mathbf{U}^T\mathbf{M} \\
&= \mathbf{V}\mathbf{S}(\mathbf{S}^2 + \lambda\mathbf{I})^{-1}\mathbf{U}^T\mathbf{M}
\end{aligned}$$

λ compares to the squared singular values of G

Take λ as a percentage of the max singular value http://youtu.be/Uxr5Pz7JPrs



time=0.00 ms

Alexandre Gramfort, Telecom ParisTech, CEA/Neurospin

Beyond L2 priors

$\phi(\mathbf{X})$ with M/EEG data: L1

L1 priors a.k.a. Minimum current estimate (MCE) :

$$\phi(\mathbf{X}) = \|\mathbf{X}\|_1 = \sum_i |x_i| \quad \text{with} \quad d_t = 1$$

[Matsuura et al. 95]

$\phi(\mathbf{X})$ is **convex**, **non differentiable** and has **no closed form solution**.

Remarks:

- It's the **LASSO** problem in the Machine Learning community [Tibshirani 96]
- It's the **Basis Pursuit** problem in Signal Processing [Chen Donoho Saunders 99]
- Matsuura uses linear programming but other algorithms exist, e.g., LARS [Efron 2004], Homotopy [Osborne 2000], coordinate descent, **IRLS**, **proximal iterations** etc.

Iterative Least Squares (IRLS)

Idea: $\|\mathbf{X}^*\|_1 = \sum_i |x_i^*| = \sum_i \frac{(x_i^*)^2}{w_i} = \|\mathbf{X}^*\|_{w,2}$ **when** $w_i = |x_i^*|$

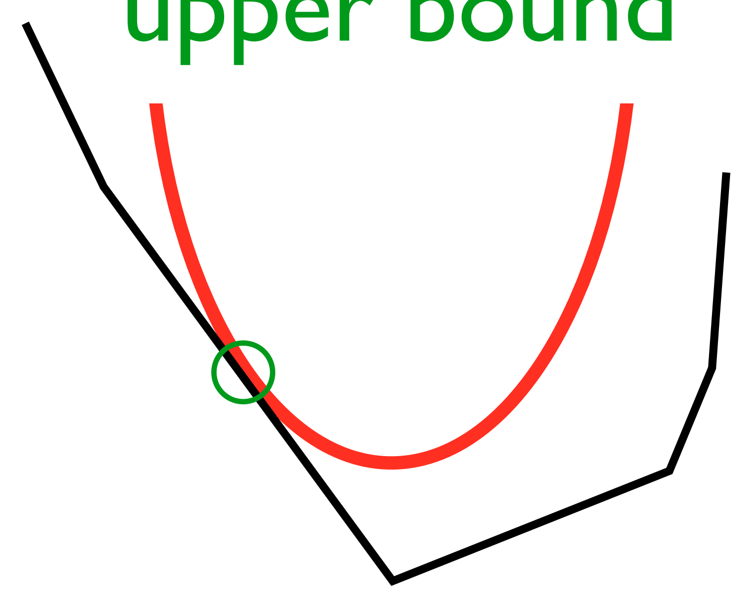
Proof

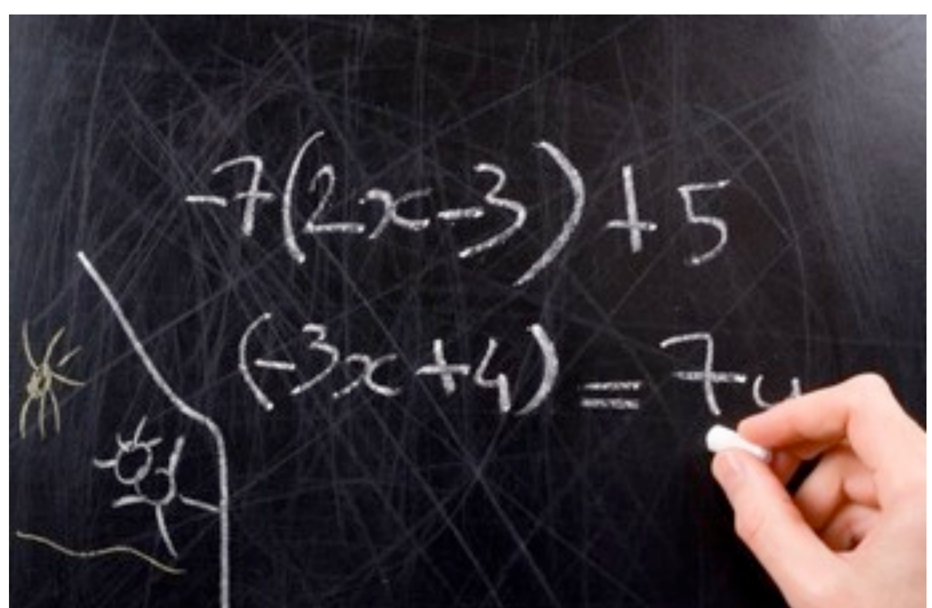
$$\min_{\mathbf{x}} \frac{1}{2\lambda}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \|\mathbf{x}\|_1$$

$$= \min_{\mathbf{x}} \frac{1}{2\lambda}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \sum_i |x_i|$$

$$= \min_{\mathbf{x},\mathbf{w}} \frac{1}{2\lambda}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \frac{1}{2}\sum_i \left(\frac{(x_i)^2}{w_i} + w_i\right)$$

Quadratic upper bound





Iterative Least Squares (IRLS)

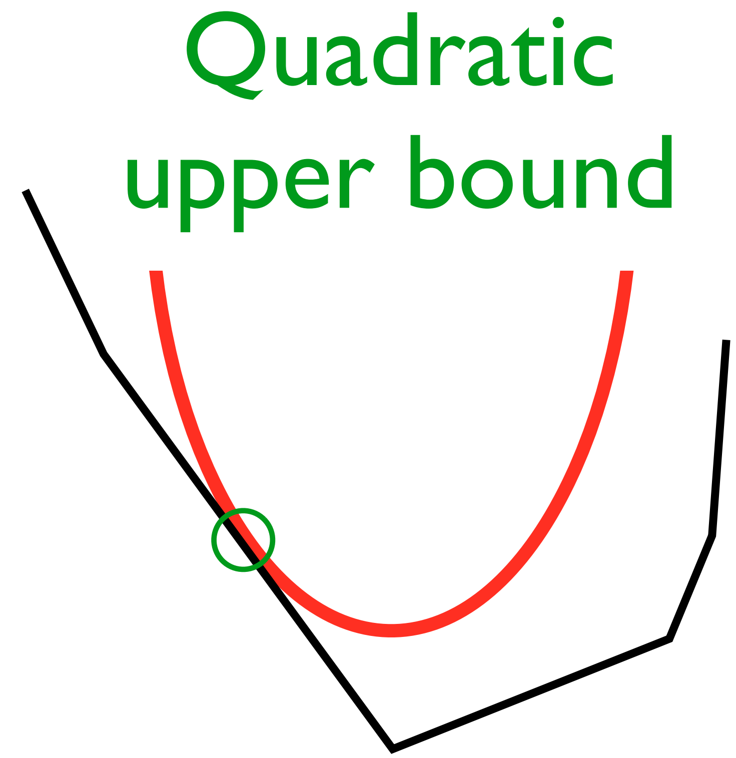
Idea: $\|\mathbf{X}^*\|_1 = \sum_i |x_i^*| = \sum_i \frac{(x_i^*)^2}{w_i} = \|\mathbf{X}^*\|_{w,2}$ when $w_i = |x_i^*|$

Proof

$$\min_{\mathbf{x}} \frac{1}{2\lambda}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \|\mathbf{x}\|_1$$

$$= \min_{\mathbf{x}} \frac{1}{2\lambda}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \sum_i |x_i|$$

$$= \min_{\mathbf{x},\mathbf{w}} \frac{1}{2\lambda}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \frac{1}{2}\sum_i \left(\frac{(x_i)^2}{w_i} + w_i\right)$$



Algorithm

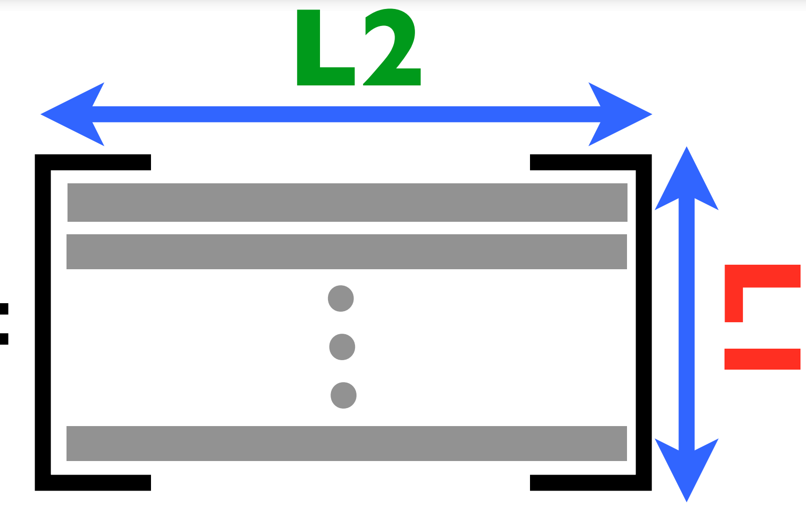
- Initialization: $\mathbf{W}^{(0)} = \mathbf{Id}$
- Compute: $\mathbf{x}^{(k+1)} = (\mathbf{W}^{(k)}) \quad \mathbf{G}^T(\mathbf{G}(\mathbf{W}^{(k)}) \quad \mathbf{G}^T + \lambda\mathbf{Id})^{-1}\mathbf{m}$

  Least square
- Update the weights: $w_i^{(k+1)} = |x_i^{(k+1)}|$
- Stop if $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|$ is smaller than a fixed tolerance value.

$\phi(\mathbf{X})$ with M/EEG data: L21

$$\phi(\mathbf{X}) = \|\mathbf{X}\|_{21} = \sum_i \sqrt{\sum_t |x_{i,t}|^2}$$

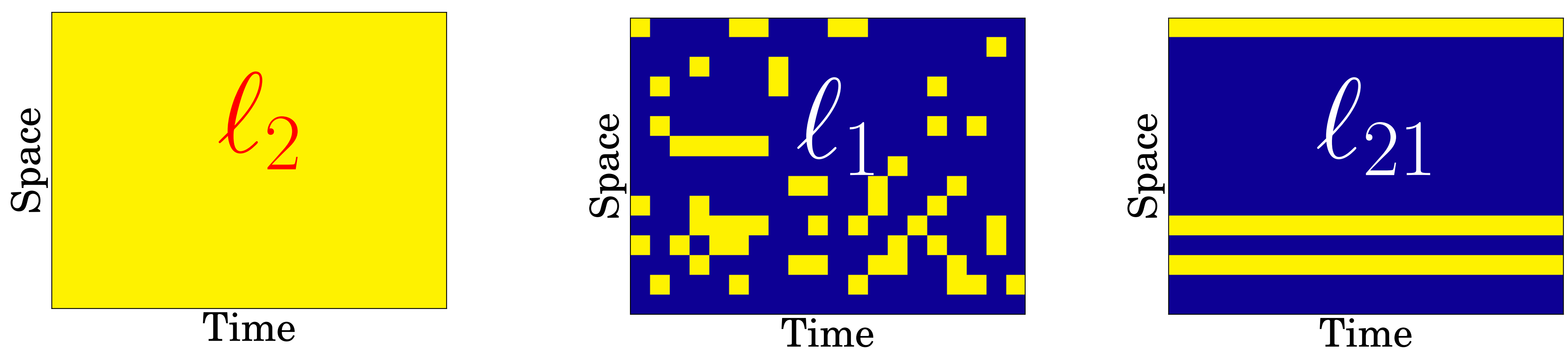
X =



2-level mixed-norm

[Ou et al. Neuroimage 2009]

- It introduces **temporal structure** in the prior
- It guarantees that the active sources are the same over time

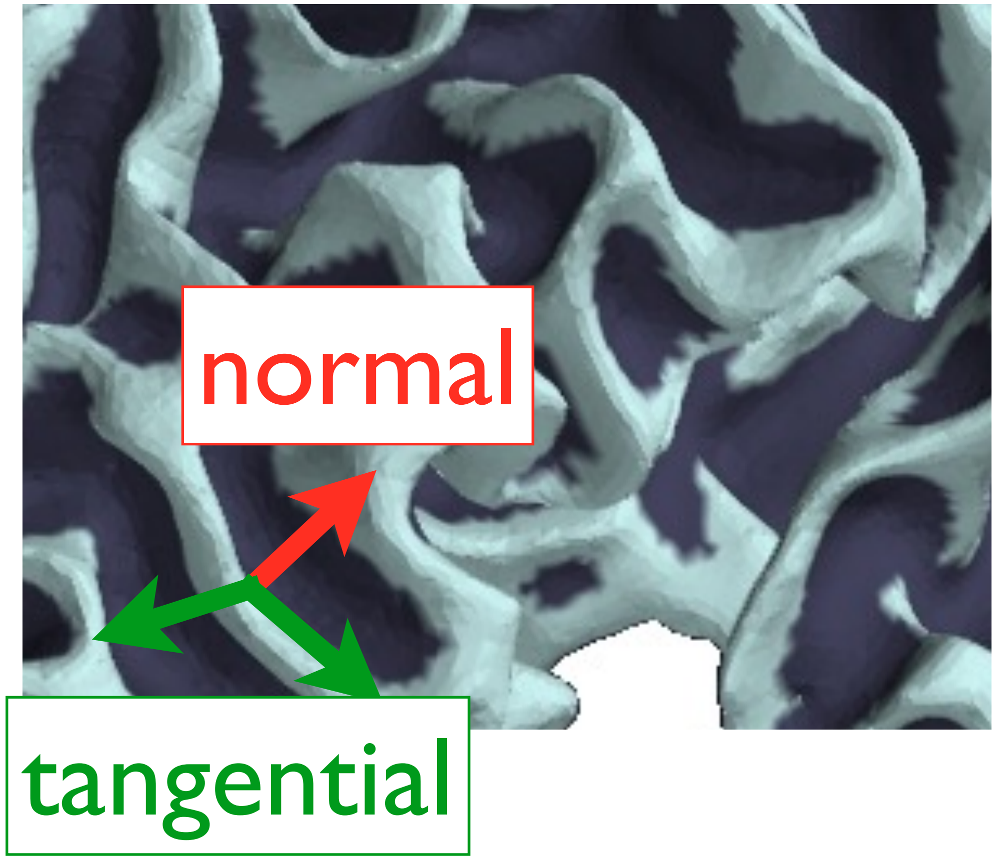


Remark : It is known as Group Lasso in Machine Learning & «joint feature selection»

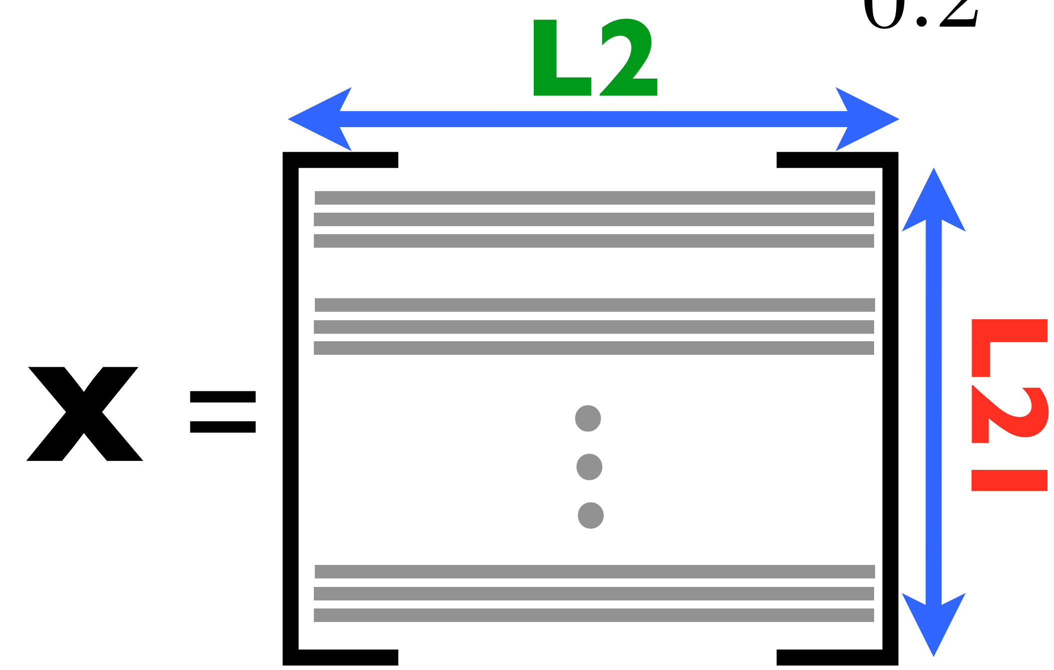
[Yuan et al. 2006, Obozinski 2009 ...]

L21 with loose orientation

$$\phi(\mathbf{X}) = \|\mathbf{X}\|_{21} = \sum_i \sqrt{\sum_t |x_{i,t}^{normal}|^2 + \rho |x_{i,t}^{tang1}|^2 + \rho |x_{i,t}^{tang2}|^2}$$



with for example $\rho = \frac{1}{0.2}$

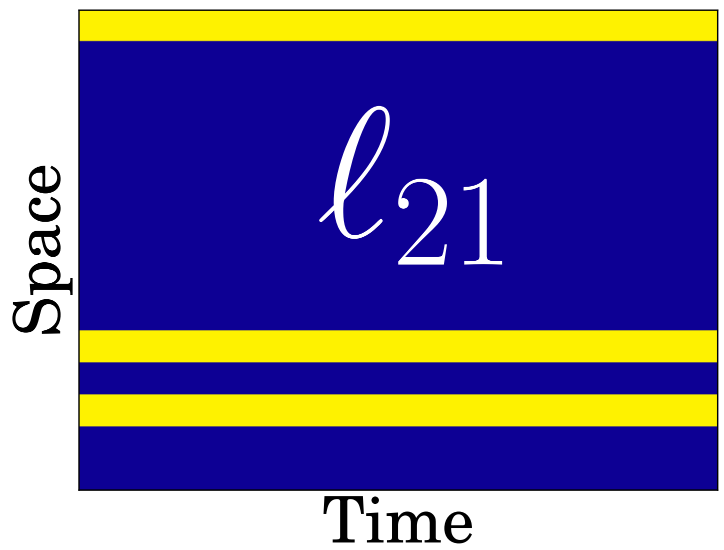


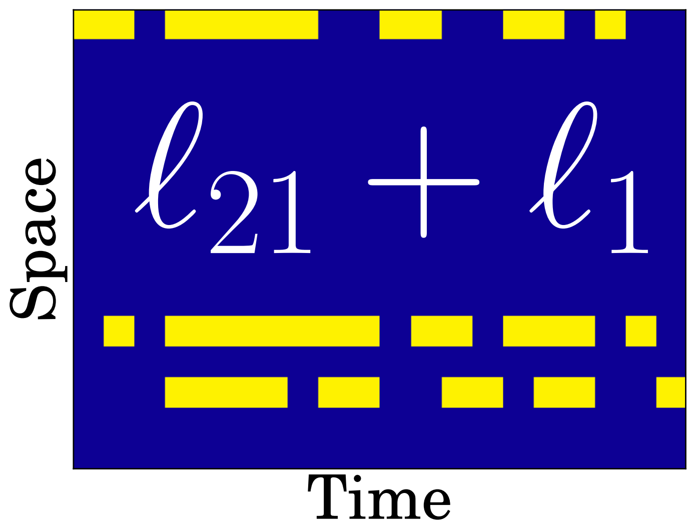
custom but still a 2-level mixed-norm

THM: you need custom sparse solvers adapted to M/EEG

But… the brain is not stationary

L21 like any other sparse solver available today
it imposes the sources to be the same over the entire time interval

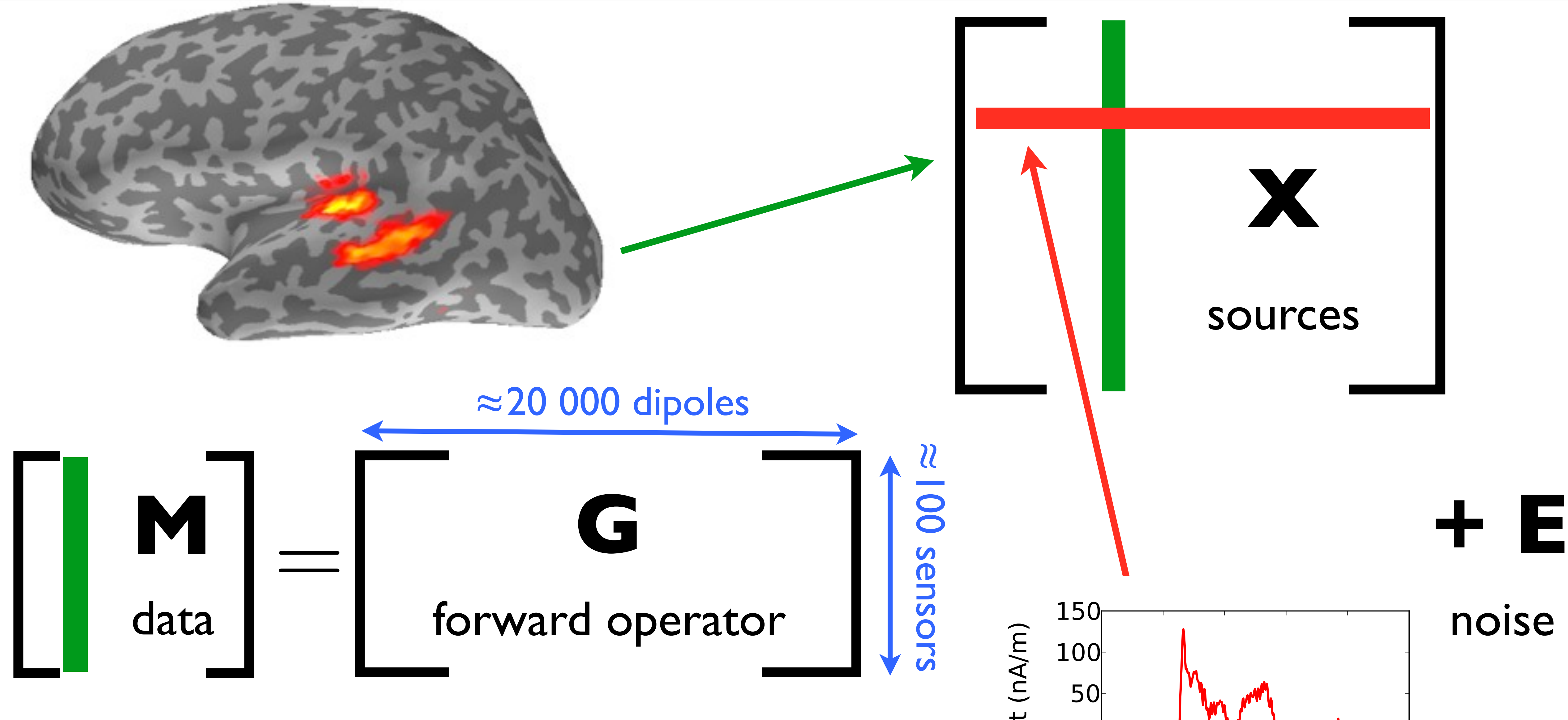


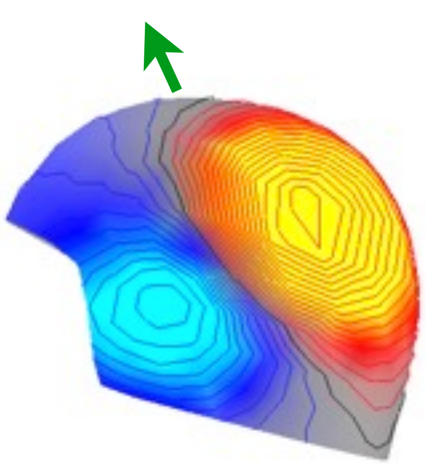


Challenge:

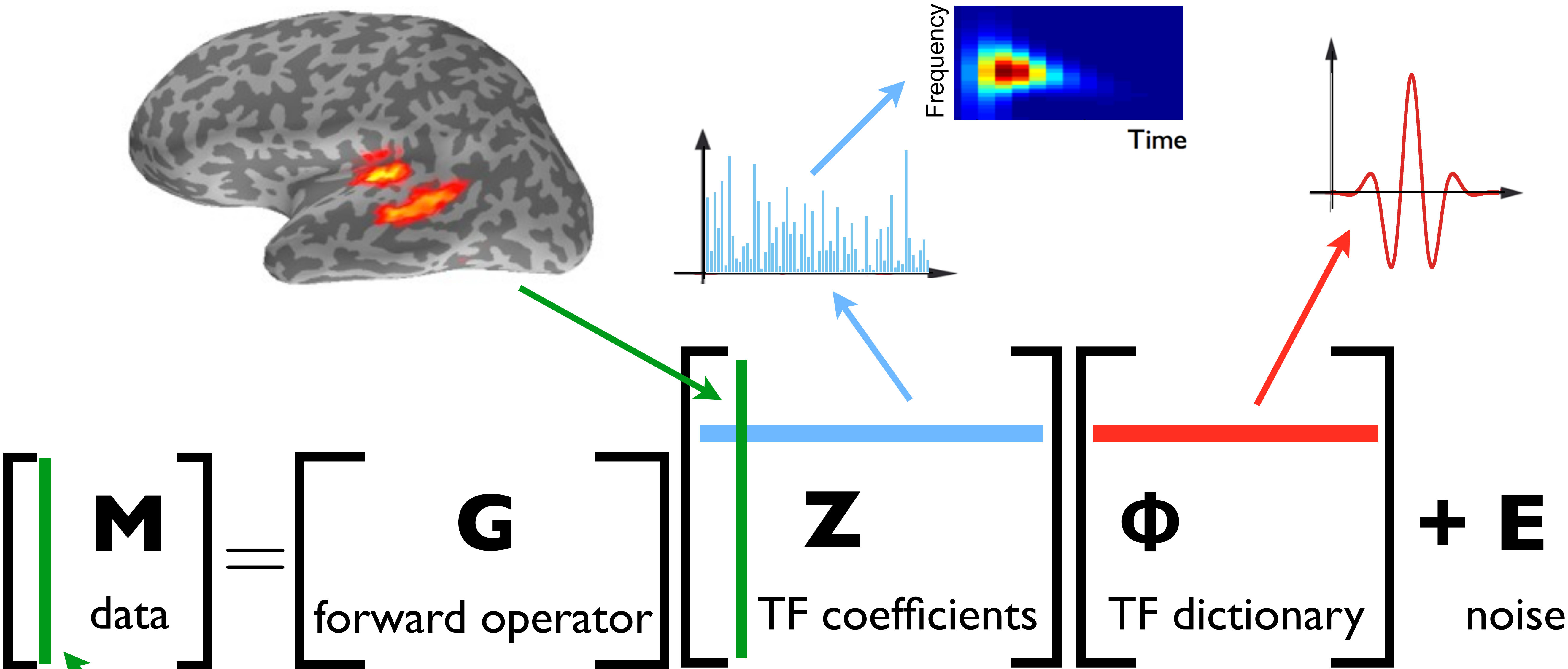
How do you promote sparse solutions with non-stationary sources?

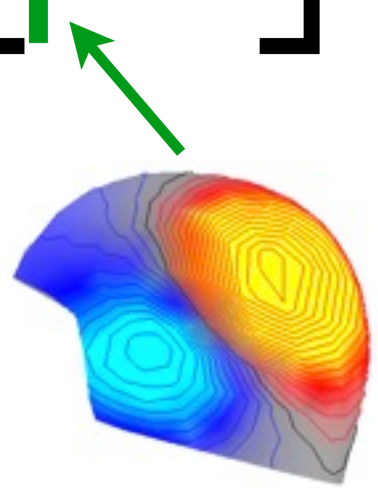
back to M = G X + E





Objective: estimate X given M





Objective: estimate Z given M

Time-frequency (TF) prior

The classical approach [MNE, dSPM, sLORETA]:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \underbrace{\|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F^2}_{\text{data fit}} + \underbrace{\lambda\phi(\mathbf{X})}_{\text{prior}}, \ \ \lambda > 0$$

we propose:

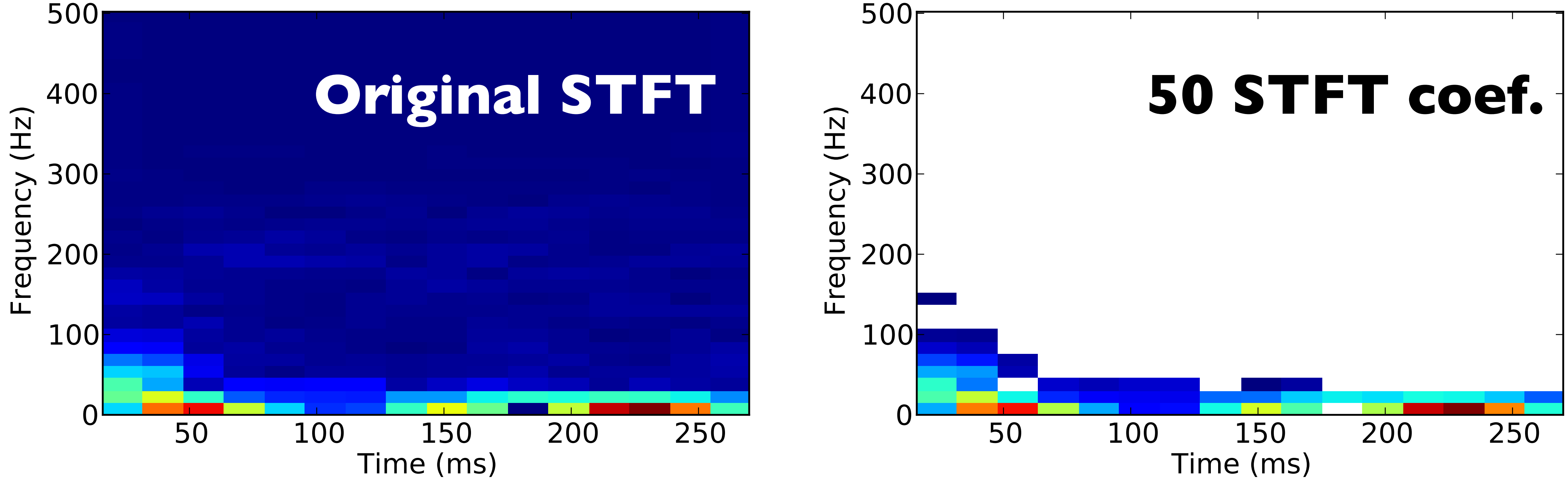
$$\hat{\mathbf{Z}} = \arg\min_{\mathbf{Z}} \|\mathbf{M} - \mathbf{G}\mathbf{Z}\mathbf{\Phi}^{\mathcal{H}}\|_F^2 + \lambda\phi(\mathbf{Z}), \ \text{then} \ \hat{\mathbf{X}} = \hat{\mathbf{Z}}\mathbf{\Phi}^{\mathcal{H}}$$

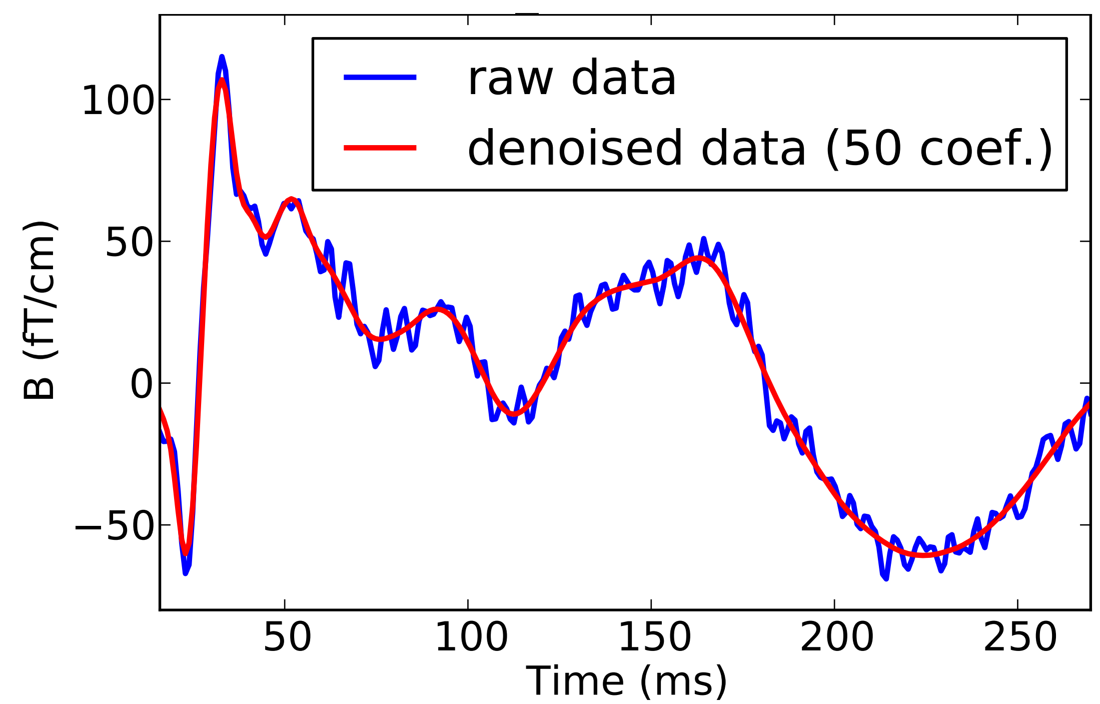
- $\mathbf{\Phi}$: is a **TF dictionary** of Gabor atoms
- $\mathbf{Z}$: **coefficients** of the **TF transform** of the sources

Advantage:
localization in
space, time and frequency
in one step

Why does it make sense?

and why a sparse prior shall work ?





[«Denoising by soft-thresholding» Donoho 95]

Time frequency dictionaries

discrete version of the
complex Gabor transform = short time fourier transform (STFT)

- It is **invertible**
- It is **translation invariant**

  (not like classical dyadic wavelets)
- It can capture **non-stationary signals** (not like FFT)

  (It is classically used in M/EEG on sensor measurements)
- It is **relatively fast** to compute

Time frequency dictionaries

complex Gabor transform = discrete version of the short time fourier transform (STFT)

- It is **invertible**
- It is **translation invariant**

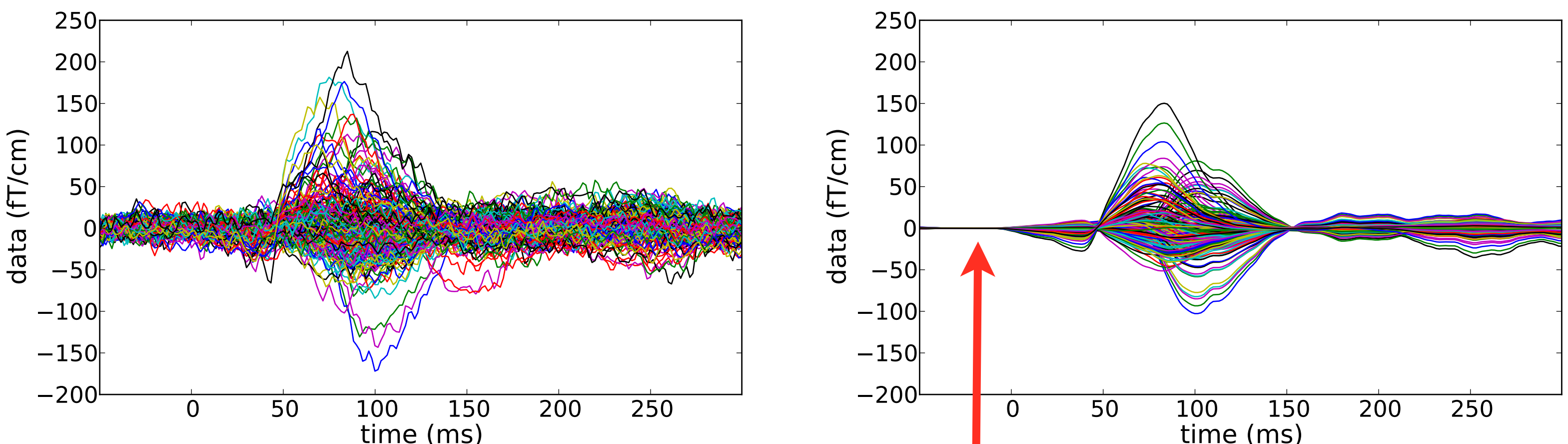
  (not like classical dyadic wavelets)
- It can capture **non-stationary signals** (not like FFT)

  (It is classically used in M/EEG on sensor measurements)
- It is **relatively fast** to compute

What is a good prior on Z?

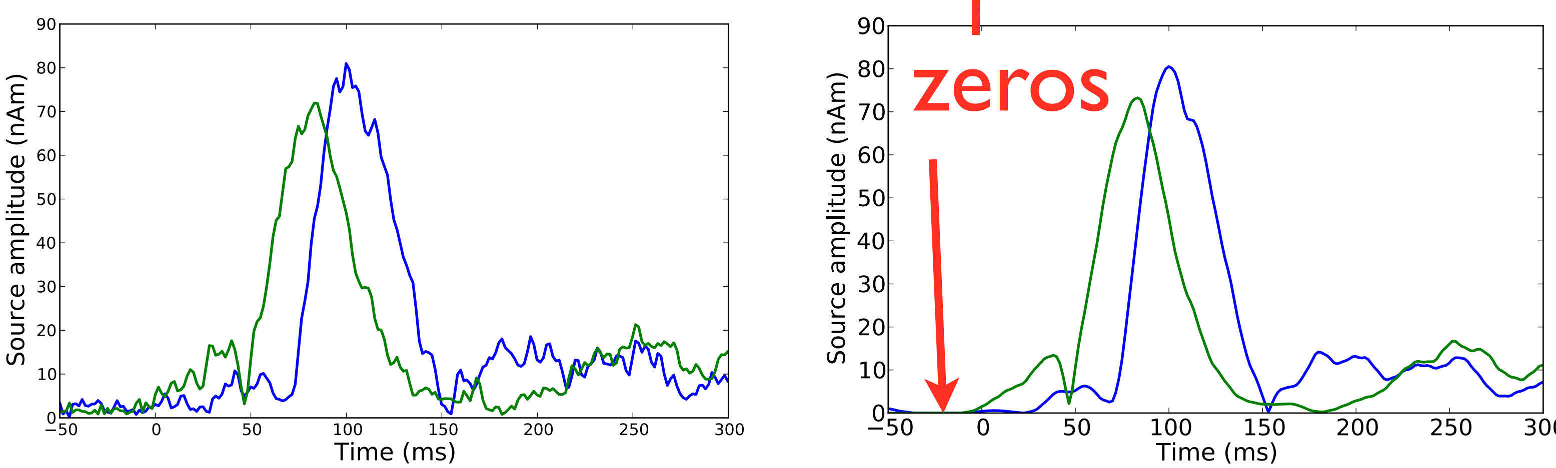
MEG Auditory data

Protocol: 50 epochs of auditory tones in left ear (305 MEG, 59 EEG channels)



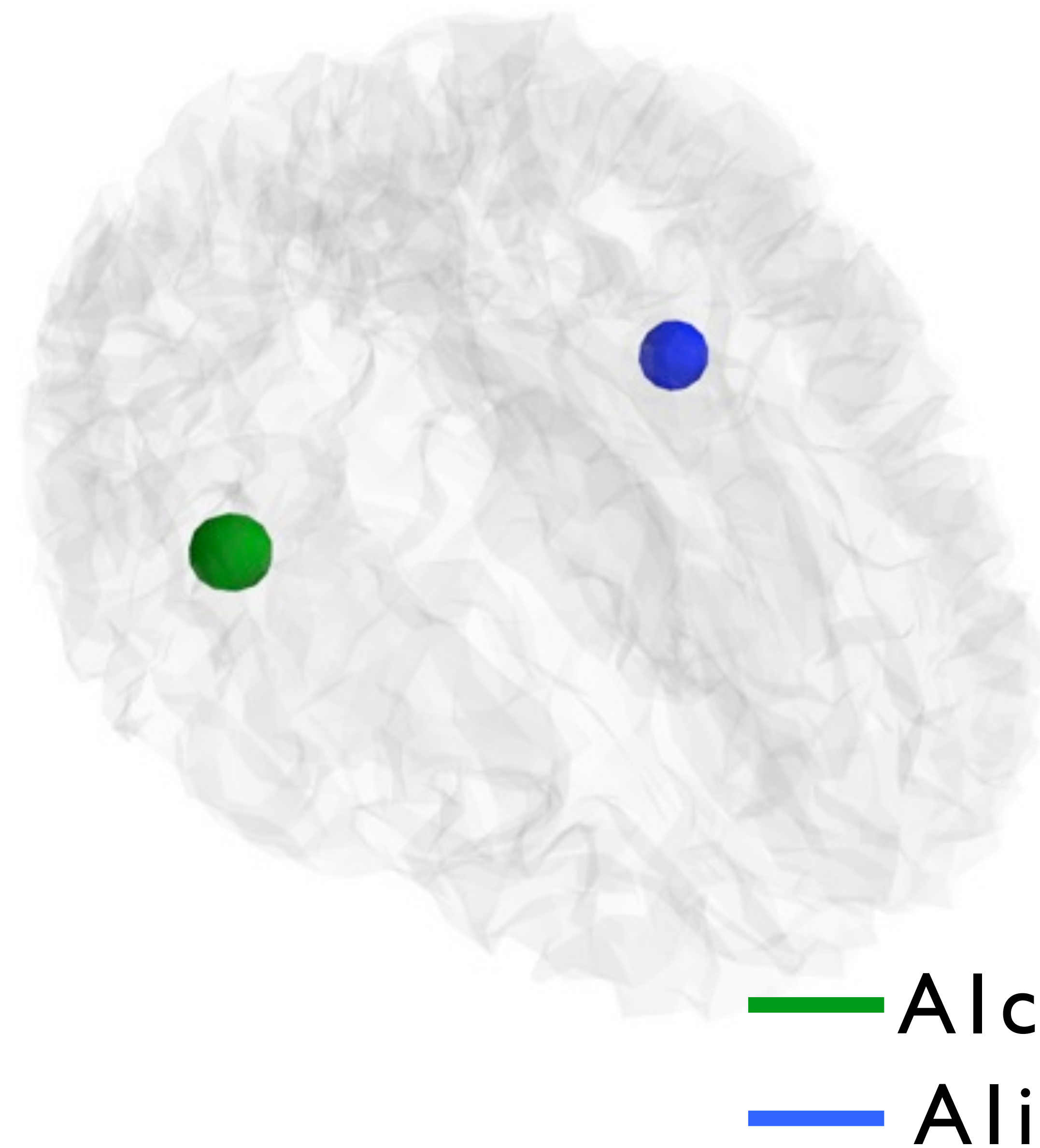
(a) MEG data (Gradiometers only)

(b) $\mathbf{G}\mathbf{X}^{\star}_{\text{TF-MxNE}}$ (explained data)

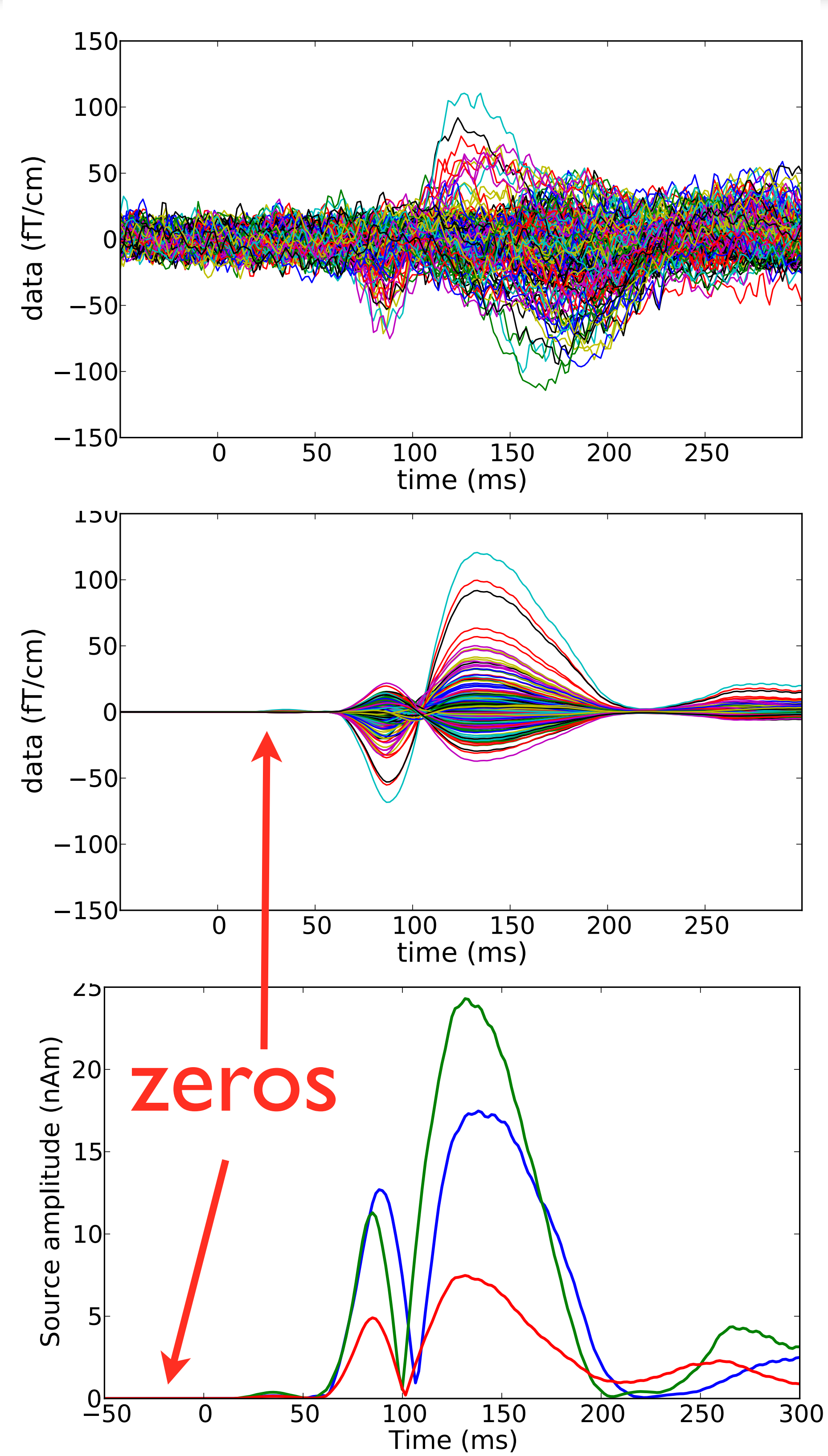


(c) $\mathbf{X}^{\star}_{\text{MxNE}}$

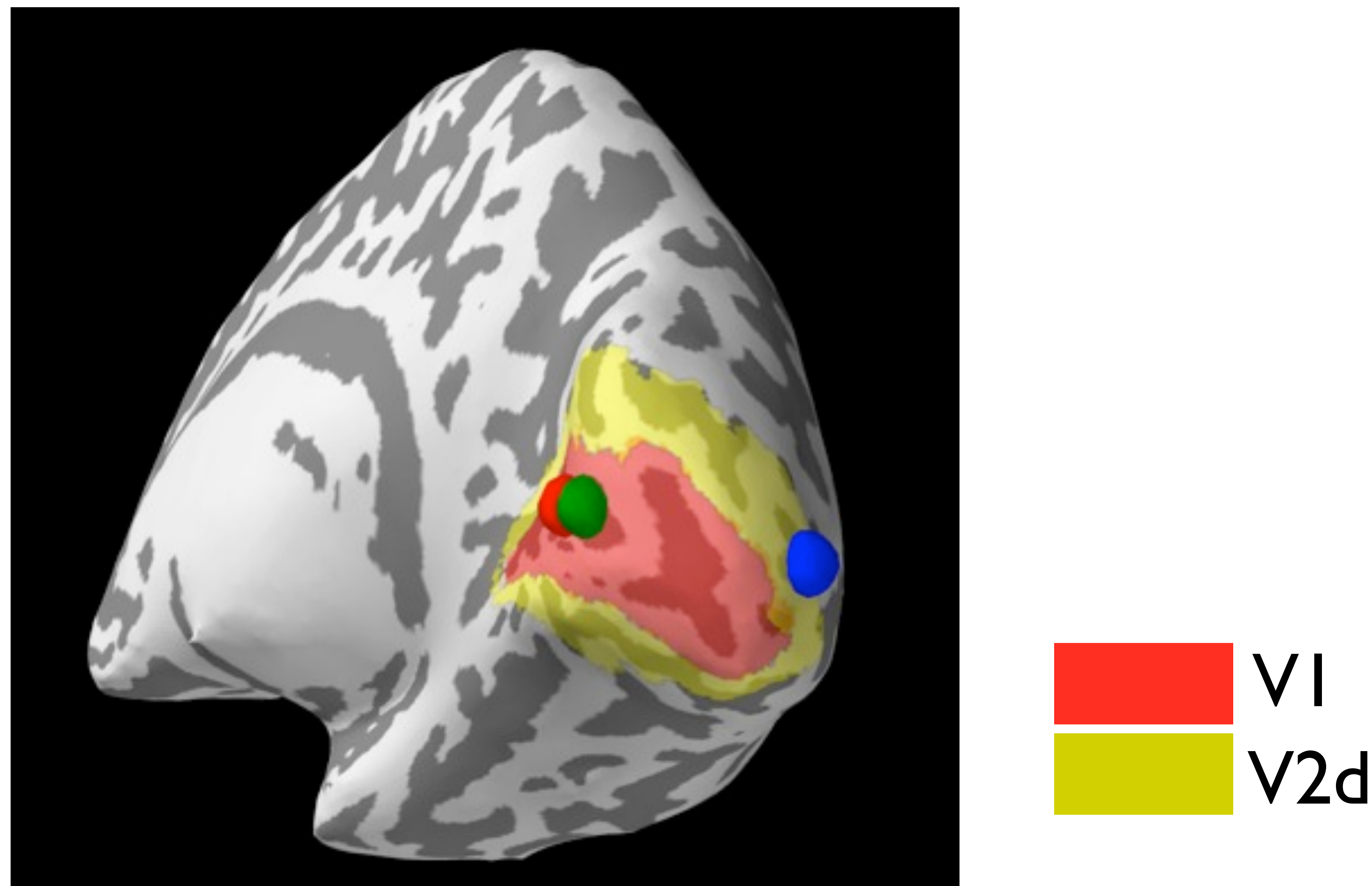
(d) $\mathbf{X}^{\star}_{\text{TF-MxNE}}$

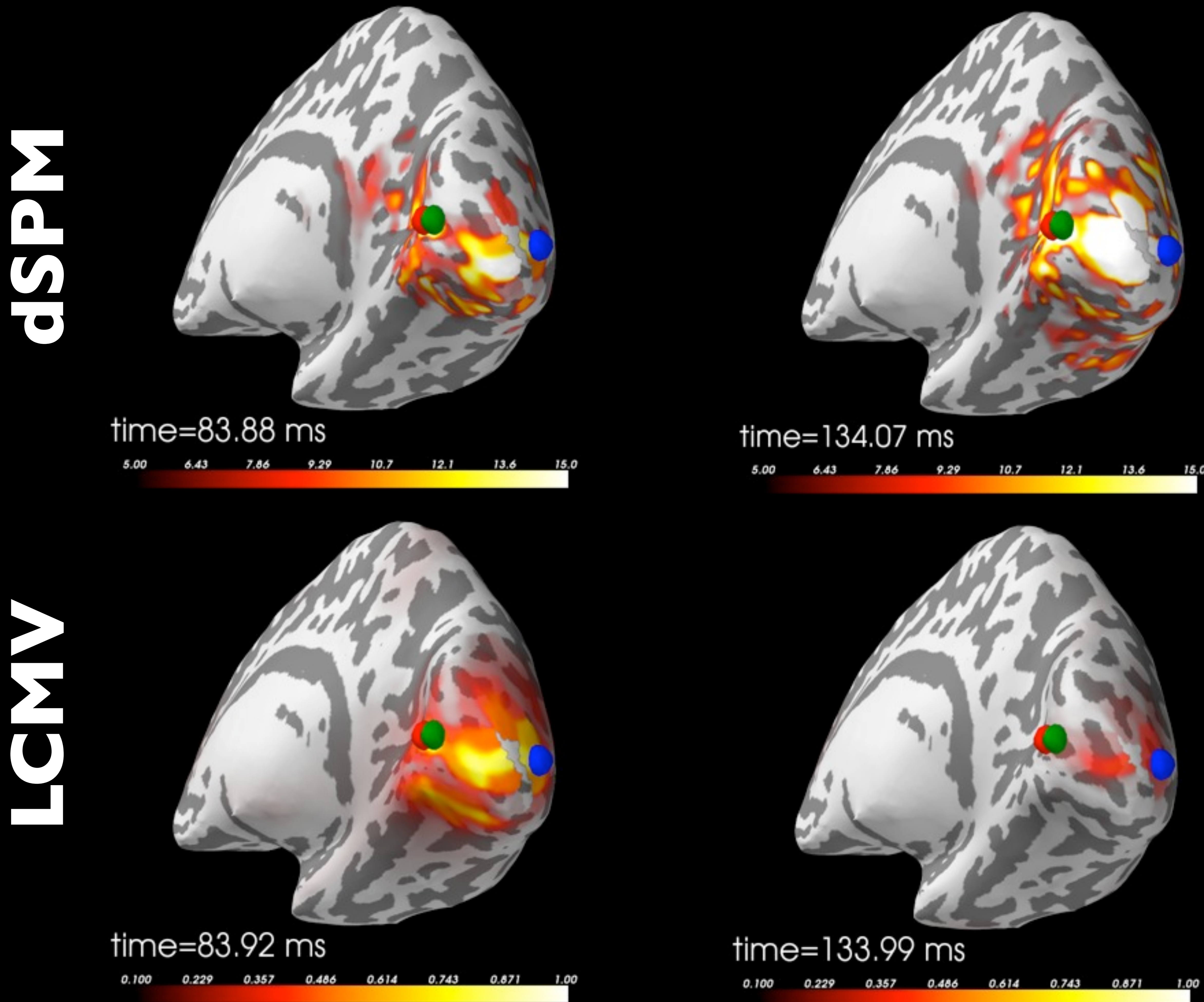


MEG Visual data



Protocol: 50 epochs of visual flash in left hemi-field (305 MEG, 59 EEG channels)





dSPM
LCMV
time=83.88 ms
5.00 6.43 7.86 9.29 10.7 12.1 13.6 15.0
time=134.07 ms
5.00 6.43 7.86 9.29 10.7 12.1 13.6 15.0
time=83.92 ms
0.100 0.229 0.357 0.486 0.614 0.743 0.871 1.00
time=133.99 ms
0.100 0.229 0.357 0.486 0.614 0.743 0.871 1.00

Conclusion

To sum up

- MEG and EEG measure the electrical activity of local assemblies of neurons (post-synaptic potentials)
- Can be used for: clinical applications (epilepsy, sleep), cognitive studies or BCI
- Acquisitions: physics
- Forward problem: image (segmentation), maths (PDE, numerical solvers)
- Inverse problem: statistics, optimization, signal processing