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

Alexandre Gramfort
alexandre.gramfort@telecom-paristech.fr

2013
It’s the study of the **brain activity** through functional imaging devices.

- **Temporal resolution (ms)**: 1, 10, 100, 1000, 10000, 100000

- **Spatial resolution (mm)**: 1, 5, 10, 15, 20

**Electrophysiology**
- sEEG
- EEG
- MEG

**Optical Imaging**
- nIRS

**Functional MRI**
- fMRI
- PET
- SPECT

*Functional neuroimaging*
Brain anatomy

2 hemispheres
4 Lobes

Brain mesh obtained by MRI segmentation

Source: Gray's anatomy (public domain)
Relation between location and function?

- Premotor cortex
- Motor cortex
- Primary somatic sensory cortex
- Intra-parietal sulcus (IPS)
- Wernicke's area
- Primary visual cortex
- Lateral occipital complex (LOC)
- Gustatory area
- Smell
- Prefrontal cortex (Emotion, behavior)
- Broca's area
- Primary auditory cortex
Relation between location and function?

Left hemisphere in medial view

Primary visual cortex (V1)

Source: Gray’s anatomy (public domain)
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.

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

White matter

Gray matter

Neurons in the gray matter

Source: dartmouth.edu
**APs** (action potentials) & **PSPs** (post-synaptic potentials)

- **Action Potentials**: Initiated in the Neuron body and propagated along the Axon via Synapse.
- **Dendrite**: Site of PSP generation with + and - signs indicating potential changes.
- **PSP**: Positive and negative potentials accumulate along the Dendrite.
- **Spike initiation zone**: Location where APs are generated after sufficient PSP accumulation.

Alexandre Gramfort  
Functional Brain Imaging with MEG/EEG/sEEG
APs (action potentials) & PSPs (post-synaptic potentials)

Action Potentials

Axon

Synapse

Dendrite

Neuron body

Temporal dynamics

100 mV

10 mV

1 ms

10 ms

AP

PSP

Spike initiation zone
APs (action potentials) & PSPs (post-synaptic potentials)

Pre-synaptic

Action potentials:
fields diminish too rapidly to sum

Post-synaptic

Postsynaptic currents:
fields diminish gradually
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

\[ Q = I \times d \]

(10 to 100 nAm) with the equivalent current dipole (ECD) model

**Gray matter**

**White matter**

**Electric Field** \( \vec{E} \)

**Magnetic Field** \( \vec{B} \)

Neural Current (post synaptic)

**Equivalent Current Dipole**
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
Hence the importance of shielding...
Magnetic shielding

Magnetically Shielded Room (MSR)

- Magnetic shielding
- 3-ply µ-metal room

External magnetic field
A machine (Neromag vectorview)

- No Magnet
- Quiet
  
  Machine makes no noise
- Participant can sit or lay down
- Can record 128 EEG simultaneously
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

EEG:
• ≈ 100 sensors

MEG:
• ≈ 150 to 300 sensors

Sampling between 250 and 1000 Hz

High temporal resolution but what about spatial resolution?

Sample EEG measurements
At each time instant EEG sensors measure a potential field.

Remark: Such a smooth potential field confirms the presence of current generators within the head.
MEG topography exhibits also a dipolar field but MEG has a **better spatial resolution**.
M/EEG Measurements

MEG
- tangential
- radial
- tilted

EEG
- tangential
- radial
- tilted

Courtesy of Prof. Matti Hämäläinen, Harvard
M/EEG Measurements: Notation

\[ M = \begin{bmatrix} \text{MEG and/or EEG} \end{bmatrix} \in \mathbb{R}^{d_m \times d_t} \]

- \( d_m \): Nb of sensors
- \( d_t \): Nb of time points

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?
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

Stimuli (if any)
- auditory
- visual
- somatosensory

Task
- attend/ignore
- detect + react

Behavorial responses
- limb/finger movement
- speech

MEG/EEG
- evoked responses
- spontaneous data

Source: nih.gov
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

- drift
- eye blink
- Buzz = Line noise 60Hz
- cardiac

Time frame: 10 seconds
Raw continuous data

Time frame: 10 seconds
Filtered 1-40Hz

Time frame: 10 seconds
Artifact correction

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

Alexandre Gramfort

Functional Brain Imaging with MEG/EEG/sEEG
To get clean data...

Time frame: 10 seconds
Source localisation with M/EEG: The forward and inverse problems
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 Equations with quasi-static approximation

\[
\begin{align*}
\nabla \times \vec{E} &= 0 \\
\nabla \cdot \vec{B} &= 0 \\
\n\nabla \times \vec{B} &= \mu_0 \vec{J} \\
\n\nabla \cdot \vec{E} &= \frac{\rho}{\epsilon_0}
\end{align*}
\]

Remark: quasi-static implies no temporal derivatives and no propagation delay
Maxwell Equations with **quasi-static** approximation:

\[
\begin{align*}
\nabla \times \vec{E} &= 0 \\
\nabla \cdot \vec{B} &= 0 \\
\n\nabla \times \vec{B} &= \mu_0 \vec{J} \\
\n\nabla \cdot \vec{E} &= \frac{\rho}{\varepsilon_0}
\end{align*}
\]

**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
- \( \sigma \) Tissue conductivity

With quasi-static approximation:

- All currents
- Electric potential
- Tissue conductivities

What is MEG? From Maxwell to the gain matrix.
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

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

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

Observations:
• B is obtained after V
• B decreases in \(1/R^2\)
• B is due both to primary currents and volume currents

where \[ \vec{B}_0 = \frac{\mu_0}{4\pi} \int \vec{J}_p \times \frac{r - r'}{||r - r'||^3} \, dr' \]
Head models

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

**Hypothesis**: The conductivities are **piecewise constant**

Sphere models

Realistic models

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

[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

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

Equivalent Current Dipole Technique
<table>
<thead>
<tr>
<th>Median Nerve Dipole Fitting Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 sensors</td>
</tr>
<tr>
<td><img src="circle.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="dipole.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="results.png" alt="Image" /></td>
</tr>
<tr>
<td><strong>Median Nerve Dipole Fitting</strong></td>
</tr>
<tr>
<td><strong>Results</strong></td>
</tr>
<tr>
<td><strong>99.7%</strong></td>
</tr>
<tr>
<td><strong>99.2%</strong></td>
</tr>
<tr>
<td><strong>98.2%</strong></td>
</tr>
<tr>
<td><strong>84.6%</strong></td>
</tr>
<tr>
<td><strong>97.6%</strong></td>
</tr>
<tr>
<td><strong>85.8%</strong></td>
</tr>
<tr>
<td><strong>84.6%</strong></td>
</tr>
<tr>
<td><strong>97.6%</strong></td>
</tr>
<tr>
<td><strong>85.8%</strong></td>
</tr>
</tbody>
</table>

*circle = size of confidence volume*
Time course of SEF
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)

\[
G = \begin{bmatrix}
\end{bmatrix}
\]

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

one column = Forward field of one dipole

EEG forward field on the electrodes

MEG forward field on sensors
\[ M = GX + E \]

**Linear forward model**, i.e.,

- **M** is the sum of the contributions of all the sources
  
  (Superposition principle)

**Symbols**:
- \( M \in \mathbb{R}^{d_m \times d_t} \): M/EEG Measurements
- \( X \in \mathbb{R}^{d_x \times d_t} \): Source amplitudes (Unkowns)
- \( G \in \mathbb{R}^{d_m \times d_x} \): Leadfield (or Gain) matrix
- \( E \in \mathbb{R}^{d_m \times d_t} \): additive noise
Scanning methods

\[ M = GX + E \]

**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 \( X \)...

---

**CHAPTER 3. THE INVERSE PROBLEM WITH DISTRIBUTED SOURCE MODELS**

Inverse matrices cannot be explicitly computed. We need for each pair \((\mu, \mu')\) to run an iterative solver which can make the GCV and L9Curve methods particularly time consuming.

**3.3 LEARNING BASED METHODS**

In previous sections, the \( \| \mu \|_2^2 \) priors used in the penalization of the inverse problem are defined a priori. Following the explanations in section “…” this means that the proposed methods assume a predefined covariance matrix for the sources.

In the following paragraphs, we will present inverse solvers that aim at designing a prior based on the data. The source covariance matrix, i.e. the weights in the \( \| \mu \|_2^2 \) penalization term, is “learned.” We will also say that the model is learned from the data [,'”].

For simplicity, we will present the following method in the context of instantaneous inverse computation.

The methods presented in this section use the Bayesian formulation of the inverse problem. We recall the Bayesian framework from section “…”:

\[
p(X | M) = \frac{p(M | X) p(X)}{p(M)}.
\]

If \( E \) and \( X \) are known, \( X \) is obtained by maximizing the likelihood which leads to:

\[
X = \arg \min X M = GX + E
\]

which leads to:

\[
X = \left( G^T (G G^T + E) \right)^{-1} M.
\]

In this framework the prior is an \( \| \mu \|_2^2 \) norm and learning the prior means learning \( X \), i.e., the source covariance matrix. One may also want to learn the noise covariance matrix \( E \). Note that in the WMN framework, learning \( X \) consists in learning the weights.

In the case where \( X \) and \( E \) are not fixed a priori, these parameters define the model commonly denoted \( M \). Bayes’ rule can be rewritten:

\[
p(X | M, M) = \frac{p(M | X, M) p(X | M)}{p(M | M)}.
\]

\( p(X | M, M) \) is called the *posterior*, \( p(M | X, M) \) is called the *likelihood*, \( p(X | M) \) is called the *prior*, \( p(M | M) \) is called the *model evidence*.
\[ M = GX + E : \text{An ill-posed problem} \]

\[
\begin{bmatrix}
\text{M}
\end{bmatrix} = \begin{bmatrix}
\text{G}
\end{bmatrix}
\]

\(d_x \approx 10000\)

\(d_m \approx 100\)

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

Alexandre Gramfort  Functional Brain Imaging with MEG/EEG/sEEG
An optimization problem:

\[
X^* = \arg \min_X \| M - GX \|^2_F + \lambda \phi(X), \quad \lambda > 0
\]

- **Data fit**
- **Prior (penalization)**

\(\lambda\) : Trade-off between the data fit and the prior

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

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

Examples for \(\phi(X)\) : \(\ell_1\), \(\ell_2\), \(\ell_p\) with \(p \geq 1\), entropy . . .

Remark: If \(\phi(X)\) is strictly convex we have a unique minimizer (sufficient but not a necessary condition)
Definition: Convex function

\[ f : \mathbb{R}^n \to \mathbb{R} \]

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:

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

Data fit

\[\text{convex} \quad + \quad \text{convex} = \text{convex}\]

- Data fit is \textit{quadratic} hence \textit{convex}
- If \(\phi(X)\) is \textit{convex}, then it’s a \textit{convex} optimization problem
Smooth or non-smooth

- **Smooth:**
  - L2 (regularized Least-squares, Tikhonov)
  - Entropy based methods
  - etc.

- **Non-smooth:**
  - L1
  - Total-Variation
  - etc.

\[
\phi(X) = \|X\|_2^2 = \sum_{i,j} x_{i,j}^2
\]

\[
\phi(X) = \|X\|_1 = \sum_{i,j} |x_{i,j}|
\]

\[
\phi(X) = TV(X) = \|\nabla_{surf} X\|_1
\]
Activation in left-auditory cortex

L2 result

L1 result
Simple L2 (Tikhonov):

\[
X^* = \arg \min_X E(X) = \arg \min_X \|M - GX\|_F^2 + \lambda \|X\|_F^2, \lambda > 0
\]
Quiz: Complexity and Computing times

- Complexity of matrix multiplication

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

- 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:

\[ A x_i = b_i, \ i = 1, \ldots, d_n \]
L2 a.k.a. Minimum Norm Estimates (MNE)

\[ \phi(X) = \|WX\|_F^2 = \sum_{i,j} w_i^2 x_{ij}^2 = \|X\|_{\Sigma,2}^2 \]

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

\[ X^* = \Sigma^{-1} G^T (G \Sigma^{-1} G^T + \lambda \text{Id})^{-1} M \]

[Tikhonov et al. 77, Wang et al. 92, Hämäläinen et al. 94]
L2 a.k.a. Minimum Norm Estimates (MNE)

\[ \phi(X) = \| WX \|_F^2 = \sum_{i,j} w_i^2 x_{ij}^2 = \| X \|_\Sigma^2 \]

Leads to a closed form solution (matrix multiplication):

\[ X^* = \Sigma^{-1} G^T (G \Sigma^{-1} G^T + \lambda \text{Id})^{-1} M \]

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 \(\lambda\) based on SNR, loose orientation, ...)

**THM:** A lot of domain knowledge to make it work

[Tikhonov et al. 77, Wang et al. 92, Hämäläinen et al. 94]
How do I set the regularization parameter?
The L-curve

(Log-Log plot)

\[ \begin{align*}
\text{residual norm } & \| M - GX \| \\
\text{solution norm } & \| X \| \\
\lambda & \text{ small} \\
\lambda & \text{ big}
\end{align*} \]

[Ref: Hansen 92]
A naïve but efficient approach

Compute the SVD (Singular Value Decomposition) of $G$:

$$G = USV^T$$

with $UU^T = U^TU = I$ and $VV^T = V^TV = I$. $S$ is diagonal and has non-zero entries.

$$-7(2x - 3) + 5 = 7u$$

$$(-3x + 4) = 7u$$
A naïve but efficient approach

Compute the SVD (Singular Value Decomposition) of $G$:

$$G = USV^T$$

with $UU^T = U^TU = I$ \quad $VV^T = V^TV = I$ \quad $S$ diagonal + zeros

Replace the SVD in:

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

$$X^* = G^T(US^2U^T + \lambda I)^{-1}M$$

$$= G^T(U(S^2 + \lambda I)U^T)^{-1}M$$

$$= G^TU(S^2 + \lambda I)^{-1}U^TM$$

$$= VS(S^2 + \lambda I)^{-1}U^TM$$

$\lambda$ compares to the squared singular values of $G$

Take $\lambda$ 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
L1 priors a.k.a. Minimum current estimate (MCE):

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

[Matsuura et al. 95]

\( \phi(X) \) 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: \[ \|X^*\|_1 = \sum_i |x_i^*| = \sum_i \frac{(x_i^*)^2}{w_i} = \|X^*\|_{w,2} \text{ when } w_i = |x_i^*| \]

\[
\begin{align*}
\min_x \frac{1}{2\lambda} \|m - Gx\|_2^2 + \|x\|_1 \\
= \min_x \frac{1}{2\lambda} \|m - Gx\|_2^2 + \sum_i |x_i| \\
= \min_{x,w} \frac{1}{2\lambda} \|m - Gx\|_2^2 + \frac{1}{2} \sum_i \left( \frac{(x_i)^2}{w_i} + w_i \right)
\end{align*}
\]

Proof

Quadratic upper bound

\[ \begin{align*}
-7(2x-3) + 5 \\
(4x + 4) = 7u
\end{align*} \]
Iterative Least Squares (IRLS)

Idea: \[ \|X^*\|_1 = \sum_i |x_i^*| = \sum_i \frac{(x_i^*)^2}{w_i} = \|X^*\|_{w,2} \quad \text{when } w_i = |x_i^*| \]

\[
\begin{align*}
\min_{x} & \quad \frac{1}{2\lambda} \|m - Gx\|_2^2 + \|x\|_1 \\
= & \min_{x} \frac{1}{2\lambda} \|m - Gx\|_2^2 + \sum_i |x_i| \\
= & \min_{x, w} \frac{1}{2\lambda} \|m - Gx\|_2^2 + \frac{1}{2} \sum_i \left( \frac{(x_i)^2}{w_i} + w_i \right)
\end{align*}
\]

\[ \text{Proof} \]

- **Initialization:** \( W^{(0)} = \text{Id} \)
- **Compute:** \( x^{(k+1)} = (W^{(k)}) \frac{G^T(G(W^{(k)})^T + \lambda \text{Id})^{-1}m}{w_i^{(k+1)}} \) \[ \text{Least square} \]
- **Update the weights:** \( w_i^{(k+1)} = |x_i^{(k+1)}| \)
- **Stop if** \( \|x^{(k+1)} - x^{(k)}\| \) \text{ is smaller than a fixed tolerance value.}
\[ \phi(X) = \|X\|_{21} = \sum_{i} \sqrt{\sum_{t} |x_{i,t}|^2} \]

2-level mixed-norm

- 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»

[Ou et al. Neuroimage 2009]

[Yuan et al. 2006, Obozinski 2009 ...]
L21 with loose orientation

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

with for example \( \rho = \frac{1}{0.2} \)

**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?
Bayes' rule can be rewritten:

\[ p_X | M = G X + E = \frac{p_M | X \cdot p_X}{p_M} \]

That in the WMN framework, learning the source covariance matrix \( \Sigma_2 \) is commonly denoted. One may also want to learn the noise covariance matrix \( \Sigma_1 \). In this framework, the prior is an inverse computation, and \( \Sigma_2 \) cannot be explicitly computed. We need for each pair \((X, M)\) where we assume Gaussian variables:

\[ p_M = \frac{1}{(2\pi)^{|M|/2} |M| |\Sigma_2|} \cdot \exp\left(-\frac{(G X + E)^T \cdot \Sigma_2^{-1} \cdot (G X + E)}{2 |M|}\right) \]

\[ p_X = \frac{1}{(2\pi)^{|X|/2} |X| |\Sigma_1|} \cdot \exp\left(-\frac{X^T \cdot \Sigma_1^{-1} \cdot X}{2 |X|}\right) \]

We recall the Bayesian framework from section "0":

\[ \log p_X \propto -\frac{1}{2} \| G X + E \|^2 \]

If and an additive model:

\[ X = \mu + \sigma \cdot N(0, \Sigma_1) \]

Note that the model evidence is learned from the data.

\[ p_M = \frac{1}{(2\pi)^{|M|/2} |M| |\Sigma_2|} \cdot \exp\left(-\frac{(G X + E)^T \cdot \Sigma_2^{-1} \cdot (G X + E)}{2 |M|}\right) \]

For simplicity, we will present the following method in the context of instantaneous reverberant systems.

The methods presented in this section use the Bayesian formulation of the inverse problem with distributed source models. Learning the prior means learning the weights in the forward operator. The priors used in the penalization of the inverse problem are data-dependent and are not fixed a priori. These parameters define the model.

**Objective:** estimate \( X \) given \( M \)

\[ \hat{X} = \arg \min_X -\text{log } p_X + \text{log } p_M \]

\[ \hat{X} = \arg \min_X \| G X + E \|^2 \]

\( G, M, E \) are obtained by maximizing the likelihood which leads to:

\[ \hat{X} = \arg \min_X \| G X + E \|^2 \]
\[ M = G Z \Phi + E \]

**Objective:** estimate \( Z \) given \( M \)

- \( M \): data
- \( G \): forward operator
- \( Z \): TF coefficients
- \( \Phi \): TF dictionary
- \( E \): noise
Time-frequency (TF) prior

The classical approach [MNE, dSPM, sLORETA]:

\[
\hat{X} = \arg \min_X \|M - GX\|_F^2 + \lambda \phi(X), \quad \lambda > 0 \\
\text{data fit} \\
\hat{Z} = \arg \min_Z \|M - GZ\Phi^H\|_F^2 + \lambda \phi(Z), \quad \text{then } \hat{X} = \hat{Z}\Phi^H \\
\]

we propose:

\[
\hat{Z} = \arg \min_Z \|M - GZ\Phi^H\|_F^2 + \lambda \phi(Z), \quad \text{then } \hat{X} = \hat{Z}\Phi^H \\
\]

• \( \Phi \) : is a TF dictionary of Gabor atoms

• \( Z \) : coefficients of the TF transform of the sources

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

Alexandre Gramfort  Functional Brain Imaging with MEG/EEG/sEEG  74
Why does it make sense?

and why a sparse prior shall work?

Original STFT

50 STFT coef.

[«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

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

What is a good prior on Z?
Protocol: 50 epochs of auditory tones in left ear (305 MEG, 59 EEG channels)
Protocol: 50 epochs of visual flash in left hemi-field (305 MEG, 59 EEG channels)
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