Bayesian inference in EEG/MEG

Jérémie Mattout
Lyon Neuroscience Research Center

“Will it ever happen that mathematicians will know enough about the physiology of the brain, and neurophysiologists enough of mathematical discovery, for efficient cooperation to be possible”

Jacques Hadamard (mathematician, 1865-1963)
Previously...

“What I cannot create, I cannot understand”

Richard Feynman (Nobel laureate in Physics, 1918-1988)

• Bayesian or Probabilistic modelling: accounts for knowledge/uncertainty in the process of interest

• Generative or phenomenological models

Ex. Dynamic causal models (DCM)

\[ \dot{x} = f(x, \theta_{\text{evol}}, u) \]

\[ y = g(x, \theta_{\text{obs}}) + \varepsilon \]

Hidden states

Unknown parameters

Observations (data features)

Graphical representation

\[ \theta_{\text{evol}} \]

\[ \theta_{\text{obs}} \]

\[ \varepsilon \]

\[ y \]

\[ f \]

\[ g \]

Fixed
Variable
Data
Previously...

“All models are wrong but some are useful”

George E. P. Box (statistician, 1919-2013)

• Bayesian or Probabilistic modelling: accounts for knowledge/uncertainty in the process of interest
• Generative or phenomenological models
• Inference on models

\[ P(\theta|Y,M) = \frac{P(Y|\theta,M)P(\theta|M)}{P(Y|M)} \]

- Usually approximated (FE, AIC, BIC)
- Trades accuracy and complexity
- Model comparison through Bayes Factors and model posteriors
- Models can only be compared on the same data
Previously...

- Bayesian or Probabilistic modelling: accounts for knowledge/uncertainty in the process of interest
- Generative or phenomenological models
- Inference on models
- Inference on model parameters

\[ P(\theta|Y, M) = \frac{P(Y|\theta, M)P(\theta|M)}{P(Y|M)} \]

- Inference at the group level
Outline

• 1 - EEG/MEG Source reconstruction

• 2 - Inferring learning mechanisms from EEG/MEG responses

• 3 - Online Bayesian inference and design optimization
1 – EEG/MEG source reconstruction

**An ill-posed inverse problem**

Generative model $M$

$p(Y \mid \theta, M)$  
Likelihood

$p(\theta \mid M)$  
Prior

$p(\theta \mid Y, M)$  
Posterior

$p(Y \mid M)$  
Evidence

Inverse Problem

Bayesian inference

No unique solution!

Additional information is needed
1 – EEG/MEG source reconstruction

Typical model families

• Biophysical model
  Static (forward) model family
  Geometrical and physical properties of the head tissues

• Source model
  **Equivalent current dipole (ECD) model family**
  \( \Theta \): dipole locations, orientations & intensity
  \( p(\Theta | M) \): number of dipoles, initial location

  **Distributed sources or Imaging model family**
  \( \Theta \): dipole intensities
  \( p(\Theta | M) \): number of dipoles, fixed locations

\[ Y \text{: sensor data (incl. noise)} \]
\[ p(Y | \theta, M) \]
1 – EEG/MEG source reconstruction

Detailed example: the distributed source model

- **Likelihood**
  \[ p(\varepsilon|M) = N(0, C_\varepsilon) \]
  The probability of a large noise is small

- **Prior**
  \[ p(\theta|M) = N(0, C_\theta) \]

\[ Y = K.\theta + \varepsilon \]
Additive noise

**Cortical mesh (3D surface)**
\[ \theta : \text{dipole intensities} \]

**Evoked response:** Nsens x Time

**Unknown source dynamics:** Nsources x Time

**Forward operator or lead-field matrix:** Nsens x Nsources
1 – EE/G/MEE source reconstruction

Detailed example: the distributed source model

- **Likelihood**
  
  \[ p(\varepsilon|M) = N(0, C_\varepsilon) \]

- **Prior**
  
  \[ p(\theta|M) = N(0, C_\theta) \]

\[ Y = K \cdot \theta + \varepsilon \]

- **Additive noise**
- **Forward operator or lead-field matrix**: Nsens x Nsources
- **Unknown source dynamics**: Nsources x Time
- **Cortical mesh (3D surface)**: \( \theta \): dipole intensities
- **Evoked response**: Nsens x Time

The proba. of a large intensity is small
1 – EEG/MEG source reconstruction

Commonly used priors

- Priors
  \[ p(\theta|M) = N(0, C_\theta) \]

Nsources x Nsources

- i.i.d or Minimum norm
- Single dipole
- Smoothness (like LORETA)
- fMRI based
- Lead-field based (Beamformer or MSP)

References:
- Philips et al. 2005
- Mattout et al. 2005 & 2006
- Wipf and Nagarajan 2009
Comparing priors

\( Y \) : observed data

\( \hat{\theta} \) : estimated source intensities

\( \tilde{Y} \) : predicted data

Explained variance

- 11%
- 96%
- 97%
- 98%

Log-evidence
1 – EEG/MEG source reconstruction

Empirical Bayes: multiple sparse priors (MSP)

Friston et al. 2008

\[ C_\theta = \lambda_1 Q_1 + \lambda_2 Q_2 + \lambda_3 Q_3 + \cdots + \lambda_n Q_n \]
1 – EEG/MEG source reconstruction

Model comparison on other dimensions

**Anatomy**
- Individual
- Inverse-normalised
- Template

Mattout et al. 2007

**Biophysics**
- Spherical head model
- Realistic surfacic model (BEM)

**Sources**
- Mesh resolution (High / Low)
- Dipole orientation (fixed / free)

Henson et al. 2009

**Model comparison on real MEG data**

Henson et al. 2007

- No evidence in favor of individual vs. Inverse-norm mesh
- Evidence in favor of BEM head model
- Evidence in favor of high + fixed vs. low + free
Directed Acyclic Graph (DAG) representation

Graphical representation of MSP

1 – EEG/MEG source reconstruction

Friston et al. 2008
1 – EEG/MEG source reconstruction

Model extension to group level inference

Directed Acyclic Graph (DAG) representation

A two-step procedure:
- Estimating the group prior variance
- Estimating the individual source intensities

Litvak and Friston 2008
1 – EEG/MEG source reconstruction

Directed Acyclic Graph (DAG) representation

Model extension to EEG-MEG data fusion

Henson et al. 2011
1 – EEG/MEG source reconstruction

Dynamic causal model (DCM) family

- Source model

- Biophysical model

Another model family

\[ \dot{x} = f(x, \theta_{evol}, u) \]

Model parameters:
- Source intrinsic parameters (e.g. synaptic gain)
- Extrinsic coupling
- Source orientation

\[ y = g(x, \theta_{obs}) + \epsilon \]
1 – EEG/MEG source reconstruction

An MMN example

Characterizing a group response in terms of modulations of effective connectivity

Classical auditory oddball paradigm

S: standard tone
D: (frequency) deviant tone

Controls (Ctrls): Controls
Minimally Conscious State (MCS)
Vegetative State (VS)

Garrido et al. 2009

Boly et al. 2011
2 – Inferring learning mechanisms from EEG/MEG data
2 – Inferring learning mechanisms from EEG/MEG data

Neurophysiological correlates of implicit perceptual learning

- In line with the Bayesian brain hypothesis and its neuronal instantiation (predictive coding), it has been suggested that evoked responses reflect prediction error

  Friston 2005

- Deviance responses should be modulated by predictability

- Single trial modulations of evoked responses should reflect perceptual learning
2 – Inferring learning mechanisms from EEG/MEG data

Neurophysiological correlates of implicit perceptual learning

• Deviance responses should be modulated by predictability

Effect of predictability

**US**: Unpredictable sequence

**PS**: Predictable sequence

![Brain activity images](image)

- Early effect: 60 ms
- MMN: 175 ms
- P3a: 315 ms

Lecaignard et al. 2015
2 – Inferring learning mechanisms from EEG/MEG data

Meta-Bayesian approach

• Single trial modulations of evoked responses should reflect perceptual learning

- Alternative models
  \[ \dot{x} = f(x, \theta, u) \]
  \[ y = g(x, \varphi) + \varepsilon \]

  - M1: Bayesian learner with different forgetting parameter value
  - M2
  - M3
  - M4
  - M5
  - M6: Change detection models
  - M7
  - M8: Null model (noise)

  \[ p(\mu|m, l) = \frac{\Gamma(m+l)}{\Gamma(m)\Gamma(l)} \mu^m (1-\mu)^l \]

  \[ \text{Bayesian Surprise} = \text{KL}(p(\mu|m_{w_{t-1}}, l_{w_{t-1}})\|p(\mu|m_{w_t}, l_{w_t})) \]

  Ostwald et al. 2012

- Data features: individual single-trial source activity in clusters estimated from EEG-MEG group inversion

  Lecaignard et al. in prep (a)
2 – Inferring learning mechanisms from EEG/MEG data

Preliminary results

- Single trial modulations of evoked responses should reflect perceptual learning

- Model comparison in a single subject, single block (condition us)

Lecaignard et al. in prep (b)
VBA toolbox, Daunizeau et al. 2014
3 – Online Bayesian inference and design optimization
Online Bayesian inference and design optimization

Motivation for Adaptive Design Optimization

- Fine neurobiological and/or computational models of evoked responses and their modulations may tell us more about information processing in the brain.
- Given such models, we may be in a position to disentangle between alternative hypothesis about ongoing perceptual inference and learning in a given subject/patient.
- However, what design to choose? How to optimize it for each individual?

Basic idea

- To make use of our ability to process data in real-time in order to optimize design parameters online, for a given individual, in the aim of selecting the most likely model both accurately and fast.

3 – Online Bayesian inference and design optimization

General principle of Adaptive Design Optimization

Classical approach (offline)

1. Experimental Design
2. Data Collection
3. Data Analysis
4. Hypotheses Testing
5. Conclusion

Adaptive approach (online)

1. Experimental Design
2. Online Optimization
3. Data Analysis

Conclusive Hypotheses Testing?

→ “No”

→ “Yes” → Conclusion

Time
A simple example: behavioural task of word memorization

**Experimental context**

<table>
<thead>
<tr>
<th>Memory phase</th>
<th>Recall phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of words</td>
<td>Number of correct recalled words</td>
</tr>
<tr>
<td>Retention Interval (Lag Time)</td>
<td></td>
</tr>
</tbody>
</table>

\[ Y \text{ (observations)} = \text{number of recalled words} \]
\[ u \text{ (design variable)} = \text{Lag time} \]

**Hypothesis**

Models of “forgetting”

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power (POW)</td>
<td>( p = a(t + 1)^{-b} )</td>
</tr>
<tr>
<td>Exponential (EXP)</td>
<td>( p = ae^{-bt} )</td>
</tr>
</tbody>
</table>

Predictions

(Low uncertainty about model parameters)

(High uncertainty about model parameters)

Cavagnaro et al. 2009, Myung et al. 2013
3 – Online Bayesian inference and design optimization

Optimization criterion

✓ Minimizing the risk of model selection error

Minimizing the risk of model selection error

Danizeau et al. 2011
Extension to more complex, multidimensional models

\[ \dot{x} = f(x, u, \theta_{\text{evol}}) \quad \text{Evolution function } f \]

\[ y = g(x, \theta_{\text{obs}}) + \varepsilon \quad \text{Observation function } g \]

Danizeau et al. 2011
Sanchez et al. 2014


3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ A perceptual learning model family: hierarchical Gaussian filters

\[ \dot{x} = f(x, u, \theta_{evol}) \]

Mathis et al. 2011 & 2014
Sanchez et al. 2014
3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ Simulations: alternative models and predictions

<table>
<thead>
<tr>
<th>Models</th>
<th>$\omega$ values</th>
<th>$\kappa$ values</th>
<th>$\theta$ values</th>
<th>Ability to track events probabilities</th>
<th>Ability to track environmental volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>-Inf</td>
<td>0</td>
<td>-</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>M2</td>
<td>-5</td>
<td>0</td>
<td>-</td>
<td>Low learning</td>
<td>No</td>
</tr>
<tr>
<td>M3</td>
<td>-4</td>
<td>0</td>
<td>-</td>
<td>High learning</td>
<td>No</td>
</tr>
<tr>
<td>M4</td>
<td>-5</td>
<td>1</td>
<td>0.2</td>
<td>Low learning</td>
<td>Yes</td>
</tr>
<tr>
<td>M5</td>
<td>-4</td>
<td>1</td>
<td>0.2</td>
<td>High learning</td>
<td>Yes</td>
</tr>
</tbody>
</table>

![Graph showing deviations and standard](Image)

Sanchez et al. 2014
3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ Simulations: alternative models and predictions

<table>
<thead>
<tr>
<th>Models</th>
<th>( \omega ) values</th>
<th>( \kappa ) values (if ( \kappa = 0 ), no 3rd level)</th>
<th>( \vartheta ) values</th>
<th>Ability to track events probabilities</th>
<th>Ability to track environmental volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>-Inf</td>
<td>0</td>
<td>-</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>M2</td>
<td>-5</td>
<td>0</td>
<td>-</td>
<td>Low learning</td>
<td>No</td>
</tr>
<tr>
<td>M3</td>
<td>-4</td>
<td>0</td>
<td>-</td>
<td>High learning</td>
<td>No</td>
</tr>
<tr>
<td>M4</td>
<td>-5</td>
<td>1</td>
<td>0.2</td>
<td>Low learning</td>
<td>Yes</td>
</tr>
<tr>
<td>M5</td>
<td>-4</td>
<td>1</td>
<td>0.2</td>
<td>High learning</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Sanchez et al. 2014
Disentangling between learning models

Simulations: alternative models and predictions

<table>
<thead>
<tr>
<th>Models</th>
<th>( \omega ) values</th>
<th>( \kappa ) values</th>
<th>( \theta ) values</th>
<th>Ability to track events probabilities</th>
<th>Ability to track environmental volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>-Inf</td>
<td>0</td>
<td>-</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>M2</td>
<td>-5</td>
<td>0</td>
<td>-</td>
<td>Low learning</td>
<td>No</td>
</tr>
<tr>
<td>M3</td>
<td>-4</td>
<td>0</td>
<td>-</td>
<td>High learning</td>
<td>No</td>
</tr>
<tr>
<td>M4</td>
<td>-5</td>
<td>1</td>
<td>0.2</td>
<td>Low learning</td>
<td>Yes</td>
</tr>
<tr>
<td>M5</td>
<td>-4</td>
<td>1</td>
<td>0.2</td>
<td>High learning</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Sanchez et al. 2014
3 – Online Bayesian inference and design optimization

Disentangling between learning models

Simulations: alternative models and predictions

<table>
<thead>
<tr>
<th>Models</th>
<th>$\omega$ values</th>
<th>$\kappa$ values $\theta$ values</th>
<th>Ability to track events probabilities</th>
<th>Ability to track environmental volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>-Inf</td>
<td>0</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>M2</td>
<td>-5</td>
<td>0</td>
<td>Low learning</td>
<td>No</td>
</tr>
<tr>
<td>M3</td>
<td>-4</td>
<td>0</td>
<td>High learning</td>
<td>No</td>
</tr>
<tr>
<td>M4</td>
<td>-5</td>
<td>1</td>
<td>Low learning</td>
<td>Yes</td>
</tr>
<tr>
<td>M5</td>
<td>-4</td>
<td>1</td>
<td>High learning</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Deviant

Standard

\[ y = \text{data (a.u.)} \]

Trials

Sanchez et al. 2014
3 – Online Bayesian inference and design optimization

Disentangling between learning models

- Simulations: alternative models and predictions

<table>
<thead>
<tr>
<th>Models</th>
<th>(\omega) values</th>
<th>(\kappa) values ((\text{if } \kappa = 0, \text{no 3rd level}))</th>
<th>(\vartheta) values</th>
<th>Ability to track events probabilities</th>
<th>Ability to track environmental volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>-Inf</td>
<td>0</td>
<td>-</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>M2</td>
<td>-5</td>
<td>0</td>
<td>-</td>
<td>Low learning</td>
<td>No</td>
</tr>
<tr>
<td>M3</td>
<td>-4</td>
<td>0</td>
<td>-</td>
<td>High learning</td>
<td>No</td>
</tr>
<tr>
<td>M4</td>
<td>-5</td>
<td>1</td>
<td>0.2</td>
<td>Low learning</td>
<td>Yes</td>
</tr>
<tr>
<td>M5</td>
<td>-4</td>
<td>1</td>
<td>0.2</td>
<td>High learning</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Simulations: alternative models and predictions
Disentangling between learning models

Static designs to be compared with ADO
3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ Results on Monte Carlo simulations

% of non-conclusive experiments

Trials before reaching conclusion

Sanchez et al. 2014
Conclusion

• **1 - EEG/MEG Source reconstruction**
  - Bayesian inference on models and parameters
  - Group level inference
  - EEG/MEG data fusion

• **2 - Inferring learning mechanisms through EEG/MEG responses**
  - meta-Bayesian inference
  - Model comparison

• **3 - Online Bayesian inference and design optimization**
Acknowledgments

Olivier Bertrand
Gareth Barnes
Anne Caclin
Jean Daunizeau
Karl Friston
Rik Henson
Françoise Lecaignard
Emmanuel Maby
Gaëtan Sanchez