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"

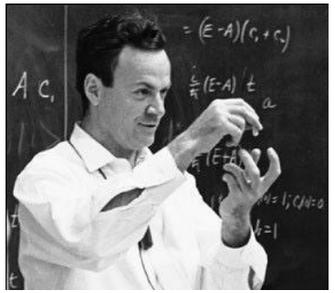
Jacques Hadamard *(mathematician, 1865-1963)*

Workshop on Bayesian approaches in neuroscience – Karolinska Institute – 2nd of Frebruary 2016

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

<u>Ex. Dynamic causal models (DCM)</u>

Hidden states

Unknown parameters

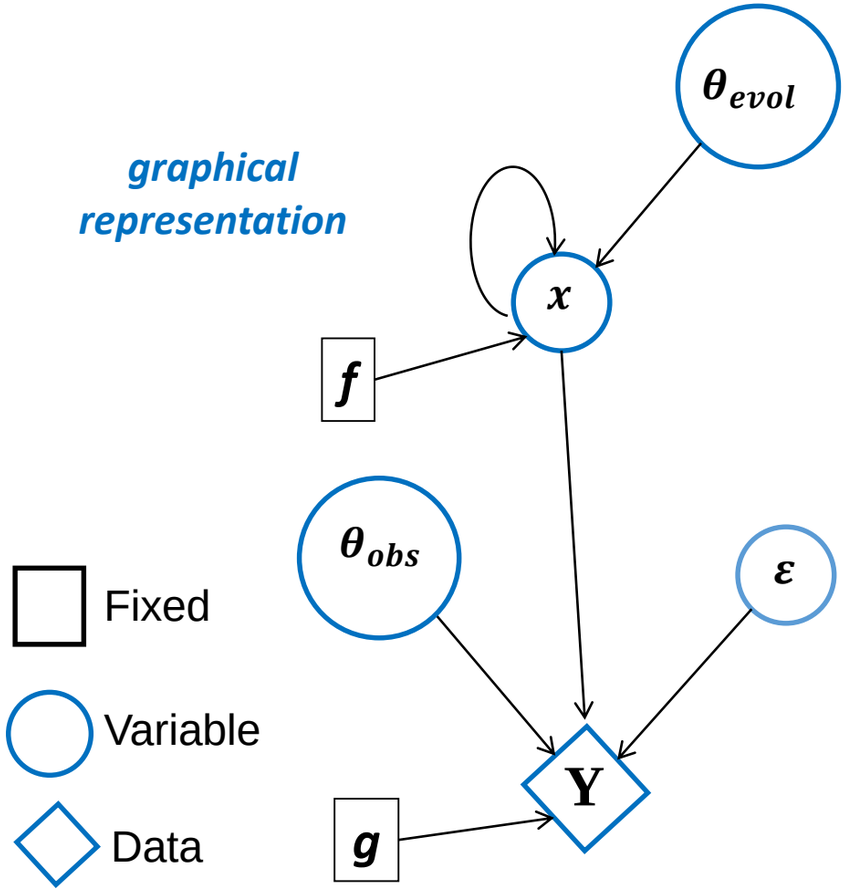
evolution $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\theta}_{evol}, \boldsymbol{u})$

observation $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\theta}_{obs}) + \boldsymbol{\varepsilon}$

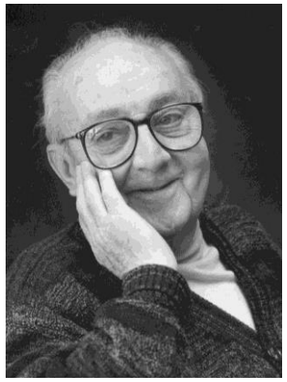
Inputs

Observations (data features)

graphical representation



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

Model M : your scientific hypothesis put into predictive equations (functions *f* and *g* + prior assumptions)

Likelihood

priors

$$P(\theta|Y,M)=\frac{P(Y|\theta,M)P(\theta|M)}{\boxed{P(Y|M)}}$$

Evidence

- 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

Most plausible model or model family

$$\boxed{P(\theta|Y,M)} = \frac{P(Y|\theta,M)P(\theta|M)}{P(Y|M)}$$

Posterior

- Inference at the group level

Outline

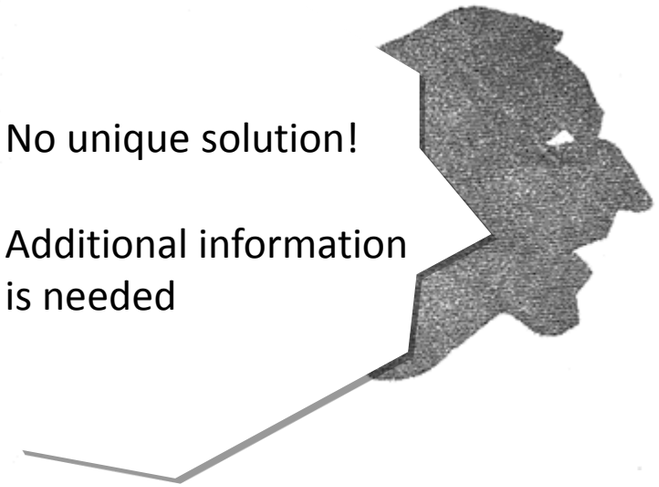
- **1 - EEG/MEG Source reconstruction**
- **2 - Inferring learning mechanisms from EEG/MEG responses**
- **3 - Online Bayesian inference and design optimization**

1 – EEG/MEG source reconstruction

An ill-posed inverse problem

No unique solution!

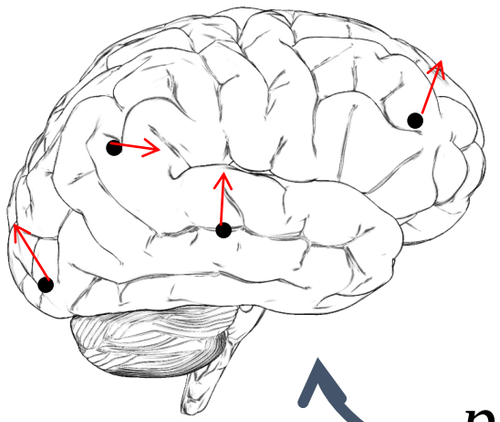
Additional information is needed



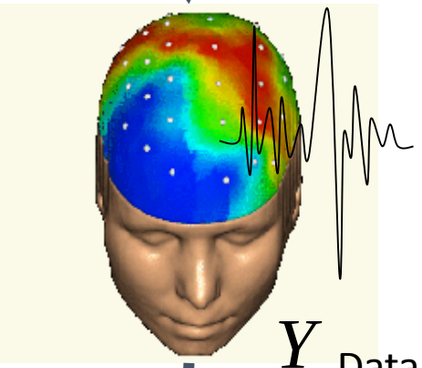
Generative model M

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

$p(\theta \mid M)$ Prior



θ Parameters



Y Data

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

Evidence $p(Y \mid M)$

Inverse Problem

Bayesian inference

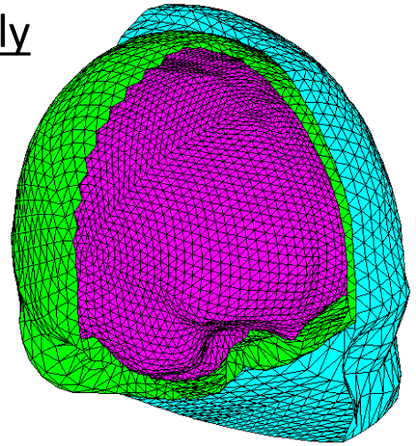
1 – EEG/MEG source reconstruction

Typical model families

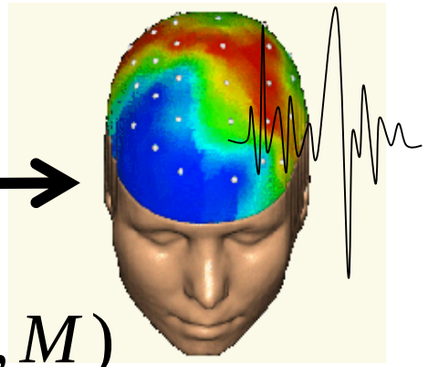
- **Biophysical model**

Static (forward) model family

Geometrical and physical properties of the head tissues



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

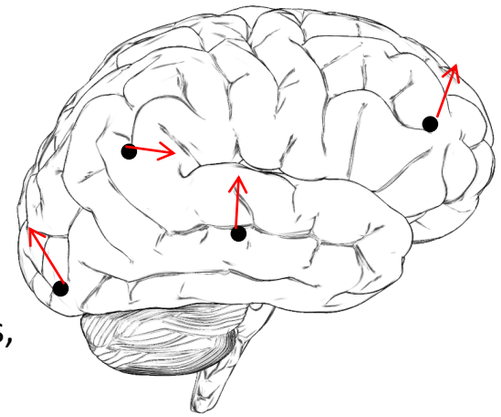


Y: sensor data (incl. noise)

- **Source model**

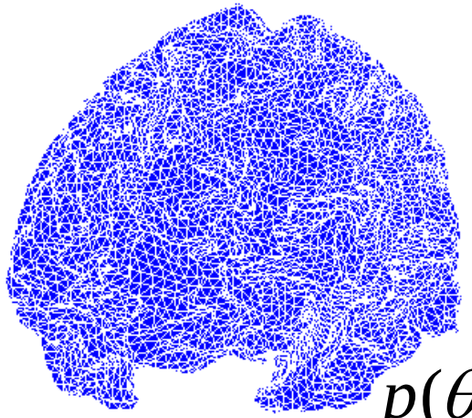
Equivalent current dipole (ECD) model family

θ : dipole locations, orientations & intensity



$p(\theta \mid M)$: number of dipoles, initial location

Distributed sources or Imaging model family

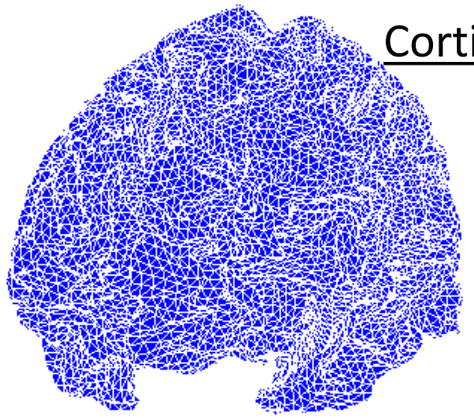


θ : dipole intensities

$p(\theta \mid M)$: number of dipoles, fixed locations

1 – EEG/MEG source reconstruction

Detailed example: the distributed source model

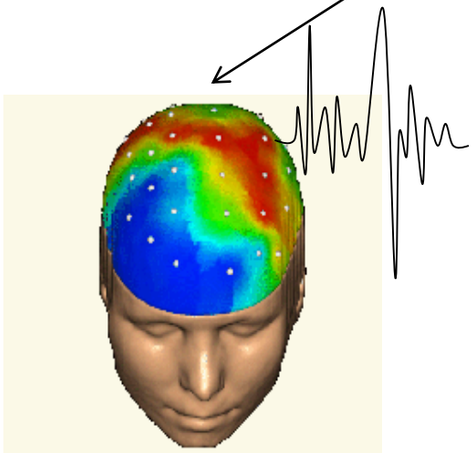


Cortical mesh (3D surface)

θ : dipole intensities

$$\boldsymbol{Y} = \boldsymbol{K}.\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

Additive noise



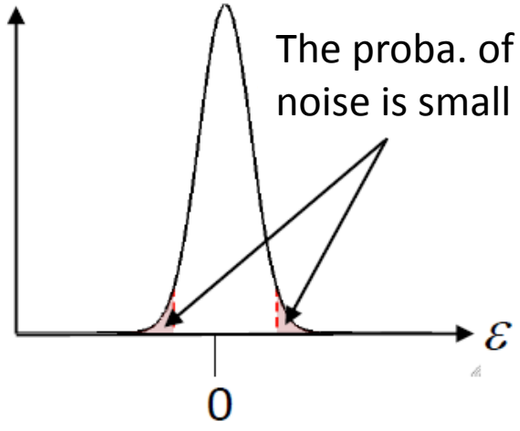
Unknown source dynamics: Nsources x Time

Forward operator or lead-field matrix: Nsens x Nsources

Evoked response: Nsens x Time

- **Likelihood**

$$\boldsymbol{p}(\boldsymbol{\varepsilon}|\boldsymbol{M}) = \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{\varepsilon}})$$



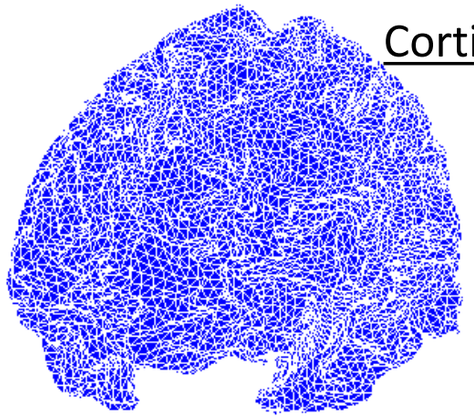
$$\boldsymbol{p}(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{M}) = \boldsymbol{N}(\boldsymbol{K}.\boldsymbol{\theta}, \boldsymbol{C}_{\boldsymbol{\varepsilon}})$$

- **Prior**

$$\boldsymbol{p}(\boldsymbol{\theta}|\boldsymbol{M}) = \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{\theta}})$$

1 – EEG/MEG source reconstruction

Detailed example: the distributed source model



Cortical mesh (3D surface)

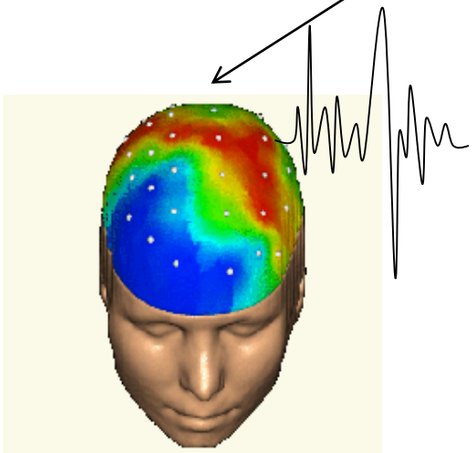
θ : dipole intensities

$$\boldsymbol{Y} = \boldsymbol{K}.\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

Additive noise

Unknown source dynamics: Nsources x Time

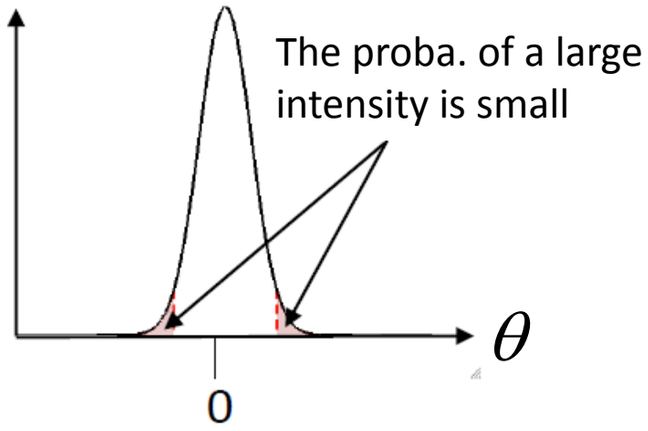
Forward operator or lead-field matrix: Nsens x Nsources



Evoked response: Nsens x Time

- **Likelihood**

$$\boldsymbol{p}(\boldsymbol{\varepsilon}|\boldsymbol{M}) = \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{\varepsilon}})$$



- **Prior**

$$\boldsymbol{p}(\boldsymbol{\theta}|\boldsymbol{M}) = \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{\theta}})$$

1 – EEG/MEG source reconstruction

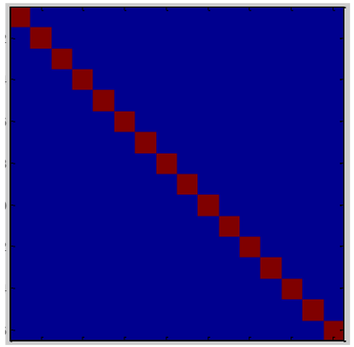
Philips et al. 2005
Mattout et al. 2005 & 2006
Wipf and Nagarajan 2009

Commonly used priors

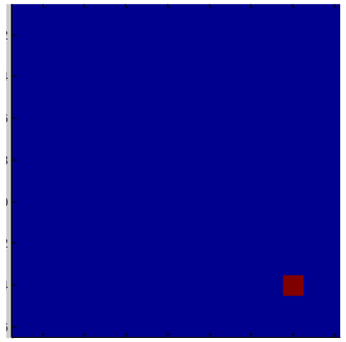
- **Priors**

$$p(\theta|M) = N(0, C_\theta)$$

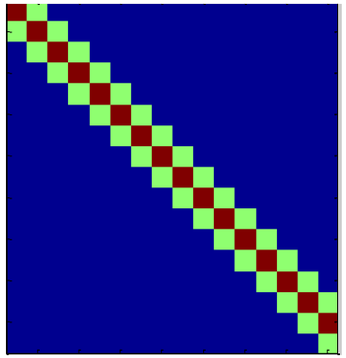
Nsources x Nsources



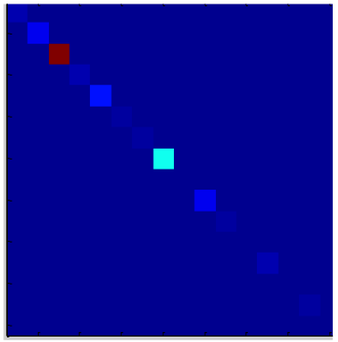
i.i.d or Minimum norm



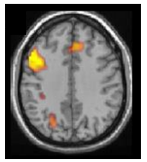
Single dipole

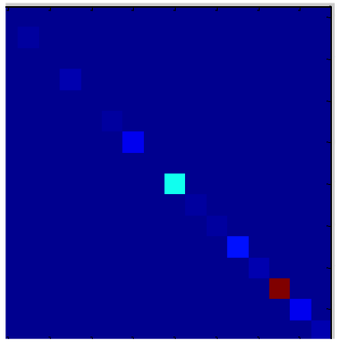


Smoothness (like LORETA)



fMRI based



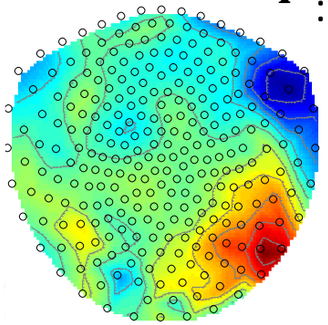


Lead-field based
(Beamformer or MSP)

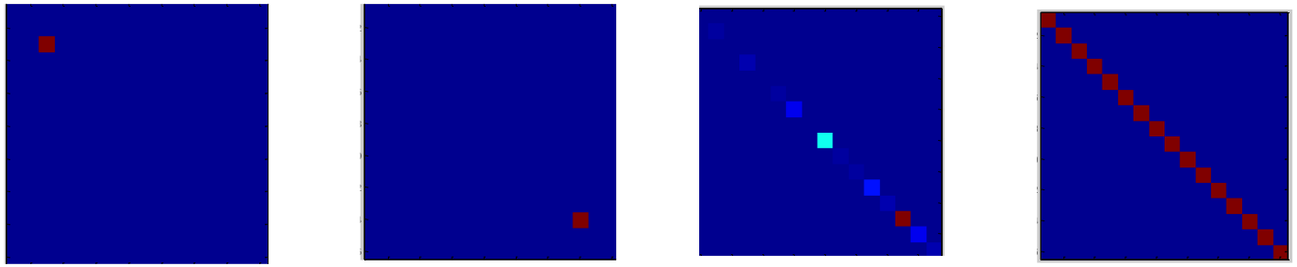
1 – EEG/MEG source reconstruction

Comparing priors

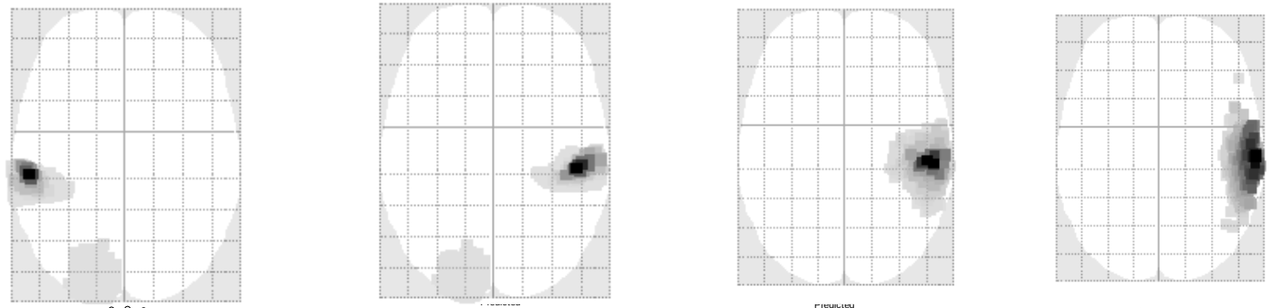
$\boldsymbol{Y}$: observed data



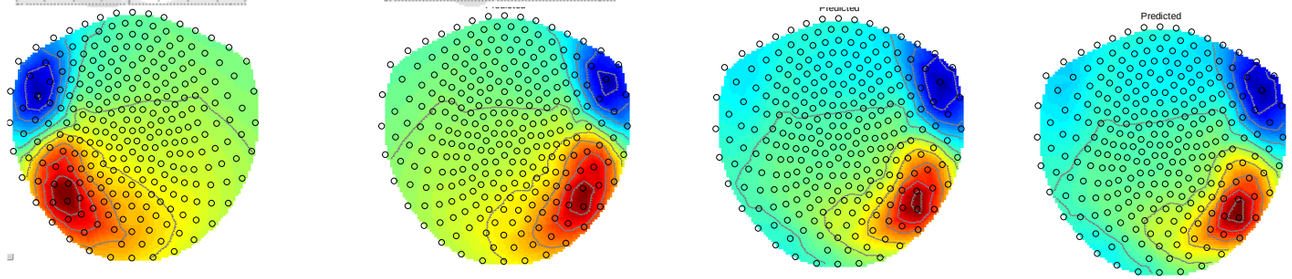
Alternative priors



$\hat{\boldsymbol{\theta}}$: estimated source intensities



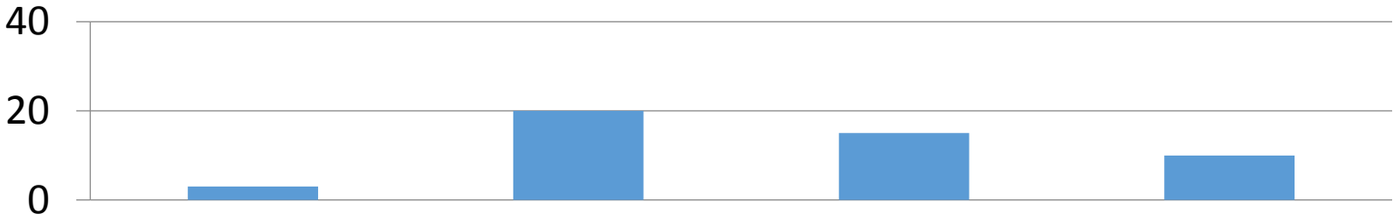
$\hat{\boldsymbol{Y}}$: predicted data



Explained variance

11 % 96% 97% 98%

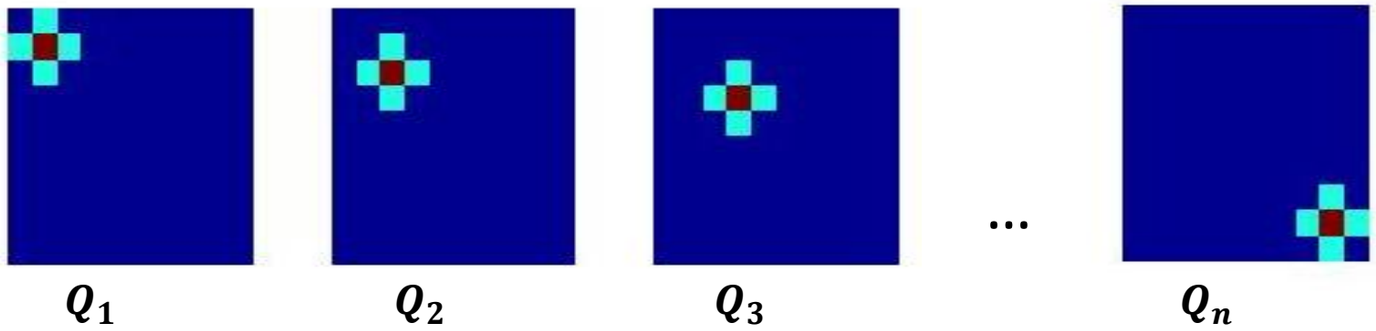
Log-evidence



1 – EEG/MEG source reconstruction

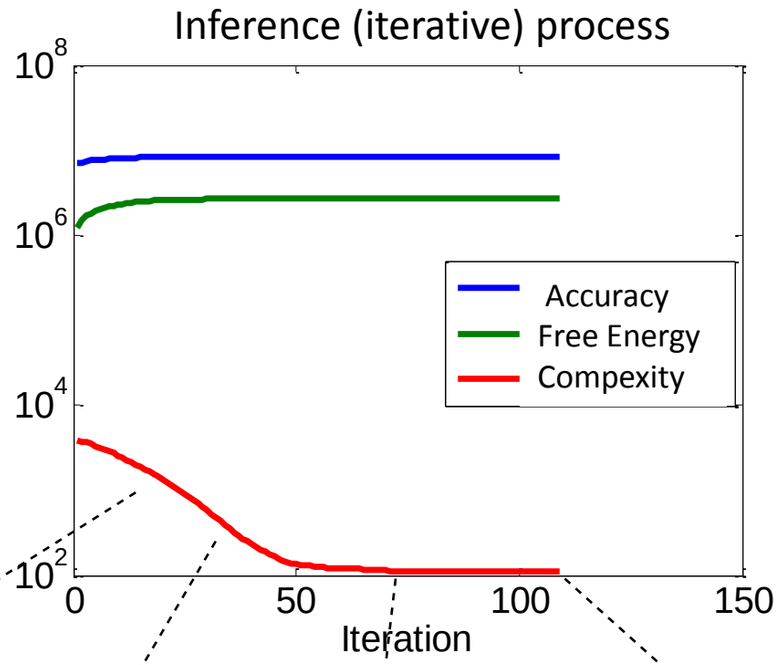
Empirical Bayes: multiple sparse priors (MSP)

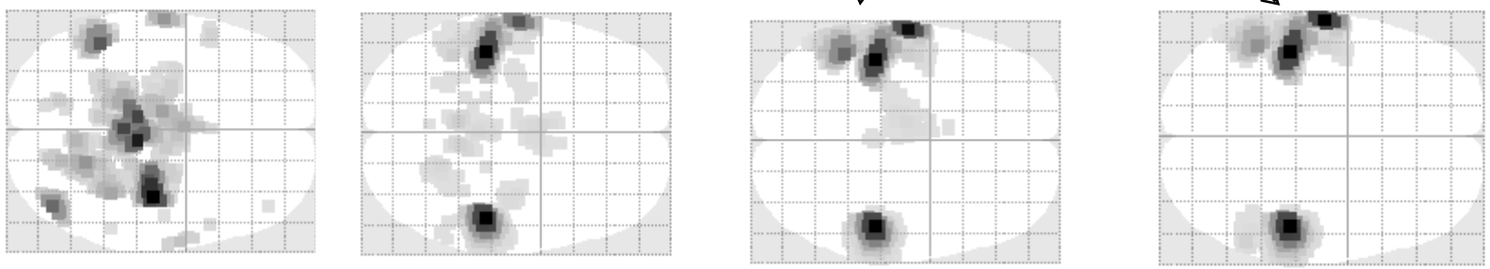
Friston et al. 2008



Model (Hyper)parameters

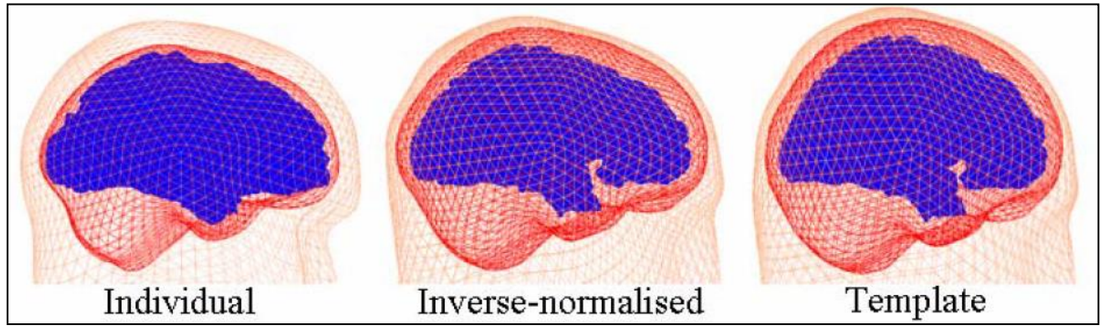
$$C_\theta = \lambda_1 . Q_1 + \lambda_2 . Q_2 + \lambda_3 . Q_3 + \cdots + \lambda_n . Q_n$$





1 – EEG/MEG source reconstruction

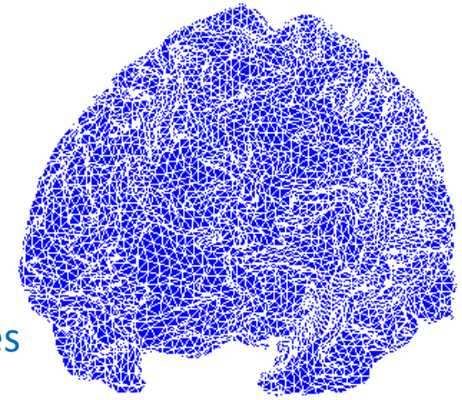
Model comparison on other dimensions



Anatomy

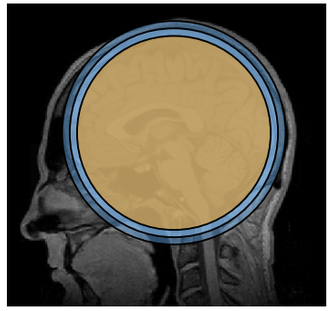
Mattout et al. 2007

Sources

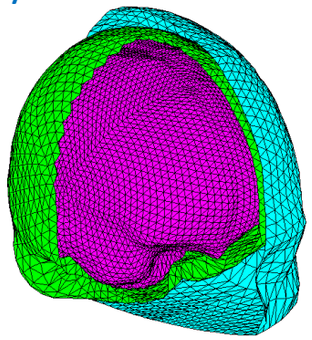


Mesh resolution (High / Low)
Dipole oreintation (fixed / free)

Biophysics



Spherical head model



Realistic surfacic model (BEM)

Model comparison on real MEG data →
Henson et al. 2007

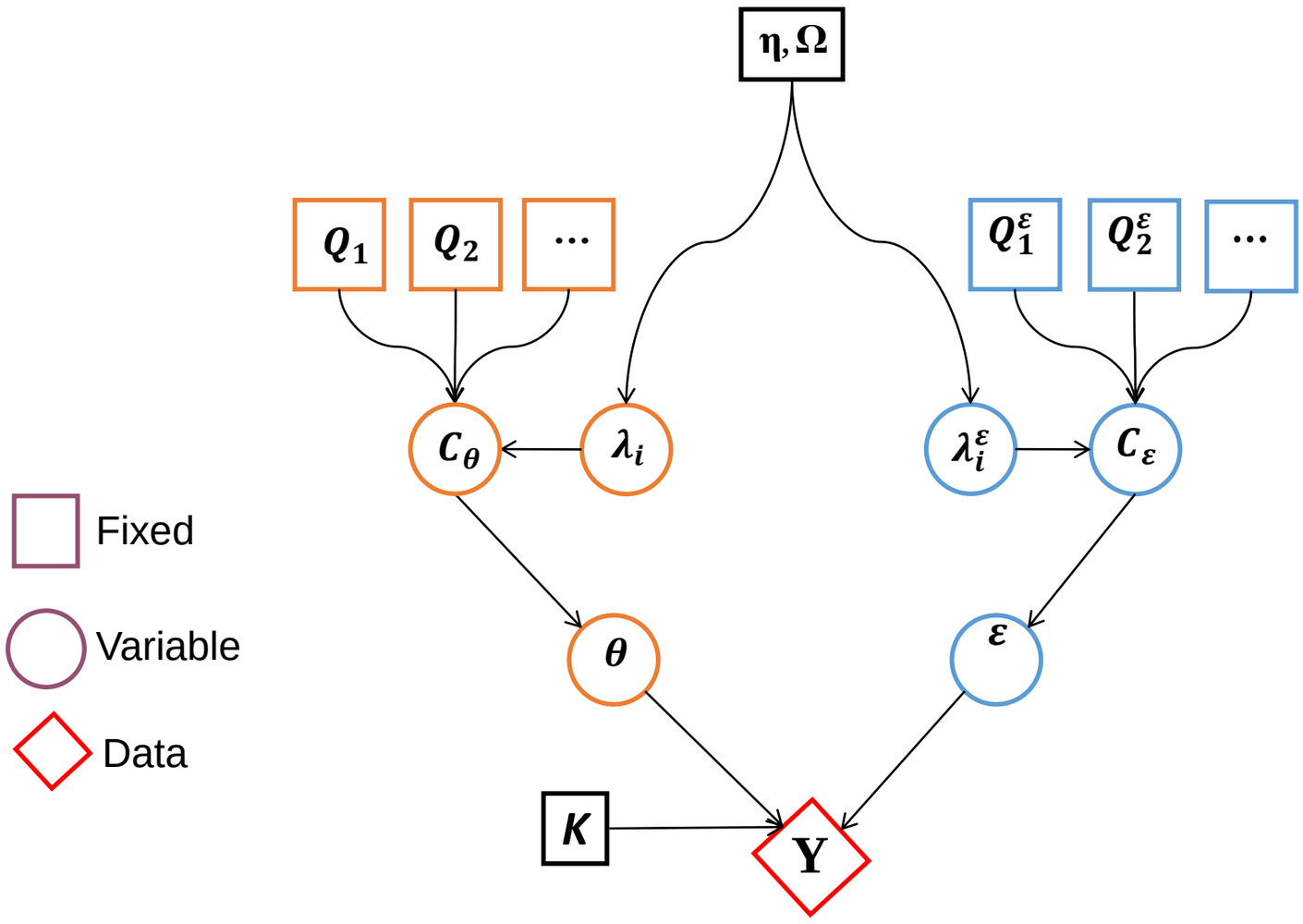
- **No evidence in favor of individual vs. Inverse-norm mesh**
- **Evidence in favor of BEM head model**
- **Evidence in favor of high + fixed vs. low + free**

Henson et al. 2009

1 – EEG/MEG source reconstruction

Graphical representation of MSP

Directed Acyclic Graph (DAG) representation



Friston et al. 2008

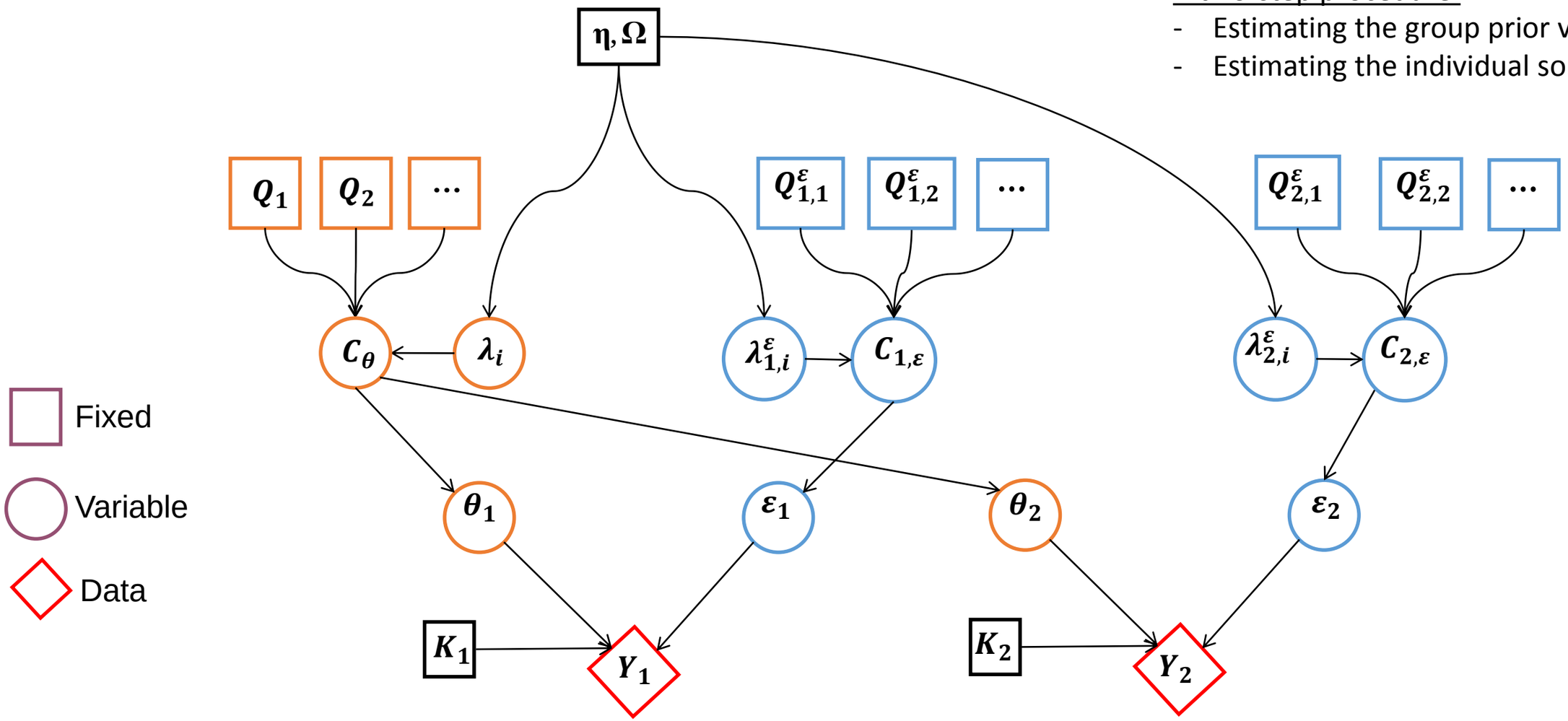
1 – EEG/MEG source reconstruction

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

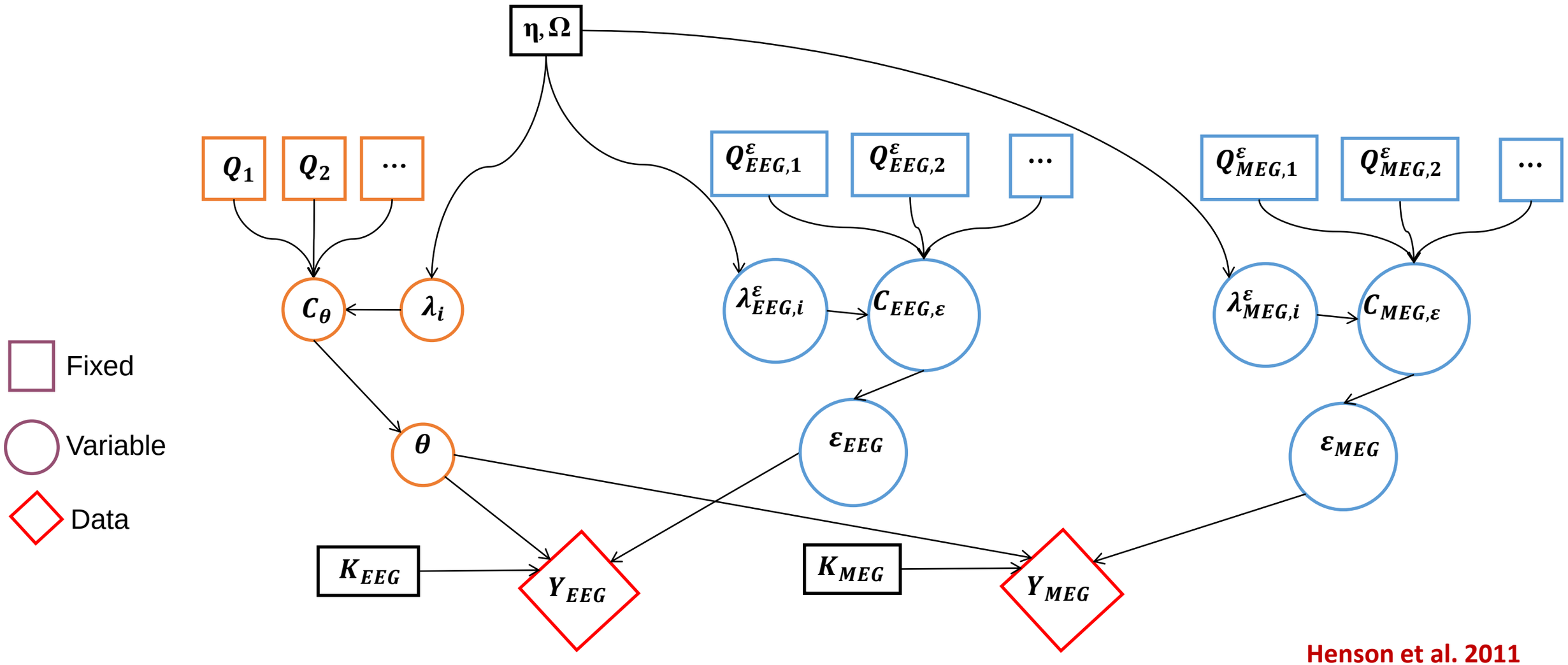


Litvak and Friston 2008

1 – EEG/MEG source reconstruction

Model extension to EEG-MEG data fusion

Directed Acyclic Graph (DAG) representation



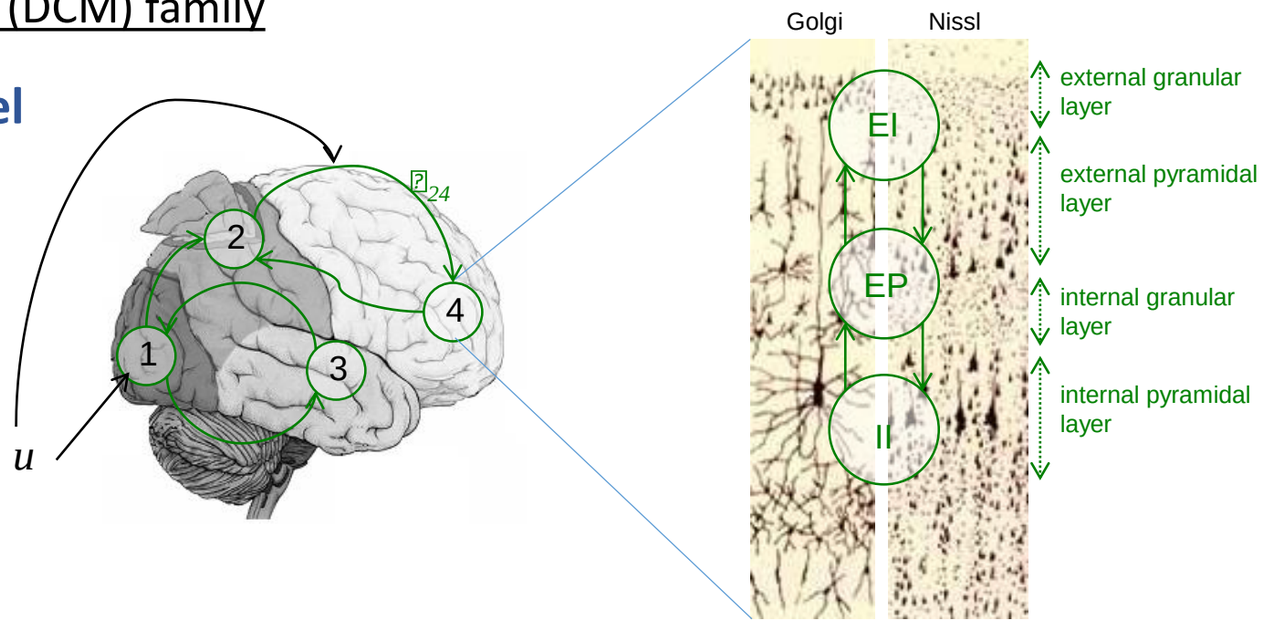
Henson et al. 2011

1 – EEG/MEG source reconstruction

Another model family

Dynamic causal model (DCM) family

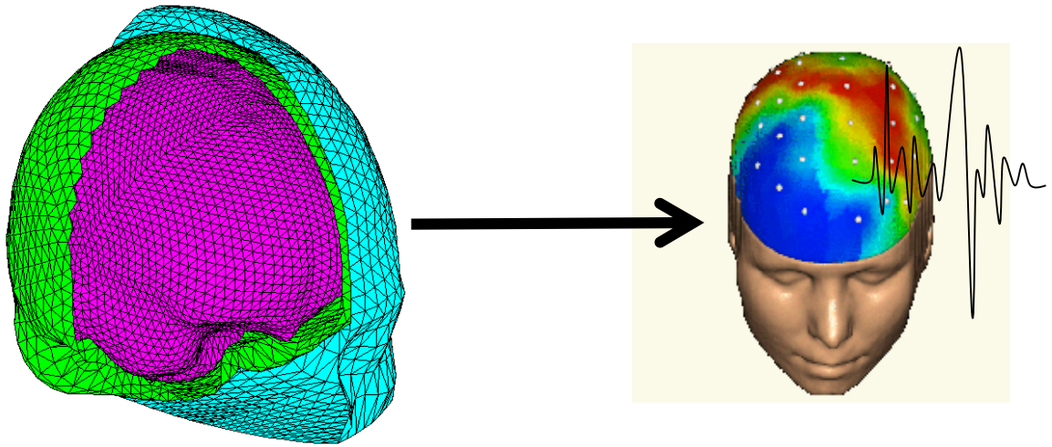
- **Source model**



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

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

- **Biophysical model**



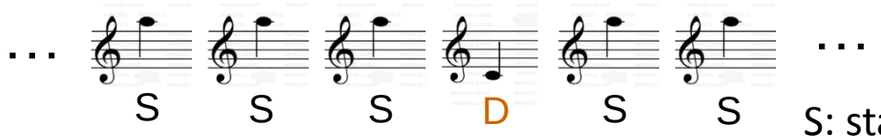
$$y = g(x, \theta_{obs}) + \varepsilon$$

1 – EEG/MEG source reconstruction

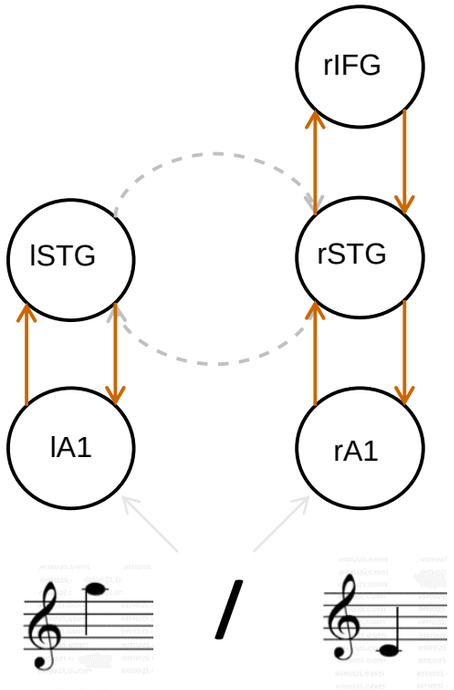
An MMN example

Characterizing a group response in terms of modulations of effective connectivity

Classical auditory oddball paradigm



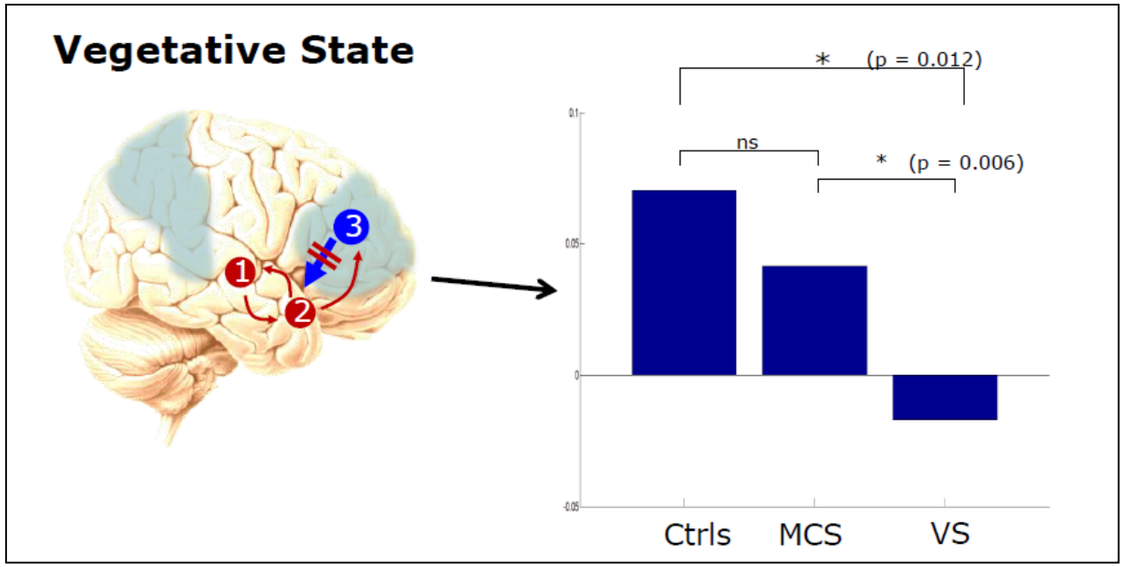
S: standard tone
D: (frequency) deviant tone



Garrido et al. 2009

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

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

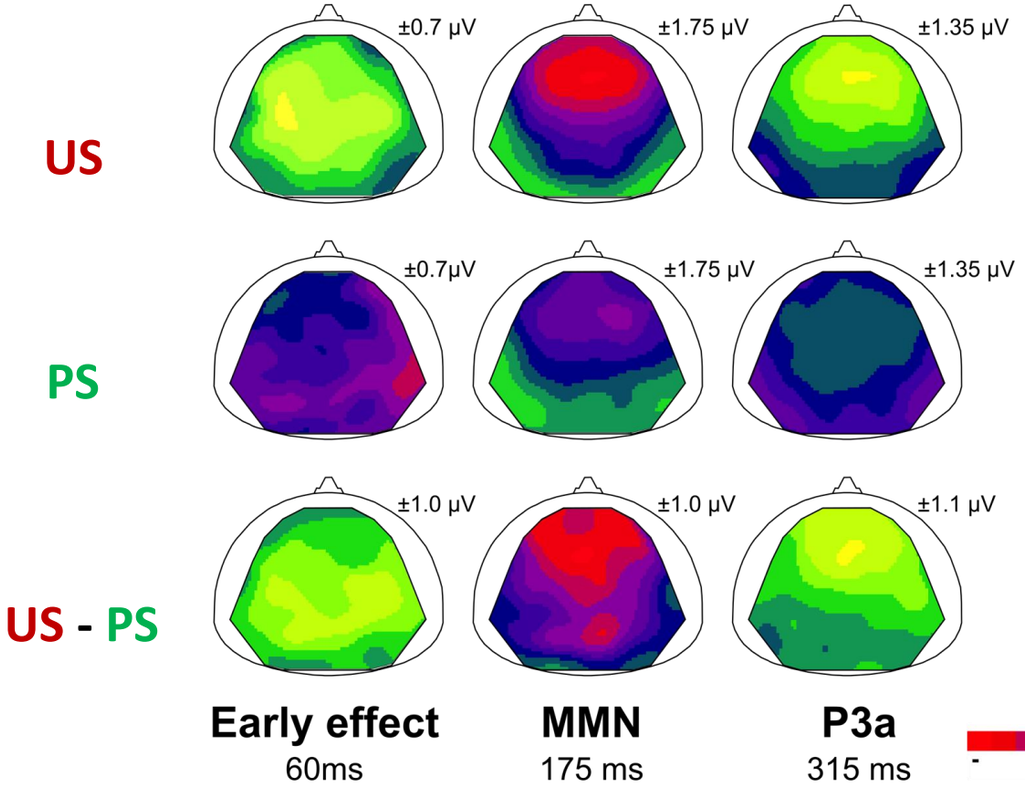
Neurophysiological correlates of implicit perceptual learning

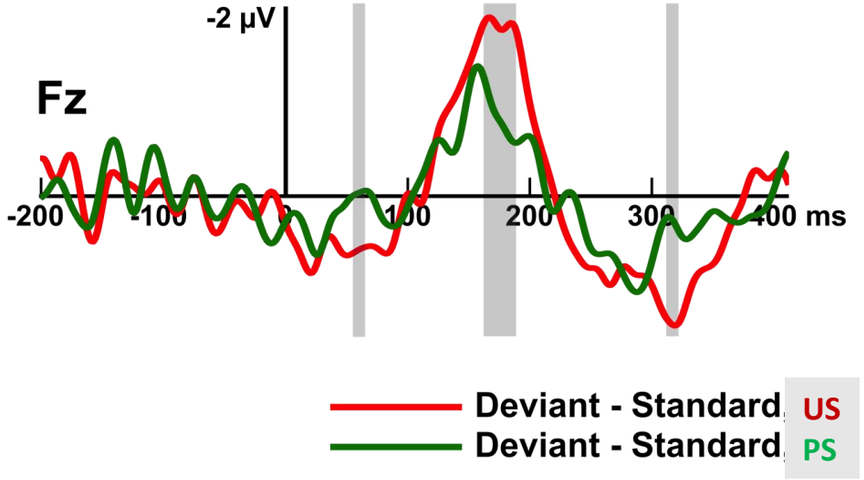
- **Deviance responses should be modulated by predictability**

Effect of predictability

US: Unpredictable sequence
PS: Predictable sequence

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

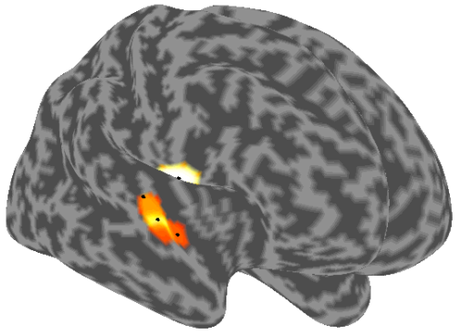
$$p(\mu|m,l) = \frac{\Gamma(m+l)}{\Gamma(m)\Gamma(l)}\mu^m(1-\mu)^l$$

Ostwald et al. 2012

Bayesian Surprise

$$= \mathrm{KL}\left(p\left(\mu|m_{w_{n-1}}, l_{w_{n-1}}\right)\|p\left(\mu|m_{w_n}, l_{w_n}\right)\right)$$

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



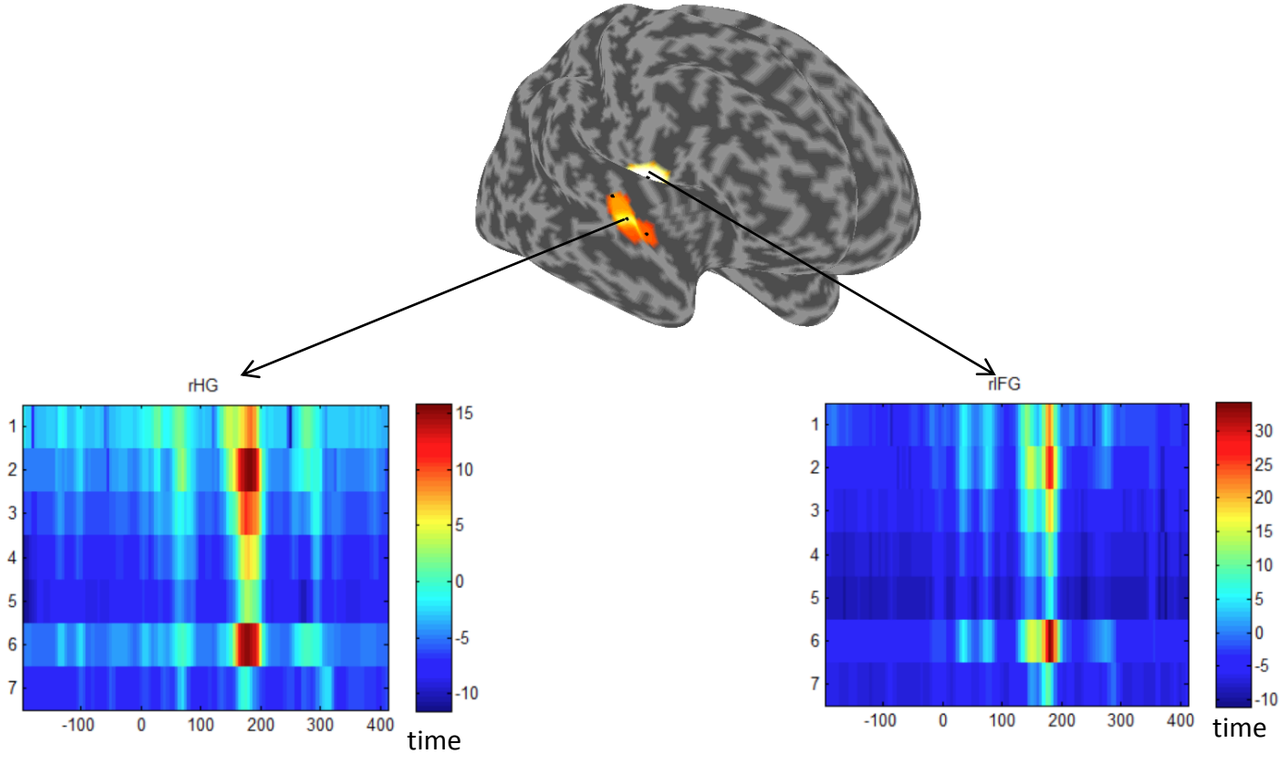
Lecaignard et al. in prep (a)

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**)



Relative Free-energies or Bayes Factors w.r.t null model (M8)

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

Lecaignard et al. in prep (b)
VBA toolbox, Daunizeau et al. 2014

3 – Online Bayesian inference and design optimization

3 – 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?

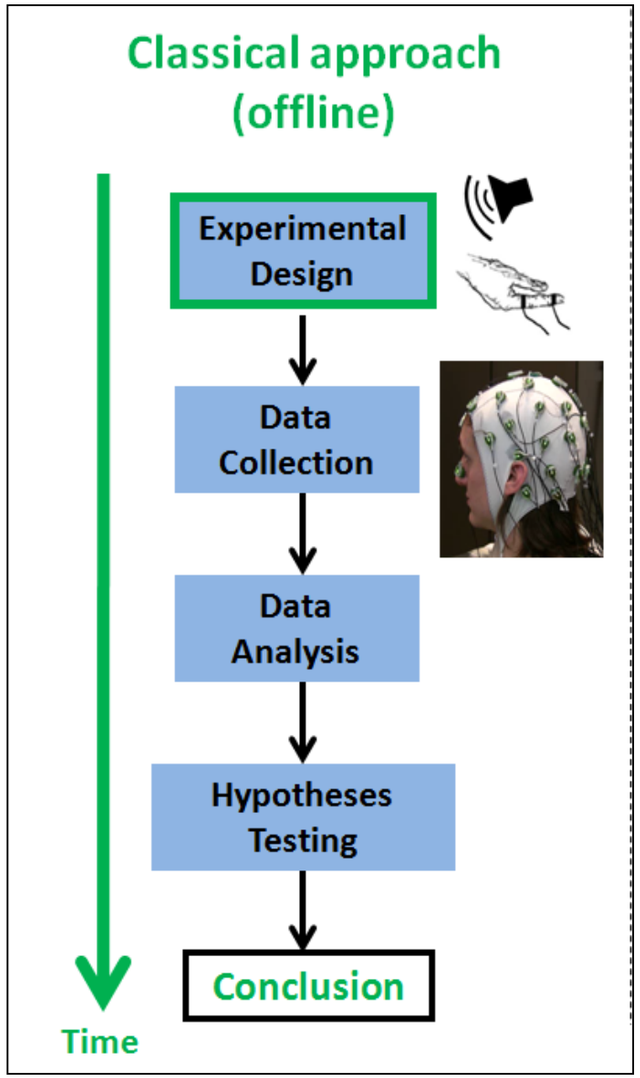
Basic idea

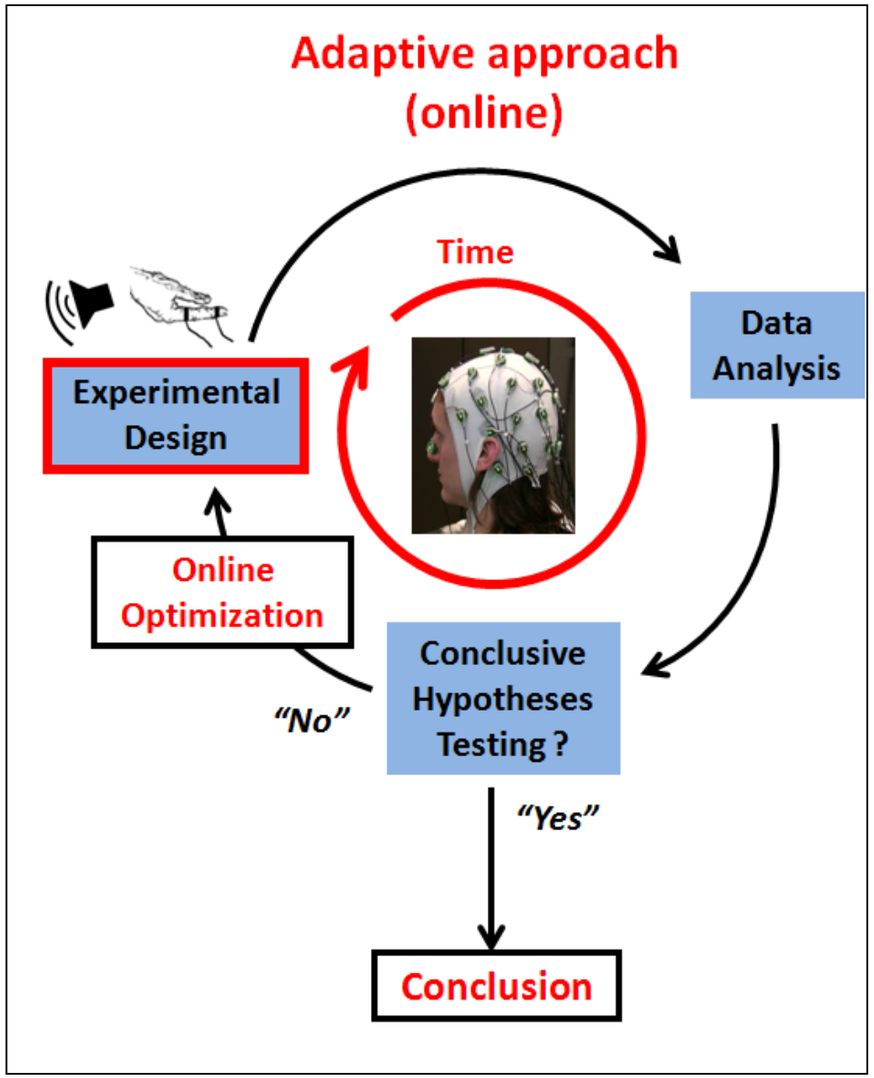
- ✓ To make use of our ability to process data in real-time in order to optimize design parameters online, for a given individual, in the aim of selecting the most likely model both accurately and fast

Wald 1947, Cavagnaro et al. 2009, Myung et al. 2013

3 – Online Bayesian inference and design optimization

General principle of Adaptive Design Optimization



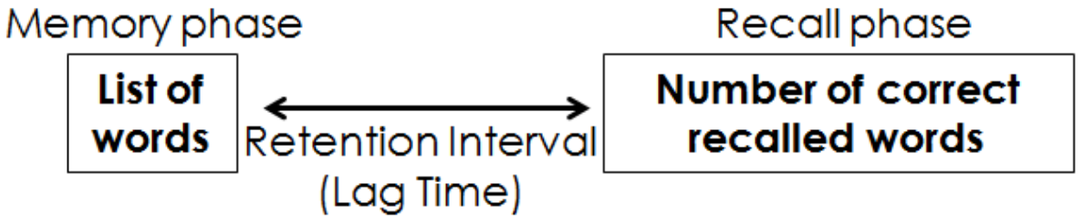


Sanchez et al. 2014

3 – Online Bayesian inference and design optimization

A simple example: behavioural task of word memorization

✓ Experimental context



u (design variable) = Lag time

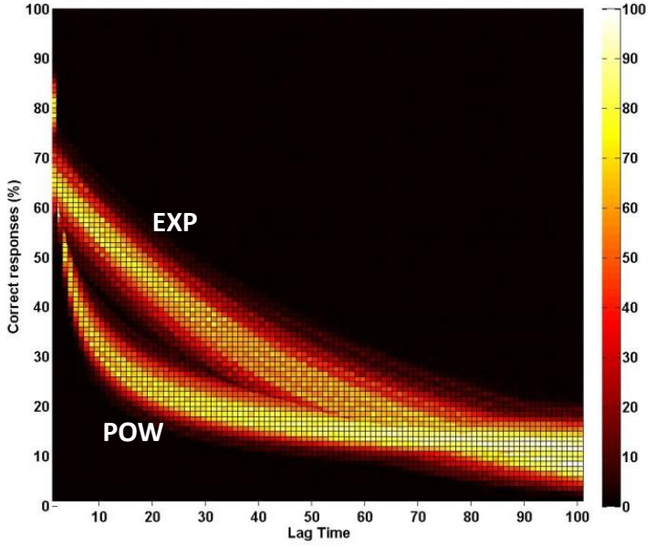
Y (observations) = number of recalled words

✓ Hypothesis

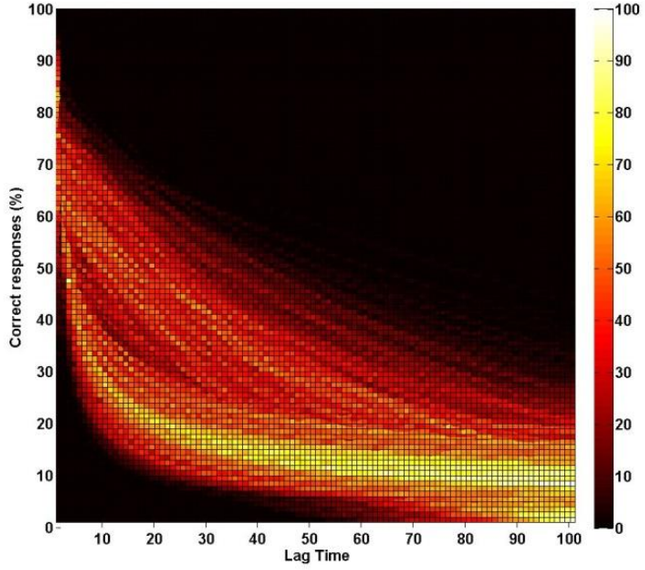
Models of "forgetting"

Model	Equation
Power (POW)	$p = a(t+1)^{-b}$
Exponential (EXP)	$p = ae^{-bt}$

Predictions
(Low uncertainty about model parameters)



Predictions
(High uncertainty about model parameters)

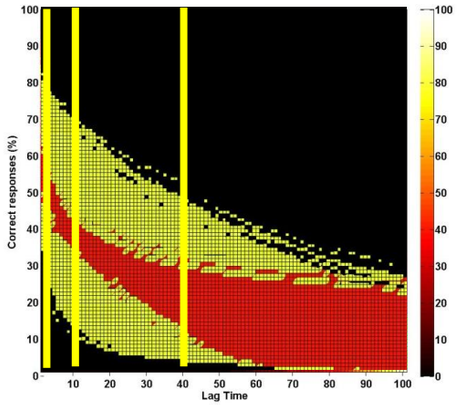


Cavagnaro et al. 2009, Myung et al. 2013

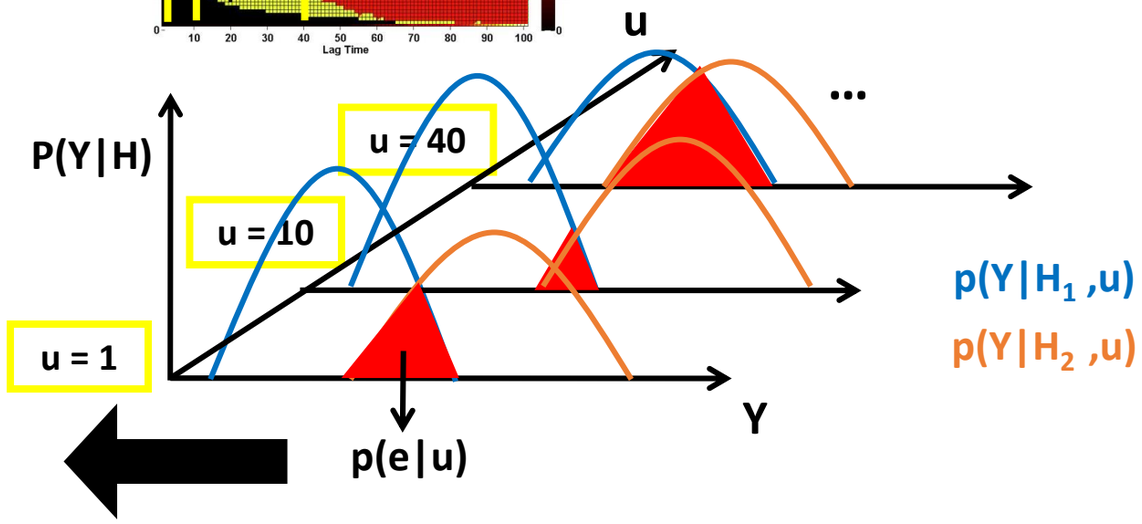
3 – Online Bayesian inference and design optimization

Optimization criterion

✓ **Minimizing the risk of model selection error**



H_1 : EXP
H_2 : POW



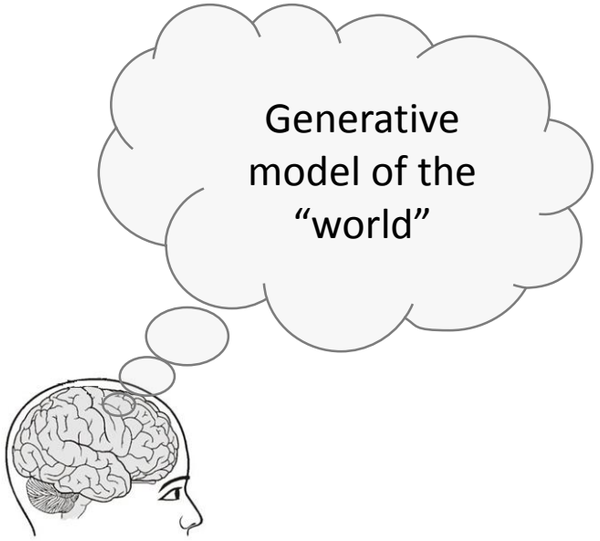
$u^* = \text{argmin}\{p(e|u)\}$

Danizeau et al. 2011

3 – Online Bayesian inference and design optimization

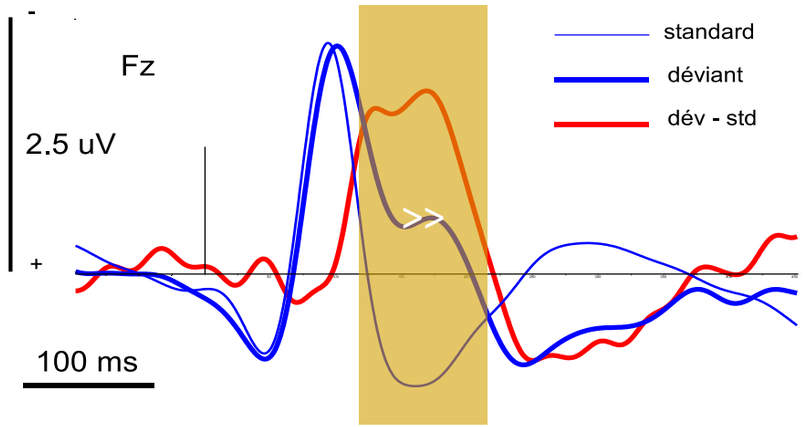
Extension to more complex (DCM-like) models

✓ **Extension to more complex, multidimensional models**



+

ERPs



Evolution function ***f***

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

Observation function ***g***

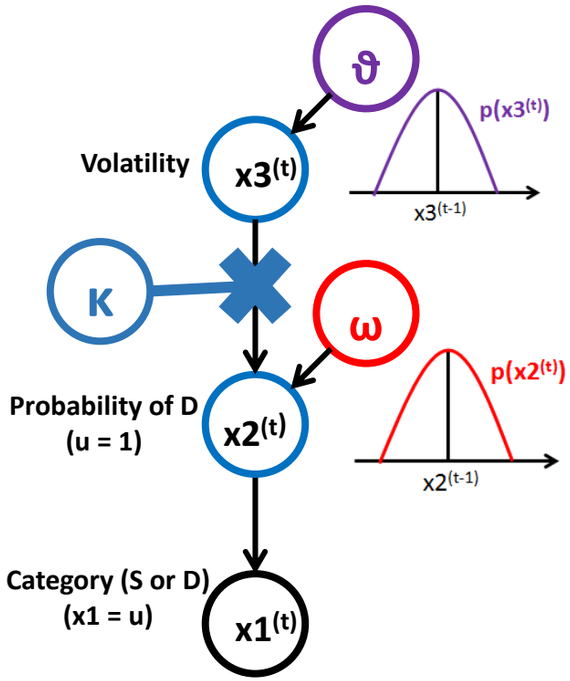
$$y = g(x, \theta_{obs}) + \varepsilon$$

Danizeau et al. 2011
Sanchez et al. 2014

3 – Online Bayesian inference and design optimization

Disentangling between learning models

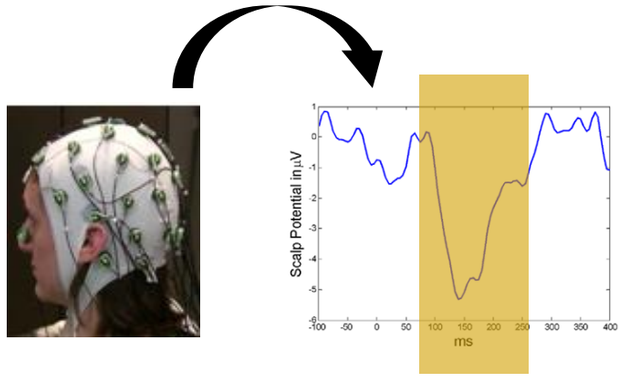
✓ **A perceptual learning model family: hierarchical Gaussian filters**



Evolution function ***f***

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

+



KL-divergence
between prior and post beliefs

Observer function ***g***

$$y = g(x, \theta_{obs}) + \varepsilon$$

Mathis et al. 2011 & 2014

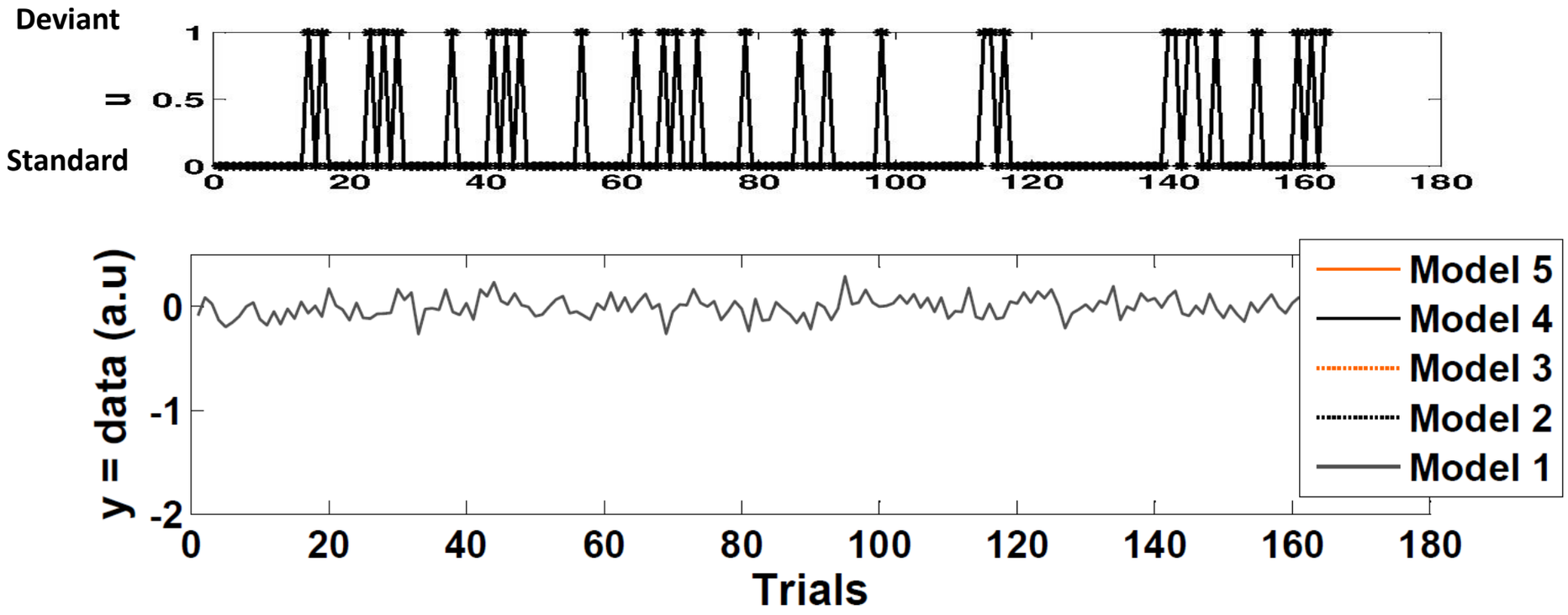
Sanchez et al. 2014

3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ **Simulations: alternative models and predictions**

Models	ω values	κ values (if κ = 0, no 3rd level)	ϑ values	Ability to track events probabilities	Ability to track environmental volatility
M1	-Inf	0	-	No	No
M2	- 5	0	-	Low learning	No
M3	- 4	0	-	High learning	No
M4	- 5	1	0.2	Low learning	Yes
M5	- 4	1	0.2	High learning	Yes



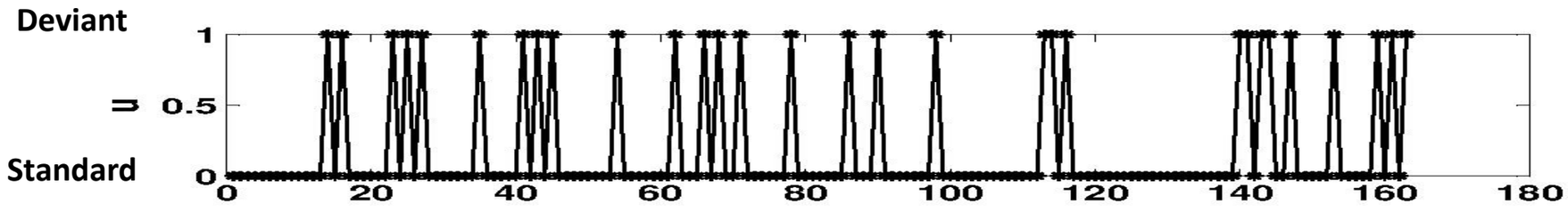
Sanchez et al. 2014

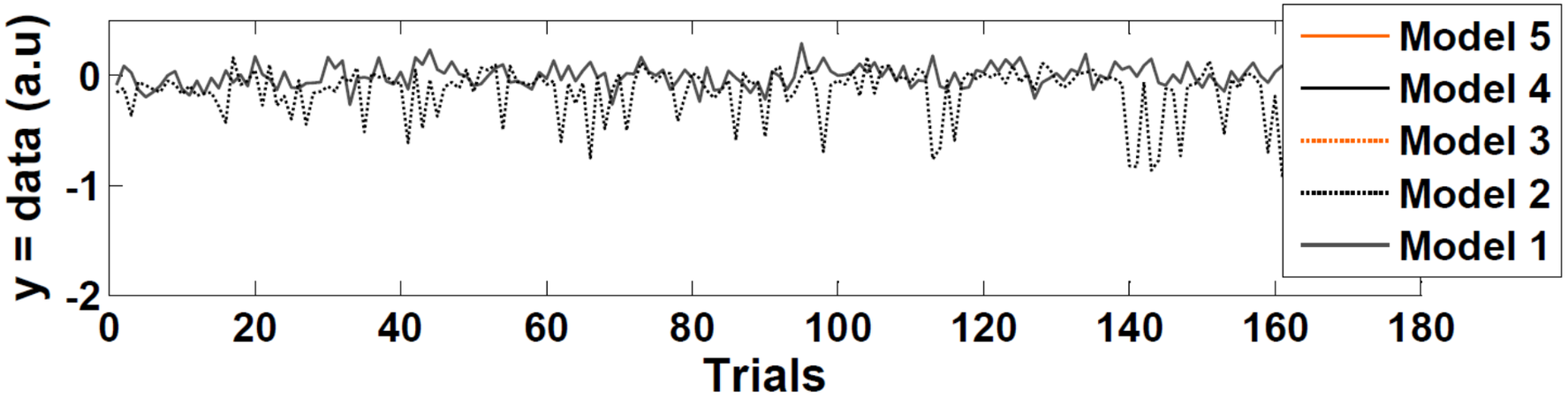
3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ **Simulations: alternative models and predictions**

Models	ω values	κ values (if κ = 0, no 3rd level)	ϑ values	Ability to track events probabilities	Ability to track environmental volatility
M1	-Inf	0	-	No	No
M2	- 5	0	-	Low learning	No
M3	- 4	0	-	High learning	No
M4	- 5	1	0.2	Low learning	Yes
M5	- 4	1	0.2	High learning	Yes





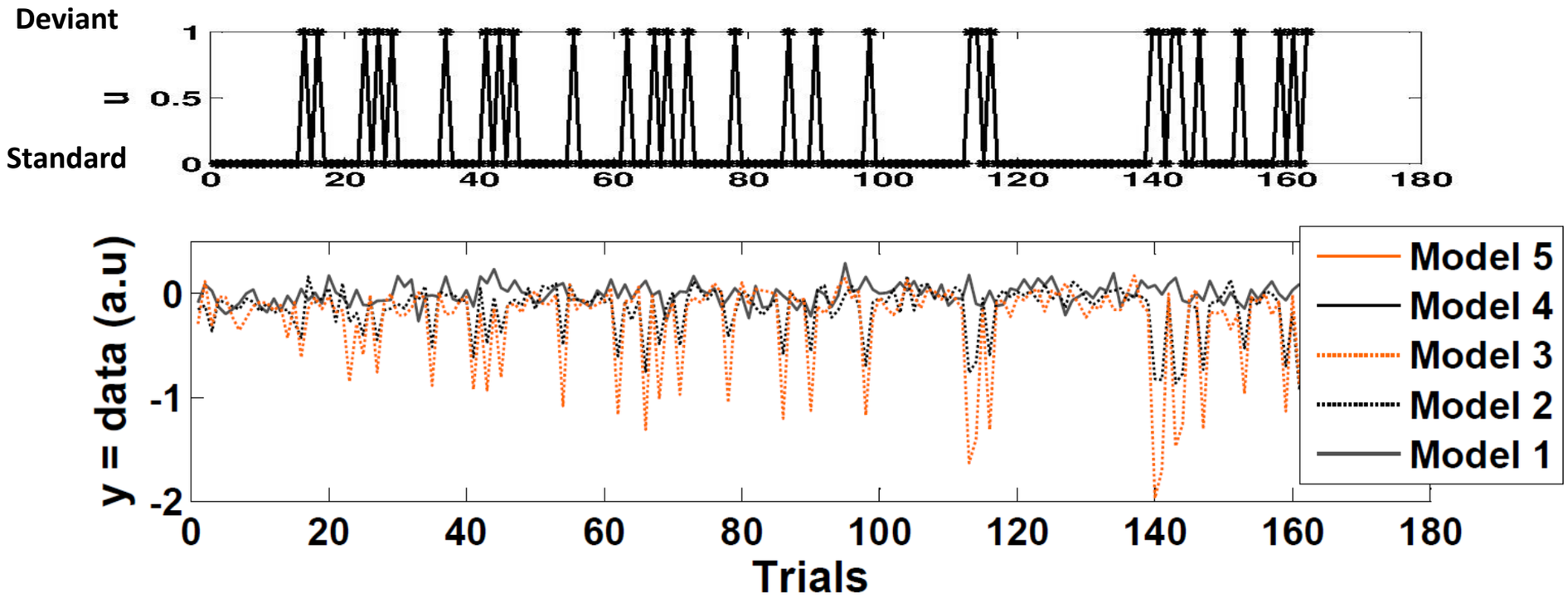
Sanchez et al. 2014

3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ **Simulations: alternative models and predictions**

Models	ω values	κ values (if κ = 0, no 3rd level)	ϑ values	Ability to track events probabilities	Ability to track environmental volatility
M1	-Inf	0	-	No	No
M2	- 5	0	-	Low learning	No
M3	- 4	0	-	High learning	No
M4	- 5	1	0.2	Low learning	Yes
M5	- 4	1	0.2	High learning	Yes



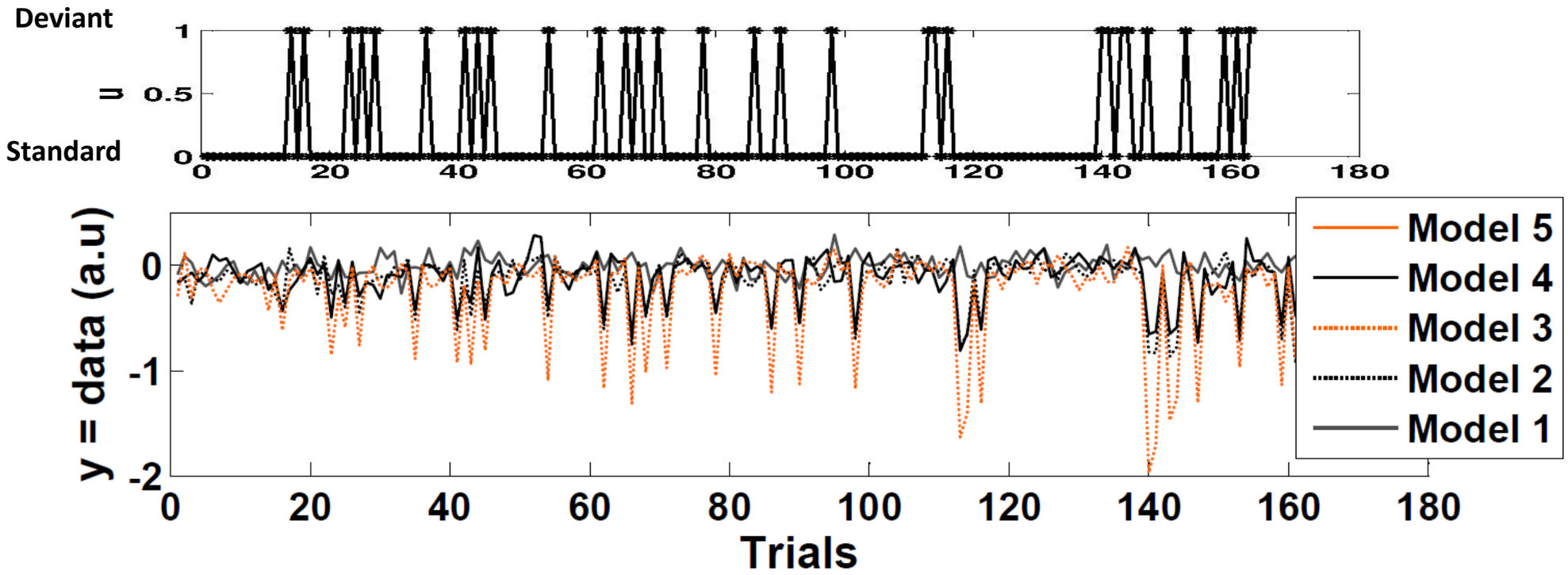
Sanchez et al. 2014

3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ **Simulations: alternative models and predictions**

Models	ω values	κ values (if κ = 0, no 3rd level)	ϑ values	Ability to track events probabilities	Ability to track environmental volatility
M1	-Inf	0	-	No	No
M2	- 5	0	-	Low learning	No
M3	- 4	0	-	High learning	No
M4	- 5	1	0.2	Low learning	Yes
M5	- 4	1	0.2	High learning	Yes



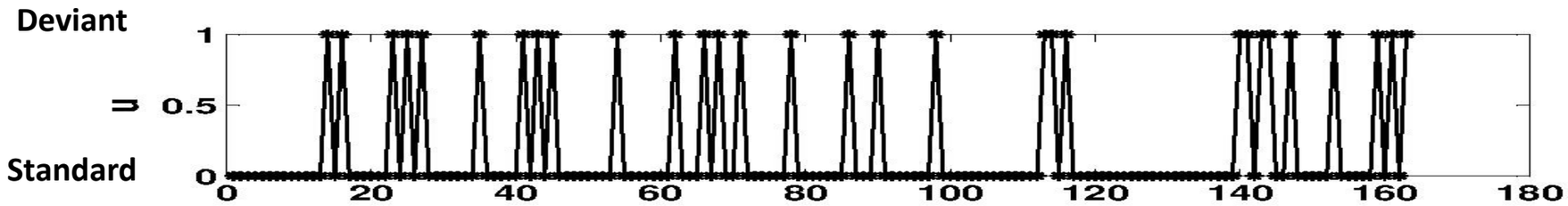
Sanchez et al. 2014

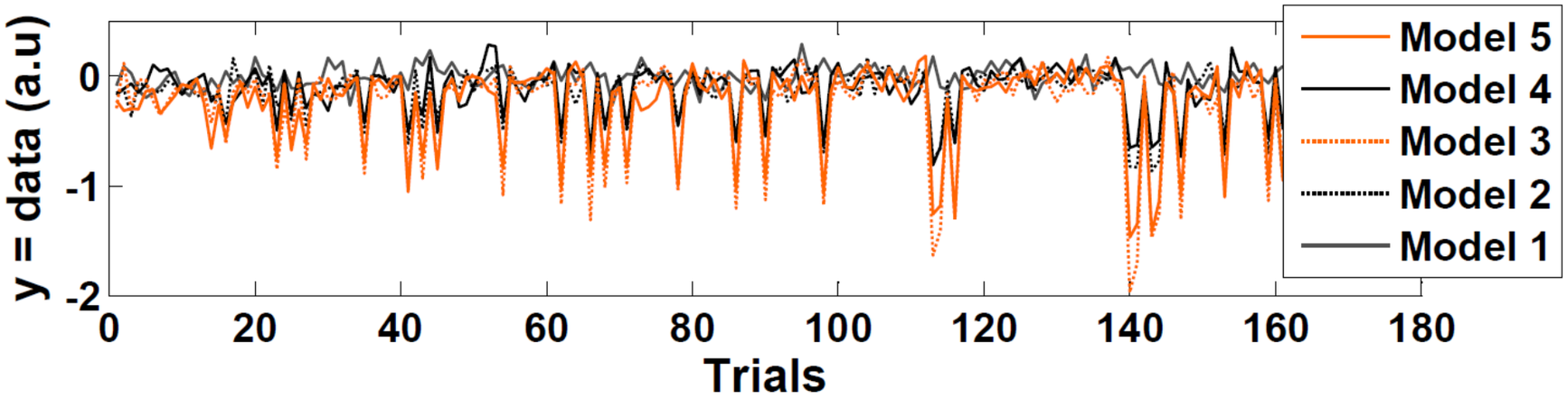
3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ **Simulations: alternative models and predictions**

Models	ω values	κ values (if κ = 0, no 3rd level)	ϑ values	Ability to track events probabilities	Ability to track environmental volatility
M1	-Inf	0	-	No	No
M2	- 5	0	-	Low learning	No
M3	- 4	0	-	High learning	No
M4	- 5	1	0.2	Low learning	Yes
M5	- 4	1	0.2	High learning	Yes



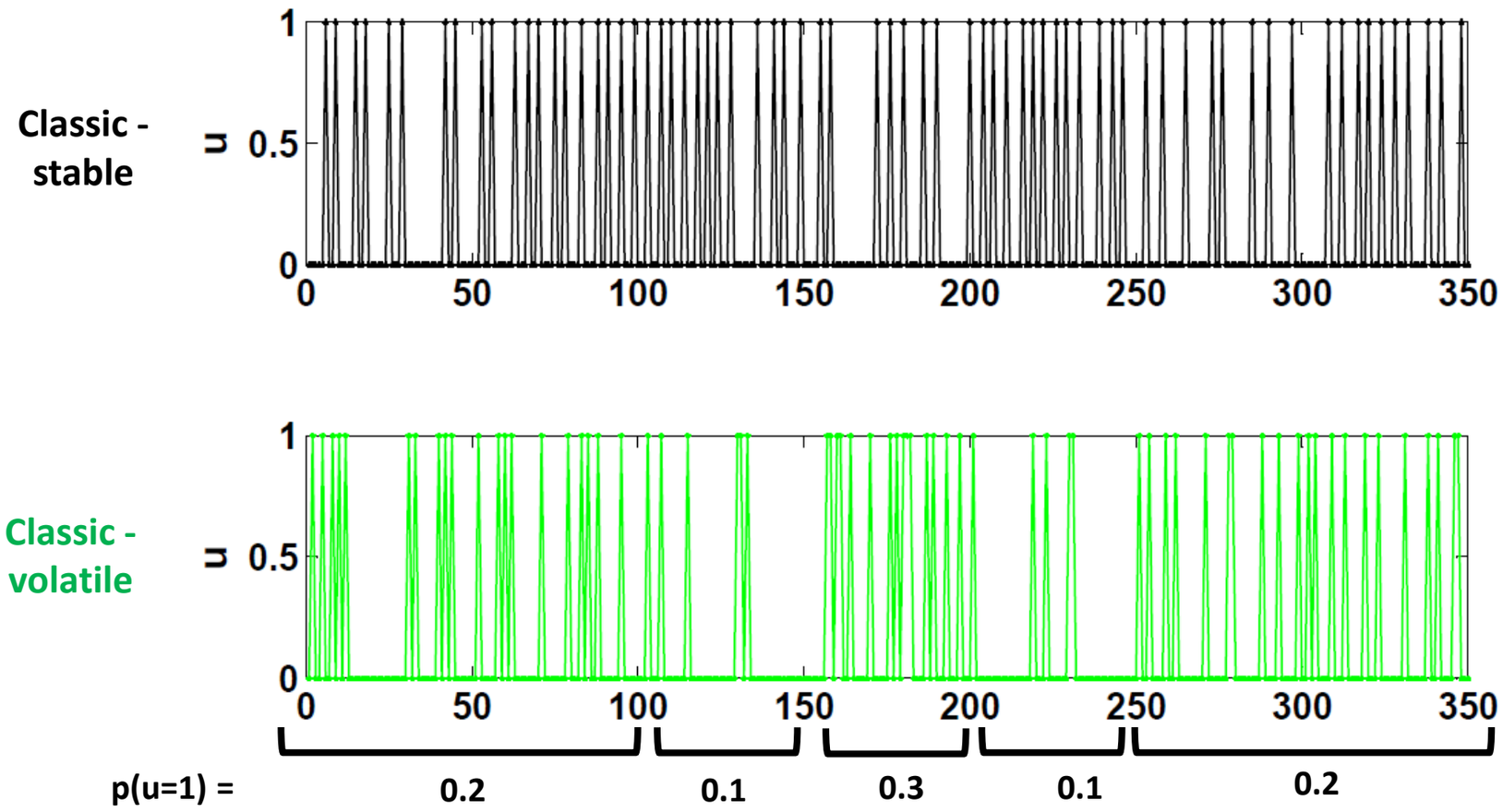


Sanchez et al. 2014

3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ **Static designs to be compared with ADO**

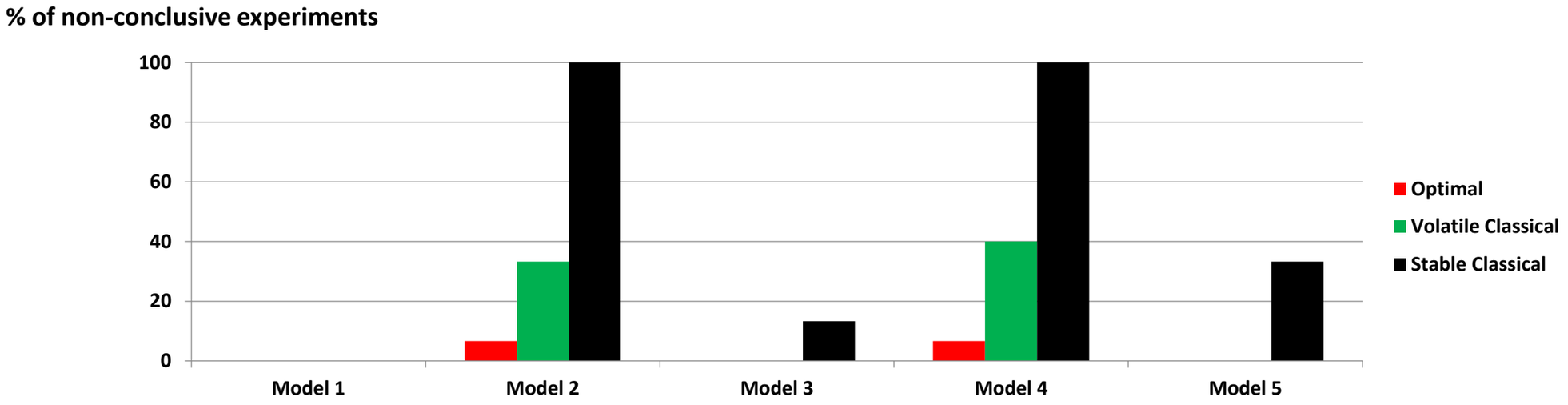


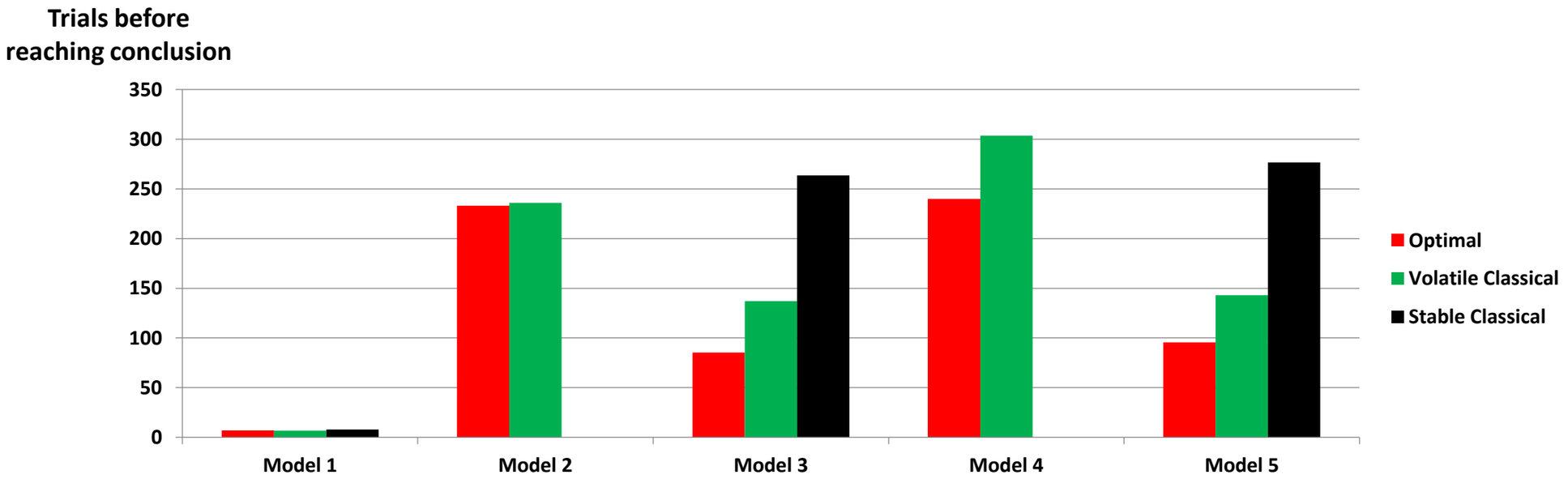
Sanchez et al. 2014

3 – Online Bayesian inference and design optimization

Disentangling between learning models

✓ **Results on Monte Carlo simulations**





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