Principles of Dynamic Causal Modelling for EEG/MEG

Ryszard Auksztulewicz

Wellcome Trust Centre for Neuroimaging

University College London

Overview

1 DCM: introduction

2 Differential equations

3 Neural states dynamics

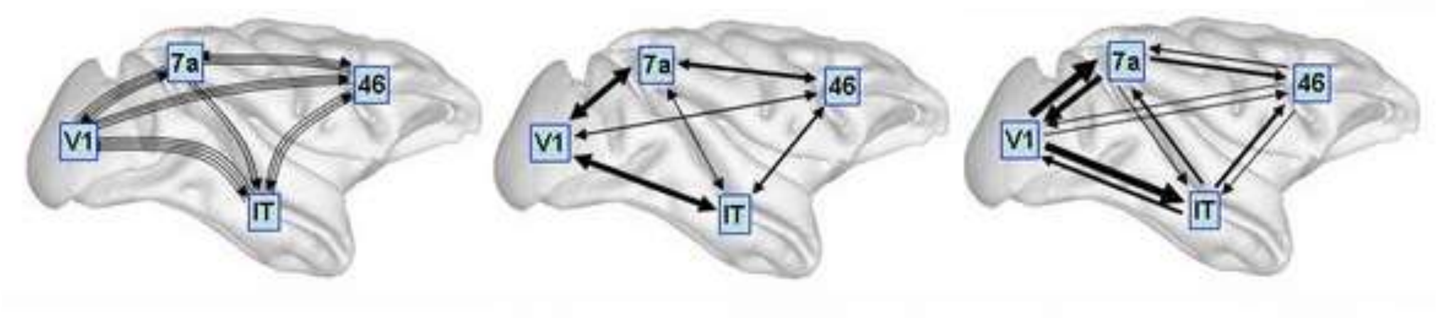
4 Bayesian inference

5 Conclusion

Introduction

structural, *functional* and *effective* connectivity

structural connectivity | functional connectivity | effective connectivity



O. Sporns 2007, *Scholarpedia*

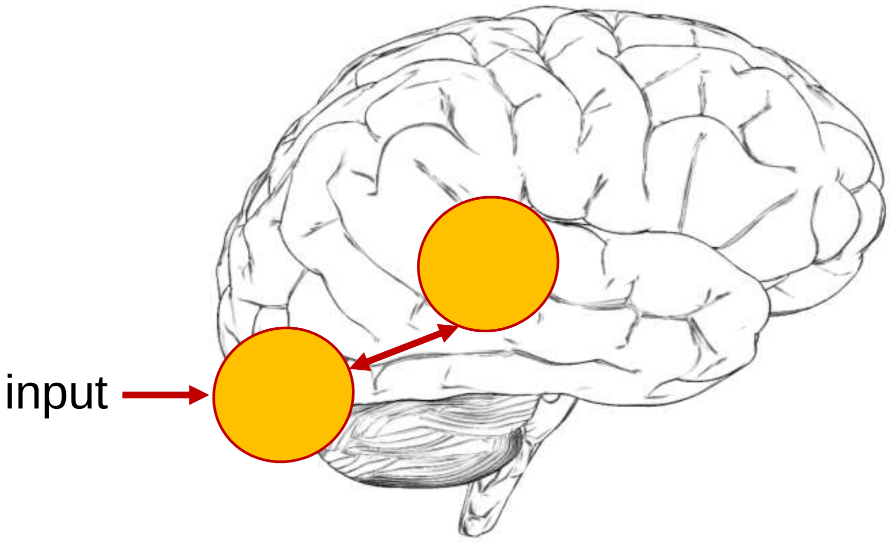
- ***structural* connectivity**
  = presence of axonal connections

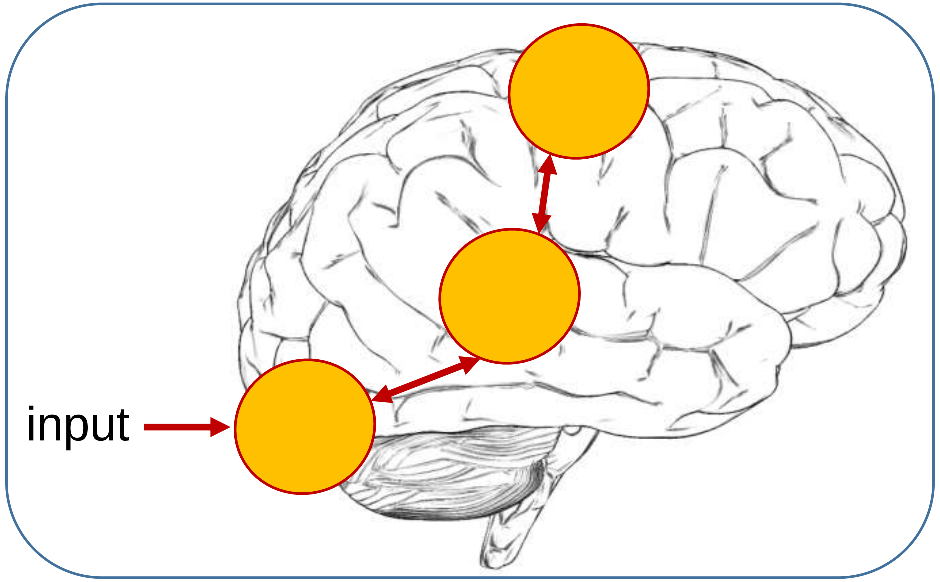
- ***functional* connectivity**
  = statistical dependencies between regional time series

- ***effective* connectivity**
  = causal (directed) influences between neuronal populations

! connections are recruited in a *context-dependent* fashion

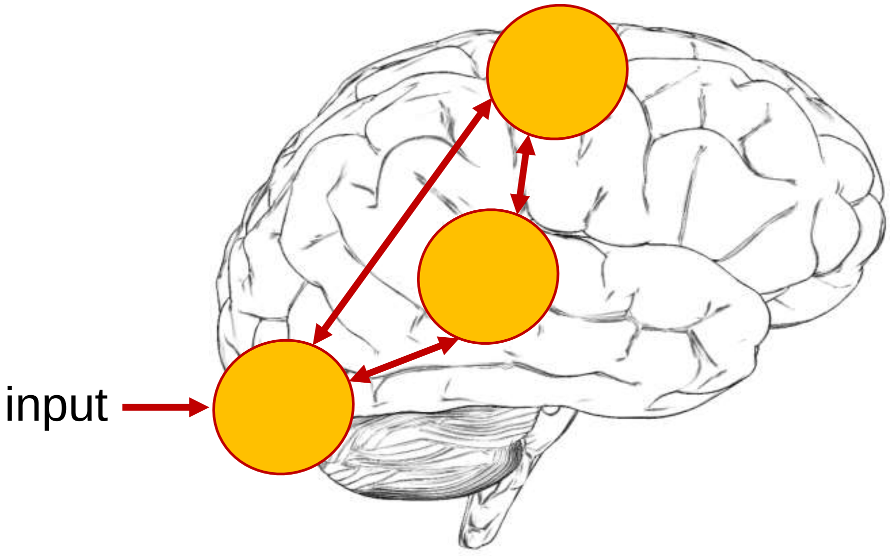
Does network XYZ explain my data better than network XY?

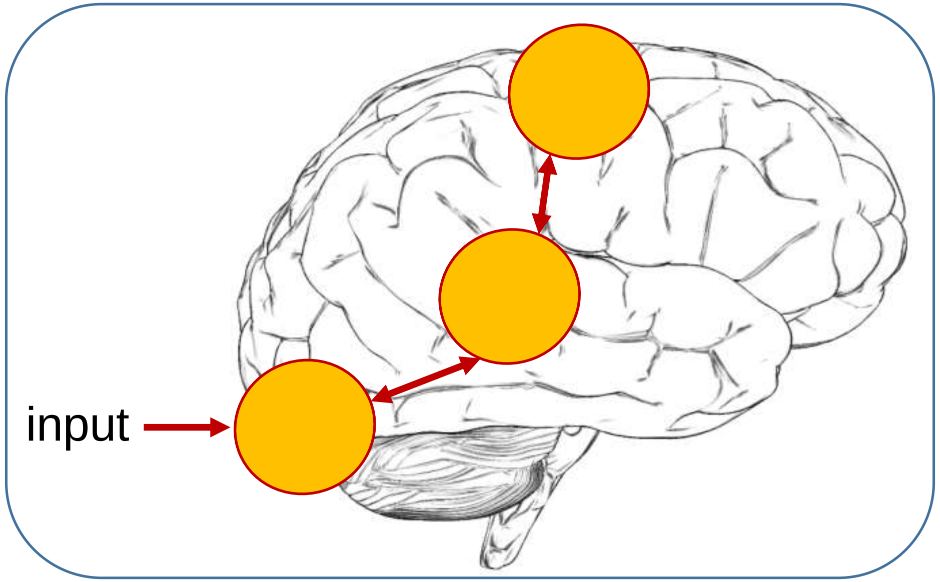




Does network XYZ explain my data better than network XY?

Which XYZ connectivity structure best explains my data?

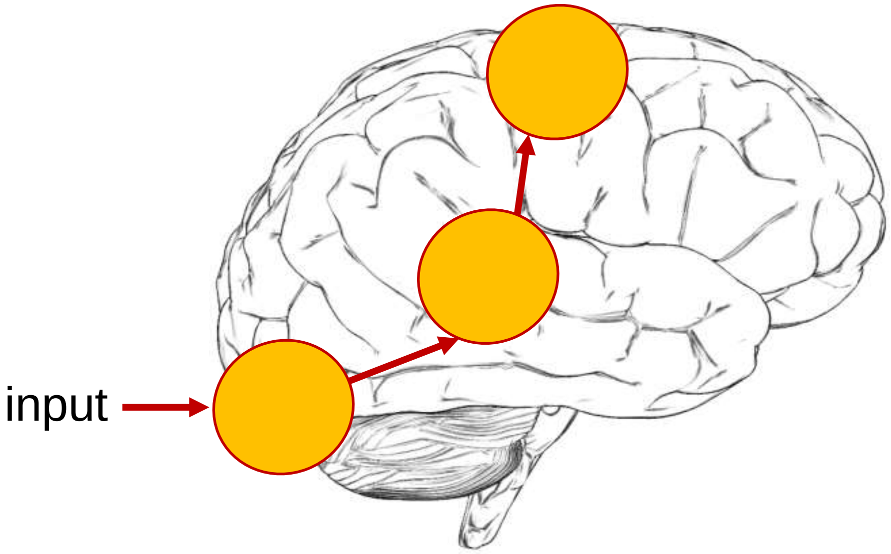


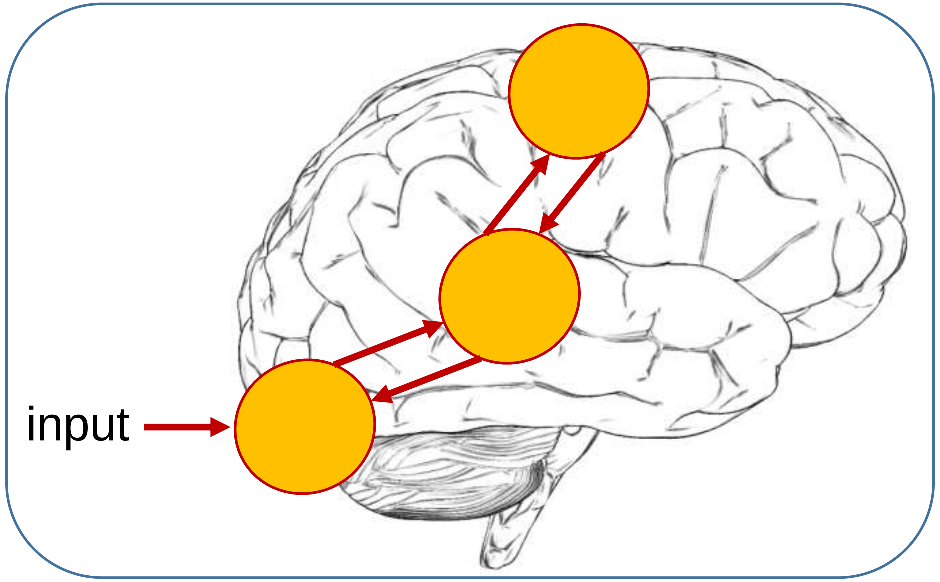


Does network XYZ explain my data better than network XY?

Which XYZ connectivity structure best explains my data?

Are X & Y linked in a bottom-up, top-down or recurrent fashion?



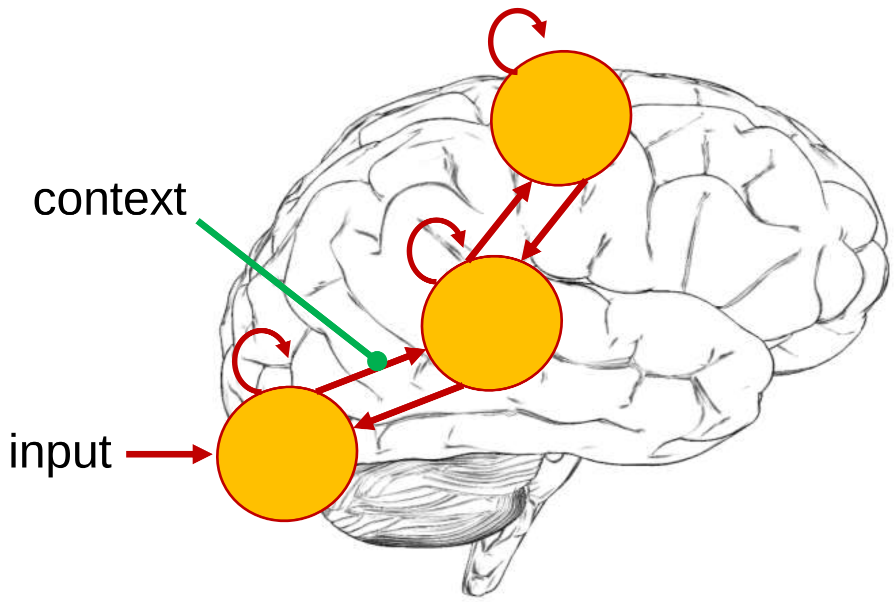


Does network XYZ explain my data better than network XY?

Which XYZ connectivity structure best explains my data?

Are X & Y linked in a bottom-up, top-down or recurrent fashion?

Is my effect driven by extrinsic or intrinsic connections?



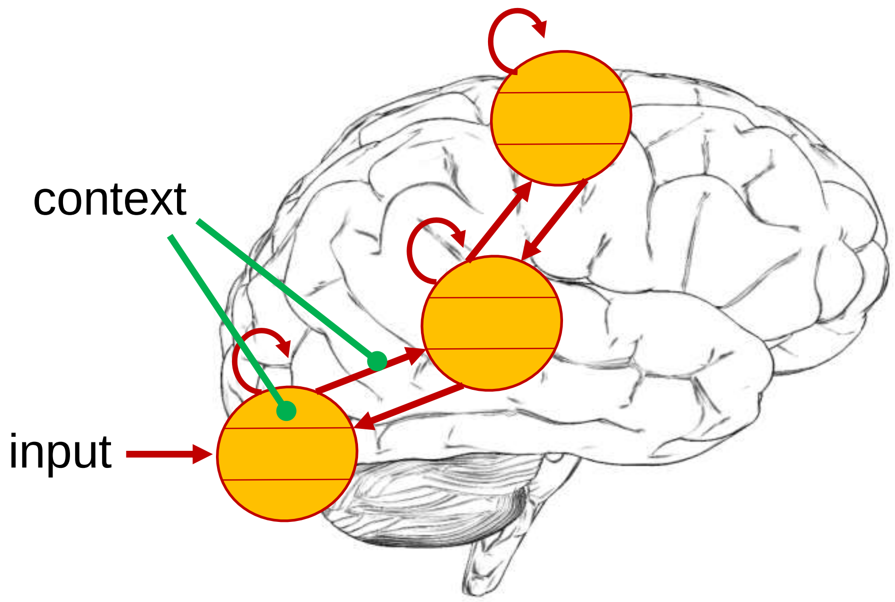
Does network XYZ explain my data better than network XY?

Which XYZ connectivity structure best explains my data?

Are X & Y linked in a bottom-up, top-down or recurrent fashion?

Is my effect driven by extrinsic or intrinsic connections?

Which neural populations are affected by contextual factors?



Does network XYZ explain my data better than network XY?

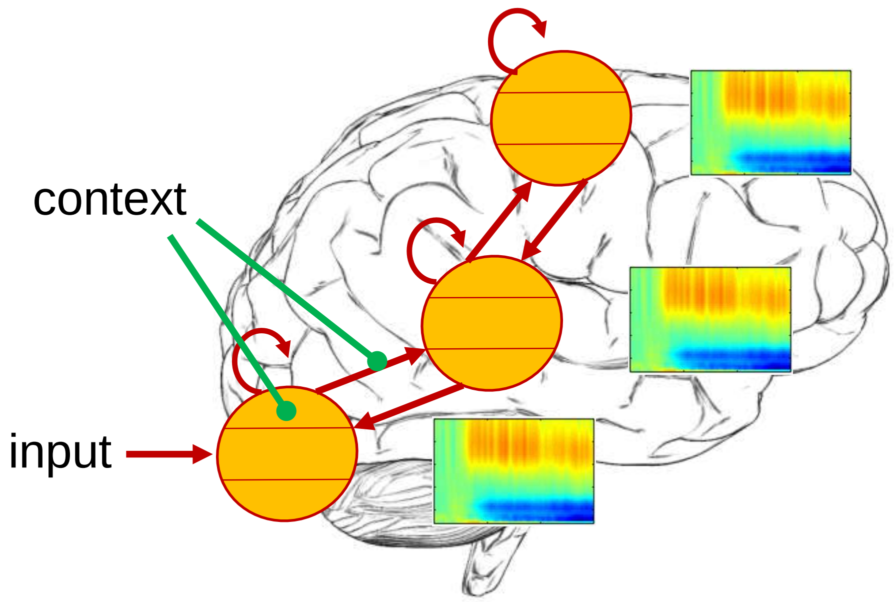
Which XYZ connectivity structure best explains my data?

Are X & Y linked in a bottom-up, top-down or recurrent fashion?

Is my effect driven by extrinsic or intrinsic connections?

Which neural populations are affected by contextual factors?

Which connections determine observed frequency coupling?



Does network XYZ explain my data better than network XY?

Which XYZ connectivity structure best explains my data?

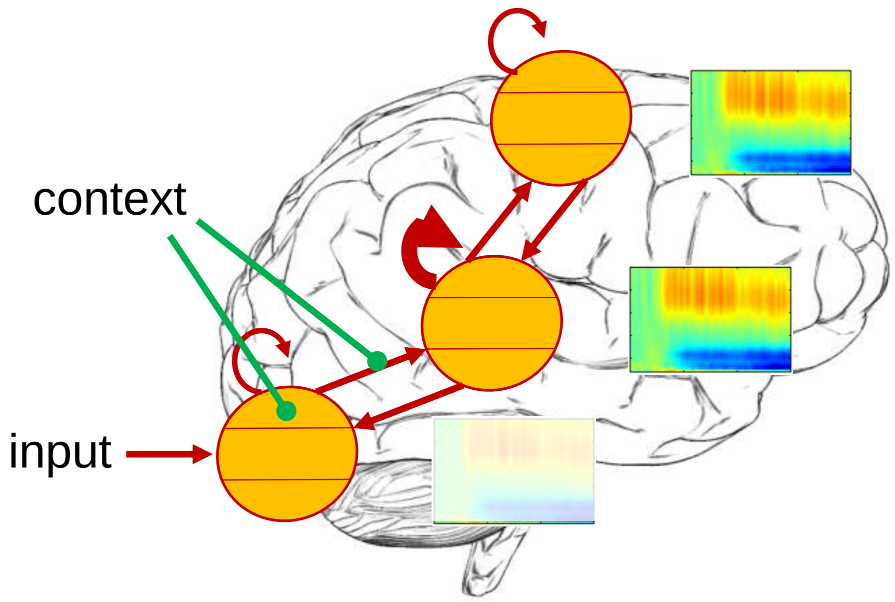
Are X & Y linked in a bottom-up, top-down or recurrent fashion?

Is my effect driven by extrinsic or intrinsic connections?

Which neural populations are affected by contextual factors?

Which connections determine observed frequency coupling?

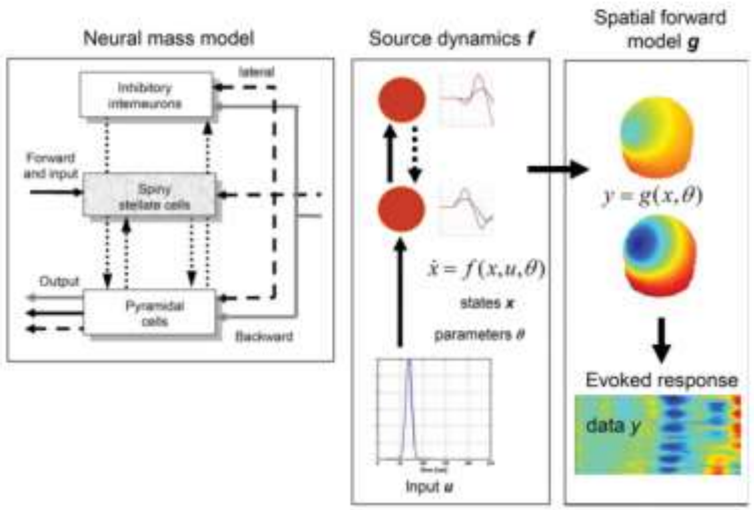
How changing a connection/parameter would influence data?



DCM for EEG/MEG

Physiological

Neurophysiological model



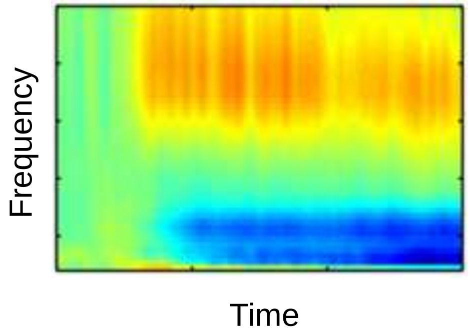
Electromagnetic forward model included

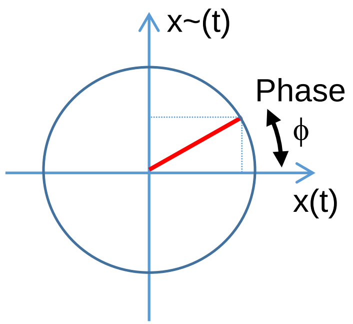
States *x* different from data *y*

- DCM for event-related potentials
- DCM for cross-spectral density

Phenomenological

Models a particular data feature



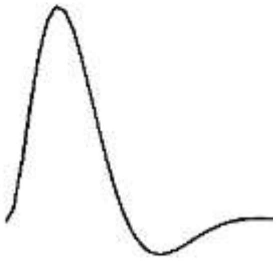


Source locations not optimized

States *x* and data *y* in the same “format”

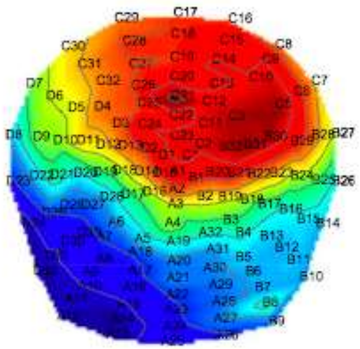
- DCM for Induced Responses
- DCM for Phase Coupling

Evolution and observation mappings



Hemodynamic observation model: temporal convolution

Electromagnetic observation model: spatial convolution

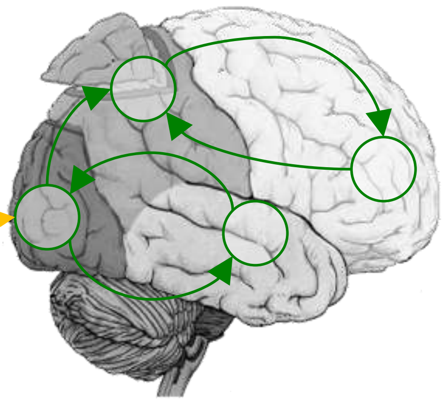


neural states dynamics

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

fMRI

EEG/MEG



inputs

- simple neuronal model
- realistic observation model

- realistic neuronal model
- simple observation model

Forward models and their inversion

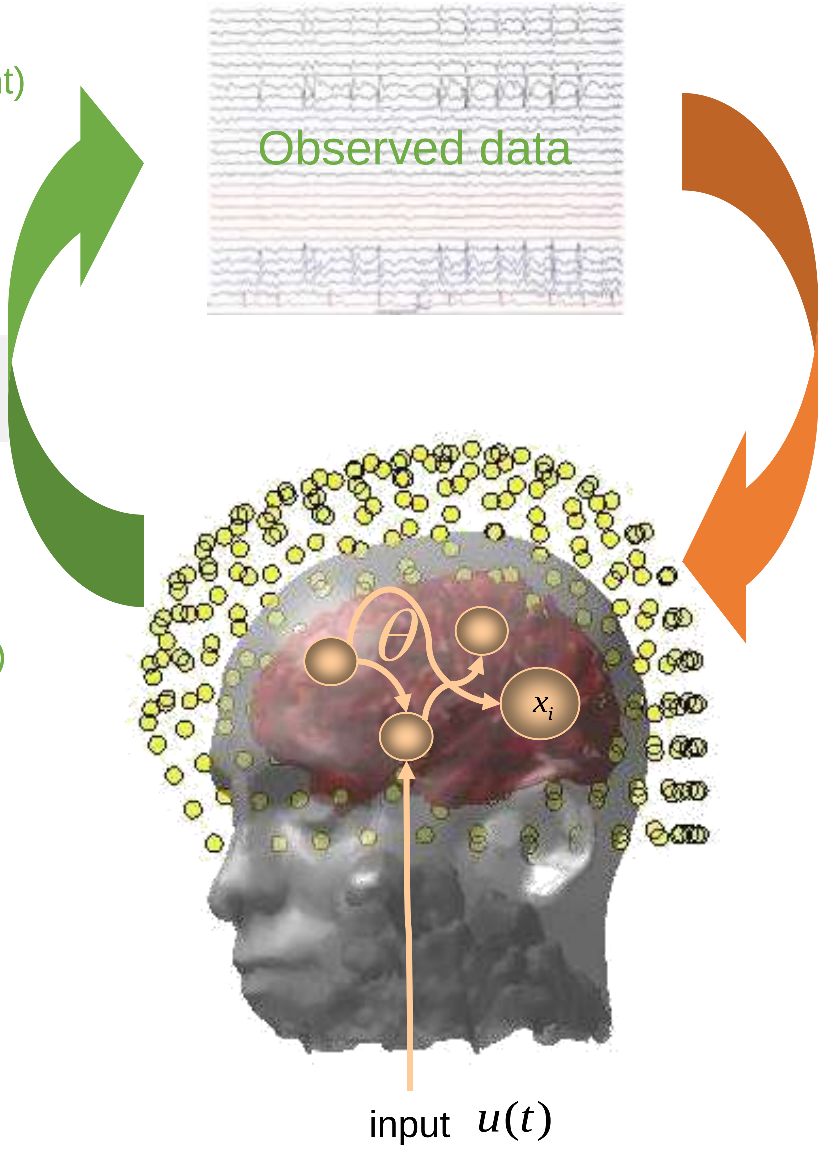
Forward model (measurement)

$$y = g(x,\theta) + \varepsilon$$

$$p(y \mid x,\theta,u,m)$$

Forward model (neuronal)

$$\dot{x} = f(x,u,\theta) + \omega$$



Model inversion

$$p(x,\theta \mid y,u,m)$$

Model specification and inversion

$$u(t)$$

$$\downarrow$$

Neural dynamics

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

Observer function

$$y = g(x, \theta) + \varepsilon$$

$$p(y \mid \theta, m) = \mathrm{N}\,(g(\theta), \Sigma(\theta))$$

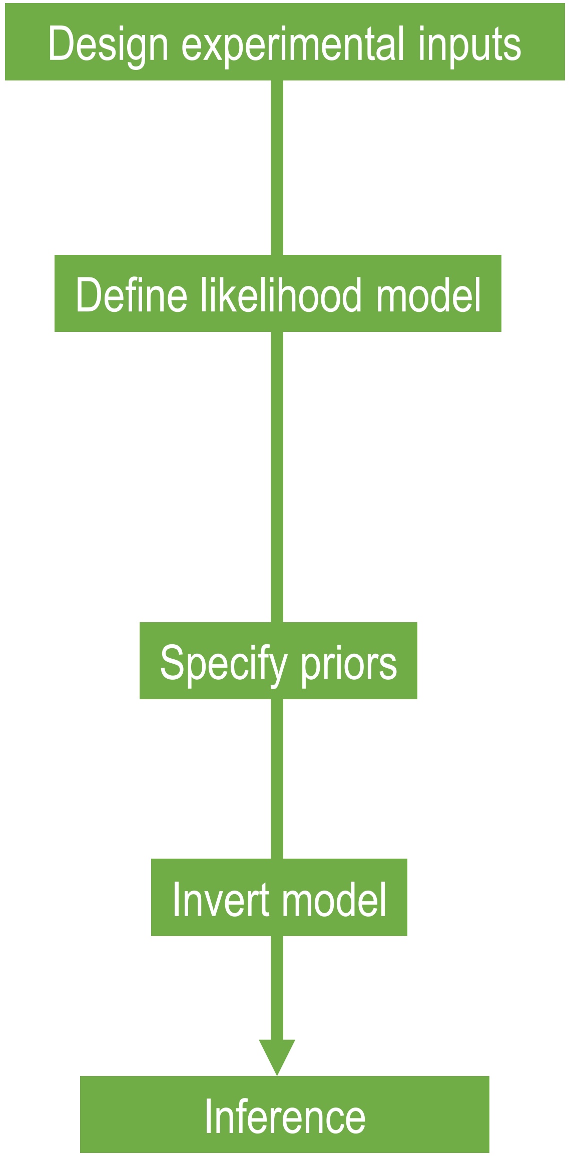
$$p(\theta, m) = \mathrm{N}\,(\mu_\theta, \Sigma_\theta)$$

Inference on models

$$p(y \mid m) = \int p(y \mid \theta, m)\, p(\theta)\, d\theta$$

Inference on parameters

$$p(\theta \mid y, m) = \frac{p(y \mid \theta, m)\, p(\theta, m)}{p(y \mid m)}$$



Overview

1 DCM: introduction

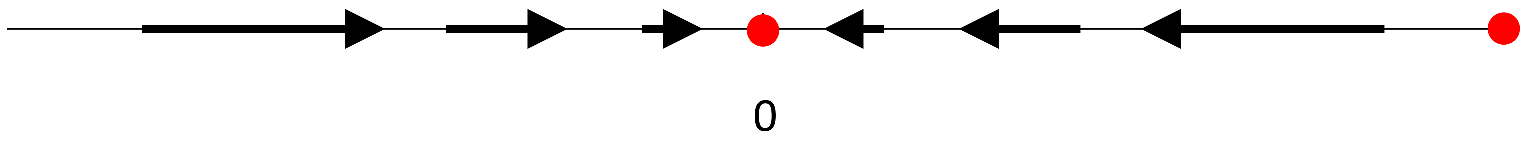
2 Differential equations

3 Neural states dynamics

4 Bayesian inference

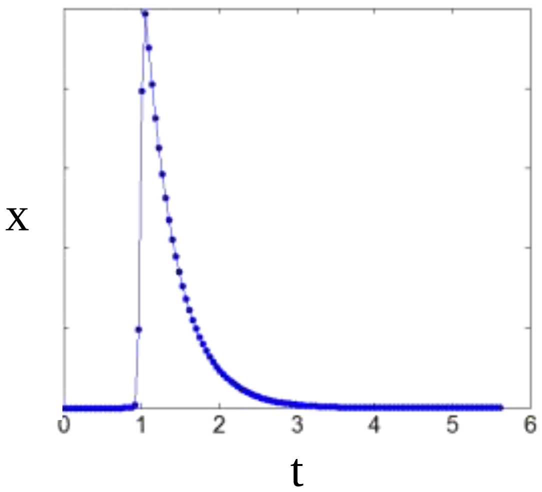
5 Conclusion

$$\dot{x} = -k \cdot x$$



Analytic solution

$$x(t) = x_0 e^{-kt}$$



Numerical solution

$$x(0) = x_0$$

$$x(0+\Delta t) = x(0) - k \cdot x(0) \cdot \Delta t$$

$$x(0+2\Delta t) = x(0+\Delta t) - k \cdot x(0+\Delta t) \cdot \Delta t$$

$$\ldots$$

‘Neural’ equation (exponential decay)

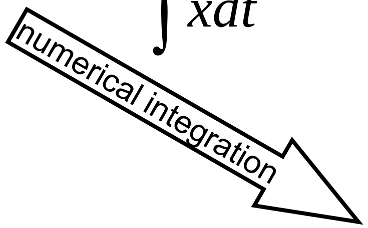
$$\dot{x} = -k \cdot x + U$$

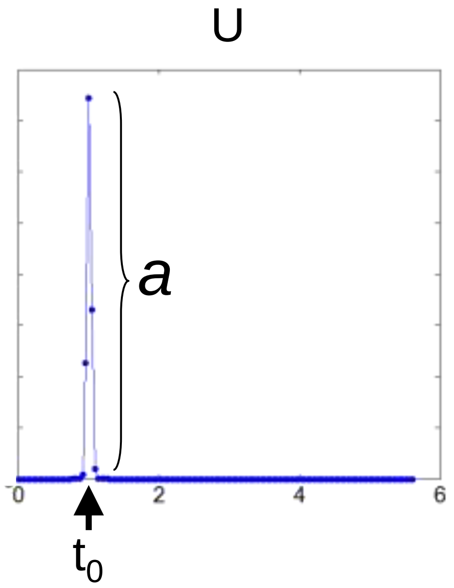
$$U = a \cdot e^{-\frac{(t-t_0)^2}{\Delta t^2}}$$

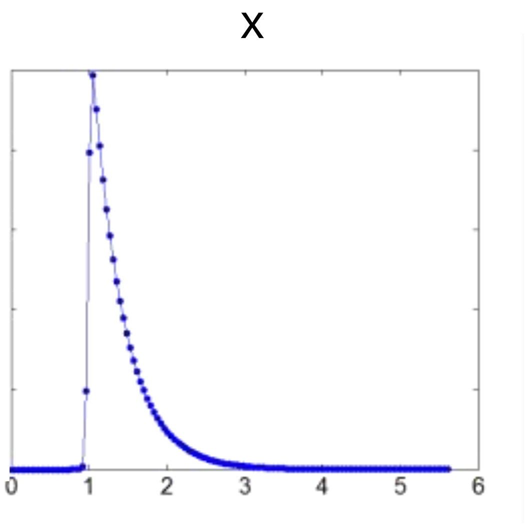
Observation equation

$$y = G \cdot x$$

$$\int \dot{x}\,dt$$







‘Neural’ equation (exponential decay)

$$\dot{x} = -2.47 \cdot x + U$$

$$U = 750 \cdot e^{-\frac{(t - t_0)^2}{\Delta t^2}}$$

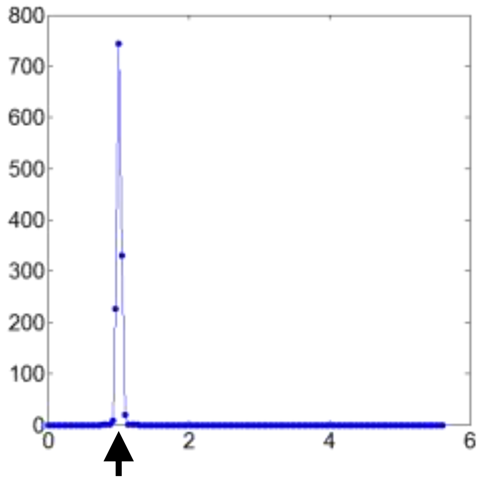
Observation equation

$$y = 1 \cdot x$$

$$\int \dot{x} dt$$

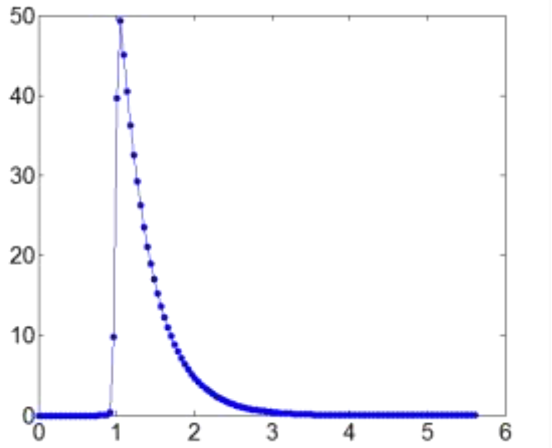
numerical integration

U



t_0 - the only free parameter

x



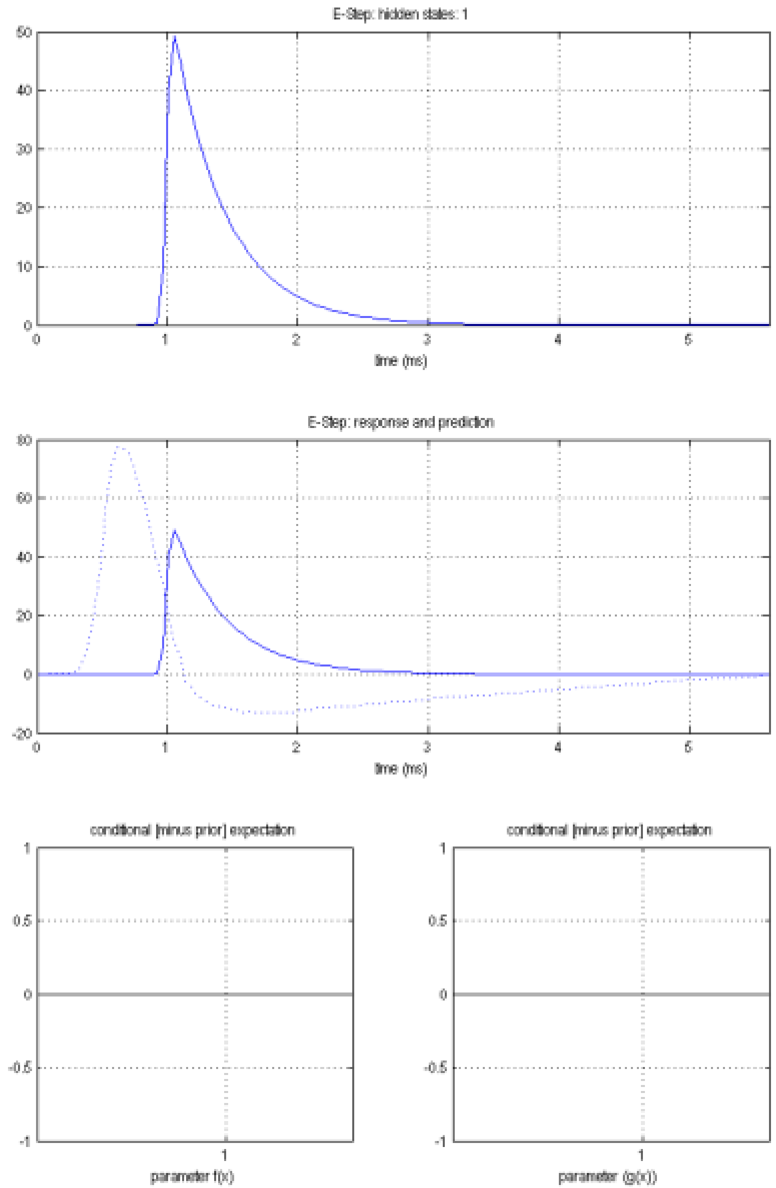
Optimization scheme for fitting the parameters to the data

- The objective function for optimization is the free energy which approximates the (log) model evidence:

$$p(y|m) = \int p(y|\vartheta, m)\, p(\vartheta|m)\, d\vartheta$$

- There are many possible schemes based on different assumptions. Present DCM implementations in SPM use variational Bayesian scheme.

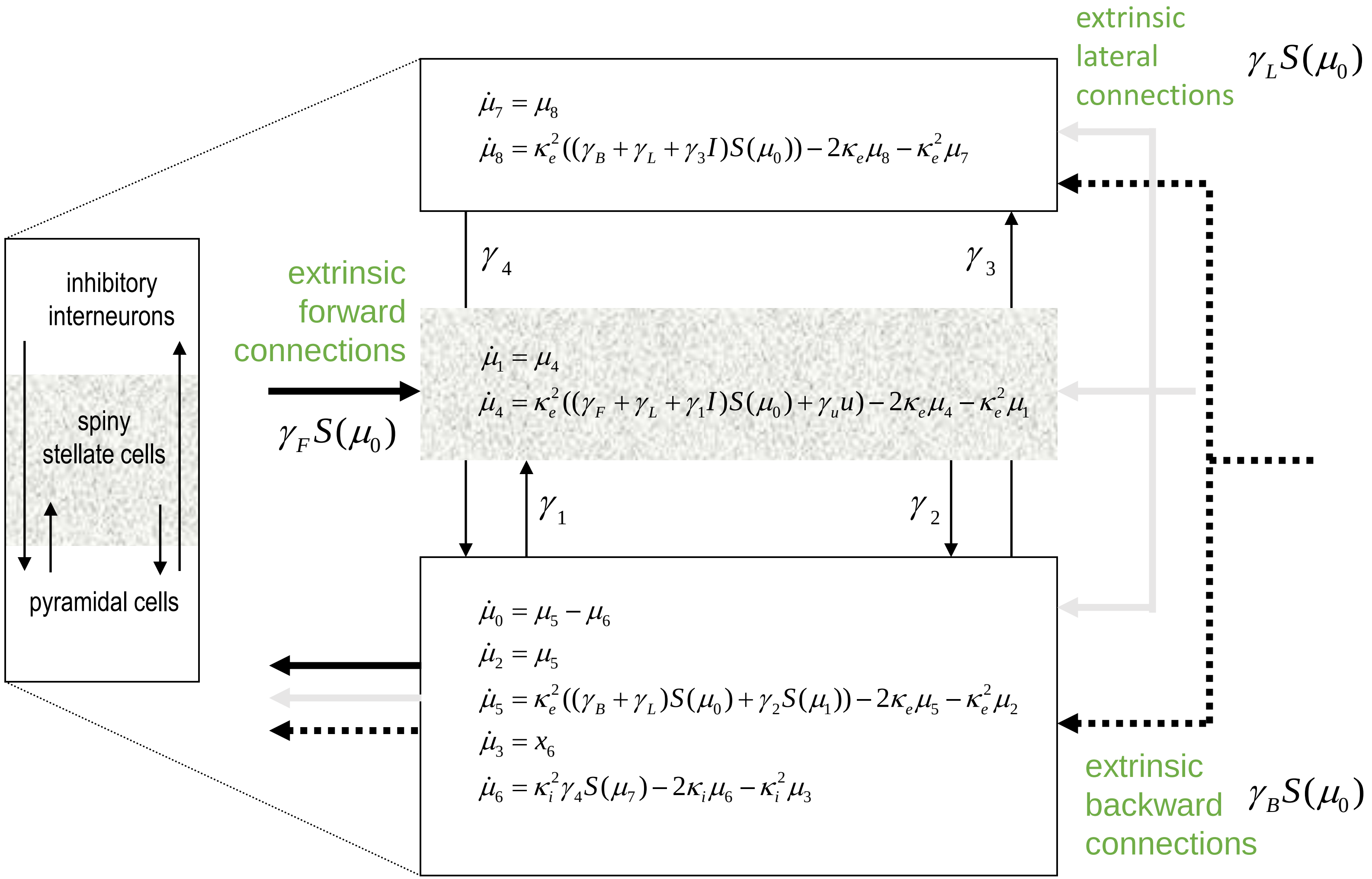
- Once the scheme converges it yields
  - The highest value of free energy the scheme could attain
  - Posterior distribution of the free parameters
  - Simulated data as similar to the original data as the model could generate

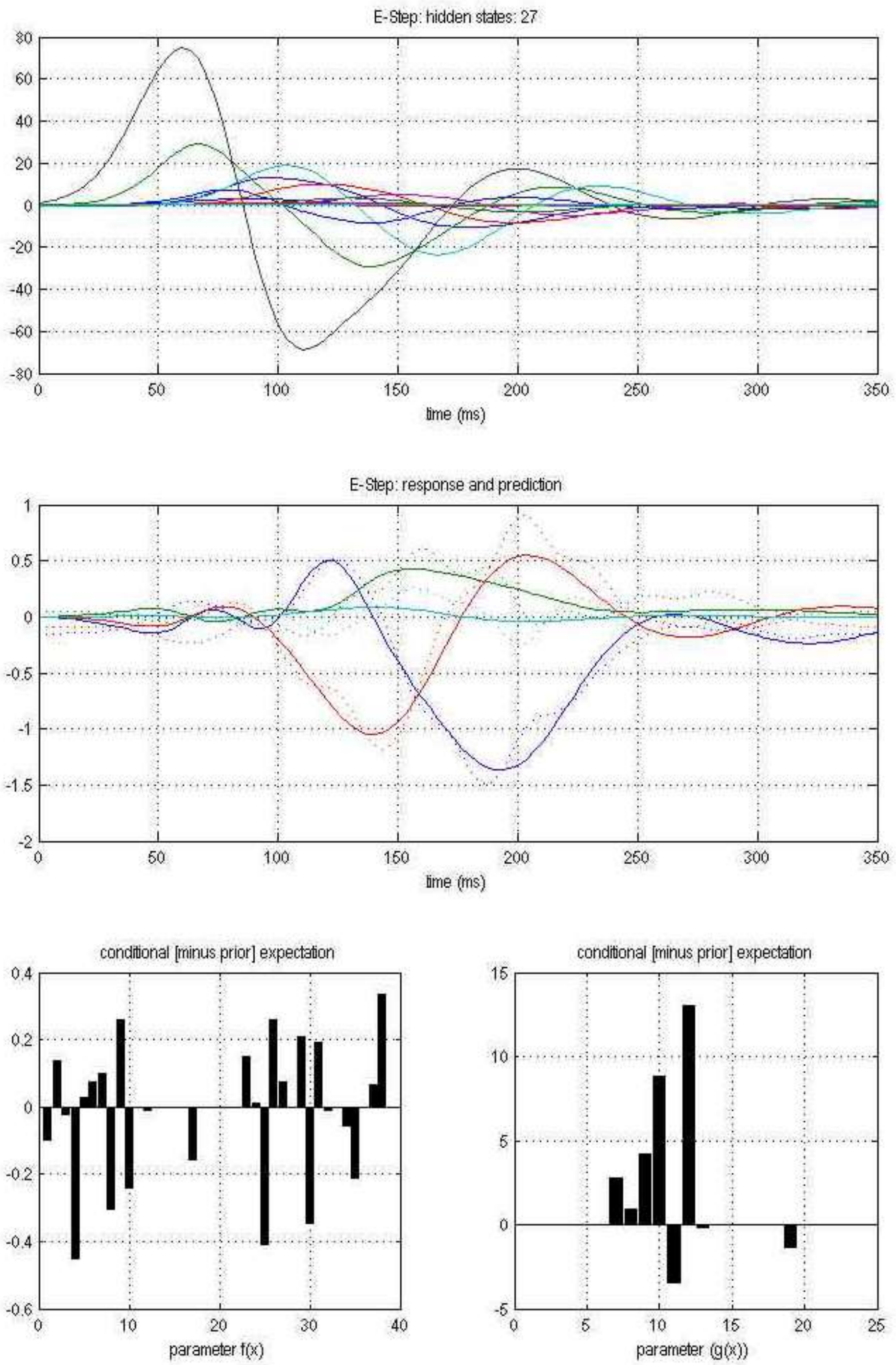

E-Step: hidden states: 1
time (ms)
E-Step: response and prediction
time (ms)
conditional [minus prior] expectation
parameter f(x)
conditional [minus prior] expectation
parameter (g(x))

Neural ensembles dynamics

DCM for M/EEG: *extrinsic connections between brain regions*

$$\dot{x} = -k \cdot x$$




E-Step: hidden states: 27
time (ms)
E-Step: response and prediction
time (ms)
conditional [minus prior] expectation
parameter f(x)
conditional [minus prior] expectation
parameter (g(x))

Overview

1 DCM: introduction

2 Differential equations

3 Neural states dynamics

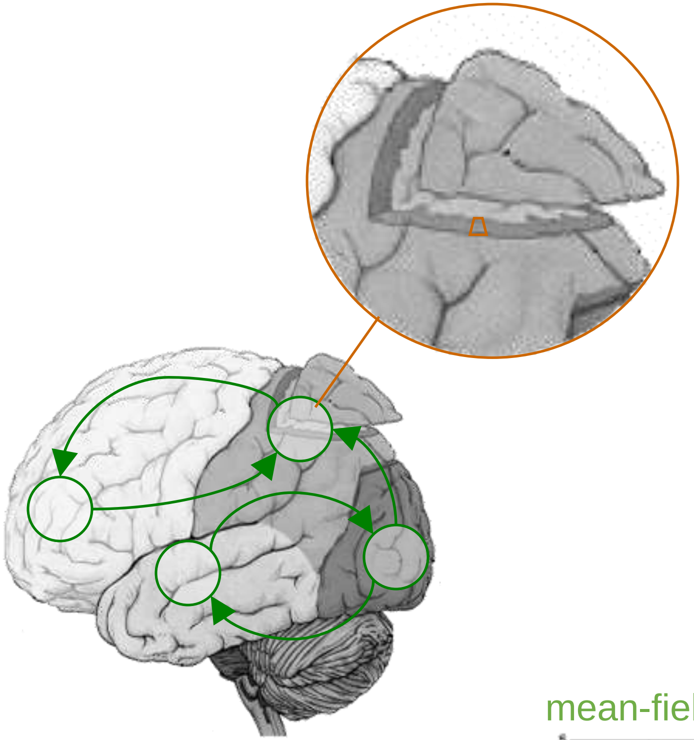
4 Bayesian inference

5 Conclusion

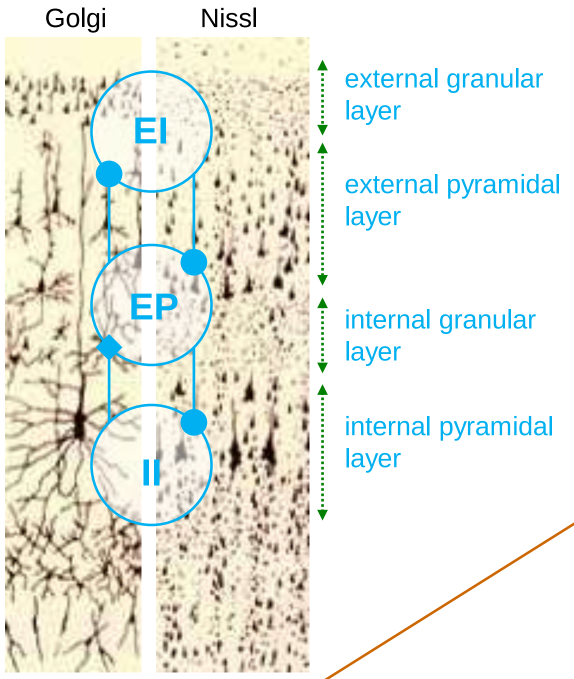
Neural ensembles dynamics

DCM for M/EEG: *systems of neural populations*

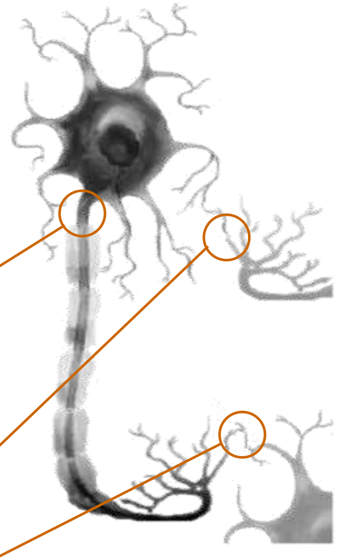
macro-scale



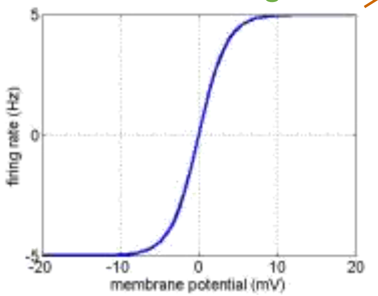
meso-scale



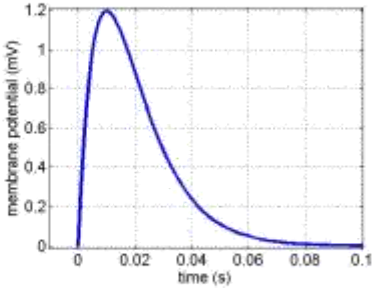
micro-scale



mean-field firing rate

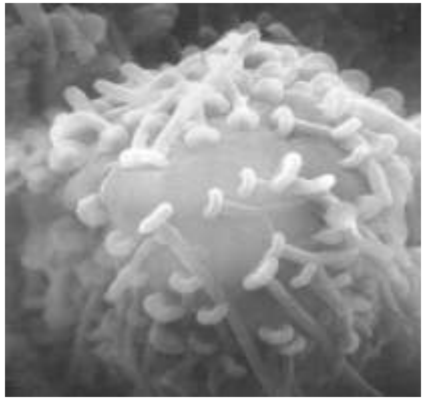


synaptic dynamics



Neural ensembles dynamics

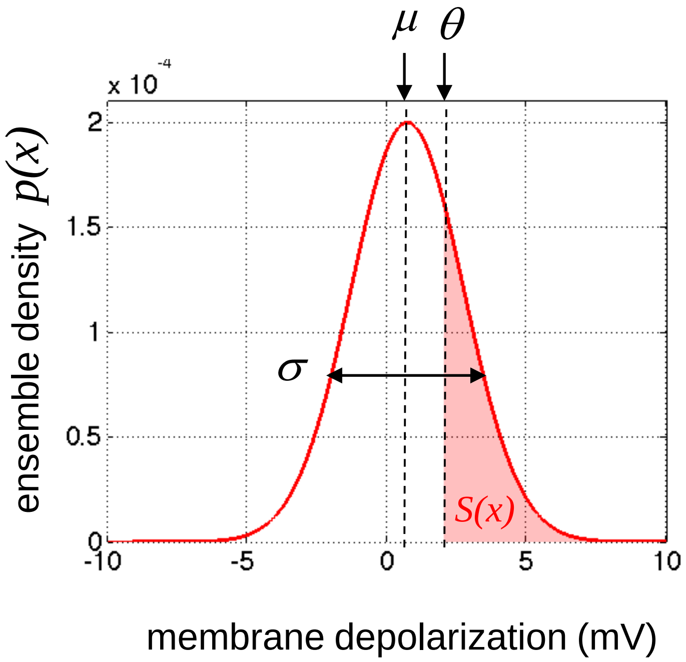
DCM for M/EEG: *from micro- to meso-scale*

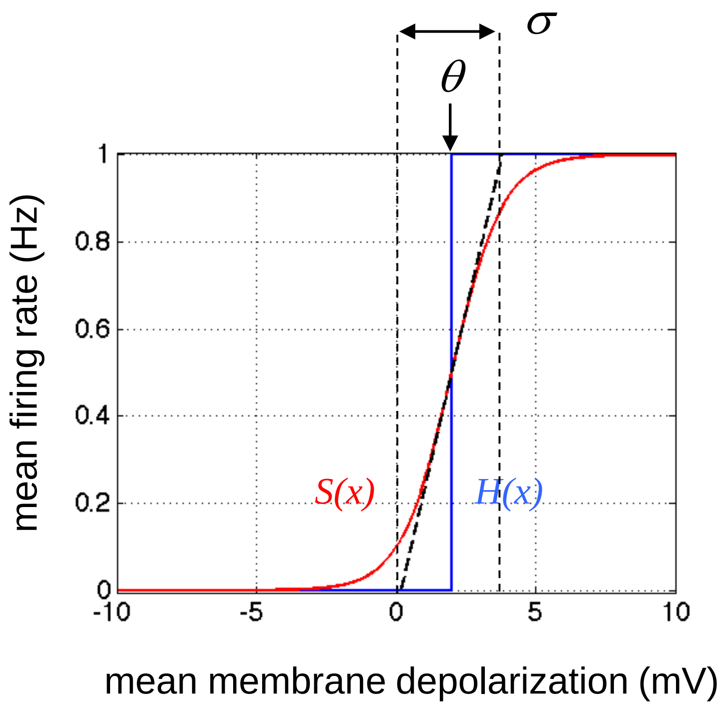


$x_j(t)$: post-synaptic potential of j^{th} neuron within its ensemble

$$\frac{1}{N-1}\sum_{j'\neq j} H\left(x_{j'}(t)-\theta\right) \xrightarrow{N\to\infty} \int H\left(x(t)-\theta\right) p\left(x(t)\right) dx$$

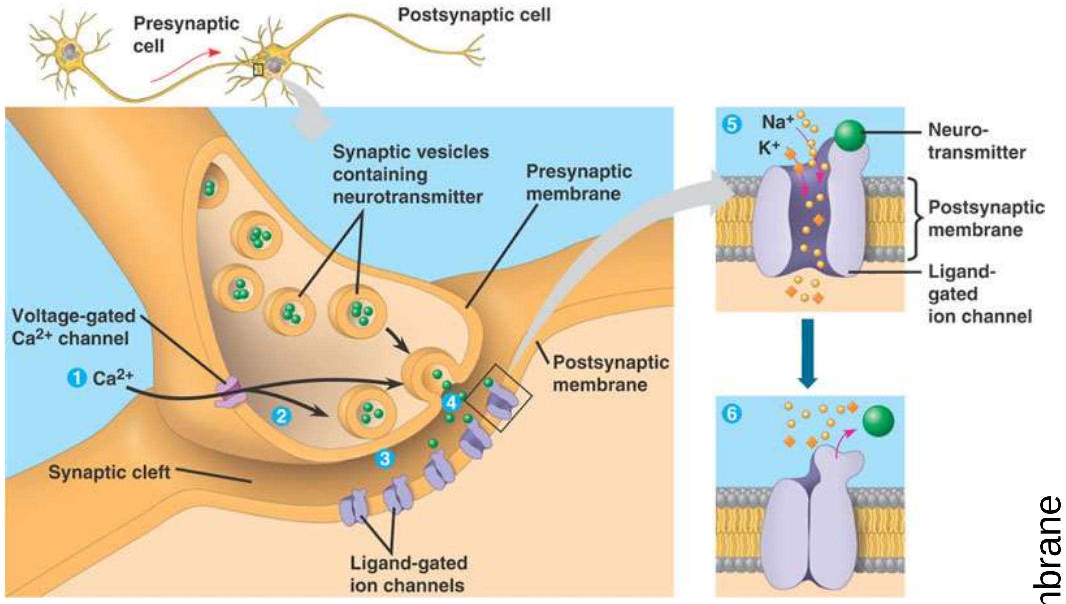
$\approx S(\mu)$ mean-field firing rate



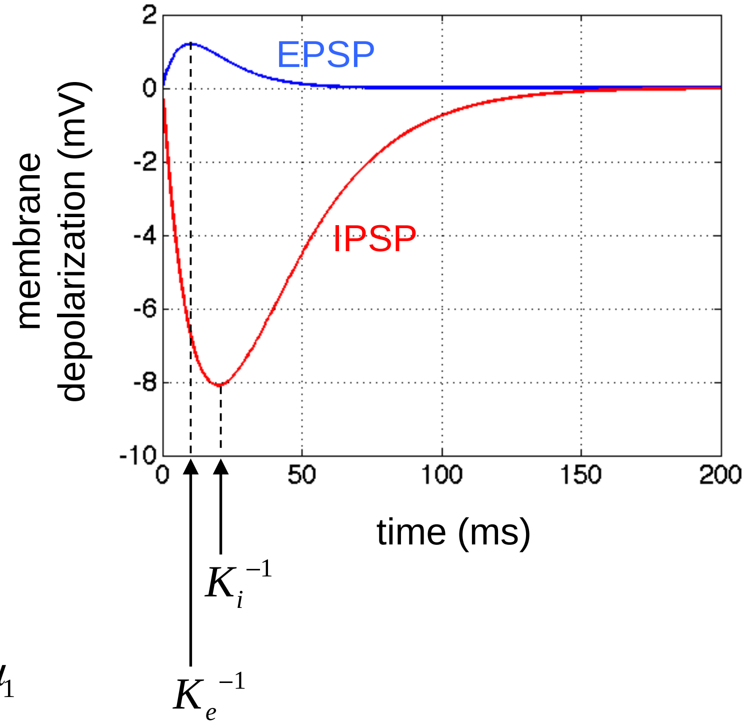


Neural ensembles dynamics

DCM for M/EEG: *synaptic dynamics*



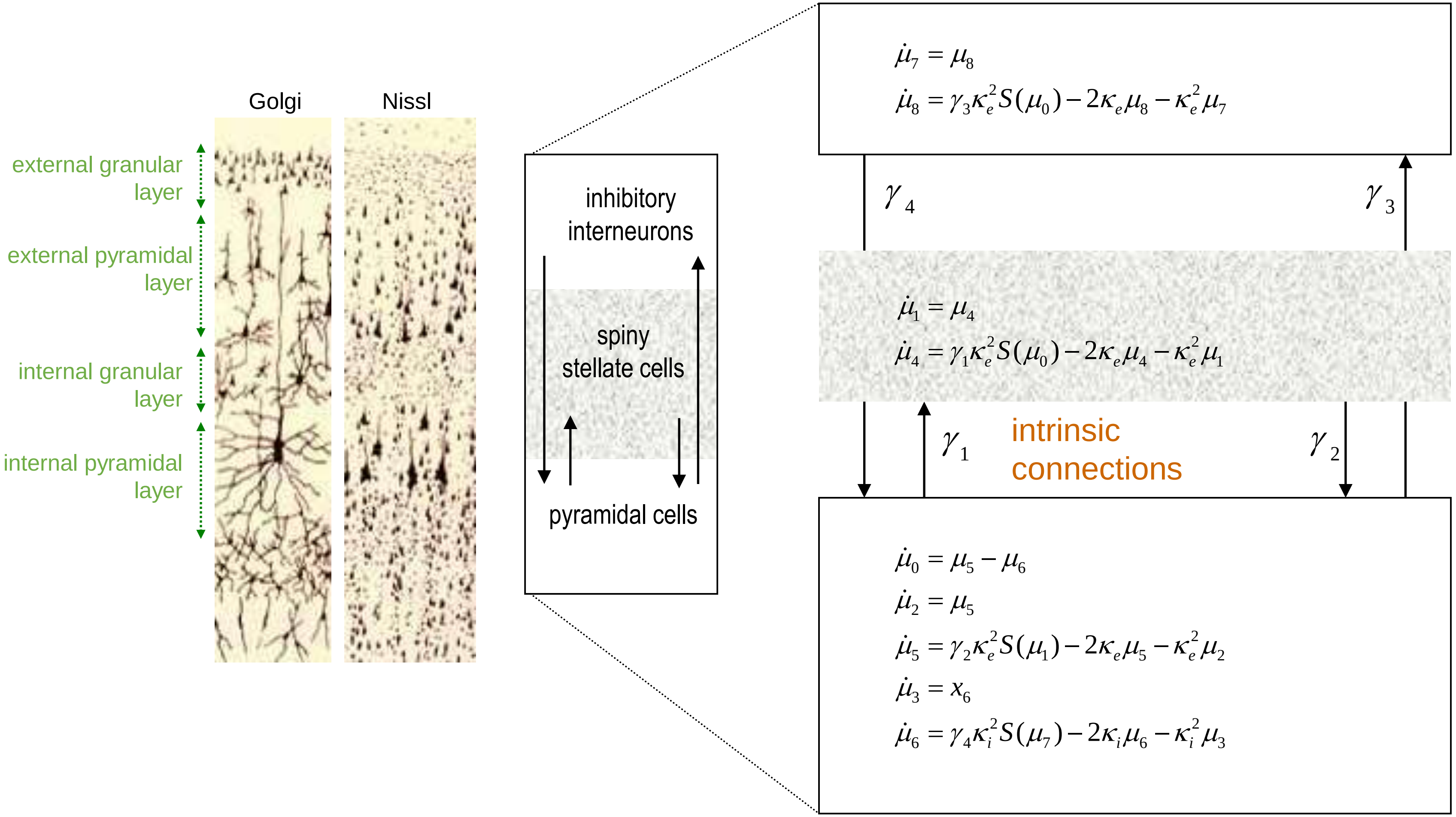
post-synaptic potential



$$\begin{cases} \dot{\mu}_1 = \mu_2 \\ \dot{\mu}_2 = \kappa_{i/e}^2 S(\mu_0) - 2\kappa_{i/e}\,\mu_2 - \kappa_{i/e}^2\,\mu_1 \end{cases}$$

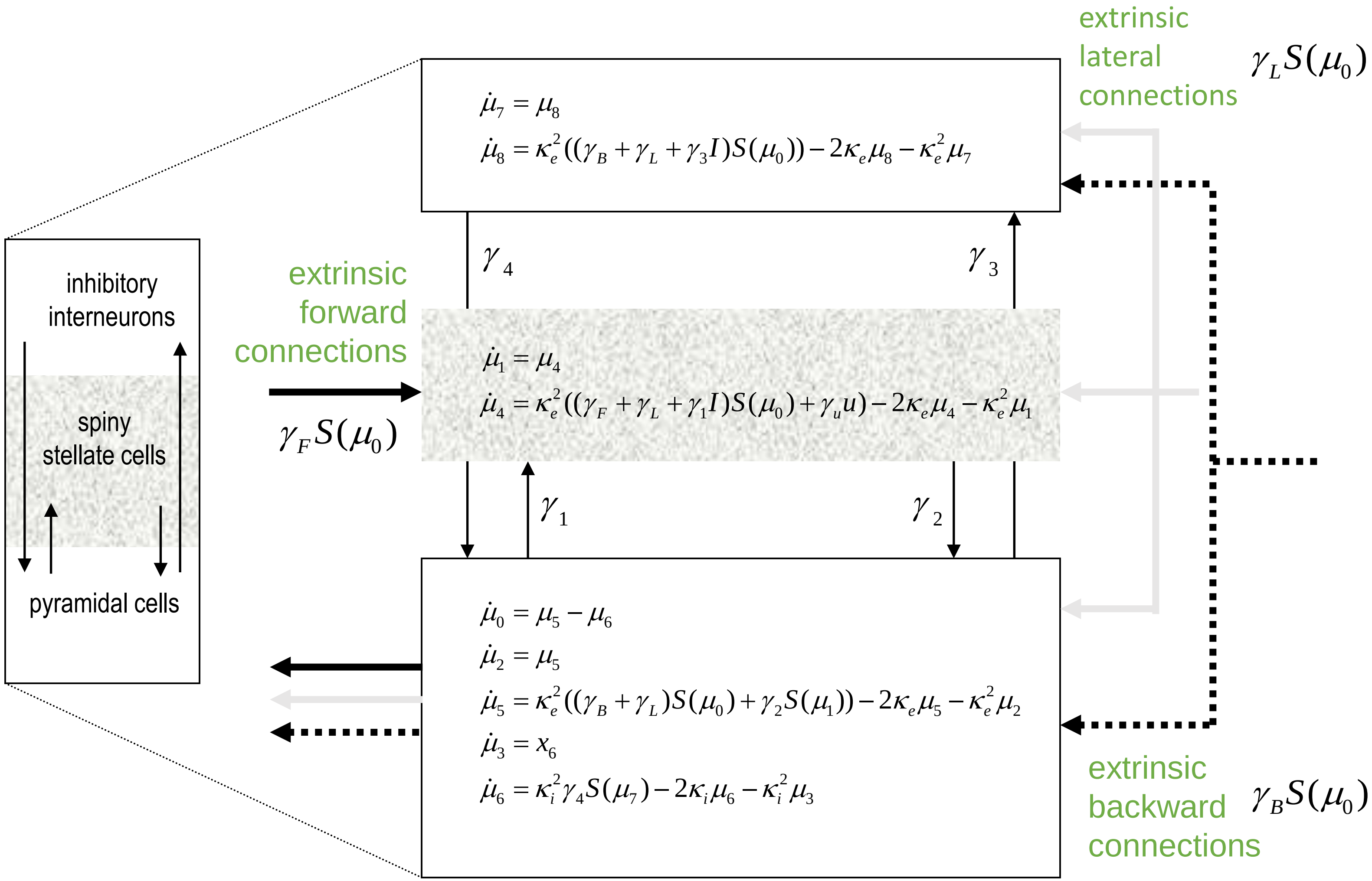
Neural ensembles dynamics

DCM for M/EEG: *intrinsic connections within the cortical column*



Neural ensembles dynamics

DCM for M/EEG: *extrinsic connections between brain regions*



Overview

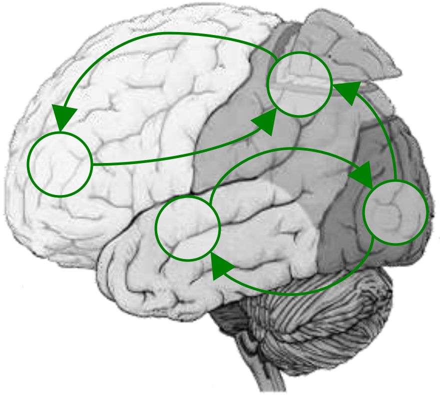
Bayesian inference

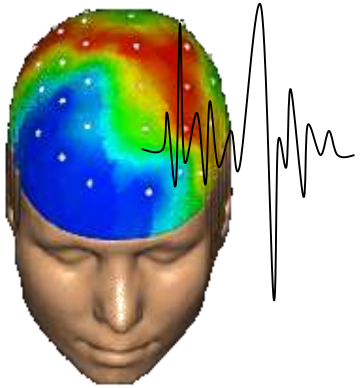
forward and inverse problems

forward problem

$$p\left(y \middle| \vartheta, m\right)$$

likelihood





posterior distribution

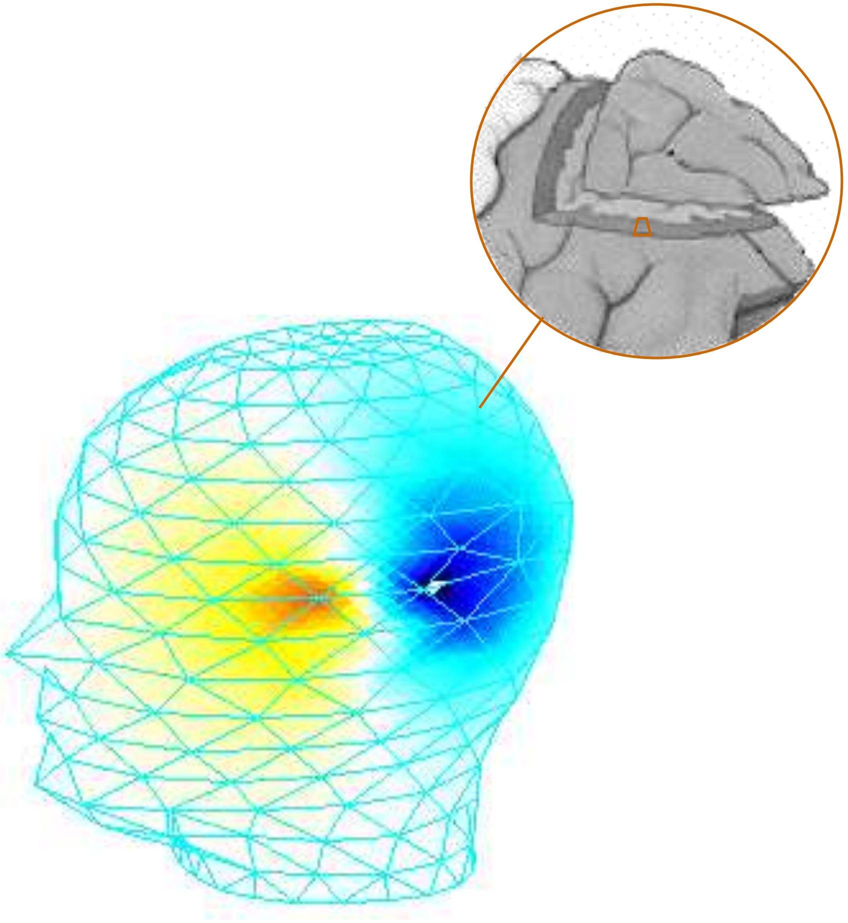
$$p\left(\vartheta \middle| y, m\right)$$

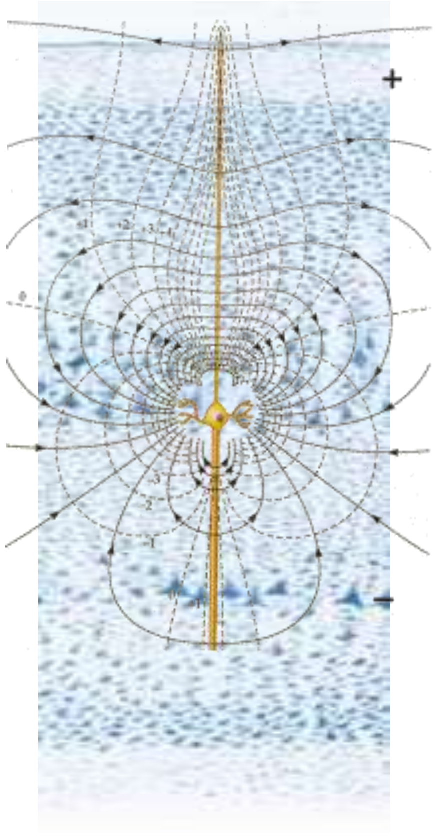
inverse problem

Bayesian inference

the electromagnetic forward problem

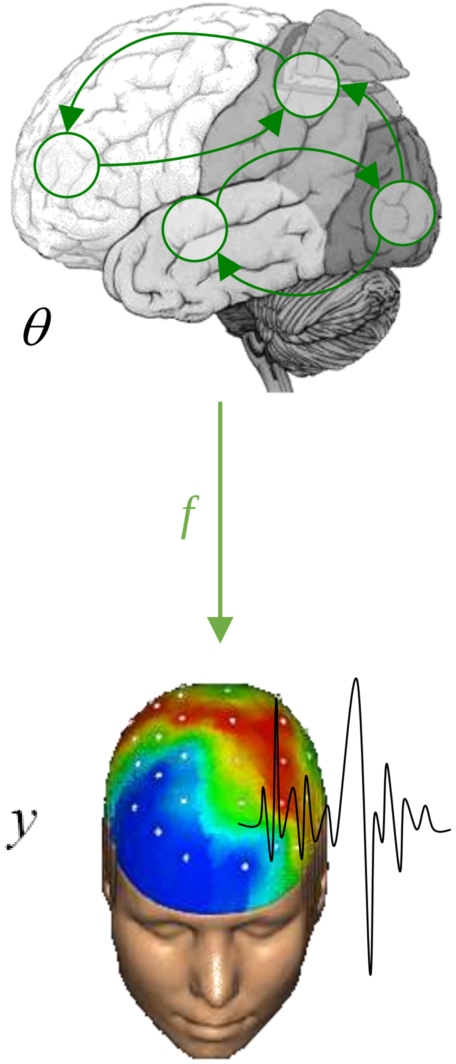
$$\mathbf{y}(t) = \sum_i \mathbf{L}^{(i)} \mathbf{w}_0^{(i)} \sum_j \beta_j \mu^{(ij)}(t) + \varepsilon(t)$$





Bayesian paradigm

deriving the likelihood function



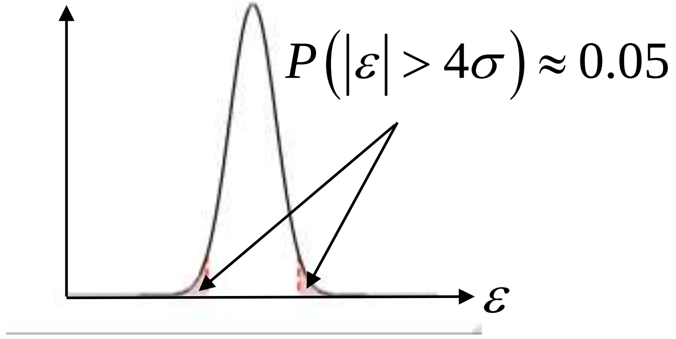
- Model of data with unknown parameters:

$$y = f(\theta) \qquad \text{e.g., GLM:} \quad f(\theta) = X\theta$$

- But data is noisy: $y = f(\theta) + \varepsilon$

- Assume noise/residuals is 'small':

$$p(\varepsilon) \propto \exp\left(-\frac{1}{2\sigma^2}\varepsilon^2\right)$$

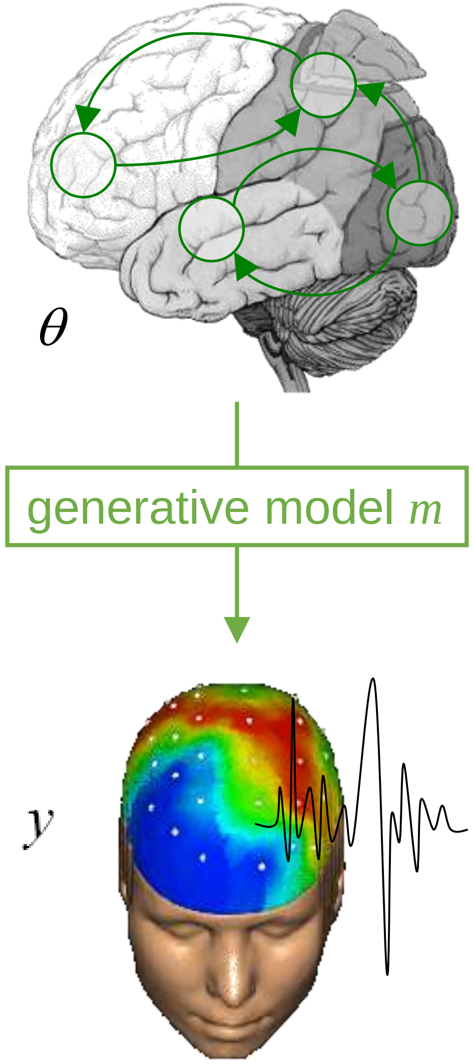


→ Distribution of data, *given fixed parameters*:

$$p(y|\theta) \propto \exp\left(-\frac{1}{2\sigma^2}(y - f(\theta))^2\right)$$

Bayesian paradigm

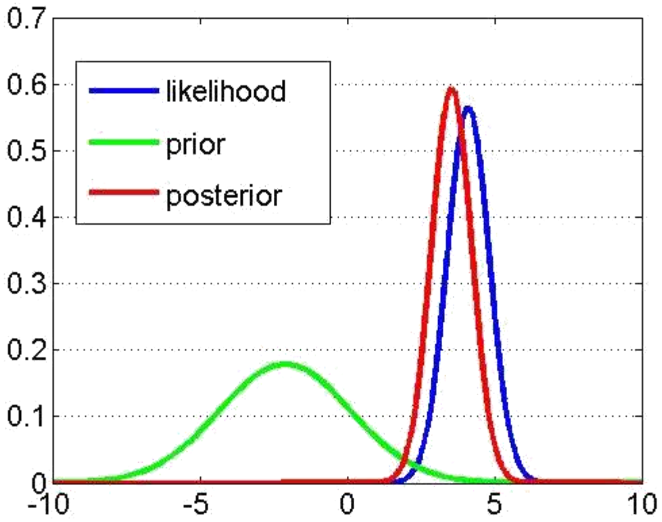
likelihood, priors and the model evidence



Likelihood: $p(y|\theta, m)$

Prior: $p(\theta|m)$

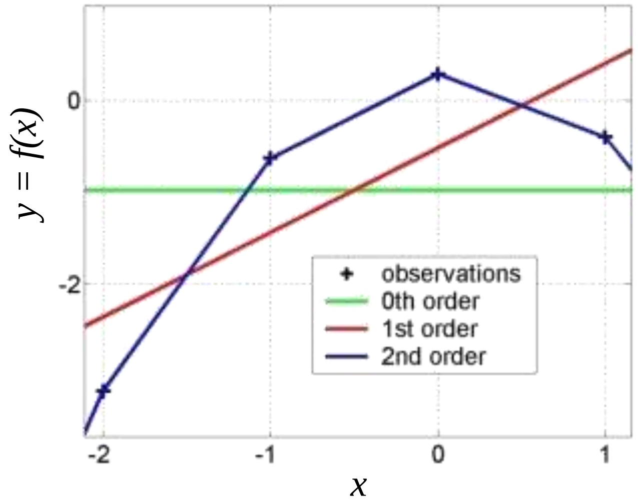
Bayes rule: $p(\theta|y,m) = \dfrac{p(y|\theta,m)\,p(\theta|m)}{p(y|m)}$

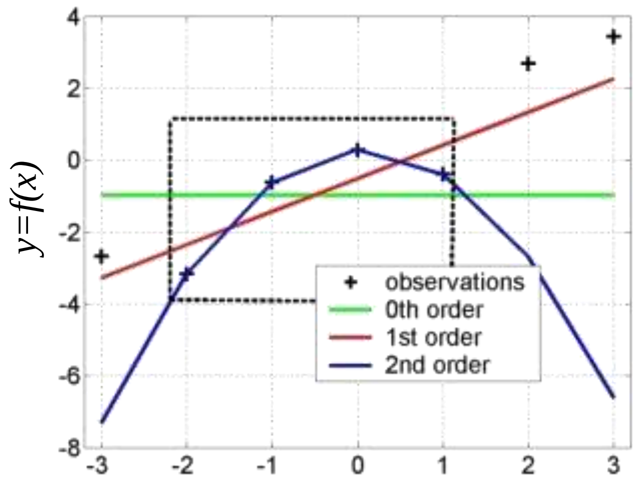


Bayesian inference

model comparison

Principle of parsimony :
« plurality should not be assumed without necessity »

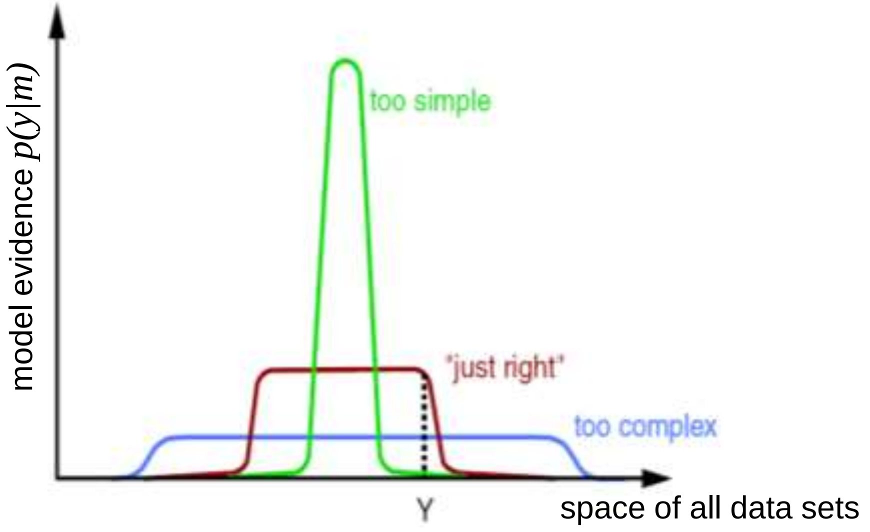




Model evidence:

$$p(y|m) = \int p(y|\vartheta, m)\, p(\vartheta|m)\, d\vartheta$$

"Occam's razor" :



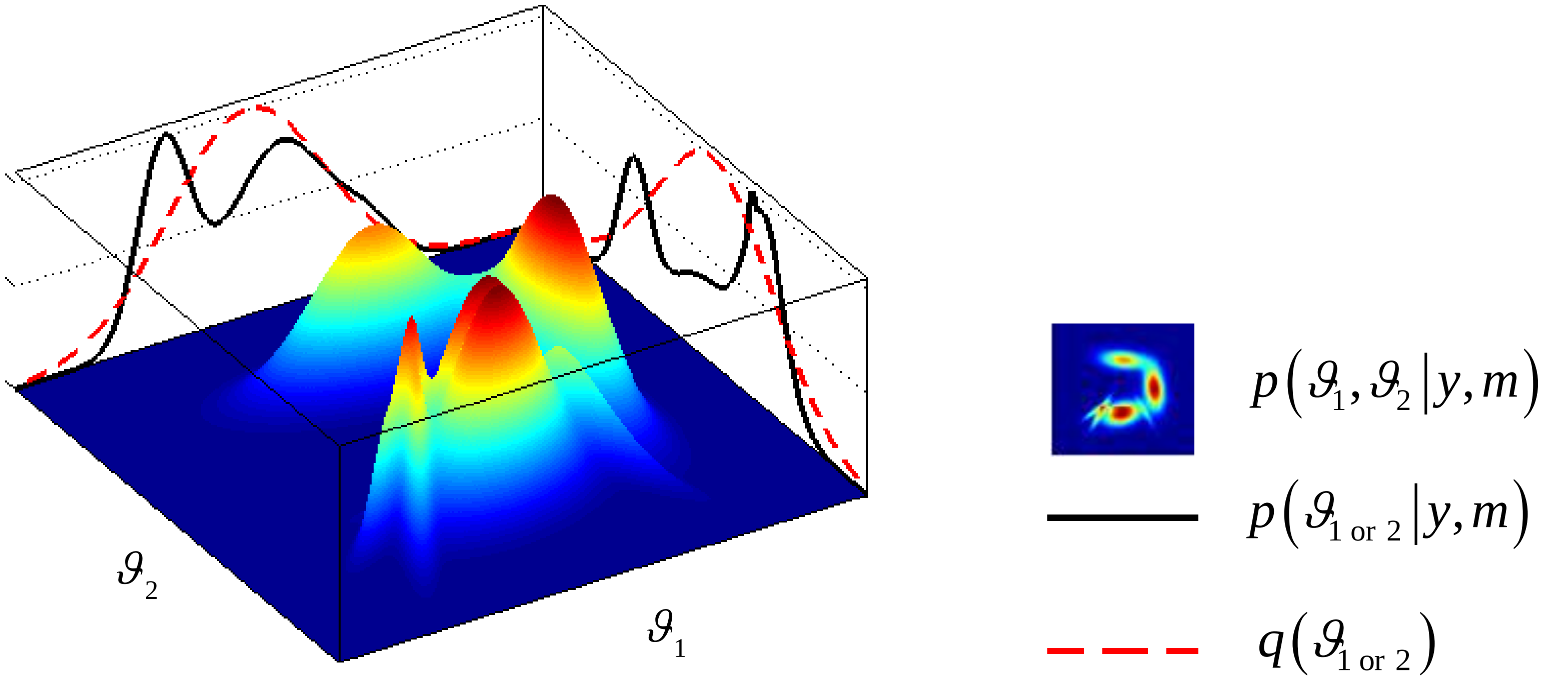
Bayesian inference

the variational Bayesian approach

$$\ln p(y|m) = \underbrace{\left\langle \ln p(\vartheta, y|m) \right\rangle_q + S(q)}_{\text{free energy : functional of } q} + D_{KL}\left(q(\vartheta); p(\vartheta|y,m)\right)$$

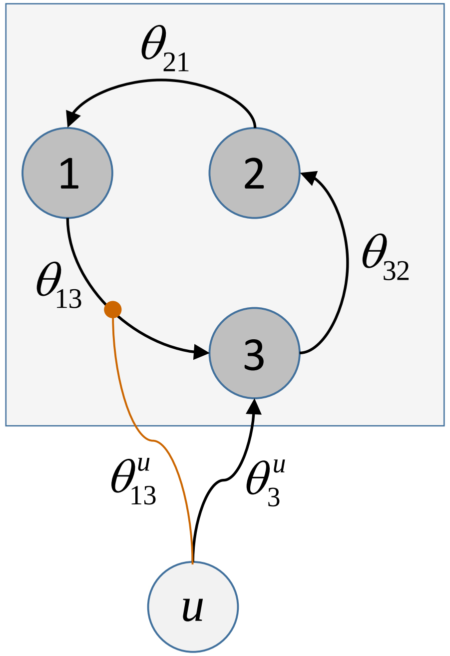
free energy : functional of q

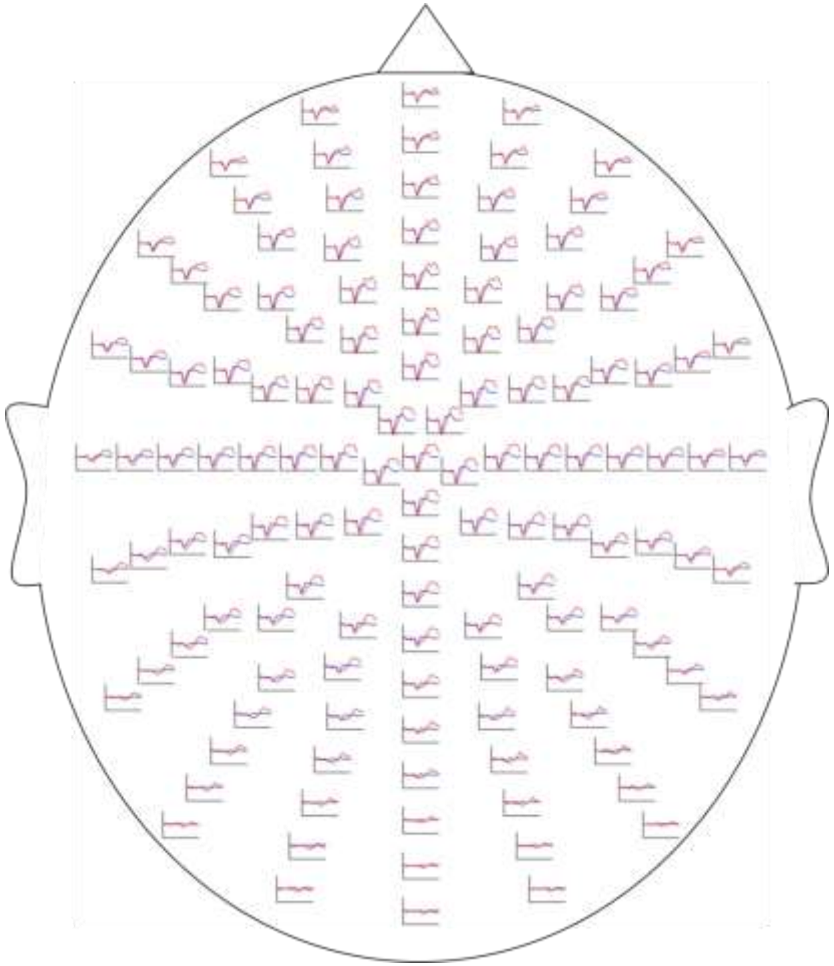
mean-field: approximate marginal posterior distributions: $\left\{ q(\vartheta_1), q(\vartheta_2) \right\}$



Bayesian inference

DCM: key model parameters





$(\theta_{21}, \theta_{32}, \theta_{13})$ state-state coupling

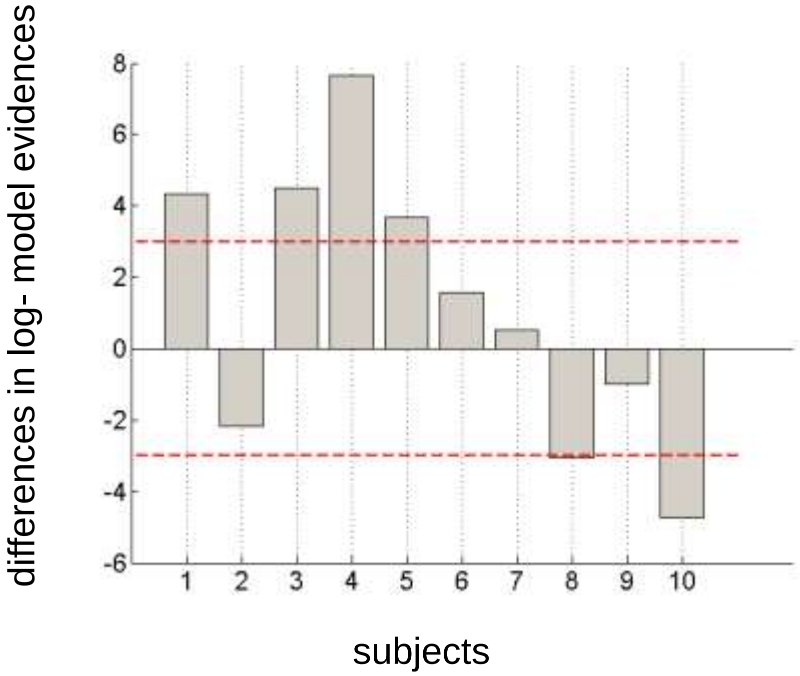
θ_3^u input-state coupling

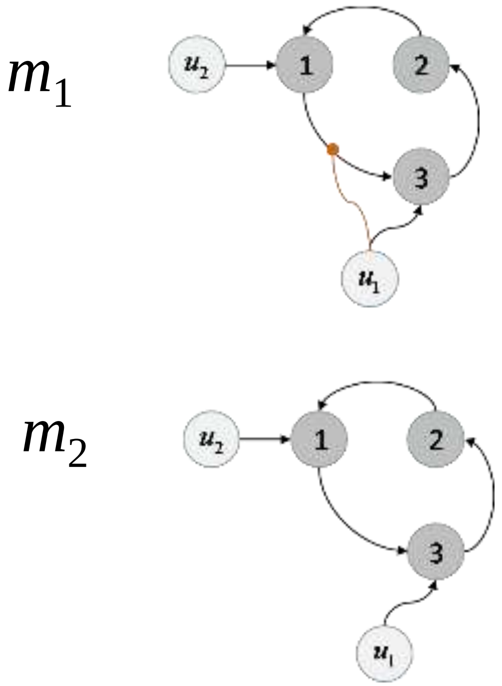
θ_{13}^u input-dependent modulatory effect

Bayesian inference

model comparison for group studies

$$\ln p(y|m_1) - \ln p(y|m_2)$$





fixed effect — assume all subjects correspond to the same model

random effect — assume different subjects might correspond to different models

Overview

1 DCM: introduction

2 Dynamical systems theory

3 Neural states dynamics

4 Bayesian inference

5 Conclusions

Conclusions

The main principle of DCM is the use of data and generative models in a Bayesian framework to infer parameters and compare models.

Implementation details may vary – e.g. variational Bayes vs. sampling methods

Model inversion is an optimization procedure where the objective function is the free energy which approximates the model evidence.

Model evidence is the goodness of fit expected under the prior parameter values.

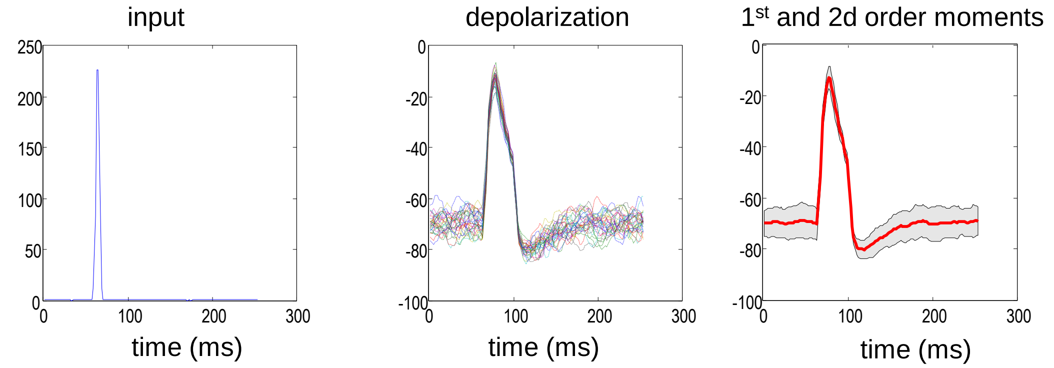
The best model is the one with precise priors that yield good fit to the data.

Different models can be compared as long as they were fitted to the same data.

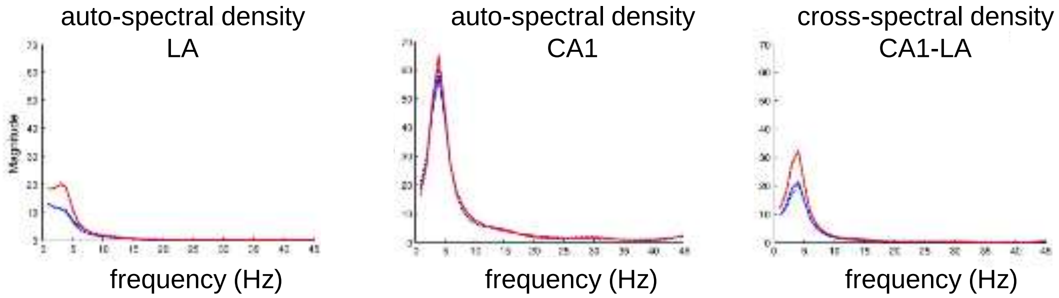
Models and priors can be gradually refined from one study to the next, making it possible to use DCM as an integrative framework in neuroscience.

DCM for EEG/MEG: variants

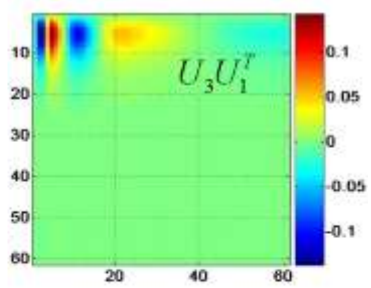
- mean-field DCM for evoked responses
- second-order mean-field DCM



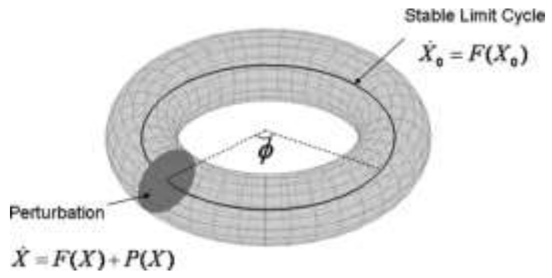
- DCM for steady-state responses



- DCM for induced responses



- DCM for phase coupling



Many thanks to:

Karl J. Friston (UCL, London, UK)
Jean Daunizeau (ICM, Paris, France)
Vladimir Litvak (UCL, London, UK)
Will D. Penny (UCL, London, UK)
Klaas E. Stephan (UZH, Zurich, Switzerland)
Stefan Kiebel (MPI, Leipzig, Germany)