

# Dynamic causal modelling for fMRI

*André Marreiros*



Friday 18<sup>th</sup> May 2012

**Statistical Parametric Mapping  
for fMRI  
2012 course**

# Overview

---

## Brain connectivity: types & definitions

*Anatomical connectivity*

*Functional connectivity*

*Effective connectivity*

## Dynamic causal models (DCMs)

*Neuronal model*

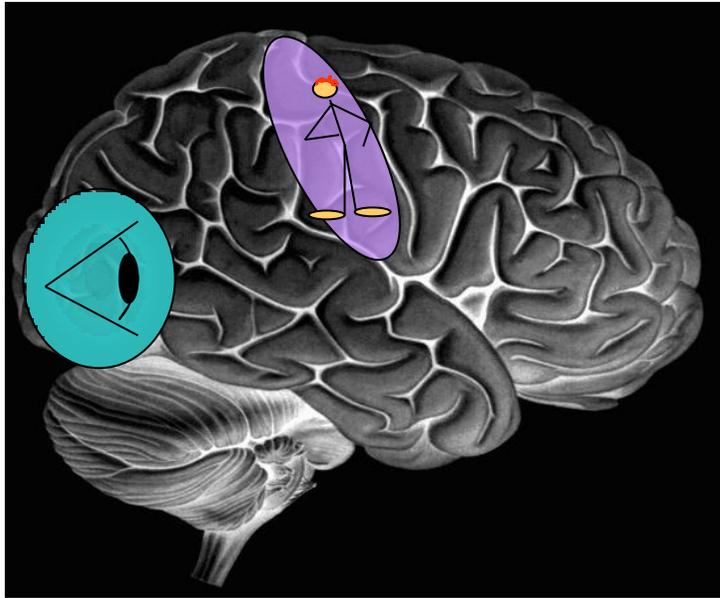
*Hemodynamic model*

*Estimation: Bayesian framework*

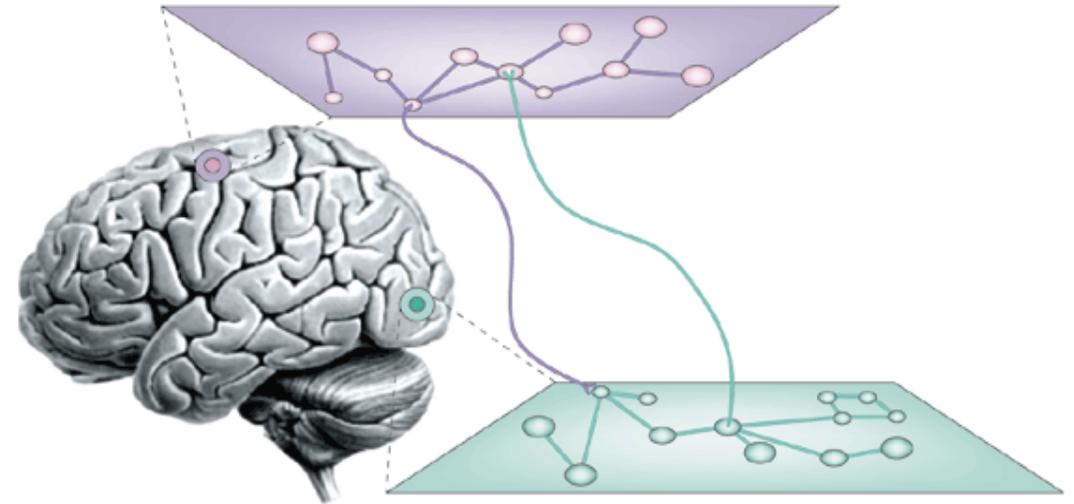
Applications & extensions of DCM to fMRI data

# Principles of Organisation

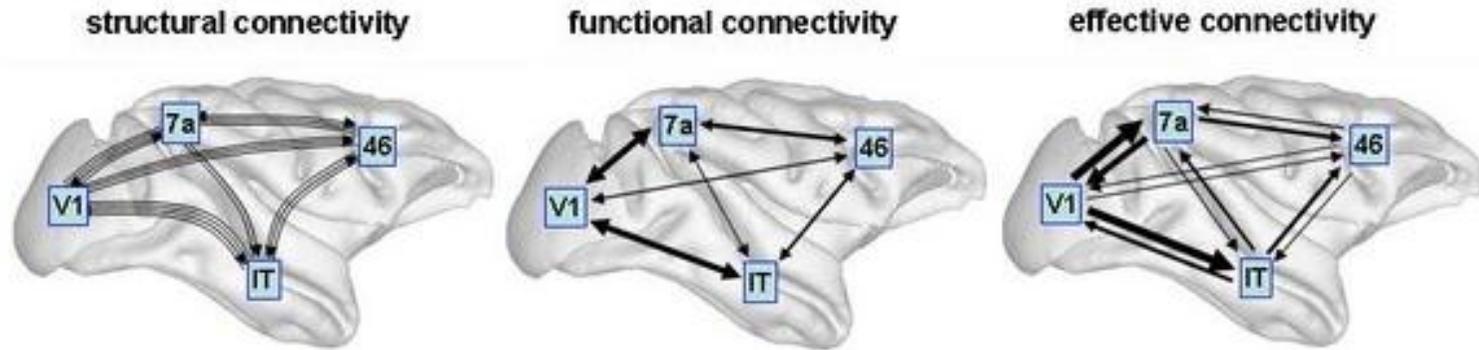
Functional specialization



Functional integration



# Structural, functional & effective connectivity



Sporns 2007, *Scholarpedia*

- **anatomical/structural connectivity**  
= presence of axonal connections
- **functional connectivity**  
= statistical dependencies between regional time series
- **effective connectivity**  
= causal (directed) influences between neurons or neuronal populations

**MECHANISM-FREE**

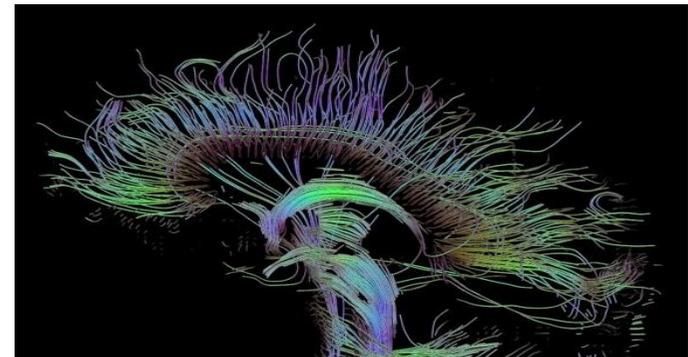
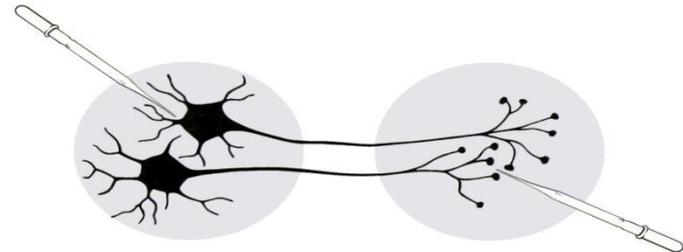
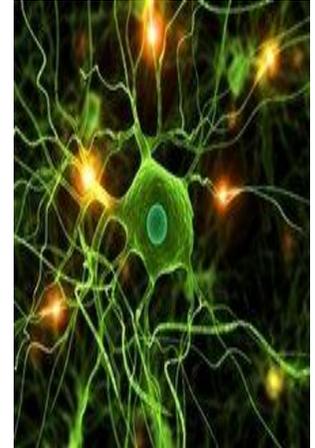
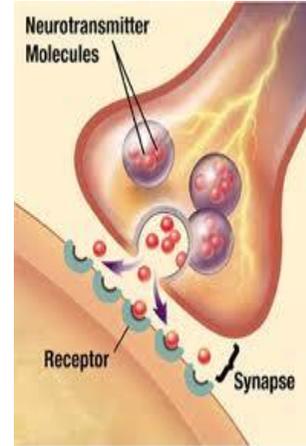
**MECHANISTIC MODEL**

# Anatomical connectivity

*Definition:*

*presence of axonal connections*

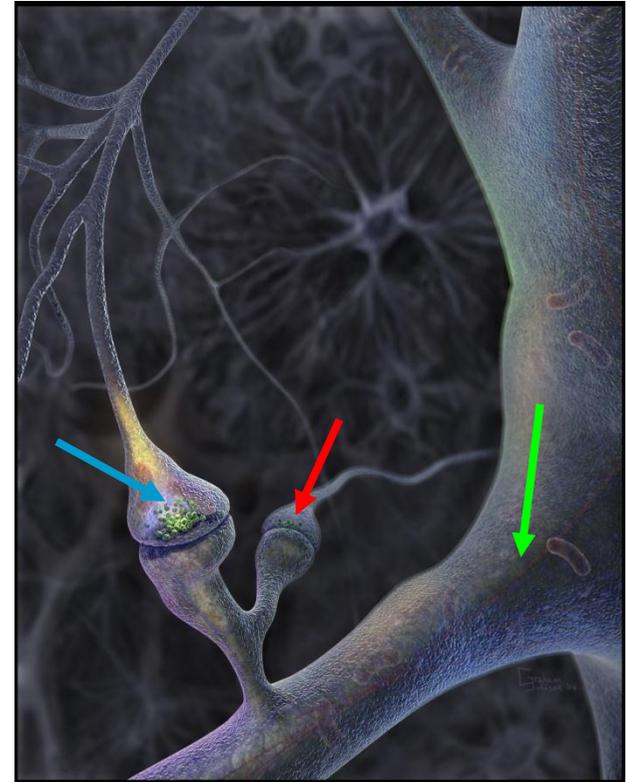
- neuronal communication via synaptic contacts
- Measured with
  - tracing techniques
  - diffusion tensor imaging (DTI)



# Knowing anatomical connectivity is not enough...

- Context-dependent recruiting of connections :
  - Local functions depend on network activity
- Connections show synaptic plasticity
  - change in the structure and transmission properties of a synapse
  - even at short timescales

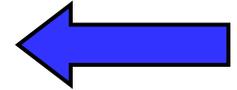
→ Look at functional and effective connectivity



# Functional connectivity

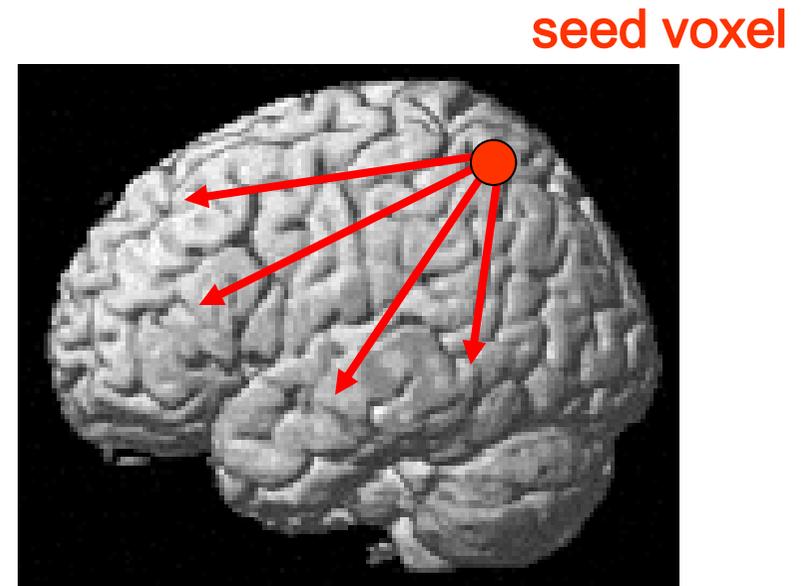
*Definition: statistical dependencies between regional time series*

- Seed voxel correlation analysis
- Coherence analysis
- Eigen-decomposition (PCA, SVD)
- Independent component analysis (ICA)
- any technique describing statistical dependencies amongst regional time series



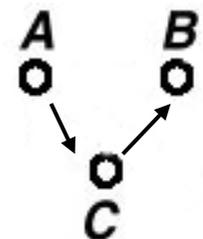
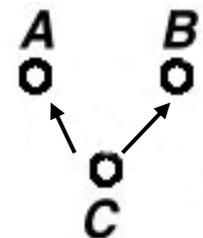
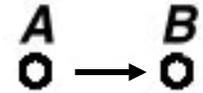
# Seed-voxel correlation analyses

- hypothesis-driven choice of a seed voxel
- extract reference time series
- voxel-wise correlation with time series from all other voxels in the brain



# Pros & Cons of functional connectivity analysis

- Pros:
  - useful when we have no experimental control over the system of interest and no model of what caused the data (e.g. sleep, hallucinations, etc.)
- Cons:
  - interpretation of resulting patterns is difficult / arbitrary
  - no mechanistic insight
  - usually suboptimal for situations where we have a priori knowledge / experimental control

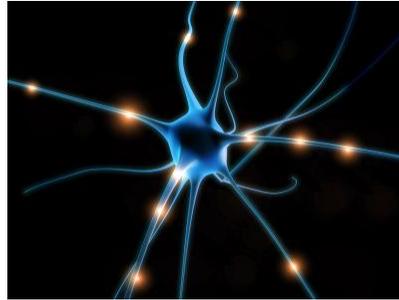


→ Effective connectivity

# Effective connectivity

*Definition: causal (directed) influences between neurons or neuronal populations*

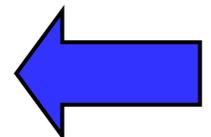
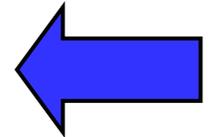
- *In vivo* and *in vitro* stimulation and recording



- Models of **causal interactions** among neuronal populations
  - explain *regional effects* in terms of *interregional connectivity*

# Some models for computing effective connectivity from fMRI data

- Structural Equation Modelling (SEM)  
McIntosh et al. 1991, 1994; Büchel & Friston 1997; Bullmore et al. 2000
- regression models  
(e.g. psycho-physiological interactions, PPIs)  
Friston et al. 1997
- Volterra kernels  
Friston & Büchel 2000
- Time series models (e.g. MAR, Granger causality)  
Harrison et al. 2003, Goebel et al. 2003
- Dynamic Causal Modelling (DCM)  
*bilinear*: Friston et al. 2003; *nonlinear*: Stephan et al. 2008



# Psychophysiological interaction (PPI)

- bilinear model of how the psychological context **A** changes the influence of area **B** on area **C** :

$$B \times A \rightarrow C$$

- A PPI corresponds to differences in regression slopes for different contexts.

# Psycho-physiological interaction (PPI)

		Task factor	
		Task A	Task B
Stimulus factor	Stim 1	A1	B1
	Stim 2	A2	B2

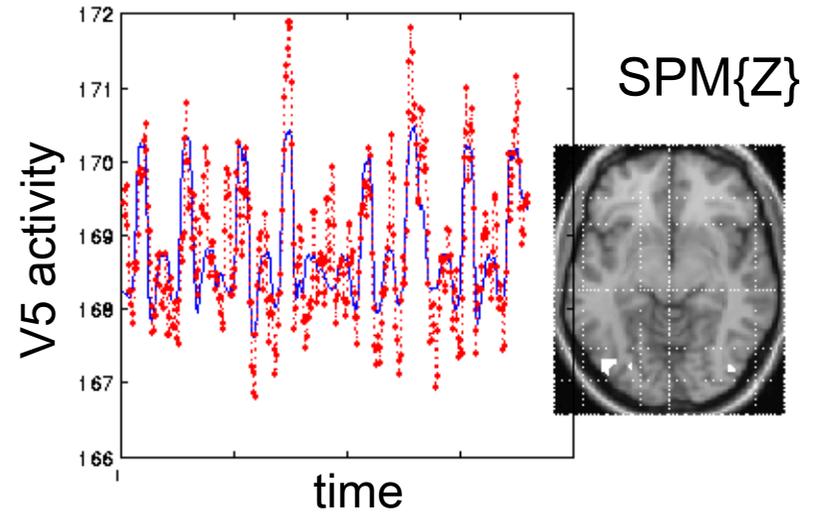
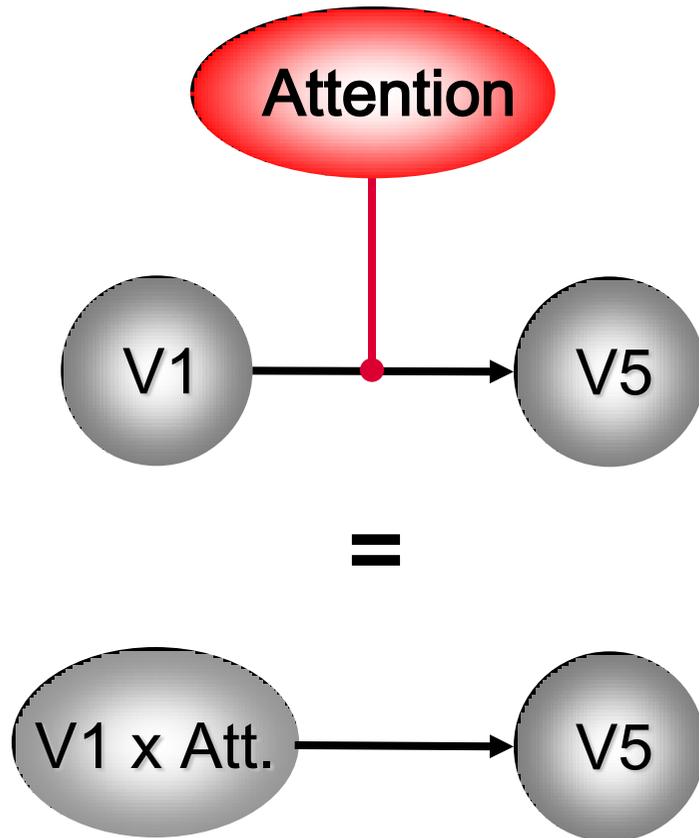
We can replace one main effect in the GLM by the time series of an area that shows this main effect.

GLM of a 2x2 factorial design:

$$\begin{aligned}
 y = & (T_A - T_B) \beta_1 && \leftarrow \text{main effect of task} \\
 & + (S_1 - S_2) \beta_2 && \leftarrow \text{main effect of stim. type} \\
 & + (T_A - T_B) (S_1 - S_2) \beta_3 && \text{interaction} \\
 & + e
 \end{aligned}$$

$$\begin{aligned}
 y = & (T_A - T_B) \beta_1 && \leftarrow \text{main effect of task} \\
 & + V1 \beta_2 && \leftarrow \text{V1 time series} \\
 & + (T_A - T_B) V1 \beta_3 && \leftarrow \text{psycho-physiological interaction} \\
 & + e
 \end{aligned}$$

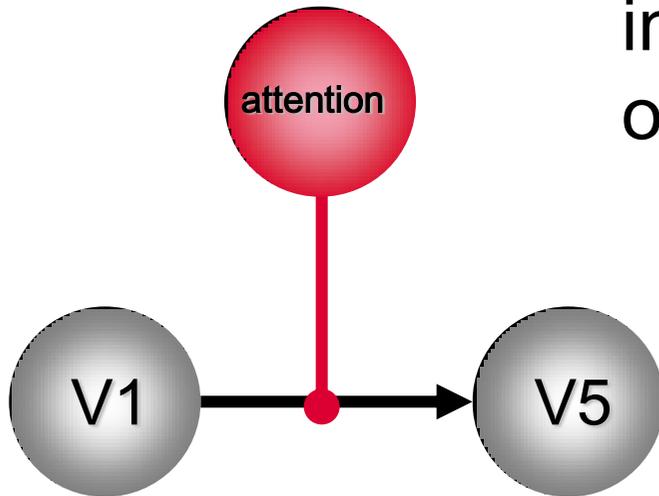
# Example PPI: Attentional modulation of V1→V5CC



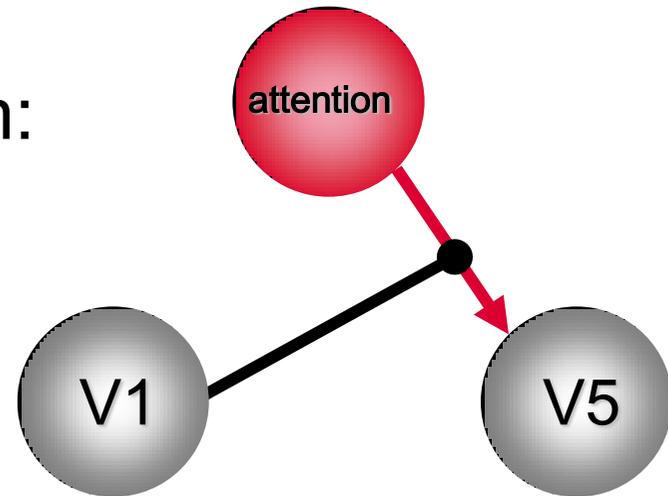
# PPI: Interpretation

$$y = (T_A - T_B) \beta_1 + V1\beta_2 + (T_A - T_B)V1\beta_3 + e$$

Two possible interpretations of the PPI term:



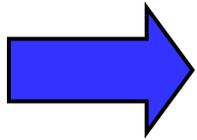
Modulation of V1→V5 by attention



Modulation of the impact of attention on V5 by V1

# Pros & Cons of PPIs

- Pros:
  - given a single source region, we can test for its context-dependent connectivity across the entire brain
  - easy to implement
- Cons:
  - only allows to model contributions from a single area
  - operates at the level of BOLD time series (SPM 99/2).  
SPM 5/8 deconvolves the BOLD signal to form the proper interaction term, and then reconvolves it.
  - ignores time-series properties of the data



*Dynamic Causal Models*

needed for more robust statements of effective connectivity.

# Overview

---

## Brain connectivity: types & definitions

*Anatomical connectivity*

*Functional connectivity*

*Effective connectivity*

## Dynamic causal models (DCMs)

*Basic idea*

*Neuronal model*

*Hemodynamic model*

*Parameter estimation, priors & inference*

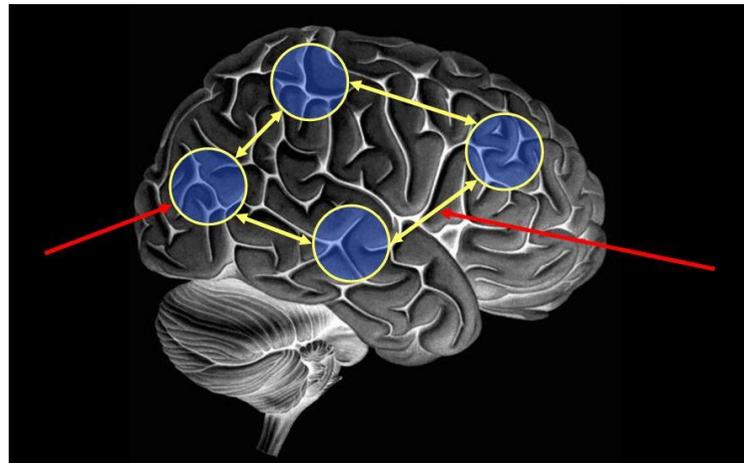
Applications & extensions of DCM to fMRI data

# Basics of Dynamic Causal Modelling

**DCM allows us to look at how areas within a network interact:**

Investigate functional integration & modulation of specific cortical pathways

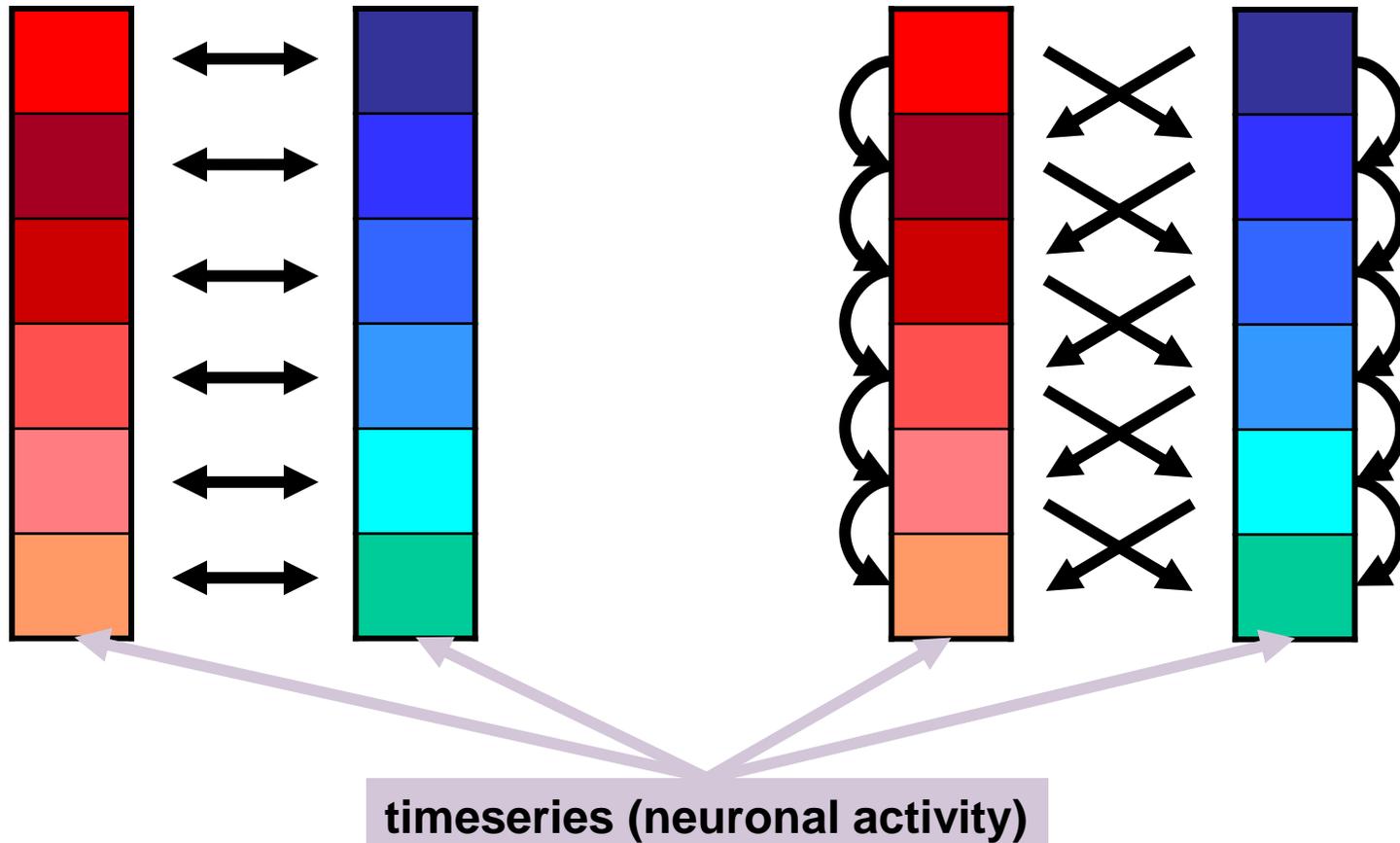
- Temporal dependency of activity within and between areas (causality)



# Temporal dependence and causal relations

Seed voxel approach, PPI etc.

Dynamic *Causal* Models

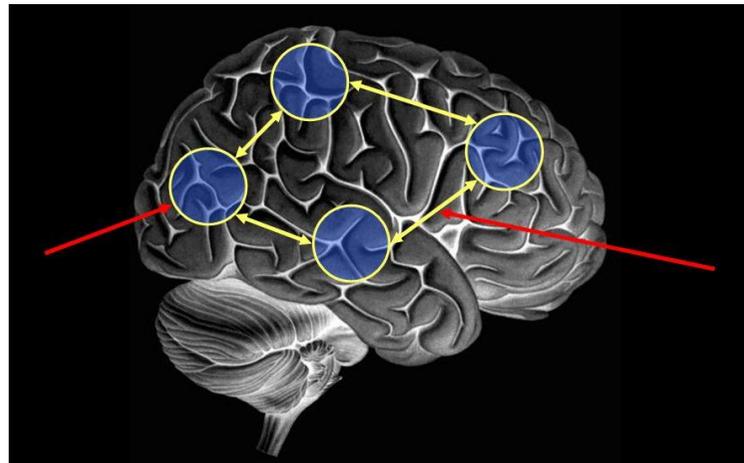


# Basics of Dynamic Causal Modelling

**DCM allows us to look at how areas within a network interact:**

Investigate functional integration & modulation of specific cortical pathways

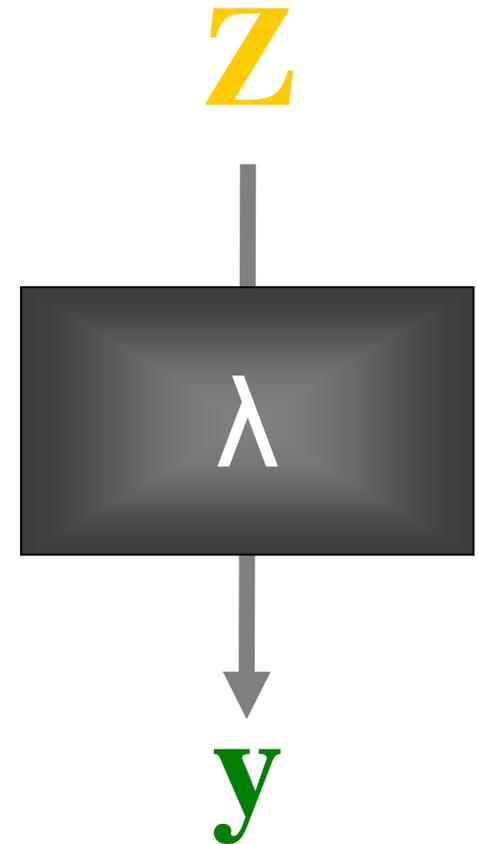
- Temporal dependency of activity within and between areas (causality)
- Separate neuronal activity from observed BOLD responses



# Basics of DCM: Neuronal and BOLD level

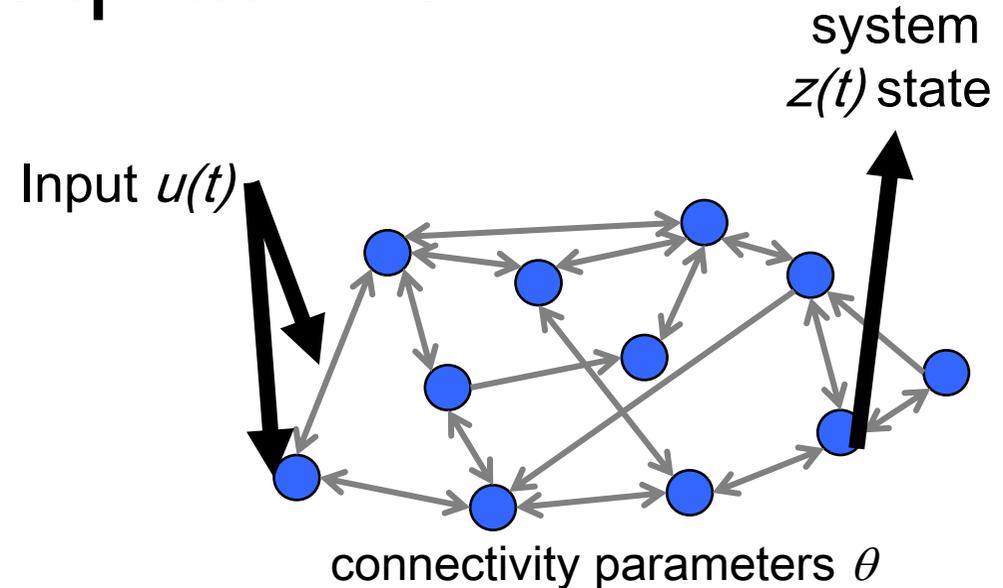
- Cognitive system is modelled at its underlying neuronal level (not directly accessible for fMRI).
- The modelled neuronal dynamics ( $Z$ ) are transformed into area-specific BOLD signals ( $y$ ) by a hemodynamic model ( $\lambda$ ).

The aim of DCM is to estimate parameters at the neuronal level such that the modelled and measured BOLD signals are optimally similar.



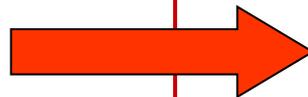
# Neuronal systems are represented by differential equations

A System is a set of elements  $z_n(t)$  which interact in a spatially and temporally specific fashion



State changes of the system states are dependent on:

- the current state  $z$
- external inputs  $u$
- its connectivity  $\theta$
- time constants & delays



$$\frac{dz}{dt} = F(z, u, \theta)$$

# DCM parameters = rate constants

*Generic solution to the ODEs in DCM:*

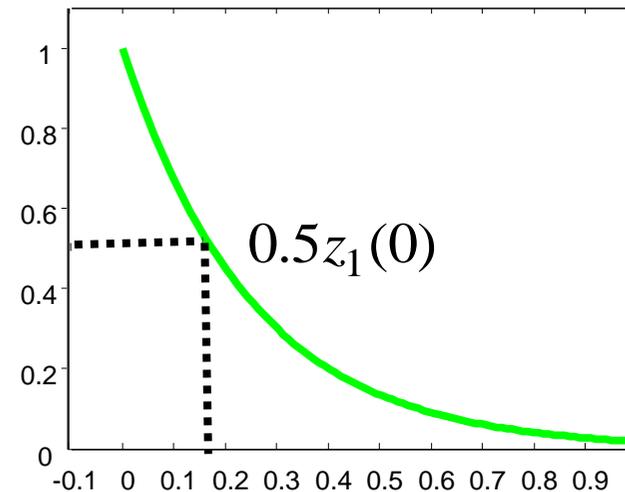
  $\frac{dz_1}{dt} = -sz_1 \quad \longrightarrow \quad z_1(t) = z_1(0) \exp(-st), \quad z_1(0) = 1$

*Half-life  $\tau$*

$$\begin{aligned} z_1(\tau) &= 0.5 z_1(0) \\ &= z_1(0) \exp(-s\tau) \end{aligned}$$

  $s = \ln 2 / \tau$

*Decay function*

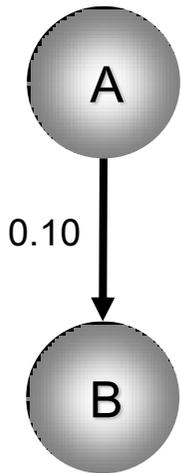


$$\tau = \ln 2 / s$$

# DCM parameters = rate constants

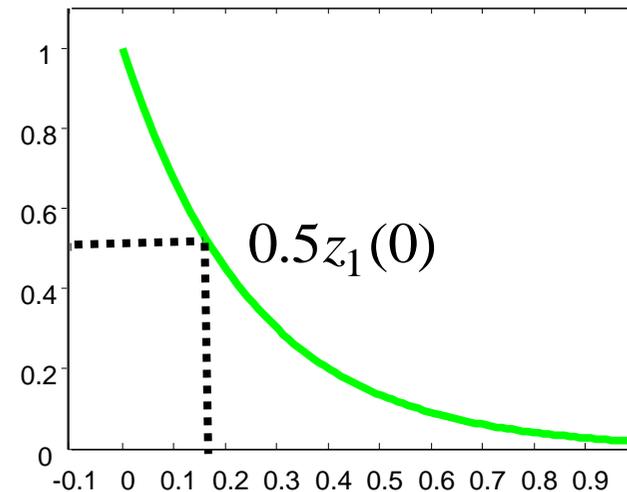
*Generic solution to the ODEs in DCM:*


$$\frac{dz_1}{dt} = -sz_1 \quad \longrightarrow \quad z_1(t) = z_1(0) \exp(-st), \quad z_1(0) = 1$$



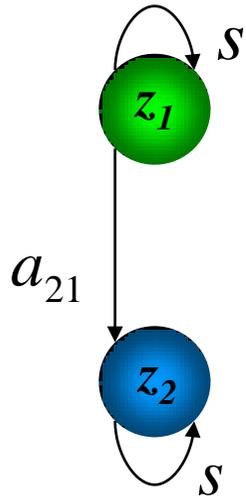
If  $A \rightarrow B$  is  $0.10 \text{ s}^{-1}$  this means that, per unit time, the increase in activity in B corresponds to 10% of the activity in A

*Decay function*



$$\tau = \ln 2 / s$$

# Linear dynamics: 2 nodes



$$\dot{z}_1 = -sz_1$$

$$\dot{z}_2 = s(a_{21}z_1 - z_2)$$

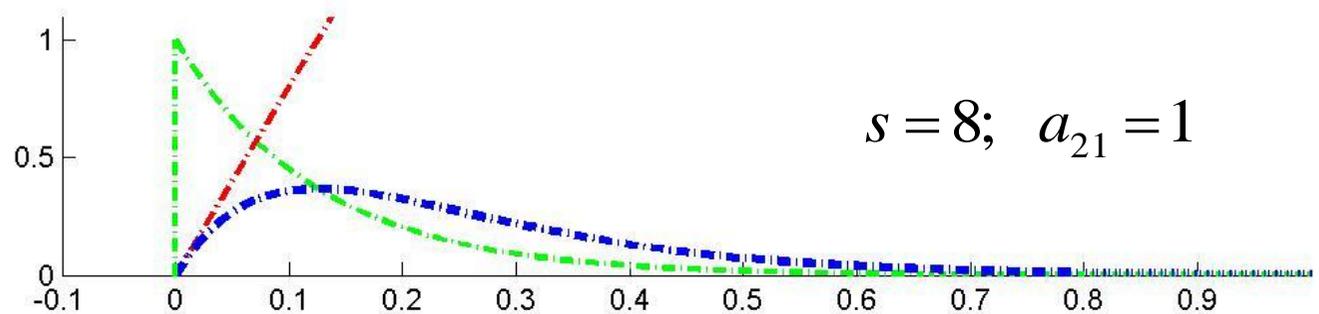
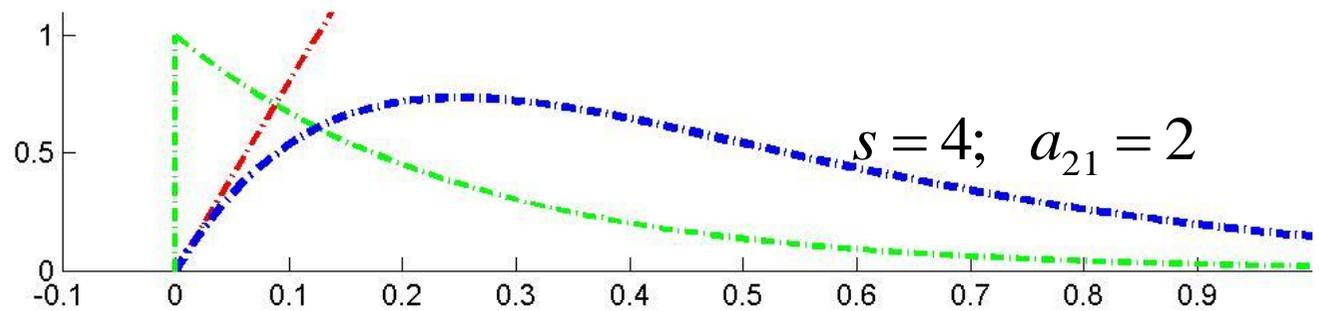
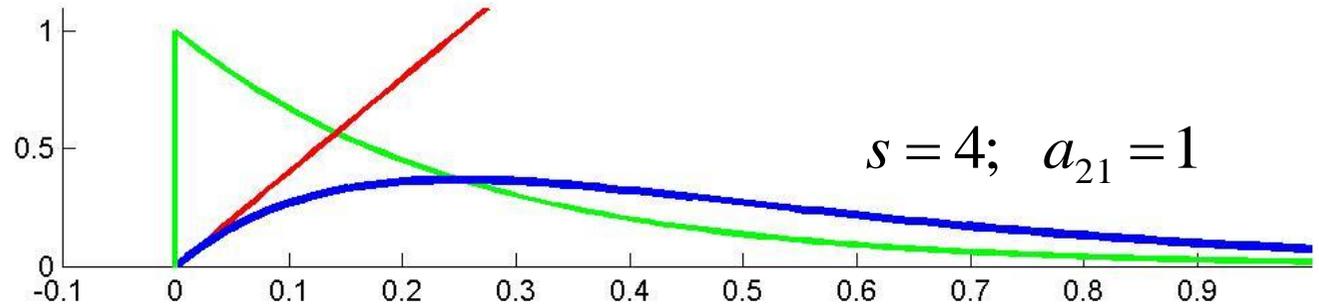
$$z_1(0) = 1$$

$$z_2(0) = 0$$

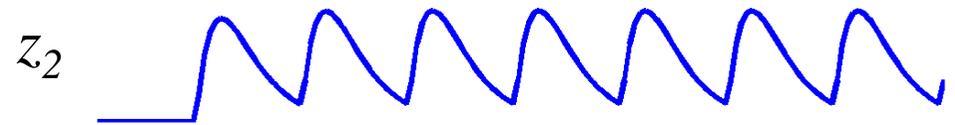
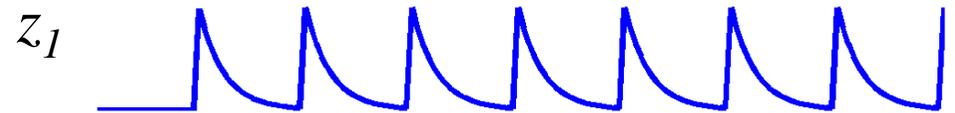
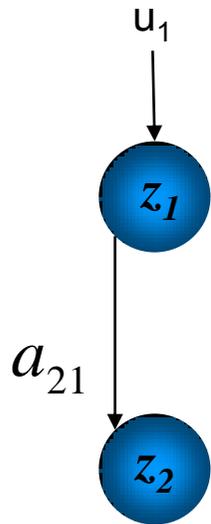
$$z_1(t) = \exp(-st)$$

$$z_2(t) = sa_{21}t \exp(-st)$$

$$a_{21} > 0$$



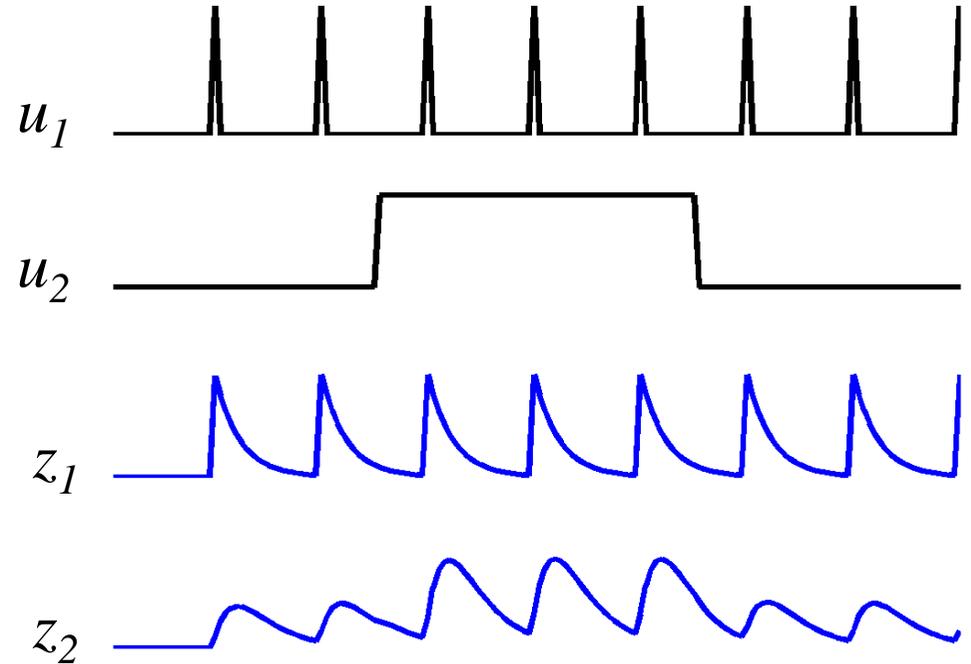
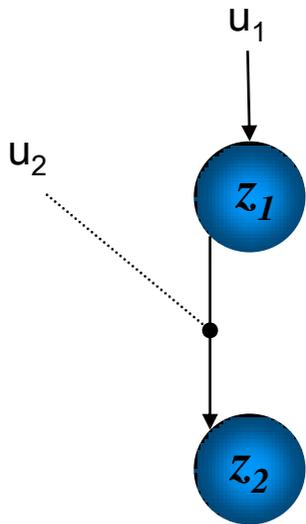
# Neurodynamics: 2 nodes with input



$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = s \begin{bmatrix} -1 & 0 \\ a_{21} & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} c \\ 0 \end{bmatrix} u_1 \quad a_{21} > 0$$

activity in  $z_2$  is coupled to  $z_1$  via coefficient  $a_{21}$

# Neurodynamics: positive modulation



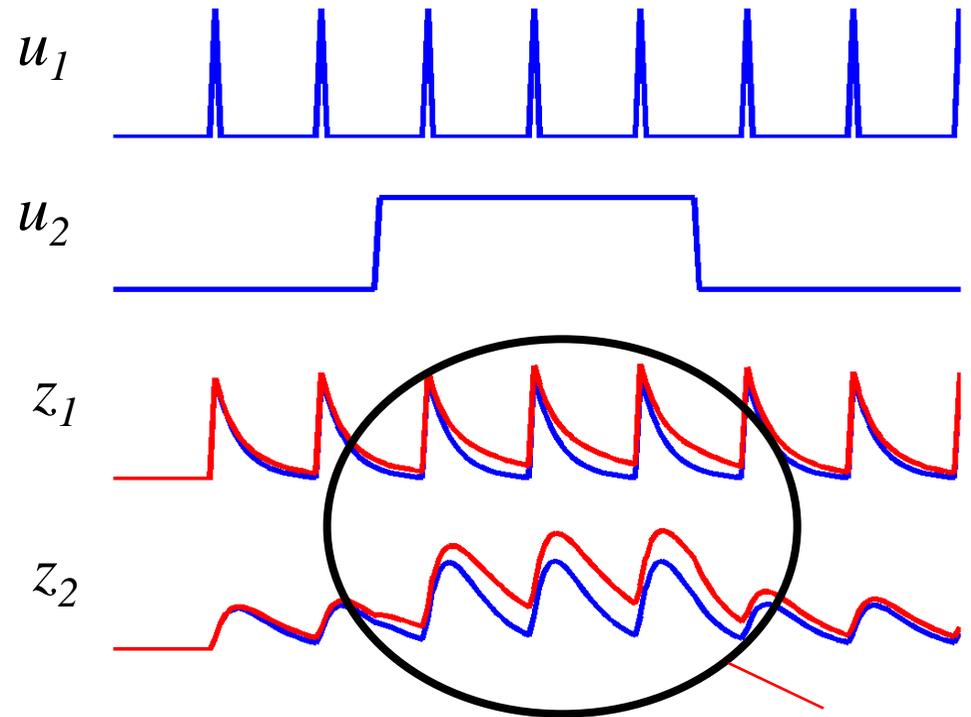
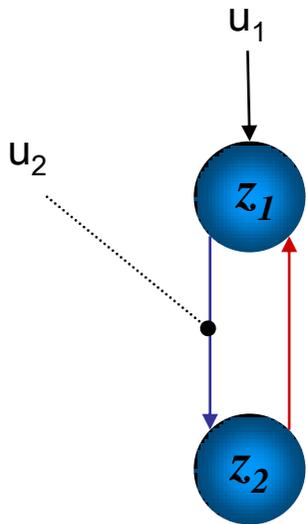
$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = s \begin{bmatrix} -1 & 0 \\ a_{21} & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + u_2 \begin{bmatrix} 0 & 0 \\ b_{21}^2 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} c \\ 0 \end{bmatrix} u_1$$

index, not squared

$$b_{21}^2 > 0$$

modulatory input  $u_2$  activity through the coupling  $a_{21}$

# Neurodynamics: reciprocal connections

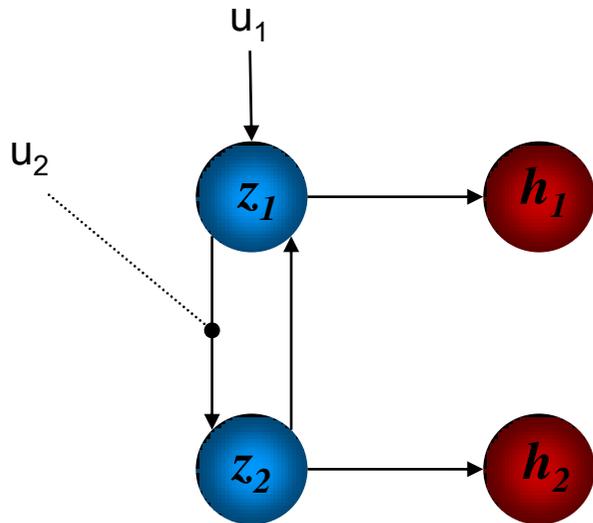


reciprocal connection  
disclosed by  $u_2$

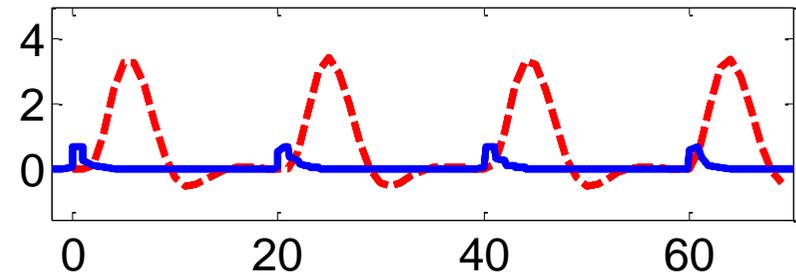
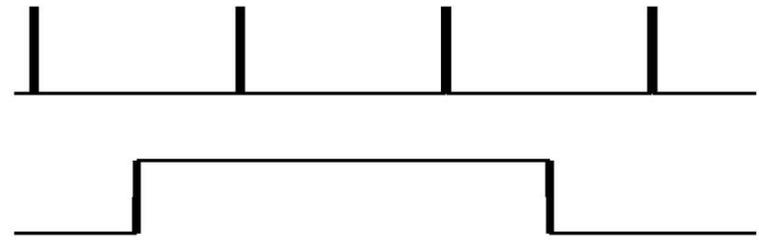
$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = s \begin{bmatrix} -1 & a_{12} \\ a_{21} & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + u_2 \begin{bmatrix} 0 & 0 \\ b_{21}^2 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} c \\ 0 \end{bmatrix} u_1$$

$$a_{12}, a_{21}, b_{21}^2 > 0$$

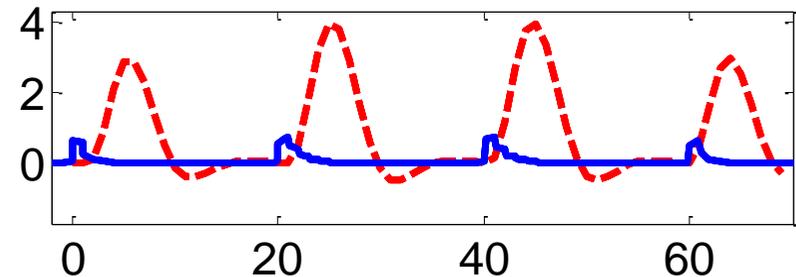
# Haemodynamics: reciprocal connections



BOLD  
(without noise)



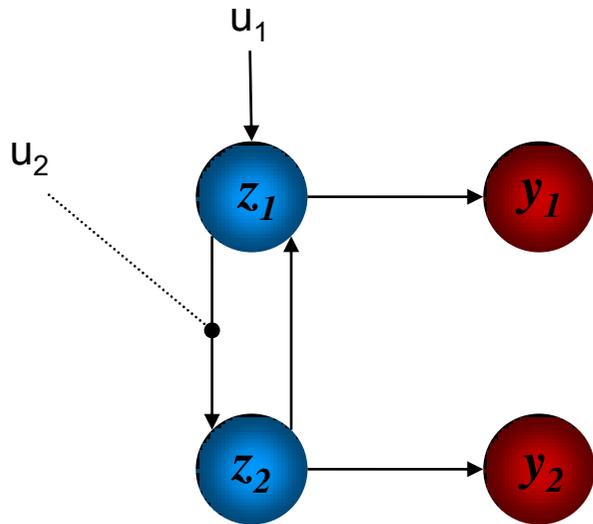
BOLD  
(without noise)



blue: neuronal activity  
red: bold response

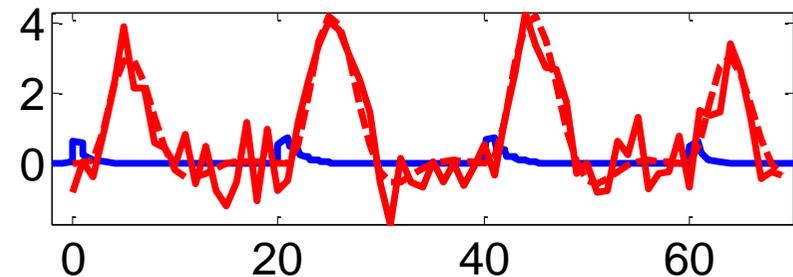
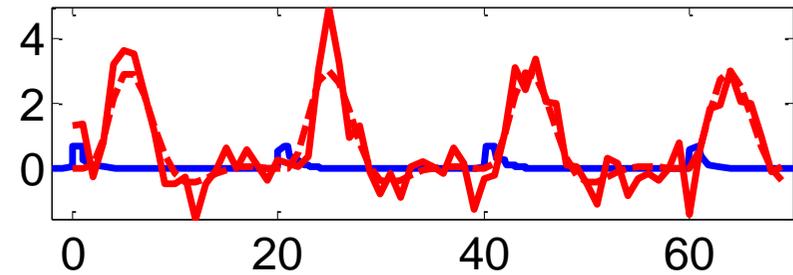
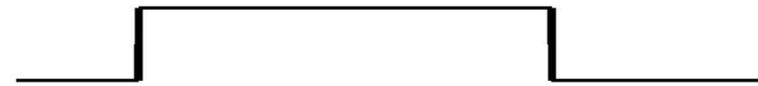
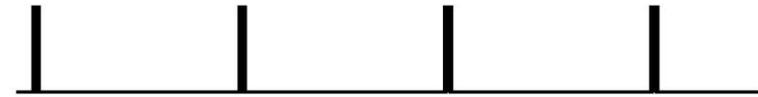
$h(u, \theta)$  represents the BOLD response (balloon model) to input

# Haemodynamics: reciprocal connections



BOLD  
with  
Noise added

BOLD  
with  
Noise added



seconds

blue: neuronal activity  
red: bold response

$y$  represents simulated observation of BOLD response, i.e. includes noise

$$y = h(u, \theta) + e$$



# Conceptual overview

Neuronal state equation  $\dot{z} = F(z, u, \theta^n)$

The bilinear model  $\dot{z} = (A + \sum u_j B^j)z + Cu$

effective connectivity

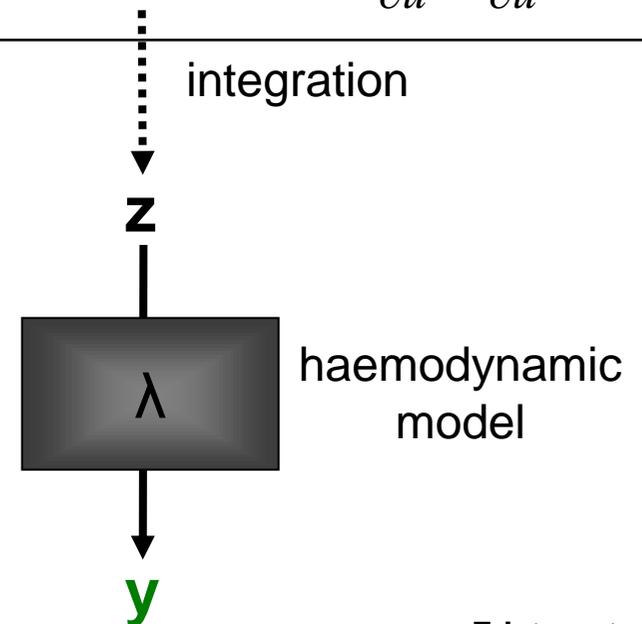
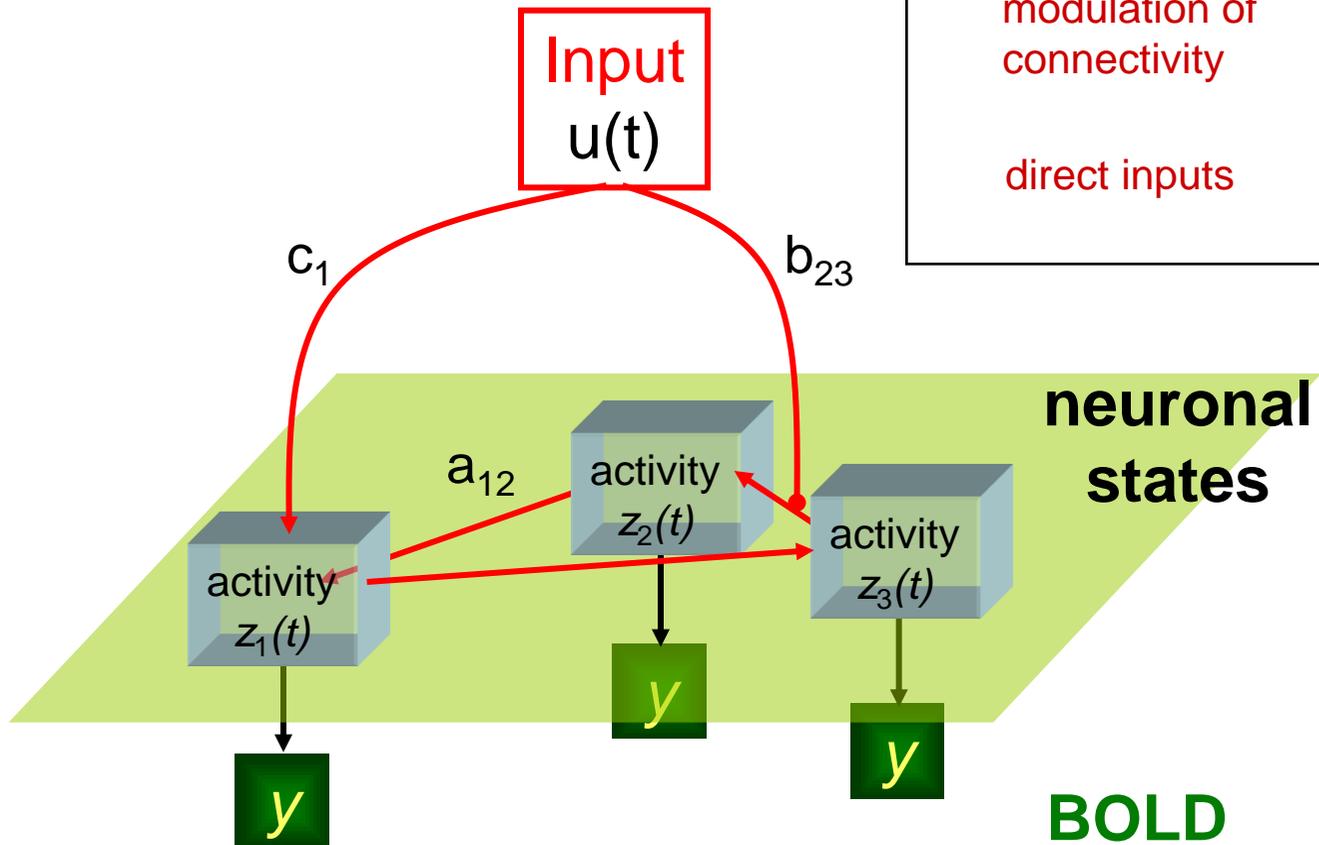
$$A = \frac{\partial F}{\partial z} = \frac{\partial \dot{z}}{\partial z}$$

modulation of connectivity

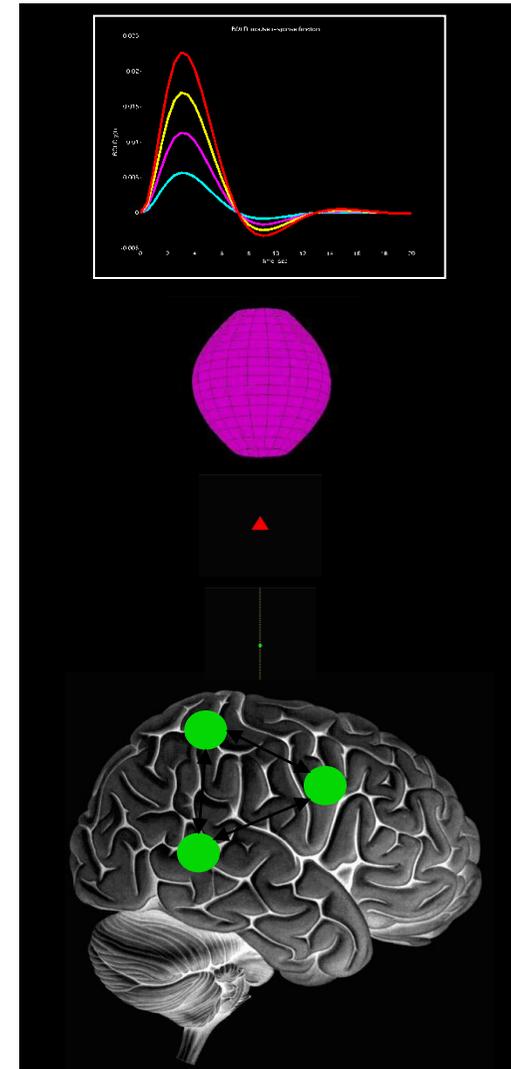
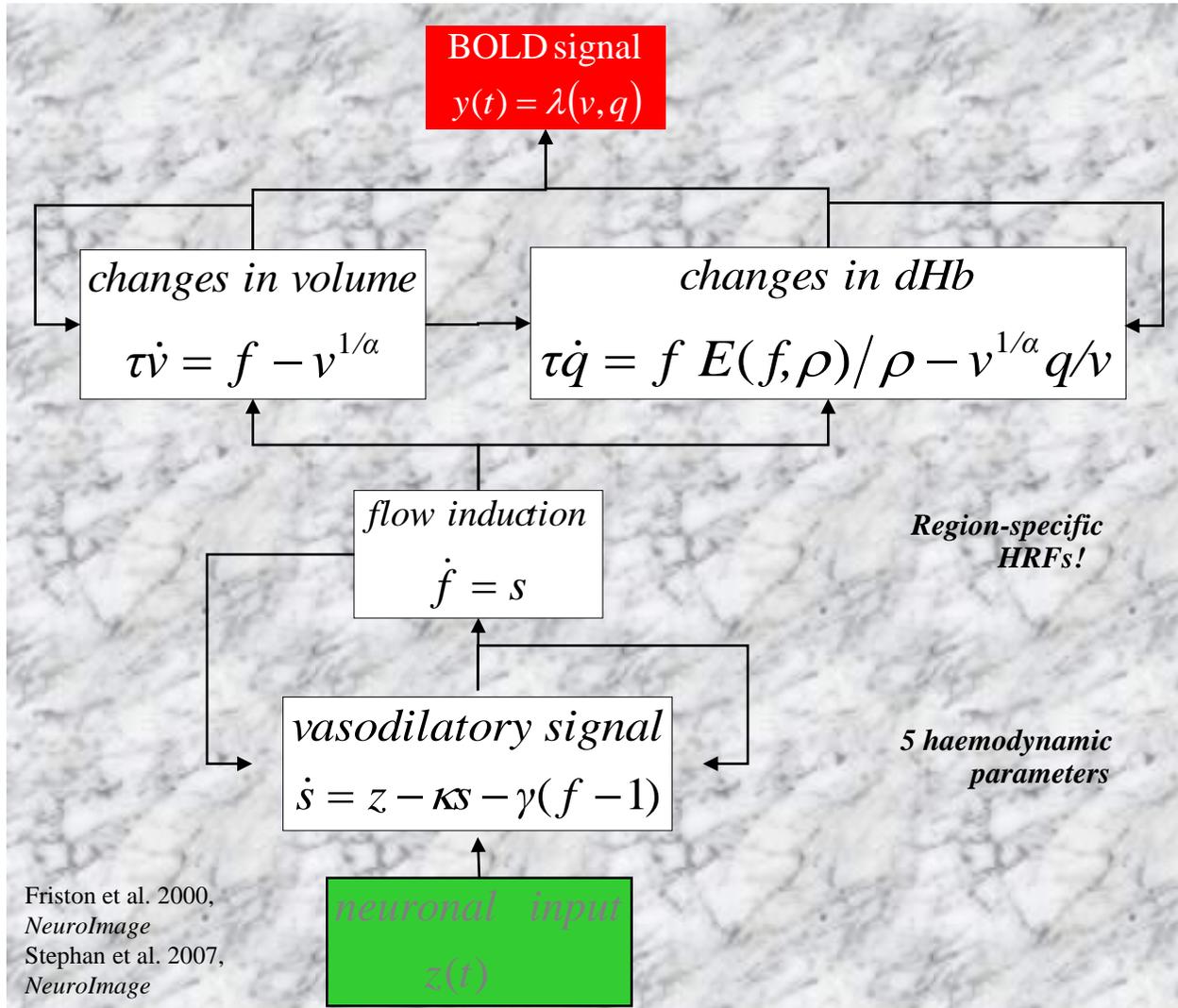
$$B^j = \frac{\partial^2 F}{\partial z \partial u_j} = \frac{\partial}{\partial u_j} \frac{\partial \dot{z}}{\partial z}$$

direct inputs

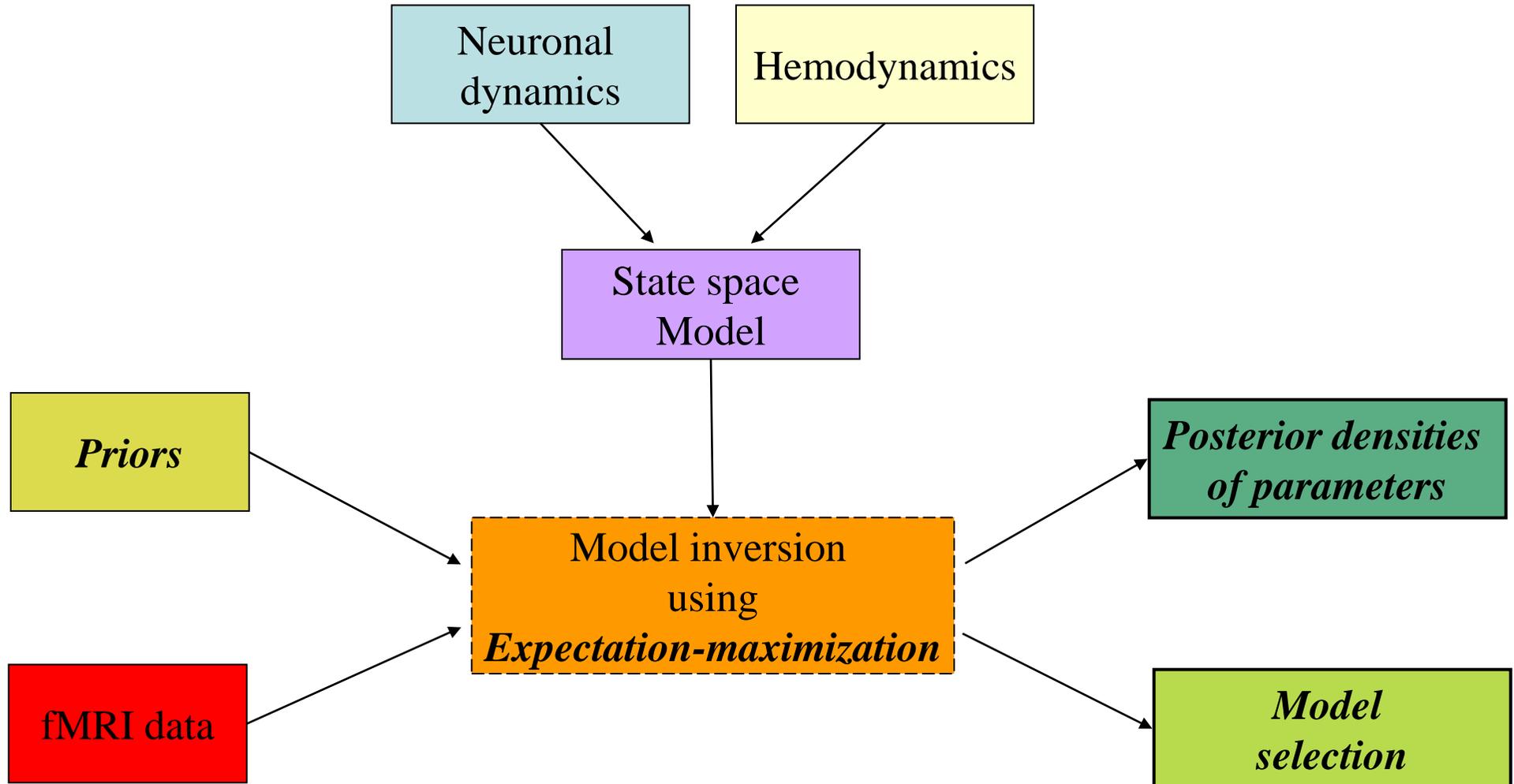
$$C = \frac{\partial F}{\partial u} = \frac{\partial \dot{z}}{\partial u}$$



# The hemodynamic “Balloon” model



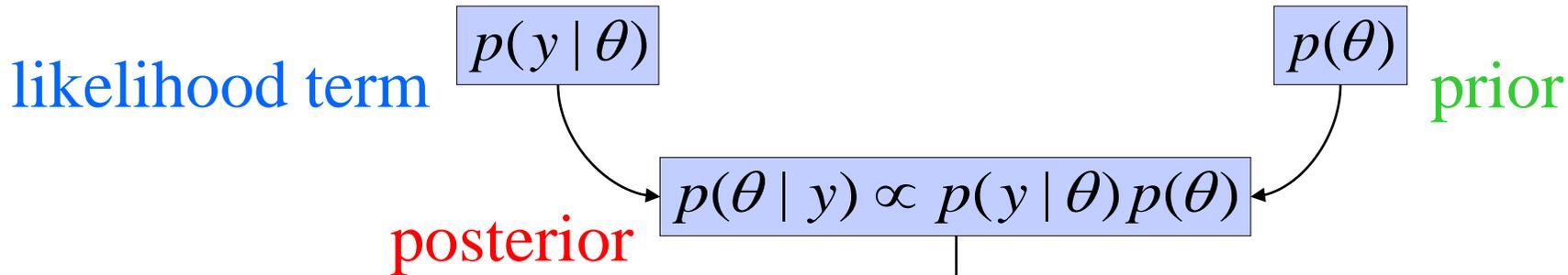
# DCM roadmap



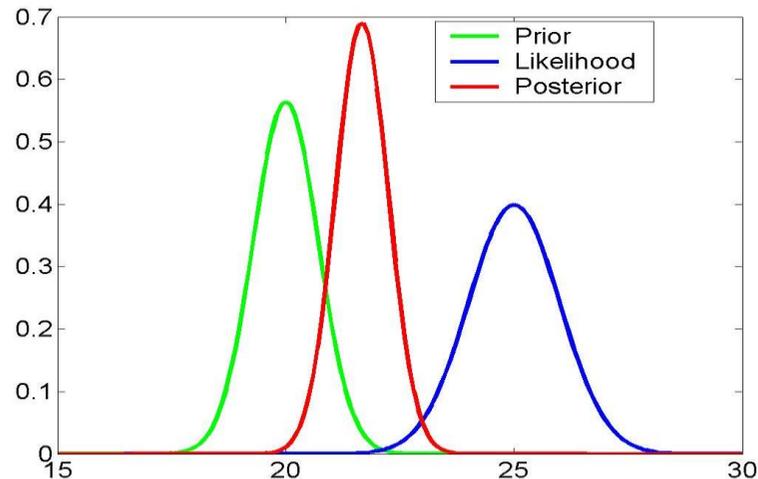
# Estimation: Bayesian framework

- Models of
- Haemodynamics in a single region
  - Neuronal interactions

- Constraints on
- Haemodynamic parameters
  - Connections

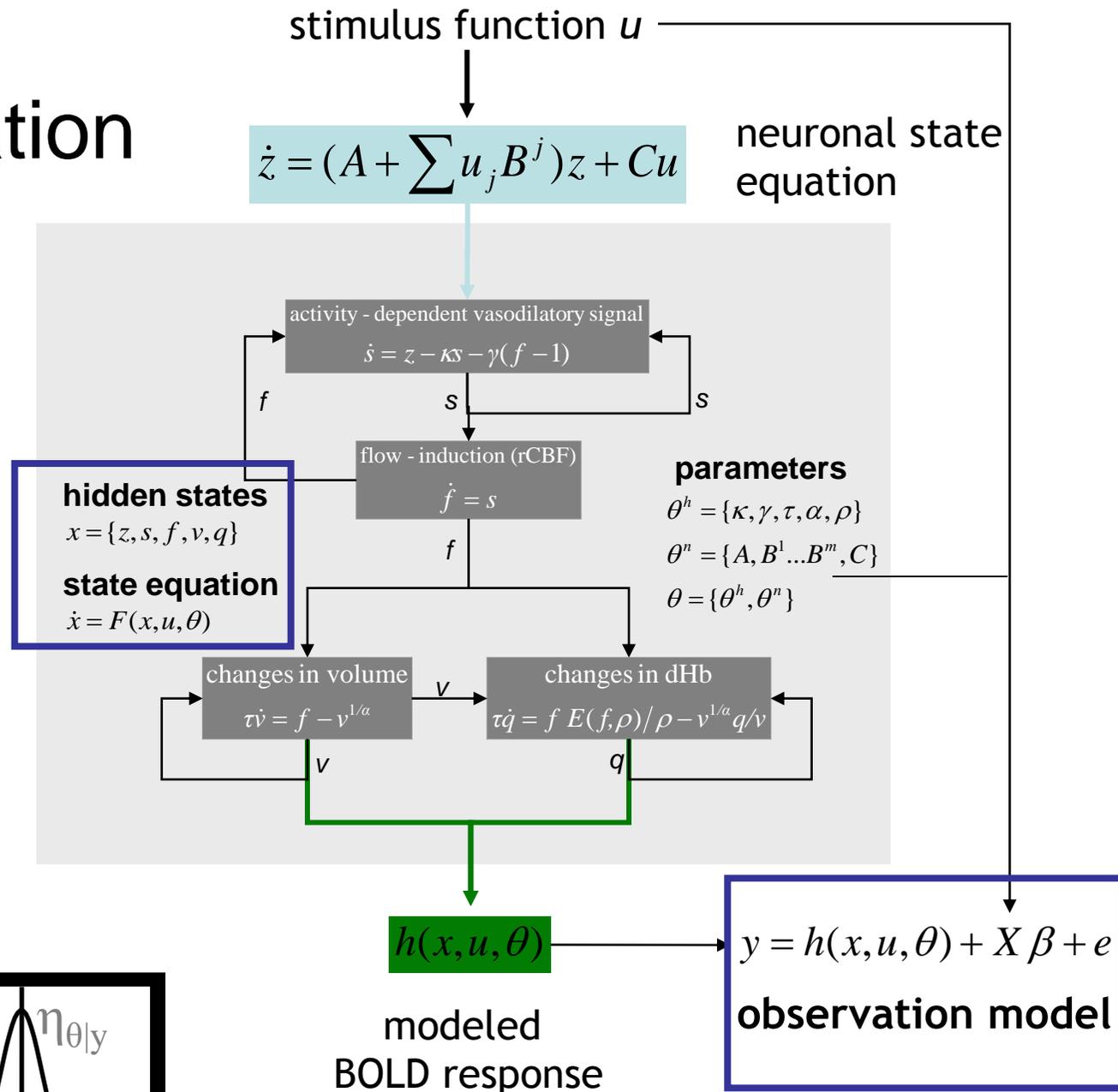
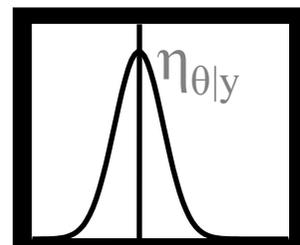


Bayesian estimation

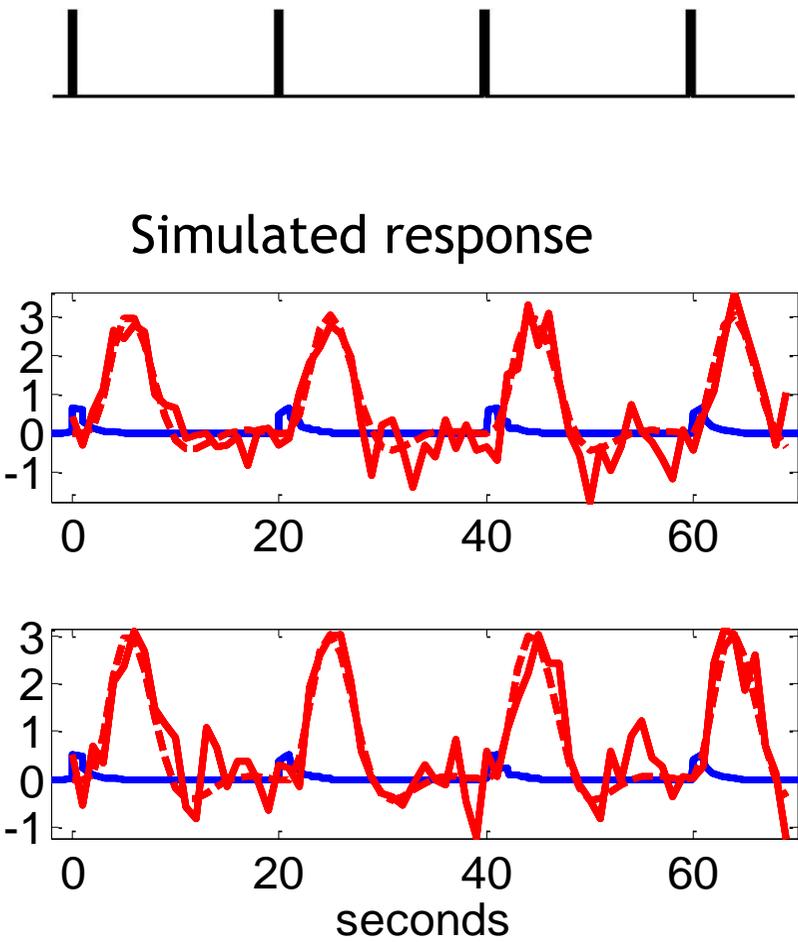
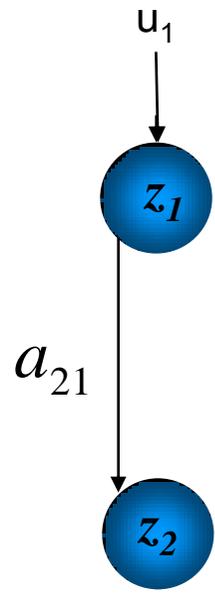


# Overview: parameter estimation

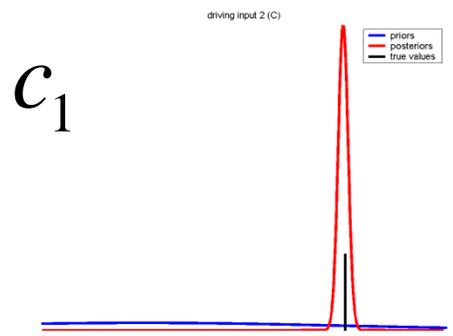
- Specify model (neuronal and haemodynamic level)
- Make it an observation model by adding measurement error  $e$  and confounds  $X$  (e.g. drift).
- Bayesian parameter estimation using expectation-maximization.
- Result:  
(Normal) posterior parameter distributions, given by mean  $\eta_{\theta|y}$  and Covariance  $C_{\theta|y}$ .



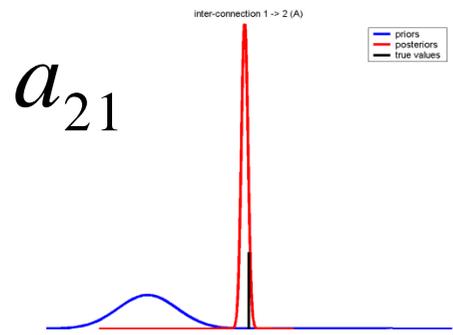
# Parameter estimation: an example



Input coupling,  $c_1$



Forward coupling,  $a_{21}$

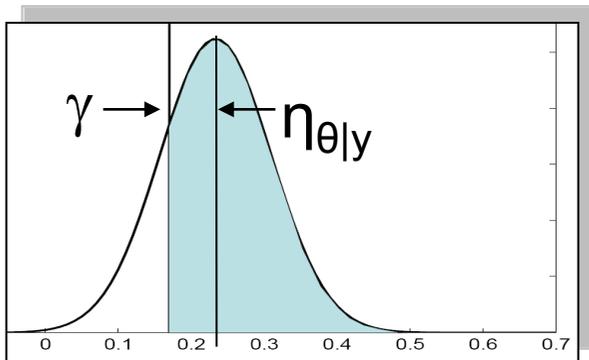


Prior density    —    Posterior density    —    true values    —

# Inference about DCM parameters

## Bayesian single subject analysis

- The model parameters are distributions that have a mean  $\eta_{\theta|y}$  and covariance  $C_{\theta|y}$ 
  - Use of the cumulative normal distribution to test the probability that a certain parameter is above a chosen threshold  $\gamma$ :

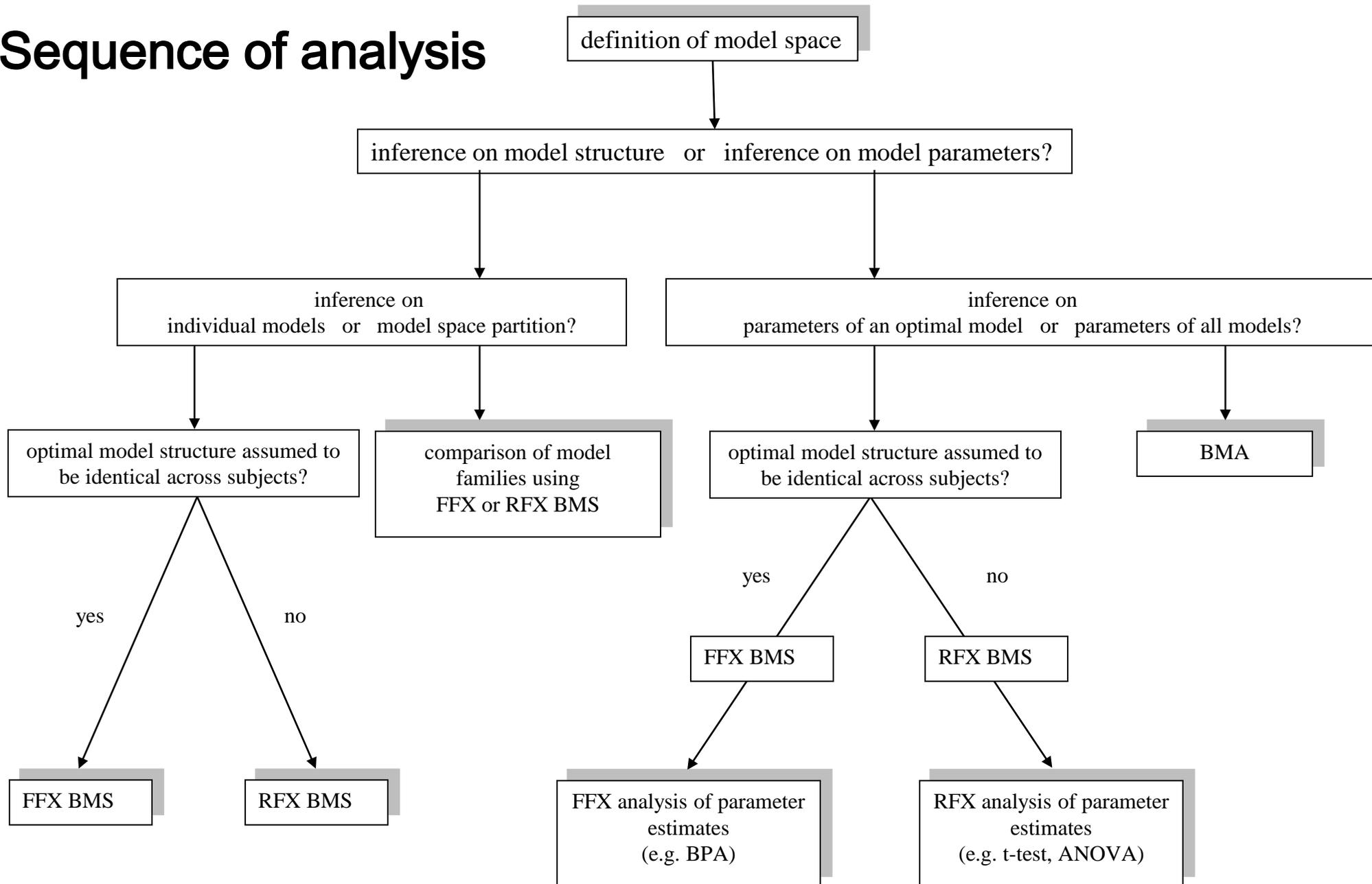


## Classical frequentist test across groups

- Test summary statistic: mean  $\eta_{\theta|y}$ 
  - One-sample t-test: Parameter  $> 0$ ?
  - Paired t-test: parameter 1  $>$  parameter 2?
  - rmANOVA: e.g. in case of multiple sessions per subject

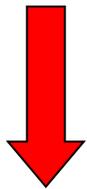
## Bayesian parameter averaging

# Sequence of analysis

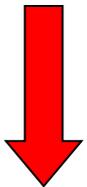


# Model comparison and selection

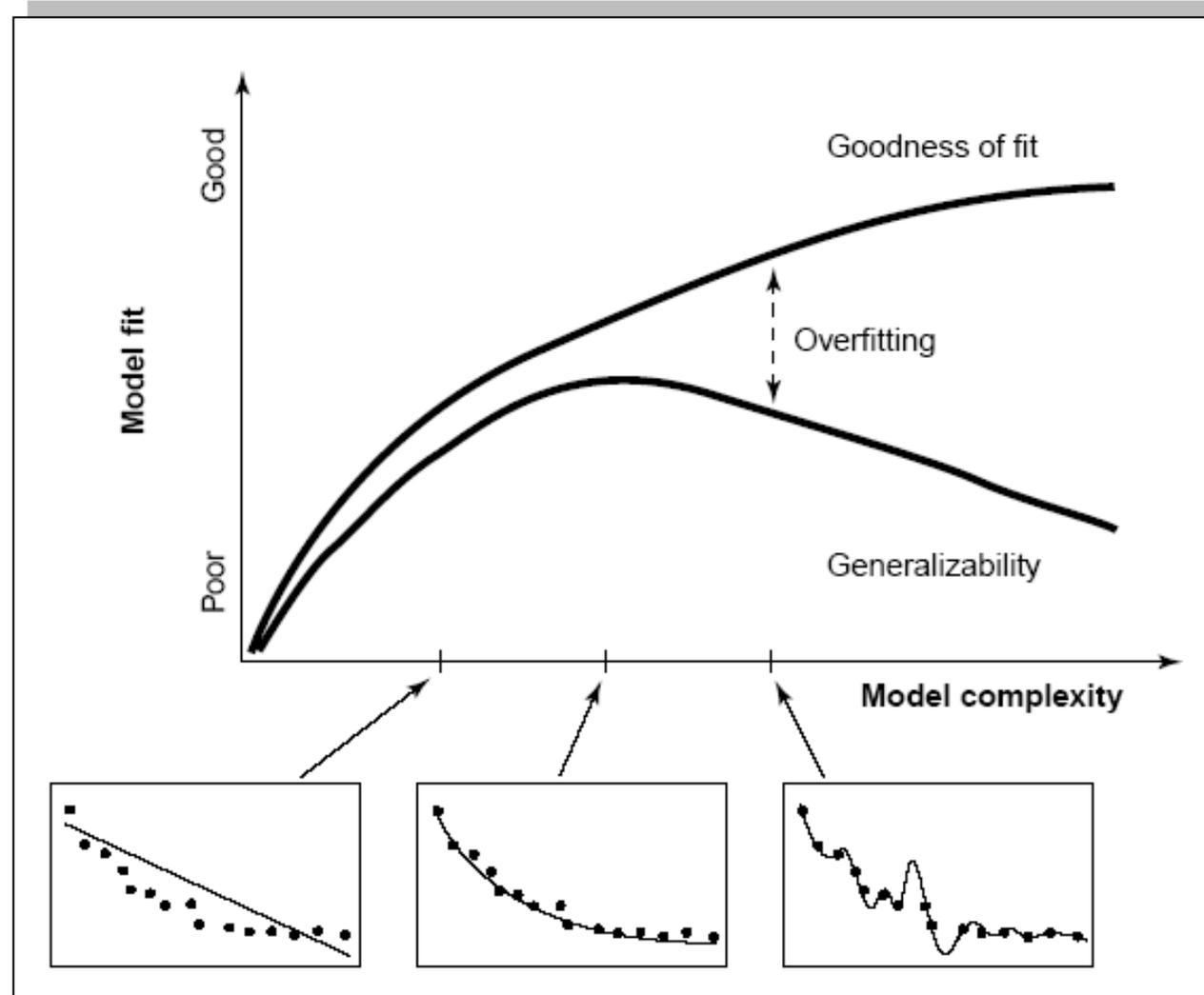
Given competing hypotheses on structure & functional mechanisms of a system, which model is the best?



Which model represents the best balance between model fit and model complexity?



For which model  $m$  does  $p(y|m)$  become maximal?



# Approximations to the model evidence in DCM

Logarithm is a  
monotonic function



Maximizing log model evidence  
= Maximizing model evidence

Log model evidence = balance between fit and complexity

$$\begin{aligned}\log p(y | m) &= \textit{accuracy}(m) - \textit{complexity}(m) \\ &= \log p(y | \theta, m) - \textit{complexity}(m)\end{aligned}$$

In SPM2 & SPM5, interface offers 2 approximations:

Akaike Information Criterion:  $AIC = \log p(y | \theta, m) - p$

No. of  
parameters

No. of  
data points

Bayesian Information Criterion:  $BIC = \log p(y | \theta, m) - \frac{p}{2} \log N$



AIC favours more complex models,  
BIC favours simpler models.

# The negative free energy approximation

- The negative free energy  $F$  is a lower bound on the log model evidence:

$$F = \log p(y | m) - KL[q(\theta), p(\theta | y, m)]$$

$F$  comprises the expected log likelihood and the Kullback-Leibler (KL) divergence between conditional and prior densities

# The complexity term in $F$

- In contrast to AIC & BIC, the complexity term of the negative free energy  $F$  accounts for parameter interdependencies. Under gaussian assumptions:

$$KL[q(\theta), p(\theta | m)] \\ = \frac{1}{2} \ln |C_{\theta}| - \frac{1}{2} \ln |C_{\theta|y}| + \frac{1}{2} (\mu_{\theta|y} - \mu_{\theta})^T C_{\theta}^{-1} (\mu_{\theta|y} - \mu_{\theta})$$

- The complexity term of  $F$  is higher
  - the more independent the prior parameters
  - the more dependent the posterior parameters
  - the more the posterior mean deviates from the prior mean
- NB: SPM8 only uses  $F$  for model selection !

# Bayes factors

For a given dataset, to compare two models, we compare their evidences.

positive value,  $[0; \infty[$

$$B_{12} = \frac{p(y | m_1)}{p(y | m_2)}$$

Kass & Raftery classification:

$B_{12}$	$p(m_1 y)$	Evidence
1 to 3	50-75%	weak
3 to 20	75-95%	positive
20 to 150	95-99%	strong
$\geq 150$	$\geq 99\%$	Very strong

or their log evidences

$$\ln(B_{12}) \approx F_1 - F_2$$

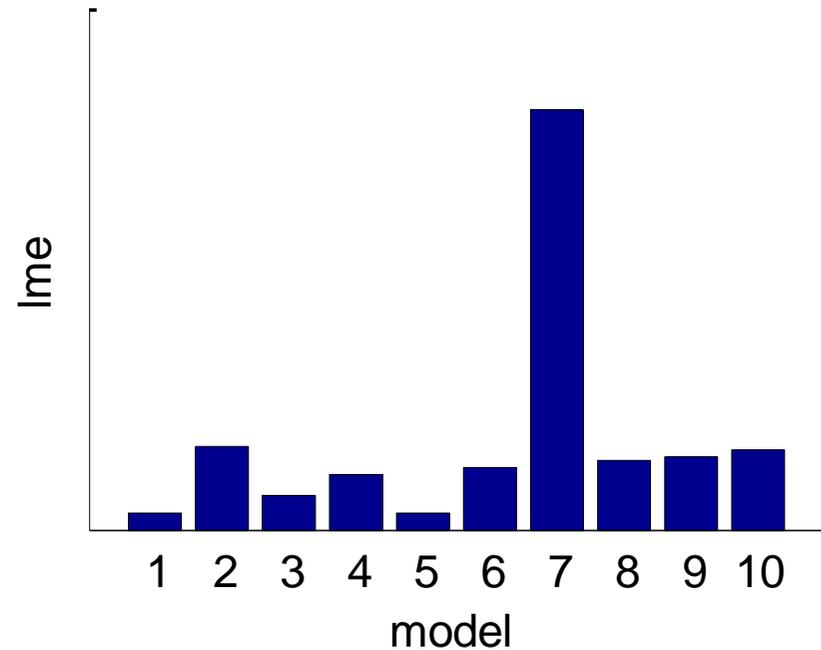
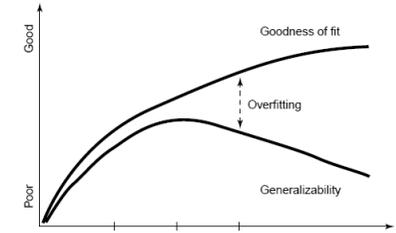
Kass & Raftery 1995, *J. Am. Stat. Assoc.*

# Inference on model space

Model evidence: The optimal balance of fit and complexity

Comparing models

- Which is the best model?



# Inference on model space

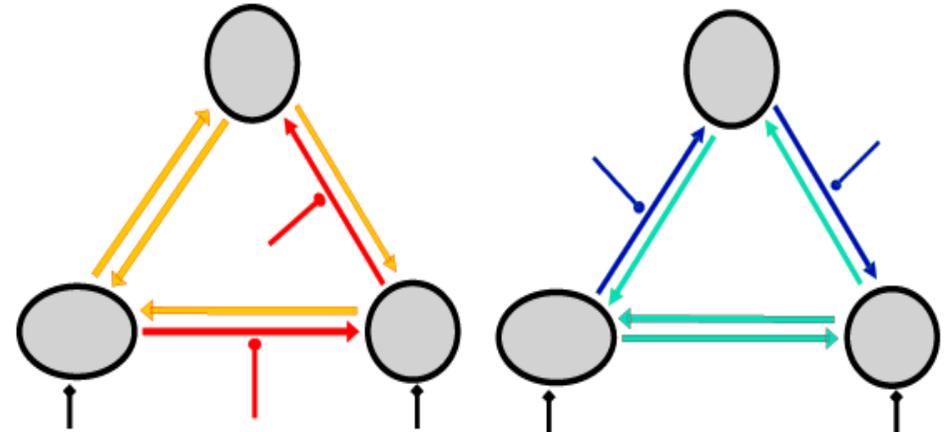
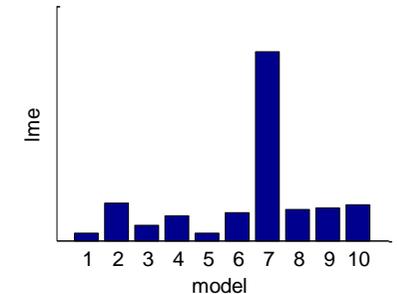
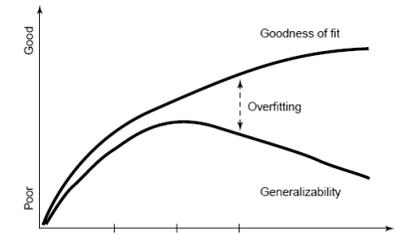
Model evidence: The optimal balance of fit and complexity

Comparing models

- Which is the best model?

Comparing families of models

- What type of model is best?
  - Feedforward vs feedback
  - Parallel vs sequential processing
  - With or without modulation



# Inference on model space

Model evidence: The optimal balance of fit and complexity

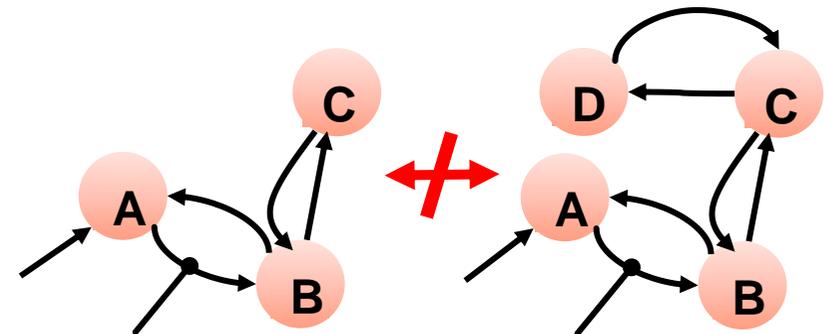
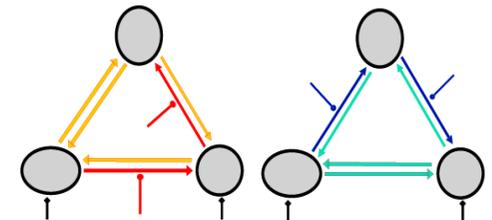
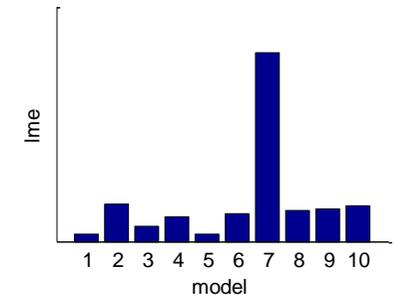
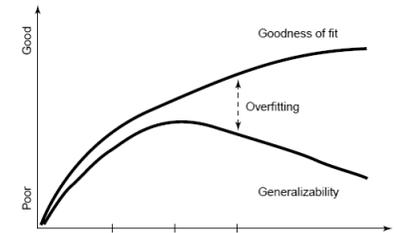
Comparing models

- Which is the best model?

Comparing families of models

- What type of model is best?
  - Feedforward vs feedback
  - Parallel vs sequential processing
  - With or without modulation

Only compare models with the same data



# To recap... so, DCM....

- enables one to **infer hidden neuronal processes** from fMRI data
- allows one to **test mechanistic hypotheses** about observed effects
  - uses a deterministic differential equation to model neuro-dynamics (represented by matrices A,B and C).
- is informed by anatomical and physiological principles.
- uses a **Bayesian framework** to estimate model parameters
- is a generic approach to modelling experimentally perturbed dynamic systems.
  - provides an observation model for neuroimaging data, e.g. fMRI, M/EEG
  - DCM is **not model or modality specific** (Models can change and the method extended to other modalities e.g. ERPs, LFPs)

# Overview

---

## Brain connectivity: types & definitions

*Anatomical connectivity*

*Functional connectivity*

*Effective connectivity*

## Dynamic causal models (DCMs)

*Neuronal model*

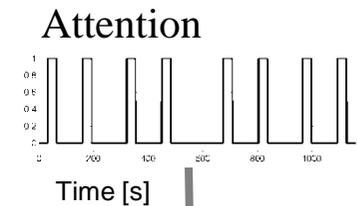
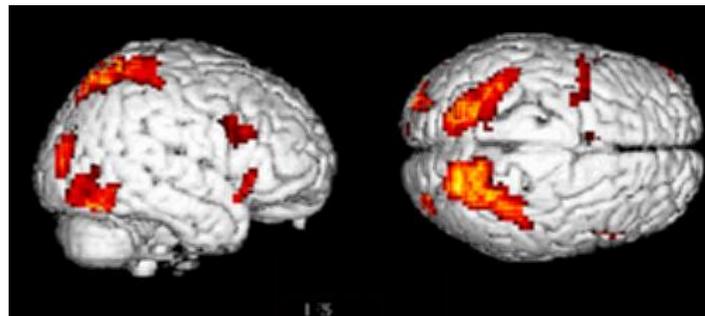
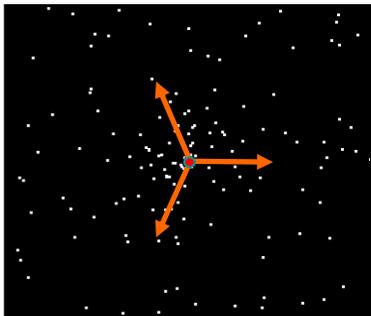
*Hemodynamic model*

*Estimation: Bayesian framework*

Applications & extensions of DCM to fMRI data

# Attention to motion in the visual system

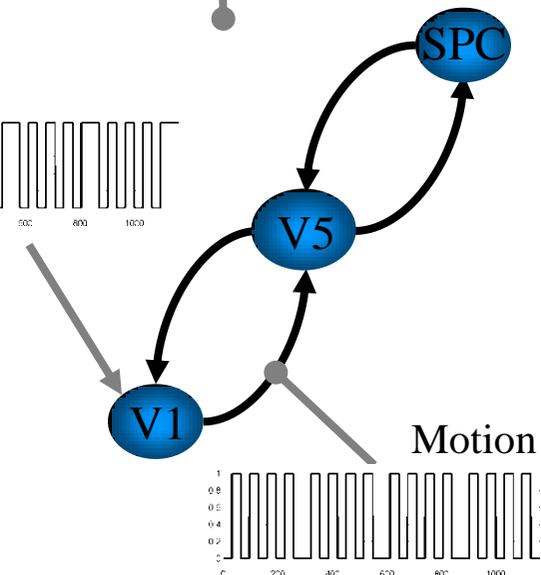
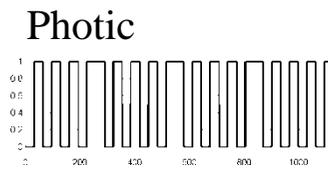
We used this model to assess the site of **attention modulation** during *visual motion processing* in an fMRI paradigm reported by *Büchel & Friston*.



- fixation only
- observe static dots
- observe moving dots
- task on moving dots

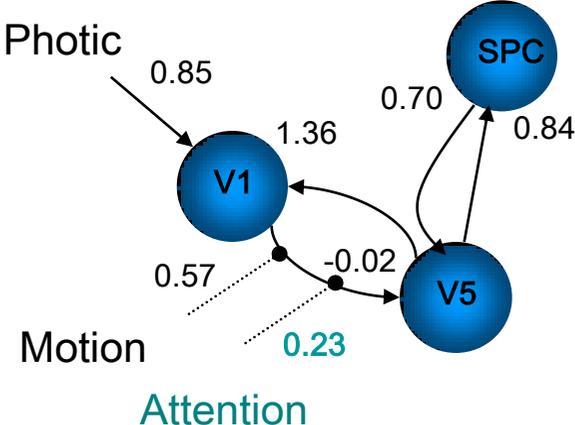
- + photic
- + motion
- + attention

- V1
- V5
- V5 + parietal cortex

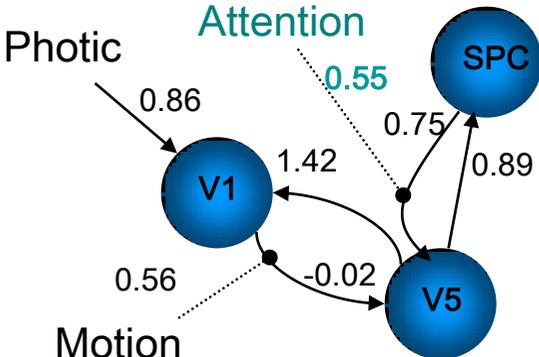


# Comparison of two simple models

**Model 1:**  
attentional modulation  
of V1→V5



**Model 2:**  
attentional modulation  
of SPC→V5



Bayesian model selection:

Model 1 better than model 2

$$\log p(y | m_1) \gg \log p(y | m_2)$$

→ Decision for model 1:

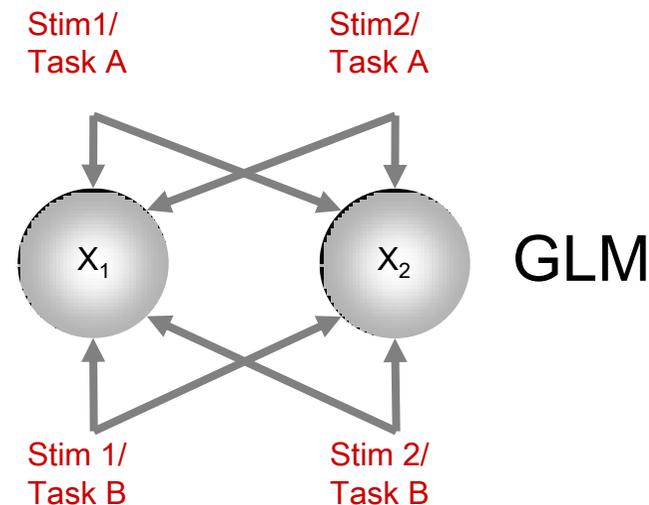
in this experiment, attention  
primarily modulates V1→V5

# Planning a DCM-compatible study

- Suitable experimental design:
  - any design that is suitable for a GLM
  - preferably multi-factorial (e.g. 2 x 2)
    - e.g. one factor that varies the driving (sensory) input
    - and one factor that varies the contextual input
- Hypothesis and model:
  - Define specific *a priori* hypothesis
  - Which parameters are relevant to test this hypothesis?
  - If you want to verify that intended model is suitable to test this hypothesis, then use simulations
  - Define criteria for inference
  - What are the alternative models to test?

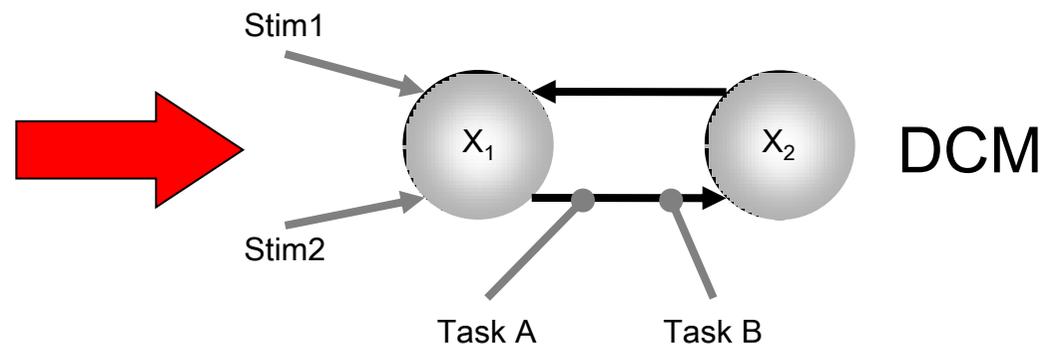
# Multifactorial design: explaining interactions with DCM

		Task factor	
		Task A	Task B
Stimulus factor	Stim 1	A1	B1
	Stim 2	A2	B2

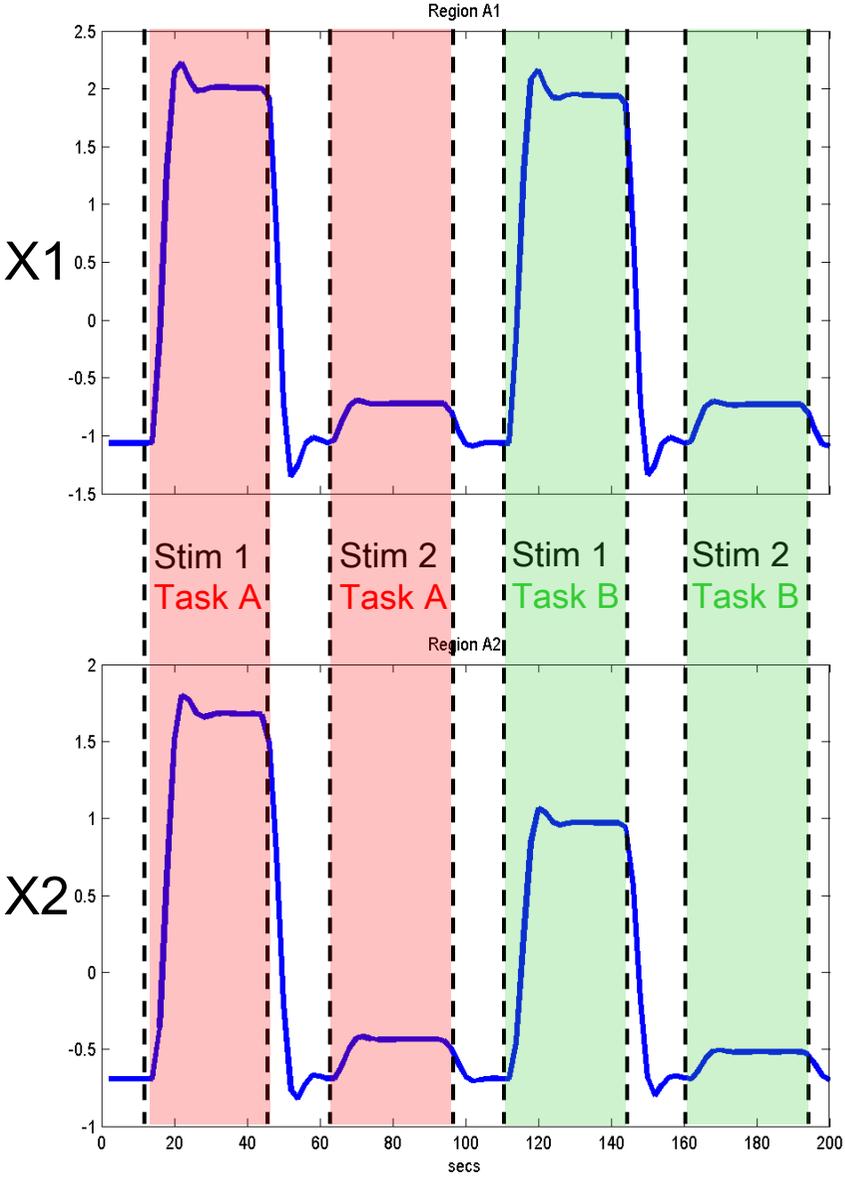
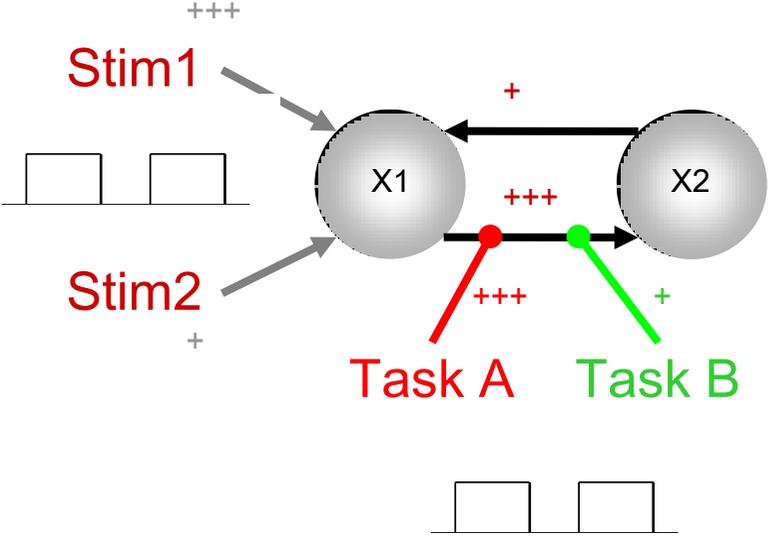


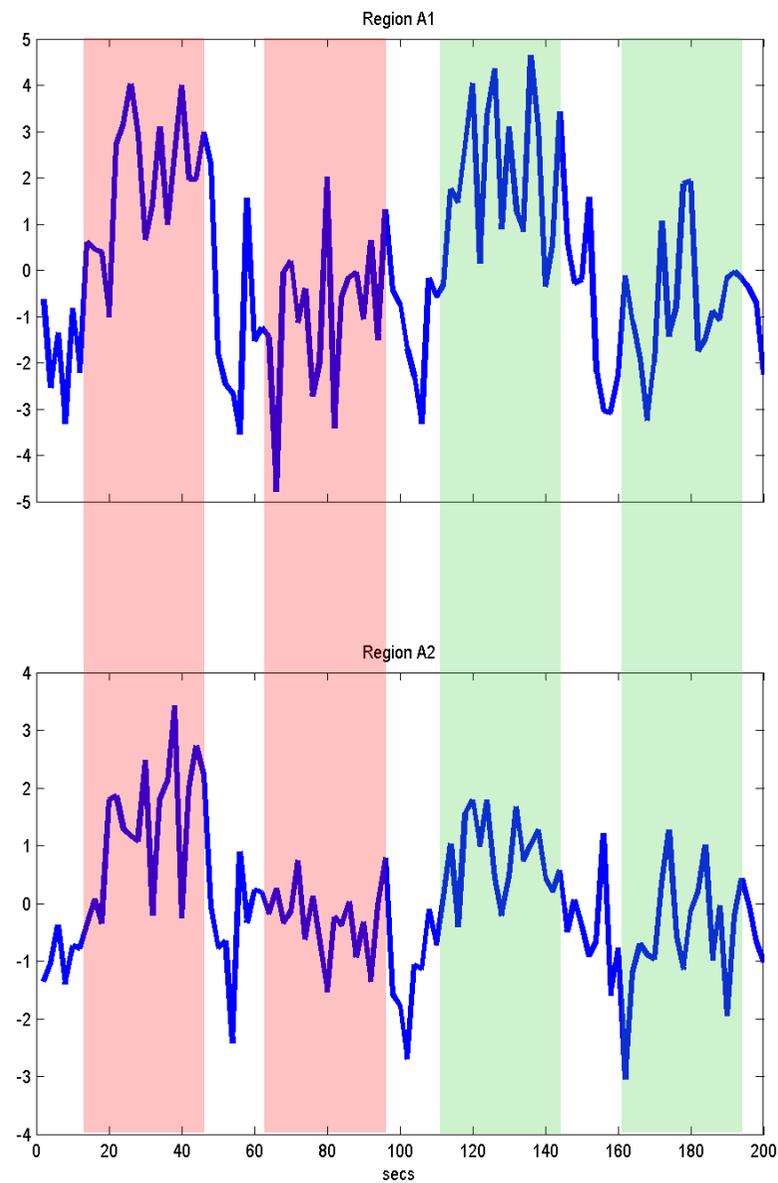
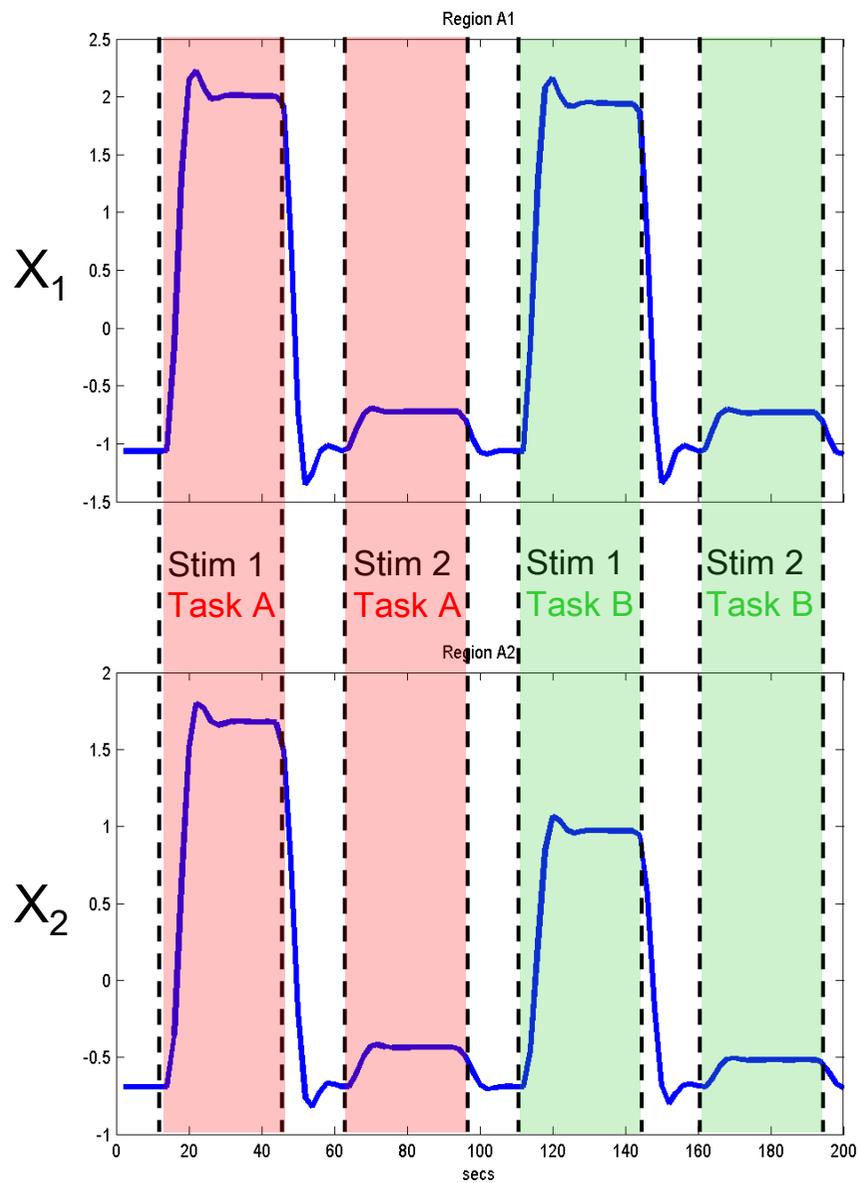
Let's assume that an SPM analysis shows a main effect of stimulus in  $X_1$  and a stimulus  $\times$  task interaction in  $X_2$ .

How do we model this using DCM?



# Simulated data

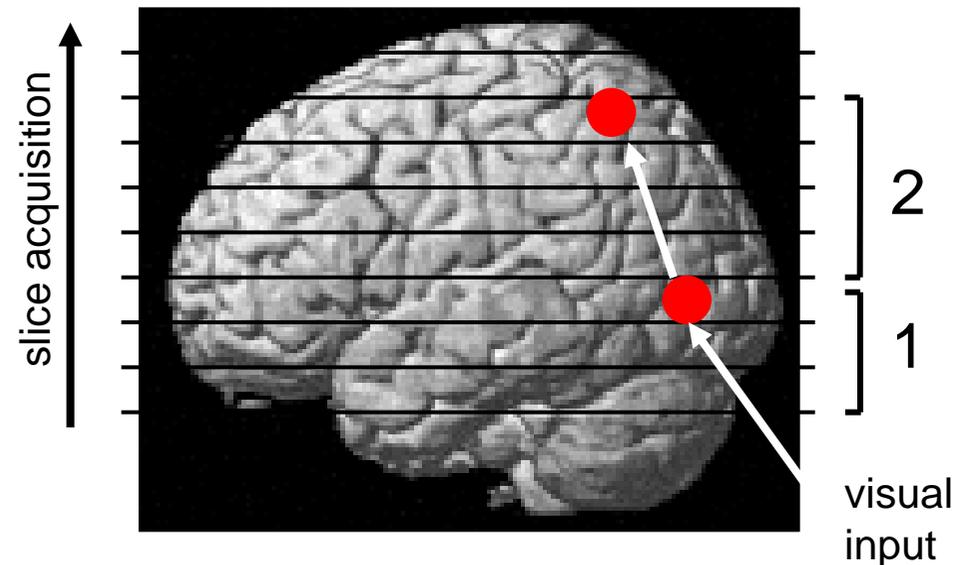




plus added noise (SNR=1)

# Slice timing model

- potential timing problem in DCM:  
temporal shift between regional time series because of multi-slice acquisition



- Solution:
  - Modelling of (known) slice timing of each area.

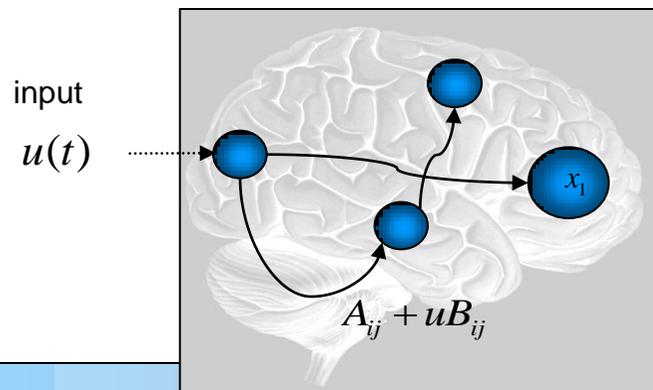
Slice timing extension now allows for any slice timing differences!

**Long TRs (> 2 sec) no longer a limitation.**

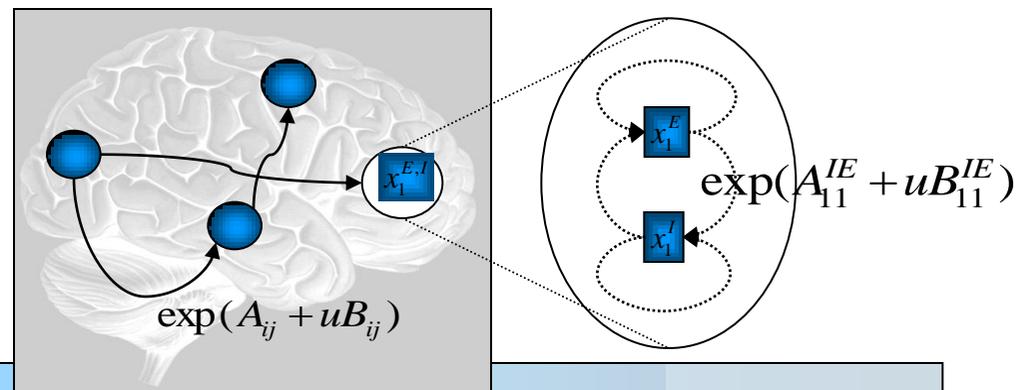
(Kiebel et al., 2007)

# Two-state model

Single-state DCM



Two-state DCM



$$\frac{\partial x}{\partial t} = (A + uB)x + Cu$$

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_{NN} \end{bmatrix} \quad x(t) = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

$$\dot{z} = Az + \sum u_j B_j z + Cu$$

$$\frac{\partial x}{\partial t} = (AB^u)x + Cu$$

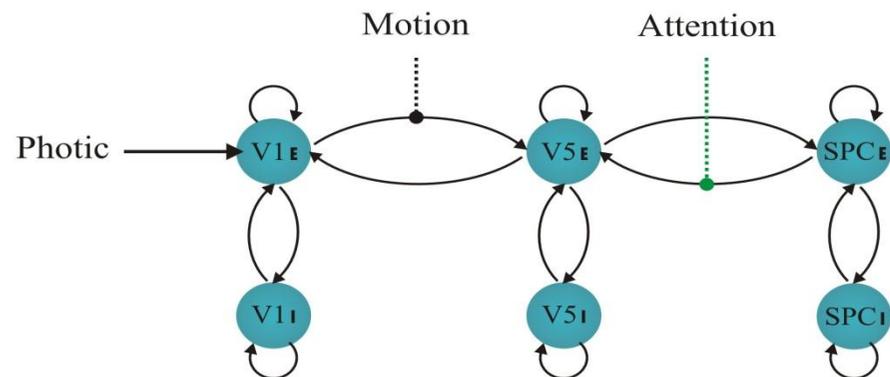
$$A = \begin{bmatrix} -e^{A_{11}^{EE}} & -e^{A_{11}^{EI}} & \cdots & e^{A_{1N}} & 0 \\ e^{A_{11}^{IE}} & -e^{A_{11}^{II}} & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ e^{A_{N1}} & 0 & & -e^{A_{NN}^{EE}} & -e^{A_{NN}^{EI}} \\ 0 & 0 & \cdots & e^{A_{NN}^{IE}} & -e^{A_{NN}^{II}} \end{bmatrix} \quad x(t) = \begin{bmatrix} x_1^E \\ x_1^I \\ \vdots \\ x_N^E \\ x_N^I \end{bmatrix}$$

Extrinsic (between-region) coupling

Intrinsic (within-region) coupling

# Example: Two-state Model Comparison

## Model 1 - BCW

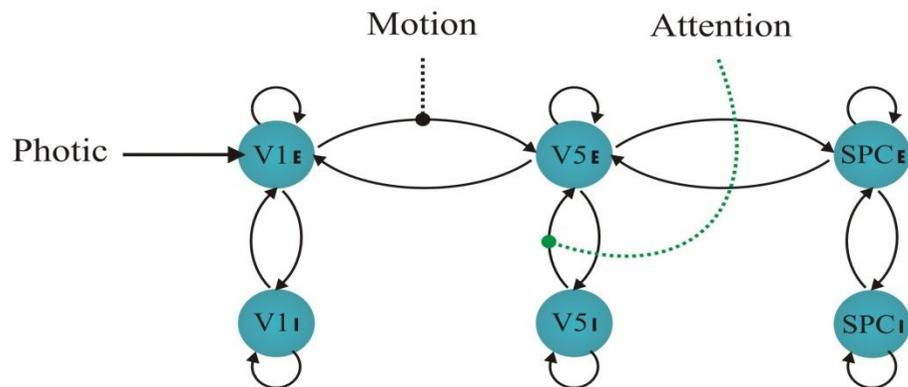


## DCM for *Büchel & Friston*

$$A = \begin{bmatrix} -EE_{V1} & -EI_{V1} & E_{V1}E_{V5} & 0 & 0 & 0 \\ IE_{V1} & -II_{V1} & 0 & 0 & 0 & 0 \\ E_{V5}E_{V1} & 0 & -EE_{V5} & -EI_{V5} & E_{V5}E_{SPC} & 0 \\ 0 & 0 & IE_{V5} & -II_{V5} & 0 & 0 \\ 0 & 0 & E_{SPC}E_{V5} & 0 & -EE_{SPC} & -EI_{SPC} \\ 0 & 0 & 0 & 0 & IE_{SPC} & -II_{SPC} \end{bmatrix}$$

Note: In the matrix above, the terms  $E_{V5}E_{SPC}$  and  $-EI_{V5}$  are highlighted with a green dashed box, and a green dotted line labeled 'Attention' points to the  $-EI_{V5}$  term.

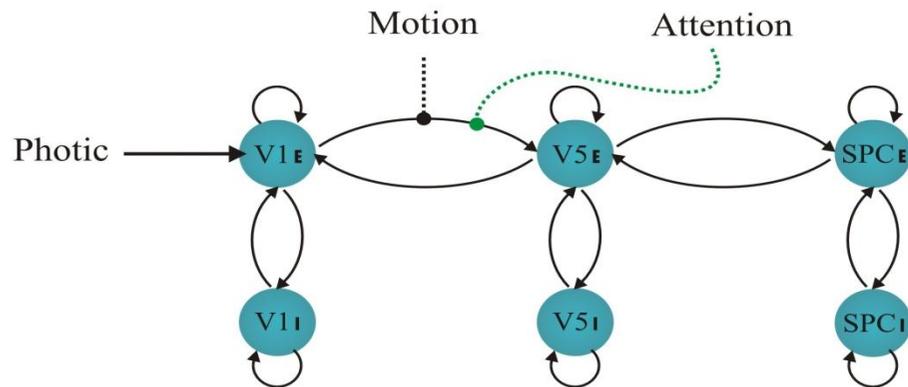
## Model 2 - Intr



$$A = \begin{bmatrix} -EE_{V1} & -EI_{V1} & E_{V1}E_{V5} & 0 & 0 & 0 \\ IE_{V1} & -II_{V1} & 0 & 0 & 0 & 0 \\ E_{V5}E_{V1} & 0 & -EE_{V5} & -EI_{V5} & E_{V5}E_{SPC} & 0 \\ 0 & 0 & IE_{V5} & -II_{V5} & 0 & 0 \\ 0 & 0 & E_{SPC}E_{V5} & 0 & -EE_{SPC} & -EI_{SPC} \\ 0 & 0 & 0 & 0 & IE_{SPC} & -II_{SPC} \end{bmatrix}$$

Note: In the matrix above, the terms  $-EI_{V5}$  and  $E_{V5}E_{SPC}$  are highlighted with a green dashed box, and a green dotted line labeled 'Attention' points to the  $-EI_{V5}$  term.

## Model 3 - FWD

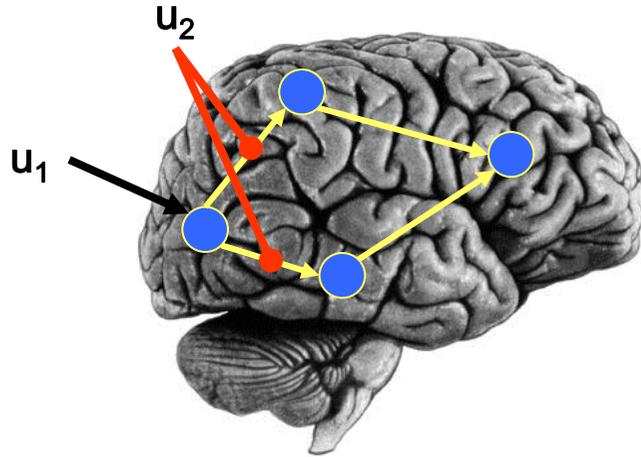


$$A = \begin{bmatrix} -EE_{V1} & -EI_{V1} & E_{V1}E_{V5} & 0 & 0 & 0 \\ IE_{V1} & -II_{V1} & 0 & 0 & 0 & 0 \\ E_{V5}E_{V1} & 0 & -EE_{V5} & -EI_{V5} & E_{V5}E_{SPC} & 0 \\ 0 & 0 & IE_{V5} & -II_{V5} & 0 & 0 \\ 0 & 0 & E_{SPC}E_{V5} & 0 & -EE_{SPC} & -EI_{SPC} \\ 0 & 0 & 0 & 0 & IE_{SPC} & -II_{SPC} \end{bmatrix}$$

Note: In the matrix above, the terms  $E_{V5}E_{V1}$  and  $-EI_{V5}$  are highlighted with a green dashed box, and a green dotted line labeled 'Attention' points to the  $-EI_{V5}$  term.

# Nonlinear DCM for fMRI

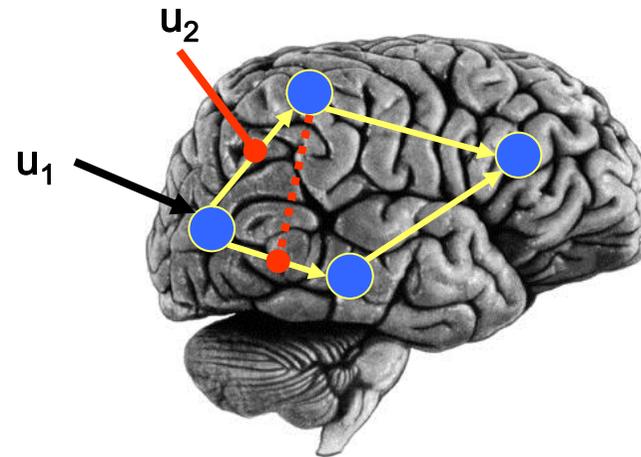
bilinear DCM



Bilinear state equation

$$\frac{dx}{dt} = \left( A + \sum_{i=1}^m u_i B^{(i)} \right) x + Cu$$

nonlinear DCM



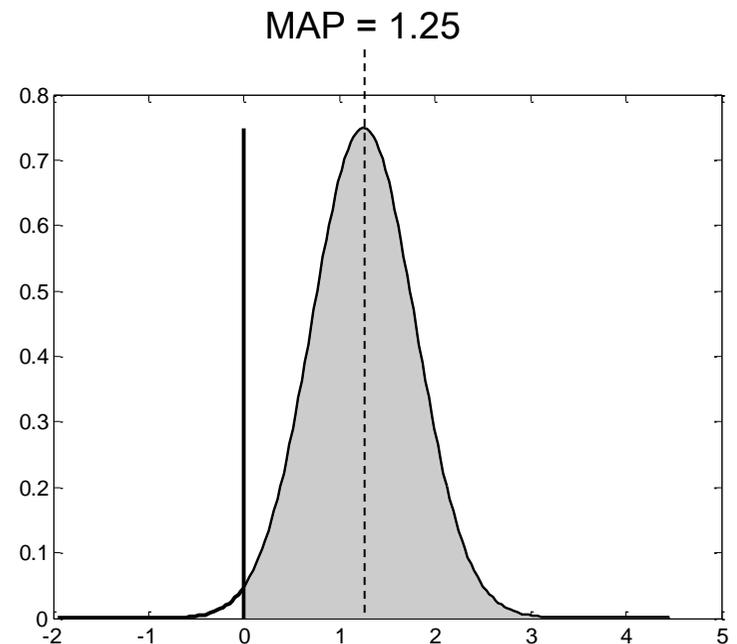
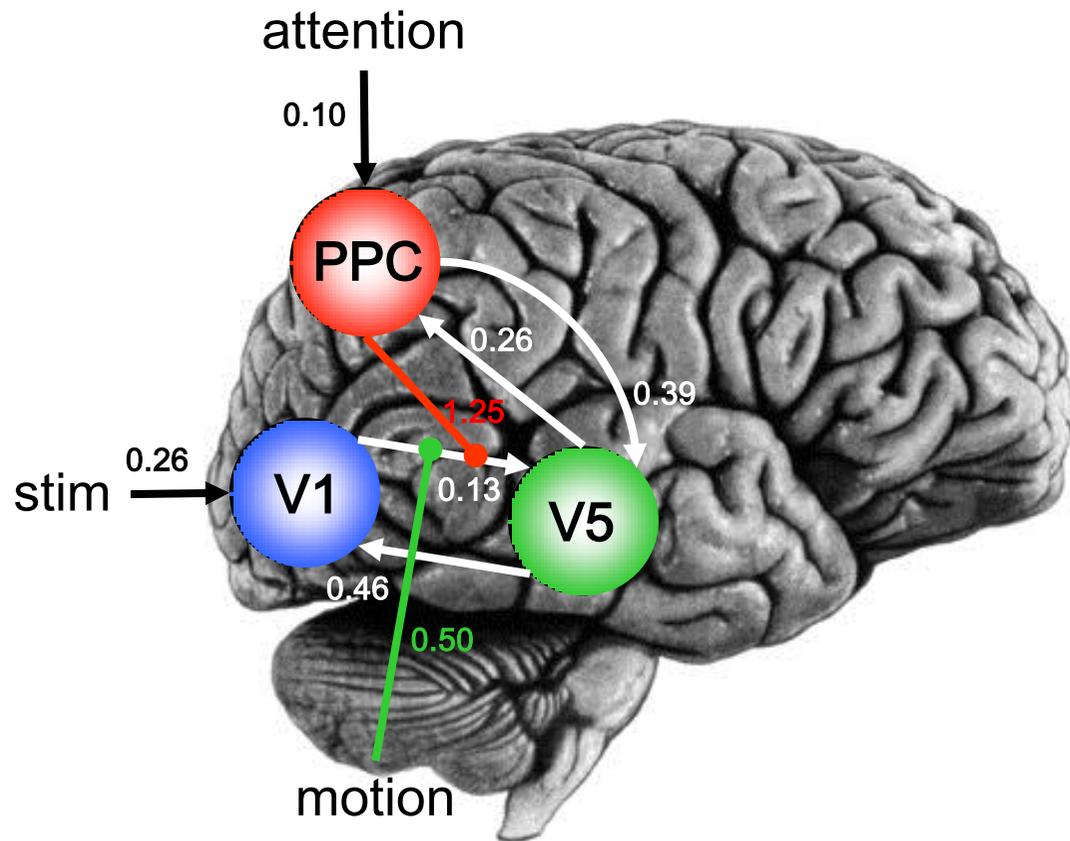
Nonlinear state equation

$$\frac{dx}{dt} = \left( A + \sum_{i=1}^m u_i B^{(i)} + \sum_{j=1}^n x_j D^{(j)} \right) x + Cu$$

Here DCM can model activity-dependent changes in connectivity; how connections are enabled or gated by activity in one or more areas.

# Nonlinear DCM for fMRI

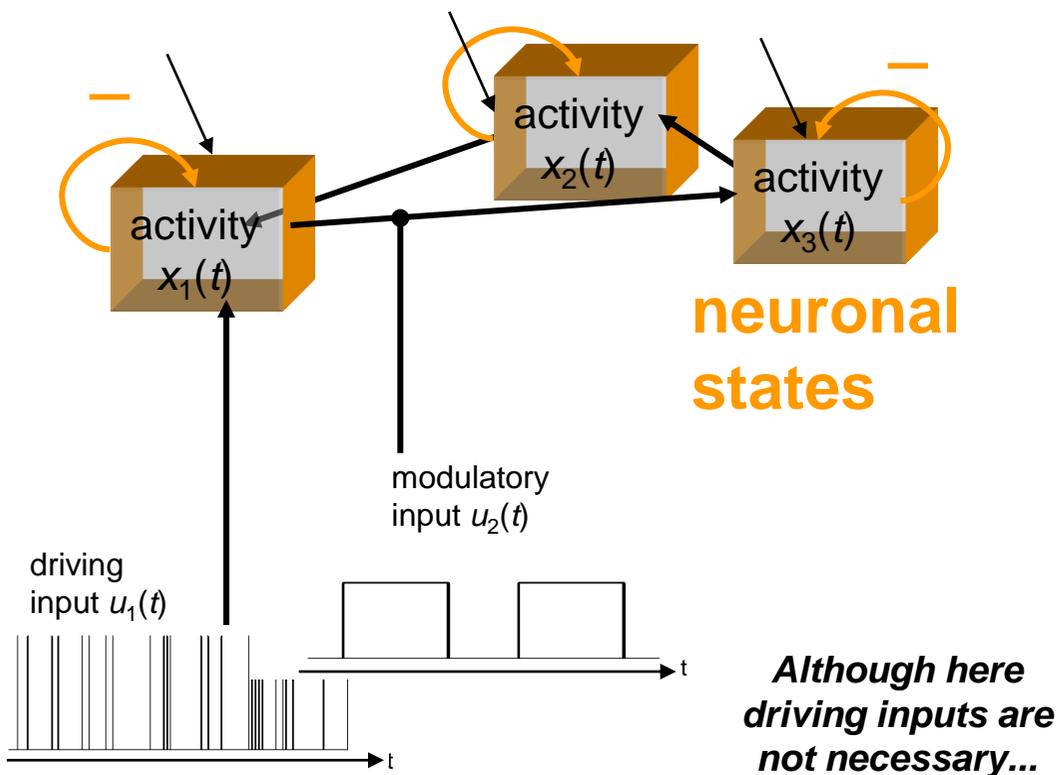
Can V5 activity during attention to motion be explained by allowing activity in PPC to modulate the V1-to-V5 connection?



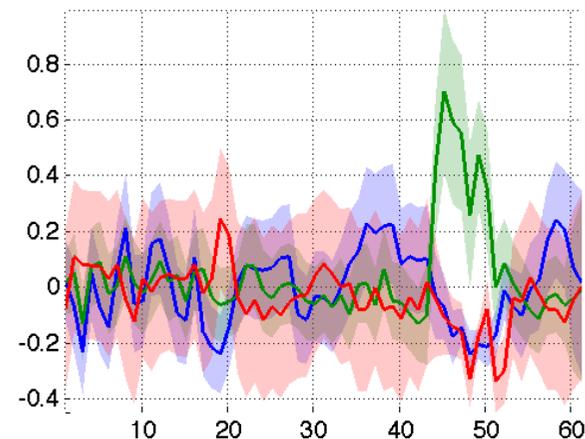
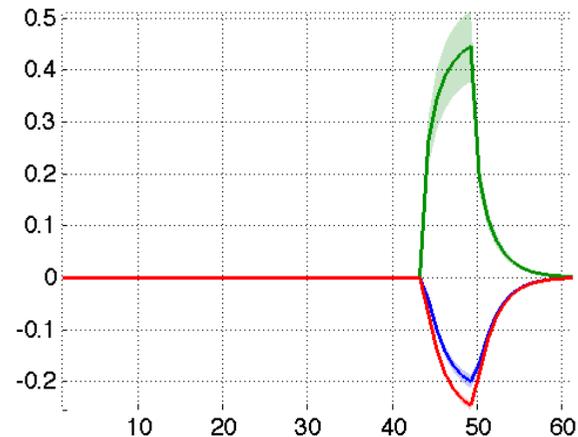
$$p(D_{V5,V1}^{PPC} > 0 | y) = 99.1\%$$

# Stochastic DCMs

Stochastic innovations: variance hyperparameter



**Inversion:** Generalised filtering (under the Laplace assumption)



Daunizeau et al, 2009  
Friston et al, 2008

# Bayesian Model Selection for large model spaces

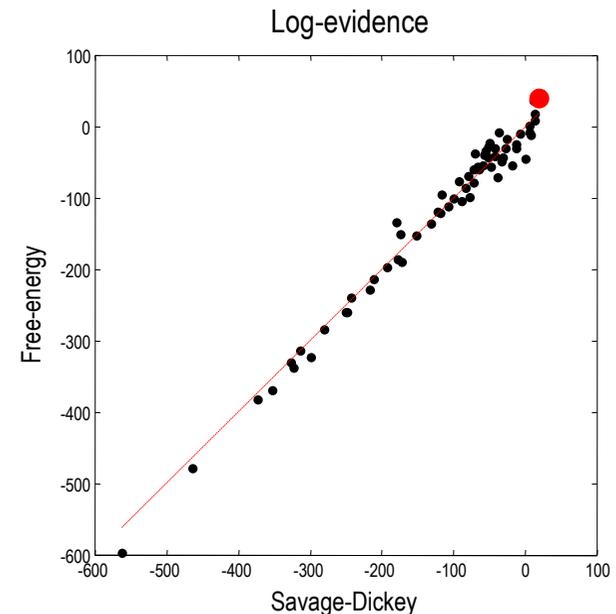
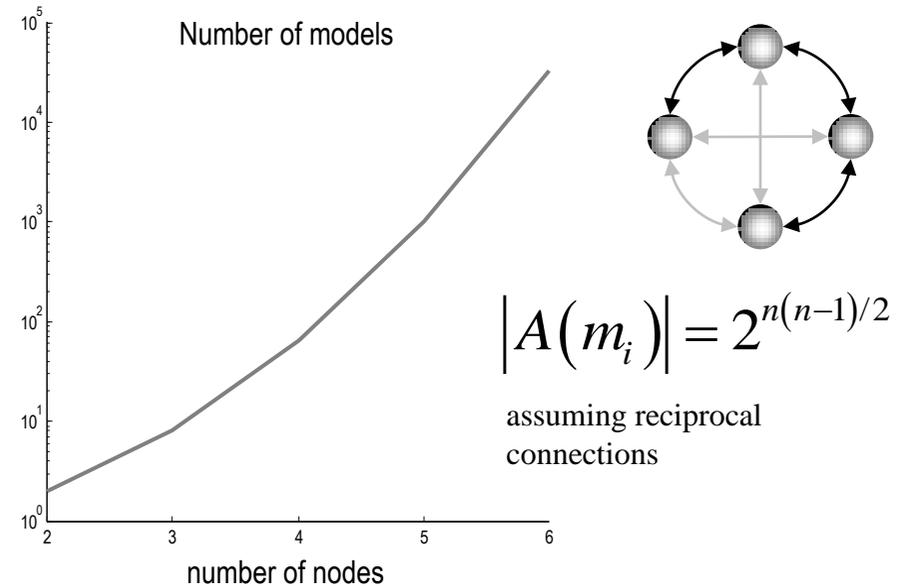
- for less constrained model spaces, search methods are needed
- fast model scoring via the Savage-Dickey density ratio:

$$\ln p(y | m_i)$$

$$\approx \ln q(\theta_i = 0 | m_i) - \ln p(\theta_i = 0 | m_F)$$

**Life easier for more  
exploratory approaches!**

Friston et al. 2011, *NeuroImage*  
Friston & Penny 2011, *NeuroImage*



# The evolution of DCM in SPM

- DCM is not one specific model, but a framework for Bayesian inversion of dynamical systems
- The default implementation in SPM is evolving over time:
  - better numerical routines for inversion
  - changes in prior to cover new variants

To enable replication of your results, you should state which SPM version you are using!

# Some introductory references

- **The first DCM paper:** Dynamic Causal Modelling (2003). Friston et al. *NeuroImage* 19:1273-1302.
- **Physiological validation of DCM for fMRI:** Identifying neural drivers with functional MRI: an electrophysiological validation (2008). David et al. *PLoS Biol.* 6 2683–2697
- **Hemodynamic model:** Comparing hemodynamic models with DCM (2007). Stephan et al. *NeuroImage* 38:387-401
- **Nonlinear DCMs:** Nonlinear Dynamic Causal Models for FMRI (2008). Stephan et al. *NeuroImage* 42:649-662
- **Two-state model:** Dynamic causal modelling for fMRI: A two-state model (2008). Marreiros et al. *NeuroImage* 39:269-278
- **Group Bayesian model comparison:** Bayesian model selection for group studies (2009). Stephan et al. *NeuroImage* 46:1004-10174
- **10 Simple Rules for DCM** (2010). Stephan et al. *NeuroImage* 52.

***Thank you for your attention!!!***