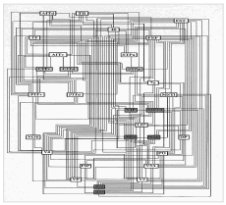
DCM: Dynamic Causal Modelling for fMRI



Mohamed Seghier

Wellcome Trust Centre for Neuroimaging,
University College London, UK

Wellcome Trust Centre for Neuroimaging wellcome trust

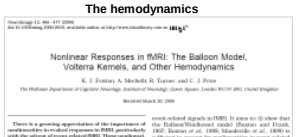
DCM is a <u>generative model</u>
= a quantitative / mechanistic description of how observed data are generated.

<u>Key features:</u>
1- Dynamic
2- Causal
3- Neuro-physiologically motivated
4- Operate at hidden neuronal interactions
5- Bayesian in all aspects.

<u>DCM [default] implementation:</u>

Deterministic	**Stochastic [Daunizeau et al. 2009]**
Bilinear	**Nonlinear [Stephan et al. 2008]**
The one-state neuronal	**The two-state [Marreiros et al. 2008]**

The hemodynamics



Nonlinear Responses in fMRI: The Balloon Model, Volterra Kernels, and Other Hemodynamics

K. J. Friston, A. Mechelli, R. Turner, and C. J. Price

[Friston et al. 2000 *Neuroimage*]

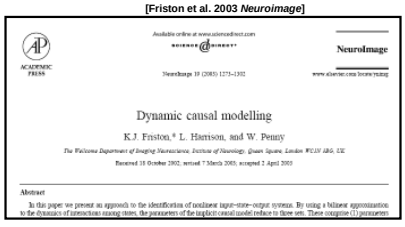
Deterministic dynamical systems

Bayesian Estimation of Dynamical Systems: An Application to fMRI

K. J. Friston

[Friston 2002 *Neuroimage*]

[Friston et al. 2003 *Neuroimage*]



Available online at www.sciencedirect.com

NeuroImage

Dynamic causal modelling

K.J. Friston,* L. Harrison, and W. Penny

Abstract

"The central idea behind dynamic causal modelling (DCM) is to treat the brain as a deterministic nonlinear dynamic system that is subject to inputs and produces outputs."

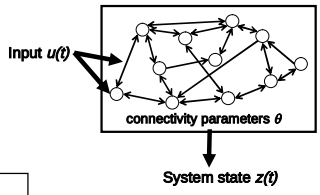
"DCM assumes the responses are driven by designed changes in inputs."

"DCM is used to test the specific hypothesis that motivated the experimental design. It is not an exploratory technique […]; the results are specific to the tasks and stimuli employed during the experiment."

[Friston et al. 2003 *Neuroimage*]

What is a system?

System =
a set of elements which interact in a spatially and temporally specific fashion



State changes of a system are dependent on:
- the current state
- external inputs
- its connectivity
- time constants & delays

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

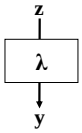
(evolution equation)

Basic idea of DCM for fMRI

♣ Effective connectivity is parameterised in terms of coupling among unobserved brain states (e.g., neuronal activity in different regions). The objective is to estimate these parameters by perturbing the system and measuring the response.

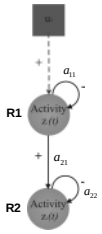
♣ A cognitive system is modelled as a bilinear model of neural population dynamics (z).

♣ The modelled neuronal dynamics (z) is transformed into area-specific BOLD signals (y) by a hemodynamic forward model (λ).

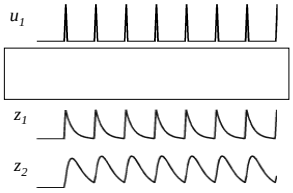


Aim: to estimate the parameters of a reasonably realistic neural model such that the predicted regional blood oxygen level dependent (BOLD) signals, correspond as closely as possible to the observed BOLD signals.

Neurodynamics: 2 nodes with input



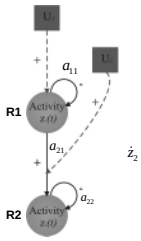
$$\dot{z}_1 = a_{11} z_1 + c u_1$$
$$\dot{z}_2 = a_{22} z_2 + a_{21} z_1$$



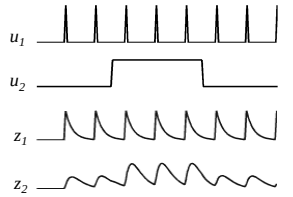
$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ (a_{21}) & a_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} c \\ 0 \end{bmatrix} u_1 \qquad a_{21} > 0$$

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

Neurodynamics: positive modulation



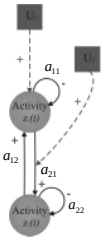
$$\dot{z}_1 = a_{11} z_1 + c u_1$$
$$\dot{z}_2 = a_{22} z_2 + a_{21} z_1 + b_{21}^2 u_2 z_1$$

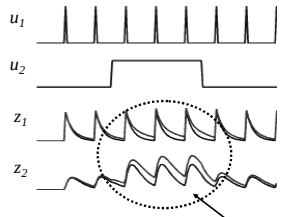


$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + u_2 \begin{bmatrix} 0 & 0 \\ (b_{21}^2) & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} c \\ 0 \end{bmatrix} u_1 \qquad b_{21}^2 > 0$$

modulatory input u_2 activity through the coupling a_{21}

Neurodynamics: reciprocal connections

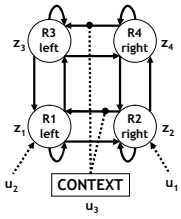




$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & (a_{12}) \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + u_2 \begin{bmatrix} 0 & 0 \\ b_{21}^2 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} c \\ 0 \end{bmatrix} u_1 \qquad a_{21} > 0 \quad a_{12} > 0 \quad b_{21}^2 > 0$$

reciprocal connection disclosed by u_2

bilinear dynamic system



$$\begin{bmatrix}\dot{z}_1\\ \dot{z}_2\\ \dot{z}_3\\ \dot{z}_4\end{bmatrix} = \left\{\begin{bmatrix}a_{11} & a_{12} & a_{13} & 0\\ a_{21} & a_{22} & 0 & a_{24}\\ a_{31} & 0 & a_{33} & a_{34}\\ 0 & a_{42} & a_{43} & a_{44}\end{bmatrix} + u_3\begin{bmatrix}0 & b_{12}^3 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & b_{34}^3\\ 0 & 0 & 0 & 0\end{bmatrix}\right\}\begin{bmatrix}z_1\\ z_2\\ z_3\\ z_4\end{bmatrix} + \begin{bmatrix}0 & c_{12} & 0\\ c_{21} & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}u_1\\ u_2\\ u_3\end{bmatrix}$$

Bilinear state equation in DCM for fMRI

The neural state equation

state changes | connectivity | modulation of connectivity | state vector | direct inputs | external inputs

$$\begin{bmatrix}\dot{z}_1\\ \vdots\\ \dot{z}_n\end{bmatrix} = \left\{\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{bmatrix} + \sum_{j=1}^{m} u_j \begin{bmatrix}b_{11}^j & \cdots & b_{1n}^j\\ \vdots & \ddots & \vdots\\ b_{n1}^j & \cdots & b_{nn}^j\end{bmatrix}\right\}\begin{bmatrix}z_1\\ \vdots\\ z_n\end{bmatrix} + \begin{bmatrix}c_{11} & \cdots & c_{1m}\\ \vdots & \ddots & \vdots\\ c_{n1} & \cdots & c_{nm}\end{bmatrix}\begin{bmatrix}u_1\\ \vdots\\ u_m\end{bmatrix}$$

n regions | *m* inputs (mod.) | *m* inputs (driv.)

$$\dot{z} = (A + \sum_{j=1}^{m} u_j B^j)z + Cu$$

$$\dot{z} = (A + \sum_{j=1}^{m} u_j B^j)z + Cu$$

"C", the direct or driving effects:
- extrinsic influences of inputs on neuronal activity.

"A", the intrinsic coupling or the latent connectivity:
- fixed or endogenous effective connectivity;
- first order connectivity among the regions in the absence of input.

"B", the bilinear term, modulatory effects, or the induced connectivity:
- context-dependent change in connectivity;
- eq. a second-order interaction between the input and activity in a source region when causing a response in a target region.

[Units]: rates, [Hz];
Strong connection = an effect that is influenced quickly or with a small time constant.

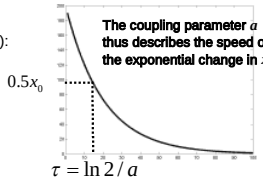
DCM parameters = rate constants

Integration of a first-order linear differential equation gives an exponential function:

$$\frac{dx}{dt} = ax \quad \Rightarrow \quad x(t) = x_0 \exp(at)$$

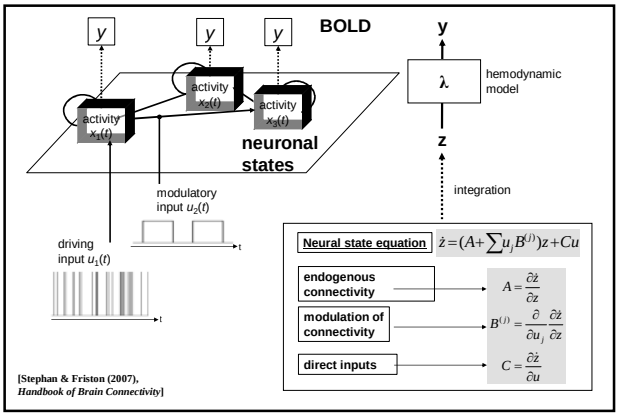
Coupling parameter a is inversely proportional to the half life τ of x(t):

$$x(\tau) = 0.5x_0$$
$$= x_0 \exp(a\tau)$$
$$\Rightarrow \boxed{a = \ln 2 / \tau}$$

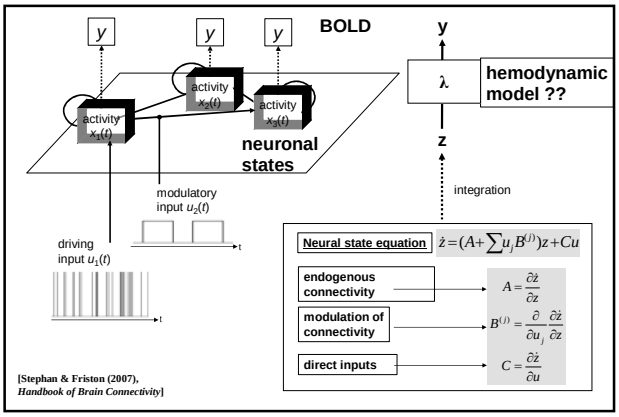


The coupling parameter a thus describes the speed of the exponential change in $x(t)$

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

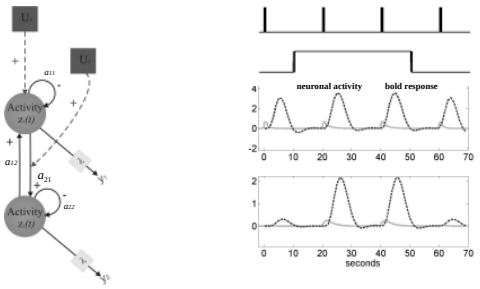


[Stephan & Friston (2007), *Handbook of Brain Connectivity*]

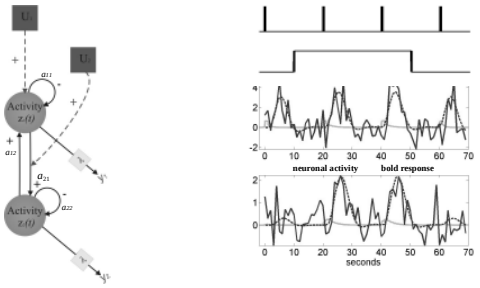


[Stephan & Friston (2007), *Handbook of Brain Connectivity*]

Hemodynamics: the indirect link

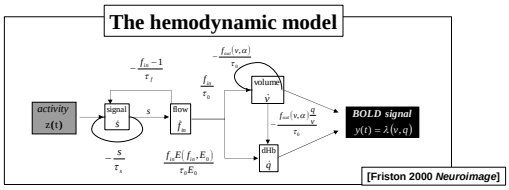


$$\begin{bmatrix}\dot{z}_1\\ \dot{z}_2\end{bmatrix}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}z_1\\ z_2\end{bmatrix}+u_2\begin{bmatrix}0 & 0\\ b_{21}^2 & 0\end{bmatrix}\begin{bmatrix}z_1\\ z_2\end{bmatrix}+\begin{bmatrix}c\\ 0\end{bmatrix}u_1 \qquad a_{21}>0 \qquad a_{12}>0 \qquad b_{21}^2>0$$



$$\begin{bmatrix}\dot{z}_1\\ \dot{z}_2\end{bmatrix}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}z_1\\ z_2\end{bmatrix}+u_2\begin{bmatrix}0 & 0\\ b_{21}^2 & 0\end{bmatrix}\begin{bmatrix}z_1\\ z_2\end{bmatrix}+\begin{bmatrix}c\\ 0\end{bmatrix}u_1 \qquad a_{21}>0 \qquad a_{12}>0 \qquad b_{21}^2>0$$

The hemodynamic model



[Friston 2000 *Neuroimage*]

State Equations

Flow component:
s : activity-dependent signal;
f : flow inducing signal

Balloon component:
v : the rate of change of volume;
q : the change in deoxyhemoglobin

Output function: a mixture of intra- and extra-vascular signal

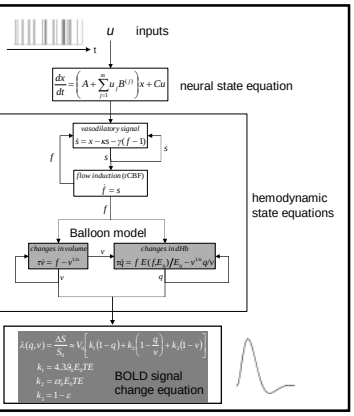
The hemodynamic model

- 6 hemodynamic parameters:

$$\theta^h = \{\kappa, \gamma, \tau, \alpha, \rho, \varepsilon\}$$

important for model fitting, but of no interest for statistical inference

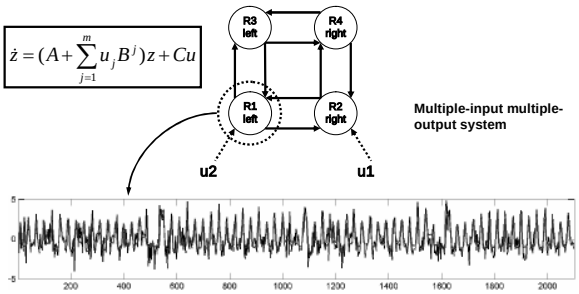
- Empirically determined *a priori* distributions.
- Area-specific estimates (like neural parameters) → **region-specific** HRFs!



[Friston et al. 2000, *NeuroImage*]
[Stephan et al. 2007, *NeuroImage*]

Example: modelled BOLD signal

$$\dot{z} = (A + \sum_{j=1}^{m} u_j B^j) z + Cu$$



black: observed BOLD signal
red: modelled BOLD signal

Priors & parameter estimation

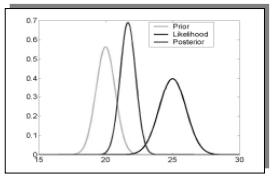
Bayesian statistics (inversion)

Bayes theorem allows us to express our **prior knowledge** or "belief" about parameters of the model.

new data $p(y|\theta)$ prior knowledge $p(\theta)$

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

posterior ∝ likelihood • prior



The **posterior** probability of the parameters given the data is an optimal combination of prior knowledge and new data, weighted by their relative precision.

Priors in DCM

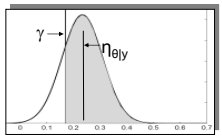
Constraints on parameter estimation:

- hemodynamic parameters: **empirical priors**
- coupling parameters of self-connections: **principled priors**
- coupling parameters other connections: **shrinkage priors**

Inference about DCM parameters: Bayesian inversion

- Gaussian assumptions about the posterior distributions of the parameters
- Use of the cumulative normal distribution to test the probability that a certain parameter (or contrast of parameters $c^T \eta_{\theta|y}$) is above a chosen threshold γ:

$$p = \phi_N\left(\frac{c^T\eta_{\theta|y} - \gamma}{\sqrt{c^T C_{\theta|y} c}}\right)$$



- By default, γ is chosen as zero ("does the effect exist?").

Bayesian parameter estimation by means of expectation-maximisation (EM)

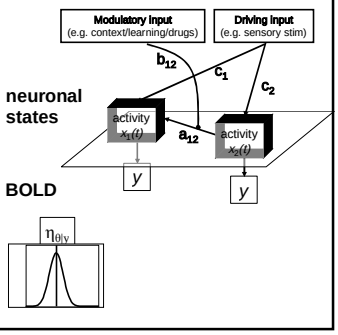
[Friston 2002 *Neuroimage*]

DCM: practical steps

Select areas you want to model

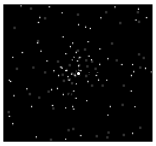
- Extract timeseries of these areas (x(t))
- Specify at neuronal level
  - what drives areas (c)
  - how areas interact (a)
  - what modulates interactions (b)
- State-space model with 2 levels:
  - Hidden neural dynamics
  - Predicted BOLD response
- Estimate model parameters:

  Gaussian a posteriori parameter distributions, characterised by mean $\eta_{\theta|y}$ and covariance $C_{\theta|y}$.



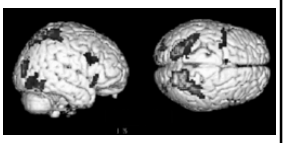
Attention to motion in the visual system

Stimuli 250 radially moving dots at 4.7 degrees/s



Pre-Scanning
5 x 30s trials with 5 speed changes (reducing to 1%)
Task - detect change in radial velocity

Scanning (no speed changes)
6 normal subjects, 4 x 100 scan sessions;
each session comprising 10 scans of 4 different conditions

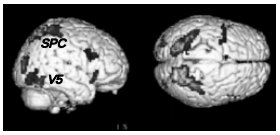


F A F N F A F N S .....

F - fixation point only
A - motion stimuli with attention (detect changes)
N - motion stimuli without attention
S - no motion

[Büchel & Friston 1997, *Cereb. Cortex*]
[Büchel et al. 1998, *Brain*]

How we can interpret, mechanistically, the increase in activity of area V5 by attention when motion is physically unchanged.



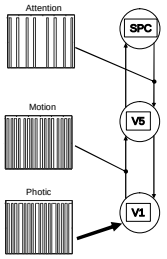
Choice of areas and time series extraction.
➔ Three ROIs: V1, V5, and SPC.

Definition of driving inputs.
➔ All visual stimuli/conditions (photic: A N S)

Definition of modulatory inputs.
➔ The effects of motion and attention (A N)

Building the model:
1- how to connect regions (intrinsic connections "A");
2- how the driving inputs enter the system (extrinsic effects "C");
3- define the context-dependent connections (modulatory effects "B").

- Visual inputs drive V1.
- Activity then spreads to hierarchically arranged visual areas.
- Motion modulates the strength of the V1→V5 forward connection.
- Attention modualtes the strength of the SPC→V5 backward connection.

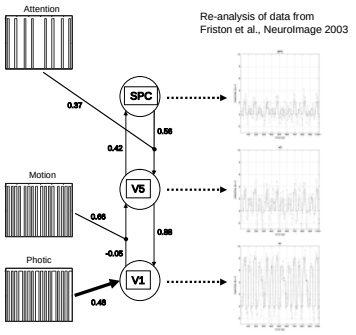


Re-analysis of data from
[Friston et al., 2003 *NeuroImage*]

After DCM estimation:

- Motion modulates the strength of the V1→V5 forward connection.
- The intrinsic connection V1→V5 is insignificant in the absence of motion (a_{21}=-0.05 Hz).
- Attention increases the backward-connection SPC→V5.

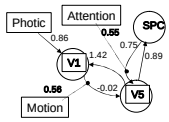
Are there other plausible/alternative models?



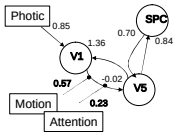
Re-analysis of data from Friston et al., NeuroImage 2003

Alternative models (hypothesis-driven approach):

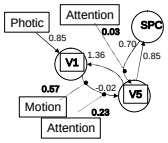
Model 1:
attentional modulation of V1→V5



Model 2:
attentional modulation of SPC→V5



Model 3:
attentional modulation of V1→V5 and SPC→V5



How we can compare between competing hypotheses?
➔ BMS (Bayesian Model Selection)

Model evidence and selection

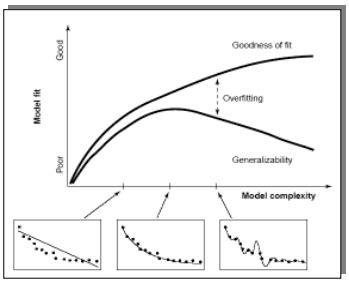
Given competing hypotheses on functional mechanisms of a system, which model is the best?

⇩

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

⇩

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



[Pitt and Miyung 2002 *TICS*]

Bayesian model selection (BMS)

Bayes' rule: $p(\theta \mid y,m) = \frac{p(y \mid \theta,m)\,p(\theta \mid m)}{p(y \mid m)}$

Model evidence: $p(y \mid m) = \int p(y \mid \theta,m) \cdot p(\theta \mid m)\, d\theta$

Model evidence: probability of generating data y from parameters θ that are randomly sampled from the prior $p(m)$.

Maximum likelihood: probability of the data y for the specific parameter vector θ that maximises $p(y|\theta,m)$.

⇒ accounts for both accuracy and complexity of the model

⇒ allows for inference about structure (generalisability) of the model

⇒ integral usually not analytically solvable, approximations necessary

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

$$\log p(y \mid m) = accuracy\,(m) - complexity\,(m)$$
$$= \log p(y \mid \theta,m) - complexity\,(m)$$

The negative variotional free energy (F) approximation [Penny et al. 2004, *NeuroImage*] [Penny et al. 2010, *PLoS Comp Biol*]

Under Gaussian assumptions about the posterior (Laplace approximation), the negative free energy *F* is a lower bound on the log model evidence:

$$\Rightarrow F = \log p(y \mid m) - KL\left[q(\theta), p(\theta \mid y,m)\right]$$

Kullback-Leibler (KL) divergence

The complexity term in *F*

- The negative free energy *F* accounts for parameter interdependencies.

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

[Penny et al. 2004 *Neuroimage*] [Stephan et al. 2009 *Neuroimage*]

- The complexity term of *F* is higher
  - the more independent the prior parameters (↑ effective DFs)
  - 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

To compare two models, we can just compare their log evidences.

But: the log evidence is just some number – not very intuitive!

A more intuitive interpretation of model comparisons is made possible by Bayes factors:

positive value, [0; ∞[

$$BF_{12} = \frac{p(y \mid m_1)}{p(y \mid m_2)}$$

Kass & Raftery classification:

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

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

Bayesian Model Selection in group studies.

Fixed effects BMS at group level

Group Bayes factor (GBF) for 1...*K* subjects:

$$GBF_{ij} = \prod_k BF_{ij}^{(k)} \qquad BF_{ij} = \frac{p(y \mid m_i)}{p(y \mid m_j)}$$

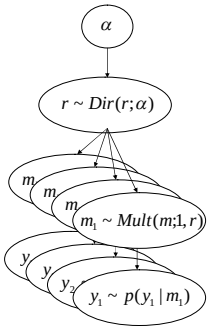
Average Bayes factor (ABF):

$$ABF_{ij} = \sqrt[K]{\prod_k BF_{ij}^{(k)}}$$

Problems:

- ► blind with regard to group heterogeneity;
- ► sensitive to outliers.

Random effects BMS for group studies



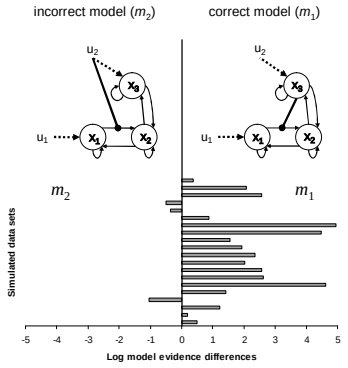
Dirichlet parameters
= "occurrences" of models in the populations

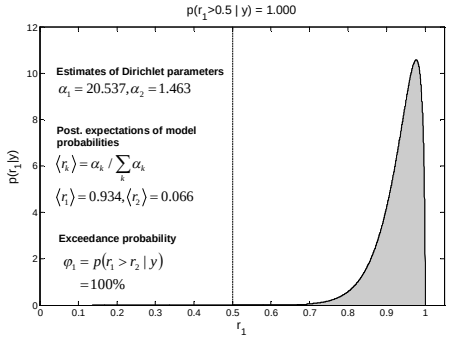
Dirichlet distribution of model probabilities

Multinomial distribution of subject-specific models

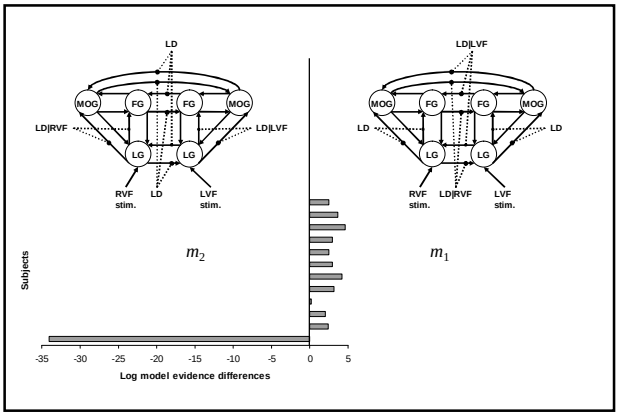
Measured data

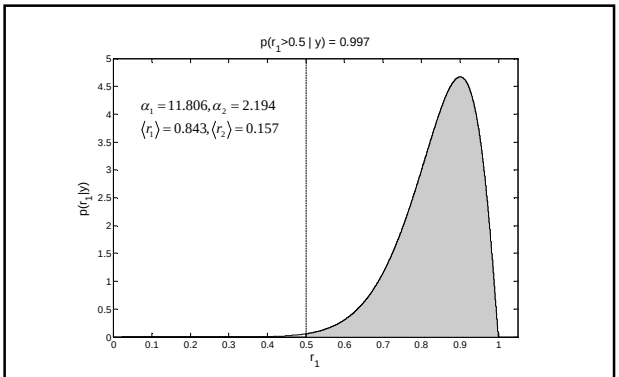
[Stephan et al. 2009, *Neuroimage*]





[Stephan et al. 2009, *Neuroimage*]



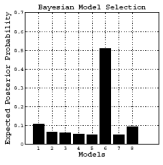


[Stephan et al. 2009, *Neuroimage*]

Interface in SPM8

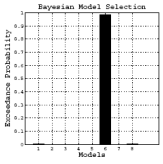
Post. expectations of model probabilities

$$\langle r_k \rangle = \alpha_k \Big/ \sum_k \alpha_k$$



Exceedance probability

$$\forall j \in \{1...K\} \setminus k:$$
$$\varphi_k = p\left(r_k > r_j \mid y\right)$$



Levels of inference: Group/population level

-- Family level --
-- System/model level --
-- Parameter/connection level --

FFX: subjects assumed to use identical systems.
RFX: optimal models vary across subjects.

[Penny et al. 2010, *PLoS Comp Biol*]

Variational Bayes: fast/accurate Nmod < N_sub.
Gibbs sampling: optimal N_mod >> N_sub.

♣ **Family level:**
- Useful when no clear winning model // models have common characteristics .

Models assigned to subsets (families) with shared parameters
➔**Inference: a class of models that best explains the data.**

♣ **System level:**
- Useful when a clear winning model can be identified (BMS).

➔**Inference: the best combination of inputs+connections that explains the data.**

♣ **Connection level:**
- Useful if interested in connectivity parameters (e.g. modulations).

➔**Inference: Bayesian parameters averaging (BPA) or t-test on DCM parameters.**
➔**Inference: BMA on the winning family (or the whole model space).**

Stephan *et al.* (2010). Ten Simple Rules for DCM. *NeuroImage*

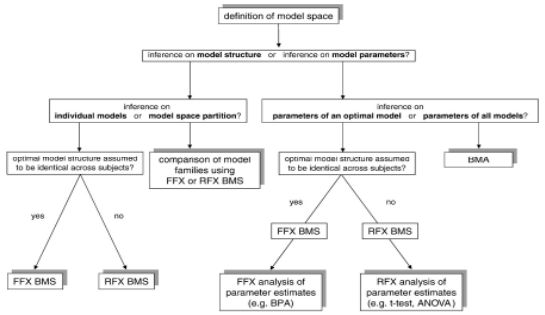


Fig. 1. This schematic summarizes the typical sequence of analysis in DCM, depending on the question of interest. Abbreviations: FFX=fixed effects, RFX=random effects, BMS=Bayesian model selection, BPA=Bayesian parameter averaging, BMA=Bayesian model averaging, ANOVA=analysis of variance.

<u>BMS has nothing to say about the "true" models</u>.
find the most plausible (useful) model, given a set of alternatives.
Best model = best balance between accuracy and complexity.

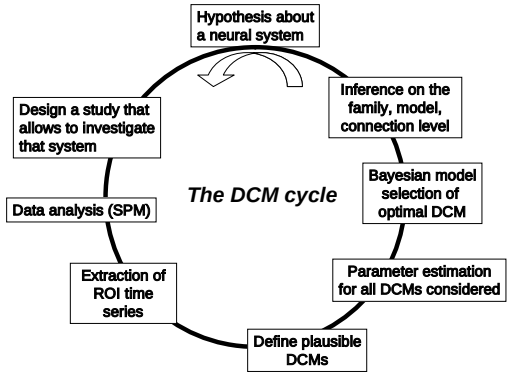
<u>BMS cannot be applied to models fitted to different data!</u>
(Only models with the same ROIs can be compared using BMS).

It is helpful to constrain your DCM model space.
number of ROIs limited to 8 in SPM8 (GUI).
(e.g., 5 ROIs, fully connected, 1 Billion alternatives for modulations!).

Define sets of models that are plausible, in a systematic way, given prior knowledge about the system (e.g. anatomical, TMS, previous studies).
→ Bad models will affect your BMS results (BMS = a "relative" space)!

Extensions in DCM for fMRI (SPM8):

• **Bayesian Model Selection BMS**	**[Penny et al. 2004 *Neuroimage*].**
• **Slice specific sampling**	**[Kiebel et al. 2007 *Neuroimage*].**
• **Refined hemodynamic model**	**[Stephan et al. 2007 *Neuroimage*].**
• **The two-state DCM**	**[Marreiros et al. 2008 *Neuroimage*].**
• **The non-linear DCM**	**[Stephan et al. 2008 *Neuroimage*].**
• **Random-effects BMS (VB)**	**[Stephan et al. 2009 *Neuroimage*].**
• **Random-effects BMS (Gibbs)**	**[Penny et al. 2010 *PLoS Comp Biol*].**
• **Stochastic DCM**	**[Daunizeau et al. 2009 *Physica D*].**
• **Anatomical-based priors for DCM**	**[Stephan et al. 2009 *Neuroimage*].**
• **Family level inference BMS**	**[Penny et al. 2010 *PLoS Comp Biol*].**
• **Bayesian model averaging BMA**	**[Penny et al. 2010 *PLoS Comp Biol*].**



Theoritical reviews:

Stephan *et al.* (2010). Ten Simple Rules for DCM**. *NeuroImage***

Daunizeau et al. (2010). DCM: a critical review of the biophysical and statistical foundations**. *NeuroImage***

Friston (2009). Causal modelling and brain connectivity in fMRI**. *PLoS Biol***

Applications: (recent examples of DCM-fMRI at the FIL)

- Word reading via the putamen:
Seghier and Price (2010) *Cerebral Cortex*.

- Intelligible speech perception:
Leff et al. (2008) *J Neurosci.*

- Associative learning and prediction error:
den Ouden et al. (2009) *Cerebral Cortex.*