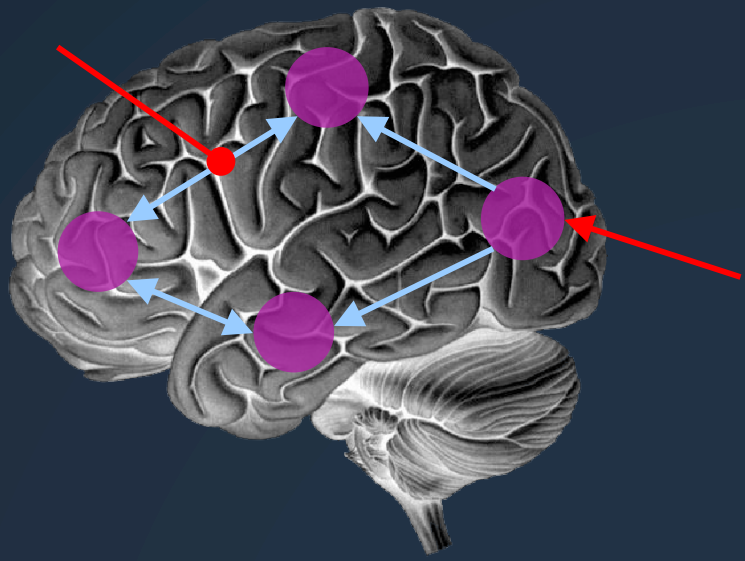
Dynamic Causal Modelling : Advanced Topics

Rosalyn Moran

Virginia Tech Carilion Research Institute

Department of Electrical & Computer Engineering, Virginia Tech



Zurich SPM Course, Feb 19th 2016

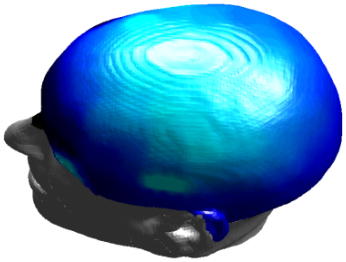
Overview

- Bayesian Model Selection
  i. Machinery
  ii. Statistics
- Non-vanilla options in DCM for fMRI
- Models in Models
  I. Computational Renderings of Behaviour in DCMS
  II. Clinical Translational Applications

Overview

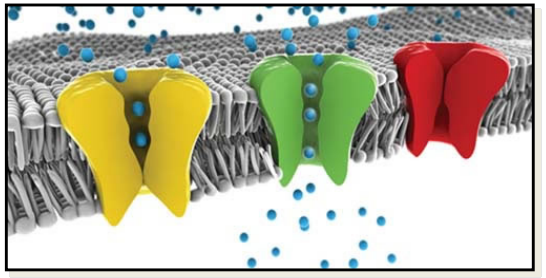
- Bayesian Model Selection
  i. Machinery
  ii. Statistics
- Non-vanilla options in DCM for fMRI
- Models in Models
  I. Computational Renderings of Behaviour in DCMS
  II. Clinical Translational Applications

Generative model



$$p(y \mid \theta, m) \cdot p(\theta \mid m)$$

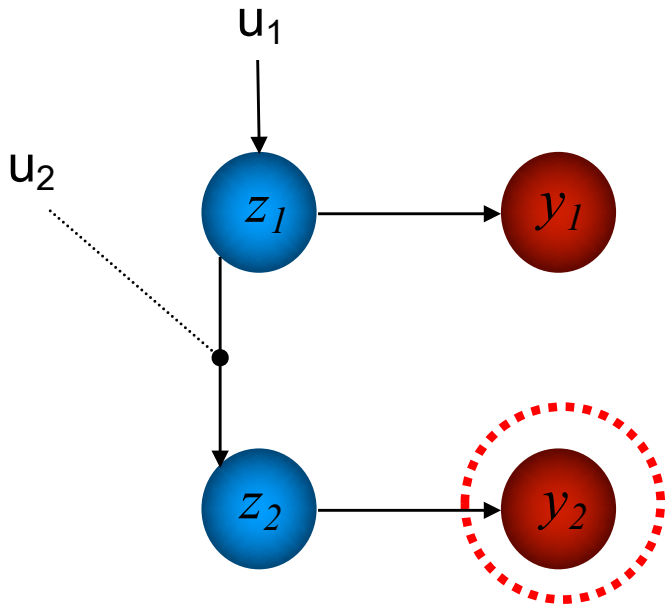
$$p(\theta \mid y, m)$$



1. enforces mechanistic thinking: how could the data have been caused? Where do we look next?
2. generate synthetic data (observations) by sampling from the prior – can model explain certain phenomena at all?
3. inference about model structure: formal approach to disambiguating mechanisms → p(y|m)
4. inference about parameters → p(θ|y)
5. A natural framework to test interventions → adjusting the optimized state space

Parameter estimation: Bayesian inversion

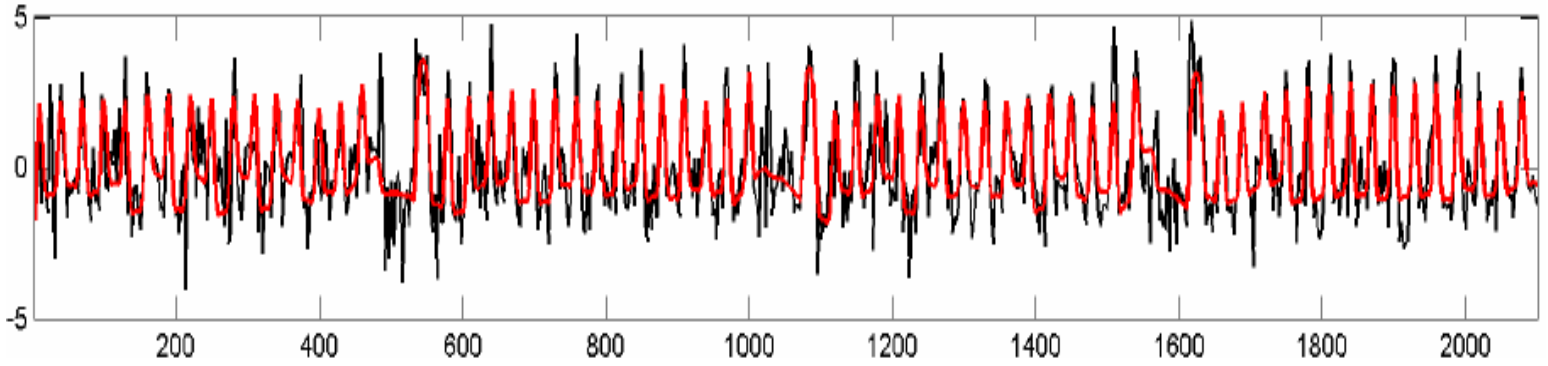
Estimate neural & hemodynamic parameters such that the **MODELLED** and **MEASURED** BOLD signals are similar (model evidence is optimised), using variational EM under Laplace approximation



... Noise Assumption (Gaussian with temporal correlations) leads to a probabilistic model.

So for one data point :

$$p(y \mid \theta, m) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(y-ym)^2}{2\sigma^2}}$$

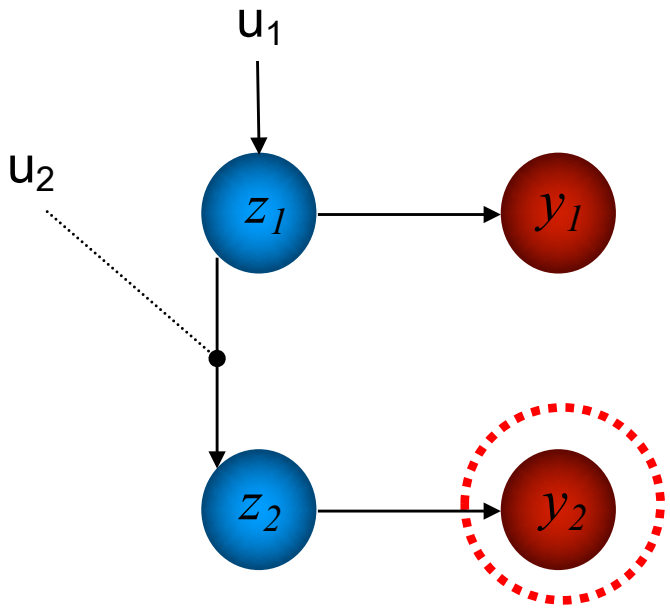


Parameter estimation: Bayesian inversion

Have Data

Have Likelihood

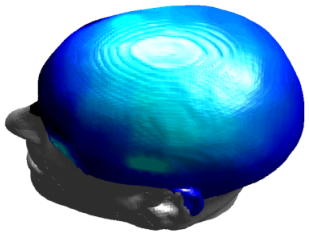
Specify Prior



Given by model form & parameters

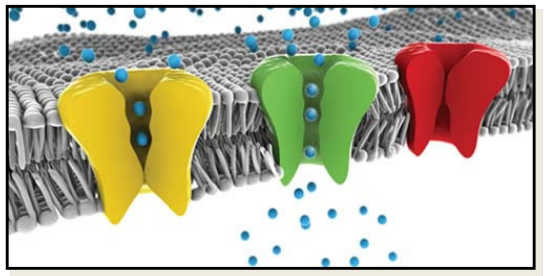
$$p(y \mid \theta, m) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(y-ym)^2}{2\sigma^2}}$$

Hyperparameter in the model



$$p(y \mid \theta, m) \cdot p(\theta \mid m)$$

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



Parameter estimation: Bayesian inversion

Have Data
Have Likelihood
Specify Prior
The Posterior:

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

Marginal Likelihood – tricky integral... try it !

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

Instead – let's assume a partition over parameter space is possible

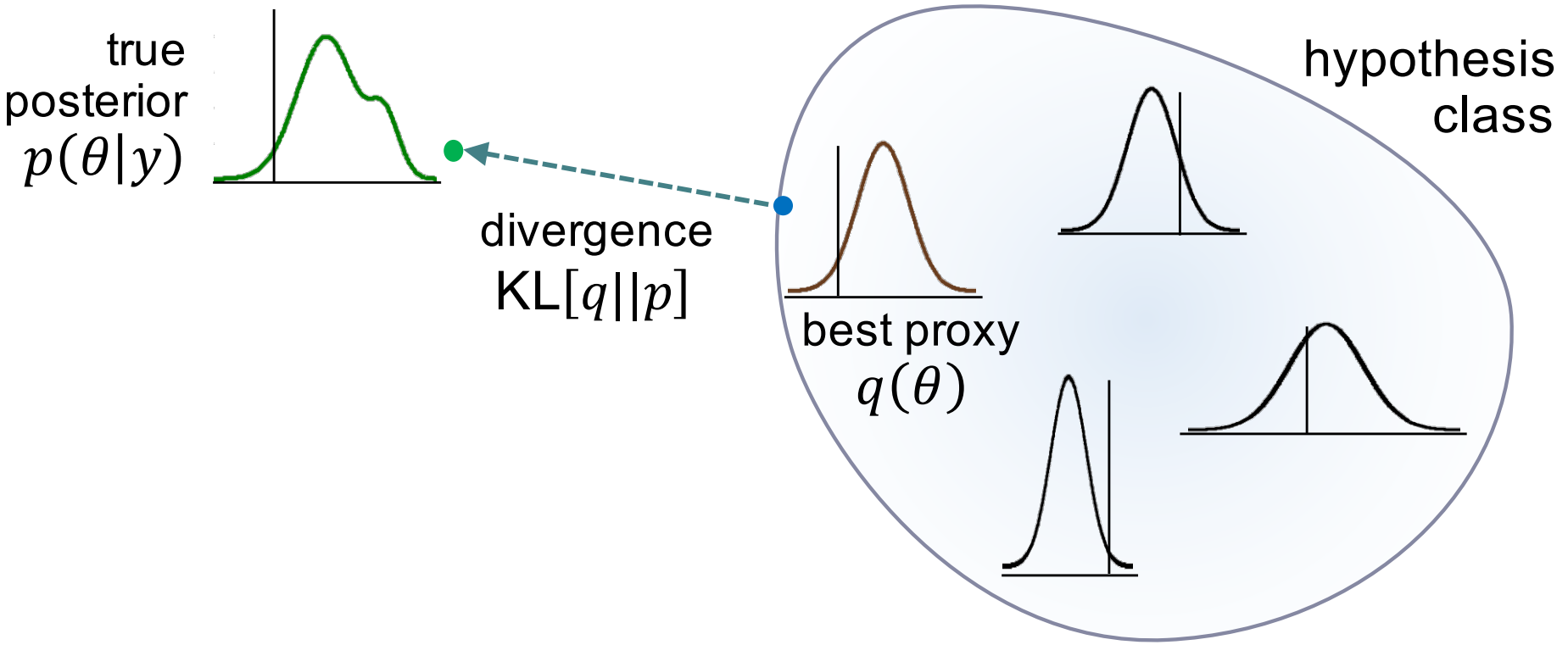
1. Split hyperparameter from parameters
2. Assume some form of posterior

$$p(\theta,\lambda \mid y) \approx q(\theta,\lambda) = q(\theta)\,q(\lambda)$$

Variational Bayes (VB)

Idea: find an approximation $q(\theta)$ to the true posterior $p(\theta|y)$.

Often done by assuming a particular form for q and then optimizing its sufficient statistics (fixed form VB).



Parameter estimation: Variational Bayes

Free-energy approximation to marginal likelihood

$$F(q) = lnp(y|m) - KL(q(\theta,\lambda), p(\theta,\lambda|y))$$

Turned an integral problem into a simpler optimization (find max) where standard procedures can be applied

Maximise variational energies via gradient ascent

$$q(\theta) \propto \exp(I_\theta) = \exp\left[\left\langle \ln p(y,\theta,\lambda)\right\rangle_{q(\lambda)}\right]$$

$$q(\lambda) \propto \exp(I_\lambda) = \exp\left[\left\langle \ln p(y,\theta,\lambda)\right\rangle_{q(\theta)}\right]$$

Parameter estimation & model evidence: Variational Bayes

Regional responses

Specify generative forward model
(with prior distributions of parameters)

Variational Expectation-Maximization algorithm

Iterative procedure:

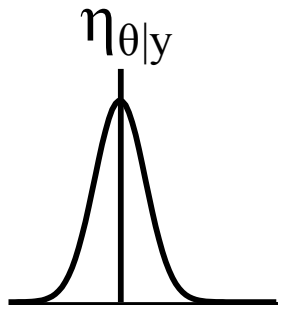
1. Compute model response using current set of parameters
2. Compare model response with data
3. Improve parameters, if possible

1. Gaussian posterior distributions of parameters

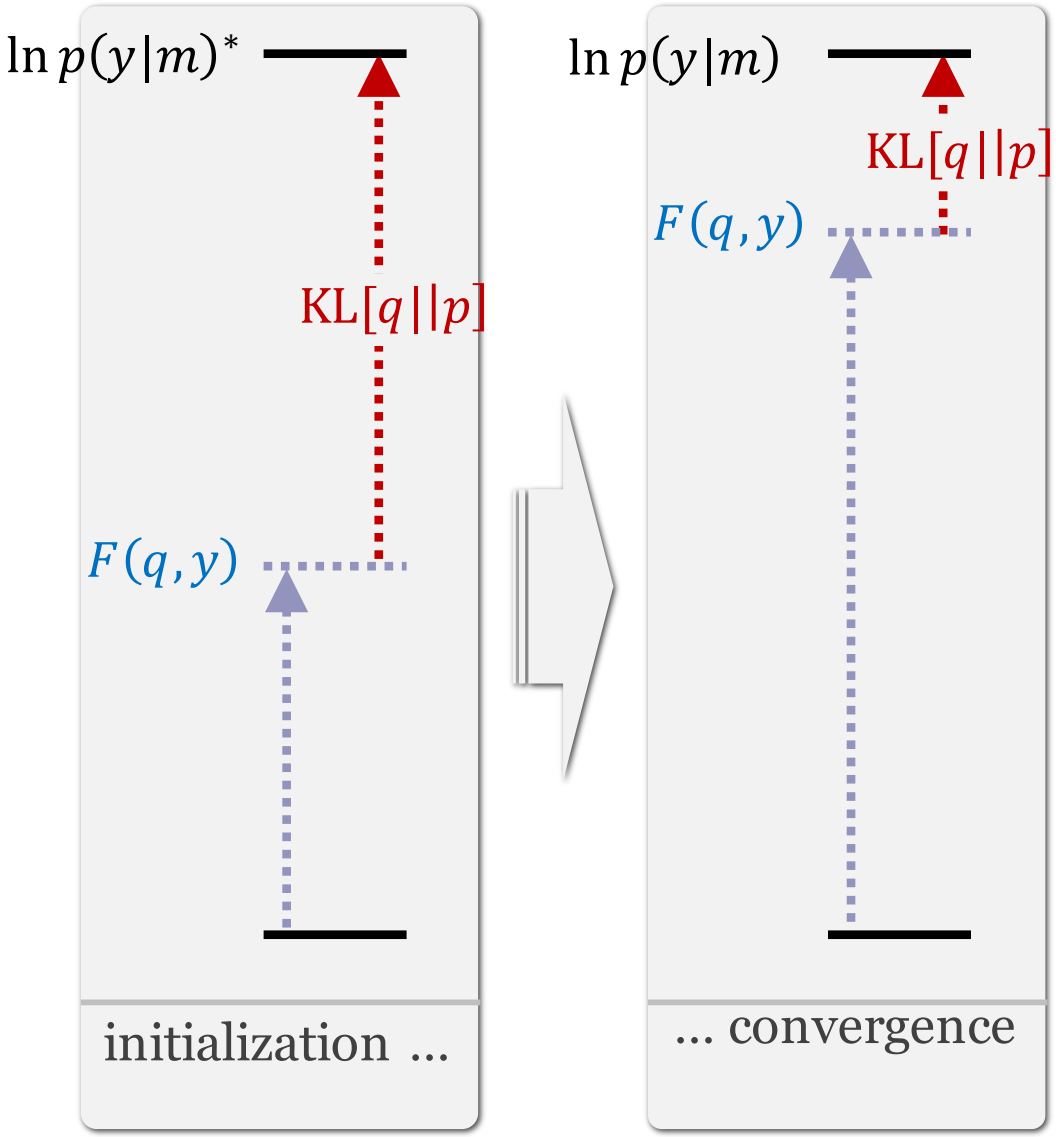
2. Model evidence

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

$p(y \mid m)$



Parameter estimation & model evidence: Variational Bayes



Alternative Schemes for Parameter Estimation

NeuroImage 125 (2016) 1107–1118

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Technical Note

Gradient-based MCMC samplers for dynamic causal modelling

Biswa Sengupta *, Karl J. Friston, Will D. Penny

Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, UK

ARTICLE INFO

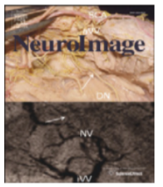
Article history:
Received 4 October 2014
Accepted 16 July 2015
Available online 23 July 2015

ABSTRACT

In this technical note, we derive two MCMC (Markov chain Monte Carlo) samplers for dynamic causal models (DCMs). Specifically, we use (a) Hamiltonian MCMC (HMC-E) where sampling is simulated using Hamilton's equation of motion and (b) Langevin Monte Carlo algorithm (LMC-R and LMC-E) that simulates the Langevin diffusion of samples using gradients either on a Euclidean (E) or on a Riemannian (R) manifold. While LMC-R requires minimal tuning, the implementation of HMC-E is heavily dependent on its tuning parameters. These parameters are therefore optimised by learning a Gaussian process model of the time-normalised sample correlation matrix. This allows one to formulate an objective function that balances tuning parameter exploration and exploitation, furnishing an intervention-free inference scheme. Using neural mass models (NMMs)—a class of biophysically motivated DCMs—we find that HMC-E is statistically more efficient than LMC-R (with a Riemannian metric); yet both gradient-based samplers are far superior to the random walk Metropolis algorithm, which proves inadequate to steer away from dynamical instability.

Direct

/locate/ynimg



Gradient-free MCMC methods for dynamic causal modelling

Biswa Sengupta *, Karl J. Friston, Will D. Penny

Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, UK

ARTICLE INFO

Article history:
Received 3 October 2014
Accepted 6 March 2015
Available online 14 March 2015

ABSTRACT

In this technical note we compare the performance of four gradient-free MCMC samplers (random walk Metropolis sampling, slice-sampling, adaptive MCMC sampling and population-based MCMC sampling with tempering) in terms of the number of independent samples they can produce per unit computational time. For the Bayesian inversion of a single-node neural mass model, both adaptive and population-based samplers are more efficient compared with random walk Metropolis sampler or slice-sampling; yet adaptive MCMC sampling is more promising in terms of compute time. Slice-sampling yields the highest number of independent samples from the target density — albeit at almost 1000% increase in computational time, in comparison to the most efficient algorithm (i.e., the adaptive MCMC sampler).

Overview

- Bayesian Model Selection
  i. Machinery
  ii. Statistics
- Non-vanilla options in DCM for fMRI
- Models in Models
  I. Computational Renderings of Behaviour in DCMS
  II. Clinical Translational Applications

Statistics

Effects of Acute Tryptophan Depletion on Prefrontal-Amygdala Connectivity While Viewing Facial Signals of Aggression

Passamonti et al. Biological Psychiatry 2012

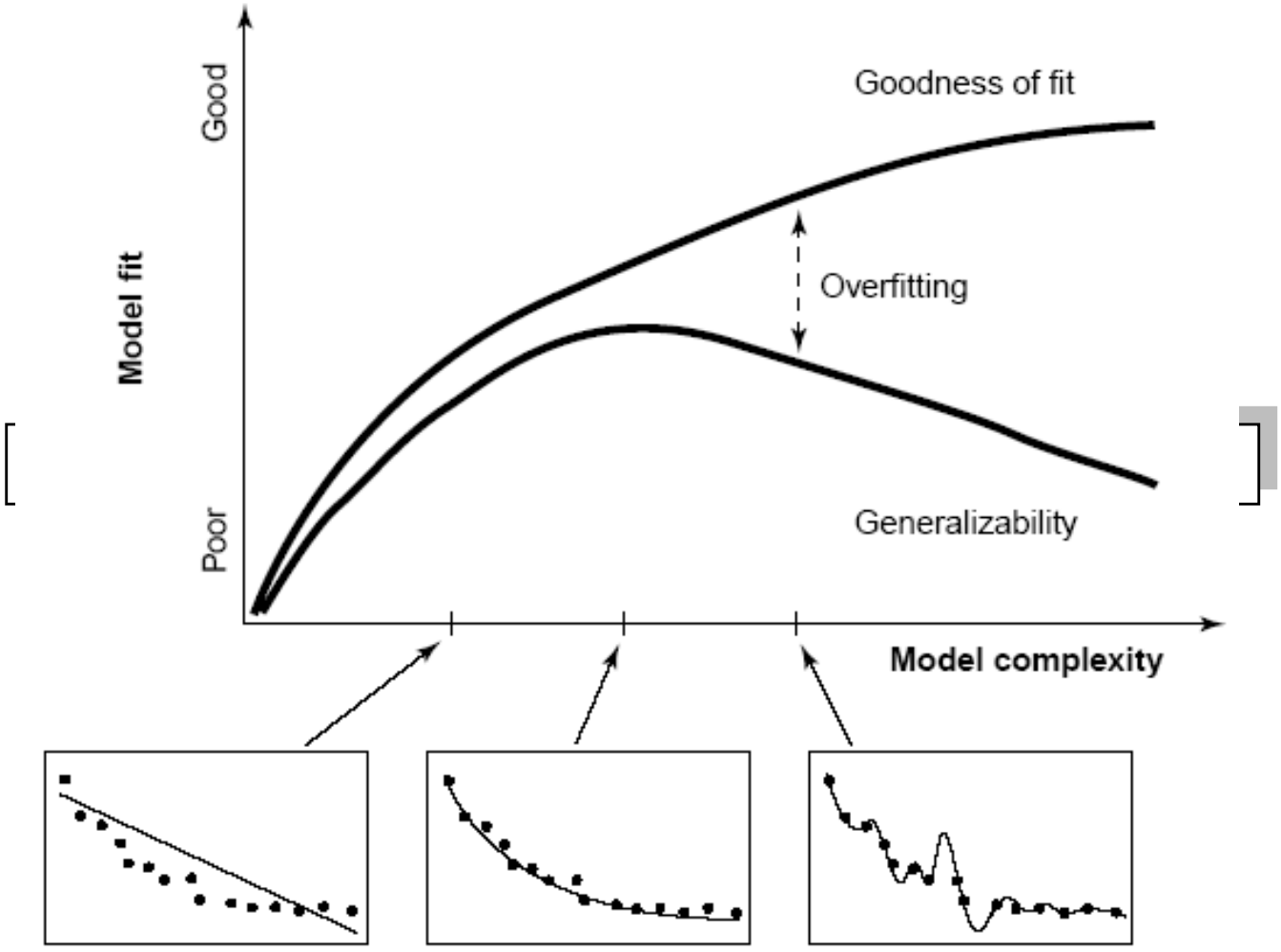
" Analysis of Effective Connectivity 2: DCM
For placebo, RFX-BMS indicated evidence favoring model C1.1. Driving inputs (all faces) entered the system via the amygdala alone, whereas the angry vs. neutral modulator affected bidirectional connections in all three pathways. Hence, during placebo, the effect of the task is distributed within internal PFC circuitry and across PFC–amygdala connections.
Under ATD, the expected and exceedance probabilities of C1.1 were reduced with increased expected and exceedance probabilities of the two models (C2.1, C3.1) in which the contextual modulator acted on two or one bidirectional connections. Furthermore, during ATD, another family of six models became more likely than under placebo, in which driving inputs "perturbed" the network via either the VLPFC or vACC alone.
To summarize, ATD reduced not only the number of PFC–amygdala pathways affected by processing angry faces but also the location where face information entered the network."

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?

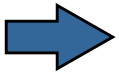
For which model *m* does $p(y|m)$ become maximal?



Pitt & Miyung (2002) *TICS*

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

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

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

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

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

No. of parameters

No. of data points

Penny et al. 2004a, *NeuroImage*

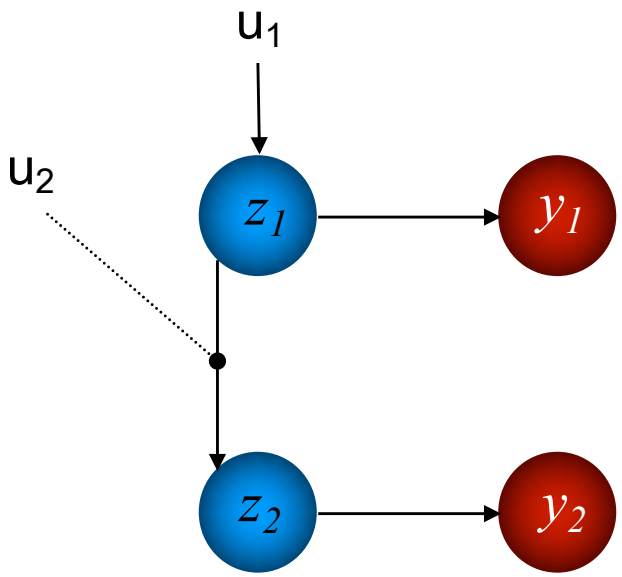
Bayesian Model Comparison

The model goodness: Negative Free Energy

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

Accuracy - Complexity

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



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

Bayes factors

To compare two models, we could 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; ∞[

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

Kass & Raftery classification:

B_{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

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

Model Comparison at the Group Level

- Prior to/ instead of inference on parameters
- Which of various mechanisms / models best explains my data
- Use model evidence

→ accounts for both accuracy and complexity of the model

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

Fixed Effects Model selection via

log Group Bayes factor:

$$BF_{1,2} = \sum_k \ln p(y|m_1) - \sum_k \ln p(y|m_2)$$

Random Effects Model selection

via Model probability:

$$p(r \mid y, \alpha)$$

$$\langle r_k \rangle_q = \alpha_k / (\alpha_1 + \ldots + \alpha_K)$$

Fixed effects BMS at group level‘

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

$$GBF_{ij} = \prod_k BF_{ij}^{(k)}$$

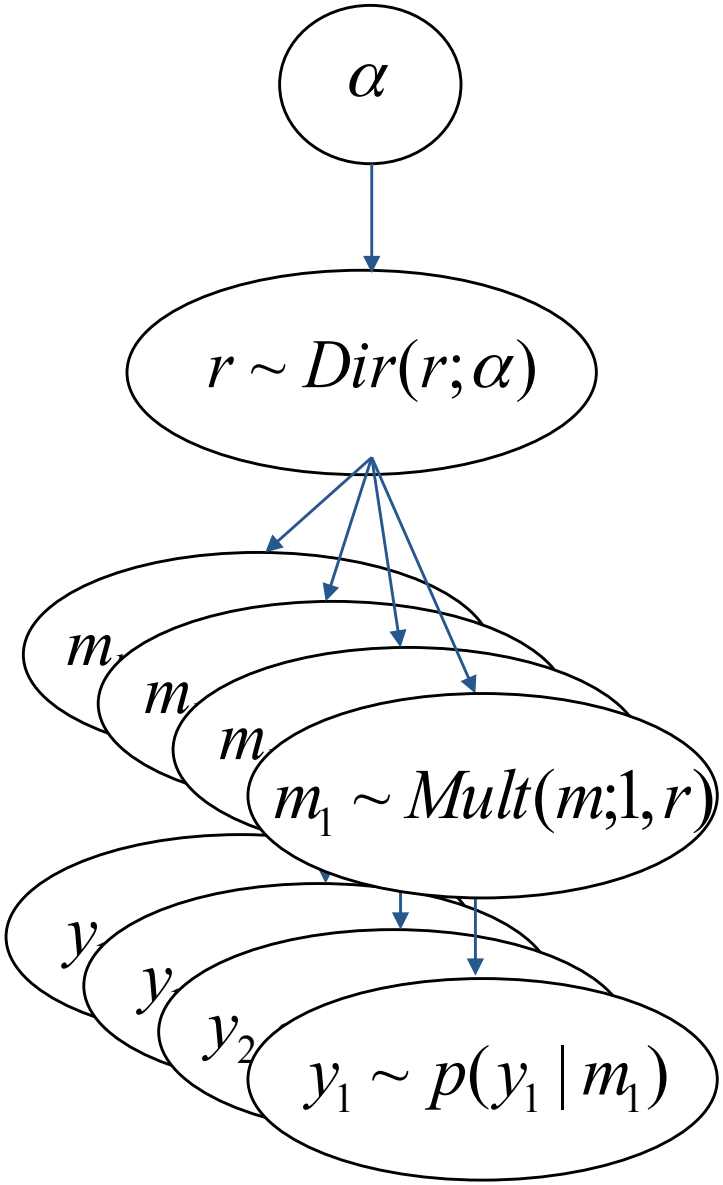
Average Bayes factor (ABF):

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

Problems:

- blind with regard to group heterogeneity
- sensitive to outliers

Random effects BMS for heterogeneous groups



Dirichlet parameters α
= "occurrences" of models in the population

Dirichlet distribution of model probabilities r

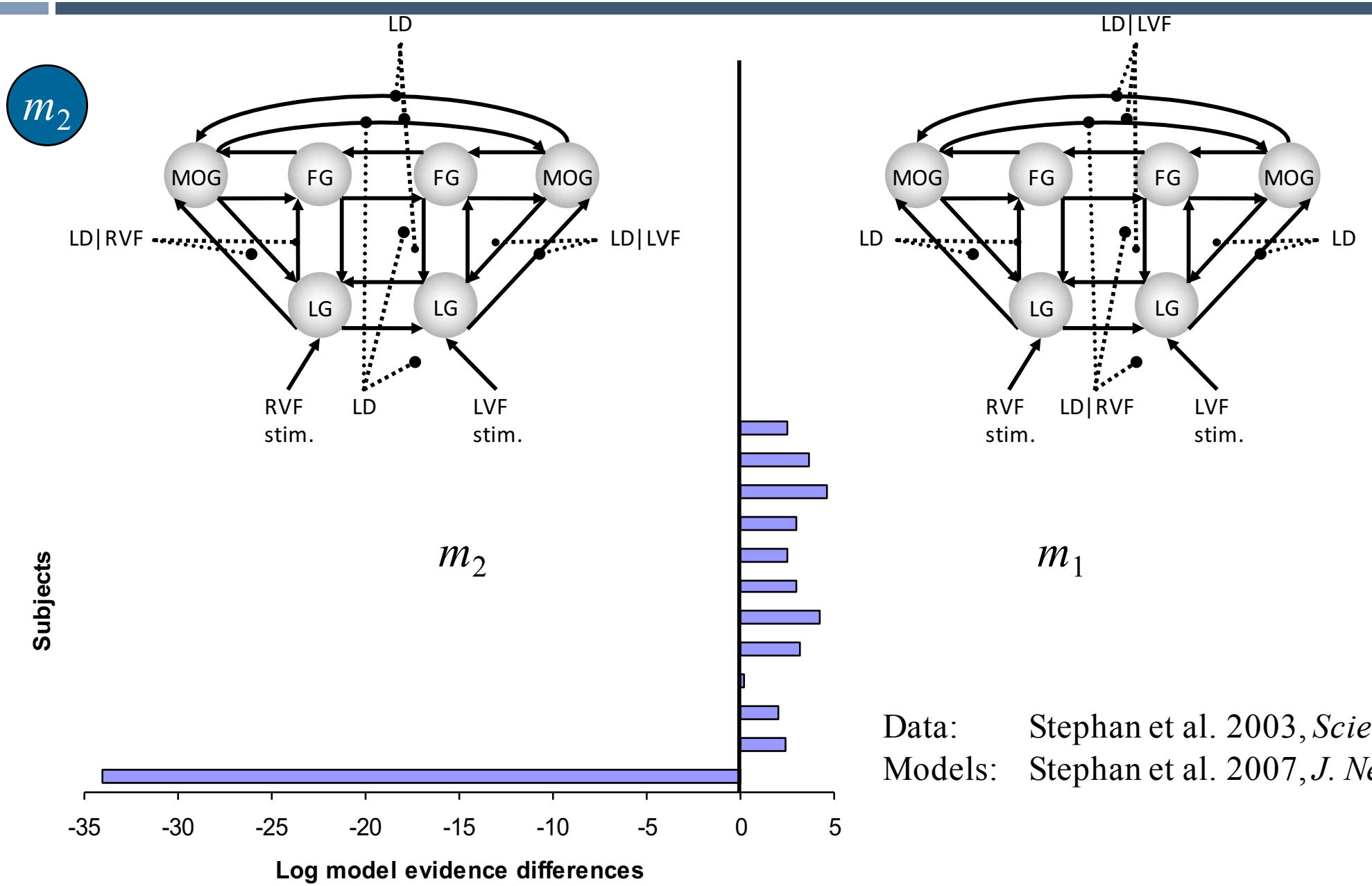
Multinomial distribution of model labels m

Model inversion by Variational Bayes (VB) or MCMC

Measured data y

Stephan et al. 2009a, *NeuroImage*
Penny et al. 2010, *PLoS Comp. Biol.*

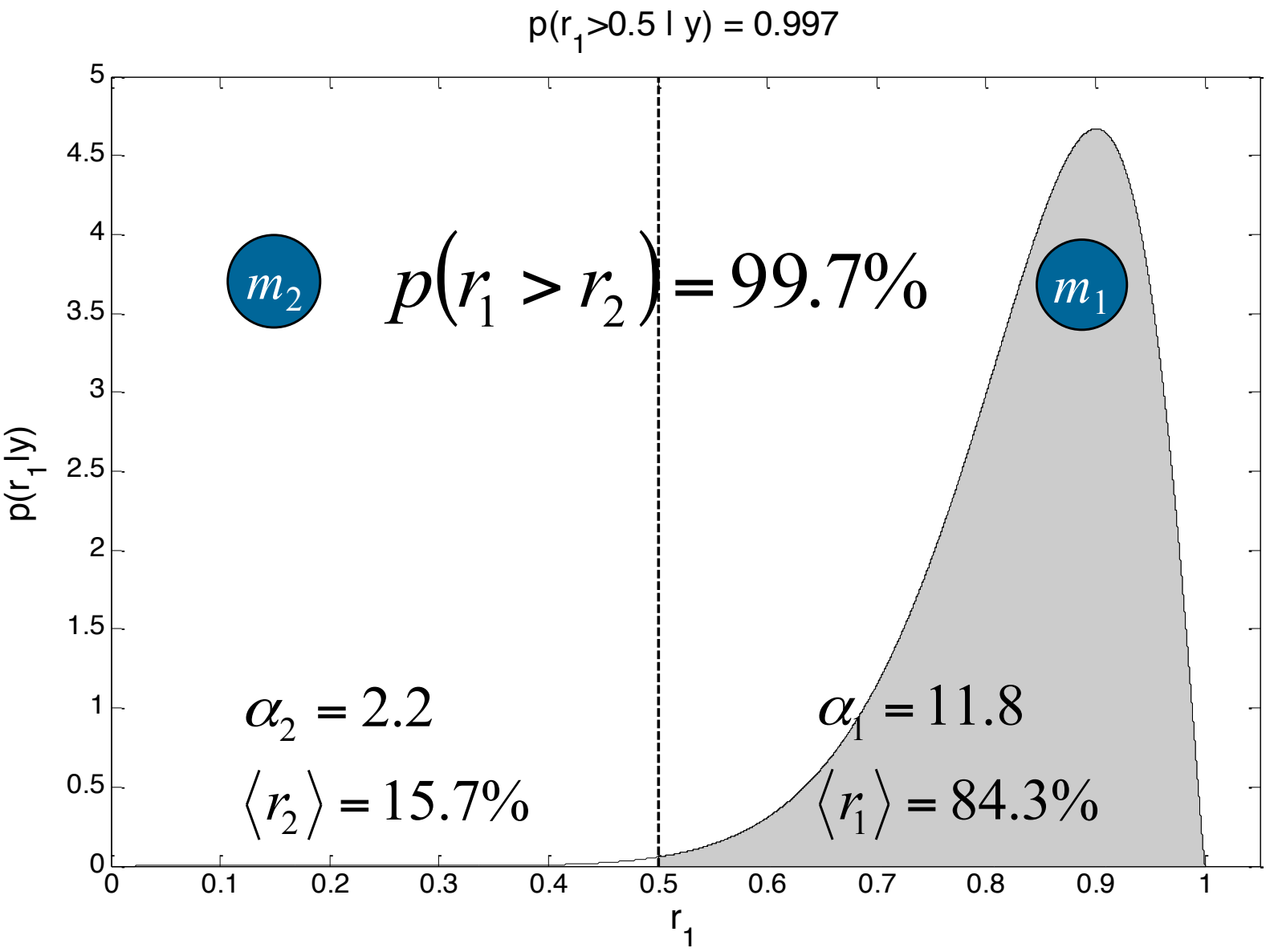
Random effects BMS for heterogeneous groups



Data: Stephan et al. 2003, *Science*

Models: Stephan et al. 2007, *J. Neurosci.*

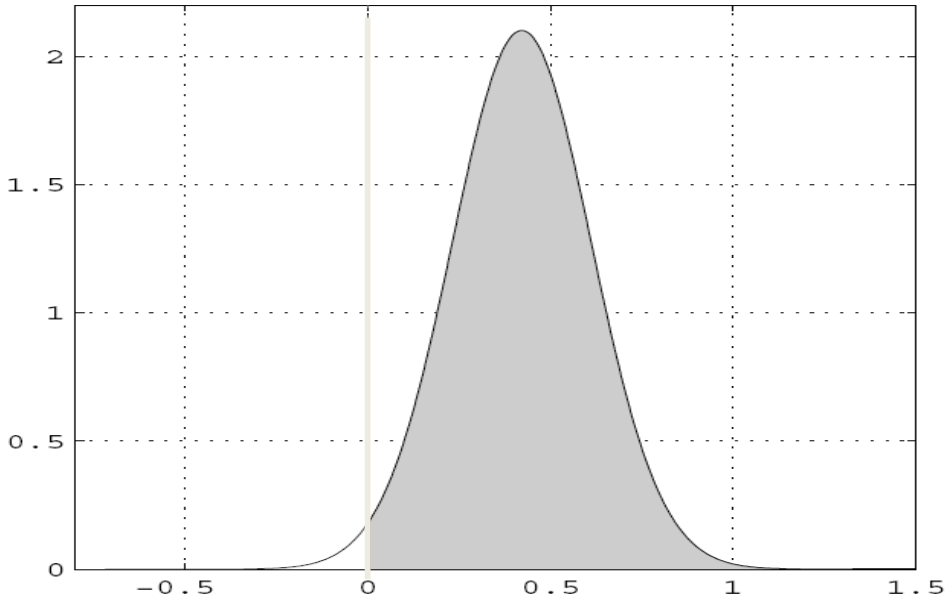
Random effects BMS for heterogeneous groups



Stephan et al. 2009a, *NeuroImage*

Inference about DCM parameters: Bayesian single subject analysis

- Gaussian assumptions about the posterior distributions of the parameters
- posterior probability that a certain parameter (or contrast of parameters) is above a chosen threshold γ:
- By default, γ is chosen as zero – the prior ("does the effect exist?").



Inference about DCM parameters: Bayesian parameter averaging

FFX group analysis:

- Likelihood distributions from different subjects are independent
- Under Gaussian assumptions, this is easy to compute
- Simply 'weigh' each subject's contribution by your certainty of the parameter

group posterior covariance

individual posterior covariances

$$C^{-1}_{\theta|y_1,\ldots,y_N} = \sum_{i=1}^{N} C^{-1}_{\theta|y_i}$$

$$\eta_{\theta|y_1,\ldots,y_N} = \left(\sum_{i=1}^{N} C^{-1}_{\theta|y_i} \eta_{\theta|y_i} \right) C_{\theta|y_1,\ldots,y_N}$$

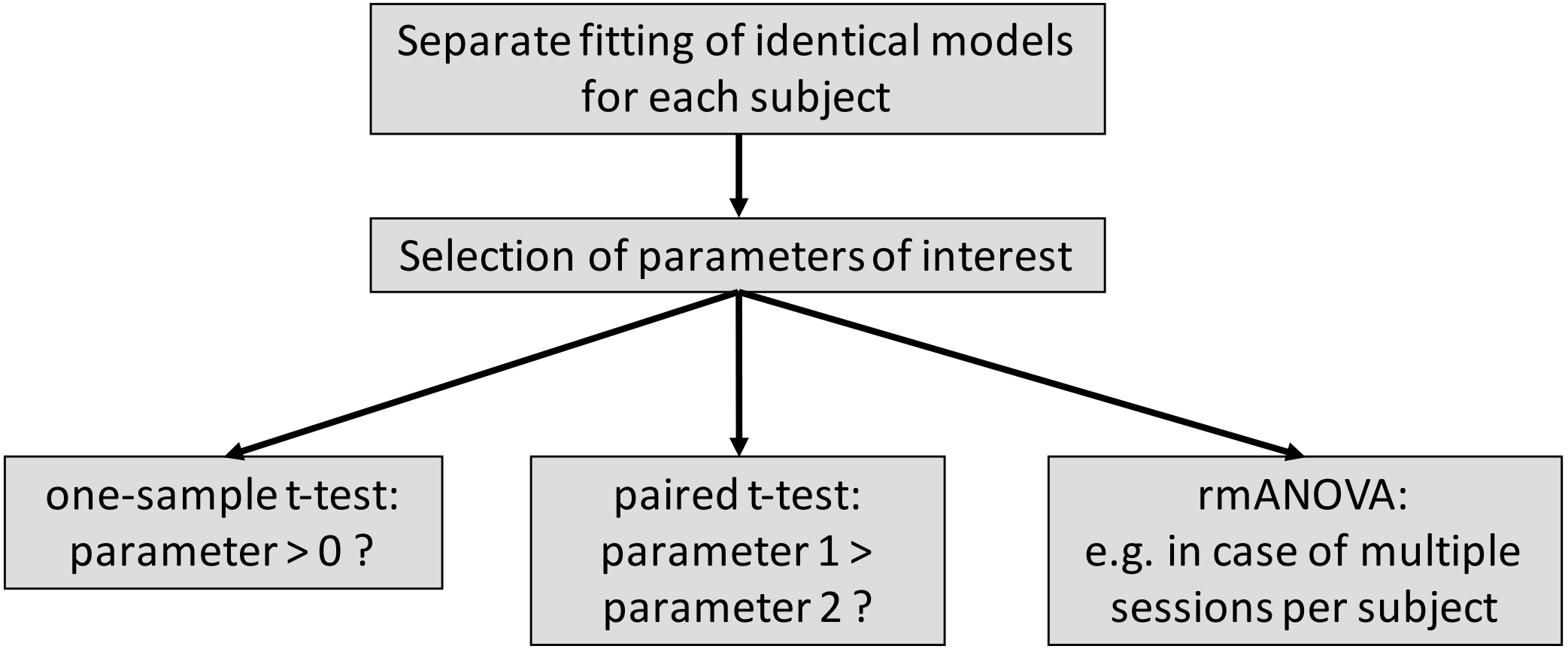
group posterior mean

individual posterior covariances and means

Useful but use carefully!

Inference about DCM parameters: RFX analysis (frequentist)

RFX Approach:
Analogous to 'random effects' analyses in SPM, 2nd level analyses can be applied to DCM parameters



Putting model and parameter space together

Bayesian Model Averaging (BMA)

- abandons dependence of parameter inference on a single model
- uses the entire model space considered (or an optimal family of models)
- computes average of each parameter, weighted by posterior model probabilities
- represents a particularly useful alternative
  - when none of the models (or model subspaces) considered clearly outperforms all others
  - when comparing groups for which the optimal model differs

$$p\left(\theta_n \mid y_{1..N}\right)$$
$$= \sum_m p\left(\theta_n \mid y_n, m\right) p\left(m \mid y_{1..N}\right)$$

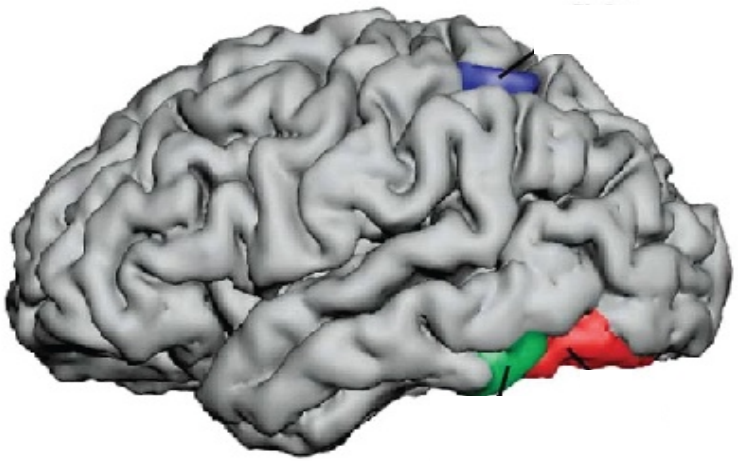
NB: $p(m|y_{1..N})$ can be obtained by either FFX or RFX BMS

Penny et al. 2010, *PLoS Comput. Biol.*

Example 2: Brain Connectivity in Synesthesia

Study: van Leeuwen, den Ouden & Hagoort, 2011

- Specific sensory stimuli lead to unusual, additional experiences
- Grapheme-color synesthesia: **color**
- Involuntary, automatic; stable over time, prevalence ~4%
- Potential cause: aberrant **cross-activation** between
  - grapheme encoding area (letter-shape area)
  - color area V4
  - superior parietal lobule (SPL)



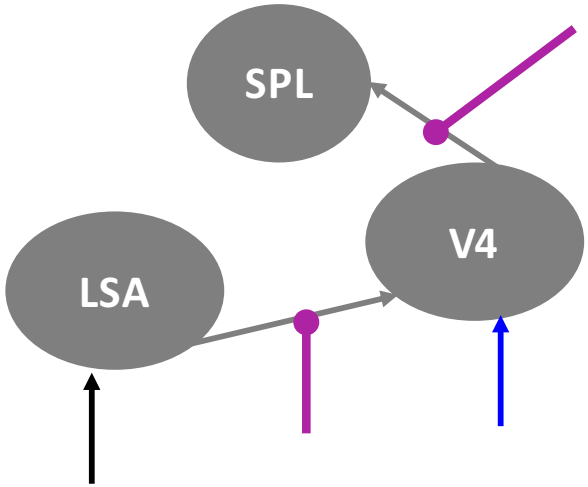
Hubbard, 2007

Can changes in effective connectivity explain synesthesia activity in V4?

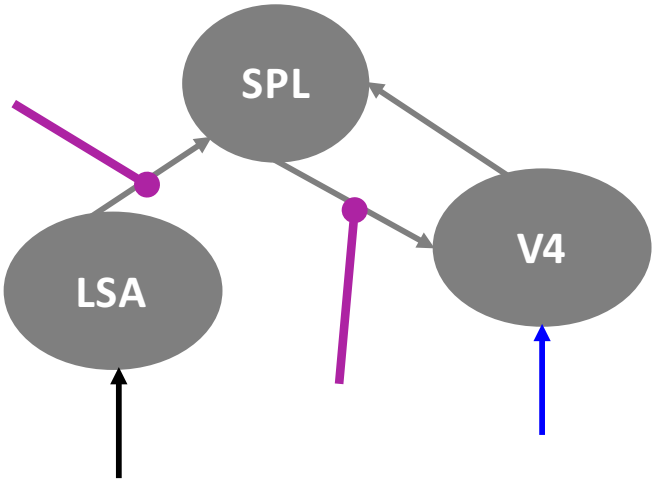
Example 2: Brain Connectivity in Synesthesia

Competing Model Hypotheses

Model 1 Grapheme color synesthesia induced by direct cross-activation in ventral-occipital cortex



Model 2 Disinhibition feedback hypothesis, aberrant feedback from multimodal region SPL, activating V4

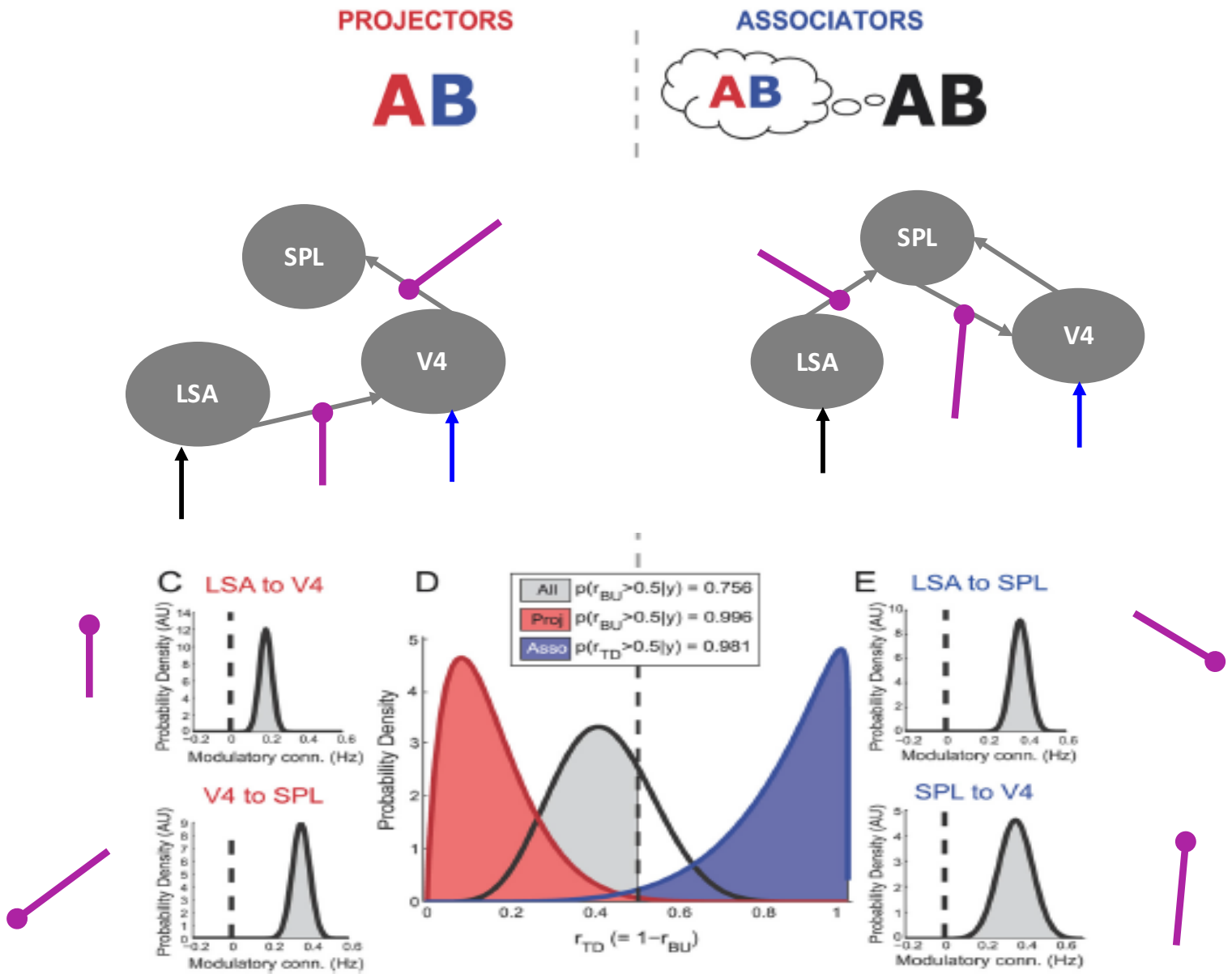


Stimuli: Graphemes (G)

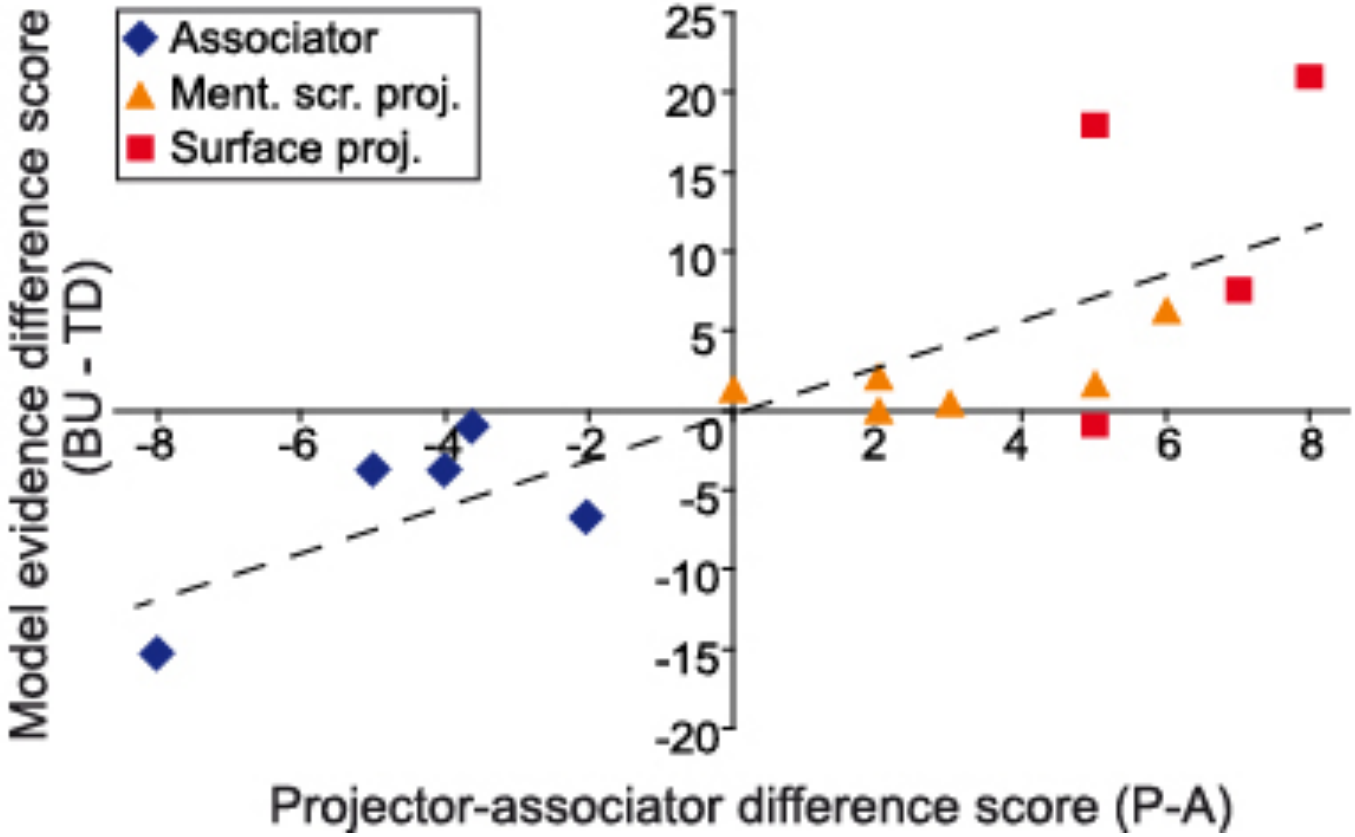
- Synesthesia-inducing graphemes (SG)
- Colored Control graphemes (CG)

Initially no winning model

Example 2: Brain Connectivity in Synesthesia



Relative model evidence predicts sensory experience



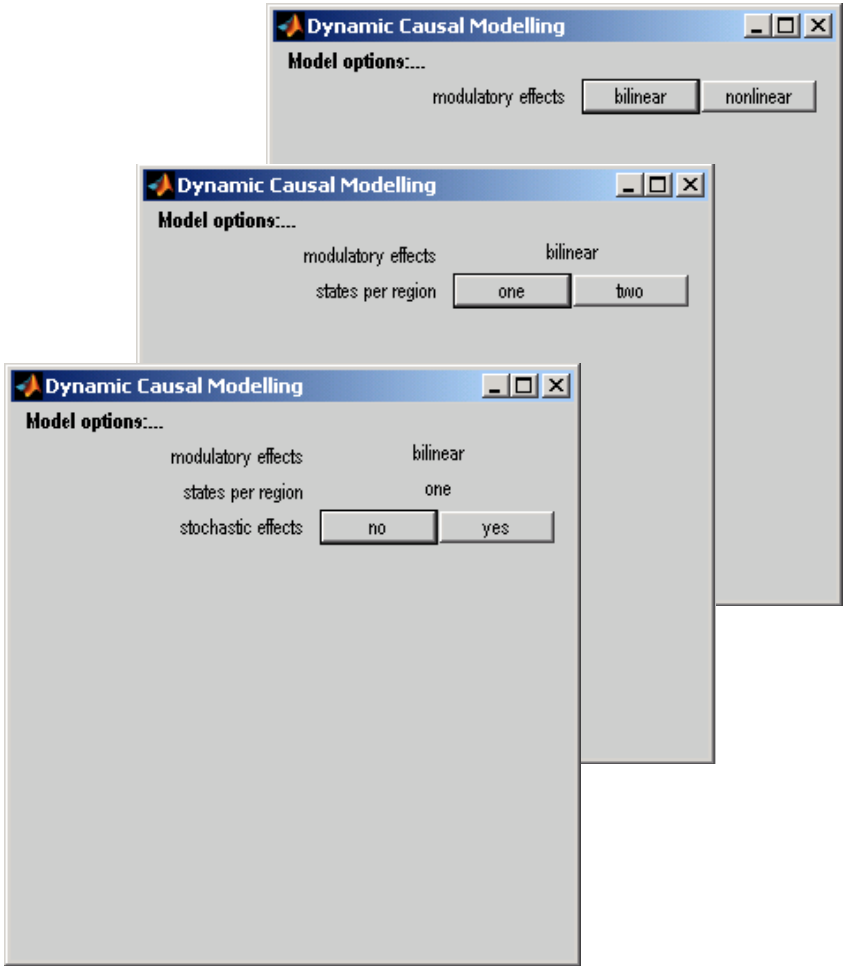
Van Leeuwen, den Ouden, Hagoort (2011) *JNeurosci*

Overview

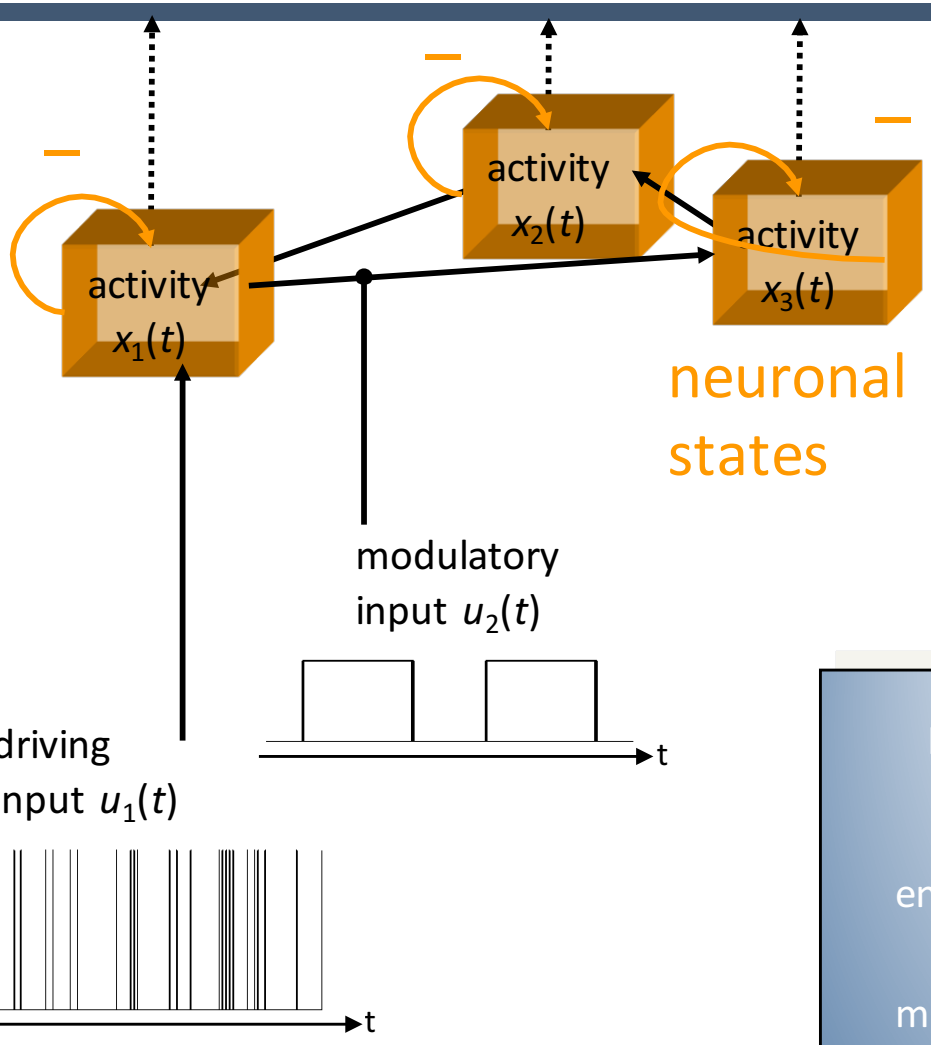
- Bayesian Model Selection
  i. Machinery
  ii. Statistics
- Non-vanilla options in DCM for fMRI
- Models in Models
  I. Computational Renderings of Behaviour in DCMS
  II. Clinical Translational Applications

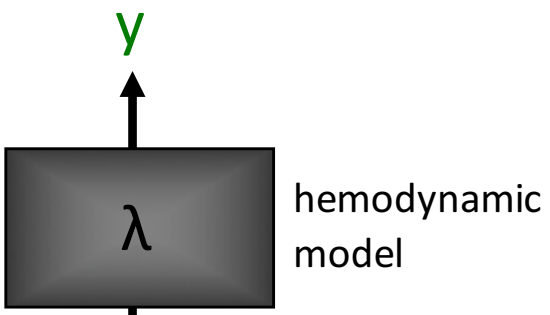
Factorial structure of model specification

- Three dimensions of model specification:
  - bilinear vs. nonlinear
  - single-state vs. two-state (per region)
  - deterministic vs. stochastic



bilinear DCM





integration

Neural state equation

$$\dot{x} = (A + \sum u_j B^{(j)})x + Cu$$

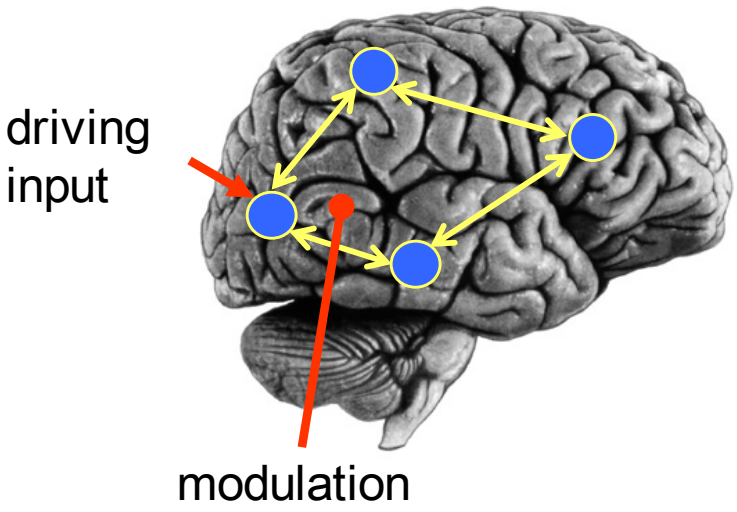
endogenous connectivity → $A = \frac{\partial \dot{x}}{\partial x}$

modulation of connectivity → $B^{(j)} = \frac{\partial}{\partial u_j}\frac{\partial \dot{x}}{\partial x}$

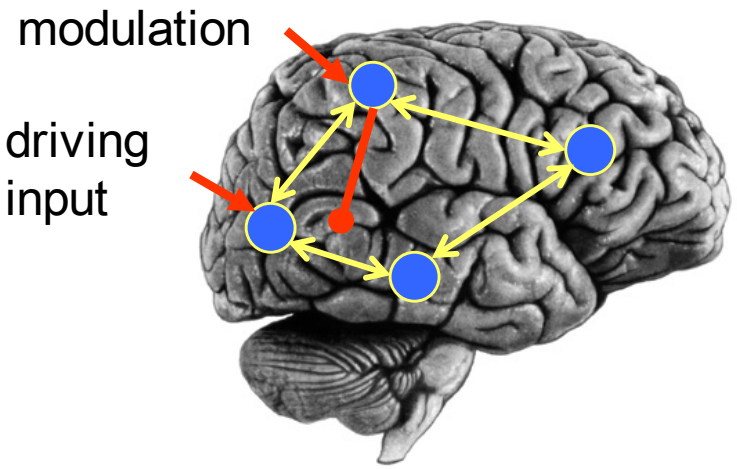
direct inputs → $C = \frac{\partial \dot{x}}{\partial u}$

The classical DCM:
a deterministic, one-state, bilinear model

bilinear DCM



non-linear DCM



Two-dimensional Taylor series (around $x_0=0$, $u_0=0$):

$$\frac{dx}{dt} = f(x,u) \approx f(x_0,0) + \frac{\partial f}{\partial x}x + \frac{\partial f}{\partial u}u + \frac{\partial^2 f}{\partial x \partial u}ux + \frac{\partial^2 f}{\partial x^2}\frac{x^2}{2} + \ldots$$

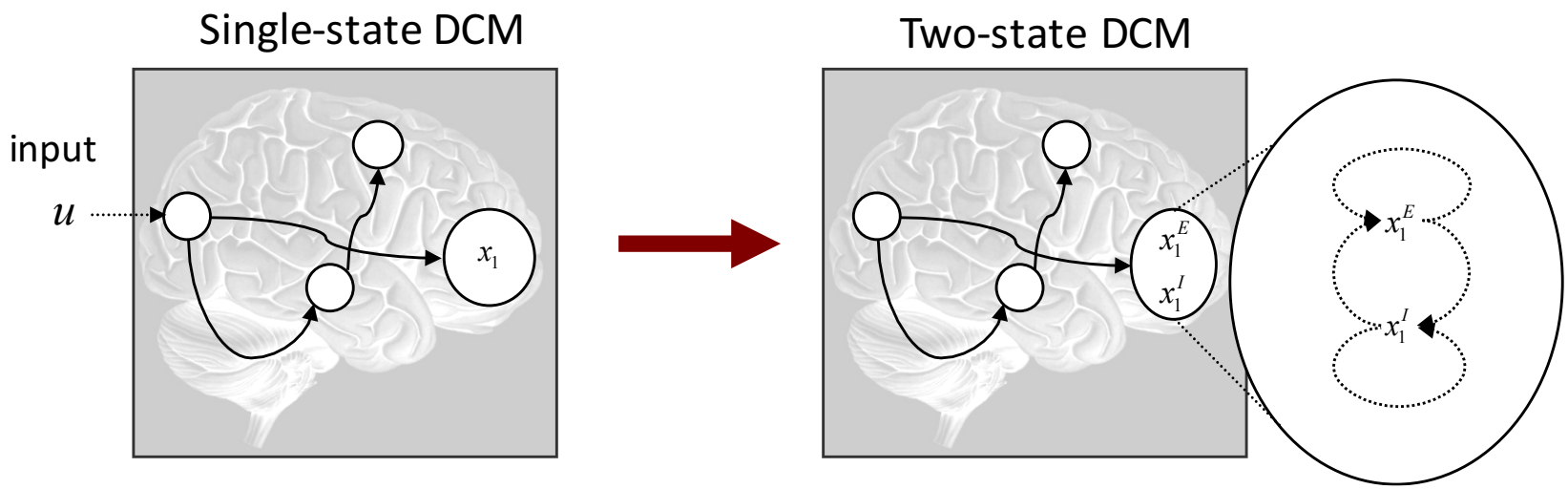
Bilinear state equation:

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

Nonlinear state equation:

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

Two-state DCM



$$\dot{x} = \Im x + Cu$$

$$\Im_{ij} = A_{ij} + uB_{ij}$$

$$\Im = \begin{bmatrix} \Im_{11} & \cdots & \Im_{1N} \\ \vdots & \ddots & \vdots \\ \Im_{N1} & \cdots & \Im_{NN} \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

$$\Im = \begin{bmatrix} \Im_{11}^{EE} & \Im_{11}^{EI} & \cdots & \Im_{1N}^{EE} & 0 \\ \Im_{11}^{IE} & \Im_{11}^{II} & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ \Im_{N1}^{EE} & 0 & & \Im_{NN}^{EE} & \Im_{NN}^{EE} \\ 0 & 0 & \cdots & \Im_{NN}^{IE} & \Im_{NN}^{II} \end{bmatrix} \quad x = \begin{bmatrix} x_1^E \\ x_1^I \\ \vdots \\ x_N^E \\ x_N^I \end{bmatrix}$$

Extrinsic (between-region) coupling

Intrinsic (within-region) coupling

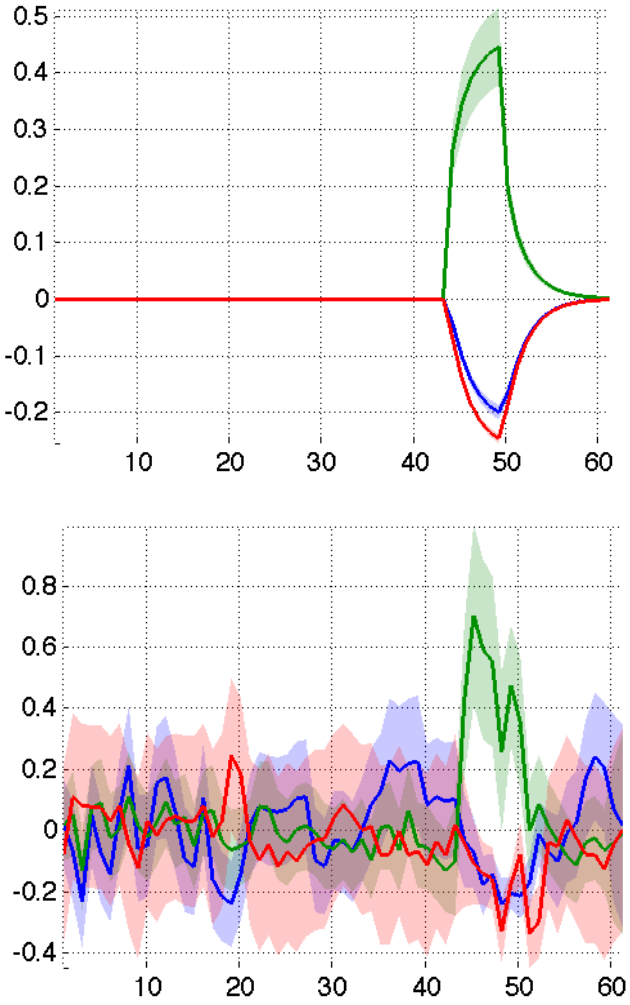
Marreiros et al. 2008, *NeuroImage*

Stochastic DCM

$$\frac{dx}{dt} = (A + \sum_j u_j B^{(j)})x + Cv + \omega^{(x)}$$

$$v = u + \omega^{(v)}$$

- all states are represented in generalised coordinates of motion
- random state fluctuations $w^{(x)}$ account for endogenous fluctuations,
  have unknown precision and smoothness
  → two hyperparameters
- fluctuations $w^{(v)}$ induce uncertainty about how inputs influence neuronal activity
- can be fitted to "resting state" data but not necessarily
- DEM optimization scheme



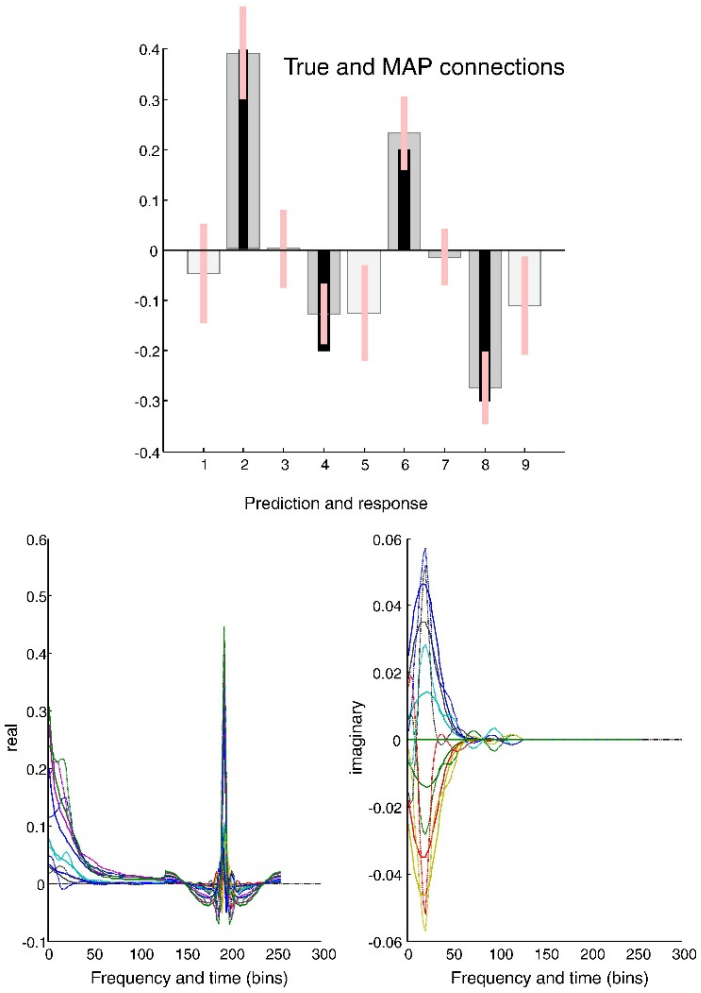
Daunizeau et al, 2009

Friston et al, 2008

Li et al. 2011, *NeuroImage*

Spectral DCM for ‚Resting Stare'

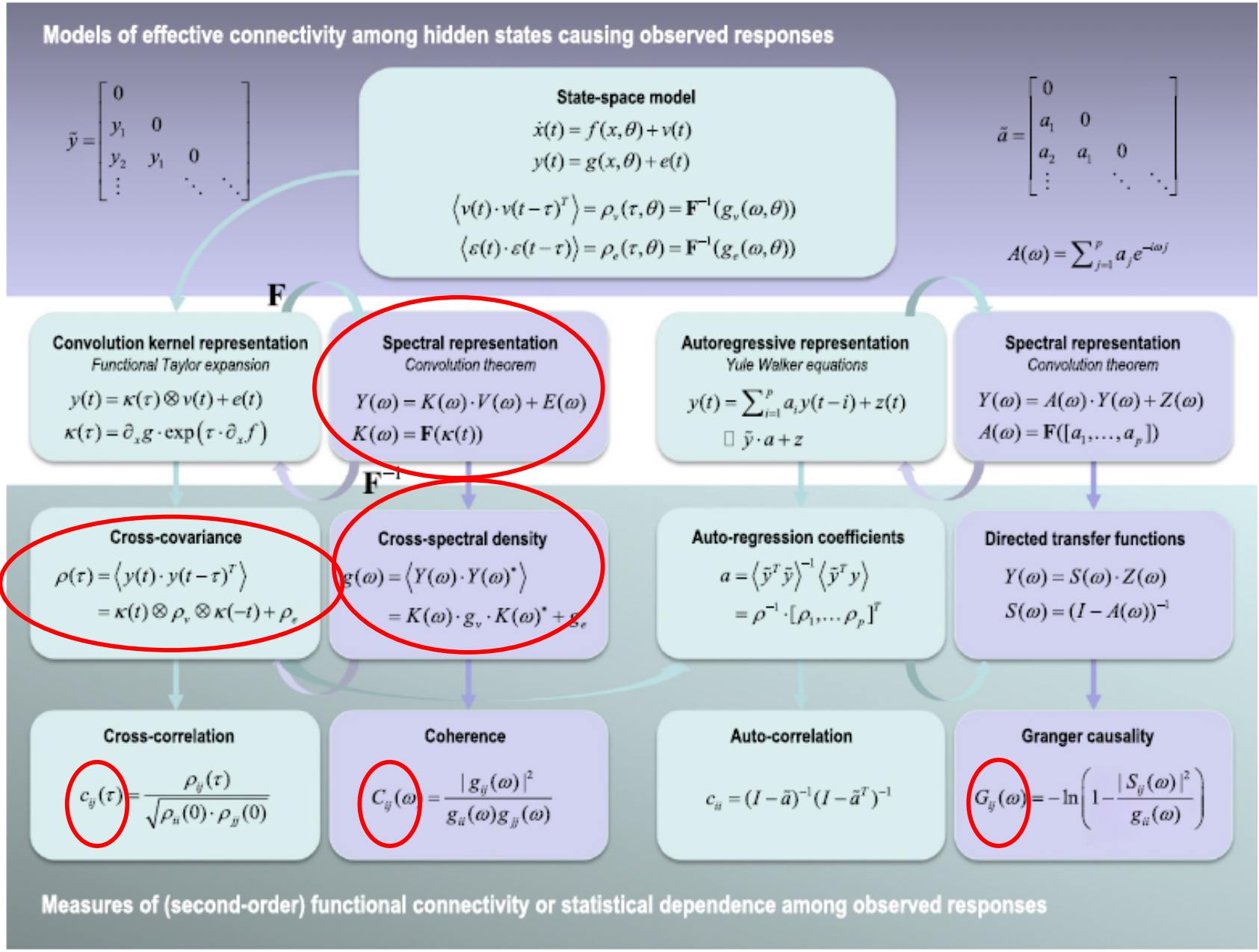
- <u>Alternative Approach to Stochastic Data</u>
- Examine 2nd order statistics in response to noise input
- deterministic model that generates predicted crossed spectra in a distributed neuronal network or graph
- finds the *effective connectivity among hidden neuronal states that best explains the observed functional connectivity* among hemodynamic responses
- advantage:
  - replaces an optimisation problem wrt. stochastic differential equations with a deterministic approach from linear systems theory → computationally very efficient
- disadvantages:
  - assumes stationarity (as do other rsfMRI methods)



Friston et al. 2014, *NeuroImage*

cross-spectra
= inverse Fourier transform of cross-correlation

cross-correlation
= generalized form of correlation (at zero lag, this is the conventional measure of functional connectivity)



Overview

- Bayesian Model Selection

i. Machinery

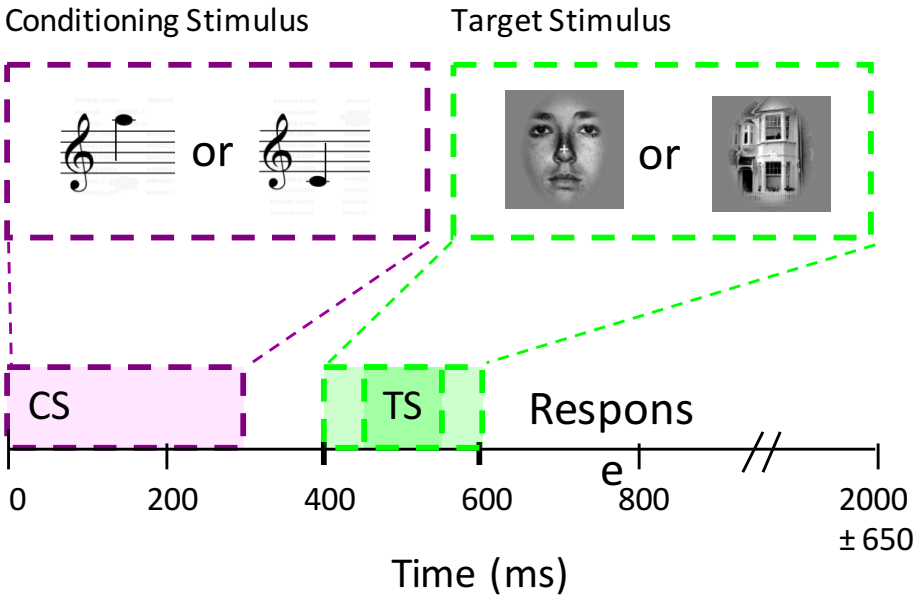
ii. Statistics

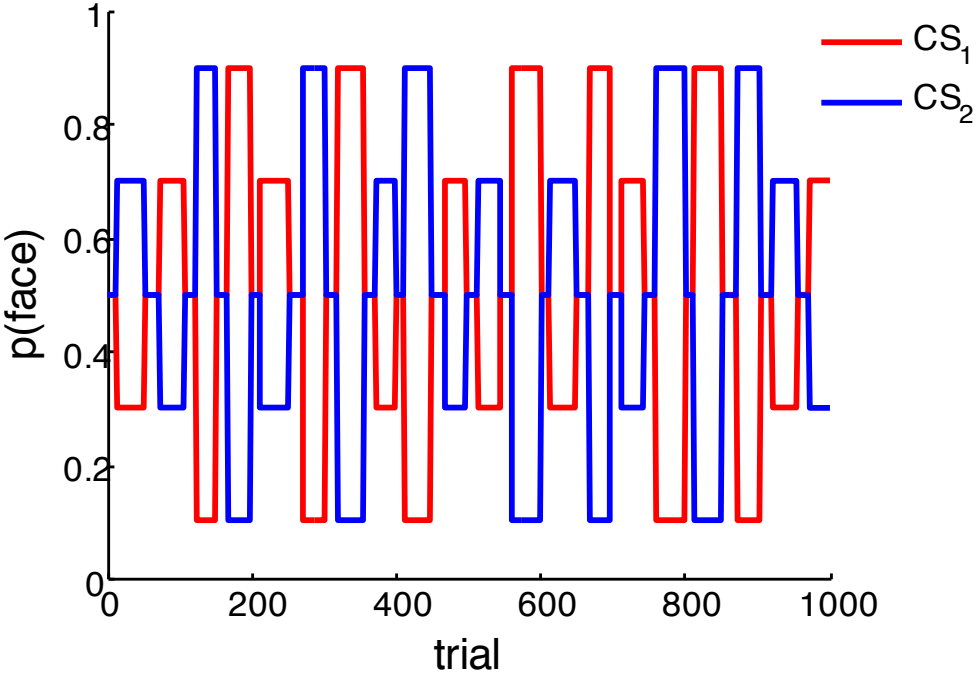
- Non-vanilla options in DCM for fMRI
- Models in Models

I. Computational Renderings of Behaviour in DCMS

II. Clinical Translational Applications

Learning of dynamic audio-visual associations





den Ouden et al. 2010, *J. Neurosci.*

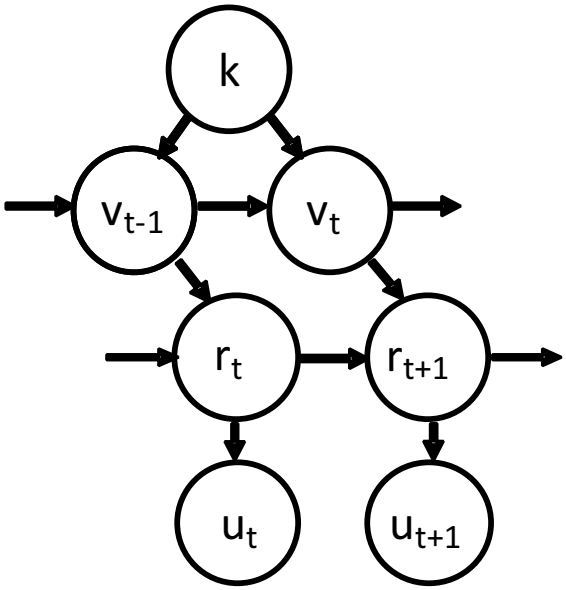
Hierarchical Bayesian learning model

prior on volatility

volatility

probabilistic association

observed events



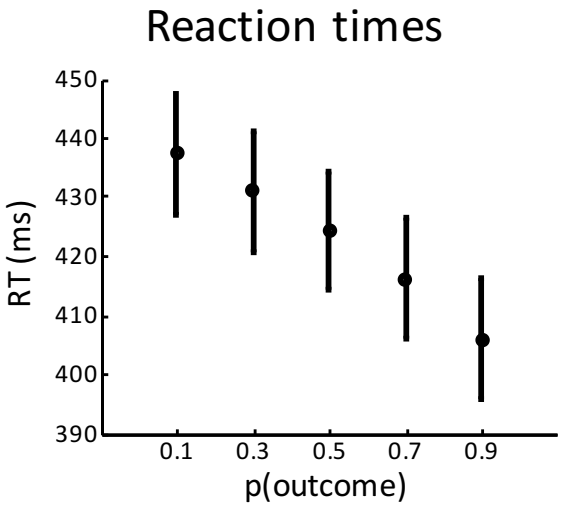
$$p(k) \propto 1$$

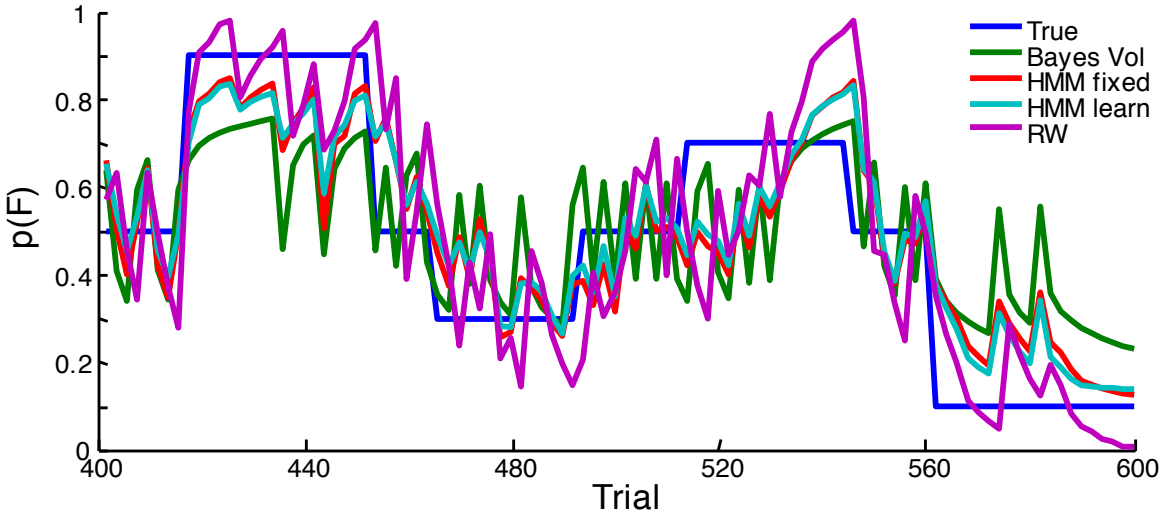
$$p(v_{t+1} \mid v_t, k) \sim N(v_t, \exp(k))$$

$$p(r_{t+1} \mid r_t, v_t) \sim Dir(r_t, \exp(v_t))$$

Behrens et al. 2007, *Nat. Neurosci.*

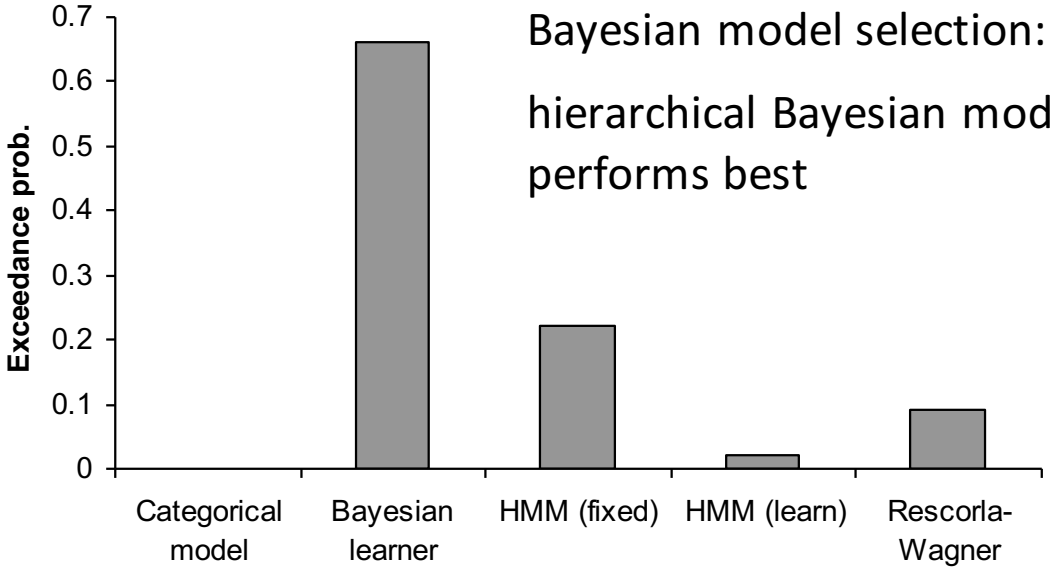
Explaining RTs by different learning models





5 alternative learning models:

- categorical probabilities
- hierarchical Bayesian learner
- Rescorla-Wagner
- Hidden Markov models (2 variants)



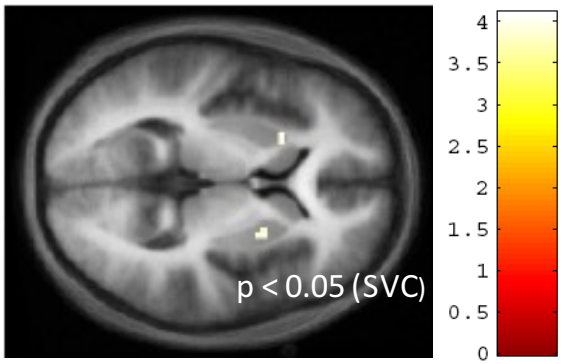
Bayesian model selection:

hierarchical Bayesian model performs best

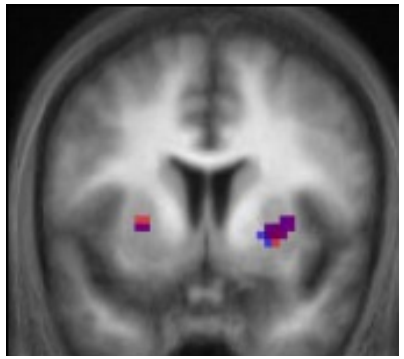
den Ouden et al. 2010, *J. Neurosci.*

Prediction error (PE) activity in the putamen

PE during active sensory learning

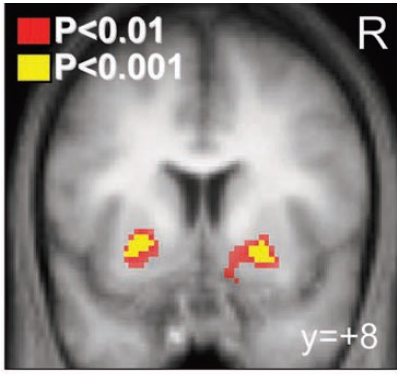


PE during incidental sensory learning



den Ouden et al. 2009, *Cerebral Cortex*

PE during reinforcement learning



O'Doherty et al. 2004, *Science*

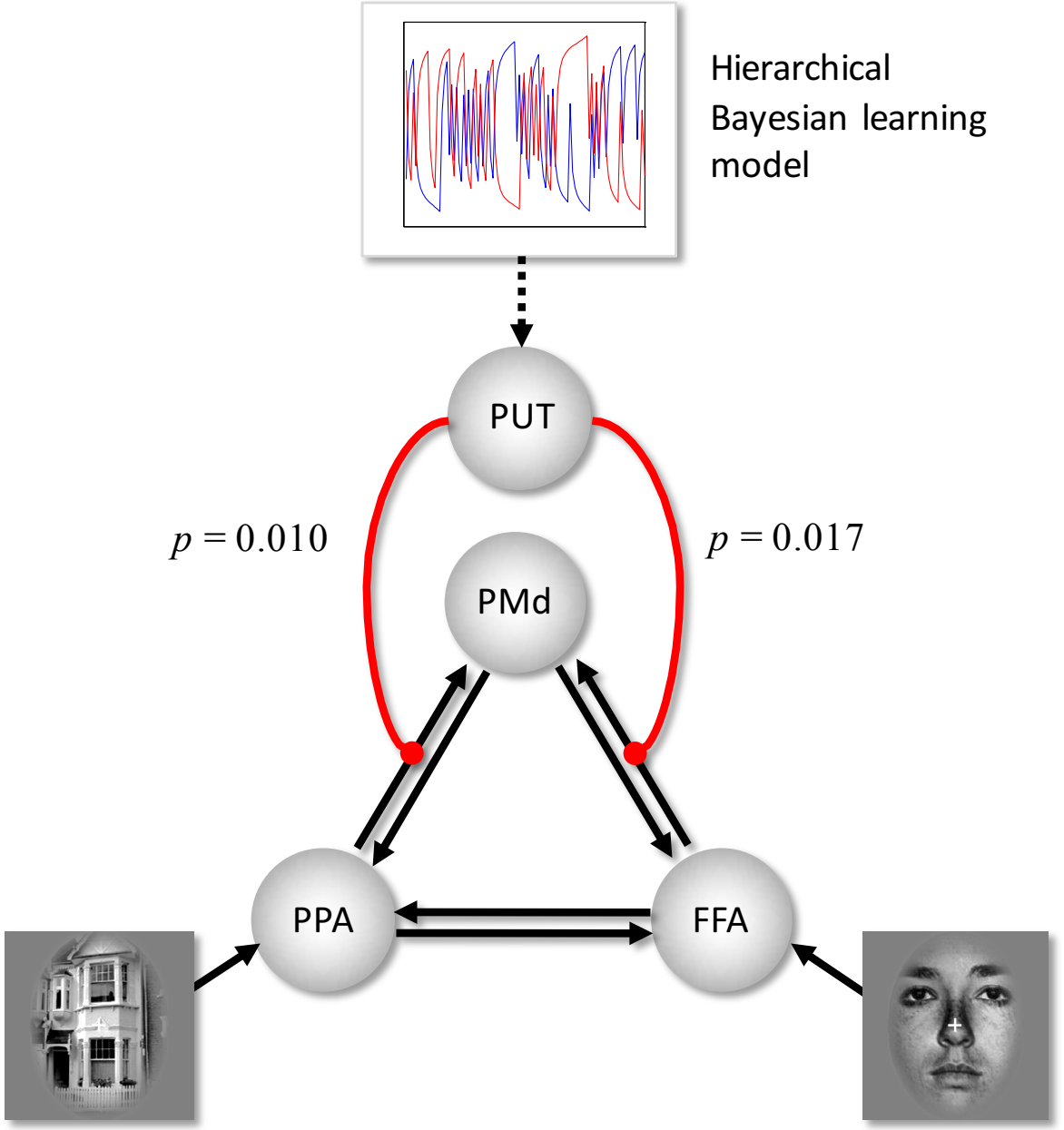
PE = "teaching signal" for synaptic plasticity during learning

Could the putamen be regulating trial-by-trial changes of task-relevant connections?

Prediction errors control plasticity during adaptive cognition

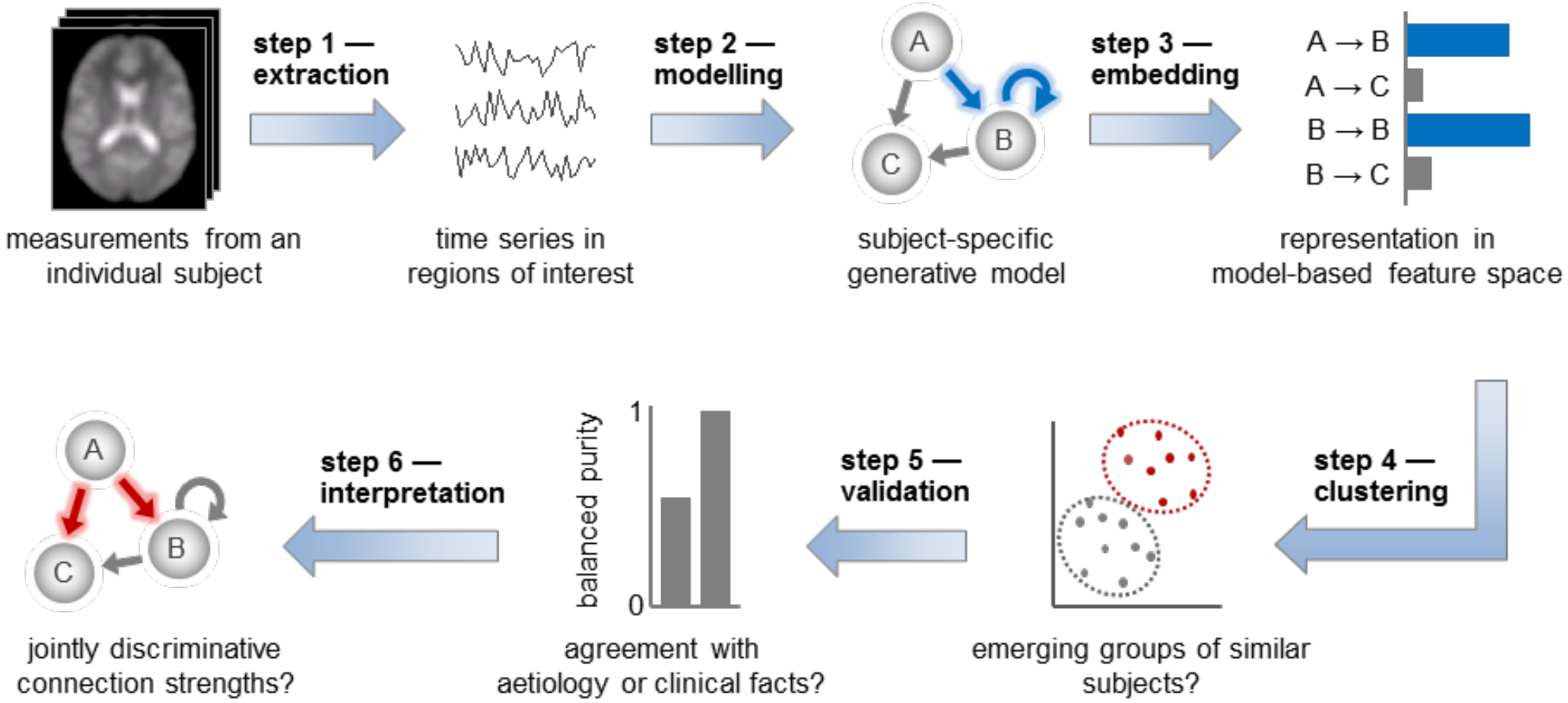
- Modulation of visuo-motor connections by striatal prediction error activity

- Influence of visual areas on premotor cortex:
  - stronger for surprising stimuli
  - weaker for expected stimuli



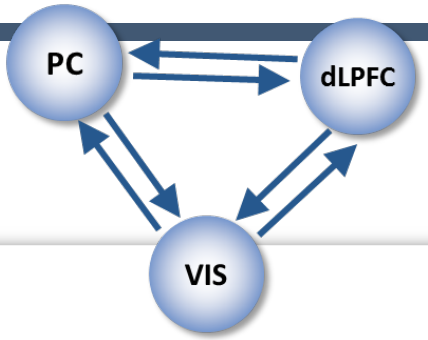
den Ouden et al. 2010, *J. Neurosci* .

Generative embedding (unsupervised): detecting patient subgroups

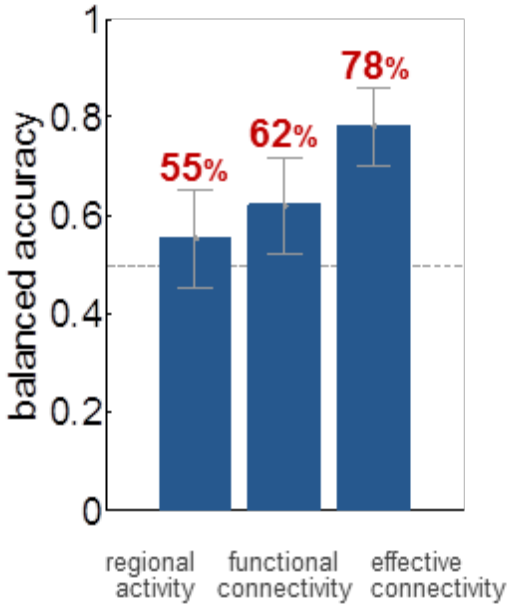


Brodersen et al. 2014, *NeuroImage: Clinical*

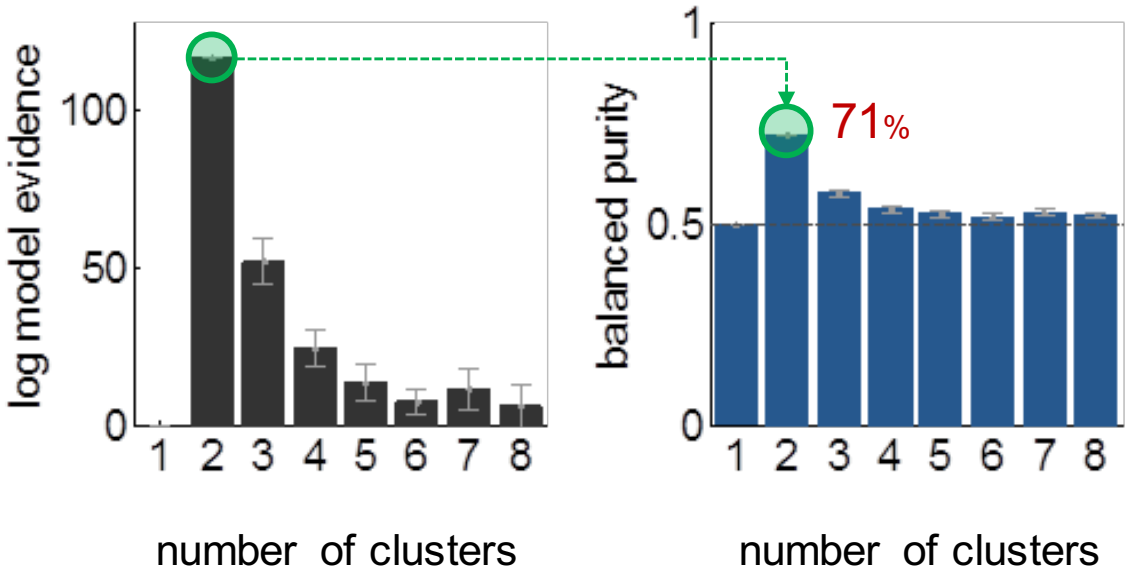
Generative embedding: DCM and variational Gaussian Mixture Models



Supervised:
SVM classification



Unsupervised:
GMM clustering

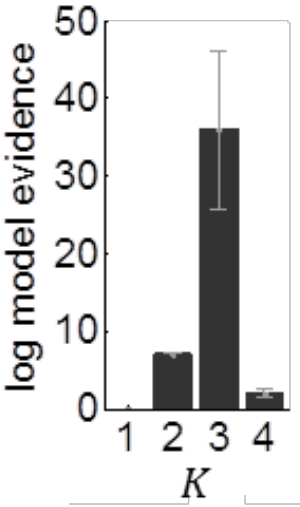


- 42 controls vs. 41 schizophrenic patients
- fMRI data from working memory task (Deserno et al. 2012, J. Neurosci)

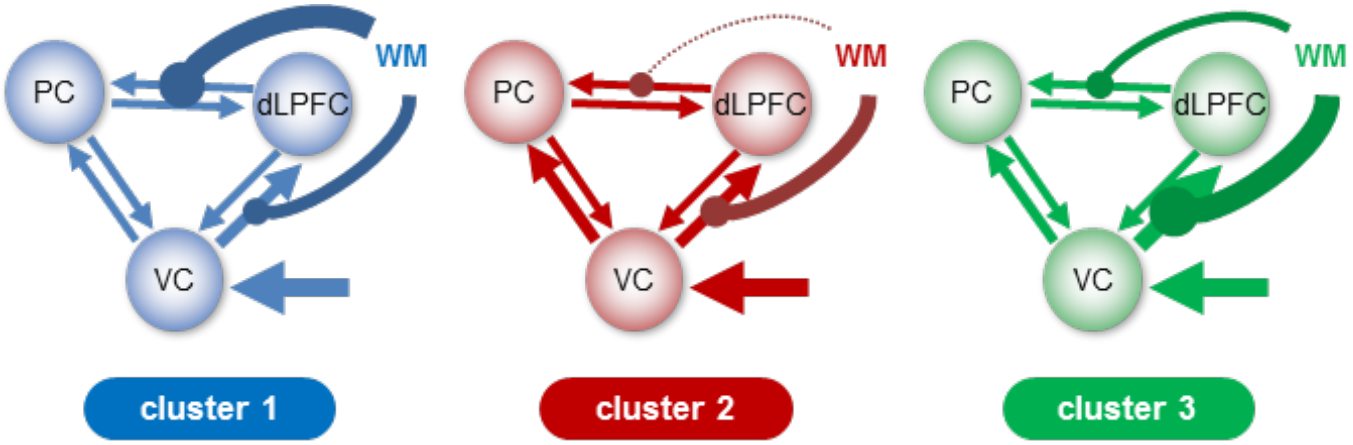
Brodersen et al. (2014) *NeuroImage: Clinical*

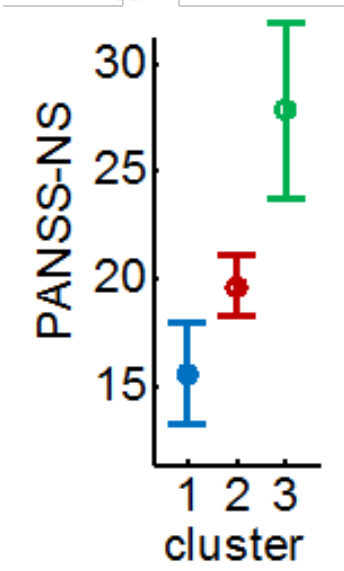
Detecting subgroups of patients in schizophrenia

- three distinct subgroups (total N=41)
- subgroups differ ($p < 0.05$) wrt. negative symptoms on the *positive and negative symptom scale* (PANSS)



Optimal cluster solution





Brodersen et al. (2014) *NeuroImage: Clinical*

Key methods papers: Advanced DCM for fMRI and BMS – part 1

- Brodersen KH, Schofield TM, Leff AP, Ong CS, Lomakina EI, Buhmann JM, Stephan KE (2011) Generative embedding for model-based classification of fMRI data. PLoS Computational Biology 7: e1002079.
- Daunizeau J, David, O, Stephan KE (2011) Dynamic Causal Modelling: A critical review of the biophysical and statistical foundations. NeuroImage 58: 312-322.
- Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. NeuroImage 19:1273-1302.
- Friston K, Stephan KE, Li B, Daunizeau J (2010) Generalised filtering. Mathematical Problems in Engineering 2010: 621670.
- Friston KJ, Li B, Daunizeau J, Stephan KE (2011) Network discovery with DCM. NeuroImage 56: 1202–1221.
- Friston K, Penny W (2011) Post hoc Bayesian model selection. Neuroimage 56: 2089-2099.
- Kasess CH, Stephan KE, Weissenbacher A, Pezawas L, Moser E, Windischberger C (2010) Multi-Subject Analyses with Dynamic Causal Modeling. NeuroImage 49: 3065-3074.
- Kiebel SJ, Kloppel S, Weiskopf N, Friston KJ (2007) Dynamic causal modeling: a generative model of slice timing in fMRI. NeuroImage 34:1487-1496.
- Li B, Daunizeau J, Stephan KE, Penny WD, Friston KJ (2011). Stochastic DCM and generalised filtering. NeuroImage 58: 442-457
- Marreiros AC, Kiebel SJ, Friston KJ (2008) Dynamic causal modelling for fMRI: a two-state model. NeuroImage 39:269-278.
- Penny WD, Stephan KE, Mechelli A, Friston KJ (2004a) Comparing dynamic causal models. NeuroImage 22:1157-1172.
- Penny WD, Stephan KE, Mechelli A, Friston KJ (2004b) Modelling functional integration: a comparison of structural equation and dynamic causal models. NeuroImage 23 Suppl 1:S264-274.

Key methods papers: Advanced DCM for fMRI and BMS – part 2

- Penny WD, Stephan KE, Daunizeau J, Joao M, Friston K, Schofield T, Leff AP (2010) Comparing Families of Dynamic Causal Models. PLoS Computational Biology 6: e1000709.
- Penny WD (2012) Comparing dynamic causal models using AIC, BIC and free energy. Neuroimage 59: 319-330.
- Stephan KE, Harrison LM, Penny WD, Friston KJ (2004) Biophysical models of fMRI responses. Curr Opin Neurobiol 14:629-635.
- Stephan KE, Weiskopf N, Drysdale PM, Robinson PA, Friston KJ (2007) Comparing hemodynamic models with DCM. NeuroImage 38:387-401.
- Stephan KE, Harrison LM, Kiebel SJ, David O, Penny WD, Friston KJ (2007) Dynamic causal models of neural system dynamics: current state and future extensions. J Biosci 32:129-144.
- Stephan KE, Weiskopf N, Drysdale PM, Robinson PA, Friston KJ (2007) Comparing hemodynamic models with DCM. NeuroImage 38:387-401.
- Stephan KE, Kasper L, Harrison LM, Daunizeau J, den Ouden HE, Breakspear M, Friston KJ (2008) Nonlinear dynamic causal models for fMRI. NeuroImage 42:649-662.
- Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009a) Bayesian model selection for group studies. NeuroImage 46:1004-1017.
- Stephan KE, Tittgemeyer M, Knösche TR, Moran RJ, Friston KJ (2009b) Tractography-based priors for dynamic causal models. NeuroImage 47: 1628-1638.
- Stephan KE, Penny WD, Moran RJ, den Ouden HEM, Daunizeau J, Friston KJ (2010) Ten simple rules for Dynamic Causal Modelling. NeuroImage 49: 3099-3109.

Thank you and the FIL Methods Group

& in particular

Klaas Enno Stephan
Hanneke den Ouden