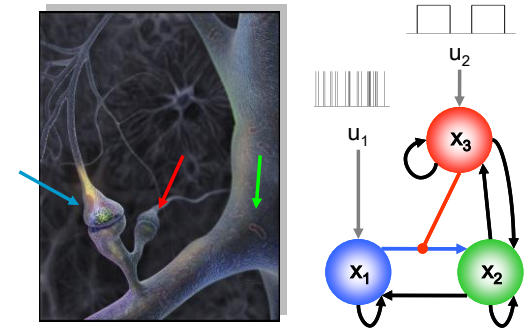
Dynamic Causal Modeling (DCM)

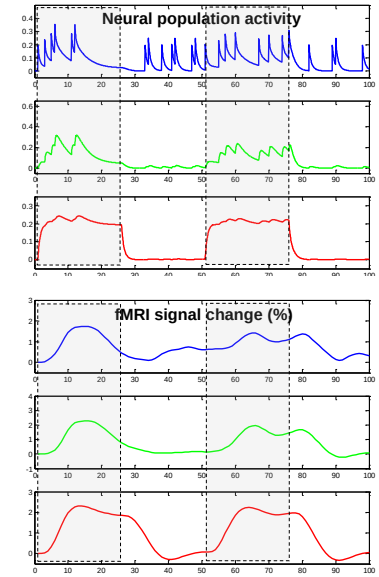
Klaas Enno Stephan

Translational Neuromodeling Unit (TNU)
Institute for Biomedical Engineering
University of Zurich & ETH Zurich

Wellcome Trust Centre for Neuroimaging
Institute of Neurology
University College London

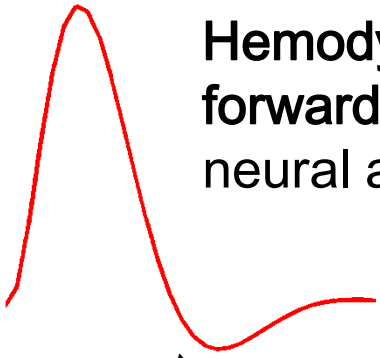


$$\frac{dx}{dt} = \left(A + \sum_{i=1}^{m} u_i B^{(i)} + \sum_{j=1}^{n} x_j D^{(j)} \right) x + Cu$$



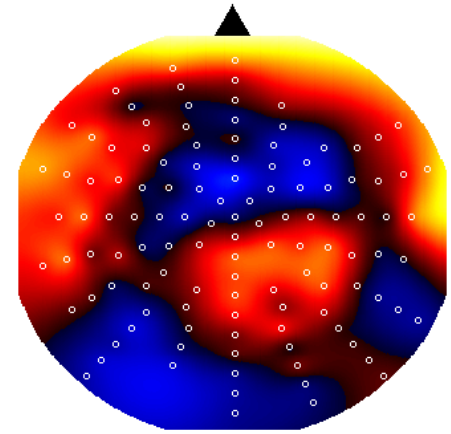
MPS-UCL Symposium
September 2012

Dynamic Causal Modeling (DCM)



Hemodynamic forward model:
neural activity→BOLD

Electromagnetic forward model:
neural activity→EEG
MEG
LFP



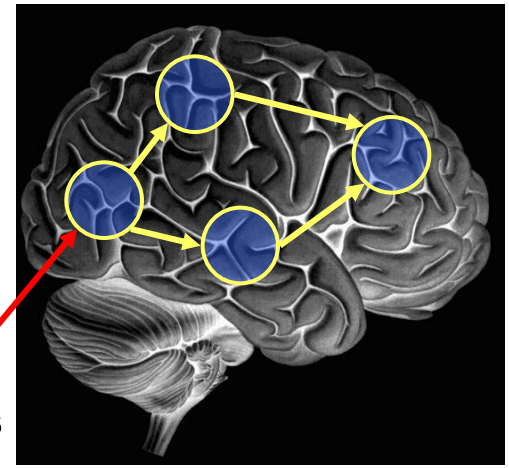
Neural state equation:

$$\frac{dx}{dt} = F(x,u,\theta)$$

fMRI

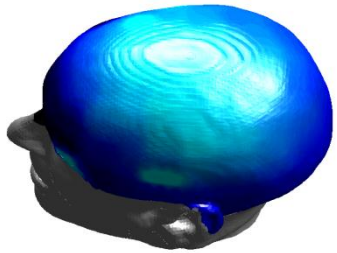
EEG/MEG

simple neuronal model
complicated forward model



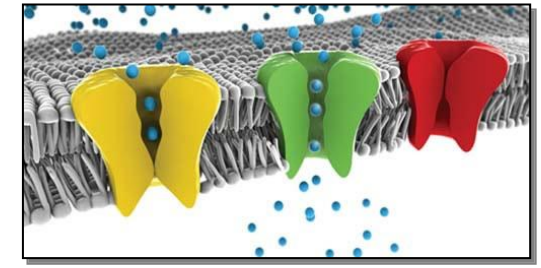
complicated neuronal model
simple forward model

A mathematical microscope

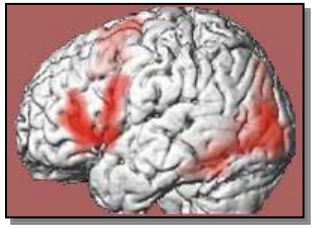


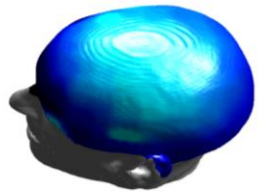
$$p(y \mid \theta, m)$$

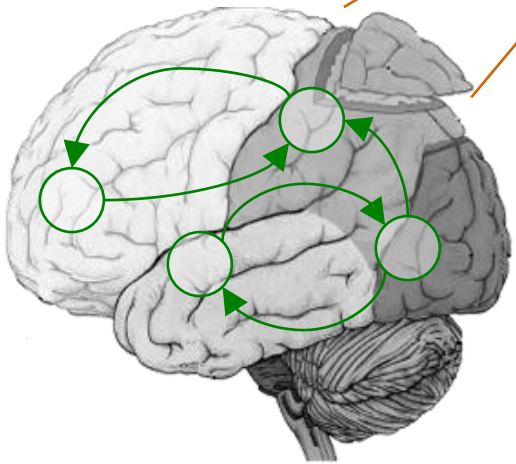
$$p(\theta \mid y, m)$$



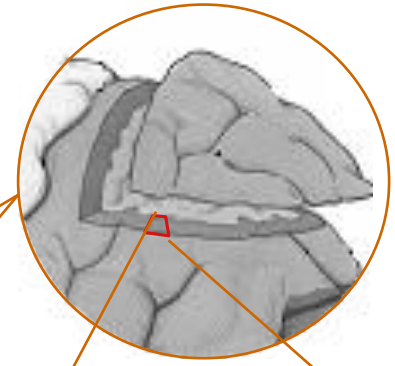
macro-scale

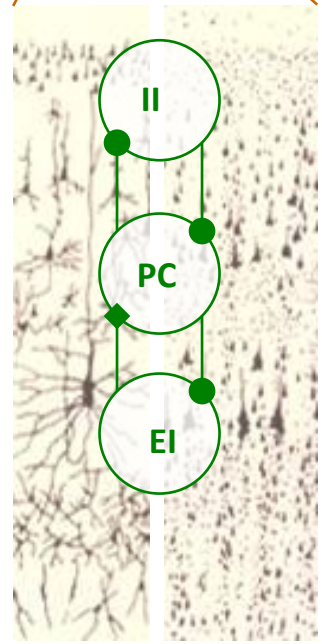






meso-scale



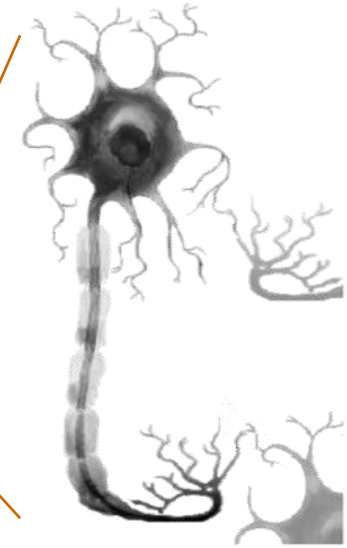


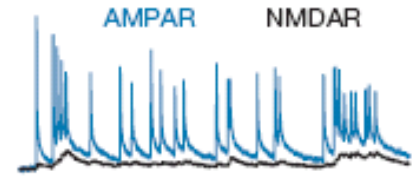
inhibitory interneurons

pyramidal cells

spiny stellate cells

micro-scale





Bayesian system identification

Neural dynamics

$$dx/dt = f(x,u,\theta)$$

Observer function

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

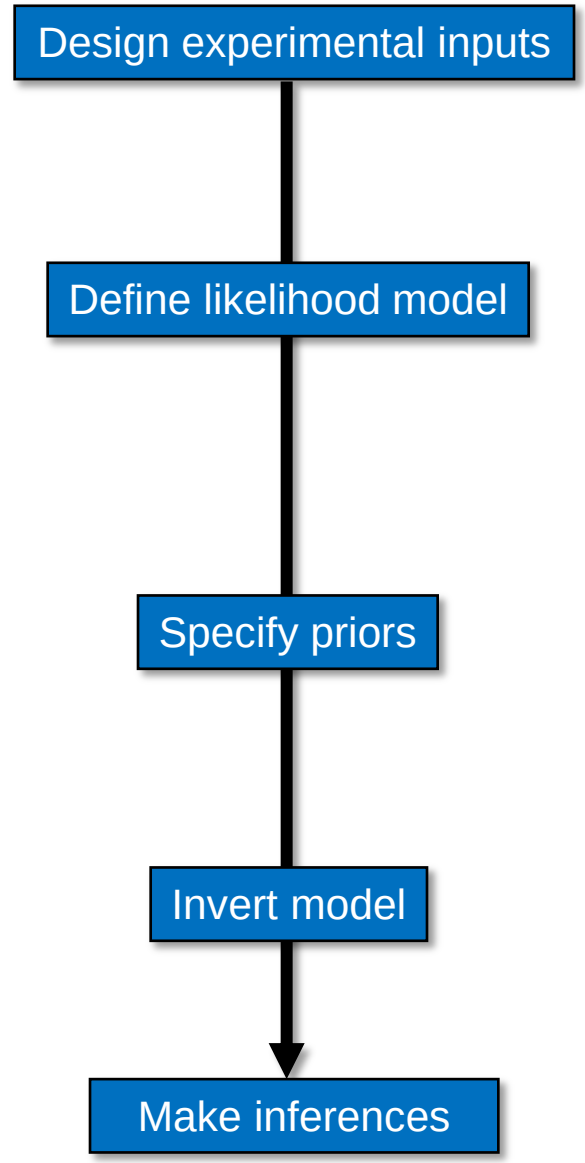
$$p(y \mid \theta, m) = N(g(\theta), \Sigma(\theta))$$
$$p(\theta, m) = N(\mu_\theta, \Sigma_\theta)$$

Inference on model structure

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

Inference on parameters

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



VB in a nutshell (mean-field approximation)

❶ Neg. free-energy approx. to model evidence.

$$\ln p(y \mid m) = F + KL\left[q(\theta,\lambda), p(\theta,\lambda \mid y)\right]$$

$$F = \left\langle \ln p(y,\theta,\lambda) \right\rangle_q - KL\left[q(\theta,\lambda), p(\theta,\lambda \mid m)\right]$$

❷ Mean field approx.

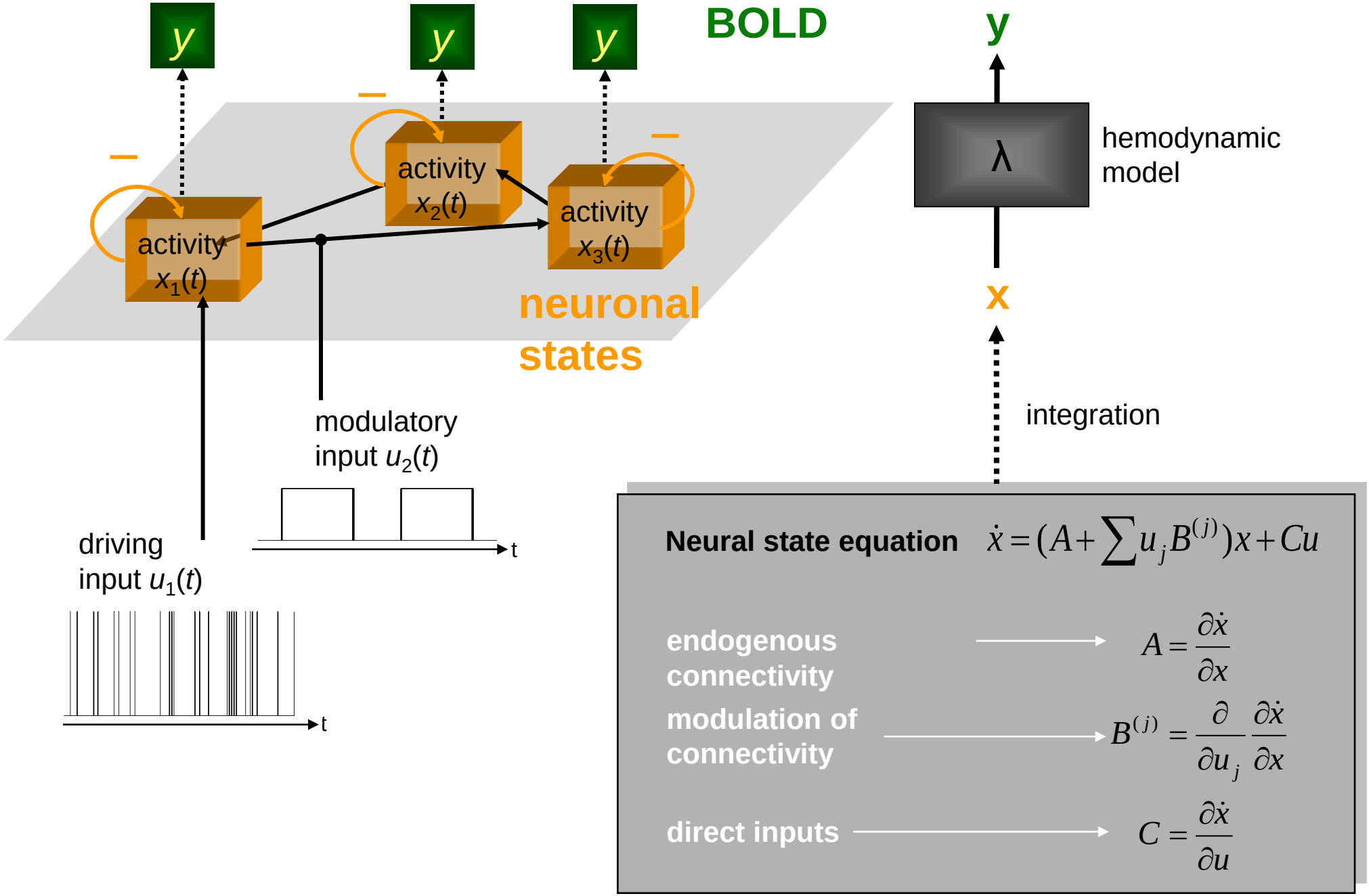
$$p(\theta,\lambda \mid y) \approx q(\theta,\lambda) = q(\theta)q(\lambda)$$

❸ Maximise neg. free energy wrt. q = minimise divergence, by maximising variational energies

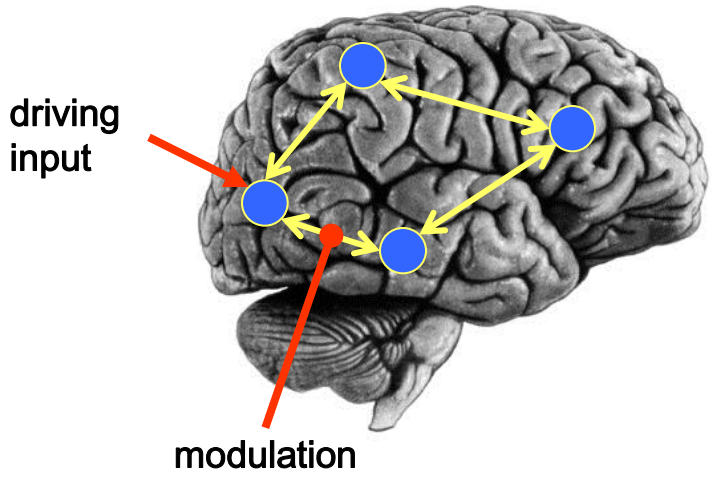
$$q(\theta) \propto \exp(I_\theta) = \exp\left[\left\langle \ln p(y,\theta,\lambda) \right\rangle_{q(\lambda)}\right]$$

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

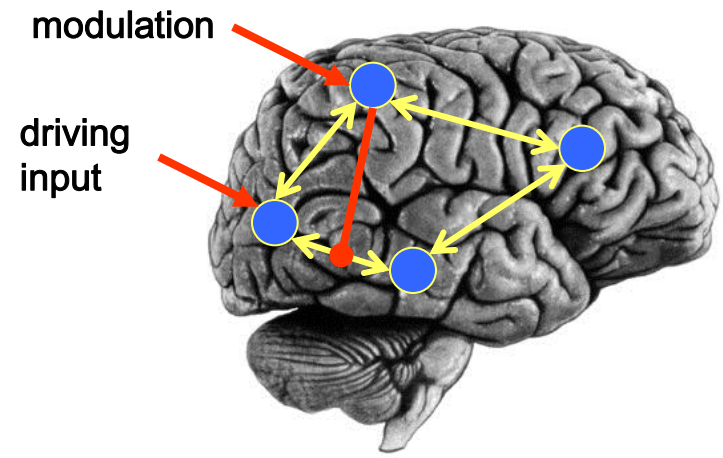
❹ Iterative updating of sufficient statistics of approx. posteriors by gradient ascent.


y
y
y
BOLD
activity
$x_1(t)$
activity
$x_2(t)$
activity
$x_3(t)$
neuronal
states
y
λ
hemodynamic
model
x
integration
modulatory
input $u_2(t)$
t
driving
input $u_1(t)$
t
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
$C = \frac{\partial \dot{x}}{\partial u}$

bilinear DCM



non-linear DCM



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

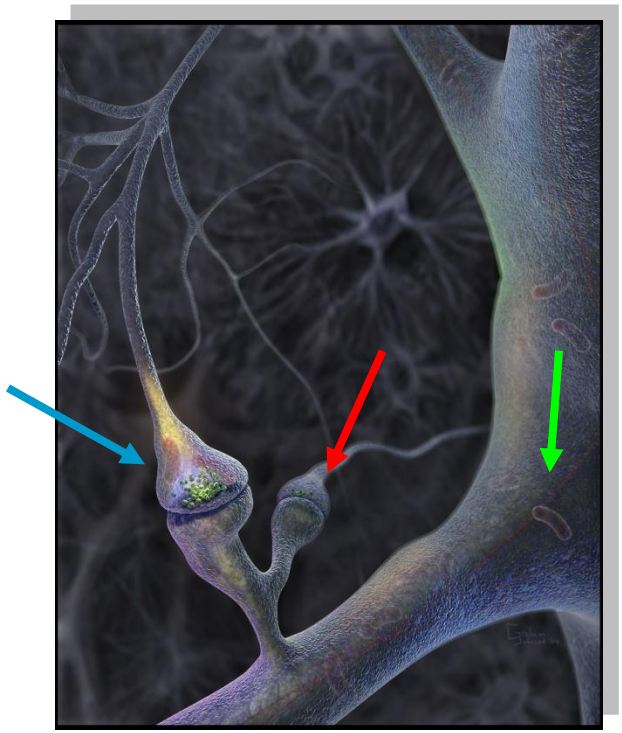
$$\frac{dx}{dt} = f(x,u) \approx f(x_0,0) + \frac{\partial f}{\partial x}x + \frac{\partial f}{\partial u}u + \frac{\partial^2 f}{\partial x \partial u}ux + \frac{\partial^2 f}{\partial x^2}\frac{x^2}{2} + \ldots$$

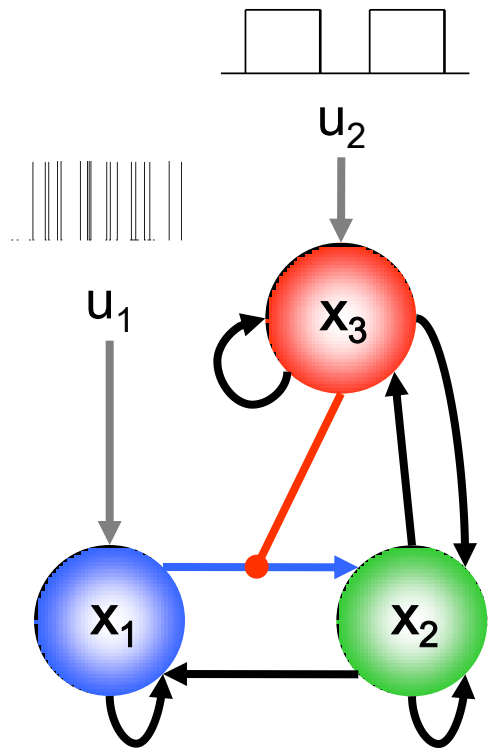
Bilinear state equation:

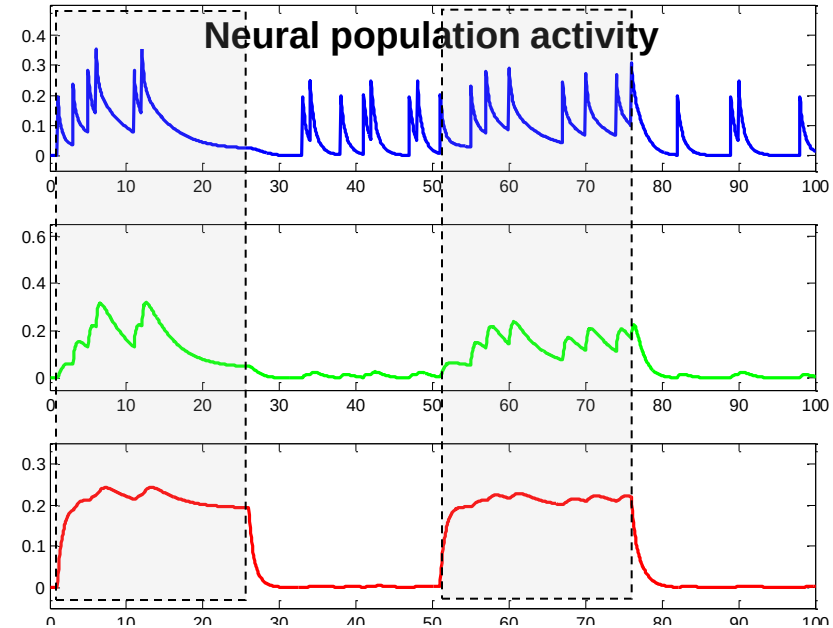
$$\frac{dx}{dt} = \left(A + \sum_{i=1}^{m} u_i B^{(i)}\right)x + Cu$$

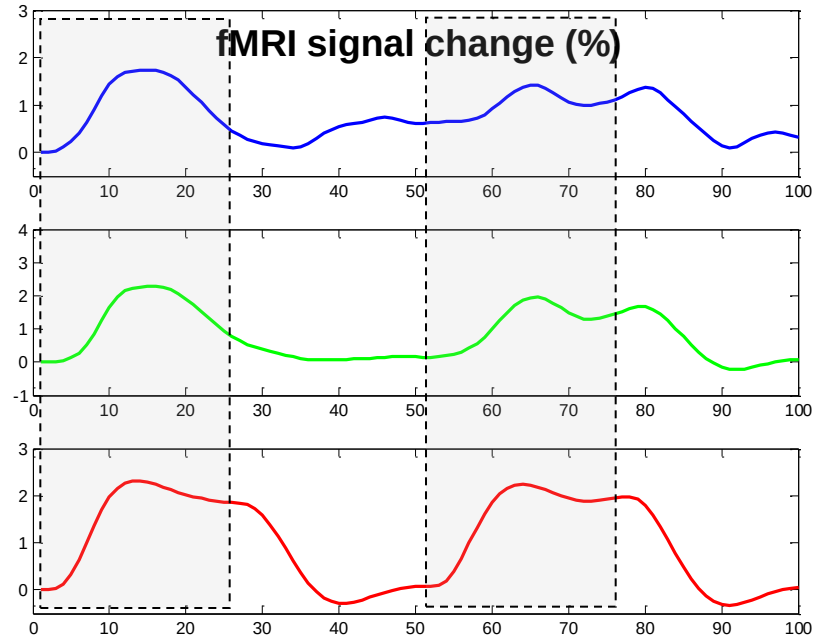
Nonlinear state equation:

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







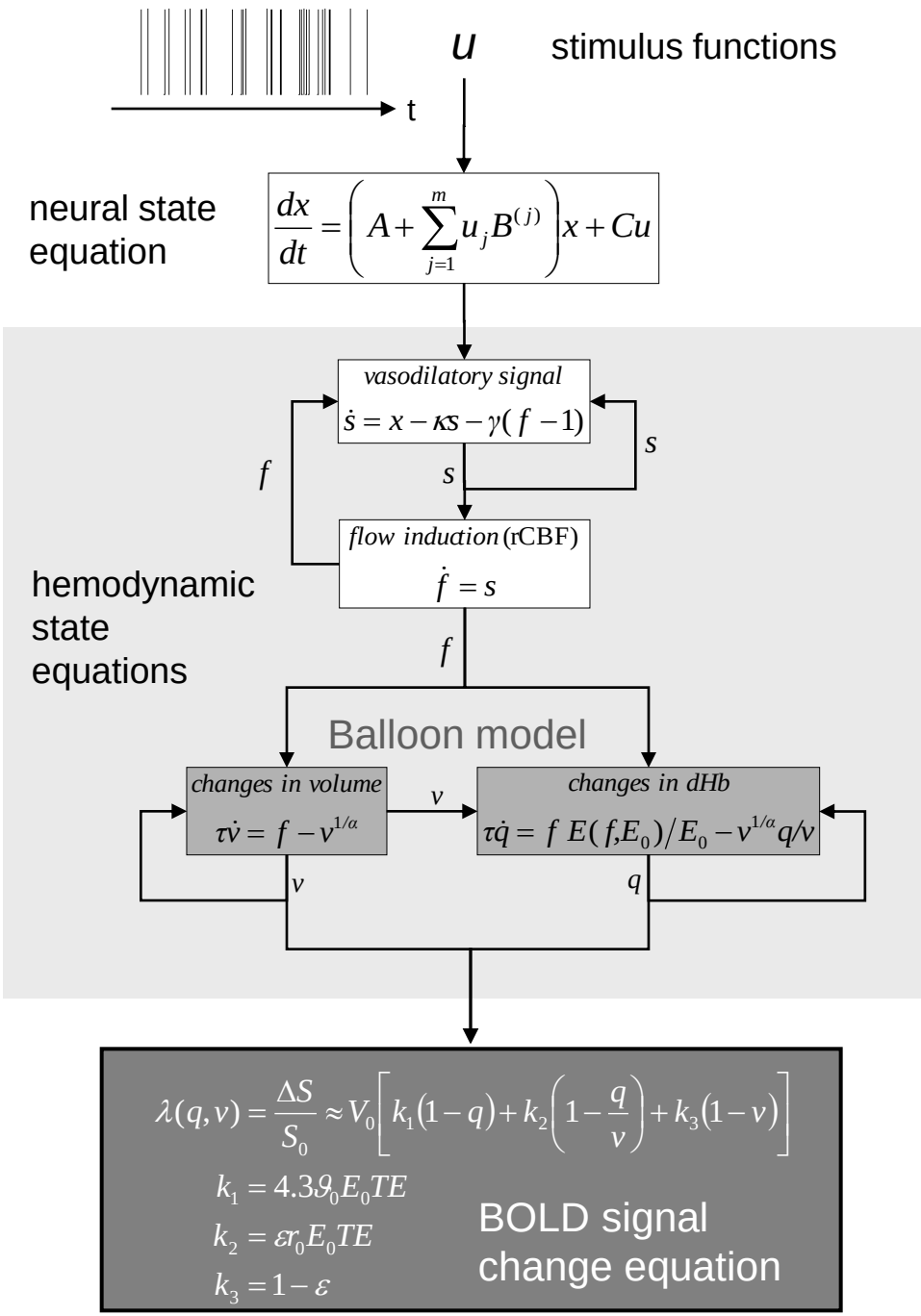


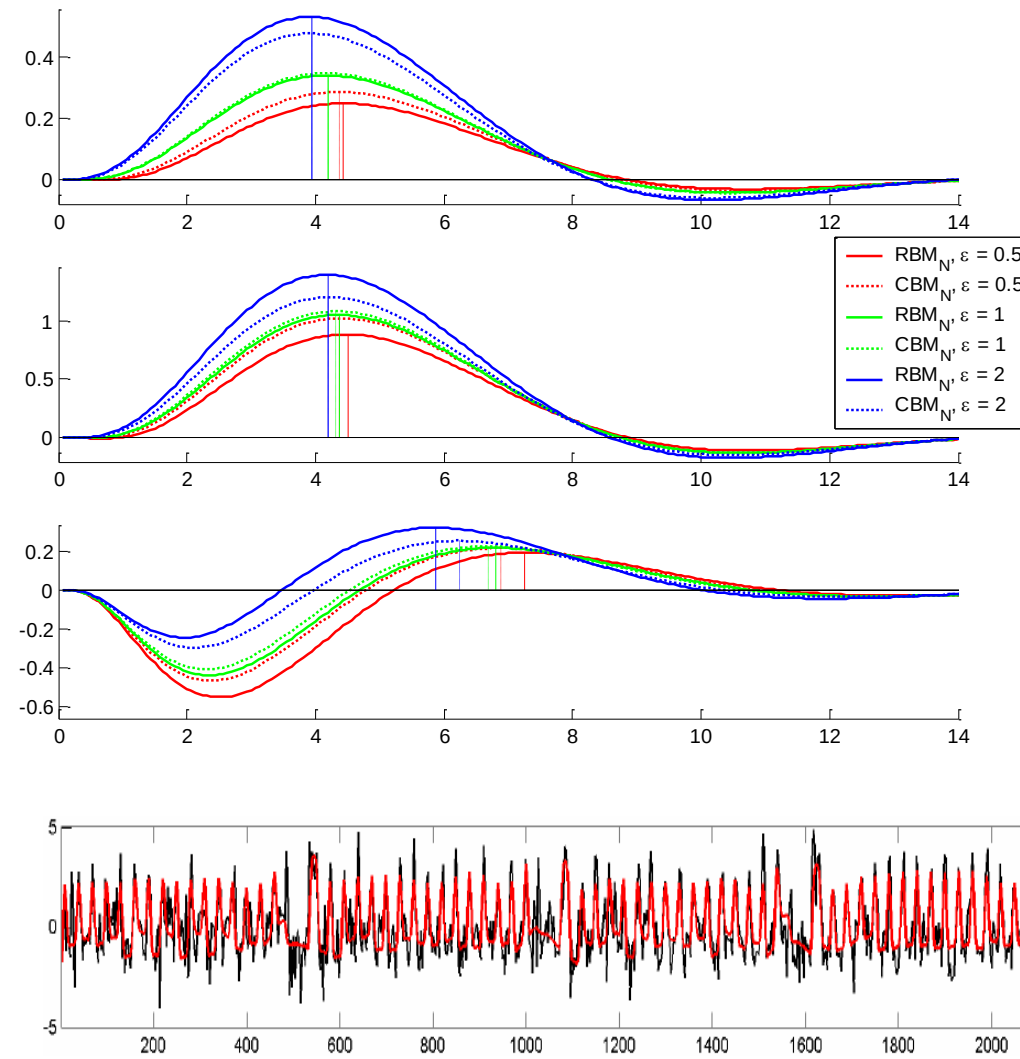
Nonlinear Dynamic Causal Model for fMRI

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

Stephan et al. 2008, *NeuroImage*

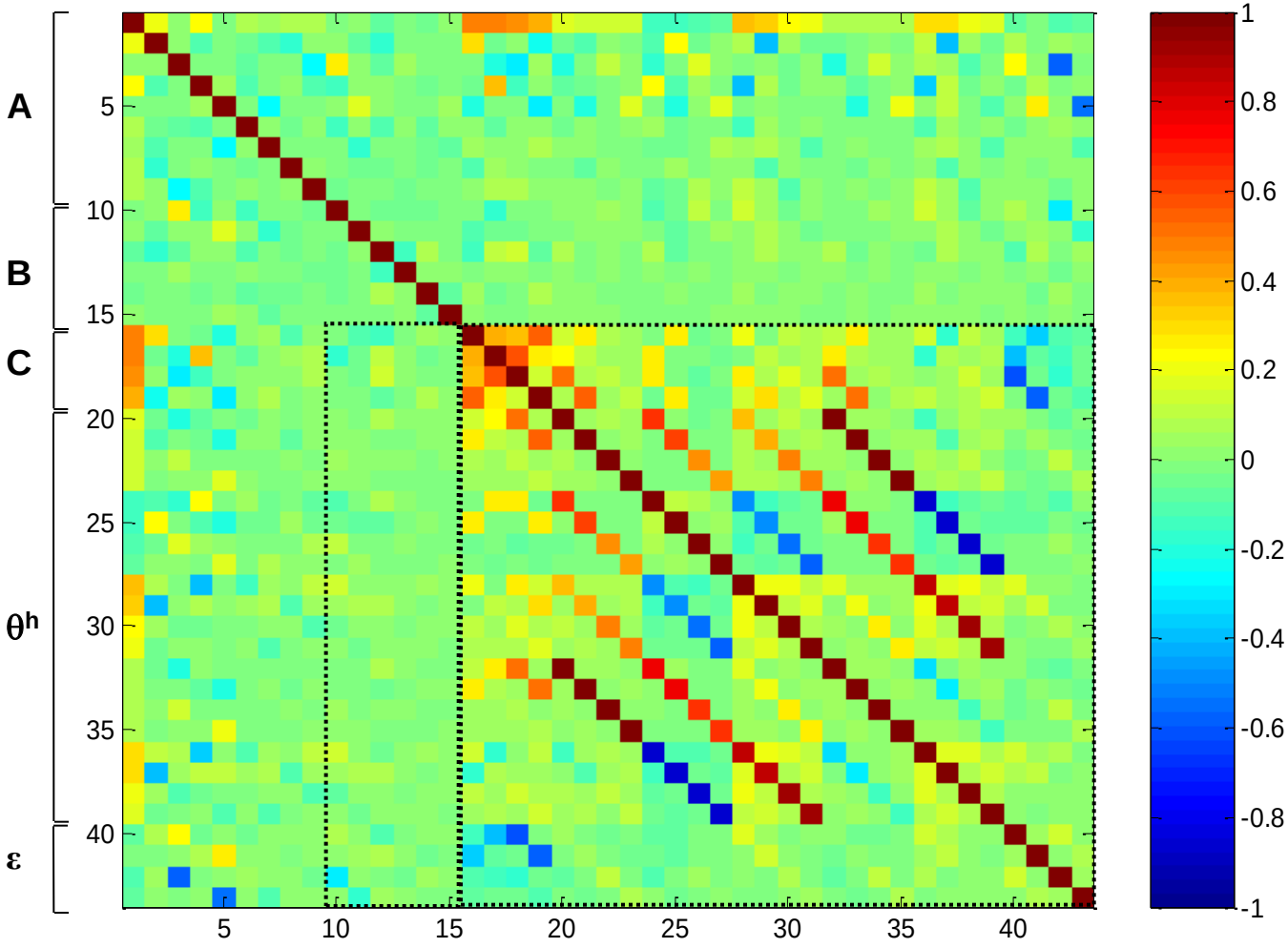
The hemodynamic model in DCM

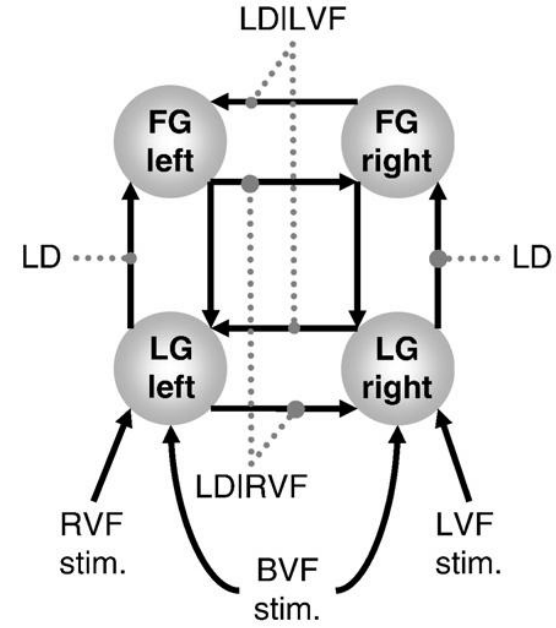




Stephan et al. 2007, *NeuroImage*

How interdependent are neural and hemodynamic parameter estimates?



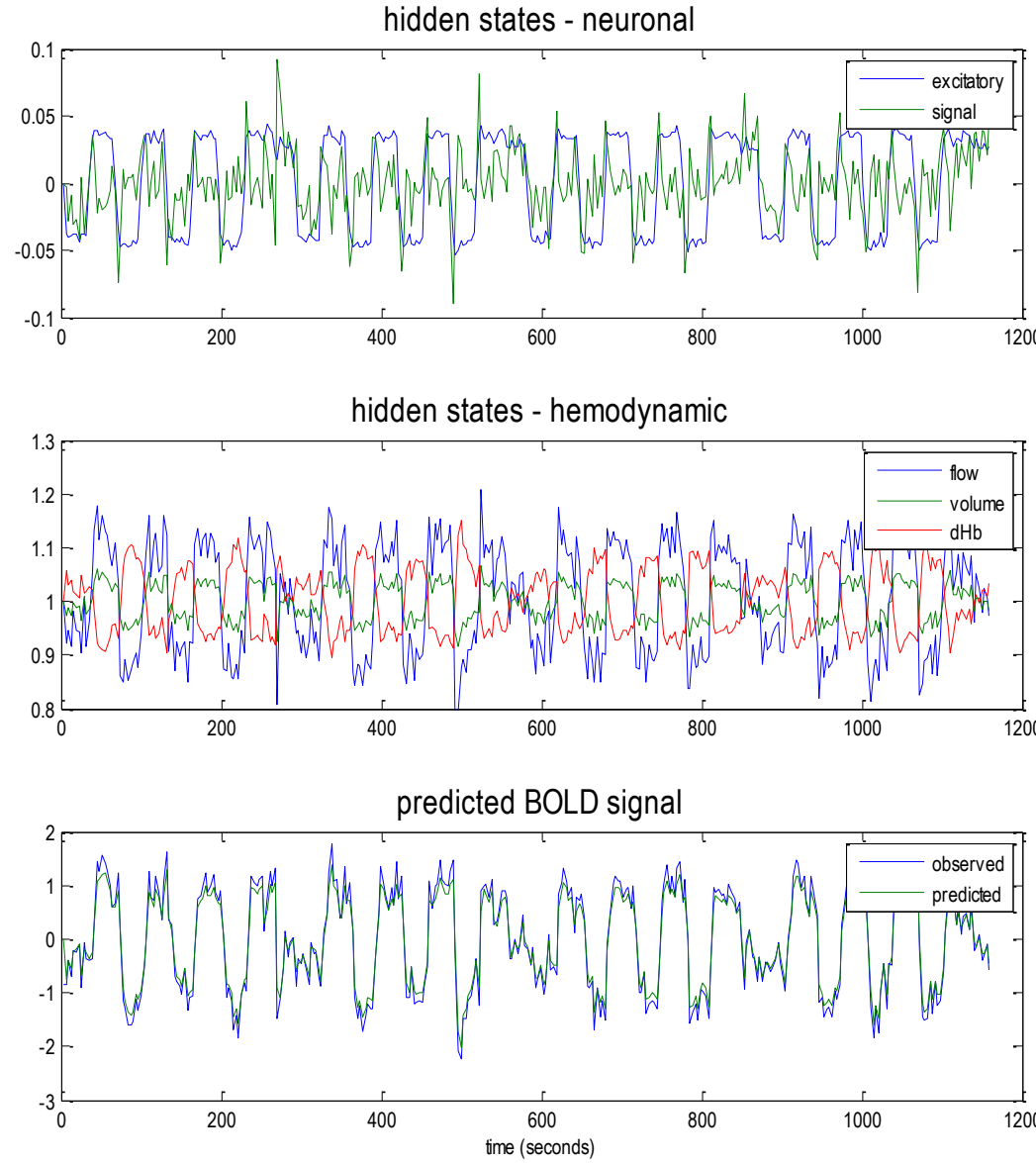


Stephan et al. 2007, *NeuroImage*

Stochastic DCM

$$\frac{dx}{dt} = f\left(x, u, \theta\right) + \omega$$

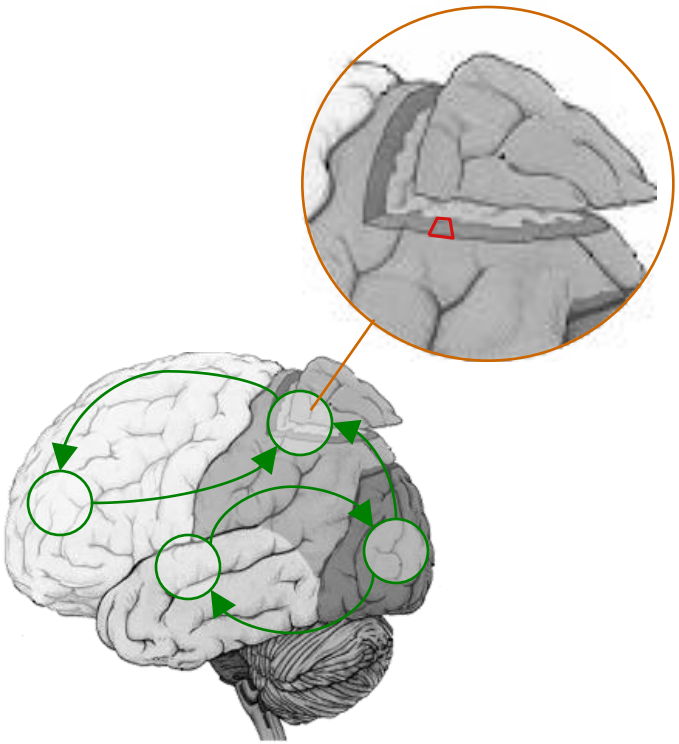
- accounts for stochastic neural fluctuations
- can be fitted to resting state data
- ω has unknown precision and smoothness
  $\rightarrow$ additional hyperparameters



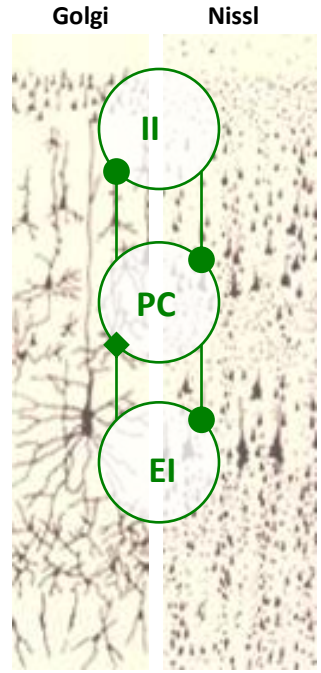
Friston et al. (2008, 2011) *NeuroImage*
Daunizeau et al. (2009) *Physica D*
Li et al. (2011) *NeuroImage*

DCM for EEG, MEG & local field potentials (LFPs)

macro-scale



meso-scale

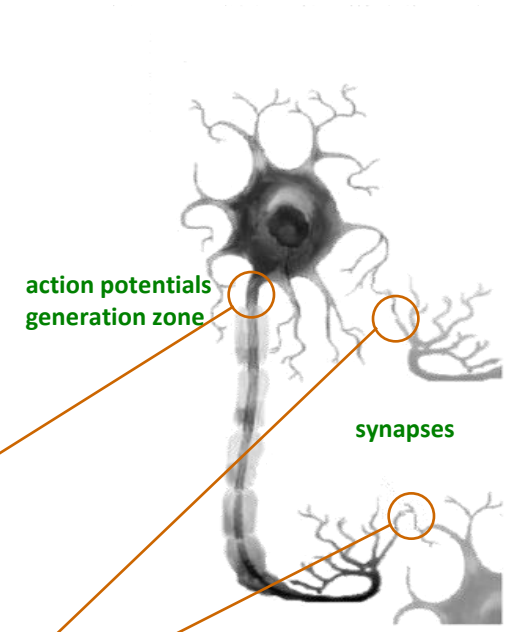


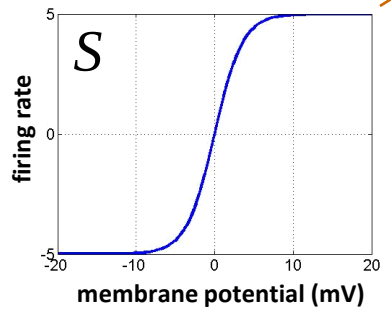
inhibitory interneurons

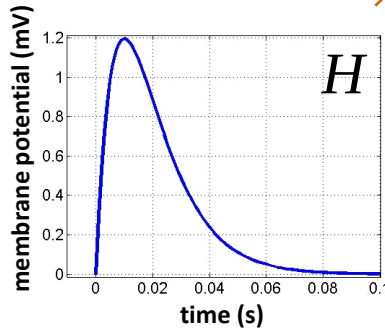
pyramidal cells

spiny stellate cells

micro-scale





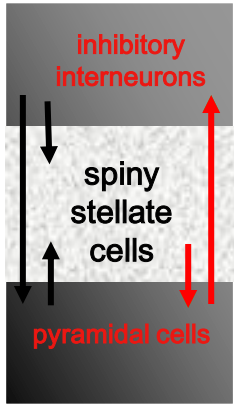


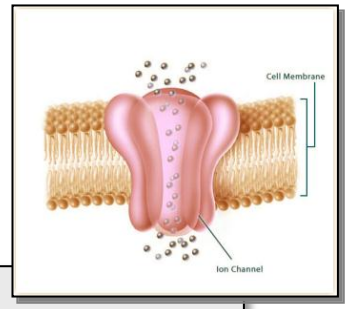
$$\frac{dx_1}{dt} = x_2$$

$$\frac{dx_2}{dt} = \kappa H\left[(A+\gamma)S(x_1)\right] - 2\kappa x_2 - \kappa^2 x_1$$

Jansen & Rit 1995, *Biol. Cybern.*
David et al. 2006, *NeuroImage*
Kiebel et al. 2007, *NeuroImage*
Moran et al. 2009, *NeuroImage*

Conductance-based DCMs





$$C\dot{V} = \sum g_i \left(V_i^0 - V_i\right)$$

$$C\dot{V}^{(2)} = g_L(V_L - V^{(2)}) + g_E^{(2)}(V_E - V^{(2)}) + g_{NMDA} f_{Mg}(V^{(2)})(V_E - V^{(2)}) + \Gamma_V$$
$$\dot{g}_E^{(2)} = \kappa_E(\gamma_{2,3}^{(E)}\sigma(\mu_V^{(3)} - V_R, \Sigma^{(3)}) - g_E^{(2)}) + \Gamma_E$$
$$\dot{g}_{NMDA}^{(2)} = \kappa_{NMDA}(\gamma_{2,3}^{(I)}\sigma(\mu_V^{(3)} - V_R, \Sigma^{(3)}) - g_{NMDA}^{(2)}) + \Gamma_{NMDA}$$

$\gamma_{1,2}^{(I)}$ $\gamma_{2,3}^{(E)}$

$$C\dot{V}^{(1)} = g_L(V_L - V^{(1)}) + g_E^{(1)}(V_E - V^{(1)}) + g_I^{(1)}(V_I - V^{(1)}) + u + \Gamma_V$$
$$\dot{g}_E^{(1)} = \kappa_E(\gamma_{1,3}^{(E)}\sigma(\mu_V^{(3)} - V_R, \Sigma^{(3)}) - g_E^{(1)}) + \Gamma_E$$
$$\dot{g}_I^{(1)} = \kappa_I(\gamma_{1,2}^{(E)}\sigma(\mu_V^{(2)} - V_R, \Sigma^{(2)}) - g_I^{(1)}) + \Gamma_I$$

$\gamma_{3,2}^{(I)}$ $\gamma_{1,3}^{(E)}$ $\gamma_{3,1}^{(E)}$

$$C\dot{V}^{(3)} = g_L(V_L - V^{(3)}) + g_E^{(3)}(V_E - V^{(3)}) + g_I^{(3)}(V_I - V^{(3)}) + g_{NMDA} f_{Mg}(V^{(3)})(V_E - V^{(3)}) + \Gamma_V$$
$$\dot{g}_E^{(3)} = \kappa_E(\gamma_{3,1}^{(E)}\sigma(\mu_V^{(1)} - V_R, \Sigma^{(1)}) - g_E^{(3)}) + \Gamma_E$$
$$\dot{g}_I^{(3)} = \kappa_I(\gamma_{3,2}^{(I)}\sigma(\mu_V^{(2)} - V_R, \Sigma^{(2)}) - g_I^{(3)}) + \Gamma_I$$
$$\dot{g}_{NMDA}^{(3)} = \kappa_{NMDA}(\gamma_{3,1}^{(I)}\sigma(\mu_V^{(1)} - V_R, \Sigma^{(1)}) - g_{NMDA}^{(3)}) + \Gamma_{NMDA}$$

Marreiros et al. 2010, *NeuroImage*
Moran et al. 2011, *NeuroImage*

Generative models, model selection & model validation

- any DCM = a particular generative model of how the data (may) have been caused

- modelling = comparing competing hypotheses about the mechanisms underlying the data

  $\rightarrow$ careful definition of model space (hypothesis space) is crucial

- model selection $\neq$ model validation!

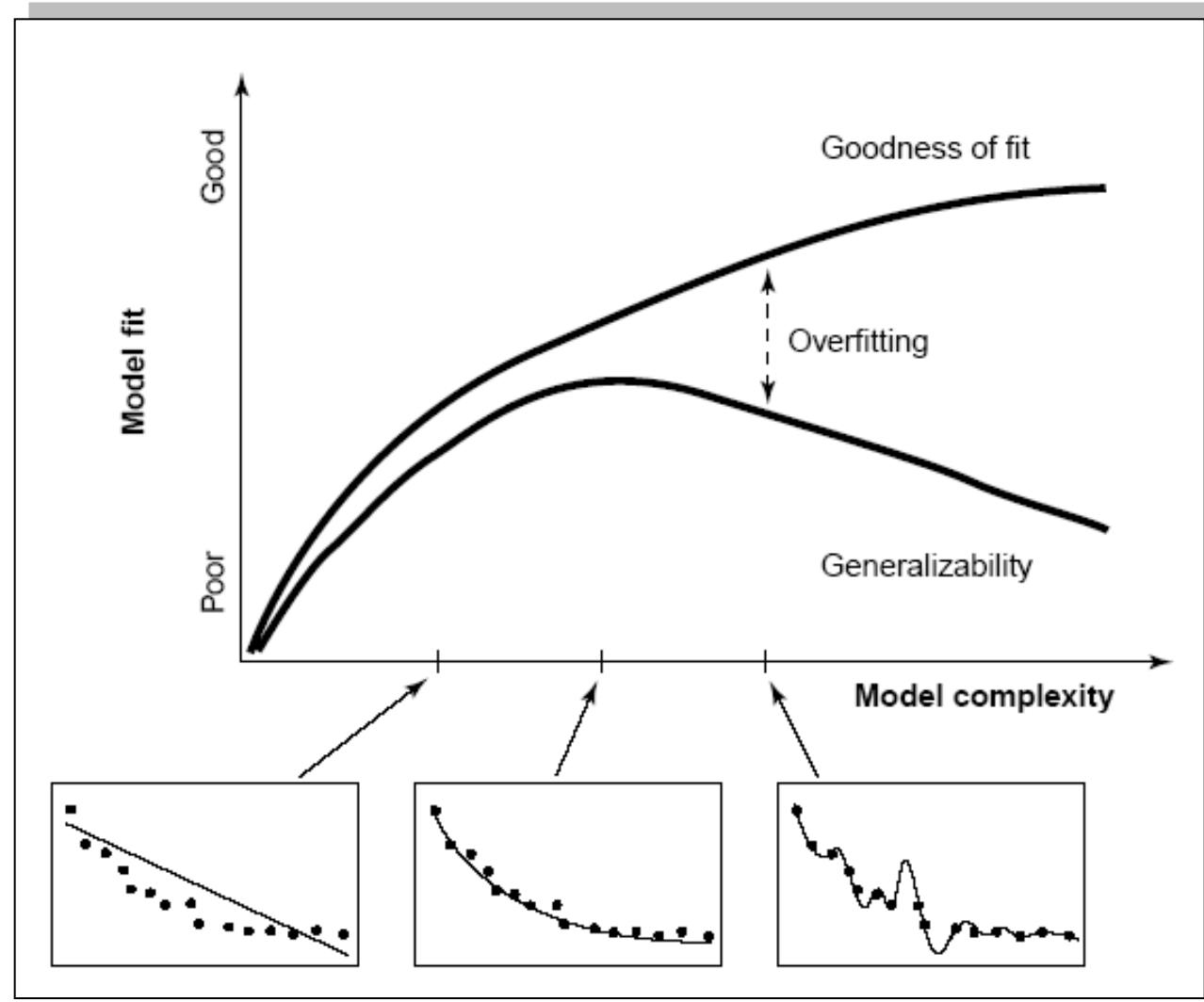
  $\rightarrow$ model validation requires external criteria (external to the measured data)

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*

Bayesian model selection (BMS)

Model evidence:

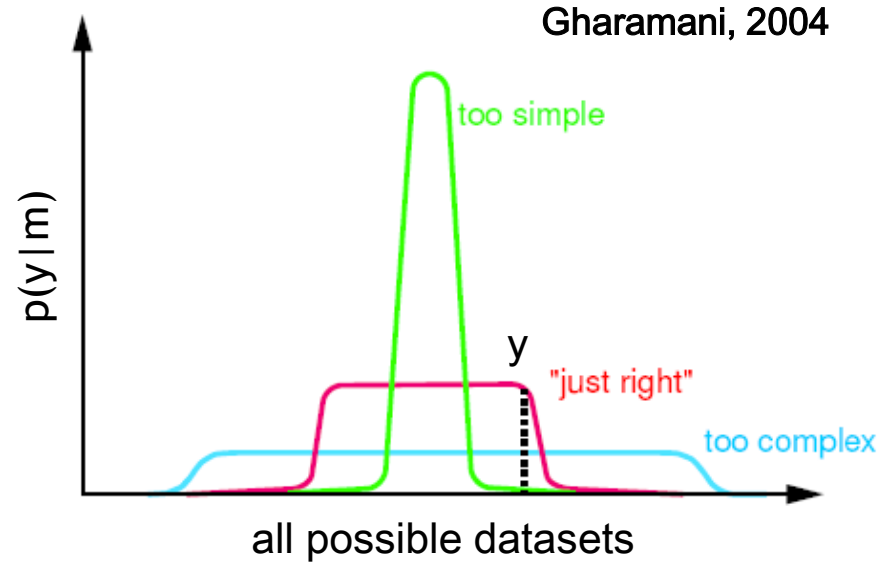
$$p(y|m) = \int p(y|\theta,m)\,p(\theta|m)\,d\theta$$

$$\log p(y|m) = \left\langle \log p(y|\theta,m) \right\rangle - KL\left[q(\theta), p(\theta|m)\right] + KL\left[q(\theta), p(\theta|y,m)\right]$$

→ accounts for both accuracy and complexity of the model

→ a measure of generalizability

Gharamani, 2004



Various approximations, e.g.:

- negative free energy, AIC, BIC

McKay 1992, *Neural Comput.*
Penny et al. 2004a, *NeuroImage*

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

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

In SPM2 & SPM5, interface offers 2 approximations:

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

→ AIC favours more complex models,
BIC favours simpler models.

Penny et al. 2004a, *NeuroImage*

VB in a nutshell (mean-field approximation)

❶ Neg. free-energy approx. to model evidence.

$$\ln p(y \mid m) = F + KL\left[q(\theta,\lambda), p(\theta,\lambda \mid y)\right]$$

$$F = \left\langle \ln p(y,\theta,\lambda) \right\rangle_q - KL\left[q(\theta,\lambda), p(\theta,\lambda \mid m)\right]$$

❷ Mean field approx.

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

❸ Maximise neg. free energy wrt. q = minimise divergence, by maximising variational energies

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

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

❹ Iterative updating of sufficient statistics of approx. posteriors by gradient ascent.

The (negative) free energy approximation

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

$$\begin{aligned}\log p(y \mid m) &= \left\langle \log p(y \mid \theta, m) \right\rangle - KL\left[q(\theta), p(\theta \mid m)\right] + KL\left[q(\theta), p(\theta \mid y, m)\right] \\ &= F + KL\left[q(\theta), p(\theta \mid y, m)\right]\end{aligned}$$

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

The complexity term in *F*

- In contrast to AIC & BIC, the complexity term of the negative free energy *F* accounts for parameter interdependencies.

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

- 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

- NB: SPM8 only uses *F* for model selection !

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; ∞[

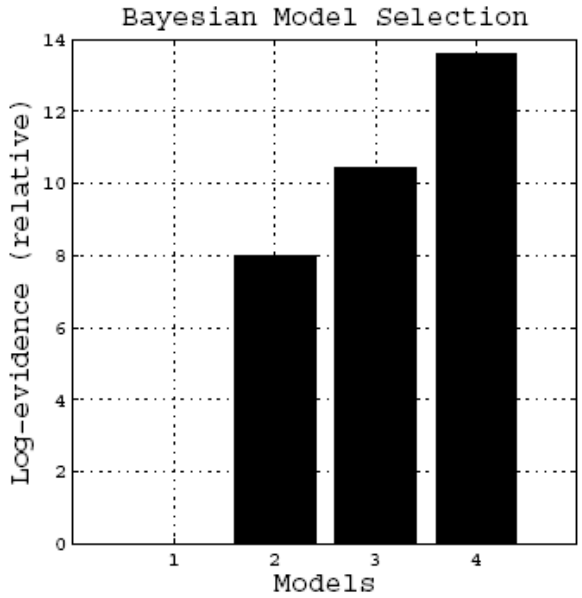
$$B_{12} = \frac{p(y \mid m_1)}{p(y \mid m_2)}$$

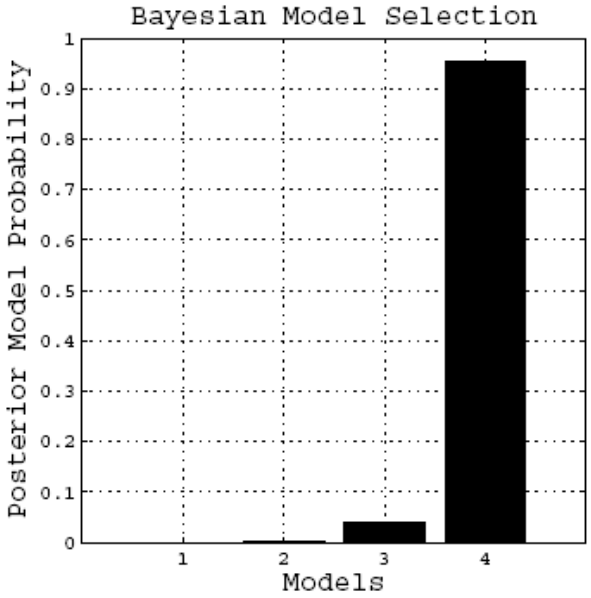
Kass & Raftery classification:

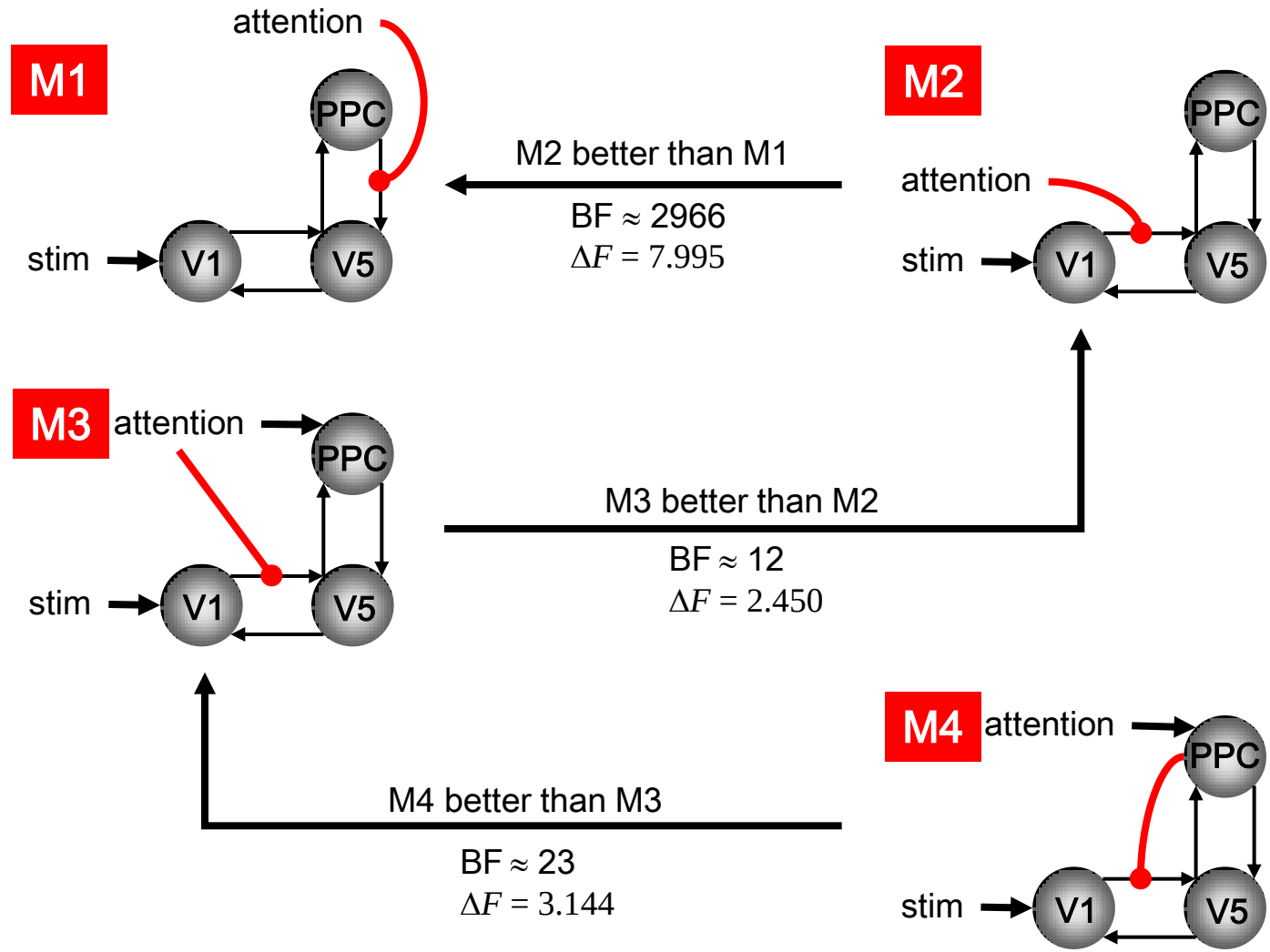
B_{12}	$p(m_1\|y)$	Evidence
1 to 3	50-75%	weak
3 to 20	75-95%	positive
20 to 150	95-99%	strong
≥ 150	≥ 99%	Very strong

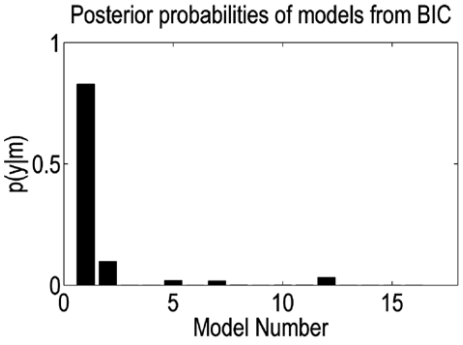
Kass & Raftery 1995, *J. Am. Stat. Assoc.*

BMS in SPM8: an example

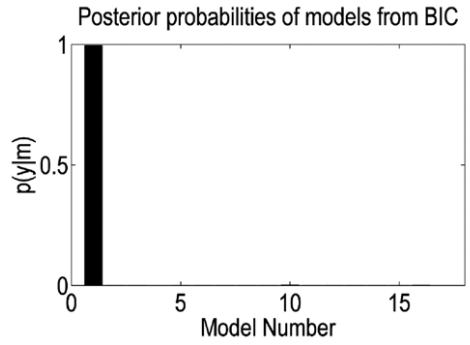




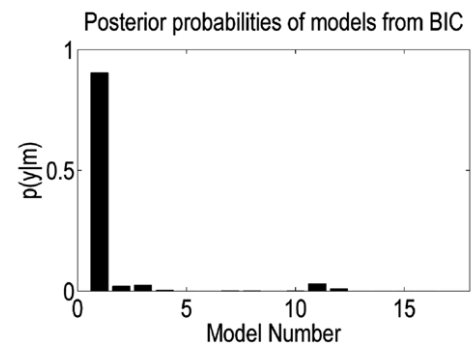




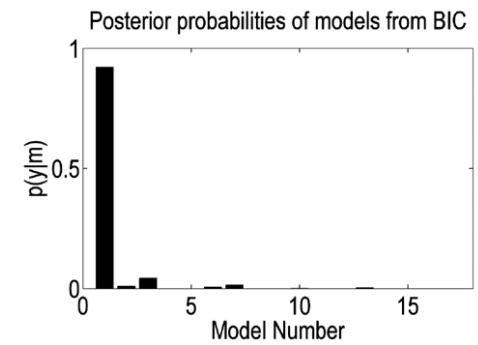
Subject 1



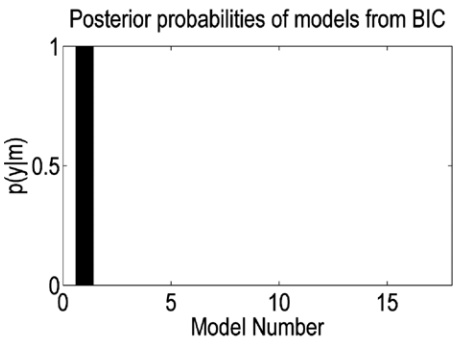
Subject 2



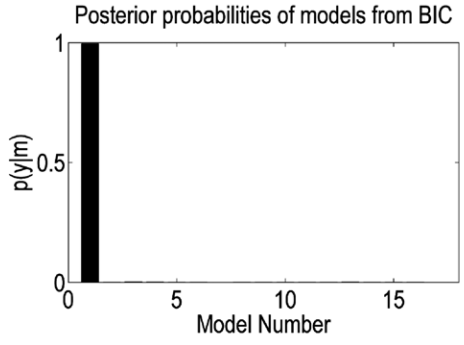
Subject 3



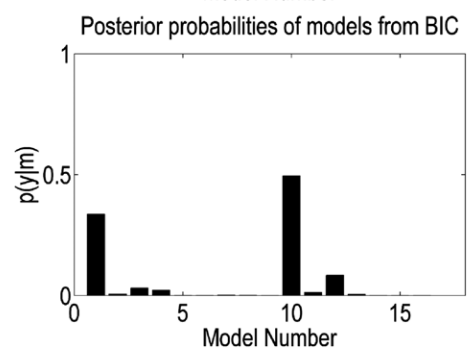
Subject 4



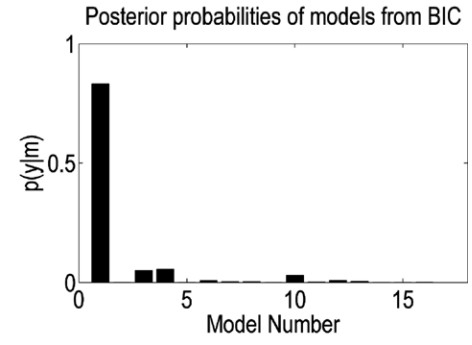
Subject 5



Subject 6



Subject 7



Subject 8

Kumar et al. 2007, *PLoS Comput. Biol.*

Fixed effects BMS at group level

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

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

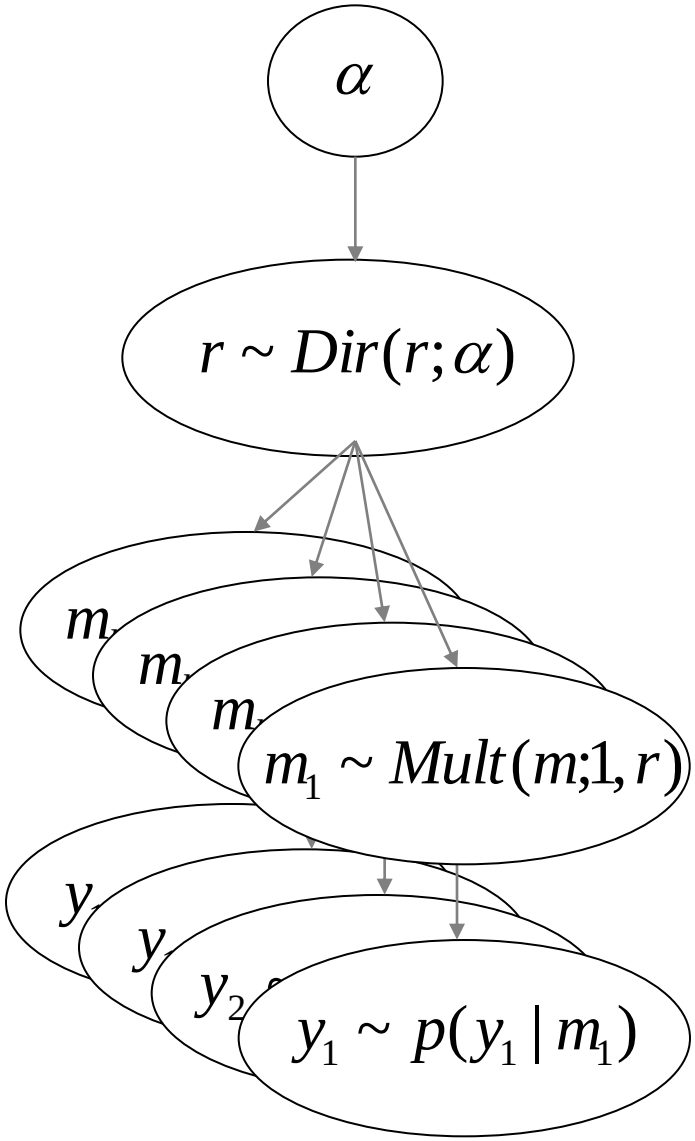
Average Bayes factor (ABF):

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

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

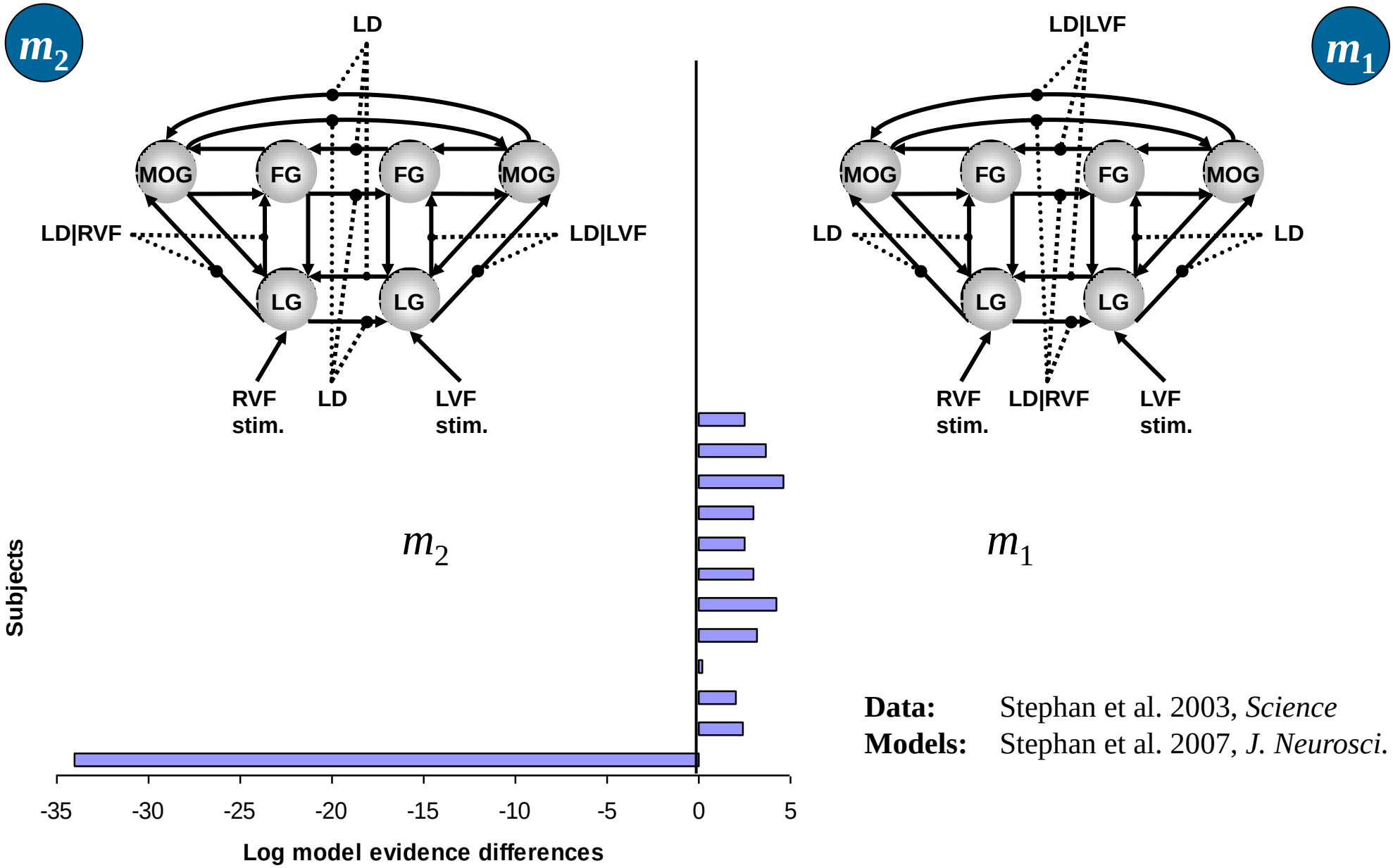
Dirichlet distribution of model probabilities r

Multinomial distribution of model labels m

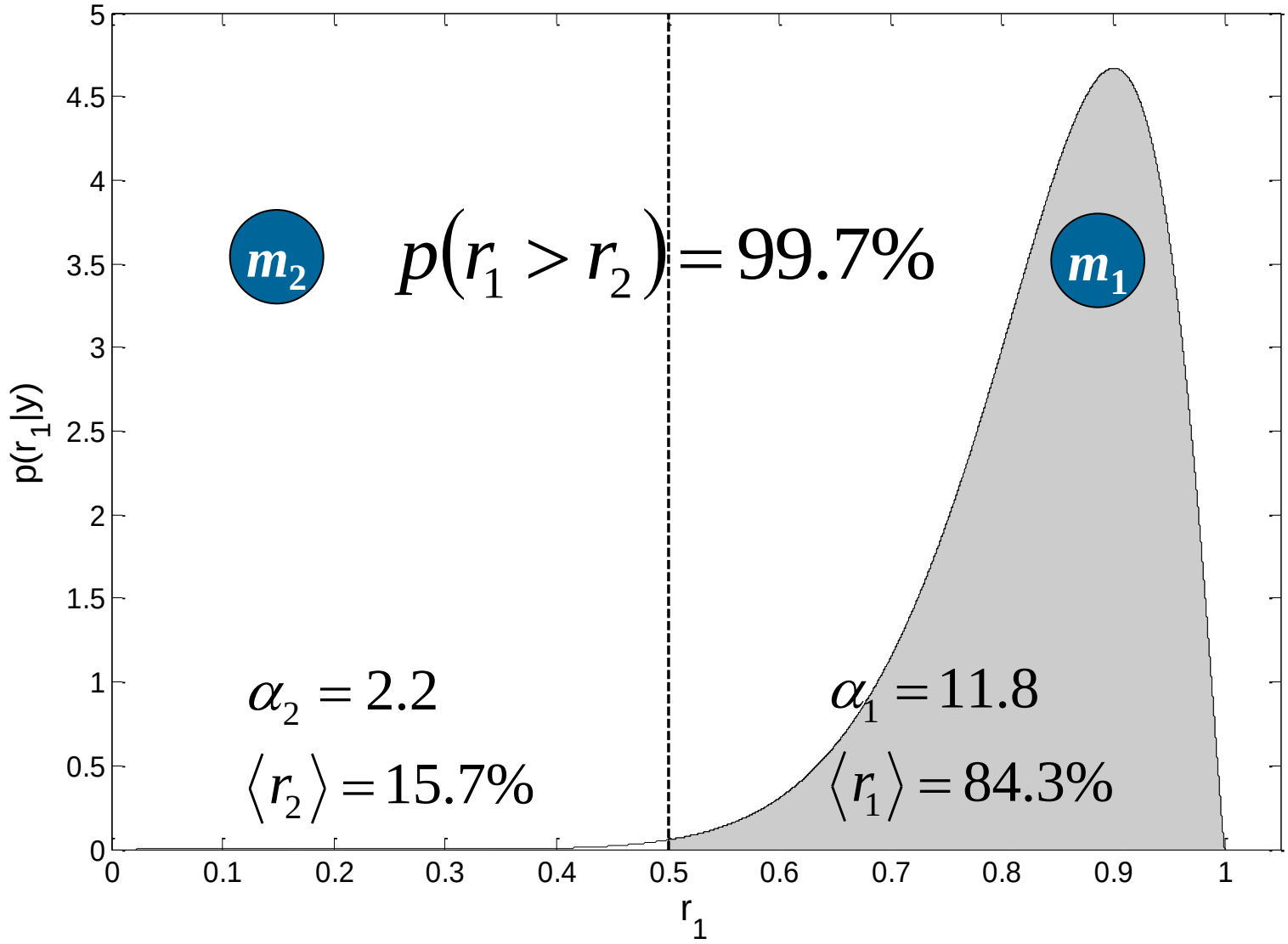
Model inversion by Variational Bayes or MCMC

Measured data y

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



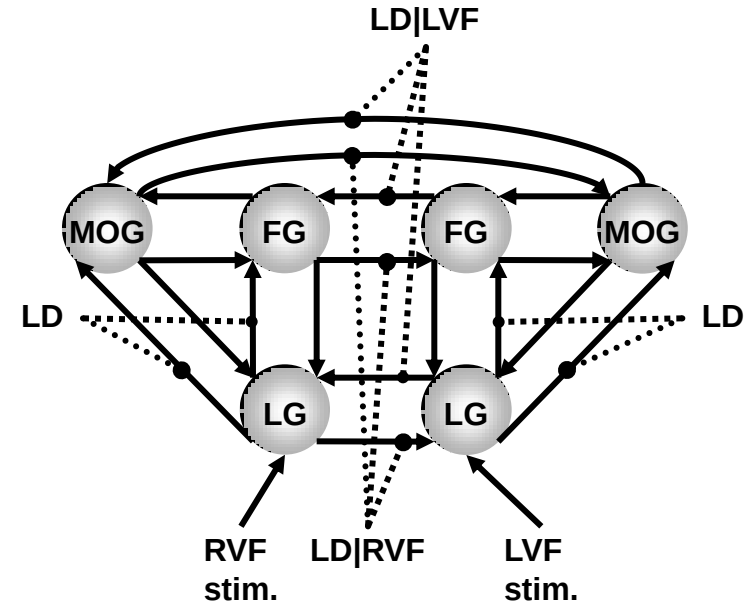
Data: Stephan et al. 2003, *Science*
Models: Stephan et al. 2007, *J. Neurosci.*



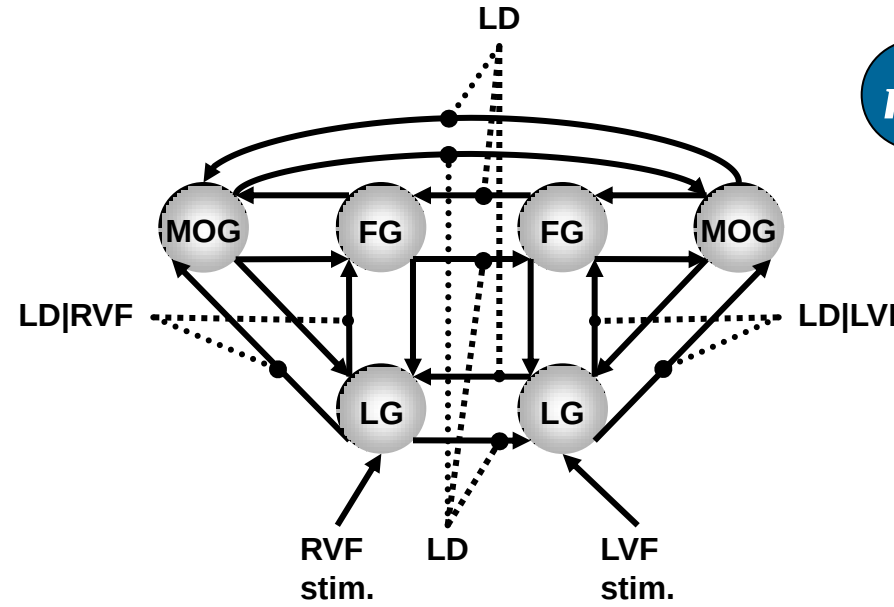
Stephan et al. 2009a, *NeuroImage*

Simulation study: random effects BMS in heterogenous populations

- Population where 70% of all subjects' data are generated by model m_1 and 30% by model m_2
- Random sampling of subjects from this population and generating synthetic data with observation noise
- Fitting both m_1 and m_2 to all data sets and performing BMS

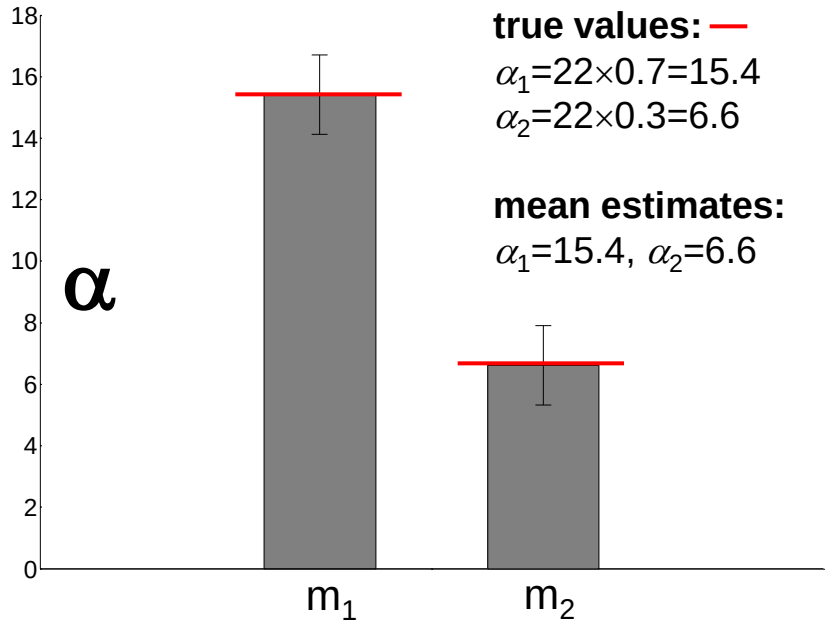


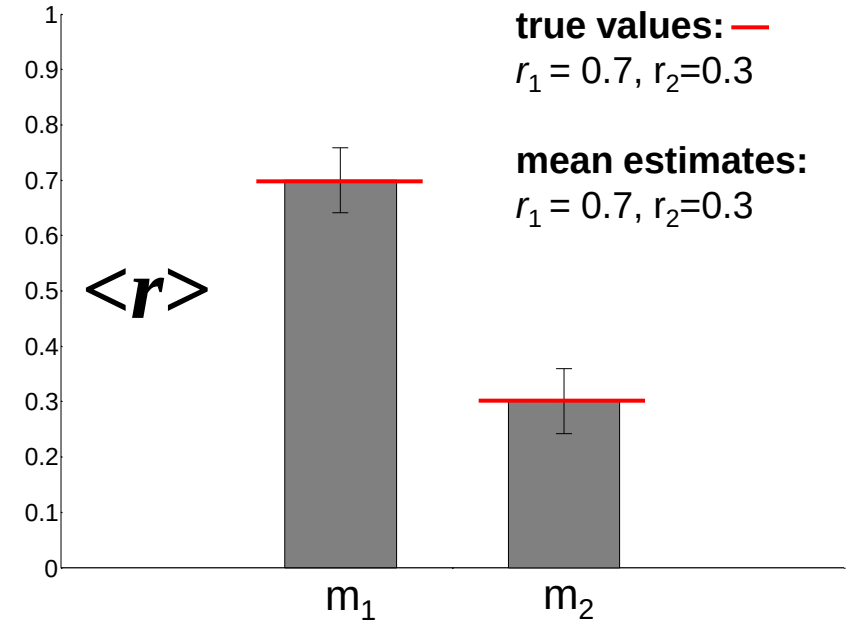
m_1

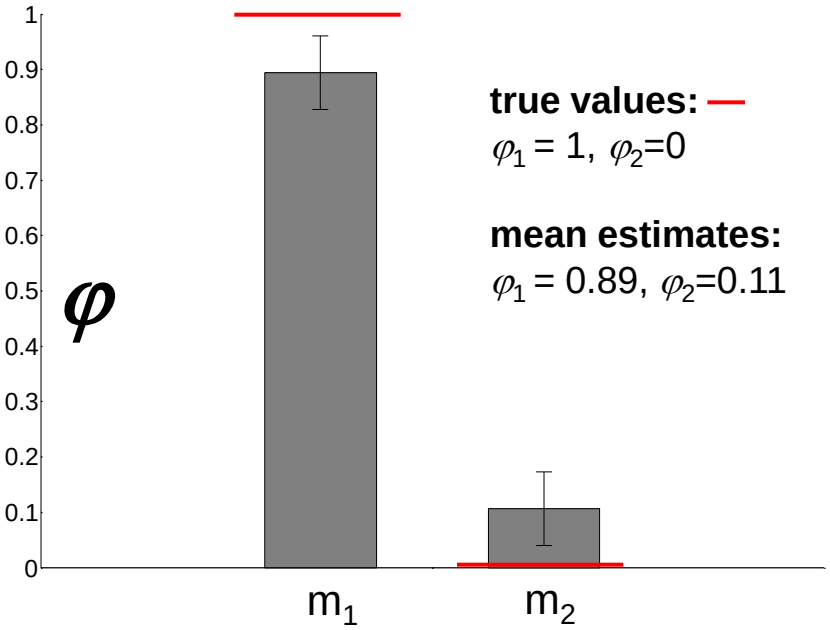


m_2

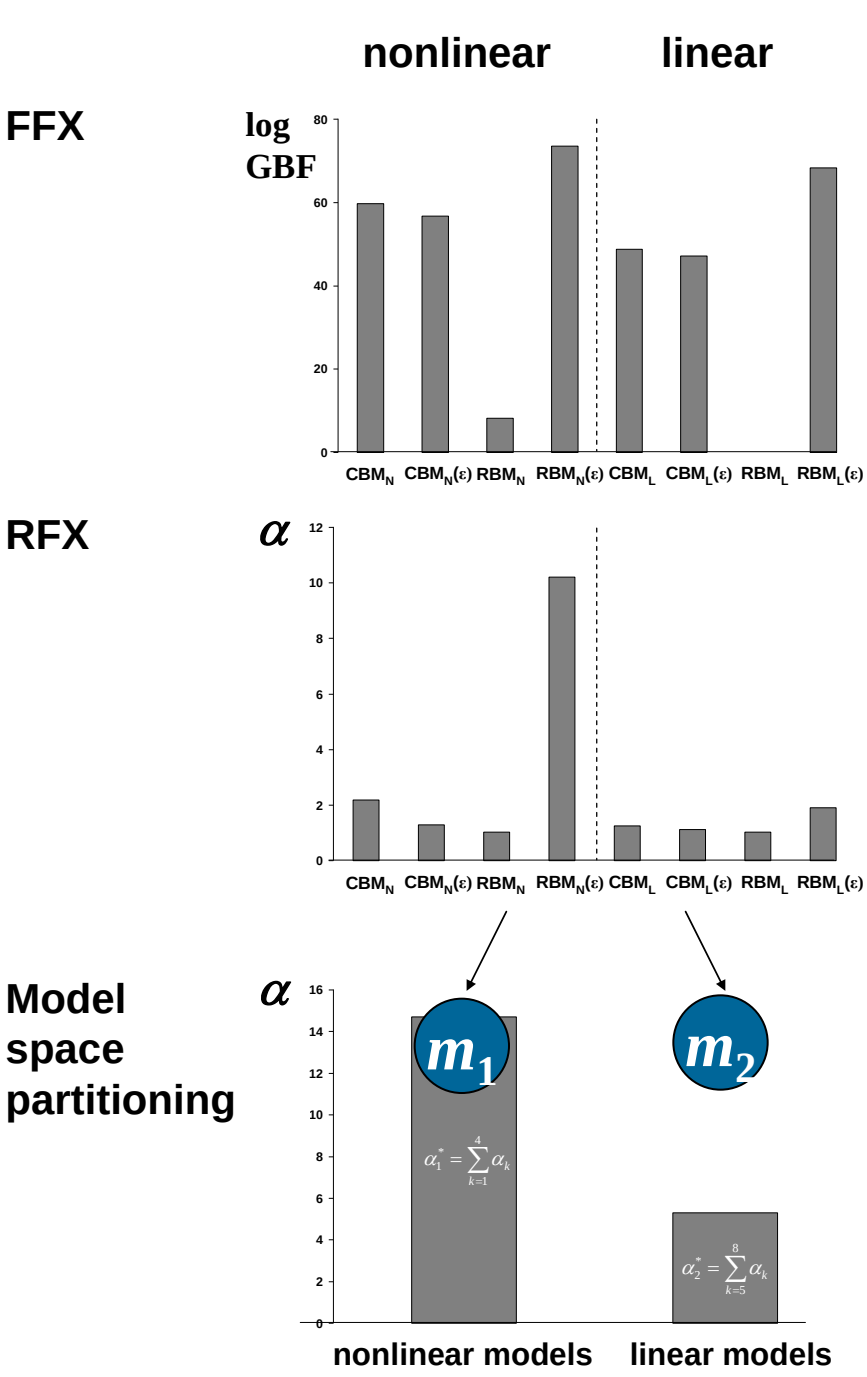
Stephan et al. 2009a, *NeuroImage*

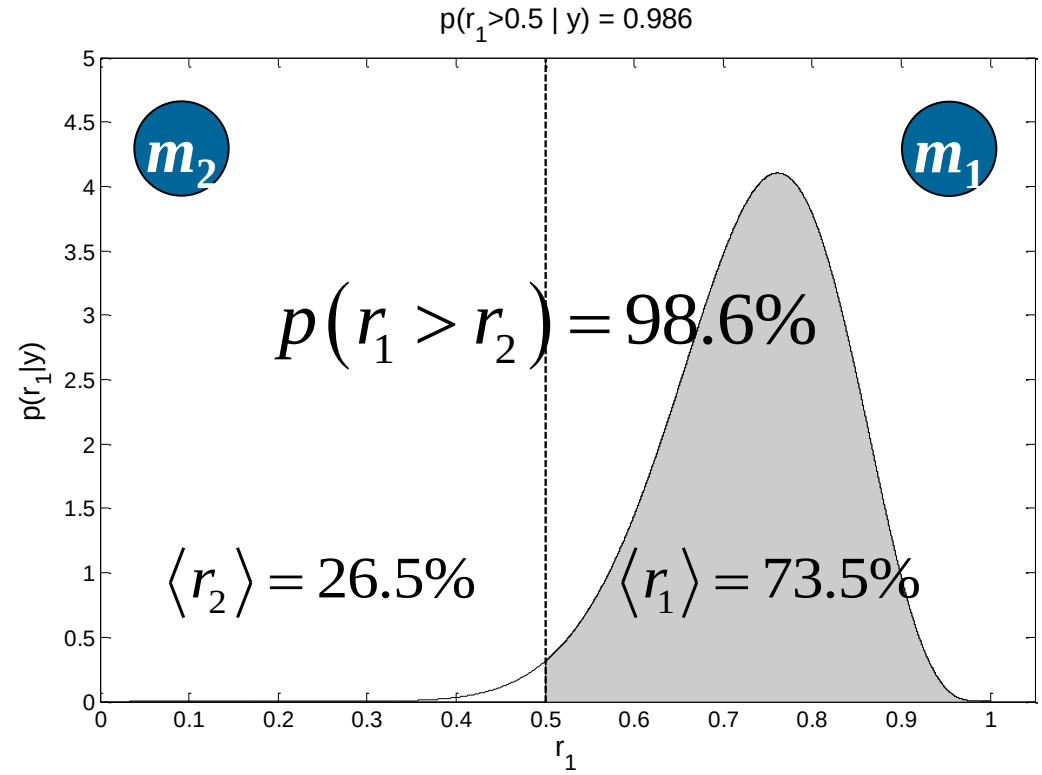

α
18
16
14
12
10
8
6
4
2
0
m1
m2
true values:
α1=22×0.7=15.4
α2=22×0.3=6.6
mean estimates:
α1=15.4, α2=6.6


<r>
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
m1
m2
true values:
r1 = 0.7, r2=0.3
mean estimates:
r1 = 0.7, r2=0.3


φ
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
m1
m2
true values:
φ1 = 1, φ2=0
mean estimates:
φ1 = 0.89, φ2=0.11

Model space partitioning: comparing model families

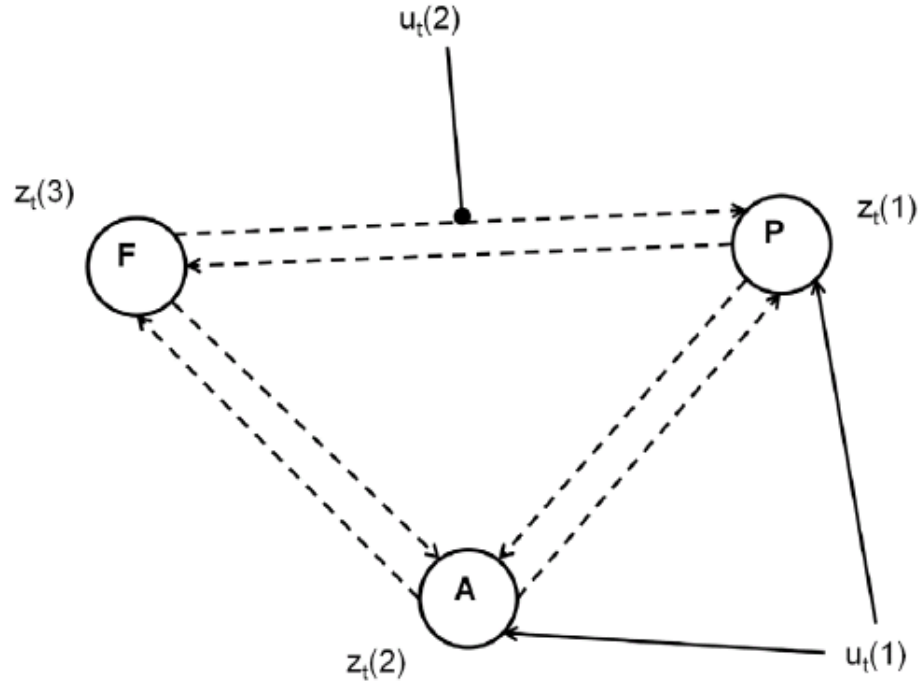


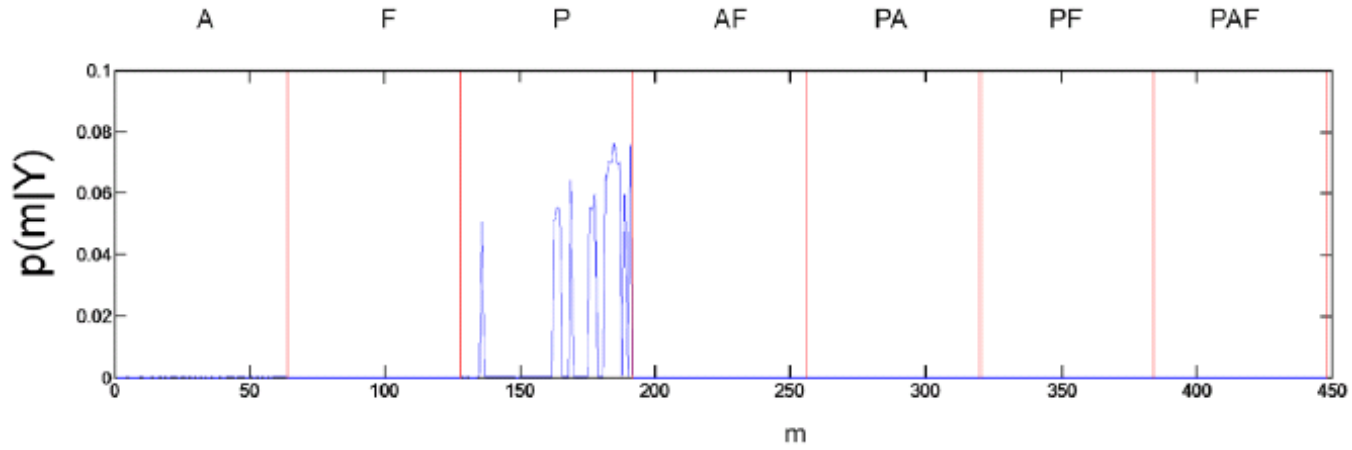


Stephan et al. 2009a, *NeuroImage*
Penny et al. 2010, *PLoS Comput. Biol.*

Comparing model families – a second example

- data from Leff et al. 2008, J. Neurosci
- one driving input, one modulatory input
- $2^6 = 64$ possible modulations
- $2^3 - 1$ input patterns
- $7 \times 64 = 448$ models
- integrate out uncertainty about modulatory patterns and ask where auditory input enters





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

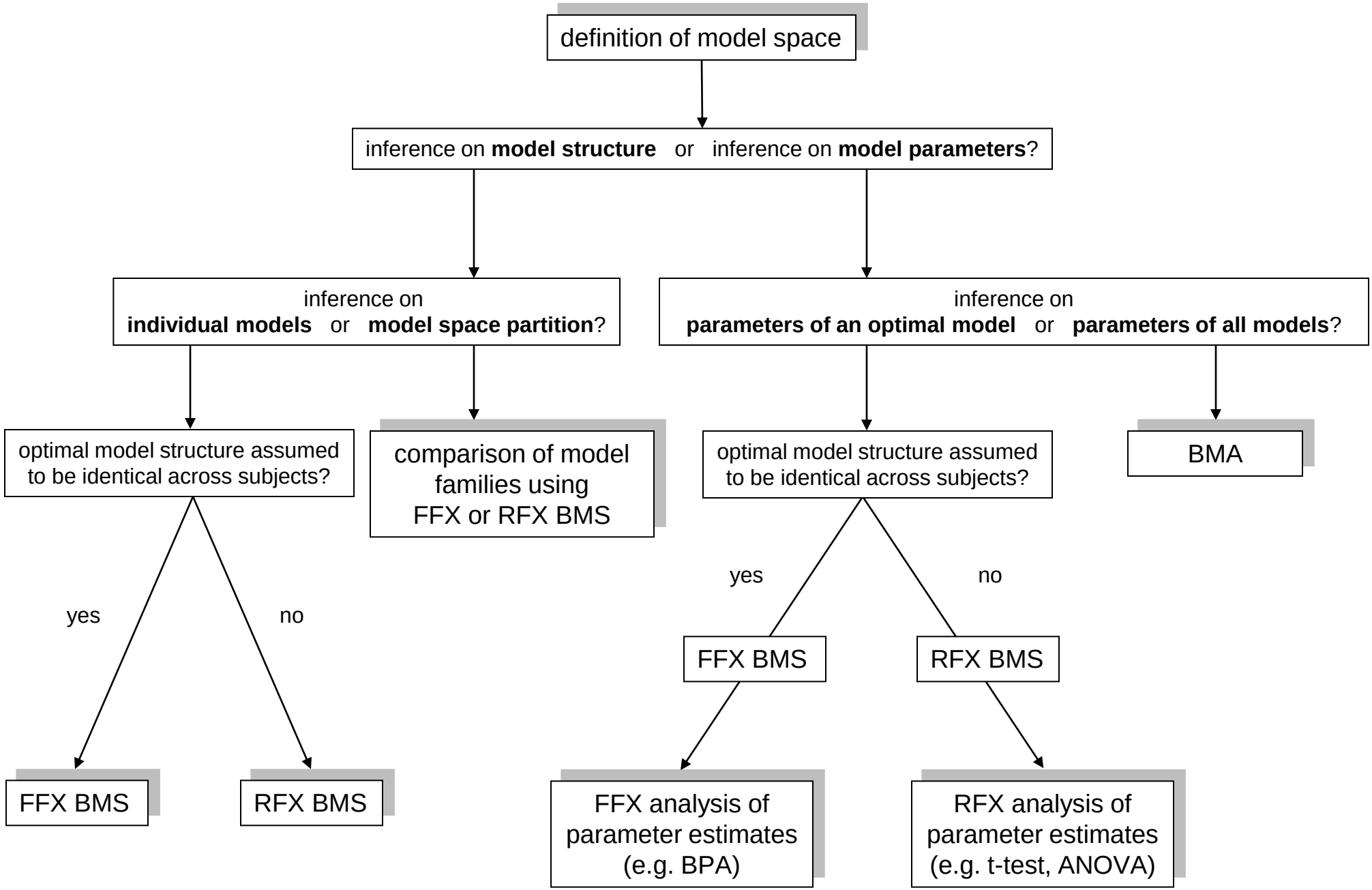
Bayesian Model Averaging (BMA)

- uses the entire model space considered (or an optimal family of models)
- averages parameter estimates, weighted by posterior model probabilities
- particularly useful alternative when
  - none of the models (subspaces) considered clearly outperforms all others
  - when comparing groups for which the optimal model differs

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

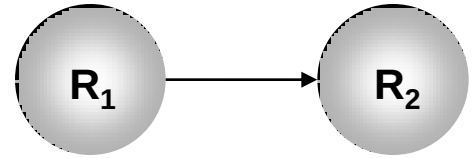
NB: $p(m|y_{1..N})$ can be obtained by either FFX or RFX BMS

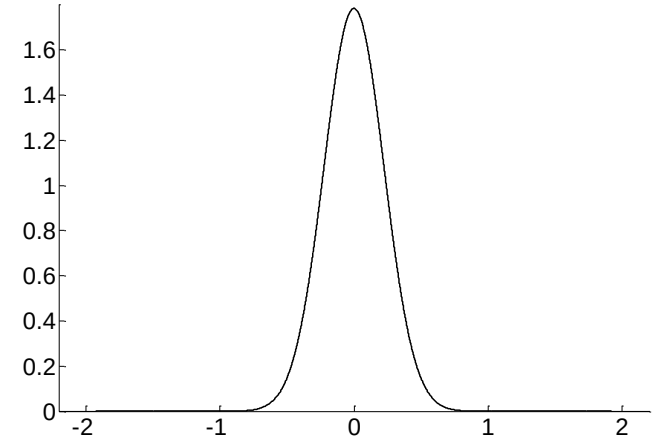
Penny et al. 2010, *PLoS Comput. Biol.*



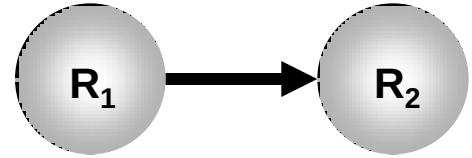
Stephan et al. 2010, *NeuroImage*

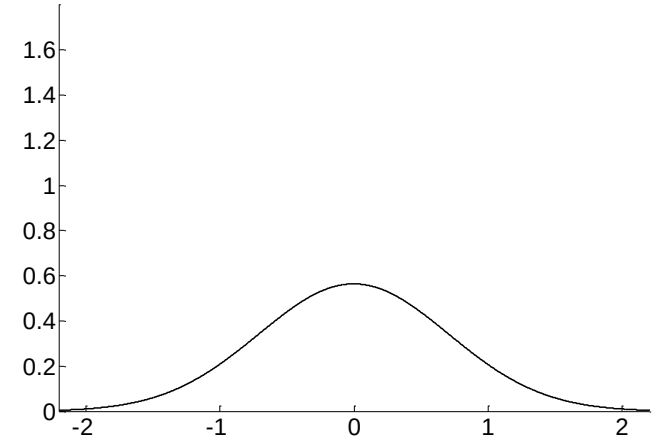
Integration of tractography and DCM





low probability of anatomical connection
→ **small prior variance** of effective connectivity parameter

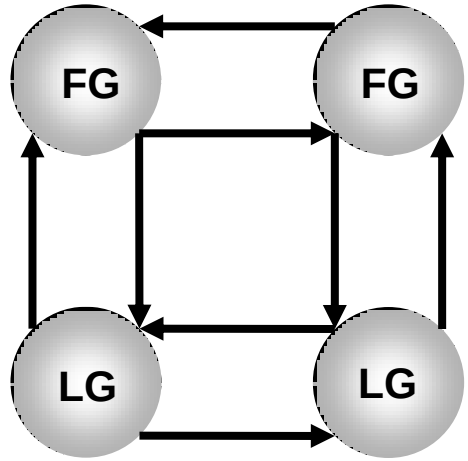




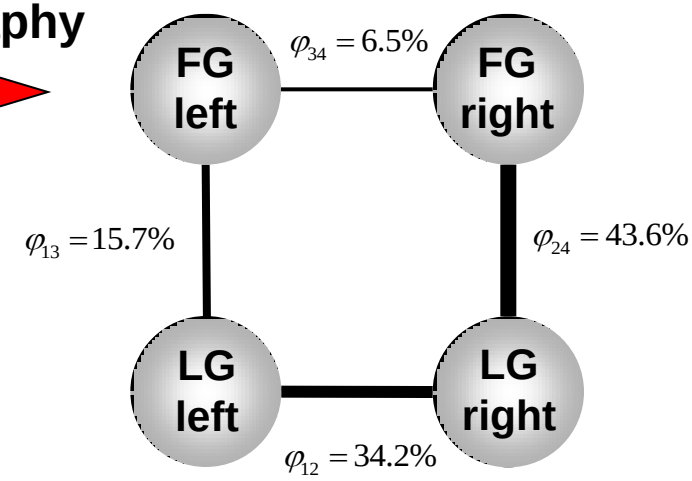
high probability of anatomical connection
→ **large prior variance** of effective connectivity parameter

Stephan, Tittgemeyer et al. 2009, *NeuroImage*

Proof of concept study

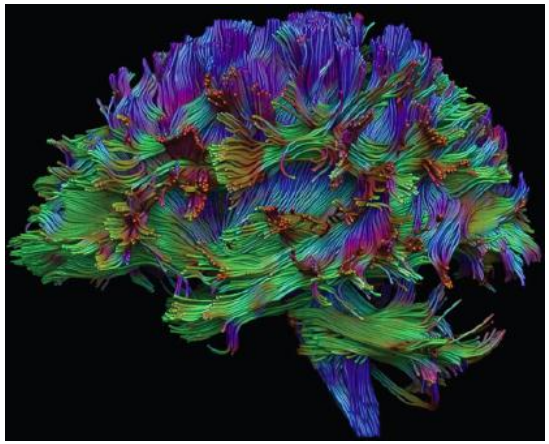


probabilistic tractography

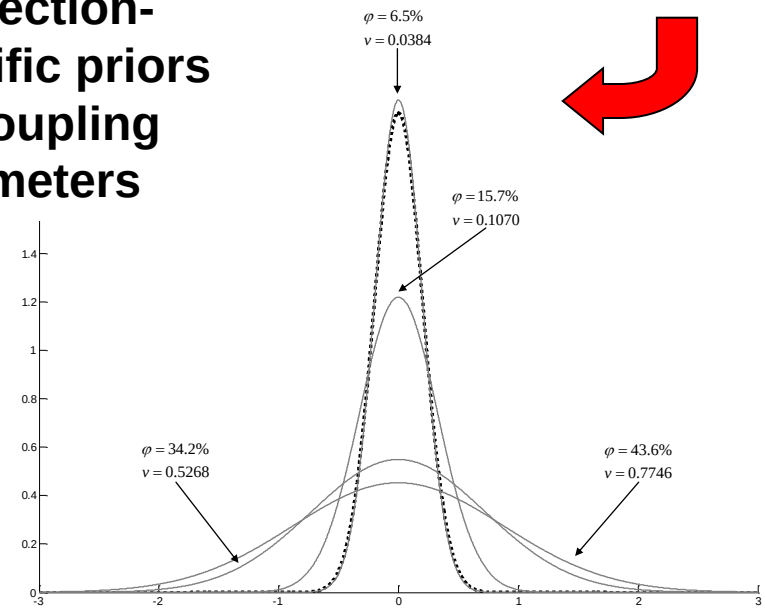


❶ DCM

❷ anatomical connectivity φ



❸ connection-specific priors for coupling parameters

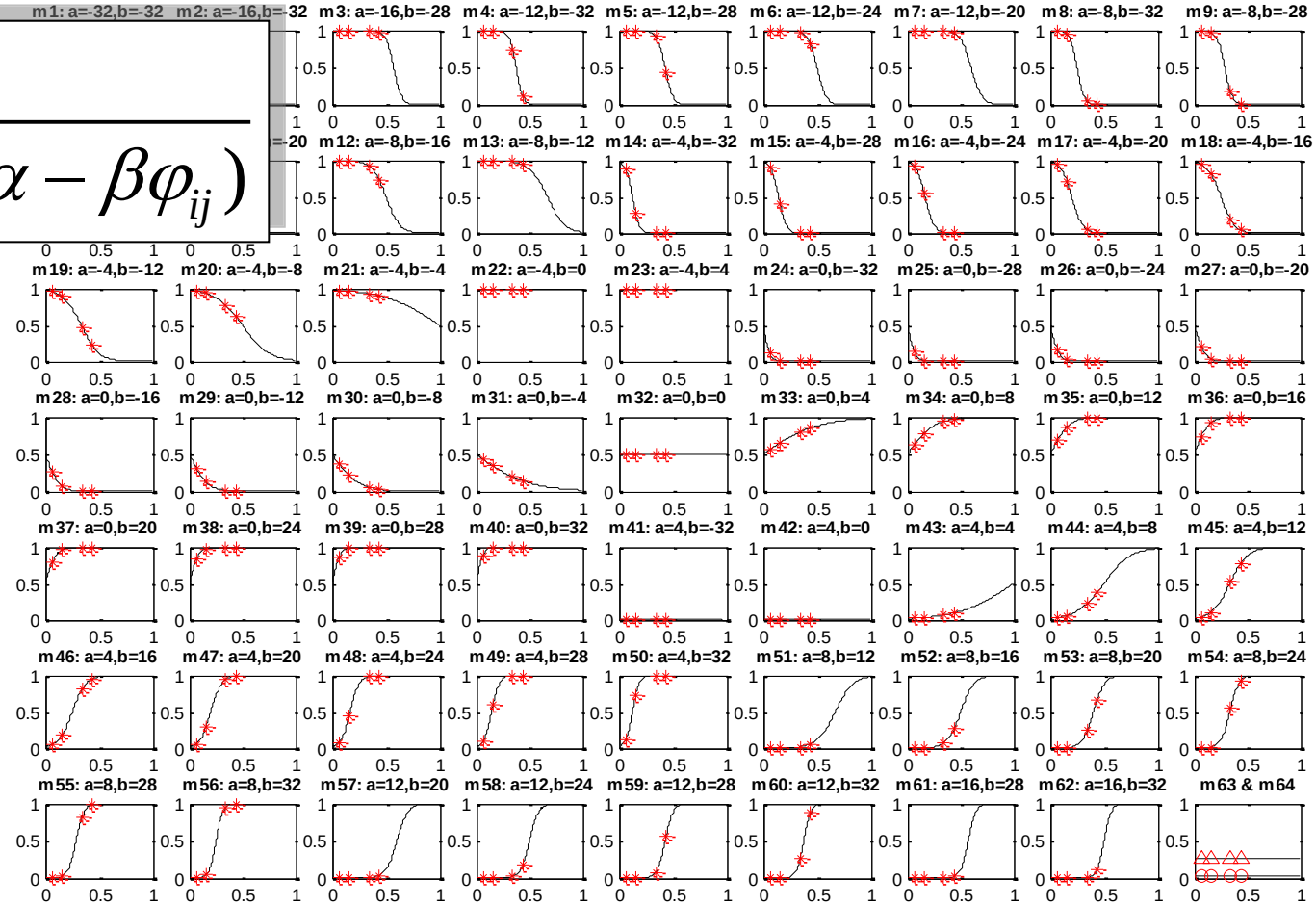


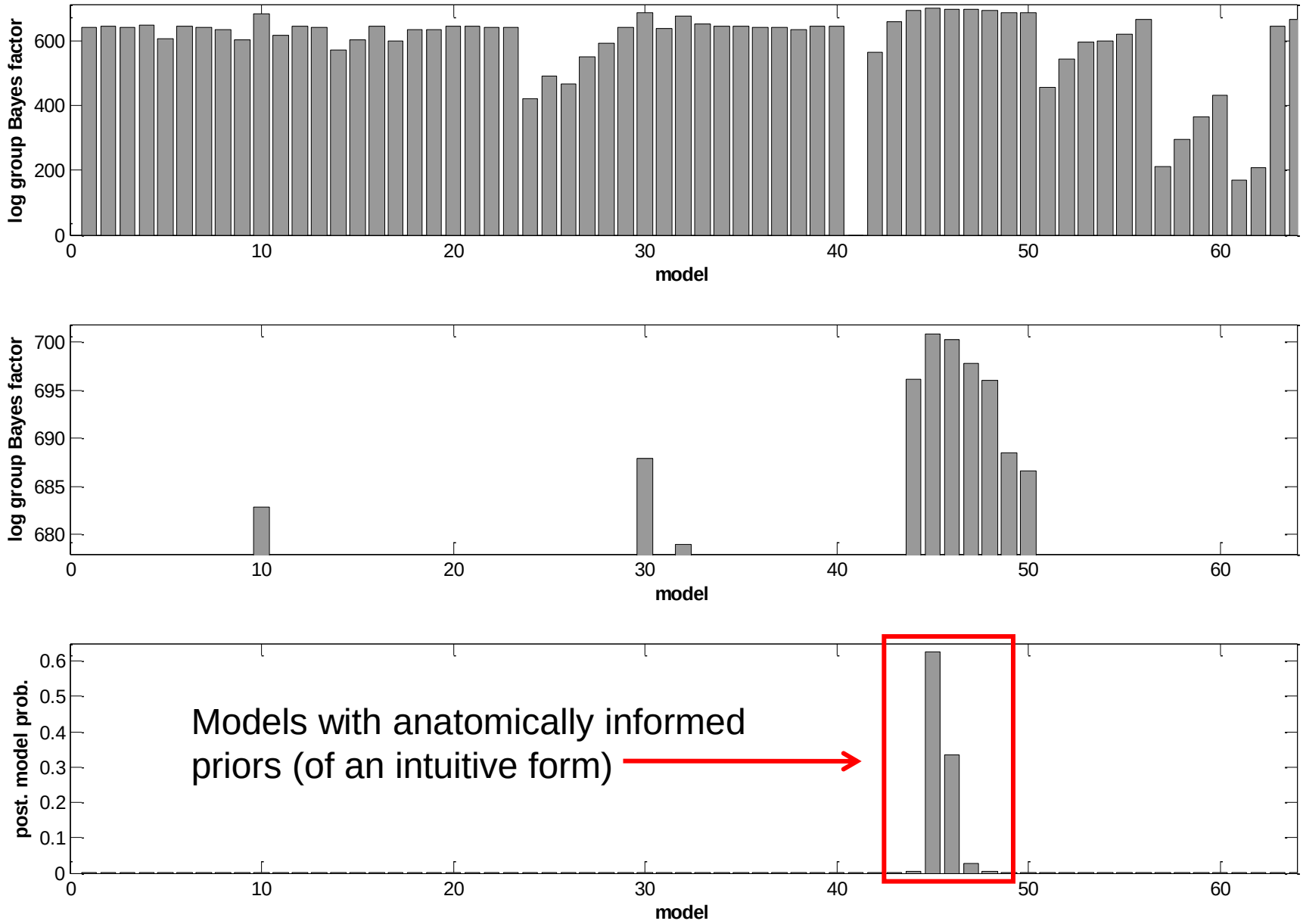
Stephan, Tittgemeyer et al. 2009, *NeuroImage*

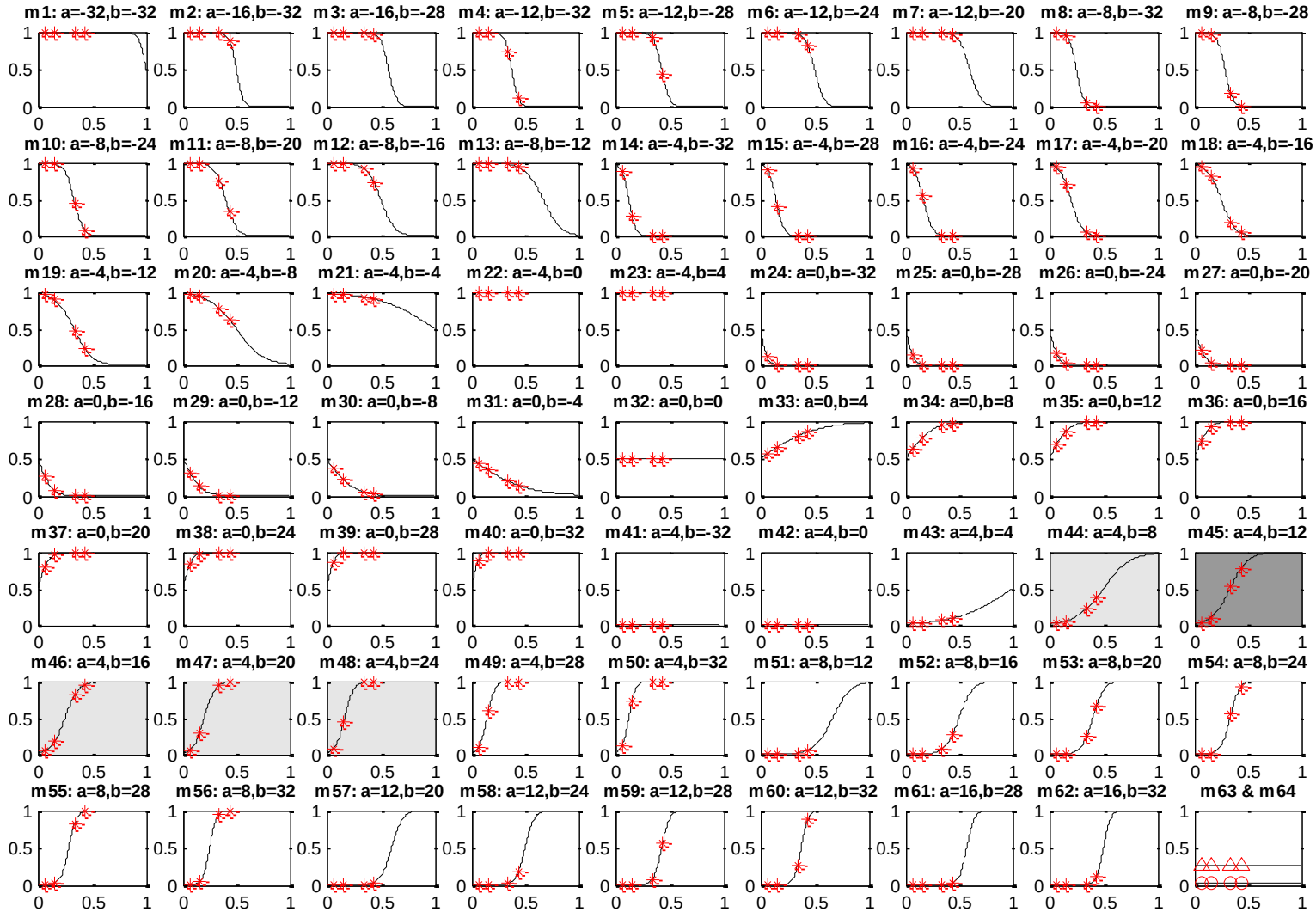
Connection-specific prior variance Σ as a function of anatomical connection probability φ

$$\Sigma_{ij} = \frac{\Sigma_0}{1 + \Sigma_0 \exp(\alpha - \beta\varphi_{ij})}$$

- 64 different mappings by systematic search across hyper-parameters α and β
- yields anatomically informed (intuitive and counterintuitive) and uninformed priors




log group Bayes factor
0
200
400
600
0
10
20
30
40
50
60
model
log group Bayes factor
680
685
690
695
700
0
10
20
30
40
50
60
model
post. model prob.
0
0.1
0.2
0.3
0.4
0.5
0.6
Models with anatomically informed
priors (of an intuitive form)
0
10
20
30
40
50
60
model



Models with anatomically informed priors (of an intuitive form) were clearly superior to anatomically uninformed ones: Bayes Factor $>10^9$

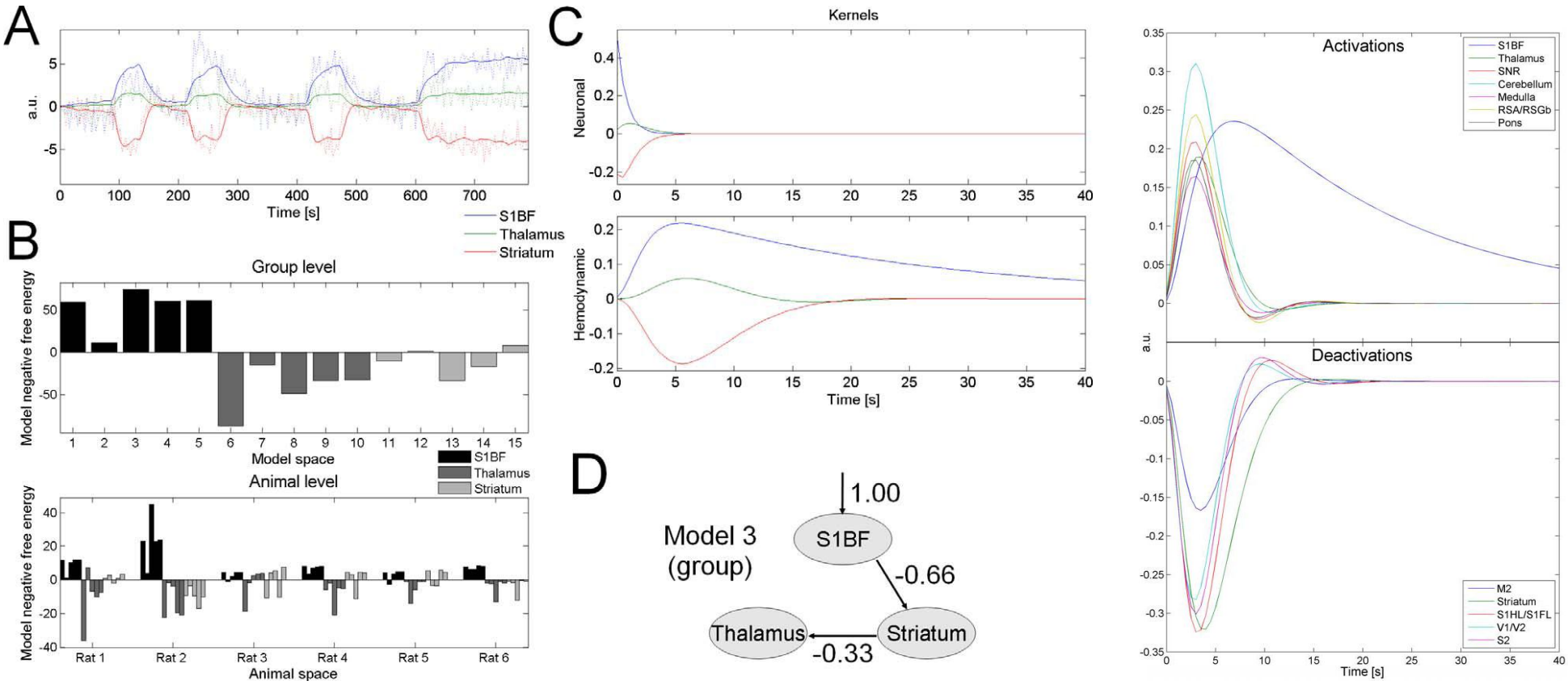
Hierarchical strategy for model validation

1	in silico	**numerical analysis & simulation studies**
2	humans	**cognitive experiments**
3	animals & humans	**experimentally controlled system perturbations**
4	patients	**clinical facts**

Previous validation studies of DCM

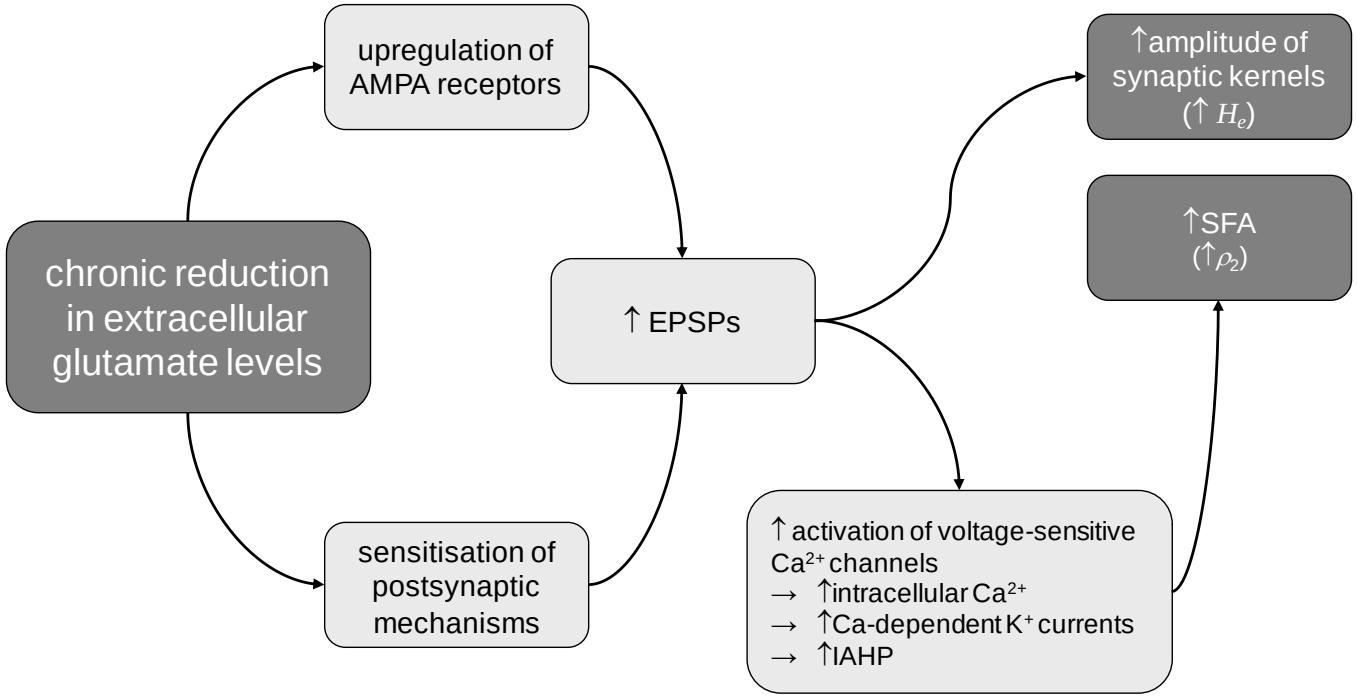
- reliability (reproducibilty)
  - parameter estimates are highly reliable across sessions (Schuyler et al. 2010)
  - model selection results are highly reliable across sessions (Rowe et al. 2010)
- face validity
  - simulations and empirical studies (Stephan et al. 2007, 2008)
- construct validity
  - comparison with SEM (Penny et al. 2004)
  - comparison with large-scale spiking neuron models (Lee et al. 2006)
- predictive validity:
  - infers correct site of seizure origin (David et al. 2008)
  - infers primary recipient of vagal nerve stimulation (Reyt et al. 2010)
  - infers synaptic changes as predicted from microdialysis (Moran et al. 2008)
  - infers conditioning-induced plasticity in amygdala (Moran et al. 2009)
  - tracks anaesthesia levels (Moran et al. 2011)
  - predicts sensory stimulation (Brodersen et al. 2010)
  - infers DA induced changes in AMPA/NMDA ratio from MEG (Moran et al. 2011)
  - predicts presence/absence of remote lesion (Brodersen et al. 2011)

Validation: Predicting origin of epileptic seizures from fMRI



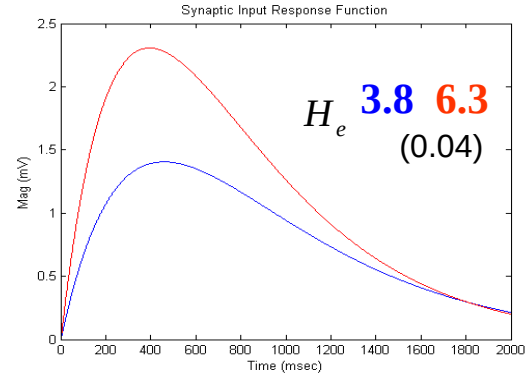
David et al. 2008, *PLoS Biol.*

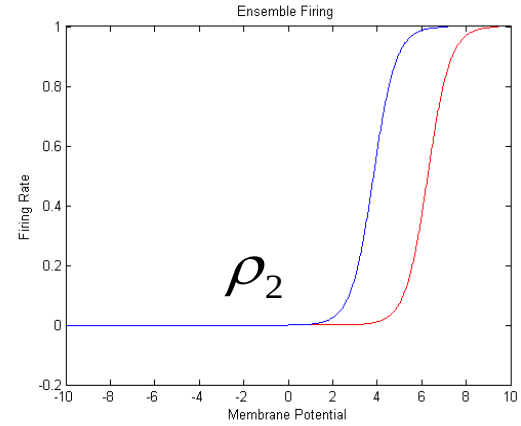
Validation: microdialysis in rat prefrontal cortex



Moran et al. 2008, *NeuroImage*

Sensitization of post-synaptic mechanisms



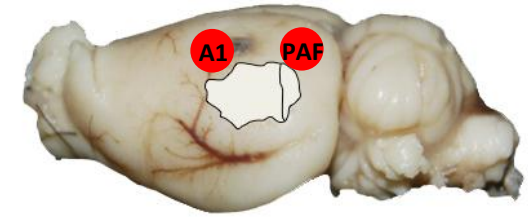


[0.76,1.34] (0.0003)

Compensatory increase in SFA

Validation: different levels of anaesthesia

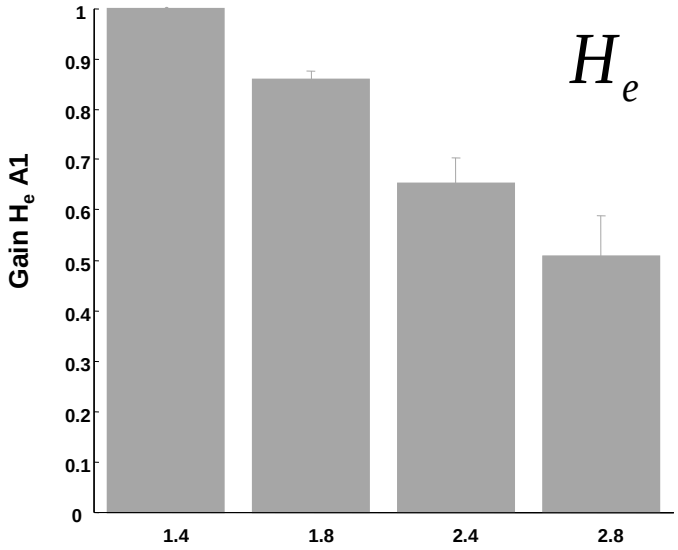
(in collaboration with MPI Cologne)



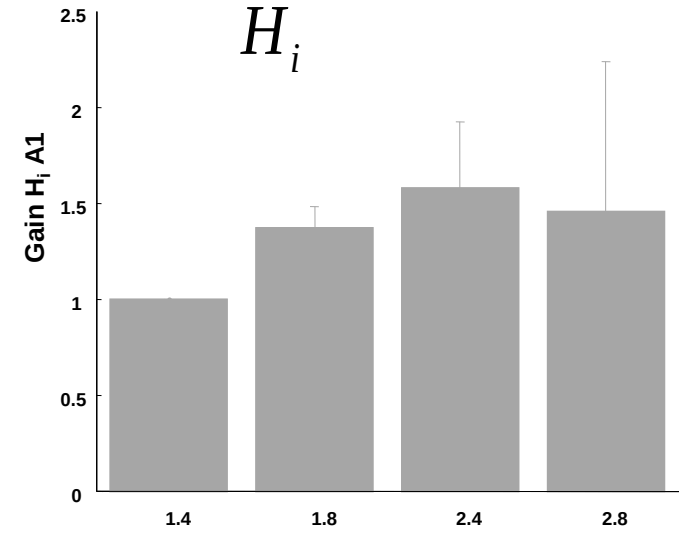
1.4 % isoflurane	1.8 % isoflurane	2.4 % isoflurane	2.8 % isoflurane

DCM fitted to LFPs from rat auditory cortex (A1 and PAF)

EPSP amplitudes

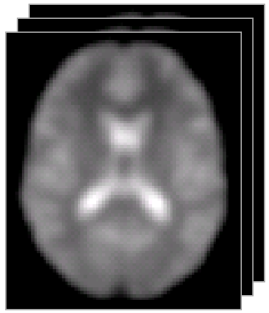


IPSP amplitudes



Moran et al. 2011, *PLoS ONE*

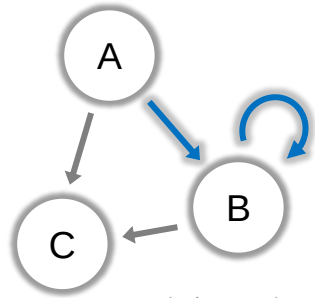
Model-based decoding by generative embedding



measurements from
an individual subject

**step 1 —
model inversion**

$\mathcal{X} \to \mathcal{M}_\Theta$



$p(\theta|x, m)$

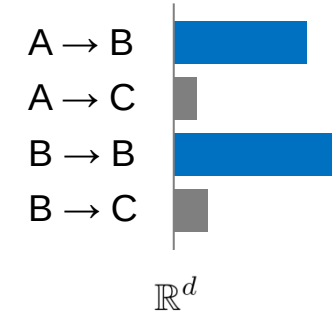
subject-specific
inverted generative model

**step 2 —
kernel construction**

$\mathcal{M}_\Theta \to \mathbb{R}^d$

$k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$

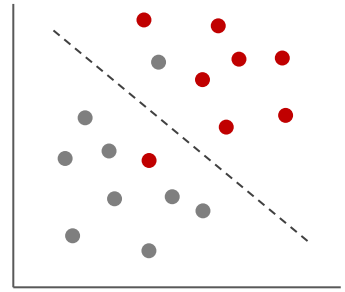
$k_\mathcal{M} : \mathcal{M}_\Theta \times \mathcal{M}_\Theta \to \mathbb{R}$



subject representation in the
generative score space

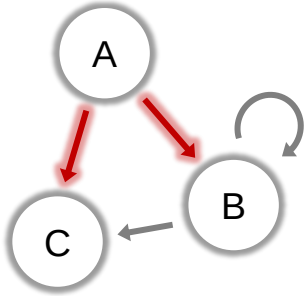
**step 3 —
support vector classification**

$$\hat{c} = \text{sgn}\left(\sum_i^n \alpha_i^* k(x_i, x) + b^*\right)$$



separating hyperplane fitted to
discriminate between groups

**step 4 —
interpretation**

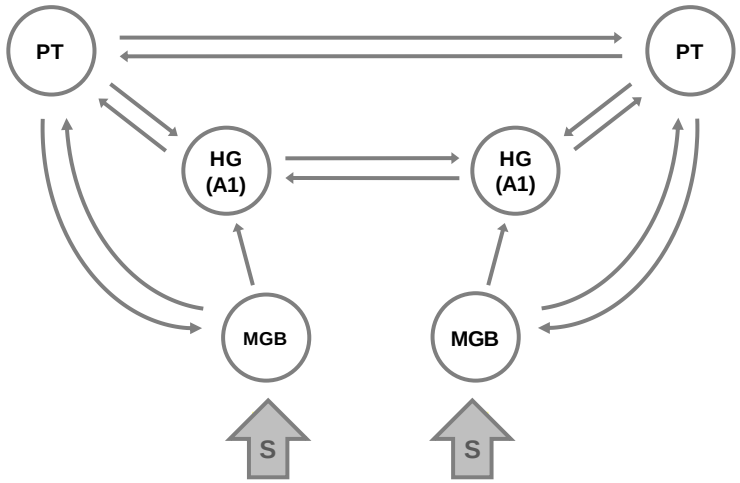


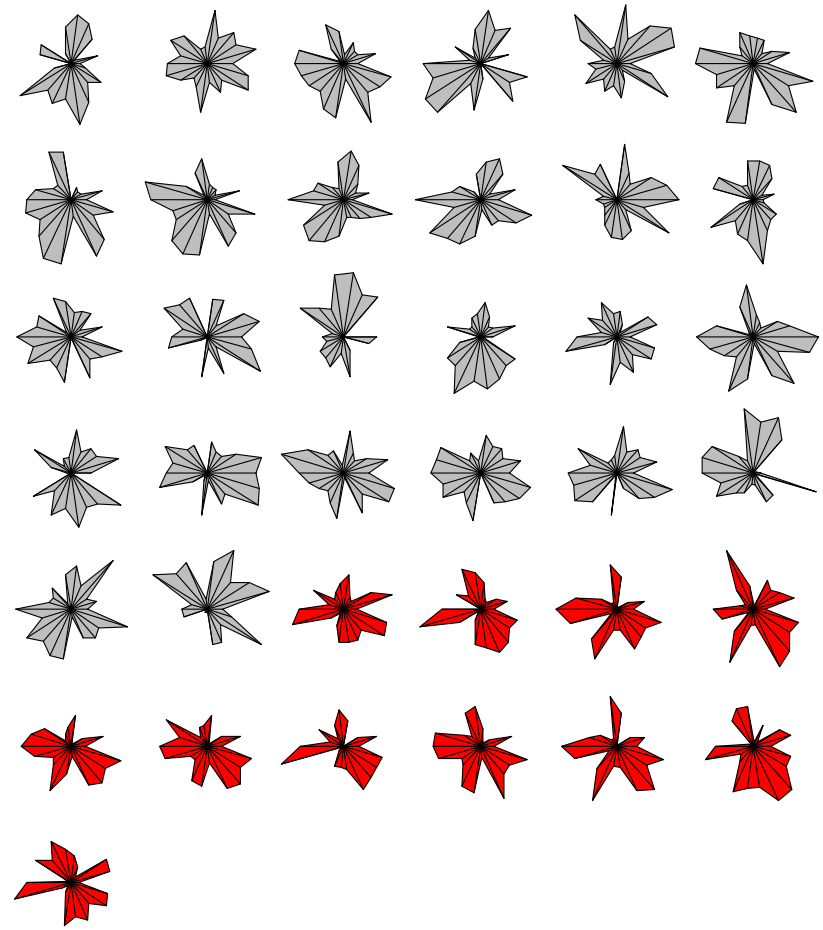
jointly discriminative
model parameters

Brodersen et al. 2011, *PLoS Comput. Biol.*

Model-based decoding of disease status: mildly aphasic patients (N=11) vs. controls (N=26)

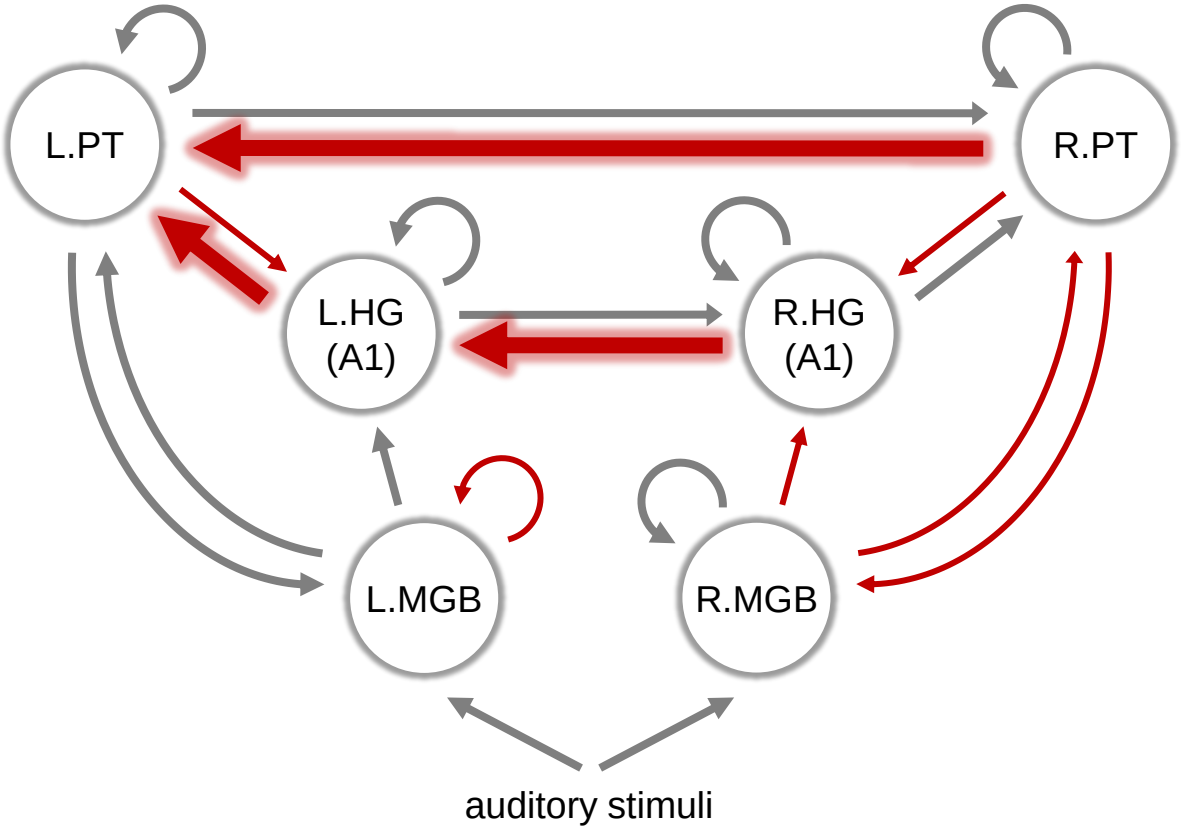
Connectional fingerprints from a 6-region DCM of auditory areas during speech perception

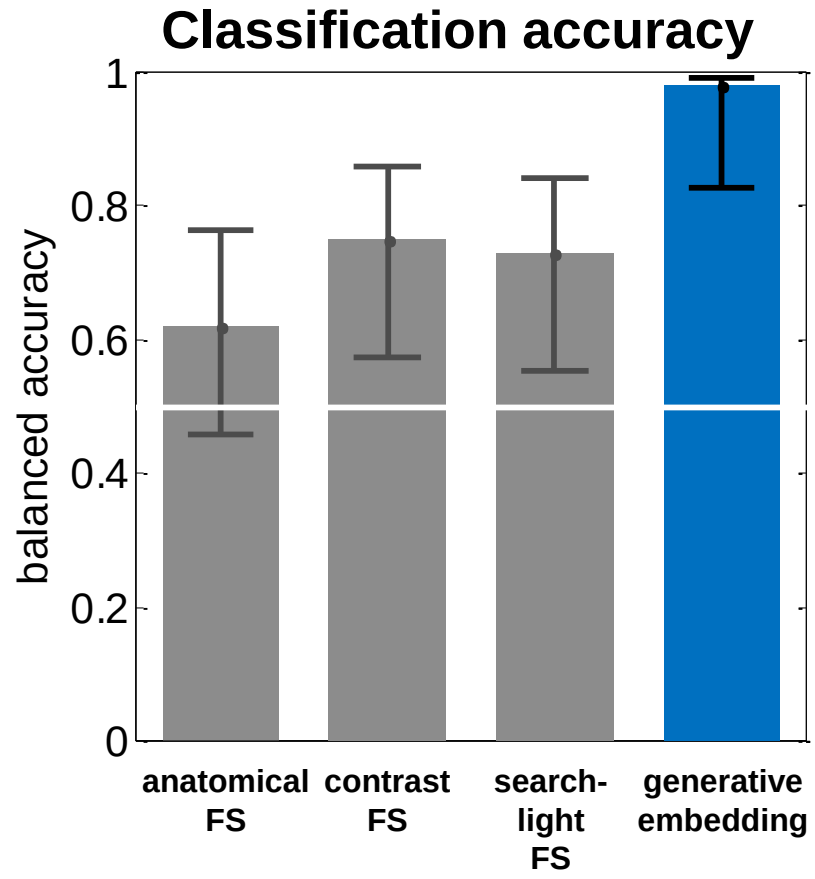




Brodersen et al. 2011, *PLoS Comput. Biol.*

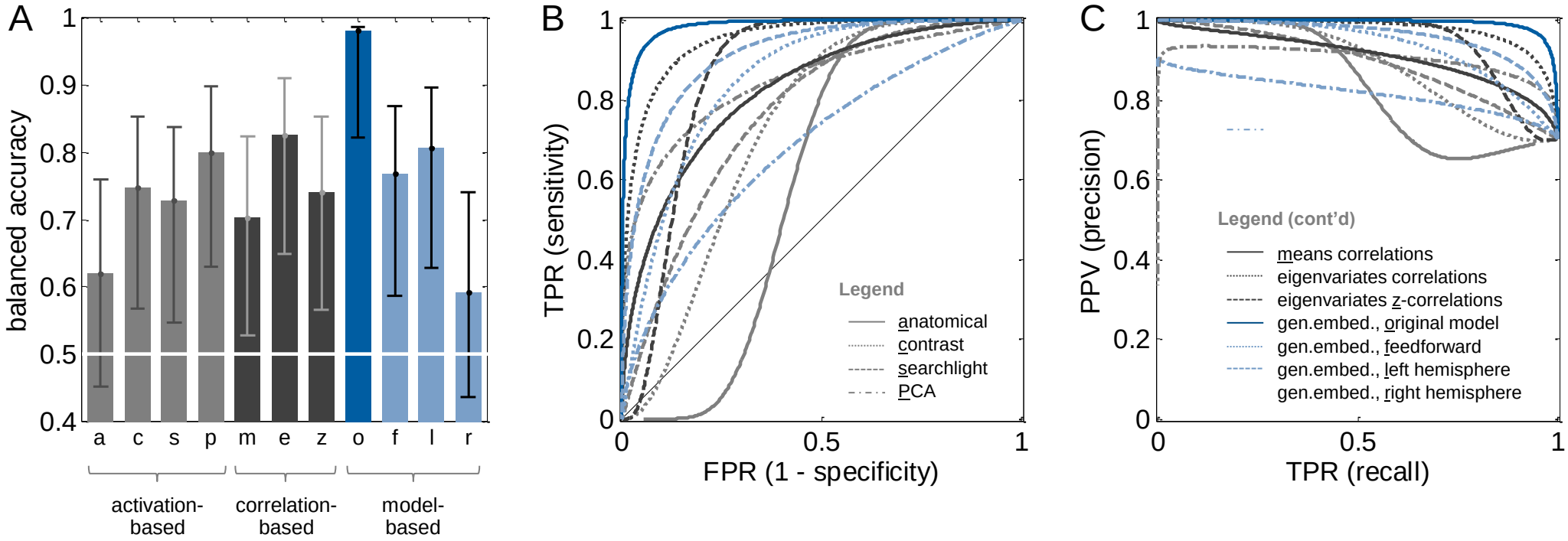
Model-based decoding of disease status: mildly aphasic patients (N=11) vs. controls (N=26)





Brodersen et al. 2011, *PLoS Comput. Biol.*

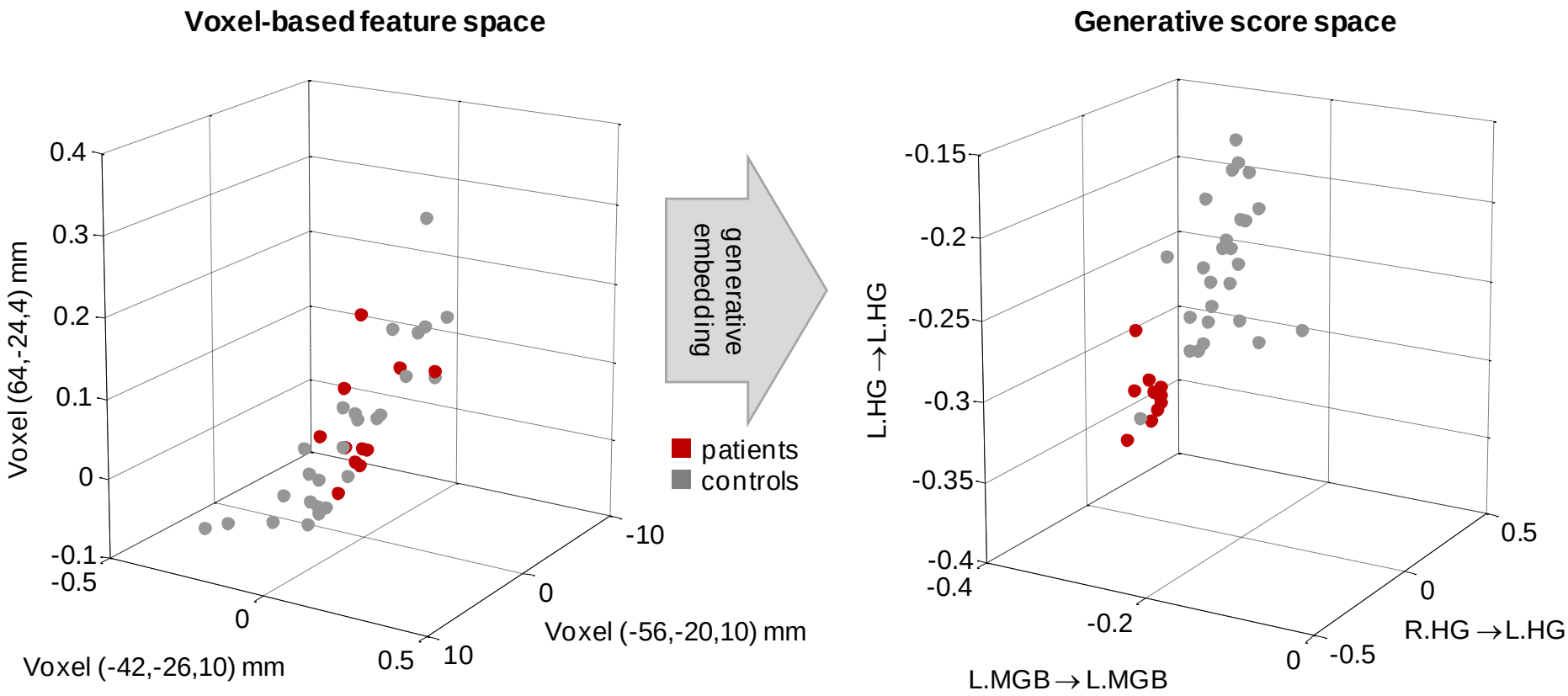
Sensitivity: 100 %
Specificity: 96.2%



Brodersen et al. 2011, *PLoS Comput. Biol.*

Multivariate searchlight classification analysis

Generative embedding using DCM



Brodersen et al. 2011, *PLoS Comput. Biol.*

Some key papers on DCM and BMS

- 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.
- David O, Kiebel SJ, Harrison LM, Mattout J, Kilner JM, Friston KJ (2006) Dynamic causal modeling of evoked responses in EEG and MEG. Neuroimage. 2006 May 1;30(4):1255-72.
- Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. NeuroImage 19:1273-1302.
- Friston K, Mattout J, Trujillo-Barreto N, Ashburner J, Penny W (2007) Variational free energy and the Laplace approximation. Neuroimage 34: 220-234.
- Friston K, Penny W (2011) Post hoc Bayesian model selection. Neuroimage 56: 2089-2099.
- Moran RJ, Stephan KE, Kiebel SJ, Rombach N, O'Connor WT, Murphy KJ, Reilly RB, Friston KJ (2008) Bayesian estimation of synaptic physiology from the spectral responses of neural masses. NeuroImage 42: 272-284.
- Moran RJ, Stephan KE, Seidenbecher T, Pape HC, Dolan RJ, Friston KJ (2009) Dynamic causal models of steady-state responses. NeuroImage 44: 796-811.
- Penny WD, Stephan KE, Mechelli A, Friston KJ (2004) Comparing dynamic causal models. NeuroImage 22:1157-1172.
- 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.
- 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 HEM, 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 (2009) Bayesian model selection for group studies. NeuroImage 46:1004-1017.
- 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