Translational Neuromodeling Unit

University of Zurich[UZH]

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Generative models of fMRI data: DCM with a special intro to rDCM

Jakob Heinzle + Stefan Frässle

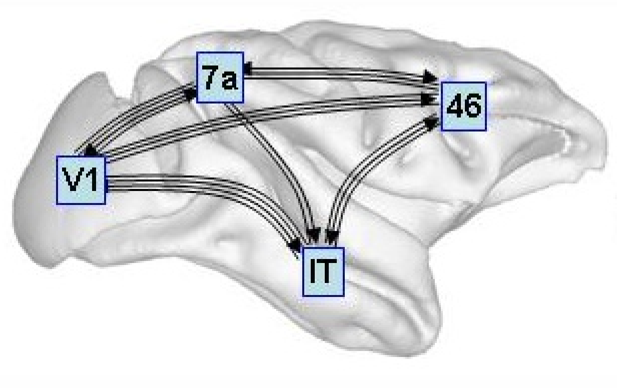
Translational Neuromodeling Unit (TNU)
University of Zurich & ETH Zurich

Translational Neuromodeling (FS 2021)

Zurich, 30 March, 2021

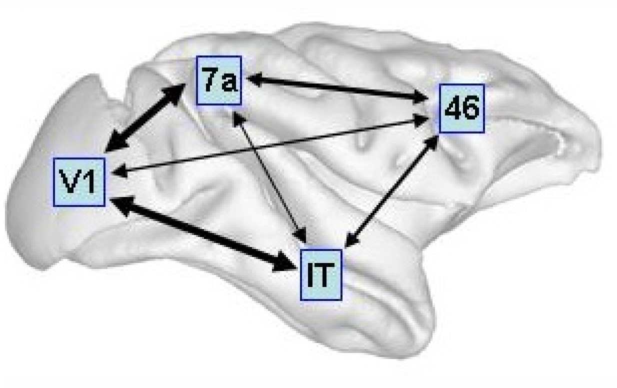
DIFFERENT FORMS OF BRAIN CONNECTIVITY

structural connectivity



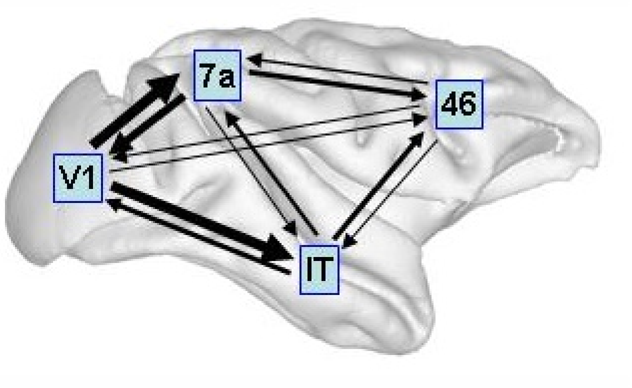
- presence of physical connections
- Diffusion weighted imaging (DWI), tractography, tracer studies

functional connectivity



- statistical dependencies between regional time series
- correlations, Independent Component Analysis (ICA)

effective connectivity

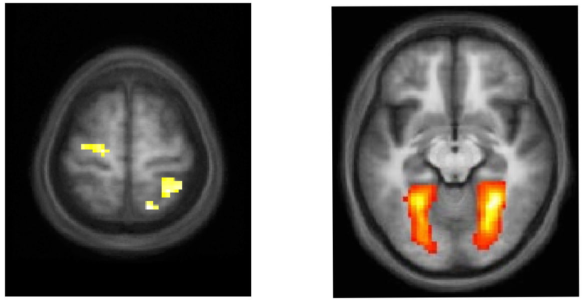


Ideally, we would like to have: Parameters of an underlying network model, including directed connectivity.

Sporns, 2007, *Scholarpedia*

From functional segregation to functional integration

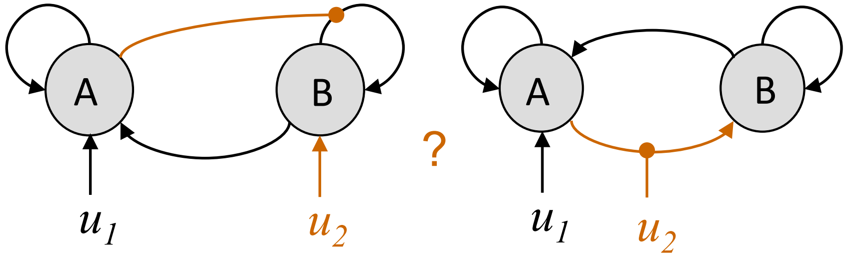
localizing brain activity:
functional segregation



u_1 $\qquad$ $u_1 X u_2$

« Where, in the brain, did my experimental manipulation have an effect? »

effective connectivity analysis:
functional integration



« How did my experimental manipulation propagate through the network? »

DYNAMIC CAUSAL MODELING

ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

NeuroImage 19 (2003) 1273–1302

NeuroImage
www.elsevier.com/locate/ynimg

Dynamic causal modelling

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

The Wellcome Department of Imaging Neuroscience, Institute of Neurology, Queen Square, London WC1N 3BG, UK

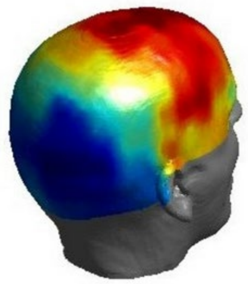
Received 18 October 2002; revised 7 March 2003; accepted 2 April 2003

- Dynamic causal modeling (DCM) for functional magnetic resonance imaging (fMRI) data was introduced in 2003 by Karl Friston, Lee Harrison and Will Penny (NeuroImage 19:1273-1302)

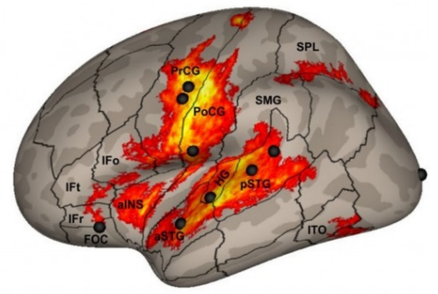
Friston et al., 2003, *NeuroImage*

DYNAMIC CAUSAL MODELING

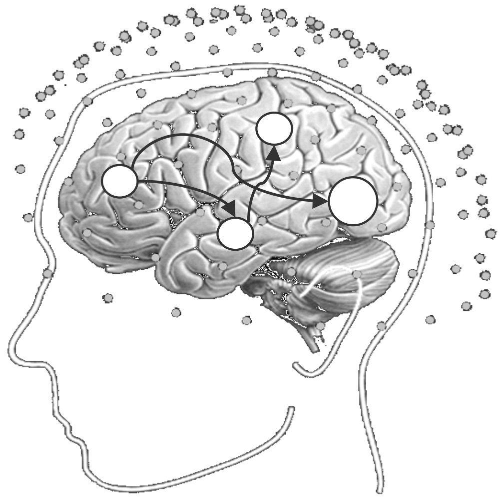
EEG, MEG



fMRT



http://sites.bu.edu/guentherlab/



Forward model:

Predicting measured activity

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

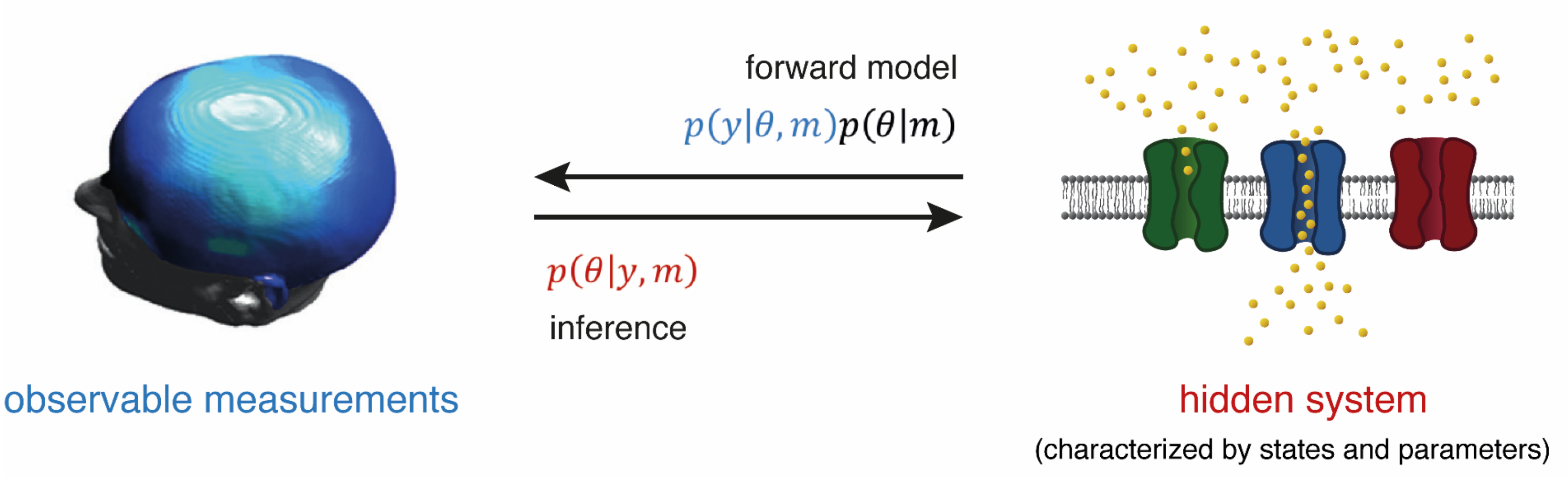
Model inversion:

Estimating neuronal mechanisms

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

Friston et al., 2003, *NeuroImage*; David et al., 2006, *NeuroImage*

GENERATIVE MODEL

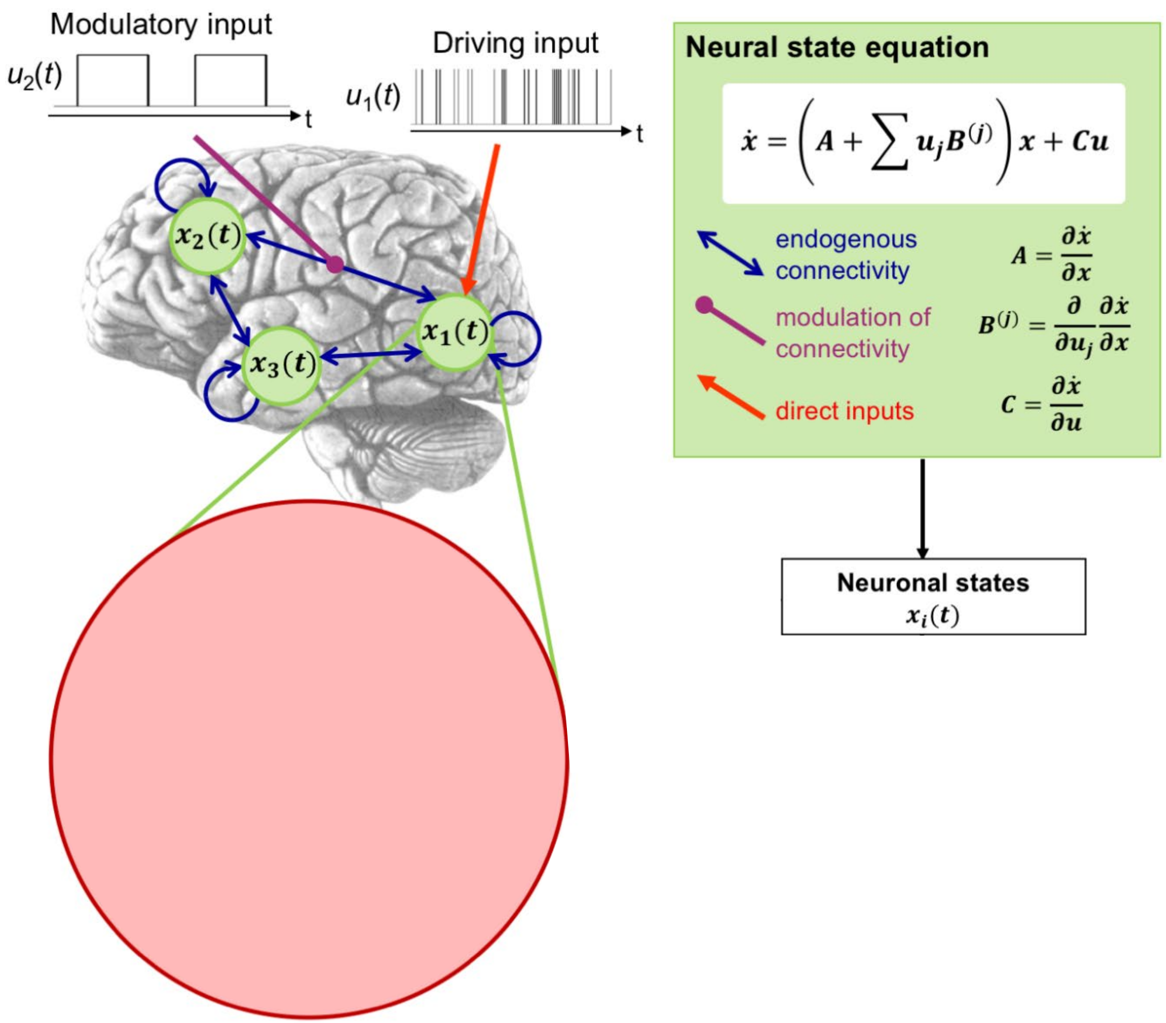


1. enforces mechanistic thinking: how could the data have been caused?
2. generate synthetic data (observations) by sampling from the prior – can the model explain certain phenomena at all?
3. inference about model structure: formal approach to disambiguating mechanisms → $p(m|y)$
4. inference about model parameters → $p(\theta|y,m)$

Stephan et al., 2016, *Front. Hum. Neurosci.*; Frässle et al., in press, *Wiley Interdiscip. Rev. Cogn. Sci.*

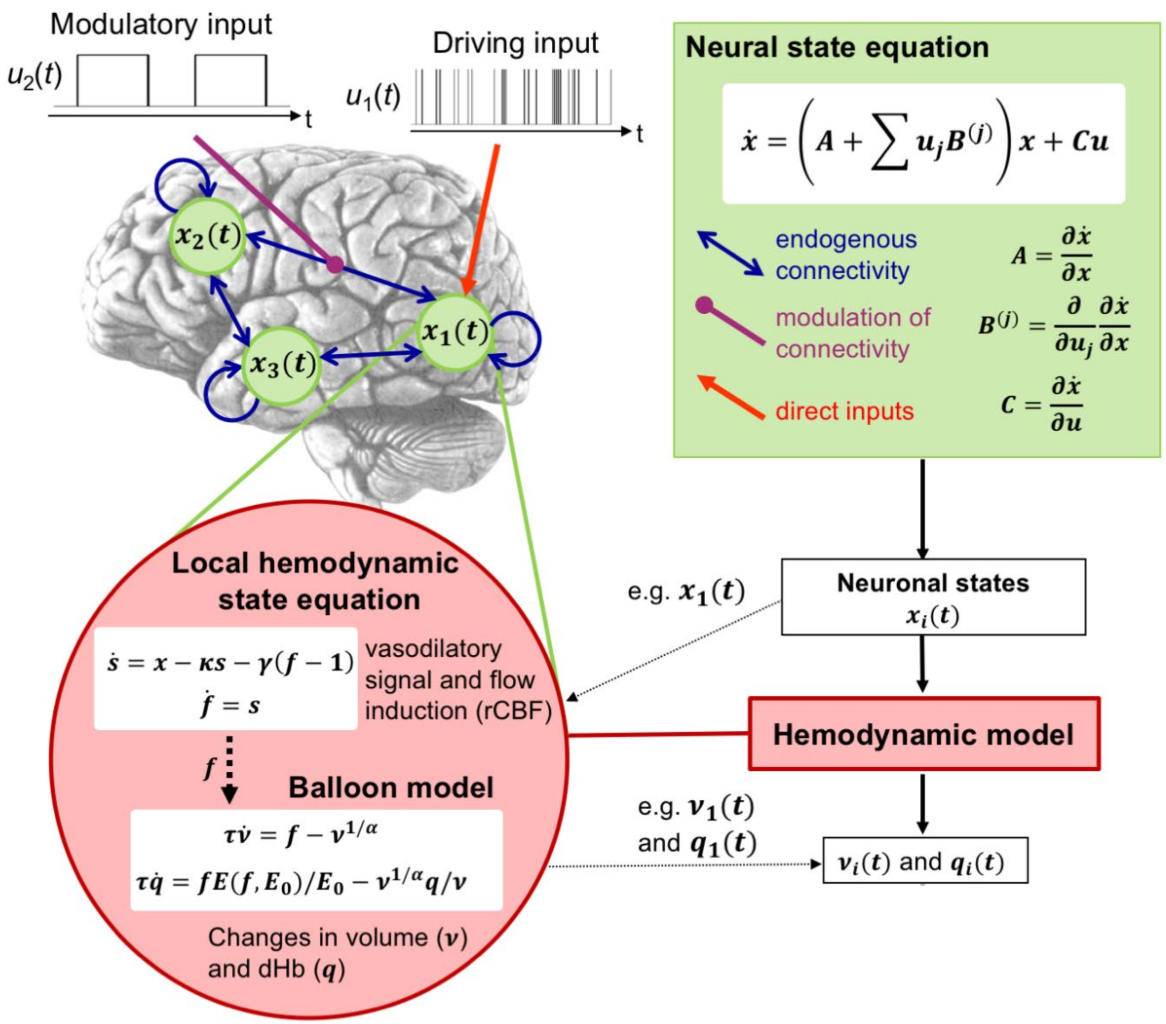
THEORY

DCM for fMRI (Overview)



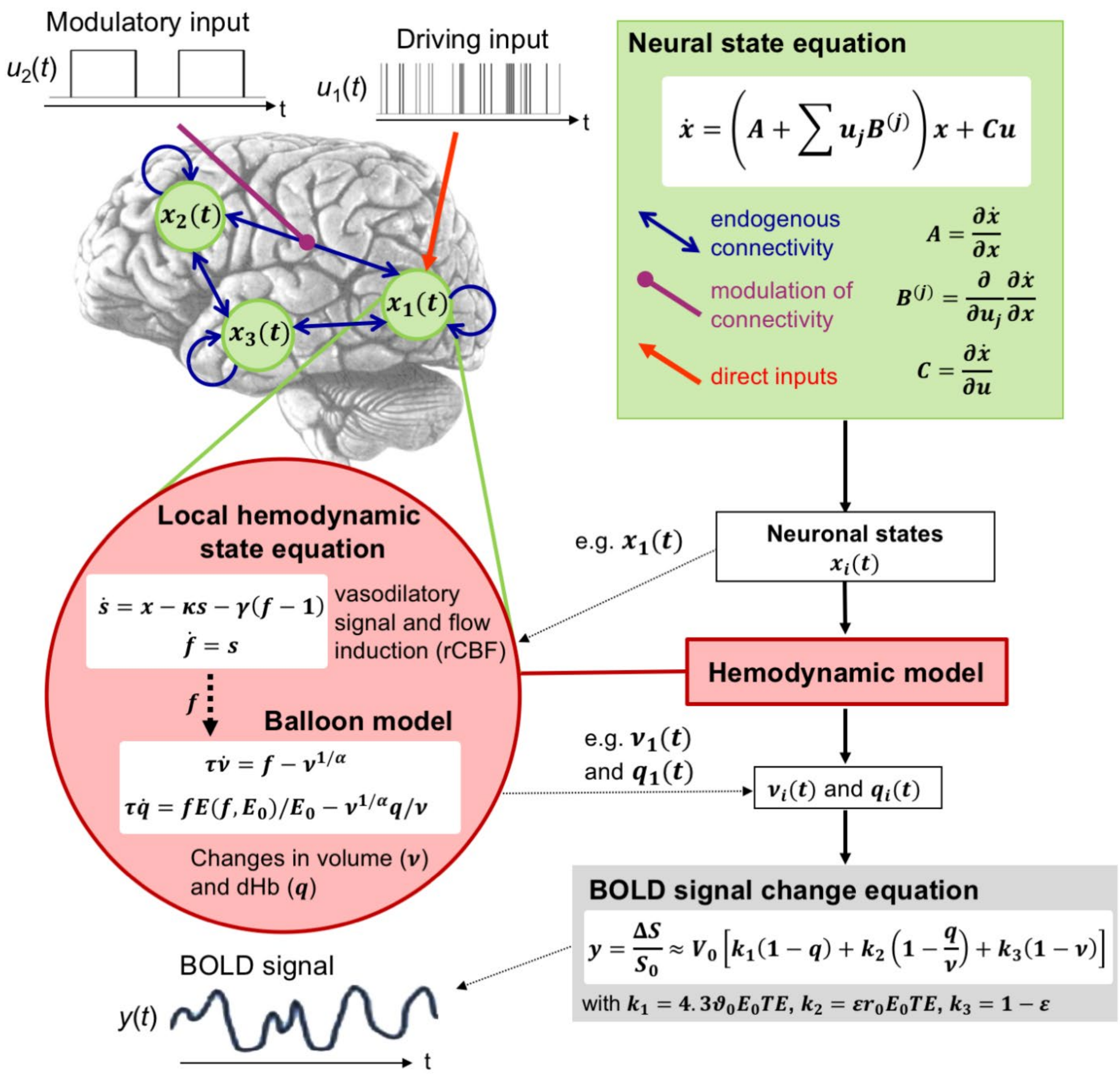
Friston et al., 2003, *NeuroImage*; Stephan et al., 2015, *Neuron*

DCM for fMRI (Overview)



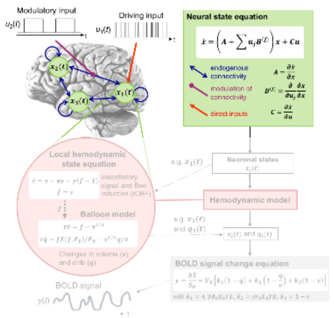
Friston et al., 2003, *NeuroImage*; Stephan et al., 2015, *Neuron*

DCM FOR FMRI (OVERVIEW)



Friston et al., 2003, *NeuroImage*; Stephan et al., 2015, *Neuron*

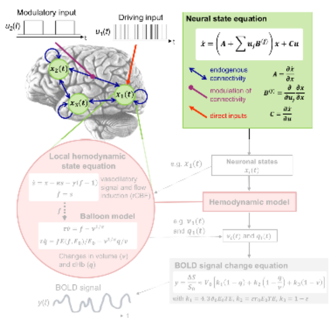
NEURONAL STATE EQUATION



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

Friston et al., 2003, *NeuroImage*; Stephan et al., 2008, *NeuroImage*

Neuronal state equation

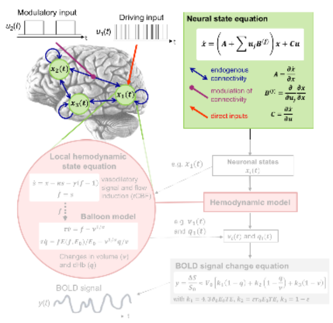


$$\frac{dx}{dt} = f(x,u) \approx \boxed{f(x_0,0) + \frac{\partial f}{\partial x}x + \frac{\partial f}{\partial u}u + \frac{\partial^2 f}{\partial x \partial u}ux} + \frac{\partial^2 f}{\partial x^2}\frac{x^2}{2} + \cdots$$

bilinear model

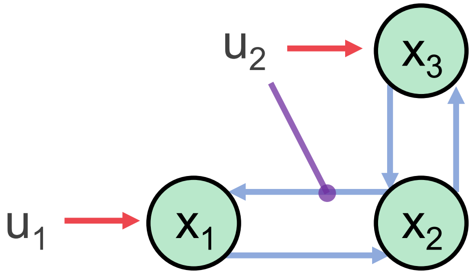
Friston et al., 2003, *NeuroImage*; Stephan et al., 2008, *NeuroImage*

Neuronal state equation

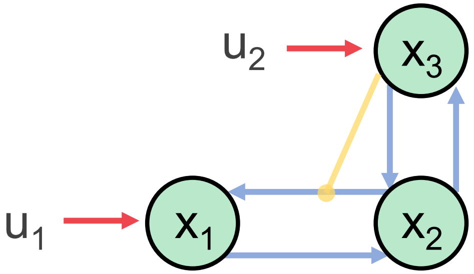


$$\frac{dx}{dt} = f(x,u) \approx f(x_0,0) + \overset{A}{\frac{\partial f}{\partial x}}x + \overset{C}{\frac{\partial f}{\partial u}}u + \overset{B}{\frac{\partial^2 f}{\partial x \partial u}}ux + \overset{D}{\frac{\partial^2 f}{\partial x^2}}\frac{x^2}{2} + \cdots$$

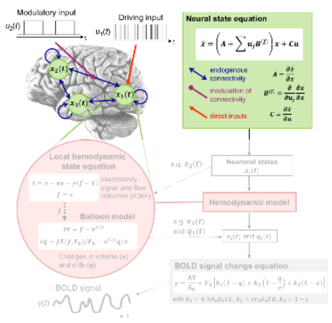
bilinear model



nonlinear model

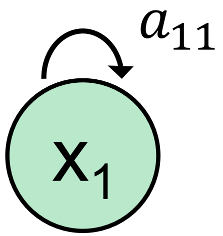


Friston et al., 2003, *NeuroImage*; Stephan et al., 2008, *NeuroImage*

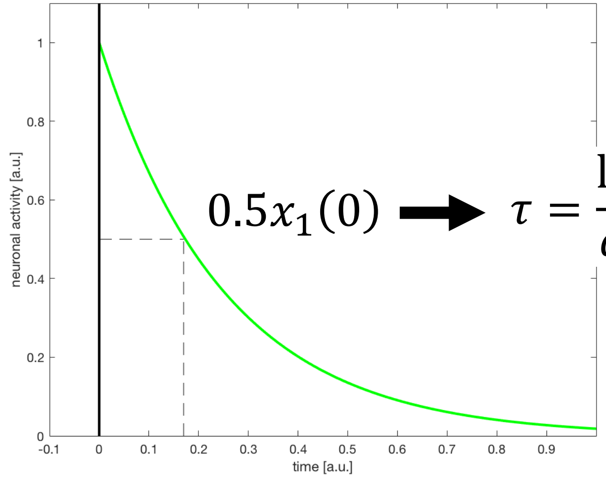


Neuronal state equation

DCM effective connectivity parameters are rate constants

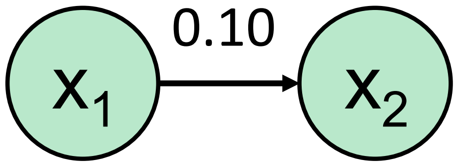


$$\frac{dx_1}{dt} = a_{11}x_1 \longrightarrow x_1(t) = x_1(0) \cdot exp(a_{11}t)$$



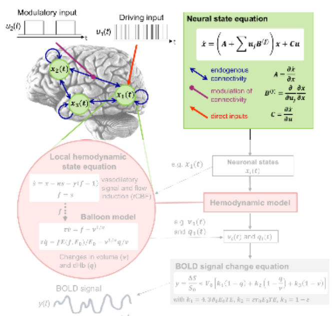
$$0.5x_1(0) \rightarrow \tau = \frac{\ln 2}{a_{11}}$$

Friston et al., 2003, *NeuroImage*

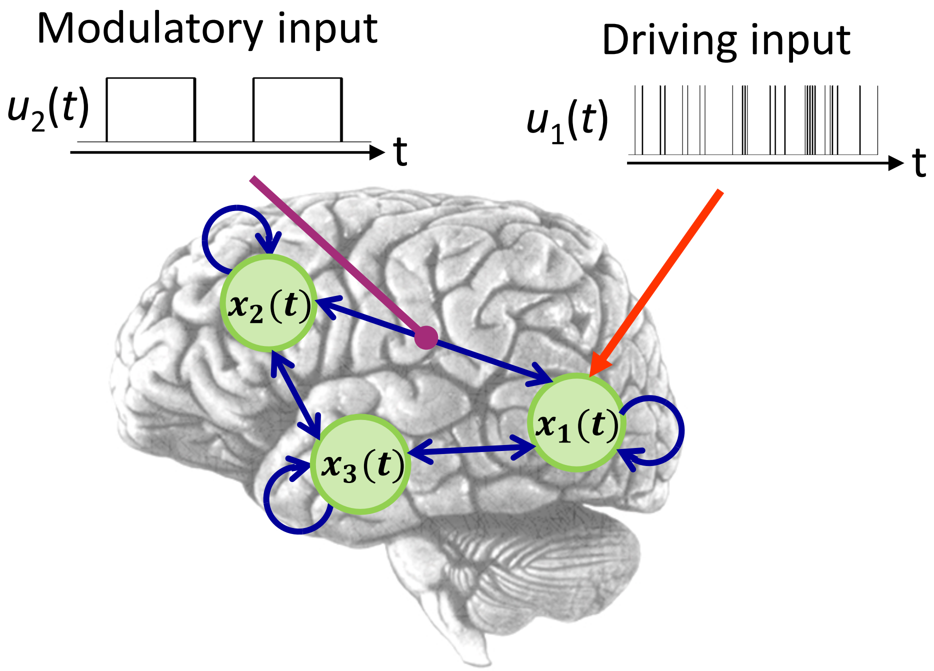


If region$_1$ → region$_2$ is 0.10s^{-1}, this means that, per unit time, the increase in activity in region$_2$ corresponds to 10% of the current activity in region$_1$

Neuronal state equation



Interim summary: bilinear neuronal state equation



Friston et al., 2003, *NeuroImage*

State change

Current state

External inputs

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

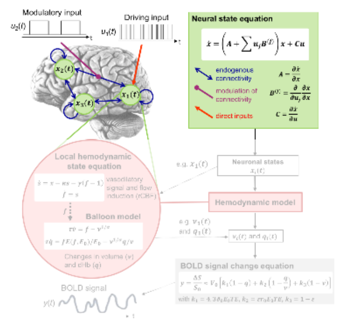
$$\theta = \{A, B^{(1)}, \ldots, B^{(m)}, C\}$$

Endogenous connectivity

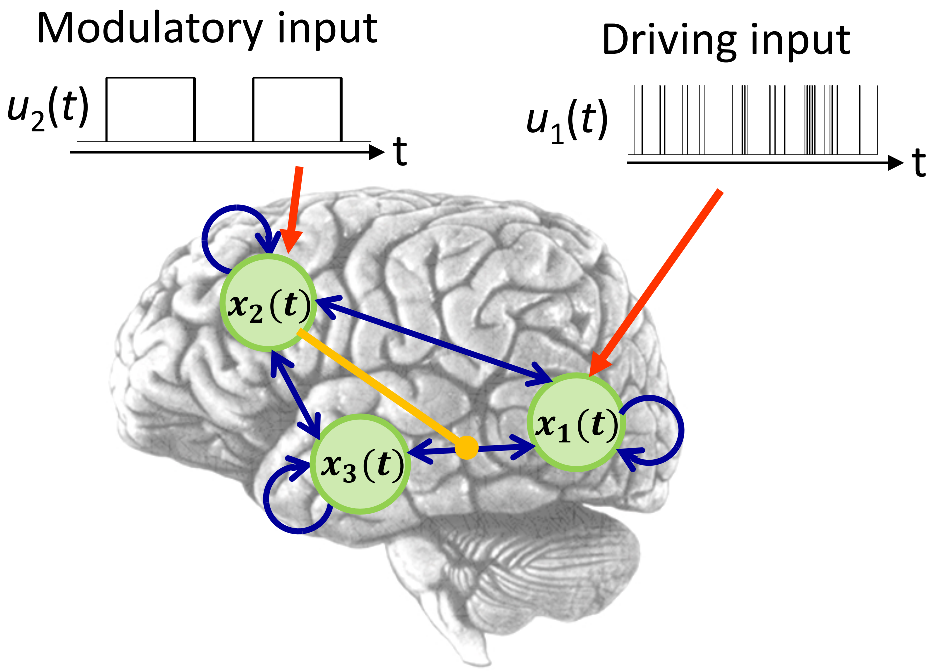
Modulatory connectivity

Driving inputs

Neuronal state equation



Interim summary: non-linear neuronal state equation



Friston et al., 2003, *NeuroImage*

State change

Current state

External inputs

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

$$\theta = \{A, B^{(1)}, \dots, B^{(m)}, C, D^{(1)}, \dots, D^{(n)}\}$$

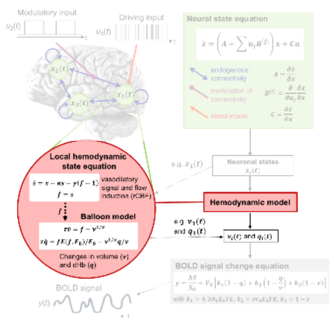
Endogenous connectivity

Modulatory connectivity

Driving inputs

Modulatory connectivity (non-linear)

HEMODYNAMIC MODEL



Neuronal dynamics only indirectly observable via hemodynamic response

↑ neuronal activity

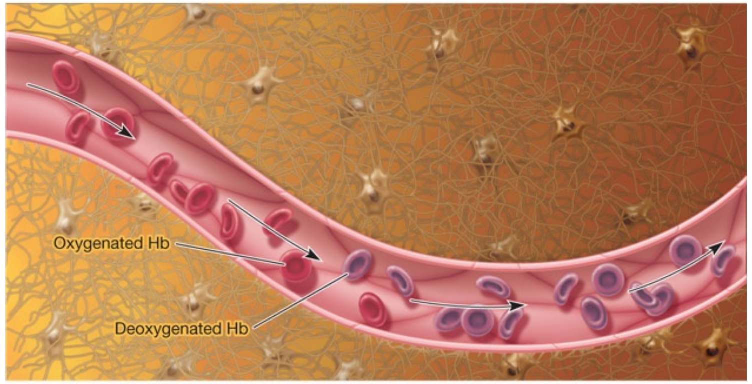
↑ blood flow

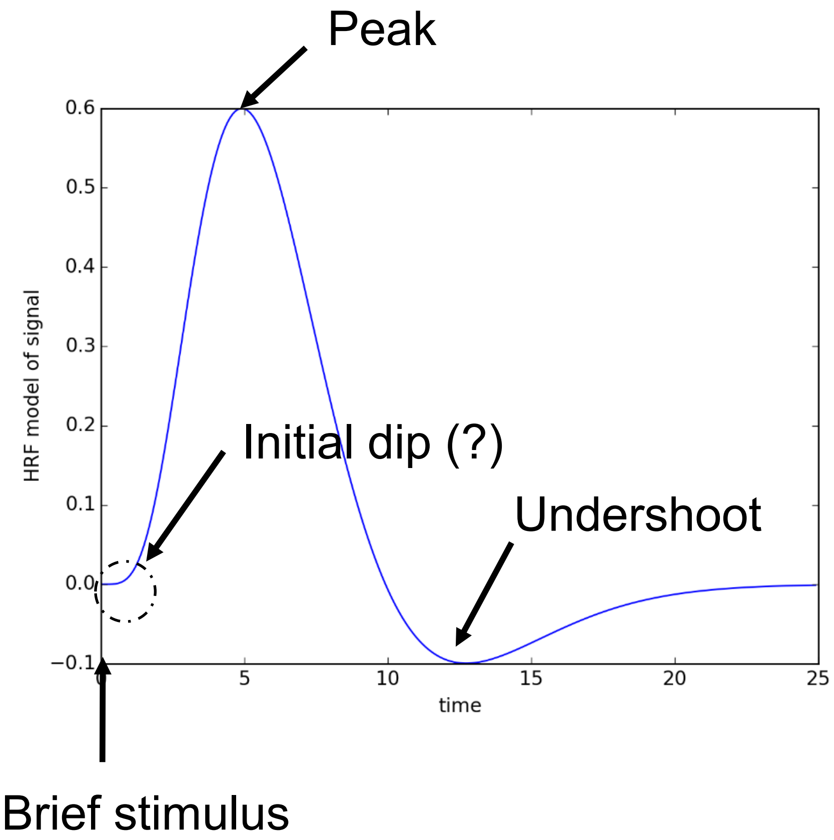
↑ oxygenated Hb

↑ T2*

↑ fMRI signal

Rest





Huettel et al., 2004, *NeuroImage*

HEMODYNAMIC MODEL

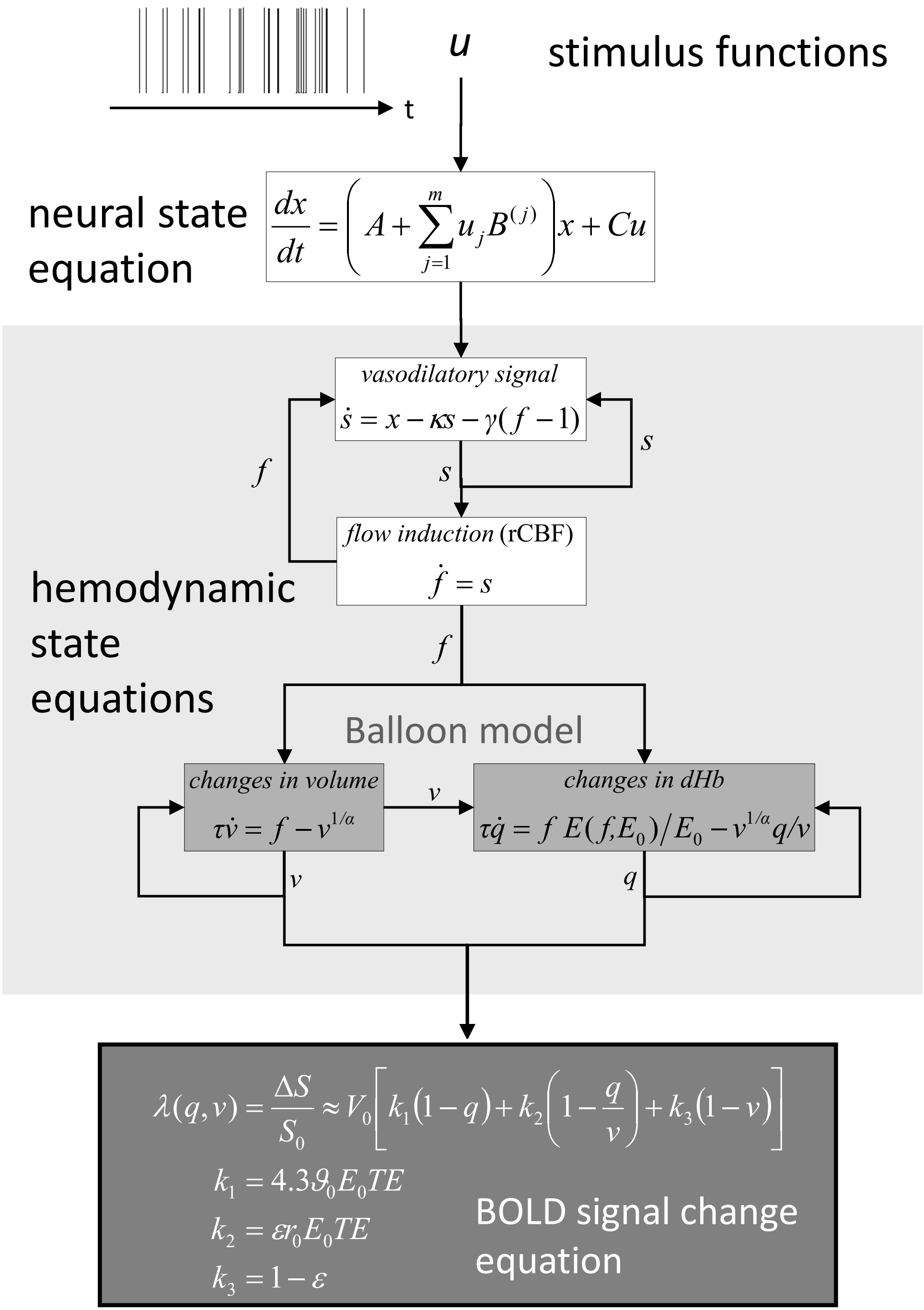
6 hemodynamic parameters:

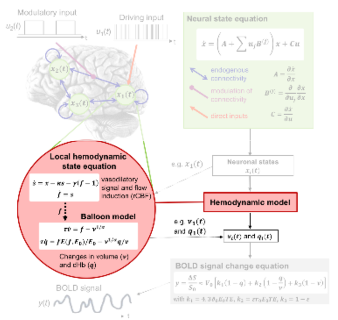
$$\theta^h = \{\kappa, \gamma, \tau, \alpha, \rho, \varepsilon\}$$

Important for model fitting, but typically of no interest for statistical inference.

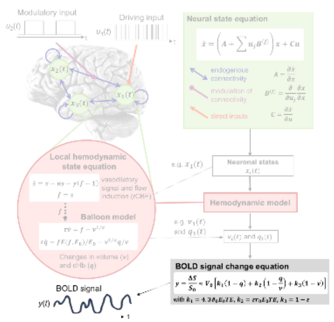
Parameters (κ,τ and ε) are computed separately for each region → region specific HRFs!

Friston et al., 2003, *NeuroImage*; Stephan et al., 2007, *NeuroImage*





BOLD SIGNAL CHANGE EQUATION



Resting blood volume

Deoxyhemoglobin content

Blood volume

$$\lambda(q,\nu) = \frac{\Delta S}{S_0} \approx V_0\left[k_1(1-q) + k_2\left(1-\frac{q}{\nu}\right) + k_3(1-\nu)\right]$$

$$k_1 = 4.3\vartheta_0 E_0 TE$$

$$k_2 = \varepsilon r_0 E_0 TE$$

$$k_3 = 1-\varepsilon$$

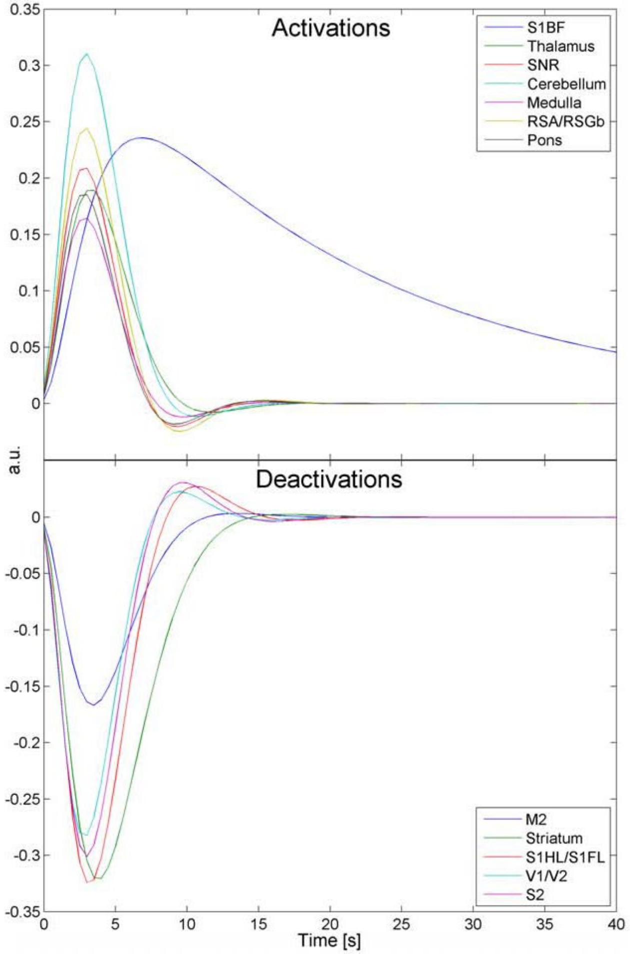
Friston et al., 2003, *NeuroImage*; Stephan et al., 2007, *NeuroImage*

Parameters

$V_0 = 0.04 \qquad E_0 = 0.4$

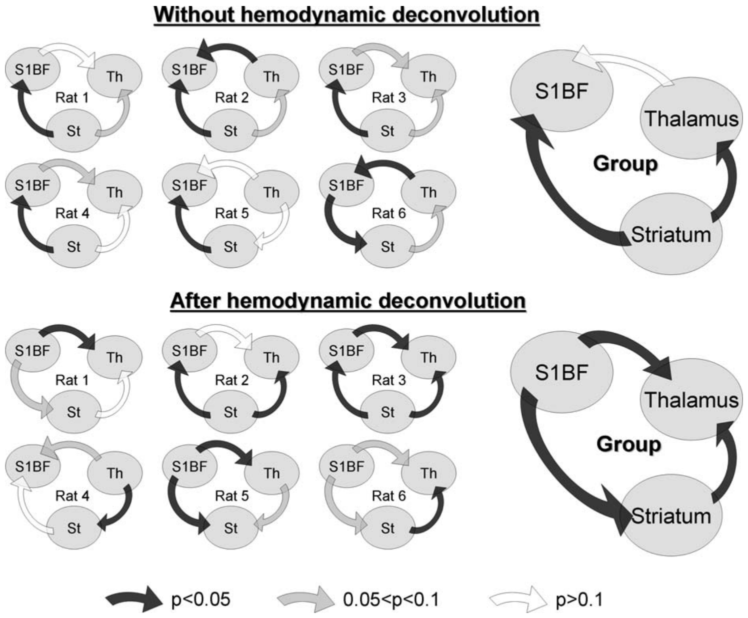
At 1.5 Tesla	At 3 Tesla	At 7 Tesla
$\vartheta_0 \approx 40.3\ \mathrm{s}^{-1}$	$\vartheta_0 \approx 80.6\ \mathrm{s}^{-1}$	$\vartheta_0 \approx 188\ \mathrm{s}^{-1}$
$r_0 \approx 25\ \mathrm{s}^{-1}$	$r_0 \approx 110\ \mathrm{s}^{-1}$	$r_0 \approx 340\ \mathrm{s}^{-1}$
$TE \approx 0.04\ s$	$TE \approx 0.035\ s$	$TE \approx 0.025\ s$
$\varepsilon \approx 1.28$	$\varepsilon \approx 0.47$	$\varepsilon \approx 0.026$

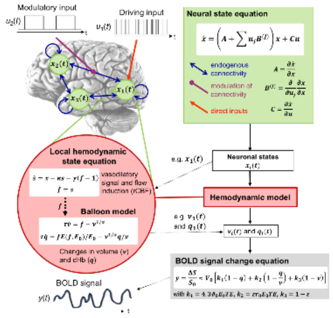
HEMODYNAMICS ARE IMPORTANT



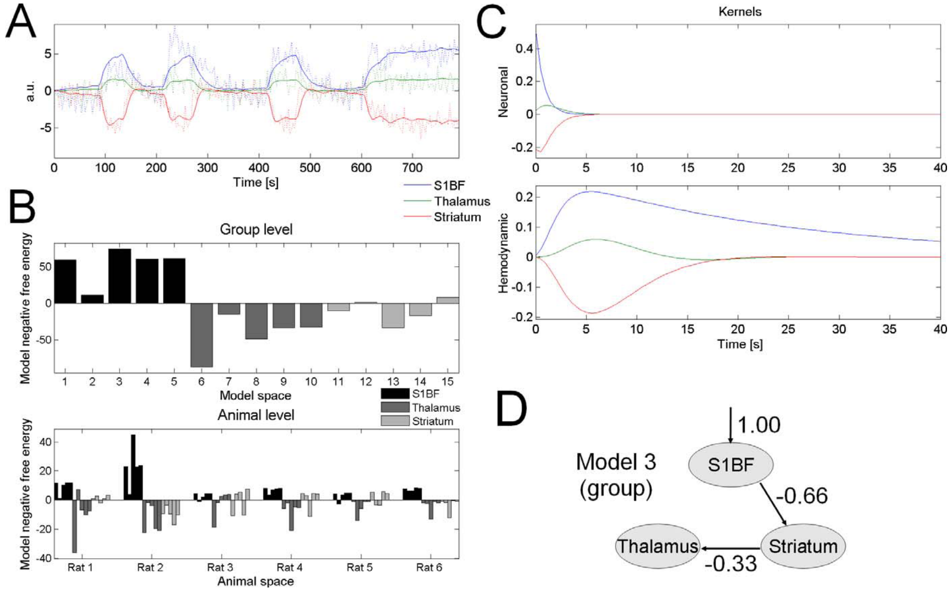
David et al., 2008, *PLoS Biol.*

Granger causality

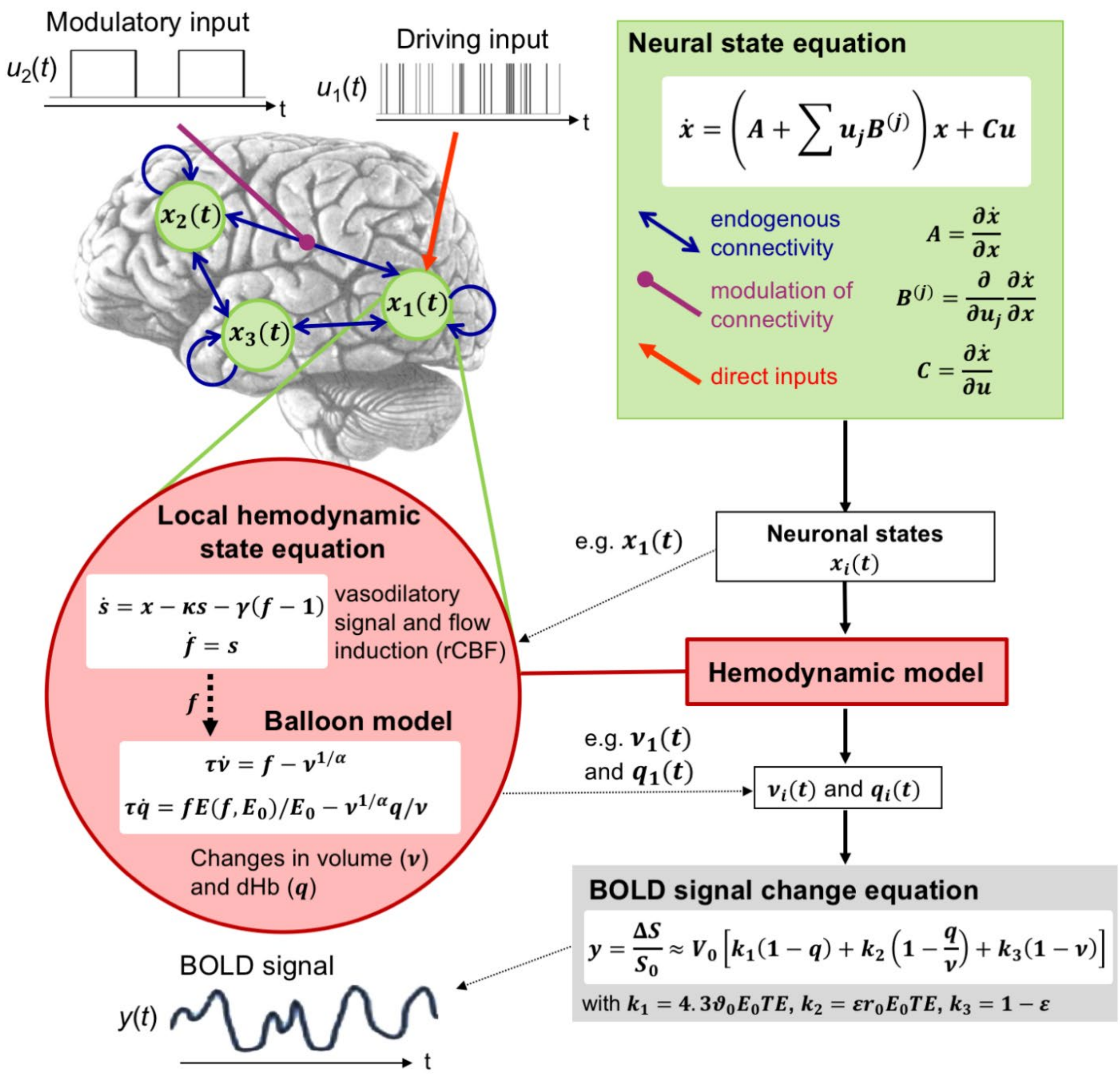




DCM



DCM for fMRI (Overview)

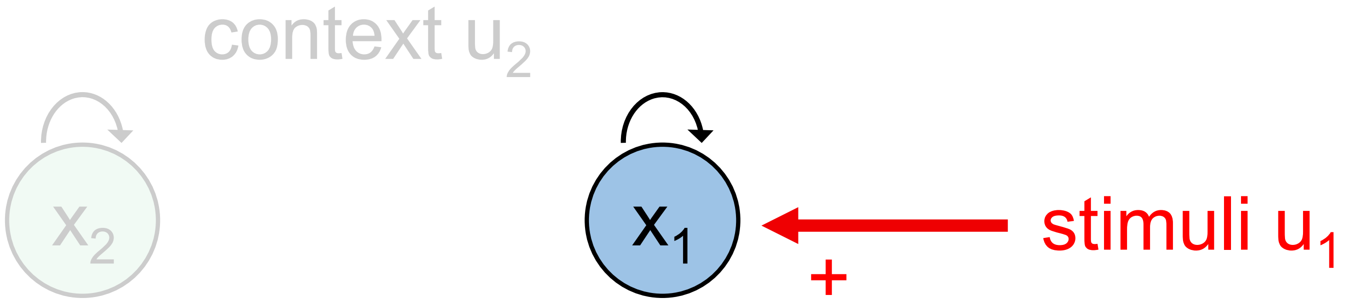


Friston et al., 2003, *NeuroImage*; Stephan et al., 2015, *Neuron*

SIMULATIONS

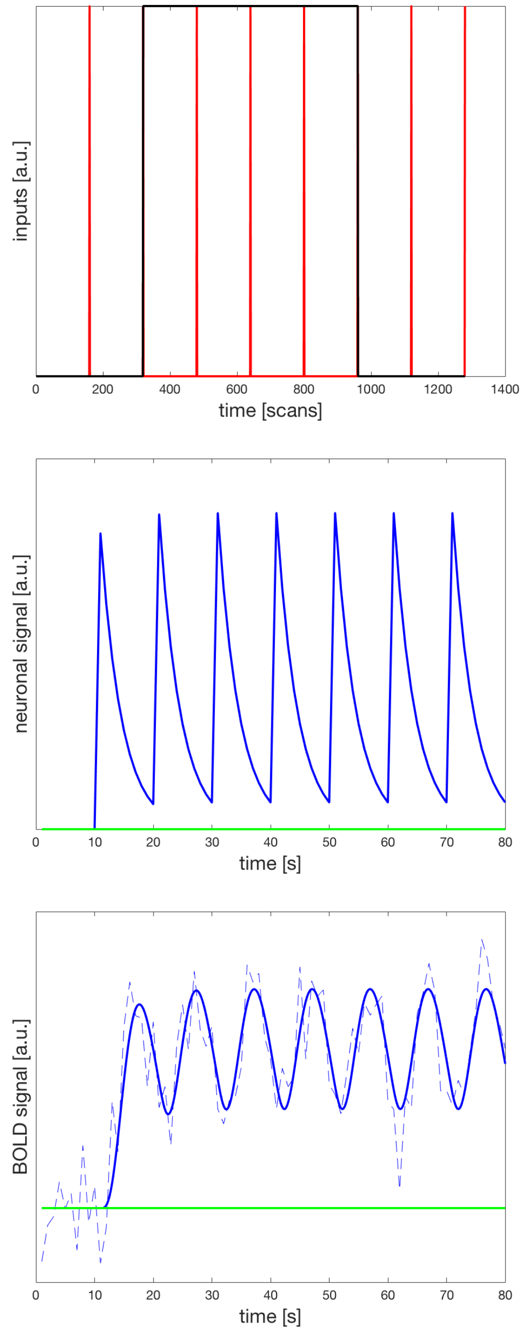
What can DCM explain?

Example: single node



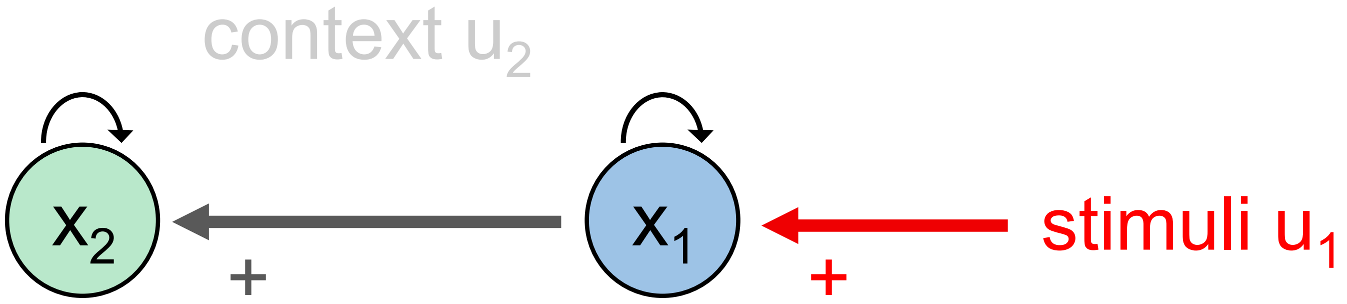
$$\frac{dx}{dt} = Ax + u_2 B^{(2)} x + C u_1$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + u_2 \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} c_{11} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$



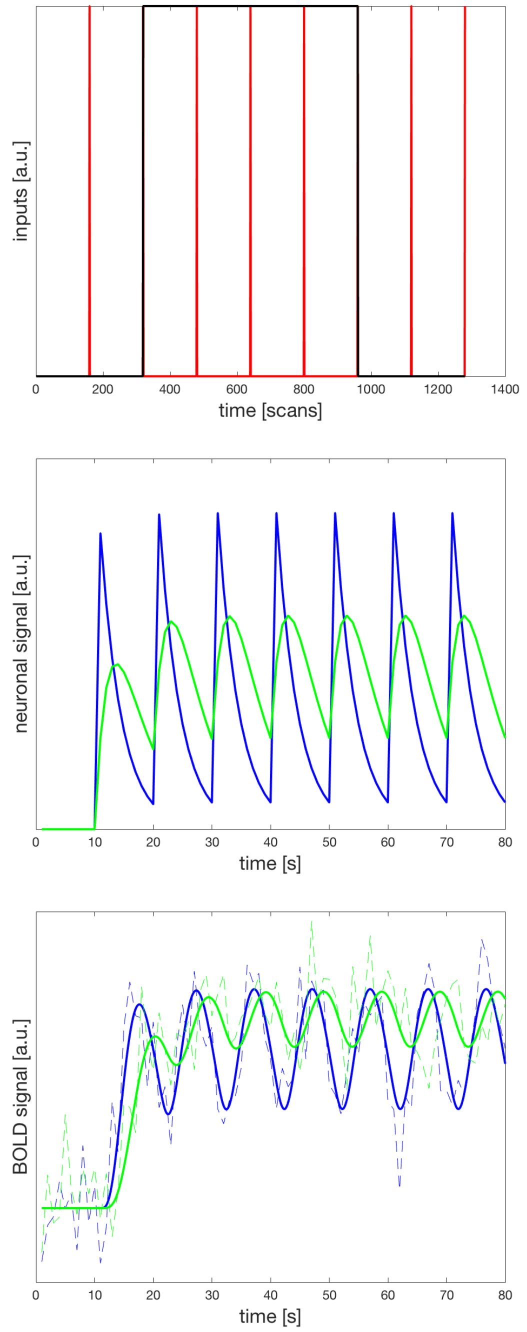
What can DCM explain?

Example: two connected node



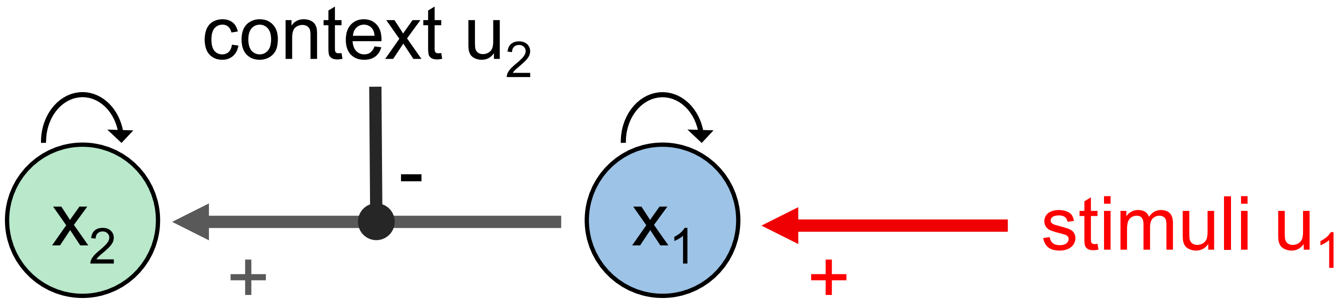
$$\frac{dx}{dt} = Ax + u_2 B^{(2)} x + C u_1$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + u_2 \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} c_{11} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$



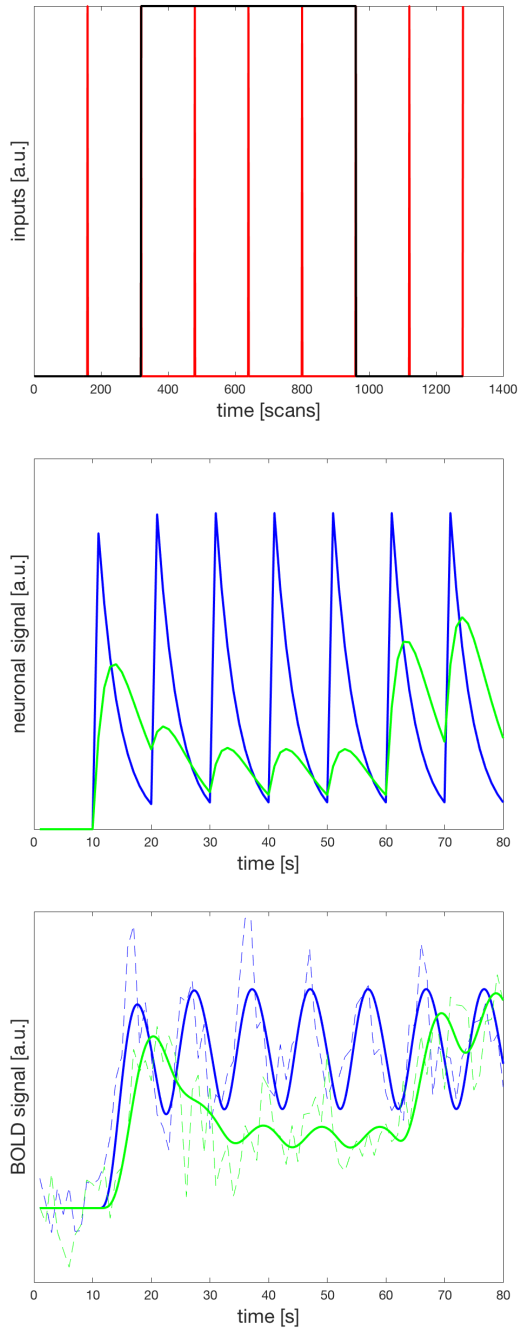
What can DCM explain?

Example: modulation of connection



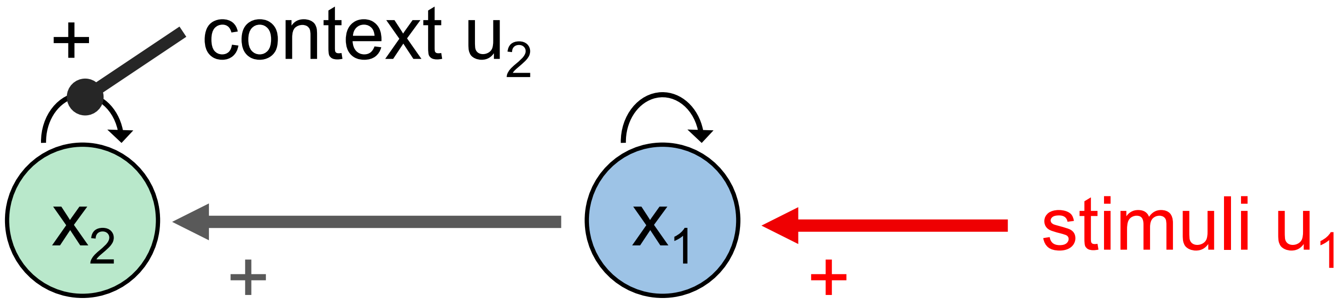
$$\frac{dx}{dt} = Ax + u_2 B^{(2)} x + C u_1$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + u_2 \begin{bmatrix} 0 & 0 \\ b_{21}^{(2)} & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} c_{11} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$



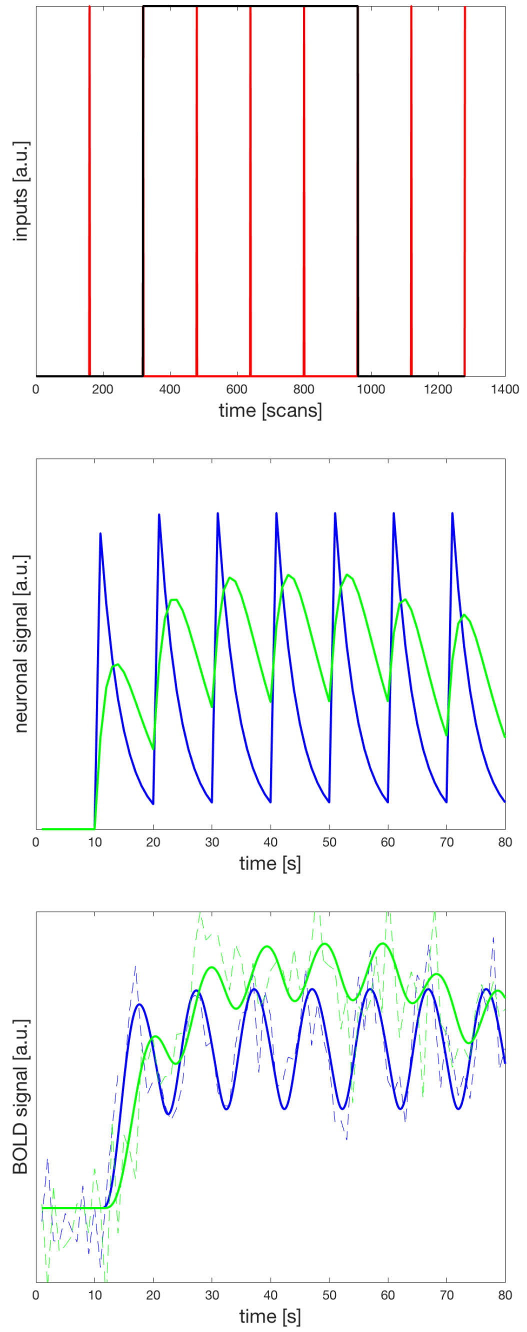
What can DCM explain?

Example: modulation of inhibitory self-connection



$$\frac{dx}{dt} = Ax + u_2 B^{(2)} x + C u_1$$

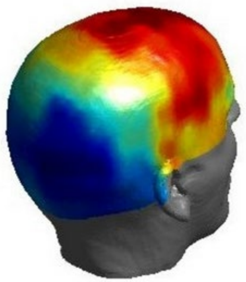
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + u_2 \begin{bmatrix} 0 & 0 \\ 0 & b_{22}^{(2)} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} c_{11} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$



Model inversion / Inference

Dynamic causal modeling

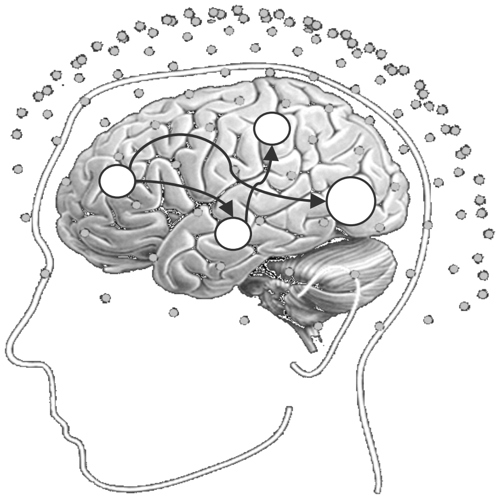
EEG, MEG



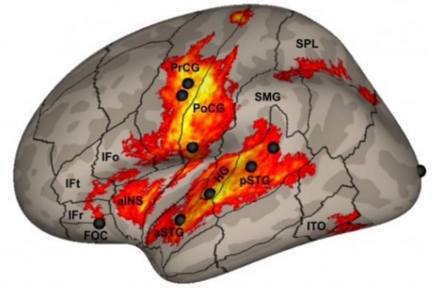
Forward model:

Predicting measured activity

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



fMRT



http://sites.bu.edu/guentherlab/

Model inversion:

Estimating neuronal mechanisms

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

Friston et al., 2003, *NeuroImage*; David et al., 2006, *NeuroImage*

BAYES THEOREM

Reverend Thomas Bayes
(1702-1761)

Bayes theorem gives a recipe for evaluating the posterior density by combining new data (likelihood) and prior knowledge

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

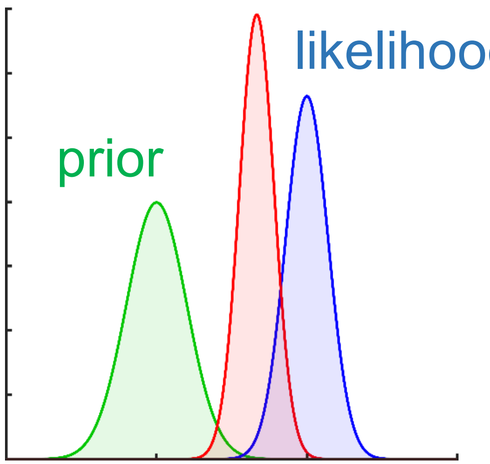
The posterior probability of the parameters is an optimal combination of our prior knowledge and the new data that we have acquired

BAYES THEOREM

Reverend Thomas Bayes
(1702-1761)

Bayes theorem gives a recipe for evaluating the posterior density by combining new data (likelihood) and prior knowledge

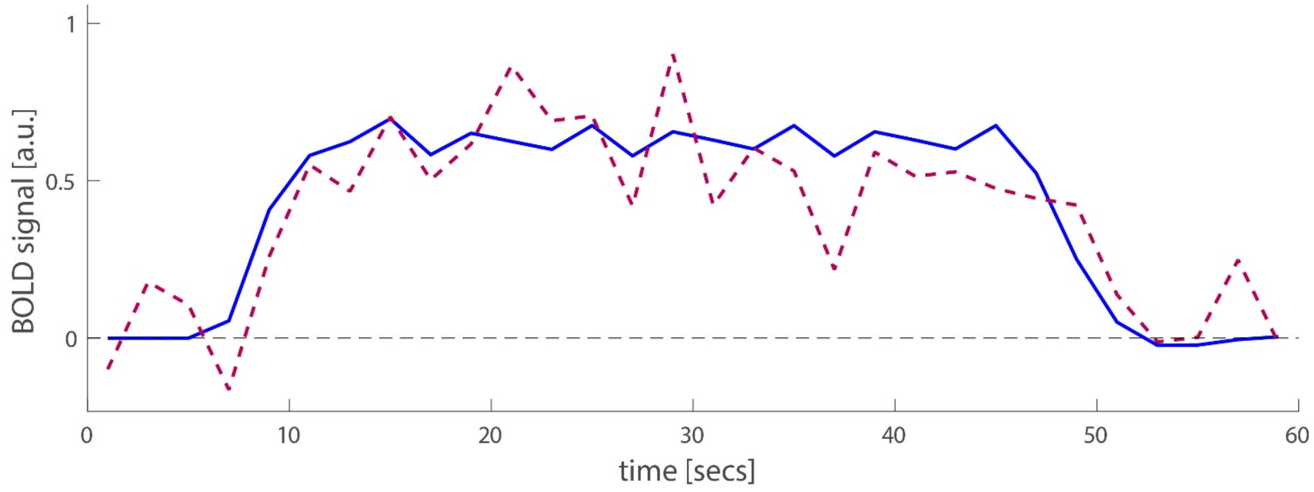
$$\underbrace{p(\theta|y,m)}_{\text{posterior}} = \frac{\overbrace{p(y|\theta,m)}^{\text{likelihood}}\,\overbrace{p(\theta|m)}^{\text{prior}}}{\underbrace{p(y|m)}_{\text{model evidence}}}$$



The posterior probability of the parameters is an optimal combination of our prior knowledge and the new data that we have acquired

LIKELIHOOD FUNCTION

$$\underbrace{p(y(t)|\theta, m)}_{\text{likelihood}} = \mathcal{N}\left(y(t); g(\theta^n, \theta^h, u), \theta^\sigma\right)$$

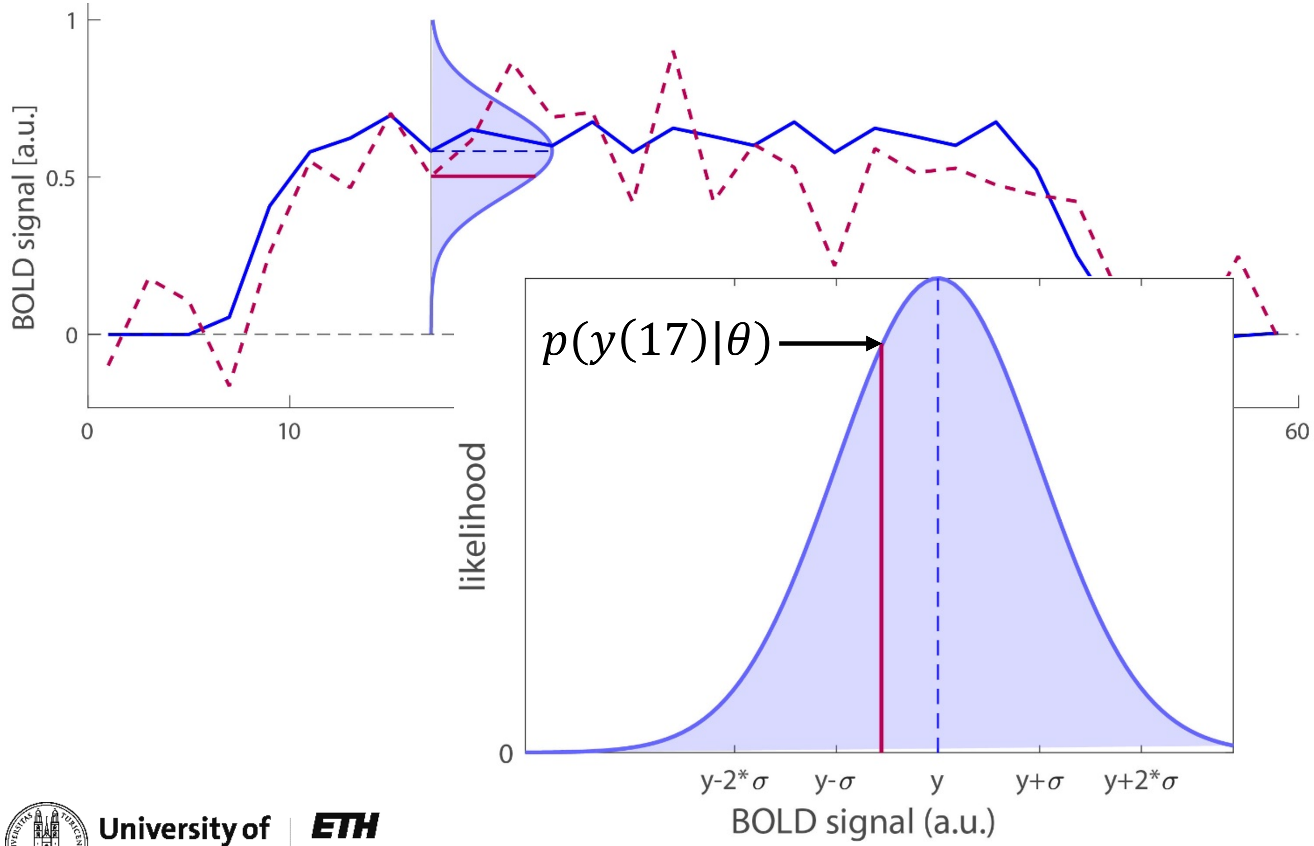


Assume data is normally distributed around the prediction from the dynamical model (Gaussian noise):

LIKELIHOOD FUNCTION

$$p(y(t)|\theta, m) = \mathcal{N}\left(y(t); g\left(\theta^{n}, \theta^{h}, u\right), \theta^{\sigma}\right)$$

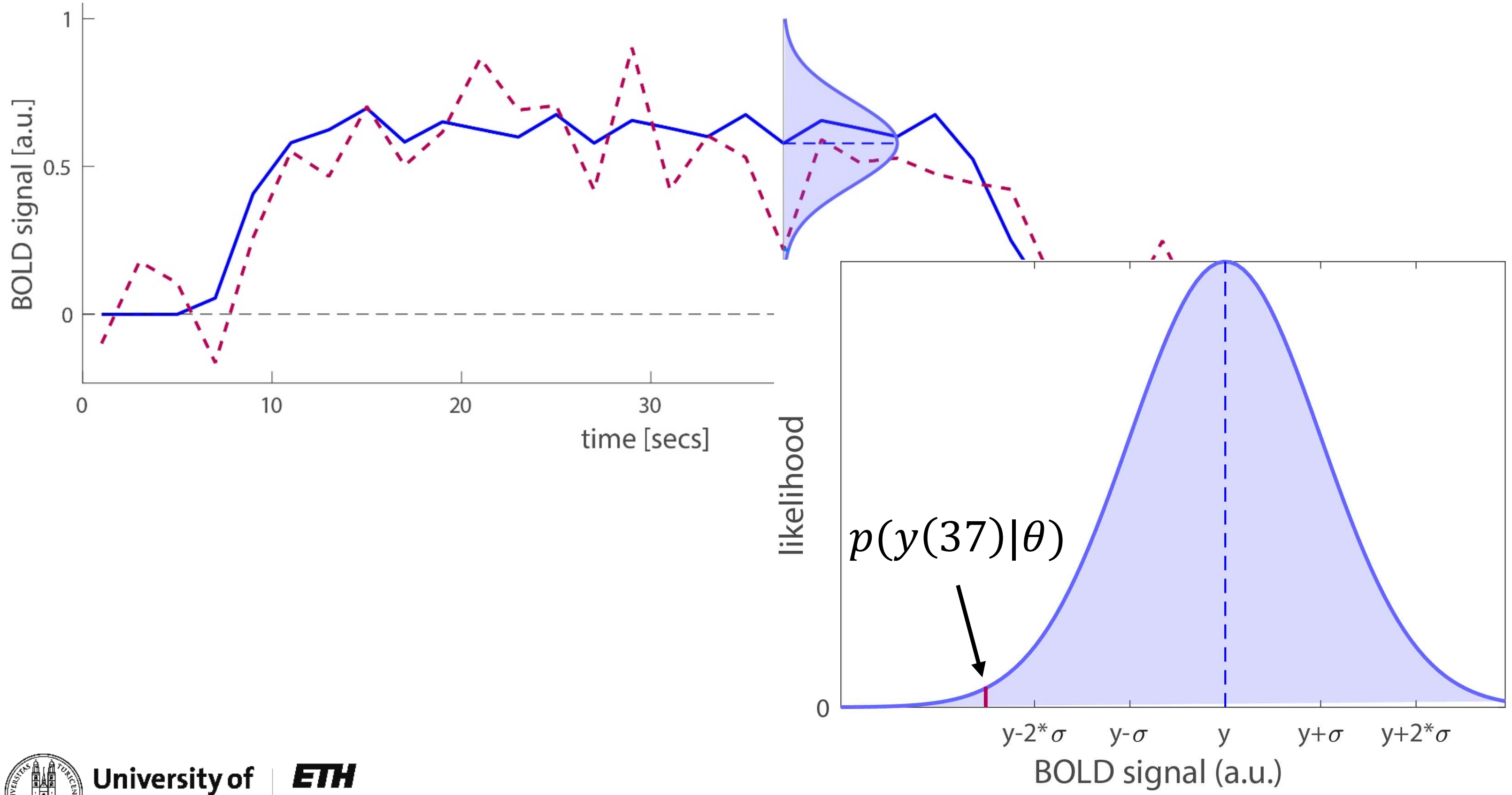
likelihood



LIKELIHOOD FUNCTION

$$p(y(t)|\theta, m) = \mathcal{N}\left(y(t); g\left(\theta^n, \theta^h, u\right), \theta^\sigma\right)$$

likelihood



Priors

Bayes theorem gives a recipe for evaluating the posterior density by combining new data (likelihood) and prior knowledge

$$p(\theta|y,m) = \frac{p(y|\theta,m)\overbrace{p(\theta|m)}^{\text{prior}}}{p(y|m)}$$

Neuronal parameters:

- self-connections: principled (to ensure that the system is stable)
- other parameters (between—region connections, modulation, inputs): shrinkage priors

Hemodynamic parameters:

- empirical

PRIORS

Types of priors:

– Explicit priors on *model parameters* (e.g., connection strengths)
– Implicit priors on *model functional form* (e.g., system dynamics)
– Choice of “interesting” *data features* (e.g., regional time-series vs. ICA analysis)

Role of priors (on model parameters):

– Resolving the *ill-posedness* of the inverse problem
– Avoiding *overfitting* (cf. generalization error)

Impact of priors:

– On parameter posterior distributions (cf. “shrinkage to the mean” effect)
– On model evidence (cf. “Occam’s razor”)
– On free-energy landscape (cf. Laplace approximation)

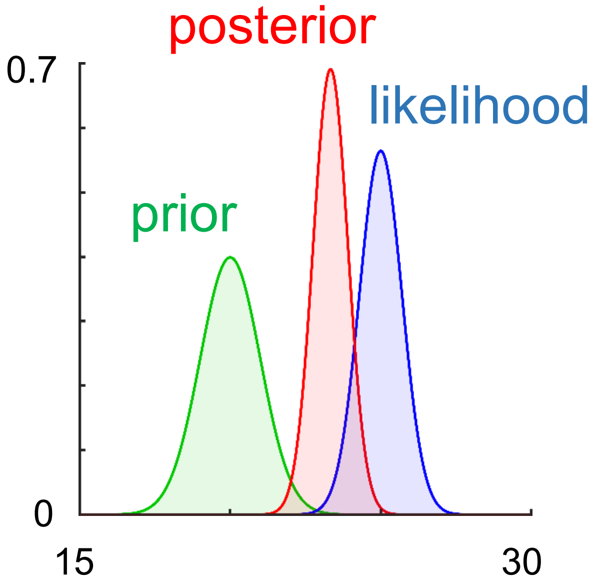
BAYES THEOREM

Reverend Thomas Bayes
(1702-1761)

Bayes theorem gives a recipe for evaluating the posterior density by combining new data (likelihood) and prior knowledge

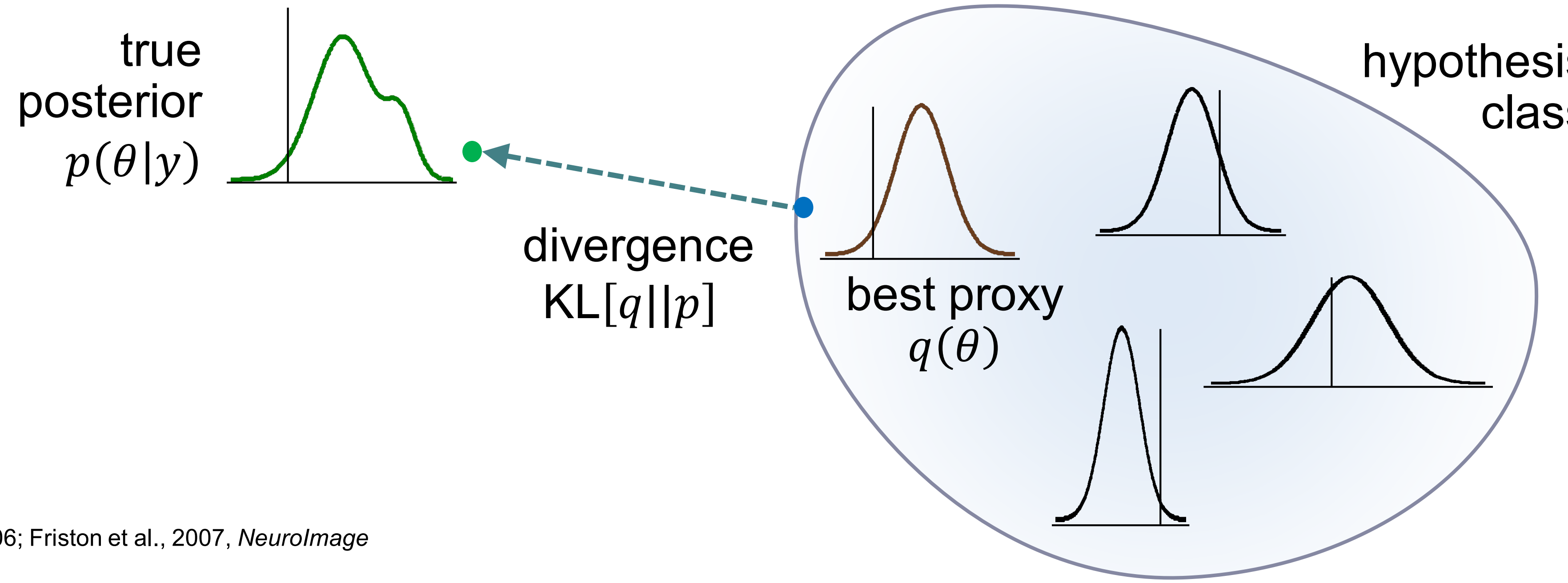
$$\underbrace{p(\theta|y,m)}_{\text{posterior}} = \frac{\overbrace{p(y|\theta,m)}^{\text{likelihood}}\overbrace{p(\theta|m)}^{\text{prior}}}{\underbrace{p(y|m)}_{\text{model evidence}}}$$

The posterior probability of the parameters is an optimal combination of our prior knowledge and the new data that we have acquired



Variational Bayes (VB)

Idea: find an approximate density $q(\theta)$ that is maximally similar to the true posterior $p(\theta|y)$. This is often done by assuming a particular form for q (fixed form VB) and then optimizing its sufficient statistics.



Bishop, 2006; Friston et al., 2007, *NeuroImage*

NEGATIVE FREE ENERGY

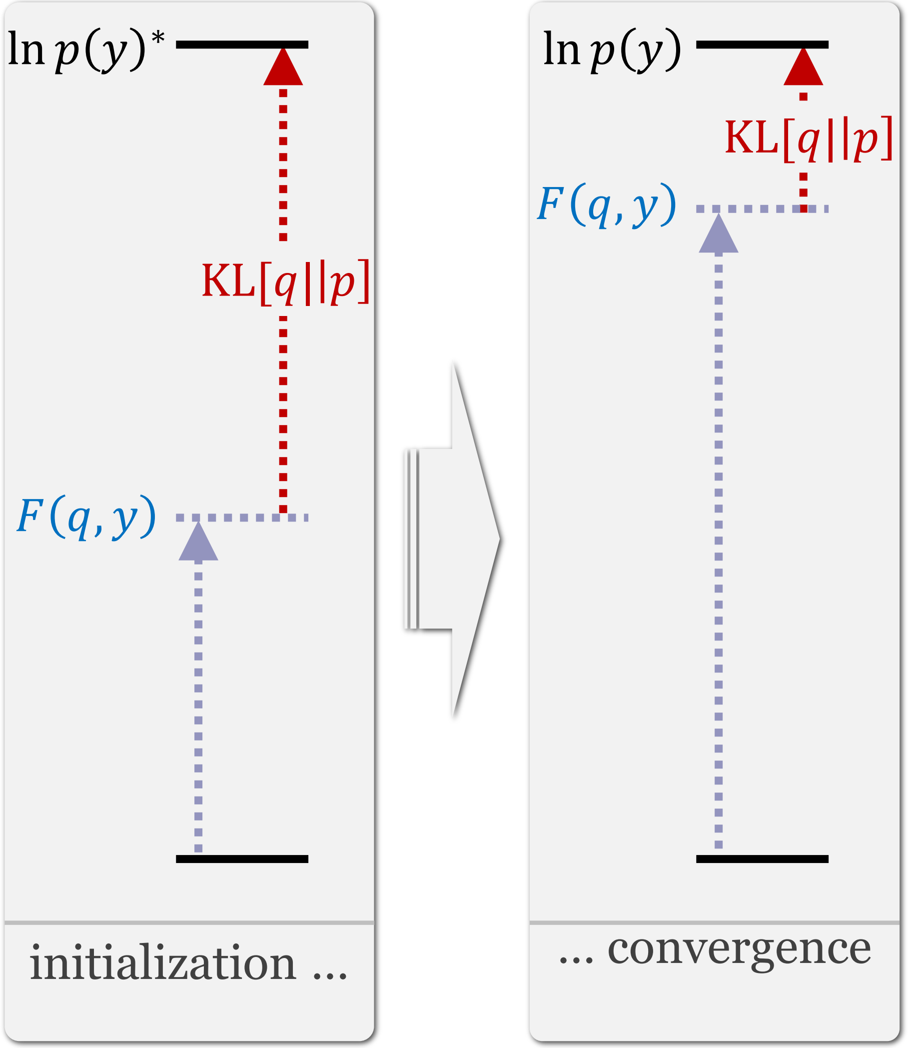
$$\ln p(y) = \underbrace{\mathrm{KL}[q||p]}_{\substack{\textbf{divergence} \geq 0 \\ \text{(unknown)}}} + \underbrace{F(q,y)}_{\substack{\textbf{neg. free energy} \\ \text{(easy to evaluate} \\ \text{for a given } q)}}$$

$F(q,y)$ is a functional with respect to the approximate posterior $q(\theta)$.

Maximizing $F(q,y)$ is equivalent to:

- minimizing $KL[q||p]$
- tightening $F(q,y)$ as a lower bound on the log model evidence

When $F(q,y)$ is maximized, $q(\theta)$ is our best estimate of the true posterior.



Bishop, 2006; Friston et al., 2007, *NeuroImage*

NEGATIVE FREE ENERGY – A CLOSER LOOK

The **negative free energy** represents a trade-off between the accuracy and complexity of a model:

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

Bishop, 2006; Friston et al., 2007, *NeuroImage*

Negative free energy – a closer look

The **negative free energy** represents a trade-off between the accuracy and complexity of a model:

$$F = \underbrace{\langle \log p(y|\theta,m)\rangle_q}_{\text{accuracy}} - \underbrace{KL[q(\theta)||p(\theta|m)]}_{\text{complexity}}$$

accuracy
(expected log likelihood)

complexity
(KL divergence between approximate posterior and prior)

Bishop, 2006; Friston et al., 2007, *NeuroImage*

NEGATIVE FREE ENERGY – A CLOSER LOOK

The **negative free energy** represents a trade-off between the accuracy and complexity of a model:

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

In contrast to “simple” criteria (e.g., AIC & BIC), the complexity term of the negative free energy accounts for parameter interdependencies and is a much richer description:

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

Bishop, 2006; Friston et al., 2007, *NeuroImage*

NEGATIVE FREE ENERGY – A CLOSER LOOK

The **negative free energy** represents a trade-off between the accuracy and complexity of a model:

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

In contrast to “simple” criteria (e.g., AIC & BIC), the complexity term of the negative free energy accounts for parameter interdependencies and is a much richer description:

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

complexity **higher** the more independent prior parameters

Bishop, 2006; Friston et al., 2007, *NeuroImage*

NEGATIVE FREE ENERGY – A CLOSER LOOK

The **negative free energy** represents a trade-off between the accuracy and complexity of a model:

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

In contrast to “simple” criteria (e.g., AIC & BIC), the complexity term of the negative free energy accounts for parameter interdependencies and is a much richer description:

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

complexity **higher** the more dependent posterior parameters

Bishop, 2006; Friston et al., 2007, *NeuroImage*

NEGATIVE FREE ENERGY – A CLOSER LOOK

The **negative free energy** represents a trade-off between the accuracy and complexity of a model:

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

In contrast to “simple” criteria (e.g., AIC & BIC), the complexity term of the negative free energy accounts for parameter interdependencies and is a much richer description:

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

complexity **higher** the more posterior deviates from prior mean

Bishop, 2006; Friston et al., 2007, *NeuroImage*

Bayesian system identification

$$u(t)$$

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

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

Design experimental inputs

Define likelihood model

Specify priors

Invert model

Make inferences

Bayesian model selection (BMS)

The **negative free energy** as a lower bound approximation to the log model evidence is the current gold standard for Bayesian model selection (BMS).

Generative modeling: comparing competing hypotheses about the mechanisms underlying observed data.

- *a priori* definition of hypothesis set (model space) is crucial
- determine the most plausible hypothesis (model), given the data

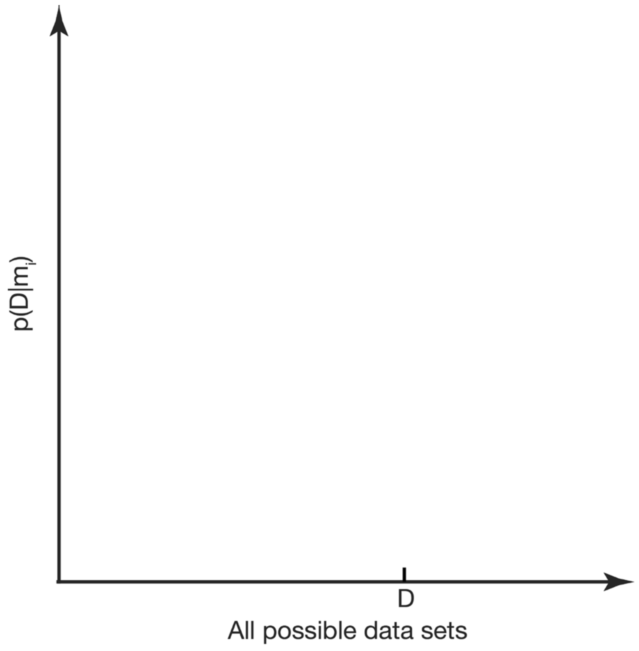
Note: **Model selection is not equal to model validation** and only allows to compare the relative goodness of competing hypotheses within the pre-specified model space!

→ Model validation requires external criteria (external to the measured data).

OVERFITTING AT THE LEVEL OF MODELS

But: There is a very large number of possible models for a given dataset. Wouldn't we need to search the entire model space and test all possible models?

No! With more models included in the model space, the risk of overfitting (at the level of models) increases, too.

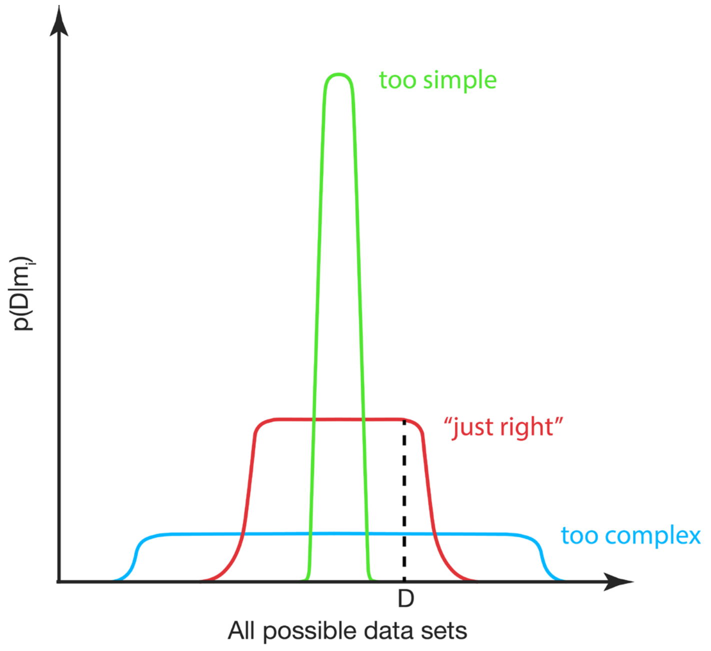


Ghahramani, 2004

OVERFITTING AT THE LEVEL OF MODELS

But: There is a very large number of possible models for a given dataset. Wouldn't we need to search the entire model space and test all possible models?

No! With more models included in the model space, the risk of overfitting (at the level of models) increases, too.

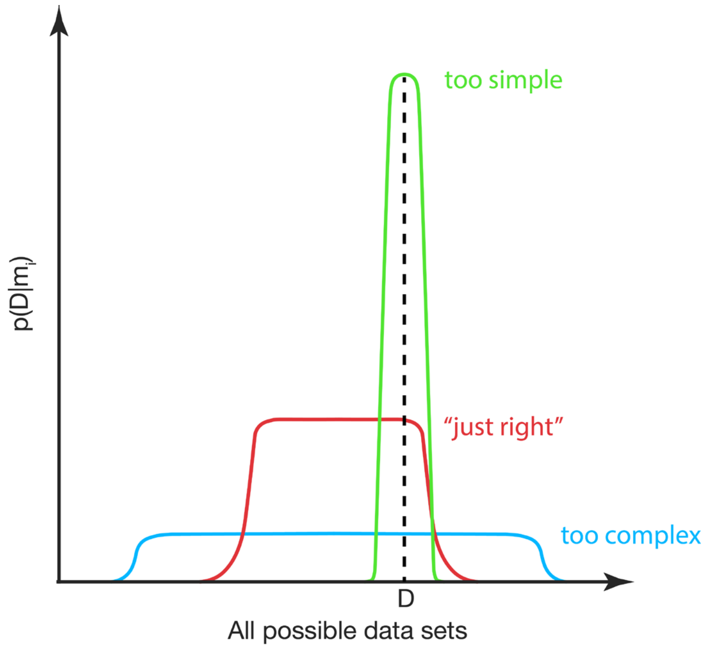


Ghahramani, 2004

Overfitting at the level of models

But: There is a very large number of possible models for a given dataset. Wouldn't we need to search the entire model space and test all possible models?

No! With more models included in the model space, the risk of overfitting (at the level of models) increases, too.



Ghahramani, 2004

Overfitting at the level of models

But: There is a very large number of possible models for a given dataset. Wouldn't we need to search the entire model space and test all possible models?

No! With more models included in the model space, the risk of overfitting (at the level of models) increases, too.

Solutions:

- regularization: definition of model space (i.e., specify priors $p(m)$ over models)
- family-level Bayesian model selection
- Bayesian model averaging (BMA)

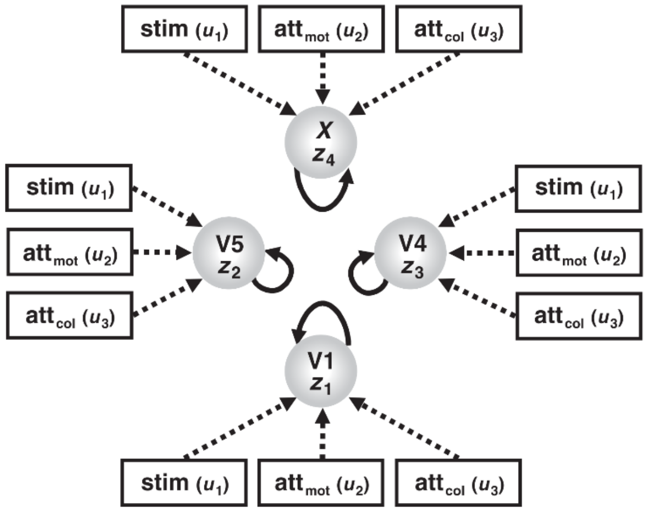
Ghahramani, 2004

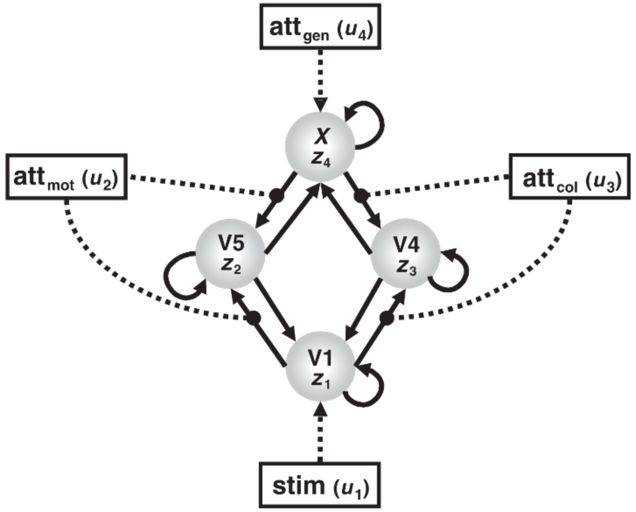
Note: GLM vs. DCM

DCM tries to model the same phenomena (i.e., local BOLD responses) as a GLM, just in a different way (via connectivity and its modulations).

No activation detected by a GLM → no motivation to include this region in a deterministic DCM.

However, a stochastic DCM (that incorporates a noise term in the neuronal state equation and can thus account for endogenous fluctuations) could be applied in the absence of a local activation.

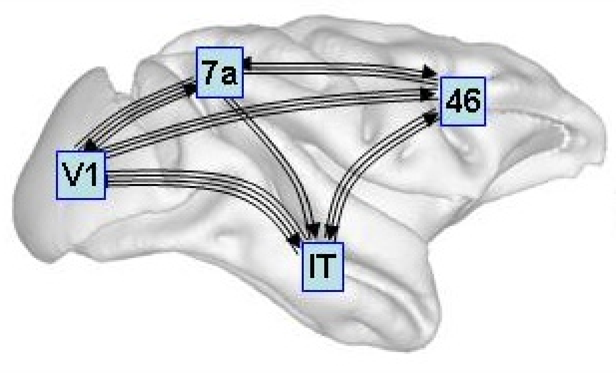




Stephan, 2004, *J. Anat.*

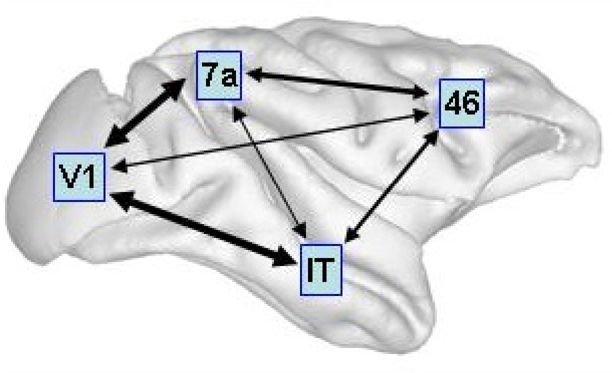
INTERPRETATION: DIFFERENT FORMS OF BRAIN CONNECTIVITY

structural connectivity



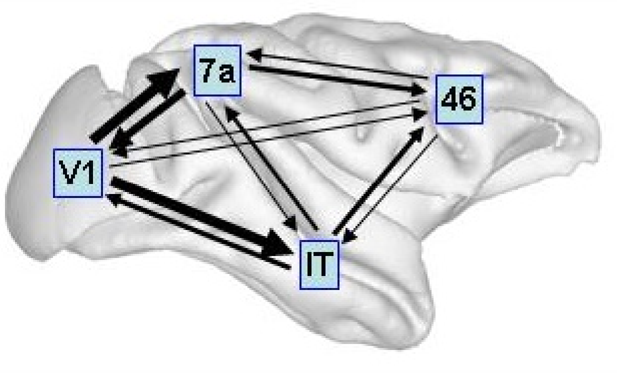
- presence of physical connections
- Diffusion weighted imaging (DWI), tractography, tracer studies

functional connectivity



- statistical dependencies between regional time series
- correlations, Independent Component Analysis (ICA)

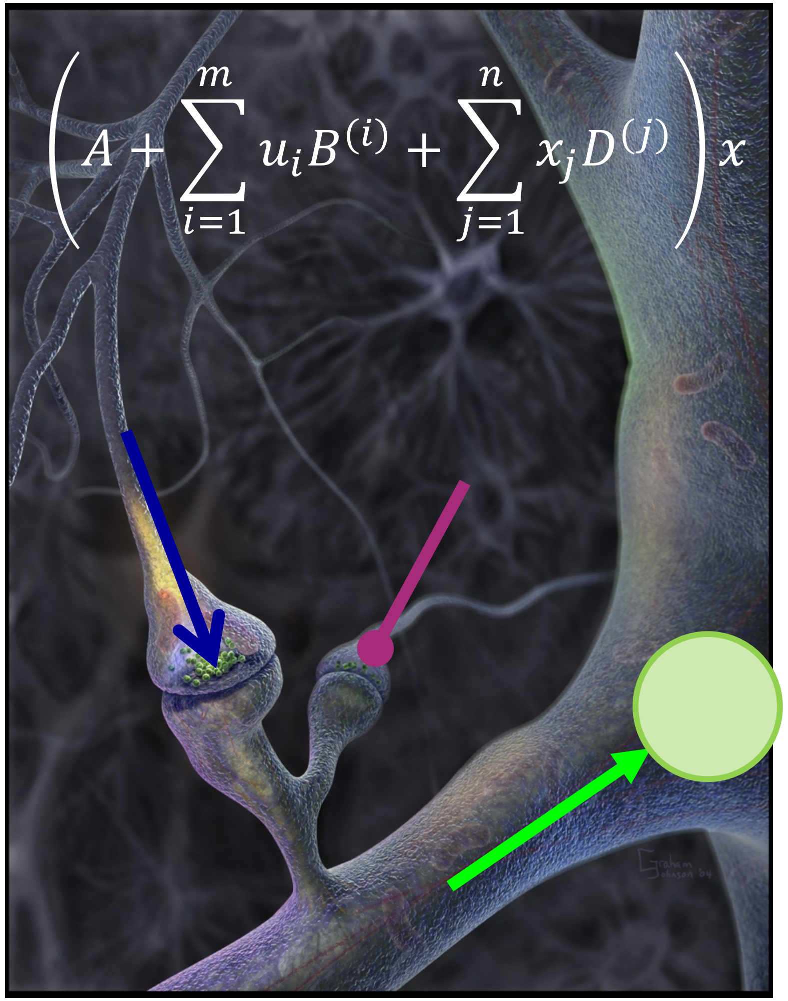
effective connectivity



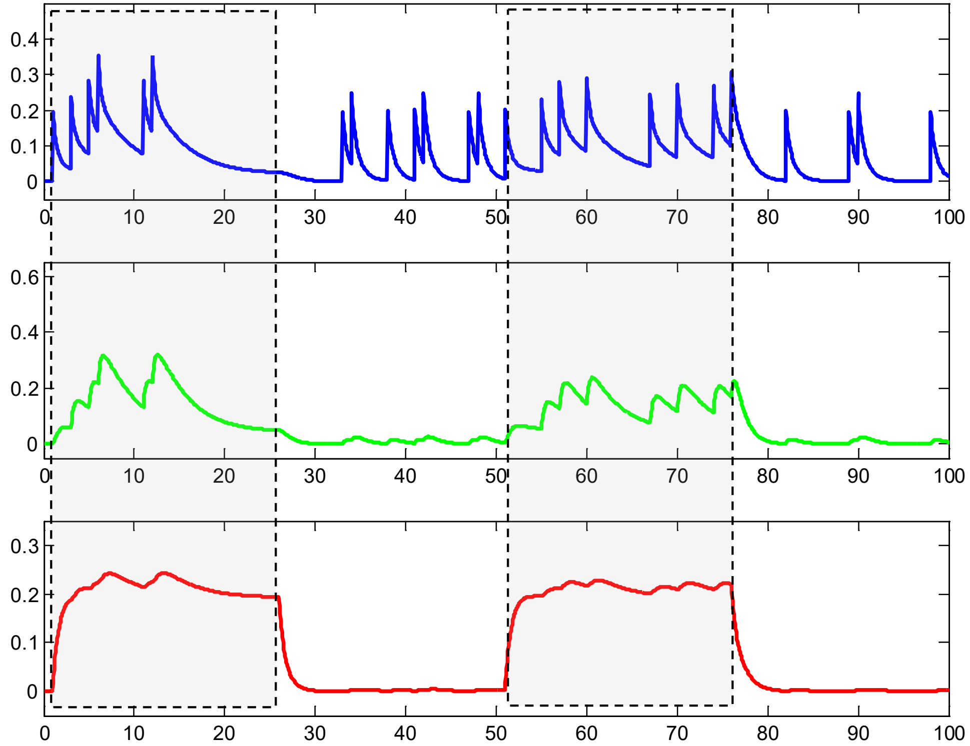
- (causal) directed influences between neuronal populations
- Dynamic causal modeling (DCM)

Sporns, 2007, *Scholarpedia*

INTERPRETATION: NEUROMODULATION



Stephan et al, 2008, *Neuroimage*

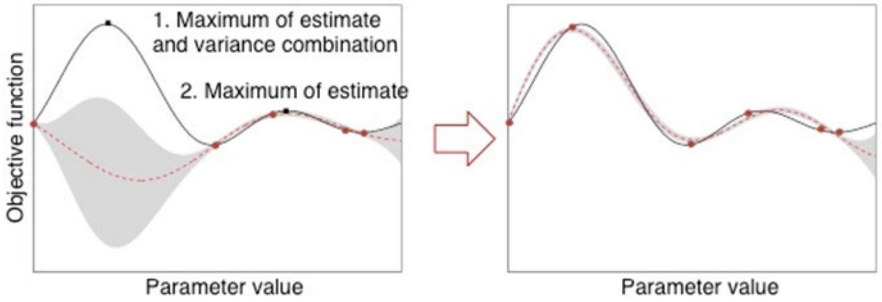


Synaptic strengths are context-sensitive:
They depend on spatio-temporal patterns of network activity.

Methodological developments of DCM

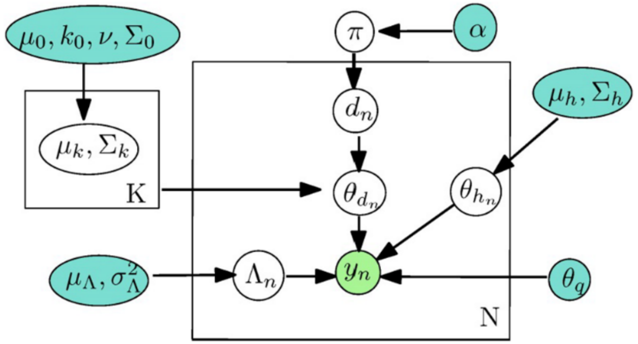
Global optimization schemes for model inversion

- Markov Chain Monte Carlo (MCMC) sampling (Sengupta et al., 2015, *NeuroImage*)
- Gaussian process (GP) regression (Lomakina et al., 2015, *NeuroImage*)



Sampling-based estimates of model evidence

- Aponte et al. 2016, *J. Neurosci. Meth.*
- Raman et al., 2016, *J. Neurosci. Meth.*



Choice of priors → empirical Bayes

- Friston et al. 2016, *NeuroImage*
- Raman et al. 2016 and 2018, *J. Neurosci. Meth.*
- Yao et al. 2018, *Neuroimage*

Sengupta et al, 2015, *NeuroImage*; Lomakina et al., 2015, *NeuroImage*; Aponte et al., 2016, *J. Neurosci. Meth.*; Friston et al., 2016, *NeuroImage*; Raman et al., 2016, *J. Neurosci. Meth.;* Raman et al., 2018, *J Neurosci Meth*; Yao et al., 2018, *NeuroImage*

APPLICATIONS

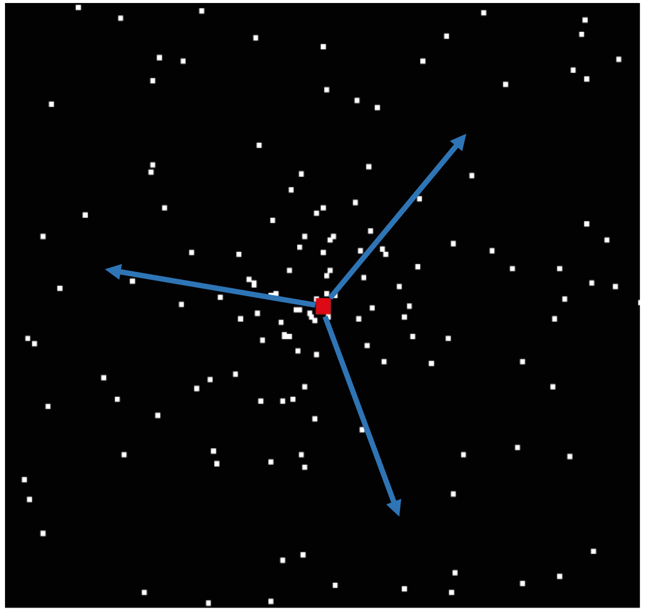
Simple example: Attention to motion

Stimuli: radially moving dots were presented.

Pre-scanning: 5x30s trials with 5 speed changes. Subjects were asked to detect the change in radial velocity.

Scanning: No actual speed changes. Conditions:

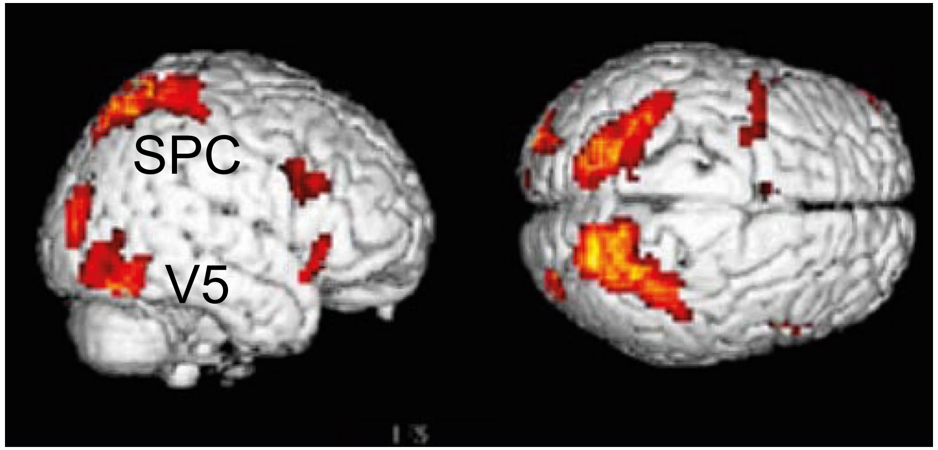
- F: fixation
- S: static dots
- M: moving dots
- A: attend moving dots



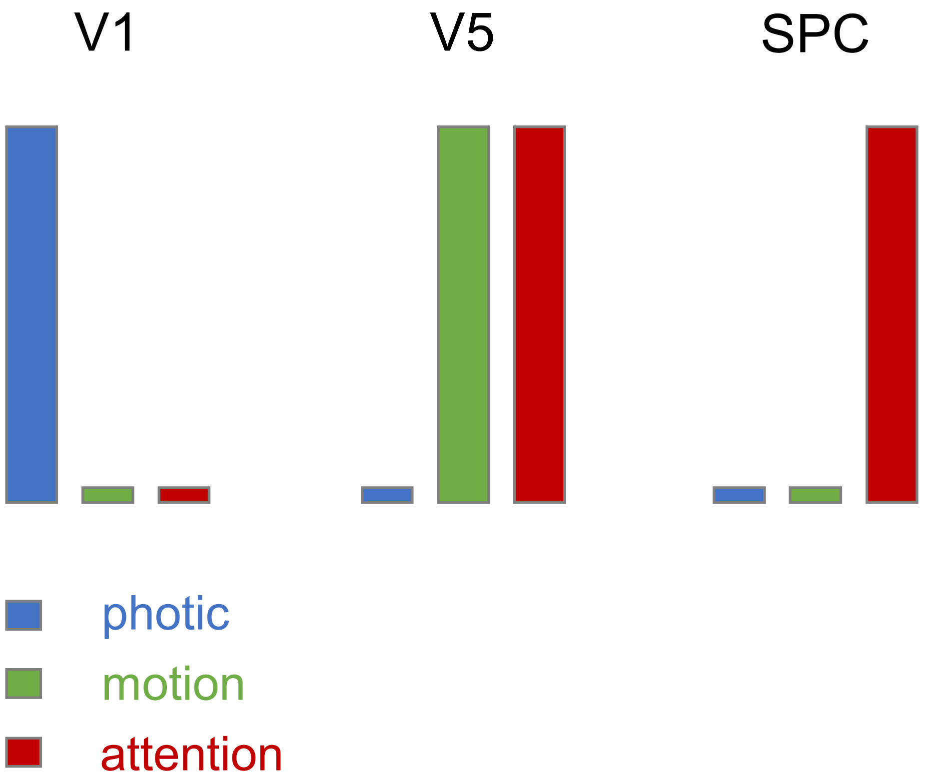
Büchel and Friston, 1997, *Cerebral Cortex*; Friston et al., 2003, *NeuroImage*

Simple example: Attention to motion

Single-subject results: BOLD activation patterns

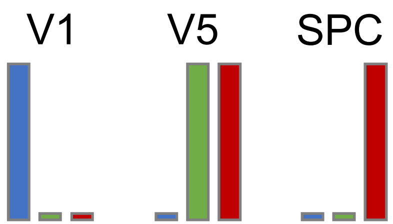


Linear contrast: attention > no attention

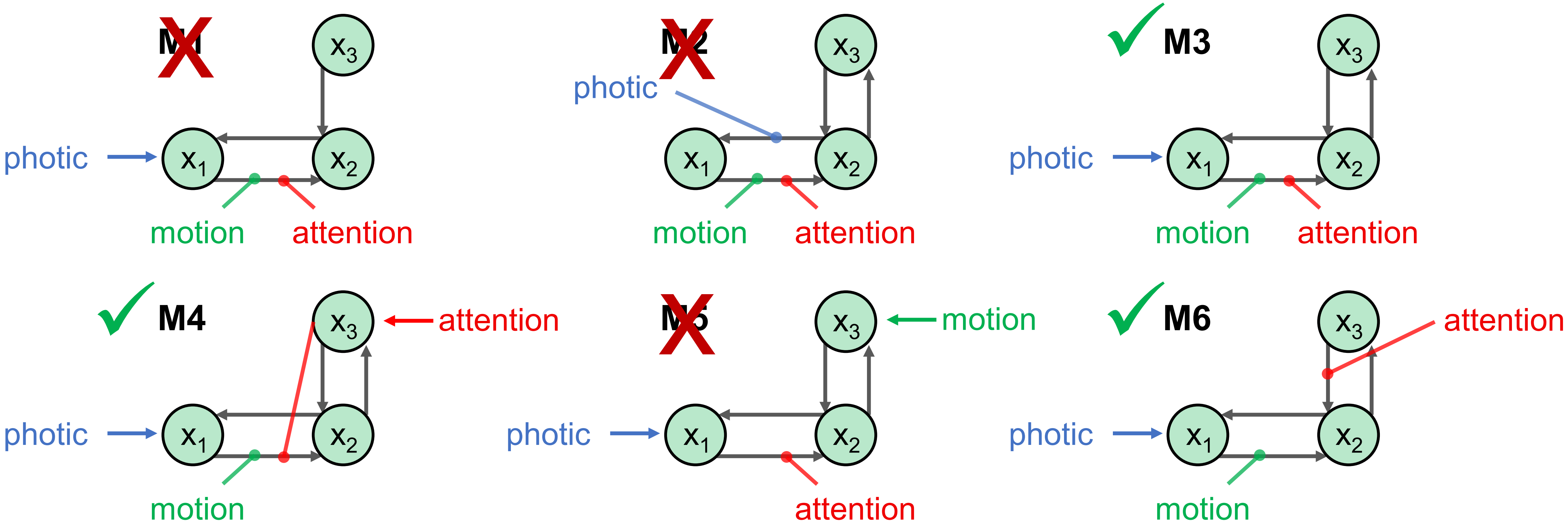


Büchel and Friston, 1997, *Cerebral Cortex*; Friston et al., 2003, *NeuroImage*

Simple example: Attention to motion



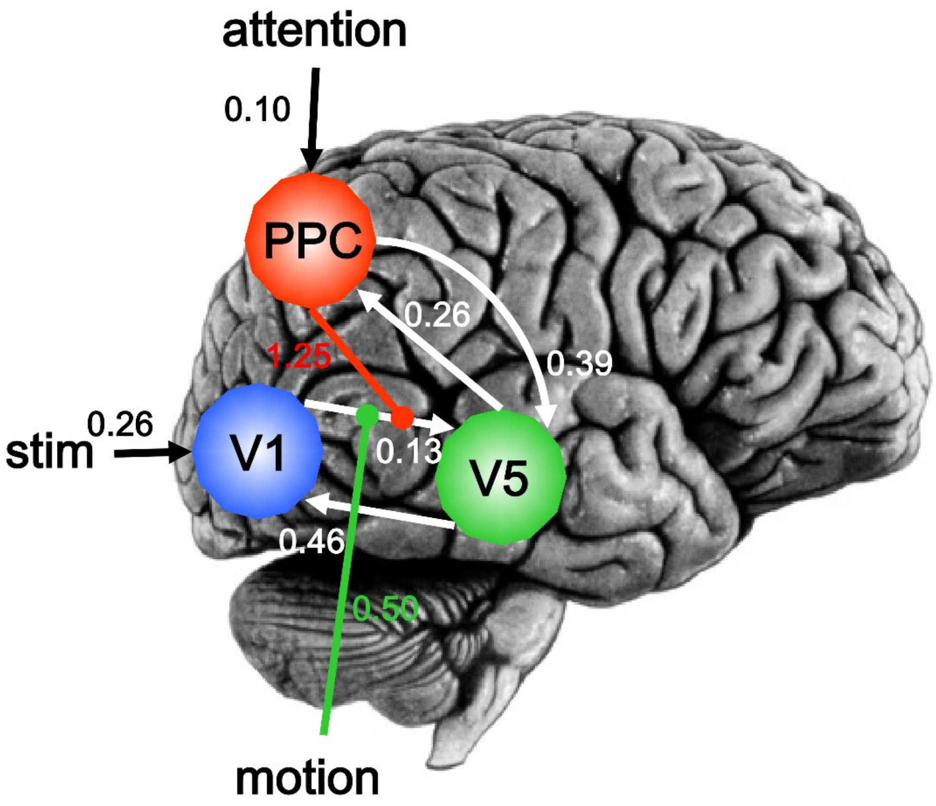
Model space definition – which models can explain the data (Quiz)?

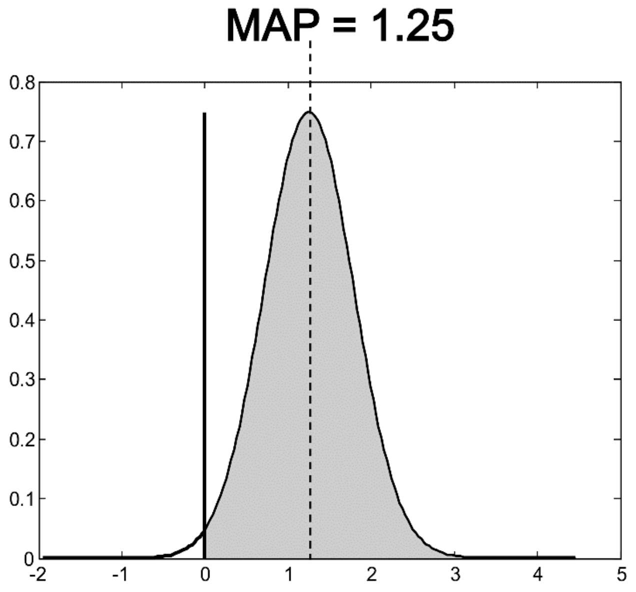


Büchel and Friston, 1997, *Cerebral Cortex*; Friston et al., 2003, *NeuroImage*

Simple example: Attention to motion

Single-subject results: DCM effective connectivity





$$p(D^{PPC}_{V5,V1} > 0 \mid y) = 99.1\%$$

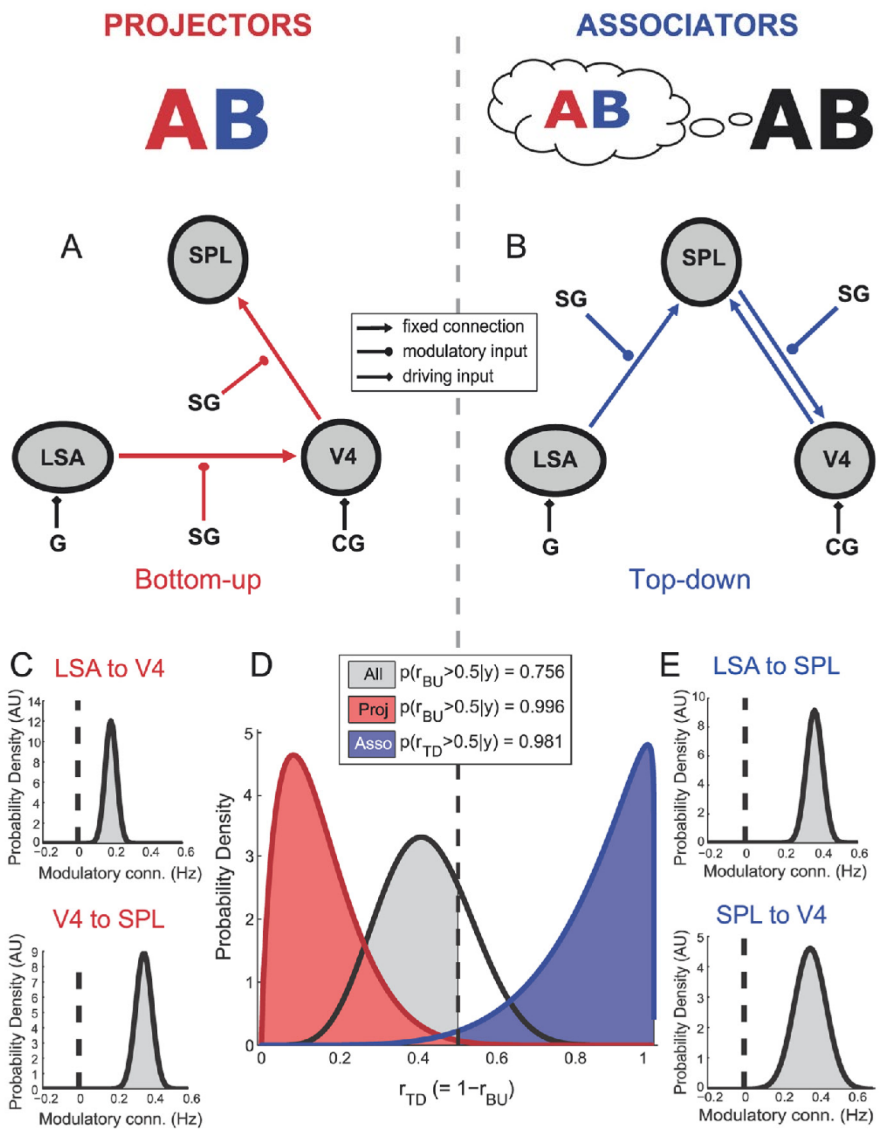
Büchel and Friston, 1997, *Cerebral Cortex*; Friston et al., 2003, *NeuroImage*; Stephan et al., 2008, *NeuroImage*

Bayesian model selection: Synesthesia (Supplemental Info)

Individuals with different forms of color-grapheme synesthesia were tested and effective connectivity in the relevant neural circuits was assessed using DCM.

Bayesian model selection (BMS) as a formal approach to differential diagnosis in clinical applications

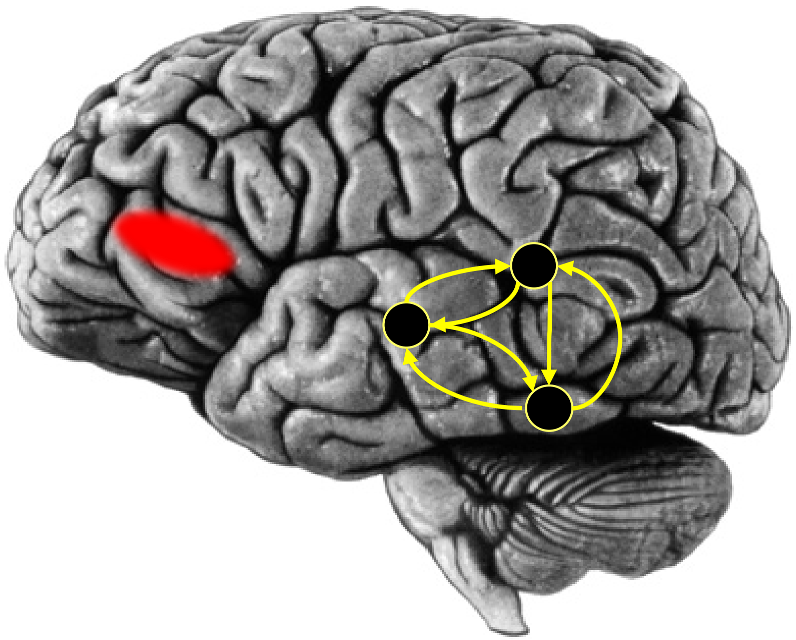
(Note: Here, different forms of synesthesia were tested. This is not a clinical condition, but simply a specific cognitive trait.)



Van Leeuwen et al., 2011, *J. Neurosci.*

GENERATIVE EMBEDDING: APHASIA

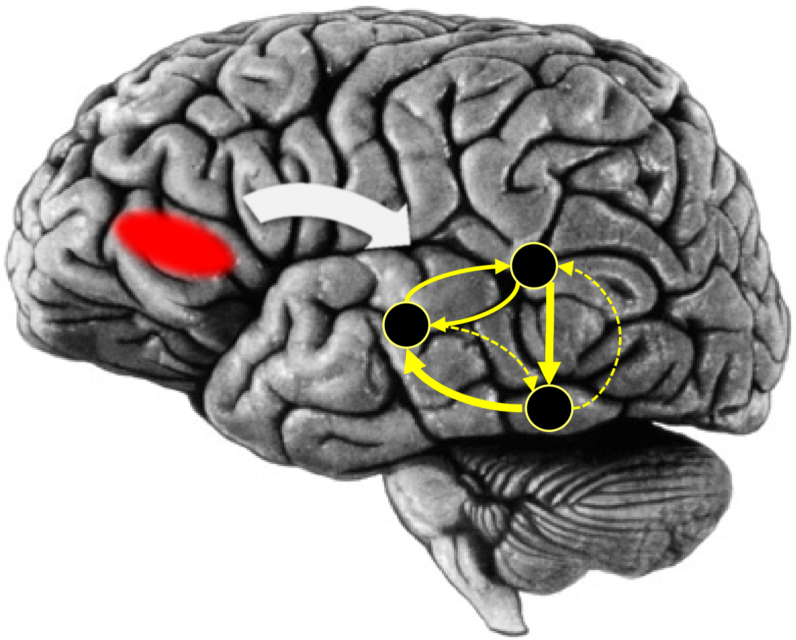
Dissociating aphasic patients (N=11) and healthy controls (N=26)



Schofield et al., 2012, *J. Neurosci.*; Brodersen et al., 2011, *PLoS Comp. Biol.*

GENERATIVE EMBEDDING: APHASIA

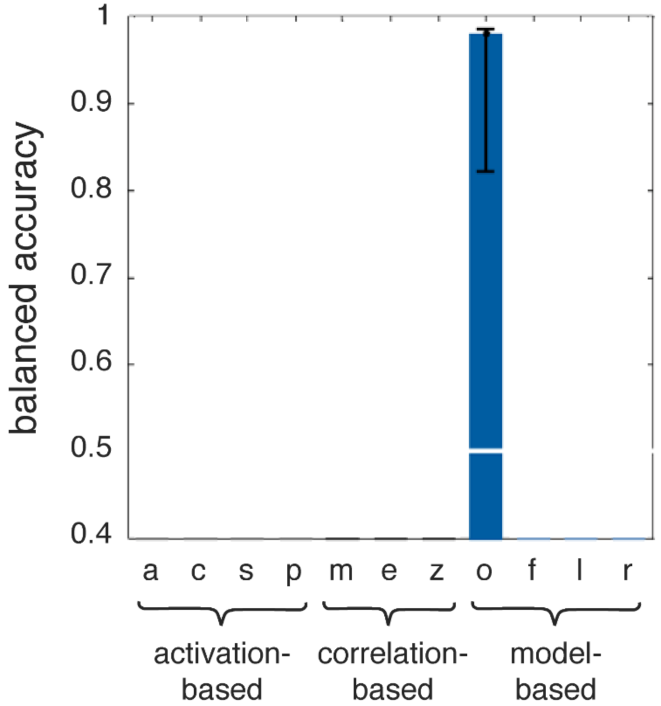
Dissociating aphasic patients (N=11) and healthy controls (N=26)



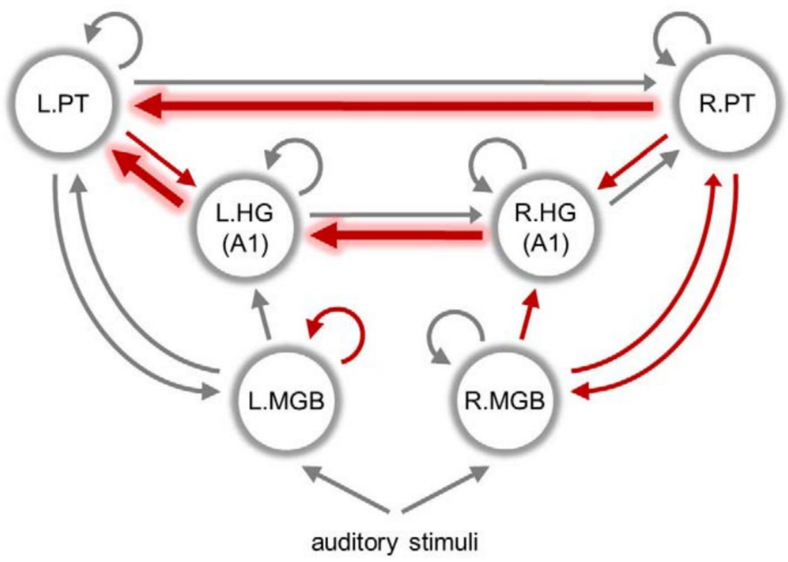
Schofield et al., 2012, *J. Neurosci.*; Brodersen et al., 2011, *PLoS Comp. Biol.*

Generative embedding: Aphasia

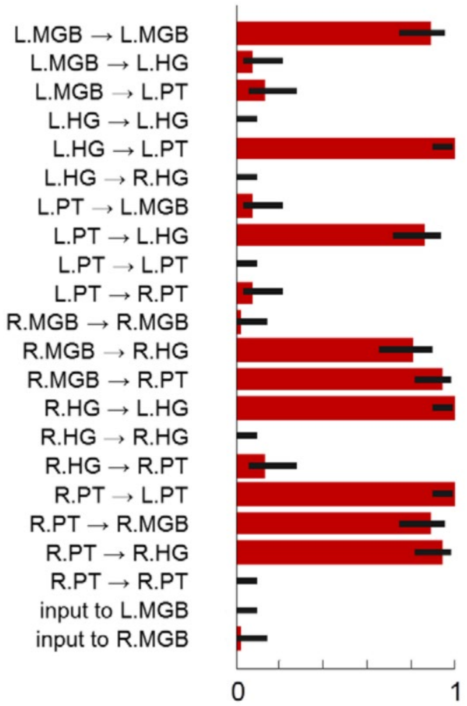
Dissociating aphasic patients (N=11) and healthy controls (N=26)



support vectors



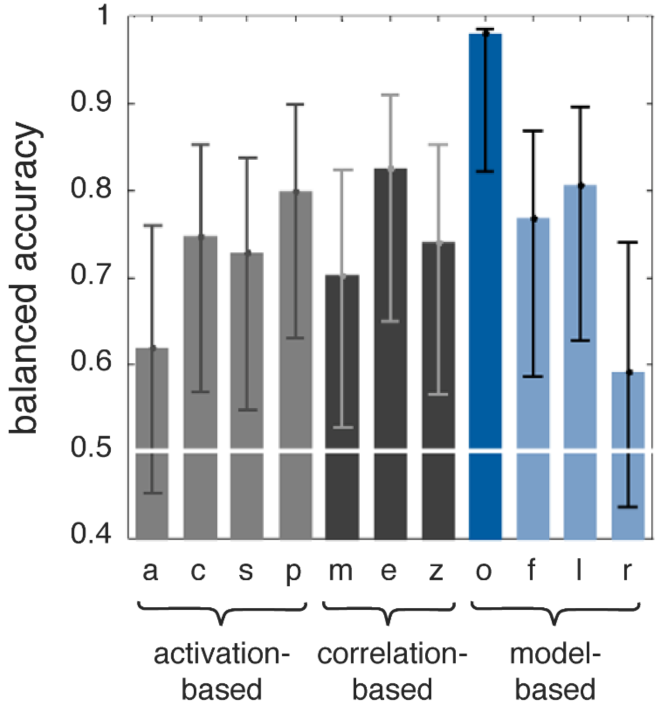
frequency



Schofield et al., 2012, *J. Neurosci.*; Brodersen et al., 2011, *PLoS Comp. Biol.*

Generative embedding: Aphasia

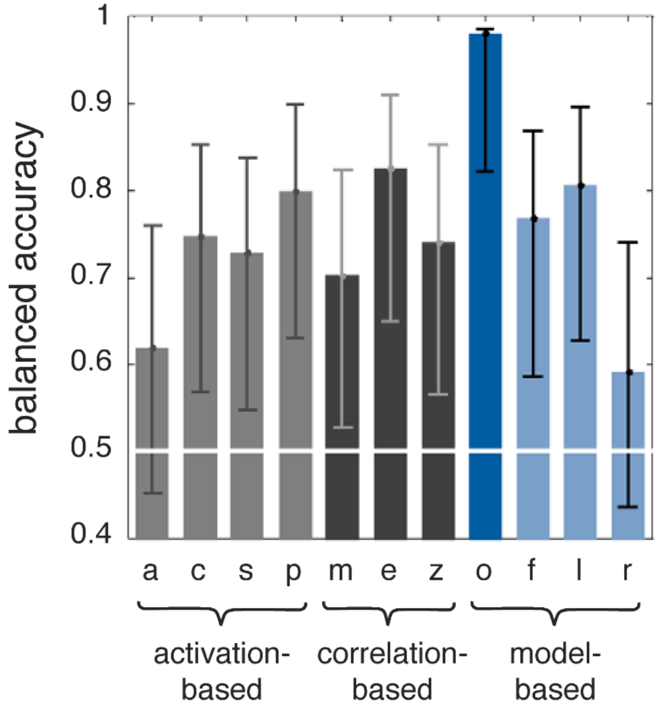
Dissociating aphasic patients (N=11) and healthy controls (N=26)

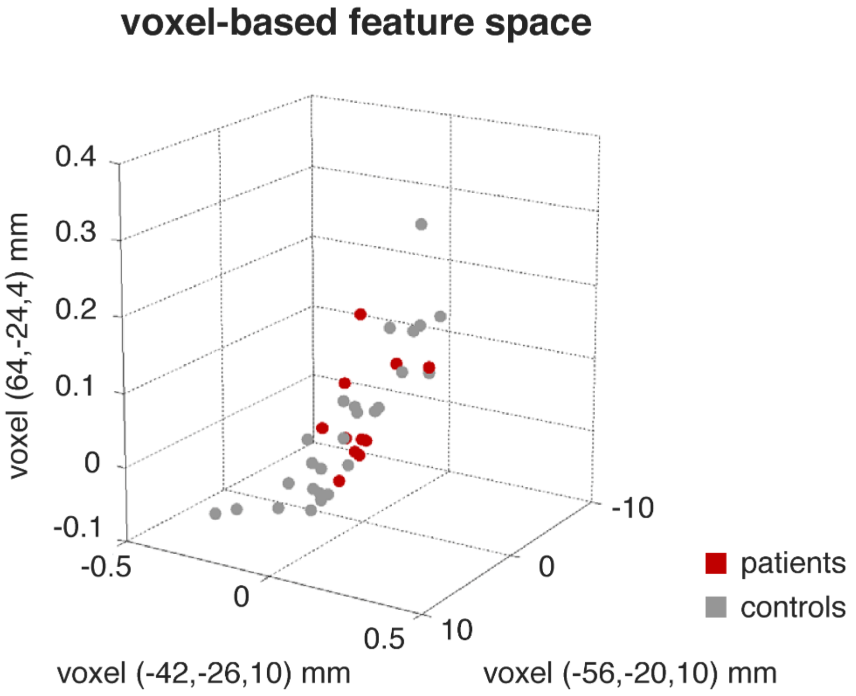


Schofield et al., 2012, *J. Neurosci.*; Brodersen et al., 2011, *PLoS Comp. Biol.*

Generative embedding: Aphasia

Dissociating aphasic patients (N=11) and healthy controls (N=26)

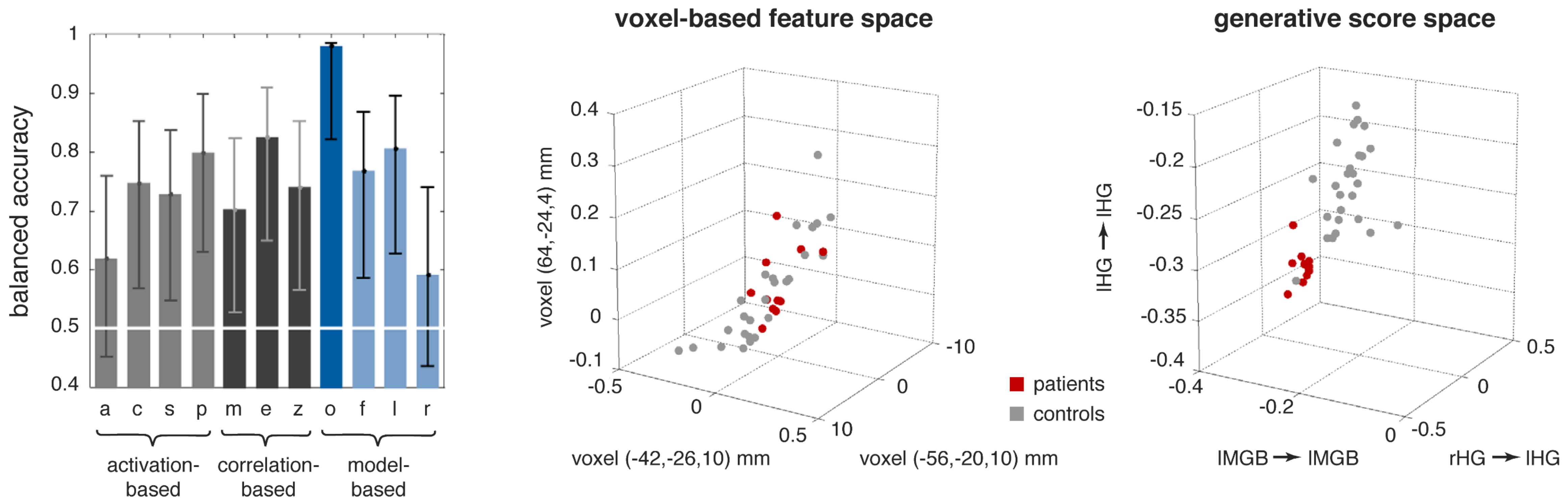




Schofield et al., 2012, *J. Neurosci.*; Brodersen et al., 2011, *PLoS Comp. Biol.*

Generative embedding: Aphasia

Dissociating aphasic patients (N=11) and healthy controls (N=26)

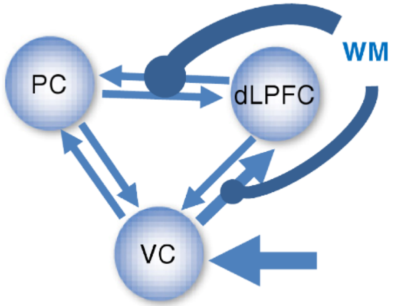


Schofield et al., 2012, *J. Neurosci.*; Brodersen et al., 2011, *PLoS Comp. Biol.*

Generative embedding: Schizophrenia

Detecting subgroups of patients in schizophrenia (N=41)

neurophysiological characterisation

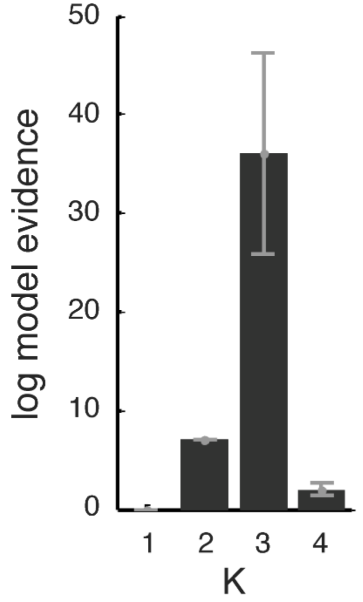


Deserno et al., 2012, *J. Neurosci.*; Brodersen et al., 2014, *NeuroImage: Clinical*

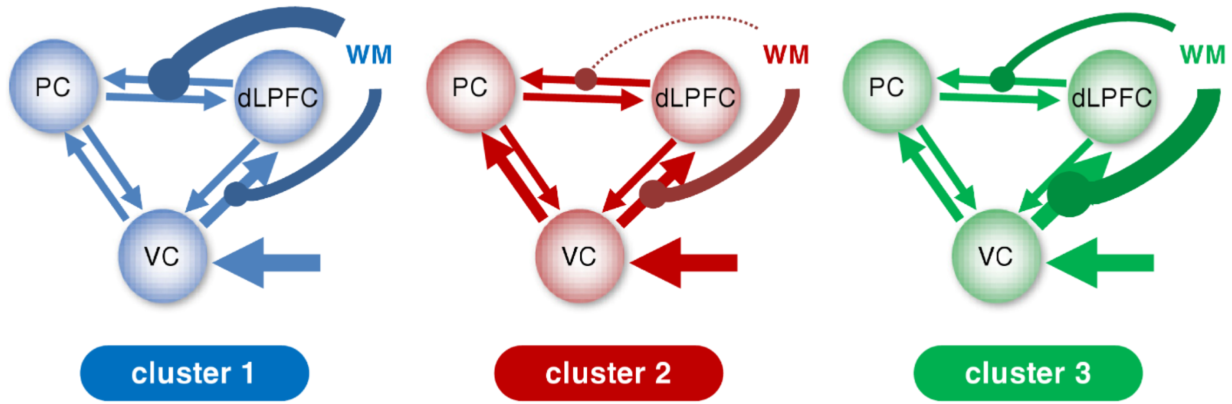
GENERATIVE EMBEDDING: SCHIZOPHRENIA

Detecting subgroups of patients in schizophrenia (N=41)

model selection



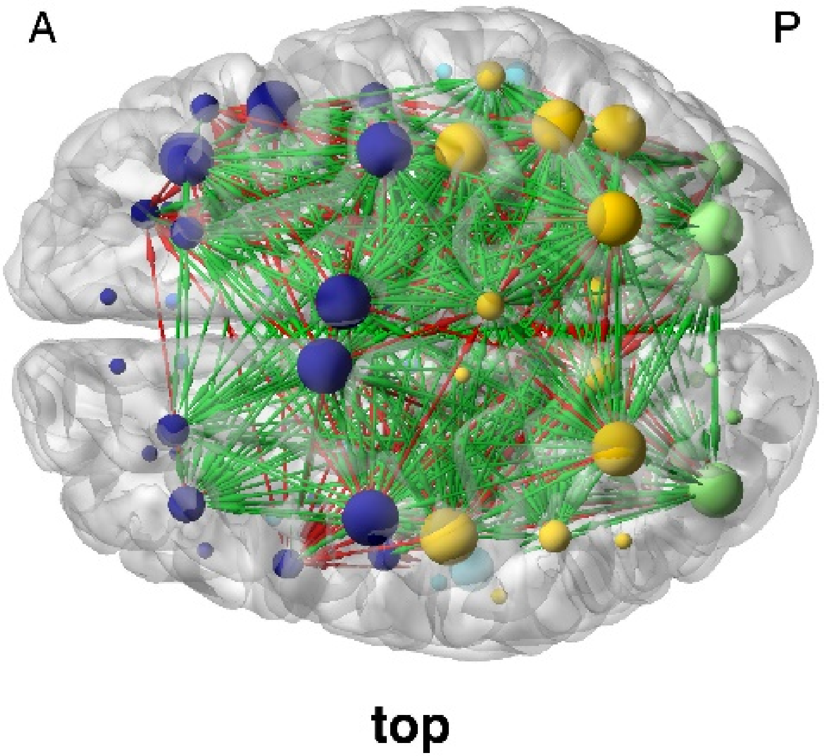
neurophysiological characterisation

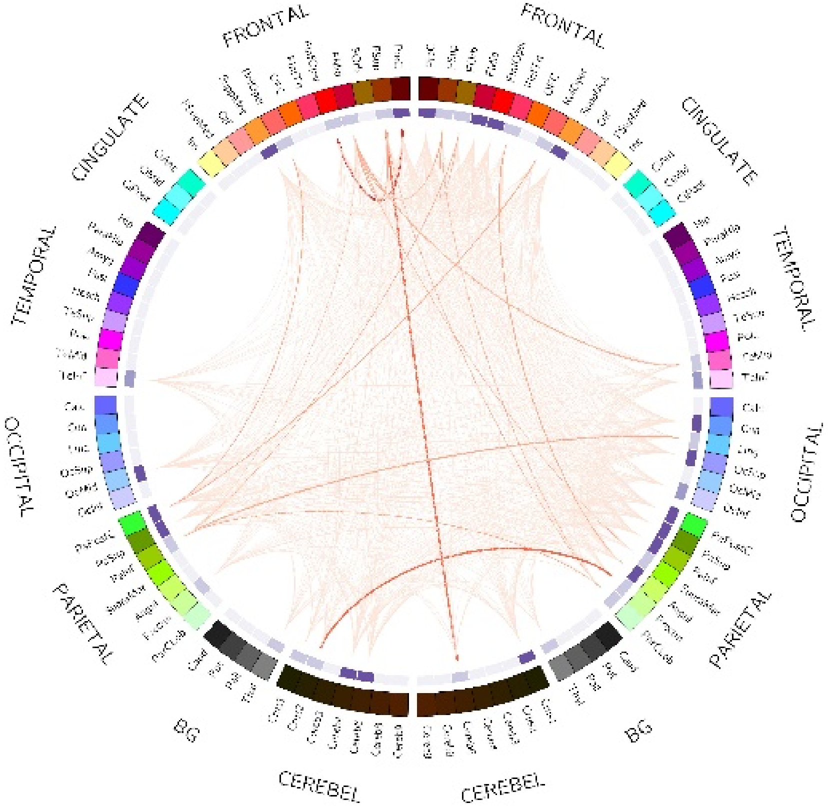


Deserno et al., 2012, *J. Neurosci.*; Brodersen et al., 2014, *NeuroImage: Clinical*

DCM Extensions

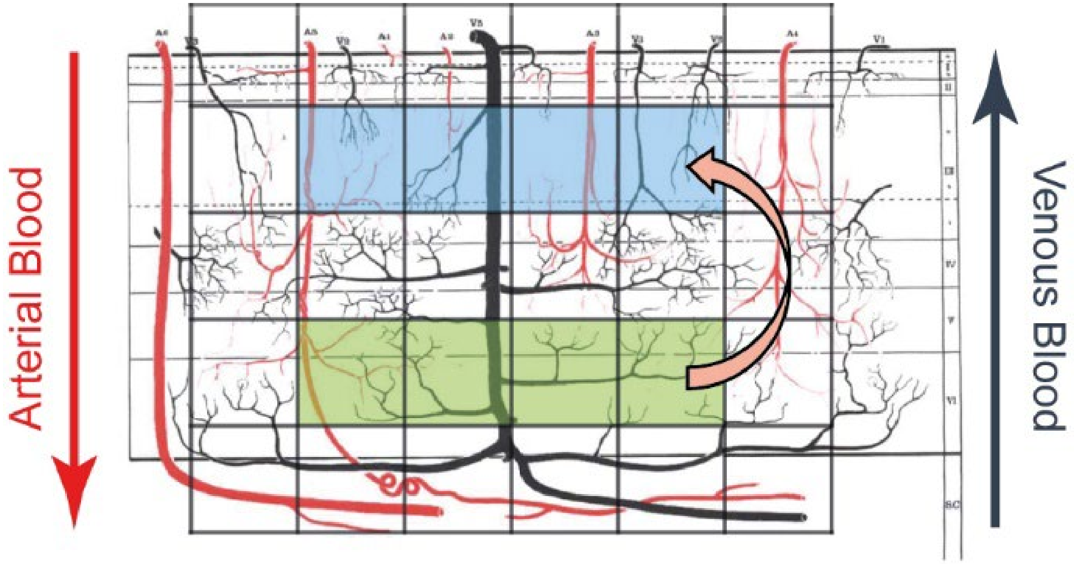
Regression DCM

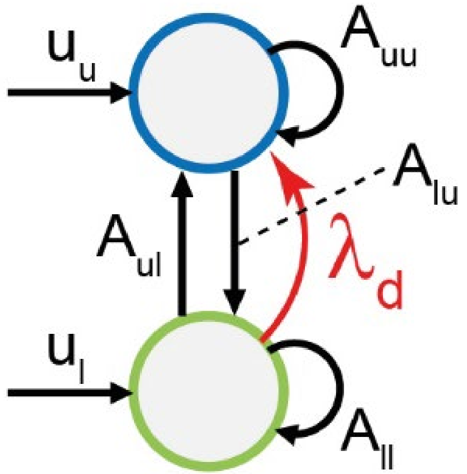




Frässle et al., 2017, *NeuroImage*; Frässle et al., 2018, *NeuroImage*

DCM for Layered fMRI





$$\tau_u \frac{dv_u}{dt} = f_u - v_u^{\frac{1}{\alpha_u}} + \lambda_d v_l^*$$

$$\tau_u \frac{dq_u}{dt} = f_u \frac{1-(1-E_0)^{1/f_u}}{E_0} - {v_u}^{1/\alpha_u} \frac{q_u}{v_u} + \lambda_d q_l^*$$

$$\tau_l \frac{dv_l}{dt} = f_l - v_l^{\frac{1}{\alpha_l}}$$

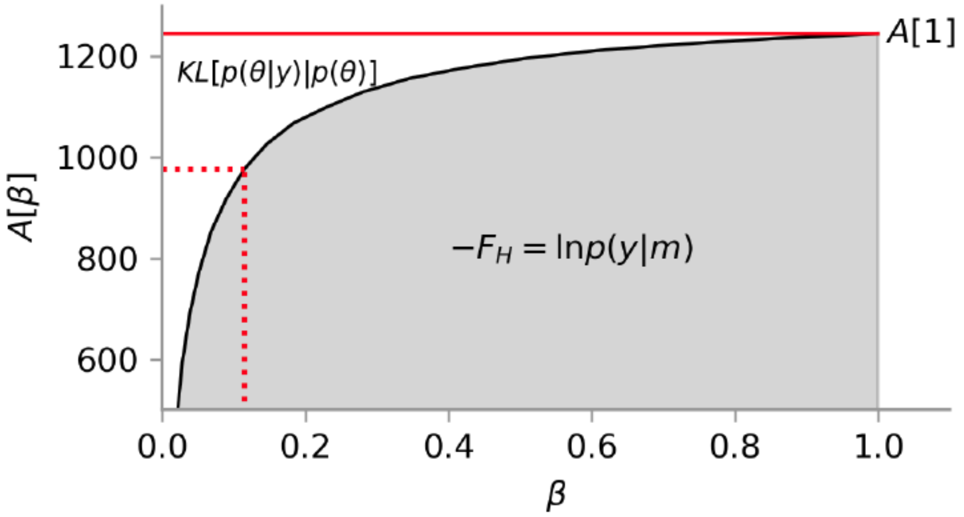
$$\tau_l \frac{dq_l}{dt} = f_l \frac{1-(1-E_0)^{1/f_l}}{E_0} - {v_l}^{1/\alpha_l} \frac{q_l}{v_l}$$

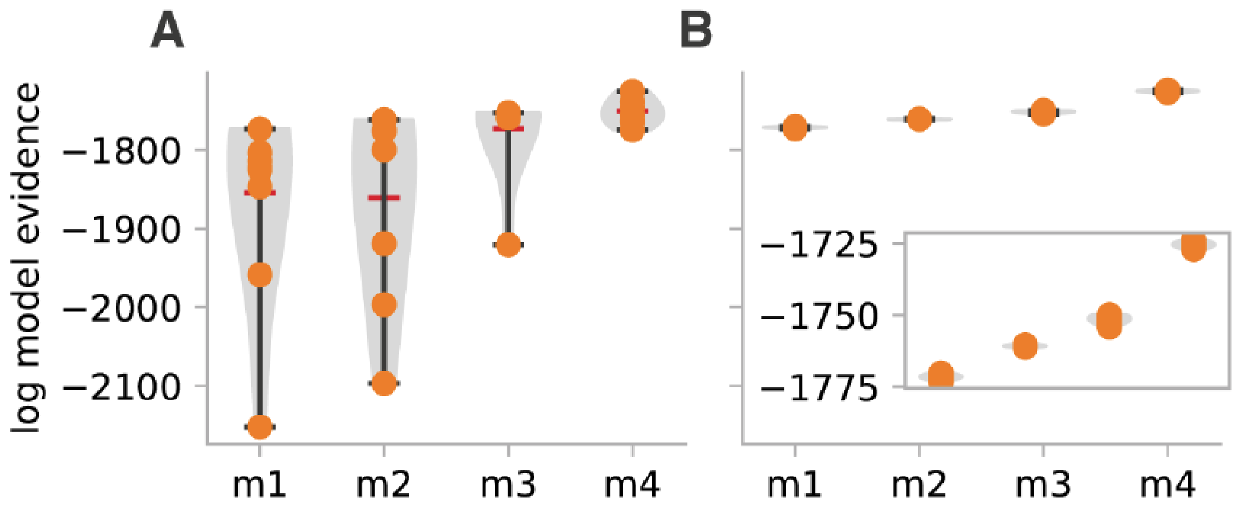
$$\tau_d \frac{dv_l^*}{dt} = -v_l^* + (v_l - 1)$$

$$\tau_d \frac{dq_l^*}{dt} = -q_l^* + (q_l - 1)$$

Heinzle et al., 2017, *NeuroImage*

Thermodynamic Integration for DCM





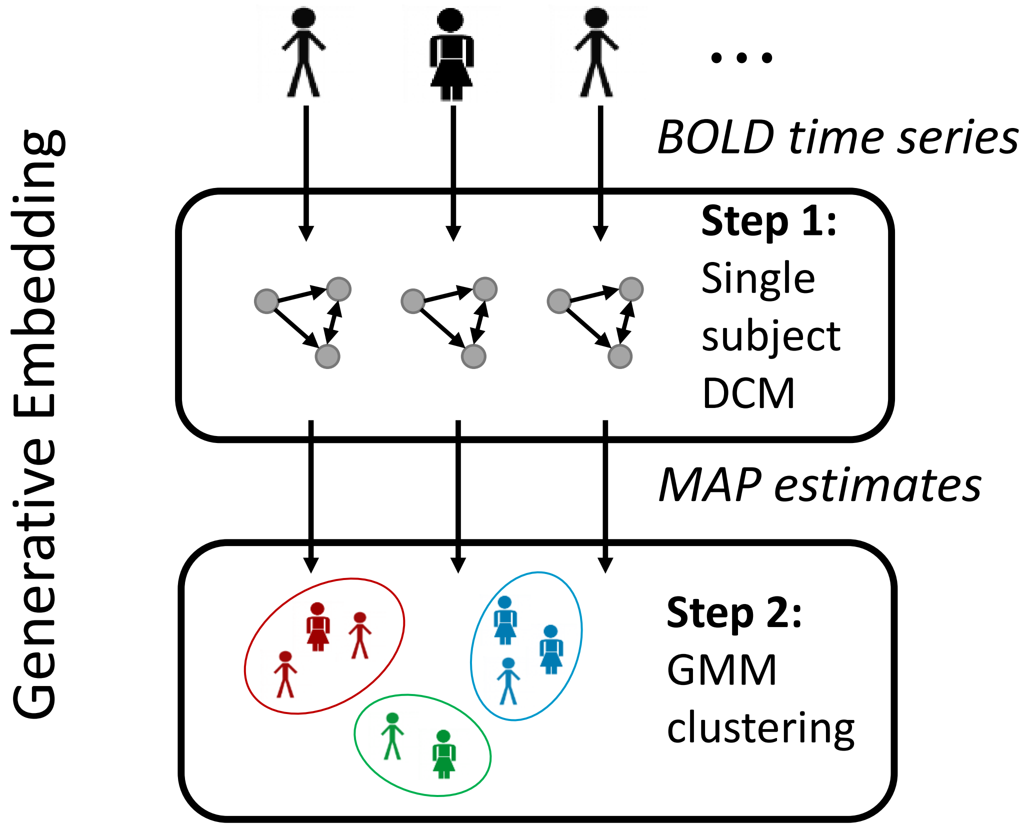
Aponte et al., 2016, *J Neurosci Meth*; Aponte et al., 2018, *biorXiv*;

Translational Neuromodeling Unit

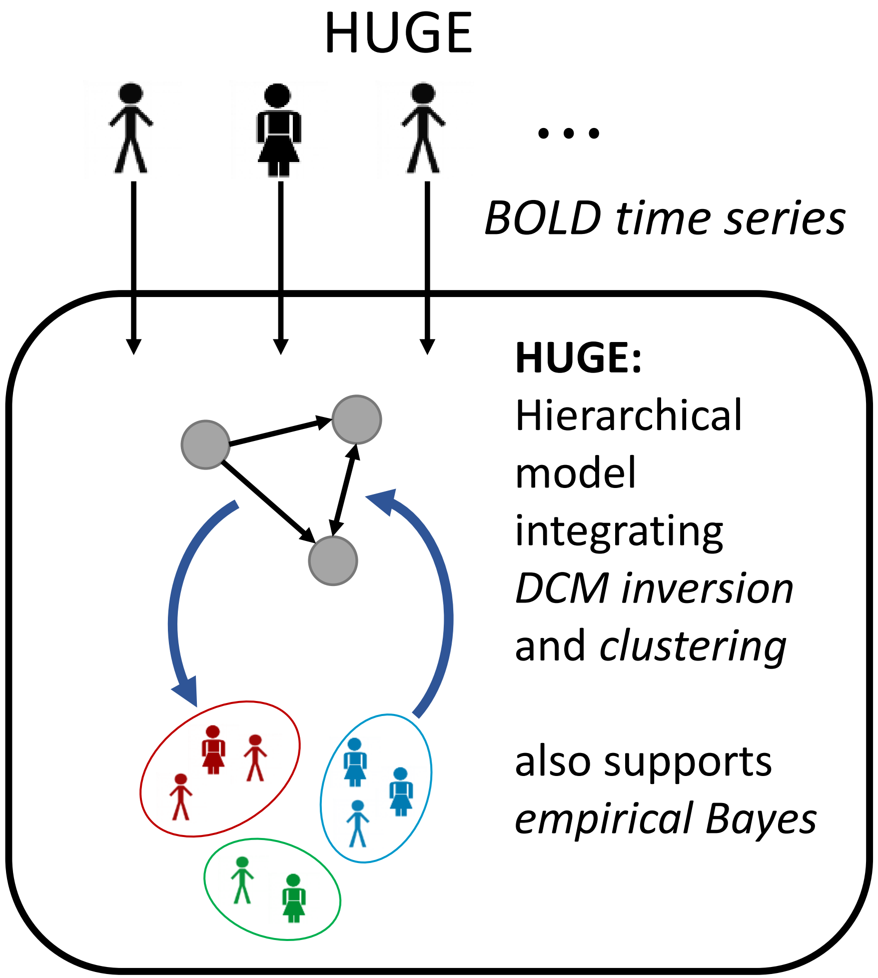
University of Zurich[UZH]

ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

HUGE – Hierarchical Unsupervised Generative Embedding



Raman et al., 2018, *J Neurosci Meth*; Yao et al., 2018, *NeuroImage*



Stochastic DCM

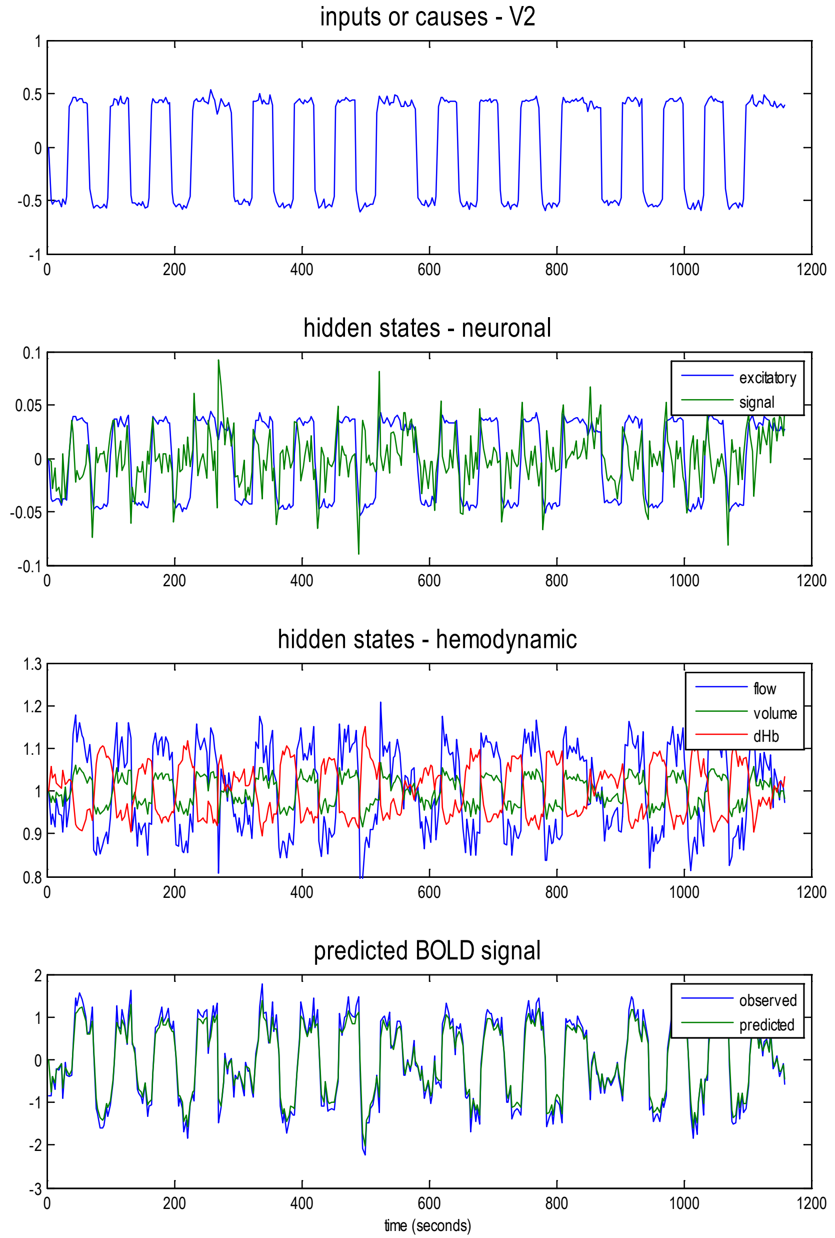
$$\frac{dx}{dt} = (A + \sum_j u_j B^{(j)})x + Cv + \omega^{(x)}$$

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

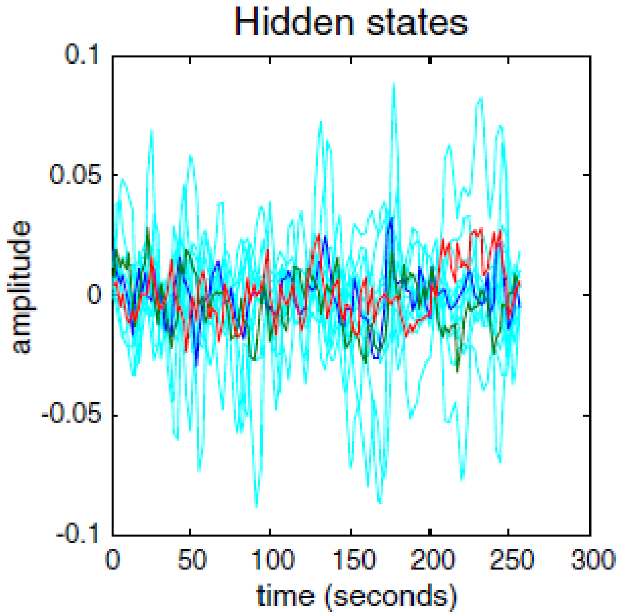
- all states are represented in generalised coordinates of motion
- random state fluctuations $w^{(x)}$ account for endogenous fluctuations, have unknown precision and smoothness $\rightarrow$ two hyperparameters
- fluctuations $w^{(v)}$ induce uncertainty about how inputs influence neuronal activity
- can be fitted to resting state data

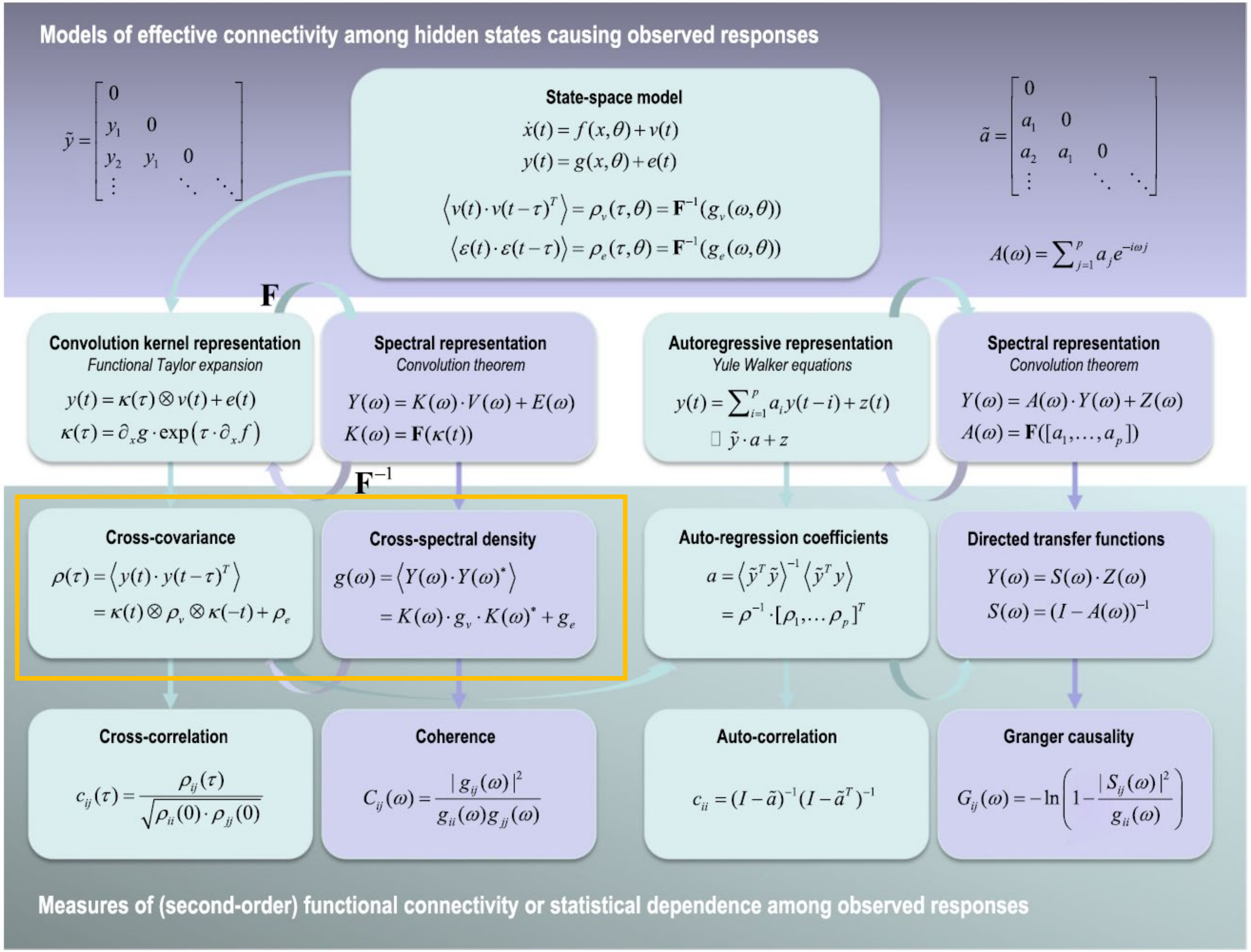
Li et al., 2011, *NeuroImage*

Estimates of hidden causes and states
(Generalised filtering)



Spectral DCM – Resting State





Friston et al, 2014, *Neuroimage*

Spectral DCM – Resting State

Cross-covariance

$$\rho(\tau) = \left\langle y(t) \cdot y(t-\tau)^T \right\rangle$$
$$= \kappa(t) \otimes \rho_v \otimes \kappa(-t) + \rho_e$$

Cross-spectral density

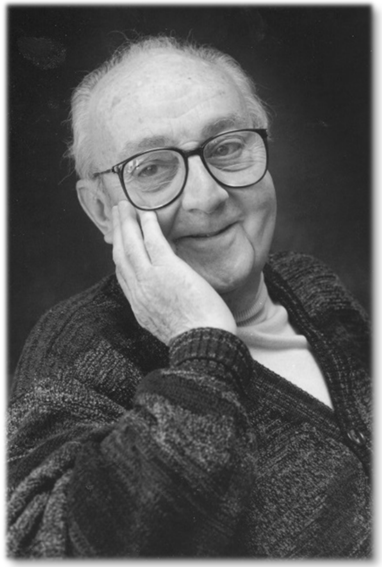
$$g(\omega) = \left\langle Y(\omega) \cdot Y(\omega)^* \right\rangle$$
$$= K(\omega) \cdot g_v \cdot K(\omega)^* + g_e$$

Friston et al, 2014, *Neuroimage*

ALL MODELS ARE WRONG

BUT SOME ARE USEFUL

George Edward Pelham Box
(1919-2013)



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 utility

For DCM: >15 published validation studies (incl. 6 animal studies):

- infers site of seizure origin (David et al. 2008)
- infers primary recipient of vagal nerve stimulation (Reyt et al. 2010)
- infers synaptic changes as predicted by microdialysis (Moran et al. 2008)
- infers fear conditioning induced plasticity in amygdala (Moran et al. 2009)
- tracks anesthesia levels (Moran et al. 2011)
- predicts sensory stimulation (Brodersen et al. 2010)

FURTHER READINGS (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.
- Brodersen KH, Deserno L, Schlagenhauf F, Lin Z, Penny WD, Buhmann JM, Stephan KE (2014) Dissecting psychiatric spectrum disorders by generative embedding. NeuroImage: Clinical 4: 98-111
- Daunizeau J, David, O, Stephan KE (2011) Dynamic Causal Modelling: A critical review of the biophysical and statistical foundations. NeuroImage 58: 312-322.
- Daunizeau J, Stephan KE, Friston KJ (2012) Stochastic Dynamic Causal Modelling of fMRI data: Should we care about neural noise? NeuroImage 62: 464-481.
- Frässle S, Lomakina EI, Kasper L, Manjaly ZM, Leff A, Pruessmann KP, Buhmann JM, Stephan KE (2018) A generative model of whole-brain effective connectivity. Neuroimage 179:505-529
- Frässle S, Lomakina EI, Razi A, Friston KJ, Buhmann JM, Stephan KE (2017) Regression DCM for fMRI. NeuroImage 155:406-421.
- Frässle S, Yao Y, Schöbi D, Aponte EA, Heinzle J, Stephan KE (2018) Generative models for clinical applications in computational psychiatry. Wiley Interdiscip Rev Cogn Sci 9(3):e1460.
- Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. NeuroImage 19:1273-1302.
- Friston KJ, Mattout J, Trujill-Barreto, Ashburner J, Penny W (2007) Variational free energy and the Laplace approximation. NeuroImage 34: 220-234.
- Friston K, Stephan KE, Li B, Daunizeau J (2010) Generalised filtering. Mathematical Problems in Engineering 2010: 621670.

FURTHER READINGS (2)

- 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.
- Friston KJ, Kahan J, Biswal B, Razi A (2014) A DCM for resting state fMRI. Neuroimage 94:396-407.
- Heinzle J, Koopmans PJ, den Ouden HEM, Raman S, Stephan KE (2016) A hemodynamic model for layered BOLD signals. Neuroimage 125:556-570.
- 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.
- 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.
- Rigoux L, Stephan KE, Friston KJ, Daunizeau J (2014). Bayesian model selection for group studies – revisited. NeuroImage 84: 971-985.

Further Readings (3)

- 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.
- Stephan KE, Mathys C (2014). Computational approaches to psychiatry. Current Opinion in Neurobiology 25: 85-92.
- Yao Y, Raman SS, Schiek M, Leff A, Frässle S, Stephan KE (2018) Variational Bayesian inversion for hierarchical unsupervised generative embedding (HUGE). Neuroimage 179:604-619.

MANY THANKS TO STEFAN FRÄSSLE,
KLAAS ENNO STEPHAN, HANNEKE DEN OUDEN
AND JEAN DAUNIZEAU FOR MANY OF THE SLIDES!

Translational Neuromodeling Unit

University of Zurich[UZH]

ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich