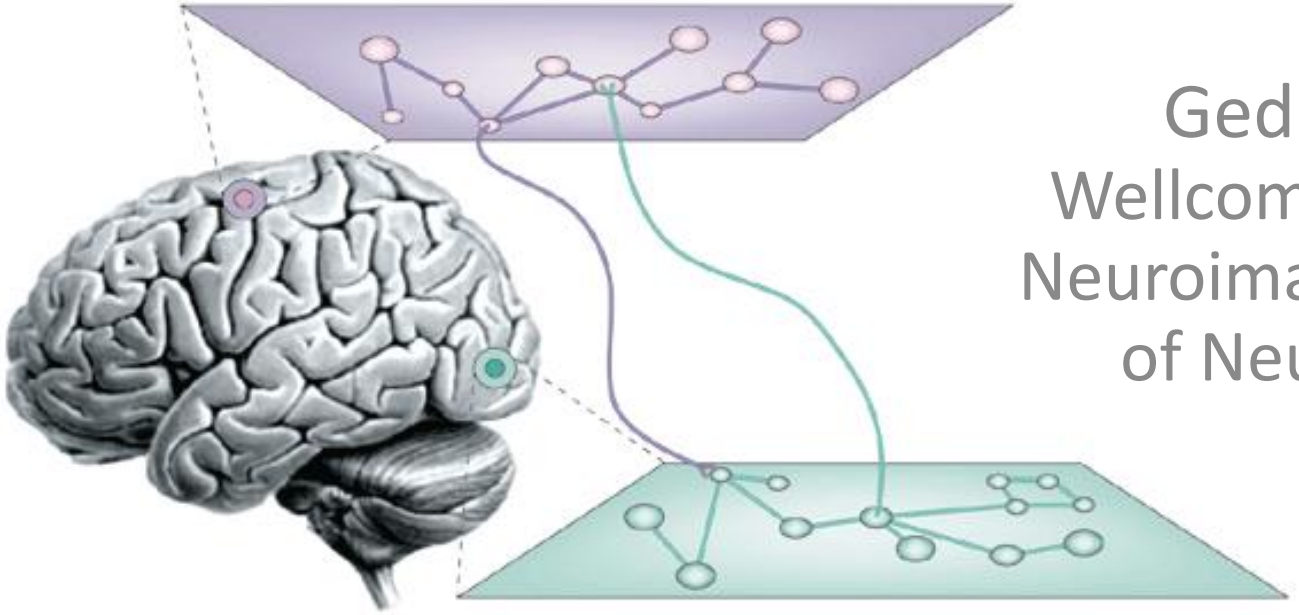
Stochastic Dynamic Causal Modelling for resting-state fMRI



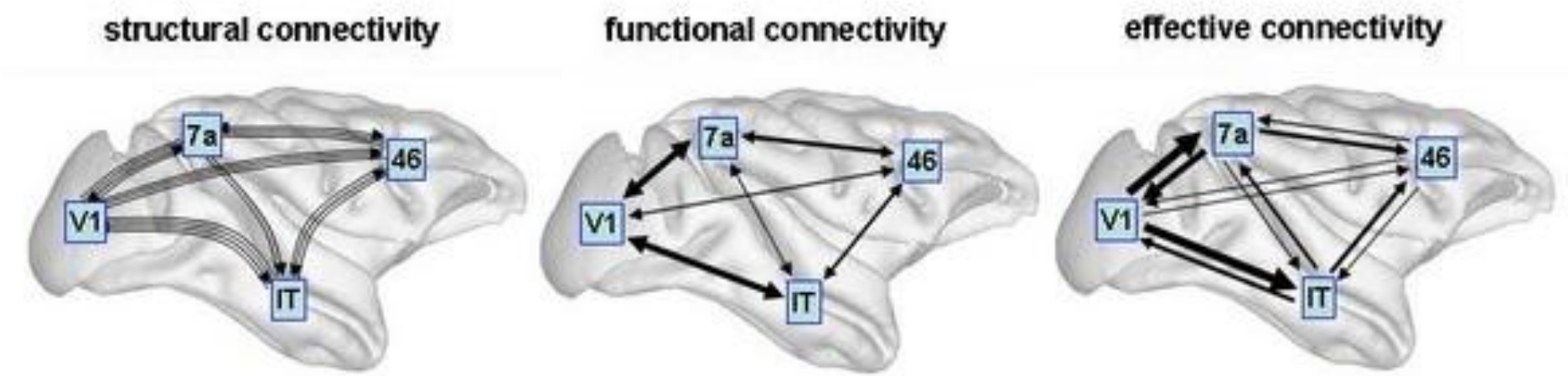
Ged Ridgway, FIL,
Wellcome Trust Centre for Neuroimaging, UCL Institute of Neurology, London

Overview

- Connectivity in the brain
- Introduction to Dynamic Causal Modelling
- Bayes, prior knowledge, and model evidence

- Connectivity in disease
- Motivation for resting-state fMRI in pharma
- Stochastic DCM and resting-state fMRI
- Pros and cons of sDCM for rs-fMRI in pharma

Connectivity in the brain



Sporns 2007, *Scholarpedia*

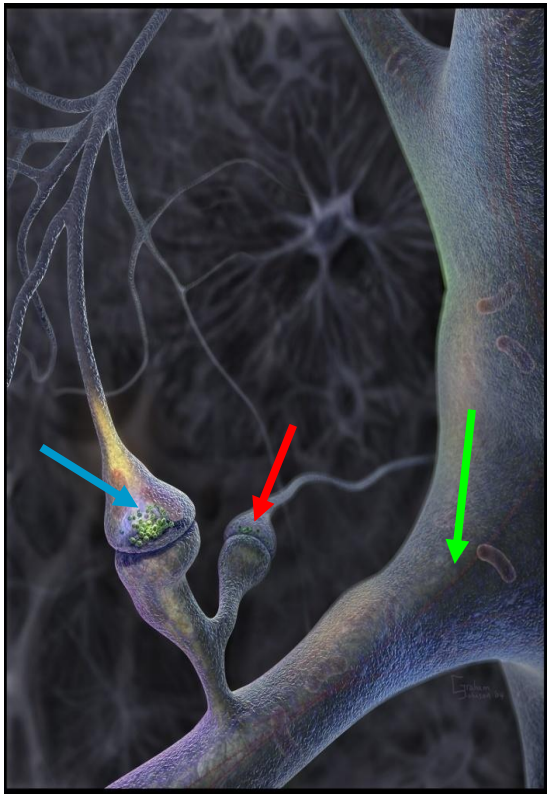
- **structural / anatomical connectivity**
  = presence of axonal connections (from tracing or dMRI)

- **functional connectivity**
  = statistical dependencies between regional time series

- **effective connectivity**
  = causal (directed) influences between neuronal populations

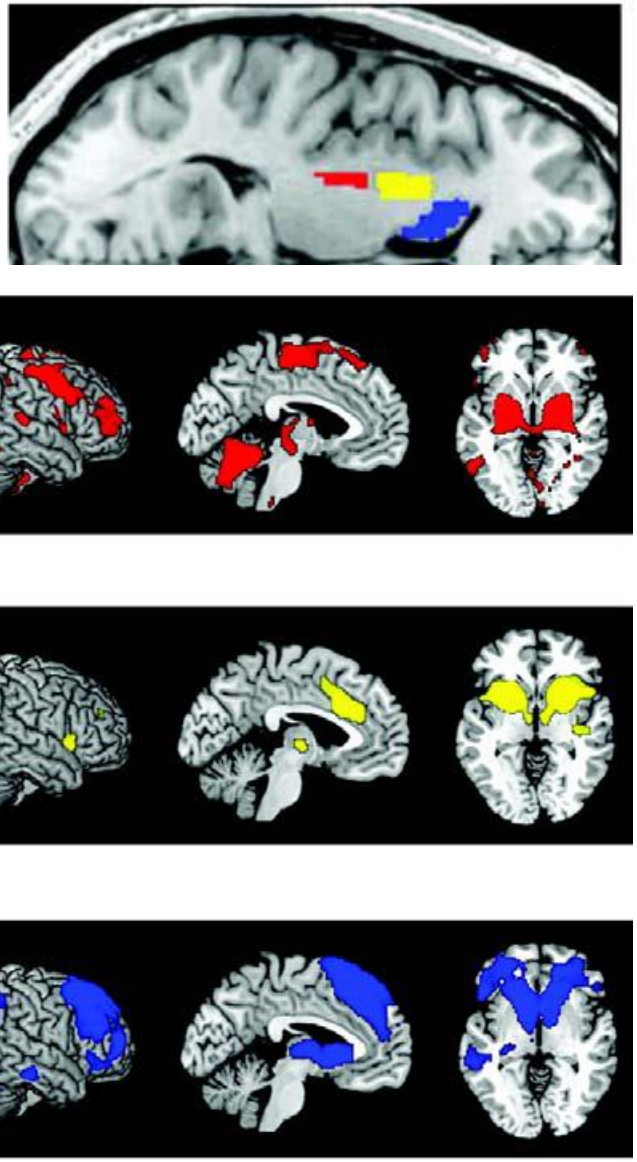
Functional and effective connectivity are dynamic

- Context-dependent recruitment and gating of connections
  - Synaptic depression over millisec
  - Long-term potentiation over weeks
- Even structural connectivity changes
  - Microscopic and macroscopic (developmental) levels
  - (Friston, 2011, *Brain Connectivity*)
- Pharmacological manipulations



Analysis of functional connectivity

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



Helmich et al. (2009) *Cerebral Cortex*

Analysis of effective connectivity

- To get beyond descriptive statistical measures requires a model; parameterise connectivity
  - "modelling -> understanding"
- The model defines what is meant by (effective) direct/directed causal influence
- Model inversion yields estimated connectivity
- Generative models cause the observed data
  - "better to use an original than a derived measure"

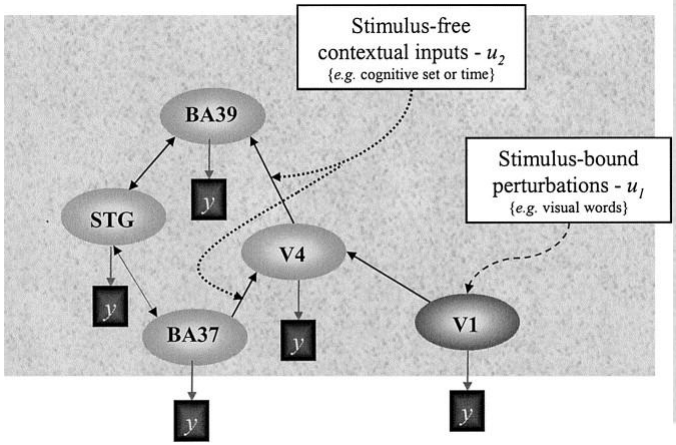
Generic time-series models

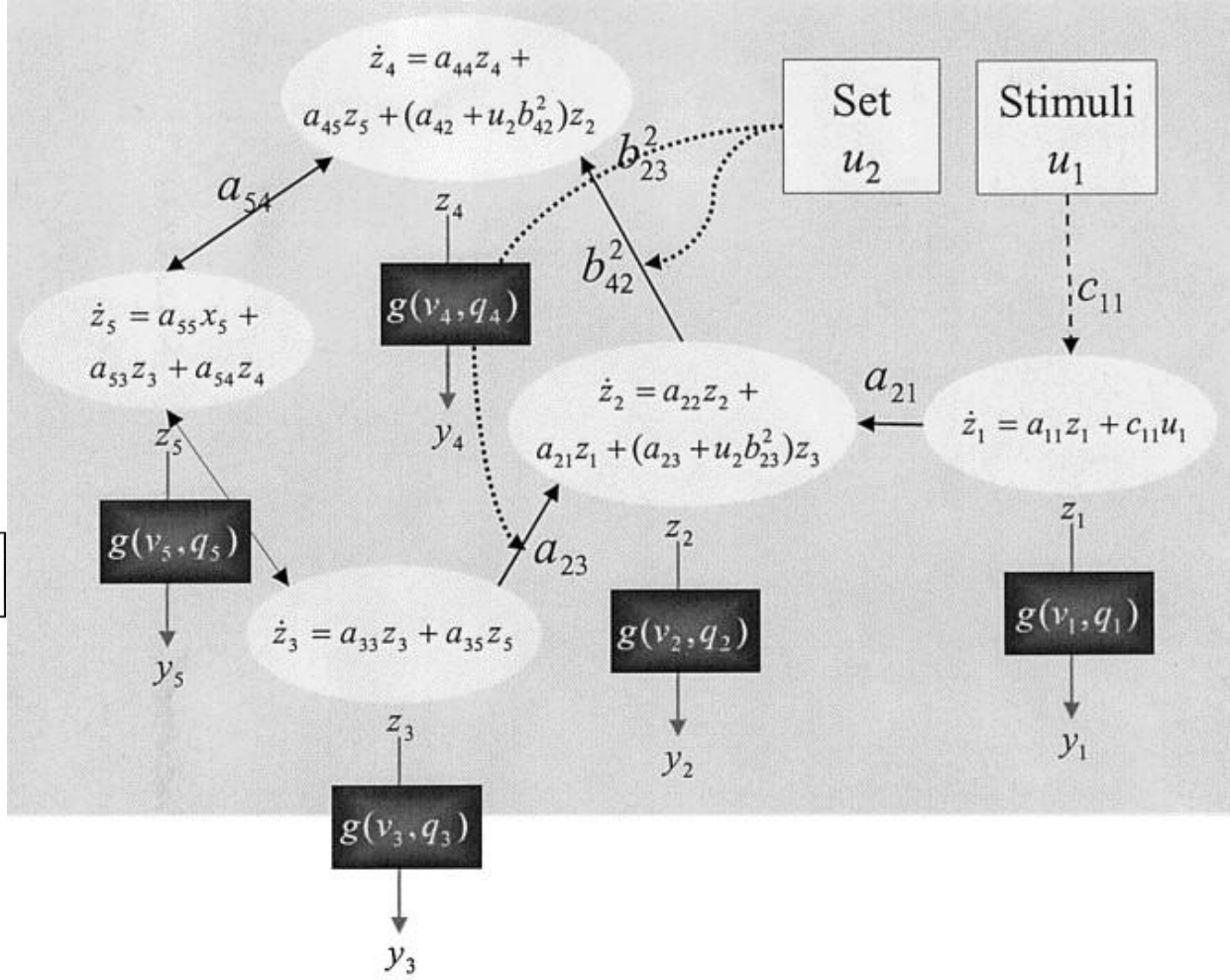
- Discrete-time “auto-regressive” models
  - next states = f(previous states, inputs, parameters)
  - $x(k+1) = f(x(k), u, \theta)$
  - Underlies Granger Causality
    - Very roughly, if current x_1 and x_2 explain next x_1 better than x_1 does alone, then x_2 Granger-causes x_1
- Continuous-time dynamical systems models
  - rate of change = f(current states, inputs, parameters)
  - $dx/dt = f(x(t), u, \theta)$
  - Used in Dynamic Causal Modelling
    - Bayesian model comparison **accounting for complexity**
    - Friston (2011) *Brain Connectivity*

Dynamic Causal Modelling

- Neurodynamic model (state evolution model)
  - Underlying (hidden) neuronal states x (or often z)
  - $dx_i/dt = f(\{x_1, \ldots, x_n\}, \{u_1, \ldots, u_m\}, \{\theta_1, \ldots, \theta_p\})$
  - Linear state-coupling terms: $a_{i1}\, x_1 + \ldots + a_{in}\, x_n = \Sigma_k\, a_{ik}\, x_k$
  - Linear input terms: $c_{i1}\, u_1 + \ldots + c_{im}\, u_m = \Sigma_j\, c_{ij}\, u_j$
  - Bilinear input-modulated coupling terms: $\Sigma_j\, \Sigma_k\, u_j\, B_{ijk}\, x_k$
  - $d\mathbf{x}/dt = A\mathbf{x} + C\mathbf{u} + \Sigma_j\, u_j B^{(j)}\, \mathbf{x}$ [A, B and C in interface]
- Haemodynamic model (observation model)
  - Response = f(state, parameters) + confounds + noise
  - $y_i = g(x_i, \{\theta_h\}) + X\beta + \varepsilon$

DCM





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

The bilinear model

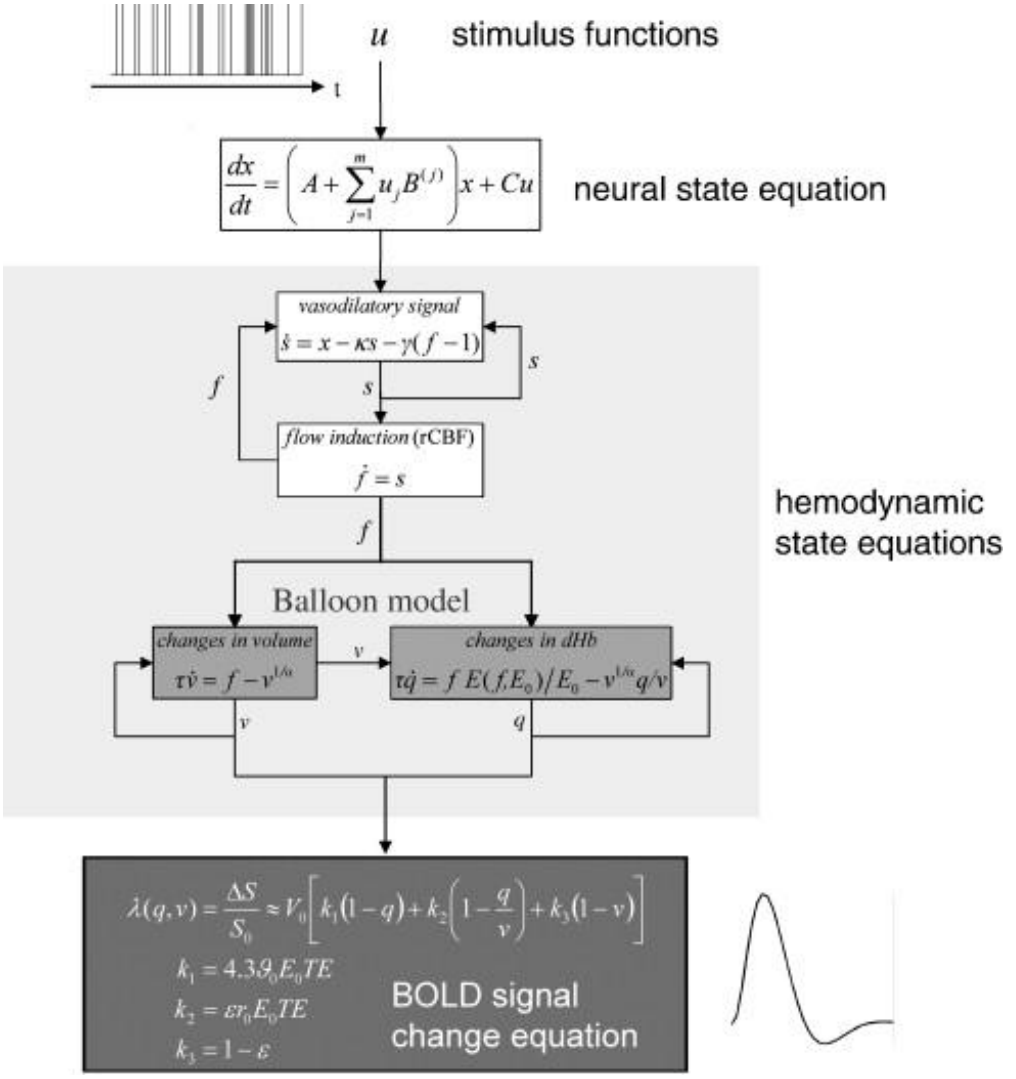
latent connectivity induced connectivity

$$\begin{bmatrix} \dot{z}_1 \\ \vdots \\ \dot{z}_5 \end{bmatrix} = \left\{ \begin{bmatrix} a_{11} & \cdots & 0 \\ a_{21} \; a_{22} \; a_{23} & & \\ \vdots \quad\; a_{33} & & a_{35} \\ a_{42} & a_{44} & a_{45} \\ 0 \quad \cdots \; a_{53} & a_{54} & a_{55} \end{bmatrix} + u_2 \begin{bmatrix} 0 & \cdots & 0 \\ & b^2_{23} & \\ \vdots & \ddots & \vdots \\ b^2_{42} & & \\ 0 & \cdots & 0 \end{bmatrix} \right\} \begin{bmatrix} z_1 \\ \vdots \\ z_5 \end{bmatrix} + \begin{bmatrix} c_{11} & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Forward, backward & self

DCM – haemodynamic model

- Generalises Buxton's balloon model
- Complete generative model including noise
- Bayesian inference allows prior constraints (& model comparison)
- Region specific
- Subject specific
- Treatment specific



Stephan et al., 2007, *Neuroimage;* **now revisiting for 7T**

DCM and Bayesian inference

- Generative or “forward” model (with noise distribution assumptions) gives “likelihood”:
  p(data | parameters, model)
- To estimate parameters given observed data need to “invert” model:
  p(parameters | data, model)
- Bayesian inference enables this inversion using “prior” information about parameters

Bayesian inference

- Bayes rule:
  - $p(A, B) = p(A|B)\, p(B) = p(B|A)\, p(A)$
  - **$p(B|A) = p(A|B)\, p(B) \,/\, p(A)$**
  - $p(A) = \Sigma_b\, p(A, B=b) = \Sigma_b\, p(A|B=b)\, p(B=b)$
- Bayes rule for DCM:
  - $p(\text{ parameters} \mid \text{data, model})$
    $= p(\text{ data} \mid \text{parameters, model})$
    $\times\, p(\text{ parameters} \mid \text{model})$
    $/\, p(\text{ data} \mid \text{model})$

Bayesian model comparison

- The denominator, p(data | model), in turn gives p(model | data) via Bayes rule
- Allows computation of "Bayes factor" to compare p($model_a$ | data) / p($model_b$ | data)
  - Note: same data; no absolute p($model_a$ | data)
- Known as the model evidence and also the marginal likelihood, because parameters are marginalised / integrated out
  - Recall: $p(A) = \Sigma_b\ p(A, B=b)$
- Accounts for complexity (favours parsimony)

Bayesian model comparison

- Can be extended to encompass
  - Random effects model selection over subjects, allowing heterogeneity and outliers (Stephan et al. 2009, *NeuroImage*)
  - Bayesian parameter averaging and Bayesian model averaging accounting for uncertainty over models (Stephan et al. 2010, *NeuroImage*)
  - Comparison of families of models, e.g. top-down/ bottom-up (Penny et al. 2010, PLoS Comput Biol)
  - Optimal experimental design (Daunizeau et al. 2011, *PLoS Comput Biol*)

Free energy in DCM (and the brain!)

- $p(\text{data} \mid \text{model}) = \int p(\text{ data} \mid \theta, \text{model}\ )\ p(\theta)\ d\theta$
- However… the integration is impossible in practice
- We can optimise a lower bound on the model evidence known as the "free energy"
- Using "variational" calculus (variational Bayes)
- The optimised "proposal distribution" tends to the posterior distribution of interest
- Unlike other methods (e.g. Monte Carlo), could be implemented biologically – the Bayesian brain
  - (Friston, 2010, *Nat Rev Neurosci*)

Overview

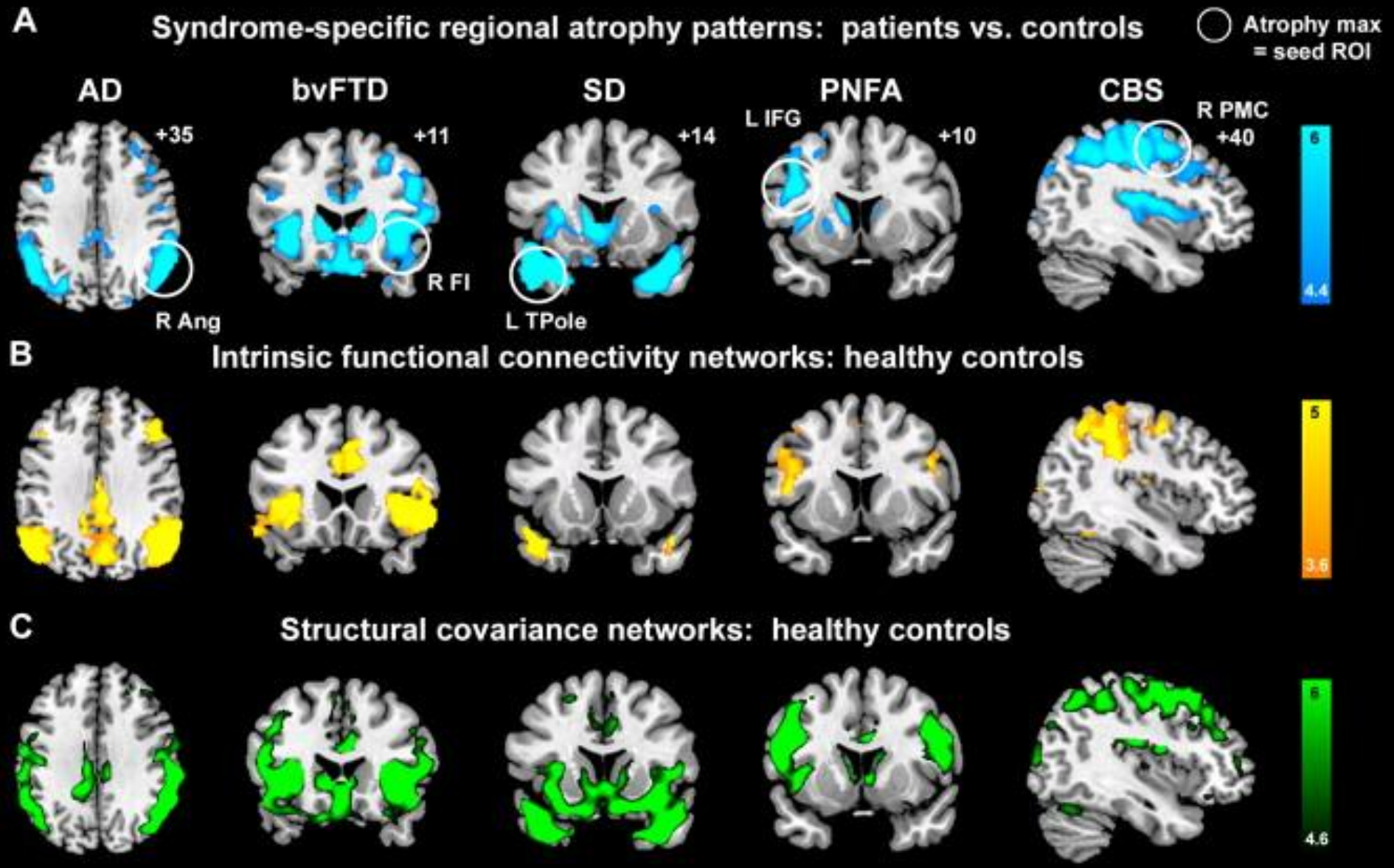
- Connectivity in the brain
- Introduction to Dynamic Causal Modelling
- Bayes, prior knowledge, and model evidence

- Connectivity in disease
- Motivation for resting-state fMRI in pharma
- Stochastic DCM and resting-state fMRI
- Pros and cons of sDCM for rs-fMRI in pharma

Connectivity and disease

- “Dysconnection in Schizophrenia …”
  - Stephan et al. (2009) *Schizophr Bul*
- “Autism spectrum disorders: developmental disconnection syndromes”
  - Geschwind et al. (2007) *Curr Opin Neurobiol*
- “Neurodegenerative Diseases Target Large-Scale Human Brain Networks”
  - Seeley et al. (2009) *Neuron*

Seeley et al. (2009)



Promising results / example applications

- Alzheimer's disease (and risk factors)
  - AD and MCI (Binnewijzend et al., in press, *Neurobiol Aging*)
  - Amyloid positive healthy elderly (Hedden et al., 2009, *J Neurosci;* Sheline et al., 2010, *Biol Psych*)
  - APOE e4 carrying elderly (Sheline et al., 2010, *J Neurosci*)
  - APOE e4 carrying under 35s! (Filippini et al., 2009, *PNAS*)
- Parkinson's disease
  - Rowe et al. (2010) *NeuroImage*:
  - "DCM model selection is robust and sensitive enough to study clinical populations and their pharmacological treatment"

Advantages of rs-fMRI for pharma

- Sensitivity to early/mild change
  - E.g. preceding structural atrophy
- Generality for multiple diseases and severities
  - No need for relevant (and implementable) task
  - No issue of task-difficulty, floor/ceiling effects, etc.
- Ease of standardisation, practicality
  - No special hardware or expertise required
  - Short scan, repeatable given problems

DCM for resting state data ?

- Neurodynamic model without inputs **u**
- d**x**/dt = A**x**
- Stability requires (roughly) negative feedback
  - More precisely, negative real eigenvalues of A
- In the absence of input/perturbation **x** decays
- Without dynamics of **x** cannot have coupling!
- Require endogenous stochastic fluctuations
  - State noise – but differentiable rather than Markovian
  - d**x**/dt = A**x** + **ω**

Stochastic DCM

- Applicable to both task-driven and resting-state fMRI
- Uses variational Bayesian "generalised filtering" (Friston et al., 2010, *Math Probl Eng*)
- More complicated than usual state noise (cf. Kalman)
  - "separation of dynamics into a slow, low-dimensional flow on an attracting manifold and a fast (analytic) fluctuating part that describes perturbations"
  - "only the slow dynamics are communicated among nodes, which means we can model distributed activity with a small number of macroscopic variables (e.g. one per node) with fast fluctuations that are specific to each node" – (Friston et al., 2011, *NeuroImage*)

Regions/nodes for (s)DCM

- ROIs can come from prior hypotheses with anatomical atlases (though see “cons” later…)
- Or from functional connectivity analyses
  - E.g. distinct clusters from seed-correlation analysis
  - Or parts from ICA modes, or entire components from a high-dimensional ICA decomposition
- Nodes needn’t be regions, can be distributed
  - E.g. distinct networks (such as default and exec.)
  - Note that (spatial) ICs can have dependencies…

sDCM of rs-fMRI for pharma – Cons

- Need for relatively strong hypotheses
  - Which ROIs, what topology, which aspects to test
- Definition of ROIs in individual subjects
  - Smith et al. (2011) *NeuroImage*, recommends against use of anatomical atlases for generic ROIs
  - Time-consuming, error-prone, less reproducible
- Validity of priors for pathology and/or drug

- Though all to some extent also cons for more general fMRI in pharma (assumptions = priors)

sDCM Cons – revisited

- Need for relatively strong hypotheses
  + Savage-Dickey facilitates network discovery
- Definition of ROIs in individual subjects
  + High-dimensional registration improving all the time (Dartel, LDDMM, ANTS, Nifty-Reg, Geodesic Shooting)
  + Atlas fusion strategies can help (STAPLE, MAPS, LEAP)
- Validity of priors for pathology and/or drug
  + Evaluating priors using model evidence (Moran et al.)

sDCM of rs-fMRI for pharma – Pros

- Connectivity from neuronal model parameters more interpretable than correlations or components; perhaps also more sensitive
- Potential for modelling concomitant neuronal and haemodynamic treatment effects
- Principled model selection, random effects inference (outliers, etc.), families of models
- Can be applied to regions within a network and/or to interacting networks
- Recent and on-going work enabling more nodes

Some useful references

- **The first DCM paper:** Dynamic Causal Modelling (2003). Friston et al. *NeuroImage* 19:1273
- **Physiological validation of DCM for fMRI:** Identifying neural drivers with functional MRI: an electrophysiological validation (2008). David et al. *PLoS Biol.* 6 2683
- **Hemodynamic model:** Comparing hemodynamic models with DCM (2007). Stephan et al. *NeuroImage* 38:387
- **Group Bayesian model comparison:** Bayesian model selection for group studies (2009). Stephan et al. *NeuroImage* 46:1004
- **Ten Simple Rules for Dynamic Causal Modelling** (2010). Stephan et al. *NeuroImage* 49(4):3099
- **Network discovery with DCM**. Friston et al., *NeuroImage* 56(3):1202
- Generalised filtering and **stochastic DCM** for fMRI. Li et al., *NeuroImage* 58(2):442