Novel and Computationally Efficient Bayesian Methods in Neuroimaging

Global Tractography within a Bayesian Framework

Theo Kypraios
http://www.maths.nott.ac.uk/∼tk

University of Nottingham, UK

Joint Statistical Meeting, 3–8 August, Montreal

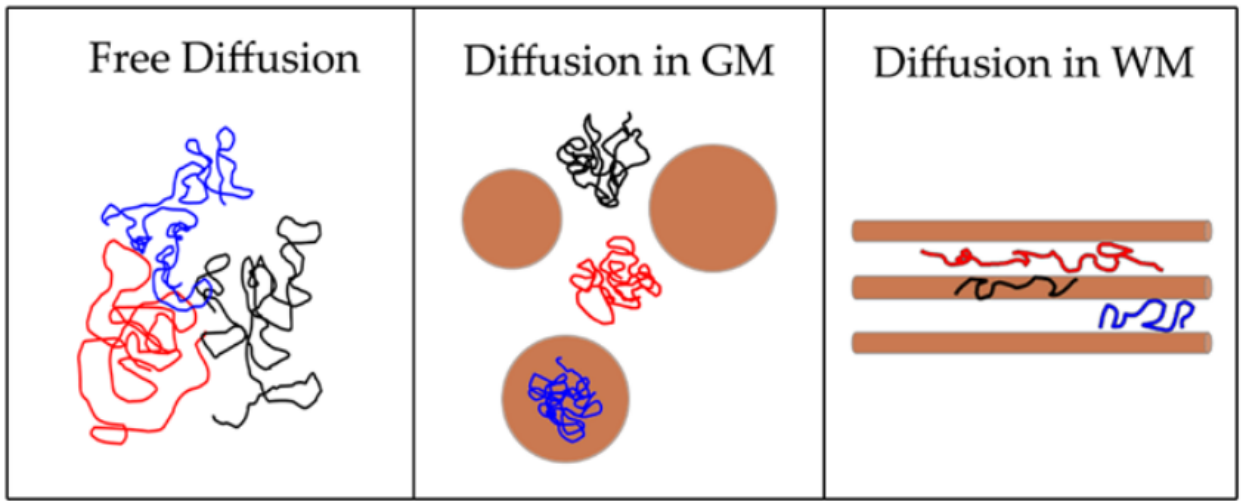
Outline

- Background in Neuroimaging/MRI.
- Motivation.
- Modelling.
- Inference.
- Model Selection/Assessment.
- Conclusions.

Diffusion Magnetic Resonance Imaging (MRI)

- Diffusion magnetic resonance imaging (MRI) provides a unique probe into the micro-structure of materials.
- The method observes the displacements of particles that are subject to Brownian motion with a sample material.
- In particular, it measures the probability density function $p(\cdot)$ of particle displacements $\mathbf{x}$ over fixed time t.
- The micro-structure of the material determines the mobility of the particles within and thus determines $p(\cdot)$.
- Conversely, features of $p(\cdot)$ provide information about the material micro-structure.

- Water is a major constituent of biological tissue.
- In biomedical diffusion MRI the particles of interest are usually water molecules which within tissue undergo random motion due to thermal fluctuations.
- The brain has a complex architecture of grey–matter areas ("functional centres") connect by white–matter fibres ("wires").
- Diffusion–Weighted MRI is currently the only tool that allows study and reconstruction of these tracts, non-invasively and in-vivo, via a process known as tractography.



Infer Structural Connectivity

Task: Given a brain scan image, can we infer which brain regions are *anatomically* connected with each other?

1. Suppose that for each voxel (e.g. a "3D pixel") we have got some measurements;
   - (e.g. signals which are related to the Fourier transform of the displacements in axial direction).

2. We may infer the fibre orientation within that voxel;
   - (e.g. by fitting some physically–motivated models on measurements).

3. Use *tractography* to reconstruct the tracts in the brain by using the inferred fibre orientation in each voxel.
   - (e.g. deterministic, probabilistic).

Methods for Tractography

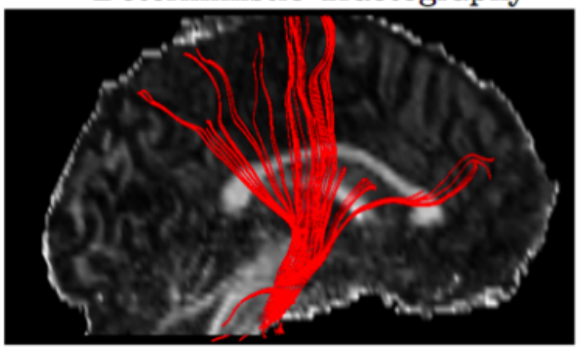
Deterministic tractography: Only uses the inferred fibre orientation in each voxel (e.g. LS, MLE . . .) and follow the fibre orientations.

- Does not take into account the uncertainty in the inferred orientation;
- Any other voxel is either connected or not connected to the seed.

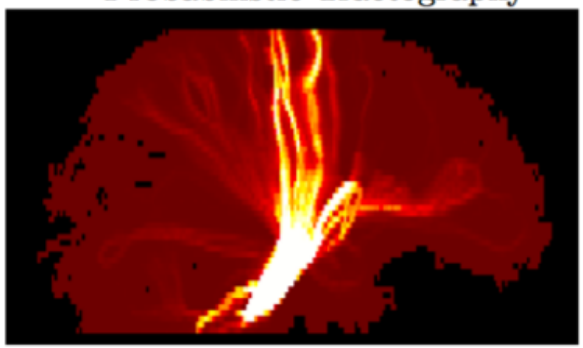
Probabilistic tractography: Draw a sample of orientations (e.g. from a posterior distribution), start from a seed and generate N streamlines, count how many of them pass through the end voxel . . .

- Uncertainty is taken into account.
- However, within small regions uncertainty can cause the pathways to deflect.

Deterministic Tractography



Probabilistic Tractography



Challenges

1. If we were to employ, for example, probabilistic tractography, how do we efficiently infer the fibre orientation for large a number of voxels (e.g. typically in the order of 64^3)? [parameter estimation]

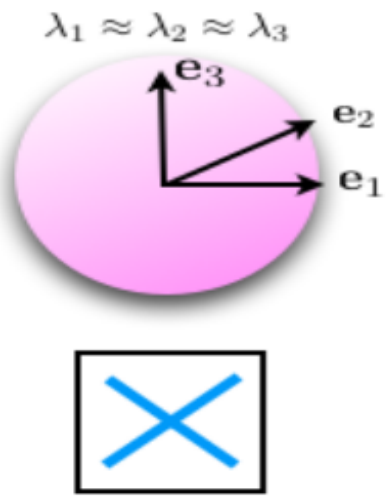
2. A wide variety of models which enable modelling the diffusion of water molecules, e.g. *the Diffusion Tensor Model*, *Partial Volume ("Ball and Sticks") Model* to name but a few.

   How do we ensure that the fitted models are appropriate? [model selection/assessment]

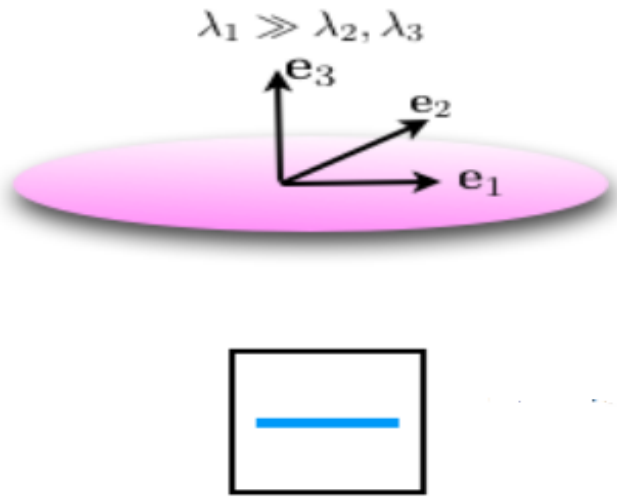
3. Are there any better alternatives to existing approaches to tractography? [global tractography]

DTI and Crossing Fibres

- Diffusion tensor imaging (DTI) has been the most commonly used DW-MRI method [why?]
- Suffers from the assumption that the diffusion scatter pattern exhibits a single directional maxima.



- Assumes that water molecules diffuse according a trivariate Gaussian distribution $(\cdot, 2D)$.
- Fine with very isotropic diffusion profile (i.e. no coherent structure is present) and regions with highly anisotropic profile.



Partial Volume ("Ball and Sticks") Model

The partial volume model (Behrens *et al.* 2003) is a special case of the multi–tensor model (Tuch *et al.*, 2002).

The tensor mixture comprises of a perfectly isotropic (the "ball") and N perfectly anisotropic compartments (the "sticks"), one per fibre population.

The predicted diffusion signal ($\boldsymbol{\mu}$) for a given voxel can be written as

$$\boldsymbol{\mu} \quad = \quad S_0 \left(\left((1 - \sum_{i=j}^{N} f_j \right) \exp\left(-bd\right) + \sum_{j=1}^{N} f_j \exp\left(-bd(\mathbf{g} \cdot \mathbf{v}_j)^2\right) \right)$$

with d being the diffusivity and $\mathbf{v}_j$ a vector describing the jth fibre orientation.

Parameter Estimation for the Partial Volume ("Ball and Sticks") Model

Suppose that N is fixed, e.g. one "ball" and one "stick".

- "Gold–standard" method (implemented in FSL): Bayesian inference using MCMC algorithm {Random walk Metropolis, 2,500 iterations (!), 500 iterations burn-in, thin sample by 2} [why?]
- Parameters of interest: $\boldsymbol{\omega} = (S_0, d, \mathbf{v}_j, f_j, j = 1, \ldots N)^T$, Target density: $\pi(\boldsymbol{\omega}|\mathbf{y})$
  - How to ensure good mixing across the different voxels? How to monitor/check this?
  - How to choose the proposal variance in the RwM? [Adaptive MCMC?]
  - Need good mixing properties and convergence in relatively short MCMC runs.

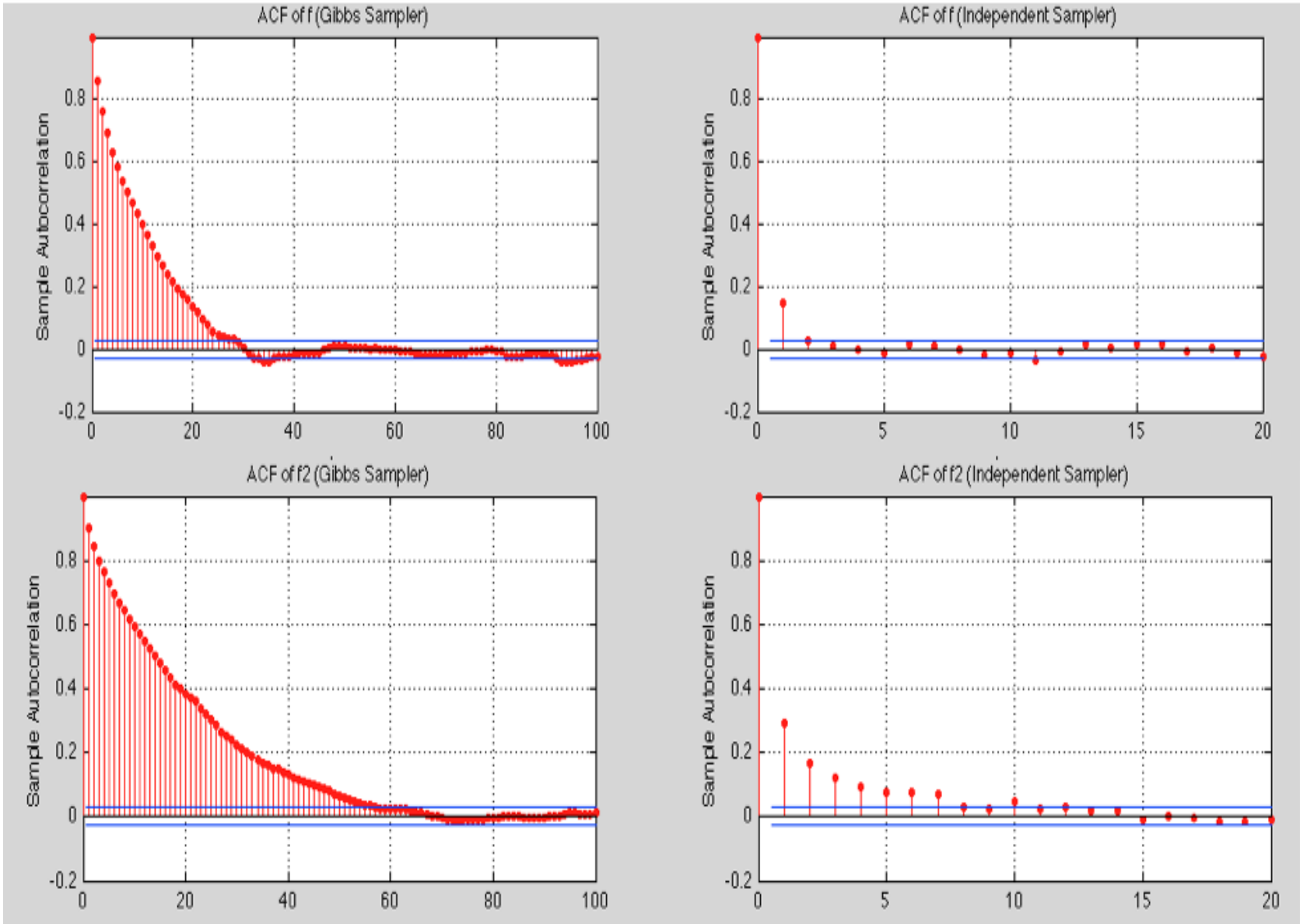
Laplace (Gaussian) Approximation to $\pi(\boldsymbol{\omega}|\mathbf{y})$

- The idea is to bypass RwM and use a tuning-free algorithm, e.g. an independence sampler.
- Need a good proposal distribution to be efficient.
- An obvious choice is to approximate $\pi(\boldsymbol{\omega}|\mathbf{y})$ with a (multivariate) Gaussian distribution:

$$\widehat{\pi(\boldsymbol{\omega}|\mathbf{y})} \equiv N\left(\widehat{\boldsymbol{\omega}}, \Sigma(\widehat{\omega})\right)$$

  where $\widehat{\boldsymbol{\omega}}$ is the MAP estimate (posterior mode) and $\Sigma(\widehat{\omega})$ is the inverse of the Hessian of $-\log \pi(\boldsymbol{\omega}|\mathbf{y})$ evaluated at $\widehat{\boldsymbol{\omega}}$.
- Question is whether or not that approximation is any good ...

RwM vs Independence Samppler



Laplace (Gaussian) Approximation to $\pi(\boldsymbol{\omega}|\mathbf{y})$

- This (simple) Gaussian (Laplace) approximation is pretty good; it seems that there is no need to include higher orders in the Taylor expansion.
- Computationally is very cheap and fast; only need to locate MAP and the Hessian can be derived analytically.
- Since the approximation seems to be pretty good, one can bypass MCMC completely.
- Compromise between Monte Carlo and approximation error.
- Recent work shows that the Laplace approximation is robust to alternative distributions on the noise, e.g. Rician and Non–central χ {Knock, K, Morgan, Sotiropoulos *et al.*, 2013}

How Many Sticks – Model Selection

- When the number of sticks is unknown then N has to be estimated by the data in each voxel.
- The huge number of voxels makes the use methods such as RJ-MCMC prohibitive.
- One option: is to choose a relatively large N by assigning some sparsity inducing priors to the weights $f_j, \quad j = 1, \ldots, N$; for example, Automatic Relevance Determination (ARD) priors [Behren's *et al.* 2007]:

  $f_j|\eta \sim Beta(1, \eta), \quad \pi(\eta) \propto \eta^{-1},$ [issues: MCMC, threshold]

- Another Option: Utilize the Laplace approximation:
  - Approximate the marginal likelihood $\pi(\mathbf{y})$ and compute Bayes Factors and posterior model probabilities.
  - Do (approximate) Bayesian Model Averaging (BMA) (K, Jones, Sotiropoulos, 2013)

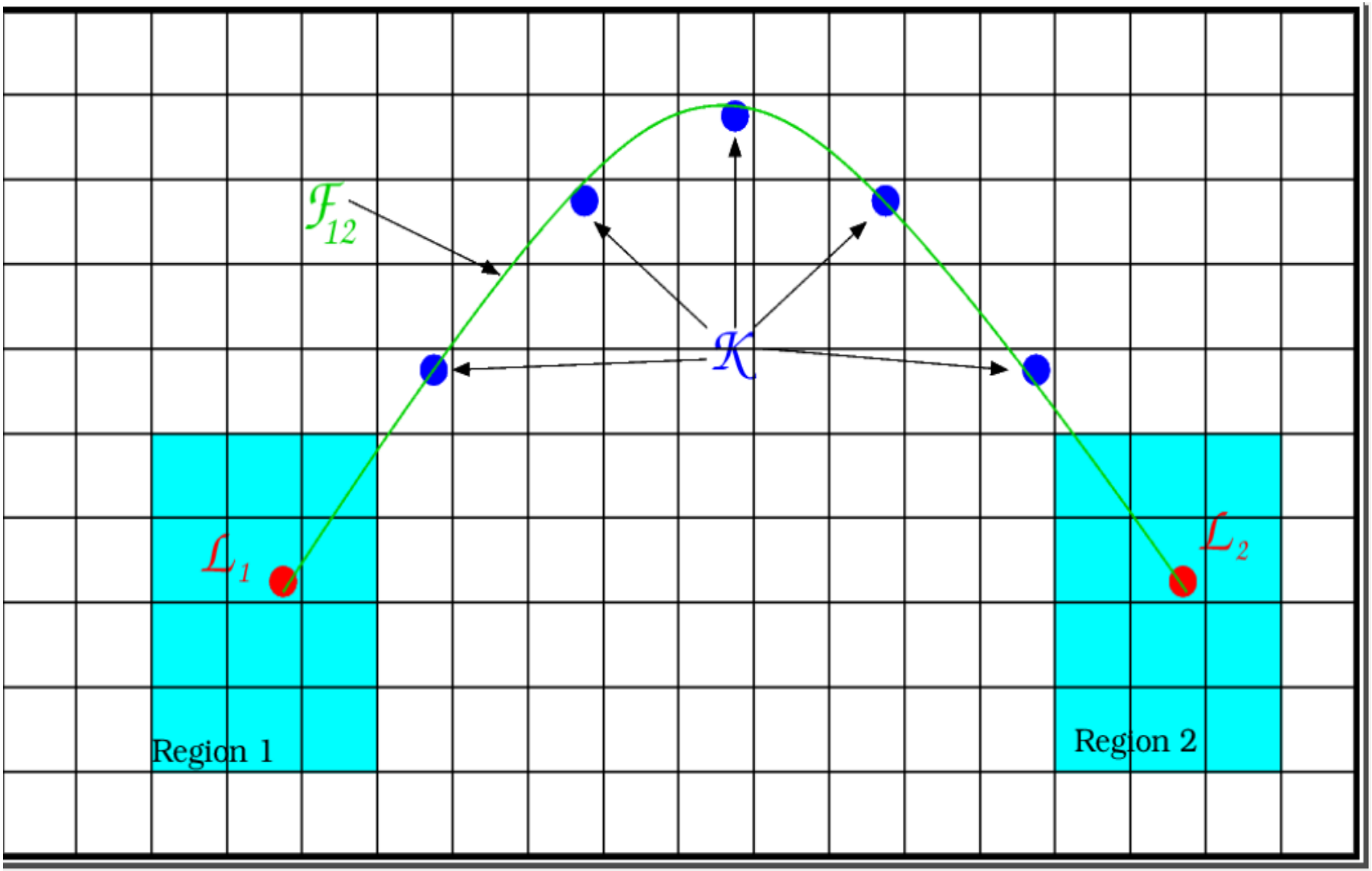
Global Tractography: Motivation

Parametrise the connections between two brain regions at a global level:

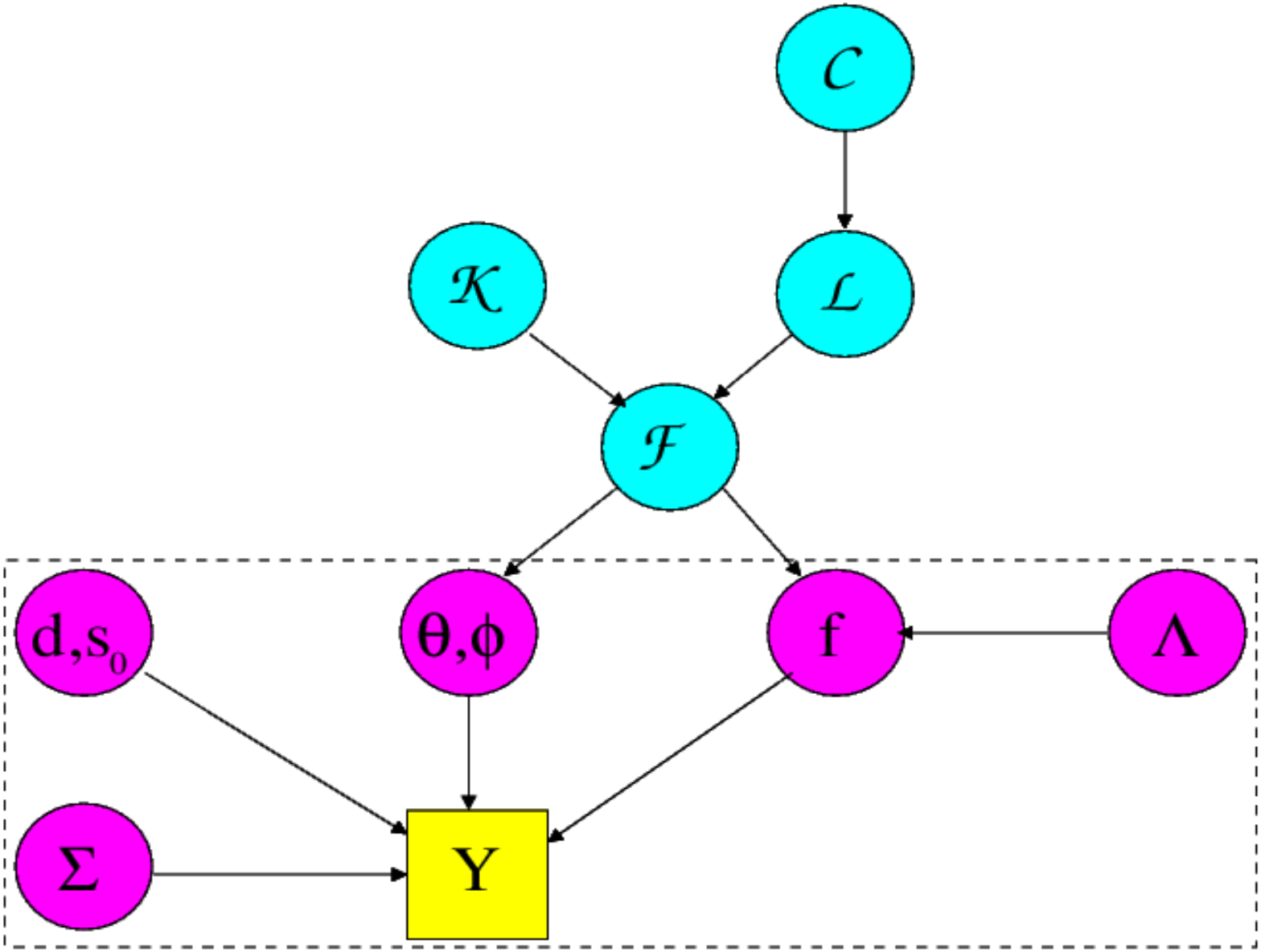
- Reduce the sensitivity to local noise and modelling error when doing tractography
- Acknowledge a priori any known connections. (e.g. from ATLAS).
- Increase robustness of connectivity–based parcellations.
- A step towards the joint inference on functional and anatomical connectivity.

{Introduced by Jbabdi *et al.*, 2007}

Global Tractography: Graphical Representation in 2D



Global Tractography: Graphical Model



Global Tractography: Challenges

- Highly computationally intensive task.
- Initializing the chain can be pretty hard!
- Need a model to describe the connection between regions (e.g. splines) and a model for the orientation within each voxel (e.g. "ball and stick").
- Infer both local and global parameters (e.g. splines).
- High non–linear dependencies between local and global parameters; off–the–self algorithms very inefficient.

Global Tractography

Possible ways to overcome the above issues

- Initialization: Use Laplace approximation to infer fibre orientation in each voxel and the do deterministic tractography.
- MCMC update: do a block update to ensure that the proposed local parameters agree with the global ones in conjunction with parallel tempering/annealing.
- Avoid RJ-MCMC and use Thermodynamic Integration to compute Bayes Factors:

$$\log \pi(\mathbf{y}) = \int_0^1 \mathrm{E}_{\boldsymbol{\omega}|\mathbf{y},t} \log \pi(\boldsymbol{\omega}|\mathbf{y}) \, \mathrm{d}t$$

where $\pi(\boldsymbol{\omega}|\mathbf{y}, t) \propto \pi(\mathbf{y}|\ \boldsymbol{\omega})^t \pi(\boldsymbol{\omega})$.

Conclusions

- Neuroimaging poses many statistical and computational challenges. Often ad-hoc methods which are not well justified are used.
- The high dimensionality of the problem and the non-linear correlation between parameters cause problems to standard MCMC algorithms.
- Need for fast, robust and efficient (but not necessarily exact!) methods for inference and model selection.
- So many modalities to draw inference from: Diffusion MRI, Functional MRI, MEG, EEG . . .: need methods to simultaneously consider all these types of information.
- Global tractography is a step toward the joint inference on functional and anatomical connectivity.

Acknowledgements

People

- Lisa Mott @ University of Nottingham
- Dr Stamatis Sotiropoulos @ FMRIB at University of Oxford
- David Jones @ Harvard University
- Dr Paul Morgan @ University of Nottingham

Funding