present

Mathematical Models of Alzheimer's Disease

Pavan Chaggar, July 2021

Joint work with Alain Goriely, Saad Jbabdi, Travis Thompson and the *OxMBM* group

Overview and Introduction

Alzheimer's is a neurodegenerative disorder with characteristics that lend itself nicely mathematical modelling. Given the clinical nature of such investigation, validation against data is crucial. Here, I will present an overview of the following:

- Alzheimer's disease (AD)
- Mathematical models of AD
- Preliminary results of Bayesian inference using patient data

Alzheimer's Disease – A Brief Summary

AD is characterised by gradual neurodegeneration associated with pervasive spreading of toxic protein species.

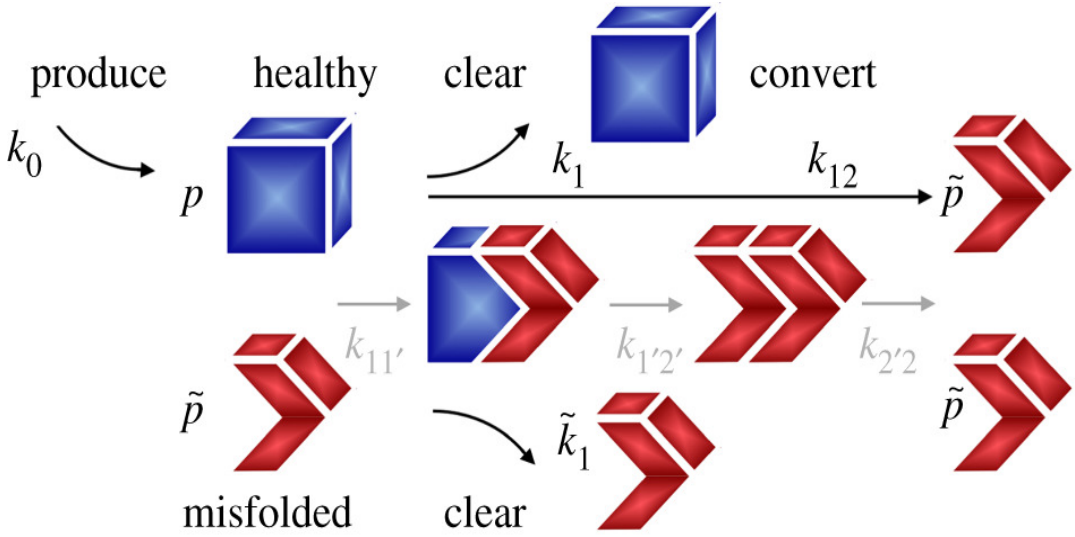
In particular, two proteins, Amyloid beta (Aβ) and tau-protein (τP) are believed to underlie and drive the development of pathology.

Historically, Aβ was primarily investigated as the primary cause of AD. However More recent work has focussed on τP, in part because it spreads very predictably and is more tightly coupled with atrophy and symptom onset.

A Pernicious Pair of Predictable Prion Proteins

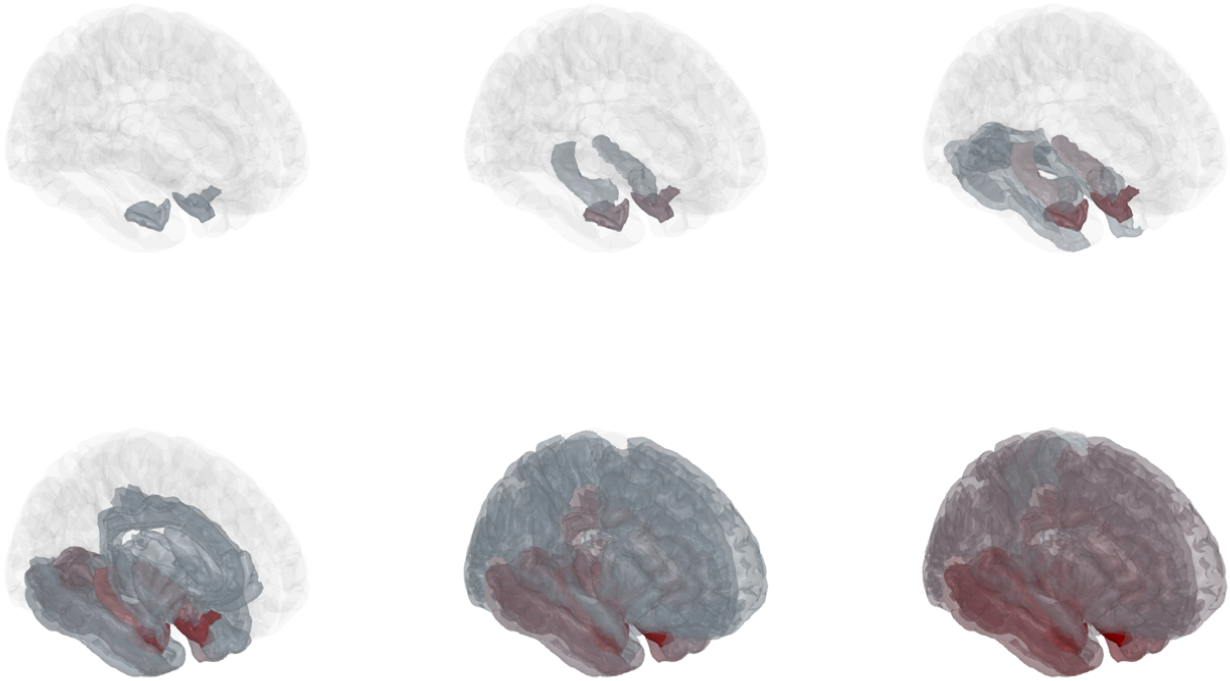
Both Aβ and τP grow via an autocatalytic process resmebling that displayed by prions.

This process is summarised as:



Braak Stages of Tau protein

In most AD cases, τP follows a predictable pattern of neurodegeneration, starting in the entorhinal cortex before spreading through the hippocampal regions, lateral cortex and finally into the neocortex. Atrophy tends to follow closely the spreading pattern of Tau, more so than than of Aβ.



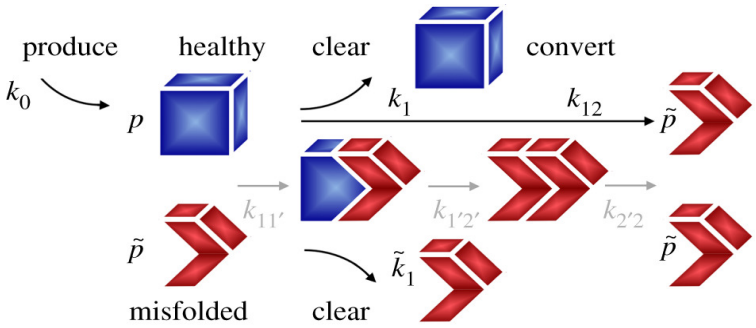
Modelling: What Is It and Why do We Care?

Mathematical models are an unreasonably effective tool for understanding complex processes with some basic ingredients and assumptions. In the case of AD, there are sufficiently many properties, such as tau propogation, that lend themselves to methods in mathematical biology.

(It's also what my funding grant is for)

The Heterodimer Model

Recall the autocatalytic process:

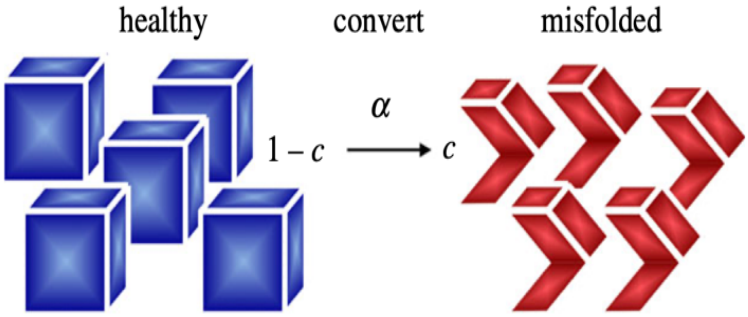


We can describe this process with the following reaction-diffusion equations, where the rates k_{ij} correspond to the rates above.

$$\begin{aligned}
\frac{\partial \mathbf{p}}{\partial t} &= \nabla \cdot (\mathbf{D} \nabla \mathbf{p}) + k_0 &&- k_1 \mathbf{p} - k_{12} \mathbf{p} \hat{\mathbf{p}} \\
\frac{\partial \hat{\mathbf{p}}}{\partial t} &= \nabla \cdot (\mathbf{D} \nabla \hat{\mathbf{p}}) &&- \hat{k}_1 \hat{\mathbf{p}} + k_{12} \mathbf{p} \hat{\mathbf{p}}
\end{aligned}$$

The FKPP Model

The protein dynamics (here shown for a single species) can be simplified in the following way:



With the appropriate diffusion term, we arrive at:

$$\frac{\partial \mathbf{p}}{\partial t} = \nabla \cdot (\mathbf{D} \nabla \mathbf{p}) + \alpha \mathbf{p}(1 - \mathbf{p})$$

Avoiding The Continuum

We typically want to avoid the heavy computations necessary to solve continuous equations on a sufficiently fine mesh. We can do this by assuming that all protein diffusion necessary for AD pathology can be described by transport through axonal pathways. With such an assumption, we can transform our continuum equation:

$$\frac{\partial \mathbf{p}}{\partial t} = \nabla \cdot (\mathbf{D} \nabla \mathbf{p}) + \alpha \mathbf{p}(1 - \mathbf{p})$$

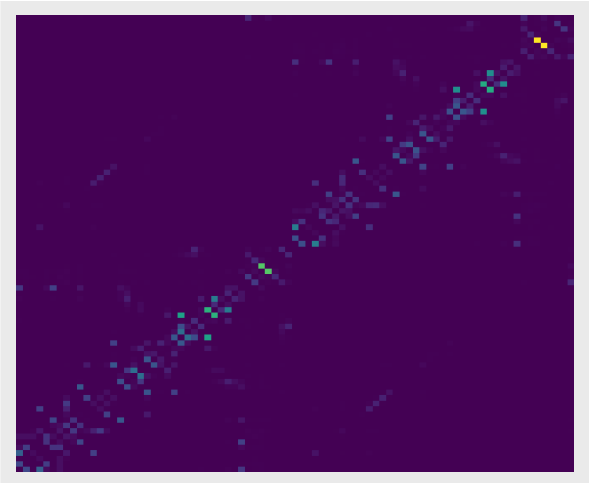
To a discrete equation:

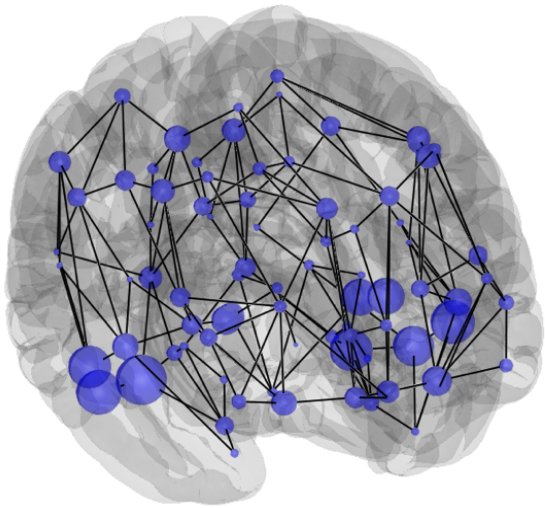
$$\frac{\partial \mathbf{p}_i}{\partial t} = -\rho \sum_{j=1}^{N} \mathbf{L}_{ij}^{\omega} \mathbf{p}_j + \alpha \mathbf{p}_i (1 - \mathbf{p}_i)$$

Where L is the *Laplacian* matrix.

Hey, Where Did You Get Your Laplacian?

This brings with it another problem, finding the Laplacian matrix. Since the Laplacian is $L = A - D$, we can construct it from a connectivity matrix, commonly obtained using tractography. However, these bring with them a high degree of uncertainty.





Simulating the models

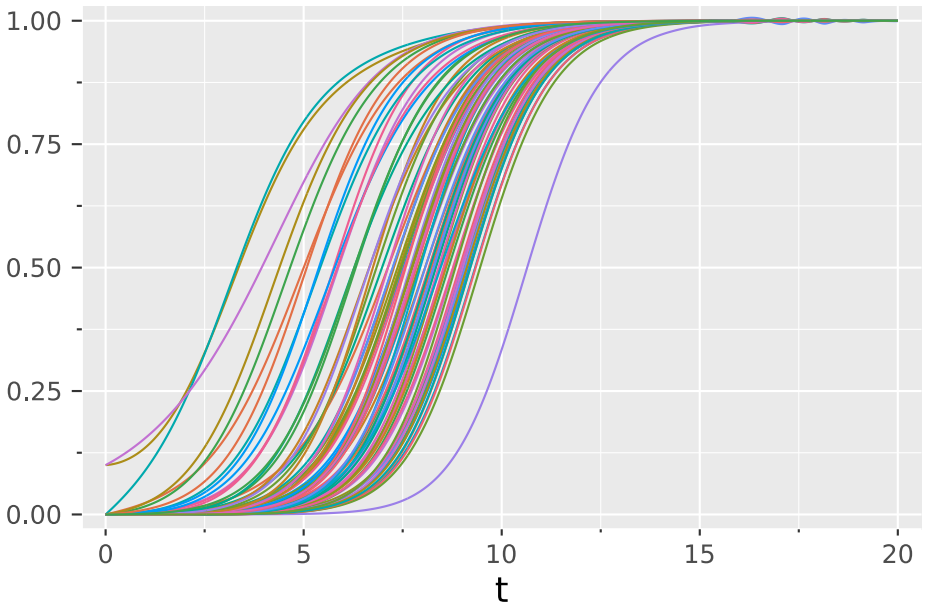
We can (and should) simulate the models to see if the dynamics the models produce align with the expected behaviour of AD pathology

FKPP Dynamics

The simulation to the right shows the dynamics of the FKPP model with seeding in the entorhinal cortex.

By varying the parameters, we can see how the dynamics change and heuristically determine regions of high sensitivity.

Notice how there are larger changes at smaller values of ρ and α.

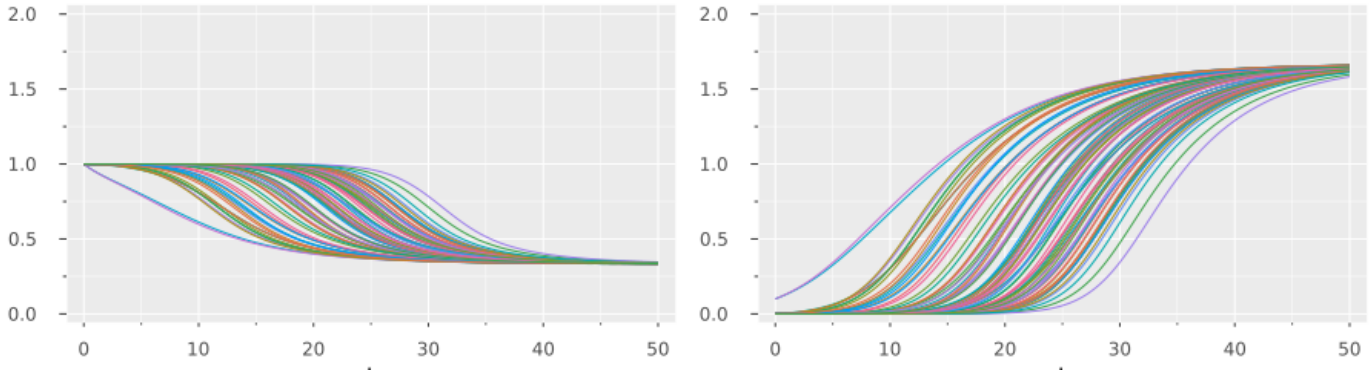


$$\frac{d\mathbf{p}_i}{dt} = -\rho \sum_{j=1}^{N} \mathbf{L}_{ij}^{\omega} \mathbf{p}_j + \alpha \mathbf{p}_i (1 - \mathbf{p}_i)$$

ρ = 1

α = 1

Heterodimer Dynamics



$$\frac{d\mathbf{p}_i}{dt} = -\rho \sum_{j=1}^{N} \mathbf{L}^{\omega}_{ij}\mathbf{p}_j + k_0 - k_1\mathbf{p} - k_{12}\mathbf{p}\hat{\mathbf{p}}$$

$$\frac{d\hat{\mathbf{p}}_i}{dt} = -\rho \sum_{j=1}^{N} \mathbf{L}^{\omega}_{ij}\hat{\mathbf{p}}_j - \hat{k}_1\hat{\mathbf{p}} + k_{12}\mathbf{p}\hat{\mathbf{p}}$$

ρ₁ = 0.1
k₀ = 0.1
k₁ = 0.1
k₁₂ = 0.1
k̂₁ = 0.1

Inference!

Now that we have some models, how do we fit them to data?

Or, more importantly, how do we fit them to data and account for sources of uncertainty?

Inverse Problems using Bayes-Price-Laplace

For observations $\mathbf{x} = x_{1:n}$ and latent variables $\theta = \theta_{1:m}$, we have a joint density

$$p(\mathbf{x}, \theta)$$

To evalulate a particular hypothesis, we need to evaluate the posterior $p(\theta \mid \mathbf{x})$, thus we decompose the joint distribution:

$$p(\mathbf{x}, \theta) = p(\mathbf{x} \mid \theta)p(\theta) = p(\theta \mid \mathbf{x})p(\mathbf{x})$$

Dividing through by the *evidence*, we obtain the Bayes-Price-Laplace rule:

$$p(\theta \mid \mathbf{x}) = \frac{p(\mathbf{x} \mid \theta)p(\theta)}{p(\mathbf{x})}$$

Reverand, What Does it Mean?

$$\underbrace{p(\theta \mid \mathbf{x})}_{posterior} = \frac{\overbrace{p(\mathbf{x} \mid \theta)p(\theta)}^{likelihood\ prior}}{\underbrace{p(\mathbf{x})}_{evidence}}$$

- Likelihood: Probability that a particular set of parameter values generate the observations.
- Prior: Probability representing our initial beliefs about the parameter values.
- Evidence: Normalising factor; probability of observing our data (given our model). Otherwise known as the marginal likelihood.
- Posterior: Probability that some data are *caused* by some set of parameters.

Why is Bayesian Inference Hard? Because Integration is hard!

It's almost always the case that we cannot do Bayesian inference analytically. The principal reason for this comes from the evidence term:

$$p(\mathbf{x}) = \int p(\mathbf{x}, \theta) d\theta$$

As θ becomes larger, the complexity of integration grows exponentially with the dimensionality. Hence, numerical integration methods (such as quadrature) become computationally infeasible.

Thus, to do Bayesian inference, we either need some approximate methods or some clever algorithms for exploring posterior space and integrating wisely.

We tested two methods: variational inference and Hamiltonian Monte Carlo.

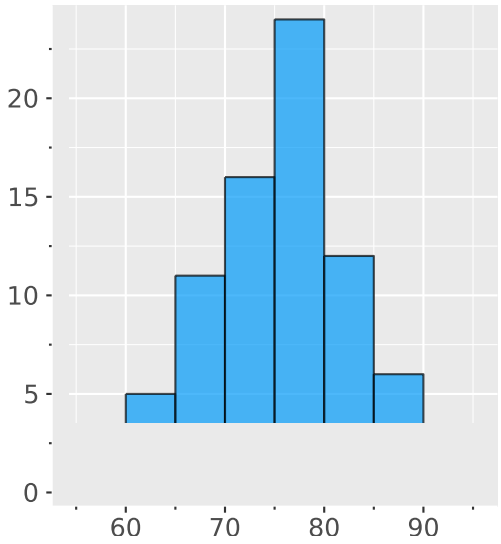
Show me the Results!

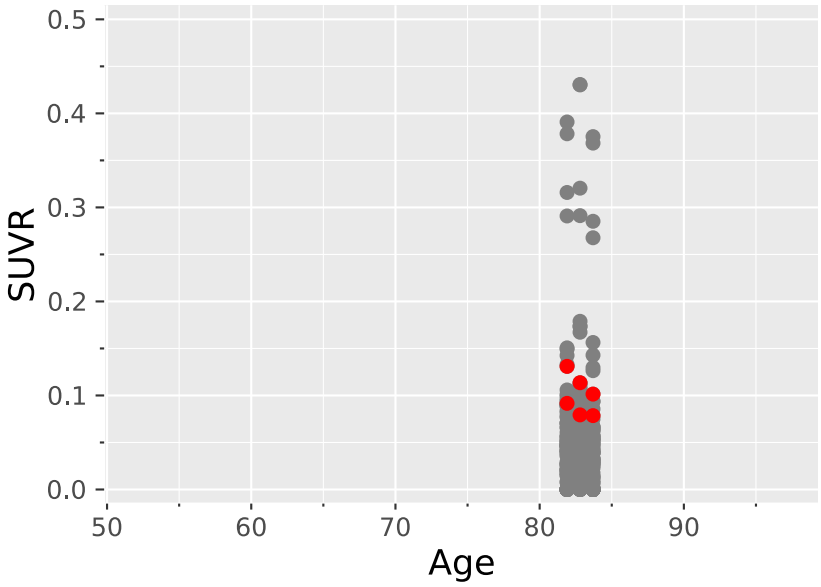
'Pavan, have you actually done any *real* work?'

Comparing Inference Methods for Patient Data

Ultimately, for our dynamical systems $f(u0, \theta, t)$, we want to infer likely values of θ. The dynamical systems proposed earlier in the talk describe the flow of proteins on a network. Thus, we need data that is apprpriate for the model.

Fortunately, such data is available from the Alzheimer's Disease Neuroimaging Initiative (ADNI) in the form of Aβ and τP PET imaging.





subject = 1

Defining a Probabilistic Model

Now that we have data, we want to fit a model using Bayesian inference. We assume the following model structure:

$$\mathbf{x} = f(\mathbf{u0}, \theta) + \mathcal{N}(0, \sigma)$$

We're going to fit the Network FKPP model, the priors for which will be:

$$\sigma \approx \Gamma^{-1}(2, 3)$$

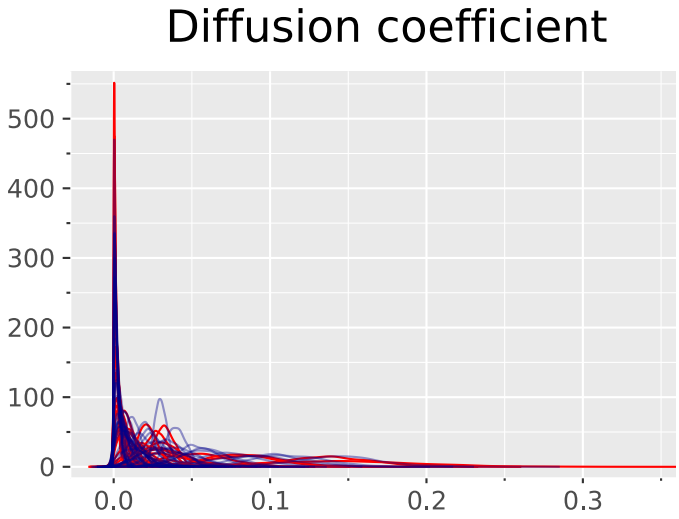
$$\rho \approx \mathcal{N}^{+}(0, 5)$$

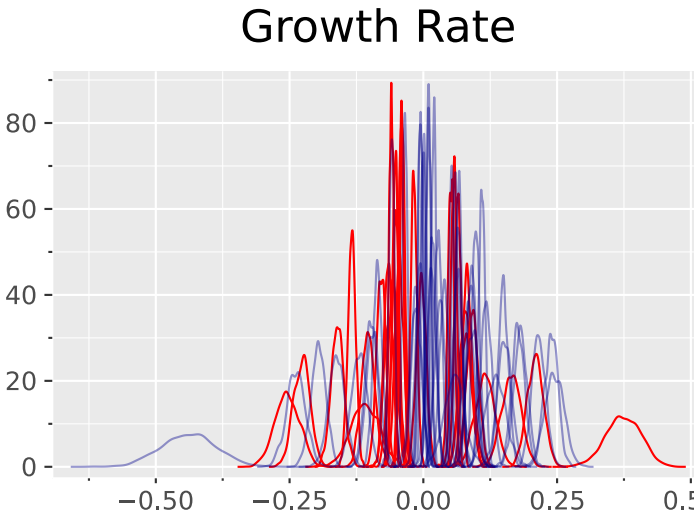
$$\alpha \approx \mathcal{N}(0, 5)$$

Results

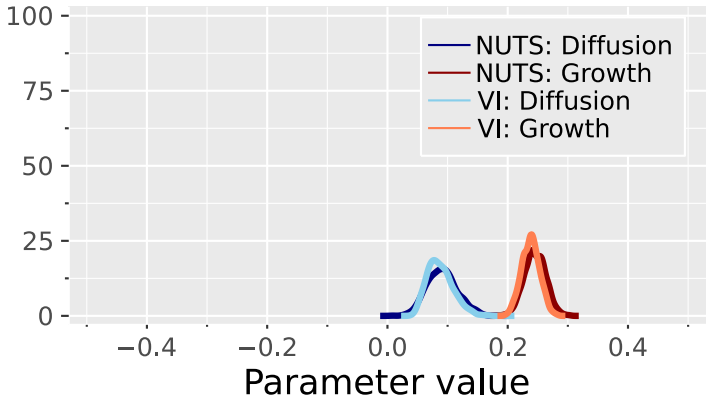
We first ran a NUTS sampler on all single subjects, seperating them based on whether they were Aβ+ (blue) or Aβ- (red). The distributions for the diffusion coefficient, ρ, indicate a *very* slow rate of diffusion, on the order of mm/year, with no substantial differences between Aβ groups.

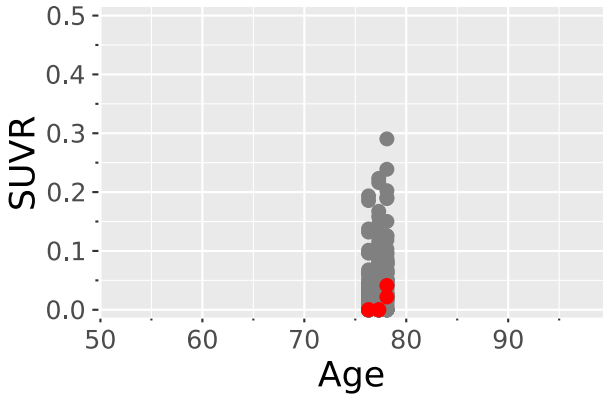
On the other hand, the distributions for the growth rate, α, are largely centered around 0 and suggesting a multimodal distribution.





Comparison with Mean-Field Variational Inference

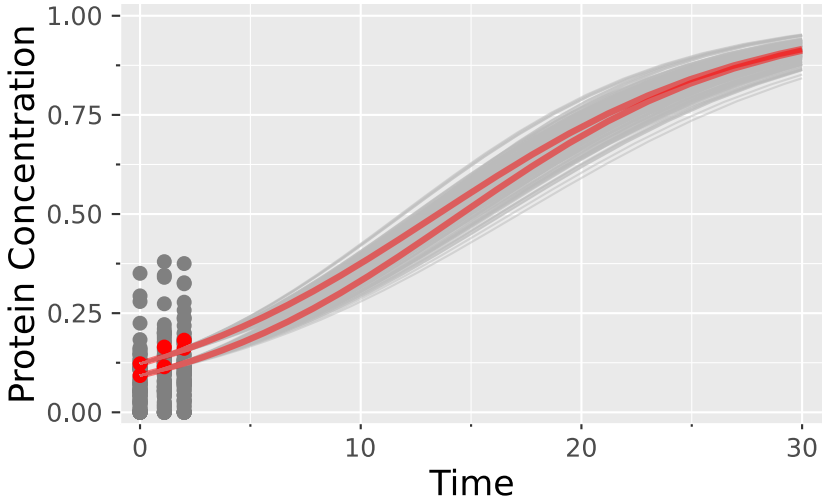


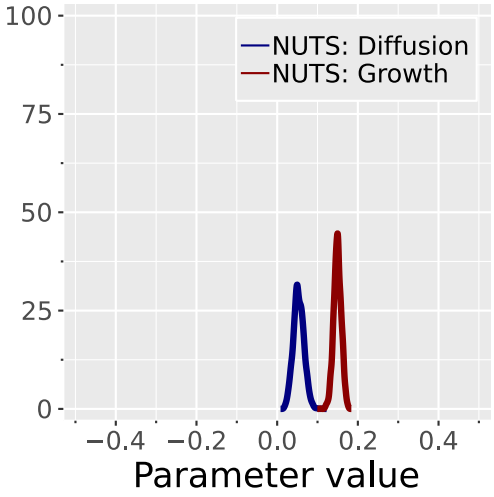


sub = 36

Projecting Uncertainty Forward

A key advantage of Bayesian modelling is the quantification of parametric uncertainty. Using the estimates of parameter distributions, we can project forward in time to simulate potential disease outcomes.





sub = 1

Integrating it All Together

There are a number of sources of uncertainty present in the modelling process:

- Numerical error
- Tractography and the graph Laplacian
- Parametric uncertainty and measurement error
- Structural/model uncertainty

Challenges and Next Steps

The most immediate step will be toward structural or model uncertainty, addressing the question of model selection. This will help determine at least two important factors:

1. Whether one model is preferred over another model
2. Whether there is enough information in the data to distinguish between two dynamic models

Following this, the next big source of uncertainty to quantify, and hopefully resolve, will be that of brain connectivity. We have seen that different graph structures yield significantly different dynamics. Determining which features are important for modelling the diffusion process and how reproducible these are will be necessary before modelling of Alzheimer's can be used for clinical purposes