Summary of Models

Propagation of misfolded proteins in time follows at least two broad mechanisms: transport and growth/decay.

$$\frac{dp_i}{dt} = \frac{\text{change in concentration at node i}}{\text{change in time}} \tag{1}$$

$$= \text{transport process} + \text{growth/decay process} \tag{2}$$

where $\mathbf{i} = 1, 2, \dots N$, for N distinct brain regions. For the DKT atlast, $N = 83$.

The network diffusion model (NDM) is an autonomous linear ordinary differential equation describing **transport** only:

$$\frac{d\mathbf{p}}{dt} = -\rho \mathbf{L} \mathbf{p} \tag{3}$$

where $\mathbf{L}$ is the graph Laplacian determined by the brains connectome. The graph Laplacian approximates the continuous Laplacian operator, Δ, which is used to describe diffusion. Intuitively, you can think about it as applying the L operation to a particular set of connected nodes with different are closer to the average of their neighbors. The stationary point therefore, i.e. the point after which repeated application of L to p have no effect, is when all nodes have the same value. That value is simply the mean concentration.

This process **conserves** the total concentration of protein over time and so is an inadequate description of toxic protein dynamics over the time because it does not describe a growth process. Depending on disease stage, this may approximate the *true* model, but in general, does not describe diseae dynamics as we know them.

The simplest expansion of the model to include growth is the FKPP model:

$$\frac{dp_i}{dt} = -\rho L_{ij} p_j + \alpha p_i (1 - p_i). \tag{4}$$

The first term here, $\rho L_{ij} p_j$, is just the i-th component of the NDM. The second term, $\alpha p_i (1 - p_i)$, describes protein growth. Note that it is not simply αp_i, which would entail exponential, **unbounded** growth. The quadratic term is necessary to provide a description of sigmoidal, **bounded** growth. This is the first justification for FKPP model. Secondly, it is a linearisation of the heterodimer kinetic model proposed by Pruisner (see report). Here, the α parameter bundles up lots of information and can be interpreted as a ratio of protein growth due to natural production and autocatalysis and natural clearance.Therefore, if α is negative, toxic protein concentration will **decreaes** toward 0, if it is positive, it will **increase** toward 1.

Variations and expansions of these models aim to describe these processes in more detail. For example, the conspiracy model (Thompson et al., 2020), describes a mechanism for $A\beta$ induced proliferation of tau, whereas the aggregation models in Fornari et al., 2019, describe an infinite dimensional reaction-diffusion process for oligomers of different sizes, the complexity of which prohibits robust analysis against data (for now, at least).

Relation to ESM model.

The ESM model is a bit hard to parse but hopefully I can elaborate on some of the similarites and differences. My interpretation of the ESM model is a much more complicated and less elegant version of FKPP. It has three terms that can be interpreted as:

$$\frac{dP_i}{dt} = (\text{accumulation and growth process at i}) - \text{clearance at i} + \text{anything else.} \quad (5)$$

$$= (1 - P_i(t))\epsilon_i(t) - \delta_i(t)P_i(t) + \mathbf{N} \quad (6)$$

$$\epsilon_i(t) = P_A(j,i)\beta_i^E(t - \tau_{ij})P_j(t - \tau_{ij}) + P_A(i,i)\beta_i^I(t)P_i(t) \quad (7)$$

The first term is given by equation 2 in Iturria-Medina et al., 2014, which in turn has two terms, given by $\epsilon(t)$, the first essentially represents the diffusion process of the NDM model and the second represents a local **unbounded** growth process. You can think of the NDM model as a deterministic (i.e. not based on probabilities) version of this first diffusion term. Both are based on an anatomical connectome and weighted by some features of connections, e.g. length, number of connections, which modulate the time it takes for protein to go from region i to j, read more at Putra et al., 2021. Next is the growth term, if you bundle this up with the clearance term, second part of equation 1 in the paper and equation 3), this describe a similar process as the non-linear term in the FKPP model.

The third term of the equation, $\mathbf{N}$, is just additive noise, describing random fluctuations, parameterised by a Normal distribution with mean μ and s.d. σ.

So, there are 4 parameters per region, $\{\beta, \delta, \mu, \sigma\}$. There are a few sources that add to the complexity of this model. First, this appears to be a non-autonomous system, i.e. the parameters depend on time. However, these arent parameterised in the paper and I believe are fitted based on individual time points. Secondly, each model parameter is regionally specific. So, if I understand that correctly, this means there are $4 \times \text{number of regions} \times \text{number of time points}$ parameters. Each of which is personalised to a particular subject, given a common seeding location. This is a huge number of parameters relative to the amount of data. In theory, before fitting a model, one should check the identifiability of parameters. My guess would be that this model would not be identifiable from data!

The FKPP model is much simpler while still capturing the important mechanistic insights.